Conversion of the Kunitz-type module of collagen VI into a highly active trypsin inhibitor by site-directed mutagenesis

Eddie KOHFELDT[1], Walter GÖHRING[1], Ulrike MAYER[1], Markus ZWECKSTETTER[1], Tad A. HOLAK[1], Mon-Li CHU[2,3] and Rupert TIMPL[1]

[1] Max-Planck-Institut für Biochemie, Martinsried, Germany
[2] Department of Biochemistry and Molecular Biology, Jefferson Institute of Molecular Medicine, Thomas Jefferson University, Philadelphia, USA
[3] Department of Dermatology and Cutaneous Biology, Jefferson Institute of Molecular Medicine, Thomas Jefferson University, Philadelphia, USA

The recombinant Kunitz protease inhibitor module (domain C5) of human collagen $\alpha3(VI)$ chain was previously shown to lack inhibitory activity for proteases with trypsin-like specificity and some other proteases. We have now prepared mutants in the binding loop region including the P1' site (D2889→A), the P2' site (F2890→R) and the P3 site (T2886→P) and in a more remote region (W2907→V) either as individual substitutions or combinations of them. These mutants were analyzed for their kinetics of binding to trypsin by surface plasmon resonance and for their capacity to inhibit various proteases. Single substitutions (D→A, T→P, W→V) showed an effect only for D→A which bound to trypsin with $K_d = 0.25 \mu M$. A 25−100-fold increase in affinity was observed for the double mutants T→P/D→A and F→R/D→A and approached the affinity of aprotinin ($K_d \approx 0.01$ nM) in two different triple mutants. These affinities correlated well with the inhibitory capacities of the mutants for trypsin in the cleavage of a large protein and a small peptide substrate. A similar but not completely identical improvement in inhibitory capacity was also observed for leucocyte elastase but not for thrombin. These data could be interpreted in terms of steric interferences or lack of hydrogen bonding of a few critical residues based on three-dimensional structures available for the C5 domain.

Keywords: collagen VI; Kunitz module; trypsin; elastase; protein modeling.

The 50−60-residue Kunitz-type module is not only characteristic for a distinct family of protease inhibitors (Bode and Huber, 1992) but also found in larger proteins such as in a neural amyloid precursor protein (Ponte et al., 1988; Tanzi et al., 1988; Kitaguchi et al., 1988) and at the C-terminals of collagen $\alpha3(VI)$ chain (Bonaldo and Colombatti, 1989; Chu et al., 1990) and $\alpha1(VII)$ chain (Greenspan, 1993). The sequences of the amyloid precursor module and the $\alpha3(VI)$ chain module contained an arginine at the P1 equivalent position of the binding loop indicating their potential to inhibit trypsin or trypsin-like proteases. This prediction was confirmed for the amyloid protein domain in several recombinant studies (Oltersdorf et al., 1989; Kido et al., 1990; Castro et al., 1990; Wagner et al., 1992; Oyama et al., 1993; van Nostrand et al., 1994). A recombinant Kunitz-type module of $\alpha3(VI)$ chain (domain C5, see Chu et al., 1990), however, failed to show any inhibitory activity for trypsin, thrombin, kallikrein, leucocyte elastase and mast-cell chymase (Mayer et al., 1994). This failure was thought to be due to the presence of unfavorable residues in the binding loop (P3, P1' positions) as well as in the first β strand of the core domain of the module. Recent studies by NMR (Zweckstetter et al., 1996)

and by X-ray crystallography (Arnoux et al., 1995), however, demonstrated for the recombinant domain C5 a secondary and tertiary structure very close to that of the same module of various protease inhibitors (Bode and Huber, 1992) and of the amyloid precursor (Hynes et al., 1990).

These observations indicated that domain C5 could be a suitable scaffold for studying structure/function relationships of such modules by subtle variation of the amino acid sequence. In the present study we have addressed this question by targeted mutagenesis of just four residues of domain C5 using the trypsin inhibitor aprotinin as a lead structure for designing the substitutions. The data demonstrated no or only small effects of single substitutions but dramatic changes after introducing three substitutions in two different combinations. These two triple mutants were in fact indistinguishable from bovine aprotinin with respect to their binding affinity and inhibitory capacity for trypsin. Furthermore, several of the effects could be explained at the level of the solution structure of domain C5 which was available through NMR studies (Zweckstetter et al., 1996).

MATERIALS AND METHODS

Construction of an episomal expression vector for domain C5. The eukaryotic episomal expression vector pCEP4 (10.4 kb, Invitrogen) was used to produce a modified vector pCEP-Sh (9.23 kb) in order to replace the hygromycin by the phleomycin resistance gene to allow selection both in bacteria and mammalian cells. An 8.2-kb fragment containing the cytomegalovirus (CMV) promoter was generated from pCEP4 by

restriction with *Nru*I and partial cleavage with *Sal*I for the selective release of the hygromycin resistence gene. A 989-bp *Hpa*I–*Bam*HI fragment of the pUT 626 vector (Eurogentech) containing the SV40 and EM7 promoters, the phleomycin resistance gene (Sh-ble) and the SV40 poly-(A) signal was inserted into the pCEP4 fragment by blunt end ligation at one site *(Nru*I*)* and the use of an annealed 16 bp adaptor (5′-TCGAGGAAGGGGT-TCG and 5′-GATCCGAACCCCTTCC) for connection of the *Sal*I and *Bam*HI sites. This novel pCEP-Sh vector was cloned in *Escherichia coli* on Lurian-Bertani plates containing 5 µg/ml phleomycin (Eurogentech) and used for generating various expression vectors.

The expression vector pRc/Ac (Mayer et al., 1994) was used as a source of domain C5 cDNA connected to the BM-40 signal peptide. The 0.4-kb insert was released with *Hind*III/*Not*I, purified by agarose gel electrophoresis and inserted into the polylinker region of pCEP-Sh 3′ next to its CMV promoter to obtain vector pCEP-Sh/Ac. Human embryonic kidney cells which constitutively express the EBNA-1 protein from Epstein-Barr virus (H293 EBNA, Invitrogen) were used for transfection (Nischt et al., 1991) in order to increase the replication rate of the plasmid. The initial selection of transfected cells was accomplished in the presence of 25 µg/ml phleomycin which was then reduced to 7 µg/ml during the phase of collecting serum-free culture medium.

Construction of C5 mutant expression vectors. Mutants were generated from the pCEP-Sh/Ac vector by PCR with Vent polymerase (Biolabs) following the supplier's instruction. Commercially obtained oligonucleotides (Pharmacia) FP-Not 5′-CTGGATCCGGCCTTGCCGGCC for the 3′end and FP-Hind 5′-GAAGCTGGGTACCAGCTG for the 5′end together with two overlapping oligonucleotides (see below) matching the region designated for mutation were used to produce two overlapping subfragments of C5 which were then purified by agarose gel electrophoresis. A mixture of these fragments was denatured, annealed and used as a template to obtain a full-length cDNA by five cycles of PCR. Subsequent addition of the terminal primers (FP-Not, FP-Hind) was followed by 25 cycles of PCR to amplify the product. These PCR products were restricted with *Hind*III and *Not*I, purified on a Sephacryl S-400 micro spin column following the supplier's protocol (Pharmacia) and inserted into the *Hind*III/*Not*I-restricted pCEP-Sh vector and cloned in *E. coli*. The correct insertion and mutation was then verified by cycle sequencing on a 373A DNA sequencer (Applied Biosystems).

The following primers were used for introducing single amino acid mutations (mutated nucleotides are underlined) and neutral mutations which introduced additional restriction sites: primer 1, 5′-GATATATGCAAGTTGCCGAAAGACGAAGG-ACCCTGCAGGG; primer 2, 5′-CGGCAACTTGCATATATC; primer 3, 5′-GGAACTTGCAGGGCTTTCATACTTAAGTGG-TACTATGATCCAA; primer 4, 5′-TATGAAAGCCCTGCAAG-TTCCTT; primer 5, 5′-ACACCAAAAGCTGTGCACGATT-CGTGTATGGAGGTTGTGGTGG; primer 6, 5′-GCACAGCT-TTTGGTGTTTG; primer 7, 5′-GCAGGGCTAGGATATTAA-AATGGTACTATGATCC; primer 8, 5′-AATATCCTAGCCCT-GCAAGTTCC.

These primers were used with pCEP-Sh/Ac as template to produce mutant T→P (primers 1 and 2), mutant D→A (primers 3 and 4) and mutant W→V (primers 5 and 6). Two amino acid substitutions were introduced with the template of mutant D→A using the primer combinations 1 and 2 (mutant T→P/D→A) and 7 and 8 (mutant D→A/F→R). Two different triple amino acid substitutions were constructed using the mutant T→P/D→A template and primers 5 and 6 (mutant T→P/D→A/W→V) and the mutant D→A/F→R template and primers 1

and 2 (mutant T→P/D→A/F→R). H293-EBNA cells were then used for transfections with the mutated vectors as described above.

Purification and sources of proteins and proteases. Recombinant C5 and its mutants were initially purified from serum-free culture medium (0.5–1 l) of transfected cells on DEAE-cellulose (Mayer et al., 1994). None of recombinant products bound to the column and were dialysed against 0.2 M NH_4HCO_3 then lyophilized. They were then passed over a reverse-phase C18 column equilibrated in 0.1% trifluoroacetic acid (Mayer et al., 1991) and could be eluted as distinct peaks in a narrow range of the acetonitrile gradient (0–70%). Recombinant mouse nidogen was prepared as described (Fox et al., 1991). Reference protease inhibitors used were aprotinin (Trasylol^R, Bayer AG) and the thrombin inhibitor rhodniin (Friedrich et al., 1993) kindly provided by Dr W. Bode (Martinsried). Trypsin treated with *N*-α-tosyl-L-phenylalanine chloromethane (Worthington) or of sequencing grade (Boehringer Mannheim) and thrombin (Boehringer Mannheim) were purchased and human leucocyte elastase was a gift of Dr H. P. Nick (Ciba-Geigy, Basel).

Surface plasmon resonance assay. Assays were performed in a BIAcore apparatus equipped with CM 5 chips from Pharmacia Biosensor (Fägerstam et al., 1992). Trypsin was immobilized on the CM 5 chip via amine coupling procedure. After activation of the carboxymethylated dextran layer by addition of 35 µl of a mixture of 0.05 M *N*-hydroxysuccinimide and 0.2 M *N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide at a flow rate of 5 µl/min, 50 µl of trypsin solution in 0.1 M sodium acetate pH 4.5 (1 mg/ml) was added. Residual activated carboxyl groups of the chip were saturated by reaction with 35 µl 1 M ethanolamine pH 8.5. This resulted in 4800 resonance units of immobilized trypsin, which corresponds to a surface concentration of about 5 ng/mm^2. Binding assays were performed in 10 mM Hepes pH 7.4, 150 mM NaCl, 3.4 mM EDTA and 0.05% BIAcore P20 surfactant at a flow rate of 5 µl/min. Soluble ligands were applied in the concentration range 3–100 µM, depending on the binding constants. This resulted in signals of about 200–1100 resonance units. Kinetic constants were calculated from the association phase and the dissociation phase, respectively, using BIAevaluation software version 2.1. The binding curves were fitted according to the one-to-one model A + B = AB. Experiments with buffer alone revealed a small baseline drift indicating some leakage of the immobilized trypsin. This constant drift was used to correct both the association and dissociation curves.

Protease inhibition assays. Nidogen (200 µg/ml) was dissolved in 0.05 M Tris/HCl pH 7.4, 0.1 M NaCl, 2 mM $CaCl_2$ and incubated with individual proteases for 4–24 h at 37°C (2–20 µg/ml). Protease inhibitors were added to the proteases at 5- or 100-fold molar excess and incubated for 1 h at 37°C prior to mixing with nidogen. Aliquots were then taken for examination of fragment patterns by electrophoresis (Mayer et al., 1994).

The peptide substrate carbobenzoxy-valyl-glycyl-arginine 4-nitrilanilide acetate (Chromozyme^R TRY, Boehringer Mannheim) was used for the determination of inhibitory constants (K_i). A 10-µl aliquot of trypsin in 0.1 mM HCl (2.5 µg/ml) was incubated (5 min, 22°C) with 10 µl inhibitor (2.5–20 µg/ml) dissolved in 0.1 M Tris/HCl pH 7.4, 20 mM $CaCl_2$. This short incubation prevented interference by trypsin autolysis but still allowed sufficient binding of the inhibitor. After addition of 780 µl 0.1 M Tris/HCl pH 7.4, 20 mM $CaCl_2$ containing 10 or 20 µl substrate (10 mM), protease activity was monitored at 405 nm and 22°C in a Shimadzu UV-260 spectrophotometer for 5–10 min. K_i values were then determined according to the Mi-
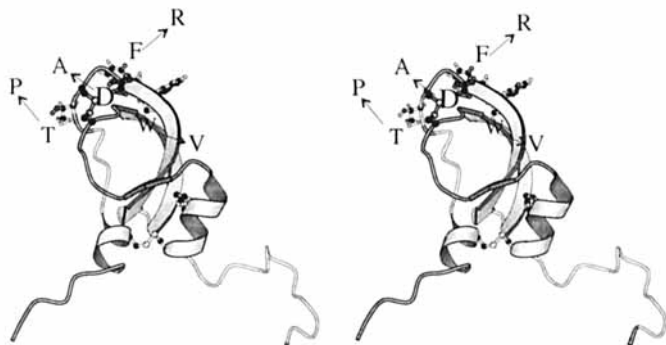
Fig. 1. Ribbon diagram of recombinant domain C5 and identification of the four residues changed by site-directed mutations. These residues (T2886, D2889, F2890, W2907) are shown with their side chains. Arrows denote the substitutions introduced. P1 denotes the scissile amide bond between Arg2888 (P1) and Asp2889(P1′). The structure is based on NMR data (Zweckstetter et al., 1996).

chaelis-Menton equation for competitive inhibition (Todhunter, 1979):

$$K_i = \frac{[I]}{\alpha - 1} \qquad \alpha = \left(\frac{V_{max}}{v_o} - 1\right) \cdot \frac{[S]}{K_m}.$$

Analytical methods. Concentrations of protein solutions were determined on a LC 3000 amino acid analyzer (Biotronik) after hydrolysis (6 M HCl, 16 h, 110°C). SDS gel electrophoresis and ELISA followed standard protocols. The rabbit antisera against domain C5 have been previously described (Mayer et al., 1994). Electrospray ionisation mass spectrometry was performed following established methods (Covey et al., 1988; Fenn et al., 1989).

Molecular modelling. The model molecular dynamics calculations of C5 were performed in the presence of NMR constraints with the program X-PLOR 3.1 (Brünger, 1993) using protocols described previously (Zweckstetter et al., 1996; Holak et al., 1989). Starting structures were the final structures obtained by Zweckstetter et al. (1996). The ϕ/ψ angles of the P4 to P4′ residues of C5 were set to the values found in aprotinin in the present calculations, with tight boundaries of ±10°. No violations of distance constraints larger than 0.06 nm were observed in the new structures. Surface potentials were calculated using DELPHI (Gilson and Honig, 1987) and displayed using the program GRASP (Nicholls et al., 1991).

RESULTS

Production and properties of site-directed mutants of domain C5. Recombinant domain C5 from the collagen VI $\alpha 3$ chain (positions 2869−2943) was previously shown to lack inhibitory activity for several proteases including some with trypsin-like specificity (Mayer et al., 1994). This was interpreted to be due to the presence of unfavorable residues in the P3, P1′ and P2′ position of the binding loop and a bulky Trp in the adjacent β strand (Fig. 1) which are atypical for Kunitz inhibitor modules. We have now prepared seven C5 mutants including a T→P (P3), a D→A (P1′), a F→R (P2′) and a W→V mutation (Fig. 1) in order to increase the similarity of C5 to the trypsin inhibitor aprotinin (Laskowski and Kato, 1980) at several critical sites. They were designed to introduce either one, two or three amino acid substitutions into C5 (Table 1). These mutants will be referred to in the following by the one-letter code for the amino acids exchanged. All mutants could be produced in an episomal expression system which ensured their proper folding

Table 1. Identification of the C5 mutants used and analysis of their molecular mass (*m*) by mass spectrometry. *m* was calculated for the disulfide-bonded form.

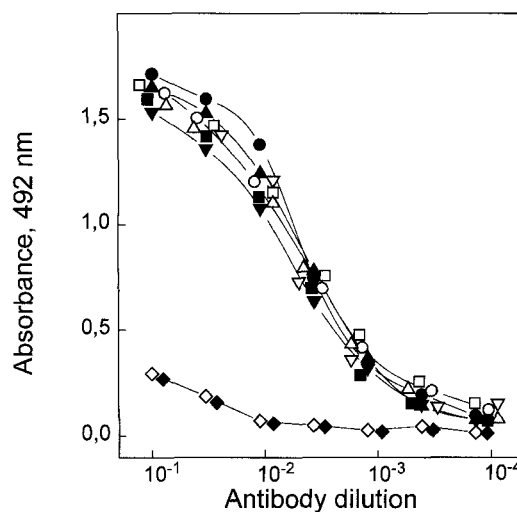| Mutant	*m*	
	calculated	found
	Da	
T→P	8 341.57	8 341.5
D→A	8 301.54	8 301.7
W→V	8 258.48	8 257.6
T→P/D→A	8 297.55	8 298.0
D→A/F→R	8 310.55	8 308.8
T→P/D→A/W→V	8 210.48	8 209.8
T→P/D→A/F→R	8 306.57	8 306.2

Fig. 2. ELISA binding of an antiserum against domain C5 to various C5 mutants. Immobilized antigens were domain C5 (●), mutant W→V (■), mutant T→P (▲), mutant D→A (▼), mutant T→P/D→A (○), mutant D→A/F→R (□), mutant T→P/D→A/F→R (△), mutant T→P/D→A/W→V (▽), aprotinin (◇) and serum albumin as a negative control (◆).

and secretion from transfected mammalian cells. They were purified in two chromatographic steps and obtained in 1−2-mg yields from 1 l serum-free culture medium.

Electrospray mass spectrometry demonstrated a molecular mass for the mutants consistent with the formation of three disulfide bridges (Table 1). The purified mutants showed a major band of about 8−10 kDa in SDS gel electrophoresis (data not shown but see Fig. 4) as found before for recombinant C5 (Mayer et al., 1994). As a conformational probe, we used antibodies against C5 which, as previously shown, failed to bind unfolded C5 (Mayer et al., 1994). ELISA titration demonstrated no or only little differences between the mutants and C5 (Fig. 2) indicating no substantial loss of conformational epitopes. Aprotinin, however, which has only about 50% sequence identity, did not bind the antibodies.

Binding to trypsin in surface plasmon resonance assay. This assay allows kinetic measurements in real time and the determination of K_d values based on association and dissociation rates (Fägerstam et al., 1992). Domain C5 and its single-site mutants T→P and W→V showed no binding to immobilized trypsin indicating a K_d distinctly larger than 1 µM. All other mutants
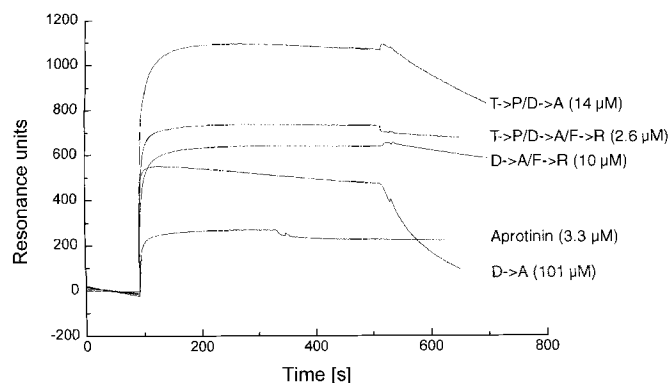
Fig. 3. Overlay plot of sensorgrams of the interaction of C5 mutants and aprotinin with trypsin obtained by surface plasmon resonance assay. The mutants are identified by the letters as described in Table 1. The ligands were injected at about 100s. The dissociation phase started at about 500s, except for aprotinin (330s).

Table 2. Kinetic parameters and dissociation equilibrium constants (K_d) of the binding of C5 mutants to immobilized trypsin. Binding was measured by surface plasmon resonance at the concentrations indicated.

Soluble ligand	Concn	k_{ass}	k_{diss}	K_d
	μM	$M^{-1} s^{-1}$	s^{-1}	nM
C5	7	no binding		
T→P	12	no binding		
W→V	6	no binding		
D→A	0.06−4.1	3.0×10^4	7.5×10^{-3}	250
T→P/D→A	14.1	1.6×10^4	1.8×10^{-4}	11
D→A/F→R	10.3	2.5×10^4	7.5×10^{-5}	2.8
T→P/D→A/W→V	13.1	8.6×10^4	4.0×10^{-6}	0.046
T→P/D→A/F→R	2.6	1.2×10^5	2.7×10^{-6}	0.023
Aprotinin	3.3	1.1×10^5	3.4×10^{-6}	0.031

showed a comparably fast association which reached saturation levels after 2.5−5 min but differed considerably in the speed of dissociation as illustrated by some typical sensorgrams in Fig. 3. This was confirmed by a quantitative determination of rate constants (Table 2) which showed only a threefold difference in the lowest and highest association rates but a high variation over three orders of magnitude for the dissociation rates. The latter variation results in a similar range of affinities for the mutants (Table 2). The affinity was lowest ($K_d = 250$ nM) for mutant D→A but increased 25−100-fold by adding a second substitution in the mutants T→P/D→A and D→A/F→R. A further increase by two orders of magnitude was achieved by adding a third mutation to mutant T→P/D→A in one of two other positions (mutants T→P/D→A/W→V and T→P/D→A/F→R). These triple mutants had rate constants and affinities indistinguishable from those of aprotinin within the limits of analytical error.

Inhibitory properties of C5 mutants for proteases. The inhibitory activity of C5 mutants was analyzed with several proteases to which they were added in 5-fold or 100-fold molar excess prior to the addition of nidogen as a high-molecular-mass, highly protease-sensitive substrate. The digestion patterns were then examined at different time points by SDS gel electrophoresis (Fig. 4). No inhibition of trypsin was observed for domain C5 and mutants W→V or T→P at either inhibitor/protease ratio while complete inhibition occurred with mutants T→P/D→A/
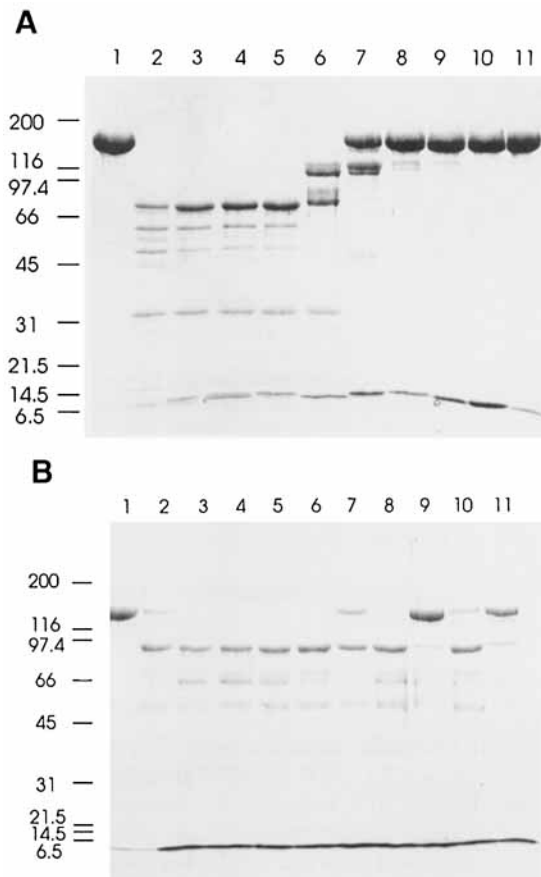

Fig. 4. Effects of C5 mutants on trypsin (A) and leucocyte elastase (B) cleavage of nidogen. Digestion was carried out for 4 h at 37°C and examined by SDS gel electrophoresis. Lane 1, undigested nidogen control; lane 2, protease digest of nidogen without inhibitor; lanes 3−11, identical digests to lane 2 in the presence of C5 (lane 3), mutant W→V (lane 4), mutant T→P (lane 5), mutant D→A (lane 6), mutant T→P/D→A (lane 7), mutant D→A/F→R (lane 8), mutant T→P/D→A/W→V (lane 9), mutant T→P/D→A/F→R (lane 10) and aprotinin (lane 11). Inhibitors at 5-fold (trypsin) or 100-fold (elastase) molar excess were incubated with the proteases prior to the addition of nidogen. The gels were calibrated with marker proteins denoted by their molecular mass (in kDa). The extra band around 8 kDa in lanes 3−11 is due to the added inhibitors.

W→V and T→P/D→A/F→R and aprotinin at 5-fold inhibitor excess. Mutants D→A, T→P/D→A and D→A/F→R showed only partial inhibition which increased in accordance with their K_d values (Fig. 4A).

Leucocyte elastase with a K_d of about 10 μM for aprotinin (Lestienne and Bieth, 1978) was used as a second, low-affinity protease to analyze the activity of the mutants in the nidogen cleavage assay (Fig. 4B). Aprotinin showed a distinct but only partial inhibition which is explained by its concentration being only fourfold higher than the K_d value. A similar activity was only observed for mutant T→P/D→A/W→V. Mutants T→P/D→A/F→R and T→P/D→A showed weaker activities and none was found for C5 or any other mutants. Thus the inhibitory activities of the mutants show distinct differences between leucocyte elastase and trypsin (Fig. 4A). A further difference was noted for thrombin, which readily cleaves nidogen into a 100-kDa and a 50-kDa fragment (Mayer et al., 1993). When used at the same concentrations as in the leucocyte elastase assay neither aprotinin, C5 nor any of its mutants showed inhibitory activity (data not shown). Cleavage could, however, be completely inhibited by a fivefold molar excess of rhodniin.

Table 3. Inhibitory constants (K_i) of C5 mutants for trypsin cleavage of a small peptide substrate. Values were determined at substrate concentrations of 0.125 mM and 0.25 mM using 1.3 nM trypsin and inhibitors in the concentration range 3.8—30.1 nM.

Inhibitor	K_i at [S]:	
	0.125 mM	0.25 mM
	nM	
C5	no inhibition	
T → P	no inhibition	
W → V	no inhibition	
D → A	48	480
T → P/D → A	7.7	7.7
D → A/F → R	0.32	0.32
T → P/D → A/W → V	0.38	0.17
T → P/D → A/F → R	0.12	0.13
Aprotinin	0.13	0.079

Table 4. Reactive-site loop conformation of domain C5 and aprotinin. The data of aprotinin were taken from Deisenhofer and Steigemann (1975) and those of domain C5 from Zweckstetter et al. (1996). Asterisks denote in C5 residues which were changed in the various mutants. For definition of the backbone angles ϕ/ψ, see Branden and Tooze (1991) page 8.

Loop position	Domain C5		Aprotinin	
	residue	ϕ/ψ angles	residue	ϕ/ψ angles
		deg.		deg.
P5	Glu	58± 6/ 12± 6	Thr	87/ 170
P4	Gly	85± 2/ 30± 3	Gly	78/ 178
P3	Thr *	76± 2/ 31±22	Pro	−77/−32
P2	Cys	−97±34/−160±10	Cys	−70/ 154
P1	Arg	−114± 5/ 42± 8	Lys	−120/ 49
P1'	Asp *	−120± 5/ 55± 4	Ala	−94/ 170
P2'	Phe *	50±10/ 120±20	Arg	−112/ 79
P3'	Ile	−130±40/ 140±10	Ile	−105/ 120

A small peptide substrate with a single trypsin cleavage site was used for comparison of K_i and K_d values. Two substrate concentrations were chosen which were 1.5 or 3 times higher than the K_m value of the reaction (0.0824 mM), and showed identical K_i values within the limits of analytical error except with mutant D→A (Table 3). The data demonstrated no activity for domain C5 and its mutants T→P and W→V. The K_i values for the other mutants and aprotinin agreed within one order of magnitude with the equilibrium dissociation constants determined by the surface plasmon resonance assay (Tables 2 and 3). This demonstrated that direct binding and inhibitory capacity are determined by the dissociation rate constants. This was also consistent with the nature of the Lineweaver-Burk plots (not shown) which demonstrated competitive inhibition for all active C5 mutants.

Molecular modelling of the C5 binding loop. The present NMR structure calculation of C5 was basically performed in an analogous manner to that previously described (Zweckstetter et al., 1996). The only difference was the introduction of additional constraints for dihedral angles in the binding loop which simulated its conformation in a complex with trypsin. Two assumptions were made: first, that the binding loop was less variable upon binding to trypsin; second, that the conformation of the binding loop was canonical, i.e. the ϕ/ψ angles of P4 to P4' residues of C5 were set to those found in aprotinin (Table 4). The best superposition of P5 to P5' residues of the adjusted C5 and aprotinin in a simulated or actual complex with trypsin is shown in Fig. 5. It can be seen from Fig. 5 that it was possible to adjust the conformation of the backbone of the binding loop to the canonical conformation of aprotinin without significant violation of NMR constraints. In addition, most of the side chains in such an adjusted loop pointed in the same directions as those of aprotinin. The notable exception was Phe (P2' site), for which the χ_1 angle was 180° relative to that of the Arg in P2' of aprotinin. In this conformation, the aromatic ring clashes with residues Gln192 and Gly193 of trypsin (for example, the distance C5 Phe Cγ to trypsin Gln192 C was only 0.1 nm). If the χ_1 angle were to be set to the value found in P2'(Arg) of aprotinin, a number of NOEs between protons of the ring of Phe and neighbouring or distant residues of C5 would be violated. These NOEs must be removed during the calculations for the ring to flip into the χ_1 orientation found in aprotinin. Even if such NOE contacts did not exist in the hypothetical complex of
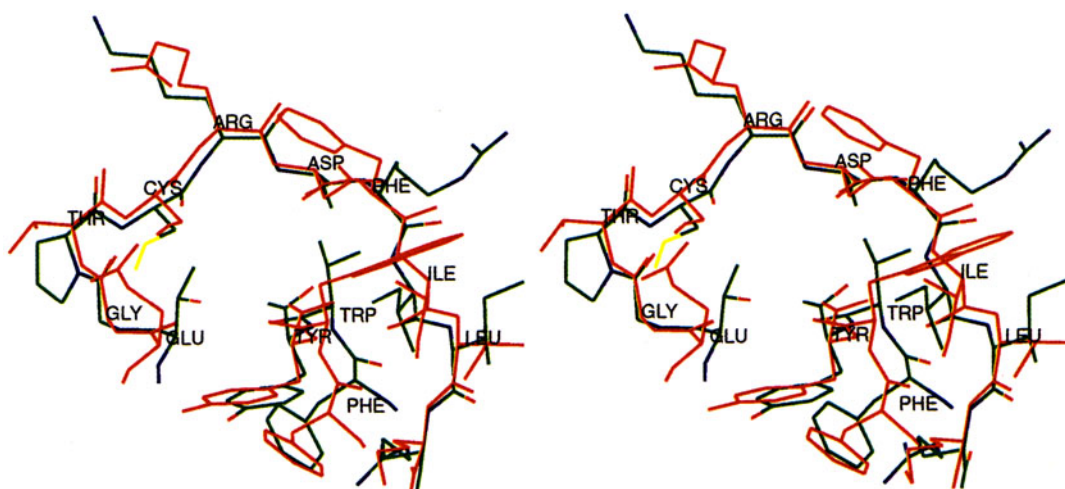

Fig. 5. Superposition of the binding loops and part of the first β strand of the modelled structure of domain C5 and of aprotinin in a complex with trypsin. The structure of C5 is shown in red and its residues are denoted in black. The aprotinin structure is multicolored showing carbons in green, nitrogens in blue oxygens in red and disulfide bonds in yellow. The β strand shows the surrounding of the tryptophan (W) which was subjected to mutation in C5 (W→V).
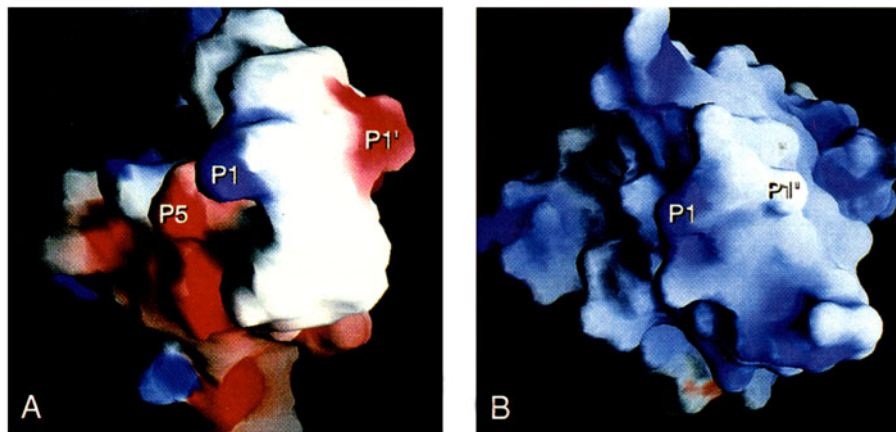
Fig. 6. Electrostatic potential surfaces of domain C5 (A) and aprotinin (B). The inhibitors are viewed from the top of their binding loops and a few critical positions are marked (P). Blue regions represent positive, red negative and white neutral potentials.

trypsin-C5, the hydrophobic Phe at P2' in C5 does not permit hydrogen bonding found in aprotinin, where imino protons of the side chain of Arg make bonds to oxygen donors of trypsin.

DISCUSSION

Domain C5 at the C-terminus of human type VI collagen $\alpha3$(VI) chain corresponds to a Kunitz-type module and some additional residues at the C-terminal (Chu et al., 1990). As shown by immunological analyses, this domain is immediately released probably by proteolytic processing after assembly and secretion of type VI collagen has been accomplished (Mayer et al., 1994). The released form persists in the circulation but could not be detected in the extracellular matrix. Recombinant domain C5 was further shown to lack inhibitory activity for trypsin and several other proteases despite having an Arg in the P1 position of the binding loop. The biological function of domain C5 therefore remains to be established but the possibility remains that it could be an inhibitor of proteases specialized for certain purposes.

The lack of predicted inhibitory function led us to examine the structural basis for this observation by site-directed mutagenesis. The mutations were restricted to four residues based on sequence comparisons (Mayer et al., 1994) which retrospectively were also in agreement with predictions derived from the X-ray structure of domain C5 (Arnoux et al., 1995). Three mutations were localized in the binding loop and one in the adjacent β strand (Fig. 1). Single mutations (T→P, D→A, W→V), except for D2889-A, showed no significant increase in affinity compared to domain C5 in direct binding and inhibition assays with trypsin. A combination of two mutations (T→P/D→A, D→A/F→R) or of three mutations (T→P/D→A/F→R, T→P/D→A/W→V) showed in each case a considerable increase in affinity by about two orders of magnitude. Yet it was surprising that two different substitutions in the triple mutants, either in the binding region (F→R) or the β strand (W→V), produced the same effects. Thus a very limited change in the sequence identity of domain C5 and aprotinin, which does not exceed 50%, caused a crucial change in the functional potential.

Most of the effects caused by mutations of C5 can now be interpreted in terms of the tertiary structure in solution which was recently completed (Zweckstetter et al., 1996). This demonstrated that the 57-residue core of domain C5 has the characteristic fold of all members of Kunitz-type protease inhibitors (Fig. 1). The NMR study, however, also showed a highly dynamic structure for C5 where about 44% of the core residues can

exist in multiple conformations. Such sites included a substantial portion of the binding loop which is defined by the P3, P2, P1, P1', P2' and P3' sites around the scissile peptide bond between P1 and P1' (see Table 4) and three of the four positions chosen for mutations (D→A, F→R, W→V). A similar variability of the binding loop was also found in NMR studies of other protease inhibitors such as aprotinin (Wagner et al., 1987; Otting et al., 1993), leech-derived tryptase inhibitor (Mühlhahn et al., 1994), the trypsin inhibitor from squash seeds (Holak et al., 1989) and eglin c (Hyberts et al., 1992). Thus lack of trypsin inhibition by domain C5, despite a favorable Arg at the P1 position, cannot be accounted for by conformational variability and very likely arises by the specific nature of the side chains of a few critical residues within or close to the binding loop. This is particularly evident for Asp at the P1' site of domain C5, which is usually occupied by Ala or Gly in other inhibitors. This exposes a negative charge on the surface of C5 not found in aprotinin (Fig. 6) which should be incompatible with the negatively charged binding cleft of trypsin and, in addition, causes steric interference with His57 of trypsin. These predictions were verified with the D→A mutant, which was the only single C5 substitution yielding a measurable trypsin binding activity. Thus Asp seems to be the most crucial residue in the binding loop which determines the inhibitory activity for trypsin-like proteases.

Other substitutions worked only in the context of an existing D→A mutation. These included a T→P mutation at the P3 site which caused a 25-fold improvement in activity. The Pro at the aprotinin P3 site forms a main-chain hydrogen bond to Gly216 of trypsin whereas, if it is changed to Thr in C5, it probably forms hydrogen bonds to a spatially adjacent Cys in C5 as shown by NMR analysis at pH 2.7. This is in agreement with the crystal structure of domain C5, which demonstrated that the direction of the carbonyl oxygen of Thr is opposite to that of Pro in aprotinin and is dependent on a different chirality of the adjacent disulfide bridge (Arnoux et al., 1995). Apparently, the formation of either an inter- or intra-molecular hydrogen bond as a consequence of the rotation of the carbonyl group is the most likely interpretation of our experimental observation.

A more complex situation seems to exist for the Phe at the P2' site of C5, which, when changed to Arg as in aprotinin, also improves binding activity. The aromatic ring seems to be fixed into a hydrophobic pocket consisting of F2890, W2907 and G2909 (Fig. 5). Furthermore, the presence of a hydrophobic residue at the P2' site should show severe steric hindrance with Gln192 and Gly193 of trypsin and abolish side-chain hydrogen bonds of the Arg imino protons to oxygen donors of trypsin. In

order to relieve these constraints, the side chain of F2890 needs to flip into the χ_1 orientation as found in aprotinin. This seems, however, not possible in domain C5 because of unfavorable van der Waals contacts with the bulky ring of W2907, which is located on the neighbouring β strand just below the binding loop (Figs 1 and 5). This dualism provides a perfect explanation as to why two different third mutations (F→R, W→V) to the T→P/D→A mutant structure cause the same improvement in trypsin inhibitory activity which is then equivalent to aprotinin.

Aprotinin is also a weak inhibitor of leucocyte elastase (Lestienne and Bieth, 1978), in spite of a Lys in the P1 position which, when changed to Val, causes an increase in activity by five orders of magnitude (Wenzel et al., 1986). Several of the C5 mutants also showed inhibitory activity but it did not fully correlate with the relative increases in trypsin affinity when compared in a semiquantitative assay. Most remarkable was the identical activity observed for mutant T→P/D→A/W→V and aprotinin while both mutants T→P/D→A/F→R and T→P/D→A had a similar lower activity. This indicates distinct differences in the binding sites of the two proteases which remain to be fully determined.

REFERENCES

Arnoux, B., Merigeau, K., Saludjian, P., Norris, F., Norris, K., Björn, S., Olsen, O., Peterson, L. & Ducruix, A. (1995) The 1.6 Å structure of Kunitz-type domain from the α3 chain of human type VI collagen, *J. Mol. Biol. 246*, 609−617.

Bode, W. & Huber, R. (1992) Natural proteinase inhibitors and their interaction with proteinases, *Eur. J. Biochem. 204*, 433−451.

Bonaldo, P. & Colombatti, A. (1989) The carboxyl terminus of the chicken α3 chain of collagen VI is a unique mosaic structure with glycoprotein Ib-like, fibronectin type III, and Kunitz modules, *J. Biol. Chem. 264*, 20235−20239.

Branden, C. & Tooze, J. (1991) *Introduction to protein structure*, Garland Publishing, New York.

Brünger, A. T. (1993) *XPLOR version 3.1 manual*, Yale University, New Haven CT.

Castro, M., Marks, C. B., Nilsson, B. & Anderson, S. (1990) Does the Kunitz domain from the Alzheimer's amyloid β protein precursor inhibit a kallikrein responsible for post-translational processing of nerve growth factor precursor? *FEBS Lett. 267*, 207−212.

Chu, M.-L., Zhang, R.-Z., Pan, T., Stokes, D., Conway, D., Kuo, H.-J., Glanville, R., Mayer, U., Mann, K., Deutzmann, R. & Timpl, R. (1990) Mosaic structure of globular domains in the human type VI collagen α3 chain: similarity to von Willebrand factor, fibronectin, actin, salivary proteins and aprotinin type protease inhibitors, *EMBO J. 9*, 385−393.

Covey, T. R., Bronner, R. F., Shushan, B. I. & Henion, J. (1988) The determination of protein, oligonucleotide and peptide molecular masses by ion-spray mass spectrometry, *Rapid Commun. Mass Spectrom. 2*, 249−256.

Deisenhofer, J. & Steigemann, W. (1975) Crystallographic refinement of the structure of bovine pancreatic trypsin inhibitor at 1.5 Å resolution, *Acta Crystallogr. B 31*, 238−250.

Fägerstam, L. G., Frostell-Karlsson, A., Karlsson, R., Persson, B. & Rönnberg, I. (1992) Biospecific interaction analysis using surface plasmon resonance detection applied to kinetic, binding site and concentration analysis, *J. Chromatogr. 597*, 397−410.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules, *Science 246*, 64−71.

Fox, J. W., Mayer, U., Nischt, R., Aumailley, M., Reinhardt, D., Wiedemann, H., Mann, K., Timpl, R., Krieg, T., Engel, J. & Chu, M.-L.

(1991) Recombinant nidogen consists of three globular domains and mediates binding of laminin to collagen type IV, *EMBO J. 10*, 3137−3146.

Friedrich, T., Kroeger, B., Bialojan, S., Lemaire, H. G., Hoeffken, H. W., Reuschenbach, P., Otte, M. & Dodt, J. (1993) A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*, *J. Biol. Chem. 268*, 16216−16222.

Gilson, M. & Honig, B (1987) Calculation of electrostatic potential in an enzyme active site, *Nature 330*, 84−86.

Greenspan, D. S. (1993) The carboxy-terminal half of type VII collagen, including the non collagenous NC-2 domain and intron/exon organization of the corresponding region of the Col7A1 gene, *Hum. Mol. Genet. 2*, 273−278.

Holak, T. A., Gondol, D., Otlewski, J. & Wilusz, T. (1989) Determination of the complete 3-dimensional structure of the trypsin-inhibitor from squash seeds in aqueous-solution by nuclear magnetic-resonance and a combination of distance geometry and dynamical simulated annealing, *J. Mol. Biol. 210*, 635−648.

Hyberts, S. G., Goldberg, M. S., Havel, T. F. & Wagner, G. (1992) The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures, *Protein Sci. 1*, 736−751.

Hynes, T. R., Randal, M., Kennedy, L. A., Eigenbrot, C. & Kossiakoff, A. A. (1990) X-ray crystal structure of the protease inhibitor domain of Alzheimer's amyloid β-protein precursor, *Biochemistry 29*, 10018−10022.

Kido, H., Fukutomi, A., Schilling, J., Wang, Y., Cordell, B. & Katunuma, N. (1990) Protease-specificity of Kunitz inhibitor domain of Alzheimer's disease amyloid protein precursor, *Biochem. Biophys. Res. Commun. 167*, 716−721.

Kitaguchi, N., Takahashi, Y., Tokushima, Y., Shiojiris, S. & Ito, H. (1988) Novel precursor of Alzheimer's disease amyloid protein shows protease inhibitory activity, *Nature 331*, 530−532.

Laskowski, M. & Kato, I. (1980) Protein inhibitors of proteinases, *Annu. Rev. Biochem. 49*, 593−626.

Lestienne, P. & Bieth, J. G. (1978) The inhibition of human leucocyte elastase by basic pancreatic trypsin inhibitor, *Arch. Biochem. Biophys. 190*, 358−360.

Mayer, U., Aumailley, M., Mann, K., Timpl, R. & Engel, J. (1991) Calcium-dependent binding of basement membrane protein BM-40 (osteonectin, SPARC) to basement membrane collagen type IV, *Eur. J. Biochem. 198*, 141−150.

Mayer, U., Mann, K., Timpl, R. & Murphy, G. (1993) Sites of nidogen cleavage by proteases involved in tissue homeostasis and remodelling, *Eur. J. Biochem. 217*, 877−884.

Mayer, U., Pöschl, E., Nischt, R., Specks, U., Pan, T.-C., Chu, M.-L. & Timpl, R. (1994) Recombinant expression and properties of the Kunitz-type protease inhibitor module from human type VI collagen α3(VI) chain, *Eur. J. Biochem. 225*, 573−580.

Mühlhahn, P., Czisch, M., Morenweiser, R., Habermann, B., Engh, R. A., Sommerhoff, C. P., Auerswald, E. A. & Holak, T. A. (1994) Structure of leech-derived tryptase inhibitor (LDTI-C) in solution, *FEBS Lett. 355*, 290−296.

Nicholls, A., Sharp, K. & Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins 11*, 281−296.

Nischt, R., Pottgiesser, J., Krieg, T., Mayer, U., Aumailley, M. & Timpl, R. (1991) Recombinant expression and properties of the human calcium-binding extracellular matrix protein BM-40, *Eur. J. Biochem. 200*, 529−536.

Oltersdorf, T., Fritz, L. C., Schenk, D. B., Lieberburg, I., Johnson-Wood, K. L., Beatti, E. C., Ward, P. J., Blacher, R. W., Dovey, H. F. & Sinha, S. (1989) The secreted form of the Alzheimer's amyloid precursor protein with the Kunitz domain is protease nexin-II, *Nature 341*, 144−147.

Otting, G., Liepinsh, E. & Wüthrich, K. (1993) Disulfide bond isomerization in BPTI and BPTI(G36S): an NMR study of correlated mobility in proteins, *Biochemistry 32*, 3571−3582.

Oyama, F., Hashino, K., Oyama, R., Kato, I. & Titani, K. (1993) Improved method for expression of Kunitz-type serine proteinase inhibitor domain of β-amyloid protein precursor in *Escherichia coli* and characterization of disulfide bonds of the product, *J. Biochem. (Tokyo) 114*, 813−819.

Ponte, P., Gonzales-DeWhitt, P., Schilling, J., Miller, J., Hsu, D., Greenberg, B., Davis, K., Wallace, W., Lieberburg, I., Fuller, F. & Cordell, B. (1988) A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibitors, *Nature 331*, 525−527.

Tanzi, R. E., McClatchey, A. I., Lamperti, E. D., Villaa-Komaroff, L., Gusella, J. F. & Neve, R. L. (1988) Protease inhibitor domain encoded by an amyloid protein precursor mRNA associated with Alzheimer's disease, *Nature 331*, 528−330.

Todhunter, J. A. (1979) Reversible enzyme inhibition, *Methods Enzymol. 63*, 383−411.

Van Nostrand, W. E., Schmaier, A. H., Neiditch, B. R., Siegel, R. S., Raschke, W. C., Sisodia, S. S. & Wagner, S. L. (1994) Expression, purification, and characterization of the Kunitz-type proteinase inhibitor domain of the amyloid β-protein precursor-like protein-2, *Biochim. Biophys. Acta 1209*, 165−170.

Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N. & Wüthrich, K. (1987) Protein structures in solution by NMR and distance geometry. The polypeptide fold of the basic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN, *J. Mol. Biol. 196*, 611−639.

Wagner, S. L., Siegel, R. S., Vedvick, T. S., Raschke, W. C. & Van Nostrand, W. E. (1992) High level expression, purification, and characterization of the Kunitz-type protease inhibitor domain of protease nexin-2/amyloid β-protein precursor, *Biochem. Biophys. Res. Commun. 186*, 1138−1145.

Wenzel, H. R., Beckman, J., Mehlich, A., Schnabel, E. & Tschesche, H. (1986). Semisynthetic conversion of the bovine trypsin inhibitor (Kunitz) into an efficient leukocyte-elastase inhibitor by specific valine for lysine substitution in the reactive site, in *Chemistry of peptides and proteins*, vol. 3 (W. Voelter, Y. A. Orchinnikov & E. Bayer, eds) pp. 105−116, Walter de Gruyter, Berlin.

Zweckstetter, M., Czisch, M., Mayer, U., Chu, M.-L., Timpl, R. & Holak, T. A. (1996) Structure and multiple conformations of the Kunitz-type domain from human type VI collagen α3(VI) chain in solution, *Structure* 4, 195−209.