Published in final edited form as:

J Mol Biol. 2012 January 27; 415(4): 699–715. doi:10.1016/j.jmb.2011.11.036.

A Reverse Binding Motif that Contributes to Specific Protease Inhibition by Antibodies

Eric L. Schneider^{1,*}, Melody S. Lee^{2,*}, Aida Baharuddin¹, David H. Goetz¹, Christopher J. Farady³, Mick Ward⁴, Cheng-I Wang^{1,5}, and Charles S. Craik^{1,#}

¹Department of Pharmaceutical Chemistry, University of California, San Francisco, 600 16th St. Genentech Hall, San Francisco, CA. 94143-2280, USA.

²Graduate Group in Chemistry and Chemical Biology, University of California, San Francisco, 600 16th St. Genentech Hall, San Francisco, CA. 94143-2240, USA.

³Graduate Group in Biophysics, University of California, San Francisco, 600 16th St. Genentech Hall, San Francisco, CA. 94143-2240, USA.

⁴Genencor, Inc.

Abstract

The type II transmembrane serine protease (TTSP) family consists of eighteen closely related serine proteases that are implicated in multiple functions. To identify selective, inhibitory antibodies against one particular TTSP, matriptase (MT-SP1), a phage display library with a natural repertoire of Fabs from human naïve B cells was created. Fab A11 was identified with a 720 picomolar inhibition constant and high specificity for matriptase over other trypsin-fold serine proteases. A *Trichoderma reesei* system expressed A11 with ~200 mg/L yield. The crystal structure of A11 in complex with matriptase has been determined and compared to the crystal structure of another antibody inhibitor (S4) in complex with matriptase. Previously discovered from a synthetic scFv library, S4 is also a highly selective and potent matriptase inhibitor. The crystal structures of the A11/matriptase and S4/matriptase complexes were solved to 2.1 Å and 1.5 Å respectively. Although these antibodies, discovered from separate libraries, interact differently with the protease surface loops for their specificity, the structures reveal a similar novel mechanism of protease inhibition. Through the insertion of the H3 variable loop in a reverse orientation at the substrate-binding pocket, these antibodies bury a large surface area for potent inhibition and avoid proteolytic inactivation. This discovery highlights the critical role the antibody scaffold plays in positioning loops to bind and inhibit protease function in a highly selective manner. Additionally, Fab A11 is a fully human antibody that specifically inhibits matriptase over other closely related proteases, suggesting this approach could be useful for clinical applications.

© 2011 Elsevier Ltd. All rights reserved.

Corresponding Author.

⁵Current address: Singapore Immunology Network, BMSI, A-STAR, Singapore 138648

*These authors contributed equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Accession numbers

Coordinates and structure factors have been deposited in the Protein Data Bank with accession numbers 3SO3 (A11/matriptase complex) and 3NPS (S4/matriptase complex).

Keywords

antibody; specificity; matriptase; structure; protease inhibitor

Introduction

The use of antibodies as therapeutics has increased significantly over the past decade.¹ The high specificity and tight binding characteristics commonly attributed to antibodies give them enormous potential as therapeutics. This specificity allows for precise targeting of protein functions, which helps minimize side effects resulting from off-target binding. Currently, therapeutic antibodies in clinical use function through three modes of action: (1) as inducers of immune system cytotoxicity, (2) as carriers of a specific cytotoxic agent, and (3) as inhibitors of the target protein function.² To date, the majority of therapeutic antibodies have fallen into this last group, acting as antagonists of proteins in disease-related signaling pathways such as Avastin for vascular endothelial growth factor, Erbitux for epidermal growth factor receptor and Humira for tumor necrosis factor.¹ By combining selectivity with a large binding footprint, antibodies have proven to be ideal for creating sufficient steric hindrance by direct or indirect means to block ligand/receptor interactions, thereby inhibiting downstream signaling cascades and functions involved in disease progression.

Many diseases are dependent upon dysregulated enzyme function for their pathogenesis. Proteases in particular have been implicated in a number of functions essential for cancer progression from extracellular matrix remodeling and release of cytokines to loss of apoptotic response.³⁻⁵ For example, certain members of the type II transmembrane serine protease (TTSP) family are thought to play a role in cancer progression.⁵ The localization of these proteases at the extracellular surface of the plasma membrane suggests their involvement in the regulation of cellular signaling events. One protein from this family that is of specific interest is Membrane Type-Serine Protease 1 (MT-SP1) or matriptase.⁶ Matriptase is over-expressed on the surface of epithelial cells involved in breast, colon and prostate cancers. This protease is involved in the activation of other proteases, growth factors and receptors, all of which result in extracellular matrix remodeling, angiogenesis and invasive growth.⁷⁻¹⁰

Protease activity is normally regulated post-translationally by zymogen activation, cofactor binding or expression of cognate inhibitors. These cognate inhibitors often compete with substrates for binding to the protease substrate-binding pocket. Since these binding pockets are generally conserved among members of a family of proteases, using the macromolecular scaffold of cognate inhibitors to achieve selective inhibition of a single protease is difficult. Therefore, recent studies have instead investigated the use of antibodies as selective inhibitors of protease function in the clotting cascade and cancer.¹¹⁻¹⁶ These antibody inhibitors block substrate-binding through steric hindrance or cause conformational changes after binding at allosteric sites.^{11,13,14,17} Additionally, the molecular basis of inhibition for some of these antibodies has been determined from the crystal structures of the antibody/protease complexes for matriptase and hepatocyte growth factor activator (HGFA).¹⁶⁻¹⁸ These structures show that the interactions with protease surface loops are necessary for antibody specificity, while the insertion of a heavy chain variable loop into the substrate-binding pocket inhibits enzyme machinery in a way reminiscent of small molecule inhibitors. Two of these antibodies, E2 and Ab58, inhibit matriptase and HGFA protease activity respectively through direct interaction with the substrate-binding pocket.^{16,18}

Here we describe two new structures of antibody inhibitors in complex with the target protease matriptase. The antibodies are highly potent, selective inhibitors of protease activity: one (A11) newly discovered from a fully humanized fragment antigen binding (Fab) phage display library; the second (S4) previously discovered from a synthetic single-chain variable fragment (scFv) phage display library.^{15,19} The structures of A11 and S4 in complex with matriptase confirm biochemical studies that the antibodies use different interactions with the protease surface loops to selectively inhibit the protease. However, these antibodies directly inhibit the substrate-binding pocket with a similar, newly identified mechanism. These studies demonstrate the utility of the antibody scaffold to identify different approaches for protease inhibition and reveal the significance of a novel reverse orientation inhibitory motif that was identified from both the natural and synthetic antibody fragment phage display library. Furthermore, the discovery of A11 from a fully humanized Fab library provides a selective agent with potential clinical utility for specific protease inhibition in human cancer.

Results

Identification of a potent matriptase inhibitor from a human Fab library

A Fab phage display library created from human naïve B cells was used to identify inhibitory antibodies against the protease domain of human matriptase. Seven unique antibodies were identified that exhibited inhibitory activity against matriptase in preliminary activity assays with purified protein (**Figure 1**). Of these, Fab A11 demonstrated the most potent inhibition of matriptase. Analysis of the amino acid sequence shows that A11 has a V_H3 heavy chain template and a V_κ3 light chain. The sequences of the hypervariable regions of A11 are shown in **Figure 2**. A11 has a 16 residue heavy chain complementarity determining region 3 (H3) loop, which is longer than the average 12-14 residue H3 loop found in most human antibodies.^{20,21} The unusual length of the H3 loop suggested that it may play a key role in binding to matriptase.

Expression and purification of anti-matriptase Fabs

S4 was raised from a fully synthetic combinatorial scFv phage display library and biochemically characterized previously.¹⁵ The scFv construct proved unsuitable for structural studies, so the variable fragment of S4 was transferred to a Fab scaffold by ligating the variable region to a human Fab constant region.²² The conversion from a scFv to Fab scaffold had minimal effect on the inhibitory potency of the antibody (data not shown). Both recombinant A11 and S4 Fabs were periplasmically expressed in *E. coli* BL21(DE3) cells utilizing the original phagemid vector.²² Purification of the periplasmic fraction over a Ni²⁺ column followed by a size exclusion column yielded approximately 3 mg of protein per L of growth media. The purified protein was determined to be > 98% pure by SDS-PAGE analysis. To boost the production levels of the A11 Fab for subsequent structural studies, a *Trichoderma reesei* system was used. This expression system significantly increased the yield of A11 compared to the *E. coli* system by 60-fold, resulting in a final yield of ~200 mg/L of culture from the growth media that was >98% pure by SDS-PAGE analysis. The expression level achieved is higher than the majority of expression levels reported for Fabs and is at the upper end of Fab expression in *T. reesei*, affirming that *T. reesei* offers a relatively simple, low cost system for high expression of Fab antibodies.^{23,24}

Steady state kinetics show A11 is a potent and specific protease inhibitor

Steady state kinetics experiments were performed to investigate the inhibition of matriptase by A11. A11 binds tightly to matriptase and competitively inhibits turnover of a synthetic peptide substrate (Spectrozyme® tPA) with a K_1 of 720 pM (**Table 1**). To determine the specificity of A11 against a panel of related proteases, assays were performed with the

serine proteases factor Xa (fXa), thrombin, plasmin, tissue plasminogen activator (tPA), urokinase plasminogen activator (uPA), HGFA, hepsin, and prostasin. A11 showed no inhibition of these proteases at a concentration of 1 μ M. These results confirm the potency and selectivity of A11.

The mouse homologue of matriptase, epithin, shares 93% sequence similarity to matriptase in the protease domain; however, the K_I of A11 measured against epithin was 87 nM, 120-fold lower than for matriptase. The surface of matriptase contains six surface loops that differentiate it from other TTSPs in sequence and structure (**Figure 3a**). Epithin contains important sequence differences in the 60s, 140s, and 220s protease surface loops, where A11 binds matriptase (see below), that may contribute to reduced A11 potency. The sequence of the 60s loop (QDDK~~N~~F~~K~~YSDY~~T~~M) in epithin contains the most differences in comparison with matriptase (IDDR~~G~~F~~R~~YSDPTQ). In the 140s loop, the presence of K145 and E146 charged residues in epithin (Q145 and Y146 in matriptase) may affect the potency of A11 due to its proximity to the S1 pocket and the H3 loop. In the 220s loop, instead of the D217 residue found in matriptase, epithin contains a E217 residue that has been shown previously to be an important interaction point for improving the cross-reactivity of the E2 matriptase antibody to epithin.²⁵

Mutational analysis of matriptase reveals surface loop binding by A11

In order to map the surface interactions of A11 with matriptase, A11 inhibition of a library of matriptase alanine mutants was tested. The differences in the proteolytic activity of most of these matriptase mutants versus wild-type were less than twofold.¹⁹ The T98A mutant was sixfold less efficient than wild-type while the D217A mutant had a threefold decrease in proteolytic activity.¹⁹ The results indicate that side chains from four of the six matriptase surface loops are important for binding (**Table 2, Figure 3b**). Mutations at Asp96 and Phe97 had the largest effect on apparent K_I , increasing it to > 1 μ M in both cases which is more than 10⁴-fold over wild-type. This partially explains the high selectivity of A11 for matriptase over other related serine proteases. Of the nine additional proteases tested, HGFA has a Phe residue at position 97 and tPA has an Asp residue at position 96. The mouse homologue of matriptase, epithin, is the only protease tested that contains both Asp96 and Phe97. Additionally, mutation of Asp217 showed a large increase in apparent K_I to 99 nM, an approximately 250-fold increase. Other point mutants revealed that residues Ile60, Asp60a, Asp60b, Asn95, Tyr146, Glu169, and Lys224 are important for antibody recognition of matriptase; however, these interactions play a much smaller role in A11/matriptase binding, resulting in a change in apparent K_I of between three and sixfold.

The apparent K_I values for S4 inhibition against the same library of matriptase alanine mutants were significantly different than those from A11 inhibition (**Table 2**). The results show that the S4/matriptase interaction involves a number of important residues on the six protease loops surrounding the active site.¹⁹ However, unlike A11 no surface loop residue contributes critical binding energy. When the fold apparent K_I values are mapped onto the surface of matriptase, different binding modes and areas of key importance are observed for A11 and S4 (**Figures 3b and 3c**). While these point mutants indicate that binding to the surface loops of matriptase is critical for the antibody inhibitor, this is not the case for the canonical serine protease inhibitor bovine pancreatic trypsin inhibitor (BPTI).¹⁹ Mutations to the matriptase surface loops minimally affect the inhibitory activity of BPTI, allowing for the broad specificity of these Kunitz-type inhibitors against serine proteases. The importance of the protease surface loops for A11 binding offers an explanation for its highly specific interaction with matriptase.

Point mutations of the H3 loops of A11 and S4 suggest a non-canonical binding motif

The longer H3 loop length and presence of an Arg-Arg sequence motif in the H3 loop of A11 is similar to that of the E2 Fab, another antibody inhibitor of matriptase (**Figure 2**). This suggests that ArgH100a and ArgH100b (Kabat numbering) of A11 may play a significant role in the binding to matriptase and the direct interaction with the protease S1 substrate-binding pocket in a manner reminiscent of E2/matriptase interactions.¹⁸ Mutations were made to the Arg-Arg motif to investigate the possible role of these residues in matriptase active site binding. The mutations ArgH100aAla, ArgH100bAla and ArgH100bLys were made and the K_I values were measured (**Table 1**). For ArgH100aAla, the K_I was determined to be 1.5 nM, about twofold greater than the wild-type inhibitor, indicating the limited matriptase binding energy contributed by this residue. However, both the ArgH100bAla and ArgH100bLys point mutants had K_I values of 180 nM against matriptase. These K_I values are 250-fold higher than the wild-type A11, indicating that ArgH100b plays an important role in matriptase binding.

In contrast, the H3 loop of S4 has only one Arg residue ArgH100b that can interact with the S1 pocket. Mutation of ArgH100b to Ala and Lys residues also exhibited significant effects on protease inhibition with K_I values 93- and 110-fold higher, respectively. The change in K_I resulting from the more extreme Arg to Ala mutation seen with both S4 and A11 demonstrates that ArgH100b is the key residue that interacts with the S1 pocket. However, the more conservative Arg to Lys mutation for S4 and A11 also increased the K_I values 110- and 250-fold, respectively. This loss of binding can be related to the fact that the original ArgH100b residue is not fully inserted and instead makes a water-mediated hydrogen bond with the Asp residue at the bottom of the S1 pocket (see below). Furthermore, the same rigidity in the H3 loop that positions the ArgH100b residue in this partially inserted position prevents the shorter Lys residue from interacting with the Asp residue.

Matriptase does not cleave A11 under digestive conditions

To further test whether the H3 loop binds in a substrate-like fashion, A11 was incubated with matriptase at both pH 8.0 and pH 6.0. Previous experiments have shown that the incubation of standard mechanism inhibitors with the target protease at low pH can result in cleavage.^{19,26,27} For A11, matriptase was unable to cleave the Fab at both the standard reaction pH (8.0) and the more acidic pH (6.0), indicating that A11 either does not insert a loop into the active site of matriptase or that it is not bound in a substrate-like manner. Conditions lower than pH 6.0 were not tested because matriptase does not tolerate more acidic pH levels. These results along with the findings from the A11 H3 loop mutations suggest that the H3 loop may bind at the active site in a non-substrate-like fashion, thereby avoiding cleavage. S4 was also previously shown to be not cleaved by matriptase under these conditions,¹⁹ further indicating that they may be utilizing similar binding mechanisms.

3D structures of the Fab/matriptase complexes confirm surface loop binding and identify specificity interactions

To understand the molecular basis for the potency and specificity of these antibodies, the crystal structures of A11 and S4 in complex with matriptase were determined. Both A11/matriptase and S4/matriptase crystallized with one copy of the complex in the asymmetric unit. The structure was determined by molecular replacement to 2.1 Å and 1.5 Å for A11/matriptase and S4/matriptase, respectively. As suggested by the mutational analysis above, the inhibitors bind to matriptase and cap the active site through numerous interactions with the protease surface loops (**Figures 4a and 4b, Table 2**). Furthermore, binding of the antibodies to matriptase did not result in remote conformational changes at either the oxyanion hole or the activating Val16-Asp194 salt bridge.

Both the light chain and heavy chain loops of A11 make significant contacts with the protease surface. In total, A11 buries 1,216 Å² of surface area (**Figure 4c**). The light chain buries 538 Å² of surface area while the heavy chain is responsible for burying a similar 678 Å² of surface area. For the heavy chain, the long H3 loop is responsible for the majority of contacts made with matriptase, burying 466 Å² of surface area (69% of heavy chain total). S4 on the other hand has a total buried surface area of 1080 Å², and the heavy chain accounts for 71% of the total buried surface area (767 Å²) with the H3 loop contributing 89% of this total. Although the overall surface area buried for S4 is less than A11, S4 is more potent than A11 (K_i for S4 = 70.4 pM¹⁹ and K_i for A11 = 720 pM). The potency of S4 could be due to the increased binding of the H3 loop surface area (682 Å²) to the protease surface and key interactions between the light chain and the 140s loop of the matriptase.

The heavy chain loops of A11 make contacts with the 60s and 90s loops of matriptase while inserting the H3 loop into the substrate-binding pocket (**Figure 4a**). The heavy chain complementarity determining region 1 (H1) loop interacts mainly with the 60s protease loop with the O γ of SerH30 and SerH31 forming hydrogen bonds with the Asp60b side chain. The side chain of AlaH33 also in the H1 loop is involved in a hydrophobic interaction with Phe97 contributing to the importance of this residue within the 90s loop. The heavy chain complementarity determining region 2 (H2) loop also makes hydrophobic interactions with Phe97 through AlaH50 and TyrH58. Additional hydrogen bonds are made from the O γ of SerH52, SerH53 and SerH56 and the side chain of TyrH58 in the H2 loop to the side chains of Asn95 and Asp96 in the 90s loop. On the other half of A11, the V κ 3 architecture of the A11 light chain allows for an extended light chain complementarity determining region 3 (L3) loop, which binds in a groove between the 90s and 170s loop on the protease. TrpL94 in particular makes significant interactions with the protease 170s loop. The indole side chain of TrpL94 interacts with Thr177 and the ring of Pro178, while the backbone amide makes a strong hydrogen bond with the backbone carbonyl of Gln174 (2.9 Å). In addition, TyrL96 makes a hydrophobic interaction with Phe97. The light chain complementarity determining region 1 (L1) loop also makes several contacts with the protease 170s and 220s loops. The O γ of SerL31 hydrogen bonds to Asp217 O γ 2 (2.6 Å) and makes a water-mediated hydrogen bond to Lys224 on the protease 220s loop, and the backbone carbonyl of SerL27a forms a 3.0 Å hydrogen bond with N ϵ of the protease Gln174. The light chain complementarity determining region 2 (L2) loop is positioned so that it makes no interactions with matriptase.

Unlike A11, the heavy chain of S4 interacts significantly with the 90s, 140s, and 220s protease surface loops while the light chain interacts mainly with the 140s loop and to a small extent the 60s loop (**Figure 4b**). No extensive interactions were observed between the S4 heavy or light chain and the 170s loop or the 90s loop, although a hydrogen bond (2.7 Å) is formed between SerH31 O γ of the H1 loop and Asp96 O γ 2 and weak hydrophobic interactions are found between IleH53 and PheH54 of the H2 loop and Phe97. The L1 loop binds Tyr146 in the 140s loop via a pi-stacking interaction with TyrL32 and a water-mediated hydrogen bond with the AsnL31 O γ 1, while the H3 loop contributes a water-mediated hydrogen bond with the backbone carbonyl of ArgH100b to Tyr146. The L1 loop also contributes to 140s loop binding via a strong hydrogen bond (2.6 Å) between the TyrL32 O γ and His143 N δ 1.

The A11 structure elucidates why Phe97 and Asp96 are hotspots for inhibitor binding, which was suggested by the analysis of the matriptase alanine mutants. The L3 and H2 loops combine to grab the 90s protease loop containing the critical Phe97 side chain in matriptase. Phe97 binds into a hydrophobic cavity formed by TyrH58 of H2 and ProL95a of L3 with an additional cation-pi interaction made between the ArgL91 of L3 and the phenyl ring of Phe97. The A11/matriptase structure reveals that Asp96 is rotated 180° from its position in

the apo matriptase structure where it forms the bottom of the S4 pocket. The H1 loop has two serine residues SerH52 and SerH53 that make hydrogen bonds with Asp96 in the 90s loop. The structural basis for the Asp217 hotspot revealed by the matriptase mutational analysis is not as clear. Although the residue forms a hydrogen bond with the SerL31 O γ of L1, this single interaction alone cannot explain the 250-fold change in the apparent K_I ; however, the threefold decrease in proteolytic activity of this mutant¹⁹ may help account for this large change. Structural changes to the entire 220s loop from the Ala mutation may explain the interruption of additional interactions with A11.

Unlike the interaction between A11 and matriptase residue Phe97 with the antibody loops forming a clamp around this residue, the S4/matriptase complex does not show a similar binding hotspot. The Phe97Ala mutation had little effect on inhibition by S4¹⁹ since the hydrophobic interactions of the H1 and H2 loops of S4 were fairly weak. In contrast, matriptase residues Phe146 and His143 in the 140s loop appear to be anchor points that interact with residues from both the S4 light and heavy chains, which is in agreement with the positions exhibiting the largest changes in the mutational studies.¹⁹ The interaction of the light chain with the 140s loop also explains why the H143A and Y146A mutations exhibit 24- and 20-fold increased apparent K_I values compared to the wild-type protease. These changes were not observed for A11 and E2. The moderate increase in the apparent K_I (13-fold increase) for D96A is due to the loss of a strong hydrogen bond between SerH31 O γ of H1 loop and Asp96 O δ 2. The mutation to D217A also caused a moderate increase in apparent K_I (12-fold) due to the loss of a hydrogen bond with ArgH100b carbonyl mediated by two water molecules.

A11 and S4 inhibit matriptase with novel presentation of the H3 loops at the protease substrate-binding pocket

The H3 loops of both A11 and S4 adopt similar conformations in the protease active site (**Figures 5a and 5b**). The H3 loop of A11 forms a β -hairpin turn that reaches into the protease active site but makes few contacts with the protease. AlaH99, AlaH100, and ValH100d combine to bury 174 Å² of surface area in hydrophobic interactions with the protease as the β -hairpin strand is extended into the active site. At the apex of the turn, A11 has two arginine residues. The C-terminal arginine (ArgH100b) binds sub-optimally, relative to a substrate, in the S1 pocket, while the N-terminal arginine (ArgH100a) extends towards the prime side of the protease active site. This conformation results in the binding of the putative scissile bond in a reverse orientation in the active site, thus explaining the lack of cleavage at this position by the protease (**Figures 6a and 6b**). In the S1 pocket, ArgH100b makes a 2.8 Å water-mediated hydrogen bond to Asp189 of matriptase (**Figure 5c**). This positioning is in contrast to the typical salt bridge formed by P1 arginine substrates and mimics such as benzamidine,²⁸ and is instead similar to the binding of a shorter P1 Lys residue observed in BPTI. The water-mediated interaction explains why the ArgH100bLys mutation is so deleterious to A11 inhibition (**Table 1**). The lysine side chain is one carbon shorter than the arginine side chain; therefore, ArgH100bLys cannot make a similar water-mediated hydrogen bond to the Asp residue at the bottom of the S1 pocket. The H3 loop residues do not occupy any additional substrate-binding pockets due to the sharp turn of the loop out of the protease. The potency and specificity of A11 are supplemented by the H1, H2 and L3 loop interactions with protease surface loops.

Similarly, the H3 loop of S4 approaches the matriptase S1 pocket in a way that also extends residue ArgH100b into the binding pocket and presents the putative scissile bond in a reverse orientation (**Figure 5b**). The positioning of residue ArgH100b in the active site is nearly identical to that seen in the A11/matriptase structure with the side chain making the same water-mediated hydrogen bond with Asp189 on matriptase (**Figure 5c**). S4 has a longer H3 loop than A11 and achieves this reverse orientation positioning with a much more

gradual turn to allow for greater interaction between the side chains and matriptase. This results in a higher total buried surface area due to the interactions of the longer side chains of ArgH99, TyrH100a, and ArgH100b residues with matriptase (**Figure 5d**). The C-terminus of the loop approaches the S1 pocket with ArgH100b followed by a turn that exposes TyrH100a to solvent and extends ArgH100 into the prime side at the P2' pocket. The next residue ArgH99 skips the S2 pocket and partially inserts into the S3 pocket before exiting the binding pocket (**Figure 6c**).

This novel presentation of the H3 loop at the protease active site in an orientation reverse of that for a standard substrate allows for potent inhibition of matriptase and avoidance of proteolytic inactivation of the A11 and S4 Fabs. In the other antibody/protease structures available in the PDB (PDB code 2R0K – HGFA and Ab58; PDB code 2R0L – HGFA and Fab75; PDB code 2ZCH – prostate specific antigen and activating antibody; PDB code 3BN9 – matriptase and E2; PDB code 3K2U – HGFA and Fab-40; PDB code 3R1G – BACE1 and exosite-binding antibody), this reverse orientation of the H3 loop is not seen nor is the H3 loop involved in inhibition in some cases. The antibodies in 2R0L, 2ZCH, 3K2U, and 3R1G do not target the protease substrate-binding pocket, with the 2R0L, 3K2U, and 3R1G antibodies acting as allosteric inhibitors instead. The H2 loop of Ab58 interacts with a portion of the HGFA substrate-binding pocket but not the S1 pocket.¹⁶ The long H3 loop of E2 inserts into the matriptase substrate-binding pocket, but binds in a canonical fashion (**Figure 6d**) that leads to proteolysis of the loop under acidic conditions.^{18,19}

Peptides based on H3 loop sequences reveal the importance of the antibody scaffold for matriptase inhibition

Given the importance of the H3 loops in inhibiting matriptase, peptides containing the H3 loop sequences of A11, S4, and E2 were synthesized to test for inhibition of matriptase activity independently of the Fab scaffold. A circularized version of the H3 loop peptides was synthesized for analysis. The IC₅₀ values for the A11, S4, and E2 peptides against matriptase indicate that they are generally weak inhibitors in the circularized version at 0.41 ± 0.01 mM, 0.032 ± 0.005 mM, and 1.3 ± 0.1 mM, respectively (**Figure 7**). This inhibition is substantially weaker than the Fab version and suggests that the Fab scaffold plays an important role in inhibition potency. Follow-up analysis of the peptide IC₅₀ values after 24 hour incubation with matriptase showed a marked change for the A11 and S4 peptides. The E2 peptides showed little change in IC₅₀ values after 24 hours. Mass spectrometric analysis of the A11 peptides confirmed that the change in IC₅₀ was the result of near complete cleavage of the A11 peptide by matriptase, with the cleavage event C-terminal to ArgH100a. This result indicates that the A11 peptide may not bind in a reverse orientation as seen in the A11 Fab. The S4 peptide showed modest improvement in inhibition after 24 hours with a decreased IC₅₀ value of 0.010 ± 0.002 mM. However, partial digestion of the S4 peptide was observed, with the major cleavage event C-terminal to ArgH100b. The slight increase in inhibition over time may be a result of product inhibition from the cleaved peptide. The E2 peptide showed no cleavage after 24 hours, supporting the unchanged IC₅₀.

Since the H3 loop of A11 makes limited contacts with the protease outside the S1 pocket, it is unlikely that the A11 peptide would adopt the reverse binding orientation without the scaffold and therefore unsurprising that the peptide is cleaved. Furthermore, the H3 loop of A11 buries a small percentage (38%) of surface area relative to the total buried surface area in the matriptase substrate-binding pocket when compared to S4 (63%) and E2 (68%). The H3 loops of S4 and E2 are assisted by more extensive interactions with the substrate-binding pocket,¹⁸ and these contacts make it more likely that the circularized peptide would act as an inhibitor. The weak inhibition exhibited by the E2 peptide in the absence of any cleavage indicates that the sequence may bind in a similar orientation to that of the H3 loop of the Fab

allowing it to avoid cleavage, presumably due to the rigidity provided by the proline residues in the sequence. Since the S4 peptide was the most potent inhibitor out of the three, this suggests that its interactions with the substrate-binding pocket may contribute more significantly to the inhibition of matriptase independent of the entire scaffold, though whether it does so in a reverse orientation is difficult to determine. The cleavage of some of the S4 peptide by matriptase over time suggests that it may not adopt this novel conformation consistently or at all without the aid of the scaffold. While the H3 loop alone contributes to a certain degree to protease inhibition, these results demonstrate that the antibody scaffold plays a significant role in the correct orientation of the H3 loop for inhibition and that additional interactions provided by the remaining Fab loops are necessary to achieve potent inhibition.

Discussion

As antibodies continue to demonstrate their usefulness in many therapeutic areas, one particular field of recent interest is protease inhibition for treatment of cancer progression. Since proteases have historically been difficult to target selectively with small molecules and macromolecular cognate inhibitors, the high affinity and specificity of antibodies are ideal for this purpose.^{29,30} An emphasis of the current study was to identify selective antibodies that competitively inhibit proteolytic activity. As a result, three different inhibitory antibodies of matriptase that all utilize a long H3 loop with a P1-Arg residue to partially fill the S1 pocket were isolated and characterized. The interaction between the Arg and the Asp residue at the bottom of the binding pocket is mediated by a water molecule, resulting in a binding mode that prevents proteolysis of the Fab. These antibodies are different from the HGFA antibodies that are mostly allosteric inhibitors which contain shorter H3 loops.^{16,17} Here the structural characteristics of specific protease inhibition using A11 and S4 Fabs are revealed. As with many enzyme-inhibiting antibodies, A11 and S4 inhibit substrate binding by capping the active site through interactions with the surrounding surface loops (**Figure 4**).^{14,16,18} These loops are areas of low homology among trypsin-fold serine proteases and can be targeted with high selectivity by antibodies. Since these antibodies use the surface loop interactions for specificity, this differentiates them from naturally-occurring macromolecular protease inhibitors that bind a broader array of proteases by interacting with the highly conserved protease active site.³¹

The crystal structures of the A11/matriptase and S4/matriptase complexes reveal a number of similarities with other antibody protease inhibitors for which structures have recently been reported, in particular the HGFA inhibitor Ab58¹⁶ and the matriptase inhibitor E2.¹⁸ All of these antibodies depend on the protease surface loops for recognition and binding to their target protease. For A11, E2, and Ab58, the 90s loop of the protease and the Phe97 residue in particular acts as a key anchor point for these antibodies. However, upon protease capping the buried surface area of the Ab58/HGFA interaction is considerably less (890 Å²)¹⁶ when compared to A11/matriptase (1,216 Å²) and E2/matriptase (1,241 Å²) (**Figure 4c**), which are both significantly greater than a typical antibody/antigen interaction (< 900 Å²).^{32,33} Ab58 only partially occludes the substrate-binding pocket at the S2 and S3 pocket.¹⁶ Conversely, the matriptase inhibitors fully occlude the substrate-binding pocket and interact with the S1 pocket by inserting a long H3 loop, resulting in the enhanced interactions.

Focusing on the antibody inhibitors of matriptase, a closer similarity in antibody/protease interactions is observed with A11 and E2. A11 and E2 have identical H1 and H2 loops even though A11 was isolated from a naïve human B cell phage display library and E2 from a fully synthetic phage display library.¹⁵ These loops are critical for anchoring the antibody between the 60s and 90s loops on the surface of matriptase (**Figure 4a**), creating very

similar binding footprints for A11 and E2. Most of the interactions in the A11/matriptase and E2/matriptase complexes are made by non-polar residues. In contrast, S4 binds matriptase in a different mode and does not utilize the interaction with the 90s loop (**Figure 4b**). S4/matriptase interactions are instead formed by polar residues of the H3 and L1 loops interacting with the 140s loop via a strong network of hydrogen bonds and pi-stacking interactions. S4 is the most potent matriptase inhibitor out of the three and this may be the result of the increased buried surface area localized to the H3 loop ($682 \text{ \AA}^2$) found in the substrate-binding pocket (**Figure 4c**). Instead of relying on the S1 pocket for critical binding energy and inhibition as with A11 and E2, S4 adopts a unique mechanism where the ArgH99 residue forms a strong salt bridge with Gln175 in the S3 pocket to position the H3 loop in the most energetically stable orientation (**Figures 5b and 5d**). On the other hand, the H3 loop of E2 is the longest of our matriptase inhibitors and adopts a more extended conformation in the matriptase active site when compared to the A11 H3 loop. As a result, the H3 loop of A11 buries less surface area ($466 \text{ \AA}^2$) compared to the H3 loop of E2 ($847 \text{ \AA}^2$) (**Figure 4c**). To compensate for this decrease in buried surface area, A11 utilizes a different light chain. The $V_{\kappa}3$ light chain of A11 buries $538 \text{ \AA}^2$ of surface area, while the light chain of E2 buries just $175 \text{ \AA}^2$ (**Figure 4c**). Attempts to make a chimera composed of the A11 light chain and E2 heavy chain that should bury a larger total surface area resulted in reduced affinity (data not shown), indicating that the loops are precisely positioned for each antibody and not interchangeable despite the similarities in the two antibodies.

The discovery and characterization of antibodies that inhibit through H3 loop interactions with the protease active site represent a new approach to specific inhibition of protease activity. The novel, reverse orientation of the H3 loop in the protease active site demonstrates that the antibody scaffold is capable of revealing completely new inhibition mechanisms, suggesting that there are still undiscovered approaches to inhibiting these enzymes. In the case of A11 and S4, interactions with the protease surface loops orient the antibody so that the long H3 loop can be inserted into the protease active site. Interaction of the H3 loop with the active site occurs in a unique fashion, with the H3 loop binding in a reverse orientation while partially inserting an Arg residue into the S1 pocket (**Figure 6**). The reverse orientation combined with the shallow arginine binding at the S1 pocket prevents hydrolysis of the H3 loop while at the same time providing critical binding energy. A few cases of reverse binding of peptides near the active site have been recorded, most notably inhibition of cathepsin pro-peptide and BIR2 domain inhibition of caspase 3 and caspase 7.³⁴ Although they share some similarities with A11 and S4, the mechanism utilized by our antibodies remains unique, relying upon a direct interaction with the primary substrate-binding pocket of the protease. Even so, the resemblance of this reverse orientation in the mechanism used by BIR2 inhibition³⁴ of the cysteine proteases caspase 3 and caspase 7 indicates that this approach could also be used in identifying selective active site inhibitors of protease classes beyond the serine proteases. For example, this approach could be extremely beneficial for the study of metallo- and aspartyl proteases, two classes that have been historically difficult to design specific inhibitors against.

Selective protease inhibition is particularly useful in the clinic to target misregulated protease activity responsible for disease pathophysiology. For example, the FDA-approved recombinant polypeptide ecallantide (Kalbitor) is used to specifically inhibit the serine protease plasma kallikrein that causes acute hereditary angioedema attacks in patients.³⁵ Moreover, selective protease inhibition may be particularly important in treating and monitoring cancer progression. Matriptase is over-expressed in breast, colon and prostate cancers. It is normally regulated by its cognate inhibitor hepatocyte growth factor activator inhibitor-1 (HAI-1); however, dysregulation of matriptase activity promotes cancer progression.³⁶ Using matriptase specific antibodies, protease activity can be monitored in disease to develop a novel approach to follow early staged malignancies. In the recent work

by Darragh et al, the A11 antibody was used as a non-invasive imaging probe to characterize protease activity *in vivo* in order to define matriptase as an early biomarker to visualize epithelial cancers in pre-clinical mouse models.³⁷ Furthermore, the recent discovery of the role of matriptase in squamous cell carcinoma³⁸ highlights the need for agents that can selectively inhibit protease activity to pharmacologically probe the pathophysiological role of the enzyme and to provide potential therapeutic applications. Here we have shown that antibodies can provide novel solutions for the selective inhibition of proteases. Our discovery highlights the importance of the antibody scaffold to uncover unique and unpredictable positioning of the inhibitory loops to bind and inhibit protease function in a highly selective manner. The identification of a fully human, inhibitory recombinant antibody A11 validates this approach and reaffirms the use of antibodies for selective inhibition of protease targets in cancer.

Materials and Methods

Identification of inhibitory Fabs from a human phage display library

A Fab library created from naïve B cells was used to identify inhibitory antibodies against the human matriptase protease domain (hMT-SP1).³⁹ Active matriptase was immobilized in wells of a 96-well ELISA plate. The panning was accomplished in three rounds with increasing stringency against hMT-SP1 adsorbed to wells. ELISAs were performed to verify binding of the identified Fabs to hMT-SP1. ELISA positive clones were expressed, purified and tested for inhibition of matriptase. Individual clones were sequenced to verify their uniqueness.

Protein expression and purification from *E. coli*

Matriptase and matriptase mutants were expressed in *Escherichia coli* and purified as previously described.^{6,19} S4 was cloned into the Fab scaffold following a procedure similar to that described in Farady et al.¹⁸ A11 and S4 Fabs were expressed in *E. coli* BL21 DE3 cells. Cultures were grown in 1 L of 2xYT containing 100 µg/ml ampicillin and 0.1% glucose at 37 °C and 250 rpm to an OD₆₀₀ of 0.6-0.8. The temperature was then reduced to 25 °C and the cultures were induced with the addition of 0.5 mM IPTG. After 18 hours of growth, the bacteria were harvested and pelleted by centrifugation. The cells were resuspended in 25 mL of buffer containing 0.2 M Tris pH 8.0, 0.5 mM EDTA and 0.5 M sucrose. The resuspended solution was left on ice for 1 hour. The solution was then pelleted and the periplasmic fraction was run over a Ni²⁺ column prewashed with wash buffer (50 mM Tris pH 8.0, 250 mM NaCl). The Ni²⁺ column was then washed with 10 column volumes of the wash buffer and the Fab was eluted with 250 mM imidazole in 50 mM Tris pH 8.0 and 100 mM NaCl. Size exclusion chromatography was carried out on the eluted A11 using a Superdex S75 26/60 with a 50 mM Tris pH 8.0, 100 mM KCl and 5% glycerol buffer.

Site-directed mutagenesis of A11 and S4 antibodies

A11 mutants ArgH100aAla, ArgH100bAla, and ArgH115bLys and S4 mutants ArgH100bAla and ArgH100bLys were all created using the QuikChange™ kit from Stratagene. Sequences were verified by DNA sequencing. Expression and purification of A11 and S4 mutants were carried out as described above.

Trichoderma reesei expression vector construction

Two independent expression vectors were constructed: one for expression of the Fab heavy chain (pCBHixFabA11 H1) and one for the light chain (pCBHixFabA11 L1). In each case, the Fab chains were produced as fusion proteins with the *T. reesei* CBHI (cellobiohydrolase

I, cel7a) catalytic core and linker region. A Kex2 cleavage site (Val Ala Val Tyr Lys Arg) was positioned between the CBHI and the Fab chain to allow cleavage of the fusion protein after the Arg residue and release of the Fab chain during secretion.

The following segments of DNA were assembled in the construction of pCBHIxFabA11 H1 and pCBHIxFabA11 L1. The *T. reesei cbh1* promoter and coding region start at a naturally occurring *Xba*I site approximately 1500 bp upstream of the coding region. The synthetic, codon optimized coding region for each Fab chain was fused to the end of the CBHI linker region at a created *Spe*I restriction site (see below). Immediately after the Fab stop codon was an *Asc*I restriction site followed by the *T. reesei cbh1* terminator region (356 bp). This was followed by a 2.75 kb fragment of *Aspergillus nidulans* genomic DNA, including the promoter, coding region and terminator of the *amdS* (acetamidase) gene. The above DNA fragments were inserted in pNEB193 (New England Biolabs, Inc., USA) between the *Xba*I and *Kpn*I sites of the multiple cloning site.

The following changes were made within the *cbh1* open reading frame. The codon for amino acid 212 of the mature CBHI protein was changed from GAG (Glutamic acid) to CAG (Glutamine) resulting in production of an inactive form of CBHI. Within the coding region for the CBHI linker a change was made to create a *Spe*I restriction site. This altered the DNA sequence from ACCCAG to ACTAGT, changing the amino acid sequence at the end of the CBHI linker region from Thr Gln to Thr Ser. The Gln in this sequence represents the first amino acid of the cellulose binding domain of CBHI.

***T. reesei* transformation**

Trichoderma reesei GICC20000150 was derived from strain RL-P37⁴⁰ by sequential deletion of the genes encoding the four major secreted cellulases (cel7a, cel6a, cel7b and cel5a). Transformation was performed using a Bio-Rad Laboratories, Inc. (Hercules, CA) model PDS-1000/He biolistic particle delivery system according the manufacturer's instructions. Transformants were selected on solid medium containing acetamide as the sole nitrogen source. For antibody production, transformants were cultured in a liquid minimal medium containing lactose as carbon source as described previously,⁴¹ except that 100 mM piperazine-*N,N'*-bis (3-propanesulfonic acid) (Calbiochem) was included to maintain the pH at 5.5. In order to produce Fab, it was necessary for transformants to have taken up both the heavy and light chain expression vectors. However, both expression vectors had the same *amdS* selectable marker so it was not immediately possible to recognize co-transformants. Culture supernatants were analyzed by SDS-PAGE under reducing conditions and those that contained the highest level of a 25 kDa band (representing heavy and/or light chain) and an apparent 60 kDa band (representing the CBHI core and linker) were selected for further analysis.

Purification of A11 from *T. reesei* expression

Media from the *Trichoderma* expression was adjusted to pH 5.5. For an initial crude purification, the media was run over an SP sepharose column equilibrated with Wash Buffer 1 (100 mM MES pH 5.5, 50 mM NaCl). The column was then washed with 5 column volumes of Wash Buffer 1, followed by 5 column volumes of Wash Buffer 2 (50 mM Tricine pH 8.0). A11 was eluted with 3 column volumes of 50 mM Tricine pH 8.0, 500 mM NaCl. The elution was buffer exchanged into 100 mM MES pH 5.5, 50 mM NaCl and loaded onto a MonoS HR 5/5 column. The column was washed with Wash Buffer 1 followed by Wash Buffer 2. Elution was then carried out in a 0-100% gradient of Wash Buffer 2 to Wash buffer 2 containing 500 mM NaCl. Further purification was carried out on a Superdex 75 26/60 size exclusion column with a 50 mM Tris pH 8.0, 100 mM KCl, 5% glycerol buffer.

Steady-state kinetics

Kinetics were carried out as previously described.¹⁹ Briefly, all reactions were carried out in 50 mM Tris, pH 8.8, 50 mM NaCl, 0.01% Tween-20 in a 96-well, medium binding, flat-bottomed plate (Corning) where cleavage of substrate Spectrozyme® tPA (hexahydroxytyrosyl-Gly-Arg-pNA, American Diagnostica, Greenwich, CT) was monitored in a UVmax Microplate Reader (Molecular Devices Corporation, Palo Alto, CA.). K_I values were measured using the tight-binding inhibition equations of Williams and Morrison.⁴² When measuring the effect mutations to matriptase had on the strength of the interaction between the protease and inhibitor, IC_{50} values were used instead of K_I values. Reactions to determine the IC_{50} values were carried out by incubating 0.2 nM enzyme with inhibitor for > 5 hours to assure steady-state behavior of the system. Apparent K_I values were then calculated from IC_{50} values as shown previously in order to normalize the IC_{50} with respect to the strength of the protease/substrate interaction.⁴³ Inhibitory activity against related proteases was measured using a similar assay monitoring the cleavage of a p-nitroanilide substrate. 10 nM Thrombin, fXa, and plasmin (Haematologic Technologies, Inc., Essex Junction, VT.) were incubated with 1 μ M Fab, and the reaction was monitored using 1 mM of the substrate T1637 (Sigma, St. Louis, MO.). 10 nM tPA and uPA (American Diagnostica) were incubated with 1 μ M Fab, and the reaction was monitored using 1 mM Spectrozyme® tPA and 400 mM Spectrozyme® UK (American Diagnostica), respectively. 1 nM Hepsin and 5 nM Protsasin (R&D Systems, Minneapolis, MN) were incubated with 1 μ M Fab, and the reaction was monitored using 0.5 mM Spectrozyme® tPA. Kaleidagraph 3.6 was used to fit all graphs and equations (Synergy Software, Reading, PA).

A11 inhibitor digestion

The digestion of A11 by matriptase was carried out as previously described.¹⁹ A11 was incubated at 2 μ M with 0.1 nM matriptase in either 100 mM MES pH 6.0, 100 mM NaCl buffer or 50 mM Tris pH 8.0, 100 mM NaCl. After 120 hours, the samples were run on a 4-20% Tris-Glycine SDS-PAGE gel (Invitrogen) to visualize.

Crystallization and data collection

A11 and S4 was incubated with matriptase at 1:1 molar ratio and the complex was purified by gel filtration using a Superdex S75 26/60 column in a buffer containing 50 mM Tris pH 8.0, 100 mM NaCl, 5% glycerol. The purified complex (A11/matriptase and S4/matriptase) was then concentrated to ~15 mg/ml. Initial crystallization conditions were discovered using a nanoliter-scale Mosquito robot (TTP Labtech). The A11/matriptase complex was crystallized in 16% PEG 3350, 0.23 M $MgSO_4$, 0.4% isopropanol, 3% glycerol and 0.12 M $AmSO_4$ in hanging drop by vapor diffusion. Crystals belonging to the hexagonal space group $P6_4$ ($a = b = 130.6$ Å and $c = 96.9$ Å; $\beta = 90.0^\circ$) grew in three days and were cryoprotected in the mother liquor supplemented with 30% sucrose. Needles like crystals of the S4/matriptase complex were grown from 0.2 M NaCl, 25% PEG 3355 and 1% ethyl acetate as additive at 25°C using the hanging drop vapor diffusion technique. These crystals belong to the monoclinic space group $P2_1$, ($a = 39.2$ Å, $b = 84.0$ Å, $c = 101.4$ Å; $\beta = 91.5^\circ$), diffract x-rays to beyond 1.5 Å resolution, and have one molecule in the asymmetric unit. The S4/matriptase crystals were transferred to 0.2 M NaCl, 25% PEG 3355, 1.0% ethyl acetate, 25% ethylene glycol. Diffraction data were collected at beamline 8.3.1 at the Advanced Light Source at LBNL (see **Table 3**). Diffraction images for A11/matriptase and S4/matriptase were reduced and integrated using DENZO and the intensities were merged and scaled using SCALEPACK from the HKL-2000 suite.⁴⁴

Structure determination and refinement

The structure of A11/matriptase and S4/matriptase were solved by molecular replacement using Phaser⁴⁵ in CCP4,⁴⁶ first searching for matriptase (using 1EAX.pdb as search model), then searching for the Fab fragment with its H3 loop truncated (using 2HFF.pdb as search model). After molecular replacement, automatic building in ARP/wARP⁴⁷ and manual building followed by restrained refinements cycles using REFMAC5⁴⁸ yielded the final structures. TLS refinement was applied at the last stage of the refinement for A11/matriptase. In the final model building/refinement cycles, water molecules were inserted at stereochemically reasonable sites that had features in electron density, and individual restrained *B*-values were refined. Five percent of all reflections were omitted from the refinement to calculate R_{free} ; the final *R* and R_{free} are 18.8 and 22.8%, respectively for S4/matriptase. For A11/matriptase, the final *R* and R_{free} are 16.0 and 19.4%, respectively. The final model of both complexes converged rapidly, yielding a model with excellent parameters (see Table 3). In the final structure of A11/matriptase, there was no density for the heavy chain residues 129-133, 213-215, or protease Ala204. There was no side chain density for A11 light chain residues Glu1, Glu143, Lys188, and Glu213, or heavy chain residues Glu1, Lys201, and Lys210, so the side chains were not modeled. These regions are often disordered in Fab structures. The whole main chain of the S4/matriptase catalytic domain is in appropriate electron density. In S4/matriptase, the side chain occupancy was set to 0.5 since the side chains exhibit two conformations. The quality of the final structures was assessed using Molprobit.⁴⁹ Buried surface area calculations were performed using PISA.³³

Synthesis of H3 loop peptides

The H3 loop sequences of A11 (circularized: GLGIAARRFVSGEG), S4 (circularized: GFHIRRYRSGYYEG), and E2 (circularized: GLTYPQRRGPQNVSEG) were synthesized by standard solid phase Fmoc chemistry on a Protein Technologies Symphony Quartet automated peptide synthesizer using Wang resin preloaded with the C-terminal amino acid (Novabiochem). An ODmab protected glutamic acid residue was inserted near the C-terminus (Novabiochem). To circularize the peptide, the ODmab protecting group was selectively removed by 2% hydrazine in DMF followed by Fmoc deprotection of the N-terminal residue and a standard coupling reaction with HOBt/HBTU/DIPEA. The finished peptides were deprotected and cleaved from the resin using 95% TFA and purified by HPLC on a C-18 column. The peptide sequences were confirmed by MALDI mass spectrometry.

Peptide IC₅₀ determination

The IC₅₀ of each peptide were carried out under conditions similar to those for the steady state kinetics described earlier. Assays were run containing 200 μ M substrate (Spectrozyme® tPA), 0.2 nM matriptase, and varying peptide concentrations. A 50 mM stock solution of each peptide in DMSO was used for all assays. Additional DMSO was added to all assays so that the DMSO concentration was equal throughout. Assays were run for 30 minutes and monitored on a fluorescent plate reader at 405 nm. Data was plotted using Kaleidagraph.

Acknowledgments

This work was funded in part by grants from the National Institutes of Health CA128765 (CSC) and GMO82250 (CSC). We also thank Professors Robert Fletterick and Matthew Jacobsen for helpful discussions and Dr. Gregory Lee for careful reading of the manuscript.

References

1. Carter PJ. Potent antibody therapeutics by design. *Nat Rev Immunol.* 2006; 6:343–57. [PubMed: 16622479]
2. Brekke OH, Sandlie I. Therapeutic antibodies for human diseases at the dawn of the twenty-first century. *Nat Rev Drug Discov.* 2003; 2:52–62. [PubMed: 12509759]
3. Chambers AF, Matrisian LM. Changing views of the role of matrix metalloproteinases in metastasis. *J Natl Cancer Inst.* 1997; 89:1260–70. [PubMed: 9293916]
4. Coussens LM, Werb Z. Matrix metalloproteinases and the development of cancer. *Chem Biol.* 1996; 3:895–904. [PubMed: 8939708]
5. Netzel-Arnett S, Hooper JD, Szabo R, Madison EL, Quigley JP, Bugge TH, Antalis TM. Membrane anchored serine proteases: a rapidly expanding group of cell surface proteolytic enzymes with potential roles in cancer. *Cancer Metastasis Rev.* 2003; 22:237–58. [PubMed: 12784999]
6. Takeuchi T, Shuman MA, Craik CS. 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 U S A.* 1999; 96:11054–61. [PubMed: 10500122]
7. Bhatt AS, Welm A, Farady CJ, Vasquez M, Wilson K, Craik CS. Coordinate expression and functional profiling identify an extracellular proteolytic signaling pathway. *Proc Natl Acad Sci U S A.* 2007; 104:5771–6. Epub 2007 Mar 27. [PubMed: 17389401]
8. Kilpatrick LM, Harris RL, Owen KA, Bass R, Ghorayeb C, Bar-Or A, Ellis V. Initiation of plasminogen activation on the surface of monocytes expressing the type II transmembrane serine protease matriptase. *Blood.* 2006; 108:2616–23. Epub 2006 Jun 22. [PubMed: 16794252]
9. Suzuki M, Kobayashi H, Kanayama N, Saga Y, Lin CY, Dickson RB, Terao T. Inhibition of tumor invasion by genomic down-regulation of matriptase through suppression of activation of receptor-bound pro-urokinase. *J Biol Chem.* 2004; 279:14899–908. Epub 2004 Jan 27. [PubMed: 14747469]
10. Takeuchi T, Harris JL, Huang W, Yan KW, Coughlin SR, Craik CS. Cellular localization of membrane-type serine protease 1 and identification of protease-activated receptor-2 and single-chain urokinase-type plasminogen activator as substrates. *J Biol Chem.* 2000; 275:26333–42. [PubMed: 10831593]
11. Colwell NS, Blinder MA, Tsiang M, Gibbs CS, Bock PE, Tollefsen DM. Allosteric effects of a monoclonal antibody against thrombin exosite II. *Biochemistry.* 1998; 37:15057–65. [PubMed: 9790668]
12. Devy L, Huang L, Naa L, Yanamandra N, Pieters H, Frans N, Chang E, Tao Q, Vanhove M, Lejeune A, van Gool R, Sexton DJ, Kuang G, Rank D, Hogan S, Pazmany C, Ma YL, Schoonbroodt S, Nixon AE, Ladner RC, Hoet R, Henderikx P, Tenhoor C, Rabbani SA, Valentino ML, Wood CR, Dransfield DT. Selective inhibition of matrix metalloproteinase-14 blocks tumor growth, invasion, and angiogenesis. *Cancer Res.* 2009; 69:1517–26. Epub 2009 Feb 10. [PubMed: 19208838]
13. Dickinson CD, Shobe J, Ruf W. Influence of cofactor binding and active site occupancy on the conformation of the macromolecular substrate exosite of factor VIIa. *J Mol Biol.* 1998; 277:959–71. [PubMed: 9545384]
14. Petersen HH, Hansen M, Schousboe SL, Andreasen PA. 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.* 2001; 268:4430–9. [PubMed: 11502203]
15. Sun J, Pons J, Craik CS. Potent and selective inhibition of membrane-type serine protease 1 by human single-chain antibodies. *Biochemistry.* 2003; 42:892–900. [PubMed: 12549907]
16. Wu Y, Eigenbrot C, Liang WC, Stawicki S, Shia S, Fan B, Ganesan R, Lipari MT, Kirchhofer D. Structural insight into distinct mechanisms of protease inhibition by antibodies. *Proc Natl Acad Sci U S A.* 2007; 104:19784–9. Epub 2007 Dec 5. [PubMed: 18077410]
17. Ganesan R, Eigenbrot C, Wu Y, Liang WC, Shia S, Lipari MT, Kirchhofer D. Unraveling the allosteric mechanism of serine protease inhibition by an antibody. *Structure.* 2009; 17:1614–24. [PubMed: 20004165]

18. Farady CJ, Egea PF, Schneider EL, Darragh MR, Craik CS. Structure of an Fab-protease complex reveals a highly specific non-canonical mechanism of inhibition. *J Mol Biol.* 2008; 380:351–60. Epub 2008 May 11. [PubMed: 18514224]
19. Farady CJ, Sun J, Darragh MR, Miller SM, Craik CS. The mechanism of inhibition of antibody-based inhibitors of membrane-type serine protease 1 (MT-SP1). *J Mol Biol.* 2007; 369:1041–51. Epub 2007 Apr 4. [PubMed: 17475279]
20. Wu TT, Johnson G, Kabat EA. Length distribution of CDRH3 in antibodies. *Proteins.* 1993; 16:1–7. [PubMed: 8497480]
21. Zemlin M, Klinger M, Link J, Zemlin C, Bauer K, Engler JA, Schroeder HW Jr, Kirkham PM. Expressed murine and human CDR-H3 intervals of equal length exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures. *J Mol Biol.* 2003; 334:733–49. [PubMed: 14636599]
22. de Haard HJ, van Neer N, Reurs A, Hufton SE, Roovers RC, Henderikx P, de Bruine AP, Arends JW, Hoogenboom HR. A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *J Biol Chem.* 1999; 274:18218–30. [PubMed: 10373423]
23. Arbabi-Ghahroudi M, Tanha J, MacKenzie R. Prokaryotic expression of antibodies. *Cancer Metastasis Rev.* 2005; 24:501–19. [PubMed: 16408159]
24. Keranen S, Penttilä M. Production of recombinant proteins in the filamentous fungus *Trichoderma reesei*. *Curr Opin Biotechnol.* 1995; 6:534–7. [PubMed: 7579664]
25. Farady CJ, Sellers BD, Jacobson MP, Craik CS. Improving the species cross-reactivity of an antibody using computational design. *Bioorg Med Chem Lett.* 2009; 19:3744–7. Epub 2009 May 7. [PubMed: 19477127]
26. McGrath ME, Hines WM, Sakanari JA, Fletterick RJ, Craik CS. The sequence and reactive site of ecotin. A general inhibitor of pancreatic serine proteases from *Escherichia coli*. *J Biol Chem.* 1991; 266:6620–5. [PubMed: 2007606]
27. Ozawa K, Laskowski M Jr. The reactive site of trypsin inhibitors. *J Biol Chem.* 1966; 241:3955–61. [PubMed: 5950769]
28. Friedrich R, Fuentes-Prior P, Ong E, Coombs G, Hunter M, Oehler R, Pierson D, Gonzalez R, Huber R, Bode W, Madison EL. Catalytic domain structures of MT-SP1/matrilysin, a matrix-degrading transmembrane serine proteinase. *J Biol Chem.* 2002; 277:2160–8. Epub 2001 Nov 5. [PubMed: 11696548]
29. Coussens LM, Fingleton B, Matrisian LM. Matrix metalloproteinase inhibitors and cancer: trials and tribulations. *Science.* 2002; 295:2387–92. [PubMed: 11923519]
30. Lee GM, Craik CS. Trapping moving targets with small molecules. *Science.* 2009; 324:213–5. [PubMed: 19359579]
31. Laskowski M, Kato I. Protein inhibitors of proteinases. *Annu Rev Biochem.* 1980; 49:593–626. [PubMed: 6996568]
32. Huang M, Syed R, Stura EA, Stone MJ, Stefanko RS, Ruf W, Edgington TS, Wilson IA. 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.* 1998; 275:873–94. [PubMed: 9480775]
33. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol.* 2007; 372:774–97. Epub 2007 May 13. [PubMed: 17681537]
34. Riedl SJ, Renatus M, Schwarzenbacher R, Zhou Q, Sun C, Fesik SW, Liddington RC, Salvesen GS. Structural basis for the inhibition of caspase-3 by XIAP. *Cell.* 2001; 104:791–800. [PubMed: 11257232]
35. Cicardi M, Levy RJ, McNeil DL, Li HH, Sheffer AL, Campion M, Horn PT, Pullman WE. Ecallantide for the treatment of acute attacks in hereditary angioedema. *N Engl J Med.* 2010; 363:523–31. [PubMed: 20818887]
36. Oberst MD, Johnson MD, Dickson RB, Lin CY, Singh B, Stewart M, Williams A, al-Nafussi A, Smyth JF, Gabra H, Sellar GC. Expression of the serine protease matrilysin and its inhibitor HAI-1 in epithelial ovarian cancer: correlation with clinical outcome and tumor clinicopathological parameters. *Clin Cancer Res.* 2002; 8:1101–7. [PubMed: 11948120]

37. Darragh MR, Schneider EL, Lou J, Phojanakong PJ, Farady CJ, Marks JD, Hann BC, Craik CS. Tumor detection by imaging proteolytic activity. *Cancer Res.* 2010; 70:1505–12. Epub 2010 Feb 9. [PubMed: 20145119]
38. Szabo R, Rasmussen AL, Moyer AB, Kosa P, Schafer JM, Molinolo AA, Gutkind JS, Bugge TH. c-Met-induced epithelial carcinogenesis is initiated by the serine protease matriptase. *Oncogene.* 2011; 30:2003–16. Epub 2011 Jan 10. [PubMed: 21217780]
39. Duriseti S, Goetz DH, Hostetter DR, LeBeau AM, Wei Y, Craik CS. Antagonistic anti-urokinase plasminogen activator receptor (uPAR) antibodies significantly inhibit uPAR-mediated cellular signaling and migration. *J Biol Chem.* 2010; 285:26878–88. Epub 2010 May 25. [PubMed: 20501655]
40. Shier-Neiss G, Montenecourt BS. Characterization of the secreted cellulases of *Trichoderma reesei* wild type and mutants during controlled fermentations. *Applied Microbiology and Biotechnology.* 1984; 20:46–53.
41. Ilmen M, Saloheimo A, Onnela ML, Penttilä ME. Regulation of cellulase gene expression in the filamentous fungus *Trichoderma reesei*. *Appl Environ Microbiol.* 1997; 63:1298–306. [PubMed: 9097427]
42. Williams JW, Morrison JF. The kinetics of reversible tight-binding inhibition. *Methods Enzymol.* 1979; 63:437–67. [PubMed: 502865]
43. Chou T. Relationships between inhibition constants and fractional inhibition in enzyme-catalyzed reactions with different numbers of reactants, different reaction mechanisms, and different types and mechanisms of inhibition. *Mol Pharmacol.* 1974; 10:235–47. [PubMed: 4212316]
44. Otwinowski Z, Minor W, Carter CW. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 1997; 276:307–326.
45. Read RJ. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr D Biol Crystallogr.* 2001; 57:1373–82. Epub 2001 Sep 21. [PubMed: 11567148]
46. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr.* 1994; 50:760–3. [PubMed: 15299374]
47. Evrard GX, Langer GG, Perrakis A, Lamzin VS. Assessment of automatic ligand building in ARP/wARP. *Acta Crystallogr D Biol Crystallogr.* 2007; 63:108–17. Epub 2006 Dec 13. [PubMed: 17164533]
48. Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr D Biol Crystallogr.* 1999; 55:247–55. Epub 1999 Jan 1. [PubMed: 10089417]
49. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins.* 2003; 50:437–50. [PubMed: 12557186]
50. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 1993; 26:283–291.

Highlights

- Specific matriptase inhibitor with reverse binding motif.
- Fully human recombinant antibody A11.
- Antibody-protease crystal structures.
- *T. reesei* expression system generating high antibody yields.

Figure 1. Inhibition of matriptase by identified Fabs. The relative inhibition of matriptase by eight Fabs identified by phage display is shown. The most potent inhibitor A11 was further investigated and fully characterized.

Light Chain																																									
	CDR L1					CDR L2					CDR L3																														
	24	25	26	27	27a	27b	28	29	30	31	32	33	34	50	51	52	53	54	55	56	88	89	90	91	92	93	94	95	95a	95b	96	97									
A11	RASQS-VSSSYLA						GASTRAT					QQRSNWPPGYT																													
S4	SGSSSNIGSNYVS						DNNQRPS					QSRDISQ--YV																													
E2	RASQ--GISSYLA						AASSLQS					QQHGNLP--YT																													
Heavy Chain																																									
	CDR H1					CDR H2					CDR H3																														
	26	27	28	29	30	31	32	33	34	35	35a	50	51	52	52a	53	54	55	56	57	58	59	95	96	97	98	99	100	100a	100b	100c	100d	100e	100f	100g	100h	100i	100j	100k	101	102
A11	GFTTFSSYAMS-						AISGSGGSTY					DLGIAARRFVSGAR---DI																													
S4	SGGTFSSYAIS						GIIPIFGTAN					TFHIRRYRSGYYDKM--DH																													
E2	GFTTFSSYAMS-						AISGSGGSTY					PYLTPQRRGPONVSPFDN																													

Figure 2.

Comparison of the amino acid sequences (Kabat numbering) of the A11, S4, and E2 heavy (H1, H2, H3) and light chain (L1, L2, L3) variable loops.

Figure 3.

Matriptase alanine point mutants and their effect on protease inhibition by A11 and S4. (a) The six matriptase surface loops surround the protease active site, which consists of a binding cleft and the catalytic triad (sticks, in yellow). (b) Matriptase surface map of K_I changes with A11. (c) Matriptase surface map of K_I changes with S4. Point mutants that had minimal effect on protease inhibition are shaded in gray, mutations that had a three to tenfold increase in inhibitor K_I are shaded in pink, and point mutants that increased inhibitor K_I by more than tenfold are shaded in red. Point mutants that decreased inhibitor K_I are shaded in green. The point-mutant/inhibitor K_I values are given in Table 2. Structural figures were generated using PyMOL (Schrodinger LLC).

Figure 4.

Close-up view of the H3 loop in the (a) A11/matriptase and (b) S4/matriptase complexes. The Fab (heavy chain, light pink; light chain, light orange) caps matriptase surface loops (space-filling models) with the 60s loop shaded (red), 37s loop (purple), 140s loop (orange), 220s loop (dark pink), 170s loop (light blue) and 90s loop (green) as indicated. The H3 loops (dark blue) of A11 and S4 are inserted directly into the active site of matriptase (catalytic triad indicated in yellow) while the remaining variable loops (L1, L2 and L3, red; H1 and H2, sky blue) interact with the protease surface loops. Structural figures were generated using PyMOL. (c) Buried surface area as determined by PISA³³ shows the contribution of the heavy and light chains to the antibody-matriptase interactions. The number in parentheses represents the % of the total buried surface area.

Figure 5.

Interaction of the H3 variable loops of A11 and S4 with the matriptase substrate-binding pocket. (a) The H3 loop of A11 (blue) accounts for the majority of the buried surface area contributed by the heavy chain variable loops. The loop inserts ArgH100b into the active site at the S1 pocket while making very few additional contacts. A11 is superimposed with chloromethylketone inhibitor D-FPR-cmk (purple, PDB code 2FIR). (b) The H3 loop of S4 (green) inserts into the matriptase substrate-binding pocket and makes contacts with a number of the matriptase residues. S4 is superimposed with D-FPR-cmk (purple). (c) ArgH100b makes a water-mediated hydrogen bond with Asp189. Benzamidine (red, ball and stick, PDB code 1EAX) bound to matriptase adapts the expected binding orientation for an arginine containing substrate, forming a salt bridge with Asp189 at the bottom of the S1 pocket of matriptase (distance of 3.1 Å). H3 loop of A11 in blue and H3 loop of S4 in green. (d) Residue ArgH99 of S4 is a critical binding residue for matriptase inhibition and ArgH100b of S4 makes the same water-mediated hydrogen bond with Asp189 as A11.

Figure 6. Cartoon model highlighting the reverse orientation of the H3 loops of A11 and S4 at the matriptase active site. (a) BPTI adopts a substrate orientation at the matriptase active site. By comparison, the reverse orientation of the (b) H3 loop of A11 and (c) H3 loop of S4 in the matriptase active site provides an additional factor to prevent cleavage of the H3 loop. (d) H3 loop of E2 also adopts a substrate orientation at the matriptase active site. The catalytic serine residue (OH) and the matriptase substrate pockets (S4, S3, S2, S1, S1', S2') are also shown in the schematic.

Figure 7.

IC_{50} plots for the A11, S4, and E2 circularized H3 peptides against matriptase. (a) IC_{50} of circularized A11 H3 peptide at time zero (red) and 24 hours (blue). The IC_{50} increases over time due to digestion of the peptide. (b) IC_{50} of circularized S4 H3 peptide at time zero (red) and 24 hours (blue). The IC_{50} improves modestly over time. (c) IC_{50} of circularized E2 H3 peptide at time zero (red) and 24 hours (blue). The IC_{50} remains the same.

Table 1

K_I values for A11 and S4 point mutants against matriptase.

	K_I (nM)	Fold difference
A11	0.72	
A11 H R100aA	1.5	2
A11 H R100bA	180	250
A11 H R100bK	180	250
S4	1	
S4 H R100bA	93	93
S4 H R100bK	110	110

Table 2Apparent K_I values for A11 and S4 with matriptase surface mutants.

	A11		S4 ¹⁹	
	$K_{I,app}$ (pM)	Fold difference	$K_{I,app}$ (pM)	Fold difference
wt-Matriptase	392		70.4	
Q38A	383	0.98	73.6	1
I41A	989	2.5	208	3
I60A	2055	5.2	40.2	0.57
D60aA	2176	5.6	125	1.8
D60bA	1369	3.5	427	6.1
R60cA	172	0.44	11.7	0.17
F60eA	1053	2.7	102	1.4
R60fA	615	1.6	88.9	1.3
Y60gA	942	2.4	151	2.1
R87A	592	1.5	54.5	0.77
F94A	280	0.71	1036	15
N95A	1590	4.1	108	1.5
D96A	> 1uM	10000	897	13
F97A	> 1uM	10000	154	2
T98A	226	0.58	239	3.4
H143A	369	0.94	1671	24
Q145A	336	0.86	116	1.6
Y146A	3119	8	1405	20
T150A	206	0.53	94.6	1.3
L153A	563	1.4	116	1.6
E169A	1376	3.5	199	2.8
Q174A	614	1.6	63.7	0.9
Q175A	711	1.8	246	3.5
D217A	98658	252	838	12
Q221aA	289	0.74	65.7	0.93
R222A	744	1.9	61.1	0.87
K224A	2052	5.2	59.1	0.84

 K_I is calculated from the IC₅₀ value; all errors < 6%.

Table 3

Summary of crystallographic information, 3NPS and 3SO3

<i>Data collection and processing</i>	S4/matriptase 3NPS	A11/matriptase 3SO3
Number of crystals used	1	1
Wavelength (Å)	1.000	1.000
Space group	P2 ₁	P6 ₄
Unit cell parameters		
a, b, c (Å)	39.2, 84.0, 101.4	130.6, 130.6, 96.9
α, β, γ (°)	90.0, 91.5, 90.0	90.0, 90.0, 120.0
Matthews coefficient (Å ³ /Da)	2.27	3.23
Solvent content (%)	45.9	61.9
Molecules per asymmetric unit	3	3
Beamline	ALS 8.3.1	ALS 8.3.1
<i>Diffraction Data</i>		
Resolution range (Å)	43.4 - 1.4 (1.46 - 1.40)	113.0 - 2.1 (2.21 - 2.10)
Unique reflections	104800 (2150)	54872 (7998)
<i>R</i> (I) _{sym} (%) ^a	5.8 (17.6)	9.4 (93.2)
Completeness (%)	84.5 (80.0)	100.0 (100.0)
Redundancy	3.7 (3.1)	7.3 (6.9)
<i>I</i> /σ(<i>I</i>)	16.0 (8.0)	11.7 (2.2)
<i>Refinement</i>		
Resolution range (Å)	28.6 - 1.5	19.9 - 2.1
Reflections used in refinement (work/free)	88531/4447	51923/2774
Final <i>R</i> values for all reflections (work ^b / free ^c) (%)	18.8 / 22.8	16.1 / 19.4
No. atoms		
Proteins	5157	5109
Ligands / ion ^d	46	35
Water	676	462
<i>RMSDs</i>		
Bonds (Å)	0.014	0.017
Angles (°)	1.61	1.64
<i>Ramachandran parameters (%)</i> ^e		
Residues in most favored regions	89.3	91.6
Residues in additional allowed regions	10.1	8.3
Residues in generously allowed regions	0.2	0.0
Residues in disallowed regions	0.4	0.2
<i>Mean B-factor (Å²)</i>		
Protein	19.8	59.6
Ligand / ion	48.3	57.7
Water	31.4	63.4

Numbers in parentheses represents the highest-resolution shell.

ALS, Advanced Light Source

^a $R_{sym} = (\sum |I - \langle I \rangle|) / \sum I$, where $\langle I \rangle$ is the average intensity of multiple measurements

^b $R_{work} = (\sum |F_{obs} - F_{calc}|) / \sum |F_{obs}|$

^c $R_{free} = R_{work}$ based on ~ 1000 (at least 5%) of reflections excluded from refinement

^d Ligands/ion are ethylene glycol, Na, Cl in S4/matriptase and glycerol, sucrose in A11/matriptase

^e Calculated using PROCHECK⁵⁰