Journal of Computer-Aided Molecular Design, **14:** 317–327, 2000.
KLUWER/ESCOM

Structural analysis of the KGD sequence loop of barbourin, an $\alpha_{IIb}\beta_3$-specific disintegrin

Hervé Minoux, Christophe Chipot, David Brown & Bernard Maigret*
Laboratoire de Chimie théorique, UMR CNRS 7565 – Structure et Réactivité des Systèmes Moléculaires Complexes, Institut Nancéien de Chimie Moléculaire, Université Henri Poincaré, BP 239, F-54506 Vandoeuvre-lès-Nancy Cedex, France

Received 8 October 1998; Accepted 4 November 1999

Key words: cation-π interactions, integrin, KGD, molecular dynamics

Summary

Disintegrins constitute a class of small proteins that inhibit platelet aggregation by binding to the fibrinogen receptor, also referred to as integrin $\alpha_{IIb}\beta_3$. Contrarily to other disintegrins that bind to a series of integrins via their Arg-Gly-Asp domain, the recognition site of barbourin contains a Lys-Gly-Asp sequence that ensures its specificity towards $\alpha_{IIb}\beta_3$. In this article, a three-dimensional model of barbourin is proposed using homology modeling and large-scale molecular dynamics simulations. The conformations of the Lys-Gly-Asp sequence of barbourin are analyzed and compared to those of peptidomimetics that exhibit similar specificity towards $\alpha_{IIb}\beta_3$. The tryptophan residue following the Lys-Gly-Asp sequence of the binding domain is shown to play a crucial role in the biological activity and the specificity of barbourin. Our results suggest that this disintegrin anchors to the binding pocket of the γ-chain of fibrinogen rather than to those of the Arg-Gly-Asp sequence.

Introduction

Disintegrins are a family of low molecular weight proteins, containing between 39 (for decorsin [1]) and 84 (for gabonin [2]) amino acid residues, and originally extracted from the venom of snakes belonging to the *viperidae* family [3], but later from the saliva of certain ticks and from leeches [4]. These proteins are known to inhibit the adhesive functions of some receptors belonging to the family of integrins [3].

Integrins form one of the four families of receptors for cell adhesion, and consist of two subunits: α and β [5]. Integrin $\alpha_{IIb}\beta_3$, also referred to as glycoprotein IIb-IIIa, is located exclusively in platelets and megakaryocytes, and is essential for the binding between platelets and fibrinogen during the platelet aggregation process, which is responsible for coronary thrombosis [6]. In contrast, integrin $\alpha_v\beta_3$ is present in many cells and is involved in the cell migration associated with many pathophysiological phenomena such as angiogenesis, restenosis, tumor cell invasion and atherosclerosis. These two integrins share the common β_3 subunit, and all their ligands, e.g. fibrinogen, the von Willebrand factor, fibronectin or vitronectin, contain the Arg-Gly-Asp (RGD) sequence, hypothesized to be essential for the binding of agonists or antagonists to several integrins, including $\alpha_{IIb}\beta_3$ and $\alpha_v\beta_3$ [4, 7, 8].

All known disintegrins, except two, contain the Arg-Gly-Asp sequence, located at the top of a loop protruding from the core of the protein. These small proteins are cysteine-rich and include many disulfide bridges that constrain the Arg-Gly-Asp loop into an optimal conformation for binding to integrins. Yet, none of these Arg-Gly-Asp containing proteins is specific towards integrin $\alpha_{IIb}\beta_3$ [9].

Barbourin is a disintegrin formed by 73 amino acids, that possesses two interesting peculiarities: (i) it contains a Lys-Gly-Asp sequence rather than an Arg-Gly-Asp one, and (ii) it is is specific towards integrin $\alpha_{IIb}\beta_3$. Lysine for arginine substitution has been as-

*To whom correspondence should be addressed. E-mail: maigret@incm.u-nancy.fr

sumed to be the only structural feature responsible for the specificity of barbourin [10]. On this basis, small cyclic peptides containing the Lys-Gly-Asp sequence, or an analog to this sequence, have been developed [11–16]. Among those compounds, one of them, eptifibatide or integrilin, prevents platelet aggregation with a high selectivity for $\alpha_{IIb}\beta_3$, and is currently in clinical trial [17]. Other compounds containing the Lys-Gly-Asp sequence have been developed, but the weak selectivity of these molecules to integrin $\alpha_{IIb}\beta_3$ tends to indicate that a lysine substitution is not sufficient for integrin selectivity, and that other structural features can contribute to this selectivity [9]. A second disintegrin specific towards $\alpha_{IIb}\beta_3$, ussutistatin 2, has been discovered recently, and contains also the Lys-Gly-Asp sequence [18].

The goal of the present contribution is to provide the structural and electrostatic criteria that condition both activity and selectivity of barbourin to integrin $\alpha_{IIb}\beta_3$. Because no structural data is available for barbourin, the three-dimensional model of this protein was built by homology and a molecular dynamics (MD) simulation of this model was carried out in bulk water over a significant time scale. The resulting conformations of this protein were then analyzed. They show a constrained core containing all six disulfide bridges, from which the more flexible Lys-Gly-Asp loop and the two N- and C-terminal parts protrude. In addition, no regular secondary structures over a significant stretch of this protein were found, as has been found previously for other disintegrins [9]. The resulting conformations of the Lys-Gly-Asp loop were, also, compared to those of small compounds inhibiting specifically integrin $\alpha_{IIb}\beta_3$.

Materials and methods

Building the starting structure of barbourin

The program Fasta3 [19, 20] allowed us to find among all protein structures present in the protein data bank (PDB) [21, 22] the one that exhibits the highest similarity, at the amino acid sequence level, with barbourin, i.e. the disintegrin flavoridin [23, 24]. From the 73 residues that form barbourin and the 70 of flavoridin, 55 are common to the two proteins (see Figure 1). The PDB file 1fvl contains 18 proposed conformations of flavoridin, elucidated from ^{1}H-NMR data [23, 24]. The structure of flavoridin is highly constrained by disulfide bridges and the most important differences between the 18 conformations occur in the N- and C-terminal parts and, to a lower extent, in the Arg-Gly-Asp-containing loop. The distance RMSDs between 18 conformations forming this set, and computed over the C_α carbons, range from 1.4 to 4.3 Å. The potential energy of each of these 18 conformations was computed using the CVFF force field [25], and conformation 13, corresponding to the lowest energy computed for this set of 18 experimental conformations, was selected to build a first three-dimensional model of barbourin. An energy minimization in vacuum, using a distant dependent dielectric constant for each of the 18 conformations, was also performed and conformation 3, corresponding to the lowest energy after optimization, was selected to build a second model for barbourin. The distance RMSD between these two models, and computed over the C_α carbons, is 4.2 Å.

Using the selected structures as a reference and the program Homology [26], the 15 amino acids differentiating flavoridin from barbourin were replaced, two residues were added at the N-terminus, and a third one was added at the C-terminal part to yield the two models of barbourin. The distribution of disulfide bridges along the sequence is problematic. As has been shown [9], pairing of cysteine residues can be different between disintegrins. An examination of known distributions of disulfide bridges for other disintegrins than barbourin, i.e. echistatin [27, 28], kistrin [29] and flavoridin [23, 24], reveals a well-conserved pattern for these bonds, except for the smaller echistatin. The distribution of disulfide bridges in flavoridin was, thus, conserved in our model of barbourin, as can be seen in Figure 2. Considering the high sequence similarity between barbourin and flavoridin, it is not surprising that the two models constructed show structural similarities with respect to flavoridin.

Preparation of the system

After preliminary, partial energy minimization using the CVFF force field, which is necessary to eliminate the structural artifacts introduced into the models of barbourin during the substitution of the residues, the resulting conformations were placed at the center of a cubic, periodic cell of side length equal to 80 Å, filled with explicit CVFF water molecules [30]. The electric neutrality of the cells, imposed by Ewald lattice summation simulations [31], was obtained by substituting four water molecules, lying away from the protein, by four sodium cations. The systems, contain-

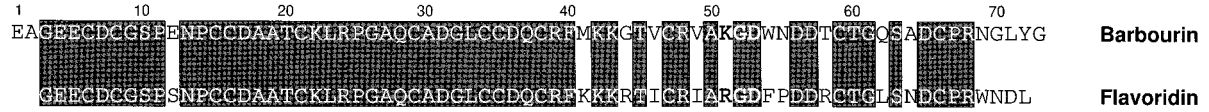


Figure 1. Comparison between the amino acid sequences of barbourin and flavoridin using the one-letter code. Common residues are highlighted in a box.

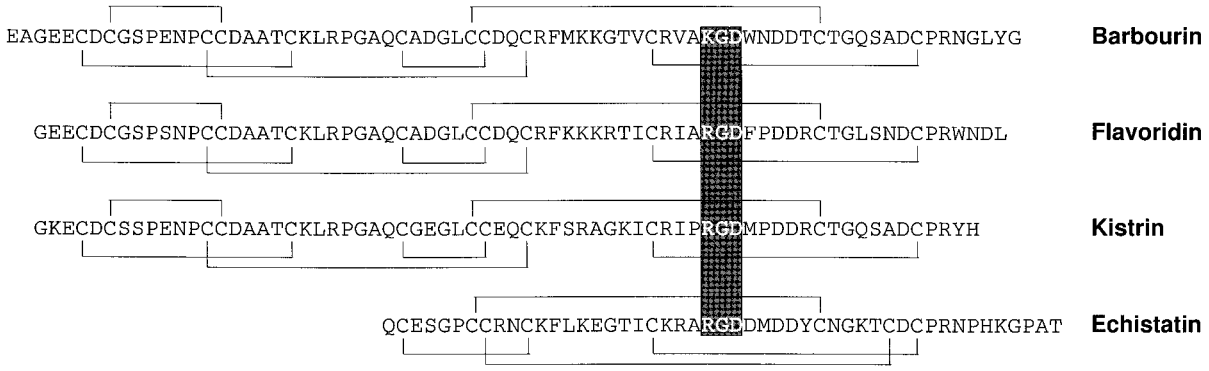


Figure 2. Comparison between the amino acid sequences and the distributions of disulfide bridges in echistatin, kistrin, flavoridin and barbourin, using the one-letter code.

ing 994 atoms for barbourin, four cations and 16611 water molecules, i.e. 50831 atoms, were employed as starting points of the MD simulations.

Molecular dynamics simulation

The parallel code Ddgmq [32, 33] was employed to study the dynamics of barbourin in an aqueous solution. The potential energy function used in this program slightly differs from CVFF in the expressions of angle bending and out-of-plane terms [32, 33]. Use was made of a parallel version of Shake to constrain all bond lengths to their equilibrium value [34].

Coulombic energies and forces were determined using an Ewald lattice summation technique [31]. The corresponding parameters, i.e. the separation parameter α, the real space cut-off and the number of *k*-vectors were 0.19 $Å^{-1}$, 11.5 Å and 1000, respectively. A time-step of 1 fs was utilized to integrate the equations of motion. The main cell was split into 27 cubic domains assigned to 27 processors of a 64-processor Silicon Graphics Origin 2000. The CPU time per time-step was 1.5 s.

The systems were equilibrated during 300 ps whilst maintaining the temperature and the pressure at 298 K and 1 bar, respectively. Subsequently, the system containing the first model of barbourin evolved in the canonical ensemble, over a simulation length equal to 2 ns, from which configurations were extracted every 5 ps, resulting in a sample of 400 configurations. For the second model of barbourin, a 1 ns simulation in the same ensemble was used to produce 200 configurations. This second simulation was only carried out for comparison purposes.

Ab initio calculations

Separate ab initio quantum chemical calculations were carried out to estimate the energetics of the complex formed by 3-methylindole and the ammonium cation, which represents a relevant model for the study of the interaction of the tryptophan and the lysine side chains. The minimum energy structure was determined at the MP2/6-31G** level with full counterpoise correction, optimizing not only the distance separating the nitrogen atom of the ammonium from the centroid of the C_6 ring of 3-methylindole, but also the relative orientation of the two species. Single point energy calculations for the optimized dimer were then performed at the MP2/6-311++G** level. All ab initio calculations were carried out using the Gaussian 94 suite of programs [35].

Results and discussion

Structure and dynamics of barbourin

A clustering analysis was used to compare the 400 configurations generated in the course of the first

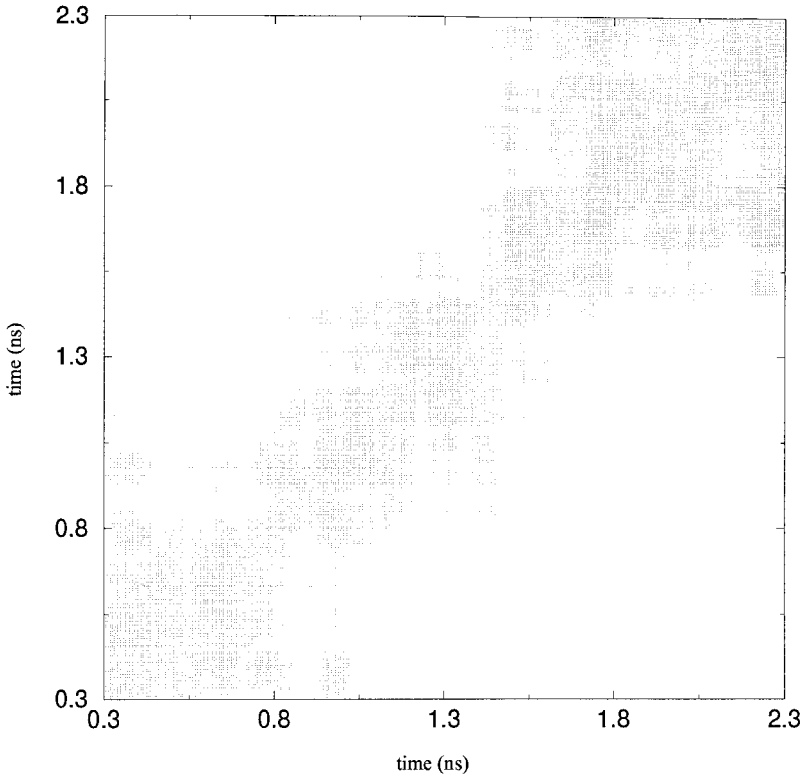


Figure 3. Cluster graph constructed from the 400 configurations of barbourin stored during the 2 ns MD simulation, using only the trace of the peptide. Only those points having a distance RMSD value lower than 2.4 Å are plotted.

MD trajectory. Figure 3 shows the cluster graph for the 73 C_α atoms constituting the trace of barbourin. The plotted points, corresponding to a distance root mean square deviation (RMSD) lower than or equal to 2.4 Å, highlight the existence of, broadly speaking, 4 structural families. A representative conformation was extracted from each of them, and all the traces were superimposed, emphasizing (i) a rigid core that corresponds to residues Cys^6-Cys^{38} and Cys^{47}-Cys^{66} including the six disulfide bonds, (ii) a more flexible loop corresponding to residues Arg^{39}-Val^{46} with the Lys-Gly-Asp sequence at its top, and (iii) a small N-terminal part, Glu^1-Glu^5, and a longer, more flexible C-terminal part, Pro^{67}-Gly^{73}.

The distance RMSD values for the trace of these 4 conformations range from 1.27 Å to 5.22 Å. The corresponding values for the 18 conformations of flavoridin elucidated from ^{1}H-NMR data are 1.41 Å and 4.32 Å, respectively. The major structural differences between these conformations, for both flavoridin and barbourin, lie in the N- and C-terminal parts and, to a lesser extent, in the Arg/Lys-Gly-Asp-containing loop. The several disulfide bridges rigidify these two small proteins in a similar way, thereby limiting noticeably their flexibility. Even if the simulation length is not long enough to explore thoroughly the conformational space of barbourin, the resulting conformations seem, in fact, to be sufficiently different to allow a detailed study of the Lys-Gly-Asp loop. It is worth pointing out that similar calculations have been performed on echistatin, and have been compared successfully to NMR spectroscopy data, thereby validating the approach [36].

Examination of the backbone for the 4 resulting conformations shows little regular secondary structure for this hydrated protein. Small patterns of antiparallel β-sheets are, however, present between Val^{46}-Arg^{48} and Asp^{57}-Asp^{58} for most of the 4 conformations, and between Cys^{16}-Asp^{17} and Lys^{22}-Leu^{23} in only 5 conformations. In addition, turns at the top of the loops formed by the disulfide bridges are present in all 4 extracted conformations. These structural properties are characteristic of disintegrins and were found in other members of this family [9].

The Lys-Gly-Asp loop

A second clustering analysis was performed to segregate the different structural families for all heavy atoms of the Lys-Gly-Asp sequence and of its two flanking residues, i.e. Ala^{50} and Trp^{54}. Among the 400 configurations sampled, six structural groups were isolated for this sequence, and a representative conformation for each family was extracted. The superimposition of these six conformations shows a rigid structure for the backbone, except for the Trp^{54} residue, which adopts two distinct conformations: one corresponding to a pseudo β-turn, where the hydrogen atom of the amino group of Trp^{54} is close to the oxygen atom of Lys^{51}, and another one corresponding to a pseudo γ-turn, where the hydrogen atom of the amino group of Trp^{54} is close to the oxygen atom of Gly^{52}, the latter being preferred, as shown in Figure 4.

A similar study has been conducted previously for echistatin, in which the Arg-Gly-Asp sequence is located at the top of a loop constrained by disulfide bridges and hydrogen bonds linking the two backbones to form such a loop. In barbourin, there is only one strong hydrogen bond between these two backbones, i.e. between the oxygen atom of Arg^{48} and the hydrogen atom of the amino group of Asp^{57}, and two short-lived hydrogen bonds, i.e. between the oxygen atom of Val^{49} and the hydrogen atom of the amino group of Lys^{51}, and between the oxygen atom of Ala^{50} and the hydrogen atom of the amino group of Gly^{52}. The hairpin motif at the top of the Lys-Gly-Asp loop must, thus, be less constrained than the one in echistatin.

It is expected that the distance separating the charged moieties in the Arg-Gly-Asp sequence is pivotal for the inhibitory activity of compounds containing this sequence or mimicking it [37]. The variation of the distance between the N_ζ atom of Lys^{51} and the C_γ atom of Asp^{53} as a function of time was, therefore, plotted (see Figure 4). This distance is, however, always greater than the corresponding one found in small, specific antagonists of integrin $\alpha_{IIb}\beta_3$, i.e. between 6.2 and 9.3 Å [38].

In addition, the charged moiety of Lys^{51} is highly mobile, due to the flexibility of its side chain. Furthermore, the N_ζ atom of Lys^{51} frequently appears in the vicinity of the aromatic Trp^{54} side chain (see Figure 4), which is suggestive of possible cation-π interactions. The variation of the distance between the C_γ atom of Asp^{53} and the N_ϵ atom of Trp^{54} as a function of time is displayed in Figure 4. For the most part, this distance ranges between 6.2 and 9.3 Å. This indicates that the Trp^{54} side chain might play a key role in the selectivity and the biological activity of barbourin. To assess this assumption, some non-peptide specific antagonists of integrin $\alpha_{IIb}\beta_3$ were superimposed to the Ala-Lys-Gly-Asp-Trp binding sequence of barbourin.

The aforementioned observations were reproduced in the trajectory obtained using the second model of barbourin, hence, suggesting that the simulation time utilized was appropriate to describe the majority of the relevant conformations of the Lys-Gly-Asp loop, regardless of the initial structure. In particular, considering the evolution of the distance between the Lys and Trp side chains as obtained from the second sample, it appears that the cation-π interaction described previously (see Figure 6) is by no means an artifact of either the methodology used or the limited sampling.

Superimposition of specific antagonists of integrin $\alpha_{IIb}\beta_3$

In a previous work [38] a common pattern for several specific antagonists of integrin $\alpha_{IIb}\beta_3$ was elucidated. Here, the conformations in an aqueous solution for the four non-peptide molecules amongst the series of antagonists investigated in Reference 36 are compared to those of the binding sequence of barbourin, which shows similar biological activity and selectivity. The molecules investigated here are: MK0383 [39], Ro 43-5054 [40], Ro 43-8857 [40] and FR 144633 [41] (see Figure 5). In order to compare the structures of these small molecules with those of the Lys-Gly-Asp loop of barbourin, superimpositions were performed via a rigid reorientation of each molecule with respect to the Lys-Gly-Asp-Trp loop, in order to minimize the RMSD between selected atoms characterizing each pharmacophore group.

These compounds and the conformations of the Lys-Gly-Asp sequence of barbourin do not superimpose well when the N_ζ atom of Lys^{51} serves as the positive charge reference for matching the structures. This matching, enforced in constrained energy minimizations, is destroyed in MD simulations. The constrained structure of barbourin, when submitted to a few hundred picoseconds of unconstrained MD, returns to conformations that do not exhibit the starting motif.

When superimposing (i) one of the carboxylate groups of each small compound, viz. marked 'negative 1' in Figure 5 and that of the Asp^{53} pertaining

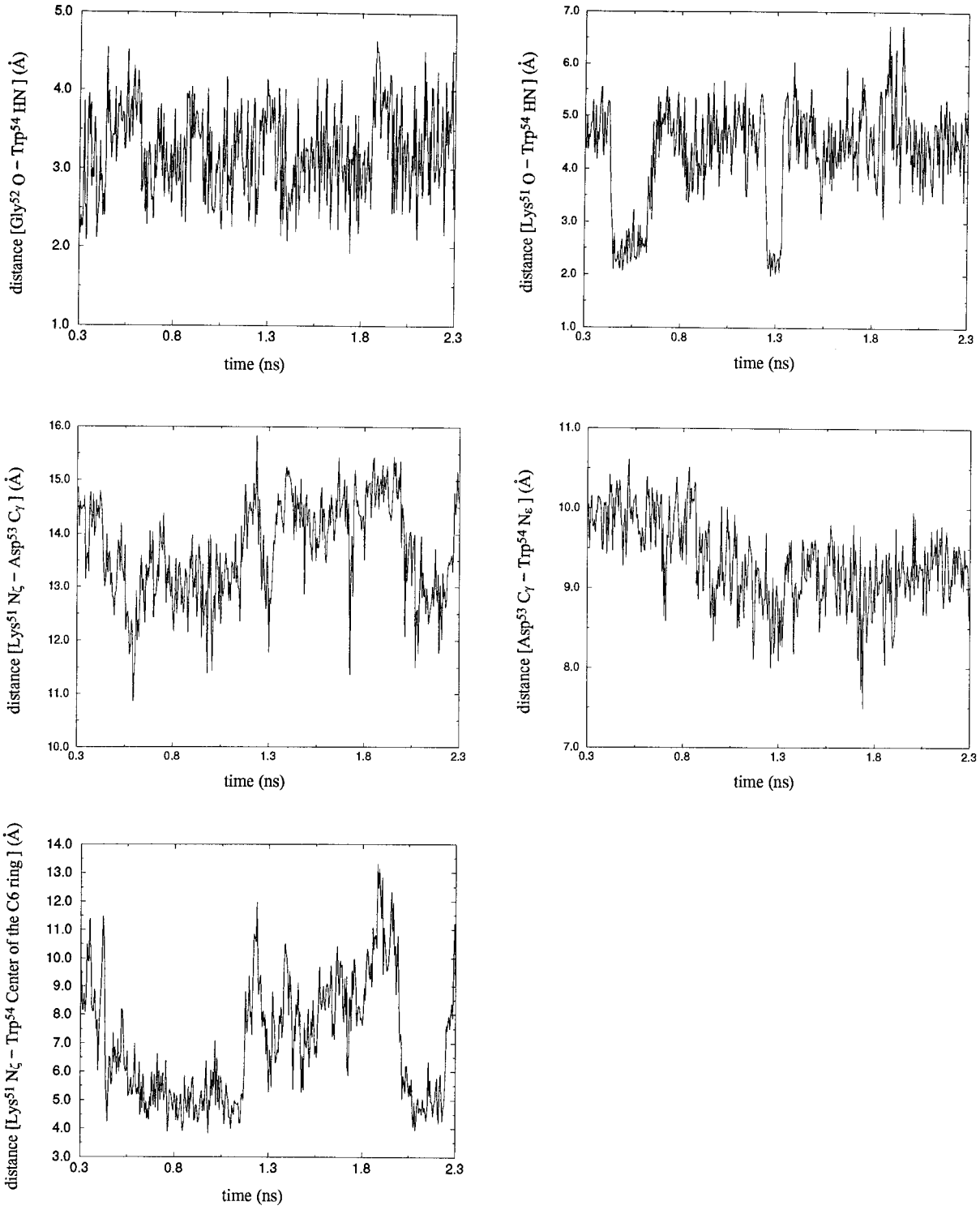
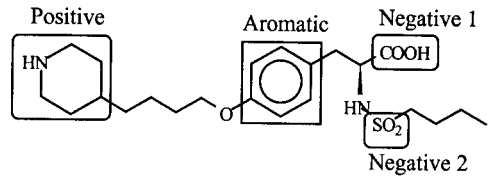
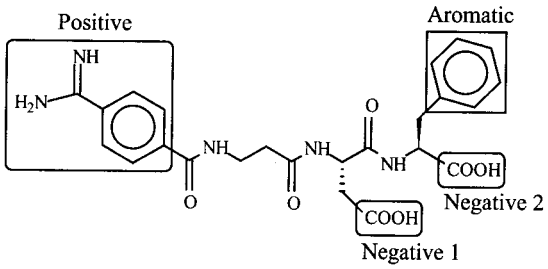


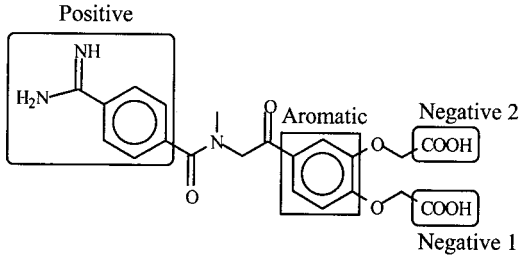
Figure 4. Distances (in Å) between the specified atoms or pseudo-atoms of the Lys^{51}-Gly^{52}-Asp^{53}-Trp^{54} sequence, shown as a function of time (in ns) during the 2 ns MD simulation.



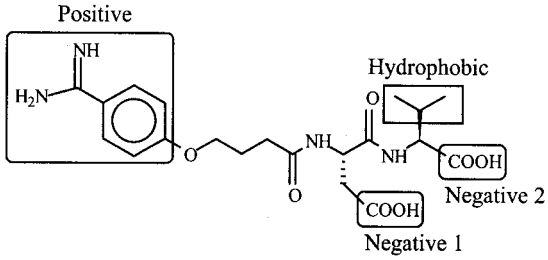
(1) MK0383 :
IC_{50} [$\alpha_{IIb}\beta_3$] = 36 nM
IC_{50} [$\alpha_v\beta_3$] = 0.1 mM



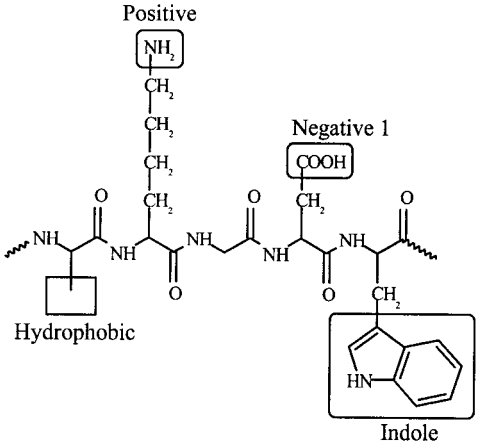
(2) Ro 43-5054 :
IC_{50} [$\alpha_{IIb}\beta_3$] = 60 nM
IC_{50} [$\alpha_v\beta_3$] = 42.3 μM



(3) Ro 43-8857 :
IC_{50} [$\alpha_{IIb}\beta_3$] = 70 nM
IC_{50} [$\alpha_v\beta_3$] = 348 μM



(4) FR 144633 :
IC_{50} [$\alpha_{IIb}\beta_3$] = 78 nM
IC_{50} [$\alpha_v\beta_3$] >1 mM



(5) Barbourin :
IC_{50} [$\alpha_{IIb}\beta_3$] = 15 nM
IC_{50} [$\alpha_v\beta_3$] >1 μM

Figure 5. The structures of the four non-peptide antagonists of $\alpha_{IIb}\beta_3$ and of the Ala-Lys-Gly-Asp-Trp sequence of barbourin. Functional groups are highlighted in a box. The IC_{50} [$\alpha_{IIb}\beta_3$] and IC_{50} [$\alpha_v\beta_3$] values correspond to the potency to inhibit integrin $\alpha_{IIb}\beta_3$ and the potency to inhibit the vitronectin receptor, respectively.

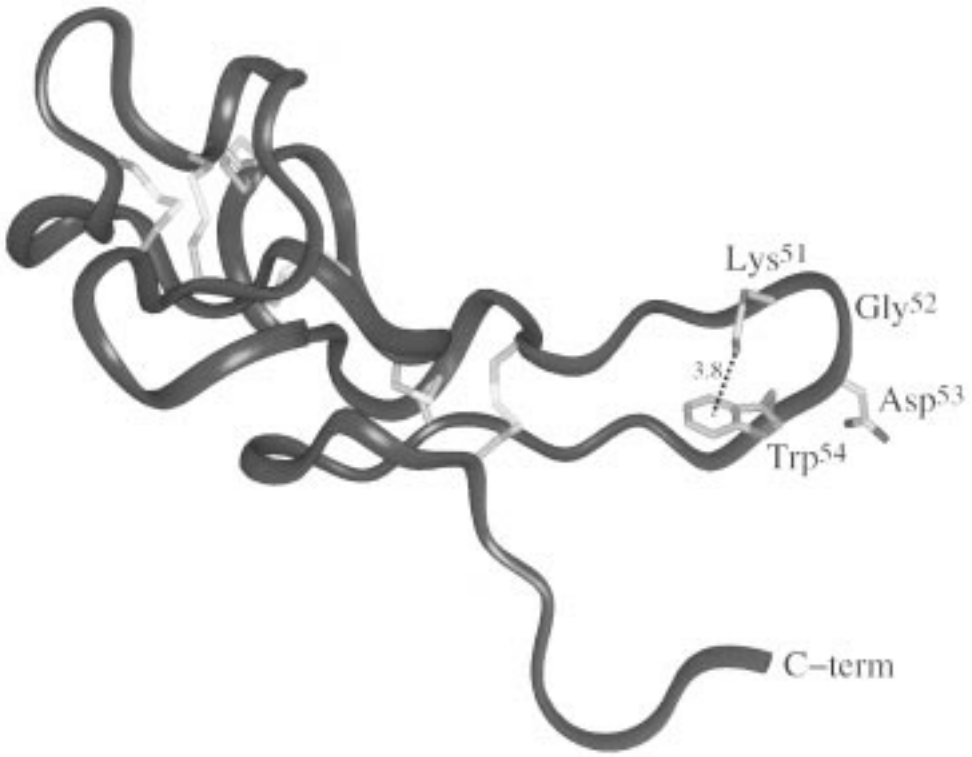


Figure 6. Example of a conformation of barbourin that exhibits a typical cation-π interaction formed by the Lys and Trp side chains.

to barbourin, (ii) the positively charged moiety, viz. marked 'positive' in Figure 5, and the N_ϵ atom of Trp^{54}, and (iii) the 'aromatic' or 'hydrophobic' moiety in Figure 5 and the side chain of Ala^{50}, not only do these three moieties match well, but the second carboxylate – or sulfone – group, viz. marked 'negative 2' in Figure 5, matches one carbonyl group of the backbone. Compounds Ro 43-5054 and FR 144633 appear to be the best mimetics of the Lys-Gly-Asp loop of barbourin, not only in terms of functional groups (i)–(iii), but also with respect to the backbone to which they are bound.

The present superimposition highlights the important role played by the Trp^{54} residue in the specific binding mechanism of barbourin towards integrin $\alpha_{IIb}\beta_3$. We note in passing that a tryptophan residue follows also the Lys-Gly-Asp sequence in ussutistatin 2, which selectively binds $\alpha_{IIb}\beta_3$ [18]. The increased affinity of disintegrin mutants possessing a tryptophan residue after the Arg-Gly-Asp sequence has been mentioned previously for echistatin [42] and kistrin [9]. In addition, disintegrins containing the Arg-Gly-Asp-Trp sequence have been found to be more selective inhibitors of $\alpha_{IIb}\beta_3$ than those containing the Arg-Gly-Asp-Asn-Pro sequence [11].

The study of the structure–activity relationship in small cyclic peptides derived from the Lys-Gly-Asp sequence of barbourin has shown that: (i) the X*aa* residues corresponding to the highest potency in the Lys-Gly-Asp-X*aa* sequence are those that are both hydrophobic and aromatic, the hydrophobic characteristic being the most important one, (ii) replacement of lysine by any residue other than arginine, or substitution of glutamic acid for aspartic acid in Lys-Gly-Asp-Trp-Pro-containing peptides leads to non-active analogs, (iii) D-Lys for L-Lys substitution does not modify significantly the inhibitory potency, whereas D-Asp substitution leads to a loss of inhibition, and (iv) modifications of the lysine side chain, for less basic moieties, lead to highly potent and selective antagonists of $\alpha_{IIb}\beta_3$ [12].

In addition, peptidomimetics bearing an amidino phenyl moiety exhibit a much greater activity for inhibiting $\alpha_{IIb}\beta_3$ than those containing a guanidinium moiety [43].

In fact, numerous non-peptide selective antagonists of integrin $\alpha_{IIb}\beta_3$ contain a piperidinyl or an

amidinophenyl moiety as a substitute for the arginine side chain.

Finally, contrarily to the initial assumption that lysine for arginine substitution in barbourin is the only structural feature that conditions its specificity [10], it should be noted that a variant of kistrin containing the Lys-Gly-Asp-Met-Pro sequence inhibits both integrins $\alpha_v\beta_3$ and $\alpha_{IIb}\beta_3$ [9]. Other residues in the Lys-Gly-Asp loop of barbourin, thus, play a role in the selectivity of this protein.

We, consequently, propose that the hydrophobic and/or aromatic groups located next to the positively charged moieties in the small antagonists of Figure 5, mimic in fact the hydrophobic side chain of tryptophan. The lysine side chain is necessary for the activity of the protein, and the charged moiety of this residue can interact with the aromatic ring of tryptophan. Additive force fields like the one used here, are, however, notoriously not very well adapted for an accurate description of cation-π interactions. In particular, in the absence of explicit polarization terms, the energetics associated with such interactions is usually underestimated [44]. It is, therefore, likely that the fluctuations in the distance separating the N_ζ atom of Lys^{51} from the center of the C_6 ring of Trp^{54} should be somewhat smaller than that reported herein.

Ab initio quantum chemical calculations were performed to investigate the interaction of the tryptophan and the lysine side chains [45]. The model adopted consisted of 3-methylindole supplemented by an ammonium cation. Preliminary geometry optimization at the MP2/6-31G** level of approximation with full counterpoise correction indicates that the bidentate motif, whereby two hydrogen atoms of ammonium point towards the C_6 ring of 3-methylindole, is energetically favored. In addition, single point energy calculations were performed at the MP2/6-311^{++}G** level on the complex at both the minimum of the potential energy surface, viz. ammonium nitrogen atom to C_6 ring centroid separation of 2.85 Å, and at a greater separation, viz. 8.00 Å. The MP2/6-311^{++}G**//MP2/6-31G** interaction energy of the contact pair, amounting to -23.41 kcal/mol, compared to the corresponding value obtained with the CVFF potential energy function of -10.52 kcal/mol emphasizes the inappropriateness of conventional, additive force fields to describe cation-π interactions. This is also illustrated by the greater separation of the onium group from the π-system by 0.44 Å, when employing the CVFF potential energy function.

Figures 4 and 6 suggest that the charged moiety of the Lys^{51} residue is sufficiently close to the tryptophan side chain, but also to the positively charged groups of the small, superimposed peptidomimetics. Using an MP2/6-311^{++}G** wavefunction, both the Mulliken and the potential derived charges borne by the N-H moiety of 3-methylindole were determined. Comparison of these charges when the separation of ammonium from the centroid of the C_6 ring decreases from 8.00 to 2.85 Å indicates an increase in the polarity of the N-H bond. This suggests a similar increase in the N_ϵ-H bond of Trp^{54} when the latter interacts with Lys^{51}. The side chain of lysine being engaged in a cation-π interaction, the N_ϵ-H moiety could, in turn, interact favorably with the receptor. It should be noted, however, that in the absence of tryptophan, just like in small Lys-Gly-Asp-X*aa* variants [12], the side chain of lysine is still flexible enough to adopt the appropriate conformation for binding to the receptor. This is in agreement with the high potency and selectivity of small cyclic peptides incorporating acetimidyl-lysine and phenylimidil-lysine analogs, i.e. residues that are less basic than arginine or lysine [12]. Moreover, lysine or arginine are of paramount importance in the sequence to guarantee biological activity [12]. The high flexibility of the lysine side chain is highlighted by the inhibitory activity of small D-Lys-Gly-Asp-Trp-containing peptides, compared to D-aspartic acid replacements, that lead to inactive compounds [12].

Another point of interest concerns the superimposition of the groups marked ‘hydrophobic’ or ‘aromatic’ for each molecule depicted in Figure 5. It suggests the presence of a hydrophobic site in the binding pocket, corresponding to the location where ‘hydrophobic’ or ‘aromatic’ groups match. In addition, all the carbon atoms bearing a carboxylic acid group, marked ‘negative 1’ in Figure 5, superimpose well. This is in total agreement with the absence of activity for compounds in which aspartic acid is replaced by glutamic acid [12].

All these structural agreements favor a unique structure for inhibiting integrin $\alpha_{IIb}\beta_3$ selectively. Interestingly enough, the four non-peptide compounds were developed separately and from different backgrounds and perspectives. Yet, they all show strikingly similar structural features, in particular for the relative positioning of both the hydrophobic/aromatic and the charged moieties. This result is pivotal, as it implies only one way of inhibiting integrin $\alpha_{IIb}\beta_3$ specifically. This receptor shares with the integrin $\alpha_v\beta_3$ the β_3 subunit, which is known to contain binding domains for

the Arg-Gly-Asp motif, whereas the α_{IIb} subunit does not recognize Arg-Gly-Asp-containing ligands, but only the His^{400}-His^{401}-Leu^{402}-Gly^{403}-Gly^{404}-Ala^{405}-Lys^{406}-Gln^{407}-Ala^{408}-Gly^{409}-Asp^{410}-Val^{411} sequence located at the C-terminal part of the γ-chain of fibrinogen [46]. In addition, it would seem that the binding domains of β_3 and α_{IIb} are either common or mutually exclusive [47]. Further investigations on the binding of platelets to fibrinogen during the platelet aggregation process have concluded that: (i) the Arg-Gly-Asp sequences of the α-chain of fibrinogen are not required for aggregation, (ii) only the C-terminal part of the γ-chain of fibrinogen is required for platelet to bind fibrinogen, and (iii) the Arg-Gly-Asp sequence of fibrinogen binds to $\alpha_v\beta_3$ [47–50].

We, therefore, propose, in agreement with the previous hypothesis put forward in Reference 38, that all compounds in Figure 5 bind to the binding site of the C-terminal part of the γ-chain of fibrinogen, rather than the Arg-Gly-Asp-binding site located on the β_3 subunit. This assertion is supported by both the unique structure of specific antagonists of $\alpha_{IIb}\beta_3$, and the lack of common structural features with other Arg-Gly-Asp-containing peptides, such as echistatin, in which the Arg-Gly-Asp sequence adopts a cup-shaped motif without any hydrophobic moiety between the charged groups. Yet, despite the absence of a full structural study of the C-terminal part of the γ-chain of fibrinogen, which would require extensive CPU time on account of the flexibility of this peptide, structural similarities between the binding sequence of fibrinogen and the Lys-Gly-Asp loop of barbourin should be noted, i.e. (i) they both contain a lysine and an aspartic acid residue, (ii) they both include an alanine amino acid near the latter charged groups, and (iii) the side chain of residues Trp^{54} and Gln^{407} contains a polar amino functional group. These features tend to favor the inhibition by barbourin and other antagonists specific of integrin $\alpha_{IIb}\beta_3$ of the binding site of the γ-chain of fibrinogen, rather than the binding site characteristic of the Arg-Gly-Asp sequence.

Conclusions

In this paper, the structure and dynamics of barbourin in an aqueous solution were investigated using large-scale MD simulations. From this study, a model three-dimensional structure is proposed. A noteworthy feature described herein lies in the structural similarities between barbourin and other disintegrins elucidated by NMR spectroscopy.

The particularity of barbourin is rooted in its Lys-Gly-Asp sequence, hypothesized to confer on it a high selectivity for integrin $\alpha_{IIb}\beta_3$, in comparison with other Arg-Gly-Asp-dependent integrins. The results of the MD simulations presented in this paper reveal that the Lys-Gly-Asp binding sequence is well exposed towards the aqueous environment and is located at the top of a mobile loop. The various conformations of this sequence extracted from the complete trajectory show little differences, except for the flexible side chain of lysine. The importance of the tryptophan residue following either the Arg-Gly-Asp or the Lys-Gly-Asp sequence, for the selectivity of disintegrins, is emphasized in the remarkable superimposition of residue Trp^{54} with moieties of peptidomimetics specific towards $\alpha_{IIb}\beta_3$, that include both a positively charged group and a hydrophobic/aromatic one. In addition, we observe an equally remarkable match for both Ala^{50} and Asp^{53} with, respectively, the hydrophobic/aromatic and the negatively charged groups of the same peptidomimetics.

Farrell et al. [48] and Liu et al. [49] have shown that the sequence of the γ-chain of the C-terminal part of fibrinogen is the only protein domain that conditions platelet aggregation. In the light of the present work, we propose that barbourin and other peptidomimetics investigated here bind to the site of the C-terminal part of fibrinogen, and not to those specific towards the Arg-Gly-Asp sequence, located in several agonists of integrins. This assumption is in complete agreement with the high selectivity exhibited by all molecules examined herein for integrin $\alpha_{IIb}\beta_3$.

Acknowledgements

H.M. is indebted to the Laboratoire de Chimie théorique for his doctoral fellowship. We acknowledge the Centre Charles Hermite in Nancy, France, for the provision of generous computer time on the SGI Origin 2000 machine.

References

1. Seymour, J.L., Henzel, W.J., Nevins, B., Stults, J.T. and Lazarus, R.A., J. Biol. Chem., 265 (1990) 10143.
2. Huang, T.F., Peng, H.C., Peng, I.S., Teng, C.M. and Ouyang, C., Arch. Biochem. Biophys., 298 (1992) 13.

3. Gould, R.J., Polokoff, M.A., Friedman, P.A., Huang, T.F., Holt, J.C., Cook, J.J. and Niewiarowski, S., Proc. Soc. Exp. Biol. Med., 195 (1990) 168.
4. Cox, D., Aoki, T., Seki, J., Motoyama, Y. and Yoshida, K., Med. Res. Rev., 14 (1994) 195.
5. Hynes, R.O., Cell, 69 (1992) 11.
6. Phillips, D.R., Charo, I.F. and Scarborough, R.M., Cell, 65 (1991) 359.
7. Pierschbacher, M.D. and Ruoslahti, E., Proc. Natl. Acad. Sci. USA, 81 (1984) 5985.
8. Pierschbacher, M.D. and Ruoslahti, E., Nature 309 (1984) 5963.
9. Lazarus, R.A. and McDowell, R.S., Curr. Opin. Biotechnol., 4 (1993) 438.
10. Scarborough, R.M., Rose, J.W., Hsu, M.A., Phillips, D.R., Fried, V.A., Campbell, A.M., Nannizzi, L. and Charo, I.F., J. Biol. Chem., 266 (1991) 9359.
11. Scarborough, R.M., Rose, J.W., Naughton, M.A., Phillips, D.R., Nannizzi, L., Arfsten, A., Campbell, A.M. and Charo, I.F., J. Biol. Chem., 268 (1993) 1058.
12. Scarborough, R.M., Naughton, M.A., Teng, W., Rose, J.W., Phillips, D.R., Nannizzi, L., Arfsten, A., Campbell, A.M. and Charo, I.F., J. Biol. Chem., 268 (1993) 1066.
13. Cheng, S., Craig, W.S., Mullen, D., Tschopp, J.F., Dixon, D. and Pierschbacher, M.D., J. Med. Chem., 37 (1994) 1
14. Muller, G., Gurrath, M. and Kessler, H., J. Comput.-Aided Mol. Design., 8 (1994) 709
15. Craig, W.S., Cheng, S., Mullen, D.G., Blevit, J. and Pierschbacher, M.D., Biopolymers, 37 (1995) 157
16. Tselepis, V.H., Green, L.J. and Humphries, M.J., J. Biol. Chem., 272 (1997) 21341
17. Phillips, D.R. and Scarborough, R.M., Am. J. Cardiol., 80 (1997) 11B.
18. Oshikawa, K. and Terada, S., J. Biochem., 125 (1999) 31
19. Pearson, W.R. and Lipman, D.J., Proc. Natl. Acad. Sci. USA, 85 (1988) 2444.
20. Pearson, W.R., Methods Enzymol., 183 (1990) 63.
21. Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F. and Weng, J., In Allen, F.H., Bergerhoff, G. and Sievers, R. (Eds), Protein Data Bank, in Crystallographic Databases-Information Content, Software Systems, Scientific Applications, Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester, 1987, pp. 107–132.
22. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer Jr., E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
23. Senn, H. and Klaus, W., J. Mol. Biol., 232 (1993) 907.
24. Klaus, W., Broger, C., Gerber, P. and Senn, H., J. Mol. Biol., 232 (1993) 897.
25. Dauber-Osguthorpe, P., Roberts, V.A., Osguthorpe, D.J., Wolff, J., Genest, M. and Hagler, A.T., Proteins, 4 (1988) 31.
26. Homology, Molecular Simulations Inc., San Diego, CA, 1997.
27. Saudek, V., Atkinson, R.A. and Pelton, J.T., Biochemistry, 30 (1991) 7369.
28. Atkinson, R.A., Saudek, V. and Pelton, J.T., Int. J. Pept. Protein. Res., 43 (1994) 563.
29. Adler, M., Lazarus, R.A., Dennis, M.S. and Wagner, G., Science, 253 (1991) 445.
30. The force field parameters for the CVFF water model are: $q_O = -0.82$, $q_H = 0.41$, A(O) = 629358 kcal mol^{-1} Å^{12}, B(O) = 625.5 kcal mol^{-1} Å^6, $r_{OH} = 0.96$ Å, HOH = 104.5° (with the usual notations).
31. Ewald, P., Ann. Phys., 64 (1921) 253.
32. Brown, D., Minoux, H. and Maigret, B., Comput. Phys. Commun., 103 (1997) 170.
33. Brown, D. The gmq User Manual, 1998; available at: http://www.lctn.u-nancy.fr/Chercheurs/David.Brown
34. Brown, D., Clarke, J.H.R., Okuda, M. and Yamazaki, T., Comput. Phys. Commun., 83 (1994) 1; *ibid* Erratum Comput. Phys. Commun., 86 (1995) 312.
35. Frish, M.J., Trucks, G.W., Schlegel, H.B., Gill, P.M.W., Johnson, B.G., Robb, M.A., Cheeseman, J.R., Keith, T., Peterson, G.A., Montgomery, J.A., Raghavachari, K., Al-Laham, M.A., Zakrzewski, V.G., Ortiz, J.V., Foresman, J.B., Peng, C.Y., Ayala, P.Y., Chen, W., Wong, M.W., Andres, J.L., Replogle, E.S., Gomperts, R., Martin, R.L., Fox, D.J., Binkley, J.S., Defrees, D.J., Baker, J., Stewart, J.P., Head-Gordon, M., Gonzales, C. and Pople, J.A., Gaussian 94. Gaussian Inc., Pittsburgh, PA, 1995.
36. Minoux, H., Brown, D., Chipot, C. and Maigret, B., unpublished results.
37. Greenspoon, N., Hershkoviz, R., Alon, R., Varon, D., Shenkman, B., Marx, G., Federman, S., Kapustina, G. and Lider, O., Biochemistry, 32 (1993) 1001.
38. Minoux, H., Moitessier, N., Chapleur, Y. and Maigret, B., J. Comput.-Aided Mol. Design, 12 (1998) 533.
39. Hartman, G.D., Egbertson, M.S., Halczenko, W., Laswell, W.L., Duggan, M.E., Smith, R.L., Naylor, A.M., Manno, P.D., Lynch, R.J., Zhang, G., Chang, C.T.-C. and Gould, R.J., J. Med. Chem., 35 (1992) 4640.
40. Alig, L., Edenhofer, A., Hadvary, P., Hurzeler, M., Knopp, D., Muller, M., Steiner, B., Trzeciak, A. and Weller, T.J., J. Med. Chem., 35 (1992) 4393.
41. Cox, D., Aoki, T., Seki, J., Motoyama, Y. and Yoshida, K., Thromb. Haemostas., 69 (1993) 706.
42. Marcinkiewicz, C., Vijay-Kumar, S., McLane, M.A. and Niewiarowski, S., Blood, 90 (1997) 1565.
43. Zablocki, J.A., Miyano, M., Garland, R.B., Pireh, D., Schretzman, L., Rao, S.N., Lindmark, R.J., Panzer-Knodle, S.G., Nicholson, N.S., Taite, B.B., Salyers, A.K., King, L.W., Campion, J.G. and Feigen, L.P., J. Med. Chem., 36 (1993) 1811.
44. Chipot, C., Maigret, B., Pearlman, D.A. and Kollman, P.A., J. Am. Chem. Soc., 118 (1996) 2998.
45. Minoux, H. and Chipot, C., J. Am. Chem. Soc., 121 (1999) 10366.
46. Loftus, J.C., Smith, J.W. and Ginsberg, M.H., J. Biol. Chem., 269 (1994) 25235.
47. Rahman, S., Lu, X., Kakkar, V.V. and Authi, K.S., Biochem. J., 312 (1995) 223.
48. Farrell, D.H., Thiagarajan, P., Chung, D.W. and Davie, E.W., Proc. Natl. Acad. Sci. USA, 89 (1992) 10729.
49. Liu, Q., Matsueda, G., Brown, E. and Frojmovic, M., Biochim. Biophys. Acta, 1343 (1997) 316.
50. Smith, J.W., Ruggeri, Z.M., Kunicki, T.J. and Cheresh, D.A., J. Biol. Chem., 265 (1990) 12267.