THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 291, NO. 27, pp. 14340–14355, July 1, 2016

Crystal Structure of a Two-domain Fragment of Hepatocyte Growth Factor Activator Inhibitor-1

*FUNCTIONAL INTERACTIONS BETWEEN THE KUNITZ-TYPE INHIBITOR DOMAIN-1 AND THE NEIGHBORING POLYCYSTIC KIDNEY DISEASE-LIKE DOMAIN**

Received for publication, November 30, 2015, and in revised form, April 15, 2016 Published, JBC Papers in Press, May 6, 2016, DOI 10.1074/jbc.M115.707240

Zebin Hong[‡], Laura De Meulemeester[‡], Annemarie Jacobi[‡], Jan Skov Pedersen[§], J. Preben Morth[¶], Peter A. Andreasen[‡], and Jan K. Jensen[‡1]

From the [‡]Department of Molecular Biology and Genetics, Danish-Chinese Centre for Proteases and Cancer, Aarhus University, Gustav Wieds Vej 10C, 8000 Aarhus C, Denmark, the [§]Department of Chemistry and iNANO Interdisciplinary Nanoscience Center, Aarhus University, 8000 Aarhus C, Denmark, and the [¶]Norwegian Center of Molecular Medicine (NCMM), University of Oslo, NO-0316 Oslo, Norway

Hepatocyte growth factor activator inhibitor-1 (HAI-1) is a type I transmembrane protein and inhibitor of several serine proteases, including hepatocyte growth factor activator and matriptase. The protein is essential for development as knockout mice die *in utero* due to placental defects caused by misregulated extracellular proteolysis. HAI-1 contains two Kunitz-type inhibitor domains (Kunitz), which are generally thought of as a functionally self-contained protease inhibitor unit. This is not the case for HAI-1, where our results reveal how interdomain interactions have evolved to stimulate the inhibitory activity of an integrated Kunitz. Here we present an x-ray crystal structure of an HAI-1 fragment covering the internal domain and Kunitz-1. The structure reveals not only that the previously uncharacterized internal domain is a member of the polycystic kidney disease domain family but also how the two domains engage in interdomain interactions. Supported by solution small angle x-ray scattering and a combination of site-directed mutagenesis and functional assays, we show that interdomain interactions not only stabilize the fold of the internal domain but also stimulate the inhibitory activity of Kunitz-1. By completing our structural characterization of the previously unknown N-terminal region of HAI-1, we provide new insight into the interplay between tertiary structure and the inhibitory activity of a multidomain protease inhibitor. We propose a previously unseen mechanism by which the association of an auxiliary domain stimulates the inhibitory activity of a Kunitz-type inhibitor (*i.e.* the first structure of an intramolecular interaction between a Kunitz and another domain).

Hepatocyte growth factor activator inhibitor-1 (HAI-1)[2] is a ~53-kDa multidomain type I transmembrane glycoprotein and inhibitor of several serine proteases. HAI-1 has been shown to be essential for placental development by regulating extracellular proteolytic activities (1), and HAI-1-deficient mice die *in utero* due to undifferentiated chorionic trophoblasts (1–3). HAI-1 plays an important part in maintaining postnatal tissue homeostasis, including keratinization of the epidermis, hair development, and colonic epithelium integrity (4–6). Also, mouse genetic studies have now convincingly established that improper inhibition of matriptase activity by HAI-1 induces carcinogenesis (7). HAI-1 was initially identified as an inhibitor co-purifying with hepatocyte growth factor activator (HGFA) (8) and was later purified from human milk in complex with the type II transmembrane serine protease matriptase (9). The physiological importance of the HAI-1-matriptase relationship was cemented in genetic mouse studies, where simultaneous ablation of matriptase rescues the placental defect caused by a lack of HAI-1 (3). It has also been proposed that HAI-1 inhibits the glycophosphatidylinositol membrane-anchored serine protease prostasin (10). Mouse studies also support a tight regulatory relationship between matriptase HAI-1 and prostasin (11). *In vitro* inhibition of additional membrane-associated serine proteases hepsin (12), TMPRSS13 (13), HAT (14), and HATL5 (15) as well as the soluble proteases trypsin, plasmin, and plasma kallikrein have also been reported (16).

The mature HAI-1 protein consists of, from the N terminus, a MANEC domain; a region proposed to contain at least one domain termed the "internal domain"; the first of two Kunitz-type inhibitor domains (Kunitz-1); a low density lipoprotein receptor (LDLR) class A domain; Kunitz-2; a single-pass transmembrane region; and finally a short C-terminal cytoplasmic tail (Fig. 1, *A* and *B*). HAI-1 also contains three potential *N*-glycosylation sites, two of which are located in the extracellular region at asparagine 66 and 235, in the MANEC domain and internal domain, respectively (8, 17). The primary inhibitory reactivity of HAI-1 toward all known target proteases has been shown to reside exclusively in Kunitz-1 (16–18).

Several soluble truncated forms of HAI-1, including matriptase·HAI-1 complexes, have been identified in human

* This work was supported by the Danish Cancer Society. The authors declare that they have no conflicts of interest with the contents of this article.
The atomic coordinates and structure factors (code 5EZD) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] To whom correspondence should be addressed. Tel.: 45-87-15-55-42; E-mail: jkj@mbg.au.dk.
[2] The abbreviations used are: HAI-1, hepatocyte growth factor activator inhibitor-1; HGFA, hepatocyte growth factor activator; MANEC, motif at N terminus with eight cysteines; Kunitz, Kunitz-type inhibitor domain; LDLR, low density lipoprotein receptor; PKD, polycystic kidney disease; Ni-NTA, nickel-nitrilotriacetic acid; SPR, surface plasmon resonance; IK1, HAI-1 protein fragment consisting of internal (N235Q) domain and Kunitz-1; SAXS, small angle x-ray scattering; PDB, Protein Data Base.

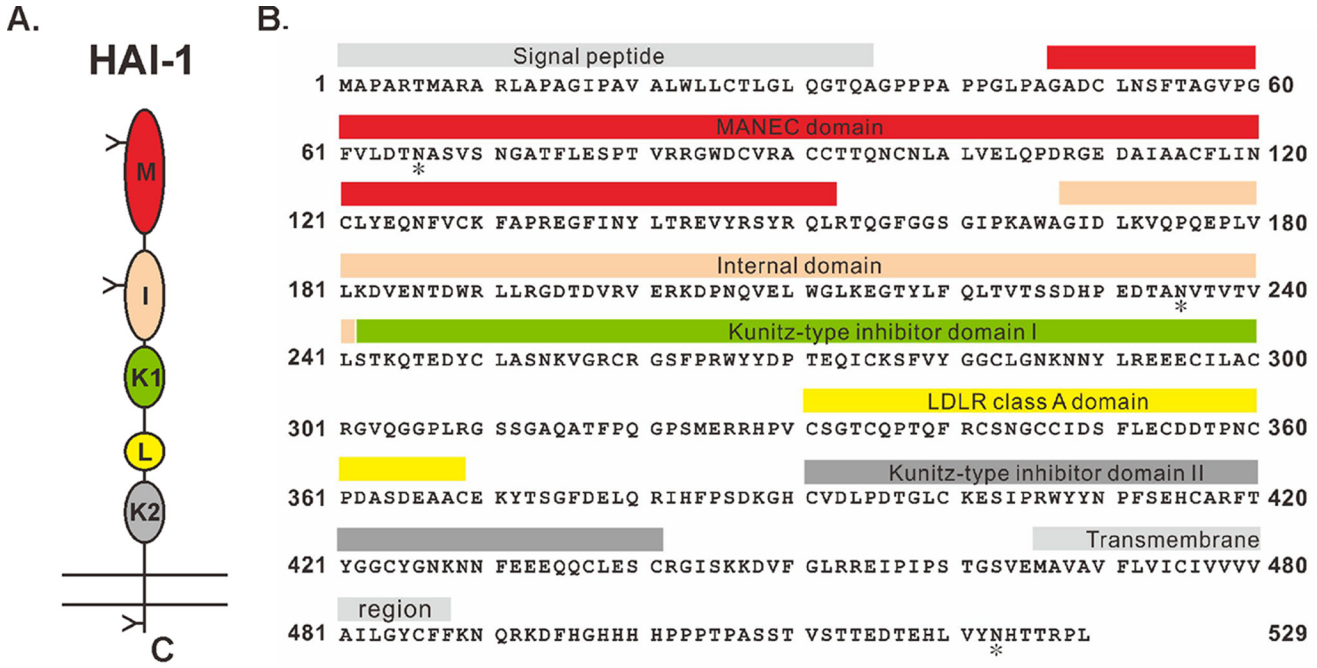

FIGURE 1. **Schematic and protein sequence presentation of HAI-1.** *A*, the human HAI-1 protein is depicted as attached to the cell membrane by the C-terminal single-pass transmembrane region. Domains and linkers are drawn to scale according to predicted and confirmed domain sequences (see *B*). Domain labels are as follows: MANEC (*M*), internal domain (*I*), Kunitz-1 (*K1*), LDLR class A domain (*L*), and Kunitz-2 (*K2*). *Y*, positions of the three potential *N*-glycosylation sites. *B*, HAI-1 full-length protein sequence with domain locations. The signal sequence, domain architecture, and transmembrane region are shown as *labeled boxes*, with *colors* matching the schematic presentation in *A*, *above* the corresponding amino acid residues (isoform 1, Uniprot accession number O43278-1). The locations of the signal sequence, Kunitz-type inhibitor domains, LDLR class A domain, and the transmembrane region are as reported previously (8). The location of the MANEC is as reported (22). The internal domain sequences correspond to that determined in this study. *, positions of the three potential *N*-glycosylation sites.

milk and conditioned medium. These must result from different proteolytic cleavages of HAI-1 following its translocation to the plasma membrane, catalyzed by yet unknown proteases (19–21). The different truncated soluble forms of HAI-1 were shown to have different inhibitory activities against the natural targets matriptase and HGFA (17, 21). In a pioneering study concerning the inhibitory activity of different truncated forms of rat HAI-1, it was further hypothesized that the non-inhibitory domains (*i.e.* MANEC, internal domain, the LDLR class A domain, and Kunitz-2) are all important for the regulation of the inhibitory activity of HAI-1, most likely by participating in intramolecular interdomain interactions resulting in different tertiary structures with corresponding varying inhibitory activities (17). As a first step to provide structural evidence of how the non-inhibitory domains of HAI-1 may contribute to the overall function of the protein, we recently completed a structural characterization of the N-terminal MANEC domain, defining it as a new subclass of the PAN/apple family of domains with numerous examples of regulatory properties as mediators of inter- and intramolecular protein-protein interactions (22).

The purpose of the present study was to complete the structural characterization of the N-terminal region of HAI-1, from the proposed internal domain through Kunitz-1. Our x-ray crystal structure analysis revealed that the internal domain belongs to the family of polycystic kidney disease (PKD)-like domains. More importantly, the structure presented a condensed v-shaped tertiary structure formed by interdomain interactions between the internal domain and Kunitz-1. We show that the association of the internal domain stimulates the inhibitory activity of Kunitz-1 and that the effect is independent of the target protease. Moreover, the structure provides not only a tool for structure-guide studies of the internal domain of HAI-1 but also the first structural evidence of a functional role for a non-inhibitory domain of HAI-1. Thus, we propose a new mechanism in which the association of an auxiliary domain stimulates the inhibitory capacity of a Kunitz-type inhibitor (*i.e.* the first structure of an intramolecular interaction between a Kunitz and another domain). Our biochemical analysis also illuminated the role of the 60-loop in matriptase as a rate-limiting factor for macromolecular access to the active site. Finally, our structure revealed the location of a potential integrin binding site in the internal domain with potential implications for the cell surface localization of HAI-1.

Experimental Procedures

Recombinant Protein Production

Pichia pastoris—To express protein with a well defined N terminus, a Kex2 cleavage site was inserted between the α-factor signal peptide and the protein sequence. The PCR product was subcloned into the pPICZαA expression vector (Invitrogen) using the XhoI and SalI restriction sites and linearized by SacI digestion before transformation into *P. pastoris* X-33 strain (Invitrogen). Protein-producing clones were stored at −80 °C in 15% glycerol. Protein production followed the manufacturer's recommendations (Invitrogen). The HAI-1 fragments internal domain (Gly^{168}–Tyr^{249}), IK1 (Gly^{161}–Val^{303},

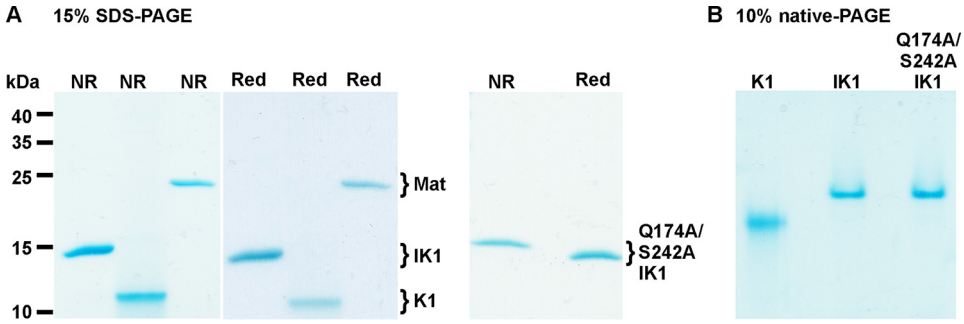

FIGURE 2. **SDS-PAGE analysis of the purified proteins.** *A*, a 15% SDS-PAGE analysis of 2 μg of K1, IK1, Q174A/S242A IK1, and catalytic domain of matriptase (*Mat*), respectively. HAI-1 fragments were produced in *P. pastoris*, whereas the catalytic domain of matriptase was purified from *E. coli* lysates and refolded, and the active form was affinity-purified on benzamidine-Sepharose. All samples were analyzed both under non-reducing (*NR*) and reducing (*Red*) conditions. The identities of the protein bands are given as *labels* on the *right side* of the corresponding gel figure. The migration of a prestained protein marker mixture (Fermentas) is indicated to the *left*. *B*, 10% native PAGE of 2 μg of K1, IK1, and Q174A/S242A IK1, respectively.

Gly^{168}–Val^{303}), and K1 (Gln^{245}–Val^{303}) were amplified from human cDNA of full-length HAI-1 splice variant A (Uniprot accession number O43278-1). The N235Q substitution was generated by site-directed mutagenesis to remove the potential *N*-glycosylation site within the C terminus of the internal domain. Due to the expression tags, the resulting purified HAI-1 fragments contain a Lys-Arg extension in the N terminus and an Asp followed by His_6 in the C terminus.

To produce the catalytic domain of matriptase, the DNA sequence covering Val^{615}–Val^{855} was amplified from the human matriptase cDNA (Uniprot accession number Q9Y5Y6). The deletion of the 60-loop (chymotrypsin template numbering) was accomplished by substitution of DDRGFR (Asp^{660}–Arg^{665}) with GG using site-directed mutagenesis. To introduce the StrepII tag in the C terminus of the matriptase catalytic domain, the StrepII tag (SAWSHPQFEK) was fused to the C terminus with an additional $(GS)_3$ linker.

Escherichia coli—For bacterial production of the internal domain, a purified pT7-I vector containing the human HAI-1 fragment Gly^{168}–Tyr^{249} was transformed into competent BL21 (DE3). Cells were cultured at 37 °C in 2× TY medium containing 100 μg/ml ampicillin. Expression of matriptase was induced by 1 mM isopropyl 1-thio-β-D-galactopyranoside for 4 h. Cells were collected by centrifugation at 5000 × *g* for 20 min.

For the bacterial production of the catalytic domain of matriptase, a pT7 vector containing human matriptase (Gly^{596}–Val^{855}) was transformed into competent Rosetta BL21 (DE3), following the protocol above with the addition of 34 μg/ml chloramphenicol.

Recombinant Protein Purification

Purification and Refolding of the Internal Domain from E. coli—Cell pellets were resuspended in sonication buffer (50 mM Tris-HCl, pH 8.0, 0.15 M NaCl) and sonicated on ice. The filtrated supernatant was incubated with nickel-nitrilotriacetic acid (Ni-NTA)-agarose containing 20 mM imidazole that had been equilibrated in equilibrium buffer (50 mM Tris-HCl, pH 8.0, 0.15 M NaCl) by slow stirring on ice for 1 h. The bound protein was eluted with 400 mM imidazole in equilibrium buffer and dialyzed extensively against equilibrium buffer at 4 °C. A protein purity of >95% was verified by SDS-PAGE (data not shown). Protein concentrations were determined by $A_{280\,nm}$ using the extinction coefficient 18,350 M^{-1} cm^{-1} (ProtParam).

For the refolding trials, the internal domain protein was denatured by 6 M urea and adjusted to 0.1–0.2 mg/ml. The protein sample was dialyzed overnight at 4 °C against refolding buffer (50 mM Tris-HCl, pH 8.0, 10% glycerol) and concentrated in 3.5 kDa cut-off spin columns (Amicon).

Purification of the Internal Domain, IK1, and K1 from P. pastoris—Cleared, filtrated medium from yeast was loaded onto an Ni-NTA-agarose column equilibrated with 20 mM Tris-HCl, pH 7.4, 150 mM NaCl. The column-bound protein was washed with equilibration buffer supplemented with 20 mM imidazole and eluted with 300 mM imidazole. The eluted protein fraction was dialyzed against 20 mM Tris-HCl, pH 7.4, 150 mM NaCl and concentrated before size exclusion chromatography (Superdex75 column, GE Healthcare) in the same buffer. Protein fractions were collected, pooled and dialyzed against 20 mM Tris-HCl, pH 7.4, 50 mM NaCl. A protein purity of >95% was verified by SDS-PAGE under reducing and non-reducing conditions (Fig. 2*A*). Protein folding homogeneity of the proteins was verified by migrating as a single band on native PAGE (Fig. 2*B*). Protein concentrations of IK1 and K1 were determined by $A_{280\,nm}$ using the extinction coefficients 13,940 and 12,090 M^{-1} cm^{-1} (ProtParam) and verified by amino acid analysis (in-house service).

Purification of the Catalytic Domain of Matriptase (Wild Type), Matriptase (Δ60-Loop), and Matriptase (C-terminally StrepII-tagged) from P. pastoris—Cleared, filtrated medium from the yeast was loaded onto a benzamidine column (GE Healthcare) equilibrated with 20 mM Tris-HCl, pH 7.4, 150 mM NaCl and washed with equilibration buffer. Bound protein was eluted by 10 mM glycine-HCl, pH 2.5, and neutralized by 10% (v/v) 1.0 M Tris-HCl, pH 9.0. A protein purity of >95% was verified by SDS-PAGE. Proteins containing fractions were pooled, quantified by $A_{280\,nm}$ using the extinction coefficient 53,200 M^{-1} cm^{-1} (ProtParam), and stored at −80 °C.

Purification and Refolding of the Catalytic Domain of Matriptase (Wild Type) from E. coli—Cell pellets were resuspended in sonication buffer containing 50 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 10% glycerol, 1 mM β-mercaptoethanol, and 1 mM EDTA and sonicated on ice. Inclusion bodies were pelleted at 10,000 × *g*, washed repeatedly using wash buffer 1 (50 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 10% glycerol, 1 mM β-mercaptoethanol, 1 mM EDTA, 1% (v/v) Triton X-100), wash buffer 2

(wash buffer 1 with 0.25% (v/v) Triton X-100), and wash buffer 3 (wash buffer 1 without Triton X-100), each step followed by centrifugation. The insoluble matter was resuspended in denaturing buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10 mM β-mercaptoethanol, 1 mM EDTA, and 6 M urea) by slowly stirring on ice for 30 min. The suspension was cleared by centrifugation at 10,000 × g, and the supernatant was incubated with Ni-NTA-agarose in the presence of 20 mM imidazole under slow stirring on ice for 1 h. Bound protein was eluted with 400 mM imidazole in denaturing buffer and dialyzed extensively against denaturation buffer at room temperature. The protein concentration was adjusted to 0.1–0.2 mg/ml using denaturation buffer and dialyzed overnight at 4 °C against refolding buffer (50 mM Tris, pH 8.0, 10% glycerol, 1 mM β-mercaptoethanol, and 3 M urea). Complete autoactivation of the protease preparation was observed as a complete band shift on SDS-PAGE (data not shown), due to the loss of N-terminal residues 596–614 following activation cleavage, during the second dialysis step against storage buffer (50 mM Tris-HCl, pH 8.0, 10% glycerol) at 4 °C overnight. The active protease was captured on a benzamidine column equilibrated with storage buffer and eluted with storage buffer supplemented with 1 M arginine. Proteins were further purified by size exclusion chromatography on a Superdex75 column equilibrated with storage buffer. Proteins containing fractions were pooled, quantified by $A_{280\,nm}$, and stored at −80 °C. A protein purity of >95% was verified by SDS-PAGE (Fig. 2*A*). The *E. coli* produced matriptase catalytic domain behaved indistinguishably from the similar protein produced in *P. pastoris* and was used exclusively for the K_i determinations due to unrestricted availability. Hepsin and plasma kallikrein were prepared using a similar strategy.

Crystallography

Protein Crystallization—Protein samples of IK1(N235Q) and IK1(N235Q) extended N-terminally to Gly^{161} were concentrated to 11 mg/ml using 3.5 kDa cut-off spin columns (Amicon). More than 1000 premade crystallization conditions (Molecular Dimensions Ltd.) were screened in a 96-well plate sitting drop setup (Molecular Dimensions Ltd.) using robotics (TTP Labtech's mosquito Crystal) with 200 + 200-nl and 200 + 100-nl drops. Diffraction quality crystals were grown from two initial screen conditions after transfer to a hanging drop vapor diffusion setup at 20 °C in 24-well plates (Molecular Dimensions Ltd.) against 1 ml of reservoir solution. The IK1(N235Q) crystals appear in 2–3 days after mixing 2 μl of protein solution with 1 μl of reservoir solution containing 28% (w/v) PEG 2000 MME, 0.1 M sodium acetate, pH 5.5, 0.2 M potassium bromide or 25% (w/v) PEG 2000 monomethyl ether, 0.1 M sodium cacodylate, pH 6.5, 0.2 M calcium acetate. For the IK1 with the extended N terminus, crystals appeared in 2–3 days after mixing 1 μl of protein solution with 1 μl of reservoir solution containing 22% (w/v) PEG 6000, 10% (v/v) 2-propanol, 0.1 M sodium acetate trihydrate, pH 4.0.

Data Collection, Structure Model Building, and Validation—Following optimization trials, we obtained protein crystals that diffracted to a resolution of 2.1 Å (Table 1). For x-ray data collection, a single crystal was transferred to a cryoprotectant solution and flash-frozen in liquid nitrogen. The cryoprotectants used were either 15% glycerol, 35% (w/v) PEG 2000 monomethyl ether, 0.1 M sodium acetate, pH 5.5, 0.2 M potassium bromide or, in the case of IK1 with an extended N terminus, 22% (w/v) PEG 6000, 10% (v/v) 2-propanol, 0.1 M sodium acetate trihydrate, pH 4.0. The diffraction data were indexed, integrated, and scaled using the Xia2, XDS, and CCP4 program packages (23–25). The initial phase information was obtained by molecular replacement using PHASER (26) and the available crystal structure of Kunitz-1 (PDB code 4ISN) with an initial R_{free} of ~44%. Including the I-TASSER (27) model of the internal domain as an additional search model resulted in a further improvement of R_{free} to ~40%. All subsequent model building was performed in COOT (28) and refined using Phenix (29). The final model includes two molecules in the asymmetric unit. The A and B chain were kept restrained with a single non-crystallographic symmetry operator throughout the refinement. In addition, TLS was used throughout refinement. The quality of the final model was validated with the wwPDB Validation Server. The final structure coordinates (PDB code 5EZD) and corresponding diffraction data were uploaded to the RCSB Protein Data Bank. In the case of IK1 with the extended N terminus, the IK1 structure was used as the initial model in PHASER (data not shown).

TABLE 1
Crystallographic parameters of IK1

Parameters	Values
PDB code	5EZD
Data acquisition	
Wavelength (Å)	1.000
Resolution (Å)[a]	45.8–2.10 (2.15–2.10)[b]
Space group	P 21 21 21
Unit cell parameters	
a, *b*, *c* (Å)[a]	57.65, 75.14, 75.26
$\alpha = \beta = \gamma$ (degrees)	90.0
Total reflections[a]	104,438 (7672)
Unique reflections[a]	36,267 (2652)
Multiplicity[a]	5.3 (5.4)
Completeness (%)[a]	99.6 (97.5)
Mean $I/\sigma(I)$[a]	22.8 (2.29)
R_{meas}[a,b]	0.038 (0.716)
Wilson *B*-factor (Å^2)	56.8
Refinement	
R_{work}[c] (%)	23.4
R_{free} (%)	26.5
RMSD[d]	
Bond lengths (Å)	0.008
Bond angles (degrees)	1. 4
Ramachandran plot[e]	
Favored (%)	97
Allowed (%)	3

[a] $R_{meas} = \Sigma_{hkl}(N/(N-1))^{1/2}\Sigma_i|I_{i,hkl} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\Sigma_I|I_{i,hkl}|$.
[b] The values in parentheses indicate the shell of highest resolution.
[c] $R_{work} = \Sigma_{hkl}|F_{obs} - F_{calc}|/\Sigma_{hkl}F_{obs}|$.
[d] R_{free} is calculated like R_{work} but with 5% (random subset) reflections excluded from refinement.
[e] Root mean square of deviations from ideal values.

Chromogenic Assay

Determination of Inhibitory Equilibrium Constant (K_i) Values under Pre-steady-state Conditions (Slow Reacting Inhibitor)—To start the enzymatic reaction, 50 pM matriptase catalytic domain, prepared by the *E. coli* protocol, was added to reaction mixtures containing 500 μM chromogenic substrate S2288 (DiaPharma) and varying concentrations of inhibitor. A final reaction volume of 200 μl in 10 mM HEPES-NaOH, pH 7.4,

150 mM NaCl, 1 mM $CaCl_2$ supplemented with 0.1% PEG 8000 was used throughout at 37 °C. Progress curves were analyzed using the pre-steady-state Equation 1, which describes the reaction toward the equilibrium between the enzyme and the inhibitor (30),

$$[P] = v_s t + \frac{v_i - v_s}{k_{obs}}[1 - e^{-k_{obs}t}] \quad \text{(Eq. 1)}$$

where [P] is the concentration of *para*-nitroaniline released by substrate hydrolysis, v_i and v_s are the initial and steady-state velocities, respectively, in the presence of inhibitor, and k_{obs} is the apparent first order rate constant for the interconversion between v_i and v_s.

The mechanism is that of simple reversible binding, determined by the association and dissociation rate constants k_{on} and k_{off}, and the competitive binding with the substrate gives Equation 2,

$$k_{obs} = k_{off} + \frac{k_{on}[I]}{(1 + [S]/K_M)} \quad \text{(Eq. 2)}$$

where [S] is the substrate concentration. A K_m value of 536 ± 13 μM for S2288 conversion by matriptase was determined by Lineweaver-Burk plots as described previously (31). Finally, the inhibitory constant K_i is defined by Equation 3,

$$K_i = \frac{k_{off}}{k_{on}} \quad \text{(Eq. 3)}$$

Determination of K_i Values under Steady State (No Time Dependence)—For determination of K_i values for the inhibition of the various enzymes under steady-state conditions, a fixed concentration of Δ60-loop matriptase (100 pM), HGFA (3 nM active form; Sino Biological Inc.), hepsin (500 pM), and plasma kallikrein (2 nM) was incubated with the concentration series of inhibitors for 10 min in 10 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 1 mM $CaCl_2$, 0.1% PEG 8000 at 37 °C. Chromogenic substrates (500 μM S2288, matriptase, and Δ60-loop matriptase; 125 μM S2302, plasma kallikrein; 200 μM S2366, hepsin (all from DiaPharma); 200 μM Spectrozyme FVIIa, HGFA (Sekisui Diagnostics) were added to start the reaction. The initial reaction velocities were monitored as a change in $A_{405\,nm}$. The K_i values were determined from the non-linear regression analyses of plots of the relative initial rates v_i/v_0 *versus* [I], using Equation 4, derived under the assumption of competitive inhibition (30)

$$\frac{v_i}{v_o} = \frac{K_i(K_m + [S])}{K_i[S] + K_m(K_i + [I])} \quad \text{(Eq. 4)}$$

where [I] is the inhibitor concentration.

Surface Plasmon Resonance (SPR)

All SPR experiments were performed on a BIACORE T200 system (GE Healthcare) at 25 °C. Anti-mouse IgG antibody (GE Healthcare) was coupled covalently to a CM5 sensor surface at 30 μg/ml in 10 mM sodium acetate (pH 5.0) to a level of ~10,000 response units using the standard amine coupling protocol (Biacore, GE Healthcare). First capture, anti-StrepII mouse antibody (10 μg/ml) was captured on the sensor surface in running buffer (10 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 1 mM $CaCl_2$, 0.05% Tween 20) to a level of ~700 response units at a flow rate of 5 μl/min. Second capture, C-terminal StrepII-tagged wild type or Δ60-loop matriptase produced in *P. pastoris* was captured directly from the medium on the sensor surface in running buffer to a level of ~60 response units. Concentration series of the IK1 or K1 protein preparations were injected at a flow rate of 60 μl/min. The sensor chip surface was regenerated with a 3-min pulse of 10 mM glycine-HCl, pH 1.7, according to the kit protocol (Biacore, GE Healthcare). To correct for buffer effects, the response from a reference surface with no matriptase capture was subtracted. The data from all SPR experiments could be fit to a 1:1 binding model in the Biacore evaluation software. The estimated kinetic parameters (k_{on} and k_{off}) were used to determine the corresponding dissociation constant (K_D) from the relationship $K_D = k_{off}/k_{on}$.

SAXS Data Collection and Analysis

The SAXS data were collected on the laboratory-based instrument at the Interdisciplinary Nanoscience Center, Aarhus University (32). Samples were placed in reusable quartz capillaries at a controlled temperature of 25 °C. The acquisition time was 3600 s for a protein sample of 1.0 mg/ml in 10 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, and supplemented with 10 mM L-tryptophan to reduce aggregation. Corresponding buffer data were collected. Background buffer subtraction and conversion of the data to absolute scale by use of water as a primary standard was performed using the SUPERSAXS program package.[3] The instrumental sample-to-detector distance was set to 640 cm, giving a q range of 0.01–0.345 $Å^{-1}$, where q is the length of the scattering vector, defined as $q = 4\pi\sin(\theta)/\lambda$, where λ is the x-ray wavelength at 1.54 Å and 2θ is the scattering angle between the incident and scattered beam. The final intensity is $I(q)$, in units of cm^{-1}. Indirect Fourier transformation analysis was performed using the approach by Glatter *et al.* (33) implemented in the WIFT program (34) to obtain the pair distance distribution function, $p(r)$, along with the characteristic parameters: the maximum diameter (D_{max}), the radius of gyration (R_g), and the forward scattering, $I(q = 0)$. The theoretical scattering profile of the crystal structure was computed and compared with the experimental data using the program CRYSOL (35). To obtain a structural model of the scattering molecule in solution without the use of any assumptions other than a starting value for D_{max}, *ab initio* modeling was performed using the program DAMMIF (36). To generate an *ab initio* envelope, a minimum of 10 calculated DAMMIF solutions were compared, aligned, averaged, and filtered using the subroutines damsel, damsup, damaver, and damfilt, respectively, from the DAMAVER program package (37). Finally, our crystal structure was directly compared with the *ab initio* SAXS model using the program SUPCOMB from the same program package (38). To optimize the crystal structure model to fit the solution SAXS data better, the program BUNCH (39) was used. The four residues (Val^{240}–Thr^{243}) that link the two domains of the crystal structure model were deleted and replaced by four

[3] J. S. Pedersen and C.L.P. de Oliveira, unpublished data.

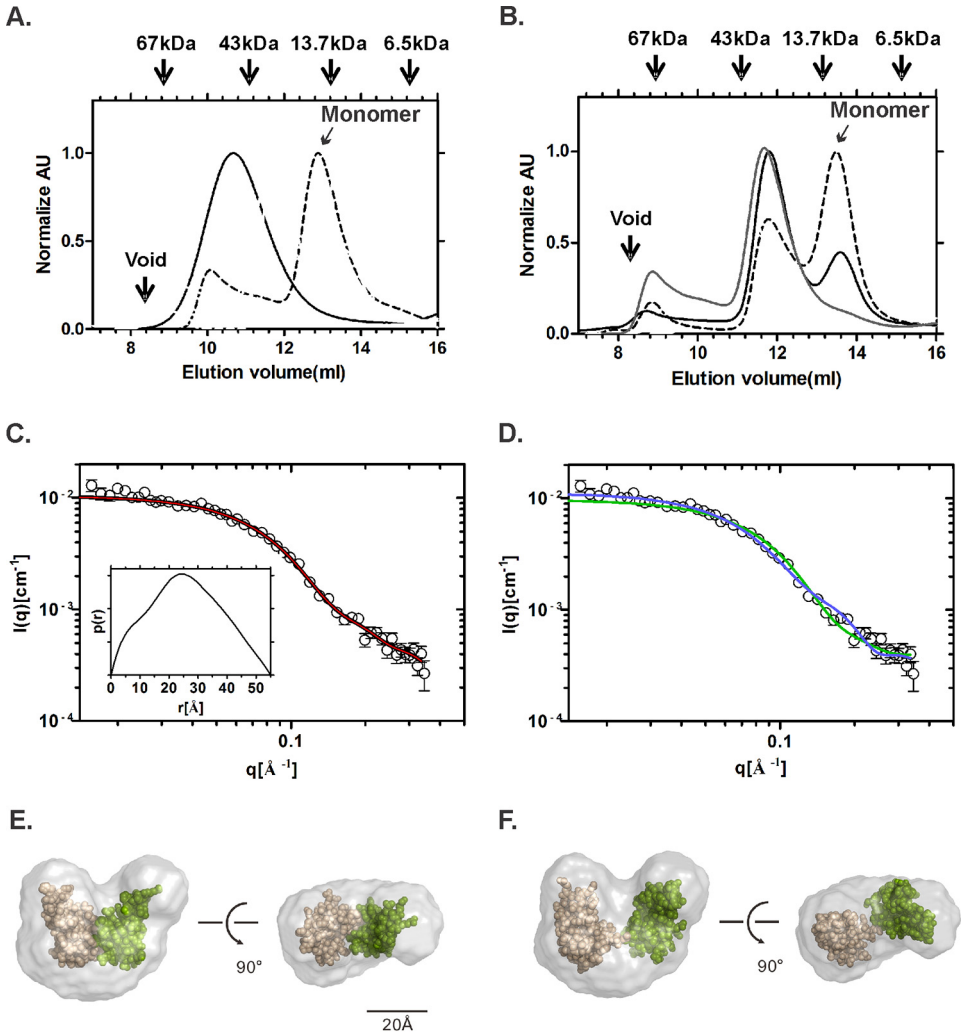

FIGURE 3. **Initial characterization of IK1.** *A*, elution profiles from a Superdex75 size exclusion column of the internal domain alone produced using *P. pastoris* (*solid black line*) and IK1 produced using *P. pastoris* (*broken black line*). *B*, elution profiles from a Superdex75 size exclusion column of the internal domain alone produced using *E. coli* (*solid gray line*), immediately following refolding (*broken black line*) and refolded material stored >1 day at 4 °C (*solid black line*). *C*, the SAXS data on the IK1 monomer produced in yeast (*black circles*) with corresponding IFT fit (*solid black line*) and fit of the *ab initio* model computed with DAMMIF (*solid red line*). The small *inset* shows the computed distance distribution function $p(r)$ of IK1 obtained from GNOM (54) with a D_{max} of 55 Å. *D*, the SAXS data with theoretical scattering curves of our IK1 crystal structure (*solid green line*) and reference BUNCH structure (*solid blue line*) as calculated by CRYSOL. *E*, the combined volume of the *ab initio* molecule shapes of the IK1 protein is shown in a *gray-colored surface presentation*. The *ab initio* shape was constructed from the experimental SAXS data using the program DAMMIF and superimposed with the IK1 crystal structure, represented by *spheres* using SUPCOMB. The internal domain and Kunitz-1 are *colored wheat* and *green*, respectively. *F*, combined volume of *ab initio* shapes as in *E* superimposed with the rigid body modeling reference structure calculated using BUNCH with a four-residue (Val^{240}–Thr^{243}) flexible interdomain linker. The crystal structure figures were prepared using PyMOL.

dummy residues to allow for flexibility of the structure. The program SUPCOMB was used for selecting the most representative model from 10 runs.

Results

Kunitz-1 Is Required for Solution Stability of the Internal Domain—Standard online protein sequence search engines, such as BLAST and SMART, failed to provide any predictive information on the location, extent, or three-dimensional structure of the internal domain of HAI-1. Based on our recent structure of the MANEC domain (22) and the two existing structures of Kunitz-1 in complex with HGFA and matriptase (40, 41), the sequence gap between the MANEC and Kunitz-1 (Arg^{153}–Tyr^{249}) was analyzed by the more advanced sequence-based homology search server, I-TASSER (Fig. 5*B*). From I-TASSER, a weak sequence homology was picked up to an NMR structure of the fourth PKD domain from the uncharacterized human protein KIAA1837 (PDB code 2YRL) to a region covering Gly^{168} to Thr^{243} in HAI-1 (for sequence reference, see Fig. 1*B*). To allow for structural and functional characterization of the proposed internal domain, the protein fragment was produced at high (>10 mg/liter) yield from both bacterial and yeast cultures and agreed well with the expected mass (9.4 kDa). Although the purified protein from both sources was soluble at high concentration under standard buffer conditions, size exclusion chromatography suggested protein aggregation with a prevailing molecular mass >30 kDa (Fig. 3, *A* and *B*). The behavior was irrespective of glycosylation status (data not shown). Although it was possible to refold a predominantly monomeric form from the *E. coli* material, the refolded mono-

mer fraction reverted back to higher order aggregates in a time-dependent manner as monitored by repeated analytical size exclusion chromatography (Fig. 3*A*). Revisiting the predicted I-TASSER structure, a weakly associated N-terminal extension to Gly^{161} can be seen. Unfortunately, inclusion of the seven additional residues had no effect on the propensity of aggregation (data not shown). Due to the inherent instability of the proposed internal domain, we decided to investigate it in the context of its neighboring domains. The MANEC-internal domain construct was unsuccessful in producing detectable amounts of protein in yeast cultures, although a similar approach had successfully produced the isolated MANEC (22). However, a construct covering the internal domain and Kunitz-1 (IK1) expressed well in yeast and the purified protein eluted from repeated size exclusion chromatography experiments as a stable monomer (Fig. 3*B*).

To verify the monomeric status of the protein in solution, we subjected the IK1 sample to SAXS analysis. From forward scattering ($I(q) = 0$), calculated from the direct Fourier transform of the SAXS data (Fig. 3*C*), a molecular mass of the IK1 was determined to 17 ± 2 kDa, confirming the monomeric state of IK1 in solution with a theoretical molecular mass of 16.8 kDa. From the same analysis, a D_{max} (the longest observed distance within the molecule) of 55 ± 5 Å and a radius of gyration of 25.1 ± 1.5 Å were calculated. A classical Guinier plot (not shown) displayed a linear behavior of the low q data and gave a similar value. The dimensions clearly suggested a compact IK1 molecule when compared with the dimensions of the available Kunitz-1 structure (PDB code 4ISN) of 20–30 Å, representing only ~40% of the total mass of IK1. The compactness was further confirmed by a Kratky plot of $q^2 I(q)$ *versus* q (not shown) that displayed a maximum, however, with some increase of the data toward the highest q, which can either be due to uncertainty in the background subtraction due to the very low (1 mg/ml) concentration or due to flexibility of some of the loops. In summary, the presence of Kunitz-1 improves the stability of the internal domain, resulting in a well behaved monomeric form of IK1.

The Internal Domain Improves the Inhibitory Properties of Kunitz-1—To study the effect of the internal domain on the inhibitory properties of Kunitz-1, we compared it with an HAI-1 fragment covering only Kunitz-1. Representative progress curves from the inhibition of the isolated catalytic domain of human matriptase by IK1 and Kunitz-1 are shown in Fig. 4. The pre-steady-state (time-dependent inhibition) analysis revealed that Kunitz-1 alone inhibited the catalytic domain of matriptase with a subnanomolar K_i of 0.38 ± 0.07 nM, whereas the addition of the internal domain resulted in a 3–4-fold higher affinity, with a K_i of 0.11 ± 0.02 nM (Table 2). To support these observations, we monitored the real-time binding of the two HAI-1 fragments to the catalytic domain of matriptase captured on a SPR sensor chip surface. The resulting K_D from the kinetic analysis agreed well with the K_i values despite the differences in experimental setup and reaction temperatures, with a K_D value of 0.60 ± 0.01 and 0.18 ± 0.04 nM for Kunitz-1 and IK1, respectively (Table 2). Interestingly, the primary determinant for the lower K_D value of IK1 was observed in the association phase, and the measured rate of association (k_{on}) was more than 3 times faster than for K1 alone (Table 2). To ensure that our 3–4-fold effects on the determined K_i and K_D (k_{on} in particular) values were not influenced by errors in the protein concentrations, we confirmed all protein concentrations by amino acid analysis. Our data propose a bidirectional relationship between the internal domain and Kunitz-1, resulting in improved solution stability of the internal domain and at the same time improved inhibitory function of Kunitz-1 manifested by a faster association with the protease.

Determination of the IK1 Crystal Structure Reveals the Structural Identity of a PKD-like Internal Domain and Intramolecular Interactions between Internal Domain and Kunitz-1—The crystal structure of IK1 includes two molecules in the asymmetric unit. The rotation axis between the two non-crystallographic symmetry copies is almost parallel to a higher symmetry crystal axis that would have allowed using the higher symmetry space group $P4_32_12$. The structure contains several disordered regions, which in part explains the relatively high refinement statistic as compared with what is expected at this resolution. Several rounds of refinement and density modification trials were performed to check for potential frame shifts but without success. Although it was possible to trace the entire main chain of the internal domain from Gly^{168} in the calculated electron density maps, three flexible loop regions within the internal domain were excluded from the final structure model (Fig. 5*A*): residues 182–188, 203–206, and 225–232. The structure of IK1 revealed two domains with different folds but similar overall dimensions. Interestingly, the two domains form a compact v-shaped tertiary structure with extensive interdomain interactions (Fig. 5*A*). The previously uncharacterized internal domain folds up in a β-sandwich fold consisting of a four-stranded mixed-type sheet sandwiched with an antiparallel three-stranded sheet of shorter strands. Using the structure-based homology search routines from the DALI server (42), we were able to confirm the structural homology (root mean square deviation ~1.8 Å) to the NMR structure of PKD-4 domain (Fig. 5*C*) from KIAA1837, as predicted earlier by I-TASSER (Fig. 5*B*). Although the homology of the overall fold is evident from the structural overlay, several areas are structurally different, such as the stretches covering β-strand 2 to 3 and β-strand 6 to 7, where the extent of loops and strands varies (Fig. 5, *C* and *D*). A comparison of our IK1 structure with the I-TASSER model and the homologous PKD-4 structure indicated that the IK1 structure may be missing a short seven-residue N-terminal stretch, as discussed earlier. Crystallization of a construct with the seven extra residues resulted in crystals of poorer diffraction quality; however, the final structure was indistinguishable from the IK1 structure of the shorter construct (data not shown). We concluded that the stretch from Arg^{153} to Ala^{167} is a flexible linker between MANEC and the internal domain in full-length HAI-1. It can be noted that the potential glycosylation site at Asn^{235} is located in the N-terminal part of β-strand 7 on the "back side" pointing away from Kunitz-1. In summary, the crystal structure of IK1 shows two structurally independently folded domains fixed to each other in a tertiary structure.

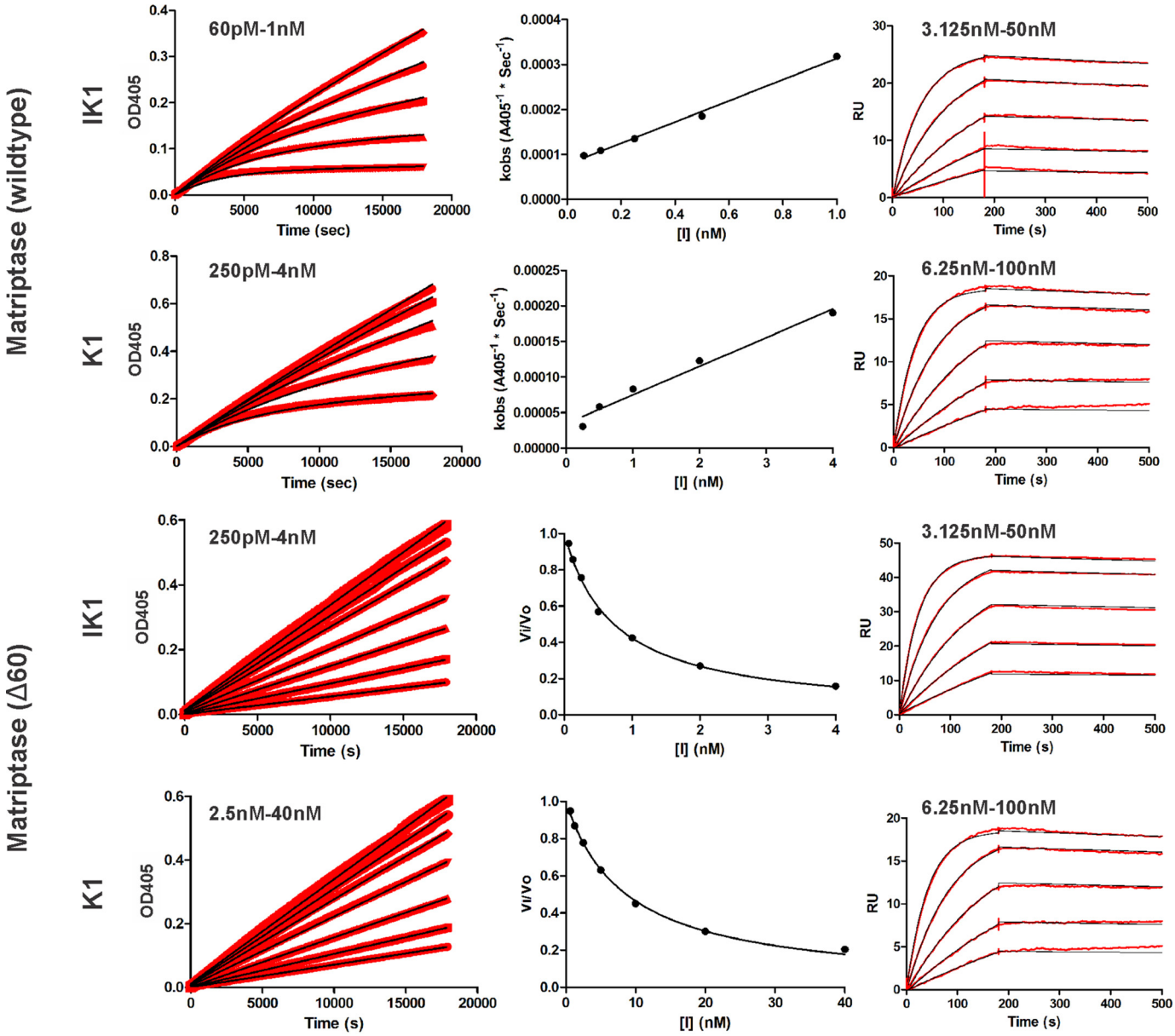

FIGURE 4. **Inhibition and SPR binding analysis of K1 and IK1 toward matriptase.** Representative progress curves (*red traces*) from chromogenic assays to study the inhibition of matriptase are shown in the *right column*. The conversion of the chromogenic substrate is monitored as a change in $A_{405\ nm}$. To start the reactions, 50 pM catalytic domain of wild type matriptase was added to reaction mixtures containing 500 μM substrate (S2288) and an increasing concentration of IK1 or K1. The *black traces* represent the best fits to Equation 1, from which k_{obs} values were obtained. Representative plots of k_{obs} as a function of inhibitor concentration are shown in the *middle column*. The *black line* represents the best fit to Equation 2 to determine k_{on} and k_{off} and subsequently the K_i. For the Δ60-loop matriptase (*two bottom panels*), 100 pM catalytic domain was incubated with the concentration series of IK1 or K1 for 10 min, followed by the addition of 500 μM chromogenic substrate (S2288) to start the reaction. The initial reaction velocities (v_i) were calculated from linear regression (*black traces*). The inhibition constants (K_i) were subsequently determined from the non-linear regression analysis of v_i/v_0 *versus* the inhibitor concentration, using Equation 4. In the *third column*, representative SPR data from the kinetic binding analysis are shown. C-terminally StrepII-tagged wild type or Δ60-loop variant matriptase was captured directly from the yeast culture medium on the sensor surface by anti-StrepII tag antibody. Concentration series of the IK1 or K1 protein preparations were injected. Corresponding fit to a 1:1 binding model using the evaluation software provided by the manufacturer is shown as *black traces*. All calculated values are shown in Table 2.

SAXS Analysis and Rigid Body Modeling Support the IK1 Tertiary Structure in Solution—SAXS analysis was used to support the monomeric form and the tertiary structure of IK1 from crystallography in solution. From the resulting pair distance distribution $p(r)$ function (Fig. 3*C*, *inset*), an *ab initio* (no prior knowledge) low resolution molecular model was constructed using DAMMIF. The combined volume of the *ab initio* models of IK1 showed a condensed flat overall shape with two distinct protruding features (Fig. 3*E*, *gray shape*). A theoretical scattering curve profile calculated from the IK1 crystal structure using CRYSOL and allowing a background to vary agreed reasonably with the experimental SAXS data (Fig. 3*D*) with a χ^2 value of 2.17 although with some systematic deviations that suggest that the size of the protein is larger in solution. To optimize the structure, we used the program BUNCH with a flexible linker between the two domains. We ran the program 10 times and selected the most representative model. The agreement with the SAXS data (Fig. 3*D*) using CRYSOL with a variable background is very good with a χ^2 value of 1.44. The crystal structure model and the optimized model were also compared with the

TABLE 2
The inhibitory constant (K_i) and SPR binding parameters between wild type matriptase or Δ60-loop matriptase and IK1 and Kunitz-1

	Chromogenic assay			SPR kinetic analysis		
	K_i	k_{on} (× 10^5)	k_{off} (× 10^{-5})	k_{on} (× 10^5)	k_{off} (× 10^{-5})	K_D
	n_M	$_M{}^{-1}\,s^{-1}$	s^{-1}	$_M{}^{-1}\,s^{-1}$	s^{-1}	n_M
Matriptase (WT)						
IK1	0.11 ± 0.02 (3)[a]	6.6 ± 0.3 (3)	6.9 ± 0.8 (3)	2.4 ± 0.5 (3)	4.3 ± 0.2 (3)	0.18 ± 0.04 (3)
K1	0.38 ± 0.07 (3)	1.0 ± 0.1 (3)	3.9 ± 0.6 (3)	0.72 ± 0.08 (3)	4.3 ± 0.5 (3)	0.60 ± 0.01 (3)
Matriptase (Δ60)						
IK1	0.46 ± 0.02 (3)	NA[b]	NA	5.7 ± 0.5 (3)	14 ± 4 (3)	0.24 ± 0.07 (3)
K1	5.5 ± 0.2 (3)	NA	NA	2.6 ± 0.2 (3)	25 ± 2 (3)	0.96 ± 0.4 (3)

[a] Values represents an average with corresponding S.D. value calculated from a number (in parenthesis) of independent determinations.
[b] NA, not applicable when using the steady-state-model.

ab initio envelope of the combined volume of the 10 solutions (Fig. 3, *E* and *F*, respectively). Both models have good agreement with the *ab initio* molecule shape, in clear support of a compact v-shaped tertiary structure. As expected, the optimized model fills out the *ab initio* envelope better than the crystal structure model, and it should also be noted that the orientation of the two domains in the two models is very similar. To summarize, although some conformational dynamics appears to be present in the interdomain interaction, overall, the SAXS data support the observed tertiary structure in solution.

The Stimulatory Effect of the Internal Domain on the Inhibitory Activity of Kunitz-1 Is Protease-independent—By a structural alignment of the IK1 structure to the available structure of the matriptase·Kunitz-1 complex (41), it was possible to propose a model of how the internal domain would position itself relative to the protease domain in the inhibitory complex. In this model, the v-shaped tertiary structure of the IK1 structure appears to "pinch" the 60-loop (chymotrypsin template numbering) protruding from the surface of the catalytic domain of matriptase (Fig. 6*A*). In the crystal structure of the matriptase·Kunitz-1 complex, the 60-loop was postulated to mediate an additional interaction between Asp^{60b} of the protease and Arg^{265} of the bound Kunitz-1 domain. In our model, the 60-loop appears to be the only area in the catalytic domain with which the internal domain can potentially interact. To verify whether the observed stimulatory effect of the internal domain is a consequence of additional interactions between the 60-loop and the internal domain, we prepared and analyzed a variant of matriptase in which the 60-loop had been deleted. The Δ60-loop variant was prepared by deleting the 6 residues of the stretch Asp^{660}–Arg^{665} and replacing them with two glycines. This was decided based on comparisons with structures of other HAI-1-inhibited serine proteases, such as HGFA (PDB code 1YC0), hepsin (PDB code 1P57), and plasma kallikrein (PDB code 2ANW), all of which have 4–5-residue-shorter 60-loops (Fig. 7). The Δ60-loop variant expressed and purified from *P. pastoris* indistinguishably from the wild type protease domain. The enzymatic properties of the Δ60-loop variant were close to that of the wild type protein, with a modest increase in the K_m for the conversion of the chromogenic substrate S2288 from 536 ± 13 to 852 ± 25 μM. Data from the K_i experiments of the HAI-1 fragments with the Δ60-loop variant fit well to a steady-state inhibition model, an obvious change in the kinetic profile compared with wild type matriptase (Fig. 4). Removal of the 60-loop resulted in a general increase in the K_i values for both K1 and IK1 by more than 14- and 4-fold, respectively (Table 2), compatible with the loss of the proposed interaction between K1 and the 60-loop. Results from a parallel SPR analysis supported the conclusion about a general drop in affinity upon removal of the 60-loop, although not as significant as observed for the K_i experiments. The dissociation rates were increased as expected from a loss of potential interaction with the 60-loop, but interestingly, the dissociation rates were increased by ~3-fold for both K1 and IK1. Irrespective of the 60-loop, the presence of the internal domain improved inhibition and interaction with the protease.

To test whether the stimulatory effect of the internal domain was restricted to matriptase, we included in the analysis three other serine proteases known to be inhibited by HAI-1: HGFA, hepsin, and plasma kallikrein. In all cases, the addition of the internal domain improved the K_i between 3- and 12-fold. The smallest enhancement was observed for the inhibition of matriptase, but matriptase is also inhibited by the lowest K_i value among the four tested proteases to begin with (Table 3). In summary, the addition of the internal domain stimulated the inhibitory activity of Kunitz-1 by a mechanism that appears to be predominantly independent of the targeted protease.

Analysis of Interdomain Contacts Supports the Hypothesis of Interdomain Cross-talk—Having established the interdomain interface as a key regulatory element, we made a detailed analysis of its components. The v-shaped IK1 complex is stabilized by an ~484 $Å^2$ (PISA server) interdomain interface. The interdomain interface involves residues primarily from β-strands 1 and 2 of the internal domain, both α-helices of Kunitz-1, and the covalent interdomain linkage. The interdomain interaction can be split into two main contact areas, one primarily composed of hydrogen bonding contacts and a second of hydrophobic interactions. Several hydrogen bonds, of main chain and side chain origin, criss-cross the interface between Val^{173}, Gln^{174}, and Leu^{241} of the internal domain and Ser^{242}, Lys^{244}, Gln^{245}, Leu^{298}, and Ala^{299} of Kunitz-1 (Fig. 6*B*). Leu^{298} from Kunitz-1 interacts with a hydrophobic patch on the internal domain consisting of Leu^{171}, Gln^{174}, Gln^{176}, Leu^{179}, and Leu^{181} (Fig. 6*C*). Although originally thought of as two independent domains, our structure clearly shows the absence of a "real" flexible interdomain linker. Following the main chain exiting from the C-terminal β-strand 7 of the internal domain, only the two residues Leu^{241} and Ser^{242} span the gap to the next secondary structure element, the

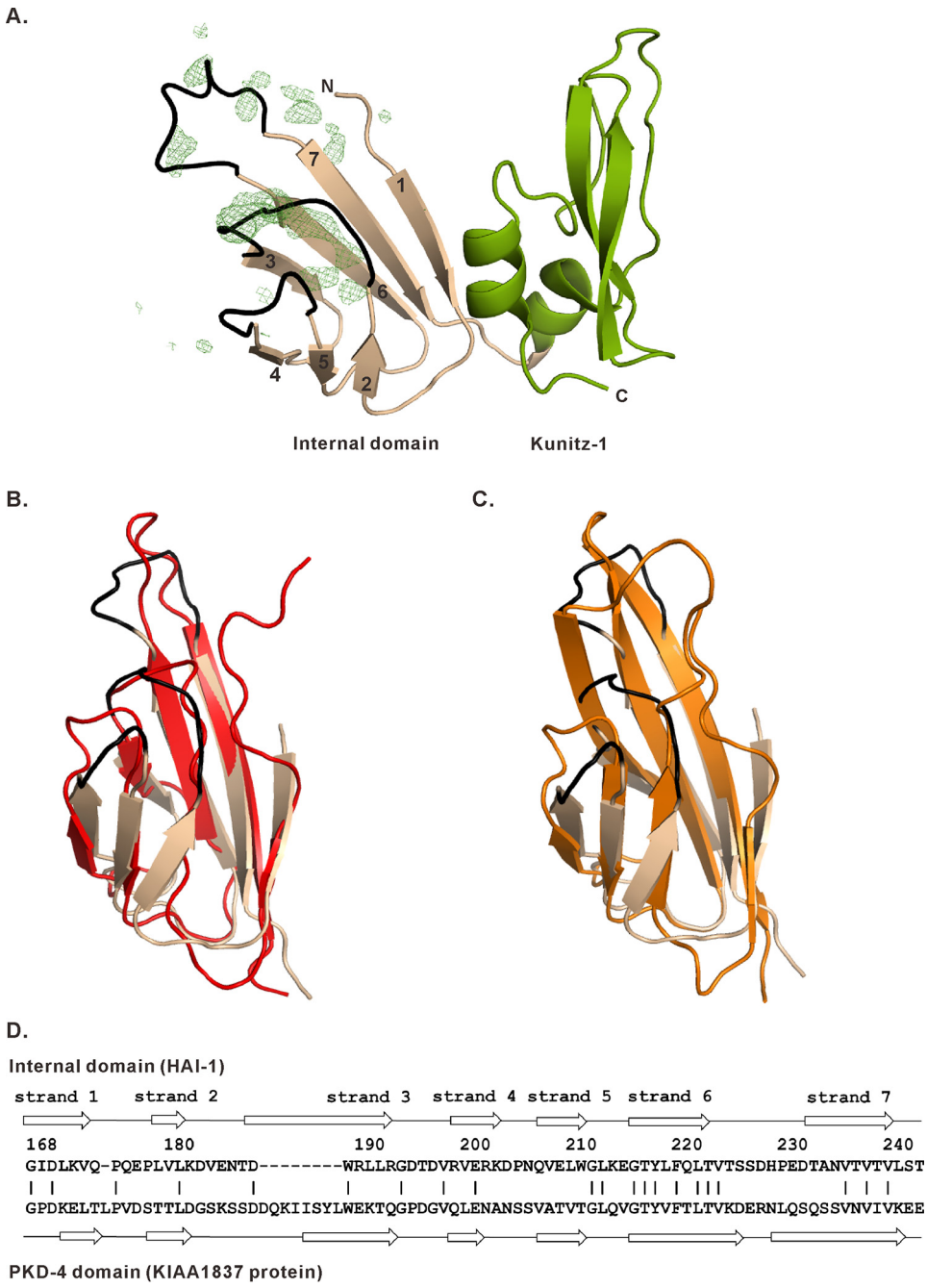

FIGURE 5. **Analysis of the IK1 crystal structure.** *A*, the IK1 crystal structure model (Gly^{168}–Val^{303}) is shown in a schematic (PyMOL cartoon) with the internal domain and Kunitz-1 *colored wheat* and *green*, respectively. The β-strands of the internal domain are *labeled 1–7*. Loop regions removed during the final structure refinement are *labeled* in *black*. The $F_o - F_c$ electron density (contoured at 3.0σ) corresponding to the removed loop regions is shown (*green mesh*). *B*, alignment of the internal domain from the IK1 crystal structure in *wheat color* with the *in silico* model of the internal domain (Gly^{161}–Thr^{244}) with the extended N terminus from I-TASSER in *red*. *C*, alignment of the internal domain in *wheat color* with the structural homologous (structure-based DALI server search) PKD-4 domain of the KIAA1837 protein (PDB code 2YRL) in *orange*. *D*, sequence alignment of the internal domain of HAI-1 and PKD-4 domain of the KIAA1837 protein. Secondary structure elements are indicated *above* the sequence: β-strands as *arrows* and *loops* and unstructured regions as *solid lines*. All schematic structures were prepared using PyMOL. The sequence alignment was prepared in I-TASSER.

first α-helix of Kunitz-1 (Fig. 6*B*). Both Leu^{241} and Ser^{242} are involved in interactions with other amino acids besides juxtaposing neighbors (*i.e.* Leu^{241} packs into a hydrophobic cavity formed in the interface between the two domains, and the carbonyl of Ser^{242} participates in α-helical-type hydrogen bonding with the amide of Thr^{246}).

To test our proposal of a function-transducing domain-domain interface, we mutated residues with central positions in the interface. Based on the crystal structure, we initially performed single alanine substitutions of Gln^{174} and Ser^{242} of the internal domain situated in the hydrogen bonding patch as well as Leu^{298} of Kunitz-1 that insert in the hydrophobic contact area on the internal domain. Neither of the single substitutions resulted in changes to the K_i or K_D as compared with the wild type IK1 protein (Table 4). However, substitution of both Gln^{174} and Ser^{242} resulted in a variant with a K_i and K_D value

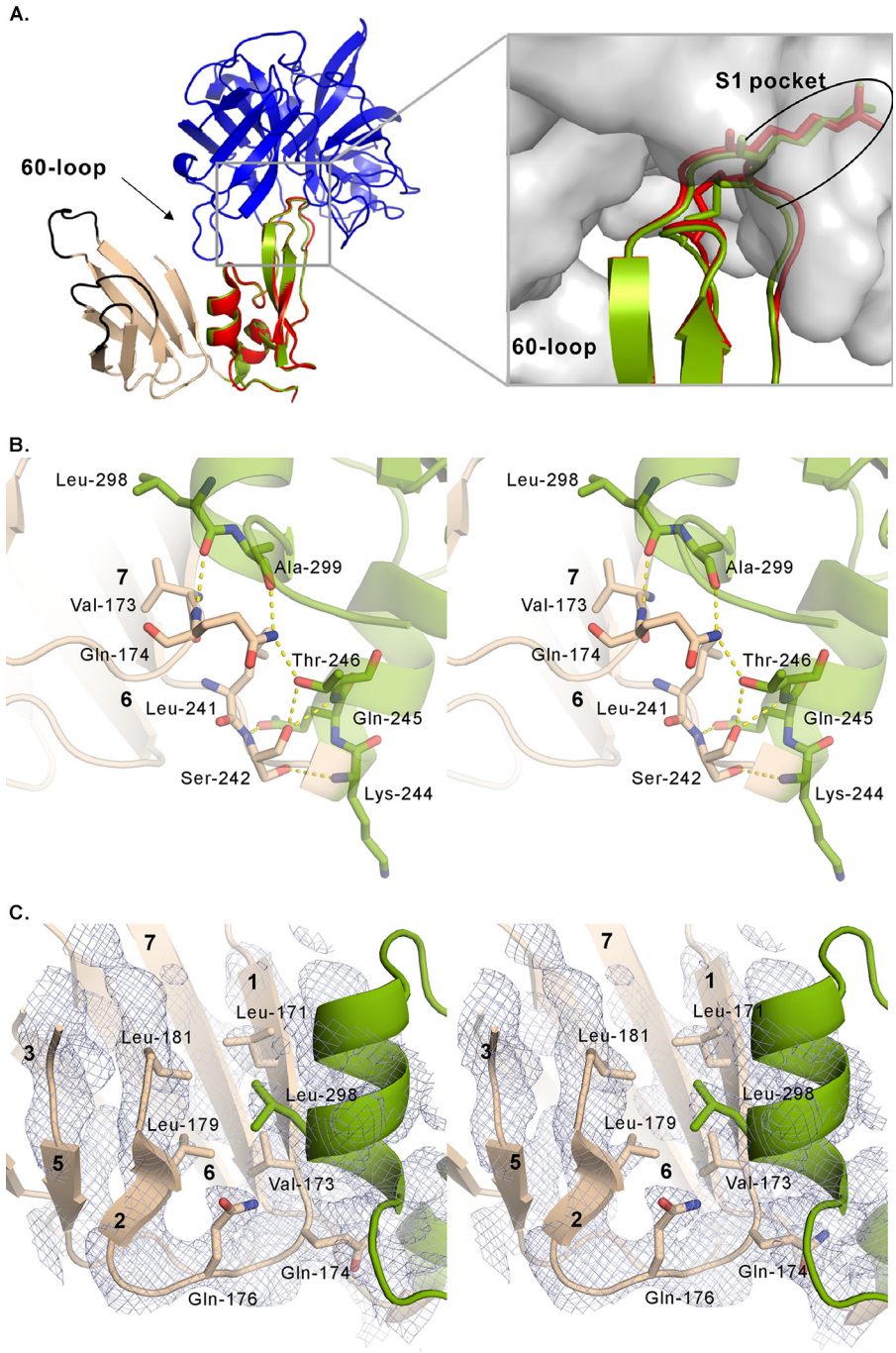

FIGURE 6. **Structural analysis of a modeled IK1·matriptase complex and the IK1 interdomain interaction.** *A*, model of the IK1·matriptase complex shown as a schematic in the *left panel*. IK1 is *colored* as in Fig. 3 with the catalytic domain of matriptase in *blue*. The structure model of the IK1·matriptase complex was modeled by an alignment of the Kunitz-1 domain from the available structure of matriptase·Kunitz-1 complex (PDB code 4ISN). The high overall structural similarity of the Kunitz domains has been emphasized by an *enlargement* of to the areas in direct interaction with the catalytic site of matriptase (*right*). *B* and C, *enlargement* of the regions of major importance for the interdomain interface between internal domain and Kunitz-1. Interacting residues are shown as *sticks* and highlighted by *labels* and *colored* by the corresponding domain *color*: internal domain as *wheat* and Kunitz-1 as *green*. Visible β-strands of the internal domain are *labeled* with *numbers*. *B*, stereo image of the area dominated by hydrogen bonding contacts. Hydrogen bonds are shown as *yellow dashed lines*. *C*, stereo image of the area dominated by hydrophobic interactions. The secondary structure elements close to the interface are shown as a schematic. The $2F_o - F_c$ electron density contoured at 1.0σ covering the entire region is shown as *blue mesh*. All structure figures were prepared using PyMOL.

about 10-fold larger than wild type Kunitz-1 alone (Table 4). The association rate of the double mutant was correspondingly reduced to a value similar to that of Kunitz-1 alone. Interestingly, the double substitution did not cause direct observable changes to the solution stability of the purified protein, although a ~10-fold drop in the expression level could indicate compromised folding and stability. In conclusion, simultaneous disruption of several interdomain interactions between the two domains reduces IK1 to an inhibitor that is even poorer than Kunitz-1, which supports that the observed stimulatory effect is being caused by a direct interdomain interaction with the internal domain.

Discussion

The Kunitz domain is a small 50–60-residue protein with a signature fold stabilized by three highly conserved disulfides. The Kunitz family includes well described inhibitors of serine proteases, although inhibition of some cysteine and aspartyl proteases has also been reported. Kunitz inhibitors are often regarded as single domain proteins, although proteins with between 2 and as many as 12 Kunitz domains have been described. The Kunitz inhibitory mechanism is described as "standard mechanism" inhibition in which the Kunitz forms a tight noncovalent complex with the target protease, blocking its active site by insertion of an exposed scissile bond loop as a nearly perfect substrate. The inhibitory mechanism is characterized by an essentially non-hydrolyzable scissile bond and exceptionally slow rates of dissociation, resulting in effectively irreversible inhibition of the protease (for a review, see Ref. 43). The Kunitz scaffold has received considerable pharmaceutical interest due to the relative ease of engineering protease inhibitors with high affinity and specificity and the applicability to directed evolution strategies (for a review, see Ref. 44). Commonly, the Kunitz is thought of as an isolated structural domain performing the inhibition of the target protease independently of external stimuli and without conformational changes during the reaction (rigid docking). We propose here that the molecular mechanism behind the inhibitory activity of Kunitz-1 from HAI-1 is an "exception to the rule."

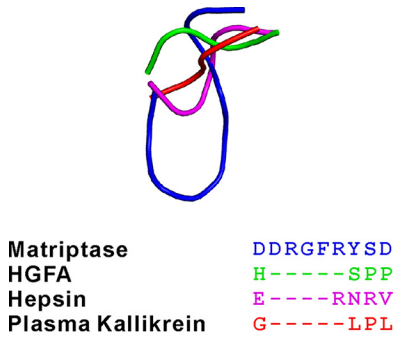

FIGURE 7. **Structural alignment of the 60-loop of matriptase with that of HGFA, hepsin, and plasma kallikrein.** Shown is a structural alignment, in a schematic, of the 60-loop in the catalytic domain from the indicated proteases all in complex with Kunitz-1 from HAI-1. The proteases are highlighted by different colors: matriptase (PDB code 4IS5) in *blue*, HGFA (PDB code 1YC0) in *green*, hepsin (PDB code 1P57) in *magenta*, and plasma kallikrein (PDB code 2ANW) in *red*. The structures were prepared using PyMOL. An alignment of the corresponding protein sequences within the 60-loop of the proteases is shown *below* in *colors* matching the structural alignment.

TABLE 3

The K_i values of the IK1 and Kunitz-1 HAI-1 fragments for the inhibition of the four indicated proteases

	K_i		
	IK1	Kunitz-1	Kunitz-1/IK1[a]
	nM		
Matriptase	0.11 ± 0.02 (3)[b]	0.38 ± 0.07 (3)	3.5
HGFA	0.87 ± 0.03 (3)	3.9 ± 0.2 (3)	4.5
Hepsin[c]	~0.1 (3)	~1.0 (3)	~10
Plasma kallikrein	9 ± 1 (3)	105 ± 6 (3)	11.7

[a] The ratio of the K_i values from K1 and IK1 inhibition of each protease is given in the rightmost column.
[b] Values represent an average with corresponding S.D. calculated from a number (in parenthesis) of independent determinations.
[c] The values are rough estimates due to non-ideal high protease concentration in the chromogenic assay.

The present project was initiated with the aim to expand the structural and functional knowledge base on the non-inhibitory domains of HAI-1. Contrary to expectations that each domain behaves as an independent structural unit, like "beads on a string," we now present data that show that Kunitz-1 is not independent of its neighboring PKD-like internal domain. In the single other study concerning the non-inhibitory domains of HAI-1 (17), fragments of rat HAI-1 were produced, and the K_i values for inhibition of full-length rat matriptase were determined. The K_i values of IK1 and K1 were determined to be 0.038 ± 0.003 and 0.3 ± 0.2 nM, respectively. Also, a >4-fold increase in the rate of association of Kunitz-1 (k_{on} derived from the K_i experiments) was observed in the presence of the internal domain. It was concluded that the internal domain somehow improved the availability of Kunitz-1 for reaction with the protease. Our biochemical data on human HAI-1 (K_i(IK1) = 0.11 ± 0.02 nM, K_i(K1) = 0.38 ± 0.007 nM) agree well with the data of the previous study. Taken together, the data support the conclusion that the internal domain stimulates the inhibitory activity of Kunitz-1 and that the mechanism is conserved between at least two mammalian species.

To understand the stimulatory effect of the internal domain, we performed a comparison of our structure model of Kunitz-1 with the previously solved structure model of Kunitz-1 in complex with the protease domain of matriptase (41) and HGFA (40). It revealed that the conformation of Kunitz-1 from our IK1 structure is essentially identical to the Kunitz-1 structure in both inhibitory complexes (Fig. 6*A*) with a root mean square deviation as low as 0.38 and 0.27 Å to the HGFA and matriptase complex, respectively. Even for the regions in direct contact with the protease domains, for which we observe well defined

TABLE 4

Binding analysis of the indicated IK1 interface mutants with matriptase

The Kunitz-1 data from Table 2 are reproduced here as a reference.

Variant	Chromogenic assay, K_i	SPR kinetic analysis		
		k_{on} (× 10^5)	k_{off} (× 10^{-5})	K_D
	nM	M^{-1} s^{-1}	M^{-1} s^{-1}	nM
IK1				
WT	0.11 ± 0.02 (3)[a]	2.4 ± 0.5 (3)	4.3 ± 0.2 (3)	0.18 ± 0.04 (3)
Q174A	0.12 ± 0.02 (3)	4.1 ± 0.5 (3)	7.5 ± 1.8 (3)	0.18 ± 0.02 (3)
S242A	0.12 ± 0.01 (3)	3.8 ± 0.3 (3)	5.2 ± 3.8 (3)	0.14 ± 0.1 (3)
L298A	0.11 ± 0.01 (3)	3.9 ± 0.2 (3)	7.4 ± 3.4 (3)	0.19 ± 0.1 (3)
Q174A/S242A	1.1 ± 0.2 (3)	0.62 ± 0.09 (3)	14 ± 5 (3)	2.3 ± 0.9 (3)
Kunitz-1				
WT	0.38 ± 0.07 (3)	0.72 ± 0.08 (3)	4.3 ± 0.5 (3)	0.60 ± 0.01 (3)

[a] Values represent an average with corresponding S.D. value calculated from a number (in parenthesis) of independent determinations.

electron density in our diffraction data, the conformations are identical (Fig. 6*A*, *enlargement*). Thus, the Kunitz-1 in our structure is in an optimal inhibitory conformation.

At this point, we could not rule out the possibility that our model of the IK1·matriptase complex was wrong and that the internal domain could interact with the protease domain outside the 60-loop. When we observed similar and even stronger stimulation by the internal domain on the inhibition of the three other target proteases, HGFA, hepsin, and plasma kallikrein, we concluded that the stimulating effect of the internal domain must be a direct effect on Kunitz-1 itself because it seemed unlikely that a similarly conserved interaction surface exists between the internal domain and the four different proteases. Whether the stimulation is a result of a beneficial entropic effect affecting the overall thermodynamics of the unbound and bound states or an allosterically induced conformational change resulting in a presentation of Kunitz-1 in a conformation that is more optimal for inhibition, remains an open question as no structure of the Kunitz-1 in the free form is yet available for comparison.

Why has additional stimulation of an already potent Kunitz domain-1 evolved? In the case of matriptase, a protease thought to be autoactivating and on the pinnacle of several important but also potentially harmful proteolytic and signaling cascades (for a review, see Ref. 45), a fine-tuned inhibition may be crucial to maintain a perfect balance of proteolytic events by controlling the time window between protease activation and inhibition. Although the matriptase/HAI-1 reaction most likely takes place between the two membrane-attached proteins on the same plasma membrane, which in itself increases the chance of a favorable interaction due to enhanced proximity and reduced diffusion, time-limiting rearrangements of the respective tertiary structures of these large multidomain proteins may have required a further optimization of Kunitz-1 to ensure efficient inhibition.

In the process of studying a potential additional direct interaction between the internal domain and the catalytic domain of matriptase, we discovered that the removal of the 60-loop also accelerated the inhibition, so that the binding kinetics could be described by simple steady-state kinetics. Similarly, the three additional proteases, HGFA, hepsin, and plasma kallikrein, were all inhibited by Kunitz-1 and IK1 with kinetics agreeing with a steady-state model (Fig. 8), compatible with the shorter 60-loops of all three proteases not posing a rate-limiting factor for the inhibition by macromolecular inhibitors as we now propose for matriptase (Table 2). An analogy can be drawn to a related serine protease thrombin, where a 60-loop of 10 residues, even longer than that of matriptase, appears to play a role in the specificity and affinity of macromolecular inhibitors targeting the active site of the protease (46).

After completing the IK1 structure, we identified the internal domain of human HAI-1 as a structural homolog of the PKD-4 domain of the uncharacterized human protein KIAA1837 (PDB code 2YRL). The low sequence identity of only 25% explains why the simple sequence-based homology searches had not previously been able to predict homology to the PKD domain. A study on a fish homolog of human HAI-1 was recently published (47). In this study, a PKD homology was predicted using the BLAST server to the similar sequence region between MANEC and Kunitz-1 (47). This shows how critical the case-specific sequence conservation and other search parameters are for the sequence-based homology search approach. Moreover, the existence of the predicted PKD homology in the fish variant seems to agree with a functional importance of an evolutionarily conserved internal domain of HAI-1.

The PKD domain, which shares some similarity to the Ig-like fold, was first identified in the polycystic kidney disease protein, polycystin-1. Polycystin-1 is a glycoprotein consisting of a long N-terminal extracellular region, multiple transmembrane domains, and a cytoplasmic tail. The N-terminal extracellular region contains 16 PKD domains thought to play a part in the mechanical stress-sensing function of polycystin-1 (48). Several naturally occurring mutations within polycystin-1 result in autosomal dominant polycystic kidney disease (49). Some of these mutations have been shown to affect the structural stability of the first PKD of polycystin-1 (50). This apparent direct link between structural integrity, externally induced changes, and the function of the PKD domain fits our observation that an internal domain gains structural integrity via its neighboring domain. Some studies have also proposed PKD domains as mediators of ligand interactions. In chitinase A from the Gram-negative bacteria *Alteromonas* sp. strain O-7, a PKD domain has been shown to bind the cell wall polysaccharide chitin (51). The PDK domains of certain surface layer proteins have also been suggested to take part in the regulation of cell adhesion (52). Our working hypothesis is that the internal domain of HAI-1, apart from its stimulatory interactions with Kunitz-1, may also function as a mediator of protein-ligand interactions based on these other reported findings.

To understand the stabilizing effect of Kunitz-1 on the monomeric form of the internal domain, we will need to understand what maintains the scaffold of the PKD fold of this domain. In the absence of disulfides, the major determinant for the stability of the folded domain is the protection of an extensive hydrophobic core, consisting of the aliphatic and aromatic residues Leu^{179}, Leu^{180}, Trp^{189}, Val^{208}, Tyr^{218}, Phe^{220}, Leu^{222}, Val^{236}, and Val^{238}, from energetically unfavorable solvent interactions. This was also observed in the first published structure of a PKD domain (49). The structure confirms an open PKD fold of the internal domain with a major groove/opening between β-strand 1 and 2. Association of the Kunitz-1 domain appears to stabilize this opening in the β-folded structure, at the same time preventing the exposure of a large hydrophobic area ($\sim$450 $Å^2$) on the surface of the internal domain. Removal of the Kunitz-1 may lead to a concerted polymerization of partly folded internal domains via hydrophobic interactions in agreement with the observation of soluble oligomers of defined molecular masses from size exclusion chromatography.

During the detailed structural analysis of the internal domain, we observed a characteristic sequence (Arg^{193}-Gly^{194}-Asp^{195}) in a solvent-exposed loop exiting from β-strand 3. The specific RGD sequence and its conformation make it a strong candidate as a potential interaction site for RGD-specific members of the cell surface receptor family of integrins (for a review, see Ref. 53). If

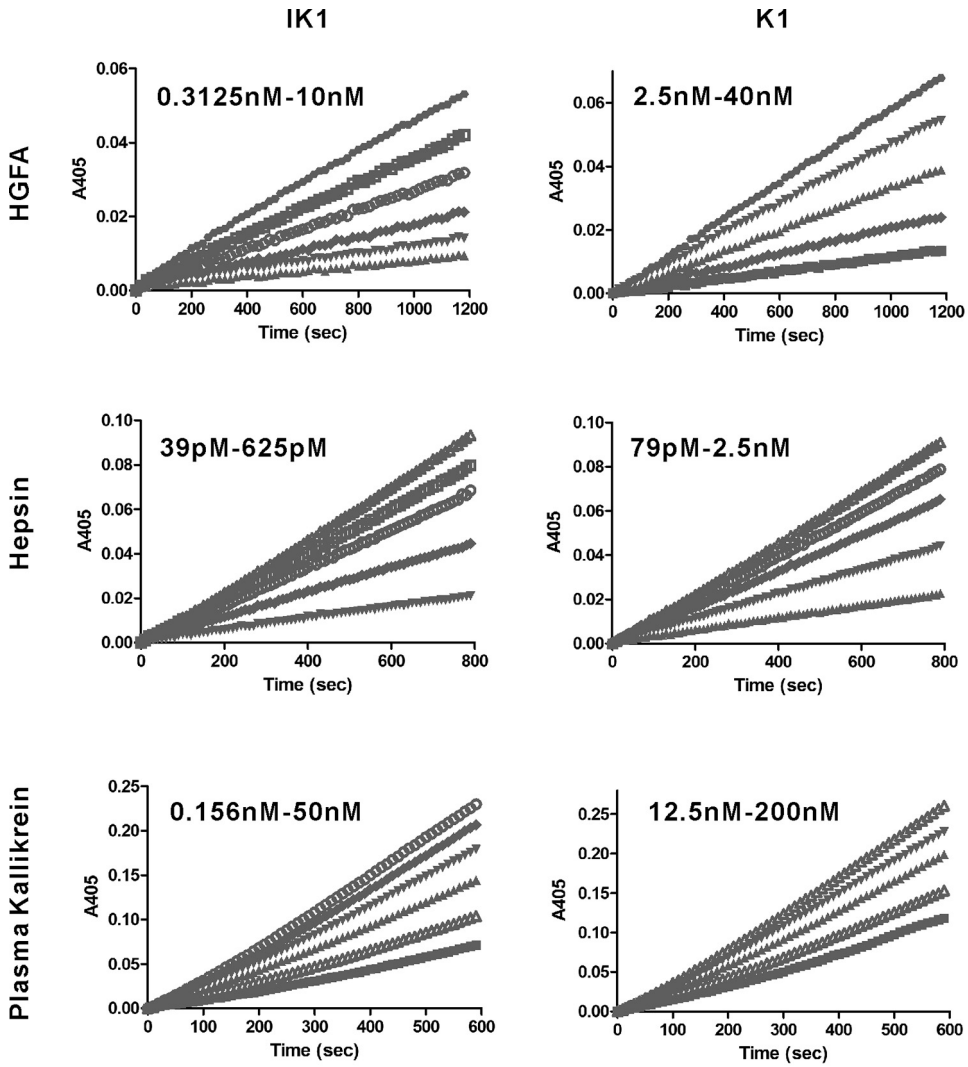

FIGURE 8. **Representative curves from K_i determination of K1 and IK1 toward HGFA, hepsin, and plasma kallikrein.** To determine the K_i values for the inhibition of the indicated serine proteases by IK1 and K1, a fixed concentration of HGFA (3 nM), hepsin (500 pM), or plasma kallikrein (3 nM) was incubated with the concentration series of the inhibitor (range indicated for each experiment) for 10 min. To start the assay reaction, the relevant chromogenic substrate (plasma kallikrein, S2302, 125 μM; hepsin, S2366, 200 μM; HGFA, Spectrozyme FVIIa, 200 μM) was added in a concentration corresponding to the K_m value of the specific substrate-protease pair. The initial reaction velocities were monitored as a change in absorbance at 405 nm. The K_i values were subsequently determined from the non-linear regression analyses of plots for v_i/v_0 *versus* [I], using Equation 4 (see "Experimental Procedures").

further investigations show that this is in fact the case, it may help the study of HAI-1 plasma membrane localization and possibly reveal other cellular functions.

In conclusion, from the x-ray crystal structure analysis, we could confirm that the internal domain belongs to the abundant but underinvestigated family of PKD domains. Our combined data confirmed the existence of an unexpected domain-domain interface and that the internal domain and Kunitz-1 form a defined tertiary structure. In the present work, we propose a novel mechanism, in which the association of an auxiliary domain, in this case the internal domain, improves the inhibitory activity of the neighboring Kunitz-1. The stimulatory effect, which is independent of the target protease, appears to stem from a stabilization of the inhibitory conformation of Kunitz-1 before its encounter with the protease. Moreover, the presented structure provides the first structural evidence of a functional role of a non-inhibitory domain of HAI-1 (*i.e.* provides new insight to the detailed molecular interplay between tertiary structure and the inhibitory mechanism of HAI-1). To the best of our knowledge, our study provides not only the first example of a Kunitz inhibitory reaction improved by external stimuli but also the first report of an intramolecular interaction between a Kunitz domain and another domain. Our study shows the 60-loop in matriptase to be a rate-limiting factor for macromolecular access to the active site. Finally, our structure reveals the presence of a potential interaction site for integrins within HAI-1.

Author Contributions—Z. H. carried out the bulk of the experimental work and data analysis; L. D. M. performed initial Ki experiments on wild type matriptase; A. J. prepared and crystallized the IK1 construct with the extended N terminus; J. S. P. provided SAXS instrumentation and participated in SAXS data analysis; J. P. M. participated in x-ray diffraction data interpretation and IK1 structure validation; P. A. A. provided reagents, instruments, and critical reading of the manuscript; J. K. J. designed experiments and analyzed data; Z. H. and J. K. J. wrote the manuscript.

Acknowledgments—We thank Tobias Kroman-Hansen and Peng Xu for generously providing purified catalytic domains of hepsin and plasma kallikrein, respectively. We thank Christine R. Schar for careful editing of the manuscript.

References

1. Tanaka, H., Nagaike, K., Takeda, N., Itoh, H., Kohama, K., Fukushima, T., Miyata, S., Uchiyama, S., Uchinokura, S., Shimomura, T., Miyazawa, K., Kitamura, N., Yamada, G., and Kataoka, H. (2005) Hepatocyte growth factor activator inhibitor type 1 (HAI-1) is required for branching morphogenesis in the chorioallantoic placenta. *Mol. Cell. Biol.* **25,** 5687–5698
2. Fan, B., Brennan, J., Grant, D., Peale, F., Rangell, L., and Kirchhofer, D. (2007) Hepatocyte growth factor activator inhibitor-1 (HAI-1) is essential for the integrity of basement membranes in the developing placental labyrinth. *Dev. Biol.* **303,** 222–230
3. Szabo, R., Molinolo, A., List, K., and Bugge, T. H. (2007) Matriptase inhibition by hepatocyte growth factor activator inhibitor-1 is essential for placental development. *Oncogene* **26,** 1546–1556
4. Kawaguchi, M., Takeda, N., Hoshiko, S., Yorita, K., Baba, T., Sawaguchi, A., Nezu, Y., Yoshikawa, T., Fukushima, T., and Kataoka, H. (2011) Membrane-bound serine protease inhibitor HAI-1 is required for maintenance of intestinal epithelial integrity. *Am. J. Pathol.* **179,** 1815–1826
5. Nagaike, K., Kawaguchi, M., Takeda, N., Fukushima, T., Sawaguchi, A., Kohama, K., Setoyama, M., and Kataoka, H. (2008) Defect of hepatocyte growth factor activator inhibitor type 1/serine protease inhibitor, Kunitz type 1 (Hai-1/Spint1) leads to ichthyosis-like condition and abnormal hair development in mice. *Am. J. Pathol.* **173,** 1464–1475
6. Szabo, R., Kosa, P., List, K., and Bugge, T. H. (2009) Loss of matriptase suppression underlies spint1 mutation-associated ichthyosis and postnatal lethality. *Am. J. Pathol.* **174,** 2015–2022
7. List, K., Szabo, R., Molinolo, A., Sriuranpong, V., Redeye, V., Murdock, T., Burke, B., Nielsen, B. S., Gutkind, J. S., and Bugge, T. H. (2005) Deregulated matriptase causes ras-independent multistage carcinogenesis and promotes ras-mediated malignant transformation. *Genes Dev.* **19,** 1934–1950
8. Shimomura, T., Denda, K., Kitamura, A., Kawaguchi, T., Kito, M., Kondo, J., Kagaya, S., Qin, L., Takata, H., Miyazawa, K., and Kitamura, N. (1997) Hepatocyte growth factor activator inhibitor, a novel Kunitz-type serine protease inhibitor. *J. Biol. Chem.* **272,** 6370–6376
9. Lin, C. Y., Anders, J., Johnson, M., and Dickson, R. B. (1999) Purification and characterization of a complex containing matriptase and a Kunitz-type serine protease inhibitor from human milk. *J. Biol. Chem.* **274,** 18237–18242
10. Fan, B., Wu, T. D., Li, W., and Kirchhofer, D. (2005) Identification of hepatocyte growth factor activator inhibitor-1B as a potential physiological inhibitor of prostasin. *J. Biol. Chem.* **280,** 34513–34520
11. Szabo, R., Uzzun Sales, K., Kosa, P., Shylo, N. A., Godiksen, S., Hansen, K. K., Friis, S., Gutkind, J. S., Vogel, L. K., Hummler, E., Camerer, E., and Bugge, T. H. (2012) Reduced prostasin (CAP1/PRSS8) activity eliminates HAI-1 and HAI-2 deficiency-associated developmental defects by preventing matriptase activation. *PLoS Genet.* **8,** e1002937
12. Herter, S., Piper, D. E., Aaron, W., Gabriele, T., Cutler, G., Cao, P., Bhatt, A. S., Choe, Y., Craik, C. S., Walker, N., Meininger, D., Hoey, T., and Austin, R. J. (2005) Hepatocyte growth factor is a preferred *in vitro* substrate for human hepsin, a membrane-anchored serine protease implicated in prostate and ovarian cancers. *Biochem. J.* **390,** 125–136
13. Hashimoto, T., Kato, M., Shimomura, T., and Kitamura, N. (2010) TMPRSS13, a type II transmembrane serine protease, is inhibited by hepatocyte growth factor activator inhibitor type 1 and activates pro-hepatocyte growth factor. *FEBS J.* **277,** 4888–4900
14. Kato, M., Hashimoto, T., Shimomura, T., Kataoka, H., Ohi, H., and Kitamura, N. (2012) Hepatocyte growth factor activator inhibitor type 1 inhibits protease activity and proteolytic activation of human airway trypsin-like protease. *J. Biochem.* **151,** 179–187
15. Miller, G. S., Zoratti, G. L., Murray, A. S., Bergum, C., Tanabe, L. M., and List, K. (2014) HATL5: a cell surface serine protease differentially expressed in epithelial cancers. *PLoS One* **9,** e87675
16. Kirchhofer, D., Peek, M., Li, W., Stamos, J., Eigenbrot, C., Kadkhodayan, S., Elliott, J. M., Corpuz, R. T., Lazarus, R. A., and Moran, P. (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
17. Kojima, K., Tsuzuki, S., Fushiki, T., and Inouye, K. (2008) Roles of functional and structural domains of hepatocyte growth factor activator inhibitor type 1 in the inhibition of matriptase. *J. Biol. Chem.* **283,** 2478–2487
18. Denda, K., Shimomura, T., Kawaguchi, T., Miyazawa, K., and Kitamura, N. (2002) Functional characterization of Kunitz domains in hepatocyte growth factor activator inhibitor type 1. *J. Biol. Chem.* **277,** 14053–14059
19. Domoto, T., Takino, T., Guo, L., and Sato, H. (2012) Cleavage of hepatocyte growth factor activator inhibitor-1 by membrane-type MMP-1 activates matriptase. *Cancer Sci.* **103,** 448–454
20. Kataoka, H., Shimomura, T., Kawaguchi, T., Hamasuna, R., Itoh, H., Kitamura, N., Miyazawa, K., and Koono, M. (2000) Hepatocyte growth factor activator inhibitor type 1 is a specific cell surface binding protein of hepatocyte growth factor activator (HGFA) and regulates HGFA activity in the pericellular microenvironment. *J. Biol. Chem.* **275,** 40453–40462
21. Shimomura, T., Denda, K., Kawaguchi, T., Matsumoto, K., Miyazawa, K., and Kitamura, N. (1999) Multiple sites of proteolytic cleavage to release soluble forms of hepatocyte growth factor activator inhibitor type 1 from a transmembrane form. *J. Biochem.* **126,** 821–828
22. Hong, Z., Nowakowski, M., Spronk, C., Petersen, S. V., Andreasen, P. A., Koźmiński, W., Mulder, F. A., and Jensen, J. K. (2015) The solution structure of the MANEC-type domain from hepatocyte growth factor activator inhibitor-1 reveals an unexpected PAN/apple domain-type fold. *Biochem. J.* **466,** 299–309
23. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132
24. Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43,** 186–190
25. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763
26. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674
27. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER suite: protein structure and function prediction. *Nat. Methods* **12,** 7–8
28. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
29. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221
30. Copeland, R. A. (2000) *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*, 2nd Ed., Wiley-VCH Inc., New York
31. Price, N. C. (1985) The determination of K_m values from Lineweaver-Burk plots. *Biochem. Educ.* **13,** 81
32. Pedersen, J. S. (2004) A flux- and background-optimized version of the NanoSTAR small-angle x-ray scattering camera for solution scattering. *J. Appl. Crystallogr.* **37,** 369–380
33. Glatter, O. (1977) New method for evaluation of small-angle scattering data. *J. Appl. Crystallogr.* **10,** 415–421
34. Pedersen, J. S., Hansen, S., and Bauer, R. (1994) The aggregation behavior of zinc-free insulin studied by small-angle neutron-scattering. *Eur. Biophys. J.* **22,** 379–389
35. Svergun, D., Barberato, C., and Koch, M. H. J. (1995) CRYSOL: a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28,** 768–773
36. Franke, D., and Svergun, D. I. (2009) DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42,** 342–346
37. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36,** 860–864
38. Kozin, M. B., and Svergun, D. I. (2001) Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* **34,** 33–41

39. Petoukhov, M. V., and Svergun, D. I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89,** 1237–1250
40. Shia, S., Stamos, J., Kirchhofer, D., Fan, B., Wu, J., Corpuz, R. T., Santell, L., Lazarus, R. A., and Eigenbrot, C. (2005) Conformational lability in serine protease active sites: structures of hepatocyte growth factor activator (HGFA) alone and with the inhibitory domain from HGFA inhibitor-1B. *J. Mol. Biol.* **346,** 1335–1349
41. Zhao, B., Yuan, C., Li, R., Qu, D., Huang, M., and Ngo, J. C. (2013) Crystal structures of matriptase in complex with its inhibitor hepatocyte growth factor activator inhibitor-1. *J. Biol. Chem.* **288,** 11155–11164
42. Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38,** W545–W549
43. Ranasinghe, S., and McManus, D. P. (2013) Structure and function of invertebrate Kunitz serine protease inhibitors. *Dev. Comp. Immunol.* **39,** 219–227
44. Viswanathan, M., Comeau, S. R., and Ladner, R. C. (2009) Engineered protein protease inhibitors. *Curr. Enzyme Inhibit.* **5,** 87–98
45. List, K. (2009) Matriptase: a culprit in cancer? *Future Oncol.* **5,** 97–104
46. Rezaie, A. R., and Yang, L. (2005) Deletion of the 60-loop provides new insights into the substrate and inhibitor specificity of thrombin. *Thromb. Haemost.* **93,** 1047–1054
47. Bathige, S. D., Umasuthan, N., Godahewa, G. I., Jayasinghe, J. D., Whang, I., Noh, J. K., and Lee, J. (2015) A homolog of Kunitz-type serine protease inhibitor from rock bream, *Oplegnathus fasciatus*: molecular insights and transcriptional modulation in response to microbial and PAMP stimulation, and tissue injury. *Fish Shellfish Immunol.* **46,** 285–291
48. Retailleau, K., and Duprat, F. (2014) Polycystins and partners: proposed role in mechanosensitivity. *J. Physiol.* **592,** 2453–2471
49. Bycroft, M., Bateman, A., Clarke, J., Hamill, S. J., Sandford, R., Thomas, R. L., and Chothia, C. (1999) The structure of a PKD domain from polycystin-1: implications for polycystic kidney disease. *EMBO J.* **18,** 297–305
50. Ma, L., Xu, M., Forman, J. R., Clarke, J., and Oberhauser, A. F. (2009) Naturally occurring mutations alter the stability of polycystin-1 polycystic kidney disease (PKD) domains. *J. Biol. Chem.* **284,** 32942–32949
51. Orikoshi, H., Nakayama, S., Hanato, C., Miyamoto, K., and Tsujibo, H. (2005) Role of the N-terminal polycystic kidney disease domain in chitin degradation by chitinase A from a marine bacterium, *Alteromonas* sp. strain O-7. *J. Appl. Microbiol.* **99,** 551–557
52. Jing, H., Takagi, J., Liu, J. H., Lindgren, S., Zhang, R. G., Joachimiak, A., Wang, J. H., and Springer, T. A. (2002) Archaeal surface layer proteins contain β propeller, PKD, and β helix domains and are related to metazoan cell surface proteins. *Structure* **10,** 1453–1464
53. Ruoslahti, E. (1996) RGD and other recognition sequences for integrins. *Annu. Rev. Cell Dev. Biol.* **12,** 697–715
54. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25,** 495–503