J Mol Model (2010) 16:725–735
DOI 10.1007/s00894-009-0590-2

ORIGINAL PAPER

Thrombin allosteric modulation revisited: a molecular dynamics study

Hermes Luís Neubauer de Amorim · Paulo Augusto Netz · Jorge Almeida Guimarães

Received: 24 June 2009 / Accepted: 2 September 2009 / Published online: 9 October 2009
© Springer-Verlag 2009

Abstract The regulatory properties of thrombin are derived predominantly from its capacity to produce different functional conformations. Functional studies have revealed that two antagonistic thrombin conformations exist in equilibrium: the fast (procoagulant) and slow (anticoagulant) forms. The mechanisms whereby thrombin activity is regulated by the binding of different effectors remain among the most enigmatic and controversial subjects in the field of protein function. In order to obtain more detailed information on the dynamic events originating from the interaction with the Na^+ effector and ligand binding at the active site and anion binding exosite 1 (ABE1), we carried out molecular dynamics simulations of thrombin in different bound states. The results indicated that Na^+ release results in a more closed conformation of thrombin, which can be compared to the slow form. The conformational changes induced by displacement of the sodium ion from the Na-binding site include: (1) distortion of the 220- and 186-loops that constitute the Na-binding site; (2) folding back of the Trp148 loop towards the body of the protein, (3) a 180° rotation of the Asp189 side-chain, and (4) projection of the Trp60D loop toward the solvent accompanied by the rearrangement of the Trp215 side chain toward the 95–100 loop. Our findings correlate well with the known structural and recognition properties of the slow and fast forms of thrombin, and are in accordance with the hypothesis that there is communication between the diverse functional domains of thrombin. The theoretical models generated from our MD simulations complement and advance the structural information currently available, leading to a more detailed understanding of thrombin structure and function.

Keywords Thrombin · Allosterism · Specificity · Molecular dynamics

Abbreviations

ABE1	Anion binding exosite 1
PC	Protein C
PAR1	Protease-activated receptor 1
PDB	Protein data bank
MD	Molecular dynamics
RMSD	Root mean square deviation
Rg	Radius of gyration
NHb	Total number of hydrogen bonds
SASA	Solvent-accessible surface area
RMSF	Atom-positional root mean square fluctuation

Introduction

Thrombin generation, initiated by cellular and vascular damage, is the central biochemical reaction in the processes

H. L. N. de Amorim (✉)
Laboratório de Bioinformática Estrutural (LaBiE),
Universidade Luterana do Brasil (ULBRA),
Sala 122, Prédio 01, Av. Farroupilha 1001,
Canoas, RS, Brazil 92450-900
e-mail: hamorim@cbiot.ufrgs.br

P. A. Netz
Departamento de Físico-Química, Instituto de Química,
Universidade Federal do Rio Grande do Sul,
Av. Bento Gonçalves 9500,
CEP 91501-970 Porto Alegre, RS, Brazil

J. A. Guimarães
Laboratório de Bioquímica Farmacológica,
Centro de Biotecnologia,
Universidade Federal do Rio Grande do Sul (UFRGS),
Lab 202, Bloco IV, Prédio 43421, Av. Bento Gonçalves, 9500,
Porto Alegre, RS, Brazil P.O.Box 15.005

of normal and altered hemostasis and thrombosis [1, 2]. Thrombin is a multifunctional serine protease, unique among the clotting cascade serine proteases in its ability to interact with a large number of structurally distinct substrates [3 (and references therein), 4, 5]. The regulatory properties of thrombin are derived predominantly from its capacity to produce different functional conformations. Unlike other protease–inhibitor interactions, thrombin interacts with both inhibitors and substrates through an induced-fit mechanism of recognition rather than a rigid body association. Functional studies have revealed that thrombin exists in two types of antagonistic conformation in equilibrium, the so-called "fast" (procoagulant) and "slow" (anticoagulant) forms [6]. The procoagulant form is responsible for: (1) conversion of fibrinogen into fibrin; (2) activation of platelets through cleavage of the platelet receptor PAR1; (3) enhancement of prothrombin–thrombin conversion through feedback activation of factors V, VII and XI; and (4) activation of the transglutaminase factor XIII. On the other hand, anticoagulant form of thrombin shows higher specificity for thrombomodulin (TM), a cofactor on the membrane of endothelial cells. The thrombin–TM complex acts as an anticoagulant through activation of protein C (PC), which cleaves and inactivates the procoagulant factors Xa and IXa that are required for thrombin generation.

It was proposed that the transition between the "slow" and "fast" forms of thrombin is triggered by the binding of a sodium ion at a region of thrombin acknowledged as the Na^+ binding site [5, 7–10]. The mechanism whereby thrombin activity is regulated by the binding of the Na^+ cation remains one of the most enigmatic and controversial subjects in the field of protein function, and has been addressed in recent reviews [5, 10].

In another conformational feature, interaction of (macro) molecules at thrombin anion binding exosite 1 [ABE1 or fibrin(ogen) secondary binding exosite] is accompanied by coupled folding transitions in both thrombin and the binding molecule [11, 12]. Moreover, a relationship between the binding of inhibitors to ABE1 and the hydrolysis of several amide substrates was also demonstrated [13]. Therefore, it is conceivable that the residues contained in the active site cleft are also involved in binding of inhibitors at ABE1. Due to the interest in thrombin, and its great importance, the allosteric properties of this enzyme have been exploited by several researchers. Previous studies have addressed the role of Glu192 [14], Trp60D [15], ABE1 [16], and the presence of the sodium ion at the Na binding loop [6, 8] as major players in defining the specificity of thrombin through allosteric modulation.

Although there is considerable experimental support indicating which residues, regions and effectors are involved in the allosteric modulation of thrombin specificity, the characterization of the precise dynamics of protein atoms is far from complete due to the difficulty of measuring dynamics with atomic detail. Conclusions derived from X-ray crystallographic data, for instance, are based on a static description of protein conformation, while the use of fluorescence spectroscopy and circular dichroism provide only indirect information. In addition, several thrombin structures have been determined under different crystallographic conditions, sometimes resulting in differing interpretations [17]. Regarding crystallographic data, it is important to note that, as observed by Huntington and Esmon [18],

> "...despite the efforts to obtain crystals of the slow form of thrombin through either soaking out the Na^+ after crystallization or crystallization in the absence of Na^+, the constraints imposed upon thrombin within the crystal resulted in thrombin with its active site pocket adopting the fast conformation"

Although the Huntington group have subsequently obtained a crystallographic structure of wild-type human thrombin in the Na^+-free state [19], which could be used to explain many biochemical properties of the slow form, it must be remembered that crystallographic models are generated through conditions in which crystal contacts and other crystallization artifacts may be present. As a consequence, the crystallographic structures of Na^+-free thrombin should be interpreted with caution.

In this case, limitations due to experimental methods can be circumvented by the use of computational tools, such as molecular dynamics (MD) simulations, which have proved useful in describing conformational events in proteins [20–24]. The main advantage of this technique is that it can model aspects of protein structure in solution.

A preliminary survey of conformational events in thrombin was described in the work of De Filippis and co-workers [25], who conducted a MD simulation of free thrombin in explicit water. The simulation showed that, after ~5 ns, thrombin undergoes a conformational transition from a more open conformation (proposed to be related to the fast form) to a more compact and closed conformation (proposed to be related to the slow form). Even so, the model obtained was not capable of correlating several structural properties subsequently discovered, such as the changes in fluorescence (associated with the movement of tryptophan residues) induced by Na^+ binding [26].

In order to contribute to increased knowledge of thrombin allosteric modulation, in this work we present the results of four MD simulations of the following systems: (1) sodium-bound thrombin complexed with fragments related to segments 7–16 of fibrinopeptide A and 54–65 of hirudin 2; (2) sodium-bound thrombin complexed with hirudin 2 fragment; (3) sodium-bound thrombin in the absence of any molecular ligand; and (4) the simulation of thrombin in the absence of

any molecular ligands and/or sodium ions. For all these systems, there are experimental data that can be used in the analysis and validation of our simulation results.

The main differences between the thrombin-averaged structures generated from the last 500 ps of the four simulations performed in this work are located in areas known to be key regions for protein function. More specifically, important changes were observed in (1) the conformation of Asp189 at the S_1 specificity site; (2) the position of the 60 insertion loop (which defines the S_2 specificity site); (3) the conformation of the Trp60D side-chain at the 60 insertion loop; (4) the position of Trp215 at the S_3 specificity site; and (5) the structure of the Na^+ binding site. The present study differs from earlier simulation studies of thrombin with respect to the force field used, the simulation set-up, and the high degree of correlation with crystallographic and spectroscopic results, which also allow clarification of some conflicting experimental results.

Materials and methods

Visualization and manipulation of molecular structures were made using the molecular modeling software SPDBV [27] and VMD [28]. Molecular dynamics simulations were carried out using GROMOS96 53a6 force field [29] implemented in the GROMACS package [30], version 3.3.1. The simulations are abbreviated in the text as follows: (1) THR-NaFH2 is the simulation of sodium-bound thrombin complexed with fragments related to segments 7–16 of fibrinopeptide A (DFLAEGGGVR) and 54–65 of hirudin 2 (GDFEEIPEEYLQ); (2) THR-NaH2 is the simulation of sodium-bound thrombin complexed with hirudin 2 fragment; (3) THR-Na is the simulation of sodium-bound thrombin in the absence of any molecular ligand; and (4) THR-nlig is the simulation of thrombin in the absence of any molecular ligands or ions. The starting coordinates of thrombin were taken from the 2.2 Å resolution crystal structure of alfa-thrombin ternary complex with hirudin 2 and fibrinopeptide A [entry 1FPH in the Protein Data Bank (PDB)]. In the THR-NaFH2, THR-NaH2 and THR-Na simulations, the atomic coordinates of the sodium ion were taken from the 2.3 Å resolution crystal structure of the procoagulant fast form of thrombin (PDB entry 1SG8). In this procedure, the thrombin-complex structure from 1FPH was superimposed on the thrombin structure obtained from 1S8G. The remaining coordinates in the 1S8G structure were then removed. After this procedure, the sodium ion was located 3.96 Å, 2.33 Å and 2.21 Å from the carbonyl oxygen atoms of Tyr184, Arg221 and Lys224, respectively. All systems were simulated in NPT ensemble and periodic boundary conditions (truncated octahedron). The dimensions of the central box were chosen in such way that the minimum distance of any protein atom to the closest box wall was 12 Å.

The simulations were carried out using explicit solvent water molecules described by the simple point charge extended (SPC/E) model [31]. Initially, the protein structure in each system was submitted to a maximum of 500 steps of steepest descent energy minimization. To relax strong solvent–solvent and solvent–protein non-bonded interactions, 1 ps MD simulation was performed restraining the protein structure. Initial velocities were assigned according to Maxwell distribution. All simulations were performed for 10 ns using an integration time step of 2 fs. Each system was heated with gradual increments in the following temperatures: 100 K (10 ps), 150 K (5 ps), 200 K (5 ps), and 250 K (5 ps). Thereafter, the temperatures of the systems were adjusted to 310 K. The first 5 ns of each simulation was considered as part of the heating (0.025 ns) and the equilibration (4.975 ns) steps and were not used in the data analysis. The temperatures of solvent and solutes (protein and sodium ion) were independently coupled to a thermal bath with a relaxation time of 0.1 ps using a Berendsen thermostat [32]. The pressure in the systems was weakly coupled to a pressure bath of 1 atm by applying an isotropic scaling and 0.5 ps relaxation time using the Parrinello-Rahman barostat [33, 34]. Bond lengths were constrained using the LINCS algorithm [35] with 4th order expansion. Electrostatic interactions among non-ligand atoms were evaluated by the particle mesh Ewald (PME) method [36] with a charge grid spacing of approximately 1.0 Å. The charge grid was interpolated on a cubic grid with the direct sum tolerance set to 4.0×10^{-6}. Lennard-Jones interactions were evaluated using a 10 Å atom-based cutoff. The pair list was updated every 10 steps.

Results and discussion

Global structural parameters

The convergence of calculated properties, and consequently the reliability of simulations, was checked by monitoring thermodynamic and structural parameters. The ratio of root mean square deviation (RMSD) fluctuations in the total energy to those in the kinetic energy was found to be less than 1%, which is satisfactory for this type of calculation [37]. A check was made for systematic drift in the total energy over time by examining the averages and RMSD fluctuations of the total energy in time intervals over 500-ps-long sections of the total run (data not shown). From this analysis, no significant gradients were found.

A better picture of stability and convergence of MD simulations can be obtained from the analysis of the time

dependence of the following parameters: the atomic positional RMSD with respect to the initial crystallographic structure, the radius of gyration (Rg), the total number of hydrogen bonds (NHb), and the total solvent-accessible surface area (SASA). These dynamic properties are shown in Fig. 1 and the average values calculated from the last 5 ns of the simulations are presented in Table 1.

Least-squares fitting of atomic coordinates for the calculation of atom position RMSD of thrombin was based on all atoms (Fig. 1a). There is a slight tendency for RMSD to increase as the number of bound ligands decreases. In Table 1, the average RMSD in the last 5 ns of the simulations ranges from a minimum of 0.32 nm for THR-NaFH2 to a maximum of 0.37 for THR-nlig. The highest degree of deviation observed in the simulation of unbound thrombin (THR-nlig) stems from the fact that the crystallographic starting structure was resolved complexed with two ligands—fibrinopeptide A and hirudin 2 fragment—and in the presence of Na^+ ions.

The Rg (Fig. 1b), which provides a rough measure of the compactness of a structure, shows an increase with time, indicating a slight expansion of thrombin conformation in all simulations. Regardless of the number of ligands bound to thrombin, the average values of Rg were the same for the last 5 ns of all simulations, ranging from 1.82 nm^3 at the beginning (starting structure) to an average of 1.85 nm^3 in the final period of the simulations. The increase in Rg is related to thrombin relaxation in water, which leads to projection of the residues located on the protein surface toward the solvent—an effect mediated by the establishment of protein–solvent hydrogen bonds–and due to the new temperature condition (310 K).

Hydrogen bonds (Fig. 1c) were calculated using a geometric criterion [28]. To detect a hydrogen bond, the hydrogen–donor–acceptor angle must have a maximum value of 30° and the donor–acceptor distance must have a maximum value of 0.35 nm. Considering the standard deviation, the average number of hydrogen bonds is almost the same for all systems in the last 5 ns of simulations (Table 1). As observed in Fig. 1c, the number of intramolecular hydrogen bonds decreases with simulation time, ranging from 242 at the beginning of the simulations to a minimum of 202 in the THR-NaFH2 and THR-nlig simulations. In fact, intramolecular hydrogen bonds are replaced

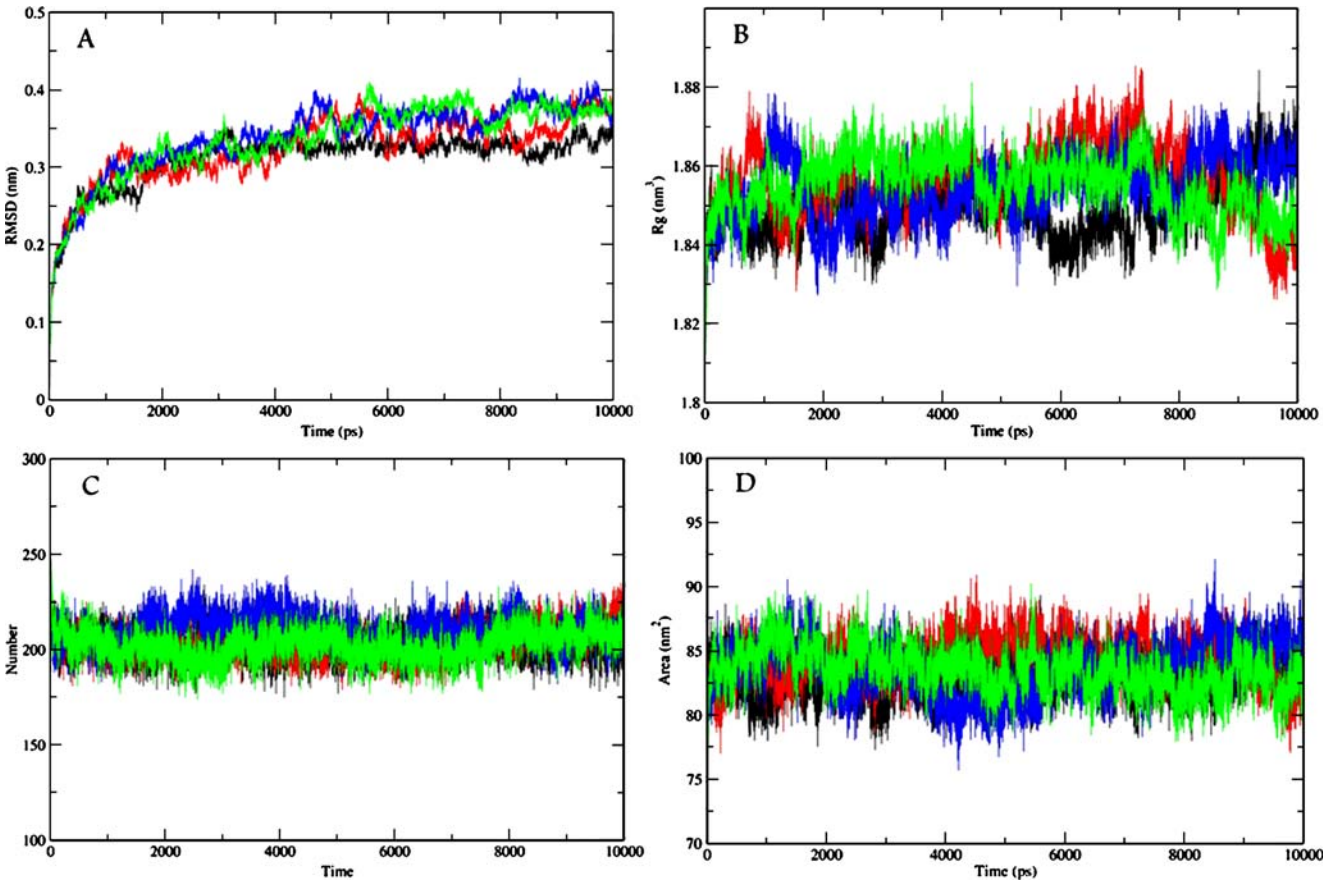


Fig. 1 **a**–**d** Time-dependence of structural parameters from molecular dynamics (MD) simulations with respect to the initial X-ray structure. **a** Atom-positional root mean-square deviation (RMSD) of all atoms. **b** Radius of gyration (Rg). **c** Total number of hydrogen bonds (NHb). **d** Total solvent-accessible surface area (SASA). *Black* THR-NaFH2, *red* THR-NaH2, *blue* THR-Na, *green* THR-nlig

Table 1 Average structural parameters of thrombin from the last 5 ns of molecular dynamics (MD) simulations. *RMSD* Root mean square deviation, *Rg* radius of gyration, *NHb-intra* number of hydrogen bonds (intramolecular), *SASA* solvent-accessible surface area

Simulation	RMSD (nm)[a]	Rg (nm^3)	NHb-intra	SASA (nm^2)
Crystallographic[b]	0	1.82	242	77.0
THR-NaFH2	0.32±0.00	1.85±0.00	202±6.9	83.2±1.5
THR-NaH2	0.35±0.01	1.85±0.01	205±8.4	84.5±1.7
THR-Na	0.35±0.01	1.85±0.00	208±8.0	86.7±2.1
THR-nlig	0.37±0.01	1.85±0.01	202±8.2	82.8±1.5

[a] RMSD of all thrombin atoms from starting coordinates [crystallographic structure: Protein Data Bank (PDB) accession code 1FPH]

[b] Values calculated for starting coordinates

gradually by hydrogen bonds with the solvent. This, in turn, is correlated with an increase in both the Rg (Fig. 1b) and SASA of thrombin (Fig. 1d).

As shown in Table 1, the SASA trajectory increased from 77 nm^2 at the beginning of the simulations to average values of 82.8±1.5 nm^2 (THR-nlig), 83.2±1.5 nm^2 (THR-NaFH2), 84.5±1.7 nm^2 (THR-NaH2) and 86.7±2.1 nm^2 (THR-Na). These results can be interpreted in the light of experimental data assuming that the more closed conformation (lower SASA) in the THR-nlig simulation resembles the slow form, while the more open conformation (largest SASA) in the THR-Na simulation resembles the fast form. On the other hand, as expected for proteins that interact through mechanisms involving a high degree of plasticity, binding can be associated with the partial involvement of the ligand by protein amino acid side-chains, as reflected in the intermediate SASA values obtained for THR-NaFH2 and THR-NaH2 simulations.

The same trend of stability found associated with RMSD, Rg, and NHb was observed for secondary structure elements. The secondary structure assignment of thrombin for the set of structures extracted every 100 ps from the simulation trajectories was done according to the DSSP rules proposed by Kabsch and Sander [38]. According to the high degree of conservation of secondary structures (Fig. 2, a1–d1) and the atom-positional RMS fluctuation (RMSF) for each residue (Fig. 2, a2–d2) during the simulations, it is clear that interaction of thrombin with effectors and ligands does not cause any drastic global changes, either in protein structure or in protein flexibility. Analysis of RMSF reveals that important regions for thrombin function and allostery, more specifically those defined by the Trp148 loop, Trp60 loop (S2 site), 220-loop and 186-loop (both involved in nesting of the sodium ion), show a similar degree of flexibility in the four simulations.

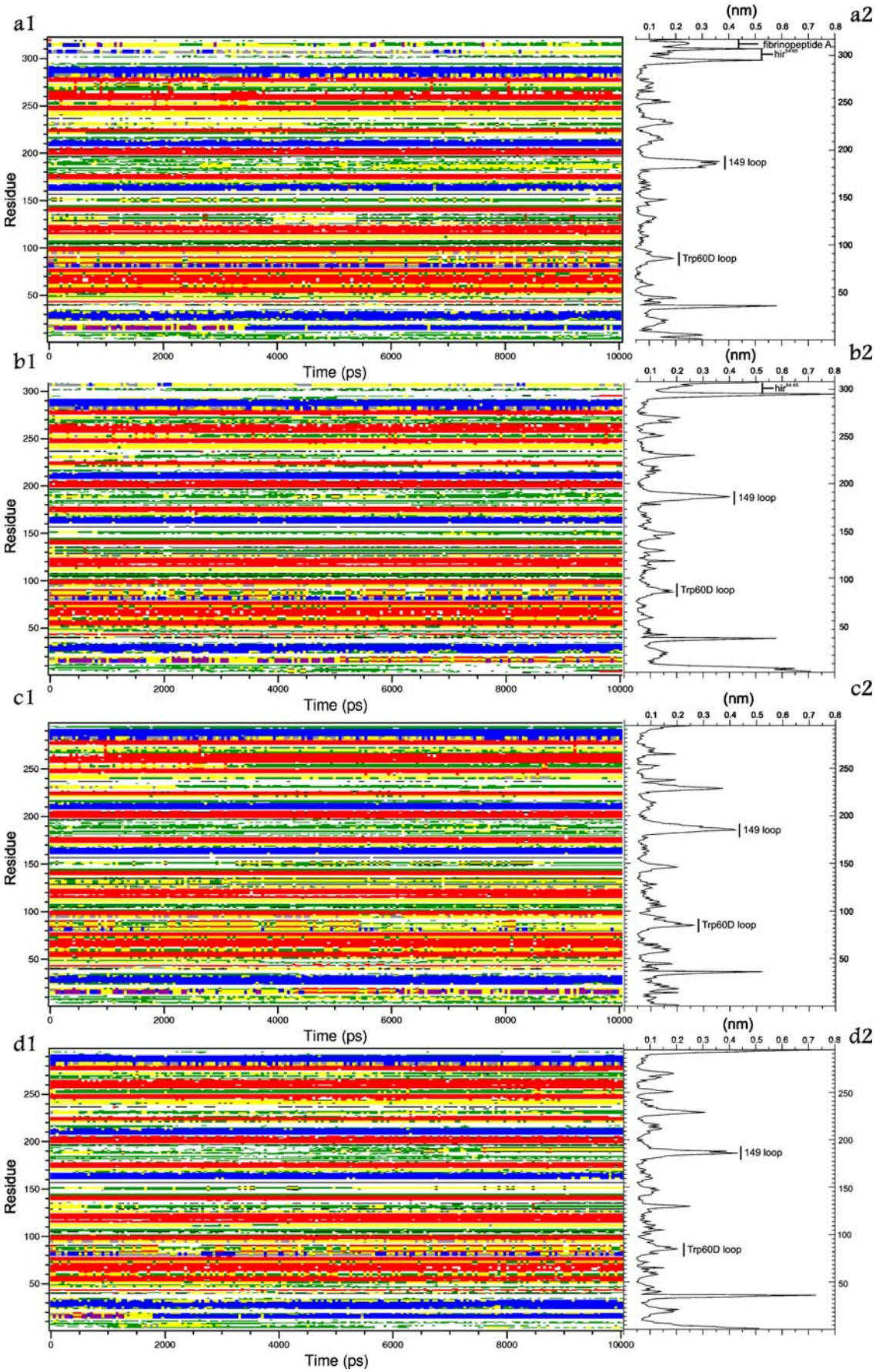
Despite the conservation of structural parameters, during the simulations significant conformational changes in important regions may be observed from visual inspection of time-averaged structures generated from the last 500 ps of each simulation. As will be discussed below, the implications of the observed structural modifications are in excellent agreement with the experimental data available for thrombin.

Na^+ binding site

The thrombin Na^+ binding site is contained within a cylindrical cavity formed by three antiparallel β-strands of the B-chain (Met180–Tyr184a, Lys224–Tyr228, and Val213–Gly219) diagonally crossed by the Glu188–Glu192 strand [6]. Unlike what was observed in a previous simulation of free thrombin [25], the thrombin Arg187–Asp222 ion-pair, which connects the 220- and 186-loops, is disrupted in the THR-nlig simulation at approximately 400 ps of simulation. As a consequence, the 220 loop slid laterally, folding back toward the Trp148 loop. Our data indicate that the site is not spontaneously prearranged for Na^+ binding, meaning that the binding of the metal ion can be entropically expensive (i.e., requires an energy cost of conformational reorganization to receive the sodium ion). These results are in agreement with the spectroscopic data obtained by De Filippis and colleagues [39]. Using far- and near-UV techniques, these latter authors observed signal changes upon Na^+-binding which, in turn, can be interpreted as an alteration in the environment of clusters of aromatic residues and/or disulfide bonds. The correlation between changes in position of the tryptophan side-chains and the type of ligand/effector associated to thrombin is discussed below.

Trp148 loop (autolysis loop)

Figure 3 presents representations of the molecular surface (solvent accessible) of the average thrombin structures generated from the last 500 ps of each simulation, color-coded according to electrostatic potential and key residues. The presence of the hirudin 2 fragment at ABE1 (THR-NaFH2 and THR-NaH2 simulations) stabilizes the average position of the Trp148 loop near the crystallographic conformation (Fig. 3, a1–a2 and b1–b2). In the absence of the ligands but with the presence of a sodium ion (THR-Na


a1
a2
b1
b2
c1
c2
d1
d2
Residue
Time (ps)
(nm)
fibrinopeptide A
149 loop
Trp60D loop

◀ **Fig. 2** *Left panels* Secondary structure of thrombin as a function of time in the simulations THR-NaFH2 (*a1*), THR-NaH2 (*b1*), THR-Na (*c1*), and THR-nlig (*d1*). *Red* α-Helix, *yellow* 3_{10}-helix, *orange* π-helix, *blue* bend, *green* extended β-strand, *violet* turn, according to the definitions of secondary structure by Kabsch and Sander [38]. *Right panels* Atom positional root-mean-square fluctuation (RMSF) of the Cα atoms as derived from the simulations THR-NaFH2 (*a2*), THR-NaH2 (*b2*), THR-Na (*c2*), and THR-nlig (*d2*)

simulation), the Trp148 loop adopts a conformation more exposed to the solvent (Fig. 3, c1–c2), while in the absence of sodium ion and ligands (THR-nlig simulation) the Trp148 loop folds back an exposed region towards the body of the protein (Fig. 3, d1–d2). In this new position, the Trp148 side-chain is positioned in front of Glu192, leading to narrowing of the cleft that allows communication between ABE1 and the active site. Additionally, according to De Filippis et al. [25], this conformational transition may restrict 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

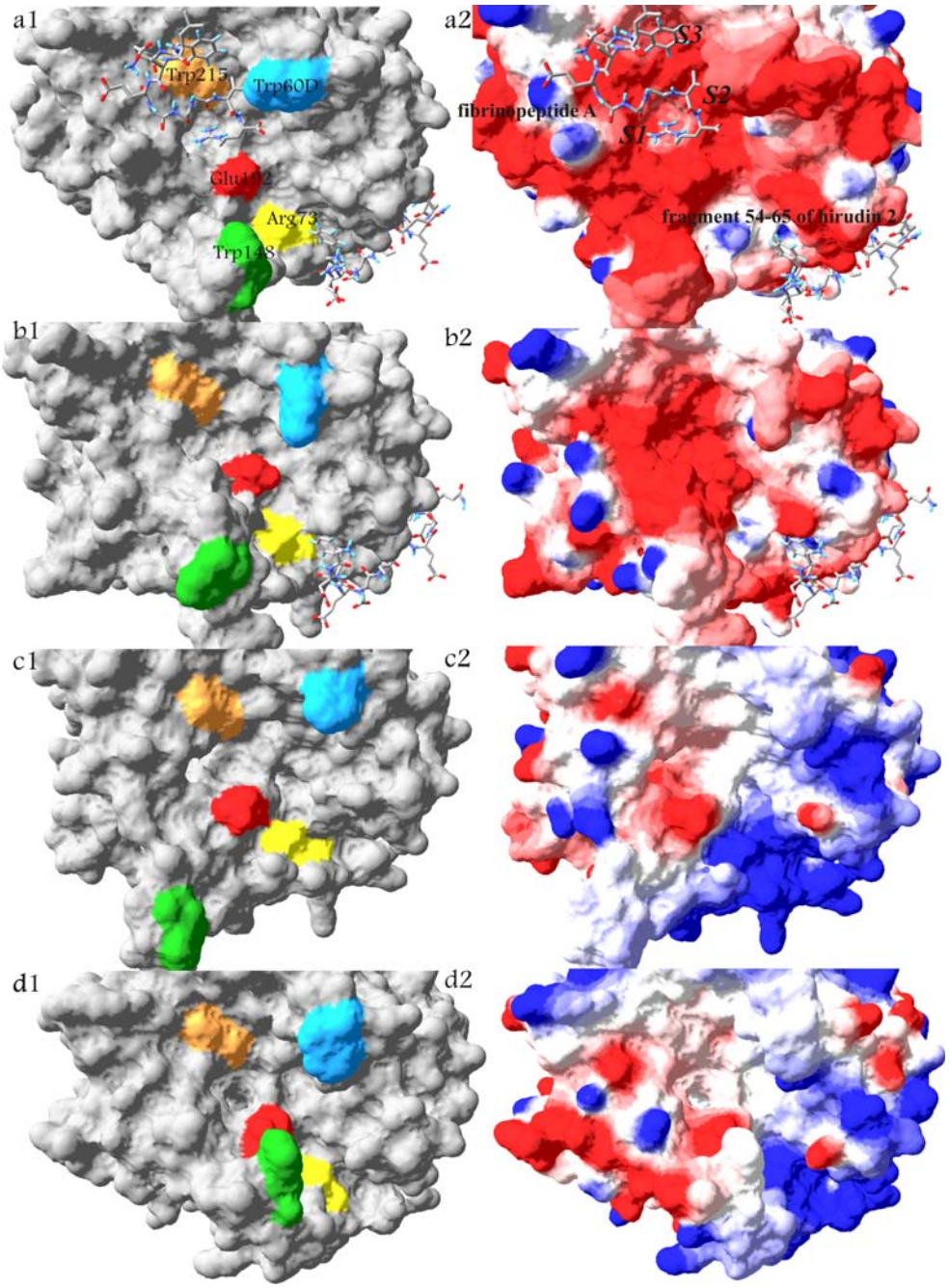


Fig. 3 Surface representations of the time-average structures of thrombin obtained from the last 500 ps of MD simulations. *Left panels* Surfaces colored according to selected residues (*blue* Trp60D, *yellow* Arg73, *green* Trp148, *red* Glu192, *orange* Trp215). *Right panels* Surfaces colored by electrostatic potential (*blue* positive, *red* negative). Simulations: *a1*, *a2* THR-NaFH2; *b1*, *b2* THR-NaH2; *c1*, *c2* THR-Na; *d1*, *d2* THR-nlig. *S1*, *S2* and *S3* indicate regions of specificity subsites

crevice that leads to the catalytic pocket. In fact, in the average structure of the THR-nlig simulation the guanidinium group of Arg73 is projected and exposed on the surface of the crevice, remaining 1.80 Å away from the same group in the THR-Na simulation which, in turn, is partially occluded. Compared with the initial coordinates and with the crystallographic fast form of thrombin, the guanidinium group of Arg73 in the THR-nlig simulation suffered a 160° rotation in the χ_3 angle (Fig. 4). In this case, reorientation of the Arg73 side-chain is associated with a displacement of 2.26 Å in the Glu39 side-chain, which leaves the carboxyl oxygen closer to the NH2 group of Arg73, contributing to the narrowing of the cleft that moves ABE1 closer to the active site.

Although the Trp148 loop is characterized by an intrinsically high conformational flexibility, it was only in the THR-nlig simulation that this specific region folded onto the protein surface, which suggests the possible participation of the Trp148 loop in the fast→slow transition. The results of these simulations are in perfect agreement with the previous assumption that Na^+ binding induces the transition of the Trp148 loop from a more closed and flexible conformation (i.e., the slow form) to a more open and rigid one (i.e., the fast form) [39].

S_1 specificity site

Thrombin prefers to cleave the peptide bond after positively charged lysine or arginine side chains due to a negatively charged aspartic acid (Asp189) located at the base of the S1 pocket. In the thrombin average structure of the THR-nlig simulation, the carboxylate of Asp189 suffered a 180° rotation, which leaves atom Oδ1 quite distant from its optimal coupling with the guanidinium group of Arg at the P1 position of the substrate (Fig. 4). On the other hand, in the THR-Na simulation, Asp189 remained in a position close to that found in the crystallographic starting structure [40] (entry 1FPH of PDB), and also in the structure reported as the thrombin fast form [8] (entry 1SG8 of PDB). In this context, it was proposed that, in Na-bound thrombin, one of the structural water molecules bridges Na^+ to the side chain of Asp189 [41]. When Na^+ is released from its site, this water molecule is relocated; the side chain of Asp189 acquires mobility and rearranges to a position that is no longer optimal for the electrostatic coupling with the guanidinium group of the Arg side chain at the P1 position of the substrate. The main significance of this observation is the fact that the rearrangement of Asp189 upon Na^+ binding to the correct position enhances substrate specificity by improving the k_{cat}/K_m [42]. The structure of thrombin in the THR-nlig and THR-Na simulations supports the hypothesis that Asp189 plays a key role in both Na^+ coordination and, as a consequence, in the enhancement of catalytic activity [8], which is also consistent with results from mutagenesis data [42].

S_2 and S3 specificity sites

The active site of thrombin is occluded by the Trp60D and 95–100 insertion loops, which allow only specific ligands to access the active site pocket. These insertion loops are unique to thrombin [43]. In the crystal structure of thrombin, a nine-residue insertion loop at position 60 (Trp60D loop) forms a lid over the S2 specificity pocket of thrombin [43, 44].

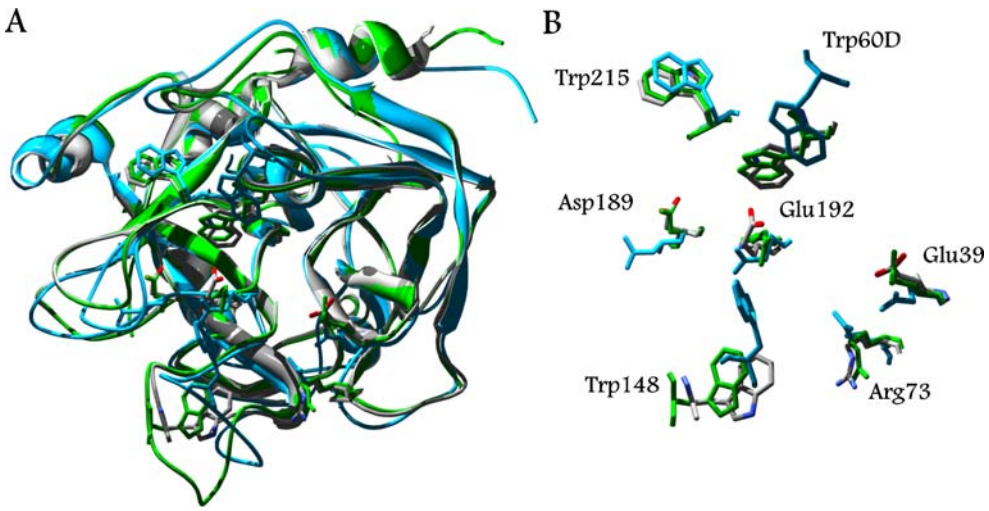


Fig. 4 **a**,**b** Structural changes induced by the Na^+ release of thrombin depicted by the superimposition of the following structures: the crystallographic structure of Na^+-bound fast form (PDB accession code 1SG8, *gray*), the crystallographic starting structure of the simulations (PDB accession code 1FPH, *green*) and the average structure from the THR-nlig simulation (*blue*). The ion and ligands were removed from the crystallographic starting structures structure for clarity. **a** Superimposition of ribbon diagrams reveal an overall conserved fold, but with conformational changes in certain regions. **b** Superimposition of specific residues illustrate the perturbations caused by Na^+ release

Moreover, allosteric linkage between the 148- and the 60-loop (S2 specificity site) has been reported to affect the functional properties of the enzyme [45–47]. The S3 specificity pocket (aryl-binding site), in its turn, is shaped by Trp215, Leu99 (located at the 95–100 insertion loop), and Ile174. In addition, the Glu192 residue, at the base of the active site pocket, affects the S3 specificity of thrombin [48, 49].

In both non-complexed Na^+-free and Na^+-bound thrombin simulations (THR-nlig and THR-Na, respectively), the Trp60D loop protruded into the solvent while in the complexes it remains occluded (Fig. 3). The results are in agreement with the hypothesis that the position of the entire 60 insertion loop is quite flexible [15], contrary to early crystallographic observations [50]. The movement of the Trp60D loop in THR-Na and THR-nlig simulations is accompanied by the rearrangement of the Trp215 side chain toward to the 95–100 loop. Although this was observed in the two simulations of the non-complexed thrombin, in the Na^+-bound thrombin simulation, the Trp60D loop remained more distant from the Trp215 side-chain than in the Na^+-free thrombin simulation, resulting in a more open active-site. The conformational linkage between Trp60D and Trp215 is also consistent with experimental data [18, 19].

In the THR-NaH2 simulation, the presence of a ligand at ABE1 promotes an increase in volume of the S2 specificity site through a slight opening of the Trp60D loop together with the rotation of the Trp60D side chain (Fig. 3, cf. c1–c2 with b1–b2). This slight increase in the volume of the active site must be related to the first stage of the mechanism of fibrinogen recognition, in which binding to thrombin ABE1 induces the partial projection of the Trp60D loop into the solvent in order to accommodate the cleavable bond of fibrinogen.

The effect of ligand binding at ABE1 can be illustrated by the results obtained by Di Cera [51], which indicate that, after binding of TM to the fast form of thrombin, the enzyme specificity increases by a factor of two, while binding of TM to the slow form increases enzyme specificity by a factor of 15. As a result, TM binding to thrombin seems to increase or diminish with the specificity differences between the two conformations. Although the crystallographic structure of thrombin bound to TM does not show any active-site conformational change [52], the presence of an active-site inhibitor does not allow us to conclude that no conformational change would otherwise be observed. Indeed, reports suggest that thrombin active-site interactions play a central role in improving recognition of protein C when thrombin is bound to TM [53].

In addition, the reaction rate for hydrolysis of small synthetic substrates is smaller for free thrombin than for thrombin complexed in the ABE1 with TM or with peptides derived from the C-terminal region of hirudin [15]. The same phenomenon was observed with bothrojaracin, a C-type lectin isolated from the venom of the snake *Bothrops jararaca*, which also binds to the ABE1 of α-thrombin [54]. In the above-mentioned cases, it is thought that the residues situated in the region of the active-site of free thrombin (THR-Na and THR-nlig simulations) are not positioned in the best geometry for association with the substrate and, consequently, for hydrolysis of the amide bond [15]. Thus, the free energy of free thrombin is higher than that of thrombin complexed at ABE1 and, as a consequence, a large number of collisions with substrate are required to effect one reaction, resulting in a decrease in the catalytic efficiency (k_{cat}/K_m). Moreover, binding at ABE1 also affects thrombin specificity. For instance, the capacity to discriminate between the tripeptides Phe-Pro-Arg*p*NA (FPR) and Phe-Gly-Arg*p*NA (FGR) of procoagulant conformation (fast) of thrombin complexed with $hir^{55\text{-}65}$ at ABE1 $\left(k_{cat}/K_{m(FPR)}/k_{cat}/K_{m(FGR)} = 19.0\right)$ is greater than that of procoagulant conformation (fast) of free thrombin $\left(k_{cat}/K_{m(FPR)}/k_{cat}/K_{m(FGR)} = 13.5\right)$ [15]. Despite the small differences between specificity constants, it should be noted that these were obtained with synthetic peptides of reduced size. In the case of natural macromolecular substrates, these differences could acquire greater significance.

The high value of the association constant (k_{on}) observed for thrombin inhibition by hirudin results, in part, from an electrostatic complementary effect. While the active-site of thrombin presents a characteristic negative electrostatic potential, the region surrounding it is filled by a positive electrostatic potential. These fields are responsible for the orientation of macromolecular substrates docking into the surface of the enzyme [55]. The positions of Glu192—highlighted by Marque et al. [13] as an important player in terms of specificity—and Trp60D are modulated by ligand binding at ABE1. In contrast to the simulations of non-complexed thrombin, in the simulations of the complexes, Glu192 seems to be better positioned to effect exclusion of the unfavorable P_3-$P_{3'}$ side-chains. Additionally, regardless of the presence of a ligand at the active-site, the binding of hirudin 2 fragment at ABE1 maintains Trp60D close to its original crystallographic position. The findings are consistent with the observations that the binding of hirudin C-terminal fragment 55–65 at ABE1 affects the kinetics (k_{cat}/K_m) of interaction of small substrates with the active site of thrombin [15].

In fact, the contribution of Glu192 to the allosteric control of thrombin was pointed out previously in a study by Nienaber and collaborators [56]. In this latter work, it was proposed that the position of Glu192 is determinant for generation of the electrostatic potential of the S3 specificity pocket of α-thrombin. According to these authors, Glu192 can adopt either an *up* or a *down* position in the S3 specificity pocket. When Glu192 is situated in the *up* position, thrombin displays a marked affinity for substrates with neutral and hydrophobic residues located at P_3-$P_{3'}$

positions. On the other hand, when Glu192 is in the *down* position, thrombin favors substrates with positively charged residues at P_3-$P_{3'}$. As seen from the electrostatic potential analysis, the region around Glu192 is predominantly negative in the average structures generated for Na^+/ligand-bonded (Fig. 3, a2 and b2) and Na^+-bonded (Fig. 3, c2) thrombin, while in the non-bonded state of the protein, the electrostatic potential of this region is not defined. Also, Glu192 of non-bonded thrombin structure is partially occluded by the hydrophobic side-chain of Tpr148 (Fig. 3, d2). Therefore, it is conceivable that an electrostatic field modified by the reorientation of Glu192 and relocation of the Trp148 loop allows for better access/orientation of substrates containing negative residues at P_3-$P_{3'}$ or, conversely, blocks access/orientation for substrates containing residues positively charged at P_3-$P_{3'}$. In the same way, the positioning of Trp60D can act synergistically with Glu192 in the control of the specificity induced for the interaction of effectors at the ABE1 region.

Conclusions

The results of MD simulations highlighted the most significant events that occur in thrombin after the binding of effectors at ABE1 and at the Na-binding site. Our findings correlate well with the known structural and recognition properties of the slow and fast forms of thrombin, and are in accordance with the hypothesis that the diverse functional domains of thrombin communicate [57]. From MD simulations, it is clear that, except for SASA, which is lower for Na^+-unbound free thrombin, the interaction of thrombin with effectors and ligands does not cause drastic global changes in protein structure or protein flexibility. The most notable structural changes are specific, involving certain loops and the side-chains of some important residues. Our results showed that Na^+ release results in a more closed conformation of thrombin, which could represent the slow form. The conformational changes provoked by displacement of the sodium ion from the Na-binding site results in: (1) distortion of the 220- and 186-loops that comprise the Na-binding site; (2) folding back of the Trp148 loop towards the body of protein, in such a way that the Trp148 side-chain remains positioned in front of Glu192, leading to narrowing of the cleft that allows communication between ABE1 and the active site; (3) a 180° rotation of the Asp189 side-chain, which leaves the Oδ1 atom substantially distant from its optimal coupling with the guanidinium group of Arg at the P1 position of the substrate; and (4) projection of the Trp60D loop toward the solvent accompanied by rearrangement of the Trp215 side chain towards the 95–100 loop. On the other hand, the Na^+-bound forms of thrombin can be compared to the fast conformation. For both forms, the presence of sodium at the binding site confers a greater structural rigidity to the 220- and 186-loops than in the Na^+-free form. The interaction of $hir^{54–65}$ with thrombin ABE1 stabilizes the average position of the Trp148 loop—a condition similar to that described in the crystallographic conformation. In free Na^+-bound thrombin, the Trp148 loop adopts a conformation more exposed to solvent. The presence of $hir^{54–65}$ in ABE1 seems to provide a better position for Glu192 in order to originate the exclusion of the unfavorable P_3-$P_{3'}$ side-chains at the entrance to the active-site. The theoretical models generated from these MD simulations reinforce, but also clarify, the currently available structural information, which is based mainly on crystallographic, kinetic and spectroscopic data. MD simulations also efficiently overcome some constraints intrinsically imposed uniquely by crystallographic studies.

References

1. Mann KG, Butenas S, Brummel K (2003) Arterioscler Thromb Vasc Biol 23(1):17–25
2. Crawley JT, Zanardelli S, Chion CK, Lane DA (2007) J Thromb Haemost 5:95–101
3. Stubbs MT, Bode W (1995) Trends Biochem Sci 20(1):23–28
4. Bode W (2006) Semin Thromb Hemost 32:16–31
5. Di Cera E (2008) Mol Aspects Med 29(4):203–254
6. Di Cera E, Guinto ER, Vindigni A, Dang QD, Ayala YM, Wuyi M, Tulinsky A (1995) J Biol Chem 270(38):22089–22092
7. Di Cera E, Vindigni A, Dang QD (1995) Proc Natl Acad Sci USA 92:5977–5981
8. Pineda AO, Carrell CJ, Bush LA, Prasad S, Caccia S, Chen ZW, Mathews FS, Di Cera E (2004) J Biol Chem 279(30):31842–31853
9. Di Cera E, Page MJ, Bah A, Bush-Pelc LA, Garvey LC (2007) Phys Chem Chem Phys 9(11):1291–1306
10. Huntington JA (2008) Biol Chem 389(8):1025–1035
11. Ayala YM, Alessandro V, Nayal M, Spolar RS, Di Cera E (1995) J Mol Biol 253:787–798
12. Mengwasser KE, Bush LA, Shih P, Cantwell AM, Di Cera E (2005) J Biol Chem 280:23997–24003
13. Marque EP, Spuntarelli R, Juliano L, Aiach M, Le Bonniec BF (2000) J Biol Chem 275(2):809–816
14. Parry MAA, Stone SR, Hofsteenge J, Jackman MP (1993) Biochem J 290:665–670
15. Guinto ER, Di Cera E (1997) Biophys Chem 64:103–109
16. Priestle JP, Rahuel J, Rink H, Tones M, Gruetter MG (1993) Protein Sci 2(10):1630–1642
17. Huntington JA (2005) Biochem J 390:e1–e3
18. Huntington JA, Esmon CT (2003) Structure 11:469–479
19. Johnson DJ, Adams TE, Li W, Huntington JA (2005) Biochem J 392:21–28
20. De Mori GM, Meli M, Monticelli L, Colombo G (2005) Mini Rev Med Chem 5(4):353–359
21. Day R, Daggett V (2003) Adv Protein Chem 66:373–403
22. Karplus M, Kuriyan J (2005) Proc Natl Acad Sci USA 102(19):6679–6685
23. Morra G, Meli M, Colombo G (2008) Curr Protein Pept Sci 9(2):181–196
24. Dodson GG, Lane DP, Verma CS (2008) EMBO Rep 9(2):144–150

25. De Filippis V, Colombo G, Russo I, Spadari B, Fontana A (2002) Biochemistry 41(46):13556–13569
26. Bah A, Garvey LC, Ge J, Di Cera E (2006) J Biol Chem 281 (52):40049–40056
27. Guex N, Peitsch MC (1997) Electrophoresis 18:2714–2723
28. Humphrey W, Dalke A, Schulten K (1996) J Mol Graph 14(1):33–38
29. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) J Comput Chem 25(13):1656–1676
30. van der Spoel D, Lindahl E, Hess B, van Buuren AR, Apol E, Meulenhoff PJ, Tieleman DP, Sijbers ALT, Feenstra KA, van Drunen R, Berendsen HJC (2005) Gromacs User Manual version 33, www.gromacs.org. Accessed 19 Sep 2007
31. Berendsen HJC, Grigera JR, Straatsma TP (1997) J Phys Chem 91 (24):6269–6271
32. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) J Chem Phys 81(8):3684–3690
33. Parrinello M, Rahman A (1981) J Appl Phys 52:7182–7190
34. Nosé S, Klein ML (1983) Mol Phys 50:1055–1076
35. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) J Comp Chem 18:1463–1472
36. Darden T, York D, Pedersen L (1993) J Chem Phys 98:10089–10092
37. Caves LSD, Nguyen DT, Hubbard RE (1990) Conformation vari ability of insulin A molecular dynamics analysis. In: Goodfellow JM (ed) Molecular dynamics: applications in molecular biology. CRC, Boca Raton, pp 27–68
38. Kabsch W, Sander C (1983) Biopolymers 22:2577–2637
39. De Filippis V, De Dea E, Lucatello F, Frasson R (2005) Biochem J 390:485–492
40. Stubbs MT, Oschkinat H, Mayr I, Huber R, Angliker H, Stone SR, Bode W (1992) Eur J Biochem 206:187–195
41. Pineda AO, Savvides SN, Waksman G, Di Cera E (2002) J Biol Chem 277:40177–40180
42. Prasad S, Cantwell AM, Bush LA, Shih P, Xu H, Di Cera E (2004) J Biol Chem 279:10103–10108
43. Bode W, Mayr I, Baumann U, Huber R, Stone SR, Hofsteenge J (1989) EMBO J 8(11):3467–3475
44. Stubbs MT, Bode W (1993) Thromb Res 69(1):1–58
45. De Cristofaro R, Landolfi R (1999) Eur J Biochem 260:97–102
46. Wang SX, Esmon CT, Fletterick RJ (2001) Biochemistry 40:10038
47. van der Locht A, Bode W, Huber R, Le Bonniec BF, Stone SR, Esmon CT, Stubbs MT (1997) EMBO J 16:2977–2984
48. Rezaie AR, Esmon CT (1993) J Biol Chem 268(27):19943–19948
49. Rezaie AR, Esmon CT (1996) Eur J Biochem 242(3):477–484
50. Bode W, Turk D, Karshikov A (1992) Protein Sci 1(4):426–471
51. Di Cera E, Dang QD, Ayala YM (1997) Cell Mol Life Sci 53:701–730
52. Fuentes-Prior P, Iwanaga Y, Huber R, Pagila R, Rumennik G, Seto M, Morser J (2000) Nature 404:518–525
53. Lu G, Chhum S, Krishnaswamy S (2005) J Biol Chem 280: 15471–15478
54. Monteiro RQ, Rapôso JG, Wisner A, Guimarães JA, Bon C, Zingali RB (1999) Biochem Biophys Res Commun 262:819–822
55. Myles T, Le Bonniec BF, Betz A, Stone SR (2001) Biochemistry 40(16):4972–4979
56. Nienaber VL, Mersinger LJ, Kettner CA (1996) Biochemistry 35(30):9690–9699
57. Hall SW, Nagashima M, Zhao L, Morser J, Leung LL (1999) J Biol Chem 274(36):25510–25516