ELSEVIER

Chemistry & Biology 8 (2001) 1107–1121

Chemistry & Biology

www.elsevier.com/locate/chembiol

Research Paper

Engineering inhibitors highly selective for the S1 sites of Ser190 trypsin-like serine protease drug targets

Bradley A. Katz*, Paul A. Sprengeler, Christine Luong, Erik Verner, Kyle Elrod, Matt Kirtley, James Janc, Jeffrey R. Spencer, J. Guy Breitenbucher, Hon Hui, Danny McGee, Darin Allen, Arnold Martelli, Richard L. Mackman [1]

Axys Pharmaceutical Corporation, 385 Oyster Point Boulevard, South San Francisco, CA 94080, USA

Received 3 April 2001; revisions requested 26 June 2001; revisions received 23 July 2001; accepted 13 August 2001; First published online 9 October 2001

Abstract

Background: Involved or implicated in a wide spectrum of diseases, trypsin-like serine proteases comprise well studied drug targets and anti-targets that can be subdivided into two major classes. In one class there is a serine at position 190 at the S1 site, as in urokinase type plasminogen activator (urokinase or uPA) and factor VIIa, and in the other there is an alanine at 190, as in tissue type plasminogen activator (tPA) and factor Xa. A hydrogen bond unique to Ser190 protease–arylamidine complexes between $O\gamma_{Ser190}$ and the inhibitor amidine confers an intrinsic preference for such inhibitors toward Ser190 proteases over Ala190 counterparts.

Results: Based on the structural differences between the S1 sites of Ser190 and Ala190 protease–arylamidine complexes, we amplified the selectivity of amidine inhibitors toward uPA and against tPA, by factors as high as 220-fold, by incorporating a halo group ortho to the amidine of a lead inhibitor scaffold. Comparison of K_i values of such halo-substituted and parent inhibitors toward a panel of Ser190 and Ala190 proteases demonstrates pronounced selectivity of the halo analogs for Ser190 proteases over Ala190 counterparts. Crystal structures of Ser190 proteases, uPA and trypsin, and of an Ala190 counterpart, thrombin, bound by a set of ortho (halo, amidino) aryl inhibitors and of non-halo parents reveal the structural basis of the exquisite selectivity and validate the design principle.

Conclusions: Remarkable selectivity enhancements of exceptionally small inhibitors are achieved toward the uPA target over the highly similar tPA anti-target through a single atom substitution on an otherwise relatively non-selective scaffold. Overall selectivities for uPA over tPA as high as 980-fold at physiological pH were realized. The increase in selectivity results from the displacement of a single bound water molecule common to the S1 site of both the uPA target and the tPA anti-target because of the ensuing deficit in hydrogen bonding of the arylamidine inhibitor when bound in the Ala190 protease anti-target.

Keywords: Selectivity at S1 site; H_2O displacement; Urokinase type plasminogen activator; Tissue type plasminogen activator; Ser190/Ala190 protease; Structure-based drug design

1. Introduction

Trypsin-like serine proteases comprise a set of enzyme targets involved or implicated in disease states. A significant finding germane to the development of trypsin-like serine protease inhibitors as drugs is that considerable specificity can be achieved at the S1 site alone [1–3]. Such specificity facilitates development of small molecule inhibitors whose low molecular weight may promote bioavailability. Despite the high structural similarity in the S1 sites of trypsin-like serine proteases, these sites are differentiated from one another by two important features that can be targeted for selectivity development – their depth, and the identity of residue 190. Simple amidine inhibitors tend to be selective for Ser190 proteases (urokinase type plasminogen activator (uPA), trypsin and tryptase) rather than Ala190 proteases (tissue type plasminogen activator (tPA), thrombin, and factor Xa), because of an additional

* Corresponding author.

[1] Also corresponding author.

E-mail addresses: brad_katz@axyspharm.com (B.A. Katz), richard_mackman@axyspharm.com (R.L. Mackman).

1074-5521/01/$ – see front matter
PII: S1074-5521(01)00084-9

hydrogen bond between the amidine and $O\gamma_{Ser190}$ in the former class [1].

Urokinase is a Ser190 trypsin-like serine protease implicated in cancer progression [4–8], and may be a tractable drug target for antimetastic therapy [1,2,9–14]. Therapeutic inhibitors of uPA should not significantly inhibit or disrupt the function of the closely related blood-clotting enzyme, tPA, an anti-target [14]. Both uPA and tPA share similar three dimensional structures [9,15], and nearly identical S1 sites except for the side-chain of residue 190 [1]. The preference of small amidine inhibitors for uPA over tPA is significant but not huge, ranging from 7.7-fold for benzamidine to 80-fold for 4-iodobenzo[*b*]thiophene-2-carboxamidine [1]. Thus it would be advantageous if the S1 recognition element of the inhibitor could be engineered to exploit more fully the Ser190 → Ala difference between uPA and tPA in order to produce small molecule scaffolds with greater intrinsic specificity toward uPA and against anti-targets such as tPA.

One reason why the selectivity of small amidine inhibitors toward Ser190 proteases over Ala190 counterparts is not larger is that a water tightly bound at S1, H_2O1_{S1}, bridges the inhibitor with the Ala190 enzymes via hydrogen bonds. Often times this water also plays a similar role in Ser190 protease complexes, where H_2O1_{S1} as well as $O\gamma_{Ser190}$ can accept a hydrogen bond from the amidine N1 nitrogen atom in a three-centered hydrogen bond network (Fig. 1a). The N1–H_2O1_{S1} interaction of this network in many trypsin– and uPA–amidine complexes is long (>3.1 Å), and thus either absent or auxiliary to the $O\gamma_{Ser190}$–H_2O1_{S1} hydrogen bond. Although the N1–$O\gamma_{Ser190}$ hydrogen bond is absent in the Ala190 protease–amidine complexes, the N1–H_2O1_{S1} interaction is a full hydrogen bond (Fig. 1b), typically less than 3.0 Å in the thrombin complexes. We reasoned that an amidine inhibitor capable of displacing H_2O1_{S1} would have differential specificity for Ser190 versus Ala190 proteases (Fig. 1c,d), and discovered that binding to an Ala190 anti-target leaves the amidine N1 atom without its full complement of hydrogen bonds (Fig. 1d). Neither $O\gamma_{Ser190}$ nor H_2O1_{S1} is present to complete the hydrogen bonding of N1 (Fig. 1d). Such a hydrogen bonding deficit markedly disfavors binding because the full hydrogen bonding capacity of the inhibitor is realized in solution with associated water molecules and/or ions.

Here we describe the remarkable specificity amplification (up to 220-fold) for Ser190 over Ala190 proteases achieved from a single atom substitution (replacement of a hydrogen ortho to the amidine with a halogen, Fig. 1c) on a unique uPA inhibitor scaffold. Final selectivity ratios as high as 6700-fold are thus obtained. Crystal structures of the Ser190 proteases, uPA and trypsin, and of the Ala190 protease, thrombin, bound by such halo-substituted inhibitor scaffolds are determined and compared with one another and with corresponding complexes involving non-halo counterparts. The structures and inhibition data show that the extraordinary specificity can be achieved through displacement of a water bound at the S1 site (Fig. 1c,d). Thus a one atom substitution on a relatively non-specific, small molecule inhibitor scaffold yields a comparably small scaffold with vastly improved specificity for the uPA target over tPA and other Ala190 anti-targets.

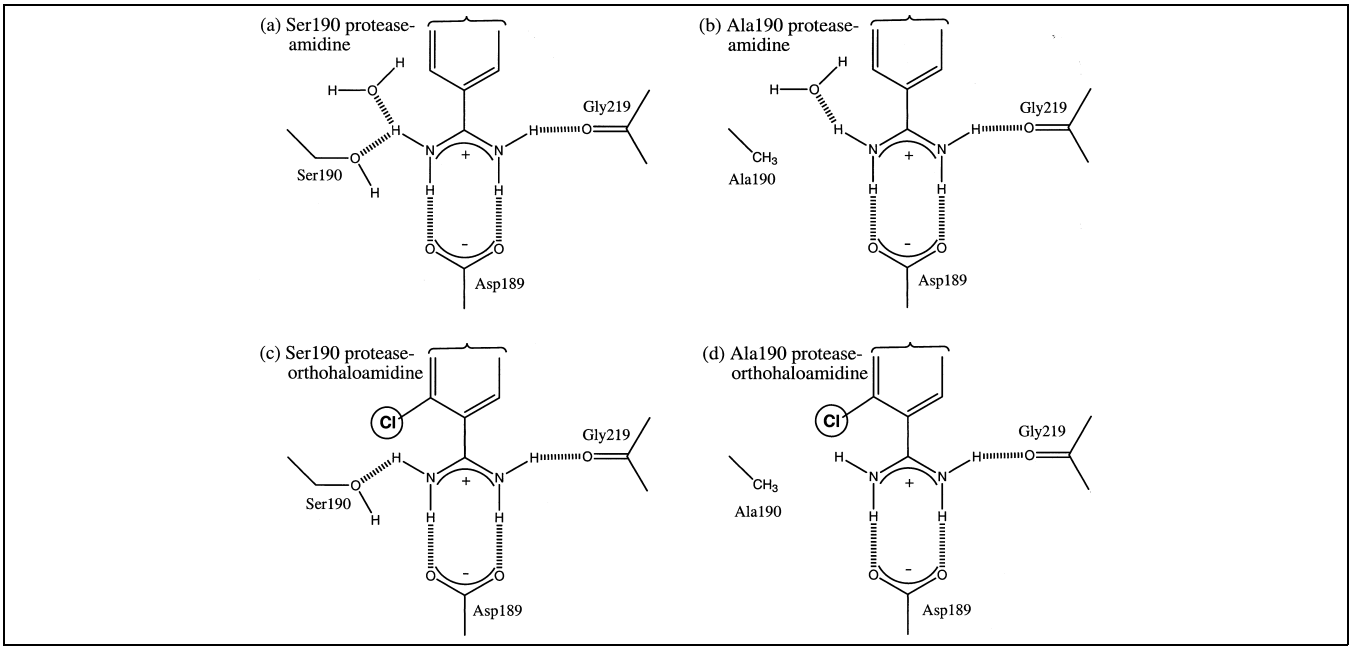


Fig. 1. a: Ser190 protease–APC-8696; b: Ala190 protease–APC-8696; c: Ser190 protease–APC-10302; d: Ala190 protease–APC-10302. H_2O1_{S1} is co-bound in all trypsin-like serine protease complexes of typical amidine inhibitors and in the apo-enzymes [1].

2. Results

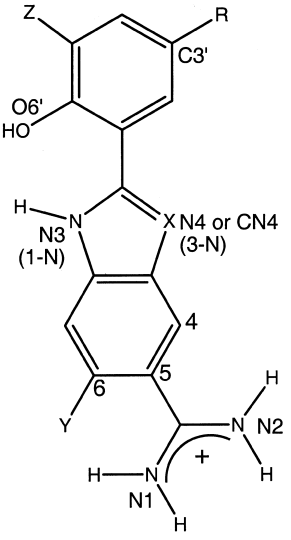
2.1. Choice of inhibitors for selectivity enhancement and of protease–inhibitor complexes for crystallography

We chose one of the most potent, small molecule, reversible uPA inhibitors reported, APC-8696 [16] (Fig. 2a), as one of the scaffolds for selectivity augmentation through incorporation of a halo ortho to the S1 amidine (at the 6 position). APC-8696 belongs to a novel class of active site-directed, small molecule inhibitors that make a network of short hydrogen bonds at the active site [16,17].

We also synthesized and assayed at pH 7.4 other analogs of APC-8696, both indoles and benzimidazoles (Fig. 2a, Table 1a). Because the benzimidazole scaffolds are capable of recruiting Zn^{2+} to mediate binding at the active site [17], we assayed them both in the presence (Table 1d) and absence (Table 1c) of Zn^{2+} at pH 8.2. Finally we studied the effect on selectivity of halo substitution ortho to the amidine of another uPA inhibitor scaffold (APC-7136, Table 1b, Fig. 2b).

For crystallographic visualization and comparison of enzyme–inhibitor interactions at the S1 site we used trypsin [17–21] and thrombin [17,22] as commercially available, readily crystallizable Ser190 and Ala190 protease archetypes, along with the Ser190 uPA target produced in house [1,17]. The thrombin complex with APC-7806, the isosteric benzimidazole analog of APC-8686, yielded larger, better diffracting crystals than did thrombin–APC-8696, and thus the structure of thrombin–APC-7806 was determined and is described. At physiological pH the

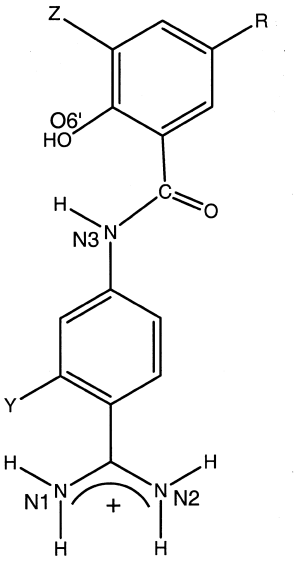
(a) APC-8696 series



APC# (R=H, except where noted)

7806: X= N; Y=H; Z=ϕ
8696: X=CH; Y=H; Z=ϕ
9008: X=CH; Y=H; Z=Cl; R=Cl
9850: X=CH; Y=Cl; Z=Cl; R=Cl
10302: X=CH; Y=Cl; Z=ϕ
10655: X= N; Y=H; Z=O-*cyclo*-C_5H_9
10433: X=CH; Y=H; Z=O-*cyclo*-C_5H_9
10991: X=CH; Y=Cl; Z=O-*cyclo*-C_5H_9
10656: X= N; Y=H; Z=O-CH_2-$CH(CH_3)_2$
11417: X=CH; Y=F; Z=ϕ
10950: X= N; Y=F; Z=O-CH_2-$CH(CH_3)_2$
10762: X= N; Y=Cl; Z=ϕ
10818: X= N; Y=F; Z=ϕ
10972: X= N; Y=Cl; Z=O-CH_2-$CH(CH_3)_2$
11092: X= N; Y=F; Z=O

(b) APC-7136 series



APC#

7136: Y=H; Z=H; R=H
10484: Y=H; Z=Br; R=CH_3
10605: Y=H; Z=I; R=CH_3
11421: Y=Cl; Z=I; R=CH_3
11922: Y=F; Z=I; R=CH_3

Fig. 2. a: APC-8696 scaffolds; b: APC-7136 scaffolds. Non-standard atom names are given in this and other figures. Standard names are given for the benzimidazole nitrogens in brackets.

structures of trypsin–APC-7806 and trypsin–APC-8696 are nearly identical, and the corresponding thrombin complexes are expected to be very similar to one another. The structures of many bound inhibitors were determined in two or all three of the trypsin, thrombin and uPA systems (Table 2). We determined the trypsin-bound structures of all the compounds in Fig. 2, both in the presence and absence of Zn^{2+} for the benzimidazoles.

2.2. *Displacement of H_2O1_{S1} through a single atom substitution on a non-specific inhibitor amplifies selectivity toward Ser190 versus Ala190 proteases*

Although APC-8696 (Fig. 2a) is a potent inhibitor of uPA ($K_i = 0.008$ μM), it has little selectivity (only 4-fold) against tPA ($K_i = 0.035$ μM, Table 1a). However, addition of a 6-chloro, in APC-10302, dramatically amplifies the selectivity toward the Ser190 protease uPA over the Ala190 protease tPA by 220-fold, resulting in a final selectivity ratio of 980 at pH 7.4, without decreasing potency (Table 1a). Similarly, this one atom change in the inhibitor also increases the selectivity toward the Ser190 protease trypsin over the Ala190 counterpart thrombin by 110-fold to a final selectivity ratio of 260 at pH 7.4 (Table 1a).

Comparison of the structures of Ser190 and Ala190 complexes of APC-10302 and APC-8696 (or of the APC-7806 isostere) illustrates the structural basis for selectivity enhancement resulting from displacement of H_2O1_{S1} by the bound inhibitor. In trypsin–APC-8696 (Fig. 3a) there are a total of eight protease–inhibitor or protease–water–inhibitor hydrogen bonds at the active site and oxyanion hole, three of which are very short (shown in cyan), in-

Table 1
Inhibition constants (K_i values, μM) of inhibitors from (a) APC-7806 (benzimidazole) and APC-8696 (indole) series, pH 7.4, (b) APC-7136 series, pH 7.4, (c) APC-7806 series, pH 8.2, in the absence of Zn^{2+}, and (d) APC-7806 series, pH 8.2, in the presence of Zn^{2+}

APC#	Halo[a]	Ser190 protease			Ala190 protease		
		uPA	Trypsin	Plasmin	tPA	Thrombin	Factor Xa
(a) pH 7.4, EDTA							
8 696		0.008	0.13	0.10	0.035	0.32	0.078
7 806		**0.45**	4.4	4.0	0.16	5.5	0.95
10 655		0.22	**0.73**	0.6	0.55	15	1.4
10 656		0.50	3.3	2.5	1.6	30	4.6
9 008		**0.20**	0.85	1.5	0.35	0.21	0.070
10 818	F	0.59	10	6.4	26	47	12
10 950	F	**0.11**	10	2.7	61	23	31
11 092	F	0.011	**6.0**	1.1	29	11	34
11 417	F	**0.020**	1.1	0.35	3.6	7.5	1.3
9 850	Cl	0.31	9.5	1.9	32	17	4.7
10 302	Cl	0.009	0.23	0.11	8.8	60	19
10 762	Cl	2.1	42	7.5	21	120	18
10 972	Cl	1.8	14	2.1	13	90	33
(b) pH 7.4, EDTA							
7 136		**3.8**	**3.2**	6.6	12	**23**	**24**
10 484		**0.077**	**0.28**	4.4	0.29	**0.15**	**0.41**
10 605		**0.33**	**0.61**	3.2	0.19	**0.53**	**2.2**
11 922	F	0.27	2.8	5.5	1.4	0.41	0.50
11 421	C1	6.0	110	>75	210	240	110
(c) pH 8.2, EDTA							
7 806		0.29	3.7	1.7	0.65	5.5	0.95
10 656		0.18	1.3	1.0	2.7	16.0	3.4
10 818	F	0.25	13	7.5	35	110	37
10 950	F	0.16	11	31	50	>450	>450
11 092	F	0.025	5.5	0.95	47	14	25
10 762	Cl	0.85	33	3.9	45	260	50
10 972	Cl	0.49	12	2.3	31	160	55
(d) pH 8.2, Zn^{2+}							
7 806		0.30	0.15	1.0	1.1	0.45	0.43
10 656		0.39	0.026	0.48	5.5	0.55	0.37
10 818	F	0.55	10	4.5	>450	1.2	2.6
10 950	F	2.0	2.3	9.9	370	1.3	1.7
11 092	F	0.14	8.0	2.3	>450	0.60	1.2
10 762	Cl	8.5	21	60	>450	250	21
10 972	Cl	6.5	0.65	2.7	160	55	4.3

Values in bold indicate that competitive inhibition was demonstrated and K_i values determined rigorously by varying both the substrate and the inhibitor concentrations as described [1,17]. The other K_i values were determined at one substrate concentration (per enzyme) set at or near the determined K_m value, while the inhibitor concentration was varied [1,17].

[a]Ortho to the amidine, Fig. 2a.

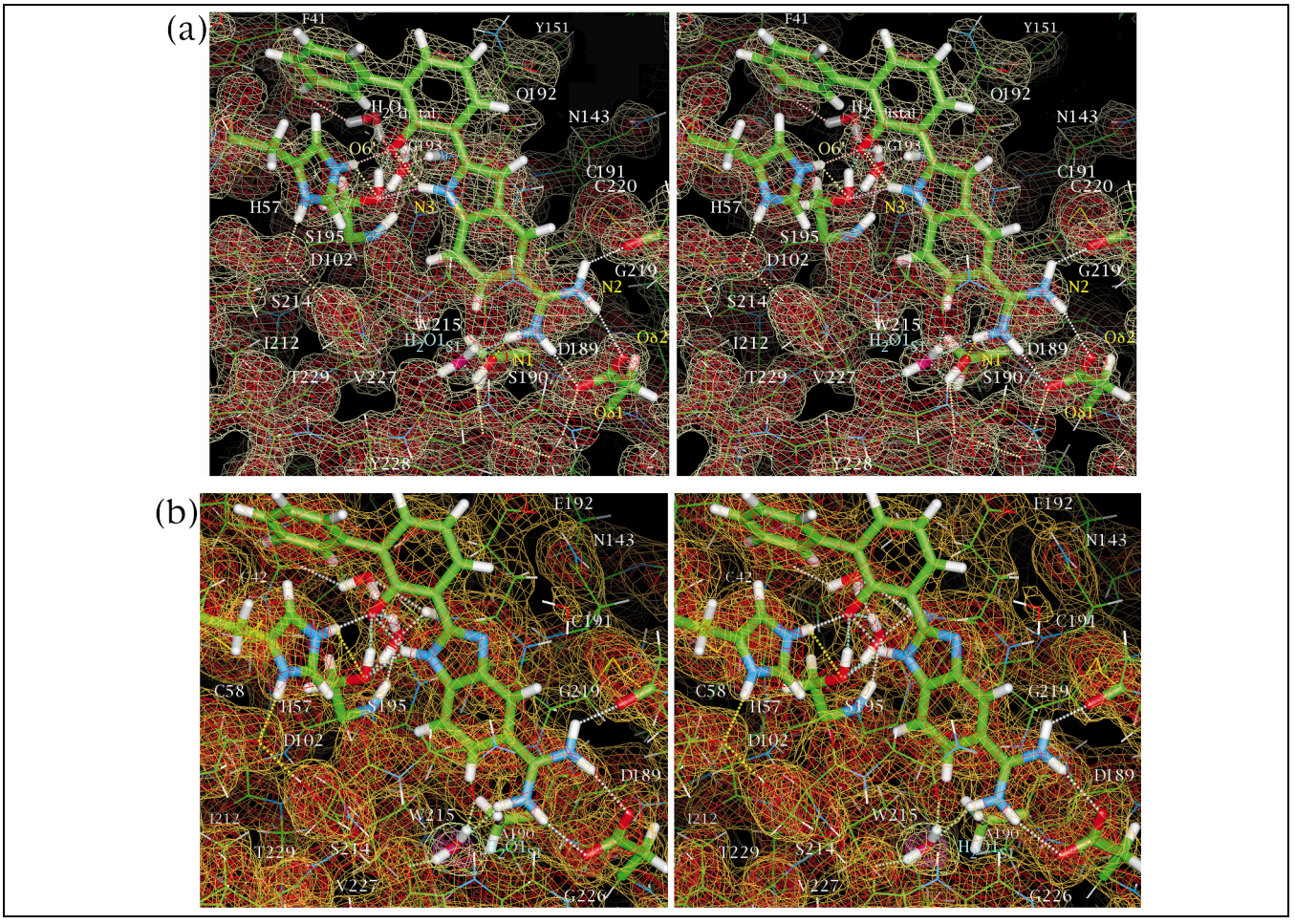


Fig. 3. a: Structure and associated ($2|F_o|-|F_c|$), α_c map of $P3_121$ (form c) trypsin–APC-8696, pH 7.07, 1.37 Å resolution, contoured at 1.0 σ (beige) and 2.4 σ (red). Residue names are labeled white, selected atom names light yellow, and H_2O1_{S1} light blue. Short hydrogen bonds at the active site are cyan, and hydrogen bonds involving H_2O1_{S1} at the S1 site are cyan. The directionalities of the hydrogen bonds involving $O\gamma_{Ser190}$ and $O\eta_{Tyr228}$ can be inferred from the following: a well ordered water (not shown) donates one proton to O_{Ala183}, and the other to $O\delta1_{Asp189}$. This water must therefore accept the hydrogen bond from $O\eta_{Tyr228}$. It also follows that $O\gamma_{Ser190}$ donates a (partial) hydrogen bond to $O\eta_{Tyr228}$, and to H_2O1_{S1}. In addition to the hydrogen bonds accepted from H_2O1_{S1}, the carbonyl oxygens of residues 215 and 227 also accept β-sheet hydrogen bonds, with better angular components, from the peptide nitrogens of residues 227 and 215, respectively, in these and other structures. b: Structure and ($2|F_o|-|F_c|$), α_c map for thrombin–APC-7806, pH 8.68, 1.75 Å resolution.

volving the hydroxyl of Ser195, the phenolate of the inhibitor (O6′) and a water co-bound in the oxyanion hole (H_2O_{oxy}). In thrombin–APC-7806 the inhibitor is also firmly anchored at the active site and oxyanion hole by an array of short and normal hydrogen bonds (Fig. 3b). A summary of hydrogen bond parameters at the active site for over 150 complexes involving the inhibitors of this study is available upon request.

The S1 site hydrogen bonding interactions for trypsin–APC-8696 are typical for trypsin–amidine inhibitor complexes [1,18]. There are two hydrogen-bonded salt bridges from the amidine to the Asp189 carboxylate, one hydrogen bond from N2 to O_{Gly219}, and two partial hydrogen bonds, N1–$O\gamma_{Ser190}$ and N1–H_2O1_{S1}, in a three-centered hydrogen bonding network (Figs. 1a and 3a). H_2O1_{S1} is tightly bound in trypsin–APC-8696 ($B = 13$ Å^2 at pH 8.0) and in other protease–amidine complexes (Table 3b). It receives long hydrogen bonds from N1 and from $O\gamma_{Ser190}$ and donates hydrogen bonds to O_{Trp215} and to O_{Val227} (Fig. 3a).

The interactions at the S1 site of the analogous Ala190 complex, thrombin–APC-7806, are also typical (Fig. 3b, Table 3). The $O\gamma_{Ser190}$–amidine and $O\gamma_{Ser190}$–H_2O1_{S1} hydrogen bonds in the trypsin complex are absent in the Ala190 thrombin counterpart. H_2O1_{S1} makes only three hydrogen bonds in thrombin–APC-7806, to O_{Val227}, to O_{Trp215}, and from N1 (Table 3b).

Introduction of a chloro ortho to the amidine in APC-10302 (Fig. 2a) results in displacement of H_2O1_{S1} upon inhibitor binding to trypsin (Fig. 4a) and to thrombin (Fig. 4b). The exclusion of H_2O1_{S1} from the protease complexes eliminates the N1–H_2O1_{S1} hydrogen bond (Fig. 4). The two partial hydrogen bonds (N1–$O\gamma_{Ser190}$ and N1–H_2O1_{S1}) of the three-centered hydrogen bond network in

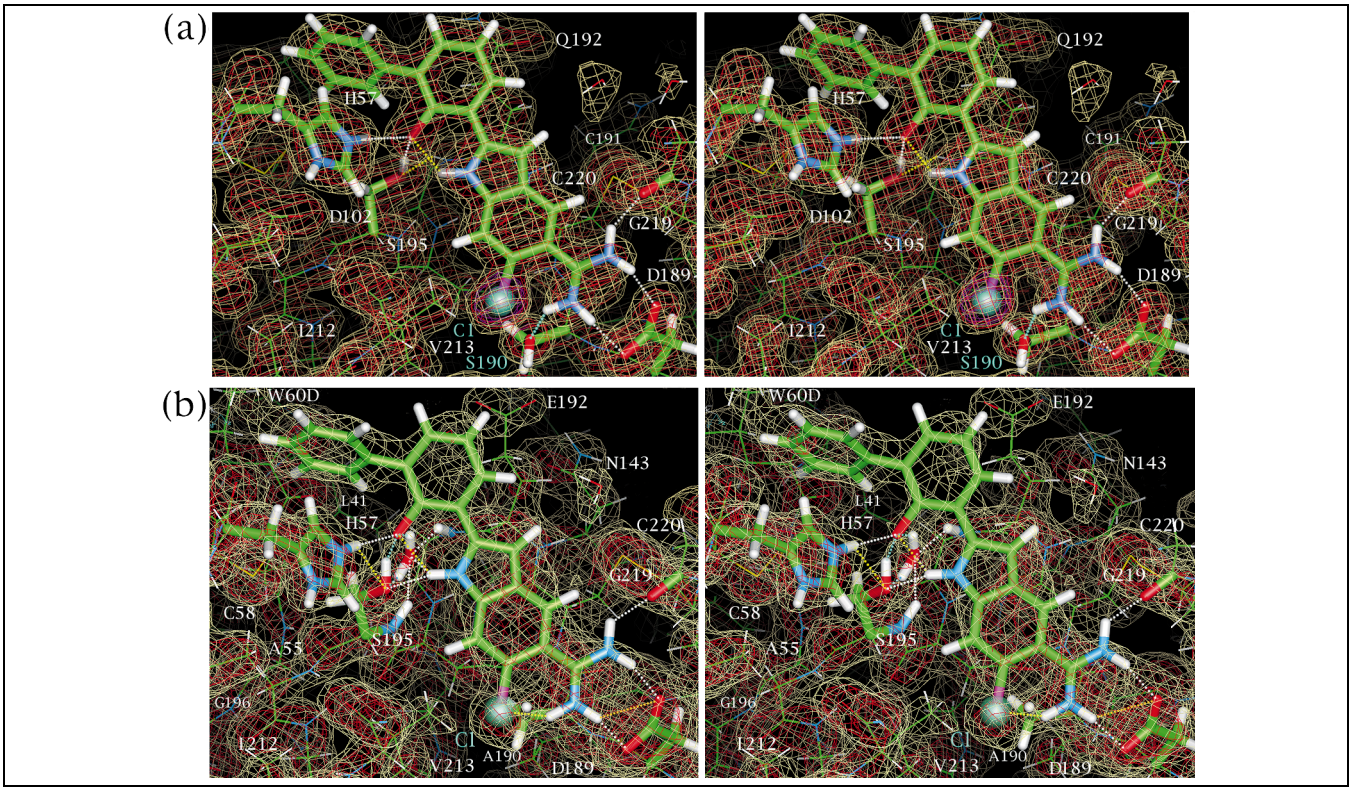


Fig. 4. a: Structure and $(2|F_o|-|F_c|)$, α_c map for trypsin–APC-10302, pH 9.05, 1.50 Å resolution. b: Thrombin–APC-10302, pH 9.0, 1.81 Å resolution. Because of the poor K_i and limited solubility of APC-10302 in the Ala190 thrombin anti-target, the inhibitor is bound with partial occupancy (0.80). The chloro makes contacts with $C\gamma1_{Val213}$ (3.25 Å), $C\beta_{Ala190}$ (3.70 Å), O_{Trp215} (3.83 Å), and $C\alpha_{Gly226}$ (4.14 Å). Hydrogen bonds involving the indole N3 atom are not short, as in thrombin–APC-7806.

trypsin–APC-8696 are replaced with a full hydrogen bond ($N1–O\gamma_{Ser190}$) in trypsin–APC-10302, where the amidine still has a sufficient complement of hydrogen bonds. But removal of the sole $N1–H_2O1_{S1}$ hydrogen bond in thrombin–APC-7806 by displacement of H_2O1_{S1} in thrombin–APC-10302 leaves the amidine H1 atom without a conventional hydrogen bond acceptor (Fig. 4b). The loss of this hydrogen bond, observed in thrombin–APC-7806 and inferred in APC-8696 complexes of thrombin, factor Xa and tPA, results in a sizable decrease in potency of APC-10302 compared to APC-8696 for thrombin, factor Xa, and tPA, by 190-, 240-, and 250-fold, respectively (Table 1a).

2.3. *Bound APC-10302 forms direct hydrogen-bonded salt bridges with Asp189*

The binding to uPA of inhibitors like APC-8696 that form short hydrogen bonds at the active site is atypical for amidine inhibitors of trypsin-like serine proteases in that there are no direct hydrogen bonds from the amidine to Asp189 at the S1 site. Instead a bound water, H_2O2_{S1}, bridges the amidine with one of the carboxylate oxygens of Asp189 [17] (Fig. 5a), emphasizing the firm anchorage of the inhibitor at the active site via the intricate short and normal hydrogen bond network.

Table 2
Complexes for which the structure of a bound inhibitor was determined in multiple enzymes

X	APC#	Resolution (Å)		
		uPA	Trypsin	Thrombin
(a) APC-8696 scaffold				
	7 806	1.90	1.43	1.73
	8 696	1.73	1.41	
	10 655	1.78	1.58	2.00
F	11 417	1.86	1.50	
F	11 092	1.80	1.46	
F	10 950	1.64	1.45	
Cl	10 302	1.50	1.50	1.81
Cl	9 850	1.84	1.43	1.60
(b) APC-7136 scaffold				
	7 136	1.56	1.37	1.50
	10 484	1.60	1.37	1.70
	10 605	1.83	1.38	1.41
F	11 922	2.20	1.50	2.05
Cl	11 421	1.75	1.40	1.67

X is the halo group ortho to the amidine, Fig. 2a. R_{cryst} and other crystallographic diffraction and refinement statistics for a subset of these structures are given in Table 4. All of the structures were determined at physiological pH values. For complexes involving structures at multiple pH values, the resolution of the best diffracting crystal is given.

An unexpected difference between the structures of uPA–APC-8696 and uPA–APC-10302 is the displacement from the latter complex of the amidine–Asp189 bridging water (H_2O2_{S1}), in addition to displacement of the other S1 water (H_2O1_{S1}) that is common to the uPA, trypsin, and thrombin complexes of APC-8696. Displacement of these waters results in direct amidine–Asp189 carboxylate hydrogen-bonded salt bridges in uPA–APC-10302 (Fig. 5b).

2.4. Structural differences at the S1 site incurred by incorporation of the chloro ortho to the amidine

In the APC-10302–protease complexes the indole portion of the inhibitor is oriented somewhat differently from that in APC-8696 counterparts, and is 0.8–1.1 Å deeper in the S1 site, placing the chloride well into the cavity produced by removal of H_2O1_{S1} (Fig. 5c). In uPA–APC-10302 the Asp189 side-chain is shifted from its position in uPA–APC-8696, by ~0.5 Å toward the inhibitor, to form hydrogen bonds with the amidine in uPA–APC-10302 (Fig. 5c).

The large 6-chloro group in APC-10302 sterically favors a more non-planar indole–aryl system. There is a significant change in the indole–aryl dihedral angle, e.g. from $0 \pm 7°$ in trypsin–APC-8696 to $29 \pm 2°$ in trypsin–APC-10302 (Table 3a), incurred by the chloro substitution and by the change in position and orientation of the inhibitor. In thrombin–APC-10302, however, the amidine is coplanar with the indole (dihedral angle = $-1 \pm 5°$), allowing the H1 amidine proton to make a hydrogen bond of 2.7 Å with the chloro. This hydrogen bond is made at the cost of steric strain from close intramolecular contacts resulting from the planarity, estimated by ab initio MO calculations [1] to be 5.0 kcal/mol. The relatively short N1–$O\delta1_{Asp189}$ and N2–$O\delta2_{Asp189}$ hydrogen bonds (2.67 Å and 2.65 Å, Table 3a), and the diagonal N1–$O\delta2_{Asp189}$ hydrogen bond (3.15 Å) may all serve to compensate for the absence of a more favorable, conventional hydrogen bond involving the amidine H1 atom.

Table 3
Hydrogen bond lengths and angles, inhibitor dihedrals, and bound water *B*-factors at the S1 site of uPA, trypsin and thrombin complexes of APC-8696 analogs

(a) Parameters not involving bound waters

	X	APC#	pH	N2–$O\delta2_{Asp189}$	N2–O_{Gly219}	N1–$O\delta1_{Asp189}$	N1–$O\delta2_{Asp189}$	N1–$O\gamma_{Ser190}$	N1–H1–$O\gamma_{Ser190}$	$O\gamma_{Ser190}$–$O\eta_{Tyr228}$	X–$O\gamma_{Ser190}$	Inhibitor dihedral	*n*
Urokinase	Cl	10302	6.5	3.13	2.73	2.96	3.25	2.75	131	2.85	3.34	31	1
	F	11092	6.5, 8.2	3.37(03)	2.70(01)	3.15(10)	3.23(08)	2.83(01)	117(4)	3.12(00)	2.93(01)	4(2)	2
Trypsin	H	8696	4.6–10.9	2.98(09)	2.68(06)	2.93(08)	3.33(09)	2.95(09)	127(5)	2.78(07)		0(7)	17
	F	10818	4.6–10.9	2.96(09)	2.65(05)	2.88(07)	3.27(09)	2.89(04)	134(3)	2.72(09)	2.93(05)	18(4)	8
	Cl	10302	4.4–10.7	2.90(08)	2.70(03)	2.89(05)	3.24(03)	2.84(09)	131(3)	2.67(09)	3.23(11)	29(2)	8
Thrombin	H	7806	7.8–10.8	2.95(11)	2.70(12)	3.04(14)	3.56(12)					−16(7)	5
	Cl	10302	7.8–9.0	2.67(08)	2.61(02)	2.65(16)	3.15(10)					−1(5)	2

(b) Parameters involving H_2O1_{S1}

	X	APC#	pH	N1–$H_2O\mathbf{1}_{S1}$	N1–H1–$H_2O\mathbf{1}_{S1}$	$O\gamma_{Ser190}$–$H_2O\mathbf{1}_{S1}$	$O_{Val/Phe227}$–$H_2O\mathbf{1}_{S1}$	O_{Trp215}–$H_2O\mathbf{1}_{S1}$	$O\gamma_{Ser190}$–$O\eta_{Tyr228}$	B $H_2O\mathbf{1}_{S1}$	F–$H_2O\mathbf{1}_S$	*n*
Urokinase	F	10950	6.5	3.49	170	2.92	3.13	3.52	3.02	28	2.76	1
Trypsin	H	8696[a]	4.6–10.9	3.30(20)	141(5)	3.23(14)	2.87(11)	3.16(10)	2.78(08)	29(9)		17
Thrombin	H	7806	7.8–10.8	2.94(03)	151(4)		2.96(11)	3.11(11)		15(5)		5

(c) Parameters involving H_2O-mediated H-bonds with Asp189

	X	APC#	pH	N2–$H_2O\mathbf{2}_{S1}$	$H_2O\mathbf{2}_{S1}$–$O\delta2_{Asp189}$	N2–O_{Gly219}	N1–$H_2O\mathbf{2}_{S1}$	N1–$O\gamma_{Ser190}$	N1–H1–$O\gamma_{Ser190}$	$O\gamma_{Ser190}$–$O\eta_{Tyr228}$	Inhibitor dihedral	*n*
Urokinase	H	8696, 7806	6.5	2.54(08)	2.62(01)	2.57(03)	3.00(07)	3.04(03)	103(1)	2.88(03)	−21(1)	2
Urokinase	F	10950	6.5	3.06	2.68	2.53	2.67	3.30	115	3.02	−19	1
Urokinase	F	11417	6.5	2.84	2.49	2.48	2.76	2.76	133	3.25	−6	1

	X	APC#	pH	N1–$H_2O\mathbf{1}_{S1}$	N1–H1–$H_2O\mathbf{1}_{S1}$	$O\gamma_{Ser190}$–$H_2O\mathbf{1}_{S1}$	$O_{Val/Phe227}$–$H_2O\mathbf{1}_{S1}$	O_{Trp215}–$H_2O\mathbf{1}_{S1}$	B $H_2O\mathbf{1}_{S1}$	B $H_2O\mathbf{2}_{S1}$	F–$H_2O\mathbf{1}_{S1}$	F–$O\gamma_{Ser190}$	*n*
Urokinase	H	8696, 7806	6.5	2.82(03)	146(5)	3.08(05)	3.08(03)	3.12(11)	27(4)	32(4)			2
Urokinase	F	10950	6.5	3.49	170	2.92	3.13	3.52	28	16	2.76	3.08	1
Urokinase	F	11417[b]	6.5	3.27	143	2.76	3.04	3.35	25	7	2.52	2.77	1

An expanded version of this table containing a summary of the S1 site parameters for 185 structures involving both the APC-8696 and APC-7136 scaffolds is available upon request.

[a] In some cases, where N1–H_2O1_{S1} is long (> 3.4 Å), H_2O1_{S1} makes a hydrogen bond with the π system of Tyr228.

[b] There is a three-centered H-bond network involving F, H_2O1_{S1} and $O\gamma_{Ser190}$. In this structure $O\gamma_{Ser190}$ does not donate a hydrogen bond to $O\eta_{Tyr228}$ (distance = 3.25 Å) as it does in many other uPA and trypsin structures.

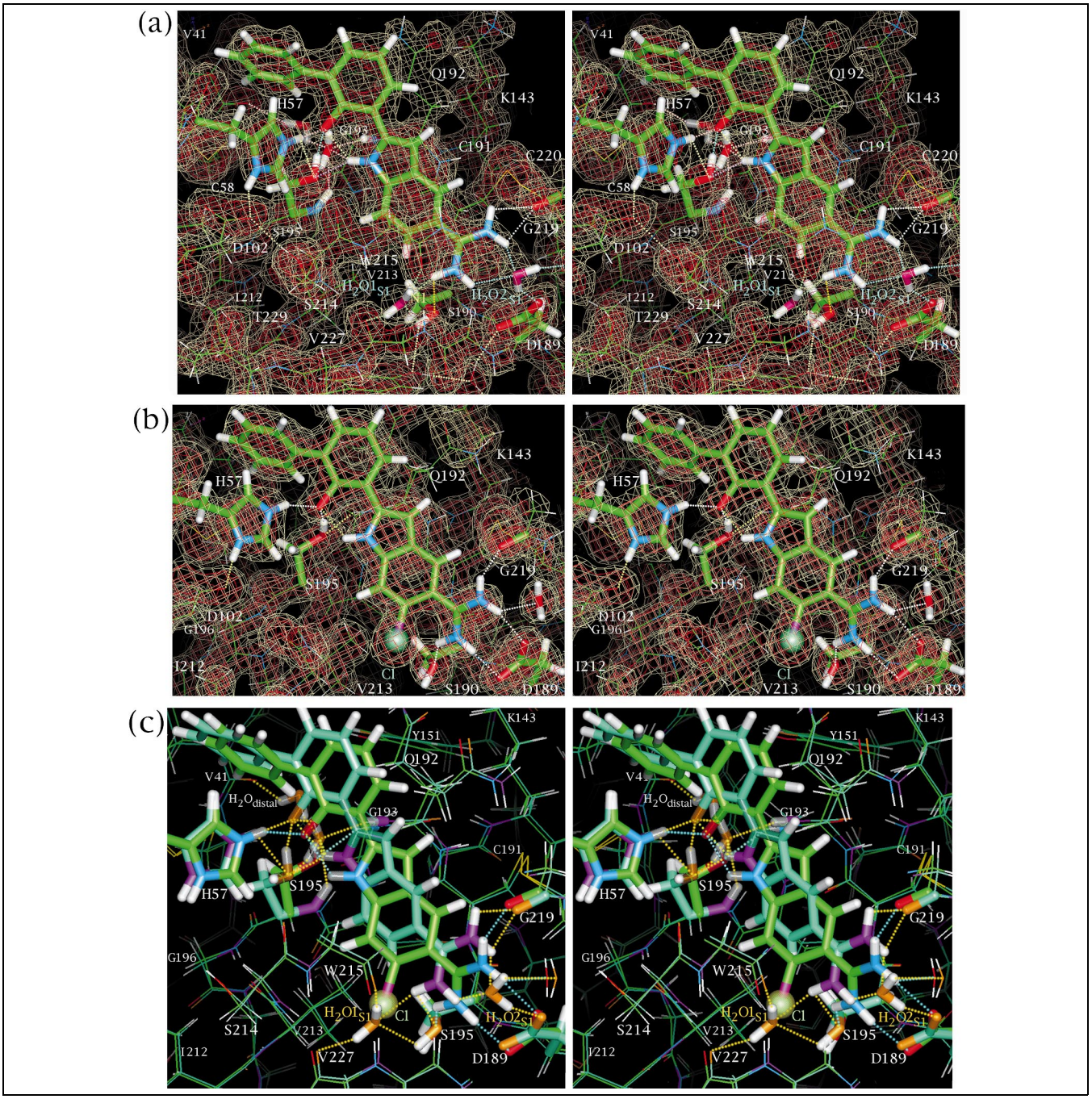


Fig. 5. Structure and $(2|F_o|-|F_c|)$, α_c map for (a) uPA–APC-8696, pH 6.50, 1.73 Å resolution, and (b) uPA–APC-10302, pH 6.50, 1.50 Å resolution. c: Superimposed structures of uPA–APC-8696 and uPA–APC-10302.

2.5. *Absence of waters (H_2O_{oxy} and H_2O_{distal}) bound at the oxyanion hole of uPA and trypsin complexes of APC-10302*

Replacement of bound APC-8696 (or APC-7806) by APC-10302 in uPA or trypsin is associated with a conformational change in Ser195 which shifts $O\gamma_{Ser195}$ by over 1 Å into the oxyanion hole. The shift in $O\gamma_{Ser195}$ breaks the N3–$O\gamma_{Ser195}$ and $N\varepsilon2_{His57}$–$O\gamma_{Ser195}$ hydrogen bonds and displaces H_2O_{oxy} and H_2O_{distal} (Fig. 4a; Fig. 5b). Loss of H_2O_{oxy} upon binding of APC-10302 to uPA converts the three-centered short hydrogen bonding array in uPA–APC-8696 to a single $O\gamma_{Ser195}$–O6′ hydrogen bond of normal length (2.75 Å) in uPA–APC-10302.

2.6. *Favorable interactions of the uPA-bound APC-10302 scaffold*

A total of four ordered waters (H_2O1_{S1}, H_2O2_{S1}, H_2O_{oxy}, and H_2O_{distal}) present in uPA-8696 are absent in

uPA–APC-10302. These waters appear well oriented through the many associated hydrogen bonds (Fig. 5a), and the first three have relatively low *B*-factors. The entropy change upon the displacement of some of these waters may be close to maximal, 2.0 kcal/mol per water [23].

Displacement of H_2O2_{S1} by bound APC-10302 results in a conversion of H_2O2_{S1}-bridged amidine–Asp189 hydrogen bonds in uPA–APC-8696 (Fig. 5a) to direct and presumably stronger ones in uPA–APC-10302 (Fig. 5b). Also in the latter complex, $O\delta1_{Asp189}$ has a full complement of hydrogen bonds, whereas in the H_2O2_{S1}-bridged binding mode of uPA–APC-8696, it does not (Fig. 5a).

The indole scaffolds, like APC-10302, exhibit significantly greater intrinsic affinity over the benzimidazole counterparts toward all the proteases in Table 1. Conversion of the non-halo (APC-7806), the 6-fluoro (APC-10818) and 6-chloro (APC-10762) benzimidazoles to the corresponding indoles (APC-8696, APC-11417, and APC-10302, respectively) significantly increases inhibitor potency for all enzymes in Table 1a by average factors of 27, 13 and 82, respectively. Increased potency of the indole scaffold reflects, in part, a less costly change in internal energy to convert the lowest energy solution conformer to the binding conformer than for the benzimidazole [17]. The benzimidazole conformer with the nitrogen at the 1-position unprotonated (the non-binding tautomer) is more stable in the unbound state than the other tautomer due to withdrawal (by resonance) of electrons from the nitrogen at position 1 by the amidine. The indole, with only one possible internal hydrogen bond possible, is locked into the binding conformation. The amplification of this affinity difference between the indole and benzimidazole scaffolds by the 6-chloro group is thought to occur through its donation of electrons by resonance to the carbons adjacent to the nitrogen at position 3 and withdrawal of electrons by induction from the nitrogen at position 1 of the benzimidazole. Higher pK_a values of the amidine and phenol of the indoles may also increase hydrogen bond strengths at the S1 and active sites.

2.7. *Dual binding potential of ortho (fluoro, amidino) aryl inhibitor scaffolds in uPA*

Replacement of the chloro of uPA-bound APC-10302 with a fluoro in the APC-11417 isostere converts a direct hydrogen-bonded amidine–Asp189 S1 binding mode into one mediated by two waters (H_2O1_{S1} and H_2O2_{S1}). In this

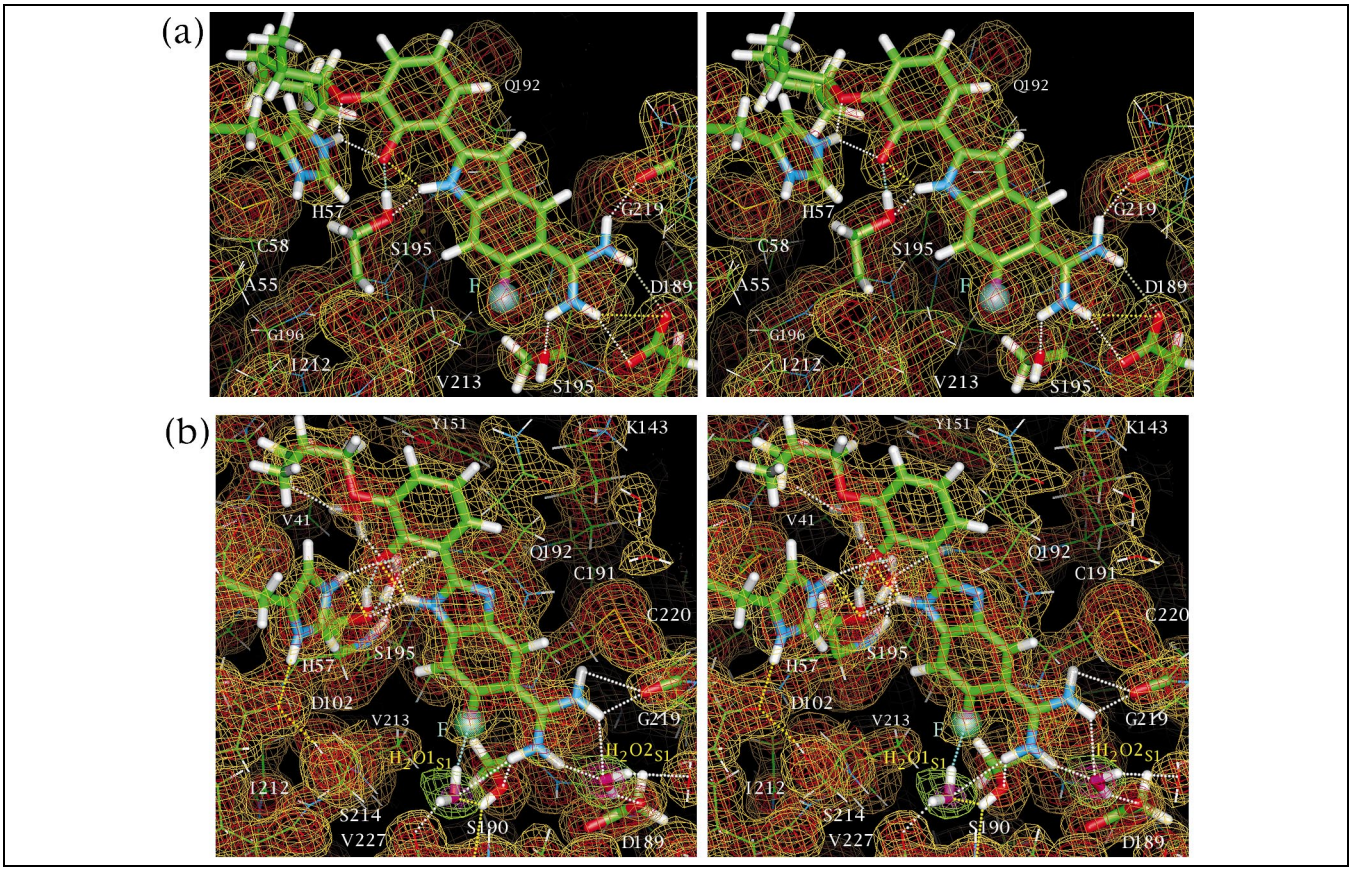


Fig. 6. a: Structure and $(2|F_o|-|F_c|)$, α_c map for uPA–APC-11092, 1.80 Å resolution. The long N1–$O\delta2_{Asp189}$ and N2–$O\delta2_{Asp189}$ interactions (3.17 Å and 3.40 Å, respectively, at pH 6.5) are shown in yellow and transparent light yellow, respectively. b: Structure and $(2|F_o|-|F_c|)$, α_c map for uPA–APC-10950, 1.64 Å resolution. The H_2O1_{S1}–F hydrogen bond (2.76 Å) is formed at the expense of the H_2O1_{S1}–O_{Trp215} interaction, the length of which increases from 3.07 Å in uPA–APC-8696 to 3.52 Å in uPA–APC-10950 (Table 3c).

water-mediated S1 binding mode, APC-11417 sits higher in the S1 site, like the uPA-bound parent compound, APC-8696. In uPA–APC-11417 H_2O1_{S1} donates a relatively short hydrogen bond (2.52 Å) to the fluoro (Table 3c).

Such a change in binding can also be effected by a change in the inhibitor group probing S1′ in cases where a fluoro is ortho to the amidine. Thus APC-11092, with a methylcyclohexyl ether group ortho to the phenol hydroxyl, makes direct hydrogen-bonded salt bridges with Asp189 (Fig. 6a), whereas APC-10950, with an isobutyl ether group ortho to the phenol hydroxyl, makes water-mediated hydrogen bonds with Asp189 (Fig. 6b). In the water-mediated S1 binding mode of APC-10950, the ether oxygen does not make hydrogen bonds (Fig. 6b), whereas in the direct amidine–Asp189 hydrogen-bonded mode of APC-11092, this oxygen receives a hydrogen bond from $N\varepsilon2_{His57}$ (Fig. 6a). While the relatively large S1 site of uPA is associated with two S1 binding modes of 6-fluoro-5-amidinoindole or 6-fluoro-5-amidinobenzimidazole inhibitors, depending on distal interactions at S1′, only the direct hydrogen-bonded amidine–Asp189, H_2O1_{S1}-displaced mode of binding of such inhibitors is compatible with the smaller S1 site of trypsin.

2.8. *Effects on potency, selectivity, and structure incurred by incorporating a halo ortho to the amidine of another scaffold*

Another uPA inhibitor scaffold, APC-7136 (Fig. 2b), can be considered an analog of APC-8696 in which the five-membered ring is opened: in APC-7136 an amide group connects the phenol with the S1 phenylamidine (compare Fig. 2a and b). The APC-7136 scaffold exhibits two distinct active site binding modes depending on both the inhibitor and the protease (Fig. 7a,b).

In APC-11421 and APC-11922 (Fig. 2b) there is a chloro or fluoro, respectively, ortho to the amidine, as for the halo-substituted APC-8696 analogs, described above. Incorporation of the fluoro in APC-11922 does not result in Ser190/Ala190 protease selectivity; APC-11922 exhibits about the same potency toward all proteases in Table 1 (within a factor of 20) as the parent, APC-10605. The crystal structures of the uPA, trypsin and thrombin complexes of APC-11922 show that H_2O1_{S1} is not displaced from the S1 site (Fig. 7a) in any of these three systems. In each complex the fluoro makes an unusually short hydrogen bond with the retained water (2.5–2.7 Å). These structures demonstrate that lack of se-

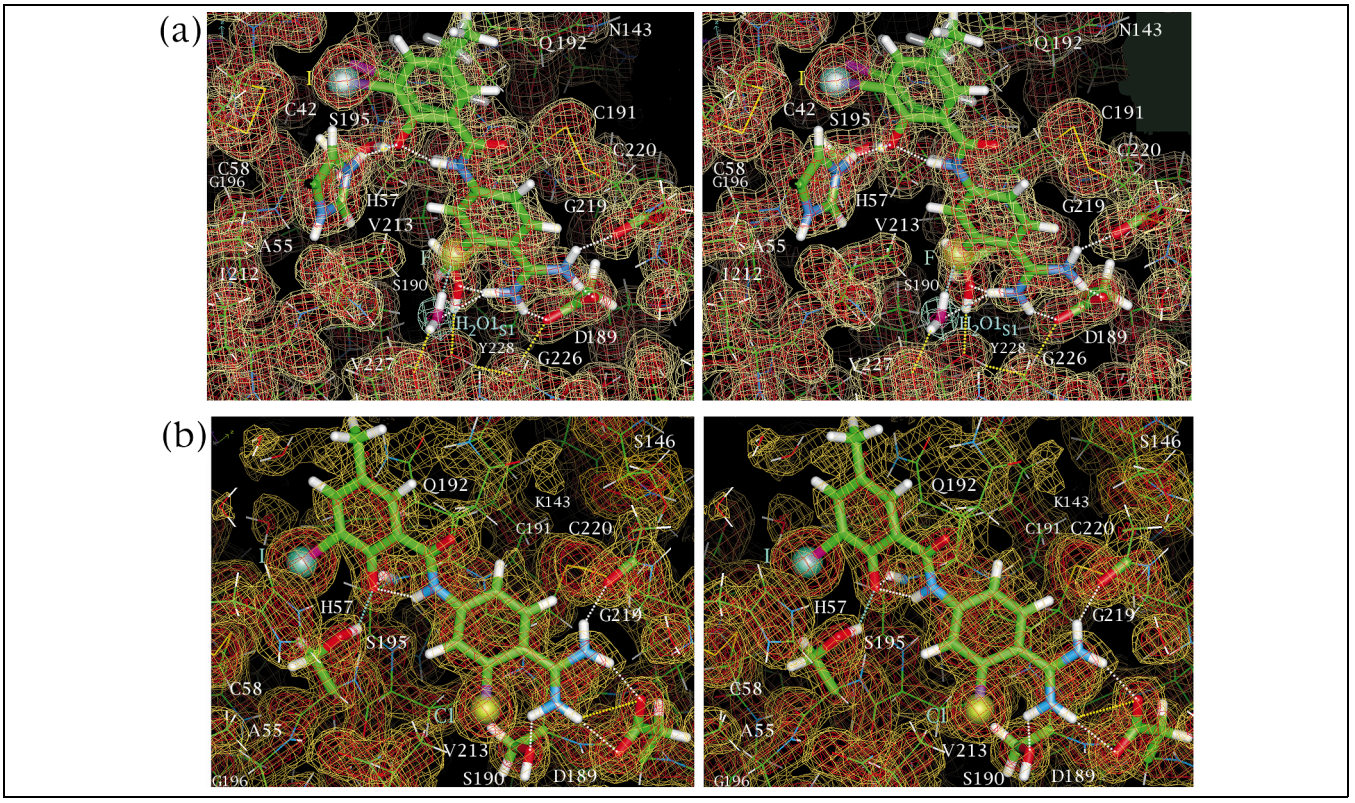


Fig. 7. a: Structure and $(2|F_o|-|F_c|)$, α_c map of trypsin–APC-11922, 1.50 Å resolution. For the major inhibitor conformer (opaque sticks) there is a hydrogen bond between the inhibitor phenol (O6′) and $N\varepsilon2_{His57}$, and a short O6′–$O\gamma_{Ser195}$ hydrogen bond. b: Structure and $(2|F_o|-|F_c|)$, α_c map of uPA–APC-11421, 1.75 Å resolution. In this and all other uPA complexes of the APC-7136 analogs in Fig. 2b, the phenol hydroxyl is at or near the oxyanion hole, receiving hydrogen bonds from N_{Gly193} and from the inhibitor NH amide group. In many trypsin and thrombin complexes the inhibitor is discretely disordered between the two binding modes in (a) and (b).

lectivity is correlated with failure of the fluoro to displace H_2O1_{S1}.

Displacement of H_2O1_{S1} by the larger chloro group is observed in the trypsin and uPA complexes of APC-11421 (Fig. 7b), and expected in the other APC-11421 complexes of the proteases in Table 1. Unlike the ortho (halo, amidino) benzimidazoles or indoles the ortho (chloro, amidino) phenyl compound, APC-11421, does not show general selectivity for Ser190 versus Ala190 proteases; it is moderately selective for uPA only, by 35-fold over tPA. Reduced affinity for Ala190 proteases like tPA is expected to occur for the same reason as for APC-10302, through a hydrogen bonding deficit involving the bound amidine when H_2O1_{S1} is displaced. However reduction in binding to a particular protease, either Ser190 or Ala190, can also occur due to loss of favorable interactions at the active site, especially when the inter-relationship of the S1 and active sites prohibits optimal enzyme–inhibitor interactions at both loci simultaneously, as observed in trypsin–APC-11421.

The S1–active site separation is smaller in trypsin than in uPA by $\sim$0.5 Å. Thus, whereas in uPA–APC-11421 the phenol hydroxyl (O6′) occupies the oxyanion hole, in the trypsin complex O6′ is shifted out of it by 0.7 Å. The O6′–$O\gamma_{Ser195}$ hydrogen bond (2.75 Å) is therefore not as short as in the uPA complex (2.39 Å). Van der Waals interactions of the inhibitor iodide are also poorer in the trypsin complex. This comparison emphasizes the observation that structural features in addition to the side-chain of residue 190 differentiate the S1 sites of trypsin-like serine proteases and constitute a unique signature of each protease [1]. The depth, width, and spatial relationship of the S1 and active site differ significantly from one protease to another, even when the local sequences that form these sites are the same [1].

2.9. *Halo selectivity switch and Zn^{2+}-mediated binding of benzimidazole scaffolds*

The potency at pH 8.2 of the non-halo benzimidazoles is essentially unaltered by Zn^{2+}, toward uPA, tPA and plasmin, or improved, by up to 50-fold, toward trypsin, thrombin and factor Xa (Table 1c,d). Crystallographically determined Zn^{2+}-mediated binding modes have been described for trypsin and thrombin [17], and determined for all the trypsin complexes of the halo and non-halo benzimidazoles in Fig. 2a. Zn^{2+} can exert pronounced effects on the activity of the 6-halo benzimidazoles, but these effects can be adverse as well as beneficial. The dramatic selectivity for uPA over tPA in the absence of Zn^{2+} (by up to 1900 at pH 8.2, Table 1c) is preserved in the presence of Zn^{2+} (by up to >3200, Table 1d). However the selectivity of the 6-fluoro-5-amidinobenzimidazoles against the other Ala190 proteases, thrombin and factor Xa is lost due to increased affinity (by up to 350-fold) in the presence of Zn^{2+}, possibly through retention of H_2O1_{S1}. While the 6-fluoro analogs are selective against only one of the Ala190 proteases (tPA) in the presence of Zn^{2+}, the 6-chloro analogs are selective against two of them (tPA and thrombin) (Table 1d). The potential participation and often significant effect of Zn^{2+} on the activity of benzimidazole inhibitors that can operate in a Zn^{2+}-dependent mode introduces another diverse inhibitor binding manifold along with unique opportunities to consider for in vivo inhibition.

3. Discussion

3.1. *Selectivity enhancement through displacement of a bound water at the S1 site via a one atom substitution*

Displacement of a tightly bound water molecule by a rigid inhibitor group is a common strategy to increase inhibitor affinity [24], due to the increase in entropy (up to 2.0 kcal/mol [23]) upon the release of the water. If the water is present in a target but absent in an anti-target an increase in selectivity is thereby achieved. Thus, displacement by an appropriate inhibitor group of a tetrahedrally hydrogen-bonded water present in HIV protease but absent in related aspartic acid proteases such as pepsins proved to be a successful strategy in the design of potent, bioavailable HIV protease inhibitors [25,26].

We are unaware, however, of any previous examples where inhibitor specificity is achieved from displacement of a tightly bound water that is conserved in both the target and the anti-target. Our successful design of inhibitors highly selective for the S1 site of uPA relies on the fact that the enzyme–H_2O1_{S1}–inhibitor interactions are disposable in the Ser190 uPA target but essential in the Ala190 tPA anti-target for these types of inhibitors. Thus the selectivity enhancements we achieved result primarily from a large decrease in inhibitor affinity toward the anti-target, rather than from an increase toward the target.

Through a single atom substitution we achieved increases in selectivity toward the uPA Ser190 protease target over the Ala190 protease anti-targets tPA, thrombin, and factor Xa, of 220-, 170-, and 220-fold, respectively, at pH 7.4. Final selectivity ratios as high as 6700 were thus realized. The displacement of a firmly bound water (H_2O1_{S1}) and ensuing loss of a critical hydrogen bond from the inhibitor amidine in the Ala190 protease complexes is a novel and effective strategy for achieving exceptional selectivity toward uPA and other Ser190 protease targets over Ala190 counterparts.

In many uPA, trypsin and other Ser190 protease–inhibitor complexes, H_2O1_{S1}, $O\gamma_{Ser190}$, and an amidine nitrogen (N1) participate in a three-centered hydrogen bonding network (Fig. 1a). Upon displacement of H_2O1_{S1} from the Ser190 protease complexes by the halo substituent of the bound inhibitor, the amidine still has a sufficient complement of hydrogen bonds: the two partial hydrogen bonds,

N1–H_2O1_{S1} and N1–$O\gamma_{Ser190}$, are replaced with a single full hydrogen bond, N1–$O\gamma_{Ser190}$ (Fig. 1c). In Ala190 proteases such as tPA, thrombin, and factor Xa, displacement of H_2O1_{S1} greatly disfavors inhibitor binding, because it leaves one of the protons on N1 without a conventional hydrogen bond (Fig. 1d).

By exploiting a subtle difference between the highly similar S1 sites of Ser190 and Ala190 proteases the design principle achieves potent, selective inhibitors with only a modest increase in molecular weight, thereby increasing the potential for bioavailability. A focus on small S1 recognition elements is also compatible with the design of relatively rigid inhibitors that are not as susceptible to the detrimental loss of conformational entropy upon binding as are larger, more flexible inhibitors.

3.2. *Specificity enhancement by displacement of H_2O1_{S1} from the Ala190 proteases only*

The relatively small atom size and bond length involving fluoro allow its incorporation at the relatively large S1 site of uPA without H_2O1_{S1} displacement in some of the complexes such as uPA–APC-11417 and uPA–APC-10950 (Fig. 6b). However, APC-11417 and APC-10950 still show specificity for Ser190 over Ala190 proteases, up to 450-fold for APC-10950 toward uPA over tPA, and up to 380-fold for APC-11417 toward uPA over thrombin at pH 7.4 (Table 1a). In the tPA and thrombin complexes these inhibitors are expected to make direct hydrogen bonds with Asp189 (the prevalent mode for trypsin-like serine proteases), involving displacement of H_2O1_{S1} by the fluoro, as observed in the structures of the corresponding trypsin complexes. Thus, to generate Ser190/Ala190 protease selectivity, H_2O1_{S1} needs not be displaced in the Ser190 enzyme, only in the Ala190 counterparts.

The crystal structures and K_i values involving the phenylamidine APC-7136 scaffold demonstrate the importance of H_2O1_{S1} displacement for specificity generation by ortho (fluoro, amidino) phenyl inhibitors. The fluoro analog, APC-11922, shows no specificity for Ser190 over Ala190 proteases, similar to the parent analog, APC-10605 (Table 1b). The crystal structures of the APC-11922 complexes of the Ser190 proteases trypsin and uPA and of the Ala190 counterpart thrombin show that this lack of specificity reflects failure to displace H_2O1_{S1} in these complexes.

3.3. *Preference for Ser190 versus Ala190 proteases of small, S1-directed amidines versus those directed to both the S1 and active sites*

Whereas a general, modest preference of small, non-active site-directed amidine inhibitors for Ser190 proteases over Ala190 counterparts was observed and attributed to an additional, $O\gamma_{Ser190}$–amidine hydrogen bond in the former complexes [1], this preference is not apparent in the active site-directed, non-halo inhibitors of this study. Interactions at the active site or elsewhere can mask the effects at S1. For example, many of the indole or benzimidazole amidine inhibitors in this study make a short hydrogen bond with H_2O_{oxy} [17]. Because H_2O_{oxy} binds to apo-thrombin but not to apo-trypsin [17], there is a less favorable entropy change, involving co-binding and ordering of H_2O_{oxy}, accompanying binding of the active site-directed inhibitors to trypsin than to thrombin. Opposing free energy changes at different loci provide a rationale for why the selectivity trends for simple amidine inhibitors [1] do not generally translate to the more elaborated, active site-directed non-halo arylamidine inhibitors of this study (Table 1).

3.4. *Effect on binding to trypsin versus uPA of ortho halo substitution on the two inhibitor scaffolds*

In uPA an enthalpy gain accompanies the binding of APC-10302 over uPA-8696 through formation of direct hydrogen bonds, whereas in trypsin there is no such change because both inhibitors make direct hydrogen bonds with Asp189. There is also no entropy gain in trypsin from displacement of H_2O2_{S1}, which is absent in all trypsin complexes. The net effect of the changes in interactions at S1 (and elsewhere) between binding of APC-10302 and APC-8696 on inhibitor affinity is small ($<$ 2-fold) for both uPA and trypsin.

Similar considerations apply to the APC-7136 scaffold (Fig. 2b), except both halo and non-halo analogs make typical hydrogen-bonded salt bridges at the S1 site of uPA (and of trypsin and thrombin). The active site binding modes of the APC-10605 parent and the APC-11421 chloro analog are very similar in uPA, but the K_i for APC-11421 is 18-fold higher. This comparison suggests that the net effect of displacement of H_2O1_{S1} by the halo, in the absence of other changes at S1 or at the active site, is somewhat unfavorable.

3.5. *Exploiting interactions at S1′ for potency and selectivity*

The scaffolds of this study are ideal for probing and optimizing interactions at the S1′ site from aromatic (Figs. 3–5) or aliphatic (Fig. 6) substituents. The ether linkages in compounds like APC-11092 (Fig. 6a) and APC-10950 (Fig. 6b) offer the opportunity for a library approach to optimization of S1′-directed substituents. Thus APC-11092, developed by this approach, is $\geq$ 100-fold selective for uPA against all other proteases in Table 1.

Even single halo atom substituents directed at S1′ in the phenylamidine scaffold can modulate specificity through differential interaction with Cys42–Cys58 disulfide at S1′. Although this disulfide is a conserved feature of trypsin-like serine proteases, its precise location and orientation relative to the active and S1 sites are a distinguishing fea-

ture of each protease. Thus the bromo group in APC-10484 increases potency by 150-fold toward thrombin but by only 1.5-fold toward plasmin (Table 1b).

3.6. Cooperativity of S1, active site, and distal binding elements in determining inhibitor potency and bound structure

Modulation of the size of the halo group or of the size and character of an alternate group (such as OH, CH_3, or OCH_3 [27]) ortho to the S1 amidine could be a powerful way of tuning selectivity toward individual Ser190 targets based on subtle differences in the size of their S1 sites. The K_i values and corresponding structures of the complexes of this study show how subtle changes in depth, width, and spatial relationship of the S1 and active sites can affect inhibitor potency. Thus, while dramatic selectivity can be achieved against Ala190 proteases through displacement of H_2O1_{S1}, such a change can also compromise binding to individual members of the Ser190 protease class because of their unique S1 sites, as observed for APC-11421, which retains reasonable affinity (6.0 μM) for uPA only. Similarly, cooperative interplay of interactions at the S1, S1′ and active sites results in high affinity (11 nM) and high selectivity (≥100-fold) of APC-11092 toward uPA only. Notable selectivity within the set of Ser190 proteases (up to 57-fold), and within the set of Ala190 proteases (up to 750-fold) is also observed for some of the benzimidazoles in the Zn^{2+}-mediated active site binding mode.

From crystallographic study of several similar inhibitors in multiple proteases, we observed a remarkable diversity in inhibitor binding, underscoring the importance of high-throughput crystallography for structure-based development of selective inhibitors. We discovered many distinct active site and S1 site hydrogen bonding modes, in which the relative location of a particular inhibitor scaffold at these sites could vary by over 1 Å. Observation of multiple, distinct binding modes in each of the lead scaffolds demonstrates tight coupling of the S1, S1′ and active sites and of the corresponding inhibitor recognition elements in determining the bound structure. Subtle differences in these structural features among the proteases can impart inhibitor discrimination and thus afford opportunities for engineering selectivity.

4. Significance

Drug development of inhibitors of trypsin-like serine proteases involved in pathogenesis can provide therapy for associated disease states. Recently we reported an intrinsic preference of arylamidine inhibitors toward Ser190 trypsin-like serine proteases over Ala190 counterparts, attributed to an additional hydrogen bond between the amidine and $O\gamma_{Ser190}$ in the former protease class. Here we have shown specifically how to harness the Ser/Ala190 difference to amplify this selectivity as high as 220-fold, with a one atom substitution on the S1 binding element of a unique uPA inhibitor scaffold. Overall selectivity values as high as 6700 are thus achieved.

In Ser190 protease complexes the inhibitor amidine N1 atom donates a hydrogen bond to $O\gamma_{Ser190}$, and may also donate a hydrogen bond to H_2O1_{S1}. Upon displacement of H_2O1_{S1} from the Ser190 protease complexes by the halo of the bound inhibitor the amidine still has a satisfactory complement of hydrogen bonds: in cases where H_2O1_{S1} participates in a three-centered hydrogen bond network the two partial hydrogen bonds, N1–H_2O1_{S1} and N1–$O\gamma_{Ser190}$, are replaced with a single full hydrogen bond, N1–$O\gamma_{Ser190}$. In corresponding inhibitor complexes of Ala190 counterparts such as tPA, thrombin and factor Xa, displacement of H_2O1_{S1} seriously compromises inhibitor binding because it leaves one of the protons on N1 without a conventional hydrogen bond.

A focus on structural differences at the S1 site for selectivity engineering allows development of exceptionally small inhibitors to increase the potential for bioavailability. Thus we have designed and developed low nanomolar, low molecular weight inhibitors, remarkably selective for the Ser190 protease cancer target, uPA, and against the Ala190 anti-targets tPA, thrombin, and factor Xa. The unique basis of specificity, displacement of a conserved water firmly bound at the same location of the S1 site of both the target and the anti-target to remove a critical hydrogen bond in the anti-target, exploits a subtle difference between closely related and highly structurally similar enzymes. This selectivity switch for discriminating between uPA and tPA and between other Ser190 targets and Ala190 anti-targets is a unique and powerful tool for structure-based drug design of potent selective protease inhibitors.

5. Materials and methods

Synthesis of inhibitors is described in detail elsewhere [16,27].

5.1. Enzymology

K_i values were determined at pH 7.4 or at 8.2 in the presence of 100 μM Zn^{2+} or 1.0 mM EDTA as described [1,16,17,28]. Apparent inhibition constants, K_i' values, were calculated from the velocity data collected at various inhibitor concentrations using the software package *BatchKi* (developed and provided by Dr. Petr Kuzmic, Biokin Ltd., Madison, WI, USA) using methodology similar to that described for tight binding inhibitors [29]. K_i' values were converted to K_i values by the formula $K_i = K_i'/(1+S/K_m)$.

5.2. Crystallography

Production and crystallization of mutagenically deglycosylated

Table 4
Crystallography of trypsin (tryp), thrombin (thro) and uPA complexes

PDB accession code (uPA complexes)		1gjb	1gjc	1gj7	1gj9	1gj8
Inhibitor and pH		7806, 6.5	8696, 6.5	10302, 6.5	11092, 6.5	10950, 6.5
No. reflections, R_{merge} (%)[a]		13323, 10.1	17604, 9.2	27459, 6.4	14468, 7.9	35750, 6.1
R_{cryst}, completeness[b]	7.0–2.00 Å	20.5, 76	17.7, 83	16.5, 93	16.6, 74	16.7, 84
	7.00–1.75 Å		20.0, 70	18.2, 86		18.2, 75
	to highest resolution	20.9, 69	20.2, 69	20.3, 71	18.1, 64	19.0, 68
	highest resolution shell (Å)	1.98–1.90	1.81–1.73	1.57–1.50	1.88–1.80	1.71–1.64
Completeness at highest resolution		27.7	33.4	36.3	31.2	34.3
Free R^{c}_{cryst}, to highest resolution		25.9	24.2	23.0	22.2	22.1
Protease, PDB accession code		uPA, 1gja	uPA, 1gjd	thro, 1gj5	thro, 1gj4	tryp, 1gj6
Inhibitor, pH		7136, 6.5	11421, 6.5	7806, 8.7	10302, 9.0	10302, 9.0
No. reflections, R_{merge} (%)[a]		25786, 5.4	16414, 9.1	26014, 6.3	19152, 9.6	30523, 8.3
R_{cryst}, completeness[b]	7.0–2.00 Å	15.9, 93	16.6, 81	17.6, 85	18.7, 68	15.0, 91
	7.00–1.75 Å	17.4, 86	17.8, 68	19.4, 71		16.6, 83
	to highest resolution	18.7, 74	18.0, 68	19.7, 70	20.7, 59	18.9, 68
	highest resolution shell (Å)	1.63–1.56	1.83–1.75	1.81–1.73	1.89–1.81	1.57–1.50
Completeness at highest resolution		31.0	30.4	32.5	30.4	29.5
Free R^{c}_{cryst}, to highest resolution		20.9	18.8	23.3	23.6	20.6

F/σ cutoffs ranged from 1.6 to 2.0, depending on the data-set. Root mean square deviations of bond lengths from ideality ranged from 0.015 to 0.018 Å.

[a] $R_{merge} = \Sigma_h \Sigma_i \ |I(h)_i - \langle I(h)_i \rangle| / \Sigma_h \Sigma_i I(h)_i$.

[b] The first value is the conventional R-factor (in percent), $R_{cryst} = \Sigma(||F_o| - |F_c||)/\Sigma|F_o|$. The second is the overall completeness in percent.

[c] Cross validation R-factor using 10% of the data withheld from refinement.

low molecular weight uPA have been described [1]. X-ray diffraction data were collected with an R-AXIS IV image plate as described [1,17]. The completeness of each data-set before rejection of weak reflections ($F_o/\sigma < 1.7$–2.0) was $> 98.5\%$ for trypsin crystals and $> 96\%$ for the lower symmetry thrombin and uPA crystals. Initial structures were determined by analysis of $(|F_o| - |F_c|)$, α_c and $(2|F_o| - |F_c|)$, α_c maps and then refined with X-PLOR. Subsequent manual rebuilding and water structure refinement throughout the protein were based on systematic analysis of positive and negative peaks with magnitudes greater than 2.8 σ in $(|F_o| - |F_c|)$, α_c maps. During refinement force field energy terms were removed for atoms involved in short hydrogen bonds. For comparison structures were superimposed based on a set of atoms at and around the S1, S1′ and active sites. Averages and standard deviations in hydrogen bond lengths and other parameters were calculated for multiply determined structures. A subset of 10 structures has been deposited into the RCSB Protein Data Bank (PDB). X-ray diffraction and refinement statistics and PDB accession codes for these structures are provided in Table 4.

References

[1] B.A. Katz, R. Mackman, C. Luong, K. Radika, A. Martelli, P.A. Sprengeler, J. Wang, H. Chan, L. Wong, Structural basis for selectivity of a small molecule, S1-binding, sub-micromolar inhibitor of urokinase type plasminogen activator, Chem. Biol. 7 (2000) 299–312.

[2] E. Zeslawska, A. Schweinitz, A. Karcher, P. Sondermann, S. Sperl, J. Stürzebecher, W. Jacob, Crystals of the urokinase type plasminogen activator variant βc-uPA in complex with small molecule inhibitors open the way towards structure-based drug design, J. Mol. Biol. 301 (2000) 465–475.

[3] H. Yang, J. Henkin, Selective inhibition of urokinase by substituted phenylguanidines: quantitative structure–activity relationship analyses, J. Med. Chem. 33 (1990) 2956–2961.

[4] K. Dano, P.A. Andreasen, J. Grondahl-Hansen, P. Kristensen, L.S. Nielsen, L. Shriver, Plasminogen activators, tissue degradation, and cancer, Adv. Cancer Res. 44 (1985) 139–266.

[5] P.A. Andreasen, L. Kjøller, L. Christensen, M.J. Duffy, The urokinase-type plasminogen activator system in cancer metastasis: A review, Int. J. Cancer 72 (1997) 1–22.

[6] S.A. Rabbani, P. Harakidas, D.J. Davidson, J. Henkin, A.P. Mazar, Prevention of prostate-cancer metastasis in vivo by a novel synthetic inhibitor of urokinase-type plasminogen activator (uPA), Int. J. Cancer 63 (1995) 840–845.

[7] S.A. Rabbani, R.H.M. Xing, Role of urokinase (uPA) and its receptor (uPAR) in invasion and metastasis of hormone-dependent malignancies, Int. J. Oncol. 12 (1998) 911–920.

[8] M.S. Stack, S.M. Ellerbroek, D.A. Fishman, The role of proteolytic enzymes in the pathology of epithelial ovarian carcinoma (review), Int. J. Oncol. 12 (1998) 569–576.

[9] C. Magill, B.A. Katz, R.L. Mackman, Emerging therapeutic targets in oncology: urokinase-type plasminogen activator system, Emerg. Ther. Targets 3 (1999) 109–133.

[10] V. Nienaber, J. Wang, D. Davidson, J. Henkin, Re-engineering of human urokinase provides a system for structure-based drug design at high resolution and reveals a novel structural subsite, J. Biol. Chem. 275 (2000) 7239–7248.

[11] V.L. Nienaber, D. Davidson, R. Edalji, V.L. Giranda, V. Klinghofer, J. Henkin, P. Magdalinos, R. Mantei, S. Merrick, J.M. Everin, R.A. Smith, K. Stewart, K. Walter, J. Wang, M. Wendt, M. Weitzberg, X. Zhao, Structure 8 (2000) 553–563.

[12] S. Sperl, U. Jacob, N. Arroyo de Prada, J. Stürzebacher, O.G. Wilhelm, W. Bode, V. Magdolen, R. Huber, L. Moroder, (4-Aminomethyl)phenylguanidine derivatives as nonpeptidic highly selective inhibitors of human urokinase, Proc. Natl. Acad. Sci. USA 97 (2000) 5113–5118.

[13] P.J. Hajduk, S. Boyd, D. Nettesheim, V. Nienaber, J. Severin, R. Smith, D. Davidson, T. Rockway, S.W. Fesik, Identification of novel inhibitors of urokinase via NMR-based screening, J. Med. Chem. 43 (2000) 3862–3866.

[14] K.T. Sabapathy, M.S. Pepper, F. Kiefer, U. Möhle-Steinlein, F. Tacchini-Cottier, I. Fetka, G. Breier, W. Risau, P. Carmeliet, R. Mon-

tesano, E.F. Wagner, Polyoma middle T-induced vascular tumor formation: the role of the plasminogen activator/plasmin system, J. Cell Biol. 137 (1997) 953–963.

[15] D. Lambda, M. Bauer, R. Huber, S. Fischer, R. Rudolph, U. Kohnert, W. Bode, The 2.3 Å crystal structure of the catalytic domain of recombinant two-chain human tissue-type plasminogen activator, J. Mol. Biol. 258 (1996) 117–135.

[16] E. Verner, B.A. Katz, J. Spencer, D. Allen, J. Hataye, W. Hruzewicz, H. Hui, A. Kolesnikov, Y. Li, C. Luong, A. Martelli, K. Radika, R. Rai, M. She, W. Shrader, P. Sprengeler, S. Trapp, J. Wang, W. Young-Convery, R.L. Mackman, Development of serine protease inhibitors displaying a multi-centered short (<2.3 Å) hydrogen bond binding mode: Inhibitors of urokinase-type plasminogen activator and factor Xa, J. Med. Chem. 44 (2001) 2753–2771.

[17] B.A. Katz, K. Elrod, C. Luong, M. Rice, R.L. Mackman, P.A. Sprengeler, J. Spencer, J. Hataye, J. Janc, J. Link, J. Litvak, R. Rai, K. Rice, S. Sideris, E. Verner, W. Young, A novel serine protease inhibition motif involving a multi-centered short hydrogen bonding network at the active site, J. Mol. Biol. 307 (2001) 1451–1486.

[18] M. Renatus, W. Bode, R. Huber, J. Stürzebecher, M.T. Stubbs, Structural and functional analyses of benzamidine-based inhibitors of factor Xa, tPA, and urokinase, J. Med. Chem. 41 (1998) 5445–5456.

[19] M. Krieger, L.M. Kay, R.M. Stroud, The structure and specific binding of trypsin: A comparison of inhibited derivatives and a model for substrate binding, J. Mol. Biol. 83 (1974) 209–230.

[20] W. Bode, P. Schwager, The refined crystal structure of bovine β-trypsin at 1.8 Å resolution. II. Crystallographic refinement, calcium binding site, benzamidine binding site and active site at pH 7.0, J. Mol. Biol. 98 (1975) 693–717.

[21] W.F. Mangel, P.T. Singer, D.M. Cyr, T.C. Umland, D.L. Toledo, R.M. Stroud, J.W. Pflugrath, R.M. Sweet, Structure of an acyl-enzyme intermediate during catalysis: (guanidinobenzoyl)trypsin, Biochemistry 29 (1990) 8351–8357.

[22] E. Skrzpczak-Jankun, V.E. Carperos, K.G. Ravichandran, A. Tulinsky, M. Westbrook, J.M. Maraganore, Structure of the hirugen and hirulog 1 complexes of α-thrombin, J. Mol. Biol. 221 (1991) 1379–1393.

[23] J.D. Dunitz, The cost of bound water in crystals and biomolecules, Science 264 (1994) 670.

[24] J.E. Ladbury, Just add water! The effect of water on the specificity of protein–ligand binding sites and its potential application to drug design, Chem. Biol. 3 (1996) 973–980.

[25] P.Y.S. Lam, P.K. Jadhav, C.J. Eyermann, C.N. Hodge, Y. Ru, L.T. Bacheler, J.L. Meek, M.J. Otto, M.M. Rayner, Y.N. Wong, C.-H. Chang, P.C. Weber, D.A. Jackson, T.R. Sharpe, S. Erickson-Viitanen, Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors, Science 263 (1994) 380–384.

[26] G.V. De Lucca, P.K. Jadhav, R.E. Waltermire, B.J. Aungst, S. Erickson-Viitanen, P.Y. Lam, De novo design and discovery of cyclic HIV protease inhibitors capable of displacing the active-site structural water molecule, Pharm. Biotechnol. 11 (1998) 257–284.

[27] R.L. Mackman, B.A. Katz, G. Breitenbucher, H. Hui, E. Verner, C. Luong, P. Sprengeler, Exploiting subsite S1 of trypsin-like serine proteases for selectivity: inhibitors of urokinase-type plasminogen activator, J. Med. Chem. (2001), in press.

[28] J.W. Janc, J.M. Clark, R.L. Warne, K.C. Elrod, B.A. Katz, W.R. Moore, A novel approach to serine protease inhibition: kinetic characterization of inhibitors whose potencies and selectivities are dramatically enhanced by zinc(II), Biochemistry 39 (2000) 4792–4800.

[29] P. Kuzmic, S. Sideris, L.M. Cregar, K.C. Elrod, K.D. Rice, J.W. Janc, High-throughput screening of enzyme inhibitors: Automatic determination of tight-binding inhibition constants, Anal. Biochem. 281 (2000) 62–67.