



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova
Department of Chemical Sciences



Vrije Universiteit Amsterdam
Department of Theoretical Chemistry

Administrative seat: University of Padova
Doctoral Course in Molecular Sciences
Curriculum: Chemical Sciences
Cycle: XXXI
Coordinator: Prof. Dr. Leonard Jan Prins

Vrije Universiteit Amsterdam
Doctoral Programme in Chemistry

The Role of Selenium in Glutathione Peroxidase: Insights from Molecular Modeling

Supervisor: Prof. Dr. F. Matthias Bickelhaupt
Co-Supervisor: Prof. Dr. Laura Orian

Ph. D. Candidate: Marco Bortoli

This manuscript has been presented to jointly opt for the doctoral degree from
the University of Padova and the Vrije Universiteit Amsterdam

The Role of Selenium in Glutathione Peroxidase:
Insights from Molecular Modeling

*Alla mia famiglia
che mi ha fatto diventare quello che sono*

Contents

| | |
|---|-----------|
| Preface | xv |
| 1 Introduction | 1 |
| 1.1 A brief historical account | 1 |
| 1.1.1 Selenium and biology: from poison to prevention | 3 |
| 1.1.2 Glutathione peroxidase: the non-existing enzyme | 6 |
| 1.2 Selenocysteine: legacy or novelty? | 8 |
| 1.3 Glutathione peroxidase: Not only an antioxidant | 10 |
| 1.4 GPx mimics: harnessing the power of glutathione peroxidase | 14 |
| 2 Theory and methods | 19 |
| 2.1 Quantum mechanical calculations | 19 |
| 2.1.1 Density functional theory | 20 |
| 2.1.2 Activation strain model and energy decomposition analysis | 22 |
| 2.1.3 Solvent effects | 25 |
| 2.2 Classical calculations | 27 |
| 2.2.1 Molecular dynamics simulations | 28 |
| 2.3 Notation convention and software | 30 |

| | | |
|----------|---|-----------|
| 3 | The chalcogen-π interaction | 31 |
| 3.1 | Introduction | 31 |
| 3.2 | Methods | 33 |
| 3.3 | Results and discussion | 35 |
| 3.3.1 | Geometrical parameters | 35 |
| 3.3.2 | Bonding analysis | 37 |
| 3.4 | Conclusions | 46 |
| | Appendices | 49 |
| A | <i>Ab initio</i> benchmark | 51 |
| B | Orbital analysis | 55 |
| 4 | The GPx oxidative phase | 57 |
| 4.1 | Introduction | 57 |
| 4.2 | Methods | 61 |
| 4.2.1 | MD simulations | 61 |
| 4.2.2 | DFT Calculations | 62 |
| 4.3 | Results and discussion | 64 |
| 4.4 | Conclusions | 76 |
| | Appendices | 79 |
| C | Parameterization protocol | 81 |
| D | Benchmarks | 87 |
| 5 | Mimicking the GPx reductive phase | 91 |
| 5.1 | Introduction | 91 |

| | | |
|---|--|------------|
| 5.2 | Methods | 94 |
| 5.3 | Results and discussion | 95 |
| 5.3.1 | General gas-phase profiles | 96 |
| 5.3.2 | Trends in gas-phase reactivity | 102 |
| 5.3.3 | Effects of solvation | 104 |
| 5.3.4 | Consequences for undesirable thiol exchange reactions . . | 109 |
| 5.4 | Conclusions | 111 |
| Appendices | | 113 |
| E Torsional barriers | | 115 |
| 6 Oxidation of organoselenides mimics of GPx by H₂O₂ | | 119 |
| 6.1 | Introduction | 119 |
| 6.2 | Methods | 120 |
| 6.3 | Results and discussion | 124 |
| 6.3.1 | Oxidation of diselenides and ditellurides by H ₂ O ₂ | 124 |
| 6.3.2 | Activation strain analysis for the oxidation of RXXR (R=H, CH ₃ , Ph) by H ₂ O ₂ | 126 |
| 6.3.2.1 | Effect of the chalcogen | 128 |
| 6.3.2.2 | Effect of the substituent | 132 |
| 6.3.3 | Redox isomerization to anhydride | 135 |
| 6.4 | Conclusions | 139 |
| Appendices | | 141 |
| F Alternative mechanism | | 143 |
| G Energy decomposition analysis | | 145 |

| | | |
|----------|--------------------------------|------------|
| 7 | Conclusions | 149 |
| 7.1 | Summary | 149 |
| 7.2 | Sommario | 153 |
| 7.3 | Concluding remarks | 157 |
| 7.4 | Acknowledgments | 159 |
| 7.5 | List of Publications | 160 |
| | References | 161 |

List of Figures

| | | |
|-----|---|----|
| 1.1 | Excerpt from the letter from Berzelius to Berthollet, front cover and excerpt from the first scientific publication where selenium is described and named. | 4 |
| 1.2 | Mechanism of organic hydroperoxides reduction catalyzed by Cys/Sec-GPx. | 11 |
| 1.3 | Structure of human GPx4 and of the tetramer of human GPx1 with the four monomers in different colors. | 12 |
| 1.4 | Mechanism for the reduction of HO by an organic thiol in the catalytic cycles of ebselen and 2-(N,N-(dimethylamino)-methyl)benzenediselenide. | 16 |
| 2.1 | ASM model applied to an S _N 2 model reaction | 24 |
| 3.1 | Examples of the optimized structures of four complexes: F₂O...2but , Cl₂S...et , Br₂Se...et and I₂Te...ac | 36 |
| 3.2 | HOMOs of the unsaturated substrates. | 43 |
| 4.1 | Mechanism of organic hydroperoxides reduction catalyzed by Cys/Sec-GPx. Oxidative step highlighted. | 58 |

| | | |
|-----|---|-----|
| 4.2 | Cluster of amino acids extracted from Cys-GPx and superposition of the optimized geometries of the Cys-, Sec- and Tec-GPx clusters (b). | 63 |
| 4.3 | MD results: position of the backbone atoms of Cys-GPx, Sec-GPx and Tec- GPx after 500 ns and root mean square fluctuation for each residue along the 500 ns dynamics. | 65 |
| 4.4 | Geometries of the optimized adducts E-H ₂ O ₂ ·H ₂ O and RMSDs with respect to the initially optimized clusters. | 67 |
| 4.5 | Mechanism of formation of tellurinic acid via a direct pathway and a hydroxy perhydroxy tellurane intermediate. | 71 |
| 4.6 | Optimized Tec-cluster in the tellurinic acid form and relevant interatomic distances. | 72 |
| 4.7 | Different partitioning of the Cys and Sec clusters. | 76 |
| C.1 | Atom numbering scheme of the parametrized residues: Sec and Tec. | 82 |
| C.2 | Fitting obtained with Paramfit of energies of the structures generated with dihedral scans for Sec and Tec. | 84 |
| 5.1 | Mechanism of organic hydroperoxides reduction catalyzed by Cys/Sec-GPx. Second reductive step highlighted. | 92 |
| 5.2 | Model reactions scheme. | 93 |
| 5.3 | Structures of the stationary points for reactions S⁻ + SSe and S⁻ + SeS | 96 |
| 5.4 | Solvent effect on reaction energy profiles in the gas-phase and in water for nucleophilic attack at tellurium selenium and sulfur. . . | 105 |
| 5.5 | Gas phase and condensed phase (water) energy profiles for the reactions S⁻ + SS and S⁻ + SSe | 108 |

| | | |
|-----|--|-----|
| 5.6 | Molecular structures of ebselen and selenenyl sulfide as it appears in the ebselen catalytic cycle. | 110 |
| E.1 | Dihedral scans of CH_3SSCH_3 , $\text{CH}_3\text{SeSeCH}_3$ and $\text{CH}_3\text{TeTeCH}_3$ <i>in vacuo</i> at the scalar ZORA-OLYP/TZ2P level of theory. | 116 |
| E.2 | Dihedral scan of the TC of the $\text{S}^- + \text{SS}$ reaction <i>in vacuo</i> at the scalar ZORA-OLYP/TZ2P level of theory. | 117 |
| 6.1 | Oxidation of 2-(N,N-(dimethylamino)-methyl)benzenediselenide by H_2O_2 followed by product isomerization to anhydride. | 121 |
| 6.2 | Reaction mechanism for the direct oxidation and subsequent possible redox isomerization of diphenyl diselenides and ditellurides. | 124 |
| 6.3 | $(\text{PhSe})_2$ optimized structure. | 125 |
| 6.4 | Reactant complex (RCox), transition state (TSox) and product complex (PCox) structures for the oxidation of $(\text{PhSe})_2$ | 127 |
| 6.5 | Activation strain analysis along the reaction path for the oxidation of $(\text{HX})_2$, without and with empirical correction for dispersion. | 130 |
| 6.6 | Activation strain analysis along the reaction path for the oxidation of $(\text{CH}_3\text{X})_2$ and $(\text{PhX})_2$ | 131 |
| 6.7 | Examples of the “closed” and “open” structures found in aryl dichalcogenides. | 133 |
| 6.8 | Stationary points and transition state structures for the isomerization of $(\text{PhSe})_2$ | 137 |

List of Tables

| | | |
|-----|--|----|
| 3.1 | Chalcogen-carbon distance and Voronoi deformation densities on chalcogen atom. | 38 |
| 3.2 | Activation strain analysis of the model systems. | 39 |
| 3.3 | Energy decomposition analysis of the model systems. | 40 |
| 3.4 | HOMO-LUMO energy gap and overlap in selected model systems. | 42 |
| 3.5 | HOMO-LUMO energy gap , and overlap in selected model systems. | 43 |
| 3.6 | HOMO-LUMO energy gap and overlap in selected model systems. | 45 |
| A.1 | Chalcogen- π bond energies computed at DFT and CC levels of theory. | 52 |
| A.2 | Chalcogen- π bond energies computed at DFT and CC levels of theory. | 53 |
| B.1 | HOMO-LUMO energy gap and overlap in all model systems. . . . | 56 |
| 4.1 | Gibbs free energies of the intermediates and transition states in gas phase and in condensed phase. | 68 |
| 4.2 | Gibbs free energies of the intermediates and transition states for the oxidation and isomerization of the Tec-cluster in gas phase and in condensed phase. | 70 |

| | | |
|-----|---|-----|
| 4.3 | Gibbs free energies for the two possible mechanisms for tellurinic acid formation. | 71 |
| 4.4 | Activation strain analysis and energy decomposition analysis for Cys and Sec clusters. | 75 |
| C.1 | RESP charges derived for Sec and Tec. | 82 |
| C.2 | New force field parameters for Sec and Tec residues. | 85 |
| D.1 | Activation strain analysis for Cys and Sec clusters. | 88 |
| D.2 | Activation strain analysis and energy decomposition analysis for Cys and Sec clusters. | 89 |
| 5.1 | Energies of stationary points of the model reactions in gas phase. | 97 |
| 5.2 | Selected interatomic distances, activation strain analysis and energy decomposition analysis for all model reactions | 99 |
| 5.3 | Selected interatomic distances and energies of stationary points of the model reactions in water. | 106 |
| 6.1 | Structure and naming scheme of the studied disulfides, diselenides and ditellurides. | 122 |
| 6.2 | Relative energies of stationary points for the oxidation of RXXX by H ₂ O ₂ | 127 |
| 6.3 | Relative energies of the stationary points for the oxidation of substituted diselenides and ditellurides by H ₂ O ₂ | 134 |
| 6.4 | Relative energies and activation strain analysis for the first oxidation of substituted diphenyl diselenides and diphenyl ditellurides by H ₂ O ₂ | 136 |
| 6.5 | Relative energies for the isomerization of the selenoxides and telluroxides. | 138 |

| | | |
|-----|---|-----|
| F.1 | Energies and dihedral amplitude of stationary points for the oxidation of (PhSe)₂ and (PhTe)₂ in two proposed mechanisms. . . | 143 |
| G.1 | Energy decomposition analysis for the first oxidation of model diphenyl disulfides. | 145 |
| G.2 | Energy decomposition analysis for the first oxidation of model diphenyl diselenides. | 146 |
| G.3 | Energy decomposition analysis for the first oxidation of model diphenyl ditellurides. | 147 |

Preface

Why did Nature choose selenium? This apparently simple question fascinated me since the beginning of my academic career. It concerns the reason why some proteins contain, in place of sulfur, the biologically rare element selenium which gives them some advantageous properties. The reason seems quite simple, as stated, selenium is more advantageous than sulfur. But here is the catch: selenium is extremely complicated to insert into the protein structure. Therefore, in the end, is it worth it?

It all started during my bachelor degree, when I decided to write an essay on nucleophilic substitution at organochalcogenides. Since then, the main focus of my academic research, which culminated during my Ph.D. student years in Padova (mostly) and Amsterdam, was always centered around this peculiar element: selenium. I had never imagined that it would accompany me from the first time I entered my then would-be supervisor's office as an undergraduate, to this day as I approach a Doctoral degree. Through the study of its chemistry I have understood how a chemical element can be both toxic and essential, that the most advantageous attribute of a molecule can also be its Achille's heel and that sometimes, also in chemistry, appearances can be deceiving and what is at a first glance an easy matter may reveal to be much more complicated.

But chemistry was only a small (albeit essential) part of my Ph.D. I was

taught many other lessons during this “adventure”, some about chemistry others about life. I had the guidance of brilliant teachers that helped me progress in my research by giving me invaluable scientific advice as well as by boosting my morale, whenever needed. I also learned to guide less experienced students to achieve their first goals as young scientists. On the other hand, I also had to overcome the natural difficulties that arise when working with people with a completely different mindset and baggage of experience than your own. Some might say it is not a rewarding experience: most of the times it is surely true, but each small achievement I was able to obtain was worth the time and effort. There were good times, there were bad times, there were “meh” times and there were times in which I could not tell the one from the other because I was too busy to bother. All in all it was a marvelous journey.

All these experiences may seem inconsequential, but it is thanks to them that I am now writing this Thesis. Their presence may not transpire from the Chapters you are going to read but I can assure you they are there, as if they all had a little part in the writing of this Thesis, which is not just the presentation of some scientific results but it is the condensation, the quintessence of the windy path that made me become, let me say it, *a scientist*.

It might be the first stone of a brilliant career in academia or it might represent a turning point towards another destination.

*Tu ne quaesieris, scire nefas, quem mihi, quem tibi
finem di dederint...*¹

I'll leave you to the interesting part.

Marco Bortoli
Padova, November 7, 2018

¹This is the beginning line of Horace's *Carpe Diem* which translates to “Do not ask, it is forbidden to know, what fate is reserved for me and you”

Chapter 1

Introduction

1.1 A brief historical account

STOCKHOLM, 1818. On a chilly morning in the Swedish Winter Jöns Jacob Berzelius is writing a letter to his colleague, chemist Claude Louis Berthollet¹.
[1]

“I owe you many thanks for your obliging letter of July 21. I delayed answering it because I wanted to share with you some results of the research done in Sweden in our favorite science. This time I have something interesting to tell you; it is the discovery of a metallic substance, the oxide of which is a new fixed alkali, and that of another acidifiable metallic substance, more analogous to sulfur, than to any other substance [...].”

Two new substances were discovered in Berzelius’ lab during the previous year. The former the scientist refers to, is what he will name *lithium*. But it is in the

¹The quoted text of the letter, originally in French, has been translated by the author of this thesis.

latter that my personal interest grew over the last few years and around which this Thesis revolves. The discovery of the new substance was, as in many other cases, accompanied by a fair dose of serendipity:

“[...] The acidifiable metallic substance was discovered in the following manner: In a sulfuric acid factory, where sulfur is removed from the pyrites of the Fahlun mine, on the bottom of the large lead chamber a reddish mass is deposited which consists mainly of sulfur. [...] the reddish precipitate struck us. We examined it, and finding that it gives on burning a very strong smell of horseradish (*raphanus sativus*), we thought it possible to conclude that the precipitate in question was a mixture of sulphuret of tellurium (*tellurium sulfide*) with sulfur. However we could not extract any tellurium. I took a small quantity with me in Stockholm, where I examined it more closely. I found then that this sulfur contained a foreign substance, very volatile, very easily reducible, but which can not be precipitated by alkalis, and I succeeded, after a few unsuccessful attempts, in isolating this substance. [...]”

After a description of the appearance and physical properties of the new material, Berzelius strongly refutes the possibility that the substance contained in the red mass is tellurium:

“[...] in contrast, if the flame of a candle is directed there, or if it is blown with a blow-torch, it colors the flame of a beautiful azure blue color, diffusing a smell of horseradish [...]. However, neither the purified tellurium, nor its oxide, nor its combinations with the metals, produce this odor. It was only when I locked a piece of tellurium in a small ball of thin glass, and blowing on it with the blowtorch until the gasified tellurium made a hole in the softened glass, that I could

produce the odor in question: it was then entirely the same as that of the new substance.”

Nonetheless, it is this strong relationship with tellurium which prompted the scientist to choose its name:

“However, to recall the relationship of the latter with tellurium, I named it *selenium*.”

The following year the first publication in a scientific journal on selenium appeared in the Swedish *Afhandlingar i Fysik, Kemi och Mineralogi* (Theses in Physics, Chemistry and Mineralogy, Figure 1.1) and the chemical history of selenium began. A history which has still many unwritten chapters and unfinished paragraphs...

1.1.1 Selenium and biology: from poison to prevention

Berzelius had already understood that selenium could be toxic, as he verified when investigating health problems at his sulfuric acid production plant. In fact, as he changed the ores used in the process, which contained only trace amounts of selenium, the health of the workers greatly improved. In the following century, the interest on selenium was very low and almost no further research was conducted on the element. The only notable use was in the first solar cells built in New York by Charles Fritts in the early 1880s. [2] But it was in the late 1930s that selenium gathered the spotlight again, and for a second time it was because of its toxicity. It was recognized as an industrial hazard [3, 4] and as a the cause of the insurgence of various cattle and poultry diseases. [5–7] Moreover, in addition to the already established toxicity of selenium, it was also observed that the element could have teratogenic effects in birds [8] and possibly in human beings. [9, 10]

Only in the mid 50s, the group of Jane Pinsent discovered the first “benign” function of selenium. [11] Studying the bacterium *Escherichia Coli*, Pinsent found

LETTRE de M. Berzelius à M. Berthollet sur deux
Métaux nouveaux.

Stockholm, le 9 février 1818.

« Je vous dois bien des remerciemens pour votre obligeante lettre du 21 juillet dernier. J'ai retardé à y répondre pour avoir à vous communiquer quelques résultats des recherches faites en Suède dans notre science favorite. Pour cette fois, j'en ai de bien intéressans à vous mander ; c'est la découverte d'une substance métallique dont l'oxide est un nouvel alcali fixe, et celle d'une autre substance métallique acidifiable et plus analogue au soufre qu'à tout autre corps.

» Le nouvel alcali a été découvert par M. Arfvedson, jeune chimiste très-habile, qui depuis un an travaille dans mon laboratoire. Il a trouvé cet alcali dans une

pand cette même odeur. Cependant, ni le tellure purifié, ni son oxide, ni ses combinaisons avec les métaux, ne produisent cette odeur. Ce n'est que lorsque j'ai enfermé un morceau de tellure dans une petite boule de verre mince, et soufflant dessus avec le chalumeau jusqu'à ce que le tellure gazéifié fit un trou dans le verre amolli, que j'ai pu produire l'odeur en question : elle était alors entièrement la même que celle de la nouvelle substance. Je ne déciderai point si cette odeur leur appartient à tous les deux, ou si le tellure est souvent accompagné de la substance nouvelle. Cependant, pour rappeler les rapports de cette dernière avec le tellure, je l'ai nommée *selenium*.

(a) Afhandlingar
i
Fysik, Kemi
och
Mineralogi.

Utgifne

af

J. AFZELIUS, N. W. ALMROTH, A. ARFVEDSON,
J. BERZELIUS, H. P. EGGERTZ, J. AF FORSELLES,
J. G. GAHN, W. HISINGER, P. LAGERHJELM,
M. AF PONTIN, E. ROTHOFF, N. G. SEFSTRÖM,
P. STRÖM, E. T. SVEDESTIERNA,
P. WALMSTEDT.

— O — O —
SJETTE DELEN.

STOCKHOLM,

Tryckte hos Dir. H. A. NORDSTRÖM,
1818.

(b) Det bruna ämnet, som vid ammoniak-
falternas sönderdelning afskilt sig, blef nu
ett föremål för undersökningen, och befanns
genom de försök, som i det följande skola
beskrivas, vara en egen, hittills okänd,
brännbar mineral kropp, hvilken jag, för
att utmärka dess släktkap i egenkapet med
tellurium, kallat Selenium, af Σελήνη, måna.
Den ligger för öfrigt i detta hänföende midt
VI Del. D emel.

Figure 1.1: Excerpt from the letter from Berzelius to Berthollet (left) and front cover (a) and excerpt from the first scientific publication where selenium is described and named (b), right. These resources are freely available through the Internet via Google Books.

that trace amounts of selenite were necessary for the enzyme formic acid dehydrogenase to work properly. It was already 1957 when two groups independently proved that selenium was essential also to animals. [12, 13] In particular Schwarz's group was working on identifying micronutrients that are essential to life. In their studies, they fed mice a diet consisting only of sugar and yeast. The observed effect was that most of the specimens ended up dying of liver necrosis. This condition could be prevented if a supplement of methionine or vitamin E was added. Moreover, if a particular kind of yeast (American brewer's yeast) was used in the diet these pathological effects did not occur. Therefore, Schwarz concluded that there was another substance that could prevent the disease and called it "Factor 3". [14] An insightful suggestion by Dr. DeWitt Stetten, [15, 16] prompted Schwarz to check the content of selenium in a sample of Factor 3. Indeed, not only did the sample contain selenium, but it was also proved that it was that particular element that prevented liver necrosis in rats. [17]

The discovery of the essential role of selenium in biology contributed to an increased interest in the subject that resulted in the identification of some human diseases caused by selenium deficiency. The first one to be discovered was the so called Keshan disease, which is a cardiomyopathy. [18] Some time later also the Kaschin-Beck disease (a disorder similar to rheumatoid arthritis) and myxoedematous cretinism were related to selenium deficiency, although that not being the only cause for the insurgence of these diseases. These findings confirmed that the element was essential also to humans. Moreover, since the discovery of the first selenium-caused disease, research was conducted also to see if selenium supplementation could be related to the incidence of cancer. From the very rich literature on the subject (for an exhaustive list of references see [19] and references therein), a quite homogeneous perspective emerged, which seemed to favor the conclusion that a low intake of selenium resulted in a higher cancer incidence. This so called "selenium-cancer" hypothesis was further expanded by a

large-scale study by Clark and Combs which showed that $200 \mu\text{g day}^{-1}$ of selenium resulted in a significant reduction in colon, prostate and lung cancer. [20] On the wave of this promising result, the US National Health Institute promoted another trial on an even larger scale. [21] Unfortunately, the results of Clark and Combs were not confirmed in this study. In fact, no clear evidence that could relate selenium supplementation with cancer prevention was found. [22]

The different outcomes of the two major studies on the relationship between selenium and cancer did not contribute to establish a clear role of the element as a protective agent. In addition, research on some selenoproteins showed how their overexpression could help promote cancer growth. [23–25] These factors, combined with the narrow span between the amount of selenium needed by the organism ($26 \mu\text{g day}^{-1}$ to $35 \mu\text{g day}^{-1}$ [26]) and the threshold above which it becomes toxic ($900 \mu\text{g day}^{-1}$ [26]), make selenium a multi-faceted element which was aptly named “the essential poison”. [27]

1.1.2 Glutathione peroxidase: the non-existing enzyme

Among the various proteins that make use of selenium in the human body, the family of glutathione peroxidases (GPx) is one of the most studied and well-characterized. A brief overview of the different classes of selenium compounds found in the human body will be given in the next Section. In this sub-section the steps that led to GPx discovery and characterization are outlined. For a more exhaustive reading on the subject, the reader is referred to a review by L. Flohé (which contains also funny anecdotes and experiences of the author). [16]

The first evidence for GPx is found in a paper of 1957 by Mill that states:

“Studies with hydrogen peroxide indicate that this enzyme catalyzes the oxidation of reduced glutathione by hydrogen peroxide, and thus may be termed a glutathione peroxidase.”

At that time not only was the presence of selenium in GPx still unknown, but a few years after Mill's paper, part of the scientific American community deemed the existence of GPx very unlikely. [16] Nonetheless, in Europe, research on the subject continued, albeit in a modest way, and culminated in 1971 with the work of Flohé. Starting from cow blood, he managed to extract, purify and characterize a compound that was confirmed to be GPx. [28, 29] The first studies on the activity and mechanism of GPx as a peroxidase brought about by these new findings, although published in a review in German, [29] reached the group of William Hoekstra in the US. They combined these discoveries with their work on the ability of selenium to prevent oxidative damage in selenium deficient rats [30] and managed to discover a direct relationship between selenium and GPx activity. [31] If this evidence proved unequivocally the need for selenium in the catalytic mechanism of GPx, it was still not clear if the element was present in the protein itself or if it acted as a co-factor in another form. The dilemma was conclusively resolved by Flohé who managed to detect Se in a GPx sample that was first subjected to neutron bombardment and then was analyzed with a gamma spectrometer. The analysis undoubtedly confirmed the presence of selenium in GPx in the amount of 1 atom per subunit of the enzyme. [32] In the following years, the presence of a selenium-containing GPx was discovered also in sheep, rats, fish and humans. [33–37] With these findings, the importance of selenium in GPx was finally recognized by the whole scientific community, which, from that point onward, dedicated much effort to clarify the role of selenium in GPx and other selenium-containing proteins, trying to explain the actual benefits deriving from its presence as opposed to the more common sulfur.

1.2 Selenocysteine: legacy or novelty?

THE discovery of selenium in proteins was a great achievement. However, as with all scientific discoveries, it clarified many doubts but opened many more questions. First of all, how is selenium inserted into proteins? It took three years to solve this issue and the first conclusive evidence on the structure of a protein-bound amino acid containing selenium was published only in 1976. [38] It was unambiguously showed that the moiety that contained the element was a cysteine residue in which sulfur was substituted by selenium (hence selenocysteine). Another ten years had to pass before it was eventually recognized that selenocysteine (Sec) could be rightfully called the 21st amino acid. [39]

The major contributions that fueled the recognition of Sec as a proteinogenic amino acid were the discoveries that two genes, coding for two different selenoproteins, contained an in-frame TGA termination codon. [40, 41] This led to the novel idea that the codon UGA, which normally is read as a stop codon during translation, could, under “particular circumstances” signal the insertion of Sec. These “particular circumstances” were seen to be a complicated machinery involving multiple co-factors that are needed to differentiate the meaning of the UGA codon. [42] While all the steps building up the process of Sec insertion were being elucidated, the prevailing ideas about the presence of selenium in biology leaned towards the opinion that it was a legacy of the anaerobic world. [43, 44] The complexity and costliness of the reactions involved in the process clearly supported this view, which was corroborated by the higher than sulfur sensitivity towards oxygen.

Later on, studies on protein homology between prokaryotes and eukaryotes led to the abandonment of that theory in favor of a new point of view: selenium is an evolutionary sign that differentiates mammals from lower eukaryotes and it has a unique and specific role and a particular advantage over sulfur that enables it

to enhance the enzymatic reactions in which it is involved. [45] This consideration seemed to be the natural result from all the studies revolving around the mutation of Sec to Cys in selenoenzymes or vice versa. [46–49] However there have been a few instances in which the catalytic role of selenium was difficult to justify. [50, 51] From these seemingly controversial results yet another possibility for the role of selenium was proposed. If the presence of selenium was still viewed as an evolutionary sophistication, the main advantage of selenium over sulfur was ascribed not to its superior catalytic activity but to its more favorable redox properties. In other terms, due to its intrinsic atomic properties, selenium is better suited to efficiently sustain the high oxidative stress levels present in most of the environments in which selenoenzymes work. [19, 44]

Despite the limited number of reactions in which selenium is involved, which is a result of evolutionary pressure, some advantage over sulfur is expected, which, in theory, could replicate the physiological role of selenium. After 200 years since its discovery, the knowledge of the function of selenium inside the cell is still not exhaustive. Selenocysteine shows different chemical properties when compared to cysteine: lower pK_a (5.2 vs 8.3) [52], superior electron acceptor and leaving group ability, [44] but these differences do not clearly explain why in some cases the choice of nature fell on selenium rather than sulfur. [53] Moreover, in some complex eukaryotes it has been seen that the replacement of Sec with Cys causes no appreciable activity loss, [51] suggesting that a deeper investigation on the chemical, biochemical and biological role of selenium is still needed.

Thus, far from being completely understood, the presence of selenium in biology with its complex insertion machinery relies on different hypotheses. Lately, the idea of a superior redox ability and a higher resistance to overoxidation seems to provide a convincing argument to disentangle this ongoing debate.

1.3 Glutathione peroxidase: Not only an antioxidant

SINCE its status of an actual, real enzyme had been confirmed (see 1.1.2), glutathione peroxidase has stirred a consistent interest in the scientific community of selenophiles.² The term glutathione peroxidase is however somewhat limiting because over the years a whole family of proteins with similar properties was discovered. [54, 55]

GPx enzymes catalyze the reduction reaction of hydrogen peroxides and organic hydroperoxides by glutathione (GSH). [56] The GPx enzymatic mechanism involves three steps (Figure 1.2): the oxidation of selenol (E) to selenenic acid (F), accompanied by the reduction of the hydroperoxide; thereafter, two equivalents of GSH are consumed in the two subsequent reductive steps to form first a selenenylsulfide intermediate (G) and then the regenerated enzyme in its starting selenol form together with oxidized glutathione (GSSG). [57, 58] The kinetics of this mechanism were seen to be more complicated than a simple Michaelis-Menten process in which an overall kinetic constant can be defined. [59] The cycle can be divided in two parts: the first, which corresponds to step I in Figure 1.2, in which the enzyme perform its peroxidatic role reducing the hydroperoxide and the second, which comprises steps II and III of Figure 1.2, in which GPx is reduced by GSH to its resting state. Although the kinetic constant for the reductive part (k_2) is usually 2 to 3 orders of magnitude smaller than that of the oxidative step (k_1), the much higher concentration of GSH *in vivo* compared to peroxide usually makes the overall reaction independent of the concentration of the reducing agent.

There are at least eight types of GPx that have been characterized and they originated from three evolutionary groups that have a common Cys-containing

²A search in the Web of Science database for records from 2010 to 2018 containing the word “glutathione peroxidase” returned about 18000 results.

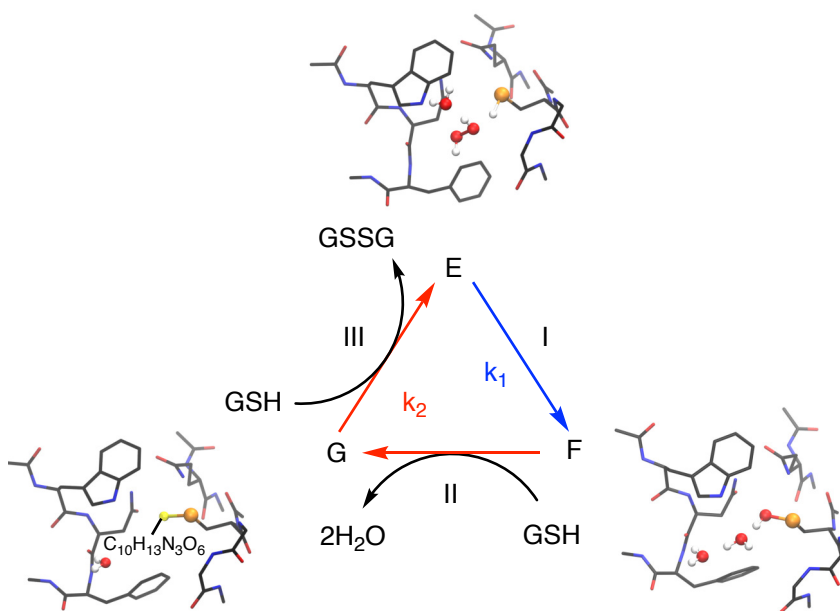


Figure 1.2: Mechanism of organic hydroperoxides reduction catalyzed by Cys/Sec-GPx. E is the reduced enzyme with Cys/Sec in the thiol/selenol form, F represents the oxidized intermediate, *i.e.* the sulfenic/selenenic acid form, G is the disulfide/selenosulfide form. The first (oxidative) part is highlighted in blue and the second (reductive) part in red.

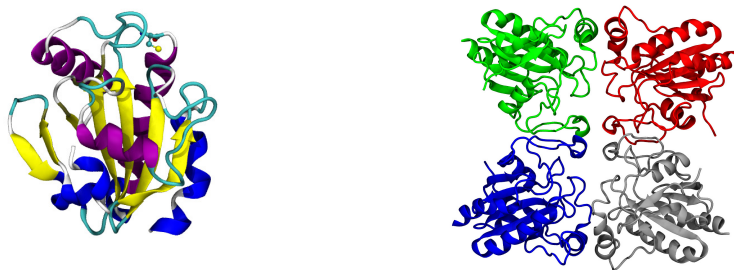


Figure 1.3: Structure of human GPx4 (left) and of the tetramer of human GPx1 with the four monomers in different colors (right). Structures taken from the RCSB PDB database (www.rcsb.org).

ancestor. [60, 61] Four of these proteins (GPx1 through GPx4) contain selenium whereas GPx6 is a selenoprotein only in human and the others (GPx5, GPx7 and GPx8) contain sulfur. A conserved structure of the active site was found throughout the family [62] consisting of four amino acids that form the so-called “catalytic tetrad” (Sec, Gln, Trp and Asn) with only 3 exceptions. [60] The different variety of GPx can be found as homotetramers (as GPx1, 2, 3, 5 and 6, Figure 1.3, right) or as monomers (GPx4, 7 and 8, Figure 1.3 left) and this might steer the reactivity towards different substrates [55]: tetrameric structures have a higher affinity for small soluble peroxides whereas monomers can react also with more complex lipid hydroperoxides.

Among the selenium containing GPxs, GPx1 (also called cytosolic GPx) is tetrameric and can reduce a wide variety of organic hydroperoxides but is highly specific for glutathione as a reducing substrate. [63] GPx2 (also known as gastrointestinal GPx) is also tetrameric and as a high chain similarity (65%) with GPx1 and similar substrate specificity. [64] GPx3 is a glycoprotein found mainly in human plasma. Its role as an extracellular antioxidant has been questioned due

to the low concentration of GSH in plasma, though it seems it can function in conjunction with thioredoxin and glutaredoxin as electron donors. [65] Finally, GPx4 (or phospholipid hydroperoxide GPx), [58, 66] is instead monomeric and its peculiarity is that it can be reduced by a variety of substrates other than glutathione. An overview of the specific functions of all the GPx types is beyond the scope of this work, however a brief description of the role of GPx4, which has been used as model in this study, will be presented.

GPx4 exists in three isoforms: cytosolic (cGPx4), mitochondrial (mGPx4) and sperm nuclear (snGPx4) and was initially recognized as a lipid peroxidation inhibiting protein. [67] Moreover, it was seen that, in case of GSH deficiency, other cysteine residues could be used as reducing substrates to restore the enzyme. Further studies revealed that GPx4 might have functions other than that of an antioxidant. It is present in spermatozoa as an inactive structural protein [54] and has a crucial role in regulating brain development and neuronal function. [68] In addition, it was observed that a knockout of cGPx4 is lethal in embryos and, when induced, in adult animals. Hypotheses have been formulated on the origin of the essentiality of GPx4. The most creditable is that the enzyme has a central role in the regulation of the activity of 12,15-lipoxygenase. [55] 12,15-lipoxygenase is an enzyme that catalyzes hydroperoxide formation in cellular membranes regulating different signaling pathways, among which apoptosis.³ GPx4 provides the required balance in lipid peroxidation which would be completely disrupted in its absence causing eventually cell death.

These ulterior *raison d'être* greatly widen the role GPx4 has in biology. However, this Thesis will not deal with them, primarily focusing on the (historically) classical role of the enzyme as a preventer of high oxidative stress levels through the reduction of different peroxides.

³Apoptosis is a form of programmed cell death that occurs in multicellular organisms.

1.4 GPx mimics: harnessing the power of glutathione peroxidase

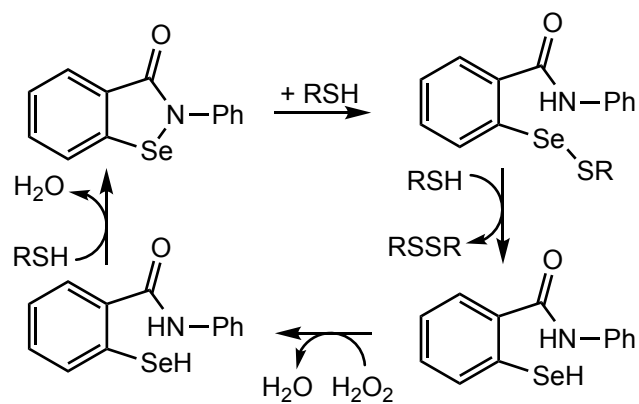
THE knowledge of the GPx structure and its reaction mechanism encouraged the scientific community to try and replicate the antioxidant properties of the enzyme in small easily synthesizable molecules. These compounds would then be used to control the oxidative stress inside the cellular environment. The study on the pharmacological effects of potential GPx mimics started inevitably with molecules containing selenium.

The first structure that was seen to display some kind of peroxidase activity was 2-phenyl-1,2-benzisoselenazol-3(2H)-one, more commonly known as ebselen. [69, 70] The confirmation of the similarity of its mechanism with that of GPx promoted many studies on similar molecules with a high activity for the reduction of hydroperoxides. [71] This class of compounds is classically divided into three subgroups consisting of (i) cyclic selenenyl amides, (ii) diaryl diselenides and (iii) aromatic or aliphatic monoselenides. [72] The steps of their catalytic mechanism (in Figure 1.4 are shown those for classes (i) and (ii)) show a different order than those of GPx and mimics with comparable reaction rates have not been discovered yet. [73]

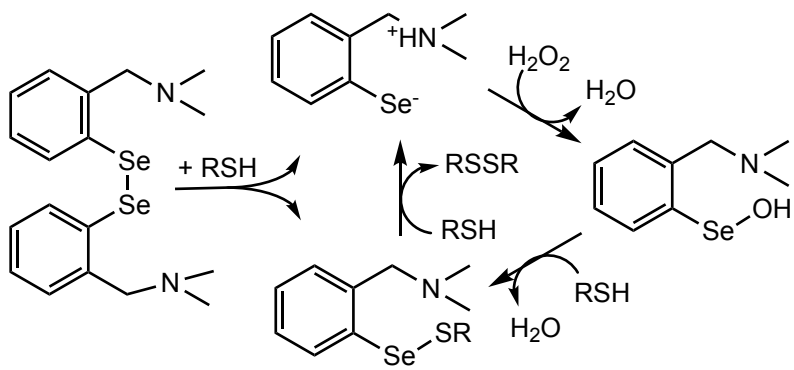
Numerous organoselenides have been synthesized and tested as potential antioxidant drugs, [72, 74, 75] highlighting the prominent role of diaryl diselenides, especially in preventing liver damage caused, for example, by a deficiency in vitamin E. [17] Therefore, intensive focus was centered on this class of compounds to assess the possible uses of diaryl diselenides as therapeutic agents. Evaluating the pharmacological advantages and drawbacks of organoselenium compounds is beyond the scope of this Thesis. However, the reader can refer to the literature for an exhaustive view on the subject. [75–78] The key feature of these mimics is the presence of a Se–Se bond which can be broken upon oxidation, resulting

in the formation of two catalytic sites. Although theoretical mechanistic studies are still fragmentary, [73, 79] evidence show that the first step in the GPx-like cycle involves the cleavage of the Se–Se bond after reaction with a thiol and the formation of a selenyl sulfide and selenol (Figure 1.4). [80] In this particular case, the zwitterionic form of the selenol is stabilized. A computational study of the direct reaction of the same compound with H₂O₂, showed how this pathway is disfavored as it results in the formation of a selenolseleninate with a highly distorted structure. [81] The product of this direct oxidation has not been completely clarified yet: experimental studies on ebselen [69, 70, 82] suggest that, after dimerization, the reaction of its diselenide with H₂O₂ results in the formation of a selenenic anhydride. [83, 84] On the other hand, *in silico* studies lean towards the initial formation of a selenolseleninate. [85, 86] An isomerization could then lead to the formation of the anhydride.

The main issue that seems to impair the catalytic activity of GPx mimics is a scrambling reaction that occurs after the selenosulfide has formed: in fact, to restore the active catalytic species from the state of selenosulfide a second reaction with a thiol is needed. This process is a nucleophilic substitution with the entering thiol as the nucleophile and the sulfur of the selenosulfide in the catalyst as the substrate. If in the GPx enzyme the substitution is completely directed towards the sulfur of the selenosulfide, in the small mimics the attack at selenium has been seen to be preferred. [72, 87] In this way after reaction with a second molecule of thiol we are not left with the regenerated catalyst, but with the starting selenosulfide which is catalytically inactive. This results in relative activities far lower than that of GPx, *e.g.* ebselen has a relative activity of about 10⁻⁴, [88] that is one of the reasons preventing the actual use of these compounds as active antioxidant drugs. Moreover, the toxicity of organoselenium compounds and their metabolites, although lower than that of inorganic ones, [73, 77] is another concerning issue that limits the *in vivo* employment of these



(a)



(b)

Figure 1.4: Mechanism for the reduction of H_2O_2 by an organic thiol in the catalytic cycles of ebselen (a) and 2-(N,N-(dimethylamino)-methyl)benzenediselenide (b).

molecules and has yet to be completely clarified. The pro-oxidant and thiol depleting effect observed for ebselen in absence of peroxides [78] and the non-substrate specificity of many mimics [73] are two of the main drawbacks of which they suffer. However, the toxicology of GPx mimics has not been sufficiently explored yet and if isolated clinical trials reported no acute toxicity, [77] the long term effect of these molecules is still not clear.

A sibling to selenium, tellurium has no established biological role. Nevertheless, recent findings on semi-natural telluroproteins [88, 89] showed that a high peroxidase activity is displayed when Sec is substituted with tellurocysteine (Tec). Moreover, the possibility to use organotellurides in pharmacology has been the subject of research in the last 20 years [78, 90] but the lack of conclusive results on the toxicity of organotellurium compounds is still the limiting factor, as their pro-oxidant activity could potentially outweigh their beneficial antioxidant function. [74, 76]

This brief overview of the properties of organoselenides and tellurides outlines a somewhat disappointing, albeit promising, picture because, if in organic synthesis the use of these compounds (in particular of organoselenides) is common practice, their employment in pharmacology and medicine is very limited. Therefore, mechanistic studies to understand the traits that could enhance the reactivity of these compounds have to go hand in hand with investigation on the long and short-term toxicology of selenium and tellurium organocompounds to provide the knowledge required to rationally design efficient and biologically viable mimics that can display a high GPx-like activity and are biocompatible to be used as anti-oxidant drugs in the prevention of elevated levels of oxidative stress.

Chapter 2

Theory and methods

THIS Thesis is a computational work that involves different methods and calculation types. Computations that are based on quantum mechanics (QM) as well as simulations that are treated according to the laws of classical mechanics have been used. In this chapter a detailed, albeit concise presentation will illustrate the features of the different methodologies employed in this work.

2.1 Quantum mechanical calculations

SINCE the studies on the nature of matter at the end of the XIX century combined with those of Max Planck on the quantization of radiation, it was seen that the best possible description for systems that are of atomic size cannot be based on classical mechanics. Therefore, nowadays it seems natural to apply the laws of quantum mechanics to accurately describe the behavior of atomistic systems. But since the formulation of Schrödinger's equation, [91] it was immediately evident that the complexity of the problem was a very hard obstacle to overcome. The impossibility of having analytical solutions for even the simplest

systems limited the early application of this theory to describe structures of practical interest. The development of the first automated calculators (computers) from the early 50s certainly expanded the boundaries of the systems that was possible to investigate but, at the same time, the computational resources that were progressively becoming faster and more reliable with each passing day prompted for the development of more accurate models that were also inevitably more computationally expensive.

The main drawback of the methods based on Schrödinger's equation (also called wavefunction methods, post-Hartree–Fock methods or *ab initio* methods) is that they rely on the solution of a wavefunction that depends on the number of electrons (N) involved and has therefore $4N$ degrees of freedom (three spatial coordinates and a spin coordinate for each electron) which quickly increase as additional atoms are added to the system. Moreover, the computational cost of these models is nowhere near linear with respect to N but it has a stronger dependence that is usually of N^5 or above for most methods. This limit of the *ab initio* computations urged the development of an alternative theory which could lead to the same results but at a reduced cost. The answer that changed the perspective completely came in 1964 with the formulation of the Hohenberg–Kohn theorems and the introduction of the density functional theory (DFT). [92]

2.1.1 Density functional theory

Although the formal description of the theory was developed by Hohenberg and Kohn in the 60s, DFT has a much older origin. Its history begins in the 20s with the theories of Thomas and Fermi on the behavior of an electron gas moving in an external potential. [93–95] The idea behind the theory was to be able to determine the energy of the ground state of a collection of electrons affected by an external (nuclear) potential in terms of the electron density alone. This would

reduce the number of degrees of freedom of the resulting equation from $4N$ to only 4 (three spatial and one spin coordinates). This theory was further developed by Dirac. But it is with the aforementioned Hohenberg–Kohn theorems that it assumed a defined shape. These two theorems state that:

1. The ground state energy can be completely defined by the corresponding electron density.
2. If a well-behaved density trial function is selected, the energy functional yields an energy that is greater than or equal to that of the exact ground state electron density.

Mathematically, they can be expressed by two very simple relationships:

$$E = E[\rho(\mathbf{r}, s)] \quad (2.1)$$

$$E[\rho_{\text{trial}}] \geq E[\rho(\mathbf{r}, s)] \quad (2.2)$$

The simplicity of Equations (2.1) and (2.2) can be deceiving. In fact, the Hohenberg–Kohn theorems just *prove* the existence of a functional of the energy but they *do not provide* any clue as to how this functional is structured. Therefore, another step was required before DFT could be effectively put to work. A year after the Hohenberg–Kohn theorems, Kohn and Sham proposed a computational scheme based on DFT. [96] The main aspect of this scheme was that the electrons were treated as noninteracting particles and the inter-electronic and electron-nuclear interactions were recovered using an effective external potential (the Kohn–Sham potential) instead of the nuclear one :

$$\left[-\frac{1}{2}\nabla^2 + v_S(\mathbf{r}) \right] \phi_i = \epsilon_i \phi_i \quad (2.3)$$

The advantage of this formulation is that the exact density can be recovered from the sum over the densities of the first N orbitals, i.e. $\rho(\mathbf{r}) = \sum |\phi_i|^2$, and thus the exact energy can be calculated as $E[\rho(\mathbf{r})]$. Although the problem is now reduced to an effective one-electron formulation, the issue to address revolves now around $v_S(\mathbf{r})$. This quantity does not account only for the attractive nucleus-electron potential and the repulsive electron-electron potential, but it also contains all exchange and correlation contributions. Since there is no analytical formula in term of the density for this exchange correlation term, approximations have to be made to obtain a potential that fairly approximates the exact density. To this end, different approaches have been developed such as the local density approximation (LDA) or the generalized gradient approximation (GGA).

In addition to those schemes, other *hybrid* functionals have been developed, in which part of the exchange energy is calculated from the exact Hartree–Fock exchange.

2.1.2 Activation strain model and energy decomposition analysis

Many biological processes would not take place without the presence of enzymes. These molecules catalyze reactions that would otherwise be kinetically inert. In a catalytic cycle, a fundamental part is played by the different transition states (TSs) that are present in the mechanism, and more specifically by the barriers that must be overcome to go from one intermediate to another. Although an accurate kinetic analysis of enzymatic cycles is out of the scope of this Thesis, tuning the thermodynamics that characterize those reactions to obtain lower barriers and thus, possibly, a more favorable reaction, is instead one of the main objective of this work. To that end, an intimate understanding of the energetics of the system is required. The activation strain model (ASM) [97, 98] and the energy decomposition analysis (EDA) [99–101] are two tools that, in combination with

the Kohn–Sham molecular orbital model, allow a quantitative description of the different energy contributions that account for the different stabilization of the stationary points in a reaction. Not only does their simple formulation give the ability to describe a particular reaction, but it also permits to predict the necessary actions to improve it. This section gives a brief overview of these two schemes.

The activation strain model (ASM) of chemical reactivity [97, 98] is based on the idea that the fragment approach, generally used for stable molecules, can be extended to reactions, in fact to any point along a potential energy surface (ΔE). Thus it constitutes a model for understanding reactivity in terms of the original reactants involved in a chemical process. If a suitable reaction coordinate (ζ) in this potential energy surface is selected, the resulting energy profile $\Delta E(\zeta)$ can be decomposed into two contributions:

$$\Delta E(\zeta) = \Delta E_{\text{strain}}(\zeta) + \Delta E_{\text{int}}(\zeta) \quad (2.4)$$

The first term on the right hand side of Eq. (2.4) refers to the energetic contributions needed to perform the geometrical rearrangements during the reaction mechanism. It depends on the rigidity of the fragments involved, *e.g.* on the nature and strength of the bond that must break and on the flexibility of the angles that get deformed. Due to its definition ΔE_{strain} is usually positive (*i.e.* destabilizing) and it is the main cause for the presence of an energy barrier in activated processes.

The second term in Eq. (2.4) is usually stabilizing and refers to the mutual interaction that the reactants experience and depends upon their electronic structure and the orientation they have as they come together during the reaction. In general, the first part of a reaction is dominated by ΔE_{strain} up to a transition state after which ΔE_{int} takes over leading to the formation of the products (Figure

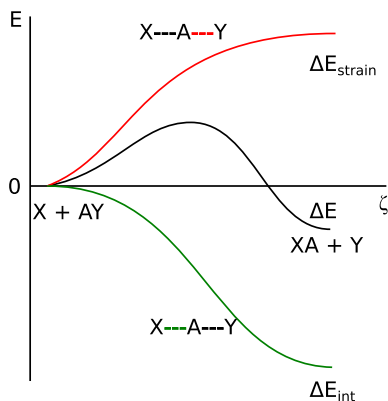


Figure 2.1: ASM model applied to an S_N2 model reaction

2.1). Since the considered transition state represents a first order saddle point on the whole potential energy surface (PES) of the structure, for which the first derivative with respect to a particular reaction coordinate is zero, if we take this derivative of Eq. (2.4) it can be easily seen that the transition state occurs when the interaction and strain energy have identical first derivative with opposite sign:

$$\frac{d\Delta E_{\text{strain}}}{d\zeta} = -\frac{d\Delta E_{\text{int}}}{d\zeta} \quad (2.5)$$

Therefore, the fine tuning of the strain and interaction through chemical modification of the reactants can directly affect the reaction barrier making the reaction less or more favorable.

The interpretation of strain energy is pretty simple and it is fairly easy to predict the effects that modification on the structure can have on the reactivity. Interaction, on the other hand, is a bit more complicated. It accounts in fact

for a variety of contributions that are most of the times intimately connected and dependent on each other. To extricate this apparent tangle, the energy decomposition analysis (EDA) scheme has been developed within the Kohn–Sham molecular orbital framework. It quantitatively divides the total interaction energy in three contributions:

$$\Delta E_{\text{int}} = \Delta V_{\text{elstat}} + \Delta E_{\text{Pauli}} + \Delta E_{\text{oi}} \quad (2.6)$$

In this way, the total interaction energy can be decomposed into terms that are directly related to a certain component: ΔV_{elstat} defines the coulombic attraction between unperturbed charge densities of the reacting fragments, ΔE_{Pauli} quantifies the repulsive term arising from the overlap of filled orbitals and finally ΔE_{oi} accounts for the stabilizing orbital interaction that take place between a filled and an empty orbital. If an empirical term to calculate the dispersion contributions is used in the calculation, it is added also in the EDA scheme to Eq. (2.6). As for the activation strain model, energy decomposition analysis can also be performed along a suitable reaction coordinate:

$$\Delta E_{\text{int}}(\zeta) = \Delta V_{\text{elstat}}(\zeta) + \Delta E_{\text{Pauli}}(\zeta) + \Delta E_{\text{oi}}(\zeta) \quad (2.7)$$

making it possible to study a reaction throughout its entire path.

2.1.3 Solvent effects

A somewhat mythical view of the chemist is that of a scientist that works in a lab trying to mix different *liquids* to obtain the desired product. The chemistry of liquids, or alternatively in the *liquid* phase has always represented the preferred

way to conduct reactions. Solutions have always offered a very advantageous way to efficiently mix reactants in a cheap and relatively safe and efficient way.

However, the first quantum theoretical models were based on reactions in the gas phase because treatment of a condensed phase was seen to be too complicated. Due to the increasing computational demands upon the addition of new molecules to the system, it was clear from the start that an explicit treatment of the solvent (*i.e.* each solvent molecule with its electrons) would be, besides difficult to theorize, extremely expensive in computational resources. Therefore, due to the impracticability of the explicit scheme, theoretical models were developed to account for the solvent effects without explicitly considering its atomistic structure. A solution to this problem was the introduction of the continuum solvation models: the solvent is treated like a bulk dielectric in which the solute is immersed. Since the electronic structure of the solvent is lost, the only effect that it has on the solute is of electrostatic nature and the main parameters that are relevant in these schemes are the dielectric constant of the solvent (ϵ) and the surface area of the cavity formed by the solute atoms.

The two solvation models used in this Thesis are the conductor-like screening model (COSMO) [102] and the SMD model [103]. They have been proved to have a comparable accuracy, [103, 104] and the choice of two different models was solely determined from the fact that they were the best methods available in the two QM softwares employed in the computations, namely ADF and Gaussian09. The general form of the standard-state free energy of solvation can be expressed as:

$$\Delta G_S^0 = \Delta G_{ENP} + G_{CDS} + \Delta G_{\text{conc}}^0 \quad (2.8)$$

in which ΔG_{ENP} is a term including the electronic, nuclear and polarization

components of the free energy change, ΔG_{CDS} accounts for the energy change caused by solvent cavitation, dispersion and changes in solvent structure and ΔG_{conc}^0 is the correction stemming from the different gas-phase and solvent standard states. COSMO and SMD use slightly different algorithms to calculate the bulk electrostatic contributions in presence of the solvent (ΔG_{ENP}), namely the eponymous scheme in the case of COSMO [102, 105] and the IEF-PCM formalism [106–108] in the case of SMD. However, these two schemes were seen to lead to very similar results and the main difference between these two methods resides in how they calculate the non electrostatic terms (ΔG_{CDS}). In COSMO, this contribution depends only on the surface area of the cavity and on a function the dielectric constant of the medium, whereas in SMD a more complex scheme is employed in which other macroscopic descriptors of the solvent, such as its refractive index or macroscopic surface tension, are included.

2.2 Classical calculations

THE distinguishing feature of the good scientist is the endless hunger for knowledge. As soon as a problem has been solved, a new one, oftentimes more complex, must be addressed. This loathing of idleness distinguished also computational chemists, who, since the dawn of this discipline, wanted to address more and more complicated problems with the new tools that the development of the computational hardware was offering. Therefore, the need of calculations of static properties of small structures grew rapidly into the necessity to compute the dynamical behavior of huge molecular complexes. Since these very large molecules were usually taken from biology (proteins, nuclear acid chains, cellular membranes. . .) and their study could lead to discoveries that could be beneficial in the medical and pharmaceutical sciences, a strong impulse was imparted to develop suitable models that could describe these systems. Quantum mechanical

models were immediately discarded because, as mentioned in Section 2.1, their computational cost rapidly increases with the number of electrons. Thus, the natural choice was to neglect the electronic structure of the molecule. However, without electrons there cannot be any quantum model. Classical mechanics was therefore the only option to describe a molecule that was eventually defined as a collection of (hard) spheres (the atoms) connected by an ensemble of springs (the chemical bonds).

2.2.1 Molecular dynamics simulations

The loss of the electronic structure in classical calculations makes it impossible to study reactive processes. The information gathered from a study of the dynamics of *e.g.*, a protein can however give useful information on bulk properties of the system: transport coefficients, time-dependent responses, rheological properties and spectra of various nature can be obtained also from a calculation relying on the laws of classical mechanics.

The theory supporting molecular dynamics (MD) simulations is rather simple and consists in the numerical solution of the classical equations of motion:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i \quad (2.9)$$

$$\mathbf{f}_i = -\frac{\partial}{\partial \mathbf{r}_i} V \quad (2.10)$$

in which V represents a potential energy from which the forces acting on each single atom are derived. The V function consists of many terms that can be broadly divided into two categories: bonding and non-bonding potentials.

Different implementations of this model led to slightly different definitions of the various terms of which the bonded and non-bonded interactions are them-

selves composed. This Thesis will focus on those employed in the Assisted Model Building with Energy Refinement (AMBER) molecular dynamics package. [109] The AMBER potential is expressed as:

$$\begin{aligned}
 V_{\text{AMBER}} = & \sum_{i < j}^{n'_{\text{atoms}}} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \sum_{i < j}^{n'_{\text{atoms}}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \\
 & + \sum_i^{n_{\text{bonds}}} b_i (r_i - r_{i,\text{eq}})^2 + \sum_i^{n_{\text{angles}}} a_i (\theta_i - \theta_{i,\text{eq}})^2 \\
 & + \sum_i^{n_{\text{dihedral}}} \sum_i^{n_{i,\text{max}}} \frac{V_{i,n}}{2} [1 + \cos(n\phi_i - \gamma_{i,n})]
 \end{aligned} \tag{2.11}$$

The first line of Eq. (2.11) contains the sum of two non-bonded terms¹, namely the Coulomb pairwise electrostatic potential and the Lennard-Jones potential. The second and third lines show the bonded interactions of the AMBER potential. They are three terms corresponding to the harmonic potentials defined for bonds and angles and to the truncated Fourier series defined for dihedral angles.

The parameters contained in Eq. (2.11) are usually derived from QM mechanical calculations and as a whole they define a *force field* which is essentially a collection of parameters that completely define a molecule. Different types of force fields were developed over the years, many of them designed specifically to accurately describe a particular class of molecules. For example, the force field used in all the simulations presented in this Thesis is ff14SB which was specifically

¹The summation for this two terms considers n' atoms due to its definition that considers only atoms on different molecules or that are separated by at least 3 bonds.

designed to run molecular dynamics of proteins. [110]

2.3 Notation convention and software

THROUGHOUT all the Thesis a coherent notation will be used for the definition of the different levels of theory employed in the QM calculations²: *modifier*³ - *functional* / *basis set* (e.g. COSMO-BLYP/TZ2P). If a different level of theory is used for geometry optimization and final energy calculation, the two will be separated by a double slash (/), with the method of the single point preceding the one used in the geometry optimization. For example if a molecule is optimized in gas phase at the BLYP/TZ2P level of theory but its final energy is calculated in water using the COSMO model with the BLYP functional using a QZ4P basis set, the whole level of theory will be denoted as: COSMO-BLYP/QZ4P//BLYP/TZ2P.

Three software packages have been employed for most of the calculations reported in this work. For the DFT part, the main programs used are the Amsterdam density functional (ADF) package [111–113] and Gaussian09. [114] All calculations using a basis set of Slater type functions (e.g. TZ2P) were carried out with ADF. On the other hand, calculations employing Pople basis sets (e.g. 6-311G(d,p)) or correlation-consistent basis sets (e.g. cc-pVDZ) were carried out with Gaussian09.

For the classical molecular dynamics simulations the Amber 2016 [115] software package comprising Amber16 and AmberTools16 programs was employed. Usage of any other software in specific cases will be detailed in the text.

²The level of theory of a QM calculation is the combination of functional and basis set employed.

³The modifier can be the inclusion of relativistic effects or of a solvation model

Chapter 3

The chalcogen- π interaction

Adapted from

Bortoli, M.; Ahmad, S. M.; Hamlin, T. A.; Bickelhaupt, F. M.; Orian, L.

Nature and strength of chalcogen- π interactions

Physical Chemistry Chemical Physics, **2018**, DOI: 10.1039/C8CP05922E

3.1 Introduction

"WITH corageous simplification, one might assert hat the chemistry of the last century was largely the chemistry of covalent bonding, whereas that of the present century is more likely to be the chemistry of non-covalent binding." [116] This quote by Prof. Schneider denotes how the study of non-covalent interactions has become the center of interest of modern chemistry. Hydrogen bond is without any doubt the most famous and extensively studied, [117] but many others belong to this group, *e.g.* dipole-dipole, ion-dipole and

π - π interactions. An analogous mechanism to that of H-bond, in which an electropositive atom (bond donor) interacts with a Lewis base (bond acceptor) to form a weak interaction, has been found also for other classes of atoms such as halogens, [118–120] chalcogens [121–123] and pnictogens. [124–126] Among the siblings to hydrogen bond, chalcogen bond has lately received much attention due to its promising features that makes it optimal to be used in many applications such as catalysis, ion transport and material and drug design. [127–129] Moreover, a very recent survey of the PDB database showed that many protein-ligand complexes in which the ligand contains a chalcogen atom adopt structures that allow the formation of chalcogen bonds. [130] In this Chapter, the discussion will be focused on a selection of chalcogen bond acceptors, *i.e.* small organic molecules having a double or triple bond. Binding energies that range between -3.3 and -6.6 kcal mol $^{-1}$ were calculated for similar complexes and the main energy contribution was attributed to a charge transfer from a π orbital of the bond acceptor, which is localized around the C–C multiple bond, to a σ^* of the bond donor localized around the chalcogen. In addition to the bond acceptor, the substituents directly attached to the chalcogen atoms were seen to affect the strength of the chalcogen bond. [131] Therefore, a systematic study on the extent and magnitude of this effect on the bond strength can provide useful insight into its nature and give precious information on the possibility to fine tune these interactions.

Density functional theory (DFT) calculations in combination with quantitative activation strain analysis (ASA) and energy decomposition analysis (EDA) were performed to analyze the formation of a bond between a chalcogen atom and chalcogen- π bond in a set of complexes of general formula $\mathbf{X}_2\mathbf{D}\cdots\mathbf{A}$ in which \mathbf{D} is the chalcogen bond donor (O, S, Se or Te), \mathbf{X} is a halogen and \mathbf{A} is a small organic molecule containing an unsaturated bond (2-butyne, ethylene and acetylene were selected for this study) that gives rise to an electron-rich π system. All the possible

combinations between **D**, **A** and **X** were systematically explored to assess the particular effect each element has on the stabilization obtained upon formation of the chalcogen bond.

3.2 Methods

THE geometries of the chalcogenides, unsaturated hydrocarbons, and the complexes were optimized imposing C_S symmetry. The functional BLYP [132–135] in combination with TZ2P basis was employed for all the elements. The TZ2P basis set is a large uncontracted set of Slater-type orbitals (STOs) of triple- ζ quality and has been augmented with two sets of polarization functions on each atom that is, 2p and 3d on H, 3d and 4f on C, S, F and Cl, 4d and 4f on Se and Br, and 5d and 4f on Te and I. The frozen-core approximation was adopted for the core electrons: up to 1s for C and F, up to 2p for S and Cl, up to 3p for Se and Br, and up to 4p for Te and I. An auxiliary set of s, p, d, f, and g STOs was used to fit the molecular density and to represent the Coulomb and exchange potentials accurately in each SCF cycle. Dispersion corrections were included employing the D3 scheme with the Becke-Johnson damping (D3(BJ)) developed by Grimme *et al.* [136] Scalar relativistic effects were accounted for through the zeroth-order regular approximation (ZORA). [137] This level of theory is referred to as ZORA-BLYP-D3(BJ)/TZ2P. Frequency calculations were employed to confirm the nature of the stationary points.

The activation strain model [97] was employed to quantitatively decompose the contributions to the chalcogen- π bonding energy in the complexes under investigation. The relative energy of a molecular complex can be divided, through activation strain analysis (ASA), as the sum of strain contribution (ΔE_{strain}) and an interaction contribution (ΔE_{int}). ΔE_{strain} is the energy required for the geometrical deformation of the reacting species when they are brought from

infinity to the geometry they acquire after complex formation and ΔE_{int} is the actual interaction energy between the fragments. This can be further divided, through energy decomposition analysis (EDA), [99] into the electrostatic interaction (ΔV_{elstat}), Pauli repulsion (ΔE_{Pauli}), and orbital interactions (ΔE_{oi}) contributions. Since in the case of BLYP-D3(BJ) functional, an empirical correction to account for dispersion interaction is added, the term ΔE_{disp} is also present. For a more detailed treatment of ASA and EDA see Chapter 2, Section 2.1.2. To perform the ASA and the EDA, single point energy calculations were run on the previously optimized geometries with a quadruple- ζ quality basis set. This level of theory is denoted as ZORA-BLYP-D3(BJ)/QZ4P//ZORA-BLYP-D3(BJ)/TZ2P.

The electron density distribution was analyzed using the Voronoi deformation density (VDD) method for computing fragment charges. [138, 139] The Voronoi deformation density method was chosen because it is basis set independent, unlike Mulliken charges, which are heavily dependent on basis choice. [139] The VDD method calculates the amount of electronic density that flows to or from a certain atom due to the bond formation, by spatial integration of the deformation density over the atomic Voronoi cell.

For selected cases, we have computed highly correlated *ab initio* reference data through CCSD(T) single-point calculations performed on ZORA-BLYP-D3(BJ)/TZ2P optimized geometries. In view of its accuracy and efficiency, [140] we chose the domain localized pair of natural orbitals coupled cluster with singles and double excitations treated explicitly and triple excitations treated perturbatively (DLPNO-CCSD(T)) as implemented in the Orca 4.0.0 software package. [141–143] Relativistic effects were accounted for through the second order scalar Douglas-Kroll-Hess method [144] and a quintuple- ζ basis sets, designed for relativistic calculations, was employed. For the heaviest elements (Te and I) this basis set was not available in Orca. Therefore, for these two elements, the aug-cc-pVQZ-DK basis set was used instead of the quintuple- ζ basis. All the benchmark data are

reported in Appendix A. All energies were calculated with and without the counterpoise correction to account for the basis set superposition error (BSSE). [145] Comparison between DFT and DLPNO-CCSD(T)/aug-cc-pV5Z level of theory for a selection of complexes resulted in a smaller average mean absolute error for the DFT uncorrected energies (see Appendix A for details), likely because of a fortuitous cancellation of the error in the uncorrected values. [146] Therefore, unless specified, uncorrected energies are used throughout the Chapter.

3.3 Results and discussion

A total of 48 complexes result from all the possible permutations of **D**, **A**, and **X**. They were all optimized imposing C_s symmetry and the relative stabilization with respect to the free reactants was systematically investigated, employing quantitative energy decomposition schemes (ASA and EDA). Throughout the rest of this Section results on the structure and energetics of these complexes will be presented. To organize the data in a clear manner, attention will be centered sequentially on the effect the variation of the chalcogen, substrate and halogen has on the total stability. Complexes will be grouped in *series* of compounds. A *series* is a set of compounds through which only **D**, **A** or **X** is varied.

3.3.1 Geometrical parameters

The general formula of the chosen model compounds is $X_2D\cdots A$ in which **D** is a chalcogen bond donor (O, S, Se or Te), **A** is the chalcogen bond acceptor and **X** is a halogen atom. Three small unsaturated organic molecules were selected as bond acceptor: 2-butyne (**2but**), acetylene (**ac**) and ethylene (**et**). The optimized geometries of all the complexes show a high degree of similarity (an example is shown in Figure 3.1). Sulfur, selenium and tellurium compounds all show a

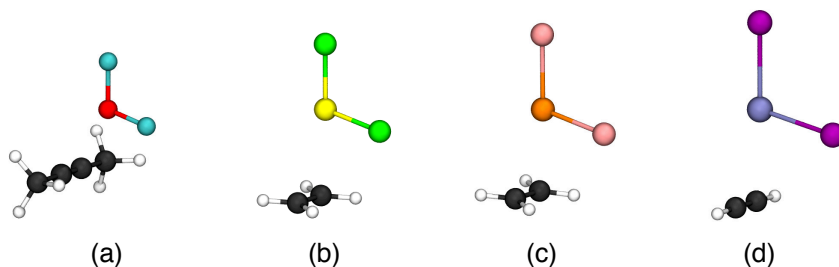


Figure 3.1: Examples of the optimized structures of four complexes: $\text{F}_2\text{O}\cdots 2\text{but}$ (a), $\text{Cl}_2\text{S}\cdots \text{et}$ (b), $\text{Br}_2\text{Se}\cdots \text{et}$ (c) and $\text{I}_2\text{Te}\cdots \text{ac}$ (d).

chalcogen- π bond that is collinear to one of the two chalcogen-halogen bonds. Moreover, the two chalcogen halogen bonds are not found to be of equal length in the complex but the one collinear to the chalcogen- π bond is calculated to be moderately longer. On the other hand, oxygen complexes have a slightly different geometry as they adopt a Y shaped structure with chalcogen-halogen bonds of equal length.

The shortest donor-acceptor distance (d , measured as the distance between the chalcogen atom and a carbon atom involved in the unsaturated bond) is found in $\text{F}_2\text{O}\cdots 2\text{but}$ (2.59 Å) and the longest one in $\text{I}_2\text{Te}\cdots \text{ac}$ (3.21 Å) with a span of 0.62 Å.

For all the chalcogenides, Voronoi deformation densities were calculated on the chalcogen atom (Table 3.1). Due to their definition, VDD charges do not indicate an absolute value of charge on a particular atom but a relative depletion or increase of charge during the formation of a molecule. This value can be very useful to compare structures with the same chalcogen to see which one undergoes the greater charge depletion or increase upon formation of the various

chalcogenides that can then react with the substrates to form the π complexes. To keep the effect of the different substrates out of the picture, VDD charges were calculated on the isolated chalcogenides at the geometry they adopt in the final complexes. A common trend found for VDD charges in all the series shows that a higher charge depletion takes place as the difference in electronegativity between the chalcogen and the halogen increases. For example deformation densities go from 0.264 to 0.043 a.u. in the $\mathbf{X}_2\mathbf{Se}\cdots\mathbf{ac}$ series as the halogen changes from F to I and from 0.287 to 0.108 a.u. in the $\mathbf{F}_2\mathbf{D}\cdots\mathbf{ac}$ series, ascending the group from Te to O.

3.3.2 Bonding analysis

Activation strain analyses (Table 3.2) and energy decomposition analyses (Table 3.3) were employed to quantify the different contributions that determine the stabilization derived from the formation of the chalcogen bond. The interacting complexes were divided into two fragments, one consisting of the chalcogenide and the other of the unsaturated substrate. The effects on the bonding stabilization obtained with the modification of the chalcogen atom, of the accepting substrate and of the halogen atoms were investigated, and data will be presented focusing on one component at a time.

Effect of the chalcogen. The electropositivity of the chalcogen atom increases down the group, from O to Te. The same trend is found in computed total stabilization energies for the π bonded complexes. For example ΔE goes from $-1.3 \text{ kcal mol}^{-1}$ in the case of $\mathbf{F}_2\mathbf{O}\cdots\mathbf{ac}$ to -4.3 , -7.5 and $-9.4 \text{ kcal mol}^{-1}$ for the analogous compounds of S, Se and Te, respectively, in the same series. This trend stems from the increasing interaction energy computed when going from oxygen to tellurium, which, when decomposed through EDA, shows how all stabilizing contributions, *i.e.* ΔV_{elstat} , ΔE_{oi} and ΔE_{disp} , become stronger with the size of

Table 3.1: Chalcogen-carbon distance (in Å) and Voronoi deformation densities (VDD, in a.u.) on chalcogen atom (D).^a

| | <i>d</i> | VDD | | <i>d</i> | VDD |
|-------------------------------|----------|--------|--------------------------------|----------|-------|
| F₂O...2but | 2.59 | 0.103 | F₂Se...2but | 2.69 | 0.254 |
| Cl₂O...2but | 2.78 | -0.039 | Cl₂Se...2but | 2.87 | 0.176 |
| Br₂O...2but | 2.72 | -0.089 | Br₂Se...2but | 2.89 | 0.125 |
| I₂O...2but | 2.83 | -0.129 | I₂Se...2but | 2.96 | 0.046 |
| F₂O...et | 2.90 | 0.105 | F₂Se...et | 2.68 | 0.272 |
| Cl₂O...et | 2.98 | -0.046 | Cl₂Se...et | 2.94 | 0.181 |
| Br₂O...et | 2.88 | -0.098 | Br₂Se...et | 2.97 | 0.128 |
| I₂O...et | 2.93 | -0.136 | I₂Se...et | 3.06 | 0.044 |
| F₂O...ac | 2.97 | 0.108 | F₂Se...ac | 2.76 | 0.264 |
| Cl₂O...ac | 3.10 | -0.044 | Cl₂Se...ac | 3.01 | 0.178 |
| Br₂O...ac | 3.05 | -0.097 | Br₂Se...ac | 3.05 | 0.126 |
| I₂O...ac | 3.08 | -0.135 | I₂Se...ac | 3.13 | 0.043 |
| F₂S...2but | 2.77 | 0.189 | F₂Te...2but | 2.81 | 0.264 |
| Cl₂S...2but | 2.92 | 0.122 | Cl₂Te...2but | 2.96 | 0.232 |
| Br₂S...2but | 2.93 | 0.073 | Br₂Te...2but | 3.00 | 0.197 |
| I₂S...2but | 2.95 | 0.002 | I₂Te...2but | 3.07 | 0.124 |
| F₂S...et | 2.86 | 0.195 | F₂Te...et | 2.82 | 0.297 |
| Cl₂S...et | 3.10 | 0.120 | Cl₂Te...et | 3.00 | 0.244 |
| Br₂S...et | 3.09 | 0.070 | Br₂Te...et | 3.05 | 0.206 |
| I₂S...et | 3.16 | -0.004 | I₂Te...et | 3.14 | 0.127 |
| F₂S...ac | 2.93 | 0.193 | F₂Te...ac | 2.83 | 0.287 |
| Cl₂S...ac | 3.15 | 0.120 | Cl₂Te...ac | 3.07 | 0.239 |
| Br₂S...ac | 3.17 | 0.069 | Br₂Te...ac | 3.12 | 0.201 |
| I₂S...ac | 3.21 | -0.003 | I₂Te...ac | 3.21 | 0.123 |

a) Computed at ZORA-BLYP-D3(BJ)/QZ4P//ZORA-BLYP-D3(BJ)/TZ2P.

Table 3.2: Activation strain analysis (electronic energies in gas phase, in kcal mol⁻¹) of the model systems. ^a

| | ΔE | ΔE_{strain} | ΔE_{int} | | ΔE | ΔE_{strain} | ΔE_{int} |
|-------------------------------|------------|----------------------------|-------------------------|--------------------------------|------------|----------------------------|-------------------------|
| F₂O...ac | -1.3 | 0.2 | -1.5 | F₂Se...ac | -7.5 | 0.8 | -8.3 |
| Cl₂O...ac | -1.0 | 0.1 | -1.1 | Cl₂Se...ac | -5.4 | 0.2 | -5.6 |
| Br₂O...ac | -1.3 | 0.1 | -1.4 | Br₂Se...ac | -5.0 | 0.2 | -5.2 |
| I₂O...ac | -1.5 | 0.0 | -1.5 | I₂Se...ac | -4.4 | 0.1 | -4.5 |
| F₂O...et | -1.7 | 0.5 | -2.2 | F₂Se...et | -9.1 | 1.6 | -10.7 |
| Cl₂O...et | -1.6 | 0.2 | -1.8 | Cl₂Se...et | -6.7 | 0.5 | -7.2 |
| Br₂O...et | -2.2 | 0.2 | -2.4 | Br₂Se...et | -6.4 | 0.4 | -6.8 |
| I₂O...et | -2.4 | 0.1 | -2.5 | I₂Se...et | -5.7 | 0.2 | -5.9 |
| F₂O...2but | -3.9 | 2.6 | -6.5 | F₂Se...2but | -11.4 | 2.0 | -13.4 |
| Cl₂O...2but | -3.3 | 0.6 | -3.9 | Cl₂Se...2but | -9.5 | 1.0 | -10.5 |
| Br₂O...2but | -4.2 | 0.5 | -4.7 | Br₂Se...2but | -9.3 | 0.7 | -10.0 |
| I₂O...2but | -4.3 | 0.2 | -4.5 | I₂Se...2but | -8.6 | 0.5 | -9.1 |
| F₂S...ac | -4.3 | 0.3 | -4.6 | F₂Te...ac | -9.4 | 1.2 | -10.6 |
| Cl₂S...ac | -3.2 | 0.2 | -3.4 | Cl₂Te...ac | -6.8 | 0.4 | -7.2 |
| Br₂S...ac | -3.3 | 0.1 | -3.4 | Br₂Te...ac | -6.2 | 0.3 | -6.5 |
| I₂S...ac | -2.9 | 0.1 | -3.0 | I₂Te...ac | -5.4 | 0.2 | -5.6 |
| F₂S...et | -5.1 | 0.6 | -5.7 | F₂Te...et | -11.7 | 2.0 | -13.7 |
| Cl₂S...et | -4.0 | 0.3 | -4.3 | Cl₂Te...et | -8.5 | 0.8 | -9.3 |
| Br₂S...et | -4.1 | 0.2 | -4.3 | Br₂Te...et | -7.9 | 0.6 | -8.5 |
| I₂S...et | -3.9 | 0.1 | -4.0 | I₂Te...et | -7.0 | 0.3 | -7.3 |
| F₂S...2but | -7.3 | 1.3 | -8.6 | F₂Te...2but | -13.2 | 2.4 | -15.6 |
| Cl₂S...2but | -6.5 | 0.7 | -7.2 | Cl₂Te...2but | -11.2 | 1.2 | -12.4 |
| Br₂S...2but | -6.8 | 0.6 | -7.4 | Br₂Te...2but | -10.7 | 0.9 | -11.6 |
| I₂S...2but | -6.6 | 0.4 | -7.0 | I₂Te...2but | -9.8 | 0.6 | -10.4 |

a) Computed at ZORA-BLYP-D3(BJ)/QZ4P//ZORA-BLYP-D3(BJ)/TZ2P.

Table 3.3: Energy decomposition analysis (electronic energies in gas phase, in kcal mol⁻¹) of the model systems.^a

| | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} | ΔE_{disp} | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} | ΔE_{disp} |
|-------------------------------|---------------------------|----------------------------|------------------------|--------------------------|--------------------------------|----------------------------|------------------------|--------------------------|
| F₂O...2but | 16.3 | -8.4 | -12.0 | -2.4 | F₂Se...2but | 42.6 | -26.5 | -23.8 |
| Cl₂O...2but | 10.5 | -4.7 | -5.8 | -4.0 | Cl₂Se...2but | 28.1 | -17.2 | -14.6 |
| Br₂O...2but | 12.9 | -5.5 | -7.2 | -4.9 | Br₂Se...2but | 27.3 | -16.4 | -13.8 |
| I₂O...2but | 11.8 | -4.9 | -5.4 | -6.0 | I₂Se...2but | 23.7 | -13.8 | -11.1 |
| F₂O...et | 6.6 | -3.4 | -4.0 | -1.4 | F₂Se...et | 45.9 | -25.7 | -26.4 |
| Cl₂O...et | 6.4 | -2.7 | -3.0 | -2.6 | Cl₂Se...et | 23.7 | -13.4 | -12.4 |
| Br₂O...et | 8.6 | -3.5 | -4.2 | -3.2 | Br₂Se...et | 22.6 | -12.6 | -11.5 |
| I₂O...et | 8.8 | -3.5 | -3.8 | -3.9 | I₂Se...et | 18.9 | -10.3 | -9.0 |
| F₂O...ac | 4.5 | -2.6 | -2.4 | -1.1 | F₂Se...ac | 32.0 | -18.9 | -17.8 |
| Cl₂O...ac | 3.8 | -1.6 | -1.5 | -1.9 | Cl₂Se...ac | 16.5 | -10.1 | -8.2 |
| Br₂O...ac | 4.7 | -1.9 | -1.9 | -2.3 | Br₂Se...ac | 15.6 | -9.3 | -7.5 |
| I₂O...ac | 5.1 | -1.9 | -1.8 | -2.9 | I₂Se...ac | 13.0 | -7.6 | -5.7 |
| F₂S...2but | 27.3 | -16.3 | -14.8 | -4.8 | F₂Te...2but | 46.3 | -30.2 | -25.1 |
| Cl₂S...2but | 19.5 | -11.1 | -9.6 | -6.0 | Cl₂Te...2but | 32.4 | -20.8 | -16.3 |
| Br₂S...2but | 19.8 | -11.0 | -9.7 | -6.5 | Br₂Te...2but | 30.5 | -19.3 | -14.8 |
| I₂S...2but | 20.3 | -10.8 | -9.0 | -7.4 | I₂Te...2but | 27.0 | -16.6 | -12.3 |
| F₂S...et | 22.3 | -12.3 | -12.1 | -3.7 | F₂Te...et | 52.7 | -30.7 | -30.4 |
| Cl₂S...et | 12.6 | -6.8 | -5.9 | -4.1 | Cl₂Te...et | 30.9 | -18.0 | -16.3 |
| Br₂S...et | 13.5 | -7.0 | -6.3 | -4.5 | Br₂Te...et | 28.2 | -16.2 | -14.3 |
| I₂S...et | 12.2 | -6.1 | -5.3 | -4.9 | I₂Te...et | 23.4 | -13.2 | -11.2 |
| F₂S...ac | 15.4 | -9.2 | -8.0 | -2.9 | F₂Te...ac | 39.1 | -23.7 | -21.8 |
| Cl₂S...ac | 9.1 | -5.3 | -4.0 | -3.2 | Cl₂Te...ac | 21.4 | -13.3 | -10.8 |
| Br₂S...ac | 8.8 | -5.0 | -3.8 | -3.3 | Br₂Te...ac | 19.5 | -12.0 | -9.4 |
| I₂S...ac | 8.8 | -4.7 | -3.4 | -3.7 | I₂Te...ac | 16.0 | -9.7 | -7.2 |

a) Computed at ZORA-BLYP-D3(BJ)/QZ4P//ZORA-BLYP-D3(BJ)/TZ2P.

the chalcogen center. The better electrostatic interaction of structures with the heavier chalcogens can be explained on the basis of the VDDs (Table 3.1): for $\mathbf{F}_2\mathbf{O}\cdots\mathbf{ac}$ a deformation density of 0.108 a.u. is computed on the chalcogen atom which increases to 0.193, 0.264 and 0.287 a.u. for sulfur, selenium and tellurium compounds in this series. A larger overlap between the frontier molecular orbitals of the interacting fragments, *i.e.* the HOMO of \mathbf{A} and the LUMO of \mathbf{D} , is the cause of a higher orbital interaction as the chalcogen becomes more electropositive (Table 3.4). The computed values of 0.06, 0.13, 0.17 and 0.20 for the $\mathbf{F}_2\mathbf{D}\cdots\mathbf{ac}$ series, oxygen through tellurium, well reflect ΔE_{oi} trend: $\mathbf{F}_2\mathbf{O}\cdots\mathbf{ac}$ is the complex with the weakest orbital interaction at $-2.4 \text{ kcal mol}^{-1}$ followed by $\mathbf{F}_2\mathbf{S}\cdots\mathbf{ac}$, $\mathbf{F}_2\mathbf{Se}\cdots\mathbf{ac}$ and $\mathbf{F}_2\mathbf{Te}\cdots\mathbf{ac}$ whose orbital interaction energy is calculated to be -8.0 , -17.8 and $-21.8 \text{ kcal mol}^{-1}$, respectively. Finally, dispersion contributions are computed to be higher for the heaviest chalcogens due to the harder nature of oxygen and sulfur, whose ΔE_{disp} are of -1.0 and $-2.8 \text{ kcal mol}^{-1}$ in the $\mathbf{F}_2\mathbf{D}\cdots\mathbf{ac}$ series, and the increased polarizability of selenium and tellurium, whose dispersion contributions amount to -3.6 and $-4.2 \text{ kcal mol}^{-1}$ for $\mathbf{F}_2\mathbf{Se}\cdots\mathbf{ac}$ and $\mathbf{F}_2\mathbf{Te}\cdots\mathbf{ac}$, respectively.

Effect of the bond acceptor. The second major player in the formation of the chalcogen- π interaction is, without doubt, the bond acceptor. In this case, 2-butyne is found to be the one which forms the complexes with the highest stabilization, followed by ethylene and acetylene, thanks to a gradual increase in electrostatic, orbital and dispersion interactions which contribute in determining the trend. The less favorable ΔV_{elstat} is due to a systematic increase in the donor-acceptor distance passing from **2but** to **et** and **ac**. In the case of the $\mathbf{Cl}_2\mathbf{Se}\cdots\mathbf{A}$ series, it measures 2.87 Å with **2but**, 2.94 Å with **et** and 3.01 Å with **ac** (Table 3.1) and this directly affects the electrostatic contribution which diminishes from -17.2 to -13.4 and $10.1 \text{ kcal mol}^{-1}$ (Table 3.3). Complexes with 2-butyne also have the strongest orbital interaction. This is mostly due to the different energy of the HOMOs of the three molecules (Figure 3.2): **2but** has the highest lying HOMO

Table 3.4: HOMO-LUMO energy gap ($\Delta\epsilon$, in eV) and overlap (S) in selected model systems.^a

| | $\Delta\epsilon$ | S | | $\Delta\epsilon$ | S |
|------------------------------|------------------|------|------------------------------|------------------|------|
| F₂O...ac | 1.84 | 0.06 | Br₂O...ac | 1.97 | 0.07 |
| F₂S...ac | 3.58 | 0.13 | Br₂S...ac | 2.84 | 0.10 |
| F₂Se...ac | 3.03 | 0.17 | Br₂Se...ac | 2.75 | 0.13 |
| F₂Te...ac | 2.89 | 0.20 | Br₂Te...ac | 2.99 | 0.14 |
| Cl₂O...ac | 2.01 | 0.06 | I₂O...ac | 2.14 | 0.07 |
| Cl₂S...ac | 3.14 | 0.10 | I₂S...ac | 2.74 | 0.08 |
| Cl₂Se...ac | 2.91 | 0.14 | I₂Se...ac | 2.71 | 0.09 |
| Cl₂Te...ac | 3.11 | 0.18 | I₂Te...ac | 2.92 | 0.12 |

^a Computed at ZORA-BLYP-D3(BJ)/QZ4P//ZORA-BLYP-D3(BJ)/TZ2P. For data of all model systems, see Appendix B.

among the three substrates and thus can have a more favorable energy match, *i.e.* a smaller HOMO-LUMO gap, with the chalcogenide. Calculations show that the HOMO-LUMO gap increases from 1.63 to 2.39 and 3.03 eV (Table 3.5) and this translates into an orbital contribution of -14.6 , -12.4 and -8.2 kcal mol⁻¹ (Table 3.3). Finally, due to the size of the molecule which makes it more sensitive to van der Waals forces, 2-butyne gains the strongest stabilization from dispersion. Although this variation is more modest than that of ΔV_{elstat} or ΔE_{oi} , it is still appreciable with calculated ΔE_{disp} of, for example, -6.8 , -5.1 and -3.8 kcal mol⁻¹ in the case of **Cl₂Se...2but**, **Cl₂Se...et** and **Cl₂Se...ac** respectively (Table 3.3).

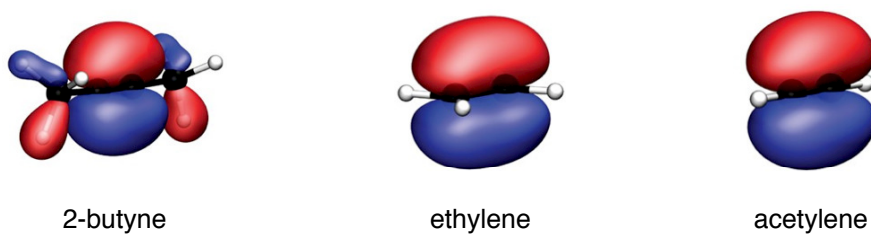


Figure 3.2: HOMOs of the unsaturated substrates. Isodensity: 0.03 a.u.

Table 3.5: HOMO-LUMO energy gap ($\Delta\epsilon$, in eV) and overlap (S) in selected model systems.^a

| | $\Delta\epsilon$ | S | | $\Delta\epsilon$ | S |
|-------------------------------|------------------|------|--------------------------------|------------------|------|
| Cl₂O...ac | 1.84 | 0.06 | Cl₂Se...ac | 3.03 | 0.17 |
| Cl₂O...et | 1.24 | 0.07 | Cl₂Se...et | 2.39 | 0.19 |
| Cl₂O...2but | 0.19 | 0.07 | Cl₂Se...2but | 1.63 | 0.17 |
| Cl₂S...ac | 3.58 | 0.13 | Cl₂Te...ac | 2.89 | 0.20 |
| Cl₂S...et | 3.00 | 0.16 | Cl₂Te...et | 2.27 | 0.22 |
| Cl₂S...2but | 2.09 | 0.14 | Cl₂Te...2but | 1.55 | 0.18 |

^a Computed at ZORA-BLYP-D3(BJ)/QZ4P//ZORA-BLYP-D3(BJ)/TZ2P. For data of all model systems, see Appendix B.

Effect of the halogen. Lastly, the variation of the halogen atom in the complexes has the lowest impact on the strength of the chalcogen- π bond. Nevertheless, a decrease in bond strength is computed going from the most electronegative (F) to the least electronegative (I) halogen. The reason behind this behavior can be found in the diminution of electrostatic and orbital interactions when going from fluorine to iodine. Taking as an example the $\mathbf{X}_2\mathbf{Se}\cdots\mathbf{et}$ series, ΔV_{elstat} is affected by the VDD on the chalcogen and the donor-acceptor distance in the complex: the former is reduced from 0.254 to 0.176, 0.125 and 0.046 a.u. and the latter lengthens from 2.69 to 2.87, 2.89 and 2.96 Å. Orbital analysis reveals that a decrease in the chalcogenide's LUMO amplitude on the chalcogen atom and thus a decrease of the overlap with the HOMO of the substrate is the main cause for diminished orbital interactions. Example values (reported for the $\mathbf{X}_2\mathbf{Se}\cdots\mathbf{2but}$) go from 0.17 ($\mathbf{F}_2\mathbf{Se}\cdots\mathbf{2but}$) to 0.12, 0.11 and 0.08 for the analogous complexes with chlorine, bromine and iodine, respectively. Consequently, a decrease in the total ΔE_{oi} is found for all the series which together with the weakening of the electrostatic interaction contributes to an overall destabilization of the complexes as the halogen bonded to the chalcogen center becomes less electronegative. For example, in the $\mathbf{X}_2\mathbf{Se}\cdots\mathbf{2but}$ series, it is found to be $-11.4 \text{ kcal mol}^{-1}$ in $\mathbf{F}_2\mathbf{Se}\cdots\mathbf{2but}$ but it amounts to $-8.6 \text{ kcal mol}^{-1}$ in $\mathbf{I}_2\mathbf{Se}\cdots\mathbf{2but}$.

Thus, the $\mathbf{X}_2\mathbf{D}\cdots\mathbf{A}$ bond strength is governed by the difference in electronegativity between chalcogen (D) and halogen (X) as this determines the atomic charge and the amplitude of the $\mathbf{X}_2\mathbf{D}$ LUMO on the chalcogen atom. This fits nicely with all main trends we compute: (i) the *decreasing* chalcogen bond strength as X becomes less electronegative along F, Cl, Br and I; and (ii) the *increasing* chalcogen bond strength as the chalcogen becomes more electropositive along O, S, Se and Te. Note that the oxygen complexes $\mathbf{X}_2\mathbf{O}\cdots\mathbf{A}$ are all relatively weakly bound with only minor and therefore less systematic variations in bond strength, *i.e.*, within a range of about 1 kcal mol^{-1} , along the various halogen substituents X.

Table 3.6: HOMO-LUMO energy gap $\Delta\epsilon$ (in eV) and overlap (S) in selected model systems.^a

| Complex | $\Delta\epsilon$ | S |
|--------------------------------|------------------|------|
| F₂S...2but | 2.09 | 0.14 |
| Cl₂S...2but | 1.77 | 0.09 |
| Br₂S...2but | 1.49 | 0.08 |
| I₂S...2but | 1.42 | 0.08 |
| F₂Se...2but | 1.63 | 0.17 |
| Cl₂Se...2but | 1.55 | 0.12 |
| Br₂Se...2but | 1.41 | 0.11 |
| I₂Se...2but | 1.40 | 0.08 |
| F₂Te...2but | 1.55 | 0.18 |
| Cl₂Te...2but | 1.75 | 0.16 |
| Br₂Te...2but | 1.66 | 0.12 |
| I₂Te...2but | 1.62 | 0.09 |

a) Computed at ZORA-BLYP-D3(BJ)/QZ4P//ZORA-BLYP-D3(BJ)/TZ2P. For data on all model systems, see Appendix B.

3.4 Conclusions

In this work, a series of chalcogen- π bonded complexes was analyzed *in silico* to obtain a clear picture of the contributions that determine the different interaction strength. Moreover, since through quantitative energy decomposition analysis these contributions can directly be related to the intrinsic properties of the fragments taking part in the bond, results from this study can be employed not only to describe the selected systems, but also to make predictions on the general behavior of the chalcogen- π interaction that can be extended to many other complexes. Our calculations showed that the stabilization of the complexes increases with the difference in electronegativity between the chalcogen and the halogen substituents. Therefore the complexes with the highest non covalent interaction strength are those that contain the most electropositive chalcogen, *i.e.* tellurium, and fluorine, the most electronegative halogen. In addition to that, **2-but** was seen to form the strongest chalcogen- π bond among the selected unsaturated substrates.

Orbital analyses of the selected complexes reveal that the main cause of this behavior is that an increase either in the chalcogen electropositivity or in the halogen substituent electronegativity clearly enhances the amplitude of the chalcogenide fragment's LUMO of the chalcogen atom and thus the overlap with the HOMO of the unsaturated substrate resulting in an overall stronger chalcogen- π bond. Also, the different bond strengths computed for the various substrates are explained by orbital analyses, as differences mainly stem from the smaller HOMO-LUMO energy gap found in complexes of **2but**. Consequently, stronger orbital interactions result in complexes with shorter chalcogen- π bond distances and thus a more favorable ΔV_{elstat} term which provides an additional stabilization in complexes with high chalcogen-halogen electronegativity differences and in those with 2-butyne.

Results on the selected model complexes showed that chalcogen- π interaction can be rationalized through quantitative energy decomposition and orbital analyses. These outcomes can be used to predict the strength of the chalcogen bonds in biological system based on the chemical nature of the atoms involved and may be exploited to create powerful design tools to improve the effectiveness of novel artificial compound based on rational and finely tuned non-covalent interactions.

Appendices

Appendix A

***Ab initio* benchmark**

Table A.1: Chalcogen- π bond energies ΔE (in kcal mol⁻¹) computed at DFT and CC levels of theory.^a

| Complex | ΔE (DFT) | ΔE (CC) | Error (DFT-CC) |
|-------------------------------|------------------|-----------------|----------------|
| F₂O...ac | -1.3 | 0.2 | -1.5 |
| F₂S...ac | -4.3 | -2.2 | -2.1 |
| F₂Se...ac | -7.5 | -3.7 | -3.8 |
| F₂Te...ac | -9.4 | -5.7 | -3.7 |
| Cl₂Se...ac | -5.3 | -2.7 | -2.6 |
| Br₂Se...ac | -5.0 | -2.5 | -2.5 |
| I₂Se...ac | -4.3 | -2.1 | -2.2 |
| F₂Se...et | -9.1 | -4.2 | -4.9 |
| F₂Se...2but | -11.3 | -6.4 | -4.9 |
| Average Error | | -3.1 | |

a) Computed with correction for basis-set superposition error. DFT = ZORA-BLYP-D3(BJ)/QZ4Pae//ZORA-BLYP-D3(BJ)/TZ2Psc. CC = DLPNO-CCSD(T)/aug-cc-pV5Z-DK, aug-cc-pVQZ-DK.

Table A.2: Chalcogen- π bond energies ΔE (in kcal mol⁻¹) computed at DFT and CC levels of theory.^a

| Complex | ΔE (DFT) | ΔE (CC) | Error (DFT-CC) |
|-------------------------------|------------------|-----------------|----------------|
| F₂O...ac | -1.3 | 0.1 | -1.4 |
| F₂S...ac | -4.3 | -2.3 | -2.0 |
| F₂Se...ac | -9.4 | -8.4 | -1.0 |
| F₂Te...ac | -7.5 | -4.6 | -2.9 |
| Cl₂Se...ac | -5.4 | -3.4 | -2.0 |
| Br₂Se...ac | -5.1 | -3.5 | -1.4 |
| I₂Se...ac | -4.4 | -3.8 | -0.6 |
| F₂Se...et | -9.1 | -5.6 | -3.5 |
| F₂Se...2but | -11.4 | -7.9 | -3.5 |
| Average error | | -2.0 | |

a) Computed without correction for basis-set superposition error. DFT = ZORA-BLYP-D3(BJ)/QZ4Pae//ZORA-BLYP-D3(BJ)/TZ2Psc. CC = DLPNO-CCSD(T)/aug-cc-pV5Z-DK, aug-cc-pVQZ-DK.

Appendix B

Orbital analysis

Table B.1: HOMO-LUMO energy gap $\Delta\epsilon$ (in eV) and, overlap (S) in all model systems.^a

| Complex | $\Delta\epsilon$ | S | Complex | $\Delta\epsilon$ | S |
|-------------------------------|------------------|------|--------------------------------|------------------|------|
| F₂O...2but | 0.19 | 0.07 | F₂Se...2but | 1.63 | 0.17 |
| Cl₂O...2but | 0.58 | 0.06 | Cl₂Se...2but | 1.55 | 0.12 |
| Br₂O...2but | 0.57 | 0.07 | Br₂Se...2but | 1.41 | 0.11 |
| I₂O...2but | 0.83 | 0.07 | I₂Se...2but | 1.40 | 0.08 |
| F₂O...et | 1.24 | 0.07 | F₂Se...et | 2.39 | 0.19 |
| Cl₂O...et | 1.51 | 0.07 | Cl₂Se...et | 2.40 | 0.15 |
| Br₂O...et | 1.46 | 0.08 | Br₂Se...et | 2.25 | 0.14 |
| I₂O...et | 1.66 | 0.08 | I₂Se...et | 2.22 | 0.13 |
| F₂O...ac | 1.84 | 0.06 | F₂Se...ac | 3.03 | 0.17 |
| Cl₂O...ac | 2.01 | 0.06 | Cl₂Se...ac | 2.91 | 0.14 |
| Br₂O...ac | 1.97 | 0.07 | Br₂Se...ac | 2.75 | 0.13 |
| I₂O...ac | 2.14 | 0.07 | I₂Se...ac | 2.71 | 0.09 |
| F₂S...2but | 2.09 | 0.14 | F₂Te...2but | 1.55 | 0.18 |
| Cl₂S...2but | 1.77 | 0.09 | Cl₂Te...2but | 1.75 | 0.16 |
| Br₂S...2but | 1.49 | 0.08 | Br₂Te...2but | 1.66 | 0.12 |
| I₂S...2but | 1.42 | 0.08 | I₂Te...2but | 1.62 | 0.09 |
| F₂S...et | 3.00 | 0.16 | F₂Te...et | 2.27 | 0.22 |
| Cl₂S...et | 2.67 | 0.11 | Cl₂Te...et | 2.57 | 0.20 |
| Br₂S...et | 2.39 | 0.11 | Br₂Te...et | 2.49 | 0.16 |
| I₂S...et | 2.28 | 0.10 | I₂Te...et | 2.44 | 0.13 |
| F₂S...ac | 3.58 | 0.13 | F₂Te...ac | 2.89 | 0.20 |
| Cl₂S...ac | 3.14 | 0.10 | Cl₂Te...ac | 3.11 | 0.18 |
| Br₂S...ac | 2.84 | 0.10 | Br₂Te...ac | 2.99 | 0.14 |
| I₂S...ac | 2.74 | 0.08 | I₂Te...ac | 2.92 | 0.12 |

a) Computed at ZORA-BLYP-D3(BJ)/QZ4Pae//ZORA-BLYP-D3(BJ)/TZ2Psc.

Chapter 4

The GPx oxidative phase

Adapted from

Bortoli, M.; Torsello, M.; Bickelhaupt, F. M.; Orian, L.

Role of the Chalcogen (S, Se, Te) in the Oxidation Mechanism of the Glutathione Peroxidase Active Site.

ChemPhysChem, **2017**, *18*, 2990-2998

4.1 Introduction

THIS Chapter presents an *in silico* investigation of the first step of the human glutathione peroxidase catalytic cycle (Figure 4.1) through a combined classical and quantum mechanical approach. The mechanistic details of the Cys-, Sec- and Tec-GPx active sites were systematically investigated to determine the different role of the three chalcogens in the oxidation step.

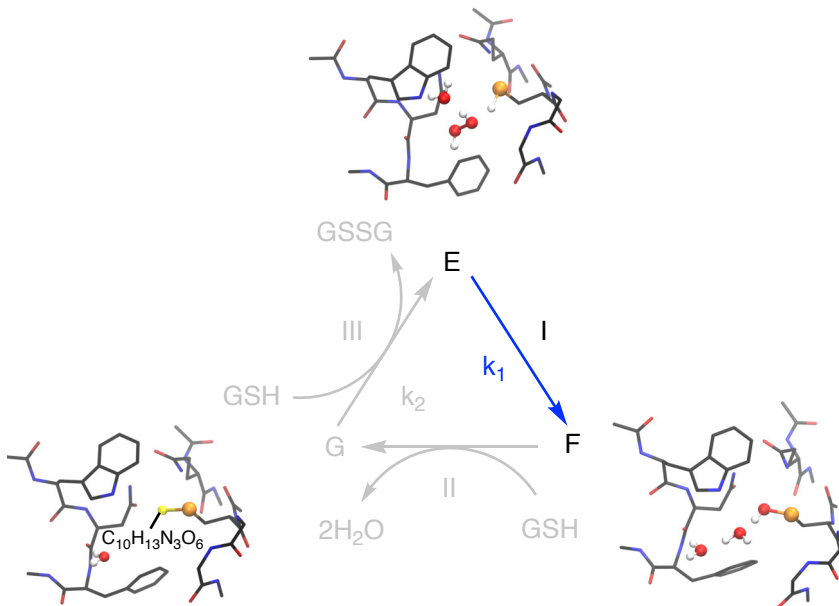


Figure 4.1: Mechanism of organic hydroperoxides reduction catalyzed by Cys/Sec-GPx. E is the reduced enzyme with Cys/Sec in the thiol/selenol form, F represents the oxidized intermediate, *i.e.* the sulfenic/selenenic acid form, G is the disulfide/selenosulfide form. The oxidative step is highlighted in blue.

Studies on a mutant Cys-GPx showed an analogous mechanism than that of Figure 4.1, although with a lower antioxidant activity. [147] In addition to that, recent computational studies on an enzymatic cluster of seven residues forming the active site of GPx supported the idea that a similar mechanism for Cys- and Sec-GPx can be predicted on the basis of DFT calculations. The calculated reaction profiles show that the presence of selenium results in lower energy barrier and in an overall rate acceleration. [148] Morokuma *et. al.* conducted pioneering investigation on the Sec-GPx mechanism using pure QM calculations on a simplified model cluster based on human GPx3, [149] as well as calculations on the entire enzyme employing hybrid quantum mechanics/molecular mechanics (QM/MM) schemes. [150] Their choice of enzymatic cluster in the pure QM calculations is a structure containing five other residues on top of selenocysteine (Gly50, Leu51, Tyr48, Gln83 and Trp157) plus one or two molecules of water. Two possible mechanisms were predicted: a direct one and a two-step one, involving first a proton transfer from the selenol to Gln83 and then the reduction of the peroxide with concomitant formation of F. Activation energies range from 17.1 to 19.1 kcal mol⁻¹ depending on the mechanism, which is in good agreement with an estimate of the experimental value, calculated from kinetic constants, of 14.9 kcal mol⁻¹. [151]

If selenium is indeed rare in biological systems (known selenoproteins in vertebrates amount to 40 and this limited presence is ascribed to the high reactivity of selenium in biological environment) [75], tellurium is altogether absent. However, incorporation of tellurocysteine (Tec) into subtilisin [89] and replacement of Ser9 of glutathione transferase from *Lucilla cuprina* with Tec [88] resulted in two semi-natural proteins that displayed remarkable peroxidase activity. This reinforced the promising results obtained employing organotellurides as antioxidants and opened the possibility of their usage as drugs. [78, 90] In fact, toxicological studies on these compounds concluded that their toxicity is comparable to

that of organoselenides, [74] which represent the canonical class of GPx mimics and have been studied for a few decades. High GPx activity was found also in cyclodextrin-derived tellurium compounds. [152]

To support and clarify results from experimental studies, structural and reactivity investigation of selenium and tellurium enzymes can be carried out with the combination of classical and quantum mechanical (QM) techniques. This approach must be backed by the application of suitable protocols, *i.e.*, in this case, accurate force field parameters for the molecular dynamics (MD) simulations, to arrive at a reliable relaxed structure of the whole enzyme, and a high level of theory for the QM part, involving the choice of state-of-the-art functionals and sufficiently large and flexible basis sets.

Standard force field databases do not contain any parameter for selenium or tellurium and simulations are usually carried out using the parameters derived for sulfur, with the sole modification of the atomic charges, [153–156] or, when parameterization is carried out, it is not specific for a single form of the residue but encompasses multiple oxidation states. [157, 158] This method can be sufficient to perform QM/MM calculations, as those of Morokuma *et. al.*, [150] but is not accurate enough to discern the different behavior of the chalcogen in the active site during MD simulations. Therefore, to obtain the results presented in this Chapter, first an accurate parameterization of the Sec and Tec residues was carried out and two novel sets of parameters of their AMBER force field were obtained (for details see Appendix C). These parameters were used in the MD simulation of GPx4 starting from its crystallographic structure, but thanks to their flexibility they may be employed in the future for the simulation of other seleno- and telluroproteins, both natural and artificial. From an MD snapshot, an enzymatic cluster representing the active site was extracted and focus was centered on the initial step of the peroxidase cycle, *i.e.* the oxidation of the Cys, Sec or Tec and the reduction of a H₂O₂ molecule. In particular, the different

role of the chalcogen on the energetics and mechanism will be discussed and analyzed through density functional theory (DFT) calculations and the application of activation strain analyses (ASAs) and energy decomposition analyses (EDAs).

4.2 Methods

IN this chapter, a combined classical and quantum mechanical work will be presented. In the following Sections a brief description of the methodologies employed in the molecular MD simulations and in the DFT calculations will be outlined. For a more exhaustive reference see Chapter 2.

4.2.1 MD simulations

The GPx4 structure was chosen as model structure for a peroxidase enzyme (pdb entry: 2OBI). [159] To run MD simulations, we chose the Assisted Model Builder with Energy Refinement (AMBER) Hamiltonian (Equation (2.11)) coupled with the ff14SB set of parameters, [110] implemented in the Amber 2016 software package, [115] due to its good performance in treating biomolecules .

Simulations were straightforward for the Cys-GPx because all the necessary parameters were already available in ff14SB. [110] On the other hand, for Sec and Tec based enzymes, no suitable parameters were available. Thus, we chose to derive them for these residues, adapting a protocol and tools recently employed to derive GAFF (General AMBER Force Field) parameters for a series of organochalcogen compounds. [160] A detailed description of this protocol is reported in Appendix C.

The new parameters were employed in the MD simulations of the Sec-GPx and Tec-GPx, whose initial structures, together with that of Cys-GPx, were solvated with TIP3P water (6942 water molecules for Cys-GPx, 6946 for Sec-GPx and 7121

for Tec-GPx). [161] Afterwards, the systems were minimized and equilibrated for 200 ps at 300 K and 1 bar. The MD simulations were carried out for 500 ns in the same conditions of the equilibration, using the Langevin thermostat and a Monte Carlo barostat.

4.2.2 DFT Calculations

A suitable enzymatic cluster was chosen to represent the GPx active site. Seven residues (Cys/Sec/Tec46, Gly47, Lys48, Gln81, Trp136, Asn137 and Phe138) were selected, in agreement with previous work. [148] The initial structures were retrieved from a snapshot of the MD simulations (Figure 4.2 a). The extracted residues were suitably capped using ACE (C(O)CH₃) and NME (N(H)CH₃) groups at the N- and C-terminal positions, respectively, to simulate the removed protein chain. Lys48 was substituted with a Gly residue to reproduce a peptide bond connecting Gly47 to the non-conserved following residues (Figure 4.2 b). [148] They were optimized keeping the backbone atoms frozen, to avoid loss of the structural effects of the large portion of GPx excluded from QM computations. The dispersion-corrected hybrid B3LYP-D3(BJ) [132, 162–164] functional was used. It comprises the Becke three parameter functional [162] and the Lee, Yang, and Parr [132–134] correlation term with an added dispersion contribution term developed by Grimme *et al.* [136]

The 6-311G(d,p) basis set was used for all atoms, except selenium and tellurium, for which Dunning's correlation consistent basis set of triple zeta quality (cc-pVTZ) was used; for tellurium, pseudo potentials were also included (cc-pVTZ-PP). The level of theory employed for partially constrained geometry optimizations is denoted B3LYP-D3(BJ)/6-311G(d,p),cc-pVTZ(-PP). Frequencies calculations were subsequently performed to verify the absence of imaginary frequencies for minima and the presence of a single imaginary frequency for transitions states, associated

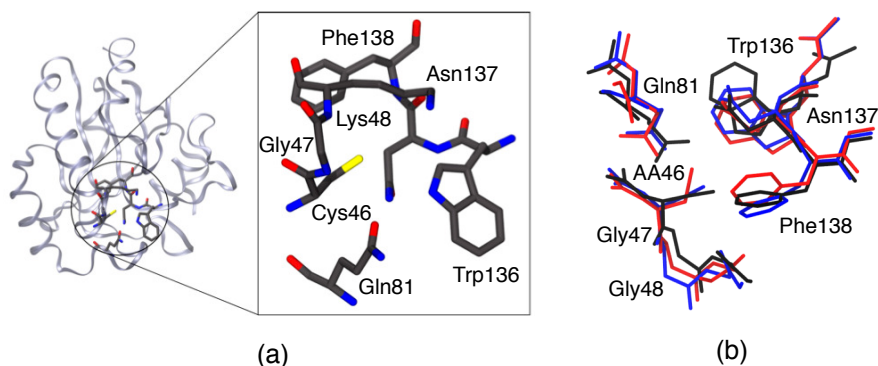


Figure 4.2: Cluster of amino acids extracted from Cys-GPx (a) and superposition of the optimized geometries of the Cys- (black), Sec- (red) and Tec-GPx (blue) clusters (b); AA46= Cys46, Sec46, Tec46; level of theory: B3LYP-D3(BJ)/6-311G(d, p), cc-pVTZ(-PP).

to the correct vibrational mode. Single point energy calculations were carried out in gas phase at B3LYP-D3(BJ)/6-311+G(d,p),cc-pVTZ(-PP) and the correction for the condensed phase was computed using the SMD solvation model (SMD-B3LYP-D3(BJ)/6-311+G(d,p),cc-pVTZ(-PP)) as implemented in Gaussian09. [103, 165] A dielectric constant of 4.0 was employed to reproduce the protein environment, in agreement with previous studies. [166] Thermodynamic corrections resulting from gas phase frequency calculations (same level of theory employed in the geometry optimizations) were added; they refer to a temperature of 298.15 K and a pressure of 1.00 atm. [167] If not differently stated, Gibbs free energies in condensed phase are discussed in the Results section.

Activation strain analyses (ASAs) [98, 99] were employed to quantitatively analyze the reaction potential energy surface (PES) of the oxidation of Cys-E and Sec-E. In this model, the system is partitioned into two fragments and the total energy, (E), is decomposed into the sum of strain energy ΔE_{strain} and the

interaction energy ΔE_{int} . ΔE_{int} was further analyzed in the framework of the Kohn-Sham molecular orbital model using a quantitative energy decomposition analysis (EDA, see Chapter 2 Section 2.1.2). In these calculations, scalar relativistic effects were accounted for through the zeroth-order regular approximation (ZORA). [137] The BLYP [132–135] functional was used, in combination with the TZ2P basis set for all elements. The TZ2P basis set is a large uncontracted set of Slater-type orbitals (STOs), of triple- ζ quality and augmented with two sets of polarization functions on each atom: 2p and 3d in the case of hydrogen, 3d and 4f in the case of carbon, nitrogen, oxygen and sulfur, and 4d and 4f in the case of selenium. Dispersion corrections were included with the D3 scheme with inclusion of the Becke Johnson damping (D3(BJ)), developed by Grimme *et al.* [136] This level of theory is denoted ZORA-BLYP-D3(BJ)/TZ2P.

4.3 Results and discussion

AFTER an MD simulation of 500 ns (Figure 4.3 a), results show that substitution of the chalcogen atom in residue 46 has limited effects on the atoms positions. Calculated mean square deviations for the backbone of Sec- and Tec-GPx with respect to Cys-GPx are equal to 0.99 Å and 1.18 Å, respectively. Analysis of the root mean square fluctuations (Figure 4.3 b) reveals that the residues have a similar local flexibility, with the exception of residue 79 in Tec-GPx, which is a Gly in the proximity of the catalytic pocket. Such a high variability in a residue that has a low conformational freedom arises from the fact that the loop in which Gly79 is inserted in Tec-GPx rotates back and forth by almost 90° during the whole simulation. Moreover, an important rearrangement during the dynamics is seen in the loop 156-160 of Sec-GPx which is located far from the catalytic active site. Furthermore, water molecules enter and exit the active site during the whole simulation, supporting the idea that the catalytic pocket is solvent accessible and

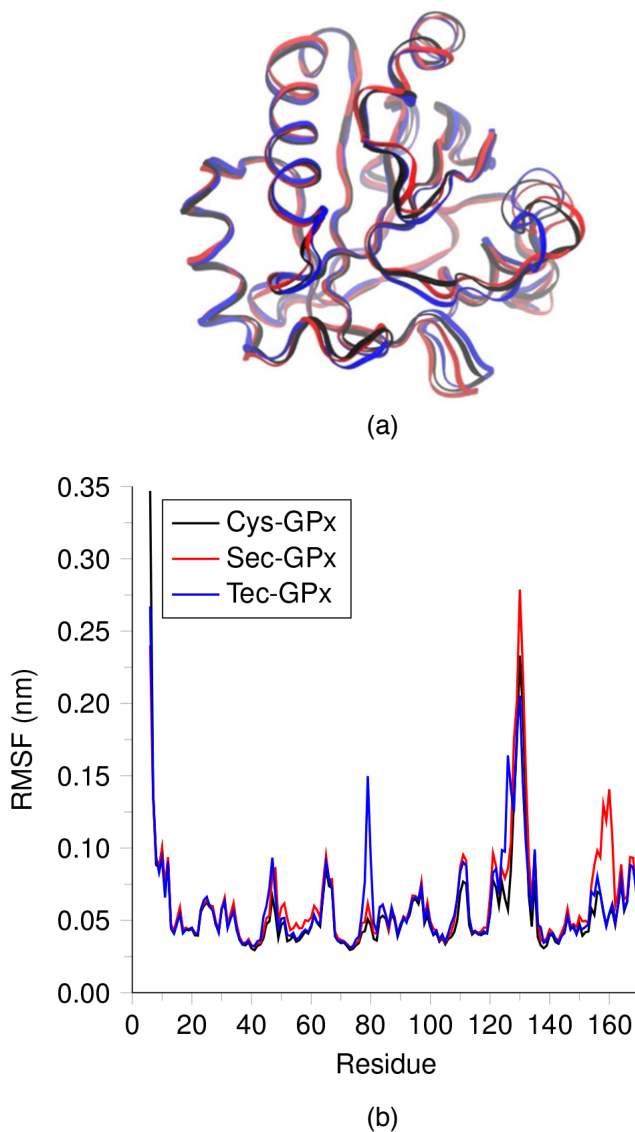


Figure 4.3: MD results: position of the backbone atoms of Cys-GPx (black), Sec-GPx (red) and Tec-GPx (blue) after 500 ns (a) and root mean square fluctuation for each residue along the 500 ns dynamics of Cys-GPx (black), Sec-GPx (red) and Tec-GPx (blue) (b).

the choice to add a water molecule in the enzymatic cluster extracted from the simulation to allow the deprotonation of the selenol, also in view of the fact that at physiological pH selenolate is the most abundant species. [148]

Seven amino acids were selected to model the active site [62] and a cluster comprising these residues was extracted from each MD simulation (Cys-cluster, Sec-cluster and Tec-cluster). Recently, a pure DFT study of the Cys- and Sec-GPx mechanism was carried out employing an analogous cluster, fully optimized in gas phase at the B3LYP/6-31G(d, p) level of theory. [148] Two additional molecules (H_2O_2 and H_2O) were manually inserted in the position that qualitatively provided the best interactions with the Cys/Sec/Tec residue. The resulting adducts ($\text{E}\cdot\text{H}_2\text{O}_2\cdot\text{H}_2\text{O}$) were then optimized¹ (Figure 4.4). Optimized structures have a relative energy of -2.4 , -8.9 and -8.2 kcal mol⁻¹, respectively, compared to the free reactants, *i.e.* the enzymatic cluster (E) plus hydrogen peroxide and water (Table 4.1), which suggests that the active site is designed to favorably contain and coordinate the substrate and water, forming a hydrogen bond network with the surrounding residues (*vide infra*). A substantial difference is found between the Tec-cluster and those with Cys or Sec. In fact, the distance of the chalcogenol proton to the closest oxygen of H_2O_2 is found to be of 4.03 Å in the Tec-cluster, whereas a smaller value is found for the Cys- (2.29 Å) and Sec-clusters (2.21 Å). This precludes any interaction between the tellurol and the H_2O_2 molecule and has important effects on the mechanism of Tec-GPx (*vide infra*).

The position of H_2O and H_2O_2 is found to be similar in the Cys- and Sec-clusters (Figure 4.4 a and b). The latter is always found between the chalcogen atom and the indole moiety of Trp136 but, if in the case of the Cys-cluster it has a favorable interaction with a carbonyl oxygen of the peptide bond between Asn137 and Phe138 in the Sec-cluster the distance from these two residues precludes the

¹As mentioned in Section 4.2.2 all optimization are carried out maintaining the protein backbone frozen.

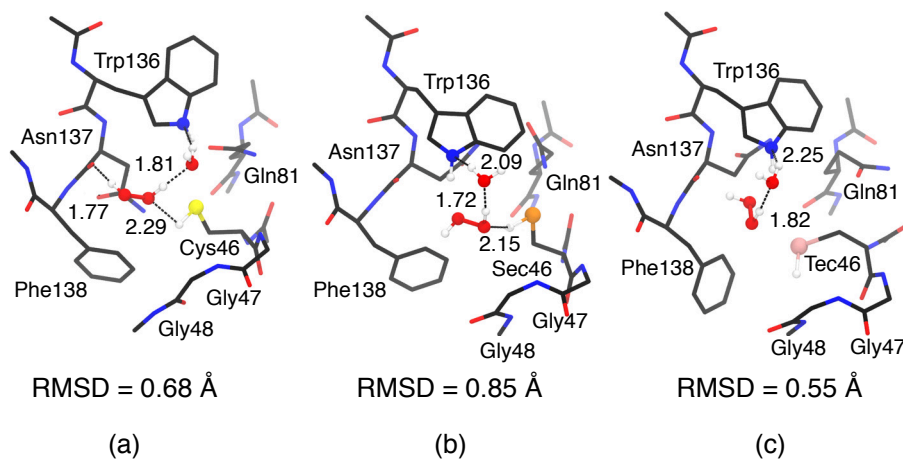


Figure 4.4: Geometries of the optimized adducts E-H₂O₂·H₂O; E=Cys-E (a), Sec-E (b) and Tec-E (c) and RMSDs (hydrogens not included) with respect to the initially optimized clusters; level of theory: B3LYP-D3(BJ)/6-311G(d,p), cc-pVTZ(-PP).

formation of a hydrogen bond. Nevertheless, the slightly different geometry of the Sec-cluster allows the formation of a non covalent bond between an oxygen of H₂O₂ and the hydrogen atom bonded to the N atom in the indole of Trp136. The H₂O molecule is found between the hydrogen peroxide and the indole moiety of Trp136 in both clusters. Its oxygen atom forms a hydrogen bond with an H of H₂O₂ and one of its hydrogen atoms points directly to the N atom of the indole in Trp136. This arrangement of water and H₂O₂ creates a network of hydrogen bonds in both clusters suggesting the possibility of a long range proton transfer from Cys/Sec to Trp136. In fact, computations show how the thiol/selenol proton is transferred to H₂O₂ which in turn transfers one of its protons to the nitrogen in the indole of Trp136. The water molecule is crucial in this step as it acts as a bridge to efficiently transport a proton from H₂O₂ to the indole. The barrier computed for the formation of this charge separated intermediate (E(CS)-H₂O₂·H₂O) is

Table 4.1: Gibbs free energies (kcal mol⁻¹) of the intermediates and transition states in gas phase and in condensed phase.^a

| | Cys-cluster | Sec-cluster | Tec-cluster |
|---|-------------|-------------|-------------|
| Gas phase | | | |
| E | 0.0 | 0.0 | 0.0 |
| E·H ₂ O ₂ ·H ₂ O | -0.4 | -4.0 | -0.8 |
| TS ₁ | 25.3 | 16.6 | |
| E(CS)·H ₂ O ₂ ·H ₂ O | 16.8 | 5.4 | |
| F | -50.2 | -73.6 | |
| Condensed phase | | | |
| E | 0.0 | 0.0 | 0.0 |
| E·H ₂ O ₂ ·H ₂ O | -2.4 | -8.9 | -8.2 |
| TS ₁ | 24.7 | 12.7 | |
| E(CS)·H ₂ O ₂ ·H ₂ O | 15.9 | 2.5 | |
| F | -51.3 | -77.3 | |

a) Computed at (SMD-)B3LYP-D3(BJ)/6-311+G(d,p),cc-pVTZ(-PP)//B3LYP-D3(BJ)/6-311G(d,p),cc-pVTZ(-PP)

lower in the Sec-cluster than in the Cys-cluster (21.6 vs 27.1 kcal mol⁻¹) and the formed structure is less destabilized in the presence of the heavier chalcogen. No subsequent transition state could be located, so we concluded that conversion to F proceeds with no barrier. A similar behavior was found in a recent study in which an analogous cluster was employed. [148]

Mechanistic studies of the Tec-cluster reveal that a pathway involving a proton transfer is not feasible. In fact, all attempts to optimize either the charge separated intermediate or the transition state leading to it systematically failed. Tellurium electronegativity (or lack thereof) plays an important role here: if for sulfur and selenium the S-H and Se-H bond are definitely polarized toward the more electronegative chalcogens, in the case of tellurium its lower electronegativity

(slightly smaller than that of hydrogen) results in an almost pure covalent Te–H bond in which the hydrogen atom has no tendency to move to an adjacent oxygen. Nevertheless, an alternative mechanism for the oxidation of Tec-cluster is possible. It involves a direct nucleophilic attack of the tellurium atom to the hydrogen peroxide molecule. This new pathway has a much smaller activation energy ($10.5 \text{ kcal mol}^{-1}$, Table 4.2) compared to that needed for the formation of the charge separated intermediates in the Cys- and Sec-clusters (27.1 and $21.6 \text{ kcal mol}^{-1}$, respectively), and results in the formation of a highly stabilized telluroxide (E–Te(O)H, $-43.6 \text{ kcal mol}^{-1}$). An analogous behavior was found in the case of several selenides and diselenides, which can be oxidized via a direct mechanism. [168–171] To arrive at the same product of the oxidation of the Cys- and Sec-clusters (F, a sulfenic/selenenic acid) an intramolecular isomerization is needed. The transition state for this process (TS_{iso}) is found at $-11.5 \text{ kcal mol}^{-1}$ and that means that a barrier of $32.1 \text{ kcal mol}^{-1}$ has to be overcome to arrive at the final tellurenic acid (F) at $-72.0 \text{ kcal mol}^{-1}$. However, in the few experimental studies employing tellurium in semi-natural enzymes, [88, 89] the resting form suggested for the Tec residue is that of a tellurinic acid (E–Te(O)OH). Since the catalytic cycle can proceed via the telluroxide or either the tellurenic or tellurinic acid, the formation of this latter species, which can be significant in excess of peroxide, was investigated.

Two different mechanisms are proposed for its formation: one proceeding in a single step to the final structure (Figure 4.5 a) and the other involving the intermediate formation of a hydroxy perhydroxy tellurane (Figure 4.5 b). The results show that the direct mechanism is favored over the stepwise oxidation (Table 4.3) which is in agreement with a recent study on the oxidation of organic selenides. [171] The activation barriers computed for the two reactions are $16.7 \text{ kcal mol}^{-1}$ for the direct mechanism versus $33.1 \text{ kcal mol}^{-1}$ for the formation of the hydroxy perhydroxy tellurane, which is found to be higher in energy than

Table 4.2: Gibbs free energies (kcal mol⁻¹) of the intermediates and transition states for the oxidation and isomerization of the Tec-cluster in gas phase and in condensed phase.^a

| | Gas phase | Condensed phase |
|-----------------------------------|-----------|-----------------|
| E + H ₂ O ₂ | 0.0 | 0.0 |
| E·H ₂ O ₂ | -4.3 | -4.4 |
| TS _{direct} | 6.6 | 6.1 |
| E-Te(O)H | -41.9 | -43.6 |
| TS _{iso} | -10.4 | -11.5 |
| F | -72.5 | -72.0 |

a) Computed at (SMD-)B3LYP-D3(BJ)/6-311+G(d,p),cc-pVTZ-PP//B3LYP-D3(BJ)/6-311G(d,p),cc-pVTZ-PP

the starting telluroxide by 6.7 kcal mol⁻¹. In both cases the most stable structure in the overall mechanism is the final tellurinic acid, lying at -52.6 kcal mol⁻¹ (Figure 4.6). These results suggest that either the telluroxide or the tellurinic acid can be viable starting structures for the continuation of the catalytic cycle, via nucleophilic attack by GSH.

The information on the energetics of the Cys- and Sec-cluster, while highlighting a more favorable thermodynamics in the case of the latter, does not give any insight on the cause of the smaller barrier for selenium. Therefore to understand why this is the case, ASA and EDA were applied, in the gas-phase, as single point calculations on previously B3LYP- D3(BJ)/6-311G(d, p),cc-pVTZ optimized geometries.² The clusters were fragmented into two parts (Figure 4.7 a): the first one composed by Cys/Sec46, Gly47 and Gly48 (f1) and the second containing the rest of the cluster, namely Gln81, Trp136, Asn137, Phe138, the

²See Section 4.2.2 for details on the level of theory employed in this instance and Section 2.1.2 of Chapter 2 for a general explanation of the ASA and EDA models.

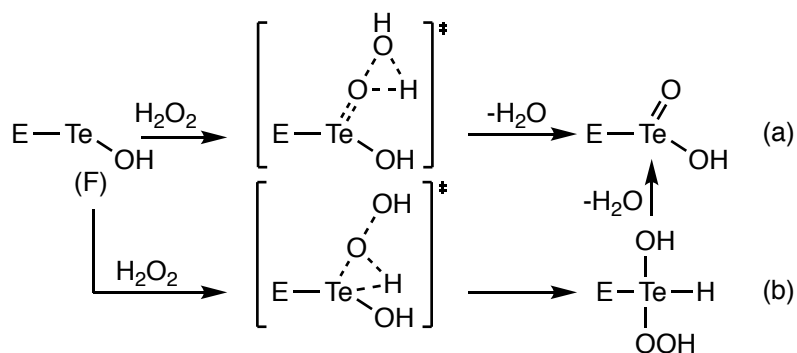


Figure 4.5: Mechanism of formation of tellurinic acid via a direct pathway (a) and a hydroxy perhydroxy tellurane intermediate (b).

Table 4.3: Gibbs free energies (kcal mol^{-1}) for the two possible mechanisms for tellurinic acid formation shown in Figure 4.5.^a

| | Path | |
|-----------------------------------|-------|-------|
| | A | B |
| F + H ₂ O ₂ | 0.0 | 0.0 |
| F·H ₂ O ₂ | -3.8 | -3.8 |
| TS | 12.9 | 29.3 |
| E-Te(OH)(OOH)H | | 2.9 |
| E-Te(O)OH | -52.6 | -52.6 |

a) Computed at SMD-B3LYP-D3(BJ)/6-311+G(d, p),cc-pVTZ-PP//B3LYP-D3(BJ)/6-311G(d, p),cc-pVTZ-PP

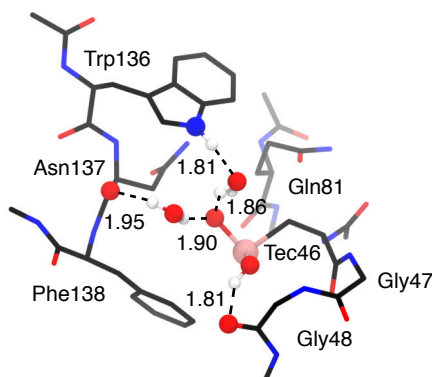


Figure 4.6: Optimized Tec-cluster in the tellurinic acid form (E-Te(O)OH) and relevant interatomic distances (Å). Level of theory: B3LYP-D3(BJ)/6-311G(d,p),cc-pVTZ-PP.

water molecule and the hydrogen peroxide molecule (f2). Analyses were performed only on adducts (E·H₂O₂·H₂O) and TS₁s as the main goal was to establish the reason behind the different activation energy between the sulfur and the selenium containing clusters. The chosen method was ZORA-BLYP-D3(BJ)/TZ2P as it was proved to be reliable in the calculation of energetics of organochalcogen compounds.³ [172] Considering the whole Cys- and Sec- clusters (Figure 4.7 a), we find that E·H₂O₂·H₂O is more stabilized in the Sec-cluster (−19.2 kcal mol^{−1}) than in the Cys-cluster (−12.9 kcal mol^{−1}) with respect to the free reactants. A much smaller strain contribution of 3.0 vs 11.8 kcal mol^{−1} (Table 4.4) is the main cause. Nevertheless, a favorable interplay in the catalytic pocket between the amino acid

³A preliminary benchmark was conducted to compare the results among different levels of theory that were good candidates to be used in the analyses. On top of the optimization level that was in all the cases B3LYP-D3(BJ)/6-311G(d,p), the tested candidates used in the final energy calculations were B3LYP-D3(BJ)/6-311+G(d,p), cc-pVTZ, ZORA-BLYP-D3(BJ)/TZ2P and ZORA-B3LYP/TZ2P. Comparable results were found for all the three levels of theory (see Appendix D)

residues, the substrate and water is present since both these adducts results lower in energy than the free reactants. Again, when looking at the relative energies of the two transition states a very different strain amount is computed: much smaller in the case of the Sec-cluster (59.5 vs 96.0 kcal mol⁻¹). The larger strain computed for the formation of the sulfur containing transition state is caused by the greater variation of the S–H bond length, amounting to 0.66 Å, compared to that found for the Se–H which results to be of 0.47 Å and is thus related to the nature of the chalcogen taking part in the reaction. Moreover, ΔE_{int} compensates for the destabilization caused by strain only in the case of the Sec-cluster since an interaction energy of -62.0 kcal mol⁻¹ makes also TS₁ stabler than the free reactants. This does not happen in the Cys-cluster, for which a computed ΔE_{int} of -87.6 kcal mol⁻¹ is not enough to balance the strain contribution and leaves TS₁ at a relative energy of 8.4 kcal mol⁻¹. This disparity in interaction magnitude originates from a higher ΔE_{oi} contribution in the case of the Sec-cluster (-188.8 vs -143.8 kcal mol⁻¹) and ultimately results in a lower activation barrier (16.7 vs 21.3 kcal mol⁻¹) for the selenium containing enzyme.

A particular interest in the role of the amino acids that form the catalytic pocket, in particular in Gln81, Trp136 and Asn137 which, together with Cys/Sec46 have been experimentally identified as the “catalytic tetrad”, [62] prompted for a further analysis with a simplified f2. First, all the protein residues were removed, leaving a new fragment comprising water and hydrogen peroxide (f2', Figure 4.7 b), then a third structure was considered (f2'', Figure 4.7 c) consisting only of the oxidant molecule. No modifications were applied to f1. The activation strain and energy decomposition analyses were then repeated on these new structures, employing the same geometries used for the whole cluster. Calculations on the Cys-cluster using f2' and f2'' show how the initial adducts are not stabler than the free reactants as they lie at 1.9 and 3.1 kcal mol⁻¹ respectively. Conversely, in the case of the Sec-cluster the stabilization of the adducts is diminished. Nevertheless,

they result lower in energy than the free reactants at -4.0 and -2.5 kcal mol⁻¹, respectively. Activation strain analyses clarify that the stabilization found in the Sec-cluster is due to a higher interaction contribution and a smaller strain energy. Activation barriers are found to be much higher for the reduced fragments, with the Sec-cluster having the lowest ones (33.3 vs 42.9 kcal mol⁻¹ for **f2'** and 53.4 vs 63.5 for **f2''**), confirming the importance of the protein residues and the water molecule in lowering the activation energy for the proton transfer process. In fact, transition state energies are computed to be very different. They increase of a considerable amount for both clusters as they shift from 8.4 to 44.8 and 68.6 kcal mol⁻¹ for the Cys-cluster and -2.5 to 29.3 and 50.9 kcal mol⁻¹ for the Sec-cluster when going from **f2** to **f2'** and **f2''**. From ASA it can be concluded that this much higher energy of the transition states is due to a strong diminution of ΔE_{int} in both structures which weakens significantly from -87.6 to -48.4 and -18.4 kcal mol⁻¹ for the Cys-clusters and from -62.0 to 34.8 and -8.5 kcal mol⁻¹ for the Sec-clusters. On the other hand, strain decreases only moderately when reducing **f2** both in the sulfur and the selenium enzyme, proving that most of this energetic demand arises from the chalcogen-hydrogen bond stretch present in **f1**.

Table 4.4: Activation strain analysis and energy decomposition analysis (electronic energies in gas phase in kcal mol⁻¹) for Cys and Sec clusters.

| Cys-cluster | | | | | | | |
|-----------------|------------|----------------------------|-------------------------|---------------------------|----------------------------|------------------------|--------------------------|
| | ΔE | ΔE_{strain} | ΔE_{int} | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} | ΔE_{disp} |
| f2 | | | | | | | |
| Adduct | -12.9 | 11.8 | -24.7 | 36.9 | -22.9 | -12.8 | -25.9 |
| TS ₁ | 8.4 | 96.0 | -87.6 | 246.8 | -119.6 | -188.8 | -26.0 |
| f2' | | | | | | | |
| Adduct | 1.9 | 5.6 | -3.7 | 6.8 | -5.2 | -2.4 | -2.9 |
| TS ₁ | 44.8 | 93.2 | -48.4 | 207.1 | -95.7 | -157.0 | -2.8 |
| f2'' | | | | | | | |
| Adduct | 3.1 | 4.8 | -1.7 | 5.7 | -3.4 | -2.1 | -1.9 |
| TS ₁ | 68.6 | 87.0 | -18.4 | 206.0 | -84.1 | -138.3 | -2.0 |
| Sec-cluster | | | | | | | |
| | ΔE | ΔE_{strain} | ΔE_{int} | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} | ΔE_{disp} |
| f2 | | | | | | | |
| Adduct | -19.2 | 3.0 | -22.2 | 35.1 | -23.1 | -13.7 | -20.5 |
| TS ₁ | -2.5 | 59.5 | -62.0 | 207.6 | -105.2 | -143.8 | -20.6 |
| f2' | | | | | | | |
| Adduct | -4.0 | 2.2 | -6.2 | 9.4 | -8.4 | -4.2 | -3.0 |
| TS ₁ | 29.3 | 64.1 | -34.8 | 176.5 | -85.9 | -121.8 | -3.6 |
| f2'' | | | | | | | |
| Adduct | -2.5 | 1.2 | -3.7 | 10.4 | -6.7 | -3.2 | -4.2 |
| TS ₁ | 50.9 | 59.4 | -8.5 | 173.1 | -73.1 | -105.2 | -3.3 |

a) Computed at ZORA-BLYP-D3(BJ)/TZ2P//B3LYP-D3(BJ)/6-311G(d, p),cc-pVTZ.

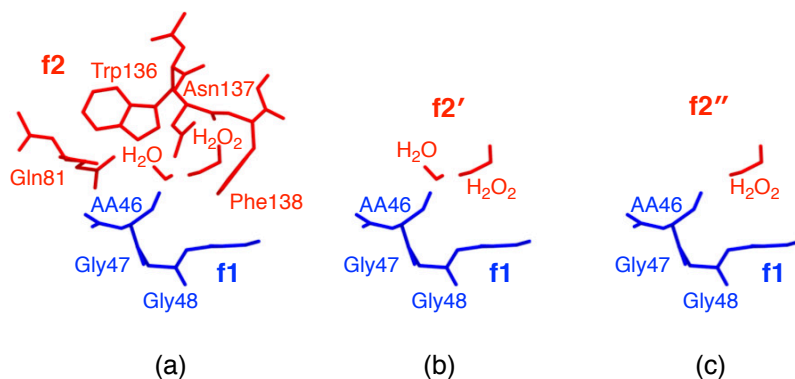


Figure 4.7: Different partitioning of the Cys and Sec clusters; AA46= Cys46, Sec46.

4.4 Conclusions

In the oxidation step of GPx, a similar mechanism is found in the case of the Cys-enzyme and Sec-enzyme. It involves a long range proton transfer and the formation of a charge separated intermediate and was computed to be more favorable for the selenium enzyme. In contrast, all attempts to model an analogous mechanism for the Tec-cluster systematically failed. Therefore an alternative mechanism was modeled. It involves the direct oxidation of Tec to a telluroxide which can be attacked directly by GSH or can isomerize to the tellurenic acid (TeOH). An ulterior oxidation is possible and leads to the formation of to the highly oxidized tellurinic acid, which is the resting state of both artificial telluroenzymes characterized so far. The mechanistic differences and the favorable energetics encountered for Tec- GPx explains the strong peroxidase activity reported for artificial Tec enzymes and prompts for more effort in this research field.

Activation strain analyses performed on the Cys- and Sec-cluster show a more favorable initial oxidation in the case of the selenium enzyme due to the balance between strain and interaction energies that leads to a smaller activation barrier.

The main reason for this is a lower strain calculated for the Sec-cluster originating from the smaller relative and absolute elongation of the Se–H compared to that of the S–H.

Appendices

Appendix C

Parameterization protocol

The following procedure, developed to obtain reliable parameters for the AMBER force fields, is an adaptation/improvement of the one found in ref. [160]. The models to which this will be applied in the current Appendix are the Sec and Tec residues (Figure C.1), but the method is in general suitable for the derivation of the parameters of any non standard small organic molecule.

The aim of this procedure is to derive the necessary parameters to insert into the AMBER Hamiltonian (Equation (2.11)). The first step is to obtain an optimized geometry of the residues, using a QM method. To simulate the peptidic bonds present in the protein, the residues were capped at the N- and C- termini with the ACE (C(O)CH₃) and NME (N(H)CH₃) groups. To cover the energetically relevant conformations of Sec and Tec, 10 conformers were selected and optimized to obtain a set of charges and parameters that can be applied in every occurrence of Sec and Tec, also in proteins other than GPx. These optimizations were carried out at the B3LYP/cc-pVTZ with the addition of a pseudo potential in the case of the tellurium atom (B3LYP/cc-pVTZ-PP).

Atomic charges were calculated with the help of R.E.D. Server Development

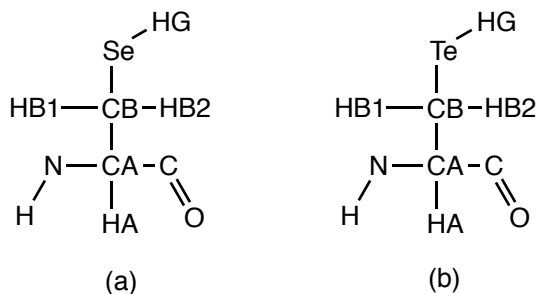


Figure C.1: Atom numbering scheme of the parametrized residues: Sec (a) and Tec (b). Capping groups not shown.

Table C.1: RESP charges derived for Sec and Tec.

| | Se/Te | HG | CA | HA | C |
|-----|------------|------------|-----------|------------|-----------|
| Sec | -0.215 132 | 0.135 408 | 0.072 191 | 0.092 757 | 0.498 517 |
| Tec | -0.090 980 | 0.049 881 | 0.234 828 | 0.179 239 | 0.483 719 |
| | O | N | H | CB | HB2=HB3 |
| Sec | -0.500 977 | -0.289 441 | 0.180 185 | -0.266 657 | 0.146 575 |
| Tec | -0.513 483 | -0.312 666 | 0.179 239 | -0.362 456 | 0.149 589 |

(<http://upjv.q4md-forcefieldtools.org/REDSERVER-Development/>), which is a web-based interface that employs the PyRED program and R.E.D. tools. To obtain highly reproducible charges two orientations were considered for each of the 10 conformers of Sec and Tec, using the rigid-body reorientation algorithm implemented in R.E.D Tools. Calculations were carried out at the Hartree-Fock level using the 6-31G(d) basis set for all the atoms except tellurium for which the Stuttgart Dresden effective core potential basis set was employed. The final charge value for each atom resulted from the simultaneous Restricted Electrostatic Potential (RESP) fit of 18 structures (Table C.1).

Lennard-Jones σ and ϵ parameters were set at 2.1200 Å and 0.2910 kcal mol⁻¹ for Se, and 2.2600 Å and 0.3980 kcal mol⁻¹ for tellurium, consistently with the work in ref. [160]. During the simulations, they were combined according to the Lorentz-Berthelot mixing rules.

Constrained scans (*i.e.* calculations where only the relevant parameter was changed and the rest of the molecule was kept frozen) were carried out at B3LYP/cc-pVTZ (with the addition of a pseudo potential for Tec) on the stablest conformer of Sec and Tec to obtain the force constants of the harmonic contributions (*i.e.* a_i and b_i terms in Equation (2.11)). Stretching and bending parameters were thus computed for CB-X and X-HG bonds (X=Se, Te) using a step of 0.02 Å and 20 scan points, and for CA-CB-X, CB-X-HG and HB2-CB-X angles using a step of 1° and 20 scan points, respectively (see Figure C.1 for the atom labels). The obtained energies were fit with a parabolic equation to obtain the a_i and b_i parameters. The equilibrium distances and amplitudes ($r_{i,eq}$ and $\theta_{i,eq}$ in Equation (2.11)) were taken from the optimized conformer with the lowest energy.

Relaxed scans for each of the ten conformers used in atomic charges derivation were performed for the dihedrals CA-CB-X-HG and HB2-CB-X-HG, the parameters of which were absent in ff14SB. Moreover, although generalized parameters for dihedrals N-CA-CB-X and C-CA-CB-X, in which the chalcogen atom is in terminal position, were already present in ff14SB, we decided to improve these data refitting the coefficients. Calculations were carried out with a step of 10° and 36 steps to complete a full rotation of the dihedral angle. The energies of the resulting ensemble of 720 structures were simultaneously fit with Paramfit, a tool included in AmberTools16, using six Fourier terms for each dihedral. The results of this fitting are shown in Figure C.2. The whole sets of the newly derived parameters for Sec and Tec are shown in Table C.2.

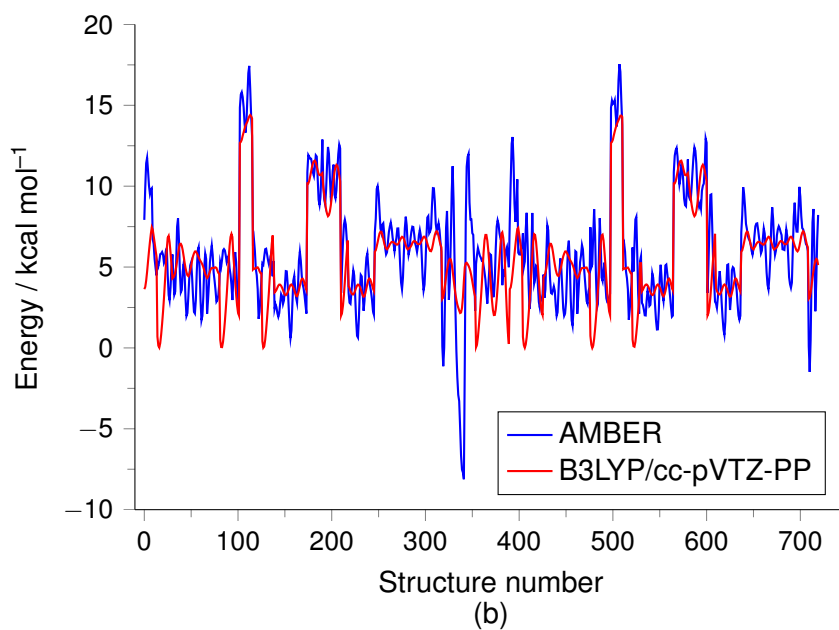
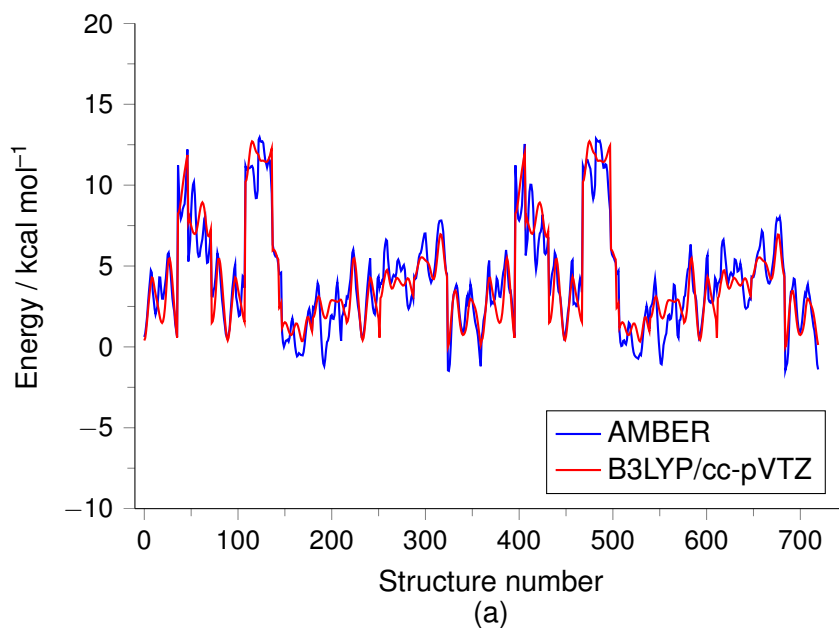


Figure C.2: Fitting obtained with Paramfit of energies of the structures generated with dihedral scans for Sec (a) and Tec (b).

Table C.2: New force field parameters for Sec and Tec residues: equilibrium distances (r_{eq} , in Å), bond force constants (b, in kcal mol⁻¹ Å⁻²), equilibrium angle amplitudes (θ_{eq} , in °), angle force constants (a in kcal mol⁻¹ rad⁻²) and dihedral Fourier series terms coefficients (V_i , in kcal mol⁻¹). See Figure C.1 for atom nomenclature.

| Sec | | | | | | |
|-----------------------|--------------------|--------------------|--------------------|-------------------|-------------------|-------------------|
| Bond | r_{eq} | b | Angle | | θ_{eq} | a |
| CB-Se | 2.00 | 85.3 | CA-CB-Se | | 115.2 | 58.9 |
| Se-HG | 1.46 | 171.5 | CB-Se-HG | | 95.0 | 50.7 |
| | | | HB2-CB-Se | | 108.3 | 48.6 |
| Dihedral ^a | $V_1/2$ | $V_2/2$ | $V_3/2$ | $V_4/2$ | $V_5/2$ | $V_6/2$ |
| CA-CB-Se-HG | 7.0309 (0.0) | 3.5059 (180.0) | 1.0414 (180.0) | 0.1670 (180.0) | 0.0389 (0.0) | 0.0389 (180.0) |
| HB2-CB-Se-HG | 6.7668 (0.0) | 5.1189 (180.0) | 0.6648 (0.0) | 0.3342 (180.0) | 0.0883 (0.0) | 0.0921 (0.0) |
| N-CA-CB-Se | 4.2636 (0.0) | 3.1691 (0.0) | 9.2590 (0.0) | 3.6892 (0.0) | 1.5985 (0.0) | 2.1255 (0.0) |
| C-CA-CB-Se | 9.8947 (180.0) | 8.8968 (180.0) | 4.5907 (180.0) | 2.4628 (180.0) | 3.3382 (180.0) | 0.1956 (0.0) |
| Tec | | | | | | |
| Bond | r_{eq} | b | Angle | | θ_{eq} | a |
| CB-Te | 2.24 | 138.1 | CA-CB-Te | | 116.2 | 59.9 |
| Te-HG | 1.27 | 260.5 | CB-Te-HG | | 93.8 | 47.0 |
| | | | HB2-CB-Te | | 107.8 | 44.1 |
| Dihedral ^a | $V_1/2$ | $V_2/2$ | $V_3/2$ | $V_4/2$ | $V_5/2$ | $V_6/2$ |
| CA-CB-Te-HG | 15.7894 (0.0) | 0.1853 (180.0) | 1.2620 (180.0) | 0.6188 (0.0) | 0.0168 (180.0) | 0.0215 (0.0) |
| HB2-CB-Te-HG | 14.3900 (0.0) | 1.3997 (180.0) | 0.5583 (0.0) | 0.1955 (0.0) | 0.2367 (0.0) | 0.0040 (0.0) |
| N-CA-CB-Te | 4.6772 (0.0) | 8.3891 (0.0) | 15.3909 (180.0) | 0.9541 (180.0) | 5.0529 (0.0) | 5.1427 (0.0) |
| C-CA-CB-Te | 18.0549 (180.0) | 12.7329 (180.0) | 6.9290 (180.0) | 2.5147 (180.0) | 0.1196 (180.0) | 0.4417 (0.0) |

a) Corresponding γ values in °, in parentheses.

Appendix D

Benchmarks

Table D.1: Activation strain analysis (electronic energies in gas phase, in kcal mol⁻¹) for Cys and Sec clusters.^a

| | ΔE_{strain} | ΔE_{int} | ΔE |
|---|----------------------------|-------------------------|------------|
| E-SH·H ₂ O ₂ ·H ₂ O | 12.5 | -26.8 | -14.3 |
| TS ₁ | 104.6 | -97.3 | 7.3 |
| E-SeH·H ₂ O ₂ ·H ₂ O | 3.6 | -24.0 | -20.4 |
| TS ₁ | 65.4 | -67.2 | -1.8 |

a) Computed at B3LYP-D3(BJ)/6-311+G(d,p),cc-pVTZ//B3LYP-D3(BJ)/6-311G(d,p),cc-pVTZ.

Table D.2: Activation strain analysis and energy decomposition analysis (electronic energies in gas phase, in kcal mol⁻¹) for Cys and Sec clusters.^a

| Cys-cluster | | | | | | |
|-----------------|------------|----------------------------|-------------------------|---------------------------|----------------------------|------------------------|
| | ΔE | ΔE_{strain} | ΔE_{int} | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} |
| f2 | | | | | | |
| Adduct | 3.9 | 6.9 | -3.0 | 31.4 | -22.4 | -12.0 |
| TS ₁ | 30.4 | 101.3 | -70.9 | 233.3 | -118.3 | -185.9 |
| f2' | | | | | | |
| Adduct | 3.0 | 4.5 | -1.5 | 5.9 | -5.2 | -2.2 |
| TS ₁ | 51.1 | 100.1 | -49.0 | 199.8 | -95.3 | -153.5 |
| f2'' | | | | | | |
| Adduct | 3.9 | 4.0 | -0.1 | 5.1 | -3.4 | -1.8 |
| TS ₁ | 80.5 | 96.4 | -15.9 | 199.4 | -82.6 | -132.7 |
| Sec-cluster | | | | | | |
| | ΔE | ΔE_{strain} | ΔE_{int} | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} |
| f2 | | | | | | |
| Adduct | -2.6 | 2.6 | -5.2 | 30.3 | -22.6 | -12.9 |
| TS ₁ | 19.8 | 67.3 | -47.5 | 198.8 | -104.6 | -141.7 |
| f2' | | | | | | |
| Adduct | -1.7 | 2.2 | -3.9 | 8.3 | -8.4 | -3.8 |
| TS ₁ | 36.1 | 69.8 | -33.7 | 172.5 | -86.1 | -120.1 |
| f2'' | | | | | | |
| Adduct | -0.4 | 1.8 | -2.2 | 7.9 | -7.1 | -3.0 |
| TS ₁ | 61.4 | 67.6 | -6.2 | 169.4 | -72.9 | -102.7 |

a) Computed at ZORA-B3LYP/TZ2P// B3LYP-D3(BJ)/6-311G(d,p),cc-pVTZ.

Chapter 5

Mimicking the GPx reductive phase

Adapted from

Bortoli, M.; Wolters, L. P.; Orian, L.; Bickelhaupt, F. M.

Addition–Elimination or Nucleophilic Substitution? Understanding the Energy Profiles for the Reaction of Chalcogenolates with Dichalcogenides

Journal of Chemical Theory and Computation, **2016**, *12*, 2752-2761

5.1 Introduction

IN recent years, much attention was directed towards organoselenium compounds. They can be used in organic chemistry as efficient catalysts, *e.g.* in the synthesis of alkenes via a β -elimination, [173] and were seen to be effective

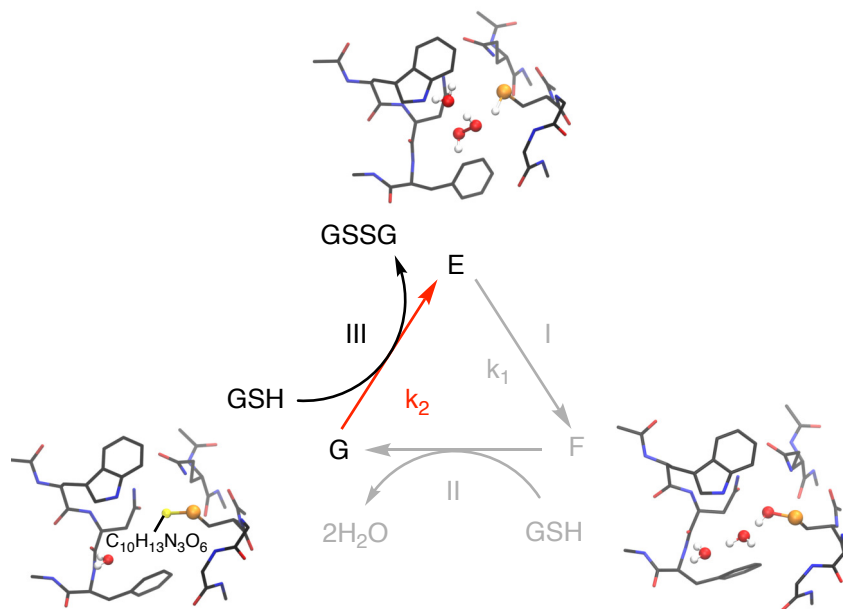


Figure 5.1: Mechanism of organic hydroperoxides reduction catalyzed by Cys/Sec-GPx. E is the reduced enzyme with Cys/Sec in the thiol/selenol form, F represents the oxidized intermediate, *i.e.* the sulfenic/selenenic acid form, G is the disulfide/selenosulfide form. The second reductive step is highlighted in red.

also in ecofriendly solvents, such as water, [174, 175]. Besides an industrial oriented use, a pharmacological potential of these molecules has been recognized. [71] Among their many applications in the medicinal field, their ability to reduce hydrogen peroxide and organic hydroperoxides, mimicking the activity of glutathione peroxidase (GPx), is certainly one of the most interesting.

However, small organoselenium compounds still suffer from many limitations that prevent their employment as efficient anti-oxidant drugs. Experimental [176–179] and computational [73, 79] studies on these compounds and on their

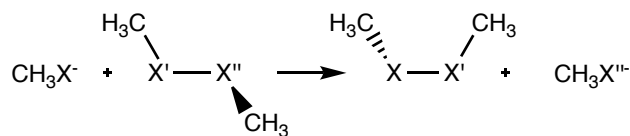


Figure 5.2: Model reactions scheme. X, X', X'' = S, Se, Te

sulfur analogs were mainly focused on nucleophilic attack at the chalcogen by a thiolate or a selenolate. The former reaction occurs also in the GPx mechanism in the two reductive phases (Figure 5.1) but, since attack at selenium seems to be favored, in the small mimics a thiol exchange reaction is expected in the last reductive step which prevents the regeneration of the active molecule. Moreover, the attacking thiol nature may have a dramatic effect on reactivity. [180, 181]

Siblings to organoselenium compounds, also small organotellurides possess a high peroxidase activity. [74, 78, 182, 183] Recently some promising works on semi-natural tellurium enzymes showed how mutated proteins containing tellurocysteine can display GPx-like activity. [89, 88] However, since mechanistic studies on tellurium based GPx mimics are still rare and their biological and pharmacological effects mostly unknown, their application in pharmacology and medicine is still under debate and more extensive studies on the subject are needed.

In this Chapter, a quantum mechanical analysis of the mechanism and energetics of the substitution reaction between a methyl chalcogenolate, acting as nucleophile, and a dimethyl dichalcogenide, acting as substrate, using relativistic density functional theory is presented (Figure 5.2), with the aim to elucidate: (a) the mechanism of group (ii) GPx mimics, (b) the effect of the variation of the substrate and/or nucleophile and the addition of solvent and (c) the fundamental differences between sulfur, selenium and tellurium redox chemistry which can be used to investigate more complex systems.

5.2 Methods

CALCULATIONS were done with scalar relativistic effects accounted for using the zeroth-order regular approximation (ZORA). [184] The OLYP [132–134, 185] density functional was used, in combination with the TZ2P basis set for all elements. The TZ2P basis set is a large uncontracted set of Slater-type orbitals (STOs), is of triple- ζ quality, and has been augmented with two sets of polarization functions on each atom: 2p and 3d in the case of H, 3d and 4f in the case of C, and 3d and 4f (S), 4d and 4f (Se), and 5d and 4f (Te) for the chalcogens. An auxiliary set of s, p, d, f, and g STOs was used to fit the molecular density and to represent the Coulomb and exchange potentials accurately in each SCF cycle. The frozen-core approximation was employed: up to 1s for C, up to 2p for S, up to 3p for Se, and up to 4p for Te. The level of theory is thus indicated as ZORA-OLYP/TZ2P.

Equilibrium and transition-state geometries were fully optimized using analytical gradient techniques. All structures were verified by frequency calculations. For minima, all normal modes have real frequencies, whereas transition states have one normal mode with an imaginary frequency. The character of the normal mode associated with the imaginary frequency was analyzed to ensure that the correct transition state was found. For a representative set of reactions, intrinsic reaction coordinate (IRC) calculations have been performed to obtain the potential energy surfaces (PES) along a selected reaction coordinate.

Solvent effects were included using the conductor-like screening model (COSMO). [102] We have used a solvent-excluding surface with an effective radius for water of 1.93 Å, derived from the macroscopic density and the molecular mass, and a relative dielectric constant of 78.39. The empirical parameter in the scaling function in the COSMO equation was chosen as 0.0. The radii of the atoms were taken to be MM3 radii, [186] divided by 1.2, giving 1.350 Å for H, 1.700 Å for C, 1.792 Å for S, 1.908 Å for Se, and 2.033 Å for Te.

Finally, the model systems were partitioned into two fragments (the nucleophilic methyl chalcogenolate and the dichalcogenide) and analyzed with the aid of the activation strain model and the energy decomposition analysis to arrive at a better understanding on how barriers of key elementary reactions depend on the nature of the reactants.

5.3 Results and discussion

EIGHTEEN model reactions, like those of Figure 5.2, were considered. A shorthand notation was chosen to refer to different reactions in a clear manner and will be used in the following. Each compound is denoted by the chalcogen atom(s) it contains and its charge, omitting the methyl groups. Moreover for the dimethyl chalcogenides the first atoms mentioned is the one who is subject to the nucleophilic attack. For example the reaction in which a methyl thiolate attacks a dimethyl selenosulfide would be written as $\text{S}^- + \text{SeS}$ if the attack occurs at selenium, whereas it will be referred as $\text{S}^- + \text{SSe}$ if the substitution happens at sulfur.

The general behavior in gas-phase of these processes will be studied together with the effects the presence of the solvent (in this case water) has on the reaction mechanism. A more detailed discussion will be concentrated on the $\text{S}^- + \text{SS}$, $\text{S}^- + \text{SSe}$ and $\text{S}^- + \text{SeS}$ reactions (Figure 5.3), which are the closest models to the second reductive step of sulfur and selenium GPx (Figure 5.1) and the unwanted thiol exchange reaction, respectively.

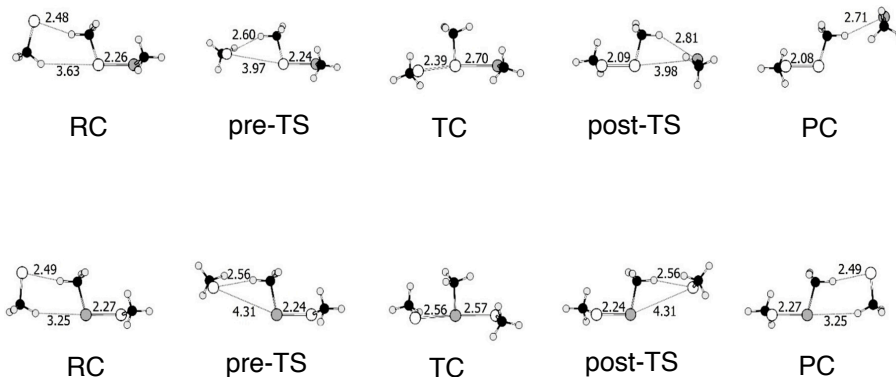


Figure 5.3: Structures of the stationary points for reactions $S^- + SSe$ (a) and $S^- + SeS$ (b)

5.3.1 General gas-phase profiles

The energies of the stationary points of all the model reactions are reported in Table 5.1. The general mechanism found for these reactions in the gas phase is that of addition-elimination or single-well S_N2 . Similar findings have been reported by Bachrach *et al.* [187] and analogous reaction schemes were found in the case of third- and higher-period electrophiles such as silicon, phosphorus and heavier group 14 elements. [188–191] The reactions begin with the formation of a weakly bound reactant complex (RC) in which the chalcogenolate (nucleophile) is coordinated to a methyl group of the dichalcogenide (substrate). These structures are found to be 6–12 kcal mol⁻¹ more stable than the free reactants (R). The structure is then modified as the nucleophile moves away from the methyl group and toward the X' chalcogen forming a first transition state (pre-TS) in which the distance of the chalcogenides of the nucleophile and substrate ranges from 3.75 to 4.66 Å. The subsequent transition complex (TC) is lower in energy with the lowest structure being that of $S^- + TeTe$ with a relative energy of -28.2 kcal mol⁻¹ (Table

Table 5.1: Energies relative to reactants (in kcal mol⁻¹) of stationary points of the model reactions in gas phase.^a

| | RC | pre-TS | TC | post-TS | PC | P |
|------------------------------|-------|--------|-------|---------|-------|-------|
| S⁻ + SS | -8.5 | -7.2 | -10.5 | -7.2 | -8.5 | 0.0 |
| S⁻ + SSe | -9.7 | -8.1 | -13.5 | -11.3 | -12.6 | -5.5 |
| S⁻ + STe | -10.5 | -8.5 | -14.8 | -14.4 | -15.9 | -10.4 |
| S⁻ + SeS | -10.0 | -9.3 | -19.5 | -9.3 | -10.0 | 0.0 |
| S⁻ + SeSe | -10.9 | -10.1 | -20.9 | -12.0 | -12.1 | -4.1 |
| S⁻ + TeS | -11.5 | -11.0 | -27.8 | -11.0 | -11.5 | 0.0 |
| S⁻ + TeTe | | | -28.2 | | | -7.5 |
| Se⁻ + SS | -7.1 | -5.8 | -8.0 | -2.6 | -4.2 | 5.5 |
| Se⁻ + SSe | -8.1 | -6.6 | -10.9 | -6.6 | -8.1 | 0.0 |
| Se⁻ + STe | -8.9 | -7.1 | -12.1 | | | -4.8 |
| Se⁻ + SeS | -8.0 | -7.9 | -16.8 | -5.9 | -6.8 | 4.1 |
| Se⁻ + SeSe | -9.1 | -8.4 | -18.3 | -8.4 | -9.1 | 0.0 |
| Se⁻ + TeS | | | -24.8 | | | 3.5 |
| Se⁻ + TeTe | | | -25.2 | | | -4.1 |
| Te⁻ + SS | -5.5 | -3.9 | -4.4 | 1.9 | -0.1 | 10.4 |
| Te⁻ + TeS | | | -20.6 | | | 7.5 |
| Te⁻ + TeSe | | | -21.2 | | | 4.1 |
| Te⁻ + TeTe | | | -21.2 | | | 0.0 |

a) Computed at ZORA-OLYP/TZ2P.

5.1). The optimized geometry shows that all the methyl groups adopt a skewed conformation in the TCs. This can be explained looking, for example, at the $\text{S}^- + \text{SS}$ reaction in which SS results already skewed for electronic reasons, as each methyl group binds to a different π^* SOMO of the triplet disulfide fragment, and the incoming nucleophile tries to minimize steric repulsion with the substituent of the central S.

For selected reactions, *i.e.* those of biological interest, namely $\text{S}^- + \text{SS}$, $\text{S}^- + \text{SSe}$ and $\text{S}^- + \text{STe}$ the stability of the complexes with coplanar methyls was studied. Minima for TCs with two coplanar *trans* methyl groups were located. The differences in energy with the analogous skewed structures is small and only in the case of $\text{S}^- + \text{SS}$ this structure is more stable than the skewed one (see Table 5.2). Attempts to optimize also the reactant substrate with the methyl groups in a coplanar geometry were not successful because, if in the dichalcogenide the torsional barriers relative to the disulfide bond are found to amount to 6.5 and 12.5 kcal mol⁻¹ (see Appendix E) in the corresponding TCs they are found to be much lower (0.5 and 3.5 kcal mol⁻¹) due to the changes in the electronic and steric environment caused by the presence of the additional substituent bonded to the central chalcogenide. However, an in-depth characterization of the different conformers is beyond the scope of the present chapter and, due to the negligible differences between analogous structures, was not attempted.

The final stages of the reaction in Figure 5.2 involve the formation of a second transition state (post-TS) and a PC that leads, eventually, to the elimination of the methyl thiolate group. These last two structures resemble the pre-TS and RC of the inverse reaction, respectively. Therefore the main focus will be set on the reactions up until the TC because the second part is already considered among the studied mechanisms, although in reverse sense.

Energetics of these systems in the gas phase reveal two possible mechanisms: triple- or single-well energy profiles. Most of the reactions proceed via a triple

well energy profile (wells correspond to RCs, TCs and PCs) with some having TCs that are moderately stabilized with respect to their pre-TS, such as $\text{S}^- + \text{SS}$ in which the two structures have a relative energy of -7.5 and -10.5 kcal mol $^{-1}$, and other that more closely resemble a single well reaction due to the larger span of the energies of the stationary points that for, *e.g.*, $\text{S}^- + \text{TeS}$ are found between -11.0 and -27.8 kcal mol $^{-1}$. A few reactions display a true single well energy profile instead. They all involve attack at tellurium and since no RC, PC or TS could be optimized, free reactants are expected to form TC without any energetic barrier.

Table 5.2: Selected interatomic distances (in Å), activation strain analysis and energy decomposition analysis (electronic energies in gas phase, in kcal mol $^{-1}$ for all model reactions).^{a,b}

| | $d_{XX'}$ | $d_{X'X''}$ | ΔE | ΔE_{strain} | ΔE_{int} | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} |
|---|-----------|-------------|------------|----------------------------|-------------------------|---------------------------|----------------------------|------------------------|
| $\text{S}^- + \text{SS}$ | | | | | | | | |
| RC | 4.78 | 2.08 | -8.5 | 0.9 | -9.4 | 12.2 | -9.8 | -11.8 |
| pre-TS | 3.75 | 2.10 | -7.2 | 0.6 | -7.8 | 12.2 | -7.6 | -12.4 |
| TC | 2.46 | 2.48 | -10.5 | 17.3 | -27.9 | 109.3 | -61.9 | -75.3 |
| TC ^c | 2.54 | 2.38 | -10.8 | 14.6 | -25.5 | 96.1 | -53.9 | -67.7 |
| $\text{S}^- + \text{SSe}$ | | | | | | | | |
| RC | 5.00 | 2.26 | -9.7 | 1.2 | -10.9 | 12.2 | -9.2 | -13.9 |
| pre-TS | 3.97 | 2.24 | -8.1 | 0.3 | -8.4 | 11.2 | -7.3 | -12.3 |
| TC | 2.39 | 2.70 | -13.5 | 22.3 | -35.8 | 142.2 | -79.2 | -98.8 |
| TC ^c | 2.50 | 2.57 | -13.5 | 14.9 | -28.5 | 107.2 | -58.9 | -76.8 |
| post-TS | 2.09 | 3.98 | -11.3 | 69.8 | -81.1 | 295.4 | -166.7 | -209.8 |
| PC | 2.08 | 5.00 | -12.6 | 69.3 | -81.8 | 308.6 | -172.6 | -217.9 |
| $\text{S}^- + \text{STe}$ | | | | | | | | |
| RC | 4.69 | 2.47 | -10.5 | 1.7 | -12.1 | 12.7 | -9.0 | -15.8 |
| pre-TS | 4.07 | 2.46 | -8.5 | 0.8 | -9.3 | 8.6 | -5.3 | -12.5 |

continued overleaf

Table 5.2: Selected interatomic distances (in Å), activation strain analyses and energy decomposition analyses (electronic energies in gas phase, in kcal mol⁻¹ for all model reactions).^{a,b} (continued)

| | $d_{XX'}$ | $d_{X'X''}$ | ΔE | ΔE_{strain} | ΔE_{int} | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} |
|-----------------------------|-----------|-------------|------------|----------------------------|-------------------------|---------------------------|----------------------------|------------------------|
| TC | 2.28 | 3.11 | -14.8 | 30.2 | -45.0 | 187.0 | -102.2 | -129.8 |
| TC ^c | 2.47 | 2.81 | -14.7 | 15.7 | -30.5 | 116.6 | -62.2 | -84.8 |
| S⁻ + SeS | | | | | | | | |
| RC | 4.74 | 2.27 | -10.0 | 1.2 | -11.2 | 12.3 | -9.9 | -13.7 |
| pre-TS | 4.31 | 2.24 | -9.3 | 0.4 | -9.7 | 11.1 | -9.5 | -11.3 |
| TC | 2.56 | 2.57 | -19.5 | 13.7 | -33.2 | 108.5 | -70.8 | -70.9 |
| post-TS | 2.24 | 4.31 | -9.3 | 72.2 | -81.5 | 246.8 | -158.0 | -170.3 |
| PC | 2.27 | 4.74 | -10.0 | 67.0 | -77.0 | 233.0 | -149.2 | -161.0 |
| S⁻ + SeSe | | | | | | | | |
| RC | 4.83 | 2.42 | -10.9 | 1.3 | -12.2 | 12.2 | -9.4 | -14.9 |
| pre-TS | 4.46 | 2.38 | -10.1 | 0.2 | -10.3 | 10.0 | -8.5 | -11.8 |
| TC | 2.54 | 2.72 | -20.9 | 14.3 | -35.2 | 115.5 | -73.6 | -77.1 |
| S⁻ + TeS | | | | | | | | |
| RC | 4.73 | 2.49 | -11.5 | 1.8 | -13.4 | 15.1 | -11.4 | -17.1 |
| pre-TS | 4.66 | 2.46 | -11.0 | 0.8 | -11.9 | 10.3 | -10.0 | -12.2 |
| TC | 2.71 | 2.71 | -27.8 | 10.1 | -37.9 | 106.4 | -78.2 | -66.0 |
| S⁻ + TeTe | | | | | | | | |
| TC | 2.69 | 3.09 | -28.2 | 10.7 | -38.9 | 114.9 | -80.6 | -73.2 |
| Se⁻ + SS | | | | | | | | |
| RC | 5.00 | 2.08 | -7.1 | 0.7 | -7.7 | 9.8 | -8.2 | -9.3 |
| pre-TS | 3.98 | 2.09 | -5.8 | 0.5 | -6.2 | 10.9 | -7.0 | -10.2 |
| TC | 2.70 | 2.39 | -8.0 | 13.9 | -21.8 | 86.4 | -49.5 | -58.7 |
| Se⁻ + SSe | | | | | | | | |
| RC | 5.07 | 2.20 | -8.1 | 0.8 | -8.8 | 8.9 | -7.3 | -10.4 |
| pre-TS | 4.05 | 2.25 | -6.6 | 0.4 | -7.0 | 10.5 | -6.5 | -11.0 |
| TC | 2.60 | 2.62 | -10.9 | 15.5 | -26.4 | 103.8 | -58.4 | -71.7 |

continued overleaf

Table 5.2: Selected interatomic distances (in Å), activation strain analyses and energy decomposition analyses (electronic energies in gas phase, in kcal mol⁻¹ for all model reactions).^{a,b} (continued)

| | $d_{XX'}$ | $d_{X'X''}$ | ΔE | ΔE_{strain} | ΔE_{int} | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} |
|------------------------------|-----------|-------------|------------|----------------------------|-------------------------|---------------------------|----------------------------|------------------------|
| post-TS | 2.25 | 4.05 | -6.6 | 70.4 | -77.0 | 247.0 | -142.2 | -181.7 |
| PC | 2.20 | 5.07 | -8.1 | 68.8 | -76.9 | 247.6 | -141.5 | -182.9 |
| Se⁻ + STe | | | | | | | | |
| RC | 4.93 | 2.47 | -8.9 | 1.4 | -10.3 | 9.8 | -7.4 | -12.7 |
| pre-TS | 4.32 | 2.18 | -7.1 | 0.6 | -7.8 | 7.1 | -4.5 | -10.5 |
| TC | 2.55 | 2.91 | -12.1 | 20.4 | -32.5 | 123.6 | -67.6 | -88.5 |
| Se⁻ + SeS | | | | | | | | |
| RC | 5.11 | 2.24 | -8.0 | 0.5 | -8.5 | 10.4 | -8.9 | -10.1 |
| pre-TS | 4.56 | 2.24 | -7.9 | 0.1 | -8.0 | 9.7 | -8.4 | -9.3 |
| TC | 2.72 | 2.54 | -16.8 | 12.5 | -29.4 | 99.0 | -65.3 | -63.1 |
| post-TS | 2.38 | 4.46 | -5.9 | 71.2 | -77.2 | 221.7 | -143.7 | -155.3 |
| PC | 2.42 | 4.83 | -6.8 | 67.3 | -74.1 | 204.7 | -133.4 | -145.4 |
| Se⁻ + SeSe | | | | | | | | |
| RC | 5.22 | 2.41 | -9.1 | 0.9 | -9.9 | 7.9 | -6.9 | -10.9 |
| pre-TS | 4.62 | 2.39 | -8.4 | 0.2 | -8.6 | 8.5 | -7.3 | -9.8 |
| TC | 2.70 | 2.71 | -18.3 | 12.2 | -30.5 | 103.4 | -66.8 | -67.0 |
| Se⁻ + TeS | | | | | | | | |
| TC | 2.85 | 2.71 | -24.8 | 10.2 | -35.1 | 99.5 | -73.4 | -61.2 |
| Se⁻ + TeTe | | | | | | | | |
| TC | 2.83 | 3.08 | -25.2 | 10.6 | -35.8 | 108.8 | -76.9 | -67.6 |
| Te⁻ + SS | | | | | | | | |
| RC | 5.45 | 2.06 | -5.5 | 0.4 | -5.8 | 5.5 | -5.7 | -5.6 |
| pre-TS | 4.20 | 2.09 | -3.9 | 0.6 | -4.5 | 9.3 | -5.7 | -8.1 |
| TC | 3.11 | 2.28 | -4.4 | 8.0 | -12.4 | 51.4 | -28.8 | -35.1 |
| Te⁻ + TeS | | | | | | | | |
| TC | 3.09 | 2.69 | -20.6 | 9.0 | -29.6 | 90.4 | -66.2 | -53.8 |

continued overleaf

Table 5.2: Selected interatomic distances (in Å), activation strain analyses and energy decomposition analyses (electronic energies in gas phase, in kcal mol⁻¹ for all model reactions).^{a,b} (continued)

| | $d_{XX'}$ | $d_{X'X''}$ | ΔE | ΔE_{strain} | ΔE_{int} | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} |
|------------------------------|-----------|-------------|------------|----------------------------|-------------------------|---------------------------|----------------------------|------------------------|
| Te⁻ + TeSe | | | | | | | | |
| TC | 3.08 | 2.83 | -21.2 | 8.4 | -29.6 | 91.4 | -65.6 | -55.3 |
| Te⁻ + TeTe | | | | | | | | |
| TC | 3.07 | 3.06 | -21.2 | 8.3 | -30.5 | 94.6 | -66.4 | -58.6 |

a) For reactions with a true single-well energy profile, only TC is reported.

b) Computed at ZORA-OLYP/TZ2P. c) Values in italics refer to geometries with two adjacent coplanar methyl groups; if not differently specified, the methyl groups in plane belong to the nucleophile and to the substrate, supposing that the skewed conformation of the substrate is maintained.

5.3.2 Trends in gas-phase reactivity

Although different kinds of energy profiles (single, double or triple-well) are found, general considerations on the behavior of these molecules can still be drawn and trends can be observed within the set of our model reactions. They will be discussed in this subsection.

A variation in the chalcogenide of the nucleophile causes a destabilization of the stationary points as we move from **S⁻** to **Se⁻** and **Te⁻**. Taking as an example the attack at a disulfide (**X⁻ + SS**) this tendency is clearly seen in RCs, which have a relative energy of -8.5, -7.1 and -5.5 kcal mol⁻¹ for the three chalcogenides respectively, pre-TSs, that are found at -7.2 kcal mol⁻¹ in the case of **S⁻** at -5.8 kcal mol⁻¹ in the case of **Se⁻** and -3.9 kcal mol⁻¹ in the case of **Te⁻**, and TCs, which go from -10.5 to -8.8 and -4.4 kcal mol⁻¹ when **X⁻ = S⁻, Se⁻ or Te⁻**. Activation strain analyses showed that the main contribution to ΔE_{int} in the TCs

originates from a mixture of the lone pair orbital on \mathbf{X}^- and the $\sigma_{\mathbf{X}'\mathbf{X}''}^*$ antibonding orbital on the substrate. The charge donation from the nucleophile to a σ^* orbital has a fundamental role in all these processes. The variation of the nucleophile down in the periodic table group results in a weakening of this interaction due to the donating orbitals becoming lower in energy (1.9, 1.7 and 1.3 eV for \mathbf{S}^- , \mathbf{Se}^- and \mathbf{Te}^- respectively). [190, 192] At the same time, ΔV_{elstat} becomes smaller because the negative charge becomes more diffuse on the bigger chalcogens. From the activation strain analyses, it can be concluded that the decreased stability is mainly caused by a weaker interaction term, although a less destabilizing ΔE_{strain} is computed. The latter originates primarily from the stretching of the S-S bond in the SS substrate, which amounts to 0.44, 0.35 and 0.24 Å along this series of model reactions.

A comparison of a set of three reactions involving the attack of \mathbf{S}^- at **SS**, **SeS** and **TeS** shows that the stationary points become lower in energy as the atom that receives the nucleophilic attack in the substrate becomes more electropositive. For RC and pre-TS, in which the nucleophile is still well separated from the substrate, the energy differences are modest. The RCs have relative energies of -8.5, -10.0 and -11.5 kcal mol⁻¹, respectively, and the pre-TSs are at -7.2, -9.3 and -11.0 kcal mol⁻¹, respectively. For the TC, in which the fragments are strongly interacting, the trend is more evident; the relative energies decrease in steps of almost 10 kcal mol⁻¹ from -10.5 to -19.5 and -27.8 kcal mol⁻¹ when the substrate is varied from **SS** to **SeS** and **TeS**. An analogous effect is found also when \mathbf{Se}^- is the nucleophile, although of a lesser magnitude. This nicely confirms that the validity of the principles observed for S_N2 reactions at carbon and heavier group 14 atoms, as well as for hydrogen- and halogen-bonded complexes, are rather general. [189, 193]

Finally the effect of the change of the leaving group is addressed. As an example the reaction of \mathbf{S}^- with either **SS**, **SSe** or **STe** will be considered. Again,

similar conclusions could be drawn if Se^- is considered as nucleophile. Computed energies (Table 5.2) show how the stationary points along the reaction become slightly more stable as the leaving group contains a heavier chalcogen. This trend is most evident in the TCs which have a relative energy of -10.5 , -13.5 and -14.8 kcal mol $^{-1}$. Decomposition of the total energy of these structures into strain and interaction terms indicates that this increased stabilization comes from a stronger ΔE_{int} . On the basis of extensive activation strain analyses in this and others works, [190, 193] it is fairly safe to say that the increased stability is a consequence of the weaker sulfur-chalcogen bond found in the three different substrates. As X'' becomes heavier, lower bond dissociation energies are computed for the substrates (60.0, 54.3 and 51.1 kcal mol $^{-1}$ for **SS**, **SSe** and **STe** respectively¹). This allows for a further stretch in the $X'-X''$ bond that is accompanied by a less destabilizing ΔE_{strain} . In turn, the higher distance with the leaving group grants the central atom the possibility to build a stronger interaction with the nucleophile in the formation of the TC, which is computed to be of -27.9 , -35.8 and -45.0 kcal mol $^{-1}$ for **SS**, **SSe** or **STe** respectively (Table 5.2).

5.3.3 Effects of solvation

The energy profiles of the reactions in Table 5.3 were studied in the presence of a polar solvent, *i.e.* water, described as a continuum dielectric using the COSMO model. [102] In reactions where charged species are present, the addition of a solvation medium stabilizes more the separated reactants than the formed complexes because in the former the charge is more localized. [194] The overall effect is that the central region around the TC is destabilized, with respect to the free reactants. [195] Results of the computations on the model reactions confirm this; if in the gas phase the transition complexes were found to be strongly

¹Pure electronic bond dissociation energies at ZORA-OLYP/TZ2P.

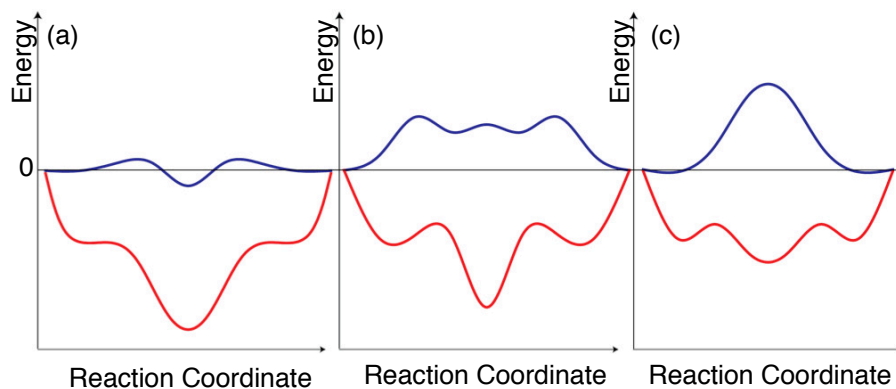


Figure 5.4: Solvent effect on reaction energy profiles (relative to free reactants) in the gas-phase (blue line) and in water (red line) for nucleophilic attack at tellurium (a), selenium (b), and sulfur (c).

stabilized with respect to the reactants, this is no longer true when solvation is taken into account.

For reactions that in the gas phase showed deep potential wells, such as $\text{S}^- + \text{TeTe}$, the TCs remain stable structures but pass from $-28.2 \text{ kcal mol}^{-1}$ of the gas phase to $0.6 \text{ kcal mol}^{-1}$ in the continuum medium, denoting how the destabilization of the central part of the PES can lead to the appearance of reaction barriers between the central TC and the isolated reactants. A pictorial representation of this situation can be seen in Figure 5.4 (a).

On the other hand, a completely different scenario is found in the case of $\text{S}^- + \text{SS}$ (schematically depicted in Figure 5.4 (c)). The relative destabilization of the polar solvent is strong enough to completely change the energy profile, making the modest gas phase barriers disappear and creating a unique central barrier of $11.4 \text{ kcal mol}^{-1}$. In this way, the mechanism has been shifted from addition-elimination in the gas phase to a concerted $\text{S}_{\text{N}}2$ in solution (Figure 5.5). This change in mechanism, which was also reported for the reaction of a thiol with

Table 5.3: Selected interatomic distances (in Å) and energies relative to free reactants (in kcal mol⁻¹) of stationary points of the model reactions in water.^a

| | $d_{XX'}$ | $d_{X'X''}$ | R | TS ^{SN2} ^b | TC | P |
|------------------------------|-------------|-------------|------------|--------------------------------|------|-------------|
| S⁻ + SS | 2.47 | 2.45 | 0.0 | 11.4 | | 0.0 |
| | <i>2.72</i> | <i>2.30</i> | <i>0.0</i> | <i>10.9</i> | | <i>0.0</i> |
| S⁻ + SSe | 2.47 | 2.60 | 0.0 | 10.5 | | -0.2 |
| | <i>2.31</i> | <i>2.84</i> | <i>0.0</i> | <i>10.0</i> | | <i>-0.2</i> |
| S⁻ + SeS | 2.40 | 2.80 | 0.0 | | 3.7 | 0.0 |
| S⁻ + SeSe | 2.55 | 2.71 | 0.0 | | 4.0 | 1.0 |
| S⁻ + TeTe | 2.79 | 3.00 | 0.0 | | 0.6 | 4.6 |
| S⁻ + STe | 2.44 | 2.85 | 0.0 | 12.2 | | 2.8 |
| | <i>2.31</i> | <i>3.06</i> | <i>0.0</i> | <i>12.0</i> | | 2.8 |
| | 2.43 | 2.87 | 0.0 | 11.8 ^c | | 2.8 |
| S⁻ + TeS | 2.71 | 2.71 | 0.0 | | -2.6 | 0.0 |
| Se⁻ + SS | 2.60 | 2.47 | 0.0 | 10.7 | | 0.2 |
| | <i>2.84</i> | <i>2.31</i> | <i>0.0</i> | <i>10.2^d</i> | | <i>0.2</i> |
| Se⁻ + SSe | 2.60 | 2.63 | 0.0 | 9.7 | | 0.0 |
| | <i>2.45</i> | <i>2.87</i> | <i>0.0</i> | <i>9.4</i> | | <i>0.0</i> |
| Se⁻ + SeS | 2.71 | 2.55 | 0.0 | | 3.1 | -1.0 |
| Se⁻ + SeSe | 2.53 | 2.96 | 0.0 | | 3.1 | 0.0 |
| Se⁻ + TeTe | 2.91 | 3.01 | 0.0 | | -0.1 | 3.2 |
| Se⁻ + STe | 2.59 | 2.85 | 0.0 | 11.2 | | 3.0 |
| | <i>2.71</i> | <i>2.74</i> | <i>0.0</i> | <i>11.1</i> | | <i>3.0</i> |
| | 2.44 | 3.12 | 0.0 | 10.8 ^e | | 3.0 |
| Se⁻ + TeS | 2.83 | 2.73 | 0.0 | | -3.1 | -1.4 |
| Te⁻ + SS | 2.85 | 2.44 | 0.0 | 9.4 | | -2.8 |
| | <i>3.06</i> | <i>2.31</i> | <i>0.0</i> | <i>9.2^f</i> | | <i>-2.8</i> |
| | 2.87 | 2.43 | 0.0 | 9.0 ^c | | -2.8 |
| Te⁻ + TeS | 3.00 | 2.79 | 0.0 | | -4.1 | -4.6 |
| Te⁻ + TeSe | 3.01 | 2.91 | 0.0 | | -3.3 | -3.2 |
| Te⁻ + TeTe | 3.07 | 3.07 | 0.0 | | -1.2 | 0.0 |

a) Computed at COSMO-ZORA-OLYP/TZ2P.

b) Values in italics refer to geometries with two adjacent coplanar methyl groups; if not differently specified, the methyl groups in plane belong to the nucleophile and to the substrate, supposing that the skewed conformation of the substrate is maintained.

c) This TS^{SN2} has all methyl groups in the same plane.

d) This TS^{SN2} has the conformation of the TS^{SN2} of **S⁻ + SSe**.

e) In this TS^{SN2}, the methyl groups bonded to S and Te are coplanar.

f) This TS has the conformation of the TS^{SN2} of **S⁻ + STe**.

a disulfide, [196–198] occurs also for the reactions $\text{S}^- + \text{SSe}$, $\text{S}^- + \text{STe}$, $\text{Se}^- + \text{SS}$, $\text{Se}^- + \text{SSe}$ and $\text{Se}^- + \text{STe}$. Some authors investigated the effect of microsolvation, *i.e.* solvation with a number of explicit water molecules, or bulk solvation on the mechanism on nucleophilic substitution at sulfur in disulfides and found that the mechanism depends on the substituent of the atom under attack. [197, 199, 200] These studies report that when the substituent is small (hydrogen) the mechanism is always addition-elimination, in gas phase as well as in solution. When a methyl group is bonded to the atom under attack, the mechanism is sensitive to the computational approach and was always found to be addition-elimination, in gas phase, also with addition of the microsolvation solvent molecules, and $\text{S}_{\text{N}}2$ in bulk solvent, with and without the explicit microsolvated structures. The results reported in Table 5.3, obtained with the COSMO model, nicely agree with these findings since attack occurs at the central sulfur bearing in all cases a methyl.

In all the concerted $\text{S}_{\text{N}}2$ TSs ($\text{TS}^{\text{S}_{\text{N}}2}$), the methyl groups are skewed. As for the gas phase, also in this case there is the possibility to have structures with coplanar methyls (Table 5.3). They are slightly more stable than the skewed counterparts and for $\text{S}^- + \text{STe}$ a transition state with all three methyl groups coplanar is the most stable.

In addition to the two extreme situations portrayed in Figure 5.4 (a) and (c) there is a number of reactions for which the presence of the solvent leads to results that are less clear. Inclusion of the continuum medium creates a rather broad transition plateau in the central PES region. On this plateau, probably caused by the flexibility of the chalcogen-chalcogen bond with respect to stretching, [160, 172] various local minima and transition states might be present. For example, if the reaction $\text{Te}^- + \text{TeTe}$ is taken as reference for the addition-elimination mechanism, inspection of the different structures show how the Te–Te bond lengths in the TC are equal in gas phase and in solvent, whereas for a reaction like $\text{Se}^- + \text{SeSe}$, which shows Se–Se bonds of equal length in the gas phase TC, the optimized minima in

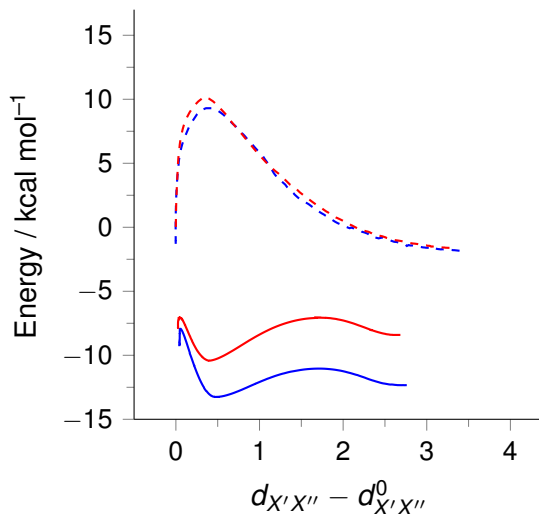


Figure 5.5: Gas-phase energy profiles for the reactions $\text{S}^- + \text{SS}$ (red) and $\text{S}^- + \text{SSe}$ (blue) and energy profiles in water for the reactions $\text{S}^- + \text{SS}$ (red dash) and $\text{S}^- + \text{SSe}$ (blue dash); $d_{X'X''}^0$ refers to the interchalcogen distance in the isolated dichalcogenide.

solvent display quite different Se–Se distances (2.53 and 2.96 Å). This suggests that at least two stable, symmetric minima must exist, which are separated by a TS with equal Se–Se bond lengths. Study of all the structures appearing on this plateau is computationally challenging and it was considered not useful to provide novel insight on the topic. However, the important result of this investigation is the recognition of the mechanism sketched in Figure 5.4 which represents a transitional regime that links the addition-elimination and the $\text{S}_{\text{N}}2$ mechanisms.

5.3.4 Consequences for undesirable thiol exchange reactions

In this last subsection, the focus is shifted to the reactions that have a biological importance, *i.e.* they resemble the last reductive step in the mechanism of GPx4. Namely, they are $\text{S}^- + \text{SSe}$ which is the reaction naturally occurring that leads to the regeneration of the enzyme and $\text{S}^- + \text{SeS}$ which represents instead an unwanted scramble reaction. From the results previously obtained, $\text{S}^- + \text{SeS}$ results energetically favored, having lower transition states and arriving at a more stabilized TC (-19.5 vs -13.5 kcal mol $^{-1}$). Since in the formation of these intermediates the same nucleophile is involved and the bond that has to be broken is always a S–Se bond, the difference in stability can be ascribed to the stronger interaction of the nucleophile with the more electropositive selenium center. Conversely, as the reaction proceeds from TC to PC and then final products $\text{S}^- + \text{SSe}$ is seen to be thermodynamically favored as it has to overcome a lower barrier (post-TSs are at -11.3 vs -9.3 kcal mol $^{-1}$) and results in the formation of a stabler PC (-12.6 vs -10.0 kcal mol $^{-1}$) and stabler products (-5.5 vs 0.0 kcal mol $^{-1}$) (in Figure 5.5 the profile for $\text{S}^- + \text{SSe}$ is shown in blue). When solvent is added, $\text{S}^- + \text{SSe}$ proceeds via a S_N2 mechanism with a barrier of 10.5 kcal mol $^{-1}$ and the resulting products are found at -0.2 kcal mol $^{-1}$, whereas for the unwanted scrambling reaction a minimum on the PES was located at 3.7 kcal mol $^{-1}$. To compare these results obtained for small models with the mechanistic features of GPx some aspects have to be taken into account: (i) there is no conclusive evidence that scrambling occurs in the enzyme and if this is the case it is of minor importance and (ii) a three-center transition state was never observed in the catalytic pocket, *i.e.* the last reductive step does not occur via an addition-elimination mechanism. Therefore, in the G intermediate (Figure 5.1), selenium is protected by the surrounding residues that prevent the attack of the second molecule of GSH by precluding the formation of a bulky TC with selenium as a

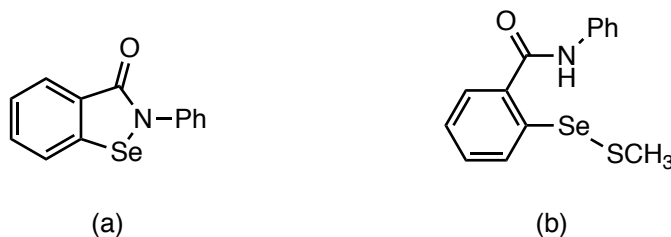


Figure 5.6: Molecular structures of (a) ebselen and (b) selenenyl sulfide as it appears in the ebselen catalytic cycle.

central atom.

Furthermore, a brief exploration of the well-known GPx mimic ebselen (Figure 5.6 (a)) was conducted. We tried to find three-center intermediates, similar to those of the model reactions, using ebselen as a model substrate. In this case, the methyl group bonded to selenium has been extended to resemble the selenenyl sulfide as it appears after ring opening (Figure 5.6 (b)). In the catalytic cycle of Figure 5.1 this corresponds to the last reductive step. Results in gas phase show that no stable intermediate is found for the attack of the nucleophile at the sulfur side, in agreement with previous results, [176] whereas the reaction at the selenium center leads to a highly stabilized minimum at $-30.9 \text{ kcal mol}^{-1}$. Optimization of this structure in water leads to another stable geometry which resembles more a reactant complex, since S–Se bonds measure 2.34 and 2.96 Å, respectively. A favorable hydrogen bond between the NH moiety and the nucleophile seems to further stabilize this molecule. However, this minimum is slightly destabilized with respect to the free reactants, with an energy of $2.8 \text{ kcal mol}^{-1}$, suggesting that in solvent this reaction is shifted towards an $\text{S}_{\text{N}}2$ mechanism. Therefore, also in ebselen the unwanted scrambling reaction is not prevented and only the presence of an enzymatic pocket can hamper this process.

5.4 Conclusions

The study here presented discloses that in gas phase all the model reaction proceed via an addition-elimination mechanism characterized by a triple-well energy profile. Exceptions are found when the substrate is a ditelluride or when the nucleophilic attack of a methyl tellenolate occurs at a tellurium atom of the substrate. In these cases a single-well PES is found for the reactions in the gas-phase. If a continuum polar solvent, such as water, is added three different situations arise, caused by the destabilization of the central region of the energy profile. An S_N2 mechanism is computed when the attack of a methyl chalcogenolate occurs at a sulfur atom. On the other hand, a single-well energy profile without any barrier is found for attack at a tellurium atom. For the attack at a selenium atom, our calculation predict a transitional regime characterized by an almost flat central plateau at positive energy values, where at least two minima and a transition state are located. However the very small energy differences between the structures on this region of the PES make the exploration of this area very lengthy and difficult.

More importantly, some reactions were designed to replicate the last step of the enzymatic cycle in Se- ($S^- + SSe$) and S-containing GPx ($S^- + SS$). Only through the addition of the solvent, a reasonable agreement with the experimental mechanism is achieved. In fact, in gas phase this two reactions proceed via an addition-elimination reaction that involves the formation of a stable three-center TC that was never observed experimentally in the enzyme. Calculations in the condensed phase show a change of the mechanism to S_N2 which is in agreement with the proposed mechanism for GPx. Comparing reaction barriers for the two reactions shows how $S^- + SSe$ has a lower activation energy than $S^- + SS$ making selenium more advantageous in the last reductive step.

Analysis on the $S^- + SeS$ reaction shows how its mechanism remains addition-elimination also with the addition of the solvent. Moreover, this reaction, which

represents the unwanted thiol scrambling, remains always thermodynamically favored over the desired substitution reaction, $\mathbf{S}^- + \mathbf{SSe}$. This is a clue that in the enzymatic pocket there must exist specific intermolecular interactions that prevent the attack at selenium in the last reductive step of the catalytic cycle.

Concerning the GPx mimics, we have found that, in the mechanism of the well-known ebselen, a three-center intermediate forms neither in gas phase nor in water. This is in perfect agreement with previous studies dedicated to the antioxidant mechanism of ebselen.

Finally reactions $\mathbf{S}^- + \mathbf{STe}$ and $\mathbf{S}^- + \mathbf{TeS}$ can be used to model the last step and the unwanted scrambling process of a semi-natural Te-containing GPx catalytic cycle. In water, $\mathbf{S}^- + \mathbf{STe}$ proceeds via an S_N2 mechanism whereas $\mathbf{S}^- + \mathbf{TeS}$ proceeds through the formation and dissociation of a slightly stabilized TC. Interestingly, comparing $\mathbf{S}^- + \mathbf{STe}$ with the corresponding model reaction of Se-containing GPx ($\mathbf{S}^- + \mathbf{SSe}$), reveals a higher barrier for the former, which is not a promising result for the design of Te-based antioxidants. However, a rigorous investigation of semi-natural enzymes and mimics based on tellurium is necessary, considering also more complex systems, and possibly the whole peroxidase mechanism.

Appendices

Appendix E

Torsional barriers

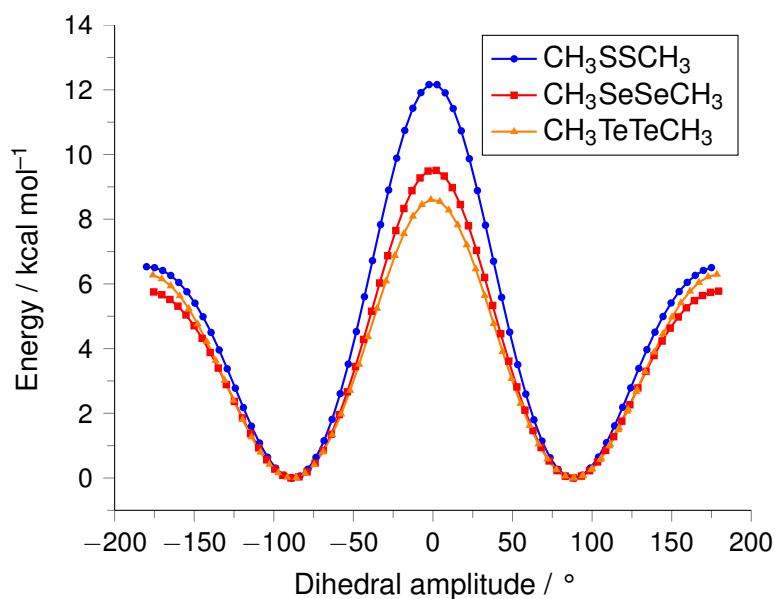


Figure E.1: Dihedral scans of CH₃SSCH₃, CH₃SeSeCH₃ and CH₃TeTeCH₃ *in vacuo* at the scalar ZORA-OLYP/TZ2P level of theory; 72 points and a 5° step were used for each scan. The energies are relative to the global minimum for each molecule (skewed conformation).

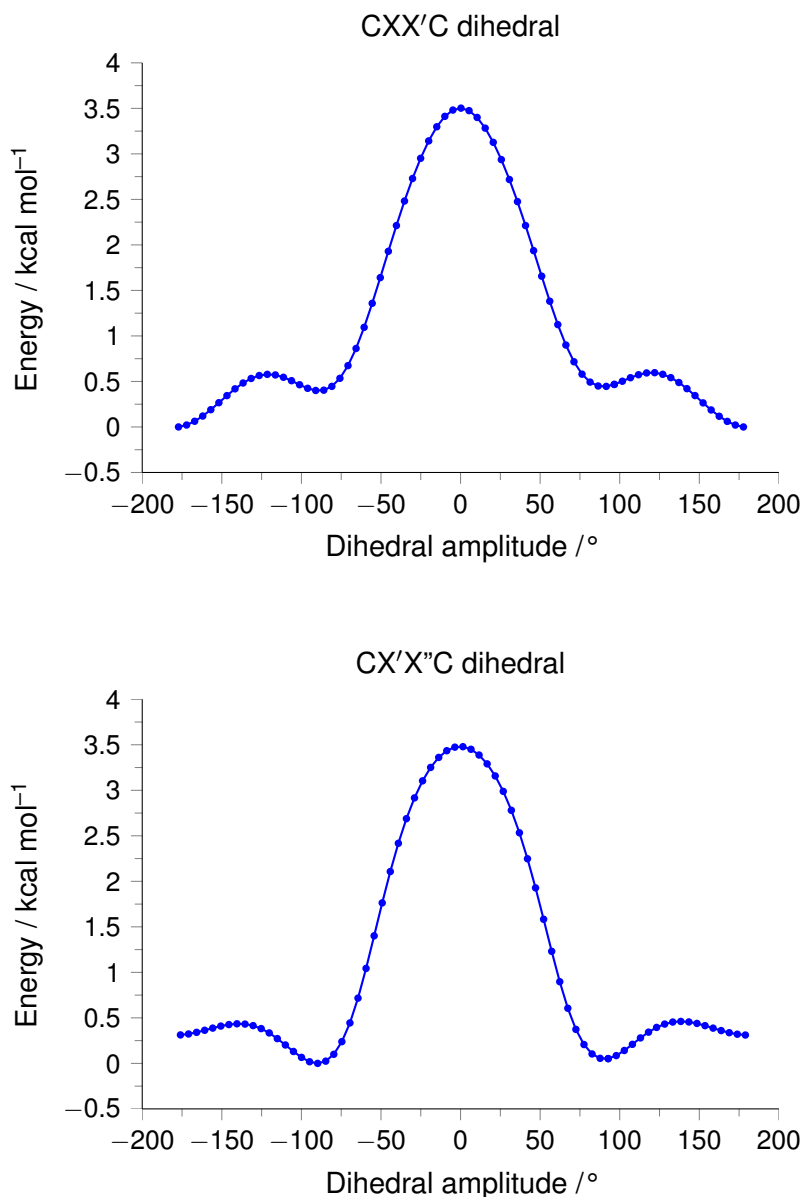


Figure E.2: Dihedral scan of the TC of the $S^- + SS$ reaction *in vacuo* at the scalar ZORA-OLYP/TZ2P level of theory; 72 points and a 5° step were used for each scan. The energies are relative to the global minimum. Both torsions involving the two S-S bonds were scanned: CXX'C (top) and CX'X''C (bottom)

Chapter 6

Oxidation of organoselenides mimics of GPx by H₂O₂

Adapted from

Bortoli, M.; Zaccaria, F.; Dalla Tiezza, M.; Bruschi, M.; Fonseca Guerra, C.; Bickelhaupt, F.M.; Orian, L.

Oxidation of organic diselenides and ditellurides by H₂O₂ for bioinspired catalyst design

Physical Chemistry Chemical Physics, **2018**, *20*, 20874-20885

6.1 Introduction

THE design of efficient antioxidant GPx-like molecular mimics is possible only through the investigation of key elementary biological reactions involving chalcogen centers. Moreover, insights from the study of these reactions can help

better understand the catalytic activity of the enzyme and can also be used in organic chemistry where Se-based compounds are actively used as catalysts.

The brief overview of the properties of organoselenides and tellurides of Section 1.4 outlines a somewhat dissatisfying picture, because the employment of diselenides and ditellurides in pharmacology and medicine is very limited. Therefore systematic studies on the behavior of these molecules and their catalytic activity, which are still scarce in literature, could greatly help improving our knowledge of these systems. To this goal, state-of-the-art computational methodologies provide an optimal method to study their reactivity toward oxidation. Density functional theory (DFT) based quantum mechanical (QM) calculations were seen to be suitable to accurately describe the energetics of organochalcogens [172] and have been used in multiple instances to study these molecules. [148, 149, 201, 202] This Chapter presents an investigation on the mechanism and energetics of the oxidation of model diselenides and tellurides. These compounds show a high antioxidant activity making them very promising GPx mimics (Figure 6.1). Computational study of the oxidation of model dichalcogenides of general formula RXXR (R=H, CH₃, Ph; X=S, Se, Te) by H₂O₂ is carried out through the application of the activation strain model along the reaction coordinate; disulfides are added for completeness.

6.2 Methods

THE REACTION of H₂O₂ with nine different diselenides and ditellurides of general formula RXXR has been investigated *in silico*; the structures and the naming scheme of the chosen compounds can be seen in Table 6.1. For the simplest compounds (i.e. those with R=H, CH₃ and Ph) the sulfur variants were also investigated for completeness.

The OLYP [132–134, 185] functional was employed in conjunction with the

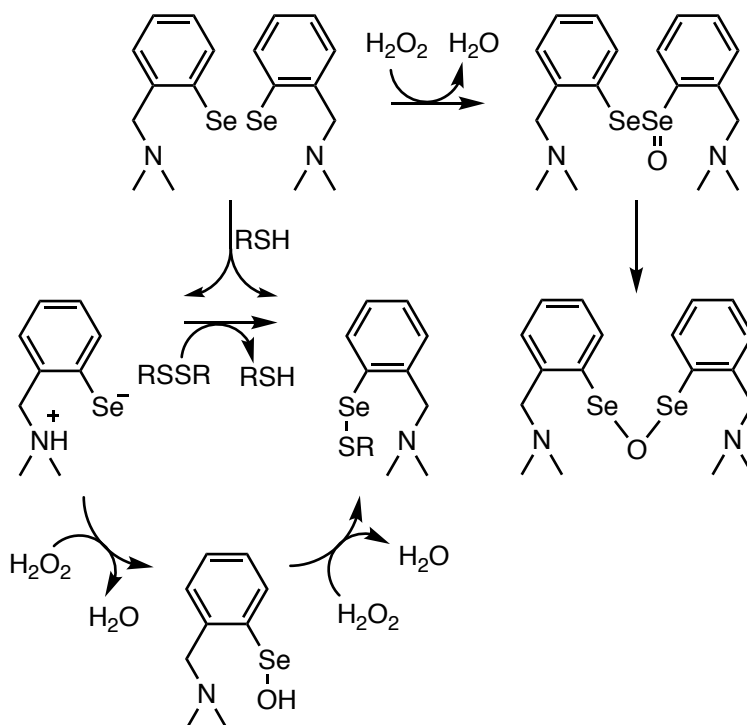
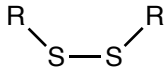
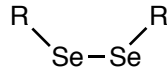
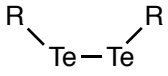


Figure 6.1: Oxidation of 2-(N,N-(dimethylamino)-methyl)benzenediselenide by H_2O_2 followed by product isomerization to anhydride; the reaction mechanism with a thiol (RSH), which is energetically favored, is also shown.

Table 6.1: Structure and naming scheme of the studied disulfides, diselenides and ditellurides.

| |  |  |  |
|-----------------------|---|---|--|
| R | S | Se | Te |
| H | (HS)₂ | (HSe)₂ | (HTe)₂ |
| CH ₃ | (CH₃S)₂ | (CH₃Se)₂ | (CH₃Te)₂ |
| Ph | (PhS)₂ | (PhSe)₂ | (PhTe)₂ |
| p-CH ₃ Ph | | (CH₃PhSe)₂ | (CH₃PhTe)₂ |
| p-CH ₃ OPh | | (CH₃OPhSe)₂ | (CH₃OPhTe)₂ |
| p-ClPh | | (ClPhSe)₂ | (ClPhTe)₂ |
| p-NH ₂ Ph | | (NH₂PhSe)₂ | (NH₂PhTe)₂ |
| p-CNPh | | (CNPhSe)₂ | (CNPhTe)₂ |
| p-NO ₂ Ph | | (NO₂PhSe)₂ | (NO₂PhTe)₂ |

TZ2P basis set, which is a large and uncontracted set of Slater-type orbitals (STOs) of triple- ζ quality, augmented with two sets of polarization functions for each atom: 2p and 3d for hydrogen, 3d and 4f in the case of carbon, nitrogen, oxygen and chlorine, 4d and 4f for selenium and 5d and 4f for tellurium. The frozen core approximation was employed: up to 1s for carbon, nitrogen and oxygen, up to 2p for chlorine, up to 3p for selenium and up to 4p for tellurium. Relativistic effects were included in the calculations using the scalar relativistic zeroth-order regular approximation (ZORA). [184] We will refer to this level of theory as ZORA-OLYP/TZ2P. For a selected number of cases, dispersion corrections were added to the calculation. In these computations, the BLYP [132–135] functional in combination with the TZ2P basis set was employed. Dispersion corrections were taken into account with the D3 scheme with inclusion of the Becke Johnson damping (D3(BJ)), developed by Grimme *et al.* [136] This level of

theory is denoted ZORA-BLYP-D3(BJ)/TZ2P. The performance of DFT methods in describing nucleophilic substitutions involving organochalcogen compounds has been assessed in early studies in benchmarks including HF, MP2 and MP4 methods. [200, 203–206] In presence of phenyl groups, there is also the possibility of π -stacking and other weak interactions that critically depend on dispersion forces, but more recent extensive benchmarks show that DFT-D performs very well as compared to highly correlated *ab initio* methods such as CCSD(T). [207, 208] Stationary points were fully optimized and frequency calculations were used to verify the results. For all energy minima, only real frequencies associated with the vibrational normal modes were found. In the case of the transition states, only one imaginary frequency resulted from the computation, corresponding to the normal mode associated with the reaction under investigation. To obtain a better picture regarding the contributions that account for the differences in the energy barriers, activation strain analyses (ASAs) were carried out. In our case, the fragments are the dichalcogenide substrate and H₂O₂. In this scheme, the total bonding energy (E) is decomposed into two contributions ΔE_{strain} and ΔE_{int} .

This latter term can be further decomposed, in the framework of Kohn-Sham molecular orbital theory, into electrostatic attraction (ΔV_{elst}) between the unperturbed fragments, Pauli repulsion (ΔE_{Pauli}), that is the repulsive interaction between occupied orbitals) and orbital interaction (ΔE_{oi}) using a quantitative energy decomposition analysis (EDA).¹

In the case of the oxidation of the smaller compounds (i.e. (HX)₂, (CH₃X)₂ and (PhX)₂), we performed ASAs along the reaction coordinate from the reactant complex (RCox) until the oxidation transition state (TSox) to evaluate the different contributions to the activation barrier. All the geometries of the non-stationary points, employed in ASA and EDA calculations, were obtained from an intrinsic reaction coordinate (IRC) computation, starting from the transition state

¹For a more detailed treatment of the ASA and EDA schemes see Chapter 2 Section 2.1.2.

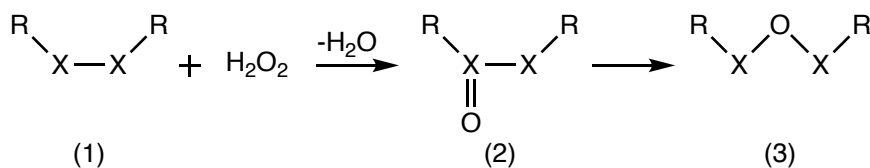


Figure 6.2: Reaction mechanism for the direct oxidation and subsequent possible redox isomerization of diphenyl diselenides and ditellurides; the definitions of R and X are given in Table 6.1

following the path along the normal mode associated with the negative frequency with a steepest descent algorithm. ASA and EDA values were then calculated from the IRC geometries using the program PyFrag. [209] For the remaining compounds, the analysis was restricted to the stationary points (i.e. **RCox** and **TSox**).

6.3 Results and discussion

THE oxidation of model dichalcogenides was investigated at the ZORA-OLYP/TZ2P level of theory according to the reaction depicted in Figure 6.2.

6.3.1 Oxidation of diselenides and ditellurides by H_2O_2

The first part of the discussion will focus on the smaller compounds, *i.e.* **(HX)₂**, **(CH₃X)₂** and **(PhX)₂**. The optimized geometries for the reactants show a high similarity upon varying the different chalcogen and the substituents, with a dihedral angle ψ (Figure 6.3) ranging from 91° to 83° for **(HSe)₂** and **(PhTe)₂** respectively, in agreement with some recent calculations [160, 172] and completely adhering to the bonding mechanism established for chalcogens, [210, 211] which predicts a skewed conformation to be stabler than a structure in which the R

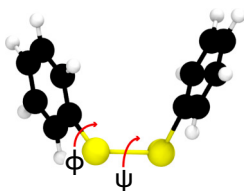


Figure 6.3: $(\text{PhSe})_2$ optimized structure; ϕ and ψ dihedral angles are shown.

group are eclipsed (ψ angle close to 0°) even when the substituents are phenyl groups. Inclusion of an empirical term to account for dispersion forces in the functional shows a stabilization of the structures with stacked R groups that are nevertheless never found to be energy minima since molecules with a ψ angle close to 90° are always found to be the most stable. Therefore, dispersion forces, although having a stabilizing effect on all the structures, do not play a crucial part in defining the mechanism of the observed reactions.

The oxidation starts with the formation of a reactant complex (**RCox**) which has a lower energy than the free reactants (Table 6.2). The subsequent transition state (**TSox**) shows an elongation of the O–O bond in H_2O_2 which takes one of the oxygens closer to the chalcogen and an increase of the ψ dihedrals which measure from 161° to 168° , positioning the two R groups in a trans arrangement with respect to the interchalcogen bond (Figure 6.4). The resulting product complexes (**PCoxs**) are found to lie lower in energy than the starting reactant complexes: they keep the trans geometry in all the cases with a chalcogen-oxygen bond length of 1.49, 1.65 and 1.82 Å in the case of sulfur, selenium and tellurium, respectively and a very small variation (0.01 Å at most) of this bond length among structures with the same chalcogen but different substituents. For $(\text{PhX})_2$ complexes, an alternative pathway was explored, in which both **TSox** and **PCox** display a ψ dihedral smaller than 90° . However, the energies of these new structures were found to be

different from the previous ones by an amount lower than the chemical accuracy of the method employed in the calculation (see Appendix F). Therefore, no further investigation was conducted on this alternative mechanism.

The stabilization of the reactant complexes is modest: it ranges from $-0.5 \text{ kcal mol}^{-1}$ of **(HS)₂** to a maximum of $-2.7 \text{ kcal mol}^{-1}$ in the case of **(CH₃Se)₂**. Moreover, among the molecules with the same chalcogen, this difference additionally thins to become about 1 kcal mol^{-1} (Table 6.2). On the other hand, the transition states and the product complexes display a broader distribution of energies. The former ones are computed to be lower in energy for tellurium, effectively making those compounds the easiest to oxidize, whereas the latter show that the reactions with the disulfides are the most exothermic, having the lowest lying **PCoxs**.

Substituents effects on the energy of the dichalcogenides are limited in most of the optimized structures. **(HX)₂** complexes show the least stable **RCoxs**, but differences fall in a 1 kcal mol^{-1} range. A similar behavior is found also in the case of **TSoxs**, whose energy is only slightly affected by the presence of the different substituents. Consequently, reaction barriers are almost not affected by the change of substituents. A more relevant effect is found in the product complexes, instead: for all three chalcogens **(CH₃X)₂** and **(PhX)₂** show a **PCox** that is approximately 5 kcal mol^{-1} lower in energy than that of **(HX)₂**.

6.3.2 Activation strain analysis for the oxidation of RXXR (R=H, CH₃, Ph) by H₂O₂

The description of the energetics of these processes can explain *how* dichalcogenides react with H₂O₂, but it is not enough to clarify *why* they do so. Thus, ASA and EDA were applied to our model reactions along the reaction path until the formation of **TSox**. The choice to stop at the transition state was dictated by the fact that focus was centered on the formation of the oxidation barrier

Table 6.2: Relative energies (kcal mol⁻¹) of stationary points for the oxidation of RXXR (X=S,Se,Te; R=H, CH₃,Ph) by H₂O₂.^a

| X | | R | | | |
|----|-------------|-------|-------------------------|-----------------|-------|
| | | H | | CH ₃ | Ph |
| S | RCox | -0.5 | <i>-4.4^b</i> | -1.7 | -1.7 |
| | TSox | 25.2 | <i>16.5</i> | 23.8 | 24.1 |
| | PCox | -48.4 | <i>-49.7</i> | -53.4 | -52.9 |
| Se | RCox | -1.5 | <i>-5.9</i> | -2.7 | -2.0 |
| | TSox | 20.2 | <i>10.3</i> | 17.5 | 19.1 |
| | PCox | -36.1 | <i>-40.2</i> | -42.3 | -42.4 |
| Te | RCox | -1.5 | <i>-5.4</i> | -2.3 | -1.9 |
| | TSox | 14.0 | <i>4.0</i> | 12.4 | 12.4 |
| | PCox | -38.5 | <i>-43.9</i> | -43.1 | -44.7 |

a) Computed at ZORA-OLYP/TZ2P.

b) Values in italics are computed at ZORA-BLYP-D3(BJ)/TZ2P.

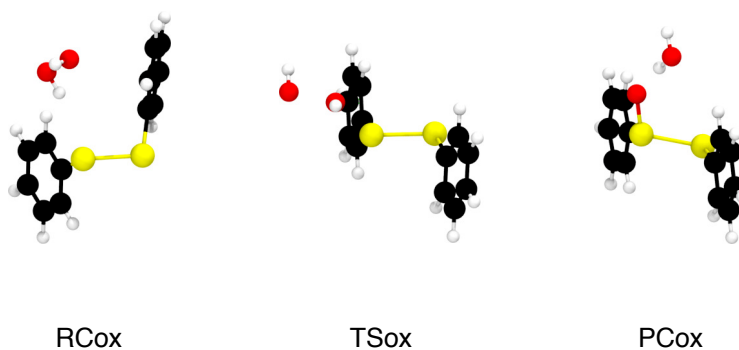


Figure 6.4: Reactant complex (**RCox**), transition state (**TSox**) and product complex (**PCox**) structures for the oxidation of (**PhSe**)₂.

which is not affected by the second half of the reaction. As a suitable reaction coordinate the O–O bond length in H₂O₂ was chosen since it is independent from the chalcogen and at the same time it is a good indicator of the progress of the reaction. IRC calculations were performed starting at **TSox** and arriving at **RCox** and on each intermediate point activation strain analysis was carried out (Figures 6.5 and 6.6). In addition to that for the stationary points (**RCox** and **TSox**) a further quantitative energy decomposition analysis was conducted partitioning the interaction energy into the electrostatic interactions, orbital interactions and Pauli repulsion (see Chapter 2 Section 2.1.2).

6.3.2.1 Effect of the chalcogen

First the effect of the chalcogen atom will be discussed. A preliminary benchmark was conducted on the simplest compounds (**HX**)₂ to evaluate the effects of the inclusion of dispersion contributions to the functional. Calculations at ZORA-BLYP-D3(BJ) level of theory showed a quite homogeneous lowering of the energies of all the structures along the reaction coordinate by about 5 kcal mol⁻¹ when compared to the ones without dispersion. This effect is registered for all the three chalcogens and even if it results in a more favorable oxidation process, it does not alter the trends found with no dispersion contributions as the energy shift is constant in all the cases (Figure 6.5).

Comparing the curves in Figure 6.5 left and Figure 6.6 left, it is evident how the strain profile are similar to each other. This can be ascribed to the fact that most of the deformation occurs in the elongation of the O–O bond which is almost not affected by the presence of the different chalcogen. Phenyl compounds (Figure 6.6 right) show a more complex behavior as the profiles are quite different when we go from (**PhS**)₂ to (**PhSe**)₂ and (**PhTe**)₂, particularly in the late stages of the oxidation. Interestingly the ordering does not follow chalcogen size: (**PhSe**)₂ is

the structure with the highest strain followed by **(PhS)₂** and **(PhTe)₂**.

A peculiar feature is spotted if we look at the ΔE_{int} curves. In fact, they do not decrease monotonously but there is an interval, at the beginning of the reaction, in which they increase. This means that computed interaction energies are positive (*i.e.* destabilizing) which is not usually the case. This feature is present in all the model reactions: it is evident in those without dispersion (Figure 6.5 left and Figure 6.6) and it is only modest (albeit still appreciable) in the cases for which dispersion is included. It can be ascribed to an initial rotation of the H₂O₂ molecule which results in the breaking of a hydrogen bond and the formation of a structure in which an electron rich oxygen points toward an equally electron rich chalcogen atom. This causes a high ΔE_{Pauli} that makes the overall interaction positive until when the two fragments arrive to an orientation and distance where favorable orbital interactions come into play determining a lowering in ΔE_{oi} , which becomes the dominant term, and an overall negative interaction. Moreover, it is clearly seen that for the hydrogen and methyl substituted compounds the total interaction energy decreases (*i.e.* becomes more stabilizing) with the size of the chalcogen. This is mostly due to the different energy of the frontier orbitals of the various molecules. As the chalcogen gets bigger, HOMO energy is seen to increase, translating into a more favorable energy match with the LUMO of H₂O₂. For the phenyl substituted structures the story is a little bit different: an analogous behavior to the hydrogen and methyl substituted compounds is found only at the beginning of the reaction. In fact, an inversion is seen at a very early point in the reaction that results in **(PhSe)₂** gaining a lower ΔE_{int} than **(PhTe)₂** for most of the oxidation.

Although slightly different trends are found in the strain and interaction contributions, their combination into the total energy along the reaction path produces always the same order in energy profiles and activation barriers: ditellurides are the most easily oxidized complexes followed by diselenides and disulfides.

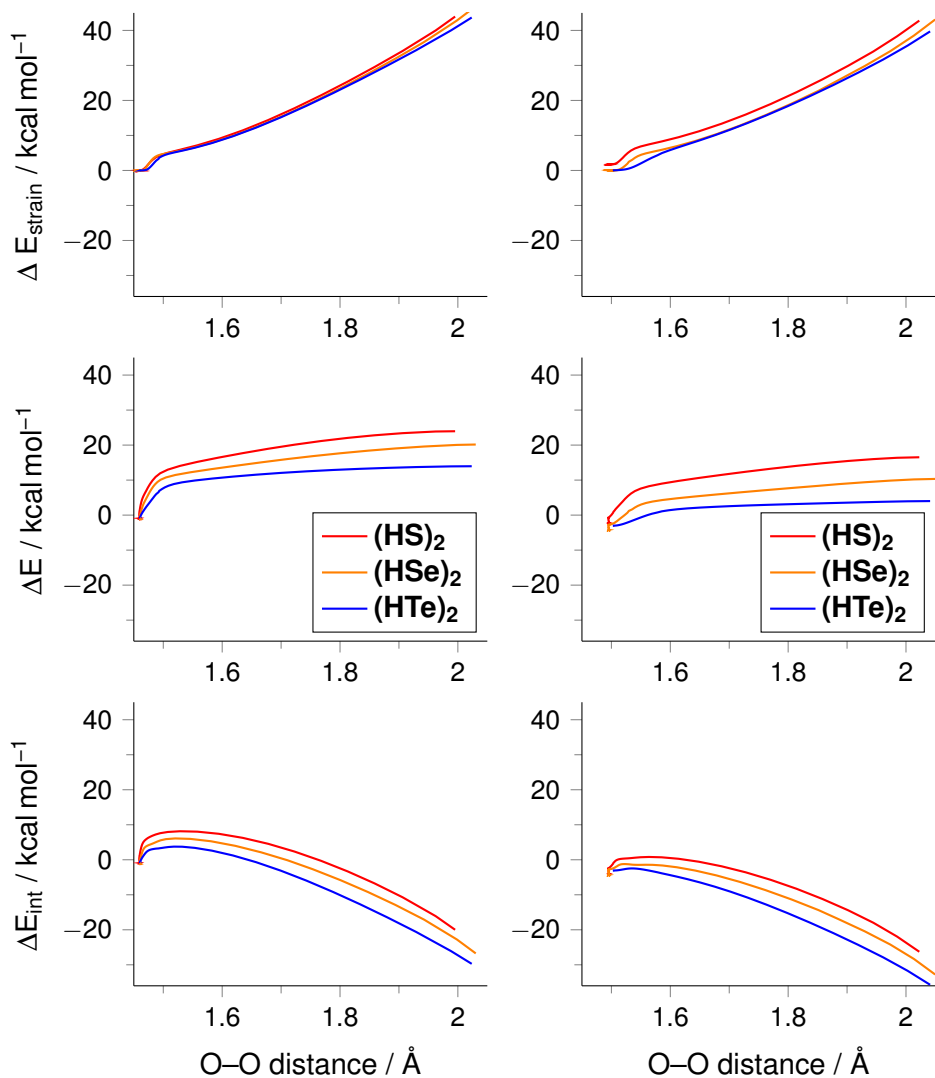


Figure 6.5: Activation strain analysis along the reaction path for the oxidation of $(HX)_2$, without (left) and with (right) empirical correction for dispersion. Pure electronic energies are relative to the free reactants. Computed at ZORA-OLYP/TZ2P (left) and ZORA-BLYP-D3(BJ)/TZ2P (right).

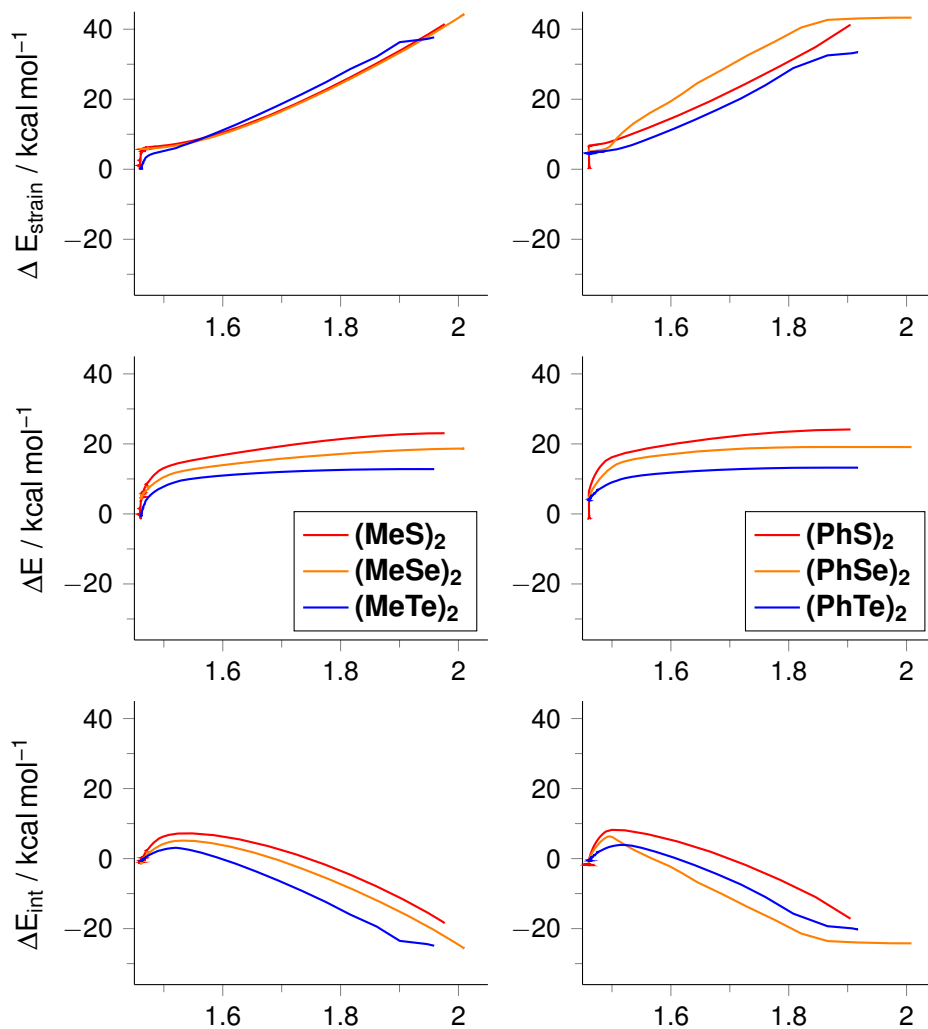


Figure 6.6: Activation strain analysis along the reaction path for the oxidation of $(\text{CH}_3\text{X})_2$ (left) and $(\text{PhX})_2$ (right). Pure electronic energies are relative to the free reactants. Computed at: ZORA-OLYP/TZ2P.

6.3.2.2 Effect of the substituent

The role of the R substituent on the oxidation energetics is found to be more modest than that of the chalcogen. Methyl and phenyl substituted compounds have very close energies (spanning a range of about 1 kcal mol^{-1} , Table 6.2) whereas HXXH complexes display a lower stabilization in **RCoxs** and a higher energy in **TSoxs** making them the most difficult to oxidize. However, the largest difference between two oxidation barriers is of $1.5 \text{ kcal mol}^{-1}$ in the case of **(HSe)₂** and **(CH₃Se)₂**.

Phenyl substituted selenium and tellurium compounds were selected to conduct a more extensive investigation on the role of the substituent and its ability to affect the reaction energetics through the modification of the electronic environment around the chalcogen atoms. Therefore, a broader selection of phenyl substituted diselenides and ditellurides was employed to study how the presence of different *para* moieties on the ring can modify the oxidation barriers (see Table 6.1 for a list of all the *para* groups). Sulfur compounds were not included because the focus of this study centers on the activity of organoselenides and tellurides.

Optimizations of the *para*-phenyl substituted selenides and tellurides show a φ angle close to 90° , with the two phenyls lying in almost parallel planes. If compared to saloon doors, the aromatic rings can be viewed as in an “open” conformation (Figure 6.7, right). A few structures, such as **(CNPhSe)₂** and **(NO₂X)₂**, showed a different geometry with a φ angle close to 0° which can be compared to “closed” saloon doors (Figure 6.7, left). The same mechanistic features of the previously discussed hydrogen and methyl compounds are present also in these new structures which, along the reaction, have geometries similar to those of Figure 6.4.

Computed oxidation barriers for diselenides show an increase in activation energy going from $20.3 \text{ kcal mol}^{-1}$ of **(NH₂PhSe)₂** to $24.1 \text{ kcal mol}^{-1}$ of **(NO₂PhSe)₂**

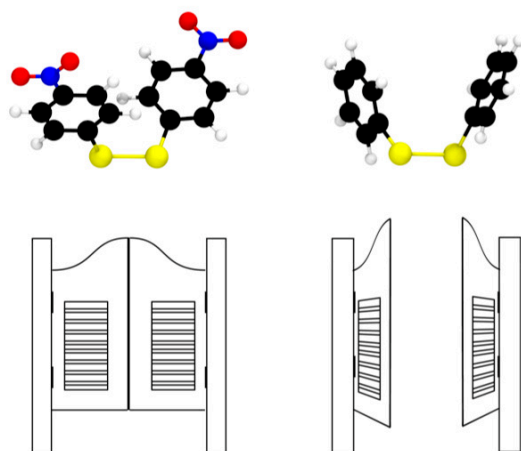


Figure 6.7: Examples of the “closed” (left) and “open” (right) structures found in aryl dichalcogenides.

Table 6.3: Relative energies (kcal mol⁻¹) of the stationary points for the oxidation of substituted diselenides and ditellurides by H₂O₂.^a

| X | R | | | | | | |
|----|--------------------|---------------------|--------------------|-------|-------|--------------------|-------|
| | NH ₂ Ph | CH ₃ OPh | CH ₃ Ph | ClPh | CNPh | NO ₂ Ph | |
| Se | RCox | -2.7 | -2.2 | -2.9 | -1.8 | -1.0 | -1.6 |
| | TSox | 17.6 | 18.1 | 17.8 | 19.9 | 21.7 | 22.5 |
| | PCox | -43.4 | -43.1 | -43.4 | -41.9 | -40.6 | -39.7 |
| Te | RCox | -2.5 | -2.8 | -2.1 | -1.7 | -1.5 | -0.6 |
| | TSox | 11.9 | 12.5 | 12.8 | 13.9 | 15.0 | 15.9 |
| | PCox | -45.6 | -45.5 | -44.9 | -44.2 | -43.5 | -42.5 |

a) Computed at ZORA-OLYP/TZ2P.

(Table 6.3). All the other structures fall between these two extreme values, in a series that follows the electron donor character of the moiety bonded to the phenyl ring with an increasing barrier in the order NH₂<CH₃O<CH₃<Cl<CN<NO₂. A large stabilization is computed for the final **PCoxs** which lie at -40 kcal mol⁻¹ ca., making the process energetically favored. Ditellurides display an analogous behavior with activation energies following the same trend but with values lower by approximately 5 kcal mol⁻¹ (Table 6.3). These results confirm the nucleophilic nature of the chalcogen, since oxidation is favored with electron donating substituents, but at the same time highlight the weak sensitivity of these reactions to a change in the electronic density due to the fact that the highest and lowest barriers differ only by a small amount: 3.8 kcal mol⁻¹ in the case of diselenides and 2.1 kcal mol⁻¹ in the case of ditellurides, respectively.

Activation strain analyses were conducted on **RCoxs** and **TSoxs** (Table 6.4) and the results show a different behavior for the two chalcogens. In the case of the diselenides a slight destabilization of the **RCoxs** with the most electron withdrawing substituents is computed (Table 6.3). Nevertheless, this cannot

overcome the increase of ΔE_{strain} in **TSoxs** from 41.3 kcal mol⁻¹ of **(NH₂PhSe)₂** to 46.7 kcal mol⁻¹ of **(NO₂PhSe)₂** (Table 6.4), which is the main cause behind the higher activation barrier, since the stabilization gain between the two extreme cases is only of 0.5 kcal mol⁻¹. On the other hand, in the case of ditellurides, a more significant variation of ΔE_{int} is found when comparing the different **TSoxs**. However the enhanced interaction cannot overcome the increased strain which determines an overall higher activation barrier for the most electron withdrawing substituents (Table 6.4).

Orbital analysis can help understand the trend found in the oxidation barriers magnitude. Upon inspection of the frontier orbitals, a systematic decrease in the HOMO energy of the diselenide/telluride is found going from **(NH₂PhX)₂** to **(NO₂PhX)₂**. This behavior correlates well with the total activation energy as dichalcogenides with a higher HOMO (and thus a smaller HOMO-LUMO gap) are more reactive towards H₂O₂.

Results from activation strain analysis for the series of substituted aromatic diselenides and ditellurides here considered confirm what found for the simpler aliphatic compounds, *i.e.* that these reactions are not much affected by the changes in electronic density around the chalcogen but, since the chalcogen atom acts as nucleophile, substituents which contribute to increase the electron density around it help to lower the required activation energy for the oxidation process through a destabilization of the HOMO of the dichalcogenide.

6.3.3 Redox isomerization to anhydride

The possible isomerization reaction to a selenenic or tellurenic anhydride (compound **(3)** in Figure 6.2) of the oxidized dichalcogenides (compound **(2)** in Figure 6.2) is investigated. The product of this reaction is a structure which represents an intermediate to highly oxidized compounds which play an important role, at least

Table 6.4: Relative energies and activation strain analysis (electronic energies in gas phase, in kcal mol⁻¹) for the first oxidation of substituted diphenyl diselenides and diphenyl ditellurides by H₂O₂.^a

| R | | Se | | | | Te | | | |
|--------------------|-------------|----------------------------|-------------------------|------------|---------------------|----------------------------|-------------------------|------------|---------------------|
| | | ΔE_{strain} | ΔE_{int} | ΔE | ΔE^\ddagger | ΔE_{strain} | ΔE_{int} | ΔE | ΔE^\ddagger |
| NH ₂ | RCox | 0.1 | -2.8 | -2.7 | | 0.1 | -2.6 | -2.5 | |
| | TSox | 41.3 | -23.7 | 17.6 | 20.4 | 32.3 | -20.4 | 11.9 | 14.4 |
| CH ₃ O | RCox | 0.1 | -2.3 | -2.2 | | 0.2 | -3.0 | -2.8 | |
| | TSox | 41.9 | -23.8 | 18.1 | 20.3 | 30.5 | -18.0 | 12.5 | 15.3 |
| CH ₃ Ph | RCox | -0.7 | -2.2 | -2.9 | | 0.1 | -2.2 | -2.1 | |
| | TSox | 42.0 | -24.2 | 17.8 | 20.7 | 31.4 | -18.4 | 12.8 | 14.9 |
| Cl | RCox | 0.1 | -1.9 | -1.8 | | 0.1 | -1.8 | -1.7 | |
| | TSox | 44.0 | -24.1 | 19.9 | 21.7 | 36.5 | -22.6 | 13.9 | 15.6 |
| CN | RCox | 0.5 | -1.5 | -1.0 | | 0.1 | -1.6 | -1.5 | |
| | TSox | 45.9 | -24.2 | 21.7 | 22.7 | 40.2 | -25.2 | 15.0 | 16.5 |
| NO ₂ | RCox | 0.3 | -1.9 | -1.6 | | 0.9 | -1.3 | -0.6 | |
| | TSox | 46.7 | -24.2 | 22.5 | 24.1 | 41.7 | -25.8 | 15.9 | 16.5 |

a) Computed at ZORA-OLYP/TZ2P.

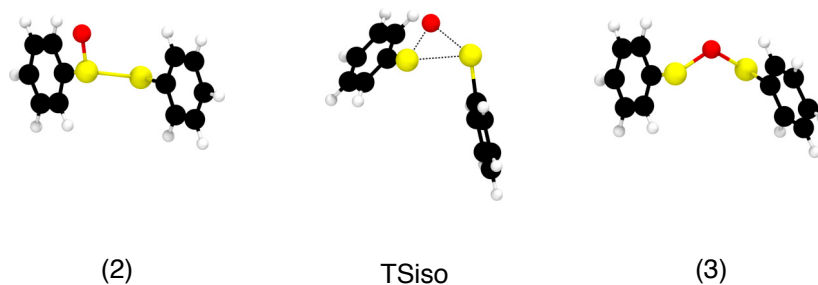


Figure 6.8: Stationary points and transition state structures for the isomerization of $(\text{PhSe})_2$.

in the case of selenium, in the catalysis of organic reactions. [212] The isomerization was studied for diselenides and ditellurides with all the R combinations (also H and CH_3) and the water molecule formed after the first oxidation was not included. For all the dichalcogenides transition states showing similar structures were fully optimized. The reaction begins with a closing of the O-X-X angle that progressively inserts the oxygen atom between the two chalcogens. At the same time, one of the two R groups rotates by almost 90° so that at **TSiso** a ψ dihedral of almost 90° is measured (Figure 6.8). In the last part of the reaction leading to the formation of the final anhydride, the R group rotates back, widening the ψ dihedral again and taking a position very similar to that of (2).

The energetics of this process show that it is always disfavored if compared to the initial oxidation with computed barriers more than 20 kcal mol^{-1} higher (Table 6.5). A broader energy span is calculated for the diselenides going from $41.1 \text{ kcal mol}^{-1}$ of $(\text{NH}_2\text{PhSe})_2$ to $45.5 \text{ kcal mol}^{-1}$ of $(\text{HSe})_2$, whereas the ditellurides have a more limited range that goes from 26.9 to $29.5 \text{ kcal mol}^{-1}$ in the case of $(\text{NO}_2\text{PhTe})_2$ and $(\text{HTe})_2$, respectively. In general, ditellurides show lower activation energies compared to diselenides, in analogy to what found for the first oxidation. Moreover, isomerization is calculated to be always unfavorable since

Table 6.5: Relative energies (with respect to **(1)** + H₂O₂, in kcal mol⁻¹) for the isomerization of the selenoxides and telluroxides.^a

| X | | R | | | | |
|----|--------------|-------|-----------------|--------------------|---------------------|--------------------|
| | | H | CH ₃ | NH ₂ Ph | CH ₃ OPh | CH ₃ Ph |
| Se | (2) | -32.6 | -37.8 | -38.6 | -38.5 | -39.0 |
| | TSiso | 12.9 | 6.6 | 2.5 | 3.2 | 3.1 |
| | (3) | -23.0 | -28.5 | -29.9 | -29.5 | -29.7 |
| Te | (2) | -34.0 | -37.8 | -39.9 | -39.8 | -39.6 |
| | TSiso | -4.5 | -9.1 | -11.9 | -11.6 | -11.3 |
| | (3) | -29.0 | -32.8 | -35.1 | -34.7 | -34.0 |

| | | R | | | |
|----|--------------|-------|-------|-------|--------------------|
| | | Ph | ClPh | CNPh | NO ₂ Ph |
| Se | (2) | -38.1 | -37.7 | -36.6 | -35.8 |
| | TSiso | 4.1 | 4.1 | 4.7 | 5.3 |
| | (3) | -28.7 | -28.7 | -29.0 | -29.0 |
| Te | (2) | -38.6 | -38.4 | -38.4 | -37.5 |
| | TSiso | -11.2 | -11.2 | -11.4 | -10.6 |
| | (3) | -33.7 | -33.8 | -34.4 | -34.3 |

a) Computed at ZORA-OLYP/TZ2P.

the anhydrides are less stable than the initial oxides. Finally, no systematic trend that could correlate the electron withdrawing or electron donating nature of the substituent with the isomerization barrier height is seen: the structures with the highest activation energy are HXXH for both chalcogens, and, even among the phenyl substituted complexes, a behavior similar to that found for the oxidation was not registered. This is a strong indication that this process is not influenced by the electronic environment around the chalcogen center, but might be more sensitive towards steric effects, that might be present, for example in the case of the diphenyl dichalcogenides, when adding a substituent to the ring closer to the chalcogen atom, in *ortho* or *meta* position.

6.4 Conclusions

The results presented in this Chapter about the oxidation of organodichalcogenides by H_2O_2 show an increased reactivity towards hydrogen peroxide mainly caused by a stronger interaction energy calculated in the presence of the heavier chalcogens. HOMO-LUMO interactions with H_2O_2 are favored in selenium and especially in tellurium compounds because of the higher energy of their HOMOs which can more favorably interact with the LUMO of hydrogen peroxide resulting in a lower activation energy. A noteworthy result obtained from ASA along the reaction coordinate is that dimethyl and diphenyl dichalcogenides show a very similar oxidation energy profile. They react in a perfectly analogous manner despite the very different nature of the R group suggesting that in the design of novel antioxidant drugs these two classes of compounds could be interchangeable and the only factor that would go in favor of one or the other could be the toxicity of its metabolites. Therefore, studies on diphenyl diselenides could be prioritized over the analogous methyl compounds which have a higher toxicity. Toxicological studies on ditellurides are still fragmentary, but their enhanced activity with respect to diselenides has prompted a more intense effort to extend the knowledge on the behavior of these compounds.

Finally, the investigation of a series of differently *para*-substituted diphenyl diselenides and ditellurides pointed out that the initial oxidation leads to a selenoxide/telluroxide without any difference in mechanism for all the complexes and that this oxidation is easier for ditellurides and only slightly affected by the nature of the phenyl substituents. Moreover, the subsequent isomerization of the selenoxide/telluroxide to the corresponding selenenic/tellurenic anhydride, while being possible, is found to be thermodynamically unfavored and requires a high activation energy. These findings could explain the difficulty in detecting these species with an experimental setup.

Appendices

Appendix F

Alternative mechanism

Table F.1: Energies (in kcal mol⁻¹) and ψ dihedral amplitude (°) of stationary points for the oxidation of **(PhSe)₂** and **(PhTe)₂** in two proposed mechanisms.^a

| | | <i>cis</i> | | <i>trans</i> | |
|---------------------------|-------------|------------|------------|--------------|------------|
| | | ψ | ΔE | ψ | ΔE |
| (PhSe)₂ | RCox | 82 | -2.0 | 82 | -2.0 |
| | TSox | 70 | 21.0 | -169 | 19.1 |
| | PCox | 51 | -41.1 | -163 | -42.3 |
| (PhTe)₂ | RCox | 83 | -1.9 | 83 | -1.9 |
| | TSox | 72 | 12.4 | -162 | 13.2 |
| | PCox | 41 | -43.6 | -167 | -44.7 |

a) Computed at ZORA-OLYP/TZ2P.

Appendix G

Energy decomposition analysis

Table G.1: Energy decomposition analysis (electronic energies in gas phase, in kcal mol⁻¹) for the first oxidation of model diphenyl disulfides.^a

| R | | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} |
|-----------------|-------------|---------------------------|----------------------------|------------------------|
| H | RCox | 2.4 | -2.4 | -1.7 |
| | TSox | 186.5 | -81.4 | -125.2 |
| CH ₃ | RCox | 4.0 | -3.9 | -2.7 |
| | TSox | 169.5 | -75.4 | -112.6 |
| Ph | RCox | 3.3 | -3.3 | -2.4 |
| | TSox | 165.1 | -72.2 | -110.0 |

a) Computed at ZORA-OLYP/TZ2P.

Table G.2: Energy decomposition analysis (electronic energies in gas phase, in kcal mol⁻¹) for the first oxidation of model diphenyl diselenides.^a

| R | | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} |
|--------------------|-------------|---------------------------|----------------------------|------------------------|
| H | RCox | 2.5 | -2.3 | -1.7 |
| | TSox | 159.1 | -73.1 | -112.7 |
| CH ₃ | RCox | 4.5 | -4.0 | -2.9 |
| | TSox | 151.2 | -70.1 | -106.7 |
| Ph | RCox | 2.9 | -2.8 | -2.2 |
| | TSox | 146.6 | -67.0 | -103.9 |
| NH ₂ | RCox | 3.8 | -3.8 | -2.8 |
| | TSox | 136.0 | -61.9 | -97.9 |
| CH ₃ O | RCox | 3.5 | -3.2 | -2.5 |
| | TSox | 138.8 | -63.1 | -99.4 |
| CH ₃ Ph | RCox | 3.1 | -3.0 | -2.3 |
| | TSox | 143.5 | -65.5 | -102.2 |
| Cl | RCox | 2.9 | -2.6 | -2.1 |
| | TSox | 149.4 | -68.2 | -105.3 |
| CN | RCox | 2.3 | -2.0 | -1.7 |
| | TSox | 156.7 | -71.9 | -109.0 |
| NO ₂ | RCox | 0.5 | -1.9 | -0.5 |
| | TSox | 158.2 | -72.6 | -109.8 |

a) Computed at ZORA-OLYP/TZ2P.

Table G.3: Energy decomposition analysis (electronic energies in gas phase, in kcal mol⁻¹) for the first oxidation of model diphenyl ditellurides.^a

| R | | ΔE_{Pauli} | ΔV_{elstat} | ΔE_{oi} |
|--------------------|-------------|---------------------------|----------------------------|------------------------|
| H | RCox | 2.4 | -2.1 | -1.9 |
| | TSox | 147.6 | -72.0 | -105.3 |
| CH ₃ | RCox | 4.2 | -3.7 | -2.9 |
| | TSox | 128.1 | -62.7 | -89.9 |
| Ph | RCox | 2.9 | -2.6 | -2.3 |
| | TSox | 117.8 | -57.2 | -80.8 |
| NH ₂ | RCox | 3.3 | -3.3 | -2.6 |
| | TSox | 110.4 | -53.2 | -77.6 |
| CH ₃ O | RCox | 4.1 | -4.1 | -2.9 |
| | TSox | 106.4 | -51.5 | -72.9 |
| CH ₃ Ph | RCox | 2.9 | -2.7 | -2.4 |
| | TSox | 110.3 | -53.6 | -75.3 |
| Cl | RCox | 2.8 | -2.4 | -2.2 |
| | TSox | 127.5 | -62.0 | -88.1 |
| CN | RCox | 2.5 | -2.1 | -2.0 |
| | TSox | 139.8 | -68.2 | -96.8 |
| NO ₂ | RCox | 2.5 | -1.9 | -1.9 |
| | TSox | 142.2 | -69.3 | -98.5 |

a) Computed at ZORA-OLYP/TZ2P.

Chapter 7

Conclusions

7.1 Summary

The role of selenium as an antioxidant, in particular as a key component in the enzymatic activity of glutathione peroxidase, was described and analyzed with a computational methodology, employing state-of-the art quantum mechanical techniques combined with classic calculations. Density functional theory methods were the main approach employed to obtain structural, energetic and mechanistic information on model systems. To include the effect of part of the systems that were left out of the QM calculations, classic molecular dynamics simulations were carried out using a recently developed force field tailored to effectively model protein structures. Finally, the application of quantitative models for energy decomposition (activation strain model and energy decomposition analysis) allowed an in-depth analysis of the formation of the reaction barriers and their underlying causes. These *in silico* techniques made the study of the intrinsic properties of selenium and of the other chalcogens possible.

Three scenarios were selected and tested: the ability of chalcogenides to form

weak non-covalent bonds (Chapter 3), their thermodynamics and reactivity in S_N2 substitution processes (Chapter 5) and their reactivity toward H_2O_2 in redox reactions (Chapters 4 and 6).

Non covalent chalcogen bonds have stirred a lot of interest lately for their importance in protein interaction and their possible application to catalysis, ion transport and rational drug design. A series of halogen substituted chalcogenides were chosen as model systems. Three small unsaturated organic molecules were selected as Lewis bases for the formation of the weak interaction. The resulting chalcogen- π bond increases in strength as the chalcogen-halogen electronegativity difference becomes larger. This is due to a combination of electrostatic and orbital effects. The bond acceptor can also modify the strength of the interaction based on the energy of its HOMO: substrates with a high energy HOMO interact more favorably with the chalcogenides resulting in stronger bonds. These considerations can be applied to a broad range of chalcogen- π complexes and will help understand their behavior and direct any rational design attempts towards the most favorable scenario.

Maximizing the efficiency of each step in a catalytic cycle is the keystone of the rational design of any catalyst and while the noncovalent interactions are predominant in the initial formation reactant complexes, the redox reactions in which selenium is involved, when acting as an antioxidant, require the breaking and formation of covalent bonds. The catalytic cycle of human glutathione peroxidase 4 was chosen as a model to determine which factors contribute to the enzymatic activity in order to establish the required features in an efficient artificial GPx mimic and to unravel the fundamental traits of the three chalcogens (S, Se and Te) which could be exploited to rationally design the most befitting molecule for a targeted application. Two steps of the whole GPx4 mechanism were thoroughly analyzed: the first oxidative reaction and the last reductive phase.

The oxidative part was studied with a combined classical/QM method in which the MD simulations served as a starting point for the subsequent DFT calculations. Results on an enzymatic cluster representing the active site show that selenium is thermodynamically favored over sulfur due to the weaker Se–H bond present in the initial structure that can more easily undergo the proton transfer needed to begin the reaction. Moreover, tellurium displays a different mechanism and can act differently from the other chalcogens as the catalytic cycle of the semi-natural tellurium-GPx4 enzyme could pass through highly oxidized tellurium states.

The final reductive step was modeled as a nucleophilic substitution of a methyl chalcogenolate on a dimethyl dichalcogenide. The choice of such a small system allows the focus to be centered on the intrinsic properties of the different chalcogens employed (S, Se and Te) to evaluate their particular behavior. Results in the gas phase show how nucleophilic substitutions have always an addition-elimination mechanism with the formation of a three-center transition complex, but those at tellurium show a single-well profile, whereas those at the other chalcogens have a triple-well profile. The inclusion of the solvent in the calculations as a dielectric continuum, aims to simulate an environment closer to that of the *in vivo* enzyme. In the condensed phase mechanistic features of this reactions depend upon the chalcogen that undergoes nucleophilic attack: in the case of sulfur the reaction is seen to be an S_N2 , in the case of tellurium it retains the single-well profile of the gas-phase, albeit with a less stabilized transition complex, and in the case of selenium a transitional mechanism is proposed which shows an energy profile with an almost flat central region inside which multiple intermediates and transition states are located. This mechanistic difference could be one of the key reasons why selenium is present in the active site of GPx, as the formation of the three-center complex seems to be unlikely in the enzyme due to the steric hindrance caused by the surrounding residues. Otherwise, the

regeneration of the enzyme in the last reductive step would be greatly hampered since the nucleophilic attack at selenium, which would result in an unwanted thiol scrambling reaction, is thermodynamically favored over that at sulfur.

The modeling of enzymatic cycles can offer great insights on the “secrets” of the efficiency and effectiveness of these catalysts. However, transferring of the information obtained for an enzyme to a smaller and easily synthesizable molecules that should reproduce the catalytic activity is never straightforward. A complete understanding of the processes in which these small mimics are involved is key to correctly translate and apply the results obtained for larger systems. Therefore in the final part of this Thesis an in-depth study of the reactivity of a very promising class of GPx mimics towards H_2O_2 is detailed. Organic diselenides are extensively employed in synthesis as catalysts and show promising features to be employed as efficient antioxidant compounds in pharmacology and medicine. Recently, organic tellurides were also studied for their antioxidant properties, although the lack of a complete picture of the toxicology of these compounds limits their employment as therapeutic agents. The oxidation reaction of selected diselenides and ditellurides by hydrogen peroxide was investigated with DFT techniques. For the simplest compounds activation strain analyses were performed along the reaction coordinate from the starting reactant complex to the transition state. Results show an increased reactivity towards H_2O_2 going from sulfur to tellurium leading to the formation of a sulfoxide/selenoxide/telluroxide and a very similar reaction profile for dimethyl and diphenyl chalcogenides. Further investigations on *para* substituted diphenyl diselenides and ditellurides show that, although the ring substituent has a modest effect on the oxidation barrier, electron donating moieties favor the oxidation process, decreasing the activation energy. Finally, the isomerization of the chalcogenoxide obtained after oxidation was investigated in light of the fact that the resulting anhydride seems to be an important intermediate to reach highly oxidized states that play a key role

in the catalysis of redox reactions, at least in the case of diselenides. Computed barriers are always larger than those for the initial oxidation, with ditellurides having smaller isomerization barriers compared to diselenides. Moreover, the resulting anhydrides lie at a higher energy than the starting oxide making the reaction thermodynamically unfavorable.

7.2 Sommar

Il ruolo del selenio come antiossidante, in particolare come componente chiave dell'attività enzimatica della glutatione perossidasi, è stato descritto e analizzato attraverso una metodologia computazionale che utilizza tecniche quantomeccaniche allo stato dell'arte combinate a calcoli di meccanica classica. Il principale approccio utilizzato per ottenere informazioni di tipo strutturale, energetico e meccanicistico sfrutta metodi basati sulla teoria del funzionale di densità (DFT). Per includere l'effetto della parte del sistema che è stata esclusa dai calcoli quantomeccanici, simulazioni di dinamica molecolare classica sono state condotte utilizzando un campo di forza sviluppato di recente e progettato per riprodurre efficacemente le strutture proteiche. Infine, l'applicazione di modelli quantitativi per la decomposizione dell'energia (il modello dell'*activation strain* e l'*energy decomposition analysis*) ha permesso un'analisi approfondita delle cause sottostanti alla formazione delle barriere di reazione. Queste tecniche *in silico* hanno reso possibile lo studio delle proprietà intrinseche del selenio e degli altri calcogeni.

Tre scenari sono stati selezionati ed esaminati: l'abilità dei calcogenuri di formare interazioni deboli di tipo non covalente (Capitolo 3), la loro termodinamica e reattività in sostituzioni S_N2 (Capitolo 5) e la loro reattività verso H_2O_2 in reazioni di ossidoriduzione (Capitoli 4 e 6).

Legami calcogeno di tipo non covalente hanno suscitato un grande interesse negli ultimi tempi per la loro importanza nelle interazioni proteiche e la loro

possibile applicazione nella catalisi, nel trasporto di ioni e nel design razionale di farmaci. Un gruppo di calcogenuri con alogeni come sostituenti sono stati selezionati come sistemi modello. Tre piccole molecole organiche insature sono state scelte come basi di Lewis per la formazione del legame debole. Il legame chalcogeno- π che ne risulta si rafforza con l'aumento della differenza in elettronegatività tra alogeno e chalcogeno. Questo deriva da una combinazione di effetti elettrostatici e orbitalici. Anche l'accettore del legame può modificare la forza del legame attraverso l'energia del suo orbitale occupato a più alta energia (HOMO): substrati con un HOMO ad alta energia hanno un'interazione più favorevole con i calcogenuri e formano perciò legami più forti. Queste considerazioni possono essere applicate ad un ampio spettro di complessi chalcogeno- π e aiuteranno nella comprensione del loro comportamento, direzionando verso lo scenario più favorevole le proposte di design razionale.

Massimizzare l'efficienza di ogni passaggio in un ciclo catalitico è la chiave di volta del design razionale di qualsiasi catalizzatore e, sebbene le interazioni non covalenti siano predominanti nella formazione dei complessi reagente iniziali, le reazioni di ossidoriduzione nelle quali il selenio è coinvolto, nella sua azione di antiossidante, richiedono la formazione e la rottura di legami covalenti. Il ciclo catalitico della glutatione perossidasi 4 umana è stato scelto come modello per determinare quali fattori contribuiscano all'attività enzimatica al fine di stabilire le proprietà necessarie per un efficiente mimetico artificiale della GPx e per far luce sulle caratteristiche intime dei tre calcogeni (S, Se e Te) che potrebbero essere sfruttate, attraverso il design razionale, per ottenere molecole che esprimano la massima efficacia per le applicazioni per le quali sono state progettate. Due fasi dell'intero meccanismo della GPx4 sono state analizzate in dettaglio: la prima reazione di ossidazione e l'ultima fase riduttiva.

La parte ossidativa è stata studiata con un metodo combinato che prevede l'utilizzo di calcoli quantomeccanici e classici in cui le simulazioni di dinamica

molecolare sono servite come punto di partenza per i successivi calcoli DFT. I risultati ottenuti su un cluster enzimatico che rappresenta il sito attivo dell'enzima mostrano che il selenio è termodinamicamente favorito rispetto allo zolfo a causa del più debole legame Se-H presente nella struttura iniziale che rende più facile il trasferimento di un protone necessario per iniziare le reazioni. Inoltre, il tellurio agisce secondo un meccanismo differente e può funzionare in maniera diversa dagli altri calcogeni in quanto il ciclo catalitico dell'enzima semi-naturale tellurio-GPx4 potrebbe passare attraverso alti stati di ossidazione del tellurio.

Il processo riduttivo finale è stato modellato come una sostituzione nucleofila di un metil calcogenolato su un dimetil dicalcogenuro. La scelta di un sistema così piccolo permette di centrare l'attenzione sulle proprietà intrinseche dei differenti calcogeni utilizzati (S, Se, Te) per valutare il loro individuale comportamento. I risultati ottenuti in fase gas mostrano che le sostituzioni nucleofile hanno sempre un meccanismo di tipo addizione-eliminazione che prevede la formazione di un complesso di transizione a tre centri, ma le reazioni al tellurio rivelano un profilo a singolo minimo, mentre quelle che avvengono agli altri calcogeni risultano avere un profilo a tre minimi. L'inclusione del solvente nei calcoli, come un mezzo dielettrico continuo, vuole simulare un ambiente più simile a quello dell'enzima *in vivo*. In fase condensata gli aspetti meccanicistici di queste reazioni dipendono dal calcogeno che subisce l'attacco nucleofilo: nel caso dello zolfo la reazione si vede essere del tipo S_N2 , nel caso del tellurio mantiene il profilo a singolo minimo della fase gas, sebbene con un complesso di transizione meno stabilizzato, e nel caso del selenio viene proposto un meccanismo intermedio che mostra un profilo energetico che presenta una regione centrale quasi piatta all'interno della quale si trovano più intermedi e stati di transizione. Questa differenza nel meccanismo potrebbe essere una delle ragioni principali per cui il selenio è presente nel sito attivo della GPx, poiché la formazione del complesso a tre centri sembra poco probabile nell'enzima a causa dell'ingombro sterico causato

dai residui circostanti. In caso contrario, la rigenerazione dell'enzima nella fase riduttiva finale sarebbe fortemente impedita dal momento che l'attacco nucleofilo al selenio, che risulterebbe in una reazione non voluta di scambio di tioli, è favorito termodinamicamente rispetto a quello allo zolfo.

Modellizzare i cicli enzimatici può offrire la possibilità di comprendere i “segreti” dell'efficienza ed efficacia di questi catalizzatori. Purtroppo, trasferire le informazioni ottenute per un enzima a una piccola molecola facilmente sintetizzabile, che dovrebbe riprodurre l'attività catalitica, non è mai un processo lineare. Una completa comprensione delle reazioni in cui questi piccoli mimetici sono coinvolti è fondamentale per tradurre e applicare correttamente i risultati ottenuti per sistemi più estesi. Perciò, nella parte finale di questa Tesi viene esposto in dettaglio lo studio della reattività verso H_2O_2 di una promettente classe di mimetici della GPx. I diseleniuri organici sono impiegati estensivamente in sintesi come catalizzatori e possiedono delle caratteristiche promettenti per essere utilizzati come efficienti antiossidanti in farmacologia e medicina. Recentemente, anche i ditellururi organici sono stati studiati per le loro proprietà antiossidanti, sebbene la mancanza di un quadro completo sulla loro tossicità limiti il loro impiego come agenti terapeutici. La reazione di ossidazione di alcuni diseleniuri e ditellururi a carico del perossido di idrogeno è stata studiata con tecniche di tipo DFT. Per i composti più semplici le analisi di *activation strain* sono state condotte lungo la coordinata di reazione dal complesso reagente iniziale allo stato di transizione. I risultati mostrano un aumento della reattività verso H_2O_2 passando da zolfo a tellurio che porta alla formazione di un solfossido/selenossido/tellurossido e un profilo di reazione molto simile per i dimetil e i difenil dicalcogenuri. Ulteriori studi su difenil diseleniuri e ditellururi sostituiti in posizione *para* mostrano che, sebbene il sostituente sull'anello abbia un effetto modesto sulla barriera di ossidazione, gruppi elettron donatori favoriscono il processo di ossidazione, diminuendo l'energia di attivazione. Infine, l'isomerizzazione del calcogenossido ottenuto

dopo l'ossidazione è stata studiata dal momento che l'anidride risultante sembra essere un intermedio importante per raggiungere alti stati di ossidazione che giocano un ruolo chiave nella catalisi di reazioni di ossidoriduzione, almeno nel caso dei diseleniuri. Le barriere calcolate sono sistematicamente più alte di quelle delle ossidazioni iniziali, più basse per i ditellururi rispetto ai diseleniuri. Inoltre, le anidridi prodotte risultano ad una energia più alta dell'ossido iniziale rendendo la reazione termodinamicamente sfavorita.

7.3 Concluding remarks

SINCE its dawn, computational chemistry tried to provide an alternative way to obtain results on systems that could hardly (and sometimes they simply could not) be studied by traditional chemistry. The intricacies of the biological world have always been one of the preferred subject of computational chemists: having the possibility to actually “see” what happens during many of the processes that permit the existence of life and owning the tools to determine how each slight modification of the various components directly affects the properties of the system is very appealing. And yet, the possibility to accurately describe even the most basic biological processes that exceed the elementary chemical reactions realm, is still beyond our grasp. Nevertheless, the continuous research and development of new methods, joined by the ever growing possibilities offered by new computational technologies represent a very promising future which will see problems that seem now completely out of reach become routine work.

This Thesis uses different computational methodologies to provide useful insights on the properties and characteristics of the antioxidant (bio)chemistry of selenium, aimed at rational catalyst design. The plethora of different components involved in the efficiency of biological catalytic cycles makes the understanding of these processes a very complex matter. Moreover, the finesse acquired by Nature

in millions of years of evolution led to the creation of molecules whose structure and properties are difficult to replicate artificially. The complete understanding of the ruling principles underlying a process and the ability to transfer them to suitable structures that can be readily available are two key concepts in rational catalyst design. In the particular case of GPx mimics, not only should the results explain the characters that make selenium “special” among the chalcogens but they should also grant the ability to design chalcogen-based structures that could reproduce the function of enzymes. This is a very ambitious and challenging task and, although the perfect mimic has not been discovered yet, this work presents some useful insights that can be used as a good starting point for the successful design of a catalyst.

Why did Nature choose selenium? A simple answer to such an ambitious question is not possible but joining together the work of many, piece by piece, each one of them adding something to get the complete picture, will eventually provide us with a satisfying response. This Thesis wants to be one of these small pieces that can help complete this still unfinished puzzle.

*Considerate la vostra semenza:
fatti non foste a viver come bruti
ma per seguir virtute e canoscenza*
Dante Alighieri, *Divine Comedy*¹

¹This is a stanza from the dialogue between Dante and Ulysses in the XXVI Canto of the Inferno of the Divine Comedy. The lines report part of a speech given by Ulysses to his companions to convince them to go beyond the Pillar of Hercules (Strait of Gibraltar), anciently thought to be the end of the world and has been translated by Henry Wadsworth Longfellow as: "Consider ye the seed from which ye sprang / Ye were not made to live like unto brutes, / But for pursuit of virtue and of knowledge."

7.4 Acknowledgments

This work would not have been possible without the help of many people. Some of them contributed to its writing, while other had nothing to do with it but they accompanied me throughout the Ph.D. years. To me, they are all essential.

First, I would like to thank my colleagues of the Theoretical Chemistry group at the University of Padova, and in particular Francesco, Enrico Alessandro and Marco without whom days in “Sala Calcolo” would have been abysmal. A special thanks goes naturally to Prof. Orian which supervised during my entire (short) academic career, who constantly pushed and prodded me to improve the quality of my work.

Second, I am grateful to the Theoretical Chemistry group at the Vrije Universiteit Amsterdam for their friendliness and good spirits. They made me feel at ease right from the beginning, transforming my stay in Amsterdam in a pleasant “holiday”. I would particularly like to thank Dr. Trevor Hamlin who helped me on a day-to-day basis and Prof. Bickelhaupt who oversaw my project for the priceless suggestions and invaluable teachings.

Third, I must recognize the support of all my friends outside the academic world who rejoiced with me for my most important achievements, let me burden them with my sorrows in the worst days and, most important of all, always managed to put a smile on my face. They are many, but I’ll try to remember them all (in no particular order): Alessio, Francesco, Davide, Matteo, Laura, Enrico, Giulio, Martina, Giorgia, Cristian, Anna, Alessia, Manuel, Andrea, Marco, Malindi, Davide, Paolo, Maria, Monica, Giacomo, Lavinia, Alessia, Laura. Instead of writing many empty words I would just like to offer you not a toast, but *our* toast: “*Prosit!*”

Last, but certainly not least my family. *Mamma, Papà e Davide*: there are no words that can fully describe what you have given me in all my life. To you I just dedicate this simple words: “*Grazie. Grazie di tutto*”.

7.5 List of Publications

- M. Bortoli, L. P. Wolters, L. Orian, F. M. Bickelhaupt *Addition–elimination or nucleophilic substitution? Understanding the energy profiles for the reaction of chalcogenolates with dichalcogenides* J. Chem. Theory Comput., **2016**, 12, 2752–2761
- G. Ribauda, M. Bellanda, I. Menegazzo, L. P. Wolters, M. Bortoli, G. Ferrer-Sueta, G. Zagotto, L. Orian, *Mechanistic insight into the oxidation of organic phenylselenides by H₂O₂*, Chem. - Eur. J., **2017**, 23, 2405–2422.
- M. Bortoli, M. Torsello, L. Orian, F. M. Bickelhaupt *Role of the chalcogen (S, Se, Te) in the oxidation mechanism of the glutathione peroxidase active site*, ChemPhysChem, **2017**, 18, 2990-2998.
- M. Bortoli, F. Zaccaria, M. Dalla Tiezza, M. Bruschi, C. Fonseca Guerra, F. M. Bickelhaupt, L. Orian *Oxidation of organic diselenides and ditellurides by H₂O₂ for bioinspired catalyst design*, Phys. Chem. Chem. Phys., **2018**, 20, 20874-20885
- P. Calligari, M. Torsello, M. Bortoli, L. Orian, A. Polimeno *Insights on the Ca²⁺-promoted structural effects in wild type and post-translationally modified Connexin26*, manuscript submitted
- M. Bortoli, S. M. Ahmad, T. A. Hamlin, F. M. Bickelhaupt, L. Orian *Nature and strength of chalcogen- π interactions*, Phys. Chem. Chem. Phys., **2018**, DOI:10.1039/C8CP05922E

References

- [1] J. J. Berzelius, *Ann. Chim. Phys* **1818**, 7, 199–206.
- [2] J. Perlin, *Let It Shine: The 6,000-Year Story of Solar Energy*, New World Library, **2013**.
- [3] H. Dudley, *Public Health Rep.* **1938**, 53, 281–292.
- [4] P. Armor, A.J.; Pringle, *Bull. Hyg. (Lond.)* **1945**, 20, 239–241.
- [5] K. W. Franke, *J. Nutr.* **1934**, 8, 597.
- [6] O. Beath, H. Eppson, C. Gilbert, *Wyoming Agricultural Experiment Station Bulletin* **1935**, 206, 55–73.
- [7] E. P. Painter, *Chem. Rev.* **1941**, 28, 179–213.
- [8] K. W. Franke, A. Moxon, W. Poley, W. Tully, *Anat. Rec.* **1936**, 65, 15–22.
- [9] A. A. Nelson, O. G. Fitzhugh, H. O. Calvery, *Cancer Res.* **1943**, 3, 230–236.
- [10] D. Robertson, *The Lancet* **1970**, 295, 518–519.
- [11] J. Pinsent, *The Biochemical journal* **1954**, 57, 10–16.
- [12] E. L. Patterson, R. Milstrey, E. L. R. Stokstad, *Exp. Biol. Med.* **1957**, 95, 617–620.

- [13] K. Schwarz, J. G. Bieri, G. M. Briggs, M. L. Scott, *Exp. Biol. Med.* **1957**, *95*, 621–625.
- [14] K. Schwarz, *Exp. Biol. Med.* **1951**, *78*, 852–856.
- [15] K. Schwarz, C. M. Foltz, *J. Am. Chem. Soc.* **1957**, *79*, 3292–3293.
- [16] L. Flohé, *Biochimica et Biophysica Acta (BBA) - General Subjects* **2009**, *1790*, 1389–1403.
- [17] K. Schwarz, C. M. Foltz, *The Journal of biological chemistry* **1958**, *233*, 245–251.
- [18] J. A. Tan, S. F. Hou, W. Y. Zhu, R. B. Li, D. X. Zheng, W. M. Y., (G. o. E. Disease), Endemic, *Acta Geogr. Sin.* **1979**, *46*, 85.
- [19] H. J. Reich, R. J. Hondal, *ACS Chem. Bio.* **2016**, *11*, 821–841.
- [20] L. C. Clark, G. F. Combs, B. W. Turnbull, E. H. Slate, D. K. Chalker, J. Chow, L. S. Davis, R. A. Glover, G. F. Graham, E. G. Gross, A. Krongrad, J. L. Lesher, H. K. Park, B. B. Sanders, C. L. Smith, J. R. Taylor, *JAMA* **1996**, *276*, 1957.
- [21] S. M. Lippman, P. J. Goodman, E. A. Klein, H. L. Parnes, I. M. Thompson, A. R. Kristal, R. M. Santella, J. L. Probstfield, C. M. Moinpour, D. Albanes, P. R. Taylor, L. M. Minasian, A. Hoque, S. M. Thomas, J. J. Crowley, J. M. Gaziano, J. L. Stanford, E. D. Cook, N. E. Fleshner, M. M. Lieber, P. J. Walther, F. R. Khuri, D. D. Karp, G. G. Schwartz, L. G. Ford, C. A. Coltman, *JNCI Journal of the National Cancer Institute* **2005**, *97*, 94–102.
- [22] S. M. Lippman, E. A. Klein, P. J. Goodman, M. S. Lucia, I. M. Thompson, L. G. Ford, H. L. Parnes, L. M. Minasian, J. M. Gaziano, J. A. Hartline, J. K. Parsons, J. D. Bearden, E. D. Crawford, G. E. Goodman, J. Claudio, E. Winquist, E. D. Cook, D. D. Karp, P. Walther, M. M. Lieber, A. R. Kristal, A. K. Darke, K. B. Arnold, P. A. Ganz, R. M. Santella, D. Albanes, P. R.

- Taylor, J. L. Probstfield, T. J. Jagpal, J. J. Crowley, F. L. Meyskens, L. H. Baker, C. A. Coltman, *JAMA* **2009**, *301*, 39.
- [23] M.-H. Yoo, X.-M. Xu, B. A. Carlson, A. D. Patterson, V. N. Gladyshev, D. L. Hatfield, *PLoS One* **2007**, *2*, (Ed.: M. Blagosklonny), e1112.
- [24] D. L. Hatfield, V. N. Gladyshev, *Mol. Interv.* **2009**, *9*, 18–21.
- [25] C. D. Davis, P. A. Tsuji, J. A. Milner, *Annu. Rev. Nutr.* **2012**, *32*, 73–95.
- [26] World Health Organization, *Guidelines for drinking-water quality: fourth edition incorporating the first addendum*. World Health Organization, Geneva, **2017**, p. 415.
- [27] T. H. Jukes, *J. Appl. Biochem.* **1983**, *5*, 233–4.
- [28] L. Flohé, B. Eisele, A. Wendel, *Hoppe-Seyler's Zeitschrift für physiologische Chemie* **1971**, *352*, 151–158.
- [29] L. Flohé, E. Schaich, W. Voelter, A. Wendel, *Hoppe-Seyler's Zeitschrift für Physiologische Chemie* **1971**, *352*, 170–180.
- [30] J. T. Rotruck, A. L. Pope, H. E. Ganther, W. G. Hoekstra, *The Journal of Nutrition* **1972**, *102*, 689–696.
- [31] J. T. Rotruck, W. G. Hoekstra, A. L. Pope, H. E. Ganther, A. B. Swanson, D. Hafemann, *Fed. Proc.* **1972**, *31*, 691.
- [32] L. Flohé, E. A. Günzler, H. H. Schock, *FEBS Lett.* **1973**, *32*, 132–137.
- [33] S.-H. Oh, H. E. Ganther, W. G. Hoekstra, *Biochemistry* **1974**, *13*, 1825–1829.
- [34] W. Nakamura, S. Hosoda, K. Hayashi, *Biochimica et Biophysica Acta (BBA) - Enzymology* **1974**, *358*, 251–261.
- [35] Y. C. Awasthi, E. Beutler, S. K. Srivastava, *J. Biol. Chem.* **1975**, *250*, 5144–5149.

- [36] Y. C. Awasthi, D. D. Dao, A. K. Lal, S. K. Srivastava, *The Biochemical journal* **1979**, *177*, 471–6.
- [37] J. Bell, C. B. Cowey, A. Youngson, *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1984**, *795*, 91–99.
- [38] J. E. Cone, R. M. D. Río, J. N. Davis, T. C. Stadtman, *Proceedings of the National Academy of Sciences* **1976**, *73*, 2659–2663.
- [39] A. Böck, K. Forchhammer, J. Heider, W. Leinfelder, G. Sawers, B. Veprék, F. Zinoni, *Mol. Microbiol.* **1991**, *5*, 515–520.
- [40] I. Chambers, J. Frampton, P. Goldfarb, N. Affara, W. McBain, P. R. Harrison, *The EMBO journal* **1986**, *5*, 1221–7.
- [41] F. Zinoni, A. Birkmann, T. C. Stadtman, A. Böck, *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 4650–4.
- [42] S. C. Low, M. J. Berry, Knowing when not to stop: Selenocysteine incorporation in eukaryotes, **1996**.
- [43] V. N. Gladyshev, G. V. Kryukov, *BioFactors* **2001**, *14*, 87–92.
- [44] R. J. Hondal, E. L. Ruggles, *Amino Acids* **2011**, *41*, 73–89.
- [45] S. Osawa, T. H. Jukes, K. Watanabe, A. Muto, *Microbiol. Rev.* **1992**, *56*, 229–64.
- [46] M. J. Axley, A. Böck, T. C. Stadtman, *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 8450–4.
- [47] M. J. Berry, A. L. Maia, J. D. Kieffer, J. W. Harney, P. R. Larsen, *Endocrinology* **1992**, *131*, 1848–1852.
- [48] S. Hazebrouck, L. Camoin, Z. Faltin, A. D. Strosberg, Y. Eshdat, *The Journal of biological chemistry* **2000**, *275*, 28715–21.
- [49] H.-Y. Kim, V. N. Gladyshev, *PLoS Biol.* **2005**, *3*, (Ed.: R. Matthews), e375.

- [50] G. M. Lacourciere, T. C. Stadtman, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 44–8.
- [51] S. M. Kanzok, *Science* **2001**, *291*, 643–646.
- [52] R. E. Huber, R. S. Criddle, *Arch. Biochem. Biophys.* **1967**, *122*, 164–173.
- [53] E. S. Arnér, *Exp. Cell Res.* **2010**, *316*, 1296–1303.
- [54] J. R. Arthur, *Cell. Mol. Life Sci.* **2000**, *57*, 1825–1835.
- [55] R. Brigelius-Flohé, M. Maiorino, *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013**, *1830*, 3289–3303.
- [56] F. M. Maiorino, R. Brigelius-Flohé, K. D. Aumann, A. Roveri, D. Schomburg, L. Flohé, *Methods Enzymol.* **1995**, *252*, 38–48.
- [57] O. Epp, R. Ladenstein, A. Wendel, *Eur. J. Biochem.* **1983**, *133*, 51–69.
- [58] F. Ursini, M. Maiorino, C. Gregolin, *BBA - Gen. Subj.* **1985**, *839*, 62–70.
- [59] L. Flohé, S. Toppo, G. Cozza, F. Ursini, *Antioxid. Redox Signaling* **2011**, *15*, 763–780.
- [60] S. Toppo, S. Vanin, V. Bosello, S. C. E. Tosatto, *Antioxid. Redox Signaling* **2008**, *10*, 1501–14.
- [61] M. Mariotti, P. G. Ridge, Y. Zhang, A. V. Lobanov, T. H. Pringle, R. Guigo, D. L. Hatfield, V. N. Gladyshev, *PLoS One* **2012**, *7*, (Ed.: V. Laudet), e33066.
- [62] S. C. E. Tosatto, V. Bosello, F. Fogolari, P. Mauri, A. Roveri, S. Toppo, L. Flohé, F. Ursini, M. Maiorino, *Antioxid. Redox Signaling* **2008**, *10*, 1515–1526.
- [63] G. C. Mills, *J. Biol. Chem.* **1957**, *229*, 189–197.
- [64] F. F. Chu, J. H. Doroshov, R. S. Esworthy, *J. Biol. Chem.* **1993**, *268*, 2571–2576.

- [65] M. Bjornstedt, J. Y. Xue, W. H. Huang, B. Akesson, A. Holmgren, *J. Biol. Chem.* **1994**, *269*, 29382–29384.
- [66] M. Maiorino, V. Bosello-Travain, G. Cozza, G. Miotto, L. Orian, A. Roveri, S. Toppo, M. Zaccarin, F. Ursini in *Selenium: its Molecular Biology and Role in Human Health*, (Eds.: D. L. Hatfield, U. Schweizer, P. A. Tsuji, V. N. Gladyshev), Springer International Publishing, **2016**, pp. 223–234.
- [67] F. Ursini, M. Maiorino, M. Valente, L. Ferri, C. Gregolin, *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1982**, *710*, 197–211.
- [68] N. E. Savaskan, C. Ufer, H. Kühn, A. Borchert, *Biol. Chem.* **2007**, *388*, 1007–1017.
- [69] A. Wendel, M. Fausel, H. Safayhi, G. Tiegs, R. Otter, *Biochem. Pharmacol.* **1984**, *33*, 3241–3245.
- [70] A. Müller, E. Cadenas, P. Graf, H. Sies, *Biochem. Pharmacol.* **1984**, *33*, 3235–3239.
- [71] G. Mugesh, W. W. Du Mont, H. Sies, *Chem. Rev.* **2001**, *101*, 2125–2179.
- [72] K. P. Bhabak, G. Mugesh, *Acc. Chem. Res.* **2010**, *43*, 1408–1419.
- [73] L. Orian, S. Toppo, *Free Rad. Biol. Med.* **2014**, *66*, 65–74.
- [74] M. Ibrahim, W. Hassan, D. F. Meinerz, M. dos Santos, C. V. Klimaczewski, A. M. Deobald, M. S. Costa, C. W. Nogueira, N. B. V. Barbosa, J. B. T. Rocha, *Mol. Cell. Biochem.* **2012**, *371*, 97–104.
- [75] J. B. T. Rocha, B. C. Piccoli, C. S. Oliveira, *Arkivoc* **2017**, 457–491.
- [76] D. F. Meinerz, J. Allebrandt, D. O. Mariano, E. P. Waczuk, F. A. Soares, W. Hassan, J. B. T. Rocha, *PeerJ* **2014**, *2*, e290.
- [77] C. W. Nogueira, J. B. T. Rocha, *Arch. Toxicol.* **2011**, *85*, 1313–1359.
- [78] C. W. Nogueira, G. Zeni, J. B. T. Rocha, *Chem. Rev.* **2004**, *104*, 6255–6285.

- [79] L. P. Wolters, L. Orian, *Curr. Org. Chem.* **2016**, *20*, 189–197.
- [80] M. Iwaoka, S. Tomoda, *J. Am. Chem. Soc.* **1994**, *116*, 2557–2561.
- [81] G. S. Heverly-Coulson, R. J. Boyd, *J. Phys. Chem. A* **2010**, *114*, 1996–2000.
- [82] M. Maiorino, A. Roveri, M. Coassin, F. Ursini, *Biochem. Pharmacol.* **1988**, *37*, 2267–2271.
- [83] R. S. Glass, F. Farooqui, M. Sabahi, K. W. Ehler, *J. Org. Chem.* **1989**, *54*, 1092–1097.
- [84] G. R. Haenen, B. M. de Rooij, N. P. Vermeulen, A. Bast, *Mol. Pharmacol.* **1990**, *37*, 412–422.
- [85] J. K. Pearson, R. J. Boyd, *J. Phys. Chem. A* **2006**, *110*, 8979–8985.
- [86] J. K. Pearson, R. J. Boyd, *J. Phys. Chem. A* **2007**, *111*, 3152–3160.
- [87] G. Mugesh, H. B. Singh, *Chem. Soc. Rev.* **2000**, *29*, 347–357.
- [88] X. Liu, L. A. Silks, C. Liu, M. Ollivault-Shiflett, X. Huang, J. Li, G. Luo, Y. M. Hou, J. Liu, J. Shen, *Angew. Chem. Int. Ed.* **2009**, *48*, 2020–2023.
- [89] S. Mao, Z. Dong, J. Liu, X. Li, X. Liu, G. Luo, J. Shen, *J. Am. Chem. Soc.* **2005**, *127*, 11588–11589.
- [90] E. R. T. Tiekink, *Dalton Trans.* **2012**, *41*, 6390.
- [91] E. Schrödinger, *Ann. Phys.* **1926**, *80*, 437–490.
- [92] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864–B871.
- [93] L. H. Thomas, *Math. Proc. Cambridge Philos. Soc.* **1927**, *23*, 542.
- [94] E. Fermi, *Rendiconti Accademia Nazionale dei Lincei* **1927**, *6*, 602–607.
- [95] E. Fermi, *Zeitschrift für Physik* **1928**, *48*, 73–79.
- [96] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133–A1138.

- [97] F. M. Bickelhaupt, K. N. Houk, *Angewandte Chemie - International Edition* **2017**, *56*, 10070–10086.
- [98] I. Fernández, F. M. Bickelhaupt, *Chem. Soc. Rev.* **2014**, *43*, 4953.
- [99] F. M. Bickelhaupt, E. J. Baerends in *Reviews in Computational Chemistry*, Vol. 15, (Eds.: K. B. Lipkowitz, D. B. Boyd), Wiley-VCH, New York, **2000**, Chapter Vol. 15, pp. 1–86.
- [100] K. Morokuma, *J. Chem. Phys.* **1971**, *55*, 1236–1244.
- [101] T. Ziegler, A. Rauk, *Theor. Chim. Acta* **1977**, *46*, 1–10.
- [102] A. Klamt, G. Schüürmann, *J. Chem. Soc. Perkin Trans. 2* **1993**, *0*, 799–805.
- [103] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378–6396.
- [104] A. Klamt, C. Moya, J. Palomar, *J. Chem. Theory Comput.* **2015**, *11*, 4220–4225.
- [105] A. Klamt, *WIREs Comput. Mol. Sci.* **2018**, *8*, e1338.
- [106] E. Cancès, B. Mennucci, *J. Math. Chem.* **1998**, *23*, 309–326.
- [107] E. Cancès, B. Mennucci, J. Tomasi, *The Journal of Chemical Physics* **1998**, *107*, 3032.
- [108] J. Tomasi, B. Mennucci, E. Cancès, *J. Mol. Struct. THEOCHEM* **1999**, *464*, 211–226.
- [109] D. Case, J. Berryman, R. Betz, D. Cerutti, T. Cheatham III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K. Merz, D. York, P. Kollman, *AMBER 2015*, University of California, San Francisco, San Francisco, **2015**.

- [110] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- [111] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931–967.
- [112] C. Fonseca Guerra, J. G. Snijders, G. te Velde, E. J. Baerends, *Theor. Chem. Acc.* **1998**, *99*, 391–403.
- [113] E. J. Baerends, T. Ziegler, A. J. Atkins, J. Autschbach, D. Bashford, A. Bérces, F. M. Bickelhaupt, C. Bo, P. M. Boerritger, L. Cavallo, D. P. Chong, D. V. Chulhai, L. Deng, R. M. Dickson, J. M. Dieterich, D. E. Ellis, M. van Faassen, A. Ghysels, A. Giammona, S. J. A. van Gisbergen, A. W. Götz, S. Gusarov, F. E. Harris, P. van den Hoek, C. R. Jacob, H. Jacobsen, L. Jensen, J. W. Kaminski, G. van Kessel, F. Kootstra, A. Kovalenko, M. Krykunov, E. van Lenthe, D. A. McCormack, A. Michalak, M. Mitoraj, S. M. Morton, J. Neugebauer, V. P. Nicu, L. Noodleman, V. P. Osinga, S. Patchkovskii, M. Pavanello, C. A. Peebles, P. H. T. Philipsen, D. Post, C. C. Pye, W. Ravenek, J. I. Rodríguez, P. Ros, R. Rüger, P. R. T. Schipper, H. van Schoot, G. Schreckenbach, J. S. Seldenthuis, M. Seth, J. G. Snijders, S. Miquel, M. Swart, D. Swerhone, G. te Velde, P. Vernooijs, L. Versluis, L. Visscher, O. Visser, F. Wang, T. A. Wesolowski, E. M. van Wezenbeek, G. Wiesenekker, S. K. Wolff, T. K. Woo, A. L. Yakovlev, ADF2016, Amsterdam, The Netherlands, **2016**.
- [114] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd,

- E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, J. A. Montgomery Jr., *Gaussian 09, Rev D.01*, Gaussian, Inc., Wallingford CT, **2009**.
- [115] D. Case, R. Betz, D. Cerutti, T. Cheatham III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, I. Omelyan, D. Onufriev, A. andRoe, A. Roitberg, C. Sagui, C. Simmerling, B.-S. W. M., J. Swails, R. C. Walker, R. M. Wang, J. andWolf, X. Wu, L. Xiao, P. Kollman, *AMBER 2016*, University of California, San Francisco, San Francisco, **2016**.
- [116] H. J. Schneider, *Angew. Chem. Int. Ed.* **2009**, *48*, 3924–3977.
- [117] G. Gilli, P. Gilli, *The Nature of the Hydrogen Bond: Outline of a Comprehensive Hydrogen Bond Theory*, Oxford University Press, Oxford, UK, **2009**, pp. 1–336.
- [118] P. Politzer, P. Lane, M. C. Concha, Y. Ma, J. S. Murray, *J. Mol. Model.* **2007**, *13*, 305–311.
- [119] P. Auffinger, F. A. Hays, E. Westhof, P. S. Ho, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16789–94.
- [120] P. Politzer, J. S. Murray, T. Clark, *Phys. Chem. Chem. Phys.* **2013**, *15*, 11178.
- [121] W. Wang, B. Ji, Y. Zhang, *J. Phys. Chem. A* **2009**, *113*, 8132–8135.

- [122] D. J. Pascoe, K. B. Ling, S. L. Cockroft, *J. Am. Chem. Soc.* **2017**, *139*, jacs.7b08511.
- [123] V. Oliveira, D. Cremer, E. Kraka, *J. Phys. Chem. A* **2017**, *121*, 6845–6862.
- [124] V. Oliveira, E. Kraka, *J. Phys. Chem. A* **2017**, *121*, 9544–9556.
- [125] A. Bauzá, D. Quiñero, P. M. Deyà, A. Frontera, *CrystEngComm* **2013**, *15*, 3137–3144.
- [126] K. T. Mahmudov, M. N. Kopylovich, M. F. C. Guedes da Silva, A. J. Pombeiro, *Coordination Chemistry Reviews* **2017**, *345*, 54–72.
- [127] S. Benz, J. López-Andarias, J. Mareda, N. Sakai, S. Matile, J. López-Andarias, J. Mareda, N. Sakai, S. Matile, *Angewandte Chemie - International Edition* **2017**, *56*, 812–815.
- [128] S. Benz, M. Macchione, Q. Verole, J. Mareda, N. Sakai, S. Matile, *J. Am. Chem. Soc.* **2016**, *138*, 9093–9096.
- [129] K. T. Mahmudov, M. N. Kopylovich, M. F. C. Guedes da Silva, A. J. L. Pombeiro, *Dalton Trans.* **2017**, *46*, 10121–10138.
- [130] K. Kříž, J. Fanfrlík, M. Lepšík, *ChemPhysChem* **2018**, DOI 10.1002/cphc.201800409.
- [131] M. D. Esrafilí, M. Vakili, *Mol. Phys.* **2014**, *112*, 2746–2752.
- [132] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
- [133] B. G. Johnson, P. M. W. Gill, J. A. Pople, *J. Chem. Phys.* **1993**, *98*, 5612–5626.
- [134] T. V. Russo, R. L. Martin, P. J. Hay, *J. Med. Phys.* **1994**, *101*, 7729–7737.
- [135] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.
- [136] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- [137] E. van Lenthe, E. J. Baerends, J. G. Snijders, *J. Chem. Phys.* **1994**, *101*, 9783.

- [138] F. L. Hirshfeld, *Theor. Chim. Acta* **1977**, *44*, 129–138.
- [139] C. Fonseca Guerra, J. W. Handgraaf, E. J. Baerends, F. M. Bickelhaupt, *J. Comput. Chem.* **2004**, *25*, 189–210.
- [140] C. Riplinger, F. Neese, *J. Chem. Phys.* **2013**, *138*, 034106.
- [141] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.
- [142] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *J. Chem. Phys.* **2013**, *139*, 134101.
- [143] C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, F. Neese, *J. Chem. Phys.* **2016**, *144*, 024109.
- [144] M. Reiher, *WIREs Comput. Mol. Sci.* **2012**, *2*, 139–149.
- [145] S. F. Boys, F. Bernardi, *Molecular Physics* **1970**, *19*, 553–566.
- [146] M. Mentel, E. J. Baerends, *J. Chem. Theory Comput.* **2014**, *10*, 252–267.
- [147] S. Toppo, L. Flohé, F. Ursini, S. Vanin, M. Maiorino, *Biochim. Biophys. Acta Gen. Subj.* **2009**, *1790*, 1486–1500.
- [148] L. Orian, P. Mauri, A. Roveri, S. Toppo, L. Benazzi, V. Bosello-Travain, A. De Palma, M. Maiorino, G. Miotto, M. Zaccarin, A. Polimeno, L. Flohé, F. Ursini, *Free Rad. Biol. Med.* **2015**, *87*, 1–14.
- [149] R. Prabhakar, T. Vreven, K. Morokuma, D. G. Musaev, *Biochemistry* **2005**, *44*, 11864–11871.
- [150] R. Prabhakar, T. Vreven, M. J. Frisch, K. Morokuma, D. G. Musaev, *J. Phys. Chem. B* **2006**, *110*, 13608–13613.
- [151] G. Roy, M. Nethaji, G. Mugesh, *J. Am. Chem. Soc.* **2004**, *126*, 2712–2713.
- [152] X. Ren, Y. Xue, J. Liu, K. Zhang, J. Zheng, G. Luo, C. Guo, Y. Mu, J. Shen, *ChemBioChem* **2002**, *3*, 356–363.

- [153] H. Zhong, E. W. Taylor, *J. Mol. Graph. Model.* **2004**, *23*, 223–231.
- [154] D. Dimastrogiovanni, M. Anselmi, A. E. Miele, G. Boumis, L. Petersson, F. Angelucci, A. Di Nola, M. Brunori, A. Bellelli, *Proteins: Struct. Funct. Bioinf.* **2010**, *78*, 259–270.
- [155] O. Gutiérrez-Sanz, M. C. Marques, C. S. A. Baltazar, V. M. Fernández, C. M. Soares, I. A. C. Pereira, A. L. De Lacey, *J. Biol. Inorg. Chem.* **2013**, *18*, 419–427.
- [156] V. H. Teixeira, A. S. C. Capacho, M. Machuqueiro, *Proteins: Struct. Funct. Bioinf.* **2016**, *84*, 1836–1843.
- [157] S. T. Ali, S. Jahangir, S. Karamat, W. M. F. Fabian, K. Nawara, J. Kóňa, *J. Chem. Theory Comput.* **2010**, *6*, 1670–1681.
- [158] J. Kóňa, W. M. F. Fabian, *J. Chem. Theory Comput.* **2011**, *7*, 2610–2616.
- [159] P. Scheerer, A. Borchert, N. Krauss, H. Wessner, C. Gerth, W. Höhne, H. Kuhn, *Biochemistry* **2007**, *46*, 9041–9049.
- [160] M. Torsello, A. C. Pimenta, L. P. Wolters, I. S. Moreira, L. Orian, A. Polimeno, *J. Phys. Chem. A* **2016**, *120*, 4389–4400.
- [161] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
- [162] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.
- [163] S. H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200–1211.
- [164] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- [165] J. Antony, R. Sure, S. Grimme, *Chem. Commun.* **2015**, *51*, 1764–1774.
- [166] L. Li, C. Li, Z. Zhang, E. Alexov, *J. Chem. Theory Comput.* **2013**, *9*, 2126–2136.

- [167] R. F. Ribeiro, A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2011**, *115*, 14556–14562.
- [168] E. A. Cowan, C. D. Oldham, S. W. May, *Arch. Biochem. Biophys.* **2011**, *506*, 201–207.
- [169] K. Arai, K. Dedachi, M. Iwaoka, *Chemistry - Eur. J.* **2011**, *17*, 481–5.
- [170] A. S. Hodage, P. P. Phadnis, A. Wadawale, K. I. Priyadarsini, V. K. Jain, *Org. Biomol. Chem.* **2011**, *9*, 2992–8.
- [171] G. Ribaud, M. Bellanda, I. Menegazzo, L. P. Wolters, M. Bortoli, G. Ferrer-Sueta, G. Zagotto, L. Orian, *Chemistry - Eur. J.* **2017**, *23*, 2405–2422.
- [172] F. Zaccaria, L. P. Wolters, C. Fonseca Guerra, L. Orian, *J. Comput. Chem.* **2016**, *37*, 1672–1680.
- [173] Y. Nishibayashi, S. Uemura, *Top. Curr. Org. Chem* **2000**, *208*, 201–233.
- [174] S. Santoro, J. B. Azeredo, V. Nascimento, L. Sancineto, A. L. Braga, C. Santi, *RSC Adv.* **2014**, *4*, 31521–31535.
- [175] C. Santi, A. Zając, A. L. Braga, B. Movassag, C. Tidei, C. Santi, T. Cristina, B. Desirée, D. Pietrella, F. Marini, F. Galli, M. Giuseppe, J. Ścianowski, J. Lewkowski, J. Drabowicz, L. Bagnoli, P. Marta, M. Iwaoka, M. Navidi, O. E. D. Rodrigues, P. Pokora-Sobczak, P. Kielbasiński, R. S. Schwab, S. Hayashi, C. Silvia, S. Sternativo, T. Murai, W. Nakanishi, Z. Rafiński, *Organoselenium Chemistry: Between Synthesis and Biochemistry*, (Ed.: C. Santi), Betham, **2014**.
- [176] A. Fava, A. Iliceto, E. Camera, *J. Am. Chem. Soc.* **1957**, *79*, 833–838.
- [177] L. Senatore, E. Ciuffarin, A. Fava, G. Levita, *J. Am. Chem. Soc.* **1973**, *95*, 2918–2922.

- [178] D. A. Keire, E. Strauss, W. Guo, B. Noszal, D. L. Rabenstein, *J. Org. Chem.* **1992**, *57*, 123–127.
- [179] D. Steinmann, T. Nauser, W. H. Koppenol, *J. Org. Chem.* **2010**, *75*, 6696–6699.
- [180] T. G. Back, Z. Moussa, *J. Am. Chem. Soc.* **2003**, *125*, 13455–13460.
- [181] T. G. Back, Z. Moussa, M. Parvez, *Angew. Chem. Int. Ed.* **2004**, *43*, 1268–1270.
- [182] C.-M. Andersson, A. Hallberg, R. Brattsand, I. A. Cotgreave, L. Engman, J. Persson, *Bioorg. Med. Chem. Lett.* **1993**, *3*, 2553–2558.
- [183] L. Engman, D. Stern, M. Pelcman, C. M. Andersson, *J. Org. Chem.* **1994**, *59*, 1973–1979.
- [184] R. van Leeuwen, E. van Lenthe, E. J. Baerends, J. G. Snijders, *J. Chem. Phys.* **1994**, *101*, 1272–1281.
- [185] N. C. Handy, A. J. Cohen, *Mol. Phys.* **2001**, *99*, 403–412.
- [186] N. L. Allinger, X. Zhou, J. Bergsma, *J. Mol. Struct. THEOCHEM* **1994**, *312*, 69–83.
- [187] S. M. Bachrach, D. W. Demoin, M. Luk, J. V. Miller, *J. Phys. Chem. A* **2004**, *108*, 4040–4046.
- [188] M. A. van Bochove, M. Swart, F. M. Bickelhaupt, *J. Am. Chem. Soc.* **2006**, *128*, 10738–10744.
- [189] A. P. Bento, F. M. Bickelhaupt, *Chem. Asian J.* **2008**, *3*, 1783–1792.
- [190] A. P. Bento, F. M. Bickelhaupt, *J. Org. Chem.* **2008**, *73*, 7290–7299.
- [191] A. P. Bento, F. M. Bickelhaupt, *J. Org. Chem.* **2007**, *72*, 2201–2207.
- [192] W. J. van Zeist, Y. Ren, F. M. Bickelhaupt, *Sci. China: Chemistry* **2010**, *53*, 210–215.

- [193] L. P. Wolters, F. M. Bickelhaupt, *ChemistryOpen* **2012**, *1*, 96–105.
- [194] J. K. Laerdahl, E. Uggerud, *Int. J. Mass Spectrom.* **2002**, *214*, 277–314.
- [195] M. A. van Bochove, F. M. Bickelhaupt, *Eur. J. Org. Chem.* **2008**, *2008*, 649–654.
- [196] P. A. Fernandes, M. J. Ramos, *Chemistry - Eur. J.* **2004**, *10*, 257–266.
- [197] R. D. Bach, O. Dmitrenko, C. Thorpe, *J. Org. Chem.* **2008**, *73*, 12–21.
- [198] A. T. Carvalho, M. Swart, J. N. Van Stralen, P. A. Fernandes, M. J. Ramos, F. M. Bickelhaupt, *J. Phys. Chem. B* **2008**, *112*, 2511–2523.
- [199] J. M. Hayes, S. M. Bachrach, *J. Phys. Chem. A* **2003**, *107*, 7952–7961.
- [200] S. M. Bachrach, J. M. Hayes, T. Dao, J. L. Mynar, *Theor. Chem. Acc.* **2002**, *107*, 266–271.
- [201] S. Antony, C. A. Bayse, *Inorg. Chem.* **2011**, *50*, 12075–12084.
- [202] G. S. Heverly-Coulson, R. J. Boyd, O. M6, M. Y6nez, *Chemistry - Eur. J.* **2013**, *19*, 3629–3638.
- [203] S. M. Bachrach, D. C. Mulhearn, *The Journal of Physical Chemistry* **1996**, *100*, 3535–3540.
- [204] D. C. Mulhearn, S. M. Bachrach, *Journal of the American Chemical Society* **1996**, *118*, 9415–9421.
- [205] S. M. Bachrach, B. D. Gailbreath, *Journal of Organic Chemistry* **2001**, *66*, 2005–2010.
- [206] S. M. Bachrach, A. C. Chamberlin, *Journal of Organic Chemistry* **2003**, *68*, 4743–4747.
- [207] T. van der Wijst, C. Fonseca Guerra, M. Swart, F. M. Bickelhaupt, B. Lippert, *Angewandte Chemie International Edition*, *48*, 3285–3287.

-
- [208] C. Fonseca Guerra, T. van der Wijst, J. Poater, M. Swart, F. M. Bickelhaupt, *Theoretical Chemistry Accounts* **2010**, *125*, 245–252.
- [209] W. J. van Zeist, C. Fonseca Guerra, F. M. Bickelhaupt, *Journal of computational chemistry* **2007**, *28*, 73–86.
- [210] F. M. Bickelhaupt, M. Solà, P. von Ragué Schleyer, *J. Comput. Chem.* **1995**, *16*, 465–477.
- [211] M. El-Hamdi, J. Poater, F. M. Bickelhaupt, M. Solà, *Inorg. Chem.* **2013**, *52*, 2458–2465.
- [212] J. Młochowski, H. Wójtowicz-Młochowska, *Molecules* **2015**, *20*, 10205–10243.