

Functionality map analysis of the active site cleft of human thrombin

Peter D.J. Grootenhuis^a and Martin Karplus^{b,*}

^aDepartment of Computational Medicinal Chemistry, N.V. Organon, P.O. Box 20, 5340 BH Oss, The Netherlands

^bDepartment of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138, U.S.A.

Received 24 August 1995

Accepted 4 December 1995

Keywords: Thrombin inhibitors; Molecular modelling; Rational drug design; MCSS method

Summary

The Multiple Copy Simultaneous Search methodology has been used to construct functionality maps for an extended region of human thrombin, including the active site. This method allows the determination of energetically favorable positions and orientations for functional groups defined by the user on the three-dimensional surface of a protein. The positions of 10 functional group sites are compared with those of corresponding groups of four thrombin–inhibitor complexes. Many, but not all features, of known thrombin inhibitors are reproduced by the method. The results indicate that certain aspects of the binding modes of these inhibitors are not optimal. In addition, suggestions are made for improving binding by interaction with functional group sites on the thrombin surface that are not used by the thrombin inhibitors.

Introduction

Thrombin is a key serine proteinase in the blood coagulation cascade [1–5]. It is involved in the generation of the two major constituents of thrombi: fibrin and activated platelets. Inhibition of thrombin is a promising approach towards prevention and treatment of thrombosis. Although considerable effort has been invested over the years by various research groups to design and synthesize selective thrombin inhibitors, only recently have some successes been reported [6a–e].

There has been a dramatic increase in knowledge of the structural features of blood coagulation proteins in the last few years [7]. The first X-ray crystallographic studies of thrombin–inhibitor complexes were published recently [8–15]. Also, the structure of the thrombin–hirudin complex revealed a completely new (nonsubstrate-like) mode of binding between proteinase and inhibitor [16,17]. Conformational [18] and theoretical [19,20] analyses of the interaction between proteinases and their substrates

and inhibitors have been made, but our understanding of the important contributions to the binding is still incomplete. The availability of structural data on thrombin and its complexes opens the way for the application of modern molecular modelling techniques to the analysis of the known interactions between the enzyme and its inhibitors and for the suggestion of additional sites that might be used to improve inhibitors.

The well-defined nature of the active site cleft of thrombin (see Fig. 1) makes it an excellent candidate for application of the Multiple Copy Simultaneous Search (MCSS) method [21], which was developed for constructing functionality maps of binding sites. As illustrated in Fig. 1, this study is concerned with an extended region that includes the active site. The MCSS method allows the determination of energetically favorable positions and orientations for the ligating functional groups on the surface of proteins with known three-dimensional structure. In the original paper describing the MCSS method, it was applied to the sialic binding site of hemagglutinin.

*To whom correspondence should be addressed.

Abbreviations: MCSS, multiple copy simultaneous search; PPACK, D-phenylalanyl-L-propyl-L-arginine chloromethane; NAPAP, N^{α} -(2-naphthylsulfonylglycyl)-D-para-amidinophenylalanyl-piperidine; argatroban, (2R,4R)-4-methyl-1-[N^{α} -(3-methyl-1,2,3,4-tetrahydro-8-quinolinylsulfonyl)-L-arginyl]-2-piperidine carboxylic acid; rms, root mean square. The thrombin residues are numbered according to the chymotrypsin-based numbering by Bode et al. [8]. P1, P2, P3, etc., denote the peptide inhibitor residues on the amino-terminal side of the scissile peptide bond, and S1, S2, S3, etc., the corresponding subsites of thrombin.

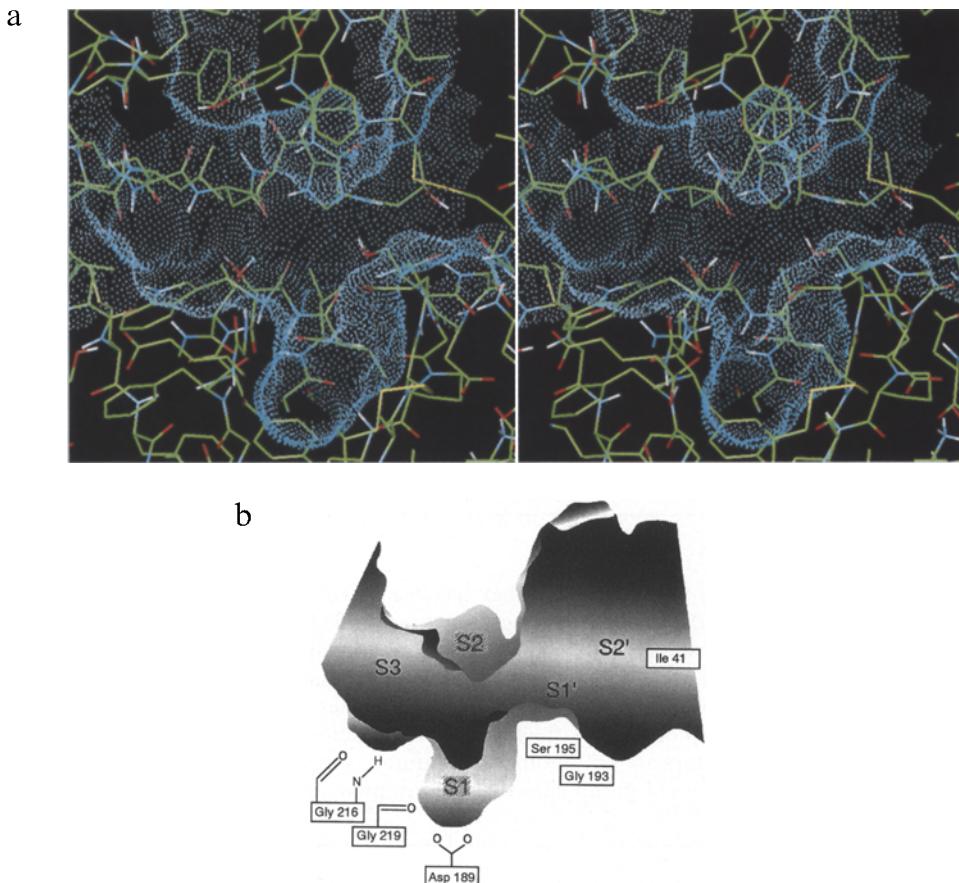


Fig. 1. (a) Stereoview of the active site cleft of thrombin; (b) schematic view of the active site cleft, showing the active site cleft with some of the binding pockets and several of the residues mentioned in the text.

Since then, the method has been used in a detailed analysis of inhibitor binding to HIV-1 protease [22] and of the binding pocket of the FK506-binding protein, FKBP, as a starting point for the dynamic ligand design (DLD) method [23].

There is a growing interest in systematic computation methods that can assist in the design of inhibitors. In the present study, we evaluate the MCSS functionality maps

for thrombin by comparing them with the interactions observed in known inhibitors. We consider the small-molecule inhibitors PPACK, NAPAP, TAPAP and argatroban and the interactions found in the hirudin–thrombin complex. A schematic representation of the small-molecule inhibitors and their interactions with the active site cleft of thrombin is given in Scheme 1.

The available structures of human and bovine throm-

TABLE I
FUNCTIONAL GROUP MODELS AND THE CORRESPONDING NUMBER OF MINIMA AND CLUSTERS OF MINIMA FOUND IN EACH MCSS RUN

Functional group model	Number of group minima	Hydration enthalpy (kcal/mol) ^a	Normalized number of minima ^b
Methyl ammonium	46	-75.0	29
Methyl guanidinium	34	-30.0	31
Acetate	21	-93.0	12
<i>N</i> -methylacetamide	125	-10.0	-125
Methanol	115	-10.6	90
Phenol	203	-13.6	103
Acetonitrile	39	-15.5	30
Methane	46	-3.3	40
Propane	136	-5.4	134
Benzene	173	-7.6	35

^a Values were taken from Refs. 21 and 35.

^b Normalization was obtained by deleting the minima with interaction energies less than one half of the hydration enthalpy of the particular functional group.

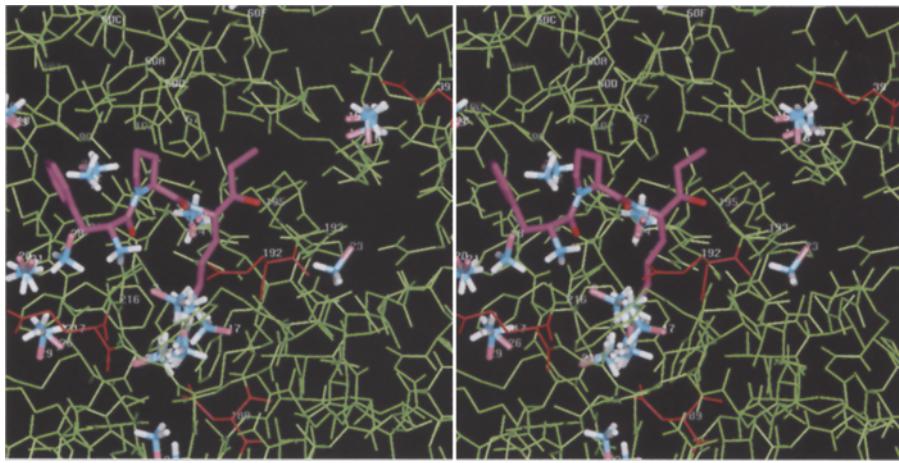


Fig. 2. Stereoview (wall-eyed) of the active site cleft of thrombin (carbons are depicted in green) with PPACK (purple carbons) and high-scoring methylammonium minima (pink). The C^α atoms of thrombin are labeled with their number; labels on the minima denote their ranking. A similar color and labelling scheme is used in Figs. 3–7. Acidic residues discussed in the text are colored red.

bin complexed with inhibitors [9] are quite similar. The overall root-mean-square (rms) deviations between the C^α atoms are on the order of 0.5 Å [9]; the active site residues directly involved in the binding of the various inhibitors show shifts of less than 1 Å [8]. For this reason, we applied the MCSS method to the human thrombin structure of the PPACK complex [8] and used the same results in the analysis of the other inhibitors.

Computational methodology

The MCSS program (v. 1.0) [21] was used to analyze the extended active site region of human thrombin, starting with the 1.92 Å resolution X-ray structure of the human thrombin–PPACK complex. Polar hydrogens were added in positions calculated by using the HBUILD option of CHARMM [24]. A sphere with a radius of 12 Å was defined around the coordinates of the C^α atom of the arginine residue of PPACK. This PPACK residue was

then removed from the structure and a total of 5000 replicas of each functional group were generated and minimized according to the MCSS methodology [21]. The polar hydrogen model was used for the MCSS functional groups [24]. The resulting MCSS minima were sorted according to a criterion based on the interaction energy [21], i.e., minima with interaction energies less than one half of the hydration enthalpy of the particular functional group were removed. This cutoff provides an approximate procedure for selecting sites that could contribute to the net free energy of binding. The selected minima were clustered by visual inspection into groups with similar positions and binding characteristics; i.e., groups that bind in the same way to the same amino acid(s) (e.g. involving the same hydrogen bonds) were put in a cluster. It is the behavior of the clusters rather than individual minima that was analyzed. Values for the parameters required for the MCSS minimizations were taken from Ref. 21.

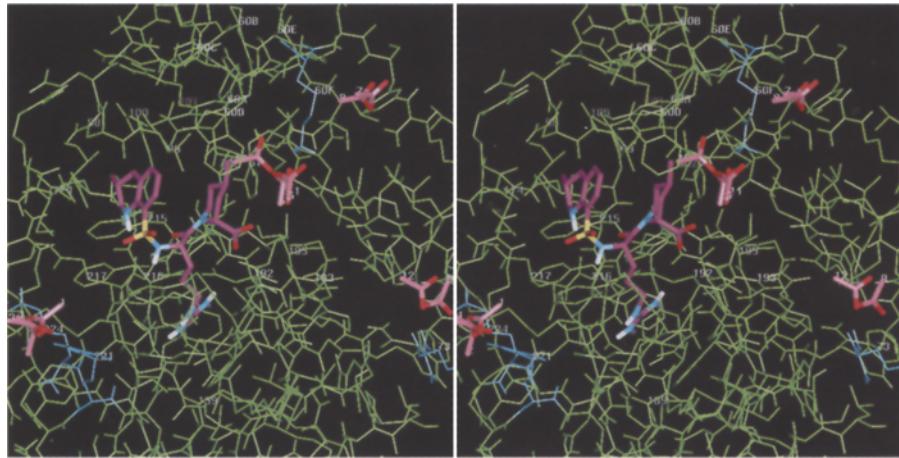


Fig. 3. Stereoview of the active site cleft of thrombin with argatroban and high-scoring acetate minima. Basic residues discussed in the text are colored blue.

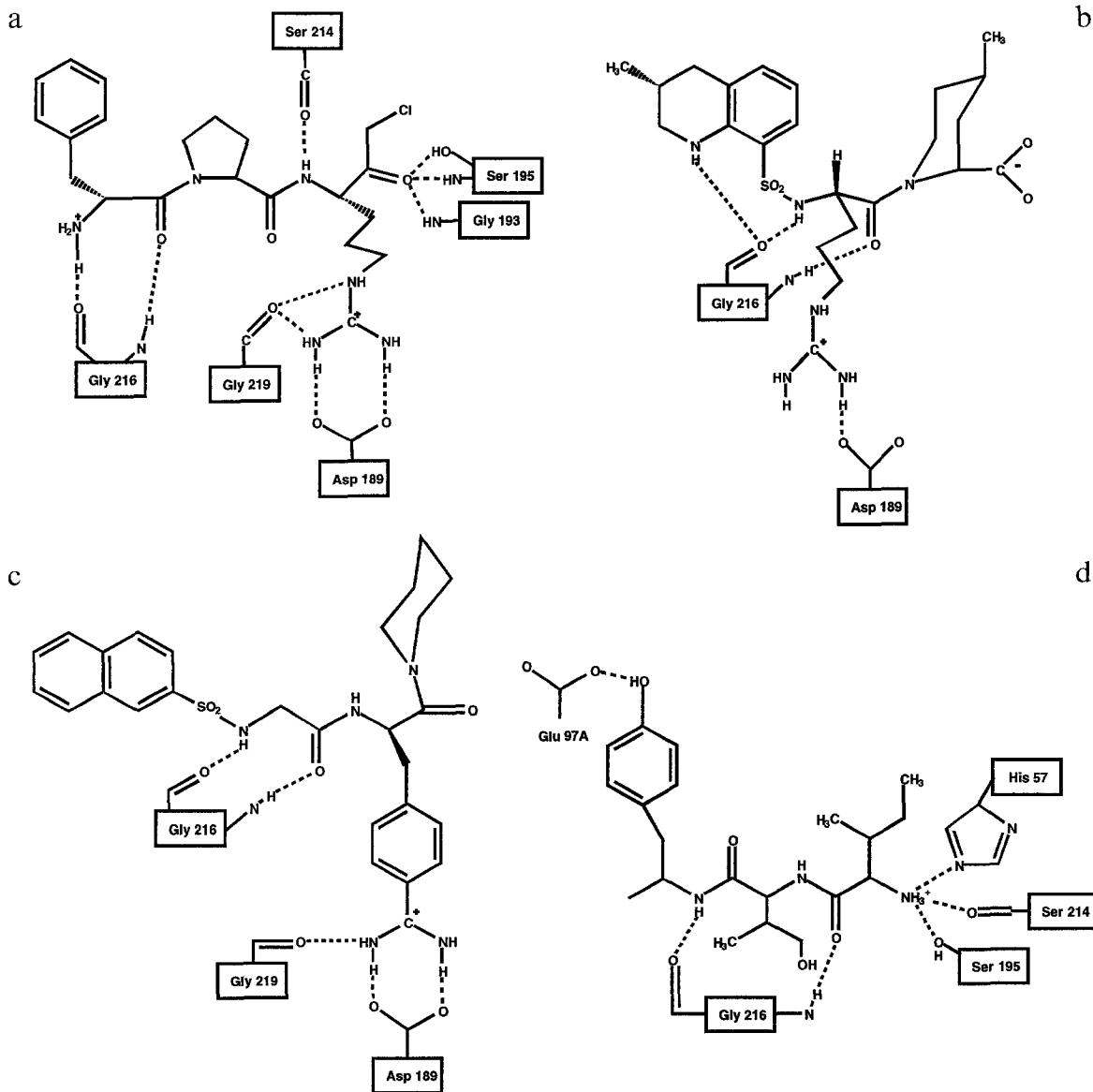
To facilitate the analysis of the MCSS results, three-dimensional models were constructed of the complexes of the PPACK–thrombin structure with NAPAP, argatroban and hirudin. This was done by transferring the inhibitor structures from their corresponding complexes with thrombin to the PPACK–thrombin structure. The resulting complexes were relaxed by a constrained energy minimization in which the thrombin structure was kept fixed.

CHARMM (v. 22.0b) was used for all energy minimizations. Residue topology files for NAPAP and argatroban were generated using the Molecular Editor module of QUANTA [25].

Results and Discussion

The MCSS method was applied to the active site re-

gion of human thrombin using the coordinates of the complex with PPACK. The functional groups employed in the present study are listed in Table 1. Charged groups (methyl ammonium, methyl guanidinium, acetate), uncharged polar groups (*N*-methylacetamide, methanol, phenol, acetonitrile), and uncharged apolar groups (methane, propane, benzene) were used. Table 1 also lists the number of minima (total number and those that satisfy the energy criterion) for each functional group. In most cases (with the exception of benzene), the selected subset is equal to one half or more of the total number. Only a small number of benzene minima are kept; this is due to the use of a polar hydrogen model, which treats each CH group as a neutral sphere centered on the carbon and does not include a representation of the benzene quadrupole moment. Below we describe the minima that provide



Scheme 1. Schematic representations of the complexes between thrombin and (a) PPACK; (b) argatroban; (c) NAPAP; and (d) the three N-terminal residues of hirudin, i.e. Ile¹-Thr²-Tyr³.

significant insights and, when possible, compare them with the corresponding functionalities in the four inhibitors.

Methyl ammonium

When the methyl ammonium group was used as a probe, the minimum displaying the most favorable interaction energy (#1: -114.9 kcal/mol; the interaction energies between the minima and thrombin are given in parentheses) was located near the bottom of the S1 pocket (see Fig. 2). The three small-molecule inhibitors examined in this study all have a similar positively charged (guanidinium or amidinium) group in close proximity to this minimum. In analogy to trypsin [26,27], the specificity of thrombin for protolysis of peptide bonds at the C-terminal end of arginine residues has been attributed to the electrostatic (and steric) features of the S1 pocket. In particular, Asp¹⁸⁹, located at the bottom of the S1 pocket, serves as the counterion for stabilization of the positively charged P1 residue. Minima #9, #10 and #11 (-96.7 to -94.7 kcal/mol) were also located at the bottom of the S1 pocket. Although they have slightly different orientations, they all interact with Asp¹⁸⁹.

At the side of the S1 pocket, minima #3, #4, #5 and #17 (-107.5 to -68.1 kcal/mol) are involved in a twofold hydrogen bonding interaction with the carboxylate group of Glu¹⁹² and the backbone carbonyl of Gly²¹⁹. The carbonyl group of the latter residue also participates in H-bonding interactions with the arginine and benzamidine moieties of PPACK and NAPAP, respectively (see Scheme 1). Although none of the inhibitors is involved in direct interactions with Glu¹⁹², a water-mediated interaction between the N-terminus of PPACK and the carboxylate of Glu¹⁹² has been observed. This indicates that Glu¹⁹² is a residue that could be used for designed thrombin inhibitors. A direct interaction through an appropriate linker group, rather than one mediated by water, would be expected to yield better binding. One reason for this is that an entropic gain is expected because no confined water molecule is required.

A number of lower ranking minima, at a somewhat larger distance from the S1 pocket, are close to some positively charged amino acid side chains of hirudin in the complex structure. Minima #14, #15 and #16 (-78.1 to -70.9 kcal/mol) interact with Glu³⁹ of thrombin. Minima #18 and #19 (-61.8 to -61.3 kcal/mol) adopt a position close to Lys²⁴ of hirudin. This lysine residue forms hydrogen bonds to the backbone carbonyls of residues 97 and 97A of thrombin. Minima #20 and #21 are indicative of the interaction of the backbone amide of Val²¹ of hirudin with Glu²¹⁷ of thrombin.

Methyl guanidinium

As expected, the minima found for the methyl guanidinium group are generally similar to the methyl ammonium

minima. Again, the minimum with the most favorable interaction energy (#1: -97.1 kcal/mol) is located near the bottom of the S1 pocket. It interacts with the carboxylate of Asp¹⁸⁹ and the backbone carbonyl of Gly²¹⁹, in agreement with the structural data for the thrombin-inhibitor complexes [7–15]. Two clusters of minima interact with the carboxylate of Glu¹⁹², again suggesting the potential importance of this amino acid residue for inhibitor binding.

Acetate ion

Of the three small thrombin inhibitors used in this study, only one (argatroban) has a carboxylate functionality. This carboxylate is not involved in direct bonding interactions. Banner and Hadváry have mentioned a possible water-mediated interaction with Lys^{60F} [9]. Although a cluster of minima (#5, #6 and #11, -43.7 to -38.1 kcal/mol) for the acetate group was located near Lys^{60F}, none of these minima was in a position close to the argatroban carboxylate group (see Fig. 3). This result is in line with the above-mentioned observation that the carboxylate group of argatroban is not essential to binding. The derivative that lacks a carboxylate group is still a potent inhibitor, with a K_i of 30 nM, compared with a K_i of 19 nM for argatroban [28]. The carboxylate group was introduced into the molecule because it led to a reduced toxicity [29,30]. A linker to the carboxylate position found by MCSS would be possible.

The cluster containing the four minima with the most favorable interaction energies (-87 to -84 kcal/mol) is positioned between the positively charged side chains of Arg²²¹ and Lys²²⁴. In hirudin, residue Asp⁵ forms salt bridges with the same residues. It has recently been found that Arg²²¹ is involved in the Na⁺ binding site of thrombin. Its side chain appears to adopt another conformation when interacting with the P3 aspartate residue of the substrate protein C [31]. Minima #17 and #18 are positioned in close proximity to Asp⁵ of hirudin. Minima #9 and #12 interact with Arg⁷³ in a manner similar to that of Asp⁵⁵ of hirudin.

N-methylacetamide

This group is a model for the peptide moieties that are present in all the inhibitors under study. A total of 125 minima are found with the MCSS method. Apparently, the active site cleft of thrombin has many positions where it is possible to bind peptides. To simplify the analysis we only consider the 20 best minima; they range in energy from -49.4 to -34.1 kcal/mol. Since the peptide backbone of PPACK (and also those of NAPAP and argatroban) has two hydrogen bonds that form an antiparallel β strand with Gly²¹⁶, we expected a number of minima in this region. To our surprise, only two minima (#19 and #22, -61.3 and -57.5 kcal/mol) adopt positions that are relatively close to the peptide moieties present in the

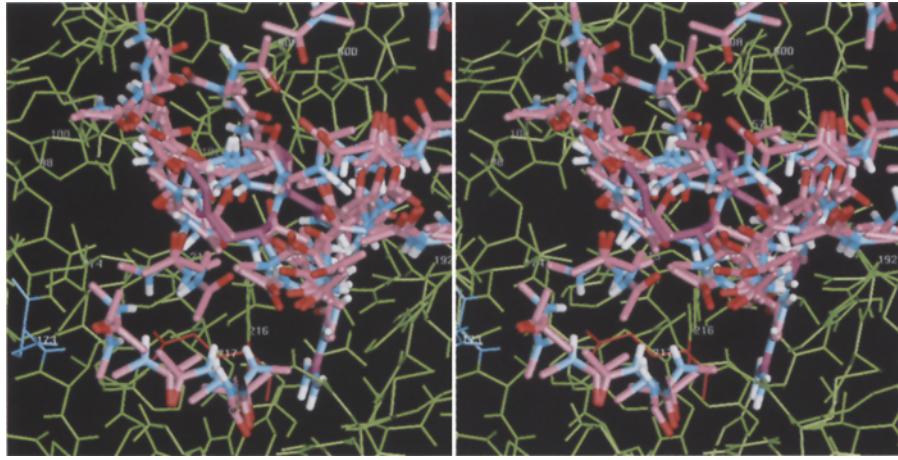


Fig. 4. Stereoview of the active site cleft of thrombin with PPACK and the high-scoring *N*-methylammonia minima.

backbone of PPACK (see Fig. 4). This observation prompted us to analyze the interaction between thrombin and PPACK in more detail. Interaction energies between thrombin and the various parts of the PPACK inhibitor were calculated with the CHARMM potential. The results of the calculations are given in Table 2. The interactions between the backbone atoms of PPACK and thrombin are seen to be unfavorable, due to a repulsive electrostatic contribution involving the carbonyl groups of PPACK and the protein. This is in accord with the small number of minima near the PPACK backbone atoms. It suggests that a gain in binding energy could be obtained by substituting the backbone atoms by other atoms or altering the position of the backbone. This is confirmed by the results for NAPAP and argatroban. Both these inhibitors have other backbone atoms and the characteristic shape of the PPACK backbone is not observed (Fig. 5). This has already been pointed out by Banner and Hadváry [9], who distinguished a substrate (PPACK) from an inhibitor (NAPAP, argatroban) mode

of binding. The two binding modes differ in the way the interacting moieties are presented toward their binding partners in the thrombin active site. The backbones of NAPAP and argatroban are at a larger distance from the surface of the protein and are in closer proximity to many of the MCSS-generated NMA minima. The backbone of the three N-terminal residues of hirudin follows the path indicated by the minima even more closely.

The two highest ranked minima, showing interacting energies of -49.4 and -46.9 kcal/mol, are positioned between Arg¹⁷³ and Glu²¹⁷; the carbonyl of the functional group forms a hydrogen bond to the guanidinium of Arg¹⁷³, while the amide hydrogen atom points towards the carboxylate of Glu²¹⁷.

Methanol

About 90 methanol minima were found by the MCSS procedure. The minima displaying the highest interaction energies bind to charged groups. Thus, #1 (-31.1 kcal/mol) and #2 (-30.9 kcal/mol) interact simultaneously with

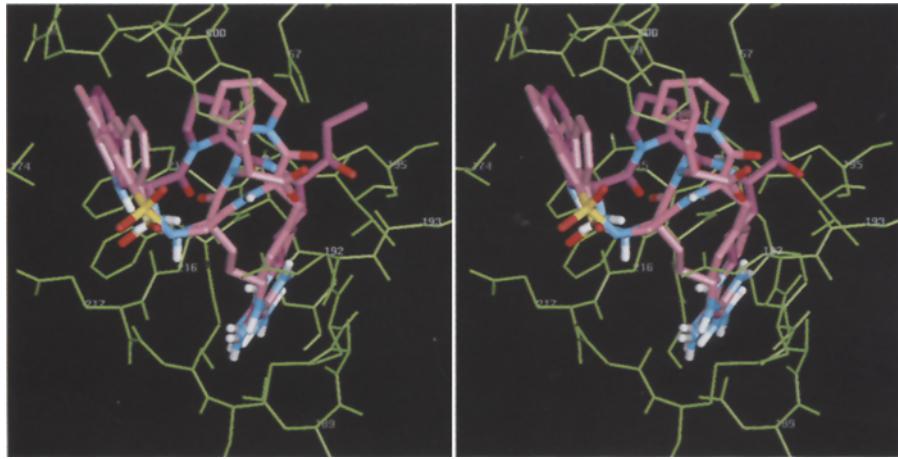


Fig. 5. Stereoview of the active site cleft of thrombin with PPACK (purple), NAPAP (magenta), and argatroban (magenta), illustrating differences between the inhibitor binding modes.

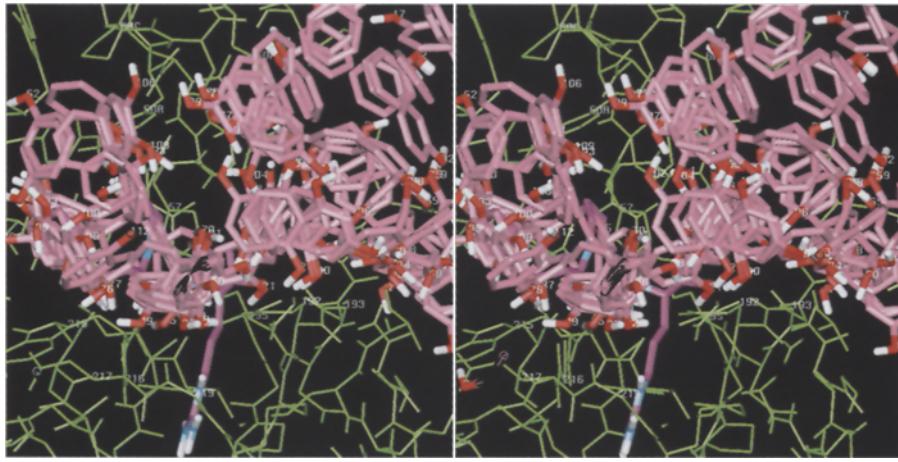


Fig. 6. Stereoview of the active site cleft of thrombin with PPACK and the high-scoring propane minima.

Glu²¹⁷ and Lys²²⁴, rather distal from the active site cleft. Minima #17 (-22.2 kcal/mol) and #18 (-22.1 kcal/mol) are positioned deep in the S1 pocket and interact with Asp¹⁸⁹. Other important interaction sites (e.g. with Gly²¹⁶) were also located by MCSS (#34: -17.4 kcal/mol ; #62: -11.5 kcal/mol). None of the inhibitors under study has a hydroxyl group that interacts with thrombin. This is also true for Thr² of hirudin, which is involved in a van der Waals interaction with the hydrophobic S2 pocket, rather than in a polar interaction where its hydroxyl group plays a role. We note that in the back of the S2 pocket, the backbone carbonyl group of Ala²¹⁵ is in a favorable position for hydrogen bonding, as indicated by minimum #43 (-14.7 kcal/mol).

Phenol

None of the high-scoring phenol copies ended up in a position close to the hirudin Tyr³ residue. However, a number of low-scoring copies were observed in the hydrophobic S3 pocket (ranked #40, #52 and #53, -16.8 to -15.1 kcal/mol) near the expected position (Fig. 6). Simi-

lar to the methanol probe, the most favorable interacting phenols, #1 (-33.4 kcal/mol) and #2 (-31.9 kcal/mol), form hydrogen bonds with Glu²¹⁷ and Lys²²⁴. A large cluster, consisting of eight phenols (including #3–#7, -31.8 to -29.5 kcal/mol), accepts a hydrogen bond from Lys^{60F}. Some phenols have positions close to the PPACK molecule. Phenol #21 (-22.5 kcal/mol) is positioned in the middle of the active site, accepting a hydrogen bond from the amide hydrogen atom of Gly¹⁹³ in the oxyanion hole and at the same time donating a hydrogen bond to the hydroxyl group of Ser¹⁹⁵. The aromatic moiety of #19 (-23.8 kcal/mol) is oriented in the same direction as the N-terminus and the hydroxyl group forms a hydrogen bond to Glu¹⁹². This suggests that addition of a para-substituted phenol moiety to the C^a atom of D-Phe may be a feasible way to obtain a direct interaction with Glu¹⁹².

Acetonitrile

Since the methyl group of acetonitrile is represented by a single extended atom, acetonitrile is mainly involved in polar interactions. The minima are scattered over the

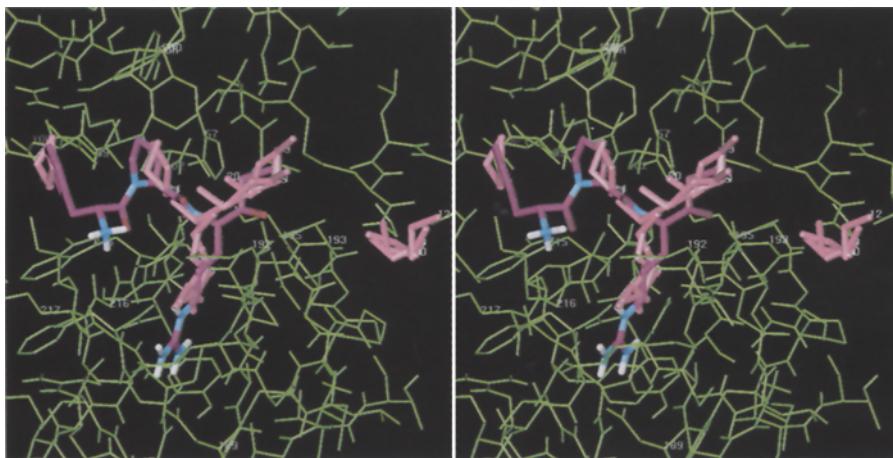


Fig. 7. Stereoview of the active site cleft of thrombin with PPACK and the high-scoring phenol minima.

TABLE 2
INTERACTION ENERGIES (kcal/mol) BETWEEN THROMBIN AND THE VARIOUS PARTS OF PPACK IN THE CRYSTAL STRUCTURE, WITH POLAR ATOMS ADDED

Part of PPACK	No. of atoms	Interaction energy		
		van der Waals	Electrostatic	Total
All atoms	40	-56.0	-309.3	-365.3
H-D-Phe	14	-14.7	-94.1	-108.8
Pro	7	-9.9	8.1	-1.8
Arg-CH ₂ Cl	19	-31.4	-223.3	-254.7
Backbone atoms	12	-15.9	56.0	40.1
Side-chain atoms	28	-40.1	-365.2	-405.4

whole active site cleft. Clusters of minima interacting with the same residues are not observed. The most favorable interaction (#1: -26.8 kcal/mol) occurs between the nitrogen of acetonitrile and the ammonium group of Lys^{60F}. Three acetonitrile copies are positioned in the specificity pocket, i.e., #13, #14 and #22 (-15.0 to -10.6 kcal/mol). However, no direct interaction with Asp¹⁸⁹ is present.

Methane

The methane group serves as a probe for the hydrophobic areas of the active site cleft. It can only be involved in van der Waals contacts, so that the interaction energies are small compared to those calculated for the charged and polar groups. A total of 43 minima was found. The highest scoring ones were located far from the active site groove but, interestingly, #3 (-3.8 kcal/mol) and #5 (-3.6 kcal/mol) were found in the S1 pocket. Minima #3 and #5 were positioned at distances of 0.87 Å and 0.16 Å, respectively, from the C^γ and C^δ of the arginine side chain of PPACK; this corresponds to a nearly optimal fit. Thus, although the bottom of the S1 pocket requires a positively charged group to interact with the negatively charged carboxylate of Asp¹⁸⁹, the upper parts of the S1 pocket are quite lipophilic. This observation is in agreement with the fact that thrombin shows a selectivity for cleaving at the C-terminal site of arginine rather than at the less hydrophilic lysine residue. The other methane minima were scattered over the active site cleft and outside it (see propane).

Propane

As for methane, the calculated minima for propane are found in the established hydrophobic areas of the thrombin active site groove, i.e., the upper part of the S1 pocket and the S2 and S3 pockets. Figure 7 shows the superposition of the 25 highest scoring propane groups (-7.9 to -5.6 kcal/mol) and the inhibitors PPACK, NAPAP and argatroban. The correspondence between the hydrophobic areas of these inhibitors and the minima calculated by MCSS is very good. The highest scoring minima #1 (-7.9 kcal/mol) to #6 (-7.1 kcal/mol) are all found in the upper part of the S1 pocket, again emphasizing its lipophilic character. Interestingly, a cluster consisting of propane minima (#14, #19, #25) is found in the putative S' region of the active site, in a pocket formed by Lys^{60F}, Cys⁴², Cys⁵⁸ and Leu⁴¹. This pocket was suggested by Stubbs and Bode to be the S2' pocket [7b]. Our calculations confirm the lipophilic character of this pocket.

ing its lipophilic character. Interestingly, a cluster consisting of propane minima (#14, #19, #25) is found in the putative S' region of the active site, in a pocket formed by Lys^{60F}, Cys⁴², Cys⁵⁸ and Leu⁴¹. This pocket was suggested by Stubbs and Bode to be the S2' pocket [7b]. Our calculations confirm the lipophilic character of this pocket.

Benzene

The benzene minima, like the methane minima, are distributed throughout the active site cleft. Minimum #6 (-8.0 kcal/mol) is positioned in close proximity to the phenyl ring of the D-Phe residue of PPACK. In the S3 pocket, a number of minima are found. This is in agreement with the aromatic moieties of the inhibitors that are usually located in this pocket (see Scheme 1). Aromatic moieties have been added to the P3 position of PPACK-based ketomethylene transition-state analogues that increase the binding towards thrombin [32].

None of the benzene minima was positioned in the specificity pocket. The same was true for the phenol probe considered above. This is in contrast with experiment, e.g., NAPAP has a benzamidine moiety that is positioned deep in the specificity pocket. It is possible that the specificity pocket in the latter complex is somewhat wider than in the thrombin-PPACK complex. Another possibility is that the united atom van der Waals parameters used to represent the aromatic moieties result in too large a volume.

Conclusions

In the present study we have used the MCSS method to systematically characterize the active site cleft of thrombin. A number of the minima found by MCSS are clearly relevant, since the inhibitor molecules PPACK, NAPAP and argatroban, as well as the three N-terminal residues of hirudin, have corresponding functional groups positioned in close proximity to the minima. Indeed, most of the crucial features of the interactions with thrombin, including a positively charged group near Asp¹⁸⁹ and aromatic moieties in the S3 pocket, are reproduced by the MCSS method. However, the highest scoring minima did

not always turn out to be those of most interest. In practical ligand design applications, it will be necessary to examine a range of minima and not only the 'best' ones. Some potentially useful clues have been obtained by application of MCSS; examples are Glu²¹⁹, Lys^{60F} and the backbone carbonyl of Ala²¹⁵, all of which appear to be important interaction points in the active site cleft of thrombin that could be used in modified inhibitors.

A particularly interesting outcome of the MCSS run with the *N*-methylacetamide probe is that it led to the realisation that the peptide backbone of PPACK may not be optimal. Interaction energy calculations demonstrated that the PPACK backbone electrostatic interaction with thrombin is unfavorable and that, therefore, chemical modification of the backbone could lead to improved binding [33].

An important feature of MCSS is its versatility: it can be used to handle any functional group. In that respect this study has not been exhaustive, and additional functional groups can be proposed that would highlight other features of thrombin inhibition. It would be attractive to design thrombin inhibitors that do not have positively charged groups positioned in the thrombin specificity pocket [34]. Such uncharged compounds are more easily absorbed by the intestines and may have reduced side effects (arginine moieties often lead to effects on the heart rate and blood pressure). It is clear from the present study that the upper part of the specificity pocket is quite hydrophobic, since propane and methane probes were observed to bind. This is in agreement with experimental observations that benzamidine (present in NAPAP) binds in this pocket. The fact that MCSS was unable to position aromatic moieties in this pocket needs further study. An all-atom representation of benzene may help to solve this problem.

In the experimental structures, ordered water molecules mediate between interaction points on the enzyme and the inhibitor. One way to mimic such interactions would be to include water molecule copies in the MCSS approach. This is being done by making up functional groups that include one or more water molecules as part of the groups of interest (D. Joseph and M. Karplus, work in progress).

In conclusion, this study shows that MCSS offers a practical way – biased only by the choice of the functional groups – to characterize extended active site clefts. The visualization of clusters of minima provided by MCSS for the various functional groups is highly suggestive. Ideas for chemical modifications of existing inhibitors to arrive at more potent or specific inhibitors originate more-or-less automatically from the MCSS functionality maps. Despite the fact that many features of thrombin inhibitors were reproduced well by MCSS, the real test for its validity and usefulness has to await the synthesis and testing of 'improved' compounds.

References

- 1 Wallis, R.B., Drugs Today, 25 (1989) 597.
- 2 Hauptmann, J. and Markwardt, F., Semin. Thromb. Hemost., 18 (1992) 200.
- 3 Davie, E.W., Fujikawa, K. and Kisiel, W., Biochemistry, 30 (1991) 10363.
- 4 Berliner, L.J., Thrombin, Structure and Function, Plenum Press, New York, NY, 1992.
- 5 Fareed, J., Hoppensteadt, D., Walenga, J.M. and Pifaré, R., In Pifaré, R. (Ed.) Anticoagulation, Hemostasis, and Blood Preservation in Cardiovascular Surgery, Hanley & Belfus, Inc., Philadelphia, PA, 1993, pp. 111–128.
- 6 a Jakubowski, J.A., Smith, G.F. and Sall, D.J., Annu. Rep. Med. Chem., 27 (1992) 99.
b Tapparelli, C., Metternich, R., Ehrhardt, C. and Cook, N.S., Trends Pharmacol. Sci., 14 (1993) 366.
c Maffrand, J.P., Nouv. Rev. Fr. Hematol., 34 (1992) 405.
d Harker, L.A., Blood Coagul. Fibrinolysis, 5 (1994) S47.
e Lyle, T.A., Persp. Drug Discov. Design, 1 (1993) 453.
- 7 a Tulinsky, A., Thromb. Haemostasis, 66 (1991) 16.
b Stubbs, M.T. and Bode, W., Curr. Opin. Struct. Biol., 4 (1994) 823.
c Stubbs, M.T. and Bode, W., Persp. Drug Discov. Design, 1 (1993) 431.
- 8 Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S.R. and Hofsteenge, J., EMBO J., 8 (1989) 3467.
- 9 Banner, D.W. and Hadváry, P., J. Biol. Chem., 266 (1991) 20085.
- 10 Brandstetter, H., Turk, D., Hoeffken, H.W., Grosse, D., Stürzebecher, J., Martin, P.D., Edwards, B.F.P. and Bode, W., J. Mol. Biol., 226 (1992) 1085.
- 11 Rydel, T.J., Ravichandran, K.G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C. and Fenton, J.W., Science, 249 (1990) 277.
- 12 Gütter, M.G., Priestle, J.P., Rahuel, J., Grossenbacher, H., Bode, W., Hofsteenge, J. and Stone, S.R., EMBO J., 9 (1990) 2361.
- 13 Skrzypczak-Jankun, E., Carperos, V.E., Ravichandran, K.G., Tulinsky, A., Westbrook, M. and Maraganore, J.M., J. Mol. Biol., 221 (1991) 1379.
- 14 Martin, P.D., Robertson, W., Turk, D., Huber, R., Bode, W. and Edwards, B.F.P., J. Biol. Chem., 267 (1992) 7911.
- 15 a. Stubbs, M.T., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S.R. and Bode, W., Eur. J. Biochem., 206 (1992) 187.
b. Zdanov, A., Wu, S., DiMaio, J., Konishi, Y., Li, Y., Wu, Z., Edwards, B.F.P., Martin, P.D. and Cygler, M., Proteins, 17 (1993) 252.
c. Maryanoff, B.E., Qiu, X., Padmanabhan, K.P., Tulinsky, A., Almond, H.R., Andrade-Gordon, P., Greco, M.N., Kauffman, J.A., Nicolaou, K.C., Liu, A., Brungs, P.H. and Fusetani, N., Proc. Natl. Acad. Sci. USA, 90 (1993) 8048.
d. Mathews, I.I., Padmanabhan, K.P., Ganesh, V., Tulinsky, A., Ishii, M., Chen, J., Turck, C.W., Coughlin, S.R. and Fenton, J.W., Biochemistry, 33 (1994) 3266.
- e. Weber, P.C., Lee, S.-L., Lewandowski, F.A., Schadt, M.C., Chang, C.H. and Kettner, C.A., Biochemistry, 34 (1995) 3750.
- f. Tabernero, L., Chang, C.Y., Ohringer, S.L., Lau, W.F., Iwanowicz, E.J., Han, W.C., Wang, T.C., Seiler, S.M., Roberts, D.G.M. and Sack, J.S., J. Mol. Biol., 246 (1995) 14.
- 16 Bode, W. and Huber, R., Eur. J. Biochem., 204 (1992) 433.
- 17 Bode, W. and Huber, R., Curr. Opin. Struct. Biol., 1 (1991) 45.
- 18 Hubbard, S.J., Campbell, S.F. and Thornton, J.M., J. Mol. Biol., 220 (1991) 507.

- 19 Daggett, V., Schröder, S. and Kollman, P.A., *J. Am. Chem. Soc.*, 113 (1991) 8926.
- 20 Bemis, G.W., Carlson-Golab, G. and Katzenellenbogen, J.A., *J. Am. Chem. Soc.*, 114 (1992) 570.
- 21 Miranker, A. and Karplus, M., *Proteins*, 11 (1991) 29.
- 22 Cafisch, A., Miranker, A. and Karplus, M., *J. Med. Chem.*, 36 (1993) 2142.
- 23 Miranker, A. and Karplus, M., *Proteins*, 23 (1995) 472.
- 24 Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M., *J. Comput. Chem.*, 4 (1983) 187.
- 25 QUANTA/CHARMM, v. 3.2, Polygen Corp., Waltham, MA, 1991.
- 26 Janin, J. and Chothia, C., *J. Mol. Biol.*, 100 (1976) 197.
- 27 Huber, R. and Bode, W., *Acc. Chem. Res.*, 11 (1978) 114.
- 28 Kikumoto, R., Tamao, Y., Tezuka, T., Tonomura, S., Hara, H., Ninomiya, K., Hijikata, A. and Okamoto, S., *Biochemistry*, 23 (1984) 85.
- 29 Hijikata, A. and Okamoto, S., *Semin. Thromb. Hemost.*, 18 (1992) 135.
- 30 Grootenhuis, P.D.J. and Van Helden, S.P., In Wipff, G. (Ed.) *Computational Approaches in Supramolecular Chemistry*, Kluwer, Dordrecht, 1994, pp. 137–149.
- 31 Di Cera, E., Guinto, E.R., Vindigni, A., Dang, Q.C., Ayala, Y.M., Wuyi, M. and Tulinsky, A., *J. Biol. Chem.*, 270 (1995) 22089.
- 32 Cheng, L., Goodwin, C.A., Schully, M.F., Kakkar, V.V. and Claeson, G., *J. Med. Chem.*, 35 (1992) 3364.
- 33 Jetten, M., Peters, C.A.M., Visser, A., Grootenhuis, P.D.J., Van Nispen, J.W. and Ottenheijm, H.J.C., *Bioorg. Med. Chem.*, 3 (1995) 1099.
- 34 For a recent example, see Deadman, J.J., Elgendi, S., Goodwin, C.A., Green, D., Baban, J.A., Patel, G., Skordalakes, E., Chino, N., Claeson, G., Kakkar, V.V. and Scully, M.F., *J. Med. Chem.*, 38 (1995) 1511.
- 35 Cabani, S., Gianni, P., Mollica, V. and Lepori, L., *J. Solut. Chem.*, 10 (1981) 563.