



Structure of an Fab–Protease Complex Reveals a Highly Specific Non-canonical Mechanism of Inhibition

Christopher J. Farady¹, Pascal F. Egea³, Eric L. Schneider²,
Molly R. Darragh¹ and Charles S. Craik^{1,2*}

¹Graduate Group in Biophysics,
University of California,
San Francisco, 600 16th St.
Genentech Hall, San Francisco,
CA 94143-2240, USA

²Department of Pharmaceutical
Chemistry, University of
California, San Francisco,
600 16th St. Genentech Hall,
San Francisco, CA 94143-2280,
USA

³Department of Biochemistry
and Biophysics, University of
California, San Francisco, 600
16th St. Genentech Hall, San
Francisco, CA 94143-2240, USA

Received 28 March 2008;
received in revised form
1 May 2008;
accepted 2 May 2008
Available online
11 May 2008

The vast majority of protein protease inhibitors bind their targets in a substrate-like manner. This is a robust and efficient mechanism of inhibition but, due to the highly conserved architecture of protease active sites, these inhibitors often exhibit promiscuity. Inhibitors that show strict specificity for one protease usually achieve this selectivity by combining substrate-like binding in the active site with exosite binding on the protease surface. The development of new, specific inhibitors can be aided greatly by binding to non-conserved regions of proteases if potency can be maintained. Due to their ability to bind specifically to nearly any antigen, antibodies provide an excellent scaffold for creating inhibitors targeted to a single member of a family of highly homologous enzymes. The 2.2 Å resolution crystal structure of an Fab antibody inhibitor in complex with the serine protease membrane-type serine protease 1 (MT-SP1/matriptase) reveals the molecular basis of its picomolar potency and specificity. The inhibitor has a distinct mechanism of inhibition; it gains potency and specificity through interactions with the protease surface loops, and inhibits by binding in the active site in a catalytically non-competent manner. In contrast to most naturally occurring protease inhibitors, which have diverse structures but converge to a similar inhibitory archetype, antibody inhibitors provide an opportunity to develop divergent mechanisms of inhibition from a single scaffold.

© 2008 Elsevier Ltd. All rights reserved.

Edited by I. Wilson

Keywords: antibody; serine protease; protease inhibitor; substrate specificity; structure

Introduction

Proteolytic activity *in vivo* is carefully regulated by spatial and temporal localization, zymogen activation, autolysis, and through the inhibition of proteases by macromolecular inhibitors. Despite divergent targets and different mechanisms of inhibition, most protease inhibitors bind a critical portion of the inhibitor in the protease active site in a substrate-like manner. Though an effective paradigm for protease inhibition, substrate-like binding in the active site often leads to

inhibitors that can potently inhibit more than one target protease. This promiscuity is evidenced by the fact that 115 annotated human protease inhibitors are capable of regulating the activity of the 612 known human proteases.¹ The few specific protease inhibitors found in biology, such as rhodniin, a thrombin inhibitor from *Rhodnius prolixus*, have gained specificity by combining substrate-like inhibition with exosite binding. Rhodniin has two domains, one of which binds and inhibits the protease via a canonical mechanism, and a second domain evolved to bind to exosite I, resulting in a potent and specific thrombin inhibitor.² Dysregulated proteolytic activity plays a role in many disease states, often caused by a single member of highly homologous protease families. As such, there is a need for selective inhibitors. Traditional attempts to develop small molecule or protein protease inhibitors have had mixed results;^{3,4} difficulties have primarily been due to specificity issues arising from the similarity of

*Corresponding author. Department of Pharmaceutical Chemistry, University of California, San Francisco, 600 16th St. Genentech Hall, San Francisco, CA 94143-2280, USA. E-mail address: craik@cgl.ucsf.edu.

Abbreviations used: uPA, urokinase plasminogen activator; HGFA, hepatocyte growth factor activator.

protease active sites. Therefore, there is a need for more diverse methods for developing specific inhibitors to single members of these highly similar enzymes.

Due to their ability to bind closely related antigens selectively, antibodies provide a particularly attractive scaffold on which to develop specific enzyme inhibitors. Of the antibody-based protease inhibitors that have been reported,^{5–11} most work by interfering with protein–protein interaction sites rather than interacting with the active site of the enzyme. Previously, we used a phage-displayed single-chain antibody library to develop potent and specific inhibitors of membrane type serine protease 1 (MT-SP1/matriptase), but the molecular details of the inhibitory mechanism were unclear.^{12,13} MT-SP1 is a cell-anchored serine protease involved in cell signaling pathways and protease activation, and has been implicated in cancer progression,^{14–16} and is a member of a large family of closely related enzymes, the trypsin-fold serine proteases. Here, we report the crystal structure at 2.2 Å resolution of E2, the most potent previously described antibody inhibitor, in complex with the catalytic domain of MT-SP1. E2 has a distinct mechanism of inhibition; it gains potency and specificity through interactions with the protease surface loops, and binds in the active site in a catalytically non-competent manner.

Results

Characterization of inhibitory Fab

E2 was raised from a phage-displayed fully synthetic human combinatorial scFv library with mod-

ular consensus frameworks and randomized CDR3 as described.¹⁷ We have reported the biochemical characterization of E2,¹³ but the scFv construct proved unsuitable for structural studies, so the Fv was transferred to an Fab scaffold by ligating the variable region to a human Fab constant region.¹⁸ The conversion from an scFv to Fab scaffold had minimal effect on the inhibitory potency of the antibody, which had a K_i of 15 pM against MT-SP1 (data not shown).

E2/MT-SP1 structure

The E2/MT-SP1 structure was determined to 2.2 Å, with two copies of the complex in the asymmetric unit. The antibody caps the protease active site and makes numerous interactions with the surface loops of the protease (Fig. 1). These loops surround the substrate-binding groove of the protease, modulate macromolecular substrate recognition, and are sites of great diversity among the well-conserved family of trypsin-like serine proteases.¹⁹ The Fv heavy chain packs against the 60 s and 90 s loop (standard chymotrypsinogen numbering) while the light chain interacts with the 170 s and 220 s loop on the opposite side of the protease (Fig. 1; Table 1). The E2 light chain makes very few interactions with MT-SP1, burying only 175 Å², or about 15%, of the total buried surface area of the Fab-protease interaction.

The antibody hypervariable loops bind to the protease surface loops either by packing against them or by “grabbing” the loop by stacking it between two CDR loops. This grabbing phenomenon is seen most clearly in the antibody’s interaction with the 90 s loop of MT-SP1, which is stacked between the hypervariable H2 and H3 loops (Fig. 2a). The 90 s loop buries the Phe97 side chain in the hydrophobic core of the

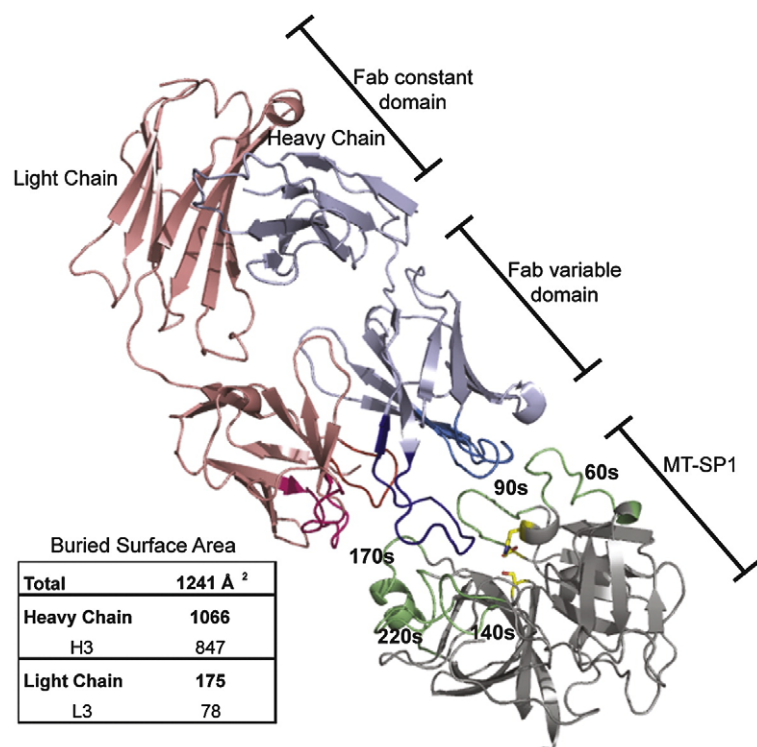


Fig. 1. Structure of the E2/MT-SP1 complex. The Fab (heavy chain, light blue; light chain, light red) caps MT-SP1 (grey) at the active site through interactions with the surface loops (green). H3 of E2 (dark blue) is inserted directly into the active site (catalytic triad indicated in yellow), while the remaining hypervariable loops (L1 and L2, pink; L3, red; H1 and H2, sky blue) interact with the protease surface loops. All figures were prepared using PyMol [<http://www.pymol.sourceforge.net>].

Table 1. E2 / MT-SP1 interactions

E2 residue and CDR Loop	MT-SP1 residue	Type of interaction	Distance (Å)	Buried surface area (Å ²)
L1 SerL30	Pro173	Hydrophobic	3.7	27
L1 TyrL32	Pro173	Hydrophobic	4.2	66
L3 GlyL92	Gln174	Hydrophobic	3.9	18
L3 AsnL93 O ^{δ2}	Gln174 O ^{ε1}	H-bond	3.5	28
L3 TyrL96	Phe97	Hydrophobic	4.0	18
H1 ThrH28 O ^{γ1}	Arg60C NH1	H-bond	3.1	60
H1 SerH30 O ^γ	Arg60C N	H-bond	3.0	34
H1 SerH31 N	Asp60D O ^{δ1}	Polar	3.6	17
H1 AlaH33	Phe97	Hydrophobic	3.8	12
H2 SerH52 O ^γ	Asp96 O ^{δ1}	H-bond	2.7	28
H2 SerH53 N	Asp96 O ^{δ2}	H-bond	3.2	46
H2 SerH56 O ^γ	Asn95 N ^{δ2}	H-bond	2.7	41
H2 TyrH58	Phe97	Hydrophobic	3.3	60
H2 TyrH58 OH	Asn95 N ^{δ2}	H-bond	3.2	
H3 TyrH99	Tyr146	Hydrophobic	3.5	121
H3 TyrH99 OH	Gln221a N ^{ε2}	H-bond	3.0	
H3 ProH100 N	Tyr146 OH	H-bond	3.1	5
H3 GlnH100a N	Tyr146 OH	H-bond	2.9	82
H3 ArgH100b NH2	Ser190 O ^γ	H-bond	3.3	187
H3 ArgH100b NH1	Ser190 O	H-bond	3.2	
H3 ArgH100c NH2	His57 O	Polar	3.6	153
H3 GlyH100d	Phe99	Hydrophobic	3.7	29
H3 ProH100e	Trp215	Hydrophobic	3.5	106
H3 ProH100e	Phe99	Hydrophobic	4.0	
H3 GlnH100f N	Phe97 O	H-bond	3.1	22
H3 AsnH100g	Phe97	Hydrophobic	3.5	85
H3 AsnH100g N ^{δ2}	Gln174 O ^{ε1}	H-bond	2.9	
H3 ValH100h	Gln175	Hydrophobic	3.7	50

Buried surface area determined by PISA.²⁷

antibody hypervariable region, where it is stacked between H2 residue TyrH58 (Kabat numbering) and the AsnH100g side chain of H3. These two crucial interactions alone bury 145 Å² of surface area (Table 1). In the apo MT-SP1 structure,²⁰ Asp96 forms the bottom of the S4 pocket, allowing a positively charged substrate P4 residue. In the antibody structure, however, Asp96 is rotated 180° around the side chain C^β, allowing it to interact with the Fab H2. This rotation allows the Asp96 carboxylic acid moiety to H-bond to the backbone amide of SerH53 (2.8 Å) and the side chain O^γ atoms of SerH53 and SerH56. The 170 s loop is also grabbed by E2. It is bound between the L1 loop, which makes van der Waals contacts with the protease Pro173, the L3 loop, which makes an H-bond between AsnL93 O^{δ2}, and the protease Gln174 side chain, and H3 residues AsnH100g and ValH100h, which make a H-bond to Gln174 O^{ε1} and van der Waals contacts with Gln175, respectively.

The most striking feature of the E2 structure is the extended binding conformation of the CDR H3 loop in the protease active site (Fig. 2b). E2 has a long 18 residue H3 loop, which bridges the active site and buries 847 Å² of surface area, about 68% of the total

buried surface area of the Fab-protease interaction. Though lacking secondary structure, the loop is rigidified by the presence of four prolines. The N-terminus of the loop bends over and packs against the 140 s loop, which regulates the approach of H3 towards the active site. The protease Tyr146 side chain is stacked between the antibody side chains TyrH99 and GlnH100a, and the Tyr146 OH residue forms H-bonds with the backbone nitrogens of ProH100 and GlnH100a, at 3.1 Å and 2.9 Å, respectively (Fig. 3). As the H3 loop approaches the catalytic residues, it inserts the first of two arginines (ArgH100b) in the S1 pocket†. The H3 loop then makes a sharp turn, allowing the side chain of the second arginine (ArgH100c) to extend into the prime side of the active site. The next two residues, P2' and P3', bind in the S2 and S3 pockets, respectively, before bending up and back out of the active site (Fig. 2b). The P2' and P3' residues are GlyH100d and ProH100e, respectively, and though they bind in the substrate pockets, they do not bind in a substrate like manner, as the chloromethyl ketone inhibitor D-FPR-cmk binds in the fVIIa active site.²¹ They are oriented slightly above D-FPR-cmk, and do not extend side chains into the binding pockets (Fig. 4a). Earlier, we showed that E2 could be processed between the two arginine residues when incubated with MT-SP1 at pH 6.0 for an extended period of time. This is a hallmark of standard mechanism serine protease inhibitors, and suggested that the inhibitor bound in some sort of substrate-like manner. In light of the structure, it is clear that E2 is in fact an extremely poor substrate due to the 7.5 Å distance between the catalytic Ser195 O^γ and the carbonyl carbon atom of the scissile bond (defined as the peptide bond in closest proximity to Ser195) and can be cleaved only when the complex is compromised at low pH.²²

The ArgH100b side chain of E2 binds in the S1 pocket of MT-SP1 in an unexpected manner (Fig. 2c). The side chain guanidino group makes a 3.1 Å hydrogen bond to Ser190 O^γ and a 2.9 Å water-mediated hydrogen bond to Asp189 O^{δ1} at the base of the S1 pocket. This is similar to the binding mode of the terminal amine of Lys15 of basic pancreatic trypsin inhibitor (BPTI), but different from that of benzamidine,²⁰ an inhibitor that mimics arginine-binding in the S1 pocket by making a salt-bridge interaction with Asp189 (Fig. 2c). P1 arginine binding in the S1 pocket is thus sub-optimal. The side chain — and by extension the scissile bond — are restricted from binding more deeply in the active site due to the geometric constraints placed upon them by the orientation of the H3 loop.

E2 binds preferentially to the active form of MT-SP1

The significant interactions E2 makes with the protease active site have another benefit. E2 interacts

† The scFv residues corresponding to ArgH100b and ArgH100c were previously referred to as Arg131 and Arg132, respectively.¹³

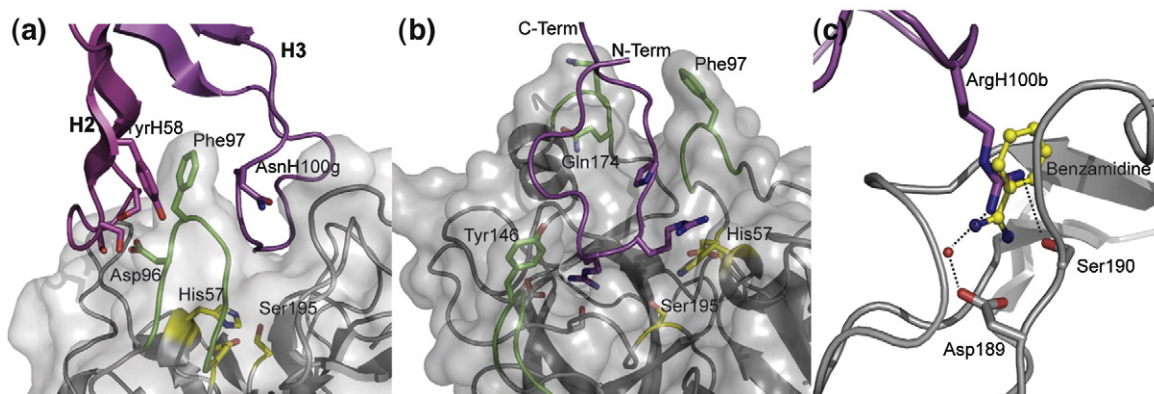


Fig. 2. Critical determinants of E2 (magenta) inhibition of MT-SP1 (gray). (a) The 90 s (green) loop of MT-SP1 is bound between the H2 and H3 loops of E2. Phe97 is critical to binding,¹³ and is stacked between TyrH58 of the H2 loop and AsnH100g of the H3 loop, while Asp96 hydrogen bonds to SerH52 and SerH53 of E2. (b) The H3 loop of E2 bridges the MT-SP1 active site, and makes contacts with the 90 s, 140 s, and 170 s loop of the protease. ArgH100b is bound in the S1 pocket, ArgH100c is bound in the S1' pocket, while ProH100e is bound in the S3 pocket. (c) ArgH100b is bound sub-optimally in the S1 pocket. Benzamidine (ball and stick, PDB code 1EAX) adapts a substrate-like binding orientation, making a salt-bridge with Asp189 at the bottom of the S1 pocket of MT-SP1 (distance of 3.1 Å). By comparison, ArgH100b NH₂ of E2 is at a distance of 4.4 Å from Asp189 O^{δ1} of MT-SP1 and alternatively makes a water-mediated hydrogen bond of 2.9 Å to Asp189.

with the 220 s loop, 140 s loop, and in the S1 pocket. These residues are all part of the activation domain of trypsin-fold serine proteases, which become ordered upon zymogen activation.²³ Surface plasmon resonance shows that E2 does not bind to the zymogen form of the enzyme (Fig. 6). At 500 nM, the zymogen does not bind to immobilized E2, while the active protease gives a robust binding signal. A K_D for the zymogen could not be determined as the zymogen could not be concentrated enough to see binding. E2 can thus preferentially bind to the active form of the protease, further showing that antibody inhibitors can be specific enough to target distinct conformations of a single enzyme.

Discussion

The structure of the antibody-based inhibitor E2 in complex with the serine protease MT-SP1 reveals the basis of its potency and specificity. In contrast to the vast majority of naturally occurring protein protease inhibitors, which bind primarily in the protease ac-

tive site, this antibody-based inhibitor binds to the surface loops flanking the protease active site. The sequences and conformations of these loops are highly degenerate among families of proteases, and thus they are sites of natural diversity. This is analogous to naturally occurring inhibitors that show a high degree of specificity, such as anti-coagulation inhibitors from blood-meal parasites, which gain selectivity by binding to exosites on the surface of the protease. Although relatively few protease exosites have been discovered, the diversity of protease surface loops makes them attractive areas to target to build specificity into an inhibitor.

The potency of E2 is striking, and our structure reveals why; the inhibitor buries a large surface area, and the antibody scaffold orients the inhibitor H3 loop in a non-substrate-like conformation in the active site to inhibit the protease. The canonical serine protease inhibitor BPTI inhibits trypsin with a K_i of 0.6 pM;²⁴ a potency that arises from exquisite shape and charge complementarity between enzyme and inhibitor.²⁵ In contrast, this inhibitor binds in a sub-optimal manner in the active site. E2 does not make

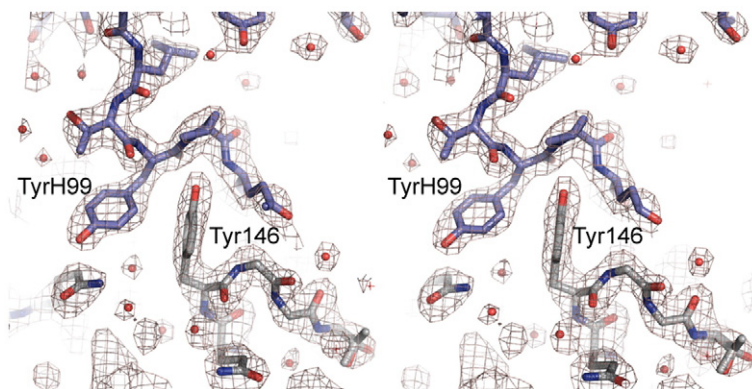


Fig. 3. Stereoview of $2F_o - F_c$ map at 2σ of the E2 H3 (blue) interacting with Tyr146 of MT-SP1 (gray). The protease Tyr146 side chain is stacked between the side chains of the antibody TyrH99 and GlnH100a, and the Tyr146 OH residue forms H-bonds with the backbone nitrogens of ProH100 and GlnH100a, at 3.1 Å and 2.9 Å, respectively.

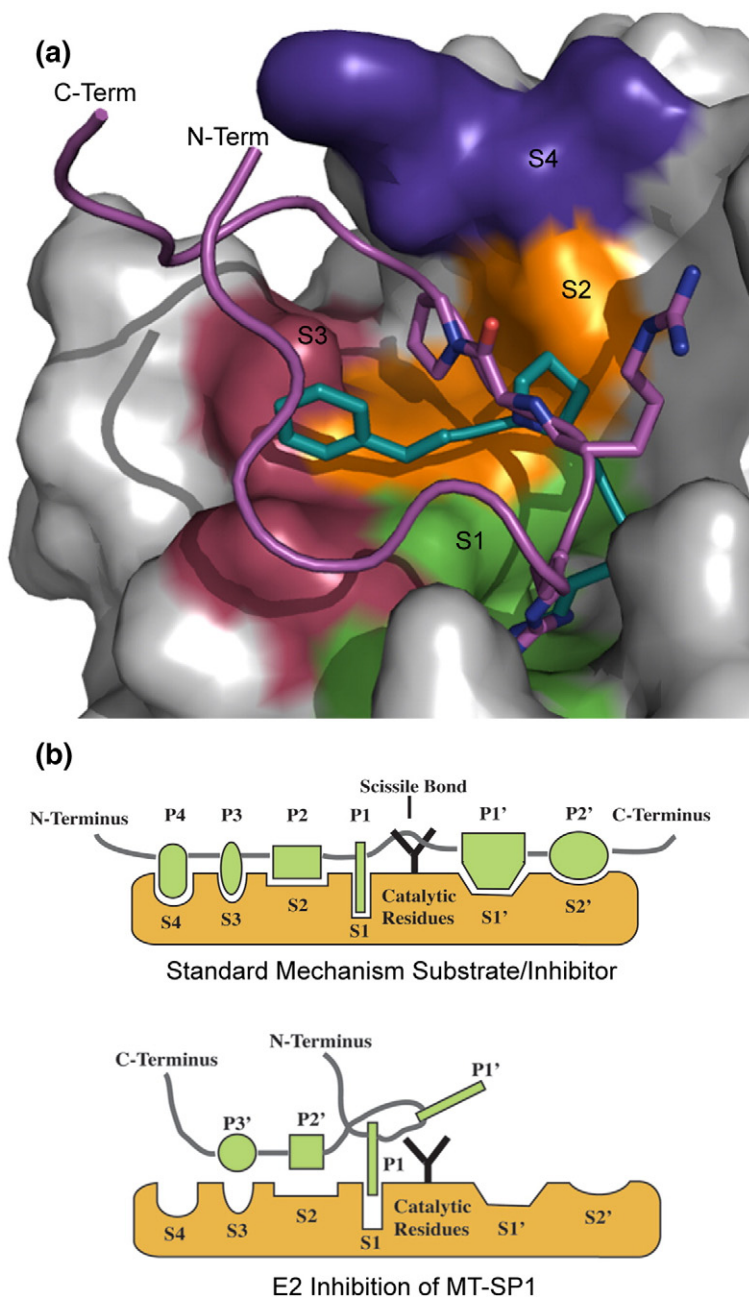


Fig. 4. E2 adopts a unique conformation in the MT-SP1 active site. (a) When the chloromethyl ketone inhibitor D-FPR-cmk (teal, bound to the serine protease fVIIa, PDB code 2FIR) is overlaid in the MT-SP1 active site, the amino acid side chains are buried in the substrate-binding sites, S1 (green), S2 (orange), S3 (light magenta) and S4 (purple). By contrast, E2 inserts ArgH100b into the S1 pocket, but then bends above the catalytic residues and binds the P2' GlyH100d and P3' ProH100e in a reverse orientation in the S2 and S3 residues, respectively. This unexpected conformation (b) allows the H3 loop to conform to the unique shape of the MT-SP1 binding cleft and make numerous beneficial interactions in the active site, but prevents the loop from being readily cleaved by the enzyme.

an energetically favored salt-bridge with Asp189 in the S1 pocket of the protease and though it makes significant contacts with other protease subsites, they are not optimal. The P2' and P3' residues are glycine and proline, respectively, and make only modest backbone interactions with the protease (Fig. 4a). Despite this, E2 has a K_i in the low picomolar range. Active site binding is responsible for some of this binding energy — the ArgH100b binding in the S1 pocket provides 5 kcal/mol binding energy for E2¹³ — but the many interactions with the surface loops are critical as well. The 1241 Å² of surface area that E2 buries on MT-SP1 is larger than the typical antibody/protein antigen interaction, which averages about 875 Å².^{26,27} This interaction area is also large for a protease inhibitor; TIMP inhibitors bury a similar surface area on matrix metalloproteases,²⁸

while stefin inhibitors of cysteine proteases²⁹ and canonical serine protease inhibitors³⁰ have interfaces of about 900 Å². Much effort has gone into using phage display and rational design to impart selectivity to naturally occurring protease inhibitors.^{31,32} This is difficult, since standard mechanism inhibitors have been evolved for maximum inhibitory efficiency. Some of this potency has to be sacrificed for these inhibitors to gain specificity; however, with protein scaffolds engineered to be protease inhibitors, potency and specificity are necessarily linked. If a protein can be engineered to bind to a large surface area and interfere with the catalytic machinery of an enzyme, it will most likely be specific.

A number of macromolecular MT-SP1 inhibitors have been described, all of which bind in the active site in a substrate like manner.^{20,33–35} Using the

structural and kinetic data presented here, we can describe a broader "inhibitope" for MT-SP1, a set of crucial contacts and interactions that, when linked together on the correct scaffold, lead to potent and specific inhibition. For E2, maximal MT-SP1 inhibition depends on interactions with the 90 s loop, 140 s loop, an arginine side chain in the S1 pocket, and non-substrate-like binding in the protease active site. Inhibitors of the closely related protease urokinase plasminogen activator (uPA) have a different inhibitope. Cyclic peptide uPA inhibitors have a strict requirement for an arginine P1,³⁶ but anti-uPA mAbs⁹ make many more significant interactions with the 37 s and 60 s loops of uPA, on the prime side of the protease active site, suggesting these loops provide possible anchor points for inhibition.

Recently, the structures of two inhibitory antibodies of the serine protease hepatocyte growth factor

activator (HGFA) were reported.³⁷ One inhibitor, Ab75, appeared to be an allosteric inhibitor, while the most potent inhibitor, Ab58, had some similarities to E2 inhibition of MT-SP1. Ab58 buries the Phe97 residue of HGFA between two hypervariable loops, and uses the 90 s loop as an anchor point, but binds to and inhibits the protease very differently. While E2 caps the MT-SP1 active site, Ab58 approaches the active site at an angle (Fig. 5a). The H1 and H2 loops of Ab58 are located in the substrate-binding groove, but the inhibitors do not approach the S1 pocket or the catalytic residues (Fig. 5b). A similar mechanism of inhibition would probably not be possible for MT-SP1, which has a deeper and more occluded active site cleft than HGFA. Thus, using a similar antibody scaffold, it is possible to develop specific protease inhibitors with completely novel mechanisms of inhibition. Whether aspects of these

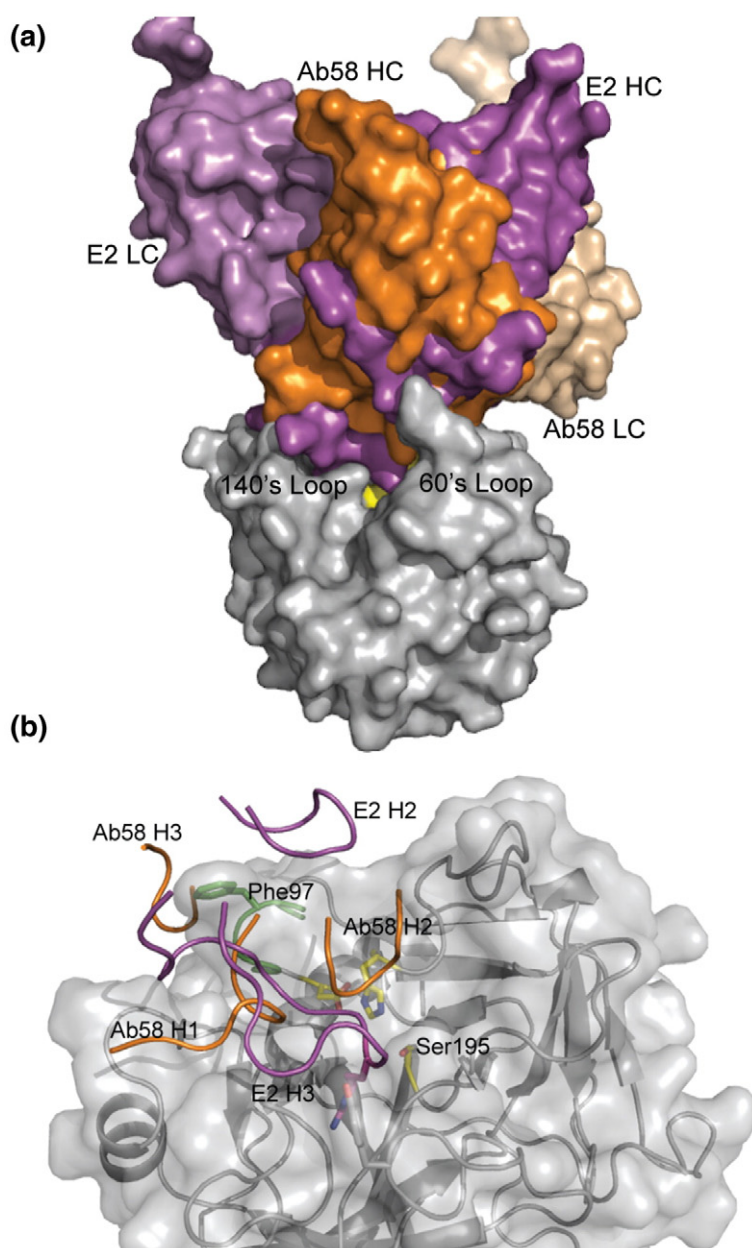


Fig. 5. Comparison of E2 with HGFA antibody inhibitor Fab58 (PDB code 2R0K). The HGFA structure was aligned with MT-SP1. (a) Fab58 (heavy chain, orange, light chain, light brown) approaches the protease (gray) active site from a different angle than E2 (heavy chain, magenta, light chain, light magenta). The Fab58 light chain also makes very few interactions with the protease. (b) The heavy chain hypervariable loops of both inhibitors bind in the substrate-binding cleft of the protease. Fab58 (orange) binds the H2 and H1 loops in the S2 and S3 substrate pockets, respectively, while the E2 (magenta) H3 loop also binds in the S1 pocket. Both inhibitors "grab" the protease 90 s loop, Fab58 with H1 and H3, E2 with H2 and H3.

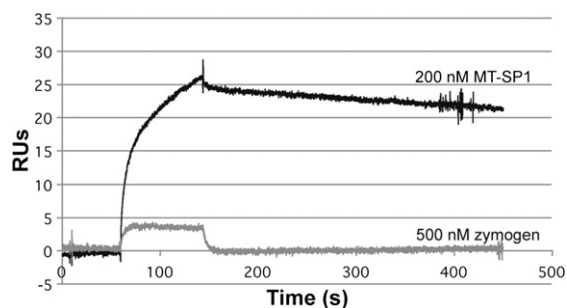


Fig. 6. SPR binding curves of MT-SP 1 (black) and the catalytically inactive mutant R15A (gray) show a lack of binding by the zymogen to E2. Analysis of binding curves for 100 nM, 200 nM and 500 nM MT-SP1 R15A did not yield a reliable fit using the BIAcore Evaluation Software, indicating that changes in RU are due to transient association of the zymogen with the chip surface or general buffer effects.

inhibitopes and mechanisms can guide the rational design of new inhibitors is yet to be seen but, given an appropriate scaffold, an antigen with a three-dimensional epitope, and an inhibitor library with sufficient diversity, it should be possible to develop specific inhibitors with novel mechanisms of action.

The structure presented here helps define the mechanism of inhibition of a potent and specific antibody inhibitor of a serine protease. The mechanism was unexpected. It would be difficult to predict which protease residues or loops would be amenable to binding, or that the H3 loop would be able to adopt a non-substrate-like conformation in the active site that would allow for potent inhibition of MT-SP1. By utilizing robust binding scaffolds and combinatorial selection techniques to identify unique inhibitopes, we have developed selective inhibitors with mechanisms that are specific for MT-SP1. This provides the opportunity to develop potent and selective inhibitors against individual enzymes, and precisely monitor and modulate a wide array of proteolytic processes.

Materials and Methods

Protein expression, purification, and mutagenesis

MT-SP1 and its mutants were expressed in *Escherichia coli* and purified as described.^{13,38} The zymogen was created by an R15A substitution, which prevented protease activation. It elutes from a gel-filtration column at the same time as the active protease, but shows no enzymatic activity. For crystallization purposes, the surface Cys122 residue was mutated to serine using the Stratagene Quickchange kit (Stratagene, La Jolla, CA). The E2 scFv was converted to an Fab by using overlap extension PCR³⁹ between the scFv and the humanized constant region from an Fab phage displayed library. The overlapped region corresponded to residues 104–113 in the heavy chain and 98–107 in the light chain. It was verified by DNA sequencing, expressed in *E. coli*, and purified as described.¹²

Steady-state kinetics

Kinetics were carried out as described.¹³ Briefly, reactions were carried out in 50 mM Tris (pH 8.8), 50 mM NaCl, 0.01% (v/v) Tween-20 in 96-well, medium-binding, flat-bottomed plates (Corning), and cleavage of the substrate (Spectrazyme-tPA (hexahydrotyrosyl-Gly-Arg-pNA), American Diagnostica, Greenwich, CT) was monitored in a UVmax Microplate Reader (Molecular Devices Corporation, Palo Alto, CA.). K_i was measured using the tight-binding inhibition equations of Williams and Morrison.⁴⁰ All graphs and equations were fit using Kaleidagraph 3.6 (Synergy Software, Reading, PA).

Surface plasmon resonance

The association and dissociation curves for MT-SP1 and the inactive zymogen MT-SP1 R15A were obtained by surface plasmon resonance using a BIAcore Biosensor T100 (GE Healthcare). The E2 Fab (ligand), in 25 mM sodium acetate buffer (pH 5.0), was covalently immobilized onto a CM5 chip according to the manufacturer's protocol with a final immobilization level of ~120 RU. The reference channel was treated using the same chemistry as the ligand-coupled surface. Enzymes (analytes) were washed in HBS-EP buffer (10 mM Hepes (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% Tween 20) and injected in varying concentrations (0.4–400 nM for MT-SP1, 100 nM–20 μ M for MT-SP1 R15A) across the chip surface at 25 μ l/min. Surface regenerations were performed with 100 mM glycine (pH 2.2), allowing a complete return to baseline. The sensorgram of the reference surface was subtracted from the ligand-conjugated surface for each injection. Multiple injections of HBS-EP were also used to remove noise from the data.

Crystallization and data collection

E2 was incubated with MT-SP1 in 1:1 molar ratio, the complexes were purified by gel-filtration in a buffer containing 50 mM Tris (pH 8.0), 100 mM NaCl, 5% (v/v) glycerol, and then concentrated to 15–20 mg/ml. High-throughput crystallization screening was performed using a nanoliter-scale Mosquito robot (TTP Labtech) in hanging drops by vapor diffusion. The E2/MT-SP1 complex crystallized in 16% (w/v) PEG 5000 MME, 0.21 M AmSO_4 and 0.1 M Tris (pH 8.0). Crystals belonging to the orthorhombic space group $P2_12_12_1$ ($a=48.63$ Å, $b=163.28$ Å and $c=201.16$ Å) grew in two days, and were cryoprotected in the mother liquor supplemented with 20% (v/v) ethylene glycol. Diffraction data were collected at beamline 8.3.1 at the Advanced Light Source at LBNL (see Table 2). E2/MT-SP1 data were reduced and scaled using Elves.⁴¹

Structure determination and refinement

The structure was solved by molecular replacement using Phaser⁴² in CCP4⁴³, first searching for MT-SP1 (using 1EAX as search model), then searching for the Fab fragment with its H3 loop truncated (using 2HFF as search model) for the MT-SP1 complex. Molecular replacement was followed by automatic building in ARP/wARP⁴⁴ and manual building cycles. Restrained refinement cycles were done using Phenix⁴⁵ and TLS refinement was applied in the last stages of the refinement. Fab residues 127–137 and 186–192 of heavy chain D and 127–131 of heavy chain F had no density, and were left out of the refinement model.

Table 2. Data collection and refinement statistics

	E2 / MT-SP1
A. Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	48.3, 163.1, 201.2
α, β, γ (°)	90, 90, 90
Resolution (Å)	127–2.18 (2.3–2.17)
No. reflections	774,451
No. reflections (unique)	74,977
R_{merge}	0.091 (0.569)
$I / \sigma I$	7.5 (1.7)
Completeness (%)	89 (77)
Redundancy	3.0 (2.5)
B. Refinement	
Resolution (Å)	85–2.18 (2.23–2.17)
No. reflections	74,828 (4314)
$R_{\text{work}} / R_{\text{free}}$	22.3 / 26.7
No. atoms	
Protein	10,226
Ligand/ion	62 (2 sulfate, 13 ethylene glycol)
Water	824
B-factors	
Protein (Å ²)	41.8
Ligand/ion (Å ²)	51.2
Water (Å ²)	38.8
r.m.s. deviation from ideal	
Bond lengths (Å)	0.02
Bond angles (°)	1.45
Ramachandran plot	
Favored regions (%)	96.5
Allowed regions (%)	99.9

Values in parentheses are for the highest-resolution shell.
The test set was 7.5% of total reflections.

These regions are often disordered in Fab structures, and make no interactions with the protease. Furthermore, light chain E residues LysH145, LysH190, AsnH210, light chain C residues LysH190 and PheH209, heavy chain D residues LysH209, LysH210, and LysH214, and heavy chain F residue ThrH191 and LysH214 had no side chain density and were truncated at C^β.

For both structures, analysis of the thermal motion parameters with TLSMD[‡] revealed anisotropic motions between the constant and variable regions for the complexes present in the asymmetric unit. The proteases were treated as single groups and antibody chains were treated as two separate groups with the boundary defined at the hinge between the constant and variable regions. The quality of the final structure was assessed using Molprobability.⁴⁶ Buried surface area calculations were performed using PISA.²⁷

Protein Data Bank accession code

Coordinates and structure factors have been deposited with the Protein Data Bank (code 3BN9).

Acknowledgements

The authors thank Professor Robert Fletterick for critical reading of the manuscript, Dr James Holton

and Dr George Meigs for assistance at ALS beamline 8.3.1, and Jeremy Wilbur and Dr Peter Hwang for technical assistance with the Biacore. This work was funded by NIH grants CA072006 and GM082250 (to C.S.C), Department of Defense grant BC043431 (to C.J.F), and an NSF graduate research fellowship (to M.R.D.).

References

1. Rawlings, N. D., Morton, F. R. & Barrett, A. J. (2006). MEROPS: the peptidase database. *Nucleic Acids Res.* **34**, D270–D272.
2. Roussel, A., Mathieu, M., Dobbs, A., Luu, B., Cambillau, C. & Kellenberger, C. (2001). Complexation of two proteic insect inhibitors to the active site of chymotrypsin suggests decoupled roles for binding and selectivity. *J. Biol. Chem.* **276**, 38893–38898.
3. Coussens, L. M., Fingleton, B. & Matrisian, L. M. (2002). Matrix metalloproteinase inhibitors and cancer: trials and tribulations. *Science*, **295**, 2387–2392.
4. Turk, B. (2006). Targeting proteases: successes, failures and future prospects. *Nature Rev. Drug Discov.* **5**, 785–799.
5. Rezacova, P., Lescar, J., Brynda, J., Fabry, M., Horejsi, M., Sedlacek, J. & Bentley, G. A. (2001). Structural basis of HIV-1 and HIV-2 protease inhibition by a monoclonal antibody. *Structure*, **9**, 887–895.
6. Puchi, M., Quinones, K., Concha, C., Iribarren, C., Bustos, P., Morin, V. *et al.* (2006). Microinjection of an antibody against the cysteine-protease involved in male chromatin remodeling blocks the development of sea urchin embryos at the initial cell cycle. *J. Cell Biochem.* **98**, 335–342.
7. Fukuoka, Y. & Schwartz, L. B. (2006). The B12 anti-trypsin monoclonal antibody disrupts the tetrameric structure of heparin-stabilized beta-trypsin to form monomers that are inactive at neutral pH and active at acidic pH. *J. Immunol.* **176**, 3165–3172.
8. Matias-Roman, S., Galvez, B. G., Genis, L., Yanez-Mo, M., de la Rosa, G., Sanchez-Mateos, P. *et al.* (2005). Membrane type 1-matrix metalloproteinase is involved in migration of human monocytes and is regulated through their interaction with fibronectin or endothelium. *Blood*, **105**, 3956–3964.
9. Petersen, H. H., Hansen, M., Schousboe, S. L. & Andreasen, P. A. (2001). Localization of epitopes for monoclonal antibodies to urokinase-type plasminogen activator: relationship between epitope localization and effects of antibodies on molecular interactions of the enzyme. *Eur. J. Biochem.* **268**, 4430–4439.
10. Xuan, J. A., Schneider, D., Toy, P., Lin, R., Newton, A., Zhu, Y. *et al.* (2006). Antibodies neutralizing hepsin protease activity do not impact cell growth but inhibit invasion of prostate and ovarian tumor cells in culture. *Cancer Res.* **66**, 3611–3619.
11. Obermajer, N., Premzl, A., Zavasnik Bergant, T., Turk, B. & Kos, J. (2006). Carboxypeptidase cathepsin X mediates beta2-integrin-dependent adhesion of differentiated U-937 cells. *Expt. Cell Res.* **312**, 2515–2527.
12. Sun, J., Pons, J. & Craik, C. S. (2003). Potent and selective inhibition of membrane-type serine protease 1 by human single-chain antibodies. *Biochemistry*, **42**, 892–900.
13. Farady, C. J., Sun, J., Darragh, M. R., Miller, S. M. & Craik, C. S. (2007). The mechanism of inhibition of antibody-based inhibitors of membrane-type serine protease 1 (MT-SP1). *J. Mol. Biol.* **369**, 1041–1051.

‡ <http://skuld.bmsc.washington.edu/~tlsmd/>

14. List, K., Szabo, R., Molinolo, A., Sriuranpong, V., Redeye, V., Murdock, T. *et al.* (2005). Deregulated matriptase causes ras-independent multistage carcinogenesis and promotes ras-mediated malignant transformation. *Genes Dev.* **19**, 1934–1950.
15. Bhatt, A. S., Welm, A., Farady, C. J., Vasquez, M., Wilson, K. & Craik, C. S. (2007). Coordinate expression and functional profiling identify an extracellular proteolytic signaling pathway. *Proc. Natl Acad. Sci. USA*, **104**, 5771–5776.
16. Welm, A. L., Sneddon, J. B., Taylor, C., Nuyten, D. S., van de Vijver, M. J., Hasegawa, B. H. & Bishop, J. M. (2007). The macrophage-stimulating protein pathway promotes metastasis in a mouse model for breast cancer and predicts poor prognosis in humans. *Proc. Natl Acad. Sci. USA*, **104**, 7570–7575.
17. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G. *et al.* (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* **296**, 57–86.
18. de Haard, H. J., van Neer, N., Reurs, A., Hufton, S. E., Roovers, R. C., Henderikx, P. *et al.* (1999). A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *J. Biol. Chem.* **274**, 18218–18230.
19. Perona, J. J. & Craik, C. S. (1997). Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* **272**, 29987–29990.
20. Friedrich, R., Fuentes-Prior, P., Ong, E., Coombs, G., Hunter, M., Oehler, R. *et al.* (2002). Catalytic domain structures of MT-SP1/matriptase, a matrix-degrading transmembrane serine proteinase. *J. Biol. Chem.* **277**, 2160–2168.
21. Bajaj, S. P., Schmidt, A. E., Agah, S., Bajaj, M. S. & Padmanabhan, K. (2006). High resolution structures of p-aminobenzamidine- and benzamidine-VIIa/soluble tissue factor: unpredicted conformation of the 192–193 peptide bond and mapping of Ca^{2+} , Mg^{2+} , Na^+ , and Zn^{2+} sites in factor VIIa. *J. Biol. Chem.* **281**, 24873–24888.
22. Luthy, J. A., Praissman, M., Finkenstadt, W. R. & Laskowski, M., Jr (1973). Detailed mechanism of interaction of bovine -trypsin with soybean trypsin inhibitor (Kunitz). I. Stopped flow measurements. *J. Biol. Chem.* **248**, 1760–1771.
23. Bode, W., Schwager, P. & Huber, R. (1978). The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. The refined crystal structures of the bovine trypsinogen-pancreatic trypsin inhibitor complex and of its ternary complex with Ile-Val at 1.9 Å resolution. *J. Mol. Biol.* **118**, 99–112.
24. Gebhard, W., Tschesche, T. & Fritz, H. (1986). Biochemistry of aprotinin and aprotinin-like inhibitors. In *Proteinase Inhibitors* (Barrett, J.A. & Salvesen, G., eds), pp. 55–152, Elsevier, New York.
25. Lawrence, M. C. & Colman, P. M. (1993). Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946–950.
26. Huang, M., Syed, R., Stura, E. A., Stone, M. J., Stefanko, R. S., Ruf, W. *et al.* (1998). The mechanism of an inhibitory antibody on TF-initiated blood coagulation revealed by the crystal structures of human tissue factor, Fab 5G9 and TF.G9 complex. *J. Mol. Biol.* **275**, 873–894.
27. Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797.
28. Maskos, K., Lang, R., Tschesche, H. & Bode, W. (2007). Flexibility and variability of TIMP binding: X-ray structure of the complex between collagenase-3/MMP-13 and TIMP-2. *J. Mol. Biol.* **366**, 1222–1231.
29. Jenko, S., Dolenc, I., Guncar, G., Dobersek, A., Podobnik, M. & Turk, D. (2003). Crystal structure of Stefin A in complex with cathepsin H: N-terminal residues of inhibitors can adapt to the active sites of endo- and exopeptidases. *J. Mol. Biol.* **326**, 875–885.
30. Scheidig, A. J., Hynes, T. R., Pelletier, L. A., Wells, J. A. & Kossiakoff, A. A. (1997). Crystal structures of bovine chymotrypsin and trypsin complexed to the inhibitor domain of Alzheimer's amyloid beta-protein precursor (APPI) and basic pancreatic trypsin inhibitor (BPTI): engineering of inhibitors with altered specificities. *Protein Sci.* **6**, 1806–1824.
31. Dennis, M. S. & Lazarus, R. A. (1994). Kunitz domain inhibitors of tissue factor-factor VIIa. II. Potent and specific inhibitors by competitive phage selection. *J. Biol. Chem.* **269**, 22137–22144.
32. Komiyama, T., VanderLugt, B., Fugere, M., Day, R., Kaufman, R. J. & Fuller, R. S. (2003). Optimization of protease-inhibitor interactions by randomizing adventitious contacts. *Proc. Natl Acad. Sci. USA*, **100**, 8205–8210.
33. Kirchhofer, D., Peek, M., Li, W., Stamos, J., Eigenbrot, C., Kadkhodayan, S. *et al.* (2003). Tissue expression, protease specificity, and Kunitz domain functions of hepatocyte growth factor activator inhibitor-1B (HAI-1B), a new splice variant of HAI-1. *J. Biol. Chem.* **278**, 36341–36349.
34. Stoop, A. A. & Craik, C. S. (2003). Engineering of a macromolecular scaffold to develop specific protease inhibitors. *Nature Biotechnol.* **21**, 1063–1068.
35. Li, P., Jiang, S., Lee, S. L., Lin, C. Y., Johnson, M. D., Dickson, R. B. *et al.* (2007). Design and synthesis of novel and potent inhibitors of the type II transmembrane serine protease, matriptase, based upon the sunflower trypsin inhibitor-1. *J. Med. Chem.* **50**, 5976–5983.
36. Zhao, G., Yuan, C., Wind, T., Huang, Z., Andreasen, P. A. & Huang, M. (2007). Structural basis of specificity of a peptidyl urokinase inhibitor, upain-1. *J. Struct. Biol.* **160**, 1–10.
37. Wu, Y., Eigenbrot, C., Liang, W.-C., Stawicki, S., Shia, S., Fan, B. *et al.* (2007). Structural insight into distinct mechanisms of protease inhibition by antibodies. *Proc. Natl Acad. Sci. USA*, **104**, 19784–19789.
38. Takeuchi, T., Shuman, M. A. & Craik, C. S. (1999). Reverse biochemistry: use of macromolecular protease inhibitors to dissect complex biological processes and identify a membrane-type serine protease in epithelial cancer and normal tissue. *Proc. Natl Acad. Sci. USA*, **96**, 11054–11061.
39. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51–59.
40. Williams, J. W. & Morrison, J. F. (1979). The kinetics of reversible tight-binding inhibition. *Methods Enzymol.* **63**, 437–467.
41. Holton, J. & Alber, T. (2004). Automated protein crystal structure determination using ELVES. *Proc. Natl Acad. Sci. USA*, **101**, 1537–1542.
42. Read, R. J. (2001). Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr. D*, **57**, 1373–1382.
43. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.

-
44. Evrard, G. X., Langer, G. G., Perrakis, A. & Lamzin, V. S. (2007). Assessment of automatic ligand building in ARP/wARP. *Acta Crystallogr. D*, **63**, 108–117.
 45. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. *et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D*, **58**, 1948–1954.
 46. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G. *et al.* (2003). Structure validation by α geometry: ϕ , ψ and χ deviation. *Proteins: Struct Funct. Genet.* **50**, 437–450.