Volume 287, number 1,2, 133–138 FEBS 09998 August 1991
© 1991 Federation of European Biochemical Societies 00145793/91/$3.50
ADONIS 001457939100711X

Geometry of binding of the Nα-tosylated piperidides of *m*-amidino-, *p*-amidino- and *p*-guanidino phenylalanine to thrombin and trypsin

X-ray crystal structures of their trypsin complexes and modeling of their thrombin complexes

Dušan Turk[1], Jörg Stürzebecher[2] and Wolfram Bode[1]

[1]*Max-Planck-Institut für Biochemie, D(W)-8033 Martinsried, Germany* and [2]*Medizinische Akademie Erfurt, D(O)-5010 Erfurt, Germany*

Received 4 May 1991

The X-ray crystal structures of the complexes formed with bovine trypsin and the Nα-tosylated piperidides of *m*-amidino-, *p*-amidino- and *p*-guanidino-D,L-phenylalanine (3-TAPAP, 4-TAPAP and 4-TGPAP) were determined with data to 1.8 Å resolution. The L-stereoisomer of 3-TAPAP binds as a compact entity into the active site of trypsin, with the amidino and the carbonyl groups of the central amidinophenylalanyl residue hydrogen-bonded to Gly^{216} of trypsin. According to modeling and energy minimization, 3-TAPAP fits perfectly in this conformation to the more restrictive thrombin active site also (Bajusz et al. (1978) Int. J. Pept. Prot. Res. 12, 217–221); the piperidine moiety extends into the cage-like S2 subsite of thrombin, but leaves room for additional substituents which might help to improve binding and pharmacological properties. In contrast, 4-TAPAP and 4-TGPAP bind only weakly and in an extended conformation to trypsin; their considerably enhanced affinities for thrombin would suggest a more compact binding to thrombin.

Thrombin; Antithrombotic; Caogulation; X-Ray crystal structure; Inhibitor complex

1. INTRODUCTION

Thrombin, a trypsin-like serine proteinase, is the final enzyme in both coagulation cascades and plays a central role in thrombosis and hemostasis. Clot formation in various thrombotic diseases has been directly attributed to the action of thrombin. Its key role in inducing diseases like coronary thrombosis, stroke and pulmonary embolism points to use appropriate inhibitors as therapeutic agents. A large number of thrombin-specific synthetic inhibitors have been found and tested, most of them derived from arginine, benzamidine and D-PheProArg [2–6]. The experimental structure of human α-thrombin in complex with D-PheProArgCh-2Cl [1] allows better understanding of the inhibitory properties and a more rational approach to the design of new inhibitors.

We have recently published the X-ray structure of the complex formed between NAPAP (for abbreviations see Fig. 1), the tightest binding benzamidine-based inhibitor of thrombin, with bovine trypsin [7]. This experimental structure allowed reliable modeling of the NAPAP–thrombin complex providing a plausible explanation for the extraordinary selectivity and high affinity of NAPAP for thrombin. In the following we describe the structure determinations of similar trypsin complexes with *m*-amidino-, *p*-amidino- and *p*-guanidino-D,L-phenylalanine derivatives (3-TAPAP, 4-TAPAP, 4-TGPAP: for abbreviations see Fig. 1), and present the results of modeling with thrombin. The structural results are correlated with the experimental binding data.

Correspondence address: D. Turk, Max-Planck-Institut für Biochemie, D(W)-8033 Martinsried, Germany.

2. MATERIALS AND METHODS

The crystallization of bovine trypsin and the crystal parameters of a new, more 'open' orthorhombic trypsin crystal form have been described elsewhere [7]. Large trypsin crystals were soaked for 2 to 3 days at 20°C in a calcium-saturated 2.5 M ammonium sulphate solution, buffered with 0.1 M phosphate to pH 8, containing about 1 mg/cm^3 of 3-TAPAP, 4-TAPAP, or 4-TGPAP (solubilized using small amounts of dimethylsulphoxide). 3- and 4-TAPAP (Fig. 1) were synthesized as described [8], 4-TGPAP (S-2576; Fig. 1) was purchased from Kabi Vitrum, Sweden. The compounds were D,L-racemates.

X-ray data were collected with the FAST television area detector (Enraf-Nonius). Evaluation (Table I), modeling and refinement with X-PLOR [9] were done as described for NAPAP-trypsin [7]. The molecular models of 3-TAPAP, 4-TAPAP and 4-TGPAP (Fig. 1) were constructed and built into electron density using MAIN [10]. The structures were crystallographically refined with energy restraints using X-PLOR version 1.5 using X-PLOR force field parameters (PARAM19X.PRO) and sulphur-atom bond lengths and angles as described previously [7]. The final mode and refinement parameters are given in Table II.

To model the thrombin complexes, the trypsin-bound inhibitor models were transformed to the thrombin active site [1] and energy-

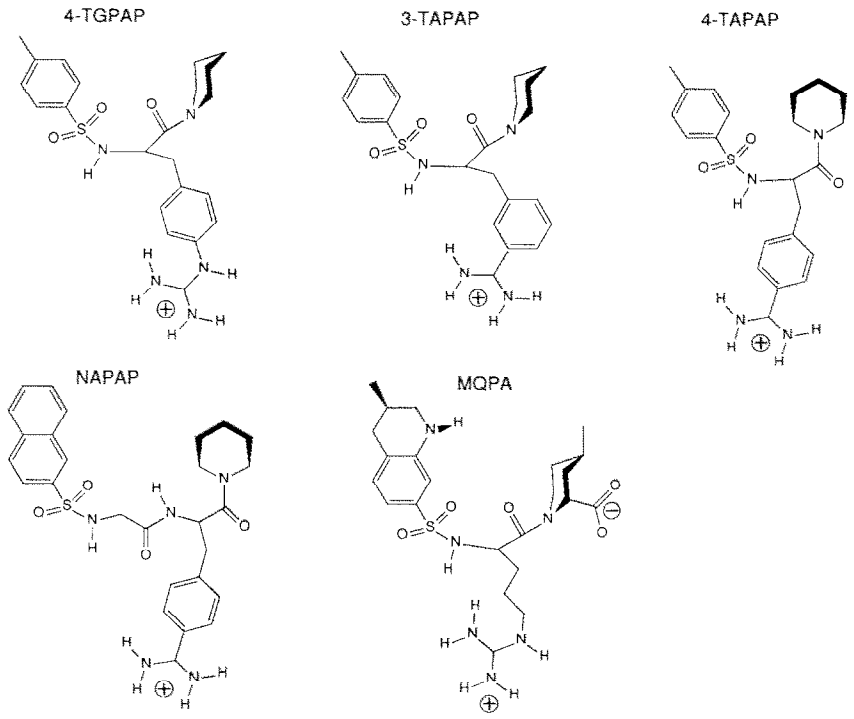

Fig. 1. Chemical formulae of 3-TAPAP (*Nα*-[4-toluene sulphonyl]-D,L-*m*-amidino-phenylalnyl-piperidine), 4-TGPAP (*Nα*-(4-toluene sulphonyl]-*p*-guanidino-phenylalanyl-piperidine), 4-TAPAP (*Nα*-[4-toluene sulphonyl]-D,L-*p*-amidino-phenylalanyl-piperidine), NAPAP (*Nα*-[2-Naphthyl-sulphonyl-glycyl]-D,L-*p*-amidino-phenylalanyl-piperidine) and MQPA ([2*R*, 4*R*]-4-methyl-1-[*Nα*(3-methyl-1,2,3,4- tetrahydro-8-quinolene-sulphonyl)-L-arginyl]-2-piperidine carboxylic acid).

minimized (excluding electrostatic energies) with MAIN using the force field parameters of X-PLOR as described previously [7].

3. RESULTS

3.1. *The cyrstal structure of the trypsin-3-TAPAP complex*

The topologies of the three inhibitors investigated together with their chemical names are given in Fig. 1. The 3-TAPAP molecule bound to trypsin is well defined by electron density (Fg. 2). It is the L-stereoisomer which nestles into the trypsin active site in a rather compact conformation. Its peptide chain runs antiparallel to trypsin segment $Ser^{214}-Ser^{217}$, forming two hydrogen bonds from the amino and carbonyl group of the central *m*-amidinophenylalanine residue to Gly^{216} of trypsin.

Table I

Reflection Parameters
($R_{merge} = \Sigma(I - <I>) / \Sigma I$)

Parameter	3-TAPAP	4-TAPAP	4-TGPAP
Maximal resolution	1.95 Å	1.75 Å	1.80 Å
Number of measurements	77873	52111	61407
R_{merge}	0.075	0.066	0.092
Number of independent reflections	18614	17191	19464
Measured/possible reflections to maximal resolution	0.88	0.57	0.73

The amidinophenyl moiety is arranged in a similar manner as benzamidine [11] and amidinophenyl pyruvic acid [12] i.e. sandwiched between trypsin segments $Ser^{190}-Gln^{192}$ and $Trp^{215}-Gly^{216}$, which together with the interconnecting disulfide bridge 191–220 form the 'base' and 'ceiling' of the trypsin specificity pocket. The amidino group faces the carboxylate group of Asp^{189}, and is in contact with $Gly^{219}O$ and a characteristic 'buried' solvent molecule at the back of the pocket, thus fully exploiting its hydrogen bonding capacity with optimal charge compensation.

The tosyl group, with slightly enhanced mobility, stands almost perpendicular to the indole moiety of Trp^{215}. Both oxygen atoms of the sulphonyl group are well resolved in the electron density and are directed towards $Gly^{219}N$ and away from the trypsin surface. The piperidine ring, which is clearly confined to one of the two possible chair conformations, is placed between the imidazole ring of His^{57} and the toluene ring. The planar amido group between the *m*-amidinophenylalanine and the piperidine ring is quite distant from Ser^{195} ($O\gamma$-C:5.8 Å); the 'oxyanion hole' of trypsin is occupied by an ordered solvent molecule (Sol^{281}, Fig. 2). As with NAPAP-trypsin [7], the side chains of Ser^{195} and of His^{57} are in an almost perfect hydrogen bonding arrangement ($O\gamma$-$N\epsilon = 2.9$Å).

Table II

Final model and refinement parameters. The R-value is defined as $(\Sigma(|F_{obs}| - |F_{cal}|) / \Sigma|F_{obs}|)$

Parameter	3-TAPAP	4-TAPAP	4-TGPAP
Number of non-hydrogen protein atoms	1629	1629	1629
Number of non-hydrogen inhibitor atoms	30	30	31
Number of solvent molecules	214	147	243
Calcium ion	1	1	1
SO_4^{2-}	1		
Overall B-value	11.7 Å^2	12.5 Å^2	13.8 Å^2
Rms standard deviation from target values			
Bond lengths	0.011 Å	0.012 Å	0.012 Å
Bond angles	2.43°	2.57°	2.67°
Number of unique reflections used for refinement	18199	14781	18091
Resolution range	8.0–1.90 Å	8.0–1.90 Å	8.0–1.90 Å
R-value	0.167	0.178	0.169

3.2. *Crystal structures of trypsin-4-TGPAP and trypsin-4-TAPAP*

The active site of the trypsin-4-TGPAP complex is shown in Fig. 3 together with the electron density calculated with phases obtained with the refined model. In contrast to the trypsin complex formed with 3-TAPAP, the density representing 4-TGPAP is weaker and less structured, so that no unequivocal inhibitor model interpretation is possible. The 4-TGPAP model shown in Fig. 3 represents the best fit to the density obtained from crystallographic refinement.

(i) The density within the specificity pocket (Fig. 3) can account for the guanidino phenylalanyl side chain; Fig. 3 shows just one of two alternative orientations of the guanidyl group.

(ii) The density around its α-carbon seems only to be in agreement with the D-stereoconfiguration.

(iii) Appropriate density exists to localize the sulphonyl moiety (with one of both S-O groups in hydrogen bond distance from $Gly^{216}N$).

(iv) There is, however, no significant density which could account for the toluene group even at low contouring levels. A similar situation of an undefined, i.e. mobile toluene group has been found for N-tosyl-lysyl-chloromethyl-ketone-trypsin [13].

(v) The density extending along Ser^{195} (opposite to the sulphonyl group, Fig. 3) has been assigned to the carbonyl group and the adjacent piperidine ring. However

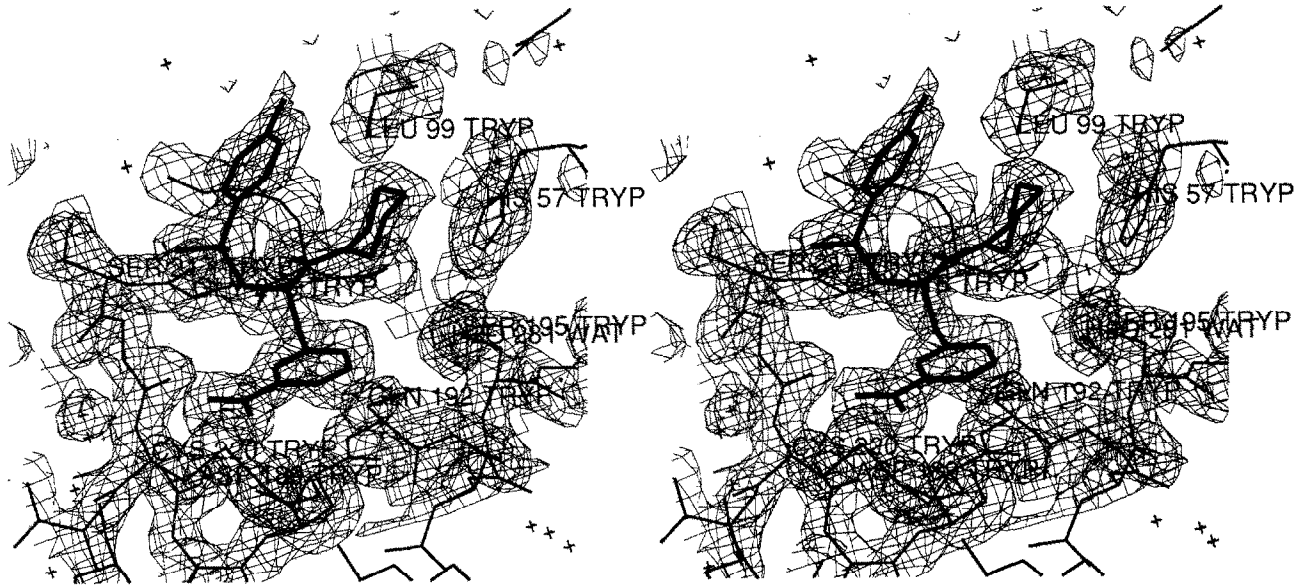

Fig. 2. Active-site region of the complex formed between 3-TAPAP (thick connections) and bovine trypsin (thin connections) superimposed with the final $2F_{obs}$-F_{calc} electron density. The view is towards the trypsin surface; the active-site residues are to the right, the specificity pocket is at the back. Crosses indicate localized solvent molecules. The density is calculated with phases omitting the inhibitor atoms, contoured at 1.0σ.

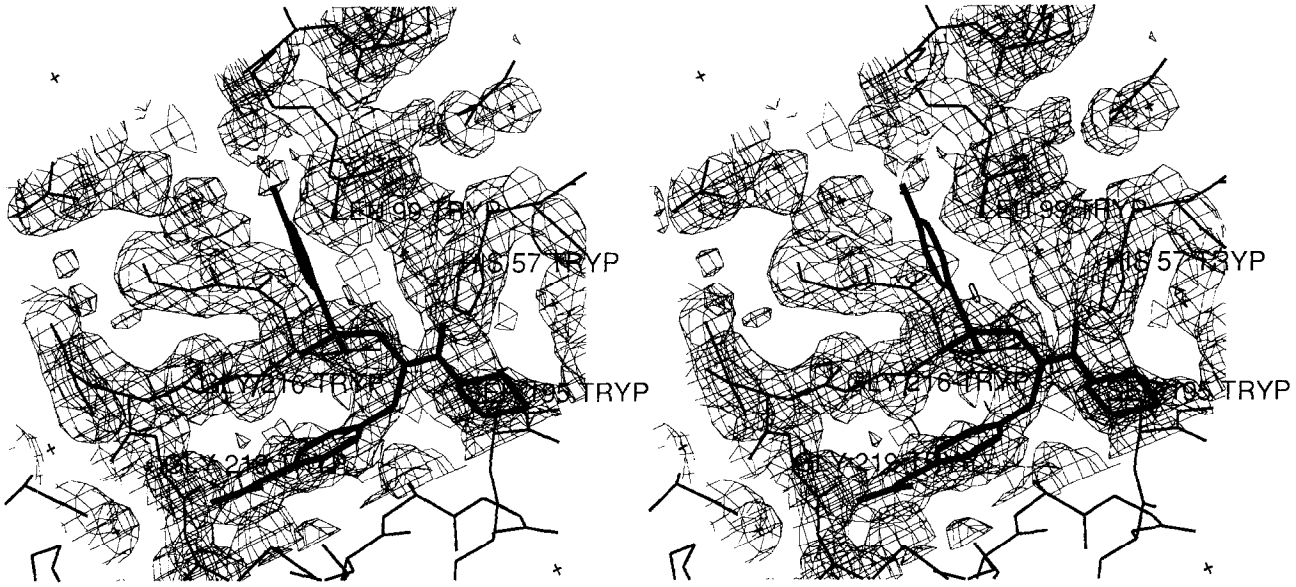

Fig. 3. Active-site region of the complex formed between 4-TGPAP (thick conections) and bovine trypsin (thin connections) as in Fig. 2.

its placement along the His^{57} imidazole ring (i.e. in the vertical density stretch in Fig. 3) is also possible.

The density for crystals containing 1 mM D,L-4-TAPAP was even more broken and difficult to interpret. The trypsin binding site should be completely saturated at this concentration, if the dissociation constant (Table III) is not affected by the crystal (salt and contact) conditions. The weak density presumably indicates that 4-TAPAP does not bind with one preferred arrangement to the active site of trypsin.

3.3. *Models of thrombin complexes*

The active-site region of the human α-thrombin structure [1] overlaid with the transformed model of 3-TAPAP (after energy minimization) is shown in Fig. 4 together with the thrombin Connolly surface. The 3-TAPAP molecule fits almost perfectly in its trypsin binding conformation (Fig. 2) to the more restrictive thrombin active site. The amidinophenyl moiety of 3-TAPAP is similarly placed in the thrombin specificity pocket as that of NAPAP [7]; the arrangement of the 'main chain' atoms of 3-TAPAP (with a phenylalaline L-stereoisomer) and the two hydrogen bonds formed with Gly^{216} are, however, relatively similar to that found for MQPA [7,14]. The toluene group of 3-TAPAP can be arranged as in the trypsin structure. In thrombin, however, it finds a much more notched and hydrophobic cleft (the 'aryl binding site', made up by Ile^{174}, Trp^{215} and Leu^{99}) (Fig. 4). The piperidine group of 3-TAPAP extends into the thrombin S2-cavity, but does not fill it up tightly.

For the corresponding modelling of the thrombin complexes with 4-TGPAP and 4-TAPAP several alternative conformations are conceivable. The 'external' positioning of their piperidine rings (as shown in Fig. 3 for the trypsin complex) seems in thrombin less probable. Their considerably stronger binding to thrombin compared with trypsin suggests that both *p*-substituted inhibitors interact more tightly with the furrowed active-side surface of thrombin.

Table III

Inhibition constants (K_i[μmol/l]) of bovine trypsin, human and bovine thrombin

Inhibitor	Human thrombin	Bovine thrombin	Bovine trypsin
3-TAPAP	0.54	0.34	1.2
4-TAPAP	0.64	1.3	64.0
4-TGPAP	0.45	0.57	100.0
MQPA	0.019	–	5.0
NAPAP	0.0066	0.006	0.69

4. DISCUSSION

The crystal structures of the trypsin complexes and the model structures of the thrombin complexes presented in this and in the preceding paper [7] reveal some more general rules for the arrangement of the amidino- and guanidino phenylalanine derivatives. In these complexes two interactions are observed (for both, trypsin as well as thrombin).

(i) The amidino or the guanidino group is placed adjacent to carboxylate group of Asp^{189} allowing a favourable hydrogen bond/salt bridge interaction.

(ii) The inhibitor is connected through at least one, but optimally two hydrogen bonds to both projecting polar groups of Gly^{216}. In the case of NAPAP [7] the glycine

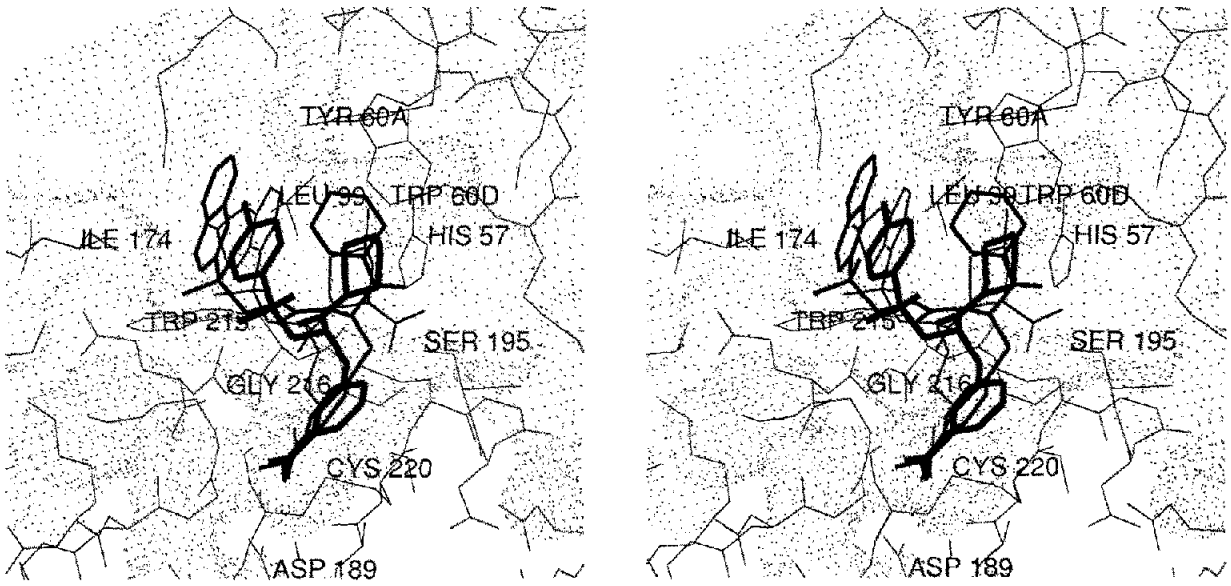

Fig. 4. Active-site region of thrombin including Connolly surface with proposed models of MQPA (thin line inhibitor), NAPAP (thicker line inhibitor) and 3-TAPAP (the thickest line inhibitor).

'spacer' forms these two hydrogen bonds with Gly^{216} (Fig. 4), while in the case of 3-TAPAP and MQPA (and probably also 3-TGAP – for explanation see legend to Fig. 5) [7,14] the amido and the carbonyl group of the central phenylalanine residue mediate this interaction (Fig. 5).

The formation of one (to Gly^{216}N) or two hydrogen bonds and the simultaneous favourable insertion of the phenylalanine moiety into the specificity pocket require an L-configuration of the central phenylalanyl residue in the '3-TAPAP-like' arrangement, but a D-configuration in the 'NAPAP-like' arrangement (Fig. 4). Model building studies suggest that binding of the corresponding enantiomeric forms in either position would be extremely unfavourable.

These interactions at the central inhibitor residue(s) are equally possible in trypsin and in thrombin. Thrombin, however, offers different (additional) docking sites

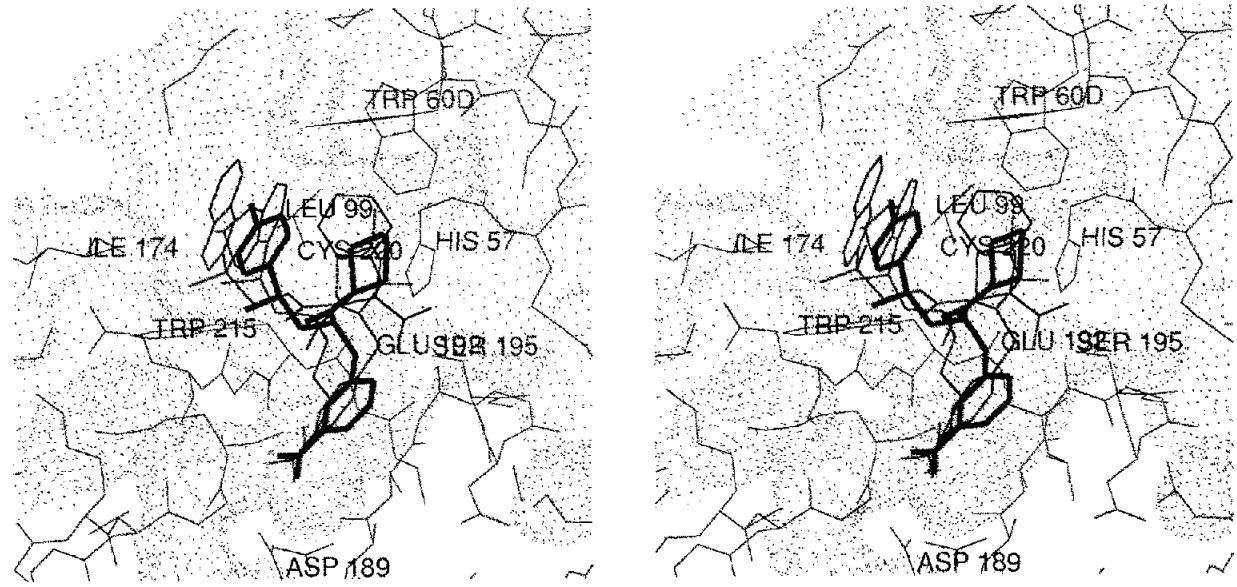

Fig. 5. Active-site region of thrombin including Connolly surface with proposed models of MQPA (thin line inhibitor), 3-TAPAP (thicker line inhibitor) and 3-TGPAP (the thickest line inhibitor). The geometry of 3-TGPAP model is based on the assumption that the binding configuration at $C\alpha$ is L. The $C\alpha$ chirality was postulated as L, since it is possible to achieve a similar distance between the central carbon atom of the guanidino group to the $C\alpha$ as in MQPA and 3-TAPAP. The 3-TGPAP guanidino group is supposed to bind in an orientation similar to MQPA, with only one nitrogen atom pointing towards the Asp^{189} (the 3-TGPAP has, as far as we know, not been synthesized yet).

for interaction with the peripheral inhibitor groups, namely (i) a much more notched hydrophobic 'aryl-binding site' [7], and (ii) a cavity-like S2-subsite.

NAPAP takes advantage of both thrombin sites through interaction with its naphtyl and unsubstituted piperidine moieties (see Fig. 4); these additional favourable interactions account for the enhanced thrombin binding compared with trypsin [7] (Table III).

The toluene group of 3-TAPAP can likewise interact favourably with the aryl-binding site of thrombin; its piperidine ring fills the S2-subsite only loosely; presumably, this improper fit results in an only moderately increased affinity to thrombin compared with trypsin (Table III). There would be some empty space for additional substitution of a small hydrophobic group in the equatorial 4-position (Fig. 4), which could render a 3-TAPAP derivative a tighter binding thrombin inhibitor; a likewise favourable substitution has been found for MQPA [2] (see Fig. 4). 3-TAPAP seems to be a particularly attractive starting molecule for derivatizations, as the position of its piperidine in thrombin allows a further substitution in the 2-position (with D-configuration) by alkyl chains, which in turn could be equipped with polar groups to allow further favourable interactions with the oxyanion hole and with the S1′-subsite of thrombin.

Binding in the '3-TAPAP-like' position (with L-configuration) is for the *p*-substituted phenylalanine derivatives (4-TAPAP and 4-TGPAP) unlikely due to primarily unfavourable amidino/guanidino group orientations with respect to Asp^{189}, but also to the large distances between the carbon atom of the charged group (amidino, guanidino) and the Cα of the phenylalanyl moieties (6.2 Å in 4-TAPAP and NAPAP, above 7.0 Å in 4-TGPAP, compared with 5.7 Å in 3-TAPAP). 4-TAPAP and 4-TGPAP in a 'NAPAP-like' position (with D-configuration) could only form one hydrogen bond (between the sulphonyl and Gly^{216}N, see Fig. 4); their weak interaction with the trypsin active site probably gives rise to the weak density (Fig. 3) and to the low affinities for trypsin (Table III). The additional tight interactions of the toluene and the piperidine moieties and perhaps another binding mode with thrombin would probably result in a stronger fixation and in higher affinities (Table III). This suggests that the binding of 4-TAPAP and 4-TGPAP towards thrombin is not equivalent to the binding to trypsin (in contrast to the arrangement of 3-TAPAP, NAPAP and MQPA). Certainty can probably only be achieved by determining crystalline thrombin complexes; corresponding experiments are in progress.

Acknowledgements: The constant interest and support of Prof. Dr R. Huber is greatly acknowledged. We thank Dr. M. Stubbs for careful proof reading. This work has been supported by the SFB 207 of the Universität München and by the Fonds der Chemischen Industrie.

REFERENCES

[1] Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S.R. and Hofsteenge, J. (1989) EMBO J. 8, 3467–3475.
[2] Okamoto, S., Hijikata, A., Kikumoto, R., Tonomura, S., Hara, N., Ninomiya, K., Maruyama, A., Sugano, M. and Tamao, Y. (1981) Biochem. Biophys. Res. Commun. 101, 440–446.
[3] Stürzebecher, J., Walsmann, P., Voigt, B. and Wagner, B. (1984) Thromb. Res. 36, 457–465.
[4] Kettner, C. and Shaw, E. (1979) Thromb. Res. 14, 969–973.
[5] Bajusz, S., Barabas, E., Tolnay, P., Szell, E. and Bagdy, D. (1978) Int. J. Peptide Protein Res. 12, 217–221.
[6] Kettner, C., Mersinger, L. and Knabb, R., (1990) J. Biol. Chem. 265, 30, 18289–18297.
[7] Bode, W., Turk, D. and Stürzebecher, J. (1990) Eur. J. Biochem. 193, 175–182.
[8] Wagner, G., Horn, H., Richter, P., Vieweg, H., Lischke, I. and Kazimirowski, H.G. (1981) Pharmazie 36, 597–603.
[9] Brünger, A.T., Kuriyan, K. and Karplus, M. (1987) Science 235, 458–460.
[10] Turk, D. (unpublished results).
[11] Bode, W. and Schwager, P. (1975) J. Mol. Biol. 98, 693–717.
[12] Walter, J. and Bode, W. (1983) Hoppe-Seyler's Z. Physiol. Chem. 364, 949–959.
[13] Li, X., Bode, W. (unpublished results)
[14] Matzusaki, T., Sasaki, C., Okumura, C. and Umeyama, H. (1989) J. Biochem. (Tokyo) 105, 949–952.