Probing the Hirudin−Thrombin Interaction by Incorporation of Noncoded Amino Acids and Molecular Dynamics Simulation[†,‡]

Vincenzo De Filippis,[*,§] Giorgio Colombo,[||] Ilaria Russo,[§] Barbara Spadari,[§] and Angelo Fontana[§]

Department of Pharmaceutical Sciences and CRIBI Biotechnology Center, University of Padua, via F. Marzolo 5, I-35131 Padua, Italy, CNR Institute for Biocatalysis and Molecular Recognition, via M. Bianco, Milan, Italy, and Department of Biophysical Chemistry, University of Groningen, Nijenborgh, The Netherlands

Received May 9, 2002; Revised Manuscript Received August 15, 2002

ABSTRACT: Thrombin is a primary target for the development of novel anticoagulants, since it plays two important and opposite roles in hemostasis: procoagulant and anticoagulant. All thrombin functions are influenced by Na^+ binding, which triggers the transition of this enzyme from an anticoagulant (slow) form to a procoagulant (fast) form. In previous studies, we have conveniently produced by chemical synthesis analogues of the N-terminal fragment 1−47 of hirudin HM2 containing noncoded amino acids and displaying up to ∼2700-fold more potent antithrombin activity, comparable to that of full-length hirudin. In the work presented here, we have exploited the versatility of chemical synthesis to probe the structural and energetic properties of the S3 site of thrombin through perturbations introduced in the structure of hirudin fragment 1−47. In particular, we have investigated the effects of systematic replacement of Tyr3 with noncoded amino acids retaining the aromatic nucleus of Tyr, as well as similar hydrophobic and steric properties, but possessing different electronic (e.g., *p*-fluoro-, *p*-iodo-, or *p*-nitro-Phe), charge (*p*-aminomethyl-Phe), or conformational (*homo*-Phe) properties. Our results indicate that the affinity of fragment 1−47 for thrombin is proportional to the desolvation free energy change upon complex formation, and is inversely related to the electric dipole moment of the amino acid side chain at position 3 of hirudin. In this study, we have also identified the key features that are responsible for the preferential binding of hirudin to the procoagulant (fast) form of thrombin. Strikingly, shaving at position 3, by Tyr → Ala exchange, abolishes the differences in the affinity for thrombin allosteric forms, whereas a bulkier side chain (e.g., β-naphthylalanine) improves binding preferentially to the fast form. These results provide strong, albeit indirect, evidence that the procoagulant (fast) form of thrombin is in a more open and accessible conformation with respect to the less forgiving structure it acquires in the slow form. This view is also supported by the results of molecular dynamics simulations conducted for 18 ns on free thrombin in full explicit water, showing that after ∼5 ns thrombin undergoes a significant conformational transition, from a more open conformation (which we propose can be related to the fast form) to a more compact and closed one (which we propose can be related to the slow form). This transition mainly involves the Trp148 and Trp60D loop, the S3 site, and the fibrinogen binding site, whereas the S1 site, the Na^+-binding site, and the catalytic pocket remain essentially unchanged. In particular, our data indicate that the S3 site of the enzyme is less accessible to water in the putative slow form. This structural picture provides a reasonable molecular explanation for the fact that physiological substrates related to the procoagulant activity of thrombin (fibrinogen, thrombin receptor 1, and factor XIII) orient a bulky side chain into the S3 site of the enzyme. Taken together, our results can have important implications for the design of novel thrombin inhibitors, of practical utility in the treatment of coagulative disorders.

Thrombin is a serine protease whose function is central to the control of the blood coagulation process (*1*), and for this reason, it remains the primary target for the development of novel anticoagulants (*2*). As often occurs in highly regulated biological systems, thrombin plays two opposite roles in hemostasis: procoagulant and anticoagulant. The procoagulant activity of thrombin involves the conversion of soluble fibrinogen into an insoluble fibrin clot, platelet aggregation via activation of protease-activated receptors (PARs),[1] stabilization of the newly formed fibrin clot through activation of factor XIII, and amplification of its own production through activation of factors V, VIII, and XI. The anticoagulant role involves the thrombomodulin-assisted activation of protein C, which in turn downregulates the conversion of prothrombin into active thrombin, through proteolytic inactivation of factor Va (*3*, *4*). All thrombin functions are influenced by Na^+ binding, which triggers the

[†] This work was supported in part by Grant PRIN-2000 from the Italian Ministry of University and Scientific Research to V.D.F.

[‡] Part of this work was presented at the 4th European Symposium of the Protein Society, Paris, France, April 18−22, 2001.

* To whom correspondence should be addressed. Telephone: +39-49-827-5698. Fax: +39-49-827-5366. E-mail: vincenzo.defilippis@unipd.it.

[§] University of Padua.

[||] CNR Institute for Biocatalysis and Molecular Recognition and University of Groningen.

10.1021/bi0203482 CCC: $22.00
Published on Web 10/22/2002

transition of this enzyme from an anticoagulant (slow) form to a procoagulant (fast) form. The Na^+-bound form cleaves more specifically fibrinogen and PARs, whereas the Na^+-free (slow) form preferentially hydrolyzes protein C (*4*). Under physiological conditions of temperature (37 °C) and Na^+ concentration (145 mM), the two allosteric forms of thrombin are almost equally populated, whereas at 25 °C and 0.2 M NaCl or ChCl, the enzyme exists predominantly (> 90%) in the fast or slow form, respectively (*4*). In this view, molecules capable of affecting the equilibrium between these two forms can have great potential therapeutic impact in the treatment of coagulative disorders (*5*).

Among natural anticoagulants, hirudin is the most potent and specific inhibitor of thrombin, with a dissociation constant in the range of 20−200 fM (*6*, *7*). Structural studies conducted on hirudin in the free (*8*−*10*) and thrombin-bound state (*11*) 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 penetrates into the specificity pockets of the enzyme. Notably, the N-terminal tripeptide makes approximately half of the total contacts observed for the binding of fragment 1−47 to thrombin. 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 (*6*). Although far less active (at least 2×10^5-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 (*12*, *13*), suggesting that the structural determinants for this behavior are specified by the amino acid residues along the sequence of the N-terminal domain. Hence, mutational studies on hirudin fragment 1−47 can be useful not only in investigating the physicochemical determinants that are responsible for the extraordinary affinity and specificity of hirudin for thrombin but also in probing the structural properties of the enzyme in the slow or fast forms.

The advent of recombinant DNA techniques made it possible to site-specifically insert, delete, or mutate almost any amino acid in a given protein, significantly improving our knowledge of protein structure, stability, and function (*14*−*16*). Nevertheless, a quantitative description of the physical and chemical basis that makes a polypeptide chain efficiently fold into a stable and functionally active conformation is still elusive. 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 natural amino acids, thus making it difficult, if not impossible, to unequivocally relate the change in protein stability or function to the variation of physicochemical properties caused by amino acid exchange(s). In this view, incorporation of noncoded amino acids with tailored side chains, allowing fine tuning of the structure at a protein site, would facilitate dissection of the effects of a given mutation in terms of one or a few physicochemical properties, thus much expanding the scope of physical organic chemistry in the study of proteins (*17*). Over the past three decades, incorporation of noncoded amino acids has been widely exploited in the study of bioactive peptides (*18*, *19*). 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 past few years of serendipity-based approaches in peptide design and drug discovery (*20*). 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. In this respect, advancements of peptide synthetic chemistry made possible the incorporation of noncoded amino acids into even long polypeptide chains, allowing researchers to carry out mutational studies on proteins. An advantage of this approach resides in the fact that proteins, contrary to peptides, are generally characterized by a well-defined three-dimensional structure, and thus, they could serve 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 physicochemical properties at the mutation site(s).

In previous studies, we have used this approach to stabilize the C-terminal domain of thermolysin against thermal denaturation (*21*) and to produce analogues of the N-terminal fragment 1−47 of hirudin HM2 displaying more potent antithrombin activity (*12*, *22*). We have also shown that the effects of multiple amino acid replacements are nicely additive, and in the case of fragment 1−47, cumulative amino acid exchanges yielded a highly potent analogue binding to the fast and slow forms of thrombin with affinities 2700 and 6800 times higher than that of the natural product, respectively, and comparable to that of full-length hirudin (*13*). In this work, we have extended our previous results and exploited the versatility of chemical synthesis to probe the structural and energetic properties of the specificity sites of thrombin through perturbations introduced into hirudin structure (Figure 1). In particular, we have investigated the effects of systematic replacement of Tyr3 with noncoded amino acids retaining the aromatic nucleus of Tyr, as well as similar hydrophobic and steric characteristics, but possessing different electronic (e.g., *p*-fluoro-, *p*-iodo-, and *p*-nitro-Phe), charge (*p*-aminomethyl-Phe), or conformational (*homo*-Phe) properties. Our results indicate that the affinity

[1] Abbreviations: ASA, accessible surface area; CD, circular dichroism; ChCl, choline chloride; DIEA, diisopropylethylamine; DMF, *N*,*N*-dimethylformamide; DTT, dithiothreitol; EDT, ethanedithiol; ESI, electrospray ionization; Fmoc, 9-fluorenylmethyloxycarbonyl; Fmoc-ONSu, (9-fluorenylmethyloxycarbonyl)-*O*-*N*-hydroxysuccinimide; HBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HOBt, 1-hydroxybenzotriazole; HPLC, high-pressure liquid chromatography; β-ME, β-mercaptoethanol; PAR, protease-activated receptor; pdb, Protein Data Bank; PEG, polyethylene glycol; RP, reverse-phase; rt, room temperature; TFA, trifluoroacetic acid; UV, ultraviolet; TOF, time-of-flight; HM2 and HV1, hirudin variants isolated from the leeches *H. manillensis* and *H. medicinalis*, respectively. Standard three-letter abbreviations are used for all natural amino acids. Tyr3Ala refers to the synthetic analogue of fragment 1−47 of hirudin HM2 in which Tyr3 has been replaced with Ala, while Tyr3Cha and Tyr3βNal correspond to the analogues obtained by replacing Tyr3 with cyclohexylalanine (Cha) and β-naphthylalanine (βNal), respectively.

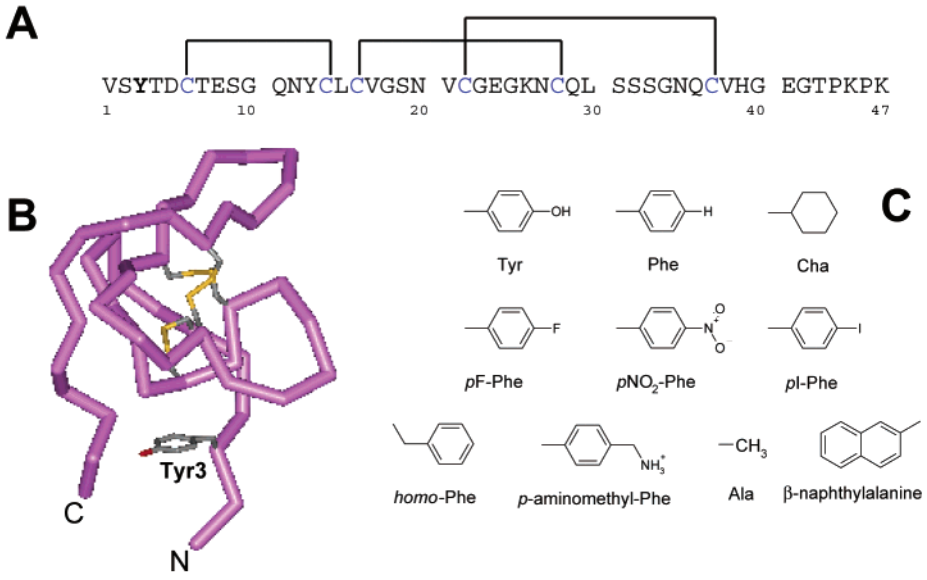

FIGURE 1: Amino acid sequence (A) (*70*) and schematic representation (B) of the solution structure (*10*) of N-terminal fragment 1−47 of hirudin HM2 from *H. manillensis*. Position 3 is indicated in bold, and disulfide bonds are represented with solid lines. The ribbon drawing was generated using the program WebLab ViewerPro 4.0 (Molecular Simulations Inc., 2000). (C) Chemical structure of the amino acid side chains at position 3 of hirudin fragment 1−47.

of fragment 1−47 for thrombin is proportional to desolvation free energy change upon complex formation, and is inversely related to the electric dipole moment of the amino acid side chain at position 3 of hirudin. In this study, we have also attempted to identify the key features that are responsible for the preferential binding of hirudin to the procoagulant (fast) form of thrombin, by introducing coarse variations into hirudin sequence and by molecular dynamics (MD) simulation of the free thrombin molecule in full explicit water solvent. Strikingly, shaving at position 3 (i.e., Tyr → Ala exchange) abolishes the differences in the affinity for thrombin allosteric forms ($\Delta G_c \sim 0$), while a bulkier side chain (e.g., βNal) slightly improves binding of the fast form. These findings are in agreement with those obtained by MD calculations providing a picture of the procoagulant fast form of thrombin being in a more open and accessible conformation with respect to that of the anticoagulant slow form.

EXPERIMENTAL PROCEDURES

Materials

Fragment 1−47 of hirudin from *Hirudinaria manillensis* (variant HM2) was obtained by limited proteolysis with trypsin of hirudin HM2 (*6*, *22*), which was generously provided by the Bioscience Center of Pharmacia-Upjhon (Milan, Italy). Bovine trypsin (EC 3.4.21.4) and human α-thrombin (EC 3.4.21.5) were purchased from Sigma (St. Louis, MO). L-Arginine-*p*-nitroanilide and, when commercially available, Fmoc derivatives of non-natural amino acid were from Bachem AG (Bubendorf, Switzerland) or Advanced Chemtech (Louisville, KY). Other protected amino acids, solvents, and reagents for peptide synthesis, as well as those for peptide and protein sequence analysis, were from Applied Biosystems (Foster City, CA). Trifluoroacetic acid (TFA), formic acid, and β-mercaptoethanol (β-ME) were purchased from Pierce (Rockford, IL); polyethylene glycol (PEG) 8000, ethanedithiol (EDT), and dithiothreitol (DTT) were from Fluka (Basel, Switzerland). All other reagents, buffers, and organic solvents were analytical grade and obtained from Fluka or Merck (Darmstadt, Germany).

Methods

Peptide Synthesis. Fmoc derivatives of *p*I-Phe and *homo*-Phe were prepared using Fmoc-ONSu as an acylating agent, according to the method of Ten Kortenaar et al. (*23*). The analogues of hirudin HM2 fragment 1−47 were synthesized by the solid-phase Fmoc method (Atherton and Sheppard, 1987) in two sequential steps involving automated synthesis of segment 6−47 on a Lys-derivatized (0.48 mmol/g of resin) *p*-alkoxybenzyl ester polystyrene resin (*24*) and elongation of the peptide by manual solid-phase procedure, as previously described (*22*). The crude reduced peptides, obtained after precipitation with ice-cold diethyl ether, were dissolved (2 mg/mL) in 0.1 M $NaHCO_3$ buffer (pH 8.3) and allowed to fold for 24 h under air oxidation conditions in the presence of 100 μM β-mercaptoethanol (*22*, *25*). The refolded disulfide-oxidized peptides were purified by preparative RP-HPLC on a Vydac (The Separations Group, Hesperia, CA) C18 column (1 cm × 25 cm, 10 μm particle size), eluted with a linear acetonitrile/0.05% TFA gradient, and their chemical identity was established by automated N-terminal sequence analysis and high-resolution mass spectrometry on a Mariner ESI-TOF instrument from Perseptive Biosystems, which gave mass values in agreement with the expected amino acid composition within 10 ppm mass accuracy. For some analogues of hirudin fragment 1−47, the correctness of disulfide pairing was established by enzymatic fingerprint analysis, as previously described (*22*).

To obtain experimental estimates of the hydrophobic properties of hirudin analogues mutated at position 3, the retention times (t_R) for all the analogues have been determined by reversed-phase HPLC analysis. Aliquots (20−30 μg) of the synthetic peptides in 5 mM Na_2HPO_4 buffer (pH 6.4) were loaded (100 μL) onto a Vydac C18 (4.6 mm × 150 mm, 5 μm) column equilibrated with the same phosphate buffer and eluted at room temperature with a linear aceto-

nitrile gradient from 5 to 25% in 30 min at a constant flow rate of 0.8 mL/min. The absorbance of the effluent was recorded at 226 nm. All the peptides gave a single, sharp, and symmetrical peak, with errors in t_R being typically less than ±0.2 min.

The chromogenic substrate D-Phe-Pro-Arg-*p*-nitroanilide (FPR) was synthesized by liquid-phase Fmoc chemistry using the HOBt/HBTU activation procedure, as previously described (*26*).

Spectroscopic Measurements. The concentration of the analogues of fragment 1−47 was determined by ultraviolet (UV) absorption at 280 nm on a double-beam model Lambda-2 spectrophotometer from Perkin-Elmer (Norwalk, CT). The extinction coefficient at 280 nm (ϵ_{280}) for the natural fragment was calculated to be 2920 M^{-1} cm^{-1}, utilizing molar absorption coefficients of 1280 M^{-1} cm^{-1} for Tyr and 120 M^{-1} cm^{-1} for Cys (*27*). For those peptide analogues possessing at position 3 nonabsorbing amino acids (i.e., Ala, Phe, *p*F-Phe, *p*-aminomethyl-Phe, *homo*-Phe, and Cha), the value of ϵ_{280} was taken to be 1640 M^{-1} cm^{-1}. The extinction coefficients for the synthetic peptides Tyr3*p*I-Phe and Tyr3*p*NO_2-Phe were taken to be 1900 and 10 740 M^{-1} cm^{-1}, respectively. The ϵ_{280} values for *p*I-Phe and *p*NO_2-Phe were determined experimentally to be 260 and 9100 M^{-1} cm^{-1}, respectively, by measuring the absorbance of amino acid solutions with known concentrations. ϵ_{280} for Tyr3βNal analogue was calculated to be 7020 M^{-1} cm^{-1}, utilizing a molar absorption coefficient of 5380 M^{-1} cm^{-1} for βNal (*12*, *28*).

Circular dichroism (CD) spectra were recorded on a Jasco (Tokyo, Japan) model J-710 spectropolarimeter equipped with a thermostated cell holder and a NesLab (Newington, NH) model RTE-110 water circulating bath. Far- and near-ultraviolet CD spectra were recorded at 2 °C in 10 mM phosphate buffer (pH 7.0) at a peptide concentration ranging from 20 to 100 μM, using 1 or 5 mm path length quartz cells in the far- or near-ultraviolet region, respectively. Flexibility and/or stability measurements were carried out by recording the ellipticity value at 265 nm ($[\theta]_{265}$) as a function of the temperature increase in the cuvette. Data are reported as the $[\theta]/[\theta]°$ ratio, where $[\theta]°$ is the ellipticity value measured at the starting temperature (2 °C).

Thrombin Inhibitory Activity. The dissociation constants, K_d, for enzyme−inhibitor complexes were determined by using the same experimental procedure and theoretical framework we have already used in our previous studies on hirudin analogues (*12*, *13*). This method has been described in detail by Di Cera and co-workers (*29*) and also used to derive inhibitory properties of the hirudin HV1 variant and its proteolytic fragments 1−49 and 1−43 (*30*). Substrate concentration, product inhibition, and the effects of sodium ion on the enzymatic activity of thrombin were taken into account in determining K_d values. Binding of hirudin fragment 1−47 analogues was quantified using progress curves started by addition of thrombin. Measurements were carried out at 25 °C in 5 mM Tris (pH 8.0) containing 0.1% (w/w) PEG 8000 in the presence of a constant concentration (20 μM) of the chromogenic substrate, in the presence of varying concentrations of the inhibitor. The reaction was started by addition of 50 pM thrombin for studying the fast form and 200 pM thrombin for the slow form. The release of *p*-nitroaniline from D-Phe-Pro-Arg-*p*NA was assessed by recording the increase in absorbance 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. The properties of the fast form were derived as the extrapolation of $[Na^+] \rightarrow \infty$ at a constant ionic strength of 200 mM. Kinetic constants k_m and k_{cat} used in this study were determined to be 0.48 ± 0.03 μM and 44 ± 2 s^{-1} for the fast form, respectively, and 1.9 ± 0.1 μM and 5.7 ± 0.3 s^{-1} for the slow form, respectively. These values yield specificity constants ($s = k_{cat}/k_m$) of 90 and 3.0 μM^{-1} s^{-1} for the fast and slow forms, respectively, identical to those reported by Di Cera and co-workers for the same substrate (*4*).

Computational Methods. The structure of the synthetic analogues of fragment 1−47 of hirudin HM2 in the free state was modeled on the NMR solution structure of natural fragment 1−47 (*10*), which can almost be superimposed with that of the corresponding segment 1−49 in the intact hirudin HV1 variant (*8*) or with that of HV1 fragment 1−51 (*9*). The best-representative model in the NMR ensemble was selected using the program OLDERADO (http://neon.chem.le.ac.uk/). The structure of the synthetic analogues in the thrombin-bound state was modeled on the 2.1 Å resolution crystallographic structure of the hirudin−thrombin complex (*11*; PDB entry 4htc).

The contribution of hydrophobicity to the free energy change of binding (ΔG_b) was estimated by calculating the difference in the free energy change due to desolvation (ΔG_{desolv}) of polar and apolar groups that are shielded from water upon hirudin−thrombin association, and is given by the equation $\Delta G_{desolv} = G_{desolv}(\text{complex}) - [G_{desolv}(\text{thrombin}) + G_{desolv}(\text{hirudin})] = -\sum_i \Delta\sigma_i \Delta ASA_i$ (*31*, *32*). ΔASA_i is the change in the accessible surface area (ASA) for the atom *i* upon binding of hirudin to thrombin, and $\Delta\sigma_i$ (also called the atomic solvation parameter) is the solvation free energy change per unit area of atom *i* that becomes buried upon binding. The desolvation free energy change of binding for a mutated hirudin analogue (M) is calculated relative to that of the wild-type fragment 1−47 (WT) as $\Delta\Delta G_{desolv} = \Delta G_{desolv}(M) - \Delta G_{desolv}(WT)$. Atomic solvation parameters for carbon (C), nitrogen (N), oxygen (O), charged nitrogen (N^+), and charged oxygen (O^-) are those reported by Eisenberg and McLachlan (*33*) [$\Delta\sigma(C) = 16 \pm 2$ cal mol^{-1} $Å^{-2}$, $\Delta\sigma(N/O) = -6 \pm 4$ cal mol^{-1} $Å^{-2}$, $\Delta\sigma(N^+) = -50 \pm 9$ cal mol^{-1} $Å^{-2}$, and $\Delta\sigma(O^-) = -24 \pm 10$ cal mol^{-1} $Å^{-2}$] and based on water → octanol transfer data for amino acid residues at 25 °C (*34*). $\Delta\sigma$ values for fluorine [$\Delta\sigma(F) = 12$ cal mol^{-1} $Å^{-2}$], iodine [$\Delta\sigma(I) = 27$ cal mol^{-1} $Å^{-2}$], and a nitro group [$\Delta\sigma(NO_2) = 3$ cal mol^{-1} $Å^{-2}$] were estimated by the equation $\Delta\sigma(X) = [\Delta\Delta G_{exp}(X) - \Delta\sigma(C)\Delta ASA(C)]/ASA(X)$, where X is F, I, or NO_2. $\Delta\Delta G_{exp}(X)$ is given by the equation $\Delta\Delta G_{exp}(X) = -2.30RT\pi_X$, where *R* is the gas constant, *T* equals 298.15 K, and $\pi(X)$ is the difference between the logarithm of the water → octanol partition coefficient (log *P*) of the *p*-X-substituted toluene derivative (ϕ-X) and that of toluene (ϕ-H) [$\pi(X) = \log P_{\phi-X} - \log P_{\phi-H}$]. $\Delta\sigma(C) = 16 \pm 2$ cal mol^{-1} $Å^{-2}$, and $\Delta ASA(C) = ASA(C)_{\phi-X} - ASA(C)_{\phi-H}$. The values of the water → octanol partition coefficients for toluene (log $P = 2.73$) and *p*-nitrotoluene (log $P = 2.42$) were those reported by Sangster (*35*), while log *P* values for *p*-iodo- and *p*-fluoro-substituted toluene derivatives were calculated by adding the contribu-

tion of the methyl group [π(-CH_3) = 0.60 (*36*)] to the corresponding values determined experimentally for fluorobenzene (log $P_{\phi-F}$ = 2.27) and iodobenzene (log $P_{\phi-I}$ = 3.25) (*35*). ASA calculations were carried out on the natural and synthetic peptide analogues in the free and thrombin-bound state by using a computer program available on-line at http://molbio.info.nih.gov/structbio/basic.html (*37*). Volumes of pockets and cavities were calculated with the computer program CASTp (*38*) which is available on-line at http://cast.engr.uic.edu/cast1/. ASA and volume calculations were carried out by using a water probe radius of 1.4 Å. When the additivity properties of electric dipole moments, μ, are considered, the values for the *p*-X-substituted toluene derivatives (Me-ϕ-X) are estimated with the equation μ(Me-ϕ-X) = μ(ϕ-X) + μ(Me-ϕ), where μ(ϕ-X) and μ(Me-ϕ) are the values of μ determined experimentally at 25 °C in benzene for X-substituted benzenes (ϕ-X) and for toluene (Me-ϕ), respectively (*39*). μ values are measured in Debye (D), where 1 D = 3.336×10^{-30} C/m^{-1}. Side chain molecular volumes were calculated using PC-Spartan-Pro version 1.0.1 (Wavefunction, Inc.).

Molecular Dynamics Simulation Protocol. The starting structure for the simulation was that of the PPACK-inhibited thrombin, determined by Bode et al. (*40*) at 1.9 Å resolution (PDB entry 1ppb). Comparison of the α-carbon tracing of PPACK-inhibited thrombin with that of the hirudin−thrombin complex, crystallized in the presence of sodium ion as well (*11*; PDB entry 4htc), reveals that approximately two-thirds of all Cα atoms differ by only 0.26 Å in rmsd, while larger differences are in the Trp148 and Trp60D insertion loops, as well as in the N- and C-terminal ends (*40*). This highly conserved structural similarity between the two thrombin structures allows a more realistic comparison of the results obtained experimentally from amino acid exchanges on hirudin structure with those inferred from MD simulations on thrombin.

The covalently bound inhibitor D-Phe-Pro-Arg-methyl ketone was removed, and the protein was solvated with water in a periodic truncated octahedron large enough to contain the protein and 0.8 nm of solvent on all sides. All solvent molecules within 0.15 nm of any protein atom were removed. The total charge of the protein was +3, and no counterions were added. As water has a high dielectric constant, the inclusion of no counterions was considered a better approximation of the experimental conditions. The resulting system was composed of 3071 protein atoms and 15 423 water molecules. The system was then energy minimized with a steepest descendent method for 1000 steps to remove all the possible bad contacts. During the simulation, the protein and the rest of the system were coupled separately to a temperature bath maintained close to the intended values by weak coupling to an external temperature bath with a coupling constant of 0.1 ps (*41*). The GROMOS-96 force field was used (*42*) for the minimization procedure, with a time step of 2 fs. Water molecules were modeled according to the "simple point charge" (SPC) model (*43*). The LINCS algorithm (*44*) was used to constrain all bond lengths. A twin range cutoff was used to calculate the nonbonded interactions. The short-range cutoff radius was set to 0.8 nm and the long-range cutoff radius to 1.4 nm for both Coulombic and Lennard-Jones interactions. The interactions within the short range were updated every time step, while interactions within the long-range cutoff were updated every five time steps together with the pairlist. All atoms were given an initial velocity obtained from a Maxwellian distribution at the desired initial temperature. The density of the system was adjusted by performing the first equilibration runs under NPT conditions by weak coupling to a bath of constant pressure (P = 1 bar, coupling time τ_p = 0.5 ps). The simulation, starting from the crystal structure, was equilibrated by 50 ps of MD run with position restraints on the protein to allow relaxation of the solvent molecules. This first equilibration run was followed by another 50 ps run without position restraint on the protein. The production run was carried out under NVT conditions for 18 ns, after equilibration. All the MD runs and the analysis of the trajectories were performed using the GROmacs software package (*45*) run on a DEC-alpha PC cluster.

Essential dynamics (ED) analysis (*46*) was also carried out to separate the motions of a protein into an essential subspace describing most of the functional motions, and into a physically constrained subspace, merely describing irrelevant local fluctuations. The essential subspace is defined by a set of orthonormal vectors, associated with the larger eigenvalues of the atomic positional fluctuation covariance matrix. Eigenvectors defining the directions of higher displacement are generally listed in decreasing order of the value of the corresponding eigenvalue, and the first few of them can capture most of the essential motions of the protein. The correlation of the essential motions of two segments of the protein was estimated by calculating the cross-correlation function between the projection of the two protein segments onto the first eigenvector obtained from the ED analysis of both segments.

RESULTS

Selection of Amino Acid Replacements. The X-ray structure of the complex of human α-thrombin with the hirudin HV2 variant (*11*) indicates that the three N-terminal residues of hirudin are completely buried in the active site of the enzyme and form a parallel β-sheet with Ser214−Gly219 of thrombin. In particular, the amino terminus forms a hydrogen bond with the OH group of Ser195 and points toward the S2 site, shaped by the Leu59−Leu62 insertion loop; the amino acid at position 2 is located at the entrance of, but does not penetrate into, the S1 binding site, and finally, Tyr3 interacts with the apolar S3 site, lined by Trp215, Ile174, and Leu99.

In a previous work (*12*), we have shown that a rough correlation exists between the increase in hydrophobicity at position 3 of hirudin fragment 1−47 and the affinity of the resulting inhibitors for thrombin. Taking advantage of the versatility of chemical synthesis, in this study, we decided to keep the hydrophobic and steric properties of the amino acids at position 3 rather constant and change their electric and conformatioanl characteristics. To this end, we have introduced electron-withdrawing groups (i.e., F, I, and NO_2) at the para position of the phenyl ring. We have also introduced a net positive charge at position 3 by replacing Tyr3 with *p*-aminomethyl-Phe. This phenylalanine analogue carries a primary amino group with a pK_a of 9.36 (*35*), and for this reason, it can be considered a suitable rigid analogue of Lys. Thereafter, to explore the influence of side chain flexibility and orientation on thrombin affinity, we have

Table 1: Physicochemical Properties of Amino Acid Side Chains at Position 3 of Hirudin Fragment 1−47

side chain at position 3	volume[a] (Å^3)	log P^b (π)	F^c	R^c	μ^d (D)
Tyr (wild type)	138	1.97 (0.96)	0.29	−0.64	−1.57 2.02[e]
Phe	127	2.73 (1.79)	0.00	0.00	0.36
Cha	146	3.88 (2.72)			0.0
*p*F-Phe	133	2.87	0.43	0.34	1.79
pNO_2-Phe	157	2.42 (1.96)	0.67	0.16	4.38
*p*I-Phe	160	3.85	0.40	−0.19	1.72
homo-Phe	147	3.15 (2.10)			0.39
p-aminomethyl-Phe	165	−0.80			na[f]
Ala	38	0.60 (0.31)			0.0
βNal	180	4.00 (3.15)			0.44

[a] Side chain molecular volumes are referenced to the volume within the van der Waals surface, and are calculated for the corresponding organic compounds taken to be suitable models for the amino acid side chains (e.g., *p*-cresol for Tyr and toluene for Phe). [b] Log *P* values are those of the corresponding model compounds determined experimentally at 25 °C and reported by Sangster (*35*). When available, the hydrophobic substituent constants, π, of amino acid side chains are given in parentheses (*34*). [c] *F* and *R* are the field and resonance contributions to the Hammett electronic substituent constant, σ_p, respectively (*52*). Electron-withdrawing groups possess positive *F* and *R* values, whereas electron-releasing groups have negative *R* values. [d] μ is the electric dipole moment of organic model compounds of the amino acid side chains determined at 25 °C in benzene (*39*, *62*). [e] μ value of Tyr corrected for the absence of resonance effects (see the Discussion). A positive sign of μ stands for a negative end pointing away from the phenyl ring, toward the *p*-X substituent. A detailed description of the procedures employed to obtain estimates of the physicochemical properties of the side chains is reported in Experimental Procedures. [f] Not available.

replaced Phe with its higher homologue *homo*-Phe, which is more hydrophobic than Phe and has a greater side chain conformational entropy, due to the presence of an extra rotatable bond between Cα and the phenyl ring. The physicochemical properties of the organic compounds (i.e., toluene derivatives), taken to be suitable models of the corresponding amino acid side chains, are reported in Table 1.

Finally, we have investigated the features responsible for the selectivity of hirudin toward the procoagulant (fast) form of thrombin by evaluating the effects of coarse variations in the side chain volume of the amino acid residue at position 3. To this end, we have decided to lessen or enlarge the size of the side chain at this position by replacing Tyr3 with Ala or β-naphthylalanine. These amino acids possess similar conformational propensities (*47*), but very different molecular volumes and hydrophobicity values (Table 1).

Synthesis and Characterization. Chemical synthesis of peptide analogues of the N-terminal domain 1−47 of hirudin was carried out by a combination of manual and automated solid-phase synthesis using the synthetic strategy previously described (*12*, *22*). The crude peptides with the Cys residues in the reduced state were allowed to fold (2 mg/mL) under air-oxidizing conditions in bicarbonate buffer (pH 8.3) in the presence of 100 μM β-mercaptoethanol (*22*, *25*). As an example, the RP-HPLC analyses of the crude synthetic analogue Tyr3*homo*-Phe in the reduced and disulfide-oxidized state are reported in Figure 2. The chemical identity of the disulfide folded peptides was established by N-terminal sequence analysis and ESI-TOF mass spectrometry, which gave experimental values in agreement with theoretical amino

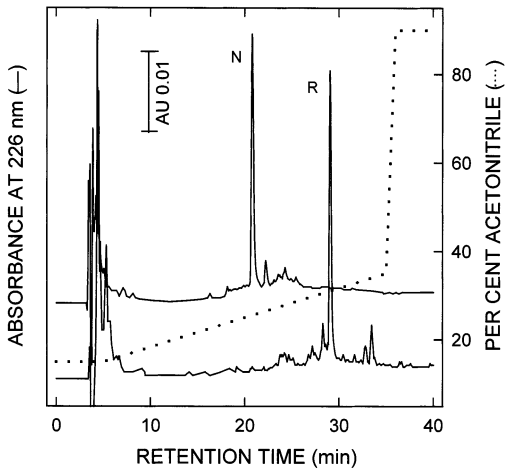

FIGURE 2: RP-HPLC analysis of the crude synthetic analogue Tyr3*homo*-Phe in the reduced (lower trace) and disulfide-oxidized states (upper trace). The chromatographic separations were conducted on a Vydac C18 analytical column (4.6 mm × 250 mm, 5 μm) eluted with a linear acetonitrile/0.05% TFA (···) gradient at a flow rate of 0.8 mL/min and recording the absorbance of the effluent at 226 nm.

acid composition within 10 ppm mass accuracy. For some of the synthetic peptides, the exact topology of disulfide bonds was established by using the enzymatic fingerprint analysis previously reported for the natural fragment (*22*). All the peptides were purified by preparative RP-HPLC, lyophilized, and used for subsequent conformational and functional characterization.

The conformational properties of the synthetic analogues of hirudin fragment 1−47 were investigated by CD spectroscopy in the far- and near-UV region. The far-UV CD spectra of all the synthetic peptides share common features typical of β-like secondary structure, with a negative minimum centered at ∼217 nm and a positive band at 200 nm (not shown). However, as outlined previously (*12*, *22*), the CD signal of these small proteins in the far-UV region is strongly influenced by the contribution of the aromatic amino acids along the polypeptide chain, and thus, it is of limited utility in the analysis of protein conformation. Conversely, near-UV CD is a very sensitive probe of protein tertiary structure, and it is usually taken as a fingerprint of the chemical environment in which aromatic amino acids and disulfide bonds are located within protein structure (*48*). As shown in Figure 3, the near-UV CD spectra of the synthetic hirudin analogues are dominated by a prominent positive band centered at 260 nm, corresponding to the absorbance of the three disulfide bonds in the right-handed chirality (*22*), in agreement with the NMR solution structures of the hirudin N-terminal domain (*8*−*10*), while the fine structure observed in the range of 280−300 nm is due to the contribution of aromatic amino acids. The spectrum of the natural fragment 1−47 has shape and molar ellipticity values closely similar to those of the full-length hirudin HM2 variant from *H. manillensis* (unpublished), as well as that of the more studied hirudin HV1 variant from *Hirudo medicinalis* (*49*), its sequence being ∼75% homologous with that of hirudin HM2 (*70*). The moderately different spectral properties of the Tyr3Cha analogue with respect to those of wild-type fragment 1−47 are almost exclusively assigned to the lack of Tyr3 contribution, as also given by the fact that the difference spectrum obtained by subtracting the

Table 2: Thermodynamic Data for the Binding of Synthetic Analogues of Fragment 1−47 to the Procoagulant (Fast) and Anticoagulant (Slow) Forms of Thrombin[a]

	fast form		slow form		
side chain at position 3	K_d (nM)	$\Delta\Delta G_b$ (kcal/mol)	K_d (nM)	$\Delta\Delta G_b$[b] (kcal/mol)	ΔG_c[c] (kcal/mol)
Tyr (wild type)	42 ± 0.5	−	1460 ± 20	−	−2.12
Phe	4 ± 0.1[d]	−1.37	203 ± 3	−1.18	−2.31
Cha	49 ± 0.6[d]	0.09	1740 ± 23	−0.09	−2.12
*p*F-Phe	18 ± 0.3	−0.48	583 ± 6	0.56	−2.04
pNO_2-Phe	185 ± 3.0	0.88	2035 ± 2	0.22	−1.46
*p*I-Phe	24 ± 0.2	−0.33	1020 ± 12	−0.23	−2.22
homo-Phe	157 ± 2.5	0.79	3290 ± 38	0.46	−1.79
p-aminomethyl-Phe	140 ± 2.0	0.72	1590 ± 22	0.03	−1.44
Ala	2750 ± 23	2.47	2600 ± 30	0.32	0.03
*β*Nal	1.1 ± 0.1[d]	−2.14	94 ± 4	−1.64	−2.64

[a] All measurements were carried out at 25 °C in 5 mM Tris (pH 8.0) containing 0.1% PEG in the presence of 200 mM NaCl for the fast form or 200 mM ChCl for the slow form (Experimental Procedures). [b] $\Delta\Delta G_b$ 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 more tightly to thrombin than the natural fragment. Errors are less than or equal to ±0.1 kcal/mol. [c] ΔG_c is the free energy of coupling to thrombin measured as $\Delta G_c = \Delta G_{b,fast} - \Delta G_{b,slow}$ (*29*). The value of ΔG_c is negative if the inhibitor binds with higher affinity to the fast form than to the slow form. [d] For these analogues, thrombin inhibition data were determined experimentally in this study for a second time and found to be in close agreement with those reported in our previous work (*12*).

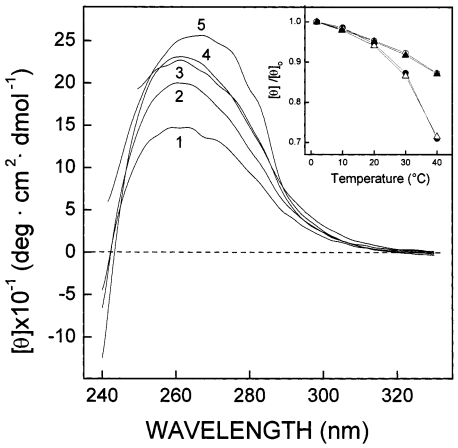

FIGURE 3: Near-UV CD spectra of the natural fragment 1−47 and its synthetic analogues mutated at position 3: Tyr3*homo*-Phe (curve 1), Tyr3Ala (curve 2), Tyr3pI-Phe (curve 3), Tyr3Cha (curve 4), and natural fragment 1−47 (curve 5). All spectra were recorded at 2 °C, at a protein concentration of 0.8−1.0 mg/mL in 10 mM sodium phosphate buffer (pH 7.0). The inset shows the temperature dependence of the $[\theta]/[\theta]°$ ratio for the natural fragment 1−47 (○) and the analogues Tyr3*p*I-Phe (▲), Tyr3Ala (●), and Tyr3*homo*-Phe (△).

spectrum of the Tyr3Cha analogue (curve 4 in Figure 3) from that of the wild-type (Tyr3-containing) fragment (curve 5) resembles that typical of Tyr (not shown). This also holds for the *p*I-Phe analogue (curve 3) and the other para-substituted phenylalanine derivatives (i.e., *p*F-Phe and *p*NO_2-Phe), which (like Cha) give only a weak absorption in the range of 250−320 nm. In the case of the hirudin analogues containing Ala (curve 2) or *homo*-Phe (curve 1), the even lower intensity of the near-UV CD signal is caused by the increased conformational flexibility of these fragment 1−47 derivatives, in addition to merely spectroscopic effects. This view is supported by the steeper temperature dependence of the ellipticity, $[\theta]$, observed for Tyr3Ala and Tyr3*homo*-Phe in the range of 2−40 °C (see the inset of Figure 3) with respect to that of the wild-type species, which is identical to that of Tyr3*p*I-Phe and to those the other para-substituted analogues (not shown).

These findings allow us to interpret the differences in affinity for thrombin for most of the synthetic analogues (i.e., Cha and *p*X-Phe derivatives) herein reported on the basis of the variation of the physicochemical properties at the mutation site, but also suggest that the influence of the conformational flexibility of the inhibitor (i.e., Tyr3Ala and Tyr3*homo*-Phe) in the free state should also be adequately taken into account for a correct interpretation of binding data.

Antithrombin Activity. The inhibitory potency of the synthetic analogues toward the procoagulant (fast) and anticoagulant (slow) forms of 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 and the difference in the free energy change of binding, $\Delta\Delta G_b$, relative to wild-type fragment 1−47 have been obtained (Table 2).

Tyr3 → Phe and Tyr3 → Cha. Thrombin binding data indicate that Tyr3 → Phe exchange improves affinity for thrombin by 10-fold, whereas saturation of the aromatic ring of Phe (by Phe → Cha substitution) reduces the level of binding by a similar amount. If hydrophobicity were the dominant driving force for binding, the Tyr3Cha analogue would be expected to be a more potent inhibitor than the wild-type and Tyr3Phe derivative, since Cha has a side chain volume very similar to that of Tyr or Phe, but is significantly more hydrophobic (see Table 1). These findings confirm our previous results (*12*) and can be explained on the basis of the three-dimensional structure of the hirudin−thrombin complex (*11*), showing that the δ^+ hydrogens on the edge of the 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 4). These kinds of aromatic−aromatic interactions are known to contribute 0.6−1.3 kcal/mol to protein stability or binding energy (*50*). Therefore, the lack of this edge-to-face interaction in the Tyr3Cha analogue may result in looser binding to thrombin.

Tyr3 → pX-Phe. The presence of F, I, or NO_2 groups on the phenyl ring of the amino acid at position 3 is expected to enhance binding to thrombin. In fact, these para-substituted Phe derivatives are slightly more hydrophobic than Tyr or

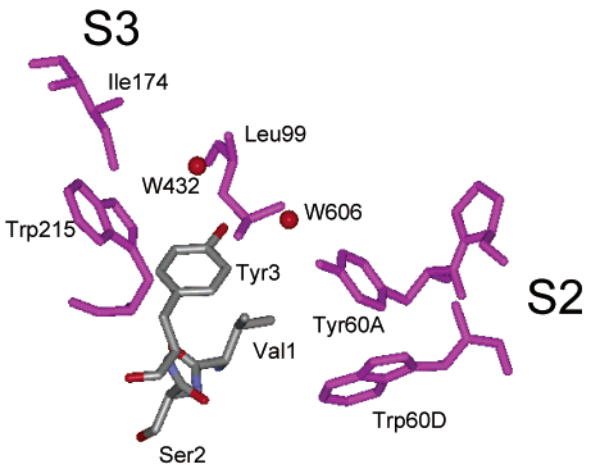

FIGURE 4: Schematic representation of the interaction of the N-terminal tripeptide of hirudin (gray) with thrombin recognition sites (magenta). Tyr60A and Trp60D of the S2 site and Trp215, Leu99, and Ile174 forming the apolar S3 site on thrombin are shown. Water molecules at the hirudin−thrombin interface in the S2 and S3 sites are represented with red spheres. The OH group of Tyr3 in hirudin interacts with water molecule W432 (*B*-factor of 36 Å^2, occupancy of 1.0) and is connected to Tyr60A in the S2 site through a water bridge involving W606 (*B*-factor of 22 Å^2, occupancy of 0.52).

Phe, possess very similar conformational propensities, and (as shown by model building studies) can be easily accommodated at the S3 site of thrombin without clashes. More importantly, the electron-attracting character of these substituents, as given by the values of *F* and *R* constants [the field and resonance contributions, respectively, to the Hammett electronic substituent constant, σ_p (*51*)] (see Table 1), would make the hydrogen atoms of the aromatic nucleus at position 3 more electron-deficient, thus reinforcing the electrostatic interaction with the δ^- π-electron cloud of Trp215. Contrary to expectations, electron-withdrawing substituents failed to improve thrombin binding. Conversely, their presence resulted in a significant reduction in affinity of 4 (i.e., Tyr3*p*F-Phe) to 46-fold (i.e., Tyr3*p*NO_2-Phe) with respect to that of Tyr3Phe (Table 2).

Rather surprisingly, the presence of a positive charge at position 3, by Tyr3 → *p*-aminomethyl-Phe exchange, reduced the affinity for the fast form of thrombin by only 3-fold with respect to wild-type fragment 1−47 and did not affect the binding strength to the slow form. Notably, analysis of model structure and ASA calculations of the Tyr3*p*-aminomethyl-Phe analogue complexed to thrombin reveal that the charged nitrogen atom of *p*-aminomethyl-Phe is shielded from water and buried in the apolar S3 site, formed by Trp215, Leu99, and Ile174 (*11*), and traditionally termed the "aryl binding site" (*52*). Theoretical calculations (*53*) estimated an energy cost of ∼10 kcal/mol to bury a charge in the protein interior, and experimental results obtained on phage T4 lysozyme indicated that replacement of Met102 with Lys in the protein core resulted in a protein having a three-dimensional structure that can be superimposed with that of the wild-type species, but less stable by ∼9 kcal/mol (*54*). In our case, the difference in hydrophobicity (see Table 1) between *p*-aminomethyl-Phe and Tyr or Phe would yield an unfavorable increase in ΔG_b to thrombin of 3.8 or 4.8 kcal/mol, respectively. Hence, the low energetic penalty of 0.72 or 0.03 kcal/mol experimentally derived for the binding of Tyr3*p*-aminomethyl-Phe to the fast or slow form of the enzyme, respectively, would lead us to reconsider the apolar properties of the S3 site or, more reasonably, to invoke specific interactions of the charged amino group within the S3 site (see below) that would compensate for its unfavorable desolvation properties.

Tyr3 → homo-Phe. The presence of *homo*-Phe at positon 3 of hirudin fragment 1−47 reduces antithrombin activity by 4- and 40-fold with respect to those of the wild type and Tyr3Phe analogue, respectively. Therefore, although *homo*-Phe is more hydrophobic than Phe, the corresponding fragment 1−47 derivative binds to the fast form of thrombin with an unfavorable change in ΔG_b of 2.1 kcal/mol (Table 2). These findings may be explained by considering that the Tyr3*homo*-Phe analogue, while retaining the overall conformational properties of wild-type fragment 1−47, is significantly more flexible than the wild type and Tyr3Phe analogue in the free state, as given by the steeper temperature dependence of ellipticity, [θ] (see the inset of Figure 3). The greater flexibility of Tyr3*homo*-Phe in the free state results in a larger entropic cost (ΔS_b) for binding of the inhibitor to thrombin and thus in a lower (absolute) value of ΔG_b. In addition, analysis of the model structure of Tyr3*homo*-Phe complexed to thrombin indicates that, due to steric clashes, the side chain of *homo*-Phe cannot establish favorable contacts with Trp215, but it is more likely that *homo*-Phe points toward Tyr60A and Trp60D in the S2 site, leaving an uncompensated cavity at the S3 site. Hence, the reduced antithrombin activity of Tyr3*homo*-Phe is caused either by an improper interaction with thrombin recognition sites or by a flexibilization of the inhibitor in the unbound state.

Tyr3 → Ala and Tyr3 → βNal. The results reported above indicate that position 3 is crucial for binding to thrombin and that even subtle modifications at this site can significantly alter the affinity for the enzyme. These findings prompted us to investigate the effects of coarse variations in the side chain volume at position 3 on the binding to the fast or slow form. Thrombin inhibition data reported in Table 2 indicate that shaving of Tyr3 at Cβ reduces the affinity for the fast form of thrombin by 65-fold, compared to the affinity of the wild-type species, and by 680-fold with respect to that of the Tyr3Phe analogue. Conversely, replacement of Tyr3 with the more hydrophobic and bulkier β-naphthylalanine (βNal) enhanced the affinity for the procoagulant (fast) form of thrombin by ∼40-fold. These data indicate that via a single amino acid replacement (Ala3 → βNal) it is possible to change the affinity of hirudin fragment 1−47 for thrombin by ∼2750-fold, with a difference in ΔG_b of 4.6 kcal/mol.

The remarkable increase in the affinity for the fast form, observed upon Tyr3 → βNal exchange, is mainly caused by the unusually high hydrophobicity of βNal, which in the free inhibitor is almost fully exposed to solvent. Moreover, the additional benzene ring of βNal can be accommodated at the S3 site of thrombin without bumps, can still interact in an energetically favorable T-shaped conformation with Trp215, and can harbor Ile174 and Leu99 through van der Waals interactions. On the other hand, the reduction in affinity for thrombin in the fast form, observed for the Ty3Ala analogue, can be reasonably explained either by the lower hydrophobicity of Ala compared to that of Tyr or βNal (see Table 1) or by the loss of numerous van der Waals contacts with the S3 site of the enzyme. Analysis of the

model structure of Tyr3Ala complexed to thrombin reveals that shaving of the Tyr3 side chain creates an ~150 Å^3 cavity at the inhibitor−enzyme interface within the S3 region, which becomes more accessible to water by ~85 Å^2. Penetration of water at the binding interface increases the local dielectric constant value and thus weakens inhibitor−enzyme electrostatic interactions, with a resulting destabilization of the hirudin−thrombin complex. In addition, as in the case of Tyr3*homo*-Phe, the Ala-containing analogue in the free state is significantly more flexible than wild-type fragment 1−47 (see Figure 3), and thus, its binding to thrombin is disfavored on entropic grounds as well.

Strikingly, the presence at position 3 of a small-sized amino acid like Ala reduces the affinity for the fast form by 65-fold and for the slow form by only 1.8-fold, compared to that of the natural product, and actually abolishes the coupling free energy for the binding of hirudin fragment 1−47 to the fast or slow form of thrombin ($\Delta G_c = 0.03$ kcal/mol). Conversely, the replacement of Tyr3 with the bulkier βNal improves affinity preferentially for the fast form, with a coupling free energy of −2.6 kcal/mol (see Table 2). These findings shed new light on the structural requirements that are responsible for the 30-fold preferential binding of hirudin (and its N-terminal domain 1−47) to the fast form of thrombin, and suggest that the S3 site of the enzyme in the procoagulant (fast) form is in a more open and accessible conformation with respect to the less forgiving structure it adopts in the slow form.

Molecular Dynamics Simulation. In this study, we have used molecular dynamics (MD) simulations to obtain a structural model of thrombin capable of reasonably accounting for the different recognition properties of the procoagulant (fast) and anticoagulant (slow) forms of the enzyme toward the hirudin analogues reported above as well as its natural substrates (e.g., fibrinogen and protein C). The starting structure of thrombin used in the MD simulations reported herein is that determined at 1.9 Å resolution in the crystal of the enzyme covalently bound to D-Phe-Pro-Arg-methyl ketone (PPACK) (*40*). The thrombin molecule (depleted of the inhibitor) was allowed to evolve in fully explicit water solvent for a time period of 18 ns, using the simulation protocol reported in Experimental Procedures.

Global Properties. The overall secondary structure elements are in general stable along the whole trajectory (not shown). However, despite this general characteristic, analysis of rmsd fluctuation for each residue during the simulation (Figure 5A) reveals that the regions defined by the Trp148 loop (residues 180−190 in the sequence numbering used in this study) and Trp60D loop (S2 site), as well as amino acids 93−101, show a high degree of flexibility. To identify the motions which are actually relevant for the thrombin conformation, and possibly for its function, separating them from irrelevant fluctuations, we applied the essential dynamics (ED) analysis (*46*) to key regions of the thrombin molecule. This kind of analysis is a valid method for assessing the correlation existing between different parts of the protein and their motions, ruling out possible artifacts that are derived from a simple rmsd analysis. Hence, it was possible to identify two well-defined conformational minima on the energy landscape. In particular, our MD simulations indicate that after ~5 ns, thrombin undergoes a significant

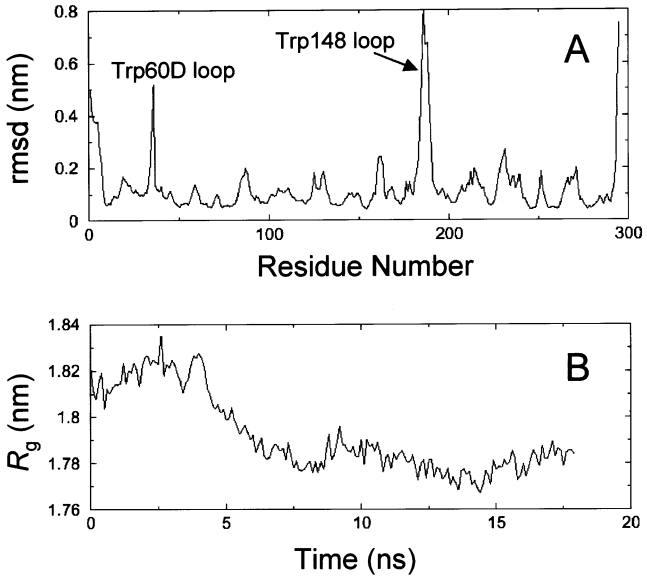

FIGURE 5: (A) Root-mean-square fluctuation of thrombin residues during the molecular dynamics simulation. (B) Variation of the radius of gyration of thrombin, R_g, as a function of time. R_g is defined as the mass-weighted rms distance of a collection of atoms from their common center of mass. Lower R_g values are compatible with an overall more compact protein conformation.

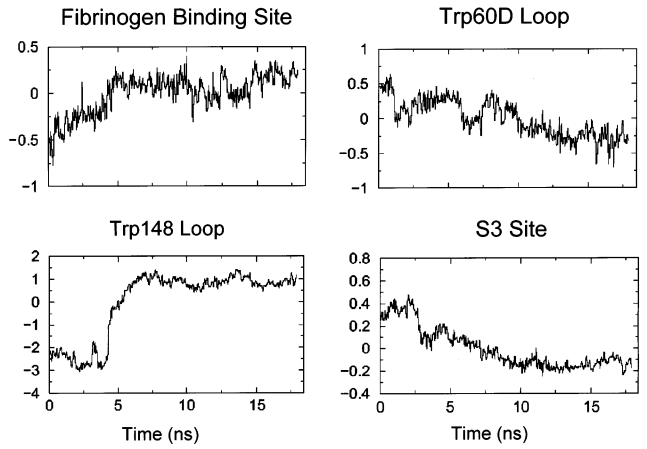

FIGURE 6: Essential dynamics analysis of key regions of thrombin. Motions along the first eigenvector of the selected regions of thrombin in the time frame of the simulation are reported.

conformational transition mainly involving the Trp148 and Trp60D loops, the S3 site, and the fibrinogen binding site, whereas the primary specificity site S1, the Na^+-binding site, and the catalytic pocket remain essentially unchanged. This transition leads the thrombin molecule from a more open and accessible conformation, which we propose can be related to the procoagulant (fast) form, to a more compact and closed form, which we propose can be related to the anticoagulant (slow) form. The more compact structure of the putative slow form is also inferred from the 230 Å^2 smaller surface area it exposes to water with respect to that of the more open fast form, and by the reduction of the radius of gyration (R_g) occurring after simulation for ~5 ns (Figure 5B). Moreover, the results of ED analysis outline the high degree of correlation that exists between the motions of all these regions of thrombin, all occurring after ~5 ns (Figure 6), and suggest that a structural network is present, capable of communicating conformational changes between the different structural domains of the enzyme.

In the following, analysis of key regions that are relevant for the functional properties of thrombin in the two allosteric forms is reported.

Trp148 Loop. Our data indicate that the Trp148 loop is a region of high conformational mobility and that after simulation for ~5 ns it moves from a solvent-exposed conformation to one in which the loop is partially folded onto the protein surface. This transition can have an important influence on other regions of thrombin. In particular, it may restrict the access to the catalytic pocket, by closing it on the lower rim, and to the fibrinogen binding site, by triggering Arg73 to protrude into a crevice leading to the catalytic pocket. Although this picture is attractive and correlates well with the known recognition properties of the fast and slow forms of thrombin (*4*), as well as with our results obtained on hirudin analogues (see obove), great caution should be used in defining the role of the Trp148 loop in the slow → fast transition. In fact, our MD simulations also indicate that the Trp148 loop can sample a wide range of conformations, which cannot yet be reached by all-atom simulations. Moreover, this loop is characterized by an intrinsically high conformational flexibility, as well documented by a number of X-ray crystal structures showing that it acquires different conformations when thrombin is complexed to different inhibitors (*40*).

Catalytic Pocket. The geometry of the catalytic triad (His57, Asp102, and Ser195) remains unchanged during the simulation and close to the atomic positions it has in the starting structure as the protein could be expected to remain active. A hydrogen bond is always present between His57 and Asp102; the hydrogen bond between Asp102 and Ser195 is present only in the first 4 ns of the simulation, while it is absent in the remaining part of the simulation, where the conformation of thrombin may be related to that of the slow form. This fact may have some implication in determining the observed reduced catalytic efficiency of the enzyme in the slow form (*4*).

Na^+-Binding Site. The Na^+-binding site of thrombin is located within a cylindrical cavity formed by three antiparallel β-strands of the B chain (Met180−Tyr184A, Tyr225−Tyr228, and Val213−Cys220), diagonally crossed by the Glu188−Glu192 segment, and shaped by the Asp221−Lys224 loop connecting the last two β-strands. The Na^+ ion is coordinated octahedrally to the carbonyl oxygen atoms of Tyr184A, Arg221A, and Lys224, and by four water molecules (*55*). Interestingly, the results of MD simulations indicate that the conformation of the Na^+-binding site remains stable during the simulation. The site is thus "preorganized" for Na^+ binding, with an obvious energetic advantage. In this respect, binding of the metal ion would not be so entropically expensive, and thus, it might provide the necessary amount of energy to drive the conformation of thrombin from the slow to the fast form.

S2 and S3 Specificity Sites. The stereochemical properties of the insertion loops have been recognized as the major determinants of the unique specificity of thrombin, compared to other serine proteases (*40*). The conformational change of the amino acid residues within the S2 pocket is evident from the dihedral angle transitions and from the rmsd calculations for this group of atoms (Figure 5A). In particular, the rmsd value for both the Cα and all-atom calculations increases from 0.1 to 0.3 Å, after simulation for ~1 ns. Of note, the solvent accessible area of the S2 site does not show any particular variation during the simulation. The same type of analysis was carried out for the residues defining the S3 site, namely, Trp215, Leu99, and Ile174. The values for the side chain torsional angles of Trp215 do not show sensitive variations. Nonetheless, the solvent accessible surface area shows a well-defined variation from 150 to 50 $Å^2$, occurring after ~5 ns and coincident with the conformational transition described above.

DISCUSSION

Structure−Activity Relationships. Given the apolar character of the S3 site, traditionally termed the "aryl binding site" of thrombin (*52*), we have attempted to relate the affinity of the synthetic analogues of hirudin fragment 1−47 for the fast form of thrombin to the hydrophobicity value of the amino acid side chain at position 3. As shown in Figure 7A, plotting the change in binding free energy (ΔG_b) versus the logarithm of the octanol → water partition coefficient (log *P*) of the amino acid side chain at position 3 yields a poor correlation (see the legend of Figure 7 and Table 1). Even worse behavior is obtained when the RP-HPLC retention time (t_R) is used (Figure 7B) as an experimental measure of the surface hydrophobicity of hirudin analogues. In this respect, RP-HPLC has been widely used to derive lipophilic properties of low-molecular weight pharmaceuticals (*56*) or even surface properties of globular proteins under mildly denaturing conditions (*57*). In the case of hirudin fragment 1−47, synthetic analogues having t_R values similar to or even higher than that of the wild-type species (e.g., Tyr3*p*NO_2-Phe and Tyr3*homo*-Phe) display remarkably lower affinities for thrombin.

To better investigate the role of hydrophobicity in driving hirudin binding to thrombin, we used the formalism introduced by Eisenberg and McLachlan (*33*). This approach assumes that the desolvation free energy change of binding $\{\Delta G_{desolv} = G_{desolv}(\text{complex}) - [G_{desolv}(\text{receptor}) + G_{desolv}(\text{ligand})]\}$ is related to the amount of polar and apolar surface area of both the ligand and receptor that becomes buried upon formation of the complex (*31, 32*) (see also Experimental Procedures). Data reported in Figure 7C clearly indicate that the experimental ΔG_b is linearly related ($r = 0.92$) to the variation in desolvation free energy (ΔG_{desolv}) calculated on theoretical grounds. Only two data points strongly deviate from the regression line, namely, Tyr3*p*I-Phe and Tyr3*p*-aminomethyl-Phe. In particular, they were found to bind thrombin much more tightly than predicted by desolvation free energy calculations. Structural analysis of the corresponding complexes with thrombin reveals that both the iodine atom of *p*I-Phe and the charged nitrogen of *p*-aminomethyl-Phe can productively interact with the aromatic nucleus of Trp215 at the S3 site, thus compensating for (or overwhelming) their unfavorable desolvation free energy of binding. In the case of the Tyr3*p*I-Phe analogue, the favorable interaction of iodine with Trp215 is predominantly driven by dispersive forces related to the high polarizability of iodine. In this regard, I_2 forms rather stable complexes with benzene (*58*). On the other hand, the surprisingly high affinity of the Tyr3*p*-aminomethyl-Phe analogue can be rationalized by taking into account a specific charge−π interaction between the protonated amino group of *p*-aminomethyl-Phe and the π-electrons of Trp215. Since

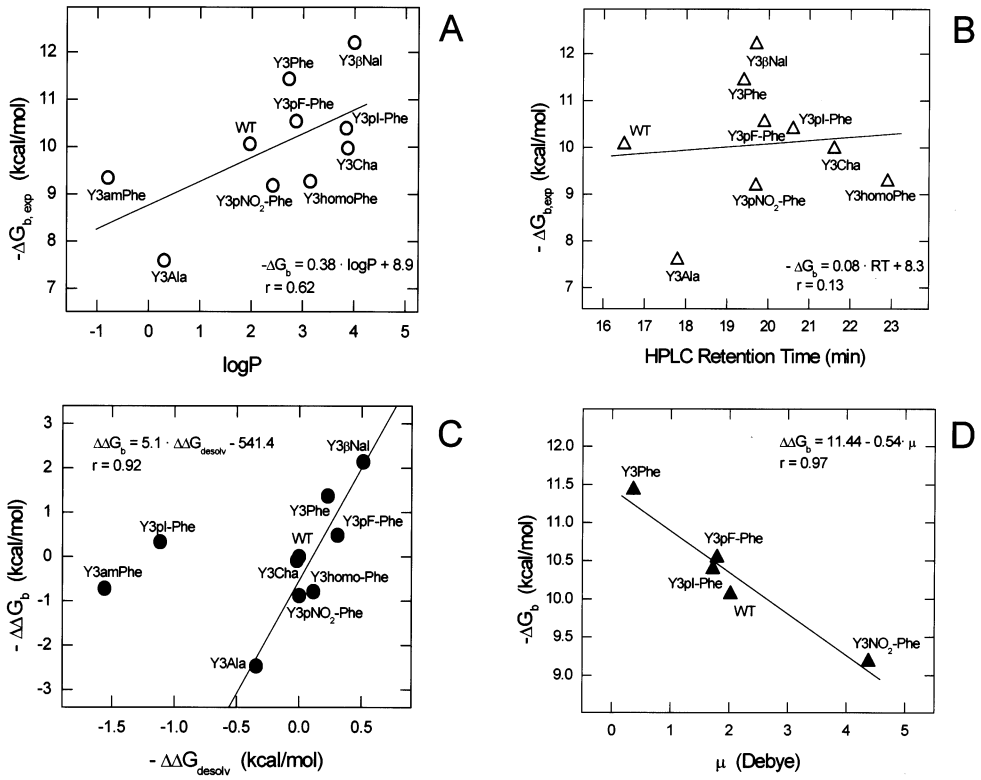

FIGURE 7: (A) Plot of the experimental free energy change of binding to the fast form of thrombin (ΔG_b) for the synthetic analogues of fragment 1−47 mutated at position 3 as a function of the log *P* value of the amino acid side chain. (B) Plot of ΔG_b vs the retention time (t_R) determined at pH 6.4 for Tyr3 → X analogues of fragment 1−47. The error in the determination of t_R values was less than ±0.1 min. (C) Plot of experimental $\Delta\Delta G_b$ vs $\Delta\Delta G_{desolv}$. Both $\Delta\Delta G_b$ and $\Delta\Delta G_{desolv}$ are calculated relative to the natural species. (D) Plot of ΔG_b vs the electric dipole moment, μ, of the amino acid side chain at position 3. A positive sign of μ stands for a negative end point away from the phenyl ring, toward the *p*-X substituent. The μ value for Tyr has been corrected for the absence of resonance effects (see Table 1). A detailed description of the procedures used to obtain estimates of the physicochemical properties of the side chains is reported in Experimental Procedures.

the seminal work of Burley and Petsko (*50*) on "weakly polar interactions", a great deal of experimental evidence has been accumulated over the years, leading to the conclusion that "nonconventional" hydrogen bonds of the type X-H···π and cation−π interactions can play a key role in protein structure and stability, as well as in ligand recognition (see ref *59* for a recent review). In this respect, theoretical calculations indicate that, among many other possible combinations of cations and aromatics, the interaction of NH_4^+ (or K^+) with indole provides the most favorable interaction energy (*60*). Recently, these predictions have been confirmed by a comprehensive structural analysis of X-H···π hydrogen bonding, performed on 592 high-resolution (≤1.6 Å) protein structures, in which the indolyl nucleus of Trp was found by far to be the most frequent π-acceptor of hydrogen and the Lys−Trp pair the most represented donor−acceptor pair (*61*).

Clearly, the results reported herein indicate that the relative change in the affinity of hirudin 1−47 analogues can be estimated by desolvation effects quite well, but on the other hand, they also emphasize the importance of "hidden interactions" in hirudin−thrombin recognition. In an attempt to unravel the physical nature of these forces, we plotted the experimental ΔG_b of hirudin analogues versus the electric dipole moment, μ, of the para-X-substituted (X is H, F, I, OH, or NO_2) amino acid side chain at position 3 reported in Table 1. Given the additive properties of electric dipole moments, the μ value of tyrosine (2.02 D) in wild-type fragment 1−47 was estimated by adding the electric dipole moment of toluene (0.36 D) to that of a single C−O bond (1.66 D) (*62*), thus assuming that the mesomeric (or resonance) contribution, *R*, of the hydroxyl group in the bound form is negligible (see Table 1). This assumption is supported by the fact that in the thrombin-bound form, the electron pairs of the phenolic oxygen atom are hydrogen bonded to two structural water molecules (W432 and W606), characterized by low *B*-factor (36 and 22 $Å^2$, respectively) and high occupancy values (0.52 and 1.00, respectively) (*11*), and thus are not available for delocalizing onto the aromatic ring of Tyr. The data shown in Figure 7D indicate that the affinity of hirudin fragment 1−47 for thrombin is inversely related ($r = 0.97$) to the value of μ. Moreover, these data suggest that the hirudin−thrombin interaction is destabilized by the presence of a partial negative charge at the para position of the phenyl ring, as in the case of *p*NO_2-Phe, where the two oxygen atoms are strongly electron dense. A reasonable explanation for these results may be found in the specific electrostatic interactions with the electronic π-system of Trp215, which for *p*-aminomethyl-Phe are attractive, whereas they are repulsive for *p*NO_2-Phe.

Structural and Dynamical Properties of Thrombin Allosteric Forms. The results reported above indicate that position 3 of hirudin is not only crucial for modulating the strength of hirudin−thrombin interaction but also responsible for the (>30-fold) preferential binding of this inhibitor to the procoagulant (fast) form of the enzyme. As shown in Table 2, the presence at position 3 of a small-sized amino acid like Ala reduces the affinity for the fast form by 65-fold,

but does not essentially affect the properties of binding of the natural fragment to the slow form of thrombin. Conversely, the replacement of Tyr3 with the bulkier βNal improves affinity preferentially for the fast form.

These data can be analyzed within the theoretical framework of site-specific thermodynamics (*63*, *64*) and used to extract structural information about the S3 site of thrombin in the slow and fast forms. According to the following thermodynamic cycle (*29*)

$$\begin{array}{ccc} & \Delta G^{s} & \\ S & \Leftrightarrow & SI \\ \Delta G^{0} \Updownarrow & & \Updownarrow \Delta G^{1} \\ F & \Leftrightarrow & FI \\ & \Delta G^{f} & \end{array}$$

the slow (S) and fast (F) forms bind the inhibitor (I) with standard free energy changes ΔG^{s} and ΔG^{f}, while $\Delta G°$ and ΔG^{1} represent the free energy changes for switching from the slow to the fast form in the absence or presence of the inhibitor. The coupling free energy (ΔG_{c}) for the cycle is given by the equation $\Delta G_{c} = \Delta G^{f} - \Delta G^{s} = \Delta G^{1} - \Delta G°$, where ΔG^{f} and ΔG^{s} can be determined experimentally. The guiding idea is that structural information about the S3 site of thrombin in the two allosteric forms can be gained from the physicochemical properties of the mutated residue at position 3 of hirudin and from the effects of the introduced perturbations on the value of ΔG_{c}. The preferential loss (or gain) in affinity of a mutated inhibitor for one of the two allosteric forms of the enzyme is a measure of the energetic contribution of the interactions being lost (or gained) upon mutation and provides strong, albeit indirect, means for identifying those regions on thrombin that have different structural features in the slow or fast form. On the other hand, mutations that affect both forms to the same extent reveal that the perturbations introduced into the inhibitor are important for binding to thrombin and that the site is not involved in the slow → fast transition or, otherwise, that the entity of the introduced perturbation is too small to elicit a different behavior at that site in the two allosteric forms. In addition, since Na^{+} binds the hirudin−thrombin complex with ~20-fold higher affinity than the free enzyme ($\Delta G° \neq \Delta G^{1}$) (*29*), then from the linkage principles mentioned above it follows that the inhibitor must bind with different affinities to the slow and fast forms ($\Delta G^{f} \neq \Delta G^{s}$), ruling out other possibilities.

In the case of Tyr3 → Ala exchange, elimination of the interactions of the Tyr side chain beyond Cβ strongly reduces the affinity for the fast form, while it is practically ineffective on the slow form, suggesting that the presence of Tyr is crucial to enhancing binding exclusively to the procoagulant fast form. On the other hand, the presence of the larger side chain of βNal at position 3 enhances the affinity for the fast form by ~40-fold and by only 15-fold for the slow form. Taken together, these results suggest that the S3 site of thrombin in the procoagulant (fast) form is in a more open and accessible conformation with respect to the less forgiving structure it acquires in the anticoagulant slow form. This conclusion is consistent with the results of molecular dynamics simulation conducted for 18 ns in explicit water, showing that after ~5 ns a conformational transition occurs, leading thrombin from a more open, permissible conformation (possibly related to the fast form) to a more compact, closed one (possibly related to the slow form). In this respect, the putative slow form exposes to water a surface area that is ~230 Å^2 smaller than that of the fast form, and the total volume of pockets and cavities in the anticoagulant (slow) form is reduced by ~170 Å^3. In particular, our data indicate that the S3 site of the enzyme is significantly less accessible to water (from 150 to 50 Å^2) in the slow form, in agreement with the fact that large side chains (e.g., Tyr and βNal) at position 3 enhance affinity preferentially for the fast form, whereas small side chains (e.g., Ala) actually abolish ΔG_{c}, without affecting the affinity for the slow form (Table 2).

The structural picture of the thrombin allosteric forms proposed above may also provide a reasonable explanation for the fact that those substrates that are related to the procoagulant activities of thrombin (i.e., fibrinogen, PAR-1, and factor XIII) orient a bulky side chain deep into the S3 site of the enzyme and, as expected, are cleaved by the fast form of thrombin 20−40-fold more efficiently than by the slow form (*4*). High-resolution structures of thrombin complexed to peptides encompassing the cleavable sites of the corresponding macromolecular targets reveal that fibrinogen interacts at the S3 site of the enzyme through the bulky Phe8 having a side chain volume of 127 Å^3 [PDB entry 1bbr (*65*)], PAR-1 through Leu38 (Leu, 100 Å^3) [PDB entry 1nrs (*66*)], and factor XIII through two Val residues (Val29 and Val34; Val, 79 Å^3) [PDB entry 1de7 (*70*)]. On the other hand, protein C, which is related to the anticoagulant function of thrombin, does not extensively interact with the S3 site of the enzyme, and it is cleaved with similar specificity by either the slow or fast form ($\Delta G_{c} = 0.2$ kcal/mol) (*4*). In fact, structural differences possibly existing at this site in the two allosteric forms are expected to be inconsequential in protein C recognition and hydrolysis by thrombin. In this regard, several lines of evidence indicate that the S3 site of thrombin is not significantly involved in binding of protein C. First, given the sequence similarity of PAR-1 (TL^{38}DPR^{41}↓SFLL) and protein C (QV^{166}DPR^{169}↓LIDG) at the cleavage site (↓), it can be inferred from model building studies (not shown) that Val166 points to, but does not fill, the S3 site. Second, enlargement of the S3 pocket of thrombin by Trp215 → Ala exchange (*68*) reduces the specificity for fibrinogen and D-Phe-Pro-Arg-*p*NA by 500- and 100-fold, respectively, while the specificity for protein C drops by only 3-fold. Third, Asp-Pro-Arg-*p*NA, lacking Val at P4, is still a good substrate for thrombin (*4*).

CONCLUSIONS

It is widely accepted that even the knowledge of the structure of a ligand−receptor complex provides only partial information for predicting ligand binding energetics (*69*). To possibly tackle this problem in this study, we have attempted to transfer the quantitative structure−activity relationship (QSAR) approach, traditionally applied to low-molecular weight bioactive compounds, to the study of recognition in macromolecular systems, such as the hirudin−thrombin complex. Central to our work was the possibility to systematically and finely tune the properties of hirudin N-terminal fragment 1−47 at a desired position through incorporation of noncoded amino acids with tailored side chains. The results reported in this study provide unprecedented information about the hidden interactions involved in hirudin−

thrombin recognition and the structural features responsible for the 30-fold preferential binding of hirudin for the fast form of thrombin over the slow form. These findings also agree with a picture of thrombin inferred from MD simulation, showing that the fast form of the enzyme is in a more open and accessible conformation than the slow form, being in a more compact and less permissible conformation. Given the importance of the slow → fast transition in controlling thrombin function in hemostasis (*4*) and the fact that the crystallographic structure of thrombin in the slow form is not yet available, our findings can help in shedding light on the interaction of thrombin with its numerous natural substrates and can have important consequences in the design of novel thrombin inhibitors.

ACKNOWLEDGMENT

The support and helpful suggestions of Dr. Giacomo Carrea and Dr. Alan Mark are gratefully acknowledged. V.D.F. also gratefully thanks Dr. Daniele Dalzoppo for critically reading the manuscript, Dr. Stefano Moro for calculating amino acid side chain volumes, and Prof. Enrico Di Cera for insightful comments.

REFERENCES

1. Davie, E. W., Fijikawa, K. A., and Kisiel, W. (1991) *Biochemistry 30*, 10363−10370.
2. Sanderson, P. E., and Naylor-Olsen, A. M. (1998) *Curr. Med. Chem. 5*, 289−304.
3. Esmon, C. T. (1995) *FASEB J. 9*, 946−955.
4. Di Cera, E., Dang, Q. D., and Ayala, Y. M. (1997) *Cell. Mol. Life Sci. 53*, 701−730.
5. Berg, D. T., Wiley, M. R., and Grinnell, B. W. (1996) *Science 273*, 1389−1391.
6. Braun, P. J., Dennis, S., Hofsteenge, J., and Stone, S. R. (1988) *Biochemistry 27*, 6517−6522.
7. Vindigni, A., De Filippis, V., Zanotti, G., Visco, C., Orsini, G., and Fontana, A. (1994) *Eur. J. Biochem. 226*, 323−333.
8. Folkers, P. J. M., Clore, G. M., Driscoll, P. C., Dodt, J., Kohler, S., and Gronenborn, A. M. (1989) *Biochemistry 28*, 2601−2617.
9. Szyperski, T., Güntert, P., Stone, S. R., and Wütrich, K. (1992) *J. Mol. Biol. 228*, 1193−1205.
10. Nicastro, G., Baumer, L., Bolis, G., and Tatò, M. (1997) *Biopolymers 41*, 731−749.
11. Rydel, T. J., Tulinsky, A., Bode, W., and Huber, R. (1991) *J. Mol. Biol. 221*, 583−601.
12. De Filippis, V., Quarzago, D., Vindigni, A., Di Cera, E., and Fontana, A. (1998) *Biochemistry 37*, 13507−13515.
13. De Filippis, V., Russo, I., Vindigni, A., Di Cera, E., Salmaso, S., and Fontana, A. (1999) *Protein Sci. 8*, 2213−2217.
14. Matthews, B. W. (1995) *Adv. Protein Chem. 46*, 249−278.
15. Fersht, A. R. (1997) *Curr. Opin. Struct. Biol. 15*, 79−84.
16. Wells, J. A. (1994) *Curr. Opin. Cell Biol. 6*, 163−173.
17. Mendel, D., Cornish, V. W., and Schultz, P. G. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 435−462.
18. Fauchère, J.-L. (1986) *Adv. Drug Res. 15*, 29−69.
19. Giannis, A., and Rübsam, F. (1997) *Adv. Drug Res. 29*, 1−78.
20. Hogan, J. C., Jr. (1997) *Nat. Biotechnol. 15*, 328−330.
21. De Filippis, V., De Antoni, F., Frigo, M., Polverino de Laureto, P., and Fontana, A. (1998) *Biochemistry 37*, 1686−1696.
22. De Filippis, V., Vindigni, A., Altichieri, L., and Fontana, A. (1995) *Biochemistry 34*, 9552−9564.
23. Ten Kortenaar, P. B. W., Van Dijk, B. G., Peeters, M. J., Raabe, B. J., Adams, P. J., and Tesser, G. I. (1986) *Int. J. Pept. Protein Res. 27*, 398−400.
24. Wang, S. S. (1973) *J. Am. Chem. Soc. 95*, 1328−1333.
25. Chatrenet, B., and Chang, J.-Y. (1992) *J. Biol. Chem. 267*, 3038−3043.
26. Vindigni, A., and Di Cera, E. (1998) *Protein Sci. 7*, 1728−1733.
27. Gill, S. G., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
28. Buckler, D. R., Haas, E., and Scheraga, H. A. (1993) *Anal. Biochem. 209*, 20−31.
29. Ayala, Y. M., and Di Cera, E. (1994) *J. Mol. Biol. 235*, 733−746.
30. Ayala, Y. M., Vindigni, A., Nayal, M., Spolar, R. S., Record, M. T., Jr., and Di Cera, E. (1995) *J. Mol. Biol. 253*, 787−798.
31. Horton, N., and Lewis, M. (1992) *Protein Sci. 1*, 169−181.
32. Juffer, A. H., Eisenhaber, F., Hubbard, S. J., Walther, D., and Argos, P. (1995) *Protein Sci. 4*, 2499−2509.
33. Eisenberg, D., and McLachlan, A. D. (1986) *Nature 319*, 199−203.
34. Fauchère, J.-L., Charton, M., Kier, L. B., Verloop, A., and Pliska, V. (1988) *Int. J. Pept. Protein Res. 32*, 269−278.
35. Sangster, J. (1989) *J. Phys. Chem. Ref. Data 18*, 1111−1229.
36. Buchwald, P., and Bodor, N. (1998) *Curr. Med. Chem. 5*, 353−380.
37. Gerstein, M. (1992) *Acta Crystallogr. A48*, 271−276.
38. Liang, J., Edlsbrunner, H., and Woodward, C. (1998) *Protein Sci. 7*, 1884−1897.
39. Lien, E. J., Guo, Z.-R., Li, R.-L., and Su, C.-T. (1982) *J. Pharm. Sci. 71*, 641−655.
40. Bode, W., Turk, D., and Karshikov, A. (1992) *Protein Sci. 1*, 426−472.
41. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Di Nola, A., and Haak, J. R. (1984) *J. Chem. Phys. 81*, 3684.
42. van Gunsteren, W. F., Daura, X., and Mark, A. E. (1998) *Encycl. Comput. Chem. 2*, 1211−1216.
43. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. (1981) in *Intermolecular Forces* (Pullman, B., Ed.) pp 331−342, Reidel, Dordrecht, The Netherlands.
44. Hess, B., Bekker, H., Fraaije, J., and Berendsen, H. J. C. (1997) *J. Comput. Chem. 18*, 1463−1472.
45. Lindahl, E., Hess, B., and van der Spoel, D. (2001) *J. Mol. Model. 7*, 306−317.
46. Amadei, A., Linssen, A. B., and Berendsen, H. J. (1993) *Proteins: Struct., Funct., Genet. 17*, 412−425.
47. Sisido, M., Egusa, S., and Imanishi, Y. (1983) *J. Am. Chem. Soc. 105*, 4077−4082.
48. Kahn, P. C. (1979) *Methods Enzymol. 61*, 339−378.
49. Otto, A., and Seckler, R. (1991) *Eur. J. Biochem. 202*, 67−73.
50. Burley, S. K., and Petsko, G. A. (1988) *Adv. Protein Chem. 39*, 125−189.
51. Hansch, C., Leo, A., Unger, S. H., Kim, K. H., Nikaitani, D., and Lien, E. J. (1973) *J. Med. Chem. 16*, 1207.
52. Berliner, L. J., and Shen, Y. Y. L. (1977) *Biochemistry 16*, 4622−4626.
53. Gilson, M. K., Rashin, A., Fine, R., and Honig, B. (1985) *J. Mol. Biol. 184*, 503−516.
54. Dao-Pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W., and Matthews, B. W. (1991) *Biochemistry 30*, 115121−11529.
55. Di Cera, E., Guinto, E. R., Vindigni, A., Dang, Q. D., Ayala, Y. M., Wuyi, M., and Tulinsky, A. (1995) *J. Biol. Chem. 270*, 22089−22092.
56. Kubinyi, H. (1993) in *QSAR: Hansch analysis and related approaches* (Mannhold, R., Krogsgaard-Larsen, P., and Timmerman, H., Eds.) Vol. 1, VCH Publishers, New York.
57. Purcell, A. W., Aguilar, M.-I., and Hearn, M. T. W. (1999) *Anal. Chem. 71*, 2440−2451.
58. Legon, A. C. (1999) *Angew. Chem., Int. Ed. 38*, 2686−2698.
59. Kim, K. S., Tarakeshwar, P., and Lee, J. Y. (2000) *Chem. Rev. 100*, 4145−4185.
60. Mecozzi, S., West, A. P., and Dougherty, D. A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10566−10571.
61. Steiner, T., and Koellner, G. (2001) *J. Mol. Biol. 305*, 535−557.
62. Li, W.-Y., Guo, Z.-R., and Lien, E. J. (1984) *J. Pharm. Sci. 73*, 553−558.
63. Wyman, J. (1968) *Q. Rev. Biophys. 1*, 35−80.
64. Di Cera, E. (1998) *Adv. Protein Chem. 51*, 59−119.
65. Stubbs, M. T., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S. R., and Bode, W. (1992) *Eur. J. Biochem. 206*, 187−195.
66. Mathews, I. I., Padmanabhan, K. P., Ganesh, V., Tulinsky, A., Ishii, M., Chen, J., Turck, C. W., Coughlin, S. R., and Fenton, J. W., II (1994) *Biochemistry 33*, 3266−3279.
67. Sadasivan, C., and Yee, V. C. (2000) *J. Biol. Chem. 275*, 36942−36948.

68. Arosio, D., Ayala, Y. M., and Di Cera, E. (2000) *Biochemistry 39*, 8095−8101.
69. Engh, R. A., Brandstetter, H., Sucher, G., Eichinger, A., Bauman, U., Bode, W., Huber, R., Poll, T., Rudolph, R., and von der Saal, W. (1996) *Structure 4*, 1353−1362.
70. Scacheri, E., Nitti, G., Valsasina, B., Orsini, G., Visco, C., Ferreia, M., Sawyer, R. T., and Sarmientos, P. (1993) *Eur. J. Biochem. 214*, 295−304.

BI0203482