



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department of Chemical Sciences

Graduate School in Molecular Sciences

Curriculum: Pharmaceutical Sciences

Cycle XXXVIII

Accelerating Drug Discovery Through Predictions of Ligand Binding Affinity

This thesis was carried out with the financial support of Evotec SE

Coordinator: Ch.mo Prof. Stefano Corni

Supervisor: Ch.mo Prof. Stefano Moro

Co-Supervisor: Dott.ssa Veronica Salmaso

Ph.D. Student: Silvia Menin

Summary

1. Abstract	3
2. Introduction	5
2.1 Ligand-Receptor Binding Affinity	7
2.2 LBDD Approaches for Calculating Binding Affinity.....	9
2.2.1 Quantitative Structure-Activity Relationships (QSAR)	9
2.2.2 Pharmacophore Modeling	9
2.2.3 Machine Learning and AI Integration.....	9
2.3 SBDD Approaches for Calculating Binding Affinity.....	11
2.3.1 Docking.....	11
2.3.2 Interaction Fingerprints (IFPs).....	11
2.3.3 Molecular Dynamics (MD)	12
2.4 Binding Affinity and Molecular Dynamics.....	14
2.4.1 Free Energy Perturbation (FEP)	14
References.....	18
3. Aim of the Work	23
3.1 Development of TTMD for Complex Stability Assessment	24
3.2 Application of FEP to Membrane Protein Targets.....	25
References.....	27
4. Scientific Publications.....	29
Qualitative Estimation of Protein–Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations	33
Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement.....	53
A Comprehensive Study Of SARS-Cov-2 Main Protease (M ^{pro}) Inhibitor-Resistant Mutants Selected In A VSV-Based System.....	75
Thermal Titration Molecular Dynamics (TTMD): Shedding Light On The Stability Of RNA-Small Molecule Complexes	107
A second life for the crystallographic structure of Berenil-dodecanucleotide complex: a computational revisitation thirty years after its publication	129

Post-Docking Refinement of Peptide or Protein-RNA Complexes Using Thermal Titration Molecular Dynamics (TTMD): A Stability Insight.....	141
Thermal Titration Molecular Dynamics (TTMD): the Revenge of the Fragments	157
5. Unpublished Works	189
Decoding GPCR-Ligand Interactions with FEP: Evaluating Key Parameters for Reliable Predictions.....	191
Adenosine receptors in neuroinflammation and neurodegeneration	205
6. Conclusions	237

1. Abstract

Predicting ligand binding affinity is a central challenge in drug discovery. This thesis focuses on two molecular dynamics-based approaches for estimating binding strength: Thermal Titration Molecular Dynamics (TTMD) and Free Energy Perturbation (FEP). TTMD, developed and optimized in this work, is a qualitatively-based, but fast and easy to settle method that evaluates complex stability under increasing thermal stress, allowing effective ligand ranking and binding pose validation across proteins, nucleic acids, and fragment-based systems. In parallel, FEP methods were explored in collaboration with Evotec to support membrane protein targets such as GPCRs. This work focused on adapting and benchmarking protocols, with the aim of integrating FEP for membrane proteins into Evotec's drug discovery pipelines. While quantitative results using Absolute Binding Free Energy (A3FE) are not yet available, the initial adaptation establishes a foundation for future industrial applications. Together, these studies highlight the potential of molecular dynamics-based approaches to predict binding affinities, providing both a fast qualitative tool (TTMD) and a framework for rigorous quantitative calculations in drug discovery.

La predizione dell'affinità di legame dei ligandi rappresenta una sfida centrale nella scoperta di farmaci. Questa tesi si concentra su due approcci basati sulla dinamica molecolare per stimare la forza di legame: Thermal Titration Molecular Dynamics (TTMD) e Free Energy Perturbation (FEP). TTMD, sviluppato e ottimizzato in questo lavoro, è un metodo rapido e privo di parametri che valuta la stabilità dei complessi sotto stress termico crescente, permettendo una classificazione efficace dei ligandi e la validazione dei loro siti di legame in proteine, acidi nucleici e sistemi basati su frammenti. Parallelamente, i metodi FEP sono stati esplorati in collaborazione con Evotec per supportare target di proteine di membrana, come i GPCR. Questo lavoro si è concentrato sull'adattamento e il benchmarking dei protocolli, con l'obiettivo di integrare FEP per le proteine di membrana nelle pipeline di scoperta farmaceutica di Evotec. Sebbene i risultati quantitativi mediante Absolute Binding Free Energy (A3FE) non siano ancora disponibili, l'adattamento iniziale costituisce una base per future applicazioni industriali. Complessivamente, questi studi evidenziano il potenziale degli approcci basati sulla dinamica molecolare per predire l'affinità di legame, fornendo sia uno strumento qualitativo rapido (TTMD) sia un quadro per calcoli quantitativi rigorosi nella scoperta di farmaci.

2. Introduction

Discovering a new therapeutic compound from scratch is anything but simple. It usually takes more than a decade, often 10 to 15 years, and requires the use of enormous financial resources^{1–3}. For a long time, drug discovery was driven by a combination of chance and trial-error methods, with consequent major time and cost implications. Moreover, this way of working goes hand in hand with the very high attrition that still characterizes this field: close to 90% of candidates entering clinical trials never make it to regulatory approval^{4–6}. Most of these failures can be caused by unexpected problems with absorption, distribution, metabolism, excretion, and toxicity (ADMET), or, simply, to insufficient therapeutic effects⁷. For this reason, the pharmaceutical industry has been pushed toward more rational-based strategies. One of the clearest outcomes of this shift is the growing role of Computer-Aided Drug Design (CADD). By using computational tools to study and predict molecular behavior *in silico*, CADD allows researchers to focus on the most promising parts of chemical space^{8,9}, thereby lowering costs, and shortening drug discovery timelines by up to 50%¹⁰.

CADD emerged in the early 1980s as a branch of computational chemistry, marking a shift from purely descriptive models toward predictive approaches, with the aim of designing new molecules with specific biological properties¹¹. This change was driven by two major developments: the rise of structural biology, which made three-dimensional (3D) structures of biological targets widely available, and the exponential growth in computational power¹². A notable early success was the structure-based

design of Captopril, an ACE inhibitor approved in 1981^{13–15}. The 1990s further demonstrated the practical potential of this approach, with the discovery of HIV protease inhibitors (Saquinavir), and the carbonic anhydrase inhibitor Dorzolamide, both approved in 1995, highlighting the real-world impact of computation in drug discovery^{14–17}. From the 2000s onward, CADD has become an essential tool in the pharmaceutical research¹⁸. Over the past decade, CADD has continued to evolve through several innovations. Modern computational methods allow for more accurate predictions of bioactivity, ADMET properties, and toxicity. Moreover, artificial intelligence (AI) generative models can even assist in designing entirely new molecules optimized for specific targets^{19,20}. Structural biology has also made tremendous improvements: the exponential growth of the number of experimentally determined 3D protein structures in the Protein Data Bank (PDB), combined with AI-assisted tools such as AlphaFold2, now allows great accuracy in protein structure prediction, opening the door to structure-based design for targets that were previously difficult or impossible to access^{21,22}. These advances are further supported by the increasing availability of High-Performance Computing (HPC) clusters and Graphic Processing Units (GPUs), which make longer and more complex simulations possible^{11,23}.

2.1 Ligand-Receptor Binding Affinity

Accurately assessing ligand-receptor binding determinants in the earliest stages of drug discovery is essential to decide whether a molecule is likely to be active toward a central target, and a central metric in this evaluation is binding affinity, quantified by the equilibrium dissociation constant K_D ²⁴.

At equilibrium, K_D is equivalent to the ligand concentration at which half of the target molecules are bound²⁵. It can be calculated as the ratio between free solvate ligand ($[L]$) and protein ($[P]$), and bound complex ($[PL]$) concentrations:

$$K_D = \frac{[L][P]}{[PL]}$$

Thermodynamically, affinity is described by the relation:

$$\Delta G_{bind} = RT \cdot \ln K_D$$

where R is the gas constant, and T the temperature²⁶. A more negative ΔG_{bind} (a larger free energy difference between the bound and unbound states) corresponds to a lower K_D , and, thus, tighter binding.

Equivalently, K_D can be written in terms of the kinetic rate constants:

$$K_D = \frac{k_{off}}{k_{on}}$$

where k_{off} and k_{on} are the macroscopic association and dissociation rate constants.

Since pharmacological activity endures only while a drug remains tethered to its target²⁷, k_{off} assumes a special role between determinants. In fact, it can be also represented by the residence time (τ), which is the time the ligand remains bound in the pocket, and is typically represented as the inverse dissociation rate²⁸:

$$\tau = \frac{1}{k_{off}}$$

Consequently, binding affinity reflects both how favorably (ΔG_{bind}), and how long a compound binds to its target (k_{off} , τ).

Experimental techniques remain the gold standard for measuring binding affinity. Isothermal Titration Calorimetry (ITC) directly determines ΔG_{bind} ²⁹. Surface Plasmon Resonance (SPR) and Bio-Layer Interferometry (BLI) monitor binding in real time to extract K_D , k_{on} and k_{off} ³⁰. Radioligand binding assays can quantify equilibrium affinities and dissociation rates. Fluorescence-based techniques provide both equilibrium and kinetic data in homogeneous or live-cell formats³¹.

While these experimental methods provide very accurate thermodynamic and kinetic data, they require each compound to be synthesized and the target protein to be produced and validated, a process that is slow, and expensive. This is why a key aim of computer-aided drug design is to estimate these same properties directly in silico, allowing large virtual libraries to be screened and only the most promising candidates to move forward to experiments^{32,33}.

Among CADD, two main strategies are used: Ligand-Based Drug Design (LBDD) and Structure-Based Drug Design (SBDD), each described in the following sections.

Modern CADD workflows often combine both ligand- and structure-based approaches rather than using them in isolation. At the start of a project, ligand-based methods can quickly narrow down huge chemical libraries by flagging scaffolds that resemble known active compounds or that fit simple predictive models^{14,34}. The filtered set of candidates can then be refined with structure-based tools, such as docking and free-energy calculations, which provide atomic-level insight into binding modes and prioritize the strongest leads^{35,36}. This stepwise integration allows researchers to balance speed and scalability, helping to explore chemical space efficiently³⁷.

2.2 LBDD Approaches for Calculating Binding Affinity

Ligand-based drug design builds on what is already known about molecules that interact with a specific target. By analyzing the chemical patterns shared by active compounds, it can highlight promising new structures. Since it does not require knowledge of the target structure, LBDD can screen very large libraries quickly. However, its power on predicting binding affinity depends strongly on having enough diverse ligands to build reliable models⁹.

2.2.1 Quantitative Structure-Activity Relationships (QSAR)

QSAR modeling represents one of the most employed LBDD techniques for affinity prediction. It consists of translating molecular structures into a set of descriptors, such as physicochemical, topological, or electronic properties, and correlating these with experimentally determined binding data. Then, statistical models are trained to map descriptor space to activity space³⁸. These models yield predictive equations that approximate free-energy contributions of different molecular features, thereby guiding chemical modifications to potentially enhance compounds affinity³⁹.

2.2.2 Pharmacophore Modeling

Pharmacophore approaches offer a complementary route to affinity estimation by abstracting ligands into sets of essential chemical features, like hydrogen bond donors or acceptors, hydrophobic groups, aromatic rings, etc., that drive molecular recognition^{40,41}. When multiple ligands are aligned, shared pharmacophoric features can be identified and mapped in three-dimensional space, highlighting interactions that could be critical for binding. Screening chemical libraries against pharmacophore models allows rapid detection of candidates expected to reproduce favorable binding interactions³⁷. Although pharmacophores do not yield explicit free energy values, their role in filtering compounds based on interaction complementarity makes them an effective first step in affinity-based prioritization³⁷.

2.2.3 Machine Learning and AI Integration

Recent advances in machine learning and artificial intelligence have substantially increased the predictive power of LBDD for binding affinity estimation⁴². Such techniques are particularly well-suited to capture dependencies between molecular descriptors and activity data. Beyond prediction, generative models can design new molecules optimized directly against desired affinity profiles⁴³.

Although LBDD techniques cannot directly compute thermodynamic binding free energies, they provide statistically grounded surrogates for these values. LBDD depends critically on the quantity, and diversity of experimental data. Robust models require sufficiently large sets of confirmed active and inactive molecules to avoid overfitting and ensure an appropriate applicability domain. When only a few active compounds are known, models become unreliable, and can mislead further researches^{9,44}.

2.3 SBDD Approaches for Calculating Binding Affinity

Structure-based drug design (SBDD) is a rational approach that uses the three-dimensional structure of a biological target, experimentally obtained or computationally predicted, to design and optimize molecules by modeling their interactions within the binding site⁴⁵. Unlike LBDD strategies, which require prior knowledge of active compounds, SBDD can be applied even when little or no information about known ligands is available, making it especially valuable for novel or poorly characterized targets.

2.3.1 Docking

Molecular docking is a computational method used to predict plausible ligand poses inside the binding site of a biological target, and to rank them with a scoring function intended to approximate binding affinity^{46,47}. Since this technique can nowadays evaluate huge numbers of molecules quickly, it has become the core engine behind Virtual Screening (VS). In a typical VS campaign, libraries containing millions, or sometimes even billions, of structures are computationally screened against the target. The results are ranked through their scores, allowing researchers to reduce this vast chemical space to a manageable number of promising candidates, that can be experimentally tested^{48–50}.

The docking process involves two main steps: a search algorithm, which explores possible orientations and conformations of the ligand in the binding site, and a scoring function, which estimates how favorable each pose is^{51,52}. Early approaches assumed both ligand and receptor were rigid, following a simple “lock-and-key” model. This was quickly recognized as unrealistic, since proteins and other biomolecules are dynamic and often adapt when binding partners interact⁵³. Therefore, modern docking software accounts for ligand flexibility and, to some degree, target flexibility, capturing the so-called “induced fit” effect^{54,55}.

Although the docking process is efficient, scoring functions struggle to precisely estimate binding affinities, since they are developed for prioritizing the rapid screening speed over accurate prediction of binding affinity^{56,57}. For this reason, docking is often combined with more advanced methods, which provide more reliable predictions and deeper insight into molecular interactions^{58,59}.

2.3.2 Interaction Fingerprints (IFPs)

An alternative way to analyze large databases is the use of Interaction Fingerprints (IFPs). This method translates the complex three-dimensional (3D) information of a ligand-target complex into a simplified digital format⁶⁰. By systematically identifying all

significant non-covalent contacts between a ligand and specific residues in the binding site, IFPs generate a unique “barcode” that characterizes a particular binding mode.

In practice, each position in a one-dimensional bit vector corresponds to a specific interaction type with a given residue. If the interaction is present, the bit is set to 1 (on); if absent, it remains 0 (off)⁶⁰. This simple binary representation enables rapid comparison of binding modes across large datasets.

The advantages of IFPs are particularly evident in virtual screening workflows, where they allow efficient post-processing of compounds that share similar binding features, thereby facilitating the identification of promising scaffolds⁶¹. Beyond static docking results, IFPs are also invaluable in the analysis of molecular dynamics (MD) simulations, providing insights into binding kinetics and conformational flexibility^{62,63}. However, it is important to note that, as for docking scoring functions, IFPs do not provide a quantitative estimate of binding affinity⁵⁷.

2.3.3 Molecular Dynamics (MD)

Molecular Dynamics (MD) simulations provide an atom-level way to study how biomolecules interact, by capturing both thermodynamic and kinetic aspects⁶⁴. At its core, MD works by applying Newton’s laws to represent how atoms move and change velocity, by following molecular mechanics force field rules that accounts for bonded interactions (bonds, angles, dihedrals) and non-bonded interactions (van der Waals and Coulomb forces)⁶⁵. The resulting trajectories reveal how targets, ligands, and solvent molecules fluctuate over time⁶⁶. Unlike docking approaches, MD fully captures flexibility, allowing observation of effects like induced fit, conformational selection, and solvent rearrangements⁶⁷.

At its core, MD is governed by Newton’s equations of motion:

$$a_i = \frac{F_i(t)}{m_i} = \frac{d^2x_i(t)}{dt^2}$$

where the acceleration a_i of the atom i is computed; m_i and x_i are the mass and position vector of the considered atom i , and F_i is the force exerted on the atom i at time t ^{68,69}.

The force F_i is derived from the gradient of the potential energy function U of the whole system, which is computed using molecular mechanics, and depends on the positions of all N atoms ($x_1, x_2, \dots, x_N$) within the system:

$$F_i = -\nabla_i U(x_1, x_2, \dots, x_N)$$

Thus, the trajectory of the system is obtained by integrating these equations for all N atoms over time. The integration updates atomic positions and velocities at each timestep, which is typically on the order of femtoseconds (2-4 fs).

Molecular mechanics approximate the behavior of molecules to classical mechanics rules, in order to reduce computational effort, for calculating the potential energy of the system. The most commonly used potential energy function are described within the used force field, and can be described using the following equation:

$$U(x) = U_{Bonded} + U_{Non-Bonded}$$

The potential energy derived from covalent and non-covalent interactions consists of a sum of energy contributions that are depicted in Figure 1, and represented by the following equations:

$$U_{Bonded} = \sum_{Bonds} k_b(b - b_0)^2 + \sum_{Angles} k_\theta(\theta - \theta_0)^2 + \sum_{Torsions} k_\phi[\cos(n\phi + \delta) \pm 1]$$

where k , b_0 , and θ_0 terms represents all the parameters that constitutes the forcefield, n is the multiplicity, and δ the phase of the torsional periodic function^{70,71}.

For the non-bonded interactions, the energy contribution is calculated for each ij pair of atoms, and is constituted by an electrostatic factor (based on Coulomb's law), and a Lennard-Jones 6-12 potential^{70,71}:

$$U_{Non-Bonded} = \sum_{Non-Bond\ Pairs} \left[\frac{q_i q_j}{\epsilon x_{ij}} + \frac{A_{ij}}{x_{ij}^{12}} - \frac{C_{ij}}{x_{ij}^6} \right]$$

where q terms represent the atoms partial charges, and ϵ represents the dielectric constant, while A and C terms represent repulsive and attractive contributions, respectively.

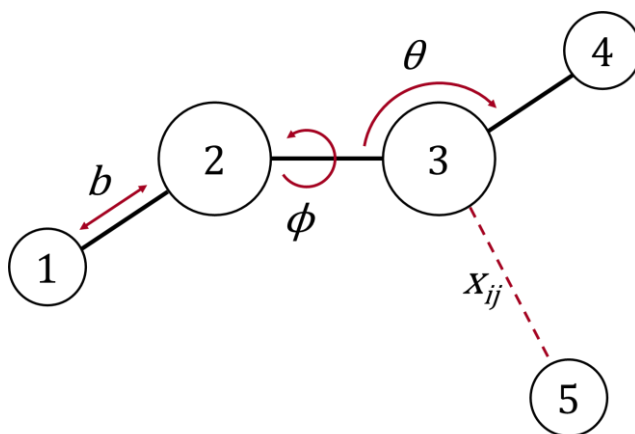


Figure 1. Schematic representation of forcefield interactions. Bonded interactions are indicated by solid lines (1-2, 2-3, 3-4), non bonded interactions (3-5) by a dashed line⁷¹.

Therefore, MD provides the mechanistic detail required to make affinity predictions, albeit at a significant computational cost⁷²⁻⁷⁶. However, the MD predictive reliability is affected by the number of approximations characterizing molecular mechanics and depends on the accuracy of the force field used^{77,78}.

2.4 Binding Affinity and Molecular Dynamics

Some biological processes are inherently slow, like large conformational rearrangements in proteins, while others are fast, but occur only sporadically, due to their high activation energy^{33,79}. Conventional molecular dynamics (MD) struggles with both regimes: typical simulations span nanoseconds, but many biologically relevant events require millisecond or longer timescales⁸⁰, as per the binding or the unbinding of a ligand with its biological target. Thus, it is not easy to study these processes for determining binding affinity with classical MD simulations.

To bridge this gap, a spectrum of MD-based strategies has emerged to estimate ligand binding affinities, each trading off speed and accuracy in different ways. The simplest example is characterized by endpoint approaches such as MM-PBSA and MM-GBSA, most often applied as post-processing on standard MD trajectories^{26,74,75}. Instead, with an increased computational effort, enhanced-sampling techniques, like steered MD⁸¹, scaled MD⁸², metadynamics⁸³, and other biased techniques, can accelerate rare binding or unbinding events^{26,84,85}. For more rigorous affinity predictions, umbrella sampling⁸⁶ and alchemical free-energy perturbation (FEP)^{26,70,87–90} explicitly reconstruct the binding free-energy landscape. When sampling is adequate, they yield quantitative estimates of the binding free energy.

However, binding free energy alone does not capture a drug's affinity. A compound remains pharmacologically active only while it stays bound to its target, making residence time equally crucial^{28,79}. To study this kinetic dimension, specialized MD approaches, such as τ -RAMD⁹¹ deliberately accelerate unbinding, enabling residence-time estimates far beyond the reach of conventional MD.

Taken together, these methods deliver a more complete view of the thermodynamics and kinetics of ligand-receptor interactions, guiding the design of molecules with both high affinity, and great ligand-target complex duration.

2.4.1 Free Energy Perturbation (FEP)

Alchemical free energy methods enable the calculation of binding free energies by simulating non-physical transformations between two different molecular states (A, and B)^{92,93}. These states can be represented by potential energy functions U_A and U_B , and their free energy difference can be calculated as

$$\Delta G = -\beta^{-1} \ln \frac{Z_B}{Z_A} = \left(\frac{\text{Zwanzig Equation}}{-\ln \langle \exp [-\beta(U_B - U_A)] \rangle_A}{\beta} \right)$$

where $\beta=1/RT$, while Z represents the configurational partition function of the relevant state, and the $\langle \rangle$ brackets denote an ensemble average of configurations subject to

potential U_A .⁹⁴ The Zwanzig equation requires significant overlap between the probability distributions of the two systems, U_A and U_B , to produce reliable results⁹⁵. Applying the equation directly from a single simulation using only the U_A potential is generally ineffective, since there is not enough phase space overlap. To overcome this, a coupling parameter λ is used to create a sequence of intermediate potential functions between the two potentials. These intermediates ensure sufficient overlap between neighboring states, allowing the Zwanzig equation to be applied incrementally to calculate accurate free energy differences.

In the context of computer-aided drug design, when calculating binding free energy with free energy perturbation (FEP), state A represents the ligand of interest. This state is transformed into the B state, which can vary depending on the specific FEP method used.

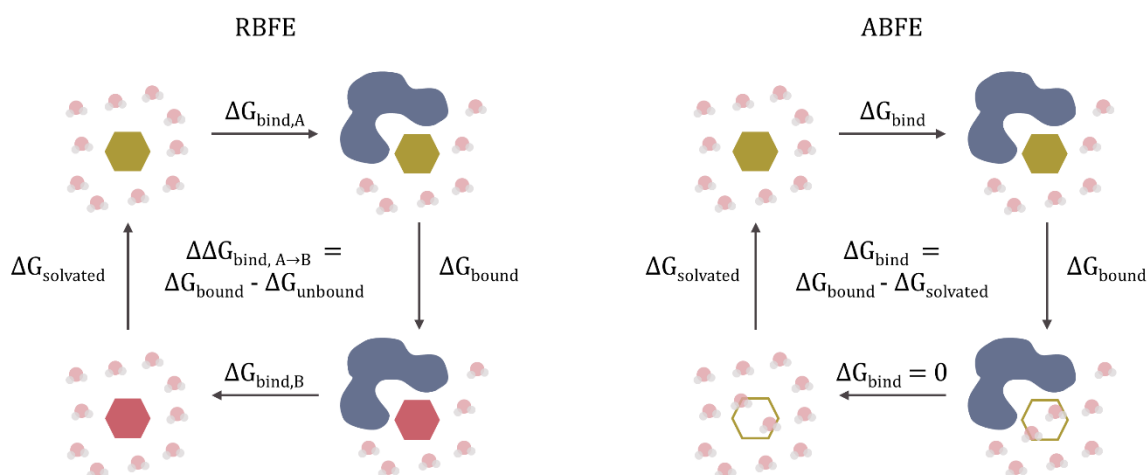


Figure 2. Schematic representation of a Relative (RBFE) and Absolute (ABFE) Binding Free Energy calculation using the corresponding alchemical transformation approach.

Relative Binding Free Energy (RBFE) methods estimate the difference in binding free energy between two structurally similar ligands (A and B are states with two different ligands) by simulating the transformation of one ligand into the other⁹⁶. Structural similarity is essential to ensure sufficient overlap in phase space, thereby minimizing the number of λ intermediate states required, since a large number of intermediates would significantly increase computational cost. This transformation is carried out in both the bound state (ligand-target complex) and the unbound state (ligand in solvent). The difference between the free energy changes in these two environments yields the relative binding free energy (Figure 2A). The accuracy and efficiency of RBFE calculations are highly dependent on the chemical similarity between the two ligands (i.e. A and B), which need to be congener. Greater similarity facilitates better phase space overlap, reduces sampling difficulties, and improves convergence. Because reliable results are typically obtained only when comparing similar compounds, ligands are often organized

into a perturbation network on the basis of their similarity, as illustrated in Figure 3. Hysteresis, whether observed within cycles of the perturbation network or between forward and backward transformations, can be monitored to identify problematic perturbations.

RBFE is especially valuable during lead optimization, where incremental chemical changes are introduced to a shared molecular scaffold, enabling precise comparisons of binding affinities.

On the other hand, ABFE methods attempt to calculate the true binding free energy value of a ligand without referencing it to a similar compound^{97,98}. This is achieved by gradually decoupling the ligand from the protein binding site and solvent environment (Figure 2B), often using a series of restrained simulations to control for translational and rotational degrees of freedom⁹⁹. A critical consideration during this process is the solvation of the cavity previously occupied by the ligand, which can be a significant source of error, especially in buried binding sites that are poorly accessible to solvent molecules¹⁰⁰. While ABFE calculations are more computationally demanding and sensitive to sampling challenges, they offer the advantage of being applicable to chemically diverse ligands, including cases where RBFE is not suitable, a feature that can be very useful during the hit-to-lead stage of the drug discovery process.

To summarize, both RBFE and ABFE require rigorous sampling, accurate force fields, and carefully designed protocols to produce reliable results. RBFE is generally more efficient and well-suited for congeneric series, whereas ABFE provides a more flexible framework, for estimating absolute binding affinities, at the cost of increased computational effort^{93,101}. The continued development of these methods and the software that implements them^{102,103}, coupled with increasing computational power, promises to further enhance their predictive accuracy and broaden their impact on the rational design of new medicines.

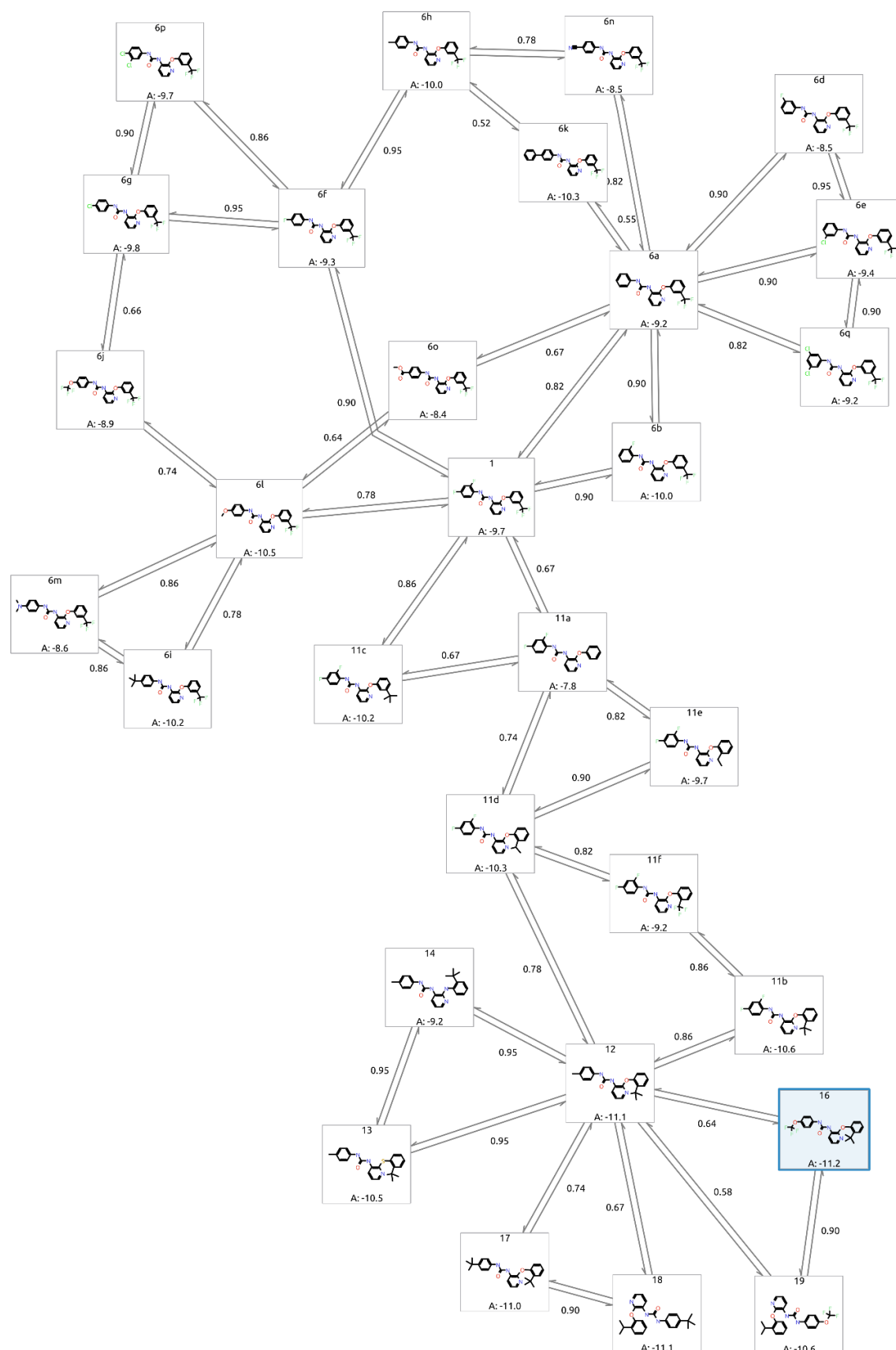


Figure 3. Example of a ligand network used in Relative Binding Free Energy (RBFE) protocols. (Image produced with Flare FEP¹⁰⁴)

References

1. Sarvagalla, S.; Kondapuram, S. K.; Devi, R. V.; Coumar, M. S. Chapter 9 - Resources for Docking-Based Virtual Screening. In *Molecular Docking for Computer-Aided Drug Design*; Coumar, M. S., Ed.; Academic Press, 2021; pp 179–203. <https://doi.org/10.1016/B978-0-12-822312-3.00002-3>.
2. Myers, S.; Baker, A. Drug Discovery—an Operating Model for a New Era. *Nat. Biotechnol.* 2001, 19 (8), 727–730. <https://doi.org/10.1038/90765>.
3. Moses, H.; Dorsey, E. R.; Matheson, D. H. M.; Thier, S. O. Financial Anatomy of Biomedical Research. *JAMA* 2005, 294 (11), 1333–1342. <https://doi.org/10.1001/jama.294.11.1333>.
4. Leelananda, S. P.; Lindert, S. Computational Methods in Drug Discovery. *Beilstein J. Org. Chem.* 2016, 12 (1), 2694–2718. <https://doi.org/10.3762/bjoc.12.267>.
5. Dowden, H.; Munro, J. Trends in Clinical Success Rates and Therapeutic Focus. *Nat. Rev. Drug Discov.* 2019, 18 (7), 495–496. <https://doi.org/10.1038/d41573-019-00074-z>.
6. Choudhary, H. B.; Patekar, R. R.; Rudrapal, M. Chapter 11 - Fundamentals of Drug Design and Discovery. In *Biochemical and Molecular Pharmacology in Drug Discovery*; Rudrapal, M., Egbuna, C., Cho, W. C., Eds.; Drug Discovery Update; Elsevier, 2024; pp 239–249. <https://doi.org/10.1016/B978-0-443-16013-4.00011-7>.
7. Das, N. R.; Jena, G. K. Chapter 15 - Bioinformatics and CADD Approaches in Drug Discovery. In *Biochemical and Molecular Pharmacology in Drug Discovery*; Rudrapal, M., Egbuna, C., Cho, W. C., Eds.; Drug Discovery Update; Elsevier, 2024; pp 313–321. <https://doi.org/10.1016/B978-0-443-16013-4.00015-4>.
8. Mihai, D. P.; Nitulescu, G. M. Computer-Aided Drug Design and Drug Discovery. *Pharmaceuticals* 2025, 18 (3), 436. <https://doi.org/10.3390/ph18030436>.
9. Purohit, D.; Makhija, M.; Pandey, P.; Kumar, S.; Kumar, S.; Dutt, R.; Kaushik, D.; Kumar, P.; Kumar, S. Role of Computer-Aided Drug Design in the Discovery and Development of New Medicinal Agents a Review. *J. Med. Pharm. MAY-JUNE 2022 (VOLUME 11 – ISSUE 3 MAY-JUNE 2022)*. <https://doi.org/10.55522/jmpas.V11i3.2300>.
10. Anwar, T.; Kumar, P.; Khan, A. U. Chapter 1 - Modern Tools and Techniques in Computer-Aided Drug Design. In *Molecular Docking for Computer-Aided Drug Design*; Coumar, M. S., Ed.; Academic Press, 2021; pp 1–30. <https://doi.org/10.1016/B978-0-12-822312-3.00011-4>.
11. Manas, E. S.; Green, D. V. S. CADD Medicine: Design Is the Potion That Can Cure My Disease. *J. Comput. Aided Mol. Des.* 2017, 31 (3), 249–253. <https://doi.org/10.1007/s10822-016-0004-3>.
12. Rajkishan, T.; Rachana, A.; Shruti, S.; Bhumi, P.; Patel, D. Computer-Aided Drug Designing. In *Advances in Bioinformatics*; Singh, V., Kumar, A., Eds.; Springer: Singapore, 2021; pp 151–182. https://doi.org/10.1007/978-981-33-6191-1_9.
13. Cushman, D. W.; Cheung, H. S.; Sabo, E. F.; Ondetti, M. A. Design of Potent Competitive Inhibitors of Angiotensin-Converting Enzyme. Carboxyalkanoyl and Mercaptoalkanoyl Amino Acids. *Biochemistry* 1977, 16 (25), 5484–5491. <https://doi.org/10.1021/bi00644a014>.
14. Kumar, B.; Chawla, P. A.; Chawla, V. Computer Aided Drug Design. In *Computer Aided Pharmaceutics and Drug Delivery: An Application Guide for Students and Researchers of Pharmaceutical Sciences*; Saharan, V. A., Ed.; Springer Nature: Singapore, 2022; pp 505–541. https://doi.org/10.1007/978-981-16-5180-9_16.
15. Elengoe, A. Chapter 11 - Programming, Tools, and Software. In *Systems Biology and In-Depth Applications for Unlocking Diseases*; Sokouti, B., Ed.; Academic Press, 2025; pp 149–155. <https://doi.org/10.1016/B978-0-443-22326-6.00011-0>.
16. Baldwin, J. J.; Ponticello, G. S.; Anderson, P. S.; Christy, M. E.; Murcko, M. A.; Randall, W. C.; Schwam, H.; Sugrue, M. F.; Gautheron, P. Thienothiopyran-2-Sulfonamides: Novel Topically Active Carbonic Anhydrase Inhibitors for the Treatment of Glaucoma. *J. Med. Chem.* 1989, 32 (12), 2510–2513. <https://doi.org/10.1021/jm00132a003>.
17. James, J. S. Saquinavir (Invirase): First Protease Inhibitor Approved—Reimbursement, Information Hotline Numbers. *AIDS Treat. News* 1995, No. no 237, 1–2.
18. Muegge, I.; Bergner, A.; Kriegl, J. M. Computer-Aided Drug Design at Boehringer Ingelheim. *J. Comput. Aided Mol. Des.* 2017, 31 (3), 275–285. <https://doi.org/10.1007/s10822-016-9975-3>.
19. Danel, T.; Łęski, J.; Podlowska, S.; Podolak, I. T. Docking-Based Generative Approaches in the Search for New Drug Candidates. *Drug Discov. Today* 2023, 28 (2), 103439. <https://doi.org/10.1016/j.drudis.2022.103439>.
20. Yang, W.; Wang, Y.; Han, D.; Tang, W.; Sun, L. Recent Advances in Application of Computer-Aided Drug Design in Anti-COVID-19 Virals Drug Discovery. *Biomed. Pharmacother.* 2024, 173, 116423. <https://doi.org/10.1016/j.biopha.2024.116423>.
21. Niazi, S. K.; Mariam, Z. Computer-Aided Drug Design and Drug Discovery: A Prospective Analysis. *Pharmaceuticals* 2024, 17 (1), 22. <https://doi.org/10.3390/ph17010022>.
22. Wu, Z.; Chen, S.; Wang, Y.; Li, F.; Xu, H.; Li, M.; Zeng, Y.; Wu, Z.; Gao, Y. Current Perspectives and Trend of Computer-Aided Drug Design: A Review and Bibliometric Analysis. *Int. J. Surg.* 2024, 110 (6), 3848. <https://doi.org/10.1097/JS9.0000000000001289>.
23. Fukunishi, Y.; Mashimo, T.; Misoo, K.; Wakabayashi, Y.; Miyaki, T.; Ohta, S.; Nakamura, M.; Ikeda, K. Miscellaneous Topics in Computer-Aided Drug Design: Synthetic Accessibility and GPU Computing, and Other Topics. *Curr. Pharm. Des.* 22 (23), 3555–3568. <https://doi.org/10.2174/1381612822666160414142547>.
24. Kairys, V.; Baranauskienė, L.; Kazlauskienė, M.; Matulis, D.; Kazlauskas, E. Binding Affinity in Drug Design: Experimental and Computational Techniques. *Expert Opin. Drug Discov.* 2019, 14 (8), 755–768. <https://doi.org/10.1080/17460441.2019.1623202>.

25. Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E. Molecular Determinants of Drug-Receptor Binding Kinetics. *Drug Discov. Today* 2013, *18* (13–14), 667–673. <https://doi.org/10.1016/J.DRUDIS.2013.02.007>.
26. Decherchi, S.; Cavalli, A. Thermodynamics and Kinetics of Drug-Target Binding by Molecular Simulation. *Chem. Rev.* 2020, *120* (23), 12788–12833. <https://doi.org/10.1021/acs.chemrev.0c00534>.
27. Ehrlich, P. Address in Pathology, ON CHEMIOTHERAPY. 1913. <https://doi.org/10.1136/bmj.2.2746.353>.
28. Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug-Target Residence Time and Its Implications for Lead Optimization. *Nat. Rev. Drug Discov.* 2006 *59* 2006, *5* (9), 730–739. <https://doi.org/10.1038/nrd2082>.
29. Turnbull, W. B.; Daranas, A. H. On the Value of c : Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry? *J. Am. Chem. Soc.* 2003, *125* (48), 14859–14866. <https://doi.org/10.1021/ja036166s>.
30. Guo, D.; Hillger, J. M.; Ijzerman, A. P.; Heitman, L. H. Drug-Target Residence Time—A Case for G Protein-Coupled Receptors. *Med. Res. Rev.* 2014, *34* (4), 856–892. <https://doi.org/10.1002/MED.21307>.
31. Bernetti, M.; Masetti, M.; Rocchia, W.; Cavalli, A. Kinetics of Drug Binding and Residence Time. *Annu. Rev. Phys. Chem.* 2019, *70*, 143–171. <https://doi.org/10.1146/ANNUREV-PHYSCHEM-042018-052340>.
32. Anderson, A. C. The Process of Structure-Based Drug Design. *Chem. Biol.* 2003, *10* (9), 787–797. <https://doi.org/10.1016/j.chembiol.2003.09.002>.
33. Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. Kinetics of Protein-Ligand Unbinding: Predicting Pathways, Rates, and Rate-Limiting Steps. *Proc. Natl. Acad. Sci.* 2015, *112* (5), E386–E391. <https://doi.org/10.1073/pnas.1424461112>.
34. Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science* 2004, *303* (5665), 1813–1818. <https://doi.org/10.1126/science.1096361>.
35. Schaller, D.; Šribar, D.; Noonan, T.; Deng, L.; Nguyen, T. N.; Pach, S.; Machalz, D.; Bermudez, M.; Wolber, G. Next Generation 3D Pharmacophore Modeling. *WIREs Comput. Mol. Sci.* 2020, *10* (4), e1468. <https://doi.org/10.1002/wcms.1468>.
36. Giordano, D.; Biancaniello, C.; Argenio, M. A.; Facchiano, A. Drug Design by Pharmacophore and Virtual Screening Approach. *Pharmaceutics* 2022, *15* (5), 646. <https://doi.org/10.3390/ph15050646>.
37. Leach, A. R.; Gillet, V. J.; Lewis, R. A.; Taylor, R. Three-Dimensional Pharmacophore Methods in Drug Discovery. *J. Med. Chem.* 2010, *53* (2), 539–558. <https://doi.org/10.1021/jm900817u>.
38. Acharya, C.; Coop, A.; Polli, J. E.; MacKerell, A. D. Recent Advances in Ligand-Based Drug Design: Relevance and Utility of the Conformationally Sampled Pharmacophore Approach. *Curr. Comput. Aided Drug Des.* 7 (1), 10–22. <https://doi.org/10.2174/157340911793743547>.
39. Ekins, S.; Mestres, J.; Testa, B. In Silico Pharmacology for Drug Discovery: Methods for Virtual Ligand Screening and Profiling. *Br. J. Pharmacol.* 2007, *152* (1), 9–20. <https://doi.org/10.1038/sj.bjp.0707305>.
40. Cottrell, S. J.; Gillet, V. J.; Taylor, R.; Wilton, D. J. Generation of Multiple Pharmacophore Hypotheses Using Multiobjective Optimisation Techniques. *J. Comput. Aided Mol. Des.* 2004, *18* (11), 665–682. <https://doi.org/10.1007/s10822-004-5523-7>.
41. Chang, Y.; Hawkins, B. A.; Du, J. J.; Groundwater, P. W.; Hibbs, D. E.; Lai, F. A Guide to In Silico Drug Design. *Pharmaceutics* 2023, *15* (1), 49. <https://doi.org/10.3390/pharmaceutics15010049>.
42. Dai, J.; Zhou, Z.; Zhao, Y.; Kong, F.; Zhai, Z.; Zhu, Z.; Cai, J.; Huang, S.; Xu, Y.; Sun, T. Combined Usage of Ligand- and Structure-Based Virtual Screening in the Artificial Intelligence Era. *Eur. J. Med. Chem.* 2025, *283*, 117162. <https://doi.org/10.1016/j.ejmech.2024.117162>.
43. Mouchlis, V. D.; Afantitis, A.; Serra, A.; Fratello, M.; Papadiamantis, A. G.; Aidinis, V.; Lynch, I.; Greco, D.; Melagraki, G. Advances in De Novo Drug Design: From Conventional to Machine Learning Methods. *Int. J. Mol. Sci.* 2021, *22* (4), 1676. <https://doi.org/10.3390/ijms22041676>.
44. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W. Computational Methods in Drug Discovery. *Pharmacol. Rev.* 2014, *66* (1), 334–395. <https://doi.org/10.1124/pr.112.007336>.
45. Gurung, A. B.; Ali, M. A.; Lee, J.; Farah, M. A.; Al-Anazi, K. M. An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19. *BioMed Res. Int.* 2021, *2021* (1), 8853056. <https://doi.org/10.1155/2021/8853056>.
46. Pinzi, L.; Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *Int. J. Mol. Sci.* 2019, *20* (18), 4331. <https://doi.org/10.3390/ijms20184331>.
47. Torres, P. H. M.; Sodero, A. C. R.; Jofily, P.; Silva-Jr, F. P. Key Topics in Molecular Docking for Drug Design. *Int. J. Mol. Sci.* 2019, *20* (18), 4574. <https://doi.org/10.3390/ijms20184574>.
48. Adelusi, T. I.; Oyedele, A.-Q. K.; Boyenle, I. D.; Ogunlana, A. T.; Adeyemi, R. O.; Ukachi, C. D.; Idris, M. O.; Olaoba, O. T.; Adedotun, I. O.; Kolawole, O. E.; Xiaoxing, Y.; Abdul-Hammed, M. Molecular Modeling in Drug Discovery. *Inform. Med. Unlocked* 2022, *29*, 100880. <https://doi.org/10.1016/j.imu.2022.100880>.
49. Bender, B. J.; Gahbauer, S.; Luttens, A.; Lyu, J.; Webb, C. M.; Stein, R. M.; Fink, E. A.; Balias, T. E.; Carlsson, J.; Irwin, J. J.; Shoichet, B. K. A Practical Guide to Large-Scale Docking. *Nat. Protoc.* 2021, *16* (10), 4799–4832. <https://doi.org/10.1038/s41596-021-00597-z>.
50. Samanta, R.; Pradhan, K. K.; Sen, D.; Kar, S.; Ghosh, M. Structure-Based Drug Design-Guided Identification of Estrogen Receptor Binders. *Mol. Divers.* 2024, *28* (3), 1291–1303. <https://doi.org/10.1007/s11030-023-10657-z>.
51. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nat. Rev. Drug Discov.* 2004, *3* (11), 935–949. <https://doi.org/10.1038/nrd1549>.
52. Huang, S.-Y.; Zou, X. Advances and Challenges in Protein-Ligand Docking. *Int. J. Mol. Sci.* 2010, *11* (8), 3016–3034. <https://doi.org/10.3390/ijms11083016>.

53. Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput. Aided-Drug Des.* 2012, 7 (2), 146–157. <https://doi.org/10.2174/157340911795677602>.
54. De Vivo, M.; Cavalli, A. Recent Advances in Dynamic Docking for Drug Discovery. *WIREs Comput. Mol. Sci.* 2017, 7 (6), e1320. <https://doi.org/10.1002/wcms.1320>.
55. Salmaso, V.; Sturlese, M.; Cuzzolin, A.; Moro, S. Combining Self- and Cross-Docking as Benchmark Tools: The Performance of DockBench in the D3R Grand Challenge 2. *J. Comput. Aided Mol. Des.* 2018, 32 (1), 251–264. <https://doi.org/10.1007/s10822-017-0051-4>.
56. Chaput, L.; Mouawad, L. Efficient Conformational Sampling and Weak Scoring in Docking Programs? Strategy of the Wisdom of Crowds. *J. Cheminformatics* 2017, 9 (1), 37. <https://doi.org/10.1186/s13321-017-0227-x>.
57. Wang, D. D.; Chan, M.-T.; Yan, H. Structure-Based Protein–Ligand Interaction Fingerprints for Binding Affinity Prediction. *Comput. Struct. Biotechnol. J.* 2021, 19, 6291–6300. <https://doi.org/10.1016/j.csbj.2021.11.018>.
58. Yu, W.; MacKerell, A. D. Computer-Aided Drug Design Methods. In *Antibiotics: Methods and Protocols*; Sass, P., Ed.; Springer: New York, NY, 2017; pp 85–106. https://doi.org/10.1007/978-1-4939-6634-9_5.
59. Lazim, R.; Suh, D.; Choi, S. Advances in Molecular Dynamics Simulations and Enhanced Sampling Methods for the Study of Protein Systems. *Int. J. Mol. Sci.* 2020, 21 (17), 6339. <https://doi.org/10.3390/ijms21176339>.
60. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Implementing a Scoring Function Based on Interaction Fingerprint for Autogrow4: Protein Kinase CK1δ as a Case Study. *Front. Mol. Biosci.* 2022, 9, 629. <https://doi.org/10.3389/fmolb.2022.909499>.
61. Calenda, S.; Catarzi, D.; Varano, F.; Vigiani, E.; Volpini, R.; Lambertucci, C.; Spinaci, A.; Trevisan, L.; Grieco, I.; Federico, S.; Spalluto, G.; Novello, G.; Salmaso, V.; Moro, S.; Colotta, V. Structural Investigations on 2-Amidobenzimidazole Derivatives as New Inhibitors of Protein Kinase CK1 Delta. *Pharmaceuticals* 2024, 17 (4), 468. <https://doi.org/10.3390/ph17040468>.
62. Palazzotti, D.; Bissaro, M.; Bolcato, G.; Astolfi, A.; Felicetti, T.; Sabatini, S.; Sturlese, M.; Cecchetti, V.; Barreca, M. L.; Moro, S. Deciphering the Molecular Recognition Mechanism of Multidrug Resistance Staphylococcus Aureus NorA Efflux Pump Using a Supervised Molecular Dynamics Approach. *Int. J. Mol. Sci.* 2019, 20 (16), 4041. <https://doi.org/10.3390/ijms20164041>.
63. Spinaci, A.; Buccioni, M.; Catarzi, D.; Cui, C.; Colotta, V.; Dal Ben, D.; Cescon, E.; Francucci, B.; Grieco, I.; Lambertucci, C.; Marucci, G.; Bassani, D.; Pavan, M.; Varano, F.; Federico, S.; Spalluto, G.; Moro, S.; Volpini, R. “Dual Anta-Inhibitors” of the A2A Adenosine Receptor and Casein Kinase CK1delta: Synthesis, Biological Evaluation, and Molecular Modeling Studies. *Pharmaceuticals* 2023, 16 (2), 167. <https://doi.org/10.3390/ph16020167>.
64. van Gunsteren, W. F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D. P.; Glättli, A.; Hünenberger, P. H.; Kastenholz, M. A.; Oostenbrink, C.; Schenk, M.; Trzesniak, D.; van der Vegt, N. F. A.; Yu, H. B. Biomolecular Modeling: Goals, Problems, Perspectives. *Angew. Chem. Int. Ed.* 2006, 45 (25), 4064–4092. <https://doi.org/10.1002/anie.200502655>.
65. Ganesan, A.; Coote, M. L.; Barakat, K. Molecular Dynamics-Driven Drug Discovery: Leaping Forward with Confidence. *Drug Discov. Today* 2017, 22 (2), 249–269. <https://doi.org/10.1016/j.drudis.2016.11.001>.
66. Katiyar, R. S.; Jha, P. K. Molecular Simulations in Drug Delivery: Opportunities and Challenges. *WIREs Comput. Mol. Sci.* 2018, 8 (4), e1358. <https://doi.org/10.1002/wcms.1358>.
67. Śledź, P.; Cafilisch, A. Protein Structure-Based Drug Design: From Docking to Molecular Dynamics. *Curr. Opin. Struct. Biol.* 2018, 48, 93–102. <https://doi.org/10.1016/j.sbi.2017.10.010>.
68. Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* 2018, 9. <https://doi.org/10.3389/fphar.2018.00923>.
69. Santamaria, R. Foundations of Newtonian Dynamics. In *Molecular Dynamics*; Santamaria, R., Ed.; Springer Nature Switzerland: Cham, 2023; pp 23–40. https://doi.org/10.1007/978-3-031-37042-7_2.
70. Lamb, M. L.; Jorgensen, W. L. Computational Approaches to Molecular Recognition. *Curr. Opin. Chem. Biol.* 1997, 1 (4), 449–457. [https://doi.org/10.1016/S1367-5931\(97\)80038-5](https://doi.org/10.1016/S1367-5931(97)80038-5).
71. Ponder, J. W.; Case, D. A. Force Fields for Protein Simulations. In *Advances in Protein Chemistry*; Protein Simulations; Academic Press, 2003; Vol. 66, pp 27–85. [https://doi.org/10.1016/S0065-3233\(03\)66002-X](https://doi.org/10.1016/S0065-3233(03)66002-X).
72. Wong, C. F.; Andrew McCammon, J. Computer Simulation and the Design of New Biological Molecules. *Isr. J. Chem.* 1986, 27 (2), 211–215. <https://doi.org/10.1002/ijch.198600030>.
73. Durrant, J. D.; McCammon, J. A. Molecular Dynamics Simulations and Drug Discovery. *BMC Biol.* 2011, 9 (1), 71. <https://doi.org/10.1186/1741-7007-9-71>.
74. Gallicchio, E.; Levy, R. M. Recent Theoretical and Computational Advances for Modeling Protein–Ligand Binding Affinities. In *Advances in Protein Chemistry and Structural Biology*; Christov, C., Ed.; Computational chemistry methods in structural biology; Academic Press, 2011; Vol. 85, pp 27–80. <https://doi.org/10.1016/B978-0-12-386485-7.00002-8>.
75. Procacci, P. Alchemical Determination of Drug-Receptor Binding Free Energy: Where We Stand and Where We Could Move To. *J. Mol. Graph. Model.* 2017, 71, 233–241. <https://doi.org/10.1016/j.jmgm.2016.11.018>.
76. Li, Z.; Huang, Y.; Wu, Y.; Chen, J.; Wu, D.; Zhan, C.-G.; Luo, H.-B. Absolute Binding Free Energy Calculation and Design of a Subnanomolar Inhibitor of Phosphodiesterase-10. *J. Med. Chem.* 2019, 62 (4), 2099–2111. <https://doi.org/10.1021/acs.jmedchem.8b01763>.
77. Mackerell Jr., A. D. Empirical Force Fields for Biological Macromolecules: Overview and Issues. *J. Comput. Chem.* 2004, 25 (13), 1584–1604. <https://doi.org/10.1002/jcc.20082>.

78. Cole, D. J.; Tirado-Rives, J.; Jorgensen, W. L. Molecular Dynamics and Monte Carlo Simulations for Protein–Ligand Binding and Inhibitor Design. *Biochim. Biophys. Acta BBA - Gen. Subj.* 2015, 1850 (5), 966–971. <https://doi.org/10.1016/j.bbagen.2014.08.018>.
79. Copeland, R. A. The Drug–Target Residence Time Model: A 10-Year Retrospective. *Nat. Rev. Drug Discov.* 2015 152 2015, 15 (2), 87–95. <https://doi.org/10.1038/nrd.2015.18>.
80. Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and Mechanism of Drug Binding to G-Protein-Coupled Receptors. *Proc. Natl. Acad. Sci.* 2011, 108 (32), 13118–13123. <https://doi.org/10.1073/pnas.1104614108>.
81. Park, S.; Schulten, K. Calculating Potentials of Mean Force from Steered Molecular Dynamics Simulations. *J. Chem. Phys.* 2004, 120 (13), 5946–5961. <https://doi.org/10.1063/1.1651473>.
82. Sinko, W.; Miao, Y.; De Oliveira, C. A. F.; McCammon, J. A. Population Based Reweighting of Scaled Molecular Dynamics. *J. Phys. Chem. B* 2013, 117 (42), 12759. <https://doi.org/10.1021/JP401587E>.
83. Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci.* 2002, 99 (20), 12562–12566. <https://doi.org/10.1073/pnas.202427399>.
84. Deganutti, G.; Moro, S.; Reynolds, C. A. A Supervised Molecular Dynamics Approach to Unbiased Ligand-Protein Unbinding. *J. Chem. Inf. Model.* 2020, 60 (3). <https://doi.org/10.1021/ACS.JCIM.9B01094>.
85. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Qualitative Estimation of Protein–Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2022, 62 (22), 5715–5728. <https://doi.org/10.1021/acs.jcim.2c00995>.
86. Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* 1977, 23 (2), 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
87. Adcock, S. A.; McCammon, J. A. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.* 2006, 106 (5), 1589–1615. <https://doi.org/10.1021/cr040426m>.
88. Michel, J.; Essex, J. W. Prediction of Protein–Ligand Binding Affinity by Free Energy Simulations: Assumptions, Pitfalls and Expectations. *J. Comput. Aided Mol. Des.* 2010, 24 (8), 639–658. <https://doi.org/10.1007/s10822-010-9363-3>.
89. Zara, L.; Efrém, N.-L.; van Muijlwijk-Koezen, J. E.; de Esch, I. J. P.; Zarzycka, B. Progress in Free Energy Perturbation: Options for Evolving Fragments. *Drug Discov. Today Technol.* 2021, 40, 36–42. <https://doi.org/10.1016/j.ddtec.2021.10.001>.
90. Muegge, I.; Hu, Y. Recent Advances in Alchemical Binding Free Energy Calculations for Drug Discovery. *ACS Med. Chem. Lett.* 2023, 14 (3), 244–250. <https://doi.org/10.1021/acsmchemlett.2c00541>.
91. Kokh, D. B.; Amaral, M.; Bomke, J.; Grädler, U.; Musil, D.; Buchstaller, H. P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F.; Bianciotto, M.; Rak, A.; Wade, R. C. Estimation of Drug–Target Residence Times by τ -Random Acceleration Molecular Dynamics Simulations. *J. Chem. Theory Comput.* 2018, 14 (7), 3859–3869. <https://doi.org/10.1021/ACS.JCTC.8B00230>.
92. Mobley, D. L.; Klimovich, P. V. Perspective: Alchemical Free Energy Calculations for Drug Discovery. *J. Chem. Phys.* 2012, 137 (23), 230901. <https://doi.org/10.1063/1.4769292>.
93. Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* 2017, 57 (12), 2911–2937. <https://doi.org/10.1021/acs.jcim.7b00564>.
94. Jespers, W.; Åqvist, J.; Gutiérrez-de-Terán, H. Free Energy Calculations for Protein–Ligand Binding Prediction. In *Protein–Ligand Interactions and Drug Design*; Ballante, F., Ed.; Methods in Molecular Biology; Springer US: New York, NY, 2021; pp 203–226. https://doi.org/10.1007/978-1-0716-1209-5_12.
95. Pohorille, A.; Jarzynski, C.; Chipot, C. Good Practices in Free-Energy Calculations. *J. Phys. Chem. B* 2010, 114 (32), 10235–10253. <https://doi.org/10.1021/jp102971x>.
96. Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* 2015, 137 (7), 2695–2703. <https://doi.org/10.1021/ja512751q>.
97. Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Accurate Calculation of the Absolute Free Energy of Binding for Drug Molecules. *Chem. Sci.* 2015, 7 (1), 207–218. <https://doi.org/10.1039/C5SC02678D>.
98. Khalak, Y.; Tresadern, G.; Aldeghi, M.; Baumann, H. M.; Mobley, D. L.; de Groot, B. L.; Gapsys, V. Alchemical Absolute Protein–Ligand Binding Free Energies for Drug Design. *Chem. Sci.* 2021, 12 (41), 13958–13971. <https://doi.org/10.1039/d1sc03472c>.
99. Borech, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. *J. Phys. Chem. B* 2003, 107 (35), 9535–9551. <https://doi.org/10.1021/jp0217839>.
100. Clark, F.; Robb, G.; Cole, D. J.; Michel, J. Comparison of Receptor–Ligand Restraint Schemes for Alchemical Absolute Binding Free Energy Calculations. *J. Chem. THEORY Comput.* 2023. <https://doi.org/10.1021/acs.jctc.3c00139>.
101. Gapsys, V.; Michielssens, S.; Seeliger, D.; de Groot, B. L. Accurate and Rigorous Prediction of the Changes in Protein Free Energies in a Large-Scale Mutation Scan. *Angew. Chem. Int. Ed.* 2016, 55 (26), 7364–7368. <https://doi.org/10.1002/anie.201510054>.
102. DeLano, W. L. Unraveling Hot Spots in Binding Interfaces: Progress and Challenges. *Curr. Opin. Struct. Biol.* 2002, 12 (1), 14–20. [https://doi.org/10.1016/S0959-440X\(02\)00283-X](https://doi.org/10.1016/S0959-440X(02)00283-X).

103. Gohlke, H.; Kiel, C.; Case, D. A. Insights into Protein–Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras–Raf and Ras–RalGDS Complexes. *J. Mol. Biol.* 2003, *330* (4), 891–913. [https://doi.org/10.1016/S0022-2836\(03\)00610-7](https://doi.org/10.1016/S0022-2836(03)00610-7).
104. *FlareTM FEP* / *Cresset*. <https://cresset-group.com/software/flare-fep/> (accessed 2025-09-28).

3. Aim of the Work

The driving question behind this research was simple: is it possible to computationally predict how tightly a ligand binds to its target?

Thus, the central objective of this thesis was to investigate, optimize and develop new techniques to computationally anticipate the molecular features that govern the interaction between a ligand and its receptor, a fundamental determinant of drug efficacy. Since nearly 90% of drug candidates fail in clinical trials because they don't work well enough or have poor pharmacokinetics, there's a real need for computational methods that can predict binding strength quickly and accurately.

Computational techniques available nowadays either lack accuracy, require considerable computational power and time, or have limited spectrum of applicability. So, the current work inserts in this context, on one side with the task of developing a novel technique, called Thermal Titration Molecular Dynamics (TTMD), qualitatively-based, but fast and easy to settle, and on the other to explore the applicability spectrum of a state-of-the-art technique such as Free Energy Perturbation, with specific insight into membrane receptors.

3.1 Development of TTMD for Complex Stability Assessment

The first part of this thesis was dedicated to the development and optimization of Thermal Titration Molecular Dynamics (TTMD), a novel enhanced-sampling method designed to indirectly evaluate ligand binding affinity by assessing the stability of ligand-target complexes. The core idea behind TTMD is that ligands that maintain their native binding mode under progressive thermal stress are likely to exhibit stronger binding affinities. By incrementally increasing the simulation temperature, TTMD accelerates the destabilization of weaker complexes while allowing more stable ones to persist.

One of the strengths of TTMD is its simplicity of use. Unlike other enhanced-sampling methods that require the careful selection of collective variables, biasing schemes, or multiple user-defined parameters, TTMD is essentially parameter-free from the user's perspective. Once the system is prepared, the protocol can be run in a standardized fashion, making it accessible even to non-specialists and easily applicable across different projects.

In the present work, TTMD was developed and tested across a broad spectrum of molecular systems, extending beyond the classical case of protein-small molecule interactions. The method was evaluated on protein-protein complexes, protein-nucleic acid systems, and nucleic acid-small molecule complexes, demonstrating its versatility in handling diverse types of biological recognition events.

The method was deployed for two different tasks: Ligand ranking, where TTMD was used to order ligands by binding stability against the same target, and Pose validation, where the method was able to discriminate between native-like binding poses and non-native decoys produced by docking simulations.

Development and application of TTMD are highlighted in the [Scientific Publications](#) section.

3.2 Application of FEP to Membrane Protein Targets

The second part of the thesis addressed a more rigorous approach to affinity estimation, namely Free Energy Perturbation (FEP). Conducted in collaboration with Evotec Verona, this work focused on evaluating FEP approaches for the optimization of ligands that recognize membrane proteins.

Membrane proteins represent a substantial portion of the human proteome. These targets are central to many physiological processes, including signal transduction, molecular transport, and cellular communication¹⁻³. They are involved in many diseases, making them important targets for drug discovery. Among these, G Protein-Coupled Receptors (GPCRs) are particularly important, mediating diverse cellular responses to extracellular stimuli⁴⁻⁶. GPCRs alone are targeted by over 30% of approved drugs, highlighting their pharmacological significance^{7,8}.

Despite their importance, membrane proteins remain particularly challenging to study both experimentally and computationally due to their intrinsic conformational flexibility, heterogeneity, and the influence of the surrounding lipid bilayer^{9,10}. Unlike soluble globular proteins, which have abundant structural and thermodynamic data, computational drug design for membrane proteins is less developed, limiting applications in the major part of the drug discovery process stages^{3,10}.

At Evotec, the Flare FEP software¹¹ is currently employed to perform RBFE calculations. However, the software has not yet been validated or optimized for membrane protein systems. Indeed, the aim of this part of the work was to establish best practices for performing RBFE calculations in the context of membrane proteins. A benchmark study was carried out to determine optimal simulation parameters and identify the most reliable simulation protocols. These results, though not yet published, are reported in the [Unpublished Works](#) section of this thesis.

While RBFE methods are well-suited for the lead optimization stage, where congeneric ligand series are available for systematic comparison, they are not applicable to earlier stages of drug discovery, such as hit discovery or hit-to-lead, where chemically diverse ligands are considered. To extend the applicability of FEP methods to this earlier stage, some preliminary work was devoted to Absolute Binding Free Energy (ABFE) approaches.

In particular, the focus was placed on A3FE, an automated ABFE software developed in Julien Michel's laboratory at the University of Edinburgh. A3FE enables automatized ABFE calculations across chemically diverse ligand sets, thus providing a means to quantitatively evaluate binding affinities even when no congeneric series is available. This makes it a promising tool for identifying hit compounds.

At the start of this project, A3FE was not yet configured to handle membrane proteins within Evotec's discovery pipelines. To fill this gap, targeted training and software adaptation were undertaken. From February to April 2025, I joined Julien Michel's group at the University of Edinburgh, where the software was originally developed. This period was useful for learning in detail how A3FE operates, and for adapting its workflows to the specific requirements of membrane proteins.

Following this training, the next step was to integrate A3FE into the industrial research environment. Since one of the goals of this work was to establish reliable and transferable protocols that could be routinely applied within Evotec, I carried out an industrial internship at Verona's site between June and August 2025. There, my focus was on exploiting and validating A3FE within the Evotec high-performance computing (HPC) infrastructure, ensuring that the package could run efficiently, and deliver reproducible results in a production setting. Unfortunately, here my mission was affected by several issues, both with the software and the HPC cluster. Thus, there are no results from this work yet. Hopefully, the collaboration with Evotec and Julien Michel's lab will allow the full implementation and validation of the protocol in the near future, providing a solid foundation for quantitative binding affinity predictions for membrane protein targets in industrial drug discovery.

References

1. Latorraca, N. R.; Venkatakrishnan, A. J.; Dror, R. O. GPCR Dynamics: Structures in Motion. *Chem. Rev.* 2017, 117 (1), 139–155. <https://doi.org/10.1021/acs.chemrev.6b00177>.
2. Basith, S.; Cui, M.; Macalino, S. J. Y.; Park, J.; Clavio, N. A. B.; Kang, S.; Choi, S. Exploring G Protein-Coupled Receptors (GPCRs) Ligand Space via Cheminformatics Approaches: Impact on Rational Drug Design. *Front. Pharmacol.* 2018, 9. <https://doi.org/10.3389/fphar.2018.00128>.
3. Chan, H. C. S.; Li, Y.; Dahoun, T.; Vogel, H.; Yuan, S. New Binding Sites, New Opportunities for GPCR Drug Discovery. *Trends Biochem. Sci.* 2019, 44 (4), 312–330. <https://doi.org/10.1016/j.tibs.2018.11.011>.
4. Lee, Y.; Basith, S.; Choi, S. Recent Advances in Structure-Based Drug Design Targeting Class A G Protein-Coupled Receptors Utilizing Crystal Structures and Computational Simulations. *J. Med. Chem.* 2018, 61 (1), 1–46. <https://doi.org/10.1021/acs.jmedchem.6b01453>.
5. Yuan, X.; Raniolo, S.; Limongelli, V.; Xu, Y. The Molecular Mechanism Underlying Ligand Binding to the Membrane-Embedded Site of a G-Protein-Coupled Receptor. *J. Chem. Theory Comput.* 2018, 14 (5), 2761–2770. <https://doi.org/10.1021/acs.jctc.8b00046>.
6. Nakliang, P.; Lazim, R.; Chang, H.; Choi, S. Multiscale Molecular Modeling in G Protein-Coupled Receptor (GPCR)-Ligand Studies. *Biomolecules* 2020, 10 (4), 631. <https://doi.org/10.3390/biom10040631>.
7. Santos, R.; Ursu, O.; Gaulton, A.; Bento, A. P.; Donadi, R. S.; Bologa, C. G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T. I.; Overington, J. P. A Comprehensive Map of Molecular Drug Targets. *Nat. Rev. Drug Discov.* 2017, 16 (1), 19–34. <https://doi.org/10.1038/nrd.2016.230>.
8. Sloop, K. W.; Emmerson, P. J.; Statnick, M. A.; Willard, F. S. The Current State of GPCR-Based Drug Discovery to Treat Metabolic Disease. *Br. J. Pharmacol.* 2018, 175 (21), 4060–4071. <https://doi.org/10.1111/bph.14157>.
9. Amaro, R. E.; Mulholland, A. J. Multiscale Methods in Drug Design Bridge Chemical and Biological Complexity in the Search for Cures. *Nat. Rev. Chem.* 2018, 2 (4), 0148. <https://doi.org/10.1038/s41570-018-0148>.
10. Heifetz, A.; James, T.; Southey, M.; Morao, I.; Fedorov, D. G.; Bodkin, M. J.; Townsend-Nicholson, A. Analyzing GPCR-Ligand Interactions with the Fragment Molecular Orbital (FMO) Method. In *Quantum Mechanics in Drug Discovery*; Heifetz, A., Ed.; Springer US: New York, NY, 2020; pp 163–175. https://doi.org/10.1007/978-1-0716-0282-9_11.
11. FlareTM FEP | Cresset. <https://cresset-group.com/software/flare-fep/> (accessed 2025-09-28).

4. Scientific Publications

Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Qualitative Estimation of Protein–Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2022, 62 (22), 5715–5728.

<https://doi.org/10.1021/acs.jcim.2c00995>.

The study introduces thermal titration molecular dynamics (TTMD) as a computational method to assess protein-ligand binding stability. TTMD works by running molecular dynamics simulations at progressively increasing temperatures and evaluating stability through a scoring function based on protein–ligand interaction fingerprints. The method was tested on four pharmaceutically relevant proteins (CK1δ, CK2, PDK2, and SARS-CoV-2 main protease) with ligands of varying size, structure, and affinity. Across all cases, TTMD effectively discriminated between high- and low-affinity compounds, ranking ligands according to their binding strength. This demonstrates TTMD's potential as a screening tool for prioritizing ligands in drug discovery campaigns.

Menin, S.; Pavan, M.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement. *International Journal of Molecular Sciences* 2023, 24 (4), 3596.

<https://doi.org/10.3390/ijms24043596>.

This work reports the first application of thermal titration molecular dynamics (TTMD) as a refinement step for molecular docking. TTMD was used to distinguish the correct, native-like binding mode from decoy poses generated by docking. TTMD evaluated the stability of binding poses. Applied to four pharmaceutically relevant targets (CK1δ, CK2,

PDK2, and SARS-CoV-2 main protease), the method successfully retrieved the native binding mode, demonstrating TTMD's potential as a robust post-docking filter.

Costacurta, F.; Dodaro, A.; Bante, D.; Schöppe, H.; Sprenger, B.; Moghadasi, S. A.; Fleischmann, J.; Pavan, M.; Bassani, D.; Menin, S.; Rauch, S.; Krismer, L.; Sauerwein, A.; Heberle, A.; Rabensteiner, T.; Ho, J.; Harris, R. S.; Stefan, E.; Schneider, R.; Kaserer, T.; Moro, S.; Laer, D. von; Heilmann, E. A Comprehensive Study of SARS-CoV-2 Main Protease (M^{pro}) Inhibitor-Resistant Mutants Selected in a VSV-Based System. *bioRxiv* October 4, 2023, p 2023.09.22.558628.

<https://doi.org/10.1101/2023.09.22.558628>.

This study investigates resistance to nirmatrelvir, the first clinically approved SARS-CoV-2 main protease (M^{pro}) inhibitor, by selecting and characterizing M^{pro} mutations under drug pressure. Mutants were analyzed using cell-based, biochemical, and replicon assays, complemented with molecular modeling. Thermal titration molecular dynamics (TTMD) was employed to probe the effect of mutations on drug binding, specifically by comparing the dynamic stability of mutant M^{pro}-inhibitor complexes relative to the wild type. Through this, TTMD helped identify how resistance mutations alter binding stability, providing mechanistic insights into M^{pro} inhibitor resistance and guiding the design of next-generation antivirals.

Dodaro, A.; Pavan, M.; Menin, S.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Shedding Light on the Stability of RNA-Small Molecule Complexes. *Front. Mol. Biosci.* 2023, 10.

<https://doi.org/10.3389/fmolb.2023.1294543>.

This study reports the first application of thermal titration molecular dynamics (TTMD) to RNA-ligand complexes. Since RNA poses unique challenges for docking, such as high structural flexibility and complex interaction patterns, scoring often lacks reliability. Here, TTMD was used as a post-docking refinement tool to identify native binding modes among decoy poses. Applied to multiple pharmaceutically relevant RNA targets, TTMD successfully recognized the correct binding conformations, demonstrating its utility in RNA-targeted drug discovery.

Novello, G.; Dodaro, A.; Menin, S.; Cavastracci Strascia, C.; Sturlese, M.; Salmaso, V.; Moro, S. A Second Life for the Crystallographic Structure of Berenil-Dodecanucleotide Complex: A Computational Revisitation Thirty Years after Its Publication. *Med Chem Res* 2024.

<https://doi.org/10.1007/s00044-024-03299-9>.

This study revisits the classic DNA-ligand complex of Berenil bound to a dodecanucleotide, originally solved crystallographically by Professor Neidle and co-workers. Using modern computational methods, we re-examined the system's structural dynamics, with particular attention to the stabilizing role of water molecules. Thermal titration molecular dynamics (TTMD) was applied to probe the unbinding processes of Berenil in both the minor and major groove binding modes, allowing comparison of their

relative strength and specificity. Combined with SuMD binding trajectories and AquaMMapS hydration analysis, TTMD contributed to a comprehensive picture of how water molecules shape DNA-ligand recognition and stability, extending insights beyond the original structural study.

Dodaro, A.; Novello, G.; Menin, S.; Cavastracci Strascia, C.; Sturlese, M.; Salmaso, V.; Moro, S. Post-Docking Refinement of Peptide or Protein-RNA Complexes Using Thermal Titration Molecular Dynamics (TTMD): A Stability Insight. *J. Chem. Inf. Model.* 2025.

<https://doi.org/10.1021/acs.jcim.4c01393>.

This study extends the application of thermal titration molecular dynamics (TTMD) to RNA-peptide complexes, building on its previous successful use in protein-ligand and RNA-small-molecule systems. TTMD was employed as a post-docking refinement tool to overcome the scoring limitations of docking, enabling the identification of native binding poses among decoys. Applied to pharmaceutically relevant RNA-peptide targets, TTMD proved effective in refining docking results and demonstrated its broader applicability for studying complex RNA-protein interactions relevant to drug discovery.

Pavan, M.; Menin, S.; Dodaro, A.; Novello, G.; Cavastracci Strascia, C.; Sturlese, M.; Salmaso, V.; Moro, S. Thermal Titration Molecular Dynamics: The Revenge of the Fragments. *J. Chem. Inf. Model.* 2025.

<https://doi.org/10.1021/acs.jcim.4c01681>.

This study introduces a TTMD protocol tailored for fragment-based drug discovery, addressing the challenges of predicting binding modes for very small, low-complexity ligands. By applying TTMD as a post-docking refinement method across 12 fragment-protein test cases on eight pharmaceutically relevant targets, the approach was able to discriminate native binding poses from decoys. TTMD successfully retrieved the correct binding mode in over 80% of cases, clearly outperforming docking alone. These results demonstrate TTMD's effectiveness as a tool to improve fragment screening and guide fragment-to-lead optimization.



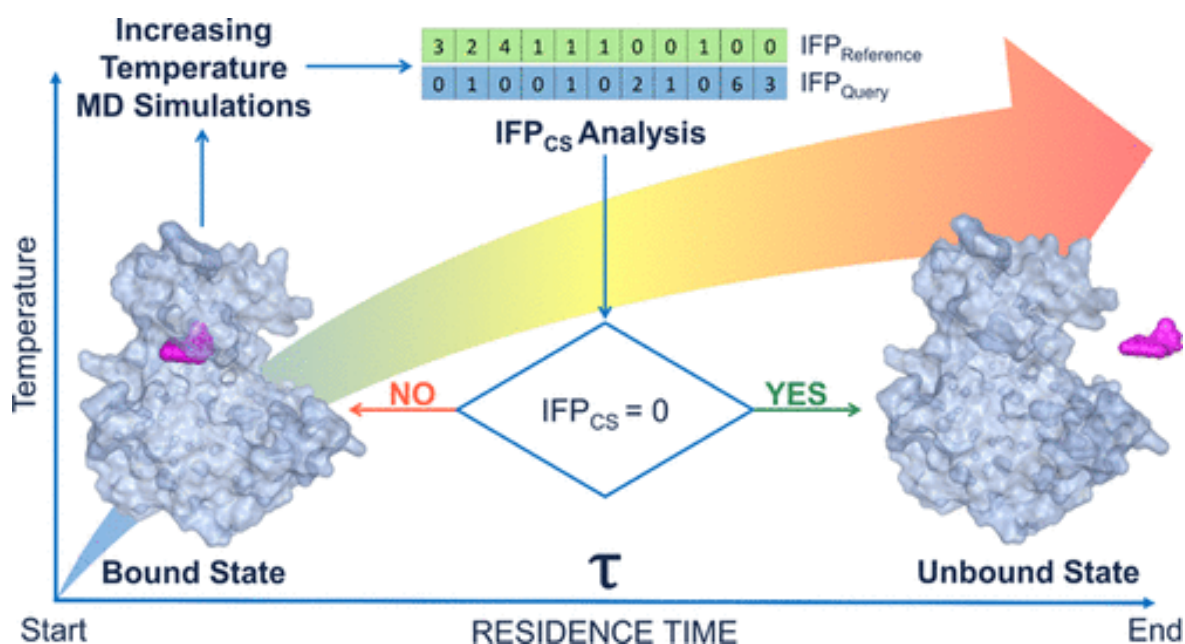
ACS Publications
Most Trusted. Most Cited. Most Read.

www.acs.org

Qualitative Estimation of Protein–Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations

Matteo Pavan, Silvia Menin, Davide Bassani, Mattia Sturlese, and Stefano Moro

Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Qualitative Estimation of Protein–Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2022, 62 (22), 5715–5728. <https://doi.org/10.1021/acs.jcim.2c00995>.



Abstract

The prediction of ligand efficacy has long been linked to thermodynamic properties such as the equilibrium dissociation constant, which considers both the association and the dissociation rates of a defined protein–ligand complex. In the last 15 years, there has been a paradigm shift, with an increased interest in the determination of kinetic properties such as the drug–target residence time since they better correlate with ligand efficacy compared to other parameters. In this article, we present thermal titration molecular dynamics (TTMD), an alternative computational method that combines a series of molecular dynamics simulations performed at progressively increasing temperatures with a scoring function based on protein–ligand interaction fingerprints for the qualitative estimation of protein–ligand-binding stability. The protocol has been applied to four different pharmaceutically relevant test cases, including protein kinase CK1δ, protein kinase CK2, pyruvate dehydrogenase kinase 2, and SARS-CoV-2 main protease, on a variety of ligands with different sizes, structures, and experimentally determined affinity values. In all four cases, TTMD was successfully able to distinguish between high-affinity compounds (low nanomolar range) and low-affinity ones (micromolar), proving to be a useful screening tool for the prioritization of compounds in a drug discovery campaign.

Introduction

At the beginning of the 20th century, Ehrlich's famous quote "Corpora non agunt nisi fixata" marked a pivotal moment in the history of modern pharmacology, molecular medicine, and drug development¹. His statement, combined with independent observations about "receptive substances" by Langley², defined the birth of the receptor theory of drug action, which postulates that a drug can only work as long as it is

bound to its target receptor^{3,4}.

Although the basic ideas of this cornerstone principle were formulated more than 100 years ago⁵, it was only in the 1970s that molecular receptors could be successfully isolated and purified^{6,7}. This allowed for the development of different biochemical and cellular assays for the direct determination of the extent to which a drug is bound to its receptor under thermodynamic equilibrium conditions, that is, the binding affinity^{8,9}. Traditionally, this parameter is quantified either through the equilibrium dissociation constant (K_d) or through other proxy metrics such as the drug concentration responsible for the half-maximal inhibition/effect (IC_{50}/EC_{50}) and the inhibition constant (K_i)¹⁰.

Although, in principle, these measurements are all adequate predictors for in vivo efficacy, that is, the capability of the drug to induce the desired response, they are all related to in vitro assays portrayed under closed system conditions¹¹. Since in an open in vivo system, the drug concentration is not a fixed variable and indeed varies over time because of various physiological processes, several authors thus suggested that the observables related to drug–receptor binding kinetics, such as the association (k_{on}) and dissociation (k_{off}) constants, could be better descriptors for drug efficacy^{12–15}. Accordingly, while the binding affinity only depends on the free energy difference between the bound and unbound states, which can be directly correlated to K_d , association and dissociation rates depend on the energy barriers that separate those states¹⁰.

Thermodynamics and kinetics of bindings are interlinked by the equation $K_d = k_{off}/k_{on}$ ¹⁰. While, in theory, both kinetic constants should equally contribute to the determination of K_d , physicochemical and pharmacological limitations on the k_{on} value¹⁶ render the in vivo duration of a receptor–ligand complex almost entirely dependent on the k_{off} value¹¹. Based on this observation, Copeland et al. first suggested that the key determinant of in vivo pharmacological activity and duration is not the binding affinity but, instead, the lifetime of the receptor–ligand complex, defined as the residence time¹². Furthermore, Copeland et al. proposed a mathematical formulation for the quantification of this parameter, defining it as the reciprocal of the k_{off} ($\tau = 1/k_{off}$)¹².

From an experimental perspective, a plethora of methods for the determination of binding kinetics are available^{17–20}. Each of them relies on monitoring the time-dependent evolution of a signal in response to the binding event²¹. The first strategy revolves around the radio-²² and spectroscopic^{20–23} labeling of ligands and includes techniques such as fluorescent resonance energy transfer²⁴ and bioluminescence resonance energy transfer²⁵. An alternative approach revolves around the exploitation of label-free approaches such as surface plasmon resonance^{26,27}, nuclear magnetic resonance²⁸, surface acoustic wave methods²⁹, and various declinations of isothermal titration calorimetry^{30,31}. Finally, another possible method is based on following enzymatic reactions, usually through the monitoring of spectroscopic parameters³².

Alongside the aforementioned experimental protocols, various computational approaches exist that can flank and expand on the information that they provide by showcasing mechanistic information about the underlying process at an atomic level of detail^{21,33,34}. Particularly, molecular dynamics (MD) simulations have been exploited to estimate thermodynamic properties such as the binding affinity for protein–ligand complexes, and due to the growing interest in the study of kinetics for drug discovery, they have recently been applied to the estimation of kinetic properties as well³⁵. Although it would theoretically be possible to exploit unbiased MD simulations for the determination of kinetics' observables, biologically relevant events such as drug–target unbinding occur at much longer timescales than those of typical MD simulations, heavily restricting their limitations in terms of computational resources' availability³⁶ and neglecting any real-world application of the technique³⁷. For this reason, several different methods have been developed throughout the years that implement smart sampling strategies to reduce the required computational effort, such as various instances of metadynamics^{38–41}, which are based on repeatedly “filling” the potential energy of the system by a sum of Gaussians centered along the trajectory, followed by an appropriately chosen ensemble of collective variables (CVs)⁴², scaled MD, which relies on smoothing the potential energy surface by applying an appropriate scaling factor^{43–45}, and τ -random acceleration MD, in which a small randomly oriented force vector is applied to the ligand^{46–48}.

In the present study, we present the first application of thermal titration MD (TTMD), an alternative MD-based approach for the qualitative estimation of protein–ligand complex stability. The method relies on evaluating the conservation of the native binding mode for a ligand of interest throughout a series of

MD trajectories performed at progressively increasing temperature values. For validation purposes, the protocol has been applied to four different biomolecular targets of pharmaceutical interest: casein kinase 1 δ (CK1 δ), casein kinase 2 (CK2), pyruvate dehydrogenase kinase 2 (PDK2), and SARS-CoV-2 main protease (M_{pro}).

Materials And Methods

Hardware Overview

Each general molecular modeling operation, such as the preparation of protein–ligand complex structures, the setup for MD simulations, and trajectory analyses, was conducted on a 20 CPU Linux workstation equipped with an Intel Core i9-9820X 3.3 GHz processor. All MD simulations were carried out on a GPU cluster composed of 20 NVIDIA drivers ranging from GTX980 to RTX2080Ti.

Structure Preparation

The three-dimensional coordinates of the protein–ligand complexes used in this study were retrieved from the Protein Data Bank (PDB)⁴⁹ and processed before MD simulations through several tools provided by the Molecular Operating Environment (MOE) 2019.01 suite⁵⁰. Four different macromolecular targets were considered in this work: CK1 δ , CK2, PDK2, and SARS-CoV-2 main protease (M_{pro}). For each macromolecular target, the considered structures are reported in Table 1.

Table 1. List of the Protein-Ligand Complex Structures Used in This Work^a

CK1 δ	3UZP ⁵¹	4TN6 ⁵²	5IH5 ⁵³	5IH6 ⁵³	5MQV ⁵⁴
CK2	2ZJW ⁵⁵	3H30 ⁵⁶	3PE1 ⁵⁷	3PE2 ⁵⁷	6HOU ⁵⁸
PDK2	4MP2 ⁵⁹	4V25 ⁶⁰	5JZ1 ⁶¹	5M4M ⁶²	7EA0 ⁶³
M _{pro}	6M2N ⁶⁴	7LTJ ⁶⁵	7M8P ⁶⁶	7N91 ⁶⁶	7N44 ⁶⁷

^a Complexes are grouped by macromolecule targets.

Each protein–ligand system was simulated in the monomeric form, except for SARS-CoV-2 M_{pro}, which was simulated in the dimeric form by applying a symmetric crystallographic transformation to each asymmetric unit. First, all structures were pre-processed using the “structure preparation” tool, assigning alternates to the highest occupancy conformation, rebuilding missing loops through homology modeling, and correcting inconsistencies between the primary sequence and the tertiary structure. Second, the “Protonate3D” tool was exploited to add missing hydrogens to the system and to determine the most probable protonation state of titratable residues at pH = 7.4. Finally, every non-protein and non-ligand atom of the system was removed before saving the structure for further calculations, except for water molecules within 4.5 Å of the ligand that were not removed and were indeed considered in the simulations. Concerning the protonation state of the ligand, the most abundant protomer at pH 7.4 according to the “protomers” tool was considered in the calculations, besides CK2 complex 2ZJW, where two different protonation states were considered. Particularly, in the case of 2ZJW, the predominant form at pH 7.4 should be the neutral, noncharged one. However, in the context of the binding pocket, the interaction network of the hydroxyl in position 3 (the one facing Lys68 and the conserved water molecule W1) suggests the prevalence of a monocharged, ionized form. Since experimental data published in the literature does not clarify the correct protonation state for ellagic acid in the context of CK2 recognition^{55,68,69}, we opted to consider both hypotheses equally relevant (50/50).

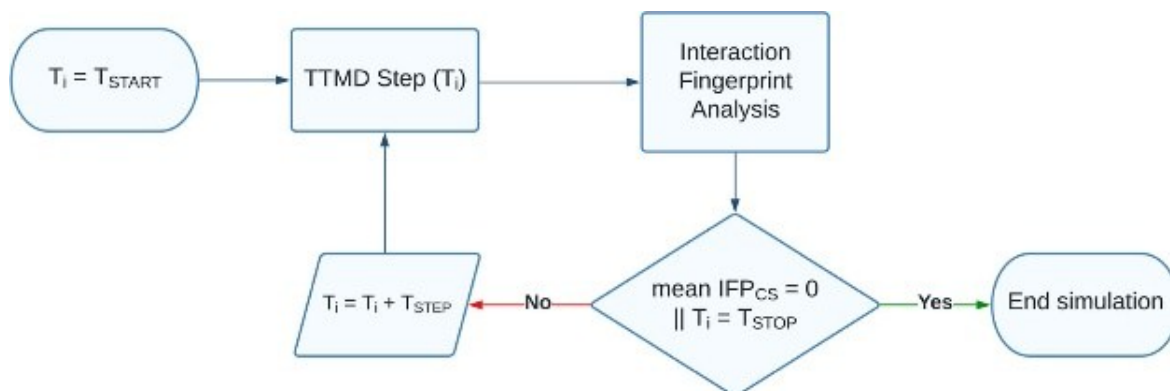


Figure 1. Computational workflow for a TTMD simulation.

System Setup for MD Simulations and Equilibration Protocol

Each protein–ligand complex is prepared as described before and further processed through various tools from Visual Molecular Dynamics (VMD) 1.9.2⁷⁰ and the Ambertools14⁷¹ suite. Protein atoms were parametrized through the ff14SB⁷² force field, while the general Amber force field⁷³ was utilized to parametrize the ligands. Partial charges were attributed to the ligand through the AM1-BCC method⁷⁴. Each investigated system was solvated in a cubic box with a padding of 15 Å, utilizing the TIP3P⁷⁵ model for water molecules. A proper number of sodium and chloride ions were added to neutralize the system and reach a salt concentration of 0.154 M. Before undergoing MD simulations, each system was energy minimized for a total of 500 steps with the conjugate-gradient algorithm to remove clashes and bad contacts.

Afterward, each minimized system was subjected to a two-step equilibration protocol. During the first stage, a 0.1 ns simulation in the canonical ensemble (NVT) was performed, with harmonic positional restraints (5 kcal mol^{−1} Å^{−2} force constant) applied on both protein and ligand atoms. The second stage, instead, consisted of a 0.5 ns simulation carried out in the isothermal–isobaric ensemble (NPT), applying the same restraints only to the ligand position and the protein backbone.

Each MD simulation presented in this work, both in the equilibration and the production stage, was performed using an integration timestep of 2 fs, keeping the temperature at a constant value of 310 K through a Langevin thermostat⁷⁶, constraining the length of bonds involving hydrogen bonds through the M-SHAKE algorithm⁷⁷, exploiting the particle-mesh Ewald⁷⁸ method to compute electrostatic interactions using cubic spline interpolation and a 1 Å grid spacing and setting a 9.0 Å cutoff for the calculation of Lennard-Jones interactions. Simulations in the NPT ensemble were carried out keeping the pressure at a constant 1 atm value by making use of a Monte Carlo barostat⁷⁹.

All MD simulations were run through the ACEMD 3⁸⁰ engine, which is based upon the open-source library for molecular simulations OpenMM 7⁸¹.

TTMD Simulations

TTMD is an alternative enhanced sampling MD approach for the qualitative estimation of protein–ligand complex stability. The method relies on evaluating the conservation of the native binding mode for a ligand of interest throughout a series of MD trajectories performed at progressively increasing temperature values. The protocol described herein is implemented as a Python 3.10 code, which relies on the NumPy, MDAnalysis^{82,83}, Open Drug Discovery Toolkit⁸⁴, and Scikit-learn libraries. The workflow for a TTMD simulation is reported in Figure 1 and detailed hereafter.

In detail, the task is accomplished through a series of short, unbiased MD simulations performed at different, progressively increasing temperatures in the NVT ensemble with the ACEMD3 engine. For each TTMD run, the duration of each simulation window (defined as “TTMD-step”) is fixed and user-defined (10 ns, in this case). The starting and final temperature values as well as the temperature increase between each “TTMD-step” are also defined by the user based on prior knowledge of the target, particularly regarding the conservation of the protein fold at higher simulation temperatures (which, in the context of

this article, is carried out by monitoring the protein backbone RMSD throughout the simulation). In this work, the starting temperature was set to 300 K, the ending temperature was set to 450 K, while the temperature increase between each “TTMD-step” was set to 10 K.

The progress of the simulation is monitored through a scoring function based on protein–ligand interaction fingerprints. The scoring function, defined as IFP_{CS} and originally described in previous scientific work from our laboratory⁸⁵, exploits the Open Drug Discovery Toolkit Python library to calculate protein–ligand interaction fingerprints for each frame of the TTMD trajectory and compares them through the cosine similarity metric as implemented in the Scikit-learn Python module to a reference fingerprint based on the last trajectory frame extracted from the second and last equilibration stage. Specifically, each protein–ligand interaction fingerprint is an integer vector composed of $r \times 8$ elements, where r is the number of protein residues. Each protein residue is encoded into eight bits of information, one for each type of intermolecular interaction considered (hydrophobic contacts, aromatic face to face, aromatic edge to face, hydrogen bonds with the protein acting as a donor, hydrogen bonds with the protein acting as an acceptor, salt bridge with the protein acting as the positively charged member, salt bridge with the protein acting as the positively negative member, and an ionic bond with a metal ion, respectively). The mathematic formulation of the IFP_{CS} scoring function is reported in Equation 1.

Equation 1. Mathematic formulation of the IFP_{CS} scoring function.

$$IFPCS = \frac{A \cdot B}{\|A\| \|B\|} \times -1$$

The IFP_{CS} value ranges from -1 , indicating a total superposition between the reference and the query fingerprint, to zero, which indicates that every interaction determinant of the reference fingerprint is lost in the query.

At the end of each “TTMD-step”, the average IFP_{CS} score for the step is calculated: if the value is null, indicating that for the whole duration of the step the original binding mode was not sampled, the TTMD trajectory is terminated, while if the value is not null, the simulation proceeds to the next “TTMD-step”.

TTMD Trajectory Analyses

Each TTMD trajectory is analyzed by making use of an in-house Python 3.10 script. The root-mean-square deviation (RMSD) of atomic coordinates for both the ligand and the protein backbone is calculated for each frame through the MDAnalysis package. The per-residue decomposition of the protein–ligand interaction energy is computed for each frame by exploiting the NAMD Energy plugin (version 1.4)⁸⁶ for VMD. Three different plots are then generated, making use of the Matplotlib and Seaborn Python packages. The first plot (“titration profile”) reports the average IFP_{CS} value for each TTMD step as a function of the step temperature. A straight line joining the start and final states of the simulation is also drawn in the graph, and its slope is reported in the legend and stored for further analysis. The second graph illustrates the time-dependent per-residue decomposition of the interaction energy, with the 25 most contacted residues alongside the TTMD trajectory being considered. The third and final plots report the time-dependent evolution of the ligand and protein backbone RMSD and the IFP_{CS} value.

MS Coefficient Determination

For each TTMD simulation, a proxy value for the protein–ligand complex stability based on the conservation of the binding mode throughout the trajectory is calculated as reported in Equation 2.

Equation 2. Mathematic formulation of the MS score.

$$MS = \frac{meanIFP_{CS}^{T_{end}} - (-1)}{T_{end} - T_{start}}$$

The MS coefficient is the slope of the straight line that interpolates the first and last points of the “titration profile” plot described in the previous paragraph. In Equation 2, $meanIFP_{CS}^{T_{end}}$ is the average IFP_{CS} value for the last temperature explored in the TTMD trajectory, -1 is the IFP_{CS} value for the initial state of the simulation, and T_{end} and T_{start} are the final and starting temperatures of the simulation. Values are positive and can vary between 0 (indicative of a strong binding) and 1 (related to a weak binding).

For each ligand, five independent TTMD simulations are performed, and the average MS coefficient is then calculated based on three of them, discarding the highest and the lowest value.

Results

To test and validate the applicability of the TTMD protocol, we performed four different case studies on four different pharmaceutically relevant targets of interest for our laboratory, specifically CK1 δ , CK2, PDK2, and SARS-CoV-2 main protease (M^{Pro}). For each protein target, five different protein–ligand complexes were chosen based on the availability of binding affinity data. A list of all protein–ligand complexes used in the present work can be found in Table 1 at the beginning of the Materials and Methods section, while detailed information about the ligands utilized in each test case can be found in Tables S1–S4 (Supporting Information). For each protein–ligand complex investigated in the article, five independent TTMD simulations were carried out. The results for each test case are reported hereafter in separate paragraphs and discussed aggregately in the Discussion section of the article. For each target, the conservation of the protein fold throughout the simulation is carried out by monitoring the time-dependent evolution of the protein backbone RMSD, as reported in the detailed analysis for each representative replicate illustrated in Figures S1–S21 (Supporting Information).

Protein Kinase CK1 δ

Protein kinase CK1 δ is a serine–threonine kinase that belongs to the family of CK1 kinases (casein kinase 1)⁸⁷. Due to its pleiotropic nature (about 140 substrates have been reported so far), this kinase is involved in the regulation of several different cellular pathways^{87,88}. Particularly relevant from a medicinal chemistry perspective is its involvement in several neurodegenerative diseases such as Alzheimer’s disease, Parkinson’s disease, and amyotrophic lateral sclerosis by phosphorylating protein targets such as the tau protein, α -synuclein, and TDP-43 (transactivate response DNA-binding protein 43)⁸⁹. 34 crystal structures of CK1 δ , among which several protein–ligand complexes can be found, are deposited in the PDB, with the affinity of co-crystallized inhibitors ranging over 3 orders of magnitude, making it a suitable target for the application of our computational protocol. The results of TTMD simulations performed on CK1 δ crystal complexes are summarized in Table 2 and Figure 2, while a detailed analysis of a representative trajectory for each protein–ligand complex (the one highlighted in green in Table 2) is reported in Figures S1–S5 in the Supporting Information.

Table 2. Results for the TTMD Simulations Performed on the Five Investigated CK1 δ Complexes^a

PDB ID	LIG ID	IC ₅₀ (nM)	MSMD1	MSMD2	MSMD3	MSMD4	MSMD5	MS _{average}
3UZP	OCK	13	0.00203	0.00333	0.00228	0.00212	0.00275	0.0024
4TN6	PFO	3.9	0.00300	0.00345	0.00307	0.00214	0.00309	0.0031
5IH5	AUE	500	0.00539	0.00446	0.00357	0.00909	0.00347	0.0045
5IH6	AUG	2500	0.00380	0.00769	0.00420	0.00557	0.00594	0.0052
5MQV	D5Q	9	0.00283	0.00295	0.00317	0.00256	0.00340	0.0030

^a For each protein–ligand complex, the PDB accession code, the ligand three-letter code, the experimentally determined affinity value, the MS coefficient for each simulation, and the average MS coefficient are reported. In each row, the lowest MS value is highlighted in red, while the highest value is highlighted in blue: both values were discarded for the calculation of the average MS coefficient reported in the last column. The most representative replicate, the one with the nearest MS coefficient to the average MS, is highlighted in green.

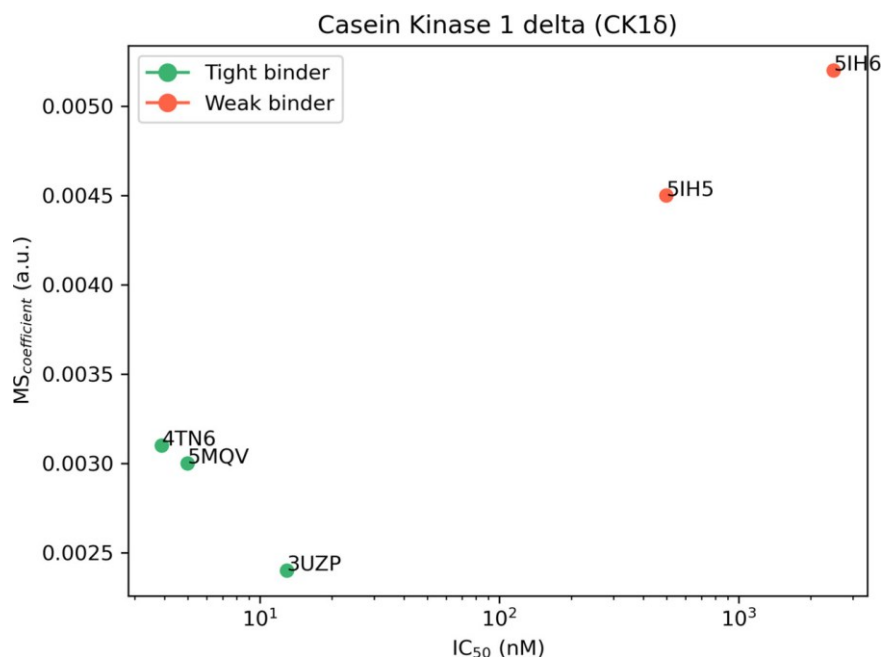


Figure 2. Aggregate results of the TTMD simulations performed on the five investigated CK1 δ complexes. On the horizontal axis, the experimentally determined affinity value (expressed as IC₅₀) is reported, while on the vertical axis, the average MS coefficient is indicated. Each dot is color-coded as green or red and classified as a tight or weak binder based on the MS cutoff value of 0.004.

As can be deduced by the analysis of the data extracted from the various TTMD simulations, the ligands respond differently to the protocol based on the experimental affinity value. As can be noticed in Figure 5, complexes 5IH5 and 5IH6, which are characterized by the lowest affinity values (500 and 2500 μ M, respectively), are the ones with the highest MS coefficient value, indicating a loss of the native binding mode throughout the simulations. On the contrary, ligands with a good experimental affinity toward the target (in the low nanomolar range) are associated with good conservation of the native binding mode, as highlighted by the lower MS coefficient value. Based on this observation, a cutoff MS value of 0.004, able to distinguish between the tight and weak binders for CK1 δ , can be determined. The detailed trajectory analyses provided in Figures S1–S5 illustrate how the loss of the native binding mode is primarily driven by the loss of crucial hydrogen bond interactions with the hinge region, particularly with Leu85 and Glu83. This evidence is in agreement with previously published work from our laboratory, which indicates how using an appropriate pharmacophore filter that takes into account the crucial hydrogen bond with the backbone of Leu85 leads to good results in virtual screening^{85,90,91}. For visual reference, a comparison between the representative replicates for the 3UZP and 5IH6 is reported in Video S1.

Protein Kinase CK2

Protein kinase CK2 is a serine–threonine kinase and represents one of the first identified protein kinases⁹². Similar to CK1 δ , CK2 can phosphorylate a plethora of different substrates and is therefore involved in the regulation of several biological pathways⁹³. The variety of biologically relevant scenarios in which CK2 is involved makes it a hot target from a pharmaceutical perspective, being related to several types of cancer, different neurodegenerative diseases (similarly to CK1 δ), and also viral infections⁹⁴. As of today, 214 crystal structures of CK2, among which several protein–ligand complexes can be found, are deposited in the PDB, with the affinity of co-crystallized inhibitors ranging over 5 orders of magnitude, making it also a suitable target for the application of our computational protocol. The results of TTMD simulations carried out on CK2 crystal complexes are summarized in Table 3 and Figure 3, while a detailed analysis for a representative trajectory for each protein–ligand complex (the one highlighted in green in Table 3) is reported in Figures S6–S11 in the Supporting Information. As explained in the Materials and Methods section, two different ligand protonation states are considered for the 2ZJW complex: the neutral one and the negatively charged one. Although they are reported separately in both Table 3 and Figures S6 and S7,

they are considered as a single entity in Figure 3, where the average MS value between the two different protonation states is reported.

Table 3. Results for the TTMD Simulations Performed on the Five Investigated CK2 Complexes^a

PDB ID	LIG ID	K _d (nM)	MS _{MD1}	MS _{MD2}	MS _{MD3}	MS _{MD4}	MS _{MD5}	MS _{average}
2ZJW (0)	REF	40	0.00318	0.00282	0.00269	0.00391	0.00205	0.0029
2ZJW (-)	REF	40	0.00192	0.00189	0.00227	0.00205	0.00133	0.0020
3H30	RFZ	13000	0.00504	0.00484	0.00395	0.00450	0.00453	0.0046
3PE1	3NG	1.5	0.00178	0.00231	0.00198	0.00219	0.00373	0.0022
3PE2	E1B	2.3	0.00179	0.00179	0.00171	0.00169	0.00170	0.0017
6HOU	V55	53400	0.0035	0.00667	0.00769	0.00364	0.00667	0.0057

^a For each protein–ligand complex, the PDB accession code, the ligand three-letter code, the experimentally determined affinity value, the MS coefficient for each simulation, and the average MS coefficient are reported. In each row, the lowest MS value is highlighted in red, while the highest value is highlighted in blue: both values were discarded for the calculation of the average MS coefficient reported in the last column. The most representative replicate, the one with the nearest MS coefficient to the average MS, is highlighted in green. For the complex 2ZJW, two different protonation states were independently considered in the simulations and are reported separately.

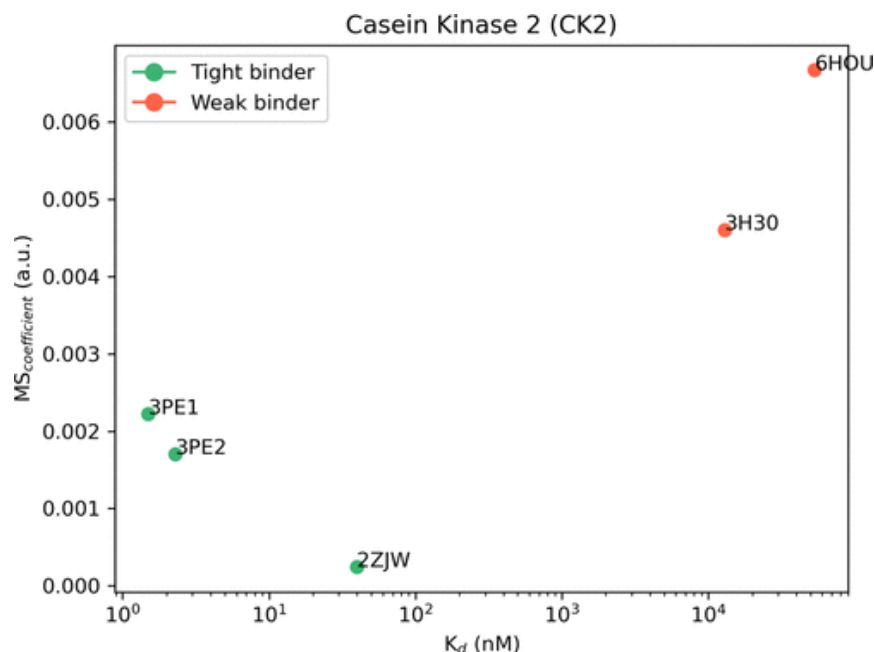


Figure 3. Aggregate results of the TTMD simulations performed on the five investigated CK2 complexes. On the horizontal axis, the experimentally determined affinity value (expressed as K_d) is reported, while on the vertical axis, the average MS coefficient is indicated. Each dot is color-coded as green or red and classified as a tight or weak binder based on the MS cutoff value of 0.004. For the complex 2ZJW, two different ligand protonation states were considered, but only one aggregate result (the average of the two states) is reported in the plot.

As in the case of CK1 δ , the investigated ligands show a different behavior during the TTMD simulations based on their experimental affinity value. The complexes characterized by a lower protein–ligand binding affinity (13 μ M for 3H30 and 53.4 μ M for 6HOU) are also the ones characterized by the highest MS coefficient value (0.0046 and 0.0057, respectively). On the contrary, as observed for CK1 δ , ligands with a binding affinity in the low nanomolar range are characterized by a conservation of the native binding mode throughout the simulation, resulting in a lower MS coefficient (below 0.003). Once again, an empirical threshold MS value of 0.004 can be extracted from this test set and utilized to distinguish between tight and weak binders. As can be noticed by the evolution of the interaction pattern between ligands and the binding pocket throughout the trajectories (Figures S6–S11), the ligands with the most stable binding mode are the ones that tightly interact with Lys68 and Ile174: this is particularly noticeable in the case of ellagic acid (2ZJW), for which two different protonation states have been simulated. In the neutral form, the ellagic acid-binding mode is less stable throughout the trajectory, while in the monocharged form, the ellagic acid-binding mode is very stable even at high simulation temperatures due to a favorable interaction with Lys68. A comparison between the representative replicates for complexes 3PE2 and 6HOU is shown in Video S2.

Pyruvate Dehydrogenase Kinase 2

PDK2 is a pivotal enzyme in cellular energy metabolism that has previously been implicated in cancer⁹⁵. PDK2 is a member of the GHKL ATPase/kinase superfamily and exerts its activity by phosphorylating and regulating the pyruvate dehydrogenase complex, which is a central control point in cellular energy metabolism since it links glycolysis with the tricarboxylic acid cycle^{96,97}. Due to its involvement in the regulation of the energetic metabolism of cells, it is a drug target both from a metabolic and an antitumoral perspective. At the present moment, 33 crystal structures of PDK2, among which several protein–ligand complexes can be found, are deposited in the PDB, with the affinity of co-crystallized inhibitors ranging over 6 orders of magnitude, making it a suitable target for the application of our computational protocol. The results of TTMD simulations performed on PDK2 crystal complexes are summarized in Table 4 and Figure 4, while a detailed analysis for a representative trajectory for each protein–ligand complex (the one highlighted in green in Table 4) is reported in Figures S12–S16 in the Supporting Information.

Table 4. Results for the TTMD Simulations Performed on the Five Investigated PDK2 Complexes^a

PDB ID	LIG ID	K _d (nM)	MS _{MD1}	MS _{MD2}	MS _{MD3}	MS _{MD4}	MS _{MD5}	MS _{average}
4MP2	PV1	3570	0.00603	0.01429	0.025	0.01111	0.00579	0.0105
4V25	SZ6	150	0.00216	0.00247	0.00194	0.00221	0.00332	0.0023
5J71	P35	110	0.00296	0.00296	0.00269	0.00243	0.00331	0.0029
5M4M	7FW	1	0.0028	0.0025	0.00286	0.00261	0.00318	0.0027
7EA0	W6P	958000	0.00347	0.00769	0.00667	0.00769	0.00665	0.0070

^a For each protein–ligand complex, the PDB accession code, the ligand three-letter code, the experimentally determined affinity value, the MS coefficient for each simulation, and the average MS coefficient are reported. In each row, the lowest MS value is highlighted in red, while the highest value is highlighted in blue: both values were discarded for the calculation of the average MS coefficient reported in the last column. The most representative replicate, the one with the nearest MS coefficient to the average MS, is highlighted in green.

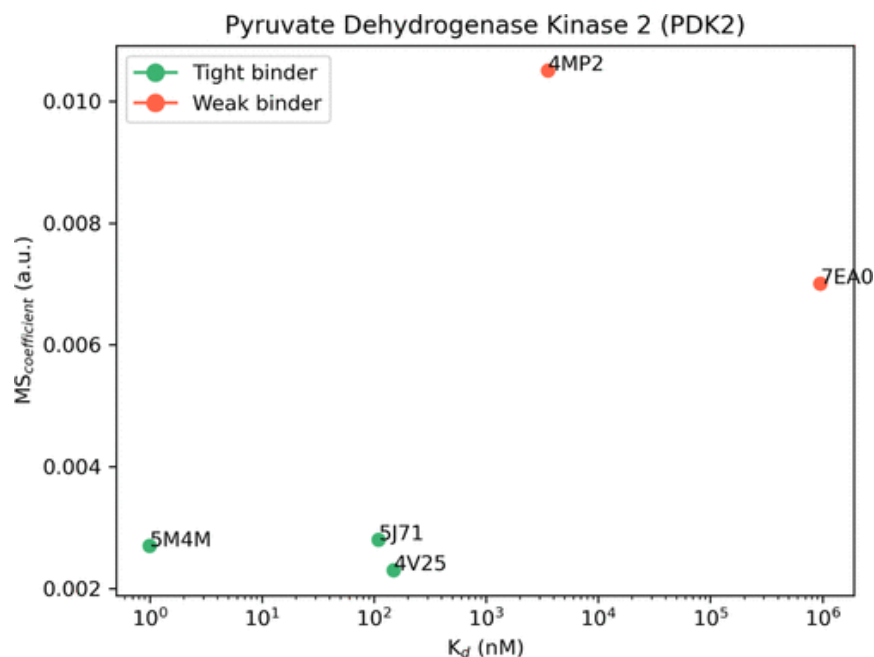


Figure 4. Aggregate results of the TTMD simulations performed on the five investigated PDK2 complexes. On the horizontal axis, the experimentally determined affinity value (expressed as K_d) is reported, while on the vertical axis, the average MS coefficient is indicated. Each dot is color-coded as green or red and classified as a tight or weak binder based on the MS cutoff value of 0.004.

The analysis of results for the TTMD simulations performed on PDK2 protein–ligand complexes matches the ones already shown for CK1 δ and CK2. Indeed, the ligands with the lowest binding affinity (3.57 μ M for 4MP2 and 958 μ M for 7EA0) are the ones with the highest MS coefficient (0.0105 and 0.007, respectively), while on the contrary, ligands characterized by a good binding affinity are also the ones characterized by the lowest MS coefficient (below 0.003). The same threshold value used for previous cases ($MS < 0.004$) can also be utilized in this case to distinguish between the weak and tight binders. Looking at the evolution of the interaction pattern of various ligands throughout the simulations, it can be noticed that tight binders are characterized by persistent attractive interactions with Asp290 and Thr354. These residues are buried within the binding pocket, which contributes to the persistence of their interaction with the ligand compared to other more solvent-exposed residues such as Asn255, Arg258, and Glu262 which seem to be less relevant in retaining the ligand within the binding site. Noticeably, in the case of complex 4MP2, a repulsive interaction with Asp290 is present at the beginning of the simulation, and this could be a possible explanation for the low persistence of the native binding mode. Moreover, as can be noticed in Figures S14 and S15, in the case of complexes 5J71 and 5M4M, the fraction of the ligand which interacts with Asp290 and Thr354 barely moves from the starting position, fully retaining this interaction for the whole duration of the simulation, while partially losing the interactions with other more exposed residues, which increases the ligand's RMSD despite most of the binding determinants being conserved. A comparison between the representative replicate for complexes 4V25 and 4MP2 is illustrated in Video S3.

SARS-CoV-2 Main Protease (M^{pro})

SARS-CoV-2 is a betacoronavirus responsible for the COVID-19 pandemic which, to date, has caused the death of more than 6.5 million people around the world^{98,99}. A pivotal enzyme in the viruses' replication cycle is represented by their main protease (M^{pro}), a cysteine peptidase that is involved in the proteolytic cleavage of the pp1a/pp1ab polyproteins into several mature nonstructural proteins^{100,101}. Due to its crucial role in the ability of the virus to replicate itself, the main protease is a validated antiviral target¹⁰² and, as such, has become the focus of several different drug discovery campaigns^{103–105}, leading to 613 experimentally solved structures deposited on the PDB, a marketed drug (Paxlovid, therapeutic association of nirmatrelvir and ritonavir)^{106,107}, and several inhibitors, with affinity values ranging from low nanomolar

to micromolar and beyond. Its pharmaceutical relevance and the abundance of structural data make the SARS-CoV-2 main protease a good target for the validation of the TTMD protocol. The results of TTMD simulations performed on SARS-CoV-2 M^{pro} crystal complexes are summarized in Table 5 and Figure 5, while a detailed analysis for a representative trajectory for each protein–ligand complex (the one highlighted in green in Table 5) is reported in Figures S17–S21 in the Supporting Information.

Table 5. Results for the TTMD Simulations Performed on the Five Investigated SARS-CoV-2 M^{pro} Complexes^a

PDB ID	LIG ID	IC ₅₀ (nM)	MSMD1	MSMD2	MSMD3	MSMD4	MSMD5	MS _{average}
6M2N	3WL	940	0.00714	0.00714	0.00557	0.00714	0.00909	0.0071
7LTJ	YD1	4200	0.00909	0.00642	0.00595	0.00667	0.00664	0.0066
7M8P	YSJ	20	0.00401	0.00336	0.00379	0.00317	0.00254	0.0034
7M91	YU4	25	0.00316	0.00437	0.00315	0.00313	0.00383	0.0034
7N44	06I	42	0.00325	0.00598	0.00398	0.00384	0.00383	0.0039

^a For each protein–ligand complex, the PDB accession code, the ligand three-letter code, the experimentally determined affinity value, the MS coefficient for each simulation, and the average MS coefficient are reported. In each row, the lowest MS value is highlighted in red, while the highest value is highlighted in blue: both values were discarded for the calculation of the average MS coefficient reported in the last column. The most representative replicate, the one with the nearest MS coefficient to the average MS, is highlighted in green.

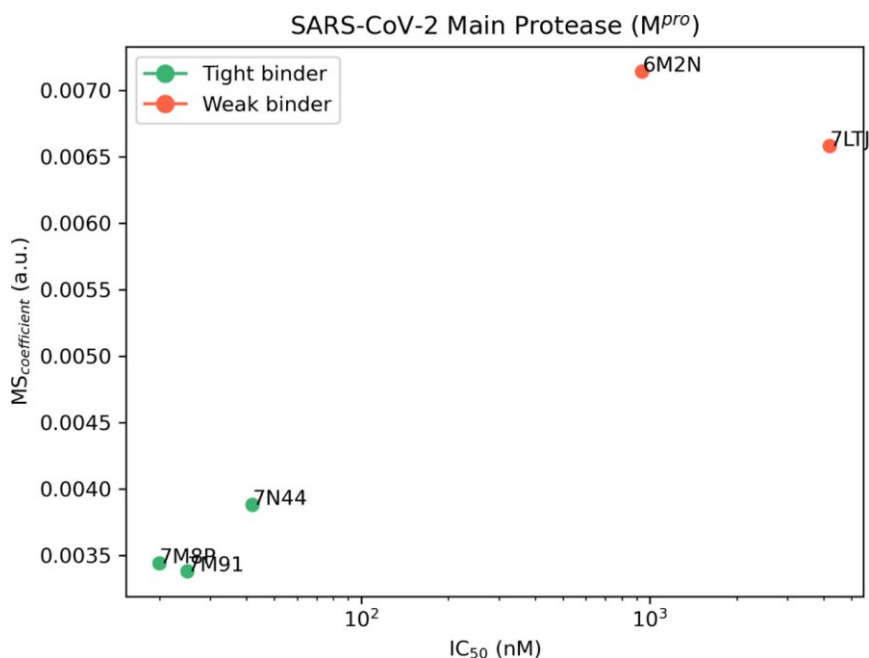


Figure 5. Aggregate results of the TTMD simulations performed on the five investigated SARS-CoV-2 M^{pro} complexes. On the horizontal axis, the experimentally determined affinity value (expressed as IC₅₀) is reported, while on the vertical axis, the average MS coefficient is indicated. Each dot is color-coded as green or red and classified as a tight or weak binder based on the MS cutoff value of 0.004.

The analysis of the TTMD simulations performed on M^{pro} protein–ligand complexes matches the one

already shown for the previous cases. Once again, the ligands characterized by the lowest experimental binding affinity (0.94 μ M for 6M2N and 4.2 μ M for 7LTJ) are the ones associated with the highest MS coefficient (0.0071 and 0.0066, respectively). Accordingly, the ligands that present the highest binding affinity are associated with low MS coefficients, indicative of a persistent binding mode. Even for the SARS-CoV-2 M^{Pro}, it is possible to reutilize the previously determined threshold value ($MS < 0.004$) to separate the strong and the weak binders. Regarding the evolution of the interaction pattern for the protein–ligand complexes that were investigated, it is possible to notice how the most persistent ligands are characterized by strong and stable interactions with key residues, such as Met164–Glu166, located in a β -sheet that constitutes the central portion of the binding site, lining several subpockets that precede the catalytic dyad such as S1, S2, and S3, and residues Leu141–Cys145, which line the S1 subpocket and constitute the so-called oxyanion loop, a structure that plays a crucial role in the catalytic cycle of the protease¹⁰¹. On the contrary, interactions with residues lining the S2 and S4 subpocket seem to not be pivotal for the interaction with the catalytic site. As can be seen in Figures S20 and S21, for example, the ligands partially lose their interactions with residues Asp187–Gln192: this causes a slight increase in the ligand RMSD toward the end of the simulation without overall altering the conservation of the native binding mode, as depicted by the interaction fingerprint analysis. Interestingly, in the case of complex 6M2N, a repulsive interaction with Glu166 is present at the beginning of the simulation: considering the pivotal role that this residue portrays both in the dimerization process¹⁰⁸ (it forms a salt bridge through its side chain with the side chain of Ser1 of the second protomer) and in the binding of ligands, this could well explain the low persistence of the native binding mode for this ligand throughout the simulation. A comparison between the representative replicates for complexes 7LTJ and 7M91 is illustrated in Video S4.

Discussion

The TTMD method is an alternative protocol for the qualitative estimation of the protein–ligand complex stability based on the persistence of the native binding mode throughout a series of MD simulations performed at progressively increasing temperatures. To evaluate the protocol capabilities, we performed four different case studies on an equal number of pharmaceutically relevant test cases, that is, protein kinase CK1 δ , protein kinase CK2, PDK2, and SARS-CoV-2 main protease (M^{Pro}). Despite its simplicity, the TTMD workflow was able to correctly discriminate between tight binders (with affinity values in the low nanomolar range) and weak binders (the ones with affinity values superior to the micromolar threshold) by applying an appropriate MS coefficient cutoff. This classification was performed on ligands with different scaffolds and different interaction features, making its application interesting in real-world drug-discovery campaigns, where compounds from different chemical classes are usually identified in the early stages and are then subjected to an iterative optimization of their binding affinity toward the target of interest through chemical modification of their structures.

Contrary to most protocols that aim at predicting or estimating the drug–target residence time or other proxy values for the ligand affinity, TTMD does not require simulating the full unbinding event. Although this results in a rawer prediction compared to other similar protocols, this approach has two major advantages. The first one is that the simulation time is limited and can be accurately estimated right from the start. This is particularly useful in the case of a batch application of the protocol across a library of different compounds resulting from a screening campaign or an optimization process. Moreover, it facilitates the automatization of the process and its incorporation into existing drug-discovery pipelines. The second advantage is that by determining a relative metric (the MS coefficient) rather than an absolute one (i.e., the time required to observe the unbinding), there is no need for the definition of an arbitrary cutoff value for the detection of the unbinding event. Concerning this, most protocols exploit geometric descriptors such as the distance between the center of masses of the ligand and the binding site or the distance between the ligand and the protein to define whether the ligand detached from the binding site^{43,46}. This poses the problem of choosing the right distance because, in the case of deep and buried binding sites, the chosen cutoff value could not consider the whole unbinding process, leading once again to an underestimation of the residence time. On the contrary, arbitrarily increasing the distance could elongate the simulation time without improving the prediction accuracy. Furthermore, using an interaction fingerprint-based metric instead of the standard RMSD for monitoring the evolution of the binding mode

throughout the simulation results in lower sensitivity toward the chemical structure of the ligand: as highlighted in some of our trajectory analyses (Figures S14 and S15, e.g.), the presence of some ligand moieties that do not directly interact with the binding pocket or are slightly solvent exposed, leading to a less stable interaction with the target, could lead to an increase in the ligand's RMSD without compromising the key binding determinants of the compounds. This could lead to a false perception of the unbinding event, causing errors in the evaluation of the persistence of the receptor–ligand complex, especially if very low cutoffs are utilized; such is the case in some studies in which classic MD simulations are used as a way to refine docking results and distinguish between native-like poses and decoys^{109,110}.

Another major advantage is related to the accessibility of the method. First, although the protocol in its current form exploits the ACEMD3 program to run MD simulations, it can be easily and readily adapted to be utilized with any other major MD engine such as OpenMM, GROMACS, or AMBER. Second, compared to other approaches, TTMD is easier to implement. For example, contrary to metadynamics-based approaches, where the choice of the CV to monitor is not trivial³⁸, in the case of TTMD, the user only needs to choose a temperature ramp that ensures the conservation of the protein fold by monitoring a simple geometric descriptor such as the protein backbone RMSD. Although some attempts at optimizing the temperature ramp to decrease the simulation time without reducing the accuracy of the method are already going on in our laboratory, the temperature ramp proposed in this article should represent a good starting point for the third-party implementation of the method. Theoretically, increasing the simulation time for each “TTMD step” should provide an increase in the resolution of the technique, but would also result in an increased computational effort. On the contrary, reducing the simulation time for each TTMD step would reduce the computational effort, making the protocol more affordable, especially for those setups where a large number of different ligands are evaluated at a given time but would also flatten the difference in the MS coefficient between the ligands, thus decreasing the sensitivity of the technique. One possible solution could be to use different simulation times at different temperatures, for example, simulating longer steps at lower temperatures and shorter ones at higher temperatures. The pool of test cases provided in this article should, in principle, aid the user in the choice of a non-default temperature ramp since the user could compare the results of its custom temperature ramp with the ones originally obtained and evaluate on its own the performance of a different ramp.

Other than estimating the protein–ligand binding affinity, the TTMD protocol could be easily adapted to perform mechanistic evaluations on the unbinding event by appropriately tuning the temperature ramp and the simulation time to carry on the simulation until the native binding mode is completely lost. Although, in its current form, this protocol is not specifically designed for this purpose, it can already be used to discriminate between different protein–ligand interactions based on their effect on the binding affinity. This indication could be very useful in the generation and refinement of pharmacophore filters, which are commonly used to reduce the false positive rates in docking-based virtual screening campaigns. Another possible application of the TTMD protocol could be to distinguish the native binding pose from a set of decoy ones. This could be particularly useful in the case of fragment compounds, which can have several plausible binding modes and are usually evolved into mature ligands by rationally modifying their scaffold to expand on their binding determinants without altering the existing interaction profile. Work in this sense is already going on in our laboratory and will be the scope of a future paper.

The last element that needs to be addressed in the nearest future is the applicability of the method to membrane proteins: so far, the protocol has been applied only to globular targets, but a wide variety of pharmaceutically relevant targets are membrane systems. Membrane systems are intrinsically more complicated to manage because, other than monitoring the protein fold throughout the simulations, one has to decide how to manage the membrane. A possible solution could be to remove the membrane and treat the protein as soluble, possibly with the implementation of restraints on the atomic positions of atoms outside the binding site⁴³. Evaluations in this sense are already going on in our laboratory to tune the protocol to also be utilized for this class of targets.

Conclusions

In this scientific work, we presented the first application of TTMD, an alternative protocol for the qualitative estimation of protein–ligand complex stability, by monitoring the conservation of the native

ligand-binding mode throughout a series of classic MD simulations performed at progressively increasing temperatures through a scoring function based on protein–ligand interaction fingerprints. Four different test cases regarding the application of the technique to three different pharmaceutically relevant targets were presented. For each case, TTMD simulations were able to distinguish between tight (low nanomolar) and weak (micromolar) binders. The simplicity of the protocol, particularly regarding the choice of user-defined parameters to run the simulations, the agnosticism concerning the selection of the MD engine, and the limited simulation time make it a viable choice for various medicinal chemistry projects, especially as a screening tool in the early stages of drug discovery campaigns. Further work is needed to extend the applicability domain of the technique to membrane proteins, and evaluations in this sense are already going on in our laboratory.

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.2c00995>.

- Additional TTMD use case illustrations including principal chemical properties of the CK1δ, CK2, PDK2, and M^{pro} ligands and analyses performed on representative TTMD trajectories for the complex deposited in the Protein Data Bank with their respective accession codes ([PDF](#))
 - Comparison between the representative TTMD replicate for a weak and a strong binder in the case of protein target CK1δ ([MP4](#)),
 - comparison between the representative TTMD replicate for a weak and a strong binder in the case of protein target CK2 ([MP4](#)),
 - comparison between the representative TTMD replicate for a weak and a strong binder in the case of protein target PDK2 ([MP4](#)), and
 - comparison between the representative TTMD replicate for a weak and a strong binder in the case of protein target SARS-CoV-2 M^{pro} ([MP4](#))
-
- [ci2c00995_si_001.pdf](#)
 - [ci2c00995_si_002.mp4](#)
 - [ci2c00995_si_003.mp4](#)
 - [ci2c00995_si_004.mp4](#)
 - [ci2c00995_si_005.mp4](#)

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Data and Software Availability

All molecular structures utilized in this work have been retrieved from the publicly available PDB (www.rcsb.org). The TTMD.py Python code to reproduce the simulations performed in this work, as well as a YAML file to reconstitute the appropriate virtual environment to run them, is available at github.com/molecularmodelingsection/TTMD and released under a permissive MIT license. The TTMD.py script carries on each step of a TTMD simulation, including the system setup and parametrization, the equilibration protocol, the production runs, and associated analyses.

References

1. Ehrlich, P. Address in Pathology, ON CHEMIOTHERAPY. 1913. <https://doi.org/10.1136/bmj.2.2746.353>.

2. Langley, J. N. On the Reaction of Cells and of Nerve-Endings to Certain Poisons, Chiefly as Regards the Reaction of Striated Muscle to Nicotine and to Curari. *J. Physiol.* 1905, 33 (4–5), 374–413. <https://doi.org/10.1113/jphysiol.1905.sp001128>.
3. Kaufmann, S. H. E. Paul Ehrlich: Founder of Chemotherapy. *Nat. Rev. Drug Discov.* 2008, 7 (5), 373–373. <https://doi.org/10.1038/nrd2582>.
4. Maehle, A.-H. "Receptive Substances": John Newport Langley (1852–1925) and His Path to a Receptor Theory of Drug Action. *Med. Hist.* 2004, 48 (2), 153–174. <https://doi.org/10.1017/S0025727300000090>.
5. Bennett, M. R. The Concept of Transmitter Receptors: 100 Years On. *Neuropharmacology* 2000, 39 (4), 523–546. [https://doi.org/10.1016/S0028-3908\(99\)00137-9](https://doi.org/10.1016/S0028-3908(99)00137-9).
6. Gourley, D. R. H. Isolation and Characterization of Membrane Drug Receptors. In *Progress in Drug Research/Fortschritte der Arzneimittelforschung/Progrès des recherches pharmaceutiques*; Jucker, E., Ed.; Birkhäuser: Basel, 1976; pp 323–346. https://doi.org/10.1007/978-3-0348-7094-8_10.
7. Rang, H. P. The Receptor Concept: Pharmacology's Big Idea. *Br. J. Pharmacol.* 2006, 147 (S1), S9–S16. <https://doi.org/10.1038/sj.bjp.0706457>.
8. Borea, P. A.; Varani, K.; Gessi, S.; Gilli, P.; Dalpiaz, A. Receptor Binding Thermodynamics as a Tool for Linking Drug Efficacy and Affinity. *Il Farm.* 1998, 53 (4), 249–254. [https://doi.org/10.1016/S0014-827X\(98\)00017-2](https://doi.org/10.1016/S0014-827X(98)00017-2).
9. Kairys, V.; Baranauskienė, L.; Kazlauskienė, M.; Matulis, D.; Kazlauskas, E. Binding Affinity in Drug Design: Experimental and Computational Techniques. *Expert Opin. Drug Discov.* 2019, 14 (8), 755–768. <https://doi.org/10.1080/17460441.2019.1623202>.
10. Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E. Molecular Determinants of Drug-Receptor Binding Kinetics. *Drug Discov. Today* 2013, 18 (13–14), 667–673. <https://doi.org/10.1016/J.DRUDIS.2013.02.007>.
11. Copeland, R. A. The Drug–Target Residence Time Model: A 10-Year Retrospective. *Nat. Rev. Drug Discov.* 2015 152 2015, 15 (2), 87–95. <https://doi.org/10.1038/nrd.2015.18>.
12. Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug–Target Residence Time and Its Implications for Lead Optimization. *Nat. Rev. Drug Discov.* 2006 59 2006, 5 (9), 730–739. <https://doi.org/10.1038/nrd2082>.
13. Swinney, D. C. The Role of Binding Kinetics in Therapeutically Useful Drug Action. *Curr. Opin. Drug Discov. Dev.* 2009, 12 (1), 31–39.
14. Zhang, R.; Monsma, F. The Importance of Drug-Target Residence Time. *Curr. Opin. Drug Discov. Dev.* 2009, 12 (4), 488–496.
15. Lu, H.; Tonge, P. J. Drug–Target Residence Time: Critical Information for Lead Optimization. *Curr. Opin. Chem. Biol.* 2010, 14 (4), 467–474. <https://doi.org/10.1016/j.cbpa.2010.06.176>.
16. Hill, T. L. Effect of Rotation on the Diffusion-Controlled Rate of Ligand-Protein Association. *Proc. Natl. Acad. Sci.* 1975, 72 (12), 4918–4922. <https://doi.org/10.1073/pnas.72.12.4918>.
17. Guo, D.; Hillger, J. M.; Ijzerman, A. P.; Heitman, L. H. Drug–Target Residence Time—A Case for G Protein-Coupled Receptors. *Med. Res. Rev.* 2014, 34 (4), 856–892. <https://doi.org/10.1002/MED.21307>.
18. Cusack, K. P.; Wang, Y.; Hoemann, M. Z.; Marjanovic, J.; Heym, R. G.; Vasudevan, A. Design Strategies to Address Kinetics of Drug Binding and Residence Time. *Bioorg. Med. Chem. Lett.* 2015, 25 (10), 2019–2027. <https://doi.org/10.1016/J.BMCL.2015.02.027>.
19. Bernetti, M.; Cavalli, A.; Mollica, L. Protein–Ligand (Un)Binding Kinetics as a New Paradigm for Drug Discovery at the Crossroad between Experiments and Modelling. *MedChemComm* 2017, 8 (3), 534–550. <https://doi.org/10.1039/C6MD00581K>.
20. Meyer-Almes, F. J. Kinetic Binding Assays for the Analysis of Protein–Ligand Interactions. *Drug Discov. Today Technol.* 2015, 17, 1–8. <https://doi.org/10.1016/J.DDTEC.2015.08.004>.
21. Bernetti, M.; Masetti, M.; Rocchia, W.; Cavalli, A. Kinetics of Drug Binding and Residence Time. *Annu. Rev. Phys. Chem.* 2019, 70, 143–171. <https://doi.org/10.1146/ANNUREV-PHYSCHEM-042018-052340>.
22. Hulme, E. C.; Trevethick, M. A. Ligand Binding Assays at Equilibrium: Validation and Interpretation. *Br. J. Pharmacol.* 2010, 161 (6), 1219–1237. <https://doi.org/10.1111/j.1476-5381.2009.00604.x>.
23. Sridharan, R.; Zuber, J.; Connelly, S. M.; Mathew, E.; Dumont, M. E. Fluorescent Approaches for Understanding Interactions of Ligands with G Protein Coupled Receptors. *Biochim. Biophys. Acta* 2014, 1838 (1 0 0), 15. <https://doi.org/10.1016/J.BBAMEM.2013.09.005>.
24. Ilien, B.; Franchet, C.; Bernard, P.; Morisset, S.; Odile Weill, C.; Bourguignon, J.-J.; Hibert, M.; Galzi, J.-L. Fluorescence Resonance Energy Transfer to Probe Human M1 Muscarinic Receptor Structure and Drug Binding Properties. *J. Neurochem.* 2003, 85 (3), 768–778. <https://doi.org/10.1046/j.1471-4159.2003.01717.x>.
25. Stoddart, L. A.; Johnstone, E. K. M.; Wheal, A. J.; Goulding, J.; Robers, M. B.; Machleidt, T.; Wood, K. V.; Hill, S. J.; Pflieger, K. D. G. Application of BRET to Monitor Ligand Binding to GPCRs. *Nat. Methods* 2015, 12 (7), 661–663. <https://doi.org/10.1038/nmeth.3398>.
26. Núñez, S.; Venhorst, J.; Kruse, C. G. Target–Drug Interactions: First Principles and Their Application to Drug Discovery. *Drug Discov. Today* 2012, 17 (1), 10–22. <https://doi.org/10.1016/j.drudis.2011.06.013>.
27. Patching, S. G. Surface Plasmon Resonance Spectroscopy for Characterisation of Membrane Protein-Ligand Interactions and Its Potential for Drug Discovery. *Biochim. Biophys. Acta* 2014, 1838 (1 Pt A), 43–55. <https://doi.org/10.1016/J.BBAMEM.2013.04.028>.
28. Millet, O.; Bernadó, P.; Garcia, J.; Rizo, J.; Pons, M. NMR Measurement of the off Rate from the First Calcium-Binding Site of the Synaptotagmin I C2A Domain. *FEBS Lett.* 2002, 516 (1–3), 93–96. [https://doi.org/10.1016/S0014-5793\(02\)02508-5](https://doi.org/10.1016/S0014-5793(02)02508-5).
29. Gronewold, T. M. A.; Baumgartner, A.; Hierer, J.; Sierra, S.; Blind, M.; Schäfer, F.; Blümer, J.; Tillmann, T.; Kiwitz, A.; Kaiser, R.; Zabe-Kühn, M.; Quandt, E.; Famulok, M. Kinetic Binding Analysis of Aptamers Targeting HIV-1 Proteins by

- a Combination of a Microbalance Array and Mass Spectrometry (MAMS). *J. Proteome Res.* 2009, 8 (7), 3568–3577. <https://doi.org/10.1021/pr900265r>.
30. Burnouf, D.; Ennifar, E.; Guedich, S.; Puffer, B.; Hoffmann, G.; Bec, G.; Disdier, F.; Baltzinger, M.; Dumas, P. kinITC: A New Method for Obtaining Joint Thermodynamic and Kinetic Data by Isothermal Titration Calorimetry. *J. Am. Chem. Soc.* 2012, 134 (1), 559–565. <https://doi.org/10.1021/ja209057d>.
31. Li, D.; Chen, L.; Wang, R.; Liu, R.; Ge, G. Synergetic Determination of Thermodynamic and Kinetic Signatures Using Isothermal Titration Calorimetry: A Full-Curve-Fitting Approach. *Anal. Chem.* 2017, 89 (13), 7130–7138. <https://doi.org/10.1021/acs.analchem.7b01091>.
32. Acker, M. G.; Auld, D. S. Considerations for the Design and Reporting of Enzyme Assays in High-Throughput Screening Applications. *Perspect. Sci.* 2014, 1 (1), 56–73. <https://doi.org/10.1016/j.pisc.2013.12.001>.
33. Bruce, N. J.; Ganotra, G. K.; Kokh, D. B.; Sadiq, S. K.; Wade, R. C. New Approaches for Computing Ligand–Receptor Binding Kinetics. *Curr. Opin. Struct. Biol.* 2018, 49, 1–10. <https://doi.org/10.1016/j.sbi.2017.10.001>.
34. Nunes-Alves, A.; Kokh, D. B.; Wade, R. C. Recent Progress in Molecular Simulation Methods for Drug Binding Kinetics. *Curr. Opin. Struct. Biol.* 2020, 64, 126–133. <https://doi.org/10.1016/j.sbi.2020.06.022>.
35. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 2016, 59 (9), 4035–4061. <https://doi.org/10.1021/acs.jmedchem.5b01684>.
36. Pan, A. C.; Xu, H.; Palpant, T.; Shaw, D. E. Quantitative Characterization of the Binding and Unbinding of Millimolar Drug Fragments with Molecular Dynamics Simulations. *J. Chem. Theory Comput.* 2017, 13 (7), 3372–3377. <https://doi.org/10.1021/acs.jctc.7b00172>.
37. Schuetz, D. A.; de Witte, W. E. A.; Wong, Y. C.; Knasmueller, B.; Richter, L.; Kokh, D. B.; Sadiq, S. K.; Bosma, R.; Nederpelt, I.; Heitman, L. H.; Segala, E.; Amaral, M.; Guo, D.; Andres, D.; Georgi, V.; Stoddart, L. A.; Hill, S.; Cooke, R. M.; De Graaf, C.; Leurs, R.; Frech, M.; Wade, R. C.; de Lange, E. C. M.; IJzerman, A. P.; Müller-Fahrnow, A.; Ecker, G. F. Kinetics for Drug Discovery: An Industry-Driven Effort to Target Drug Residence Time. *Drug Discov. Today* 2017, 22 (6), 896–911. <https://doi.org/10.1016/j.drudis.2017.02.002>.
38. Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. Kinetics of Protein–Ligand Unbinding: Predicting Pathways, Rates, and Rate-Limiting Steps. *Proc. Natl. Acad. Sci.* 2015, 112 (5), E386–E391. <https://doi.org/10.1073/pnas.1424461112>.
39. Bortolato, A.; Deflorian, F.; Weiss, D. R.; Mason, J. S. Decoding the Role of Water Dynamics in Ligand–Protein Unbinding: CRF1R as a Test Case. *J. Chem. Inf. Model.* 2015, 55 (9), 1857–1866. <https://doi.org/10.1021/acs.jcim.5b00440>.
40. Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A. Substrate Binding Mechanism of HIV-1 Protease from Explicit-Solvent Atomistic Simulations. *J. Am. Chem. Soc.* 2009, 131 (33), 11811–11818. <https://doi.org/10.1021/ja903045y>.
41. Sun, H.; Li, Y.; Shen, M.; Li, D.; Kang, Y.; Hou, T. Characterizing Drug-Target Residence Time with Metadynamics: How to Achieve Dissociation Rate Efficiently without Losing Accuracy against Time-Consuming Approaches. *J. Chem. Inf. Model.* 2017, 57 (8), 1895–1906. <https://doi.org/10.1021/ACS.JCIM.7B00075>/ASSET/IMAGES/LARGE/CI-2017-00075P_0004.JPEG.
42. Bussi, G.; Laio, A. Using Metadynamics to Explore Complex Free-Energy Landscapes. *Nat. Rev. Phys.* 2020, 2 (4), 200–212. <https://doi.org/10.1038/s42254-020-0153-0>.
43. Mollica, L.; Decherchi, S.; Zia, S. R.; Gaspari, R.; Cavalli, A.; Rocchia, W. Kinetics of Protein-Ligand Unbinding via Smoothed Potential Molecular Dynamics Simulations. *Sci. Rep.* 2015, 5 (1), 11539. <https://doi.org/10.1038/srep11539>.
44. Schuetz, D. A.; Bernetti, M.; Bertazzo, M.; Musil, D.; Eggenweiler, H.-M.; Recanatini, M.; Masetti, M.; Ecker, G. F.; Cavalli, A. Predicting Residence Time and Drug Unbinding Pathway through Scaled Molecular Dynamics. *J. Chem. Inf. Model.* 2019, 59 (1), 535–549. <https://doi.org/10.1021/acs.jcim.8b00614>.
45. Bernetti, M.; Rosini, E.; Masetti, M.; Pollegioni, L.; Recanatini, M. Binding Residence Time through Scaled Molecular Dynamics: A Prospective Application to hDAAO Inhibitors. *J. Chem. Inf. Model.* 2018, 58 (11), 2255–2265. <https://doi.org/10.1021/ACS.JCIM.8B00518>.
46. Kokh, D. B.; Amaral, M.; Bomke, J.; Grädler, U.; Musil, D.; Buchstaller, H. P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F.; Bianciotto, M.; Rak, A.; Wade, R. C. Estimation of Drug-Target Residence Times by τ -Random Acceleration Molecular Dynamics Simulations. *J. Chem. Theory Comput.* 2018, 14 (7), 3859–3869. <https://doi.org/10.1021/ACS.JCTC.8B00230>.
47. Nunes-Alves, A.; Kokh, D. B.; Wade, R. C. Ligand Unbinding Mechanisms and Kinetics for T4 Lysozyme Mutants from τ RAMD Simulations. *Curr. Res. Struct. Biol.* 2021, 3, 106–111. <https://doi.org/10.1016/j.crsb.2021.04.001>.
48. Kokh, D. B.; Wade, R. C. G Protein-Coupled Receptor–Ligand Dissociation Rates and Mechanisms from τ RAMD Simulations. *J. Chem. Theory Comput.* 2021, 17 (10), 6610–6623. <https://doi.org/10.1021/acs.jctc.1c00641>.
49. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* 2000, 28 (1), 235–242. <https://doi.org/10.1093/NAR/28.1.235>.
50. Molecular Operating Environment (MOE), 2019.01; Chemical Computing Group ULC: 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2021. https://www.chemcomp.com/Research-Citing_MOE.Htm (Accessed Jan 19, 2021).
51. Long, A.; Zhao, H.; Huang, X. Structural Basis for the Interaction between Casein Kinase 1 Delta and a Potent and Selective Inhibitor. *J. Med. Chem.* 2012, 55 (2), 956–960. <https://doi.org/10.1021/JM201387S>.
52. Bank, R. P. D. RCSB PDB - 4TN6: CK1d in complex with inhibitor. <https://www.rcsb.org/structure/4TN6> (accessed 2025-09-08).
53. Ursu, A.; Illich, D. J.; Takemoto, Y.; Porfetye, A. T.; Zhang, M.; Brockmeyer, A.; Janning, P.; Watanabe, N.; Osada, H.; Vetter, I. R.; Ziegler, S.; Schöler, H. R.; Waldmann, H. Epiblastin A Induces Reprogramming of Epiblast Stem Cells

- into Embryonic Stem Cells by Inhibition of Casein Kinase 1. *Cell Chem. Biol.* 2016, 23 (4), 494–507. <https://doi.org/10.1016/J.CHEMBIOL.2016.02.015>.
54. Halekotte, J.; Witt, L.; Ianes, C.; Krüger, M.; Bührmann, M.; Rauh, D.; Pichlo, C.; Brunstein, E.; Luxemburger, A.; Baumann, U.; Knippschild, U.; Bischof, J.; Peifer, C. Optimized 4,5-Diarylimidazoles as Potent/Selective Inhibitors of Protein Kinase CK1 δ and Their Structural Relation to P38 α MAPK. *Molecules* 2017, 22 (4), 522. <https://doi.org/10.3390/molecules22040522>.
 55. Sekiguchi, Y.; Nakaniwa, T.; Kinoshita, T.; Nakanishi, I.; Kitaura, K.; Hirasawa, A.; Tsujimoto, G.; Tada, T. Structural Insight into Human CK2 α in Complex with the Potent Inhibitor Ellagic Acid. *Bioorg. Med. Chem. Lett.* 2009, 19 (11), 2920–2923. <https://doi.org/10.1016/j.bmcl.2009.04.076>.
 56. Raaf, J.; Brunstein, E.; Issinger, O. G.; Niefind, K. The CK2 α /CK2 β Interface of Human Protein Kinase CK2 Harbors a Binding Pocket for Small Molecules. *Chem. Biol.* 2008, 15 (2), 111–117. <https://doi.org/10.1016/J.CHEMBIOL.2007.12.012>.
 57. Battistutta, R.; Cozza, G.; Pierre, F.; Papinutto, E.; Lolli, G.; Sarno, S.; O'Brien, S. E.; Siddiqui-Jain, A.; Haddach, M.; Anderes, K.; Ryckman, D. M.; Meggio, F.; Pinna, L. A. Unprecedented Selectivity and Structural Determinants of a New Class of Protein Kinase CK2 Inhibitors in Clinical Trials for the Treatment of Cancer. *Biochemistry* 2011, 50 (39), 8478–8488. <https://doi.org/10.1021/bi2008382>.
 58. Cozza, G.; Zonta, F.; Dalle Vedove, A.; Venerando, A.; Dall'Acqua, S.; Battistutta, R.; Ruzzene, M.; Lolli, G. Biochemical and Cellular Mechanism of Protein Kinase CK2 Inhibition by Deceptive Curcumin. *FEBS J.* 2020, 287 (9), 1850–1864. <https://doi.org/10.1111/FEBS.15111>.
 59. Tso, S. C.; Qi, X.; Gui, W. J.; Wu, C. Y.; Chuang, J. L.; Wernstedt-Asterholm, I.; Morlock, L. K.; Owens, K. R.; Scherer, P. E.; Williams, N. S.; Tambar, U. K.; Wynn, R. M.; Chuang, D. T. Structure-Guided Development of Specific Pyruvate Dehydrogenase Kinase Inhibitors Targeting the ATP-Binding Pocket. *J. Biol. Chem.* 2014, 289 (7), 4432–4443. <https://doi.org/10.1074/JBC.M113.533885>.
 60. Moore, J. H.; Margreitter, C.; Janet, J. P.; Engkvist, O.; de Groot, B. L.; Gapsys, V. Automated Relative Binding Free Energy Calculations from SMILES to $\Delta\Delta G$. *Commun. Chem.* 2023, 6 (1), 1–12. <https://doi.org/10.1038/s42004-023-00859-9>.
 61. Tso, S. C.; Lou, M.; Wu, C. Y.; Gui, W. J.; Chuang, J. L.; Morlock, L. K.; Williams, N. S.; Wynn, R. M.; Qi, X.; Chuang, D. T. Development of Dihydroxyphenyl Sulfonylisindoline Derivatives as Liver-Targeting Pyruvate Dehydrogenase Kinase Inhibitors. *J. Med. Chem.* 2017, 60 (3), 1142–1150. <https://doi.org/10.1021/ACS.JMEDCHEM.6B01540>.
 62. Brough, P. A.; Baker, L.; Bedford, S.; Brown, K.; Chavda, S.; Chell, V.; D'Alessandro, J.; Davies, N. G. M.; Davis, B.; Le Strat, L.; Macias, A. T.; Maddox, D.; Mahon, P. C.; Massey, A. J.; Matassova, N.; McKenna, S.; Meissner, J. W. G.; Moore, J. D.; Murray, J. B.; Northfield, C. J.; Parry, C.; Parsons, R.; Roughley, S. D.; Shaw, T.; Simmonite, H.; Stokes, S.; Surgenor, A.; Stefaniak, E.; Robertson, A.; Wang, Y.; Webb, P.; Whitehead, N.; Wood, M. Application of Off-Rate Screening in the Identification of Novel Pan-Isoform Inhibitors of Pyruvate Dehydrogenase Kinase. *J. Med. Chem.* 2017, 60 (6), 2271–2286. <https://doi.org/10.1021/ACS.JMEDCHEM.6B01478>.
 63. Akaki, T.; Bessho, Y.; Ito, T.; Fujioka, S.; Ubukata, M.; Mori, G.; Yamanaka, K.; Orita, T.; Doi, S.; Iwanaga, T.; Ikegashira, K.; Hantani, Y.; Nakanishi, I.; Adachi, T. Fragment-Based Lead Discovery to Identify Novel Inhibitors That Target the ATP Binding Site of Pyruvate Dehydrogenase Kinases. *Bioorg. Med. Chem.* 2021, 44. <https://doi.org/10.1016/J.BMC.2021.116283>.
 64. Su, H.; Yao, S.; Zhao, W.; Li, M.; Liu, J.; Shang, W.; Xie, H.; Ke, C.; Hu, H.; Gao, M.; Yu, K.; Liu, H.; Shen, J.; Tang, W.; Zhang, L.; Xiao, G.; Ni, L.; Wang, D.; Zuo, J.; Jiang, H.; Bai, F.; Wu, Y.; Ye, Y.; Xu, Y. Anti-SARS-CoV-2 Activities in Vitro of Shuanghuanglian Preparations and Bioactive Ingredients. *Acta Pharmacol. Sin.* 2020, 41 (9), 1167–1177. <https://doi.org/10.1038/s41401-020-0483-6>.
 65. Clyde, A.; Galanie, S.; Kneller, D. W.; Ma, H.; Babuji, Y.; Blaiszik, B.; Brace, A.; Brettin, T.; Chard, K.; Chard, R.; Coates, L.; Foster, I.; Hauner, D.; Kertesz, V.; Kumar, N.; Lee, H.; Li, Z.; Merzky, A.; Schmidt, J. G.; Tan, L.; Titov, M.; Trifan, A.; Turilli, M.; Van Dam, H.; Chennubhotla, S. C.; Jha, S.; Kovalevsky, A.; Ramanathan, A.; Head, M. S.; Stevens, R. High-Throughput Virtual Screening and Validation of a SARS-CoV-2 Main Protease Noncovalent Inhibitor. *J. Chem. Inf. Model.* 2022, 62 (1), 116–128. <https://doi.org/10.1021/acs.jcim.1c00851>.
 66. Deshmukh, M. G.; Ippolito, J. A.; Zhang, C.-H.; Stone, E. A.; Reilly, R. A.; Miller, S. J.; Jorgensen, W. L.; Anderson, K. S. Structure-Guided Design of a Perampanel-Derived Pharmacophore Targeting the SARS-CoV-2 Main Protease. *Structure* 2021, 29 (8), 823–833.e5. <https://doi.org/10.1016/j.str.2021.06.002>.
 67. Zhang, C.-H.; Stone, E. A.; Deshmukh, M.; Ippolito, J. A.; Ghahremanpour, M. M.; Tirado-Rives, J.; Spasov, K. A.; Zhang, S.; Takeo, Y.; Kudalkar, S. N.; Liang, Z.; Isaacs, F.; Lindenbach, B.; Miller, S. J.; Anderson, K. S.; Jorgensen, W. L. Potent Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular Sculpting of the Drug Perampanel Guided by Free Energy Perturbation Calculations. *ACS Cent. Sci.* 2021, 7 (3), 467–475. <https://doi.org/10.1021/acscentsci.1c00039>.
 68. Cozza, G.; Bonvini, P.; Zorzi, E.; Poletto, G.; Pagano, M. A.; Sarno, S.; Donella-Deana, A.; Zagotto, G.; Rosolen, A.; Pinna, L. A.; Meggio, F.; Moro, S. Identification of Ellagic Acid as Potent Inhibitor of Protein Kinase CK2: A Successful Example of a Virtual Screening Application. *J. Med. Chem.* 2006, 49 (8), 2363–2366. <https://doi.org/10.1021/jm060112m>.
 69. Cozza, G.; Gianoncelli, A.; Bonvini, P.; Zorzi, E.; Pasquale, R.; Rosolen, A.; Pinna, L. A.; Meggio, F.; Zagotto, G.; Moro, S. Urolithin as a Converging Scaffold Linking Ellagic Acid and Coumarin Analogues: Design of Potent Protein Kinase CK2 Inhibitors. *ChemMedChem* 2011, 6 (12), 2273–2286. <https://doi.org/10.1002/cmdc.201100338>.
 70. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* 1996, 14 (1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).

71. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* 2005, 26 (16), 1668–1688. <https://doi.org/10.1002/JCC.20290>.
72. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 2015, 11 (8), 3696–3713. https://doi.org/10.1021/ACS.JCTC.5B00255/SUPPL_FILE/CT5B00255_SI_001.PDF.
73. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 2004, 25 (9), 1157–1174. <https://doi.org/10.1002/JCC.20035>.
74. Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* 2002, 23 (16), 1623–1641. <https://doi.org/10.1002/jcc.10128>.
75. Jorgensen, W. L.; Tirado-Rives, J. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimization for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* 1988, 110 (6), 1657–1666. <https://doi.org/10.1021/ja00214a001>.
76. Davidchack, R. L.; Handel, R.; Tretyakov, M. V. Langevin Thermostat for Rigid Body Dynamics. *J. Chem. Phys.* 2009, 130 (23). <https://doi.org/10.1063/1.3149788/925280>.
77. Krättiler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A Fast SHAKE Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. *J. Comput. Chem.* 2001, 22 (5), 501–508. [https://doi.org/10.1002/1096-987X\(20010415\)22:5<501::AID-JCC1021>3.0.CO;2-V](https://doi.org/10.1002/1096-987X(20010415)22:5<501::AID-JCC1021>3.0.CO;2-V).
78. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 1995, 103 (19), 8577–8593. <https://doi.org/10.1063/1.470117>.
79. Faller, R.; Pablo, J. J. D.; Chem, J.; Phys, (. Constant Pressure Hybrid Molecular Dynamics–Monte Carlo Simulations. *J. Chem. Phys.* 2002, 116 (1), 55–59. <https://doi.org/10.1063/1.1420460>.
80. Harvey, M. J.; Giupponi, G.; Fabritius, G. D. ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* 2009, 5 (6), 1632–1639. https://doi.org/10.1021/CT9000685/ASSET/IMAGES/LARGE/CT-2009-000685_0004.JPEG.
81. Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLOS Comput. Biol.* 2017, 13 (7), e1005659. <https://doi.org/10.1371/journal.pcbi.1005659>.
82. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* 2011, 32 (10), 2319–2327. <https://doi.org/10.1002/jcc.21787>.
83. Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domański, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python Sci. Conf.* 2016, 98–105. <https://doi.org/10.25080/Majora-629e541a-00e>.
84. Wójcikowski, M.; Zielenkiewicz, P.; Siedlecki, P. Open Drug Discovery Toolkit (ODDT): A New Open-Source Player in the Drug Discovery Field. *J. Cheminformatics* 2015, 7 (1), 26. <https://doi.org/10.1186/S13321-015-0078-2>.
85. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Implementing a Scoring Function Based on Interaction Fingerprint for Autogrow4: Protein Kinase CK1δ as a Case Study. *Front. Mol. Biosci.* 2022, 9, 629. <https://doi.org/10.3389/fmolb.2022.909499>.
86. Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* 2020, 153 (4), 44130. https://doi.org/10.1063/5.0014475/16709547/044130_1_ACCEPTED_MANUSCRIPT.PDF.
87. Knippschild, U.; Gocht, A.; Wolff, S.; Huber, N.; Löhler, J.; Stöter, M. The Casein Kinase 1 Family: Participation in Multiple Cellular Processes in Eukaryotes. *Cell. Signal.* 2005, 17 (6), 675–689. <https://doi.org/10.1016/j.cellsig.2004.12.011>.
88. Xu, P.; Ianes, C.; Gärtner, F.; Liu, C.; Burster, T.; Bakulev, V.; Rachidi, N.; Knippschild, U.; Bischof, J. Structure, Regulation, and (Patho-)Physiological Functions of the Stress-Induced Protein Kinase CK1 Delta (CSNK1D). *Gene* 2019, 715, 144005. <https://doi.org/10.1016/j.gene.2019.144005>.
89. Perez, D. I.; Gil, C.; Martinez, A. Protein Kinases CK1 and CK2 as New Targets for Neurodegenerative Diseases. <https://doi.org/10.1002/med.20207>.
90. Cescon, E.; Bolcato, G.; Federico, S.; Bissaro, M.; Valentini, A.; Ferlin, M. G.; Spalluto, G.; Sturlese, M.; Moro, S. Scaffold Repurposing of In-House Chemical Library toward the Identification of New Casein Kinase 1 δ Inhibitors. *ACS Med. Chem. Lett.* 2020, 11 (6), 1168–1174. <https://doi.org/10.1021/acsmchemlett.0c00028>.
91. Bolcato, G.; Cescon, E.; Pavan, M.; Bissaro, M.; Bassani, D.; Federico, S.; Spalluto, G.; Sturlese, M.; Moro, S. A Computational Workflow for the Identification of Novel Fragments Acting as Inhibitors of the Activity of Protein Kinase CK1δ. *Int. J. Mol. Sci.* 2021, 22 (18), 9741. <https://doi.org/10.3390/ijms22189741>.
92. Venerando, A.; Ruzzene, M.; Pinna, L. A. Casein Kinase: The Triple Meaning of a Misnomer. *Biochem. J.* 2014, 460 (2), 141–156. <https://doi.org/10.1042/BJ20140178>.
93. Meggio, F.; Pinna, L. A. One-Thousand-and-One Substrates of Protein Kinase CK2? *FASEB J.* 2003, 17 (3), 349–368. <https://doi.org/10.1096/fj.02-0473rev>.
94. Borgo, C.; D'Amore, C.; Sarno, S.; Salvi, M.; Ruzzene, M. Protein Kinase CK2: A Potential Therapeutic Target for Diverse Human Diseases. *Signal Transduct. Target. Ther.* 2021, 6 (1), 183. <https://doi.org/10.1038/s41392-021-00567-7>.

95. Atas, E.; Oberhuber, M.; Kenner, L. The Implications of PDK1–4 on Tumor Energy Metabolism, Aggressiveness and Therapy Resistance. *Front. Oncol.* 2020, 10. <https://doi.org/10.3389/fonc.2020.583217>.
96. Zhang, S.; Hulver, M. W.; McMillan, R. P.; Cline, M. A.; Gilbert, E. R. The Pivotal Role of Pyruvate Dehydrogenase Kinases in Metabolic Flexibility. *Nutr. Metab.* 2014, 11 (1), 10. <https://doi.org/10.1186/1743-7075-11-10>.
97. Patel, M. S.; Korotchikina, L. G. Regulation of the Pyruvate Dehydrogenase Complex. *Biochem. Soc. Trans.* 2006, 34 (2), 217–222. <https://doi.org/10.1042/BST0340217>.
98. Wang, C.; Horby, P. W.; Hayden, F. G.; Gao, G. F. A Novel Coronavirus Outbreak of Global Health Concern. *The Lancet* 2020, 395 (10223), 470–473. [https://doi.org/10.1016/S0140-6736\(20\)30185-9](https://doi.org/10.1016/S0140-6736(20)30185-9).
99. COVID - Coronavirus Statistics - Worldometer. <https://www.worldometers.info/coronavirus/> (accessed 2025-09-11).
100. Dai, W.; Zhang, B.; Jiang, X.-M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F.; Li, C.; Li, Y.; Bai, F.; Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L.-K.; Xu, Y.; Yang, H.; Liu, H. Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease. *Science* 2020, 368 (6497), 1331–1335. <https://doi.org/10.1126/science.abb4489>.
101. Fornasier, E.; Macchia, M. L.; Giachin, G.; Sosis, A.; Pavan, M.; Sturlese, M.; Salata, C.; Moro, S.; Gatto, B.; Bellanda, M.; Battistutta, R. A New Inactive Conformation of SARS-CoV-2 Main Protease. *Acta Crystallogr. Sect. Struct. Biol.* 2022, 78 (3), 363–378. <https://doi.org/10.1107/S2059798322000948>.
102. Ullrich, S.; Nitsche, C. The SARS-CoV-2 Main Protease as Drug Target. *Bioorg. Med. Chem. Lett.* 2020, 30 (17), 127377. <https://doi.org/10.1016/j.bmcl.2020.127377>.
103. Consortium, T. C. M.; Achdout, H.; Aimon, A.; Bar-David, E.; Barr, H.; Ben-Shmuel, A.; Bennett, J.; Bilenko, V. A.; Boby, M. L.; Borden, B.; Bowman, G. R.; Brun, J.; Bvnbs, S.; Calmiano, M.; Carbery, A.; Carney, D.; Cattermole, E.; Chang, E.; Chernyshenko, E.; Chodera, J. D.; Clyde, A.; Coffland, J. E.; Cohen, G.; Cole, J.; Contini, A.; Cox, L.; Cvitkovic, M.; Dias, A.; Donckers, K.; Dotson, D. L.; Douangamath, A.; Duberstein, S.; Dudgeon, T.; Dunnett, L.; Eastman, P. K.; Erez, N.; Eyermann, C. J.; Fairhead, M.; Fate, G.; Fearon, D.; Fedorov, O.; Ferla, M.; Fernandes, R. S.; Ferrins, L.; Foster, R.; Foster, H.; Gabizon, R.; Garcia-Sastre, A.; Gawriljuk, V. O.; Gehrtz, P.; Gileadi, C.; Giroud, C.; Glass, W. G.; Glen, R.; Glinert, I.; Godoy, A. S.; Gorichko, M.; Gorrie-Stone, T.; Griffen, E. J.; Hart, S. H.; Heer, J.; Henry, M.; Hill, M.; Horrell, S.; Huliak, V. D.; Hurley, M. F. D.; Israely, T.; Jajack, A.; Jansen, J.; Jnoff, E.; Jochmans, D.; John, T.; Jonghe, S. D.; Kantsadi, A. L.; Kenny, P. W.; Kiappes, J. L.; Kinakh, S. O.; Koekemoer, L.; Kovar, B.; Krojer, T.; Lee, A.; Lefker, B. A.; Levy, H.; Logvinenko, I. G.; London, N.; Lukacik, P.; Macdonald, H. B.; MacLean, B.; Malla, T. R.; Matviuk, T.; McCorkindale, W.; McGovern, B. L.; Melamed, S.; Melnykov, K. P.; Michurin, O.; Mikolajek, H.; Milne, B. F.; Morris, A.; Morris, G. M.; Morwitzer, M. J.; Moustakas, D.; Nakamura, A. M.; Neto, J. B.; Neyts, J.; Nguyen, L.; Noske, G. D.; Oleinikovas, V.; Oliva, G.; Overheul, G. J.; Owen, D.; Pai, R.; Pan, J.; Paran, N.; Perry, B.; Pingle, M.; Pinjari, J.; Politi, B.; Powell, A.; Psenak, V.; Puni, R.; Rangel, V. L.; Reddi, R. N.; Reid, S. P.; Resnick, E.; Ripka, E. G.; Robinson, M. C.; Robinson, R. P.; Rodriguez-Guerra, J.; Rosales, R.; Rufa, D.; Saar, K.; Saikatendu, K. S.; Schofield, C.; Shafeev, M.; Shaikh, A.; Shi, J.; Shurrush, K.; Singh, S.; Sittner, A.; Skyner, R.; Smalley, A.; Smeets, B.; Smilova, M. D.; Solmesky, L. J.; Spencer, J.; Strain-Damerell, C.; Swamy, V.; Tamir, H.; Tennant, R.; Thompson, W.; Thompson, A.; Tomasio, S.; Tsurupa, I. S.; Tumber, A.; Vakonakis, I.; Rij, R. P. van; Vangeel, L.; Varghese, F. S.; Vaschetto, M.; Vitner, E. B.; Voelz, V.; Volkamer, A.; Delft, F. von; Delft, A. von; Walsh, M.; Ward, W.; Weatherall, C.; Weiss, S.; White, K. M.; Wild, C. F.; Wittmann, M.; Wright, N.; Yahalom-Ronen, Y.; Zaidmann, D.; Zidane, H.; Zitzmann, N. Open Science Discovery of Oral Non-Covalent SARS-CoV-2 Main Protease Inhibitor Therapeutics. *bioRxiv January 30, 2022*, p 2020.10.29.339317. <https://doi.org/10.1101/2020.10.29.339317>.
104. Luttens, A.; Gullberg, H.; Abdurakhmanov, E.; Vo, D. D.; Akaberi, D.; Talibov, V. O.; Nekhotiaeva, N.; Vangeel, L.; De Jonghe, S.; Jochmans, D.; Krambrich, J.; Tas, A.; Lundgren, B.; Gravenfors, Y.; Craig, A. J.; Atilaw, Y.; Sandström, A.; Moodie, L. W. K.; Lundkvist, Å.; van Hemert, M. J.; Neyts, J.; Lennerstrand, J.; Kihlberg, J.; Sandberg, K.; Danielson, U. H.; Carlsson, J. Ultralarge Virtual Screening Identifies SARS-CoV-2 Main Protease Inhibitors with Broad-Spectrum Activity against Coronaviruses. *J. Am. Chem. Soc.* 2022, 144 (7), 2905–2920. <https://doi.org/10.1021/jacs.1c08402>.
105. Ghahremanpour, M. M.; Saar, A.; Tirado-Rives, J.; Jorgensen, W. L. Computation of Absolute Binding Free Energies for Noncovalent Inhibitors with SARS-CoV-2 Main Protease. *J. Chem. Inf. Model.* 2023, 63 (16), 5309–5318. <https://doi.org/10.1021/acs.jcim.3c00874>.
106. Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R. D.; Carlo, A.; Coffman, K. J.; Dantonio, A.; Di, L.; Eng, H.; Ferre, R.; Gajiwala, K. S.; Gibson, S. A.; Greasley, S. E.; Hurst, B. L.; Kadar, E. P.; Kalgutkar, A. S.; Lee, J. C.; Lee, J.; Liu, W.; Mason, S. W.; Noell, S.; Novak, J. J.; Obach, R. S.; Ogilvie, K.; Patel, N. C.; Pettersson, M.; Rai, D. K.; Reese, M. R.; Sammons, M. F.; Sathish, J. G.; Singh, R. S. P.; Steppan, C. M.; Stewart, A. E.; Tuttle, J. B.; Updyke, L.; Verhoest, P. R.; Wei, L.; Yang, Q.; Zhu, Y. An Oral SARS-CoV-2 Mpro Inhibitor Clinical Candidate for the Treatment of COVID-19. *Science* 2021, 374 (6575), 1586–1593. <https://doi.org/10.1126/science.abl4784>.
107. Pavan, M.; Bolcato, G.; Bassani, D.; Sturlese, M.; Moro, S. Supervised Molecular Dynamics (SuMD) Insights into the Mechanism of Action of SARS-CoV-2 Main Protease Inhibitor PF-07321332. *J. Enzyme Inhib. Med. Chem.* 2021, 36 (1), 1645–1649. <https://doi.org/10.1080/14756366.2021.1954919>.
108. Cheng, S.-C.; Chang, G.-G.; Chou, C.-Y. Mutation of Glu-166 Blocks the Substrate-Induced Dimerization of SARS Coronavirus Main Protease. *Biophys. J.* 2010, 98 (7), 1327–1336. <https://doi.org/10.1016/j.bpj.2009.12.4272>.
109. Proctor, E. A.; Yin, S.; Tropsha, A.; Dokholyan, N. V. Discrete Molecular Dynamics Distinguishes Nativelike Binding Poses from Decoys in Difficult Targets. *Biophys. J.* 2012, 102 (1), 144–151. <https://doi.org/10.1016/j.bpj.2011.11.4008>.
110. Liu, K.; Watanabe, E.; Kokubo, H. Exploring the Stability of Ligand Binding Modes to Proteins by Molecular Dynamics Simulations. *J. Comput. Aided Mol. Des.* 2017, 31 (2), 201–211. <https://doi.org/10.1007/s10822-016-0005-2>.

Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement

Silvia Menin, Matteo Pavan, Veronica Salmaso, Mattia Sturlese and Stefano Moro

Menin, S.; Pavan, M.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement. *International Journal of Molecular Sciences* 2023, 24 (4), 3596.
<https://doi.org/10.3390/ijms24043596>.

Abstract

Molecular docking is one of the most widely used computational approaches in the field of rational drug design, thanks to its favorable balance between the rapidity of execution and the accuracy of provided results. Although very efficient in exploring the conformational degrees of freedom available to the ligand, docking programs can sometimes suffer from inaccurate scoring and ranking of generated poses. To address this issue, several post-docking filters and refinement protocols have been proposed throughout the years, including pharmacophore models and molecular dynamics simulations. In this work, we present the first application of Thermal Titration Molecular Dynamics (TTMD), a recently developed method for the qualitative estimation of protein-ligand unbinding kinetics, to the refinement of docking results. TTMD evaluates the conservation of the native binding mode throughout a series of molecular dynamics simulations performed at progressively increasing temperatures with a scoring function based on protein-ligand interaction fingerprints. The protocol was successfully applied to retrieve the native-like binding pose among a set of decoy poses of drug-like ligands generated on four different pharmaceutically relevant biological targets, including casein kinase 1 δ , casein kinase 2, pyruvate dehydrogenase kinase 2, and SARS-CoV-2 main protease.

Introduction

Until the 1980s, the discovery of new drugs was mainly related to empirical observations about the symptoms-relief properties of natural products or serendipity^{1,2}. Nowadays, thanks to the high availability of experimentally determined structures for pharmaceutically relevant biomolecules, the drug discovery process shifted towards a more rational approach³. Specifically, novel therapeutic entities are usually designed by taking into consideration the topological features of the interacting partner, in a process defined as “structure-based drug design”⁴.

Alongside advancements in the field of biophysical approaches for the experimental determination of protein-ligand complex structures, a pivotal role in the last two decades has been played by computational methodologies⁵. These “in silico” approaches not only flank experimental techniques in rationalizing the mechanism of action of active molecules but are routinely used as a less expensive and quicker alternative in the earliest stages of drug discovery pipelines^{6,7}.

Among these methods, one of the most widely adopted is certainly molecular docking, a program that is used by medicinal chemists to predict the most plausible ligand accommodation within the active site of a target biomolecule⁸. Molecular docking programs are composed of two main parts: a search algorithm, which is responsible for exploring the ligand’s conformational degrees of freedom and generating a series of possible ligand orientations within the binding site, and a scoring function, which numerically estimates the quality of protein-ligand interactions for poses generated by the search algorithm⁹.

Although search algorithms are usually very proficient in exploring the conformational space available to the ligand, docking can sometimes produce inaccurate results, due to the limitations of scoring functions¹⁰. For instance, there is a natural variance in docking performances across different biological targets, since each binding site has its peculiar interaction features, which may differ from the ones included in the training set, thus leading to an incorrect pose ranking¹¹.

To mitigate these known issues of molecular docking, several refinement protocols have been proposed throughout the years¹². A first possible strategy is to filter docking poses based on prior knowledge of the binding features that are required for potent and selective binding to the target¹³. This approach can be carried out by creating a pharmacophore model or a protein-ligand interaction fingerprint based on the binding pattern of known binders, i.e., true active compounds, and then retaining only those poses which match this set of features¹⁴. Although this strategy is successful in diminishing the false positive rate in docking calculations, it is not a first-principle method; therefore, its usage is limited to those cases where the target and its interaction features have already been characterized¹⁵.

Another approach is to rescore docking poses with more computationally demanding methods such as the molecular mechanics/generalized Born surface area (MM/GBSA) method for free binding energy determination¹⁶. This method is usually a good compromise between rapidity of execution and quality of calculations, and it can be a useful tool for many cases. Nevertheless, its accuracy is limited by some factors, such as the implicit consideration of solvent molecules and the sensitivity to-wards the polarity and charge properties of both the ligand and the binding site^{17,18}.

Finally, another commonly adopted strategy is to use short-scale molecular dynamics simulations to evaluate the time-dependent pose stability¹⁹. This approach has the advantage of considering both the protein flexibility and the explicit treatment of solvent, which are usually neglected in most docking calculations^{20,21}. Although in principle the MD-based post-docking refinement seems like a sound method to distinguish poses based on their persistence, two aspects can impair its usefulness. First, the short timescale that is used for most applications (tens of nanoseconds) is insufficient to sample the unbinding event for most poses, flattening the difference between the native pose and wrong but stable enough decoys²². Furthermore, in most cases, simple geometric descriptors such as the ligand RMSD or the center-of-mass distance between the ligand and the protein are used to monitor the pose evolution throughout the simulation, for which there is no clear threshold value to define the unbinding event, other than being heavily influenced by the geometry of the binding site and the ligand structure^{23,24}.

In the current work, we present the first application of Thermal Titration Molecular Dynamics (TTMD) as an alternative MD-based post-docking refinement tool. Recently developed as an enhanced sampling MD protocol for the qualitative estimation of protein-ligand unbinding kinetics²⁵, TTMD evaluates the persistence of the native binding mode in a certain protein-ligand complex of interest by combining a series of short MD simulations performed at progressively increasing temperatures with a scoring function based on protein-ligand interaction fingerprints²⁶. The protocol has been applied to twelve different protein-ligand complexes from four different pharmaceutically relevant biological targets, including protein kinase CK1 δ , protein kinase CK2, pyruvate dehydrogenase kinase 2, and SARS-CoV-2 main protease.

Results

To extend the applicability domain of the TTMD protocol as a post-docking refinement tool, 12 different test cases on four different pharmaceutically relevant targets of interest for our laboratory were investigated: CK1 δ , CK2, PDK2, and M^{pro}. Test cases were chosen coherently with the original work, picking up two "tight binders" (IC₅₀/K_d values in the low nanomolar range) and one "weak binder" (IC₅₀/K_d above the low micromolar threshold). For each biomolecular target, three different protein-ligand complexes available on the Protein Data Bank were chosen. A list of each system investigated in this work can be found in Table 6, while details about ligands' physicochemical properties (Tables S1, S4, S7, and S10) and relevant statistics for each simulated system (Table S13) can be found in the Supplementary Materials.

At first, for each protein-ligand system, a self-docking experiment with the PLANTS docking software was executed. The first five docking poses according to the ChemPLP scoring function were used as a starting point for the TTMD simulations. For each investigated docking pose, five independent TTMD replicates

were performed. The results for each test case are reported separately and discussed aggregately afterward in the Discussion section.

Table 1. PDB accession codes for protein-ligand complex structures employed in this work^a.

CK1δ	3UZP ²⁷	5IH6 ²⁸	5MQV ²⁹
CK2	3PE2 ³⁰	6HOU ³¹	3PE1 ³⁰
PDK2	4V25 ³²	7EA0 ³³	5M4M ³⁴
M ^{pro}	7LTJ ³⁵	7M91 ³⁶	7M8P ³⁶

^aStructures are grouped by the target.

Protein Kinase CK1δ

The first test case regarded protein kinase 1 δ, a serine-threonine kinase belonging to the family of CK1 kinases³⁷ that is involved in several neurodegenerative illnesses like Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis³⁸ due to its pleiotropic nature³⁹. The result of the TTMD refinement carried out on docking poses generated for each CK1δ protein-ligand complex is summarized in Figures 1–3, while detailed analysis about representative replicates (Figures S1–S4), and per-replicate reports (Tables S2 and S3) can be found in the Supplementary Materials. For reference, a comparison between the representative TTMD trajectory for the best (pose 1, MD4), and worst (pose 5, MD5) scoring pose for complex 5MQV, according to the MS coefficient, is highlighted in Video S1.

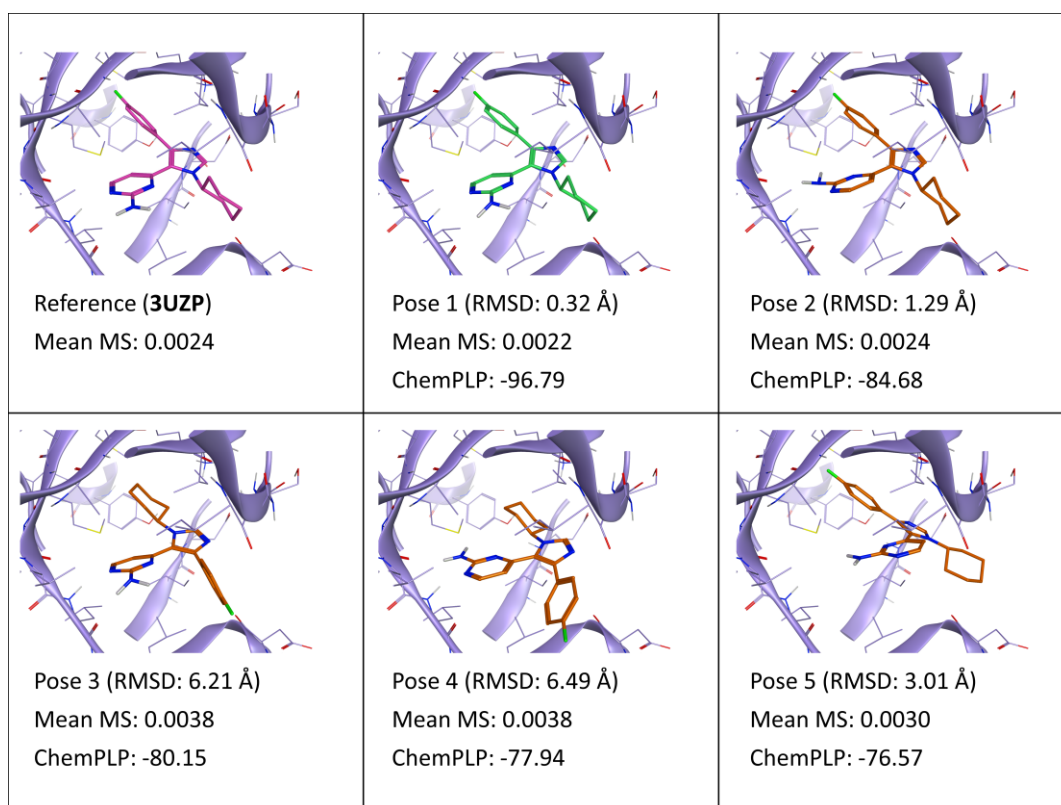


Figure 1. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand 0CK within the catalytic site of the protein kinase CK1δ (PDB ID: 3UZP). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

As can be seen in Figure 1, which summarizes the results of TTMD refinement performed on docking poses of ligand 0CK (PDB ID: 3UZP), although poses are very similar and plausible, the MS value correctly identifies the most correct/native-like poses among the five ones. Particularly, poses 1 and 2, which are practically indistinguishable from a binding pattern perspective (both orientations of the 2-aminopyrimidine moiety ensure a bivalent hydrogen bond interaction with the hinge region of the kinase, while the rest of the molecule is superimposable) have the lowest MS value, which is comparable to the one of the experimentally determined binding mode, coherently with the low RMSD value to the reference. On the contrary, poses 3 and 4, which insert the cyclohexyl moiety within the hydrophobic “selectivity pocket” located behind the hinge region, in the innermost part of the binding site, present a higher MS value, indicating a lower persistence of the native binding mode. Intriguingly, pose 5, which does not show any interaction with the hinge region, has a better MS value compared to poses 3 and 4: this can be explained by the fact that, due to the conserved orientation compared to poses 1 and 2, pose 5 can quite easily interconvert into pose 2 by a slight rotation towards the hinge region, gaining the directional interaction with the hinge without losing any hydrophobic interaction provided by the rest of the molecule.

A similar trend compared to the first case can also be observed for ligand D5Q (PDB ID: 5MQV). As reported in Figure 2, poses 1–2, which only differ for the relative orientation of the solvent-exposed dimethoxyphenylpyrrole moiety, present the lowest MS value, similar to the reference. Pose 3, which is also quite similar to the reference and the first two poses, also displays a good MS coefficient, albeit slightly higher than the first two poses. A possible explanation for this phenomenon is that, although the relative orientation of the dimethoxyphenylpyrrole portion of the ligand is similar to the reference, contributing to a marginally lower RMSD value compared to pose 2, the 2-carboxamide-pyridin core, which provides the bidentate interaction with the hinge region, is mildly shifted away from the hinge. A little rearrangement of the ligand core through the course of the simulation allows interconversion into a reference-like pose, accompanied by an alteration of interaction pattern, which explains the reason why the pose presents good stability but with lesser conservation of its initial binding pattern. As already seen for ligand 0CK, in this case also the MS score penalizes poses that do not insert an aromatic ring within the hydrophobic selectivity pocket, despite having an optimal interaction with the hinge (pose 4). Finally, the MS score heavily penalizes pose 5, which does not retain any interaction found in the reference.

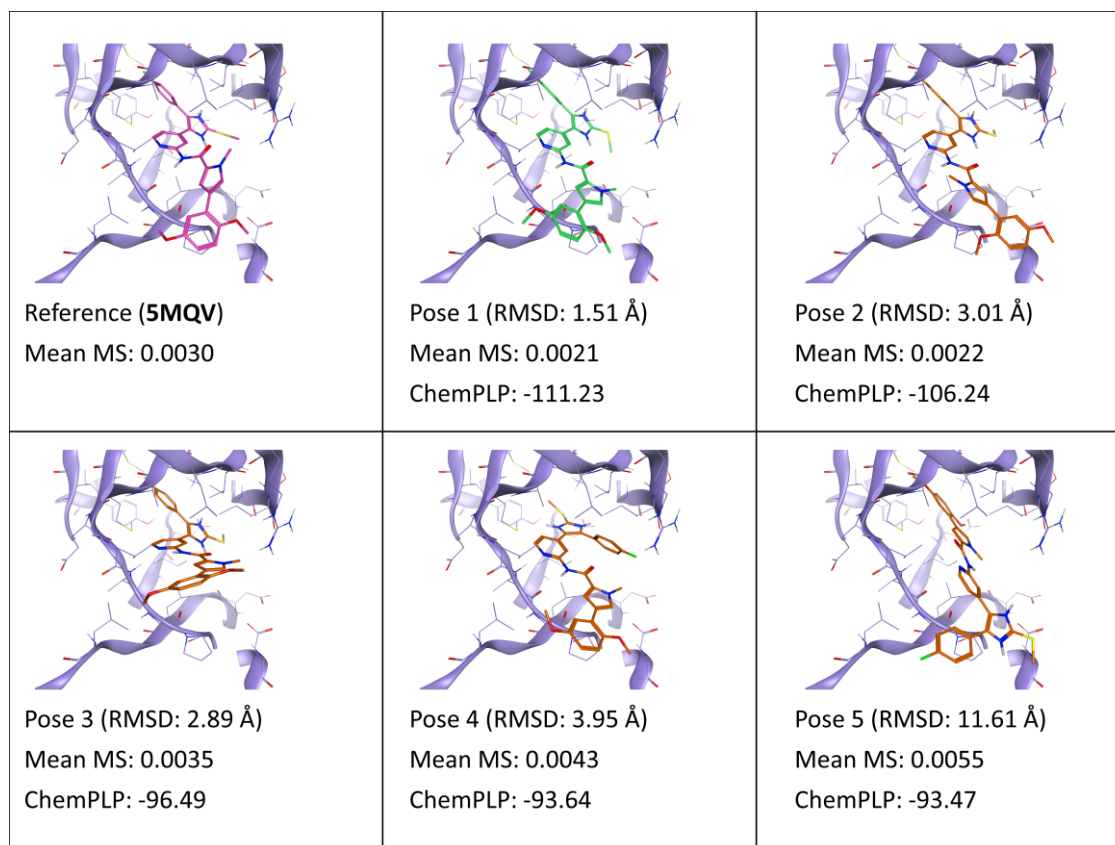


Figure 2. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand D5Q within the catalytic site of the protein kinase CK1 δ (PDB ID: 5MQV). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

Differently from the first two cases, involving “tight binding” ligands, the results of the TTMD refinement were inconclusive for ligand AUG (PDB ID: 5IH6). As can be deduced from Figure 3, the best scoring conformation according to the MS value is pose 2, which presents a reasonable although incorrect binding mode, characterized by three hydrogen bonds with the hinge region. This orientation of the triamine-pteridine core, although allowing for a comparable interaction pattern with the hinge, leads to a suboptimal burial of the hydrophobic core of the molecule within the internal portion of the pocket, especially concerning the bromophenyl moiety. However, in this case, the TTMD refinement is not able to discriminate between the native binding mode and other incorrect ones, as noticeable by the fact that the crystal reference and pose 1, which are practically superimposable, present a high MS value, comparable with pose 3, which is incorrect. Intriguingly, as denoted by Figure S4 (Supplementary Materials), throughout the simulation, pose 1 interconverts into pose 2, which is the final state of the simulation. Considering that the two poses have a similar quality of interaction with the target, it is possible that the two conformations are in equilibrium and that kinetic factors influence the shift towards either of the two. Nevertheless, another plausible explanation for this phenomenon could be that the temperature ramp utilized, which was the same as the one employed in the original publication to discriminate between strong and weak binders, is too aggressive for ligands with low affinity for the targets. The aggressiveness of the ramp, which was designed to facilitate the observation of unbinding events, could flatten the differences between different poses of the same ligand when the crystal pose itself is unstable in those conditions. If this is the case, a different ramp for this category of binders would have to be developed.

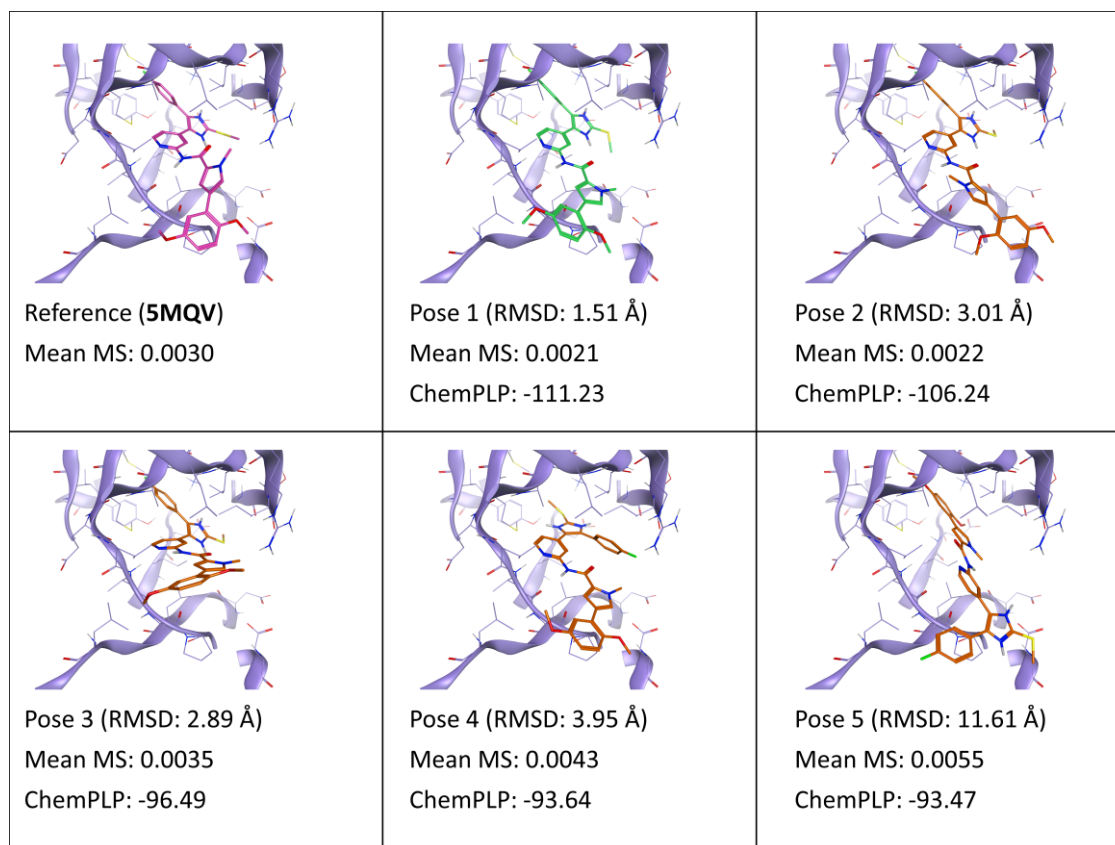


Figure 3. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand AUG within the catalytic site of the protein kinase CK1 δ (PDB ID: 5IH6). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

Protein Kinase CK2

The second test case regarded protein kinase 2, a serine-threonine kinase that represents one of the first identified protein kinases⁴⁰. Similar to CK1, thanks to its pleiotropic nature⁴¹ this kinase is involved in several illnesses such as different types of cancer, various neurodegenerative diseases, and viral infections⁴². The result of the TTMD refinement carried out on docking poses generated for each CK2 protein-ligand complex is summarized in Figures 4–6, while detailed analysis about representative replicates (Figures S5–S8), and per-replicate reports (Tables S5 and S6) can be found in the Supplementary Materials. For reference, a comparison between the representative TTMD trajectory for the best (pose1, MD3) and worst (pose4, MD4) scoring pose for complex 3PE1, according to the MS coefficient, is highlighted in Video S2.

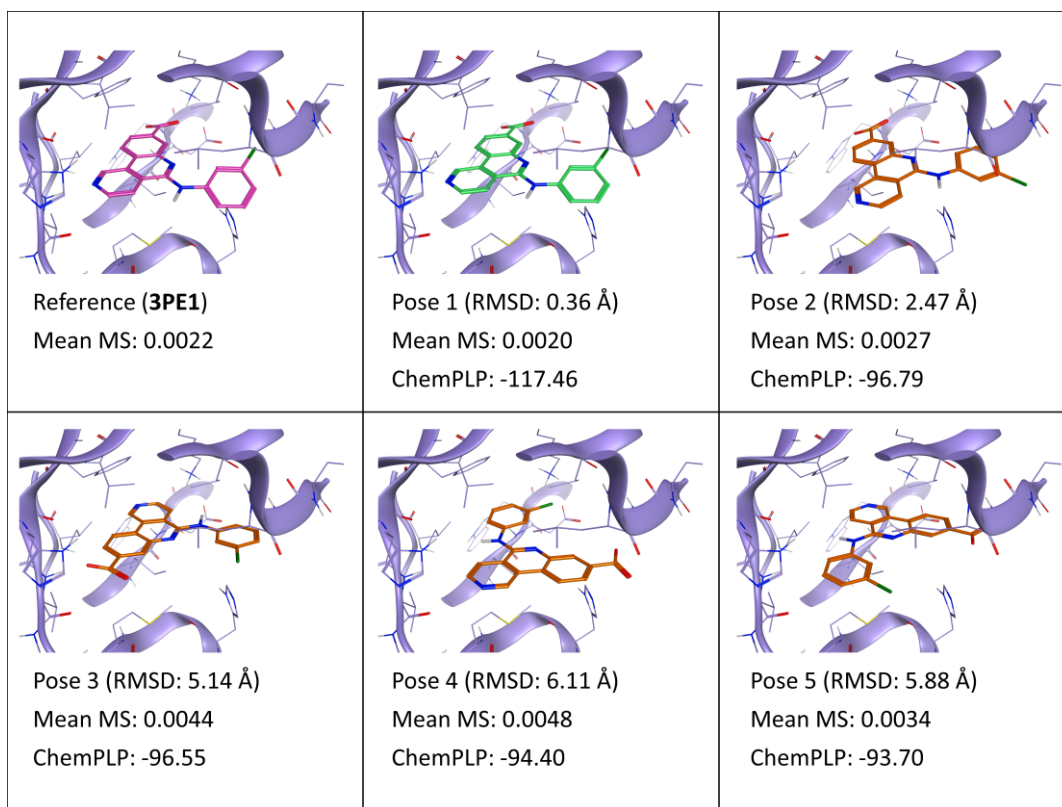


Figure 4. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand 3NG within the catalytic site of the protein kinase CK2 (PDB ID: 3PE1). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

As highlighted by Figure 4, which synthesized the results of the TTMD refinement for ligand 3NG (PDB ID: 3PE1), a similar trend to what has been observed for CK1δ can also be found. Pose 1, which is superimposable to the crystal reference, is associated with the lowest MS value, which is practically identical to the one of the reference. Pose 2, which differs from pose 1 for a slight shift of the naphthyridine core away from the hinge region of the kinase and for the orientation of chlorine atom on the 3-chlorophenyl substituent, displays the second lowest MS value overall, coherently with its suboptimal but overall native-like orientation within the binding site. Pose 3 is penalized by the MS score due to the loss of the pivotal network of direct and water-mediated interactions involving the ligand carboxylate and residues deeply buried in the inner portion of the binding site such as Lys68 and Asp175. Similarly, pose 4 is also heavily penalized by the TTMD-derived score for the blatantly incorrect ligand conformation within the active site: in this case, interactions involving the carboxylate are lost, but the whole naphthyridine core is misplaced. Finally, pose 5 shows an intermediate MS score between the native-like poses and the wrong ones: a possible interpretation for this is that the naphthyridine core of the molecule is deeply buried in a hydrophobic, solvent-shielded area of the binding site, thus compensating for the lack of directional interactions provided by the carboxylate.

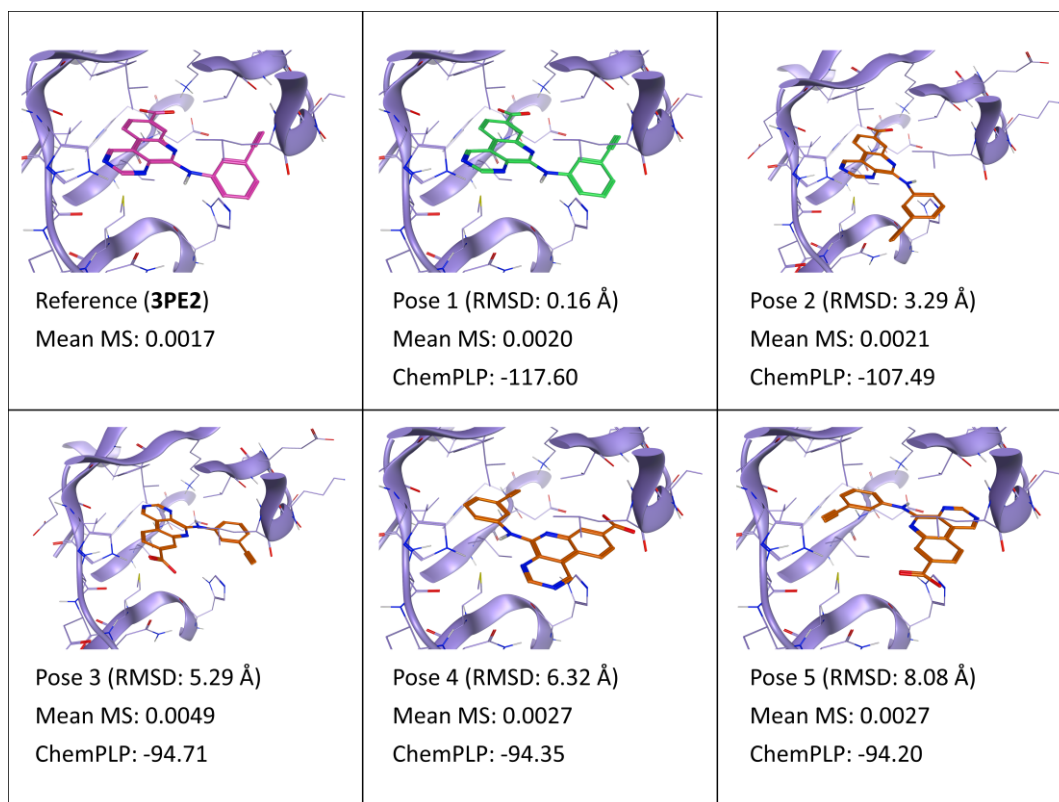


Figure 5. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand E1B within the catalytic site of the protein kinase CK2 (PDB ID: 3PE2). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

As illustrated in Figure 5, results for the TTMD refinement in the case of ligand E1B (PDB ID: 3PE2) recall the ones for ligand 3NG. This is not unexpected, due to the similarities between the two ligands. Particularly, pose 1, which is identical to the reference, presents the lowest MS value possible, which is very similar to the value calculated for the experimentally determined binding mode. The same is true also for pose 2, which only differs for the orientation of the ethynylphenyl substituent but has the same binding pattern for the pyrimidoquinoline core. On the contrary, pose 3, which locates its core near the hinge region of the kinase, as for poses 1 and 2, but presents an incorrect orientation of the carboxylate, leading to the loss of crucial directional interactions with Lys68/Asp175, is linked with a higher MS value. Intriguingly, poses 4 and 5, which are the most different from the reference, present an intermediate MS score, not much higher than the one of native-like poses. A putative explanation for this is that these compounds, 3NG and E1B, both present very rigid and planar structures, enhancing the shape complementarity with the active site and increasing the persistence of short-range hydrophobic and dipolar interactions. Accordingly, despite the lack of directional interactions with the target, both poses 4 and 5 display a good shape and electrostatic complementarity with the pocket, as also indicated by the docking score which is practically identical to that of pose 3, which contributes to the high persistence of the pose despite the lack of the directional interactions provided by the carboxylate.

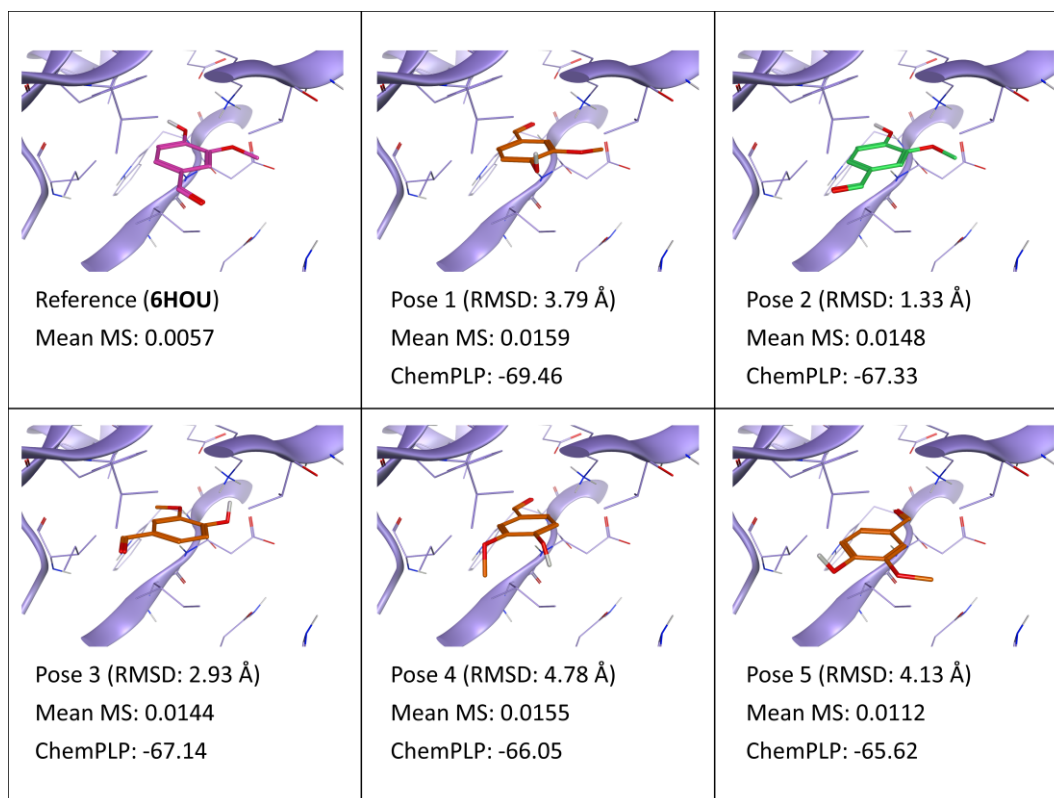


Figure 6. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand V55 within the catalytic site of the protein kinase CK2 (PDB ID: 6HOU). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

In agreement with the results of the TTMD refinement of the CK1δ weak binder, results are inconclusive also for CK2 homologous, ligand V55 (PDB ID: 6HOU). As observable in Figure 6, the MS value is very high for each of the five docking poses, as it is for the reference. High MS values, such as those obtained for V55 poses, indicate that the ligand has rapidly and completely lost its native binding mode in all cases, as also confirmed by detailed trajectory analyses reported in Figures S7 and S8 (Supplementary Materials). This reinforces the hypothesis that the default temperature ramp is too aggressive to be useful in the refinement of poses for weak binders.

Pyruvate Dehydrogenase Kinase 2

The third test case involved pyruvate dehydrogenase kinase 2, a member of the GHKL ATPase/kinase superfamily that has a central control role upon cellular energy metabolism^{43,44} and is involved both in metabolic and neoplastic diseases⁴⁵. The result of the TTMD refinement carried out on docking poses generated for each PDK2 protein-ligand complex is summarized in Figures 7–9, while detailed analysis about representative replicates (Figures S9–S12), and per-replicate reports (Tables S8 and S9) can be found in the Supplementary Materials. For reference, a comparison between the representative TTMD trajectory for the best (pose1, MD1) and worst (pose4, MD2) scoring pose for complex 4V25, according to the MS coefficient, is highlighted in Video S3.

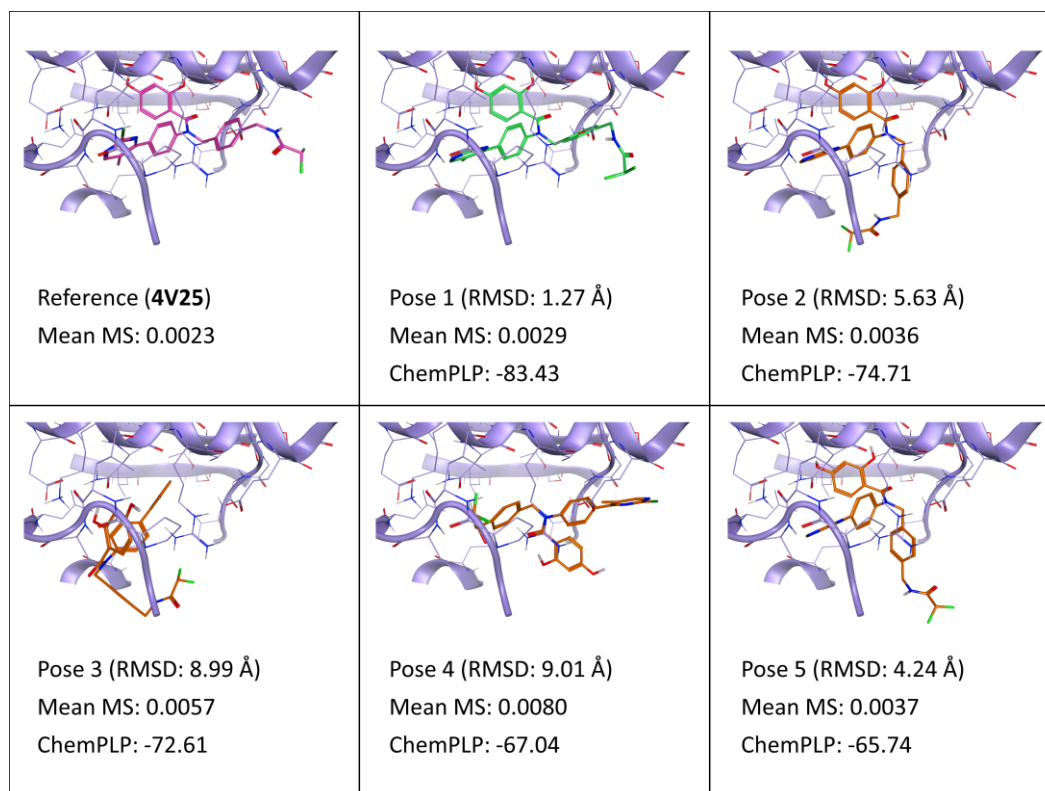


Figure 7. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand SZ6 within the catalytic site of the pyruvate dehydrogenase kinase 2 (PDB ID: 4V25). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

As depicted in Figure 7, which reports the results for the TTMD post-docking refinement performed on docking poses of ligand SZ6 (PDB ID: 4V25), once again the MS score can identify the native-like docking pose among a set of decoys. Particularly, pose 1 (which is the closest to the reference), pose 2, and pose 5, all share the same correct orientation of the core (spanning from the dihydroxybenzamide moiety to the chloromethylpyrimidine one) within the catalytic site, and present significantly lower MS value compared to poses 3 and 4, where the binding pattern is completely divergent compared to the reference. This can be attributed to the ability of these poses to preserve the pivotal network of direct and water-bridged interactions between the dihydroxybenzamide moiety and residues Asp290, Thr354, and Gly294. As mentioned when discussing some of the previous cases, poses 2 and 5, despite having an incorrect orientation of the difluoroacetylaminomethylbenzyl tail, can easily interconvert throughout the simulation into pose 1 and the crystal reference, because this portion of the molecule is completely solvent-exposed and does not contribute to the interaction with the binding site.

Accordingly, like the case of complex 5MQV, where the core of the molecule was placed correctly in most poses and the difference was mainly related to the tail of the molecule which extended out of the binding site, in this case the MS coefficient was also able to distinguish poses with a high RMSD but native-like interaction pattern from poses with high RMSD but incorrect interaction features.

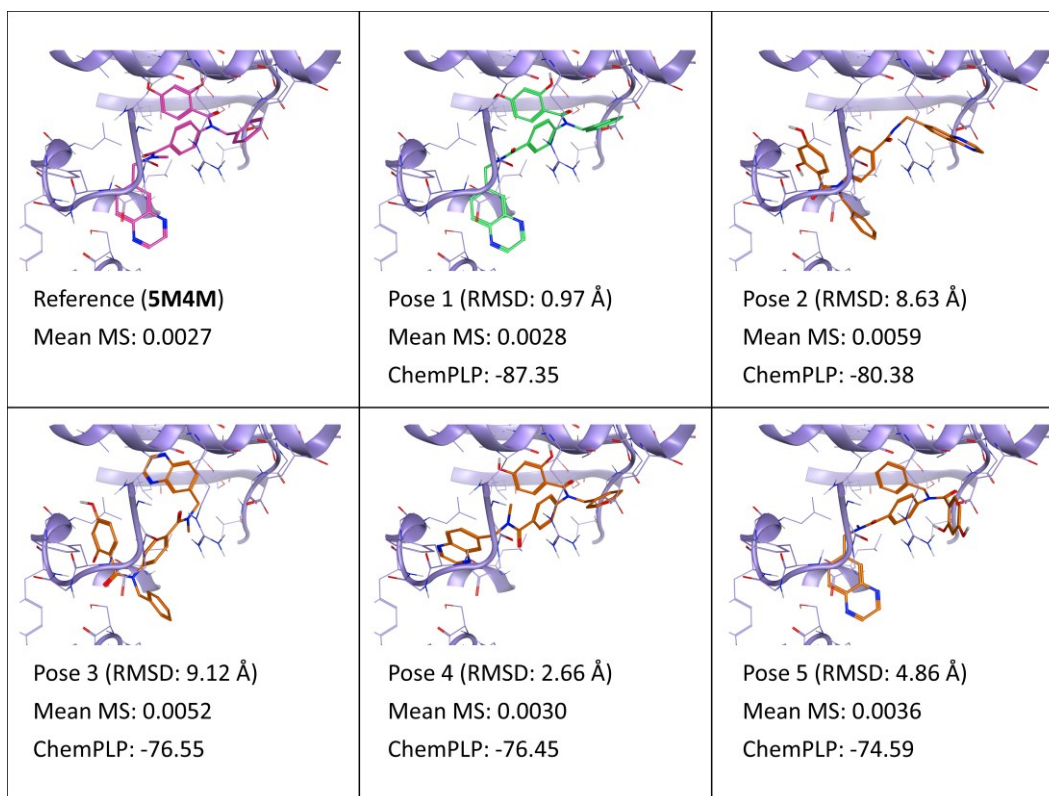


Figure 8. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand 7FW within the catalytic site of the pyruvate dehydrogenase kinase 2 (PDB ID: 5M4M). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

Similar behavior can be observed also for ligand 7FW (PDB ID: 5M4M), as noticeable in Figure 8. In this case, the two native-like poses (poses 1 and 4) outscore the other ones, according to the MS coefficient. In both cases, the ligand's dyhydroxybenzamide moiety is correctly located within a subpocket where it can establish multiple directional interactions with Asp290, Thr54, and Gly294, although in the case of pose 4, the quinoxaline moiety is flipped compared to the reference and pose 1 and interacts with the backbone of Leu330. Interestingly, pose 5 (which occupies the pocket with a similar orientation to poses 1 and 4, but inserts an undecorated phenyl ring within the aforementioned pocket instead of the dyhydroxymethoxybenzamide) has a higher but still reasonable MS score, contrary to poses 2 and 3 where the complementarity of shape and electrostatic properties between the ligand and the pocket are inferior, and most of the interaction between the ligand the binding site occurs in the lower portion, the subpocket, which in the reference pose is occupied by the quinoxaline group.

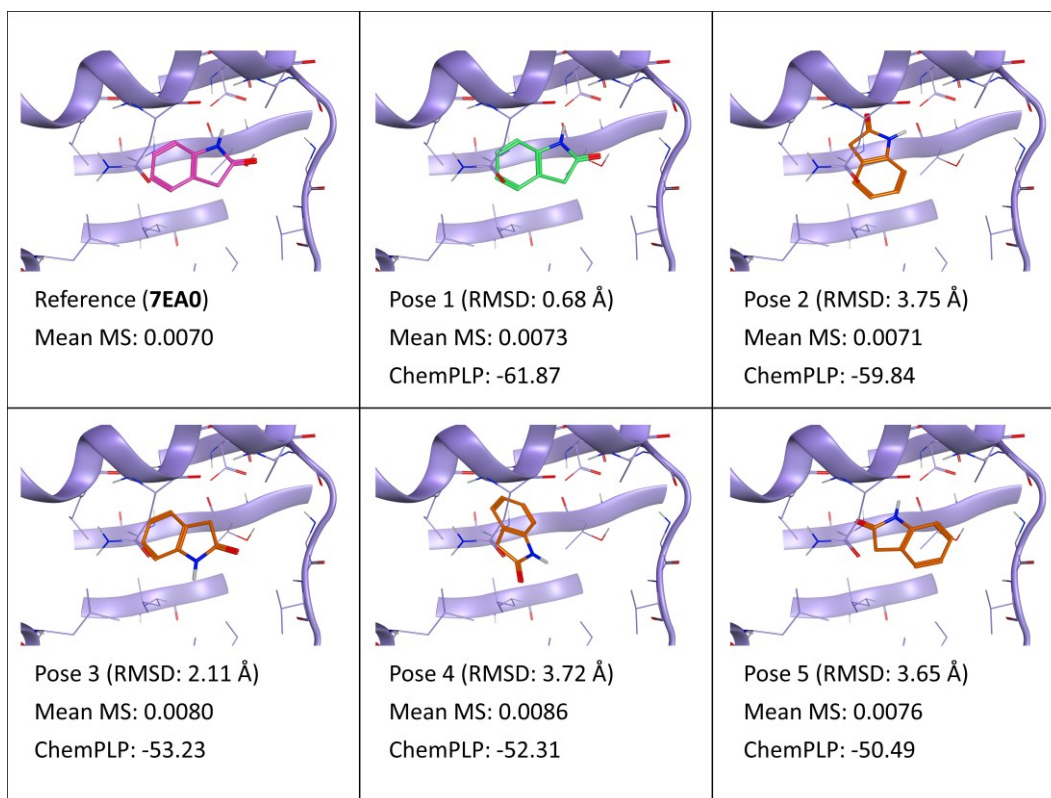


Figure 9. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand W6P within the catalytic site of the pyruvate dehydrogenase kinase 2 (PDB ID: 7EA0). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

Contrary to the first two complexes involving strong binders and coherently with what was observed for complexes 5IH6 and 6HOU, the TTMD refinement was also inconclusive for poses generated for ligand W6P (PDB ID: 7EA0). Once again, the poses were all unstable under the chosen simulation conditions (see also Figures S11 and S12, Supplementary Materials, for reference), leading to high MS values and concealment of the subtle differences between different ligand conformation and interaction patterns. The third observance of this phenomenon is a further reaffirmation that the default temperature ramp is too aggressive for elucidating small differences in binding pose stability such as the ones of weak binders.

SARS-CoV-2 Main Protease

The fourth and final test case involved the SARS-CoV-2 main protease (M^{pro}), an enzyme playing a pivotal role in this virus replication cycle⁴⁶ thanks to its ability to drive the maturation of several nonstructural proteins through the proteolytic cleavage of the pp1a/pp1ab polyproteins⁴⁷. The result of the TTMD refinement carried out on docking poses generated for each M^{pro} protein-ligand complex is summarized in Figures 10–12, while detailed analysis of representative replicates (Figures S13–S16), and per-replicate reports (Tables S11 and S12) can be found in the Supplementary Materials. For reference, a comparison between the representative TTMD trajectory for the best (pose1, MD5) and worst (pose2, MD4) scoring pose for complex 7M8P, according to the MS coefficient, is highlighted in Video S4.

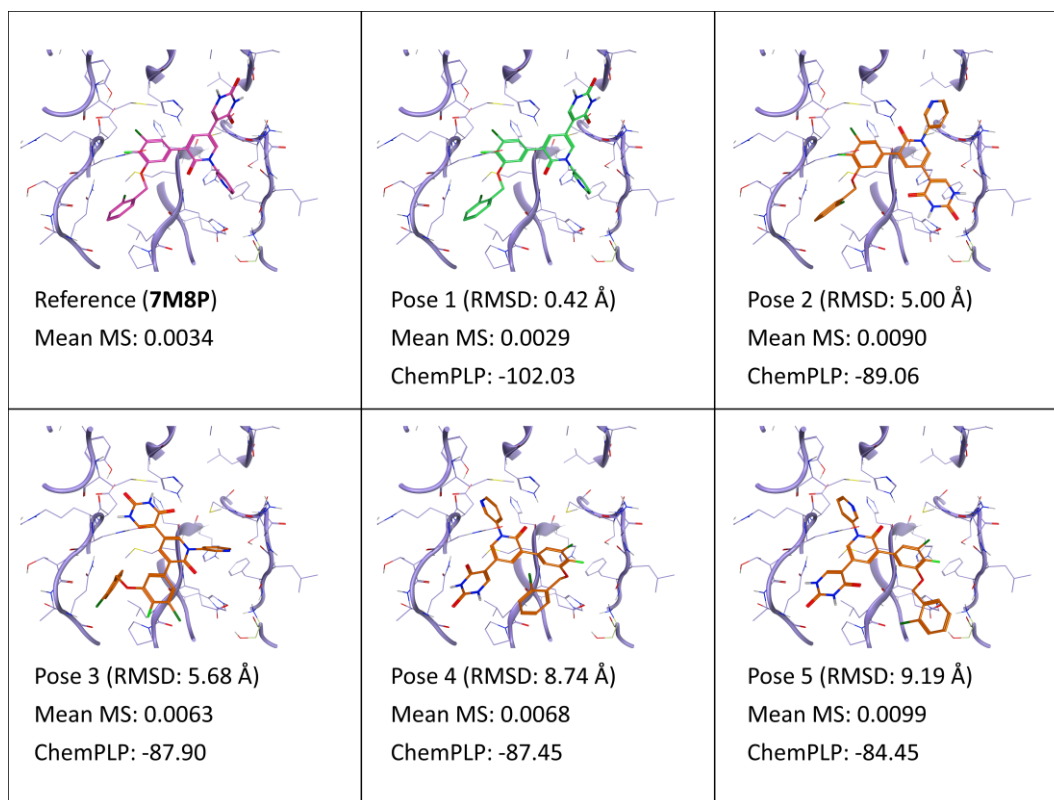


Figure 10. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand YSJ within the catalytic site of the SARS-CoV-2 main protease (PDB ID: 7M8P). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

In the case of ligand YSJ (PDB ID: 7M8P), the MS value determined through the TTMD post-docking refinement indicates pose 1, which is superimposable to the crystal reference, as the most stable pose (see also Video S5, Supplementary Materials). The evident difference between the stability of pose 1 compared to the other ones can be related to the fact that it is the only docking pose that contemporarily satisfies both hydrogen bond interactions with the sidechain of His163 and the backbone of Glu166 (see also Figure S13, Supplementary Materials), which are known to be essential for binding at the catalytic site of the protease⁴⁸. Another possible explanation for the enhanced gap in poses' stability compared to previously discussed cases is that, in the case of kinases, the binding pocket is usually very druggable, implying that there are several possible ligand conformations with reasonable complementarity of shape and electrostatic properties with the cavity, hence with high stability during MD simulations. On the contrary, M^{pro} catalytic site is very shallow and solvent-exposed⁴⁹, implying more constraints on possible ligand orientations that lead to a stable binding. Indeed, short-range nondirectional interactions play a lesser important role in ligand recognition at this site, with more importance being given to a reduced number of high-quality interactions centered upon the S1 subpocket.

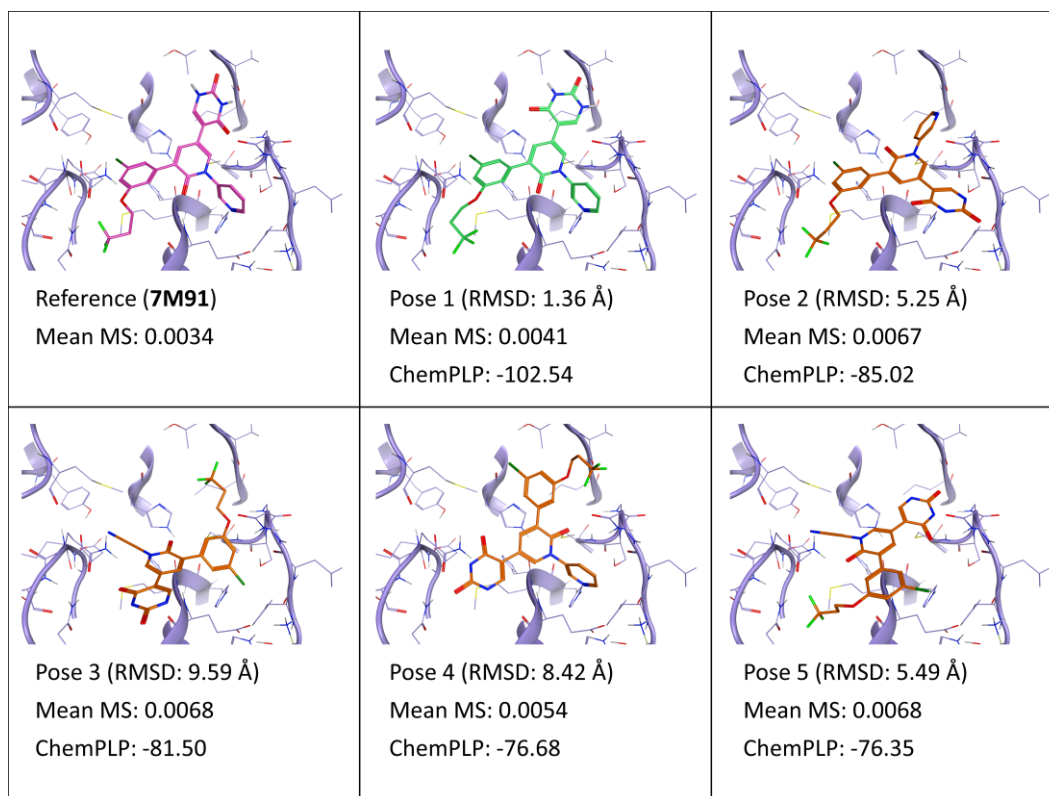


Figure 11. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand YU4 within the catalytic site of the SARS-CoV-2 main protease (PDB ID: 7M91). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

Accordingly, similar results are obtained also for ligand YU4 (PDB ID: 7M91). As depicted in Figure 11, the MS value once again correctly discriminates between the native-like ligand conformation (pose 1) and incorrect binding poses (2–5). Interestingly, pose 4, which presents one of the two conserved hydrogen bonds (the one with His163), is the second most stable pose under simulation conditions, further highlighting the importance of this interaction in anchoring the ligand to the catalytic site. Furthermore, pose 4 is more stable than poses 2, 3, and 5, despite being the only one without a hydrophobic/aromatic ring within the S2 subpocket, another important binding feature of many crystal ligands, including natural peptide substrates⁵⁰.

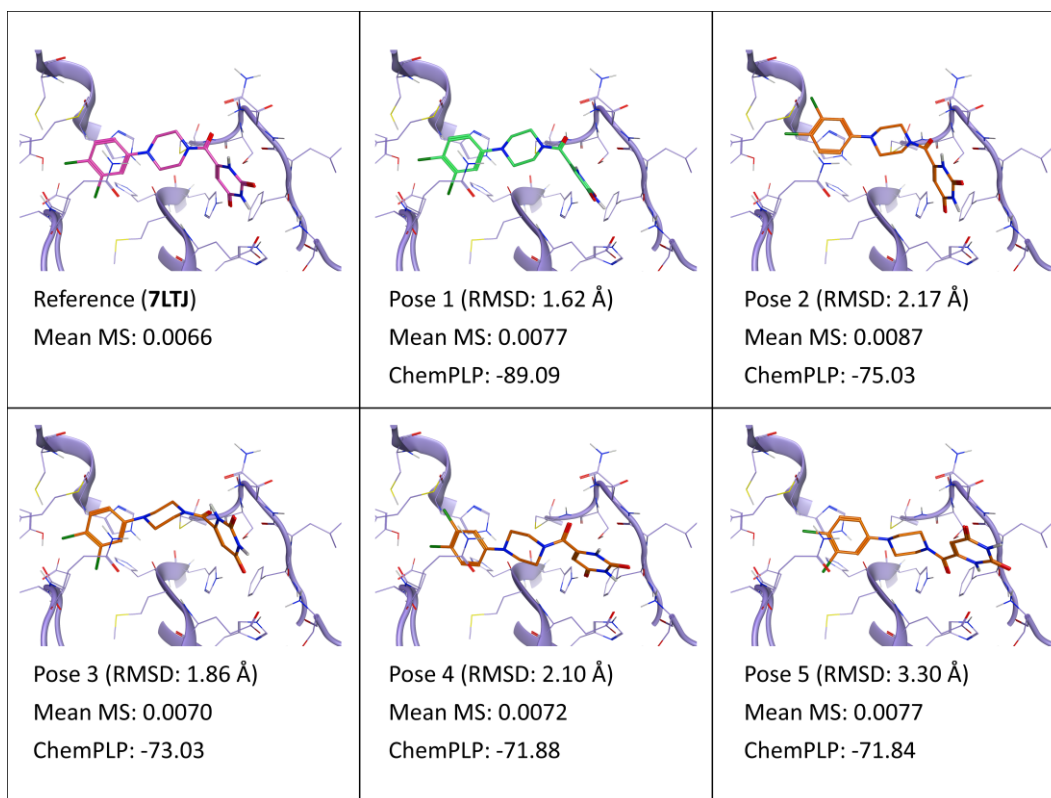


Figure 12. This panel encompasses the results of the TTMD post-docking refinement of poses generated for ligand YD1 within the catalytic site of the SARS-CoV-2 main protease (PDB ID: 7LTJ). Protein residues within 5 Å of the ligand (orange) position are shown (violet). For each pose, the RMSD in comparison with the reference, the average MS coefficient resulting from TTMD simulations, and the ChemPLP docking score are reported. As a reference, the experimentally determined binding mode (magenta) is also reported, as well as the MS coefficient derived from its TTMD refinement, while the best docking pose (according to the RMSD to the reference) is highlighted in green.

Finally, as observed for all previous biological targets, in this case also the TTMD refinement of docking poses generated for the weak binder (ligand YD1, PDB ID: 7LTJ) was inconclusive. The MS values are high and quite similar although expectable due to the similarity among this set of poses and the crystal reference and the low experimentally determined IC₅₀ value.

Discussion

In the present work, we presented the first application of Thermal Titration Molecular Dynamics (TTMD), an alternative enhanced sampling MD-based method for the qualitative estimation of protein-ligand unbinding kinetics, as a post-docking refinement protocol. TTMD, which combines a series of molecular dynamics simulations performed at progressively increasing temperatures with a scoring function based on protein-ligand interaction fingerprints, was used to rank a series of docking poses based on the persistence of the native binding mode. Four different pharmaceutically relevant biological targets, including protein kinase CK1δ, protein kinase CK2, pyruvate dehydrogenase kinase 2, and SARS-CoV-2 main protease were investigated, with three different protein-ligand complexes for each target used in this work. Test cases were chosen among the same protein-ligand complexes used in the original paper²⁵, as well as the same simulation conditions, i.e., the adopted temperature ramp.

Results reported in this manuscript show how TTMD can be readily used as a post-docking refinement tool for mature, drug-like ligands. The results of the TTMD refinement performed on docking poses generated for this class of ligands showed how the MS score was consistently able to identify the native-

like binding pose among the other ones. Furthermore, the MS score demonstrated a good sensitivity in discriminating between poses with partially overlapping interaction features, which could be tricky to evaluate based only on docking calculations.

Most MD-based post-docking refinement protocols involve performing short simulations (tens of nanoseconds) at a fixed temperature to distinguish between stable and unstable poses. The discrimination is usually based on simple geometric descriptors such as the ligand RMSD or the center-of-mass distance between the ligand and the protein. Although this strategy seems legitimate and could work in some cases, based on our experience and as can be observed by the detailed trajectory analyses provided in the Supplementary Materials, these types of simulations are rarely successful for this purpose. First, short windows of classic MD simulations are not sufficient to appreciate significant differences between plausible, near-native poses and the native binding mode, since they probably satisfy a good number of interactions with the target allowing them to occupy the binding site for a sufficient time to be considered stable, even if incorrect. Second, there is no clear cutoff to use for a geometric descriptor to define if the starting binding mode is lost or not: depending on the ligand structure and topological properties of the binding site, values would have to be adjusted in a case-by-case scenario. Moreover, geometric descriptors do not discriminate well if the deviation from the starting point of the simulation regards the whole ligand or only a portion of it: in some cases, as shown in the Results section, some ligand moieties do not interact with the protein active site and are therefore relatively unconstrained. These ligand chunks contribute to increasing the ligand RMSD without causing a loss of interaction with the target, falsely leading to a perception of pose instability.

On the contrary, thanks to its simple but effective enhanced sampling strategy based on increasing the temperature in between each simulation window, TTMD can emphasize stability differences between relatively similar or partially overlapping docking poses in simulation times comparable to those of classical post-docking refinements (see Table S13 for statistics about calculation speed on each investigated system). Moreover, the use of the fingerprint-based scoring function (IFP_{CS}) instead of the ligand RMSD allows for determining a clear cutoff value (zero) that can univocally define whether the loss of interactions found in the starting protein-ligand complex happened. This threshold value is used as an early termination criterion, avoiding spending time simulating the ligand floating around the simulation box thus increasing the feasibility of using such an approach in a high-throughput fashion. Indeed, as can be deduced by Tables S3, S6, S9, and S12 (Supplementary Materials), several simulations end with the total detachment of the ligand from the binding site (considerably high RMSD values). This fact has two main consequences: first, two poses that detach at different times/temperatures would be considered equally unstable if the final/average RMSD would be used as a metric. Second, one would have to simulate for a long time even in those cases where it is not necessary.

Although the TTMD protocol can be useful as is, two questions remain unaddressed still. The first issue with the current protocol is related to the applicability to weak binders, i.e., compounds with affinity values that are above the micromolar threshold, including fragments, which are the starting point of many drug discovery campaigns these days. As reported in the Results section of the manuscript, the current temperature ramp is too aggressive to be applied to weakly binding ligands. Particularly, the aggressiveness of the ramp causes both native-like and wrong poses to rapidly lose their binding features and eventually detach from the binding site. This situation could be addressed by using an alternative ramp, focused on lower temperature ranges. Furthermore, as highlighted by Figure S3 (Supplementary Materials) in the case of complex 5IH6, but also in Video S5 (Supplementary Materials) which regards complex 7M8P, it is possible that poses interconvert between each other during the TTMD simulation, especially for low-affinity compounds such as fragments. To better keep track of this behavior, an alternative analysis protocol could be to analyze the frequency by which a certain pose is sampled across the whole set of replicates for a given ligand, regardless of the starting ligand conformation. For both these latest aspects, work in this direction is currently being carried out in our laboratory and will be the scope of a future paper. The second problem is to make the protocol suitable also for investigating membrane proteins. The question is not trivial, since these types of targets are difficult to simulate accurately in standard conditions and therefore deserve special care when dealing with unusual simulation conditions such as the high temperatures explored with the current protocol. A possible solution, which is currently being evaluated in our laboratory, is to simulate these proteins as soluble, without introducing variables related to the

membrane behavior.

Materials and Methods

Hardware Overview

Most of the molecular modeling tasks, like the preparation of protein and ligand structures, docking calculations, the setup for MD simulations, and MD trajectory analyses, were carried out on a Linux workstation, equipped with a 20 cores Intel Core i9-9820X 3.3 GHz processor, running Ubuntu 20.04 as its operating system. An in-house GPU cluster, composed of 20 NVIDIA drivers ranging from GTX980 to RTX2080Ti, has been exploited for carrying out each MD simulation.

Structure Preparation

The three-dimensional coordinates of each protein-ligand complex used in this study were retrieved from the Protein Data Bank (PDB)⁵¹ and processed through various tools embedded within the Molecular Operating Environment (MOE) 2022.02 suite⁵². In the present work, four different biological targets were considered: casein kinase 1 isoform δ (CK1 δ), casein kinase 2 (CK2), pyruvate dehydrogenase kinase 2 (PDK2), and SARS-CoV-2 main protease (M^{pro}). For each macromolecule, three different protein-ligand complex structures were considered: the accession codes for each structure are summarized in Table 1. Protein-ligand systems were simulated as monomers, aside from SARS-CoV-2 M^{pro}, for which the catalytically competent homodimer form was reconstituted through symmetric crystallographic transformations. Initially, each structure was preprocessed through the “structure preparation” tool, to assign residues with alternative conformations to the highest occupancy one, to rebuild broken loops through homology modeling, and to fix inconsistencies between the aminoacidic sequence and the experimentally solved structure. Afterward, missing hydrogen atoms were added to the system according to the most probable protonation and tautomeric state at pH = 7.4 through the “Protonate3D” tool. Lastly, each solvent, co-solvent, and ion atom was removed before storing the prepared complex for subsequent calculations.

Docking Calculations

For each protein-ligand complex, a self-docking experiment was conducted. Each ligand was prepared using the “Wash” tool of the MOE 2022.02 suite, assigning the most probable protomeric/tautomeric state at pH = 7.4, rebuilding its three-dimensional coordinates, and attributing the correct partial charges, according to the MMFF94 force field⁵³. The binding site was defined using a sphere of radius 10.5 Å centered upon the center of mass of the co-crystallized ligand in the experimentally solved structure. Docking calculations were carried out using the Protein-Ligand ANT System (PLANTS)^{54,55} program, which uses an Ant Colony Optimization (ACO) search algorithm and is free to use for academics. The first five poses according to the ChemPLP⁵⁶ scoring function were stored for further calculations.

System Setup for MD Simulations and Equilibration Protocol

Each protein-ligand complex obtained through docking calculations was prepared for MD simulations through the combination of different packages included in both Visual Molecular Dynamics (VMD⁵⁷) 1.9.2 and the AmberTools22^{58,59} suite. Parameters for protein, water molecules, and ions were attributed according to the ff14SB⁶⁰ force field, while ligands were parametrized utilizing the general Amber force field (GAFF⁶¹). Ligand partial charges were calculated through the AM1-BCC⁶² method. Each protein-ligand system was solvated within a rectangular base prism, ensuring a 15 Å padding value between the box border and the nearest protein atom and using the TIP3P⁶³ model for water molecules. The proper number of monovalent ions, i.e., sodium and chlorine, were then added to neutralize the net charge of the box and to reach a salt concentration of 0.154 M. The prepared system was then subjected to 500 steps of energy minimization with the conjugate-gradient algorithm before undergoing MD simulations, to remove clashes and bad contacts.

Each MD simulation presented in this work was conducted using the ACEMD⁶⁴ 3.5 engine, which is based upon OpenMM⁶⁵, an opensource Python library for molecular simulations. Both in the equilibration and

the production stage, an integration timestep of 2 fs was used, the length of bonds involving hydrogen atoms was constrained through the M-SHAKE algorithm, electrostatic interactions were computed through the particle-mesh Ewald⁶⁶ method, using cubic spline interpolation and a 1 Å grid spacing, while a 9.0 Å cutoff was used for calculating Lennard–Jones interactions.

Before productive MD simulations, each system was equilibrated in a two-step process. In the first equilibration run, a 0.1 ns simulation in the canonical ensemble was carried out, imposing harmonic positional restraints on both the protein and ligand atoms, leaving the water molecules and the ions unconstrained. In the second equilibration run, a 0.5 ns simulation in the isothermal-isobaric ensemble was performed, with harmonic positional restraints applied only on the protein backbone and on the ligand atoms, leaving the protein sidechain unconstrained. In both cases, a force constant of 5 kcal mol^{−1} Å^{−2} was applied for the whole duration of each equilibration run on each restrained atom. For both equilibration stages, the temperature was kept at a constant 300 K value through a Langevin thermostat⁶⁷, while in the second stage, the pressure was maintained at the constant value of 1 atm through a Monte Carlo barostat⁶⁸.

Thermal Titration Molecular Dynamics (TTMD) simulations

As thoroughly described in the work of Pavan et al.²⁵, Thermal Titration Molecular Dynamics (TTMD) is an alternative enhanced sampling MD approach originally developed for the qualitative estimation of protein-ligand unbinding kinetics. The sampling method consists of a series of short classic MD simulations (defined as TTMD-steps), performed at progressively increasing temperature values. The progress of the simulation is monitored by evaluating the conservation of the starting binding mode for a certain protein-ligand complex (a docking pose, in this case) through an interaction fingerprint-based scoring function²⁶, originally implemented in opensource Autogrow4⁶⁹ program, a semi-automatized workflow for de novo drug design. The protocol described afterward is implemented as a Python 3.10 code and relies on the NumPy, MDAnalysis^{70,71}, Open Drug Discovery Toolkit⁷², and Scikit-learn external libraries. The code to run TTMD simulations is available free of charge, under a permissive MIT license, at github.com/molecularmodelingsection/TTMD.

The user must define the “temperature ramp”, i.e., the temperature values that are sampled throughout the simulation and the time spent simulating at each temperature. Different from the original version of the code, in the current one, the length of each simulation window is not fixed, so different temperatures can be sampled for a different amount of time. In the present work, for consistency with the previous paper, the same temperature ramp was used: the starting temperature was set at 300 K, the end temperature was set at 450 K, the temperature increment between each TTMD-step was set at 10 K, while the length of each TTMD-step was 10 ns. The extension of the temperature ramp must be carefully chosen based on prior knowledge of the target, especially concerning the conservation of the native protein fold throughout the simulation (which, in this case, was carried out by monitoring the backbone RMSD). As anticipated, the progress of the simulation is evaluated by a scoring function based on protein-ligand interaction fingerprints. This scoring protocol, named IFP_{CS}, exploits the Scikit-learn Python module to calculate the cosine similarity between two vectors, i.e., the protein-ligand interaction fingerprints of a reference and a query, where the reference is the protein-ligand complex extracted from the last frame of the second equilibration stage, and the query is each protein-ligand complex obtained from each frame of a certain TTMD-step. Specifically, each interaction fingerprint is an integer vector of $r \times 8$ length, where r represents the number of protein residues and 8 represents the possible type of protein-ligand interaction that can be encoded in the vector (hydrophobic contacts, aromatic face-to-face, aromatic edge-to-face, hydrogen bonds with the protein acting as a donor, hydrogen bonds with the protein acting as an acceptor, salt bridge with the protein acting as the positively charged member, salt bridge with the protein acting as the positively negative member, and an ionic bond with a metal ion, respectively). The cosine similarity value is then multiplied by -1 to comply with most scoring functions, where negative values indicate high affinity towards the target and low, near-zero, values indicate lower affinity.

The values of IFP_{CS} can range from -1 , highlighting a total convergence between the reference (the native binding mode) and the query (the binding pose sampled at a certain time during the simulation), to zero, indicating that each interaction feature of the reference has been lost during the simulation.

After the conclusion of each TTMD-step, the average IFP_{CS} calculated across each frame of the step is

calculated: if this value is null, implying that the native binding mode has not been sampled for the whole duration of the step, the TTMD simulation is terminated; otherwise, the simulation proceeds with the next TTMD-step.

Trajectory Analyses and MS Coefficient Determination

TTMD trajectories were analyzed through an in-house Python script. Specifically, the RMSD of the protein backbone and the ligand are calculated at each frame by exploiting the appropriate function provided by the MDAnalysis package. A time-dependent per-residue decomposition of the protein-ligand interaction energy is calculated by exploiting the NAMD Energy plugin 1.4⁷³ for VMD. Afterward, three analysis plots are generated by exploiting the Matplotlib and Seaborn Python packages.

The first plot, defined as the “titration profile”, correlates the average IFP_{CS} value with the temperature of each step: the slope of the straight line linking the initial and the final state of the titration is defined as the MS coefficient and is used as a proxy measure for estimating the “residence time” of the investigated pose.

Equation 1. Mathematical formulation of the MS Coefficient.

$$MS = \frac{meanIFP_{CS}^{T_{end}} - (-1)}{T_{end} - T_{start}}$$

The MS coefficient (Equation 1) ranges from zero (indicative of a tight binding/high persistence of the native binding mode) to 1 (which, on the contrary, indicates low persistence of the starting binding pose). For each investigated protein-ligand complex, five independent TTMD simulations are carried out, with the average MS coefficient being calculated across three different replicates, after discarding the highest and the lowest value.

The second plot reports the time-dependent per-residue decomposition of the protein-ligand interaction energy for the 25 protein residues that are most frequently contacted by the ligand throughout the trajectory (using a distance cutoff of 4.5 Å). The third and final plot reports the time-dependent evolution of the ligand and protein backbone RMSD and the IFP_{CS} score.

Supplementary Materials

The following supporting information can be downloaded at:

<https://www.mdpi.com/article/10.3390/ijms24043596/s1>.

Author Contributions

Conceptualization, M.P. and S.M. (Stefano Moro); methodology, S.M. (Silvia Menin) and M.P.; software, S.M. (Silvia Menin) and M.P.; validation, S.M. (Silvia Menin) and M.P.; formal analysis, S.M. (Silvia Menin) and M.P.; writing—original draft preparation, S.M. (Silvia Menin) and M.P.; writing—review and editing, V.S., M.S. and S.M. (Stefano Moro); supervision, S.M. (Stefano Moro). All authors have read and agreed to the published version of the manuscript.

References

1. Harvey, A. Natural Products in Drug Discovery. *Drug Discov. Today* 2008, 13 (19–20), 894–901. <https://doi.org/10.1016/j.drudis.2008.07.004>.
2. Ban, T. A. The Role of Serendipity in Drug Discovery. *Dialogues Clin. Neurosci.* 2006, 8 (3), 335–344. <https://doi.org/10.31887/DCNS.2006.8.3/tban>.
3. Morphy, R.; Kay, C.; Rankovic, Z. From Magic Bullets to Designed Multiple Ligands. *Drug Discov. Today* 2004, 9 (15), 641–651. [https://doi.org/10.1016/S1359-6446\(04\)03163-0](https://doi.org/10.1016/S1359-6446(04)03163-0).
4. Anderson, A. C. The Process of Structure-Based Drug Design. *Chem. Biol.* 2003, 10 (9), 787–797. <https://doi.org/10.1016/j.chembiol.2003.09.002>.
5. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W. Computational Methods in Drug Discovery. *Pharmacol. Rev.* 2014, 66 (1), 334–395. <https://doi.org/10.1124/pr.112.007336>.

6. Leelananda, S. P.; Lindert, S. Computational Methods in Drug Discovery. *Beilstein J. Org. Chem.* 2016, 12 (1), 2694–2718. <https://doi.org/10.3762/bjoc.12.267>.
7. Kapetanovic, I. M. Computer-Aided Drug Discovery and Development (CADD): In Silico-Chemico-Biological Approach. *Chem. Biol. Interact.* 2008, 171 (2), 165–176. <https://doi.org/10.1016/j.cbi.2006.12.006>.
8. Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput. Aided-Drug Des.* 2012, 7 (2), 146–157. <https://doi.org/10.2174/157340911795677602>.
9. Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of Docking: An Overview of Search Algorithms and a Guide to Scoring Functions. *Proteins Struct. Funct. Bioinforma.* 2002, 47 (4), 409–443. <https://doi.org/10.1002/prot.10115>.
10. Chaput, L.; Mouawad, L. Efficient Conformational Sampling and Weak Scoring in Docking Programs? Strategy of the Wisdom of Crowds. *J. Cheminformatics* 2017, 9 (1), 37. <https://doi.org/10.1186/s13321-017-0227-x>.
11. Chen, Y.-C. Beware of Docking! *Trends Pharmacol. Sci.* 2015, 36 (2), 78–95. <https://doi.org/10.1016/j.tips.2014.12.001>.
12. Salmasso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* 2018, 9 (AUG), 393738. <https://doi.org/10.3389/FPHAR.2018.00923/BIBTEX>.
13. Ferreira, L. G.; Santos, R. N. D.; Oliva, G.; Andricopulo, A. D. Molecular Docking and Structure-Based Drug Design Strategies. *Mol.* 2015 Vol 20 Pages 13384–13421 2015, 20 (7), 13384–13421. <https://doi.org/10.3390/MOLECULES200713384>.
14. Tan, L.; Batista, J.; Bajorath, J. Computational Methodologies for Compound Database Searching That Utilize Experimental Protein–Ligand Interaction Information. *Chem. Biol. Drug Des.* 2010, 76 (3), 191–200. <https://doi.org/10.1111/j.1747-0285.2010.01007.x>.
15. Muthas, D.; Sabnis, Y. A.; Lundborg, M.; Karlén, A. Is It Possible to Increase Hit Rates in Structure-Based Virtual Screening by Pharmacophore Filtering? An Investigation of the Advantages and Pitfalls of Post-Filtering. *J. Mol. Graph. Model.* 2008, 26 (8), 1237–1251. <https://doi.org/10.1016/j.jmgm.2007.11.005>.
16. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discov.* 2015, 10 (5), 449–461. <https://doi.org/10.1517/17460441.2015.1032936>.
17. Tuccinardi, T. What Is the Current Value of MM/PBSA and MM/GBSA Methods in Drug Discovery? *Expert Opin. Drug Discov.* 2021, 16 (11), 1233–1237. <https://doi.org/10.1080/17460441.2021.1942836>.
18. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* 2019, 119 (16), 9478–9508. <https://doi.org/10.1021/acs.chemrev.9b00055>.
19. Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining Docking and Molecular Dynamic Simulations in Drug Design. *Med. Res. Rev.* 2006, 26 (5), 531–568. <https://doi.org/10.1002/med.20067>.
20. Karplus, M.; McCammon, J. A. Molecular Dynamics Simulations of Biomolecules. *Nat. Struct. Biol.* 2002, 9 (9), 646–652. <https://doi.org/10.1038/nsb0902-646>.
21. Durrant, J. D.; McCammon, J. A. Molecular Dynamics Simulations and Drug Discovery. *BMC Biol.* 2011, 9 (1), 71. <https://doi.org/10.1186/1741-7007-9-71>.
22. De Vivo, M.; Masetti, M.; Bottegioni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 2016, 59 (9), 4035–4061. <https://doi.org/10.1021/acs.jmedchem.5b01684>.
23. Proctor, E. A.; Yin, S.; Tropsha, A.; Dokholyan, N. V. Discrete Molecular Dynamics Distinguishes Nativelike Binding Poses from Decoys in Difficult Targets. *Biophys. J.* 2012, 102 (1), 144–151. <https://doi.org/10.1016/j.bpj.2011.11.4008>.
24. Liu, K.; Watanabe, E.; Kokubo, H. Exploring the Stability of Ligand Binding Modes to Proteins by Molecular Dynamics Simulations. *J. Comput. Aided Mol. Des.* 2017, 31 (2), 201–211. <https://doi.org/10.1007/s10822-016-0005-2>.
25. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Qualitative Estimation of Protein–Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2022, 62 (22), 5715–5728. <https://doi.org/10.1021/acs.jcim.2c00995>.
26. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Implementing a Scoring Function Based on Interaction Fingerprint for Autogrow4: Protein Kinase CK1δ as a Case Study. *Front. Mol. Biosci.* 2022, 9, 629. <https://doi.org/10.3389/fmolb.2022.909499>.
27. Long, A.; Zhao, H.; Huang, X. Structural Basis for the Interaction between Casein Kinase 1 Delta and a Potent and Selective Inhibitor. *J. Med. Chem.* 2012, 55 (2), 956–960. <https://doi.org/10.1021/JM201387S>.
28. Ursu, A.; Illich, D. J.; Takemoto, Y.; Porfetye, A. T.; Zhang, M.; Brockmeyer, A.; Janning, P.; Watanabe, N.; Osada, H.; Vetter, I. R.; Ziegler, S.; Schöler, H. R.; Waldmann, H. Epiblastin A Induces Reprogramming of Epiblast Stem Cells into Embryonic Stem Cells by Inhibition of Casein Kinase 1. *Cell Chem. Biol.* 2016, 23 (4), 494–507. <https://doi.org/10.1016/J.CHEMBIOL.2016.02.015>.
29. Halekotte, J.; Witt, L.; Ianes, C.; Krüger, M.; Bührmann, M.; Rauh, D.; Pichlo, C.; Brunstein, E.; Luxenburger, A.; Baumann, U.; Knippschild, U.; Bischof, J.; Peifer, C. Optimized 4,5-Diarylimidazoles as Potent/Selective Inhibitors of Protein Kinase CK1δ and Their Structural Relation to P38α MAPK. *Molecules* 2017, 22 (4), 522. <https://doi.org/10.3390/molecules22040522>.
30. Battistutta, R.; Cozza, G.; Pierre, F.; Papinutto, E.; Lolli, G.; Sarno, S.; O'Brien, S. E.; Siddiqui-Jain, A.; Haddach, M.; Anderes, K.; Ryckman, D. M.; Meggio, F.; Pinna, L. A. Unprecedented Selectivity and Structural Determinants of a New Class of Protein Kinase CK2 Inhibitors in Clinical Trials for the Treatment of Cancer. *Biochemistry* 2011, 50 (39), 8478–8488. <https://doi.org/10.1021/bi2008382>.
31. Cozza, G.; Zonta, F.; Dalle Vedove, A.; Venerando, A.; Dall'Acqua, S.; Battistutta, R.; Ruzzene, M.; Lolli, G. Biochemical and Cellular Mechanism of Protein Kinase CK2 Inhibition by Deceptive Curcumin. *FEBS J.* 2020, 287 (9), 1850–1864. <https://doi.org/10.1111/FEBS.15111>.

32. Moore, J. D.; Staniszevska, A.; Shaw, T.; D'Alessandro, J.; Davis, B.; Surgenor, A.; Baker, L.; Matassova, N.; Murray, J.; Macias, A.; Brough, P.; Wood, M.; Mahon, P. C. VER-246608, a Novel Pan-Isoform ATP Competitive Inhibitor of Pyruvate Dehydrogenase Kinase, Disrupts Warburg Metabolism and Induces Context-Dependent Cytostasis in Cancer Cells. *Oncotarget* 2014, 5 (24), 12862–12876. <https://doi.org/10.18632/oncotarget.2656>.
33. Akaki, T.; Bessho, Y.; Ito, T.; Fujioka, S.; Ubukata, M.; Mori, G.; Yamanaka, K.; Orita, T.; Doi, S.; Iwanaga, T.; Ikegashira, K.; Hantani, Y.; Nakanishi, I.; Adachi, T. Fragment-Based Lead Discovery to Identify Novel Inhibitors That Target the ATP Binding Site of Pyruvate Dehydrogenase Kinases. *Bioorg. Med. Chem.* 2021, 44. <https://doi.org/10.1016/J.BMC.2021.116283>.
34. Brough, P. A.; Baker, L.; Bedford, S.; Brown, K.; Chavda, S.; Chell, V.; D'Alessandro, J.; Davies, N. G. M.; Davis, B.; Le Strat, L.; Macias, A. T.; Maddox, D.; Mahon, P. C.; Massey, A. J.; Matassova, N.; McKenna, S.; Meissner, J. W. G.; Moore, J. D.; Murray, J. B.; Northfield, C. J.; Parry, C.; Parsons, R.; Roughley, S. D.; Shaw, T.; Simmonite, H.; Stokes, S.; Surgenor, A.; Stefaniak, E.; Robertson, A.; Wang, Y.; Webb, P.; Whitehead, N.; Wood, M. Application of Off-Rate Screening in the Identification of Novel Pan-Isoform Inhibitors of Pyruvate Dehydrogenase Kinase. *J. Med. Chem.* 2017, 60 (6), 2271–2286. <https://doi.org/10.1021/ACS.JMEDCHEM.6B01478>.
35. Clyde, A.; Galanie, S.; Kneller, D. W.; Ma, H.; Babuji, Y.; Blaiszik, B.; Brace, A.; Brettin, T.; Chard, K.; Chard, R.; Coates, L.; Foster, I.; Hauner, D.; Kertesz, V.; Kumar, N.; Lee, H.; Li, Z.; Merzky, A.; Schmidt, J. G.; Tan, L.; Titov, M.; Trifan, A.; Turilli, M.; Van Dam, H.; Chennubhotla, S. C.; Jha, S.; Kovalevsky, A.; Ramanathan, A.; Head, M. S.; Stevens, R. High-Throughput Virtual Screening and Validation of a SARS-CoV-2 Main Protease Noncovalent Inhibitor. *J. Chem. Inf. Model.* 2022, 62 (1), 116–128. <https://doi.org/10.1021/acs.jcim.1c00851>.
36. Deshmukh, M. G.; Ippolito, J. A.; Zhang, C.-H.; Stone, E. A.; Reilly, R. A.; Miller, S. J.; Jorgensen, W. L.; Anderson, K. S. Structure-Guided Design of a Perampanel-Derived Pharmacophore Targeting the SARS-CoV-2 Main Protease. *Structure* 2021, 29 (8), 823–833.e5. <https://doi.org/10.1016/j.str.2021.06.002>.
37. Knippschild, U.; Gocht, A.; Wolff, S.; Huber, N.; Löhler, J.; Stöter, M. The Casein Kinase 1 Family: Participation in Multiple Cellular Processes in Eukaryotes. *Cell. Signal.* 2005, 17 (6), 675–689. <https://doi.org/10.1016/j.cellsig.2004.12.011>.
38. Perez, D. I.; Gil, C.; Martinez, A. Protein Kinases CK1 and CK2 as New Targets for Neurodegenerative Diseases. <https://doi.org/10.1002/med.20207>.
39. Xu, P.; Ianes, C.; Gärtner, F.; Liu, C.; Burster, T.; Bakulev, V.; Rachidi, N.; Knippschild, U.; Bischof, J. Structure, Regulation, and (Patho-)Physiological Functions of the Stress-Induced Protein Kinase CK1 Delta (CSNK1D). *Gene* 2019, 715, 144005. <https://doi.org/10.1016/j.gene.2019.144005>.
40. Venerando, A.; Ruzzene, M.; Pinna, L. A. Casein Kinase: The Triple Meaning of a Misnomer. *Biochem. J.* 2014, 460 (2), 141–156. <https://doi.org/10.1042/BJ20140178>.
41. Meggio, F.; Pinna, L. A. One-Thousand-and-One Substrates of Protein Kinase CK2? *FASEB J.* 2003, 17 (3), 349–368. <https://doi.org/10.1096/fj.02-0473rev>.
42. Borgo, C.; D'Amore, C.; Sarno, S.; Salvi, M.; Ruzzene, M. Protein Kinase CK2: A Potential Therapeutic Target for Diverse Human Diseases. *Signal Transduct. Target. Ther.* 2021, 6 (1), 183. <https://doi.org/10.1038/s41392-021-00567-7>.
43. Zhang, S.; Hulver, M. W.; McMillan, R. P.; Cline, M. A.; Gilbert, E. R. The Pivotal Role of Pyruvate Dehydrogenase Kinases in Metabolic Flexibility. *Nutr. Metab.* 2014, 11 (1), 10. <https://doi.org/10.1186/1743-7075-11-10>.
44. Patel, M. S.; Korotchkina, L. G. Regulation of the Pyruvate Dehydrogenase Complex. *Biochem. Soc. Trans.* 2006, 34 (2), 217–222. <https://doi.org/10.1042/BST0340217>.
45. Atas, E.; Oberhuber, M.; Kenner, L. The Implications of PDK1–4 on Tumor Energy Metabolism, Aggressiveness and Therapy Resistance. *Front. Oncol.* 2020, 10. <https://doi.org/10.3389/fonc.2020.583217>.
46. Ullrich, S.; Nitsche, C. The SARS-CoV-2 Main Protease as Drug Target. *Bioorg. Med. Chem. Lett.* 2020, 30 (17), 127377. <https://doi.org/10.1016/j.bmcl.2020.127377>.
47. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature* 2020, 582 (7811), 289–293. <https://doi.org/10.1038/s41586-020-2223-y>.
48. Shitrit, A.; Zaidman, D.; Kalid, O.; Bloch, I.; Doron, D.; Yarnitzky, T.; Buch, I.; Segev, I.; Ben-Zeev, E.; Segev, E.; Kobiler, O. Conserved Interactions Required for Inhibition of the Main Protease of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). *Sci. Rep.* 2020, 10 (1), 20808. <https://doi.org/10.1038/s41598-020-77794-5>.
49. Fornasier, E.; Macchia, M. L.; Giachin, G.; Susic, A.; Pavan, M.; Sturlese, M.; Salata, C.; Moro, S.; Gatto, B.; Bellanda, M.; Battistutta, R. A New Inactive Conformation of SARS-CoV-2 Main Protease. *Acta Crystallogr. Sect. Struct. Biol.* 2022, 78 (3), 363–378. <https://doi.org/10.1107/S2059798322000948>.
50. Dai, W.; Zhang, B.; Jiang, X.-M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F.; Li, C.; Li, Y.; Bai, F.; Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L.-K.; Xu, Y.; Yang, H.; Liu, H. Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease. *Science* 2020, 368 (6497), 1331–1335. <https://doi.org/10.1126/science.abb4489>.
51. Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* 2000, 28 (1), 235–242. <https://doi.org/10.1093/nar/28.1.235>.
52. Molecular Operating Environment (MOE), 2022.02, C. G. Molecular Operating Environment (MOE), 2022.02. https://www.chemcomp.com/Research-Citing_MOE.htm.
53. Halgren, T. A. Force Fields: MMFF 94. In *Encyclopedia of Computational Chemistry*; Von Ragué Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., Schreiner, P. R., Eds.; Wiley, 1998. <https://doi.org/10.1002/0470845015.cma012m>.

54. Korb, O.; Stützle, T.; Exner, T. E. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design. *Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.* 2006, 4150 LNCS, 247–258. https://doi.org/10.1007/11839088_22/COVER.
55. Korb, O.; Stützle, T.; Exner, T. E. An Ant Colony Optimization Approach to Flexible Protein–Ligand Docking. *Swarm Intell.* 2007, 1 (2), 115–134. <https://doi.org/10.1007/s11721-007-0006-9>.
56. Korb, O.; Stützle, T.; Exner, T. E. Empirical Scoring Functions for Advanced Protein–Ligand Docking with PLANTS. *J. Chem. Inf. Model.* 2009, 49 (1), 84–96. <https://doi.org/10.1021/ci800298z>.
57. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* 1996, 14 (1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
58. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* 2005, 26 (16), 1668–1688. <https://doi.org/10.1002/JCC.20290>.
59. Case, D. A. AmberTools22, 2022.
60. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 2015, 11 (8), 3696–3713. https://doi.org/10.1021/ACS.JCTC.5B00255/SUPPL_FILE/CT5B00255_SI_001.PDF.
61. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 2004, 25 (9), 1157–1174. <https://doi.org/10.1002/JCC.20035>.
62. Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* 2002, 23 (16), 1623–1641. <https://doi.org/10.1002/jcc.10128>.
63. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 1983, 79 (2), 926–935. <https://doi.org/10.1063/1.445869>.
64. Harvey, M. J.; Giupponi, G.; Fabritiis, G. D. ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* 2009, 5 (6), 1632–1639. https://doi.org/10.1021/CT9000685/ASSET/IMAGES/LARGE/CT-2009-000685_0004.JPEG.
65. Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLOS Comput. Biol.* 2017, 13 (7), e1005659. <https://doi.org/10.1371/journal.pcbi.1005659>.
66. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 1995, 103 (19), 8577–8593. <https://doi.org/10.1063/1.470117>.
67. Davidchack, R. L.; Handel, R.; Tretyakov, M. V. Langevin Thermostat for Rigid Body Dynamics. *J. Chem. Phys.* 2009, 130 (23). <https://doi.org/10.1063/1.3149788/925280>.
68. Faller, R.; Pablo, J. J. D.; Chem, J.; Phys, (. Constant Pressure Hybrid Molecular Dynamics–Monte Carlo Simulations. *J. Chem. Phys.* 2002, 116 (1), 55–59. <https://doi.org/10.1063/1.1420460>.
69. Spiegel, J. O.; Durrant, J. D. AutoGrow4: An Open-Source Genetic Algorithm for de Novo Drug Design and Lead Optimization. *J. Cheminformatics* 2020, 12 (1), 25. <https://doi.org/10.1186/s13321-020-00429-4>.
70. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* 2011, 32 (10), 2319–2327. <https://doi.org/10.1002/jcc.21787>.
71. Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domański, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python Sci. Conf.* 2016, 98–105. <https://doi.org/10.25080/Majora-629e541a-00e>.
72. Wójcikowski, M.; Zielenkiewicz, P.; Siedlecki, P. Open Drug Discovery Toolkit (ODDT): A New Open-Source Player in the Drug Discovery Field. *J. Cheminformatics* 2015, 7 (1), 26. <https://doi.org/10.1186/S13321-015-0078-2>.
73. Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* 2020, 153 (4), 44130. https://doi.org/10.1063/5.0014475/16709547/044130_1_ACCEPTED_MANUSCRIPT.PDF.

A Comprehensive Study Of SARS-Cov-2 Main Protease (M^{pro}) Inhibitor-Resistant Mutants Selected In A VSV-Based System

Francesco Costacurta¹, Andrea Dodaro², David Bante¹, Helge Schöppe³, Ju-Yi Peng⁴, Bernhard Sprenger⁵, Xi He⁴, Seyed Arad Moghadasi⁶, Lisa Maria Egger⁷, Jakob Fleischmann^{8,9}, Matteo Pavan², Davide Bassani², Silvia Menin², Stefanie Rauch¹, Laura Krismer¹, Anna Sauerwein¹, Anne Heberle¹, Toni Rabensteiner¹, Jose Ho¹⁰, Reuben S. Harris^{11,12}, Eduard Stefan^{8,9}, Rainer Schneider⁵, Theresia Dunzendorfer-Matt⁷, Andreas Naschberger¹³, Dai Wang⁴, Teresa Kasere³, Stefano Moro², Dorothee von Laer¹, Emmanuel Heilmann^{1,13*}

Costacurta, F.; Dodaro, A.; Bante, D.; Schöppe, H.; Peng, J.-Y.; Sprenger, B.; He, X.; Moghadasi, S. A.; Egger, L. M.; Fleischmann, J.; Pavan, M.; Bassani, D.; Menin, S.; Rauch, S.; Krismer, L.; Sauerwein, A.; Heberle, A.; Rabensteiner, T.; Ho, J.; Harris, R. S.; Stefan, E.; Schneider, R.; Dunzendorfer-Matt, T.; Naschberger, A.; Wang, D.; Kaserer, T.; Moro, S.; Laer, D. von; Heilmann, E. A Comprehensive Study of SARS-CoV-2 Main Protease (M^{pro}) Inhibitor-Resistant Mutants Selected in a VSV-Based System. *PLOS Pathogens* 2024, 20 (9), e1012522. <https://doi.org/10.1371/journal.ppat.1012522>.

Abstract

Nirmatrelvir was the first protease inhibitor specifically developed against the SARS-CoV-2 main protease (3CL^{pro}/M^{pro}) and licensed for clinical use. As SARS-CoV-2 continues to spread, variants resistant to nirmatrelvir and other currently available treatments are likely to arise. This study aimed to identify and characterize mutations that confer resistance to nirmatrelvir. To safely generate M^{pro} resistance mutations, we passaged a previously developed, chimeric vesicular stomatitis virus (VSV-M^{pro}) with increasing, yet suboptimal concentrations of nirmatrelvir. Using Wuhan-1 and Omicron M^{pro} variants, we selected a large set of mutants. Some mutations are frequently present in GISAID, suggesting their relevance in SARS-CoV-2. The resistance phenotype of a subset of mutations was characterized against clinically available protease inhibitors (nirmatrelvir and ensitrelvir) with cell-based, biochemical and SARS-CoV-2 replicon assays. Moreover, we showed the putative molecular mechanism of resistance based on in silico molecular modelling. These findings have implications on the development of future generation M^{pro} inhibitors, will help to understand SARS-CoV-2 protease inhibitor resistance mechanisms and show the relevance of specific mutations, thereby informing treatment decisions.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of COVID-19 (coronavirus disease-19), has established itself as a permanent human and animal pathogen worldwide. Vaccines, alongside monoclonal antibodies, have drastically reduced hospitalization and/or mortality, especially in immunocompromised individuals, the elderly, and people with pre-existing medical conditions (1). As vaccines do not confer complete immunity against infection, the virus continues to spread

effectively due to easier host-to-host transmission (2), and immune-escaping variants, such as Beta, Delta and Omicron (3–5).

In November 2021, the FDA granted emergency use authorization to Paxlovid, an orally administered medication that combines nirmatrelvir (6), the active component, and ritonavir, a pharmacokinetic enhancer. Nirmatrelvir is a highly potent protease inhibitor (PI) against the SARS-CoV-2 3CL^{pro}/M^{pro}, and is most relevant in the clinical setting, as it has been shown to decrease hospitalization, in-hospital disease progression and death (7). One year after the licensing of Paxlovid, another protease inhibitor, ensitrelvir, was approved in Japan through the emergency regulatory approval system under the commercial name Xocova (8), and recently received Fast Track designation by the FDA. Recently, leritrelvir (9) has been approved by the China National Medical Products Administration. Other oral M^{pro} inhibitors that have entered clinical trials include bofutrelvir (FB2001) (10), pomotrelvir (PBI-0451 (11), halted (12)), simnotrelvir (13), EDP-235 (14), and HS-10517 (GDDI-4405). SARS-CoV-2 continues to spread, and it is expected that the use of protease inhibitors could eventually lead to the selection of drug-resistant M^{pro} variants, with serious consequences for individuals who cannot benefit from vaccines, for example due to immune defects or pre-existing comorbidities.

To date, several studies have addressed the issue of nirmatrelvir-resistant/escaping variants using different approaches: highlighting mutational hotspots (15, 16), in silico investigation of specific mutations (17), studying resistance phenotypes (18–21), addressing which mutations are the most prone to decrease inhibitor susceptibility (22–24) or more comprehensive studies describing either mutation resistance, fitness costs, or both (25–28). Notably, a few groups have performed gain-of-function selection experiments using the Wuhan-1 strain of SARS-CoV-2. While this represents the most straightforward system to generate and study drug-resistant mutants, it requires government approval, BSL-3 facilities, and demands absolute caution to avoid biosafety breaches and potential release of these protease inhibitor-resistant variants.

Furthermore, owing to safety concerns, most of the studies that employed live SARS-CoV-2 were not performed with the Omicron variant (18, 19, 21). This precaution was taken, because protease resistant pre-Omicron variants would likely be unable to compete with the current Omicron strains, if released accidentally. In a preprint by Lan et al. (29), the authors validated the antiviral resistance of certain M^{pro} resistant variants in the Omicron-M^{pro} context by using SARS-CoV-2 replicons, circumventing potential safety issues for mutated transmissible variants. It is worth noting that the Omicron signature mutation (P132H) does not confer resistance to common inhibitors (30–33) but alters the thermal stability of the protease, as reported by Sacco et al. (34). This alteration, as well as other structural and biochemical features caused by the P132H substitution, could potentially affect the mechanism of resistance development and the relevance of mutations selected with Wuhan-1 M^{pro} for Omicron.

Recently, we described a BSL-2 mutation selection system based on a M^{pro}-dependent chimeric vesicular stomatitis virus (VSV-3CL^{pro}/VSV-M^{pro}) (35). Replication of the chimeric VSV was effectively inhibited by nirmatrelvir. Using this system, we were able to select a panel of mutants and characterize them through computational and cellular methods. In the present study, we deepened our understanding of the SARS-CoV-2 M^{pro} resistance mutation landscape by selecting mutations in the Omicron M^{pro} and increasing the resistance phenotypes of an already resistant variant, namely VSV-L167F-M^{pro}. By generating and characterizing the mechanism of Omicron-M^{pro} and L167F-derived mutations, individually or in combination, this study aims to advance the understanding of protease-inhibitor-resistant M^{pro} variants and provide valuable information for treatment decision-making.

Results

A safe method to select protease inhibitor resistant M^{pro} variants

In a previous study, we developed a safe alternative to using un-attenuated SARS-CoV-2 for the selection of inhibitor-resistant variants (18, 19, 35). The technology is based on a chimeric VSV-M^{pro} that encodes an artificial, non-functional polyprotein (G-M^{pro}-L), and relies on M^{pro} activity for its replication (Fig. 1A,B). In the absence of an inhibitor, the protease processes the polyprotein, releasing G and L, and the virus can replicate. However, when a protease inhibitor is applied, the protease is inhibited, and the virus can no longer replicate.

Wuhan-1 and Omicron VSV-M^{pro} variants are equally susceptible to nirmatrelvir

To gain a deeper understanding of the potential of the Omicron-M^{pro} to acquire resistance mutations differently than the Wuhan-1 M^{pro}, we introduced the Omicron-M^{pro} signature mutation P132H into chimeric VSV-M^{pro}, generating VSV-Omicron-M^{pro} (VSV-O-M^{pro}) for subsequent selection experiments with the protease inhibitor nirmatrelvir. First, we demonstrated that nirmatrelvir is equally effective against both the wild-type and Omicron M^{pro} also in our system, according to literature^{23–26}. To that end, we performed a nirmatrelvir dose response experiment and found comparable sensitivities to the inhibitor (Figure S1).

VSV-M^{pro} supports the main protease evolution of Omicron variant and variants with existing resistance to escape nirmatrelvir

Following-up on our first study (35), a resistant variant of VSV-M^{pro} (VSV-L167F-M^{pro}), which was previously identified in selection experiments, was used. This mutation was of particular interest, because it was also found in resistance studies with authentic SARS-CoV-2 (18, 19, 21). Then, selection experiments using both VSV-O-M^{pro} and VSV-L167F-M^{pro} were performed (Fig. S2A). By passaging both VSV-L167F-M^{pro} and VSV-O-M^{pro} in the presence of suboptimal concentrations of nirmatrelvir, M^{pro} mutations that were generated by the error-prone VSV polymerase were selected (36, 37). Samples from the first two pilot experiments with VSV-L167F-M^{pro} and VSV-O-M^{pro} were sequenced via Sanger sequencing. Samples from subsequent selection experiments were deep-sequenced (Nanopore) (Fig. 1C) on the same target region (G_{Cterm}-M^{pro}-L_{Nterm}) (Table S1). We obtained 29 distinct non-synonymous mutations in double-, triple-, and quadruple mutated VSV-L167F-M^{pro} variants and 47 distinct non-synonymous single- and double-mutated VSV-O-M^{pro} variants, as schematically shown by the unrooted phylogenetic tree (Fig. 1D).

To achieve multiple-mutated viruses, several interesting mutants were selected (Fig. 2A) according to specific parameters for further selection experiments, where we increased the concentration of the inhibitor, imposing stronger selection pressure at each passage. An overview on the M^{pro}-genome location of the generated mutations is displayed in Fig. S2B. Several variants were chosen for additional passaging in the presence of nirmatrelvir according to following criteria. First, the proximity of the amino acid substitution to the inhibitor (within 5 Å of the catalytic site or near the catalytic site, within 5 – 10 Å). Second, the prevalence in the GISAID database of specific mutations (38–40). Especially for further testing and modelling we considered substitutions whose count in the GISAID database was above 500 entries. Third, the frequency of a specific mutation occurring in different samples independently.

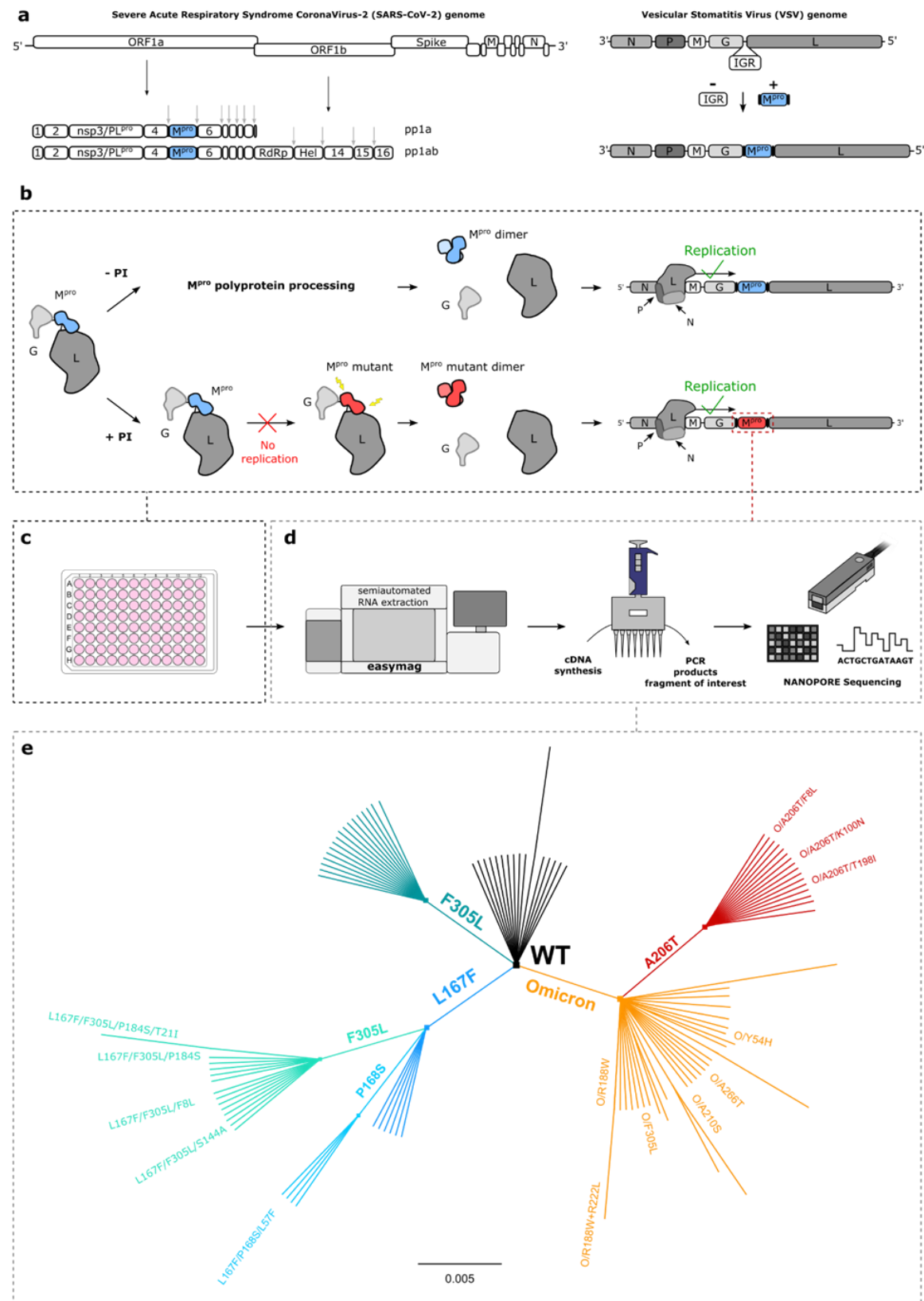


Figure 1. VSV-G-M^{pro}-L construct: molecular mechanism, sequencing workflow, and mutant lineage phylogeny. (A) Schematic representation of SARS-CoV-2 genome, polyproteins pp1a and pp1ab and the VSV-based mutation selection tool. The InterGenic Region (IGR) between proteins G and L of WT VSV was replaced with the SARS-CoV-2 M^{pro} Wuhan-1 sequence and its cognate autocleavage

sites. M^{pro} genome positions in SARS-CoV-2 and in VSV-M^{pro} are highlighted in light blue. (B) The virus is fully dependent on M^{pro} for replication. Upon translation of G-M^{pro}-L, two outcomes are possible: 1. without an inhibitor, M^{pro} is free to process the polyprotein, and the transcription and replication complexes can assemble; 2. with an inhibitor, M^{pro} is inhibited, the polyprotein is not processed, and the virus is thus not able to replicate, unless it acquires a mutation rendering the M^{pro} less susceptible to the inhibitor. Then, the virus can replicate despite the inhibitor. (C) Selection experiments workflow: cells are infected with VSV-M^{pro} and treated with a protease inhibitor (nirmatrelvir), supernatants of cytopathic effect (CPE) positive wells are used to isolate viral RNA, synthesize cDNA, PCR amplify the region of interest (M^{pro}) that will be sequenced using Nanopore sequencing. (D) Unrooted phylogenetic tree showing the relationship between original/parental viruses and mutants. M^{pro} variants belonging to the same parental virus are coloured accordingly: WT (black), F305L (sea green), L167F (blue), L167F/F305L (light sea green), L167F/P168S (light blue), Omicron (orange), Omicron/A206T (red). For clarity, only the names of the mutants investigated in this work are displayed. *Previously generated/investigated set of mutants in our first study (35).

Some mutations combined more than one criterion. For example, some substitutions located in the catalytic site were also frequent in the GISAID repository. However, residues located in this area were generally more conserved compared to other coronavirus main proteases (Fig. S3) (41, 42), whereas residues outside the catalytic site were more frequent in our selection experiments and more frequent in GISAID. We found ten substitutions with more than 500 entries: T21I, P184S, K100N, T198I, F8L, A234T, A194S, A210S, A266T, and P168S (Fig. 2B) and broke them down according to their appearance in previous variants of concern (Table S2).

Despite not reaching 500 entries in GISAID, we also included the F305L mutation for further study. We recently described F305L as an autocleavage site optimization mutant with similarities to T304I selected in authentic SARS-CoV-2 resistance studies. All mutations occurred during selection experiments, except for T21I and R222L. R222L arose during plaque purification of VSV-Omicron-R188W-M^{pro}. T21I appeared during the generation of VSV-L167F/F305L/P184S from its plasmid (also referred to as “rescue” (43)). This virus variant had to be produced from its plasmid, as it could not be plaque purified.

Based on the abovementioned parameters (proximity, GISAID frequency and/or independent occurrences), we chose two VSV-L167F-M^{pro} variants, L167F/F305L and L167F/P168S for further selection experiments. From VSV-O-M^{pro}, the variants O/A206T, O/R188W, O/Y54H, O/A210S, and O/A266T were chosen for further study. Among these variants, we selected O/A206T for additional selection experiments, as it occurred 17 times independently in VSV-O-M^{pro} selection experiments (Fig. 2C). To confirm that this variant was not present in our VSV-Omicron-M^{pro} stocks already, Nanopore sequencing was performed on them. We confirmed the absence of any VSV-O-M^{pro} variant subpopulation to the extent of sensitivity that Nanopore sequencing provides. Subsequently, the next round of selection experiments with the variants L167F/F305L and L167F/P168S was carried out, in which we selected the triple-mutant variants L167F/F305L/F8L, L167F/F305L/S144A, L167F/F305L/P184S, and L167F/P168S/L57F. From VSV-O/A206T-M^{pro}, the mutants O/A206T/F8L, O/A206T/K100N and O/A206T/T198I were selected. To assess whether mutants generated from the selection experiments and selected for further analyses were not detrimental for viral propagation, the Ultrafast Sample placement on Existing tRee (USHER) tool was applied. Using USHER, we generated pedigrees of the most frequently represented mutants in the GISAID database, T21I, T198I, P184S, K100N, and F8L (Fig. S4A,B). The prevalence of these mutations before and after the Omicron surge (21st December 2021, up to 18th January 2023) (Fig. S4C) was also examined. From this analysis we observed that M^{pro} variants bearing these substitutions appear to not be harmful to the spread of the virus.

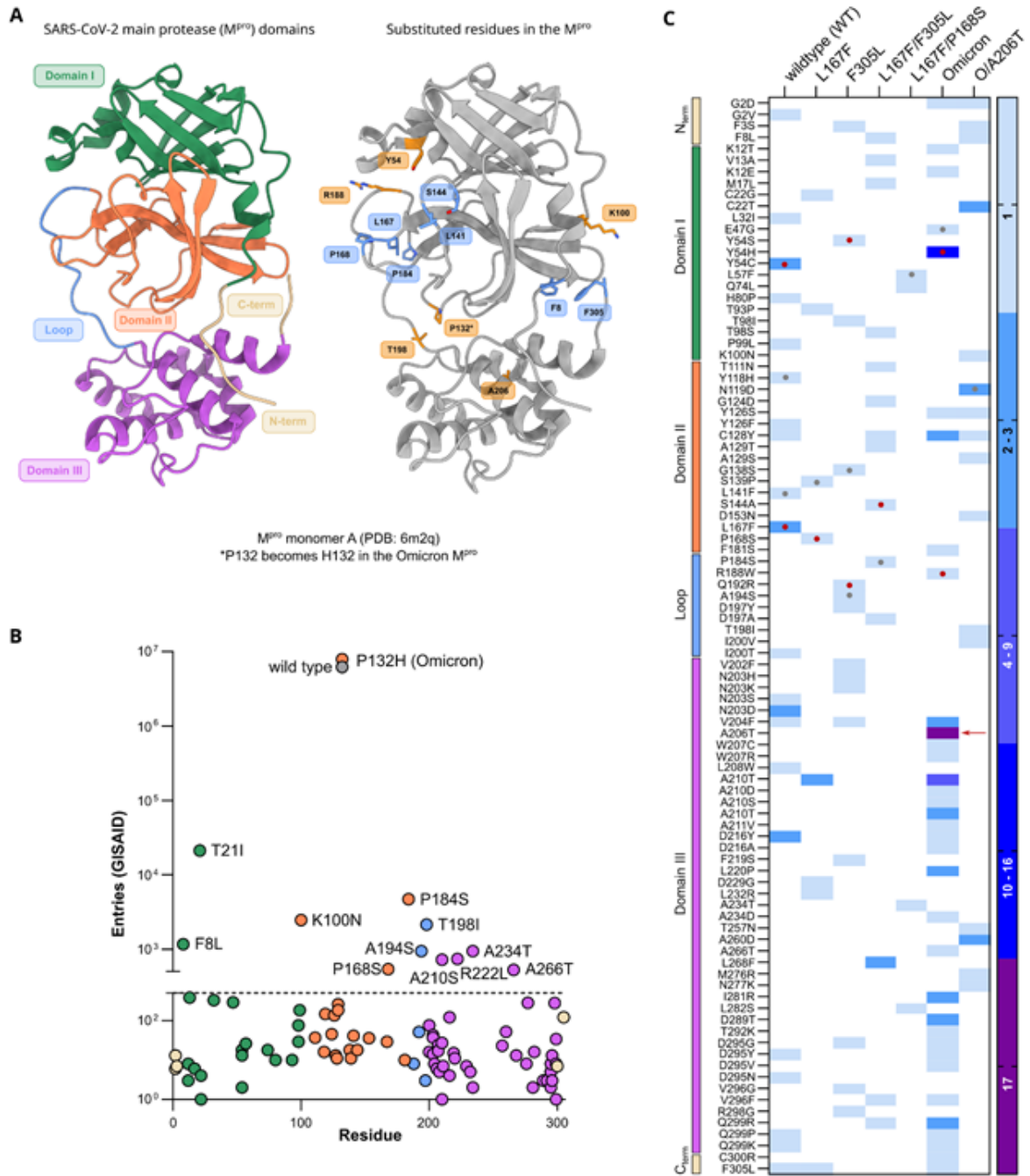


Figure 2. M^{pro} structure, Frequency of M^{pro} mutations in the GISAID database and generated mutations. (A) 3D structure of the M^{pro} monomer A (PDB: 6M2Q). On the left side of the panel, M^{pro} is sub-divided into domains. On the right side of the panel, the same M^{pro} structure is shown highlighting residues that were found to be mutated after selection experiments: residues found mutated in the WT background are coloured blue; mutants in the Omicron background orange. (B) The total number of $nsp5/M^{pro}$ GISAID substitution entries for each mutation found after selection experiments. Entries are coloured according to domains as displayed in panel A. WT and Omicron- M^{pro} sequences are displayed in grey and orange, respectively. The dotted line at 500 represent the cut-off value. (C) Heat map representing the number of a specific substitution's independent occurrences from different samples of VSV- M^{pro} selection experiments. Color-coded domains are indicated as in A. The red arrow indicates the substitution A206T. Red dots indicate residues within the catalytic site and grey dots indicate residues near the catalytic site.

Mutations selected in VSV-G-Mpro-L confer resistance to nirmatrelvir & ensitrelvir

To quantify the resistance of the selected variants with VSV-M^{pro} (Fig. S5A), the most interesting mutations were introduced into two previously described live-cell-based protease activity assays, namely 3CL/M^{pro}-On and 3CL/M^{pro}-Off (44) (Fig. S5B,C). These two protease activity measurement tools rely on replication-incompetent VSV-dsRed variants missing either the phosphoprotein (P) or the polymerase (L), which were replaced with the red fluorescent protein dsRed. Briefly, cells were transfected either with a plasmid encoding the P protein modified with an INTRAmolecular-M^{pro}-tag (P_{Nterm}:M^{pro}:P_{Cterm}) or an INTERmolecular fusion protein made of the green fluorescent protein, M^{pro} and L (GFP-M^{pro}-L). The cells were then infected with VSV-ΔP or VSV-ΔL, respectively. The P_{Nterm}:M^{pro}:P_{Cterm} intramolecular tag in combination with VSV-ΔP-dsRed constitutes a gain-of-signal assay also called “M^{pro}-On”, whereas the artificial polyprotein (or fusion protein) GFP-M^{pro}-L in combination with VSV-ΔL-dsRed constitutes a loss-of-signal assay also called “M^{pro}-Off”.

First, the M^{pro}-On system was used to quantify the resistance phenotype of four triple mutants that arose from the WT protease: L167F/F305L/F8L, L167F/P168S/L57F, L167F/F305L/P184S, and L167F/F305L/S144A, against nirmatrelvir and ensitrelvir.

We also included the related single and double mutants to investigate the contribution of each mutation to the resistance phenotype. Single substitutions conferred low resistance to protease inhibitors, but the resistance increased by combining them. Mutations that arose during selection experiments with the Omicron protease (P132H) were also first characterized with this gain-of-signal system. In Fig. S6, non-linear regression analyses of M^{pro}-On dose-response experiments are shown. However, there were limitations in quantifying the resistance of some WT-related combinations using the gain-of-signal assay. This was largely observed in multiple mutant variants, where the gain in signal was completely suppressed by the high resistance of these protease variants and IC₅₀ values could not be quantified. Since inhibitor concentrations of 100 μM or higher were cytotoxic, thereby interfering with signal generation, we did not increase the inhibitor concentration further.

Instead, to attain better resistance quantification, we cloned and tested the same set of WT M^{pro} mutations in the more sensitive loss-of-signal assay (Fig. S5B). As expected, while single mutants had a mild effect on M^{pro} susceptibility to nirmatrelvir and ensitrelvir, triple mutants showed increased resistance compared to single mutants (Fig. S7). Initially, the most prominent differences between PIs were observed for L167F/F305L/F8L with 18.4-fold and 40.5-fold, and L167F/P168S/L57F, with 25.8-fold changes and 34.6-fold changes for nirmatrelvir and ensitrelvir (Fig. 3A and Fig. S7), respectively. When T21I was added to L167F/F305L/P184S, the resistance phenotype substantially increased, with 49.0- and 35.5-fold changes in the IC₅₀ values of nirmatrelvir and ensitrelvir, respectively (Fig. 3A and Fig. S8A). M^{pro}-O/A206T-Off, M^{pro}-O/A206T/K100N-Off, M^{pro}-O/A206T/T198I-Off, M^{pro}-O/A206T/F8L-Off and M^{pro}-O/Y54H/F305L-Off showed IC₅₀ fold-changes of 27.3, 26.0, 75, 100.2 and 11.6, respectively. Moreover, we observed differences between responses against nirmatrelvir and ensitrelvir, in particular for M^{pro}-O/Y54H/F305L-Off, which was more resistant against ensitrelvir with an IC₅₀ fold-change of 20.1 (Fig. 3B and Fig. S8B).

VSV-M^{pro} selected mutations are also resistant in a non-infectious SARS-CoV-2 replicon

To support the resistance phenotypes that have been observed in the VSV-based systems, we used a previously described SARS-CoV-2 replicon assay (45). By removing essential genes from SARS-CoV-2, the replicon allows testing authentic viral phenotypes, such as responses to inhibitors or the effect of mutations, without the risk of genetically altering a dangerous pathogen (Fig. 4A).

The mutations selected by VSV-M^{pro} also conferred resistance in the SARS-CoV-2 replicon against nirmatrelvir and ensitrelvir (Fig. 4B,C). Overall, the resistance phenotypes against nirmatrelvir in the VSV-based system (M^{pro}-Off) and the replicon correlated more than against ensitrelvir (Fig. 4D). Although the patterns of resistance against nirmatrelvir were amplified in the replicon, they were in general similar with the notable exception of S144A (Fig. 4E). Likewise, S144A displayed a stronger phenotype against ensitrelvir in the replicon compared to VSV (Fig. 4F).

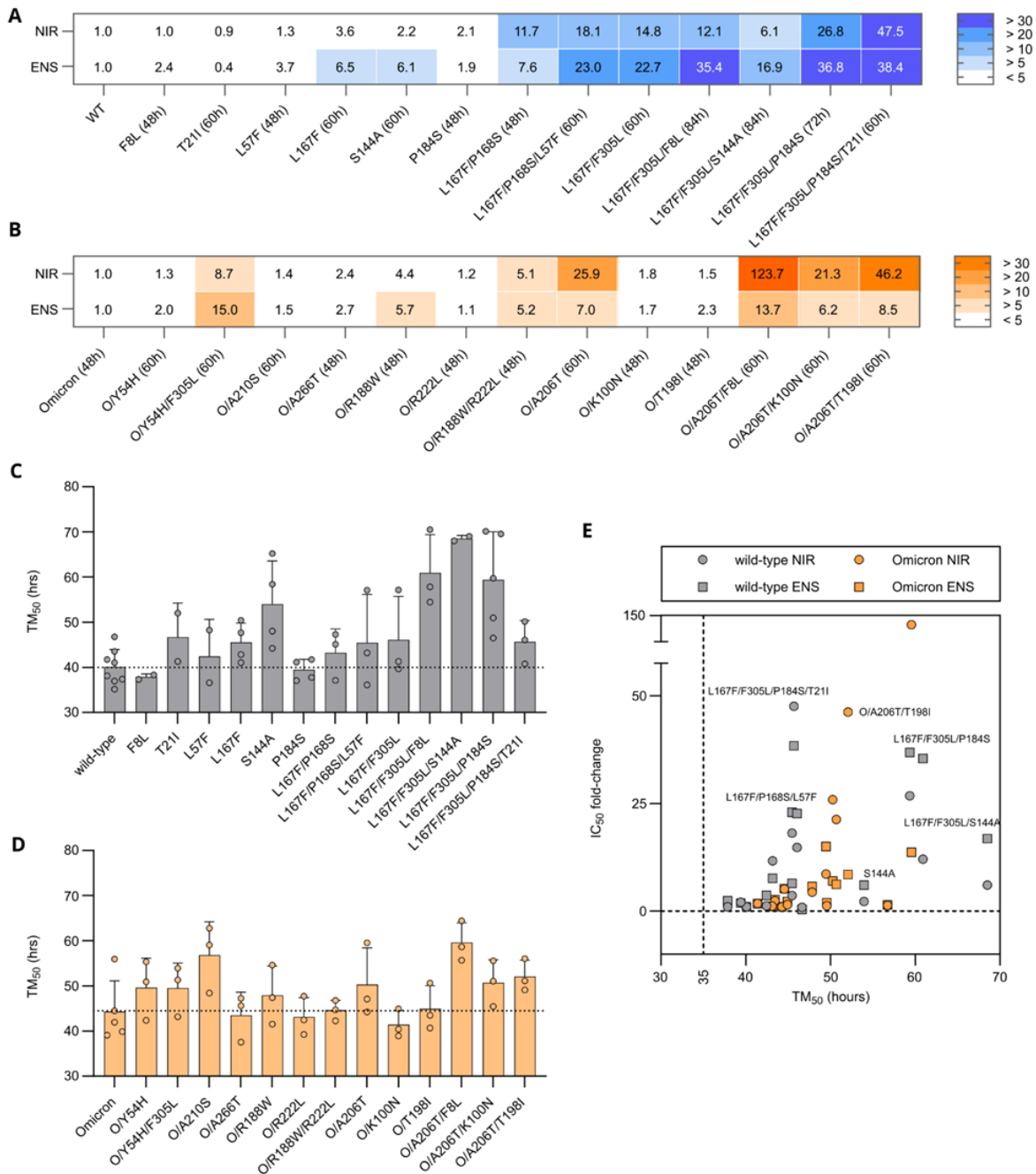


Figure 3. Resistance data and viral replication kinetics of WT, Omicron and mutant main proteases (A) IC_{50} fold-changes of nirmatrelvir and ensitrelvir tested against WT-derived mutants. Fold-changes were calculated by dividing the IC_{50} of each mutant with the IC_{50} of the WT protease. Data is displayed as mean of $n = 2$ / $n = 3$ / $n = 4$ independent experiments. (B) IC_{50} fold-changes of nirmatrelvir and ensitrelvir tested against WT-derived mutants. Fold-changes are calculated by dividing the IC_{50} of each mutant with the IC_{50} of the Omicron protease. Data is displayed as mean of $n = 2$ / $n = 3$ independent experiments. (C) Bar plot showing the TM_{50} (hours post infection) of WT and derived mutants. Data is displayed as mean $\pm$ SD of $n = 2$ to $n = 5$ independent experiments (with the exception for WT with $n = 8$, used as internal control in each experiment). Each kinetic experiment comprised $n = 6$ / $n = 8$ or $n = 12$ technical replicates. (D) Bar plot showing the TM_{50} (hours post infection) of Omicron and derived mutants. Data is displayed as mean $\pm$ SD of $n = 2$ / $n = 3$ independent experiments (with the exception for Omicron with $n = 5$). Each kinetic experiment comprised $n = 6$ / $n = 8$ or $n = 12$ technical replicates. (E) Scatter plot showing the relationship

between acquired resistance and the change in replication kinetic. Mean TM_{50} values and mean IC_{50} fold-changes are plotted on the x- and the y-axis, respectively (WT, grey; Omicron, orange; circles, nirmatrelvir; squares, ensitrelvir).

Purification and resistance characterization of M^{pro} variants through dose response experiments via a fluorescent-based biochemical assay

To further support our findings, recombinant proteases of several selected mutants (WT, L57F, L167F/P168S, L167F/P168S/L57F and Omicron, O/A206T, O/A206T/T198I) were produced as shown in Fig. S10A,B. The activity of the purified WT proteases was tested with the recombinant proteases and a reporter bearing the nsp4/5 or N-terminal M^{pro} auto-cleavage sequence of the SARS-CoV-2 polyprotein (Fig. S10C). The amount of the proteolytically processed fragment of the reporter in a western blot decreased with adding up mutations (Fig. S10D). Additionally, we tested the WT and catalytically inactive M^{pro} C145A with nirmatrelvir and ensitrelvir in a fluorescence cleavage-based assay, where the activity of the purified enzyme leads to an increased release of fluorescent 7-Amino-4-Methylcoumarin (AMC), whereas inhibition blunts it (Fig. S10E). We also applied an HPLC-based substrate cleavage assay that facilitated higher substrate doses necessary for kinetic experiments (Fig. S10F). With the AMC-based assay, we observed increasing resistance phenotypes of Wuhan-1 single to triple mutants against nirmatrelvir and ensitrelvir (Fig. S11A,B, Table 1). Omicron WT and derived mutants did not show a clear resistance in this assay (Fig. S11C,D). Kinetics performed with the HPLC-based assay revealed that Wuhan-1 and Omicron WT processed more fluorogenic substrate than its mutants (Fig. S11E,F). Accordingly, the 'net output' as determined by the ratio of k_{cat}/K_m of Wuhan-1 and Omicron WTs was higher than their respective mutants (Table 1).

Table 1. Purified M^{pro} dose response and kinetic data.

Protease	FC Nir	FC Ens	V_{MAX} [$\mu M \cdot min^{-1}$]	K_M [μM]	k_{cat} [min^{-1}]	k_{cat}/K_M [$\mu M^{-1} \cdot min^{-1}$]
Wildtype (WT)	1	1	74.4	329.8	37.2	0.1128
L57F	1.8	4.1	24.3	208.9	12.2	0.0283
L167F/P168S	9.4	9.8	76.5	528.1	38.2	0.0724
L167F/P168S/L57F	25.6	69.9	24	648.7	12	0.0185
Omicron	1	1	426.6	356.8	213.3	0.5978
O/A206T	2	1.3	2526	2467	252.6	0.1024
O/A206T/T198I	2.3	1.6	1966	1885	196.6	0.1043

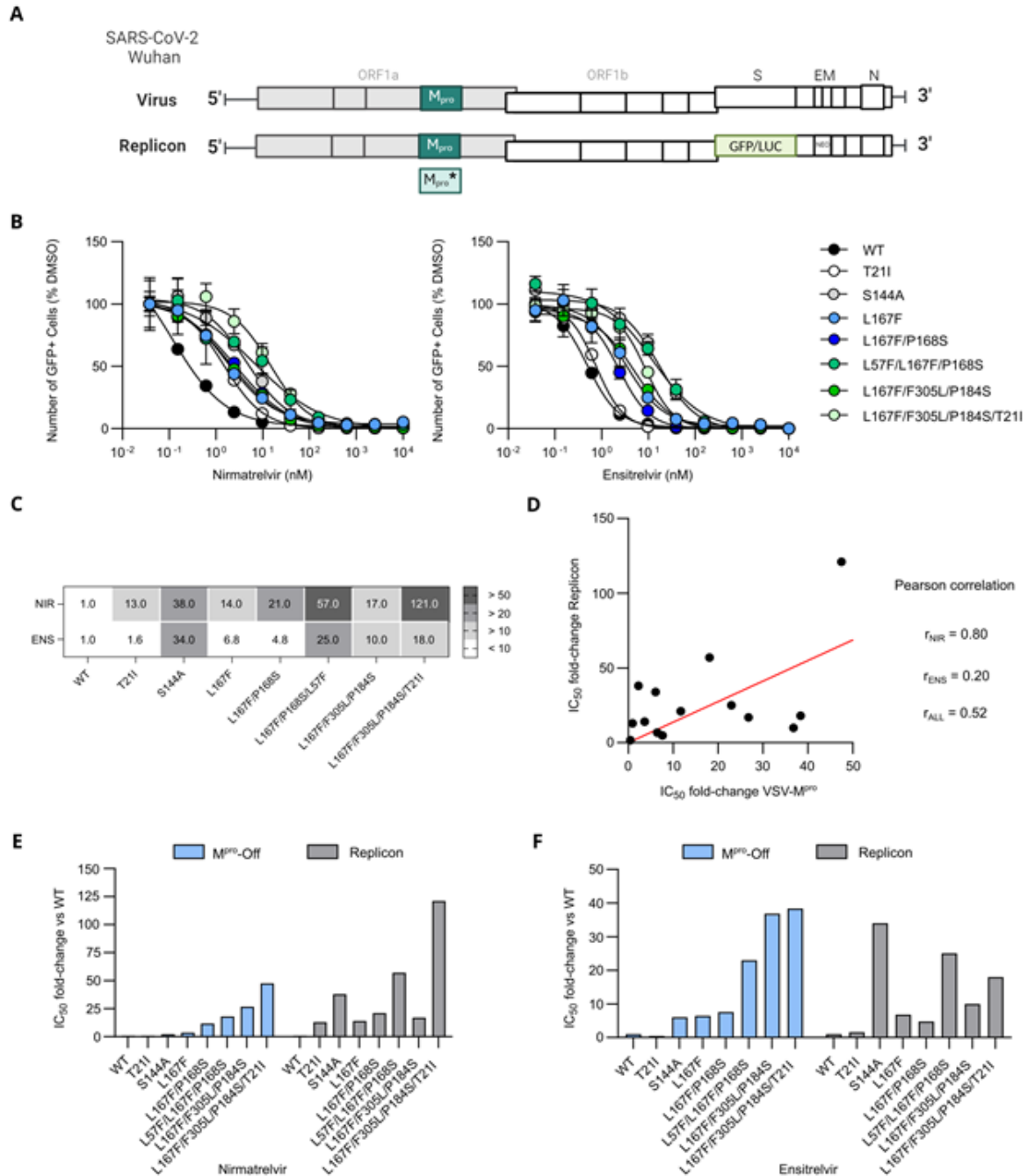


Figure 4. SARS-CoV-2 replicon dose responses of WT and mutants against nirmatrelvir (A) Schematic view of the SARS-CoV-2 genome and the replicon system that was generated. Spike and E and M proteins were exchanged with GFP/LUC and NEO (B) Dose responses experiments of nirmatrelvir (left) and ensitrelvir (right) against M^{pro} mutants tested with the replicon system. Data is presented as $\pm$ SEM of $n = 3$ technical replicates. (C) SARS-CoV-2 replicon IC₅₀ fold-changes compared to WT. (D) Linear regression analysis of resistance data generated by performing M^{pro}-Off and SARS-CoV-2 replicon assays. M^{pro}-Off IC₅₀ fold-changes are plotted in the x-axis and replicon IC₅₀ fold-changes are plotted on the y-axis. Pearson's correlation coefficients are displayed below the plots (nirmatrelvir and ensitrelvir grouped: $r_{ALL} = 0.52$; nirmatrelvir only: $r_{NIR} = 0.80$, ensitrelvir only: $r_{ENS} = 0.20$). Overall, the two systems have a moderate correlation. A strong correlation was found for nirmatrelvir, whereas a moderate correlation was seen for ensitrelvir. Bar plot of the comparison between VSV-based assay to replicon resistance phenotypes (IC₅₀ fold changes) of nirmatrelvir (E) and ensitrelvir (F).

Computational analyses reveal destabilization of protease-inhibitor complex formation and provide insights on protease stability and dimerization affinity upon mutation

To provide potential mechanisms of resistance, we performed molecular modelling with Bioluminate (46–49). This software uses already existing structural data to calculate the impact of a mutation on the stability of a specific protein conformation. The software returned models of mutant structures and delta stability values (Δ _Stability in kcal/mol), indicating whether a mutation stabilizes (negative values) or destabilizes (positive values) the investigated conformation. Many of the investigated mutations returned positive values, indicating destabilization of the nirmatrelvir-binding conformation in the mutant WT and Omicron-M^{Pro} structures, (PDB entries 8DZ2 (50) and 7TLL (51)) (Fig. S11G). The destabilization of the overall structure by mutations correlated with the destabilization of the interaction between M^{Pro} and the inhibitor nirmatrelvir as well as M^{Pro} and one of its natural substrates nsp5/6, suggesting that the inhibitor and natural substrate preferably bind to a certain, stable confirmation of the enzyme.

Nirmatrelvir is a covalent peptidomimetic inhibitor, whereas ensitrelvir a non-covalent, non-peptide inhibitor. Although they both bind to the active site of M^{Pro}, they locate to different subsites (N – S1, S2 and S4; E – S1, S2 and S1') and these distinct binding modes are affected differently by mutations such as L167F (Fig. 5A) (50) (52). Indeed, when modelling the mutant L167F, we obtained a higher Δ _Stability value for ensitrelvir (+235.7 kcal/mol) than for nirmatrelvir (+61.9 kcal/mol). A structurally related variant is the triple mutant L167F/F305L/F8L, which we investigated in more detail, as both N- and C-terminal regions are mutated (F8L and F305L, respectively). The folding of M^{Pro} causes the residues F8 and F305 to be close one another, interacting with a pi-pi T-stack. In our M^{Pro} assays, mutations F305L and F8L alone did not confer resistance (Fig. S6,7), and in our previous work the F305L mutant showed an increase in the replication kinetics (35). We observed a similar stability for F8L (+1.09 kcal/mol) and an increase in stability for F305L (-5.7 kcal/mol). To investigate the effect of F305L on M^{Pro} in more detail, we performed dimerization affinity prediction calculations using Osprey 3.3 (53). This program exploits rotamers of residue side chains to provide an estimation of binding affinity between two interaction partners, with higher Log₁₀ K* scores indicating stronger and lower Log₁₀ K* scores weaker binding. An increase in the binding affinity of +9.3 in the F305L mutant was calculated compared to the WT, suggesting an increase of M^{Pro} dimerization affinity and formation of the mature dimer. F8L alone, conversely, did not influence dimerization affinity (Fig. S11H). Upon mutation to L305, the shorter leucine side chain is located further away from the backbone of I152 on protomer A and S123 on protomer B, potentially reducing the electronic repulsion between protomers A and B (Fig. 5B). In addition, upon mutation of F8 to L8 and F305 to L305, the previously mentioned T-stack interaction is lost. The simultaneous presence of mutations F8L and F305L likely causes an increase in C-terminus flexibility that could compete with the ligand for the binding site by interfering with its entry or acting as a 'wiper' by occasionally moving towards the active site (Fig. 5C, Movie S6).

A comparison of recent M^{Pro} crystal structures harboring mutations in the domain II-III linker (A193P, E166V, L167F) and the Omicron signature mutation (P132H) revealed high structural variation in this region (Fig. 5D). These conformational changes seem to allosterically affect the S1', S2 and S4 pockets, in which several of the mutants selected in this study are located. Mutations P168S, P184S from the triple and quadruple variants WT-L167F/P168S/L57F and WT-L167F/F305L/P184S/(T21I) as well as T198I and A206T from O-A206T/T198I are located in or near the domain II-III linker, affecting the conformation of the linker and adjacent binding pockets.

Triple mutant L167F/P168S/L57F was mildly more resistant to ensitrelvir than nirmatrelvir. Structural modelling as well as the proximity of F57 to the S1' pocket suggest that L57F destabilizes ensitrelvir binding preferentially. Stability calculations on L167F/P168S/L57F resulted in a large destabilization for nirmatrelvir (235.7 kcal/mol) and even stronger destabilization of ensitrelvir (417.4 kcal/mol).

O/A206T was frequently observed in our selection experiments, prompting us to investigate it in detail. The larger, polar side chain of the threonine mutant is not compatible with the hydrophobic local environment, sterically clashing with the residues nearby and requiring conformational changes to accommodate the side chain of threonine, affecting both domain III and the domain II-III linker (residues 185-200) (Fig. 5D) (54).

In addition to their impact on drug sensitivity, some mutations showed changes of their TM₅₀ (Fig. 3C,D),

indicating impaired protease activity. Therefore, we performed overall protein stability calculations of several interesting mutants. Thus, we predicted a moderate to high decrease of protein stability for most mutated protein conformations, with the exception of F305L (increase), F8L (comparable/mild decrease) and T21I (nirmatrelvir: mild increase; ensitrelvir, mild decrease) (Table S3).

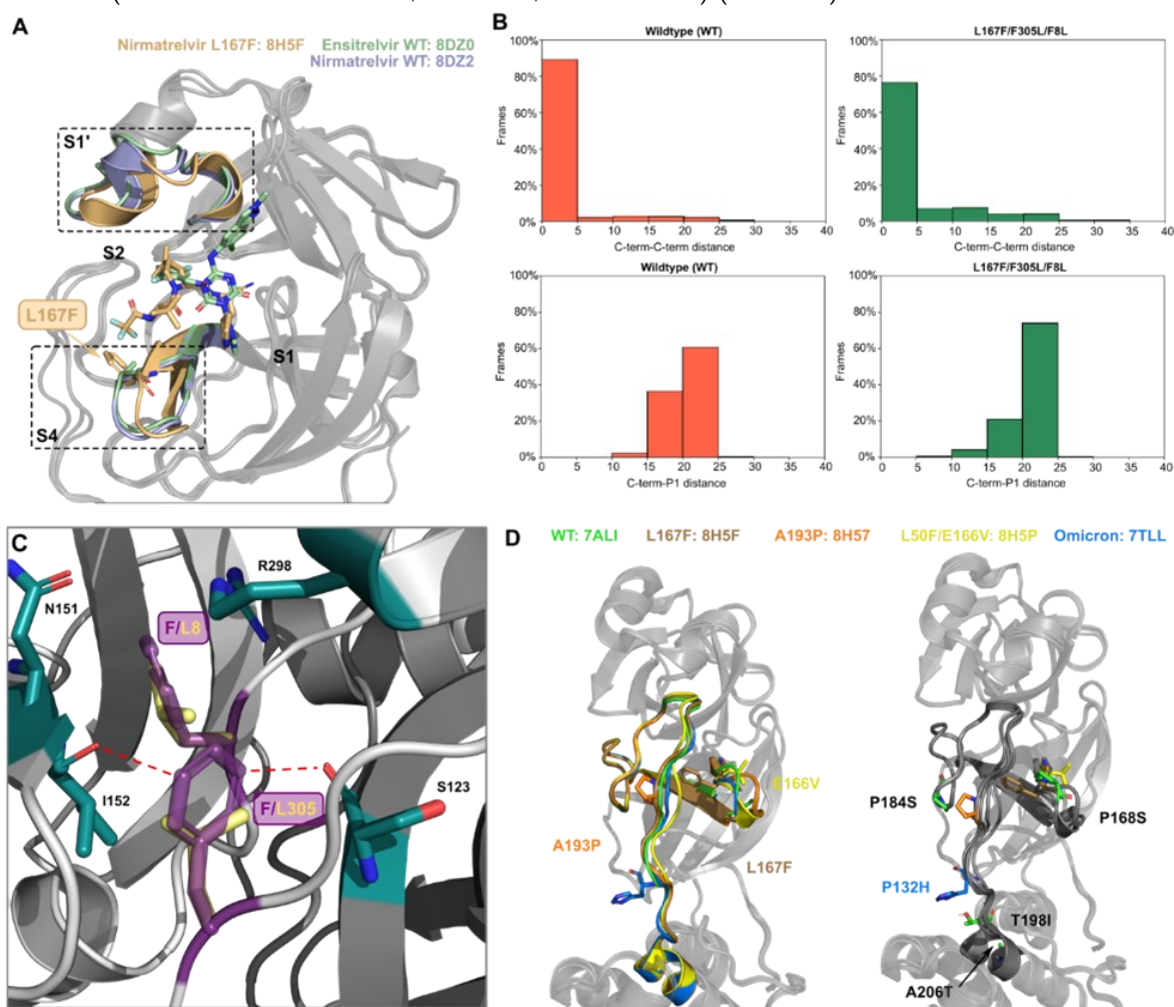


Figure 5. Structural analysis of mutants. (A) Superimposition of nirmatrelvir-WT (8DZ0), ensitrelvir-WT (8DZ2) and nirmatrelvir-L167F bound (8H5F) crystal structures. L167F is labeled in light brown and the two L167F-affected catalytic subpockets are highlighted with dotted-boxes (B) Bar plots showing the percentage of frames against C-term/C-term distance or C-term/P1 distance for WT (left) and for the L167F/F305L/F8L mutant (right). (C) The F8 and F305 aromatic side chains (dark violet) form a pi-pi T-stack and F8 in addition interacts with N151, I152, and R298 (green sticks, PDB entry 7ALI (57)). F8 is located within the homodimer interface, with chains A and B colored light violet and blue, respectively. The conjugated p-orbitals of the F305 side chain are located nearby of the I152 and S123 backbone oxygen atoms, potentially leading to electronic repulsion. In the F305L mutant, the distance of the leucine side chain (yellow sticks) to these oxygen atoms is considerably increased. (D) Superimposition of recent M^{pro} and Omicron crystal structures harboring different mutations (A193P (PDB entry: 8H57 (52)), or near (E166V (PDB entry: 8H5P (52)), L167F (PDB entry: 8H5F (52)) the domain II-III linker loop, and the Omicron signature mutation (P132H (PDB entry: 7TLL (51)).

To investigate the protease-ligand complex stability while considering protein dynamic effects, we performed Thermal Titration Molecular Dynamics (TTMD) (58) simulations. These simulations return the so-called MS coefficient, which is a measure for how long the protein-ligand complex is

stably formed. This measure facilitates a qualitative comparison of the different protein-ligand complex stabilities, for example, nirmatrelvir bound to different mutant M^{Pro} variants.

Five catalytic site mutants (Fig. 2A, Fig. 3A) that arose from the WT-L167F-M^{Pro} were selected for this modelling and compared to WT M^{Pro}-nirmatrelvir. Three M^{Pro}-ensitrelvir complexes were modelled (Table 2). Each TTMD simulation was repeated five times and an average calculated from these in-silico replicates. To illustrate the stabilities of the different complexes, we ranked the mutants according to MS coefficients, showing that the nirmatrelvir-WT complex is the most stable, followed by L167F/F305L/S144A > L167F/F305L/P184S/T21I > L167F/F305L/F8L > L167F/F305L/P184S and > L167F/P168S/L57F. TTMD simulations were also performed on ensitrelvir-M^{Pro} complexes, were again we observed that the ensitrelvir-WT complex is the most stable, followed by L167F/P168S/L57F > L167F/F305L/P184S > L167F/F305L/S144A.

Table 2. MS coefficients of WT and mutants. MS coefficients were determined for each TTMD replicate. The average MS value was calculated from the replicates. Stabilities were ranked increasingly from highest (1), which was the Wuhan-1 WT.

	rank	Average	RUN 1	RUN 2	RUN 3	RUN 4	RUN 5
Nirmatrelvir							
Wildtype (WT)	1	0.0038	0.0029	0.0028	0.0048	0.0037	0.0047
L167F/P168S/L57F	6	0.0065	0.0055	0.0069	0.0111	0.0038	0.0045
L167F/F305L/S144A	2	0.0043	0.0029	0.0044	0.0052	0.0042	0.0047
L167F/F305L/P184S	5	0.0050	0.0049	0.0071	0.0050	0.0054	0.0026
L167F/F305L/F8L	4	0.0049	0.0046	0.0041	0.0038	0.0039	0.0083
L167F/F305L/P184S/T21I	3	0.0045	0.0040	0.0044	0.0052	0.0049	0.0038
Ensitrelvir							
Wildtype (WT)	1	0.0029	0.0024	0.0019	0.0029	0.0026	0.0043
L167F/P168S/L57F	2	0.0033	0.0033	0.0028	0.0046	0.0023	0.0033
L167F/F305L/P184S	3	0.0038	0.0037	0.0041	0.0040	0.004	0.0032
L167F/F305L/S144A	4	0.0041	0.0061	0.0033	0.0047	0.0037	0.0027

TTMD simulations generate further useful metrics that can be visualized, most importantly 'rainbow plots', root mean square deviation (RMSD) plots, titration profiles, interaction energy plots and inhibitor-protease structures. Interaction energy plots and inhibitor-protease structures of Wuhan-1 WT M^{Pro} show that the interaction pattern (per-residue interaction energy) changes and the active site constantly reshapes throughout the simulation (Fig. 6A, Movie S1). The active site reshaping is particularly strong in the most resistant mutant structures L167F/F305L/P184S/T21I and L167F/P168S/L57F in the area affected by the mutations (Fig. 6B-F, Movies S2-S5). As displayed in the rainbow and root mean square deviation (RMSD) plots as well as the titration profile of Wuhan-1 WT M^{Pro}, the interaction with nirmatrelvir became unstable at 80 ns, 375 K (Fig. S12A). These metrics help to assess the dynamic instability of mutated M^{Pro} variants relative to the WT. In mutant L167F/F305L/P184S/T21I the TTMD indicated a stabilizing effect of the T21I mutation (Fig. 6B, Movie S5). L167F/P168S/L57F led to a loss of interaction with the backbone of crucial catalytic site residues (164-166) by altering the floor of the binding site, causing a drag effect on nirmatrelvir (Fig. 6C and Movie S1). The resistance mechanism of the L167F/F305L/S144A mutant combines an increased instability of the oxyanion loop (residues 138-145) caused by S144A. This potentially increases plasticity of the floor of the binding site due to the L167F mutation, which finally leads to the loss of the hydrogen bond between nirmatrelvir and H163, subsequently resulting in the loss of interaction with residue Q189 (61) (Fig. S11B and Movie S2). L167F/F305L/F8L has only one mutation within the catalytic site (L167F) and two other ones located at the dimeric interface, F8L and F305L (Fig. S11D). As mentioned above, the C-terminus is freer to move in the available space, potentially competing with the ligand in a 'wiper'-like motion (Movie S6). Ensitrelvir

binds to a different subsite of the catalytic pocket of M^{Pro}, which is reflected in its interaction energy plot (Fig. 6D). In L167F/F305L/P184S, the P184S mutation enhances the plasticity of the loop spanning from residue S184 to Q192, thereby reducing the ligand-pocket shape complementarity, which is essential for both nirmatrelvir and ensitrelvir positioning. The higher plasticity of the loop affects the stability and the number of water molecules hosted between domains I-II, including the stable water commonly interacting and connecting H164, H41, and D187 residues, a water molecule that holds not only a structural role but has also been described as a third member of the catalytic dyad (Fig. 6E and Movie S3). The mutations L167F/P168S/L57F and L167F/F305L/S144A reshaped the catalytic pocket and strongly destabilized ensitrelvir interactions in the S2 pocket, leading to a loss of the hydrogen bond with E166 and the π - π stacking with H41 (Fig. 6E, Fig. S12D).

Discussion

In this study, we used a previously developed, safe, VSV-based tool to select for protease inhibitor resistance mutations. We demonstrated the capability of our selection tool to achieve SARS-CoV-2 M^{Pro} variants with multiple substitutions and potential Omicron-related mutants relevant to the current pandemic situation. We confirmed the resistance of several of our mutations in gold-standard biochemical and SARS-CoV-2 replicon assays. Lastly, we provide comprehensive static and dynamic modelling to suggest resistance mechanisms.

Focus on specific mutants

First, we generated a large number of mutants, which we subsequently filtered on a three-parameter basis for further characterization: proximity to the inhibitor, prevalence in the GISAID database, and frequency of a specific mutation in our selection experiments. Substitutions that appeared close to the inhibitor are likely to yield a meaningful, structural explanations that can then inform structural derivatization of inhibitors. Highly frequent substitutions in GISAID are relevant because they occur in viruses that can efficiently propagate in humans and therefore should not be detrimental to the enzymatic activity. Lastly, mutations that occur frequently in selection experiments could be favoured routes of evolution for the M^{Pro} enzyme (62) and thus the mechanism might be of interest.

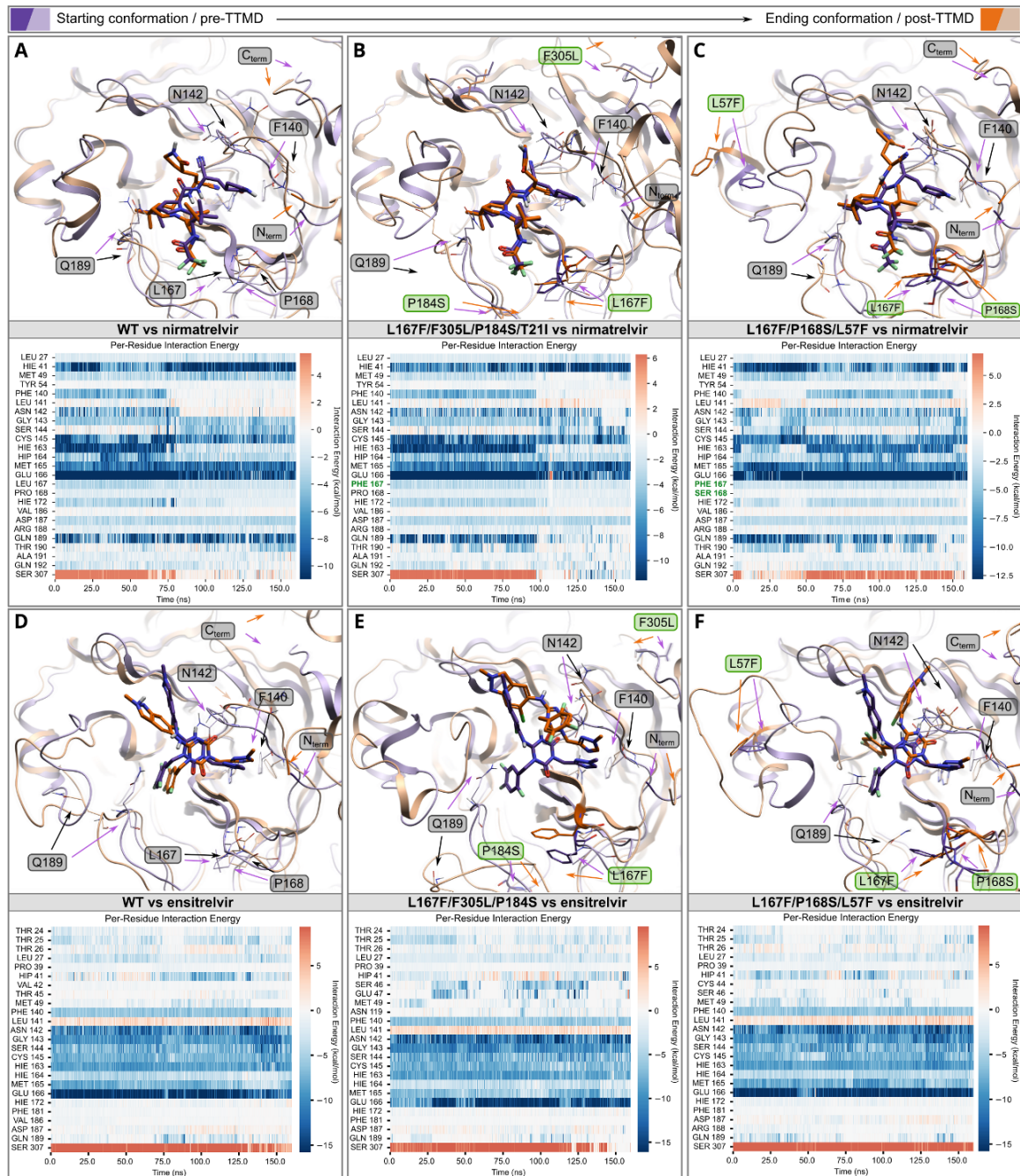


Figure 6. Thermal Titration Molecular Dynamics (TTMD) experiments of nirmatrelvir and ensitrelvir M^{pro} complexes. TTMD simulation data of the nirmatrelvir-WT- M^{pro} (A), nirmatrelvir-L167F/F305L/P184S/T21I- M^{pro} (B) and nirmatrelvir- L167F/P168S/L57F- M^{pro} (C) complexes. Top: overlay of M^{pro} structure at the beginning (violet/purple) and at the end of the simulation (light brown/orange); bottom: heat map of interaction energies between ligand and surrounding residues. TTMD simulation data of the ensitrelvir-WT- M^{pro} (D), ensitrelvir- L167F/F305L/P184S- M^{pro} (E) and ensitrelvir- L167F/P168S/L57F- M^{pro} (F) complexes. Top: overlay of M^{pro} structure at the beginning (violet/purple) and at the end of the simulation (light brown/orange); bottom: heat map of interaction energies between ligand and surrounding residues. Mutated residues and residues that are mentioned in the results/discussion are labelled in green and grey, respectively.

Wuhan-1 M^{pro} resistance mutations

Based on the VSV-L167F- M^{pro} variant, additional substitutions close to the inhibitor were selected, namely

S144A, P184S and L57F. In some of those mutants, we observed differences in the susceptibility to the two protease inhibitors nirmatrelvir and ensitrelvir. Interestingly, the mutant L167F/F305L/F8L demonstrated a considerably greater degree of resistance to ensitrelvir than to nirmatrelvir. Both L167F/F305L/P184S and L167F/F305L/L57F showed a similar trend. L167F/P168S/S144A had a comparable resistance profile against both protease inhibitors. Mutations of S144 can simultaneously cause an increase in drug resistance and a decrease in the catalytic activity of the protease. Consistently, in our kinetic experiments and in the literature (18) this mutation impaired replication.

As mentioned above, the second filtering parameter we applied was the absolute frequency of a specific nsp5/M^{pro} substitution in the GISAID database. Residues in catalytic sites tend to be more conserved (41, 42), and we expected them to not be frequently represented. Indeed, most high-frequent mutants were observed in positions further away from the active pocket. P184S is an exception as it is frequently represented (4713 entries, as of 01/06/2023) and is near the catalytic pocket. During the rescue (generation of a virus from a plasmid) of L167F/F305L/P184S, one sample had acquired an additional mutation, namely T21I. The resistance profile of L167F/F305L/P184S changed from more resistant against ensitrelvir to stronger resistance against nirmatrelvir upon introduction of the T21I mutation. However, T21I alone neither conferred resistance against nirmatrelvir nor ensitrelvir in both our gain- and loss-of-signal assays. Consistently, it has been reported in the literature as a compensatory mutation for the fitness loss caused by other resistance mutations, such as E166V (18). T21I could support or compensate the increased plasticity of L167F/F305L/P184S we found in TTMD simulations by an allosteric, re-stabilizing effect (25, 63).

Taken together, it seems that L167F, P168S, S144A and P184S mutations deform the binding pocket and loosen the electrostatic interactions between nirmatrelvir and the protease. This observation indicates that not only mutations of residues directly interacting with the inhibitor can decrease susceptibility, but mutations of residues within the catalytic pocket can trigger conformational changes in the enzymatic cleft that impact in the inhibitor-protease complex formation as well, a finding that is also supported by recent structural data (52).

The triple mutants L167F/F305L/F8L and L167F/P168S/L57F are among our most resistant M^{pro} variants and showed the most marked difference in susceptibility between the two inhibitors nirmatrelvir and ensitrelvir. The residues F305 and F8 are in close contact with each other in the folded protein and this is expected to be a stabilizing interaction for the dimer interface. However, other residues are viable at this position, for example leucine (L) has been shown to be the preferred residue instead of phenylalanine (F) (42). In fact, we calculated an increase in the dimerization affinity from F305 to L305, suggesting an improved formation of the fully functional dimeric protein upon mutation. We hypothesize that the simultaneous mutation of both F305 and F8 residues into leucines results in a modified interaction between the C-term of the first protomer and the N-term of the second, increasing the conformational degree of freedom of the C-term, thereby leading to a “wiper” mechanism that could compete with the inhibitor for the catalytic site.

Omicron-M^{pro} resistance mutations

The Omicron-M^{pro} sequence has a signature substitution at amino acid position 132, exchanging proline for histidine (P132H). This mutation has been shown to not confer resistance to approved protease inhibitors, such as nirmatrelvir and ensitrelvir (30–33), in agreement with our findings, but altered thermal stability (34). To reduce biosafety risks, previous studies based on authentic SARS-CoV-2 to select resistance mutations have used early variants of SARS-CoV-2 (18, 19, 21). Protease inhibitor resistant mutants derived from such early variants would unlikely be able to compete with Omicron variants if released accidentally. This caveat does not apply to mutation selection performed with our VSV-M^{pro} system. Consequently, we selected protease inhibitor resistant Omicron M^{pro} mutations within and outside the catalytic site. The catalytic site mutant O/Y54H, which occurred 10 times independently, and O/R188W showed only mild resistance alone as well as in combination with R222L. O/Y54H/F305L showed an increased, selective resistance against ensitrelvir, with a 20-fold-change of the IC₅₀ in the loss-of-signal assay (M^{pro}-Off). The mutants K100N, T198I, A210S, and A266T were frequent in GISAID, therefore we picked them for resistance characterization. Double mutants O/A206T/F8L and O/A206T/T198I showed strong selective resistance against nirmatrelvir (up to 100-fold-change), while

being less effective for ensitrelvir. We speculate that in O/A206T/F8L, the contribution of F8L to resistance is the same as in mutant L167F/F305L/F8L (“wiper” effect). O/A206T and O/A206T/K100N also showed considerable resistance (~26-fold), and the computed destabilization correlated with resistance.

Most Omicron and WT substitutions, however, occurred in the amino acid sequence that constitutes domain III (residues 200–306) of M^{pro}. We chose to characterize the A206T mutation due to the unusually high frequency in the selected VSV-O-M^{pro} pool of mutants. A206T only fits our third filtering parameter: frequent selection in our experiments, as it is far from the catalytic site and scarcely represented in the GISAID database. Likewise, some mutations may confer an advantage for viral replication only in the presence of a protease inhibitor but could be unfavourable in an untreated individual. In a study on SARS-CoV M^{pro} polyprotein maturation (64), the authors proposed that despite the presence of deleterious mutations that hinder mature dimerization, M^{pro} retains its ability to cleave its N-terminal cleavage site owing to a weak, immature dimerization between two polyprotein monomers (1 and 2) catalysed by domain III (1) – domain III (2) interaction. We speculate that mutations of residues between 200 and 306, most frequently found in VSV-O-M^{pro}, positively contribute to this interaction. As with A206T, we hypothesize that destabilization of the mature M^{pro} to which nirmatrelvir binds, is the resistance mechanism.

Viral replication kinetics

In kinetic studies, we observed that the gain in signal was slower for some mutants, which might be an indirect read-out of M^{pro} activity. The L167F/F305L/S144A mutant was among the slowest variants and the time to reach the plateau was 1.75 times longer than for the WT M^{pro}. However, mutations that were described as compensatory in the literature, such as T21I (18), could partially reverse the kinetic attenuation. T21I appeared after generating VSV-L167F-F305L-P184S-M^{pro} from its plasmid (also called “rescue”), and R222L after plaque purification of VSV-O-R188W-M^{pro}. T21I and R222L are likely compensating for substitutions that occur in the catalytic site. For example, T21I seems to compensate for the loss of replicative fitness of the previously described, highly resistant M^{pro}-E166V variant (18). Furthermore, T21I is a very frequently represented substitution in the GISAID database (21248 entries, 1st of June 2023). From mutations generated in the VSV-O-M^{pro} variant, viral replication kinetics exhibited a different behaviour in comparison to VSV-L167F-M^{pro} variants. Overall, mutations related to the Omicron-M^{pro} did not decrease the viral replication rate to the same extent as in the L167F-M^{pro}. The slowest variant was O/A206T/F8L, with an increase of ~16 hours to reach the plateau. As expected, most mutations that have been computationally predicted to decrease protease stability indeed had a negative impact on viral replication, with higher TM₅₀ values related to M^{pro} variants bearing such mutations. Furthermore, they were less represented in isolates in the GISAID repository (1 or 2 digit entries), whereas those that had similar TM₅₀ values compared to the WT such i.e. F8L, P184S or K100N and T198I alone, were more frequent (3 or 4 digit entries). These correlations support the validity of computational predictions

Phenotype validation with SARS-CoV-2 replicons and gold-standard enzymatic assays

To validate resistance phenotypes in independent assay systems, we used a previously described SARS-CoV-2 replicon (45) and found that the resistances between the two assay systems correlated. A notable exception was the mutation S144A, of which the resistance was underestimated by the VSV-based assays. We suspect that this phenotype is caused by the particularly strong attenuation of this mutant M^{pro} that we observed in the VSV-based kinetic Off assay. Likewise, we could observe that introducing selected mutations into purified M^{pro} enzymes increases the resistance, while at the same time reducing enzymatic fitness, as was also described before.

Caveats of the study

Chimeric VSV-Spike, where the VSV glycoprotein G was replaced by the SARS-CoV-2 spike, has been used previously to predict antibody escape mutations (65–67). Before that, VSV was shown many times to be promiscuous in the context of its glycoprotein, which could easily be replaced for example with those of Ebola and Marburg viruses (EBOV, MARV) (68) or lymphocytic choriomeningitis virus (LCMV) (69).

To replace one of the VSV intergenic regions with a protease and thereby use VSV as a protease mutation tool, however, was a previously untapped area of research. Therefore, our resulting mutations need careful evaluation. We observed many mutations in chimeric VSV-M^{pro}, some of which aligned exactly with those identified in SARS-CoV-2 gain-of-function experiments reported in the literature (18, 19), whereas others were completely different. One explanation may be the artificial system used in our experiments, which uses cis-cleavage, the only requirement for M^{pro} processing in VSV-M^{pro} replication. In contrast, in experiments with authentic SARS-CoV-2, both cis- and trans-cleavage must be preserved in the development of resistance mutations. As described above, during auto-cleavage there is an intermediate dimerization state, different from the mature protease. This intermediate state may require specific interactions that differ from those involved in the mature dimer. Another reason is the difference between the natural polyprotein of coronaviruses and the artificial G-M^{pro}-L polyprotein expressed by chimeric VSV-M^{pro}. The observed divergence between SARS-CoV-2 and VSV may also be attributed to the differences between two viral polymerases: SARS-CoV-2 RNA-dependent RNA polymerase (RdRp) and VSV polymerase L. It has been reported in the literature that the SARS-CoV-2 polymerase can proofread (70), whereas VSV lacks such mechanisms (37), leading to a much higher error rate of 1/10,000 nucleotides, as described previously (36). However, this could also be seen as an advantage as resistance mutations may develop faster in the error prone VSV-M^{pro} replication system. Lastly, VSVs replication cycle is much faster than that of SARS-CoV-2, in part owed to the simpler structure of the VSV genome and replication strategy (REFs). Therefore, VSV proteins are expressed faster and in greater number in a shorter time, which accounts for the higher amounts of M^{pro} inhibitor needed when treating VSV-M^{pro} vs. SARS-CoV-2.

Materials and Methods

Study Design

The overall rationale of the study was to use a previously developed mutation selection tool based on VSV to select and characterize a comprehensive collection of SARS-CoV-2 main protease inhibitor escape mutants. The study was performed on cell lines, in biochemical settings and in-silico, and no animal husbandry or human participants were involved. Human cell lines with replicating BSL-1 and -2 viruses were treated with protease inhibitors to observe resistance phenotypes in appropriate facilities. Viral titres were determined using TCID₅₀. Measurement readouts were fluorescence-based, detected by ELISpot/FluoroSpot and multi-well readers. Autofluorescent fibers were excluded automatically and manually from spot counting in the ELISpot readout. Experiments were neither blinded nor randomly distributed to experimenters. We chose sample sizes empirically based on experience from former studies. At least two and up to four biologically independent replicates were performed per condition. Biologically independent meant distinct wells with the same condition, not multiple measurements of the same wells (technical replicates). Resistance phenotypes were reproduced at least twice, usually more often and in different combinations. Representative measurements were chosen to compile graphs and figures.

Cloning strategies

The chimeric VSV variant with M^{pro} instead of the intergenic region between G and L was cloned as previously described (35). VSV-G-M^{pro}-L carrying the Omicron signature mutation, P132H (VSV-O-M^{pro}) was cloned as follows: N-terminal fragment comprising part of G and M^{pro} was amplified by PCR using primers 33n-before-KpnI-for (Q195) and Omicron rev VSV-G-M^{pro}-L as a template, and the C-terminal fragment comprising the remaining part of M^{pro} and part of L was amplified by PCR using primers Omicron for and 33n-after-HpaI-rev (Q176). N- and C-terminal fragments were ligated together by Gibson assembly in a VSV-backbone vector digested with KpnI and HpaI. VSV-G-M^{pro}-L L167F (VSV-L167F-M^{pro}) was plaque purified as it was generated in a previous work.

M^{pro}-Off and -On point mutants were generated by point directed mutagenesis on parental plasmids (GenBank accession codes: M^{pro}-Off: ON262565; M^{pro} pro-On: ON262564) with mutation primers and the Herculase II Fusion DNA Polymerase. Herculase is a polymerase that can overcome difficult (GC-rich) sequences and amplify large plasmids. However, primers could not be chosen for optimal design because

the mutation site was fixed. For this reason, this simple point directed mutagenesis did not work for each construct. For M^{pro}-On and M^{pro}-Off plasmids, where point-directed mutagenesis did not work, we used mutagenic Gibson assembly as previously described (35). Cloning primers used in this study are shown in Table S3.

Cell lines

BHK-21 cells (American Type Culture Collection, ATCC) were cultured in Glasgow Minimum Essential Medium (GMEM) (Lonza) supplemented with 10 % fetal calf serum (FCS), 5 % tryptose phosphate broth, and 100 units/ml penicillin and 0.1 mg/ml streptomycin (P/S) (Gibco). 293T cells (293tsA1609neo, ATCC), and 293-VSV (293 expressing N, P-GFP and L of VSV) (73) were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % FCS, P/S, 2 % glutamine, 1x sodium pyruvate and 1x non-essential amino acids (Gibco). Vero E6 (ATCC CRL-1586) was cultured in DMEM supplemented with 5 % FCS (VWR) and 1 % penicillin–streptomycin–glutamine (PSG) solution (Corning). Huh-7 cells (ATCC) were cultured in DMEM supplemented with 10 % FCS .

Virus recovery ('rescue')

VSV-G-M^{pro}-L WT and a few mutants were rescued in 293T cells by CaPO₄ transfection of whole-genome VSV plasmids together with T7-polymerase, N-, P-, M-, G- and L expression plasmids as helper plasmids (43). Briefly, genome and helper plasmids were transfected into 293T in the presence of 10 µM chloroquine to avoid lysosomal DNA degradation. After 6 to 16 hours, chloroquine was removed, and cells were cultured until cytopathic effects occurred. M and G proteins were used as helper plasmids; although these proteins are optional in the recovery of VSV, they were chosen here as a precaution to support the rescue of a potentially attenuated virus variant. After the rescue, viruses were passaged on 293-VSV cells and plaque purified twice on BHK-21 cells. ΔP and ΔL VSV variants expressing dsRed were produced on replication supporting 293-VSV cells. VSV-G-M^{pro}-L was fully replication competent and produced on BHK-21 cells.

Mutation selection assay

10⁴ BHK-21 cells/well were seeded in a 96-well plate one day before infection with the chimeric VSV-M^{pro} viruses. VSV-L167F-M^{pro} and VSV-O-M^{pro} infected cells (MOI = 0.01) were treated with nirmatrelvir with concentration ranging from 10 to 100 µM. The resistant variants VSV-L167F/P168S-M^{pro} and VSV-L167F/F305L-M^{pro} at MOI = 0.01 were passaged 5 times, and with increasing concentrations of inhibitor (from 50 to 100 µM). VSV-O-M^{pro} and VSV-O/A206T-M^{pro} at MOI ranging from 0.001 and 0.01 were passaged only once with increasing concentrations of inhibitor (from 10 to 100 µM). Each virus variant occupied from 24 wells to 96 of the 96-well plate. Supernatants from wells that displayed cytopathic effect after 48-72 hours were collected for downstream viral RNA isolation, cDNA synthesis, PCR amplification and Sanger or Nanopore sequencing as described in this section.

Viral RNA isolation, cDNA synthesis and M^{pro} sequencing

VSV-G-M^{pro}-L RNA was isolated either by using the E.Z.N.A. Viral RNA Kit (Omega Bio-Tek Inc.), the NucleoSpin RNA Virus (Macherey-Nagel GmbH) or by semiautomated RNA extraction using the Easymag (BioMerieux™). BHK-21 cells were infected with the respective VSV-G-M^{pro}-L variant in 96-well plates as described above. Virus-containing supernatants were collected from individual 96-wells and the RNA was purified from the supernatants according to manufacturers' instructions (Easymag). Then, cDNA was synthesized from isolated viral RNA by RevertAid RT Reverse Transcription Kit (Thermo Fisher Scientific). G_{Cterm}-M^{pro}-L_{Nterm} fragment sequence was amplified by PCR with primers (primer_for: CTCAGGTGTTCTCGAACATCCTCAC and primer_rev: GATGTTGGGATGGGATTGGC) and either sent for Sanger sequencing (MicroSynth AG) or sequenced using Nanopore (as described below). For Nanopore sequenced samples, the fraction of mutated virus of the full population is shown in Table S4. Obtained sequences were mapped to the L167F-M^{pro} or Omicron-M^{pro} reference sequences in Geneious Prime 2023.0.1 and examined for mutations.

Nanopore sequencing workflow

After the first two selection experiments where we sequenced our samples with Sanger sequencing (MicroSynth AG), we used Nanopore sequencing. First, PCR amplification products of the G_{Cterm}-M^{Pro}-L_{Nterm} fragments, as described above, were used. The sequencing libraries were prepared using the Rapid Barcoding Kit SQK-RBK110.96 (Oxford Nanopore Technologies, ONT), and up to 96 samples were multiplexed and sequenced on a MinION Mk1B sequencer with R9.4.1 flowcells (ONT). Raw data in the form of electrical signals are translated into nucleotide sequences (base-calling) and saved in fastq files. Basecalling using the super high accuracy model, demultiplexing, as well as barcode and adapter sequence trimming, was performed in Guppy (version 6.1.5, ONT).

Raw reads were filtered to a PHRED quality score of $\geq$ Q15 and length between 200 bp and 1800 bp using SeqKit (version 2.4.0) (74), aligned to the reference sequence using minimap2 (version 2.22) (75), followed by sorting and indexing using SAMtools (version 1.13) (76). SAMtools depth was used to check sufficient depth. To find SNVs, we used LoFreq for variant calling with --min-cov option set to 300 (v2.1.3.1) (77).

The resulting VCF files were imported to Geneious Prime 2023.0.1 and called variants manually checked for plausibility.

Unrooted phylogenetic tree generation and visualization

Mutations had to be manually inserted into the main protease sequence for each variant. The phylogenetic tree was generated using the online tool <http://www.phylogeny.fr/> with standard parameters. The output as .newick file was imported to Geneious Prime 2023.0.1, visualized and exported as PDF. Graphic rearrangements and overhauls were made using Inkscape version 1.1.

Collection of GISAID frequencies of generated mutations

To obtain updated nsp5 frequencies, metadata from the EpiCoV database (correct as of 1st June 2023) was used; specifically, the `all\_mutations` field from a JSON dump obtained with GISAID credentials. Using Python 3.9 and the pyarrow and pandas' libraries, sequences were classified into WT (not assigned Omicron lineage, and not carrying the P132H mutation) and Omicron (either assigned Omicron lineage or carrying the P132H mutation). Then, for both these groups of sequences, the counts of 95 nsp5 mutations of interest were aggregated.

SARS-CoV-2 variant phylogenetic analyses using Ultrafast Sample placement on Existing tRees (USHER)

Phylogenetic trees were generated for the resistance variants of interest using patient derived sequences deposited in the GISAID database. For each variant of interest, viral genomes harboring the corresponding mutation in Nsp5 were retrieved using the GISAID EpiCoV web server after filtering to consider only viruses descending from the Omicron lineage and removing genomes with low sequence coverage. These sequences were subsequently uploaded to the USHER (78) web server to generate phylogenetic trees using all the sequences available in the GISAID database, visualized using the Auspice.us web application and edited to highlight the variants of interest using Adobe Illustrator. The frequency of each resistance mutation within Omicron and Pre-Omicron lineages was similarly determined using the GISAID EpiCoV web server by dividing the number of occurrences of the corresponding Nsp5 mutation within Omicron or all other SARS-CoV-2 lineages by the total number of viral genomes deposited belonging to each respective lineage. Histograms comparing the frequency of the variants within the Omicron and Pre-Omicron lineages were generated using GraphPad Prism 9.

Screening assay with ELISpot/FluoroSpot read-out

3×10^5 cells were seeded per well in 6-well plates and transfected one day after seeding with 3CL^{Pro} plasmids using TransIT-PRO (Mirus Bio LLC) and incubated 8-9 hours for M^{Pro}-On and 10 hours for M^{Pro}-Off assays. Then, cells were seeded into a 96-well plate with 2×10^4 cells per well in 50 μ l complete growth medium. Directly after seeding, compounds and virus (MOI 0.1) were added in 50 μ l complete growth medium to wells. After 48 to 72 hours, supernatants were removed, and fluorescent spots counted in a

Fluoro/ImmunoSpot counter (CTL Europe GmbH). For experiments where FluoroBrite™ medium was used, supernatants were not removed, and the signal was read out every 12 hours. The manufacturer-provided software CTL switchboard 2.7.2. was used to scan 90% (70% when FluoroBrite™ medium was used) of each well area concentrically to exclude reflection from the well edges, and counts were normalized to the full area. Automatic fiber exclusion was applied while scanning. The excitation wavelength for dsRed was 570 nm, and the D_F_R triple band filter was used to collect fluorescence. To increase comparability between M^{pro}-On and -Off signals, we normalized dsRed events with the following strategies. In M^{pro}-On, the highest compound concentrations would not reach the same value due to the different responses of each mutant. Therefore, we normalized to the highest mean of the experiment. In M^{pro}-Off, we normalized the signal to either the wild type or the Omicron proteases. For both assays, fluorescent spot count (y-axis) data were plotted against PI concentration (x-axis) and IC₅₀ values were extrapolated as described later in this section.

Replication kinetic experiments

For replication kinetic experiments with WT M^{pro}, Omicron and mutants, we adapted the 3CL/M^{pro}-Off assay to enable multiple read-outs over time every 12 hours. We exchanged the standard DMEM supplemented with 10 % FCS, P/S, 2 % L-glutamine, 1x sodium pyruvate and 1x non-essential amino acids (Gibco), with FluoroBrite™ DMEM, equally supplemented. FluoroBrite™ DMEM has the same composition as DMEM, except it lacks phenol red which interferes with spot recognition. We experimented as described in the previous paragraph. After seeding 2×10^4 transfected cells in 50 µl of FluoroBrite™ complete medium into 96-well plates, only virus (MOI 0.1) was added (50 µl). The signal was read out at 12 hours intervals. As the virus starts replicating, dsRed is expressed by the infected cells, and the expression of dsRed is correlated with the amount of viral progeny and therefore with the number of spot count. We generally stopped signal acquisition after 84 up to 108 hpi, because the cells die due to virus-induced cytopathy, and the spot count decreases. The manufacturer-provided software CTL switchboard 2.7.2. was used as described above (for FluoroBrite™ medium). Replication kinetic curves were normalized individually to the highest mean of the construct. We normalized fluorescent spot count data by the highest mean of each dataset. Then, we plotted the fluorescent dsRed signal (spot count) against time (hpi). We then performed non-linear regression analyses using the built-in function “Sigmoidal 4PL, X is concentration” in GraphPad Prism 9.5, where concentration was replaced with “hours” and extrapolated the TM₅₀ (time-maximum fifty) value, which represents the time required to achieve half of the maximum signal/plateau.

Protein expression and purification

M^{pro} with a C-terminal Hexahistidine tag (M^{pro}-6H) was expressed in E.coli followed by a tag removal with the human rhinovirus 14 3C protease (HRV3C protease).

WT M^{pro}-6H construct was codon optimized for the expression in E. coli and ordered from Thermo Scientific (Waltham, MA, USA). Point mutations to generate the mutant variants were introduced using the Q5® Site-Directed Mutagenesis Kit purchased from New England Biolabs (NEB, Ipswich, MA, USA). Primers were acquired from Sigma Aldrich (St. Louis, MO, USA).

M^{pro}-6H variants were transformed into electrocompetent E. coli BL21(DE3) cells. An overnight culture was inoculated 1:50 in 1L TB medium and cells were grown at 37°C, 220 rpm to an OD_{600nm} of 1. After induction with 0.4 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) temperature was lowered to 25°C for 4 hours. Cells were isolated by centrifugation at 3200 g, 20 minutes at 4°C, and cell pellets frozen at -20°C.

For purification, frozen pellets were resuspended in 50 mM Tris/50mM NaCl, pH 7.5 buffer and disrupted with a French press (Thermo Scientific). Lysed cells were centrifuged at 20000 g, 4°C, 20 min and the clarified supernatant was adjusted to 50 mM Tris/300 mM NaCl/20 mM Imidazole. IMAC purification was carried out with a ÄKTApurifier (GE Healthcare Life Sciences; Little Chalfon, United Kingdom) and a HisTrap FF Crude 1 mL (GE Healthcare) column. Unbound proteins washed out using 4 column volumes (CV) 50 mM Tris/300 mM NaCl/20 mM Imidazole, pH 7.5 and 1 CV 50 mM Tris/300 mM NaCl/40 mM Imidazole, pH 7.5. The protein-containing fractions were pooled, and the buffer was changed to 50 mM Tris/50mM NaCl, pH 7.5 using a HiTrap Desalting 5 mL (GE Healthcare) column. Protein concentration

was determined with a NanoDrop ND-1000 Spectrometer (Thermo Scientific).

Hexahistidine Tag removal by negative IMAC

The highly soluble Hexahistidine tagged NT*-HRV3Cprotease was expressed as previously described (79) and purified the same way as M^{pro}. M^{pro}-6H was incubated together with NT*-HRV3protease at 4°C overnight followed by negative IMAC purification with a manually packed Ni-NTA column. The flowthrough yields the mature M^{pro} with native N- and C-termini without His tag, because the His tag as well as the NT*-HRV3C protease are bound to the Ni-NTA resin.

Cross validation with biochemical M^{pro} inhibition assay

The biochemical assay used to confirm mutations was based on the 3CL^{pro} activity assay from BPS Biosciences, catalog number #78042-2. The M^{pro} in the kit was replaced by an in-house produced M^{pro} and mutants thereof (L57F, L167F/P168S, L167F/P168S/L57F), as described in the M^{pro} purification paragraph. Solutions of WT and mutant M^{pro} variants – 84.5 ng/reaction in 30 µl buffer – (20 mM Tris/HCl pH = 8, 150 mM NaCl, bovine serum albumin 0.1 mg/ml, 1 mM dithiothreitol – DTT) were prepared according to the kit's manual, to reach a final concentration of 50 nM in 50 µl. 10 µl of five-fold excess to tested nirmatrelvir concentrations (180 µM) were added to the 30 µl of M^{pro} solution and incubated for 30 minutes. Then, 10 µl of a fluorescent substrate (Ac-Abu-Tle-Leu-Gln-MCA) were added to get a final concentration of 40 µM. This generates a 1:5 dilution of the excess of protease inhibitors – and therefore final concentrations – and the reactions were incubated for 4 hours. Fluorescence was measured by excitation at 365 nm and read-out at 415-445 nm emission with a Glomax Explorer fluorometer (Promega).

Expression constructs for M^{pro} cutting reporter

M^{pro} cutting reporter: Oligonucleotides for the “short” (GCAGTGCTCCAAAGCGGATTTTCGC) and the “long” (ATCACGAGTGCTGCTCCAAAGCGGATTTTCGCAAATGGCC) cleavage sequence of M^{pro} (nsp4/nsp5 autocleavage site), flanked by N-terminal AgeI and a C-terminal HpaI restriction enzyme sites were ordered at Eurofins. The annealed oligos were ligated into a vector presenting N-terminally the fragment 1 (F[1]) and C-terminally the fragment 2 (F[2]) of the Renilla luciferase harboring the corresponding restriction enzyme sites (pcDNA3.1 backbone vector) as previously described (80, 81). Between each fragment and the cleavage sequence we inserted interjacent 10-aa linkers.

Preparation of lysates for western blotting

The cells expressing the M^{pro} cutting reporter and the purified proteins were prepared as described before. 50 µl of the protein suspension (5 µg of protein) were pre-incubated with inhibitor (10µM) or DMSO for 15 minutes at room temperature (R.T.). After that, 25 µl of cell suspension were added and the mixture was incubated for 3 hours at R.T. The cleavage reaction was stopped by the addition of 5x SDS-Loading buffer. Before the SDS-Page the samples were cooked at 95°C for 10 minutes. Western blot images were obtained using the ChemiDoc MP from Bio-Rad.

TCID₅₀ assay and dose responses

For initial dose response experiments, 5 × 10⁴ BHK-21 cells per well were seeded in 48-well plates one day before infection. Cells were infected in duplicates with a MOI of 0.05 of VSV-M^{pro} WT or VSV-O-M^{pro} and indicated concentrations of nirmatrelvir were added to the wells. After 48 hours, supernatants were collected and titrated with TCID₅₀. For quantification, the 50 % tissue culture infective dose (TCID₅₀) assay was performed as described previously (82). In short, 100 µl of serial dilutions of virus were added in octuplicates to 10³ BHK-21 cells seeded in a 96-well plate. Six days after infection, the TCID₅₀ were read out and titers were calculated according to the Kaerber method (83).

Construction of SARS-CoV-2 M^{pro} mutant replicons

To introduce M^{pro} mutations to the SARS-CoV-2 replicon, BAC recombineering was performed as previously described (84, 85). First, pSMARTBac-T7-SARS-CoV-2 replicon was transformed into SW102 E. coli host along with a fragment containing an ampicillin (Amp) cassette flanked by AscI restriction sites

and adjacent homologous sequences of M^{Pro}, resulting in the creation of pSMARTBac-T7-SARS-CoV-2-M^{Pro}-Amp. Next, pSMARTBAC-T7-SARS-CoV-2-replicon-Amp DNA was prepared using NucleoBond BAC 100 kit (Takara). A fragment containing the M^{Pro} mutation sequence was synthesized with a 30-bp overlap with the upstream and downstream regions of the M^{Pro} sequence for assembly. Following Ascl (NEB) digestion, the pSMARTBAC-T7-SARS-CoV-2-replicon-AMP DNA was ligated to the fragment containing the M^{Pro} mutations using Gibson Assembly kit. This resulted in the generation of pSMARTBAC-T7-SARS-CoV-2-M^{Pro} mutants-replicons.

Protease inhibition assay using SARS-CoV-2 replicon mutants

SARS-CoV-2 replicons RNA were prepared as previously described(45). Briefly, 1×10^6 Huh-7 cells were electroporated with 2 μ g of replicon RNA and subsequently plated into 96-well plates (Corning, cat. 3904) with a seeding density of 1×10^4 /well. Immediately after electroporation, a four-fold serial dilution of protease inhibitors (nirmatrelvir or ensitrelvir) was added to the wells. DMSO only was added to the control wells and served as solvent control. At 48 hours post-transfection, the reporter activities of in the cells treated with inhibitors were measured by quantifying the number of green cells in each well using an Acumen eX3 scanner and normalized it to the number of green cells in DMSO only wells (control wells).

Pearson's correlation coefficient calculation of M^{Pro}-Off and SARS-CoV-2 replicon data

A simple R script (R version 4.2.2) was used to calculate the Pearson's correlation coefficients for the whole dataset (M^{Pro}-Off with nirmatrelvir and ensitrelvir, and nirmatrelvir or ensitrelvir only).

IC₅₀ calculations

In this study, different assay systems were used to generate resistance data, namely VSV-based cellular assays with FluoroSpot read out, a luminescence-based cellular assay that employs recombinant M^{Pro}, as well as a peptide-MCA cleavage-based biochemical assay and SARS-CoV-2 mutant replicons assays. IC₅₀ calculations and statistical analysis for all assays were performed with GraphPad Prism 10. To calculate IC₅₀s, we used GraphPad's pre-set sigmoidal models: 4PL, X is concentration.

$$Y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + \left(\frac{\text{IC}_{50}}{X}\right)^{\text{HillSlope}}}$$

However, under some circumstances, slightly bell-shaped curves arise due to an initial signal in- or decrease (On and Off, respectively), followed by a signal peak and then a de- or increase. In the On assay, the decrease at the end occurs because of inhibitor toxicity at high concentrations. In the Off assay, the initial increase can be attributed to the fast death of cells in the absence of any inhibitor, whereas cells die slower at small doses, allowing for more cell division and thereby more "substrate" (cells and virus) generating dsRed. The bell-shaped curves can cause overfitting with the sigmoidal models, leading to extremely steep de- or increasing curves. Although in general, this does not lead to important changes in IC₅₀s, we still chose to compensate for such unrealistically steep slopes by constraining the HillSlope to 3 or lower (positive slope) for the gain-of-signal assay and minus 3 or higher (negative slope) for the loss-of signal assay.

Hardware overview (TTMD simulations)

Most of the molecular modeling operations, such as the structure preparation, the setup for molecular dynamics (MD) simulations, and trajectory analyses, were performed on a Linux workstation, equipped with a 20 cores Intel Core i9-9820X 3.3 GHz processor, running the Ubuntu 20.04 operating system. Molecular dynamics simulations were carried out on an in-house Linux GPU cluster composed of 20 NVIDIA devices ranging from GTX1080Ti to RTX3090.

Structure preparation

A computational study was conducted to rationalize the effect of SARS-CoV-2 M^{Pro} mutations on Nirmatrelvir resistance. In detail, a recently developed enhanced sampling molecular dynamics method named Thermal Titration Molecular Dynamics (TTMD) (58, 60) was exploited to compare the stability of the non-covalent complex between Nirmatrelvir and both the WT and mutated proteases, based on the

assumption that the formation of the reversible, non-covalent complex, is the rate-determining step that leads to the formation of the irreversible (or very slowly reversible), covalent complex (86). The top five mutants concerning their experimentally determined drug resistance were investigated.

The experimentally determined three-dimensional coordinates of the WT nirmatrelvir-M^{Pro} complex were retrieved from the Protein Data Bank (PDB) (87) and prepared for further calculations exploiting various tools provided within the Molecular Operating Environment (MOE) 2022.02 suite (88). Particularly, the complex deposited with accession code 7RFW (6) was used for describing the WT protease, while mutants were generated through the “protein builder” module of MOE.

The catalytically competent dimeric conformation of the protease was reconstructed by applying symmetric crystallographic transformations. Inconsistencies in the experimental structure were checked and fixed through the “structure preparation” tool. Missing hydrogen atoms were added according to the predominant tautomeric and protomeric state of each residue at pH = 7.4 through the “Protonate3D” module. Finally, each water molecule was removed before storing the prepared complex structure for subsequent calculations. In the end, a 1:1 protein-ligand non-covalent complex between nirmatrelvir and various M^{Pro} proteases was considered in the simulated systems.

System setup for MD simulations and Equilibration Protocol

Each complex obtained in the previous step was prepared for MD simulation making use of various tools provided within Visual Molecular Dynamics 1.9.2 (VMD) (89) and the AmberTools 22 (90, 91) suite. All parameters were attributed according to the ff14SB (92) force field, except for the ligand ones which were instead assigned through the general Amber force field (GAFF) (93). Partial charges for the ligand were calculated through the AM1-BCC method (94). Each one of the protein-ligand systems was solvated using the TIP3P (95) water model into a rectangular base box, ensuring a minimum 15 Å distance between the box border and the nearest protein or ligand atom. To neutralize the net charge of the system and to reach a salt concentration of 0.154 M, the proper number of sodium and chlorine ions was added. Finally, to remove clashes and unfavorable contacts, 500 steps of energy minimization through the conjugate gradient algorithm were performed.

Before the production phase, a two-step equilibration process was performed. First, 0.1 ns of simulation in the canonical ensemble were carried out, imposing harmonic positional restraints on both the protein and ligand atoms. Then, 0.5 ns of simulation in the isothermal-isobaric ensemble were performed, with harmonic positional restraints applied only on the protein backbone and ligand atoms. In each case, a 5 kcal mol⁻¹ Å⁻² force constant was imposed on each restrained atom for the whole duration of the simulation, and the temperature was maintained at a constant value of 300 K through a Langevin thermostat (96). In the second equilibration stage, the pressure was kept constant at 1 atm through a Monte Carlo barostat (97).

All MD simulations were conducted exploiting the ACEMD16 3.5 engine, which is based upon the open-source Python library for biomolecular simulations OpenMM (98). A 2 fs integration timestep was used, and the M-SHAKE algorithm was used to constrain the length of bonds involving hydrogen atoms. The particle-mesh Ewald method was used to compute electrostatic interactions, using cubic spline interpolation and a 1 Å grid spacing. A 9.0 Å cutoff was applied to the computation of Lennard-Jones interactions.

Thermal Titration Molecular Dynamics (TTMD) simulations

As thoroughly described in the work of Pavan et al. (99), Thermal Titration Molecular Dynamics (TTMD) is an enhanced sampling MD protocol originally developed for the qualitative estimation of protein-ligand unbinding kinetics. This approach allows the characterization in high detail of the binding mode and residence time of a ligand within the catalytic site of a given protein-ligand complex. It thereby enables the comparison between: 1. different ligands bound on the same protein (58); 2. different ligand conformations within the same binding site; 3. different mutants of the same protein in complex with the same ligand, as in the present work. The simulation consists of a series of classic molecular dynamics simulations performed at progressively increasing temperatures. Whether the ligand remains bound or not, is monitored through a scoring function based on protein-ligand interaction fingerprints (IFP_{CS}) (59).

The TTMD workflow relies on a series of short classic molecular dynamics simulations, defined as “TTMD-

steps", performed at progressively increasing temperatures. The temperature increase is used to augment the kinetic energy of the system, thus shortening the simulation time required to observe protein-ligand unbinding events compared to classic MD simulations. To monitor the progress of the simulation, the conservation of the native binding mode is evaluated through a protein-ligand interaction fingerprint (PLIF) based scoring function (59). The protocol described hereafter is implemented as a Python 3.10 code relying on the NumPy, MDAnalysis (100, 101), Open Drug Discovery Toolkit (ODDT) (102) and Scikit-learn packages. The code is released under a permissive MIT license and available free of charge at github.com/molecularmodelingsection/TTMD.

In detail, the user must define a "temperature ramp", i.e., the number, the temperature, and the length of each "TTMD-step". Consistently with previous works on the target (58, 60), in this case, the starting and end temperatures were set at 300K and 450K respectively, with a temperature increase between each "TTMD-step" of 10K and the duration of each simulation window set at 10 ns. The extension of the temperature ramp was determined based on the conservation of the native fold of the protein throughout the simulation, monitored through the backbone RMSD.

Structure preparation (analyses with Bioluminate and Osprey)

The PDB entries 7ALI (57), 8DZ0 (50), 8DZ2 (50), 7DVW (54) and 7TLL (51) were used to investigate the impact of mutations on apo WT, ensitrelvir bound WT, nirmatrelvir bound WT, NSP 5-6 bound H41A mutant and nirmatrelvir bound Omicron dimers. All structures were prepared with the default settings of the Protein Preparation Wizard in Maestro version 2022-3 (103) except that selenomethionines were converted to methionines and no water molecules were deleted. Please note, only the structure 7ALI (57) and 7DVW (54) includes residue 305 in both chains, 8DZ0 (50) and 8DZ2 (50) only includes residue 305 in chain B, and 7TLL (51) does not include residue 305 at all. Structure visualization and figure creation were performed in PyMol (104) version 2.5.0.

Stability prediction

The stability of all the mutations was calculated starting from the WT or Omicron M^{Pro} structures using the default settings of the `residue_scanning_backend.py` module in Bioluminate (46–49) release 2022-3. Mutations were introduced in combination in the two protomers.

Affinity prediction

For the affinity predictions in Osprey version 3.3 (53, 105–107) water molecules were not considered. Osprey calculates so called Log10 K* scores (106) which provide an estimation of binding affinity. The required yaml and frcmod input files were created as described in detail in the Guerin et al. STAR Protocol (108). For the dimer affinity calculations, chain B was considered as the ligand. The stability threshold was disabled, epsilon was defined as 0.03 and WT and mutant side chain conformations were calculated as continuously flexible. Osprey is an open source software and is available for free at <https://github.com/donaldlab/OSPNEY3>.

Supplementary material

Supplementary Material can be found at

[A comprehensive study of SARS-CoV-2 main protease \(M^{Pro}\) inhibitor-resistant mutants selected in a VSV-based system | PLOS Pathogens](#)

Author Contributions

Conceptualization: Seyed Arad Moghadasi, Theresia Dunzendorfer-Matt, Emmanuel Heilmann.

Data curation: Francesco Costacurta, Andrea Dodaro, David Bante, Helge Schöppke, Ju-Yi Peng, Bernhard Sprenger, Xi He, Seyed Arad Moghadasi, Lisa Maria Egger, Jakob Fleischmann, Matteo Pavan, Davide Bassani, Silvia Menin, Stefanie Rauch, Laura Krismer, Anna Sauerwein, Anne Heberle, Toni Rabensteiner, Theresia Dunzendorfer-Matt, Emmanuel Heilmann.

Formal analysis: Francesco Costacurta, Andrea Dodaro, David Bante, Helge Schopppe, Ju-Yi Peng, Bernhard Sprenger, Xi He, Seyed Arad Moghadas, Lisa Maria Egger, Jakob Fleischmann, Matteo Pavan, Davide Bassani, Silvia Menin, Laura Krismer, Toni Rabensteiner, Theresia Dunzendorfer-Matt, Emmanuel Heilmann.

Funding acquisition: Reuben S. Harris, Eduard Stefan, Rainer Schneider, Theresia Dunzendorfer-Matt, Dai Wang, Teresa Kaserer, Stefano Moro, Dorothee von Laer, Emmanuel Heilmann.

Investigation: Francesco Costacurta, Andrea Dodaro, David Bante, Helge Schopppe, Ju-Yi Peng, Bernhard Sprenger, Xi He, Lisa Maria Egger, Jakob Fleischmann, Matteo Pavan, Davide Bassani, Silvia Menin, Stefanie Rauch, Anna Sauerwein, Jose Ho, Theresia Dunzendorfer-Matt, Emmanuel Heilmann.

Methodology: Francesco Costacurta, Andrea Dodaro, David Bante, Helge Schopppe, Ju-Yi Peng, Bernhard Sprenger, Seyed Arad Moghadas, Lisa Maria Egger, Jakob Fleischmann, Davide Bassani, Silvia Menin, Anna Sauerwein, Anne Heberle, Toni Rabensteiner, Jose Ho, Theresia Dunzendorfer-Matt, Emmanuel Heilmann.

Project administration: Theresia Dunzendorfer-Matt, Dai Wang, Emmanuel Heilmann. Resources: Reuben S. Harris, Eduard Stefan, Rainer Schneider, Theresia Dunzendorfer-Matt, Andreas Naschberger, Dai Wang, Teresa Kaserer, Stefano Moro, Dorothee von Laer, Emmanuel Heilmann.

Software: Andrea Dodaro, David Bante, Matteo Pavan, Davide Bassani, Silvia Menin.

Supervision: Reuben S. Harris, Eduard Stefan, Rainer Schneider, Theresia Dunzendorfer-Matt, Andreas Naschberger, Dai Wang, Teresa Kaserer, Stefano Moro, Dorothee von Laer, Emmanuel Heilmann.

Validation: Francesco Costacurta, Andrea Dodaro, David Bante, Helge Schopppe, Bernhard Sprenger, Seyed Arad Moghadas, Jakob Fleischmann, Matteo Pavan, Davide Bassani, Silvia Menin, Stefanie Rauch, Laura Krismer, Theresia Dunzendorfer-Matt, Stefano Moro, Emmanuel Heilmann.

Visualization: Francesco Costacurta, Andrea Dodaro, David Bante, Helge Schopppe, Bernhard Sprenger, Seyed Arad Moghadas, Matteo Pavan, Davide Bassani, Emmanuel Heilmann.

Writing – original draft: Francesco Costacurta, Emmanuel Heilmann.

Writing – review & editing: Francesco Costacurta, Ju-Yi Peng, Theresia Dunzendorfer-Matt, Emmanuel Heilmann.

Affiliations

1 Institute of Virology, Medical University of Innsbruck, Innsbruck, Tyrol, Austria

2 Molecular Modeling Section (MMS), Department of Pharmaceutical and Pharmacological Sciences, University of Padua, Padova, Italy

3 Institute of Pharmacy / Pharmaceutical Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innsbruck, Tyrol, Austria

4 Department of Infectious Diseases and Vaccines Research, MRL, Merck & Co., Inc., Rahway, New Jersey, United States of America

5 Institute of Biochemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innsbruck, Austria

6 Department of Biochemistry, Molecular Biology and Biophysics, Institute for Molecular Virology, University of Minnesota, Minneapolis, Minnesota, United States of America

7 Institute of Molecular Biochemistry, Biocentre, Medical University of Innsbruck, Innsbruck, Austria

8 Institute of Molecular Biology, University of Innsbruck, Innsbruck, Tyrol, Austria

9 Tyrolean Cancer Research Institute (TKFI), Innsbruck, Tyrol, Austria

10 Bioinformatics Institute, Agency for Science Technology and Research, Singapore, Singapore

11 Department of Biochemistry and Structural Biology, University of Texas Health San Antonio, San Antonio, Texas, United States of America

12 Howard Hughes Medical Institute, University of Texas Health San Antonio, San Antonio, Texas, United States of America

13 Division of Biological and Environmental Sciences and Engineering (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

References

1. Mohammed, A. Nauman, P. Paul, S. Ganesan, K. H. Chen, S. M. S. Jalil, S. H. Jaouni, H. Kawas, W. A. Khan, A. L. Vattoth, Y. A. Al-Hashimi, A. Fares, R. Zeghlache, D. Zakaria, The efficacy and effectiveness of the COVID-19 vaccines in reducing infection, severity, hospitalization, and mortality: a systematic review. *Hum Vaccin Immunother* 18 (2022).
2. N. G. Davies, S. Abbott, R. C. Barnard, C. I. Jarvis, A. J. Kucharski, J. D. Munday, C. A. B. Pearson, T. W. Russell, D. C. Tully, A. D. Washburne, T. Wenseleers, A. Gimma, W. Waites, K. L. M. Wong, K. van Zandvoort, J. D. Silverman, K. Diaz-Ordaz, R. Keogh, R. M. Eggo, S. Funk, M. Jit, K. E. Atkins, W. J. Edmunds, Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* (1979) 372 (2021).
3. T. Kustin, N. Harel, U. Finkel, S. Perchik, S. Harari, M. Tahor, I. Caspi, R. Levy, M. Leshchinsky, S. Ken Dror, G. Bergerzon, H. Gadban, F. Gadban, E. Eliassian, O. Shimron, L. Saleh, H. Ben-Zvi, E. Keren Taraday, D. Amichay, A. Ben-Dor, D. Sagas, M. Strauss, Y. Shemer Avni, A. Huppert, E. Kepten, R. D. Balicer, D. Netzer, S. Ben-Shachar, A. Stern, Evidence for increased breakthrough rates of SARS-CoV-2 variants of concern in BNT162b2-mRNA-vaccinated individuals. *Nat Med* 27, 1379–1384 (2021).
4. Rössler, L. Riepler, D. Bante, D. von Laer, J. Kimpel, SARS-CoV-2 Omicron Variant Neutralization in Serum from Vaccinated and Convalescent Persons. 1–3 (2021).
5. M. Carabelli, T. P. Peacock, L. G. Thorne, W. T. Harvey, J. Hughes, T. I. de Silva, S. J. Peacock, W. S. Barclay, T. I. de Silva, G. J. Towers, D. L. Robertson, SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nat Rev Microbiol* 21, 162–177 (2023).
6. D. R. Owen, C. M. N. Allerton, A. S. Anderson, L. Aschenbrenner, M. Avery, S. Berritt, B. Boras, R. D. Cardin, A. Carlo, K. J. Coffman, A. Dantonio, L. Di, H. Eng, R. Ferre, K. S. Gajiwala, S. A. Gibson, S. E. Greasley, B. L. Hurst, E. P. Kadar, A. S. Kalgutkar, J. C. Lee, J. Lee, W. Liu, S. W. Mason, S. Noell, J. J. Novak, R. S. Obach, K. Ogilvie, N. C. Patel, M. Pettersson, D. K. Rai, M. R. Reese, M. F. Sammons, J. G. Sathish, R. S. P. Singh, C. M. Steppan, A. E. Stewart, J. B. Tuttle, L. Updyke, P. R. Verhoest, L. Wei, Q. Yang, Y. Zhu, An oral SARS-CoV-2 M pro inhibitor clinical candidate for the treatment of COVID-19. *Science* (1979) 374, 1586–1593 (2021).
7. K. H. Wong, I. C. H. Au, K. T. K. Lau, E. H. Y. Lau, B. J. Cowling, G. M. Leung, Real-world effectiveness of molnupiravir and nirmatrelvir plus ritonavir against mortality, hospitalisation, and in-hospital outcomes among community-dwelling, ambulatory patients with confirmed SARS-CoV-2 infection during the omicron wave in Hong Kong: a. *The Lancet* 400, 1213–1222 (2022).
8. Y. Unoh, S. Uehara, K. Nakahara, H. Nobori, Y. Yamatsu, S. Yamamoto, Y. Maruyama, Y. Taoda, K. Kasamatsu, T. Suto, K. Kouki, A. Nakahashi, S. Kawashima, T. Sanaki, S. Toba, K. Uemura, T. Mizutare, S. Ando, M. Sasaki, Y. Orba, H. Sawa, A. Sato, T. Sato, T. Kato, Y. Tachibana, Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CL Protease Inhibitor Clinical Candidate for Treating COVID-19. *J Med Chem*, doi: 10.1021/acs.jmedchem.2c00117 (2022).
9. X. Chen, X. Huang, Q. Ma, P. Kuzmič, B. Zhou, J. Xu, B. Liu, H. Jiang, W. Zhang, C. Yang, S. Wu, J. Huang, H. Li, C. Long, X. Zhao, H. Xu, Y. Sheng, Y. Guo, C. Niu, L. Xue, Y. Xu, J. Liu, T. Zhang, J. Spencer, W. Deng, S.-H. Chen, X. Xiong, Z. Yang, N. Zhong, Inhibition mechanism and antiviral activity of an α -ketoamide based SARS-CoV-2 main protease inhibitor. *bioRxiv*, 2023.03.09.531862 (2023).
10. W. Shang, W. Dai, C. Yao, L. Xu, X. Tao, H. Su, J. Li, X. Xie, Y. Xu, M. Hu, D. Xie, H. Jiang, L. Zhang, H. Liu, In vitro and in vivo evaluation of the main protease inhibitor FB2001 against SARS-CoV-2. *Antiviral Res* 208, 105450 (2022).
11. J. R. Spengler, S. R. Welch, J. Deval, B. G. Gentry, A. Brancale, K. Carter, J. Moffat, C. Meier, K. L. Seley-Radtke, L. M. Schang, "Meeting report: 35th International Conference on Antiviral Research in Seattle, Washington, USA – March 21–25, 2022" in *Antiviral Research* (Elsevier B.V., 2023; <https://doi.org/10.1016/j.antiviral.2022.105521>)vol. 211, p. 105521.
12. Pardes Biosciences Announces Top-line Results from Phase 2 Trial Evaluating Pomotrelvir for the Treatment of COVID-19. [Preprint] (2023).
13. China National Medical Products Administration (NMPA), NMPA conditionally approved the marketing of the combined packaging of Simnotrelvir Tablets/Ritonavir Tablets and Renmindevir Hydrobromide Tablets for COVID-19 treatment (2023); <https://www.nmpa.gov.cn/zhuant/yqjzxd/yqjzxd/20230129143814163.html>.
14. Y. Li, T. I. Zang, L. Xu, D. Leonard, K. Hoang, T. Greizer, S. Zhang, C. Foster, M. Huang, J. Kibel, J. Klaene, Y. S. Or, L.-J. Jiang, 1123. EDP-235, A Potent and Once-daily Oral Antiviral, Demonstrates Excellent Penetration into Macrophages and Monocytes, with the Potential for Mitigation of Cytokine Storm in High-Risk COVID-19 Patients. *Open Forum Infect Dis* 9, ofac492.962 (2022).
15. K. Padhi, T. Tripathi, Hotspot residues and resistance mutations in the nirmatrelvir-binding site of SARS-CoV-2 main protease: Design, identification, and correlation with globally circulating viral genomes. *Biochem Biophys Res Commun* 629, 54–60 (2022).
16. Y. Hu, E. M. Lewandowski, H. Tan, X. Zhang, R. T. Morgan, X. Zhang, L. M. C. Jacobs, S. G. Butler, M. V. Gongora, J. Choy, X. Deng, Y. Chen, J. Wang, Naturally Occurring Mutations of SARS-CoV-2 Main Protease Confer Drug Resistance to Nirmatrelvir. *ACS Cent Sci*, doi: 10.1021/acscentsci.3c00538 (2023).
17. J. Clayton, V. M. de Oliveira, M. F. Ibrahim, X. Sun, P. Mahinthaichan, M. Shen, R. Hilgenfeld, J. Shen, Integrative Approach to Dissect the Drug Resistance Mechanism of the H172Y Mutation of SARS-CoV-2 Main Protease. *J Chem Inf Model*, doi: 10.1021/acs.jcim.3c00344 (2023).
18. S. Iketani, H. Mohri, B. Culbertson, S. J. Hong, Y. Duan, M. I. Luck, M. K. Annavajhala, Y. Guo, Z. Sheng, A.-C. C. Uhlemann, S. P. Goff, Y. Sabo, H. Yang, A. Chavez, D. D. Ho, Multiple pathways for SARS-CoV-2 resistance to nirmatrelvir. *Nature* 613, 558–564 (2022).
19. Jochmans, C. Liu, K. Donckers, A. Stoycheva, S. Boland, S. K. Stevens, C. De Vita, B. Vanmechelen, P. Maes, B. Trüeb, N. Ebert, V. Thiel, S. De Jonghe, L. Vangeel, D. Bardiot, A. Jekle, L. M. Blatt, L. Beigelman, J. A. Symons, P. Raboisson,

- P. Chaltin, A. Marchand, J. Neyts, J. Deval, K. Vandyck, The Substitutions L50F, E166A, and L167F in SARS-CoV-2 3CLpro Are Selected by a Protease Inhibitor In Vitro and Confer Resistance To Nirmatrelvir. *mBio* 14 (2023).
20. S. A. Moghadas, E. Heilmann, A. M. Khalil, C. Nnabuike, F. L. Kearns, C. Ye, S. N. Moraes, F. Costacurta, M. A. Esler, H. Aihara, D. von Laer, L. Martinez-Sobrido, T. Palzkill, R. E. Amaro, R. S. Harris, "Transmissible SARS-CoV-2 variants with resistance to clinical protease inhibitors" (2023); <https://doi.org/10.1126/sciadv.ade8778>.
 21. Y. Zhou, K. A. Gammeltoft, L. A. Ryberg, L. V. Pham, H. D. Tjørnelund, A. Binderup, C. R. D. Hernandez, C. Fernandez-Antunez, A. Offersgaard, U. Fahnøe, G. H. J. Peters, S. Ramirez, J. Bukh, J. M. Gottwein, Nirmatrelvir-resistant SARS-CoV-2 variants with high fitness in an infectious cell culture system. *Sci Adv* 8 (2022).
 22. M. Shaqra, S. N. Zvornicanin, Q. Y. J. Huang, G. J. Lockbaum, M. Knapp, L. Tandeske, D. T. Bakan, J. Flynn, D. N. A. Bolon, S. Moquin, D. Dovala, N. Kurt Yilmaz, C. A. Schiffer, Defining the substrate envelope of SARS-CoV-2 main protease to predict and avoid drug resistance. *Nat Commun* 13, 1–11 (2022).
 23. L. Parigger, A. Krassnigg, T. Schopper, A. Singh, K. Tappler, K. Köchl, M. Hetmann, K. Gruber, G. Steinkellner, C. C. Gruber, Recent changes in the mutational dynamics of the SARS-CoV-2 main protease substantiate the danger of emerging resistance to antiviral drugs. *Front Med (Lausanne)* 9, 1–12 (2022).
 24. K. S. Yang, S. Z. Leeuwon, S. Xu, W. R. Liu, Evolutionary and Structural Insights about Potential SARS-CoV-2 Evasion of Nirmatrelvir. *J Med Chem* 65, 8686–8698 (2022).
 25. J. M. Flynn, N. Samant, G. Schneider-Nachum, D. T. Barkan, N. K. Yilmaz, C. A. Schiffer, S. A. Moquin, D. Dovala, D. N. A. Bolon, Comprehensive fitness landscape of SARS-CoV-2 Mpro reveals insights into viral resistance mechanisms. *Elife* 11, 1–27 (2022).
 26. J. M. Flynn, Q. Y. J. Huang, S. N. Zvornicanin, G. G. Schneider-Nachum, A. M. Shaqra, N. K. Yilmaz, S. A. Moquin, D. Dovala, C. A. Schiffer, D. N. A. Bolon, Systematic Analyses of the Resistance Potential of Drugs Targeting SARS-CoV-2 Main Protease. *ACS Infect Dis*, doi: 10.1021/acscinfecdis.3c00125 (2023).
 27. J. Ou, E. M. Lewandowski, Y. Hu, A. A. Lipinski, A. Aljasser, M. Colon-Ascanio, R. T. Morgan, L. M. C. Jacobs, X. Zhang, M. J. Bikowitz, P. R. Langlais, H. Tan, J. Wang, Y. Chen, J. S. Choy, A yeast-based system to study SARS-CoV-2 Mpro structure and to identify nirmatrelvir resistant mutations. *PLoS Pathog* 19, e1011592 (2023).
 28. S. Iketani, S. J. Hong, J. Sheng, F. Bahari, B. Culbertson, F. F. Atanaki, A. K. Aditham, A. F. Kratz, M. I. Luck, R. Tian, S. P. Goff, H. Montazeri, Y. Sabo, D. D. Ho, A. Chavez, Functional map of SARS-CoV-2 3CL protease reveals tolerant and immutable sites. *Cell Host Microbe* 30, 1354–1362.e6 (2022).
 29. S. Lan, G. Neilsen, R. L. Slack, W. A. Cantara, A. E. Castaner, Z. C. Lorson, N. Lulkin, H. Zhang, J. Lee, M. E. Cilento, P. R. Tedbury, S. G. Sarafianos, Nirmatrelvir Resistance in SARS-CoV-2 Omicron_BA.1 and WA1 Replicons and Escape Strategies. *bioRxiv*, 2022.12.31.522389 (2023).
 30. L. Vangeel, W. Chiu, S. De Jonghe, P. Maes, B. Slechten, J. Raymenants, E. André, P. Leyssen, J. Neyts, D. Jochmans, Remdesivir, Molnupiravir and Nirmatrelvir remain active against SARS-CoV-2 Omicron and other variants of concern. *Antiviral Res* 198, 10–12 (2022).
 31. P. Li, Y. Wang, M. Lavrijsen, M. M. Lamers, A. C. de Vries, R. J. Rottier, M. J. Bruno, M. P. Peppelenbosch, B. L. Haagmans, Q. Pan, SARS-CoV-2 Omicron variant is highly sensitive to molnupiravir, nirmatrelvir, and the combination. *Cell Res* 32, 322–324 (2022).
 32. S. Kawashima, Y. Matsui, T. Adachi, Y. Morikawa, K. Inoue, S. Takebayashi, H. Nobori, M. Rokushima, Y. Tachibana, T. Kato, Ensitrelvir is effective against SARS-CoV-2 3CL protease mutants circulating globally. *Biochem Biophys Res Commun* 645, 132–136 (2023).
 33. S. Ullrich, K. B. Ekanayake, G. Otting, C. Nitsche, Main protease mutants of SARS-CoV-2 variants remain susceptible to nirmatrelvir. *Bioorg Med Chem Lett* 62, 128629 (2022).
 34. M. D. Sacco, Y. Hu, M. V. Gongora, F. Meilleur, M. T. Kemp, X. Zhang, J. Wang, Y. Chen, The P132H mutation in the main protease of Omicron SARS-CoV-2 decreases thermal stability without compromising catalysis or small-molecule drug inhibition. *Cell Res* 32, 498–500 (2022).
 35. Heilmann, F. Costacurta, S. A. Moghadas, C. Ye, M. Pavan, D. Bassani, A. Volland, C. Ascher, A. K. H. Weiss, D. Bante, R. S. Harris, S. Moro, B. Rupp, L. Martinez-Sobrido, D. von Laer, SARS-CoV-2 3CLpro mutations selected in a VSV-based system confer resistance to nirmatrelvir, ensitrelvir, and GC376. *Sci Transl Med* 15, 1–19 (2023).
 36. J. W. Drake, J. J. Holland, Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A* 96, 13910–13913 (1999).
A. Steinhauer, E. Domingo, J. J. Holland, Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* 122, 281–288 (1992).
 37. Y. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data – from vision to reality. [Preprint] (2017). <https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494>.
 38. S. Elbe, G. Buckland-Merrett, Data, disease and diplomacy: GISAID's innovative contribution to global health. *Global Challenges* 1, 33–46 (2017).
 39. S. Khare, C. Gurry, L. Freitas, M. B. Schultz, G. Bach, A. Diallo, N. Akite, J. Ho, R. T. C. Lee, W. Yeo, G. C. C. Team, S. Maurer-Stroh, GISAID's Role in Pandemic Response. *China CDC Wkly* 3, 1049–1051 (2021).
 40. K. Anand, J. Ziebuhr, P. Wadhvani, J. R. Mesters, R. Hilgenfeld, (3CL pro) Structure: Basis for Design of Anti-SARS Drugs. *Science* (1979) 1763, 1763–1768 (2014).
 41. S. Ullrich, C. Nitsche, The SARS-CoV-2 main protease as drug target. *Bioorg Med Chem Lett* 30, 127377 (2020).
 42. S. E. Witko, C. S. Kotash, R. M. Nowak, J. E. Johnson, L. A. C. Boutlier, K. J. Melville, S. G. Heron, D. K. Clarke, A. S. Abramovitz, R. M. Hendry, M. S. Sidhu, S. A. Udem, C. L. Parks, An efficient helper-virus-free method for rescue of recombinant paramyxoviruses and rhadoviruses from a cell line suitable for vaccine development. *J Virol Methods* 135, 91–101 (2006).
 43. Heilmann, F. Costacurta, S. Geley, S. A. Mogadashi, A. Volland, B. Rupp, R. S. Harris, D. von Laer, A VSV-based assay quantifies coronavirus Mpro/3CLpro/NSp5 main protease activity and chemical inhibition. *Commun Biol* 5, 1–12 (2022).

44. X. He, S. Quan, M. Xu, S. Rodriguez, S. L. Goh, J. Wei, A. Fridman, K. A. Koeplinger, S. S. Carroll, J. A. Grobler, A. S. Espeseth, D. B. Olsen, D. J. Hazuda, D. Wang, Generation of SARS-CoV-2 reporter replicon for high-throughput antiviral screening and testing. *Proc Natl Acad Sci U S A* 118 (2021).
45. K. Zhu, T. Day, D. Warshaviak, C. Murrett, R. Friesner, D. Pearlman, Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. *Proteins: Structure, Function and Bioinformatics* 82, 1646–1655 (2014).
46. N. K. Salam, M. Adzhigirey, W. Sherman, D. A. Pearlman, D. Thirumalai, Structure-based approach to the prediction of disulfide bonds in proteins. *Protein Engineering, Design and Selection* 27, 365–374 (2014).
47. Beard, A. Cholleti, D. Pearlman, W. Sherman, K. A. Loving, Applying physics-based scoring to calculate free energies of binding for single amino acid mutations in protein-protein complexes. *PLoS One* 8, 1–11 (2013).
48. Schrödinger Release 2023-2: BioLuminate, Schrödinger, LLC, New York, NY, 2023.
49. D. Noske, E. de Souza Silva, M. O. de Godoy, I. Dolci, R. S. Fernandes, R. V. C. Guido, P. Sjö, G. Oliva, A. S. Godoy, Structural basis of nirmatrelvir and ensitrelvir activity against naturally occurring polymorphisms of the SARS-CoV-2 main protease. *Journal of Biological Chemistry* 299 (2023).
50. S. E. Greasley, S. Noell, O. Plotnikova, R. A. Ferre, W. Liu, B. Bolanos, K. Fennell, J. Nicki, T. Craig, Y. Zhu, A. E. Stewart, C. M. Steppan, Structural basis for the in vitro efficacy of nirmatrelvir against SARS-CoV-2 variants. *Journal of Biological Chemistry* 298, 1–7 (2022).
51. Y. Duan, H. Zhou, X. Liu, S. Iketani, M. Lin, X. Zhang, Q. Bian, H. Wang, H. Sun, S. J. Hong, B. Culbertson, H. Mohri, M. I. Luck, Y. Zhu, X. Liu, Y. Lu, X. Yang, K. Yang, Y. Sabo, A. Chavez, S. P. Goff, Z. Rao, D. D. Ho, H. Yang, Molecular mechanisms of SARS-CoV-2 resistance to nirmatrelvir. *Nature* 622, 376–382 (2023).
52. M. A. Hallen, J. W. Martin, A. Ojewole, J. D. Jou, A. U. Lowegard, M. S. Frenkel, P. Gainza, H. M. Nisonoff, A. Mukund, S. Wang, G. T. Holt, D. Zhou, E. Dowd, B. R. Donald, OSPREY 3.0: Open-source protein redesign for you, with powerful new features. *J Comput Chem* 39, 2494–2507 (2018).
53. Y. Zhao, Y. Zhu, X. Liu, Z. Jin, Y. Duan, Q. Zhang, C. Wu, L. Feng, X. Du, J. Zhao, M. Shao, B. Zhang, X. Yang, L. Wu, X. Ji, L. W. Guddat, K. Yang, Z. Rao, H. Yang, Structural basis for replicase polyprotein cleavage and substrate specificity of main protease from SARS-CoV-2. *Proc Natl Acad Sci U S A* 119, e2117142119 (2022).
54. G. D. Noske, E. de Souza Silva, M. O. de Godoy, I. Dolci, R. S. Fernandes, R. V. C. Guido, P. Sjö, G. Oliva, A. S. Godoy, Structural basis of nirmatrelvir and ensitrelvir activity against naturally occurring polymorphisms of the SARS-CoV-2 main protease. *Journal of Biological Chemistry* 299, 103004 (2023).
55. E. Costanzi, M. Kuzikov, F. Esposito, S. Albani, N. Demitri, B. Giabbai, M. Camasta, E. Tramontano, G. Rossetti, A. Zaliani, P. Storici, Structural and biochemical analysis of the dual inhibition of MG-132 against SARS-CoV-2 main protease (Mpro/3CLpro) and human cathepsin-L. *Int J Mol Sci* 22 (2021).
56. E. Costanzi, M. Kuzikov, F. Esposito, S. Albani, N. Demitri, B. Giabbai, M. Camasta, E. Tramontano, G. Rossetti, A. Zaliani, P. Storici, Structural and biochemical analysis of the dual inhibition of MG-132 against SARS-CoV-2 main protease (Mpro/3CLpro) and human cathepsin-L. *Int J Mol Sci* 22 (2021).
57. M. Pavan, S. Menin, D. Bassani, M. Sturlese, S. Moro, Qualitative Estimation of Protein-Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J Chem Inf Model* 62, 5715–5728 (2022).
58. M. Pavan, S. Menin, D. Bassani, M. Sturlese, S. Moro, Implementing a Scoring Function Based on Interaction Fingerprint for Autogrow4: Protein Kinase CK1δ as a Case Study. *Front Mol Biosci* 9, 1–17 (2022).
59. S. Menin, M. Pavan, V. Salmaso, M. Sturlese, S. Moro, Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement. *Int J Mol Sci* 24 (2023).
60. E. Fornasier, M. L. Macchia, G. Giachin, A. Sosic, M. Pavan, M. Sturlese, C. Salata, S. Moro, B. Gatto, M. Bellanda, R. Battistutta, A new inactive conformation of SARS-CoV-2 main protease. *Acta Crystallogr D Struct Biol* 78, 363–378 (2022).
61. D. Bloom, R. A. Neher, Fitness effects of mutations to SARS-CoV-2 proteins. *Virus Evol*, doi: doi.org/10.1093/ve/vead055 (2023).
62. S. Albani, E. Costanzi, G. L. Hoang, M. Frings, N. Ansari, N. Demitri, T. Toan, Unexpected single ligand occupancy and negative cooperativity in the SARS-CoV-2 Main protease. *ChemRxiv*, doi: 10.26434/chemrxiv-2023-n1x1n D O I: 10.26434/chemrxiv-2023-n1x1n [opens in a new tab] (2023).
63. S. Chen, F. Jonas, C. Shen, R. Higenfeld, Liberation of SARS-CoV main protease from the viral polyprotein: N-terminal autocleavage does not depend on the mature dimerization mode. *Protein Cell* 1, 59–74 (2010).
64. F. Schmidt, Y. Weisblum, M. Rutkowska, D. Poston, J. Da Silva, F. Zhang, E. Bednarski, A. Cho, D. J. Schaefer-Babajew, C. Gaebler, M. Caskey, M. C. Nussenzweig, T. Hatziioannou, P. D. Bieniasz, High genetic barrier to SARS-CoV-2 polyclonal neutralizing antibody escape. *Nature* 600, 512–516 (2021).
65. Y. Yahalom-Ronen, N. Erez, M. Fisher, H. Tamir, B. Politi, H. Achdout, S. Melamed, I. Glinert, S. Weiss, I. Cohen-Gihon, O. Israeli, M. Izak, M. Mandelboim, Y. Caraco, N. Madar-Balakirski, A. Mechaly, E. Shinar, R. Zichel, D. Cohen, A. Beth-Din, A. Zvi, H. Marcus, T. Israely, N. Paran, Neutralization of SARS-CoV-2 Variants by rVSV-ΔG-Spike-Elicited Human Sera. [Preprint] (2022). <https://doi.org/10.3390/vaccines10020291>.
66. Z. Liu, L. A. VanBlargan, L. M. Bloyet, P. W. Rothlauf, R. E. Chen, S. Stumpf, H. Zhao, J. M. Errico, E. S. Theel, M. J. Liebeskind, B. Alford, W. J. Buchser, A. H. Ellebedy, D. H. Fremont, M. S. Diamond, S. P. J. Whelan, Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe* 29, 477–488.e4 (2021).
67. T. W. Geisbert, H. Feldmann, Recombinant vesicular stomatitis virus-based vaccines against Ebola and marburg virus infections. *Journal of Infectious Diseases* 204, 1075–1081 (2011).

68. Muik, I. Kneiske, M. Werbizki, D. Wilflingseder, T. Giroglou, O. Ebert, A. Kraft, U. Dietrich, G. Zimmer, S. Momma, D. von Laer, Pseudotyping Vesicular Stomatitis Virus with Lymphocytic Choriomeningitis Virus Glycoproteins Enhances Infectivity for Glioma Cells and Minimizes Neurotropism. *J Virol* 85, 5679–5684 (2011).
69. M. Romano, A. Ruggiero, F. Squeglia, G. Maga, R. Berisio, A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA Synthesis, Proofreading and Final Capping. *NLM (Medline) [Preprint]* (2020). <https://doi.org/10.3390/cells9051267>.
70. Z. Sun, L. Wang, X. Li, C. Fan, J. Xu, Z. Shi, H. Qiao, Z. Lan, X. Zhang, L. Li, X. Zhou, Y. Geng, An extended conformation of SARS-CoV-2 main protease reveals allosteric targets. *Proc Natl Acad Sci U S A* 119, 1–9 (2022).
71. M. Westberg, Y. Su, X. Zou, P. Huang, A. Rustagi, J. Garhyan, Design of SARS-CoV-2 protease inhibitors with improved affinity and reduced sensitivity to mutations. doi: doi.org/10.1101/2023.07.19.549739 (2023).
72. D. Panda, P. X. Dinh, L. K. Beura, A. K. Pattnaik, Induction of Interferon and Interferon Signaling Pathways by Replication of Defective Interfering Particle RNA in Cells Constitutively Expressing Vesicular Stomatitis Virus Replication Proteins. *J Virol* 84, 4826–4831 (2010).
73. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 11, 1–10 (2016).
74. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100 (2018).
75. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, Twelve years of SAMtools and BCFtools. *Gigascience* 10, 1–4 (2021).
76. Wilm, P. P. K. Aw, D. Bertrand, G. H. T. Yeo, S. H. Ong, C. H. Wong, C. C. Khor, R. Petric, M. L. Hibberd, N. Nagarajan, LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res* 40, 11189–11201 (2012).
77. Y. Turakhia, B. Thornlow, A. S. Hinrichs, N. De Maio, L. Gozashti, R. Lanfear, D. Haussler, R. Corbett-Detig, Ultrafast Sample placement on Existing tRees (USHER) enables real-time phylogenetics for the SARS-CoV-2 pandemic. *Nat Genet* 53, 809–816 (2021).
78. E. H. Abdelkader, G. Otting, NT*-HRV3CP: An optimized construct of human rhinovirus 14 3C protease for high-yield expression and fast affinity-tag cleavage. *J Biotechnol* 325, 145–151 (2021).
79. R. Röck, J. E. Mayrhofer, O. Torres-Quesada, F. Enzler, A. Raffener, P. Raffener, A. Feichtner, R. G. Huber, S. Koide, S. S. Taylor, J. Troppmair, E. Stefan, BRAF inhibitors promote intermediate BRAF(V600E) conformations and binary interactions with activated RAS. *Sci Adv* 5, 14–16 (2019).
80. E. Mayrhofer, F. Enzler, A. Feichtner, R. Röck, J. Fleischmann, A. Raffener, P. Tschaikner, E. Ogris, R. G. Huber, M. Hartl, R. Schneider, J. Troppmair, O. Torres-Quesada, E. Stefan, Mutation-oriented profiling of autoinhibitory kinase conformations predicts RAF inhibitor efficacies. *Proc Natl Acad Sci U S A* 117, 31105–31113 (2020).
81. C. Dold, C. Rodríguez Urbíola, G. Wollmann, L. Egerer, A. Muik, L. Bellmann, H. Fiegl, C. Marth, J. Kimpel, D. Von Laer, Application of interferon modulators to overcome partial resistance of human ovarian cancers to VSV-GP oncolytic viral therapy. *Mol Ther Oncolytics* 3, 16021 (2016).
82. G. Kärber, Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche. *Naunyn Schmiedebergs Arch Exp Pathol Pharmacol* 162, 480–483 (1931).
83. S. Warming, N. Costantino, D. L. Court, N. A. Jenkins, N. G. Copeland, Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res* 33, e36–e36 (2005).
84. X. He, S. Quan, M. Xu, S. Rodriguez, S. L. Goh, J. Wei, A. Fridman, K. A. Koeplinger, S. S. Carroll, J. A. Grobler, A. S. Espeseth, D. B. Olsen, D. J. Hazuda, D. Wang, Generation of SARS-CoV-2 reporter replicon for high-throughput antiviral screening and testing. *Proceedings of the National Academy of Sciences* 118, e2025866118 (2021).
85. Singh, R. C. Pette, T. A. Baillie, A. Whitty, The resurgence of covalent drugs. *Nat Rev Drug Discov* 10, 307–317 (2011).
86. H. M. Berman, The Protein Data Bank. *Nucleic Acids Res* 28, 235–242 (2000).
87. Molecular Operating Environment (MOE), Molecular Operating Environment (MOE), 2019.01; Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite. *Scientific Computing & Instrumentation*, 32 (2009).
88. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual Molecular Dynamics. *J Mol Graph* 14, 33–38 (1996).
89. D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, The Amber biomolecular simulation programs. *J Comput Chem* 26, 1668–1688 (2005).
90. D. A. Case, H. M. Aktulga, K. Belfon, I. Y. Ben-Shalom, J. T. Berryman, S. R. Brozell, D. S. Cerutti, T. E. I. Cheatham, G. A. Cisneros, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, M. C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K. M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K. A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rhanamoun, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C. L. Simmerling, N. R. Skrynnikov, J. Smith, J. Swails, R. C. Walker, H. Wang, J. Wang, H. Wei, R. M. Wolf, X. Wu, Y. Xiong, Y. Xue, D. M. York, S. Zhao, P. A. Kollman, Amber 2022, University of California, San Francisco. (2022).
91. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* 11, 3696–3713 (2015).
92. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general Amber force field. *J Comput Chem* 25, 1157–1174 (2004).
93. Jakalian, D. B. Jack, C. I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* 23, 1623–1641 (2002).
94. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79, 926–935 (1983).

95. R. L. Davidchack, R. Handel, M. V. Tretyakov, Langevin thermostat for rigid body dynamics. *Journal of Chemical Physics* 130 (2009).
96. R. Faller, J. J. De Pablo, Constant pressure hybrid molecular dynamics-Monte Carlo simulations. *Journal of Chemical Physics* 116, 55–59 (2002).
97. P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L. P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, V. S. Pande, OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol* 13, 1–17 (2017).
98. Pavan, S. Menin, D. Bassani, M. Sturlese, S. Moro, Qualitative Estimation of Protein-Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J Chem Inf Model* 62, 5715–5728 (2022).
99. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, MDAAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* 32, 2319–2327 (2011).
100. R. Gowers, M. Linke, J. Barnoud, T. Reddy, M. Melo, S. Seyler, J. Domański, D. Dotson, S. Buchoux, I. Kenney, O. Beckstein, MDAAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proceedings of the 15th Python in Science Conference*, 98–105 (2016).
101. Wójcikowski, P. Zielenkiewicz, P. Siedlecki, Open Drug Discovery Toolkit (ODDT): A new open-source player in the drug discovery field. *J Cheminform* 7, 1–6 (2015).
102. G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* 27, 221–234 (2013).
103. The PyMOL Molecular Graphics System, Version 2.5.0 Schrödinger, LLC.
104. Georgiev, R. H. Lilien, B. R. Donald, The minimized dead-end elimination criterion and its application to protein redesign in a hybrid scoring and search algorithm for computing partition functions over molecular ensembles. *J Comput Chem* 29, 1527–1542 (2008).
105. R. H. Lilien, B. W. Stevens, A. C. Anderson, B. R. Donald, A Novel Ensemble-Based Scoring and Search Algorithm for Protein Redesign and Its Application to Modify the Substrate Specificity of the Gramicidin Synthetase A Phenylalanine Adenylation Enzyme. *Journal of Computational Biology* 12, 740–761 (2005).
106. E. Roberts, P. R. Cushing, P. Boisguerin, D. R. Madden, B. R. Donald, Computational design of a PDZ domain peptide inhibitor that rescues CFTR activity. *PLoS Comput Biol* 8 (2012).
107. Guerin, A. Feichtner, E. Stefan, T. Kaserer, B. R. Donald, Resistor: An algorithm for predicting resistance mutations via Pareto optimization over multistate protein design and mutational signatures. *Cell Syst* 13, 830–843.e3 (2022).

Thermal Titration Molecular Dynamics (TTMD): Shedding Light On The Stability Of RNA-Small Molecule Complexes

Andrea Dodaro, Matteo Pavan, Silvia Menin, Veronica Salmaso, Mattia Sturlese and Stefano Moro*

Dodaro, A.; Pavan, M.; Menin, S.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Shedding Light on the Stability of RNA-Small Molecule Complexes. *Front. Mol. Biosci.* 2023, 10. <https://doi.org/10.3389/fmolb.2023.1294543>.

Abstract

Ribonucleic acids are gradually becoming relevant players among putative drug targets, thanks to the increasing amount of structural data exploitable for the rational design of selective and potent binders that can modulate their activity. Mainly, this information allows employing different computational techniques for predicting how well would a ribonucleic-targeting agent fit within the active site of its target macromolecule. Due to some intrinsic peculiarities of complexes involving nucleic acids, such as structural plasticity, surface charge distribution, and solvent-mediated interactions, the application of routinely adopted methodologies like molecular docking is challenged by scoring inaccuracies, while more physically rigorous methods such as molecular dynamics require long simulation times which hamper their conformational sampling capabilities. In the present work, we present the first application of Thermal Titration Molecular Dynamics (TTMD), a recently developed method for the qualitative estimation of unbinding kinetics, to characterize RNA-ligand complexes. In this article, we explored its applicability as a post-docking refinement tool on RNA in complex with small molecules, highlighting the capability of this method to identify the native binding mode among a set of decoys across various pharmaceutically relevant test cases.

Introduction

Following the “Central Dogma” of molecular biology, RNA has been historically perceived as a bridging element between DNA genetic information and protein biosynthesis ^{1,2}. However, discoveries in the last decades completely shifted this paradigm, showcasing the strikingly complex genetic and catalytic functions of RNA ^{3,4}. Therefore, this biomolecule is no longer considered just a carrier of information, instead, it is thought to be one of the most pluripotent chemical entities in molecular biology ³.

Considering that only a small percentage (3%) of the genome is translated into proteins ⁵ non-coding RNAs (ncRNAs) are under the spotlight as they are involved in epigenetics ^{4,6}, playing a pivotal role in the etiopathogenesis of both cancer and neurodegenerative diseases ⁶.

The identification of RNAs as possible therapeutic targets and the increased availability of experimentally solved three-dimensional structures of RNA complexes paves the way for the application of structure-based drug design (SBDD) techniques to characterize the interaction between small molecules and RNA, allowing the possibility of rationally discover new hits and steer their development into mature leads ^{2,4}. Although being an appealing perspective, applying routinely adopted molecular modeling protocols to the study of RNA systems is not trivial. Historically, these techniques have been optimized to study the recognition process between small organic molecules as ligands and proteins as receptors ^{7–9}. Nevertheless,

the obvious structural differences between proteins and RNAs, such as the peculiar surface charge properties portrayed by the polyanionic phosphate backbone, the ions' role in the structural stability and folding of RNA, the role of the solvent in mediating structural stability and forming bridged interactions, other than the intrinsic structural flexibility of ribonucleic acids, limited so far the possibility to repurpose these methodologies to the study of RNA complexes ¹.

One of the most successful approaches in structure-based drug design is molecular docking, a fast and reliable protocol that provides possible binding hypotheses generated through a conformational search algorithm and ranked through a scoring function ^{10,11}. Albeit being relatively fast to perform and computationally effortless when compared to other techniques such as molecular dynamics, pose scoring can suffer from inaccuracies ¹². Moreover, docking neglects the role of the solvent and the flexibility of the RNA which cannot explore its conformational landscape ^{1,2}.

To overcome the limitations of a docking approach, molecular dynamics simulations can be theoretically exploited to characterize RNA-ligand interactions with a more physically rigorous methodology. However, the computational effort required to spontaneously sample both binding and unbinding processes during unbiased MD simulations makes this technique incompatible with the timings of modern-day drug discovery campaigns ¹³.

Enhanced sampling algorithms are therefore required to drastically cut the simulation time to observe the desired phenomenon. These protocols rely on energetically biasing the system to increase the frequency of observation of the desired event and have been successfully applied to the study of unbinding processes in protein-ligand complexes ¹⁴. Among various approaches published in the literature, Thermal Titration Molecular Dynamics (TTMD) is a recently developed pipeline that has been applied to rank protein-ligand complexes based on the persistence of the native intermolecular interaction, allowing both to rank different binding hypotheses for the same ligand upon a defined receptor binding site ¹⁵ or to rank different ligands based on their dissociation rate ¹⁶. Specifically, TTMD operates through a series of short classic molecular dynamics simulations performed at increasing temperatures while monitoring the persistence of native receptor-ligand interactions through an interaction fingerprints-based scoring function ¹⁷.

Due to the encouraging success observed in the characterization of protein-ligand complex stability, in the present work we aim at extending the applicability domain of the technique to the world of RNA-ligand complexes. Particularly, we tried to understand if TTMD can be used as a post-docking filter to steer the identification of a native-like binding pose for small organic ligands onto RNA receptor binding sites.

Materials and Methods

Hardware overview

Modeling tasks such as the structure preparation of RNA targets and respective ligands, docking, system setup for molecular dynamics (MD) simulations and subsequent trajectory analysis were carried out on a Linux workstation running Ubuntu 20.04 as its operating system equipped with a 20 cores Intel Core i9-9820X 3.3 Ghz processor. MD simulations were performed exploiting an in-house GPU cluster composed of 20 NVIDIA devices ranging from GTX1080Ti to RTX3090.

Structure preparation

The three-dimensional structures of the targets presented in this work were retrieved from the Protein Data Bank (PDB ¹⁸), and successfully processed through different built-in modules of the Molecular Operating Environment ¹⁹. At first, each inconsistency between the primary sequence and the tertiary structure was fixed through the "Structure preparation" module. Then, the "Protonate3D" tool was exploited for adding missing hydrogens according to the most probable tautomeric and protonation state of titratable groups at pH = 7.4. Finally, every non-RNA and non-ligand residue was removed, except for K⁺ ions in the G-quadruplex portion of 5BJO complex, which were retained as they play a pivotal role in the stabilization of these non-canonical structures that shape the binding site at the monomer-monomer interface ²⁰.

To avoid any bias in the docking calculation that may favor a crystal-like pose, the ligands have been prepared starting from their SMILES, exploiting Open Babel ²¹ and various QUACPAC OpenEye tools ²².

Open Babel was used to generate two-dimensional and three-dimensional coordinates, while “Tautomers” and “Fixpka” assigned the correct protomeric and tautomeric state at pH 7.00. Afterwards “Molcharge” set the partial charges according to the MMFF94 force field. Both the RNA and ligand were stored for docking calculations.

Docking calculation

A self-docking calculation was conducted through the Protein-Ligand ANT System (PLANTS^{23–25} program, one of the state-of-the-art docking tools for nucleic acids-ligand docking²⁶. The binding site was defined as a sphere of radius 10.5 Å centered around the center of mass of the crystal ligand in the experimentally solved complex. The first 5 poses according to the ChemPLP scoring function were stored for refinement through the Thermal Titration Molecular Dynamics protocol.

System setup for MD Simulations and equilibration protocol

Several packages of Visual Molecular Dynamics (VMD²⁷) 1.9.3 and AmberTools22^{28,29} were used to prepare each RNA-ligand complex for MD simulations.

Each nucleic atom was parametrized according to ff14SB force with χ modification tuned for RNA (χ_{OL3}) field^{30–32}, while ligands parameters were assigned according to General Amber Force Field (GAFF³³). Since the Corn Aptamer (PDB ID: 5BJO) presents a modified Uracil (I5-U17), a custom preparation was exploited following Amber's workflow to create modified residues employing antechamber and parmcheck2 tools, and assigning parameters from ff14SB and GAFF force field using parmcheck2.

Each RNA-ligand complex was solvated within a rectangular base prism box, with a 15 Å padding between the box border and the nearest RNA atom, using the TIP3P water model³⁴.

Sodium and chlorine monovalent ions were added to neutralize the net charge of the box, reaching the physiological salt concentration of 0.154M. Before equilibration, each system was then subjected to 500 steps of energy minimization with the conjugate-gradient method to remove clashes and bad contacts.

Following this preparation phase, a two-step equilibration process was performed. The first phase consisted of a 0.5 ns simulation in canonical ensemble (NVT) imposing a 5 kcal mol⁻¹ Å⁻² harmonic positional restraint to each RNA and ligand atom, leaving water and ions unconstrained. In the second equilibration run, a 0.5 ns simulation was conducted in the isothermal-isobaric ensemble (NPT) constraining only the backbone atoms of RNA through the same force constant used in the first stage. For each equilibration stage, the temperature was kept constant at the lowest value indicated in the temperature range through a Langevin thermostat³⁵, and for the second NPT simulation, the pressure was fixed at 1 atm through a Monte Carlo barostat³⁶.

For each MD simulation an integration timestep of 2 fs was used. The simulations were performed exploiting the proprietary ACEMD 3.5 engine³⁷, which is based on the open-source library for molecular simulations OpenMM³⁸. The M-SHAKE algorithm was used to constrain the length of bonds involving hydrogen atoms, the particle-mesh Ewald (PME³⁹) method was exploited to compute electrostatic interactions using cubic spline interpolation, and finally, a 9.0 Å cutoff was used for calculating Lennard–Jones interactions.

Thermal Titration Molecular Dynamics (TTMD) simulations

Thermal Titration Molecular Dynamics (TTMD^{15,16}) is an enhanced sampling molecular dynamics approach originally developed for the estimation of protein-ligand unbinding kinetics. This method consists in a series of short MD simulations (TTMD-steps) performed at progressively increasing temperatures. The length and the temperature for each step, defining the temperature ramp for the TTMD simulation, are defined by the user based on the knowledge of the system of interest, especially concerning the conservation of the native receptor fold throughout the duration of the simulation. In this work, two different temperature ramps were used: the first one, defined as “standard”, is the same ramp described in the original publication (starting temperature 300 K, final temperature 450 K, temperature increase 10 K, step length 10 ns), and a second one defined as “alternative” (starting temperature 73 K, final temperature 223 K, temperature increase 10 K, step length 10 ns). An interaction fingerprint-based scoring function¹⁷ is exploited to monitor the conservation of the native binding mode. Specifically, the IFP_{CS} scoring function¹⁷ is used to compare each of the binding features of each trajectory frame (encoded as protein-ligand

interaction fingerprints) to the last frame of the second equilibration stage. In the context of this work, two different PLIFs were employed, specifically the InteractionFingerprint function of the Open Drug Discovery Toolkit (ODDT ⁴⁰) and the one provided by the ProLIF package ⁴¹, while the cosine similarity metrics was used for the comparison, as implemented in the Scikit-learn Python library. The calculated value is then multiplied by -1 to comply with the scale of most scoring functions. In the end, the resulting score can range from -1 to 0, where -1 indicates total congruence of the PLIFs (total conservation of the binding mode) and 0 indicates that all native binding features are lost.

The simulation continues until the end of the temperature ramp is reached, or an early termination criterion is reached. Specifically, at the end of each TTMD-step, the average IFP_{CS} score is calculated for the last 10% of the step: if this value is above -0.05, the simulation is stopped as the ligand lost its original binding mode. The code for running TTMD simulations is open-source and available through the MIT license at github.com/molecularmodelingsection/TTMD.

Trajectory Analysis, MS and IFF Coefficients Determination

TTMD trajectories were analyzed partially through the same Python code described in the previous paragraph and partially through the SuMD-analyzer ² script available at github.com/molecularmodelingsection/SuMD-analyzer. Specifically, the MDAnalysis package ^{42,43} was exploited to calculate the RMSD of the receptor backbone, the ligand and a defined set of binding site residues.

The “titration timeline” plot reports the time-dependent evolution of both the ligand, binding site, and receptor backbone RMSD and of the IFP_{CS} score.

The “titration profile” plots the average IFP_{CS} score for each TTMD-step against the temperature at which the step was executed. In this plot, the slope of the straight line linking the first and last point of the simulation (the MS coefficient) is extracted and used as a proxy measure for the estimation of the overall complex stability. The mathematical formulation of the MS coefficient is reported in Equation 1.

Equation 1. MS mathematical formulation.

$$MS = \frac{\text{mean IFP}_{CS}^{T_{end}} - (-1)}{T_{end} - T_{start}}$$

The MS coefficient can range from zero (indicative of a tight and persistence binding / high conservation of the native binding mode) to 1 (indicative of a high volatility of the native pose). For each investigated receptor-ligand complex, five independent TTMD simulations were performed, with the average MS coefficient being calculated across three different replicates, after discarding the highest and the lowest value.

Furthermore, to complement the MS coefficient, a second one defined as IFF was calculated as defined in Equation 2.

Equation 2. IFF mathematical formulation.

$$IFF = \sqrt{\frac{\sum_t^T (IFP_{CS,t} - \langle IFP_{CS} \rangle)^2}{T}}$$

The IFF coefficient is calculated as the root mean square fluctuation of the IFP_{CS} value across the simulation in respect of the whole run mean IFP_{CS} value. As the MS coefficient only takes in account only the initial and the final state of the simulation, neglecting even significant fingerprint fluctuations that may suggest a poor complex stability, the IFF coefficient can overcome this issue, further improving the protocol accuracy.

Results

To extend the applicability domain of Thermal Titration Molecular Dynamics to the characterization of RNA-ligand complexes, six different test-cases were chosen based on the chemical diversity of the ligand and structural diversity of the receptor among the experimental complexes deposited in the Protein Data Bank. Four out of the six test-cases were drawn out from the work of Bissaro et al. ¹, specifically 2LWK, 3Q50, 5BJO and 6E1U, while 1UUD and 1UUI were added because of the challenge provided by the high conformational freedom of the P12 and P14 ligands.

Hereinafter a brief overview of the six test cases.

Influenza A virus RNA promoter

Influenza A virus RNA promoter has an important role in the modulation of transcription and replication of this group of viruses belonging to the Orthomyxoviridae family. Therefore, it is considered an interesting target for the development of antiviral drugs ⁴⁴. Through a fragment screening approach, 6,7-dimethoxy-2-(1-piperazinyl)-4-quinazolinamine (DPQ) was identified by Varani's group ⁴⁴. This compound has a micromolar affinity for the promoter region (K_D of 50.5 μ M), and consequently, it is a scaffold for further development of antivirals targeting Influenza A promoter region (PDB ID: 2LWK).

Corn Aptamer

The Corn aptamer is an *in vitro* selected aptamer that binds a fluorophore ligand (DFHO), with a nanomolar binding affinity ($K_D = 70$ nM) ⁴⁵. Thanks to his limited cytotoxicity, the Corn-DFHO could represent a valuable imaging tool. Despite being therapeutically less interesting than the other RNA structures selected for this study, its peculiar three-dimensional organization in a quasi-symmetric dimer with a non-cationic ligand buried at the monomer-monomer interface between two G-quadruplex structures stabilized by K⁺ ions makes it an appealing target to apply TTMD protocol, since its stability in molecular dynamics, and the low structural flexibility of the fluorophore ligand (3,5-difluoro-4-hydroxybenzylidene imidazolinone-2-oxime) led Bissaro et al to optimal results when applying SuMD protocol to this system¹.

Pre-queosine 1 Riboswitch

PreQ1 (Pre-queosine 1, 7-aminomethyl-7-deazaguanine) is a precursor of the hypermodified guanine nucleotide Queosine (Q), a modified nucleoside that influences anticodon-codon stability in tRNAs, hence increasing translational fidelity. Prokaryotes can synthesize PreQ1 molecule from a multienzyme pathway starting from GTP. By contrast, eukaryotes do not synthesize PreQ1 *de novo* and need to retrieve queuoine from the diet ⁴⁶. For this reason, PreQ1 riboswitches are only found in prokaryotes, where, by binding PreQ1, they play a fundamental role in the biosynthesis and transport of this molecule. Therefore, riboswitches are an appealing target for antibacterial drug development. Two PreQ1 riboswitch complexes have been selected, the first one (PDB ID: 3Q50) is bound to a PreQ1 precursor with a nanomolar binding affinity ($K_D = 2$ nM), the second one (PDB ID: 6E1U) is a hybrid riboswitch aptamer bound to a synthetic compound (HMJ) that presents a sub-micromolar affinity ($K_D = 0.5$ μ M). This complex has been obtained through the deletion of nucleobase A14 and the two vicinal ones, as a direct consequence, the binding site lacks important nucleotides and, congruently to the work of Bissaro et al. ¹, before running TTMD simulations, only the ligand was retained from this complex, and the aptamer structure was retrieved from 3Q50 crystal.50 crystal.

Tar

Tat-Tar interaction modulates the transcription process in HIV-1. Therefore, targeting Tar RNA could be an appealing pathway to follow for the development of HIV-1 therapeutics. Although combination antiretroviral therapy (cART) can significantly suppress viral load, the eradication of this virus is still challenging. Moreover, chronic long-term comorbidities can arise since viral reservoirs are not completely eliminated by cART. In this scenario blocking HIV transcription could be beneficial ⁴⁷. Several efforts have been made in the characterization of this target and various small-molecule complexes have been reported in the PDB. In this work, two Tar complexes have been reported, specifically selected to assay TTMD behavior with challenging flexible ligands (P12, P14) with high conformational freedom.

The investigated cases are summarized in Table 1.

Table 1. Summary of the six test-cases investigated in the present study.

PDB code	Ligand PDB ID	Receptor type
1UUD ⁴⁸	P14	HIV-1 TAR RNA (29 bases)
1UUI ⁴⁸	P12	HIV-1 TAR RNA (29 bases)
2LWK ⁴⁴	0EC	influenza A virus RNA promoter (32 bases)
3Q50 ⁴⁹	PRF	PreQ1 riboswitch aptamer (33 bases)
5BJO ⁴⁵	747	Corn RNA aptamer (36 bases)
6E1U ⁵⁰	HMJ	PreQ1 riboswitch (33 bases)

For each complex, at first, a self-docking experiment was conducted using the PLANTS docking program, since it has recently been reported as one of the best docking tools for RNA-ligand complexes ²⁶ and it is also free for academics. The results of the self-docking calculations are summarized in Figures 1-6.

For each complex, the top five scoring poses and the experimentally determined one were subjected to TTMD refinement using the same conditions described in the original publications ^{15,16}. Furthermore, another run of TTMD calculations was conducted using an alternative temperature ramp. Two different sets of protein-ligand fingerprints were used for monitoring the evolution of the binding mode, as well as two different metrics for the qualitative estimation of pose stability, namely the MS and IFF coefficient. Every methodological detail is reported in the Materials and Methods section. Hereafter, each TTMD experiment is discussed separately based on execution conditions, while aggregated considerations are reported in the Discussion.

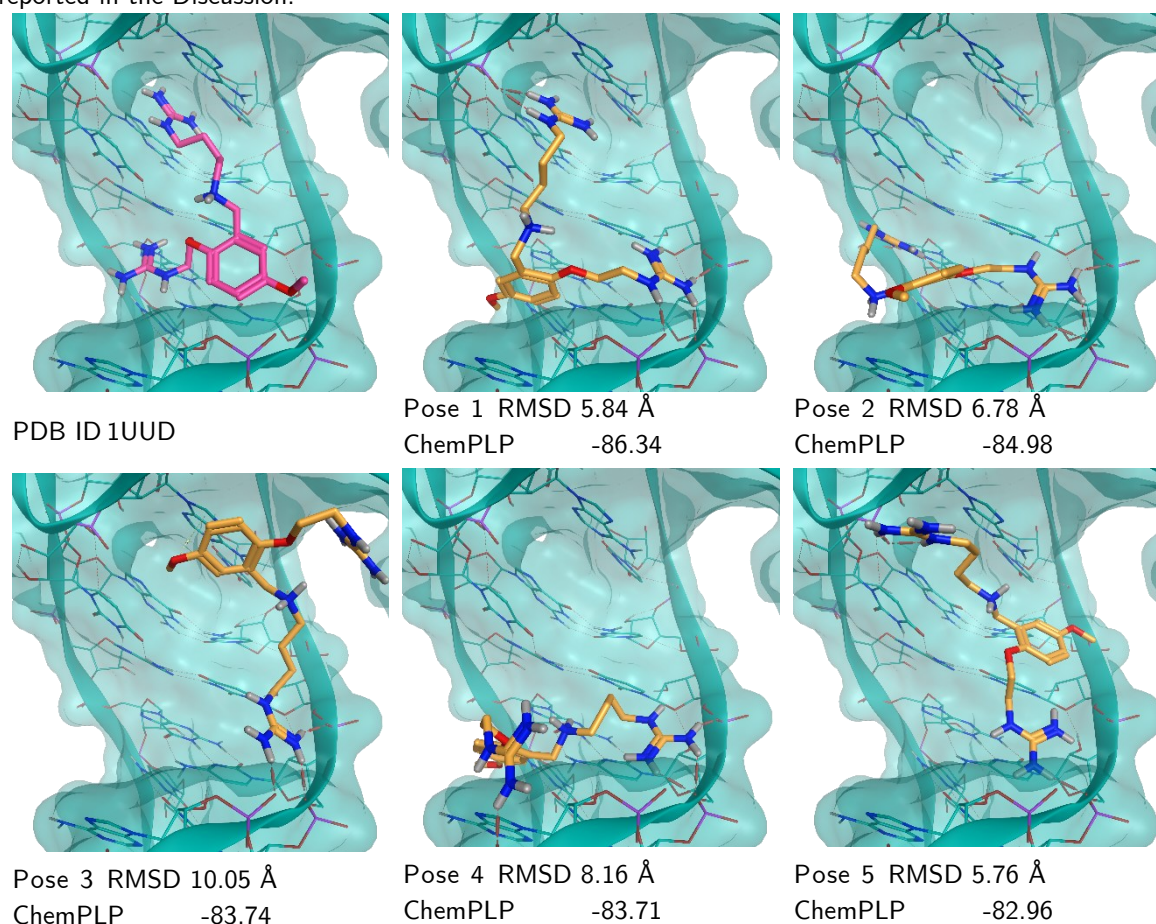


Figure 1. This panel encompasses all binding modes (experimental + docking poses) investigated in this work for complex deposited in the PDB with accession code 1UUD. For each pose, the ChemPLP docking score and RMSD to the experimental pose are reported.

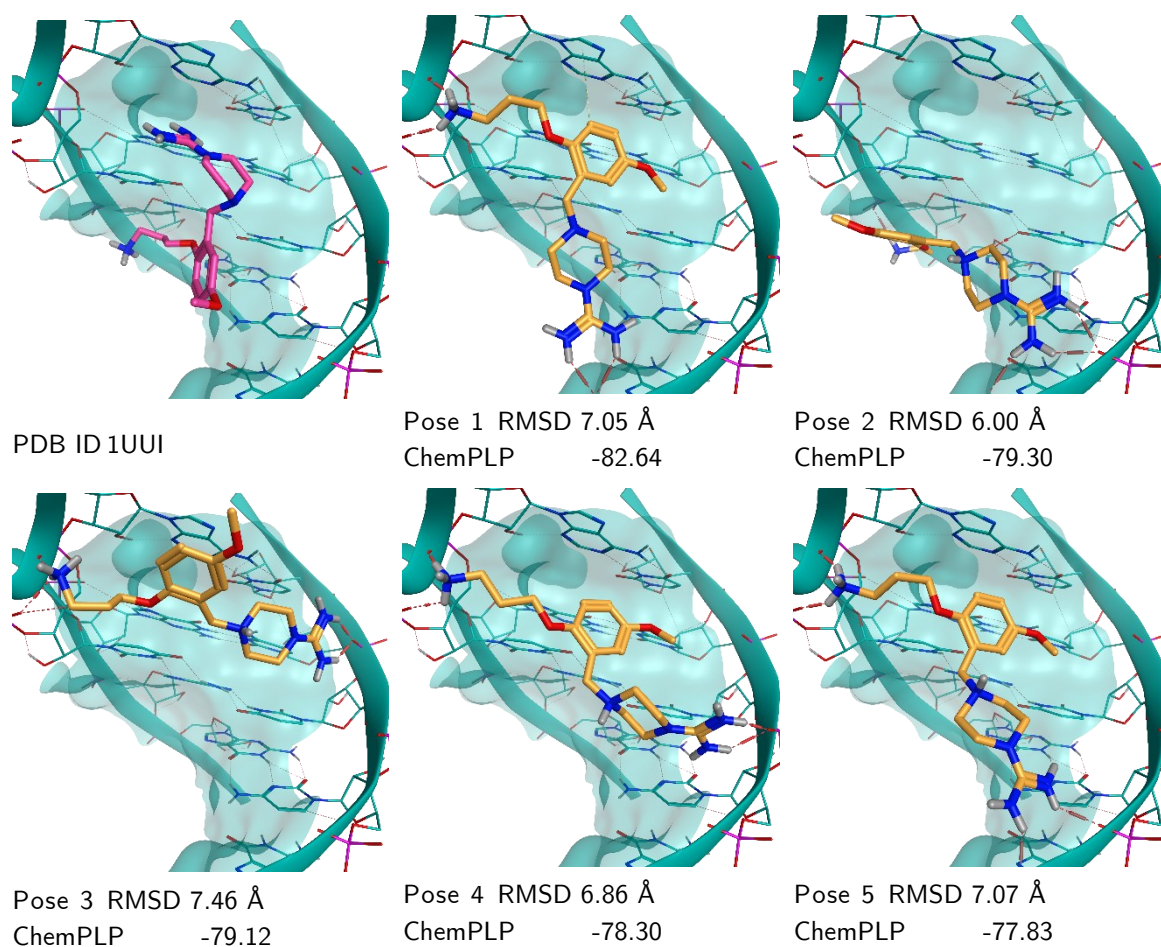


Figure 2.: this panel encompasses all binding modes (experimental + docking poses) investigated in this work for complex deposited in the PDB with accession code 1UUI. For each pose, the ChemPLP docking score and RMSD to the experimental pose are reported.

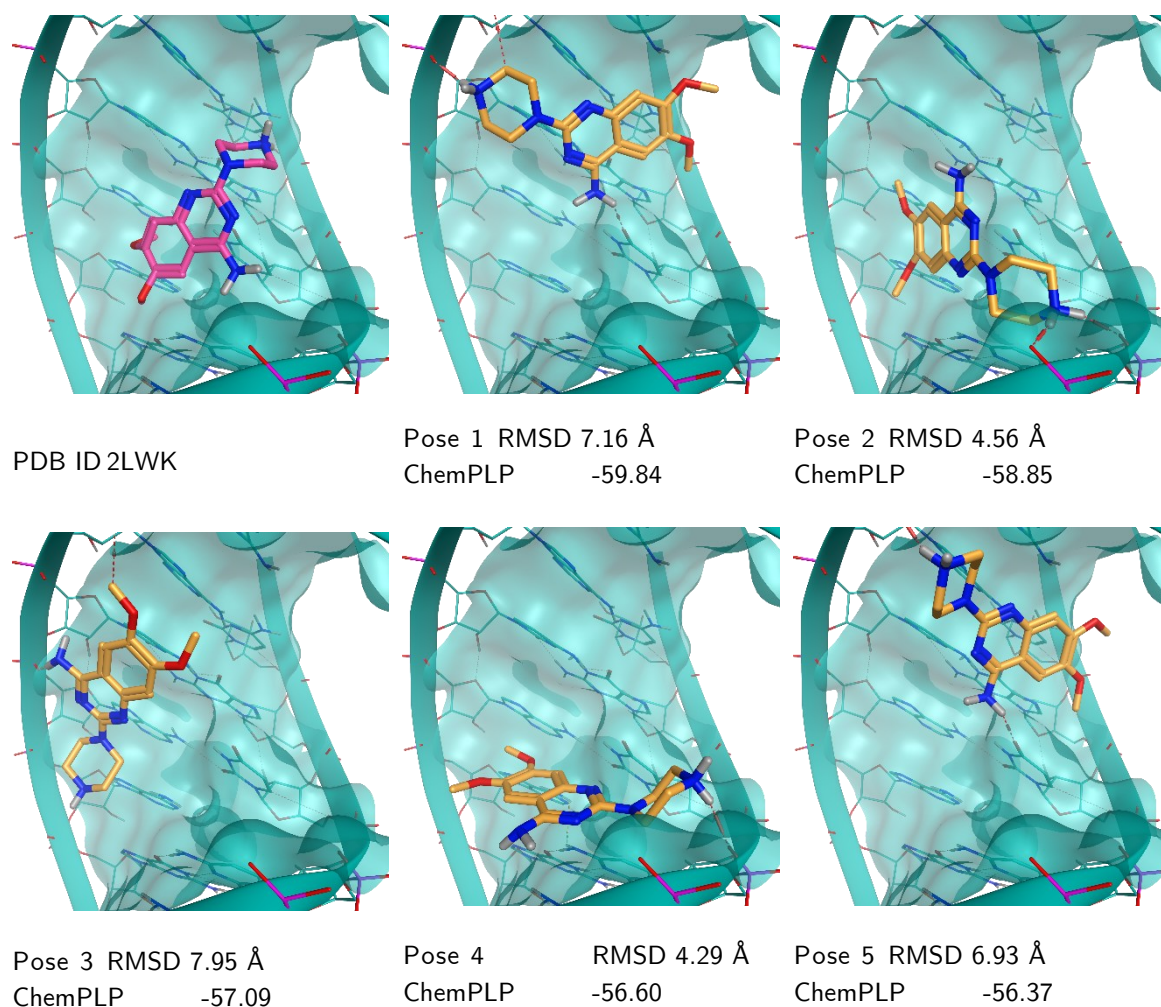


Figure 3. This panel encompasses all binding modes (experimental + docking poses) investigated in this work for complex deposited in the PDB with accession code 2LWK. For each pose, the ChemPLP docking score and RMSD to the experimental pose are reported.

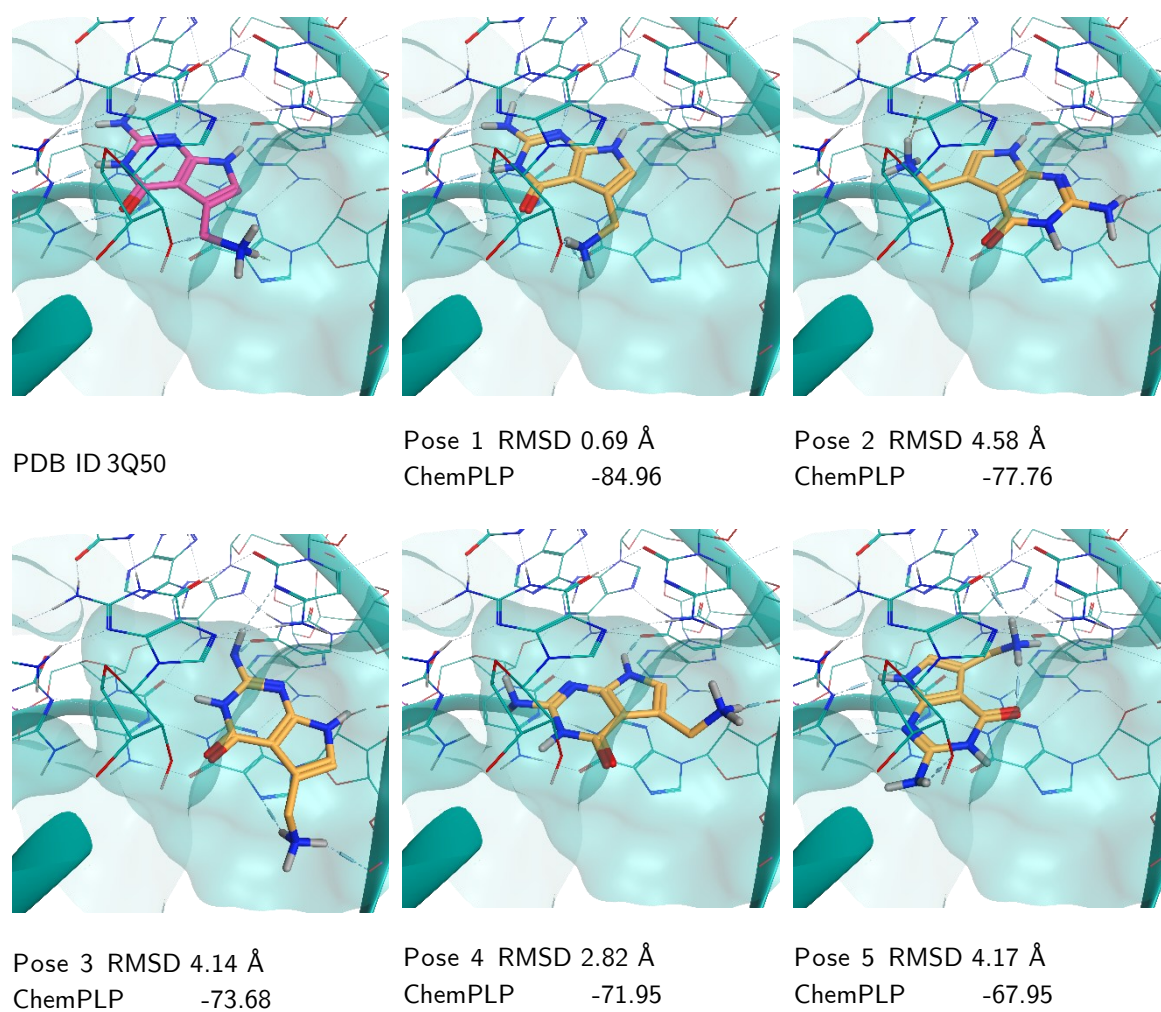


Figure 4. This panel encompasses all binding modes (experimental + docking poses) investigated in this work for complex deposited in the PDB with accession code 3Q50. For each pose, the ChemPLP docking score and RMSD to the experimental pose are reported.

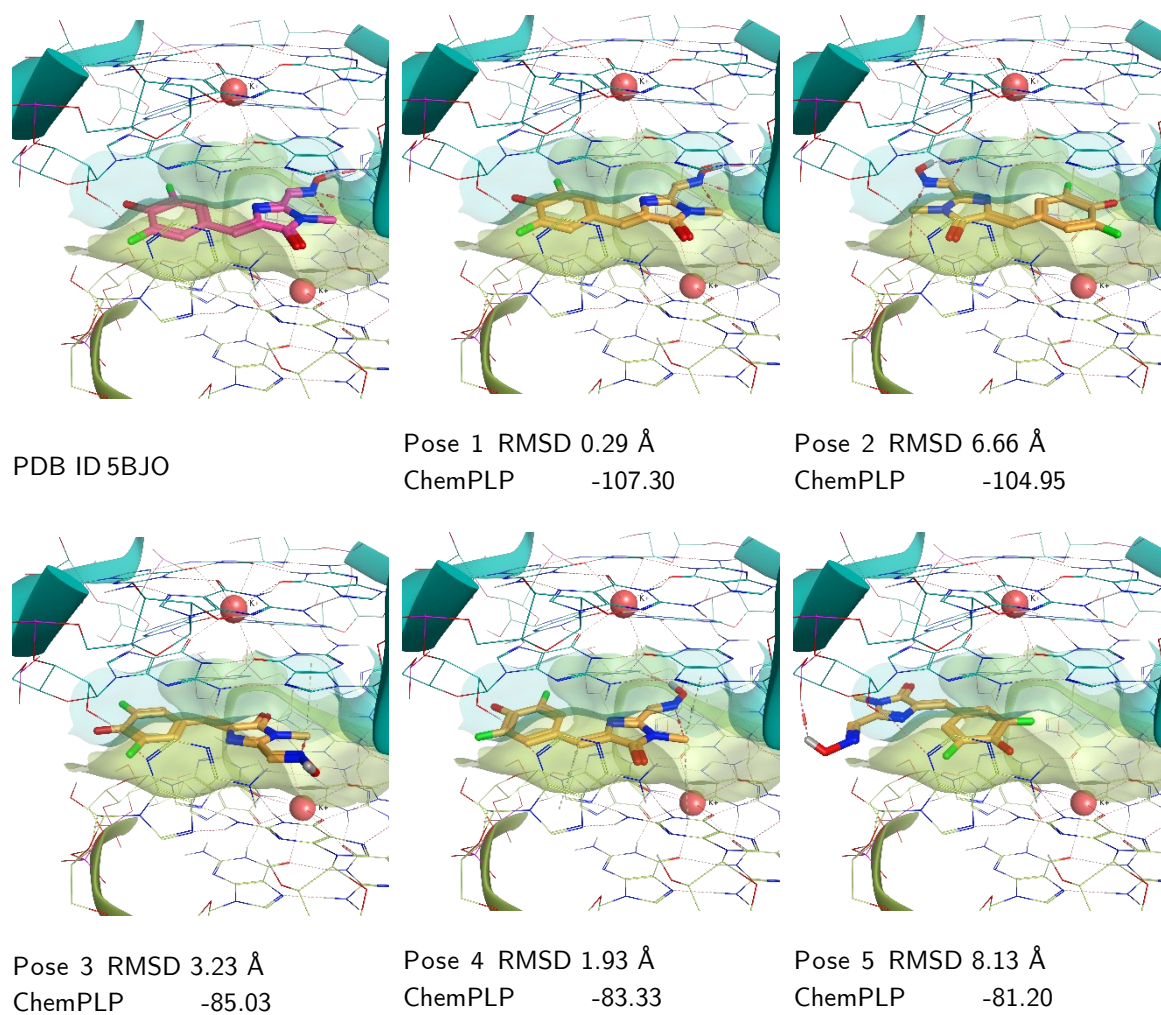


Figure 5. This panel encompasses all binding modes (experimental + docking poses) investigated in this work for complex deposited in the PDB with accession code 5BJO. For each pose, the ChemPLP docking score and RMSD to the experimental pose are reported.

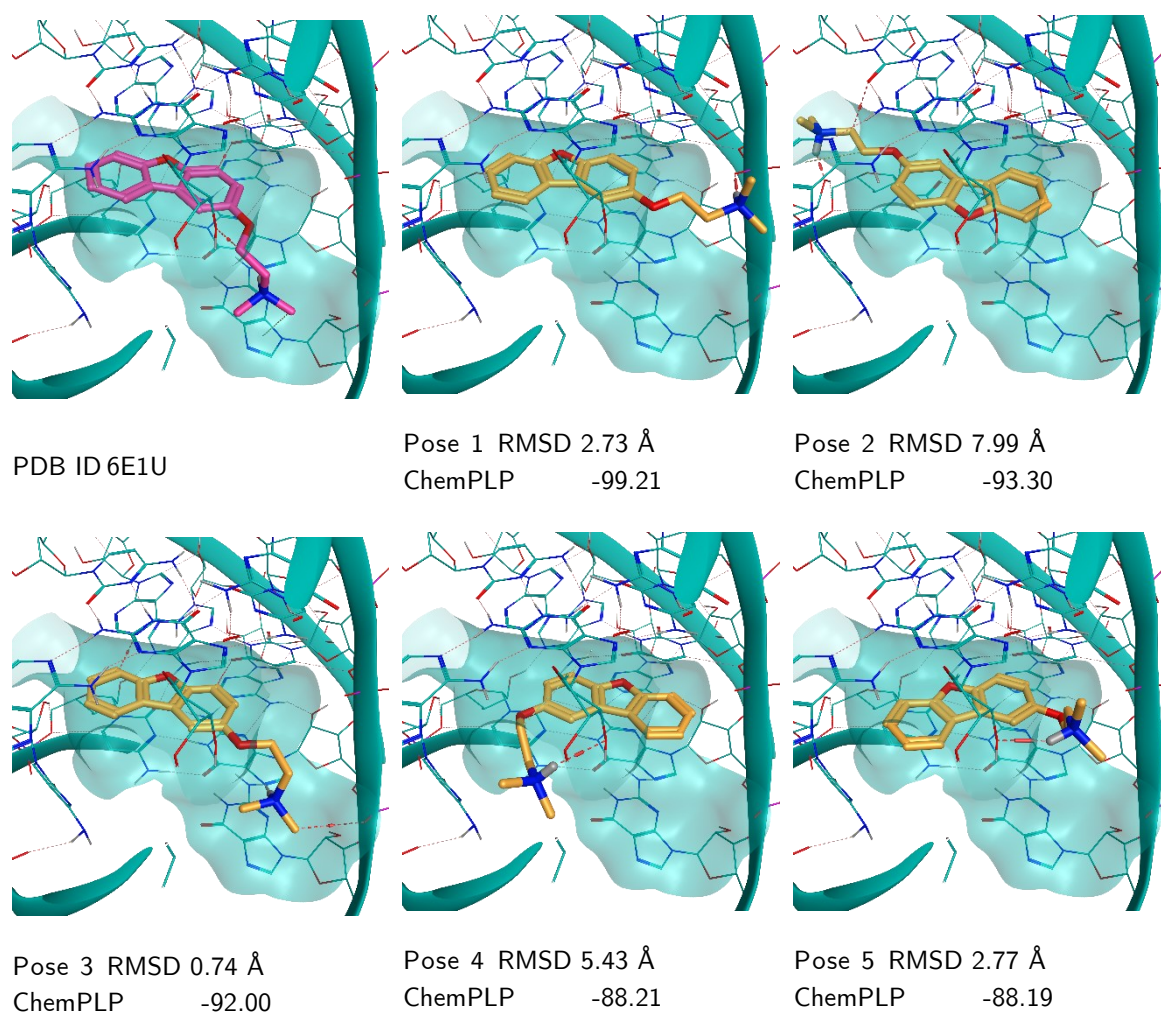


Figure 6. This panel encompasses all binding modes (experimental + docking poses) investigated in this work for complex deposited in the PDB with accession code 6E1U. For each pose, the ChemPLP docking score and RMSD to the experimental pose are reported.

TTMD protocol 1: standard temperature ramp, ODDT fingerprints, MS coefficient

As a first attempt, we tried repurposing the same TTMD workflow described in the original publications^{15,16} without any modifications. We therefore used a standard temperature ramp, setting up the starting simulation temperature at 300 K, the end temperature at 450 K, the temperature increase between each TTMD-step to 10 K and the duration of each TTMD-step at 10 ns. Poses were then ranked according to the MS coefficient, using ODDT interaction fingerprints for monitoring the evolution of the binding mode. Results for this first round of TTMD simulations are summarized in Table 2.

As can be seen in Table 2, the standard TTMD protocol that was successfully applied to the characterization of protein-ligand complexes' residence time cannot be readily repurposed to RNA-ligand complexes. What we noticed immediately, indeed, was that several replicates presented very high MS values (1, for example), implying that the ligands lose most or even all the native interaction features in the first TTMD-step. By visually inspecting the trajectory and subsequent analyses, we noticed though that this was not true, as the ligands were still in place and in a conformation within the binding site that was compatible with a similar interaction pattern as the one of the reference frame. We carefully checked the fingerprint calculation step, and we noticed indeed that ODDT interaction fingerprints cannot capture most of the interactions that exists between the ligand and the RNA receptor, especially concerning hydrogen bonds and pi stacking, which are dominant in this kind of systems, and captured just a fraction of the hydrophobic contacts. We tried manually modifying the cutoffs for the calculation of such interactions, to see if they were too strict, but the results did not dramatically change. We therefore concluded that it was an atom-typing issue, with ODDT not being able to rightfully classify all RNA atoms and depict the intermolecular interactions involving them. For this reason, we modified the TTMD code to use a different protein-ligand interaction fingerprint, provided by the ProLIF package⁴¹, since it supports DNA and RNA molecules as well. TTMD protocol 2: standard temperature ramp, PROLIF fingerprints, MS coefficient

As a second attempt, we re-analyzed existing TTMD trajectories monitoring the evolution of the receptor-ligand interaction fingerprint through the ProLIF package instead of the ODDT one. The result of this analysis is reported in Table 3.

As can be observed in Table 3, although the sensitivity of the fingerprint-based scoring function improved compared to the traditional TTMD protocol thanks to the change of the package used for the PLIF calculation, the results of the TTMD simulations are still not satisfactory. Taking aside the first three test-cases, i.e. 1UUD, 1UUI, and 2LWK for which the docking itself was not reliable thus justifying the challenging nature of the task, in the other three cases the results were promising but showed margin of improvement. Indeed, in all these three cases, the TTMD refinement would have at best excluded a couple of wrong poses from the equation but would have attributed some other wrong poses similar scores to the experimentally determined binding mode, resulting in an inconclusive post-docking refinement procedure. On the other hand, in the case of complex 3Q50, where the closest poses to the crystal reference are pose 1 and 4, respectively, the MS coefficient would have awarded pose 1 and 2, giving them scores similar to the experimentally determined binding pose, and higher scores to the other ones. Intriguingly, the pose ranking capabilities would have been very good in the case of complexes 5BJO and 6E1U, were the closest poses to the experimental data (1, 4 and 3 for 5BJO, and 3, 1, and 5 for 6E1U, respectively) would have been correctly distinguished by the MS coefficient from the more incorrect ones, other than being close in MS score to the crystal reference and being correctly ranked based on the RMSD to the native binding mode. Based on the visual inspection of TTMD trajectories, and on previous experience accumulated working on similar targets^{1,2}, we thought that a possible solution for improving the accuracy of the prediction could be to use a different temperature ramp, specifically a less aggressive one, in order to reduce the amplitude of the conformational movements of the RNA receptors to favorably decouple the evolution of the three-dimensional structure of the receptor from the diffusion of the ligand from the binding site.

Table 2. This table summarizes results for TTMD simulations carried out under the protocol 1 simulation and analysis conditions. For each investigated system, the individual MS coefficient for each TTMD replicate and the average MS value are reported.

Model	Pose	MD1	MD2	MD3	MD4	MD5	TrimMean	TrimError
1UUD	crystal	0.00759	0.00909	0.01960	0.01667	0.00514	0.01112	0.00554
1UUD	1	0.00905	0.00628	0.01200	0.00506	0.01000	0.00844	0.00251
1UUD	2	0.01595	0.03264	0.00612	0.00665	0.02471	0.01577	0.01030
1UUD	3	0.00701	0.00648	0.01990	0.01661	0.01111	0.01157	0.00528
1UUD	4	1.00000	1.00000	1.00000	1.00000	0.09940	1.00000	0.36024
1UUD	5	0.00456	0.00612	0.00421	1.00000	1.00000	0.33689	0.48747
1UUI	crystal	0.01000	0.01363	0.01000	0.01111	0.00347	0.01037	0.00336
1UUI	1	0.01419	0.01246	1.00000	1.00000	1.00000	0.67140	0.48337
1UUI	2	0.01429	0.09650	0.10000	0.01664	0.01991	0.04435	0.03989
1UUI	3	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	0.00000
1UUI	4	1.00000	0.04945	0.03333	1.00000	0.03249	0.36093	0.47111
1UUI	5	1.00000	1.00000	0.09894	1.00000	0.10000	0.70000	0.44117
2LWK	crystal	1.00000	N/A	1.00000	1.00000	0.03330	1.00000	0.41859
2LWK	1	1.00000	1.00000	1.00000	1.00000	0.01906	1.00000	0.39238
2LWK	2	1.00000	1.00000	0.00833	1.00000	N/A	1.00000	0.42941
2LWK	3	0.05000	1.00000	1.00000	0.02500	1.00000	0.68333	0.47159
2LWK	4	1.00000	0.00598	1.00000	N/A	1.00000	1.00000	0.43042
2LWK	5	0.02000	0.09918	0.00714	1.00000	0.03333	0.05084	0.38534
3Q50	crystal	0.00326	0.00237	0.00323	0.00212	0.00168	0.00257	0.00063
3Q50	1	0.00231	0.00154	0.00271	0.00311	0.00216	0.00239	0.00053
3Q50	2	0.00225	0.00137	0.00210	0.00207	0.00198	0.00205	0.00031
3Q50	3	0.00524	0.00480	0.00430	0.00540	0.00233	0.00478	0.00111
3Q50	4	0.00517	0.00404	0.00368	0.00437	0.00353	0.00403	0.00058
3Q50	5	0.00348	0.00440	0.00314	0.00398	0.00351	0.00366	0.00044
5BJO	crystal	0.00318	0.00463	0.00433	0.00305	0.00455	0.00402	0.00069
5BJO	1	0.00273	0.00457	0.00385	0.00471	0.00279	0.00374	0.00085
5BJO	2	0.00391	0.00511	0.00473	0.00431	0.00472	0.00459	0.00041
5BJO	3	0.00368	0.00474	0.00469	0.00444	0.00430	0.00448	0.00038
5BJO	4	0.00337	0.00346	0.00415	0.00257	0.00256	0.00313	0.00060
5BJO	5	0.00371	0.00392	0.00520	0.00601	0.00503	0.00472	0.00085
6EIU	crystal	0.00207	0.00073	0.00241	0.00063	0.00653	0.00174	0.00215
6EIU	1	0.00182	0.09966	0.02446	0.03333	0.00064	0.01987	0.03614
6EIU	2	0.00650	0.00667	0.00382	0.00461	0.00390	0.00500	0.00124
6EIU	3	0.00071	0.00139	0.00113	0.00040	0.00055	0.00080	0.00037
6EIU	4	0.09758	0.04980	1.00000	0.05000	0.04929	0.06579	0.37579
6EIU	5	0.09838	1.00000	0.00250	0.04971	0.00427	0.05079	0.38612

Table 3. This table summarizes results for TTMD simulations carried out under the protocol 2 simulation and analysis conditions. For each investigated system, the individual MS coefficient for each TTMD replicate and the average MS value are reported.

Model	Pose	MD1	MD2	MD3	MD4	MD5	TrimMean	TrimError
1UUD	crystal	0.00553	0.00895	0.01250	0.01404	0.00518	0.00899	0.00358
1UUD	1	0.00833	0.00538	0.00752	0.00612	0.00991	0.00732	0.00160
1UUD	2	0.00596	0.03247	0.00752	0.00666	0.01096	0.00838	0.01002
1UUD	3	0.00706	0.00585	0.00392	0.01102	0.01111	0.00798	0.00285
1UUD	4	0.01429	0.00961	0.09604	0.01193	0.09968	0.04075	0.04213
1UUD	5	0.00470	0.00731	0.00420	0.00694	1.00000	0.00632	0.39769
1UUI	crystal	0.00769	0.01427	0.01000	0.00958	0.00336	0.00909	0.00354
1UUI	1	0.00534	0.00553	1.00000	0.01363	0.05000	0.02305	0.39290
1UUI	2	0.01603	0.00627	0.10000	0.00865	0.01954	0.01474	0.03528
1UUI	3	0.00692	0.00414	1.00000	0.00894	0.09760	0.03782	0.38984
1UUI	4	1.00000	0.00709	0.02424	0.00442	0.00602	0.01245	0.39589
1UUI	5	1.00000	0.10000	0.03278	0.00587	0.01250	0.04843	0.38632
2LWK	crystal	1.00000	0.03235	0.10000	0.00529	0.02500	0.05245	0.38506
2LWK	1	1.00000	1.00000	1.00000	1.00000	0.00833	1.00000	0.39667
2LWK	2	0.00887	0.00448	0.00833	0.00566	0.00533	0.00644	0.00174
2LWK	3	0.00909	0.01588	0.01246	0.02438	1.00000	0.01757	0.39385
2LWK	4	1.00000	0.00406	0.01111	0.01374	0.10000	0.04162	0.38870
2LWK	5	0.02000	0.03268	0.00694	0.02424	0.03307	0.02564	0.00962
3Q50	crystal	0.00053	0.00035	0.00045	0.00068	0.00090	0.00055	0.00019
3Q50	1	0.00060	0.00109	0.00070	0.00067	0.00038	0.00065	0.00023
3Q50	2	0.00049	0.00054	0.00038	0.00041	0.00049	0.00047	0.00006
3Q50	3	0.00335	0.00340	0.00342	0.00202	0.00080	0.00292	0.00105
3Q50	4	0.00154	0.00086	0.00249	0.00220	0.00277	0.00208	0.00069
3Q50	5	0.00191	0.00090	0.00115	0.00054	0.00085	0.00096	0.00046
5BJO	crystal	0.00075	0.00048	0.00053	0.00041	0.00051	0.00051	0.00012
5BJO	1	0.00069	0.00083	0.00043	0.00057	0.00524	0.00070	0.00185
5BJO	2	0.00293	0.00319	0.00164	0.00397	0.00211	0.00274	0.00082
5BJO	3	0.00226	0.00141	0.00078	0.00229	0.00509	0.00199	0.00147
5BJO	4	0.00089	0.00085	0.00083	0.00075	0.00083	0.00084	0.00004
5BJO	5	0.00206	0.00139	0.00320	0.00627	0.00596	0.00374	0.00200
6EIU	crystal	0.00114	0.00039	0.00043	0.00050	0.00140	0.00069	0.00042
6EIU	1	0.00035	0.00071	0.00077	0.00124	0.00048	0.00066	0.00030
6EIU	2	0.00224	0.00663	0.00054	0.00052	0.00464	0.00247	0.00239
6EIU	3	0.00034	0.00046	0.00033	0.00037	0.00059	0.00039	0.00010
6EIU	4	0.00185	0.00298	0.00213	0.00193	0.00351	0.00235	0.00065
6EIU	5	0.00172	0.00076	0.00068	0.00119	0.00199	0.00122	0.00052

TTMD protocol 3: alternative temperature ramp, PROLIF fingerprints, MS coefficient

In the attempt of further improving the predictive capabilities of the method, we decided to test out an alternative, less aggressive, temperature ramp. Specifically, we set up a starting temperature of 73 K and an end temperature of 223 K, while maintaining the same temperature increase and step duration. The ramp was purposefully chosen within a temperature interval that could be considered “extreme cold” based on a pair of assumptions: first, with highly thermostable macromolecules such as protein receptors investigated in previous TTMD works, working within a temperature interval that is way outside the realistic representation of the system did not impair the ability to extract meaningful data out of the simulation (i.e., a qualitative estimation of the residence time), second lowering the temperature of the simulated system will reduce the atomic velocities, thus resulting in a decreased flexibility of the RNA receptor. Although temperatures in Molecular Dynamics do not have necessarily an equivalence with real-world ones, a similar concept of extracting meaningful data at non-physiological temperatures is commonly applied in Molecular Biology, for example, while determining crystallographic structures with X-ray spectroscopy and Cryo-EM techniques. The results of this set of TTMD simulations are encompassed in Table 4.

As can be deduced from Table 4, the use of an alternative temperature ramp that operates at lower temperature values did not necessarily improve the quality of the TTMD post-docking refinement. First, the quality of the prediction decreased in those cases (1UUD, 1UUI, and 2LWK) where docking was not able to generate a good docking pose: indeed, while protocol 2 was at least able to assign the lowest scores to a pool of poses that included the experimentally determined ones, here it can be noticed how the native binding mode is always among the ones with the worst scores in term of MS coefficient. Speaking of the other three cases, instead, where protocol 2 performed quite decently, here the pose ranking capabilities of TTMD stays practically identical, with the only difference being that the MS score difference between various poses is decreased, flattening the difference in score between native-like poses and wrong ones. Looking at the fingerprint profile, though, we noticed how lowering the temperature ranges not only reduced the amplitude of conformational shifts in the nucleic receptor, but also decreased the tendency of the ligand to diffuse from the binding site, thus making less efficient the protocol. Practically speaking, although the “titration timeline” profile clearly showed a difference in the amplitude of the changes in the fingerprint-based score throughout the TTMD simulation, the difference between the end and starting point of the simulation was not so marked. This observation led us to conclude that, maybe, the MS coefficient in its original formulation could not be the most sensitive metric to capture the subtle rearrangements of the ligand within the binding site and the conservation of the binding features, thus inducing us to introduce a different metric defined as IFF coefficient (see Materials and Methods for a detailed explanation on how the IFF coefficient is computed).

TTMD protocol 4: alternative temperature ramp, PROLIF fingerprints, IFF coefficient

To test if an alternative metric to the MS coefficient would be able to improve the sensitivity of the TTMD protocol using an alternative temperature ramp that operates at sub-freezing temperature, we re-analyzed those trajectories determining the IFF coefficient. The results are summarized in Table 5.

By observing Table 5, it can be seen as even the new metric is not able to improve the results of the TTMD refinement performed at such lower temperatures. Specifically, the performance of the method remains like what could be observed with the old-fashioned MS coefficient. To further prove the point, we re-analyzed also those trajectories performed with the standard temperature ramp (see Table S1, Supplementary Materials) and we found a great level of coherence between the MS and IFF coefficient regarding their ability to rank poses based on the persistence of native binding features throughout the simulation.

Table 4. This table summarizes results for TTMD simulations carried out under the protocol 3 simulation and analysis conditions. For each investigated system, the individual MS coefficient for each TTMD replicate and the average MS value are reported.

Model	Pose	MD1	MD2	MD3	MD4	MD5	TrimMean	TrimError
1UUD	crystal	0.00442	0.00314	0.00493	0.00436	0.00494	0.00457	0.00065
1UUD	1	0.00476	0.00090	0.00089	0.00125	0.00353	0.00189	0.00159
1UUD	2	0.00461	0.00397	0.00324	0.00279	0.00336	0.00352	0.00063
1UUD	3	0.00534	0.00159	0.00318	0.00144	0.00159	0.00212	0.00150
1UUD	4	0.00068	0.00221	0.00686	0.00239	0.00280	0.00247	0.00206
1UUD	5	0.00449	0.00189	0.00282	0.00440	0.00299	0.00340	0.00099
1UUI	crystal	0.00402	0.00353	0.00321	0.00191	0.00329	0.00335	0.00070
1UUI	1	0.00192	0.00175	0.00167	0.00200	0.00198	0.00188	0.00013
1UUI	2	0.00054	0.00073	0.00058	0.00109	0.00555	0.00080	0.00193
1UUI	3	0.00282	0.00349	0.00135	0.00323	0.00302	0.00302	0.00075
1UUI	4	0.00355	0.00157	0.00302	0.00255	0.00230	0.00262	0.00067
1UUI	5	0.00205	0.00437	0.00182	0.00308	0.00266	0.00259	0.00090
2LWK	crystal	0.00334	0.00336	0.00327	0.00371	0.00279	0.00332	0.00029
2LWK	1	0.00448	0.00452	0.00489	0.00118	0.00500	0.00463	0.00143
2LWK	2	0.00203	0.00170	0.00506	0.00349	0.00234	0.00262	0.00123
2LWK	3	0.00042	0.00087	0.00437	0.00571	0.00273	0.00266	0.00202
2LWK	4	0.00133	0.00142	0.00089	0.00112	0.00091	0.00112	0.00021
2LWK	5	0.00381	0.00343	0.00445	0.00199	0.00517	0.00390	0.00107
3Q50	crystal	0.00072	0.00026	0.00015	0.00060	0.00024	0.00037	0.00022
3Q50	1	0.00016	0.00017	0.00065	0.00016	0.00012	0.00016	0.00020
3Q50	2	0.00024	0.00041	0.00045	0.00033	0.00023	0.00032	0.00009
3Q50	3	0.00029	0.00037	0.00041	0.00046	0.00051	0.00041	0.00008
3Q50	4	0.00131	0.00143	0.00139	0.00164	0.00127	0.00138	0.00013
3Q50	5	0.00037	0.00049	0.00060	0.00055	0.00043	0.00049	0.00008
5BJO	crystal	0.00028	0.00020	0.00015	0.00021	0.00017	0.00019	0.00004
5BJO	1	0.00019	0.00021	0.00022	0.00016	0.00017	0.00019	0.00002
5BJO	2	0.00150	0.00290	0.00221	0.00175	0.00156	0.00184	0.00052
5BJO	3	0.00061	0.00038	0.00047	0.00051	0.00047	0.00049	0.00007
5BJO	4	0.00099	0.00099	0.00044	0.00071	0.00067	0.00079	0.00021
5BJO	5	0.00095	0.00141	0.00037	0.00161	0.00358	0.00133	0.00108
6EIU	crystal	0.00037	0.00032	0.00022	0.00016	0.00027	0.00027	0.00008
6EIU	1	0.00020	0.00055	0.00028	0.00023	0.00032	0.00028	0.00012
6EIU	2	0.00019	0.00053	0.00032	0.00042	0.00039	0.00038	0.00011
6EIU	3	0.00016	0.00016	0.00224	0.00015	0.00039	0.00024	0.00082
6EIU	4	0.00103	0.00086	0.00175	0.00167	0.00191	0.00149	0.00042
6EIU	5	0.00059	0.00049	0.00076	0.00044	0.00054	0.00054	0.00011

Table 5. This table summarizes results for TTMD simulations carried out under the protocol 4 simulation and analysis conditions. For each investigated system, the individual IFF coefficient for each TTMD replicate and the average IFF value are reported.

Model	Pose	MD1	MD2	MD3	MD4	MD5	TrimMean	TrimError
1UUD	crystal	0.08170	0.10090	0.13150	0.09300	0.10040	0.09810	0.01653
1UUD	1	0.06440	0.02760	0.04400	0.04900	0.04240	0.04513	0.01185
1UUD	2	0.20300	0.13680	0.14390	0.08850	0.15780	0.14617	0.03683
1UUD	3	0.12670	0.07270	0.09660	0.05810	0.05610	0.07580	0.02661
1UUD	4	0.09040	0.16830	0.20920	0.10370	0.15340	0.14180	0.04341
1UUD	5	0.22860	0.07600	0.11650	0.25030	0.12610	0.15707	0.06776
1UUI	crystal	0.14520	0.13290	0.10620	0.18700	0.10570	0.12810	0.03000
1UUI	1	0.07680	0.07460	0.06880	0.07180	0.06380	0.07173	0.00456
1UUI	2	0.04520	0.03630	0.08390	0.05670	0.26680	0.06193	0.08601
1UUI	3	0.12960	0.10450	0.09530	0.09370	0.15250	0.10980	0.02268
1UUI	4	0.12010	0.07550	0.06420	0.11970	0.12710	0.10510	0.02608
1UUI	5	0.10760	0.11910	0.08610	0.10380	0.11110	0.10750	0.01095
2LWK	crystal	0.10880	0.08440	0.14570	0.10430	0.23220	0.11960	0.05244
2LWK	1	0.20760	0.17670	0.16880	0.06800	0.24650	0.18437	0.05942
2LWK	2	0.09790	0.09620	0.21300	0.16450	0.10550	0.12263	0.04627
2LWK	3	0.03890	0.06550	0.24620	0.27600	0.13130	0.14767	0.09482
2LWK	4	0.07100	0.05820	0.08010	0.07680	0.06400	0.07060	0.00806
2LWK	5	0.19820	0.16620	0.19210	0.11030	0.23790	0.18550	0.04213
3Q50	crystal	0.03240	0.01460	0.01350	0.02740	0.01560	0.01920	0.00771
3Q50	1	0.01830	0.01940	0.03700	0.01800	0.02800	0.02190	0.00741
3Q50	2	0.01190	0.02160	0.01500	0.01570	0.01470	0.01513	0.00318
3Q50	3	0.01540	0.02380	0.02470	0.01890	0.02420	0.02230	0.00365
3Q50	4	0.03980	0.06500	0.06080	0.07390	0.06200	0.06260	0.01123
3Q50	5	0.03190	0.02640	0.03390	0.02570	0.02420	0.02800	0.00378
5BJO	crystal	0.01500	0.01070	0.01290	0.01690	0.01200	0.01330	0.00220
5BJO	1	0.01250	0.01070	0.01530	0.01210	0.01180	0.01213	0.00153
5BJO	2	0.05420	0.08590	0.06550	0.08580	0.05420	0.06850	0.01427
5BJO	3	0.02340	0.01610	0.02170	0.02030	0.01890	0.02030	0.00249
5BJO	4	0.03760	0.03790	0.01630	0.04140	0.02980	0.03510	0.00899
5BJO	5	0.08020	0.08120	0.05140	0.07380	0.16150	0.07840	0.03752
6EIU	crystal	0.02510	0.02530	0.02400	0.01210	0.01150	0.02040	0.00639
6EIU	1	0.02380	0.01820	0.01500	0.03590	0.01400	0.01900	0.00802
6EIU	2	0.02140	0.03310	0.01610	0.02880	0.02040	0.02353	0.00613
6EIU	3	0.00740	0.01510	0.11340	0.00900	0.01950	0.01453	0.04049
6EIU	4	0.05210	0.04850	0.08180	0.12690	0.05890	0.06427	0.02904
6EIU	5	0.02720	0.02240	0.03100	0.03700	0.02890	0.02903	0.00478

Discussion

Molecular docking represents the state-of-the-art technique for the prediction of protein-ligand complexes, thanks to the good compromise between accuracy and rapidity of execution ⁹. Although routinely used in various structure-based drug discovery campaigns, molecular docking has been specifically designed and optimized with protein-ligand receptor in mind, mainly due to pharmaceutical relevance and availability of experimentally determined structures to use as a benchmark for method development ⁷. Despite the growing interest of medicinal chemists towards targeting RNA macromolecules for therapeutic and diagnostic purposes in the last decade, the field of rational design of RNA-targeting ligands is still relatively unexplored, due to some intrinsic challenges portrayed by these molecular entities, such as their structural plasticity which makes it difficult to obtain their experimentally determined three-dimensional structure, a pivotal requirement for modern-days drug discovery campaigns ^{1,2}.

Despite all these challenges, recent works have demonstrated how, despite some intrinsic limitations, routinely used molecular docking programs such as PLANTS can be quite efficiently used for investigating RNA-ligand complexes as well ²⁶. Furthermore, we previously showcased how Supervised Molecular Dynamics (SuMD) simulations can be successfully utilized to flank molecular docking in investigating the recognition process of complexes involving RNA molecules, both as ligands ² and as receptors ¹, retrieving useful information about the whole binding process beyond the final bound state. Critically, one of the major problems of docking in all its iterations, both classic and dynamic based, is finding a good scoring metric that can discriminate native-like poses from the wrong ones ¹². The peculiarities of RNA molecules, such as the distinctive surface charge distribution, exacerbate the limitations of classical scoring functions, which have been developed specifically for protein binding sites. For all these reasons, introducing new and improved ways of investigating RNA-ligand complexes is a very pressing issue.

In the present article, we presented the first application of Thermal Titration Molecular Dynamics (TTMD), a recently developed MD-based protocol for the qualitative estimation of protein-ligand unbinding kinetics, to the characterization of RNA-ligand complexes. We investigated six different pharmaceutically relevant test-cases of different molecular complexity, both on the ligand and on the receptor side. We performed two rounds of simulations, one with the standard temperature ramp already described in the original publications ^{15,16}, and one with an alternative ramp operating at sub-freezing temperatures, to investigate the effect of the temperature ramp on the accuracy of prediction. We modified the original code to use the ProLIF package ⁴¹ for the calculation of RNA-ligand interaction fingerprint instead of the ODDT one ⁴⁰, since they exhibit a better description of these systems. We also explored a different metric, the IFF coefficient, to use alongside the MS coefficient for monitoring the conservation of the native binding determinants and calculating a proxy measure for the receptor-ligand residence time. By analyzing the whole set of simulations and analysis performed upon them, it seems clear that TTMD can be a helpful tool to refine docking results in those cases where the geometry of the binding site is well defined and maintained throughout the simulation, in a similar way to what was already observed for protein-ligand systems. 5BJO simulations are a proof of this concept as the binding site remains extremely stable throughout the course of the simulation (Video V1, Table S1, Table S2 Supplementary Materials). In those cases where the perturbation of the receptor conformation and/or the receptor fold overcomes the tendency of the ligand to diffuse from the binding site, the method accuracy drops significantly, not improving docking results and even invalidating successful ones.

Concerning this aspect, the choice of the temperature ramp is pivotal for the success of the TTMD rescoring process. Despite the high flexibility of RNA molecules, especially at high simulation temperatures, the standard temperature ramp still outperformed the alternative, sub-freezing one. A plausible explanation for this evidence is that the electrostatic component dominates the protein-ligand interaction energy, making it very unlikely for the ligand to spontaneously detach and unbind without any external aid in the shape of a temperature increase. Regardless of the choice of the metric used for estimating the poses residence time (the MS or the IFF coefficient), the ranking provided by the standard ramp was more useful than the one provided by the alternative one, due to the flattening of the score differences between good and bad poses at lower temperatures. Despite this, the exploration of the alternative ramp allowed us to determine that simulating in those conditions will not result in discarding a good pose, at least for those cases where TTMD (and docking too) perform well. The possibility to tune the temperature interval with

a certain degree of freedom based on the knowledge of the system makes it quite appealing, since in a prospective use case one could simply start with a titration in the low-temperature range, evaluate the separation in score between different poses, and eventually re-execute the titration at higher temperatures based on the results provided by the first round of simulations.

Another useful expediency that the user could implement would be to carry out a TTMD “dry run” on just the RNA molecule itself, a functionality that has been introduced purposefully in the latest version of the TTMD code used for this article, and evaluate its behavior in terms of structural plasticity, especially at the binding site level, before running the titration on systems including the ligand as well. It is important to stress that it probably exists an intermediate ramp between the two that were utilized in the present work that might be a better compromise and provide slightly better results, but at the same time we firmly believe that in order to be adopted by the largest number of user possible a method has to be easy to implement and setup, without spending too much time on optimizing parameters for its execution. For this reason, we wanted to showcase that the method can perform reasonably well almost regardless of the choice of the ramp, although the original ramp that has already proven to be successful in the case of protein-ligand complexes is still the best one in terms of accuracy. Finally, it is worth mentioning that working with macromolecules that have an intrinsic conformational flexibility such as RNA strongly limits the applicability of any given MD technique in a high-throughput fashion, due to its limited sampling capabilities, making docking still the most efficient choice for investigating these systems. On the other hand, the ever-increasing computational power available for MD simulations will make more and more appealing the use of pipelines like TTMD in the future, thus justifying the interests in the further development of this and other MD-based method for the investigation of systems involving nucleic acids³.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2023.1294543/full#supplementary-material>

Author Contributions

AD: Formal Analysis, Investigation, Writing—original draft. MP: Investigation, Methodology, Writing—original draft. SME: Data curation, Formal Analysis, Software, Visualization, Writing—original draft. VS: Methodology, Supervision, Validation, Writing—review and editing. MS: Supervision, Supervision, Writing—review and editing, Formal Analysis. SMO: Supervision, Conceptualization, Writing—review and editing.

References

1. Bissaro, M.; Sturlese, M.; Moro, S. Exploring the RNA-Recognition Mechanism Using Supervised Molecular Dynamics (SuMD) Simulations: Toward a Rational Design for Ribonucleic-Targeting Molecules? *Front. Chem.* 2020, 8, 515945. <https://doi.org/10.3389/fchem.2020.00107/BIBTEX>.
2. Pavan, M.; Bassani, D.; Sturlese, M.; Moro, S. Investigating RNA–Protein Recognition Mechanisms through Supervised Molecular Dynamics (SuMD) Simulations. *NAR Genomics Bioinforma.* 2022, 4 (4). <https://doi.org/10.1093/NARGAB/LQAC088>.
3. Sponer, J.; Bussi, G.; Krepl, M.; Banas, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jurečka, P.; Walter, N. G.; Otyepka, M. RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chem. Rev.* 2018, 118 (8), 4177. <https://doi.org/10.1021/ACS.CHEMREV.7B00427>.
4. Childs-Disney, J. L.; Yang, X.; Gibaut, Q. M. R.; Tong, Y.; Batey, R. T.; Disney, M. D. Targeting RNA Structures with Small Molecules. *Nat. Rev. Drug Discov.* 2022 2110 2022, 21 (10), 736–762. <https://doi.org/10.1038/s41573-022-00521-4>.
5. Zafferani, M.; Hargrove, A. E. Small Molecule Targeting of Biologically Relevant RNA Tertiary and Quaternary Structures. *Cell Chem. Biol.* 2021, 28 (5), 594–609. <https://doi.org/10.1016/J.CHEMBIOL.2021.03.003>.
6. Falese, J. P.; Donlic, A.; Hargrove, A. E. Targeting RNA with Small Molecules: From Fundamental Principles towards the Clinic. *Chem. Soc. Rev.* 2021, 50 (4), 2224–2243. <https://doi.org/10.1039/D0CS01261K>.
7. Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* 2018, 9 (AUG), 393738. <https://doi.org/10.3389/fphar.2018.00923/BIBTEX>.

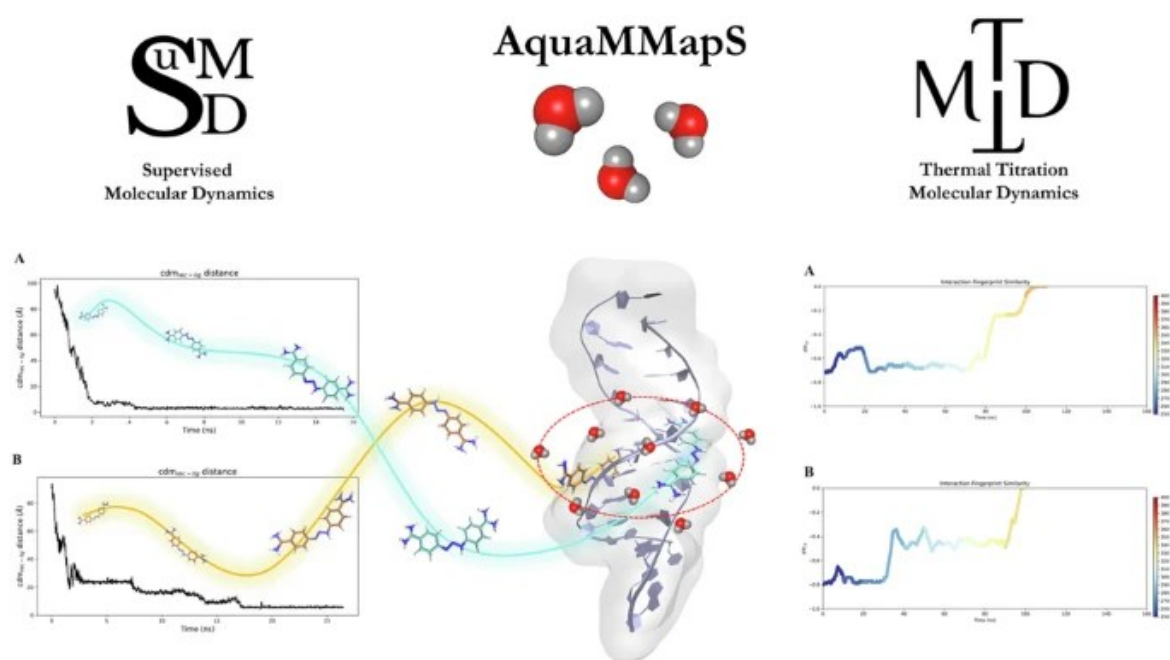
8. Bassani, D.; Moro, S. Past, Present, and Future Perspectives on Computer-Aided Drug Design Methodologies. *Mol.* 2023 Vol 28 Page 3906 2023, 28 (9), 3906. <https://doi.org/10.3390/MOLECULES28093906>.
9. Pavan, M.; Moro, S. Lessons Learnt from COVID-19: Computational Strategies for Facing Present and Future Pandemics. *Int. J. Mol. Sci.* 2023 Vol 24 Page 4401 2023, 24 (5), 4401. <https://doi.org/10.3390/IJMS24054401>.
10. Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput. Aided-Drug Des.* 2012, 7 (2), 146–157. <https://doi.org/10.2174/157340911795677602>.
11. Ferreira, L. G.; Santos, R. N. D.; Oliva, G.; Andricopulo, A. D. Molecular Docking and Structure-Based Drug Design Strategies. *Mol.* 2015 Vol 20 Pages 13384-13421 2015, 20 (7), 13384–13421. <https://doi.org/10.3390/MOLECULES200713384>.
12. Chaput, L.; Mouawad, L. Efficient Conformational Sampling and Weak Scoring in Docking Programs? Strategy of the Wisdom of Crowds. *J. Cheminformatics* 2017, 9 (1), 37. <https://doi.org/10.1186/s13321-017-0227-x>.
13. Bernardi, R. C.; Melo, M. C. R.; Schulten, K. Enhanced Sampling Techniques in Molecular Dynamics Simulations of Biological Systems. *Biochim. Biophys. Acta BBA - Gen. Subj.* 2015, 1850 (5), 872–877. <https://doi.org/10.1016/J.BBAGEN.2014.10.019>.
14. Do, T. N.; Carloni, P.; Varani, G.; Bussi, G. RNA/Peptide Binding Driven by Electrostatics - Insight from Bidirectional Pulling Simulations. *J. Chem. Theory Comput.* 2013, 9 (3), 1720–1730. https://doi.org/10.1021/CT3009914/SUPPL_FILE/CT3009914_SI_001.PDF.
15. Menin, S.; Pavan, M.; Salmasso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement. *Int. J. Mol. Sci.* 2023, 24 (4), 3596. <https://doi.org/10.3390/ijms24043596>.
16. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Qualitative Estimation of Protein–Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2022, 62 (22), 5715–5728. <https://doi.org/10.1021/acs.jcim.2c00995>.
17. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Implementing a Scoring Function Based on Interaction Fingerprint for Autogrow4: Protein Kinase CK1 δ as a Case Study. *Front. Mol. Biosci.* 2022, 9, 629. <https://doi.org/10.3389/fmolb.2022.909499>.
18. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* 2000, 28 (1), 235–242. <https://doi.org/10.1093/NAR/28.1.235>.
19. Molecular Operating Environment (MOE), 2022.02, C. C. G. Molecular Operating Environment (MOE), 2022.02. https://www.chemcomp.com/Research-Citing_MOE.htm.
20. Bhattacharyya, D.; Mirihana Arachchilage, G.; Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* 2016, 4.
21. O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminformatics* 2011, 3 (10), 1–14. <https://doi.org/10.1186/1758-2946-3-33/TABLES/2>.
22. Cadence Molecular Sciences. OpenEye. *Cadence Molecular Sciences, Santa Fe, NM.* <http://www.eyesopen.com>, 2023. <http://www.eyesopen.com>.
23. Korb, O.; Stützle, T.; Exner, T. E. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design. In *Ant Colony Optimization and Swarm Intelligence*; Dorigo, M., Gambardella, L. M., Birattari, M., Martinoli, A., Poli, R., Stützle, T., Eds.; Lecture Notes in Computer Science; Springer: Berlin, Heidelberg, 2006; pp 247–258. https://doi.org/10.1007/11839088_22.
24. Korb, O.; Stützle, T.; Exner, T. E. An Ant Colony Optimization Approach to Flexible Protein–Ligand Docking. *Swarm Intell.* 2007, 1 (2), 115–134. <https://doi.org/10.1007/s11721-007-0006-9>.
25. Korb, O.; Stützle, T.; Exner, T. E. Empirical Scoring Functions for Advanced Protein–Ligand Docking with PLANTS. *J. Chem. Inf. Model.* 2009, 49 (1), 84–96. <https://doi.org/10.1021/ci800298z>.
26. Jiang, D.; Zhao, H.; Du, H.; Deng, Y.; Wu, Z.; Wang, J.; Zeng, Y.; Zhang, H.; Wang, X.; Wu, J.; Hsieh, C. Y.; Hou, T. How Good Are Current Docking Programs at Nucleic Acid–Ligand Docking? A Comprehensive Evaluation. *J. Chem. Theory Comput.* 2023, 19, 5633–5647. https://doi.org/10.1021/ACS.JCTC.3C00507/ASSET/IMAGES/LARGE/CT3C00507_0006.JPEG.
27. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* 1996, 14 (1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
28. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* 2005, 26 (16), 1668–1688. <https://doi.org/10.1002/JCC.20290>.
29. Case, D. A. AmberTools22, 2022.
30. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* 2007, 92 (11), 3817–3829. <https://doi.org/10.1529/BIOPHYSJ.106.097782>.
31. Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E.; Jurečka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* 2011, 7 (9), 2886–2902. https://doi.org/10.1021/CT200162X/SUPPL_FILE/CT200162X_SI_001.PDF.
32. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 2015, 11 (8), 3696–3713. https://doi.org/10.1021/ACS.JCTC.5B00255/SUPPL_FILE/CT5B00255_SI_001.PDF.

33. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 2004, 25 (9), 1157–1174. <https://doi.org/10.1002/JCC.20035>.
34. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 1983, 79 (2), 926–935. <https://doi.org/10.1063/1.445869>.
35. Davidchack, R. L.; Handel, R.; Tretyakov, M. V. Langevin Thermostat for Rigid Body Dynamics. *J. Chem. Phys.* 2009, 130 (23). <https://doi.org/10.1063/1.3149788/925280>.
36. Faller, R.; Pablo, J. J. D.; Chem, J.; Phys, (. Constant Pressure Hybrid Molecular Dynamics–Monte Carlo Simulations. *J. Chem. Phys.* 2002, 116 (1), 55–59. <https://doi.org/10.1063/1.1420460>.
37. Harvey, M. J.; Giupponi, G.; Fabritiis, G. D. ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* 2009, 5 (6), 1632–1639. https://doi.org/10.1021/CT9000685/ASSET/IMAGES/LARGE/CT-2009-000685_0004.JPEG.
38. Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLOS Comput. Biol.* 2017, 13 (7), e1005659. <https://doi.org/10.1371/journal.pcbi.1005659>.
39. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 1995, 103 (19), 8577–8593. <https://doi.org/10.1063/1.470117>.
40. Wójcikowski, M.; Zielenkiewicz, P.; Siedlecki, P. Open Drug Discovery Toolkit (ODDT): A New Open-Source Player in the Drug Discovery Field. *J. Cheminformatics* 2015, 7 (1), 26. <https://doi.org/10.1186/S13321-015-0078-2>.
41. Bouysset, C.; Fiorucci, S. ProLIF: A Library to Encode Molecular Interactions as Fingerprints. *J. Cheminformatics* 2021, 13 (1), 1–9. <https://doi.org/10.1186/S13321-021-00548-6/FIGURES/4>.
42. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* 2011, 32 (10), 2319–2327. <https://doi.org/10.1002/jcc.21787>.
43. Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domański, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python Sci. Conf.* 2016, 98–105. <https://doi.org/10.25080/Majora-629e541a-00e>.
44. Lee, M. K.; Bottini, A.; Kim, M.; Bardaro, M. F.; Zhang, Z.; Pellicchia, M.; Choi, B. S.; Varani, G. A Novel Small-Molecule Binds to the Influenza A Virus RNA Promoter and Inhibits Viral Replication. *Chem. Commun.* 2013, 50 (3), 368–370. <https://doi.org/10.1039/C3CC46973E>.
45. Warner, K. D.; Sjekloa, L.; Song, W.; Filonov, G. S.; Jaffrey, S. R.; Ferré-D'Amaré, A. R. A Homodimer Interface without Base Pairs in an RNA Mimic of Red Fluorescent Protein. *Nat. Chem. Biol.* 2017 1311 2017, 13 (11), 1195–1201. <https://doi.org/10.1038/nchembio.2475>.
46. Eichhorn, C. D.; Kang, M.; Feigon, J. Structure and Function of preQ1 Riboswitches. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* 2014, 1839 (10), 939–950. <https://doi.org/10.1016/J.BBAGRM.2014.04.019>.
47. Alanazi, A.; Ivanov, A.; Kumari, N.; Lin, X.; Wang, S.; Kovalskyy, D.; Nekhai, S. Targeting Tat–Tar Rna Interaction for Hiv-1 Inhibition. *Viruses* 2021, 13 (10), 2004. <https://doi.org/10.3390/V13102004/S1>.
48. Davis, B.; Afshar, M.; Varani, G.; Murchie, A. I. H.; Karn, J.; Lentzen, G.; Drysdale, M.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T.; Aboul-Ela, F. Rational Design of Inhibitors of HIV-1 TAR RNA through the Stabilisation of Electrostatic “Hot Spots.” *J. Mol. Biol.* 2004, 336 (2), 343–356. <https://doi.org/10.1016/J.JMB.2003.12.046>.
49. Jenkins, J. L.; Krucinska, J.; McCarty, R. M.; Bandarian, V.; Wedekind, J. E. Comparison of a PreQ1 Riboswitch Aptamer in Metabolite-Bound and Free States with Implications for Gene Regulation. *J. Biol. Chem.* 2011, 286 (28), 24626–24637. <https://doi.org/10.1074/JBC.M111.230375>.
50. Connelly, C. M.; Numata, T.; Boer, R. E.; Moon, M. H.; Sinniah, R. S.; Barchi, J. J.; Ferré-D'Amaré, A. R.; Schneekloth, J. S. Synthetic Ligands for PreQ1 Riboswitches Provide Structural and Mechanistic Insights into Targeting RNA Tertiary Structure. *Nat. Commun.* 2019 101 2019, 10 (1), 1–12. <https://doi.org/10.1038/s41467-019-09493-3>.

A second life for the crystallographic structure of Berenil-dodecanucleotide complex: a computational revisitation thirty years after its publication

Gianluca Novello, Andrea Dodaro, Silvia Menin, Chiara Cavastracci, Mattia Sturlese, Veronica Salmaso, Stefano Moro*

Novello, G.; Dodaro, A.; Menin, S.; Cavastracci Strascia, C.; Sturlese, M.; Salmaso, V.; Moro, S. A Second Life for the Crystallographic Structure of Berenil-Dodecanucleotide Complex: A Computational Revisitation Thirty Years after Its Publication. *Med Chem Res* 2024. <https://doi.org/10.1007/s00044-024-03299-9>.



Abstract

This study revisits the pioneering work of Professor Neidle, and co-workers, on the crystal structure of complexes formed between groove binders and DNA sequences. The original research revealed a DNA-ligand complex consisting of a dodecanucleotide bound with Berenil [1,3-bis(4'-amidinophenyl)-triazene] an anti-trypanocidal drug. This article aims to delve deeper into the structural dynamics of this system, showcasing the role played by water molecules in stabilizing the interaction between the ligand and the DNA. With this work, we reevaluate the findings from the original crystallographic study by employing modern molecular dynamics techniques, including Supervised Molecular Dynamics (SuMD) for generating binding trajectories, Thermal Titration Molecular Dynamics for assessing unbinding events, and AquaMMapS to identify regions occupied by stationary water molecules. The study addresses a minor and a major groove binding mode and assesses their strength and specificity using TTMD simulations, generating unbinding trajectories. This comprehensive approach integrates the understanding of the interaction of this DNA-ligand complex, which originated with

the valuable work of Professor Neidle, resulting in an in-depth insight into the pivotal role of water molecules with this DNA, a behavior detected and extendable even to other nucleic acid complexes.

Introduction

In this work, we revisited one of the first central contributions of Professor Neidle ¹ to the field of small groove ligand-DNA complexes crystal structures. The complex deposited in the Protein Data Bank ² with the ID code 2DBE, and core of the original research, consists of a dodecanucleotide d(CGCGAATTCGCG) bound with the anti-trypanocidal drug Berenil (Figure 1), a compound extensively studied for its applications as a cytotoxic and antiviral agent in animals ^{3,4}.

Ligands, like the cited Berenil, exhibit both intercalative and minor groove binding properties with DNA and RNA duplexes ⁵. This compound innovative crystallographic work not only showcases the complex structure of the DNA-ligand interaction but also highlights the essential role of water molecules, for the first time, in the recognition process between a minor-groove ligand and DNA. Coherently to the footprint analysis reported in the past ⁶, also the analysis performed in the study of Professor Neidle and his research team, suggests that strong binding sequences cover at least three base pairs, leading to the hypothesis of a possible preference for alternating AT sites. In the original published article, thanks to molecular modeling and perturbation energy studies, it was found that the use of energy minimization emphasized the binding mode obtained in the crystal, suggesting that Berenil's interaction with the eight-base duplex involved a broad enthalpy well extending over the 4 bp d(. . AATT . .) sequence, rather than a single global minimum. Building on these foundational insights, the original molecular modeling studies proposed a structural rationale for the interaction of Berenil with B-DNA via the minor groove, spanning two base pairs and forming preferential hydrogen bonds with adjacent thymine bases through their O2 atoms ⁷⁻⁹. In homage to the significant contributions of Professor Neidle, this study embarks on revisiting his landmark research, employing well-established molecular dynamics techniques developed by Professor Stefano Moro and his research team. Techniques such as Supervised Molecular Dynamics (SuMD) ¹⁰ for generating binding trajectories, Thermal Titration Molecular Dynamics ^{11,12} for assessing unbinding trajectories, and AquaMMapS ¹³ to identify zones occupied by stationary water molecules during MD simulations.

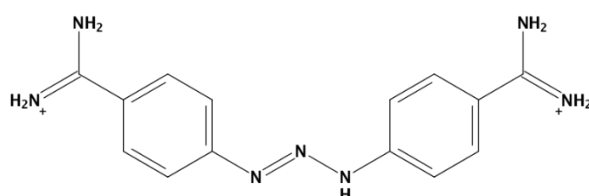


Figure 1. Chemical structure of Berenil [1,3-bis(4'-amidinophenyl)-triazene].

The goal of this study is to dive deeper into the structural dynamics of the Berenil-DNA complex, giving an insight into the role of water in the stabilization of the interaction between these two molecular objects, while elucidating and integrating the findings from the original crystallographic study, emphasizing the role of bridging water molecules and interactions within the DNA backbone

The focus of this work is to utilize information gathered from previous studies and reevaluate them through the perspective of modern molecular dynamics approaches, commonly employed in nucleic acid research ¹⁴⁻¹⁶. Employing both minor and major groove binding trajectories targeted toward the AT binding site, assessments will be made regarding the strength and specificity of interactions using Thermal Titration Molecular Dynamics (TTMD) simulations in unbinding trajectories. Through this comprehensive approach, an additional advance in the understanding of Berenil's interaction with DNA was obtained, contributing overall to the strategic and promising field of nucleic acid-small molecule interaction studies.

Results and Discussion

Starting with the structural information provided by the Berenil crystallography complex and the nucleotide sequences cataloged in the PDB under the 2DBE code, an attempt was made to generate a binding trajectory that could replicate the crystallographic binding mode, and its key interactions with the nucleic acid. To this end, Supervised Molecular Dynamics (SuMD) simulations have been conducted. SuMD is a well-established computational technique used for investigating unbiased receptor-ligand recognition pathways on the nanosecond timescale, by combining unbiased MD simulations and a tabu-like algorithm^{10,17}. This technique has been used in various studies to investigate the binding between several biomolecular receptors, including proteins, peptides, and RNA aptamers (available references). As depicted in Fig. 3 (Panel A), starting 95 Å away from the AATT nucleotide sequence, to avoid possible charge attraction effects, it was possible to explore the entire binding recognition in less than 16 nanoseconds (ns) of SuMD simulation.

The binding trajectory is rapid and linear. In fact, the ligand reaches the crystallographic binding site within the first 2 ns with favorable energy values. This is followed by a second phase of recognition and positioning extending up to 4 ns, where the ligand fully inserts into the minor groove. Despite the initial energetic disfavor observed during the first phase of recognition, the subsequent positioning within the minor groove is characterized by favorable and negative energetic interactions (Figure 2 and Video S1).

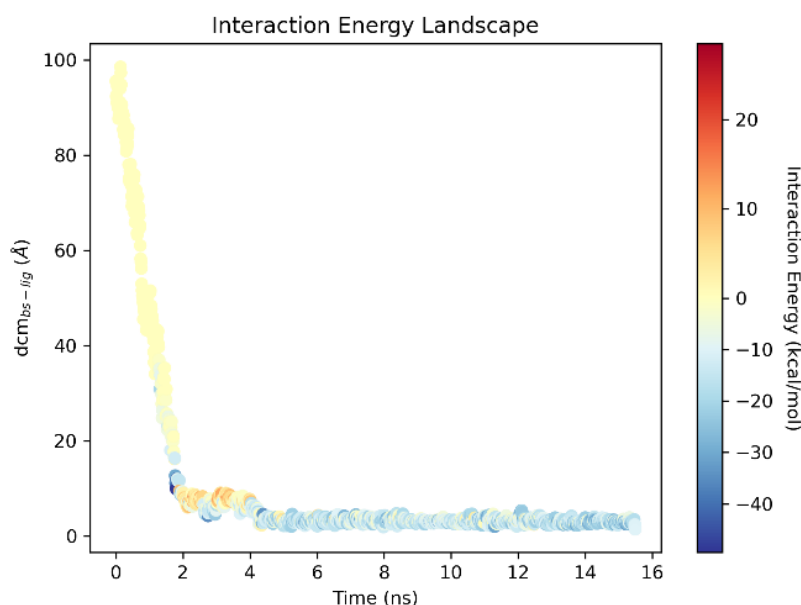


Figure 2. Illustration of the Interaction Energy profile of the minor groove molecular dynamics SuMD simulation.

By supervising the ligand proximity to the AT-rich binding site across the SuMD simulation, another stable binding mode at the level of AT-rich sequences was identified during the recognition phase. In this case, the ligand lies within the major groove displaying an alternative binding mode. The starting point of this simulation was identical to the one previously analyzed, but resulted in a different binding profile, observable in Figure 3 (Panel B). In this case, the simulation runs for approximately 26 ns, and is characterized by a less direct approach path and several interactive steps along the trajectory.

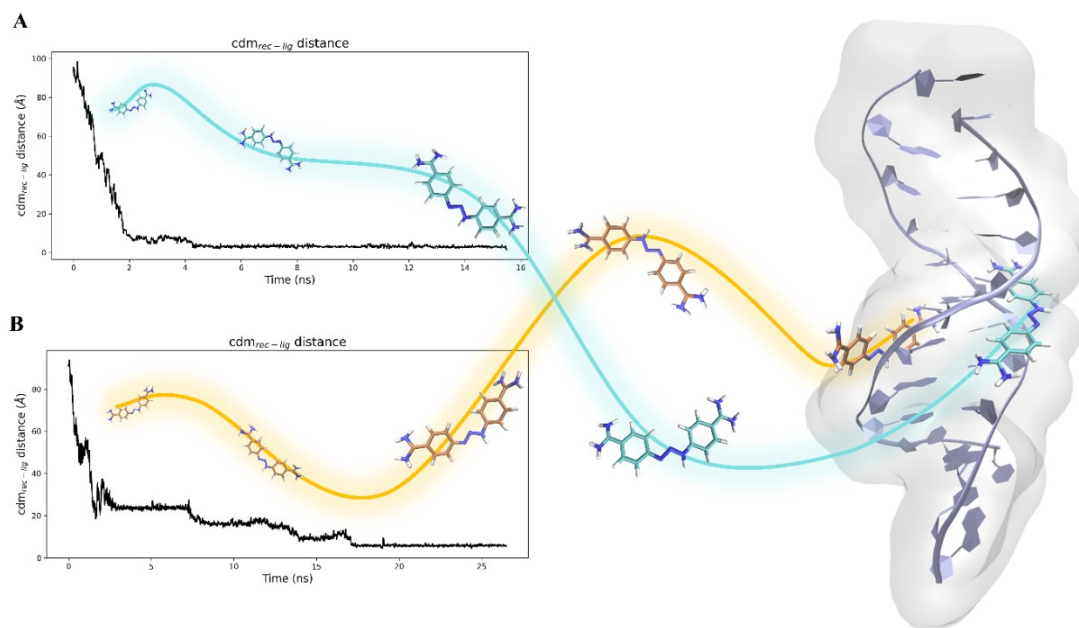


Figure 3. Schematic representation of the binding trajectory, compared between the minor-groove SuMD simulation (blue) and the major-groove one (orange). Panel A represents the approach of the ligand to the binding site in terms of distance to the center of mass of the DNA minor-groove binding site. Panel B represents the same for the major-groove binding site.

The ligand undergoes a progressive reorientation during the binding trajectory towards the AT-rich binding site. Four predominant DNA-ligand populations are observable in the energetic profile reported in Figure 4 and Video S1. The comparison of the of minor-groove and major-groove binding trajectories and their interactions analysis are available in Video S1 in the Supplementary information.

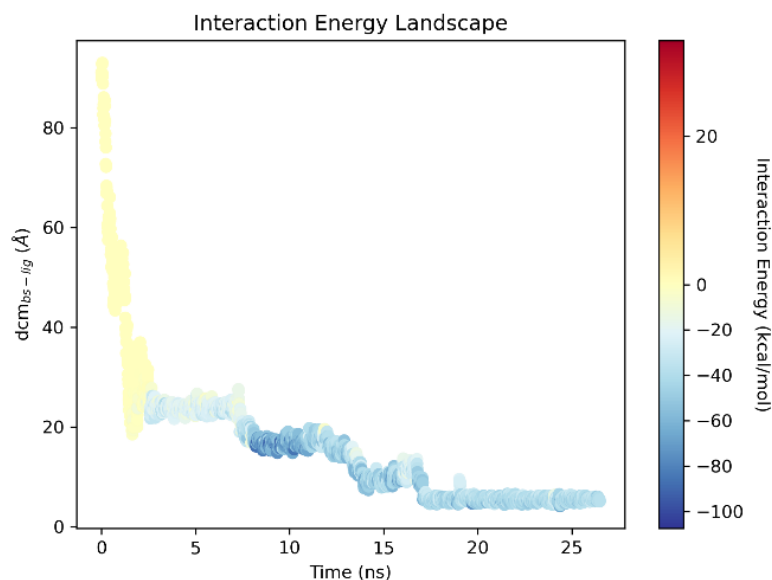


Figure 4. Illustration of the Interaction Energy profile of the major groove molecular dynamics SuMD simulation.

The final frame of the generated binding trajectories revealed favorable interactions with water molecules in both minor and major grooves. A comparison of these binding modes and their heat maps with the crystallographic ones is reported in Figure 5. The water molecules highlighted in the reported binding

modes of the last frame of the molecular dynamics (panels A and B) are also present at the crystallographic level. The heat maps illustrate the mediation of favorable electrostatic interactions with the ligand carried by the water molecule. This data further confirms the crucial arrangement of water during the dynamics, observable even in the deposited crystal, in which water bridges interactions between the ligand and the nucleic acid. The interactive profile of the crystallographic binding mode (panel B) mirrors the one generated from the final frame of the dynamics in the minor groove (panel A), both in terms of electrostatic and hydrophobic interactions. For the major groove binding mode, the heatmap reveals reduced electrostatic interaction with the water molecule and, compared to the other two heatmaps (panel C), significant hydrophobic interaction differences, with several points of interaction missing.

In parallel with the exploration of the binding trajectories, we conducted a molecular dynamics study of the unbinding event using Thermal Titration Molecular Dynamics (TTMD) to evaluate the stability of a ligand-nucleic acid complex. This technique has been successfully applied to both protein and RNA targets^{11,12,16}. For both compounds' binding modes, five simulations were conducted, starting at 250 K, increasing the temperature until the binding mode was lost, resulting in the ligand exiting the binding pocket. The starting point for the molecular dynamic simulations was the last frame of the previously obtained SuMD run. The analysis of the trajectories is shown in Figure 6.

The comparison of unbinding trajectories and their interaction analysis are available in Video S2 in Supplementary information. For the minor groove-bound ligand (Figure 6, panel A), the pose remains stable throughout the temperature gradient explored, for the large part of the simulation.

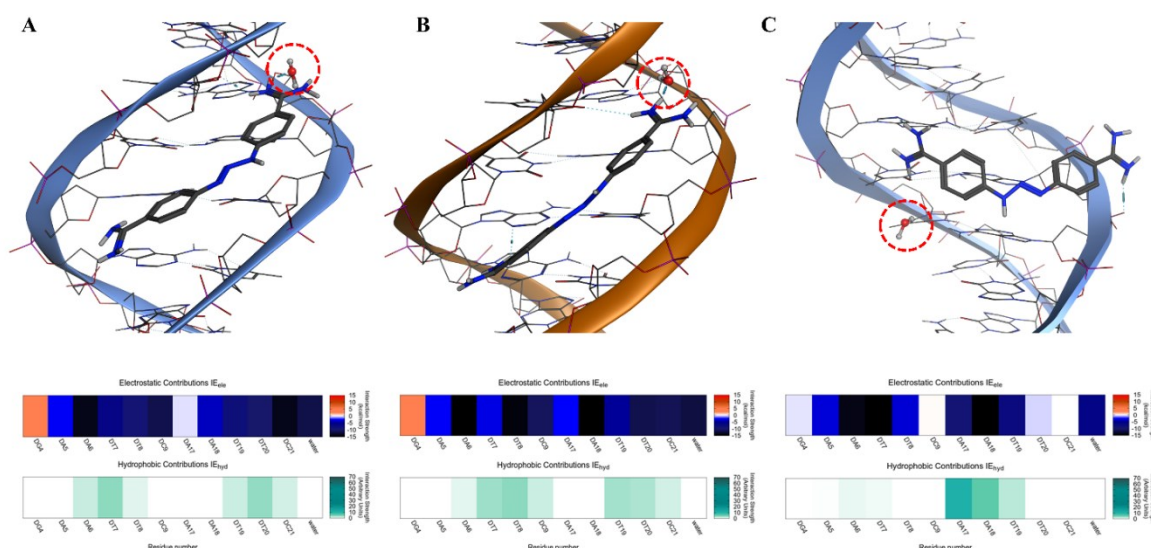


Figure 5. The proposed binding mode of the minor-groove last frame of the molecular dynamic simulation (panel A), crystallographic binding mode (panel B), and major-groove last frame of the molecular dynamic simulation (panel C); protein residues mainly involved in binding are shown as sticks. The water molecules involved in the interactions are highlighted with a red circle.

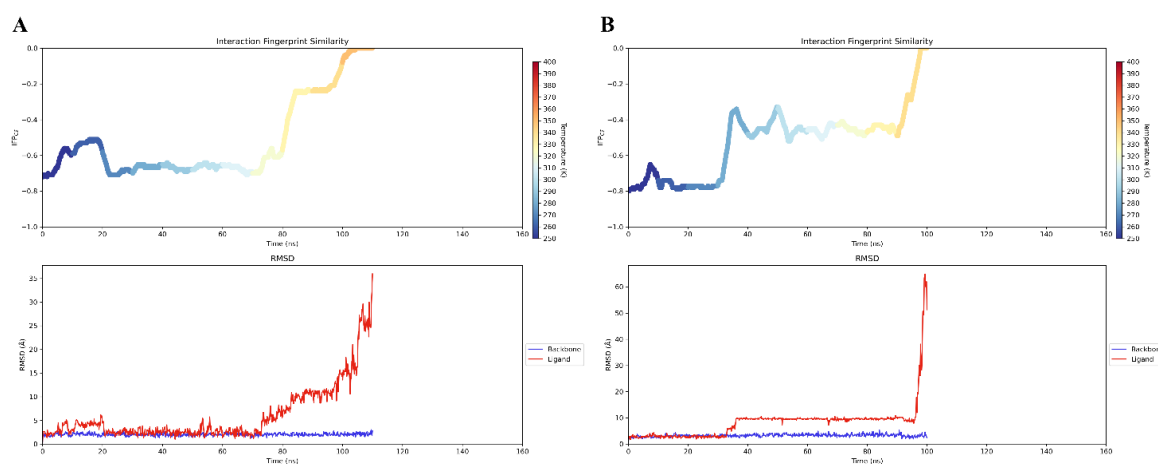


Figure 6. Interaction Fingerprint Similarity and RMSD values of minor-groove last frame of the SuMD simulation (panel A). Interaction Fingerprint Similarity and RMSD values of major-groove last frame SuMD simulation (panel B).

The Interaction Fingerprint Similarity (IFPCs) values compared to the initial reference, starting at a value below -0.7, remain constant throughout the temperature gradient explored until the ligand exits the binding site, at 320-330 K. There is only a first reassessment of the ligand in the first 20 ns of the simulation, with a slight loss of the IFPCs. This steady trend is also reflected in the ligand's RMSD values, which remain stable during the simulation. On the contrary, the major groove-bound ligand (Figure 6, panel B) shows a loss of interaction profile at lower temperatures. After 30 ns of simulation, the ligand loses its starting binding mode, indicating that it may have a shorter residence time, further confirming the preference for the minor groove. The reasons behind this propensity may be sought in a better shape complementarity as discussed in the original work of Professor Neidle, and in a steadier interaction between the nucleic structure and water molecules that can stabilize the binding mode. Considering that not only the nucleic-ligand interactions but also the solvent-mediated ones are crucial for a comprehensive understanding of both the binding and unbinding event, an insight into the stability of water molecules was carried out. For this purpose, in addition to the fingerprint similarity analysis, the AquaMMapS tool was used to map stable waters throughout the TTMD runs.

It is important to stress the fact that during the simulations and even in the equilibration phase, all the waters were not constrained. This approach facilitates the system's proper solvation and, as a result, among the most stable waters, the ones occupying a "crystal-like" position, were obtained spontaneously. Despite the rising temperatures and the exposure of the binding sites to the solvent, AquaMMapS revealed the persistence of the crystallographic water molecules identified along the unbinding trajectory during the Thermal Titration Molecular Dynamics (TTMD) simulation and also observable in the abovementioned last frame of SuMD simulation (Figure 5 panel A and C). In the selected TTMD replica of the ligand in the minor groove, two stable water molecules were identified in the position occupied by the crystallographic one, exhibiting a B-factor value of approximately 21, as shown in Figure 7 (panel A). Similarly, in the simulation of the ligand in the major groove, a stable water molecule was detected with a B-factor of approximately 25, as depicted in Figure 7 (panel B).

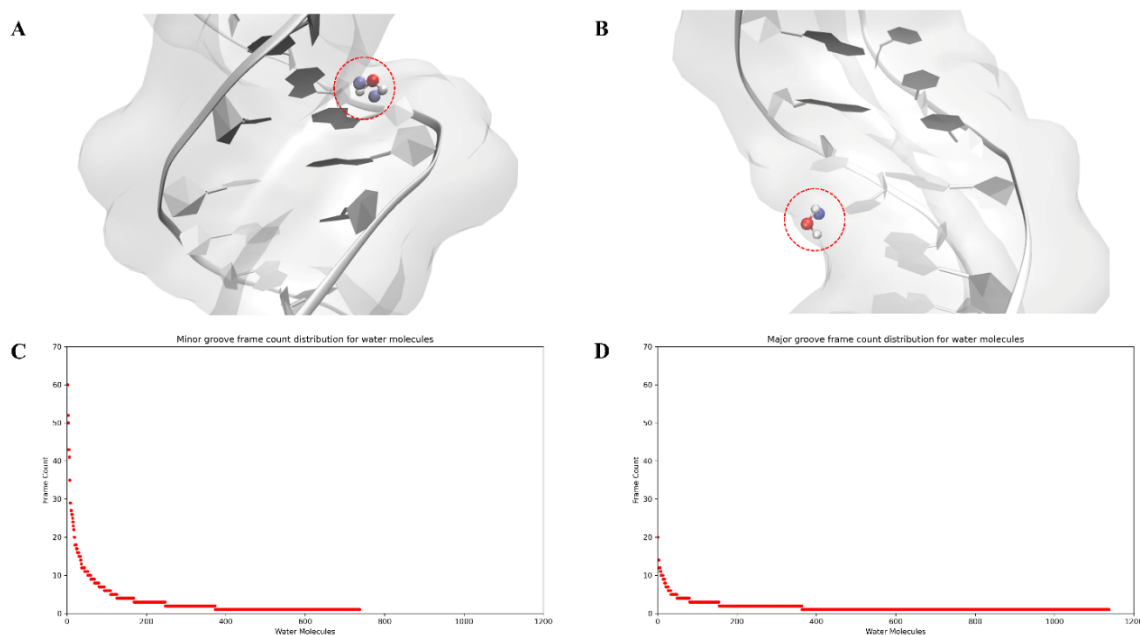


Figure 7. Representation of stable-waters predicted by AquaMMapS for the TTMD simulation on minor-groove (panel A) and major-groove (panel B). In the red circle is represent the crystallographic water molecule, in blue the AquaMMapS's predicted stable water. Panel C and D represent the minor-groove (C) and major-groove (D) frame count distribution of water molecules in the position highlighted in the red circle during the TTMD simulation.

The data highlights the crucial role of these water molecules as emphasized in Professor Neidle's original study. To further characterize the role and location of these water molecules, an in-house Python 3.10 script was used to examine the dynamic profile of water molecules that were observed to interchange in the positions identified via AquaMMapS. This behavior can be examined in Figure 7 (panels C and D). The analysis is specifically focused on counting distinct water molecules that interchange positions within the radius of a single water molecule at the chosen coordinates of the crystallographic waters. For each of these distinct water molecules, the number of simulation frames where they were detected was considered. The lower the number of distinct water molecules interchanging positions during the dynamics, and the more the number of frames where these water molecules are detected, the higher the hydration stability in that position. The results reveal that the count of distinct water molecules capable of interchanging positions in the minor groove is significantly lower compared to those in the major groove (~750 in the minor groove versus ~1200 in the major groove), with a notably increased frame count favoring the minor groove. This implies that these water molecules might be much more stable in the selected position in the minor groove than the major one.

Conclusion

In conclusion, the study provides significant insights into the dynamics of ligand-nucleic acid interactions. The simulations conducted revealed the importance of the AT-rich binding site and the role of water molecules in mediating interactions between the ligand and the nucleic acid. Thermal Titration Molecular Dynamics (TTMD) simulations further confirmed the stability of these interactions, with the minor groove-bound ligand exhibiting greater stability compared to the major groove-bound ligand.

The analysis showcased the stability of crystallographic water molecules during the binding trajectory, despite the rising temperatures and the exposure of binding sites to the solvent. The study provided a deeper understanding of the dynamic profile of these molecules, revealing that fewer distinct waters interchanged positions in the minor groove compared to the major groove, suggesting higher hydration stability in the minor groove.

This work enriches the knowledge present in the field of ligand-nucleic acid interactions, specifically

widening the expertise on the Berenil-DNA complex. In particular, the critical role of water molecules and the relevance of investigating their dynamic behavior have been pointed out, leaving precious considerations for further studies addressing this target and nucleic acids in general. These findings indeed offer valuable insights for future research endeavors, especially in investigating the potential application of Berenil or similar ligands to RNA duplex targets. They open new possibilities for probing nucleic acids' behavior and enhancing understanding of their role in various biological processes.

Materials and Methods

Hardware overview

Molecular modeling operations for the structure preparation and the setup for molecular dynamics (MD) were carried out on a Linux workstation (Linux 20.04) comprising a 25-core Intel Core i9-9820X 3.3 GHz processor. Molecular dynamics simulations (TTMD and SuMD) were computed on a GPU cluster composed of 30 NVIDIA devices ranging from GTX980 to RTX4090.

Structure preparation

The experimentally determined three-dimensional coordinates of the Berenil-Dodecanucleotide complex were retrieved from the Protein Data Bank ² (PDB ID: 2DBE ¹). The structure was prepared by exploiting several tools from the Molecular Operating Environment (MOE) 2022.02 suite ¹⁸. In particular, the "Structure Preparation" tool was adopted to check and fix the inconsistencies deriving from the original structure. Then, each water molecule was removed, and the "Protonate3D" tool was used to add missing hydrogens according to the prevailing tautomeric and protomeric state (pH = 7.4, T = 310 K, i.f. = 0.154). Finally, hydrogen atoms were energy minimized, and partial charges were assigned according to the AMBER14:ETH force field. The prepared structure was kept for molecular dynamics and docking calculations. The Berenil ligand underwent a comprehensive preparation process for computational analysis, utilizing various tools from the OpenEye suite ¹⁹. Three-dimensional coordinates of the molecule was generated with Omega ²⁰, and the protonation states of ionizable groups were then determined with Fixpka ²¹. Lastly, partial charges were assigned to the atoms within the ligand database thanks to Molcharge ²².

System setup and Equilibration Protocol

The systems obtained were processed for MD simulations, exploiting several tools from Visual Molecular Dynamics (VMD) 1.9.2 ²³, and Ambertools22 ²⁴. The parameters for each nucleic atom were assigned according to the DNA OL15 force field ²⁵. The ligand was instead parametrized according to the General Amber force field (GAFF) ²⁶. Each system was solvated in a rectangular box using the TIP3P water model ²⁷, keeping a padding of 15 Å. The net charge of the system was then neutralized through the addition of Na⁺ and Cl[−] ions, reaching a salt concentration of 0.154 M. Afterwards, 500 steps of energy minimization through the conjugate gradient algorithm were performed to reduce clashes and bad contacts.

Before entering the production phase, a two-step equilibration protocol was carried out.

The first equilibration stage of 0.1 ns was performed in the canonical ensemble (NVT), applying a 5 kcal mol^{−1} Å^{−2} harmonic positional restraint to each DNA and ligand atom. The second equilibration consists of a 0.5 ns simulation in the isothermal-isobaric ensemble (NPT), where the same force constant is used to constrain only the ligand and the backbone of the DNA, and the pressure is fixed at 1 atm through a Monte Carlo barostat ²⁸. The temperature for all the equilibration phases was kept constant through a Langevin thermostat ²⁹, keeping the value at 310 K for SuMD and 350 K for TTMD.

Each MD simulation was conducted by exploiting the ACEMD 3.5 engine ³⁰ based on the open-source Python library OpenMM ³¹. An integration timestep of 2 fs was used, and the M-SHAKE algorithm was applied to constrain the length of the bonds involving hydrogen atoms. The particle-mesh Ewald method ³² was exploited to compute electrostatic interactions using cubic spline interpolation. A 9.0 Å cutoff was applied for the calculation of Lennard-Jones interactions.

SuMD

Supervised Molecular Dynamics (SuMD)¹⁰ is an enhanced sampling MD method used to explore ligand-target recognition pathways on the nanosecond timescale. This unbiased technique allows for the sampling of infrequent events, such as the binding one, reducing the simulation time and related costs of classical MD. SuMD is based on a series of short MD simulations followed by an evaluation of the simulation progress by a tabu-like algorithm. Every subsequent short simulation defined as a SuMD-step is computed in the canonical ensemble at 310 K for 500 ps.

The distance between the ligand and the binding site is monitored during each step, specifically the distance between the center of mass of the ligand and the center of mass of the user-defined binding site residues is calculated and fitted into a linear function. If the slope of the fitted line is negative, the ligand is approaching the binding site, and the step is considered productive and retained to generate the final trajectory. In the opposite scenario, whenever the slope is positive, indicating the distancing between the ligand and the binding site, the step is discarded and repeated, exploiting the Langevin thermostat to randomly reassign particle velocities, starting from the final coordinates of the previous SuMD-step. The supervision algorithm operates until the distance between the centers of mass reaches a threshold value (9 Å, in this case). As the distance decreases, once it crosses the threshold value, the algorithm is switched off, and the simulation proceeds further with 10 ns of classical MD.

TTMD

Thermal Titration Molecular Dynamics (TTMD)^{11,12} is an enhanced sampling MD Python tool useful for the evaluation of complex unbinding kinetics. This technique, initially developed to compare protein-ligand complexes considering the persistence of the native intermolecular interactions, is now applicable to nucleic acid-small molecule complexes.

TTMD is based on a series of short classical MD simulations performed at progressively increasing temperatures, defined as TTMD-steps. The user defines the length of each step and the total explorable temperature ramp. Throughout the simulation, the program checks for changes in the native binding mode of the ligand through the evaluation of the interaction fingerprints of each frame. If the original binding mode is retained, the simulation proceeds until the end temperature is reached; on the contrary, if the binding mode is lost during a TTMD-step, the simulation is stopped.

In this work, the starting temperature, is 250 K, while the end temperature is 400 K. The temperature increase between every TTMD-step is 10 K, and the length of each simulation window is 10 ns. ProLIF³³ interaction fingerprint Python package has been exploited to compute the interaction fingerprint of the ligand for each frame, allowing the evaluation of the binding mode retention operated by a fingerprint-based scoring function (IFP_{CS})³⁴.

Trajectory analysis

SuMD. SuMD trajectories were carried out exploiting the 'NAMD Energy' plugin for VMD³⁵. The interaction energy, defined as the sum of van der Waals and electrostatic contributions, was calculated through this package according to the force field.

TTMD. During a TTMD simulation, every frame generated is computed with the ProLIF Python package, and the interaction fingerprint of each frame is compared to the last frame of the second equilibration. The cosine similarity metric is applied to evaluate the persistence of the original binding mode throughout the simulation. The value obtained is multiplied by -1, following the scale of most scoring functions. The interaction fingerprint cosine similarity (IFP_{CS}) is a value that ranges from -1 to 0, where -1 highlights the persistence of the original binding mode. On the other hand, 0 indicates that the native binding mode is lost. The IFP_{CS} is computed for the last 10% of each TTMD-step, if the value is above -0.05 the simulation is stopped as all the native interactions are lost.

The average IFP_{CS} score for each TTMD-step is plotted against the temperature in the "Titration Profile" plot. In this graph, the slope of the straight line that connects the first and last IFP_{CS} point, defined as the "MS coefficient" is a proxy measure for the stability of the complex.

This coefficient can range from 0 to 1. Specifically, 0 is indicative of a lasting conservation of the binding mode, and 1 indicates a labile binding mode.

The mathematical formulation of the MS score is reported in the equation below.

$$MS = \frac{\text{mean } IFP_{CS}^{Tend} - (-1)}{T^{end} - T^{start}}$$

For both the minor and major groove complexes, five independent TTMD simulations were carried out. Following the calculated MS score, the two simulations with the highest and lowest values were discarded, and the mean MS value was calculated among the three remaining trajectories, leading to the selection of the one with the closest value to the mean.

A second graph, defined as the “Titration timeline” is generated. This plot displays the time-dependent evolution of the IFP_{CS} scores throughout the simulation and both the ligand and the DNA RMSD computed by exploiting the MDAnalysis Python package^{36,37}.

The “Titration Timeline” plot reports the time-dependent evolution of the ligand, binding site, and receptor’s backbone RMSD and the IFP_{CS} score.

AquaMMapS. AquaMMapS¹³ is a Python tool developed to map hydration sites according to the occupancy rate during MD simulations. This program performs a grid-based inspection of the frequency of occupation of the water molecules by creating cells designed to accommodate only one water molecule. For every cell, the occupancy value is computed as the ratio between the number of frames in which a water molecule is present and the total number of frames. A second occupancy value is calculated as the ratio between the number of frames in which the cell is occupied by a stationary water molecule (RMSF below 1.4), and the total frames.

Supplementary material

[Video S1](#): Comparison of SuMD minor-groove and major-groove binding trajectory of Berenil;

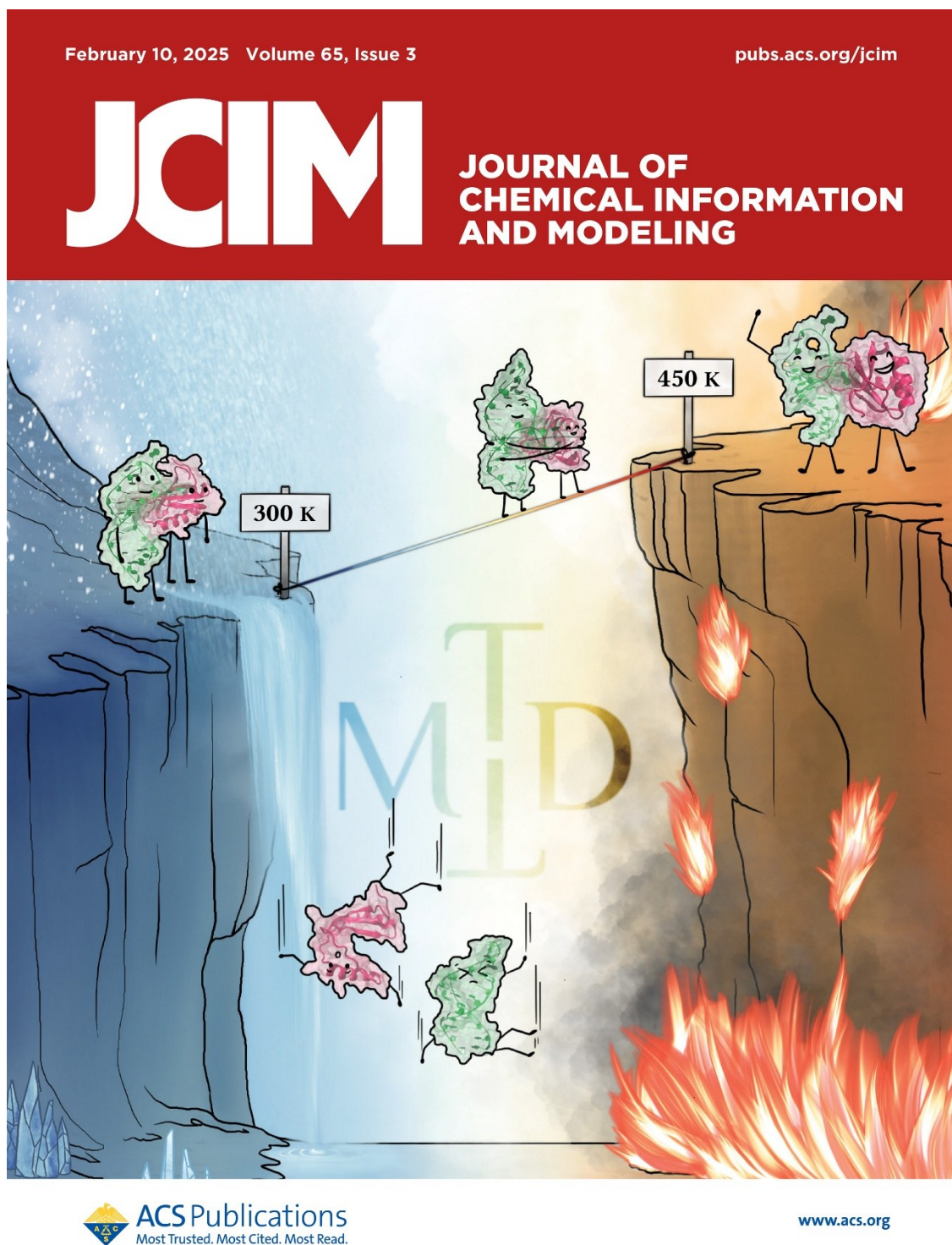
[Video S2](#): Comparison of TTMD unbinding molecular dynamics trajectories of minor-groove and major-groove SuMD last-frame binding modes of Berenil.

[Supplementary Information](#)

References

1. Brown, D. G.; Sanderson, M. R.; Skelly, J. V.; Jenkins, T. C.; Brown, T.; Garman, E.; Stuart, D. I.; Neidle, S. Crystal Structure of a Berenil-dodecanucleotide Complex: The Role of Water in Sequence-specific Ligand Binding. *EMBO J.* 1990, 9 (4), 1329–1334. <https://doi.org/10.1002/j.1460-2075.1990.tb08242.x>.
2. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* 2000, 28 (1), 235–242. <https://doi.org/10.1093/nar/28.1.235>.
3. De Clercq, E.; Dann, O. Diaryl Amidine Derivatives as Oncoviral DNA Polymerase Inhibitors. *J. Med. Chem.* 1980, 23 (7), 787–795. <https://doi.org/10.1021/jm00181a016>.
4. Kuriakose, S.; Uzonna, J. E. Diminazene Aceturate (Berenil), a New Use for an Old Compound? *Int. Immunopharmacol.* 2014, 21 (2), 342–345. <https://doi.org/10.1016/j.intimp.2014.05.027>.
5. Pilch, D. S.; Kirolos, M. A.; Liu, X.; Plum, G. E.; Breslauer, K. J. Berenil [1,3-Bis(4'-Amidinophenyl)Triazene] Binding to DNA Duplexes and to a RNA Duplex: Evidence for Both Intercalative and Minor Groove Binding Properties. *Biochemistry* 1995, 34 (31), 9962–9976. <https://doi.org/10.1021/bi00031a019>.
6. Portugal, J.; Waring, M. J. Comparison of Binding Sites in DNA for Berenil, Netropsin and Distamycin. *Eur. J. Biochem.* 1987, 167 (2), 281–289. <https://doi.org/10.1111/j.1432-1033.1987.tb13334.x>.
7. Gresh, N.; Pullman, B. A Theoretical Study of the Nonintercalative Binding of Berenil and Stilbamidine to Double-Stranded (dA-dT)_n Oligomers. *Mol. Pharmacol.* 1984, 25 (3), 452–458.
8. Gago, F.; Reynolds, C. A.; Richards, W. G. The Binding of Nonintercalative Drugs to Alternating DNA Sequences. *Mol. Pharmacol.* 1989, 35 (2), 232–241.
9. Pearl, L. H.; Skelly, J. V.; Hudson, B. D.; Neidle, S. The Crystal Structure of the DNA-Binding Drug Berenil: Molecular Modelling Studies of Berenil-DNA Complexes. *Nucleic Acids Res.* 1987, 15 (8), 3469–3478. <https://doi.org/10.1093/nar/15.8.3469>.
10. Sabbadin, D.; Moro, S. Supervised Molecular Dynamics (SuMD) as a Helpful Tool To Depict GPCR–Ligand Recognition Pathway in a Nanosecond Time Scale. *J. Chem. Inf. Model.* 2014, 54 (2), 372–376. <https://doi.org/10.1021/ci400766b>.
11. Menin, S.; Pavan, M.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement. *Int. J. Mol. Sci.* 2023, 24 (4), 3596. <https://doi.org/10.3390/IJMS24043596/S1>.
12. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Qualitative Estimation of Protein-Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2022, 62 (22), 5715–5728. https://doi.org/10.1021/ACS.JCIM.2C00995/SUPPL_FILE/Ci2C00995_SI_005.MP4.

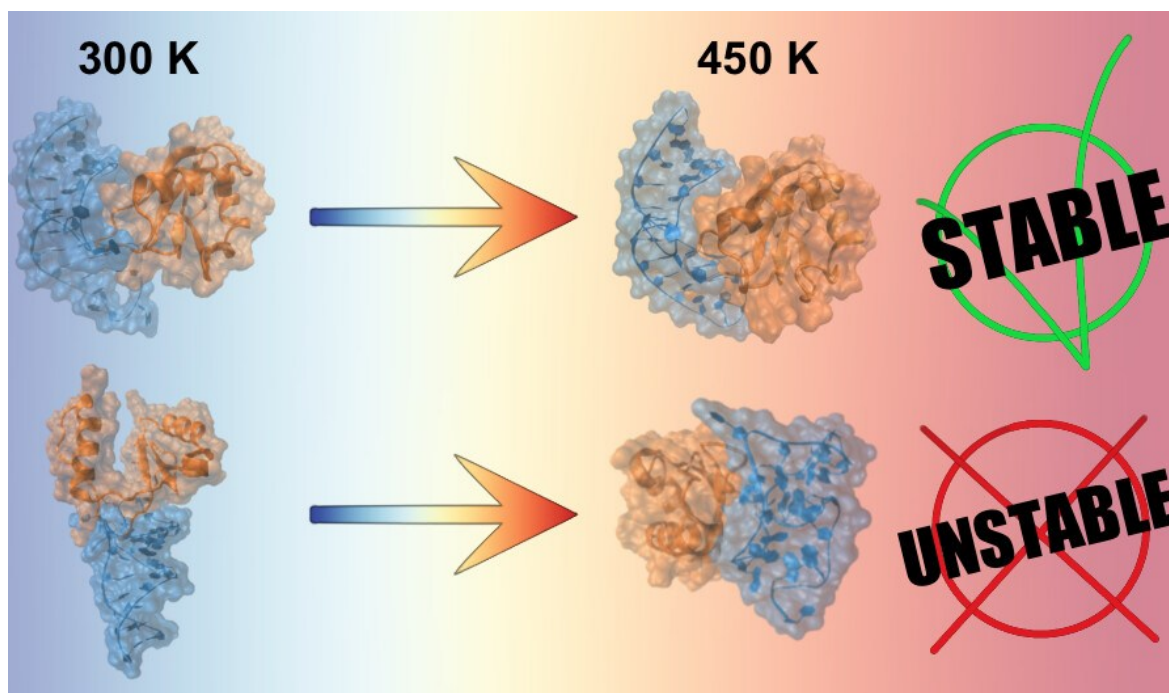
13. AquaMMapS: An Alternative Tool to Monitor the Role of Water Molecules During Protein–Ligand Association - Cuzzolin - 2018 - ChemMedChem - Wiley Online Library. <https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.201700564> (accessed 2024-03-26).
14. Bissaro, M.; Sturlese, M.; Moro, S. Exploring the RNA-Recognition Mechanism Using Supervised Molecular Dynamics (SuMD) Simulations: Toward a Rational Design for Ribonucleic-Targeting Molecules? *Front. Chem.* 2020, 8, 107. <https://doi.org/10.3389/fchem.2020.00107>.
15. Pavan, M.; Bassani, D.; Sturlese, M.; Moro, S. Investigating RNA–Protein Recognition Mechanisms through Supervised Molecular Dynamics (SuMD) Simulations. *NAR Genomics Bioinforma.* 2022, 4 (4). <https://doi.org/10.1093/NARGAB/LQAC088>.
16. Dodaro, A.; Pavan, M.; Menin, S.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Shedding Light on the Stability of RNA–Small Molecule Complexes. *Front. Mol. Biosci.* 2023, 10, 1294543. <https://doi.org/10.3389/fmolb.2023.1294543>.
17. Bissaro, M.; Federico, S.; Salmaso, V.; Sturlese, M.; Spalluto, G.; Moro, S. Targeting Protein Kinase CK1δ with Riluzole: Could It Be One of the Possible Missing Bricks to Interpret Its Effect in the Treatment of ALS from a Molecular Point of View? *ChemMedChem* 2018, 13 (24), 2601–2605. <https://doi.org/10.1002/cmdc.201800632>.
18. Molecular Operating Environment (MOE), 2022.02. C. C. G. Molecular Operating Environment (MOE), 2022.02. https://www.chemcomp.com/Research-Citing_MOE.htm.
19. Cadence Molecular Sciences. OpenEye. Cadence Molecular Sciences, Santa Fe, NM. <http://www.eyesopen.com>, 2023.
20. Conformer Generation Software | Omega. <https://www.eyesopen.com/omega> (accessed 2024-04-11).
21. FixpKa — Applications. <https://docs.eyesopen.com/applications/quacpac/fixpka/fixpka.html> (accessed 2024-04-11).
22. MolCharge — Applications. <https://docs.eyesopen.com/applications/quacpac/molcharge/molcharge.html> (accessed 2024-04-11).
23. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* 1996, 14 (1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
24. Case, D. A. AmberTools22, 2022.
25. Zgarbová, M.; Šponer, J.; Otyepka, M.; Cheatham, T. E. I.; Galindo-Murillo, R.; Jurečka, P. Refinement of the Sugar–Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. *J. Chem. Theory Comput.* 2015, 11 (12), 5723–5736. <https://doi.org/10.1021/acs.jctc.5b00716>.
26. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 2004, 25 (9), 1157–1174. <https://doi.org/10.1002/JCC.20035>.
27. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 1983, 79 (2), 926–935. <https://doi.org/10.1063/1.445869>.
28. Faller, R.; de Pablo, J. J. Constant Pressure Hybrid Molecular Dynamics–Monte Carlo Simulations. *J. Chem. Phys.* 2002, 116 (1), 55–59. <https://doi.org/10.1063/1.1420460>.
29. Davidchack, R. L.; Handel, R.; Tretyakov, M. V. Langevin Thermostat for Rigid Body Dynamics. *J. Chem. Phys.* 2009, 130 (23). <https://doi.org/10.1063/1.3149788/925280>.
30. Harvey, M. J.; Giupponi, G.; Fabritiis, G. D. ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* 2009, 5 (6), 1632–1639. https://doi.org/10.1021/CT9000685/ASSET/IMAGES/LARGE/CT-2009-000685_0004.JPEG.
31. Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLOS Comput. Biol.* 2017, 13 (7), e1005659. <https://doi.org/10.1371/JOURNAL.PCBI.1005659>.
32. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 1995, 103 (19), 8577–8593. <https://doi.org/10.1063/1.470117>.
33. Bouysset, C.; Fiorucci, S. ProLIF: A Library to Encode Molecular Interactions as Fingerprints. *J. Cheminformatics* 2021, 13 (1), 1–9. <https://doi.org/10.1186/S13321-021-00548-6/FIGURES/4>.
34. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Implementing a Scoring Function Based on Interaction Fingerprint for Autogrow4: Protein Kinase CK1δ as a Case Study. *Front. Mol. Biosci.* 2022, 9, 909499. <https://doi.org/10.3389/FMOLB.2022.909499/BIBTEX>.
35. Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* 2020, 153 (4), 44130. https://doi.org/10.1063/5.0014475/16709547/044130_1_ACCEPTED_MANUSCRIPT.PDF.
36. Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domański, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python Sci. Conf.* 2016, 98–105. <https://doi.org/10.25080/MAJORA-629E541A-00E>.
37. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* 2011, 32 (10), 2319–2327. <https://doi.org/10.1002/JCC.21787>.



Post-Docking Refinement of Peptide or Protein-RNA Complexes Using Thermal Titration Molecular Dynamics (TTMD): A Stability Insight

Andrea Dodaro, Gianluca Novello, Silvia Menin, Chiara Cavastracci Strascia, Veronica Salmaso, Stefano Moro*

Dodaro, A.; Novello, G.; Menin, S.; Cavastracci Strascia, C.; Sturlese, M.; Salmaso, V.; Moro, S. Post-Docking Refinement of Peptide or Protein-RNA Complexes Using Thermal Titration Molecular Dynamics (TTMD): A Stability Insight. *J. Chem. Inf. Model.* 2025. <https://doi.org/10.1021/acs.jcim.4c01393>.



Abstract

RNA-protein interactions drive and regulate fundamental cellular processes like transcription and translation. Despite being still limited, the growing body of structural data significantly contributes to the characterization of these interactions. However, RNA complexes involving proteins or peptides are not always available due to the structural determination challenges that this biopolymer entails. Consequently, modeling approaches like molecular docking are exploited to generate complexes relevant to structural and pharmaceutical purposes, including analysis of putative drug targets. Docking methods, despite their widespread adoption, are often hindered by limitations in scoring accuracy, which affects the ranking of the generated poses. Post-docking refining methods, including Molecular Dynamics (MD) approaches, have been developed to tackle this issue. Thermal Titration Molecular Dynamics (TTMD) is an enhanced sampling Molecular Dynamics technique that has been previously effectively applied to refine protein or RNA-small molecule docking poses. This study presents the first application of TTMD to RNA-peptide complexes, validating this method on more complex systems and extending its applicability domain. Our findings showcase the capability of this technique to refine peptide-RNA docking poses, correctly identifying native binding modes among decoys for different pharmaceutically relevant targets.

Introduction

Ribonucleic acid (RNA) is a crucial molecule within the cell, commonly considered a bridging element between genetic information and protein expression. Beyond its canonical “Central Dogma” role, RNA has emerged as a versatile biopolymer involved in several cellular processes like RNA transcription, translation,

and splicing^{1,2,3}.

Remarkably, only a fraction of the genome is translated into proteins⁴, and non-coding RNAs are now receiving significant attention as their dysregulation is related to conditions like cancer and neurodegenerative diseases⁵.

Various proteins can bind coding or non-coding RNA⁶, with RBP (RNA Binding Proteins) being about 7.5% of all proteins⁷, and flaws in protein-RNA interactions are found in various diseases, ranging from cancer to neurological disorders^{8,9}. For these reasons, studying the interactions between RNA and proteins could give insights into cellular processes and could be strategic for identifying new therapeutics.

An increasing, yet limited, number of high-resolution structures of RNA-protein complexes (ribonucleoproteins, RNPs) has been solved and deposited in the last decade; however, the experimental determination of these complexes is still a time-consuming and complex task due to the peculiar nature of RNA. High structural flexibility, the negatively charged phosphate backbone, and the transient nature of the interaction involving RNA-protein complexes hamper the resolution of these structures³.

Given the lack of deposited structures, modeling approaches could help predict and study these interactions, with molecular docking being one of the most adopted methods³.

The docking procedure is composed of two stages: ranking and scoring.

The first one is the generation of possible binding modes (poses) of a molecule at a given target by sampling its orientation and conformation, while the latter consists of ranking the generated models using a scoring function¹⁰.

One of the main issues associated with sampling is that the recognition process and the peptide-RNA complex formation often involve conformational changes. While many tools perform rigid body docking, neglecting this critical aspect in favor of speed and reducing computing costs, some semi-flexible programs can address this conformational issue by considering local rearrangements. Still, conformational changes involving backbone and loop rearrangements are considered a critical challenge in modeling macromolecular complexes¹⁰.

On the other hand, pose scoring can suffer from inaccuracies¹¹. Many docking tools were initially conceived or adapted to generate protein-protein complexes, and specific protein-RNA scoring potentials are not always available^{2,12}. A consensus-scoring approach could improve the results³, as no protein-RNA scoring functions can now reach optimal performances. Limited knowledge of protein-RNA interactions restricts the refinement process of these functions. In contrast, the situation is different for protein complexes not involving RNA, as they have been extensively characterized and studied. Many structures are available in the Protein Data Bank (PDB), and the scoring functions used to evaluate these complexes are better tuned¹³.

Given the peculiar charge distribution of RNA, another essential aspect neglected by common docking approaches is the role of water and ions¹⁴, which can be crucial for maintaining structural stability and facilitating RNA-protein interactions¹⁵.

An alternative approach based on Molecular Dynamics (MD) could solve these limitations, considering both RNA and protein flexibility, simulating the system in explicit solvent and ions, thus providing a better model of a biological environment.

However, due to the long computational times, classical MD is not effective in investigating the strength of the interactions between RNA and proteins and observing infrequent events like binding and unbinding. Enhanced sampling MD techniques, conceived to offset the time limitation of classical MD, allow the observation of such phenomena in reduced computational time. Many of these methods rely on biasing the energy of the simulated system¹⁶. Among them, a collective variable (CV) free method called Thermal Titration Molecular Dynamics (TTMD)¹⁷ has been recently applied to study the stability of RNA-small molecule complexes¹⁸. TTMD performs a series of MD simulations at increasing temperatures while monitoring the persistence of the original binding mode through an interaction fingerprint-based scoring function¹⁹: the persistency of the pose discriminates strong binders from weak ones. TTMD is a user-friendly program, and simulations can be easily carried out without a strong modeling background as they are automatically performed.

As a result of the positive outcome of the application of TTMD to RNA-small molecule complexes, this work aims to assess the applicability domain of the technique to study the stability of RNA-peptide complexes. In particular, we evaluated the application of TTMD as a post-docking filter, capable of

discriminating poses with native-like binding modes from decoys based on interaction stability.

Materials and Methods

Hardware overview

The molecular modeling studies, such as the preparation of RNA-peptide/protein complex structures, the setup for MD simulations, and trajectory analyses, were conducted on a 20 CPU Linux workstation equipped with an Intel Core i9-9820X 3.3 GHz processor. Molecular dynamics simulations were carried out on a cluster composed of 30 NVIDIA GPUs ranging from GTX980 to RTX4090.

Structure Preparation

The three-dimensional structures of the RNA-peptide/protein complexes selected for the analysis were retrieved from the Protein Data Bank (PDB)¹⁹. The structures of these complexes are reported in Tables S1-S5 (Supporting Information) with their PDB IDs reported as well in Table 1. These structures were prepared using tools provided by the Molecular Operating Environment (MOE) suite 2022.02²⁰.

The "Structure Preparation" tool of the MOE environment was utilized to address issues within the chosen protein and RNA chains. This preparation process included the addition of missing atoms, capping of N/C-termini with acetyl and N-methylamide respectively, and handling of tautomeric forms and multiple conformations of specific residues and nucleotides. Missing loops were reconstructed using a homology-based approach.

The "Protonate3D" tool was used to add hydrogen atoms and assign ionization states ($\text{pH} = 7.4$, $T = 310 \text{ K}$, $i.f. = 0.154$), determining the protonation states. This tool also calculated the most probable tautomer based on hydrogen-bonding contributions, ensuring the stability of the structure. Partial charges were computed using the AMBER14:EHT force field, and hydrogen atoms added during reconstruction underwent energy minimization using the same force field.

The prepared structures were then used for subsequent molecular dynamics and docking calculations, ensuring comprehensive and accurate preparation for computational analysis.

RNA-Protein Docking

Docking calculations, performed to generate RNA-peptide/protein complexes, were carried out using two protocols: HDockLite 1.1 for rigid docking and HADDOCK 2.5 for the flexible one.

Rigid Docking

HDockLite1.1 is a docking program derived from the HDock server²¹. Unlike the web-based version, the local software does not use a template-based approach. It employs the Fast Fourier Transform (FFT) to sample potential binding modes between proteins and/or nucleic acids. For the FFT-based translational search, an angle interval of 15° was used for rotational sampling, with a spacing of $1.2 \text{ \AA}$. The scoring function is exploited to evaluate each rotation and the binding modes corresponding to the best-scored translations are retained.

Flexible Docking

To incorporate the flexibility of RNA and peptide or protein counterparts, flexible docking was performed using a locally installed version of HADDOCK2.5 was performed.

HADDOCK2.5 is an integrative modeling docking program derived from its web server interface²². It can combine experimental and predicted information to guide the structure prediction of biomolecular complexes. The docking calculation is based on a triple consecutive run of docking using protocols of first rigid-body calculation, semi-flexible refinement, and, in the end, an explicit solvent analysis.

The docking protocol follows the method developed by Charitou and co-workers²³, designating the entire peptide or protein as passive, while defining the RNA interface based on nucleotides within a $4.5 \text{ \AA}$ distance cutoff from the target in the reference PDB structures. The three-stage docking process generates 5000 poses in the initial rigid body minimization stage, followed by 400 poses during the semi-flexible step, terminating with the final refinement in explicit solvent. For the subsequent molecular dynamics (MD)

simulations, we employed complexes generated during the semi-flexible refinement stage, since solvation and refinement were performed following our standard internal protocol for MD simulation.

System Setup for MD Simulations

The RNA-peptide or protein complexes for molecular dynamics (MD) simulations were prepared utilizing Visual Molecular Dynamics (VMD)²⁴ 1.9.3 and AmberTools22^{25,26}. Parameters for RNA atoms were assigned using the ff14SB force field with χ modification tuned for RNA (χ OL3)^{27,28}, while protein and peptide ligands were parameterized employing the ff14SB²⁹ force field. Each system was solvated in an orthorhombic box with a padding of 15 Å between the complex and the box boundary, using the TIP3P water model³⁰. Sodium (Na⁺) and chloride (Cl[−]) ions were added to neutralize the system and achieve a physiological salt concentration of 0.154 M. A 500-step energy minimization using the conjugate gradient algorithm was performed to eliminate clashes and bad contacts.

A two-step equilibration protocol was carried out before the production phase. The first step involved a 1 ns simulation in the canonical ensemble (NVT), with a 5 kcal mol^{−1} Å^{−2} harmonic positional restraint applied to each protein/peptide and RNA atom. The second equilibration step consisted of a simulation of equal length in the isothermal-isobaric ensemble (NPT), where the same restraint was applied only to the RNA and peptide/protein backbone. During the NPT phase, the pressure was maintained at 1 atm using a Monte Carlo barostat³¹, while the temperature, for all the equilibration phases, was kept constant at the initial (i.e. lowest) TTMD temperature value through a Langevin thermostat³².

All MD simulations were conducted using the ACEMD 3.5³³ engine based on the open-source Python library OpenMM³⁴. An integration timestep of 2 fs was employed, and the M-SHAKE algorithm³⁵ was utilized to constrain bonds involving hydrogen atoms. Long-range electrostatic interactions were calculated using the particle-mesh Ewald method³⁶, applying a 9.0 Å cutoff for Lennard-Jones and real-space electrostatic interactions, and a switching distance of 7.5 Å for Lennard-Jones interactions.

TTMD

Thermal Titration Molecular Dynamics (TTMD)^{17,37} is an enhanced sampling molecular dynamics protocol, developed for evaluating complex unbinding kinetics³⁷, initially used for protein-ligand complexes and also applied to RNA-small molecule structures. TTMD operates through a series of short classical MD simulations, performed at progressively increasing temperatures, defined as "TTMD-steps". Thus, the kinetic energy of the system is progressively increased, thereby reducing the simulation time required to observe unbinding events compared to conventional MD simulations.

In this study, TTMD simulations were carried out using Python 3.10, employing libraries such as NumPy, MDAnalysis^{38,39}, and Scikit-learn. The starting temperature was set at 300 K, and the simulation proceeded in 10 K increments up to the final temperature of 450 K. Each simulation window, or TTMD-step, lasted for 10 ns.

Throughout the simulation, the preservation of the native binding mode was assessed using an interaction fingerprint-based scoring function (IFP_{cs})⁴⁰.

TTMD Trajectory analysis

During a TTMD simulation, each frame generated is analyzed using the ProLIF⁴¹ Python package to compute the interaction fingerprint between the two molecular entities of the complex. The fingerprint is then compared for each frame to the last frame of the second equilibration phase using the cosine similarity metric, which helps assess the maintenance of the original binding mode during the simulation. The similarity value is multiplied by -1 to provide a negative score range, where 0 represents a complete deviation from the reference pose, i.e. loss of the reference binding mode. The interaction fingerprint cosine similarity (IFP_{cs}) score thus ranges from -1 to 0, where -1 indicates complete retention of the original binding mode, and 0 signifies a total loss of native interactions. If the IFP_{cs} score for the last 10% of a TTMD-step exceeds -0.05, indicating that all native interactions have been lost, the simulation is stopped. The average IFP_{cs} score for each TTMD-step is plotted against the temperature in the "Titration Profile" graph. The "MS coefficient" is defined by the slope of the line connecting the first and last points of the graph and is an indicator of the complex stability. This coefficient can range from 0 to 1, with 0 representing persistent conservation of the binding mode and 1 indicating an unstable binding mode that

leads to the loss of the initial interactions in the first step of the simulation. The MS score is calculated using the following equation:
Equation 1. MS formula.

$$MS = \frac{mean\ IFP_{CS}^{Tend} - (-1)}{T^{end} - T^{start}}$$

Five independent TTMD simulations were executed for each examined complex, and the average MS coefficient was computed over three replicates after discarding the ones with the highest and the lowest values.

In addition, a "Titration Timeline" plot is generated by the automatic TTMD procedure, illustrating the time-dependent evolution of IFP_{CS} scores throughout the simulation. This plot also includes the root mean square deviation (RMSD) of the protein and RNA backbone, calculated using the MDAnalysis Python package, providing a comprehensive view of the dynamics and stability of the complex during the simulation.

Results

To assess the applicability of TTMD to the study of RNA-peptide complexes, we applied this technique to five test cases, selected on the basis of availability of an experimental structure in the Protein Data Bank (PDB) and spanning over different "ligand" classes, ranging from cyclic peptides to proteins, mainly focusing on peptides. In particular, the test cases included the following complexes: Nun protein-BoxB RNA (PDB ID: 1HJI), TAR binding protein-HIV-1 TAR RNA (PDB ID: 6XH0), 30S ribosomal protein S8-RNA aptamer (PDB ID: 4PDB), REV peptide-RRE RNA (PDB ID: 1ETG) and macrocyclic peptide-pre-miR21 (PDB ID: 5UZZ).

1HJI, 1ETG, and 5UZZ were chosen to assess the technique's performance with RNA-peptide complexes with the latter characterized by a cyclic peptide, while 4PDB and 6XH0 were used to explore the limits of TTMD application with RNA-protein complexes.

A concise outline of the selected cases, summarized in Table 1, is reported hereafter.

Table 1. Summary of the investigated complexes.

PDB code	accession	Peptide/Protein	RNA	Technique	Resolution
1HJI ⁴²		Nun protein	BoxB RNA	NMR	-
6XH0 ⁴³		TAR binding protein	TAR Trans-Activation Response Element	X-Ray	3.1 Å
4PDB ⁴⁴		30S ribosomal protein S8	RNA aptamer	X-Ray	2.6 Å
1ETG ⁴⁵		REV peptide	RRE Rev Responsive Element RNA	NMR	-
5UZZ ⁴⁶		Macrocyclic Peptide (L50)	pre-miR21	NMR	-

Nun protein

Bacteriophages rely on the bacteria machinery to replicate their genomes, synthesize viral proteins, and generate offspring. They can also compete for the same host, and in this scenario some molecular mechanisms enable a lysogenic phage to exclude superinfecting phages. In particular, HK022 and superinfecting λ phage are involved in a well-studied exclusion mechanism, where HK022-encoded Nun protein acts as a transcription terminator for λ early genes, competing with the λ phage-encoded N protein for the major groove-binding of the stem-loop RNA boxB⁴⁷. These complexes have been extensively studied as models to elucidate protein-RNA recognition⁴⁸, as they both share on the protein side an arginine-rich motif (ARM) like other RNA-binding proteins as Tat and Rev⁴⁹. In this paper, a nun/boxB complex, deposited under the PDB ID: 1HJI⁴², has been subjected to docking and TTMD refinement.

TAR and REV

Human immunodeficiency virus of type 1 (HIV-1) is the causative agent of the acquired immunodeficiency syndrome (AIDS). Despite the significant efforts directed to treat and prevent this retrovirus, its eradication is still hampered by the absence of an effective vaccine and by the limited capability of Antiretroviral Therapy (ART) to completely remove the virus from the host⁵⁰, as the presence of latently infected cells could be a source of persistent residual viremia⁵¹ and long term related comorbidities can occur⁵². In this scenario, transcription inhibitors could play a key role in controlling viral reactivation, thus reducing viral reservoir. Among the targets involved in the transcription phase, the TAT-TAR complex is one of the most attractive for drug development⁵³. The elongation of HIV-1 transcription is promoted by host CDK9/cyclin T1, recruited by HIV-1 transactivation response (TAR). TAR is activated by the Tat protein, which is essential in this process as, in its absence, the transcription results in short and abortive transcripts⁵⁴. Not surprisingly, the TAR complexes deposited in the PDB mainly include peptidomimetics and cyclic peptides specifically designed to inhibit the formation of the TAR-Tat complex. In this work, a complex that involves a TAR binding protein, deposited under the PDB ID: 6XH0⁴³, has been selected. This case study, involving a protein-RNA complex, was chosen to assess the applicability of TTMD to protein-RNA complexes, with proteins presenting a larger surface of contact if compared with peptides.

TAR-TAT complex is not the only target that emerged as a potential candidate against the latent form of HIV-1: REV-RRE complex has also gathered significant attention as it is involved in the process that leads to the translation of structural proteins⁵⁵. HIV-1 REV is a regulatory protein that recognizes the Rev response element (RRE), a ~350 nucleotide sequence located in unspliced and incompletely spliced viral mRNA transcripts. The REV-RRE oligomeric recruits host Crm1/Ran-GTP nuclear export machinery to facilitate the export of these mRNAs from the nucleus to the cytoplasm^{56,57}. Subsequently, after the complex dissociation in the cytoplasm, viral RNAs are packaged for new virions or translated into essential viral proteins^{56,58}.

The interaction between a purine-rich bulge stem-loop IIb of RRE and the arginine-rich α -helix binding domain of REV drives the recognition between Rev and RRE. In the present study, a complex available under PDB code 1ETG⁴⁵ and presenting RRE complexed with a REV peptide portion that includes the arginine-rich motif (ARM) has been selected, prepared, and subjected to docking and TTMD pose refinement.

Bacillus anthracis ribosomal protein S8

The bacterial S8 ribosomal protein has been studied and taken as a model for the characterization of RNA-protein interaction. This RNA-binding protein contributes to the correct folding of the central domain of 16S rRNA⁵⁹. Apart from this function in the ribosome assembly, it has a role in the repression of the translation of the genes encoded by the *spc* operon mRNA, including the ones of S8 itself and ten other ribosomal proteins⁴⁴. Although greater affinity is commonly attributed to the 16S rRNA site, the region targeted on the mRNA is very similar⁶⁰, and many interactions between the S8 protein and the two binding sites are the same⁴⁴. The study of the interaction between RNA and S8 has also been addressed through aptamers: Davlieva and coworkers⁴⁴, through a SELEX experiment, discovered an RNA aptamer able to bind the Bacillus anthracis S8 protein with high affinity. They aimed to identify RNA sequences that, while retaining the ability to bind S8 protein, do not present the conserved features of helix 21 (16S rRNA)⁴⁴, and resulted in a valuable insight into the ability of the selected aptamer to adopt a different secondary structure upon binding, stressing the importance of RNA flexibility into RNA-protein recognition. To investigate the impact of RNA flexibility on TTMD performances, we decided to include this complex, registered with the PDB ID: 4PDB, performing both rigid and semi-flexible docking. The results of a prior work from our group also guided the selection of this test case, as a binding simulation between the aptamer and the protein was successfully performed through Supervised Molecular Dynamics (SuMD), with the last frame of the trajectory converging to the crystal structure⁶¹.

Pre-miR21

The overexpression of miR21 is implicated in the development of cancer. In many solid tumors, it is linked

with cell proliferation, differentiation, and apoptosis, contributing to tumor growth, invasion, and metastasis⁶². After the first cleavage of the primary miRNA transcripts inside the nucleus, the pre-miRNA is exported into the cytoplasm, where it is further cleaved at the apical portion of the stem-loop by another RNase III enzyme named Dicer. While the passenger strand of the obtained duplex is degraded, the guide strand forms the miRNA-induced silencing complex⁶³, which targets complementary RNAs, leading to mRNA cleavage, degradation, and translational repression⁶⁴.

For these reasons, miR21 is investigated as a potential tumor biomarker for diagnostic purposes. Moreover, targeting miRNA-21 could provide adjuvant therapy for tumors like glioblastoma, in which this RNA acts like an oncogene capable of blocking apoptosis⁶².

However, inhibition of miR-21 using small molecules has had limited success. A cyclic peptides library screening led to the identification of a cyclic peptide that is able to bind pre-miR21 at the Dicer cleavage site⁴⁶, and the same authors also deposited the structure of the complex under the PDB ID: 5UZZ⁴⁶. In this work, starting from different docking poses generated through HDOCKlite, the TTMD protocol has been applied to a pharmaceutically relevant target like miR21, evaluating the behavior of the technique in the challenging field of cyclic peptides.

For each RNA-peptide complex, different docking poses were generated by exploiting HDOCKlite and HADDOCK programs, and the obtained structures are reported in Table S1-S5 (Supporting Information). These two different algorithms were considered to compare a fully rigid docking procedure (HDOCKlite) to one including some extent of flexibility (HADDOCK). The only exception was represented by structure 5UZZ, characterized by the presence of a cyclic peptide with a D-Proline, that was not easily managed with HADDOCK, which was thus excluded from the comparison.

The five top-scoring poses were selected for each protocol and subjected to TTMD simulations. When the pose with lowest RMSD to the crystallographic reference (i.e. the crystal-like pose) was not included in the top-five structures, a lower-ranked crystal-like pose was retrieved, regardless of its ranking position, and subjected to TTMD as well. The results were compared to reference TTMD simulations performed on the experimental structure, to assess the complexes stabilities and evaluate the capability of TTMD to discriminate crystal and crystal-like poses from decoys.

As an example, the poses for 6XH0 are reported in Table 2, while the docking poses for the remaining test cases are reported in Tables S1-S5 (Supporting Information). In this case, as regards HDOCKlite, the 5 top scoring poses have been selected, with the crystal-like pose being at the top of the ranking. In the case of HADDOCK, the 5 top scoring poses were selected as well, but they did not include the crystal-like pose, which was instead at the 212th position of the ranking, and additionally submitted to TTMD simulations.

The docking pose of 6XH0, obtained exploiting HDOCKlite and HADDOCK, and the structure deposited in the PDB along with its assertion code, are reported. The RNA is colored in Blue while the protein is pictured orange.

The Prolif fingerprint package, according to the experience gained in the previous application of TTMD to RNA-small molecule complexes¹⁸, was adopted to compute the RNA-peptide interaction fingerprints. In TTMD, the cosine similarity (IFP_{cs}) between the interaction fingerprints during the simulation and the reference position (end of equilibration) is used to monitor the persistence of the binding mode during time. In this work, the starting temperature of the simulation was set to 300 K, the end temperature to 450K, and the duration of each TTMD-step to 10 ns with a temperature increase of 10 K between each step. The TTMD MS coefficient, i.e. the slope of the line connecting the IFP_{cs} at the beginning and end of the simulation (and deeply explained in the method section), was adopted to rank the poses and characterize the conservation of the binding mode between the start and the end of the simulation, with lower MS values representing more stable poses.

Tables 3-4 summarize the TTMD results, for each test case, indicating the average MS values calculated from the TTMD simulations, while the whole set of results is reported in Tables S6-S10 (Supporting Information).

Table 2. Docking results of 6XH0.

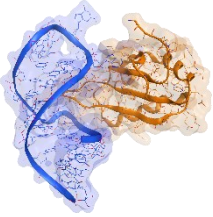
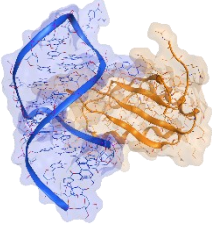
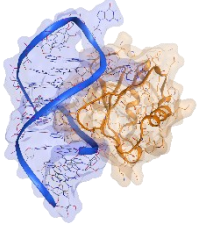
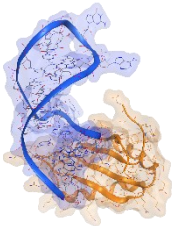
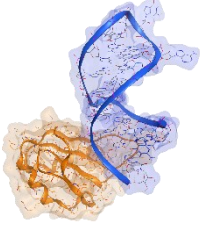
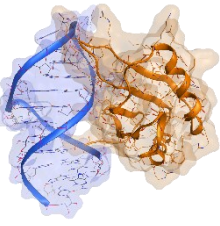
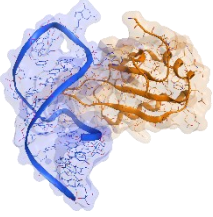
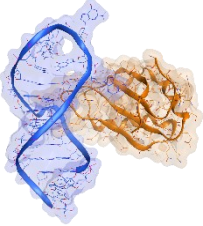
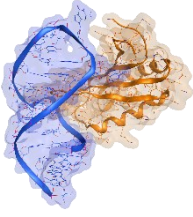
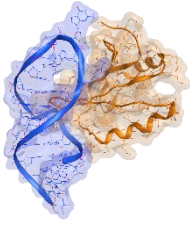
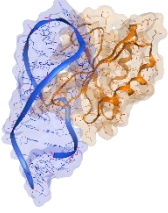
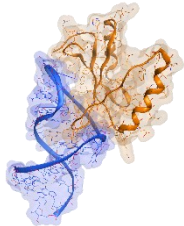
HDOCKlite		
 PDB ID 6XH0	 Pose 1 RMSD 2.81 Å Docking Score - 794.3	 Pose 2 RMSD 25.2 Å Docking Score -735.7
 Pose 3 RMSD 17.6 Å Docking Score -727.8	 Pose 4 RMSD 28.5 Å Docking Score -690.2	 Pose 5 RMSD 32.9 Å Docking Score -678.3
HADDOCK2.5		
 PDB ID 6XH0	 Pose 212 RMSD 2.32 Å Docking Score -55.71	 Pose 1 RMSD 12.9 Å Docking Score -91.97
 Pose 2 RMSD 15.5 Å Docking Score -91.02	 Pose 3 RMSD 21.7 Å Docking Score -89.14	 Pose 4 RMSD 23.8 Å Docking Score -86.74
 Pose 5 RMSD 13.8 Å Docking Score -86.48		

Table 3. TTMD simulation results for each significant replica generated with HADDOCK.

HADDOCK							
6XH0		1ETG		1HJI		4PDB	
Pose	Mean MS	Pose	Mean MS	Pose	Mean MS	Pose	Mean MS
Crystal	0,0038	Crystal	0,0041	Crystal	0,0041	Crystal	0,0032
1	0,0048	1	0,0060	1	0,0066	1	0,0034
2	0,00516	2	0,00558	2	0,00613	2	0,00388
3	0,00529	3	0,00573	3	0,00598	3	0,00417
4	0,00488	4	0,00552	4	0,00648	4	0,00455
5	0,00580	5	0,00465	5	0,00536	5	0,00444
C-like (Pose212)	0,0041	C-like (Pose247)	0,00535	C-like (Pose73)	0,0043	C-like (Pose317)	0,0038

The PDB accession code, the docking program used, and the average MS score calculated, are reported for each test case simulated. The mean MS value for the crystal and crystal-like pose, the docking protocol and the PDB code of each case are highlighted in bold.

Table 4. TTMD simulation results for each significant replica generated with HADDOCK.

HDOCKlite									
6XH0		1ETG		1HJI		4PDB		5UZZ	
Pose	Mean MS	Pose	Mean MS	Pose	Mean MS	Pose	Mean MS	Pose	Mean MS
Crystal	0,0038	Crystal	0,0041	Crystal	0,0041	Crystal	0,0032	Crystal	0,0052
1	0,0038	1	0,0054	1	0,0042	1	0,0041	1	0,0053
2	0,0053	2	0,00589	2	0,00559	2	0,00603	2	0,00563
3	0,00604	3	0,0053	3	0,00517	3	0,0061	3	0,00558
4	0,00532	4	0,00532	4	0,00464	4	0,00512	4	0,00625
5	0,00598	5	0,00484	5	0,0040	5	0,00627	5	0,00551
C-like (Pose1)	0,0038	C-like (Pose52)	0,0045	C-like (Pose1)	0,0042	C-like (Pose1)	0,0041	C-like (Pose1)	0,0053

The PDB accession code, the docking program used, and the average MS score calculated, are reported for each test case simulated. The mean MS value for the crystal and crystal-like pose, the docking protocol and the PDB code of each case are highlighted in bold.

The application of TTMD as a post-docking refinement tool with this set of test cases has brought positive results. In detail, all the crystal structures have lower MS values when compared to the derived docking poses, regardless of the protocol used to generate them. The only exception is represented by pose number 5 (RMSD 5,02 Å) of 1HJI, generated through the HDOCKlite rigid docking approach, which has an MS score slightly lower than the crystal and the crystal-like structure (pose 1, RMSD 1,63 Å), but still comparable to them.

For most of the selected test cases, the technique can assign the lowest MS value, among the different poses, to the crystal-like predicted structures. Detailed information about the docking scores for each simulation and RMSD values for each docking pose can be found in Tables S1-S5 (Supporting information).

Remarkably, even when exploiting a semi-flexible docking protocol, like HADDOCK, TTMD successfully

distinguished crystal-like poses from decoys in all selected cases except for 4PDB and 1ETG.

All HADDOCK poses obtained for 30S ribosomal protein S8 (4PDB) have reasonably low and close RMSD values as shown in Table S4 (Supporting information), and TTMD failed to discriminate their subtle differences. In this situation, TTMD identified the crystal pose as the most stable, but, among the docking poses, the lowest MS was assigned to pose 1 (RMSD of 2,52 Å) rather than pose 317 (RMSD of 1,82 Å). Considering that both RMSD values are below the crystallographic resolution (2,60 Å), the two poses can be described as crystal-like, and TTMD still discriminated them from decoys.

In the case of REV peptide (1ETG), TTMD assigned the lowest MS score to the crystal pose, while the second MS score was not assigned to the crystal-like pose (pose 247), RMSD of 5,62 Å, but to pose 5 (RMSD of 7,28 Å). However, both poses are topologically similar to the crystal structure and are assigned the lowest MS scores among the docking poses.

Despite these situations, the technique was able to identify crystal-like poses for the remaining cases. 6XH0 and 1HJI are particularly significant since, following HADDOCK's ranking, most of the top poses are characterized by high RMSD. Therefore, the application of TTMD and the evaluation of the MS score are valuable in rescoring HADDOCK's poses, identifying the more crystal-like poses regardless the ranking provided by HADDOCK score.

On the other hand, when used starting from the HDocklite-generated poses, TTMD obtained positive results, isolating the crystal-like pose in almost all the cases apart from 1HJI. As mentioned above, the crystal structure, pose 1 (crystal-like pose), and pose 5, share low and close MS score values, 0,0042, 0,00423, and 0,00414 respectively. The technique still manages to identify a stable pose among the others, but it also describes pose number 5 as the most stable. Intriguingly, pose 1 and 5 have the lowest RMSD values among the generated ones. Moreover, as observable in the Titration Timeline graphs Figure S1 (Supporting information), the fluctuation of the IFF_{cs} is higher for pose 5, whereas, for pose number 1 and the crystal reference, it is steadier over time. The IFF score¹⁸ a previously reported additional TTMD scoring metric evaluating the fingerprint fluctuation around the mean IFF_{cs} value, provides further support, suggesting a slightly stabler fingerprint fluctuation for the crystal-like pose, with an IFF value of 0,8 compared to 0,9 attributed to pose 5. The MS score is calculated on the last TTMD-step, as explained in the Materials and Methods section. For this reason, at the end of the simulation, rearrangements can lead the complex to assume native interactions previously lost during the initial part of the simulation, leading to a slightly lower MS score.

Differently from the HADDOCK-originated poses of 1ETG discussed above, in the case of HDocklite, TTMD managed to identify the crystal-like pose among the others even if it was out of the five top scoring poses.

Notably, TTMD was able to seamlessly identify the crystal-like pose among the ones generated for 4PDB, as, in contrast with HADDOCK, they are much more diversified, as can be observed in Table S4 (Supporting Information).

This specific test case was also selected because an enhanced sampling MD technique called Supervised Molecular Dynamics (SuMD) has already been used to simulate the binding event with this test case reaching a near-crystal complex⁶¹. This achievement opens the possibility of using SuMD as an alternative tool for the complex generation, coupling it with TTMD to score the obtained structures, an approach already tested for a Berenil-dodecanucleotide complex⁶⁵.

Starting from the last frame of the SuMD trajectory retrieved from a previous study, we also performed a TTMD simulation as a proof of concept. As shown in Table 5, despite being higher than the crystal one, the average MS score, calculated from the TTMD simulations carried on the last frame of the SuMD trajectory, was lower than those computed from the HDocklite and HADDOCK-generated crystal-like poses. Given the encouraging result, SuMD seems able to generate stable crystal-like complexes. Despite the need for further validation and the long computational times required, this technique could represent a viable alternative to standard docking approaches for complex generation.

Table 5. TTMD simulation results of the last frame from the SuMD trajectory.

Model	Pose	Program	MD1	MD2	MD3	MD4	MD5	TrimMean	TrimError
4PDB	Crystal		0,00303	0,00297	0,00361	0,00302	0,00341	0,0032	0,000182
4PDB	Last frame	SuMD	0,00335	0,00457	0,00416	0,00295	0,00337	0,0036	0,000377
4PDB	1 (C-like)	HDOCKlite	0,00429	0,00418	0,00374	0,00305	0,00458	0,0041	0,000238
4PDB	C-like	HADDOCK	0,00342	0,00405	0,00356	0,00417	0,0037	0,0038	0,000206

* The PDB accession code, the MD protocol or docking program used, the MS score for each simulation, and the average MS coefficient are reported for 4PDB. The replica with the nearest MS value to the average is highlighted in bold.

Discussion And Conclusion

This work showcases the first application of Thermal Titration Molecular Dynamics (TTMD) to RNA-peptide complexes.

This Python-based tool was initially developed for the qualitative estimation of protein-ligand unbinding kinetic thanks to the analysis of the conservation of the interaction fingerprint during a series of consecutive MD simulations performed at increasing temperatures. In this paper, we employed this program, previously adapted to the study of RNA-ligand complexes¹⁸, as a post-docking refinement tool that identifies crystal-like stable poses according to the conservation of the original binding mode, separating them from decoys. Five test cases have been identified, basing the decision on structural differentiation and pharmaceutical relevance, and state-of-the-art docking tools like HDOCKlite and HADDOCK have been used to generate the poses starting from the selected deposited structures. We decided to carry out both rigid and semi-flexible docking approaches to assess whether a fast posing could affect or invalidate the TTMD capability to identify crystal-like poses, rather than a sophisticated yet slower method, like HADDOCK, that keeps into account local flexibility and generates refined docking poses.

In a previous study, RNA-small molecule complexes were characterized through TTMD using different temperatures and alternative Python packages to compute the interactions, resulting in the definition of a protocol feasible for nucleic acids¹⁸. In this work, TTMD simulations were performed adopting the same protocol (see Materials and Methods).

Globally, examining Table S6-S10, TTMD seems capable of refining docking results, being able, in most of the cases, to discriminate crystal or crystal-like poses, from docking poses that consistently differ from experimental structures. As an example of the discriminatory capability of the technique, Videos S1-S4 (Supporting information) report the TTMD simulations of the crystal structure of 6XH0 compared respectively to the following poses of the same test-case: HADDOCK-generated crystal-like pose (lowest RMSD) and pose 5 (highest MS), HDOCKlite-generated pose 1 (lowest RMSD) and pose 5 (highest MS). As can be observed in the videos, the crystal structure and the crystal-like poses maintain structural integrity despite the high temperatures, while the structural instability is particularly amplified by the rising temperatures for non-crystal-like poses.

Subtle differences between the poses, such as sharing the same interactions and having low RMSD, can reduce the technique's resolution, like in the scenario where all the poses under discussion may be described as crystal-like, for instance, poses with RMSD below crystallographic resolution, a situation close to the one reported for 4PDB.

The rigid docking approach seems optimal when no local rearrangements are expected upon binding, and the structural data is available like in the current validation, even considering the short computational times: just a few minutes in the case of HDOCKlite. In fact, with this protocol, for each case a native-like pose with RMSD below 3 Å was obtained, and, except in one case, it coincided with the top-scoring pose. On the other side, in a hypothetical scenario where structure prediction programs are exploited due to the absence of a deposited structure and MD pose refinement is needed⁶⁶, it is not trivial to get such outstanding performances with a rigid docking approach, thus following a semi-flexible docking procedure, like HADDOCK, could be a better choice. Still, it is important to consider that while HADDOCK's posing seems robust and preferable in specific situations, aside from the longer computational times required

(hours), its scoring capabilities are not equally effective, with native-like poses poorly ranked, sometimes even above the 100th position, as shown in Table S6-S10 (Supporting information).

In this circumstance, TTMD could be a critical resource since, as discussed previously, it can discriminate native-like poses from decoys and refine the docking ranking results, at least for the selected test cases.

However, coupling TTMD to a docking protocol like HADDOCK, and applying it to such a high number of docking poses in a prospective study, where no experimental structure is known, is not feasible in a short amount of time, at least with currently available consumer GPUs.

Customizing the parameters used to execute HADDOCK could be beneficial in this scenario, but finding the optimal protocol could be difficult.

In general, rescoring or consensus approaches are an option to reduce the number of poses to attention and refine with TTMD, or as an alternative, the top poses generated through rigid docking, followed by MD relaxation, could be used. Despite the long simulation time, generating only a few complexes through enhanced sampling Molecular Dynamics binding tools like SuMD (Supervised Molecular Dynamics) could also be a viable solution, given the encouraging results in obtaining stable near-native complexes at the final state of the simulation, as mentioned for 4PDB in the Results section.

This work highlights the perks and the limitations of using TTMD to characterize RNA-peptide complexes, including larger systems formed by RNA and proteins, further extending the applicability domain of this method, already applied to the study of RNA-small molecule complexes¹⁸. The positive and encouraging results obtained pave the way for the future practical application of this technique as a post-docking refinement tool. However, implementing this program in a high-throughput pipeline is still hampered by the long computational times required to carry out MD simulations. Further testing is needed to determine a robust docking setup, additional TTMD code optimization and hardware progress are also necessary to reach the computational power required to execute such workflow rapidly. However, the rate of recent tech development suggests that near-future applications of enhanced sampling methods like TTMD on a large scale will be feasible⁶⁷.

Supplementary material

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.4c01393>.

Author Contributions

The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript. A.D. and G.N. contributed equally.

References

1. Mattick, J. S.; Makunin, I. V. Non-Coding RNA. *Hum. Mol. Genet.* 2006, 15 (suppl_1), R17–R29. <https://doi.org/10.1093/hmg/ddl046>.
2. Puton, T.; Kozłowski, L.; Tuszyńska, I.; Rother, K.; Bujnicki, J. M. Computational Methods for Prediction of Protein–RNA Interactions. *J. Struct. Biol.* 2012, 179 (3), 261–268. <https://doi.org/10.1016/j.jsb.2011.10.001>.
3. Nithin, C.; Ghosh, P.; Bujnicki, J. M. Bioinformatics Tools and Benchmarks for Computational Docking and 3D Structure Prediction of RNA-Protein Complexes. *Genes* 2018, 9 (9), 432. <https://doi.org/10.3390/genes9090432>.
4. Zafferani, M.; Hargrove, A. E. Small Molecule Targeting of Biologically Relevant RNA Tertiary and Quaternary Structures. *Cell Chem. Biol.* 2021, 28 (5), 594–609. <https://doi.org/10.1016/J.CHEMBIOL.2021.03.003>.
5. Falese, J. P.; Donlic, A.; Hargrove, A. E. Targeting RNA with Small Molecules: From Fundamental Principles towards the Clinic. *Chem. Soc. Rev.* 2021, 50 (4), 2224–2243. <https://doi.org/10.1039/D0CS01261K>.
6. Guilhot-Gaudeffroy, A.; Froidevaux, C.; Azé, J.; Bernauer, J. Protein-RNA Complexes and Efficient Automatic Docking: Expanding RosettaDock Possibilities. *PLOS ONE* 2014, 9 (9), e108928. <https://doi.org/10.1371/journal.pone.0108928>.
7. Gerstberger, S.; Hafner, M.; Tuschl, T. A Census of Human RNA-Binding Proteins. *Nat. Rev. Genet.* 2014, 15 (12), 829–845. <https://doi.org/10.1038/nrg3813>.
8. Cooper, T. A.; Wan, L.; Dreyfuss, G. RNA and Disease. *Cell* 2009, 136 (4), 777–793. <https://doi.org/10.1016/j.cell.2009.02.011>.

9. Lukong, K. E.; Chang, K.; Khandjian, E. W.; Richard, S. RNA-Binding Proteins in Human Genetic Disease. *Trends Genet.* 2008, **24** (8), 416–425. <https://doi.org/10.1016/j.tig.2008.05.004>.
10. Tuszyńska, I.; Matelska, D.; Magnus, M.; Chojnowski, G.; Kasprzak, J. M.; Kozłowski, L. P.; Dunin-Horkawicz, S.; Bujnicki, J. M. Computational Modeling of Protein–RNA Complex Structures. *Methods* 2014, **65** (3), 310–319. <https://doi.org/10.1016/j.ymeth.2013.09.014>.
11. Chaput, L.; Mouawad, L. Efficient Conformational Sampling and Weak Scoring in Docking Programs? Strategy of the Wisdom of Crowds. *J. Cheminformatics* 2017, **9** (1), 1–18. <https://doi.org/10.1186/S13321-017-0227-X/FIGURES/7>.
12. Jones, S. Protein–RNA Interactions: Structural Biology and Computational Modeling Techniques. *Biophys. Rev.* 2016, **8** (4), 359–367. <https://doi.org/10.1007/s12551-016-0223-9>.
13. Qiu, L.; Zou, X. Scoring Functions for Protein–RNA Complex Structure Prediction: Advances, Applications, and Future Directions. *Commun. Inf. Syst.* 2020, **20** (1), 1–22. <https://doi.org/10.4310/CIS.2020.v20.n1.a1>.
14. Xiao, W.; Wang, D.; Shen, Z.; Li, S.; Li, H. Multi-Body Interactions in Molecular Docking Program Devised with Key Water Molecules in Protein Binding Sites. *Molecules* 2018, **23** (9), 2321. <https://doi.org/10.3390/molecules23092321>.
15. Fingerhut, B. P. The Mutual Interactions of RNA, Counterions and Water – Quantifying the Electrostatics at the Phosphate–Water Interface. *Chem. Commun.* 2021, **57** (96), 12880–12897. <https://doi.org/10.1039/D1CC05367A>.
16. Do, T. N.; Carloni, P.; Varani, G.; Bussi, G. RNA/Peptide Binding Driven by Electrostatics – Insight from Bidirectional Pulling Simulations. *J. Chem. Theory Comput.* 2013, **9** (3), 1720–1730. https://doi.org/10.1021/CT3009914/SUPPL_FILE/CT3009914_SI_001.PDF.
17. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Qualitative Estimation of Protein–Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2022, **62** (22), 5715–5728. https://doi.org/10.1021/ACS.JCIM.2C00995/SUPPL_FILE/C12C00995_SI_005.MP4.
18. Dodaro, A.; Pavan, M.; Menin, S.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Shedding Light on the Stability of RNA–Small Molecule Complexes. *Front. Mol. Biosci.* 2023, **10**.
19. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* 2000, **28** (1), 235–242. <https://doi.org/10.1093/NAR/28.1.235>.
20. Molecular Operating Environment (MOE), 2022.02, C. C. G. Molecular Operating Environment (MOE), 2022.02. https://www.chemcomp.com/Research-Citing_MOE.htm.
21. Yan, Y.; Tao, H.; He, J.; Huang, S.-Y. The HDock Server for Integrated Protein–Protein Docking. *Nat. Protoc.* 2020, **15** (5), 1829–1852. <https://doi.org/10.1038/s41596-020-0312-x>.
22. Honorato, R. V.; Trellet, M. E.; Jiménez-García, B.; Schaarschmidt, J. J.; Giulini, M.; Reys, V.; Koukos, P. I.; Rodrigues, J. P. G. L. M.; Karaca, E.; van Zundert, G. C. P.; Roel-Touris, J.; van Noort, C. W.; Jandová, Z.; Melquiond, A. S. J.; Bonvin, A. M. J. J. The HADDOCK2.4 Web Server for Integrative Modeling of Biomolecular Complexes. *Nat. Protoc.* 2024, **1**–23. <https://doi.org/10.1038/s41596-024-01011-0>.
23. Charitou, V.; van Keulen, S. C.; Bonvin, A. M. J. J. Cyclization and Docking Protocol for Cyclic Peptide–Protein Modeling Using HADDOCK2.4. *J. Chem. Theory Comput.* 2022, **18** (6), 4027–4040. <https://doi.org/10.1021/acs.jctc.2c00075>.
24. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* 1996, **14** (1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
25. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* 2005, **26** (16), 1668–1688. <https://doi.org/10.1002/JCC.20290>.
26. Case, D. A. AmberTools22, 2022.
27. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* 2007, **92** (11), 3817–3829. <https://doi.org/10.1529/BIOPHYSJ.106.097782>.
28. Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E. I.; Jurečka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* 2011, **7** (9), 2886–2902. <https://doi.org/10.1021/ct200162x>.
29. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 2015, **11** (8), 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>.
30. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 1983, **79** (2), 926–935. <https://doi.org/10.1063/1.445869>.
31. Faller, R.; de Pablo, J. J. Constant Pressure Hybrid Molecular Dynamics–Monte Carlo Simulations. *J. Chem. Phys.* 2002, **116** (1), 55–59. <https://doi.org/10.1063/1.1420460>.
32. Davidchack, R. L.; Handel, R.; Tretyakov, M. V. Langevin Thermostat for Rigid Body Dynamics. *J. Chem. Phys.* 2009, **130** (23). <https://doi.org/10.1063/1.3149788/925280>.
33. Harvey, M. J.; Giupponi, G.; Fabritiis, G. D. ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* 2009, **5** (6), 1632–1639. https://doi.org/10.1021/CT9000685/ASSET/IMAGES/LARGE/CT-2009-000685_0004.JPEG.
34. Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLOS Comput. Biol.* 2017, **13** (7), e1005659. <https://doi.org/10.1371/JOURNAL.PCBI.1005659>.

35. Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A Fast SHAKE Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. *J. Comput. Chem.* 2001, **22** (5), 501–508. [https://doi.org/10.1002/1096-987X\(20010415\)22:5<501::AID-JCC1021>3.0.CO;2-V](https://doi.org/10.1002/1096-987X(20010415)22:5<501::AID-JCC1021>3.0.CO;2-V).
36. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 1995, **103** (19), 8577–8593. <https://doi.org/10.1063/1.470117>.
37. Menin, S.; Pavan, M.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement. *Int. J. Mol. Sci.* 2023, **24** (4), 3596. <https://doi.org/10.3390/IJMS24043596/S1>.
38. Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domański, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python Sci. Conf.* 2016, 98–105. <https://doi.org/10.25080/MAJORA-629E541A-00E>.
39. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* 2011, **32** (10), 2319–2327. <https://doi.org/10.1002/jcc.21787>.
40. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Implementing a Scoring Function Based on Interaction Fingerprint for Autogrow4: Protein Kinase CK1δ as a Case Study. *Front. Mol. Biosci.* 2022, **9**, 909499. <https://doi.org/10.3389/FMOLB.2022.909499/BIBTEX>.
41. Bouysset, C.; Fiorucci, S. ProLIF: A Library to Encode Molecular Interactions as Fingerprints. *J. Cheminformatics* 2021, **13** (1), 1–9. <https://doi.org/10.1186/S13321-021-00548-6/FIGURES/4>.
42. Faber, C.; Schärpf, M.; Becker, T.; Sticht, H.; Rösch, P. The Structure of the Coliphage HK022 Nun Protein-λ-Phage *boxB* RNA Complex. *J. Biol. Chem.* 2001, **276** (34), 32064–32070. <https://doi.org/10.1074/jbc.M102975200>.
43. Chavali, S. S.; Mali, S. M.; Jenkins, J. L.; Fasan, R.; Wedekind, J. E. Co-Crystal Structures of HIV TAR RNA Bound to Lab-Evolved Proteins Show Key Roles for Arginine Relevant to the Design of Cyclic Peptide TAR Inhibitors. *J. Biol. Chem.* 2020, **295** (49), 16470–16486. <https://doi.org/10.1074/jbc.RA120.015444>.
44. Davlieva, M.; Donarski, J.; Wang, J.; Shamoo, Y.; Nikonowicz, E. P. Structure Analysis of Free and Bound States of an RNA Aptamer against Ribosomal Protein S8 from *Bacillus Anthracis*. *Nucleic Acids Res.* 2014, **42** (16), 10795–10808. <https://doi.org/10.1093/nar/gku743>.
45. Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R. α Helix-RNA Major Groove Recognition in an HIV-1 Rev Peptide-RRE RNA Complex. *Science* 1996, **273** (5281), 1547–1551. <https://doi.org/10.1126/science.273.5281.1547>.
46. Shortridge, M. D.; Walker, M. J.; Pavelitz, T.; Chen, Y.; Yang, W.; Varani, G. A Macrocyclic Peptide Ligand Binds the Oncogenic MicroRNA-21 Precursor and Suppresses Dicer Processing. *ACS Chem. Biol.* 2017, **12** (6), 1611–1620. <https://doi.org/10.1021/acscchembio.7b00180>.
47. Uc-Mass, A.; Khodursky, A.; Brown, L.; Gottesman, M. E. Overexpression of Phage HK022 Nun Protein Is Toxic for *Escherichia Coli*. *J. Mol. Biol.* 2008, **380** (5), 812–819. <https://doi.org/10.1016/j.jmb.2008.05.030>.
48. Zhang, X.; Lee, S. W.; Zhao, L.; Xia, T.; Qin, P. Z. Conformational Distributions at the N-Peptide/*boxB* RNA Interface Studied Using Site-Directed Spin Labeling. *RNA* 2010, **16** (12), 2474–2483. <https://doi.org/10.1261/rna.2360610>.
49. Watnick, R. S.; Herring, S. C.; Palmer, A. G.; Gottesman, M. E. The Carboxyl Terminus of Phage HK022 Nun Includes a Novel Zinc-Binding Motif and a Tryptophan Required for Transcription Termination. *Genes Dev.* 2000, **14** (6), 731–739. <https://doi.org/10.1101/gad.14.6.731>.
50. Sadowski, I.; Hashemi, F. B. Strategies to Eradicate HIV from Infected Patients: Elimination of Latent Provirus Reservoirs. *Cell. Mol. Life Sci.* 2019, **76** (18), 3583–3600. <https://doi.org/10.1007/s00018-019-03156-8>.
51. Van Lint, C.; Bouchat, S.; Marcello, A. HIV-1 Transcription and Latency: An Update. *Retrovirology* 2013, **10** (1), 67. <https://doi.org/10.1186/1742-4690-10-67>.
52. Alanazi, A.; Ivanov, A.; Kumari, N.; Lin, X.; Wang, S.; Kovalsky, D.; Nekhai, S. Targeting Tat–Tar Rna Interaction for Hiv-1 Inhibition. *Viruses* 2021, **13** (10), 2004. <https://doi.org/10.3390/V13102004/S1>.
53. Mousseau, G.; Mediouni, S.; Valente, S. T. Targeting HIV Transcription: The Quest for a Functional Cure. In *The Future of HIV-1 Therapeutics: Resistance Is Futile?*; Torbett, B. E., Goodsell, D. S., Richman, D. D., Eds.; Springer International Publishing: Cham, 2015; pp 121–145. https://doi.org/10.1007/82_2015_435.
54. Donahue, D. A.; Kuhl, B. D.; Sloan, R. D.; Wainberg, M. A. The Viral Protein Tat Can Inhibit the Establishment of HIV-1 Latency. *J. Virol.* 2012, **86** (6), 3253–3263. <https://doi.org/10.1128/jvi.06648-11>.
55. Pai, S.; Mudgal, J.; Kamath, B. V.; Pai, K. S. R. An Insight on Promising Strategies Hoping to Cure HIV-1 Infection by Targeting Rev Protein—Short Review. *Pharmacol. Rep.* 2021, **73** (5), 1265–1272. <https://doi.org/10.1007/s43440-021-00257-9>.
56. Fernandes, J.; Jayaraman, B.; Frankel, A. The HIV-1 Rev Response Element: An RNA Scaffold That Directs the Cooperative Assembly of a Homo-Oligomeric Ribonucleoprotein Complex. *RNA Biol.* 2012, **9** (1), 6–11. <https://doi.org/10.4161/rna.9.1.18178>.
57. Sherpa, C.; Rausch, J. W.; Le Grice, S. F. J.; Hammarskjöld, M.-L.; Rekosh, D. The HIV-1 Rev Response Element (RRE) Adopts Alternative Conformations That Promote Different Rates of Virus Replication. *Nucleic Acids Res.* 2015, **43** (9), 4676–4686. <https://doi.org/10.1093/nar/gkv313>.
58. Rausch, J. W.; Grice, S. F. J. L. HIV Rev Assembly on the Rev Response Element (RRE): A Structural Perspective. *Viruses* 2015, **7** (6), 3053–3075. <https://doi.org/10.3390/v7062760>.
59. Moine, H.; Cachia, C.; Westhof, E.; Ehresmann, B.; Ehresmann, C. The RNA Binding Site of S8 Ribosomal Protein of *Escherichia Coli*: Selex and Hydroxyl Radical Probing Studies. *RNA* 1997, **3** (3), 255–268.

60. Davies, C.; Ramakrishnan, V.; White, S. W. Structural Evidence for Specific S8–RNA and S8–Protein Interactions within the 30S Ribosomal Subunit: Ribosomal Protein S8 from *Bacillus Stearotherophilus* at 1.9 Å Resolution. *Structure* 1996, 4 (9), 1093–1104. [https://doi.org/10.1016/S0969-2126\(96\)00115-3](https://doi.org/10.1016/S0969-2126(96)00115-3).
61. Pavan, M.; Bassani, D.; Sturlese, M.; Moro, S. Investigating RNA–Protein Recognition Mechanisms through Supervised Molecular Dynamics (SuMD) Simulations. *NAR Genomics Bioinforma.* 2022, 4 (4). <https://doi.org/10.1093/NARGAB/LQAC088>.
62. Huang, Y.; Yang, Y. B.; Zhang, X. H.; Yu, X. L.; Wang, Z. B.; Cheng, X. C. MicroRNA-21 Gene and Cancer. *Med. Oncol.* 2013, 30 (1), 376. <https://doi.org/10.1007/s12032-012-0376-8>.
63. Medley, J. C.; Panzade, G.; Zinovyeva, A. Y. microRNA Strand Selection: Unwinding the Rules. *WIREs RNA* 2021, 12 (3), e1627. <https://doi.org/10.1002/wrna.1627>.
64. Iwakawa, H.; Tomari, Y. Life of RISC: Formation, Action, and Degradation of RNA-Induced Silencing Complex. *Mol. Cell* 2022, 82 (1), 30–43. <https://doi.org/10.1016/j.molcel.2021.11.026>.
65. Novello, G.; Dodaro, A.; Menin, S.; Cavastracci, C.; Sturlese, M.; Salmaso, V.; Stefano, M. A Second Life for the Crystallographic Structure of Berenil-Dodecanucleotide Complex: A Computational Revisitation Thirty Years after Its Publication. April 23, 2024. <https://doi.org/10.21203/rs.3.rs-4269844/v1>.
66. Zhang, Y.; Vass, M.; Shi, D.; Abualrous, E.; Chambers, J. M.; Chopra, N.; Higgs, C.; Kasavajhala, K.; Li, H.; Nandekar, P.; Sato, H.; Miller, E. B.; Repasky, M. P.; Jerome, S. V. Benchmarking Refined and Unrefined AlphaFold2 Structures for Hit Discovery. *J. Chem. Inf. Model.* 2023, 63 (6), 1656–1667. <https://doi.org/10.1021/acs.jcim.2c01219>.
67. Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* 2018, 9 (AUG), 393738. <https://doi.org/10.3389/FPHAR.2018.00923/BIBTEX>.

February 10, 2025 Volume 65, Issue 3

pubs.acs.org/jcim

JCIM

**JOURNAL OF
CHEMICAL INFORMATION
AND MODELING**



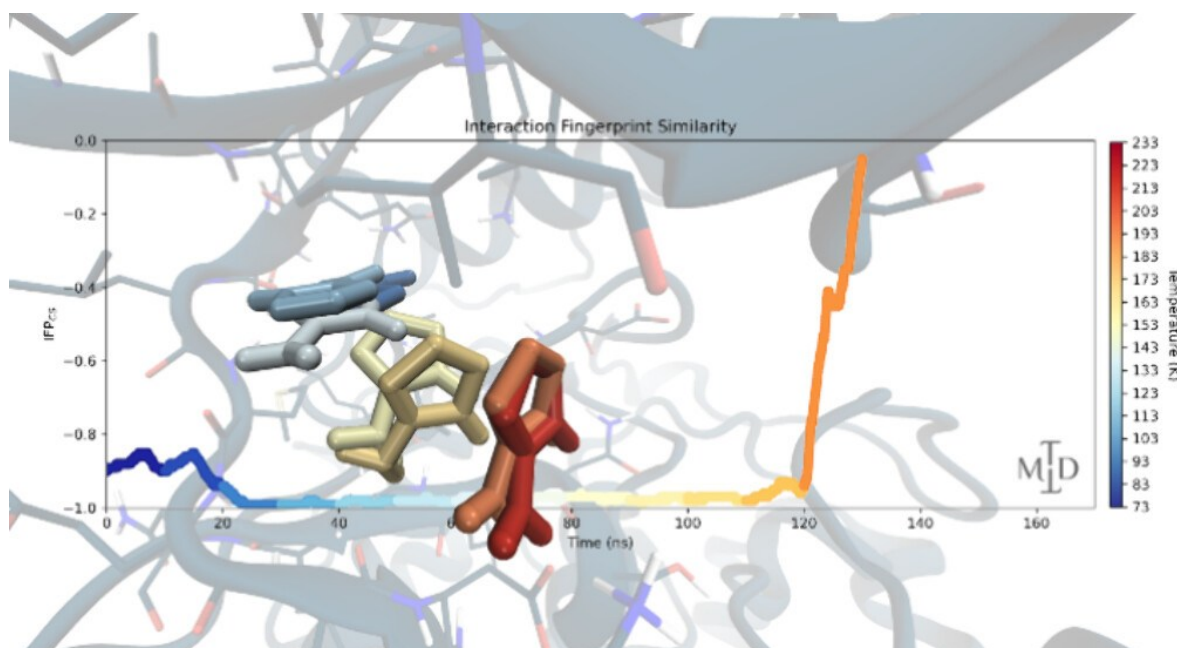
 **ACS Publications**
Most Trusted. Most Cited. Most Read.

www.acs.org

Thermal Titration Molecular Dynamics (TTMD): the Revenge of the Fragments

Matteo Pavan^{1‡}, Silvia Menin[‡], Andrea Dodaro, Gianluca Novello, Chiara Cavastracci Strascia, Mattia Sturlese, Veronica Salmaso and Stefano Moro*

Pavan, M.; Menin, S.; Dodaro, A.; Novello, G.; Cavastracci Strascia, C.; Sturlese, M.; Salmaso, V.; Moro, S. Thermal Titration Molecular Dynamics: The Revenge of the Fragments. *J. Chem. Inf. Model.* 2025. <https://doi.org/10.1021/acs.jcim.4c01681>.



Abstract

During the last 20 years, the Fragment-Based Drug Discovery (FBDD) approach gained popularity both in industrial and academic settings due to its efficient exploration of the chemical space. This bottom-up approach relies on identifying high-efficiency small ligands (fragments) that bind to a target binding site and then rationally evolve them into mature druglike compounds. To achieve such a task, researchers rely on accurate information about the ligand binding mode, usually obtained through experimental techniques such as X-ray crystallography or computer simulations. However, the physicochemical characteristics of fragments limit the accuracy and reliability of the computational prediction of their binding mode. This article presents a new thermal titration molecular dynamics (TTMD) protocol, a recently developed enhanced sampling method for qualitatively estimating protein–ligand-binding stability, specifically tuned for the refinement of fragment docking results. The protocol has been applied to eight pharmaceutically relevant targets on twelve different test cases, including ligands with very low molecular weight and structural complexity (MiniFrag / FragLites). In more than 80% of cases, TTMD successfully identified the native fragment binding mode among a set of docking poses, outperforming docking alone and proving to be a useful tool to assist the fragment screening and optimization process.

Introduction

Fragment-Based Drug Discovery (FBDD) is an emerging approach to the rational design of novel active principles^{1–3}. It relies on the so-called ‘fragments’, i.e. small organic molecules that follow the ‘rule of three’⁴, an adapted version of the famous Lipinski’s ‘rule of five’⁵. The main advantage of FBDD campaigns

is the ability to cover a wider range of chemical space compared to the canonical 'drug-like' landscape, thus increasing the hit success rate⁶⁻⁸.

Due to the limited size of fragments compared to druglike molecules, the hit identification stage is only the starting point for an iterative structure optimization process⁹. Although highly sensitive biophysical techniques like isothermal titration calorimetry (ITC)¹⁰ and surface plasmon resonance (SPR)^{11,12} are routinely adopted in fragment-based screening campaigns, they do not provide the structural insights needed to steer the fragment-to-lead process. Lately, both X-ray crystallography (XRC)^{13,14} and nuclear magnetic resonance (NMR)^{15,16} have been successfully used to perform screening campaigns in the frag-like chemical space. Although these techniques can provide structural information about the interaction between the fragment and its target, their application is limited by costs and drawbacks. For example, NMR spectroscopy requires large quantities of soluble, labeled protein, a considerable amount of time, and is mostly limited to the investigation of small proteins¹⁷. On the other hand, XRC requires crystallization of the sample, which limits its application to those targets whose size and dynamics do not hamper the process¹⁸.

Parallel to the use of experimental methodologies, several *in silico* tools started becoming part of FBDD campaigns, aiding the work of medicinal chemists at several stages, such as the identification of putative binding sites, virtual screening, binding mode characterization, and hit optimization^{19,20}. Although computer-aided drug discovery (CADD) methodologies have been a staple point of drug discovery campaigns for a long time²¹⁻²³, their application to fragments is limited so far by the peculiar physicochemical characteristics of frag-like molecules. Indeed, these molecules are characterized by low-affinity transient interactions with their targets that lead to a population of multiple binding modes that are difficult to recapitulate with traditional methodologies developed for working with druglike molecules²⁴. For example, molecular docking is widely used for predicting the binding mode of druglike compounds, but its accuracy is subpar when applied to fragments. Indeed, although molecular docking can sample the native binding mode in most cases, traditional scoring functions cannot prioritize it among a set of properly chosen decoys^{25,26}.

To compensate for the weak points of docking, several strategies can be used. For example, the best combination between the search algorithm and the scoring function can be chosen among a set of different protocols on a case-by-case scenario by retrospectively evaluating the capabilities of said protocols to reproduce experimentally determined binding modes for compounds that bind the target of interest, in a process defined as self or cross-docking^{23,27-30}. Another strategy could be to use a knowledge-based approach and filter poses based on compliance to a pharmacophore model³¹ or similarity to a reference interaction fingerprint³². Furthermore, one could use a consensus approach between orthogonal docking protocols to minimize the proneness of false positives³¹. Although effective, all these strategies do not address some underlying issues with docking, such as neglect of protein flexibility and the role of solvent, which have a pivotal role in ligand recognition, especially in the case of fragments.

Some docking protocols can account for the presence of water molecules and significantly increase their performance³³. However, one would have to accurately assign the position and orientation of water molecules in the binding site, either based on experimental data or computational prediction³⁴. Although feasible, this strategy still does not take into account the receptor flexibility, thus applying to a target with a defined shape of the binding pocket.

Molecular Dynamics (MD) simulations can bypass this limitation and provide a more accurate biophysical description of the binding event. Nevertheless, although a brute-force approach to fragment posing and screening is theoretically achievable³⁵, the time and computational resources required to perform such a task are limited to an academic setting and not to real-world drug discovery campaigns.

To find a compromise between the swiftness of molecular docking and the accuracy of MD simulations, several different enhanced sampling MD protocols were developed³⁶. These protocols exploit various strategies to reduce the simulation time required to sample the binding event without altering the predictivity of the technique.

Among these methods, Supervised Molecular Dynamics (SuMD³⁷) demonstrated to be particularly successful when applied to both fragment posing^{33,38-40} and screening⁴¹. Although different scoring strategies were proposed for fragment posing using SuMD, including geometric clustering, qualitative analysis of interaction features, MMGBSA scoring, persistence of the bound state, or a combination

between residence time and energetic favorability of the pose⁴², its main disadvantage is that it is impossible to determine a priori the simulation time (i.e., the number of replicates and the length of each one) required to sample the binding event, making its adoption troublesome.

Recently, an alternative enhanced sampling approach to the qualitative estimation of protein-ligand binding stability was developed in our lab. Thermal Titration Molecular Dynamics (TTMD) is a protocol that evaluates the persistence of the binding mode for a given receptor-ligand complex throughout a series of short MD simulations performed at increasingly high temperatures through a scoring function based on interaction fingerprints. This approach can be successfully used to rank different ligands based on their affinity for the target⁴³, evaluate the effect of receptor mutations on protein-ligand complex stability⁴⁴, or successfully rescore docking poses on both protein⁴⁵ and nucleic receptors⁴⁶. So far, the original protocol has demonstrated remarkable posing accuracy for druglike molecules, making it a good alternative to traditional MD-based refinement of docking results, but it has struggled with frag-like compounds⁴⁵. The main strong point of TTMD compared to SuMD is that the maximum required simulation time is known, making it possible to apply it in a high-throughput fashion, facilitating its incorporation in existing structure-based drug discovery pipelines.

In this article, we present the fine-tuning of the TTMD protocol for fragment posing on several pharmaceutically relevant protein targets, discussing its applicability domain, its strengths and weaknesses compared to similar methodologies.

Materials And Methods

Hardware Overview

Each general molecular modeling operation, such as the preparation of protein-ligand complex structures, the setup for MD simulations, and trajectory analyses, was conducted on a 24 CPU Linux workstation equipped with an Intel Core i9-12900 2.4 GHz processor. All MD simulations were carried out on a GPU cluster composed of 29 NVIDIA drivers ranging from GTX980 to RTX3090.

Structure Preparation

The three-dimensional coordinates of the protein-ligand complexes used in this investigation were obtained from the Protein Data Bank (PDB)⁴⁷ and underwent preprocessing using various tools within the Molecular Operating Environment (MOE) 2022.01⁴⁸. This study encompassed eight distinct macromolecular targets: Casein Kinase 1 delta (CK1δ), Casein Kinase 2 (CK2), Pyruvate Dehydrogenase Kinase 2 (PDK2), SARS-CoV-2 Main Protease (M^{Pro}), Extracellular signal-Regulated Kinase (ERK2), ATPase family AAA Domain-containing protein 2 (ATAD2), BRomodoDomain-containing protein 4 (BRD4), Heat Shock Protein 90 (HSP90), tumor protein p53 (p53), and Tyrosyl-tRNA Synthetase (TyrRS). Detailed information regarding the structures considered for each macromolecular target is presented in Table 1.

The list of protein-ligand complex structures and ligands used in this work includes PDB IDs grouped by macromolecule target. For each ligand, the following information is provided: PDB ID, molecular weight (MW), the computed sum of possible hydrogen bond acceptors and donor groups (a_accdon, descriptor from MOE2022), log octanol/water partition coefficient (SLogP descriptor from MOE2022), and its 2D structure.

Every protein-ligand complex was simulated as a monomer, except for SARS-CoV-2 M^{Pro}, which was simulated in its catalytically competent homodimeric form by applying a symmetric crystallographic transformation to each asymmetric unit.

Firstly, each structure was pre-processed using the "Structure Preparation" tool, assigning alternates to the highest occupancy conformation, rebuilding the missing loop through homology modeling, and fixing inconsistencies between the aminoacidic sequence and the experimentally solved structure. Secondly, missing hydrogen atoms were added to the system according to the most probable protonation and tautomeric state at pH = 7.4 through the "Protonate 3D" tool. Lastly, each solvent, co-solvent, and ion atom was removed before storing the prepared complex for subsequent calculations.

Docking calculations

A self-docking experiment was conducted for each protein-ligand complex; every ligand was prepared using the “Wash” tool of the MOE 2022.02⁴⁸ suite, the most probable protomeric/tautomeric state at pH = 7.4 was assigned, then its three-dimensional coordinates were rebuilt; lastly, the correct partial charges, according to the MMFF94 force field⁴⁹, were attributed.

The binding site was defined using a sphere of radius 10.5 Å centered upon the center of mass of the co-crystallized ligand in the experimentally solved structure. Docking calculations were performed using the Protein-Ligand ANT System (PLANTS)^{50,51} program, which uses an Ant Colony Optimization (ACO) search algorithm and is free to use for academics. The first five poses according to the ChemPLP⁵² scoring function have been considered for further calculations.

System Setup for Equilibration Protocol and MD Simulations

Each protein-ligand complex obtained with docking calculations was prepared for MD simulations through a combination of different packages included in both Visual Molecular Dynamics (VMD)⁶² 1.9.2 and the AmberTools22^{63,64} suite. Protein atoms were parametrized through the ff14SB force field⁶⁵, while the general Amber force field⁶⁶ was utilized to parametrize the ligands. The AM1-BCC⁶⁷ method was used to assign partial charges to the ligand atoms. Each investigated system was solvated in a rectangular box, with a padding of 15 Å, utilizing the TIP3P⁶⁸ model for water molecules. A proper number of sodium and chloride ions were added to neutralize the system and reach a salt concentration of 0.154 M. Before undergoing MD simulations, each system was energy minimized for a total of 500 steps with the conjugate-gradient algorithm to remove clashes and bad contacts.

Afterward, each minimized system was subjected to two equilibration steps. In the first stage the system was simulated for 0.1 ns in the canonical ensemble (NVT), with harmonic positional restraints (5 kcal·mol⁻¹·Å⁻² force constant) applied on both protein and ligand atoms, and the temperature was kept at a constant value of 73 K, which is the TTMD ramp starting temperature, through a Langevin thermostat⁶⁹. During the second stage, a 0.5 ns simulation in the isothermal-isobaric ensemble (NPT) was performed, applying the same restraints only to the ligand and the protein backbone atoms and using a Monte Carlo barostat⁷⁰ for keeping the pressure at a constant 1 atm value. The canonical ensemble (NVT) was employed during the productive stage, and the temperature was kept at a constant value in each TTMD trajectory step.

For both the equilibration stages and the productive stages, an integration timestep of 2 fs was used, the length of bonds involving hydrogen atoms was constrained with the M-SHAKE algorithm⁷¹, electrostatic interactions were computed through the particle-mesh Ewald method⁷², using cubic spline interpolation and a 1 Å grid spacing and setting a 9.0 Å cutoff for the calculation of Lennard-Jones interactions.

Both equilibrations and productive simulations were carried out through the ACEMD⁷³ 3.5 engine, which is based upon the open-source library for molecular simulations OpenMM⁷⁴ 7.

TTMD simulations

Thermal Titration Molecular Dynamics (TTMD⁴³) is an enhanced sampling MD approach originally developed for the qualitative estimation of protein-ligand unbinding kinetics. Its applicability domain has already been extended to the identification of native-like docking poses both on protein⁴⁵ and RNA complexes⁴⁶. Through a sequence of brief, unbiased MD simulations conducted at incrementally higher temperatures, TTMD aims to examine the resilience of these complexes under varying thermal conditions. For monitoring the TTMD trajectory, a protein-ligand interaction fingerprints-based scoring function denoted as IFP_{CS}, was used, which was originally implemented for the open-source Autogrow4³² program, a semi-automatized workflow for fragment-based de novo drug design. For every trajectory frame, the protein-ligand interaction fingerprint (assumed as vectors, A) is calculated through the Open Drug Discovery Toolkit (ODDT⁷⁵) or the Protein-Ligand Interaction Fingerprint (ProLIF⁷⁶) Python libraries, and is then compared against a reference fingerprint derived from the final frame of the equilibration stage (vector B), employing the cosine similarity metric implemented within the Scikit-learn Python module. The resulting value is then multiplied by -1 to be comparable with most scoring functions. Indeed, the final score can range from -1, which represents the maximum grade of similarity between the reference and the observed fingerprint vector, and 0, which represents the complete discrepancy between the two vectors, indicating that the initial binding mode was lost. The mathematical formulation of the IFP_{CS}

scoring function is reported in Equation 1.

$$IFP_{CS} = \frac{A \cdot B}{\|A\| \cdot \|B\|} \cdot -1$$

Equation 1.

After every TTMD-step, the IFP_{CS} calculation is performed; then, taking into account only the last 10% ("Score-Range") frames of the trajectory, the average IFP_{CS} value is calculated and, if it is greater than -0.05 ("Score-Stop", which means that the native binding mode has been almost lost at the end of the step), the run is stopped earlier, rather than reaching the final temperature step, to save valuable computational time.

Central to this approach is the definition of temperature ramp settings, including the maximum and minimum temperatures, the duration of each simulation window (known as "TTMD-step"), and the temperature increment between successive step temperatures (known as "temperature-step"). These parameters are carefully chosen based on prior insights into the target system, particularly concerning the protein's structural stability at elevated temperatures. Within this work, an alternative ramp, previously developed in the work of Dodaro et al.⁴⁶, was used, which consists of 17 TTMD step of 10 ns each (for a total of 170 ns), starting from 73 to 233 K, with a temperature-step of 10 K. Notably, ice formation does not occur under these simulation conditions due to the complexity of the biological assemblies studied, the limited time spent at low temperatures, and the use of the TIP3P water model, which is less prone to ice nucleation compared to models like TIP4P⁷⁷.

The code to run TTMD, which is open-source and available at <https://github.com/molecularmodelingsection/TTMD>.

MS and AUC Coefficient Determination

For every trajectory, two different Python packages (ODDT⁷⁵ and ProLif⁷⁶) were used to calculate IFP_{CS} values for every frame. Moreover, two different metrics were used to calculate a proxy value for the binding mode conservation: the MS coefficient and the %AUC.

The MS coefficient is calculated using Equation 2 and represents the slope of the straight-line interpolating between the trajectory's starting and ending points. The starting point is the last equilibration frame, where the IFP_{CS} value is -1 due to complete overlap with the reference. The ending point is the average IFP_{CS} value of the last 10% of frames from the final temperature step.

$$MS = \frac{meanIFP_{CS}^{T^{end}} - (-1)}{T^{end} - T^{start}}$$

Equation 2.

An alternative metric, already used in other works^{78–80} to evaluate TTMD trajectories, is the %AUC coefficient. As can be seen in Equation 3, this metric is calculated by dividing the Area Under the Curve (AUC), derived from integrating the IFP_{CS} score over every frame of the trajectory, by the maximum possible AUC value, and finally by multiplying this value by 100. The maximum AUC is theoretically obtained when the IFP_{CS} value remains at -1 throughout the entire trajectory.

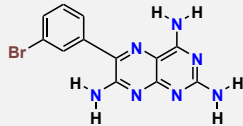
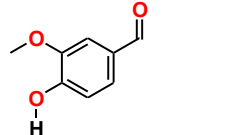
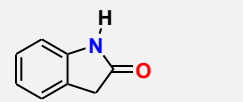
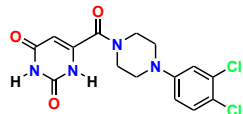
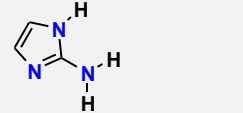
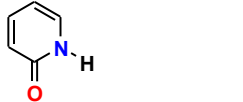
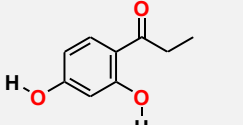
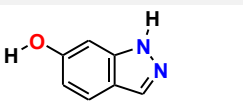
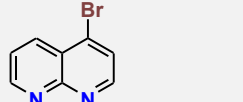
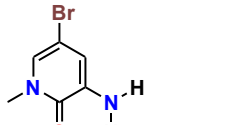
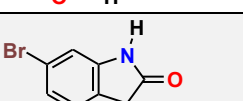
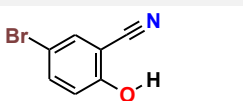
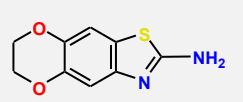
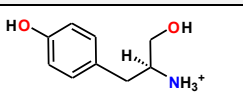
$$\%AUC = \left(\frac{\int_{t^{start}}^{t^{end}} IFP_{CS}}{AUC_{Max}} \right) \cdot 100$$

Equation 3.

A %AUC value close to 100% indicates that the protein-ligand fingerprint is well-conserved during the trajectory, whereas lower values suggest that the binding mode has been partially or completely lost.

For each ligand, five independent TTMD simulations are performed, and the four average metric coefficients are then calculated based on three of them, discarding the highest and the lowest value.

Table 1. Structures considered for each macromolecular target.

TARGET	CRYSTAL PDB ID	LIGAND ID	MW	A_DONACC	SLOGP	2D STRUCTURE
CK1Δ	5IH6 ⁵³	AUG	332.17	7	1.6	
CK2	6HOU ⁵⁴	V55	152.15	4	1.2	
PKD2	7EA0 ⁵⁵	W6P	133.15	2	1.2	
M ^{PRO}	7LTJ ⁵⁶	YD1	369.21	5	1.4	
ERK2	6Q7K ⁵⁷	2AI	83.09	5	0.0	
	6QA4 ⁵⁷	HRZ	95.10	4	0.2	
HSP90	2YE4 ⁵⁸	2FY	166.18	5	1.7	
	2YEB ⁵⁸	2K4	134.14	6	1.3	
ATAD2	7QU7 ⁵⁹	HH8	209.05	2	2.4	
	7R00 ⁵⁹	HWC	203.04	2	0.5	
BRD4	7ZE6 ⁵⁹	1P8	212.05	2	1.9	
	7ZE7 ⁵⁹	UUJ	198.02	3	2.0	
P53	2X0U ⁶⁰	X0U	208.24	4	1.6	
TYRRS	1Q11 ⁶¹	TYE	168.22	4	-0.5	

Results

Rational design of the fragment-specific temperature ramp

In our previous work⁴⁵, we assessed how the original TTMD protocol described in the work of Pavan et al.⁴³ could be readily repurposed as a post-docking refinement tool for mature, drug-like compounds. Specifically, we assessed how TTMD could successfully retrieve the native-like binding mode for a series of pharmaceutically relevant protein-ligand complexes among decoy docking poses. However, when the same protocol was applied to frag-like compounds, the results of the TTMD refinement were inconclusive, highlighting the need for the protocol's settings to be adapted to the weaker and more transient nature of complexes involving these compounds. Specifically, we observed how the previously adopted temperature ramp was too aggressive for this use case, leading to premature and spontaneous unbinding of most investigated complexes and flattening stability differences between the native and non-native poses. For this reason, a new fragment-specific temperature ramp was developed, and its performance was assessed in the present work. Specifically, since complexes involving fragments show a higher intrinsic instability under standard simulation conditions contrary to the ones involving drug-like compounds, we decided to explore lower, subzero, temperatures. Instead of aggressively heating up the systems, we opted for slowly thawing them, to exacerbate the stability differences between native and non-native poses for frag-like compounds. Contrary to the original temperature ramp (300-450 K, $dT = 10$ K, $dt = 10$ ns) which explores temperatures above the physiological threshold, the newly proposed ramp (73-233 K, $dT = 10$ K, $dt = 10$ ns) explores below freezing ones. To evaluate the performance of the new ramp, we applied the new TTMD protocol to a series of pharmaceutically relevant test-cases, starting with the ones already explored in the work of Menin et al.⁴⁵ and including some new, challenging ones selected from the literature. A list of each investigated system can be found in Table 1, while a complete report of results can be found in Supplementary Material.

Previous Test Cases

To assess the potential of the new temperature ramp, we started from the four fragment test cases examined in the previous work from Menin et al.⁴⁵, to have a direct comparison between the performance of the old and new temperature ramp.

Protein Kinase CK1 δ

The initial test case focused on Protein Kinase 1 δ , a serine-threonine kinase belonging to the CK1 kinase family⁸¹. This kinase is implicated in various neurodegenerative diseases such as Alzheimer's, Parkinson's, and amyotrophic lateral sclerosis⁸² due to its pleiotropic nature⁸³. The results of TTMD refinement conducted on docking poses generated for each CK1 δ protein-ligand complex are summarized in Figure 1 and Figure 2.

5IH6

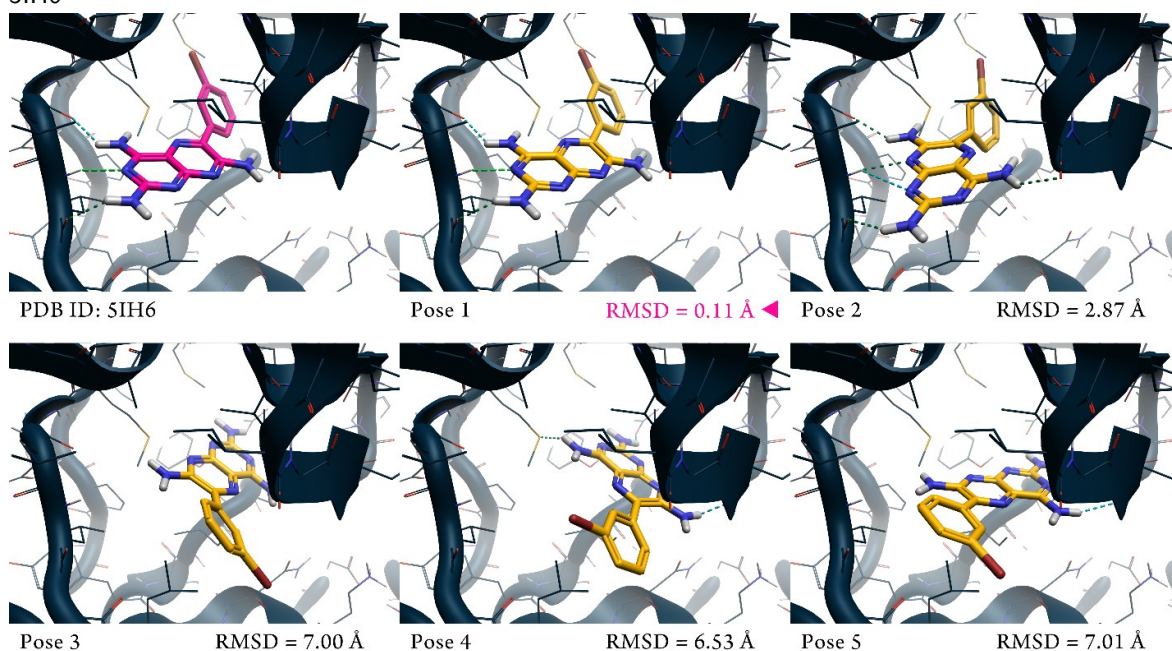


Figure 1. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 5IH6. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

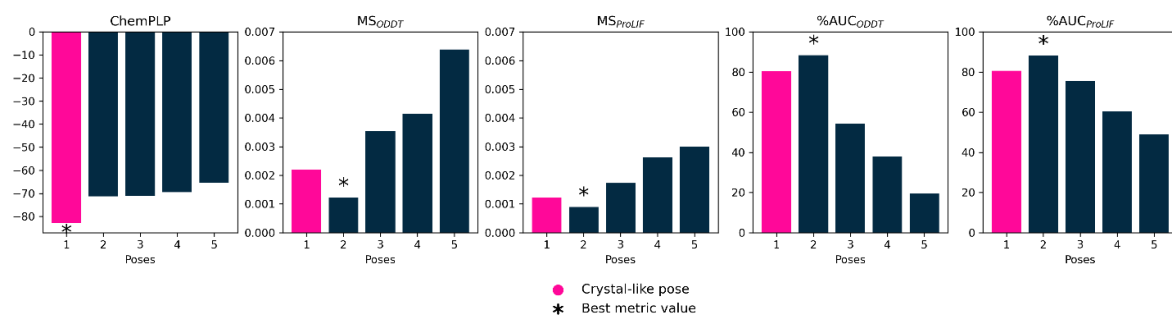


Figure 2. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 5IH6. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

In the work of Menin et al.⁴⁵, the original TTMD protocol was not able to rescue the native-like pose (pose 1) in the case of complex 5IH6. Instead, TTMD favored a flipped ligand orientation (pose 2) within the pocket that had a better electrostatic interaction due to forming a higher number of hydrogen bonds. As can be seen in Figure 1 and Table 2, the same ranking is maintained even when the new temperature ramp is adopted, regardless of the metric used for evaluating the conservation of the interaction fingerprint. However, the difference in score between the flipped conformation and the native one decreases, while the difference between the top two docking solutions and other wrong poses increases, suggesting a potential improvement in the sensitivity and accuracy of the technique when the new ramp is adopted.

Protein Kinase CK2

The second test case regarded protein kinase 2, a serine-threonine kinase that represents one of the first identified protein kinases⁸⁴. Similar to CK1, thanks to its pleiotropic nature⁸⁵ this kinase is involved in

several illnesses such as different types of cancer, various neurodegenerative diseases, and viral infections⁸⁶. The results of TTMD refinement conducted on docking poses generated for each CK2 protein-ligand complex are summarized in Figure 3 and Figure 4.

6HOU

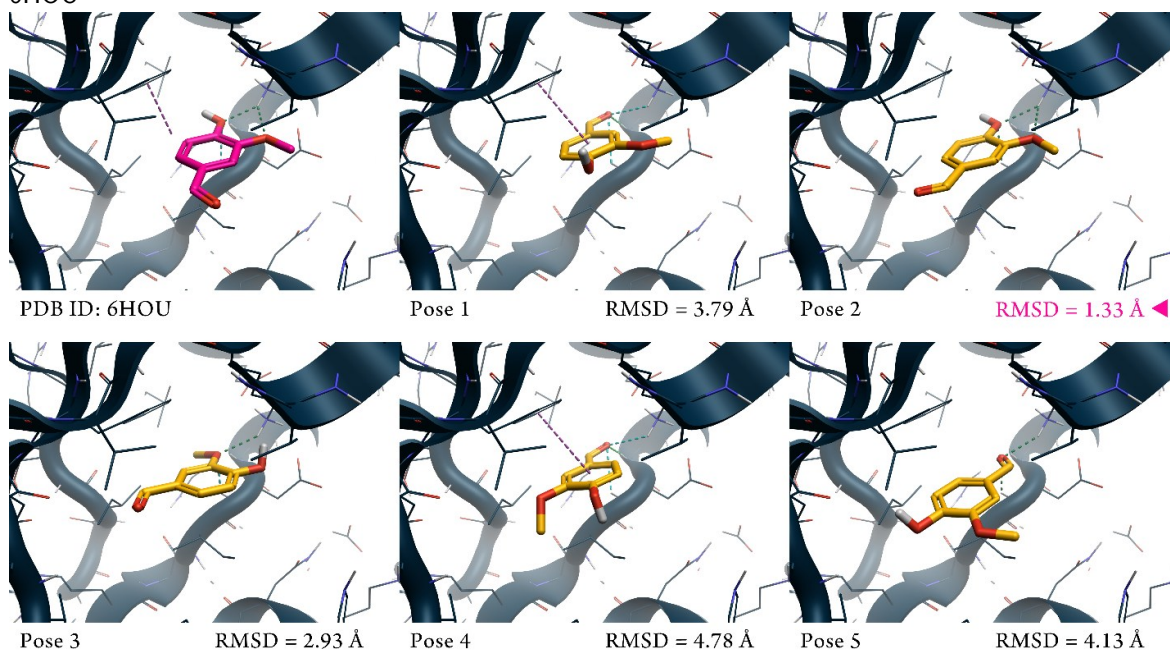


Figure 3. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 6HOU. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

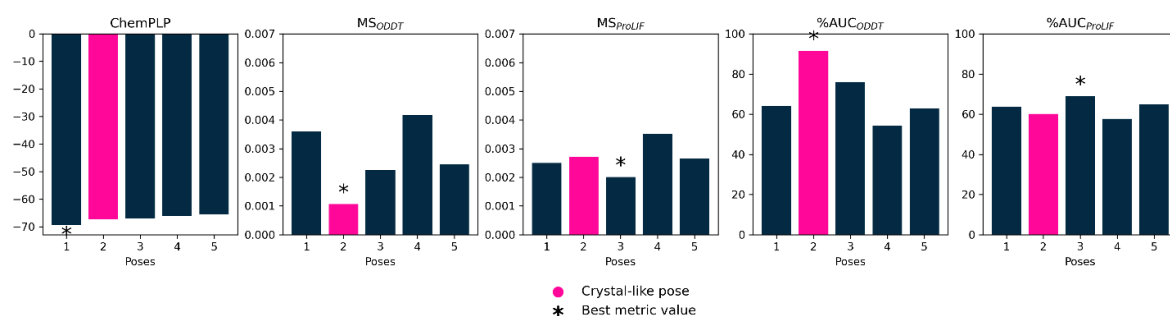


Figure 4. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 6HOU. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

In the original work, no difference in stability could be appreciated between native and non-native binding poses sampled for complex 6HOU. Regardless of the ligand orientation, the original temperature ramp would cause rapid unbinding. Interestingly, the new temperature ramp can correctly identify the native-like pose (pose 2) with both adopted metrics (MS and AUC) when ODDT fingerprints are used. On the contrary, ProLIF fingerprints would lead to identifying pose 3 as the most stable pose, which is the second-best solution according to the RMSD to the experimental conformation. Interestingly, docking alone would have prioritized pose 1, making 6HOU the first case where TTMD refinement outperforms docking alone. Regardless of the discrepancy between different protein-ligand interaction fingerprints, the ability of TTMD to correctly identify the best or second-best solution among a set of highly similar poses for such a small

compound (MW=152.15 g/mol) is a remarkable improvement over the original protocol and a hint at the potential of the newly developed temperature ramp. The reliability of both protein-ligand interaction fingerprint packages will be assessed based on the aggregate results for other test cases similar to what has been done in the case of RNA-ligand complexes explored in the work of Dodaro et al.⁴⁶.

Pyruvate Dehydrogenase Kinase 2 (PDK2)

The third test case involved pyruvate dehydrogenase kinase 2, a member of the GHKL (Gyrase, Hsp90, histidine Kinase, MutL) ATPase/kinase superfamily that has a central control role upon cellular energy metabolism⁸⁷ and is involved both in metabolic and neoplastic diseases⁸⁸. The result of the TTMD refinement carried out on docking poses generated for each PDK2 protein-ligand complex is summarized in Figure 5 and Figure 6.

7EA0

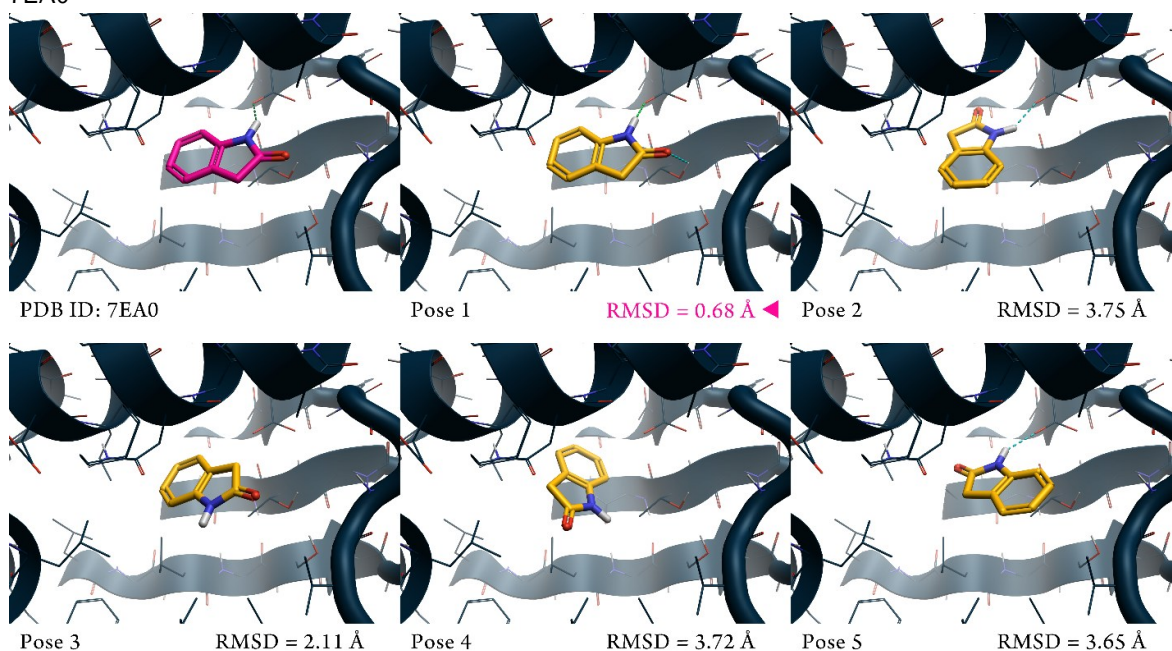


Figure 5. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 7EA0. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

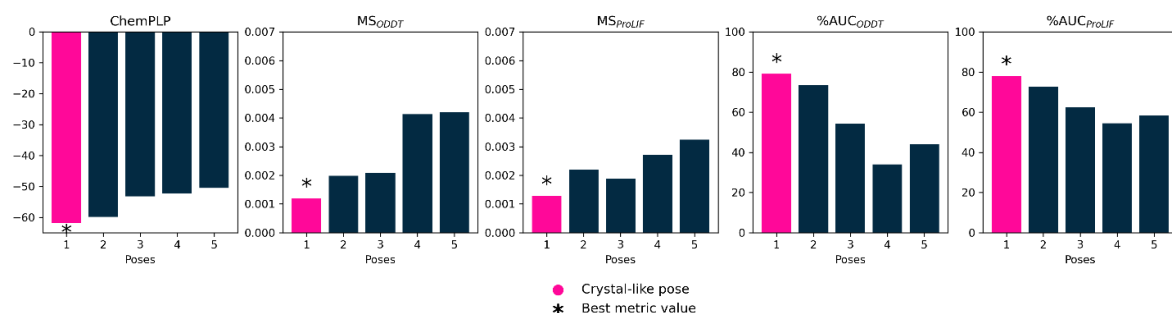


Figure 6. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 7EA0. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

In the original publication, the TTMD refinement for complex 7EA0 was inconclusive, due to the high

instability observed for all the different poses. In this case, instead, the new ramp can easily retrieve the native binding pose (pose 1) among other docking poses, supporting the ranking provided by the ChemPLP scoring function regardless of the metric and fingerprint package adopted.

SARS-CoV-2 Main Protease (M^{Pro})

The fourth and final test case derived from the work of Menin et al.⁴⁵ involved the SARS-CoV-2 main protease (M^{Pro}), an enzyme playing a pivotal role in this virus replication cycle⁸⁹ thanks to its ability to drive the maturation of several nonstructural proteins through the proteolytic cleavage of the pp1a/pp1ab polyproteins⁹⁰.

7LTJ

The result of the TTMD refinement carried out on docking poses generated for each M^{Pro} protein-ligand complex is summarized in Figure 7 and Figure 8.

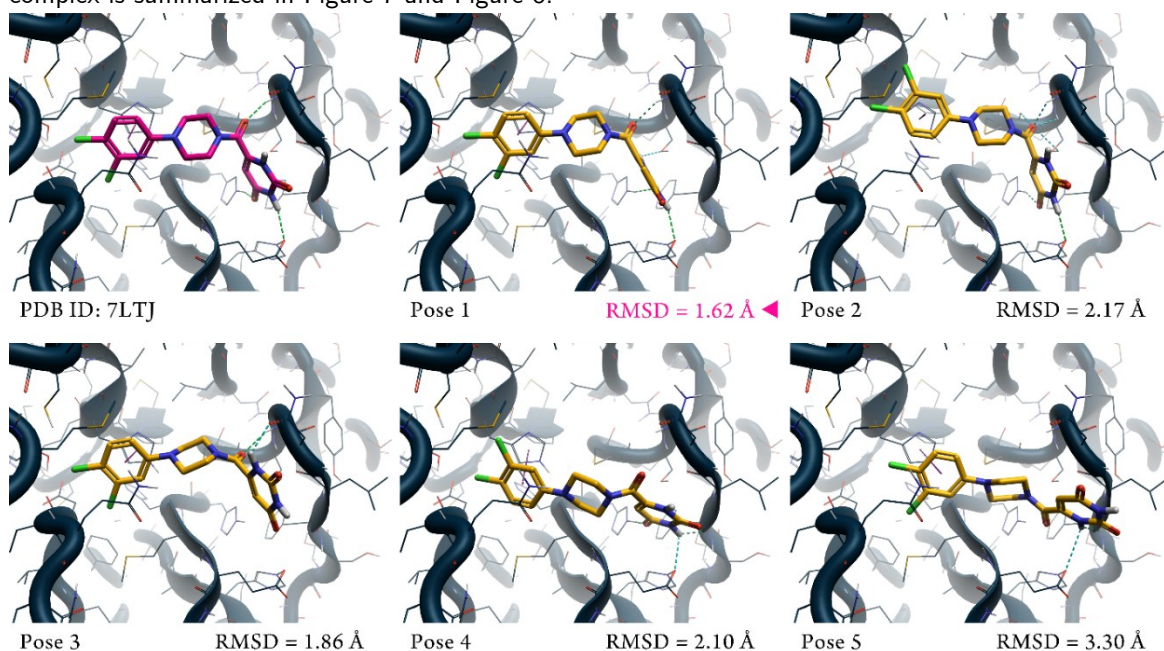


Figure 7. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 7LTJ. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

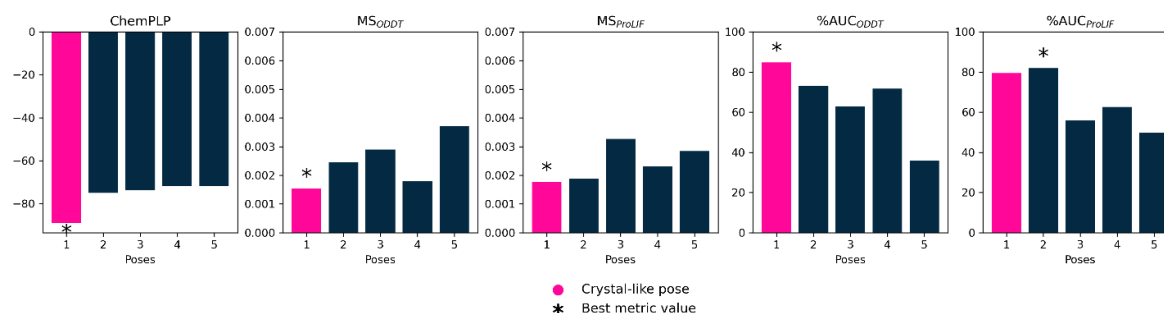


Figure 8. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 7LTJ. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

As observed for PDK2 (complex 7EA0), even in the case of SARS-CoV-2 M^{Pro} the new temperature ramp

succeeds where the original ramp failed. While the previous TTMD refinement of the pose was inconclusive, in this case, the new ramp correctly identified pose 1 as the best pose with all used metrics except for the ProLIF % AUC, which would have selected pose 2 instead. Interestingly, the difference in score between the best solution and the second top-scoring one is higher when ODDT fingerprints are used, regardless of the metric. In the case of ProLIF's fingerprint, the difference in score between pose 1 and pose 2 is small, thus explaining the discrepancy between the MS and AUC-based rankings. Due to the way the MS and AUC metrics are calculated, the MS value weights more the comparison between the final and starting state of the simulation, regardless of the path, while the AUC better encompasses the fingerprint conservation across the whole duration of the simulation. Based on the observed data, it is tempting to speculate that a consensus ranking between the two metrics could make the pose selection process more robust.

New Test Cases

To further assess the potential of the newly proposed fragment-specific temperature ramp, we decided to apply the TTMD refinement protocol to 8 additional protein-ligand complexes, 2 for each pharmaceutically relevant target.

Extracellular signal-Regulated Kinase 2 (ERK2)

Extracellular signal-Regulated Kinase 2 (ERK2) is a pleiotropic serine/threonine kinase that participates in the Ras-Raf-MEK-ERK (RAt Sarcoma- Rapidly Accelerated Fibrosarcoma kinase - MAP/ERK kinase (Mitogen Activated Protein/ERK)) signaling pathway, thus regulating a large variety of biological processes that are pivotal for the cell life cycle. As such, ERK2 is a hot target for treating multiple types of cancer^{91,92}. Due to being a well-characterized target, both from a structural and functional perspective, it has been used for a proof-of-concept crystallographic screening campaign using a library of ultra-low-molecular-weight compounds (MiniFrag)⁵⁷. The results of the TTMD refinement carried out on docking poses generated for each ERK2 protein-ligand complex are summarized below.

6Q7K

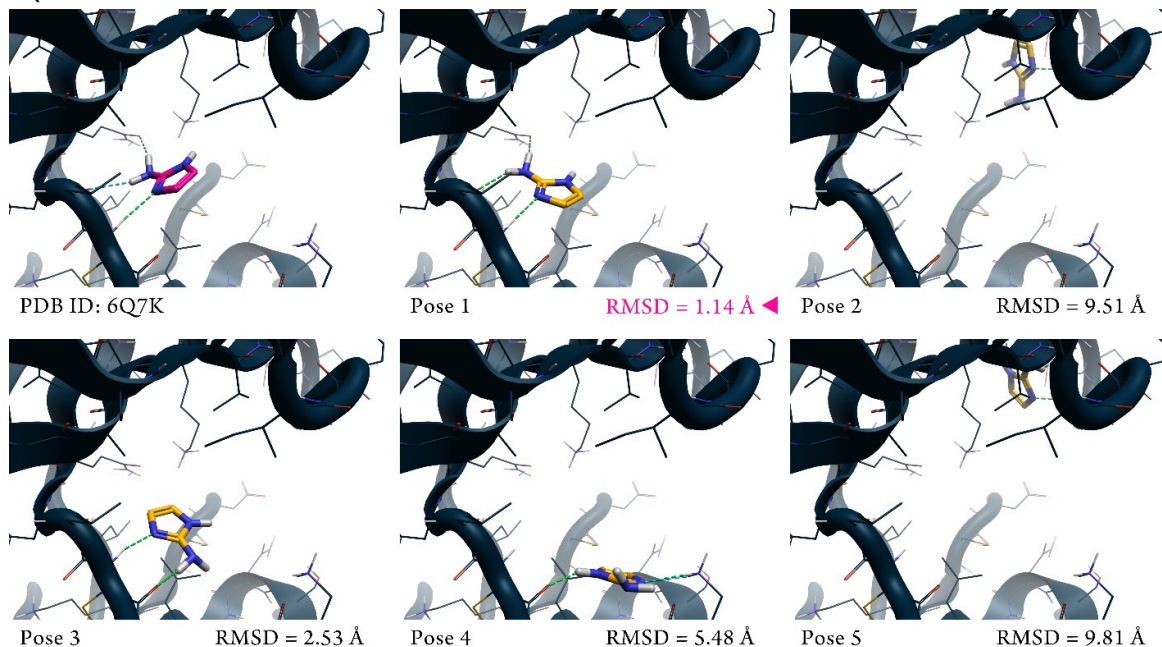


Figure 9. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 6Q7K. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

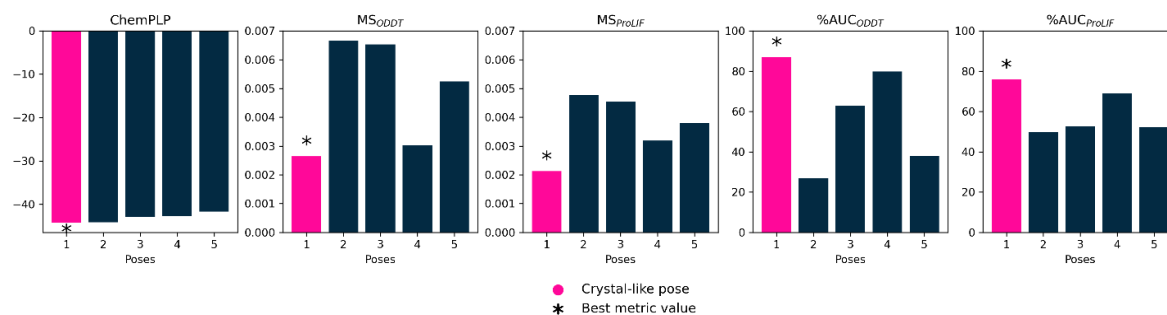


Figure 10. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 6Q7K. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

As can be seen in Figure 9 and Figure 10, both docking alone and the TTMD refinement based on the new ramp can prioritize the native binding pose in the case of complex 6Q7K, regardless of the scoring metric adopted. However, differences in score between the top solution and other decoys are amplified for TTMD compared to docking, where the close scores between the native and incorrect poses might mistakenly lead to the conclusion that they have comparable stability, regardless of parameters like solvent exposure. Interestingly, the TTMD refinement protocol can distinguish between pose 1 (native) and pose 3, a flipped ligand conformation that maintains a double hydrogen bond with the hinge region, a conserved crucial interaction for most kinase inhibitors³¹.

6QA4

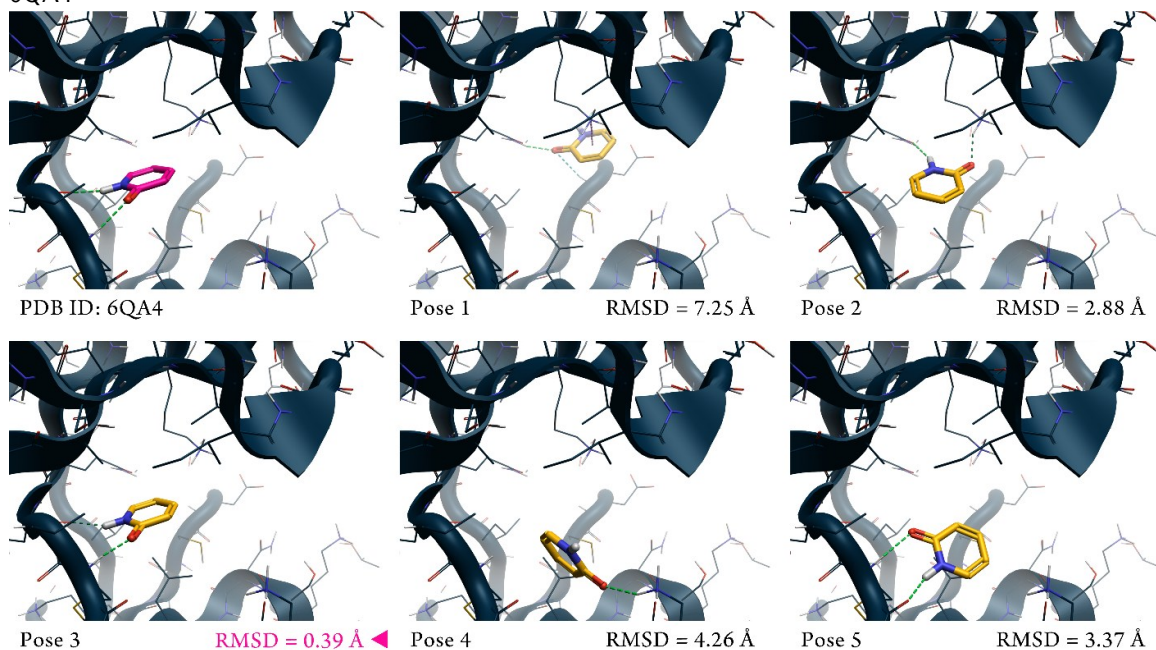


Figure 11. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 6QA4. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

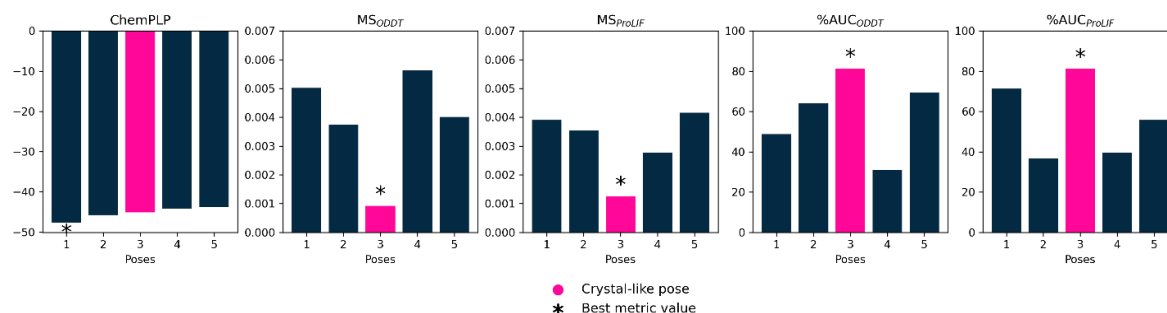


Figure 12. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 6QA4. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

Concerning complex 6QA4 (Figure 11, Figure 12), not only can TTMD once again identify the native binding mode (pose 3) with all adopted metrics, but for the second time (after complex 6HOU) TTMD outperforms docking alone. Indeed, docking would have prioritized pose 1, with the native pose being ranked as the third best. Furthermore, as for the previous case, TTMD can distinguish between the native pose and pose 5, a pose that presents an alternative hydrogen bond pattern with the hinge region. A comparison between the TTMD trajectory for pose 1 (docking top-scoring solution) and pose 3 (TTMD top-scoring solution, native pose) is reported in Video 1 (Supplementary Materials).

Heat Shock Protein 90 (HSP90)

Heat Shock Protein 90 (HSP90) is a chaperone protein member of the GHKL family of ATPases⁹³ that aids the folding of several client proteins, including multiple oncogenes and validated cancer targets⁹⁴. The most abundant cytosolic isoform is HSP90 α , whose N-terminal domain is well-characterized from a structural perspective⁹⁵ and has been targeted by several drug discovery campaigns^{96,97}, even ones involving fragments⁵⁸. The results of the TTMD refinement carried out on docking poses generated for each HSP90 protein-ligand complex are summarized below.

2YE4

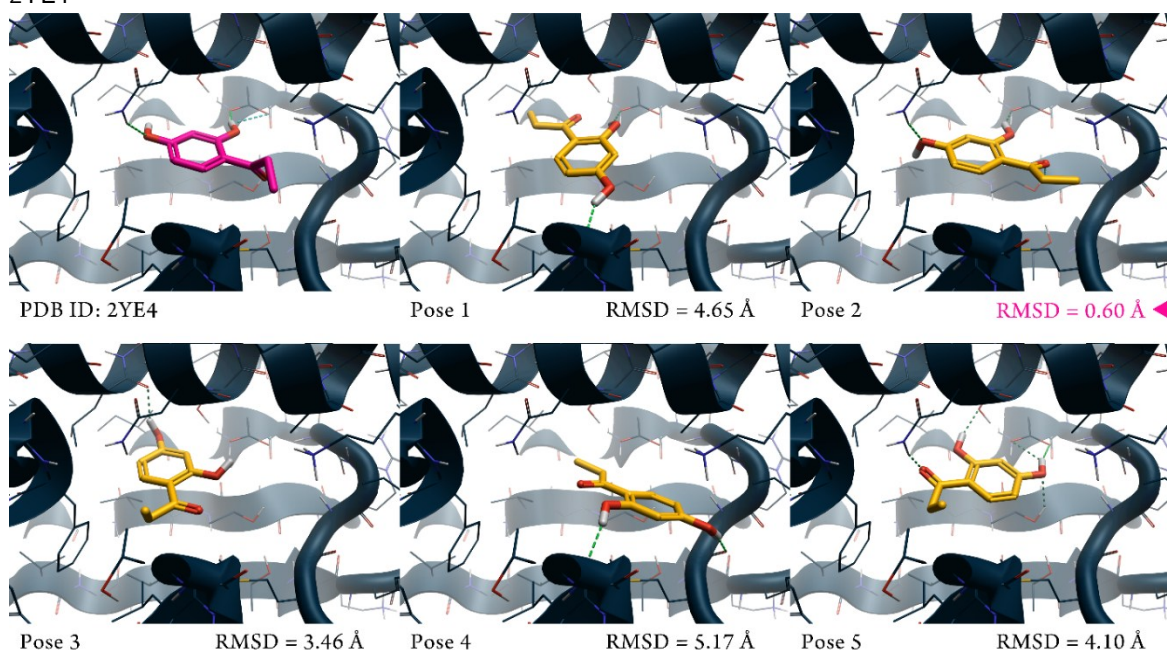


Figure 13. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 2YE4. For each pose,

the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

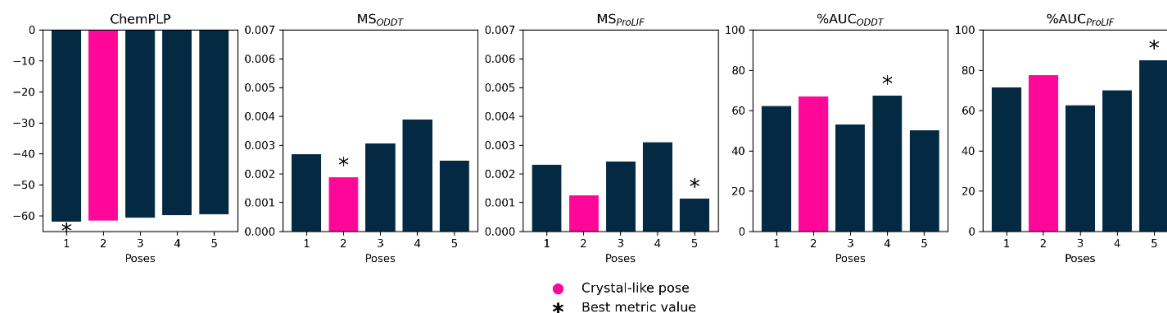


Figure 14. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 2YE4. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

As can be seen in Figure 13 and Figure 14, summarizing the results of TTMD refinement for poses related to complex 2YE4, there is no consensus between the different metrics used. While the docking score would prioritize pose 1 (fourth best according to the RMSD), the MS value computed using the ODDT fingerprints would correctly identify pose 2 as the native binding mode, which would be the second-best one by a small margin according to the ODDT AUC score. On the contrary, both metrics based on the ProLIF's fingerprints would prioritize pose 5, with a small margin over pose 2, which would be ranked as the second-best pose.

2YEB

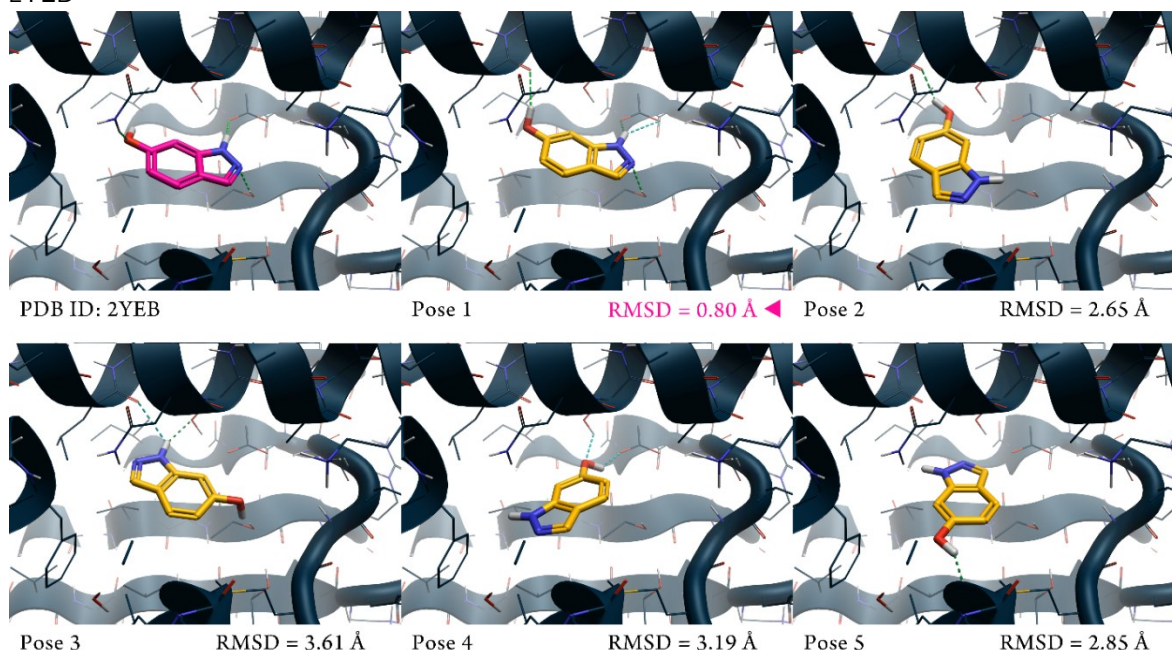


Figure 15. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 2YEB. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

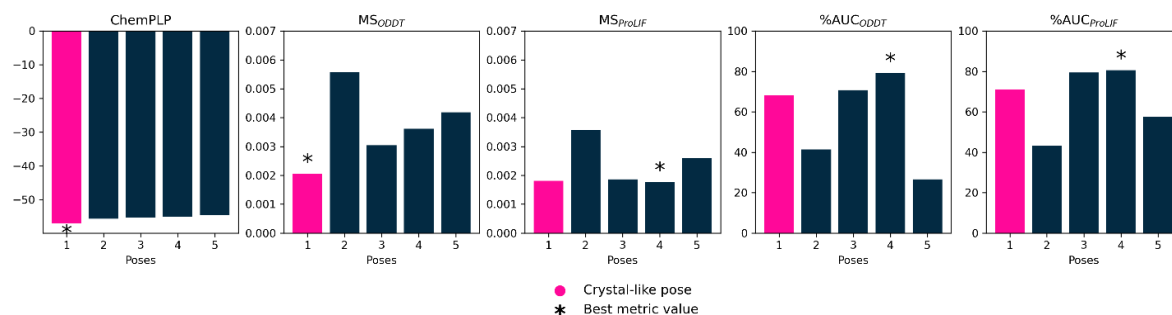


Figure 16. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 2YEB. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

Concerning case 2YEB (Figure 15, Figure 16), there is no consensus between different scoring metrics. Both the docking score and the ODDT MS value correctly identify pose 1 as being the native pose. Both AUC values would rank pose 4 as the best one, with the native pose being ranked as the third best. Interestingly, according to the ProLIF MS value, pose 1 would be the second-best, very close in score to poses 4 and 3. This data not only contradicts the previous hypothesis about a possible consensus strategy between different metrics but also seems to suggest that the MS value might be more accurate in discriminating the pose stability compared to the AUC. This might be related to the fact that, for HSP90, water-bridged interactions play an important role in ligand recognition, which might make the interaction with the pocket more dynamic and increase the fingerprint's fluctuation³³.

ATPase family AAA domain containing 2 Bromodomain (ATAD2)

Bromodomain-containing proteins are a class of epigenetic readers that recognize acetylated lysine residues, often located on histones⁹⁸. Among these, ATPase family AAA domain containing 2 Bromodomain (ATAD2) is a coactivator for several transcription factors, as well as being involved in DNA replication⁹⁹, which has been associated with poor prognosis in several types of cancer¹⁰⁰.

Despite being flagged by early computational analyses as difficult-to-drug¹⁰¹, ATAD2's bromodomain has been the target of several fragment-based screening campaigns^{16,102,103} that paved the way for the identification of multiple potent ligands⁵⁹. The results of the TTMD refinement carried out on docking poses generated for each ATAD2 protein-ligand complex are summarized below. In both examined cases for ATAD2, docking would fail to prioritize the native binding pose.

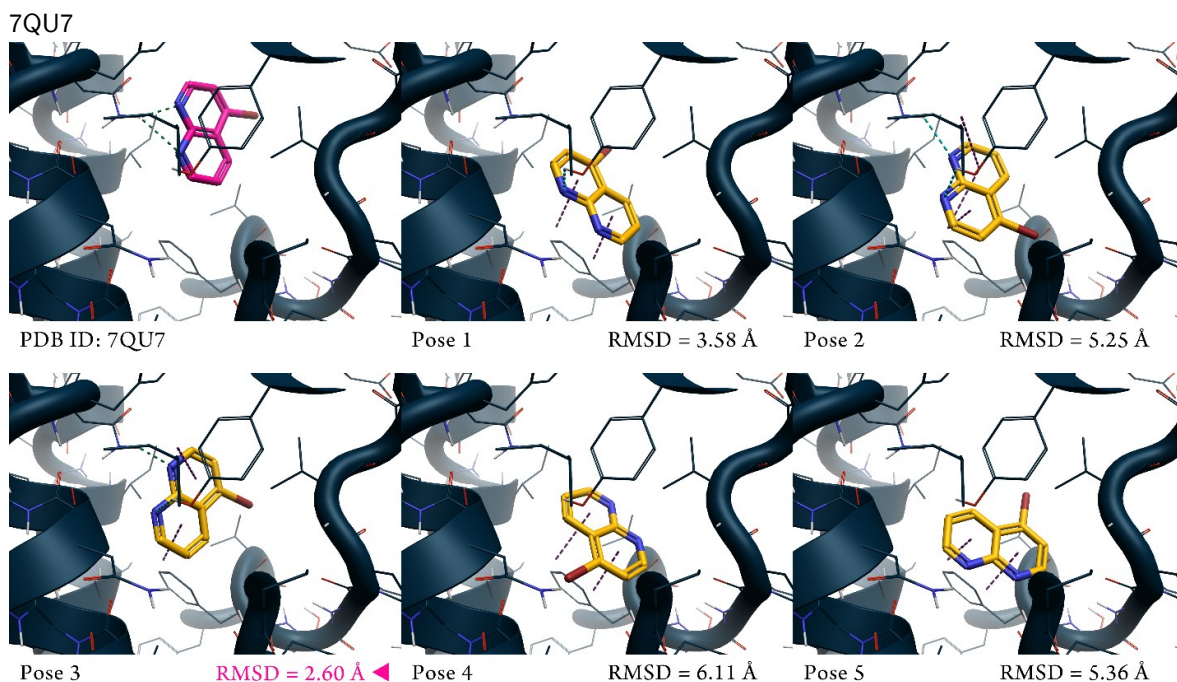


Figure 17. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 7QU7. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

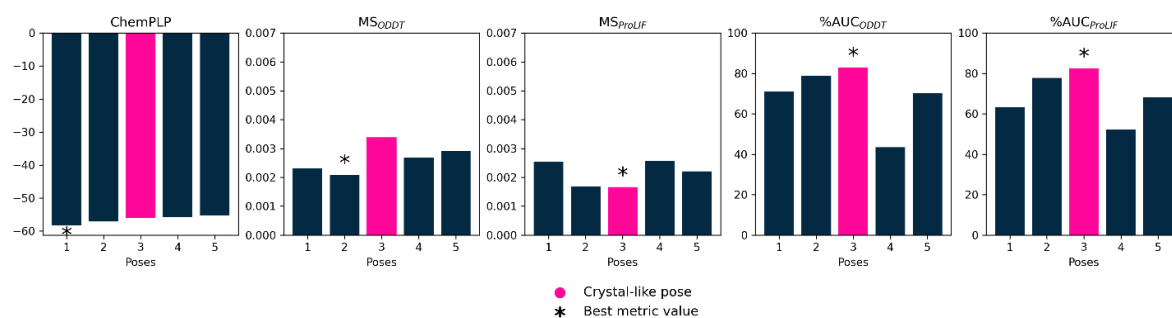


Figure 18. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 7QU7. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

For complex 7QU7 (Figure 17, and Figure 18), the native-like pose (pose 3) would be the third-ranked according to docking, while three out of four metrics adopted by TTMD (except for the ODDT MS) would be able to identify out over other solutions.

Failure of the ODDT MS score for complex 7QU7, contrary to the previous trend, might be related to the slightly higher RMSD value of the most-native docking solution (2.6 Å) compared to typical solutions. Furthermore, as for HSP90, conserved water molecules represent an important feature of the binding site for most bromodomains as well¹⁰¹.

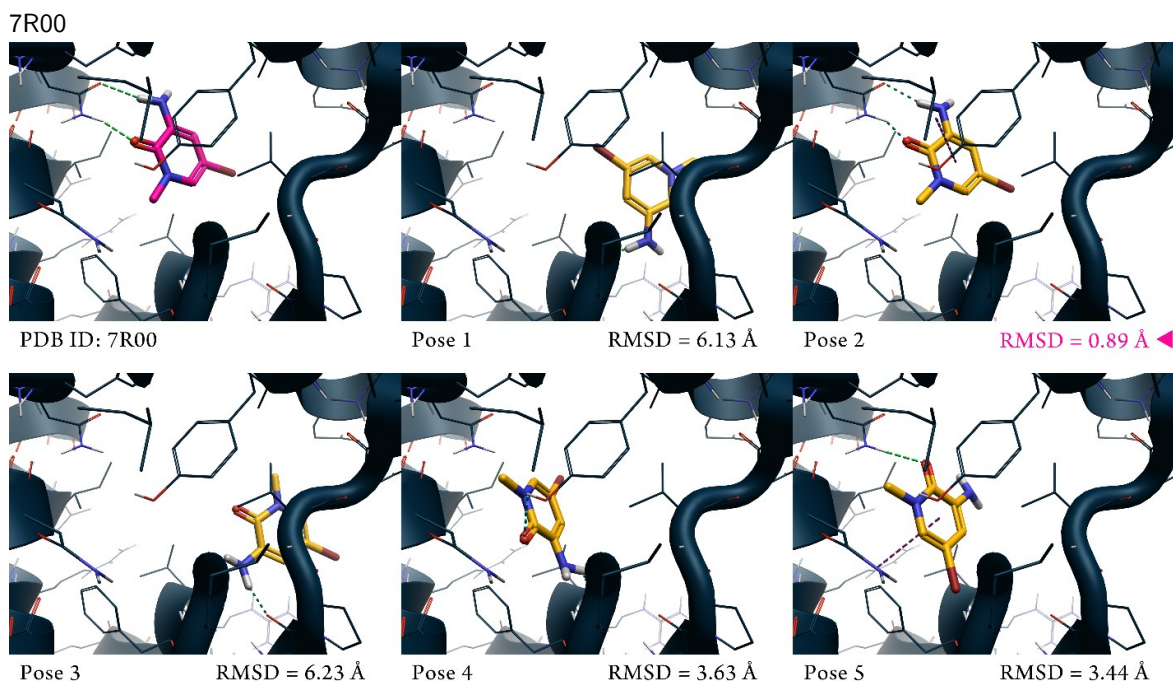


Figure 19. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 7R00. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

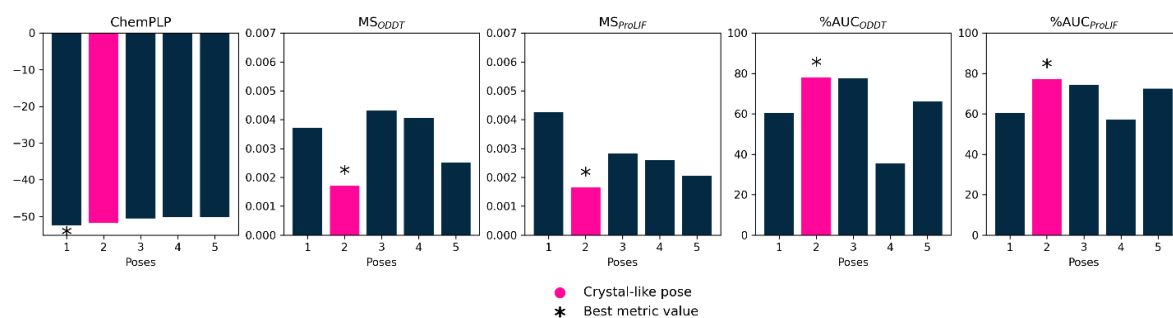


Figure 20. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 7R00. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

In the case of complex 7R00 (Figure 19, and Figure 20), the native pose (pose 2) would be correctly retrieved by all four TTMD scoring metrics. A comparison between the TTMD trajectory for pose 1 (docking top-scoring solution) and pose 2 (TTMD top-scoring solution, native pose) is reported in Video 2 (Supplementary Materials).

Bromodomain-containing protein 4 (BRD4)

Bromodomain-containing protein 4 (BRD4) is a member of the Bromodomain and Extraterminal (BET) family. It plays a pivotal role in mediating several oncogenes expression¹⁰⁴, and as such it is a validated anticancer target¹⁰⁵. It contains two bromodomains and is one of the most studied family members thanks to its higher druggability¹⁰⁶, leading to several drug discovery campaigns¹⁰⁷ including fragment-based ones⁵⁹. The results of the TTMD refinement carried out on docking poses generated for each BRD4 protein-ligand complex are summarized below.

7ZE6

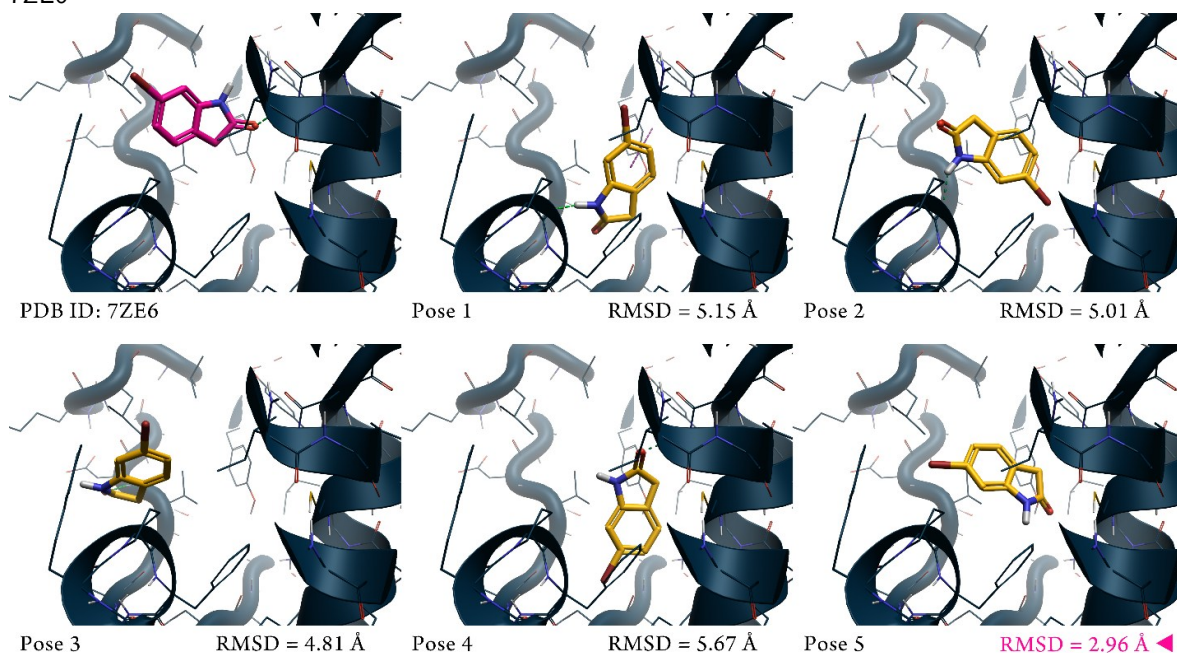


Figure 21. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 7ZE6. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

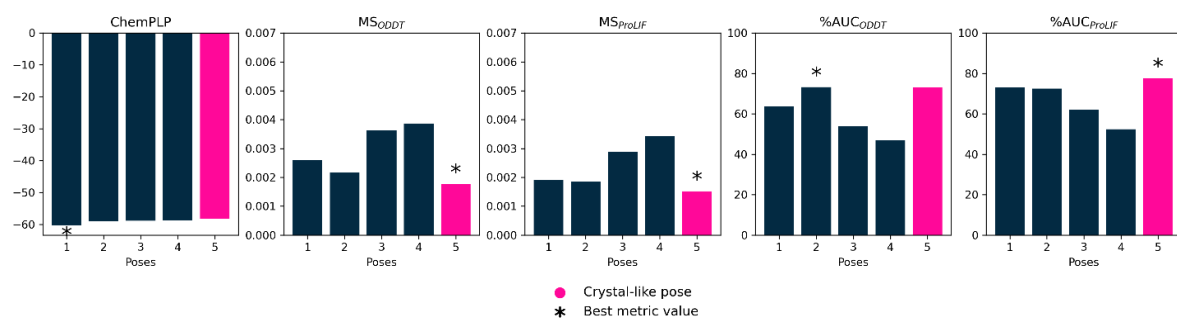


Figure 22. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 7ZE6. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

As reported in Figure 21 and Figure 22, summarizing the TTMD results for poses related to complex 7ZE6, not only would docking fail to identify the native-like pose (pose 5) as the top-scoring one, but it would also be the lowest-scoring solution. On the contrary, every TTMD scoring metric except for the ODDT AUC, would correctly prioritize pose 5 over other binding modes. Interestingly, the ODDT AUC value calculated for pose 5 is very close to the top-scoring solution (pose 2), which would make it the second-best scoring solution according to this metric.

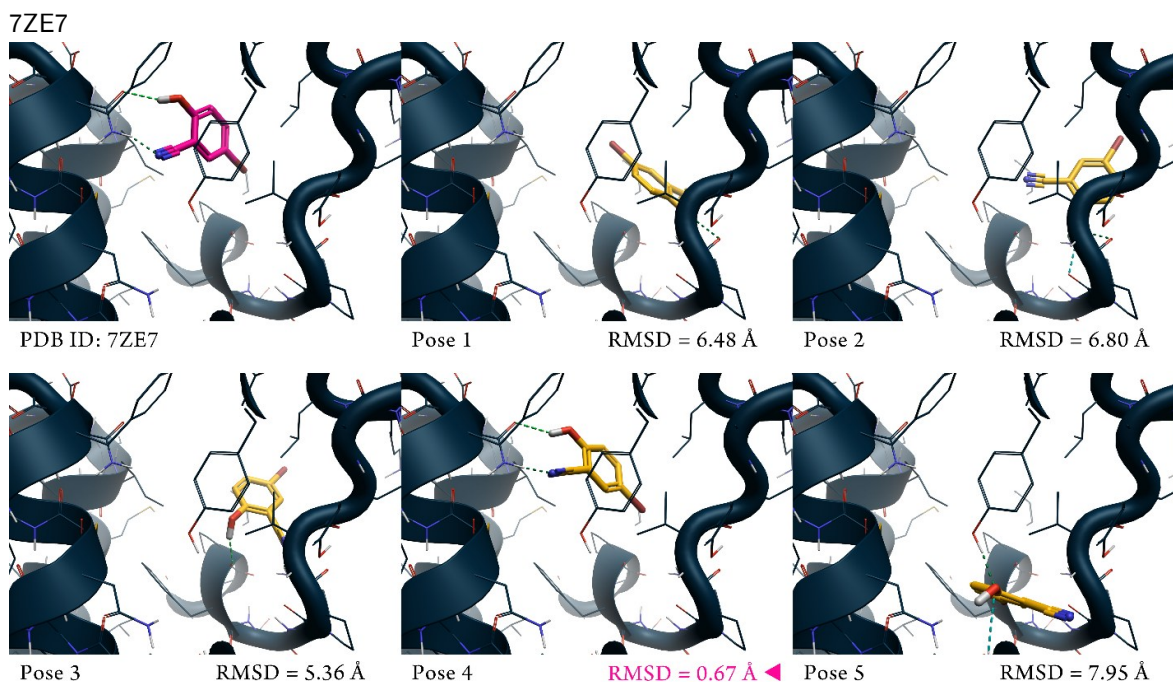


Figure 23. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 7ZE7. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

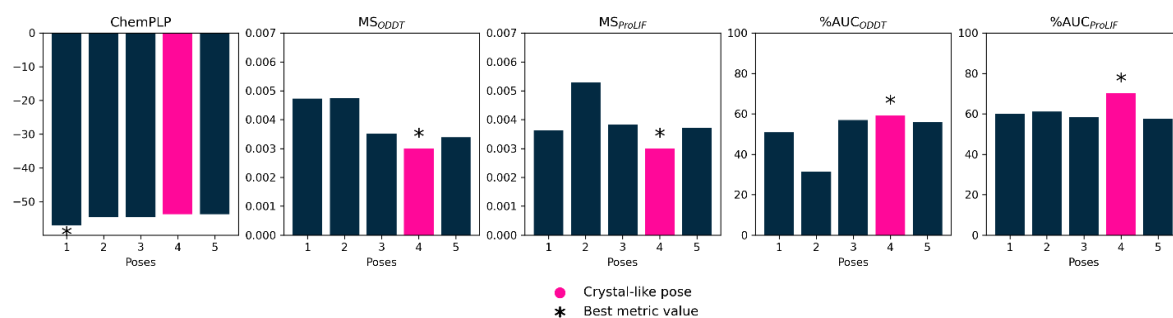


Figure 24. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 7ZE7. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

Remarkably, even in the case of 7ZE7 (Figure 23, Figure 24), docking fails to identify pose 4 as the native one, while all TTMD scoring metrics correctly highlight it as the native one. A comparison between the TTMD trajectory for pose 1 (docking top-scoring solution) and pose 4 (TTMD top-scoring solution, native pose) is reported in Video 3 (Supplementary Materials).

Tumor Suppressor p53

Tumor protein p53 (p53), also known as “the guardian of the genome”, is an onco-suppressor¹⁰⁸ that exerts its conservative role by binding to DNA and regulating gene expression, thus preventing genome mutation¹⁰⁹. Due to its pivotal involvement in preventing cancer formation¹¹⁰, it is the most frequently mutated gene in human tumors¹¹¹, making it an attractive target for anticancer therapy. Since mutations can alter p53 conformation, revealing mutant-specific pockets, fragment-based screening campaigns are an attractive way to find small molecule p53 modulators and characterize the effect of mutations on p53

structure and function¹¹². The results of the TTMD refinement carried out on docking poses generated for the examined p53 protein-ligand complex are summarized below.

2X0U

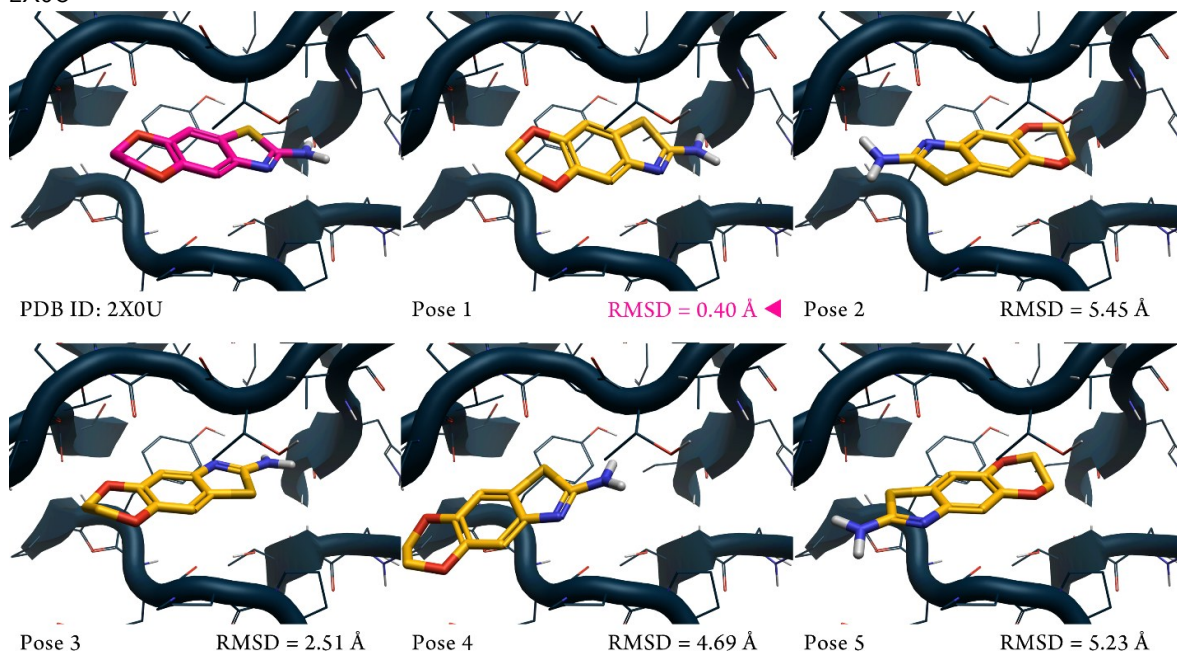


Figure 25. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 2X0U. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

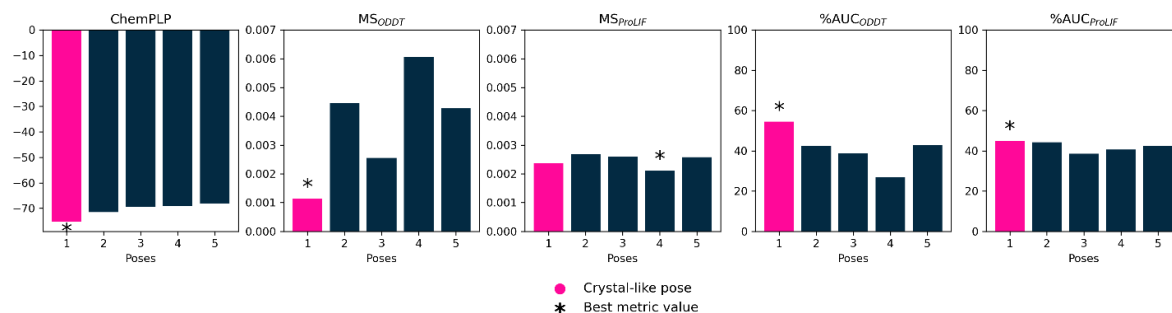


Figure 26. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 2X0U. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

As depicted in Figure 25 and Figure 26, TTMD refinement of docking poses generated for complex 2X0U would successfully identify the native pose among a set of decoys with three out of four scoring metrics, with the only exception being the ProLIF average MS, which would classify the native pose as second best behind pose 4. In this case, docking would also correctly identify the native pose as the top-ranked. However, the differences in scores between native and incorrect poses would be smaller compared to the ODDT average MS, the top-scoring TTMD metric so far, which would lower the confidence in selecting pose 1 as the native binding mode for this ligand.

Tyrosyl-tRNA Synthetase

Tyrosine-tRNA ligase, also known as tyrosyl-tRNA synthetase is an enzyme belonging to the family of the

aminoacyl-tRNA synthetases, a subtype of ligase that forms carbon-oxygen bonds in tRNA and related compounds¹¹³. Since they chemically link amino acids to their cognate tRNA, they are crucial for in vivo translation of the genetic code¹¹⁴. The results of the TTMD refinement carried out on docking poses generated for the examined tyrosine-tRNA ligase protein-ligand complex are summarized below.

1Q11

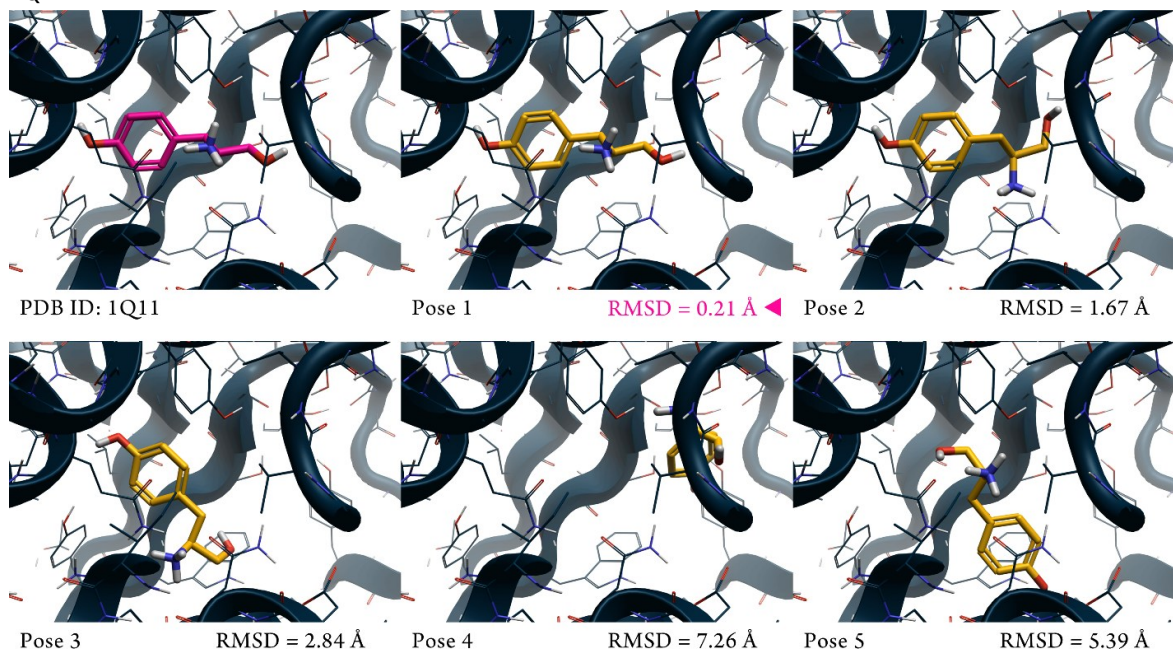


Figure 27. This panel comprises all binding modes (experimental (pink) + docking poses (orange)) investigated in this work for complex deposited in the PDB with accession code 1Q11. For each pose, the RMSD to the experimental pose is reported. The lowest RMSD value is highlighted in pink with a little triangle.

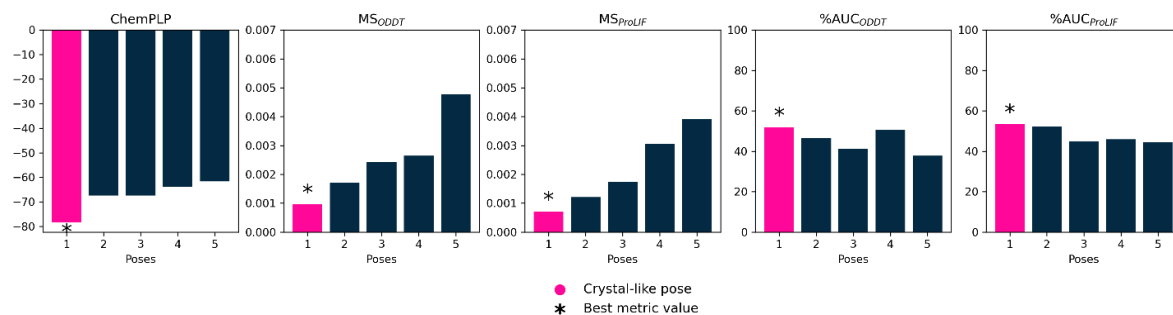


Figure 28. Results of the TTMD simulations performed on the five poses of the complex deposited in the PDB with accession code 1Q11. The bar plot reports each evaluation metric, including the ChemPLP score, ODDT average MS, ProLIF average MS, ODDT %AUC, and ProLIF %AUC. The native-like pose (with the lowest RMSD to the experimental structure) is highlighted in pink, while the top-scoring pose for each metric is marked with an asterisk (*).

As depicted in Figure 27 and 28, both docking and TTMD would be able to highlight pose 1 as the native binding mode for this complex, regardless of the adopted metric. Concerning TTMD, the MS metric once again outperforms the AUC metric. Indeed, the difference in score between the native pose and decoys like pose 4 is small, despite the latest being the worst possible solution according to the RMSD to the experimentally determined bound conformation.

Aggregate Results

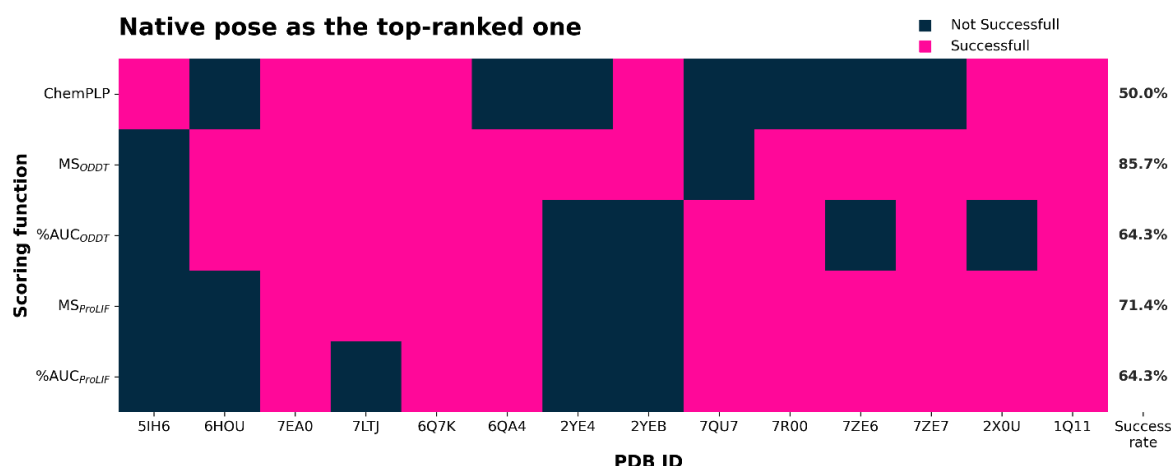


Figure 29. The success rate for the TTMD-refinement of docking poses for test cases presented in the paper. Success is defined as the ability of different scoring metrics to highlight the native-like pose (defined as the one with the lowest RMSD to the experimentally determined one) as the top-ranking solution.

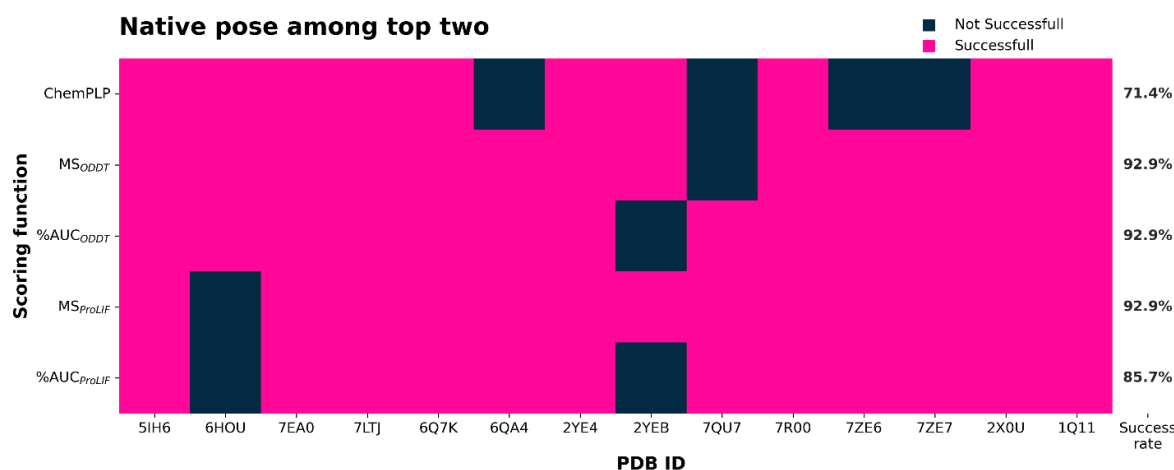


Figure 30. The success rate for the TTMD-refinement of docking poses for test cases presented in the paper. Success is defined as the ability of different scoring metrics to highlight the native-like pose (defined as the one with the lowest RMSD to the experimentally determined one) among the first two top-ranking solutions.

As reported in and Figure 29 and 30, summarizing the ability of the TTMD-refinement to highlight the native-like pose as the top or second-best scoring solution, regardless of the metric used docking is always outperformed by TTMD. Remarkably, the ODDT MS value has nearly double the accuracy (85.7% vs 50.0%) of the ChemPLP scoring function among examined test cases, as it also outperforms every other TTMD-related scoring metric. Notably, even when the success criterion is defined as the ability to prioritize the native pose among the two best-scoring solutions, the ChemPLP success rate is just 71.4%, while every TTMD-related metric is above 85% and ODDT and MS ones are above 90%.

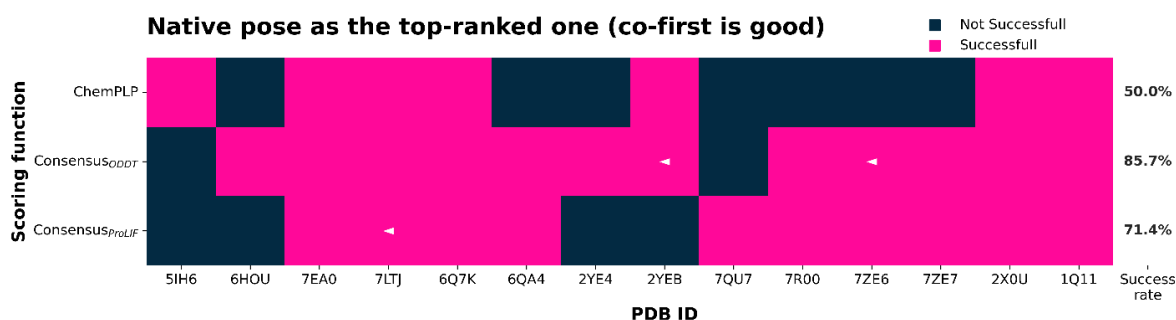


Figure 31. The success rate for the TTMD-refinement of docking poses for test cases presented in the paper. Success is defined as the ability of different scoring metrics to highlight the native-like pose (defined as the one with the lowest RMSD to the experimentally determined one) as the top-ranking solution when a consensus ranking between the MS and AUC is used. Co-first-ranked poses are highlighted with a little triangle.

Regarding the hypothesis of using a consensus ranking between the MS and the AUC metric, instead, as can be seen in Figure 31 this would not improve the accuracy of the TTMD refinement protocol, even by considering co-first-ranked solutions as a positive outcome.

Discussion

Thermal Titration Molecular Dynamics (TTMD) is an alternative enhanced sampling MD-based protocol for the qualitative estimation of receptor-ligand complex stability⁴³. Initially developed to discriminate between high and low-affinity ligands for a given target, it has been successfully used also as a post-docking refinement tool for both protein-ligand⁴⁵ and nucleic acid-ligand complexes⁴⁶. In the previous work from Menin et al.⁴⁵, we showed that TTMD could easily retrieve the native binding mode for mature, druglike ligands across a set of docking poses. The good performances of the method for mature compounds were also confirmed by its ability to interpret the effect of receptor mutations on complex stability using SARS-CoV-2 M^{Pro} as a case study⁴⁴.

However, we also noticed that the protocol could not be readily applied as it was to the characterization of complexes involving fragments, as the previously adopted temperature ramp was too aggressive and would cause premature unbinding for every complex, flattening differences in complex stability.

For this reason, in the present work, we introduced a modified version of the protocol, specifically tuned toward frag-like compounds. Specifically, in virtue of the high intrinsic instability of complexes involving fragments under classic simulation conditions, we designed a temperature ramp that explores subzero temperatures, slowly thawing the complex instead of aggressively heating it up. The necessity to explore this extreme temperature range was assessed after noticing how the intermediate temperature ramp presented in the work of Dodaro et al.⁴⁶ and applied to the study of nucleic acid-ligand complexes was still too aggressive (data not shown).

At first, we tested the performance of the new ramp on four complexes involving low-affinity frag-like ligands extracted from the work of Menin et al.⁴⁵ For all these test cases, the original TTMD refinement of docking poses either led to wrong or inconclusive results. With the new temperature ramp, instead, only one out of four cases (5IH6 - CK1δ) was still unsuccessful (albeit slightly improved compared to the original one), with TTMD being able to identify the native pose with at least one of the scoring metrics adopted. Encouraged by these preliminary results, we decided to further assess the protocol's applicability domain by considering four other challenging pharmaceutically relevant systems, examining two complexes for each target.

As reported in the Results section, we noticed how the new TTMD ramp can easily highlight the native-like pose for frag-like compounds, even ones characterized by very low structural complexity and molecular weight ('MiniFrag's⁵⁷ and 'FragLites'⁵⁹), for the vast majority of examined test cases, even in those cases where docking alone would fail. This trend is similar to the one observed by Bolcato et al.³³, who noticed how the use of Supervised Molecular Dynamics (SuMD) would outperform a state-of-the-art docking tool

like GOLD for posing fragments with molecular weight below 150 Da.

Among the different interaction fingerprints used, we noticed how ODDT fingerprints perform systematically better than ProLIF's ones, regardless of the metric. Among the different metrics, the original MS score outperforms the AUC, making the MS value calculated using ODDT fingerprints the most accurate TTMD scoring metric. This is most likely related to the higher intrinsic instability of fragment poses, which produces ample fingerprint fluctuations even in those cases where the binding mode is retained throughout the simulations.

Remarkably, among considered test cases, TTMD accuracy in identifying the native binding mode as the top-ranking pose was as high as 85.7% (compared to 50.0% for docking alone). Moreover, TTMD still outperformed docking alone when the top two scoring poses were considered (92.9 % for TTMD compared to 71.4 % for docking alone).

The docking protocol (PLANTS-ChemPLP) adopted in the present paper was chosen due to its free-for-academics nature, its accuracy over a wide variety of test cases, which has been demonstrated to be comparable to one of the state-of-the-art proprietary docking tools such as Glide and GOLD and superior to the most popular open-source docking package Vina^{27–29,115} and for consistency with previous TTMD-related works. Some docking protocols (like GOLD, for example) allow the user to incorporate explicit water molecules in the calculation, and it has been shown to improve the performance of docking for cases where the bound water molecules in the active site make a difference³³. However, despite several strategies for predicting the placement of water molecules in the active site being available, and one can certainly take advantage of insights from experimental structures, it is certainly a delicate process, especially when it comes to predicting the displaceability of said water molecules. Furthermore, it would still not address the issue of receptor flexibility. For this reason, we think that a simple MD-based method like TTMD is competitive against other methods, especially since it requires little to no setup, and can be used effectively even with a limited knowledge of the target.

Although the number of test cases is limited and the choice of them could suffer from cherry-picking, demanding a more extensive validation of the technique to give statistical meaning to the results, this proof-of-concept study clearly shows the potential for TTMD to be used as a complementary tool for fragment posing, for validating or refining docking results.

Despite the advancements in the use of experimental techniques such as X-ray crystallography or nuclear magnetic resonance certainly improved accessibility to high-resolution structures of complexes involving fragments, computational tools like TTMD can flank them and make fragment-based drug discovery campaigns more time and cost-efficient. Indeed, although the binding mode for a fragment can be preserved throughout the fragment-to-lead process, it is also true that small structural modifications can sometimes drastically alter the binding mode⁵⁸. Instead of relying purely on determining as many experimental structures as possible during the early stages of fragment evolution, one could strategically combine the experimental results with accurate computational predictions like the one provided by docking combined with TTMD to save both time and money.

Conclusions

In the present work, we presented a modified Thermal Titration Molecular Dynamics (TTMD) post-docking refinement protocol specifically tuned to characterize complexes involving fragments. Compared to the original protocol, which led to wrong or inconclusive results for this class of compounds, the new one demonstrated to be able to identify the native binding mode across a set of docking poses in more than 85% of examined test cases, outscoring a popular and accurate free-for-academics docking protocol like PLANTS-ChemPLP even when the top two scoring poses were considered. These results demonstrate the usefulness of TTMD as a post-docking refinement tool for fragment posing, complementing other existing computational and experimental methodologies in the challenging task of retrieving the correct binding pose for compounds characterized by low molecular weight and structural complexity. The new fragment-specific TTMD protocol can be easily implemented into existing fragment-based drug discovery pipelines thanks to the open-source and modular nature of the code, assisting both the fragment screening and the fragment-to-lead optimization processes.

Supplementary material

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.4c01681>.

Author Contributions

‡ M.P. and S.M. contributed equally. The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

References

- Erlanson, D. A.; McDowell, R. S.; O'Brien, T. Fragment-Based Drug Discovery. *J. Med. Chem.* 2004, 47 (14), 3463–3482. <https://doi.org/10.1021/jm040031v>.
- Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. Fragment-Based Lead Discovery. *Nat. Rev. Drug Discov.* 2004, 3 (8), 660–672. <https://doi.org/10.1038/nrd1467>.
- Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty Years on: The Impact of Fragments on Drug Discovery. *Nat. Rev. Drug Discov.* 2016, 15 (9), 605–619. <https://doi.org/10.1038/nrd.2016.109>.
- Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A 'Rule of Three' for Fragment-Based Lead Discovery? *Drug Discov. Today* 2003, 8 (19), 876–877. [https://doi.org/10.1016/S1359-6446\(03\)02831-9](https://doi.org/10.1016/S1359-6446(03)02831-9).
- Lipinski, C. A. Rule of Five in 2015 and beyond: Target and Ligand Structural Limitations, Ligand Chemistry Structure and Drug Discovery Project Decisions. *Adv. Drug Deliv. Rev.* 2016, 101, 34–41. <https://doi.org/10.1016/j.addr.2016.04.029>.
- Bembek, S. D.; Tounge, B. A.; Reynolds, C. H. Ligand Efficiency and Fragment-Based Drug Discovery. *Drug Discov. Today* 2009, 14 (5), 278–283. <https://doi.org/10.1016/j.drudis.2008.11.007>.
- Leach, A. R.; Hann, M. M. Molecular Complexity and Fragment-Based Drug Discovery: Ten Years On. *Curr. Opin. Chem. Biol.* 2011, 15 (4), 489–496. <https://doi.org/10.1016/j.cbpa.2011.05.008>.
- Schuffenhauer, A.; Ruedisser, S.; Marzinzik, A.; Jahnke, W.; Selzer, P.; Jacoby, E. Library Design for Fragment Based Screening. *Curr. Top. Med. Chem.* 5 (8), 751–762. <https://doi.org/10.2174/1568026054637700>.
- Keserü, G. M.; Erlanson, D. A.; Ferenczy, G. G.; Hann, M. M.; Murray, C. W.; Pickett, S. D. Design Principles for Fragment Libraries: Maximizing the Value of Learnings from Pharma Fragment-Based Drug Discovery (FBDD) Programs for Use in Academia. *J. Med. Chem.* 2016, 59 (18), 8189–8206. <https://doi.org/10.1021/acs.jmedchem.6b00197>.
- Holdgate, G. A. Thermodynamics of Binding Interactions in the Rational Drug Design Process. *Expert Opin. Drug Discov.* 2007, 2 (8), 1103–1114. <https://doi.org/10.1517/17460441.2.8.1103>.
- Neumann, T.; Junker, H.-D.; Schmidt, K.; Sekul, R. SPR-Based Fragment Screening: Advantages and Applications. *Curr. Top. Med. Chem.* 7 (16), 1630–1642. <https://doi.org/10.2174/156802607782341073>.
- Olaru, A.; Bala, C.; Jaffrezic-Renault, N.; Aboul-Enein, H. Y. Surface Plasmon Resonance (SPR) Biosensors in Pharmaceutical Analysis. *Crit. Rev. Anal. Chem.* 2015, 45 (2), 97–105. <https://doi.org/10.1080/10408347.2014.881250>.
- Murray, C. W.; Blundell, T. L. Structural Biology in Fragment-Based Drug Design. *Curr. Opin. Struct. Biol.* 2010, 20 (4), 497–507. <https://doi.org/10.1016/j.sbi.2010.04.003>.
- Jhoti, H.; Cleasby, A.; Verdonk, M.; Williams, G. Fragment-Based Screening Using X-Ray Crystallography and NMR Spectroscopy. *Curr. Opin. Chem. Biol.* 2007, 11 (5), 485–493. <https://doi.org/10.1016/j.cbpa.2007.07.010>.
- Favaro, A.; Sturlese, M. A Novel NMR-Based Protocol to Screen Ultralow Molecular Weight Fragments. *J. Med. Chem.* 2024, 67 (5), 3874–3884. <https://doi.org/10.1021/acs.jmedchem.3c02222>.
- Harner, M. J.; Chauder, B. A.; Phan, J.; Fesik, S. W. Fragment-Based Screening of the Bromodomain of ATAD2. *J. Med. Chem.* 2014, 57 (22), 9687–9692. <https://doi.org/10.1021/jm501035j>.
- Davis, A. M.; St-Gallay, S. A.; Kleywegt, G. J. Limitations and Lessons in the Use of X-Ray Structural Information in Drug Design. *Drug Discov. Today* 2008, 13 (19), 831–841. <https://doi.org/10.1016/j.drudis.2008.06.006>.
- Renaud, J.-P.; Chari, A.; Ciferri, C.; Liu, W.; Rémy, H.-W.; Stark, H.; Wiesmann, C. Cryo-EM in Drug Discovery: Achievements, Limitations and Prospects. *Nat. Rev. Drug Discov.* 2018, 17 (7), 471–492. <https://doi.org/10.1038/nrd.2018.77>.
- Bissaro, M.; Sturlese, M.; Moro, S. The Rise of Molecular Simulations in Fragment-Based Drug Design (FBDD): An Overview. *Drug Discov. Today* 2020, 25 (9), 1693–1701. <https://doi.org/10.1016/j.drudis.2020.06.023>.
- Kumar, A.; Voet, A.; Zhang, K. Y. J. Fragment Based Drug Design: From Experimental to Computational Approaches. *Curr. Med. Chem.* 19 (30), 5128–5147. <https://doi.org/10.2174/092986712803530467>.
- Bassani, D.; Moro, S. Past, Present, and Future Perspectives on Computer-Aided Drug Design Methodologies. *Molecules* 2023, 28 (9), 3906. <https://doi.org/10.3390/molecules28093906>.
- Pavan, M.; Moro, S. Lessons Learnt from COVID-19: Computational Strategies for Facing Present and Future Pandemics. *Int. J. Mol. Sci.* 2023, 24 (5), 4401. <https://doi.org/10.3390/ijms24054401>.
- Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* 2018, 9. <https://doi.org/10.3389/fphar.2018.00923>.

24. Murray, C. W.; Verdonk, M. L. The Consequences of Translational and Rotational Entropy Lost by Small Molecules on Binding to Proteins. *J. Comput. Aided Mol. Des.* 2002, *16* (10), 741–753. <https://doi.org/10.1023/A:1022446720849>.
25. Verdonk, M. L.; Giangreco, I.; Hall, R. J.; Korb, O.; Mortenson, P. N.; Murray, C. W. Docking Performance of Fragments and Druglike Compounds. *J. Med. Chem.* 2011, *54* (15), 5422–5431. <https://doi.org/10.1021/jm200558u>.
26. Chaput, L.; Mouawad, L. Efficient Conformational Sampling and Weak Scoring in Docking Programs? Strategy of the Wisdom of Crowds. *J. Cheminformatics* 2017, *9* (1), 37. <https://doi.org/10.1186/s13321-017-0227-x>.
27. Cuzzolin, A.; Sturlese, M.; Malvacio, I.; Ciancetta, A.; Moro, S. DockBench: An Integrated Informatic Platform Bridging the Gap between the Robust Validation of Docking Protocols and Virtual Screening Simulations. *Molecules* 2015, *20* (6), 9977–9993. <https://doi.org/10.3390/molecules20069977>.
28. Salmaso, V.; Sturlese, M.; Cuzzolin, A.; Moro, S. DockBench as Docking Selector Tool: The Lesson Learned from D3R Grand Challenge 2015. *J. Comput. Aided Mol. Des.* 2016, *30* (9), 773–789. <https://doi.org/10.1007/s10822-016-9966-4>.
29. Bolcato, G.; Cuzzolin, A.; Bissaro, M.; Moro, S.; Sturlese, M. Can We Still Trust Docking Results? An Extension of the Applicability of DockBench on PDBbind Database. *Int. J. Mol. Sci.* 2019, *20* (14), 3558. <https://doi.org/10.3390/ijms20143558>.
30. Bassani, D.; Pavan, M.; Bolcato, G.; Sturlese, M.; Moro, S. Re-Exploring the Ability of Common Docking Programs to Correctly Reproduce the Binding Modes of Non-Covalent Inhibitors of SARS-CoV-2 Protease Mpro. *Pharmaceuticals* 2022, *15* (2), 180. <https://doi.org/10.3390/ph15020180>.
31. Bolcato, G.; Cescon, E.; Pavan, M.; Bissaro, M.; Bassani, D.; Federico, S.; Spalluto, G.; Sturlese, M.; Moro, S. A Computational Workflow for the Identification of Novel Fragments Acting as Inhibitors of the Activity of Protein Kinase CK1δ. *Int. J. Mol. Sci.* 2021, *22* (18), 9741. <https://doi.org/10.3390/ijms22189741>.
32. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Implementing a Scoring Function Based on Interaction Fingerprint for Autogrow4: Protein Kinase CK1δ as a Case Study. *Front. Mol. Biosci.* 2022, *9*. <https://doi.org/10.3389/fmolb.2022.909499>.
33. Bolcato, G.; Bissaro, M.; Sturlese, M.; Moro, S. Comparing Fragment Binding Poses Prediction Using HSP90 as a Key Study: When Bound Water Makes the Difference. *Molecules* 2020, *25* (20), 4651. <https://doi.org/10.3390/molecules25204651>.
34. Cuzzolin, A.; Deganutti, G.; Salmaso, V.; Sturlese, M.; Moro, S. AquaMMapS: An Alternative Tool to Monitor the Role of Water Molecules During Protein–Ligand Association. *ChemMedChem* 2018, *13* (6), 522–531. <https://doi.org/10.1002/cmdc.201700564>.
35. Greisman, J. B.; Willmore, L.; Yeh, C. Y.; Giordanetto, F.; Shahamadtar, S.; Nisonoff, H.; Maragakis, P.; Shaw, D. E. Discovery and Validation of the Binding Poses of Allosteric Fragment Hits to Protein Tyrosine Phosphatase 1b: From Molecular Dynamics Simulations to X-Ray Crystallography. *J. Chem. Inf. Model.* 2023, *63* (9), 2644–2650. <https://doi.org/10.1021/acs.jcim.3c00236>.
36. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 2016, *59* (9), 4035–4061. <https://doi.org/10.1021/acs.jmedchem.5b01684>.
37. Sabbadin, D.; Moro, S. Supervised Molecular Dynamics (SuMD) as a Helpful Tool To Depict GPCR–Ligand Recognition Pathway in a Nanosecond Time Scale. *J. Chem. Inf. Model.* 2014, *54* (2), 372–376. <https://doi.org/10.1021/ci400766b>.
38. Deganutti, G.; Moro, S. Supporting the Identification of Novel Fragment-Based Positive Allosteric Modulators Using a Supervised Molecular Dynamics Approach: A Retrospective Analysis Considering the Human A2A Adenosine Receptor as a Key Example. *Molecules* 2017, *22* (5), 818. <https://doi.org/10.3390/molecules22050818>.
39. Bissaro, M.; Bolcato, G.; Pavan, M.; Bassani, D.; Sturlese, M.; Moro, S. Inspecting the Mechanism of Fragment Hits Binding on SARS-CoV-2 Mpro by Using Supervised Molecular Dynamics (SuMD) Simulations. *ChemMedChem* 2021, *16* (13), 2075–2081. <https://doi.org/10.1002/cmdc.202100156>.
40. Grosjean, H.; Işık, M.; Aimon, A.; Mobley, D.; Chodera, J.; von Delft, F.; Biggin, P. C. SAMPL7 Protein–Ligand Challenge: A Community-Wide Evaluation of Computational Methods against Fragment Screening and Pose-Prediction. *J. Comput. Aided Mol. Des.* 2022, *36* (4), 291–311. <https://doi.org/10.1007/s10822-022-00452-7>.
41. Ferrari, F.; Bissaro, M.; Fabbian, S.; De Almeida Roger, J.; Mammi, S.; Moro, S.; Bellanda, M.; Sturlese, M. HT-SuMD: Making Molecular Dynamics Simulations Suitable for Fragment-Based Screening. A Comparative Study with NMR. *J. Enzyme Inhib. Med. Chem.* 2021, *36* (1), 1–14. <https://doi.org/10.1080/14756366.2020.1838499>.
42. Ibrahim, P. E. G. F.; Zuccotto, F.; Zachariae, U.; Gilbert, I.; Bodkin, M. Accurate Prediction of Dynamic Protein–Ligand Binding Using P-Score Ranking. *J. Comput. Chem.* n/a (n/a). <https://doi.org/10.1002/jcc.27370>.
43. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Qualitative Estimation of Protein–Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2022, *62* (22), 5715–5728. <https://doi.org/10.1021/acs.jcim.2c00995>.
44. Costacurta, F.; Dodaro, A.; Bante, D.; Schöppe, H.; Sprenger, B.; Moghadasi, S. A.; Fleischmann, J.; Pavan, M.; Bassani, D.; Menin, S.; Rauch, S.; Krüger, L.; Sauerwein, A.; Heberle, A.; Rabensteiner, T.; Ho, J.; Harris, R. S.; Stefan, E.; Schneider, R.; Kaserer, T.; Moro, S.; Laer, D. von; Heilmann, E. A Comprehensive Study of SARS-CoV-2 Main Protease (Mpro) Inhibitor-Resistant Mutants Selected in a VSV-Based System. *bioRxiv* September 25, 2023, p 2023.09.22.558628. <https://doi.org/10.1101/2023.09.22.558628>.
45. Menin, S.; Pavan, M.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement. *Int. J. Mol. Sci.* 2023, *24* (4), 3596. <https://doi.org/10.3390/ijms24043596>.

46. Dodaro, A.; Pavan, M.; Menin, S.; Salmaso, V.; Sturlese, M.; Moro, S. Thermal Titration Molecular Dynamics (TTMD): Shedding Light on the Stability of RNA-Small Molecule Complexes. *Front. Mol. Biosci.* 2023, *10*. <https://doi.org/10.3389/fmolb.2023.1294543>.
47. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* 2000, *28* (1), 235–242. <https://doi.org/10.1093/nar/28.1.235>.
48. Molecular Operating Environment (MOE), 2022.02 Chemical Computing Group ULC, 910-1010 Sherbrooke St. W., Montreal, QC H3A 2R7, Canada, 2024.
49. Halgren, T. A. Force Fields: MMFF94. In *Encyclopedia of Computational Chemistry*; John Wiley & Sons, Ltd, 2002. <https://doi.org/10.1002/0470845015.cma012m>.
50. Korb, O.; Stützle, T.; Exner, T. E. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design. In *Ant Colony Optimization and Swarm Intelligence*; Dorigo, M., Gambardella, L. M., Birattari, M., Martinoli, A., Poli, R., Stützle, T., Eds.; Springer: Berlin, Heidelberg, 2006; pp 247–258. https://doi.org/10.1007/11839088_22.
51. Korb, O.; Stützle, T.; Exner, T. E. An Ant Colony Optimization Approach to Flexible Protein–Ligand Docking. *Swarm Intell.* 2007, *1* (2), 115–134. <https://doi.org/10.1007/s11721-007-0006-9>.
52. Korb, O.; Stützle, T.; Exner, T. E. Empirical Scoring Functions for Advanced Protein–Ligand Docking with PLANTS. *J. Chem. Inf. Model.* 2009, *49* (1), 84–96. <https://doi.org/10.1021/ci800298z>.
53. Ursu, A.; Illich, D. J.; Takemoto, Y.; Porfetye, A. T.; Zhang, M.; Brockmeyer, A.; Janning, P.; Watanabe, N.; Osada, H.; Vetter, I. R.; Ziegler, S.; Schöler, H. R.; Waldmann, H. Epiblastin A Induces Reprogramming of Epiblast Stem Cells Into Embryonic Stem Cells by Inhibition of Casein Kinase 1. *Cell Chem. Biol.* 2016, *23* (4), 494–507. <https://doi.org/10.1016/j.chembiol.2016.02.015>.
54. Cozza, G.; Zonta, F.; Dalle Vedove, A.; Venerando, A.; Dall'Acqua, S.; Battistutta, R.; Ruzzene, M.; Lolli, G. Biochemical and Cellular Mechanism of Protein Kinase CK2 Inhibition by Deceptive Curcumin. *FEBS J.* 2020, *287* (9), 1850–1864. <https://doi.org/10.1111/febs.15111>.
55. Akaki, T.; Bessho, Y.; Ito, T.; Fujioka, S.; Ubukata, M.; Mori, G.; Yamanaka, K.; Orita, T.; Doi, S.; Iwanaga, T.; Ikegashira, K.; Hantani, Y.; Nakanishi, I.; Adachi, T. Fragment-Based Lead Discovery to Identify Novel Inhibitors That Target the ATP Binding Site of Pyruvate Dehydrogenase Kinases. *Bioorg. Med. Chem.* 2021, *44*, 116283. <https://doi.org/10.1016/j.bmc.2021.116283>.
56. Clyde, A.; Galanie, S.; Kneller, D. W.; Ma, H.; Babuji, Y.; Blaiszik, B.; Brace, A.; Brettin, T.; Chard, K.; Chard, R.; Coates, L.; Foster, I.; Hauner, D.; Kertesz, V.; Kumar, N.; Lee, H.; Li, Z.; Merzky, A.; Schmidt, J. G.; Tan, L.; Titov, M.; Trifan, A.; Turilli, M.; Van Dam, H.; Chennubhotla, S. C.; Jha, S.; Kovalevsky, A.; Ramanathan, A.; Head, M. S.; Stevens, R. High-Throughput Virtual Screening and Validation of a SARS-CoV-2 Main Protease Noncovalent Inhibitor. *J. Chem. Inf. Model.* 2022, *62* (1), 116–128. <https://doi.org/10.1021/acs.jcim.1c00851>.
57. O'Reilly, M.; Cleasby, A.; Davies, T. G.; Hall, R. J.; Ludlow, R. F.; Murray, C. W.; Tisi, D.; Jhoti, H. Crystallographic Screening Using Ultra-Low-Molecular-Weight Ligands to Guide Drug Design. *Drug Discov. Today* 2019, *24* (5), 1081–1086. <https://doi.org/10.1016/j.drudis.2019.03.009>.
58. Roughley, S. D.; Hubbard, R. E. How Well Can Fragments Explore Accessed Chemical Space? A Case Study from Heat Shock Protein 90. *J. Med. Chem.* 2011, *54* (12), 3989–4005. <https://doi.org/10.1021/jm200350g>.
59. Davison, G.; Martin, M. P.; Turberville, S.; Dorman, S.; Heath, R.; Heptinstall, A. B.; Lawson, M.; Miller, D. C.; Ng, Y. M.; Sanderson, J. N.; Hope, I.; Wood, D. J.; Cano, C.; Endicott, J. A.; Hardcastle, I. R.; Noble, M. E. M.; Waring, M. J. Mapping Ligand Interactions of Bromodomains BRD4 and ATAD2 with FragLites and PepLites—Halogenated Probes of Druglike and Peptide-like Molecular Interactions. *J. Med. Chem.* 2022, *65* (22), 15416–15432. <https://doi.org/10.1021/acs.jmedchem.2c01357>.
60. Basse, N.; Kaar, J. L.; Settanni, G.; Joerger, A. C.; Rutherford, T. J.; Fersht, A. R. Toward the Rational Design of P53-Stabilizing Drugs: Probing the Surface of the Oncogenic Y220C Mutant. *Chem. Biol.* 2010, *17* (1), 46–56. <https://doi.org/10.1016/j.chembiol.2009.12.011>.
61. Yang, X.-L.; Otero, F. J.; Skene, R. J.; McRee, D. E.; Schimmel, P.; Ribas de Pouplana, L. Crystal Structures That Suggest Late Development of Genetic Code Components for Differentiating Aromatic Side Chains. *Proc. Natl. Acad. Sci.* 2003, *100* (26), 15376–15380. <https://doi.org/10.1073/pnas.2136794100>.
62. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* 1996, *14* (1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
63. Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* 2005, *26* (16), 1668–1688. <https://doi.org/10.1002/jcc.20290>.
64. *AmberTools*. <https://ambermd.org/AmberTools.php> (accessed 2024-05-02).
65. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 2015, *11* (8), 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>.
66. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 2004, *25* (9), 1157–1174. <https://doi.org/10.1002/jcc.20035>.
67. Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* 2002, *23* (16), 1623–1641. <https://doi.org/10.1002/jcc.10128>.
68. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 1983, *79* (2), 926–935. <https://doi.org/10.1063/1.445869>.

69. Davidchack, R. L.; Handel, R.; Tretyakov, M. V. Langevin Thermostat for Rigid Body Dynamics. *J. Chem. Phys.* 2009, **130** (23), 234101. <https://doi.org/10.1063/1.3149788>.
70. Faller, R.; de Pablo, J. J. Constant Pressure Hybrid Molecular Dynamics–Monte Carlo Simulations. *J. Chem. Phys.* 2002, **116** (1), 55–59. <https://doi.org/10.1063/1.1420460>.
71. Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A Fast SHAKE Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. *J. Comput. Chem.* 2001, **22** (5), 501–508. [https://doi.org/10.1002/1096-987X\(20010415\)22:5<501::AID-JCC1021>3.0.CO;2-V](https://doi.org/10.1002/1096-987X(20010415)22:5<501::AID-JCC1021>3.0.CO;2-V).
72. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 1995, **103** (19), 8577–8593. <https://doi.org/10.1063/1.470117>.
73. Harvey, M. J.; Giupponi, G.; Fabritiis, G. D. ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* 2009, **5** (6), 1632–1639. <https://doi.org/10.1021/ct9000685>.
74. Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLOS Comput. Biol.* 2017, **13** (7), e1005659. <https://doi.org/10.1371/journal.pcbi.1005659>.
75. Wójcikowski, M.; Zielenkiewicz, P.; Siedlecki, P. Open Drug Discovery Toolkit (ODDT): A New Open-Source Player in the Drug Discovery Field. *J. Cheminformatics* 2015, **7** (1), 26. <https://doi.org/10.1186/s13321-015-0078-2>.
76. Bouysset, C.; Fiorucci, S. ProLIF: A Library to Encode Molecular Interactions as Fingerprints. *J. Cheminformatics* 2021, **13** (1), 72. <https://doi.org/10.1186/s13321-021-00548-6>.
77. Matsumoto, M.; Saito, S.; Ohmine, I. Molecular Dynamics Simulation of the Ice Nucleation and Growth Process Leading to Water Freezing. *Nature* 2002, **416** (6879), 409–413. <https://doi.org/10.1038/416409a>.
78. Fisher, C. L.; Pavan, M.; Salmaso, V.; Keyes, R. F.; Wan, T. C.; Pradhan, B.; Gao, Z.-G.; Smith, B. C.; Jacobson, K. A.; Auchampach, J. A. Extrahelical Binding Site for a 1H-Imidazo[4,5-c]Quinolin-4-Amine A3 Adenosine Receptor Positive Allosteric Modulator on Helix 8 and Distal Portions of Transmembrane Domains 1 and 7. *Mol. Pharmacol.* 2024, **105** (3), 213–223. <https://doi.org/10.1124/molpharm.123.000784>.
79. Pradhan, B.; Pavan, M.; Fisher, C. L.; Salmaso, V.; Wan, T. C.; Keyes, R. F.; Rollison, N.; Suresh, R. R.; Kumar, T. S.; Gao, Z.-G.; Smith, B. C.; Auchampach, J. A.; Jacobson, K. A. Lipid Trolling to Optimize A3 Adenosine Receptor-Positive Allosteric Modulators (PAMs). *J. Med. Chem.* 2024, **67** (14), 12221–12247. <https://doi.org/10.1021/acs.jmedchem.4c00944>.
80. Tosh, D. K.; Pavan, M.; Cronin, C.; Pottie, E.; Wan, T. C.; Chen, E.; Lewicki, S. A.; Campbell, R. G.; Gao, Z.-G.; Auchampach, J. A.; Stove, C. P.; Liang, B. T.; Jacobson, K. A. 2-Substituted (N)-Methanocarba A3 Adenosine Receptor Agonists: In Silico, In Vitro, and In Vivo Characterization. *ACS Pharmacol. Transl. Sci.* 2024, **7** (7), 2154–2173. <https://doi.org/10.1021/acsptsci.4c00223>.
81. Knippschild, U.; Gocht, A.; Wolff, S.; Huber, N.; Löhler, J.; Stöter, M. The Casein Kinase 1 Family: Participation in Multiple Cellular Processes in Eukaryotes. *Cell. Signal.* 2005, **17** (6), 675–689. <https://doi.org/10.1016/j.cellsig.2004.12.011>.
82. Perez, D. I.; Gil, C.; Martinez, A. Protein Kinases CK1 and CK2 as New Targets for Neurodegenerative Diseases. *Med. Res. Rev.* 2011, **31** (6), 924–954. <https://doi.org/10.1002/med.20207>.
83. Xu, P.; Ianes, C.; Gärtner, F.; Liu, C.; Burster, T.; Bakulev, V.; Rachidi, N.; Knippschild, U.; Bischof, J. Structure, Regulation, and (Patho-)Physiological Functions of the Stress-Induced Protein Kinase CK1 Delta (CSNK1D). *Gene* 2019, **715**, 144005. <https://doi.org/10.1016/j.gene.2019.144005>.
84. Venerando, A.; Ruzzene, M.; Pinna, L. A. Casein Kinase: The Triple Meaning of a Misnomer. *Biochem. J.* 2014, **460** (2), 141–156. <https://doi.org/10.1042/BJ20140178>.
85. Meggio, F.; Pinna, L. A. One-Thousand-and-One Substrates of Protein Kinase CK2? *FASEB J.* 2003, **17** (3), 349–368. <https://doi.org/10.1096/fj.02-0473rev>.
86. Borgo, C.; D'Amore, C.; Sarno, S.; Salvi, M.; Ruzzene, M. Protein Kinase CK2: A Potential Therapeutic Target for Diverse Human Diseases. *Signal Transduct. Target. Ther.* 2021, **6** (1), 1–20. <https://doi.org/10.1038/s41392-021-00567-7>.
87. Zhang, S.; Hulver, M. W.; McMillan, R. P.; Cline, M. A.; Gilbert, E. R. The Pivotal Role of Pyruvate Dehydrogenase Kinases in Metabolic Flexibility. *Nutr. Metab.* 2014, **11** (1), 10. <https://doi.org/10.1186/1743-7075-11-10>.
88. Atas, E.; Oberhuber, M.; Kenner, L. The Implications of PDK1–4 on Tumor Energy Metabolism, Aggressiveness and Therapy Resistance. *Front. Oncol.* 2020, **10**. <https://doi.org/10.3389/fonc.2020.583217>.
89. Ullrich, S.; Nitsche, C. The SARS-CoV-2 Main Protease as Drug Target. *Bioorg. Med. Chem. Lett.* 2020, **30** (17), 127377. <https://doi.org/10.1016/j.bmcl.2020.127377>.
90. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature* 2020, **582** (7811), 289–293. <https://doi.org/10.1038/s41586-020-2223-y>.
91. Roskoski, R. ERK1/2 MAP Kinases: Structure, Function, and Regulation. *Pharmacol. Res.* 2012, **66** (2), 105–143. <https://doi.org/10.1016/j.phrs.2012.04.005>.
92. Lavoie, H.; Gagnon, J.; Therrien, M. ERK Signalling: A Master Regulator of Cell Behaviour, Life and Fate. *Nat. Rev. Mol. Cell Biol.* 2020, **21** (10), 607–632. <https://doi.org/10.1038/s41580-020-0255-7>.
93. Dutta, R.; Inouye, M. GHKL, an Emergent ATPase/Kinase Superfamily. *Trends Biochem. Sci.* 2000, **25** (1), 24–28. [https://doi.org/10.1016/S0968-0004\(99\)01503-0](https://doi.org/10.1016/S0968-0004(99)01503-0).

94. Powers, M. V.; Workman, P. Targeting of Multiple Signalling Pathways by Heat Shock Protein 90 Molecular Chaperone Inhibitors. 2006. <https://doi.org/10.1677/erc.1.01324>.
95. Pearl, L. H.; Prodromou, C. Structure and Mechanism of the Hsp90 Molecular Chaperone Machinery. *Annu. Rev. Biochem.* 2006, **75** (Volume 75, 2006), 271–294. <https://doi.org/10.1146/annurev.biochem.75.103004.142738>.
96. Biamonte, M. A.; Van de Water, R.; Arndt, J. W.; Scannevin, R. H.; Perret, D.; Lee, W.-C. Heat Shock Protein 90: Inhibitors in Clinical Trials. *J. Med. Chem.* 2010, **53** (1), 3–17. <https://doi.org/10.1021/jm9004708>.
97. Janin, Y. L. ATPase Inhibitors of Heat-Shock Protein 90, Second Season. *Drug Discov. Today* 2010, **15** (9), 342–353. <https://doi.org/10.1016/j.drudis.2010.03.002>.
98. Cochran, A. G.; Conery, A. R.; Sims, R. J. Bromodomains: A New Target Class for Drug Development. *Nat. Rev. Drug Discov.* 2019, **18** (8), 609–628. <https://doi.org/10.1038/s41573-019-0030-7>.
99. Fu, J.; Zhang, J.; Chen, X.; Liu, Z.; Yang, X.; He, Z.; Hao, Y.; Liu, B.; Yao, D. ATPase Family AAA Domain-Containing Protein 2 (ATAD2): From an Epigenetic Modulator to Cancer Therapeutic Target. *Theranostics* 2023, **13** (2), 787–809. <https://doi.org/10.7150/thno.78840>.
100. Nayak, A.; Dutta, M.; Roychowdhury, A. Emerging Oncogene ATAD2: Signaling Cascades and Therapeutic Initiatives. *Life Sci.* 2021, **276**, 119322. <https://doi.org/10.1016/j.lfs.2021.119322>.
101. Vidler, L. R.; Brown, N.; Knapp, S.; Hoelder, S. Druggability Analysis and Structural Classification of Bromodomain Acetyl-Lysine Binding Sites. *J. Med. Chem.* 2012, **55** (17), 7346–7359. <https://doi.org/10.1021/jm300346w>.
102. Demont, E. H.; Chung, C.; Furze, R. C.; Grandi, P.; Michon, A.-M.; Wellaway, C.; Barrett, N.; Bridges, A. M.; Craggs, P. D.; Diallo, H.; Dixon, D. P.; Douault, C.; Emmons, A. J.; Jones, E. J.; Karamshi, B. V.; Locke, K.; Mitchell, D. J.; Mouzon, B. H.; Prinjha, R. K.; Roberts, A. D.; Sheppard, R. J.; Watson, R. J.; Bamborough, P. Fragment-Based Discovery of Low-Micromolar ATAD2 Bromodomain Inhibitors. *J. Med. Chem.* 2015, **58** (14), 5649–5673. <https://doi.org/10.1021/acs.jmedchem.5b00772>.
103. Chaikuad, A.; Petros, A. M.; Fedorov, O.; Xu, J.; Knapp, S. Structure-Based Approaches towards Identification of Fragments for the Low-Druggability ATAD2 Bromodomain. *MedChemComm* 2014, **5** (12), 1843–1848. <https://doi.org/10.1039/C4MD00237G>.
104. Donati, B.; Lorenzini, E.; Ciarrocchi, A. BRD4 and Cancer: Going beyond Transcriptional Regulation. *Mol. Cancer* 2018, **17** (1), 164. <https://doi.org/10.1186/s12943-018-0915-9>.
105. Liang, D.; Yu, Y.; Ma, Z. Novel Strategies Targeting Bromodomain-Containing Protein 4 (BRD4) for Cancer Drug Discovery. *Eur. J. Med. Chem.* 2020, **200**, 112426. <https://doi.org/10.1016/j.ejmech.2020.112426>.
106. Duan, Y.; Guan, Y.; Qin, W.; Zhai, X.; Yu, B.; Liu, H. Targeting Brd4 for Cancer Therapy: Inhibitors and Degraders. *MedChemComm* 2018, **9** (11), 1779–1802. <https://doi.org/10.1039/C8MD00198G>.
107. Lu, T.; Lu, W.; Luo, C. A Patent Review of BRD4 Inhibitors (2013–2019). *Expert Opin. Ther. Pat.* 2020, **30** (1), 57–81. <https://doi.org/10.1080/13543776.2020.1702645>.
108. Matlashewski, G.; Lamb, P.; Pim, D.; Peacock, J.; Crawford, L.; Benchimol, S. Isolation and Characterization of a Human P53 cDNA Clone: Expression of the Human P53 Gene. *EMBO J.* 1984, **3** (13), 3257–3262. <https://doi.org/10.1002/j.1460-2075.1984.tb02287.x>.
109. Toufekhtchan, E.; Toledo, F. The Guardian of the Genome Revisited: P53 Downregulates Genes Required for Telomere Maintenance, DNA Repair, and Centromere Structure. *Cancers* 2018, **10** (5), 135. <https://doi.org/10.3390/cancers10050135>.
110. Surget, S.; Khoury, M. P.; Bourdon, J.-C. Uncovering the Role of P53 Splice Variants in Human Malignancy: A Clinical Perspective. *OncoTargets Ther.* 2013, **7**, 57–68. <https://doi.org/10.2147/OTT.S53876>.
111. Hann, B. C.; Lane, D. P. The Dominating Effect of Mutant P53. *Nat. Genet.* 1995, **9** (3), 221–222. <https://doi.org/10.1038/ng0395-221>.
112. Boeckler, F. M.; Joerger, A. C.; Jaggi, G.; Rutherford, T. J.; Veprintsev, D. B.; Fersht, A. R. Targeted Rescue of a Destabilized Mutant of P53 by an in Silico Screened Drug. *Proc. Natl. Acad. Sci.* 2008, **105** (30), 10360–10365. <https://doi.org/10.1073/pnas.0805326105>.
113. Jr, C. W. C. COGNITION, MECHANISM, AND EVOLUTIONARY RELATIONSHIPS IN AMINOACYL-tRNA SYNTHETASES. *Annu. Rev. Biochem.* 1993, **62** (Volume 62, 1993), 715–748. <https://doi.org/10.1146/annurev.bi.62.070193.003435>.
114. Ibba, M.; Morgan, S.; Curnow, A. W.; Pridmore, D. R.; Vothknecht, U. C.; Gardner, W.; Lin, W.; Woese, C. R.; Söll, D. A Euryarchaeal Lysyl-tRNA Synthetase: Resemblance to Class I Synthetases. *Science* 1997, **278** (5340), 1119–1122. <https://doi.org/10.1126/science.278.5340.1119>.
115. Salmaso, V.; Sturlese, M.; Cuzzolin, A.; Moro, S. Combining Self- and Cross-Docking as Benchmark Tools: The Performance of DockBench in the D3R Grand Challenge 2. *J. Comput. Aided Mol. Des.* 2018, **32** (1), 251–264. <https://doi.org/10.1007/s10822-017-0051-4>.
- 111.

5. Unpublished Works

Menin, S.; Salmaso, V.; Bissaro, M.; Esposito, C.; Pozzan, A.; Moro, S. Decoding GPCR-Ligand Interactions with FEP: Evaluating Key Parameters for Reliable Predictions. In preparation.

This study benchmarks the Flare FEP relative binding free energy (RBFEE) implementation in the context of membrane proteins, focusing on GPCR targets. Using the adenosine A_{2A} receptor as a primary test case, we systematically evaluated the influence of ligand placement, membrane inclusion, water treatment, and ligand force field on predictive performance. Results highlighted ligand placement and water handling as the dominant factors affecting accuracy, while force field choice (GAFF2 vs. OpenFF2) had only minor effects. To broaden the scope, additional benchmarks were conducted on alternative A_{2A} scaffolds, mGlu₅, and P2Y₁ receptors. Overall, the work establishes the robustness and transferability of Flare FEP across diverse GPCR systems, while identifying the key parameters that govern predictive quality.

Salmaso, V.; Menin, S.; Moro, S.; Spalluto, G.; Federico, S. Adenosine receptors in neuroinflammation and neurodegeneration. Submitted.

This review discusses the role of adenosine and its four receptor subtypes (A₁, A_{2A}, A_{2B}, and A₃) in neuroinflammation and neurodegenerative diseases. Neuroinflammation can have both protective and harmful effects on the central nervous system, contributing to the onset and progression of various brain disorders, including life-threatening neurodegenerative conditions. The review highlights the therapeutic potential of targeting specific adenosine receptors, noting that while some applications are already clinically established, others remain promising areas for further research.

Decoding GPCR-Ligand Interactions with FEP: Evaluating Key Parameters for Reliable Predictions

Silvia Menin¹, Maicol Bissaro², Veronica Salmaso¹, Carmen Esposito², Alfonso Pozzan², Stefano Moro^{*1}

Article in preparation.

Abstract

Predicting ligand binding free energies with high accuracy remains a key challenge in drug discovery, particularly for G protein-coupled receptors (GPCRs). Relative binding free energy (RBFE) calculations are increasingly applied, yet their performance in membrane systems has not been systematically benchmarked. Here, we assessed the Flare FEP RBFE implementation using a triazine antagonist series bound to the adenosine A₂A receptor, a well-characterized GPCR test case. We investigated four key parameters: ligand placement, membrane inclusion, water treatment, and ligand force field. Results showed that predictive quality was driven mainly by ligand placement and water handling. Force field choice (GAFF2 vs. OpenFF2) had only minor effects. To evaluate the impact of the forcefield choice, we extended the benchmark to three additional ligand series: an alternative A₂A scaffold, metabotropic glutamate receptor (mGlu₅), and purinergic receptor P2Y₁. Across systems, both force fields performed similarly, with accuracy depending more on ligand diversity and binding environment than on parameter choice.

Introduction

Reliable prediction of ligand binding affinities is a central goal in early-stage drug discovery, where even modest improvements in accuracy can accelerate lead optimization and reduce experimental costs^{1,2}. Over the past decade, alchemical free-energy methods have emerged as one of the most rigorous computational approaches for this task³. By transforming one ligand into another in both the protein binding site and bulk solvent, relative binding free energy (RBFE) calculations provide quantitative estimates of $\Delta\Delta G$ values that can be directly compared to experimental affinity⁴. These methods, when carefully applied, have demonstrated mean unsigned errors below 1 kcal/mol across diverse soluble protein targets, highlighting their potential to guide medicinal chemistry with predictive power⁵.

The practical application of RBFE has been enabled by advances in molecular dynamics force fields, GPU-accelerated hardware, and automated workflows. Several benchmark studies on soluble proteins such as kinases, proteases, and nuclear hormone receptors have established best practices for applying this methodology across multiple datasets⁵⁻⁷. Despite these successes, the extension of RBFE to membrane proteins remains less explored. Membrane-embedded receptors, particularly G protein-coupled receptors (GPCRs), present unique challenges: their orthosteric pockets are deeply buried within the lipid bilayer, ligand binding often depends on structured water molecules⁸, and receptor flexibility is strongly coupled to the surrounding membrane environment. These factors complicate system setup and sampling, and small inaccuracies in water placement, lipid representation, or ligand positioning can propagate into large errors in predicted free energies⁹.

Several commercial and open-source implementations of RBFE exist, with Schrödinger's FEP+ being the most widely validated on large soluble-protein datasets⁵. More recently, Cresset's Flare software has introduced an FEP-RBFE module built on the Sire framework and leveraging open-source tools such as

BioSimSpace¹⁰, SOMD¹¹, OpenMM¹², and AmberTools¹³. Flare provides automated setup, execution, and analysis of RBFE calculations through both a graphical interface and a Python API, lowering the barrier to entry for non-expert users. However, while Flare FEP has shown promising results on soluble proteins, its performance on membrane-bound receptors has not yet been systematically assessed¹⁴. Without such validation, users face uncertainty in how to configure key aspects of membrane simulations, including lipid bilayer representation, water placement strategies, ligand alignment, and force-field selection.

To address this gap, we carried out a Flare FEP benchmark study on the human adenosine A_{2A} receptor, a class A GPCR with extensive structural and pharmacological characterization. A_{2A} is a clinically relevant target in immuno-oncology and neurodegeneration^{15,16}, and its triazine antagonist series provides both high-resolution X-ray complexes and a rich set of binding affinity data¹⁷. This well-characterized dataset offers an ideal proving ground for testing RBFE workflows on GPCRs and allows direct comparison to previously published studies. In our benchmark, we systematically evaluated the impact of four critical parameters, such as ligand placement, membrane handling, water treatment, and ligand force field, on the accuracy of predicted $\Delta\Delta G$ values.

To extend this investigation beyond the initial triazine series of A_{2A} antagonists, we further included an additional ligand series targeting A_{2A}, but with a different chemical scaffold, thereby testing the transferability of optimized protocols across chemotypes.

We also broadened the scope to two other GPCR systems that have been previously used in benchmarking studies^{18,19}. The first was the metabotropic glutamate (mGlu₅) receptor, a GPCR that plays a central role in modulating excitatory neurotransmission in the central nervous system., and have been the focus of intense drug discovery efforts, particularly through the development of allosteric modulators that fine-tune receptor signaling²⁰.

The second additional system was the P2Y₁ receptor, a purinergic GPCR activated by ADP and involved in platelet aggregation, vascular regulation, and pain signaling²¹. Importantly, structural studies have revealed that P2Y₁ possesses two distinct ligand-binding sites: an orthosteric pocket within the transmembrane bundle and an allosteric site located at the receptor-lipid bilayer interface²². The latter is of particular interest because it represents one of the first structurally characterized examples of a GPCR allosteric modulator binding entirely outside the helical bundle, highlighting the challenges of modeling lipid-exposed allosteric sites in RBFE workflows.

By incorporating these additional ligand series, we created a diverse benchmark spanning different GPCR classes, binding site environments, and ligand series. This expanded dataset allowed us to test the robustness of Flare FEP across a broader chemical and structural space.

Materials and Methods

This benchmark was performed using a workstation equipped with 24 CPU cores for proteins and projects preparation, while the FEP calculations were conducted on a separate workstation equipped with four NVIDIA RTX 4090 GPUs.

Receptor Structure Preparation

As in the study by Deflorian and colleagues¹⁷, the crystal structure of the Adenosine receptor A_{2A} used in the primary section of this benchmark was obtained from the Protein Data Bank (PDB²³) under accession code 6GT3²⁴. To broaden the benchmark, additional receptor structures were included: a second A_{2A} receptor structure (PDB ID: 4EIY²⁵), associated with a different ligand series from the work of Lenselink et al.²⁶; the metabotropic Glutamate (mGlu₅) receptor (PDB ID: 5CGC²⁰), based on the study by Guo et al.¹⁸; and the P2Y₁ receptor (PDB ID: 4XNV²⁷), as reported by Dickson et al.¹⁹

Every receptor structure within this work has been prepared through the Molecular Operating Environment (MOE²⁸) suite: initially, co-crystallization factors were removed, except for the crystallographic ligands, water molecules and the conserved sodium ion. Subsequently, all mutated residues were reverted to the wildtype sequence based on the reference FASTA retrieved from the UniProt²⁹ database, and an energy minimization step was performed on these residues, while restraining the protein backbone, using the Protein Builder tool. After that, chimeric protein regions were deleted, and loops were rebuilt using the

Loop Modeler tool. Following this, alternate conformations were resolved using the Structure Preparation tool by selecting those with the highest occupancy, and the N- and C-termini were capped with ACE and NME groups, respectively. Finally, the structures were protonated using the Protonate 3D tool, with a pH of 7.4 and a temperature of 310 K, and missing hydrogen atoms were added. A final energy minimization, using the Amber14 EHT Forcefield, was then performed, restraining all atoms except hydrogen. The final step involved aligning the coordinates of the prepared structure with those of the corresponding entry from the OPM³⁰ database, as the Flare software requires pre-aligned structures to correctly model the membrane environment.

Based on the protocol used for performing simulations, different contents of crystallographic water molecules were kept in the receptor structure. The term Binding Site Water molecules (BS Water) refers to water molecules retained from the crystallographic structure that are within 4.5 Å of the ligand. The term Sodium Water molecules (Na⁺ Water) refers to the three crystallographic water molecules that constitute the first solvation shell of the sodium ion located in the allosteric site³¹. Further details regarding water molecules in the system are treated in the Flare FEP Project Preparation section.

Ligand Preparation

All ligand datasets used in this work were retrieved from the corresponding target paper^{17–19,26}, and prepared using the automatic import tool within Flare. 3D structures were generated by using one of the method listed below:

In-Place Drawing: the crystallographic ligand from the corresponding receptor structure served as a template. Modifications of the molecule were then made directly within the Flare Ligand Editor by substituting atoms in place, and fixing bond lengths and angles with the dedicated tool. For triazinic ligands containing non-symmetrical ring systems (e.g., 4j, 4k, 4l, 4m, 4r), both possible orientations were modeled to account for ambiguity in substitution geometry. This expanded the dataset from 10 to 15 ligands and, thus, changed the size and connectivity of the alchemical network.

PLANTS Docking: To produce docked binding poses for ligands, the binding site was defined using the center of mass of the crystallographic ligand from the corresponding receptor structure, with a spherical search space of 15 Å radius, ensuring comprehensive coverage of the orthosteric site. For each ligand, 20 poses were generated using the PLANTS³² Docking software along with the ChemPLP³³ scoring function, and the top-ranked pose was selected for further free energy calculations.

Conf Hunt and Align: The Conf Hunt and Align tool in the Flare GUI was used to generate low-energy ligand conformations and align them to a reference structure based on a shared substructure. The Conformation Hunt accuracy was set to "Very accurate and slow", and the alignment method to "Substructure". As excluded volume, only the protein chain was selected, allowing crystallographic water molecules to be potentially occupied by ligand atoms if favorable. Other settings were kept at their default value. Ligand 4q, drawn in-place using as template the crystallographic 6GT3 ligand, and representative of the shared triazine core, was used as the alignment reference for triazines.

For the additional test cases, the corresponding crystallographic ligand was selected as the alignment reference, except in the A_{2A} additional ligand set, where the pre-aligned ligand 25d was used as reference.

Flare FEP Project Preparation

Each FEP project was initialized within the Flare interface by selecting the pre-generated ligand conformations, the primary protein chain, and any additional chains required by the specific protocol. Upon project creation, experimental activity values were imported and associated with the ligands.

Alchemical networks were then constructed in benchmark mode, which fit the calculated values of relative binding free energies on the entire dataset, without anchoring results to a single reference compound. Network generation was performed using Flare's automated workflow, with the "maximum cycle size" parameter set to 5. The option "pay attention to 3D coordinates" was enabled to obtain best ligand pairs formation based on their spatial alignment.

For the triazine ligand series, three different methods were used to generate input conformations, which have been thoroughly described in the Ligand Preparation section. Accordingly, two separate networks were built to reflect differences in the number of available ligands: a 15-node network for the drawn ligand set, which included both orientations of asymmetric substituents, and a 10-node network used for both

the docked and aligned ligand sets. These networks remained unchanged across all protocols to ensure consistency in free energy comparisons.

For the additional systems (new A_{2A}, mGlu₅, and P2Y₁), only one input conformation method (alignment) was applied. Thus, a single alchemical network was constructed per system, with network size corresponding to the number of ligands available in each dataset.

Flare FEP Simulations

The final step prior to running the simulations was the selection of simulation settings. The complete simulation parameters, along with the corresponding protein chains, are summarized in Table 1.

While specific protocol configurations varied across different setups, all other simulation settings were kept at their default values.

For every protocol, three independent project replicas were carried out, and subsequently analyzed by using both metrics from the Flare FEP GUI interface, and scripts relying on the MDAnalysis^{34,35} and Matplotlib³⁶ python packages.

Predicting binding free energy with accuracy remains a central challenge in computational drug discovery. This is particularly true for ligands interacting with membrane-associated proteins, where the local environment, water networks, and conformational flexibility introduce significant complexity. In the present study, we focused on a congeneric series of triazine-based ligands, a scaffold of pharmaceutical relevance, with available experimental binding affinities. This series has previously been employed by Deflorian and colleagues in their RBEF benchmark study¹⁷ (A_{2A} study 6), providing a well-characterized dataset for systematic evaluation of computational protocols.

To explore how methodological choices affect prediction quality, we systematically evaluated ten protocols differing in ligand placement strategy (manual draw-in-place, docking, or alignment), forcefield choice (GAFF2 vs. OpenFF2), membrane inclusion, and treatment of water molecules, as summarized in Table 1.

Results and Discussion

Table 1. Simulation settings for the triazine section of the benchmark.

ID	Ligand Placement	Network Size	FF	Membrane	GCNCMC	BS Water	Na+ Water
1	Drawing	15	GAFF2	Yes	No	Yes	Yes
2	Drawing	15	GAFF2	Yes	No	No	Yes
3	Drawing	15	GAFF2	Yes	Yes	No	No
4	Drawing	15	GAFF2	Yes	Yes	No	Yes
5	Drawing	15	GAFF2	No	No	Yes	Yes
6	Drawing	15	OpenFF2	Yes	No	Yes	Yes
7	Docking	10	GAFF2	Yes	No	Yes	Yes
8	Docking	10	OpenFF2	Yes	No	Yes	Yes
9	Alignment	10	GAFF2	Yes	No	Yes	Yes
10	Alignment	10	OpenFF2	Yes	No	Yes	Yes

For every protocol (ID), the following settings are reported: Ligand Placement, Network Size (the number of nodes in the network), FF (ligand Forcefield used), Membrane (presence of the double phospholipid layer in the system), GCNCMC (application of the GCNCMC method for enhanced water equilibration), BS Water, and Na⁺ Water (indicating whether binding site water molecules and sodium-associated water molecules, respectively, are included in the system).

Performances were assessed using four complementary metrics. R^2 quantifies how well the computational model captures the overall trend. Kendall's τ evaluates the agreement between predicted and experimental rankings of ligand affinities. RMSE (Root Mean Square Error) measures the typical magnitude of prediction errors, giving insight into overall accuracy while penalizing larger deviations. MAE (Mean Absolute Error) calculates the average absolute difference between predicted and experimental values, reflecting systematic deviations without emphasizing outliers.

We began our analysis with the default setup (Protocol 1), which combined manual draw-in-place ligand placement, the GAFF2 forcefield, inclusion of the membrane, crystallographic binding site waters (BS Water), and sodium-associated water molecules (Na⁺ Water). For some ligands, placed by in-place drawing, two alternative bound conformations were possible due to the chemical asymmetry of certain triazine substitutions (Figure 1).

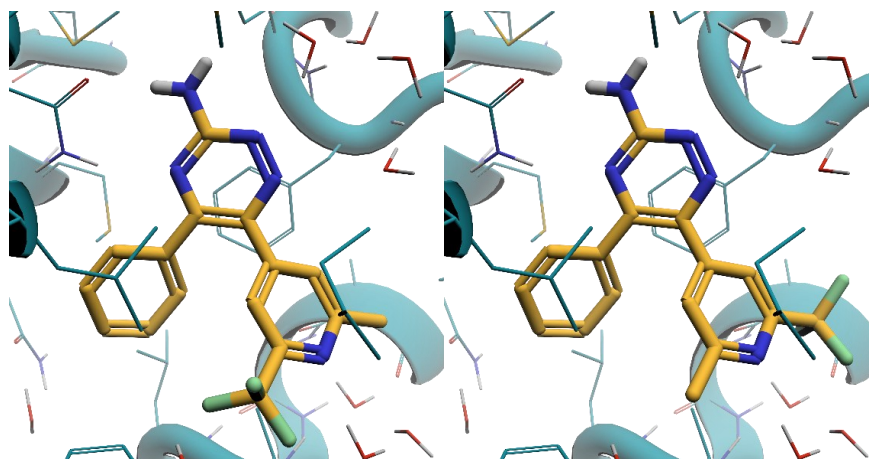


Figure 1. Depiction of two possible binding poses for ligand 4k.

In these cases, one conformation consistently produced better agreement with experimental affinities, and its orientation closely matched the poses generated by docking and alignment in the subsequent protocols. Initially, our analysis was performed by averaging the ΔG values obtained from both conformations of

each ligand, and then calculating the overall mean across replicas (“Cumulative” Predicted ΔG). However, it is plausible that in the biological system only one binding mode is preferentially stabilized, for example, due to local interactions with the protein pocket, water network rearrangements, or energetic contributions that favor one orientation over the other. For this reason, within Protocol 1 results, we repeated the analysis considering only the mean values of the conformations yielding the lowest (i.e., most favorable) predicted ΔG , hereafter referred to as the “Best” predicted ΔG . As can be seen in Figure 2, the “Best” approach consistently achieved higher correlation with experimental measurements compared to the “Cumulative” analysis, even though the latter showed slightly lower absolute errors due to averaging predictions. This indicates that while the “Cumulative” analysis smooths out deviations, it does so at the expense of correctly reproducing affinity trends. In contrast, the “Best” analysis better captures the dominant binding mode, offering a more accurate description of ligand binding in this series. For this reason, all draw-in-place results reported in this study are based on the “Best” predicted ΔG , which we consider the most reliable representation of binding energetics.

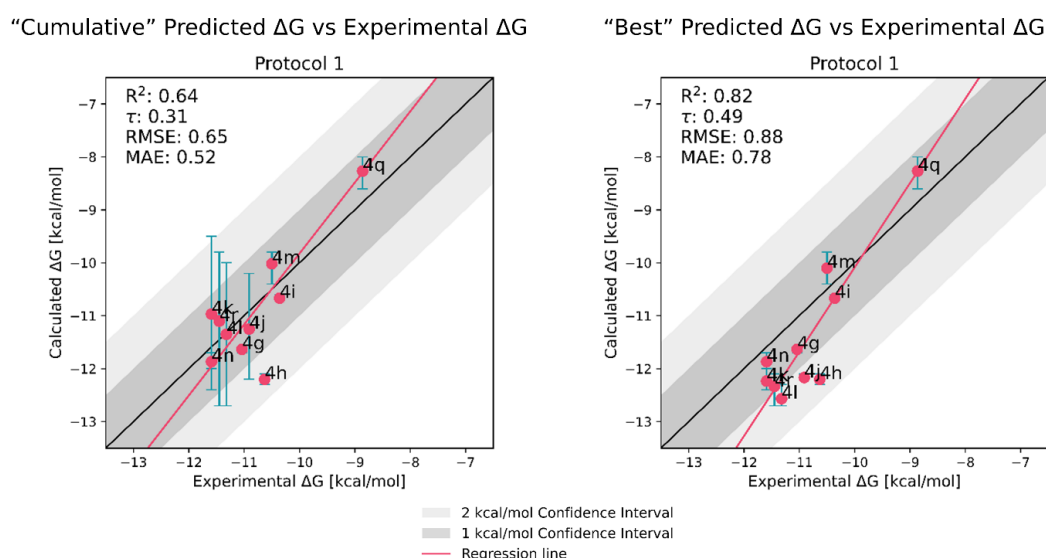


Figure 2. Comparison of predicted versus experimental binding affinities for the triazine ligand series. Left: results obtained using the “Cumulative” predicted ΔG approach (averaging all conformations and replicas values for each ligand). Right: results based on the “Best” predicted ΔG (per replica mean value considering only the conformation with the most favorable ΔG for each ligand). Each point represents a ligand, and error bars indicate the range between minimum and maximum predicted ΔG values across all replicas. Key performance metrics are reported for each plot, including R^2 , Kendall’s τ , RMSE, and MAE.

The complete set of results for the triazine benchmark is presented in Figure 3, and the detailed protocol-by-protocol discussion follows in the next sections.

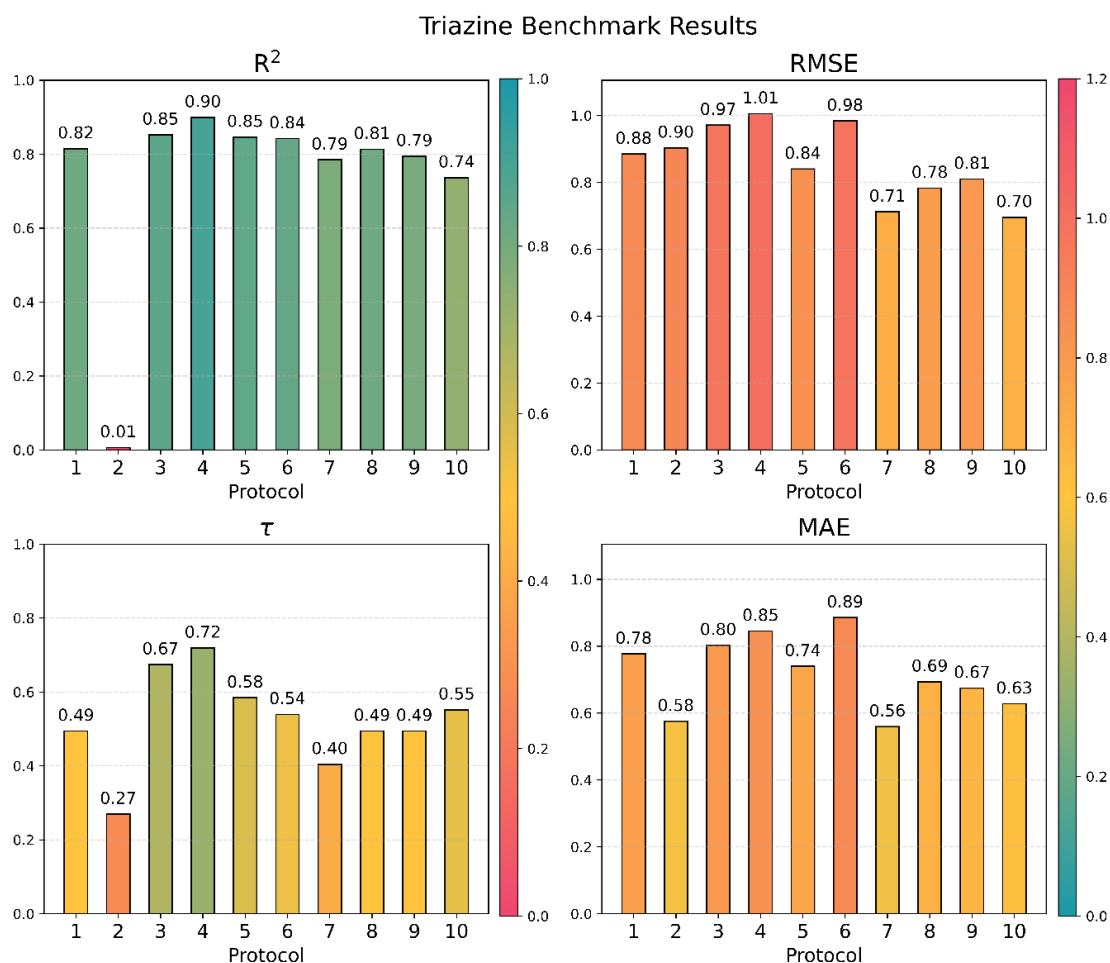


Figure 3. Comparison of statistical performance across protocols. Each panel shows one performance metric (R^2 , Kendall's τ , RMSE, and MAE). The x-axis reports the corresponding protocol numbers, while the y-axis indicates the metric value. Bars are color-coded according to performance: green/blue shades indicate better predictive quality, whereas orange shades highlight poorer performance.

Assessing the Role of Water Molecules

To evaluate the impact of hydration, we first removed all Binding Site (BS) water molecules from the system (Protocol 2). Predictive power collapsed, with R^2 dropping to 0.01 and τ to 0.27, despite RMSE (0.89) and MAE (0.57) remaining superficially similar.

We then applied GCNMC equilibration (Protocol 3), allowing water molecules to dynamically sample positions near the ligand. This markedly improved results compared to Protocol 1, with $R^2 = 0.85$, $\tau = 0.67$, RMSE = 0.97, and MAE = 0.81. These values highlight the importance of including water sampling to capture structural hydration effects. However, GCNMC sampling is restricted to a region around the ligand (4.5 Å), which prevents equilibration of waters coordinating the sodium ion.

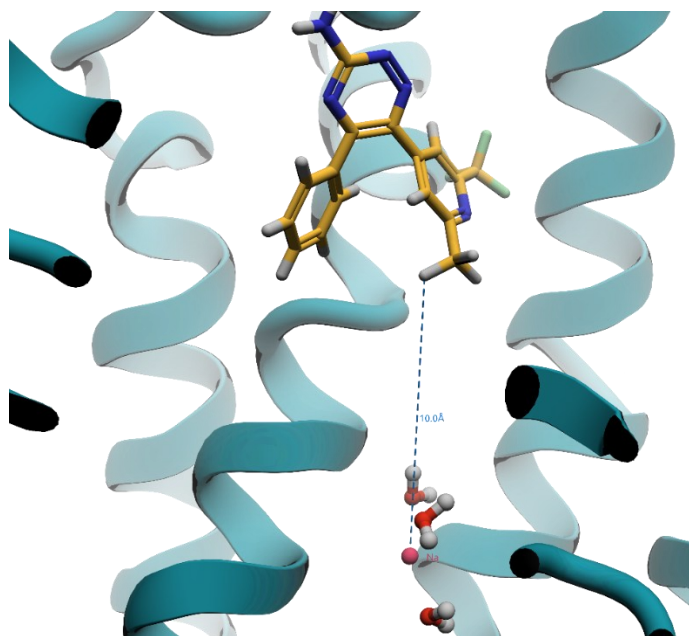


Figure 4. Annotated distance (10.0 Å) of the sodium ion (depicted in pink) from the triazine ligand (4k, orange) closest atom within the A_{2A} receptor binding site. The Na⁺ Waters are also shown.

To address this limitation, we retained the crystallographic sodium waters alongside GCNMC equilibration (Protocol 4). This combined approach delivered the best overall performance of the series, with $R^2 = 0.90$, $\tau = 0.72$, RMSE = 1.01, and MAE = 0.86. These results confirm that combining dynamic equilibration with explicit retention of key structural waters provides the most realistic representation of the binding environment.

To better understand the poor performance of Protocol 2, we carried out an RMSD analysis and compared it with the best-performing setup, Protocol 4.

As shown in Figure 5, Protocol 2 displays significantly greater fluctuations in both the ligand and the binding site residues compared to Protocol 4. This increased mobility suggests reduced structural stability under Protocol 2 conditions.

Figure 6 illustrates a representative snapshot of the A_{2A} binding site when a triazine antagonist is bound. In this configuration, a buried water pocket is typically formed, where a single water molecule bridges the pyridine nitrogen of the ligand and the side chain of His278. Under Protocol 2 conditions, however, this water molecule is absent, resulting in a vacuum-like cavity within the binding site. In several simulations, particularly those with elevated RMSD values, this void appears to draw in nearby sodium-coordinated water molecules, further disrupting the local environment. Another plausible explanation is that, due to the inaccessibility of the inner pocket once the ligand is bound, the absence of water creates an artificial cavity that destabilizes both the ligand and surrounding residues. Together, these observations point to structural instabilities in Protocol 2, which likely contribute to its poorer predictive performance.

In summary, in protocols where all key water molecules are correctly placed, the system remains stable and predictive performance improves significantly. In fact, under these conditions, the results across different protocols become more consistent and comparable, underscoring the critical role of buried water molecules in stabilizing both the ligand and the binding site architecture.

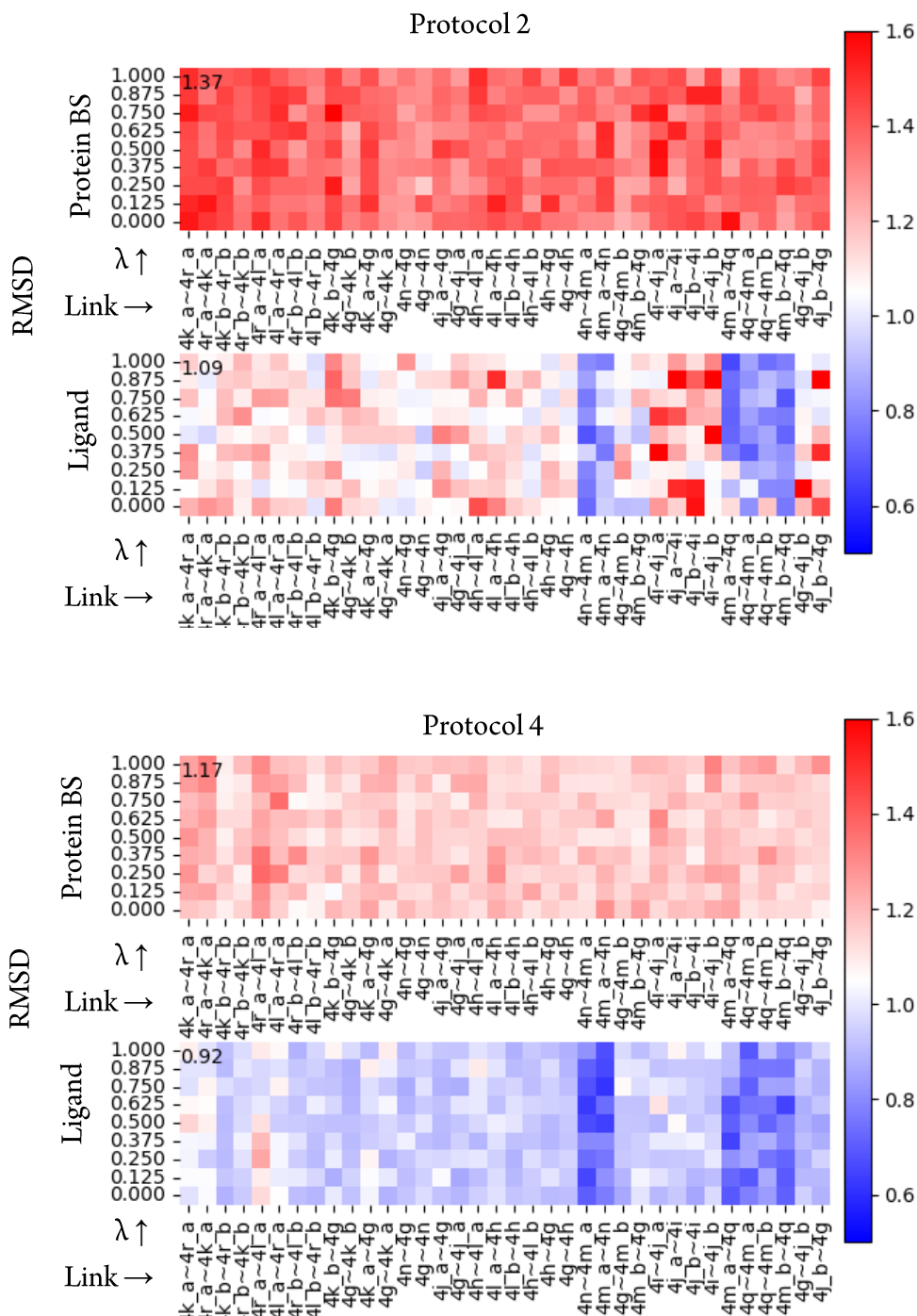


Figure 5. RMSD analysis comparing Protocol 2 (top panels) and Protocol 4 (bottom panels). For each protocol, the mean RMSD across three replicas is shown for both the protein binding-site residues (Protein BS) and the ligand, reported separately by λ -window and by network link. The overall cumulative mean RMSD value is indicated in each quadrant of the graphs.

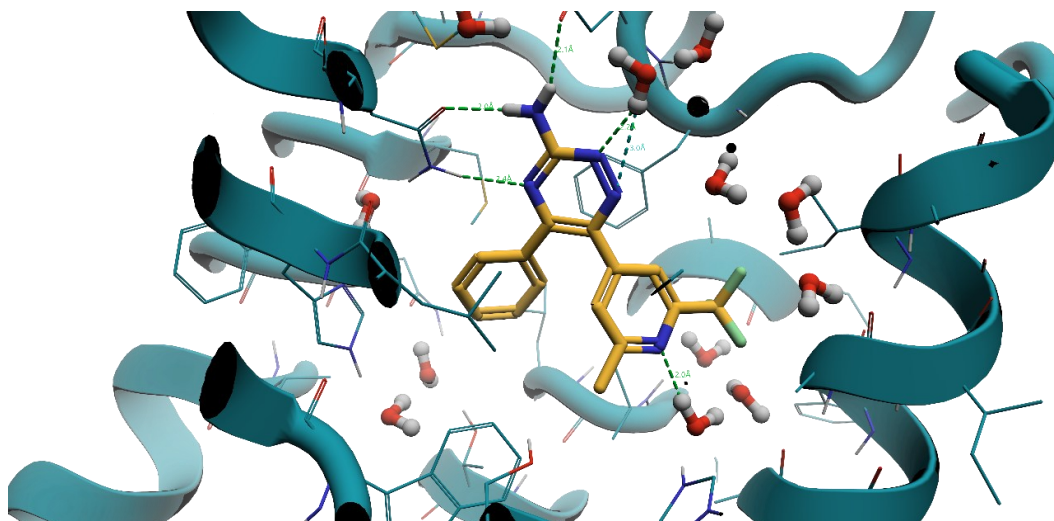


Figure 6. Depiction of the A_{2A} orthosteric binding site bound to ligand 4k. Water molecules and H-bonds are also shown.

Membrane Effects

We next examined the effect of the lipid bilayer by comparing the baseline setup (Protocol 1, $R^2 = 0.82$, $\tau = 0.49$, RMSE = 0.88, and MAE = 0.78) with a no-membrane configuration (Protocol 5). Removing the membrane yielded comparable overall predictive quality ($R^2 = 0.84$, RMSE = 0.85, MAE = 0.75). Interestingly, Protocol 5 also trended toward a slightly higher τ value (0.58 vs. 0.49), but, overall, the two τ values are effectively similar. These findings indicate that while the lipid bilayer contributes to reproducing the physiological environment of the membrane-associated protein, absolute error values are similar between the two setups. The effect of removing the membrane, which would enable the simpler and less computational expensive system setup, has not been systematically investigated across other targets, but, in specific cases, such as when the binding pocket lies at the protein-membrane interface, the bilayer forms an integral part of the ligand environment. In such cases, its removal would constitute a substantial perturbation of the binding site.

In this specific case, despite broadly comparable predictive performance, the guideline from this analysis is to continue using the membrane-inclusive, physiologically realistic system.

Ligand Placement and Forcefield Effects

We first compared the impact of ligand placement method, analyzing manual draw-in-place (Protocol 1), docking-based (Protocol 7), and alignment-based (Protocol 9) setups. Draw-in-place produced the strongest performance, with $R^2 = 0.82$, $\tau = 0.49$, RMSE = 0.88, and MAE = 0.78. Docking-based placement yielded slightly weaker results (Protocol 7: $R^2 = 0.79$, $\tau = 0.40$, RMSE = 0.71, MAE = 0.56), while alignment-based placement produced intermediate performance (Protocol 9: $R^2 = 0.79$, $\tau = 0.49$, RMSE = 0.81, MAE = 0.67). These differences likely reflect not only the intrinsic accuracy of the placement methods but also the changes in network size and topology: draw-in-place generated 15 nodes due to multiple conformations, whereas docking and alignment involved only 10 nodes. Moreover, simulation time is reduced in the latter protocols, which may partially explain their modest decrease in predictive performance. A direct comparison of the three placement strategies is provided in Figure 7.

We next examined the effect of ligand forcefield choice by comparing GAFF2 and OpenFF2 within each placement strategy. For draw-in-place, Protocol 1 (GAFF2) and Protocol 6 (OpenFF2) gave very similar outcomes ($R^2 = 0.82$ vs. 0.84; $\tau = 0.49$ vs. 0.54; RMSE = 0.88 vs. 0.98; MAE = 0.78 vs. 0.89). For docking, Protocol 7 (GAFF2) achieved $R^2 = 0.79$ and $\tau = 0.40$, while Protocol 8 (OpenFF2) slightly improved to $R^2 = 0.81$ and $\tau = 0.49$, although error values increased (RMSE = 0.71 vs. 0.78; MAE = 0.56 vs. 0.69). For alignment, Protocol 9 (GAFF2) yielded $R^2 = 0.79$, $\tau = 0.49$, RMSE = 0.81, and MAE = 0.67, while Protocol 10 (OpenFF2) showed somewhat weaker correlation but improved ranking ($R^2 = 0.74$, $\tau = 0.55$, RMSE = 0.70, MAE = 0.63). Overall, these comparisons indicate that forcefield choice

has only a minor influence on predictive quality within this series, with performance differences generally smaller than those introduced by ligand placement method.

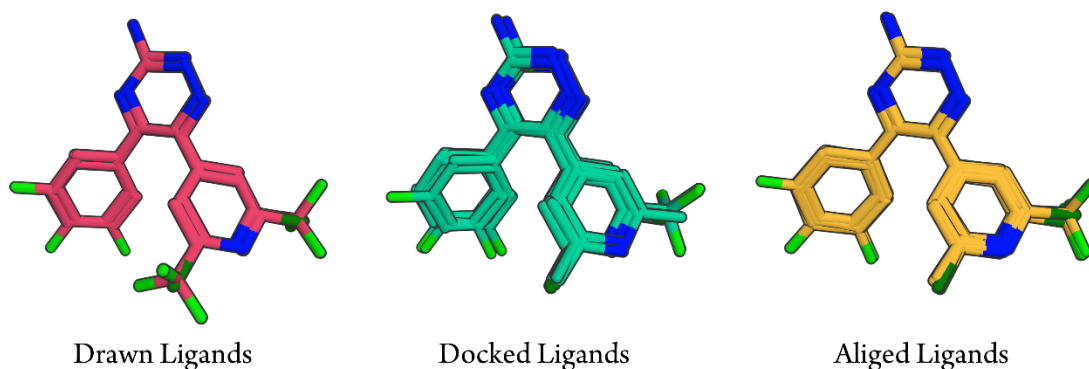


Figure 7. Depiction of results coming from different ligand placement methods.

It is also important to emphasize that the triazine ligands investigated here are relatively rigid molecules. As such, this single series is not sufficient to fully explore the potential differences between GAFF2 and OpenFF2.

To address this limitation and provide a broader assessment, we extended our benchmark to include three additional targets with chemically distinct ligand series: the adenosine A_{2A} receptor (new A_{2A}), investigated here with a ligand set different from the triazine series, the metabotropic glutamate receptor (mGlu), and the purinergic receptor P2Y1. These expanded datasets allow us to evaluate forcefield performance across a wider range of scaffolds and binding environments, offering a more generalizable perspective on predictive reliability.

All additional simulations were performed using the best performing protocol from the triazine part of the benchmark (protocol 4), and ligand input conformations were generated by alignment to a reference structure. This approach was chosen because generating accurate drawn poses for these ligands was not feasible, and previous results had shown that aligned poses consistently outperformed docked poses in terms of predictive accuracy.

In the new A_{2A} test case, both forcefields achieved only moderate predictive quality, with GAFF2 yielding $R^2 = 0.46$, $\tau = 0.39$, RMSE = 1.34, and MAE = 1.11, while OpenFF2 produced nearly identical performance ($R^2 = 0.46$, $\tau = 0.39$, RMSE = 1.33, MAE = 1.08). These results confirm that for this dataset, forcefield choice has negligible influence. For this dataset, performances were weaker than in the triazine benchmark, possibly because the chemical substitutions in this ligand series are mainly located in the outer portion of the binding site, where they are more solvent-exposed. In such regions, binding free energies are strongly influenced by local hydration patterns and the balance between ligand desolvation and protein-water interactions.

For the mGlu₅ receptor, results were somewhat stronger, with GAFF2 producing $R^2 = 0.60$, $\tau = 0.67$, RMSE = 1.51, and MAE = 1.11, while OpenFF2 improved slightly to $R^2 = 0.64$, $\tau = 0.67$, RMSE = 1.39, and MAE = 1.01. Both forcefields captured affinity trends well, particularly for strong binders, though predictions for weaker ligands remained less reliable. Here, OpenFF2 showed a modest but consistent reduction in error relative to GAFF2, suggesting a potential advantage for this more modern parametrization in certain chemical contexts.

The P2Y1 receptor benchmark further supported the notion that differences between the two forcefields are subtle and context-dependent. GAFF2 achieved $R^2 = 0.63$, $\tau = 0.62$, RMSE = 1.37, and MAE = 1.13, while OpenFF2 yielded $R^2 = 0.61$, $\tau = 0.60$, RMSE = 1.28, and MAE = 1.06. As in the mGlu₅ case, OpenFF2 slightly reduced absolute errors, while GAFF2 achieved marginally higher correlation values. Both forcefields reproduced the main binding affinity trends, though deviations remained more pronounced than in the triazine dataset.

All discussed data can be visually inspected in Figure 8.

Crucially, predictive quality appears to be driven more by the chemical diversity of the ligand series, the

complexity of the binding environment, and the quality of the input poses than by the specific forcefield employed. Across all test cases, ligand pose generation exerted a greater influence on predictive performance than forcefield choice. When high-quality conformations were available, particularly alignment-based placements, both GAFF2 and OpenFF2 yielded accurate and consistent RBFE predictions. No systematic advantage was observed for either forcefield; their relative performance varied with the system and ligand set. These findings support the viability of both GAFF2 and OpenFF2 in FEP workflows, provided that ligand poses are carefully prepared.

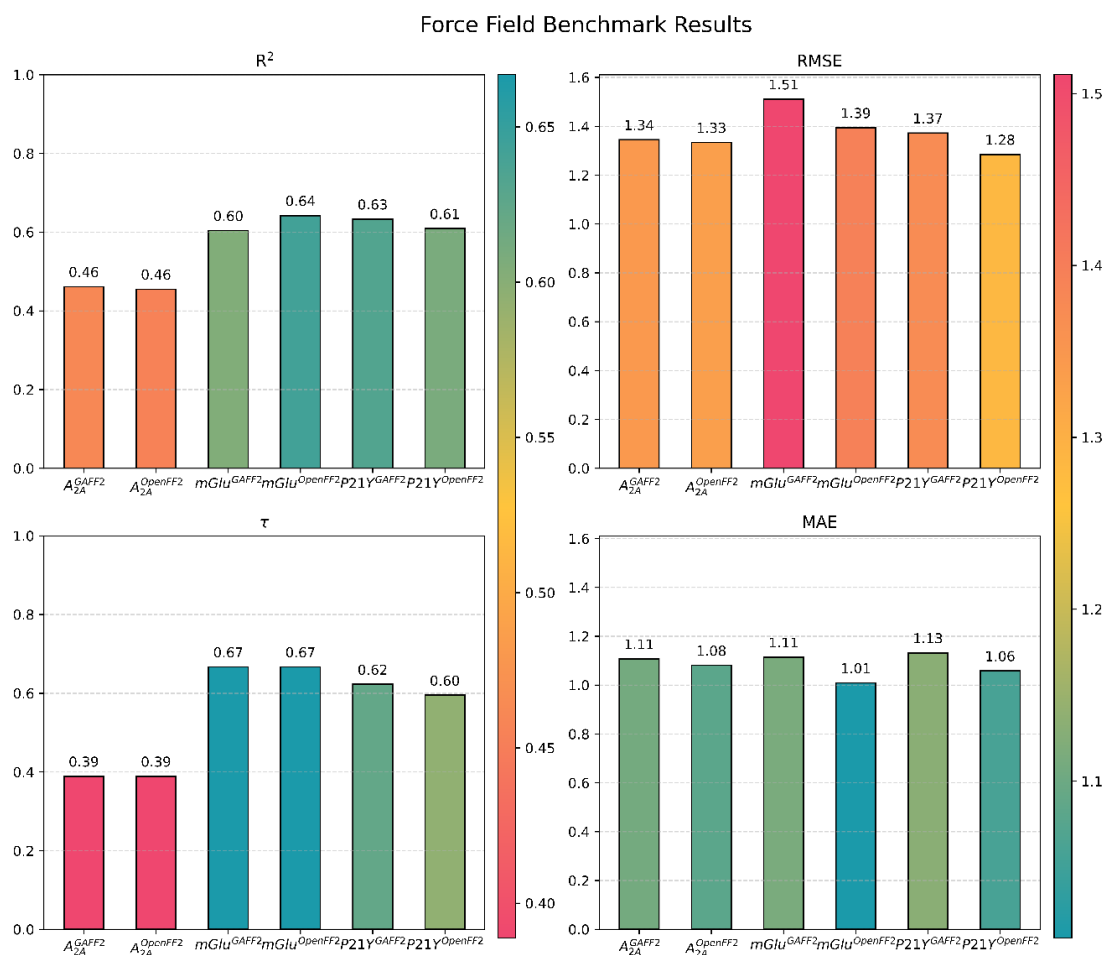


Figure 8. Force field comparison across three GPCR test cases using four statistical performance metrics. Each force field (GAFF2 and OpenFF) is evaluated across four independent replicas for each target, with the x-axis grouping results by system. The thin, black-edged bar in each group represents the mean value across replicas. The four panels report R^2 (coefficient of determination), MUE (mean unsigned error), τ (Kendall's rank correlation coefficient), and Linkwise RMSD, used to assess the accuracy and stability of free energy predictions. Test cases include: new A_{2A} , $mGlu_5$ and $P2Y1$.

Conclusions

Our benchmark highlights that Flare FEP accuracy for GPCRs is influenced more by ligand placement, and water treatment than by the choice of force field. Alignment-based strategies provided the most reliable results, while inclusion of the membrane improved realism and ranking in most cases. Both GAFF2 and OpenFF2 performed comparably across all test cases, with no systematic advantage. Overall, high-quality input poses and careful system preparation remain the primary determinants of transferable and physically meaningful Flare RBFE predictions.

Affiliations

1 Molecular Modeling Section (MMS), Department of Pharmaceutical and Pharmacological Sciences, University of Padua, Padova, Italy

2 Aptuit, an Evotec company, Via Alessandro Fleming, 4, 37135 Verona VR, Italy

References

1. Anderson, A. C. The Process of Structure-Based Drug Design. *Chem. Biol.* 2003, 10 (9), 787–797. <https://doi.org/10.1016/j.chembiol.2003.09.002>.
2. Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. Kinetics of Protein–Ligand Unbinding: Predicting Pathways, Rates, and Rate-Limiting Steps. *Proc. Natl. Acad. Sci.* 2015, 112 (5), E386–E391. <https://doi.org/10.1073/pnas.1424461112>.
3. Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* 2017, 57 (12), 2911–2937. <https://doi.org/10.1021/acs.jcim.7b00564>.
4. Jarzynski, C. Rare Events and the Convergence of Exponentially Averaged Work Values. *Phys. Rev. E* 2006, 73 (4). <https://doi.org/10.1103/PhysRevE.73.046105>.
5. Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* 2015, 137 (7), 2695–2703. <https://doi.org/10.1021/ja512751q>.
6. Song, L. F.; Lee, T.-S.; Zhu, C.; York, D. M.; Merz, K. M. Jr. Using AMBER18 for Relative Free Energy Calculations. *J. Chem. Inf. Model.* 2019, 59 (7), 3128–3135. <https://doi.org/10.1021/acs.jcim.9b00105>.
7. Kuhn, M.; Firth-Clark, S.; Tosco, P.; Mey, A. S. J. S.; MacKey, M.; Michel, J. Assessment of Binding Affinity via Alchemical Free-Energy Calculations. *J. Chem. Inf. Model.* 2020, 60 (6), 3120–3130. https://doi.org/10.1021/ACS.JCIM.0C00165/SUPPL_FILE/C10C00165_SI_002.ZIP.
8. Margiotta, E.; Deganutti, G.; Moro, S. Could the Presence of Sodium Ion Influence the Accuracy and Precision of the Ligand-Posing in the Human A2A Adenosine Receptor Orthosteric Binding Site Using a Molecular Docking Approach? Insights from Dockbench. *J. Comput. Aided Mol. Des.* 2018, 32 (12), 1337–1346. <https://doi.org/10.1007/s10822-018-0174-2>.
9. Salari, R.; Joseph, T.; Lohia, R.; Hénin, J.; Brannigan, G. A Streamlined, General Approach for Computing Ligand Binding Free Energies and Its Application to GPCR-Bound Cholesterol. *J. Chem. Theory Comput.* 2018, 14 (12), 6560–6573. <https://doi.org/10.1021/acs.jctc.8b00447>.
10. Hedges, L. O.; Mey, A. S. J. S.; Laughton, C. A.; Gervasio, F. L.; Mulholland, A. J.; Woods, C. J.; Michel, J. BioSimSpace: An Interoperable Python Framework for Biomolecular Simulation. *J. Open Source Softw.* 2019, 4 (43), 1831. <https://doi.org/10.21105/joss.01831>.
11. Woods, C. J.; Hedges, L. O.; Mulholland, A. J.; Malaisree, M.; Tosco, P.; Loeffler, H. H.; Suruzhon, M.; Burman, M.; Bariami, S.; Bosisio, S.; Calabro, G.; Clark, F.; Mey, A. S. J. S.; Michel, J. Sire: An Interoperability Engine for Prototyping Algorithms and Exchanging Information between Molecular Simulation Programs. *J. Chem. Phys.* 2024, 160 (20), 202503. <https://doi.org/10.1063/5.0200458>.
12. Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLOS Comput. Biol.* 2017, 13 (7), e1005659. <https://doi.org/10.1371/journal.pcbi.1005659>.
13. Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V. W. D.; Forouzes, N.; Giese, T. J.; Götz, A. W.; Gohlke, H.; Izadi, S.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kurtzman, T.; Lee, T.-S.; Li, P.; Liu, J.; Luchko, T.; Luo, R.; Manathunga, M.; Machado, M. R.; Nguyen, H. M.; O’Hearn, K. A.; Onufriev, A. V.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Risheh, A.; Schott-Verdugo, S.; Shajan, A.; Swails, J.; Wang, J.; Wei, H.; Wu, X.; Wu, Y.; Zhang, S.; Zhao, S.; Zhu, Q.; Cheatham, T. E. I.; Roe, D. R.; Roitberg, A.; Simmerling, C.; York, D. M.; Nagan, M. C.; Merz, K. M. Jr. AmberTools. *J. Chem. Inf. Model.* 2023, 63 (20), 6183–6191. <https://doi.org/10.1021/acs.jcim.3c01153>.
14. Promising results for three GPCR benchmarks using Flare™ FEP for accurate binding affinity calculations in membrane proteins | Cresset. <https://cresset-group.com/about/news/promising-results-three-gpcr-benchmarks/> (accessed 2025-09-29).
15. Allard, B.; Beavis, P. A.; Darcy, P. K.; Stagg, J. Immunosuppressive Activities of Adenosine in Cancer. *Curr. Opin. Pharmacol.* 2016, 29, 7–16. <https://doi.org/10.1016/j.coph.2016.04.001>.
16. Pinna, A. Adenosine A2A Receptor Antagonists in Parkinson’s Disease: Progress in Clinical Trials from the Newly Approved Istradefylline to Drugs in Early Development and Those Already Discontinued. *CNS Drugs* 2014, 28 (5), 455–474. <https://doi.org/10.1007/s40263-014-0161-7>.

17. Deflorian, F.; Perez-Benito, L.; Lenselink, E. B.; Congreve, M.; van Vlijmen, H. W. T.; Mason, J. S.; Graaf, C. de; Tresadern, G. Accurate Prediction of GPCR Ligand Binding Affinity with Free Energy Perturbation. *J. Chem. Inf. Model.* 2020, *60* (11), 5563–5579. <https://doi.org/10.1021/acs.jcim.0c00449>.
18. Guo, Y.; Zhou, Y.; Bai, Q.; Bo, Z.; Song, K.; Chang, J.; Zhang, Y.; Yang, M.; Deng, Y.; Wang, D. A Benchmark for Accurate GPCR Ligand Binding Affinity Prediction with Free Energy Perturbation. *ChemRxiv* September 20, 2023. <https://doi.org/10.26434/chemrxiv-2023-mhp2x>.
19. Dickson, C. J.; Hornak, V.; Duca, J. S. Relative Binding Free-Energy Calculations at Lipid-Exposed Sites: Deciphering Hot Spots. *J. Chem. Inf. Model.* 2021, *61* (12), 5923–5930. <https://doi.org/10.1021/acs.jcim.1c01147>.
20. Christopher, J. A.; Aves, S. J.; Bennett, K. A.; Doré, A. S.; Errey, J. C.; Jazayeri, A.; Marshall, F. H.; Okrasa, K.; Serrano-Vega, M. J.; Tehan, B. G.; Wiggin, G. R.; Congreve, M. Fragment and Structure-Based Drug Discovery for a Class C GPCR: Discovery of the mGlu5 Negative Allosteric Modulator HTL14242 (3-Chloro-5-[6-(5-Fluoropyridin-2-yl)Pyrimidin-4-yl]Benzonitrile). *J. Med. Chem.* 2015, *58* (16), 6653–6664. <https://doi.org/10.1021/acs.jmedchem.5b00892>.
21. von Kügelgen, I. Pharmacological Profiles of Cloned Mammalian P2Y-Receptor Subtypes. *Pharmacol. Ther.* 2006, *110* (3), 415–432. <https://doi.org/10.1016/j.pharmthera.2005.08.014>.
22. Chao, H.; Turdi, H.; Herpin, T. F.; Roberge, J. Y.; Liu, Y.; Schnur, D. M.; Poss, M. A.; Rehfsuss, R.; Hua, J.; Wu, Q.; Price, L. A.; Abell, L. M.; Schumacher, W. A.; Bostwick, J. S.; Steinbacher, T. E.; Stewart, A. B.; Ogletree, M. L.; Huang, C. S.; Chang, M.; Cacace, A. M.; Arcuri, M. J.; Celani, D.; Wexler, R. R.; Lawrence, R. M. Discovery of 2-(Phenoxypyridine)-3-Phenylureas as Small Molecule P2Y1 Antagonists. *J. Med. Chem.* 2013, *56* (4), 1704–1714. <https://doi.org/10.1021/jm301708u>.
23. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* 2000, *28* (1), 235–242. <https://doi.org/10.1093/NAR/28.1.235>.
24. Borodovsky, A.; Barbon, C. M.; Wang, Y.; Ye, M.; Prickett, L.; Chandra, D.; Shaw, J.; Deng, N.; Sachsenmeier, K.; Clarke, J. D.; Linghu, B.; Brown, G. A.; Brown, J.; Congreve, M.; Cheng, R. K.; Dore, A. S.; Hurrell, E.; Shao, W.; Woessner, R.; Reimer, C.; Drew, L.; Fawell, S.; Schuller, A. G.; Mele, D. A. Small Molecule AZD4635 Inhibitor of A2AR Signaling Rescues Immune Cell Function Including CD103+ Dendritic Cells Enhancing Anti-Tumor Immunity. *J. Immunother. Cancer* 2020, *8* (2), e000417. <https://doi.org/10.1136/jitc-2019-000417>.
25. Liu, W.; Chun, E.; Thompson, A. A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G. W.; Roth, C. B.; Heitman, L. H.; IJzerman, A. P.; Cherezov, V.; Stevens, R. C. Structural Basis for Allosteric Regulation of GPCRs by Sodium Ions. *Science* 2012, *337* (6091), 232–236. <https://doi.org/10.1126/science.1219218>.
26. Lenselink, E. B.; Louvel, J.; Forti, A. F.; van Veldhoven, J. P. D.; de Vries, H.; Mulder-Krieger, T.; McRobb, F. M.; Negri, A.; Goose, J.; Abel, R.; van Vlijmen, H. W. T.; Wang, L.; Harder, E.; Sherman, W.; IJzerman, A. P.; Beuming, T. Predicting Binding Affinities for GPCR Ligands Using Free-Energy Perturbation. *ACS Omega* 2016, *1* (2), 293–304. <https://doi.org/10.1021/acsomega.6b00086>.
27. Zhang, D.; Gao, Z.-G.; Zhang, K.; Kiselev, E.; Crane, S.; Wang, J.; Paoletta, S.; Yi, C.; Ma, L.; Zhang, W.; Han, G. W.; Liu, H.; Cherezov, V.; Katritch, V.; Jiang, H.; Stevens, R. C.; Jacobson, K. A.; Zhao, Q.; Wu, B. Two Disparate Ligand-Binding Sites in the Human P2Y1 Receptor. *Nature* 2015, *520* (7547), 317–321. <https://doi.org/10.1038/nature14287>.
28. Chemical Computing Group (CCG) | Computer-Aided Molecular Design. <https://www.chemcomp.com/en/index.htm> (accessed 2025-05-22).
29. The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2025. *Nucleic Acids Res.* 2025, *53* (D1), D609–D617. <https://doi.org/10.1093/nar/gkae1010>.
30. OPM. <https://opm.phar.umich.edu/proteins?search=4ey> (accessed 2024-02-14).
31. Gutiérrez-de-Terán, H.; Massink, A.; Rodríguez, D.; Liu, W.; Han, G. W.; Joseph, J. S.; Katritch, I.; Heitman, L. H.; Xia, L.; IJzerman, A. P.; Cherezov, V.; Katritch, V.; Stevens, R. C. The Role of a Sodium Ion Binding Site in the Allosteric Modulation of the A2A Adenosine G Protein-Coupled Receptor. *Structure* 2013, *21* (12), 2175–2185. <https://doi.org/10.1016/j.str.2013.09.020>.
32. Korb, O.; Stützle, T.; Exner, T. E. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design. *Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.* 2006, *4150 LNCS*, 247–258. https://doi.org/10.1007/11839088_22/COVER.
33. Korb, O.; Stützle, T.; Exner, T. E. Empirical Scoring Functions for Advanced Protein–Ligand Docking with PLANTS. *J. Chem. Inf. Model.* 2009, *49* (1), 84–96. <https://doi.org/10.1021/ci800298z>.
34. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* 2011, *32* (10), 2319–2327. <https://doi.org/10.1002/jcc.21787>.
35. Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domański, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python Sci. Conf.* 2016, 98–105. <https://doi.org/10.25080/Majora-629e541a-00e>.
36. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 2007, *9* (3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>.

Adenosine receptors in neuroinflammation and neurodegeneration

Veronica Salmaso^{1*}, Silvia Menin¹, Stefano Moro¹, Giampiero Spalluto² and Stephanie Federico^{2*}

Special Issue Adenosine and Purinergic Receptors: Regulation and Essential Role in Human Disease.
Cells
Submitted.

Abstract

Adenosine plays a crucial role in various pathophysiological conditions, including neuroinflammation and neurodegeneration. Neuroinflammation can be either beneficial or detrimental to the central nervous system, depending on the intensity and duration of the inflammatory response. Across a wide range of brain disorders, neuroinflammation contributes to both the onset and progression of disease. Notably, neuroinflammation is not limited to conditions primarily classified as neuroinflammatory but is also a key factor in other neurological disorders, including life-threatening neurodegenerative diseases. All four adenosine receptor subtypes (A_1 , A_{2A} , A_{2B} , and A_3) are implicated, to varying degrees, in these conditions. This review aims to summarize the roles of individual adenosine receptor subtypes in neuroinflammation and neurodegenerative diseases, emphasizing their therapeutic potential. While some therapeutic applications are well-established with clinically approved drugs, others warrant further investigation due to their promising potential.

Introduction

Adenosine and Adenosine Receptors

Adenosine is an endogenous purine nucleoside. Beyond its role in energy metabolism (it is the final product of the catabolism of the energetic molecule for definition, adenosine triphosphate (ATP)), adenosine modulates numerous processes in the human organism, through the recognition of its receptors (i.e. A_1 , A_{2A} , A_{2B} , A_3), and consequent cell signaling, and behaves as a neuromodulator in the central nervous system (CNS). It exerts a cytoprotective effect in response to stress, thus for instance limiting the damage provoked by excessive inflammatory response; for this reason, it has been defined as a retaliatory metabolite [1].

In physiological conditions, the extracellular concentration of adenosine falls in the 20-300 nM range, rising to the order of 10 μ M in case of metabolic stress and hypoxia, which are well-known to trigger inflammation [2,3]. This effect has been particularly studied in the CNS, where hypoxia can lead to 30 μ M extracellular adenosine concentration [3–5]. The extracellular adenosine is produced by the enzymatic catabolism of adenosine nucleotides by means of different ectonucleotidases [6], and it is also generated intracellularly and exported through nucleoside transporters or as a result of cell damage [7–9].

Adenosine exerts its effects through its four adenosine receptors (ARs) that are class A G protein-coupled receptors (GPCRs), structurally characterized by the typical 7-helices transmembrane bundle connected by three intracellular and three extracellular loops of various lengths, an N-terminal region facing the extracellular milieu and an intracellular C-terminal region, constituted by an amphipathic helix (helix 8) and a disordered tail. Currently, at least one experimental tridimensional structural has been released for each receptor subtype (some examples are reported in Figure 1), with A_{2A} AR being the first and most

explored one.

The four AR subtypes present high sequence identity, especially in the transmembrane helices (TM) region, with the couple A_{2A}-A_{2B} characterized by the highest transmembrane sequence identity (66%), followed by A₁-A₃ (57%), A₁-A_{2A} (55%), A₁-A_{2B} (52%), A_{2A}-A₃ (47%), and A_{2B}-A₃ (45%) (values computed through <https://gpcrdb.org/> [10], focusing on TM1-7). The similarity among the receptors is especially high in the orthosteric binding pocket, where Asn6.55, Phe45.52^{EL2}, Thr3.36, Ser7.42 (replaced by Thr in A₁), His7.43 (numbered according to Ballesteros-Weinstein [11]) are key residues for agonists binding (Figure 1). It is worth noting that few residues discriminate the orthosteric binding pocket of ARs. For instance, A₃ presents the following mutations compared to the other three AR subtypes: Val3.32Leu, Gln3.37His, Asn5.42Ser, Val5.47Ile, His6.52Ser, Thr6.58Ile. Also the following residues show some subtype variability: Val/Ala/Ala/Val 2.57 (respectively in A₁/A_{2A}/A_{2B}/A₃ ARs), Asn/Ser/Ser/Ser 2.65, Leu/Leu/Val/Leu 6.51, Thr/Met/Met/Leu 7.35, and the residues preceding (Glu/Leu/Leu/Gln) and following (Glu/Glu/Glu/Val) the conserved Phe45.52 on EL2. Additionally, major differences among AR subtypes affect the extracellular and intracellular regions of the receptors. A_{2A} presents an unusually long C-terminal tail, more than 80 residues longer than the other subtypes. Extracellular loop 2 (EL2), which has been observed to play a role in ligand recognition [12], is peculiarly different among the receptor subtypes: A₃ is characterized by the shortest loop (24-residues) and A_{2B} by the longest one (34-residues), with A₁ and A_{2A} being in the middle. All EL2s are connected to TM3 through a disulfide bond between the conserved Cys45.50^{EL2} and Cys3.25 on TM3, that is typical of class A GPCRs; moreover, two additional disulfide bonds connect EL2 to EL1 in the case of A_{2A} AR, causing some degree of rigidity to the loop, while the role of three additional cysteines in A_{2B}'s EL2 still needs to be fully elucidated [13,14]. A₁ loop contains a long 3-turn α -helix positioned almost perpendicularly to the membrane leaflet, in A_{2A} EL2 forms a 2.5-turn α -helix parallel to the membrane. In the case of A₃ AR, EL2 contains a short 1.5-turn alpha helix, while the conformation of the long A_{2B} AR EL2 has not been solved yet. Also EL3 is longest in the case of A_{2B}, where it outscores by 4, 2 and 1 residues, respectively, A₃, A₁, A_{2A}; and a disulfide bond between EL3 and the tip of TM6 (Cys6.61) characterizes A₁ and A_{2A}.

Adenosine has a different affinity towards the four ARs. As reported in Figure 2, adenosine shows an affinity in the high nanomolar range towards A₁, A_{2A} and A₃ ARs. Instead, A_{2B} AR presents a considerable weaker affinity (μ M) as compared to the other subtypes (medium to high nM) [15]. This behavior is observed both in humans and rats, except for A₃ AR subtype, for which adenosine displays a micromolar affinity at the rat compared to human receptor. In fact, a high interspecies difference is present between rA₃ AR and hA₃ AR [16].

Adenosine receptors are characterized by different tissue and cell distribution. A₁AR is highly expressed in the CNS: especially in the brain cortex, cerebellum, hippocampus. It is also highly distributed in heart atria, eye, adrenal glands, adipose tissue, testis, kidney, liver and others. A_{2A} AR has the highest expression levels in spleen, thymus, leukocytes (lymphocytes and granulocytes), blood platelets, in various brain regions such as striatum, nucleus accumbens and olfactory tubercles, olfactory bulb, heart, blood vessels, lungs; A_{2B} AR in cecum, colon, bladder, blood vessels [17–19]. A₃ AR is expressed in diverse tissues at relatively low levels, such as testis, spinal cord, various brain regions, bladder, lung, adipose tissue, and whole blood [20]. Moreover, the four AR subtypes are also expressed in various immune cells [21] and overexpressed in cancer cells [22–24].

Receptor distribution is important because both basal and induced adenosine influence cellular behavior depending on the specific expression patterns of adenosine receptor subtype within each cell type, but also depending on the entire dynamic (temporal) process involving the receptor expression and the adenosine release [17,25,26]. Adenosine receptors couple to different G proteins and downstream intracellular signaling pathways. G proteins are the primary intracellular partner of the receptors, but they are also involved in other signaling cascades. For instance, all AR subtypes can engage mitogen-activated protein kinase (MAPK) pathways, but the mechanisms and outcomes of MAPK activation vary depending on the receptor subtype and cell context. Here, G proteins and further relevant interacting ARs interacting partners are reported, together with organ distribution of the receptors [27,28].

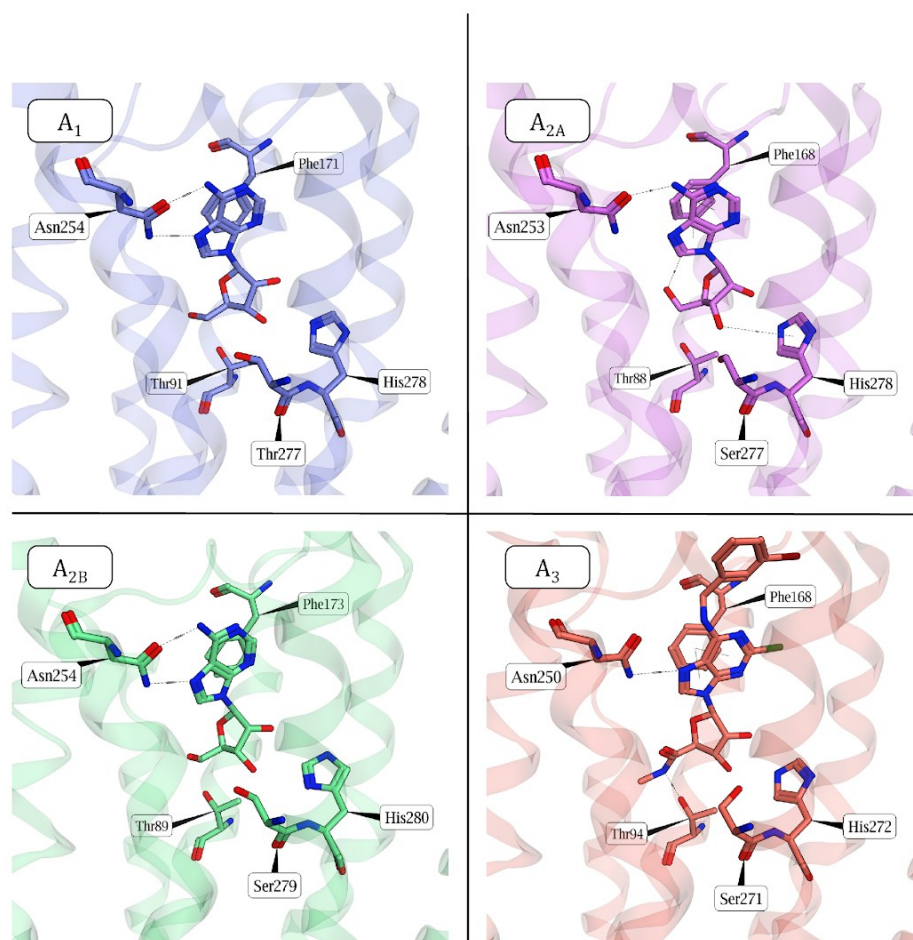


Figure 1. Experimental structures (Cryo-electron microscopy) of the four active-state adenosine receptor (AR) subtypes bound to an agonist. A₁ (PDB ID: 7LD4), A_{2A} (PDB ID: 9EE8) and A_{2B} (PDB ID: 8HDP) ARs are bound the endogenous agonist adenosine, while A₃ AR (PDB ID: 8X17) is bound to Namodenoson (no adenosine-bound structure available to date). The figure shows a focus on the orthosteric binding pocket of the receptors, and it shows the conservation of key amino acids involved on agonist binding.

A₁ and A₃ ARs couples to Gi proteins, whose activated G α i subunits inhibit adenylyl cyclase and decreases intracellular cAMP levels with negative effect on downstream effectors. The G $\beta\gamma$ subunit activates phospholipase C (PLC), with a consequent increase of inositol 1,4,5-trisphosphate (IP3) and intracellular Ca²⁺ levels, which stimulate protein kinase C (PKC) and other calcium-binding proteins [27–29].

A_{2A} and A_{2B} ARs couples to Gs proteins, which activate adenylyl cyclase, increasing cAMP levels and in turn activating protein kinase A (PKA) and its downstream signaling pathways (such as cAMP response element-binding protein, CREB). A_{2A} AR couples mainly to Gs in peripheral systems, but in the striatum, where the receptor is more abundant, it is coupled to Golf, which has a Gs-like behavior [27–29].

Pharmacological and signaling studies suggest that the A_{2A} receptor is the major subtype accounting for adenosine-induced mast cell tPA (tissue plasminogen activator) activity and provide valuable insights into regulation of A_{2A} AR expression in macrophages in response to inflammatory stimuli [30,31].

A_{2B}, and at high agonist concentration, A₃ AR can also couple to Gq protein, which activates PLC [28].

In cardiac muscles and neurons, A₁ AR can activate pertussis toxin-sensitive and K_{ATP} potassium ion channels, and inactivate Q-, P-, and N-type calcium channels, influencing excitability [29].

ARs can also activate MAPKs, such as ERK1/2 (extracellular regulated kinases, ERK) or p38 and JNK (c-Jun N-terminal kinase), thus promoting cell differentiation, survival, proliferation, and death. A₁ can activate ERK1/2 through Gi/o $\beta\gamma$ subunit, A_{2A} can activate or inhibit ERK1/2 in a Gs-dependent or -independent manner, A₃ can activate ERK1/2, and A_{2B} can activate ERK1/2 and also p38 and JNK [27].

[18] Given the many roles of adenosine and its receptors, and their broad distribution, developing drugs against adenosine receptor is challenging [32]. Moreover, pursuing selectivity is not an easy task, but still the availability of numerous pan-ligands provides useful pharmacological tools for these receptors: few examples of pan-agonists (Adenosine and NECA, N-ethylcarboxamidoadenosine) and pan-antagonists (Caffeine and Theophylline) are reported in Figure 2.

Few drugs have been approved so far acting on ARs. The A_{2A} AR agonist Regadenoson has been approved for myocardial perfusion imaging (MPI) and the A_{2A} AR antagonist istradefylline for Parkinson Disease [33].

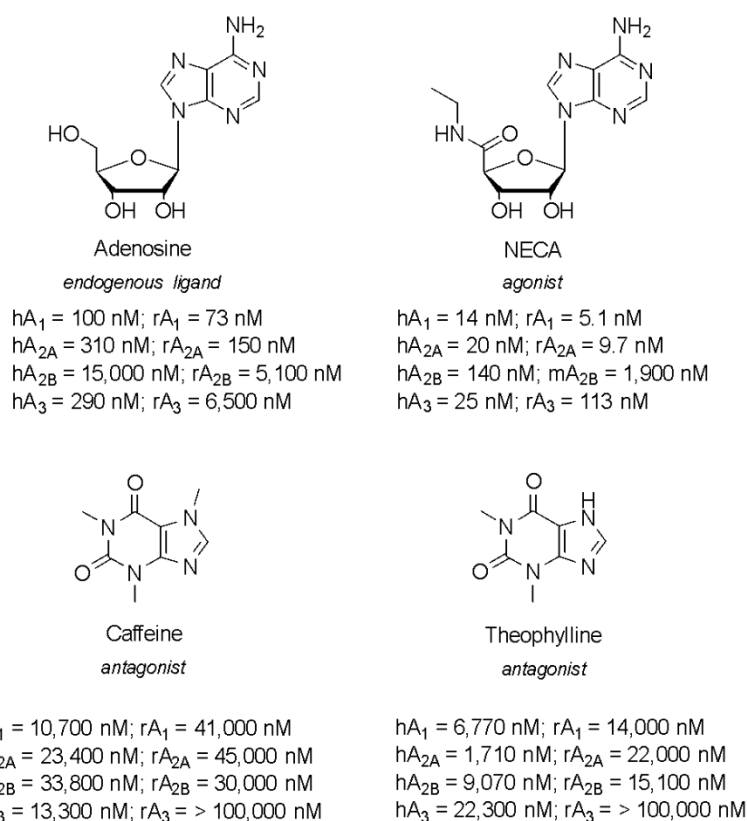


Figure 2. Structures of the endogenous ligand adenosine and of the non-selective agonist NECA and antagonists caffeine and theophylline.

Adenosine, the endogenous ligand of ARs is also a drug, used to treat supraventricular tachycardia (SVT), by slowing down electrical signals at the atrioventricular (AV) node, helping to restore a normal heart rhythm.[34] Caffeine instead, is present in different medicinal specialties as an analgesic adjuvant in combination with other analgesics [35]. Also theophylline, a non-selective AR antagonist, is a currently used drug in chronic obstructive pulmonary disease (COPD), but the therapeutic effect is mediated by inhibition of phosphodiesterases [36].

Neuroinflammation

Neuroinflammation is a complex inflammatory reaction within the CNS. As any inflammatory reaction, it is characterized by the release of mediators of inflammation, that in the case of neuroinflammation are produced by different cell types such as resident CNS glial cells (i.e. microglia and astrocytes), endothelial and peripheral immune cells infiltrated into CNS. Other hallmarks of neuroinflammation are the increased permeability of blood-brain barrier (BBB) and edema[37].

Neuroinflammation is a response to several insults, like infection, stroke, brain and spinal cord injury, traumatic and chronic stressors. It has a beneficial effect as an initial inflammatory response to insults, acting as a protective mechanism to limit tissue damage. However, in prolonged and, even worse, chronic neuroinflammation, there is a continuous release of inflammatory mediators which can cause neuronal

damage over time and major infiltration of peripheral immune cells, with a progressively destructive effect leading to neurodegenerative disease [38,39]. Hypoxia and ischemia are among the events triggering neuroinflammation. [40] Studies have highlighted the attention on the impact of chronic cerebral hypoperfusion (CCH) on the onset and progression of neurocognitive disorders such as Alzheimer's disease (AD) [41,42]. Indeed, evidence suggests a connection between vascular risk and cognitive impairment [43]. Oxygen mediates a post-transcriptional modification on hypoxia-inducible factor-1 (HIF-1)- α and HIF-2 α triggering them to proteasome degradation [44,45]. In hypoxic conditions, this modification does not occur, leading to HIF1 α and HIF2 α stabilization and translocation to the nucleus where they mediate transcriptional regulation [46]. The regulated genes include those encoding enzymes involved in energy metabolism, apoptosis, and inflammation. Notably, adenosine receptors, enzymes responsible for adenosine metabolism (such as CD73 and CD39), and adenosine transporters are among them [47,48]. Among the main actors of neuroinflammation we mention microglia and astrocytes.

Microglia cells are innate resident CNS immune phagocytic cells, characterized by low turnover. Microglia cells exist in a homeostatic (surveying) state in healthy conditions, but get activated upon perturbation of homeostasis, thus changing their morphological and protein expression patterns. Microglia express many pattern-recognition receptors (PRRs) detecting pathogen-associated (PAMPs) or danger -associated (DAMPs) molecular patterns, including Toll-like receptors (TLRs) and their coreceptors, NOD-, LRR- and Pyrin domain-containing protein 3 (NLRP3) inflammasome, nucleic acids receptors, C-Type Lectin Domain Containing 7A (CLEC7A) and others [49]. Inflammasomes are multimeric protein complexes that assemble in the cytoplasm after PAMPs and DAMPs recognition by PRRs. Inflammasomes thus recruit inactive pro-caspase-1 proteins, whose oligomerization leads to their autoproteolytic cleavage leading to active caspase-1, which, in turn, cleaves pro-IL-1 β and pro-IL-18 generating IL-1 β and IL-18. Currently, a two-signal model suggests that a first signal is required to prime NLRP3 and then a second signal is required to activate it. The first signal can consist of binding of lipopolysaccharide (LPS) to TLR, leading to NF- κ B activation and translocation into the nucleus, where it induces NLRP3 expression. Afterward, various stimuli, including ATP, pore-forming toxins, nucleic acids, hyaluronan, and fungal, bacterial or viral pathogens etc. induce NLRP3 formation [50,51].

Activation of microglia can be of the M1 type form, through TLR and interferon γ (IFN- γ) signaling pathways, with microglia exerting a pro-inflammatory and neurotoxic role including: production of proinflammatory cytokines and chemokines, such as tumor necrosis α (TNF- α), interleukin 6 (IL-6), IL-1 β , IL-12, and C-C motif chemokine ligand 2 (CCL2), and expression of nitric oxidase and nicotinamide adenine dinucleotide phosphate (NADPH) oxidase, generating respectively nitric oxide (NO) and ROS.

The alternative form of activation, triggered by different stimuli, leads to M2 microglia, characterized by an opposite anti-inflammatory and neuroprotective effect, with secretion of growth factors, and anti-inflammatory cytokines such as IL-10 and transforming growth factor β (TGF- β) [49]. However, *in vivo* there is no absolute separation between M1 and M2 active states, since different reactive phenotypes exist and appear to be context dependent [52].

Another CNS resident glial cell type playing a major role in neuroinflammation is astrocyte, which also distinguishes into A1 inflammatory and neurotoxic astrocyte, and A2 neuroprotective astrocyte. Their activity is regulated by Janus kinase-signal transducer and transcription activator 3 (JAK/STAT3), nuclear factor kappa B (NF- κ B), MAPK and calcineurin pathways [53]. In homeostatic conditions they also contribute to maintaining CNS homeostasis, while during neuroinflammation they undergo a reactive state, in which they participate to the release of pro-inflammatory cytokines. Moreover, astrocytes secrete various factors affecting tight junction proteins (e.g. morphogens such as Sonic hedgehog (Shh) and retinoic acid, growth factors such as vascular endothelial growth factor (VEGF) and many others), regulating the integrity and permeability of BBB, which can thus be altered once astrocytes are in the reactive state [54].

The alteration of BBB permeability let the migration of peripheral immune cells into CNS, such as neutrophils, which contribute to BBB impairment, monocytes, natural killer cells, dendritic cells, T-cells, and B-cells, with a consequent exacerbation of the neuroinflammation course [38].

Neurodegeneration

As previously stated, neuroinflammation, if not limited, can lead to CNS dysfunction and damage

contributing to the complex multifactorial framework causing neurodegenerative diseases. Neurodegenerative diseases are characterized by the progressive dysfunction and loss of neurons in the CNS. Their mechanism is not completely clear, but includes aging, immunity, and neuroinflammation [55].

Parkinson's disease (PD), Alzheimer's disease (AD), Huntington's disease (HD), and Amyotrophic Lateral Sclerosis (ALS) are defined as neurodegenerative disorders, whose features are neuronal dysfunction and death. All of them are characterized by common risk factors, such as oxidative stress, environmental factors, aging, and protein dysfunction. Regarding proteins, misfolding and aggregation of specific proteins is a common characteristic [56]. A brief introduction to the neurodegenerative disease under discussion in this review is provided in the following paragraphs.

Alzheimer's disease. AD, the most common form of dementia, is characterized by synapse loss and neuronal atrophy that predominately hits hippocampus first and then cerebral cortex [57]. The disease leads to a progressive loss of memory, worsened cognitive functions including language, and failure in performing simple and daily activities [58,59]. The formation of extracellular senile plaques composed of amyloid β protein ($A\beta$) and intracellular neurofibrillary tangles (NFTs) consisting of hyperphosphorylated tau (τ) protein, are the typical hallmarks of the disease [60]. $A\beta$ is generated through the proteolytic processing of amyloid-beta precursor protein ($A\beta$ PP) by β -secretase 1 (BACE1) followed by γ -secretase [57]. AD is caused by a combination of both environmental and genetic factors, thus is a complex and multifactorial disease for which several hypotheses have been postulated, such as, along with amyloid and tau protein, there are the cholinergic, inflammatory, glutamate excitotoxicity, oxidative stress, metal ion, microbiota-gut-brain axis, and abnormal autophagy [61]. The neurodegeneration characterized by neuronal loss is accompanied by the neuroinflammatory process, oxidative stress, and several cellular abnormalities [62]. Regarding neuroinflammation, both microglia and astrocytes are activated in AD. In the early stages, this activation is beneficial, as these glial cells help clear amyloid-beta ($A\beta$) deposits. However, as the disease progresses, chronically activated microglia begin to exert detrimental effects, contributing to neurodegeneration in surrounding brain regions [63]. Astrocyte reactivity also intensifies with increasing $A\beta$ accumulation and tends to co-localize with amyloid plaques. In contrast, astrocytes located away from plaques are often atrophic, leading to reduced support for neurons and synaptic connections. This loss of support is associated with early cognitive impairment [64].

Parkinson's disease. PD is a progressive disorder of the central nervous system that affects mobility. The motor symptoms of PD are due to the degeneration of the dopaminergic neurons in the nigrostriatal pathway, which in turn resulted in depletion of dopamine production [65]. The direct and indirect pathways of the basal ganglia are essential for motor control, both originating in the striatum and modulated by dopaminergic input. These pathways exert opposing effects on thalamic activity: the direct pathway facilitates movement by disinhibiting thalamic neurons, while the indirect pathway suppresses movement through increased inhibition. In Parkinson's disease (PD), the degeneration of dopaminergic neurons in the substantia nigra pars compacta disrupts this balance—leading to underactivity of the direct pathway and overactivity of the indirect pathway. This imbalance contributes to hallmark motor symptoms such as bradykinesia, rigidity, and tremor [66,67]. Lewy Bodies are intracellular inclusions that represents one of the hallmarks of PD and are composed mainly by α -synuclein (α -Syn) aggregates, which accumulation contributes to the decrease of dopamine biosynthesis by inhibition of tyrosine hydroxylase [68–71].

Other neurodegenerative disease. Other neurodegenerative diseases involving ARs discussed in this review are ALS and HD.

ALS is characterized by progressive loss and injury of motor neurons that leads to muscle weakness, followed by paralysis, and fatal respiratory failure. The main pathogenic factors of this disorder are represented by mutations in genes of Trans activation response DNA binding protein 43 (TDP-43), superoxide dismutase 1 (SOD1), Chromosome 9 Open Reading Frame 72 (C9orf72) protein, and Fused in Sarcoma (FUS, also called translocated in liposarcoma, TLP) protein [72]. These mutated proteins, and in particular TDP-43, are found in the aggregates typic of ALS [73]. Normally, TDP-43 is in the nucleus where it binds both DNA and RNA, regulating mRNA at different stages [74,75]. In ALS, TDP-43 mislocalizes from the nucleus to cytoplasm, leading to a deep deregulation of mRNAs in the cell. The trigger point that determines the degeneration of motor neurons is the aggregation of this anomalous protein in the cytoplasm within and outside stress granulates (SGs), contributing to produce oxidative

stress and neuronal toxicity [76–79]. In ALS, neuroinflammation involves both microglia and astrocytes, which acquire neurotoxic phenotypes that promote motor neuron death. Microglial activation has been observed early in disease progression, with elevated cytokines such as monocyte chemoattractant protein-1 (MCP-1) and IL-6 in the CNS and periphery. Astrocytic glutamate dysregulation and impaired support to neurons further exacerbate excitotoxicity [80].

Finally, HD is a hereditary condition marked by neuropsychiatric symptoms, a movement disorder—most commonly chorea—and progressive cognitive decline.[81] characterized by progressive striatal neurodegeneration. The mutation responsible for this disease is an abnormally expanded and unstable CAG repeat within the gene encoding huntingtin protein [82]. This is ubiquitously expressed but high levels are present in the brain. Mutated huntingtin forms toxic mutant huntingtin oligomers. GABAergic medium spiny neurons of the striatum are the most vulnerable cells in HD [81].

In the following pages, the single AR subtypes are discussed separately, trying to summarize the main discoveries in terms of both general neuroinflammation and specific neurodegenerative diseases. In addition, at the beginning of every chapter, a list of the most common and significant ligands, both agonists and antagonists, used as tools for that specific receptor, is given with the affinity data for all the receptors, when present, and, if available, also the data on the rat or mouse receptor, due to the usual employment of these species as animal models. The aim is to acknowledge the effective selectivity of the tools used, which can help to interpret data obtained from an experiment and a guide to choose the correct tool for future experiments.

A₁ AR

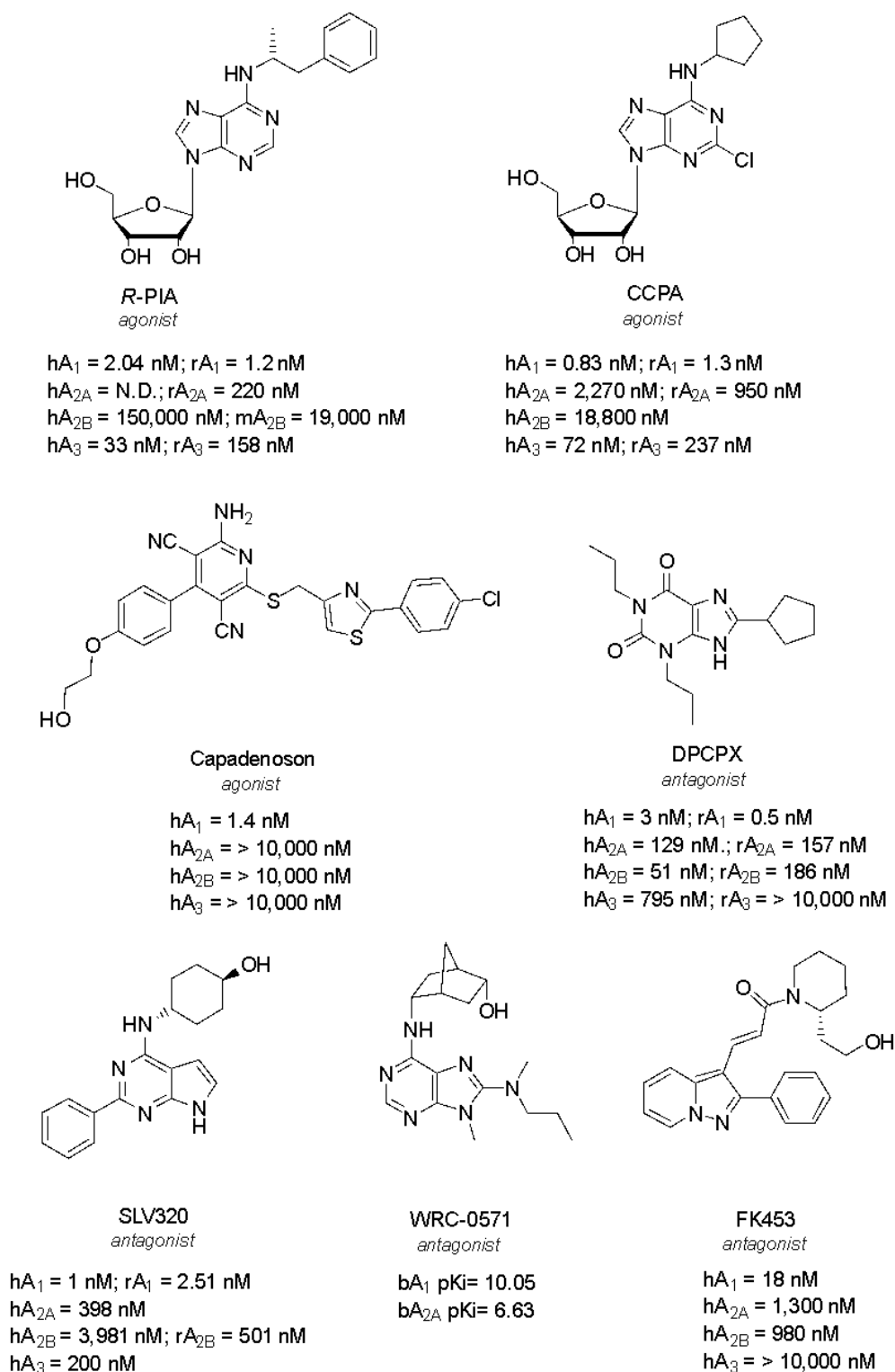
Being the first AR to be identified, most of the ligands for this receptor were discovered in the last century. The most representative agonists and antagonists used as pharmacological tools for A₁ AR are reported in Figure 3, but it is important to highlight that also several allosteric modulators have been reported for this AR subtype [77–80].

Role in neuroinflammation

A₁ AR is widely expressed in the brain and particularly in neurons and microglia [83,84]. Even if microglia activation can be considered the initial trigger of neuroinflammation[85], activation of A₁ AR in these cells was found to both promote and counteract inflammation [86–89]. Diseased A₁ AR knockout (KO) mice (e.g. by traumatic brain injury, TBI, [90] or glioblastoma [91]) showed an increased activity of microglia and neuroinflammation, suggesting an anti-inflammatory role for the A₁ AR (Figure 4). Chronic caffeine intake, which cause A₁ AR upregulation, alleviates inflammation in a model of multiple sclerosis (MS) [88]. A₁ AR agonists were found to inhibit the morphological activation of microglia, probably blocking calcium influx triggered by ATP treatment [92].

Astrocytes express all ARs on their membrane but A₁ ARs reveal the highest binding ability to adenosine, which leads to a decrease in production of IL-12, IL-23 and chemokines, generating an immunosuppressive effect, useful in conditions like autoimmune encephalomyelitis (Figure 4) [93]. On the contrary, recently, Guo et al. demonstrated that A₁ ARs in astrocytes, and not in pericytes, microglia or oligodendrocyte precursors cells, trigger the early stage of inflammatory response of microglia in a model of LPS-induced sepsis-associated encephalopathy (SAE) [94]. Another crucial point highlighted by this work is that peripheral adenosine could pass the BBB directly enhancing its central extracellular levels [95].

Noteworthy, adenosine, via A₁ AR, and glutamate signaling are strictly connected, partly explaining the adenosine-mediated neuroprotective effect [5]. In fact, presynaptic A₁ AR decreases glutamate release in synapses, while post-synaptic A₁ AR activation limits N-methyl-D-aspartate (NMDA) receptor activation and calcium current influx inward flux, suppressing excitatory action potentials (Figure 4) [86,96]. Under hypoxic conditions, this response is thought to serve as a compensatory strategy to safeguard neurons, by balancing the oxygen supply and its consumption [97]. In seizures, conditions characterized by neuronal hyperexcitability, A₁ AR was found neuroprotective [85]. Regarding the time course of its neuroprotective effect, adenosine has been shown to inhibit excitatory synaptic transmission in the hippocampus, probably via A₁ AR, both during ischemia and for a significant period afterward [4].

Figure 3. Structures of the most representative A₁ AR ligands. All the reported values are Ki.

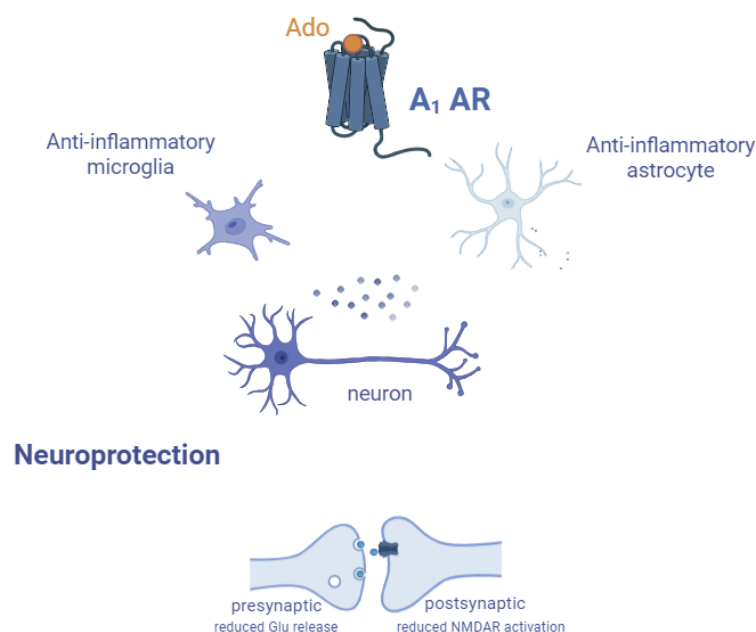


Figure 4. General neuroprotective effect mediated by A₁ AR activation. Receptor activation has an anti-inflammatory effect on both microglia and astrocytes, that along with its ability to reduce glutamate (Glu) release at presynaptic level in neurons and to reduce postsynaptic NMDA receptor activation, confers global neuroprotection.

The relation between adenosine and neuroinflammation has been investigated also in depression, a psychiatric disorder. Adenosine has been found to behave antidepressant effects, especially towards A_{2A} AR, discussed afterwards, but also by A₁ AR activation by astrocyte-derived adenosine [98–100].

As reported in the introductory part, A₁ AR is expressed in heart tissue, where the use of agonists can cause bradycardia, atrioventricular block or hemodynamic effects. Possible strategies to avoid this include using partial agonists, positive allosteric modulators or biased agonists instead of full agonists [101,102].

Role in neurodegenerative disorders

There are several theories regarding the mechanisms of neurodegeneration. One of these is the glutamate hypothesis, which postulates that neurodegeneration is driven by excitotoxicity resulting from excessive glutamate signaling. Excessive recruitment of NMDA receptors, which is associated with spreading depolarization, leads to dysregulated calcium influx, ultimately causing neuronal damage [103]. As reported before, adenosine, mainly through A₁ AR, decreases this excitatory transmission [96,104], and several studies reported a related decrease in the neurodegenerative process both in *in vitro* and *in vivo* models [105–110]. In fact, as already highlighted, adenosine, also mainly through A₁ AR affects metabolism activity not only at neurons but also at astrocytes and microglia [111–116]. These findings contrast with the documented neuroprotective effect attributed to the AR antagonist caffeine when consumed in a chronic but moderate way [117], even if it has been related to an A₁ AR upregulation [118].

The principal drawback in using full A₁ AR agonists to treat neurodegenerative diseases is the on-target side effects on heart[102]. In addition, A₁ AR agonism has been reported to be protective in several non-central tissues, suggesting a general protective mechanism that is not solely related to its role as a neuromodulator [119]. Another important aspect to consider when evaluating A₁ AR as a target in neurodegenerative diseases is that neuroprotection has only been observed when receptor activation occurs shortly before or simultaneously with the noxious stimulus. During such events, elevated levels of extracellular adenosine led to A₁ AR desensitization, resulting in a loss of neuroprotective effects and, in some cases, even worsening brain damage after prolonged exposure to adenosine [120–123]. This observation aligns with recent studies on AR KO mice, which suggest that A₁ AR plays a role primarily in the early stages of neuronal damage [110]. This supports the idea that A₁ AR activation may have a more prophylactic than therapeutic effect. Additionally, adenosine released at the site of injury can diffuse to

nearby regions, providing localized protection—a phenomenon known as brain preconditioning [124,125]. Therefore, A₁ AR agonism may be beneficial in cases of acute, repetitive, and short-term brain insults. However, in the context of chronic neurodegenerative processes, dysfunctional astrocytes lose their ability to mediate preconditioning, contributing to the spread of neurodegeneration [126].

Alzheimer's Disease. It is important to note that the role of ARs in Alzheimer's disease is under-investigated compared to Parkinson's disease and has been studied primarily in *in vitro* and *in vivo* models. However, post-mortem brain samples from AD patients showed a decreased expression of A₁ ARs in dentate gyrus, CA1 and CA3 regions of the hippocampus, which was attributed to neuronal loss in these brain regions, the most involved in learning and memory [127–130], even if a reduction of A₁ ARs in CA1 region of hippocampus has been observed also in other disorders with dementia [131]. Instead, Angulo *et al.* found a non-significative difference in mRNA levels of A₁ AR in the hippocampus of AD patients respect to healthy subjects, but they found an increase in protein levels in degenerating neurons with NFT and in dystrophic neurites of senile plaques [132]. Concerning senile plaques, in human neuroblastoma cells, A₁ AR activation increased the production of a soluble secreted form of APP, supporting the use of agonists for the treatment of AD [132–134]. Another more recent study reported that A₁ AR, along with A_{2A} AR, showed increased expression in the frontal cortex in AD [135].

Parkinson's Disease. A₁ AR is highly expressed in substantia nigra (SN), and its gene is in a locus strongly related to PD [136]. Mutation or disfunction of A₁ AR are related to a worse scenario of the disease, inhibiting dopaminergic signaling [136–138]. In fact, A₁ AR co-localize with the dopamine D1 receptors on the GABA/dynorphin output neurons where they can influence each other's function at either the receptor or intracellular signaling cascade level.[139] A₁ AR/D1 R can form functional heteromers, where A₁ AR antagonistically and specifically modulate the binding and functional characteristics of dopamine D1 receptors.[139–142] Instead, A_{2A} ARs are mostly highly expressed in intermediate spiny neurons of the striatum in close association with dopamine D2 receptors (indirect pathway) and, again, specific negative interactions have been described.[140,143,144]

As already said, chronic A₁ AR stimulation may aggravate neuroinflammation and, in PD, has been also observed to induce synucleinopathy in some rodent models [145,146]. The mechanism seems to reside in the binding of the A₁ AR to the N-terminus of α -syn promoting its misfolding [138,146], this suggests that not only A_{2A} AR is involved in this pathology and that further studies are needed to understand the exact role of this receptor in the disease.

A_{2A} AR

A_{2A} AR has been extensively studied, and a huge number of selective agonists and antagonists are available. A considerable number of them reached or are still in clinical trials, in particular for the treatment of Parkinson's disease and cancer. Among agonists: regadenoson is an A_{2A} AR agonist approved for myocardial perfusion imaging, while istradefylline, is an approved antagonists as adjuvant in classical dopaminergic PD therapy. Representative ligands are reported in Figure 5 [84,147,148].

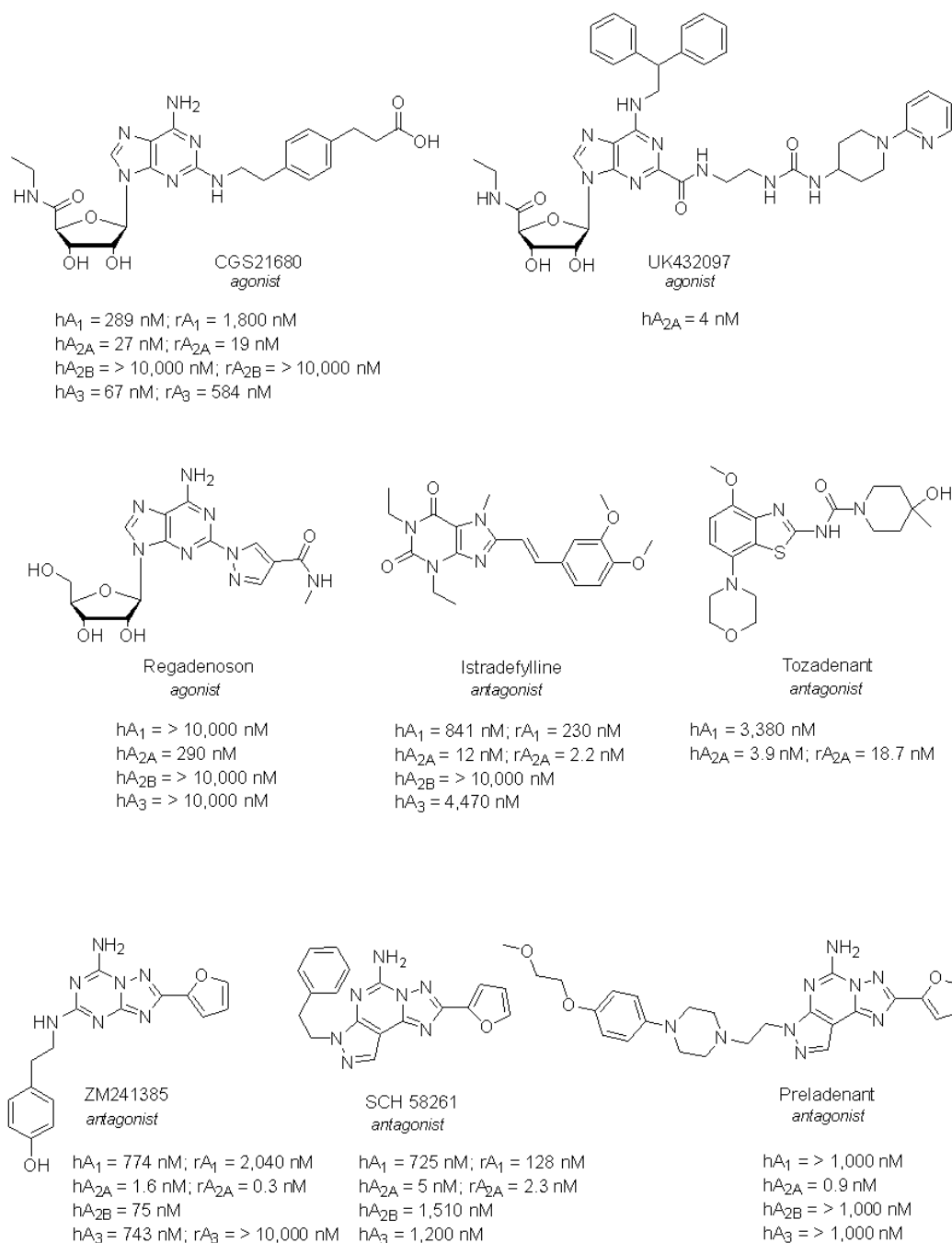


Figure 5. Structures of the most representative A_{2A} AR ligands. All the reported values are Ki.

Role in neuroinflammation

In general, adenosine receptors, especially towards A_{2A} AR, are involved in inflammasome activation; in fact, increase of cAMP levels is followed by PKA activation that leads to an augmented potassium efflux that leads to activation of protein kinase A (PKA) and which triggers caspase-1 activation [149,150]. Moreover, elevated adenosine levels in the extracellular space have been associated not only with increased caspase-1 activity but also with enhanced IL-1 β production (Figure 6) [149,151]. As for A₁ AR, also A_{2A} AR regulates both microglia and astrocytes during neuroinflammatory events, and, again, showing both pro- and anti-inflammatory effects [85,152]. In particular, Pedata *et al.* reviewed the role of A_{2A} ARs in brain ischemia suggesting that antagonists for these receptors are protective during acute injury, by alleviating hyper-excitotoxicity at a central level, while agonists allow prolonged protection by regulating leukocyte infiltration also after days from the ischemic event [153]. A sustained anti-

inflammatory behaviour of A_{2A} AR agonists was also observed after blunt spinal trauma, resulting in a reduction of long-term neurologic injury [154].

As reported earlier in this review, adenosine, via A_{2A} AR, mediated vasodilation, in fact, the agonist Regadenoson is clinically used in imaging with this purpose. Thus, the use of A_{2A} AR agonists for their anti-inflammatory (or others) effect is accompanied by vasodilation and a decrease in blood pressure, that should be taken in consideration for drug development purposes. The role of cardiac A_{2A} AR has been quite recently reviewed [155,156].

In particular, A_{2A} AR was found to have a role in hyperexcitability and excitotoxicity (i.e. by increasing glutamate release) in the CNS, not only in striatum, but also in hippocampus and cortex (Figure 6) [157–159]. The other big role of A_{2A} AR upregulation is that of control microglial activity (Figure 6) by 1) promoting their proliferation; 2) mediating microglial process retraction, an important marker of neuroinflammation [159–162]. At physiological conditions A_{2A} AR is not highly expressed on microglia, but its expression increases as a result of brain insults leading to the activation of signaling pathways different to those active at the physiological state [163]. Process retraction conveys in a modification of microglia cells from a ramified to an amoeboid shape, and specifically A_{2A} AR mediates activation to the so-called M1 phenotype, which is characterized by overexpression of inflammatory enzymes and release of pro-inflammatory mediators [160,164,165]. A_{2A} AR antagonists seems to be beneficial in animal models of perinatal brain injury and of Parkinson's disease [162,166]. A similar behavior has been observed in anxiety disorders, which are also characterized by altered microglia. These disorders can be caused by prenatal exposure to glucocorticoids, and perinatal treatment with A_{2A} AR antagonists was able to ameliorate the inflammatory state and the anxiety behavior [151,167].

Even if the principal effects of A_{2A} AR activation are observed in microglia, there are some studies also on astrocytes. This receptor regulates the uptake of both GABA and glutamate, thus fine-tuning the inhibitory and excitatory modulation at synapses [168,169]. A_{2A} AR antagonists were able to inhibit astrogliosis both *in vitro* and *in vivo*, contributing to understanding the mechanism of A_{2A} AR blockade-mediated neuroprotection (Figure 6) [170,171]. Due to the A_{2A} AR upregulation observed in AD, A_{2A} AR overexpression in primary cortical astrocytes has been studied revealing alterations in the astrocytic transcriptome with an important effect on genes relevant for cell-activation, immune response and also angiogenesis [172]. Various studies concur in attributing a significant role to astrocytic A_{2A} AR in regulating synaptic plasticity and memory, thereby opening new avenues for therapeutic interventions targeting this receptor (see discussion below) [168,173,174]. It should be noted that opposite results were instead reported by Yuan *et al.* in mice with CCH-induced white matter lesions, in which authors observed an anti-inflammatory effect by A_{2A} AR activation, that involves STAT3/YKL-40 inhibition on astrocytes [175]. In a model of oxygen glucose deprived (OGD) hippocampal slices A_{2A} AR are protective by delaying astrocyte activation along with the appearance of anoxic depolarization [176]. This is in line with the well-known role of A_{2A} ARs in cancer, which is involved in the tumor immune escape [177].

Finally, as anticipated, antidepressant effect of adenosine is mainly attributed to its action through the A_{2A} AR [98–100], even if also A_{2A} AR KO or the use of antagonists, including caffeine, demonstrated antidepressant-like effects in *in vivo* tests specific to predict antidepressant activity [178–180]. This effect is principally due to the counterbalance given by A_{2A} AR to the reduction of dopamine, via A_{2A} AR-D2 R heteromer regulation [178].

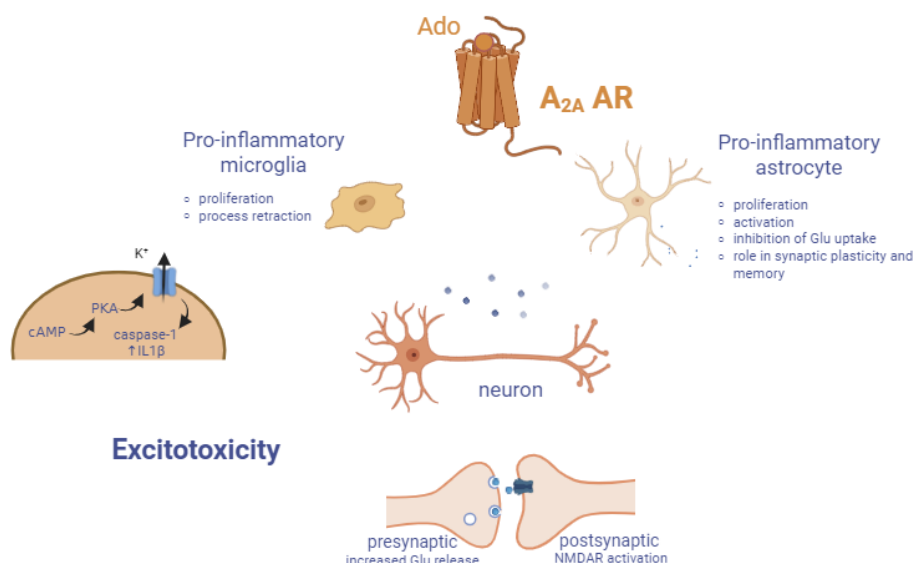


Figure 6. General proinflammatory and excitotoxic effect mediated by A_{2A} AR activation. Receptor activation had a pro-inflammatory effect on both microglia and astrocytes. In addition, A_{2A} AR, via cAMP, that after activated PKA which, by augmenting potassium efflux leads to caspase-1 activation, resulting in the increase of the proinflammatory cytokine $IL1\beta$. At neurons, presynaptic A_{2A} AR activation increases glutamate (Glu) release participating to the excitotoxic effect.

Role in neurodegenerative disorders

A_{2A} ARs, due to their ability to bind SAP102 (synapse-associated protein 102), are mainly located in glutamatergic synapses[181], where they activate neurons in the hippocampus. Specifically, A_{2A} AR activation increases the entry of calcium through voltage-sensitive calcium channels, depolarizes neurons, enhances the release of glutamate and the activity of NMDA receptors, thus promoting excitotoxicity [121]. The use of A_{2A} AR antagonists arose as an efficient therapeutic strategy in various models of neurodegenerative disorders [182]. Then, the role of A_{2A} ARs already discussed on glia cells supported this receptor as a valuable target in neurodegeneration [183]. In addition, a correlation between A_{2A} AR expression in brain with those in platelets indicated the receptor as a possible marker of AD [184].

The use of A_{2A} AR antagonists for neurodegenerative diseases could lead to side effects related to the other phenomena mediated by receptors. In particular, an antagonist could potentially give vasoconstriction, inflammation or sleep disturbance. Noteworthy, the FDA approved Istradefylline does not exhibit these side-effects. Reasonably, the use of A_{2A} AR antagonists should be safe. Some other side effects that resulted in failed clinical trials for other compounds (i.e. Tozadenant for agranulocytosis) were not clearly proved to be on-target side effects [185].

Alzheimer's Disease. In post-mortem brains of AD patients both A_1 and A_{2A} AR levels were increased in the frontal cortex [134,186]. In healthy subjects, A_{2A} AR is mainly present in striatal neurons, while in post-mortem AD samples, it has been primarily found in glial cells, not only in cortex, but also in hippocampus [132,187]. Thus, combining these data, along with that already reported concerning the A_1 AR, it can be pointed out that expression and distribution of ARs in AD are region and cell specific. In particular, A_{2A} AR in AD is highly expressed in astrocytes, where it regulates Na^+/K^+ -ATPase, responsible to sustain several processes occurring at the membrane level, like regulation of glutamate transporter 1 (GLT-1), that controls glutamate uptake [188]. A_{2A} AR genetic removal in astrocytes enhances memory performance of aged mice [189]. Besides the general neuroprotective effect observed with caffeine or the use of A_{2A} AR antagonists, caffeine demonstrated beneficial effect also in specific transgenic mouse models of AD bearing mutant forms of $A\beta$ PP and/or presenilin-1 (PS1) [190–192], reducing both synapsis loss and neurotoxicity in hippocampus of rat models [193]. All these findings indicated a crucial role of A_{2A} AR in controlling synaptic plasticity and memory, thus suggesting that antagonists are useful to prevent synaptic dysfunction and cognitive deficits like memory loss [194–196]. Temido-Ferreira et al. found that

overexpression of A_{2A} AR in rats is sufficient to impair LTD-to-LTP transition in neurons, a hallmark of memory impairment [197].

Recent studies provide compelling evidence that neuronal A_{2A} ARs dysfunction, as seen in the brain of patients, contributes to amyloid-related pathogenesis and underscores the potential of A_{2A} AR as a relevant therapeutic target for mitigating cognitive impairments in this neurodegenerative disorder [198]. In addition, very recently has been highlighted the impact of astrocytic A_{2A} AR upregulation, as seen in various neurological conditions, on the development of a detrimental multicellular response associated with memory alterations and provides an additional proof-of-concept for the value of targeting this receptor in different neurodegenerative conditions [199].

Parkinson's Disease. A_{2A} ARs play a crucial role in the pathophysiology of PD due to their interaction with the dopaminergic system in extrapyramidal regions. Specifically, A_{2A} AR antagonists are therapeutically beneficial by reducing the inhibitory output of the basal ganglia's indirect pathway, thereby helping to restore motor function [134,200,201]. In fact, at the cell bodies of medium spiny neurons in the striatum, A_{2A} AR and D₂ Rs are co-localized and interact antagonistically. This existence of A_{2A}-D₂ heteromeric complexes has been demonstrated on co-immunoprecipitation studies and on fluorescence resonance energy transfer and bioluminescence resonance energy transfer analyses. It has now become possible to show that A_{2A} and D₂ receptors also co-immunoprecipitate in striatal tissue, giving evidence for the existence of A_{2A}-D₂ heteromeric receptor complexes also in rat striatal tissue [202].

Thus, blockade of A_{2A} AR enforces the effect of the diminished D₂R-induced stimulation by pathologically low dopaminergic innervation from the damaged substantia nigra pars compacta. Reported effects of blocking A_{2A} AR in PD are improvement of motor function, attenuation of dyskinesia, reduction of α -synuclein aggregation and, alleviation of non-motor symptoms like depression and cognitive disfunction [203–208]. In addition, as already mentioned in the neuroinflammation section, there is a high level of activation of microglia in the midbrain/substantia nigra in PD, and A_{2A} AR antagonists are found to return microglia in their resting state, decreasing neuroinflammatory response [209,210]. It is important to note that A_{2A} AR antagonists are effective in PD, but as an adjuvant therapy to complement dopamine D₂ R agonists treatment [211]. Several A_{2A} AR antagonists are in human trials for Parkinson's disease based on their motor enhancement feature in various animal models, nevertheless only a compound named Istradefylline (1,3-diethyl-8-(3-methoxystyryl)-7-methylxanthine, has been approved in Japan and USA for treatment of basal ganglia disorders such as Parkinson's disease, in association with L-DOPA [204,212]. Several studies on xanthine-based A_{2A} AR antagonists for the treatment of PD have been reported so far, Petzer et al. investigated if some of these compounds, including Istradefylline, were active also on monoamine oxidase B (MAO-B). MAO-B is the major enzyme responsible for the dopamine catabolism in CNS, thus, its inhibition is another useful strategy in PD therapy. The results showed that most compounds inhibited MAO-B in the micromolar range, compared to their nanomolar affinity for the A_{2A} AR. This suggests that both *in vitro* and *in vivo* findings should be interpreted with caution, as the observed effects may result from multiple mechanisms of action. Results demonstrated that most of them inhibited MAO-B in the micromolar range (respect to nM affinity at the A_{2A} AR), revealing that results obtained *in vitro* or *in vivo* should be cautiously interpreted because they can be ascribed to a multiple mechanism of action. It is important to note that several A_{2A}/MAO-B dual targeting ligands were rationally designed in the last decades [213–216].

Huntington's Disease. In HD, A_{2A} AR is highly expressed in the vulnerable GABAergic enkephalin neurons of the basal ganglia, but not in other neurons that equally express mutant and normal huntingtin suggesting an involvement of A_{2A} ARs in HD [217]. Experimental models of HD revealed alterations in signaling of A_{2A} AR as also in their expression, last has been observed not only in neurons but also in peripheral blood cells. This higher expression has been related more from an altered receptor turnover than from an increased transcription, as A_{2A} AR mRNA levels remain unchanged [218,219]. In a mouse model of HD, a specific altered A_{2A} AR phenotype has been found in the early stages of disease [219]. Instead, in later stages, a significant decrease in A_{2A} AR transcription has been observed. However, conflicting data about the potential neuroprotective and neurodegenerative effects of these receptors in the brain have been reported, with beneficial effects shown by both A_{2A} AR agonists and antagonists [217,220].

Amyotrophic Lateral Sclerosis. A_{2A} AR antagonists have been reported to reduce motoneuron death in ALS [221–223]. The purposed mechanism is that motoneurons vulnerability to excitotoxicity is due to

tyrosine receptor kinase B (TrkB) activation by brain-derived neurotrophic factor (BDNF), and A_{2A} AR transactivates TrkB. This transactivation occurs even in absence of the ligand BDNF, it is BDNF-independent [221,222]. On the contrary, another research group advanced an opposite theory, based on the fact that neurotrophins are potent survival factors for developing and injured neurons, and that neurotrophins were not useful for therapy due to administration and side effect issues. Thus, they suggest using transactivation, like that mediated by A_{2A} AR for TrkB, to trigger a neuroprotective effect [188]. In fact, in their work A_{2A} AR agonists were able to delay disease onset in mice [188]. Authors also reported the inability of BDNF to protect motoneurons, and attributed this to the overactivation of a truncated form of the TrkB receptor rather than the full-length form transactivated by A_{2A} AR agonists [224,225]. In SOD1(G93A) mice, the classical animal model of ALS which bears a mutation in superoxide dismutase 1 (SOD1) gene, the effect of A_{2A} AR agonists has been investigated both in the pre-symptomatic and symptomatic stage of the disease (e.g. 4–6 weeks and 12–14 weeks old, respectively). Results indicated a beneficial involvement of A_{2A} AR only in pre-symptomatic individuals, which is in line with the previous evidence discussed of the A_{2A} agonist-delaying effect on the onset of ALS [226]. Symptomatic stage has been related also with an impairment of long-term potentiation (LTP), suggesting that the decreased synaptic plasticity can involve the A_{2A} AR over-activation of the early disease stages [227]. In another ALS mouse model which expressed human TDP-43, A_{2A} AR activation rescued TDP-43 cytoplasm mislocalization, which has been proved to be counteracted by activation of the co-localized D2R in motor neurons of ALS patients [228,229].

Regarding disease progression, in experiments where pharmacological inhibition and partial genetic ablation of A_{2A} ARs were performed in SOD1(G93A) mice, a significant delay in disease progression was observed. A study by Armida et al. aimed to clarify existing uncertainties by administering an A_{2A} receptor agonist or antagonist to mice starting from the presymptomatic stage. Unfortunately, in neither case was there an improvement in motor skill deterioration or survival in SOD1(G93A) mice [230].

Conversely, in the same animal model treated with a daily intake of caffeine for 70 days, mice experienced a significantly shortened survival [231].

In a study performed on 377 newly diagnosed ALS patients, compared to 754 controls, caffeine was found to reduce ALS risk, while in another study on a larger cohort of patients, this association has not been demonstrated [232,233]. A recent study aimed to investigate the effect of caffeine consumption on ALS patients failed to find a correlation with disease progression and survival, but higher caffeine consumption was found to improve significantly the cognitive performance in a specific cohort of patients, that carry the minor allele T of rs2472297 and are considered as fast metabolizers [234]. [219] Thus, high discrepancy on the role of A_{2A} AR in ALS is present, due to the complex role of the purinergic signaling in a multifactorial and multisystemic disease like ALS [134]. Vincenzi *et al.* showed an up-regulation of A_{2A} AR in lymphocytes from blood samples of ALS patients, demonstrating a positive correlation between receptor density and the scores of the Revised Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS-R) that suggest again a possible protective effect of this receptor subtype [235].

Others. As reviewed by Cunha R. A., A_{2A} AR antagonists have been proposed also for other neurodegenerative disorders like diabetic neuropathy, ataxia-related Machado-Joseph disease, experimental autoimmune encephalomyelitis, and subarachnoid hemorrhage [121].

Pharmacological studies, by Chen et al, indicate that A_{2A} ARs play a prominent role in the development of ischemic injury within brain and demonstrate the potential for anatomical and functional neuroprotection against stroke by A_{2A} receptor antagonists [236].

A_{2B} AR

A_{2B} AR, which is activated only at high concentrations of adenosine, has been a challenging target for the development of selective ligands. In fact, only in recent decades have effective pharmacological tools become available. To date, no potent and selective A_{2B} AR agonists are accessible. BAY60-6583, which was initially defined as an agonist, has since been shown to behave as a partial agonist [237]. Instead, initially several xanthines have been reported as selective antagonists (i.e. PSB603 and the xanthine-like BAY-545), but as already said, recently, new structures have been developed (i.e. OSIP339391 and LAS38096) [17,84,238,239]. Representative structures of ligands towards A_{2B} AR are displayed in Figure

7. Etrumadenant is a dual A_{2A} AR/ A_{2B} AR antagonist currently in clinical trials for cancer [240,241], but is reported here as an example of a dual-targeting ligand that, due to the complex regulation in the CNS by different AR subtypes, could be an effective strategy for the treatment of neurodegenerative diseases.

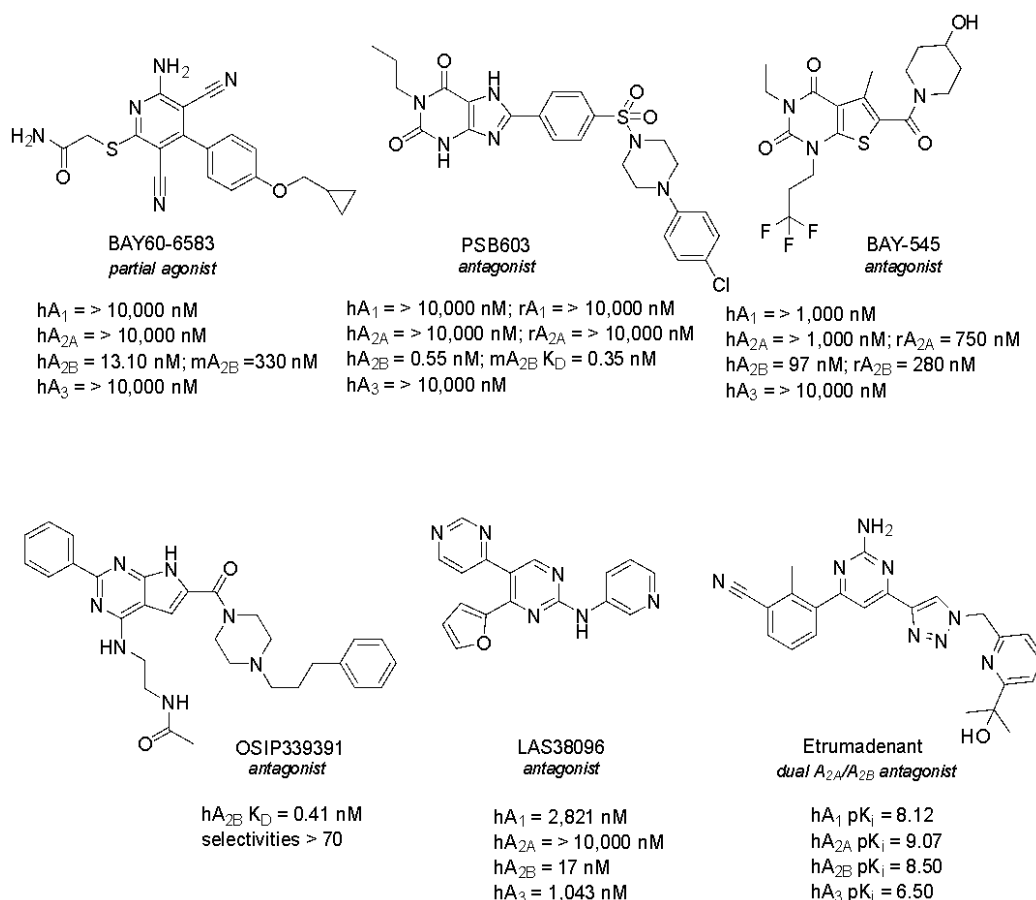


Figure 7. Structures of the most representative A_{2B} AR ligands. All values reported are K_{is} , unless stated otherwise.

Role in neuroinflammation

A_{2B} AR are widely distributed throughout the brain [242]. In primary murine microglia, through p38 MAPK kinase pathway, the A_{2B} AR has been found to promote production of both anti-inflammatory cytokine IL-10 [243], and pro-inflammatory IL-6, along with stimulation of microglia proliferation [244,245]. Downregulation of A_{2B} AR in inflamed brain is reported [246].

A_{2B} AR has been reported among the most highly expressed G protein-coupled receptors in astrocytes [247], and it was found to mediate their neuronal activity-dependent metabolic activation and a consequent involvement in sleep and memory. Some neurological diseases are characterized by an excess in excitatory synapses and A_{2B} AR reduces their number in postnatal development by decrease of metabotropic glutamate receptor 5 (mGlu5) expression in astrocytes [248]. On the other hand, A_{2B} AR antagonists were found protective in hippocampus subjected to severe hypoxic/ischemic conditions, but they were not able to counteract glutamate-induced injury, thus revealing a mechanism of action upstream to glutamate release [249]. Coppi *et al.* reviewed the role of A_{2B} AR in brain ischemia and subsequent demyelination injury. It is located on any cell type of the brain and on vascular and blood cells participating in both salvage and damaging of the tissue, including protracted demyelination [250]. These findings point out A_{2B} AR as a possible target in neurodegenerative and neurological diseases. Its low affinity for adenosine makes it particularly intriguing, as it becomes activated only under pathological conditions characterized by elevated extracellular adenosine levels—offering a degree of selectivity in targeting disease states [251]. Thus, they all represent important targets for drugs having different therapeutic time-windows after stroke.

The final protective outcome for an agonist versus antagonist compound depends on the time of administration and district of activation of the receptor targeted by the drug.

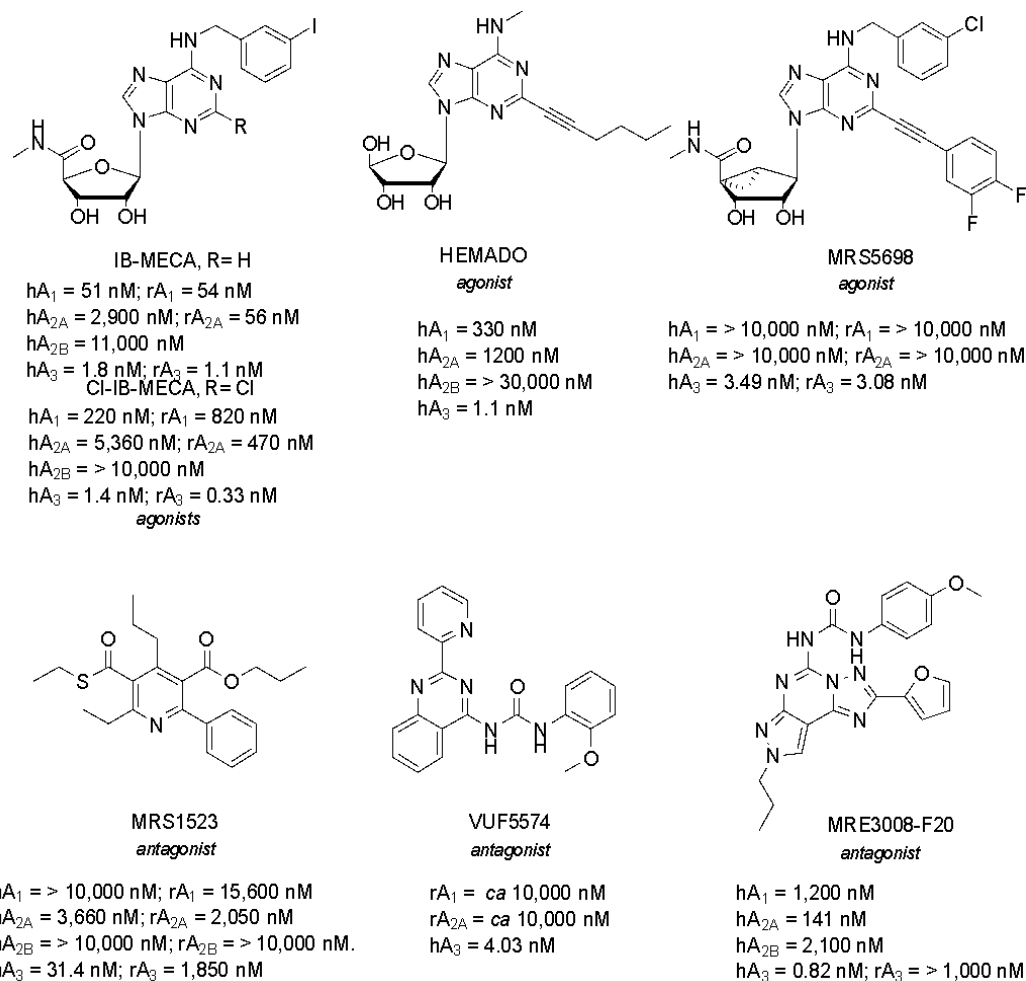
Recently, because literature points out that selective antagonism at both A_{2A} AR and A_{2B} AR delays the anoxic depolarization induced by a severe OGD insult in the hippocampus, Venturini *et al.* investigated a dual A_{2A}/A_{2B} AR antagonist, P626, which showed to be effectively neuroprotective [252]. Later, Dettori *et al.*, used another dual antagonist, MRS3997, on a different model, the rat model of transient middle cerebral artery occlusion (tMCAo), obtaining a reduction in brain ischemic damage and neuroinflammation acting at both astrocytes and microglia, thus combining the neuroprotective effect of both A_{2A} ARs and A_{2B} ARs [253].

Role in neurodegenerative disorders

There are few articles concerning the role of A_{2B} AR and A₃ AR in neurodegenerative disease. Concerning A_{2B} AR, in mice treated with A β to infer a cognitive deficit, A_{2B} AR has been found to be downregulated along with impairment of mitochondrial activity that the same authors reported to be rescued by treatment with an A_{2B} AR agonist. Noteworthy, the agonist used is NECA, that is a well-known pan-agonist for ARs [254,255]. One interesting involvement of A_{2B} AR in neurodegenerative diseases comes from the leaky gut, one of the syndromes related to AD, but also PD [256]. Recently, Ishioh *et al.* shown that intestinal permeability is regulated through basal forebrain cholinergic neurons (BFCN) mediated by A_{2B} AR and the vagal pathway [257–259]. The same group further investigated the signaling pathway finding a key role for brain H₁ receptor signaling [260]. These findings pave the way to new targets to be considered in neurodegenerative diseases, especially for AD.

A₃ AR

The most representative pharmacological tools for the A₃ AR subtype are reported in Figure 8 [84,261,262]. Up to now, no A₃ AR ligand has reached the market, but agonists IB-MECA and CI-IB-MECA are under clinical investigations for cancer and autoimmune diseases [147]. The major challenge in investigating the A₃ AR lies in the significant interspecies differences between the receptor in standard animal models, such as rats and mice, and in humans. Jacobson addressed this issue by developing MRS5698, a compound that is equipotent at both the mouse and human A₃ AR [263]. This makes it a valuable tool for studying the role of A₃ AR at the preclinical level [264]. Concerning A₃ AR antagonists, a huge number of potent and selective compounds have been developed [265], exploring a quite wide chemical space: MRS1523, VUF5574 and MRE3008-F20 represented those frequently used as pharmacological tools for this receptor, with MRS1523 displaying affinity also against mouse and rat A₃ AR, while the other two are completely inactive towards rodent receptors [266].

Figure 8. Structures of the most representative A₃ AR ligands.

Role in neuroinflammation

With respect to A₁ AR and A_{2A} AR, there are few studies concerning A₃ AR involvement in neuroinflammation. Although most studies attribute microglial process retraction to A_{2A} AR activation [162], also A₃ AR was reported as mediator of this process [267]. In contrast, Choi *et al.* showed that A₃ AR can reduce inflammatory cell migration, including that of microglia, during brain ischemic injury [268]. Furthermore, one of the mechanisms underlying A₃ AR-mediated attenuation of neuropathic pain involves inhibition of excessive microglial activation in the spinal cord dorsal horn [269,270]. A recent transcriptomic study investigated the effects of A₃ AR agonists on gene regulation in activated microglia. Although the authors identified several genes whose expression was altered by A₃ AR agonists, only a few showed statistically significant changes. Moreover, it remains difficult to attribute these changes to the receptor's ability to mediate a transition between distinct microglial phenotypes. Therefore, further experiments are necessary to clarify these findings [271]. Repeated episodes of hypoxia on A₃ AR KO mice led to a major vulnerability of hippocampal pyramidal neurons respect to A₃ AR^{+/+} mice, resulting in a decline in cognitive function [272]. Pugliese *et al.* investigated the role of A₃ AR in rat hippocampal slices under severe ischemic conditions. Treatment with antagonists or with an agonist, but for a long time that allows receptor-desensitization, led to a neuroprotective effect [273,274]. The group of Salvemini D. provides support for the beneficial effects of A₃ AR agonists to mitigate secondary tissue injury and cognitive impairment following TBI [275].

The A_{2A} AR and A₃ AR interact to form heteromeric complexes which expression is higher in neurons from the striatum than from the cortex or hippocampus and it is similar in activated or resting microglia, but it has been found up-regulated in an AD model. Between them there is a negative functional cross-talk,

thus A_{2A} AR antagonists make A₃ AR responsive to endogenous adenosine reinforcing the therapeutic interest of A_{2A} AR antagonists to fight AD [278].

Role in neurodegenerative disorders

A specific role of A₃ AR in AD has been ascribed in relation to the generation of intracellular A β . The mechanism involves the cell internalization of the precursor A β PP, with caffeine able to decrease A β PP internalization and, consequently, lowering intracellular levels of A β . The effect of caffeine has been replicated using a selective A₃ AR antagonist, and not by A₁, A_{2A} or A_{2B} ARs antagonists, suggesting that it is A₃ AR the subtype involved [279].

Conclusions

Adenosine receptors play a complex and not yet fully understood role in neuroinflammatory and neurodegenerative processes. Among them, A₁ ARs and A_{2A} ARs are the most prominently involved, affecting both neuronal and glial cells. The intricate cross-talk between these cell types varies not only spatially but also temporally, making it challenging to determine whether agonists or antagonists are more appropriate, and whether a single-target or multitarget approach should be pursued.

An intriguing area for future research is the potential role of biased agonism at these receptors in the pathophysiology of such diseases. Additionally, we reviewed literature on A_{2B} ARs and A₃ ARs, particularly in the context of neuroinflammation and, to a lesser extent, neurodegeneration. These receptors have been largely underexplored in these areas. However, with the development of more effective pharmacological tools, this gap is likely to be addressed in the coming years. This progress may lead to a more comprehensive understanding of adenosine receptor function in central nervous system disorders and potentially pave the way for novel therapeutic strategies for currently unmet medical needs.

Abbreviations

The following abbreviations are used in this manuscript:

α -Syn	α -Synuclein
A β	Amyloid beta
A β PP	Amyloid-beta precursor protein
AD	Alzheimer's disease
ALS	Amyotrophic lateral sclerosis
ALSFRS-R	Revised Amyotrophic Lateral Sclerosis Functional Rating Scale
APP	Amyloid precursor protein
AR	Adenosine receptor
ATP	Adenosine triphosphate
AV	Atrioventricular
BACE1	β -secretase 1
BBB	Blood brain barrier
BDNF	Brain-derived neurotrophic factor
BFCN	Basal forebrain cholinergic neurons
C9orf72	Chromosome 9 open reading frame 72
CCH	Chronic cerebral hypoperfusion
CCL	C-C Motif Chemokine Ligand
CLEC7A	C-Type Lectin Domain Containing 7A
CNS	Central nervous system
COPD	Chronic obstructive pulmonary disease
CREB	cAMP response element-binding protein
DAMP	Danger -associated molecular pattern
DNA	Deoxyribonucleic acid
EL	Extracellular loop
ERK	Extracellular regulated kinase
FDA	Food and drug administration
FUS	Fused in Sarcoma, also called translocated in liposarcoma protein (TLP)

GABA	Gamma-aminobutyric acid
GLT-1	Glutamate transporter 1
GPCR	G protein-coupled receptor
H ₁	Histamine 1
HD	Huntington's disease
IFN	Interferon
IL	Interleukin
JAK	Janus kinase-
JNK	C-Jun N-terminal kinase
KO	Knock out
L-DOPA	Levo-dioxyphenylalanine
LPS	Lipopolysaccharide
LTD	Long term depression
LTP	Long term potentiation
MAPK	Mitogen activated protein kinase
MAO-B	Monoamine oxidase B
mGLU5	Metabotropic glutamate receptor 5
MPI	Myocardial perfusion imaging
MS	Multiple sclerosis
NADPH	Nicotinamide adenine dinucleotide phosphate
NECA	N-ethylcarboxamidoadenosine
NF- κ B	Nuclear factor kappa B
NFT	Neurofibrillary tangle
NLRP3	NOD-, LRR- and pyrin domain-containing protein 3
NMDA	N-methyl-D-aspartate
NO	Nitric oxide
OGD	Oxygen glucose deprivation
PAMP	Pathogen-associated molecular pattern
PD	Parkinson's disease
PLC	Phospholipase C
PKA	Protein kinase A
PKC	Protein kinase C
PRR	Pattern-recognition receptors
PS1	Presenilin-1
RNA	Ribonucleic acid; mRNA (messenger RNA)
ROS	Reactive oxygen species
SAE	Sepsis associated encephalopathy
SAP102	Synapse-associated protein 102
Shh	Sonic hedgehog
SN	Substantia nigra
SOD1	Superoxide dismutase 1
STAT3	Signal transducer and transcription activator 3
SVT	Supraventricular tachycardia
TBI	Traumatic brain injury
TDP-43	Trans activation response DNA binding protein 43
TGF- β	Transforming growth factor β
TLR	Toll-like receptor
tMCAo	Transient middle cerebral artery occlusion
TNF	Tumor Necrosis Factor
TrKB	Tropomyosin receptor kinase B/tyrosine receptor kinase B
tPA	Tissue plasminogen activator
USA	United States of America
VEGF	Vascular endothelial growth factor

References

- Newby, A.C. Adenosine and the Concept of 'Retaliatory Metabolites.' *Trends Biochem Sci* 1984, 9, 42–44, doi:10.1016/0968-0004(84)90176-2.
- Fredholm, B.B. Physiological and Pathophysiological Roles of Adenosine. *Sleep Biol Rhythms* 2011, 9, 24–28, doi:10.1111/J.1479-8425.2010.00460.X.
- Pearson, T.; Nuritova, F.; Caldwell, D.; Dale, N.; Frenguelli, B.G. A Depletable Pool of Adenosine in Area CA1 of the Rat Hippocampus. *Journal of Neuroscience* 2001, 21, 2298–2307, doi:10.1523/JNEUROSCI.21-07-02298.2001.
- Pearson, T.; Damian, K.; Lynas, R.E.; Frenguelli, B.G. Sustained Elevation of Extracellular Adenosine and Activation of A₁ Receptors Underlie the Post-Ischaemic Inhibition of Neuronal Function in Rat Hippocampus in Vitro. *J Neurochem* 2006, 97, 1357–1368, doi:10.1111/J.1471-4159.2006.03823.X;REQUESTEDJOURNAL:JOURNAL:14714159;JOURNAL:JOURNAL:14714159;WGROU:STRING: PUBLICATION.
- Latini, S.; Pedata, F. Adenosine in the Central Nervous System: Release Mechanisms and Extracellular Concentrations. *J Neurochem* 2001, 79, 463–484, doi:10.1046/j.1471-4159.2001.00607.x.
- Zimmermann, H. Extracellular Metabolism of ATP and Other Nucleotides. *Naunyn Schmiedebergs Arch Pharmacol* 2000, 362, 299–309, doi:10.1007/s002100000309.
- Camici, M.; Garcia-Gil, M.; Tozzi, M. The Inside Story of Adenosine. *Int J Mol Sci* 2018, 19, 784, doi:10.3390/ijms19030784.
- Huang, Z.; Xie, N.; Illes, P.; Di Virgilio, F.; Ulrich, H.; Semyanov, A.; Verkhatsky, A.; Sperlagh, B.; Yu, S.-G.; Huang, C.; et al. From Purines to Purinergic Signalling: Molecular Functions and Human Diseases. *Signal Transduct Target Ther* 2021, 6, 162, doi:10.1038/s41392-021-00553-z.
- Pastor-Anglada, M.; Pérez-Torras, S. Who Is Who in Adenosine Transport. *Front Pharmacol* 2018, 9, doi:10.3389/fphar.2018.00627.
- Pándy-Szekeres, G.; Munk, C.; Tsonkov, T.M.; Mordalski, S.; Harpsøe, K.; Hauser, A.S.; Bojarski, A.J.; Gloriam, D.E. GPCRdb in 2018: Adding GPCR Structure Models and Ligands. *Nucleic Acids Res* 2018, 46, D440–D446, doi:10.1093/nar/gkx1109.
- Ballesteros, J.A.; Weinstein, H. [19] Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors. In: 1995; pp. 366–428.
- Seibt, B.F.; Schiedel, A.C.; Thimm, D.; Hinz, S.; Sherbiny, F.F.; Müller, C.E. The Second Extracellular Loop of GPCRs Determines Subtype-Selectivity and Controls Efficacy as Evidenced by Loop Exchange Study at A₂ Adenosine Receptors. *Biochem Pharmacol* 2013, 85, 1317–1329, doi:10.1016/j.bcp.2013.03.005.
- Schiedel, A.C.; Hinz, S.; Thimm, D.; Sherbiny, F.; Borrmann, T.; Maaß, A.; Müller, C.E. The Four Cysteine Residues in the Second Extracellular Loop of the Human Adenosine A_{2B} Receptor: Role in Ligand Binding and Receptor Function. *Biochem Pharmacol* 2011, 82, 389–399, doi:10.1016/j.bcp.2011.05.008.
- De Filippo, E.; Namasivayam, V.; Zappe, L.; El-Tayeb, A.; Schiedel, A.C.; Müller, C.E. Role of Extracellular Cysteine Residues in the Adenosine A_{2A} Receptor. *Purinergic Signal* 2016, 12, 313–329, doi:10.1007/s11302-016-9506-7.
- Fredholm, B.B.; IJzerman, A.P.; Jacobson, K.A.; Linden, J.; Müller, C.E. International Union of Basic and Clinical Pharmacology. LXXXI. Nomenclature and Classification of Adenosine Receptors—An Update. *Pharmacol Rev* 2011, 63, 1–34, doi:10.1124/pr.110.003285.
- Gao, Z.G.; Auchampach, J.A.; Jacobson, K.A. Species Dependence of A₃ Adenosine Receptor Pharmacology and Function. *Purinergic Signal* 2023, 19, 523–550, doi:10.1007/s11302-022-09910-1.
- Fredholm, B.B.; IJzerman, a P.; Jacobson, K. a; Klotz, K.N.; Linden, J. International Union of Pharmacology. XXV. Nomenclature and Classification of Adenosine Receptors. *Pharmacol Rev* 2001, 53, 527–552, doi:10.1124/pr.110.003285.1.
- Rosin, D.L.; Robeva, A.; Woodard, R.L.; Guyenet, P.G.; Linden, J. Immunohistochemical Localization of Adenosine A_{2A} Receptors in the Rat Central Nervous System. *J Comp Neurol* 1998, 401, 163–186, doi:10.1002/(SICI)1096-9861(19981116)401:2<163::AID-CNE2>3.0.CO;2-D.
- Wang, M.; Li, Z.; Song, Y.; Sun, Q.; Deng, L.; Lin, Z.; Zeng, Y.; Qiu, C.; Lin, J.; Guo, H.; et al. Genetic Tagging of the Adenosine A_{2A} Receptor Reveals Its Heterogeneous Expression in Brain Regions. *Front Neuroanat* 2022, 16, 978641, doi:10.3389/FNANA.2022.978641/BIBTEX.
- Jacobson, K.A.; Merighi, S.; Varani, K.; Borea, P.A.; Baraldi, S.; Aghazadeh Tabrizi, M.; Romagnoli, R.; Baraldi, P.G.; Cianchetta, A.; Tosh, D.K.; et al. A₃ Adenosine Receptors as Modulators of Inflammation: From Medicinal Chemistry to Therapy. *Med Res Rev* 2018, 38, 1031–1072, doi:10.1002/med.21456.
- Haskó, G.; Linden, J.; Cronstein, B.; Pacher, P. Adenosine Receptors: Therapeutic Aspects for Inflammatory and Immune Diseases. *Nat Rev Drug Discov* 2008, 7, 759–770, doi:10.1038/nrd2638.
- Ye, H.; Zhao, J.; Xu, X.; Zhang, D.; Shen, H.; Wang, S. Role of Adenosine A_{2A} Receptor in Cancers and Autoimmune Diseases. *Immun Inflamm Dis* 2023, 11, doi:10.1002/iid3.826.
- Sepúlveda, C.; Palomo, I.; Fuentes, E. Role of Adenosine A_{2b} Receptor Overexpression in Tumor Progression. *Life Sci* 2016, 166, 92–99, doi:10.1016/j.lfs.2016.10.008.
- Venugopala, K.N.; Buccioni, M. Current Understanding of the Role of Adenosine Receptors in Cancer. *Molecules* 2024, 29, 3501, doi:10.3390/molecules29153501.
- Chu, S.; Xiong, W.; Zhang, D.; Soylu, H.; Sun, C.; Albensi, B.C.; Parkinson, F.E. Regulation of Adenosine Levels during Cerebral Ischemia. *Acta Pharmacol Sin* 2013, 34, 60–66, doi:10.1038/APS.2012.127;KWRD=BIOMEDICINE.

26. Guo, Q.; Gobbo, D.; Zhao, N.; Zhang, H.; Awuku, N.O.; Liu, Q.; Fang, L.P.; Gampfer, T.M.; Meyer, M.R.; Zhao, R.; et al. Adenosine Triggers Early Astrocyte Reactivity That Provokes Microglial Responses and Drives the Pathogenesis of Sepsis-Associated Encephalopathy in Mice. *Nature Communications* 2024 15:1 2024, 15, 1–19, doi:10.1038/s41467-024-50466-y.
27. Fredholm, B.B.; IJzerman, A.P.; Jacobson, K.A.; Klotz, K.-N.; Linden, J. International Union of Pharmacology. XXV. Nomenclature and Classification of Adenosine Receptors. *Pharmacol Rev* 2001, 53, 527–552, doi:10.1016/S0031-6997(24)01511-4.
28. Borea, P.A.; Gessi, S.; Merighi, S.; Vincenzi, F.; Varani, K. Pharmacology of Adenosine Receptors: The State of the Art. *Physiol Rev* 2018, 98, 1591–1625, doi:10.1152/physrev.00049.2017.
29. Jacobson, K.A.; Gao, Z.-G. Adenosine Receptors as Therapeutic Targets. *Nat Rev Drug Discov* 2006, 5, 247–264, doi:10.1038/nrd1983.
30. Elson, G.; Eisenberg, M.; Garg, C.; Outram, S.; Ferrante, C.J.; Hasko, G.; Leibovich, S.J. Induction of Murine Adenosine A_{2A} Receptor Expression by LPS: Analysis of the 5' Upstream Promoter. *Genes Immun* 2013, 14, 147–153, doi:10.1038/GENE.2012.60;SUBJMETA.
31. Sereda, M.J.; Bradding, P.; Vial, C. Adenosine Potentiates Human Lung Mast Cell Tissue Plasminogen Activator Activity. *The Journal of Immunology* 2011, 186, 1209–1217, doi:10.4049/JIMMUNOL.1001563.
32. Chen, J.-F.; Eltzschig, H.K.; Fredholm, B.B. Adenosine Receptors as Drug Targets — What Are the Challenges? *Nat Rev Drug Discov* 2013, 12, 265–286, doi:10.1038/nrd3955.
33. Kutryb-Zajac, B.; Kawecka, A.; Nasadiuk, K.; Braczko, A.; Stawarska, K.; Caiazzo, E.; Koszałka, P.; Cicala, C. Drugs Targeting Adenosine Signaling Pathways: A Current View. *Biomedicine & Pharmacotherapy* 2023, 165, 115184, doi:10.1016/j.biopha.2023.115184.
34. Rankin, A.C.; Brooks, R.; Ruskin, J.N.; McGovern, B.A. Adenosine and the Treatment of Supraventricular Tachycardia. *Am J Med* 1992, 92, 655–664, doi:10.1016/0002-9343(92)90784-9.
35. Derry, C.J.; Derry, S.; Moore, R.A. Caffeine as an Analgesic Adjuvant for Acute Pain in Adults. *Cochrane Database Syst Rev* 2014, 2014, CD009281, doi:10.1002/14651858.CD009281.PUB3.
36. Spina, D.; Page, C.P. Xanthines and Phosphodiesterase Inhibitors. *Handb Exp Pharmacol* 2016, 237, 63–91, doi:10.1007/164_2016_71.
37. da Fonseca, A.C.C.; Matias, D.; Garcia, C.; Amaral, R.; Geraldo, L.H.; Freitas, C.; Lima, F.R.S. The Impact of Microglial Activation on Blood-Brain Barrier in Brain Diseases. *Front Cell Neurosci* 2014, 8, doi:10.3389/fncel.2014.00362.
38. Adamu, A.; Li, S.; Gao, F.; Xue, G. The Role of Neuroinflammation in Neurodegenerative Diseases: Current Understanding and Future Therapeutic Targets. *Front Aging Neurosci* 2024, 16, doi:10.3389/fnagi.2024.1347987.
39. DiSabato, D.J.; Quan, N.; Godbout, J.P. Neuroinflammation: The Devil Is in the Details. *J Neurochem* 2016, 139, 136–153, doi:10.1111/jnc.13607.
40. Sha, S.; Tan, J.; Miao, Y.; Zhang, Q. The Role of Autophagy in Hypoxia-Induced Neuroinflammation. *DNA Cell Biol* 2021, 40, 733–739, doi:10.1089/DNA.2020.6186;REQUESTEDJOURNAL:JOURNAL:DNA;ISSUE:ISSUE:10.1089/DNA.2021.40.ISSUE-6;WGROU:STRING:PUBLICATION.
41. Zhao, Y.; Gong, C.X. From Chronic Cerebral Hypoperfusion to Alzheimer-Like Brain Pathology and Neurodegeneration. *Cell Mol Neurobiol* 2015, 35, 101–110, doi:10.1007/S10571-014-0127-9/FIGURES/1.
42. Ciacciarelli, A.; Sette, G.; Giubilei, F.; Orzi, F. Chronic Cerebral Hypoperfusion: An Undefined, Relevant Entity. *Journal of Clinical Neuroscience* 2020, 73, 8–12, doi:10.1016/J.JOCN.2020.01.026.
43. Breteler, M.M.B.; van Swieten, J.C.; Bots, M.L.; Grobbee, D.E.; Claus, J.J.; van den Hout, J.H.W.; van Harskamp, F.; Tanghe, H.L.J.; de Jong, P.T.V.M.; van Gijn, J.; et al. Cerebral White Matter Lesions, Vascular Risk Factors, and Cognitive Function in a Population-based Study. *Neurology* 1994, 44, 1246–1246, doi:10.1212/WNL.44.7.1246.
44. Jaakkola, P.; Mole, D.R.; Tian, Y.M.; Wilson, M.I.; Giibert, J.; Gaskell, S.J.; Von Kriegsheim, A.; Hebestreit, H.F.; Mukherji, M.; Schofield, C.J.; et al. Targeting of HIF- α to the von Hippel-Lindau Ubiquitylation Complex by O₂-Regulated Prolyl Hydroxylation. *Science (1979)* 2001, 292, 468–472, doi:10.1126/SCIENCE.1059796/ASSET/86437EBB-F402-4333-B056-3641116447B5/ASSETS/GRAPHIC/SE1519354005.JPEG.
45. Ohh, M.; Park, C.W.; Ivan, M.; Hoffman, M.A.; Kim, T.Y.; Huang, L.E.; Pavletich, N.; Chau, V.; Kaelin, W.G. Ubiquitination of Hypoxia-Inducible Factor Requires Direct Binding to the β -Domain of the von Hippel - Lindau Protein. *Nat Cell Biol* 2000, 2, 423–427, doi:10.1038/35017054;KWRD=LIFE+SCIENCES.
46. Semenza, G.L.; Neifelt, M.K.; Chi, S.M.; Antonarakis, S.E. Hypoxia-Inducible Nuclear Factors Bind to an Enhancer Element Located 3' to the Human Erythropoietin Gene. *Proc Natl Acad Sci U S A* 1991, 88, 5680–5684, doi:10.1073/PNAS.88.13.5680;WGROU:STRING:PUBLICATION.
47. Bowser, J.L.; Lee, J.W.; Yuan, X.; Eltzschig, H.K. The Hypoxia-Adenosine Link during Inflammation. *J Appl Physiol* 2017, 123, 1303–1320, doi:10.1152/JAPPLPHYSIOL.00101.2017/ASSET/IMAGES/LARGE/ZDG0101723690003.JPEG.
48. Stockwell, J.; Jakova, E.; Cayabyab, F.S. Adenosine A₁ and A_{2A} Receptors in the Brain: Current Research and Their Role in Neurodegeneration. *Molecules* 2017, Vol. 22, Page 676 2017, 22, 676, doi:10.3390/MOLECULES22040676.
49. Colonna, M.; Butovsky, O. Microglia Function in the Central Nervous System During Health and Neurodegeneration. *Annu Rev Immunol* 2017, 35, 441–468, doi:10.1146/annurev-immunol-051116-052358.
50. Guo, H.; Callaway, J.B.; Ting, J.P.-Y. Inflammasomes: Mechanism of Action, Role in Disease, and Therapeutics. *Nat Med* 2015, 21, 677–687, doi:10.1038/nm.3893.

51. Zhang, W.; Xiao, D.; Mao, Q.; Xia, H. Role of Neuroinflammation in Neurodegeneration Development. *Signal Transduct Target Ther* 2023, *8*, 267, doi:10.1038/s41392-023-01486-5.
52. Ransohoff, R.M. A Polarizing Question: Do M1 and M2 Microglia Exist? *Nat Neurosci* 2016, *19*, 987–991, doi:10.1038/nn.4338.
53. Giovannoni, F.; Quintana, F.J. The Role of Astrocytes in CNS Inflammation. *Trends Immunol* 2020, *41*, 805–819, doi:10.1016/j.it.2020.07.007.
54. Verkhatsky, A.; Pivoriūnas, A. Astroglia Support, Regulate and Reinforce Brain Barriers. *Neurobiol Dis* 2023, *179*, 106054, doi:10.1016/j.nbd.2023.106054.
55. Fakhoury, M. Role of Immunity and Inflammation in the Pathophysiology of Neurodegenerative Diseases. *Neurodegener Dis* 2015, *15*, 63–69, doi:10.1159/000369933.
56. Kurtishi, A.; Rosen, B.; Patil, K.S.; Alves, G.W.; Møller, S.G. Cellular Proteostasis in Neurodegeneration. *Mol Neurobiol* 2019, *56*, 3676–3689, doi:10.1007/S12035-018-1334-Z/FIGURES/3.
57. Sheppard, O.; Coleman, M. Alzheimer's Disease: Etiology, Neuropathology and Pathogenesis. *Alzheimer's Disease: Drug Discovery* 2020, 1–22, doi:10.36255/EXONPUBLICATIONS.ALZHEIMERSDISEASE.2020.CH1.
58. Weller, J.; Budson, A. Current Understanding of Alzheimer's Disease Diagnosis and Treatment. *F1000Res* 2018, *7*, 1–9, doi:10.12688/f1000research.14506.1.
59. Rios-Romenets, S.; Lopera, F.; Sink, K.M.; Hu, N.; Lian, Q.; Guthrie, H.; Smith, J.; Cho, W.; Mackey, H.; Langbaum, J.B.; et al. Baseline Demographic, Clinical, and Cognitive Characteristics of the Alzheimer's Prevention Initiative (API) Autosomal-Dominant Alzheimer's Disease Colombia Trial. *Alzheimer's and Dementia* 2020, *16*, 1023–1030, doi:10.1002/alz.12109.
60. Decourt, B.; D'Souza, G.X.; Shi, J.; Ritter, A.; Suazo, J.; Sabbagh, M.N. The Cause of Alzheimer's Disease: The Theory of Multipathology Convergence to Chronic Neuronal Stress. *Aging Dis* 2022, *13*, 37, doi:10.14336/AD.2021.0529.
61. Zhang, J.; Zhang, Y.; Wang, J.; Xia, Y.; Zhang, J.; Chen, L. Recent Advances in Alzheimer's Disease: Mechanisms, Clinical Trials and New Drug Development Strategies. *Signal Transduction and Targeted Therapy* 2024 9:1 2024, *9*, 1–35, doi:10.1038/s41392-024-01911-3.
62. Blennow, K.; de Leon, M.J.; Zetterberg, H. Alzheimer's Disease. *Lancet* 2006, *368*, 387–403, doi:10.1016/S0140-6736(06)69113-7.
63. Kaur, D.; Sharma, V.; Deshmukh, R. Activation of Microglia and Astrocytes: A Roadway to Neuroinflammation and Alzheimer's Disease. *Inflammopharmacology* 2019 27:4 2019, *27*, 663–677, doi:10.1007/S10787-019-00580-X.
64. Lopes, C.R.; Cunha, R.A.; Agostinho, P. Astrocytes and Adenosine A2A Receptors: Active Players in Alzheimer's Disease. *Front Neurosci* 2021, *15*, 666710, doi:10.3389/FNINS.2021.666710/XML/NLM.
65. Gholami, A.; Mortezaee, K. Neurotransmitters in Neural Circuits and Neurological Diseases. *ACS Chem Neurosci* 2025, doi:10.1021/ACSCHMNEURO.5C00426.
66. Olanow, C.W.; Stern, M.B.; Sethi, K. The Scientific and Clinical Basis for the Treatment of Parkinson Disease (2009). *Neurology* 2009, *72*, doi:10.1212/WNL.0B013E3181A1D44C/ASSET/8D6884A6-DD41-48A8-A27B-6B20E3608EE4/ASSETS/GRAPHIC/1TT26.JPEG.
67. McGregor, M.M.; Nelson, A.B. Circuit Mechanisms of Parkinson's Disease. *Neuron* 2019, *101*, 1042–1056, doi:10.1016/J.NEURON.2019.03.004.
68. Kalia, L. V.; Lang, A.E.; Shulman, G. Parkinson ' s Disease. *The Lancet* 2015, *386*, 896–912, doi:10.1016/S0140-6736(14)61393-3.
69. Tenreiro, S.; Eckermann, K.; Outeiro, T.F. Protein Phosphorylation in Neurodegeneration: Friend or Foe? *Front Mol Neurosci* 2014, *7*, 1–30, doi:10.3389/fnmol.2014.00042.
70. Li, D.W.; Liu, Z.Q.; Chen, W.; Yao, M.; Li, G.R. Association of Glycogen Synthase Kinase-3 β with Parkinson's Disease (Review). *Mol Med Rep* 2014, *9*, 2043–2050, doi:10.3892/mmr.2014.2080.
71. Leak, R.K.; Clark, R.N.; Abbas, M.; Xu, F.; Brodsky, J.L.; Chen, J.; Hu, X.; Luk, K.C. Current Insights and Assumptions on α -Synuclein in Lewy Body Disease. *Acta Neuropathol* 2024, *148*, 1–21, doi:10.1007/S00401-024-02781-3/FIGURES/1.
72. Bono, N.; Fruzzetti, F.; Farinazzo, G.; Candiani, G.; Marcuzzo, S. Perspectives in Amyotrophic Lateral Sclerosis: Biomarkers, Omics, and Gene Therapy Informing Disease and Treatment. *International Journal of Molecular Sciences* 2025, Vol. 26, Page 5671 2025, *26*, 5671, doi:10.3390/IJMS26125671.
73. Koski, L.; Ronnevi, C.; Berntsson, E.; Wärmländer, S.K.T.S.; Roos, P.M. Metals in ALS TDP-43 Pathology. *International Journal of Molecular Sciences* 2021, Vol. 22, Page 12193 2021, *22*, 12193, doi:10.3390/IJMS222212193.
74. Jo, M.; Lee, S.; Jeon, Y.M.; Kim, S.; Kwon, Y.; Kim, H.J. The Role of TDP-43 Propagation in Neurodegenerative Diseases: Integrating Insights from Clinical and Experimental Studies. *Exp Mol Med* 2020, *52*, 1652–1662, doi:10.1038/s12276-020-00513-7.
75. Sreedharan, J.; Blair, I.P.; Tripathi, V.B.; Hu, X.; Vance, C.; Rogelj, B.; Ackerley, S.; Durnall, J.C.; Williams, K.L.; Buratti, E.; et al. TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Science (1979)* 2008, *319*, 1668–1672.
76. Suk, T.R.; Rousseaux, M.W.C. The Role of TDP-43 Mislocalization in Amyotrophic Lateral Sclerosis. *Mol Neurodegener* 2020, *15*, 1–16, doi:10.1186/s13024-020-00397-1.
77. Gasset-rosa, F.; Lu, S.; Yu, H.; Shorter, J.; Cruz, S. Da; Cleveland, D.W.; Gasset-rosa, F.; Lu, S.; Yu, H.; Chen, C.; et al. Cytoplasmic TDP-43 De-Mixing Independent of Stress Granules Drives Inhibition of Nuclear Import , Loss of Nuclear TDP-43 , and Cell Death. *Neuron* 2019, *102*, 339–357, doi:10.1016/j.neuron.2019.02.038.
78. Fernandes, N.; Nero, L.; Lyons, S.M.; Ivanov, P.; Mittelmeier, T.M.; Bolger, T.A.; Buchan, J.R. Stress Granule Assembly Can Facilitate but Is Not Required for TDP-43 Cytoplasmic Aggregation. 2020, *10*, 1–20.

79. Chen, Youjun and Cohen, T.J. Aggregation of the Nucleic Acid-Binding Protein TDP-43 Occurs via Distinct Routes That Are Coordinated with Stress Granule Formation. *J Biol Chem* 2019, *294*, 3696–3706, doi:10.1074/jbc.RA118.006351.
80. Philips, T.; Rothstein, J.D. Glial Cells in Amyotrophic Lateral Sclerosis. *Exp Neurol* 2014, *262*, 111–120, doi:10.1016/J.EXPNEUROL.2014.05.015.
81. Stoker, T.B.; Mason, S.L.; Greenland, J.C.; Holden, S.T.; Santini, H.; Barker, R.A. Huntington's Disease: Diagnosis and Management. *Pract Neurol* 2022, *22*, 32–41, doi:10.1136/PRACTNEUROL-2021-003074.
82. McColgan, P.; Tabrizi, S.J. Huntington's Disease: A Clinical Review. *Eur J Neurol* 2018, *25*, 24–34, doi:10.1111/ENE.13413;SUBPAGE:STRING:FULL.
83. Johnston, J.B.; Silva, C.; Gonzalez, G.; Holden, J.; Warren, K.G.; Metz, L.M.; Power, C. Diminished Adenosine A1 Receptor Expression on Macrophages in Brain and Blood of Patients with Multiple Sclerosis. *Ann Neurol* 2001, *49*, 650–658, doi:10.1002/ANA.1007;WGROU:STRING:PUBLICATION.
84. Fredholm, B.B.; IJzerman, A.P.; Jacobson, K.A.; Linden, J.; Müller, C.E. International Union of Basic and Clinical Pharmacology. LXXXI. Nomenclature and Classification of Adenosine Receptors—An Update. *Pharmacol Rev* 2011, *63*, 1–34, doi:10.1124/pr.110.003285.
85. Ribeiro, F.F.; Xapelli, S.; Miranda-Lourenço, C.; Tanqueiro, S.R.; Fonseca-Gomes, J.; Diógenes, M.J.; Ribeiro, J.A.; Sebastião, A.M. Purine Nucleosides in Neuroregeneration and Neuroprotection. *Neuropharmacology* 2016, *104*, 226–242, doi:10.1016/J.NEUROPHARM.2015.11.006.
86. Haskó, G.; Pacher, P.; Vizi, E.S.; Illes, P. Adenosine Receptor Signaling in the Brain Immune System. *Trends Pharmacol Sci* 2005, *26*, 511–516, doi:10.1016/J.TIPS.2005.08.004.
87. Beamer, E.; Göllöncsér, F.; Horváth, G.; Beko, K.; Otrókoci, L.; Koványi, B.; Sperlágh, B. Purinergic Mechanisms in Neuroinflammation: An Update from Molecules to Behavior. *Neuropharmacology* 2016, *104*, 94–104, doi:10.1016/J.NEUROPHARM.2015.09.019.
88. Tsutsui, S.; Schnermann, J.; Noorbakhsh, F.; Henry, S.; Yong, V.W.; Winston, B.W.; Warren, K.; Power, C. A1 Adenosine Receptor Upregulation and Activation Attenuates Neuroinflammation and Demyelination in a Model of Multiple Sclerosis. *Journal of Neuroscience* 2004, *24*, 1521–1529, doi:10.1523/JNEUROSCI.4271-03.2004.
89. Illes, P.; Rubini, P.; Ulrich, H.; Zhao, Y.; Tang, Y. Regulation of Microglial Functions by Purinergic Mechanisms in the Healthy and Diseased CNS. *Cells* 2020, Vol. 9, Page 1108 2020, *9*, 1108, doi:10.3390/CELLS9051108.
90. Haselkorn, M.L.; Shellington, D.K.; Jackson, E.K.; Vagni, V.A.; Janesko-Feldman, K.; Dubey, R.K.; Gillespie, D.G.; Cheng, D.; Bell, M.J.; Jenkins, L.W.; et al. Adenosine A1 Receptor Activation as a Brake on the Microglial Response after Experimental Traumatic Brain Injury in Mice. *J Neurotrauma* 2010, *27*, 901, doi:10.1089/NEU.2009.1075.
91. Synowitz, M.; Glass, R.; Färber, K.; Markovic, D.; Kronenberg, G.; Herrmann, K.; Schnermann, J.; Nolte, C.; Van Rooijen, N.; Kiwit, J.; et al. A1 Adenosine Receptors in Microglia Control Glioblastoma-Host Interaction. *Cancer Res* 2006, *66*, 8550–8557, doi:10.1158/0008-5472.CAN-06-0365.
92. Luongo, L.; Guida, F.; Imperatore, R.; Napolitano, F.; Gatta, L.; Cristino, L.; Giordano, C.; Siniscalco, D.; Di Marzo, V.; Bellini, G.; et al. The A1 Adenosine Receptor as a New Player in Microglia Physiology. *Glia* 2014, *62*, 122–132, doi:10.1002/GLIA.22592;JOURNAL:JOURNAL:10981136;WGROU:STRING:PUBLICATION.
93. Liu, G.; Zhang, W.; Guo, J.; Kong, F.; Zhou, S.; Chen, S.; Wang, Z.; Zang, D. Adenosine Binds Predominantly to Adenosine Receptor A1 Subtype in Astrocytes and Mediates an Immunosuppressive Effect. *Brain Res* 2018, *1700*, 47–55, doi:10.1016/J.BRAINRES.2018.06.021.
94. Guo, Q.; Gobbo, D.; Zhao, N.; Zhang, H.; Awuku, N.O.; Liu, Q.; Fang, L.P.; Gampfer, T.M.; Meyer, M.R.; Zhao, R.; et al. Adenosine Triggers Early Astrocyte Reactivity That Provokes Microglial Responses and Drives the Pathogenesis of Sepsis-Associated Encephalopathy in Mice. *Nature Communications* 2024 15:1 2024, *15*, 1–19, doi:10.1038/s41467-024-50466-y.
95. Liu, S. Le; Tang, Y. Astrocytic Adenosine A1 Receptors: A New Potential Target for Treating Sepsis-Associated Encephalopathy. *Purinergic Signal* 2024, 1–3, doi:10.1007/S11302-024-10049-4/METRICS.
96. De Mendonça, A.; Sebastião, A.M.; Ribeiro, J.A. Adenosine: Does It Have a Neuroprotective Role after All? *Brain Res Rev* 2000, *33*, 258–274, doi:10.1016/S0165-0173(00)00033-3.
97. Guillemin, K.; Krasnow, M.A. The Hypoxic Response: Huffing and HIFing. *Cell* 1997, *89*, 9–12, doi:10.1016/S0092-8674(00)80176-2.
98. Hines, D.J.; Schmitt, L.I.; Hines, R.M.; Moss, S.J.; Haydon, P.G. Antidepressant Effects of Sleep Deprivation Require Astrocyte-Dependent Adenosine Mediated Signaling. *Translational Psychiatry* 2013 3:1 2013, *3*, e212–e212, doi:10.1038/tp.2012.136.
99. Beamer, E.; Göllöncsér, F.; Horváth, G.; Beko, K.; Otrókoci, L.; Koványi, B.; Sperlágh, B. Purinergic Mechanisms in Neuroinflammation: An Update from Molecules to Behavior. *Neuropharmacology* 2016, *104*, 94–104, doi:10.1016/J.NEUROPHARM.2015.09.019.
100. Kaster, M.P.; Budni, J.; Santos, A.R.S.; Rodrigues, A.L.S. Pharmacological Evidence for the Involvement of the Opioid System in the Antidepressant-like Effect of Adenosine in the Mouse Forced Swimming Test. *Eur J Pharmacol* 2007, *576*, 91–98, doi:10.1016/j.ejphar.2007.08.026.
101. Nguyen, A.T.N.; Tran, Q.L.; Baltos, J.A.; McNeill, S.M.; Nguyen, D.T.N.; May, L.T. Small Molecule Allosteric Modulation of the Adenosine A1 Receptor. *Front Endocrinol (Lausanne)* 2023, *14*, 1184360, doi:10.3389/FENDO.2023.1184360.
102. Albrecht-Küpper, B.E.; Leineweber, K.; Nell, P.G. Partial Adenosine A1 Receptor Agonists for Cardiovascular Therapies. *Purinergic Signal* 2012, *8*, 91–99, doi:10.1007/S11302-011-9274-3/FIGURES/7.

103. Lewerenz, J.; Maher, P. Chronic Glutamate Toxicity in Neurodegenerative Diseases-What Is the Evidence? *Front Neurosci* 2015, 9, 170294, doi:10.3389/FNINS.2015.00469/XML/NLM.
104. Juvenal, G.; Higa, G.S.V.; Bonfim Marques, L.; Tessari Zampieri, T.; Costa Viana, F.J.; Britto, L.R.; Tang, Y.; Illes, P.; di Virgilio, F.; Ulrich, H.; et al. Regulation of GABAergic Neurotransmission by Purinergic Receptors in Brain Physiology and Disease. *Purinergic Signalling* 2024 21:1 2024, 21, 149–177, doi:10.1007/S11302-024-10034-X.
105. Boison, D. Adenosine Kinase: Exploitation for Therapeutic Gain. *Pharmacol Rev* 2013, 65, 906–943, doi:10.1124/PR.112.006361.
106. Ribeiro, J.A.; Sebastião, A.M.; De Mendonça, A. Adenosine Receptors in the Nervous System: Pathophysiological Implications. *Prog Neurobiol* 2002, 68, 377–392, doi:10.1016/S0301-0082(02)00155-7.
107. Dunwiddie, T. V.; Masino, S.A. The Role and Regulation of Adenosine in the Central Nervous System. *Annu Rev Neurosci* 2001, 24, 31–55, doi:10.1146/ANNUREV.NEURO.24.1.31/CITE/REFWORKS.
108. von LUBITZ, D.K.J.E.; CARTER, M.F.; BEENHAKKER, M.; LIN, R.C.; JACOBSON, K.A. Adenosine: A Prototherapeutic Concept in Neurodegeneration. *Ann N Y Acad Sci* 1995, 765, 163–178, doi:10.1111/J.1749-6632.1995.TB16573.X.
109. Fredholm, B.B. Chapter 11 Adenosine and Neuroprotection. *Int Rev Neurobiol* 1996, 40, 259–280, doi:10.1016/S0074-7742(08)60723-0.
110. Lopes, C.R.; Lourenço, V.S.; Tomé, Â.R.; Cunha, R.A.; Canas, P.M. Use of Knockout Mice to Explore CNS Effects of Adenosine. *Biochem Pharmacol* 2021, 187, 114367, doi:10.1016/J.BCP.2020.114367.
111. Daval, J.L.; Nicolas, F. Non-Selective Effects of Adenosine A1 Receptor Ligands on Energy Metabolism and Macromolecular Biosynthesis in Cultured Central Neurons. *Biochem Pharmacol* 1998, 55, 141–149, doi:10.1016/S0006-2952(97)00378-X.
112. Häberg, A.; Qu, H.; Haraldseth, O.; Unsgård, G.; Sonnewald, U. In Vivo Effects of Adenosine A1 Receptor Agonist and Antagonist on Neuronal and Astrocytic Intermediary Metabolism Studied with Ex Vivo ¹³C NMR Spectroscopy. *J Neurochem* 2000, 74, 327–333, doi:10.1046/J.1471-4159.2000.0740327.X.
113. Canals, S.; Larrosa, B.; Pintor, J.; Mena, M.A.; Herreras, O. Metabolic Challenge to Glia Activates an Adenosine-Mediated Safety Mechanism That Promotes Neuronal Survival by Delaying the Onset of Spreading Depression Waves. *Journal of Cerebral Blood Flow and Metabolism* 2008, 28, 1835–1844, doi:10.1038/JCBFM.2008.71/ASSET/53E09ACE-12BB-4CD5-8EA9-C8875A382B35/ASSETS/IMAGES/LARGE/10.1038_JCBFM.2008.71-FIG5.JPG.
114. Duarte, J.M.N.; Cunha, R.A.; Carvalho, R.A. Adenosine A1 Receptors Control the Metabolic Recovery after Hypoxia in Rat Hippocampal Slices. *J Neurochem* 2016, 136, 947–957, doi:10.1111/JNC.13512;PAGEGROUP:STRING:PUBLICATION.
115. Stockwell, J.; Jakova, E.; Cayabyab, F.S. Adenosine A1 and A2A Receptors in the Brain: Current Research and Their Role in Neurodegeneration. *Molecules* 2017, Vol. 22, Page 676 2017, 22, 676, doi:10.3390/MOLECULES22040676.
116. Chang, C.P.; Wu, K.C.; Lin, C.Y.; Chern, Y. Emerging Roles of Dysregulated Adenosine Homeostasis in Brain Disorders with a Specific Focus on Neurodegenerative Diseases. *Journal of Biomedical Science* 2021 28:1 2021, 28, 1–25, doi:10.1186/S12929-021-00766-Y.
117. Nehlig, A. Effects of Coffee/Caffeine on Brain Health and Disease: What Should I Tell My Patients? *Pract Neurol* 2016, 16, 89–95, doi:10.1136/PRACTNEUROL-2015-001162.
118. Johansson, B.; Georgiev, V.; Lindström, K.; Fredholm, B.B. A1 and A2A Adenosine Receptors and A1 mRNA in Mouse Brain: Effect of Long-Term Caffeine Treatment. *Brain Res* 1997, 762, 153–164, doi:10.1016/S0006-8993(97)00378-8.
119. Park, S.W.; Chen, S.W.C.; Kim, M.; D'Agati, V.D.; Lee, H.T. Selective Intrarenal Human A1 Adenosine Receptor Overexpression Reduces Acute Liver and Kidney Injury after Hepatic Ischemia Reperfusion in Mice. *Laboratory Investigation* 2010 90:3 2010, 90, 476–495, doi:10.1038/labinvest.2009.143.
120. De Mendonça, A.; Sebastião, A.M.; Ribeiro, J.A. Adenosine: Does It Have a Neuroprotective Role after All? *Brain Res Rev* 2000, 33, 258–274, doi:10.1016/S0165-0173(00)00033-3.
121. Cunha, R.A. How Does Adenosine Control Neuronal Dysfunction and Neurodegeneration? *J Neurochem* 2016, 139, 1019–1055, doi:10.1111/jnc.13724.
122. Coelho, J.E.; Rebola, N.; Fragata, I.; Ribeiro, J.A.; De Mendonça, A.; Cunha, R.A. Hypoxia-Induced Desensitization and Internalization of Adenosine A1 Receptors in the Rat Hippocampus. *Neuroscience* 2006, 138, 1195–1203, doi:10.1016/J.NEUROSCIENCE.2005.12.012.
123. Baines, A.E.; Corrêa, S.A.L.; Irving, A.J.; Frenguelli, B.G. Differential Trafficking of Adenosine Receptors in Hippocampal Neurons Monitored Using GFP- and Super-Ecliptic PHluorin-Tagged Receptors. *Neuropharmacology* 2011, 61, 1–11, doi:10.1016/J.NEUROPHARM.2011.02.005.
124. Plamondon, H.; Blondeau, N.; Heurteaux, C.; Lazdunski, M. Mutually Protective Actions of Kainic Acid Epileptic Preconditioning and Sublethal Global Ischemia on Hippocampal Neuronal Death: Involvement of Adenosine A1 Receptors and K(ATP) Channels. *Journal of Cerebral Blood Flow and Metabolism* 1999, 19, 1296–1308, doi:10.1097/00004647-199912000-00002/ASSET/DCC1EB12-A64C-4EB6-8D40-767A7D2B7CF8/ASSETS/IMAGES/LARGE/10.1097_00004647-199912000-00002-FIG5.JPG.
125. Fredholm, B.B.; Chen, J.F.; Cunha, R.A.; Svenningsson, P.; Vaugeois, J.M. Adenosine and Brain Function. *Int Rev Neurobiol* 2005, 63, 191–270, doi:10.1016/S0074-7742(05)63007-3.
126. Trendelenburg, G.; Dirnagl, U. Neuroprotective Role of Astrocytes in Cerebral Ischemia: Focus on Ischemic Preconditioning. *Glia* 2005, 50, 307–320, doi:10.1002/GLIA.20204.

127. Ulaş, J.; Brunner, L.C.; Nguyen, L.; Cotman, C.W. Reduced Density of Adenosine A1 Receptors and Preserved Coupling of Adenosine A1 Receptors to G Proteins in Alzheimer Hippocampus: A Quantitative Autoradiographic Study. *Neuroscience* 1993, *52*, 843–854, doi:10.1016/0306-4522(93)90533-L.
128. Kalaria, R.N.; Sromek, S.; Wilcox, B.J.; Unnerstall, J.R. Hippocampal Adenosine A1 Receptors Are Decreased in Alzheimer's Disease. *Neurosci Lett* 1990, *118*, 257–260, doi:10.1016/0304-3940(90)90641-L.
129. Jansen, K.L.R.; Faull, R.L.M.; Dragunow, M.; Synek, B.L. Alzheimer's Disease: Changes in Hippocampal N-Methyl-d-Aspartate, Quisqualate, Neurotensin, Adenosine, Benzodiazepine, Serotonin and Opioid Receptors—an Autoradiographic Study. *Neuroscience* 1990, *39*, 613–627, doi:10.1016/0306-4522(90)90246-Z.
130. Braak, H.; Braak, E. Evolution of the Neuropathology of Alzheimer's Disease. *Acta Neurol Scand* 1996, *94*, 3–12, doi:10.1111/J.1600-0404.1996.TB05866.X;WGROU:STRING:PUBLICATION.
131. Deckert, J.; Abel, F.; Küning, G.; Hartmann, J.; Senitz, D.; Maier, H.; Ransmayr, G.; Riederer, P. Loss of Human Hippocampal Adenosine A1 Receptors in Dementia: Evidence for Lack of Specificity. *Neurosci Lett* 1998, *244*, 1–4, doi:10.1016/S0304-3940(98)00108-6.
132. Angulo, E.; Casadó, V.; Mallol, J.; Canela, E.I.; Viñals, F.; Ferrer, I.; Lluís, C.; Franco, R. A1 Adenosine Receptors Accumulate in Neurodegenerative Structures in Alzheimer Disease and Mediate Both Amyloid Precursor Protein Processing and Tau Phosphorylation and Translocation. *Brain Pathology* 2003, *13*, 440–451, doi:10.1111/J.1750-3639.2003.TB00475.X.
133. Stone, T.W.; Ceruti, S.; Abbracchio, M.P. Adenosine Receptors and Neurological Disease: Neuroprotection and Neurodegeneration. *Handb Exp Pharmacol* 2009, *193*, 535–587, doi:10.1007/978-3-540-89615-9_17.
134. Burnstock, G. An Introduction to the Roles of Purinergic Signalling in Neurodegeneration, Neuroprotection and Neuroregeneration. *Neuropharmacology* 2016, *104*, 4–17, doi:10.1016/J.NEUROPHARM.2015.05.031.
135. Albasanz, J.L.; Perez, S.; Barrachina, M.; Ferrer, I.; Martín, M. RESEARCH ARTICLE: Up-Regulation of Adenosine Receptors in the Frontal Cortex in Alzheimer's Disease. *Brain Pathology* 2008, *18*, 211–219, doi:10.1111/J.1750-3639.2007.00112.X.
136. Jaber, E.; Rohani, M.; Shahidi, G.A.; Nafissi, S.; Arefian, E.; Soleimani, M.; Moghadam, A.; Arzenani, M.K.; Keramatian, F.; Klotzle, B.; et al. Mutation in ADORA1 Identified as Likely Cause of Early-Onset Parkinsonism and Cognitive Dysfunction. *Movement Disorders* 2016, *31*, 1004–1011, doi:10.1002/MDS.26627;CTYPE:STRING:JOURNAL.
137. Roberts, B.M.; Lambert, E.; Livesey, J.A.; Wu, Z.; Li, Y.; Cragg, S.J. Dopamine Release in Nucleus Accumbens Is under Tonic Inhibition by Adenosine A1 Receptors Regulated by Astrocytic ENT1 and Dysregulated by Ethanol. *Journal of Neuroscience* 2022, *42*, 1738–1751, doi:10.1523/JNEUROSCI.1548-21.2021.
138. Zhao, Y.; Liu, X.; Yang, G. Adenosinergic Pathway in Parkinson's Disease: Recent Advances and Therapeutic Perspective. *Molecular Neurobiology* 2023 *60:6* 2023, *60*, 3054–3070, doi:10.1007/S12035-023-03257-3.
139. Ferré, S.; Fredholm, B.B.; Morelli, M.; Popoli, P.; Fuxe, K. Adenosine–Dopamine Receptor–Receptor Interactions as an Integrative Mechanism in the Basal Ganglia. *Trends Neurosci* 1997, *20*, 482–487, doi:10.1016/S0166-2236(97)01096-5.
140. Fuxe, K.; Marcellino, D.; Borroto-Escuela, D.O.; Guescini, M.; Fernández-Dueñas, V.; Tanganelli, S.; Rivera, A.; Ciruela, F.; Agnati, L.F. Adenosine–Dopamine Interactions in the Pathophysiology and Treatment of CNS Disorders. *CNS Neurosci Ther* 2010, *16*, e18–e42, doi:10.1111/J.1755-5949.2009.00126.X.
141. Ferré, S.; Torvinen, M.; Antoniou, K.; Irenius, E.; Civelli, O.; Arenas, E.; Fredholm, B.B.; Fuxe, K. Adenosine A1 Receptor-Mediated Modulation of Dopamine D1 Receptors in Stably Cotransfected Fibroblast Cells. *Journal of Biological Chemistry* 1998, *273*, 4718–4724, doi:10.1074/JBC.273.8.4718.
142. Cheong, S.L.; Federico, S.; Spalluto, G.; Klotz, K.N.; Pastorin, G. The Current Status of Pharmacotherapy for the Treatment of Parkinson's Disease: Transition from Single-Target to Multitarget Therapy. *Drug Discov Today* 2019, *24*, 1769–1783, doi:10.1016/J.DRUDIS.2019.05.003.
143. Ferré, S.; Fredholm, B.B.; Morelli, M.; Popoli, P.; Fuxe, K. Adenosine–Dopamine Receptor–Receptor Interactions as an Integrative Mechanism in the Basal Ganglia. *Trends Neurosci* 1997, *20*, 482–487, doi:10.1016/S0166-2236(97)01096-5.
144. Ongini, E.; Fredholm, B.B. Pharmacology of Adenosine A2A Receptors. *Trends Pharmacol Sci* 1996, *17*, 364–372, doi:10.1016/S0165-6147(96)80010-1.
145. Lv, Y.C.; Gao, A.B.; Yang, J.; Zhong, L.Y.; Jia, B.; Ouyang, S.H.; Gui, L.; Peng, T.H.; Sun, S.; Cayabyab, F. Long-Term Adenosine A1 Receptor Activation-Induced Sortilin Expression Promotes α -Synuclein Upregulation in Dopaminergic Neurons. *Neural Regen Res* 2020, *15*, 712–723, doi:10.4103/1673-5374.266916.
146. Jakova, E.; Moutaoufik, M.T.; Lee, J.S.; Babu, M.; Cayabyab, F.S. Adenosine A1 Receptor Ligands Bind to α -Synuclein: Implications for α -Synuclein Misfolding and α -Synucleinopathy in Parkinson's Disease. *Transl Neurodegener* 2022, *11*, 1–26, doi:10.1186/S40035-022-00284-3/FIGURES/12.
147. IJzerman, A.P.; Jacobson, K.A.; Müller, C.E.; Cronstein, B.N.; Cunha, R.A. International Union of Basic and Clinical Pharmacology. CXII: Adenosine Receptors: A Further Update. *Pharmacol Rev* 2022, *74*, 340–372, doi:10.1124/pharmrev.121.000445.
148. Renk, D.R.; Skraban, M.; Bier, D.; Schulze, A.; Wabbs, E.; Wedekind, F.; Neumaier, F.; Neumaier, B.; Holschbach, M. Design, Synthesis and Biological Evaluation of Tozadenant Analogues as Adenosine A2A Receptor Ligands. *Eur J Med Chem* 2021, *214*, 113214, doi:10.1016/J.EJMECH.2021.113214.
149. Chiu, G.S.; Darmody, P.T.; Walsh, J.P.; Moon, M.L.; Kwakwa, K.A.; Bray, J.K.; McCusker, R.H.; Freund, G.G. Adenosine through the A2A Adenosine Receptor Increases IL-1 β in the Brain Contributing to Anxiety. *Brain Behav Immun* 2014, *41*, 218–231, doi:10.1016/J.BBI.2014.05.018.
150. Bernier, L.P. Purinergic Regulation of Inflammasome Activation after Central Nervous System Injury. *Journal of General Physiology* 2012, *140*, 571–575, doi:10.1085/JGP.201210875.

151. Rebola, N.; Simões, A.P.; Canas, P.M.; Tomé, A.R.; Andrade, G.M.; Barry, C.E.; Agostinho, P.M.; Lynch, M.A.; Cunha, R.A. Adenosine A2A Receptors Control Neuroinflammation and Consequent Hippocampal Neuronal Dysfunction. *J Neurochem* 2011, *117*, 100–111, doi:10.1111/J.1471-4159.2011.07178.X.
152. Cheng, C.; Cheng, W.; Wang, Y.; Chen, X.; Zhang, L.; Li, Y.; Shen, F.; Yuan, D.; Hong, P.; Huang, W. The Dual Role of A2aR in Neuroinflammation: Modulating Microglial Polarization in White Matter Lesions. *eNeuro* 2025, *12*, doi:10.1523/ENEURO.0579-24.2025.
153. Pedata, F.; Pugliese, A.M.; Coppi, E.; Dettori, I.; Maraula, G.; Cellai, L.; Melani, A. Adenosine A2A Receptors Modulate Acute Injury and Neuroinflammation in Brain Ischemia. *Mediators Inflamm* 2014, *2014*, 805198, doi:10.1155/2014/805198.
154. Reece, T.B.; Okonkwo, D.O.; Ellman, P.I.; Warren, P.S.; Smith, R.L.; Hawkins, A.S.; Linden, J.; Kron, I.L.; Tribble, C.G.; Kern, J.A. The Evolution of Ischemic Spinal Cord Injury in Function, Cytoarchitecture, and Inflammation and the Effects of Adenosine A2A Receptor Activation. *J Thorac Cardiovasc Surg* 2004, *128*, 925–932, doi:10.1016/J.JTCVS.2004.08.019.
155. Guerrero, A. A2A Adenosine Receptor Agonists and Their Potential Therapeutic Applications. An Update. *Curr Med Chem* 2018, *25*, 3597–3612, doi:10.2174/0929867325666180313110254.
156. Boknik, P.; Eskandar, J.; Hofmann, B.; Zimmermann, N.; Neumann, J.; Gergs, U. Role of Cardiac A2A Receptors Under Normal and Pathophysiological Conditions. *Front Pharmacol* 2021, *11*, 627838, doi:10.3389/FPHAR.2020.627838/XML.
157. Boison, D. Adenosinergic Signaling in Epilepsy. *Neuropharmacology* 2016, *104*, 131–139, doi:10.1016/J.NEUROPHARM.2015.08.046.
158. Sebastião, A.M.; Ribeiro, J.A. Adenosine A2 Receptor-Mediated Excitatory Actions on the Nervous System. *Prog Neurobiol* 1996, *48*, 167–189, doi:10.1016/0301-0082(95)00035-6.
159. Dai, S.S.; Zhou, Y.G. Adenosine 2A Receptor: A Crucial Neuromodulator with Bidirectional Effect in Neuroinflammation and Brain Injury. *Rev Neurosci* 2011, *22*, 231–239, doi:10.1515/RNS.2011.020/MACHINEREADABLECITATION/RIS.
160. Orr, A.G.; Orr, A.L.; Li, X.J.; Gross, R.E.; Traynelis, S.F. Adenosine A2A Receptor Mediates Microglial Process Retraction. *Nat Neurosci* 2009, *12*, 872–878, doi:10.1038/NN.2341;KWRD=BIOMEDICINE.
161. Gomes, C.A.R.V.; Vaz, S.H.; Ribeiro, J.A.; Sebastião, A.M. Glial Cell Line-Derived Neurotrophic Factor (GDNF) Enhances Dopamine Release from Striatal Nerve Endings in an Adenosine A2A Receptor-Dependent Manner. *Brain Res* 2006, *1113*, 129–136, doi:10.1016/J.BRAINRES.2006.07.025.
162. Gyoneva, S.; Shapiro, L.; Lazo, C.; Garnier-Amblard, E.; Smith, Y.; Miller, G.W.; Traynelis, S.F. Adenosine A2A Receptor Antagonism Reverses Inflammation-Induced Impairment of Microglial Process Extension in a Model of Parkinson's Disease. *Neurobiol Dis* 2014, *67*, 191–202, doi:10.1016/J.NBD.2014.03.004.
163. Minghetti, L.; Greco, A.; Potenza, R.L.; Pezzola, A.; Blum, D.; Bantubungi, K.; Popoli, P. Effects of the Adenosine A2A Receptor Antagonist SCH 58621 on Cyclooxygenase-2 Expression, Glial Activation, and Brain-Derived Neurotrophic Factor Availability in a Rat Model of Striatal Neurodegeneration. *J Neuropathol Exp Neurol* 2007, *66*, 363–371, doi:10.1097/NEN.0B013E3180517477.
164. Fiebich, B.L.; Biber, K.; Lieb, K.; Van Calcar, D.; Berger, M.; Bauer, J.; Gebicke-Haerter, P.J. Cyclooxygenase-2 Expression in Rat Microglia Is Induced by Adenosine A2a-Receptors. *Glia* 1996, *18*, 152–180, doi:10.1002/(SICI)1098-1136(199610)18:2<152::AID-GLIA7>3.0.CO;2-2.
165. Müller, L.; Di Benedetto, S.; Müller, V. From Homeostasis to Neuroinflammation: Insights into Cellular and Molecular Interactions and Network Dynamics. *Cells* 2025, *Vol. 14*, Page 54 2025, *14*, 54, doi:10.3390/CELLS14010054.
166. Colella, M.; Zinni, M.; Pansiot, J.; Cassanello, M.; Mairesse, J.; Ramenghi, L.; Baud, O. Modulation of Microglial Activation by Adenosine A2a Receptor in Animal Models of Perinatal Brain Injury. *Front Neurol* 2018, *9*, 355877, doi:10.3389/FNEUR.2018.00605/BIBTEX.
167. Illes, P.; Rubini, P.; Ulrich, H.; Zhao, Y.; Tang, Y. Regulation of Microglial Functions by Purinergic Mechanisms in the Healthy and Diseased CNS. *Cells* 2020, *Vol. 9*, Page 1108 2020, *9*, 1108, doi:10.3390/CELLS9051108.
168. Merighi, S.; Borea, P.A.; Varani, K.; Vincenzi, F.; Jacobson, K.A.; Gessi, S. A2A Adenosine Receptor Antagonists in Neurodegenerative Diseases. *Curr Med Chem* 2022, *29*, 4138, doi:10.2174/092986732866621129122550.
169. Pedata, F.; Gianfriddo, M.; Turchi, D.; Melani, A. The Protective Effect of Adenosine A2A Receptor Antagonism in Cerebral Ischemia. *Neurol Res* 2005, *27*, 169–174, doi:10.1179/016164105X21913;PAGE:STRING:ARTICLE/CHAPTER.
170. Brambilla, R.; Cottini, L.; Fumagalli, M.; Ceruti, S.; Abbracchio, M.P. Blockade of A2A Adenosine Receptors Prevents Basic Fibroblast Growth Factor-Induced Reactive Astrogliosis in Rat Striatal Primary Astrocytes. *Glia* 2003, *43*, 190–194, doi:10.1002/GLIA.10243;CTYPE:STRING:JOURNAL.
171. Hindley, S.; Herman, M.A.R.; Rathbone, M.P. Stimulation of Reactive Astrogliosis in Vivo by Extracellular Adenosine Diphosphate or an Adenosine A2 Receptor Agonist. *J Neurosci Res* 1994, *38*, 399–406, doi:10.1002/JNR.490380405;WGROU:STRING:PUBLICATION.
172. Paiva, I.; Carvalho, K.; Santos, P.; Cellai, L.; Pavlou, M.A.S.; Jain, G.; Gnad, T.; Pfeifer, A.; Vieau, D.; Fischer, A.; et al. A2AR-Induced Transcriptional Deregulation in Astrocytes: An in Vitro Study. *Glia* 2019, *67*, 2329–2342, doi:10.1002/GLIA.23688;REQUESTEDJOURNAL:JOURNAL:10981136;PAGE:STRING:ARTICLE/CHAPTER.
173. Lopes, C.R.; Cunha, R.A.; Agostinho, P. Astrocytes and Adenosine A2A Receptors: Active Players in Alzheimer's Disease. *Front Neurosci* 2021, *15*, 666710, doi:10.3389/FNINS.2021.666710/XML/NLM.
174. Matos, M.; Shen, H.Y.; Augusto, E.; Wang, Y.; Wei, C.J.; Wang, Y.T.; Agostinho, P.; Boison, D.; Cunha, R.A.; Chen, J.F. Deletion of Adenosine A2A Receptors From Astrocytes Disrupts Glutamate Homeostasis Leading to Psychomotor and Cognitive Impairment: Relevance to Schizophrenia. *Biol Psychiatry* 2015, *78*, 763–774, doi:10.1016/J.BIOPSYCH.2015.02.026.

175. Yuan, J.; Chen, L.; Wang, J.; Xia, S.; Huang, J.; Zhou, L.; Feng, C.; Hu, X.; Zhou, Z.; Ran, H. Adenosine A2A Receptor Suppressed Astrocyte-Mediated Inflammation Through the Inhibition of STAT3/YKL-40 Axis in Mice With Chronic Cerebral Hypoperfusion-Induced White Matter Lesions. *Front Immunol* 2022, 13, 841290, doi:10.3389/FIMMU.2022.841290/BIBTEX.
176. Pugliese, A.M.; Traini, C.; Cipriani, S.; Gianfriddo, M.; Mello, T.; Giovannini, M.G.; Galli, A.; Pedata, F. The Adenosine A 2A Receptor Antagonist ZM241385 Enhances Neuronal Survival after Oxygen-Glucose Deprivation in Rat CA1 Hippocampal Slices. *Br J Pharmacol* 2009, 157, 818–830, doi:10.1111/J.1476-5381.2009.00218.X;SUBPAGE:STRING:FULL.
177. Congreve, M.; Brown, G.A.; Borodovsky, A.; Lamb, M.L. Targeting Adenosine A 2A Receptor Antagonism for Treatment of Cancer. *Expert Opin Drug Discov* 2018, 13, 997–1003, doi:10.1080/17460441.2018.1534825;WGROU:STRING:PUBLICATION.
178. Hodgson, R.A.; Bertorelli, R.; Varty, G.B.; Lachowicz, J.E.; Forlani, A.; Fredduzzi, S.; Cohen-Williams, M.E.; Higgins, G.A.; Impagnatiello, F.; Nicolussi, E.; et al. Characterization of the Potent and Highly Selective A2A Receptor Antagonists Preladenant and SCH 412348 [7-[2-[4-(2,4-Difluorophenyl)-1-Piperazinyl]Ethyl]-2-(2-Furanyl)-7H-Pyrazolo[4,3-e][1,2,4]Triazolo[1,5-c]Pyrimidin-5-Amine] in Rodent Models of Movement Disorders and Depression. *J Pharmacol Exp Ther* 2009, 330, 294–303, doi:10.1124/JPET.108.149617.
179. Pechlivanova, D.M.; Tchekalarova, J.D.; Alova, L.H.; Petkov, V. V.; Nikolov, R.P.; Yakimova, K.S. Effect of Long-Term Caffeine Administration on Depressive-like Behavior in Rats Exposed to Chronic Unpredictable Stress. *Behavioural Pharmacology* 2012, 23, 339–347, doi:10.1097/FBP.0b013e3283564dd9.
180. Beamer, E.; Göllöncsér, F.; Horváth, G.; Beko, K.; Otrokocsi, L.; Koványi, B.; Sperlágh, B. Purinergic Mechanisms in Neuroinflammation: An Update from Molecules to Behavior. *Neuropharmacology* 2016, 104, 94–104, doi:10.1016/J.NEUROPHARM.2015.09.019.
181. Thurner, P.; Gsandtner, I.; Kudlacek, O.; Choquet, D.; Nanoff, C.; Freissmuth, M.; Zezula, J. A Two-State Model for the Diffusion of the A 2A Adenosine Receptor in Hippocampal Neurons. *Journal of Biological Chemistry* 2014, 289, 9263–9274, doi:10.1074/jbc.m113.505685.
182. Chen, J.F.; Sonsalla, P.K.; Pedata, F.; Melani, A.; Domenici, M.R.; Popoli, P.; Geiger, J.; Lopes, L. V.; de Mendonça, A. Adenosine A2A Receptors and Brain Injury: Broad Spectrum of Neuroprotection, Multifaceted Actions and “Fine Tuning” Modulation. *Prog Neurobiol* 2007, 83, 310–331, doi:10.1016/J.PNEUROBIO.2007.09.002.
183. Boussadia, Z.; Chiodi, V.; Pazienti, A.; Martire, A. A Major Role for Adenosine A2A Receptor in the Interaction between Astrocytes and Myelinated Neurons: Possible Implications for the Therapy of Neurodegenerative Disorders. *Purinergic Signal* 2022, 18, 5–7, doi:10.1007/S11302-021-09835-1/METRICS.
184. Merighi, S.; Battistello, E.; Casetta, I.; Gragnaniello, D.; Poloni, T.E.; Medici, V.; Cirrincione, A.; Varani, K.; Vincenzi, F.; Borea, P.A.; et al. Upregulation of Cortical A2A Adenosine Receptors Is Reflected in Platelets of Patients with Alzheimer's Disease. *Journal of Alzheimer's Disease* 2021, 80, 1105–1117, doi:10.3233/JAD-201437.
185. Jacobson, K.A.; Gao, Z.G.; Matricon, P.; Eddy, M.T.; Carlsson, J. Adenosine A2A Receptor Antagonists: From Caffeine to Selective Non-Xanthines. *Br J Pharmacol* 2022, 179, 3496–3511, doi:10.1111/BPH.15103;CTYPE:STRING:JOURNAL.
186. Albasanz, J.L.; Perez, S.; Barrachina, M.; Ferrer, I.; Martín, M. Up-Regulation of Adenosine Receptors in the Frontal Cortex in Alzheimer's Disease. *Brain Pathology* 2008, 18, 211–219, doi:10.1111/J.1750-3639.2007.00112.X.
187. Woods, L.T.; Ajit, D.; Camden, J.M.; Erb, L.; Weisman, G.A. Purinergic Receptors as Potential Therapeutic Targets in Alzheimer's Disease. *Neuropharmacology* 2016, 104, 169–179, doi:10.1016/J.NEUROPHARM.2015.10.031.
188. Matos, M.; Augusto, E.; Agostinho, P.; Cunha, R.A.; Chen, J.F. Antagonistic Interaction between Adenosine A2A Receptors and Na⁺/K⁺-ATPase-A2 Controlling Glutamate Uptake in Astrocytes. *Journal of Neuroscience* 2013, 33, 18492–18502, doi:10.1523/JNEUROSCI.1828-13.2013.
189. Orr, A.G.; Hsiao, E.C.; Wang, M.M.; Ho, K.; Kim, D.H.; Wang, X.; Guo, W.; Kang, J.; Yu, G.Q.; Adame, A.; et al. Astrocytic Adenosine Receptor A2A and Gs-Coupled Signaling Regulate Memory. *Nature Neuroscience* 2015 18:3 2015, 18, 423–434, doi:10.1038/nn.3930.
190. Arendash, G.W.; Schleif, W.; Rezai-Zadeh, K.; Jackson, E.K.; Zacharia, L.C.; Cracchiolo, J.R.; Shippy, D.; Tan, J. Caffeine Protects Alzheimer's Mice against Cognitive Impairment and Reduces Brain β -Amyloid Production. *Neuroscience* 2006, 142, 941–952, doi:10.1016/J.NEUROSCIENCE.2006.07.021.
191. Dall'igna, O.P.; Porciúncula, L.O.; Souza, D.O.; Cunha, R.A.; Lara, D.R. Neuroprotection by Caffeine and Adenosine A 2A Receptor Blockade of β -Amyloid Neurotoxicity. *Br J Pharmacol* 2003, 138, 1207–1209, doi:10.1038/SJ.BJP.0705185;WEBSITE:WEBSITE:BPSPUBS;WGROU:STRING:PUBLICATION.
192. Cao, C.; Cirrito, J.R.; Lin, X.; Wang, L.; Verges, D.K.; Dickson, A.; Mamcarz, M.; Zhang, C.; Mori, T.; Arendash, G.W.; et al. Caffeine Suppresses Amyloid- β Levels in Plasma and Brain of Alzheimer's Disease Transgenic Mice. *Journal of Alzheimer's Disease* 2009, 17, 681–697, doi:10.3233/JAD-2009-1071;JOURNAL:JOURNAL:ALZA;WEBSITE:WEBSITE:SAGE;WGROU:STRING:PUBLICATION.
193. Canas, P.M.; Porciúncula, L.O.; Cunha, G.M.A.; Silva, C.G.; Machado, N.J.; Oliveira, J.M.A.; Oliveira, C.R.; Cunha, R.A. Adenosine A2A Receptor Blockade Prevents Synaptotoxicity and Memory Dysfunction Caused by β -Amyloid Peptides via P38 Mitogen-Activated Protein Kinase Pathway. *Journal of Neuroscience* 2009, 29, 14741–14751, doi:10.1523/JNEUROSCI.3728-09.2009.
194. Canas, P.M.; Cunha, R.A.; Agostinho, P. Adenosine Receptors in Alzheimer's Disease. *Receptors* 2018, 34, 259–280, doi:10.1007/978-3-319-90808-3_11.

195. Da Silva, S.V.; Haberl, M.G.; Zhang, P.; Bethge, P.; Lemos, C.; Gonçalves, N.; Gorlewicz, A.; Malezieux, M.; Gonçalves, F.Q.; Grosjean, N.; et al. Early Synaptic Deficits in the APP/PS1 Mouse Model of Alzheimer's Disease Involve Neuronal Adenosine A2A Receptors. *Nature Communications* 2016 7:1 2016, 7, 1–11, doi:10.1038/ncomms11915.
196. Chen, J.F. Adenosine Receptor Control of Cognition in Normal and Disease. *Int Rev Neurobiol* 2014, 119, 257–307, doi:10.1016/B978-0-12-801022-8.00012-X.
197. Temido-Ferreira, M.; Ferreira, D.G.; Batalha, V.L.; Marques-Morgado, I.; Coelho, J.E.; Pereira, P.; Gomes, R.; Pinto, A.; Carvalho, S.; Canas, P.M.; et al. Age-Related Shift in LTD Is Dependent on Neuronal Adenosine A2A Receptors Interplay with mGluR5 and NMDA Receptors. *Mol Psychiatry* 2020, 25, 1876–1900, doi:10.1038/S41380-018-0110-9;TECHMETA=110,13,59,64;SUBJMETA=378,45,631;KWRD=BIOCHEMISTRY,NEUROSCIENCE.
198. Gomez-Murcia, V.; Launay, A.; Carvalho, K.; Burgard, A.; Meriaux, C.; Caillierez, R.; Eddarkaoui, S.; Kilinc, D.; Siedlecki-Wullich, D.; Besegher, M.; et al. Neuronal A2A Receptor Exacerbates Synapse Loss and Memory Deficits in APP/PS1 Mice. *Brain* 2024, 147, 2691–2705, doi:10.1093/BRAIN/AWAE113.
199. Launay, A.; Carvalho, K.; Genin, A.; Gauvrit, T.; Nobili, P.; Gomez-Murcia, V.; Augustin, E.; Burgard, A.; Gambi, J.; Fourmy, D.; et al. Upregulation of Adenosine A2A Receptor in Astrocytes Is Sufficient to Trigger Hippocampal Multicellular Dysfunctions and Memory Deficits. *Mol Psychiatry* 2025, 1–15, doi:10.1038/S41380-025-03115-9;TECHMETA.
200. Morelli, M.; Carta, A.R.; Kachroo, A.; Schwarzschild, M.A. Pathophysiological Roles for Purines: Adenosine, Caffeine and Urate. *Prog Brain Res* 2010, 183, 183–208, doi:10.1016/S0079-6123(10)83010-9.
201. Schiffmann, S.N.; Fisone, G.; Moresco, R.; Cunha, R.A.; Ferré, S. Adenosine A2A Receptors and Basal Ganglia Physiology. *Prog Neurobiol* 2007, 83, 277–292, doi:10.1016/J.PNEUROBIO.2007.05.001.
202. Fuxe, K.; Ferre, S.; Canals, M.; Torvinen, M.; Terasmaa, A.; Marcellino, D.; Goldberg, S.R.; Staines, W.; Jacobsen, K.X.; Lluis, C.; et al. Adenosine A2A and Dopamine D2 Heteromeric Receptor Complexes and Their Function. *J Mol Neurosci* 2005, 26, 209–217, doi:10.1385/JMN/26.
203. Morelli, M.; Carta, A.R.; Jenner, P. Adenosine A2A Receptors and Parkinson's Disease. *Handb Exp Pharmacol* 2009, 193, 589–615, doi:10.1007/978-3-540-89615-9_18.
204. Pinna, A. Adenosine A2A Receptor Antagonists in Parkinson's Disease: Progress in Clinical Trials from the Newly Approved Istradefylline to Drugs in Early Development and Those Already Discontinued. *CNS Drugs* 2014, 28, 455–474, doi:10.1007/S40263-014-0161-7/TABLES/3.
205. Kanda, T.; Uchida, S. ichi Clinical/Pharmacological Aspect of Adenosine A2A Receptor Antagonist for Dyskinesia. *Int Rev Neurobiol* 2014, 119, 127–150, doi:10.1016/B978-0-12-801022-8.00006-4.
206. Ferreira, D.G.; Batalha, V.L.; Miranda, H.V.; Coelho, J.E.; Gomes, R.; Gonçalves, F.Q.; Real, J.I.; Rino, J.; Albino-Teixeira, A.; Cunha, R.A.; et al. Adenosine A2A Receptors Modulate α -Synuclein Aggregation and Toxicity. *Cerebral Cortex* 2017, 27, 718–730, doi:10.1093/CERCOR/BHV268.
207. Uchida, S. ichi; Kadowaki-Horita, T.; Kanda, T. Effects of the Adenosine A2A Receptor Antagonist on Cognitive Dysfunction in Parkinson's Disease. *Int Rev Neurobiol* 2014, 119, 169–189, doi:10.1016/B978-0-12-801022-8.00008-8.
208. Yamada, K.; Kobayashi, M.; Shiozaki, S.; Ohta, T.; Mori, A.; Jenner, P.; Kanda, T. Antidepressant Activity of the Adenosine A2A Receptor Antagonist, Istradefylline (KW-6002) on Learned Helplessness in Rats. *Psychopharmacology (Berl)* 2014, 231, 2839–2849, doi:10.1007/S00213-014-3454-0/FIGURES/5.
209. Gyoneva, S.; Shapiro, L.; Lazo, C.; Garnier-Amblard, E.; Smith, Y.; Miller, G.W.; Traynelis, S.F. Adenosine A2A Receptor Antagonism Reverses Inflammation-Induced Impairment of Microglial Process Extension in a Model of Parkinson's Disease. *Neurobiol Dis* 2014, 67, 191–202, doi:10.1016/J.NBD.2014.03.004.
210. Tansey, M.G.; Goldberg, M.S. Neuroinflammation in Parkinson's Disease: Its Role in Neuronal Death and Implications for Therapeutic Intervention. *Neurobiol Dis* 2010, 37, 510–518, doi:10.1016/J.NBD.2009.11.004.
211. Pinna, A. Novel Investigational Adenosine A2A Receptor Antagonists for Parkinson's Disease. *Expert Opin Investig Drugs* 2009, 18, 1619–1631, doi:10.1517/13543780903241615;WEBSITE:WEBSITE:TFOPB;PAGEGROUP:STRING:PUBLICATION.
212. Chen, J.F.; Cunha, R.A. The Belated US FDA Approval of the Adenosine A2A Receptor Antagonist Istradefylline for Treatment of Parkinson's Disease. *Purinergic Signal* 2020, 16, 167–174, doi:10.1007/S11302-020-09694-2/FIGURES/1.
213. Petzer, J.P.; Steyn, S.; Castagnoli, K.P.; Chen, J.F.; Schwarzschild, M.A.; Van Der Schyf, C.J.; Castagnoli, N. Inhibition of Monoamine Oxidase B by Selective Adenosine A2A Receptor Antagonists. *Bioorg Med Chem* 2003, 11, 1299–1310, doi:10.1016/S0968-0896(02)00648-X.
214. Preti, D.; Baraldi, P.G.; Moorman, A.R.; Borea, P.A.; Varani, K. History and Perspectives of A2A Adenosine Receptor Antagonists as Potential Therapeutic Agents. *Med Res Rev* 2015, 35, 790–848, doi:10.1002/MED.21344;PAGEGROUP:STRING:PUBLICATION.
215. Jaiteh, M.; Zeifman, A.; Saarinen, M.; Svenningsson, P.; Bréa, J.; Loza, M.I.; Carlsson, J. Docking Screens for Dual Inhibitors of Disparate Drug Targets for Parkinson's Disease. *J Med Chem* 2018, 61, 5269–5278, doi:10.1021/ACS.JMEDCHEM.8B00204/SUPPL_FILE/JM8B00204_SI_004.PDB.
216. Stöbel, A.; Schlenk, M.; Hinz, S.; Küppers, P.; Heer, J.; Gütschow, M.; Müller, C.E. Dual Targeting of Adenosine A2A Receptors and Monoamine Oxidase B by 4H-3,1-Benzothiazin-4-Ones. *J Med Chem* 2013, 56, 4580–4596, doi:10.1021/JM400336X/SUPPL_FILE/JM400336X_SI_001.PDF.
217. Popoli, P.; Blum, D.; Martire, A.; Ledent, C.; Ceruti, S.; Abbracchio, M.P. Functions, Dysfunctions and Possible Therapeutic Relevance of Adenosine A2A Receptors in Huntington's Disease. *Prog Neurobiol* 2007, 81, 331–348, doi:10.1016/J.PNEUROBIO.2006.12.005.

218. Varani, K.; Bachoud-Lévi, A.C.; Mariotti, C.; Tarditi, A.; Abbracchio, M.P.; Gasperi, V.; Borea, P.A.; Dolbeau, G.; Gellera, C.; Solari, A.; et al. Biological Abnormalities of Peripheral A2A Receptors in a Large Representation of Polyglutamine Disorders and Huntington's Disease Stages. *Neurobiol Dis* 2007, *27*, 36–43, doi:10.1016/J.NBD.2007.03.011.
219. Tarditi, A.; Camurri, A.; Varani, K.; Borea, P.A.; Woodman, B.; Bates, G.; Cattaneo, E.; Abbracchio, M.P. Early and Transient Alteration of Adenosine A2A Receptor Signaling in a Mouse Model of Huntington Disease. *Neurobiol Dis* 2006, *23*, 44–53, doi:10.1016/J.NBD.2006.01.014.
220. Lee, C. fei; Chern, Y. Adenosine Receptors and Huntington's Disease. *Int Rev Neurobiol* 2014, *119*, 195–232, doi:10.1016/B978-0-12-801022-8.00010-6.
221. Mojsilovic-Petrovic, J.; Jeong, G.B.; Crocker, A.; Arneja, A.; David, S.; Russell, D.; Kalb, R.G. Protecting Motor Neurons from Toxic Insult by Antagonism of Adenosine A2a and Trk Receptors. *Journal of Neuroscience* 2006, *26*, 9250–9263, doi:10.1523/JNEUROSCI.1856-06.2006.
222. Volonté, C.; Apolloni, S.; Parisi, C.; Amadio, S. Purinergic Contribution to Amyotrophic Lateral Sclerosis. *Neuropharmacology* 2016, *104*, 180–193, doi:10.1016/J.NEUROPHARM.2015.10.026.
223. Higashimori, H.; Tolman, M.; Yang, Y. Suppression of Adenosine 2a Receptor (A 2a R) -Mediated Adenosine Signaling Improves Disease Phenotypes in a Mouse Model of Amyotrophic Lateral Sclerosis. *Exp Neurol* 2015, *267*, 115–122, doi:10.1016/j.expneurol.2015.03.004.
224. Wiese, S.; Jablonka, S.; Holtmann, B.; Orel, N.; Rajagopal, R.; Chao, M. V.; Sendtner, M. Adenosine Receptor A2A-R Contributes to Motoneuron Survival by Transactivating the Tyrosine Kinase Receptor TrkB. *Proc Natl Acad Sci U S A* 2007, *104*, 17210–17215, doi:10.1073/PNAS.0705267104/ASSET/B6259867-D0DA-4124-8293-878EA0DDB57D/ASSETS/GRAPHIC/ZPQ0410779150004.JPEG.
225. Yanpallear, S.U.; Barrick, C.A.; Buckley, H.; Becker, J.; Tessarollo, L. Deletion of the BDNF Truncated Receptor TrkB.T1 Delays Disease Onset in a Mouse Model of Amyotrophic Lateral Sclerosis. *PLoS One* 2012, *7*, e39946, doi:10.1371/JOURNAL.PONE.0039946.
226. Nascimento, F.; Pousinha, P.A.; Correia, A.M.; Gomes, R.; Sebastião, A.M.; Ribeiro, J.A. Adenosine A2A Receptors Activation Facilitates Neuromuscular Transmission in the Pre-Symptomatic Phase of the SOD1(G93A) ALS Mice, but Not in the Symptomatic Phase. *PLoS One* 2014, *9*, e104081, doi:10.1371/JOURNAL.PONE.0104081.
227. Rei, N.; Rombo, D.M.; Ferreira, M.F.; Baqi, Y.; Müller, C.E.; Ribeiro, J.A.; Sebastião, A.M.; Vaz, S.H. Hippocampal Synaptic Dysfunction in the SOD1G93A Mouse Model of Amyotrophic Lateral Sclerosis: Reversal by Adenosine A2AR Blockade. *Neuropharmacology* 2020, *171*, 108106, doi:10.1016/J.NEUROPHARM.2020.108106.
228. Lai, C.Y.; Liu, Y.J.; Lai, H.L.; Chen, H.M.; Kuo, H.C.; Liao, Y.P.; Chern, Y. The D2 Dopamine Receptor Interferes with the Protective Effect of the A2A Adenosine Receptor on TDP-43 Mislocalization in Experimental Models of Motor Neuron Degeneration. *Front Neurosci* 2018, *12*, 336030, doi:10.3389/FNINS.2018.00187/BIBTEX.
229. Liu, Y.J.; Ju, T.C.; Chen, H.M.; Jang, Y.S.; Lee, L.M.; Lai, H.L.; Tai, H.C.; Fang, J.M.; Lin, Y.L.; Tu, P.H.; et al. Activation of AMP-Activated Protein Kinase A1 Mediates Mislocalization of TDP-43 in Amyotrophic Lateral Sclerosis. *Hum Mol Genet* 2015, *24*, 787–801, doi:10.1093/HMG/DDU497.
230. Armida, M.; Matteucci, A.; Pèzzola, A.; Baqi, Y.; Müller, C.E.; Popoli, P.; Potenza, R.L. Modulating P1 Adenosine Receptors in Disease Progression of SOD1 G93A Mutant Mice. *Neurochem Res* 2019, *44*, 1037–1042, doi:10.1007/S11064-019-02745-0/FIGURES/2.
231. Potenza, R.L.; Armida, M.; Ferrante, A.; Pèzzola, A.; Matteucci, A.; Puopolo, M.; Popoli, P. Effects of Chronic Caffeine Intake in a Mouse Model of Amyotrophic Lateral Sclerosis. *J Neurosci Res* 2013, *91*, 585–592, doi:10.1002/JNR.23185;JOURNAL:JOURNAL:10974547;PAGE:STRING:ARTICLE/CHAPTER.
232. Fondell, E.; Oreilly, É.J.; Fitzgerald, K.C.; Falcone, G.J.; Kolonel, L.N.; Park, Y.; Gapstur, S.M.; Ascherio, A. Intakes of Caffeine, Coffee and Tea and Risk of Amyotrophic Lateral Sclerosis: Results from Five Cohort Studies. *Amyotroph Lateral Scler Frontotemporal Degener* 2015, *16*, 366–371, doi:10.3109/21678421.2015.1020813.
233. Beghi, E.; Pupillo, E.; Messina, P.; Giussani, G.; Chiò, A.; Zoccollella, S.; Moglia, C.; Corbo, M.; Logroscino, G.; Al-Chalabi, A.; et al. Coffee and Amyotrophic Lateral Sclerosis: A Possible Preventive Role. *Am J Epidemiol* 2011, *174*, 1002–1008, doi:10.1093/AJE/KWR229.
234. Huin, V.; Blum, D.; Delforge, V.; Cailliau, E.; Djedjiri, S.; Dujardin, K.; Genet, A.; Viard, R.; Attarian, S.; Bruneteau, G.; et al. Caffeine Consumption Outcomes on Amyotrophic Lateral Sclerosis Disease Progression and Cognition. *Neurobiol Dis* 2024, *199*, 106603, doi:10.1016/J.NBD.2024.106603.
235. Vincenzi, F.; Corciulo, C.; Targa, M.; Casetta, I.; Gentile, M.; Granieri, E.; Borea, P.A.; Popoli, P.; Varani, K. A2A Adenosine Receptors Are Up-Regulated in Lymphocytes from Amyotrophic Lateral Sclerosis Patients. *Amyotroph Lateral Scler Frontotemporal Degener* 2013, *14*, 406–413, doi:10.3109/21678421.2013.793358;REQUESTEDJOURNAL:JOURNAL:IAFD20;WGROUP:STRING:PUBLICATION.
236. Chen, J.F.; Huang, Z.; Ma, J.; Zhu, J.M.; Moratalla, R.; Standaert, D.; Moskowitz, M.A.; Fink, J.S.; Schwarzschild, M.A. A2A Adenosine Receptor Deficiency Attenuates Brain Injury Induced by Transient Focal Ischemia in Mice. *Journal of Neuroscience* 1999, *19*, 9192–9200, doi:10.1523/JNEUROSCI.19-21-09192.1999.
237. Hinz, S.; Lacher, S.K.; Seibt, B.F.; Müller, C.E. BAY60-6583 Acts as a Partial Agonist at Adenosine A2B Receptors. *Journal of Pharmacology and Experimental Therapeutics* 2014, *349*, 427–436, doi:10.1124/JPET.113.210849.
238. Beukers, M.W.; Meurs, I.; Ijzerman, A.P. Structure-Affinity Relationships of Adenosine A2B Receptor Ligands. *Med Res Rev* 2006, *26*, 667–698, doi:10.1002/MED.20069.

239. Härter, M.; Kalthof, B.; Delbeck, M.; Lustig, K.; Gerisch, M.; Schulz, S.; Kast, R.; Meibom, D.; Lindner, N. Novel Non-Xanthine Antagonist of the A2B Adenosine Receptor: From HTS Hit to Lead Structure. *Eur J Med Chem* 2019, *163*, 763–778, doi:10.1016/J.EJMECH.2018.11.045.
240. Wainberg, Z.A.; Han, S.-W.; Lee, S.; Lee, K.-W.; Kopetz, S.; Mizrahi, J.; Hong, Y.S.; Ghiringhelli, F.; Italiano, A.; Tougeron, D.; et al. ARC-9: A Randomized Study to Evaluate Etrumadenant Based Treatment Combinations in Previously Treated Metastatic Colorectal Cancer (MCRC). *Journal of Clinical Oncology* 2024, *42*, 3508–3508, doi:10.1200/JCO.2024.42.16_SUPPL.3508.
241. Claff, T.; Schlegel, J.G.; Voss, J.H.; Vaaßen, V.J.; Weiße, R.H.; Cheng, R.K.Y.; Markovic-Mueller, S.; Bucher, D.; Sträter, N.; Müller, C.E. Crystal Structure of Adenosine A2A Receptor in Complex with Clinical Candidate Etrumadenant Reveals Unprecedented Antagonist Interaction. *Communications Chemistry* 2023 *6:1* 2023, *6*, 1–10, doi:10.1038/s42004-023-00894-6.
242. Dixon, A.K.; Gubit, A.K.; Sirinathsinghji, D.J.S.; Richardson, P.J.; Freeman, T.C. Tissue Distribution of Adenosine Receptor MRNAs in the Rat. *Br J Pharmacol* 1996, *118*, 1461–1468, doi:10.1111/J.1476-5381.1996.TB15561.X.
243. Koscsó, B.; Csóka, B.; Selmeczy, Z.; Himer, L.; Pacher, P.; Virág, L.; Haskó, G. Adenosine Augments IL-10 Production by Microglial Cells through an A2B Adenosine Receptor-Mediated Process. *The Journal of Immunology* 2012, *188*, 445–453, doi:10.4049/JIMMUNOL.1101224.
244. Merighi, S.; Bencivenni, S.; Vincenzi, F.; Varani, K.; Borea, P.A.; Gessi, S. A2B Adenosine Receptors Stimulate IL-6 Production in Primary Murine Microglia through P38 MAPK Kinase Pathway. *Pharmacol Res* 2017, *117*, 9–19, doi:10.1016/J.PHRS.2016.11.024.
245. Fiebich, B.L.; Biber, K.; Gyufko, K.; Berger, M.; Bauer, J.; Van Calter, D. Adenosine A2b Receptors Mediate an Increase in Interleukin (IL)-6 mRNA and IL-6 Protein Synthesis in Human Astrogloma Cells. *J Neurochem* 1996, *66*, 1426–1431, doi:10.1046/J.1471-4159.1996.66041426.X.
246. Rosi, S.; McGann, K.; Hauss-Wegrzyniak, B.; Wenk, G.L. The Influence of Brain Inflammation upon Neuronal Adenosine A2B Receptors. *J Neurochem* 2003, *86*, 220–227, doi:10.1046/J.1471-4159.2003.01825.X;PAGEGROUP:STRING:PUBLICATION.
247. Endo, F.; Kasai, A.; Soto, J.S.; Yu, X.; Qu, Z.; Hashimoto, H.; Gradinaru, V.; Kawaguchi, R.; Khakh, B.S. Molecular Basis of Astrocyte Diversity and Morphology across the CNS in Health and Disease. *Science (1979)* 2022, *378*, doi:10.1126/SCIENCE.ADC9020/SUPPL_FILE/SCIENCE.ADC9020_MOVIES_S1_AND_S2.ZIP.
248. Theparambil, S.M.; Kopach, O.; Braga, A.; Nizari, S.; Hosford, P.S.; Sagi-Kiss, V.; Hadjihambi, A.; Konstantinou, C.; Esteras, N.; Gutierrez Del Arroyo, A.; et al. Adenosine Signalling to Astrocytes Coordinates Brain Metabolism and Function. *Nature* 2024 *632:8023* 2024, *632*, 139–146, doi:10.1038/s41586-024-07611-w.
249. Fusco, I.; Ugolini, F.; Lana, D.; Coppi, E.; Dettori, I.; Gaviano, L.; Nosi, D.; Cherchi, F.; Pedata, F.; Giovannini, M.G.; et al. The Selective Antagonism of Adenosine A2B Receptors Reduces the Synaptic Failure and Neuronal Death Induced by Oxygen and Glucose Deprivation in Rat CA1 Hippocampus in Vitro. *Front Pharmacol* 2018, *9*, 329222, doi:10.3389/FPHAR.2018.00399/BIBTEX.
250. Coppi, E.; Dettori, I.; Cherchi, F.; Bulli, I.; Venturini, M.; Lana, D.; Giovannini, M.G.; Pedata, F.; Pugliese, A.M. A2B Adenosine Receptors: When Outsiders May Become an Attractive Target to Treat Brain Ischemia or Demyelination. *International Journal of Molecular Sciences* 2020, *Vol. 21, Page 9697* 2020, *21*, 9697, doi:10.3390/IJMS21249697.
251. Popoli, P.; Peponi, R. Potential Therapeutic Relevance of Adenosine A2B and A2A Receptors in the Central Nervous System. *CNS Neurol Disord Drug Targets* 2012, *11*, 664–674, doi:10.2174/187152712803581100.
252. Venturini, M.; Cherchi, F.; Santalmassi, C.; Frulloni, L.; Dettori, I.; Catarzi, D.; Pedata, F.; Colotta, V.; Varano, F.; Coppi, E.; et al. Pharmacological Characterization of P626, a Novel Dual Adenosine A2A/A2B Receptor Antagonist, on Synaptic Plasticity and during an Ischemic-like Insult in CA1 Rat Hippocampus. *Biomolecules* 2023, *13*, 894, doi:10.3390/BIOM13060894/S1.
253. Dettori, I.; Bulli, I.; Venturini, M.; Magni, G.; Cherchi, F.; Rossi, F.; Lee, H.; Pedata, F.; Jacobson, K.A.; Pugliese, A.M.; et al. MRS3997, a Dual Adenosine A2A/A2B Receptor Agonist, Reduces Brain Ischemic Damage and Alleviates Neuroinflammation in Rats. *Neuropharmacology* 2025, *262*, 110214, doi:10.1016/J.NEUROPHARM.2024.110214.
254. Semwal, B.C.; Garabadu, D. 5-N-Ethyl Carboxamidoadenosine Stimulates Adenosine-2b Receptor-Mediated Mitogen-Activated Protein Kinase Pathway to Improve Brain Mitochondrial Function in Amyloid Beta-Induced Cognitive Deficit Mice. *Neuromolecular Med* 2020, *22*, 542–556, doi:10.1007/S12017-020-08615-1/FIGURES/1.
255. Semwal, B.C.; Garabadu, D. Amyloid Beta (1–42) Downregulates Adenosine-2b Receptors in Addition to Mitochondrial Impairment and Cholinergic Dysfunction in Memory-Sensitive Mouse Brain Regions. *Journal of Receptors and Signal Transduction* 2020, *40*, 531–540, doi:10.1080/10799893.2020.1767136.
256. Camilleri, M. Leaky Gut: Mechanisms, Measurement and Clinical Implications in Humans. *Gut* 2019, *68*, 1516–1526, doi:10.1136/GUTJNL-2019-318427.
257. Ishioh, M.; Nozu, T.; Miyagishi, S.; Igarashi, S.; Funayama, T.; Ohhira, M.; Okumura, T. Activation of Basal Forebrain Cholinergic Neurons Improves Colonic Hyperpermeability through the Vagus Nerve and Adenosine A2B Receptors in Rats. *Biochem Pharmacol* 2022, *206*, 115331, doi:10.1016/J.BCP.2022.115331.
258. Ishioh, M.; Nozu, T.; Miyagishi, S.; Igarashi, S.; Funayama, T.; Ueno, N.; Okumura, T. Brain Histamine Improves Colonic Hyperpermeability through the Basal Forebrain Cholinergic Neurons, Adenosine A2B Receptors and Vagus Nerve in Rats. *Biochem Pharmacol* 2024, *224*, 116201, doi:10.1016/J.BCP.2024.116201.
259. Ishioh, M.; Nozu, T.; Igarashi, S.; Tanabe, H.; Kumei, S.; Ohhira, M.; Takakusaki, K.; Okumura, T. Activation of Central Adenosine A2B Receptors Mediate Brain Ghrelin-Induced Improvement of Intestinal Barrier Function through the Vagus Nerve in Rats. *Exp Neurol* 2021, *341*, 113708, doi:10.1016/J.EXPNEUROL.2021.113708.

260. Ishioh, M.; Nozu, T.; Miyagishi, S.; Funayama, T.; Ueno, N.; Takakusaki, K.; Okumura, T. Carnosine Improves Colonic Hyperpermeability through the Brain Histamine H1 Receptor, Basal Forebrain Cholinergic Neurons, Adenosine A2B Receptors and Vagus Nerve in Rats. *Eur J Pharmacol* 2025, *1002*, 177844, doi:10.1016/J.EJPHAR.2025.177844.
261. Baraldi, P.G.; Cacciari, B.; Romagnoli, R.; Merighi, S.; Varani, K.; Borea, P.A.; Spalluto, G. A3 Adenosine Receptor Ligands: History and Perspectives. *Med Res Rev* 2000, *20*, 103–128, doi:10.1002/(SICI)1098-1128(200003)20:2<103::AID-MED1>3.0.CO;2-X.
262. Moro, S.; Gao, Z.; Jacobson, K.A.; Spalluto, G. Progress in the Pursuit of Therapeutic Adenosine Receptor Antagonists. *Med Res Rev* 2006, *26*, 131–159, doi:10.1002/med.20048.
263. Tosh, D.K.; Deflorian, F.; Phan, K.; Gao, Z.; Wan, T.C.; Gizewski, E.; Auchampach, J.A.; Jacobson, K.A. Structure-Guided Design of A3 Adenosine Receptor-Selective Nucleosides: Combination of 2-Arylethynyl and Bicyclo[3.1.0]Hexane Substitutions. *J Med Chem* 2012, *55*, 4847–4860, doi:10.1021/jm300396n.
264. Tosh, D.K.; Padia, J.; Salvemini, D.; Jacobson, K.A. Efficient, Large-Scale Synthesis and Preclinical Studies of MRS5698, a Highly Selective A3 Adenosine Receptor Agonist That Protects against Chronic Neuropathic Pain. *Purinergic Signal* 2015, *11*, 371–387, doi:10.1007/S11302-015-9459-2/FIGURES/5.
265. Jacobson, K.A.; Merighi, S.; Varani, K.; Borea, P.A.; Baraldi, S.; Aghazadeh Tabrizi, M.; Romagnoli, R.; Baraldi, P.G.; Ciancetta, A.; Tosh, D.K.; et al. A3 Adenosine Receptors as Modulators of Inflammation: From Medicinal Chemistry to Therapy. *Med Res Rev* 2018, *38*, 1031–1072, doi:10.1002/med.21456.
266. Gao, Z.G.; Suresh, R.R.; Jacobson, K.A. Pharmacological Characterization of DPTN and Other Selective A3 Adenosine Receptor Antagonists. *Purinergic Signal* 2021, *17*, 737–746, doi:10.1007/S11302-021-09823-5/FIGURES/4.
267. Ohsawa, K.; Sanagi, T.; Nakamura, Y.; Suzuki, E.; Inoue, K.; Kohsaka, S. Adenosine A3 Receptor Is Involved in ADP-Induced Microglial Process Extension and Migration. *J Neurochem* 2012, *121*, 217–227, doi:10.1111/J.1471-4159.2012.07693.X.
268. Choi, I.Y.; Lee, J.C.; Ju, C.; Hwang, S.; Cho, G.S.; Lee, H.W.; Choi, W.J.; Jeong, L.S.; Kim, W.K. A3 Adenosine Receptor Agonist Reduces Brain Ischemic Injury and Inhibits Inflammatory Cell Migration in Rats. *Am J Pathol* 2011, *179*, 2042–2052, doi:10.1016/J.AJP.2011.07.006.
269. Terayama, R.; Tabata, M.; Maruham, K.; Iida, S. A3 Adenosine Receptor Agonist Attenuates Neuropathic Pain by Suppressing Activation of Microglia and Convergence of Nociceptive Inputs in the Spinal Dorsal Horn. *Exp Brain Res* 2018, *236*, 3203–3213, doi:10.1007/S00221-018-5377-1/FIGURES/7.
270. Coppi, E.; Cherchi, F.; Lucarini, E.; Ghelardini, C.; Pedata, F.; Jacobson, K.A.; Mannelli, L.D.C.; Pugliese, A.M.; Salvemini, D. Uncovering the Mechanisms of Adenosine Receptor-Mediated Pain Control: Focus on the A3 Receptor Subtype. *Int J Mol Sci* 2021, *22*, doi:10.3390/IJMS22157952.
271. Lillo, A.; Serrano-Marín, J.; Lillo, J.; Raïch, I.; Navarro, G.; Franco, R. Gene Regulation in Activated Microglia by Adenosine A3 Receptor Agonists: A Transcriptomics Study. *Purinergic Signal* 2024, *20*, 237–245, doi:10.1007/S11302-022-09916-9/FIGURES/4.
272. Fedorova, I.M.; Jacobson, M.A.; Basile, A.; Jacobson, K.A. Behavioral Characterization of Mice Lacking the A3 Adenosine Receptor: Sensitivity to Hypoxic Neurodegeneration. *Cell Mol Neurobiol* 2003, *23*, 431–447, doi:10.1023/A:1023601007518/METRICS.
273. Maria Pugliese, A.; Coppi, E.; Volpini, R.; Cristalli, G.; Corradetti, R.; Jeong, L.S.; Jacobson, K.A.; Pedata, F. Role of Adenosine A3 Receptors on CA1 Hippocampal Neurotransmission during Oxygen–Glucose Deprivation Episodes of Different Duration. *Biochem Pharmacol* 2007, *74*, 768–779, doi:10.1016/J.BCP.2007.06.003.
274. Pugliese, A.M.; Coppi, E.; Spalluto, G.; Corradetti, R.; Pedata, F. A3 Adenosine Receptor Antagonists Delay Irreversible Synaptic Failure Caused by Oxygen and Glucose Deprivation in the Rat CA1 Hippocampus in Vitro. *Br J Pharmacol* 2006, *147*, 524–532, doi:10.1038/SJ.BJP.0706646.
275. Farr, S.A.; Cuzzocrea, S.; Esposito, E.; Campolo, M.; Niehoff, M.L.; Doyle, T.M.; Salvemini, D. Adenosine A3 Receptor as a Novel Therapeutic Target to Reduce Secondary Events and Improve Neurocognitive Functions Following Traumatic Brain Injury. *J Neuroinflammation* 2020, *17*, 1–14, doi:10.1186/S12974-020-02009-7/FIGURES/7.
276. Borea, P.A.; Varani, K.; Vincenzi, F.; Baraldi, P.G.; Tabrizi, A.; Merighi, S.; Gessi, S. The A3 Adenosine Receptor: History and Perspectives. *PHARMACOLOGICAL REVIEWS Pharmacol Rev* 2015, *67*, 74–102, doi:10.1124/pr.113.008540.
277. Jacobson, K.A.; Müller, C.E. Medicinal Chemistry of Adenosine, P2Y and P2X Receptors. *Neuropharmacology* 2016, *104*, 31–49, doi:10.1016/j.neuropharm.2015.12.001.
278. Lillo, A.; Raïch, I.; Lillo, J.; Pérez-Olives, C.; Navarro, G.; Franco, R. Expression of the Adenosine A2A–A3 Receptor Heteromer in Different Brain Regions and Marked Upregulation in the Microglia of the Transgenic APPSw,Ind Alzheimer's Disease Model. *Biomedicines* 2022, Vol. 10, Page 214 2022, *10*, 214, doi:10.3390/B10MEDICINES10020214.
279. Li, S.; Geiger, N.H.; Soliman, M.L.; Hui, L.; Geiger, J.D.; Chen, X. Caffeine, Through Adenosine A3 Receptor-Mediated Actions, Suppresses Amyloid- β Protein Precursor Internalization and Amyloid- β Generation. *Journal of Alzheimer's Disease* 2015, *47*, 73–83, doi:10.3233/JAD-142223;WEBSITE:WEBSITE:SAGE;WGROU:STRING:PUBLICATON.

6. Conclusions

This thesis addresses the challenge of accelerating drug discovery through the prediction of ligand binding affinity, with a particular focus on molecular dynamics–based approaches capable of capturing the dynamic and thermodynamic determinants of molecular recognition. Binding affinity represents a central driver of drug efficacy and selectivity, yet its reliable prediction remains nontrivial. Although ligand-based and structure-based methods are indispensable for early-stage screening, their predominantly static or empirical nature limits their ability to describe conformational flexibility, solvent effects, and the dynamic stability of biomolecular complexes. By integrating methodological development with extensive applications to pharmaceutically relevant targets, this work demonstrates how advanced molecular dynamics simulations provide a more physically grounded framework, complementing, and, in specific contexts, overcoming the limitations of traditional drug design strategies.

The central contribution of this thesis is the development, validation, and systematic application of Thermal Titration Molecular Dynamics (TTMD) as a qualitative strategy for assessing biomolecular complex stability. When positioned with respect to existing approaches, TTMD occupies an intermediate and distinctive space between fast but static methods (such as docking and scoring functions) and computationally demanding quantitative techniques (such as free energy methods). Unlike docking-based approaches, TTMD explicitly accounts for time-dependent stability, protein and

ligand flexibility, solvent effects, and the progressive weakening of interactions under thermal stress. At the same time, it avoids the high computational cost and setup complexity typically associated with alchemical free energy calculations, making it particularly attractive as a post-docking refinement or prioritization tool.

Across a broad set of published studies, TTMD demonstrated consistent strengths: it reliably discriminated high- from low-affinity ligands, identified native-like binding poses among decoys, and proved robust across targets with diverse folds, binding site architectures, and ligand chemotypes. Its effectiveness was particularly evident in challenging scenarios where classical scoring functions are known to perform poorly, including fragment-based drug discovery, systems dominated by water-mediated interactions, and flexible or polyanionic targets such as nucleic acids. The extension of TTMD to RNA-small molecule, RNA-peptide, and DNA-ligand complexes represents a significant advancement, revealing the generality of stability-based criteria for molecular recognition beyond conventional protein-small molecule paradigms.

At the same time, the results obtained in this thesis allow a critical assessment of TTMD's limitations. By design, TTMD provides a qualitative or semi-quantitative estimate of complex stability rather than an absolute binding free energy. Its outcomes depend on the quality of the initial binding poses and on the definition of interaction fingerprints, which requires informed selection and careful interpretation. Moreover, while TTMD is computationally more efficient than rigorous free energy methods, it remains more demanding than docking and thus is not intended for ultra-large virtual screening campaigns. These considerations highlight that TTMD is most effective when strategically integrated within multistage computational workflows rather than used as a standalone solution.

In addition to TTMD, this thesis addressed quantitative binding affinity prediction through Free Energy Perturbation (FEP), with a particular focus on membrane proteins. The benchmarking study on GPCR targets provides valuable insights into the practical application of FEP in complex, heterogeneous environments. The analysis clearly shows that accurate ligand placement and explicit treatment of key water molecules are decisive factors for predictive reliability, often outweighing the impact of force field choice. These findings are especially relevant in the context of membrane proteins, where anisotropic environments, slow conformational dynamics, and buried solvation sites pose substantial challenges. While TTMD was not directly applied to GPCRs in this work, the conceptual framework developed here suggests that stability-based approaches could complement FEP by providing rapid qualitative insights into pose robustness and mutation or environment-induced destabilization prior to quantitative refinement.

Taken together, the results of this thesis demonstrate that TTMD and FEP should be viewed as complementary rather than competing methodologies. TTMD serves as an efficient, interpretable tool for stability assessment, pose validation, and ligand prioritization in early- and mid-stage drug discovery, particularly for difficult targets and unconventional ligands. FEP, in contrast, provides rigorous quantitative estimates suitable for late-stage optimization when chemical series are well defined. Their combined use enables a multiscale and rational strategy for binding affinity prediction, balancing computational cost, accuracy, and interpretability.

In conclusion, this work provides both methodological advances and practical guidelines for the use of molecular dynamics in computational drug discovery. By critically positioning TTMD within the broader landscape of available methods, extending its applicability to diverse biomolecular systems, and clarifying the strengths and limitations of dynamic approaches in complex environments such as membrane proteins, this thesis supports a shift toward more physically grounded, mechanism-aware, and strategically integrated computational pipelines. Such approaches are expected to play an increasingly central role in the rational, efficient, and cost-effective design of next-generation therapeutics.