Role of the Intramolecular Hydrogen Bond Network in the Inhibitory Power of Chymotrypsin Inhibitor 2[†,‡]

Evette S. Radisky, Chia-Jung Karen Lu, Gene Kwan,[§] and Daniel E. Koshland, Jr.*

Department of Molecular and Cell Biology, University of California, Berkeley, California 94720

Received December 22, 2004; Revised Manuscript Received March 9, 2005

ABSTRACT: A series of mutants of chymotrypsin inhibitor 2 (CI2), at residues involved in intramolecular interactions that shape and constrain the binding loop, were studied to determine their relative importance for inhibition of the serine protease subtilisin BPN′, and for resistance of the inhibitor to proteolysis. These functional properties were investigated in tandem with the crystal structures of the mutant inhibitor−enzyme complexes. A dense hydrogen bonding network that supports the binding loop in the vicinity of the scissile bond was found to be important both for enzyme affinity and for stability to proteolysis. Structural analysis, in combination with biochemical measurements, allows differentiation of the structural components most important for resistance to proteolysis and/or binding. The most critical participating residues in the network were found to be Thr-58, Glu-60, Arg-65, and Gly-83. Glu-60 is more important for resistance to proteolysis than for binding, while Arg-65 and two other Arg residues play a greater role in binding than in resistance to proteolysis. Structural comparisons reveal a wide variety of subtle conformational changes in response to mutation, with built-in robustness in the hydrogen bond network, such that loss of one contact is compensated by other new contacts.

Chymotrypsin inhibitor 2 (CI2)[1] inhibits the serine protease subtilisin by binding in the active site with a dissociation constant of 3.0×10^{-12}, and resisting proteolysis (*3*, *4*). It was previously thought that inhibition by CI2 and a large class of similar protein inhibitors might be caused by a barrier to reaction, perhaps due to extreme rigidity at the enzyme−inhibitor interface or misalignment of reacting groups. Our earlier work has shown, however, that CI2 acylates subtilisin BPN′ rapidly, but that further reaction is retarded because peptide religation is favored over hydrolytic attack on the acyl−enzyme complex; in Scheme 1, k_{-2} is much greater than k_3, resulting in an apparent equilibrium between the Michaelis complex and the acyl−enzyme complex (*3*). We refer to this phenomenon as the clogged gutter mechanism, in which the tight and oriented binding of the leaving group peptide (H_2N-R′ in Scheme 1) prevents acyl−enzyme hydrolysis and product dissociation, favoring the back reaction (*3*). The root of the inhibitory power of CI2 therefore lies in the resistance of the acyl−enzyme complex to hydrolysis. Our previous studies have found that side chain interactions between CI2 binding loop residues and subtilisin contribute only modestly to proteolytic resistance (*5*), while interactions with the CI2 protein core that stabilize the conformation of the binding loop backbone are the primary determinants of this resistance (*6*). Here we investigate intramolecular contacts between the CI2 core and the binding loop that control binding loop conformation and mediate inhibitory function.

Examination of the structure of the CI2−subtilisin BPN′ complex reveals that the binding loop of CI2 protrudes from the globular hydrophobic core of the inhibitor, supported and conformationally constrained by a grouping of polar and charged residues forming a dense network of hydrogen bonds and electrostatic interactions [the H-bond network (*3*, *6*−*8*)] (Figure 1). These residues serve to facilitate enzyme−inhibitor recognition by shaping the binding loop backbone into a conformation complementary to the enzyme active site. Residues of the H-bond network flank the cleavage site targeted by subtilisin, leading us to speculate that this network may play a key role in the retention and orientation of the leaving group peptide following formation of the acyl−enzyme complex, preventing hydrolysis and favoring religation.

[†] This work was supported in part by National Institutes of Health Grant DK09765 (to D.E.K.). E.S.R. was supported by a National Institutes of Health postdoctoral fellowship. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of U.S. Department of Energy, under Contract DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory.

[‡] Structure coordinates have been deposited with the Protein Data Bank as entries 1Y1K, 1Y33, 1Y34, 1Y3B, 1Y4A, 1Y4D, 1Y3C, 1Y48, and 1Y3D.

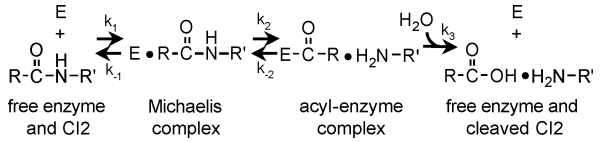
* To whom correspondence should be addressed. E-mail: dek@uclink.berkeley.edu. Phone: (510) 642-0416. Fax: (510) 643-6386.

[§] Present address: University of California at San Diego School of Medicine, La Jolla, CA 92093.

[1] Abbreviations: CI2, chymotrypsin inhibitor 2; PCR, polymerase chain reaction; HPLC, high-pressure liquid chromatography; PEG, poly(ethylene glycol); CCD, charge-coupled device; rmsd, root-mean-square deviation. CI2 mutant nomenclature follows this example, using the single-letter amino acid code: T58A refers to residue 58, formerly threonine, now mutated to alanine. G83Δ refers to a CI2 mutant in which the C-terminal residue, Gly-83, has been deleted. Throughout the text, the designation of individual protein atoms follows the standard nomenclature rules of the Protein Data Bank (*1*). Throughout the text, we use the nomenclature of Schechter and Berger (*2*) for designating substrate residues surrounding the cleavage site, and the corresponding binding subsites on the enzyme. Starting at the scissile bond, substrate residues are numbered P_1, P_2, P_3, etc., in the direction of the N-terminus and P_1', P_2', P_3', etc., in the direction of the C-terminus. The corresponding enzyme subsite binding pockets are numbered S1, S2, etc.

10.1021/bi047301w CCC: $30.25
Published on Web 04/16/2005

Scheme 1: Mechanism of Cleavage of CI2 by Subtilisin[a]



[a] Subtilisin is represented by E, and the N- and C-terminal protein fragments of CI2 are depicted as R and R′, respectively. The scissile peptide bond targeted by the enzyme links the R fragment (residues 1−59) and the R′ fragment (residues 60−83).

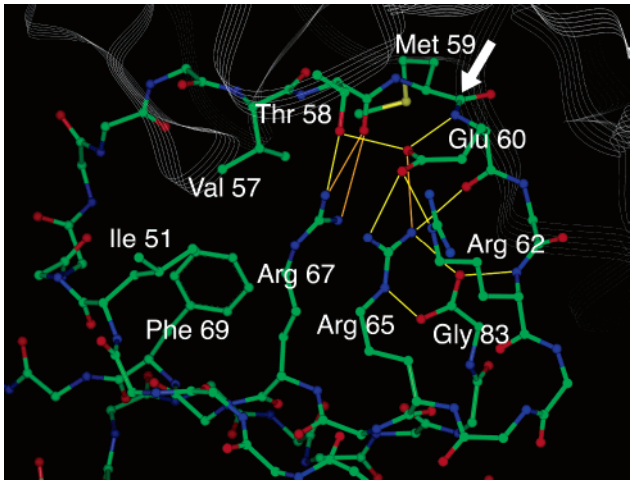


FIGURE 1: Binding loop of CI2 in complex with subtilisin BPN′. The CI2 backbone and relevant side chain atoms are colored green (carbons), red (oxygens), and blue (nitrogens); subtilisin is shown as a white ribbon. A white arrow indicates the targeted cleavage site between CI2 residues 59 and 60. Potential hydrogen bonds and electrostatic interactions of the H-bond network are colored yellow (interatomic distances of 2.5−3.0 Å) and orange (interatomic distances of 3.0−3.5 Å), respectively. Coordinates are from PDB entry 1TM1.

In this paper, we use mutagenesis to analyze the importance of each residue in the H-bond network of CI2, both for the initial binding step, as measured by enzyme−inhibitor dissociation constants, and for the subsequent clogged gutter effect, as measured by overall rates of inhibitor hydrolysis. We mutated each H-bonding side chain to Ala, including Thr-58, Glu-60, Arg-62, Arg-65, and Arg-67, and deleted the H-bonding C-terminal Gly-83. We also constructed three additional mutants, Thr-58 → Pro, Glu-60 → Ser, and the double mutant Met-59 → Arg/Glu-60 → Ser, since analogous mutations of eglin c, a homologue of CI2, have been previously reported to result in more rapid cleavage of that inhibitor (*9*). We report the effects of each of these mutations on the dissociation constants between CI2 and subtilisin, and on the rates of cleavage of CI2 by subtilisin. We also report crystal structures for the CI2 mutants in complex with subtilisin BPN′, allowing us to correlate changes in hydrolysis and binding with structural features.

EXPERIMENTAL PROCEDURES

Subtilisin BPN′. His-tagged recombinant subtilisin BPN′ was prepared as described previously (*6*). Active enzyme concentrations were determined by a kinetic assay with succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide as described previously (*6*).

CI2. Truncated recombinant CI2 was expressed and purified from *Escherichia coli* as described previously (*6*). Throughout the text, we use the amino acid numbering of the original full-length protein sequence. The CI2 expression construct was mutagenized using the PCR-based QuikChange method (Stratagene); the mutant expression constructs were verified by DNA sequencing. Mutant CI2 proteins were expressed and purified following the same protocols used for the wild-type inhibitor (*6*). Wild-type and mutant CI2 concentrations were determined by peak integration from HPLC absorbance traces and by titration with subtilisin BPN′ at a known concentration, as described previously (*6*).

Subtilisin Inhibition (binding) Studies. CI2 is a potent slow, tight-binding inhibitor of subtilisin (*4*). Inhibition constants (K_i) were calculated from the reduced rates of subtilisin hydrolysis of a chromogenic substrate, succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide, in the presence of wild-type or mutant CI2; all assay conditions, procedures, equations, and analyses have been described in detail previously (*6*). The reported K_i values represent mean values from two or more experiments.

Hydrolysis Studies. The initial linear rate of depletion of intact CI2 in time course incubations with subtilisin BPN′ was monitored by HPLC, as described previously (*6*). For each time point, HPLC injections were performed in duplicate, and remaining CI2 quantities were determined by peak integration of absorbance traces and comparison to standard curves generated from known quantities of pure CI2 (*6*).

Crystallization. Most mutant complexes were crystallized by vapor diffusion essentially as described previously (*5*). Lyophilized subtilisin BPN′ and CI2 were each dissolved in 10 mM NaOAc (pH 5.8), and were mixed in a 1:1.2 stoichiometric molar ratio, at a total protein concentration of 4−6 mg/mL. Crystals were grown at 4 °C in hanging drops, typically over a reservoir of 0.1 M sodium citrate (pH 4.6), 20% 2-propanol, and 15% PEG 2000; drops were prepared by mixing equal volumes of protein and reservoir solutions. Variations in the protocol for individual CI2 mutants were as follows: the E60A and R67A CI2 complexes were crystallized with 15% PEG monomethyl ether 750; the R62A CI2 complex was crystallized with 15% PEG 4000; and the E60S CI2 complex was crystallized with 15% PEG 4000, with 4% acetone added to the crystallization drops and reservoir. The M59R/E60S CI2 complex was crystallized in the $P2_1$ crystal form by microseeding a drop containing 1.5 mg/mL total protein, 0.1 M sodium citrate (pH 4.6), 20% 2-propanol, and 15% PEG 2000, with seeds obtained by dounce homogenization of poorer quality M59R/E60S CI2 complex crystals. The M59R/E60S CI2 complex was crystallized in the $P6_1$ crystal form by the hanging drop method over a reservoir of 25% PEG 8000 and 0.05 M potassium phosphate (pH 4.5), at a total protein concentration of 4 mg/mL. The R65A CI2 complex was crystallized by macroseeding into a drop containing 1 mg/mL total protein, 0.1 M sodium citrate (pH 4.6), 20% 2-propanol, and 15% PEG 2000, and with a subtilisin:CI2 stoichiometry of 1:3. The R65A CI2 complex macroseeds were grown by microseeding a similar drop with seeds obtained by dounce homogenization of wild-type complex crystals.

Crystal Data Collection and Processing. For each variant CI2 complex, synchrotron X-ray data were collected from a single crystal at 100 K using an ADSC Quantum 210 CCD detector at Advanced Light Source beamline 5.0.1 or 8.3.1 (Lawrence Berkeley National Laboratory, Berkeley, CA). The automation package ELVES (*10*) was used to direct

Table 1: Effect of CI2 Mutations on Subtilisin BPN′ Inhibition and Hydrolysis

	inhibition		hydrolysis	
	K_i (M)	relative	k_{cat} (s^{-1})	relative
wild type	3.0×10^{-12}	1	3.8×10^{-6}	1
T58A	3.0×10^{-10}	100	3.4×10^{-4}	90
T58P	1.7×10^{-9}	559	3.1×10^{-3}	816
E60A	5.2×10^{-10}	174	1.6×10^{-3}	426
E60S	3.1×10^{-10}	105	8.7×10^{-4}	229
M59R/E60S	1.2×10^{-9}	397	1.4×10^{-3}	374
R62A	2.5×10^{-11}	8	3.8×10^{-6}	1
R65A	1.7×10^{-9}	554	4.4×10^{-4}	116
R67A	5.6×10^{-10}	186	1.2×10^{-4}	32
G83Δ	3.4×10^{-9}	1137	4.0×10^{-3}	1053

MOLFLM (*11*) for indexing and integration and SCALA (*12*) for scaling and merging the reflections. For all structures except that of the M59R/E60S double mutant complex, crystals were of space group $P6_522$ and the same unit cell observed for a previous series of mutant complexes (*5*), and the structure of the M59K CI2−subtilisin BPN′ complex (PDB entry 1TM3) was used as a starting model for refinement. Test sets comprised of 5% of the total reflections, inherited from the M59K CI2 complex, were excluded from refinement to allow calculation of the free *R*-factors. For the M59R/E60S double mutant complex, crystals were of space group $P2_1$ or $P6_1$, with a unit cell not observed previously. The initial models for these structures were obtained by molecular replacement using EPMR (*13*), using as the search model the wild-type CI2−subtilisin complex in space group $P2_12_12_1$ (PDB entry 1LW6), with solvent molecules omitted (*3*). All models were improved through alternating cycles of manual rebuilding using the interactive graphics program O (*14*), automated refinement using REFMAC5 (*15*), and automated solvent addition using ARP/WARP (*16*). Model validation was carried out using PROCHECK (*12*) and WHATCHECK (*17*). Superpositions were performed using GEM (*18*), and structure figures were generated with InsightII (Biosym Technologies).

RESULTS

Inhibition Constants and Hydrolysis Rates of CI2 Mutants Reveal Diverse Effects that Range from Negligible to Greater Than 3 Orders of Magnitude. The data from inhibition and hydrolysis studies with the CI2 mutants are presented in Table 1; the K_i is equal to the dissociation constant K_d, and k_{cat} is the overall rate of turnover. These results are consistent with previous mutational studies of CI2 (*8*, *19*). Among the residues mutated in this study, central participants in the H-bonding network Gly-83, Glu-60, and Thr-58 prove to be the most critical for resistance to hydrolysis, followed by Arg-65. A similar pattern is seen in the effect on inhibition, but with Arg-65 playing a greater role. Mutation of Arg-62, a peripheral participant in the H-bonding network, has comparatively little effect on hydrolysis or inhibition, while mutation of Arg-67 has an intermediate effect.

All Mutants except G83Δ CI2 Were Crystallized in Complex with Subtilisin BPN′. Most of the CI2 mutant complexes were successfully crystallized with only minor variations in conditions employed previously for the CI2−subtilisin complex (*5*). The crystals were rod-shaped with hexagonal cross sections, and belonged to space group $P6_5$-22. However, three of the mutant complexes, those with the R65A, M59R/E60S, and G83Δ CI2 variants, proved to be resistant to crystallization. For these mutants, we extensively screened additional additives and attempted a variety of methods of seeding nucleation. Suspecting that the difficulty might relate to hydrolysis of these CI2 mutants by subtilisin under the conditions of crystallization, we also screened higher ratios of CI2 to subtilisin. For the R65A mutant complex, we eventually achieved success through heterogeneous microseeding with wild-type complex crystal seeds, followed by macroseeding into fresh protein drops. The cross seeding resulted in epitaxial growth in a crystal lattice indistinguishable from that of the wild-type complex. For the M59R/E60S CI2 complex, crystals eventually formed spontaneously in additive screens, but these were imperfect clusters of thin needles that diffracted poorly. Serial microseeding produced long, thin rectangular plates of space group $P2_1$, with a unique unit cell. The M59R/E60S CI2 complex also crystallized in another crystal form, belonging to space group $P6_5$, under different conditions identified in a sparse matrix screen. These crystals were star-shaped and quite fragile, with more limited diffraction resolution. For the G83Δ CI2 complex, all of the methods described above were attempted repeatedly, along with additional pH and temperature screens, as well as variations in protein purification procedures. We observed the growth of multiple crystal forms; however, these crystals were invariably composed of subtilisin alone. One of the crystal forms exhibited a low-density lattice with a cavity adjacent to the active site large enough to accommodate CI2. Mass spectrometric analysis of washed crystals revealed the presence of many small peptides derived from CI2, but no intact G83Δ CI2 was detected. All attempts to produce crystals of the intact G83Δ CI2 complex failed.

Crystal Structures of CI2 Mutant Complexes Reveal an Interconnected H-Bond Network. X-ray structures, ranging from 1.5 to 1.8 Å resolution, were obtained for complexes of subtilisin BPN′ with all the CI2 mutants described except the G83Δ mutant. A summary of data collection and refinement statistics is presented in Table 2. Global structural comparison indicates that no gross changes in the CI2 backbone structure have been introduced by the mutations. Superpositions were based on the α-carbon positions of all residues of subtilisin minus the His tag, plus the residues of the CI2 binding loop (residues I56−I61). The residues of the CI2 core were not used in superpositioning, because the junction between the CI2 core and binding loop displays a degree of flexibility, with slight hinge angle variations between the mutant complexes. Average root-mean-square (rms) deviations of the atoms used in superpositioning, from the corresponding atoms in the wild-type structure, were between 0.05 and 0.10 Å for all structures except the T58P CI2 complex (0.157 Å) and the M59R/E60S CI2 complex (0.235 Å).

In all structures, the backbone atom positioning of CI2 binding loop residues 58−63, encompassing the targeted scissile bond, is essentially unchanged. In all cases, the majority of the H-bonds are conserved, though subtle shifts in side chain positions result in small changes in H-bond donor to acceptor distances (Table 3). Though small, these differences are significant and likely to indicate changes in the strengths of the relevant interactions, since the estimated standard uncertainties of atomic positions in all structures

Table 2: Summary of Data Collection and Refinement Statistics

	T58A	T58P	E60A	E60S	M59R/E60S	M59R/E60S	R62A	R65A	R67A
PDB entry	1Y1K	1Y33	1Y34	1Y3B	1Y4A	1Y4D	1Y3C	1Y48	1Y3D
space group	$P6_522$	$P6_522$	$P6_522$	$P6_522$	$P2_1$	$P6_1$	$P6_522$	$P6_522$	$P6_522$
cell dimensions									
a (Å)	94.05	94.02	93.82	94.25	57.45	115.38	93.87	93.75	94.27
b (Å)	94.05	94.02	93.82	94.25	41.38	115.38	93.87	93.75	94.27
c (Å)	187.38	186.72	185.33	186.14	63.88	42.91	185.07	185.25	185.57
α (deg)	90	90	90	90	90	90	90	90	90
β (deg)	90	90	90	90	110.41	90	90	90	90
γ (deg)	120	120	120	120	90	120	120	120	120
resolution (Å)	1.56	1.80	1.55	1.80	1.60	2.00	1.69	1.84	1.80
no. of unique reflections	69911	44898	69997	45872	26317	22160	53250	42418	45812
completeness (%)	99.5	98.1	99.3	99.6	99.8	99.6	97.4	99.8	99.7
multiplicity	5.6	20.3	4.4	11.4	3.8	4.9	4.8	4.4	18.6
$I/SD^{a,b}$	13.9 (2.8)	20.0 (3.5)	13.0 (2.5)	15.8 (3.8)	9.4 (4.1)	7.4 (2.8)	12.4 (3.2)	10.1 (2.3)	20.8 (5.3)
$R_{merge}{}^{b,c}$ (%)	7.0 (39)	14.5 (85)	6.6 (37)	12.0 (65)	7.2 (33)	15.7 (44)	7.9 (47)	9.4 (48)	10.8 (59)
$R_{cryst}{}^{d}$ (%)	16.4	15.3	15.5	15.1	15.6	19.7	14.8	15.5	14.7
R_{free} (%)	18.5	17.9	18.0	17.8	20.3	25.2	17.8	19.6	17.3
rmsd for bonds[e] (Å)	0.020	0.022	0.018	0.021	0.019	0.023	0.021	0.023	0.022
rmsd for angles[e] (deg)	1.865	1.815	1.779	1.773	1.649	1.810	1.879	1.878	1.794
protein average *B* (Å^2)	16.7	15.8	14.7	16.0	16.1	26.9	15.9	17.6	18.0
water average *B* (Å^2)	33.6	32.8	32.1	34.7	30.7	30.6	32.8	34.3	35.8

[a] *I* is the intensity; SD is the standard deviation. [b] Outer shell values are given in parentheses. [c] $R_{merge} = \sum|I - \langle I\rangle|/\sum I \times 100\%$. [d] $R_{cryst} = \sum|F_{obs} - F_{calc}|/\sum_{obs} \times 100\%$. [e] The rms deviations from ideal geometry.

Table 3: Hydrogen Bond Distances (donor to acceptor, in angstroms) in Variant CI2 Complexes[a]

	wild type	T58A	T58P	E60A	E60S	M59R/E60S	R62A	R65A	R67A
T58 OG1−E60 OE1	2.64	−	−	(2.58)[c]	(2.94)[e]	(2.80)[e]	2.60	2.57	2.64
E60 N−E60 OE1	2.78	2.85	2.85[b]	(2.99)[c]	(3.14)[f]	(3.20)[f]	2.79	2.75	2.73
R65 NH1−E60 OE2	3.00	2.99	3.20[b]	(3.13)[d]	(3.15)[e]	(3.03)[e]	3.09	−	2.93
R65 NH2−E60 OE1	3.12	3.09	2.85[b]	(2.95)[c]	3.05[g]	2.93[g]	3.07	2.98[h]	3.19
R65 NH2−E60 O	2.98	3.03	3.17	3.00	2.99	2.97	2.89	3.06[h]	2.91
R65 NH2−G83 O	2.89	2.86	2.97	2.87	2.86	2.81	2.86	2.98[h]	2.85
R62 N−G83 O	2.84	2.81	2.82	2.84	2.88	2.83	2.88	2.83	2.88
R65 NE−G83 OXT	2.86	2.83	2.95	2.86	2.92	2.81	2.82	3.34[i]	2.81
R67 NH2−T58 OG1	2.77	−	−	2.81	2.65	2.80	2.67	3.04[l]	−
R67 NH2−T58 O	3.26	3.19	3.30	3.30	3.56	3.69	3.54	3.77	−
R67 NH1−T58 O	3.49	3.92	3.96	3.17	3.01	2.95	2.98	3.03	−
R62 NE−E60 OE2	2.98	3.40[k]	2.76[b]	−	−	−	(2.74)[j]	2.86	2.91[k]

[a] Numbers in parentheses represent distances to water molecules that spatially replace polar atoms of mutated side chains. [b] This distance is to the nearer of two alternative conformations of the E60 side chain in the T58P structure, each with 50% occupancy. [c] In the E60A mutant, distances are to a new bridging water molecule, S381, rather than to E60 OE1. [d] In the E60A mutant, the distance is to new water molecule S407, rather than to E60 OE2. [e] In the E60S and M59R/E60S mutants, distances are to a new bridging water molecule, S97, rather than to E60 OE1 or OE2. [f] In the E60S and M59R/E60S mutants, the distance is between bridging water S97 and S60 OG. [g] In the E60S and M59R/E60S mutants, the distance is between R65 NH2 and S60 OG. [h] In the R65A mutant, distances are from R67 NH2 instead of R65 NH2. [i] In the R65A mutant, the distance is from R67 NE instead of R65 NE. [j] In the R62A mutant, the distance is to new water molecule S410 instead of R62 NE. [k] This distance is to the nearer of two alternative conformations of the R62 side chain in this structure. [l] This distance is from R67 NH1 rather than R67 NH2.

are in the range of 0.04−0.07 Å based on a maximum likelihood estimation (*15*). The CI2 binding loop residues are among the more highly ordered protein regions in the complexes, suggesting that the coordinate error associated with these residues is lower than for average protein atoms. Where H-bonding functional groups have been removed, we observe a number of different structural effects stemming from the mutations. In some structures, the functional group and potential H-bonds are deleted with only very slight conformational changes in surrounding residues and solvent waters, while in other cases, new bridging waters are incorporated into the H-bond network replacing the deleted functional groups. In yet other instances, most notably with the R65A CI2 mutant, more substantial reorganization of adjacent side chains alters the H-bonding network such that the most critical missing contacts are substituted with new H-bonds (Figure 2).

Mutation of Thr-58 Results in the Loss of Loop-Stabilizing H-Bonds. The structure of the T58A CI2 complex suggests that the ∼100-fold increases in both dissociation constant and hydrolysis rate with this mutant are largely attributable to the loss of the H-bonds from Thr-58 to Glu-60 and Arg-67. We observe no significant backbone adjustment, no substantial alteration of the ordered solvent molecules, and only very subtle changes in side chain positioning within the H-bonding network, including closer contact between the side chains of Glu-60 and Arg-67, which draw nearer by 0.4 Å (Figure 2A).

The greater impact on both dissociation constant and hydrolysis rate of the T58P mutation can be attributed to the loss of an additional hydrogen bond between the N atom of Thr-58 of CI2 and the Gly-100 backbone carbonyl of subtilisin. Loss of this intermolecular backbone H-bond results in displacement of a subtilisin loop (residues 97−101) by as much as 1.2 Å, relieving a potential steric clash (Figure 3). We observe subtle adjustments in the H-bond network similar to those seen in the T58A CI2 complex, and in addition, Glu-60 displays partial occupancy of an alternative, non-H-bonding conformation.

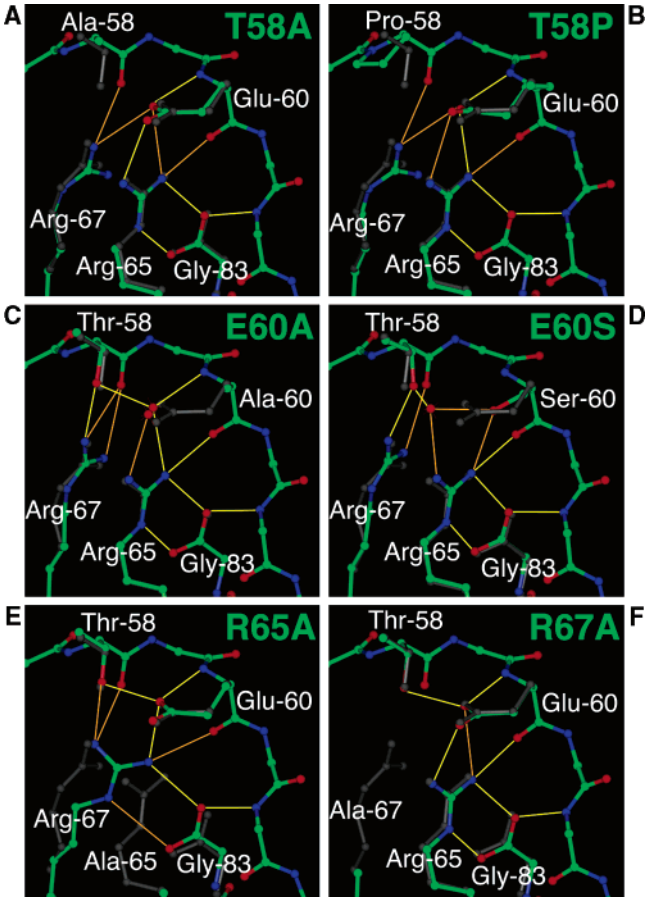


FIGURE 2: Binding loops of six CI2 mutants in complex with subtilisin BPN′. The CI2 backbone and relevant side chain atoms of the mutants are colored green (carbons), red (oxygens and bridging water molecules), and blue (nitrogens); the superposed side chains and C-terminal carboxylate of the wild-type CI2 complex are colored gray for comparison. Potential hydrogen bonds and electrostatic interactions of the H-bond network are colored yellow (interatomic distances of 2.5−3.0 Å) and orange (interatomic distances of 3.0−3.5 Å), respectively. The side chain and potential H-bonds involving Arg-62 have been omitted for clarity; these H-bonds are of minimal functional significance (see the text): (A) the T58A CI2 complex, (B) the T58P CI2 complex (only the more nativelike of two partially occupied conformations of Glu-60 is shown), (C) the E60A CI2 complex, (D) the E60S CI2 complex, (E) the R65A CI2 complex, and (F) the R67A CI2 complex.

Mutation of Glu-60 Results in Replacement of Cross-Loop Side Chain H-Bonds with Solvent-Bridged H-Bonds. In the E60A mutant complex, replacement of the Glu-60 carboxylate with two new bound solvent molecules causes the observed substantial increase in both dissociation constant and hydrolysis rate. The new waters S381 and S407 replace the Glu-60 carboxylate oxygens, forming solvent-bridged H-bonds (Figure 2C). The Arg-67 side chain also extends slightly further than in the native inhibitor, shortening the Arg-67 to Thr-58 contact by 0.3 Å.

Substitution of Glu-60 with Ser rather than Ala has a slightly weaker impact on both inhibition and hydrolysis, perhaps because the Ser-60 hydroxyl can form a direct hydrogen bond with Arg-65. A new water molecule, S97, bridges H-bonding contacts with Thr-58, Arg-65, and Ser-60 (Figure 2D). As in the E60A CI2 complex structure, the Arg-67 to Thr-58 distance is shortened by 0.5 Å.

Mutation of Arg-65 Results in Side Chain Rearrangement and Formation of New H-Bonds, Compensating for Lost Critical Contacts. Glu-60, Arg-65, and Gly-83 are held together by a web of H-bonds and appear be a functional unit. Gly-83 stabilizes the Arg-65 side chain conformation such that it in turn supports the Glu-60 side chain. Since mutation of Gly-83 or Glu-60 accelerates proteolytic hydrolysis greatly, one expects that mutation of the intervening Arg-65 will have a similarly striking effect. The more moderate rate of hydrolysis observed in the R65A CI2 mutant at first appeared to be something of an anomaly. The explanation may be found in the structure of the R65A mutant complex, in which substantial reorganization of the H-bonding network allows the Arg-67 side chain to compensate for the missing Arg-65 guanidinium. In this complex, the Arg-67 side chain shifts more than 2.5 Å to assume a position intermediate between the original positions of Arg-65 and Arg-67, making most of the critical contacts of both residues (Figure 2E). The only H-bond from the native complex that is completely eliminated and not replaced in the R65A mutant is that between Arg-65 NH1 and Glu-60 OE2. The efficient side chain restructuring, maximizing the number and strength of potential H-bonds, explains why the kinetic results with the R65A mutant give an underestimation of the true importance of the Arg-65 residue.

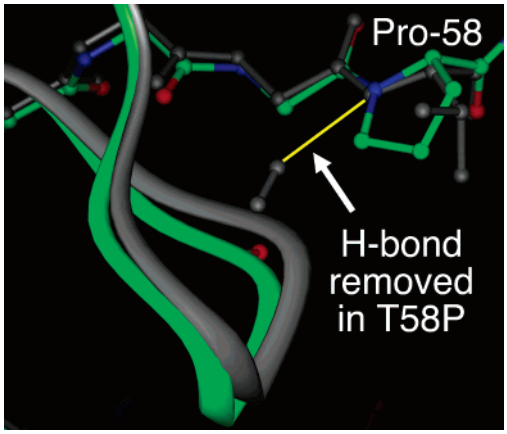


FIGURE 3: Adjustment by the subtilisin backbone upon mutation of Thr-58 of CI2 to Pro. The CI2 backbone and residue 58 side chain are colored green (carbons), red (oxygens), and blue (nitrogens) for the T58P mutant, with the wild-type complex superposed in gray. The subtilisin backbone is shown as a solid ribbon in green (T58P) or gray (wild type). The H-bond formed between the N atom of Thr-58 of CI2 and the Gly-100 backbone carbonyl of subtilisin in the wild-type complex is colored yellow.

Mutation of Arg-67 Results in the Loss of Three Cross-Loop H-Bonds. We observe no substantial structural adjustments in the R67A CI2 complex. Thus, we can attribute the modest increase in hydrolysis rate and dissociation constants with this mutant directly to loss of the Arg-67 side chain and the three H-bonds in which it normally participates (Figure 2F). Removal of the Arg-67 side chain leaves a solvent accessible gap through the center of the binding loop, through which a tube of continuous electron density threads. This density, which we have modeled as a PEG chain, is quite diffuse, suggesting that no specific and discrete H-bonds are formed between the solvent chain and the protein.

The M59R/E60S CI2 Complex Resembles the E60S Complex in the H-Bond Network, with Greater Structural Divergence Further from the Binding Loop. The double mutant, compared to the E60S single mutant, is only slightly

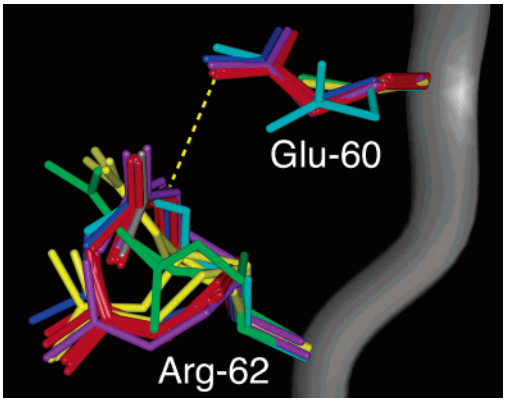


FIGURE 4: Superposition of all CI2−subtilisin BPN′ complex structures, demonstrating the conformational variability of Arg-62. The CI2 backbone is shown as a gray ribbon, and only the side chains of CI2 residues 60 and 62 are shown; in structures where either Glu-60 or Arg-62 is present in two partially occupied conformations, both conformers are shown. Side chains are colored as follows: gray for the wild type, red for all mutants of non-H-bonding residues, including Y61A and five mutants of the Met-59 P_1 residue, yellow for E60A and E60S, green for M59R/E60S (both crystal forms), cyan for T58P, blue for T58A, and purple for R65A and R67A. The potential H-bond between Arg-62 and Glu-60 identified in the wild-type complex is represented by a dashed yellow line.

more compromised with regard to inhibition and resistance to hydrolysis; this is in line with our previous results showing relatively small functional effects upon mutation of the Met-59 P_1 residue of CI2 (*5*). Despite crystallizing in a different crystal form belonging to space group $P2_1$, the structure of the double mutant reveals that the backbone in the binding loop region of residues 58−63 is indistinguishable from the other complexes. The side chain positioning and hydrogen bond network very closely resemble that found in the E60S CI2 single mutant complex. The new Arg-59 P_1 residue fits into the S_1 binding cleft of subtilisin like the wild-type Met.

The M59R/E60S CI2 backbone further from the region of contact with subtilisin shows much greater divergence from the wild-type structure. A second, somewhat lower-resolution (2.0 Å) structure of this complex, in a crystal form belonging to space group $P6_1$, shows essentially the same structure for the subtilisin-contacting residues and the H-bond network, but more distant regions of CI2 diverge substantially from both the $P2_1$ structure and the wild-type complex structure, leading to the conclusion that the structural differences further from the contact interface, rather than being propagated from the sites of mutation, are probably caused primarily by differences in crystal packing forces between the various crystal forms.

The Arg-62 Side Chain of CI2 Is of Limited Functional Importance. Inhibition and hydrolysis assays indicate that Arg-62 plays only a minimal role in the stabilizing H-bond network. Mutation of the Arg-62 side chain to Ala increases the dissociation constant by less than 1 order of magnitude, and has little effect on the resistance of CI2 to hydrolysis by subtilisin. Many other CI2 mutants, including those at residues uninvolved in the H-bond network, show multiple alternative conformations of the Arg-62 side chain (Figure 4), further indicating that the native conformation featuring an H-bond between Arg-62 and Glu-60 is only marginally favored.

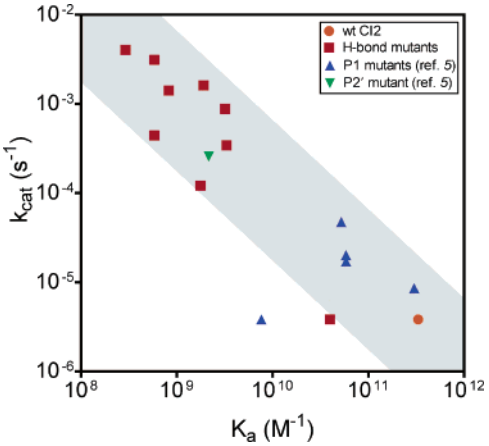


FIGURE 5: Binding−hydrolysis correlation. Hydrolysis rates k_{cat} are plotted vs association constants K_a, calculated as the reciprocal of inhibition (dissociation) constants K_i; both axes are log scaled. The results with the P_1 and P_2' mutants were reported previously (*5*).

DISCUSSION

As a general trend, the rate of CI2 hydrolysis correlates inversely with binding affinity, in agreement with our previous observations (Figure 5) (*5*). The observed correlation between the mutational effects on inhibitor binding and hydrolysis is not surprising, since the same CI2 residues near the enzyme interface are responsible both for constraining the binding loop into a conformation complementary to the enzyme active site and for providing the network of interactions that serve to retain and orient the leaving group peptide following formation of the acyl−enzyme complex, preventing hydrolysis and accelerating peptide religation. Hydrolysis of the acyl−enzyme complex requires partial dissociation of the leaving group peptide to allow access by the hydrolytic water, and inhibitors with stronger binding affinity in the intact form are likely to remain more resistant to dissociation in the cleaved form. Mutations that weaken the binding affinity of the intact inhibitor are therefore likely also to increase the rate of leaving group dissociation and hydrolysis.

While the structural determinants of binding and hydrolysis are inextricably linked, the deviation of individual mutant inhibitors from the correlation is particularly revealing, since it distinguishes residues more important for one function versus the other. The three Arg to Ala mutations, at CI2 residues 62, 65, and 67, all have a more deleterious effect on binding affinity than on hydrolysis rate, indicating that these residues are more critical for providing the correct conformation for initial binding than for retention of the leaving group peptide following formation of the acyl−enzyme complex. Once the acyl−enzyme complex forms, the covalent linkage between the enzyme and inhibitor plays a role in maintaining the conformation of the CI2 binding loop, reducing the importance of the cross-loop H-bonds involving the Arg residues. By contrast, the role of Glu-60 becomes even more critical following formation of the acyl−enzyme complex. The Glu-60 residue is unique in that the mutations at this site have a greater effect on leaving group retention and preventing hydrolysis than they do on inhibitor binding. This is perhaps unsurprising, since Glu-60 is the amino-terminal residue of the leaving group peptide, and

interactions of the glutamic acid side chain are expected to be important in orienting this residue as the attacking nucleophile in the religation reaction.

Among the complexes examined here, the diverse structural changes in response to mutation range from the very subtle to the more substantial. In some structures, new solvent molecules spatially replace truncated side chains, in some cases bridging electrostatic interactions between the remaining protein functional groups, while in other structures, no significant solvent reorganization is observed. For most mutations within the H-bond network, numerous subtle structural adjustments are propagated throughout the entire H-bond network, resulting in some interaction distances becoming shorter and presumably stronger, while others become longer and likely less important for structural stabilization. In several cases, side chains that were previously held in a single conformation become more disordered, with multiple partially occupied conformations. Finally, in some structures, side chain movement is great enough to result in a shuffling of H-bonding partners. For each mutation, the functional effects observed are created not only by simple mutation of a side chain and the loss of interactions in which it participated but also by numerous small local adjustments throughout the H-bond network. The most striking example of this phenomenon is found in the R65A CI2 mutant, where most of the H-bonds in which Arg-65 normally serves as a donor are conserved, with Arg-67 taking the place of Arg-65.

The same network of H-bonds found in CI2 appears to be conserved throughout the closely related I13 family of protease inhibitors, also known as the potato I class of inhibitors (*20*). Most have a Thr residue in the P_2 position, an acidic residue (more often Asp) in the P_1' position, Arg residues at P_3', P_6', and P_8', and a C-terminal glycine residue. Structural and mutational studies with several of these homologues hint at a conserved functional role for the H-bond network as well. Mutation of H-bonding residues in eglin c leads to accelerated hydrolysis (*9*, *21*), and mutation of Arg residues in *Curcurbita maxima* trypsin inhibitor V results in increases in both the rate of hydrolysis and conformational flexibility of the binding loop (*22*, *23*). Related inhibitors that naturally lack one or more of these conserved charged or polar residues are also likely to be more susceptible to hydrolysis; for example, chymotrypsin inhibitor 1, which contains substitutions at Thr-58 and Arg-67, is cleaved by target proteases more rapidly than CI2, although it does still function as an inhibitor (*24*, *25*). The structures of mutant CI2 complexes presented here provide models for how such inhibitors, lacking a full complement of H-bonds to stabilize the binding loop, may nevertheless function through conformational optimization of the existing electrostatic side chains.

CI2 and a few other members of the I13 family of protease inhibitors are unusual in that the stabilizing interactions between the protein scaffold and the binding loop are almost entirely comprised of electrostatic interactions between charged and polar residues, while in most other families of inhibitors, disulfide bonds play a role in constraining the binding loop conformation. However, even among disulfide-stabilized inhibitors, H-bond networks are still important for maintaining the inhibitory conformation of the binding loop. For example, the I12 family of Bowman−Birk inhibitors features a conserved Thr residue in the P2 position, and a conserved Ser residue in the P_1' position (*20*). The Thr hydroxyl group H-bonds to the Ser hydroxyl group as well as to the P_5' backbone amide, and additional backbone H-bonds further stabilize the binding loop (*26*, *27*). Studies with synthetic peptides reproducing the structural motif of the Bowman−Birk binding loop have shown that replacement of the Thr at P_2 makes the loop more susceptible to hydrolysis (*28*). Another example is found in the I1 family ovomucoid inhibitors, which, like CI2, typically feature Thr in the P_2 position H-bonding to Glu in the P_1' position, as well as an intraresidue H-bond between the Glu backbone amide nitrogen and the side chain carboxylate (*29*, *30*). The I3 family Kunitz-type inhibitors found in plants also feature a binding loop shaped and constrained by an extensive H-bond network (*31*, *32*). When an Asn residue central to the network was mutated in a member of this family, the original H-bond network was nearly conserved, through side chain conformational adjustment and inclusion of new solvent atoms (*33*), reminiscent of our observations of the CI2 mutants presented here. These and other similar inhibitors may all function via the clogged gutter mechanism that we have elaborated for CI2. Hydrogen bond networks that link inhibitor binding loops to protein scaffolds may be a common motif by which the leaving group peptide is retained following formation of the acyl−enzyme complex, resulting in inhibitor religation and slowing proteolysis.

A more general insight to be gleaned from this study concerns the importance of structural information when one wishes to interpret biochemical data from mutant proteins in terms of specific atomic interactions. It has become somewhat commonplace to see quantitative free energy values in protein folding, intermolecular binding, or enzyme catalysis attributed to a particular H-bond, electrostatic, or hydrophobic interaction on the basis of mutagenesis experiments, often in the absence of high-resolution structural information. While such studies are suggestive, and can provide order-of-magnitude estimations, it is prudent to recognize that any given mutation may represent a complex mix of subtle structural adjustments, as seen here.

The results presented here are best interpreted in light of the clogged gutter mechanism, which postulates rapid formation of an acyl−enzyme intermediate, followed by favorable religation of the peptide bond, regenerating the intact inhibitor (*3*). We previously showed that CI2 binding loop mutations at residues that directly interact with subtilisin have a modest impact on hydrolysis rates, with residues on the leaving group side of the scissile bond showing a somewhat larger effect than those on the acylating side, consistent with the idea that R′-side interactions play a role in retaining and positioning the leaving group peptide to favor religation (*5*). We demonstrated that the CI2 protein core plays a much more critical role in slowing hydrolysis and favoring religation, as evidenced by the fact that linear or cyclized peptides that mimic the isolated CI2 binding loop sequence are rapidly hydrolyzed by subtilisin (*6*). Here we have shown that the network of hydrogen bonds and electrostatic interactions connecting the CI2 binding loop to the protein core provides structural integrity and conformational stability relevant both for binding affinity and for control of inhibitor religation. The most critical interactions are those involving the Gly-83 C-terminus, the Arg-65 side

chain, the Glu-60 side chain, and the Thr-58 side chain. Mutation of these key residues accelerates hydrolysis by 2−3 orders of magnitude. The H-bond between Thr-58 and Glu-60 bridges the cleavage site, while the interactions between Gly-83, Arg-65, and Glu-60 tie the leaving group R′ peptide tightly to the protein core, assisting in leaving group retention and accelerating the religation reaction. These intraloop electrostatic interactions, combined with enzyme−inhibitor interface complementarity and the covalent linkage between the enzyme and inhibitor in the acyl−enzyme intermediate, work together to arrest and reverse proteolysis.

ACKNOWLEDGMENT

We thank James Holton and George Meigs for technical assistance with X-ray data collection.

REFERENCES

1. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The protein data bank, *Nucleic Acids Res. 28*, 235−242.
2. Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain, *Biochem. Biophys. Res. Commun. 27*, 157−162.
3. Radisky, E. S., and Koshland, D. E., Jr. (2002) A clogged gutter mechanism for protease inhibitors, *Proc. Natl. Acad. Sci. U.S.A. 99*, 10316−10321.
4. Longstaff, C., Campbell, A. F., and Fersht, A. R. (1990) Recombinant chymotrypsin inhibitor 2: Expression, kinetic analysis of inhibition with α-chymotrypsin and wild-type and mutant subtilisin BPN′, and protein engineering to investigate inhibitory specificity and mechanism, *Biochemistry 29*, 7339−7347.
5. Radisky, E. S., Kwan, G., Lu, C. K., and Koshland, D. E., Jr. (2004) Binding, proteolytic, and crystallographic analyses of mutations at the protease/inhibitor interface of the subtilisin BPN′/chymotrypsin inhibitor 2 complex, *Biochemistry 43*, 13648−13656.
6. Radisky, E. S., King, D. S., Kwan, G., and Koshland, D. E., Jr. (2003) The role of the protein core in the inhibitory power of the classic serine protease inhibitor, chymotrypsin inhibitor 2, *Biochemistry 42*, 6484−6492.
7. McPhalen, C. A., and James, M. N. G. (1988) Structural comparison of two serine proteinase-protein inhibitor complexes: Eglin-c-subtilisin Carlsberg and CI-2-subtilisin novo, *Biochemistry 27*, 6582−6598.
8. Jackson, S. E., and Fersht, A. R. (1994) Contribution of residues in the reactive site loop of chymotrypsin inhibitor 2 to protein stability and activity, *Biochemistry 33*, 13880−13887.
9. Heinz, D. W., Hyberts, S. G., Peng, J. W., Priestle, J. P., Wagner, G., and Grütter, M. G. (1992) Changing the inhibitory specificity and function of the proteinase inhibitor eglin c by site-directed mutagenesis: Functional and structural investigation, *Biochemistry 31*, 8755−8766.
10. Holton, J., and Alber, T. (2004) Automated protein crystal structure determination using ELVES, *Proc. Natl. Acad. Sci. U.S.A. 101*, 1537−1542.
11. Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data, *Joint CCP4+ESF-EAMCB Newsletter on Protein Crystallography 26*.
12. Collaborative Computational Project Number 4 (1994) The CCP4 suite: Programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
13. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search, *Acta Crystallogr. D55*, 484−491.
14. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
15. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method, *Acta Crystallogr. D53*, 240−255.
16. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol. 6*, 458−463.
17. Hooft, R. W. W., Vriend, G., Sander, C., and Abola, E. E. (1996) Errors in protein structures, *Nature 381*, 272.
18. Fauman, E. B., Rutenber, E. E., Maley, G. F., Maley, F., and Stroud, R. M. (1994) Water-mediated substrate/product discrimination: The product complex of thymidylate synthase at 1.83 Å, *Biochemistry 33*, 1502−1511.
19. Otzen, D. E., and Fersht, A. R. (1999) Analysis of protein−protein interactions by mutagenesis: Direct versus indirect effects, *Protein Eng. 12*, 41−45.
20. Rawlings, N. D., Tolle, D. P., and Barrett, A. J. (2004) MEROPS: The peptidase database, *Nucleic Acids Res. 32* (Database Issue), D160−D164.
21. Lu, W.-Y., Starovasnik, M. A., Dwyer, J. J., Kossiakoff, A. A., Kent, S. B. H., and Lu, W. (2000) Deciphering the role of the electrostatic interactions involving Gly70 in eglin *c* by total chemical protein synthesis, *Biochemistry 39*, 3575−3584.
22. Cai, M., Huang, Y., Prakash, O., Wen, L., Dunkelbarger, S. P., Huang, J.-K., Liu, J., and Krishnamoorthi, R. (1996) Differential modulation of binding loop flexibility and stability by Arg^{50} and Arg^{52} in *Curcurbita maxima* trypsin inhibitor-V deduced by trypsin-catalyzed hydrolysis and NMR spectroscopy, *Biochemistry 35*, 4784−4794.
23. Cai, M., Gong, Y.-X., Wen, L., and Krishnamoorthi, R. (2002) Correlation of binding-loop internal dynamics with stability and function in the potato I inhibitor family: Relative contributions of Arg50 and Arg52 in *Curcurbita maxima* trypsin inhibitor-V as studied by site-directed mutagenesis and NMR spectroscopy, *Biochemistry 41*, 9572−9579.
24. Boisen, S., Andersen, C. Y., and Hejgaard, J. (1981) Inhibitors of chymotrypsin and microbial serine proteases in barley grains, *Physiol. Plant. 52*, 167−176.
25. Jonassen, I., and Svendsen, I. (1982) Identification of the reactive sites in two homologous serine proteinase inhibitors isolated from barley, *Carlsberg Res. Commun. 47*, 199−203.
26. Li, Y., Huang, Q., Zhang, S., Liu, S., Chi, C., and Tang, Y. (1994) Studies on an artificial trypsin inhibitor peptide derived from the mung bean trypsin inhibitor: Chemical synthesis, refolding, and crystallographic analysis of its complex with trypsin, *J. Biochem. 116*, 18−25.
27. Luckett, S., Garcia, R. S., Barker, J. J., Konarev, Al. V., Shewry, P. R., Clarke, A. R., and Brady, R. L. (1999) High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds, *J. Mol. Biol. 290*, 525−533.
28. McBride, J. D., Brauer, A. B. E., Nievo, M., and Leatherbarrow, R. J. (1998) The role of threonine in the P_2 position of Bowman-Birk proteinase inhibitors: Studies on P_2 variation in cyclic peptides encompassing the reactive site loop, *J. Mol. Biol. 282*, 447−457.
29. Fujinaga, M., Read, R. J., Sielecki, A., Ardelt, W., Laskowski, M., Jr., and James, M. N. G. (1982) Refined crystal structure of the molecular complex of *Streptomyces griseus* protease B, a serine protease, with the third domain of the ovomucoid inhibitor from turkey, *Proc. Natl. Acad. Sci. U.S.A. 79*, 4868−4872.
30. Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I., and Laskowski, M., Jr. (1982) Crystallographic refinement of Japanese quail ovomucoid, a Kazal-type inhibitor, and model building studies of complexes with serine proteases, *J. Mol. Biol. 158*, 515−537.
31. Onesti, S., Brick, P., and Blow, D. M. (1991) Crystal structure of a Kunitz-type trypsin inhibitor from *Erythrina caffra* seeds, *J. Mol. Biol. 217*, 153−176.
32. Song, H. K., and Suh, S. W. (1998) Kunitz-type soybean trypsin inhibitor revisited: Refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from *Erythrina caffra* and tissue-type plasminogen activator, *J. Mol. Biol. 275*, 347−363.
33. Ravichandran, S., Dasgupta, J., Chakrabarti, C., Ghosh, S., Singh, M., and Dattagupta, J. K. (2001) The role of Asn14 in the stability and conformation of the reactive-site loop of winged bean chymotrypsin inhibitor: Crystal structures of two point mutants Asn14 → Lys and Asn14 → Asp, *Protein Eng. 14*, 349−357.

BI047301W