Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Atomic resolution structures of trypsin provide insight into structural radiation damage

Hanna-Kirsti Schrøder Leiros,[a] Seán M. McSweeney[b] and Arne O. Smalås[a]*

[a]Protein Crystallography Group, Department of Chemistry, Faculty of Science, University of Tromsø, N-9037 Tromsø, Norway, and [b]ESRF, 6 Rue Jules Horowitz, F-38043 Grenoble CEDEX, France

Correspondence e-mail: arne.smalas@chem.uit.no

Received 30 October 2000
Accepted 8 January 2001

PDB References: anionic salmon trypsin, 1hj8; bovine trypsin, 1hj9.

Radiation damage is an inherent problem in protein X-ray crystallography and the process has recently been shown to be highly specific, exhibiting features such as cleavage of disulfide bonds, decarboxylation of acidic residues, increase in atomic *B* factors and increase in unit-cell volume. Reported here are two trypsin structures at atomic resolution (1.00 and 0.95 Å), the data for which were collected at a third-generation synchrotron (ESRF) at two different beamlines. Both trypsin structures exhibit broken disulfide bonds; in particular, the bond from Cys191 to Cys220 is very sensitive to synchrotron radiation. The data set collected at the most intense beamline (ID14-EH4) shows increased structural radiation damage in terms of lower occupancies for cysteine residues, more breakage in the six disulfide bonds and more alternate conformations. It appears that high intensity and not only the total X-ray dose is most harmful to protein crystals.

1. Introduction

The lifetime of cryogenically cooled protein crystals exposed to synchrotron radiation has decreased using the new third-generation beamlines with insertion devices. The new beamlines are so intense that the proteins suffer badly from radiation damage; during routine data collection, loss of diffraction quality and increased sample mosaicity are often observed. In 1988, when cryocooling of crystals became feasible (Hope, 1988), the problem of radiation damage was generally thought to be eliminated. The lifetime of cryocooled crystals was improved spectacularly; however, Gonzalez *et al.* (1992) were the first to describe radiation damage in cryocooled crystals. Several investigators have tried to unravel details of the radiation-damage process (Gonzalez & Nave, 1994; Hedman *et al.*, 1985; Helliwell, 1988; Nave, 1995) and until recently radiation damage was believed to be a non-specific problem. Recent reports have given a better understanding of the problem (Teng & Moffat, 2000) and radiation-damage studies have all described the damage to be highly specific, *e.g.* on acetylcholinesterase (AChE), hen egg-white lysozyme (HEWL), a chymotrypsin inhibitor (Weik *et al.*, 2000; Ravelli & McSweeney, 2000) and myrosinase (Burmeister, 2000). All investigators have performed comprehensive studies by collecting consecutive data sets on protein crystals; between each data set the crystals were left unexposed, exposed to attenuated beam or exposed to the monochromatic (unattenuated) full beam. Comparison of successive data sets revealed an increase in atomic *B* factors, increased unit-cell parameters, slight rotation and translation of the protein molecules, breakdown of disulfide bridges and decarboxylation of aspartic and glutamic acids. It was postulated

that the changes are initiated by free radicals formed during X-ray exposure and that the absorbed X-ray dose needed to form radicals can be reached during routine data collection. Burmeister (2000) also reported specific damage (loss in electron density) to the hydroxyl groups of tyrosine and the methylthiol groups of methionine residues. The aims of these studies have been to understand the details of radiation damage, since it introduces non-isomorphism that can hamper phasing methods such as multiple isomorphous replacement (MIR) and multiwavelength anomalous dispersion (MAD).

The number of structures at atomic resolution is increasing rapidly; they have been used to provide detailed structure information on, for example, the position of the H atom on the catalytic histidine of elastase (Würtele *et al.*, 2000) and subtilisin (Kuhn *et al.*, 1998), the role of C—H···O interactions in β-sheets and deviations from well accepted dictionaries such as the Engh & Huber parameters (Ridder *et al.*, 1999; Wilson *et al.*, 1998). High-resolution structures can also give new details that are valuable for drug designers.

In the present study, we describe the structures of cold-adapted anionic salmon trypsin and bovine trypsin, which have previously been refined to 1.83 and 1.34 Å, respectively (PDB codes 1bit and 5ptp, respectively; Berglund, Smalås *et al.*, 1995; Finer-Moore *et al.*, 1992). The present structures are refined to atomic resolution, 1.00 and 0.95 Å, respectively. One data set was collected at an insertion-device beamline, ID14-EH4, while the other was collected on a bending-magnet beamline, the Swiss–Norwegian Beamline (SNBL), both at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. The observed radiation damage in both structures is described along with other findings. Trypsin has been chosen as a model system because crystals are easy to obtain and diffract to high resolution. The enzyme also has six disulfide bonds, which is more than any of the other proteins used so far in radiation-damage studies.

2. Experimental

2.1. Crystallization

Crystals of anionic salmon trypsin (AST) with benzamidine were grown by the hanging-drop method with only minor changes from the previously described conditions for AST crystal form II (Berglund, Smalås *et al.*, 1995; Smalås *et al.*, 1990). The initial protein solution contained 20 mg ml^{-1} protein and 60 m*M* benzamidine. The reservoir solution contained 0.7–0.8 *M* ammonium sulfate and 0.05 *M* citrate buffer pH 5.8 and the crystals were grown at 310 K. The crystal used for data collection had approximate dimensions of 0.15 × 0.2 × 0.3 mm and was flash-cooled in the reservoir solution containing 30% glycerol.

The bovine trypsin (BT) with aniline was also grown from a hanging drop, with the reservoir solution consisting of 25% polyethylene glycol (PEG) 8000, 0.2 *M* ammonium sulfate, 0.1 *M* Tris buffer pH 8 and 0.1 *M* aniline (Kurinov & Harrison, 1994). The drops were produced by mixing 5 µl protein solution at 15 mg ml^{-1} with 5 µl reservoir solution and the trays were stored at room temperature. The crystal had dimensions of roughly 0.15 × 0.3 × 0.4 mm and the cryoprotectant was the reservoir solution with addition of 15% glycerol.

2.2. Data collection and processing

Both data sets were collected at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France, but at two different beamlines; the procedures will be described for both experimental setups.

Data from AST with benzamidine were collected at ID14-EH4 at ESRF with an ADSC CCD detector from a single crystal cooled to 100 K. Data were collected with a crystal-to-detector distance of 120 mm, a detector offset height of 60 mm, 0.5° oscillation, 2 s exposure time and a total of 208 frames. The crystal was exposed at a second place to retain the highest resolution possible and 180 new frames were collected to 1.4 Å.

Data for BT were collected at the Swiss–Norwegian Beamline (SNBL) at 120 K from one single crystal on a MAR345 image-plate detector. First, one high-resolution data set was collected to 0.95 Å with a crystal-to-detector distance of 140 mm, 1° oscillation, a dose corresponding to 45 s exposure and a total of 150 images. The low-resolution data set was taken from the same crystal to a maximum of 1.5 Å, a dose corresponding to 4 s exposure, 2.5° oscillation per frame and 61 frames. The high-resolution data were collected with the largest MAR mode (345 mm diameter), whilst the low-resolution data set used the smallest mode available (180 mm diameter).

All data sets were autoindexed and processed with *DENZO* (Otwinowski, 1993); scaling and merging were performed in *SCALA* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

During the merging it became clear that the final frames collected (about 90 in total) from AST did not merge very well with the rest of the data set; they were therefore omitted from the final data set. AST had an $R_{\rm merge}$ per frame of 5.2% (598 reflections to 1.00 Å) for the first image and 9.8% (281 reflections to 1.40 Å) for the last frame used in the final data set. The final data is 76.5% complete, with an overall $R_{\rm merge}$ of 4.9% (Table 1). In order to increase the overall completeness we tried to incorporate previously collected data to 1.2 Å, but these data were clearly not isomorphous with the 1.00 Å data set; therefore, only images from one crystal were used. The overall temperature factor estimated from the Wilson plot (Wilson, 1942) was 8.61 Å^2.

All frames collected for BT scaled well together with an $R_{\rm merge}$ of 7.4% (5200 reflections to 0.95 Å) for the first image and 4.6% (2390 reflections to 1.5 Å) for the last. The overall $R_{\rm merge}$ was 4.8% and the overall *B* factor from the Wilson plot was 7.94 Å^2. The completeness of 99.5% and the redundancy of 6.9 are considered to be satisfactory. See Table 1 for further details.

2.3. Molecular replacement and refinement

Initial phases for the first map of AST were obtained by molecular replacement in *AMoRe* (Navaza, 1994) using the

Table 1
Data-collection statistics of anionic salmon trypsin (AST) and bovine trypsin (BT).

The outer-shell values (tenth bin) are shown in parentheses.

	AST	BT
Beamline	ID14-EH4	SNBL
Max. diffraction (Å)	1.00	0.95
Temperature (K)	100	120
Detector	ADSC CCD	MAR345
Wavelength (Å)	0.934	0.800
Unit-cell parameters (Å)	a = 72.63, b = 82.77, c = 30.92	a = 54.06, b = 56.81, c = 66.28
Space group	$P2_12_12$	$P2_12_12_1$
Solvent content (%)	36.5	42.0
No. of observations (∞–max. resolution)	174662	886215
No. of unique reflections	76713	128074
R_{merge} (%)	4.9 (23.2)	4.8 (47.2)
$I/\sigma(I)$	8.2 (3.2)	7.3 (1.6)
Completeness (%)	76.5 (53.4)	99.5 (99.6)
Wilson B factor (Å^2)	8.61	7.94
Absorbed dose (Gy)	6×10^7	7×10^6

Table 2
Final refinement statistics.

	AST	BT
Refinement programs	*X-PLOR*, *SHELXL*	*X-PLOR*, *SHELXL*
Resolution in final refinement (Å)	8.0–1.00	8.00–0.95
No. of reflections	76564	127765
R factor including all data (%)	11.86	11.71
R_{free} (%)	14.91	13.99
No. of protein atoms	1809	1689
No. of water molecules	228	255
No. of other solvent molecules†	1 Ca^{2+}, 1 SO_4^{2-}, 2 BZA	1 Ca^{2+}, 3 SO_4^{2-}, 1 ANL, 1 GOL
No. of alternative conformations	21	8
Mean B factors (Å^2)		
All atoms	12.78	13.31
Protein	11.18	11.33
Main chain	9.21	9.99
Side chain	13.46	12.96
Water molecules	24.90	25.16
Calcium ion	11.49	8.84
Inhibitor‡	11.61	26.09
R.m.s. deviations from ideal geometry		
Bond lengths (Å)	0.016	0.018
Angle distances (Å)	0.033	0.035
Luzzati (1952) coordinate error (Å)	0.056	0.050

† Benzamidine (BZA), aniline (ANL), glycerol (GOL). ‡ Located in the binding site.

room-temperature AST crystal form II (PDB code 1bit) as a search model. The cross-rotation and translation searches were performed with data in the resolution range 15–4.0 Å and an integration radius of 20 Å. The resulting Crowther angles and fractional translation vectors were α = 30.20, β = 3.11, γ = 327.77°, T_x = −0.0712, T_y = 0.0214, T_z = 0.1554. The BT phases were obtained from a refined structure of BT with benzylamine (Brandsdal & Smalås, unpublished work), only improved by rigid-body refinement.

The initial refinement included simulated annealing (3000 K) and conjugate-gradient minimization in *X-PLOR* (Brünger, 1992), accompanied by conjugate-gradient least-squares minimization with *SHELXL* (Sheldrick & Schneider, 1997). All atoms were refined with anisotropic B factors, but were restrained toward the atomic bond directions (DELU command) and according to their bond distance (SIMU command). Some residues (21 in AST and eight in BT) were refined with double conformations. At the very end of the refinements, rigid H atoms were added to the models, but not to the hydroxyl groups of Ser, Thr and Tyr, and not to the N atoms of the His side chains. All data were included in the final round. Solvent molecules were added to the models when the difference density exceeded 3.5σ.

Subsequent cycles of refinement and manual adjustments in *O* (Jones *et al.*, 1991) were performed with σ_A-weighted $F_o - F_c$ and $2F_o - F_c$ maps (Read, 1986); 5% of the data were reserved for cross validation using the free R factor.

The Ramachandran plots (Ramachandran & Sasisekharan, 1968) show 77.2 and 71.4% of residues in the inner core of the allowed regions of AST and BT, respectively, compared with the expected value of 80% (Kleywegt & Jones, 1996). Outliers in AST are Asn115 and Ser214, while Asp71 and Ser214 fall outside the expected regions in BT.

The expected value of the main-chain ω angle is 180°, but the distributions of the angle (Fig. 1) show that the values are not symmetrical around 180°. The highest deviation for BT is 21.7° for Glu70, a residue bound to the calcium ion, which is well defined in the electron-density maps. The mean value of the ω angle is 179.33° for AST and 178.80° for BT. Deviations from the ideal value of 180° for the ω angle are not uncommon in high-resolution structures (Ridder *et al.*, 1999; Walsh *et al.*, 1998; Wilson *et al.*, 1998).

3. Results

3.1. Final models

The final models of anionic salmon trypsin (AST) and bovine trypsin (BT) include 1809 and 1681 protein atoms,

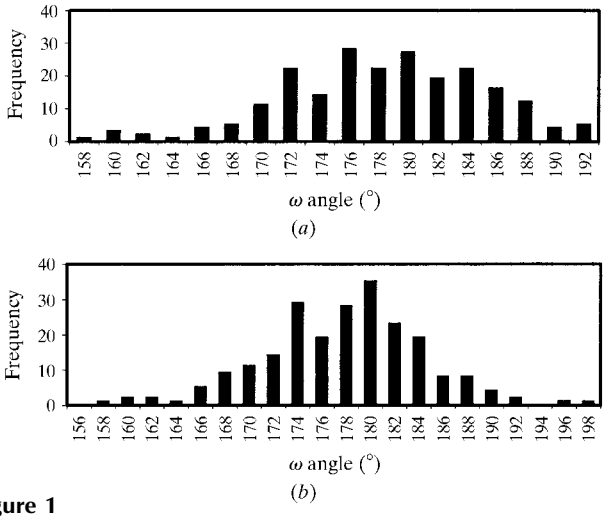


Figure 1
Histogram of ω main-chain angles (°) in (*a*) anionic salmon and (*b*) bovine trypsin.

respectively; the higher number in the former is a consequence of the greater number (21) of alternate conformations. The final R values of both trypsins are less than 12% and $R_{\rm free}$ is 14.9 and 14.0%, respectively (Table 2). The active sites are occupied by an inhibitor, benzamidine in AST and aniline in BT, and both models have one calcium ion in the calcium-binding loop. The number of solvent molecules incorporated as water molecules is 228 (AST) and 255 (BT), while there are three sulfate molecules and one glycerol in BT, and one sulfate and a second benzamidine in the AST structure. The position of the second benzamidine in AST is the same as the glycerol in BT. The benzamidine of AST stacks almost parallel to Phe82 and is adjacent to the symmetry-related Tyr217 and Tyr172. In BT, one of the glycerol O atoms is hydrogen bonded to the main-chain N atom of Phe82.

The common sulfate ion in the two structures is in the 'oxyanion hole' as observed for a number of other trypsin and elastase structures (see Schrøder *et al.*, 1998, and references therein). Some sulfates are refined with reduced occupancy.

The occupancy of the calcium ion in the AST structure was refined to 0.65 since negative difference electron density was observed at the calcium position.

The two structures superimpose very well, except for the known conformational difference of residues 145–153 in the autolysis loop (Smalås *et al.*, 1994), with an r.m.s. deviation for main-chain atoms of 0.65 Å excluding residues 145–153. In AST, residues 23–26 are poorly defined and the main-chain atoms of residues 24 and 25 are shifted about 3 Å relative to the BT structure. The side chain of Tyr25 is well defined in the AST density and the conformational shift seems to arise from aromatic interactions between Tyr25 and a symmetry-related Tyr39. The difference density observed in this area of AST can probably be explained by the high flexibility or multiple chain conformations of residues 23–26.

The number of atoms with zero occupancy is 37 in AST and 21 in BT; the higher number in the former is caused by loss of density for the last two residues of the C-terminal α-helix as in all other AST structures (Berglund, Smalås *et al.*, 1995; Helland, Berglund *et al.*, 1999; Helland *et al.*, 1998; Smalås *et al.*, 1994).

The BT structure, which has the highest number of unique reflections and the lowest estimated coordinate error (Table 2), also has the highest deviation from the Engh and Huber parameters (Engh & Huber, 1991) used in the refinement. BT possesses r.m.s. deviations of 0.018 and 0.035 Å for bond lengths and angle distances, respectively, compared with 0.016 and 0.033 Å for

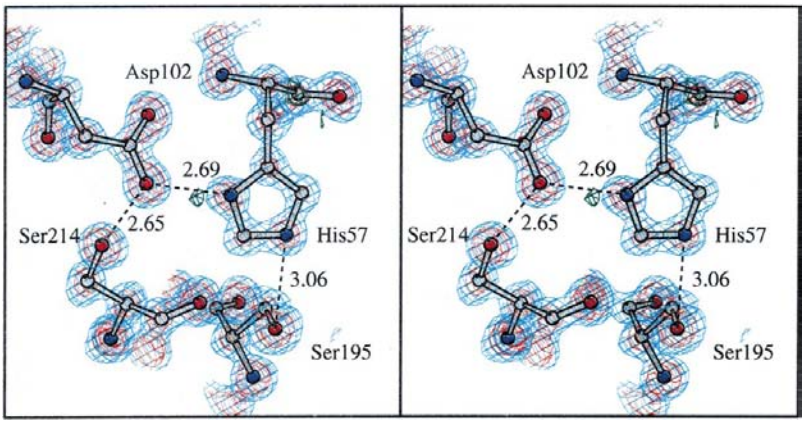


Figure 2
The catalytic hydrogen bond from His57 $N^{\delta1}$ to Asp102 $O^{\delta2}$ in anionic salmon trypsin (AST). The $2mF_o - DF_c$ map is contoured at 2.0σ (blue) and 4.0σ (red), while the $mF_o - DF_c$ map around the His57 side chain is contoured at +2.5σ (green). Ser195 and Ser214 are also included in the figure, which was created with *Bobscript* (Esnouf, 1997).

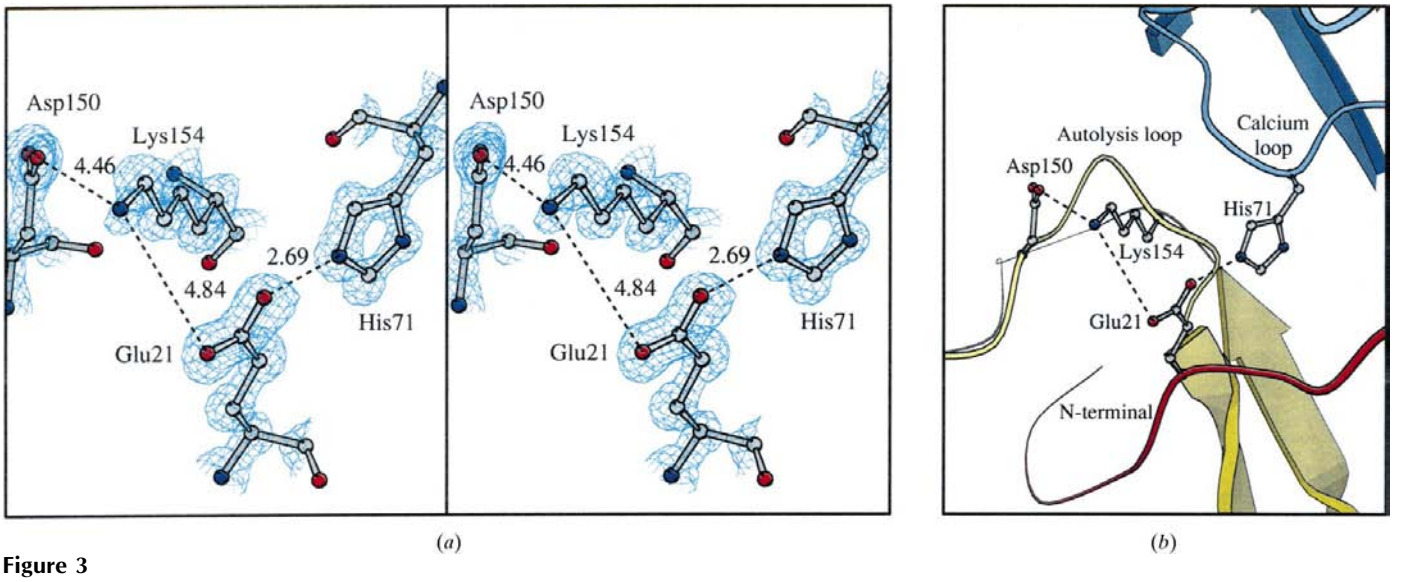


Figure 3
(*a*) Stereoview of the salt-bridge interactions that involve residues uniquely conserved in all cold-adapted trypsins (Leiros *et al.*, 1999, 2000). The $2mF_o - DF_c$ map of the side chains only is contoured at 1.8σ. (*b*) Location of the ion-pair network relative to the N-terminal (red), the calcium-binding loop (blue) and the autolysis loop (yellow).

AST. For comparison, the 1.15 Å structure of haloalkane dehalogenase reported by Ridder *et al.* (1999) possesses r.m.s. deviations of 0.014 and 0.029 Å for bond lengths and angle distances, respectively, and HEWL at 0.95 Å possesses r.m.s. deviations of 0.017 Å (bonds) and 0.044 Å (angles) (Walsh *et al.*, 1998).

The position of a H atom on the catalytic histidine has been reported in porcine pancreatic elastase (Würtele *et al.*, 2000) and in *Bacillus lentus* subtilisin (Kuhn *et al.*, 1998). In the case of AST, we were able to localize the position of the $HN^{\delta1}$ atom at 2.5σ in the difference density map (Fig. 2). The distance from the maximum of the peak to His57 $N^{\delta1}$ is 1.00 Å, whereas His57 $N^{\delta1}$ and Asp102 $O^{\delta2}$ are separated by 2.69 Å and the angle $N^{\delta1}-HN^{\delta1}-O^{\delta2}$ is 159°. These numbers are similar to those reported for the porcine elastase (Würtele *et al.*, 2000). The presence of His57 $HN^{\delta1}$ in BT is less obvious, but one hydrogen site could be assigned to Ser214 O^{γ} in the direction of Asp102 $O^{\delta2}$. In acetylcholinesterase (AChE), the active-site His440 is the histidine residue most affected by X-ray radiation. Weik *et al.* (2000) concluded that active-site residues are among the most radiation-sensitive residues, as their conformation often constitutes 'weak lines' or 'strained' configurations in protein structures. This does not seem to be the case for our trypsin structures, since all the residues in the catalytic triads are well defined in the electron-density maps.

3.2. Salt bridges

The conformational difference of the autolysis loop (residues 145–153) enables AST to form three unique ion-pair interactions involving residues Asp150, Lys154, Glu21 and His71, of which Glu21, Asp150 and Lys154 were found to be residue determinants for the cold-adapted fish trypsins (Leiros *et al.*, 1999). The ion-pair network has been pointed out to be a cold-adaptation determinant (Leiros *et al.*, 2000). In the cryocooled AST structure presented here, all residues involved are well defined in the electron-density map (Fig. 3*a*) and the distance from Lys154 N^{ζ} to Asp150 $O^{\delta2}$ is 4.46 Å and from

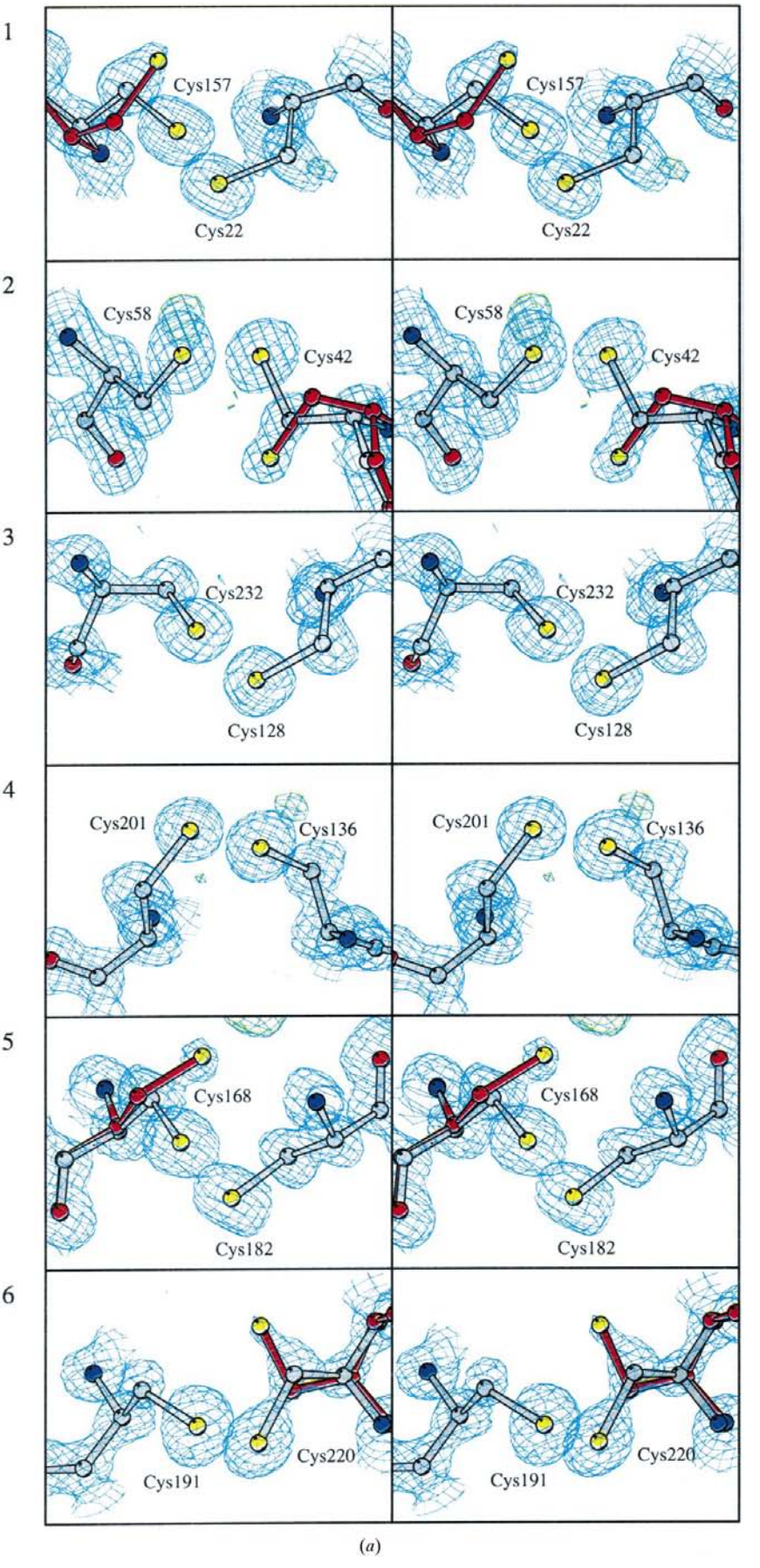
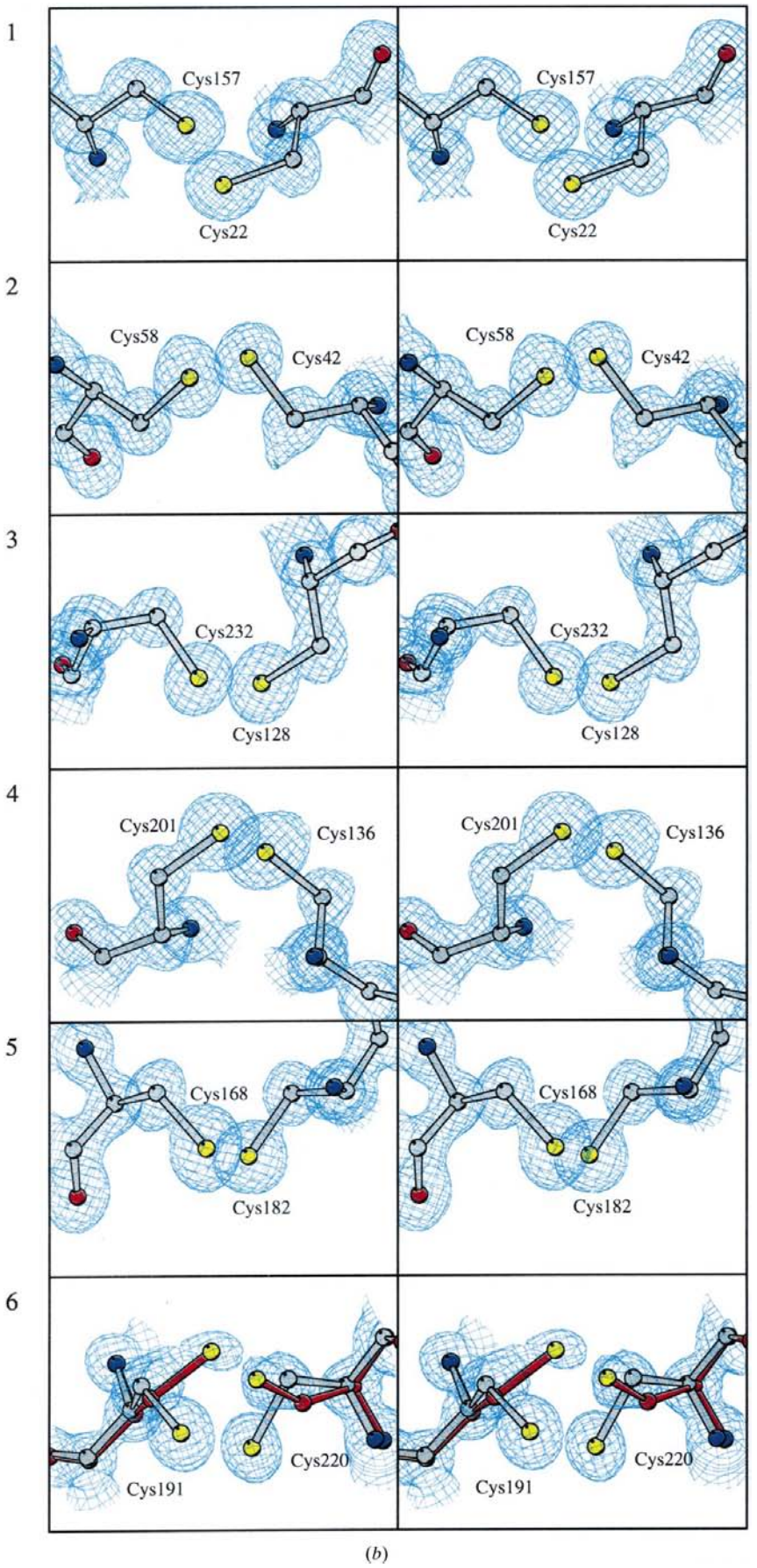


Figure 4
Electron-density maps of the disulfide bridges in (*a*) anionic salmon trypsin (AST) and (*b*) bovine trypsin (BT). The second conformations are given in red, $2mF_o - DF_c$ maps are contoured at 1.5σ (blue) and $mF_o - DF_c$ maps at 4σ (green) and −4σ (red).



(*b*)

Figure 4 (continued)

Lys154 N^{ζ} to Glu21 $O^{\varepsilon 1}$ is 4.84 Å, while it is 2.69 Å from Glu21 $O^{\varepsilon 2}$ to His71 $N^{\varepsilon 2}$. The two interactions involving Lys154 are beyond the cutoff criteria of 4.0 Å which is often used to define salt bridges; however, bearing in mind the long-range nature of electrostatic interactions, some stabilizing interactions still occur. The network connects the autolysis loop (residues 150 and 154) to the N-terminal loop (Glu21) and the N-terminal loop to the calcium-binding loop (His71) as illustrated in Fig. 3(*b*); this could stabilize the catalytically active trypsin. The network is very pH-dependent and could explain why the cold-adapted trypsins are unstable at low pH (Ásgeirsson *et al.*, 1989; Genicot *et al.*, 1988; Outzen *et al.*, 1996).

3.3. Disulfide bridges

All the recent studies on the specific nature of X-ray radiation damage (Weik *et al.*, 2000; Ravelli & McSweeney, 2000; Burmeister, 2000) reported an increase in *B* factors for cysteines, glutamic acids and aspartic acids for successive data sets of the proteins investigated. The increase in *B* factors for cysteines is caused by the breakage of disulfide bonds. A clear ordering of the disulfide bonds in terms of susceptibility to radiation damage was also observed; additionally, within each disulfide bond, one S^{γ} atom seemed to be more labile whilst the other remained in the same position.

In the two trypsin structures presented here, none of the six disulfide bridges in AST have both cysteine residues refined with full occupancy, whereas three bridges in BT have both Cys residues with full occupancy, namely 22–57, 136–201 and 168–182 (Table 3). Four of the 12 cysteines in AST are in an alternate conformation and nine of the 12 S^{γ} atoms have reduced occupancy, while BT has two double conformations and six S^{γ} atoms with reduced occupancy (Figs. 4*a* and 4*b*).

There is some correlation between high water-accessible surface area (WAS) and reduced occupancy (Table 3). All S^{γ} atoms in both models with full occupancy are totally buried with a WAS of 0 and all S atoms (except one) with WAS $\geq$ 1 $Å^2$ have reduced occupancies of 0.90–0.65.

The Cys191–Cys220 disulfide bridge is partly disordered in both trypsin structures

Table 3
Water-accessible surface area (WAS, Å^2) for the S$^\gamma$ atoms and occupancy for the cysteine residues in the six disulfide bridges of AST and BT.

WAS for the main conformation (Conf. 1) is calculated without any second conformations (2); values for the second conformations are given in parentheses.

(*a*) AST.

No.	Residue No.	WAS Conf. 1 (2)	Occupancy Conf. 1 (2)	Residue No.	WAS Conf. 1 (2)	Occupancy Conf. 1 (2)
1	Cys22	0	0.75†	Cys157	1 (1)	0.90 (0.10)
2	Cys42	7 (0)	0.75 (0.25)	Cys58	0	1.00
3	Cys128	20	0.80†	Cys232	1	0.80†
4	Cys136	0	1.00	Cys201	0	0.95†
5	Cys168	1 (1)	0.90 (0.10)	Cys182	0	1.00
6	Cys191	7	0.80†	Cys220	8 (1)	0.80 (0.20)

(*b*) BT.

No.	Residue No.	WAS Conf. 1 (2)	Occupancy Conf. 1 (2)	Residue No.	WAS Conf. 1 (2)	Occupancy Conf. 1 (2)
1	Cys22	0	1.00	Cys157	2	1.00
2	Cys42	6	0.80†	Cys58	0	0.91†
3	Cys128	28	0.70†	Cys232	2	0.70†
4	Cys136	0	1.00	Cys201	0	1.00
5	Cys168	0	1.00	Cys182	0	1.00
6	Cys191	2 (0)	0.65 (0.35)	Cys220	15 (0)	0.67 (0.33)

† S$^\gamma$ atom only.

and the bridge is also weakly defined in other trypsin structures refined in our group (R. Helland, personal communication; Helland, Otlewski *et al.*, 1999). The bridge connects the two walls of the specificity pocket of trypsin and the bridge is the second most water-exposed bridge in the structure. The most exposed connection is Cys128–Cys232, which bridges the beginning of the C-terminal α-helix and the interdomain loop (see Leiros *et al.*, 1999, for nomenclature).

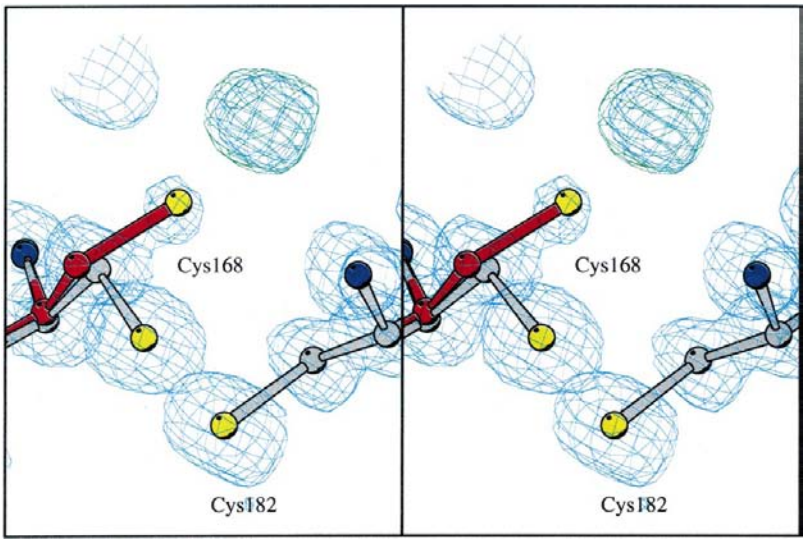


Figure 5
The disulfide bridge Cys168–Cys182 in anionic salmon trypsin (AST) and the uninterpretable density near the second conformation of Cys168. The $2mF_o - DF_c$ map is contoured at 1.5σ (blue) and the $mF_o - DF_c$ map at 4σ (green) and −4σ (red).

The disulfide bridge Cys168–Cys182 in AST has alternate conformations for residue 168 but also has one large positive peak in the difference electron-density map (Fig. 5). If this extra density had been assigned as a water molecule, it would have been 3.7 Å from the S$^\gamma$ on Cys168 but only 1.8 Å from the S$^\gamma$ in the second conformation. The role and nature of this solvent molecule is not clear.

In AST some additional electron density is seen near Cys58 S$^\gamma$ (Fig. 4*a*), but no second conformation could be assigned to this position. Also, the extra density at Cys22 and Cys136 indicates some movements or cleavage of these bridges. In BT, Cys42 has a slight indication of a second conformation, but no such indicators are seen for any of the other cysteines.

The observation of one anchor and one separator within the bridge (Ravelli & McSweeney, 2000) is not clearly seen from our results. For BT, both cysteines within a bridge are either fully occupied or both have reduced occupancy, while in AST some anchors could be assigned. AST has four bridges, with one cysteine in a second conformation; the remaining cysteines could then be denoted the anchors. This is valid for Cys42–Cys58(A), Cys168–Cys182(A) and Cys191(A)–Cys220, where A indicates anchor.

3.4. Loss of definition

The specific radiation damage has been reported as loss in definition of the carboxyl groups of aspartic and glutamic acid residues and also the loss of the hydroxyl groups of tyrosine and the methylthiol groups of methionine residues (Burmeister, 2000; Ravelli & McSweeney, 2000; Weik *et al.*, 2000). In our structures, AST has reduced occupancy for Tyr20 C$^\zeta$ and O$^\eta$ and Glu77 C$^\delta$, O$^{\varepsilon1}$ and O$^{\varepsilon2}$, occupancy 0.8 for Asp236 O$^{\delta2}$, Met104 C$^\gamma$, Met135 S$^\delta$, Met175 S$^\delta$, occupancy zero for Glu49 C$^\delta$, Glu186 O$^{\varepsilon1}$ and O$^{\varepsilon2}$, and high *B* factors (>30 Å^2) for Met145 C$^\varepsilon$, Met175 C$^\varepsilon$, Tyr184 O$^\eta$ and Tyr217 O$^\eta$. BT has poorly defined carboxyl groups (occupancy zero) for Asp153, C$^\delta$ and O$^{\varepsilon1}$ of Glu186, whereas Asp165 has occupancy 0.6 for the carboxyl group. Met104 has one alternate conformation. Residues that could be decarboxylated in AST are Glu77 and Glu186 and in BT are Asp153 and Glu186.

The calcium ion of AST has occupancy 0.65, as already mentioned, and one of the ligands, Glu77, is poorly defined in the electron-density map. The carboxyl group of Glu77 has reduced occupancy (0.64–0.88), the density of the side chain is poorer compared with the other ligands and the nearby residue Ser79 has two conformations (Fig. 6*a*). A comparison of room-temperature AST (PDB code 1bit; Berglund, Smalås *et al.*, 1995) and cryocooled AST shows that the calcium-binding mode, the ligands and Ser79 (main conformation) superimpose very well. In the room-temperature structure, Tyr25 O$^\eta$ is hydrogen bonded to Ser79 O$^\gamma$ (2.93 Å) and Tyr25 O$^\eta$ is at a distance

of 4.5 Å from the carboxyl O atoms of Glu77. These interactions are not possible in the cryocooled AST as the N-terminal loop has moved and the OH group of Tyr25 is shifted 9 Å relative to the room-temperature structure (Fig. 6*b*). In the cryocooled structure Glu77 becomes more exposed to the solvent with higher WAS, it is more poorly defined and it seems more sensitive to radiation damage and decarboxylation. A decarboxylation can further expel the calcium binding, since Glu77 is one of the calcium-binding ligands; this might explain the low occupancy (0.65) for the calcium ion.

3.5. Dose calculations

The typical flux at ID14-EH4 has been found to be of the order of 5×10^{12} photons s^{-1} through a pinhole collimator with diameter of 0.15 mm, whereas the flux at SNBL would be of the order of 1×10^{12} photons s^{-1} for a 0.5×0.5 mm focal spot. Using the crystal dimensions $0.15 \times 0.2 \times 0.3$ mm for AST and $0.15 \times 0.3 \times 0.4$ mm for BT and the densities of the crystals, we can calculate the absorbed dose as described by Ravelli & McSweeney (2000). The doses absorbed are 6×10^7 Gy (J kg^{-1}) for AST and 7×10^6 Gy for BT (Table 1). In comparison, the absorbed dose per data set for HEWL was 10^5 Gy (with attenuated beam) and 10^6 Gy (without attenuator), 10^7 Gy for AChE and 10^6 Gy for the chymotrypsin inhibitor (Weik *et al.*, 2000; Ravelli & McSweeney, 2000), and 2.5×10^{12} photons mm^{-2} ($\sim 1.1 \times 10^5$ Gy) for myrosinase (Burmeister, 2000).

Henderson (1990) estimated that cooled protein crystals could absorb a dose of 2×10^7 Gy before the crystalline diffraction would be destroyed. New experiments on HEWL have shown that when the crystal is exposed to a dose which exceeds 1×10^7 Gy the linear relation of dose to increased *B* factors, dose to increase in unit-cell volume or dose to decrease in intensities is lost. When the dose exceeds this threshold, secondary and perhaps tertiary radiation-damage effects become more significant (Teng & Moffat, 2000). For both structures presented here the dose is of the order of 1×10^7 Gy and secondary radiation-damage effects could be observed.

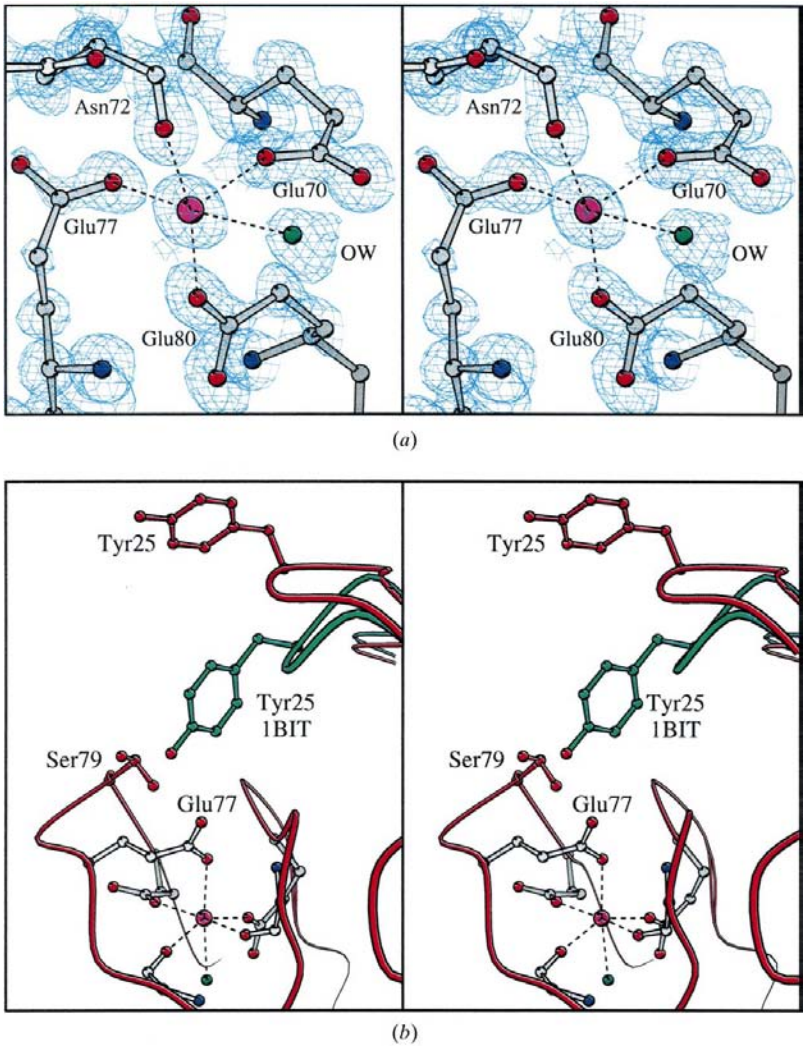


Figure 6
The arrangement of the calcium binding in AST with (*a*) the $2mF_o - DF_c$ map (blue at 1.8σ) and the calcium ligands. The main-chain O atom of Val75 is bound to the calcium ion but is omitted here in order to simplify the figure. (*b*) shows the calcium ligands as in (*a*); Ser79 is shown in two conformations and the arrangement of Tyr25 and the N-terminal loop of the cryocooled (red) and room-temperature (green; PDB code 1bit) AST is included.

4. Conclusions

Synchrotron data for two trypsin structures have been collected at two different beamlines at ESRF in Grenoble, one at the insertion device of ID14-EH4 and the other at the bending magnet of SNBL (BM01). During the data collection it was clear that the crystal of anionic salmon trypsin (AST) exposed at ID14-EH4 suffered more from radiation damage than the bovine trypsin crystal (BT) exposed at SNBL. For BT, more of the unique reflections could be recorded from one single crystal, resulting in an overall completeness of 99.5%. AST lacks both low- and high-resolution data and has an overall completeness of 76.5%. Even at this early stage of the data collection and refinement, it became clear that the trypsin crystal exposed at the insertion-device beamline (ID14-EH4) suffered significantly from the intense X-ray beam.

Structural differences confirmed the damage through a greater number of alternate conformations in AST (21) compared with BT (eight) and more breakages in the disulfide bridges. Both trypsins have six disulfide bridges and four of these have one cysteine residue with alternate conformations in AST, while in BT one disulfide bridge, Cys191–Cys220, is broken with both cysteines in dual conformations. From the remaining cysteines, AST has reduced occupancy for five of the eight S^{γ} atoms, while BT has lower occupancy for four of the ten S^{γ} atoms.

The observed structural changes could be caused by primary radiation damage such as photoelectric absorption. The incoming X-ray photon is then absorbed and a lower level core electron is ejected from the atom. The photoabsorption by sulfur is 80 times higher than for carbon and nitrogen at the energies used in the experiment (Ravelli & McSweeney, 2000) and all sulfurs in the structures should be affected by the X-ray exposure. We observed only breakage of some disulfide bonds and low occupancy for two of the six methionines in AST and no change for the two methionines of BT. Since the S atoms are not affected equally, primary radiation damage alone can not explain the observed structural changes.

When photoabsorption occurs in water, hydrated electrons (e^-_{aq}), hydroxyl and hydrogen radicals are formed, with highest yields of the two former. When these species react with the protein, secondary radiation damage occurs. The hydrated electrons (e^-_{aq}) have been found to react one to two orders of magnitude faster with cysteines compared with other amino acids (Favaudon *et al.*, 1990). If the major source for e^-_{aq} is the solvent, the most water-exposed disulfide bonds should be the most susceptible to radiation damage. We found some correlation between high water-accessible surface and low occupancy for the S atoms in the cysteine residues, as also reported by Weik *et al.* (2000). Ravelli & McSweeney (2000) found no clear correlation and concluded that if electrons are a major cause of the specific radiation damage, the electrons must be mobile both within the protein and in the solvent.

The disulfide bridges have been ordered of in terms of susceptibility (Weik *et al.*, 2000; Ravelli & McSweeney, 2000). These authors questioned whether the order of bond breakage was related to the natural stability of the bonds and/or to the order in which they are formed during protein folding. This has been shown for HEWL, where the disulfide bond formed last in the folding process (van den Berg *et al.*, 1999) is ranked second in terms of radiation-damage susceptibility. Our study should give the same order of weakness for the disulfide bonds in the two trypsin structures, but we can only rank Cys191–Cys220 as being the most fragile. This bond is refined with alternate conformations in both AST and BT and is broken in many other trypsin structures refined in our group. To our knowledge, there have not been any folding studies reported on trypsin which order the six disulfide bridges in the folding of the native enzyme and we therefore cannot correlate our observation to the folding process.

The calculated X-ray dose which the crystals have been absorbing during the data collection was estimated to be 6×10^7 Gy for AST and 7×10^6 Gy for BT. These doses were obtained during 10 min exposure time for AST, while BT reached this dose during 1 h 50 min of data collection. The comparison of the structures from two different crystals at the two different beamlines then indicates that apart from the total X-ray dose the incoming photon flux might also cause severe deterioration of protein crystals.

The 0.95 Å structure of BT has the highest r.m.s. deviations compared with the Engh and Huber parameters (Engh & Huber, 1991), but the structure also has the highest fraction of reflections per protein atom and the lowest coordinate error and should therefore be the one most accurately defined. This supports the observations of Ridder *et al.* (1999) and Wilson *et al.* (1998) that atomic resolution structures deviate from what has generally been accepted as 'ideal geometry'. It also demonstrates the necessity of reconsidering the currently used refinement libraries. The fact that the number of atomic resolution structures is increasing and that many of the dictionaries are based on small-molecule structures and not protein structures also supports this action.

Despite our hopes, the two last residues of the C-terminus of AST were not defined in this first cryocooled structure of the enzyme. Definition of the C-terminus would tell us whether the formation of two salt bridges from the C-terminal carboxyl group to Arg87 and Lys107 is possible, as found from the molecular-dynamic simulations (Heimstad *et al.*, 1995) and in the crystal structures of BT. In other trypsin structures, there are two ion pairs from the C-terminal carboxyl group to Lys87 and Lys107 and these interactions stabilize the structure by binding the C-terminal α-helix to the N-terminal barrel of the molecule. Since the ion pairs are not present in the cold-adapted AST and are also not found in the crystal structure of the cold-adapted salmon elastase (Berglund, Willassen *et al.*, 1995), it might explain their reduced stability at elevated temperatures.

The present comparative study shows that a higher amount of absorbed dose and high beam intensity yield more radiation damage such as broken disulfide bonds. Additional experiments are needed to obtain more details on the radiation-damage mechanism and also to find experimental setups that will reduce the damage in the protein crystal. Crystal cooling in liquid helium (40 K) has been tried, but did not have any dramatic effect on the observed radiation damage (Ravelli & McSweeney, 2000; Yonath *et al.*, 1998).

This work has been funded by the Norwegian Research council (HKSL, AOS); HKSL acknowledges support from the Training and Mobility of Researchers (TMR) Access to Large-Scale Facilities. The organizers at the Swiss–Norwegian Beamline (SNBL) are acknowledged for providing us with beamtime and Philip Pattison is thanked for providing details from the SNBL beamline. We would also like to thank Dr Ingar Leiros for valuable assistance during data collection and Dr Raimond Ravelli for help with the dose calculations.

References

Ásgeirsson, B., Fox, J. W. & Bjarnason, J. B. (1989). *Eur. J. Biochem.* **180**, 85–94.

Berg, B. van den, Chung, E. W., Robinson, C. V., Mateo, P. L. & Dobson, C. M. (1999). *EMBO J.* **18**, 4794–4803.

Berglund, G. I., Smalås, A. O., Hordvik, A. & Willassen, N. P. (1995). *Acta Cryst.* D**51**, 725–730.

Berglund, G. I., Willassen, N. P., Hordvik, A. & Smalås, A. O. (1995). *Acta Cryst.* D**51**, 925–937.

Brünger, A. T. (1992). *X-PLOR, Version 3.1. A System for X-ray Crystallography and NMR.* New Haven/London: Yale University Press.

Burmeister, W. P. (2000). *Acta Cryst.* D**56**, 328–341.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
Favaudon, V., Tourbez, H., Houee-Levin, C. & Lhoste, J. M. (1990). *Biochemistry*, **29**, 10978–10989.
Finer-Moore, J. S., Kossiakoff, A. A., Hurley, J. H., Earnest, T. & Stroud, R. M. (1992). *Proteins*, **12**, 203–222.
Genicot, S., Feller, G. & Gerday, C. (1988). *Comp. Biochem. Physiol. B*, **90**, 601–609.
Gonzalez, A. & Nave, C. (1994). *Acta Cryst.* D**50**, 874–877.
Gonzalez, A., Thompson, A. & Nave, C. (1992). *Rev. Sci. Instrum.* **63**, 1177–1180.
Hedman, B., Hodgson, K. O., Helliwell, J. R., Liddington, R. & Papiz, M. Z. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 7604–7607.
Heimstad, E. S., Hansen, L. K. & Smalås, A. O. (1995). *Protein Eng.* **8**, 379–388.
Helland, R., Berglund, G. I., Otlewski, J., Apostoluk, W., Andersen, O. A., Willassen, N. P. & Smalås, A. O. (1999). *Acta Cryst.* D**55**, 139–148.
Helland, R., Leiros, I., Willassen, N. P., Berglund, G. I. & Smalås, A. O. (1998). *Eur. J. Biochem.* **256**, 317–324.
Helland, R., Otlewski, J., Sundheim, O., Dadlez, M. & Smalås, A. O. (1999). *J. Mol. Biol.* **287**, 923–942.
Helliwell, J. R. (1988). *J. Cryst. Growth*, **90**, 259–272.
Henderson, R. (1990). *Proc. R. Soc. London Ser. B*, **241**, 6–8.
Hope, H. (1988). *Acta Cryst.* B**44**, 22–26.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **4**, 1395–1400.
Kuhn, P., Knapp, M., Soltis, S. M., Ganshaw, G., Thoene, M. & Bott, R. (1998). *Biochemistry*, **37**, 13446–13452.
Kurinov, I. V. & Harrison, R. W. (1994). *Nature Struct. Biol.* **1**, 735–743.
Leiros, H.-K. S., Willassen, N. P. & Smalås, A. O. (1999). *Extremophiles*, **3**, 205–219.
Leiros, H.-K. S., Willassen, N. P. & Smalås, A. O. (2000). *Eur. J. Biochem.* **267**, 1039–1049.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Nave, C. (1995). *Radiat. Phys. Chem.* **45**, 483–490.
Otwinowski, Z. (1993). *DENZO: An Oscillation Data Processing Program for Macromolecular Crystallography*. Yale University, New Haven, CT, USA.
Outzen, H., Berglund, G. I., Smalås, A. O. & Willassen, N. P. (1996). *Comp. Biochem. Physiol. B*, **115**, 33–45.
Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–437.
Ravelli, R. B. & McSweeney, S. M. (2000). *Structure Fold. Des.* **8**, 315–328.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Ridder, I. S., Rozeboom, H. J. & Dijkstra, B. W. (1999). *Acta Cryst.* D**55**, 1273–1290.
Schrøder, H.-K., Willassen, N. P. & Smalås, A. O. (1998). *Acta Cryst.* D**54**, 780–798.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **276**, 319–343.
Smalås, A. O., Heimstad, E. S., Hordvik, A., Willassen, N. P. & Male, R. (1994). *Proteins*, **20**, 149–166.
Smalås, A. O., Hordvik, A., Hansen, L. K., Hough, E. & Jynge, K. (1990). *J. Mol. Biol.* **214**, 355–358.
Teng, T.-Y. & Moffat, K. (2000). *J. Synchrotron Rad.* **7**, 313–317.
Walsh, M. A., Schneider, T. R., Sieker, L. C., Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1998). *Acta Cryst.* D**54**, 522–546.
Weik, M., Ravelli, R. B., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J. & Sussman, J. L. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 623–628.
Wilson, A. J. C. (1942). *Nature (London)*, **150**, 151–152.
Wilson, K. S., Butterworth, S., Dauter, Z., Lamzin, V. S., Walsh, M., Wodak, S., Pontius, J., Richelle, J., Vaguine, A., Sander, C., Hooft, R. W. W., Vriend, G., Thornton, J. M., Laskowski, R. A., MacArthur, M. W., Dodson, E. J., Murshudov, G., Oldfield, T. J., Kaptein, R. & Rullmann, J. A. C. (1998). *J. Mol. Biol.* **276**, 417–436.
Würtele, M., Hahn, M., Hilpert, K. & Höhne, W. (2000). *Acta Cryst.* D**56**, 520–523.
Yonath, A., Harms, J., Hansen, H. A., Bashan, A., Schlunzen, F., Levin, I., Koelln, I., Tocilj, A., Agmon, I., Peretz, M., Bartels, H., Bennett, W. S., Krumbholz, S., Janell, D., Weinstein, S., Auerbach, T., Avila, H., Piolleti, M., Morlang, S. & Franceschi, F. (1998). *Acta Cryst.* A**54**, 945–955.