Protein Engineering vol.14 no.5 pp.349–357, 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

S.Ravichandran[1], J.Dasgupta[1], C.Chakrabarti[1], S.Ghosh[2], M.Singh[2] and J.K.Dattagupta[1,3]

[1]Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics, 1/AF, Bidhan Nagar, Calcutta 700 064 and [2]Indian Institute of Chemical Biology, 4 Raja S.C. Mullick Road, Calcutta 700 032, India
[3]To whom correspondence should be addressed.
E-mail: jiban@cmb2.saha.ernet.in

A double-headed chymotrypsin inhibitor, WCI, from winged bean seeds was cloned for structural and biochemical studies. The inhibitor was subjected to two point mutations at a conserved position, Asn14. This residue, known to have a pivotal role in stabilizing the first reactive-site loop (Gln63–Phe68) of the inhibitor, is highly conserved in the sequences of the other members of Kunitz (STI) family as well as in the sequences of Kazal family of serine protease inhibitors. The mutants, N14K and N14D, were subjected to biochemical assay and their characteristics were compared with those of the recombinant inhibitor (rWCI). Crystallographic studies of the recombinant and the mutant proteins are discussed. These studies were primarily aimed at understanding the importance of the protein scaffolding towards the conformational rigidity of the reactive-site loop. Our analysis reveals that, as the Lys14 side chain takes an unusual fold in N14K and the Asp14 side chain in N14D interacts with the loop residues by water-mediated hydrogen bonds, the canonical conformation of the loop has remained effectively intact in both the mutant structures. However, minor alterations such as a 2-fold increase in the inhibitory affinity towards the cognate enzyme were observed.

Keywords: reactive-site loop/site-directed mutagenesis/winged bean chymotrypsin inhibitor/X-ray structure

Introduction

The double-headed winged bean chymotrypsin inhibitor WCI inhibits α-chymotrypsin in a 1:2 molar ratio (Kortt, 1980). It belongs to the Kunitz (STI) family of serine protease inhibitors and is composed of 183 amino acid residues (M_r = 20.2 kDa). The three-dimensional X-ray structure of WCI has been determined and refined at different resolutions, the highest being at 2.13 Å (Ravichandran *et al.*, 1999). WCI has a characteristic β-trefoil fold and its two reactive sites, Gln63–Phe68 and Asn38–Leu43 (designated as the 'first' and 'second' site, respectively), are situated on surface exposed loops, being 34.51 Å apart from each other. The inhibitor displays structural homology with other members of Kunitz (STI) family, e.g. *Erythrina caffra* trypsin inhibitor (ETI) (Onesti *et al.*, 1991), soybean trypsin inhibitor (STI) (Meester *et al.*, 1998; Song and Suh, 1998) and winged bean albumin (WBA-1) (McCoy and Kortt, 1997), as well as with some functionally unrelated proteins, e.g. the C-terminal domain of tetanus neurotoxin Hc fragment (Umland *et al.*, 1997) and interleukin-1α (Graves *et al.*, 1990).

In our previous structural studies on WCI, it was observed that the side chain of one N-terminal residue, Asn14, intrudes inside the first reactive-site loop (Figure 1A) and forms a network of hydrogen bonds involving the main-chain and the side-chain atoms of loop residues (Dattagupta *et al.*, 1996, 1999).

This asparagine residue plays a crucial role in stabilizing the reactive-site loop conformation of WCI and its position is conserved in the sequences of inhibitors that belong to the STI (Kunitz) family (Figure 1B). Like Asn14 in WCI, the side chain of the conserved asparagine in ETI (Asn12) and STI (Asn13) also plays a similar role. Even in the case of WBA-1, a seed storage protein belonging to the same family, it was reported that the putative reactive-site loop is supported through hydrogen-bonding interactions made between Asn15 and the main-chain and side-chain atoms of Ser61 at the P4 position (nomenclature according to Schecter and Berger, 1967). Interestingly, a conserved Asn residue is also observed in the Kazal family (Laskowski *et al.*, 1980), where it was shown that the residue's side chain plays a strikingly similar role in stabilizing the reactive-site loop conformation.

In this study, we wanted to see the effect of the conserved Asn residue on the stability and functionality of the first reactive-site loop of WCI. First, the cDNA of WCI was cloned (Ghosh and Singh, 1997) and Asn14 was mutated to Lys (N14K) and Asp (N14D) individually through site-directed mutagenesis. These mutations (Asn14→Lys and Asn14→Asp) were done with a view to observing the spatial and electrostatic effect of side chains at this conserved position. Although these are point mutations, by virtue of the crucial position of Asn14 and its hydrogen-bonding interactions with neighboring loop residues, they are nonetheless useful for giving an idea about the contribution of the scaffold in the stability of the reactive-site loop. The crystal structure of recombinant WCI and its two mutants N14K and N14D were determined and their structural characteristics compared. The process of mutation, expression, purification and the results of biochemical assay are also reported.

Materials and methods

Cloning, mutation and expression

The double-stranded cDNA pool synthesized by reverse transcriptase and DNA polymerase I from the poly(A)$^+$ RNA of developing winged bean seeds has been used for PCR amplification of the WCI gene (Ghosh and Singh, 1997). The recombinant pBS-WCI was subjected to PCR-directed mutagenesis using mutagenic oligonucleotide primers containing single base mismatches (Higuchi, 1990; Zhao *et al.*, 1993). The mutagenic PCR involved the generation of two PCR products that overlap in sequence containing the same mutation introduced as part of the PCR primers. A subsequent re-amplification of these fragments with cloning primers resulted in the enrichment of the full-length secondary product. A single (T/A) transition in the AAT codon generated by this

method will code for Lys instead of Asn14 when expressed upon induction by IPTG. Similarly, substitution of the first A for G in the same codon will have Asp in the same position in the expressed protein. The plasmids harboring the mutated genes were named pBS-N14K-WCI and pBS-N14D-WCI, respectively.

For the generation of DNA ends encoding the N-terminal part of the protein from the site of mutation, the M13 reverse primer and the primer with sequences corresponding to the opposite strand of the DNA containing the desired point mutation (N14K BKW and N14D BKW) were used. Again, for the amplification of C-terminal part of the protein, the mutagenic primer with the single base substitution (N14K FWD and N14D FWD) and the M13(–40) primers were used. Taking these two PCR products as template, another set of PCR was performed using WCI FWD and WCI REV primers to generate *Eco*RI and *Sal*I sites and the PCR product thus obtained was cloned into pBS-SK(–). All these products were sequenced using an ABI 377 automated DNA sequencer.

Sequences of primers used for mutation of Asn into Lys and Asp

The sequences were as follows:

N14K FWD 5′-GGTAACTTAGTTGAAAA**A**GGTGGCACA-3′
N14K BKW 5′-TGTGCCACC**T**TTTTCAACTAAGTTACC-3′
N14D FWD 5′-GGTAACTTAGTTGAA**G**ATGGTGGCACA-3′
N14D BKW 5′-TGTGCCACCAT**C**TTCAACTAAGTTACC-3′
WCI FWD 5′-GG-GAATTC-GATGATGATTTGGTCGATGC-3′
WCI REV 5′-GAGAGA-CTCGAG-GGATGAGAAGTGCTTA-ATGGC-3′

The sequence within the dashed region in the WCI FWD primer denotes the *Eco*RI recognition sequence and that in the WCI REV primer shows the *Sal*I recognition sequence. In the top four primers, the single bases, denoted by bold letters, represent the point mutations over the WCI gene sequence.

For bacterial overexpression of N14K and N14D, these sequences were subcloned into pTrc99A (Amersham Pharmacia Biotech) and induced with 0.1 mM IPTG for 1 h at 37°C. A 2.5 g amount of optimally induced *E.coli* XL1-Blue cells, harboring the chimeric plasmids, was sonicated and the soluble fraction was used for the purification of mutant proteins by single-step immunoaffinity chromatography, with an overall yield of ~98% (see Table I). The same procedure was used for the recombinant protein (rWCI) and the overall yield was similar (Ghosh and Singh, 1997). In all three cases, the specific protein content was nearly one and usually the preparations were essentially homogeneous as attested by single protein bands in SDS-PAGE. However, unlike rWCI and N14K (Figure 2, lanes b and d), in the case of N14D mutant, there were some faster moving minor bands in addition to major band (Figure 2, lane c). These were identified by immunoblot analysis as fragments of the purified protein. The proteins were quantitated using the Bradford method (Bradford, 1976).

Crystallization, diffraction and data collection

Crystals of rWCI and the two mutants were grown according to the procedure mentioned in the crystallization of native WCI (Dattagupta *et al.*, 1999). However, the reservoir solution contained 30 and 35% (w/v) ammonium sulfate (0.1 M acetate, pH 5.4) for rWCI and mutants, respectively. X-ray diffraction data were collected at the ID14-EH4 undulator MAD beamline at ESRF (Grenoble, France). First, the crystals were exposed briefly (100 ms) in order to derive the crystal quality and its

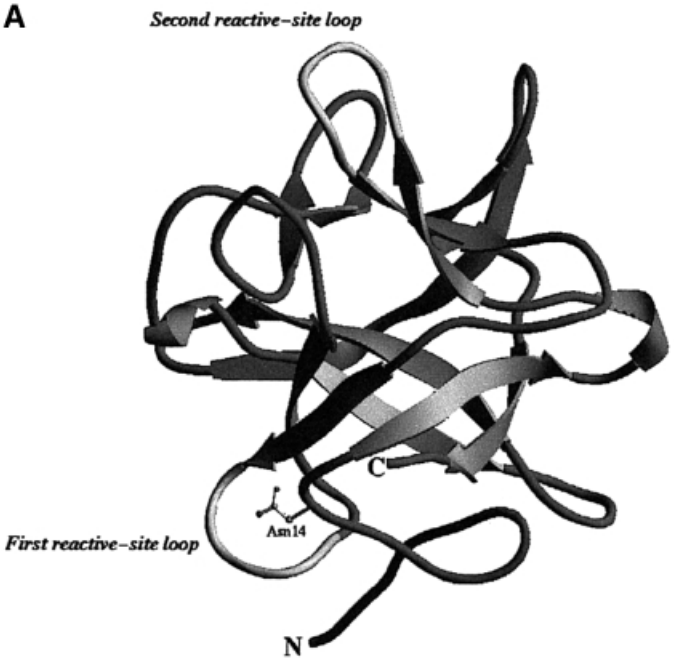

Fig. 1. (**A**). A ribbon representation of WCI. The 12 anti-parallel β strands are joined together by irregular loops and two short 3_{10} helices. The position of Asn14 side chain is also shown as a ball-and-stick model. (**B**) Sequence comparison between WCI and a few other STI (Kunitz) family inhibitors showing the location of the conserved Asn residue (boxed) near the N-terminal region. ECI: *Erythrina vareigata* chymotrypsin inhibitor (Kouzuma *et al*., 1992), WTI-1B and WTI-2: winged bean trypsin inhibitors (Yamamoto *et al*., 1983) The program ClustalW (Thompson *et al*., 1994) was used for sequence alignment.

orientation. Then the oscillation range was maintained at 0.5° (Δt = 1 s/frame and for low resolution Δt = 0.1 s/frame) and the final datasets were collected. They were processed and scaled using MOSFLM (Leslie, 1990) and SCALA (Evans, 1997), respectively. The scaled intensities were converted to structure factor amplitudes using TRUNCATE (French and Wilson, 1978). The crystals belong to the hexagonal space group $P6_122$. Data collection statistics are given in Table II.

Structure refinement

The coordinates of native WCI (PDB entry 4WBC; Ravichandran *et al.*, 1999) served as the starting model for the refinement of rWCI and the two mutants. However, the disordered regions (His137–Asp142, Lys176–His183) and the solvent molecules were excluded from the starting model. Also, in order to avoid model bias in the N-terminal region (since the cloned proteins have the additional tripeptide sequence Met1′–Glu2′–Phe3′), residues Asp1 and Asp2 were excluded from the starting model. For the mutants, Asn14 was replaced by Ala in the starting model. All initial calculations were done using X-PLOR (Brünger, 1992) and CNS (Brünger *et al.*, 1998). For model building, the programs TOM (Jones, 1985) and O (Jones *et al.*, 1991) were used.

Table I. Purification of mutant inhibitors: a summary (starting with 2.5 g of *Escherichia coli*)

Step	Total protein (mg)	N14K (mg)	N14D (mg)	Specific protein content (mg/mg)		Recovery (%)	
				N14K	N14D	N14K	N14D
Bacterial lysate	60	4.0	4.3	0.066	0.07	100	100
Affinity chromatography	4.2	4.0	4.1	0.95	0.97	100	95.3

Table II. Crystallographic parameters and data statistics

	rWCI	N14K	N14D
Unit cell parameters	$a = b = 60.76$; $c = 208.44$ Å; $\alpha = \beta = 90$; $\gamma = 120°$	$a = b = 60.86$; $c = 208.43$ Å; $\alpha = \beta = 90$; $\gamma = 120°$	$a = b = 61.24$; $c = 210.86$ Å; $\alpha = \beta = 90$; $\gamma = 120°$
Resolution range	30–1.9 Å	15–2.05 Å	18–1.90 Å
No. of reflections			
Total	124635	147330	33037
Unique	18974	13051	14503
R_{merge} (%)			
Overall	5.6	8.2	8.7
Last shell	20.1 (2.0–1.9 Å)	28.2 (2.16–2.05 Å)	42.3 (2.0–1.9 Å)
Completeness (%)			
Overall	99.7	88.1	75.7
Last shell	99.5 (2.0–1.9 Å)	91.8 (2.16–2.05 Å)	68.8 (2.0–1.9 Å)
Wilson B (Å^2)	26.9	28.1	31.4
R_{sym} (%)			
Overall	5.7	8.2	8.7
Last shell	20.2 (2.0–1.9Å)	28.2 (2.16–2.05 Å)	42.3 (2.0–1.9 Å)

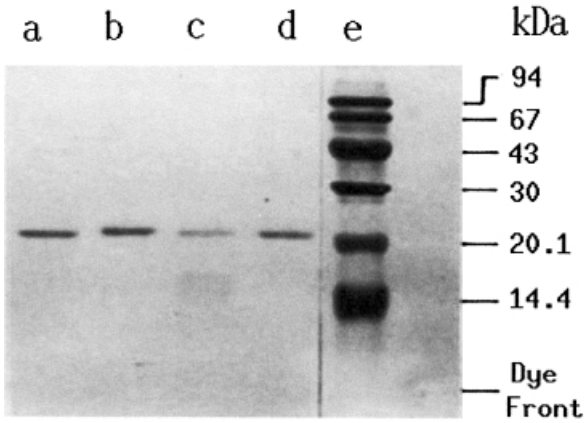

Fig. 2. Homogeneity of the affinity-purified rWCI and mutant inhibitors. The affinity purified proteins (4 μg protein) was subjected to SDS-PAGE (15% acrylamide) and stained with Coomassie Brilliant Blue. Lanes: a, purified WCI from winged bean cotyledons; b, purified rWCI; c, purified N14D: d, purified N14K; e, molecular weight markers (phosphorylase b, bovine serum albumin, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor and α-lactalbumin with their respective subunit molecular weights shown on the right margin).

Refinement of rWCI

In the case of rWCI, 15 cycles of rigid-body refinement followed by a simulated annealing procedure (Brünger *et al.*, 1990) of slow cooling from a starting temperature of 3000 K to 300 K in steps of 25 K yielded a starting R factor of 36.7% (R_{free} = 40.1%). The next 75 cycles of positional refinement involving Powell minimization further reduced the R factor to 33.4% (R_{free} = 36.8%). σ_A weighted maps (Read, 1986) with coefficients $3m|F_o| - 2D|F_c|$ and $m|F_o| - D|F_c|$ were then calculated. Water molecules were gradually incorporated and a flat bulk-solvent correction (Jiang and Brünger, 1994) was applied in all subsequent stages of the refinement. These water molecules were added to the coordinate list only if a peak greater than 3σ appeared in the $F_o - F_c$ difference map and the position of the peak implied at least one hydrogen bond within 3.6 Å, either to a protein atom or to an already defined water molecule. At this stage, a few strong peaks (>4σ height) were observed and were attributed to sulfate ions based on the chemical environment and tetrahedral shape of the electron density. Following several cycles of positional and individual temperature factor refinement, the R factor dropped to 23.7% (R_{free} = 27.1%). At this stage, the refinement was continued using the programs REFMAC (Murshudov *et al.*, 1997) and wARP (Lamzin and Wilson, 1993). More water molecules were added in the model and, finally, all solvent positions were re-confirmed using an OMIT map (Bhat, 1988) involving a modified σ_A coefficient (Vellieux and Dijkstra, 1997). Residues Met1′ and Glu178–His183 were not included in the final model owing to ill-defined electron density around these regions. For the same reason, the N-terminal residue (Glu2′) was kept as Ala in the final model. At the end of final refinement, the R factor converged to 19.5% (R_{free} = 24.6%) for 16 825 reflections in the resolution range 10.0–1.9 Å.

Refinement of N14K and N14D

In case of the mutants, 40 steps of rigid body refinement, followed by 75 steps of positional refinement yielded a starting R factor of 38.1% (R_{free} = 41.9%) and 28.2% (R_{free} = 31.2%) for N14K and N14D, respectively. Lys and Asp were

Table III. Refinement statistics

	rWCI	N14K	N14D[a]
R factor (%)	19.5 (10–1.9 Å)	19.3 (15–2.05 Å)	20.9 (10–2.00 Å)
R_{free} (%)	24.6 (10–1.9 Å)	25.3 (15–2.05 Å)	28.7 (10–2.00 Å)
No. of protein atoms	1405	1415	1392
No. of water molecules	190	176	171
No. of sulfate ions	5	5	4
No. of restraints			
Bond dist.	1470	1484	1446
Angle dist.	2007	2026	1973
Planar 1–4 dist.	495	495	471
R.m.s.d. from ideal values			
Bond lengths (Å)	0.011	0.012	0.027
Bond angles (°)	1.90	2.10	2.10

[a]The structure was refined only to 2.0 Å because the last shell (2.0–1.9 Å) R_{merge} value for the N14D dataset was high (42.3%).

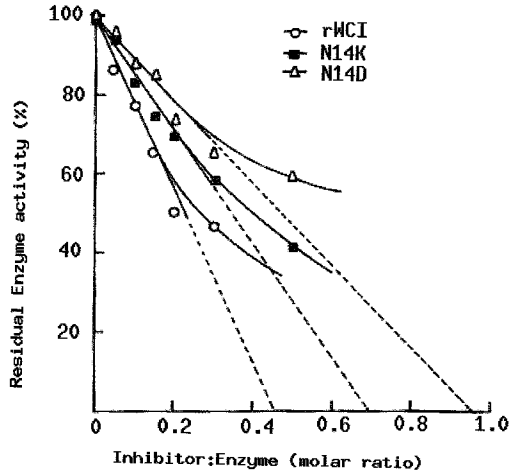

Fig. 3. Inhibition of bovine α-chymotrypsin by rWCI and mutant inhibitors. BTEE was used as the substrate for the assay of chymotrypsin. For chymotrypsin, the protein value was obtained from the extinction coefficient. The differential inhibition patterns for different mutant inhibitors are indicated.

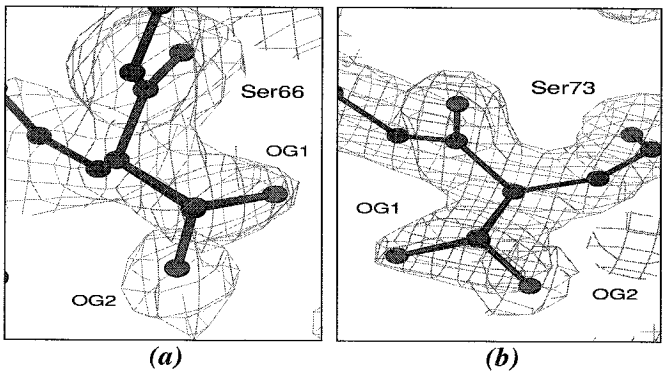

Fig. 4. The two serine residues, (**a**) Ser66 and (**b**) Ser73, display alternate conformations for the side-chain oxygen atoms. The figure indicates the strong positive density (contoured at 2σ) in the $F_o - F_c$ difference map, when only one of the two positions of the OG atom was refined with full occupancy.

incorporated in the corresponding models to satisfy the extra electron density at residue 14, which was Ala in the starting model. All solvent molecules were identified and located using the criteria mentioned before. Next, 100 steps of positional refinement and 120 steps of simulated annealing (slow cool) refinement further reduced the *R* factor (N14K, $R = 30.6\%$, $R_{free} = 39.0\%$; N14D, $R = 26\%$, $R_{free} = 30.0\%$). σ_A maps were then calculated and the models were re-built. After a few cycles of positional and individual *B* factor refinement, the *R* factor further dropped to 22.6% ($R_{free} = 25.5\%$) and 24.1% ($R_{free} = 29.1\%$) for N14K and N14D, respectively. In order to confirm the correct side-chain orientation of the mutated (fourteenth) residue and to resolve any ambiguity or model bias, OMIT maps were calculated for the two models. During this procedure, the mutated residue was omitted and a couple of other regions (Lys176–Ser177 and His137–Asp142), where high atomic *B* factors were noted, were also omitted in the two models. Finally, a few cycles of maximum likelihood (MLKF) refinement were performed using REFMAC. The refinement statistics for the final models of N14K and N14D are shown in Table III. Owing to the poor quality of the electron density map, residues Met1′ and Glu178–His183 could not be located in the final model of N14K as per the wild type. Also for the same reason, residues Met1′–Glu2′–Phe3′ and Glu178–His183 could not be located in the final model of N14D.

Results

Biochemical assay

The purified proteins, recombinant and mutants, were subjected to biochemical assay. rWCI, purified from bacterial lysate, was found to be as active against α-chymotrypsin as the naturally occurring WCI, the respective specific activities being 69 and 74 units/mg protein with *N*-benzoyl-L-tyrosine *p*-nitroanilide used as substrate; it also inhibits the proteinase in the same non-stoichiometric pattern as WCI. The specific inhibitory properties of the mutant inhibitors were then studied with α-chymotrypsin and were found to follow the same non-stoichiometric inhibition pattern as observed in the case of rWCI (Figure 3). The observed pattern of inhibition was clearly suggestive of differential strength and stoichiometry of binding of different inhibitors. Whereas the rWCI molecule was observed to inhibit α-chymotrypsin in a 1:2 ratio as seen in the figure, the N14D molecule showed an almost 1:1 inhibition ratio and N14K molecule showed an intermediate pattern. The specific activity of the mutant inhibitors as well as their dissociation constants with the enzymes were also determined as described by Bieth (1974). The specific activities were 4800 units/mg protein for rWCI, 2900 units/mg protein for N14K

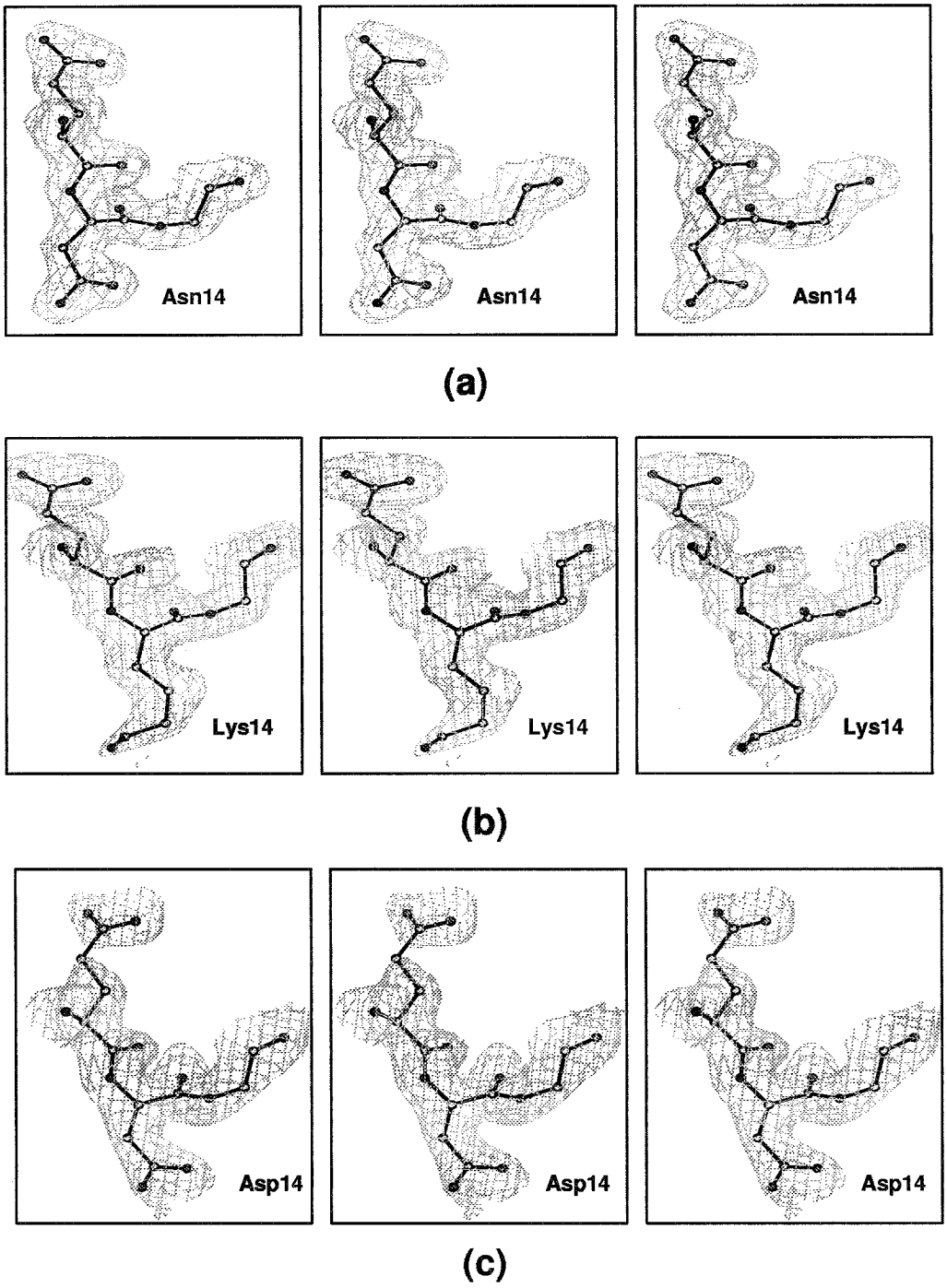

Fig. 5. Stereoscopic representation of the electron-density map shown around the side chain of the fourteenth residue in the three models: (**a**) rWCI ($2F_o - F_c$ map, contoured at 1.2σ), (**b**) N14K and (**c**) N14D ($2F_o - F_c$ OMIT maps, contoured at 1.2σ).

and 1500 units/mg protein for N14D, using BTEE as substrate. Again, for the non-stoichiometric inhibition pattern of these inhibitors, the dissociation constants (K_i) of these inhibitors were calculated using the equation

$$K_i = \frac{[E][I]}{[EI]}$$

where [*E*], [*I*] and [*EI*] are the molar equilibrium concentrations of enzyme, inhibitor and the complex, respectively. Now, if the stoichiometry of association is different from 1:1, then assuming the binding sites to be equivalent and independent, the intrinsic dissociation constant can be determined by the equation

$$\frac{[I_0]}{1-a} = \frac{1}{a} \cdot \frac{K_i}{n} + \frac{[E_0]}{n}$$

where n is the number of binding sites for the enzyme, which was taken as two in this case, $[E_0]$ and $[I_0]$ are the total concentrations of the enzyme and inhibitor, respectively, and a represents the fraction of the enzyme not bound to the inhibitor. The K_i values, obtained from the solution of the equation, for rWCI, N14K and N14D were 1.14×10^{-9}, 2.4×10^{-9}

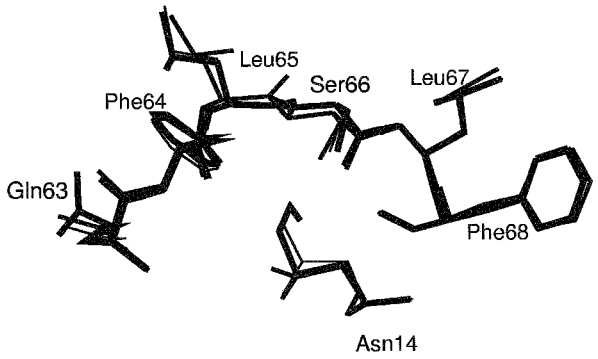

Fig. 6. Superposition of the first reactive-site loop of rWCI with N14K and N14D. The figure also indicates the corresponding side-chain superposition of the fourteenth residue in the three models.

Table IV. (ϕ/ψ) angles (°) for the residues in the first reactive-site loop

Residue	rWCI	N14K	N14D
Ser62	–89.9/131.6	–91.9/135.1	–84.0/127.2
Gln63	–71.3/172.9	–74.6/171.7	–74.8/167.7
Phe64	–97.8/–17.5	–90.3/–13.5	–98.0/–14.3
Leu65	–82.3/168.1	–92.1/172.2	–99.1/170.1
Ser66	–55.4/–9.1	–50.7/–8.9	–78.5/24.9
Leu67	–73.7/158.4	–75.9/160.4	–78.4/160.7
Phe68	–119.6/–29.2	–112.0/–32.0	–122.3/–24.4

and 2.58×10^{-9} M, respectively. The stoichiometry of chymotrypsin–inhibitor complex formation was studied for WCI, rWCI, N14D and N14K by gel permeation chromatography in an FPLC system with a Superose 12 HR 10/30 column. All four inhibitors formed both 1:1 and 1:2 complexes and, in this respect, they were indistinguishable.

Structural comparison of rWCI with native WCI

The three-dimensional structure of rWCI was first compared with the 2.13 Å model of native WCI and, as expected, there was no significant deviation in the overall topology of the two models. The superposition calculations were performed using the programs LSQKAB (Collaborative Computational Project Number 4, 1994) and TOP (Lu, 2000). The main-chain backbone atoms of the 177 amino acid residues superpose with an r.m.s. deviation of 0.5 Å while all the atoms, including the side-chain atoms, superpose with an r.m.s. deviation of 0.86 Å. However, minor conformational changes were observed in a few loop regions and the most notable change was at the β-turn region, Pro70–Ser73, where a 'peptide flip' occurred. This region of rWCI adopts a type I β-turn conformation (ϕ_{71} = –67.9°; ψ_{71} = –17.0°; ϕ_{72} = –89.2°; ψ_{72} = 9.7°), whereas in the native WCI model, the corresponding region had a type II conformation (ϕ_{71} = –54.1°; ψ_{71} = 130.5°; ϕ_{72} = 82.8°; ψ_{72} = 7.6°). Such an interconversion phenomenon, in general, depicts a stereochemically mobile site in homologous crystal structures (Gunasekharan *et al.*, 1998). In addition, conformational changes were also observed in the N-terminal region due to the presence of the extra tripeptide sequence (Met1′–Glu2′–Phe3′) in rWCI. The average r.m.s. deviation for the two N-terminal residues (Asp1 and Asp2) was 2.2 Å and the carbonyl group of Asp1 showed an ~180° peptide flip in this region. Also, double occupancies for the side-chain OG atom of two serine residues, Ser66 and Ser73, were observed (Figure 4). These were not seen in the native WCI model.

Structure of N14K and N14D mutant proteins

An overall structural comparison of the mutant proteins with rWCI reveals that the mutation at this conserved position had almost no effect on the overall topology of the protein. The main-chain CA atoms of 176 amino acid residues of N14K and N14D superpose on rWCI with r.m.s. deviation values of 0.2 and 0.4 Å, respectively. Figure 5 shows the electron-density map around the side chain of the fourteenth residue in the final models of rWCI, N14K and N14D. A superposition of the first reactive-site loop of N14K and N14D on that of rWCI is shown in Figure 6. The r.m.s. deviation values, obtained from CA superpositions of loop residues, are 0.09 and 0.27 Å for N14K and N14D, respectively. In all three structures, the reactive-site loop maintains more or less an identical canonical conformation (Bode and Huber, 1992). The values of dihedral angles (ϕ,ψ) for the loop residues have been compared in Table IV.

For N14K, the long side chain of Lys14, instead of disturbing the canonical conformation of the reactive-site loop, has itself folded back. The average temperature factor value for the side chain atoms (after Cβ) of Lys14 is 26.5 Å^2 whereas the average *B* value for all atoms of Lys14 is 21.98 Å^2. For N14D, the side-chain carboxylate group of Asp14 orients in a direction perpendicular to the corresponding amide group of Asn14 in rWCI (Figure 6). The average value of atomic temperature factor for Asp14 is 33.3 Å^2. Like rWCI, two alternative conformations are observed for the two Ser residues, Ser66 and Ser73. In addition, double occupancy is also observed for the side-chain atoms of Gln63 in N14K and Ser61 in N14D.

Discussion

In general, the loop conformation results from a rather extensive system of hydrogen bonds and hydrophobic interactions, which involve residues both from the loop and the inhibitor scaffolding (Iwanaga *et al.*, 1999; Otlewski *et al.*, 1999). Occasionally, it was also observed that a few water molecules, located in some crucial positions, mediate the interactions between the scaffold and the loop atoms, thereby providing additional stability to the loop. The nature and type of such interactions vary between different classes of proteins and these have not yet been classified in most serine protease inhibitors.

For the Kunitz (STI) family of serine protease inhibitors, the role of a conserved Asn residue coming from the N-terminal segment was emphasized owing to its crucial position (Onesti *et al.*, 1991; Meester *et al.*, 1998). In the present study on rWCI, it is observed that the interactions made by the side chain atoms of Asn14 are more pronounced than its main-chain atoms in stabilizing the conformation of the reactive-site loop. Figure 7A shows the hydrogen-bonding pattern around the first reactive-site loop of rWCI. The side-chain amide group of Asn14 makes three major hydrogen bonds with the reactive-site loop residues: Asn14 ND1 with carbonyl O atoms of Phe64 (P2; 2.91 Å) and Ser66 (P1′; 2.86 Å) and Asn14 OD2 with the main-chain N atom of Ser62 (P4; 2.80 Å); whereas the main-chain N atom of Asn14 contributes only one, i.e. with the carbonyl O atom of Leu67 (P2′; 2.81 Å). This was also observed in the crystal structure of native WCI (Ravichandran *et al.*, 1999). Amongst these, the hydrogen bond between the Asn14 side chain and the P2 residue (Phe64) is conserved in this family. For example, in ETI, Asn12 ND1 makes a hydrogen bond with the carbonyl O of Leu62 and in STI, Asn13 ND1 makes a hydrogen bond with the carbonyl

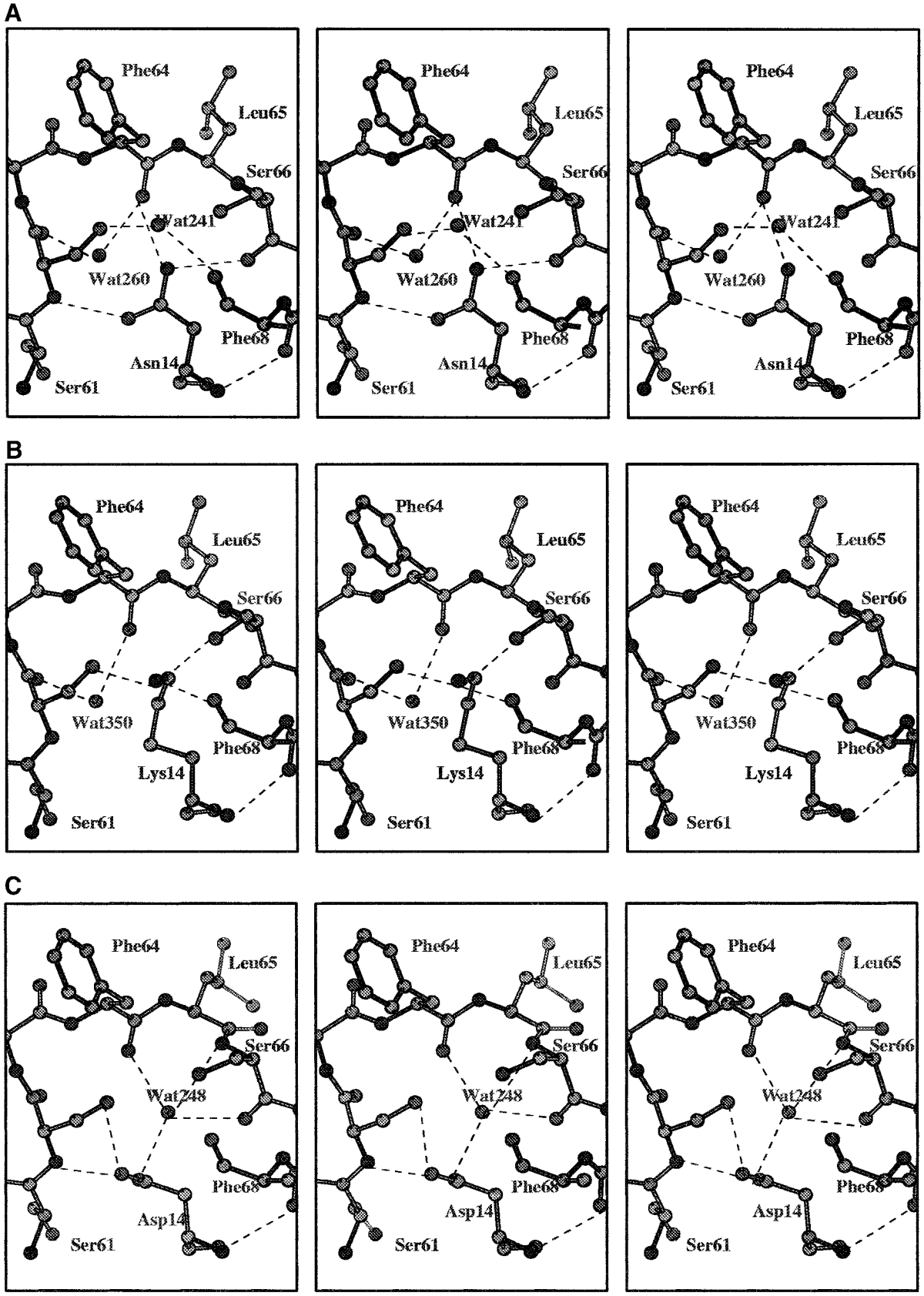

Fig. 7. (**A**) Stereoscopic representation of hydrogen-bonding interactions (shown in dotted lines) made by the side chain of Asn14 with the residues in the first reactive site loop of rWCI. The two water molecules Wat241 and Wat260, which interact with the atoms in the loop residues, are also shown. (**B**) Stereoscopic representation of hydrogen-bonding interactions (shown in dotted lines) made by the side chain of Lys14 with the residues in the first reactive-site loop of N14K. Here, the position of the water molecule Wat350 corresponds to Wat260 of rWCI. The side-chain NZ atom of Lys14 occupies a position corresponding to Wat241 in rWCI. (**C**) Stereoscopic representation of hydrogen-bonding interactions (shown in dotted lines) made by the side-chain of Asp14 with the first reactive site loop of N14D. A water molecule, Wat248, mediates the interaction between the side-chain OD1-atom of Asp14 and the loop residues.

O of Tyr62. A similar interaction is also observed in PSTI, a member of Kazal family, where Asn33 ND1 makes a hydrogen bond with the carbonyl O of P2 residue. Thus, it has been proposed that this conserved asparagine acts as a 'fixed spacer' to keep the loop in a protruded conformation (Papamokos *et al.*, 1982; Onesti *et al.*, 1991).

Also, insufficient protrusion of loop from the protein scaffolding may result in a loss of functionality (Apostoluk and Otlewski, 1998). This was also observed in some non-inhibitory proteins whose loop possesses a canonical conformation, e.g. extracellular hydrolases, toxins, cytokines and viral proteins (Jackson and Russell, 2000). In this regard, there has been considerable interest in understanding the contribution of the protein scaffolding towards the functionality of the inhibitor. Recent studies on the chimeric proteins of *Erythrina vareigata* chymotrypsin/trypsin inhibitor belonging to the STI (Kunitz) family has shown that the N-terminal segment of the protein is involved in the inhibitory activity and is therefore responsible for stabilizing the primary binding loop of ECI (Iwanaga *et al.*, 1999). Interestingly WCI, a member of the same family, has a very high (~70%) sequence homology with ECI.

To check whether Asn14 in WCI is similarly responsible for stabilizing the binding loop conformation, it was mutated to Lys and Asp. Such mutations may help in understanding the characteristics of this crucial residue, contributed by the scaffolding of the inhibitor. A long side chain such as Lys, in place of Asn14, was expected to disturb the stabilizing interactions with the loop. Similarly, a negatively charged residue such as Asp in such a position was also expected to disrupt the hydrogen-bonding pattern.

The first reactive-site loops of N14K, N14D were analyzed and compared with that of rWCI. Figure 7B and C show the hydrogen-bonding pattern around the first reactive-site loop of N14K and N14D, respectively. In rWCI (Figure 7A), apart from the Asn14 side-chain interactions mentioned above, additional stability comes from two water molecules (Wat260 and Wat241) located near the reactive-site loop. Wat260 forms hydrogen bonds with the carbonyl O atoms of P4 (Ser62) and P2 (Phe64) residues while Wat241 is hydrogen bonded to the side-chain OG atom of Ser62 and the carbonyl O atom of the P3′ residue (Phe68). In the mutants the conserved hydrogen bond between the fourteenth residue side chain and P2 residue is absent. However, as in rWCI, the main-chain N atom of the fourteenth residue of both mutants maintains a similar pattern of hydrogen bonds with the P2′ residue (Leu67). Besides this, there is a water molecule, Wat350 in N14K, which corresponds to Wat260 of rWCI and shows a similar pattern of hydrogen bonding, whereas no water molecule corresponding to Wat241 of rWCI is observed. The orientation of the long side chain of Lys14 is responsible for the absence of this water molecule. When superposed with rWCI, the side-chain NZ atom of Lys14 is found to occupy the corresponding position of Wat241 and makes hydrogen bonds (of distance 2.33 and 2.63 Å, respectively) with side-chain OG atoms of two serine residues (P4 and P1′) and the carbonyl O-atom of Phe68 (P3′; 2.61 Å). In N14D, the hydrogen-bonding network at the loop region is mediated through a water molecule, Wat248. The position of this water molecule corresponds to the side-chain ND1 atom of Asn14 in rWCI (see Figure 7A). Wat248 forms hydrogen bonds with the main-chain N and carbonyl O atoms of Ser66 (P1′; 3.18 and 2.96 Å, respectively), the carbonyl O atom of Phe64 (P2; 2.74 Å) and the carboxylate oxygen atoms of Asp14 (2.46 Å) (Figure 7C). Such water-mediated interaction between the reactive-site loop and the protein scaffold was also observed in a few protease inhibitors, e.g. in BPTI, where the interaction is mediated through a water molecule to Gly12 (Otlewski *et al.*, 1999). Apart from this, one of the conserved water molecules (Wat260 of rWCI) is absent in N14D. This is because of one of the alternating positions of the side-chain OG atom of Ser61 (P5) has occupied this region.

Finally, our observations point to the fact that the hydrogen-bonding interactions are almost conserved in rWCI, N14K and N14D. In N14K, although the conserved hydrogen bond involving the side chain of the fourteenth residue and the P2 carbonyl oxygen is not observed, the canonical conformation of the loop remains intact. The long side chain of Lys14, instead of destabilizing the conformation of the reactive-site loop, has itself folded back with an unusual non-rotameric fold. It is seen that the accessible surface area for the Lys14 side chain is 2.44 $Å^2$ and this value is the lowest among the values calculated for the remaining Lys side chains in the model. Moreover, it is reported that Lys side chains can display environment-dependent conformational changes and their flexibility becomes higher during ligand binding to proteins (Najmanovich *et al.*, 2000). Hence it can be understood that the inherent conformational flexibility of the Lys side chain and its limited accessible surface area are responsible for the unusual fold of the Lys14 side chain.

From consideration of the retention volumes of the complexes, it is evident that all three inhibitors form both types of complexes (1:1 and 1:2) with chymotrypsin, depending on the stoichiometry of the incubation mixture. This means that the replacement of Asn14 with Lys or Asp has only a minimal effect on the binding property of the inhibitor molecules. However, minor changes in the hydrogen-bonding pattern could be a plausible reason behind the 2-fold alterations in the dissociation constant values. This result also agrees with the observation of a scissile bond hydrolysis experiment performed for a point mutant, N33S of ovomucoid third domain (a Kazal family inhibitor), where the change in hydrolysis constant K_{hyd} was found to be relatively small, i.e. not exceeding a factor of 3–5 (Ardelt and Laskowski, 1991). The conformational stability of the reactive-site loop and its electrostatic environment therefore depends mainly on its residue type and the local interactions. The residues of the loop (Gln63–Phe68) demand a hydrogen-bonding network for its stability and this need is satisfied in all the structures either by the residues of the scaffold or by solvent molecules keeping the loop conformation intact. As a result, the activity of the protein remains almost unaltered. In other words, the replacement of Asn14 with polar residues marginally influences the activity but its actual role, however, can only be estimated after drastic mutations at this site.

Acknowledgements

We thank Dr S.McSweeney and Dr R.Ravelli for collecting rWCI, N14K and N14D datasets at the ESRF, Grenoble, France. We are grateful to Mr D.Mukhopadhyay and Dr U.Sen for many useful discussions. The coordinates and the structure factor data for rWCI, N14K and N14D have been deposited at the Protein Data Bank (entries 1EYL, 1FMZ and 1FN0). J.K.D. thanks the X-ray Research, Germany, for travel support to ESRF, Grenoble. This work is supported by a grant (BT/PRO0139/RandD/15/11/96) from the Department of Biotechnology, Government of India.

References

Apostoluk,W. and Otlewski,J. (1998) *Proteins: Struct. Funct. Genet.*, **32**, 459–474.

Ardelt,W. and Laskowski,M.,Jr. (1991) *J. Mol. Biol.*, **220**, 1041–1053.
Beith,J. (1974) In Fritz,H., Tschesche,H., Greene,L.J. and Truscheit,E. (eds), *Bayer Symposium V, Proteinase Inhibitors.* Springer, Berlin, pp. 463–469.
Bhat,T.N. (1988) *J. Appl. Crystallogr.*, **21**, 279–281.
Bode,W. and Huber,R. (1992) *Eur. J. Biochem.*, **204**, 433–451.
Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
Brünger,A.T. (1992) *XPLOR Version 3.1: a System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
Brünger,A.T., Krukowski,A. and Erickson,J. (1990) *Acta Crystallogr.*, **A46**, 585–593.
Brünger,A.T. *et al.* (1998) *Acta Crystallogr.*, **D54**, 905–921.
Collaborative Computational Project Number 4 (1994) *Acta Crystallogr.*, **D50**, 760–763.
Dattagupta,J.K., Chakrabarti,C., Sen,U., Dutta,S.K. and Singh,M. (1996) *Acta Crystallogr.*, **D52**, 521–528.
Dattagupta,J.K., Podder,A., Chakrabarti,C., Sen,U., Mukhopadhyay,D., Dutta,S.K. and Singh,M. (1999) *Proteins: Struct. Funct. Genet.*, **35**, 321–331.
Evans,P.R. (1997) *Joint CCP4/ESF–EACBM Newsl.*, **33**, 22–24.
French G.S. and Wilson K.S. (1978) *Acta Crystallogr.*, **A34**, 517.
Ghosh,S. and Singh,M. (1997) *Protein Express. Purif.*, **10,** 100–106.
Graves,B.J., Hatada,M.H., Hendrickson,W.A., Miller,J.K., Madison,V.S. and Satow,Y. (1990) *Biochemistry*, **29**, 2679–2684.
Gunasekharan,K., Gomathi,L., Ramakrishnan,C., Chandrasekhar,J. and Balaram,P. (1998) *J. Mol. Biol.*, **284**, 2505–2516.
Higuchi,R. (1990) In Innis,M.A., Gelfand,D.H., Sninsky,J.J. and White,T.J. (eds), *PCR Protocols.* Academic Press, San Diego, pp. 177–183.
Iwanaga,S., Nagata,R., Miyamoto,A., Kouzuma,Y., Nobuyuki,Y. and Kimura,M. (1999) *J. Biochem.*, **126**, 162–167.
Jackson,R.N. and Russell,R.A. (2000) *J. Mol. Biol.*, **296**, 325–334.
Jiang,J.S. and Brunger,A.T. (1994) *J. Mol. Biol.*, **243**, 100–115.
Jones,T.A. (1985) *Methods Enzymol.*, **115**, 157–171.
Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta Crystallogr.*, **A47**, 110–119.
Kortt,A.A. (1980) *Biochim. Biophys. Acta*, **624**, 237–248.
Kouzuma,Y., Suetake,M., Kimura,M. and Yamasaki,N. (1992) *Biosci. Biotechnol. Biochem.*, **56**, 1819–1824.
Lamzin,V.S. and Wilson,K.S. (1993) *Acta Crystallogr.*, **D49**, 129–147.
Laskowski,M.,Jr, Kato. I., Kohr,W.C., March,C.J. and Bogard,W.C., (1980) *Protides Biol. Fluids*, **28**, 123–128.
Leslie,A.G.W. (1990) *Crystallographic Computing*. Oxford University Press, Oxford.
Lu,G. (2000) *J. Appl. Crystallogr.*, **33**, 176–183.
McCoy,A.J. and Kortt,A.A. (1997) *J. Mol. Biol.*, **269,** 881–891.
Meester,P. De., Brick,P., Lloyd,L.F., Blow,M.D. and Onesti,S. (1998) *Acta Crystallogr.*, **D54,** 589–597.
Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) *Acta Crystallogr.*, **D53,** 240–255.
Najmanovich,R., Kuttner,J., Sobolev,V. and Edelman M. (2000) *Proteins: Struct. Funct. Genet.*, **39**, 261–268.
Onesti,S., Brick,P. and Blow,D.M. (1991) *J. Mol. Biol.*, **217**, 153–176.
Otlewski,J., Krowarsch,D. and Apostoluk,W. (1999) *Acta Biochem. Pol.*, **46**, 531–565.
Papamokos,E., Weber,E., Bode,W. Huber,R., Empie,M.W., Kato,I. and Laskowski,M.,Jr. (1982) *J. Mol. Biol.*, **158**, 515–537.
Ravichandran,S., Sen,U., Chakrabarti,C. and Dattagupta,J.K. (1999) *Acta Crystallogr.*, **D55**, 1814–1821.
Read,R.J. (1986) *Acta Crystallogr.*, **A42**, 140–149.
Schecter,I. and Berger,A. (1967) *Biochem. Biophys. Res. Commun.*, **27**, 157–162.
Song,H.K. and Suh,S.W. (1998) *J. Mol. Biol.*, **275,** 347–363.
Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) *Nucleic Acids Res.*, **22**, 4673–4690.
Umland,T.C., Wingert,L.M., Swaminathan,S., Furey,W.F., Schmidt,J.J. and Sax,M. (1997) *Nature Struct. Biol.*, **4**, 788–792.
Vellieux,F.M.D. and Dijkstra,B.W. (1997) *J. Appl. Crystallogr.*, **30**. 396–399.
Yamamoto,M., Hara,S. and Ikenaka,T. (1983) *J. Biochem. (Tokyo)*, **94**, 849–863.
Zhao,L-J., Zhang,Q.X. and Padmanabhan,R. (1993) *Methods Enzymol.*, **217**, 218–227.

Received September 6, 2000; revised January 22, 2001; accepted February 15, 2001