Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The 1.23 Å structure of *Pichia pastoris* lysyl oxidase reveals a lysine–lysine cross-link

Anthony P. Duff,[a]‡ Aina E. Cohen,[b] Paul J. Ellis,[b] Kim Hilmer,[c] David B. Langley,[a] David M. Dooley,[c] Hans C. Freeman[a]* and J. Mitchell Guss[a]*

[a]School of Molecular and Microbial Biosciences, University of Sydney, NSW 2006, Australia, [b]Stanford Synchrotron Radiation Laboratory, 2575 Sandhill Road, Menlo Park, CA 94025, USA, and [c]Department of Chemistry and Biochemistry, Montana State Universtiy, Bozeman, MT 59717, USA

‡ Present address: Australian Nuclear Science and Techology Organisation, Private Mail Bag 1, Menai, NSW 2234, Australia.

Correspondence e-mail: freemanh@chem.usyd.edu.au, m.guss@mmb.usyd.edu.au

Received 21 April 2006
Accepted 6 July 2006

PDB Reference: PPLO, 1w7c, r1w7csf.

The structure of *Pichia pastoris* lysyl oxidase (PPLO) in a new crystal form has been refined at 1.23 Å resolution. PPLO, a copper amine oxidase (CuAO) with a 2,4,5-trihydroxyphenylalanine quinone (TPQ) cofactor, differs from most other members of the CuAO enzyme family in having the ability to oxidize the side chain of lysine residues in a polypeptide. In the asymmetric unit of the crystals, the structure analysis has located residues 43–779 of the polypeptide chain, seven carbohydrate residues, the active-site Cu atom, an imidazole molecule bound at the active site, two buried Ca^{2+} ions, five surface Mg^{2+} ions, five surface Cl^{-} ions and 1045 water molecules. The crystallographic residuals are R = 0.112 and $R_{\rm free}$ = 0.146. The TPQ cofactor and several other active-site residues are poorly ordered, in contrast to the surrounding well ordered structure. A covalent cross-link is observed between two lysine residues, Lys778 and Lys66. The cross-link is likely to have been formed by the oxidation of Lys778 followed by a spontaneous reaction with Lys66. The link is modelled as dehydrolysinonorleucine.

1. Introduction

The lysyl oxidase of the yeast *Pichia pastoris* (PPLO) is a copper amine oxidase (CuAO) with a 2,4,5-trihydroxyphenylalanine quinone (TPQ) cofactor derived by post-translational modification of a conserved tyrosine residue (Dove *et al.*, 1996). CuAOs catalyze the oxidative deamination of primary amines,

$$R\mathrm{CH_2NH_3^+} + \mathrm{O_2} + \mathrm{H_2O} \rightarrow R\mathrm{CHO} + \mathrm{NH_4^+} + \mathrm{H_2O_2}$$

and have been found in a wide range of organisms from bacteria to humans, where they appear to perform a multitude of biological roles (Dove & Klinman, 2001). In microorganisms, amine oxidases generally have a nutritional role in the utilization of primary amines as the sole source of nitrogen or carbon. In plants, they have been implicated in wound healing. In mammals, functions such as detoxification have been identified and some CuAOs are known to be tissue-specific. In humans, an increased plasma concentration of a CuAO named vascular adhesion protein-1/semicarbazide-sensitive amine oxidase (VAP-1/SSAO) is known to be associated with metabolic and vascular diseases (Salmi & Jalkanen, 1996; Smith *et al.*, 1998). Since the physiological function of amine oxidases in higher organisms is frequently associated with the breakdown or transformation of biologically active amines, these enzymes are likely to act as regulators of physiological amine concentrations and therefore may participate in numerous biological processes (Cohen, 1988). In humans, two TPQ-containing CuAOs play critical physiological roles: the

diamine oxidase, which appears to function in histamine metabolism (Elmore *et al.*, 2002), and the 'semicarbazide-sensitive' amine oxidase (SSAO), recently identified as the vascular adhesion protein-1 (VAP-1), which mediates L-selectin-dependent lymphocyte adhesion to endothelial cells (Salmi & Jalkanen, 1996; Smith *et al.*, 1998).

We have previously reported the structure of PPLO at 1.65 Å (Duff *et al.*, 2003). The other structurally characterized CuAOs are *Escherichia coli* amine oxidase (ECAO; Parsons *et al.*, 1995), *Pisum sativum* amine oxidase (PSAO; Kumar *et al.*, 1996), *Arthrobacter globiformis* amine oxidase (AGAO; Wilce *et al.*, 1997), *Hansenula polymorpha* amine oxidase (HPAO; Li *et al.*, 1998), bovine serum amine oxidase (BSAO; Lunelli *et al.*, 2005) and the human vascular adhesion protein (VAP-1; Airenne *et al.*, 2005; Jakobsson *et al.*, 2005).

These CuAOs are homodimers of similar size and topology (Fig. 1*a*). In each subunit, the N-terminal domains D2 and D3 pack against a large core, which comprises the two symmetry-related D4 domains. Two long β-hairpin arms protrude from each D4 domain and embrace the other D4 domain. Between the two D4 domains is a large solvent-filled 'lake'. In all the structures the lake is connected with the external solvent *via* openings on and around the dimer twofold axis. ECAO alone has an additional N-terminal D1 domain in each subunit.

One active site is buried deeply in each D4 domain and is accessed by substrates *via* a channel from the surface of the enzyme. The residues that line the channel belong to the D2, D3 and D4 domains of one subunit and to the tip of one of the β-hairpin arms of the symmetry-related subunit. In PPLO the substrate channel is a broad solvent-filled funnel. Each active site contains a Cu^{II} atom and a TPQ cofactor. Three conserved histidine side chains coordinate the copper. The TPQ is formed by the copper-dependent self-processing post-translational oxidation of a conserved tyrosine side chain (Matsuzaki *et al.*, 1994; Wilce *et al.*, 1997). In the mature enzyme, the TPQ has been reported in one or both of two conformations: an 'on-copper' conformation, in which the TPQ O4 atom is a copper ligand, and an 'off-copper' conformation in which the reactive TPQ O5 atom points into the substrate-binding site. In all native CuAO structures where the TPQ is off-copper, a well ordered water molecule is observed in the position occupied by the O4 atom when TPQ is on-copper and this position is usually designated the 'axial' water site. The Cu atom and its three histidine ligands are consistently well resolved, with Cu—N distances of ~2.0 Å. A water molecule is sometimes observed as a fifth copper ligand, in what is usually termed the 'equatorial' water site. In many CuAO structures no atom is modelled at this site, but a water molecule site is modelled 3.2–4.4 Å from the copper.

Enzyme catalysis occurs in two half-reactions *via* a ping-pong mechanism (Dove & Klinman, 2001),

$$\mathrm{TPQ} + R\mathrm{CH_2} \rightarrow \mathrm{TPQ_{red}} + R\mathrm{CHO}$$

(reductive half reaction),

$$\mathrm{TPQ_{red}} + \mathrm{O_2} + \mathrm{H_2O} \rightarrow \mathrm{TPQ} + \mathrm{NH_3} + \mathrm{H_2O_2}$$

(oxidative half reaction).

In the reductive half-reaction the TPQ reacts with the amine substrate, releasing water and producing an intermediate substrate Schiff base. Several crystal structures of the molecule trapped in this state have been reported (O'Connell *et al.*,

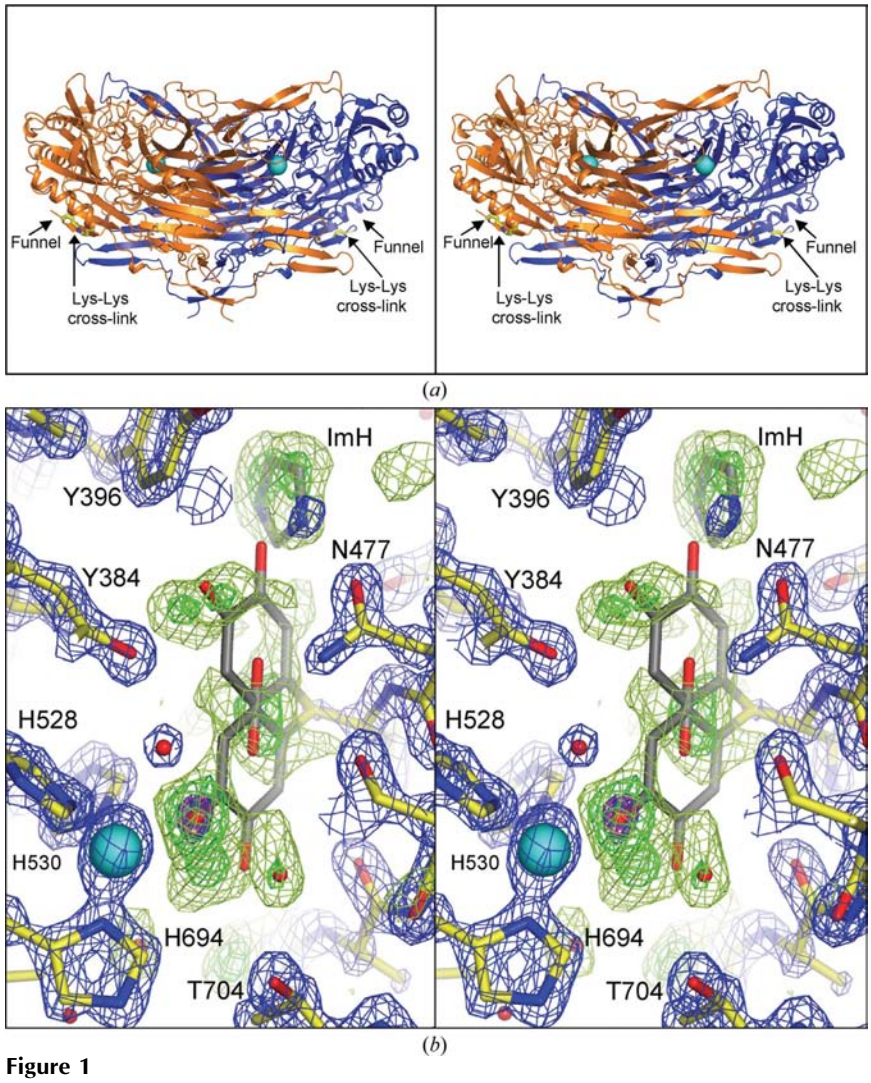

Figure 1
The structure of PPLO. (*a*) Stereoview of the dimer of PPLO with the newly identified cross-link and entrance to the active site indicated. The monomers are coloured orange and blue. The Cu atom is shown as a cyan sphere. (*b*) Order and disorder at the active site. Stereoview of unbiased OMIT electron density for the TPQ quinone ring and the imidazolium ion contoured at 9σ (magenta), 5σ (green), 2.5σ (olive) and final $2F_o - F_c$ electron density contoured at 2.75σ (blue). The TPQ side chain is shown in its standard on-copper and off-copper conformations. The TPQ beyond the C^{β} atom was not included in the refined model (see text).

2004; Wilmot *et al.*, 1997, 2004). A proton is abstracted, converting the substrate Schiff base to a product Schiff base. Hydrolysis yields the reduced aminoquinol (TPQ_{red}) and the aldehyde product, which is released. In the oxidative half-reaction, two electrons are transferred from the aminoquinol to dioxygen. The mechanism of this electron transfer remains uncertain. Hydrogen peroxide is released and a final hydrolysis reaction frees ammonia and regenerates the TPQ cofactor.

PPLO has the unusual ability to oxidize peptidyl lysine side chains, including lysine side chains in tropoelastin and collagen (Tur & Lerch, 1988). In this respect, it resembles another class of CuAOs, the lysine tyrosylquinone (LTQ) containing lysyl oxidases (Smith-Mungo & Kagan, 1998). In the absence of structural information for a true mammalian lysyl oxidase, the catalytic mechanism established for CuAOs, including PPLO, is currently the best model for the lysyl oxidases. We have previously reported the PPLO structure at 1.65 Å with R = 0.160 and R_{free} = 0.187. We now report a substantially improved structure in a new crystal form. The resolution and accuracy of the new PPLO structure are the highest reported so far for a CuAO.

2. Methods

2.1. Crystallization, X-ray data collection and structure solution

P. pastoris lysyl oxidase (PPLO) was overexpressed in *P. pastoris* GS115 cells. The expression and purification have been described previously (Dove *et al.*, 1996; Kuchar & Dooley, 2001). Crystals of a new crystal form were grown by vapour diffusion in 35% MPD, 200 m*M* $MgCl_2$ and 100 m*M* imidazole pH 8.0 at 277 K. Superior crystals were obtained by microseeding. A crystal was flash-cryocooled in a stream of N_2 gas at 100 K and data were recorded at SSRL beamline 9-2 at 12 654 eV (0.9811 Å). Data to 1.1 Å were measured, but were truncated at 1.23 Å, where $\langle I/\sigma(I)\rangle$ = 2.0. Reflections were integrated with *MOSFLM* (Leslie, 1999), scaled with *SCALA* and structure factors were obtained with *TRUNCATE* (Collaborative Computational Project, Number 4, 1994).

The structure was solved with *MOLREP* (Vagin & Teplyakov, 1997). The search model, a monomer, was derived from the original PPLO structure (PDB code 1n9e; Duff *et al.*, 2003), and contained the non-H atoms of residues *A*42–*A*774, excluding the side chain of TPQ*A*478. No ions or solvent atoms were used. Second conformers were removed and occupancies of all atoms in the search model were set to 1.

2.2. Refinement and model building

REFMAC v.5.1.24 (Murshudov *et al.*, 1997) was used for refinement. A 2% (4927 reflections) test set was used to calculate R_{free}. Following rigid-body and restrained refinement, $2F_o - F_c$, $F_o - F_c$ and anomalous difference electron-density maps were inspected with *O* (Jones *et al.*, 1991). The active-site Cu atom and the two tightly bound calcium ions (Duff *et al.*, 2003) were added. Potential water/ion sites were identified using *ARP*/*wARP* (Lamzin & Wilson, 1993). Water molecules in the active site were not added to the model at this early stage. Each *ARP*/*wARP* peak was inspected and modelled as a water molecule if it was associated with prior $2F_o - F_c$ density of least 1.5σ, had at least one hydrogen bond of length 2.6–3.3 Å and was not within 2.5 Å of any other hydrogen-bonding atom or within 3.3 Å of any C atom. Magnesium ions were identified by their regular octahedral coordination and short (∼2.1 Å) bonds to surrounding atoms. Chloride ions were identified by stronger electron density than expected for a solvent molecule, long (∼3.5 Å) hydrogen bonds and an anomalous electron-density peak (f'' = 0.3 e). Carbohydrate residues were built where their identity and orientation were unambiguous in an $F_o - F_c$ electron-density map.

Ambiguous or poorly resolved residues were assessed and rebuilt using OMIT maps. In many cases, these OMIT maps clearly indicated the presence of multiple conformations. The occupancies of alternate conformers were initially set to be equal and were then estimated using *CNS* (Brünger *et al.*, 1998). An individual atomic *B* factor for every atom in the model and grouped occupancies for the alternative conformers were refined iteratively until convergence. In cases where there were two or three alternative conformers, their occupancies were rounded to the nearest 10% or 5%, respectively. The *B* factors obtained by this procedure were discarded. A similar process was used to estimate the occupancy of the chloride and magnesium ions and an imidazolium ion. For those residues with multiple conformers but with overlapping electron density (TPQ478, Tyr480 and Lys778), refinement of the relative occupancies was not straightforward. In these cases the occupancies were fixed in steps of 5% from 0 to 100% and the models were refined with *REFMAC*. The final model was chosen based on *R* factor, residual electron density and agreement of the model with OMIT density.

When the refinement had reached the stage where R = 0.166 and R_{free} = 0.187, weakly restrained anisotropic *B*-factor refinement of all non-H atoms was introduced. Geometric restraints were loosened to give desired stereochemical values: bond-length r.m.s. deviations between 0.010 and 0.015 Å, angle r.m.s. deviations smaller than 1.5° and an average deviation $\langle\Delta\omega\rangle$ from peptide-group planarity between 6 and 7°. These changes reduced R to 0.129 and R_{free} to 0.160 and significantly improved the map quality in parts of the structure that were previously unclear. Continued re-evaluation of the model led to the modelling of additional residues with alternate conformations and of more ordered solvent. In several cases, the model was in conflict with the published sequence. Finally, with residuals R = 0.127 and R_{free} = 0.154, unbiased OMIT maps were used to model an imidazole in the active site and C-terminal residues 775–779, including an unexpected cross-link between the side chains of lysine residues 778 and 66.

MolProbity (Lovell *et al.*, 2003) was used for the detection of small modelling errors and with *WHATCHECK* (Vriend, 1990) and *PROCHECK* (Laskowski *et al.*, 1993) for the validation of the final coordinates. *VOIDOO* was used to identify

Table 1
Crystallographic data and refinement statistics.

Values in parentheses are for the highest resolution shell.

Space group	*C*2
Unit-cell parameters (Å, °)	a = 140.0, b = 67.0, c = 108.7, β = 119.0
Resolution range (Å)	24.0–1.23
Redundancy	2.8 (2.1)
$I/\sigma(I)$	12.8 (2.1)
Completeness (%)	95.4 (94.8)
$R_{\rm merge}$†	0.04 (0.51)
Wilson B (Å^2)	14.1
Average B value (Å^2)	22.8
No. of non-H atoms refined‡	7280
No. of unique reflections	242737 (17727)
$R_{\rm cryst}$§	0.112 (0.227)
$R_{\rm free}$¶	0.145 (0.259)
ESU based on maximum likelihood†† (Å)	0.027
DPI‡‡ (Å)	0.028
Ramachandran plot	
Residues in most favoured region (%)	88.2
Residues in allowed regions (%)	11.8
R.m.s.d. from standard geometry	
Bond lengths (Å)	0.012
Bond angles (°)	1.5

† $R_{\rm merge} = \sum_h \sum_i |I_{hi} - \langle I_h\rangle| / \sum_h \sum_i \langle I_h\rangle$. ‡ Includes buffer or anion atoms. § $R_{\rm cryst} = \sum_h |F_h({\rm obs}) - F_h({\rm calc})| / \sum_h F_h({\rm obs})$. ¶ $R_{\rm free}$ = $R_{\rm cryst}$ for approximately 2% of the data (4727 reflections) not used during refinement. †† Estimated uncertainty from *REFMAC* refinement (Murshudov & Dodson, 1997). ‡‡ Diffraction precision indicator (Cruickshank, 1999).

possible routes for the movement of O_2 and HOOH. Surfaces were calculated using a small (1.0 Å) probe to allow some protein flexibility, but introduced the risk of finding more cavities and tunnels than really exist. Where a residue was observed in multiple conformations, a separate calculation was carried out for each conformer. The TPQ side chain was placed in the (reactive) off-copper conformation for the surface calculations. *FLOOD* (Kleywegt & Jones, 1994) was used to estimate the size of the substrate channel in PPLO and in the other CuAOs by counting the number of water molecules that could be fitted into each channel. Figures were created with *PyMOL* (DeLano, 2002).

3. Results

The important new findings of this work are that the TPQ cofactor is disordered (Fig. 1*b*) while the rest of the active site is well ordered, that there is a gated pathway to the Cu site from the central 'lake', and that the enzyme has apparently oxidized one of its own lysine side chains with the resulting formation of an intramolecular lysine–lysine cross-link. New details of the structure include the previously unobserved C-terminal residues 775–779, a larger number of residues with multiple conformers, and an additional carbohydrate residue. Nine sequence conflicts have been identified and corrected. All these features reflect the high resolution of the data and the accuracy of the final phases. The estimated standard uncertainty based on maximum likelihood (Murshudov & Dodson, 1997) has nearly halved from 0.050 Å at 1.65 Å to 0.027 Å at 1.23 Å resolution. Similarly, the Diffraction Precision Indicator (DPI; Cruickshank, 1999), has decreased from 0.09 Å at 1.65 Å to 0.028 Å at 1.23 Å resolution.

Table 2
Sequence conflicts identified by the structure analysis.

Residue No.	Published amino acid, codon†	Amino acid as modelled	Possible codons	Mutations required at position(s)
314	Ile, ATT	Val	GT*X*	1
338	Glu, GAA	Asp	GA(T/C)	3
417	Ala, GCT	Ser	TC*X*, AG(T/C)	1
549	Lys, AAG	Gln †	CA(A/G)	1
577	Leu, TTA	Phe	TT(T/C)	3
579	Lys, AAG	Asn †	AA(T/C)	3
650	Glu, GAA	Thr	AC*X*	1 and 2
758	Leu, CTT	Val	GT*X*	1
761	Pro, CCT	Ala	GC*X*	1

† The original sequence is deposited in the SwissProt database as Q96X16 (Kuchar & Dooley, 2001).

The data and refinement statistics for the final model are summarized in Table 1. The asymmetric unit includes one monomer of the holoenzyme. Atomic coordinates have been determined for residues 43–779, with the exception of the side chain of TPQ478, which is found to be disordered (see below). Parts of the side chains of three residues, Glu44, Lys264 and Lys553, are unobserved. The unobserved atoms were included in the refinement with occupancies of 0.01. Other poorly resolved features include a surface residue (Asp310) and residues 737–743.

3.1. Evidence for sequence conflicts and microheterogeneity

Nine conflicts with the published sequence (Kuchar & Dooley, 2001) were revealed unambiguously (Table 2). OMIT maps were used to verify each of the conflicts. At these positions, the published amino acid fitted the electron density poorly and the fit deteriorated as the refinement progressed. For example, residue 577 was initially modelled and refined as leucine. At a stage when the residuals were R = 0.166 and $R_{\rm free}$ = 0.187, the electron density showed that the residue was actually phenylalanine (Fig. 2). Gln549, Asn579, Thr650 and Val758 were originally described as Lys, Lys, Glu and Leu, with alternate conformations that mimicked the branch point of the correct side chain. However, in each case neither conformer was a plausible rotamer. The former two residues had unacceptable steric clashes and the latter two simply did not fit into the observed electron density. Thr650 O$^{\gamma}$ forms a hydrogen bond that clearly identifies the correct side-chain rotamer. In the cases of Val314 and Asp338, which were originally modelled as Ile and Glu, the B factors of the incorrect additional atoms were unreasonably high. For Ser417, which was initially modelled as Ala, the side-chain O$^{\gamma}$ atom was indicated by two resolved peaks corresponding to alternate conformers. In each position, Ser417 O$^{\gamma}$ participates in a hydrogen bond and the electron density was insufficient to correspond to Val or Thr at 100% occupancy. For Ala761, which was initially modelled as Pro, there was no electron density at the positions of C$^{\gamma}$ and C$^{\delta}$, while the backbone NH group was clearly hydrogen bonded to a water molecule.

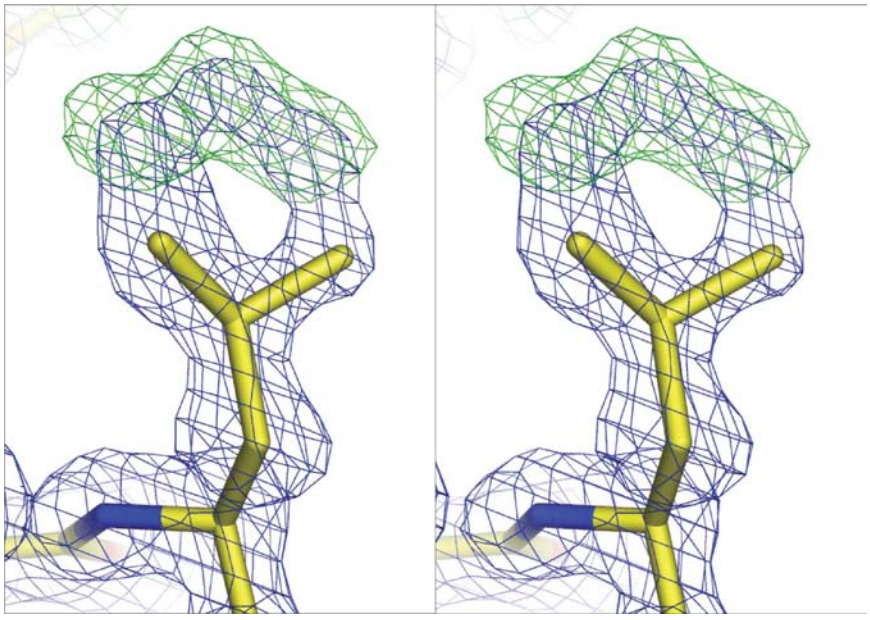

Figure 2
Evidence for a sequence conflict at Leu577. Stereoview of electron density at Leu577 at an intermediate stage of the refinement (R = 0.166, R_{free} = 0.187). $2F_o - F_c$ electron density contoured at 1.2σ (blue). Unbiased $F_o - F_c$ OMIT density contoured at 3σ (green). Residue 577 was subsequently modelled as phenylalanine.

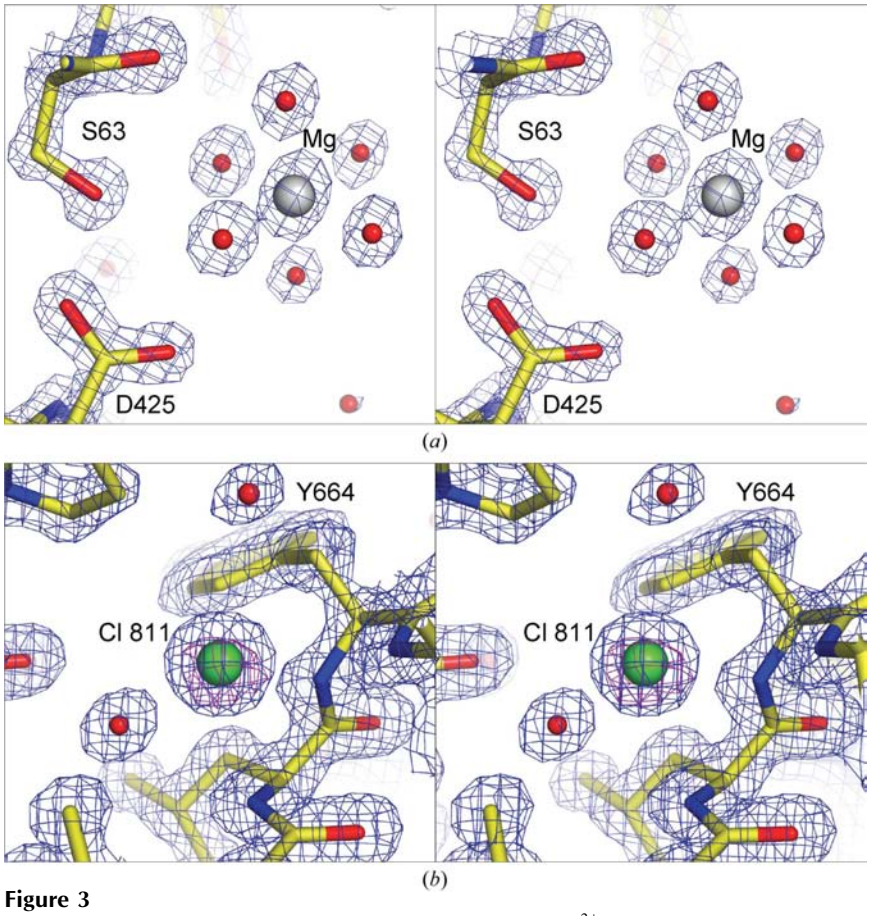

Figure 3
Stereoviews of examples of adventitious ions. (*a*) A $Mg(H_2O)_6^{2+}$ ion in a cleft on the protein surface. The $2F_o - F_c$ density is contoured at 3σ (blue). The six Mg—O distances, all ∼2.1 Å, were unrestrained during the refinement. (*b*) A Cl^- ion in a cleft on the protein surface. $2F_o - F_c$ electron density contoured at 1σ (blue). Anomalous electron density contoured at 3σ (magenta).

At seven other positions, micro-heterogeneity appeared to be present (Table 3). In each case, difference electron density indicated the presence of additional or fewer atoms consistent with a different amino acid. No allowance for the micro-heterogeneities was included in the final model.

3.2. Glycosylation

At least one carbohydrate residue is included at each of the five potential N-linked glycosylation sites identified by the presence of a consensus triplet sequence. The side chain of Asn81 is resolved in two conformations, in each of which it is attached to an *N*-acetyl-D-glucosamine (NAG) residue. Asn104 has three linearly attached carbohydrate groups: two NAG residues followed by a β-D-mannose residue (BMA). One NAG residue is observed attached to each of Asn191, Asn309 and Asn434. Residual electron density is consistent with some further glycosylation at these sites, but the density is too weak to permit accurate modelling.

3.3. Exogenous ions

Five Mg^{2+} and five Cl^- ions are observed on the protein surface, but none of these is in the lake, at the active site or in the substrate channel. The Mg^{2+} ions all show classic octahedral geometry (for an example, see Fig. 3*a*). Each of the Cl^- ions is bound to one or more exposed backbone NH groups. One of the chloride sites (residue 811; Fig. 3*b*) is in the same position as a sulfate ion (residues *A*–*D*811) in the four independent subunits of the 1.65 Å PPLO structure. A re-examination of the electron density of the 1.65 Å structure reveals that these four sulfate ions should have been modelled as chloride ions. The other sulfate sites in the 1.65 Å PPLO structure do not coincide with a chloride site in the 1.23 Å structure.

3.4. The active site

The active site is a region of contrast between order and disorder (Fig. 1*b*). The Cu atom and its histidine ligands are well ordered, but the reactive TPQ cofactor is disordered. Almost all protein atoms around the TPQ are well ordered with low *B* factors. The side chains of Tyr480 and Leu692 are observed in two conformations (Figs. 4

Table 3
Microheterogeneity suggested by the structure analysis.

Residue No.	Published amino acid, codon†	Apparent mixture	Possible codons for second residue	Mutations required at position(s)
326	Val, GTA	Val/Ile	AT(T/C/A)	1
337	Gly, GGT	Gly/Ala	GC*X*	2
362	Tyr, TAT	Tyr/Phe	TT(T/C)	2
413	Tyr, TAC	Tyr/Phe	TT(T/C)	2
469	Val, GTT	Val/Ile	AT(T/C/A)	1
498	Ala, GCT	Ala/Gly	GG*X*	2
618	Gly, GGA	Gly/Ser	TC*X*, AG(T/C)	1 and 2 or 1 and 3

† The original sequence is deposited in the SwissProt database as Q96X16 (Kuchar & Dooley, 2001).

and 5). Solvent in the active site ranges from well ordered molecules to molecules with partial occupancy and a high *B* factor.

The TPQ side chain could not be positioned using the X-ray data. The backbone and C^{β} atoms of TPQ478 were located in electron density and refined with small displacement parameters *B*, but the electron density beyond the C^{β} atom was discontinuous. Attempts to model and refine single or multiple conformations of the TPQ side chain were unsuccessful. The best agreement with an unbiased OMIT electron-density map was achieved when the TPQ was on-copper, but when the occupancy of the quinone ring in this conformation was varied, the lowest values of the residuals *R* and $R_{\rm free}$ resulted when the occupancy was zero. The omission of TPQ from the model accounts for some residual electron density. The model deposited in the PDB includes coordinates for both the on-copper and off-copper conformations of TPQ478, with zero occupancies.

3.5. Conformational flexibility at Leu692 and Tyr480

Leu692 is in two conformations; a *gauche*$^-$ (mt) rotamer (conformer *A*) and a *trans* (tp) rotamer (conformer *B*; Lovell *et al.*, 2000). The occupancies are 60 and 40%, respectively. Each conformer makes a prohibitively short contact with at least one solvent molecule, whose occupancy must therefore also be fractional (see below). Only conformer *A* was observed in the original crystal structure of PPLO at 1.65 Å (Duff *et al.*, 2003) and only conformer *B* was observed in the crystal structure of PPLO complexed with xenon (Duff *et al.*, 2004). In the latter structure, the bound Xe atom would sterically hinder conformer *A*.

OMIT electron-density maps show that the side chain of Tyr480 has two conformers with their aromatic ring planes inclined at an angle of 14° to each other (Fig. 4*a*). The distance between the O^{η} atoms of the two conformers is 1.45 Å. The positions of the side chains in the two conformers imply that they are associated with slightly different conformations of the polypeptide backbone at residues 478–482. There are no significant differences between corresponding side-chain torsion angles for the two conformers of Tyr480.

The side chain of Tyr480 in both of its conformations lies between the TPQ and the active-site base (Asp398) and makes contacts with both of them (Fig. 4*b*). The relationship is affected only slightly by whether the TPQ side chain is modelled on-copper or off-copper. When Tyr480 is in one conformation, the phenolic hydroxyl O^{η} lies 3.7–4.0 Å from the cofactor and 3.7 Å from Asp398 $O^{\delta 2}$. When Tyr480 is in its other conformation, the corresponding distances are 4.6–5.5 and 2.5 Å, respectively.

3.6. The copper-binding site

The active-site Cu atom gives rise to the strongest features in the electron density,

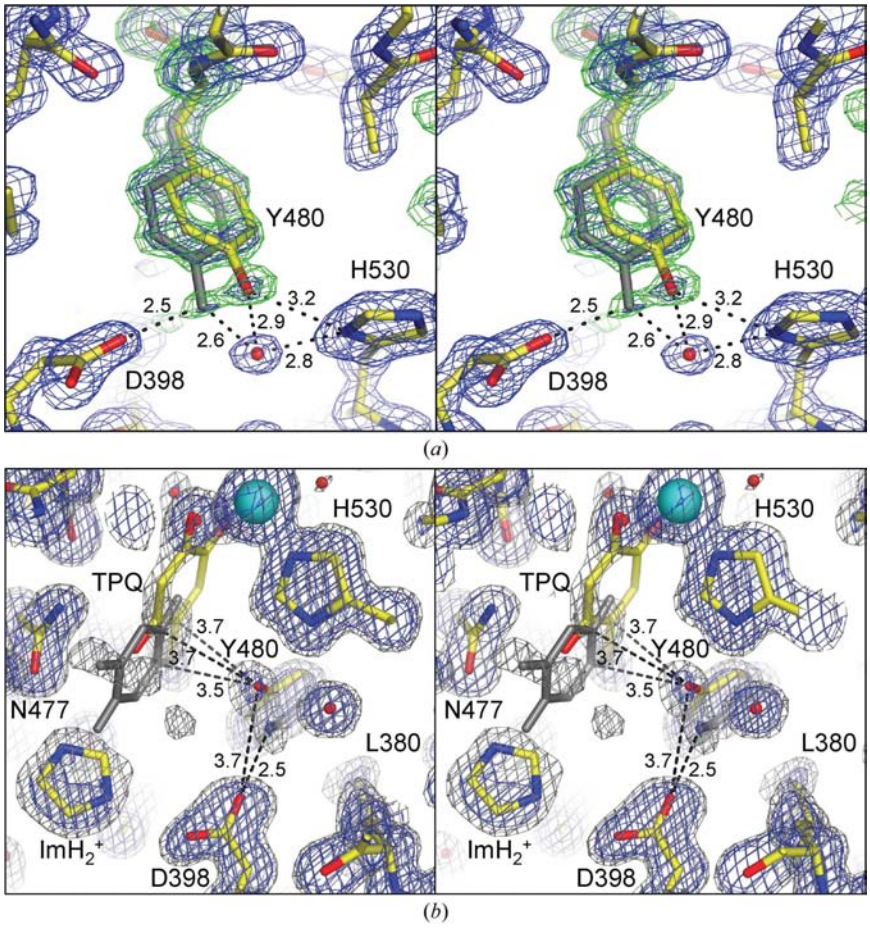

Figure 4
Stereoviews of Tyr480 resolved in two closely similar conformations. (*a*) OMIT electron density is contoured at 3.5σ (green) and final $2F_{\rm o} - F_{\rm c}$ density is contoured at 1.8σ (blue). Atoms of Tyr480 were omitted from the calculation of $F_{\rm c}$ and the structure was then re-refined before calculation of the electron-density map. Hydrogen bonds to Tyr480 OH are shown by dashed lines with distances in Å. (*b*) Final $2F_{\rm o} - F_{\rm c}$ density is contoured at 1.8σ (blue) and at 1.0σ (grey). TPQ (modelled but not refined) is shown with C atoms in the on-copper conformation (yellow) and in the off-copper conformation (grey). The Cu atom is shown as a cyan sphere. Contact distances from Tyr480 OH to the TPQ (modelled) and to Asp398 are shown in Å.

with a final $2F_o - F_c$ peak height of 30σ and the strongest anomalous peak. Four Cu ligands are identified. The three histidine ligands, His528, His530 and His694, are well ordered, with electron-density peaks corresponding to individual non-H atoms. The Cu—N distances are 2.02–2.03 Å. The presence of a fourth ligand is indicated by a difference electron-density peak of 3.7σ. This peak arises partly from a solvent atom and partly from the O4 atom of TPQ in its on-copper conformation. As the TPQ could not be modelled reliably, a water molecule is included in the model with 100% occupancy. To maintain consistency with previous literature, it is labelled O_{axial}. The refined $Cu—O_{axial}$ distance is 2.23 Å. There is no electron density corresponding to a fifth Cu ligand.

3.7. Access from the active site to the lake

Two water molecules (Z961 and Z472) were observed in the vicinity of the active site. Z961 (1σ in a $2F_o - F_c$ electron density map) is 3.4 Å from the Cu atom and <2 Å from the position that would be occupied by a fifth Cu ligand if the coordination geometry of the Cu atom were square pyramidal. Z961 was modelled with 60% occupancy, since it is 2.8 Å from Leu692 $C^{\delta2}$ when Leu692 is in conformation *B* (40% occupancy). Z961 makes a single 2.6 Å hydrogen bond to Z472 and its closest non-bonded neighbours are Ala709 C^{β} (3.4 Å) and Ala498 C^{β} (3.5 Å). The water molecule Z472 makes one additional hydrogen bond to an ordered water molecule in the central lake. The position of Z472, like that of Z961, is incompatible with the *B* conformer of Leu692.

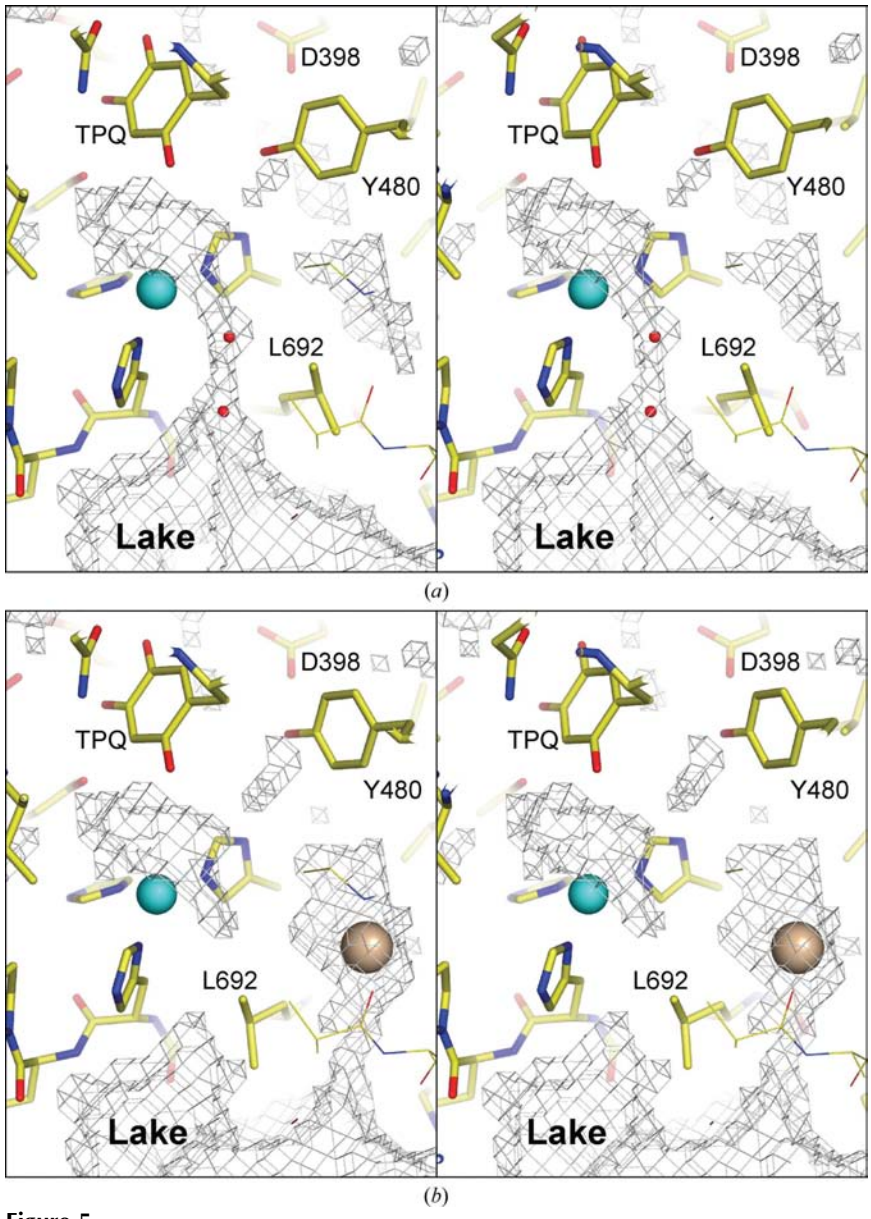

Figure 5
Stereoviews of internal cavities probed by a rolling sphere. The 1.0 Å probe-accessible surface is shown. The TPQ is in the off-copper conformation. (*a*) Surface calculated using conformation *A* of Leu692, associated residues and water molecules. In this orientation there is a single narrow tunnel connecting the copper site to the inland lake. (*b*) Surface calculated using conformation *B* of Leu692, associated residues and water molecules. In this conformation Leu692 closes the narrow tunnel seen in (*a*). On the other side of Leu692, a hydrophobic cavity, the previously identified 'common CuAO xenon site', is enlarged. The Xe atom (Xe 903) from the PPLO + Xe structure (Duff *et al.*, 2004; PDB code 1rky) is shown as a gold sphere.

3.8. An adventitious imidazolium ion

An imidazolium ion (ImH; PDB code IMD821), no doubt derived from the crystallization buffer, is located in the substrate-binding site (Fig. 1). It is well resolved and refinement indicated an occupancy of ∼70%. Each of the two N atoms of the imidazole ring makes a hydrogen bond. In one hydrogen bond, ImH821 N3···Asn477 $O^{\delta1}$, the imidazolium must be the donor since Asn477 $N^{\delta2}$ is the donor in another hydrogen bond to Tyr501 O, thus uniquely defining the orientation of the Asn477 side chain. In the other hydrogen bond, ImH821 N1···Asp398 $O^{\delta1}$, one or other of the atoms must be protonated. The position of the imidazolium ion is sterically incompatible with the off-copper conformation of the TPQ side chain. It follows that TPQ can be off-copper in at most ∼30% of the active sites in the crystal.

3.9. Measuring the active-site channels

To quantify the volumes of the active-site channels, we used *FLOOD* to fill the space with solvent molecules and counted the number of water molecules as a function of distance from O5 of TPQ when the cofactor is in its off-copper position (Fig.

6). The volume of the active-site channel is clearly larger in PPLO than in the other CuAOs (Fig. 6). Furthermore, whereas the substrate channel of PPLO opens rapidly into a broad funnel, the channels of the other CuAOs remain narrow for a longer distance from the active site (∼10 Å in PSAO, ∼14 Å in ECAO, AGAO and VAP-1 and ∼17 Å in HPAO). If we arbitrarily consider the volume of the channel from O5 of TPQ to a point 15 Å towards the protein surface, then the number of water molecules at this distance gives an estimate of the relative total volume of the channel. The narrowest substrate channel belongs to HPAO, which holds 12 water molecules within 15 Å of the TPQ. The channels in AGAO, VAP-1, PSAO and ECAO are remarkably similar, holding 21, 22, 23 and 25 water molecules, respectively, within the same distance. The channel of BSAO is larger (35 water molecules), but this can be attributed to the omission of residues 718–762 from the final refined model. Since BSAO and VAP-1 have 83% sequence identity, it is reasonable to assume that BSAO would closely resemble the structure of VAP-1 if the model of the former were complete. None of the other CuAOs has a channel volume as large as PPLO (70 water molecules). The PPLO channel begins to widen at a distance of ∼6 Å from the position of the N^{ζ} of a lysine residue modelled as a substrate bound to the enzyme. The place where the channel widens corresponds approximately to the position of the C^{α} atom of the lysine residue.

3.10. Alternate backbone conformations near the C-terminus

The electron density for the polypeptide backbone splits near the C-terminus at Tyr770, which is located on the rim of the substrate channel. Beyond this residue, the polypeptide has been modelled as conformer *A* (residues 770–779, occupancy 70%) and conformer *B* (residues 770–772, occupancy 30%). Conformer *A* is represented by clear and continuous electron density up to Lys779 N, which also contributes to the rim of the substrate channel. The electron density for conformer *B*, which points into the solvent, rapidly weakens, the last identifiable atom being Arg772 C^{α}. There is no evidence that the two conformers rejoin. Presumably, the polypeptide between the last observed residue in each conformer and the C-terminal residue 787 is disordered.

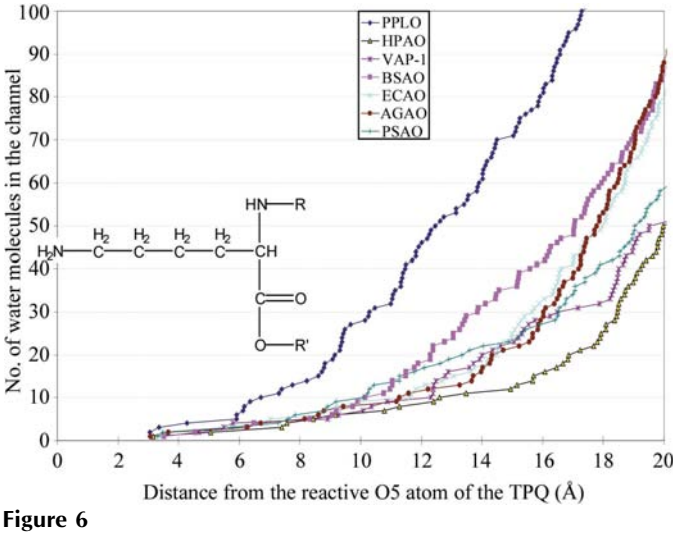

Figure 6
Volume of the substrate channel in CuAOs as a function of distance from the O5 atom of TPQ. The TPQ is assumed to be in its off-copper conformation. Each channel was loaded with water molecules using *FLOOD* (Kleywegt & Jones, 1994). The number of water molecules is plotted as a function of the distance *d* from the O5 atom of the TPQ. A lysine residue (drawn to scale) is shown in a position suitable for reaction with the TPQ.

3.11. Identification of an intramolecular lysine–lysine cross-link

When refinement of the structure had converged, continuous difference electron density remained between the side chain of Lys778 in conformer *A* (defined above) and the side chain of Lys66 of the same monomer (Fig. 7*a*). The cross-link was modelled and refined as dehydrolysinonorleucine (residue KNK778), comprising atoms derived from both Lys66 and Lys778. The electron density indicated that Lys778 of conformer *A* is partially free and partially cross-linked. The occupancies of the two components were refined with *REFMAC* in steps of 5%. The lowest values of *R* and $R_{\rm free}$ and the smallest residual electron density occurred when the free and cross-linked components of Lys778 both had occupancy 35%. In the free component of Lys778, the side-chain N^{ζ} atom is hydrogen bonded to Lys66 O (Fig. 7*b*). The side chain of Lys66 has approximately the same structure whether it is cross-linked (occupancy 35%) or not (occupancy 65%).

4. Discussion

4.1. Flexibility of TPQ

The TPQ side chain is not resolved in electron-density maps, despite the high resolution of the data and the high quality of the final phases (Fig. 1). This result is consistent with the observed orientations and mobility of the TPQ in ECAO, HPAO, AGAO and PSAO (Murray *et al.*, 1999). It is also consistent with the recently published structures of the mammalian CuAOs BSAO (Lunelli *et al.*, 2005) and SSAO (Airenne *et al.*, 2005; Jakobsson *et al.*, 2005), in which the TPQ is poorly resolved and variously observed either on-copper or off-copper. On the basis of the observations of the TPQ in the other CuAOs and of our observation of discontinuous electron density, we conclude that the TPQ exists in a broad range of energetically similar states. The flexibility of the TPQ may be important for the two biological processes in which it is involved; that is, its own biogenesis and the amine oxidase catalysis reaction. Moreover, the transitions among TPQ conformations may be rapid in the absence of external perturbations (*e.g.* substrate or inhibitor binding) that 'lock' a specific TPQ conformation into place. It is also conceivable that the TPQ dynamics vary among amine oxidases and that this accounts for differences among the detailed mechanisms of these enzymes.

4.2. Coordination geometry of the Cu atom

Unexpectedly, the dimensions of the Cu site (see supplementary material[1]) do not conform to the expected square-pyramidal coordination geometry (Kuchar & Dooley, 2001). Indeed, when the experimental ligand—Cu—ligand bond angles are compared with ideal values for regular tetrahedral, square-pyramidal and trigonal-pyramidal geometries, the smallest average deviation occurs for the last of these. It is not clear how this finding can be reconciled with the published electronic spectra (UV–Vis and EPR) for PPLO and other CuAOs (Kuchar & Dooley, 2001). One possibility is that the Cu atom in the PPLO crystal has been reduced by exposure to an intense X-ray beam. The reduction of Cu^{II} in synchrotron X-ray beams has been reported previously (Penner-Hahn *et al.*, 1989).

4.3. Intramolecular cavities and tunnels leading to the active site

The high quality of the present model and the new experimental evidence for alternate side-chain conformations make it appropriate to re-examine the routes by which small reactants travel to and from the active site. The oxidative half-reaction of the catalytic cycle requires O_2 as a reagent and produces H_2O_2 and NH_4^+ as byproducts. Since the reactions involving the amine substrate, the product aldehyde and the NH_4^+ take place on the channel side of the TPQ cofactor, it is assumed that these molecules travel to or from the active site *via* the channel. The reactions involving O_2 and H_2O_2 take place on the side of the TPQ remote from the channel and near the Cu atom. Is there a second connection between the active site and the external solvent?

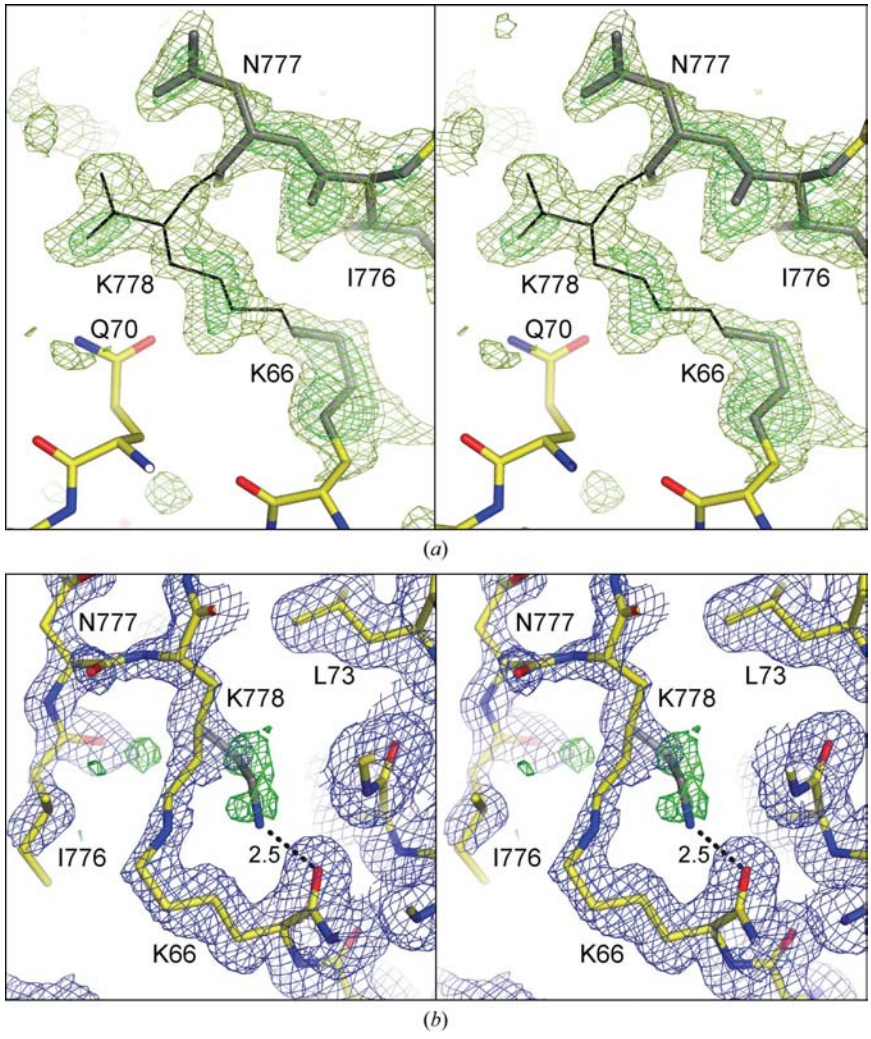

Figure 7
Stereoviews of the covalent Lys–Lys cross-link. (*a*) Continuous OMIT electron density connecting Lys778 to Lys66. $F_o - F_c$ electron density contoured at 4σ (green) and contoured at 2σ (olive). Atoms of residues Lys66, Ile776 and Asn777, connected by thick lines, were omitted from the calculation of F_c and the structure was then re-refined before calculation of the electron-density map. Lys779 N and atoms of Lys778, connected by thin black lines in their final refined positions, were not included in the model at any time prior to the calculation of this electron-density map. (*b*) The link between Lys66 and Lys778 modelled as dehydrolysinonorleucine (see Fig. 9). The final model is shown superposed on a $2F_o - F_c$ electron-density map contoured at 0.7σ (blue). Unbiased OMIT electron density for the alternate conformer of Lys778 is contoured at 3.0σ (green).

An examination of an accessible surface generated with a 1 Å probe suggests that there is no access to the Cu atom *via* the substrate channel. The bulky TPQ blocks this route. On the other hand, the probe-accessible surface reveals that there is an alternative route to and from the Cu site *via* a narrow tunnel from the central lake (Fig. 5*a*). The tunnel is open only when Leu692 is in conformation *A* and it contains the two sites partially occupied by water molecules Z961 and Z472. The wall of the tunnel is lined by the side chains of Leu692, Ala709, His694 and Thr704. The tunnel connects directly to a cavity adjacent to the Cu atom. In addition to the Cu atom, His530, His528, Tyr384, TPQ478, Tyr480, Ala498, Ile502 and main-chain atoms from residues 499 through 501 line the cavity. This cavity corresponds to the binding sites for NO and H_2O_2 in the complexes with ECAO (Wilmot *et al.*, 1999). The site proposed for oxygen binding in HPAO (Su & Klinman, 1998) is also revealed as a small cavity barely separated from the larger cavity. The present analysis shows that the TPQ blocks the connection between the Cu site and the bulk solvent *via* the substrate channel even when disordered, supporting previous proposals that small reagents travel to and from the active site *via* the lake (Li *et al.*, 1998; Su & Klinman, 1998; Wilce *et al.*, 1997).

When Leu692 is in the *B* conformation, the connection is closed, the two water molecules (Z961 and Z472) are displaced and another cavity adjacent to Leu692 is consider-

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference WD5066). Services for accessing this material are described at the back of the journal.

ably enlarged (Fig. 5*b*). This cavity is the xenon-binding site common to PPLO, AGAO, PSAO (Duff *et al.*, 2004) and BSAO (Lunelli *et al.*, 2005). The observation that Leu692 can exist in this conformation in the absence of xenon shows that the flexibility of Leu692 is real and is not an artefact of xenon binding. Other nearby residues that have multiple conformations, such as Tyr480, do not appear important for access to the Cu site.

The narrow tunnel identified above provides a plausible exit route for H_2O_2. The fact that the two solvent sites in the tunnel are only partly occupied suggests that the solvent molecules in these sites are not tightly bound, as expected for a region through which H_2O_2 passes. Viscosity experiments indicate that O_2 pre-binds to the enzyme (Su & Klinman, 1998). A suitable location for pre-binding of O_2 is the xenon-binding site common to a series of CuAOs (Duff *et al.*, 2004). This site is sufficiently far from the Cu and the TPQ to make it unlikely to be the reaction site. It is more likely to be a transient binding site. Movement of O_2 from this site requires the side chain of Leu692 to change from its *B* to its *A* conformation. In this scenario, the Leu692 side chain fills the position vacated by the O_2. The conformational change required at the Leu side chain is a 136° rotation of χ_1 from −54° (*gauche*$^-$) to 170° (*trans*). A 234° rotation in the opposite direction would sweep O_2 into the lake, but would involve the Leu side chain passing through the very unfavourable *gauche*$^+$ (pp) conformation (Lovell *et al.*, 2000)[2].

4.4. Comparison between the substrate channels in PPLO and other CuAOs

PPLO remains the only structurally characterized CuAO with a wide, open, deep and funnel-shaped substrate channel. As shown when the volumes of the channels are measured by filling them with water molecules (Fig. 6), PPLO has the largest channel. We note that according to this measurement the volume of the channel is much smaller in VAP-1 than in PPLO and that the volume in VAP-1 is closely similar to the volumes in the other CuAOs. This result disagrees with a published statement (Airenne *et al.*, 2005) that 'the active-site cavity of hVAP-1 resembles the wide active-site funnel of the PPLO structure more than the narrow substrate channel of ECAO, PSOA, AGAO and HPAO'. In fact, access to the TPQ is more restricted in VAP-1 and BSAO than in any of the other structurally characterized CuAOs. It is puzzling that despite its narrow substrate channel, VAP-1 resembles PPLO in being able to oxidize lysine side chains in tropoelastin and to form intermolecular cross-links in that protein (A. S. Weiss, personal communication).

PPLO also differs from the majority of CuAOs at the entrance to the substrate channel. Like HPAO, PPLO has large polypeptide protrusions on the rim of its substrate channel. (The protrusions were described as 'turrets' by Duff *et al.*, 2003.) In the structures of VAP-1, BSAO, AGAO, ECAO and PSAO, the molecular surface surrounding the channel is relatively flat.

A unique feature of PPLO is that its substrate channel completely lacks a constriction that limits access to the TPQ in the other CuAOs. This narrowing in the channel is defined by a residue on a β-strand of the D3 domain (Ser214 in PPLO), a residue on the β-turn preceding TPQ (Gly476 in PPLO) and a residue on the α-helix of the D4 domain (Tyr401 in PPLO) which carries the catalytic base (Asp398) (Fig. 8). In the other CuAOs, the residue

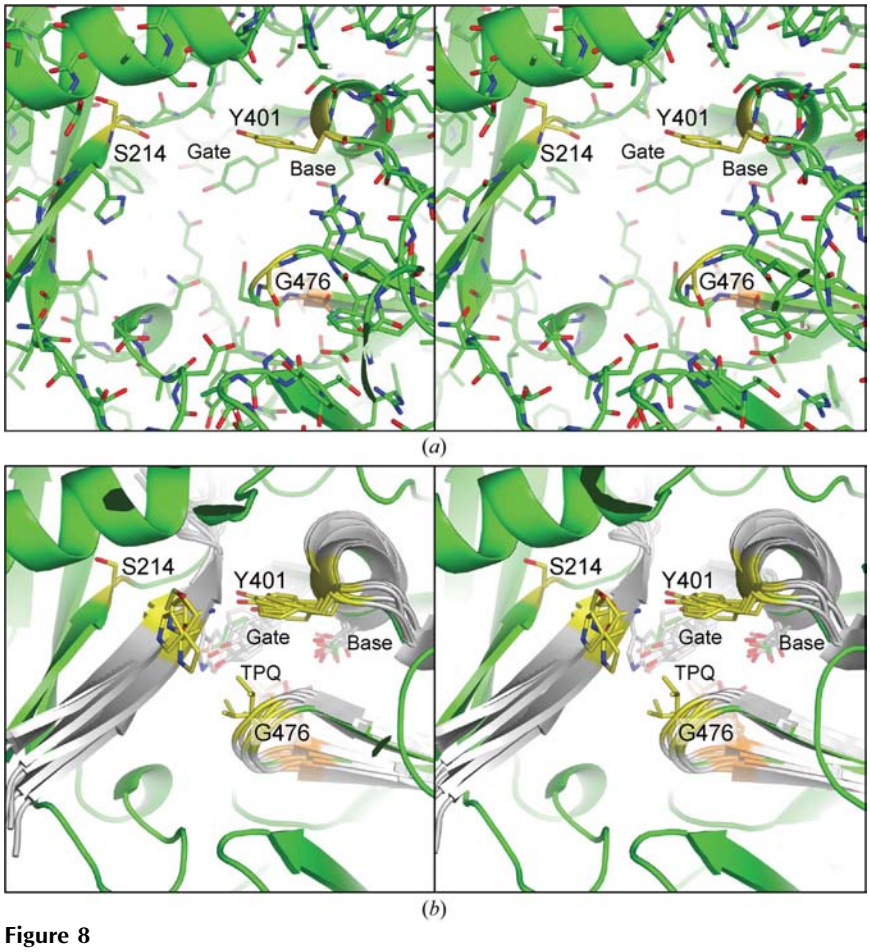

Figure 8
Stereoviews of the substrate channel. (*a*) Bottom of the substrate channel in PPLO. The selected residues are labelled and highlighted. The TPQ is obscured but is located behind Gly476. The 'gate' and 'base' denote residues Tyr396 and Asp398 in this structure, which are agreed to control access to the active site and to perform an essential catalytic function, respectively (Duff *et al.*, 2004). (*b*) Superposition of the substrate channels in seven CuAOs. The polypeptide chain of PPLO is shown in green and the others in grey. Residues forming the constriction in the channel are shown as stick bonds (yellow C atoms). The 'gate' and 'base' denote the structurally equivalent residues in the seven CuAOs.

[2] The conformations are specified using the new sign definitions of Lovell *et al.* (2000).

equivalent to Ser214 in PPLO is a residue of similar or slightly larger size (Ser, Thr or Pro). The position occupied by Tyr401 in PPLO is well conserved, the Tyr being replaced by Phe in some CuAOs. Gly476 in PPLO is unchanged in AGAO, ECAO and PSAO, but is replaced by Ala in HPAO and by Leu in BSAO and VAP-1. The bulky Leu residue in BSAO and VAP-1 means that access close to the TPQ is restricted more than in the other CuAOs. It has been noted previously that the Leu in VAP-1 completely blocks access, but that a narrow channel can be formed by relatively small movements of protein side chains (Airenne *et al.*, 2005). The combined effect of these structural factors is to fine-tune the opening in the substrate channel above the TPQ cofactor. In PPLO, a ∼6 Å movement of the β-strand of the D3 domain, which carries Ser214, away from Tyr401 and Gly476 completely removes the constriction in the substrate channel, allowing free access to the TPQ (Fig. 8).

4.5. Formation of an intramolecular lysine–lysine cross-link

The observation of a covalent link between Lys66 and Lys778 was unexpected. To the best of our knowledge, this is the first occasion when such a link has been seen crystallographically. It dramatically demonstrates the peptidyl lysyl oxidase activity of PPLO.

The chemistry of spontaneous cross-linking between a normal lysine side chain and an oxidized lysine side chain is known (Moran *et al.*, 1994). A lysyl oxidase oxidizes lysine to the corresponding aldehyde allysine. Allysine can then spontaneously react with a second lysine to form dehydrolysinonorleucine. The structural formulae of selected physiological cross-linking side chains are shown in Fig. 9.

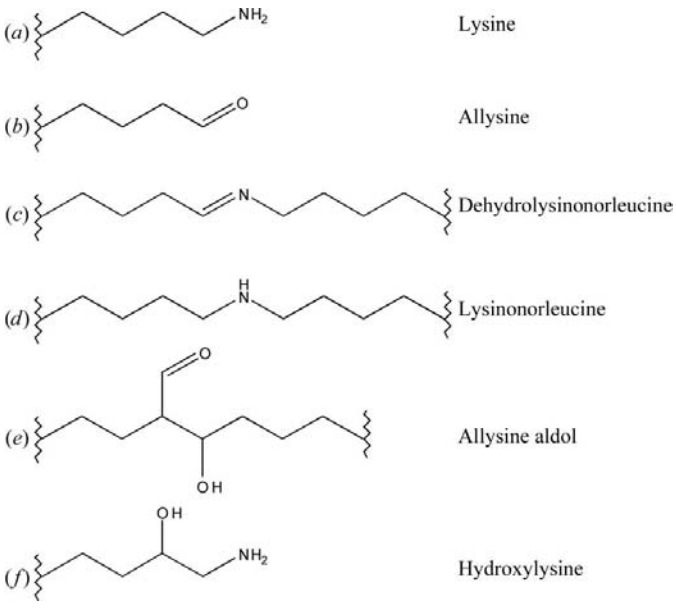

Figure 9
Lysine and some of its potential oxidation products. (*a*) Lysine. (*b*) Allysine. (*c*) Dehydrolysinonorleucine. (*d*) Lysinonorleucine. (*e*) Allysine aldol. (*f*) Hydroxylysine.

Simple modelling based on the structure of PPLO suggests that while Lys778 can readily reach into the active site of the same or neighbouring PPLO molecule, Lys66 cannot do so. Lys778 can reach the active site of PPLO by rotations in flexible torsion angles beyond residue Tyr770 in the polypeptide without moving well ordered sections of the protein. In contrast, Lys66 is well ordered in a well ordered region. Without substantial restructuring of well ordered sections of the protein, Lys66 cannot reach the active site of the same or of another PPLO molecule. An alternative lysine–lysine cross-link, allysine aldol, can be formed by the condensation of two oxidized lysine side chains. However, allysine aldol is two atoms shorter than dehydrolysinonorleucine and is clearly incompatible with the electron density observed in the present work.

The cross-link that we observe may be either dehydrolysinonorleucine (Bellingham *et al.*, 2003) or the chemically and structurally similar lysinonorleucine (Vrhovski & Weiss, 1998). Both are physiological cross-links that have been identified in biological tissues (Lent & Franzblau, 1967). Lysinonorleucine is the product of the two-electron reduction of the dehydrolysinonorleucine aldimine bond. These two cross-links are often not distinguished owing to the reactivity of dehydrolysinonorleucine and the use of the reducing agent borohydride in many sample preparations (Bailey *et al.*, 1970). Owing to the structural similarity of dehydrolysinonorleucine and lysinonorleucine, at the resolution of our data the electron density cannot be used to distinguish between the two. However, there are no reductants in the protein or crystallization buffers which would produce lysinonorleucine. On this basis, we have modelled the cross-link as dehydrolysinonorleucine.

Other variations of lysine–lysine cross-links are those that involve hydroxylysine, which has a hydroxyl group attached to C^{γ} (Moran *et al.*, 1994). Again, the electron density is not consistent with an additional hydroxyl group attached to Lys778 or Lys66. The cross-linking of Lys66 to Lys778 may be present and structurally important for PPLO *in vivo*. The presence of the cross-link should add structural stability to residues 770–779. When ordered, residues 770–779 contribute to the structure of the outermost rim of the substrate channel and they may contribute to substrate affinity and specificity in addition to total enzyme stability. In crystals of PPLO, only ∼35% of the potential Lys66–Lys778 cross-links are formed. The reason may be that the protein was chilled, purified and crystallized at 277 K very soon (0–24 h) after it was expressed. It has not been possible to test this hypothesis by producing well diffracting crystals from incubated protein.

This work was supported by the Australian Research Council (DP0208320 to HCF and JMG) and by NIH (GM27659 to DMD). Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology

Program is supported by the Department of Energy, Office of Biological and Environmental Research, by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the National Institute of General Medical Sciences.

References

Airenne, T. T., Nymalm, Y., Kidron, H., Smith, D. J., Pihlavisto, M., Salmi, M., Jalkanen, S., Johnson, M. S. & Salminen, T. A. (2005). *Protein Sci.* **14**, 1964–1974.

Bailey, A. J., Peach, C. M. & Fowler, L. J. (1970). *Biochem. J.* **117**, 819–831.

Bellingham, C. M., Lillie, M. A., Gosline, J. M., Wright, G. M., Starcher, B. C., Bailey, A. J., Woodhouse, K. A. & Keeley, F. W. (2003). *Biopolymers*, **70**, 445–455.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Cohen, S. S. (1988). Editor. *A Guide to Polyamines*. Oxford University Press.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.

DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, CA, USA.

Dove, J. E. & Klinman, J. P. (2001). *Adv. Protein Chem.* **58**, 141–174.

Dove, J. E., Smith, A. J., Kuchar, J., Brown, D. E., Dooley, D. M. & Klinman, J. P. (1996). *FEBS Lett.* **398**, 231–234.

Duff, A. P., Cohen, A. E., Ellis, P. J., Kuchar, J. A., Langley, D. B., Shepard, E. M., Dooley, D. M., Freeman, H. C. & Guss, J. M. (2003). *Biochemistry*, **42**, 15148–15157.

Duff, A. P., Trambaiolo, D. M., Cohen, A. E., Ellis, P. J., Juda, G. A., Shepard, E. M., Langley, D. B., Dooley, D. M., Freeman, H. C. & Guss, J. M. (2004). *J. Mol. Biol.* **344**, 599–607.

Elmore, B. O., Bollinger, J. A. & Dooley, D. M. (2002). *J. Biol. Inorg. Chem.* **7**, 565–579.

Jakobsson, E., Nilsson, J., Ogg, D. & Kleywegt, G. J. (2005). *Acta Cryst.* D**61**, 1550–1562.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kleywegt, G. J. & Jones, T. A. (1994). *Acta Cryst.* D**50**, 178–185.

Kuchar, J. A. & Dooley, D. M. (2001). *J. Inorg. Biochem.* **83**, 193–204.

Kumar, V., Dooley, D. M., Freeman, H. C., Guss, J. M., Harvey, I., McGuirl, M. A., Wilce, M. C. J. & Zubak, V. M. (1996). *Structure*, **4**, 943–955.

Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.

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

Lent, R. & Franzblau, C. (1967). *Biochem. Biophys. Res. Commun.* **26**, 43–50.

Leslie, A. G. W. (1999). *Acta Cryst.* D**55**, 1696–1702.

Li, R., Klinman, J. P. & Mathews, F. S. (1998). *Structure*, **6**, 293–307.

Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.

Lovell, S. C., Word, J. M., Richardson, J. S. & Richardson, D. C. (2000). *Proteins*, **40**, 389–408.

Lunelli, M., Di Paolo, M. L., Biadene, M., Calderone, V., Battistutta, R., Scarpa, M., Rigo, A. & Zanotti, G. (2005). *J. Mol. Biol.* **346**, 991–1004.

Matsuzaki, R., Fukui, T., Sato, H., Ozaki, Y. & Tanizawa, K. (1994). *FEBS Lett.* **351**, 360–364.

Moran, L. A., Scrimgeour, K. G., Horton, H. R., Ochs, R. S. & Rawn, J. D. (1994). *Biochemistry*, 2nd ed. Burlington, NC, USA: Neil Patterson Publishers.

Murray, J. M., Saysell, C. G., Wilmot, C. M., Tambyrajah, W. S., Jaeger, J., Knowles, P. F., Phillips, S. E. & McPherson, M. J. (1999). *Biochemistry*, **38**, 8217–8227.

Murshudov, G. N. & Dodson, E. J. (1997). *CCP4 Newsl.* **33**, 31–39.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

O'Connell, K. M., Langley, D. B., Shepard, E. M., Duff, A. P., Jeon, H.-B., Sun, G., Freeman, H. C., Guss, J. M., Sayre, L. M. & Dooley, D. M. (2004). *Biochemistry*, **43**, 10965–10978.

Parsons, M. R., Convery, M. A., Wilmot, C. M., Yadav, K. D., Blakeley, V., Corner, A. S., Phillips, S. E., McPherson, M. J. & Knowles, P. F. (1995). *Structure*, **3**, 1171–1184.

Penner-Hahn, J. E., Murata, M., Hodgson, K. O. & Freeman, H. C. (1989). *Inorg. Chem.* **28**, 1826–1832.

Salmi, M. & Jalkanen, S. (1996). *J. Exp. Med.* **183**, 569–579.

Smith, D. J., Salmi, M., Bono, P., Hellman, J., Leu, T. & Jalkanen, S. (1998). *J. Exp. Med.* **188**, 17–27.

Smith-Mungo, L. I. & Kagan, H. M. (1998). *Matrix Biol.* **16**, 387–398.

Su, Q. & Klinman, J. P. (1998). *Biochemistry*, **37**, 12513–12525.

Tur, S. S. & Lerch, K. (1988). *FEBS Lett.* **238**, 74–76.

Vagin, A. A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Vrhovski, B. & Weiss, A. S. (1998). *Eur. J. Biochem.* **258**, 1–18.

Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.

Wilce, M. C. J., Dooley, D. M., Freeman, H. C., Guss, J. M., Matsunami, H., McIntire, W. S., Ruggiero, C. E., Tanizawa, K. & Yamaguchi, H. (1997). *Biochemistry*, **36**, 16116–16133.

Wilmot, C. M., Hajdu, J., McPherson, M. J., Knowles, P. F. & Phillips, S. E. (1999). *Science*, **286**, 1724–1728.

Wilmot, C. M., Murray, J. M., Alton, G., Parsons, M. R., Convery, M. A., Blakeley, V., Corner, A. S., Palcic, M. M., Knowles, P. F., McPherson, M. J. & Phillips, S. E. (1997). *Biochemistry*, **36**, 1608–1620.

Wilmot, C. M., Saysell, C. G., Blessington, A., Conn, D. A., Kurtis, C. R., McPherson, M. J., Knowles, P. F. & Phillips, S. E. (2004). *FEBS Lett.* **576**, 301–305.