Molecular Recognition with Biological Receptors: Structure-Based Design of Thrombin Inhibitors

Patrick Betschmann\textsuperscript{a}, Christian Lerner\textsuperscript{a}, Stefan Sahli\textsuperscript{a}, Ulrike Obst\textsuperscript{b}, and François Diederich\textsuperscript{a*}

Abstract: Molecular recognition is at the center of biological function. Consequently, a profound understanding of the underlying nonbonding interactions is required to intervene in a rational way in biological processes. Such detailed knowledge can be gained in studies with designed artificial receptors or, more directly, with biological receptors such as the enzyme thrombin. X-ray structural information on this key enzyme in the blood coagulation cascade has guided the structure-based design of a class of active and selective non-peptidic, reversibly binding low molecular weight inhibitors. These compounds feature a conformationally rigid bi- or tricyclic core structure from which side chains diverge into the four major binding pockets (distal D, proximal P, recognition or selectivity S1, and oxyanion hole) at the thrombin active site. With their rigid central core, all inhibitors prefer similar modes of association to thrombin, and detailed information on the strength of individual intermolecular bonding interactions and their incremental contribution to the overall free enthalpy of complexation is generated in correlative binding studies. Phenylamidinium is the side chain of choice for the S1-pocket. Attempts to replace this group with less basic functional groups, which cannot undergo bidentate ionic H-bonding to the carboxylate of Asp189 at the bottom of this pocket, were unsuccessful. The P-pocket is occupied by an isopropyl group, in analogy to the natural substrate fibrinogen, which uses the side chain of a valine residue to fill this site. The large hydrophobic D-pocket was found to accommodate one and even two aromatic residues. Attempts to direct side chains bearing H-bond acceptor groups into the oxyanion hole are described. The most active inhibitor prepared in this investigation showed a $K_i$ value for thrombin inhibition of 9 nM and a 800-fold selectivity for binding thrombin over trypsin.

Keywords: Molecular recognition - Non-peptidic inhibitors - Pharmaceutical chemistry - Structure-based design - Thrombin

1. Introduction

Biological processes such as receptor-ligand, enzyme-substrate, enzyme-inhibitor, or antibody-hapten binding, and cell-surface recognition are governed by molecular recognition phenomena [1]. Understanding and controlling the underlying noncovalent bonding interactions such as H-bonding, ion pairing, or dispersive interactions as well as the complex solvation-desolvation phenomena accompanying molecular recognition processes in the liquid phase is therefore essential for rational approaches towards the treatment of diseases.

In our laboratory we have been following two distinct approaches to advance the understanding of biological molecular recognition processes. On the one hand, we prepare and study a wide variety of artificial receptors for the complexation of small biomolecules such as steroids [2], saccharides [3], or amino acid derivatives [4]. On the other, we exploit X-ray structural information available on pharmacologically interesting biological receptors to design and synthesize low molecular weight inhibitors and to determine - in systematic structure-activity relationships - incremental contributions of individual intermolecular bonding interactions to the overall binding free enthalpy. The latter approach is illustrated in this article.

High-resolution X-ray crystal structures of free enzymes and their complexes with inhibitors increasingly disclose size, shape, and molecular recognition properties of their active sites [5][6]. This kind of structural information is essential for the rational design of novel non-peptidic small molecules (de novo design) capable of binding to the sterically and electronically complementary active site and inhibiting the enzymatic reaction. It is also of tremendous help in the rational modification and improvement of hit or lead structures [7].

De novo design of non-peptidic enzyme inhibitors follows an iterative cycle comprising i) X-ray structure determination of the protein bound to an inhibitory substrate, ii) computer-aided analysis of
Thrombin is believed to be an ideal candidate for structure-based inhibitor design, as X-ray crystal structures of enzyme-inhibitor complexes show well-defined binding pockets in the active site of an overall rigid protein. Analyzing various X-ray crystal structures of thrombin-inhibitor complexes helped distinguishing four major binding pockets in the active site: a large hydrophobic pocket capable of binding aromatic or alicyclic rings (distal D-pocket), a smaller hydrophobic pocket binding small hydrophobic side chains (proximal P-pocket), a deep, narrow pocket with a carboxylate residue (side chain of Asp189) at the bottom (selectivity S1-pocket), and the catalytic site with the nucleophilic serine and the so-called oxyanion hole [18].

After studying the important interactions between thrombin and some of its most potent inhibitors, it was decided to build a new lead structure with a rigid bicyclic scaffold featuring a range of side chains reaching into the various pockets [19]. With both binding partners the enzyme active site and the inhibitor, being highly preorganized and conformationally rigid, we anticipated low entropic costs of inhibitor binding [10a].
A phenylamidinium side chain was chosen to bind to the S1-recognition pocket. Many potent thrombin inhibitors place a positive charge in the S1-pocket where the natural substrate fibrinogen binds with an Arg side chain [17][20]. Furthermore, a benzimidazolinium ion by itself is an inhibitor of thrombin [21]. Molecular modeling and biological data on a first generation of synthetic inhibitors indicated that a large piperonyl (= (1,3-benzodioxol-5-yl)methyl) residue would nicely fit into the D-pocket, which in case of the natural substrate is occupied by a Phe side chain [10a]. Finally, an isopropyl was found to be the substituent of choice for the smaller P-pocket; in the complex with fibrinogen, this pocket is occupied by a Val side chain. This way, compound (+)-1 (Fig. 2) emerged as a highly active thrombin inhibitor ($K_i = 7$ nM [10b]) which also displayed remarkable selectivity ($K_i$(trypsin)/$K_i$(thrombin) = 740) over the related serine protease trypsin [22].

The X-ray crystal structure of the complex of thrombin and (+)-1 revealed that the phenylamidinium side chain undergoes ion pairing with the COO- residue of Asp189 in the S1-pocket. The amidinium group acts as donor in four H-bonds, two with the side pairs of the COO- residue of Asp189, one with the carbonyl O-atom of Gly219, and one with a water molecule (not shown, Fig. 2). The phenyl ring in the upper part of the S1-pocket is stabilized by π-π stacking interactions with the polarizable β-sheet type peptidic surrounding. The piperonyl ring in the D-pocket forms a hydrogen bond with the phenolic OH-group of Tyr60A and takes an edge-to-face orientation with respect to the indole moiety of Trp215. This pocket can also be filled by a cyclohexyl residue without much loss in binding free enthalpy, which demonstrates that dispersion interactions and hydrophobic desolvation are more important interaction forces in this pocket than the electrostatic aromatic CH/π interaction [10a,c]. The complex of (+)-1 with thrombin is additionally stabilized by yet another hydrogen bond between its lactam carbonyl group and the NH-group of Gly216 [10b]. Defined mutations in a series of highly preorganized inhibitors suggest a contribution of 0.8 kcal mol$^{-1}$ to the overall binding free enthalpy resulting from this H-bond [10b,c].

Subsequent work, described in the following, was aimed at further enhancements of binding affinity and specificity through modification of the inhibitor side chains extending in the S1- and D-poockets and by anchoring new side chains to the central skeleton for reaching into the oxyanion hole.

### 3. Substitution of the Phenylamidinium Side Chain

Lead structure (+)-1, similar to many other thrombin inhibitors, contains a highly basic amidinium functionality, which is completely ionized under physiological conditions. Compounds of this type are very likely to exhibit poor oral bioavailability [23], defeating even the most favorable pharmacodynamic parameters such as high activity or target selectivity. A series of analogs bearing less basic or nonpolar substituents in place of the phenylamidinium moiety were therefore designed and synthesized [24], aiming at potentially orally bioavailable compounds that were to retain satisfactory inhibitory activity. Several amines, amides, a phenol, and a (hydrated) aldehyde were all predicted to fit well into the S1-pocket, exhibiting a suitable degree of steric and electronic complementarity.

The key step in a typical synthesis [10c] (Scheme 1) consisted of the 1,3-dipolar cycloaddition between an in situ generated azomethine ylide (from 5 and 6) and an N-substituted maleimide 4 (prepared from 2 and 3), yielding the tricyclic scaffold (+)-7. Chemo- and regioselective reduction of the sterically less hindered imide carbonyl group and sequential nucleophilic substitution via N-acyliminium ions led to sulfone (+)-8 and the isopropyl-substituted lactam (+)-9. Regioselective catalytic hydrogenation of the protected aniline (+)-10 gave a mixture of the cis- and trans-cyclohexylamines (+)-11 and (+)-12, which could be separated at the stage of the trifluoroacetyl derivatives (+)-13 and (+)-14.

Applying Pd-catalyzed cross-coupling methodology to the versatile aryl bromide (+)-15, which had been synthesized in a similar way to acetanilide (+)-9, gave access to phenol (+)-16 and amine (+)-10 although in inferior yield. The corresponding nitrile (+)-17 was readily reduced to benzylamine (+)-18 and hy...
drolyzed to benzamide (±)-19. Partial reduction of (±)-17 to the imine, followed by hydrolysis yielded aldehyde (±)-20 (Scheme 2) [25].

Unfortunately, almost all of the targeted molecules ((±)-9, (±)-10, (±)-11, (±)-12, (±)-16, (±)-19, and (±)-20) were found to be inactive against thrombin within the range of a standard biological assay ($K_i$ for thrombin inhibition was in each case higher than 35 μM) [22]. The only analog of this series to display any residual activity was benzylamine (±)-18 with an inhibitory strength ($K_i$ = 17 μM) three orders of magnitude worse than that of the corresponding phenylamidinium derivative (±)-1 ($K_i$ = 13 nM). Apparently, the lack of either the positive charge (e.g. in (±)-19) or steric similarity (e.g. in (±)-12) to the phenylamidinium residue in (±)-1 leads to a dramatic loss in activity. It is assumed that the rigid and itself tightly bound tricyclic scaffold, being optimized for the phenylamidinium substituent, cannot sufficiently adjust its position to allow a different S1-pocket substituent to form the predicted favorable interactions with the enzyme.

4. Optimizing the D-Pocket Substituent

A group at Merck working on thrombin inhibition discovered that the hydrophobic D-pocket in the active site should be spacious enough to easily accommodate larger substituents than piperonyl [26]. Likewise, we expected a gain in activity by introducing a second benzene ring to occupy this sub-site. The N-benzhydryl inhibitor (±)-21 and, for comparison purposes, the N-benzyl analog (±)-22 (Fig. 3) were prepared and tested. Compound (±)-23 combines the features of both (±)-1 and (±)-21. Fig. 3 shows schematic representations of the predicted interactions between thrombin and the active enantiomers of the inhibitors (±)-21, (±)-22, and (±)-23; Fig. 4 shows a computer model of the expected binding mode of the active enantiomer of inhibitor (±)-23. A binding mode similar to that found in the crystal structure of thrombin complexed to (±)-1 was assumed.

During the synthesis of (±)-21, it became apparent that the bulkiness of the N-benzhydryl substituent precluded the successful application of the established synthetic protocol as outlined in Scheme 1 [10b,c] when steric hindrance blocked the introduction of the sulfone group. It was therefore decided to change the order of the synthetic steps.

Fig. 3. Inhibitors designed for the study of the spatial capacity of the D-pocket.

Scheme 2. Synthesis of target molecules (±)-16, (±)-18, (±)-19, and (±)-20 and alternative synthesis of aniline (±)-10.
of side chain introduction and attach the isopropyl before the benzhydryl residue to the central scaffold. The need for a N-protective group resisting acidic, basic, reductive, and organometallic conditions while guaranteeing acceptable yields and selectivity in the cycloaddition and reduction steps was fortunately met by the piperonyl group (Scheme 3). Benzylic oxidation of lactam (±)-15 followed by acidic hydrolysis provided the N-deprotected lactam (±)-24 [27] which was realkylated to furnish the N-benzhydryl and isopropyl substituted inhibitor (±)-25 [28]. Conversion to nitrile (±)-26 followed by Pinner reaction gave the target molecule (±)-21 [10c]. Compound (±)-23 was prepared in the same manner by starting from lactam (±)-24.

The biological activity of the N-benzhydryl compound (±)-21 \( (K_i = 28 \text{ nM}) \) being better than that of the N-benzyl derivative (±)-22 \( (K_i = 65 \text{ nM}) \) confirmed our prediction that the D-pocket would favorably interact with a group larger than benzyl or piperonyl. While (±)-21 was just a little less active than the piperonyl analog (±)-1 \( (K_i = 13 \text{ nM}) \), its selectivity for thrombin over trypsin was found to be slightly elevated \( (K_i(\text{trypsin})/K_i(\text{thrombin}) = 1200 ((±)-21), 760 ((±)-1), 380 ((±)-22) \). These results encouraged us to continue this study by designing and synthesizing (±)-23 (Fig. 3), with two piperonyl residues. Interestingly, only a modest rise in binding strength was observed \( (K_i = 9 \text{ nM}) \) and selectivity was again a little lower \( (K_i(\text{trypsin})/K_i(\text{thrombin}) = 800) \).

 Apparently, a second benzene ring in (±)-21 compensates – at least partly – for the loss of the hydrogen bond between benzodioxol and Tyr60A in (±)-1. Introducing a second benzodioxol moiety into the D-pocket, however, does not seem to boost the biological properties as much as we had hoped. X-ray crystallographic studies are under way to clarify the exact inhibition mode of (±)-21 and (±)-23 and, in particular, to elucidate the nature of the interaction between the bulky benzhydryl groups and the D-pocket.

5. Reaching into the Oxyanion Hole

The negatively charged tetrahedral intermediates formed during thrombin-catalyzed peptide cleavage are stabilized by ionic H-bonding to NH groups of the peptidic backbone inside the so-called oxyanion hole (Fig. 5). Placing a suitable substituent, capable of H-bonding to these groups, into this pocket could sub-
stantially increase binding strength. We expected carbonyl groups to act as a good H-bond acceptors when placed into the oxyanion hole. Also, H-bonding towards Leu41 or ion pairing with Lys60F were predicted for the active enantiomers of compounds (±)-27b–j. Unfortunately, all synthesized derivatives (±)-27b–k, featuring a bicyclic scaffold with a side chain pointing in the direction of the catalytic site, were inferior to the tricyclic parent compound (±)-27a.

The syntheses of derivatives (±)27b–k usually began with the familiar azomethine ylide 1,3-dipolar cycloaddition [10c]. R² and R³ were determined by the choice of amino acid employed for the cycloaddition reaction and subsequent functional group interconversions. R² was introduced by N-substitution of the cycloadduct [10][29]. The conversion of the nitrile to the amide, usually one of the last steps, was accomplished by the Pinner reaction [30]. A typical synthesis for derivatives (±)-27i,j,k is shown in Scheme 4 [10][31]. The azomethine ylide formed from amino acid 28 and aldehyde 29 reacted with maleimide 4 to give cycloadduct (±)-30, which was N-methylated and transesterified to yield (±)-31. The subsequent Pinner reaction provided (±)-27i. Alternatively, (±)-31 was transformed into N-benzyloxylamide-substituted (±)-32, which was converted via (±)-27k into hydroxylamide (±)-27j (unfortunately, when (±)-32 was submitted to Pinner reaction conditions, the N-benzyloxylamide was largely solvolyzed by methanol, yielding 31% of undesired methyl ester (±)-27i but only 21% of hydroxylamide (±)-27j).

Successfully equipping this class of inhibitors with an oxyanion hole binding side chain that does not undergo reaction with the serine nucleophile appears to be rather difficult. Here again, the high degree of preorganization and the rigidity of the tightly bound scaffold may not allow the new substituent to assume a position required to form the predicted H-bonds. However, the high residual activity of the large N-benzyloxylamide substituted derivative 27k clearly indicates that the catalytic site and its neighborhood, the S1'-pocket, are capable of accommodating quite large groups. This again leaves room for further exploration.

6. Perspectives

This study shows that our class of inhibitors, with their rigid bi- and tricyclic scaffolds, offer very little tolerance for structural or functional group variation in the S1-recognition pocket. The phenylamidinium side chain is by far superior to all the other residues designed to bind to this selectivity pocket. The D-pocket however, is easily capable of binding at least two aromatic residues. Obviously, this opens up a plethora of conceivable substituted or fused aromatic and heteroaromatic systems. More research needs to be done on the D-pocket substituent, and it seems to be an exciting direction to follow. Similarly, the catalytic site with the oxyanion hole and the S1'-pocket in particular look very promising as largely unused pockets with the capacity of binding larger or smaller residues. Further studies will be based on substituting inhibitors that have a tricyclic scaffold as shown for the active enantiomer of (±)-28 in Fig. 6. Again, H-bonding in the oxyanion hole and/or ion pairing with Lys60F would be the goal.

Finally, a ‘fluorine scan’ is planned: a systematic replacement of one or more H-atoms in the scaffold and the side
chains of the inhibitor by fluoro substituents, hoping to elucidate the 'fluorophilicity' or 'fluorophobicity' of a structurally well-defined enzyme active site.

This study clearly demonstrates how difficult it still is to successfully predict new receptor-inhibitor bonding interactions, even if both binding partners are structurally well defined. Nevertheless, we believe that in future medicinal chemistry research, the use of structural information will become increasingly helpful to plan efficient and selective molecular recognition of novel low molecular weight inhibitors by biological receptors.

Acknowledgements

This work was supported by F. Hoffmann-La Roche AG. We thank David W. Banner and Allán D’Arcy for protein X-ray crystallographic analyses and Heidi Hoffmann Maiocchi, Pascale Blond, and Thomas Tschopp for the measurement of enzyme inhibitor constants. We also thank Holger Gohlke for his contributions towards the design of some inhibitors described.

Received: September 6, 2000

---

[1] W.-D. Woggdon, R.M. Wenger, Chimia 2000, 54, 9.
[2] a) P. Wallimann, T. Marti, A. Führer, F. Diederich, Chém. Rev. 1997, 97, 1567; b) P. Wallimann, P. Seiler, F. Diederich, Helv. Chim. Acta 1996, 79, 779; c) P. Wallimann, S. Mattei, P. Seiler, F. Diederich, Helv. Chim. Acta 1997, 80, 2368; d) S. Mattei, P. Wallimann, B. Kenda, W. Amrein, F. Diederich, Helv. Chim. Acta 1997, 80, 2391; e) T. Marti, B.R. Peterson, A. Führer, T. Mordasini-Denti, J. Zarske, B. Jaun, F. Diederich, Helv. Chim. Acta 1998, 81, 105; f) A. Führer, T. Marti, F. Diederich, H. Künzer, M. Brehm, Helv. Chim. Acta 1999, 82, 1843; g) A.E. Chris-