Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The structures of T_6, T_3R_3 and R_6 bovine insulin: combining X-ray diffraction and absorption spectroscopy

Christian Grundahl Frankær,[a] Marianne Vad Knudsen,[a] Katarina Norén,[b] Elena Nazarenko,[c] Kenny Ståhl[a] and Pernille Harris[a]*

[a]Department of Chemistry, Technical University of Denmark, Kemitorvet B207, 2800 Kgs. Lyngby, Denmark, [b]MAX-lab, Lund University, PO Box 118, 22 100 Lund, Sweden, and [c]Department of Chemistry, Biochemistry and Biophysics, University of Gothenborg, 40 530 Gothenborg, Sweden

Correspondence e-mail: ph@kemi.dtu.dk

Received 28 March 2012
Accepted 10 June 2012

PDB References: bovine insulin, T_6, 4e7t; T_3R_3, 4e7u; R_6, 4e7v.

The crystal structures of three conformations, T_6, T_3R_3 and R_6, of bovine insulin were solved at 1.40, 1.30 and 1.80 Å resolution, respectively. All conformations crystallized in space group *R*3. In contrast to the T_6 and T_3R_3 structures, different conformations of the N-terminal B-chain residue PheB1 were observed in the R_6 insulin structure, resulting in an eightfold doubling of the unit-cell volume upon cooling. The zinc coordination in each conformation was studied by X-ray absorption spectroscopy (XAS), including both EXAFS and XANES. Zinc adopts a tetrahedral coordination in all R_3 sites and an octahedral coordination in T_3 sites. The coordination distances were refined from XAS with a standard deviation of <0.01 Å. In contrast to the distances determined from the medium-resolution crystal structures, the XAS results were in good agreement with similar coordination geometries found in small molecules, as well as in other high-resolution insulin structures. As the radiation dose for XRD experiments is two orders of magnitude higher compared with that of XAS experiments, the single crystals were exposed to a higher degree of radiation damage that affected the zinc coordination in the T_3 sites in particular. Furthermore, XANES spectra for the zinc sites in T_6 and R_6 insulin were successfully calculated using finite difference methods and the bond distances and angles were optimized from a quantitative XANES analysis.

1. Introduction

Insulin is a protein hormone produced in the pancreas. It is mainly known for its regulation of the sugar level in the blood and is therefore given to patients suffering from diabetes. The insulin molecule is built up of a 21-residue A chain and a 30-residue B chain which are connected by two interchain disulfide bonds with an intrachain disulfide bond in the A chain. The sequence of bovine insulin differs from human insulin at three residues: A8, A10 and B30 (Thr, Ile and Thr in human and Ala, Val and Ala in bovine, respectively).

In the presence of Zn^{2+} ions, insulin associates into a hexamer binding two zinc ions. Both Zn^{2+} ions are located on the threefold symmetry axis passing through the hexamer and they coordinate to three symmetry-related histidine $N^{\varepsilon 2}$ atoms. In its hexameric form insulin is an allosteric complex in which the first eight residues of chain B adopt different conformations: tensed (T), in which residues B1–B8 adopt an extended conformation, or relaxed (R), in which the central B-chain α-helix is continuous from B1 to B19 (Kaarsholm *et al.*, 1989).

Three different conformations are known. The T_6 conformation occurs when all six B chains adopt a T conformation.

Table 1
Reservoir composition for single-crystal growth of T_6 and R_6 insulin using the vapour-diffusion technique.

	T_6 insulin	R_6 insulin
Sodium citrate (*M*)	0.05	0.12
Zinc acetate (m*M*)	0.5	17
Acetone [%(*v*/*v*)]	15	22
m-Cresol† [%(*v*/*v*)]	—	0.5
NaCl (*M*)	—	1.67
pH	7.2	8.4

† The *m*-cresol was dissolved in the acetone before addition.

In the presence of high lyotropic anion concentrations, one of the two trimers which constitute the hexamer will adopt the R conformation while the other will remain tensed (T_3R_3). The presence of a phenol derivative fully stabilizes the N-terminal B-chain helix and both trimers adopt the R conformation (R_6).

In the present work, all three conformations of bovine insulin have been studied by X-ray diffraction (XRD) and X-ray absorption spectroscopy (XAS). XRD and XAS complement each other well and have previously been combined for the study of metalloproteins (see Hasnain & Hodgson, 1999; Hasnain & Strange, 2003; Strange *et al.*, 2005; Arcovito *et al.*, 2007; Yano & Yachandra, 2008 and references therein). XAS is element-specific and can be used directly on proteins in solution as well as in crystals; it probes the structure around the absorbing metal atom within a radius of 5 Å. XAS covers both extended X-ray absorption fine-structure (EXAFS) spectroscopy and X-ray absorption near-edge structure (XANES) spectroscopy. Advanced fitting algorithms for XANES spectra have recently been developed and have been improved during the last decade (Benfatto *et al.*, 2001; Smolentsev & Soldatov, 2006; Sarangi *et al.*, 2008; Jacquamet *et al.*, 2009). The main advantage of quantitative XANES analysis is the possibility of obtaining information about the three-dimensional environment of the absorbing atom. Unlike the EXAFS method, which only provides the radial distribution of atoms, XANES enables both bond lengths and angles to be extracted. This information is complementary to that obtained from EXAFS spectroscopy. However, XAS does not allow direct extraction of structural information from the spectrum; it requires a starting structural model.

In contrast, XRD probes the entire protein structure, but to match the accuracy from XAS around the metal ions, ultra-high-resolution structures are required. At medium resolution (1.5–2.5 Å) bond lengths are determined with lower accuracy. Harding *et al.* (2010) reported how the accuracy of the bond distances around the metal atom depends on the resolution. Radiation damage may be more important for protein crystals containing heavy elements, as they generally absorb more radiation and therefore have shorter lifetimes (Holton, 2009).

The work presented here illustrates the complementarities between the two techniques. The structures of all three conformations of bovine insulin have been solved by single-crystal XRD. Two of the three structures that we present here are new, as only the structure of the T_6 conformation has previously been solved for bovine insulin (Smith *et al.*, 2005).

Table 2
Crystallization conditions for the preparation of microcrystals of T_6, T_3R_3 and R_6 insulin by batch crystallization.

The compositions are given as the concentration of each compound in the final crystallization solution.

	T_6 insulin	T_3R_3 insulin	R_6 insulin
Insulin (mg ml^{-1})	5.0	6.5	5.0
HCl (*M*)	0.01	0.01	0.01
Zinc acetate (m*M*)	5.0	7.0	7.5
Sodium citrate (*M*)	0.02	0.05	0.05
Acetone [%(*v*/*v*)]	15	15	20
KSCN (*M*)	—	0.2	—
m-Cresol† [%(*v*/*v*)]	—	—	0.5
NaCl (*M*)	—	—	1.0
pH	6.0	6.4	8.2

† The *m*-cresol was dissolved in the acetone before addition.

Despite the fact that insulin is a well studied protein, only one EXAFS study of the zinc sites in hexameric insulin has been reported (Bordas *et al.*, 1983). As the EXAFS technique has become more established, we have performed new studies of the zinc coordination using both EXAFS and XANES.

Finally, we have compared our XAS results with those observed for small molecules with similar coordination geometries, as well as in other insulin structures, to evaluate the accuracy of XAS.

2. Experimental

Lyophilized insulin from bovine (*Bos taurus*) pancreas was purchased from Sigma–Aldrich (catalogue No. I-5500). Other reagents used were stock chemicals of biological grade.

2.1. Growth of single insulin crystals

After several crystallization trials, it was found that the concentration of Zn^{2+} was crucial to the growth of large single crystals. Zinc ions were removed from the insulin before use. Metal-free insulin was prepared by chelation using the method of Coffman & Dunn (1988). After chelation, the supernatant was separated from the resin by syringe filtration (0.22 µm). The solution was then diluted with Milli-Q water. All concentrations were determined by UV–Vis spectrophotometry using an extinction coefficient at 280 nm (ε_{280}) of 5960 M^{-1} cm^{-1} as estimated by *ProtParam* (Gasteiger *et al.*, 1995).

For the growth of T_6 and R_6 insulin crystals, the vapour-diffusion technique was used. 2 µl of a solution consisting of 5 mg ml^{-1} insulin adjusted to pH 6.8 using aqueous HCl was mixed with 2 µl reservoir solution and equilibrated in a hanging drop against a 1 ml reservoir with the composition given in Table 1. After 7 d, crystals of dimensions 200–500 µm were observed.

T_3R_3 insulin was crystallized using the batch crystallization technique. The crystallization procedure is described in §2.2.

Single crystals with dimensions of 300–500 µm were soaked in cryoprotective solution consisting of 15%(*v*/*v*) PEG 400, 25%(*v*/*v*) glycerol and 60%(*v*/*v*) mother liquor (Smith *et al.*,

Table 3
Data-collection parameters and processing statistics.

Values in parentheses are for the outermost resolution shell.

	T_6 insulin	T_3R_3 insulin	R_6 insulin
Data collection			
Beamline	I911-2, MAX II, MAX-lab	I911-2, MAX II, MAX-lab	I911-2, MAX II, MAX-lab
Wavelength (Å)	1.04002	1.03914	1.04002
Temperature (K)	100	100	100
No. of frames	180	200	180
Oscillation range (°)	1	0.5	1
Exposure time per frame (s)	8	10	10
Detector distance (mm)	80	70	120
Resolution (Å)	1.30	1.21	1.70
Data processing			
Resolution (Å)	20.78–1.40 (1.50–1.40)	21.27–1.23 (1.30–1.23)	28.88–1.80 (2.00–1.80)
No. of reflections	86987 (16137)	70690 (7380)	365308 (96841)
No. of unique reflections	15790 (2963)	23527 (3060)	64897 (17655)
Multiplicity	5.51 (5.45)	3.00 (2.41)	5.63 (5.49)
Completeness (%)	98.1 (99.8)	93.0 (79.0)	97.5 (97.9)
$R_{\rm merge}$† (%)	4.2 (60.8)	4.6 (41.4)	7.1 (50.1)
$\langle I/\sigma(I)\rangle$	19.28 (2.62)	14.68 (3.52)	11.65 (2.95)
Space group	*R*3	*R*3	*R*3
No. of molecules per asymmetric unit	2	2	16
Unit-cell parameters			
a (Å)	80.98	79.20	156.24
c (Å)	33.49	37.22	78.88
Average mosaicity (°)	0.716	0.371	0.422
Solvent content‡ (%)	33.2	37.2	39.1

† $R_{\rm merge}$ is defined as $\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of a set of equivalent reflections. ‡ Estimated by the program *Matthews Probability Calculator* (Kantardjieff & Rupp, 2003).

2001, 2003) and stored in liquid nitrogen until diffraction analysis.

2.2. Preparation of insulin microcrystals

Microcrystal samples (crystal dimensions of less than 1 µm) of the three conformations were used for X-ray absorption spectroscopy measurements. All three conformations were prepared using the batch crystallization technique. The compositions of all three crystallization solutions are given in Table 2 as the concentration of each compound in the crystallization solution. The crystallization compounds were added in the order given in Table 2.

T_6 insulin microcrystals were obtained using a slight modification of the procedure of Schlichtkrull (1956). The crystallization solution was warmed to 323 K and the pH was adjusted to 6.0 with HCl. The solution was allowed to cool to room temperature and microcrystals were precipitated during 24 h under agitation. The microcrystals were collected by filtration using an Ultrafree-MC centrifuge filter (Amicon/Millipore) with a pore size of 0.22 µm. Excess zinc was removed by recrystallization under analogous conditions but with the exclusion of zinc acetate, resulting in a stochiometric ratio of insulin and zinc.

T_3R_3 insulin crystals were grown with thiocyanate using a slight modification of the procedure of Whittingham *et al.* (1995). After the crystallization solution had been prepared (Table 2), the pH was subsequently increased to 8.0 with 1 *M* NaOH. The solution was warmed to 323 K to ensure complete solution and maintained at approximately 323 K while the pH was lowered to 6.4 with 1 *M* HCl. The solution was placed in an incubator at 323 K and cooled to room temperature over a period of 4 d. The crystallization resulted in a few crystals with dimensions of 500 µm and several smaller rhomboid-shaped crystals. The crystals were recovered by filtration using a centrifuge filter. A large crystal was saved for single-crystal XRD, while the rest were gently crushed into a powder. The powder was washed three times with a solution consisting of 0.04 *M* sodium citrate, 12%(*v*/*v*) acetone, 0.2 *M* KSCN pH 6.0 to eliminate excess zinc in the sample.

R_6 insulin was cocrystallized with the phenol derivative *m*-cresol using a crystallization solution with an analogous composition to that used in the procedure of Smith *et al.* (2000) (see Table 2). The pH was increased to 9.4 with 1 *M* NaOH to ensure complete solution. After adjusting the final pH to 8.2, the solution was placed in an incubator at 277 K and microcrystals were observed after 2 d. The crystals were separated by filtration on a centrifuge filter and excess zinc was eliminated by washing three times with a zinc-free solution consisting of 0.035 *M* sodium citrate, 23%(*v*/*v*) acetone, 0.3% *m*-cresol, 1 *M* NaCl.

For all three samples, the correct crystal form was verified using X-ray powder diffraction (XRPD).

2.3. Single-crystal diffraction

Single-crystal diffraction data were collected on beamline I911-2, MAX-II, MAX-lab, Lund, Sweden using a MAR Research MAR165 CCD detector. Data-collection and data-processing statistics for all three crystals are summarized in Table 3. The data were processed and scaled using *XDS* and *XSCALE* (Kabsch, 2010). The processed data were tested for twinning using the online merohedral twin-detection algorithm (Padilla & Yeates, 2003). The structures were refined using *REFMAC*5 (Murshudov *et al.*, 2011) included in the *CCP*4 suite (Winn *et al.*, 2011) and *PHENIX* (Adams *et al.*, 2010). Model building and editing were carried out using *WinCoot* (Emsley & Cowtan, 2004), whereas ordered solvent was modelled using the auto-insertion procedure in *PHENIX*, which was followed by manual inspection to remove insignificant water molecules. For each data set, a random 5% subset of all reflections was reserved for validation. The structures were validated using *PROCHECK* (Laskowski *et al.*, 1993), *WHAT_CHECK* (Hooft *et al.*, 1996) and the *STAN* structure-analysis server (Kleywegt & Jones, 1996). Refinement and validation statistics are listed in Table 4.

The residue numbers are assigned according to the nomenclature used by Smith *et al.* (2000, 2001), in which the number after the point refers to the monomer number. In the deposited coordinates A, C, E, G, I, K, M, O, Q, S, U, W, Y, 1, 3 and 5 refer to A chains and B, D, F, H, J, L, N, P, R, T, V, X, Z, 2, 4 and 6 refer to B chains. For instance, PheB1.3 designates the first phenylalanine residue in the B chain of insulin monomer 3, which corresponds to PheF1 in the deposited PDB file.

2.3.1. T_6 insulin. The structure was refined using the peptide chain from the 1.0 Å resolution structure of human T_6 insulin (PDB entry 1mso; Smith *et al.*, 2003). The residues specific to bovine insulin were changed, the N-terminus of the B.2 chain (PheB1.2–GlnB4.2) was remodelled and the side chains of residues ValB12.1, LeuB17.1, CysA11.2 and LeuB17.2 were modelled in two alternate conformations. A total of 84 water molecules were inserted. Restrained refinement was carried out in *PHENIX* and H atoms were included. The two zinc ions were modelled by anisotropic refinement and the peptide chain by a combination of TLS refinement and isotropic refinement. The TLS domains were chosen in such a way that each secondary-structure element constituted a TLS domain, resulting in seven domains in total: residues 1–8 and 13–19 in the A chains, residues 9–18 in the B chains and a group containing residues 23–27 of two adjacent B chains. Other atoms were refined isotropically. Validation showed that only one residue, SerA9.1, fell into the outlier region in the Ramachandran plot using the regions defined by Kleywegt & Jones (1996). Atomic coordinates and experimental data for bovine T_6 insulin (PDB entry 4e7t) have been deposited in the PDB.

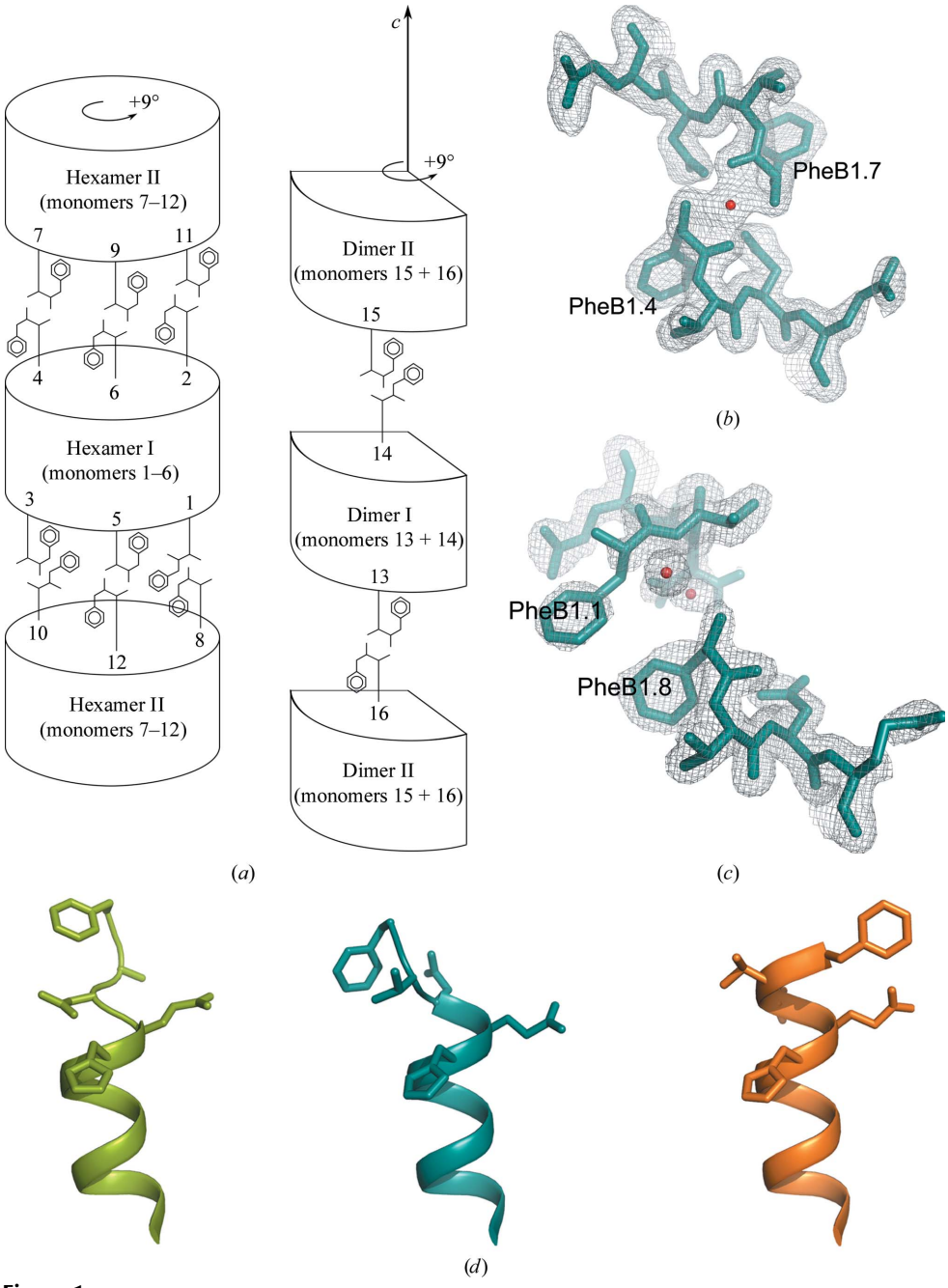

Figure 1
(*a*) Schematic representation of the overall packing of the eight dimers constituting the asymmetric unit in the R_6 bovine insulin structure; the different conformations of PheB1 are emphasized. The N-terminal residues B1–B4 are shown in (*b*) for the site where monomers 4 and 7 meet and in (*c*) for the site where monomers 1 and 8 meet. (*d*) Overview over the different R conformations: the B-chain N-terminus in the R^f conformation (green) seen in the R sites of T_3R_3 insulin [shown here for PDB entry 1ben (Smith *et al.*, 1996) as the PheB1.2 residue is disordered in the T_3R_3 structure presented here], the new conformation (blue) seen in monomers 2–9, 11–13, 15 and 16 and the pure R conformation (orange) seen in monomers 1, 10 and 14.

2.3.2. T_3R_3 insulin. The structure was refined using the peptide chain from the structure of human T_3R_3 insulin cocrystallized with 4-hydroxybenzamide (PDB entry 1ben; Smith *et al.*, 1996). Two zinc ions and one molecule of thiocyanate were included. The refinement procedure was analogous to that used for T_6 insulin. The residues PheB1.2, LysB29.2 and AlaB30.2 were disordered and could not be modelled. The final structure contained 98 water molecules and the side chains of residues GlnA5.1, LeuA16.1, GluB13.1 and SerB9.2 were modelled with

Table 4
Data-refinement and validation statistics.

	T_6 insulin	T_3R_3 insulin	R_6 insulin
Resolution cutoff (Å)	1.40	1.30	1.80
No. of atoms in the model			
Total non-H atoms	903	897	6816
Total protein atoms	817	794	6218
Total ordered water molecules	84	98	454
Total ligand atoms	0	3	136
Total Zn^{2+} ions	2	2	8
B factors† (Å^2)			
Overall	32	27	41
Main chain	26	20	36
Side chains and water molecules	34	29	45
Ligands		22	35
Zn^{2+} ions	16	11	24
R.m.s. deviation from ideal			
Bonds (Å)	0.013	0.008	0.007
Angles (°)	1.449	1.038	0.982
Ramachandran plot‡ (%)			
Residues in core regions	98.9	98.8	98.4
Outliers	1.1	1.2	1.6
R factors§			
R	0.1938	0.1439	0.2088
R_{free}	0.2285	0.1794	0.2717

† The *B*-factor analysis was performed using *BAVERAGE* included in *CCP*4 (Winn *et al.*, 2011). ‡ R and $R_{free} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated using a random 5% subset of all reflections that was excluded from refinement. § The Ramachandran plot regions were defined according to Kleywegt & Jones (1996).

alternate conformations. H atoms were included and anisotropic refinement of the atomic displacement factors was applied. Validation showed that only one residue, SerA9.1, fell into the outlier region in the Ramachandran plot. Atomic coordinates and experimental data for bovine T_3R_3 insulin (PDB entry 4e7u) have been deposited in the PDB.

2.3.3. R_6 insulin. The structure was solved by molecular replacement using an entire insulin hexamer generated from the structure of human R_6 insulin (PDB entry 1ev3; Smith *et al.*, 2000) as a template. Molecular replacement was carried out using *Phaser* (McCoy *et al.*, 2007), which found two hexamers in the asymmetric unit. According to the doubling of unit-cell parameters compared with the other conformations, 16 insulin molecules were expected to constitute the asymmetric unit. Another molecular replacement was performed on the difference map using an insulin dimer as a template. This revealed another two dimers located close to the threefold rotation axis, generating two hexamers with symmetry-related dimers. The asymmetric unit was thus found to contain a total of 16 insulin molecules distributed as two hexamers (insulin momomers 1–6 and 7–12) and two dimers (monomers 13–14 and 15–16). A schematic overview of all insulin molecules in the R_6 conformation is shown in Fig. 1(*a*).

The difference electron-density maps clearly revealed the zinc sites and showed tetrahedral coordination, as a result of which eight chloride ions were further included in the model. The residues specific to bovine insulin were changed and 16 molecules of *m*-cresol were inserted into the binding pocket located next to the CysA6 backbone O atoms. The two C-terminal residues LysB29 and AlaB30 were disordered in all monomers and were not modelled. Special attention was made to modelling the phenylalanine residues at the N-termini of all B chains, as they were found to adopt different conformations in the different chains (see Fig. 1).

Alternate conformations were assigned to residues AsnB3.1, ValB18.1, ValB18.2, GlnB4.3, ValB18.4, ValB18.6, LeuB17.7, AsnB3.10, SerA9.12, ValB18.12, AsnB3.14, LeuB17.16 and ValB18.16, with occupancies refined in *PHENIX* (the minimum occupancy was found to be 0.33). 454 water molecules were included in the structure and the atomic displacement factors were refined by a combination of TLS refinement and isotropic refinement using seven TLS domains per dimer in analogy to the T_6 structure, resulting in a total of 56 domains. All other atoms were refined isotropically. Validation showed that 11 ValB2 residues in monomers 3, 4, 5, 7, 8, 9, 11, 12, 13, 15 and 16 fell into the outlier region of the Ramachandran plot. Atomic coordinates and experimental data for bovine R_6 insulin (PDB entry 4e7v) have been deposited in the PDB.

2.4. X-ray powder diffraction

The XAFS samples were analysed by X-ray powder diffraction (XRPD) before and after the XAFS experiment to verify the conformation and to monitor the eventual degradation of the samples. Small amounts (5–7 µl) of wetted insulin microcrystal samples were mounted in a 1 mm thick sample holder for XRPD as well as XAS (Frankær *et al.*, 2011). Using the procedure described by Hartmann *et al.* (2010), the conformations were verified in 30 min on a Huber G670 diffractometer using Cu $K\alpha_1$ radiation (λ = 1.5406 Å).

2.5. X-ray absorption spectroscopy

2.5.1. Data collection. XAFS measurements were carried out at the Zn *K* edge on beamline I811 at the synchrotron at MAX-lab, Lund, Sweden (Carlson *et al.*, 2006) using an Si(111) double-crystal monochromator. The storage ring was operating with ring currents of between 150 and 300 mA and harmonics were suppressed by detuning the peak intensity by 60% at 10 660 eV.

XAS data for the insulin samples as prepared for XRPD were collected in fluorescence mode. XAS data were also collected for two model compounds representing octahedrally and tetrahedrally coordinated zinc, respectively: the six-coordinated zinc in hexakisimidazole zinc(II) chloride (Sandmark & Brändén, 1967) and the four-coordinated zinc in bisimidazole zinc(II) chloride (Edsall *et al.*, 1954). The model compounds were diluted with boron nitride in a 1:6 ratio to avoid self-absorption and were measured in transmission mode.

The samples were kept at 100 K in a liquid-nitrogen cryostat. Transmission data were collected using ion chambers and fluorescence data were collected using a PIPS PD-5000 (Passivated Implanted Planar Silicon) detector from Canberra equipped with Soller slits and a copper filter to suppress non-Zn fluorescence scattering. Data were collected in the region 9510–10 660 eV in the following intervals: pre-edge data (150–30 eV below the edge) were collected in steps of 5 eV for 1 s,

edge data (from 30 eV below to 30 eV above the edge) were collected in steps of 0.3 eV for 1 s and EXAFS data (30–1000 eV above the edge, corresponding to 16 Å^{-1} in k-space) were collected in steps of 0.05 Å^{-1} for 1–15 s. Three to four spectra were collected for each sample in order to ensure reproducibility and that no radiation damage of the sample had taken place.

2.5.2. XAS data reduction. The spectra were deglitched, averaged, energy calibrated and background corrected and the EXAFS function $\chi(k)$ was extracted using *WinXAS* (Ressler, 1998). Energy calibration was carried out using a simultaneous XAFS transmission spectrum of a Zn reference foil. The first inflection point of the foil spectrum was assigned to the Zn K edge (9659 eV). The background was modelled with a first-order polynomial and subtracted. The spectra were normalized by scaling the edge step to unity and the threshold energy E_0 was assigned to the first inflection point of the sample spectrum. The absorption of the isolated atom in the field of its neighbours without specific interactions $\mu_0(E)$ was modelled with a spline function. The spectra were converted to k-space and the EXAFS function $\chi(k)$ was extracted.

2.5.3. EXAFS analysis. Simulations of the calibrated, averaged and background-corrected EXAFS, $\chi(k)$, were carried out using *EXCURVE* (Gurman *et al.*, 1984, 1986; Binsted *et al.*, 1991). Potentials were calculated using the relaxed approximation and phase shift by the Hedin–Lundqvist method.

The starting models, consisting of atoms within a radius of 5.6 Å from the Zn atoms (which thus include the C^{β} atoms of the histidine residues), were taken from the crystal structures described in §2.3.

Multiple scattering, which occurs in the imidazole and in particular in the linear thiocyanate molecule, was included throughout all refinements. The amplitude reduction factor, S_0^2, was set to 0.95 owing to a relatively low self-absorption caused by the low zinc concentrations.

The models were first refined by a constrained refinement and subsequently by a restrained refinement (Binsted *et al.*, 1992). In the constrained refinement E_0, distances and Debye–Waller factors were refined iteratively, in which each histidine was assumed to be rigid and refined by only one distance and one angular parameter (Binsted *et al.*, 1992). The number of refinable parameters was kept low by grouping Debye–Waller factors of atoms with similar distances to the central atom (*e.g.* the $C^{\varepsilon 1}$ and $C^{\delta 2}$ atoms of the histidine residue).

During refinement, the ordered water molecules included in the outer coordination shells in the crystal structures were included stepwise in the model and justified only if the fit index, ε_v^2, decreased by more than 5% (Joyner *et al.*, 1987). The fit index, taking account of the degree of overdeterminacy in the system, is given as

$$\varepsilon_v^2 = \frac{1}{N_{\rm ind} - p}\frac{N_{\rm ind}}{N}\sum_i^N \frac{1}{\sigma_i^2}[\chi^{\rm exp}(k_i) - \chi^{\rm teo}(k_i)]^2, \qquad (1)$$

where N is the number of data points, $N_{\rm ind}$ is the number of independent data points and p is the number of parameters (Binsted *et al.*, 1992). The fit was evaluated by a residual, $R_{\rm exafs}$, defined as (Binsted *et al.*, 1992)

$$R_{\rm exafs} = \sum_i^N \frac{1}{\sigma_i}[|\chi^{\rm exp}(k_i) - \chi^{\rm teo}(k_i)]) \times 100\%. \qquad (2)$$

Restrained refinement was then performed in which restraints were set up for the histidine residue by restraining the imidazole and the C^{β} atom as defined by Engh & Huber (1991). The geometric restraints and the EXAFS were weighted by 50% each.

2.5.4. XANES analysis. For XANES analysis, the calibrated, averaged and background-corrected experimental spectra, given as a function of energy, $\mu(E)$, are compared with other experimental XANES spectra (qualitative analysis) and calculated XANES spectra (quantitative analysis). Calculation of XANES spectra was performed using the *FDMNES* code (Joly, 2001), which performs full potential calculations using finite-difference methods (FDM; Kimball & Shortley, 1934). The FDM is currently the most precise approach to calculating XANES spectra (Jacquamet *et al.*, 2009), as the method offers an approach to solving the Schrödinger equation in a free shape potential when working in the framework of the local density approximation and therefore goes further than the muffin-tin approximation.

Quantitative fitting of XANES spectra was possible for T_6 and R_6 insulin. The coordinates optimized from the EXAFS analysis for both structures were used as starting models. Full multiple scattering calculations were performed at the Zn K edge using the *FDMNES* code (Joly, 2001), testing both the muffin-tin (MT) approximation and the FDM approach. A cluster with a radius of 4.5 Å from the Zn ion was found to be suitable to reproduce the features of the XANES spectra while still minimizing the calculation time. The calculated spectra were convoluted with an energy-dependent broadening using an arctangent function to account for inelastic processes and with a Gaussian function to simulate the experimental energy resolution of the instrument.

For the fitting procedure, geometrical transformations were decomposed into parameters and selected using the criterion that each parameter must have a significant influence on the XANES spectra. Four structural parameters (two ligand–Zn distances and two angles) have been chosen for the T_6 and R_6 insulin conformers. All geometry transformations and FDM calculations were carried out and the structural parameters were optimized by multidimensional interpolation of calculated XANES spectra by *FitIt* (Smolentsev & Soldatov, 2007). The best fit was found by finding a suitable interpolation polynomial and a XANES spectrum of the optimized geometry was calculated for verification. The fit between the calculated and experimental spectra is evaluated with an R factor given as

$$R_{\rm xanes} = \sum_i^N \frac{[|\mu^{\rm exp}(E_i) - \mu^{\rm cal}(E_i)|]}{\mu^{\rm exp}(E_i)} \times 100\%. \qquad (3)$$

2.6. Comparative analysis

An analysis of Zn coordination in other insulin structures was carried out by searching the PDB. 67 X-ray structures of hexameric insulin (including mutant proteins) containing Zn were found. Structures solved at noncryogenic temperatures were discarded and only structures with a resolution of 2.00 Å or better were included in the analysis since the geometrical restraints are tighter for low-resolution structures. This resulted in a total of 31 structures. The Zn sites in these structures were classified as either a T site or an R site and the bond distances from Zn to its inner coordination sphere ligands were extracted.

Zn coordination in small molecules was analysed by searching the Cambridge Structural Database (CSD) for structures solved at temperatures below 200 K containing zinc which coordinates to three acyclic ligands through sp^2-hybridized N atoms. For the T-site and the R-site, respectively, only structures containing Zn coordinated to six atoms (three nitrogen ligands and three other atoms) and Zn coordinated to four atoms (three nitrogen ligands and one chloride) were included in the analysis. Among the structures containing hexacoordinated zinc, only Zn—O distances were included in the analysis besides the three Zn—N distances.

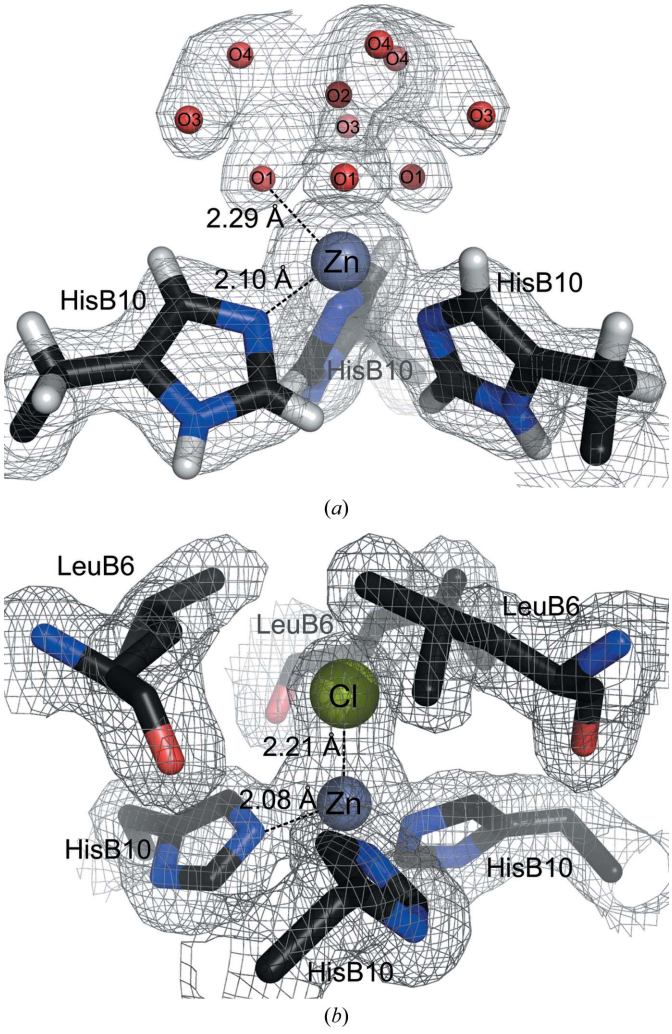

Figure 2
(*a*) Octahedrally coordinated Zn in the 1.4 Å resolution T_6 insulin structure, in which the electron density has been modelled with four symmetry-related water molecules, of which one is located on the threefold axis (O2). (*b*) Tetrahedrally coordinated Zn in the 1.8 Å resolution R_6 insulin structure, in which the LeuB6 residues encase the metal site. Distances to the first coordination sphere as determined by single-crystal X-ray diffraction are shown. As the T_6 insulin and R_6 insulin contain two and eight Zn ions per asymmetric unit, respectively, the distances shown are the average of the different Zn sites in each structure. The σ_A-weighted $2F_o - F_c$ maps are contoured at 1.0σ.

3. Results

3.1. Single-crystal X-ray diffraction structures

3.1.1. General conformation. The conformations of the B-chain N-termini, which determine the overall insulin conformation, have been verified in all three structures: both insulin molecules in the T_6 structure adopt the tensed conformation in which the backbone of residues B1–B8 is extended.

In the T_3R_3 structure the insulin molecule present in the T_3 trimer also adopts an extended conformation, whereas for the molecule involved in the R_3 trimer an R^f conformation[1] was observed analogous to that observed for human and porcine T_3R_3 insulin (Smith *et al.*, 2001; Whittingham *et al.*, 1995).

In the R_6 insulin structure, the pure R conformation, in which all residues in the range B1–B18 take part in an α-helix, was only observed in monomers 1, 10 and 14 (see Figs. 1*a*–1*c*). The three first residues B1–B3 in the remaining 13 insulin molecules were not observed to adopt either a pure R or an R^f conformation, but an intermediate conformation in which the PheB1 residue is extended (Fig. 1*d*). This conformation will be designated R^g in the following. As a consequence, the φ and ψ torsion angles of ValB2 take values around −89° and 50°, respectively, which are found to be located on the border of the allowed regions in the Ramachandran plot, thereby justifying the outlying ValB2 residues.

The PheB1 residues are located on the protein surface and are thus involved in the hexamer packing. The backbone orientation of the two hexamers and the two dimers, respectively, were observed to differ by a rotation of 9° about an axis parallel to the *c* axis (see Fig. 1*a*).

3.1.2. Geometry of the Zn sites. The coordination of the zinc sites are presented in Figs. 2 and 3. For the T_6 conformation, octahedral coordination geometry was verified at both Zn sites, fulfilled by three symmetry-related water molecules at an average distance of 2.29 Å from the Zn atom. As the Zn sites in T_6 insulin are exposed to the solvent, a large electron density was observed above the Zn ion (Fig. 2*a*) which was best modelled by water molecules connected by a hydrogen-bonding network.

The Zn sites in R_6 insulin (Fig. 2*b*) are not exposed to the solvent. They are encased by the helix formation of the N-terminus of the B chains and thus there is only room for a tetrahedral coordination with a chloride ion completing the

[1] The conformation in which residues B1–B3 are extended has been reported as the frayed R conformation, denoted R^f (Ciszak *et al.*, 1995).

coordination. The Zn—Cl distances ranged from 2.07 to 2.40 Å throughout the eight Zn sites, with an average of 2.21 ± 0.10 Å.

In the T_3R_3 structure both Zn-site geometries were observed (Fig. 3). One thiocyanate ion fulfils the tetrahedral Zn-site coordination (Fig. 3*a*) as observed previously in the porcine T_3R_3 structure (Whittingham *et al.*, 1995). A clear octahedral Zn coordination was observed at the other Zn site and the electron density was modelled by water molecules at a distance of 2.47 Å from the Zn atom (Fig. 3*b*).

3.2. Powder diffraction

Using the procedures described by Hartmann *et al.* (2010), the crystal forms were verified from the powder patterns. The unit-cell parameters were determined from a full pattern-profile fit and are different for the different conformations, *e.g.* the unit-cell parameter *c* is increased by approximately 3 Å for each trimer undergoing a T to R conformational change. This systematic variation in unit-cell parameters is therefore the primary factor in the identification of the conformations.

In order to detect whether the R_6 insulin undergoes a phase transition when cryocooled, powder patterns with longer data-collection times were collected at room temperature and at 120 K. The unit-cell doubling observed in the single-crystal experiment was detected from the XRPD patterns collected at 120 K by looking at the low-angle region (2θ = 3–5°), in which the peaks are better resolved (see Supplementary Material[2]).

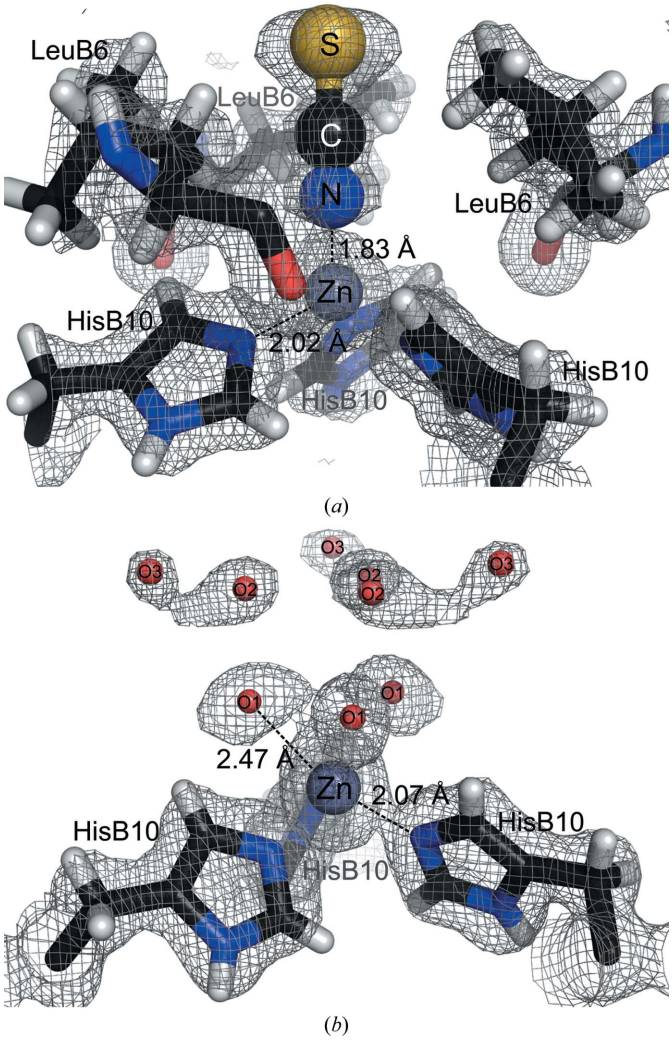

Figure 3
Zn coordination in the 1.3 Å resolution T_3R_3 insulin structure. (*a*) A thiocyanate fulfils the tetrahedral coordination and the site is further encased by LeuB6 residues in analogy to the Zn sites in R_6 insulin. (*b*) Octahedrally coordinated Zn in which the electron density has been modelled by three symmetry-related water molecules. Distances to the first coordination sphere as determined by single-crystal X-ray diffraction are shown. The σ_A-weighted $2F_o - F_c$ maps are contoured at 1.0σ.

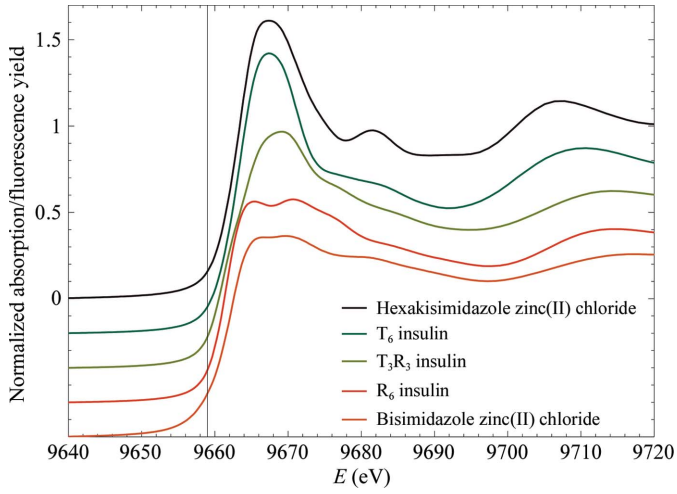

Figure 4
XANES spectra of the three protein samples and two reference compounds: hexakisimidazole zinc(II) chloride, T_6 insulin, T_3R_3 insulin, R_6 insulin and bisimidazole zinc(II) chloride. The *K* edge for metallic zinc is indicated at 9659 eV.

3.3. Qualitative XANES spectroscopy

In Fig. 4, XANES spectra of the three insulin conformations T_6, R_6 and T_3R_3, which contain six-coordinated zinc, four-coordinated zinc and a mixture of the two, respectively, are compared with those of two model compounds. The octahedrally coordinated zinc complexes with six ligands give strong 'white lines' (>1.5) with normalized edge steps, whereas tetrahedrally coordinated zinc complexes give 'white lines' (<1.5) with normalized edge steps (Feiters *et al.*, 2003). The medium intensity of the 'white line' for the T_3R_3 conformation indicates coordination of both octahedral and tetrahedral character, which is in agreement with the dual coordination observed in the crystal structure.

3.4. EXAFS spectroscopy

The extracted k^3-weighted EXAFS spectra and the modulus of the phase-corrected Fourier transforms of the three insulin conformations are presented in Fig. 5. The major single peak in the radial distribution function is observed around 2.0 Å and corresponds to the inner coordination shell, which is

[2] Supplementary material has been deposited in the IUCr electronic archive (Reference: TZ5009). Services for accessing this material are described at the back of the journal.

Table 5
Zn-coordination distances as refined from EXAFS of T_6, R_6 and T_3R_3 insulin compared with crystallographic values.

T_6 and R_6 insulin were refined with restrained refinement, while T_3R_3 insulin was refined using constrained refinement.

(*a*) T_3 sites.

	T_6 insulin			T_3R_3 insulin		
	XRD	Restrained EXAFS		XRD	Constrained EXAFS	
T	'*R*'† (Å)	*R* (Å)	$2\sigma^2$ (Å^2)	'*R*'† (Å)	*R* (Å)	$2\sigma^2$ (Å^2)
$N^{\varepsilon 2}$ (HisB10)	2.10	2.074 (3)	0.012 (1)	2.07	2.03 (1)‡	0.014 (1)
$C^{\varepsilon 1}$	3.03	3.07 (4)§	0.020 (3)	3.05	2.84	0.018 (1)
$C^{\delta 2}$	3.13	3.05 (3)§	0.020 (3)	3.05	3.16	0.018 (1)
$N^{\delta 1}$	4.17	4.22 (2)§	0.017 (3)	4.16	4.03	0.028 (4)
C^{γ}	4.26	4.22 (3)§	0.017 (3)	4.20	4.21	0.028 (4)
C^{β}	5.69	5.55 (5)§	0.017 (3)	5.61	5.66	0.030 (3)
Ow1	2.29	2.14 (1)	0.030 (1)	2.47	2.29 (2)	0.035 (5)
Ow2 (axial)	3.09	2.88 (3)	0.020 (1)			

(*b*) R_3 sites.

	R_6 insulin			T_3R_3 insulin		
	XRD	Restrained EXAFS		XRD	Constrained EXAFS	
	'*R*'† (Å)	*R* (Å)	$2\sigma^2$ (Å^2)	'*R*' (Å)	*R* (Å)	$2\sigma^2$ (Å^2)
$N^{\varepsilon 2}$ (HisB10)	2.08	2.001 (4)	0.007 (1)	2.02	1.99 (1)‡	0.014 (1)
$C^{\varepsilon 1}$	3.09	2.98 (2)§	0.010 (2)	2.99	2.99	0.018 (1)
$C^{\delta 2}$	3.04	3.04 (2)§	0.010 (2)	3.03	2.96	0.018 (1)
$N^{\delta 1}$	4.18	4.15 (1)§	0.012 (2)	4.10	4.09	0.028 (4)
C^{γ}	4.19	4.14 (2)§	0.012 (2)	4.16	4.11	0.028 (4)
C^{β}	5.60	5.53 (3)§	0.012 (2)	5.58	5.52	0.030 (3)
Cl	2.21	2.218 (3)	0.006 (1)			
N^{SCN}				1.83	1.802 (9)	0.014 (1)
C^{SCN}				2.98	2.96	0.018 (1)
S^{SCN}				4.72	4.69	0.017 (3)

† Average distances of all Zn sites in the structure. The standard deviations among the different sites are below 0.03 Å for T_6 insulin and below 0.1 Å for R_6 insulin. ‡ The rotation angle of the histidine unit around an axis orthogonal to the imidazole plane passing through the $N^{\varepsilon 2}$ atom was included in the refinement. § The φ angle (polar coordinates) was refined in order to allow some movement of the atoms in the restrained imidazole ring.

mostly dominated by the three imidazole N atoms and is common to all three conformations. The shells occurring around 3 and 4 Å verify the imidazole coordination of histidine in all conformations. The imidazole coordination can also be directly identified from the 'camel-back' feature around 4–5 Å^{-1}, which originates from the interference between the backscattered electron waves from the first-shell, second-shell and third-shell atoms of the imidazole ring (Bordas *et al.*, 1983).

By comparing the k^3-weighted $\chi(k)$ at low k values (approximately 3–7 Å^{-1}) the T_3R_3 conformation has both T_6 and R_6 character, which is in agreement with the dual tetrahedral/octahedral coordination of the zinc ions as observed from the crystal structure and the XANES spectra.

An indication of coordination of thiocyanate in the T_3R_3 conformation is seen as a slightly more intense and complex peak around 4.5–5.0 Å in the radial distribution function, whereas the chloride which fulfils the tetrahedral coordination in the R_6 conformation is seen as a more intense and slightly more broadened inner coordination shell peak at 2.0–2.2 Å.

Table 6
EXAFS refinement statistics.

T_6 and R_6 were refined with restrained refinement, while T_3R_3 was refined using constrained refinement.

	T_6	R_6	T_3R_3
E_f (eV)	−3.31	−5.47	−2.366
ε_ν^2	0.4127	0.1920	3.1572
R_{exafs} (%)	14.09	10.17	22.53
R_{dist} (%)	1.14	0.65	
R_{total} (%)	15.23	10.82	
N_p	19	19	14
k range (Å^{-1})	2.8–14.3	2.8–13.3	2.8–14.3
w_{dist}	0.5	0.5	

Table 7
Quantitative XANES fit parameters and *R* factors for the initial model (EXAFS model) and the model optimized by *FitIt* (XANES model).

	Input model (EXAFS model)	Optimized model (XANES model)
T_6 insulin		
Zn—$N^{\varepsilon 2}$(His) (Å)	2.07	2.08
Zn—Ow1 (Å)	2.14	2.13
∠(Zn—$N^{\varepsilon 2}$—$C^{\varepsilon 1}$) (°)	127	119
∠($N^{\varepsilon 2}$—Zn—Ow1) (°)	166	173
R_{xanes} (%)	3.7	2.5
R_6 insulin		
Zn—$N^{\varepsilon 2}$(His) (Å)	2.00	2.00
Zn—Cl (Å)	2.22	2.24
∠(Zn—$N^{\varepsilon 2}$—$C^{\varepsilon 1}$) (°)	130	141
∠($N^{\varepsilon 2}$—Zn—Cl) (°)	109	107
R_{xanes} (%)	3.4	2.4

The optimized distances and Debye–Waller factors are presented in Table 5 and the best fits are shown in Fig. 5 (dashed lines). Owing to the complexity of the double cluster necessary for modelling the T_3R_3 conformation, restrained refinement resulted in too many parameters and was therefore abandoned. Hence, the T_3R_3 zinc sites were refined by a constrained refinement only. Restrained refinement of the T_6 and R_6 conformations gave the best fits, in which residuals as low as 11–15% were reached. Further EXAFS refinement statistics are summarized in Table 6.

The distances to the inner coordination sphere as determined by EXAFS spectroscopy were used as restraints for refinement of the T_6 and R_6 insulin XRD structures. The EXAFS distances were given as the ideal distance and the uncertainty as the standard deviation σ in *PHENIX*. Slightly increased R_{free} values were observed for both T_6 and R_6 insulin when applying the restraints from the EXAFS model.

3.5. Quantitative fitting of XANES spectra

XANES spectra calculated using the FDM approach are shown in Fig. 6 for the T_6 and R_6 insulin conformations before and after quantitative fitting of the spectra. The parameters which were varied during the quantitative fit are listed in Table 7 together with R_{xanes} for the fits.

XANES spectra were also calculated using the MT approach, but matched the observed spectra poorly. The spectra calculated by the FDM approach were in better

agreement with the experimental spectra (see Supplementary Material).

Among the techniques used, the XANES technique is more sensitive to bond angles. Regarding the inner coordination sphere of Zn, the XANES analysis of T_6 insulin shows a more regular octahedrally coordinated Zn than was determined by XRD, as the $N^{\varepsilon 2}$—Zn—Ow1 angle was optimized from 166° to 173°. The tetrahedral Zn site in R_6 insulin was verified without significant optimization, meaning that the initial R_6 model from the EXAFS-optimized XRD structure was better than the T_6 model. However, the in-plane rotation of the histidines was optimized from 130° to 141° with a significant influence on the shape and the intensity ratio of the main double peak of the spectrum (see Fig. 6*b*).

The optimized models from XANES were used for another round of EXAFS analysis, which gave fits of similar quality to the initial models.

3.6. Comparison with other reported Zn-site geometries

The result of the analysis of bond distances involved in the zinc coordination observed in other insulin structures, as well as in small molecules with similar coordination geometries, is summarized in Table 8.

Table 8
Average bond distances and statistical standard deviations for Zn bonds to inner sphere ligands among 31 insulin structures deposited in the PDB and 39 small-molecule structures with similar Zn coordination.

(*a*) PDB analysis.

	T sites (25 Zn sites)		R sites (42 Zn sites)	
	Zn–$N^{\varepsilon 2}$(His) (Å)	Zn–Ow1 (Å)	Zn–$N^{\varepsilon 2}$(His) (Å)	Zn–Cl (Å)
All structures	2.06 (5)	2.40 (2)	2.00 (1)	2.18 (9)
XRD (this work)†	2.10 (1)	2.29 (3)	2.08 (5)	2.20 (1)
EXAFS (this work)†	2.07	2.14	2.00	2.22

(*b*) CSD analysis.

Octahedral Zn (3N + 3X) (26 structures)		Tetrahedral Zn (3N + 1Cl) (13 structures)	
Zn—N(sp^2)	Zn—O	Zn—N(sp^2)	Zn—Cl
2.16 (3)	2.16 (4)	2.01 (2)	2.26 (3)

† Only distances for T_6 and R_6 insulin are presented.

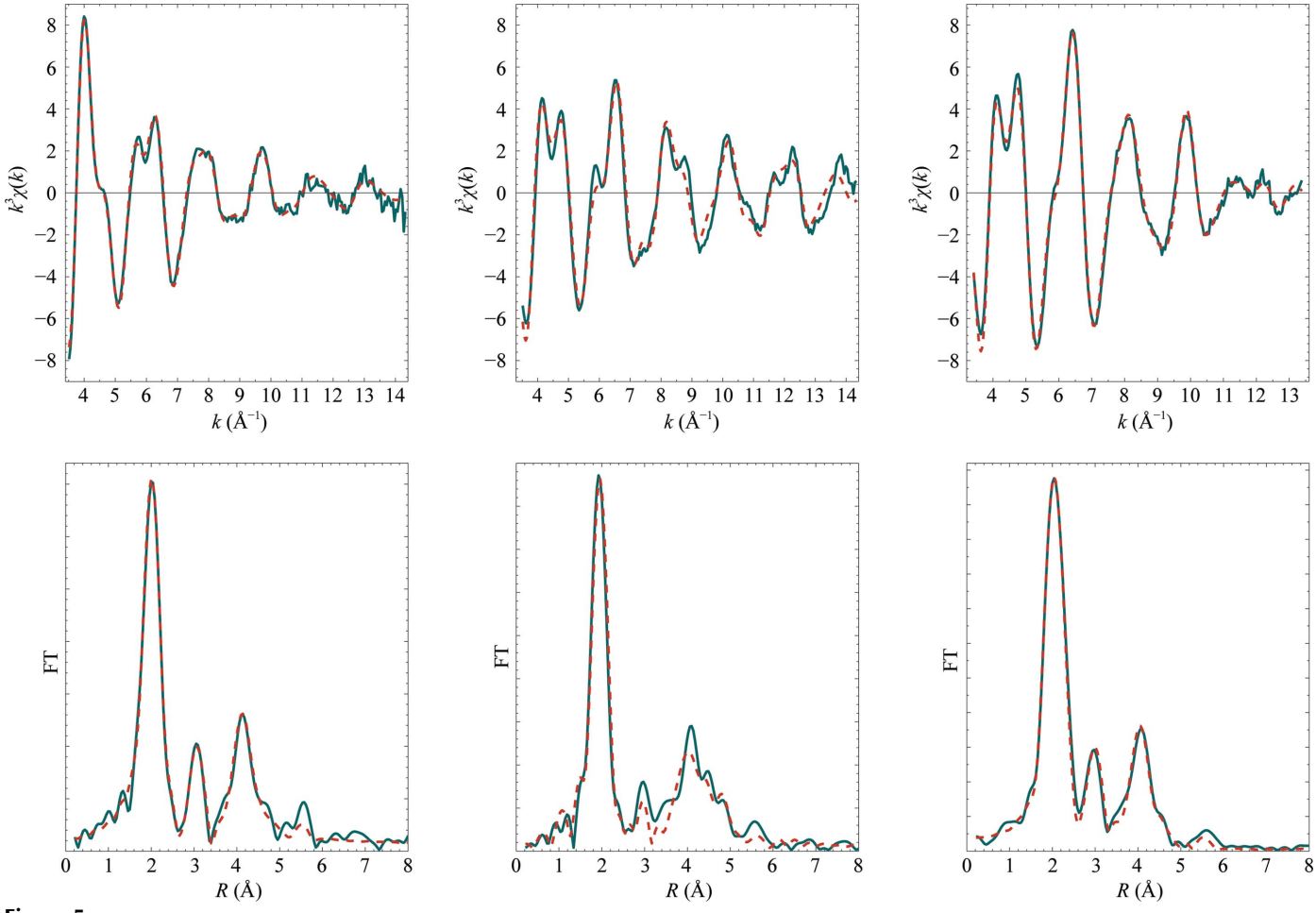

Figure 5
k^3-weighted EXAFS spectra (top) and radial distribution functions calculated as the modulus of the phase-corrected Fourier transform (bottom) of the three insulin conformations: T_6 (left), T_3R_3 (centre) and R_6 (right). Experimental spectra are shown in blue and simulated spectra in red (dashed lines), using the parameters from the restrained refinement given in Table 5 for T_6 and R_6 and the parameters from the constrained refinement given in Table 5 for T_3R_3.

The database search for small-molecule structures with a geometry that resembles the T_6 geometry revealed only a small number of structures with Zn coordinated to three nitrogen and three oxygen ligands with facial isomerism. Therefore, the three ligands besides nitrogen were unspecified in the search. For the T-site Zn—N distances there is a good agreement between the mean XRD and our EXAFS results. The small-molecule search shows a somewhat longer distance, maybe as a consequence of the difficulty in finding matching structures. For the Zn—O distances, however, EXAFS spectroscopy clearly gives results that are much closer to the small-molecule structures, while the distances from XRD are too large and have large standard deviations. A brief look at the difference electron-density maps for insulin structures containing T sites from the PDB (including our own structure) often reveals unstructured densities modelled with water molecules and at chemically unreasonable bond lengths, in particular in structures with lower resolution.

For the R sites, the Zn—N distance found by our EXAFS study is closer to the Zn—N distance found by averaging all of the PDB structures than to the structure determined in this paper. The broad spread amongst the results from XRD can be seen in Fig. 7, in which the bond distances found are plotted as a function of resolution. It is clear that the high-resolution structures show good agreement with results from the small-molecule survey and with our EXAFS result.

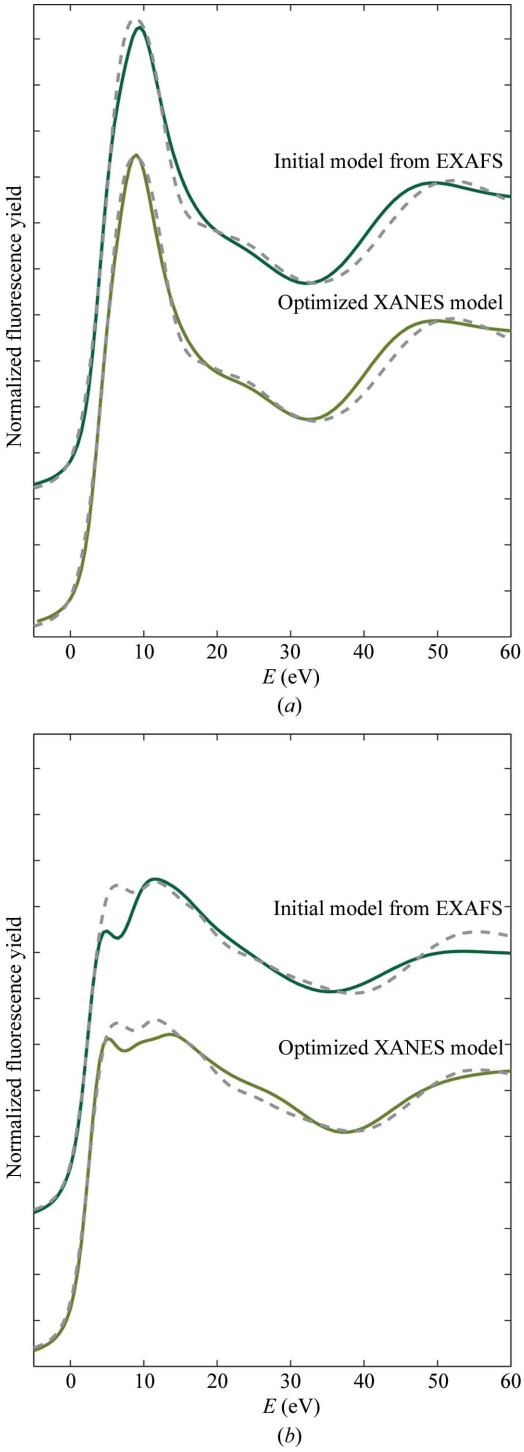

Figure 6
XANES calculated on a 4.5 Å cluster around each Zn atom using the FDM approach and compared with experimental XANES (dashed curves). Calculated spectra are shown for the initial model (EXAFS model) and the model optimized by *FitIt* (XANES model) of (*a*) T_6 insulin and (*b*) R_6 insulin. The offset of the energy scale is 9659 eV, corresponding to the *K*-edge position of metallic zinc.

4. Discussion

4.1. Single-crystal XRD structures

The three conformations of bovine insulin were crystallized in space group *H*3 and data were collected to 1.40, 1.30 and 1.80 Å resolution for T_6, T_3R_3 and R_6 insulin, respectively. The structures of bovine insulin in the T_3R_3 and R_6 conformations have not been solved previously. In analogy to what was observed for a previously reported bovine T_6 insulin structure (Smith *et al.*, 2005), the structural consequences of the

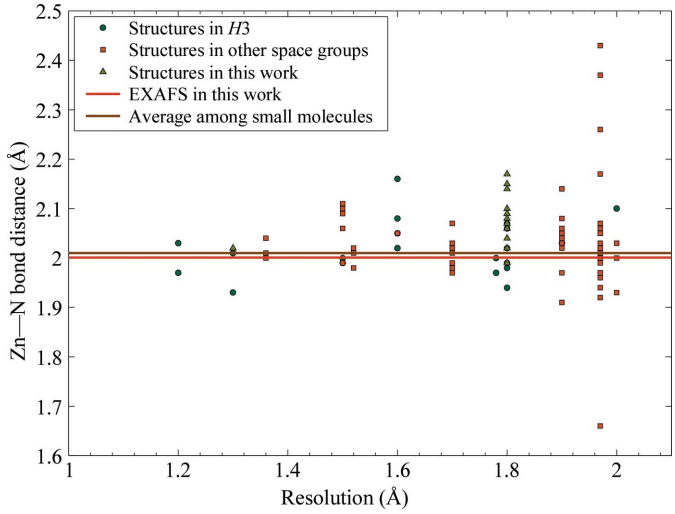

Figure 7
Zn—N distances as a function of resolution among 42 R sites deposited in the PDB. Insulin structures in space group *H*3 are distinguished from those in other space groups. Furthermore, the Zn—N distances from the XRD structures presented in this work, the Zn—N distance in R_6 insulin as determined by EXAFS spectroscopy and the average Zn—N distance among tetrahedral Zn sites in small molecules are included.

sequence changes at A8 and A10 were not significant in any of the three conformations. However, there are some differences from previously reported structures.

The structures presented in this paper were superposed with other structures deposited in the PDB. Independent T_2, TR and R_2 dimers in T_6, T_3R_3 and R_6 insulin, respectively, were superposed using a least-squares procedure using *SUPERPOSE* (Krissinel & Henrick, 2004), in which the displacements of C^α atoms of all common residues were minimized and the root-mean-square deviations (r.m.s.d.s) were calculated.

The T_6 structure in this paper shows a high resemblance to the previously published bovine structure (r.m.s.d. = 0.282 Å; PDB entry 2a3g; Smith *et al.*, 2005) and the porcine structure (r.m.s.d. = 0.239 Å; PDB entry 4ins; Baker *et al.*, 1988). A somewhat larger discrepancy from the human structure (r.m.s.d. = 1.013 Å; PDB entry 1mso; Smith *et al.*, 2003) may be explained by a difference in the conformation of the B1.2–B4.2 chain.

The T_3R_3 insulin structure closely resembles the porcine structure (r.m.s.d. = 0.377 Å; PDB entry 2tci; Whittingham *et al.*, 1995), which was also crystallized with thiocyanate.

The bovine R_6 structure displays an eightfold doubling of the unit cell compared with the human R_6 structure (PDB entry 1ev3; Smith *et al.*, 2000). This means that the asymmetric unit contains 16 monomers, in the majority of which (13 of 16 monomers) the B-chain N-termini are found to adopt the R^g conformation. Excluding PheB1 and ValB2 from the superposition resulted in an r.m.s. displacement of C^α atoms of 0.413 Å. The R^g conformation has previously been observed in one of the six molecules in a hexamer in monoclinic human R_6 insulin structures (Smith *et al.*, 2000).

Together with the observed rotation of adjacent hexamers/dimers along the *c* direction by 9°, the conformational variation of the B-chain N-termini explains the eightfold doubling of the unit cell. Using powder diffraction, we could see that when cooled the R_6 conformation undergoes a phase transition analogous to that observed in the human T_3R_3 structure by Smith *et al.* (2001), with the large unit cell being the low-temperature structure.

Both Zn sites in the T_6 insulin structure are octahedral and coordinate three water molecules at similar (but long) distances to those observed in the other T_6 structures (Baker *et al.*, 1988; Smith *et al.*, 2003, 2005). For the T_3R_3 Zn sites the three water molecules which fulfil the octahedral coordination are clearly observed in the T_3 site, in contrast to the porcine (Whittingham *et al.*, 1995) and human (Smith *et al.*, 2001) structures, in which tetrahedral coordination fulfilled by a single water molecule and a chloride ion, respectively, was modelled. However, the R_3 sites in both the T_3R_3 and R_6 conformations are always tetrahedral as one lyotropic anion fulfils the coordination geometry.

4.2. XAS

Using EXAFS spectroscopy the inner coordination sphere distances are more precisely determined (standard deviations within 0.01 Å), whereas the distances to the outer shells are determined with standard deviations of around 0.01–0.05 Å. Not surprisingly, our results demonstrate that the technique is better for mononuclear metalloproteins or metalloproteins containing several metal clusters that all have identical coordination geometry. If the samples have different coordination geometries of the same metal type, as in the T_3R_3 conformation, the complexity of the data treatment increases and reduces the amount of structural information that can be resolved.

It was clearly demonstrated that FDM methods reproduce the XANES spectra quite well. The coordination geometry of the zinc sites (in particular the bond angles) was optimized from XANES. However, the very good quality of the EXAFS spectra (high signal-to-noise ratio up to 13–14 $Å^{-1}$) allowed us to obtain a quite accurate initial structural model. Therefore, only subtle adjustments of the coordination geometries were possible with quantitative XANES fitting. The optimized bond distances were in agreement with EXAFS, and thus XANES serves as a tool for cross-validation. Furthermore, qualitative comparison of XANES spectra is a powerful tool for the identification and verification of coordination geometry.

4.3. Comparison with other reported Zn-site geometries

The quite large discrepancy between the XRD structures and the EXAFS results may originate from the resolution differences between the two techniques. Another explanation for this discrepancy could be the difference in the X-ray dose imposed on the samples. The doses for the two experiments were estimated using the photon energy, $E_{\rm photon}$, flux, Φ, and area of the beam, $A_{\rm beam}$, at beamlines I911-2 and I811 as well as the exposed mass of the sample, $m_{\rm sample}$, and the exposure time, $t_{\rm exp}$,

$$\text{dose} = \frac{E_{\rm photon}\Phi t_{\rm exp}}{A_{\rm beam} m_{\rm sample}}. \quad (4)$$

The estimation shows a higher dose by more than two orders of magnitude (2.6×10^9 Gy) for the single crystal used for XRD compared with the XAS sample (2.5×10^7 Gy); therefore, it must be assumed that the single crystals suffer a higher degree of radiation damage. It is also reasonable to assume that radiation damage influences the loosely bound water molecules in the T site more severely than the tightly bound chloride ion in the R site.

5. Conclusion

This study of hexameric bovine insulin illustrates the complementarity between XRD and XAS for studying metals in proteins.

The structures of all three conformations (T_6, T_3R_3 and R_6) of bovine insulin were solved by single-crystal XRD; two of the structures are new (T_3R_3 and R_6).

The distances between the Zn atom and its ligands were very precisely determined using EXAFS spectroscopy (standard deviations within 0.01 Å), in particular those in the first coordination sphere. The distances from EXAFS are in

agreement with those observed in small-molecule crystal structures with similar coordination geometries, as well as in high-resolution insulin structures. These observations may be coupled to a lower radiation dose for XAS experiments. Compared with XRD, this makes XAS better suited for studying metal sites in proteins, which are sensitive to radiation damage. Using XANES the coordination geometry was verified by qualitative comparison with reference compounds. XANES spectra of the T_6 and R_6 conformations were calculated using finite-difference methods and quantitative fits, from which the distances and angles could be extracted, were successfully obtained.

Portions of this research were carried out on beamlines I811 and I911, MAX-lab synchrotron-radiation source, Lund University, Sweden. Funding for the beamline I811 project was kindly provided by The Swedish Research Council and The Knut och Alice Wallenbergs Stiftelse.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.

Arcovito, A., Benfatto, M., Cianci, M., Hasnain, S. S., Nienhaus, K., Nienhaus, G. U., Savino, C., Strange, R. W., Vallone, B. & Della Longa, S. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 6211–6216.

Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N. & Vijayan, N. M. (1988). *Philos. Trans. R. Soc. Lond. Ser. B*, **319**, 369–374.

Benfatto, M., Congiu-Castellano, A., Daniele, A. & Della Longa, S. (2001). *J. Synchrotron Rad.* **8**, 267–269.

Binsted, N., Campbell, J. W., Gurman, S. J. & Stephenson, P. C. (1991). *EXCURV*92. SERC, Daresbury Laboratory, Warrington.

Binsted, N., Strange, R. W. & Hasnain, S. S. (1992). *Biochemistry*, **31**, 12117–12125.

Bordas, J., Dodson, G. G., Grewe, H., Koch, M. H. J., Krebs, B. & Randall, J. (1983). *Proc. R. Soc. Lond. B*, **219**, 21–39.

Carlson, S., Clausén, M., Gridneva, L., Sommarin, B. & Svensson, C. (2006). *J. Synchrotron Rad.* **13**, 359–364.

Ciszak, E., Beals, J. M., Frank, B. H., Baker, J. C., Carter, N. D. & Smith, G. D. (1995). *Structure*, **3**, 615–622.

Coffman, F. D. & Dunn, M. F. (1988). *Biochemistry*, **27**, 6179–6187.

Edsall, J. T., Felsenfeld, G., Goodman, D. S. & Gurd, F. R. N. (1954). *J. Am. Chem. Soc.* **76**, 3054–3061.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Feiters, M. C., Eijkelenboom, A. P. A. M., Nolting, H.-F., Krebs, B., van den Ent, F. M. I., Plasterk, R. H. A., Kaptein, R. & Boelens, R. (2003). *J. Synchrotron Rad.* **10**, 86–95.

Frankær, C. G., Harris, P. & Ståhl, K. (2011). *J. Appl. Cryst.* **44**, 1288–1290.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (1995). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.

Gurman, S. J., Binsted, N. & Ross, I. (1984). *J. Phys. C*, **17**, 143–151.

Gurman, S. J., Binsted, N. & Ross, I. (1986). *J. Phys. C*, **19**, 1845–1861.

Harding, M. M., Nowicki, M. W. & Walkinshaw, M. D. (2010). *Crystallogr. Rev.* **16**, 247–302.

Hartmann, C. G., Nielsen, O. F., Ståhl, K. & Harris, P. (2010). *J. Appl. Cryst.* **43**, 876–882.

Hasnain, S. S. & Hodgson, K. O. (1999). *J. Synchrotron Rad.* **6**, 852–864.

Hasnain, S. S. & Strange, R. W. (2003). *J. Synchrotron Rad.* **10**, 9–15.

Holton, J. M. (2009). *J. Synchrotron Rad.* **16**, 133–142.

Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.

Jacquamet, L., Traoré, D. A., Ferrer, J.-L., Proux, O., Testemale, D., Hazemann, J. L., Nazarenko, E., El Ghazouani, A., Caux-Thang, C., Duarte, V. & Latour, J. M. (2009). *Mol. Microbiol.* **73**, 20–31.

Joly, Y. (2001). *Phys. Rev. B*, **63**, 125120.

Joyner, R. W., Martin, K. J. & Meehan, P. (1987). *J. Phys. C*, **20**, 4005–4012.

Kaarsholm, N. C., Ko, H.-C. & Dunn, M. F. (1989). *Biochemistry*, **28**, 4427–4435.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.

Kimball, G. E. & Shortley, G. H. (1934). *Phys. Rev.* **45**, 815–820.

Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **4**, 1395–1400.

Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.

Ressler, T. (1998). *J. Synchrotron Rad.* **5**, 118–122.

Sandmark, C. & Brändén, C. (1967). *Acta Chem. Scand.* **21**, 993–999.

Sarangi, R., Hocking, R. K., Neidig, M. L., Benfatto, M., Holman, T. R., Solomon, E. I., Hodgson, K. O. & Hedman, B. (2008). *Inorg. Chem.* **47**, 11543–11550.

Schlichtkrull, J. (1956). *Acta Chem. Scand.* **10**, 1455–1458.

Smith, G. D., Ciszak, E., Magrum, L. A., Pangborn, W. A. & Blessing, R. H. (2000). *Acta Cryst.* D**56**, 1541–1548.

Smith, G. D., Ciszak, E. & Pangborn, W. (1996). *Protein Sci.* **5**, 1502–1511.

Smith, G. D., Pangborn, W. & Blessing, R. H. (2001). *Acta Cryst.* D**57**, 1091–1100.

Smith, G. D., Pangborn, W. A. & Blessing, R. H. (2003). *Acta Cryst.* D**59**, 474–482.

Smith, G. D., Pangborn, W. A. & Blessing, R. H. (2005). *Acta Cryst.* D**61**, 1476–1482.

Smolentsev, G. & Soldatov, A. (2006). *J. Synchrotron Rad.* **13**, 19–29.

Smolentsev, G. & Soldatov, A. V. (2007). *Comput. Mater. Sci.* **39**, 569–574.

Strange, R. W., Ellis, M. & Hasnain, S. S. (2005). *Coord. Chem. Rev.* **249**, 197–208.

Whittingham, J. L., Chaudhuri, S., Dodson, E. J., Moody, P. C. & Dodson, G. G. (1995). *Biochemistry*, **34**, 15553–15563.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yano, J. & Yachandra, V. K. (2008). *Inorg. Chem.* **47**, 1711–1726.