Journal of Computer-Aided Molecular Design, 8 (1994) 479–490
ESCOM

J-CAMD 263

Rational design of hirulog-type inhibitors of thrombin

Ursula Egner*, Georg-Alexander Hoyer and Wolf-Dieter Schleuning

Research Laboratories of Schering AG Berlin, Müllerstrasse 170–178, D-13342 Berlin, Germany

Received 10 December 1993
Accepted 24 May 1994

Key words: Molecular structure-based inhibitor design; Molecular mechanics and dynamics; Bifunctional inhibitors; Thrombin

SUMMARY

The two crystal structures of thrombin complexed with its most potent natural inhibitor hirudin and with the active-site inhibitor D-Phe-Pro-Arg-CH_2Cl [Rydel, T.J. et al., J. Mol. Biol., 221 (1991) 583; Bode, W. et al., EMBO J., 8 (1989) 3467] were used as a basis to design a new inhibitor, combining the high specificity of the polypeptide hirudin with the simpler chemistry of an organic compound. In the new inhibitor, the C-terminal amino acid residues 53–65 of hirudin are linked by a spacer peptide of four glycines to the active-site inhibitor NAPAP (N^{α}-(2-naphthyl-sulfonyl-glycyl)-DL-*p*-amidinophenylalanyl-piperidine). Energy minimization techniques served as a tool to determine the preferred configuration at the amidinophenylalanine and the modified piperidine moiety of the inhibitor. The predictions are supported by the interaction energies determined for D- and L-NAPAP in complex with thrombin, which are in good agreement with experimentally determined dissociation constants. The conformational flexibility of the linker peptide in the new inhibitors was investigated with molecular dynamics techniques. A correlation between the P1' position and the interactions of the linker peptide with the protein is suggested. Modifications of the linker peptide are proposed based on the distribution of its main-chain torsion angles in order to enhance its binding to thrombin.

INTRODUCTION

Blood coagulation proceeds from the interaction of a family of regulatory proteases, protease inhibitors and cofactors, cooperating in the activation of thrombin, the principal clotting enzyme. Human α-thrombin, a serine protease, consists of a 36-residue A-chain and a 259-residue B-chain, linked by a disulfide bridge. Many serine protease inhibitors are known, but only few of them specifically inhibit thrombin. Hence, the unique properties of thrombin can be incorporated in inhibitor design, since α-thrombin possesses, in contrast to other serine proteases, at least three independent binding sites for substrate, inhibitor and effector molecules [1]. The details of the

*To whom correspondence should be addressed.

active site and an inhibitor binding site, the so-called anion binding or fibrinogen recognition exosite, were elucidated by X-ray crystallography of thrombin–inhibitor complexes. A third, putative binding site was deduced from calculations of the electrostatic potential of thrombin and most likely represents the heparin binding site [2,3]. The active center and the fibrinogen recognition site appear to be in close vicinity, as observed in the X-ray structure of the complex between thrombin and recombinant hirudin [4,5]. Hirudin, a polypeptide of 65 amino acid residues from the leech *Hirudo medicinalis*, is the most potent thrombin inhibitor known so far [6]. It forms a very tight, noncovalent complex with thrombin ($K_d \approx 10^{-13}$ M). Its N-terminal three residues interact with the active site of thrombin, while its C-terminal residues 53–65 bind to the so-called anion binding or fibrinogen recognition exosite. As these properties are unique to hirudin, they have attracted much interest with respect to potential clinical applications in anticoagulatory therapy.

Various peptides derived from the hirudin C-terminal sequence have been investigated as anticlotting agents. Some of these, the hirulogs and hirutonins, mimic the distinctive mechanism of hirudin by simultaneously occupying both the active site and the anion binding exosite [7,8]. The binding mode of some of these inhibitors to thrombin has been confirmed by X-ray structure determination, although in some complexes no or weak electron density was found for the linker peptides [9,10]. Hirulogs are derived from the active-site inhibitor D-Phe-Pro-Arg-Pro, the C-terminal residues of hirudin 53–65, linked by a spacer peptide of variable length. The advantage of hirudin and hirulogs over active-site inhibitors is their high degree of specificity.

Active site-directed thrombin inhibitors have been thoroughly investigated [11–13]. Tripeptides of the type hydrophobic residue-Pro-Arg-aldehyde (or other nonhydrolyzable groups) have been shown to inhibit thrombin activity. Other efficient inhibitors of thrombin are for example PPAC (D-Phe-Pro-Arg-chloromethylketone), MQPA ((2*R*,4*R*)-4-methyl-1-[N^{α}-(3-methyl-1,2,3,4-tetrahydro-8-quinolinsulfonyl)-L-arginyl]-2-piperidine-carbonic acid) and NAPAP (N^{α}-(2-naphthylsulfonyl-glycyl)-DL-*p*-amidinophenylalanyl-piperidine), whose binding modes are known from X-ray structural analyses of complexes with human or bovine α-thrombin [14,15] or the homologous serine protease trypsin [16]. The disadvantage of active-site inhibitors in thrombolytic therapy is that they inhibit other serine proteases as well as thrombin, e.g. NAPAP inhibits thrombin (K_d = 6.6 nM for the racemic form [17]) and trypsin (K_d = 0.69 μM [16]). With the design of low-

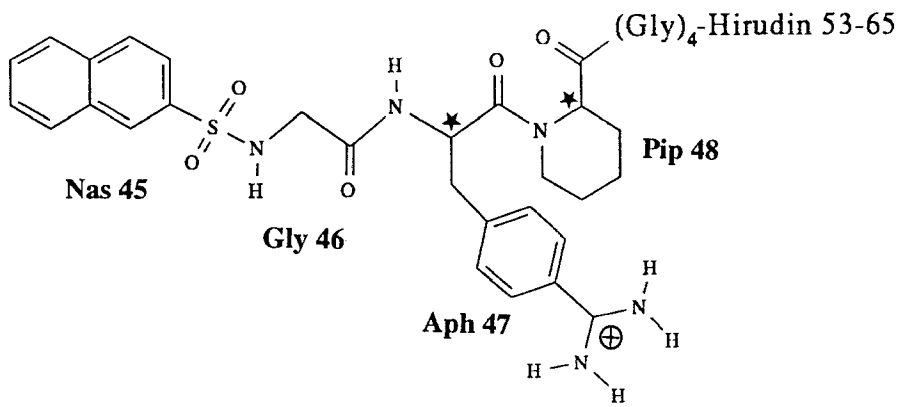

Fig. 1. Chemical structure of the 'NAPAP-hirulog'. Stereocenters discussed in the text are marked by asterisks.

molecular-weight compounds, like the bifunctional hirulog-type inhibitors, enhanced potency and specificity for thrombin is expected. Specific thrombin inhibitors may serve as useful agents to control disposition to vascular occlusive disease.

We used molecular modeling methods to investigate the possibility of replacing the active-site inhibitor D-Phe-Pro-Arg-Pro in hirulogs by NAPAP. The effect of stereochemistry in L- or D-amidinophenylalanine as well as in the piperidine-carbonic acid moiety of NAPAP was studied, since D-NAPAP has a 1000-fold reduced activity compared to its L-isomer [15]. The minimum size of the linker to bridge NAPAP with the C-terminal part of hirudin 53–65 was determined and the mobility of the linker peptide was investigated using molecular dynamics techniques. The results of this analysis provide the basis for a rational design of bifunctional inhibitors, consisting of a conventional active-site inhibitor combined with the C-terminal part of hirudin, and for modifications of the linker peptide to enhance binding to thrombin.

MATERIALS AND METHODS

The atomic coordinates of the complexes of human α-thrombin with PPAC and hirudin were provided by W. Bode (Max Planck Institute of Biochemistry, Martinsried). They were refined to maximum resolutions of 0.19 and 0.23 nm, respectively, and crystallographic R-factors of 0.156 and 0.173. Both thrombin structures were compared to investigate the influence of the so-called autolysis loop 145–150 upon inhibitor binding (our numbering of thrombin residues is according to topological equivalence with the tertiary structure of chymotrypsinogen [18]).

NAPAP–thrombin interactions were modeled with either a D- or L-amidinophenylalanine, with interactions similar to the crystal structures of the complex between human thrombin or trypsin and D-NAPAP, respectively [11–13]. Crystallographically determined water molecules were included in the study, if they did not overlap with the inhibitors.

For the modeling of a 'NAPAP-hirulog', the piperidine moiety of NAPAP was changed to piperidine-carbonic acid. Four stereoisomers of the 'NAPAP-hirulog' were considered with D- or L-conformation at the C^{α}-atoms of the amidinophenylalanine and piperidine-carbonic acid moieties (residue numbers 47 and 48, Fig. 1). Throughout this study, they will be named (D-47,L-48) for a molecule with D-conformation at position 47 and L-conformation at position 48. (D-47, D-48), (L-47,L-48) and (L-47,D-48) are used analogously.

Energy minimization and molecular dynamics details

The models were constructed using the molecular modeling package InsightII (Biosym Technologies, Inc.) running on an SGI workstation. Energy minimizations and molecular dynamics (MD) calculations were performed with Discover (Biosym Technologies, Inc.) on a Siemens Supercomputer S200. The amino acid residues aspartic acid, glutamic acid, lysine and arginine were taken in their charged forms, with the exception of some residues at the surface of thrombin (Asp^{21}, Glu^{127}, Asp^{243}) to obtain neutral protein–inhibitor complexes. All complexes comprising inhibitor, protein and crystallographically determined water molecules were soaked with a 0.3 nm layer of water molecules, resulting in an average number of 7000 atoms (all-atom model). Energy minimizations were performed for all structures using a cutoff radius of 1.5 nm and a switching function of 0.1 nm, with recalculation of the nearest-neighbour list for nonbonded interactions every 20 steps. The conjugate gradient algorithm was applied until the maximum derivative was

less than 4.2 kJ mol^{-1} nm^{-1}, with an rms deviation less than 0.4 kJ mol^{-1} nm^{-1}. A uniform dielectric constant of 1.0 was used for all calculations. The charge distributions for the inhibitor molecules were calculated with the InsightII software. An MD calculation was carried out for 100 ps at 300 K to investigate the conformational flexibility of the linker peptide; sample structures were saved every 1 ps. Interaction energies (E_{inter}) of the inhibitors and thrombin were calculated with the energies defined as $E_{inter} = E_{PI} - (E_P + E_I)$, where E_{PI} is the energy of the protein–inhibitor complex, E_P that of the protein and E_I that of the inhibitor. The energies of the protein and inhibitor were determined in the conformation adopted by these molecules in the protein–inhibitor complex.

Force field parameters

The consistent valence force field of Hagler and Lifson [19] was used throughout the calculations. As the automatic parameter assignment of the sulfonyl moieties of NAPAP did not sufficiently reproduce the bond lengths and angles found in the crystal structure, these data and the charge distributions were taken from ab initio calculations on sulfonamides [20]. The force constants were changed to fit the consistent valence force field. They were estimated for all bonds to 100 kcal mol^{-1} $Å^{-2}$ and for all angles to 30 kcal mol^{-1}, similar to the force constants found in the force field for interactions with sulfur atoms.

RESULTS AND DISCUSSION

NAPAP–thrombin complex

A comparison of the energy-minimized NAPAP structure with the ab initio bond lengths and angles for the sulfonamide moiety revealed that the potential parameters sufficiently reproduce the geometry of this group. The deviations found in bond lengths and angles (on average 0.001 nm and 5°, respectively) justify the application of these parameters to the modeling of these thrombin inhibitors.

NAPAP was fitted into the active site of thrombin taking into account hydrogen bonds and conformations of the naphthyl and piperidine moieties published for crystal structures of complexes of NAPAP with thrombin and trypsin, respectively. The naphthyl ring was placed near T-Trp^{215} (amino acid residues referring to thrombin are marked with 'T-', those of the inhibitors with 'I-'), although some mobility seems to be allowed for this ring. Obviously, the electron density in this area of the complex is weak or difficult to interpret, as Bode et al. [14] found high temperature factors for the naphthyl group in the complex with trypsin and bovine thrombin and Banner and Hadváry [15] have refined two alternative positions of NAPAP bound to human thrombin. In the alternative orientation, the naphthyl ring is rotated by 180° about the bond to sulfur.

The orientation of the piperidine moiety seems to be similar in all crystal structures, the piperidine ring being in contact with the naphthyl ring. In this conformation, the ϕ,ψ angles of the amidinophenylalanine (Aph) 47 in the complex NAPAP–bovine thrombin are 122° and –83°, respectively [14]. An alternative arrangement of the piperidine group can be modeled with ϕ,ψ angles of I-Aph 47 of 57° and 58°, respectively, which results in a more open conformation of NAPAP compared to the compact form found in crystal structures. In this orientation, the piperidine group points to the cleft between the active site and the anion binding exosite. The binding of the piperidine group is influenced by a loop comprising residues T-60A–60I. In comparison with the chymotrypsinogen structure this is an insertion loop, which protrudes into the

vicinity of the active site. The side chains of T-Tyr^{60A} and T-Trp^{60D} are in contact with the piperidine moiety; their conformations differ between the models and in the crystal structures of thrombin. Differences also occur in the trace of the main chain: distances of up to 0.1 nm between equivalent C^{α} atoms of this loop are observed in crystal structures of thrombin and in the modeled complexes. In the crystal of the complex between NAPAP and bovine α-thrombin [14], this loop is in close contact to adjacent molecules due to the space group symmetry, thus favouring the compact conformation of the inhibitor. In solution, where the protein is allowed more flexibility, the existence of the open conformation cannot be ruled out.

The stereoisomers of NAPAP

Interaction energies were calculated for thrombin complexes with D- and L-NAPAP in the open orientation and in the compact arrangement found in the crystal structures. The following rank order is found: 'open D' < 'compact D' < 'open L' < 'compact L', with relative values of 0 < 64 < 93 < 128 kJ mol^{-1}. Irrespective of the conformation of NAPAP, open or closed, the D-conformer is preferred. This result is in good agreement with experimental data, as the D-isomer of NAPAP is 1000 times more active than the L-isomer [15].

The influence of the loop 145–150

The three-dimensional structures of thrombin in complex with hirudin and PPAC are very similar, with an rms deviation of 0.03 nm for 450 N,C^{α},C-atoms in β-sheets. Differences are mainly confined to the loop comprising residues 145–150, also called the autolysis loop because of autolytic cleavage at T-Lys^{149E} [9]. Upon hirudin binding, this loop is shifted up to 0.8 nm with respect to its conformation in the PPAC–thrombin structure. X-ray analysis of other thrombin–inhibitor complexes suggests that this loop is disordered, which might be due to crystal packing effects [3,9,12]. To study the influence of this loop upon binding of NAPAP, D-NAPAP was modeled in complex with both thrombin structures and the energies were minimized. The rms deviations for 450 N,C^{α},C-atoms in β-sheets between starting model and energy-minimized structures were only 0.06 and 0.05 nm, respectively. The largest differences occur in surface regions, for example, the distance between corresponding C^{α}-atoms for T-Ala^{149A} in the autolysis loop is 0.15 nm and for T-Ser^{1E} it is 0.12 nm.

No direct van der Waals contacts can be found between NAPAP and amino acid residues from the autolysis loop. The nearest distance is observed between I-Gly^{46} C^{α} of NAPAP and T-Glu^{146} C^{δ}: 0.67 nm for thrombin originating from the complex with hirudin and 0.62 nm for thrombin originating from the complex with PPAC (nearest distance between I-Gly^{46} C^{α} and T-Trp^{148} $N^{\varepsilon 1}$). As the autolysis loop seems to have no influence upon inhibitor binding and as the C-terminal residues of hirudin are included in 'NAPAP-hirulogs', the structure of thrombin complexed with hirudin was used for all further calculations.

'NAPAP-hirulog'–thrombin complexes

Models of the 'NAPAP-hirulog'–thrombin complexes were built based on the assumption that NAPAP and the C-terminal part of hirudin do not change their binding modes compared to the interactions observed in the NAPAP–trypsin/NAPAP–thrombin and the hirudin–thrombin complexes. Only the orientation of the piperidine-carbonic acid had to be changed within the active site compared to the conformation of piperidine in the crystal structures, in order to allow the

glycines used as spacers to point to the C-terminal residues of hirudin. The conformation of NAPAP in the 'NAPAP-hirulogs' resembles that of the open form discussed in the previous section.

Hydrogen bonds to Ser^{216} of thrombin

In the crystal structures of complexes between peptidic active-site inhibitors and serine proteases, similar patterns of hydrogen bonds are observed between inhibitor and protein, i.e., two with the amino and carbonyl groups of Gly^{216} (of the proteins) in an antiparallel, slightly twisted manner (H-bonds I and II) and, in case of thrombin and trypsin, with the side-chain atoms of T-Asp^{189}, situated at the bottom of the specificity pocket (H-bond III).

Hydrogen bonds I and III are always present in the modeled complexes of thrombin with all stereoisomers of the 'NAPAP-hirulogs'. The formation of hydrogen bond II depends on the ϕ,ψ torsion angles of Gly^{46} of the NAPAP moiety of the designed hirulogs. With ϕ,ψ angles of $-90°$ and 163° (data from the (D-47,D-48)-isomer), respectively, hydrogen bond II is formed between the peptide-NH of I-Gly^{46} of the NAPAP moiety and T-Gly^{216}, with a length of 0.19 nm and an angle of 169°. If this hydrogen bond is not fixed during the energy minimizations, the ϕ,ψ torsion angles of I-Gly^{46} change to 43° and 86°, lengthening the distance between NH of I-Gly^{46} of NAPAP and the carbonyl oxygen of T-Gly^{216} to 0.5 nm.

The complex with hydrogen bond I is favoured in the energy minimizations, as the interaction energy of this complex is stabilized by 66.5 kJ mol^{-1} in comparison with the other complex with H-bonds I and II. The interactions of both inhibitors with thrombin were analyzed with respect to their stability.

The largest differences in the nonbonded energies are observed for residues 45 and 46 (Table 1), with differences of 17 and 30 kJ mol^{-1}, respectively. Differences of about 10 kJ mol^{-1} are found for residues 49–51, while the energies of residues 47 and 52–65 are very similar in both complexes. The van der Waals interactions are always more favourable in the complex with H-bond I, while

TABLE 1
NONBONDED ENERGIES OF THE 'NAPAP-HIRULOG' IN COMPLEX WITH THROMBIN WITH RESPECT TO THE NUMBER OF HYDROGEN BONDS TO GLY^{216} OF THROMBIN

	Complex with H-bond I[a]			Complex with H-bonds I,II[b]		
Residue number	van der Waals energy	Coulomb energy	Total	van der Waals energy	Coulomb energy	Total
45	169.7	−50.0	119.7	178.1	−41.2	136.9
46	−16.8	−81.5	−98.3	−8.8	−59.2	−68.0
47	121.0	−212.9	−91.9	131.0	−221.3	−90.3
48	18.9	71.8	90.7	21.4	79.0	100.4
49	−6.7	29.4	22.7	−1.7	34.4	32.7
50	−2.1	4.6	2.5	4.6	7.1	11.7
51	1.7	13.0	14.7	4.6	7.1	11.7
52	−8.0	−3.8	−11.8	−10.1	−2.1	−12.2
53–65	177.2	−2445.2	−2268.0	178.1	−2450.7	−2272.6

Energies are given in kJ mol^{-1}.

[a] A hydrogen bond between the 'NAPAP-hirulog' and the amino group of T-Gly^{216} of thrombin is observed.

[b] Hydrogen bonds between the 'NAPAP-hirulog' and the amino and carbonyl groups of T-Gly^{216} of thrombin are observed.

the Coulomb energies of residues 47 and 51 are more favourable in the complex with both H-bonds. With respect to the NAPAP moiety (residues 45–48), both the van der Waals and Coulomb energy are each about 30 kJ mol^{-1} more favourable in the complex with H-bond I than in the complex with H-bonds I and II. Interestingly, the Coulomb energy for I-Gly^{46} as the H-bond partner of T-Gly^{216} is 20 kJ mol^{-1} more favourable in the complex with H-bond I. In the 3D structure of this complex, the amino group of I-Gly^{46} is involved in an extensive hydrogen-bond network with the surrounding water. Water molecules bridge the amino group of I-Gly^{46} with the side chain of T-Glu^{192} of thrombin and the sulfonyl group of the NAPAP moiety (Fig. 2). The interactions of hirudin with thrombin are still more favourable than those of the 'NAPAP-hirulogs'. The van der Waals and Coulomb energies of hirudin are –736.9 and –4555.8 kJ mol^{-1}, respectively. The differences in comparison with the energies of the 'NAPAP-hirulogs' can be explained from the differing number of contacts of these inhibitors with thrombin. Fourty-four amino acid residues of hirudin are at van der Waals distance to thrombin, while 20 (out of 21) residues of the 'NAPAP-hirulogs' are in contact with the protein.

Stereoisomers of 'NAPAP-hirulogs'

The interaction energies of all stereoisomers of the 'NAPAP-hirulogs' were determined from complexes with a hydrogen bond to the peptide NH of T-Gly^{216} (H-bond I). The NAPAP-hirulog with the lowest (relative) interaction energy is the (L-47,D-48) isomer, the energies of the (D-47,L-48) and (D-47,D-48) molecules are less favourable by 42 kJ mol^{-1} and the weakest binding is found for the (L-47,L-48) isomer with an interaction energy of 105 kJ mol^{-1}. As these energies are higher than that of hirudin in complex with thrombin, we expect the 'NAPAP-hirulogs' to be less active than hirudin. This has to be proven with the synthesis of the compounds.

The stereochemical preference of the amidinophenylalanine moiety has changed between NAPAP and a 'NAPAP-hirulog'. While D-NAPAP is preferred in a complex with thrombin, the L-configuration seems to be favoured in a complex between thrombin and a 'NAPAP-hirulog'. This preference at the P1/P1' sites of the inhibitors (notation according to Schechter and Berger [21], corresponding to residues 47/48 in this study) seems to depend on the nature of the residues

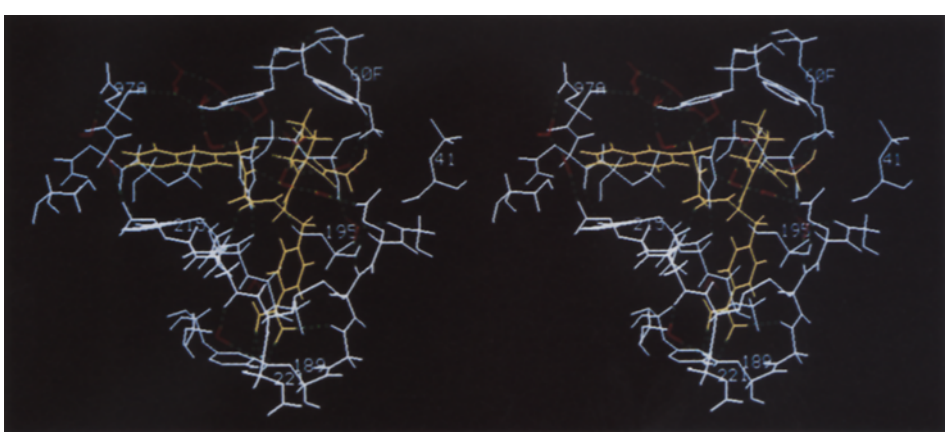

Fig. 2. Environment of the NAPAP moiety of the (D-47,D-48) 'NAPAP-hirulog' (in yellow), complexed with thrombin (in white). Water molecules are coloured red.

at these positions and perhaps on those of the linker peptides. DiMaio et al. [22] found in experiments with peptides based on the sequence of hirudin 53–65 that configurational inversion of Pro^{48} to D-Pro^{48} (topologically equivalent to Pip^{48}) in a [D-Phe^{45},Arg^{47},Pro^{48}] hirudin 53–65 results in an analog with an increased IC_{50} value, i.e., 45 nM instead of 4.1 nM for L-Pro^{48}. Although piperidine-carbonic acid is structurally related to proline, the increase in ring size might influence the stereochemistry at the P1' position of the inhibitor.

A superposition of the complexes shows that the choice of the 'NAPAP-hirulog' isomers does not influence the structure of thrombin, the maximum rms deviation for 450 N,C^{α},C-atoms in β-sheets being 0.07 nm. The smallest rms deviation was observed between the (D-47,L-48) and (L-47,L-48) isomers, i.e., 0.01 nm.

Within the inhibitors, residues 46–49 and 53–65 superimpose quite well, while the glycines in the linker peptide can adopt different conformations due to the size of the cleft between the active site and the anion binding exosite of thrombin. This observation is supported by the results of the crystal structures of some hirulog–thrombin complexes, in which the spacer peptide is disordered and only poor electron density is found for this portion of the molecule [9]. A stereopresentation of the complex between the (L-47,D-48) isomer and thrombin is shown in Fig. 3. The 'NAPAP-hirulog' extends for 3.9 nm across the thrombin molecule. The active site and the anion binding exosite can be seen at the top right and left part of the molecule, respectively.

Linker peptide

Due to their conformational flexibility, we have chosen glycine residues for the linker peptide in all 'NAPAP-hirulog' isomers. The distance which has to be spanned between the carbonic acid moiety of I-Pip^{48} and the C^{α}-atom of I-Asn^{53} of the hirudin C-terminal residues was measured to be 1.68 nm. Therefore, a minimum number of four glycines is required to bridge the active site with the anion binding exosite. This is consistent with experimental results, as inhibitors with a linker of two or three glycines loose the bifunctionality of the inhibition [7].

In all energy-minimized complexes, hydrogen bonds between the glycines and T-Arg^{73} and T-Glu^{192} of thrombin are observed. Since the cleft between the active site and anion binding exosite

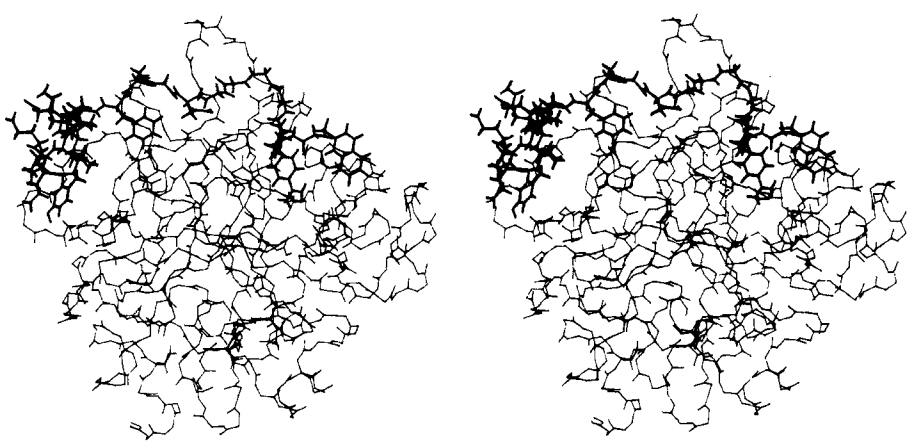

Fig. 3. Stereopresentation of the complex between (L-47,D-48) 'NAPAP-hirulog' (in bold) and thrombin (backbone atoms only).

of thrombin is quite large, the glycines have ample space to adopt different conformations. In the four 'NAPAP-hirulog' isomers, differences of up to 0.37 nm between equivalent atoms are found (Fig. 4). A 100 ps MD simulation was carried out on the (L-47,D-48) and the (D-47,D-48) hirulog to reveal the significance of this finding and to explore possible conformations of the glycines in the cleft.

The hydrogen bonds between the glycines of the hirulog and the carbonyl oxygen of T-Glu^{192} could be confirmed in both dynamics simulations. Evaluating the data of the last 50 ps, i.e., 50 different structures, and applying a distance criterion for hydrogen bonds of 0.26 nm, hydrogen bonds are observed with the peptide-NH of I-Gly^{50} and I-Gly^{52} in nearly 80% (38 structures) and 40% (17 structures) of the (D-47,D-48) hirulog structures, respectively. In the (L-47,D-48) isomer, the peptide-NH of I-Gly^{49} and I-Gly^{50} are each found as acceptor atoms for the side-chain oxygens of T-Glu^{192} in 26 structures, as well as I-Gly^{51} in eight structures. An involvement of T-Glu^{192} in the interaction with the linker peptide is proposed from the X-ray analysis of a complex between thrombin and D-Phe-Pro-β-homoArg-$(Gly)_5$-desulfato-Tyr^{63}-hirugen [10]. A critical role of T-Glu^{192} is also suggested from mutation studies on protein C activation by the thrombomodulin–thrombin complex [23].

No interaction of the linker peptide and T-Arg^{73} could be found in the (D-47,D-48) isomer, whereas a hydrogen bond between the carbonyl oxygen of I-Gly^{52} of the (L-47,D-48) hirulog and the side chain of this residue was observed in nearly all of the evaluated structures with an average distance of 0.21 nm.

No direct hydrogen bonds between the amino and carbonyl groups of I-Gly^{49} and I-Gly^{51} of the (D-47,D-48) isomer and thrombin could be found during the simulation time, but a positionally stable water molecule was detected bridging the side chain of T-Lys^{60F} of thrombin and the carbonyl group of I-Gly^{49}. While the interaction between the water molecule and T-Lys^{60F} is observed continuously, the hydrogen bond to I-Gly^{49} is measured in 36 structures. A similar stable water molecule was also found for the (L-47,D-47) isomer; it stabilizes the orientation of the side chain of T-Lys^{60F} relative to I-Pip^{48} in nearly all evaluated structures.

To analyze the conformational flexibility of the glycine linker, the distribution of their ϕ,ψ angles was evaluated. In the (D-47,D-48) isomer, I-Gly^{51} and I-Gly^{52} seem to be more flexible, as their ϕ,ψ angles are distributed over a larger area of the Ramachandran map [24] than I-Gly^{49} and I-Gly^{50}. Taking the last 50 structures into account, the average values of ϕ,ψ are (values for the (L-47,D-48) isomer are given in square brackets): (82°,–123°) and [–86°,–94°] for I-Gly^{49}, (–120°,–94°) and [–108°,–84°] for I-Gly^{50}, (149°,94°) and [94°,–90°] for I-Gly^{51} and (–118°,98°) and [155°,135°] for I-Gly^{52}. With the exception of I-Gly^{50}, the ϕ,ψ angles of the other glycines differ, which impedes predictions of substitutes for these residues.

To find possible alternatives for the linker glycines, it is important to look at ϕ,ψ angles, i.e., whether the combination is allowed for amino acid residues with side chains or requires a flexible, partly aliphatic group. For the latter replacement, an ω-amino carbonic acid seems to be suitable; an example of such a substitution is I-Gly^{49} in the (D-47,D-48) isomer. This residue is in contact with aliphatic carbon atoms of I-Aph^{47} and T-Glu^{192} and its ϕ,ψ angles are energetically unfavourable for other amino acids.

I-Gly^{51} in the (L-47,D-48) hirulog might be replaced with other amino acid residues. For example, in hirutonins [20] this position is occupied by histidine. The interactions of I-51 in the cleft are influenced by the nature of the remaining linker peptide as well as the residue in the P1'

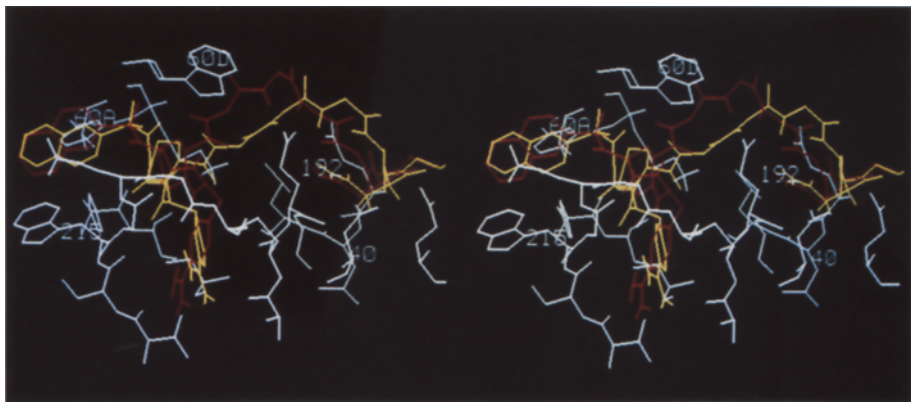

Fig. 4. Stereopresentation of the environment of residues 45–53 of the (L-47,D-48) (in red) and (D-47,D-48) (in yellow) 'NAPAP-hirulogs' in complex with thrombin (in white).

position of the inhibitor. The latter residue determines the conformation of the linker peptide, as it affects the orientation of the linker out of the active site. The exact position of the entrance of any linker into the cleft is influenced by the occupancy of the P1' position, either with proline or piperidine-carbonic acid, ketomethylene, etc.

In both isomers, the ϕ,ψ angles of I-Gly^{50} allow substitution by other amino acid residues. Evaluating the similarity coefficients between amino acid substitutions defined by Niefind and Schomburg [25], the most favoured substitution of glycine, with minimal changes in the main-chain torsion angles, is asparagine, followed by substitutions with serine and tyrosine. Taking into account the rotamer library of chi values [26], the replacement by asparagine enables a network of strong hydrogen bonds, as charged residues are involved. Hydrogen bonds are likely to be found between the side chains of the new I-Asn^{50} and T-Glu^{192}, as well as the side chain of T-Glu^{192} and the peptide-NH of I-Asn^{50}. Additionally, the orientation of I-Asn^{50} can be fixed by a water molecule, which bridges the carbonyl atoms of the side chains of I-Asn^{50} and T-Glu^{39}.

Although a substitution of I-Gly^{52} with respect to ϕ and ψ is allowed in the (D-47,D-48) isomer,

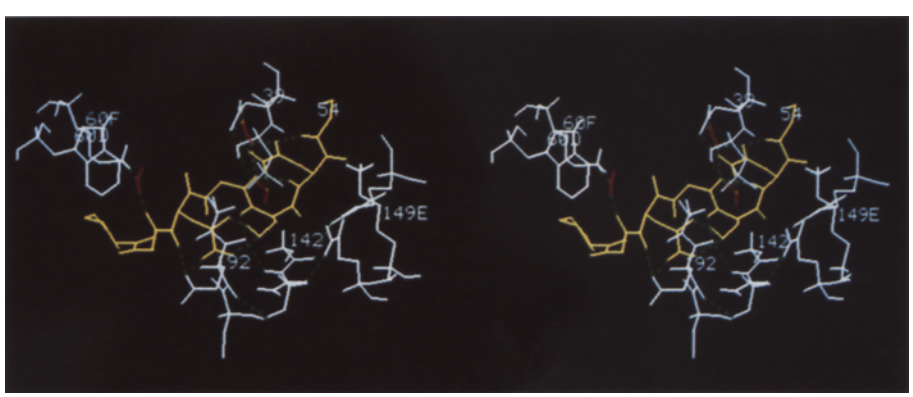

Fig. 5. Proposed binding of the Asn^{50}, D-Ser^{52} mutant of the (D-47,D-48) 'NAPAP-hirulog' (in yellow) to thrombin (in white). Water molecules are coloured red.

even a small side chain would influence the conformation of T-Asn^{143} of thrombin. The use of a D-amino acid at this position might be advantageous, e.g., D-Glu could accept hydrogen bonds from T-Asn^{143} and T-Gln^{151} of thrombin, and D-Ser could stabilize the conformation of the inhibitor by donating an intramolecular hydrogen bond to the side chain of I-Asn^{50}. Alternatively, D-Ser^{52} could accept a hydrogen bond from T-Gln^{151} of thrombin. Which amino acid is better suited to enhance activity has to be tested by synthesis of the proposed compound. Figure 5 shows one possible conformation of the (D-47,D-48) hirulog with Asn^{50} and D-Ser^{52}, including nearby thrombin residues.

CONCLUSIONS

The interaction of a bifunctional inhibitor, a 'NAPAP-hirulog', with thrombin has been studied by molecular modeling. The influence of stereochemistry of the NAPAP moiety was investigated and appropriate substitutions of the linker peptide 49–52 of the inhibitor were proposed. This linker was originally modeled as four glycines. Suitable candidates to replace these glycines are e.g. ω-amino carbonic acids, as well as L- or D-amino acid residues and combinations of these chemical groups. Efforts are made to synthesize the compounds discussed, in order to confirm the results of the calculations.

The nature of the linker peptide and the stereochemistry of the residue at the P1' position of the hirulog seem to be correlated. The residue at the P1' site determines at which position the linker peptide enters the cleft and thus interacts with the protein. This leads to the conclusion that, although a variety of residues or chemical groups seem to be allowed in the cleft, the optimal substitute of a linker peptide depends on the residue occupying the P1' site of the inhibitor.

Another area of future interest which has to be focussed on are the C-terminal residues of the hirulogs. Peptidomimetics could be incorporated here, in order to reduce biological degradation of the inhibitors. Research on bifunctional hirulog-type inhibitors covers a wide field of interest and provides the basis for rational drug design of specific and potent thrombin inhibitors.

ACKNOWLEDGEMENTS

The help of K. Hoyer during earlier stages of the work and the careful proofreading of the manuscript by Dr. Anke Müller-Fahrnow are gratefully acknowledged.

REFERENCES

1 Tulinsky, A. and Qiu, X., Blood Coagul. Fibrinolysis, 4 (1993) 305.
2 Bode, W., Turk, D. and Karshikov, A., Protein Sci., 1 (1992) 426.
3 Karshikov, A., Bode, W., Tulinsky, A. and Stone, S.R., Protein Sci., 1 (1992) 727.
4 Rydel, T.J., Tulinsky, A., Bode, W. and Huber, R., J. Mol. Biol., 221 (1991) 583.
5 Grütter, M.G., Priestle, J.P., Rahuel, J., Grossenbacher, H., Bode, W., Hofsteenge, J. and Stone, S.R., EMBO J., 9 (1990) 2361.
6 Badgy, D., Barabas, E., Graf, L., Petersen, T.E. and Magnusson, S., Methods Enzymol., 25 (1976) 669.
7 Maraganore, J.M., Bourdon, P., Jablonski, J., Ramachandran, K.L. and Fenton II, J.W., Biochemistry, 29 (1990) 7095.
8 DiMaio, J., Ni, F., Gibbs, B. and Konishi, Y., FEBS Lett., 282 (1991) 47.
9 Skrzypczak-Jankun, E., Carperos, V.E., Ravichandran, K.G., Tulinsky, A., Westbrook, M. and Maraganore, J.M., J. Mol. Biol., 221 (1991) 1379.

10 Qiu, X., Padmanabhan, K.P., Carperos, V.E., Tulinsky, A., Kline, T., Maraganore, J.M. and Fenton II, J.W., Biochemistry, 31 (1992) 11689.
11 Mattson, C.H., Erikson, E. and Nilsson, S., Folia Haematol., 109 (1992) 543.
12 Okamoto, S., Hijikata, A., Kikumoto, R., Tonomura, S., Hara, H., Ninimiya, K., Maruyama, A., Sugano, M. and Tamao, Y., Biochem. Biophys. Res. Commun., 101 (1981) 440.
13 Stuber, W., Kosina, H. and Heimburger, N., Int. J. Pept. Protein Res., 31 (1988) 63.
14 Brandstetter, H., Turk, D., Hoeffken, W., Grosse, D., Stürzebecher, J., Martin, P.D., Edwards, B.F.P. and Bode, W., J. Mol. Biol., 226 (1992) 1085.
15 Banner, D.W. and Hadváry, P., J. Biol. Chem., 266 (1991) 20085.
16 Bode, W., Turk, D. and Stürzebecher, J., Eur. J. Biochem., 193 (1990) 175.
17 Stürzebecher, J., Walsmann, P., Voigt, B. and Wagner, G., Thromb. Res., 36 (1984) 457.
18 Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S.R. and Hofsteenge, J., EMBO J., 8 (1989) 3467.
19 Hagler, A.T. and Lifson, S., J. Am. Chem. Soc., 96 (1974) 5319.
20 Nicholas, J.B., Vance, R., Martin, E., Burke, B.J. and Hopfinger, A.J., J. Phys. Chem., 95 (1991) 9803.
21 Schechter, T. and Berger, A., Biochem. Biophys. Res. Commun., 27 (1967) 157.
22 DiMaio, J., Gibbs, B., Munn, D., Lefebvre, J., Ni, F. and Konishi, Y., J. Biol. Chem., 265 (1990) 21698.
23 LeBonniec, B.F. and Esmon, C.T., Proc. Natl. Acad. Sci. USA, 88 (1991) 7371.
24 Ramachandran, G.N. and Sasisekharan, V., Adv. Protein Chem., 23 (1968) 283.
25 Niefind, K. and Schomburg, D., J. Mol. Biol., 219 (1991) 481.
26 Ponder, J.W. and Richards, F.M., J. Mol. Biol., 193 (1987) 775.