Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the subtilisin Carlsberg–OMTKY3 complex reveals two different ovomucoid conformations

Jason T. Maynes,[a]* Maia M. Cherney,[a] M. A. Qasim,[b] Michael Laskowski Jr[b] and Michael N. G. James[a]

[a]Canadian Institutes of Health Research, Group in Protein Structure and Function, Department of Biochemistry, Faculty of Medicine, University of Alberta, Edmonton, Alberta T6G 2H7, Canada, and [b]Department of Chemistry, Purdue University, 1393 Brown Building, West Lafayette, IN 47907, USA

Correspondence e-mail: jason@biochem.ualberta.ca

Received 16 December 2004
Accepted 12 February 2005

PDB Reference: subtilisin Carlsberg–OMTKY3 complex, 1yu6, r1yu6sf.

Michael Laskowski Jr has been the driving force behind the detailed physical chemical study of OMTKY3 variants binding to serine proteinases. Unfortunately, he passed away on 2 August 2004. We dedicate this paper, to which he contributed so much, to his memory.

One of the most studied protein proteinase inhibitors is the turkey ovomucoid third domain, OMTKY3. This inhibitor contains a reactive-site loop (Lys13I–Arg21I) that binds in a nearly identical manner to all studied serine proteinases, regardless of their clan or specificity. The crystal structure of OMTKY3 bound to subtilisin Carlsberg (CARL) has been determined. There are two complete copies of the complexes in the crystallographic asymmetric unit. Whereas the two enzyme molecules are virtually identical [0.16 Å root-mean-square difference (r.m.s.d.) for 274 C^{α} atoms], the two inhibitor molecules show dramatic differences between one another (r.m.s.d. = 2.4 Å for 50 C^{α} atoms). When compared with other proteinase-bound OMTKY3 molecules, these inhibitors show even larger differences. This work facilitates a re-evaluation of the importance of certain ovomucoid residues in proteinase binding and explains why additivity and sequence-based binding-prediction methods fail for the CARL–OMTKY3 complex.

1. Introduction

Many standard-mechanism (Laskowski & Kato, 1980) canonical (Bode & Huber, 1992) protein inhibitors of serine proteinases are not highly specific. Instead, they are promiscuous, forming strong inactive 1:1 complexes with many different serine proteinases[1]. Turkey ovomucoid third domain, OMTKY3, is an intensely studied member of the Kazal family (for a recent list of inhibitor families, see Laskowski & Qasim, 2000). It was chosen as the wild type for many comparative studies of natural (Apostol *et al.*, 1993) and recombinant (Lu *et al.*, 2001) variants of avian ovomucoid third domains. Its reactive-site peptide bond (P1–P1′ in Schechter & Berger, 1967, notation) is Leu18I–Glu19I (all residues of the inhibitor are designated by a 'I' in the numbering scheme). An overwhelming amount of structural evidence (Bode & Huber, 1992; Read & James, 1986) and many recent publications show that the reactive site is always intact in all standard-mechanism canonical protein–inhibitor complexes with their cognate enzymes. The strongest assertion of the standard mechanism is that the complexes made from an enzyme and either the intact or the hydrolysed inhibitor are the same substance. Therefore, the reactive-site peptide bond is resynthesized upon complex formation with hydrolysed inhibitor.

Complexes were made from hydrolysed turkey ovomucoid third domains and eight different serine proteinases (Ardelt & Laskowski, 1985). These complexes were subjected to kineti-

[1] Currently, the clear champion of such promiscuity is ecotin. Ecotin accomplishes this, at least in part, by employing two distinct and interacting binding sites (Yang *et al.*, 1998). Here, we are concerned only with inhibitors that employ a single binding site centered on the enzyme's active site in a competitive fashion.

cally controlled dissociation. In all eight cases, the predominant products of the dissociation were the free enzyme and the intact inhibitor. This proves that the Leu18I–Glu19I peptide bond serves as the reactive site in the eight serine proteinases. Since this experiment was completed, the set of serine proteinases for which Leu18I–Glu19I serves as a reactive site has been extended to 15. No exceptions have been found.

Figure 1
Representation of one of the two CARL–OMTKY3 complexes present in the crystallographic asymmetric unit. Coloring is blue at the amino-terminus to red at the carboxy-terminus of subtilisin. The ovomucoid inhibitor is shown in magenta. The active-site catalytic triad and the inhibitor P1 Leu are shown as sticks. A calcium ion bound to CARL is shown as an orange sphere. The amino- and carboxy-termini of subtilisin and OMTKY3 are labeled. All figures were created using the program *PyMOL* (DeLano, 2002).

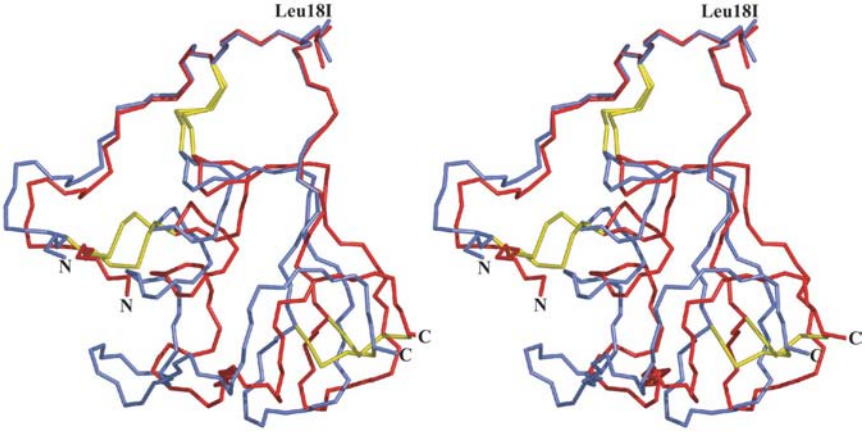

Figure 2
Overlay of the two OMTKY3 inhibitor molecules present in the asymmetric unit (C^{α} trace). The alignment was performed by aligning the two reactive-site loops (P6–P'_3, 0.15 Å C^{α} r.m.s.d. in this region) of the inhibitors, which corresponds more closely to biological function than the best global alignment. This alignment creates a distance between equivalent residues in the distal loop region of $\simeq$ 3.5 Å (*e.g.* 3.3 Å between the C^{α} atoms of the two Asn45I residues). The P1 residue Leu18I is shown as sticks and the three disulfide bridges are shown in yellow.

However, the identity of the reactive-site peptide bonds does not necessarily ensure the identity of all enzyme–inhibitor contact residues. This is most easily tested by three-dimensional structure determination of the complexes of OMTKY3 with the various enzymes. Four such structures are now available: *Streptomyces griseus* proteinase B (SGPB; Fujinaga *et al.*, 1982; Read *et al.*, 1983), human leukocyte elastase (HLE; Bode *et al.*, 1986), bovine chymotrypsin A (CHT; Fujinaga *et al.*, 1987) and a recent structure of OMTKY3 bound to subtilisin Carlsberg (Horn *et al.*, 2003). In the first three cases, the same nine contiguous residues, P6–P3′, surrounding the reactive site are in contact with the cognate enzyme. In every case, several additional discontiguous residues are also in contact with residues of the proteinase. Tabulation of these allowed us to arrive at a consensus discontiguous contact residue set as P14′, P15′ and P18′ (Apostol *et al.*, 1993). This conclusion in turn served for the construction (Lu *et al.*, 2001) of the ten consensus variable residue set [P3 and P15′ were omitted as structural and largely unvaried since Cys16I (P3) is part of a disulfide and Asn33I (P15′) donates a hydrogen bond to both the carbonyl groups of P2 and P1′ in all structures]. The additivity-based sequence-to-reactivity algorithm asserts that in Kazal-family inhibitors only changes in the ten consensus variable residues affect the association equilibrium constant. This is highly hazardous as the set was developed on the basis of only three enzyme–inhibitor structures, but the algorithm makes predictions for these three and for three additional serine proteinases.

We plan to determine the structures of complexes of OMTKY3 with subtilisin Carlsberg (CARL), porcine pancreatic elastase (PPE) and *S. griseus* proteinase A (SGPA). This is the first of these papers. The choice of CARL is obvious. There are several families of serine proteinases, but the most studied of these are the chymotrypsin-related (SA clan; Delbaere *et al.*, 1975) and the subtilisin-related (SB) clan (Barrett *et al.*, 1998). Of the six enzymes with which the sequence-to-reactivity algorithm was developed, five belong to the SA clan and thus may be expected to be similar. CARL is a subtilisin and therefore is a member of the non-homologous SB clan.

The structures of the OMTKY3 inhibitor bound to CHT, SGPB and HLE showed small accommodations by the inhibitor to match the diversity of the tertiary structure of the specific protease. A recent structure of the CARL–OMTKY3 complex showed similar small changes (Horn *et al.*, 2003). While the inhibitor structure changes slightly from enzyme to enzyme, it was

Table 1
Data-collection and refinement statistics.

Values in parentheses correspond to the highest resolution shell.

Data collection	
Unit-cell parameters (Å, °)	a = 110.2, b = 101.1, c = 115.5, $\alpha = \beta = \gamma$ = 90
Space group	$C222_1$
Wavelength (Å)	1.07
Resolution (Å)	40–1.55 (1.61–1.55)
Total No. of reflections	1497065
No. of unique reflections	88530
Completeness (%)	95.4 (77.3)
Redundancy	6.41 (5.13)
$\langle I/\sigma(I)\rangle$	14.7 (3.21)
R_{sym}† (%)	10.2 (42.2)
Mosaicity	0.700
Refinement	
Resolution (Å)	40–1.55
Protein atoms	4607
Water molecules	329
R.m.s. deviations	
Bond lengths (Å)	0.02
Bond angles (°)	1.60
Average B factors (Å^2)	
Protein atoms	13.5
Solvent	19.5
R_{cryst}‡ (%)	19.1
R_{free}§ (%)	20.5

† $R_{sym} = \sum|I - \langle I\rangle|/\sum\langle I\rangle$, where I is the observed intensity and $\langle I\rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections. ‡ $R_{cryst} = \sum||F_o| - |F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor amplitudes, respectively. § R_{free} was calculated as for R_{cryst} using the 5% of the data omitted from structural refinement.

believed that the inhibitor would bind identically to any single enzyme. The plasticity of OMTKY3 when bound to CHT or when bound to SGPB is minimal. Our crystal structure, determined at 1.7 A resolution, has two distinct complexes of CARL–OMTKY3 in the asymmetric unit. Each complex shows a different mode of binding for OMKTY3 to the proteinase, illustrating inhibitor plasticity not only between enzymes but also when binding to a single enzyme.

2. Materials and methods

2.1. Crystallization

Natural subtilisin Carlsberg was purchased from Sigma–Aldrich and subjected to extensive amino-acid analysis to ensure that the correct Carlsberg variant was indeed present. The recombinant third domain of the turkey ovomucoid proteinase inhibitor (residues 6–56) was obtained as described previously (Lu *et al.*, 2001). Crystals were formed by the hanging-drop vapour-diffusion method at room temperature. The enzyme (40 mg ml^{-1}) and inhibitor in water were mixed in a 1:2 molar ratio to a final total protein concentration of $\sim$8 mg ml^{-1} and then mixed with an equal volume of mother liquor (10% ethylene glycol, 480 m*M* sodium malate, 75 m*M* sodium citrate pH 6.1). The complex crystallized in space group $C222_1$, with unit-cell parameters a = 110.2, b = 101.1, c = 115.5 Å and two complexes per asymmetric unit. The Matthews coefficient for this crystal form is 2.1 Å^3 Da^{-1}. This is a different space group and unit cell than the previously determined structure of the CARL–OMTKY3 complex (Horn *et al.*, 2003).

Table 2
Comparison of OMTKY3 structures bound to various proteases (C^α r.m.s.d. in Å).

Values were obtained by aligning the reactive-site loop (P6–P3′) of the OMTKY3 inhibitor molecules and calculating a resultant r.m.s.d. value for the whole domain of the inhibitor. This was performed to compare the differences in how the inhibitors interact with the various proteases, rather than a global best-fit r.m.s.d.

	Cpx 1	Cpx 2	1cho	1ppf	3sgb	1r0r
CARL–OMTKY3 complex 1	—	2.4	3.6	1.7	1.4	2.4
CARL–OMTKY3 complex 2	—	—	1.6	1.6	1.6	0.67
CHT–OMTKY3 (1cho)	—	—	—	1.3	2.4	1.4
HLE–OMTKY3 (1ppf)	—	—	—	—	0.61	1.3
SGPB–OMTKY3 (3sgb)	—	—	—	—	—	1.5
CARL–OMKTY3 (1r0r)	—	—	—	—	—	—

2.2. Data collection, structure determination and refinement

The X-ray diffraction data set was collected at 100 K on beamline 8.3.1 at the Advanced Light Source in Berkeley, California on a ADSC Q210 detector. The data were processed using the programs *MOSFLM* (Powell, 1999) and *SCALA* (Collaborative Computational Project, Number 4, 1994). The structure was solved by molecular replacement with the program *AMoRe* (Navaza & Saludjian, 1997), using separately the structure of subtilisin (PDB code 1vsb) and the structure of OMTKY3 from the CHT–OMTKY3 complex (PDB code 1cho). Electron density for both the protein and the inhibitor were clear in the initial map generated from the molecular-replacement solution. The protein–inhibitor model was manually fitted using the program *XtalView* (McRee, 1999). The model was subjected to iterative rounds of macromolecular refinement using the program *REFMAC* with a maximum-likelihood target (Murshudov *et al.*, 1997). The full data set (1.5 Å resolution) was used for the entire refinement procedure, van der Waals distances were restrained and isotropic B factors were refined. Waters were chosen and refined using the program *ARP*/*wARP* (Perrakis *et al.*, 1999). The crystallographic data are listed in Table 1. The final model consisted of electron density for subtilisin residues 1–275 (274 residues; numbering analogous to subtilisin BPN′ leaves no residue numbered 56) in both complexes, OMTKY3 residues 6–56 in one complex and 7–56 in the other. The model was checked for validity using *PROCHECK* (Laskowski *et al.*, 1993) and *WHATCHECK* (Hooft *et al.* (1996). *PROCHECK* showed that 100% of residues were in allowed Ramachandran plot ranges with an overall G factor of 0.01.

3. Results

3.1. Subtilisin Carlsberg structure

The CARL structure observed here consists of a central seven-stranded parallel β-sheet with helices packed on both sides of the sheet. The active-site residues are located in the region near the C-termini of the strands (Fig. 1). The sheet is

flanked on one side by two α-helices and on the other by a grouping of four α-helices that contain active-site residues His64 and Ser221. Asp32 is on the C-terminus of β-strand 1. The two molecules of CARL that are present in the asymmetric unit of the crystal structure are virtually identical, with a C^{α} r.m.s.d. of 0.18 Å over all 274 residues that are present in the density. The structure observed for the complex between subtilisin and the turkey ovomucoid inhibitor is very similar to other CARL structures that have been determined. The complex of CARL with the inhibitor D-*para*-chlorophenyl-1-acetamido-2-boronic acid (PDB code 1avt) has a C^{α} r.m.s.d. value of 0.29 Å (average of the two subtilisin–OMKTY3 complexes in the asymmetric unit) over all 274 residues and the complex of CARL with eglin c (PDB code 1cse) has a similar r.m.s.d. value of 0.32 Å. The largest difference is in comparison to the apoenzyme (PDB code 1sbc), which has a C^{α} r.m.s.d. value of 0.53 Å (averaged over the two CARL molecules here), where the largest differences are in a surface loop consisting of residues 157–165 that flank the active site.

3.2. OMTKY3 structure

Despite the exceptional structural similarities of the two independent subtilisin molecules in the asymmetric unit, the two ovomucoids show larger and more dramatic differences (Fig. 2). The best global r.m.s.d. for the 50 common C^{α} atoms of the two inhibitors is 0.69 Å, a factor of 3.8 times larger than the equivalent number for the two subtilisin molecules in the asymmetric unit. However, this alignment does not accurately represent the differences that exist between the two inhibitor molecules, because the best overall fit causes misalignment of the reactive-site loops. If functionally aligned (by aligning the two reactive-site loops P6 to P3′ and therefore simulating binding to the enzyme), the C^{α} r.m.s.d. jumps to 2.4 Å. This number is astoundingly large given that both inhibitors bind to the active site of CARL *via* the reactive-site loop virtually identically and both have three intramolecular disulfide bridges, making them quite rigid. The differences in the two inhibitor molecules are independent of the enzyme-binding reactive-site loop (the C^{α} r.m.s.d. of the nine reactive-site loop residues is 0.15 Å), but there are large disparities in the rest of the inhibitor (3.3 Å distance between the C^{α} atoms of the two Asn45I residues, 3.4 Å C^{α} r.m.s.d. over residues 39I–51I). The major differences in the backbone regions of OMTKY3 occur in the loop region between Gly25I and Thr30I and in the coil region that links the only helix present in the inhibitor and the final strand of the sheet. The difference between the two inhibitor molecules is as large as the difference observed between these and the ovomucoids that are bound to other serine proteinases (Table 2). There is also a significant difference in some of the side chains of the reactive-site loop regions. In one complex, Glu19I is positioned to have a moderate electrostatic hydrogen-bonded interaction with Arg21I (2.99 Å), whereas in the other complex the arginine side chain is flipped into solution with subsequent movement of the Glu19I side chain, disrupting this interaction (Fig. 3). The interaction between Glu19I and Arg21I is possible because of the disruption of the hydrogen bond that normally occurs between the side chains of Thr17I and Glu19I. This latter bond is present in all other OMTKY3 complexes with proteases and is thought to maintain the rigidity of the reactive-site loop during inhibition of the cognate enzyme. Its exclusion in CARL–OMTKY3 complexes either suggests that reformation of the scissile bond may not occur as readily with CARL as with other proteinases or that the Thr17I–Glu19I interaction may not be vital for inhibition of CARL.

Table 3
Interactions within 4 Å observed between CARL and OMTKY3.

OMTKY3 residue	OMTKY3 atom	CARL residue	CARL atom	Interaction distance in complex 1†	Interaction distance in complex 2†
Tyr11I	OH	Ser101	O^{γ}	—	3.78
Lys13I	N^{ζ}	Ser101	O^{γ}	3.40	—
Lys13I	N^{ζ}	Asn97	$O^{\delta1}$	(2.98)	3.65
Pro14I	O	Gly128	C^{α}	3.17	3.46
Pro14I	C^{β}	Tyr104	C^{ζ}	3.99	—
Pro14I	C^{β}	Tyr104	$C^{\varepsilon2}$	—	3.88
Ala15I‡	N	Gly102	O	(3.07)	(3.05)
Ala15I‡	O	Gly102	N	(2.89)	(2.90)
Ala15I	C^{β}	Ile107	$C^{\delta1}$	3.21	3.21
Cys16I‡	N	Gly127	O	(3.10)	(3.00)
Cys16I‡	O	Gly127	N	(2.95)	(2.97)
Cys16I	O	Leu126	C^{α}	3.18	3.17
Thr17I‡	N	Gly100	O	(3.10)	(3.11)
Thr17I	C^{β}	His64	$C^{\delta2}$	3.95	—
Thr17I	C^{β}	His64	$N^{\varepsilon1}$	—	3.66
Thr17I	$C^{\gamma2}$	Leu96	$C^{\delta1}$	3.79	3.63
Leu18I‡	O	Asn155	$N^{\delta2}$	(2.71)	(2.77)
Leu18I‡	O	Ser221	N	(3.21)	(3.07)
Leu18I	O	Gly219	C^{α}	3.50	—
Leu18I	O	Thr220	N	3.58	3.47
Leu18I‡	N	Ser221	O^{γ}	(2.96)	(2.89)
Leu18I‡	N	Ser125	O	(3.30)	(3.25)
Leu18I	$C^{\delta1}$	Gly154	C^{α}	3.73	—
Leu18I	$C^{\delta1}$	Ala152	O	3.64	—
Leu18I	$C^{\delta1}$	Ala152	C^{β}	—	3.54
Leu18I	$C^{\delta2}$	Gly154	C^{α}	—	3.70
Glu19I	C^{δ}	Leu217	$C^{\delta2}$	3.70	—
Glu19I	$O^{\varepsilon2}$	Leu217	C^{γ}	—	3.69
Tyr20I‡	N	Asn218	O	(2.83)	(2.87)
Tyr20I	N	Gly219	C^{α}	3.88	3.84
Tyr20I	$C^{\delta2}$	Phe189	$C^{\delta2}$	3.34	3.47
Lys55I	N^{ζ}	Ser188	C^{β}	3.88	—

† Distances in parentheses are potential hydrogen-bonding distances (in Å); other distances are the closest interaction observed. Only the closest interaction is listed for each enzyme residue unless multiple hydrogen bonds were observed for that residue. ‡ These interactions are the usual seen in standard-mechanism canonical inhibitors binding to the subtilisin clan of proteases.

4. Discussion

4.1. Subtilsin–OMTKY3 interactions

Most of the typical interactions for standard-mechanism canonical inhibitors and the subtilisin-clan (SB clan) proteinases are present in both subtilisin–OMTKY3 complexes observed here (Laskowski *et al.*, 2000; Fig. 3 and Table 3). This involves eight main-chain–main-chain hydrogen-bonding interactions and two main-chain–side-chain hydrogen-bonding interactions (see Table 3 for a list of the typical intermolecular interactions). The most interesting interaction is the bifurcated hydrogen bonding between the

main-chain N atom of the P1 Leu18I residue and the backbone carbonyl of Ser125 (3.28 Å average for both complexes present) and the O^{γ} atom of the catalytic Ser221 (2.93 Å). The carbonyl O atom of Leu18I also interacts with two enzyme residues: the backbone N atom of Ser221 (3.14 Å) and the $N^{\gamma 2}$ of Asn155 (2.74 Å). The interactions of canonical standard-mechanism inhibitors with the SB clan of proteinases is expected to have two more hydrogen bonds than with members of the SA clan (chymotrypsin clan; Laskowski *et al.*, 2000). This includes the P2 residue (Thr17I peptide N atom interacts through backbone hydrogen bonds with Gly100, 3.11 Å) and the P4 residue (Ala15I interacts through both potential backbone hydrogen bonds with Gly102, 3.06 and 2.90 Å). The canonical interactions seen in the complex between CARL and OMTKY3 are all seen in the complex between CARL and eglin c and in the previous structure of CARL and OMTKY3 (Fig. 4; Horn *et al.*, 2003; Laskowski *et al.*, 2000).

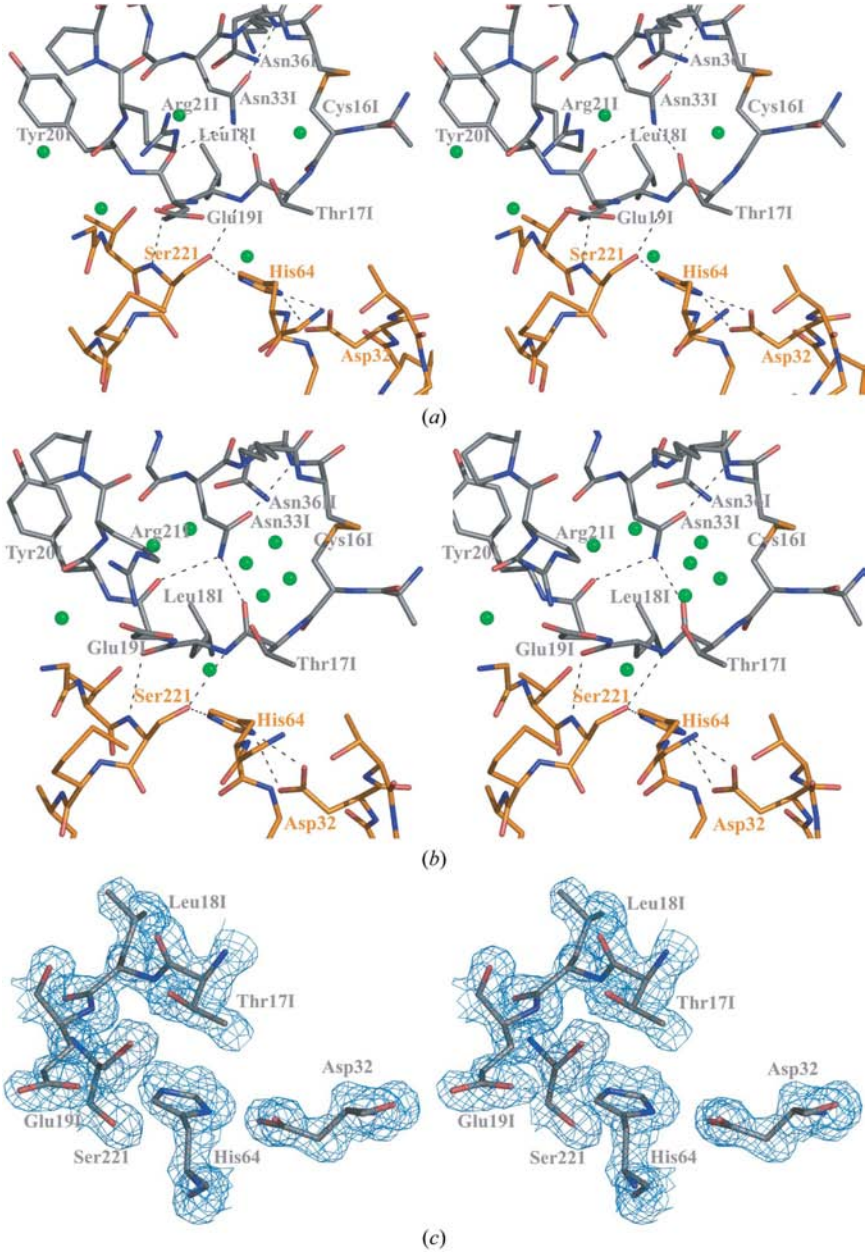

Figure 3
(*a*) Active-site representation of one of the CARL–OMKTY3 complexes in the asymmetric unit. C atoms for the CARL structure are colored orange and for OMTKY3 are colored grey. Labelling for pertinent residues follows a similar coloring scheme. Water molecules are shown as green spheres and hydrogen-bonding interactions are shown as dashed lines. (*b*) Active-site representation of the second complex of CARL–OMTKY3 present in the asymmetric unit. Coloring is as in (*a*). (*c*) Representative electron density shown in stereo for the catalytic triad of subtilisin (Asp32, His64 and Ser221) and the P_2 (Thr17I), P_1 (Leu18I) and P_1' (Glu19I) residues of the ovomucoid inhibitor. The map shown is a σ_A-weighted $2|F_o| - |F_c|$ map (α_{calc}) contoured at 1σ.

Additional interactions between the enzyme and the inhibitor involve water-mediated hydrogen bonds between the acidic group of Glu19I and the enzyme residues Asn62 and Tyr209. These interactions are not seen in any other OMTKY3 complexes owing to the alternative orientation of Glu19I in the present complexes (see below). It was noted that the P2′ residue of OMTKY3 contributes the most out of the P6 to P3′ residues to binding with subtilsin and HLE, unlike chymotrypsin or SGPB (Laskowski *et al.*, 2000). The P2′ residue (Tyr20I) has the potential to participate in a stacking hydrophobic interaction with Phe189 in subtilisin, whereas the equivalent residues in the other proteinases may not allow as strong an interaction [*i.e.* Arg41 in SGPB (some stacking may be possible in this case), Thr151 in CHT, Ile151 in HLE]. This interaction may provide a significant amount of the binding energy for the formation of the subtilisin–OMTKY3 complex ($\simeq 8$ kJ mol^{-1}; McGaughey *et al.*, 1998). This is supported by the kinetic data, which show that CARL is less tolerant of P2′ substitutions from Tyr to residues where the stacking interaction is not possible (for example, the $\Delta\Delta G$ for substitution of P2′ Tyr to Ala ranges from 8.0 to 13.2 kJ mol^{-1} for CHT, SGPB and HLE, but is 22.6 kJ mol^{-1} for CARL; Lu *et al.*, 2001). The P4 residue (Ala15I) provides a significant contribution to binding in both the SGPB and subtilisin complexes, but makes much less of a contribution in the CHT and HLE complexes. In both CARL and SGPB, the S4 site, in which the methyl group of Ala15I sits, consists of a moderately deep hydrophobic pocket. In contrast, the S4 site in CHT and HLE is less hydrophobic and does not involve a deep pocket; it is only a small patch on the surface of the protein. This insinuates that CHT and HLE can accommodate a rather large range of P4 residues both in size and charge, as

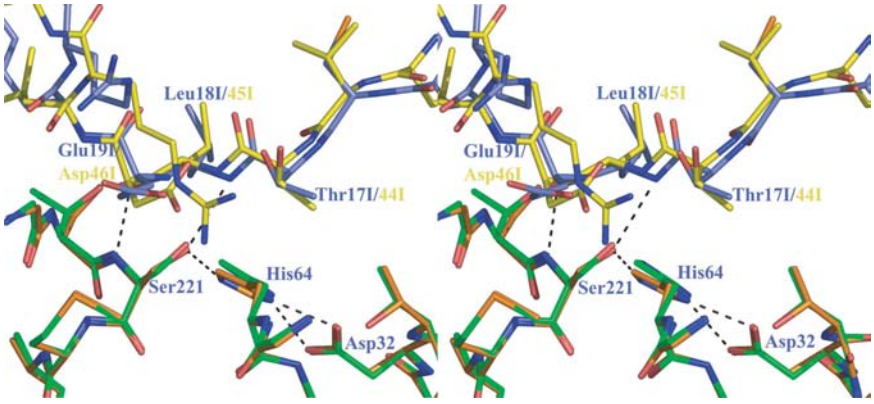

Figure 4
Overlay of the complex between CARL–OMTKY3 and CARL–eglin c (PDB code 1cse). The CARL–OMTKY3 complex is shown with orange (CARL) and blue (OMKTY) C atoms. The CARL–eglin c complex is shown with green (CARL) and yellow (eglin c) C atoms. The sequence of OMKTY3 in this region is P_2-Thr-Leu-Glu-P_1' and for eglin c the sequence is P_2-Thr-Leu-Asp-P_1'. Potential hydrogen bonds are shown as dashed lines.

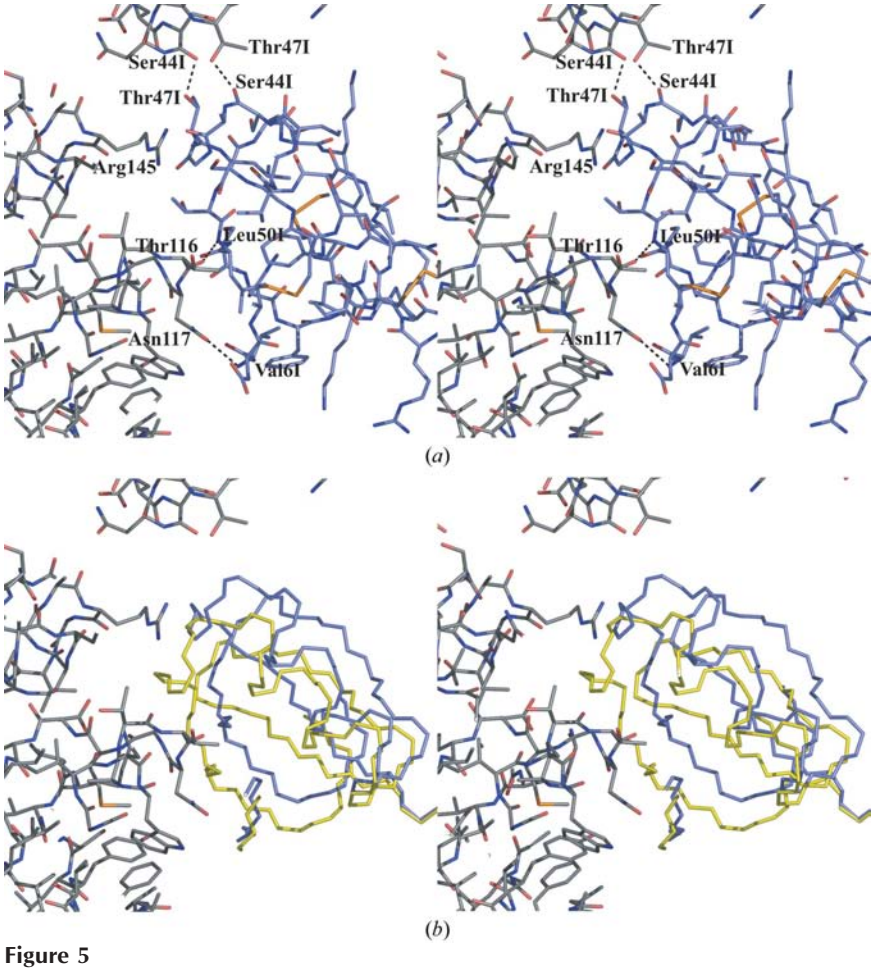

Figure 5
(*a*) Representation of one OMTKY3 molecule (shown with blue C atoms) affected by intermolecular interactions with symmetry-related molecules (shown with grey C atoms). The four pertinent hydrogen bonds between the inhibitor molecule and symmetry-related molecules are shown as dashed lines and the two loops causing steric clashes have representative residues labelled (Thr116 and Arg145). (*b*) Overlay of the two OMTKY3 molecules aligned *via* the reactive-site loops. The OMKTY3 affected by symmetry-related interactions is shown as a blue backbone trace and the unaffected OMTKY3 as a yellow backbone trace. The steric clashes are evident with loops from a symmetry-related subtilisin molecule (show with grey C atoms). The view is identical to that in (*a*).

shown by the energetically optimal P4 residues being Trp for both CHT and HLE; however, the residue having the highest specificity for the enzyme is P4 Gly in the case of CHT and P4 Arg for HLE (Laskowski *et al.*, 2000). With a higher promiscuity for the P4 residue, it is less likely that this site would be important for binding. The reason for the increased P4 specificity in CARL can be attributed to the presence of a deeper more restrictive S4 binding pocket, which leads to a more prominent role for P4 binding in CARL. The S4 pocket of SGPB is also more accommodating of the P4 Ala residue than the S4 pockets of CHT or HLE. This is represented by the larger contribution of the P4 residue in SGPB binding. The pocket is not as deep as the CARL S4 pocket, which is apparent in the lower residue specificity of SGPB than CARL at the P4 site.

4.2. Comparison of the two inhibitor molecules

The two complexes of CARL–OMTKY3 present in the crystallographic asymmetric unit show two vastly different conformations of the ovomucoid inhibitor, despite the nearly identical conformations of the CARL molecules. The difference in conformation is surprising given that there are three disulfide bonds in OMTKY3. These disulfides are presumably used to maintain the inhibitor's conformation in solution and to minimize the entropic cost of binding to cognate enzymes. For OMKTY3 complexes with SGPB, 17 of the 20 naturally occurring amino acids were placed in the P1 position of OMTKY3, along with several chemically synthesized amino acids, and their SGPB–complex structures determined (Bateman *et al.*, 2000, 2001; Huang *et al.*, 1995). Complexes with naturally occurring amino acids had a maximum r.m.s.d. of 0.27 Å for all 944 main-chain atoms (enzyme plus inhibitor) in pairwise comparisons. Even the replacement of the amide bond between P2 (Thr17I) and P1 (Leu18I) by an ester bond, thereby removing the hydrogen-bonding ability of Leu18 NH, only resulted in an r.m.s.d. value of 0.14 Å to the wild-type OMTKY3. As mentioned above, the differences observed between the two ovomucoids present here bound to the same

Table 4
Comparison of backbone torsion angles (°) for subtilisin complexes with OMTKY3 and eglin c (Bode *et al.*, 1987).

Residue	OMKTY3 complex 1 torsion angles (φ/ψ)	OMKTY3 complex 2 torsion angles (φ/ψ)	Eglin c complex torsion angles (φ/ψ)	OMKTY3 torsion angles† (φ/ψ)
P4	−91/127	−90/127	−71/140	−101/140
P3	−127/149	−127/153	−138/168	−132/156
P2	−57/156	−59/160	−62/143	−62/161
P1	−111/37	−109/37	−115/44	−106/36
P1′	−72/139	−75/130	−96/168	−77/134
P2′	−95/108	−88/103	−117/109	−96/106
P3′	−139/78	−133/83	−121/112	−142/74

† Horn *et al.* (2003).

enzyme are as large as the differences seen between ovomucoids bound to different enzymes. One of the two OMTKY3 molecules does have contacts with symmetry-related subtilisin and OMTKY3 molecules, which may facilitate the movement in the inhibitor. The change in this single inhibitor molecule seems to involve two types of interactions: steric interactions that deform the outer loops and hydrogen-bonding interactions that stabilize the new conformation (Fig. 5). The steric interactions include contact between the inhibitor and the two loops involving Thr116 and Arg145 in a symmetry-related subtilisin molecule and His52I in a symmetry-related OMTKY3 molecule (both the subtilisin and the OMKTY3 are part of a complex that is not a symmetry-related complex of the deformed OMKTY3). There are four hydrogen-bonding interactions which occur between this OMKTY3 and symmetry-related molecules that appear to stabilize the complex. Two of these are between the OMKTY3 and the same subtilisin molecule that produces the steric interactions above (Leu50I N to Thr116 O, 2.92 Å and Val6I N to Asn117 $O^{\delta1}$, 2.86 Å). The other two interactions are between the OMKTY3 and a symmetry-related OMKTY3 that is also deformed. These two interactions are interesting since they are related to one another by a twofold symmetry axis (Thr47I $O^{\gamma1}$ in one inhibitor molecule hydrogen bonds to Ser44I O of the other inhibitor and *vice versa*; both bonds are 2.67 Å in length). Since the alterations seen do not involve the backbone of the reactive-site loop, these movements presumably represent two snapshots of dynamic motions of the inhibitor that occur in solution but would not affect how the inhibitor binds to the enzyme. The larger open active-site cleft of CARL may allow a promiscuous binding mode.

4.3. Comparison of proteinase-bound ovomucoids

The differences seen here between the two inhibitor molecules are equivalent to observed differences among ovomucoids bound to other proteinases. The OMTKY3 molecules bound to chymotrypsin, HLE, SGPB and now CARL differ typically by a 1 Å C^{α} r.m.s.d.. These inhibitors all have similar placements of the respective reactive-site loops and similar interactions with the cognate enzyme. It has been observed that residues P4–P3′ in OMTKY3 all adopt similar ψ/φ angles, giving the reactive-site loops similar conformations (Bode & Huber, 1992; Laskowski *et al.*, 2000). This is again true for the subtilisin-bound OMTKY3 molecules, with the exception of the P4 residue. The typical value for φ of the P4 residue in bound OMTKY3 is $\simeq -140^{\circ}$ and the typical Ψ value is $\simeq 150^{\circ}$, whereas the corresponding values in the subtilisin complexes are approximately −90 and 125°, respectively (Table 4). The structural changes that these alterations cause can be seen in the alignment of the ovomucoids as causing the N-terminal coil region of the inhibitor to turn away from the enzyme and is similar to the previously determined CARL–OMTKY3 complex (Horn *et al.*, 2003). This variation may be caused by the backbone interaction between the Ala15I and Gly102 in a β-sheet type fashion that does not occur in the CHT, HLE and SGPB complexes.

All of the proteinase-bound OMTKY3s bind to their cognate enzymes in a similar manner using the reactive-site loop. The differences between the inhibitor molecules arise from apparent hinging on the reactive-site loop. On the N-terminal side of the reactive-site loop the hinge point is quite obvious at the P4 Ala residue, whereas on the C-terminal side the changes occur more gradually, with slight φ/ψ changes becoming apparent at P4′ Pro. This may represent some plasticity in the inhibitor that allows it to conform to a specific peptidase without compromising its active-site binding.

To facilitate binding of the OMTKY3 inhibitor to its target proteinase, it is hypothesized that the reactive-site loop is held firmly in a conformation that is complementary to the enzyme's active site in order to reduce the entropic cost of binding. This is achieved *via* an intramolecular disulfide bond (Cys16–Cys35) and four important internal hydrogen-bonding interactions (Thr17I–Glu19I and Asn33I–Asn36I side-chain hydrogen bonding and Asn33I $N^{\delta2}$ hydrogen bonding to both the backbone carbonyl O atoms of Thr17I and Glu19I). These interactions are present in all previous structures of bound OMTKY3. In the case of subtilisin-bound OMTKY3, the Asn33I–Asn36I hydrogen bond is intact but the crucial Thr17I–Glu19I bond is not formed. The loss of hydrogen bonding arises solely from a change in the side-chain conformation of Glu19I and not from a rearrangement of the reactive-site loop. The electron density for the reactive-site loop is quite clear and the positions of both Glu19I and Thr17I are not in doubt. In both complexes, Glu19I makes interactions with Tyr20I (a water-mediated hydrogen bond to the carbonyl O atom) and with Asn62 and Tyr209 of subtilisin (all water-mediated hydrogen bonds). Additionally, one of these complexes has an electrostatic interaction between Glu19I and Arg21I, the consequences of which are not clear. The interactions between Glu19I and subtilisin are somewhat surprising since the P1′ residues have not been shown to be relatively important in the binding of OMTKY3 to subtilisin (in comparison to those residues with CHT, HLE or SGPB). The loss of the hydrogen bond between Thr17I and Glu19I possibly occurs in the transition of OMKTY3 from in solution to bound to subtilisin. The loss of this interaction cannot be a consequence of the electrostatic interaction between Arg21I

and Glu19I since this new interaction only occurs in one of the two complexes. This suggests that the interaction between Thr17I and Glu19I may not be vital for inhibition of CARL. The hydrogen-bonding between Asn33I $N^{\delta2}$ and the backbone carbonyl O atom of Thr17I is intact and of similar length to other OMTKY3 complexes, as is the hydrogen bonding between Asn33I $N^{\delta2}$ and the backbone carbonyl O atom of Glu19I. When compared with the inhibition of other proteases, the Thr17I O^{γ}–Glu19I $O^{\varepsilon1}$ hydrogen bond does indeed appear to be less important to CARL. Substitution of Thr17 with Val causes an increase in the free energy of binding by 3.26 kJ mol^{-1} for CARL, whereas for CHT, SGPB and HLE the equivalent values are 13.3, 9.16 and 7.28 mol^{-1}, respectively (Lu *et al.*, 2001). The same occurs when analyzing substitutions of Glu19 with Leu, with an increase in binding free energy for CARL of 2.34 mol^{-1} and equivalent values for CHT, SGPB and HLE being 11.8, 3.22 and 4.40 mol^{-1}, respectively. Radisky & Koshland (2002) proposed that rigidity of the reactive-site loop did not contribute significantly to proteinase inhibition by protein proteinase inhibitors. While this appears to be true for OMTKY3 binding to subtilisin since ablation of the Thr17I–Glu19I bond does not affect inhibition, it most likely is not true for other proteinases bound to OMKTY3 since removal of this bond significantly affects OMTKY3 efficacy. The differences between the two subtilisin-bound OMTKY3 molecules, the change in structure at P4 when compared to other proteinase-bound ovomucoids and the loss of the hydrogen bond between Thr17I and Glu19I all suggest that this system is less optimal for the additivity-based sequence-to-reactivity algorithm since the equations derived for that algorithm require isostructural systems (Laskowski *et al.*, 2000).

4.4. Comparison of CARL–OMTKY3 complexes

With the two complexes seen in the asymmetric unit of our structure, there are now three complexes of CARL and OMKTY3 to compare. In all three complexes the protease is virtually unchanged (C^{α} r.m.s.d. of 0.25/0.24 Å for the two complexes seen here compared with that in Horn *et al.* (2003), a testament to the ability of OMKTY to inhibit subtilisin without changing the conformation of the enzyme. In contrast to this, one of the inhibitor molecules seen here is similar to the previous structure (0.67 Å C^{α} r.m.s.d.), whereas the other is vastly different (2.4 Å C^{α} r.m.s.d. for 51 C^{α} atoms). This change is independent of the binding of the inhibitor to the enzyme (0.24 Å and 0.15 Å C^{α} r.m.s.d. over the nine residues in the reactive-site loop). Indeed, it is quite remarkable that the reactive-site loop of OMKTY3 can remain in the active site of CARL, conserving important inhibitory interactions, whilst the inhibitor changes its structure so dramatically. This suggests that the dynamic movement of OMKTY3 may be extended beyond minor adjustments of the amino-acid side chains, as concluded by others (Horn *et al.*, 2003), to the polypeptide backbone as well, with the exclusion of the reactive-site loop.

This work was supported by grant MGP-37770 from the Canadian Institutes of Health Research to MNGJ and by equipment funds from the Alberta Heritage Foundation for Medical Research. JTM was supported by an MD/PhD Training Award from the Canadian Institutes of Health Research and the Canadian Gene Cure Foundation and supplemental funding from the Alberta Heritage Foundation for Medical Research and Pfizer. ML was supported by NIH grants GM10831 and GM63539-01. JTM would like to thank Thomas Earnest and James Holton at beamline 8.3.1 at the Advanced Light Source for their assistance and the Alberta Synchrotron Institute for funding data collection in Berkeley.

References

Apostol, I., Giletto, A., Komiyama, T., Zhang, W. & Laskowski, M. Jr (1993). *J. Protein Chem.* **12**, 419–433.
Ardelt, W. & Laskowski, M. Jr (1985). *Biochemistry.* **24**, 5313–5320.
Barrett, A. J., Rawlings, N. D. & Woessner, J. F. (1998). *Handbook of Proteolytic Enzymes.* London: Academic Press.
Bateman, K. S., Anderson, S., Lu, W., Qasim, M. A., Laskowski, M. Jr & James, M. N. G. (2000). *Protein Sci.* **9**, 83–94.
Bateman, K. S., Huang, K., Anderson, S., Lu, W., Qasim, M. A., Laskowski, M. Jr & James, M. N. G. (2001). *J. Mol. Biol.* **305**, 839–849.
Bode, W. & Huber, R. (1992). *Eur. J. Biochem.* **204**, 433–451.
Bode, W., Papamokos, E. & Musil, D. (1987). *Eur. J. Biochem.* **166**, 673–692.
Bode, W., Wei, A. Z., Huber, R., Meyer, E., Travis, J. & Neumann, S. (1986). *EMBO J.* **5**, 2453–2458.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *The PyMOL Users Manual*. San Carlos, CA, USA: DeLano Scientific.
Delbaere, L. T., Hutcheon, W. L., James, M. N. G. & Thiessen, W. E. (1975). *Nature (London)*, **257**, 758–763.
Fujinaga, M., Read, R. J., Sielecki, A., Ardelt, W., Laskowski, M. Jr & James, M. N. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 4868–4872.
Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M. Jr & James, M. N. G. (1987). *J. Mol. Biol.* **195**, 397–418.
Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.
Horn, J. R., Ramaswamy, S. & Murphy, K. P. (2003). *J. Mol. Biol.* **331**, 497–508.
Huang, K., Lu, W. Y., Anderson, S., Laskowski, M. & James, M. N. G. (1995). *Protein Sci.* **4**, 1985–1997.
Laskowski, M. Jr & Kato, I. (1980). *Annu. Rev. Biochem.* **49**, 593–626.
Laskowski, M. & Qasim, M. A. (2000). *Biochim Biophys Acta*, **1477**, 324–337.
Laskowski, M. Jr, Qasim, M. A. & Lu, S. M. (2000). *Protein–Protein Recognition*, ch. 8, edited by C. Kleanthous. Oxford University Press.
Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). *J. Mol. Biol.* **231**, 1049–1067.
Lu, S. M. *et al.* (2001). *Proc. Natl Acad. Sci. USA*, **98**, 1410–1415.
McGaughey, G. B., Gagne, M. & Rappe, A. K. (1998). *J. Biol. Chem.* **273**, 15458–15463.
McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–593.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Powell, H. R. (1999). *Acta Cryst.* D**55**, 1690–1695.
Radisky, E. S. & Koshland, D. E. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 10316–10321.

Read, R. J., Fujinaga, M., Sielecki, A. R. & James, M. N. (1983). *Biochemistry*, **22**, 4420–4433.
Read, R. J. & James, M. N. G. (1986). In *Proteinase Inhibitors*, edited by A. J. Barrett & G. Salvesen. Amsterdam: Elsevier.
Schechter, I. & Berger, A. (1967). *Biochem. Biophys. Res. Commun.* **27**, 157–162.
Yang, S. Q., Wang, C. I., Gillmor, S. A., Fletterick, R. J. & Craik, C. S. (1998). *J. Mol. Biol.* **279**, 945–957.