UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Farmaceutiche

Scuola di Dottorato di Ricerca in Scienze Molecolari

Indirizzo in Scienze Farmaceutiche

XXI Ciclo

Protein Engineering by Chemical and Genetic Methods

Applications to Protein Recognition and Thrombin Function

Direttore della Scuola: Ch.mo Prof. Maurizio CASARIN

Supervisore: Ch.mo Prof. Vincenzo DE FILIPPIS

Dottoranda: Roberta FRASSON

UNIVERSITY OF PADUA

Department of Pharmaceutical Sciences

Ph.D. School in Molecular Sciences
Pharmaceutical Sciences Curriculum

XXI Cycle

Protein Engineering by Chemical and Genetic Methods

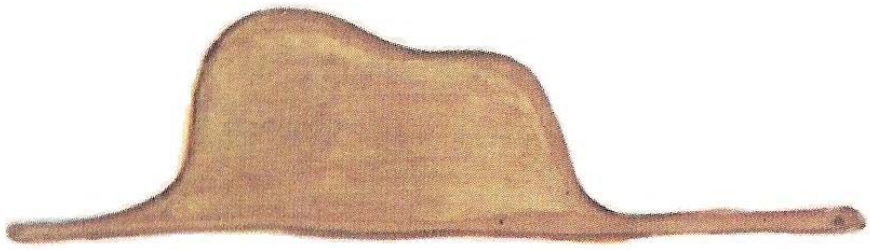
Applications to Protein Recognition and Thrombin Function

School Director: Prof Maurizio CASARIN

Supervisor: Prof. Vincenzo DE FILIPPIS

Ph.D. Student: Roberta FRASSON

A Enrico



“Addio”, disse la volpe. “Ecco il mio segreto. E’ molto semplice: non si vede bene che col cuore. L’essenziale è invisibile agli occhi. E’ il tempo che hai perduto per la tua rosa che ha fatto la tua rosa così importante. Gli uomini hanno dimenticato questa verità. Ma tu non la devi dimenticare. Tu diventi responsabile per sempre di quello che hai addomesticato. Tu sei responsabile della tua rosa...”

“Io sono responsabile della mia rosa...” ripetè il piccolo principe per ricordarselo.

Il piccolo principe, Antoine De Saint-Exupéry, Capitolo 21

Contents

Riassunto

Negli ultimi 30 anni, lo sviluppo della tecnologia del DNA ricombinante ha permesso l'alterazione sito-specifica di una determinata catena polipeptidica, espandendo enormemente lo studio dei meccanismi molecolari implicati nel *folding*, nella stabilità e nella funzione delle proteine. Questo approccio, conosciuto anche come ingegneria proteica, è basato sulla possibilità di modificare la composizione chimica di una proteina in un sola o multiple posizioni con uno dei 20 amminoacidi naturali, ottenendo così proteine ricombinanti mutate aventi struttura,stabilità e funzione alterate. L'analisi degli effetti della mutazione sulle caratteristiche delle proteine così ottenute permette di determinare in maniera quantitativa il contributo di uno specifico amminoacido che è stato mutato, nella proteina *wild-type*. Questa informazione a sua volta potrà contribuire a comprendere i determinanti chimico-fisici utilizzati dalle catene polipeptidiche per acquisire in modo spontaneo una conformazione unica, stabile e attiva dal punto di vista funzionale. Una visione generale di questi processi a livello molecolare è di vitale importanza per la progettazione di nuove proteine, cioè l'obiettivo finale dell'ingegneria proteica. Più recentemente, l'incorporazione nelle proteine di amminoacidi non naturali che mostrino proprietà chimico-fisiche o spettroscopiche peculiari si è rivelato un approccio nuovo e promettente nella chimica delle proteine. L'ingegneria proteica con amminoacidi non naturali infatti ha consentito ai ricercatori di modificare in maniera fine la struttura in un particolare sito, espandendo enormemente le opportunità della chimica fisica-organica nello studio delle proteine. A questo riguardo la sintesi peptidica su fase solida rimane il metodo più facile e veloce per incorporare in maniera sito-specifica e con alte rese, amminoacidi non naturali in polipeptidi anche lunghi (50-80 amminoacidi), avvicinandosi alle dimensioni reali delle proteine. In questa tesi di dottorato vengono presentati esperimenti di ingegneria proteica svolti sia mediante metodi genetici, che chimici.

Nell'affrontare questi studi, i quali hanno l'obiettivo di effettuare correlazioni struttura-funzione dell'α-trombina umana, si è rivelato indispensabile lo sviluppo di un sistema efficiente per la produzione di grandi quantità di proteina ricombinante, sia nella forma nativa che mutata. La trombina è una proteasi serinica che svolge un ruolo chiave nella cascata coagulativa. La trombina umana è una glicoproteina che consiste di due catene polipeptidiche, una catena B di 259 residui e una catena A più piccola di 36 residui, tenute insieme da un legame disolfuro. La pretrombina, il più piccolo precursore fisiologico della trombina a livello fisiologico è stato espressa in *E. coli*, un sistema procariotico semplice, economico e veloce rispetto ai sistemi eucariotici. Utilizzando questo sistema di espressione, si sono proposte alcune questioni al fine di condurre studi struttura-funzione sulla trombina. In primo luogo, sfruttando il fatto che *E. coli* non possiede il macchinario biochimico per coniugare catene di carboidrati alle proteine, abbiamo studiato il ruolo della glicosilazione sulla struttura, stabilità e funzione della proteina ricombinante *wild-type*. In secondo luogo, al fine di comprendere l'effetto di una mutazione naturale (i.e. Gly25Ser) posta all'interfaccia tra catena A e B sulla struttura e sulla funzione dell'enzima, abbiamo prodotto la forma ricombinante corrispondente. In terzo luogo, abbiamo prodotto due mutanti della trombina in cui residui di arginina chiave (Arg73 nell'esosito I, e Arg101 nell'esosito II) localizzate nei siti di

riconoscimento della trombina sono stati sostituiti con alanina, al fine di abrogare la capacità di *binding* di ciascun esosito. In ogni caso, le proteine ricombinanti espresse in *E. coli* vengono accumulate nei corpi di inclusione, dai quali sono state refoldate mediante rinaturazione accoppiata alla formazione dei ponti disolfuro, con rese significative per quasi tutte le specie, e pari a 10 mg di α-trombina umana pura e attiva per litro di coltura. Tutte le proteine mutanti sono state caratterizzate estensivamente rispetto alla loro identità chimica, alla loro conformazione, stabilità e funzione. Il risultato principale del nostro lavoro riguarda il fatto che l'enzima ricombinante *wild-type*, in cui la specie oligosaccaridica è assente, mostra proprietà conformazionali e funzionali identiche a quelle della trombina naturale, ma è significativamente meno stabile rispetto alla sua controparte naturale.

Un applicazione principale dell'ingegneria proteica con amminoacidi non naturali riguarda l'introduzione di particolari sonde spettroscopiche per studiare interazioni ligando-proteina e proteina-proteina. A questo proposito, la *o*-nitrotirosina (NT) assorbe la radiazione nel *range* di lunghezze d'onda dove Trp e Tyr emettono fluorescenza (300-450 nm), ed è essenzialmente non fluorescente. Perciò, la NT può funzionare come accettore in studi di *energy transfer* (FRET) per studiare interazioni ligando-proteina. Da qui, le potenzialità della NT sono state testate sul sistema irudina-trombina, una coppia inibitore-proteasi ben caratterizzata e di importanza farmacologia chiave. Noi abbiamo sintetizzato due analoghi del dominio N-terminale (1-47) dell'irudina: Y3NT, nel quale la Tyr3 è stata sostituita dalla NT, e S2R/Y3NT, contenente le sostituzioni Ser2→Arg e Tyr3→NT. Il *binding* di questi analoghi alla trombina è stato investigato a pH 8 tramite FRET e spettroscopia di assorbimento UV/Vis. In seguito a *binding* dell'irudina, la fluorescenza della trombina viene ridotta del 50% circa, a causa dell'*energy transfer* che avviene tra i residui di Trp dell'enzima, (donatori) e la singola NT dell'inibitore (cioè, l'accettore). I nostri risultati indicano che l'incorporazione di NT può essere effettivamente utilizzata per rivelare interazioni proteina-proteina con sensibilità nel *range* del basso nanomolare, per scoprire caratteristiche strutturali sottili all'interfaccia ligando-proteina e per ottenere valori di K_d affidabili per studi di correlazione struttura-attività.

Studi di *high throughput screening* di interazioni proteina-proteina rivestono un ruolo importante per applicazioni farmacologiche e biotecnologiche. In questo lavoro viene proposto l'uso di *o*-nitrotirosina e *p*-iodofenilalanina come sonde spettroscopiche in combinazione con tecniche di dicroismo circolare e *quenching* di fluorescenza, al fine di determinare l'orientazione peptidica in complessi ligando-recettore. Questo metodo è stato testato con successo su un dominio SH3 da miosina di lievito, che riconosce in maniera specifica peptidi di classe-I. Le strategie chimiche evidenziate in questo lavoro sottolineano la vasta applicabilità di amminoacidi non naturali in *screening* biotecnologici e farmacologici.

Abstract

In the past three decades, the advent of recombinant DNA technology allowed the site-specific alteration of a given polypeptide chain at a glance, thus much expanding the tools available to study the molecular mechanisms of protein folding, stability, and function. This approach, also known as protein engineering, is based on the possibility of modifying the chemical composition of a protein at a single or multiple sites with the 20 available coded amino acids, thus obtaining recombinant mutant proteins with altered structure, function, or stability properties. Evaluation of the effects of the mutation on the property under investigation (e.g., folding, stability or molecular recognition) will allow quantitative estimation of the contribution of that specific amino acid which has been mutated in the wild-type protein. This information will in turn help in understanding the physico-chemical determinants exploited by natural polypeptide chains to spontaneously acquire a unique, stable and functionally active conformation. A comprehensive view of these fundamental processes at the molecular level is of paramount importance in the successful design of novel proteins, which is the ultimate goal of protein engineering. More recently, the incorporation of noncoded amino acids, possessing unique physico-chemical or spectroscopic properties, into proteins has emerged as a novel and promising approach in protein science. Protein engineering with non-coded amino acids, in fact, allow investigators to finely tune the structure at a protein site, thus much expanding the scope of physical-organic chemistry in the study of proteins. With respect to this, stepwise solid-phase chemical synthesis remains the easiest and fastest approach to site-specifically incorporate in high yields noncoded amino acid into even long (50-80 amino acids) polypeptide chains, approaching the size of real proteins. In this doctoral Thesis, relevant applications of protein engineering experiments by both genetic and chemical methods will be presented.

During our studies aimed to dissect the structure-function relationships of human α-thrombin, we approached the task of devising an efficient system to produce large amounts of recombinant protein, either in native or mutated form. Thrombin is a serine protease that plays a pivotal role in haemostasis. Human thrombin is a glycoprotein consisting of two polypeptides, a 259-residue B-chain and a smaller 36-residue A-chain, connected by a disulfide bond. B-chain contains the catalytic triad of thrombin and three disulfide bonds. Hence, prethrombin-2, the smallest physiological single-chain precursor of α-thrombin, was expressed in E. coli, a prokaryotic system which is easy to work with, to scale up, and less time-consuming than eukaryotic systems. Using this expression system, we addressed several issues in structure-function studies on thrombin. Firstly, taking advantage of the fact that E. coli lacks the biochemical machinery for conjugating carbohydrate chains to proteins, we have studied the role of glycosylation on the structure, stability and function of the wild-type recombinant protein. Secondly, with the aim to understand the effect of natural mutations (i.e., desLys9a and Gly25Ser) in the thrombin A-chain on the structure and function of the

enzyme, we have produced the corresponding recombinant forms of the naturally occurring variants of thrombin. Thirdly, we produced two mutants of human thrombin in which key Arg-residues (i.e., Arg73 in exosite I and Arg101 in exosite II) in the positively charged exosite I and II binding sites were replaced by Ala, in order to abrogate ligand-binding at each exosite. In all cases, the recombinant proteins accumulated in the inclusion bodies, from which disulfide-coupled renaturation was achieved in significantly high yields in almost all cases, yielding about 10 mg per liter of fully active wild-type human α-thrombin. All mutant proteins were subjected to thorough characterization with respect to their chemical identity, as well as conformational, stability and functional properties. A major finding of our work was that the recombinant wild-type enzyme, lacking the carbohydrate moiety, has conformational and functional properties indistinguishable from those pertaining to the natural thrombin, but it is significantly less stable than the natural counterpart.

A major application of protein engineering with noncoded amino acids entails the incorporation of suitable spectroscopic probes for studying ligand-protein and protein-protein interactions. With respect to this, 3-nitrotyrosine (NT) in absorbs radiation in the wavelength range where Tyr and Trp emit fluorescence (300-450 nm) and it is essentially nonfluorescent. Therefore, NT may function as an energy acceptor in resonance energy transfer (FRET) studies for investigating ligand-protein interactions. Here, the potentialities of NT were tested on the hirudin-thrombin system, a well-characterized protease-inhibitor pair of key pharmacological importance. We synthesized two analogues of the N-terminal domain (residues 1-47) of hirudin: Y3NT, in which Tyr3 was replaced by NT, and S2R/Y3NT, containing the substitutions Ser2→Arg and Tyr3→NT. The binding of these analogues to thrombin was investigated at pH 8 by FRET and UV/Vis-absorption spectroscopy. Upon hirudin binding, the fluorescence of thrombin was reduced by ~50%, due to the energy transfer occurring between the Trp-residues of the enzyme (*i.e*, the donors) and the single NT of the inhibitor (*i.e.*, the acceptor). Our results indicate that the incorporation of NT can be effectively used to detect protein-protein interactions with sensitivity in the low nanomolar range, to uncover subtle structural features at ligand-protein interface, and to obtain reliable K_d values for structure-activity relationships studies.

High throughput screening of protein-protein and protein-peptide interactions is of high interest both for biotechnological and pharmacological applications. Here, we propose the use of the non-coded amino acids *o*-nitrotyrosine and *p*-iodophenylalanine as spectroscopic probes in combination with circular dichroism and fluorescence quenching techniques (i.e., collisional quenching and resonance energy trasfer) as a mean to determine the peptide orientation in protein-peptide complexes. The method was successfully tested on an SH3 domain from a yeast myosin which is known to recognize specifically class-I peptides. The chemical strategies outlined here highlights the broad applicability of noncoded amino acids in biotechnology and pharmacological screening.

1. Introduction

CHAPTER 1.1

Protein Engineering

1.1.1. Protein Engineering by Genetic Methods

Recombinant DNA technology has made it possible to produce large quantities of previously rare and unknown proteins, an in a second time to engineer specifically altered or new and novel protein molecules possessing tailored chemical and biological characteristics. Termed protein engineering, the deliberate design and construction of unique proteins with enhanced or novel molecular properties is a result of specifying the exact amino acid sequence (protein primary structure) of that protein (Richardson and Richardson, 1990, Cleland and Craik, 1996).

The primary structure affects the protein's conformation. The conformation of each and every amino acid component present in the protein influences the protein's complex three-dimensional structure. The conformational preference of the protein chain residues determines the protein's secondary structure including α-helices and β-sheets or reverse turns. The local secondary structures are folded into three-dimensional tertiary structures made up of domains. The domains are not only structural units, but are also functional units often containing intact ligand binding (in a receptor) or enzyme catalytic sites. Thus, protein engineering provides an approach to modify a native protein's structure specifically or to create a unique, new protein with a particular structure. Protein engineering has numerous powerful theoretical and practical implications for examining protein structure and function, probing enzyme mechanisms, investigating protein folding and conformation, studying protein stability, introducing detectable groups into proteins as an analytical tool, and producing improved second generation tailored biopharmaceuticals.

Site-directed mutagenesis (also called site-specific mutagenesis) is a protein engineered technique allowing specifically (site-direct) alteration (mutation) of the primary amino acid sequence of proteins to create new chemical entities, by change the base sequence of relative genes. Until the introduction of site-directed mutagenesis (Hutchinson, 1978), and subsequently protein engineering by the manipulation of genes, protein scientist were just observers and users of proteins. But by changing first individual amino acid residues and then

whole segments of proteins, precise structure-function activity experiments could be designed and new functional protein made.

Several different types of mutagenesis methods are now available in molecular biology to understand the contribution of specific amino acids to the structure and function of proteins. Saturation mutagenesis is used to generate mutations at many sites in an unbiased fashion and includes chemical mutagenesis, cassette mutagenesis, misincorporation mutagenesis, and spiked oligonucleotide primers. On the other side, oligonucleotide-directed mutagenesis is used to test with exquisite specificity and extraordinary breadth the role of particular residues in the structure, catalytic activity, and ligand-recognition of a protein. This method took advantage when pioneering work in DNA chemistry led to the routine synthesis of oligonucleotides. The classic method used synthetic oligonucleotides that were completely homologous to single-strand virus M13 DNA except for a single base change, and *E. coli* DNA polymerase. Most of the mutagenesis methods in current use are based on polymerase chain reaction (PCR). Early investigators showed in fact that single-bases mismatches in hybrids between oligonucleotide primers and the target DNA did not affect the efficiency of amplification, and that these mismatches could be converted into mutations located in the primer regions of PCR product. PCR-based methods offered important advantages, like ability to use double-stranded DNA templates, high rate of recovery of mutants and availability of commercial kits (Sambrook and Russel, 2001).

Site-specific mutagenesis was initially used to mutate genes that encode enzymes with the aim of dissecting the mechanism of catalysis, as in the case of tyrosyl-transfer RNA sinthetase (Winter et al., 1982) and serine proteases (Craik et al., 1987). Figure 1 suggests possible theoretical mutations of the active site of a model serine protease enzyme that could be engineered to probe the mechanism of action of the enzyme. Craik and co-workers have actually tested the role of the aspartic acid residue in the serine protease catalytic triad Asp, His, and Ser. They replaced Asp102 (carboxylate anion side chain) of trypsin with Asn (neutral amide side chain) by site-directed mutagenesis and observed a pH depended change in the catalytic activity compared to the wild-type parent serine protease (see Figure 1).

Applications of this approach involve also the improvement of the activity, affinity and specificity of useful biopharmaceuticals. Site-directed mutagenesis can alter, and increase the stability of the “wild type” parent protein by replacing reactive amino acids in the parent molecules with residues that resist reaction. Also, physical stability improves with the introduction of amino acids that contribute to a protein’s conformational stability (i.e., introduce site mutations that create new intramolecular hydrogen bonds or disulfide bridges).

1.1.2. Protein Engineering by Chemical Methods

During the past decade, the incorporation of noncoded amino acids into proteins has emerged as a novel and promising approach in protein science. In its infancy, the approach of protein engineering was essentially restricted to the possibility to chemically modify, in a rather unspecific fashion, particular amino acid side chains in proteins (Freedman et al., 1971). More recently, the advent of recombinant DNA technology allowed the site-specific alteration of a given polypeptide chain at a glance, thus much expanding the tools available to study the molecular mechanisms of protein folding, stability, and function (Ferscht and Winter, 1992). Nevertheless, a quantitative description of the physical and chemical basis that makes a polypeptide chain to efficiently fold into a stable and functionally active conformation is still elusive.

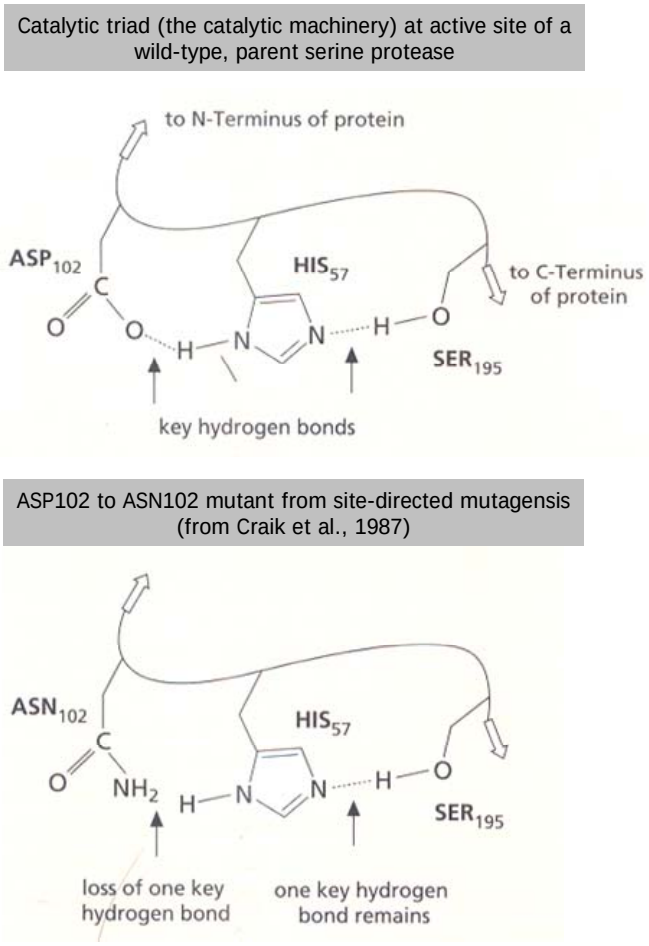


Fig. 1: A possible site directed mutation of the amino acids composing the catalytic triad of a seine protease.

This mainly originates from the fact that nature combined, in a yet unknown manner, different properties (*i.e.*, hydrophobicity, conformational propensity, polarizability, hydrogen

bonding capability, etc.) into the 20 standard protein amino acids, thus making difficult, if not impossible, to univocally relate the change in protein stability or function to the variation of physico-chemical properties caused by amino acid exchange(s). In this view, incorporation of non-coded amino acids with tailored side chains, allowing investigators to finely tune the structure at a protein site, would facilitate to dissect the effects of a given mutation in terms of one or a few physico-chemical properties, thus much expanding the scope of physical-organic chemistry in the study of proteins (Cornish et al., 1994).

Incorporation of noncoded amino acids has been widely exploited in the study of bioactive peptides (Hogan 1997). However, the results of these studies provide only a qualitative picture of the mechanism of ligand-receptor interaction, and thus are of limited predictive power, as also documented by the explosion in the last few years of serendipity-based approaches in peptide design and drug-discovery. These difficulties stem primarily from the fact that introduction of single or multiple amino acid exchanges into a short peptide is expected not only to alter the interaction energy at the binding site(s) of the receptor, but also to affect (in a yet unpredictable way) fundamental properties of the free ligand, including electrostatic potential, hydration energy and conformational entropy. Contrary to short peptides (5-20 amino acids), longer polypeptide chains (> 50 amino acids) are generally characterized by a well-defined and stable 3D structure, thus serving as macromolecular scaffolds, by keeping the global properties of the molecule rather constant and allowing the changes in binding free energy to be related in a more predictable way to the local variations of the physico-chemical properties at the mutation site(s). Hence, the incorporation of noncoded amino acids into proteins represents a further, (almost) obligatory extension of the studies aimed at elucidating the relationships existing between the structure, stability, and function in proteins (*i.e.*, protein engineering).

In this view, several different strategies have been pursued to incorporate non-coded amino acids, including peptide synthesis (Kent 1988), native chemical ligation (Dawson and Kent, 2000), enzyme-catalyzed semisynthesis (Wallace, 1993), biosynthetic incorporation *via* auxotrophic bacterial strain expression (Wong and Eftink, 1997), and nonsense suppression methodologies in cell-free or whole-cell (Cropp and Schultz, 2004, Dougherty, 2000) expression systems. Both native chemical ligation and enzymatic semisynthesis require careful tailoring of the experimental procedures, while biosynthetic incorporation methods are not site-specific and are restricted to those cases when the structure of the noncoded amino acid (e.g., *p*-fluorephenylalanine, selenomethionine, 7-azatryptophan) is similar to that the natural counterpart to be replaced. On the other hand, despite significant advancements during

the last decade, genetic methods are still limited by exceedingly low amounts of the resulting mutant protein, usually micrograms, that impair a thorough structural and functional characterization. However, stepwise solid-phase chemical synthesis remains the easiest and fastest approach to site-specifically incorporate in high yields any noncoded amino acid into even long (50-80 amino acids) polypeptide chains approaching the size of real proteins (Albericio, 2004).

A major application of protein engineering with noncoded amino acids regards the introduction into proteins of biophysical probes possessing physico-chemical properties (*e.g.*, side-chain volume, hydrophobicity) similar to those of the corresponding natural amino acids, but spectral features distinct from those of the natural counterparts and highly sensitive to the chemical environment in which the probe is located (Cornish et al., 1994). Hence, by the use of the so-called "spectrally enhanced proteins" (Wong and Eftink, 1997) it should be possible to effectively monitor the local structure and dynamics of the mutated protein during key events, such as protein folding and denaturation or ligand binding, without significantly perturbing the kinetics and equilibrium properties of the process under investigation. With respect to this, in a recent study Cohen and co-workers were able to site-specifically introduce 6-dimethylamino-2-acyl-naphthylalanine (Aladan) into the B1 domain of staphylococcal protein G to obtain estimates of the local dielectric constant of the protein at different sites (Cohen et al, 2002).

More specifically, the development of new spectroscopic tools for studying protein-protein interactions is central to many disciplines, including structural biology, biotechnology, and drug discovery. Traditionally, the change in tryptophan (Trp) fluorescence has been exploited to study ligand-protein interactions (Eftink 1997). However, the fluorescence signal of many proteins is insensitive to ligand binding, since fluorescence changes are mostly restricted to those cases where Trp-residues are embedded in the ligand-protein interface or when the ligand binding induces conformational changes in the protein, remote from the binding region and involving one or more Trp-residues. Furthermore, the presence of multiple tryptophans in proteins may lead to compensating effects that often complicate a unique structural interpretation of the fluorescence data (Lakowicz 1999).

To overcome these problems, several extrinsic spectroscopic probes, characterized by well-defined spectral properties, have been covalently bound to protein functional groups (*i.e.*, Cys and Lys), to act as energy donors or acceptors in fluorescence resonance energy transfer (FRET) studies (Wu and Brand, 1994). This approach, however, is limited by possible labeling heterogeneity, non-quantitative modification, structural alteration of the

proteins resulting from the labeling *per se*, and perturbation of the binding process, due to the large size of the fluorescent labels used and the incorporation of non-natural amino acids by solid-phase peptide synthesis still remains the major approach.

CHAPTER 1.2

Thrombin Structure and Function

Thrombin is the central protease of the vertebrate blood coagulation cascade. Prior to the advent of modern biochemistry, thrombin was identified as a substance capable of promoting the formation of a fibrous blood clot (fibrin), and it accordingly derives its name from thrombos, the Greek word for clot. Early empirical observations indicated that thrombin could be antagonized by an inhibitory substance in the blood (antithrombin), and generated when its precursor (prothrombin) was combined with lime salts (calcium). Prothrombin, in turn, was believed to be generated by another substance (thrombogen) in the presence of an activator (thrombokinase), both of which were absent until blood was shed. Although such a conception of blood coagulation may seem quaintly simplicistic in retrospect, particularly with regard to the nature of thrombogen and thrombokinase, it has served as the foundation for our current understanding of the blood coagulation cascade.

The haemostasis (blood clotting) is initiated by vascular damage to the blood vessel wall, exposing the subendothelial tissue to the blood. The subendothelial surface is rich in collagen, which binds to and localises platelets, and in tissue factor. When tissue factor (TF), a transmembrane protein, becomes exposed to circulating blood, it binds to factor VII (FVII) or FVIIa. The TF-FVIIa complex initiates a series of reactions that lead to the formation of the prothrombinase complex (Figure 2) (consisting of FVa, FXa, Ca^{2+}, and anionic phospholipids) and subsequent proteolytic conversion of prothrombin to thrombin (Davie et al., 1991).

Functions of thrombin

Thrombin acts on a variety of substrates to promote haemostasis, the best known of which is fibrinogen. Fibrinogen circulates in the plasma as a dimer of three chains, $(A\alpha B\beta\gamma)_2$, that are covalently linked by disulfide bonds. Thrombin attacks the Aα chain first to release fibrinopeptide A and to generate fibrin I monomer. Fibrin then noncovalently assembles into protofibrils that are attacked by thrombin at the Bβ chains, leading to the release of fibrinopeptide B and the assembly of thick fibers forming the scaffolds of the ensuing clot

(Mosesson, 2005). Thrombin converts the inactive form of FXIII to FXIIIa, and this transglutaminase imparts structural integrity to the fibrin network by a covalent cross-linking of glutamine and lysine residues on adjacent fibrin molecules. Clot stability is further enhanced in an indirect manner by thrombin-mediated proteolytic conversion of thrombin-activatable fibrinolysis inhibitor (TAFI) to its active form (TAFIa), a carboxi-peptidase that removes lysine residues from fibrin that are recognized by fibrinolytic enzymes. In addition, thrombin facilitates positive feedback amplification of the pathway, leading to its own generation by proteolitically converting FXI to FXIa (a serine protease of the intrinsic pathway), as well as FVIII and FV (cofactors in the generation of FXa and thrombin, respectively) (Davie et al., 2006).

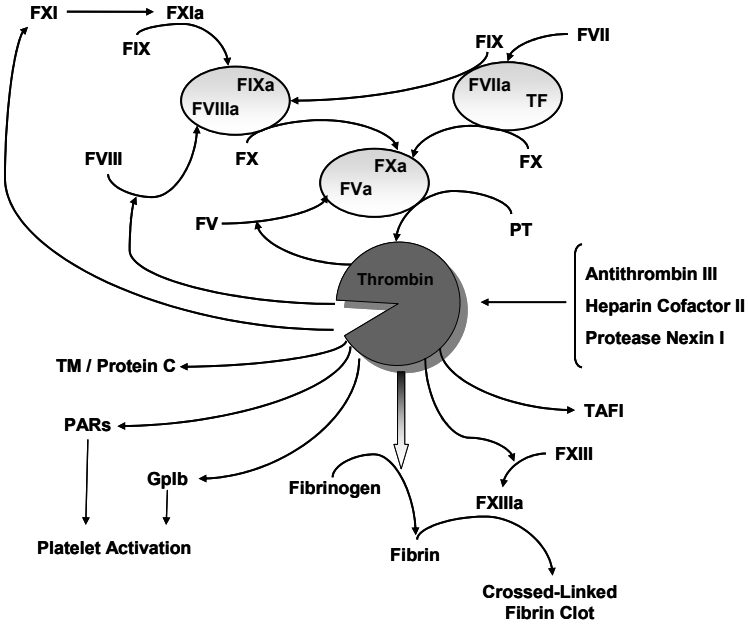


Fig. 2: The central role of thrombin in haemostasis.

In addition to its role in generating the cross-linked fibrin clot, thrombin promotes platelet activation, leading to platelet aggregation, degranulation, and surface expression of procoagulant phospholipids such as phosphatidylserine. These cellular effects are triggered by thrombin cleavage of protease-activated receptors (PARs), which are members of the G-protein-coupled receptors superfamily (Vu et al., 1991). These receptors share the same mechanism of activation: thrombin cleaves at a specific site within the extracellular N-terminus exposing a new N-terminal tethered ligand domain that binds and activates the cleaved receptor. Two PARs have been identified to be involved in haemostatic process. PAR-1 is responsible for platelet activation in humans at low thrombin concentration and its action is reinforced by PAR-4 at high enzyme concentration. Furthermore, thrombin promotes

many of the hallmarks of platelet activation through its association with the platelet surface glycoprotein GPIbα (Figure 3), which is capable of acting as a cofactor in the proteolytic activation of PAR-1 by thrombin, but which may also mediate platelet activation by thrombin in a manner independent of thrombin's catalytic activity (De Candia et al., 2001).

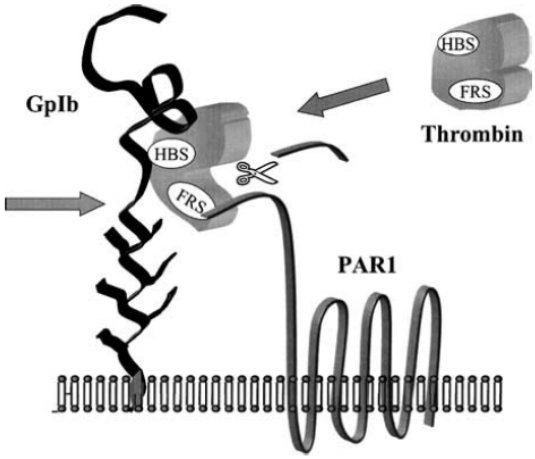


Fig. 3: Cartoon showing the interaction of thrombin with both Gp1b and PAR-1 on the platelet membrane. HBS, heparin binding site; FRS, fibrinogen recognition site.

Despite the wide variety of procoagulant substrates activated by thrombin, not all effects of thrombin promote the thrombus formation. Upon contact with thrombomodulin, which is an integral membrane protein present on the membrane of intact vascular endothelial cells, a region of thrombin critical for the recognition of its procoagulant substrates is obscured, therefore the ability of thrombin to cleave fibrinogen and PAR-1 is suppressed (Esmon, 1989). However, the specificity of the enzyme toward the zymogen protein C is enhanced greater than thousand-fold. The reaction is further enhanced by the presence of a specific endothelial cell protein C receptor. Activated protein C cleaves and inactivates FVa and FVIIIa, two essential cofactors of coagulation, FXa and FIXa that are required for thrombin generation, thereby down-regulating both the amplification and progression of the coagulation cascade. Hijacking of thrombin by thrombomodulin and activation of protein C in the microcirculation constitute the natural anticoagulant pathway that prevents massive intravascular conversion of fibrinogen into a soluble clot upon thrombin generation .

In addition to its direct role in promoting and regulating clot formation, thrombin elicits numerous cellular responses in cells other than platelets, many of which can be ascribed to its activation of PARs. Thrombin stimulates a variety of responses by endothelial cells, including cell surface expression and secretion of cellular adhesion molecules as well as the production

of growth factors and cytokines. Thrombin also promotes cytokine elaboration by smooth muscle cells and stimulates the proliferation of both smooth muscle cells and fibroblasts.

Efficient regulation of thrombin activity is essential to prevent excessive or improperly localized clot formation. Certain members of the serine protease inhibitor (serpin) superfamily inhibit the catalytic activity of thrombin, and this reaction is greatly accelerated in the presence of glycosamminoglycans such as heparin, heparan sulphate, and dermatan sulphate. The most important serpins responsible for thrombin inhibition are antithrombin (originally called antithrombin III), heparin cofactor II (HCII), and protease nexin I (PNI). Antithrombin is abundant in the blood (2.3 μM), inhibits a variety of serine proteases involved in blood coagulation, and appears to be the main inhibitor limiting intravascular clot formation. In contrast, HCII is highly specific for thrombin and appears to regulate thrombin activity in extravascular tissues following vascular injury, whereas PNI is likely to inhibit thrombin at or near the surface of a variety of cell types. In addition to this endogenous thrombin inhibitors, several potent thrombin inhibitors have been isolated from hematophagous organisms, including hirudin and haemadin from the leeches *Hirudo medicinalis* and *Haemadipsa sylvestris,* respectively.

Prothrombin biosynthesis

Thrombin is derived from its inactive zymogen form, prothrombin, a 70-kDa glycoprotein that is synthesized in the liver and secreted in the blood. The amino acid sequences of blood and human thrombin were determined in the mid 1970s (Butkowski et al., 1977), and later confirmed by determination of the sequence of their corresponding cDNAs and genes (Degen and Davie, 1987). The primary sequence of human prepro-prothrombin consists of several discrete functional units: a prepro leader sequence, a γ-carboxiglutamic acid (Gla) domain of ~ 40 residues, two kringle domain consisting of ~ 80 residues each, and a carboxi-terminal serine protease domain. Prothrombin undergoes a variety of posttranslational modifications prior to its conversion to thrombin. The signal peptide is removed from prepro-prothrombin by signal peptidase following import of the nascent polypeptide into the endoplasmatic reticulum (ER). The resulting molecule, pro-prothrombin, is then modify by γ-glutamyl carboxylase, an enzyme present in the ER that recognizes the propeptide sequence and catalyzes the vitamin K-dependent conversion of 10 glutamate residues to γ-carboxyglutamate (Gla) residues within the ~ 40 residue Gla domain. This essential modification confers upon the Gla domain the ability to coordinate Ca^{2+} ions and adopt a conformation required for

binding to anionic phospholipids surfaces at sites of vascular injury and on activated platelets. γ-Glutamyl carboxylase uses a reduced form of vitamin K (vitamin K hydroquinone) as a cofactor in the conversion of Glu residues to Gla, and in the process oxidizes vitamin K hydroquinone to vitamin K 2,3-epoxide. The epoxide form is then reduced back to the hydroquinine form by the sequential activities of vitamin K epoxide reductase, both of which are potently inhibited by warfarin, a widely used anticoagulant. The impairment of this series of reactions, collectively referred to as the vitamin K cycle, results in a decrease in the plasma level of functional carboxylated prothrombin (Davie et al., 2006).

Following γ-glutamyl carboxylation, the propeptide is removed by a furin-like proprotein convertase to generate the amino terminus of the mature zymogen. Both bovine and human prothrombin contain three N-linked carbohydrate chains, two of which are located in kringle domain 1 and one of which is present within the serine protease domain. Following the attachment of N-linked carbohydrates, prothrombin is secreted into the blood where it is found at a concentration of ~ 1.2 μM.

The covalent structure of thrombin

During activation of the coagulation cascade, prothrombin is converted to thrombin by prothrombinase, a complex consisting of the serine protease, FXa, the cofactor FVa, anionic phospholipids on the surface of activated platelets, and calcium ions. In this reaction, FXa initially cleaves prothrombin after Arg320 (prothrombin numbering) to generate meizothrombin (Figure 4). This intermediate subsequently is cleaved after Arg 271 to liberate thrombin and fragment 1.2, the latter consisting of the Gla domain and two tandem kringle domain. The removal of fragment 1.2 generates a thrombin molecule, with a molecular weight (36 kDa) that is approximately half that of prothrombin. The proteolytically active thrombin molecule consists of a light chain of 49 residues linked by a single disulfide bond to the heavy chain of 259 residues that contain three intrachain disulfide bonds. In contrast to bovine thrombin, human thrombin undergoes additional proteolyisis of the Arg13-Thr bond within the light chain, removing a 13-residue N-terminal peptide and resulting in a light chain that is 36 residues in length in α-thrombin (Downing et al., 1975). It is not known whether the former, latter, or both forms of thrombin are present at the site of vascular injury.

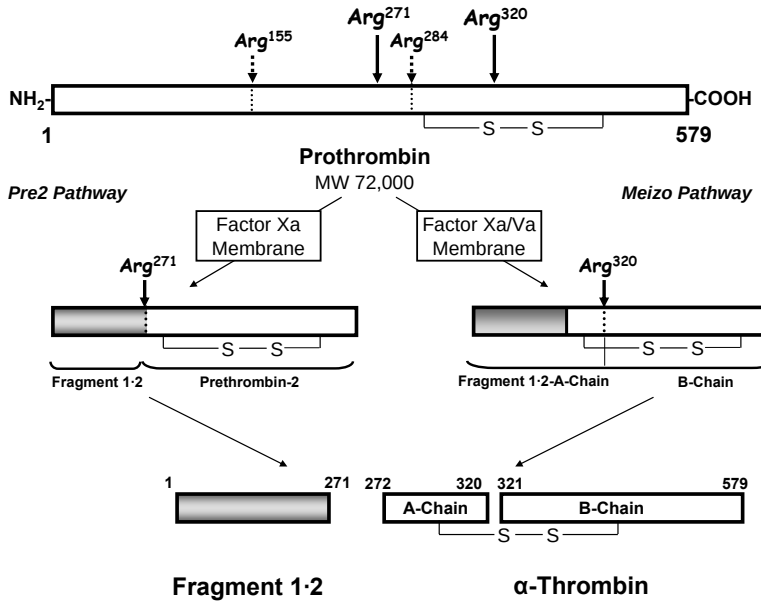


Fig. 4: Scheme of the two pathways for the activation of human prothrombin. Two factor Xa-catalyzed cleavages convert prothrombin to thrombin. Initial cleavage at Arg271 by membrane-bound factor Xa in the absence of factor Va results in the generation of fragment 1·2 and prethrombin 2. Subsequent cleavage at Arg320 results in thrombin formation (*pre2 pathway*). Initial cleavage at Arg320 by prothrombinase (factor Va bound to factor Xa on a membrane surface in the presence of Ca^{2+}) results in the production of an enzymatically active intermediate (meizothrombin). Cleavage of this intermediate at Arg271 produces thrombin and fragment 1·2 (*meizo pathway*). Thrombin also cleaves prothrombin at Arg155 and at Arg284 (*dashed arrows*).

Given the abundance of serpins in the blood, it is unlikely that any given molecule of thrombin remains in an active form for long in circulation. However, non physiologic autolytic degradation products of thrombin may be generated ex-vivo (Figure 5) (Fenton et al., 1977). Upon prolonged standing in solution, human α-thrombin is first autolytically converted to β-thrombin by the cleavage of the Arg75-Tyr76 and Arg77a-Asn78 bonds (chymotrypsinogen numbering is used throughout) in the heavy chain. β-thrombin is further converted to γ-thrombin by the additional cleavage of the Lys149e-Gly150 bond. The molecules are functionally distinct from α-thrombin. Although the catalytic activities of α-, β- and γ-thrombin are roughly similar with respect to the hydrolysis of small chromogenic substrates, β- and γ-thrombin are significantly less active toward physiologic macromolecular substrates such as fibrinogen.

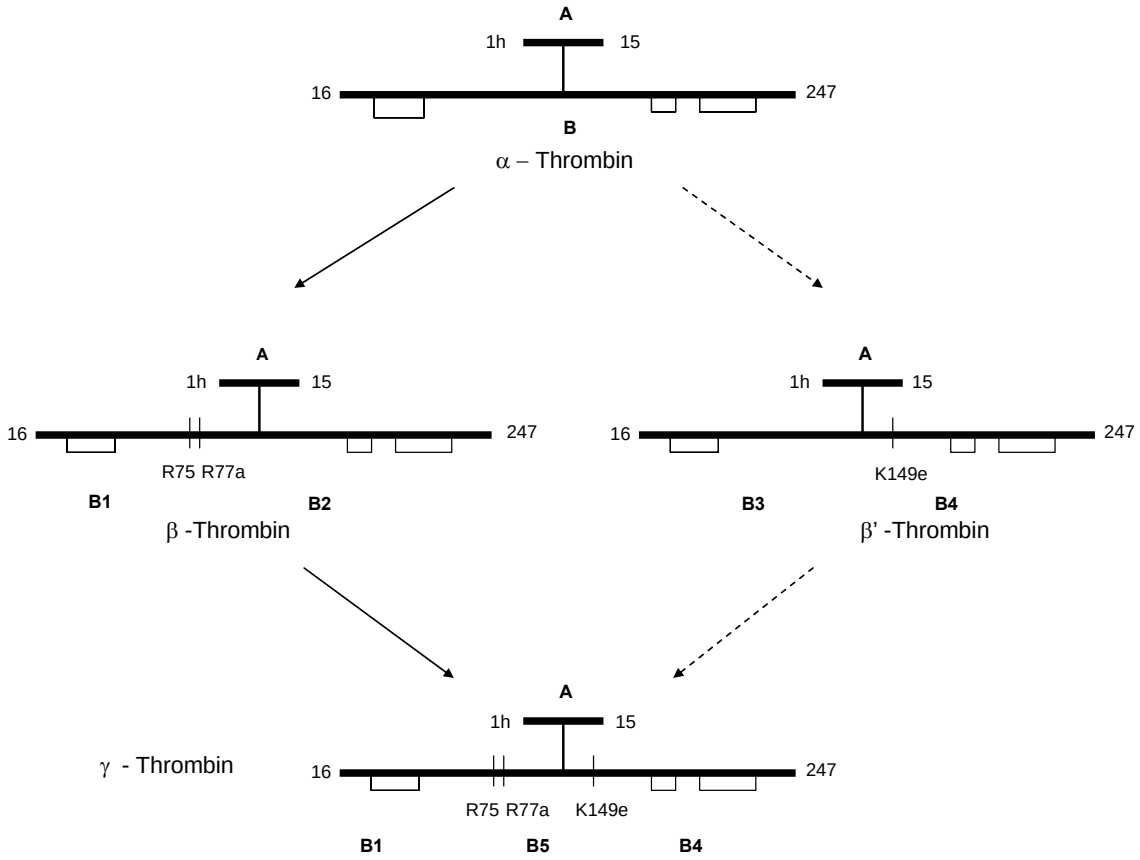


Fig. 5: Autolysis of human thrombin. α-thrombin is converted to β-thrombin by proteolysis after residues R75 or R77a. β-thrombin is subsequently converted to γ-thrombin by proteolysis after residues K149e. For reference, the locations of disulfide bonds (thin black lines) are shown.

Thrombin structure

Human α-thrombin is a serine protease belonging to the chymotrypsin family, with which show a high degree of similarity (~49%). Most notable, in the alignment of the two sequence, is the absolute conservation of the residues that constitute the catalytic apparatus, His57, Asp102, and Ser195. Thrombin sequence numbering follows the numeration of chymotrypsinogen. Since thrombin bears some insertion residues that do not find a counterpart in the prototypic chymotrypsin sequence, they are indicated in the so-called chymotrypsinogen numbering system by a number followed by a letter, that expresses the relative position along the primary sequence of the enzyme (Bode et al., 1992). This numbering system is valid only for the mature active molecule. In Appendix A, the four numeration used for thrombin are compared. In this thesis, the chymotrypsinogen numbering system will be used, whereas differently indicated.

We have previously described that α-thrombin consists of a 36-residue A- (light-) chain and a 259-residue B- (heavy-) chain covalently linked by a disulfide bridge between Cys1 and Cys122. The B-chain is organized into two adjacent barrels, similar to trypsin and the catalytic domains of related serine proteases. The active-site residues Ser195, His57 and Asp102 are present at the junction between the two barrels and the active-site cleft extends across both barrels perpendicular to this junction. The mechanism of peptide bond hydrolysis by serine protease has been worked out in great detail and involves the precise alignment of the scissile bond of the substrate [the so-called P1-P1' bond after the nomenclature of Schechter and Berger, (Schechter and Berger, 1967), where substrate residues are numbered sequentially to either side], with the catalytic residues of the proteases (Figure 6).

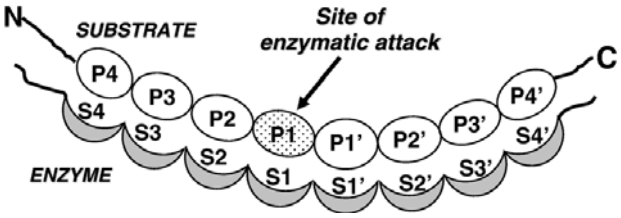


Fig. 6: Scheme of the interaction of a polypeptide substrate at the active site of an enzyme. At least an eight-residue segment of a polypeptide chain interacts with its side-chain residues (P) at a series of subsites (S) of the enzyme. Binding requires a specific stereochemical adaptation of the substrate, usually in an extended conformation. The P1 side-chain residue interacting with the S1 binding site of the enzyme is the major determinant for the enzyme specificity, but it can be not unique. Usually, proteases cleave at the carboxyl-terminal side of the scissile bond.

Proteolysis occurs trough nucleophilic attack by the side γ oxygen of Ser195 on the main carbonyl carbon of the P1 residue of the substrate. This reaction requires the potentation of Ser195 Oγ by the catalytic diad of Asp102 and His57. The resulting tetrahedral transition state is stabilized by the so-called oxyanion hole, formed by the main chain amides of Gly193 and Ser195, and eventually collapses to the acyl enzyme intermediate. After diffusion of the P1' residues out of the active site cleft, a molecule of water is potentiated by the catalytic diad for attack of the acyl enzyme to release the P end and regenerate the active protease. One of the most important factors determining the rate of hydrolysis of a substrate, and thus specificity, is the complementarity of the substrate peptide sequence with the properties of the active site cleft of the protease.

We said that thrombin bears some insertion loops on its surface, either in the A- and B-chain. Compelling evidence has demonstrated that the insertion loops of the B-chain

contribute to regulate the catalytic function of the enzyme. In particular, the 60- and the 148-insertion loops in B-chain (named according to typical or central residues contained) block the access of many macromolecular substrates or inhibitors to the catalytic residues, which is a primary cause for the narrow specificity of α-thrombin.

The boomerang-shaped A-chain lies in the groove of the B-chain in the back of the molecule, opposite the active site (Figure 7). It's organized mainly in a multiple-turn and partly helical conformation, making a smooth contour of the B-chain. The A-chain is topologically similar to the activation peptide of chtmotrypsinogen and connected in a similar manner through the Cys1-Cys122 dusilfide bridge to the B-chain. Most of the A-B chain interaction involve charged side-chains; six are buried salt bridges, and 10 are involved in interchain hydrogen bonds. Stabilization of the A-chain also occurs, mostly via polar and salt bridge interaction. The N-terminal segment of the A-chain up to Glu1c is characterized by relatively weak electron density in the –Phe-Pro-Arg-methyl ketone-thrombin crystal (Bode et al., 1992), and seems to have a high degree of conformational flexibility. The C-terminal segment up to Tyr14j is an amphiphilic α-helix forming one and one-half turns and showing a high degree of flexibility as well. The central region is the most rigid segment of the A-chain and contains strong salt-bridge, such as that involving the Ap1a-Lys9 side chains, which is characterized by a high electrostatic energy (-1.8 kcal/mol), significantly contributing to confer the boomerang-like shape to the chain. This chain may contribute to the stability of the entire thrombin molecule, and recently a renewed in A-chain relative to function and allosteric properties of thrombin has been growing up (De Cristofaro et al., 2004, Di Cera et al., 2008).

The thrombin structure has unique features that control specificity, such as loops and charged patches surrounding the pocket of the active site. A particularly striking feature of thrombin molecule is its pronounced uneven charge distribution, resulting in localized highly positive and negative electrostatic fields on the thrombin surface. Three contiguous surface patches, two positive and one of negative potential, have received particular attention (Figure 8). The negatively charged region extends around the active site of thrombin, with D189 and E192 located at the bottom of and at the entrance to the S1 pocket, respectively.

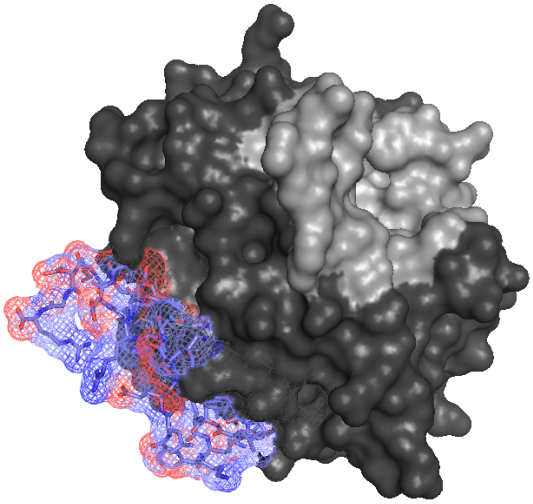


Fig. 7: Surface representation of the interface between thrombin heavy-chain (surface) and the light chain (sticks). In light gray the active-site cleft of thrombin B-chain is highlighted.

The first positively charged patch extends along the convex thrombin surface, where the active-site cleft levels off. Due to its interaction with fibrin(ogen), it had been designated as fibrin(ogen) recognition exosite or most recently as anion binding exosite I. The second positively charged surface patch, with an even stronger positive electrostatic field, extends from the intermediate helix toward the C-terminus of thrombin. Due to its interaction with heparin, this positively charged patch has been designated heparin binding site or anion binding exosite II.

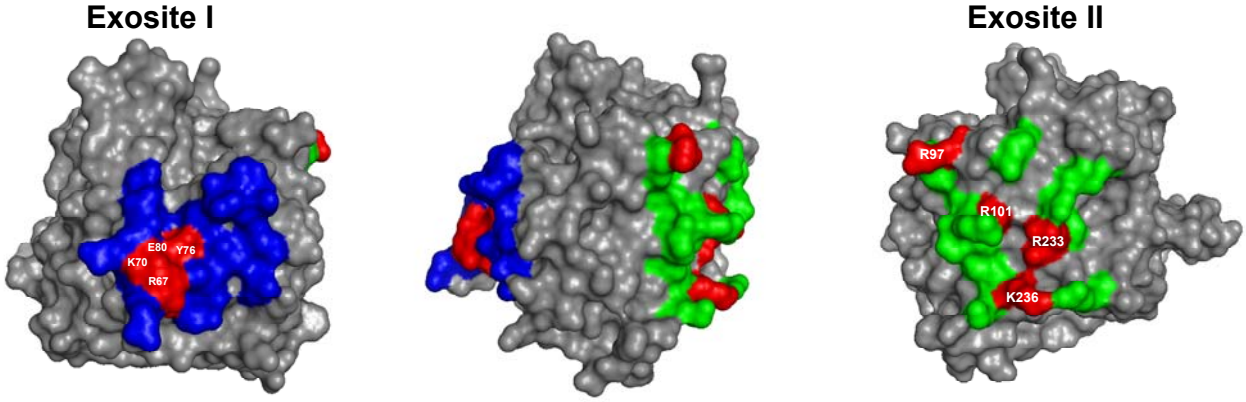


Fig. 8: Surface representation of the thrombin exosite I (on the left hand) and exosite II (on the right hand). The molecule with the exosite II highlighted is rotated by 180° in the sagittal plane with respect to molecule with the exosite I highlighted (PDB code: 1PPB, Bode et al., 1992).

These electrostatic patches create strong electrostatic surface potentials used by the thrombin molecule to select and pre-orient substrates and inhibitors. Both exosites

differentially influence the active site and mutually affect the binding properties of the opposite exosite.

Active-site cleft and direct substrate/inhibitor interaction

The canyon-like active site cleft is closed on top by the 60 loop, which forms an exposed hydrophobic lid to the active site. The Y60a-P60b-P60c-W60d sequence is unique to thrombin and is conserved in all thrombin sequences known to date. The ridge-like architecture of the YPPW peptide may represent the chemotactic domain of thrombin (Stubbs and Bode, 1993). The 60 loop opposes the five-residue 148 insertion loop near W148, which form the lower rim of the active-site cleft. This latter loop can adopt multiple conformations in some thrombin structures, including that of the original PPACK-thrombin structure (Bode et al., 1989), exhibiting either a defined or a disordered conformation. The active-site residues reside in the centre of the narrow cleft framed by both loops. On the left side is the opening of the deep S1 specificity pocket, which according to the overall shape and the acidic D189 residue at the bottom is similar to that of prototypical trypsin. However, it differs from the trypsin pocket by an A190 residue, and due to its increased flexibility, allows accommodation of more hydrophobic or even uncharged P1 groups and confers a strong preference for P1-R over K residues. Relatively unique to thrombin is the acidic E192 residue positioned at the entrance to this pocket. It is an important determinant for ligand specificity of thrombin because it assists in selecting inhibitors and substrates (Guinto et al., 1995). Based on the W215 indole moiety, a hydrophobic surface groove extends on top of the S1 pocket, which is partially delimited by the 60 loop.

Table 1: Cleavage Sequences of Thrombin Substrates or Inhibitors Around the Scissile Peptide Bond and the Reactive-Center Loop, respectively.

	P4	P3	P2	P1	P1'	P2'	P3'	Cofactor and Exosite
Fibrinogen Aα	Gly	Gly	Val	Arg	Gly	Pro	Arg	I
Fibrinogen Bβ	Phe	Ser	Ala	Arg	Gly	His	Arg	I
FV	Leu	Gly	Ile	Arg	Ser	Phe	Arg	I and II
FV	Leu	Ser	Pro	Arg	Thr	Phe	His	I and II
FV	Trp	Tyr	Leu	Arg	Ser	Asn	Asn	I and II
FVIII	Ile	Gln	Ile	Arg	Ser	Val	ala	I and II
FVIII	Ile	Glu	Pro	Arg	Ser	Phe	Ser	I and II

FVIII	Gln	Ser	Pro	Arg	Ser	Phe	Gln	I and II
FXIII	Gly	Val	Pro	Arg	Gly	Val	Asn	None
PAR-1	Leu	Asp	Pro	Arg	Ser	Phe	Leu	I
PAR-4	Pro	Ala	Pro	Arg	Gly	Tyr	Pro	None
FXI	Ile	Lys	Pro	Arg	Ile	Val	Gly	I
PC	Val	Asp	Pro	Arg	Ile	Val	Gly	TM (I)
TAFI	Val	Ser	Pro	Arg	Ala	Ser	Ala	TM (I)
AT	Ile	Ala	Gly	Arg	Ser	Leu	Asn	Heparin (II)
HCII	Phe	Met	Pro	Leu	Ser	Yhr	Gln	I and Heparin (II)

AT, antithrombin; F, factor, HC, heparin cofactor, TAFI, thrombin-activable fibrinolysis inhibitor.

The conjunction of hydrophobic residues in this loop together with the base of the active-site cleft form an apolar binding site, which is subdivided into the S2 cavity and the aryl binding site/S4 groove. The latter site is specific for the P4 side chain of all L-amino acids substrates, whereas the P3 side chain extends alongside E192, away from the cleft. The S1' site of thrombin positioned to the right of the active site is limited in size by the K60f side chain, and therefore particularly suited to accommodate small polar P1' side chain. The S2' subsite is of medium size and mainly hydrophobic, so that bulky hydrophobic P2' residues are preferred. The S3' site is open and slightly acidic, resulting in a weak preference for basic P3' side chains and an aversion for acidic residues.

Therefore, a P4 to P3' consensus sequence of an optimal thrombin polypeptide substrate should contain a P4-F/L, any P3 residue, a P2-P/V, a P1-R, a P1'-S/G, a P2'-F, and a P3'-R residue (Table 1). Indeed, the majority of thrombin natural substrates follows this pattern and posses a medium-sized to bulky hydrophobic P4 residue, a P at the P2 position, an R at P1 (except for HCII), an S at P1', a hydrophobic residue at P2', and a basic residue at P3'. Important exceptions to this consensus scheme are the thrombin cleavage site of fibrinogen Aα (FAα), factor XIII, protein C, and HCII. These substrates/inhibitors with non optimal cleavage site bind to regions removed from the active site, such as the Na^+ binding site, exosite I, and/or exosite II (Figure 9). This additional binding makes not necessary to optimally present the scissile peptide bond to the reactive S195-Oγ.

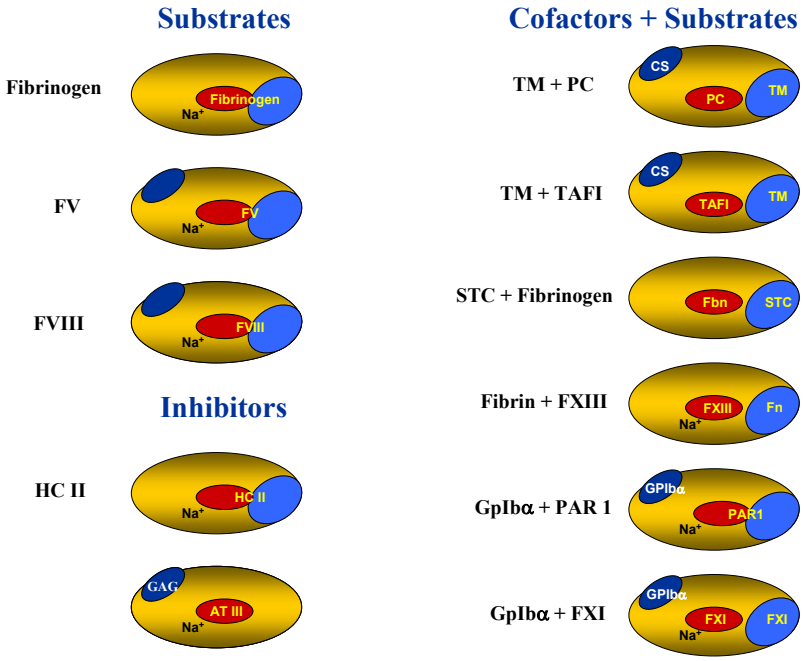


Fig. 9: Schematic interaction of thrombin (represented as a large ellipsoid) and the different surface regions that bind substrates, inhibitors, carbohydrate, and cofactors. AT, antithrombin; F, factor; Gp, glycoprotein; HC, heparin cofactor; PAR, protease-activated receptor; PC, protein C; TAFI, thrombin-activable fybrinolysis inhibitor; STC, staphylocoagulase; TM, thrombomodulin.

Na^+ binding site and allosteric effect

Sodium has been found to be an important allosteric modulator of α-thrombin. The binding of this monovalent cation to thrombin plays a relevant role in the allosteric control of the protease activities, as it causes a conformational transition from a Na^+-free form, referred to as slow, to a Na^+-bound form, referred to as fast. The slow and fast forms are significantly (2:3 ratio) populated under physiologic conditions because the K_d for Na^+ binding is 110 mM at 37°C and physiologic NaCl concentration (140 mM) is not sufficient for saturation. The two forms exhibit different relative activities toward macromolecular substrates. The fast thrombin form cleaves fibrinogen and PARs more efficiently and displays procoagulant, prothrombotic, and prosignaling properties. The slow form preferentially cleaves protein C, exhibiting more anticoagulant properties (Table 2).

The Na^+ binding site has been confined to a site in the center of the 222 loop situated in a solvent-filled cavity behind the S1 specificity pocket located near D189. The bound Na^+ ion, located more than 15 Å away from the catalytic triad, is coordinated octaedrically by the carbonyl oxigens of R R221a and K224 and by four internal water molecules, and is further stabilized by the negative charges of D221 and D222.

Table 2: Enzymatic Properties (K_{cat}/K_m, $\mu M^{-1} \cdot s^{-1}$) of the Fast and Slow Forms of Human α-Thrombin Under Physiological Conditions

	Fast	Slow	r
(D)-Phe-Pro-Arg-pNA[a]	88.9 ± 4	3.5 ± 0.5	26
Fibrinopeptide A release[b]	35 ± 4	1.5 ± 0.1	23
Fibrinopeptide B release[b]	17 ± 1	0.73 ± 0.03	23
PAR-1 cleavage[b]	54 ± 2	1.4 ± 0.1	39
Proteina C acrtivation[b]	0.21 ± 0.01	0.32 ± 0.01	0.7

[a] Measurement were performed in 5 mM Tris, 0.1% PEG 8000, pH 8.0, 25°C, containing 200 mM NaCl or 200 mM ChCl. [b] Measurements were performed in 5 mM Tris, 0.1% PEG 8000, pH 7.4, 37°C, containing 145 mM NaCl or 145 mM ChCl (Di Cera et al., 1997). r: specificity ratio between the fast and slow form.

In a recent study, which combined comprehensive mutational experiments with crystallographic results, Di Cera and co-workers identified residues D189, E217, D222, and Y225 clustering around the Na^+ site as being energetically linked to Na^+ binding and responsible for transducing Na^+ binding into enhanced catalytic activity. The authors suggested that Na^+ binding reorients the R221a carbonyl group to form the R187-D222 salt bridge, thereby favourably orienting the side chains of D189, E192, and S195 and the connecting water network. The structural movements of these and surrounding residues observed are relatively small, namely in the range of usual crystal packing effects. However, these movements may be limited due to the constraints of the crystal packing effects.

Larger structural changes observed with a human S195A thrombin mutant devoid of any active-site or exosite contacts, as well as in a E217K mutant with match residues fibrinogenase activity but similar protein C cleavage capacity, were suggested to be typical for the slow form of thrombin. These crystal structures indicated an allosteric switch mechanism, according to which the Na^+ removal induces flipping of the C168-C182 disulfide bridge and the aromatic side chains of F227, W215, and W60d, resulting in the constriction of the active site cleft, limiting access of substrates. Although many reported observations regarding the perturbations in the vicinity of aromatic residues of the S2 to S4 subsites agree with considering these structures as representing the slow thrombin form, the main problem with these conclusions is that they have Na^+ bound or are not capable of binding Na^+, respectively. Nevertheless, these structures display the enormous plasticity of the substrates interaction surface of thrombin and its capability to propagate binding effects to distal regions allosterically.

Currently available data suggest that Na^+-bound (fast) thrombin form is more stable and exhibits a more open accessible and rigid active site cleft, whereas the Na^+-free (slow) form

possesses a more closed, flexible substrate binding region (De Filippis et al., 2005). Thus the fast from of thrombin would be a better template for productive binding of the inherently flexible cleavage segment of fibrinogen and PAR-1 to its active centre, such that the scissile peptide bonds are optimally presented to the S195 Oγ and the oxyanion hole. Importantly, thrombin substrate complexes in the presence or absence of Na^+ go through transition states that are short-lived, making the structural characterization of these states difficult. Several naturally occurring mutations of the prothrombin gene, like prothrombin Frankfurt, Salakta, Greenville, Scranton, Copenhagen and Saint Denis, affect residues linked to Na^+ binding and are often associated with bleeding (Di Cera et al., 2005).

Exosite I interactions

The thrombin exosite I is a slightly notched surface depression located 20 Å away the active site, mainly placed on the R67 to I82 loop and bordered by the 37-loop and segment K109-K110. In this domain, four charged residues (R67, K70, E77, and E80) form a salt bridge cluster, which is buried well below the surface of the exosite, substantially contributing to the rigidity of the loop (Bode et al. 1992). Over this buried charged spot, several cationic residues are present in this domain: R73, R75, R77a, which form a strong electropositive field around the exosite. When the peptide bond at R77a is cleaved either autocatalytically or by other proteases, the entire loop is lost, giving rise to the derivative γ-thrombin, which is practically devoided of any fibrinogenolytic and platelet-activating ability (Hofsteenge et al., 1988, De Cristofaro et al., 1991, Rydel et al., 1994).

The particular richness in electropositive residues generates a strong electrical field, although the highest surface positive field is generated by the exosite II (Myles et al., 2001). Electrostatic forces generated by the exosite I's charges have an important influence on the biomolecular association of thrombin with some ligands. The exosite I, in fact, represents the recognition site for many macromolecular ligands, such as fibrin(ogen), thrombomodulin, PAR-1, heparin cofactor II, hirudin, coagulation factor V, VIII and XIII (Stubbs et al., 1992, Steen et al., 2002, Esmon et al., 1996, Saddasivan et al., 2000, Hall et al., 1991, Myles et al., 2001, Myles et al., 1998, Hall et al., 2001). It is hypothesized that the exosite I's electrical field could pre-orient the enzyme for a productive interaction so that the hydrophobic stacking components can subsequently stabilized the various adducts.

Exosite II interactions

Based on chemical modifications data and the concentrated clustering of positive charges (Bode, 2006), the other positively charged patch located toward the top left of the thrombin molecule had been identified as the heparin binding site. In fact, exosite II is the locale for interaction with polyanionic ligands such as heparin and glucosaminoglycans. At this surface, a small hydrophobic L234-based groove is surrounded (in clockwise order) by basic residues R93, R101, R165, R233, R126, K236, K235, K240, and H91, with most of their side chain charges not compensated by adjacent negative charges. This assignment was confirmed with exosite II mutants by measuring the heparin-catalyzed thrombin inhibition by ATIII (Gan et al., 1994), and crystallographically by a structure of the α-thrombin complex with an eight unit heparin fragment, in which each octasaccharide chain is sandwiched between two adjacent thrombin molecules (Carter et al., 2005). Exosite II is also the locale for thrombin interaction with the platelet receptor GpIb, hemadin, and the acidic moiety of the fibrinogen γ' chain (De Cristofaro et al., 2001, Lancellotti et al., 2008).

CHAPTER 1.3

Hirudin Structure and Function

Among natural anticoagulants, hirudin is the most potent and specific inhibitor of thrombin, with a dissociation constant in the 20-200 fM range. This protein was firstly derived from the salivary glands of the blood sucking leeches (Marwardt, 1991). A number of hirudin variants have been isolated, mostly from *Hirudo medicinalis*. All of them consist of a single polypeptide chain of 63-66 amino-acid residues with three disulfide bridges (Dodt et al., 1985, Tripier, 1988, Stringer and Lindenfeld, 1992, Vindigni et al., 1994). Other hirudin variants have been isolated from *Hirudinaria manillensis*, also called buffalo leech, a species which is significantly more specialized for mammalian parasitism than *H. medicinalis* (Sawyer, 1986). The primary structure of hirudin isoforms from *H. manillensis* (Steiner et al., 1992) show significant (60-75%) sequence similarity to those of *H. medicinalis* (Dodt et al., 1985, Scharf et al., 1989), and, in particular, they all contain six cysteine residues in highly conserved positions, this signifying the same pattern of disulfide bond formation (Thornton, 1981).

Due to its important pharmacological implications, the hirudin-thrombin pair has been the object of thorough structural and biochemical investigations. With respect to this, structural studies conducted on hirudin in the free (Folkers et al., 1989, Haruyama and Wuthrich, 1989) and thrombin-bound state (Rydel et al., 1991), indicate that this inhibitor is composed of a compact N-terminal region, encompassing residues 1-47 and cross-linked by three disulfide bridges, and a flexible negatively charged C-terminal tail that binds to the fibrinogen-recognition site (exosite I) on thrombin. The N-terminal domain covers the active site of thrombin and through its first three amino acids extensively penetrate into the specificity pockets of the enzyme (see Figure 10). Notably, the N-terminal tripeptide makes about half of the total contacts observed for the binding of the core domain 1-47 to thrombin (PDB code: 4HTC.pdb) and accounts for ~30% of the total free energy of binding. The importance of hirudin for studies of thrombin structure-function correlations stems from a large number of contacts made with the enzyme, and fromits extraordinary specificity. Hirudin is, in fact, an excellent probe of the allosteric state of thrombin because it covers about 12% of the available area of the enzyme upon formation of the complex (Rydel et al., 1991).

By limited proteolysis of full-length hirudin, it's possible to produce the peptide fragment corresponding to the N-terminal domain 1-47 of hirudin HM2 (De Filippis et al., 1997). Although far less active (at least 2 x 10^5-fold) than intact hirudin, this fragment (like the parent hirudin molecule) binds ~ 30-fold more tightly to the *fast* form of thrombin than to the *slow* form (Di Cera et al., 1992), suggesting that the structural determinants for this behavior are stored in the N-terminal domain. Hence, mutational studies on hirudin fragment 1-47 can be useful not only to investigate the physico-chemical determinants responsible for the extraordinary affinity and specificity of hirudin for thrombin, but also to probe the structural properties of the enzyme in the *slow* or *fast* forms.

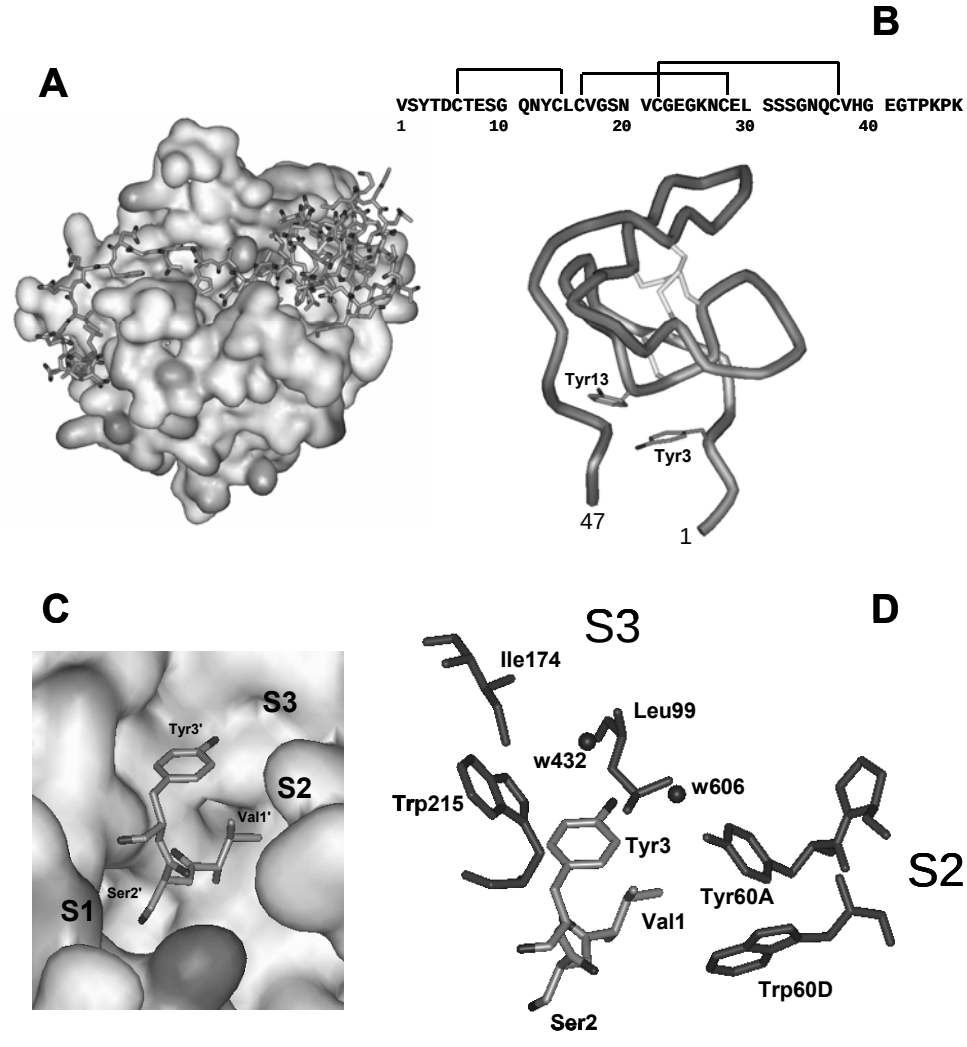


Fig. 10: (**A**) Schematic representation of the interaction of full-length hirudin (stick) with thrombin (van der Waals surface), based on the crystal structure of the hirudin-thrombin complex (Rydel et al., 1991). (**B**) Amino acid sequence (Scacheri et al., 1993) and schematic representation of the solution structure (Nicastro et al., 1997) of the N-terminal fragment 1-47 of hirudin HM2 from *H. manillensis*. Disulfide bonds are indicated by plain lines. (**C**-**D**) Structural details of the interaction of the N-terminal tripeptide

of hirudin with thrombin. Val1' of hirudin contacts the S2 site of thrombin, shaped by Tyr60a and Trp60d; Ser2' covers, but does not penetrate, the S1 site, containing Asp189 at the bottom; Tyr3' fills the apolar S3 site, formed by Trp215, Leu99, and Ile174. Structural water molecules at the hirudin-thrombin interface in the S2-S3 sites are also indicated.

The N-terminal tripeptide of hirudin

Val1. NMR studies indicate that Val1 is almost fully exposed to solvent and highly flexible in the free hirudin (Sziperski et al., 1992, Nicastro et al., 1997). Conversely, in the thrombin-bound state, Val1 is completely buried into the active site of the enzyme and fixed into a single side-chain conformation (Rydel et al., 1991). In particular, the α-NH_2 group of Val1 forms an hydrogen bond with the Oγ of the catalytic Ser195, while its side-chain makes numerous hydrophobic contacts with Tyr60A and Trp60D, shaping the S2 of the enzyme (Fig. 9). This loop, absent in other homologous trypsin-like proteases, defines the S2 specificity site on thrombin and narrows the access to the active-site such that only small-sized apolar residues are allowed.

Ser2. Analysis of the 3D-structure of hirudin-thrombin complex reveals that position 2 of hirudin is located at the entrance to, but does not enter, the primary specificity (S1) site (i.e. Asp189) (Rydel et al., 1991). Hence, the occupancy of the S1 site is not strictly required for binding.

Tyr3. Contrary to the high conformational flexibility observed for the first two residues of hirudin, Tyr3 has a well-defined conformation both in the free and bound state. In the hirudin-thrombin complex, the side-chain of Tyr3 projects into the apolar binding site of thrombin (the S3 site), formed by a large hydrophobic cavity comprising residues Trp215, Leu99 and Ile174. The importance of position 3 is confirmed by the fact that Tyr3 is highly conserved through the hirudin family (Steiner et al., 1992). Analysis of the 3D structure of hirudin-thrombin complex reveals that the δ^+ hydrogens on the edge of Tyr3 ring can favorably interact in a T-shaped conformation with the δ^- π-electron cloud of the aromatic ring of Trp215 at the S3 site on thrombin (see Figure 9D).

CHAPTER 1.4

SH3 Domain

Especially since the determination of the human genome, protein–protein interactions have been considered one of the key elements to explain the complexity of living organisms (Xenarios and Eisenberg, 2001). Among the several protein–protein interaction modules is the Src homology-3 (SH3) domain, one of the widest-spread protein modules mostly observed in eukaryotic organisms having been identified in many hundreds of signalling proteins in species from yeast to human (Mayer, 2001; Tong et al., 2002). Often found in intracellular or membrane-associated proteins (Mayer 2001), SH3 domains are known to act through their binding to specific proline-rich sequence motifs as molecular adhesives and to play a crucial role in the formation of multiprotein complexes and networks responsible for signal transduction, cytoskeletal organization and other cellular processes (Morton and Campbell, 1994; Mayer, 2001). In humans, mutations in several SH3 domains are known to cause severe malfunctions leading, among others, to inflammatory diseases and to cancer (Smithgall, 1995; Dalgarno et al., 1997; Vidal et al., 2001). SH3-domains are relatively small protein modules, usually comprising 50–70 amino acid residues. From the structural point of view, they have a common characteristic fold (Yu et al., 1992, 1994; Musacchio et al., 1994), with five to eight β-strands arranged into two tightly packed antiparallel β-sheets, linked by the RT and the n-SRC-loops.

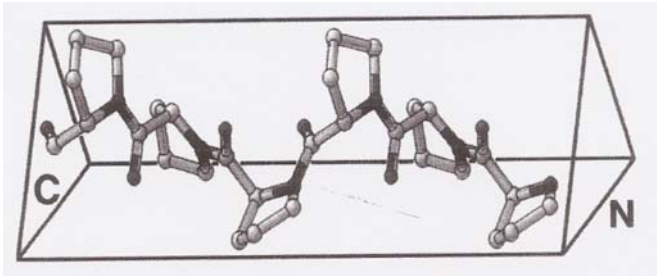


Fig. 11: The polyproline type II helix is an all-trans left handed helix, whose structure can be represented as a triangular prism. The three consecutive side chains that account for one turn of the PPII helix occupy a different edge of the prism (i.e. they are threefolf symmetric along the helical axis).

The ligands, which are usually accommodated in a groove adjacent to the RT-loop, adopt a left-handed polyproline-2 (or PPII) helix conformation (Musacchio 2002). This

conformation is due to the restricted torsion angle (ϕ) of the proline residue in a polypeptide chain (i.e. -65°). These segment occur in many prokaryotic and eukaryotic proteins and appear to be involved in the regulation of important cellular phenomena.

Because of the intrinsic symmetry of PPII helices, the target peptides can be accommodated into the SH3 cavity in both of the two possible orientations, traditionally classified into two families, according to specific consensus motifs. The so-called type-I and type-II ligands have the general consensus sequences +xxPxxP (or +x ϕPx ϕP) and xPxxPx+ (or ϕPx ϕPx+), respectively, where x is any amino acid; + is in most cases a basic residue, usually an arginine; and ϕ is a hydrophobic residue (Mayer 2001; Musacchio 2002). More recently, it has become clear, however, that such a classification might be too rigid as sequences not belonging to either of the two motifs are able to be recognized and that key positions not interacting directly with the ligand can favour one of the two orientations (Fernandez-Ballester et al., 2004). In general, the affinity of most of the known SH3 domains for their peptide ligands is relatively low, with Kd values ranging from 1–100 μM and relatively low selectivity (Feng et al., 1994; Nguyen et al., 2000). This finding opens the intriguing question of what is the physiological significance and the evolutionary advantage of maintaining such relatively low affinities and selectivity. Because of their importance at the cellular level, a deeper understanding of the target specificity of particular SH3 domains may eventually lead us to develop molecules able to inhibit specific interactions *in vivo*.

2. Results

Protein Engineering
by Genetic Methods

CHAPTER 2.1

Expression of Human α-Thrombin in *E. coli*

INTRODUCTION

Thrombin (EC 3.4.21.5) is a serine protease that plays a central role in haemostasis. Human thrombin is a glycoprotein consisting of two polypeptides, a 259-residue B-chain and a smaller 36-residue A-chain, connected by a disulfide bond between Cys1 in the A-chain and Cys122 in the B-chain. B-chain contains the catalytic triad of thrombin and is homologous to the catalytic domains of other trypsin-like serine protease including trypsin, chymotrypsin, and elastase. The A-chain lies along the side of thrombin opposite to the active site. In addition to interchain disulfide bond, there are three disulfides in the B-chain, namely Cys42-Cys58, Cys168-Cys182, and Cys191-Cys220 (chymotrypsinogen numbering). Thrombin circulates in the blood as an inactive zymogen, prothrombin, which is converted to thrombin by two sequential factor Xa cleavages at Arg271-Thr272 and Arg320-Ile321 (prothrombin numbering). In the absence of the rest of the prothrombinase complex (factor Va, calcium, and phospholipids), factor Xa has been shown to cleave initially at Arg271-Thr272. This cleavage produces prethrombin-2, the smallest physiological single-chain precursor of thrombin, and prothrombin fragment 1.2 (Esmon and Jackson, 1974).

In the interest of pursuing further structural, mutational, and microcalorimetric studies aimed at dissecting the interaction between thrombin and its ligands, we have developed an expression system for the production of recombinant thrombin, and mutants thereof. A large number of mammalian (Le Bonniec et al., 1991, Tsiang et al., 1995, Russo et al., 1997) and bacterial (Di Bella et al., 1995, Soejima et al., 2001, Jonhson et al., 2005) systems for the production of thrombin and its mutants had been reported so far. A original system had also been developed in which thrombin activity is reconstituted by combining wild-type ζ1-thrombin, a proteolitically-derived fragment of human thrombin, with recombinant ζ2-thrombin from *Escherichia coli* (Gan et al., 1991, Gan et al., 1994). ζ1-Thrombin consist of the A-chain of thrombin linked by a disulfide bond to residues 16-148 of the B-chain (chymotrypsinogen numbering will be used therefore), while ζ2-thrombin consists of residues 149 to 247 of the B-chain. Among these methods, we had chosen to use *E. coli* for over-

expression of human thrombin, using as starting point the method reported by Huntington's group (Johnson et al., 2005). *E. coli*, in fact, is a prokaryotic system which is easy to work with, to scale up, and less time-consuming than eukaryotic system. Consequently, it has the potential for producing higher yields of thrombin than those usually obtained from mammalian system. In addition protein expression in *E. coli* facilitates isotopic labelling of amino acids for NMR studies, and has also the added advantage of producing de-glycosilated thrombin which allows functional analysis of the role of carbohydrate on thrombin properties. However, since the environment of *E. coli* cytosol is reducing, proteins are often isolated as insoluble inclusion bodies and must be solubilized and refolded into their native conformation.

In developing an *E. coli* system for the production of thrombin, a decision has been made whether to work with the two-chain species, thrombin, or with one of its earlier single-chain precursor. With a single-chain species, it should be easier to form the native disulfide pairing since the entropy loss when the two chains are connected is eliminated. An additional consideration was the number of disulfide bonds which must be formed. Prethrombin-2 has only four disulfide bonds, whereas prethrombin-1 has seven and prothrombin has twelve. If disulfide bonds occur randomly, a native four-disulfide conformation statistically should be much more likely to form than one containing either seven or twelve disulfide bonds. Therefore, human prethrombin-2, the smallest single-chain precursor of thrombin, was chosen for expression in *E. coli*.

In this chapter, we described the over-expression of prethrombin-2 in *E. coli*. Prethrombin-2 was isolated as intracellular inclusion bodies, which were reduced and refolded using different buffers and dilution strategies. The method with the highest yield had been chosen for production scale-up, and chemical characterization of the recombinant protein. Our findings indicate that this is the best method for producing α-thrombin in *E. coli* reported so far. Furthermore, we demonstrated that sequence and disulfide bond topology of rThb, established by peptide mass fingerprint analysis, and mass spectrometry analysis are the same as those of natural hThb, and that the kinetic constants (K_m and k_{cat}) for the hydrolysis of the chromogenic substrate (D)-Phe-Pro-Arg-pNA are identical for both rThb and hThb. In conclusion, *E. coli* is a convenient factory for obtaining large amount of recombinant thrombin.

MATERIALS AND METHODS

Materials

Plasmid construct was made in the laboratory of Prof. James Huntington (Cambridge Institute for Medical Research, Cambridge, UK). Plasma derived α-thrombin was purchased from Haematologic technologies (Essex Junction, VT, U.S.A.).

Production of Recombinant Thrombin

Cultivation and harvesting of *E. coli.* The plasmid containing the gene for human prethrombin-2 under the control of T7 promoter was transformed into *E. coli* strain BL21Star(DE3)pLysS which expresses the T7 RNA polymerase under the inducible lacUV5 promoter. Cells were cultured at 37°C in Luria Bertani broth with 50 ug/ml ampicillin and 34 ug/ml chloramphenicol to a OD_{595} of 0.8. Expression of the T7 polymerase, and consequently, of the foreign gene, was induced by adding isopropyl β-D-thiogalactoside (IPTG) to a final concentration of 1 mM for 6 h. The cells were harvested by centrifugation for 15 min at 4°C. Harvested cells were stored at –80°C.

Purification and extraction of inclusion bodies. Thawed cell pellets (typically from 0.8 liter of cell culture) were lysed in 20 mM Tris-HCl, 1% (v/v) Triton-X, 20 mM EDTA and 20 mM dithiothreitol (DTT), pH 7.4, and sonicated on ice in 15-s bursts for a total of 3 min. After centrifugation at 4500 rpm for 15 min, inclusion body pellets were sequentially washed, then centrifugated, with 20 mM Tris-HCl, 20 mM EDTA, first containing 20 mM (DTT) and 1% (v/v) Triton-X, then containing 1M NaCl, and finally with 20 mM Tris-HCl, 20 mM EDTA alone. The washed pellets were resuspended in 1 ml of H_2O-TFA 0.1% and subdivided in aliquots of 100 µl and stored at -20°C until use.

Solubilization and refolding of prethrombin-2 from inclusion bodies. Each aliquot was then solubilized in 1 ml 6 M guanidinium-chloride, 30 mM L-cysteine, and incubated at 25°C for three hours. The solubilized protein concentration was determined using bovine serum albumin as a standard with protein assay reagent (Bio-Rad Laboratories, California), based on the method reported by Bradford (Bradford, 1976). Three refolding trials were performed on three different aliquots. In one case, refolding was firstly initiated by dilution of the solubilized protein into 5 ml of refolding buffer (i..e., 20 mM Tris-HCl, 0.5 M NaCl, 0.6 M

arginine, 20 mM Ca_2Cl, 10% (v/v) glycerol and 0.2% (v/v) Brij58, pH 8.5) containing Gnd-HCl 6 M. This solution was then diluted to 250 ml of refolding buffer, by the reverse dilution method. In the second case, the same procedure of refolding was used, but a little was changed in the refolding buffer composition, with ammonium acetate instead of Tris-HCl to keep the pH to 8.5. In the third case we use, for comparison, the same procedure used by Huntington's group in a recent work (Johnson et al., 2005).

The refolded protein was dialyzed in all cases against 10 litres of 25 mM Tris-HCl, 2 mM EDTA, 0.1% PEG 6000 and 0.12 M NaCl at 25°C, overnight. The precipitate was removed by centrifugation and filtration, before purification of the correctly folded prethrombin-2 on a 5 ml heparin-Sepharose column (GE Lifescience) eluting with a gradient of 0.12-0.95 M NaCl. The elution position of the correctly folded prethrombin-2 was determined by high resolution mass-spectrometry on a Mariner ESI-TOF instrument from Perseptive Byosystem, which gave mass values in agreement with the expected amino acid composition within 10 ppm mass accuracy. The inactive proteins were activated by incubation for 1 h at 37° with snake venom from *E. carinatus* that had been pre-treated with PMSF to inhibit serine protease activity. After 3-fold dilution with 50 mM tris-HCl, pH 7.4, active thrombin was purified on a heparin-Sepharose column as above. The concentration of active protein was estimated by measuring the absorbance at 280 nm and the samples were immediately frozen until use.

Chemical Characterization of Recombinant Thrombin

The homogeneity of proteins was established by SDS-PAGE and RP-HPLC. Standard PAGE in the presence of β-mercaptoethanol was performed use a 12% polyacrylamide gel, according to the method of Laemmli (Laemmli, 1970). The gel was stained with Coomassie Brilliant Blue R-250. RP-HPLC analysess were carried out on a C4 analytical column (4.6 x 150 mm, 5 μm particle size, 300 Å porosity, from Grace Vydac). The column was equilibrated with 0.1% H_2O-TFA and eluted with a linear 0.1% TFA-acetonitrile gradient at a flow rate of 0.8 ml/min. The absorbance of the effluent was recorded at 226 nm. The chemical identity of the purified protein, either in oxidizing and reducing conditions, was established by ESI-TOF mass spectrometry on a Mariner instrument from Perseptive Byosystem. Typically protein samples (10 μl, 10 μM) were analysed by direct injection in acqueos-acetonitrile (1:1, v/v), containing 1% (v/v) formic acid at a flow rate of 15 μl/min. The nozzle temperature was set up at 140°C and the electrostatic potential at 4.4 kV. The instrument was calibrated using the standard protein kit from Sigma.

Localization of the intrachain disulfide bonds in prethrombin-2. In order to assess the disulfide pairing in recombinant prethrombin-2 (Pre2), the protein was firstly subjected to proteolysis with subtilisin Carlsberg for 3 h at 25°C using an enzyme:protein ratio of 1:50 (w/w) in 5 mM Tris-HCl, PEG 8000 0.1%, containing 0.2 M NaCl. Proteolytic mixture was purified by RP-HPLC and the peptide material eluted in correspondence of the two major chromatographic peaks (namely F1 and F2) was lyophilized and analysed by mass spectrometry. Fragment 1, corresponding to the segment Ala-13–Ala149a and fragment 2, corresponding to the segment Asn149b-Glu247, contain two of the four disulfide bridge of prethrombin-2, each. Fragment 1 and 2 were subjected to further proteolysis with different strategies. Proteolysis of fragment 1 was performed with Glu-C proteinase V8 from *S. aureus* for 16 h at 25°C, in 50 mM ammonium bicarbonate buffer, pH 7.8, at a protease:substrate ratio of 1:10 (w/w) in the presence or not of 4 M urea. Proteolysis of fragment 2 was performed in the same experimental conditions as fragment 1, and with bovine trypsin for 16 h at 25° in 50 mM Tris-HCl, pH 8.3, at a protease:substrate ratio of 1:50 (w/w) in the presence of 4 M urea. Proteolytic mixtures were purified by RP-HPLC, and the peptide material eluted in correspondence of major chromatographic peaks was lyophilized and analysed by mass spectrometry.

(D)-Phe-Pro-Arg-pNA substrate hydrolysis by recombinant thrombin

Steady-state kinetic of the hydrolytic reaction of the (D)-Phe-Pro-Arg-*p*-nitroanilide (FPR-pNA) substrate by recombinant thrombin was studied under the experimental conditions of 5 mM Tris-HCl, 0.2 M NaCl, PEG 8000 0.1%, pH 8 at 25°C. Assay was performed by following the release of *p*-nitroaniline resulting from the hydrolysis of FPR-pNA at 405 nm. The concentration of released *p*-nitroanilide was measured using the extinction coefficient of 9,920 $M^{-1}cm^{-1}$ at 405 nm (Lottemberg and Jackson, 1983). The initial rates of FPR-pNA hydrolysis were measured using 50 pM purified enzyme with 0.5-16 μM substrate. K_m and k_{cat} were determined by fitting the data to the Henri-Michaelis-Menten equation. The enzyme active-site concentration was previously titrated with hirudin HM2.

RESULTS AND DISCUSSION

Expression, Purification and Activation of Prethrombin-2

The prethrombin-2 (Pre2) gene from human liver mRNA was cloned into pET23(+) vector. *E. coli* cells, strain BL21Star(DE3)pLysS, were transformed with pET23(+)-Pre2 vector and induced by IPTG. SDS-PAGE analysis of the expression of prethrombin shows that after addition of IPTG, the translational system of *E. coli* is exclusively direct toward to synthesis of the foreign protein (Figure 1A). Recombinant prethrombin-2 produced in *E. coli* was found to be present only as insoluble inclusion bodies as shown by SDS-PAGE (Figure 1B). From 1 litre of *E. coli* cell culture, 1 g of purified inclusion bodies was obtained. The inclusion body fraction purified as described in "Methods", was observed as a major single band on reducing SDS-PAGE (data not shown) and a major peak on RP-HPLC, with a purity as high as 85% (Figure 1C).

Inclusion bodies could be solubilized by adding 30 mM cysteine and 7 M guanidine hydrochloride, but not urea 8 M, to a final concentration of protein of 10-15 mg/ml. To optimize the refolding yield of prethrombin-2 as regard to procedures reported so far, (Di Bella et al., 1995, Soejima et al., 2001, Huntington et al., 2005), we tried for different refolding procedure. Refolding involves transfer of protein molecules from a denaturant environment to an aqueous solvent to force protein to collapse into a compact structure. Optimal procedure to reduce denaturant concentration and assistance of refolding of by solvent additives, plays a key role in determining refolding efficiency. Among current ways reported for removing denaturant agents from solubilized inclusion bodies, we choose the so called rapid dilution and reverse dilution methods. Briefly, in rapid dilution, unfolded sample goes through a rapid collapse of its structure, with no step in which intermediate concentrations of denaturant are present. On the other hand, reverse dilution is done by adding refolding buffer into the unfolded protein containing concentrated denaturant such that both the denaturant and protein concentration decrease simultaneously (Tsumoto et al., 2003). In order to determine optimum refolding conditions for prethrombin-2, we refolded thrombin using the protocol reported by Huntington and co-workers (Huntington et al. 2005), indicated as Refolding 3 (i.e., Rapid Dilution and Buffer 3) in the table below, or with some modifications either in refolding buffer composition and in refolding methods (Rfeolding 1 and Refolding 2). In particular we considered (i) the use of L-cysteine or β-mercaptoethanol as an aid in disulfide bond formation and (ii) the reverse dilution or method for removal of

denaturant. The results indicate that addition of L-cysteine is at least 3-fold more effective to prevent the formation of wrong disulfide bonds than β-mercaptoethanol.

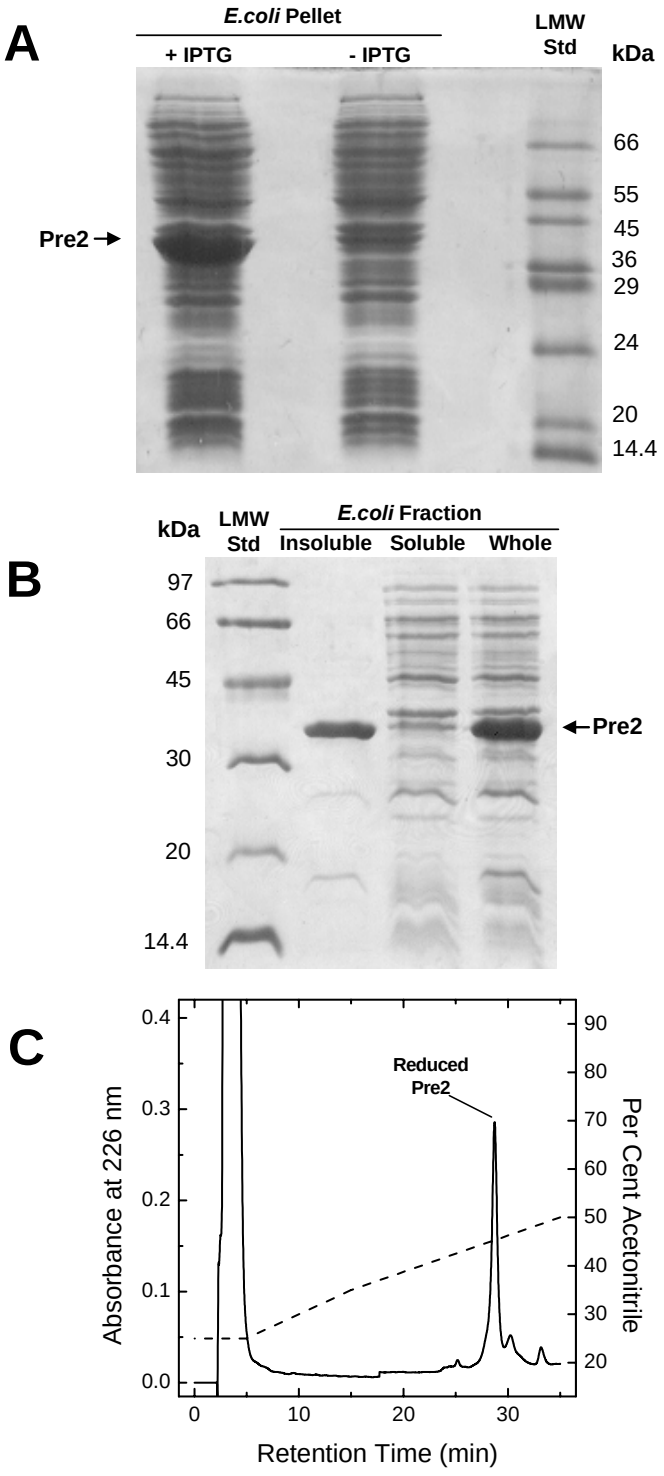


Fig. 1: Chemical analysis of protein expression by SDS-PAGE and RP-HPLC. (**A**) SDS-gel electrophoresis was conducted in a 12% acrylamide gel and proteins were visualized by Coomassie blue staining. From the left to the right: 8 µl of bacterial cell lysate from BL21Star(DE3)pLys cells transformed with the pET(+)23-Pre2 vector DNA in the absence (-) and in the presence (+) of 1 mM IPTG; molecular weight protein markers are indicated on the left-hand side. (**B**) SDS-PAGE analysis of insoluble, insoluble and whole fraction obtained after lysis of BL21Star(DE3)pLys cells induced with

IPTG. (**C**) RP-HPLC analysis of solubilized and reduced inclusion bodies of prethrombin-2. An aliquot (10 μg) of inclusion bodies was loaded onto a C4 column (4.6 x 150 mm) eluted with an acetonitrile-0.1% TFA gradient (---) from 25 to 35 % in 10 min, and from 35 to 50% in 20 min.

Furthermore, the oxido-shuffling system, such as cysteine and cystine, or reduced glutathione and glutathione disulfide reported previously (Soeyima et al., 2001, Johnson et al., 2005) is not effective. Finally, reverse dilution is a better way to remove the denaturant than rapid dilution, and the yield increases as the dilution rate increases. In summary, denatured and reduced thrombin was most efficiently refolded by single-step dilution (250-fold) at room temperature in refolding buffer containing 50 mM $NH_4^+AC^-$, 0.5 M NaCl, 1 mM EDTA, 0.6 M L-arginine, 1 mM L-cysteine, 10% (v/v) glycerol, and 0.2% (w/v) Brij-58, pH 8.5. The refolded protein was then dialyzed against 20 mM Tris-HCl, pH 7.4, 0.25 M NaCl, 2 mM EDTA, PEG 6000 0.1%. and finally purified by heparin-sepharose affinity chromatography. The final yield of refolded and purified prethrombin-2, based on the absorbance at 280 nm (estimated using the extinction coefficient $\varepsilon_M^{280} = 67{,}880\ M^{-1} \cdot cm^{-1}$), was 12% of the starting amount of solubilized protein and about 14 mg per litre of cell culture.

Table 1: Refolding Trials Performed on Recombinant Pre2

Refolding 1	Refolding 2	Refolding 3
NH^+Ac^- 0.1M pH 8.5	NH^+Ac^- 0.1M pH 8.5	Tris-HCl 50 mM pH 8.5
EDTA 1 mM	EDTA 1 mM	EDTA 1 mM
L-Arginine HCl 0.6 M	L-Arginine HCl 0.6 M	L-Arginine HCl 0.6 M
NaCl 0.5 M	NaCl 0.5 M	$CaCl_2$ 20 mM
0.2% Brij 58	0.2% Brij 58	0.2% Brij 58
10% Glycerol	10% Glycerol	10% Glycerol
1 mM L-Cysteine	1 mM -Mercaproethanol	
Reverse Dilution	*Reverse Dilution*	*Rapid Dilution*
Dilution: 250X	Diluiton: 250X	Dilution: 100X
Refolding Yield: 12 %	Refolding Yield: 5 %	Refolding Yield: 5 %

Prethrombin-2 is a precursor form of thrombin, and must be activated by an activating enzyme such as ecarin from the venom of the snake *E. carinatus*. Ecarin, in fact, is known to possess a specific and high activity for converting prethrombin-2 into α-thrombin (Morita et al., 1981), by cleaving quantitatively prethrombin-2 at the Arg320-Ile321 bond (prothrombin numbering). Ecarin-activated α-thrombin can be easily purified by further purification onto heparin-sepharose column (Figure 2). On reducing SDS-PAGE, the recombinant α-thrombin thus obtained showed a band corresponding to B-chain with apparent molecular masses of

about 30 kDa, as expected. Purified α-thrombin was obtained with a final yield of approximately 12 mg from litre of culture of *E. coli*.

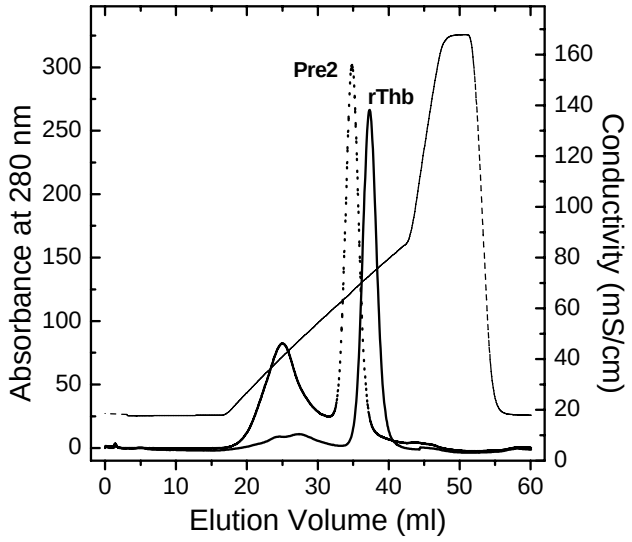
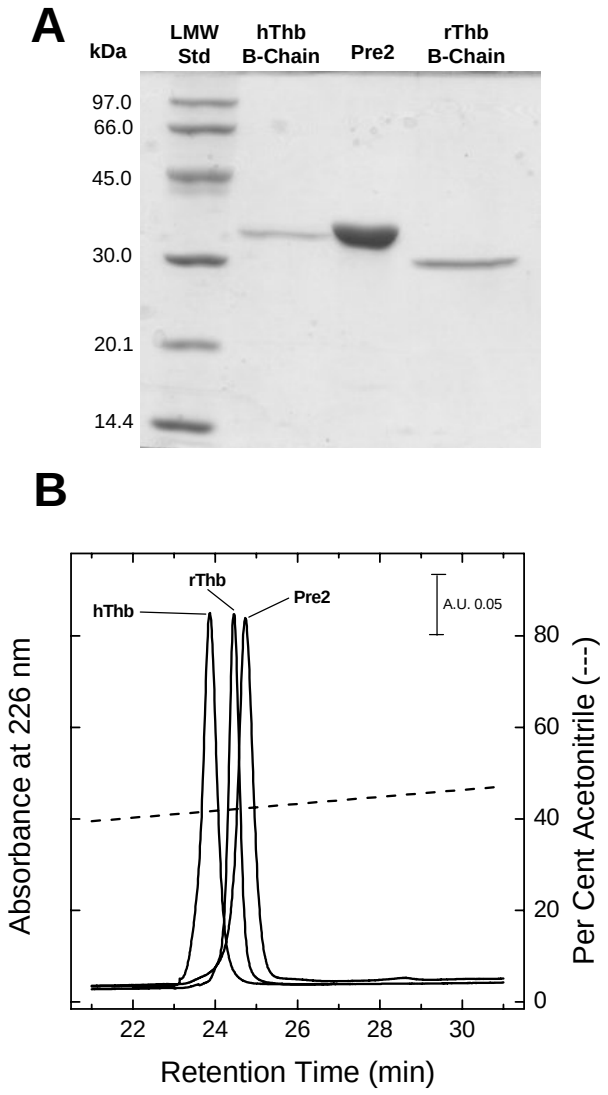


Fig. 2: Heparin-sepharose purification of recombinant prethrombin-2 (···) and α-thrombin (—). The material obtained by solubilization, refolding, dialysis and centrifugation of 100 mg of Pre2-IBs was loaded onto a 5 ml heparin-sepharose column, equilibrated at a flow-rate of 2 ml/min in Tris-HCl 50 mM, pH 7.4, containing 0.12 M NaCl. The column was eluted with a linear gradient of Tris-HCl 50 mM, pH 7.4, 2 M NaCl.(---) The eluate was recorded at 280 nm.

Chemical characterization of recombinant prethrombin-2 and α-thrombin

In order to compare the primary structure of recombinant and plasma-derived thrombin, we performed SDS-PAGE analysis on the pure proteins, together with RP-HPLC and ESI-TOF mass spectrometry analyses. Moreover, for comparison, we analyzed in the same way also the inactive recombinant protein, prethrombin-2. Prethrombin-2 and B-chain of rThb migrated at ~ 36 kDa and 30 kDa, respectively, in close agreement with its theoretical molecular weights, whereas B-chain of hThb migrated at ~ 36 kDa, likely on account of the anomalous binding of SDS to glycoproteins and subsequent altered electrophoretic mobility. Hence, the difference between the species can be ascribed to the contribution of the N-linked carbohydrate chain (Figure 3A). We also analyzed Pre2, rThb and hThb by RP-HPLC, demonstrating that the recombinant proteins eluted as a single peak on a C4 analytical column (Figure 3B) in agreement with a high homogeneity of protein preparations. Accurate molecular weight determination of Pre2 and rThb, carried out by ESI-TOF mass spectrometry (Figure 3C), yielded mass values of 35,872.1 ± 2.9. and 33,820.9 ± 2.1 a.m.u in close

agreement with those expected from the amino acid composition of these species (average masses: 33,872.9 and 33,820.7 a.m.u., respectively). The difference in mass existing between rThb and hThb (average mass: 36,030.2 ± 1.8 a.m.u.) gave the molecular weight of the oligosaccharide chain, that is 2,209.3 ± 0.5 a.m.u. Moreover, the difference between thrombin zymogen and active form is in close agreement with the deletion of the first 18 amino acids in the N-terminus of prethrombin-2 that occurs after thrombin activation, as reported previously (Downing et al., 1975). RP-HPLC (Figure 4) and mass spectrometry analyses were also performed purified on rThb that had been reduced with DTT. We obtained two species with masses 4,090.0 ± 0.5 and 29,735.1 ± 0.3 a.m.u., corresponding to the light (A-) and heavy (B-) chain of α-thrombin, respectively. These values confirm that the difference between recombinant and plasma thrombin resided on the B-chain, according to the location of the glycosilation site.



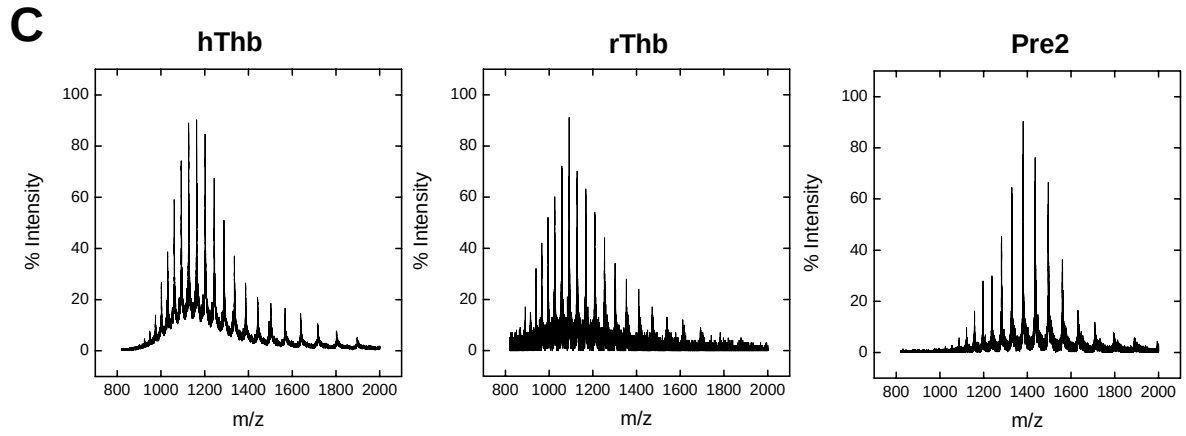


Fig. 3: (**A**) SDS-PAGE analysis under reducing conditions of pure Pre2, rThb and hThb. Molecular mass standards for SDS-PAGE (in kDa) are indicated on the left hand side. (**B**) RP-HPLC analysis of Pre2, rThb and hThb in the oxidized form. (**C**) m/z ESI-TOF mass spectra of RP-HPLC-purified rThb.

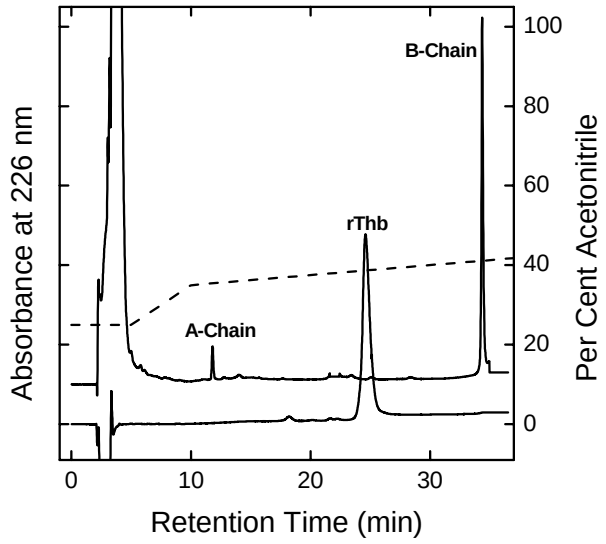


Fig. 4: RP-HPLC analysis of rThb with the four disulfide bonds in the reduced form. Briefly, an aliquot (10 mg) of pure protein in Tris-HCl 0.1 M, pH 8.3, 10 mM EDTA, 6 M Gnd-HCl, was reduced with 0.1 M DTT

It had been previously shown (De Filippis et al., 2005), that subtilisin specifically cleaves thrombin between Ala149a and Asn149b in the 148-loop consistent with the higher flexibility of this segment in the structure of the enzyme. Our findings show that subtilisin cleaves prethrombin-2 with the same specificity (data not shown), and the two fragments thus produced, called respectively 1 and 2, were suitable to subsequent proteases digestion for disulfide pairing assignment. Fragment 1 (F1) corresponds to the polypeptide Ala(-18)-Ala149a at the N-terminus of Pre2, while fragment 2 (F2) corresponds to the Asn149b-Glu259 at the C-terminus of Pre2. These fragments were purified by RP-HPLC and

lyophilized for following digestion. In order to assign the disulfide bond pairing in recombinant thrombin, and to compare it to that of plasma α-thrombin (Figure 5A), we tried to further proteolyzed the two fragments, in order to obtain smaller peptide that allowed us to identify disulfide linked segments by RP-HPLC and mass spectrometry analyses. F1 had been successfully proteolyzed using *S. aureus* V8 protease at pH 7.8, in the presence of 4 M urea. The fragments thus obtained allowed us to determine that the disulfide bond topology in recombinant F1 is identical to that present in the natural counterpart (i.e. Cys1-Cys122, and Cys42-Cys58). With respect to F2, we found it highly stable to physical and chemical denaturants as well as to proteolytic degradation by several different proteases (e.g. trypsin, *S. aureus* V8 protease, thermolysin), and therefore (despite many attempts) we failed to obtain F2 fragments suitable for disulfide pairing assignment. In any case, recombinant F2 showed a retention time in RP-HPLC identical to F2 obtained from plasma thrombin (data not shown). This is a good, albeit indirect, indication that the correct disulfide bonds (Cys168-Cys182, and Cys191-Cys220) were formed, as previously reported (Gan et al., 1991).

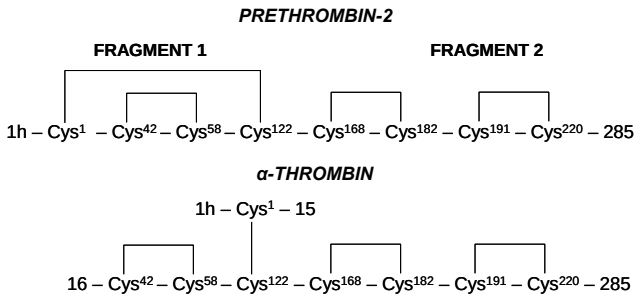


Figure 5: Disulfide bond topology in one-chain prethrombin-2 and in two-chain α-thrombin .

Functional characterization of recombinant α-thrombin

To provide a preliminary comparison between the enzymatic properties of recombinant and plasma-derived α-thrombin, we determined the steady-state parameters for the hydrolysis of the chromogenic substrate (D)-Phe-Pro-Arg-*p*-nitroanilide (Figure 6). As shown in Table 2, the k_{cat} (52.8 ± 3.1 s^{-1}) and K_m (0.61 ± 0.05 μM) determined for the hydrolysis of FPR-*p*NA by recombinant α-thrombin remained essentially unchanged to those determined for plasma-derived α-thrombin (k_{cat} = 44.8 ± 2.1 s^{-1} and K_m = 0.55 ± 0.08 μM). Although not itself definitive, the above results suggest that glycosilation doesn't inferred the catalytic function and the molecular recognition at the active site.

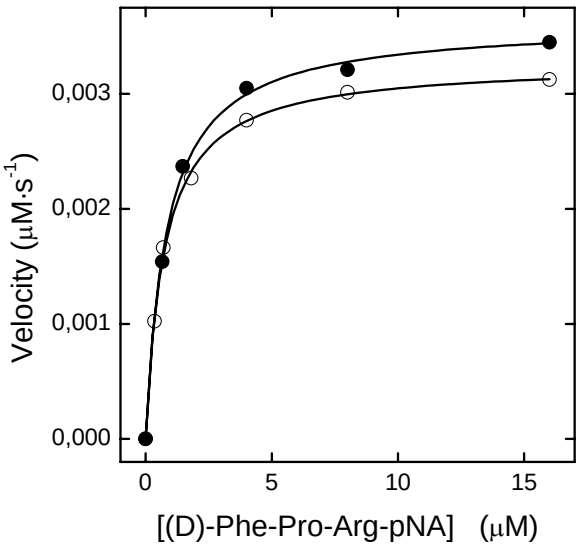


Fig. 6: Steady-state kinetic hydrolysis of D-Phe-Pro-Arg-*p*-nitroaniline substrate hydrolysis by recombinant (●) and plasma-derived (○) thrombin.

Table 2: Kinetic Parameters for Hydrolysis of D-Phe-Pro-Arg-pNA

	K_m (μM)	K_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1} \cdot s^{-1}$)
hThb	0.55 ± 0.08	44.8 ± 2.1	81.5 ± 3.2
rThb	0.61 ± 0.05	52.8 ± 3.1	86.6 ± 2.1

CONCLUSIONS

Altogether, our results demonstrate that *E. coli* is a suitable and convenient expression system for obtaining milligram quantities of pure and fully active recombinant thrombin to be used in future structural, calorimetric and functional studies. However, the N-linked glycosilation at Asn60g in plasma thrombin is absent in recombinant protein expressed in *E. coli*. The 60- loop caps the active site and plays a major role in the restricted specificity of thrombin. On this basis, recombinant thrombin produced in this study could be useful in further studies aimed to establish the effect of glycosilation in structure, stability and function of thrombin.

Chapter 1.3

Effect of Glycosilation on Structure, Stability and Function of Human α-Thrombin

INTRODUCTION

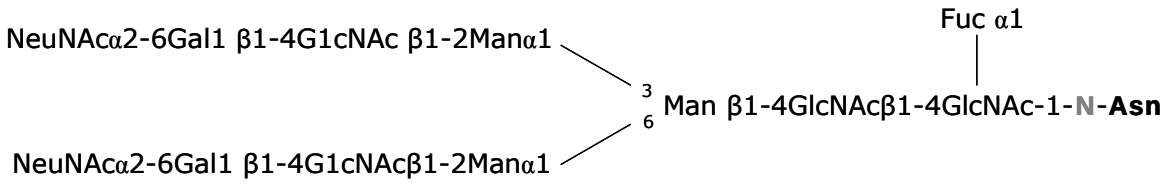
Protein glycosilation is one of the most common and most complex posttranslational processes. Databases show that more than half of all proteins are glycosilated (Apweiler et al., 1999) and N-linked glycosilation, usually to Asn residues, is by far the most common way of glycosilation. The asparagine is linked to N-acetylglucosamine (GlcNAc) residues and additional sugar residues depend on the type of glycosilation. Protein glycosilation *in vivo* involves numerous processing steps and different donor and acceptor molecules. As soon as the nascent polypeptide chain enters the lumen of the endoplasmic reticulum, prefabricated oligosaccharide units are transferred from dolicolphosphate to asparagine residues and the association of glycans finally terminates in the Golgi complex. The consensus sequence for N-linked glycosilation is Asn-Xxx-Ser/Thr, where Xxx is any amino acid other than proline.

Oligosaccharides usually extend from the surface of proteins rather that participating in their internal structures, in agreement with their hydrophilic character. Both experimental and theoretical studies indicated that protein oligosaccharides have mobile and rapidly fluctuating conformations. In addition, the N-glycosidic linkage to asparagine is flexible and, as a result, the sugars can shield large regions of the surface of the proteins, with a likely effect on the local polypeptide conformation. According to this, only in a few crystal structure electron density had been observed in correspondence to sugar group locations.

Many biological functions have been ascribed to sugar moiety in proteins (e.g., directing the protein to the appropriate cellular location, recognition by carbohydrate-binding proteins, modifications of immunological properties, etc.), but its role for protein structure, stability and folding is still debated. Many studies have demonstrated that, although N-linked glycans help proteins to fold, removal of the glycans from the folded protein has no effect on its activity, but its stability and folding kinetics its altered (Mitra et al., 2003). General roles described for glycosilation are stabilization of protein conformation, protection from

proteolysis, and enhancement in polypeptide solubility. Also, certain glycosilation sites in a given protein are more important compared with other site, suggesting that glycosilation exerts local effects on protein conformation in addition to more global effect (Mitra et al., 2003, O'Connor et al, 1996). However, the mechanism with which these effects are obtained still remains to be clarified.

This work explored the role of glycosilation on the structure, stability, and function of human α-thrombin. α-Thrombin is a serine protease that plays a central role in the coagulation cascade. In its active form, it possesses one N-linked oligosaccharide chain attached to Asn60g (chymotrypsinogen numbering) as determined by mass spectrometry studies (Nilsson et al., 1983), and the following carbohydrate structure has been proposed:



It comprises approximately 7% of the molecular weight of thrombin and consists of a single biantennary chain with two terminal sialic acid residues (Horne and Gralnick, 1983). The Asn residue to which oligosaccharide chain is attached, is located on the so called 60-loop. This loop encompasses the residues Cys58-Leu64 and bears 9 additional residues if compared with the corresponding segment of chymotrypsin. The 60-loop plays a central role in the specificity of thrombin interaction with substrates and inhibitors. In fact, the residue Tyr60 mostly contributes to define the S2 site that confers the preference of thrombin for substrates with a small and hydrophobic side chain at P2. Only the first N-acetilglucosamine residue attached to Asn60g has been observed in a few X-ray crystal structure obtained to date (Rydel et al., 1992), indicating that this chain is likely unstructured and flexible.

The only published work about carbohydrate function in thrombin (Horne and Gralnick., 1983), reported that partial removal of oligosaccharide chain had no apparent effect on any of thrombin activities explored (i.e., fibrinogen clotting, factor VIII activation, anti-thrombin-activation, and platelet aggregation). However, deglycosilation with exoglycosidase is difficult to perform.

One of the simpler ways with which non-glycosilated forms of glycoproteins are generated, is the expression of proteins in bacteria. Human α-thrombin had been expressed in *E. coli* for 7 years, and, functional (Soejima et al., 2001) and structural (Johnson et al., 2005) studies reported so far reveal that non-glycosilated thrombin shares with its natural

counterpart the same structure and functional properties. However, no solution studies aimed to determine whether the oligosaccharide chain can influence in such a way, the conformation and the stability of thrombin have been performed so far.

In the present study, recombinant human α-thrombin expressed in *E. coli* was compared with plasma-derived thrombin, with respect to the effect of glycosilation on stability, conformation, function and propensity to aggregate, either by chemical and spectroscopic techniques. Our results show that glycosilation increases the stability of the protein and inhibits aggregation, without essentially affect the global conformation, catalytic function, and molecular recognition of thrombin.

MATERIALS AND METHODS

Materials

The chromogenic substrate D-Phe-Pro-Arg-*p*-nitroanilide (FPR-pNA) was synthesized as previously described (De Filippis et al., 2002). Reagents for electrophoresis, buffers, and organic solvents were of analytical grade and obtained from Fluka or Merck (Darmstadt, Germany). Plasma derived α-thrombin was purchased from Haematologic technologies (Essex Junction, VT, U.S.A.).

Expression and Purification of Recombinant Thrombin

Human recombinant α-thrombin was expressed and refolded as previously described. Briefly, after transformation of pET23(+)-Pre2 vector into *E. coli* strain BL21Star(DE3)pLysS, cells were cultured at 37°C in LB broth with 50 ug/ml ampicillin and 34 ug/ml chloramphenicol to a OD_{595} of 0.8, followed by induction with 1 mM isopropyl β-D-thiogalactoside (IPTG) for 6 h. Harvested cells were stored at –80°C. Thawed cell pellets (typically from 0.8 liter of cell culture) were lysed in 20 mM Tris-HCl, 1% (v/v) Triton-X, 20 mM EDTA and 20 mM dithiothreitol (DTT), pH 7.4, and sonicated on ice in 15-s bursts for a total of 3 min. After centrifugation at 4500 rpm for 15 min, inclusion body pellets were sequentially washed, then centrifugated, with 20 mM Tris-HCl, 20 mM EDTA, first containing 20 mM (DTT) and 1% (v/v) Triton-X, then containing 1 M NaCl, and finally with 20 mM Tris-HCl, 20 mM EDTA alone. The washed pellets were resuspended in 1 ml of H_2O-TFA 0.1%, subdivided in aliquots of 100 µl each and stored at -20°C until use. Each aliquot

was then solubilized in 1 ml 6 M guanidinium-chloride, 30 mM L-cysteine, and incubated at 25°C for three hours. Refolding was initiated by dilution of the solubilized protein to 5 ml of Gnd-HCl 6 M, 50 mM $NH_4^+AC^-$, 0.6 M *L*-arginine, 0.5 M NaCl, 1 mM EDTA, 10% Glycerol, 0.2% Brij 58, 1 mM L-cysteine, then by a dropwise dilution with 250 ml of 50 mM $NH_4^+AC^-$, 0.6 M *L*-arginine, 0.5 M NaCl, 1 mM EDTA, 10% glycerol, 0.2% brij 58, 1 mM L-cysteine, pH 8.5. The refolded protein was dyalised against 10 litres of 25 mM tris-HCl, 2 mM EDTA, 0.1% PEG 6000 and 0.12 M NaCl at 25°C, overnight. The precipitate was removed by centrifugation and filtration before purification of the correctly folded prethrombin-2 on a 5 ml heparin-Sepharose column (GE Lifescience) eluting with a gradient of 0.12-0.95 M NaCl. The inactive protein eluted from heparin-sepharose column was collected and incubated for 1 h at 37° with snake venom from *E. carinatus* (1:10, enzyme:substrate ratio) that had been pre-treated with PMSF to inhibit serine protease activity. After 3-fold dilution with 50 mM Tris-HCl, pH 7.4, thrombin was purified on a heparin-Sepharose column as described above. The concentration of thrombin was estimated by measuring the absorbance at 280 nm and the samples were immediately frozen until use.

The purity of thrombin preparations (~98%) was established by SDS/PAGE (12% acrylamide gel) and RP-HPLC on a C4 analytical column (4.6 x 150 mm, 5 µm particle size, 300 Å porosity) from Grace-Vydac (Hesperia, CA, U.S.A.). The column was equilibrated with 0.1% (v/v) aqueous TFA and eluted with a linear 0.1% (w/w)-TFA-acetonitrile gradient at a flow rate of 0.8 ml/min. The absorbance of the effluent was recorded at 226 nm. The chemical identity of the purified proteins was established by ESI-TOF mass spectrometry on a Mariner instrument from Perseptive Biosystems (Stafford, TX, U.S.A.). The nozzle temperature was set at 140°C and the electrostatic potential at 4.4 kV. The instrument was calibrated using the standard protein kit from Sigma.

Analytical Size Exclusion Chromatography

The apparent molecular weight of thrombin proteins was estimated by SEC on a HR10/30 Superose- 12 column (1 x 30 cm) equilibrated in 5 mM Tris-HCl, pH 8.0, 25°C, containing 0.2 M NaCl and eluted with the same buffer at a flow rate of 0.3 ml/min. Absorbance of effluent was monitored at 280 nm. The column was calibrated using the low-molecular weight gel-filtration protein calibration kit (GE Healthcare). The value of void volume (V_0) and interstitial volume (V_i) were determined by loading dextran blue (2 x 10^3 kDa) and the tripeptide H-Gly-Tyr-Gly-OH.

Spectroscopic Techniques

UV absorption. All measurements were recorded at 25°C in 5 mM Tris-HCl, pH 8.0, containing 0.2 M NaCl or 0.2 M ChCl. Measurements were carried out in the correct buffer and the final spectrum resulted from subtraction of the corresponding baseline. Temperature correction was applied for Tris buffer (Stoll, and Blanchard, 1990).

Protein concentration was determined by ultraviolet (UV) absorbance at 280 nm on a Lambda-2 spectrophotometer from Perkin Elmer (Cupertino, CA, U.S.A.) using molar absorbivity values (ε) obtained experimentally, as in the case of plasma thrombin (65,935 $M^{-1} \cdot cm^{-1}$, Fenton et al., 1977), or calculated on the basis of the amino acid composition (i.e. 66,390 $M^{-1} \cdot cm^{-1}$ for rThb, and 67,880 $M^{-1} \cdot cm^{-1}$ for Pre2, (Pace et al., 1995). Second-derivative absorption spectra were recorded in the same buffer.

Circular dichroism. Circular dichroism (CD) spectra were recorded on a Jasco (Tokyo, Japan) model J-810 spectropolarimeter equipped with a thermostated cell holder, connected to a model RTE-111 (NesLab) water-circulating bath. Far-UV spectra were recorded in a 1-mm pathlength quartz cell, at scan speed of 10 nm/min, using a response time of 16 s, and were the average of four accumulations. Near-UV spectra were recorded in a 1-cm cell, at a scan-speed of 50 nm/min, with a response time of 2 s, and were the average of 16 accumulations. CD signal was expressed as the mean residue ellipticity, calculated with the formula $[\theta] = \theta_{obs} \cdot MRW/(10 \cdot l \cdot c)$, where θ_{obs} is the observed ellipticity in degree, MRW is the mean residue molecular weight, l is the optical pathlength in cm, and c the protein concentration in g/ml.

Fluorescence emission. Fluorescence measurements were carried out on a Jasco spectrofluorimeter model FP-6500, equipped with a Peltier model ETC-273T temperature control system from Jasco. Spectra were recorded in a 1-cm pathlength quartz cell, at a scan speed of 200 nm/min by exciting the protein samples at 280 nm or 295 nm.

Urea stability. Urea-mediated denaturation of α-thrombin was carried out at 25°C, at pH 8.0 in the presence of 0.2 M chloride salts, after 30-min incubation in the correct buffer, as described previously (De Filippis et al., 2005). Protein concentration was 30 nM. Spectra were recorded at a scan speed of 240 nm/min, with an excitation/emission slit of 5 nm. Since

for recombinant proteins denaturation was only partially reversible, the value of $[urea]_{1/2}$ was estimated as reported (Di Stasio et al., 2004).

Turbidimetric measurements. Measurements were carried out at 25°C on a double-beam model Lambda-2 Perkin-Elmer spectrophotometer equipped with a thermostated cell-holder, connected to a Haake model F3-C water-circulating bath. Acid-induced aggregation of thrombin proteins was monitored by adding to a solution of thrombin (2 ml, 0.1 mg/ml) in 5 mM Tris-HCl, pH 8.0, aliquots (2-5 μl) of 0.05-0.1 M aqueous HCl. The solution pH was measured at 25 ± 0.2 °C with a Metrohm (Herisau, CH) model 632 pH-meter. After HCl addition, protein samples were gently stirred for 1 min and allowed to equilibrate for 2 min and then OD at 320 nm was measured as a function of [HCl]. Absorbance values were corrected for sample dilution (< 5% of the final volume) and subtracted for the corresponding value of the buffer alone.

Functional Characterization

Steady-state kinetic of (D)-Phe-Pro-Arg-pNA substrate hydrolysis. Steady-state kinetic of the hydrolytic reaction of the (D)-Phe-Pro-Arg-*p*-nitroanilide (FPR-pNA) substrate by thrombin was studied under the experimental conditions of 5 mM Tris-HCl, 0.2 M NaCl, PEG 8000 0.1%, pH 8 at 25°C. All assays were performed by following the release of *p*-nitroaniline resulting from the hydrolysis of FPR-pNA at 405 nm. The concentration of released *p*-nitroanilide was measured using the extinction coefficient of 9,920 $M^{-1}cm^{-1}$ at 405 nm (Workman and Lundblad, 1978). The initial rates of FPR-pNA hydrolysis were measured using 50 pM purified enzyme with 0.5-16 μM substrate. K_m and k_{cat} were determined by fitting the data to the Henri-Michaelis-Menten equation. The enzyme active-site concentration was previously titrated with hirudin HM2.

Inhibition of thrombin by hirudin HM2 and its N-terminal (1-47) fragment. K_d values of hirudin HM2 and its N-terminal fragment for thrombins were estimated by classical competitive inhibition experiments of thrombin-mediated substrate hydrolysis, according to the tight-binding inhibition model (Copeland 2000), as previously detailed. This method has been described in detail by Di Cera and co-workers (Ayala and Di Cera, 1994). HM2 was incubated at 25 ± 0.2 °C for one hour with 250 pM thrombin in 5 mM Tris, pH 8.0, containing 0.1% (w/v) PEG 8000 and 0.2 M NaCl. The reaction was started by addition of S-2238 (45

μM) and the release of *p*-nitroaniline (*p*NA) was determined by recording the absorbance increase at 405 nm. Thrombin inhibition data were fitted to equation:

$V_i/V_0 = 1-\{([E]+[I]+K_I^{app})-\{([E]+[I]+K_I^{app})^2-4\cdot[E][I]\}^{1/2}\}/2\cdot[E]$

to obtain the apparent inhibition constant, $K_{I\ app}$. [E] and [I] are the total enzyme and inhibitor concentrations and V_i and V_0 are the steady-state velocities of substrate hydrolysis by thrombin, in the presence (V_i) or absence (V_0) of the inhibitor. $K_{I\ app}$ values were corrected for substrate concentration and for the K_m value of S-2238 for thrombin (2.1 μM), according to equation:

$K_I = K_I^{app}/[1+([S]/K_m)]$

where K_I is equal to the dissociation constant, K_d, of the enzyme-inhibitor complex (Copeland 2000).

Binding of [fluorescein]-hirudin$^{54–65}$(PO_3H_2) to α-thrombin. Fluorescein-conjugated and phosphorylated C-terminal hirudin-(54–65) peptide, [F]-hirudin$^{54-65}$(PO_3H_2), having the sequence GDFEEIPEEY(PO_3H_2)LQ, was synthesized as described previously (Lancellotti et al., 2008). Binding of this peptide to exosite I of thrombin was studied by monitoring the decrease of the peptide fluorescence occurring upon interaction with thrombin. Fluorescence spectra were recorded on a Jasco (Tokyo, Japan) spectrofluorometer. Excitation and emission wavelengths were 492 and 516 nm, respectively, using an excitation/emission slit of 3/5 nm. During titration experiments, the decrease of fluorescence intensity at 516 nm was recorded as a function of thrombin concentration. For all measurements, the long time measurement software (Jasco) was used. Fluorescence intensities were corrected for dilution (*i.e.* 8–10%) at the end of the titration. Data were analyzed by the following binding isotherm equation using the program Origin 7.5 (MicroCal Inc.),

$$\left(\frac{F-F_0}{\Delta F_{max}}\right)=\left(\frac{\alpha}{2}\right)\times\left(1+\frac{K_d+L}{P_0}-\sqrt{\left(\left(1+\frac{K_d+L}{P_0}\right)^2-\frac{4\times L}{P_0}\right)}\right)+1$$

where ΔF_{max} is the maximum fluorescence change; K_d is the dissociation constant; L is the total concentration of thrombin; and P_0 is the concentration of [F]-hirudin$^{54–65}$(PO_3H_2).

Binding of Na^+ monitored by fluorescence emission. Fluorescence experiments were carried out using a Jasco FP-6500-model (Tokyo, Japan) spectrophotometer. Fluorescence titration of recombinant and plasma-derived thrombin with Na^+ took place under experimental conditions of 5 mM Tris, 0.1% PEG 8000, 800 mM ionic strength, pH 8.0, at 25°C. Titrations were carried out as follows: a solution containing 30 nM enzyme, 5 mM Tris-HCl, 0.1% PEG 8000 and 800 mM NaCl was incrementally substituted to a solution containing 30 nM enzyme, 5 mM Tris, 0.1% PEG 8000 and 800 mM ChCl. Ionic strength and enzyme concentration were held constant, while the Na^+ concentration was varied. Excitation was at 280 nm and emission was measured at 334 nm. Appropriate corrections were made for buffer fluorescence. The value of thrombin intrinsic fluorescence as a function of Na^+ concentration was fitted according to the equation (Krem and Di Cera, 2003):

$$F = \frac{F_0 + F_1 \frac{x}{K_d}}{1 + \frac{x}{K_d}}$$

where F is the intrinsic fluorescence at a particular Na^+ concentration, F_0 is the intrinsic fluorescence in the absence of Na^+, F_1 is the intrinsic fluorescence under saturating Na^+, x is the Na^+ concentration and K_d is the equilibrium dissociation constant for Na^+ binding.

Hydrolysis of PAR1-(38–60) and PAR4-(44–66) Peptide by Recombinant and Plasma-Derived Thrombin. PAR1-(38–60) (^{38}LDPRSFLLRNPNDKYEPFWEDEE^{60}) and PAR4-(44–66) (^{44}PAPRGYPGQVCANDSDTLELPDS^{66}) peptides were synthesized by solid-phase peptide synthesis using the Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) method. Cleavage of these peptides by thrombin was followed by measuring the release of the peptide LDPR and PAPR from the cleavage of the N-terminus of PAR-1 and PAR-4 respectively, according to a previously described method (De Candia et al., 1997). Briefly, 0.5 μM PAR-1 peptide and 25 μM PAR-4 peptide was incubated with 50 nM thrombin in Tris-HCl 5 mM, PEG 8000 0.1%, containing 0.15 M NaCl at 25°C. At time intervals (2, 5, 8, 12, 20, 30, 60 min) the reaction was stopped with 0.3 M H_2O-HCOOH 10% (v/v), and the cleaved peptide was measured by reverse-phase HPLC, using a 250 x 4.6 mm C18 column from Vydac. Experimental concentrations of PAR-peptides cleaved at time t (P_t), were fitted to the following equation,

$$P_t = P_\infty (1 - e^{k_{obs} t})$$

where P_∞ is the peptide concentration at $t = \infty$ and k_{obs} is the pseudo-first order rate of PAR-1 hydrolysis, equal to $E_0 \times k_{cat}/K_m$ (E_0 is the thrombin concentration).

RESULTS AND DISCUSSION

Chemical Analysis of Glycosilated and Deglycosilated Forms.

Purified proteins were firstly analyzed by RP-HPLC and mass spectrometry analysis. The chromatographic analysis show that proteins were pure (i.e., ≥ 98%), and that human thrombin elutes from the column at a lower per cent of acetonitrile with respect to recombinant thrombin, in agreement with a higher hydrophilic-to-hydrophobic ratio of glycosilated proteins (Figure 1A). In fact, RP-HPLC retention time has been used as an experimental measure of the surface hydrophobicity of globular proteins under mildly denaturing conditions (Purcell et al., 1999). Prethrombin-2 (Pre2) elutes at higher per cent of acetonitrile, probably due to a higher hydrophobic surface of the zymogen with respect to the mature protein. Mass spectrometry analysis gave a mass value of 33,820.9 ± 2.1 a.m.u., in close agreement with its amino acid composition. The difference between the mass of human thrombin and the mass of recombinant thrombin (i.e. 36,030.2 ± 1.8 a.m.u.) gave the weight of oligosaccharide chain that is 2,209.3 ± 0.5 a.m.u. Reducing and non-reducing SDS-PAGE analyses showed a difference in the mobilities of recombinant (rThb) and plasma-derived (hThb) α-thrombins. In particular, rThb B-chain migrates as expected at ~29 kDa when run on a reducing polyacrilamide gel (Figure 1B), where the disulfide bond between the A- and the B- chain is in the reduced from. On the other hand, hThb B-chain (33 kDa) migrates at ~ 36 kDa, likely on account of the anomalous binding of SDS to glycoproteins and subsequent altered electrophoretic mobility. The difference between the two forms is therefore due to the contribute of the N-linked carbohydrate chain to the mass of hThb, and in agreement with the altered electroforetic mobility of glycoproteins. Similar conclusion were found when proteins run on a SDS-PAGE under non reducing conditions, where A- and B- chains are connected by a disulfide bond (Figure 1C). On the other side, Pre2 migrated at ~ 36 kDa, in perfect agreement with its theoretical molecular weight and showed identical electrophoretic mobility when run in a SDS-PAGE gel under reducing conditions, according to the presence of a single polypeptide chain.

Prethrombin, plasma, and recombinant thrombin have very similar molecular masses, and thus changes in global structure should be reflected in differences between the hydrodynamic

radii for the three species. In order to assess the global extent of changes in structure occurring upon activation of Pre2 to thrombin or between plasma and recombinant thrombin, these three species were subjected to analysis by gel filtration on a calibrated column of Superose 12. With respect to retention time in gel filtration, at 0.2 M NaCl, pH 7.5, all proteins elute with an apparent molecular weight lower than expected, likely due the symmetric structure of thrombin. However, aspecific binding effect cannot be ruled out. Moreover, as shown in Table 1, the apparent molecular weight of hThb is increased by about 1.8 fold relative to the de-glycosilated species, in agreement with the contribution of flexible oligosaccharide chains to the hydrodynamic radius of glycoproteins. Furthermore, the higher apparent molecular weight of Pre2 relative to rThb is not due to a looser conformation of the zymogen protein as previously reported (Stevens and Nesheim, 1993), but to the presence of a highly flexible 18-amino acid tail at the N-terminus.

Table 1: Deterrmination of the Apparent Molecular Weight of hThb, rThb, Pre2, and tPre2 by Size-Exclusion Chromatography.

	Estimated MW (Da)	Real MW (Da)
hThb	28,314	36,030
rThb	15,849	33,820
Pre2	22,824	35,872
tPre2	15,849	33,802

Errors on the estimation of molecular weight are about ~ 30%.

Conformational and Stability Properties of Glycosilated and Deglycosilated Forms.

Thrombin conformation and stability were investigated by a number of spectroscopic techniques, at 25°C in buffer Tris-HCl 5 mM, pH 8.0, PEG 8000 0.1%, containing 0.2 M NaCl. Similar studies were conducted on prethrombin-2, the inactive smallest single-chain precursor of thrombin, which shows high structural similarity with thrombin, but lacks any hydrolytic activity.

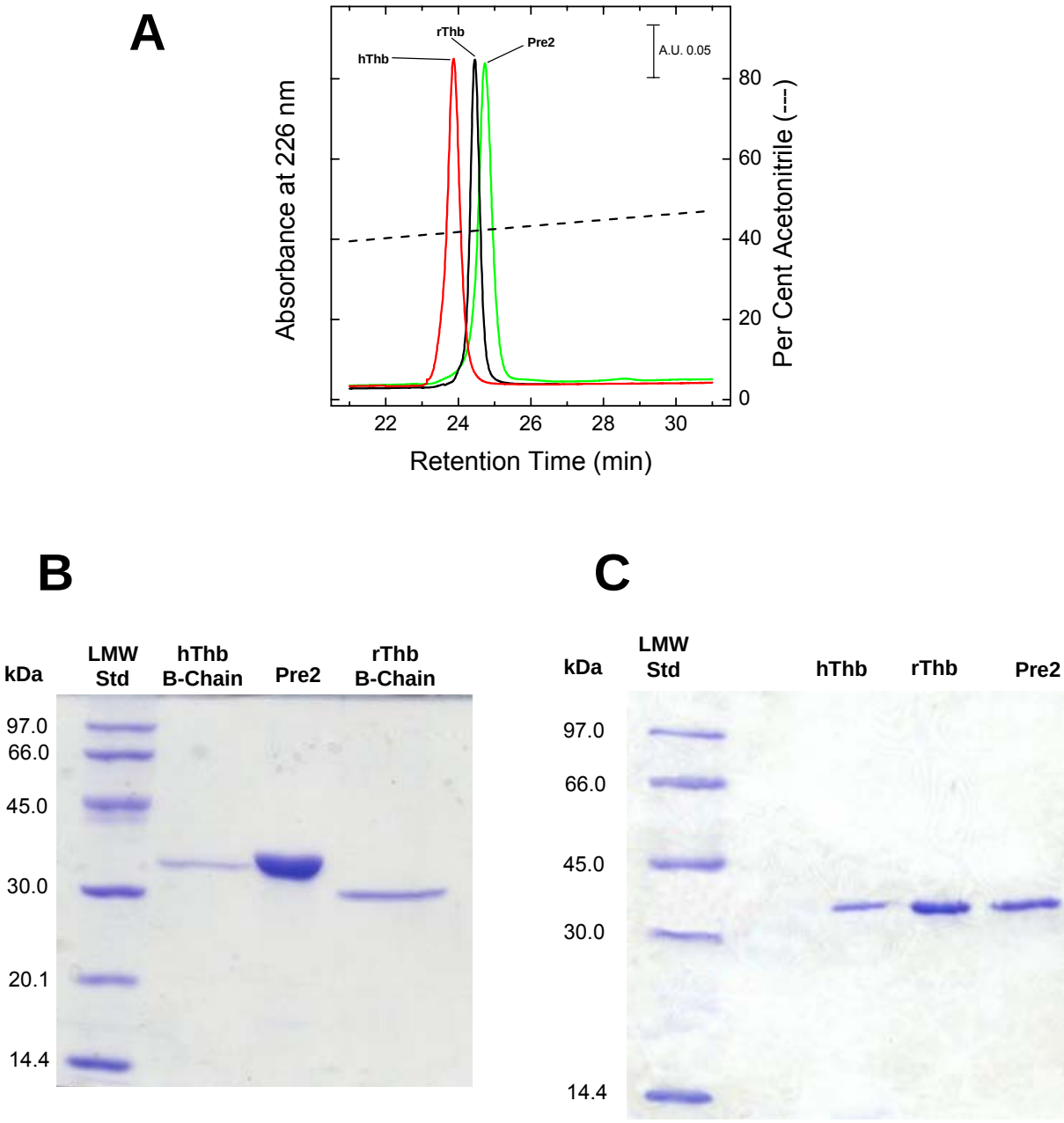


Fig. 1: Chemical analysis of rThb, Pre2, and hThb by RP-HPLC and SDS-PAGE. (**A**) RP-HPLC analysis of rThb, Pre2, and hThb. Aliquots (15 µg) of the purified proteins were loaded onto a Vydac (The Separations Group, Hesperia, CA) C4 column (4.6 x 150 mm, 5 µm particle size), eluted with a linear acetonitrile-0.1% TFA gradient, at a flow rate of 0.8 ml/min. (**B**) SDS-PAGE analysis on a 12%-acrylamide gel. Aliquots of purified proteins were loaded onto the gel under reducing conditions. (**C**) SDS-PAGE analysis on a 12%-acrylamide gel. Aliquots of purified proteins were loaded onto the gel under non reducing conditions. Acrylamide gels were stained with Coumassie Brilliant Blue R-250.

Circular dichroism. Regardless of the presence or not of glycosilation, the shape of the far-UV CD spectra of human α-thrombin resembles that of a protein possessing a significant helical content (Figure 2A). However, a deeper inspection reveals that the two minima centred at 210 and 225 nm are red-shifted by 3-5 nm compared to the bands typical of helical conformation, occurring at 220-222 nm and 208 nm. More importantly, the intensity of the CD bands is unrealistically low for a polypeptide chain with substantial helical secondary structure. Indeed, the crystallographic structure of thrombin (1PPB) yields 14% α-helix and 27% β-sheet. These observations suggest that the far-UV CD spectrum of thrombin may be also affected by the contribution of aromatic amino acids and disulfide bonds in the far-UV region. Such effects become more important for those proteins displaying low CD signal intensity and are most prominent in systems where aromatic groups are clustered in the protein structure (Brahms, and Brahms, 1980), as in the case of thrombin. Of note, three aromatic/hydrophobic clusters can be clearly identified in thrombin structure, comprising S2 and S3 specificity sites (1PPB) (De Filippis et al., 2005).

The spectra of recombinant and plasma proteins share a similar shape and signal intensity (Figure 2A), suggesting that glycosilation do not alter the secondary structure of the proteins. Little changes observed in intensity values likely involve perturbation of the environment of some aromatic amino acids. In fact, 7 aromatic amino acids (Figure 5) are located in the glycosilation loop (i.e. Trp60d, Tyr60a) or are exposed to transient interaction of the oligosaccharide chain (i.e. Trp96, Tyr94, Trp89, Trp237, Trp51). Moreover the oligosaccharide is composed by chiral elements that can cause an induced chirality on the peptide bond.

Near-UV CD is often taken as a fingerprint of aromatic side-chain and disulfide bond topology in proteins, since both shape and intensity of dichroic bands strongly depend on the exact type and orientation of interacting groups in the asymmetric protein environment and, usually, higher conformational flexibility lowers signal intensity (Strickland, 1974). The near-UV CD spectra of thrombin display extensive fine structure between 282 and 305 nm, assigned to the contribution of Trp-residues, and between 260 and 280, where Phe- and Tyr-residues absorb. Likely, the contribution of disulfide bonds is masked by the absorption of the aromatic amino acids (Strickland, 1974). Although the spectra of recombinant and plasma thrombin share common features, in the absence of glycosilation thrombin spectrum shows a more intense band at 287 nm and an higher signal in the 250-280 nm range if compared with plasma thrombin. These features likely derive from the sum of positive and negative contributions of the different chromophores, perhaps due to tyrosine residues, according to the

considerations reported for far-UV spectra. Even in this case the effect of the glycosilation chain on the aromatic residues exposed to its action must be taken in account.

Fluorescence emission. The fluorescence spectra of recombinant and plasma thrombin display a λ_{max} value at short wavelengths (i.e., 334 nm), consistent with the crystallographic structure of human thrombin (Bode et al., 1992) showing that on average Trp-residues are embedded in chemical environments of medium polarity. The reduced quantum yield in the case of recombinant protein is in agreement with far UV- and near UV-CD spectra (Figure 4).

Second derivative absorption. Finally, the conformational properties of thrombin forms were also investigated by second derivative UV spectroscopy. The similarity of the spectra of recombinant and plasma thrombin (Figure 3B) is an additional strong indication that the absence of glycosilation does not significantly alter the overall protein conformation.

Together these results are in close agreement with the tridimensional structures obtained by X-ray diffraction studies of glycosilated and de-glycosilated thrombins. The overlay of the backbone traces (Figure 6) shows that structures are grossly similar with slight differences in the loop and terminal regions.

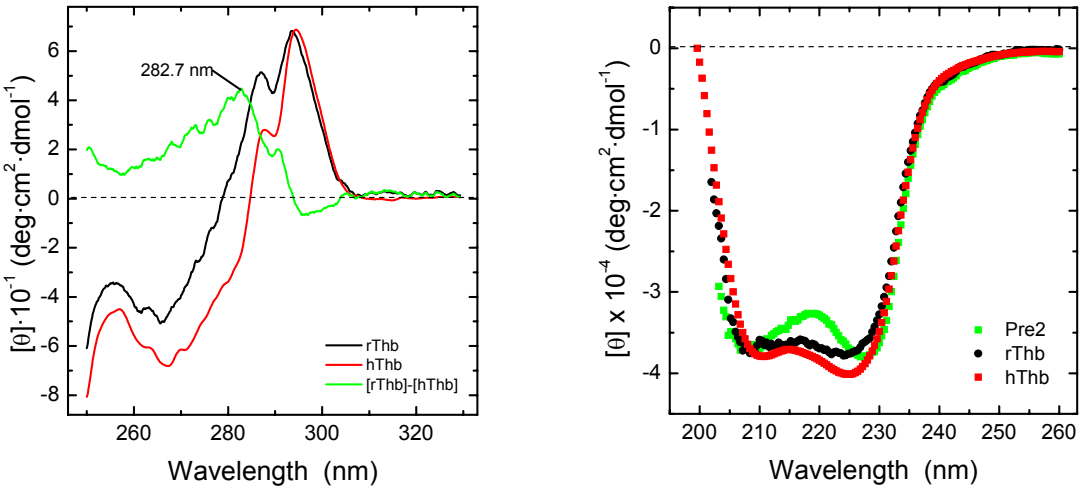


Fig. 2: Far-(A) and Near-(B) UV CD spectra of hThb (—), rThb (—) and Pre2 (—). All spectra were recorded at 25°C in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% PEG-8000, in the presence of 0.2 M NaCl (**A**) CD spectra of thrombin in the far-UV region were taken at a protein concentration of 3.5 µM for recombinant proteins and 5 µM for plasma thrombin. (**B**) CD spectra of thrombin in the near-UV region were taken at a protein concentration of 30 µM and 15 µM respectively for plasma and recombinant proteins.

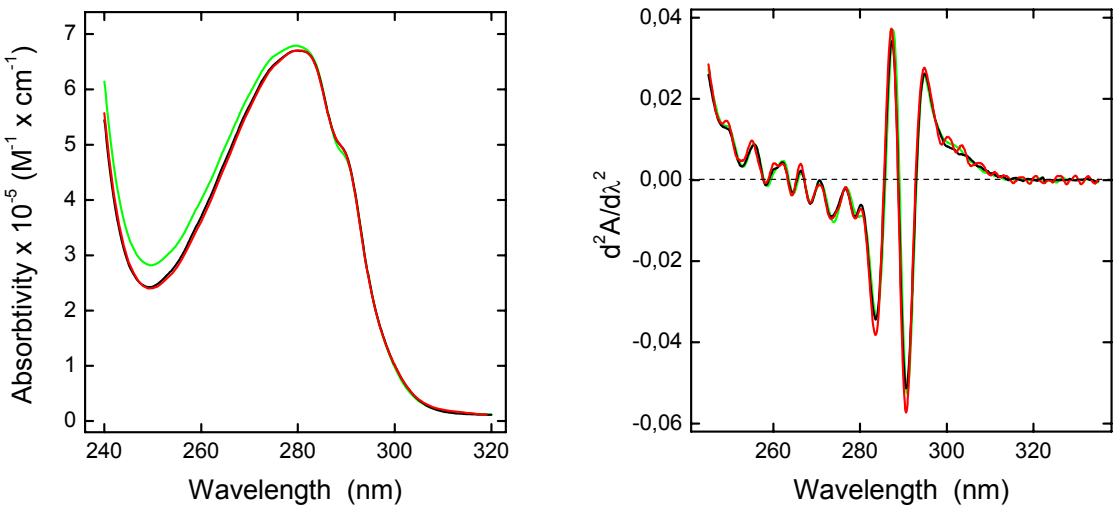


Fig. 3: UV-spectra second-derivative UV absorption spectroscopy of hThb (—), rThb (—) and Pre2 (—). (**A**) UV-Vis spectra and (**B**) second-derivative spectra of thrombin were recorded at 25°C in Tris-HCl buffer, pH 8.0, containing 0.2 M NaCl . All spectra were recorded in a 1-cm quartz cell at a scan speed of 60 nm/min.

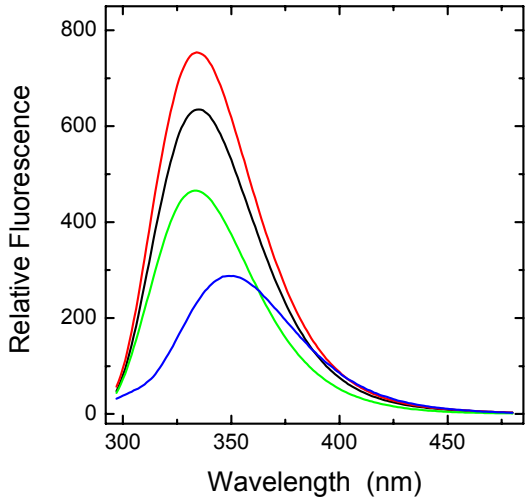


Fig. 4: Fluorescence spectra of hThb (—), rThb (—) and Pre2 (—). Fluorescence spectra were recorded at 25°C in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% PEG-8000, in the presence of 0.2 M. Thrombin samples (2 ml, 30 nM) were excited at 280 nm. The spectrum of denatured rThb in 6 M Gnd-HCl is also reported.

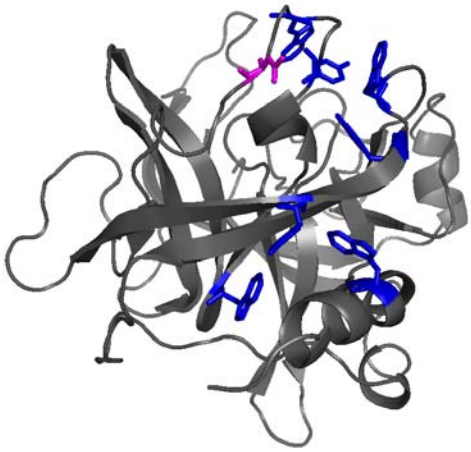


Fig. 5: Schematic representation of Trp and Tyr residues (blue) located on the side of thrombin likely in contact with fluctuations of the oligosaccharide chain linked to Asn60g (magenta) (PDB code: 1PPB, Bode et al., 1992).

Fig. 6: Overlapping of the crystal structure of recombinant α-thrombin (blue, PDB code:2AFQ, Johnson et al., 2005) and plasma α-thrombin (red, PDB code:1JOU, Huntington and Esmon, 2003).

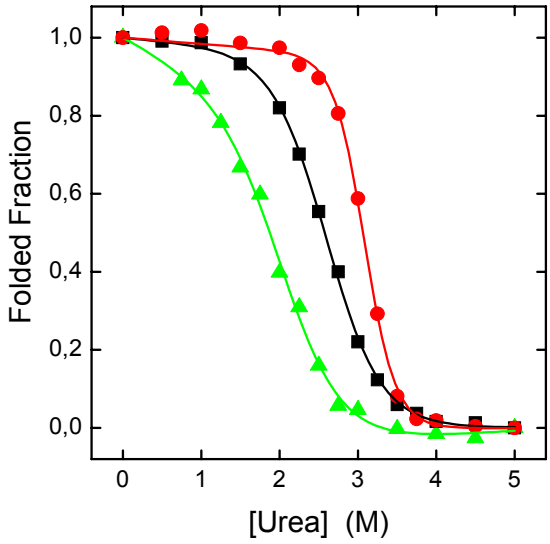


Fig. 7: Urea stability of hThb hThb (—), rThb (—) and Pre2 (—) at pH 8.0. Measurements were carried out at 25◦C in 5 mM Tris/HCl buffer, pH 8.0, containing 0.1% PEG-8000, in the presence of 0.2

M NaCl. Fluorescence measurements were carried out by exciting the samples (2 ml, 30 nM μM) at 280 nm and recording the intensity at 342 nm as a function of urea concentration.

Urea stabilities. The stability of proteins under investigation was studied by recording the fluorescence intensity (Figure 7) at the λ_{max} as a function of urea concentration. Reversibility of denaturation was less than 60%, ad reported fro plasma thrombin at pH 8.0 (De Filippis et al., 2005). Plasma thrombin, a sharp transition is observed, with values of $[urea]_{1/2}$ of about 3.2 ± 0.1 M, where $[urea]_{1/2}$ is the concentration of denaturant corresponding to the midpoint of the transition being considered. As compared to plasma thrombin, the stability of recombinant thrombin and prothrombin is decreased up to 2.6 ± 0.1 M and 1.9 ± 0.1 M, respectively. Our results indicate that glycosilation stabilizes the protein against urea denaturation.

Comparison of the Enzymatic and Recognition Properties of Recombinant and Plasma Thrombin

The enzymatic and recognition properties of recombinant α-thrombin were compared with those of plasma-derived α-thrombin in several assays. First, the steady-state parameters for the hydrolysis of the chromogenic substrate FPR-pNA were determined for both proteins. This tripeptide allowed us to probe either the catalytic machinery and the structure of the primary specificity sites of the enzyme. We found that the K_s value (86.6 ± 1.5 $\mu M^{-1} \cdot s^{-1}$) determined for recombinant α-thrombin is in reasonable agreement with the K_s (81.5 ± 3.2 $\mu M^{-1} \cdot s^{-1}$) determined for plasma-derived α-thrombin.

Table 2: Specificity Constants of PAR-1(38-60) and PAR-4(44-60) Hydrolysis by Plasma and Recombinant Thrombin

	rThb	hThb
PAR-1 (38-60) K_s (k_{cat}/K_m) $\mu M^{-1} \cdot s^{-1}$		
37°C	33.7 ± 1	30 ± 1
25°C	35.8 ± 1	39 ± 2.5
PAR-4 (44-60) K_s (k_{cat}/K_m) $\mu M^{-1} \cdot s^{-1}$		
25°C	0.024 ± 002	0.044 ± 0.006

Thrombin activation of platelets is mediated by hydrolysis of PARs. We compared the kinetic parameters for hydrolysis of PAR-1(38-60) and PAR-4(44-66) peptides by recombinant and plasma-derived thrombin. The experimental method allowed measurement

of the k_{cat}/k_m value pertaining to PAR-soluble peptide by the thrombin forms. The best-fit k_{cat}/k_m values are listed in table 2 and are in close agreement between the two species.

The anticoagulant peptide hirudin and its N-terminal and C-terminal fragments are suitable macromolecular probes for mapping thrombin in the glycosilated or recombinant form, since they makes extensive contacts with its surface, including exosite I and the active-site cleft. By limited proteolysis of full-length hirudin, we were able to produce the peptide fragment corresponding to the N-terminal domain (residues 1-47) of hirudin HM2, and by solid-phase peptide synthesis we produced the C-terminal domain (residues 54-65) of hirudin HV1.

The inhibitory potency of full-length hirudin HM2 and its N-terminal domain (1-47) toward thrombin has been determined by measuring at 405 nm the extent of release of *p*-nitroaniline from the synthetic substrate (D)-Phe-Pro-Arg-*p*NA (Experimental Procedures). From these experiments, the values of the dissociation constant, K_d, of thrombin-inhibitor complexes have been obtained. As shown in Table 2, the values of dissociation constants fully overlapped between the two species. To study the binding of the C-terminal domain of hirudin, we used [F]-hirudin$^{54-65}$(PO_3H_2) having the sequence ^{54}GDFEEIPEEY(PO_3H_2)LQ^{65}. Binding of this peptide to ABE-I of thrombin was studied by monitoring the decrease of the peptide fluorescence occurring upon interaction with thrombin, as reported previously (Verhamme et al., 2002). These experiments showed that the K_d value was not significantly changed for recombinant or plasma-derived thrombin. The equilibrium dissociation constants of the hirudin peptide were in fact equal to 18 ± 3 and 20.9 ± 1.4 nM, in the presence and absence of glycosilation, respectively. Altogether these findings suggest that the binding mode and the affinity of hirudin and its fragments to both active site and exosite I is similar for the two species under investigation.

Table 3: Dissociation Constants of Hirudin Binding to rThb, hThb and Pre2

	rThb	hThb	Pre2
HM2 Hirudin	0.32 ± 0.05	0.2 ± 0.03 pM	2.09 ± 0.07 µM
HM2 Hir(1-47)	45 ± 2 nM	40 ± 2 nM	-
HV1 [F]-Hir(54-65)PO_3H_2	20.9 ± 1.4 nM	18 ± 3 nM	n.d.

Effect Glycosilation in Slow-Fast Allosteric Transition of Thrombin

The most effective modulator of thrombin function in solution is Na^+, which triggers the transition of the enzyme from an anticoagulant ('slow') form to a procoagulant ('fast') form (Dang et al., 1995, Di Cera et al., 1997). The Na^+-bound (fast) form displays procoagulant properties, since it cleaves more specifically fibrinogen and protease-activated receptors, whereas the Na^+-free (slow) form is anticoagulant because it retains the normal activity toward Protein C, but is unable to promote acceptable hydrolysis of procoagulant substrates. Na^+ binding to thrombin elicits an increase (~18%) in the intrinsic fluorescence of the protein with no change in the λ_{max} value at 334 nm. This effect is compatible with a global rigidification of the enzyme in the Na^+-bound form, even though specific local interactions (i.e. reorientation of the Phe227–Trp215 pair) can also influence thrombin fluorescence (De Filippis et al., 2005). The fluorescence increase upon Na^+ binding allows direct measurement of a titration curve as described previously (Krem and Di Cera, 2003). Fluorescence titration of Na^+ binding yields dissociation constant (K_d) of 21.4 ± 2.8 and 18.5 ± 2.1 mM for recombinant and plasma-derived thrombin, respectively. Therefore de-glycosilation exerts no perturbation on monovalent cation binding properties of thrombin. Moreover, none of the molecular events responsible of the allosteric rigidification of the enzyme following cation binding seems to be altered.

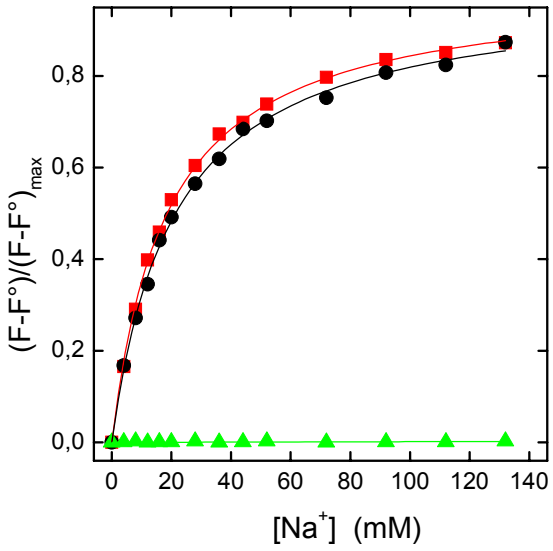


Fig. 8: Determination of the dissociation constants (K_d) of the complexes formed by Na^+ with Pre2 (▲), rThb (●) and hThb (■) thrombins. Fluorescence data were obtained as detailed in the Methods and

expressed as $(F-F°)/(F-F°)_{max}$. Continuous lines represents the best fit of the data points to the equation previously reported, that allowed us to obtain the K_d values reported in the text.

On this basis, thrombin activity and stability were investigated by fluorescence spectroscopy, and activity assays, under temperature and salt conditions in which the enzyme predominantly (> 90%) exists in the Na^+-bound (fast) form (25°C and 0.2 M NaCl) or in the Na^+-free (slow) form (25°C and 0.2 M ChCl), as already reported for plasma thrombin (De Filippis et al., 2005, Dang et al., 1995, Di Cera et al., 1997). Choline chloride was used to control ionic strength, since it is too bulky (r_{Na+} = 1.02 Å, r_{Ch+} ~ 2.90 Å) (Marcus, 1996) to interact with the enzyme and its "inert" nature on thrombin function has been previously verified (Ayala and Di Cera, 1994).

Na^+-bound (fast) thrombin cleaves the chromogenic substrate (D)-Phe-Pro-Arg-*p*-nitroanilide with 30-fold higher specificity than Na^+-free (slow) thrombin (Vindigni et al. 1997). Several studies have been conducted aimed to dissect out the recognition sites of thrombin responsible for different substrate cleavage rates between the fast and the slow form. It had been shown that the transition affects mostly the environment of the S2 site (i.e. Y60a, P60b, P60c, W60d) with modest effects on the S1 site. Moreover, in a recent paper we demonstrated that Na^+-binding "freeze" the S3 site into a more rigid and open conformation (De Filippis et al., 2005). In addition, the structures of the fast (1JOU_AB and 1JOU_CD) and slow (1JOU_EF) form reported by Huntington and Esmon (Huntington and Esmon, 2003), based on the inactive Ser195→Ala mutant, all contained a Na^+ ion bound and showed significant differences at the level of the S2 and S3 sites, which in the slow form protrude on to the protein surface and limit the access to the catalytic pocket. Since the glycosilation is located on the basis of the 60-loop, we analyzed if the allosteric switch of the S2 site is unaltered in the slow → fast transition in the deglycosilated form.

Our findings demonstrated that Na^+-binding to recombinant thrombin determines a functional transition identical to which observed for plasma-thrombin, where K_s for (D)-FPR-pNA is 86.6 ± 2.1 $\mu M^{-1} \cdot s^{-1}$ for the fast thrombin and 3.5 ± 0.2 $\mu M^{-1} \cdot s^{-1}$ for slow thrombin.

One other ligand useful to probe the recognition sites in thrombin allosteric forms is the N-terminal (1-47) domain of hirudin. It had been reported that the N-terminal domain 1-47 of hirudin extensively interacts with thrombin recognition sites and, like intact hirudin (Di Cera et al., 1997, De Filippis et al., 1999) and fibrinogen (Dang et al., 1995, Di Cera et al., 1997) binds preferentially to the fast form of the plasma derived enzyme. Therefore, it may serve as a reliable structural probe for the transition of thrombin from the slow into the fast form in the case of recombinant thrombin. Our findings demonstrated that this fragment bound ~ 30 fold

more tightly to recombinant fast thrombin (K_d = 45 ± 2 nM) than to the slow form (K_d = 1,440 ± 20 nM), as already reported for plasma thrombin (De Filippis et al., 2005). Finally, although the stability of recombinant thrombin to urea denaturation is reduced with respect to plasma thrombin, we estimated a $[urea]_{1/2}$ value of 2.3 ± 0.1 M for the slow form, and 2.6 ± 0.1 M for the fast form, in agreement with the lower stability of the slow form (2.7 ± 0.1 M) relative to fast form (3.2 ± 0.1 M) determined for plasma thrombin (De Filippis et al., 2005).

Effect of Glycosilation on the Propensity to Aggregate of Thrombin

One of the proposed role for protein glycosilation involves the inhibition of aggregation, probably by making the protein more soluble, or by steric interference of the process. Here, the propensity of recombinant and plasma derived thrombins to form insoluble aggregates was investigated by recording the increase of turbidity (i.e. the absorbance at 320-350 nm) as a function of the pH value. In fact, it had been reported in a previous work (Fenton et al., 1977) that thrombin precipitated upon titration with HCl, where heaviest precipitation occurred between pH 4.3 and 4.0, followed by clearing with decreasing pH.

Turbidity is caused by scattering of light, and detects all particles with a hydrodynamic radius greater than the wavelength of the incident light. Thus, the OD at 320 nm is proportional to the amount of insoluble forms of thrombin. To maximize the propensity of thrombin to aggregate, we performed this experiment in the absence of ionic strength, since it had been reported that thrombin has a poor solubility at lower ionic strength (Fenton et al., 1977).

At mild alkaline pH, light scattering is negligible for all thrombin species. Conversely, as the pH decrease, turbidity of recombinant thrombin, either active or in the zymogen form increases with a linear growth shape, in which the formation of precipitates sharply increases from small decrease in pH. On the other hand, a lag phase is present in plasma-derived thrombin behaviour, followed by a gaussian shape. In all cases at low pH the solution become completely cleared for recombinant products, while a slight turbidity is still present for plasma thrombin.

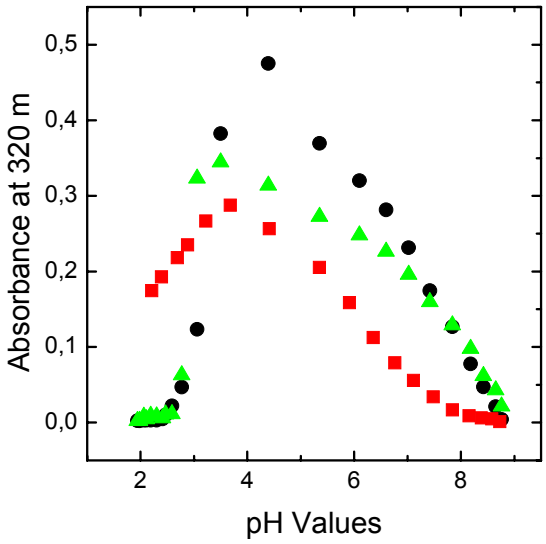


Fig. 9: Effect of pH on precipitation of rThb (●), Pre2 (▲) and hThb (■). The OD values at 320 nm of rThb, Pre2, and hThb were measured at 25°C as a function of pH in 5 mM Tris-HCl, pH 8.0, PEG 8000 0.1%. Measurements were obtained at a protein concentration of 3 μM.

CONCLUSIONS

In conclusion, chemical, spectroscopic and functional studies are consistent with insignificant structural differences between recombinant and plasma-derived thrombin in their native states. However, urea denaturation and turbidimetric studies show that the absence of glycan change the stability against urea and the propensity to aggregate of thrombin. This is consistent with results obtained for other glycoproteins, and with the role proposed for oligosaccharides to stabilized protein conformation by making it more soluble, resistant to proteases and to chemical denaturation.

CHAPTER 2.3

Production of a Natural Variant of Thrombin Containing the Mutation Gly25 → Ser

INTRODUCTION

Prothrombin congenital deficiences are a rare cause of bleeding, affecting 1:2,000,000 in the general population (Peyvandy et al., 2002). Prothrombin is encoded by a gene which is 21 kb in length and located in chromosome 11p11-q12. It is composed of 14 exons separated by 13 introns (Degen and Davie, 1987). This gene is transcribed into a prepropeptide of 622 amino acids. Exons 1 and 2 encode the prepropeptide region (residue -42 to -1), exons 2 and 3 encode the γ-carboxyglutamic (Gla) domain (residues 1-40), exons 3-7 for kringle-1 domain (residues 41-155) and exons 7 and 8 for kringle-2 domain (residues 156-271). Exons 8-9 (residues 271-320) encode the light chain of thrombin and exons 9-14 encode the catalytic serine protease domain (residues 321-579). To date, 39 single-nucleotide acid mutations in prothrombin gene have been identified, distributed in almost every part of the molecule (Figure 1). These mutation have been reported to cause prothrombin deficiency and bleeding disorders by two phenotypes: hypoprothrombinemia, with concomitantly low levels of coagulant activity and antigen, and dysprothrombinemia, with very low activity but subnormal or normal antigen levels. For some mutations, as those located at factor Xa cleavage site, the effects on the molecule's functioning are obvious; for others they are not. Natural mutants of a protein have been important models in protein chemistry since the pioneristic work of Perutz on haemoglobin (Perutz, 1970). These mutants constitute, in fact, important tools for investigating structure-function relationships in a protein.

In a recent study (Akhavan et al., 2000), a natural variant of prothrombin containing the mutation Gly25 → Ser was identified in an Italian woman with prothrombin deficiency. Gly25 residue is located at the interface between A- and B-chain of thrombin, ~ 30 Å away the catalytic site, and not in thrombin recognition sites. The Gly to Ser amino acid substitution could be hypothesized to alter the correct folding process but not the function of thrombin.

However, the phenotype of the patient consisted of a mild decrease in prothrombin antigen (40%), with a much greater decrease in prothrombin activity (23%).

Several extensive mutagenesis studies have provided an important database of information on the critical residues involved in thrombin interaction with various substrates or ligands (Tsiang et al., 1995, Hall et al., 1999, Ayala et al., 2001, Pineda et al., 2004). However, the importance of this residue for the conformation and the function of this protease has not been yet evaluated.

In a previous study aimed to dissect out the role of Gly25 in thrombin, the G25S variant had been expressed in CHO cells (Akhavan et al., 2004). However, the expression level was 50% compared with wild-type thrombin, and only preliminary analyses had been conducted.

Since this has been a typical field of medical and clinical research, a little is known about molecular and biochemical mechanism responsible of thrombin molecules with altered folding and activity. A few functional studies, in fact, exist, but to our knowledge, neither solution and crystallographic structure studies of these natural variants exist. However, the characterization of these mutants represents a potent tool for the understanding of the correlation between the structure and the activity of thrombin *in vivo*, in order to establish the critical features for the normal function and specificity of this enzyme.

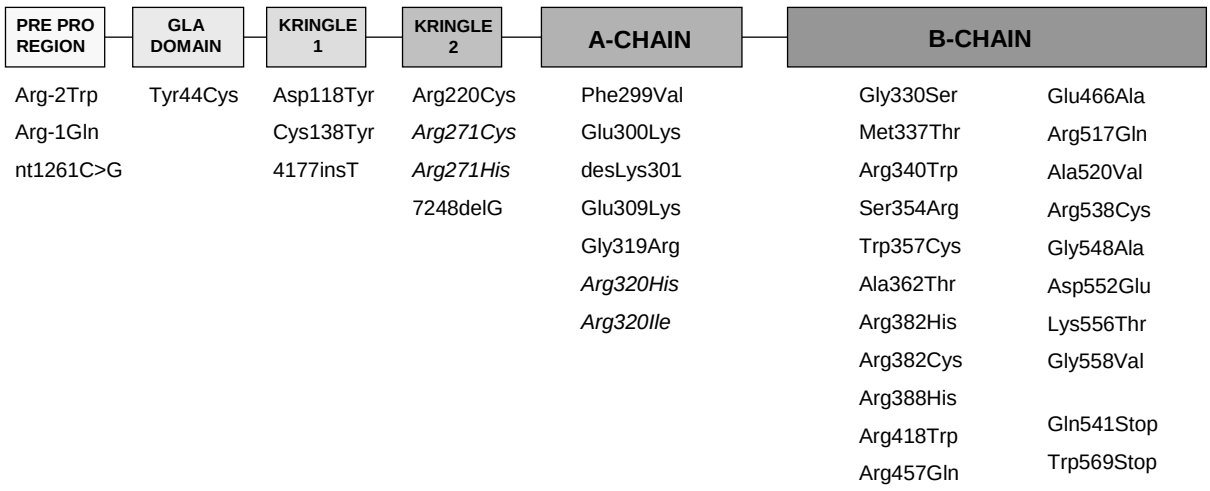


Fig. 1: Natural mutations identified in prothrombin gene.

The aim of the present project is the production of high quantity of Gly25Ser mutants in order to perform chemical, structural and functional studies with the aim of providing new insights into correlation between the chemical composition of thrombin mutants to structural modification of mutated enzymes and to their altered activity in vivo.

MATERIALS AND METHODS

Materials

The oligonucleotides were synthetized at Primm S.r.l. (Milan, Italy). The chromogenic substrate S-2238 (D-Phe-Pip-Arg-*p*-nitroanilide) was purchased from Chromogenix (Milan, Italy). Proteases, reagents for electrophoresis, buffers and organic solvents were of analytical grade and obtained from Fluka.

Site-directed mutagenesis, expression, refolding and purification of Gly25Ser

Mutation were introduced in the human prethrombin-2 gene by using the QuikChange® mutagenesis kit (Stratagene) according to the manufacturer's recommendations. Oligonucleotide 5'-TCG GAT GCA GAG ATC AGC ATG TCA CCT TGG C-3' was used to introduce a G to A point mutation coding for Ser instead of Gly25. The mutated insert was sequenced and sequencing confirmed that the desired mutation had been introduced. The homogeneity of the protein was established by SDS-PAGE, RP-HPLC and mass spectrometry analysis.

Wild-type (WT) and mutant prethrombin-2 were expressed and refolded, as previously described. The inactive purified protein was incubated for 1 h at 37° with snake venom from *E. carinatus* that had been pre-treated with PMSF to inhibit undesired serine protease activity. We used a 1:10 enzyme : substrate ratio in the case of WT thrombin, and a 1:5 enzyme:substrate ratio in the case of Gly25Ser mutant.

Steady-state kinetic of S-2238 substrate hydrolysis

Steady-state kinetic of the hydrolytic reaction of the D-Phe-Pip-Arg-*p*-nitroanilide (S-2238) substrate by WT and mutant thrombins was studied under the experimental conditions of 5 mM Tris-HCl, 0.2 M NaCl, PEG 8000 0.1%, pH 8 at 25°C. Use of polyethylene glycol 8000 prevents adsorbance of the enzyme to the cuvette walls and greatly increases the stability of the sample (De Cristofaro and Di Cera, 1990). Assays were performed by following the release of *p*-nitroaniline resulting from the hydrolysis of S-2238 at 405 nm. The concentration of released *p*-nitroanilide was measured using the extinction coefficient of 9,920 $M^{-1}cm^{-1}$ at 405 nm (Workman and Lundblad, 1978). The initial rates of S-2238

hydrolysis were measured for each of the mutants using 50 pM purified enzymes with different concentration of substrate. K_m and k_{cat} were determined by fitting the data to the Henri-Michaelis-Menten equation.

RESULTS AND DISCUSSION

In a recent study aimed at identifying novel natural variants of prothrombin, the Gly25 → Ser exchange was found. This amino acid substitution involves a glycine located ~ 30 Å away the catalytic site, at the interface between the A- and the B-chain, and not in thrombin recognition sites. A-chain of thrombin is composed of 36 amino acids, which are connected to the catalytic B-chain by a single disulfide bond, and its function has not yet been well established. In an old study carried out on bovine thrombin, it has been proposed that A-chain, although linked to B-chain through strong electrostatic and apolar bonds, does not play a major role in the specificity of the enzyme catalytic activity (Hageman et al., 1975). In contrast, a novel natural prothrombin mutant in which a lysine in A-chain is deleted, revealed an unexpected influence of this chain on the activity of human thrombin (De Cristofaro et al., 2002). With the aim of providing new insights into the function of the light chain and into the extraordinary conformational plasticity of thrombin, we had utilized the procedure optimized for wild-thrombin to produce high quantities of active Gly25Ser for following structural and functional studies.

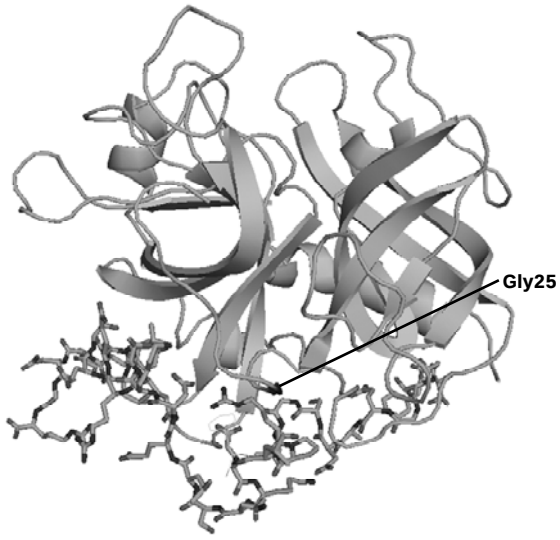


Fig. 2: Schematic representation of human α-thrombin (PDB code: 1ppb.pdb). B-chain is depicted in ribbon, A-chain in sticks. Gly25 residue is highlighted.

Expression, refolding and activation of Gly25Ser

Recombinant Gly25Ser was expressed in *E. coli*, according to previously reported procedure. The expression level of Gly25Ser was analyzed by SDS-PAGE under reducing conditions and measured with protein assay reagent (Bio-Rad Laboratories, California), based on the method reported by Bradford (Bradford, 1976). The expression level of Gly25Ser was found to be identical compared with wild-type thrombin (Figure 3A), and equal to 120 mg of protein each 0.8 litre of cell culture.

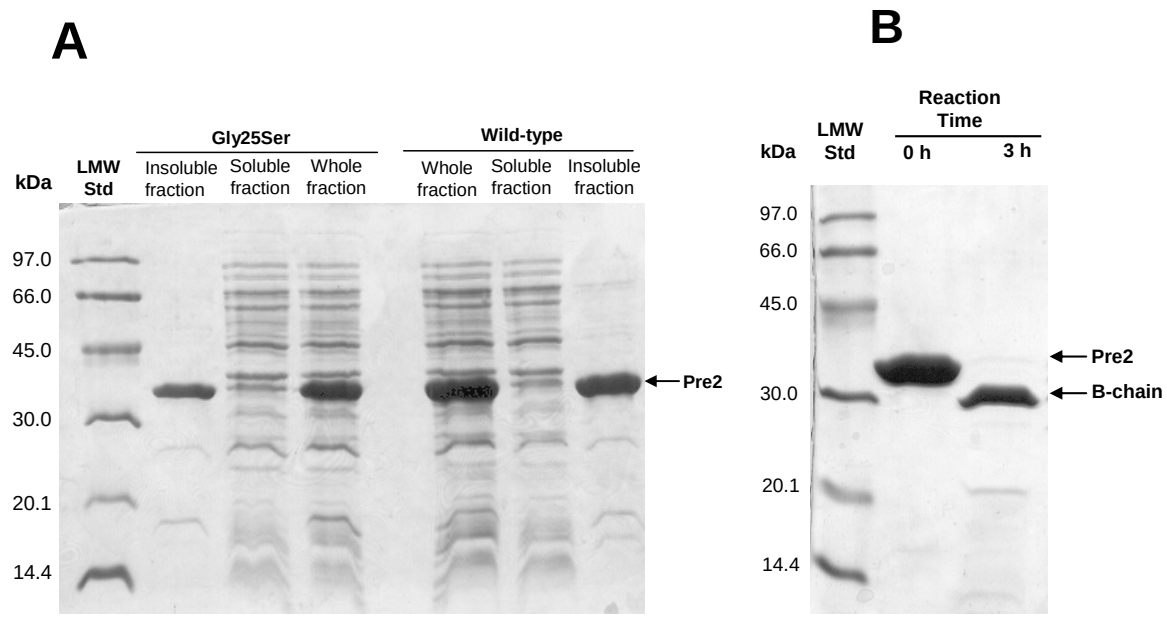


Fig. 3: (**A**) SDS-PAGE analysis of expression of wild-type and Gly25Ser thrombin in *E. coli*. (**B**) SDS-PAGE analysis of activation of Gly25Ser by ecarin. SDS-gel electrophoresis was performed on a 12% acrylamide gel and proteins were visualized by Coomassie blue staining.

Refolding was performed using the same procedure reported for wild-type thrombin. However, the refolding efficiency was at least 10-fold lower than wild-type. Moreover, activation with ecarin, which is know to cleave selectively Arg320-Ile321 peptide bond (Morita et al., 1976), has a 50% yield as shown by SDS-PAGE, and it is less efficient than for wild-type thrombin. Re-purification of active protein by heparin-sepharose chromatography showed that this mutant has a dramatically lower solubility. In fact, the material eluted from the affinity column is more than 10-fold reduced with respect to the expected, and it is not in agreement with the yield determined by SDS-PAGE.

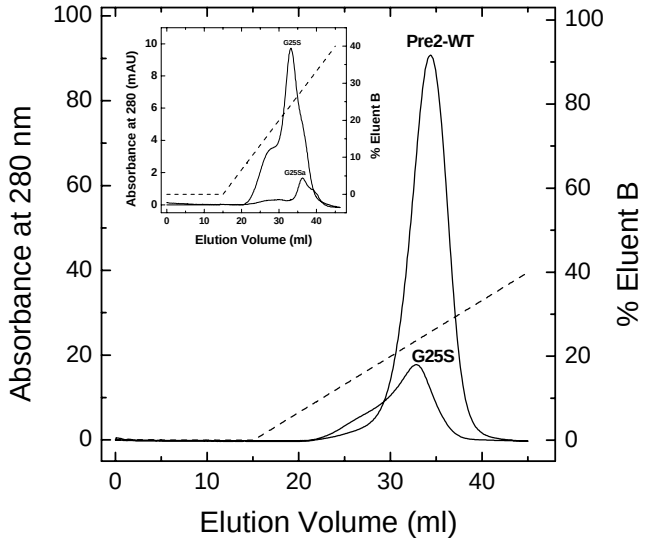


Fig. 4: Heparin-sepharose purification of Pre2-Gly25Ser relative to wild-type thrombin. *Inset:* heparin-sepharose purification of the activated Gly25Ser mutant relative to the starting material of prothrombin-G25S.

Kinetic of S-2238 hydrolysis by thrombins

We compared the enzymatic properties of Gly25Ser with those of plasma-derived α-thrombin by determining the steady-state parameters for the hydrolysis of the chromogenic substrate S-2238. As shown in Table 1, the k_{cat} (22.6 ± 1.2 s^{-1}) determined for the hydrolysis of S-2238 by Gly25Ser is characterized by a 4-fold decrease relative to k_{cat} (91.4 ± 4.6 s^{-1}) determined for plasma-derived α-thrombin. Conversely, the K_m (3.7 ± 0.2 μM) is only 1.8-fold lower with respect to wild-type (2.1 ± 0.1 μM), with an overall decrease in K_s of 7-fold (6.1 ± 0.3 $\mu M^{-1} \cdot s^{-1}$ versus 43 ± 2.2 $\mu M^{-1} \cdot s^{-1}$). The observed decrease suggests that the geometry of the catalytic residues is altered in Gly25Ser, while the binding of the substrate is substantially unaltered.

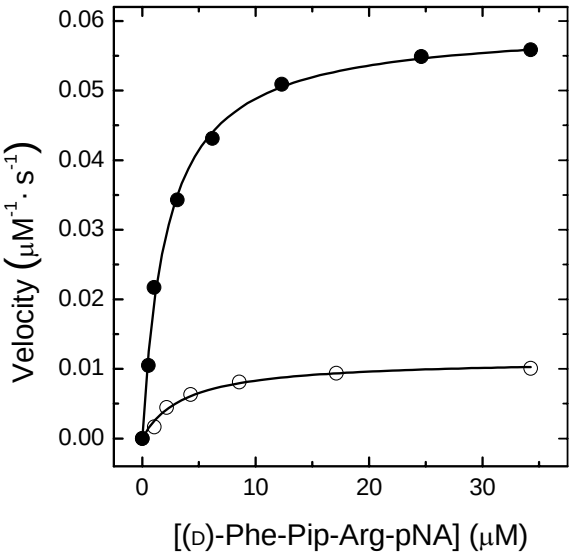


Fig 5: Henri-Michaelis-Menten plot for S-2238 hydrolysis by wild-type α-thrombin (●) and Gly25Ser (○). The assays were conducted at 25°C in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% (w/v) PEG 8000 and 0.2 M NaCl by monitoring the release of *p*NA form S-2238, and recording the increase in the absorbance at 405 nm.

Table 1: Best-fit K_m and k_{cat} Values Pertaining to the Hydrolysis of the (D)-Phe-Pip-Arg-pNA Substrate.

	Gly25Ser	Wild-type
K_m (μM)	3.7 ± 0.2	2.1 ± 0.1
k_{cat} (s^{-1})	22.6 ± 1.2	91.4 ± 4.6
k_{cat}/K_m ($\mu M^{-1} \cdot s^{-1}$)	6.1 ± 0.3	43 ± 2.2

CONCLUSIONS

In conclusion, Gly25 is important for proper expression *in vivo* and probably proper folding of the protein. This residue is located in a turn structure, and a Gly seems to be the only appropriate residue at this location. Moreover, the more bulky and polar serine can constitute a hindrance in the bending of the N-terminus of prethrombin-2 polypeptide chain that leads to the formation of the disulfide bond between Cys1 and Cys122.

This mutation confirms, once more, the extraordinary plasticity of thrombin, where a mutation at the interface between the A- and B-chain exerts a detrimental effect on the catalytic site. Future studies involve the production of high quantities of this mutant for the determination of the structure by X-ray diffraction.

CHAPTER 2.4

Production of Thrombin Mutants Containing Arg → Ala Substitutions In Exosite I and II

INTRODUCTION

The substrate specificity of thrombin is regulated by binding of macromolecular substrates and effectors, to two electropositive sites, exosite I and II, located in near-opposition on the enzyme surface. Exosites are sites on proteinase physically separated from the residues that determine the primary S1-S4 peptide substrate and inhibitor specificity. In thrombin, exosite I binds fibrinogen (Naski M.C., et al. 1990), fibrin I and II (Naski and Shafer, 1991), the carboxi-terminal domain of hirudin (Rydel et al., 1991), thrombomodulin (Ye et al., 1992), the thrombin receptors (Liu et al., 1991), and an acidic sequence on the serpin, heparin cofactor II (Van Deerlin, 1991). On the other side, exosite II binds heparin and other glycosaminoglycans (Gan et al., 1994), prothrombin activation fragment 2 (Arni et al., 1993), the chondroitin sulphate moiety of thrombomodulin (Liu et al., 1994), the leech peptide hemadin (Richardson et al., 2000), the fibrinogen γ' chain and an exosite II specific human monoclonal antibody (Colwell 1998). Factor Va, and VIII, and platelet glycoprotein Ibα have been reported to interact with both exosites (Verhamme et al, 2002). Binding of ligands to thrombin exosites is correlated with significant changes (either enhancement or inhibition) in the kinetic of hydrolysis of small chromogenic substrates. In fact, previous studies indicate that exosite binding of allosteric effectors is coupled to conformational changes affecting the S1-S3 substrate specificity subsites in the thrombin catalytic site. Importantly, all natural thrombin substrates interact with at least one exosite, either directly or through cofactor mediation (Huntington et al., 2008).

The study of exosite-mediated thrombin recognition has benefited greatly from the development of methods of site-directed mutagenesis and protein expression. In fact, numerous extensive mutagenesis studies of positively-charged residues in thrombin surface(i.e. Arg and Lys) had be performed, in order to map the secondary binding sites required for fibrinogen clotting, protein C activation (Tsiang et al., 1999, Wu et al., 1991),

serpin binding (Myles et al., 1998), TAFI activation (Hall et al., 1999), PAR-1 hydrolysis (Myles et al., 2001), thrombomodulin binding (Pineda et al., 2002), Gp1b binding (De Candia et al., 2001, De Cristofaro et al., 2001), and heparin inhibition (Sheehan and Sadler, 1994, Gan et al., 1994, Tsiang et al., 1997). In addition to these data, a large number of crystal structures of thrombin in complex with substrates, cofactors and inhibitors, allowed to strictly identify residues responsible of thrombin exosite-mediated specificity (Figure 1).

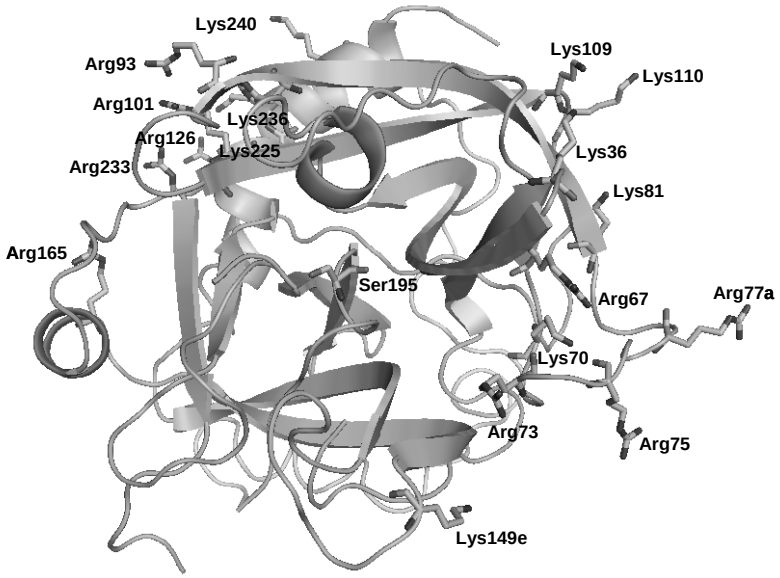


Fig. 1: Schematic representation of positive residues that constitute exosite I and exosite II of thrombin. Exosite I is shown at the right hand side, whereas exosite II is shown at the left hand side. The catalytic Ser195 is highlighted at the centre of the molecule and A-chain is on the back.

When a new modulator of thrombin is identified, it's important to define which exosite is involved in the recognition between the two proteins. Three techniques are worldwide used with the aim of unambiguously defining interaction sites, as well as hot spots, involved in protein-protein interactions. One approach relies on determination of 3D structure of the complex, through the application of X-ray crystallography; a second approach relies on the use of NMR techniques. A popular alternative strategy for identifying these sites is the mutation of key amino acids to alanine, a method known as alanine-scanning mutagenesis. In this approach, the side chain of amino acids located on the surface of the receptor are replaced with alanine by oligonucleotide-directed mutagenesis. Alanine, in fact, is compatible with both α- and β- secondary structures, is tolerated in both buried and exposed locations in proteins, and the non-polarity and small size of its side chain ensures that substitution with alanine is less likely to disrupt protein conformation (Tsiang et al. 1995). With these assumptions, it is reasonably to expect that mutation to alanine of residues at important ligand

recognition site will significantly perturb formation of the complex without compromising the properties of the free form of the receptor. The resulting loss in binding affinity can then be used to assess the contribution of the mutated residue to ligand binding (Greenspan and Di Cera, 1999).

The goal of this project is therefore the design, production and characterization of two distinct thrombin mutants unable to bind respectively to exosite I- or exosite II- specific binders, but with an intact catalytic machinery if compared with the wild-type enzyme. To this purpose a hot-spot residue in each exosite must be selected and substituted with alanine.

Considering mutagenesis and structural data reported so far, we choose the substitution Arg to Ala at position 73 and the substitution Arg to Ala at position 101 as representative, respectively, of a thrombin molecule with defective exosite I and II. Arg73 is located in the exosite I, and, together with Arg67, has been proved to be critical for fibrinogen clotting, and complex formation with hirudin, PAR-1(38-64) and HCII (Myles et al., 2001). In particular, the crystal structures of thrombin bound to hirudin (Rydel et al., 1991) and to PAR-1 peptide 38-64 (Mathews et al., 1994) indicate that this residue forms a hydrogen-bonded salt-bridge with the hirudin residue Asp55 and the PAR-1 residue Asp50. The importance of this residue in physiological interaction of thrombin is highlighted by the recent finding that the natural thrombin variant R73H is associated with mild bleeding tendency *in vivo* (d'Audigier et al., 2008), even if this mutated enzyme shows normal hydrolysis of the synthetic substrate S-2238. On the other side, Arg101 is located in the exosite II, and its importance in the binding of heparin has been assessed by two separate mutagenesis studies (Gan et al., 1994, Tsiang et al., 1997) and by the crystal structure of thrombin bound to an eight-unit heparin fragment (Carter et al., 2005). Moreover, the mutant R101A had been recently used to demonstrate that Gp1bα, which binds with high-affinity to α-thrombin, interacts with exosite II. A natural variant had been shown to occur at position 101, in which Arg is substituted with Trp (James et al., 1995). However, this substitution perhaps doesn't interfere with heparin binding, but perturbs the active site, since the more bulky tryptophan occludes the catalytic site.

In the present work, the mutant forms R73A and R101A, representative of a exosite I-, and a exosite II-defective α-thrombin, respectively, were produced at high yield in *E. coli*. These mutants were characterized by chemical techniques, and their enzymatic activity on a chromogenic substrate was studied and compared with wild-type enzyme. Moreover, the perturbed exosites were studied with ligand specific for each exosite (i.e. full-length hirudin and its C–terminal domain for exosite I, heparin and fibrinogen γ' peptide for exosite II) to evaluate the degree of exosite perturbation inferred by the substitutions thus introduced .

MATERIALS AND METHODS

Materials

Oligonucleotides for mutagenesis were synthetized at Primm S.r.l. (Milan, Italy). The chromogenic substrate S-2238 (D-Phe-Pip-Arg-*p*-nitroanilide) was purchased from Chromogenix (Milan, Italy). Proteases, reagents for electrophoresis, buffers and organic solvents were of analytical grade and obtained from Fluka.

Site-directed mutagenesis, expression, refolding and purification of recombinant mutant thrombins

Mutations in the human prethrombin-2 gene were introduced by using the QuikChange® mutagenesis kit (Stratagene) according to the manufacturer's recommendations. Oligonucleotide 5'-gggagaacctggacgcggacattgccctga-3' was used to introduce a CG to GC, coding for Ala instead of Arg101 and 5'-attggcaagcactccgccacaaggtacgagcg-3' was used to introduce a CG to GC, coding for Ala instead of Arg73. The mutated inserts were sequenced and sequencing confirmed that the desired mutations had been introduced. Prethrombin-2 mutants were expressed, refolded, purified and activated as previously described for wild-type thrombin. The concentration of WT and mutant thrombins were estimated by measuring the absorbance at 280 nm and the samples were immediately frozen until use. Experiments were carried out of one preparation each of recombinant thrombins.

Chemical Characterization of WT and Alanine-Mutants

The homogeneity of the proteins was established by SDS-PAGE and RP-HPLC (Figure 2). Standard PAGE in the presence of β-mercaptoethanol was performed use a 12% polyacrylamide gel. RP-HPLC analysis were carried out on a C4 analytical column (4.6 x 150 mm, 5 μm particle size, 300 Å porosity, from Grace Vydac). The column was equilibrated with 0.1% H_2O-TFA and eluted with a linear 0.1% TFA-acetonitrile gradient at a flow rate of 0.8 ml/min. The absorbance of the effluent was recorded at 226 nm. The chemical identity of the purified proteins was established by ESI-TOF mass spectrometry on a Mariner instrument from Perseptive Byosystem. Typically protein samples (5 μg) were analysed in acqueos-acetonitrile (1:1, v/v), containing 1% (v/v) formic acid at a flow rate of 15 μl/min. The nozzle

temperature was set up at 140°C and the electrostatic potential at 4.4 kV. The instrument was calibrated using the standard protein kit from Sigma.

Functional Characterization of WT and Alanine-Mutants

Steady-state kinetic of S-2238 substrate hydrolysis. Steady-state kinetic of the hydrolytic reaction of the D-Phe-Pip-Arg-*p*-nitroanilide (S-2238) substrate by WT and mutant thrombins was studied under the experimental conditions of 5 mM Tris-HCl, 0.2 M NaCl, PEG_{8000} 0.1%, pH 8 at 25°C. All assays were performed by following the release of *p*-nitroaniline resulting from the hydrolysis of S-2238 at 405 nm, as previously detailed. The initial rates of S-2238 hydrolysis were measured for each of the mutants using 50 pM purified enzymes with 0.75-24 μM substrate. K_m and k_{cat} were determined by fitting the data to the Henri-Michaelis-Menten equation. The enzyme active-site concentration was previously titrated with hirudin HM2.

Inhibition of recombinant thrombins by hirudin HM2. K_d values of hirudin HM2 for WT and thrombin mutants were estimated by classical competitive inhibition experiments of thrombin-mediated substrate hydrolysis, according to the tight-binding inhibition model (Copeland, 2000), as previously detailed. This method has been described in detail by Di Cera and co-workers (Ayala and Di Cera, 1994).

Affinity of recombinant thrombins for heparin-Sepharose. Heparin affinity was preliminary assessed by elution from a heparin-Sepharose column. Briefly, WT and mutant thrombins in 50 mM Tris-Cl, pH 7.4, 0.12 M NaCl, were applied at 2 ml/min to a 5 ml HiTrap heparin-Sepharose column (GE Lifescience) pre-equilibrated in the same buffer. The column was eluted at 2 ml/min with a linear-gradient of NaCl. Thrombin was detected by absorbance at 280 nm and the conductivity of the peaks were measured by AKTA purifier system (GE Lifescience).

Binding of WT and Mutant Thrombins to Fibrinogen γ'-peptide. Binding of γ' peptide was studied by recording the increase in tryptophan fluorescence of thrombin at λ_{max} (i.e. 334 nm) as a function of fibrinogen γ'-peptide. The interaction of the latter with thrombin was monitored by adding, under gentle magnetic stirring, to a solution of thrombin (1.4 ml, 50 nM) in 5 mM Tris-HCl buffer, pH 7.5, 0.1% PEG 8000, 0.15 M NaCl, aliquots (2-5 μl) of γ'

peptide. Fluorescence spectra were recorded on a Jasco (Tokyo, Japan) model FP-6500 spectrofluoromether, equipped with a Peltier model ETC-273T temperature control system from Jasco. Excitation and emission wavelength were 295 and 334 nm, respectively. For all measurements, the long time measurement software (Jasco) was used. Under these conditions, at the end of the titration, a Trp photobleaching lower than 2% was observed. The absorbance of the solution at both 295 and 334 nm was always lower than 0.05 unit, and therefore no inner filter effect occurred during titration experiments. Fluorescence intensities were corrected for dilution (2-3% at the end of titration) and subtracted for the contribution of the ligand at the indicated concentration. The fluorescence values, measured in duplicate, were analyzed as a function of the γ'-peptide concentration by a hyperbole equation to obtain the value of the F_{max} (corresponding to the fluorescence at γ' peptide concentration = ∞). This parameter was used to calculate $\Delta F_{max} = F_{max} - F°$ (where F° is the fluorescence value in the absence of the peptide). The fluorescence changes expressed as $(F_{obs} - F°)/\Delta F_{max}$ were analyzed as a function of the total γ' peptide concentration according to a single site binding isotherm. Non linear least squares fitting was performed using the program Origin 7.5 (MicroCal Inc.), which allowed us to obtain the best fitting parameter values along with their standard errors.

Inhibition of recombinant thrombins by hirudin HM2 (54-65)P_2O_3. The interaction of HV1 (54-65)P_2O_3 with thrombins was monitored by adding, under gentle magnetic stirring, to a solution of thrombin (1.8 mL, 50 nM), in 5 mM Tris-HCl buffer, pH 7.5, 0.1% PEG 8000, in the presence of 0.15 M NaCl, aliquots (2-10 μL) of inhibitor stock solutions (2-212 μM) in the same buffer. At each inhibitor concentration, protein samples were equilibrated for 2 min at 25°C and excited at 295 nm, using an excitation/emission slit of 10 and 10 nm, respectively. The increase in fluorescence intensity at the λ_{max} (334 nm) of thrombin was recorded as a function of inhibitor concentration. Fluorescence data were corrected for sample dilution (< 5% of the final volume) and expressed as $(F°-F)/\Delta F_{max}$, where $F°$ and F is the fluorescence of thrombin in the absence and presence of the inhibitor, respectively, and ΔF_{max} is the maximum fluorescence change at saturating concentrations of inhibitor. For a simple one-site binding mechanism $R + L \Leftrightarrow RL$, the fluorescence intensity (F) of the receptor (R) at a given concentration of ligand (L) is linearly related to the concentration of the complex $[RL]$, $F = [RL]\cdot F_{bound} + [R]_{free}\cdot F_{free}$. Since $[R]_{free} = [R] - [RL]$, then $(F°-F)/\Delta F_{max} = [RL]/[R]$ (Eftink 1997). The data were fitted to following equation (Copeland 2000), using the program Origin 6.0 (MicroCal Inc.):

$$[RL]/[R] = \{([R] + [L] + K_d) - \{([R] + [L] + K_d)^2 - 4\cdot[R][L]\}^{1/2}\}/2\cdot[R]$$

where K_d is the dissociation constant of complex and *[R]* is the total concentration of the receptor.

RESULTS AND DISCUSSION

Purification and Chemical Characterization of WT and Alanine-Mutants.

In a previous chapter, a method for obtaining large quantities of pure and active human thrombin by *E. coli* expression, *in vitro* refolding and heparin-sepharose purification was developed and optimized. The recombinant enzyme has been extensively characterized by chemical, spectroscopic, and functional techniques, and its properties fully overlap with the natural counterpart.

The present work describes the production by site-directed mutagenesis and expression in *E. coli* cells of two thrombin mutants, Arg73Ala and Arg101Ala, which bear a perturbation in fibrinogen-binding-site and heparin-binding-site, respectively. Previous mutagenesis and structural data show that these two residues are engaged in important interactions with thrombin substrates, cofactors, and inhibitors.

Recombinant WT and alanine-mutant proteins were refolded, purified and activated, according to the previously reported procedure. The correct production of mutant active thrombins was analysed by SDS-PAGE and RP-HPLC (see Figure 2). In SDS-PAGE analysis, all proteins migrates as a single band, with apparent molecular masses of 30 kDa, in agreement with the B-chain weights. Moreover, either before and after activation, recombinant proteins had an electrophoretic mobility identical to WT thrombin. The homogeneity of recombinant proteins was also assessed by RP-HPLC, and was found to be higher than 95%. The chemical identity of recombinant proteins was established by ESI-TOF mass spectrometry, which gave mass values in agreement with the expected amino acid composition within 50 ppm accuracy (Table 1). Moreover, the difference in average molecular mass between mutants and the wild-type enzyme is in close agreement with a Arg to Ala substitution (i.e., 85.1 a.m.u.).

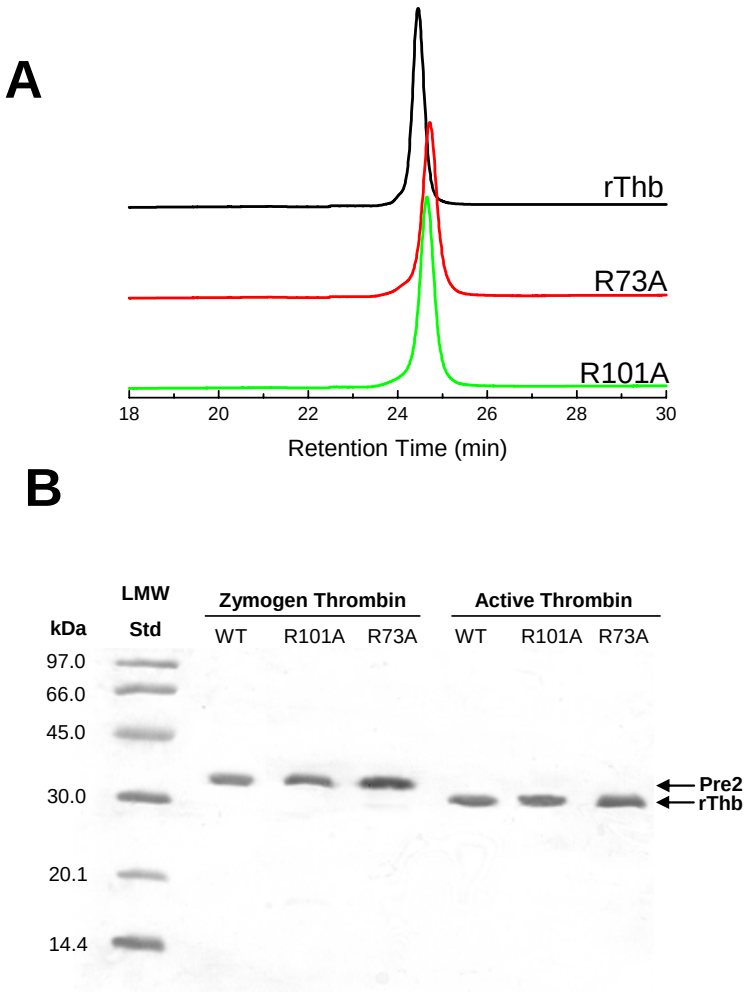


Fig. 2. Chemical analysis of purified wild-type thrombin and Ala-mutants. Wild-type human thrombin and R73A and R101A mutants were purified as described in Methods. **(A)** RP-HPLC analysis of wild-type thrombin, R73A and R101A at 25°C in the active form. **(B)** Wild-type thrombin and Ala-mutants were analyzed by SDS-PAGE: lanes 1, 2,3,4 WT, R73A and R101A prethrombin-2; lanes 5,6,7 WT, R73A, R101A. Molecular mass standards for SDS-PAGE (in kDa) are indicated in lane 1.

Table 1: Average Molecular Mass Values of WT and Ala-Mutants.

Thrombin Forms	Theoretical Molecular Weight (a.m.u)	Experimental Molecular Weight (a.m.u.)
rThb	33820.7	-
R73A	33735.6	33734.2 ± 3.3
R101A	33735.6	33735. 5 ± 3.3

Effect of Mutations on Enzymatic and Recognition Properties of Thrombin

In order to verify the enzymatic and recognition properties of thrombin mutants, their catalytic activities were investigated by classical chromogenic assays, whereas the structure of

exosite I and exosite II was studied by different probes like hirudin, its C-terminal fragment 54-65, heparin and fibrinogen γ'-peptide. In particular, hirudin HM2 covers the catalytic pocket of thrombin and binds to exosite I with its C-terminal chain (54-65), whereas heparin and fibrinogen γ'-peptide specifically bind to exosite II. Changes in affinity of these ligands are expected to provide key information on the conformational state of their putative binding sites in Ala-mutants.

Steady-state kinetics of amidase activity toward S-2238. The recombinant WT and mutant thrombins showed similar Michaelis-Menten values pertaining to the hydrolysis of the (D)-Phe-Pip-Arg-pNA substrate as showed in Figure 3 and listed in Table 2. These experiments clearly showed that mutations did not alter the catalytic machinery of the enzyme. Any other modification of the amidase activity of mutated thrombins in the presence of a ligand, will be thus attributed to alterations concerning the perturbed enzyme exosites, involved in macromolecular ligand interactions.

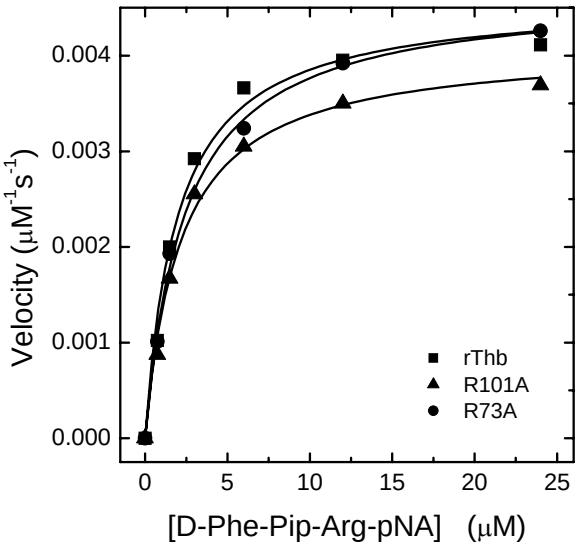


Fig. 3. Henri-Michaelis-Menten plot for S-2238 hydrolysis by wild-type α-thrombin (■) and Arg73Ala (●) and Arg101Ala (▲) mutants. The assays were conducted at 25°C in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% (w/v) PEG 8000 and 0.2 M NaCl by monitoring the release of *p*NA from S-2238, recoding the increase in the absorbance at 405 nm. The reaction was started by addition of thrombin solution (50 pM) to different concentrations of S-2238, preincubated in cuvettes at 25°C.

Table 2: Best-fit K_m and k_{cat} Values Pertaining to the Hydrolysis of the D-Phe-Pip-Arg-pNA (S-2238) Substrate by the WT and Mutant Thrombin Forms

Thrombin Forms	K_m S-2238 (μM)	k_{cat} S-2238 (s^{-1})	k_{cat}/K_m S-2238 ($\mu M^{-1} \cdot s^{-1}$)

rThb	1.9 ± 0.2	92 ± 3	47.4 ± 2.1
R73A	2.4 ± 0.3	93.4 ± 2.1	38.9 ± 1.3
R101A	2.1 ± 0.2	82.7 ± 3.2	38.2 ±1.8

Exosite I-Structure Evaluation

Inhibition of WT and Ala-mutants by hirudin HM2. As reported previously, hirudin can be a suitable probe for studying the structures of active-site and exosite I of thrombin. K_d values of hirudin HM2 for WT and thrombin mutants were estimated by classical competitive inhibition experiments of thrombin-mediated substrate hydrolysis.

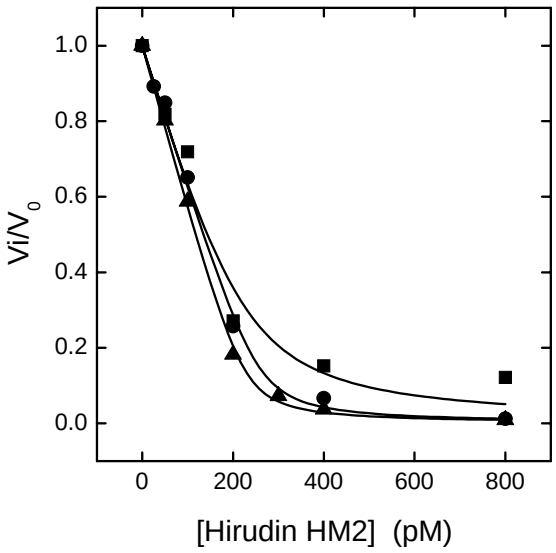


Fig. 4: Plot of fractional velocity of WT -type α-thrombin (■) and Arg73Ala (●) and Arg101Ala (▲) mutants. as a function of HM2 peptide concentration. The assays were conducted at 25°C in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% (w/v) PEG 8000 and 0.2 M NaCl, by monitoring the release of *p*NA form S-2238, recording the increase in the absorbance at 405 nm. The reaction was started by addition of the substrate (56 μM) to a thrombin solution (250 pM), preincubated with increasing concentrations of hirudin HM2 (see Methods).

The inhibition (Figure 4) of the exosite-II mutant R101A showed no major differences relative to WT thrombin (0.2 ± 0.1 pM versus 0.3 ± 0.1 pM). On the other side, the K_d of hirudin for the exosite-I mutant R73A was about 10-fold higher (3.3 ± 1.1 pM versus 0.3 ± 0.1 pM). Since the analysis of S-2238 hydrolysis gave K_m values similar for all thrombin

forms, this is a strong indication that the reduction in dissociation constant observed for Arg73Ala must be ascribed only to the binding of the C-terminal tail of hirudin to perturbed exosite I. On this basis, we successively analysed the binding of the isolated C-terminal chain of hirudin to thrombin.

Binding of WT thrombin and Ala-mutants to hirudin HV1 (54-65)P_2O_3. Hirudin HV1 (54-65)P_2O_3 peptide is rich in acidic residues (GDFEEIPEEY(PO_3H_2)LQ) and binds exclusively to exosite-1 of thrombin. This ligand is expected to bind with similar affinity to wild-type and Arg101Ala mutants, since the last has a binding site with the native structure. On the other side, the shaving of arginine side chain to Ala at position 73 removes a major positive charge for the electrostatic interaction of residue Asp55 of this ligand with thrombin. We studied the binding by monitoring the increase in fluorescence emission of thrombin at increasing concentration of HV1(54-65)P_2O_3 peptide. Binding experiments showed that Arg73 plays a central role in the interaction with HV1(54-65)P_2O_3, because the K_d decreases 22-fold in the Arg73Ala mutant relative to WT thrombin, as shown in Figure 5 (4.655 ± 0.026 μM versus 0.213 ± 0.005 μM). On the other side, the K_d for hirudin of Arg101Ala remained essentially unchanged (0.162 ± 0.011 μM versus 0.213 ± 0.005 μM), in agreement with the fact that exosite II is not engaged in thrombin interaction with hirudin.

Exosite II-Structure Evaluation

Binding of WT and mutant thrombins to heparin-sepharose chromatography. Heparin is a glycosaminnoglycan, required in vivo for efficient inhibition of thrombin by antithrombin. Heparin binding to thrombin is highly dependent fron ionic strength, and occurs at exosite-II. In particular, five alkaline residues of thrombin exosite II (Arg93, Arg101, Arg233, Lys236 and Lys240) are primarily involved in the complex formation. Binding of recombinant mutants to heparin, as assessed by elution from heparin-Sepharose column, showed a large decrease in affinity by R101A over R73A and WT thrombin (Table 3).

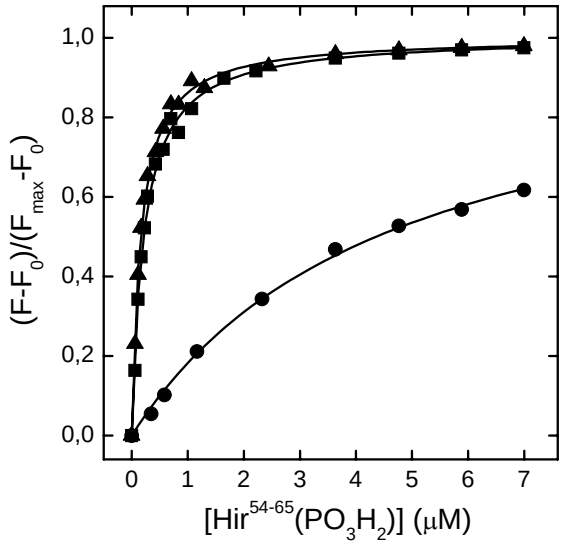


Fig. 5. Binding of hirudin HV1 (54–65)PO_3H_2 to α-thrombin (■) and Arg73Ala (●) and Arg101Ala (▲) mutants monitored by tryptophan fluorescence. The interaction of hirudin (54–65)PO_3H_2 with thrombin was monitored by adding, under gentle magnetic stirring, to a solution of thrombin (1.8 ml, 50 nM) in 5 mM Tris-HCl buffer, pH 7.5, containing 0.1% PEG 8000, and 0.15 M NaCl, aliquots (2–5 ul) of hirudin (54–65)PO_3H_2 (2-212 μM). The solid lines represent the least square fits, with best fit parameters K_d = 0.213 ± 0.005 μM (rThb), K_d = 4.655 ± 0.026 μM (R73A) and K_d = 0.162 ± 0.011 μM (R101A).

Thus, mutation of residue Arg73 in anion-binding exosite I does not affect the affinity of thrombin for heparin, while the Arg101 to Ala substitution in exosite II heavily compromises the binding of this mutant to the resin. Furthermore, the observed effect is specific for the removal of a positive charge in the exosite II, since Arg73Ala binds to the resin with the same affinity of wild-type protein.

Table 3: Heparin affinity of the various forms of thrombin

Thrombin Forms	**Conductivity** (mS/cm)
Pre2	63,7
rThb	73,5
Pre2-R73A	63,5
R73A	73,4
Pre2-R101A	48,5
R101A	53,5

Binding of WT and mutant thrombins to fibrinogen γ' peptide. Fibrinogen γ' chain is a natural variant of fibrinogen gamma chain that originates from an alternative mRNA splicing between exons 9 and 10, resulting in an elongated chain composed of 427 instead of 411 residues. The inserted region at the C-terminus shows a highly anionic C-terminal sequence (^{408}VRPEHPAETEYDSLYPEDDL^{427}), that binds to the exosite II of thrombin (Lancellotti et al. 2008). The crystal structure of the complex of α-thrombin with γ'-peptide (Pineda et al., 2007; PDB code: 2HVL.ppb), shows that γ' peptide makes many hydrogen bonds and salt bridges with exosite-II residues. Arg101 is involved in this bond network because the carboxylate group of residue Asp419 in γ' peptide makes a salt bridge with the guanidyl side chain of thrombin R101. Hence, we use this ligand as a structural probe of exosite II conformation. The binding of γ' peptide to thrombin causes a significant increase of tryptophan fluorescence (~24%), without appreciable change in the λ_{max} value. Therefore, we studied the binding by monitoring the increase in fluorescence emission of thrombin at increasing concentration of γ' peptide. The K_d value for γ' peptide binding to R73A is 9.03 ± 0.26 μM and it was found to be similar to that determined for WT enzyme (i.e. 6.23 ± 0.15 μM), whereas, the binding to R101A is severely impaired, with a dissociation constant at least 70-fold higher (Figure 6).

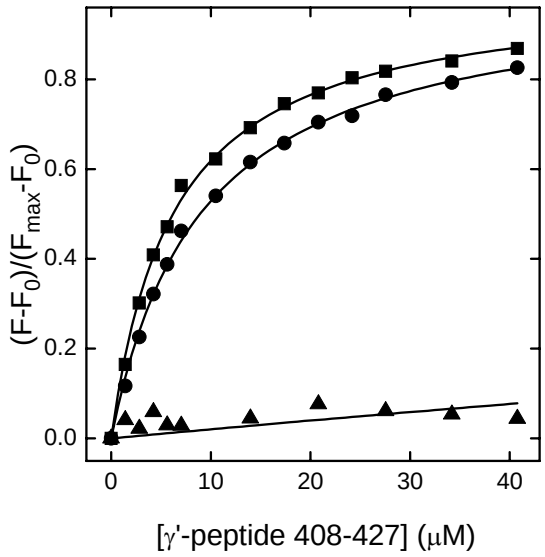


Fig. 6. Binding of γ' peptide to thrombin monitored by tryptophan fluorescence. Plot of the fluorescence intensity of thrombin as a function of γ' peptide concentration The interaction of γ' peptide with thrombin was monitored by adding, under gentle magnetic stirring, to a solution of thrombin (1. 5 ml, 50 nM) in 5 mM Tris-HCl buffer, pH 7.5, containing 0.1% PEG 8000, and 0.15 M NaCl, aliquots (2–5 ul) of γ' peptide (1.06 mM). The *solid lines* represent the least square fits, with best fit parameters K_d =

6.23 ± 0.15 μM (rThb) and K_d = 9.03 ± 0.26 μM (R73A). The K_d for R101A mutant was too high to be estimated under these experimental conditions.

These findings are in agreement with the fact that γ' peptide binds to exosite II, whereas exosite I is not engaged in thrombin interaction with this ligand. Moreover, shaving of Arg101 side chain removes an important positive charge for interaction with γ', leading to a perturbed conformation of exosite-II.

CONCLUSIONS

Together these findings underline, once again, that protein engineering is a powerful technique that permit to selectively modulate recognition properties of thrombin. We demonstrated that the shaving of the basic side chains of residues Arg101 and Arg73 severely perturbed thrombin exosite-mediated binding to important modulators of its activity, without altering catalytic machinery. In particular, Arg73Ala mutant shows drastic decrease in affinity for ligand specific for exosite I (i.e., hirudin C-terminal chain), whereas Arg101Ala shows less affinity for ligand specific for exosite II (i.e., heparin and γ' peptide). Hence, these mutants could be used as important tools in the study of novel ligands to thrombin with the aim of mapping which exosite is responsible in the formation of the complex.

3. Results

Protein Engineering by Chemical Methods

CHAPTER 3.1

o- Nitrotyrosine as a Spectroscopic Probe for Investigating Thrombin-Hirudin Interaction

INTRODUCTION

The development of new spectroscopic tools for studying protein-protein interactions is central to many disciplines, including structural biology, biotechnology, and drug discovery (Hovius et al. 2000). Traditionally, the change in tryptophan (Trp) fluorescence has been exploited to study ligand-protein interactions (Eftink, 1997). However, the fluorescence signal of many proteins is insensitive to ligand binding (Lakowicz 1999; Jameson et al. 2003), since fluorescence changes are mostly restricted to those cases where Trp-residues are embedded in the ligand-protein interface or when the ligand binding induces conformational changes in the protein, remote from the binding region and involving one or more Trp-residues. Furthermore, the presence of multiple tryptophans in proteins may lead to compensating effects that often complicate a unique structural interpretation of the fluorescence data (Eftink 1997).

To overcome these problems, several extrinsic spectroscopic probes, characterized by well-defined spectral properties, have been covalently bound to protein functional groups (*i.e.*, Cys and Lys), to act as energy donors or acceptors in fluorescence resonance energy transfer (FRET) studies (Wu and Brand 1994; Selvin 1995; Hovius et al. 2000; Yan and Marriott 2003). This approach, however, is limited by possible labeling heterogeneity, non-quantitative modification, structural alteration of the proteins resulting from the labeling *per se*, and perturbation of the binding process, due to the large size of the fluorescent labels used (Wu and Brand, 1994; Selvin, 1995). In the last decade, advances in both peptide synthetic chemistry (Dawson and Kent 2000; Albericio 2004; Nilsson et al., 2005) and genetic techniques (for recent reviews, see Dougherty 2000; Cropp and Schultz, 2004, England, 2004; Hendrickson et al., 2004) made the incorporation of noncoded amino acids into even long polypeptide chains a feasible task. With respect to this, several unnatural amino acids, possessing physico-chemical properties (e.g., size, polarity, hydrogen bonding properties)

similar to those of the corresponding natural amino acids, but distinct spectral features, were site-specifically incorporated into proteins to monitor key events, such as protein folding and binding (Cornish et al., 1994; Twine and Szabo, 2003; De Filippis et al., 2004).

Among the noncoded analogues of tyrosine, 3-nitrotyrosine (NT), which is produced *in vivo* by reaction of protein tyrosines with peroxynitrite (Halliwell, 1997), displays very interesting structural and spectroscopic properties. The NT side-chain is only 30 $Å^3$ larger than the unmodified Tyr and the presence of the electron-withdrawing nitro-group makes the phenolic hydrogen of free NT about 10^3-fold more acidic (pK_a 6.8) (Riordan et al., 1967). At pH < pKa, where the neutral form is predominant, NT is more hydrophobic than Tyr, whereas at higher pH, where NT exists in the ionized form, it is much more polar (Csizmadia et al., 1997; Abraham et al., 2000). NT can form an internal hydrogen bond (see Figure 1A) and its absorption properties are strongly pH-dependent. In particular, at basic pH, the UV/Vis spectrum of free NT displays a major band at 422 nm, characteristic of the ionized form, whereas at acidic pH a prominent band appears at 355 nm, assigned to the contribution of the neutral form (Riordan et al., 1967). NT is essentially non fluorescent and absorbs radiation in the wavelength range where Tyr and Trp emit fluorescence, with a Trp-to-NT Förster's distance (*i.e.*, the donor-acceptor distance at which the FRET efficiency is 50%) as large as 26 Å (Steiner et al. 1991). For these reasons, NT has great potential as an energy acceptor in FRET studies and indeed direct chemical nitration of Tyr was used to investigate the structural and folding properties of calmodulin (Steiner et al., 1991) and apomyoglobin (Rischel and Poulsen, 1995; Rischel et al., 1996; Tcherkasskaya and Ptitsyn, 1999). However, very little is known about the possibility of exploiting the unique spectral properties of NT to study molecular recognition (Riordan et al., 1967, Juminaga et al., 1994; Mezo et al., 2001).

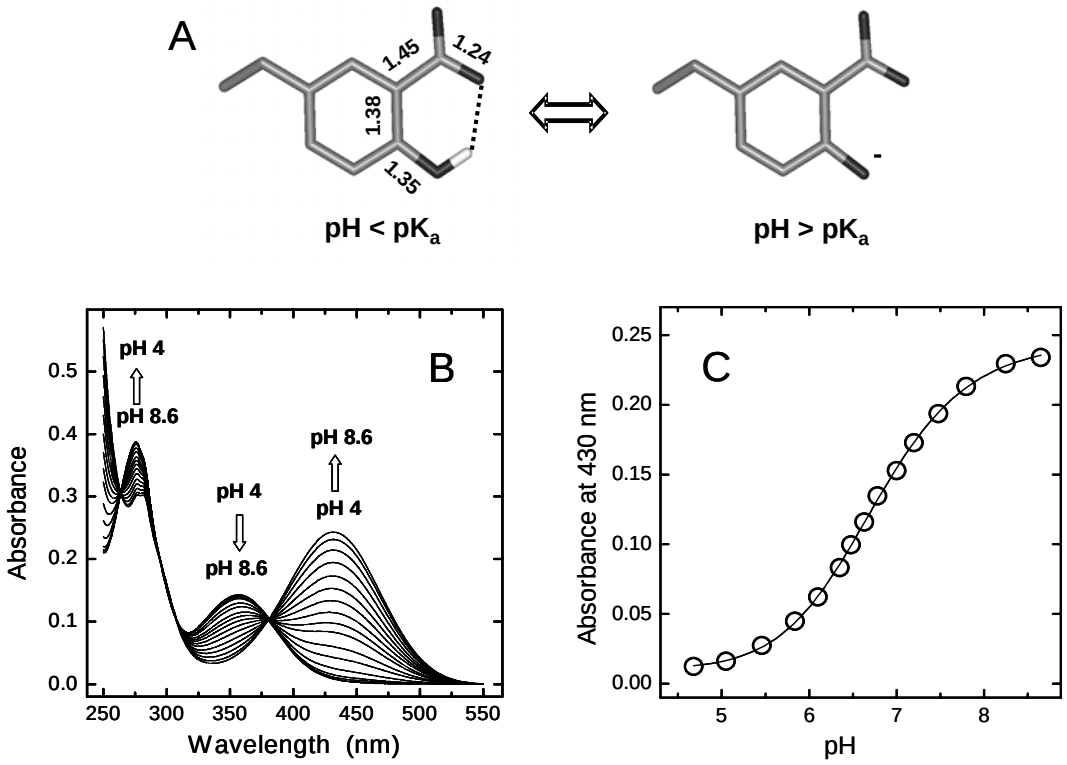


Fig. 1. (**A**) Structure of 3-nitrotyrosine. Bond distances (Å) are taken from the crystallographic structure of free NT (Mostad and Natarajan 1990). (**B**-**C**) Spectrophotometric titration of Y3NT by UV/Vis absorption spectroscopy. Absorption spectra (**B**) and plot of the absorbance values at 430 nm of Y3NT (2 mL, 48 μM) as a function of pH (**C**). Measurements were carried out in 2 mM citrate-borate-phosphate buffer, at the indicated pH. The data points were fitted to equation 1, yielding a pK_a value for NT of 6.74 ± 0.02.

Thrombin is a serine protease that plays a pivotal role in hemostasis (Davie et al. 1991). Under physiological conditions, it exists in equilibrium (~50%) between a procoagulant (fast) form and an anticoagulant (slow) form. The slow ⇔ fast transition of thrombin is triggered by Na^+ binding (for a review, see Di Cera et al., 1997), that stabilizes the enzyme into a more open and rigid structure (*i.e.*, the fast form), compared to the more closed and flexible conformation it assumes in the slow form (De Filippis et al., 2002, 2005, Johnson et al., 2005). Hirudin, a 64-amino acid polypeptide isolated from the salivary secretions of medicinal leeches, is the most potent and specific inhibitor of thrombin (Markwardt, 1994). It is composed of a C-terminal tail, interacting with the fibrinogen-binding site on thrombin, and a compact N-terminal domain, encompassing residues 1-47 and stabilized by three disulfide bonds, that covers the catalytic pocket and extensively penetrates into the enzyme specificity sites (Rydel et al., 1991).

In the present work, we chose the hirudin-thrombin system as a suitable model for evaluating the potentialities of NT as a spectroscopic probe in the study of protein-protein interactions. Due to its important pharmacological implications, the hirudin-thrombin pair has been the object of thorough biochemical and structural studies. Indeed, the high-resolution crystallographic structures of both free (Pineda et al., 2004; Johnson et al., 2005) and hirudin-bound form of thrombin are available (Rydel et al., 1991), together with several NMR structures of full-length and truncated hirudin forms (Haruyama and Wutrich, 1989; Szyperski et al., 1992; Nicastro et al., 1997). Hence, taking advantage of the synthetic procedure previously established (De Filippis et al., 1995, 1998), we synthesized two NT-containing analogues of hirudin fragment 1-47: the Y3NT analogue, in which Tyr3 was replaced by NT, and the S2R/Y3NT analogue, containing the double substitution Ser2 $\rightarrow$ Arg and Tyr3 $\rightarrow$ NT. The binding of these analogues to thrombin allosteric forms was probed by FRET and UV/Vis absorption spectroscopy, that allowed us to obtain reliable values for the dissociation constants (K_d) of the enzyme inhibitor complexes and key information on the details of hirudin-thrombin interaction in solution. On the whole, our results demonstrate that NT is a suitable spectroscopic probe for investigating ligand-protein interactions and suggest that its incorporation into proteins may have vast applications in biotechnology and pharmacological screening.

MATERIALS AND METHODS

Materials

Human α-Thrombin was purchased from Haematologic Technologies Inc. (Essex Junction, VT) or from Calbiochem. The chromogenic substrate (D)-Phe-Pro-Arg-*p*NA (FPR) was synthesized as previously described (De Filippis et al. 2002). N^{α}-Fmoc-3-nitrotyrosine and other protected amino acids, solvents and reagents for peptide synthesis were purchased from Applied Biosystems (Foster City, CA) or Bachem AG (Bubendorf, Switzerland). Nitrophenol-isomers were from Sigma, while buffers and organic solvents were of analytical grade and obtained from Fluka (Basel, Switzerland).

Peptide Synthesis

Hirudin analogues were synthesized by the solid-phase Fmoc-method in two sequential steps, involving automated synthesis of segment 6-47 on a Lys-derivatized (0.90 mmol/g of resin) *p*-alkoxybenzylester polystyrene resin and elongation of the peptide by manual solid phase procedure, as previously described (De Filippis et al. 1995, 1998). After precipitation with diethylether, the crude reduced peptides (2 mg/mL) were allowed to fold for 24 h in 0.1 M $NaHCO_3$ buffer, pH 8.3, under air-oxidation conditions in the presence of 100 μM β-mercaptoethanol (De Filippis et al. 1995), and purified by preparative RP-HPLC (see Figure S1, Supporting Information). The chemical identity of the disulfide-oxidized species was established by automated N-terminal sequence analysis and mass spectrometry on a Mariner ESI-TOF instrument from Perseptive Biosystems (Stafford, TX).

Spectroscopic Measurements

Unless otherwise specified, all measurements were carried out at 25 ± 0.2 °C in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% (w/v) PEG-8000 and 0.2 M NaCl, for the fast form, or choline chloride (ChCl), for the slow form. Temperature correction was applied for Tris buffer. All spectra were subtracted for the corresponding base lines.

Protein concentration was determined by UV absorption at 280 nm (Gill and von Hippel 1989) on a Lambda-2 spectrophotometer from Perkin-Elmer (Norwalk, CT) using a molar absorptivity value of 65770 $M^{-1}\cdot cm^{-1}$, for thrombin, and 2920 $M^{-1}\cdot cm^{-1}$, for the wild-type fragment 1-47 (WT) and S2R analogue. The concentration of NT-containing analogues was determined at 381 nm, using a molar extinction coefficient of 2200 $M^{-1}\cdot cm^{-1}$ (Tcherkasskaya and Ptitsyn 1999). The pK_a value of NT was determined by spectrophotometric titration of Y3NT at 430 nm. A solution of the hirudin analogue (2 mL, 48 μM) in 2 mM citrate-borate-phosphate buffer, pH 8.6, containing PEG-8000 0.1%, was titrated with 0.5 M aqueous HCl (2-6 μL aliquots). The solution pH was measured at 25 ± 0.2 °C with a Metrohm (Herisau, CH) model 632 pH-meter. Absorbance data were corrected for sample dilution (< 2% of the final volume) and fitted to equation 1:

$$A_{430\,nm} = \{(a - b)/[1 + 10^{(pKa - pH)}]\} + b \qquad (1)$$

where *a* and *b*, are the higher and lower absorbance limits of NT at 430 nm.

CD spectra were recorded on a Jasco (Tokyo, Japan) model J-810 spectropolarimeter. Far-UV spectra were recorded in a 1-mm cell, at a scan-speed of 10 nm/min, with a response time of 16 sec, and resulted from the average of four accumulations. Near-UV spectra were

recorded in a 1-cm cell, at a scan-speed of 50 nm/min, with a response time of 2 sec, and resulted from the average of 16 accumulations. Ellipticity data were expressed as mean-residue ellipticity. Fluorescence spectra were recorded on a Perkin-Elmer spectrofluorimeter model LS-50B. Spectra were taken using a 1-cm pathlength cuvette, at a scan speed of 120 nm/min by exciting the protein samples at 280 nm, with an excitation/emission slit of 5 nm.

Thrombin-Binding Measurements

a. Fluorescence. The interaction of NT-containing hirudin analogues with thrombin was monitored by adding, under gentle magnetic stirring, to a solution of thrombin (2 mL, 50 nM), in 5 mM Tris-HCl buffer, pH 8.0, 0.1% PEG, in the presence of 0.2 M NaCl or ChCl, aliquots (2-10 μL) of inhibitor stock solutions (2-250 μM) in the same buffer. At each inhibitor concentration, protein samples were equilibrated for 5 min at 25°C and excited at 295 nm, using an excitation/emission slit of 5 and 10 nm, respectively, and a scan speed of 240 nm/min. The decrease in fluorescence intensity at the λ_{max} (342 nm) of thrombin was recorded as a function of inhibitor concentration. Fluorescence data were corrected for sample dilution (< 5% of the final volume) and expressed as $(F°-F)/\Delta F_{max}$, where $F°$ and F is the fluorescence of thrombin in the absence and presence of the inhibitor, respectively, and ΔF_{max} is the maximum fluorescence change at saturating concentrations of inhibitor. For a simple one-site binding mechanism $R + L \Leftrightarrow RL$, the fluorescence intensity (F) of the receptor (R) at a given concentration of ligand (L) is linearly related to the concentration of the complex $[RL]$, $F = [RL]\cdot F_{bound} + [R]_{free}\cdot F_{free}$. Since $[R]_{free} = [R] - [RL]$, then $(F°-F)/\Delta F_{max} = [RL]/[R]$ (Eftink 1997). The data were fitted to equation 2 (Copeland 2000), using the program Origin 6.0 (MicroCal Inc.):

$$[RL]/[R] = \{([R] + [L] + K_d) - \{([R] + [L] + K_d)^2 - 4\cdot[R][L]\}^{1/2}\}/2\cdot[R] \qquad (2)$$

where K_d is the dissociation constant of complex and $[R]$ is the total concentration of the receptor.

b. UV/Vis absorption. Binding of S2R/Y3NT to thrombin allosteric forms was monitored by adding 8 μL of a thrombin stock solution (261 μM) to a solution of hirudin analogue (0.5 mL, 8.6 μM), in a 1-cm pathlength quartz-cuvette (0.2 x 1.0 cm). After 15-min equilibration,

spectra were taken at a scan speed of 60 nm/min. In the case of Y3NT, 15 μL of thrombin stock solution (197 μM) were added to the solution of hirudin analogue (0.5 mL, 8.6 μM).

Thrombin Inhibition Assays

The K_d values of NT-containing analogues were also estimated by classical competitive inhibition experiments of thrombin-mediated substrate hydrolysis, according to the tight-binding model (Copeland 2000), as previously detailed (De Filippis et al. 2002). The inhibitor was incubated at 25 ± 0.2 °C for one hour with 100 pM thrombin in 5 mM Tris, pH 8.0, containing 0.1% (w/v) PEG-8000 and 0.2 M NaCl or ChCl. The reaction was started by addition of FPR (20 μM) and the release of *p*-nitroaniline (*p*NA) was determined by recording the absorbance increase at 405 nm. The ionic strength was kept constant at 200 mM with NaCl for the fast form or with ChCl when the slow form was being studied. Thrombin inhibition data were fitted to the equation:

$$V_i/V_0 = 1-\{([E]+[I]+K_I^{app})-\{([E]+[I]+K_I^{app})^2-4\cdot[E][I]\}^{1/2}\}/2\cdot[E] \quad (3)$$

to obtain the apparent inhibition constant, K_I^{app}. *[E]* and *[I]* are the total enzyme and inhibitor concentrations and V_i and V_0 are the steady-state velocities of substrate hydrolyisis by thrombin, in the presence (V_i) or absence (V_0) of the inhibitor. K_I^{app} values were corrected for substrate concentration and for the K_m value of FPR for thrombin (K_m = 0.48 ± 0.03 μM) (Di Cera et al. 1997; De Filippis et al. 2002), according to equation 4:

$$K_I = K_I^{app}/[1+([S]/K_m)] \quad (4)$$

where K_I is equal to the dissociation constant, K_d, of the enzyme-inhibitor complex (Copeland 2000).

Computational Methods

The structure of the synthetic analogues in the thrombin-bound state was modeled on the crystallographic structure of hirudin-thrombin complex (4HTC.pdb) (Rydel et al. 1991), by keeping the position of all atoms unchanged and building the nitro-group on Tyr3, to obtain NT. The geometry of NT in the enzyme-inhibitor complex was optimized using the bond

length and angle parameters derived from the crystal structure of free NT (Mostad and Natarajan 1990).

RESULTS AND DISCUSSION

Synthesis and Characterization of Hirudin Analogues. The analogues of hirudin fragment 1-47, Y3NT and S2R/Y3NT, were obtained by combining automated and manual solid-phase peptide synthesis, and allowed to fold under air oxidizing conditions, as previously detailed (De Filippis et al. 1995, 1998). The disulfide oxidized species were purified by RP-HPLC (Figure 2) and their chemical identity established by enzymatic fingerprint analysis (not shown), as described elsewhere (De Filippis et al. 1995), and ESI-TOF mass spectrometry, which gave mass values in agreement with the expected amino acid composition within 50 ppm accuracy.

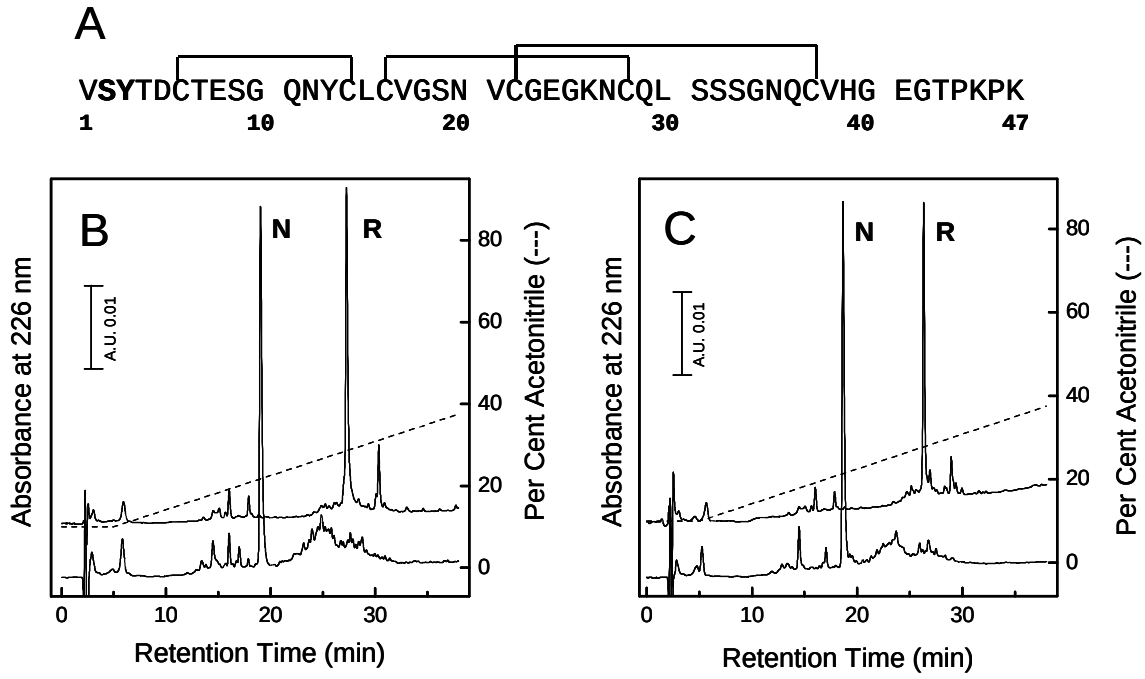


Fig. 2 (**A**) Amino acid sequence of the N-terminal domain 1-47 of hirudin HM2 from *Hirudinaria manillensis*. Mutated residues are indicated in bold and disulfide bonds, connecting the six Cys-resisues, by plain lines. RP-HPLC analysis of the crude synthetic analogues Y3NT (**B**) and S2R/Y3NT (**C**), with cysteines in the reduced (R) and disulfide oxidized (N) state. Aliquots (100 μg) of the synthetic peptides were loaded onto a Vydac (The Separations Group, Hesperia, CA) C18 column (4.6 x 150 mm, 5 μm particle size), eluted with a linear acetonitrile-0.1% TFA gradient, at a flow rate of 0.8 ml/min. The material eluted in correspondence of the major chromatographic peaks were collected and characterized by high-resolution mass spectrometry. Alternatively, the disulfide-folded species were purified for subsequent conformational and functional characterization.

The NT side-chain can form an internal hydrogen bond (Figure 1A) and the absorption properties are strongly dependent on its ionization state (Riordan et al. 1967) (Figure 1B). At low pH (< 6.5), where the neutral (–OH) form of NT is predominant, the spectrum of Y3NT is characterized by two bands centered at 276 and 357 nm and assigned to the π-π* transitions of the aromatic ring, denoted as $^1A_{1g} \rightarrow {}^1B_{1u}$ and $^1A_{1g} \rightarrow {}^1B_{2u}$, respectively (see Meloun et al. 1968). At higher pH values, where the ionized ($–O^-$) form prevails, the intensity of the 276-nm band decreases, while a new band appears at 430 nm. The position of the band at longer wavelengths (357 or 430 nm) reflects the protonation state of NT in solution and the change in the intensity of the 430-nm band allowed us to estimate the pK_a value of NT in the Y3NT analogue, calculated as 6.74 ± 0.02 (Figure 1C). This value is very similar to that of free NT-derivatives (Riordan et al. 1967) and is in keeping with the high surface exposure of Tyr3 in the structure of the natural hirudin fragment 1-47 (WT) (Nicastro et al. 1997).

The spectral properties of Y3NT were investigated by circular dichroism (CD) and fluorescence spectroscopy and compared with those of the natural fragment 1-47 (Figure 3). The close similarity of the CD spectra of the natural and modified species (Figure 3A) provides strong evidence that the secondary structure content of the protein is retained upon Tyr3 → NT exchange. The near-UV CD spectrum of WT is dominated by the contribution of the three disulfide bonds, while the spectrum of Y3NT is characterized by a broad positive band centered at 426 nm, assigned to the contribution of the ionized form of NT, indicating that NT3 is located in a rigid and asymmetric environment (Strickland 1974) (Figure 3B). The fluorescence spectrum of WT (Figure 3C), taken after excitation at 280 nm, displays a λ_{max} value at 305 nm, due to the emission of Tyr3 and Tyr13 in the hirudin sequence (Figure 2) (Lakowicz 1999), whereas the fluorescence intensity of Y3NT is nearly zero. We propose that the emission of Tyr13 (*i.e.*, the donor) is transferred, by a non-radiative energy transfer process, to NT3 (*i.e.*, the acceptor), which is essentially non-fluorescent. As a result, the fluorescence of Y3NT is negligible.

FRET is a non-radiative decay process occurring between a donor and an acceptor, which interact *via* electromagnetic dipoles transferring the excitation energy of the donor to the acceptor (Lakowicz, 1999). For a one donor-one acceptor system, the efficiency of energy transfer depends on the extent of spectral overlap of the emission spectrum of the donor with the absorption spectrum of the acceptor, on the donor quantum yield, on the inverse six power of the distance separating the donor and acceptor, and on their orientation (Selvin 1995; Lakowicz, 1999). In the case of hirudin, there is significant overlap of the emission of Tyr

with the absorption of NT (Figure 3C). Furthermore, the aromatic rings of Tyr13 and NT3 are almost parallel and separated by fairly short distance (~ 7 Å) in the inhibitor structure (Nicastro et al. 1997). These observations provide strong support to our proposal that the fluorescence of Tyr13 is quenched by NT3 *via* resonance energy transfer.

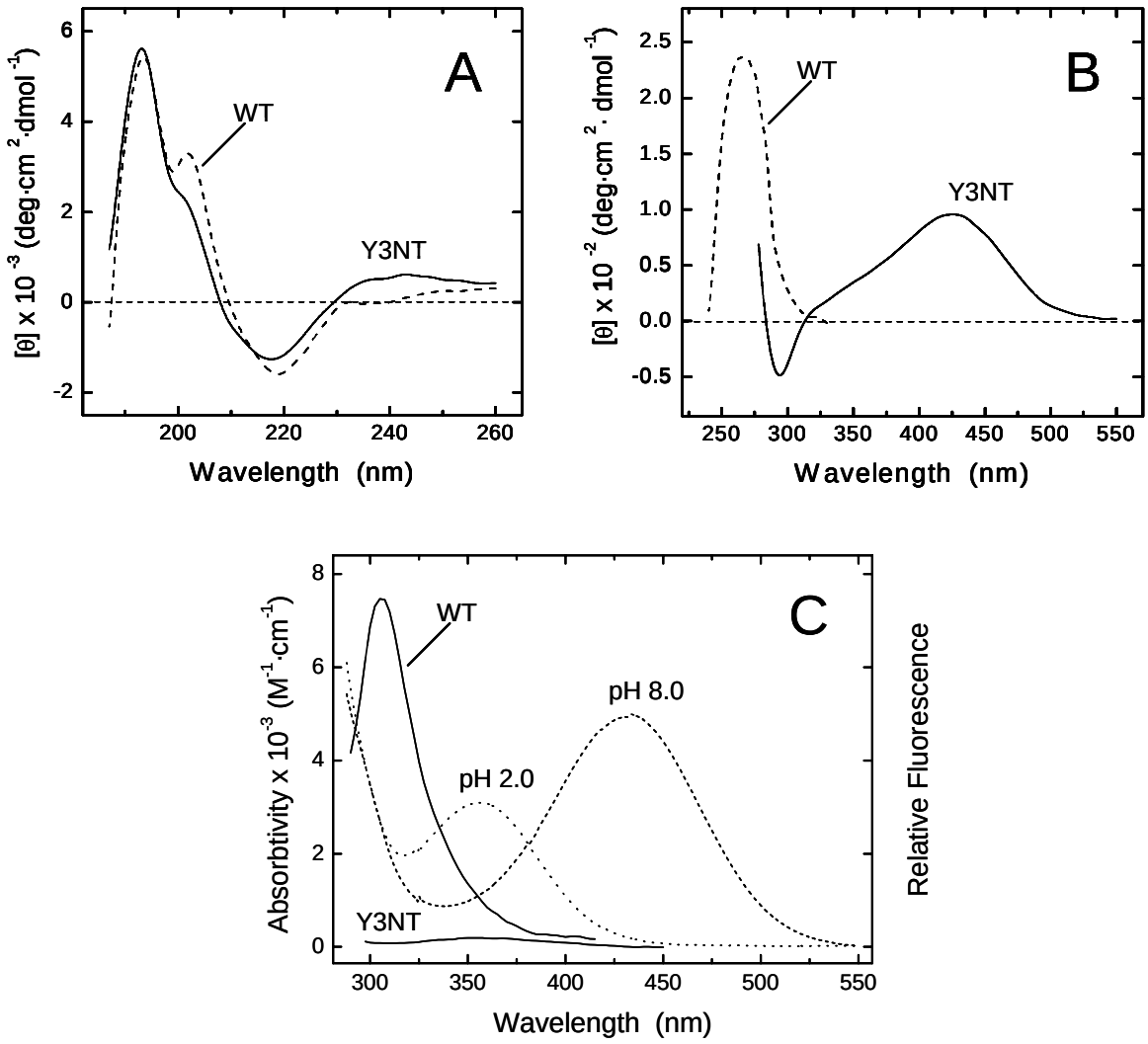


Fig. 3. Conformational characterization of wild-type hirudin fragment 1-47 (WT) and Y3NT analogue. Far- (**A**) and near-UV (**B**) CD spectra of hirudin analogues were taken at a protein concentration of 40 and 200 μM in the far- and near-UV region, respectively. (**C**) Superimposition of the fluorescence spectra (continuous lines) of WT and Y3NT with the absorption spectra (dashed/dotted lines) of Y3NT at pH 2.0 and 8.0. All measurements were carried out at 25° in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% (w/v) PEG 8000 and 0.2 M NaCl.

Probing Hirudin-Thrombin Interaction by Trp-to-NT Fluorescence Energy Transfer

The fluorescence of thrombin is characterized by a λ_{max} value at 342 nm (Wells and Di Cera 1992; De Filippis et al. 2005), in agreement with the relatively high surface exposure of tryptophans in the enzyme structure (Bode et al. 1992). Earlier Trp → Phe mutagenesis studies have shown that Trp207 is the major contributor to the fluorescence intensity of thrombin (35%), followed by Trp96 (11%), while Trp60d and Trp215 each contribute by only ~6% (Bell et al., 2000). In the Na^+-free (slow) form, the fluorescence intensity is about 18% lower than that of the enzyme in the Na^+-bound (fast) form (Wells and Di Cera 1992; De Filippis et al. 2005). This effect was assigned to the higher conformational flexibility of the slow form, compared to that of the more rigid and stable fast form, and to specific changes in the environment of Trp215 in the S3 site of the enzyme (De Filippis et al. 2005; Johnson et al. 2005). In the presence of saturating concentrations of Y3NT or S2R/Y3NT, the fluorescence of thrombin is strongly quenched (Figure 3A, B) and approaches a similar value, under either fast (0.2 M NaCl) and slow (0.2 M ChCl) conditions (Figure 3B).

As in the case of hirudin, quenching of fluorescence is mainly caused by resonance energy transfer, occurring between (some of) the Trp-residues of thrombin (i.e., the donors) and the single 3-nitrotyrosine of the inhibitors (i.e., the acceptor). With respect to this, there is an extensive overlap of the emission spectrum of the enzyme (i.e., the donor) with the absorption spectrum of the inhibitor (i.e., the acceptor) (see Figure 4C). In addition, Trp-to-NT energy transfer is also favored by the relatively short distances separating Trp-residues and NT in the enzyme-inhibitor complex (Rydel et al. 1991) (see Figure 5). To estimate the possible contribution of spectroscopic effects other than FRET (e.g, unspecific binding, dynamic or static quenching), the fluorescence of thrombin was measured in the presence of increasing concentrations of free NT. The data reported in Figure 4D indicate that NT slightly (~14%) reduces the fluorescence of the enzyme, in keeping with the notion that nitro-compounds (e.g., nitromethane and nitrobenzene) quench the emission of polycyclic aromatic hydrocarbons by a mixed static/dynamic mechanism (Sawicki et al., 1964; Dreeskamp et al., 1975; Lakowicz, 1999). However, in the case of hirudin-thrombin interaction this effect is expected to be negligible. Indeed, for either static or dynamic quenching to occur, the quencher must contact the fluorophore within van der Waals distance (Lakowicz, 1999). With respect to this, the modeled structure of Y3NT bound to thrombin (see below) reveals that NT does not directly contact any of the Trp-residues of thrombin and that even the closest Trp-residues (i.e., Trp215 and Trp60d) are located at 7-8 Å from NT in the hirudin-thrombin complex (see Figure 5). All these considerations allow us to conclude that quenching of

thrombin fluorescence by Y3NT (or S2R/Y3NT) is mainly caused by Trp-to-NT energy transfer.

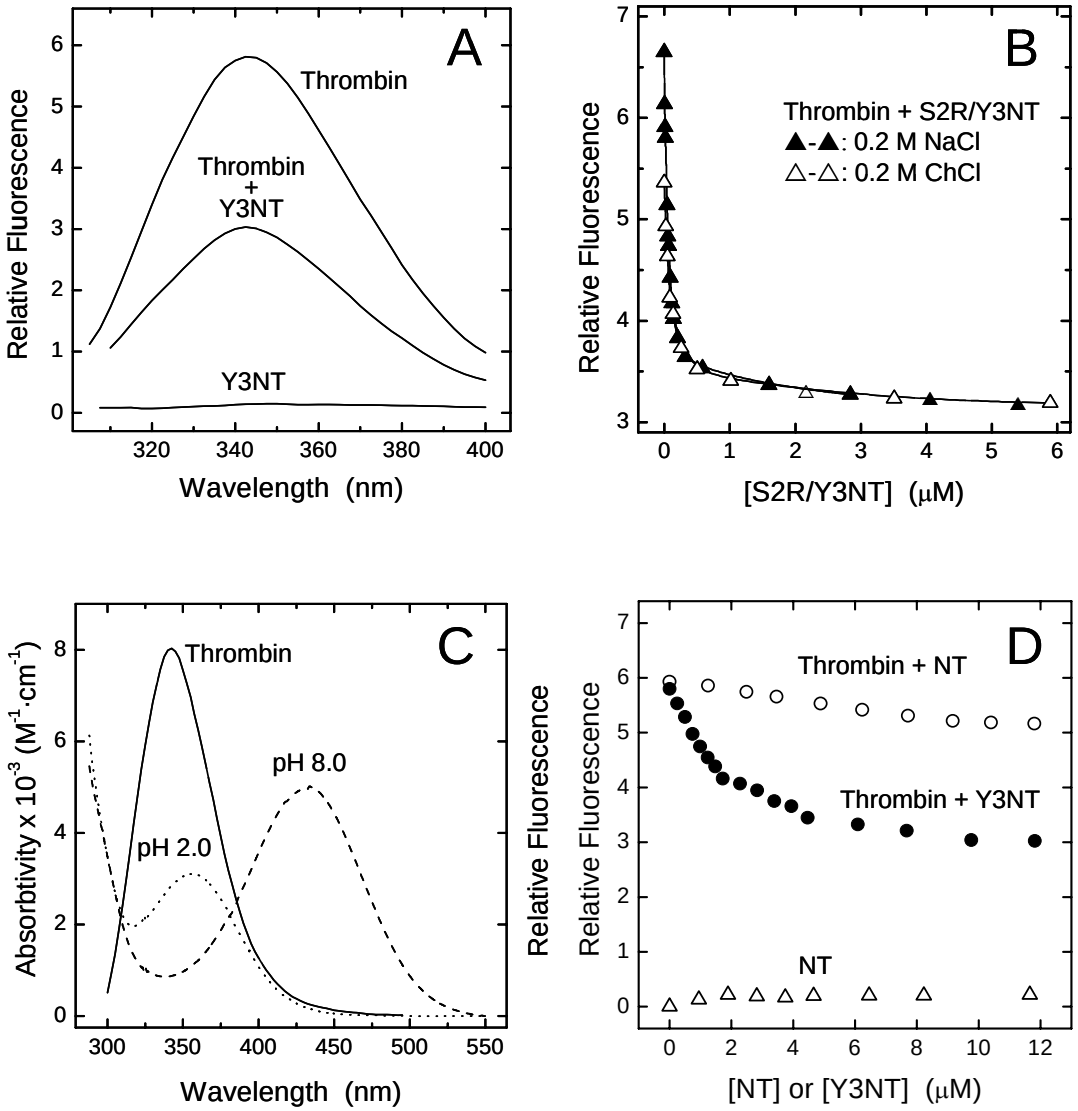


Fig. 4. Binding of Y3NT and S2R/Y3NT to thrombin, monitored by Trp-to-NT fluorescence energy transfer. (**A**) Fluorescence spectra of thrombin alone (50 nM) and in the presence of Y3NT (10 μM). For comparison, the spectrum of the free inhibitor Y3NT (10 μM) is also reported. (**B**) Change in the fluorescence of thrombin as a function of S2R/Y3NT concentration, under fast (▲-▲, 0.2 M NaCl) and slow (△-△, 0.2 M ChCl) conditions. (**C**) Superimposition of the fluorescence spectrum of thrombin (continuous line) with the absorption spectra of Y3NT at pH 2 and 8.0 (dashed/dotted lines). (**D**) Change in thrombin fluorescence as a function of Y3NT concentration (●-●). As a control, the fluorescence intensity of thrombin in the presence of free NT (○-○) is reported. The signal of Y3NT alone (△-△) is also included. All measurements were carried out at 25°C by exciting the protein samples at 295 nm in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% (w/v) PEG 8000 and 0.2 M salt, as indicated, and recording the fluorescence signal at 342 nm.

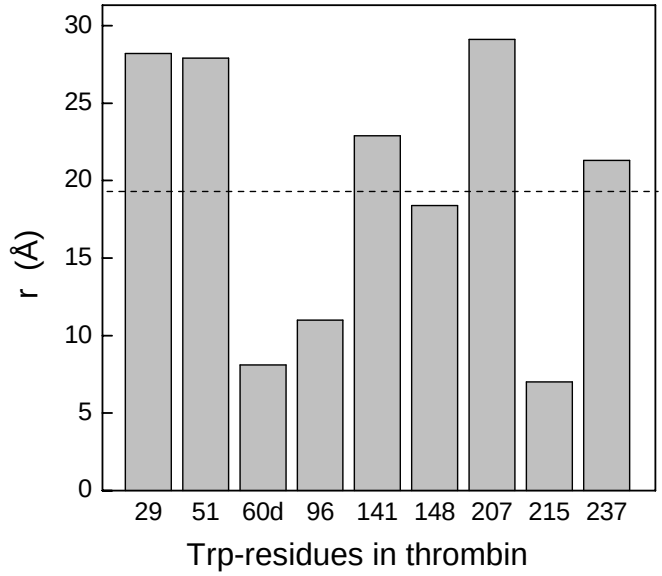


Fig. 5. Centroid distance, *r*, between the side-chain of Trp-residues in thrombin and the single NT-residue in the synthetic analogue Y3NT. Distance values are based on the modelled structure of Y3NT bound to thrombin (4HTC.pdb) (Rydel et al. 1991). The dashed line represents the average Trp-NT distance in the hirudin thrombin complex.

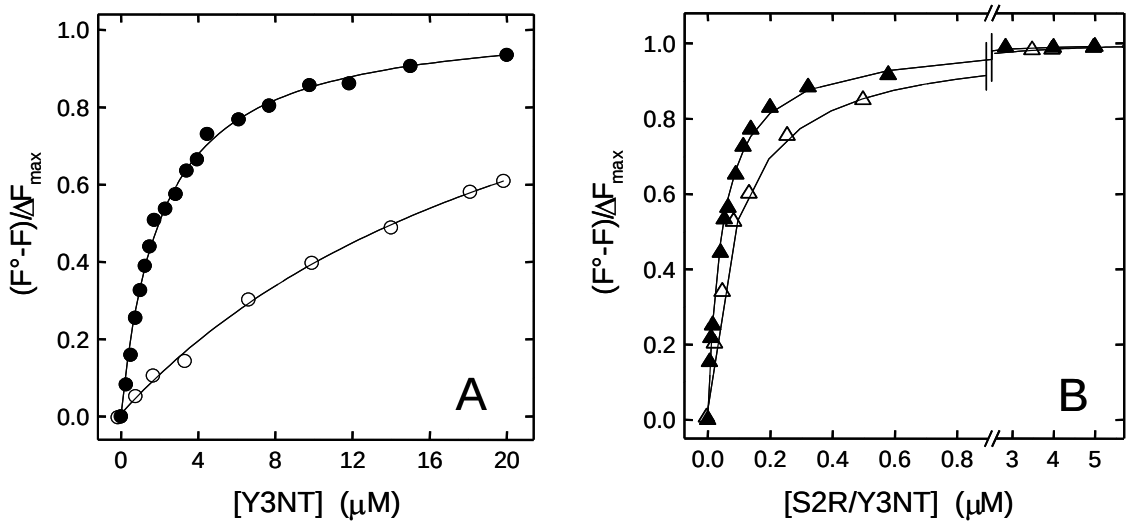


Fig. 6. Determination of the dissociation constant (K_d) of the complexes formed by the synthetic analogues Y3NT (**A**) (●, ○) and S2R/Y3NT (**B**) (▲, △) with thrombin, under fast (filled symbols) and slow (empty symbols) conditions. Fluorescence data were obtained as detailed in the Methods (see also Figure 3) and expressed as $(F°-F)/\Delta F_{max}$. Continuous lines represents the best fit of the data points to equation 2, that allowed us to obtain the K_d values reported in Table 1.

The quenching data reported above were used to obtain quantitative estimates (i.e., K_d values) of the binding strength of Y3NT and S2R/Y3NT to thrombin allosteric forms (see Figure 5 and Table 1). For both analogues, the excellent fit of the experimental data to the

curve describing one-site binding mechanism is a stringent, albeit indirect, proof of 1:1 binding stoichiometry (Copeland 2000). The replacement of Tyr3 with NT resulted in a drop in the affinity of Y3NT for thrombin, that was restored in the doubly-substituted analogue S2R/Y3NT by replacing Ser2 with Arg. The structural model of Y3NT bound to thrombin (see Figure 6), based on the crystallographic structure of the hirudin-thrombin complex (4HTC.pdb) (Rydel et al. 1991) (see Methods), reveals that the $-NO_2$ group of NT might be easily accommodated into the S2 specificity site of the enzyme without requiring steric distortion. Likely, the lower affinity of Y3NT reflects the lower hydrophobicity of NT at pH 8.0, where it exists by ~95% in the ionized form.

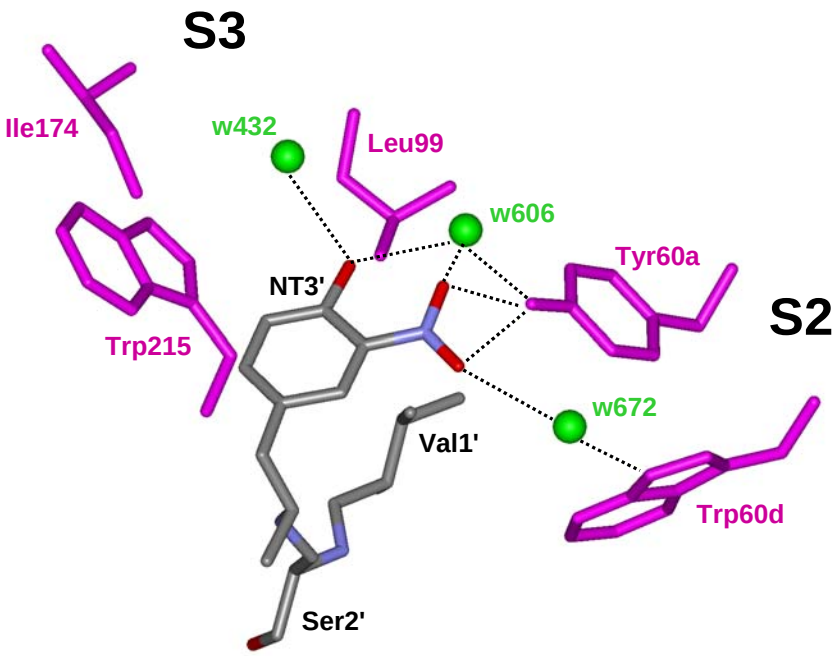


Fig. 6. Schematic representation of the interaction of the N-terminal tripeptide of Y3NT with thrombin. The inhibitor is color-coded (carbon, gray; nitrogen, blue; oxygen, red), while the relevant residues of thrombin in the S2 (Tyr60a and Trp60d) and S3 site (Trp215, Leu99 and Ile174) are shown in magenta. Tyr3' of hirudin fills the apolar S3 site of thrombin, Val1' contacts the S2 site, and Ser2' covers, but does not penetrate, the S1 site. The structure of Y3NT was modeled on the structure of hirudin-thrombin complex, crystallized under conditions stabilizing the fast form (*i.e.*, sodium acetate) (4HTC.pdb) (Rydel et al. 1991). The structural water molecules (w432, B-factor 36 $Å^2$, occupancy 1.0; w606, B-factor 22 $Å^2$, occupancy 0.52; w672, B-factor 49 $Å^2$, occupancy 0.65) are indicated by green spheres and approximately lie in the geometrical plane of NT. The nitro-group was easily accommodated at the enzyme-inhibitor interface without bump, by keeping the orientation of Tyr3 (χ_1 = -62° and χ_2 = -56°) unchanged, together with the rest of the hirudin-thrombin structure. The bond-lengths and angles of NT are derived from the crystal structure of the free amino acid (Mostad and Natarajan 1990). Relevant NT-thrombin distances, in the 2.5-3.5 Å range, are indicated by dashed lines. Of note, w606, that in the wild-type hirudin-thrombin structure connects Tyr3' of the inhibitor to Tyr60a of the enzyme, is well suited as a hydrogen bond donor to stabilize the NT ring system (see text).

With respect to this, the logP value (i.e., the logarithm of octanol/water partition coefficient, P) of 2-nitrophenol, taken as a suitable model of the NT side-chain, is –1.47 at pH 8.0 (Csizmadia et al. 1997), while that of phenol, taken as a model of Tyr, is +1.50 (Abraham et al. 2000). However, hydrophobicity alone is not sufficient to explain the remarkable decrease in the affinity of Y3NT for thrombin. Indeed, when measurements were carried out at pH 6.0, where NT predominantly (~82%) exists in the more hydrophobic neutral form (logP +1.79) (Abraham et al. 2000), the affinity of Y3NT for thrombin was only 1.4-fold higher than that determined at pH 8.0 (see Figure S3, Supporting Information). Besides hydrophobicity, the presence of the nitro-group introduces a net (i.e., at pH 8.0) or partial (i.e., at pH 6.0) negative charge at position 3 of hirudin, that can oppose binding through unfavorable electrostatic interaction with the strong negative potential of the thrombin active site (Karshikov et al. 1992), in agreement with our previous structure-activity relationship studies (De Filippis et al. 2002).

As expected from our earlier work (De Filippis et al. 1998, 2005), Ser2→Arg exchange strongly enhances the affinity of hirudin fragment for thrombin, due to a favorable electrostatic coupling of Arg2' with Asp189 in the S1 site of the enzyme. Of interest, the effects of the amino acid substitutions on the binding to thrombin are strictly additive (Wells 1990) (see legend to Table 1), as already observed with other cumulative amino acid replacements in hirudin (De Filippis et al. 1999). All these indications suggest that the binding data obtained by FRET measurements are realistic. However, the reliability of these data was verified by comparing the K_d values of Y3NT and S2R/Y3NT, obtained by FRET measurements, with those determined by classical enzyme inhibition experiments, in which the rate of thrombin-mediated substrate hydrolysis was measured as a function of inhibitor concentration (Figure S4, Supporting Information). Strikingly, the K_d values for the binding of Y3NT and S2R/Y3NT to thrombin fast form were determined as 1.4 ± 0.1 μM and 41 ± 2 nM, respectively, in agreement (8-10%) with those obtained by FRET (Table 1).

Table 1. Thrombin Binding Data of the Synthetic Hirudin Analogues, as Obtained by Fluorescence Energy Transfer [a] and Enzyme Inhibition Assays

	Fast form		**Slow form**		
1-47 analogues	**K_d** (nM)	**$\Delta\Delta G_b$** (kcal/mol)	**K_d** (nM)	**$\Delta\Delta G_b$** (kcal/mol)	[d] **ΔG_c** (kcal/mol)
WT [b] (Tyr3)	42 ± 0.5	-	1460 ± 20	-	-2.10

Y3NT [a]	1325 ± 80	2.06	17000 ± 500	1.45	-1.51
S2R [b]	1.7 ± 0.02	-1.90	12 ± 2	-2.84	-1.16
S2R/Y3NT [a]	45 ± 2.2	0.06	91 ± 4.2	-1.64	-0.41

[a] The K_d values of Y3NT and S2R/Y3NT were obtained by fitting FRET data to equation 2. [b] The inhibitory potency of the wild-type (WT) and S2R analogue was determined at 25 °C by measuring at 405 nm the release of *p*NA from the synthetic substrate FPR. Further details are reported under Methods [c] $\Delta\Delta G_b$ is the difference in the free energy change of binding to thrombin between the synthetic analogue (ΔG_b^*) and the natural fragment (ΔG_b^{wt}): $\Delta\Delta G_b = \Delta G_b^* - \Delta G_b^{wt}$. A negative value of $\Delta\Delta G_b$ indicates that the mutated species binds to thrombin more tightly than the natural fragment. Errors are ± 0.1 kcal/mol or less. [d] ΔG_c is the free energy of coupling to thrombin, measured as $\Delta G_c = \Delta G_{b,\,fast} - \Delta G_{b,\,slow}$ (Di Cera et al. 1997). The value of ΔG_c is negative if the inhibitor binds to the fast form with higher affinity than to the slow form. Additivity of mutational effects was calculated by the equation $\Delta G_I = \Delta\Delta G_b(S2R/Y3NT) - [\Delta\Delta G_b(Y3NT) + \Delta\Delta G_b(S2R)]$, where ΔG_I is the free energy term that accounts for the energetic interaction between the mutated sites (Wells 1990). ΔG_I values as low as –0.10 and –0.25 kcal/mol were calculated for the binding of S2R/Y3NT to the fast and slow form, respectively. A value of ΔG_I comparable to R·T (~ 0.6 kcal/mol, at 25°C) is usually taken as a proof of additivity (Wells 1990).

Probing Hirudin-Thrombin Interaction by UV/Vis-Absorption Spectroscopy

The absorption spectra of S2R/Y3NT, recorded in the absence and presence of thrombin, are reported in Figure 6. Under fast conditions (Figure 7A), the 430-nm band of S2R/Y3NT, assigned to the contribution of the ionized form of NT at pH 8.0, is reduced by 54%, while an additional band of similar intensity appears at about 362 nm, characteristic of NT in the neutral form. Of note, the binding of Y3NT to thrombin fast form yields very similar results (data not shown). These observations can be explained on the basis of the modeled structure of Y3NT bound to thrombin and assuming that NT interacts with the enzyme in the neutral form.

As shown in Figure 5, three structural water molecules at the enzyme-inhibitor interface (*i.e.*, w432, w606, and w672), characterized by low thermal factors and high occupancy values, can variably interact with NT. In particular, w606, that in the structure of the wild-type hirudin-thrombin complex connects Tyr3' of the inhibitor to Tyr60a of the enzyme (Rydel et al. 1991), is suitably positioned as a hydrogen bond donor to stabilize the six-membered ring system of NT (see Figure 1A and Figure 6). As a result, the contribution of the protonated NT in the bound form appears as a distinct band at about 362 nm in the absorption spectrum of the thrombin-S2R/Y3NT complex (Figure 7A). The residual intensity

of the 430-nm band is contributed by the ionized form of NT in the free inhibitor that exists in equilibrium with the thrombin-bound form. Hence, we conclude that the phenate moiety of NT in the free state becomes protonated to phenol upon binding to thrombin and that a water molecule at the hirudin-thrombin interface, likely w606, functions as a hydrogen donor. Notably, w432 and w606 are conserved in the structure of thrombin bound to hirugen (*i.e.*, the 53-64 peptide of hirudin) (1HAH.pdb; Vijayalakshmi et al. 1994), where the specificity sites of the enzyme are unoccupied, thus suggesting that these water molecules represent constant spots in the solvation shell of thrombin and, perhaps, key elements for molecular recognition, in keeping with the key role that protein-water interactions play in ligand binding (for a review, see Mattos 2002).

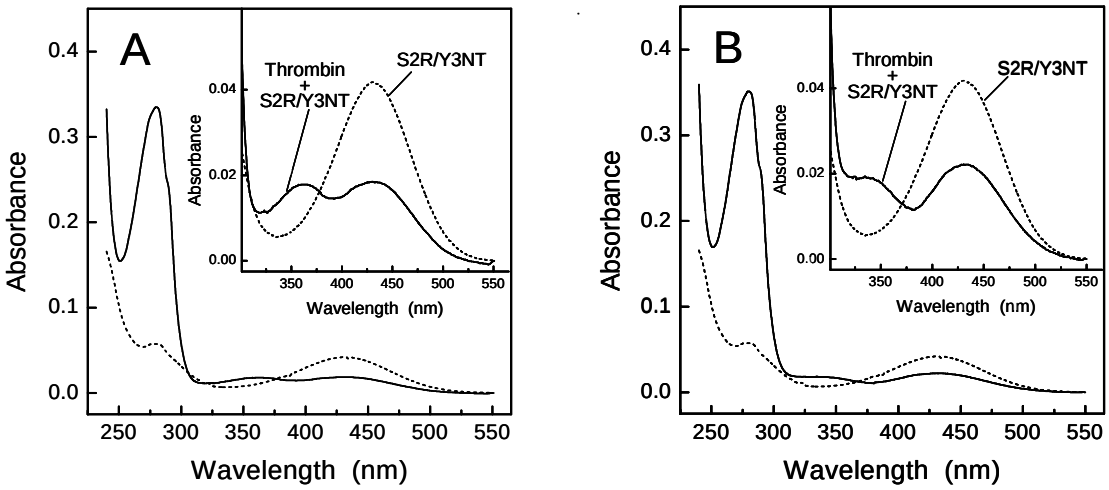


Fig. 7. Binding of the synthetic analogue S2R/Y3NT to the fast (**A**) or slow (**B**) form of thrombin, monitored by UV/Vis absorption spectroscopy. Spectra of the inhibitor (8.6 μM) were taken at 25° in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% PEG 8000 and 0.2 M NaCl, for the fast form, or 0.2 M ChCl, for the slow form, in the absence (---) and presence (—) of thrombin (4.1 μM). For clarity, the *Insets* show the spectra in the wavelength range 300-550 nm.

Similarly to what is observed for the fast form, the interaction of S2R/Y3NT with thrombin slow form reduces the intensity of the 430-nm band by 45% (Figure 7B), in agreement with the lower affinity of this analogue for the Na^+-free enzyme (see Table 1). Furthermore, a new band appears at 335 nm (Figure 7B), blue-shifted by about 27 nm compared to that observed upon binding of the inhibitor to the fast form (Figure 7A). These results suggest that NT3 interacts with the Na^+-free enzyme still in the neutral form, but that the details of this interaction (*e.g.*, side-chain orientation, hydrogen bonding, or solvation) might be slightly different. In particular, the observed spectral changes may reflect the perturbation of the internal hydrogen bond in the NT ring system, possibly caused by some

rearrangements of the water molecules at the interface and/or distortion of $–NO_2$ group out of the plane of the aromatic ring. This view is consistent with the absorption spectra of nitrophenols recorded at pH 2.0 (data not shown), showing that the λ_{max} value decreases from 350 nm of *ortho*-nitrophenol, where the internal hydrogen bond is formed, to 333 nm and 318 nm of *meta*- and *para*-nitrophenol, respectively, where the formation of the intramolecular hydrogen bond is impaired.

Structural studies, conducted either in solution (De Filippis et al. 2005) or in the crystal state (Johnson et al. 2005), highlight some differences in the conformational/flexibility properties of thrombin allosteric forms. In particular, the binding of Na^+ specifically stabilizes the enzyme in a more open and rigid conformation, compared to the more closed and flexible structure it assumes in the Na^+-free state (De Filippis et al., 2005). It is possible that these conformational changes also alter the structure of the solvation shell of the free enzyme prior to hirudin binding and/or the hydrogen-bond network at the interface in the enzyme-inhibitor complex. Strikingly, these changes are precisely pin pointed by the λ_{max} shift in the absorption of NT, from 362 to 335 nm, observed upon binding of the inhibitor to the fast or slow form of thrombin.

CONCLUSIONS

The results of our work on the hirudin-thrombin system demonstrate that NT can be used as a Tyr-analogue to detect protein-protein interactions with sensitivity in the low nanomolar range, to uncover subtle details at the ligand-protein interface, and to obtain reliable K_d values for structure-activity relationships studies. In addition, several considerations support the general applicability of NT in ligand binding studies: first, even though unnatural amino acid mutagenesis is still limited by rather complicated experimental procedures and by the exceedingly low amounts of the resulting mutated proteins (England 2004), many noncoded amino acids, including NT, have been successfully incorporated at specific sites into several different protein systems (Cropp and Schultz, 2004) (see also Introduction); secondly, a survey of the distribution frequency of amino acids in proteins reveals that the binding regions in proteins are much richer in aromatic residues *(i.e.*, especially in Trp-residues) than the average protein surface (Lo Conte et al. 1999), thus suggesting that protein interfaces could be particularly sensitive to Trp-to-NT energy transfer; thirdly, due to the large Förster's distance (26 Å) of the Trp-NT pair (Steiner et al. 1991), fluorescence energy transfer between Trp and NT would operate even when the donor(s) and acceptor are quite distant in the

macromolecular complex; fourthly, the small increase in the size of NT, compared to other bulkier fluorescence acceptors (for reviews, see Wu and Brand, 1994; Selvin, 1995), is expected to introduce only slight structural perturbation at the mutation site, thus providing a more realistic picture of the binding process under study.

In conclusion, our results demonstrate that NT is a suitable spectroscopic probe for investigating ligand-protein interactions and suggest that its incorporation into proteins may have vast applications in biotechnology and pharmacological screening.

This chapter has been adapted from:

De Filippis, V., Frasson, R., and Fontana, A. (2006) 3-Nitrotyrosine as a spectroscopic probe for investigating protein-protein interactions. *Protein Sci.* **15**, 976-86.

CHAPTER 3.2

o- Nitrotyrosine and *p*-Iodophenylalanine as Spectroscopic Probes for Structural Characterization of SH3 Complexes

INTRODUCTION

The central role played by molecular interactions in the functioning of all biological systems requires a detailed structural understanding of how recognition takes place (Jones and Thornton 1996; Pawson and Nash 2003). Complete structural determination of protein complexes can however be time-consuming, since with the current technical tools it is often necessary to undertake *de novo* structure determinations even when the structures of all the interacting partners are known individually. It is therefore increasingly important to develop new high throughput approaches which may allow to grasp quickly at least the basic features of a complex without its detailed and lengthy structural description.

Here, we discuss the use of suitable spectroscopic probes, such as *o*-nitrotyrosine (NT) or *p*-iodophenylalanine (IF), to study molecular interactions. Both IF and NT are expected to decrease protein fluorescence through two distinct mechanisms. In particular, iodine-containing molecules are known to act as collisional quenchers of Trp fluorescence by promoting non-radiative decay of the excited singlet state through intersystem crossing to an excited triplet state (Berlman 1973; Lakowicz 1999). The non-fluorescent NT, on the other hand, which absorbs radiation in a pH-dependent manner in the wavelength range where Tyr and Trp emit fluorescence, has proven as an efficient acceptor in energy transfer processes, such as those occurring in protein folding (Rischel and Poulsen 1995; Tcherkasskaya and Ptitsyn 1999) and molecular recognition (De Filippis et al. 2006).

We demonstrate that a complementary use of the two probes can be exploited for high throughput characterisation of complexes of Src homology-3 (SH3) domains with their binding peptides. SH3 domains are relatively small protein modules of 50-70 amino acid residues and one of the most widespread motifs, mostly observed in eukaryotic organisms (Kay et al. 2000; Mayer 2001; Tong et al. 2002). Through specific recognition of proline-rich

sequences, they play a crucial role in the formation of multiprotein complexes and networks responsible for signal transduction, cytoskeletal organisation and other cellular processes (Morton et al. 1994; Mayer 2001). In humans, mutations in several SH3 domains are known to cause severe malfunctions leading, amongst others, to inflammatory diseases and to cancer (Smithgall 1995; Dalgarno et al. 1997; Vidal et al. 2001).

Despite even large differences in the sequences of both components, the mode of binding of SH3/peptide complexes is highly reproducible (Larson and Davidson 2000; Di Nardo et al. 2003; Li 2005). In a minimalist approach, two parameters seem to be necessary and sufficient to describe an SH3/peptide complex: the affinity of the interaction and the directionality of the peptide in the binding groove. The target peptides, tyipically 7-10 residues long (Zarrinpar and Lim 2000; Musacchio 2002; Rath et al. 2005; Li 2005), adopt a left-handed polyproline II helix conformation (PP_{II}) and sit in a well defined groove of the SH3 domain which is formed between the so called n-Src- and RT-loops (Musacchio 2002). Because of the intrinsic symmetry of PP_{II} helices, the peptides can in principle be accommodated into the same SH3 cavity in the two possible orientations (Mayer 2001; Musacchio 2002) (Figure 1). Specific consensus motifs, denoted as class I and class II peptides, seemed at first to determine the peptide orientation. However, more recently, several non-standard sequences which are

difficult to be classified as class-I or class-II binders have been identified (Cestra et al. 1999; Mongiovì et al. 1999; Kang et al. 2000; Tong et al. 2002; Jia et al. 2005), thus suggesting that the orientation cannot be reliably predicted solely from the peptide sequence (Musi et al. 2006; Cesareni et al. 2002; Li 2005). Fast and reliable determination of the affinity and orientation should therefore be sufficient to provide at least the general features of the interaction without undergoing complete structure determination of all possible SH3/peptide complexes.

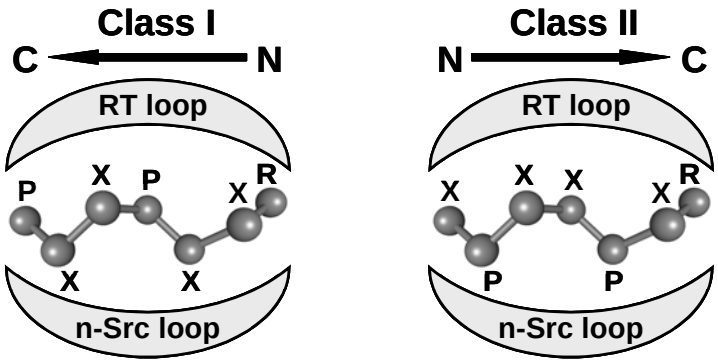


Fig. 1. Schematic representation of the mode of binding of class-I and class-II peptides into the conserved groove of SH3 domains (adapted from Mayer 2001). The consensus sequences of the proline-rich peptides are indicated, using X for any amino acid, whereas P and R indicate semi-conserved prolines and arginines.

Both IF and NT derivatives are ideally suited for characterizing the binding mode of target peptides to SH3 domains, which contain as part of their signature motif a highly conserved Trp (Fernandez-Ballester et al. 2004; Larson and Davidson 2000). The fluorescence quenching effects determined by the amino acid derivatives onto the fluorescence signal of the conserved Trp can therefore be used to define the peptide orientation. Additionally, the presence of NT, which has absorption properties totally different from those of other protein amino acids, is expected to generate a unique circular dichroism (CD) signal in the near-UV/Vis spectrum, upon complex formation.

We have chosen here the SH3 domain (Myo3-SH3) from the type-I myosin isoform Myo3 from *S. cerevisae* as a model system. Myo3-SH3 is of potential medical interest, since mutations in the corresponding human orthologs result in the Wiskott-Aldrich syndrome, a severe immunodeficiency related to alteration of key cellular processes, such as endocytosis and cytoskeleton assembly (Anderson et al. 1998; Rodal et al. 2003). The solution structure of Myo3-SH3 and a detailed analysis of its binding mode have recently been described (Musi et al. 2006). We show that, short of doing a full structural characterization of the complex, the orientation of Myo3-SH3/peptide complexes can be easily determined by fluorescence and CD through incorporation of IF or NT at the N- or C-terminal ends of the peptide. Our approach can be used as a valuable and efficient tool for obtaining a low resolution but fast and reliable description of peptide/protein interactions.

MATERIALS AND METHODS

Production of Myo3-SH3 and the synthetic analogs of P2 and P4 peptides. N^{α}-Fmoc-derivatives of NT, IF and other protected amino acids, solvents and reagents for peptide synthesis were purchased from Bachem AG (Bubendorf, Switzerland) or Applied Biosystems (Foster City, CA). Buffers and organic solvents were of analytical grade and purchased from Fluka (Basel, Switzerland).

P2 and P4 peptides, and their analogs containing NT or IF were manually synthesized by the solid-phase Fmoc-method on a chlorotrityl resin (36 mg) (Barlos et al. 1991), derivatized (0.69 μmoles/mg) with the C-terminal amino acid (i.e., Pro). The chlorotrityl resin was used to minimize the release of the C-terminal dipeptide occurring through di-ketopiperazine formation (Rothe and Mazànek 1972). For those analogs having a non-Pro residue at the C-terminus, a Wang resin (46 mg) derivatized (0.65 μmoles/mg) with the Fmoc-amino acid was used. Elongation of the peptide chain was accomplished by double coupling cycle at all steps,

using HATU as an activator (Carpino 1993) and a five-fold molar excess of protected amino acids. Removal of the Fmoc-group was carried out by a 15 min treatment with 20% pyperidine in NMP. Peptide cleavage from the resin and removal of side-chain protecting groups was carried out by treating the peptidyl resin with a mixture of TFA/EDT/H_2O (95:2.5:2.5, v/v) for 90 min at 0°C. After precipitation with ice-cold diethylether, the crude peptides were purified to homogeneity (> 98%) by RP-HPLC on a semi-preparative Vydac (Hesperia, CA) C18 column (1 x 25 cm, 5 μm particle size), equilibrated in aqueous TFA (0.1%) and eluted with an acetonitrile-TFA (0.1%) gradient from 2 to 35 %, at a flow rate of 1.5 ml/min. The absorbance of the effluent was recorded at 226 nm. The chemical identity of the purified peptides was established by high-resolution mass spectrometry on a Mariner ESI-TOF instrument from Perseptive Biosystems (Stafford, TX). Analyses were conducted according to standard manufacturer's procedures and gave mass values in agreement with amino acid composition within 50 ppm accuracy.

Myo3-SH3, spanning residues 1122-1190 of *S. cerevisae* Myo3, was expressed and purified as previously detailed (Musi *et al.*, 2006).

Determination of peptide/protein concentration. Protein/peptide concentrations were determined by UV absorption (Gill and von Hippel 1989) on a Lambda-2 spectrophotometer from Perkin-Elmer (Norwalk, CT). The molar absorptivity value of Myo3-SH3 was taken as 15,220 $M^{-1}\cdot cm^{-1}$ at 280 nm. The concentrations of the synthetic peptides containing NT (P2NT1, P2NT10, P2NT13, P4NT1) or IF (P2IF1, P2IF13) were calculated from the molar absorbance value of NT, 2200 $M^{-1}\cdot cm^{-1}$ at 381 nm (Tcherkasskaya and Ptitsyn 1999), and IF, 260 $M^{-1}\cdot cm^{-1}$ at 280 nm (De Filippis et al. 2002). The concentrations of the unmodified P2 and P4 peptides were estimated by quantitative analysis of amino acid composition, carried out on the corresponding peptide solutions (~2.5 mg/ml), prepared by weighting the lyophilized peptides.

Fluorescence measurements. Fluorescence spectra were recorded on a Perkin-Elmer spectrofluorimeter model LS-50B, equipped with a thermostatted cell-holder. The interaction of P2 and P4 peptides and their analogs (0-120 μM) with Myo3-SH3 (175 nM) was monitored by adding, under gentle magnetic stirring in a 1 cm pathlength cuvette (2 ml), aliquots (2-10 μl) of NT- and IF-peptides stock solutions (2.5-5 mM) to a solution of Myo3-SH3 (2 ml). At each peptide concentration, Myo3-SH3 samples were equilibrated for 5 min at 25 ± 0.2 °C and excited at 295 nm, using an excitation/emission slit of 5 and 10 nm, respectively, and a

scan speed of 240 nm/min. The decrease in fluorescence intensity of Myo3-SH3 at the wavelength where the maximum fluorescence change was observed (i.e., 349 nm for NT-peptides and 360 nm for the unmodified P2 and P4 peptides and IF derivatives) was recorded as a function of ligand concentration, [L]. Fluorescence data were corrected for sample dilution (< 5% of the final volume) and expressed as $\Delta F = F°-F$, where F° and F is the fluorescence of Myo3-SH3 in the absence and presence of the ligand, respectively.

For NT-peptides, fluorescence data were also corrected for IFE (see below and Supplemental Material), since fluorescence intensity is only proportional to the absorbance of the sample up to an optical density of 0.05 units, both at λ_{ex} and λ_{em} (Puchalski et al. 1991; Lakowicz 1999). In our case, IFE becomes significant for ligand concentrations higher than 20 μM. Fluorescence data were corrected using the equation:

$$F_{corr} = F_{obs} \cdot 10^{[(\Delta Aex + \Delta Aem)/2]} \quad (1)$$

where ΔA_{ex} and ΔA_{em} are the observed additional absorbance at the excitation (λ_{ex} = 295 nm) and emission (λ_{em} = 349 nm) wavelengths.

For a simple one-site binding mechanism R + L ⇔ RL, the fluorescence intensity, F, of the receptor, R, at a given concentration of ligand, L, is linearly related to the concentration of the complex [RL], according to the equation $F = [RL]\cdot F_{bound} + [R]_{free}\cdot F_{free}$. Since $[R]_{free} = [R] - [RL]$, then $\Delta F = ([RL]/[R])\cdot \Delta F_{max}$ (Eftink 1997). The data were interpolated with equation 2, using the program Origin 6.0 (MicroCal Inc., Northampton, MA) to obtain the fitting parameters ΔF_{max} and K_d:

$$\Delta F = F°-F = (\Delta F_{max}\cdot[L])/(K_d + [L]) \quad (2)$$

where K_d is the dissociation constant of the complex, RL, and ΔF_{max} is the maximum fluorescence change at infinite concentration of ligand, $[L]_\infty$. Equation 2 assumes that at equilibrium $[L]_{free} \cong [L]_{tot}$ and thus it is valid only when $K_d >> [R]$, as usually observed for SH3 complexes (Kay et al. 2000).

CD measurements. CD spectra were recorded on a Jasco (Tokyo, Japan) model J-810 spectropolarimeter. Near-UV/Vis spectra were recorded in a 0.5 cm cell, at a scan-speed of 50 nm/min, with a response time of 2 s, and resulted from the average of six accumulations.

After base line subtraction, ellipticity values, θ, were expressed in millidegree (mdeg), without further data elaboration.

Computational analysis. Structural analysis of SH3/peptide complexes was performed on 14 PDB entries (1abo, 1a0n, 1azg, 1bbz, 1fyn, 1io6, and 1rlq as examples of class-I peptides; 1aze, 1cka, 1ckb, 1efn, 1n5z, 1prm, and 1qwe as examples of class-II peptides). For NMR structures, the best-representative model in the NMR ensemble was selected using the program OLDERADO available on-line at the site http://pqs.ebi.ac.uk/pqs-nmr.html. In most cases, the Cβ-Cβ distances were measured from the conserved Trp of the SH3 domain to the N- or C-terminal amino acid of the ligand. In a few SH3 complexes, where the coordinates of the ligand peptides extend far beyond the segment that actually contacts the SH3 surface, the Cβ-Cβ distances were measured from the conserved Trp to the first residue (from the N-terminus or from the C-terminus, respectively) which is at a distance < 6 Å from any atom of the SH3 structure. The structures of P2 analogs in the SH3-bound state were modeled on the 2btt structure (Musi et al. 2006) by keeping unchanged the positions of all atoms in the lowest energy HADDOCK (Dominguez et al. 2003) structure and substituting His1' or Pro10' with the NT and IF side-chains.

RESULTS AND DISCUSSION

Strategy of the work. Structural analysis based on representative SH3/peptide complexes from the Protein Data Bank (PDB) (see Methods) shows that one of the terminal ends of the bound peptide is much closer to the conserved Trp than the other: in class-I peptides, the N-terminus is within 4.7 ± 0.9 Å (C□-C□ distance) from the conserved Trp, whereas the C-terminus is 18.8 ± 1.5 Å away (Figure 2A). This pattern is inverted in class-II peptides, although with slightly larger distance variability. For a class-I peptide, we can therefore confidently expect that incorporation of IF or NT at the N-terminal end of the binding peptide should bring these probes close enough to the conserved Trp to specifically quench its fluorescence or, as in the case of NT, generate a new signal in the near-UV/Vis CD spectrum of the SH3/peptide complex. Conversely, when NT or IF are incorporated far from the Trp pocket (i.e., at position 10 or further), the direct collisional quenching of IF should be unable to operate and the efficiency of Trp-to-NT energy transfer, which is strongly dependent on distance (Wu and Brand 1994; Dos Remedios and Moens 1995; Selvin 1995; Eftink 1997), is

expected to be reduced, together with the CD signal of NT. On the other hand, for class-II peptides, the position dependent spectroscopic effects are reversed.

Hence, we tested our strategy on a known model system formed by Myo3-SH3 and the P2 peptide. We have recently reported the solution structure of Myo3-SH3 and demonstrated that the domain recognizes class-I SH3 peptides (Musi et al. 2006). We therefore selected one of these peptides, P2, which has high affinity for Myo3-SH3 and for which we have a reliable model in complex with Myo3-SH3 (Musi et al. 2006). A distance (Cβ-Cβ) of 6 Å from the conserved Trp39 to position 1 of P2 was estimated, whereas the peptide C-terminus is at about 22 Å from Trp39 (Figure 2B).

It is important to note that Myo-SH3 contains an additional non-conserved Trp, at position 52. The presence of this extra Trp-residue should not however alter our analysis, since it is within the Trp pocket and spatially close to Trp39 (Trp39-Trp52, Cβ-Cβ distance 4.9 Å).

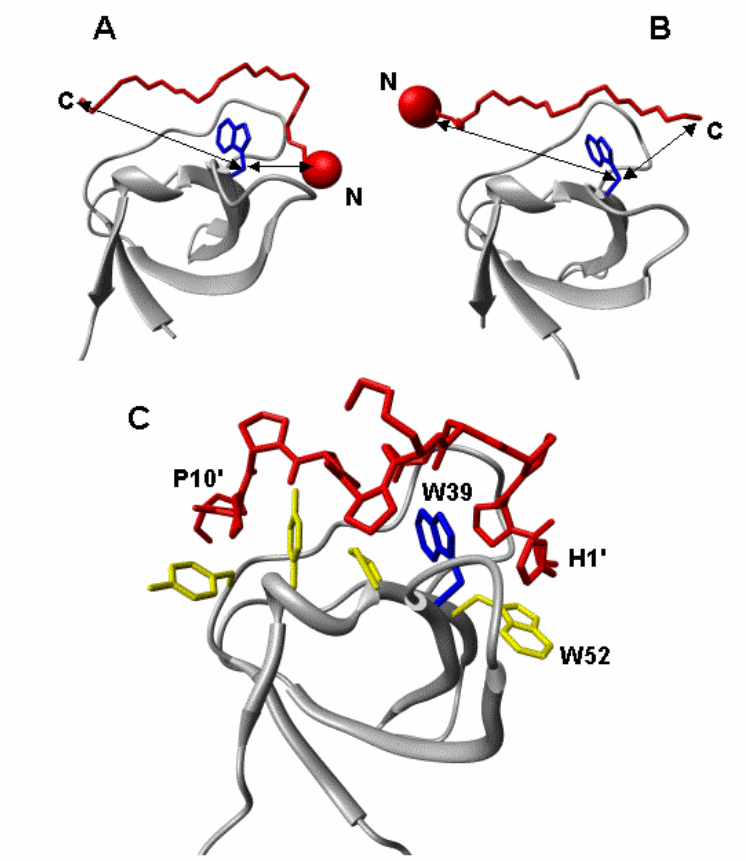


Fig. 2. Structural analysis of SH3 complexes to illustrate the strategy proposed. Structures of two representative SH3 domains (ribbon drawing, gray) bound to class-I (*A*) (1abo, Musacchio et al. 1994) and class-II (*B*) (1cka, Wu et al. 1995) peptides (Cα-trace, red). The side chain of the conserved Trp-residue in the SH3 domain is shown in blue, while the distances (Cβ-Cβ atoms) between the peptide termini and Trp are shown by arrows. The incorporation of a suitable spectroscopic probe (red sphere), like NT or IF, is expected to perturb the intrinsic fluorescence of the conserved Trp in the SH3 domain, in a way that is strongly dependent on the orientation of the binding peptide on the SH3 surface. For

instance, if the probe is incorporated at the N-terminus of the ligand peptide, the fluorescence of SH3 is expected to be quenched only by a class-I peptide, whereas it should be essentially unaffected if the ligand is a class-II peptide. (*C*) Structure of the P2/Myo3-SH3 complex (Musi *et al*. 2006), showing relevant interactions of the ligand peptide (stick, red) with the SH3 domain (ribbon drawing, gray). The side chains of key amino acids of Myo3-SH3 are also explicitly shown (stick, yellow), while the conserved Trp39 is shown in blue.

Different P2 analogs were synthesized, incorporating either NT or IF in different positions along the P2 sequence (Table 1). Unmodified P2, P2NT1, P2NT10, and P2NT13 were prepared to check the effect of the introduction of NT at the N- or C-termini of the peptide. Likewise, the analogs P2IF1 and P2IF13, containing IF at positions 1 and 13, were produced to test the effect of IF. In the case of P2NT13 and P2IF13, the P2 sequence was extended C-terminally with the highly flexible segment –G-G-X-G (X = IF or NT) to have the spectroscopic probes outside the P2/SH3 interface, to exclude possible interferences with complex formation. Finally, the unmodified and the NT1 analog of the P4 peptide, P4NT1, were synthesized as negative controls, since P4 does not bind Myo3-SH3 (Musi et al. 2006).

Incorporation of NT or IF does not alter the binding mode of P2 to Myo3-SH3. The chemical strategy outlined above requires that incorporation of NT or IF does not alter the binding mode of P2 analogs to Myo3-SH3. To test this assumption, we measured the affinities of NT- and IF-analogs of P2 by exploiting their ability to quench the fluorescence of Myo3-SH3 (see Figg. 3 and 4). As recently demonstrated for the binding of the Tyr3NT analog of hirudin to thrombin (De Filippis et al. 2006), quenching of Myo3-SH3 fluorescence by NT-analogs (Figure 3A,B) is predominantly caused by strong resonance energy transfer occurring between the Trp residues of the SH3 domain (i.e., the donor) and NT of P2-peptides (i.e., the acceptor) (Figure 3C). On the other hand, IF-analogs of P2 remarkably decrease Myo3-SH3 fluorescence (Figure 4A,B) by a specific and direct effect of iodine in the SH3 pocket, by promoting non-radiative decay of the excited state of Trp-residues through a collisional mechanism (Berlman 1973; Lakowicz 1999).

Analysis of fluorescence binding data is sometimes complicated by the contribution of the intrinsic emission of the ligand to the fluorescence of the receptor and by unspecific effects, including static and dynamic quenching, as well as Inner Filter Effect (IFE) (Selvin 1995; Eftink 1997). We carefully analysed their contribution in the Supplementary Materials to make sure that these effects do not alter our analysis. In the case of NT peptides, the emission of free NT is negligible at 349 nm (i.e., the wavelength at which the fluorescence signal was

measured) (Supplemental Fig. S1) and therefore we conclude that addition of NT peptides does not interfere with fluorescence measurements. The data shown in Figure 3 and Supplemental Figure S2 suggest that free NT, taken as a suitable model compound, only slightly affects Myo3-SH3 fluorescence by either static and dynamic quenching, as expected for nitro-compounds which, in the absence of energy transfer, are known to quench the fluorescence of aromatic hydrocarbons by a mixed static and dynamic mechanism (Sawicki et al. 1964; Dreeskamp et al. 1975). In addition, the contribution of IFE on the binding data of high affinity NT analogs (i.e., P2NT1 and P2NT13) is marginal (data not shown). In all cases, the fluorescence data were anyway corrected for IFE using equation 1 and the adequacy of the correction method was carefully established. Unspecific effects can be fully excluded for IF analogs, since IF weakly absorbs in the near-UV region and is essentially non-fluorescent. As a consequence, free IF only marginally reduces Myo3-SH3 fluorescence (Figure 4B).

The fluorescence data, corrected for IFE in the case of NT analogs, were analyzed assuming a one-site binding mechanism (see Methods). The excellent fit of the experimental data to equation 2 is a stringent, albeit indirect, proof of a 1:1 binding stoichiometry (Copeland 2000). The affinity data indicate that incorporation of NT or IF at position 1 or 13 does not change the affinity of the P2-derivatives for Myo3-SH3 (Table 1). This is in agreement with our structural analysis which suggests that, when at position 1, either NT or IF can be easily accommodated into the Trp pocket with only minimal adjustments of the highly flexible 35-37 segment of Myo3-SH3 (Musi et al. 2006). More generally, our results are consistent with the fact the Trp pocket of SH3 is quite “tolerant” in binding the target peptides, outside the polyproline consensus sequence, and this is also well documented by the high structural variability of the amino acid that in the ligand peptides interacts with the conserved Trp (Musacchio 2002; Li 2005). Likewise, incorporation of NT or IF at position 13 does not interfere with binding, since the modifications involve residues outside the P2/Myo3-SH3 interface. Only when Pro10 was replaced with NT, a two-fold decrease of the affinity of P2NT10 was observed. This effect is likely caused by the replacement of a structurally constrained Pro (MacArthur and Thornton 1991) with the more flexible NT side-chain (Searle and Williams 1992).

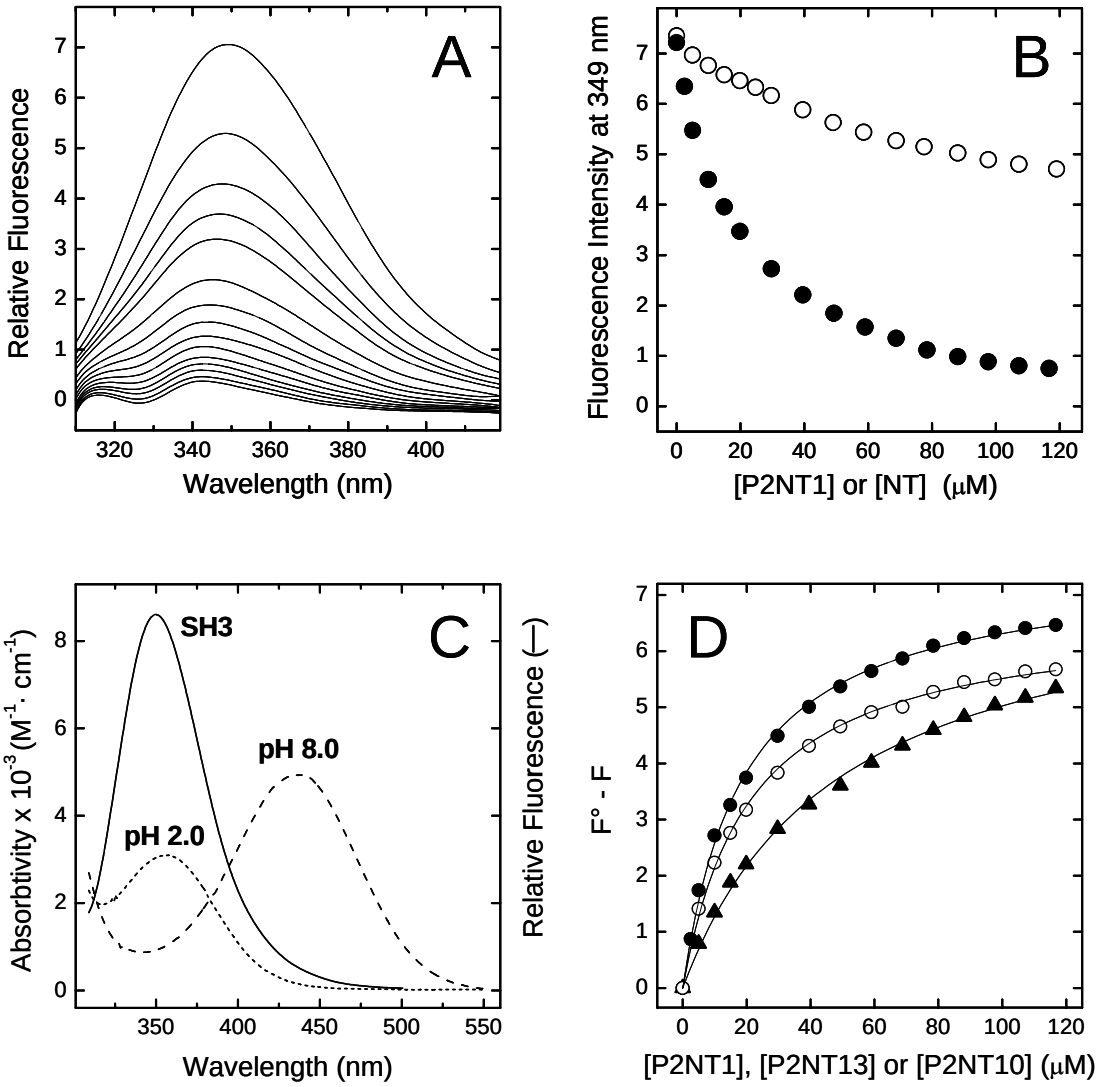


Fig. 3. Binding of NT-analogs of P2 to Myo3-SH3 monitored by Trp-to-NT fluorescence energy transfer. (*A*) Fluorescence spectra of Myo3-SH3 (175 nM) in the presence of increasing concentrations of P2NT1 (0-120 μM). (*B*) Plot of the fluorescence intensity of Myo3-SH3 as a function of P2NT1 concentration (●). As a control, the data relative to free NT (○) are also reported. Protein samples were excited at 295 nm and fluorescence data were corrected for IFE, according to equation 1 (see Methods). (*C*) Superposition of the fluorescence spectrum of Myo3-SH3 (continuous line), obtained after excitation of the sample at 295 nm, with the absorption spectrum of P2NT1 at pH 2 (dotted line) and pH 8 (dashed line). (*D*) Determination of K_d values of the Myo3-SH3 complexes with NT-analogs: P2NT1 (●-●), P2NT10 (▲-▲), and P2NT13 (○-○). Corrected fluorescence intensities were expressed as $F° - F$, where $F°$ is the intensity of Myo3-SH3 in the absence of ligand, and the data points fitted by equation 2 (continuous lines) to yield K_d and ΔF_{max} (see Methods). All measurements were carried out at 25 ± 0.2°C by exciting the protein samples at 295 nm in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% (w/v) PEG 8000 and 0.2 M NaCl.

Our data concurrently support the assumption that the mutations introduced in P2 do not alter the binding mode of the resulting analogs to the SH3 domain and allow us to interpret confidently the position-dependent effects of the mutations on the intrinsic spectroscopic properties of P2 analogs or on the spectral changes evoked in Myo3-SH3 upon ligand binding.

Probing the binding mode of P2 to Myo3-SH3 by incorporation of NT

Fluorescence Resonance Energy Transfer. The efficiency of energy transfer depends on the donor quantum yield, on the extent of spectral overlap of the emission spectrum of the donor with the absorption spectrum of the acceptor, on the donor-acceptor distance, and on their relative orientation (Lakowicz 1999). However, for a given donor-acceptor pair, the distance factor is predominant (Wu and Brand 1994; Dos Remedios and Moens 1995) and therefore we have exploited here the distance-dependence of FRET efficiency to probe the orientation of P2 in complex with Myo3-SH3.

The presence of NT at position 1 of P2 decreases Myo3-SH3 fluorescence up to 90% (Figure 3A,B), in agreement with the extensive overlap of the emission spectrum of the SH3 domain (i.e., the donor) with the absorption spectrum of P2NT1 (i.e. the acceptor) (Figure 3C) and with the short donor-acceptor distance in the SH3 complex (<6 Å). Concomitantly, the λ_{max} value is blue-shifted by 7 nm. This latter result parallels those obtained with the unmodified P2 peptide (not shown), where a similar shift of the λ_{max} was observed, and is consistent with the partial shielding of Trp residues from water upon ligand binding. An alternative possibility is that one of the two Trp-residues emits at longer wavelengths and that preferential quenching of the fluorescence of this Trp would result in the blue-shift of λ_{max}. However, this possibility is unlikely because the two Trp-residues in Myo3-SH3 are exposed to the solvent to a similar extent (Musi et al. 2006) and therefore they are expected to display identical emission properties. In addition, some other SH3 domains, like the N-terminal SH3 of Grb-2 (Vidal et al. 1999), contain only the conserved Trp and still have λ_{max} values very similar to that of Myo3-SH3. Despite the similar affinities of the two complexes, quenching of Myo3-SH3 fluorescence at saturating concentrations of ligand is lower for P2NT13 (Figure 3D), where NT is far from the Trp pocket (>22 Å). The reduced extent of quenching is a consequence of the weaker efficiency of FRET at longer Trp-to-NT distances and is even more significant if one considers that the unmodified P2 peptide is also able to quench SH3 fluorescence (Figure 4D). In the P2NT13/Myo3-SH3 complex, the efficiency of energy

transfer is reduced, but not abolished, because the Trp-to-NT distances are still within the Förster's distance of the donor-acceptor pair (26 Å) (Steiner et al. 1991). Taken together, these results are fully consistent with a class-I orientation of P2 in binding Myo3-SH3.

Circular Dichroism. Further evidence strongly supporting a class-I orientation of P2 was obtained by comparison of the near-UV/Vis CD spectra of free and SH3-bound NT-peptides (Figure 5). The CD spectrum of free SH3 displays an intense positive band centered at 268 nm and a pronounced fine structure in the 285-300 nm range, characteristic of Trp absorption (Strickland, 1974). As expected, beyond 305 nm, the signal approaches to zero. The spectrum of free P2NT1 presents a marginally weak negative signal in the 300-390 nm range and a low-intensity positive band at 280 nm, assigned to the $^1A_{1g} \rightarrow {}^1B_{1u}$ transition of NT (Meloun et al. 1968; De Filippis et al. 2006), that acquires some rotational strength, due to the presence of several rigid prolines which stabilize the PP_{II} conformation.

When SH3 and P2NT1 are mixed to form the non-covalent complex, a new band appears at about 350 nm. Considering that natural amino acids do not absorb beyond 305-310 nm, this band can solely originate from the $^1A_{1g} \rightarrow {}^1B_{2u}$ transition of NT that becomes optically active by induction of chirality of NT upon binding in the Trp pocket of the SH3 domain.

Furthermore, the λ_{max} value of the newly generated band suggests that NT binds in the Trp pocket of Myo3-SH3 in its neutral –OH form, which maximally absorbs at 350 nm (Meloun et al. 1968) (Figure 3C). Similar results have been recently reported for the binding of NT-analogs of hirudin to thrombin (De Filippis et al. 2006). Strikingly, when NT is incorporated at position 13, in the highly flexible segment –G-G-NT-G and outside the binding interface of P2 with Myo3-SH3, this band is absent (Figure 5, *Inset*).

These findings therefore provide conclusive support to the class-I orientation of P2 to Myo3-SH3 (Musi et al. 2006) and suggest that the combined use of FRET and CD techniques can be successfully exploited for characterizing peptide/protein interactions.

Table 1. *Binding Data of P2 and P4 Analogs to Myosin-3 SH3 Domain*

Peptide name	Peptide sequence	K_d (μM) [a]
P2	**H**-P-P-R-K-P-P-P-P-**P**	18.8 ± 0.9
P2NT1	**NT**-P-P-R-K-P-P-P-P-P	17.5 ± 0.8

P2NT10	H-P-P-R-K-P-P-P-P-**NT**-G	37.6 ± 1.8
P2NT13	H-P-P-R-K-P-P-P-P-P--G-G-**NT**-G	19.1 ± 0.9
P2IF1	**IF**-P-P-R-K-P-P-P-P-P	16.5 ± 0.4
P2IF13	H-P-P-R-K-P-P-P-P-P--G-G-**IF**-G	18.5 ± 0.8
P4	H-F-K-H-P-F-P-P-P-P	n.d. [b]
P4NT1	**NT**-F-K-H-P-F-P-P-P-P	n.d. [b]

[a] K_d values were obtained by fitting the fluorescence data reported in Figures 3 and 4 to equation 2.

[b] n.d., not determined because the fluorescence changes were negligible.

Probing the binding mode of P2 to Myo3-SH3 by incorporation of IF. Although the work carried out with the NT analogs indicates that it is possible to use FRET to determine the orientation of an SH3 binding peptide, some limitations might be envisaged when it is not easy to make reliable assumptions on the relative distance of the N- and C-termini from the conserved Trp. To overcome these problems, we exploited the effect of IF on the fluorescence signal of Myo3-SH3, since the radius of action of this derivative is much more limited (3-6 Å) than that observed by FRET (Berlamn 1973; Selvin 1995).

The incorporation of IF at position 1 of P2 decreases the fluorescence intensity of Myo3-SH3 by ~50%, with a concomitant blue-shift in the λ_{max} value of 6-7 nm (Figure 4A,B). As for P2NT1 binding, the blue shift in the emission of Myo3-SH3 reflects the less polar environment of Trp39 and Trp52 upon ligand binding, while the strong quenching of fluorescence is caused by the specific and direct contact of the iodine atom of IF in the Trp pocket. Conversely, when IF is incorporated in the fraying C-terminus of P2IF13, the quenching effect is remarkably lower and identical to that of the unmodified P2 (Figure 4C,D). These results highlight the strong position-dependent effect of IF on the SH3 fluorescence. When located outside the ligand-SH3 interaction surface, IF is too far from the Trp pocket to quench the intrinsic fluorescence of Myo3-SH3, thus providing key information on the orientation of the ligand peptide in the SH3 complex.

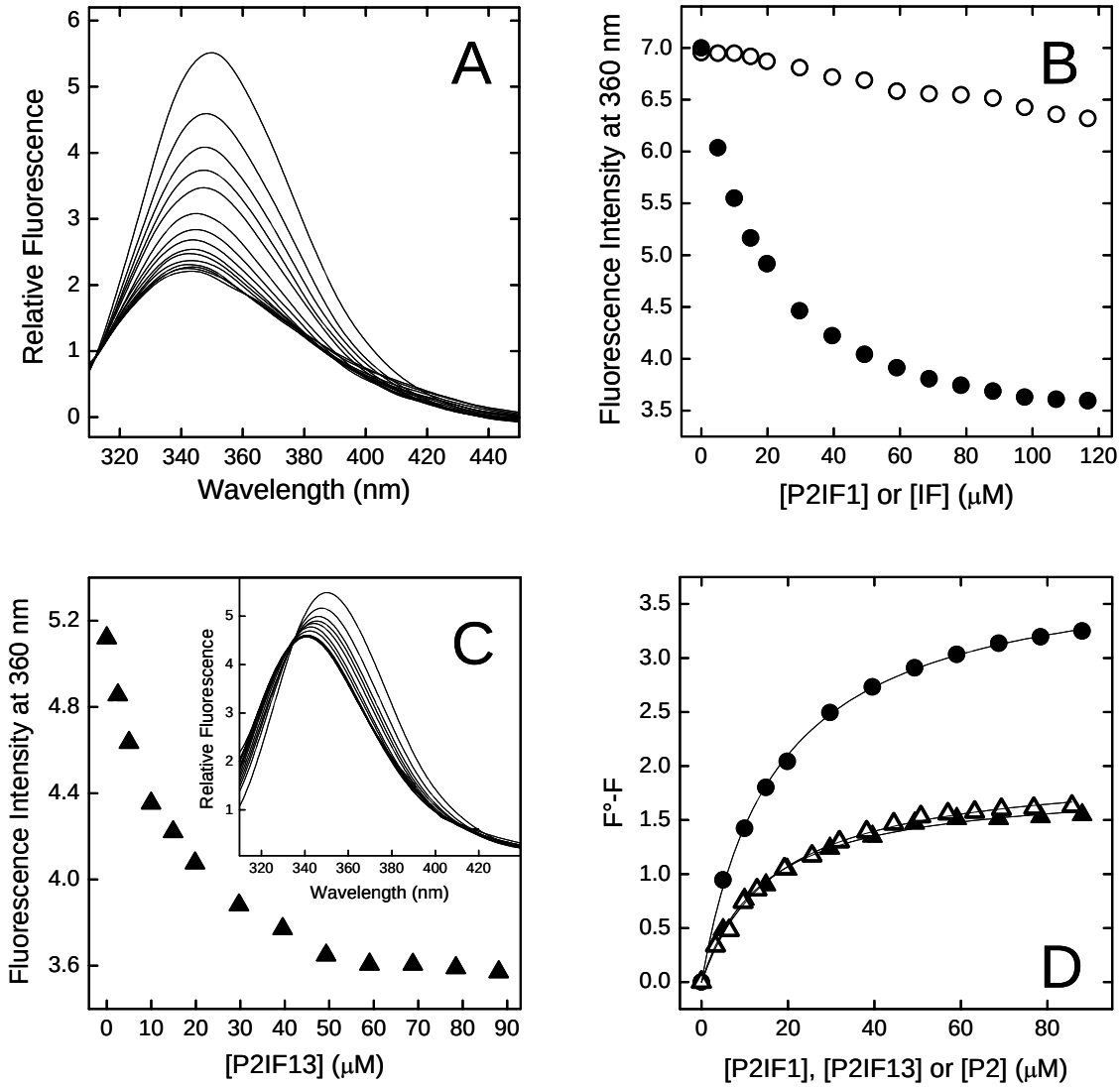


Fig. 4. Binding of IF analogs of P2 to Myo3-SH3 monitored by collisional quenching of Myo3-SH3 fluorescence. (*A*) Fluorescence spectra of Myo3-SH3 (175 nM) in the presence of increasing concentrations of P2IF1 (0-120 μM). (*B*) Change in the fluorescence intensity of Myo3-SH3 as a function of P2IF1 concentration (●). As a control, the effect of free IF is also reported (○). (*C*) Plot of the fluorescence intensity of Myo3-SH3 as a function of P2IF13 concentration (▲). *Inset*: Fluorescence spectra of Myo3-SH3 (175 nM) in the presence of increasing concentrations of P2IF13. (**D**) Determination of K_d values of SH3 complexes with IF-analogs: P2IF1 (●-●) and P2IF13 (▲-▲). For comparison, the binding data of the unmodified P2 peptide to Myo3-SH3 are also included (△-△). All spectra were recorded at 25 ± 0.2°C by exciting the protein samples at 295 nm in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% (w/v) PEG 8000 and 0.2 M NaCl. Continuous lines represent the best fit of the data points to equation 2.

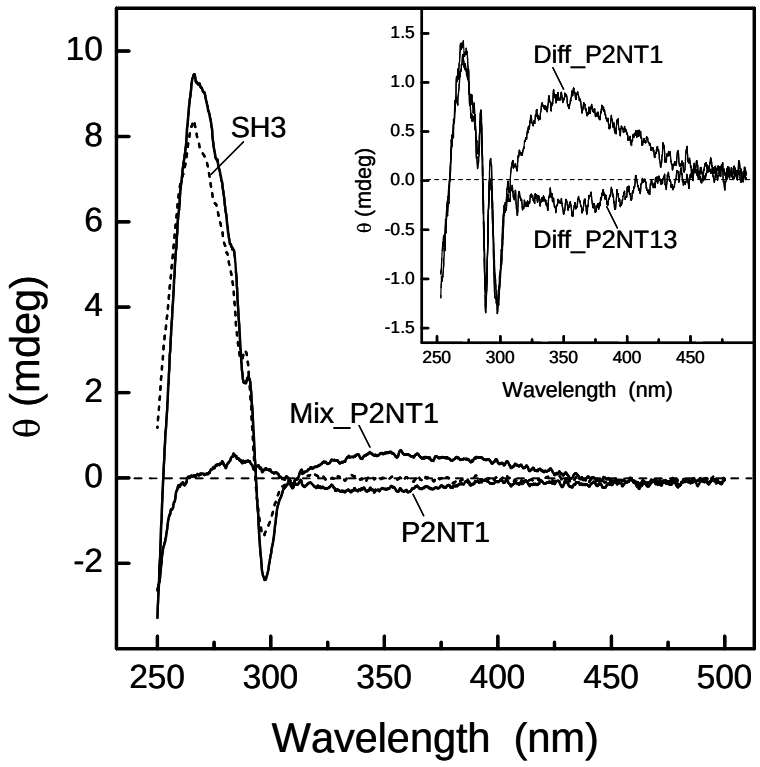


Fig. 5. Binding of NT-analogs of P2 to Myo3-SH3 monitored by circular dichroism. CD spectra of free Myo3-SH3 (65 μM, ---) and P2NT1 (168 μM, —) were recorded in 5 mM Tris-HCl buffer, pH 8.0, containing 0.1% (w/v) PEG 8000, 0.2 M NaCl. Myo3-SH3 and P2NT1 were alternatively mixed in the same molar ratio (65 μM:168 μM) to yield about 85% of bound SH3 (Mix_P2NT1, —). *Inset*: difference spectra (Diff_P2NT1 and Diff_P2NT13) obtained by subtracting the theoretical sum spectra of free SH3 and NT-peptides from the corresponding spectra of the experimental complexes (Mix_P2NT1 and Mix_P2NT13). Measurements were carried out at 25 ± 0.2°C in a 0.5 cm quartz cuvette and subtracted for the corresponding base-line. Ellipticity data, θ, were expressed in millidegree (*mdeg*), without further normalization.

CONCLUSIONS

The development of new spectroscopic tools for studying protein-protein interactions finds applications which span from structural biology to drug discovery (Hovius et al. 2000). For this purpose, several non-natural amino acids with physico-chemical properties (e.g., size, polarity, hydrogen bonding properties) similar to those of the corresponding natural amino acids, but with distinct spectral features have been used to monitor protein folding and binding processes (Cornish et al. 1994; Twine and Szabo 2003; De Filippis et al. 2004, 2006). In this study, we have shown that incorporation of NT or IF can be effectively used to obtain fast and reliable information on the affinity and the orientation of SH3 binding peptides, which are the parameters necessary and sufficient to describe, albeit at low resolution, the

structural features of peptide/SH3 complexes. While easily introduced in peptides by standard solid-phase synthetic methods, these spectroscopic probes exploit different physical processes (e.g., FRET, chiroptical properties, direct fluorescence quenching) which can be used in a complementary fashion, as also integrated by comparative modelling. The different spectroscopic effects elicited by NT or IF at different positions along the ligand peptide sequence can be confidently exploited to generate low-resolution models on a large scale with no need of solving the detailed atomic structure of each complex.

Here, we have reported the successful application of our chemical strategy to an SH3/peptide complex, but the same approach may have vast applications in screening studies also of other interacting domains of potential pharmacological interest, in which knowledge of the binding directionality is important (see Li 2005 and citations therein). Finally, we believe that the possibility of approaching the study of complex networks of interactions in a high throughput fashion will open new avenues to the development of new drugs.

This charter has been adapted from:

De Filippis, V., Draghi, A., Frasson, R., Grandi, C., Musi, V., Fontana, A., and Pastore, A. (2007) *o*-Nitrotyrosine and *p*-iodophenilalanine as spectroscopic probes for structural characterization of SH3 complexes. *Protein Sci.* **16**, 1257-1265.

Appendix

Appendix A

Abbreviations and Symbols

Å	Angstrom
aa	amino acid
Da	Dalton
DTT	Dithiothreitol
EDTA	Ethylene Diamino Tetracetic Acid
IPTG	IsoPropyl-β-D-ThioGalactopyranoside
LB	Luria Bertani liquid medium
MW	Molecular Weight
PEG	PolyEthylene Glycol
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	SDS-PolyAcrylamide Gel Electrophoresis
ESI	Electrospray ionization
GlcNAc	N-Acetylglucosamine
HPLC	High-pressure liquid chromatography
m/z	Mass to charge ratio
MS	Mass spectrometry
RP	Reverse-phase
HPLC	High-pressure liquid chromatography
SEC	Size-exclusion chromatography
TFA	Trifluoroacetic acid
TOF	Time-of-flight
w/v	Weight/volume
UV	ultraviolet
Tris	Tris(hydroxymethyl)aminomethane
Pre2	Recombinant prethrombin-2
rThb	Recombinant human α-thrombin
hThb	Plasma-derived human α-thrombin
R73A	α-Thrombin containing the Arg73 → Ala mutation
R101A	α-Thrombin containing the Arg101 → Ala mutation
G25S	α-Thrombin containing the Gly25 → Ser mutation

Amino acids

Ala	A	Alanine
Arg	R	Arginine
Asp	D	Aspartic acid
Asn	N	Asparagine
Cys	C	Cysteine
Gly	G	Glycine
Gln	Q	Glutamine
Glu	E	Glutamic acid
His	H	Histidine
Ile	I	Isoleucine
Lys	K	Lysine
Leu	L	Leucine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Tyr	Y	Tyrosine
Trp	W	Tryptophan
Val	V	Valine

Appendix B

Thrombin Numbering Scheme

Chym	T1h	F1g	G1f	S1e	G1d	E1c	A1b	D1a	C1	G2	L3	R4
Ch-A	1	2	3	4	5	6	7	8	9	10	11	12
Ch-B	-	-	-	-	-	-	-	-	-	-	-	-
ProT	285	286	287	288	289	290	291	292	293	294	295	296

Chym	P5	L6	F7	E8	K9	K10	S11	L12	E13	D14	K14a	T14b
Ch-A	13	14	15	16	17	18	19	20	21	22	23	24
Ch-B	-	-	-	-	-	-	-	-	-	-	-	-
ProT	297	298	299	300	301	302	303	304	305	306	307	308

Chym	E14c	R14d	E14e	L14f	L14g	E14h	S14i	Y14j	I14k	D14l	G14m	R15
Ch-A	25	26	27	28	29	30	31	32	33	34	35	36
Ch-B	-	-	-	-	-	-	-	-	-	-	-	-
ProT	309	310	311	312	313	314	315	316	317	318	319	320

Chym	I16	V17	E18	G19	S20	D21	A22	E23	I24	G25	M26	S27
Ch-A	37	38	39	40	41	42	43	44	45	46	47	48
Ch-B	1	2	3	4	5	6	7	8	9	10	11	12
ProT	321	322	323	324	325	326	327	328	329	330	331	332

Chym	P28	W29	Q30	V31	M32	L33	F34	R35	K36	S36a	C37	Q38
Ch-A	49	50	51	52	53	54	55	56	57	58	59	60
Ch-B	13	14	15	16	17	18	19	20	21	22	23	24
ProT	333	334	335	336	337	338	339	340	341	342	343	344

Chym	E39	L40	L41	C42	G43	A44	S45	L46	I47	S48	D49	R50
Ch-A	61	62	63	64	65	66	67	68	69	70	71	72
Ch-B	25	26	27	28	29	30	31	32	33	34	35	36
ProT	345	346	347	348	349	350	351	352	353	354	355	356

Chym	W51	V52	L53	T54	A55	A56	H57	C58	L59	L60	Y60a	P60b
Ch-A	73	74	75	76	77	78	79	80	81	82	83	84
Ch-B	37	38	39	40	41	42	43	44	45	46	47	48
ProT	357	358	359	360	361	362	363	364	365	366	367	368

Chym	P60c	W60d	D60e	K60f	N60g	F60h	T60i	E61	N62	D63	L64	L65
Ch-A	85	86	87	88	89	90	91	92	93	94	95	96
Ch-B	49	50	51	52	53	54	55	56	57	58	59	60
ProT	369	370	371	372	373	374	375	376	377	378	379	380

Chym	V66	R67	I68	G69	K70	H71	S72	R73	T74	R75	Y76	E77
Ch-A	97	98	99	100	101	102	103	104	105	106	107	108
Ch-B	61	62	63	64	65	66	67	68	69	70	71	72
ProT	381	382	383	384	385	386	387	388	389	390	391	392

Chym	R77a	N78	I79	E80	K81	I82	S83	M84	L85	E86	K87	I88
Ch-A	109	110	111	112	113	114	115	116	117	118	119	120
Ch-B	73	74	75	76	77	78	79	80	81	82	83	84
ProT	393	394	395	396	397	398	399	400	401	402	403	404

Chym	Y89	I90	H91	P92	R93	Y94	N95	W96	R97	E97a	N98	L99
Ch-A	121	122	123	124	125	126	127	128	129	130	131	132
Ch-B	85	86	87	88	89	90	91	92	93	94	95	96
ProT	405	406	407	408	409	410	411	412	413	414	415	416

Chym	D100	R101	D102	I103	A104	L105	M106	K107	L108	K109	K110	P111
Ch-A	133	134	135	136	137	138	139	140	141	142	143	144
Ch-B	97	98	99	100	101	102	103	104	105	106	107	108
ProT	417	418	419	420	421	422	423	424	425	426	427	428

Chym	V112	A113	F114	S115	D116	Y117	I118	H119	P120	V121	C122	L123
Ch-A	145	146	147	148	149	150	151	152	153	154	155	156
Ch-B	109	110	111	112	113	114	115	116	117	118	119	120
ProT	429	430	431	432	433	434	435	436	437	438	439	440

Chym	P124	D125	R126	E127	T128	A129	A129a	S129b	L129c	L130	Q131	A132
Ch-A	157	158	159	160	161	162	163	164	165	166	167	168
Ch-B	121	122	123	124	125	126	127	128	129	130	131	132
ProT	441	442	443	444	445	446	447	448	449	450	451	452

Chym	G133	Y134	K135	G136	R137	V138	T139	G140	W141	G142	N143	L144
Ch-A	169	170	171	172	173	174	175	176	177	178	179	180
Ch-B	133	134	135	136	137	138	139	140	141	142	143	144
ProT	453	454	455	456	457	458	459	460	461	462	463	464

Chym	K145	E146	T147	W148	T149	A149a	N149b	V149c	G149d	K149e	G150	Q151
Ch-A	181	182	183	184	185	186	187	188	189	190	191	192
Ch-B	145	146	147	148	149	150	151	152	153	154	155	156
ProT	465	466	467	468	469	470	471	472	473	474	475	476

Chym	P152	S153	V154	L155	Q156	V157	V158	N159	L160	P161	I162	V163
Ch-A	193	194	195	196	197	198	199	200	201	202	203	204
Ch-B	157	158	159	160	161	162	163	164	165	166	167	168
ProT	477	478	479	480	481	482	483	484	485	486	487	488

Chym	E164	R165	P166	V167	C168	K169	D170	S171	T172	R173	I174	R175
Ch-A	205	206	207	208	209	210	211	212	213	214	215	216
Ch-B	169	170	171	172	173	174	175	176	177	178	179	180
ProT	489	490	491	492	493	494	495	496	497	498	499	500

Chym	I176	T177	D178	N179	M180	F181	C182	A183	G184	Y184a	K185	P186
Ch-A	217	218	219	220	221	222	223	224	225	226	227	228
Ch-B	181	182	183	184	185	186	187	188	189	190	191	192
ProT	501	502	503	504	505	506	507	508	509	510	511	512

Chym	D186a	E186b	G186c	K186d	R187	G188	D189	A190	C191	E192	G193	D194
Ch-A	229	230	231	232	233	234	235	236	237	238	239	240
Ch-B	193	194	195	196	197	198	199	200	201	202	203	204
ProT	513	514	515	516	517	518	519	520	521	522	523	524

Chym	S195	G196	G197	P198	F199	V200	M201	K202	S203	P204	F204a	N204b
Ch-A	241	242	243	244	245	246	247	248	249	250	251	252
Ch-B	205	206	207	208	209	210	211	212	213	214	215	216
ProT	525	526	527	528	529	530	531	532	533	534	535	536

Chym	N205	R206	W207	Y208	Q209	M210	G211	I212	V213	S214	W215	G216
Ch-A	253	254	255	256	257	258	259	260	261	262	263	264
Ch-B	217	218	219	220	221	222	223	224	225	226	227	228
ProT	537	538	539	540	541	542	543	544	545	546	547	548

Chym	E217	G219	C220	D221	R221a	D222	G223	K224	Y225	G226	F227	Y228
Ch-A	265	266	267	268	269	270	271	272	273	274	275	276
Ch-B	229	230	231	232	233	234	235	236	237	238	239	240
ProT	549	550	551	552	553	554	555	556	557	558	559	560

Chym	T229	H230	V231	F232	R233	L234	K235	K236	W237	I238	Q239	K240
Ch-A	277	278	279	280	281	282	283	284	285	286	287	288
Ch-B	241	242	243	244	245	246	247	248	249	250	251	252
ProT	561	562	563	564	565	566	567	568	569	570	571	572

Chym	V241	I242	D243	Q244	F245	G246	E247
Ch-A	289	290	291	292	293	294	295
Ch-B	253	254	255	256	257	258	259
ProT	573	574	575	576	577	578	579

In the first row the chymotrypsinogen numeration of thrombin is indicated; in the second row the numeration in which the first residue of the A-chain is designated as 1 is indicated.; in the third row the numeration in which the first residue of the B-chain is designated as 1 is indicated; finally, the last row show the numeration of prothrombin.

Appendix C

Natural Mutations Occurring in Thrombin Serine Protease Domain

Prothrombin	Amino Acid Substitution	References
Suresnes	Arg299 (7) → Val	Francois et al., 2006
Denver I	Glu300 (8) → Lys	Akhavan et al., 2000 Lefkowitz et al., 2000
n.d.	des301/302 (9/10) LysK	Akhavan et al., 2000 De Cristofaro et al., 2004 De Cristofaro et al., 2006
Denver II	Glu309 (14c) → Lys	Akhavan et al., 2000 Lefkowitz et al., 2000
Segovia	Gly319 (14m) → Arg	Akhavan et al., 2000 Akhavan et al., 1999 Collados et al., 1997 Rocha et al., 1986
San Antonio	Arg320 (15) → His	Akhavan et al., 2000 Sun et al., 2000
Clamart	Arg320 (15) → Ile	Huisse et al., 1986
n.d.	Gly330 (25) → Ser	Akhavan et al., 2000 Akhavan et al., 2005
Himi I	Met337 (32) → Thr	Akhavan et al., 2000 Morishita et al., 1992 Morishita et al., 1991
n.d.	Arg340 (35) → Trp	Akhavan et al., 2000 Tamary et al., 1997
n.d.	Ser354 (48) → Arg	Akhavan et al., 2000
n.d.	Trp357 (51) → Cys	
Vellore 1	Ala362 (56) → Thr	Jayandharan et al., 2005

n.d.	Arg(67) → His	Akhavan et al., 2000 Akhavan et al., 2002
Quick 1 Corpus Christi	Arg382 (67) → Cys	Akhavan et al., 2000 Henriksen and Mann, 1988 O'Marcaigh et al., 1996
Himi II	Arg388 (73) → His	Morishita et al., 1992
Molise 1 Tokushima	Arg418 (101) → Trp	Akhavan et al.,2000 James et al. 1995
Puerto Rico I	Arg457 (137) → Gln	Lefkowitz et al., 2003
Frankfurt Salakta	Glu466 (146) → Ala	Degen et al., 1995 Di Cera et al., 2007 Bezeaud et al., 1984, 1988 Dang et al., 1997
Greenville	Arg517 (187) → Gln	Henriksen et al., 1998 Di Cera et al., 2007 Akhavan et al., 2000
Copenhagen	Ala520 (190) → Val	Stanchev et al., 2006 Di Cera et al., 2007
n.d.	Arg538 (206) → Cys	Akhavan et al., 2000
Perijà	Gly548 (216) → Ala	Sekine et al., 2002
Saint-Denis	Asp552 (221) → Glu	Rouy et al., 2006 Di Cera et al., 2007
Scranton	Lys556 (224) → Thr	Sun et al., 2001 Di Cera et al., 2007
Quick II	Gly558 (226) → Val	Akhavan et al., 2000 Henriksen and Mann, 1989

Numbering for thrombin residues is based on prothrombin sequence with corresponding chymotrypsinogen numbering in parentheses.
n.d. Name not determined

References

Abraham, M.H., Du, C.M., and Platts, J.A. (2000) Lipophilicity of the nitrophenols. *J. Org. Chem.* **65**, 7114-7118.

Akhavan, S, Mannucci, P.M., Lak, M,. Mancuso, G., Mazzucconi, M.G., Rocino, A., Jenkins, P.V., and Perkins, S.J. (2000) Identification and three-dimensional structural analysis of nine novel mutations in patients with prothrombin deficiency. *Thromb Haemost* **84**, 989–97.

Akhavan, S., Miteva, M.A., Villoutreix, B.O., Venisse, L., Peyvandi, F., Mannucci, P..M., Guillin, M.C., and Bezeaud, A. (2005) A critical role for Gly25 in the B chain of human thrombin. *J. Thromb Haemost.* **3**, 139–145.

Albericio, F. (2004) Developments in peptides and amide synthesis. *Curr. Opin. Chem. Biol.* **8**, 211-221.

Anderson, B.L., Boldogh, I., Evangelista, M., Boone, C., Greene, L.A., and Pon, L.A. (1998) The Src homology domain 3 (SH3) of a yeast type I myosin, Myo5p, binds to verprolin and is required for targeting to sites of actin polarization. *J. Cell. Biol.* **141**, 1357-1370.

Apweiler, R., Hermjakob, H., and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta.* **1473**, 4-8.

Arni, R.K., Padmanabhan, K., Padmanabhan, K.P., Wu, T.P., and Tulinsky, A. (1993) Structures of the noncovalent complexes of human and bovine prothrombin fragment 2 with human PPACK-thrombin. *Biochemistry*. **32**, 4727-4737.

Ayala, Y., and Di Cera, E. (1994) Molecular Recognition by thrombin. Role of the slow → fast transition, site-specific ion binding energetics and thermodynamic mapping of structural components. *J. Mol. Biol*. **235**, 733-746.

Ayala, Y.M., Cantwell, A.M., Rose, T., Bush, L.A., Arosio, D., and Di Cera, E. (2001) Molecular mapping of thrombin–receptor interactions. *Proteins* **45**, 107–116.

Barlos, K., Chatzi, O., Gatos, D., and Stavropoulos, G. (1991) 2-Chlorotrityl chloride resin. Studies on anchoring of Fmoc-amino acids and peptide cleavage. *Int. J. Pept. Protein Res.* **37,** 513-520.

Bell., R., Stevens, W.K., Jia, Z., Samis, J., Coté, H.C.F., MacGillivray, R.T.A., and Nesheim, M.E. (2000) Fluorescence properties and functional roles of tryptophan residues 60d, 96, 148, and 215 of thrombin. *J. Biol. Chem.* **275,** 29513-29520.

Berlman, I.B. (1973) Empirical study of heavy-atom collisional quenching of the fluorescence state of aromatic compounds in solution. *J. Phys. Chem.* **77,** 562-567.

Bode W. (2006) The structure of thrombin: A janus-headed proteinase. Semin. Thromb. *Haemost*. **32**, 16-31.

Bode, W. Maryr, I., and Baumann, U., (1989) The refined 1.9 Å crystal structure of human a-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr Pro Pro Trp insertion segment. *EMBO J.* **8**, 3467-3475.

Bode, W., Turk, D., and Karshikov, A. (1992) The refined 1.9-Å X-ray crystal structure of D-Phe-Pro-Arg-chloromethylketone-inhibited human α-thrombin: Structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure-function relationships. *Protein Sci*. **1**, 426-471.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.

Brahms, S., and Brahms, J. (1980) Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J. Mol. Biol.* **138**, 149-178.

Butkowski, R.J., Elio, J., Downing, M.R., and Mann, K.G. (1977) Primary structure of human prothrombin-2 and alpha-thrombin. *J. Biol. Chem.* **252**, 4942-4957.

Carpino, L.A. (1993) 1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive. *J. Am. Chem. Soc.* **115**, 4397-4398.

Carter, W.J., Cama, E., Huntington, J.A. (2005) Crystal structure of thrombin bound to heparin. *J. Biol. Chem.* **280**, 2745-2749.

Cesareni, G., Panni, S., Nardelli, G., and Castagnoli, L. (2002) Can we infer peptide recognition specificity mediated by SH3 domains? *FEBS Lett.* **513**, 38-44.

Cestra, G., Castagnoli, L., Dente, L., Minenkova, O., Petrelli, A., Mingine, N., Hoffmuller, U., Schneider-Mergener, J., and Cesareni, G. (1999) The SH3 domain of endophilin and amphiphysin bind to the proline rich region of synaptojanin at distinct sites that display an unconventional binding specificity. *J. Biol. Chem.* **274**, 32001-32007.

Cleland JL, Craik CS. (1996). *Protein Engineering. Principles and Practice*. Wiley-Liss, New York.

Cohen, B.E., McAnaney, T.B., Park, E.S., Jan, Y.N., Boxer, S.G., and Jan, L.Y. (2002) Probing protein electrostatics with a synthetic fluorescent amino acid. *Science **296**,* 1700-1703.

Colwell, N.S., Blinder, M.A., Tsiang, M., Gibbs, C.S., Bock, P.E., Tollefsen D.M. (1998) Allosteric effects of a monoclonal antibody against thrombin exosite II. *Biochemistry.* **37**, 15057-65.

Copeland, R.A. (2000). *Enzymes: a Practical Introduction to Structure, Mechanism, and Data Analysis.* 2nd ed., J. Wiley & Sons, New York.

Corbett, R.J., and Roche, R.S. (1984) Use of high-speed size-exclusion chromatography for the study of protein folding and stability. *Biochem.* **23**, 7299-7307.

Cornish, V.W., Benson, D.R., Altenbach, C.A., Hideg, K., Hubbell, W.L., and Schultz, P.G. (1994) Site-specific incorporation of biophysical probes into proteins. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2910-2924.

Craik, C.S., Roczniak, S., Largman, C., Rutter, W.J. (1987). The catalytic role of the active site aspartic acid in serine protease. *Science.* **237**, 909- 913.

Cropp, A.T., and Schultz, P.G. (2004) An expanding genetic code. *Trends Genet.* **20**, 625-630.

Csizmadia, F., Tsantili-Kkoulidou, A., Panderi, I., and Darvas, F. (1997) Prediction of distribution coefficient from structure. 1. Estimation method. *J. Pharm. Sci.* **7**, 865-871.

d'Audigier, C., Pasmant, E., Bournier, O., Laurian, Y., Guillin., MC., and Bezeaud. A. (2008) A natural variant with a point mutation resulting in a homozygous Arg to His substitution at position 388 in prothrombin. *Haematologica* **93**, 799-800.

Dalgarno, D.C., Botfield, M.C., and Rickles, R.J. (1997) SH3 domains and drug design: ligands, structure, and biological function. *Biopolymers* **43**, 383-400.

Dang, Q.D., Vindigni, A., and Di Cera, E. (1995) An allosteric switch controls the procoagulant and anticoagulant activities of thrombin. *Proc. Natl. Acad. Sci*. USA **92**, 5977-5981.

Davie, E. W., Fujikawa, K., and Kisiel, W. 1991. The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry* **30,** 10363-10370.

Davie, E.W., and Kulman, J.D. (2006) An overview of the structure and function of thrombin. *Sem. Thromb. Haemost.* **32**, Suppl. 1, 3-15.

Dawson, P., and Kent, S.B. (2000) Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **69**, 923-960.

De Candia, E., De Cristofaro, R., De Marco, L., Mazzucato, M., Picozzi, M., and Landolfi, R. (1997) Thrombin interaction with platelet GpIB: role of the heparin binding domain. *Thromb Haemost*. **77**, 735-740.

De Candia, E., Hall, S.W., Rutella, S., Landolfi, R., Andrews, R.K., and De Cristofaro, R. (2001) Binding of thrombin to glycoprotein Ib accelerates the hydrolysis of PAR-1 on intact platelets. *J. Biol. Chem.* **276**, 4692-4698.

De Cristofaro, R., Akhavan, S., Altomare, C., Carotti, A., Peyvandi, F., and Mannucci, P.M. (2004) A natural prothrombin mutant reveals an unexpected influence of A-chain structure on the activity of human alpha-thrombin. *J. Biol Chem.* **279**, 13035-13043.

De Cristofaro, R., Akhavan, S., Altomare, C., Carotti, A., Peyvandi, F., and Mannucci, P.M. (2004) A natural prothrombin mutant reveals an unexpected influence of A-chain structure on the activity of human alpha-thrombin. *J. Biol Chem.* **279**, 13035-13043.

De Cristofaro, R., and Di Cera, E. (1990) Effect of protons on the amidase activity of human alpha-thrombin. Analysis in terms of a general linkage scheme. *J. Mol. Biol.* **216**, 1077-1085.

De Cristofaro, R., De Candia, E., Landolfi, R., Rutella, S., and Scott W. Hall, S.W. (2001) Structural and functional mapping of the thrombin domain involved in the binding to the platelet glycoprotein Ib. *Biochemistry*. **40**, 13268-13273.

De Filippis, V., Colombo, G., Russo, I., Spadari, B., and Fontana, A. (2002) Probing hirudin-thrombin interaction by incorporation of noncoded amino acids and molecular dynamics simulation. *Biochemistry.* **43**, 1537-1550.

De Filippis, V., De Boni, S., De Dea, E., Dalzoppo, D., Grandi, C., and Fontana, A. (2004) Incorporation of the fluorescent amino acid 7-azatryptophan into the core domain 1-47 of hirudin as a probe of hirudin folding and thrombin recognition. *Protein Sci.* **13**, 1489-1502.

De Filippis, V., De Dea, E., Lucatello, F., and Frasson, R. (2005) Effect of Na^+ binding on the conformation, stability, and molecular recognition properties of thrombin. *Biochem. J.* **390,** 485-492.

De Filippis, V., Frasson, R., and Fontana, A. (2006) 3-Nitrotyrosine as a Spectroscopic Probe for Investigating Protein-Protein Interactions. *Protein Sci.* **15**, 976-986.

De Filippis, V., Quarzago, D., Vindigni, A., Di Cera, E., and Fontana, A. (1998) Synthesis and characterization of more potent analogues of hirudin fragment 1-47 containing non-natural amino acids. *Biochemistry.* **37**, 13507-13515.

De Filippis, V., Russo, I., Vindigni, A., Di Cera, E., Salmaso, S., and Fontana, A. (1999) Incorporation of noncoded amino acids into the N-terminal domain 1-47 of hirudin yields a highly potent and selective thrombin inhibitor. *Protein Sci.* **8**, 2213-2217.

De Filippis, V., Vindigni, A., Altichieri, L., and Fontana A. (1995) Core domain of hirudin from leech. *Hirudinaria manillensis.* Chemical synthesis, purification and characterization of a Trp3-analogue of fragment 1-47. *Biochemistry.* **34**, 9552-9564.

Degen, S.J., and Davie, E.W. (1987) Nucleotide sequence of the gene for human prothrombin. *Biochemistry.* **26**, 6165-6177.

Di Bella, E.E., Maurer, M.C., and Scheraga, H.A. (1995) Expression and folding of recombinant bovine prethrombin-2 and its activation to thrombin. *J. Biol. Chem.* **270**, 163-169.

Di Cera, E., Dang, Q.D., and Ayala, Y.M. (1997) Molecular mechanisms of thrombin function. *Cell Mol. Life Sci.* **53**, 701-730.

Di Nardo, A.A., Larson, S.M., and Davidson, A.R. (2003) The relationship between conservation, thermodynamic stability, and function in the SH3 domain hydrophobic core. *J. Mol. Biol.* **333**, 641-655.

Di Stasio, E., Bizzarri, P., Misiti, F., Pavoni, E., and Brancaccio, A. (2004) A fast, accurate procedure to collect and analyse unfolding fluorescence signal: the case of dystroglycan. *Biophys. Chem.* **107**, 197-211.

Dodt, J., Seemüller, U., Maschler, R., and Fritz, H. (1985) The complete covalent structure of hirudin. Localization of the disulfide bonds. *Biol Chem*. **366**, 379-385.

Dominguez, C., Boelens, R., and Bonvin, A.M. (2003) HADDOCK: A protein–protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731-1737.

Dos Remedios, C.G., and Moens, P.D.J. (1995) Fluorescence resonance energy transfer spectroscopy is a reliable "ruler" for measuring structural changes in proteins. Dispelling the problem of the unknown orientation factor. *J. Struct. Biol.* **115**, 175-185.

Dougherty, D.A. (2000) Un natural amino acids as probes of protein structure and function. *Curr. Opin. Chem. Biol.* **4**, 645-652.

Downing, M.R., Butkowski, R.J., Clark, M.M., and Mann, K.G. (1975) Human prothrombin activation. *J. Biol. Chem.* **250**, 8897-8906.

Dreeskamp, H., Koch, E., and Zander, M. (1975) On the fluorescence quenching of polycyclic aromatic hydrocarbons by nitromethane. *Z. Naturforsch.* **30a**, 1311-1314.

Eftink, M.R. (1997) Fluorescence methods for studying equilibrium macromolecule-ligand interactions. *Methods Enzymol*. **278**, 221-257.

Eftink, M.R. and Ghiron, C. (1981) Fluorescence quenching studies with proteins. *Anal. Biochem.* **114**, 199-227.

England, P.M. (2004) Un natural amino acid mutagenesis: a precise tool for probing protein structure and function. *Biochemistry* **43**, 11623-11629.

Esmon, C.T. (1989) The roles of protein C and thrombomodulin in the regulation in the regulation of blood coagulation. *J. Biol. Chem*. **264**, 4743-4746.

Esmon, C.T., and Jackson, C.M. (1974) The conversion of prothrombin to thrombin. III. The factor Xa-catalyzed activation of prothrombin. *J. Biol. Chem.* **249**, 7782-7790.

Esmon, C.T., and Lollar, P. (1996) Involvement of thrombin anionbinding exosites 1 and 2 in the activation of factor V and factor VIII. *J Biol Chem* **271**, 13882–13887.

Fenton, J.W. II, Fasco, M., and Stackrow, A.B. (1977) Human thrombins: production, evaluation, and properties of α-thrombin. *J. Biol. Chem.* **252**, 3587-3598.

Fernandez-Ballester, G., Blanes-Mira, C., and Serrano, L. (2004) The tryptophan switch: changing ligand-binding specificity from type-I to type-II in SH3 domains. *J. Mol. Biol.* **335**, 619-629.

Fersht, A., and Winter, G. (1992) Protein engineering. *Trends Biochem. Sci.* **17**, 292-295.

Folkers, P.J.M., Clore, G.M., Driscoll, P.C., Dodt, J., Kohler, S., and Gronenborn, A.M. (1989) Solution structure of recombinant hirudin and the Lys-47→Glu mutant: A nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing study. *Biochemistry* **28**, 2601-2617.

Freedman, R.B. (1971) Applications of the chemical reactions of proteins in studies of their structure and function. *Q. Rev. Chem. Soc.* **25**, 431-462.

Fuentes-Prior, P. (2000) Crystal structure of the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. *EMBO J.* **19**, 5650-5660.

Gan, Z.R., Lewis, S.D., Stone, J.R., and Shafer, J.A. (1991) Reconstitution of catalytically competent human ζ-thrombin by combination of ζ-thrombin residues A1-36 and B1-148 and an Escherichia coli expressed polypeptide corresponding to ζ-thrombin residues B149-259. *Biochemistry.* **30**, 11694-11699.

Gan, Z.R., Li., Y., Chen, Z., Lewis, S.D., and Shafer, J.A. (1994) Identification of basic amino acid residues in thrombin essential for heparin-catalyzed inactivation by antithrombin III. *J. Biol. Chem.* **269**, 1301-1305.

Gill, S.G. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182,** 319-326.

Greenspan, N.S. and Di Cera, E. (1999) Defining epitopes: It's not as easy as it seems. *Nat. Biotechnol.* **17**, 936-937.

Hageman, T.C., Endres, G.F., and Scheraga, H.A. (1975) Mechanism of action of thrombin on fibrinogen. On the role of the A chain of bovine thrombin in specificity and in differentiating between thrombin and trypsin. *Arch. Biochem. Biophys.* **17**, 1327-1336.

Hall, S.W., Nagashima, M., Zhao, L., Morser, J., and Leung, L.K. (1999) Thrombin interacts with thrombomodulin, protein C, and thrombin-activable fibrinolysis inhibitor via specific and distinct domains. *J. Biol. Chem.* **274**, 25510-25516.

Halliwell, B. (1997) What nitrates tyrosine? Is nitrotyrosine specific as a biomarker of peroxynitrite formation *in vivo*? *FEBS Lett.* **411**, 157-160.

Haruyama. H, and Wüthrich, K. (1989) Conformation of recombinant desulfatohirudin in aqueous solution determined by nuclear magnetic resonance. *Biochemistry.* **28**, 4301-4312.

Hendrickson, T.L., de Crécy-Lagard, V., and Schimmel, P. (2004) Incorporation of nonnatural amino acids into proteins. *Annu. Rev. Biochem.* **73**, 147-176.

Hofsteenge, J, Braun, P.J., and Stone, S.R. (1988) Enzymatic properties of proteolytic derivatives of human α-thrombin. *Biochemistry* **27**, 2144–2151.

Hogan, J.G. Jr. (1997) Combinatorial chemistry in drug discovery. *Nature Biotech.* **15**, 328-330.

Horne, M.K., and Gralnick, H.R. (1983) The oligosaccharide of human thrombin: investigations of functional significance. *Blood* **1**, 188-194.

Hovius, R., Vallotton, P., Wohland, T., and Vogel, H. (2000) Fluorescence techniques: shedding light on ligand-receptor interactions. *Trends Pharmacol. Sci.* **21**, 266-273.

Hutchinson III, C.A., Phillips, S., Edgell, M.H., Gillam, S., Jahnke, P., ans Smith, M. (1978) Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **253**, 6551-6560.

Huntington, J.A. (2008) How Na^+ activates thrombin – A review of the functional and structural data. *Biol. Chem.* **389**, 1025-1035.

Huntington, J.A., and Esmon, C. (2003) The molecular basis of thrombin allostery revealed by a 1.8-Å structure of the "slow" form. *Structure* **11**, 469-479.

James, H.L., Kim, D.J., Zheng, D.Q., and Girolami, A. (1995) Prothrombin Molise I: documentation of a second incidence of replacement of a critical Arg near the active site. *Thromb Res.* **80**, 363-366.

Jameson, D.M., Croney, J.C., and Moens, P.D. (2003) Fluorescence: basic concepts, practical aspects, and some anecdotes. *Methods Enzymol.* **360**, 1-43.

Jia, C.Y.H., Nie, J., Wu, C., Li, C., and Li, S.S. (2005) Novel Src homology 3 domain-binding motifs identified from proteomic screen of a Pro-rich region. *Mol. Cell. Proteomics* **4**, 1155-1166.

Johnson, D.J.D., Adams, T.E., Li, W., and Huntington, J.A. (2005) Crystal structure of wild-type human thrombin in the Na^+-free state. *Biochem*. J. **392**, 21-28.

Jones, S., and Thornton, J.M. (1996) Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13-20.

Juminaga, D., Albaugh, S.A., and Steiner, R.F. (1994) The interaction of calmodulin with regulatory peptides of phosphorylase kinase. *J. Biol. Chem.* **269**, 1660-1667.

Kang, H., Freund, C., Duke-Cohan, J.S., Musacchio, A., Wagner, G., and Rudd, C.E. (2000) SH3 domain recognition of a proline-independent tyrosine-based RkxxYxxY motif in immune cell adaptor SKAP55. *EMBO J.* **19**, 2889-2899.

Karshikov, A., Bode, W., Tulinsky, A., and Stone, S.R. (1992) Electrostatic interactions in the association of proteins: An analysis of the thrombin-hirudin complex. *Protein Sci.* **1**, 727-735.

Kay, B., Williamson, M.P., and Sudol, M. (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14**, 231-241.

Kent, S.B. (1988) Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.* **57***,* 957-989.

Krem, M.M., and Di Cera, E. (2003) Dissecting substrate recognition by thrombin using the inactive mutant S195A. *Biophys Chem.* **100**, 315-323.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lakowicz, J.R. (1999) *Principles of Fluorescence Spectroscopy* 2nd ed., Kluwer Academic/Plenum, New York.

Lancellotti, S., Rutella, S., De Filippis, V., Pozzi, N., and Rocca, B. (2008) Fibrinogen-elongated γ chain inhibits thrombin-induced platelet response, hindering the interaction with different receptors. *J. Biol. Chem.* **283**, 30193-30204.

Larson, S.M., and Davidson, A.R. (2000) The identification of conserved interactions within the SH3 domain by alignment of sequences and structures. *Protein Sci.* **9**, 2170-2180.

Le Bonniec, B.F., and Esmon, C.T. (1991) Glu-192 → Gln substitution in thrombin mimics the catalytic switch induced by thrombomodulin. *Proc. Natl. Acad. Sci.* **88**, 7371-7375.

Li, S.S. (2005) Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem. J.* **390**, 641-653.

Liu, L.W., Rezaie, A.R., Carson, C.W., Esmon. N.L., and Esmon, C.T. (1994) Occupancy of anion binding exosite 2 on thrombin determines Ca^{2+} dependence of protein C activation. *J. Biol. Chem.* **269**, 11807-11812.

Liu, L.W., Vu, T.K., Esmon, C.T., and Coughlin, S.R. (1991) The region of thrombin receptor resembling hirudin binds to thrombin and alters enzyme specificity. *J. Biol. Chem.* **266**, 16977–16980.

Lo Conte, L., Chothia, C., and Janin, J. (1999) The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177-2198.

Lottenberg, R., and Jackson, C.M. (1983) Solution composition dependent variation in extinction coefficients for *p*-nitroaniline. Biochim. Biophys. Acta. **742**, 558–564.

MacArthur, M.W., and Thronton, J.M. (1991) Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**, 397-412.

Marcus, Y. (1994) A simple empirical model describing the thermodynamics of hydration of ions widely varying charges, sizes, and shapes. *Biophys. Chem.* **51**, 111-127.

Markwardt, F., (1994) The development of hirudin as an antithrombotic drug. *Thromb. Res.* **74**, 1-23.

Mathews, I.I., Padmanabhan, K.P., Ganesh, V., Tulinsky, A., Ishii, M., Chen, J., Turck, C.W., Coughlin, S.R., Fenton, J.W. and. (1994) Crystallographic structures of thrombin complexed with thrombin receptor peptides: existence of expected and novel binding modes. *Biochemistry.* **33**, 3266-3279.

Mattos, C. (2002) Protein-water interactions in a dynamic world. *Trends Biochem. Sci.* **27**, 203-208.

Mayer, B.J. (2001) SH3 domains: complexity in moderation. *J. Cell. Sci.* **114**, 1253-1263.

Meloun, B., Frič, I., and Šorm, F. (1968) Nitration of tyrosine residues in the pancreatic trypsin inhibitor with tetranitromethane. *Eur. J. Biochem.* **4**, 112-117.

Mezo, A.R., Cheng, R.P., and Imperiali, B. (2001) Oligomerization of uniquely folded mini-protein motifs: development of a homotrimeric ββα peptide. *J. Am. Chem. Soc.* **123**, 3885-3891.

Mitra, N., Sinha, S., Ramya, T.N.C., and Surolia, A. (2006) N-linked oligosaccharides as outfitters for glycoprotein folding, form and functions. *Trends Biochem. Sci.* **31**, 156-163.

Mongiovì, A.M., Romano, P.R., Panni, S., Mendoza, M., Wong, W.T., Musacchio, A., Cesareni, G., and Di Fiore, P.P. (1999) A novel peptide-SH3 interaction. *EMBO J.* **18**. 5300-5309.

Morita, T., and Iwanaga, S. (1981) Prothrombin activator from Echis carinatus venom. *Methods Enzymol.* **80**, 303-311.

Morton, C.J. and Campbell, I.D. (1994) SH3 domains: molecular 'velcro'. *Curr. Biol.* **4**, 615-617.

Mosesson, M.W. (2005) Fibrinogen and fibrin structure and functions. *J. Thromb. Haemost.* **3**, 1894-1904.

Mostad, A., and Natarajan, S. (1990) Crystal and molecular structure of 3-nitro-4-hydroxy-phenylalanine nitrate. *Z. Kristall.* **193**, 127-136.

Musacchio, A., Saraste, M., and Wilmanns, M. (1994) High resolution crystal structure of tyrosine kinase SH3 domains complexed to proline-rich peptides. *Nature Struct. Biol.* **1**, 546-551.

Musacchio. A. (2002) How SH3 domains recognize proline. *Adv. Protein Chem.* **61**, 211-268.

Musi, V., Birdsall, B., Fernandez-Ballester, G., Guerrini, R., Salvatori, S., Serrano, L., and Pastore, A. (2006) New approaches to high-throughput structure characterization of SH3 complexes: the example of myosin-3 and myosin-5 from *S. cerevisiae*. *Protein Sci.* **15**, 795-807.

Myles, T., Church, F.C., Whinna, H.C., Monard, D., and Stone, S.R. (1998) Role of thrombin anion-binding exosite-I in the formation of thrombin-serpin complexes. *J. Biol. Chem.* **273**, 31203-31208.

Myles, T., Le Bonniec, B.F., and Stone, S.R. (2001) The dual role of thrombin's anion-binding exosite-I in the recognition and cleavage of the protease-activated receptor 1. *Eur. J. Biochem.* **268**, 70-77.

Naski, M.C., and Shafer, J.A. (1991) A kinetic model for the alpha-thrombin-catalyzed conversion of plasma levels of fibrinogen to fibrin in the presence of antithrombin III. *J Biol Chem.* **266**, 13003-13010.

Naski, M.C., Fenton, J.W. 2nd, Maraganore, J.M., Olson, S.T., Shafer, J.A. (1990) The COOH-terminal domain of hirudin. An exosite-directed competitive inhibitor of the action of alpha-thrombin on fibrinogen. *J. Biol. Chem.* **265**, 13484-13489.

Nicastro, G., Baumer, L., Bolis, G., and Tatò, M. (1997) NMR solution structure of a novel hirudin variant HM2, N-terminal 1-47 and N64 → V+G mutant. *Biopolymers* **41**, 731-749.

Nilsson, B., Horne, M.K. 3rd, and Gralnick, H.R. (1983) The carbohydrate of human thrombin: structural analysis of glycoprotein oligosaccharides by mass spectrometry. *Arch. Biochem. Biophys.* **224**, 127-1233.

Nilsson, B.L., Soellner, M.B., and Raines, R.T. (2005) Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 91-118.

O'Connor, S.E., and Imperiali, B. (1996) Modulation of protein structure and function by asparagine-linked glycosylation. *Chem. and Biol.* **3**, 803-812.

Pace, C.N., Vajdos, F., Fee, L. Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411-2423.

Papaconstantinou, M.E., Bah, A., and Di Cera E. (2008) Role of the A chain in thrombin function. *Cell Mol Life Sci.* **65**, 1943-1947.

Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445-452.

Perutz, M.F. (1970) Stereochemistry of cooperative effects in haemoglobin. Nature. **228**, 726-739.

Peyvandi, F., Duga, S., Akhavan, S., and Mannucci, P.M. (2002) Rare coagulation deficiencies. *Haemophilia.* **8**, 308-321.

Pineda, A.O., Carrell, C.J., Bush, L.A., Prasad, S., Caccia, S., Chen, Z.W., Mathews, F.S., and Di Cera, E. (2004) Molecular dissection of Na^+ binding to thrombin. *J. Biol. Chem.* **279**, 31842-31853.

Pineda, A.O., Savvides, S.N., Waksman, G., and Di Cera, E. (2002) Crystal structure of the anticoagulant slow form. *J. Biol. Chem.* **277**, 40177-40180

Protein Sci. 1997 3 :689-97. Biosynthetic incorporation of tryptophan analogues into staphylococcal nuclease: effect of 5-hydroxytryptophan and 7-azatryptophan on structure and stability.Wong CY, Eftink MR.

Puchalski, M.M., Morra, M.J., and von Wandruszka, R. (1991) Assessment of inner filter effect corrections in fluorimetry. *Fresenius J. Anal. Chem.* **340**, 341-344.

Purcell, A.W., Aguilar, M.I., and Hearn, M.T. (1999) Probing the binding behavior and conformational states of globular proteins in reversed-phase high-performance liquid chromatography. *Anal. Chem.* **71**, 2440-2451.

Rath, A., Davidson, A.R., and Deber, C.M. (2005) The structure of "unstructured" regions in peptides and proteins: role of the polyproline II helix in protein folding and recognition. *Biopolymers* **80**, 179-185.

Richardson, J,S., Richardson, D.C. (1990). The *de novo* synthesis of proteins. In *Proteins: Form and Function*. Elsevier Trends Journal. 173-182.

Richardson, J.L., Fuentes-Prior, P., Sadler, J.E., Huber, R., and Bode, W. (2002) Characterization of the residues involved in the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. *Biochemistry.* **41***,* 2535-2542.

Riordan, J.F., Sokolovsky, M., and Vallee, B.L. (1967) Environmentally sensitive tyrosyl residues. Nitration with tetranitromethane. *Biochemistry* **6**, 358-361.

Rischel, C. and Poulsen, F.M. (1995) Modification of a specific tyrosine enables tracing of the end-to-end distance during apomyoglobin folding. *FEBS Lett.* **374**, 105-109.

Rischel, C., Thyberg, P., Rigler, R., and Poulsen, F.M. (1996) Time-resolved fluorescence studies of the molten globule state of apomyoglobin. *J. Mol. Biol.* **257**, 877-885.

Rodal, A.A., Manning, A.L., Goode, B.L., and Drubin, D.G. (2003) Negative regulation of yeast WASP by two SH3 domain-containing proteins. *Curr. Biol.* **13**, 1000-1008.

Rothe, M., and Mazànek, J. (1972) Side-reactions arising on formation of cyclodipeptides in solid-phase peptide synthesis. *Angew. Chem. Int. Ed.* **11**, 293.

Russo, G., Gast, A., Schlaeger, E., Angiolillo, A., and Pietropaolo, C. (1997) Stable expression and purification of a secreted human recombinant prethrombin-2 and its activation to thrombin. *Prot. Expr. Purif.* **10**, 214-225.

Rydel, T.J., Tulinski, A., Bode, W., and Huber, R. (1991) Refined structure of the hirudin-thrombin complex. *J. Mol. Biol.* **221**, 583-601.

Rydel, T.J., Yin, M., and Padmanabhan, K.P. (1994) Crystallographic structure of human gamma-thrombin. *J. Biol. Chem.* **269**, 22000–22006.

Sadasivan, C., Yee, V.C. (2000) Interaction of the factor XIII activation peptide with alpha-thrombin. Crystal structure of its enzyme-substrate analog complex. *J. Biol. Chem.* **275**, 36942–36948.

Sawicki, E., Stanley, T.W., and Elbert, W.C. (1964) Quenchofluorometric analysis for fluoranthenic hydrocarbons in the presence of other types of aromatic hydrocarbons. *Talanta* **11**, 1433-1441.

Scacheri, E., Nitti, G., Valsasina, B., Orsini, G., Visco, C., Ferreia, M., Sawyer, R.T., and Sarmientos, P. (1993) Novel hirudin variants from the leech. *Hirudinaria manillensis*. Amino acid sequence, cDNA cloning and genomic organization. *Eur J Biochem.* **214**, 295-304.

Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain, *Biochem. Biophys. Res. Commun.* **27**, 157–162.

Searle, A.S., and Williams, D.H. (1992) The cost of conformational order: entropy changes in molecular association. *J. Amer. Chem. Soc.* **114**, 10690-10697.

Selvin, P.R. (1995) Fluorescence resonance energy transfer. *Methods Enzymol.* **246**, 300-334.

Sheehan, J. P., and Sadler, J. E. (1994) Proc. Natl. Acad. Sci. U. S. A. **91**, 5518–5522.

So, I.S., Lee, S., Kim, S.W., Hahm, K., and Kim, J. (1992) Purification and activation of recombinant human prethrombin 2 produced in E. coli. *Korean Biochem. J.* **25**, 60–65.

Soejima, K., Mimura, N., Yonemura, H., Nakatake, H., Imamura, T., and Nozaki, C. (2001) An efficient refolding method for the preparation of recombinant human prethrombin-2 and characterization of the recombinant derived α-thrombin. *J.Biochem.* **130**, 269-277.

Steen, M., and Dahlback, B. (2002) Thrombin-mediated proteolysis of factor V resulting in gradual B-domain release and exposure of the factor Xa-binding site. *J. Biol. Chem.* **277**, 38424–38430.

Steiner, R.F., Albaugh, S., and Kilhoffer, M.C. (1991) Distribution separations between groups in an engineered calmodulin. *J. Fluoresc.* **1**, 15-22.

Steiner, V., Knecht, R., Börsen, K.O., Gasssmann, E., Stone, S.R., Raschdorf, F., Schaleppi, J.-M., and Maschler, R. (1992) Primary structure and function of novel *O*-glycosylated hirudins from the leech hirudinaria manillensis. *Biochemistry* **31**, 2294-2298.

Stevens, W.K., and Nesheim,M.E. (1993) Structural changes in the protease domain of prothrombin upon activation as assessed by N-bromosuccinimmide modification of tryptophan residues in prothrombin-2 and thrombin. *Biochemistry.* **32**, 2787-2794.

Stoll, V.S., and Blanchard, J. (1990) Buffers: principles and practice. *Methods Enzymol.* **182**, 24-38.

Strickland, E. H. (1974) Aromatic contributions to circular dichroism spectra of proteins. *CRC Crit. Rev. Biochem.* **3**, 113-175.

Stringer, K.A., and Lindenfeld, J. (1992) Hirudins: antithrombin anticoagulants. *Ann Pharmacother*. **12**, 1535-1540.

Stubbs, M.T., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S.R., and Bode, W. (1992) The interaction of thrombin with fibrinogen: A structural basis for its specificity. *Eur. J. Biochem.* **206**, 187-195.

Szyperski, T., Güntert, P., Stone, S.R., and Wütrich, K. (1992) Nuclear magnetic resonance solution structure of hirudin (1-51) and comparison with corresponding three-dimensional structures determined using the complete 65-residue hirudin polypeptide chain. *J. Mol. Biol.* **228**, 1193-1205.

Tcherkasskaya, O., and Ptitsyn, O.B. (1999) Direct energy transfer to study the 3D structure of non-native proteins: AGH complex in the molten globule state of apomyoglobin. *Protein Eng.* **12**, 485-490.

Tong, A.H., Drees, B., Nardelli, G., Bader, G.D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., Quondam, M., Zucconi, A., Hogue, C.W., Fields, S., Boone, C., and Cesareni, G. (2002) A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science* **295**, 321-324.

Tsiang, M., Jain, A.K., and Gibbs, C.S. (1997) Functional requirements for inhibition of thrombin by

Tsiang, M., Jain, A.K., Dunn, K.E., Rojas, M.E., Leung, L.L., and Gibbs, C.S. (1995) Functional mapping of the surface residues of human thrombin. *J Biol Chem.* **270**, 16854-16863.

Tsumoto, K., Ejima, D., Kumagai, I., and Arakawa, T. (2003) Practical considerations in refolding proteins from inclusion bodies. *Prot. Expr. Pur.* **28**, 1-8.

Twine, S.M., and Szabo, A.G. (2003) Fluorescent amino acid analogs. *Methods Enzymol.* **360**, 104-127.

Van Deerlin, V.M., and Tollefsen, D.M. (1991) The N-terminal acidic domain of heparin cofactor II mediates the inhibition of alpha-thrombin in the presence of glycosaminoglycans. *J. Biol. Chem.* **266**, 20223-20231.

Verhamme, I.M., Olson, S.T., Tollefsen, D.M., and Bock, P.E. (2002) Binding of exosite ligands to human thrombin. Re-evaluation of allosteric linkage between thrombin exosites I and II. *J. Biol. Chem.* **277**, 6788-6798.

Vidal, M., Gigoux, V., and Garbay, C. (2001) SH2 and SH3 domains as targets for anti-proliferative agents. *Crit. Rev. Oncol. Hematol.* **40**, 175-186.

Vidal, M., Goudreau, N., Cornille, F., Cussac, D., Gincel, E., and Garbay, C. (1999) Molecular and cellular analysis of Grb2 SH3 domain mutants: interaction with Sos and dynamin. *J. Mol. Biol.* **290**, 717-730.

Vijayalakshmi, J., Padmanabhan, K.P., Mann, K.G., and Tulinsky, A. (1994) The isomorphous structures of prethrombin-2, hirugen-, and PPACK-thrombin: Changes accompanying activation and exosite binding to thrombin. *Protein Sci.* **3**, 2254-2271.

Vindigni, A., Dang, Q.D., and Di Cera, E. (1997) Site-specific dissection of substrate recognition by thrombin. *Nat. Biotechnol.* **31**, 11721-11730.

Vindigni, A., De Filippis, V., Zanotti, G., Visco, C., Orsini, G., and Fontana, A. (1994) Probing the structure of hirudin from *Hirudinaria manillensis* by limited proteolysis: Isolation, characterization and thrombin-inhibitory properties of N-terminal fragments. *Eur. J. Biochem.* **226**, 323-333.

Vu, T.K., Hung, D.T., Wheaton, V.I., and Coughlin, S.R. (1991) Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell*. **64**, 1057-1068.

Wallace, C.J. (1993). Understanding cytochrome c function: engineering protein structure by semisynthesis. *FASEB J.* **7**, 505-515.

Wang, C., Eufemi, M., Turano, C., and Giartosio, A. (1996) Influence of the carbohydrate moiety on the stability of glycoproteins. **11**, 7299-7307.

Wells, C.M. and Di Cera, E. (1992) Thrombin is a Na^+-activated enzyme. *Biochemistry* **31**, 11721-11730.

Wells, J.A. (1990) Additivity of mutational effects in proteins. *Biochemistry* **29**, 8509-8517.

Wong, C.Y., and Eftink, M.R. (1997) Biosynthetic incorporation of tryptophan analogues into staphylococcal nuclease: effect of 5-hydroxytryptophan and 7-azatryptophan on structure and stability. *Protein Sci.* **3**, 689-697.

Workman, E.F., and Lundblad, R.L. (1978) The effect of monovalent cations on the catalytic activity of thrombin. *Arch. Biochem. Biophys*. **185**, 544-548.

Wu, P. and Brand, L. (1994) Resonance energy transfer: methods and applications. *Anal. Biochem.* **18**. 1-13.

Wu, Q.Y., Sheehan, J.P., Tsiang, M., Lentz, S.R., Birktoft, J.J., and Sadler, J.E. (1991) Single amino acid substitutions dissociate fibrinogenclotting and thrombomodulin-binding activities of human thrombin. *Proc Natl Acad Sci USA* **88**, 6775-6779.

Wu, X., Knudsen, B., Feller, S.M., Zheng, J., Sali, A., Cowburn, D., Hanafusa, H., and Kuriyan, J. (1995) Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. *Structure* **3**, 215-226.

Xenarios, I., and Eisenberg, D. (2001) Protein interaction databases. *Curr Opin Biotechnol*. **12**, 334-339. Review.

Yan, Y., and Marriott, G. (2003) Analysis of protein interactions using fluorescence technologies. *Curr. Opin. Chem. Biol.* **7**, 635-640.

Ye, J., Liu, L.W., Esmon, C.T., and Johnson, A.E. (1992) The fifth and sixth growth factor-like domains of thrombomodulin bind to the anion-binding exosite of thrombin and alter its specificity. *J. Biol. Chem.* **267**, 11023-11028.

Zarrinpar, A., and Lim, W.A. (2000) Converging on proline: the mechanism of WW domain peptide recognition. *Nature Struct. Biol.* **7**, 611-613.

Grazie….

Grazie al Prof. Vincenzo De Filippis… dopo quasi 5 anni, stiamo raggiungendo di nuovo, insieme, un altro importante traguardo.

Grazie al Dr. Daniele Dalzoppo, come amico e scienziato… enciclopedia preziosa da cui attingere sempre suggerimenti indispensabili, ma soprattutto uomo dalla bontà, gioia di vivere e generosità rare.

Grazie al Dr. Capo Nicola Pozzi… ancora una volta, semplicemente grazie.

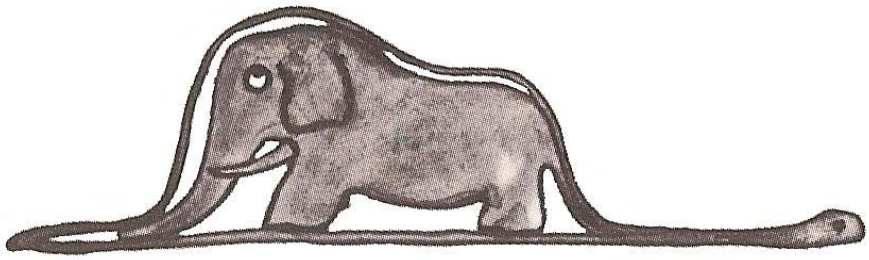
Grazie a chi ha condiviso con me la vita nel "laboratorio più bello del mondo" in questi 3 anni: Daniela Galla, Fabio Maset, Massimo Coatto, Annamaria Draghi, Marco Daniele, Ilaria Rizzi, Davide Zaramella, Lucia Dalvit, Samuele Bettin e Isabella Perissinotto.

Grazie al Prof Jim Huntington, al Dr. Dan Johnson, a Genichi Nakamura e al Dr. Masayuki Yamasaki, per avermi dato la possibilità grazie ai mesi trascorsi nel loro laboratorio di crearmi un mio piccolo filone di ricerca che mi ha ridato entusiasmo in un periodo difficile del mio dottorato.

Grazie alla biotecnologa Elisa Pasqualetto, compagna di sempre, e ai biotecnologi Marco Cavalli, Luca Persano, e Raffaele Niero per le stimulating discussions.

E infine grazie di cuore soprattutto a chi da Casa tifa per me… mamma, papà, e mio marito, Enrico, a cui dedico questo lavoro. Grazie!

Roberta



“Io ti rendo lode, o Padre, Signore del cielo e della terra, perché hai nascosto queste cose ai sapienti e agli intelligenti, e le hai rivelate ai piccoli.”

Mt 11,25