PERSPECTIVE

Dalton
www.rsc.org/dalton

A new framework for understanding substrate binding and functional diversity in haem peroxidases

Katherine H. Sharp,[a] Peter C. E. Moody[b] and Emma Lloyd Raven*[a]

[a] *Department of Chemistry and Department of Biochemistry, University of Leicester, University Road, Leicester, UK LE1 7RH. E-mail: emma.raven@le.ac.uk*

[b] *Department of Biochemistry, University of Leicester, University Road, Leicester, UK LE1 7RH*

Received 11th August 2003, Accepted 23rd September 2003
First published as an Advance Article on the web 13th October 2003

The haem-containing peroxidase enzymes catalyse the H_2O_2-dependent oxidation of a wide variety of substrates and have provided a focal point for our more general understanding of structure/function relationships in other, more complex haem enzymes. Mechanistically, the haem peroxidases are well characterised: they share a common catalytic cycle that involves formation of a high-oxidation-state (ferryl) intermediate. In contrast, our understanding of the structural features that control the diverse substrate specificity are less well defined. In this review, we discuss how recent spectroscopic and structural information for ascorbate peroxidase has provided new insight into the *modus operandi* of this enzyme and how this has helped to clarify certain aspects of the catalytic and, in particular, the substrate binding properties of the closely related cytochrome *c* peroxidase enzyme, which has been a benchmark for peroxidase catalysis over more than 20 years.

Introduction

There is an overwhelming body of evidence now available that establishes beyond doubt that the reactivity of a particular metal ion or metal complex can be altered when incorporated into a protein structure. The iron complex of protoporphrin IX, Fig. 1, is a good example. This simple macrocycle – or modifications of it – is an integral component of a very large group of proteins and enzymes and is capable of delivering an astonishing range of reactivity depending on the environment of the protein into which it is incorporated. Broadly, this includes electron transfer (the cytochromes), oxygen transport and storage (the globins) and redox catalysis (the haem enzymes). The availability of different oxidation states of the metal ion means that the haem group can be tuned into a particular reactivity according to the particular function required. Nature has been very efficient in her design and we are left to wonder precisely how this simple macrocycle, when incorporated into different protein structures, can deliver such a rich diversity of function. Or, in other words, how does the structure of the protein that surrounds the metal centre control its function?

This question has occupied the minds of many scientists over many years. We do not have a complete answer but there are several variables that are known to be influential. These include the identity of the axial ligands that occupy the fifth and sixth coordination positions of the haem, the solvent accessibility and the polarity of the haem environment. For the catalytic haem enzymes, the utilisation of different substrates, and

Katie Sharp was born in Oldham, Lancashire in 1979 and graduated with a Master of Chemistry with Drug Design and Toxicology from the University of Hull in 2001. She is currently studying for her PhD in chemistry at the University of Leicester. Her research work involves the molecular evolution of haem proteins for enhanced biocatalysis, and her PhD work is sponsored by the Institute of Applied Catalysis (iAc).

Peter Moody is a protein crystallographer with interests in enzyme structure and function. He has a BA in Biochemistry from the University of York and a PhD in Biophysics from Imperial College. Since then he has worked with William N. Lipscomb at Harvard, Andrew Leslie at Imperial College and Guy Dodson at York, before taking a post in the Biochemistry Department at the University of Leicester in 1995. He is currently a Senior Lecturer.

Emma Raven was born and brought up in Northamptonshire and obtained her BSc in Chemistry from the University of Leicester in 1988 and her PhD in Chemistry from the University of Newcastle Upon Tyne in 1991 (with Professor A. G. Sykes, FRS). She then moved to the west coast of Canada to work with Professor Grant Mauk at the University of British Columbia (Vancouver). In 1995, she returned to the UK to a Lectureship in Inorganic Chemistry at the University of Leicester and was promoted to Senior Lecturer in 2002. She currently holds a research fellowship from the Wellcome Trust.

Katie Sharp

Peter Moody

Emma Raven

DOI: 10.1039/b309648n

This journal is © The Royal Society of Chemistry 2003

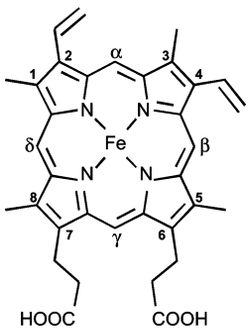


Fig. 1 The structure of iron protoporphyrin IX. The identity of the eight substituent groups bound to the porphyrin vary amongst different haem proteins,[25] but are not discussed here. Five- and six-coordinate haem proteins are known. For five-coordinate haem proteins, the axial ligand is provided by a protein residue capable of metal ligation (*e.g.* histidine, tyrosine, cysteine); six-coordinate haem proteins contain either two protein ligands, as above, or one protein ligand and an exogenous ligand (*e.g.* H_2O, O_2, CO *etc*). Haem nomenclature is non-systematic. The word haem is a general term that communicates no information about the structure of the porphyrin or the oxidation state; this nomenclature will be adopted in this Perspective.

variations in the accessibility of these substrates to the haem macrocycle, is also important.

There are various experimental routes through which one might address this question but, on the whole, isolation of any one of these variables, while keeping all others constant, is not a trivial exercise. By far the most powerful experimental approach is the application of site-directed mutagenesis methods[1,2] to the protein, or proteins, of interest. Using this technique, it is possible to alter one (or more than one) amino acid for any another of the remaining 19 amino acids at target positions in the protein architecture. This can provide unique information on the specific role(s) of individual amino acid residues in controlling function and can be used to build an overall picture of the *modus operandi*. This technique has been successfully applied to numerous haem-containing proteins and enzymes over the past 15–20 years (reviewed in refs. 3–21) and, collectively, the information that has emerged has allowed us to build a more integrated picture of haem structure/function relationships.

An absolute pre-requisite for site-directed mutagenesis work is that an expression system is available for production of the protein of interest in quantities that are sufficient for functional studies. If such a system is available, it is highly desirable (but not essential) that a high-resolution crystal structure for the protein also exists because it allows detailed rationalisation of functional data at the molecular level. A few haem proteins (cytochrome *c*, cytochrome b_5, myoglobin) met both of these criteria at an early stage and, inevitably, were therefore subjected to such intense experimental scrutiny that they provided an important benchmark against which all other haem proteins could be compared. One of these benchmark enzymes was cytochrome *c* peroxidase (C*c*P), which was first studied by mutagenesis in 1986.[22] In this Perspective, we describe how the study of the closely-related ascorbate peroxidase enzyme has provided an important new framework against which the established properties of C*c*P can be directly compared, and how very recent advances in our understanding of ascorbate peroxidase have helped to clarify our appreciation of some of the anomalous properties of C*c*P.

Haem peroxidases

We begin with a general introduction. The haem peroxidases (reviewed in refs. 6, 8, 10, 13, 14 and 23) are a very important family of enzymes that have provided a focal point for our more general understanding of haem enzymes and a workable experimental framework against which this can be measured. They are involved in a diverse range of biological processes,[24] from lignin degradation through to antibacterial and antifungal action. In general, they contain iron protoporphyrin IX, Fig. 1, as prosthetic group; although there are some exceptions to this (in the mammalian peroxidases),[25,26] these will not be considered here. They catalyse the H_2O_2-dependent oxidation of a very wide variety of substrates, usually (but not exclusively) organic substrates, and comprise two separate groups: the plant peroxidases and the mammalian peroxidases. The plant peroxidase family – which contains plant, fungal and bacterial enzymes – is by far the most well characterised, and has been classified[27] into three types. Class I peroxidases contain the prokaryotic enzymes (examples include cytochrome *c* peroxidase (C*c*P) and ascorbate peroxidase (APX)), class II peroxidases contain the fungal enzymes (*e.g.* lignin peroxidase and manganese peroxidase) and class III contain the classical secretory peroxidases (*e.g.* horseradish peroxidase (HRP)).

Structural and mechanistic aspects

From a structural and mechanistic standpoint, the haem peroxidases are very well characterised. For many years, C*c*P was the only peroxidase for which a crystal structure was available[28,29] but, after a lengthy gap, new structures emerged for other peroxidases. These included horseradish peroxidase,[30] manganese peroxidase,[31] myeloperoxidase (a mammalian enzyme),[32,33] barley peroxidase,[34] *A. ramosus* peroxidase,[35] *C. cinereus* peroxidase[36] (*A. ramosus* and *C. cinereus* peroxidases are identical), lignin peroxidase,[37–40] peanut peroxidase,[41] ascorbate peroxidase[42,43] and a (class I) catalase-peroxidase[44] enzyme. (A crystal structure for chloroperoxidase is also available,[45] but this enzyme is not part of the mammalian or plant superfamilies and has properties that liken it to cytochrome P450.) The active site structure of C*c*P is shown in Fig. 2(a). With the exception of Trp51 and Trp191, these active site residues are invariant across the plant peroxidase family and the active site topology is highly conserved.

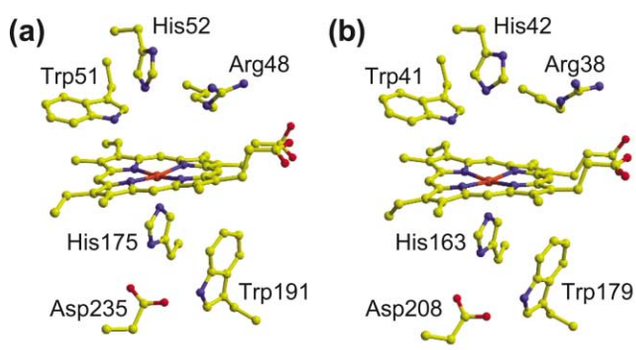
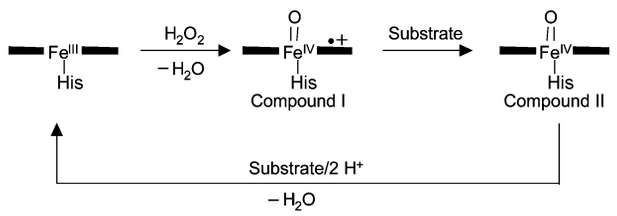


Fig. 2 The active site of (a) cytochrome *c* peroxidase[69] and (b) ascorbate peroxidase,[43] showing the haem and various active site residues.[104,105]

Their intimate mechanism has been studied in detail and site-directed mutagenesis work on C*c*P, and later HRP, has revealed an intricate picture of the roles of various active site amino acids. This work has been reviewed extensively (see, for example, refs. 10, 12–14, 21 and 46) and will not be rehearsed here. All peroxidases examined to date share a common catalytic cycle that involves reaction of the ferric haem with hydrogen peroxide in a rapid ($k \approx 10^7$ M^{-1} s^{-1}) reaction that leads to formation of a high-valent ferryl intermediate, known as Compound I, through heterolytic cleavage of the O–O peroxide bond and loss of a water molecule. Compound I is a two-electron oxidised intermediate that, for most peroxidases, contains a Fe^{IV}=O haem and a porphyrin π-cation radical. The distal histidine (His52 in C*c*P and His42 in HRP) is particularly important in Compound I formation.[10,12,14,21,46] Formation of Compound I is followed by one-electron reduction of

Compound I by substrate, first to Compound II and then back to ferric haem (although direct reduction of Compound I to ferric haem is also known, for example in the mammalian peroxidases), Scheme 1.



Scheme 1 The catalytic mechanism of haem peroxidases. The thick black line represents the haem ring. In the first step, the enzyme in the (ferric) resting state reacts with H_2O_2 in a very fast second-order reaction that leads to formation of the catalytic intermediate Compound I. Compound I is then reduced by the substrate in two sequential single electron transfer reactions, to regenerate the resting enzyme. A histidine residue (His) provides the fifth ligand to the haem iron.

A consistent picture has therefore emerged for Compound I formation in the haem peroxidases: they have broadly similar active site structures that have been designed to support a catalytic mechanism that, in all cases, favours formation of a common, high-oxidation state, haem intermediate. The precise structural features that favour stabilisation of ferryl haem, as opposed to ferrous haem (for example in the globins), have been examined extensively using mutagenesis but are not completely understood, although the presence of charged residues in the peroxidase active are thought to encourage separation of the O–O peroxide bond. This has been reviewed previously [17–19] and is not considered further here.

Substrate binding

In contrast to the structural and mechanistic aspects, the identity of the substrate that reduces Compound I, Scheme 1, varies enormously. It is this variety, and not the active site architecture (*vide supra*), that delivers such rich diversity of biological function across the family. For example, the substrate can be a metal ion (Mn^{2+} in manganese peroxidase), a halide (Cl^- in myeloperoxidase), an organic (L-ascorbate in ascorbate peroxidase) or phenolic (ferulic acid in horseradish peroxidase) substrate, or a protein molecule (cytochrome *c* in cytochrome *c* peroxidase). Quite often, the peroxidase will oxidise more than one substrate; in these cases it is difficult to establish which is the physiological substrate and to assign a biological role (horseradish peroxidase is an example).

Because the substrate identity is so varied, and because structural information for peroxidase/substrate complexes [47,31,33,48–53] has been slow to emerge compared to the mechanistic information, defining the structural features that control substrate binding and recognition – the substrate specificity – has been much more challenging. Indeed, this has been prohibitive in terms of our complete understanding of how functional diversity is achieved in this group of haem enzymes. Below, we discuss very recent advances in our understanding of substrate binding in ascorbate peroxidase, a class I haem peroxidase, and describe how this has shed new light on more general our understanding of the functional properties.

The class I haem peroxidases – cytochrome c peroxidase and ascorbate peroxidase. *Cytochrome c peroxidase.* Of all the peroxidases, cytochrome *c* peroxidase (C*c*P) was the first to have its crystal structure solved [28,29] and the first to be over-expressed in *E. coli* for mutagenesis work.[54] These developments paved the way for extensive structure/function studies that defined many mechanistic aspects of peroxidase catalysis for the first time. Because of this, C*c*P became *the* benchmark against which all other peroxidases were normally compared. However, although many of the key features of the C*c*P enzyme are duplicated across the rest of the peroxidase family, it has two anomalous features that are not typical of other peroxidases:

(i) C*c*P uses an atypical Compound I intermediate that contains a protein-based radical – localised on Trp191,[55–59] Fig. 3 – rather than the more usual porphyrin-based radical as observed in other peroxidases (*e.g.* HRP), Fig. 3;

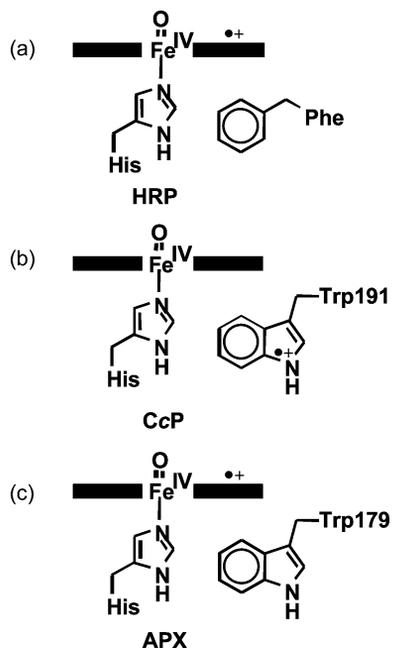


Fig. 3 The structures of the Compound I derivatives in (a) HRP, (b) C*c*P and (c) APX. The thick black line represents the haem ring and the axial histidine ligand is shown. For HRP, the second oxidising equivalent resides on the porphyrin ring (porphyrin π-cation radical); for C*c*P, an aromatic tryptophan residue, Trp191, adjacent to the haem is used ($Trp^{\bullet+}$ radical); for APX, a porphyrin π-cation radical is used despite APX containing the same Trp residue as is found in C*c*P.

(ii) C*c*P uses an atypical protein substrate, cytochrome *c*, that is not representative of other, small molecule (usually organic) peroxidase substrates.

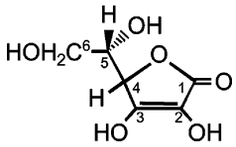
The working hypothesis that emerged over several years to account for the first of these anomalies was based on amino acid sequence comparisons. C*c*P contains an oxidisable amino acid (Trp191) close to the haem that is used in Compound I formation. Most other peroxidases lack an oxidisable residue at this position (replaced by a phenylalanine residue): oxidation of Phe is energetically less favourable (higher reduction potential) and a porphyrin π-cation is used instead, Fig. 3.

Ascorbate peroxidase. The first opportunity to directly test the above hypothesis was provided when it realised that another class I haem peroxidase, ascorbate peroxidase (APX), had high (>30%) sequence identity [60] to C*c*P, including the residue equivalent to Trp191 in C*c*P. Here we describe how functional studies on APX have helped to clarify our overall appreciation of the anomalous features of the C*c*P enzyme.

Ascorbate-dependent peroxidase activity was first reported in 1979,[61,62] but early isolation procedures did not generate workable amounts of enzyme and very little detailed progress was made for another 15 years. All APXs (reviewed in refs. 63–65) identified to date contain iron protoporphyrin IX, Fig. 1. They catalyse the H_2O_2-dependent oxidation of ascorbate in plants, algae and certain cyanobacteria, eqn. (1).

$$2\,\text{ascorbate} + H_2O_2 \longrightarrow 2\,\text{monodehydroascorbate radical} + 2\,H_2O \quad (1)$$

APX enzymes show high specificity for L-ascorbate, Scheme 2, as electron donor, but will also oxidise non-physiological (usually organic) substrates that are characteristic of the class



Scheme 2 The structure of L-ascorbic acid, showing the L configuration at C^5. The 2-OH and 3-OH groups have pK_a values of 11.3 and 4.0, respectively,[106] so that at neutral pH the 3-OH group is deprotonated and the molecule is anionic.

III peroxidases, in some cases at rates comparable to that of ascorbate itself.[63–65] Hence, and in direct contrast to the anomalous C*c*P molecule, the physiological substrate (ascorbate) in APX is much more typical of other peroxidases (this itself was a key driver for studies on APX). In fact, the substrate specificity of APX places it at the interface between the class I and class III peroxidases, a distinction that has been highlighted recently.[66] In principle, APX therefore provided a perfect opportunity to reassess the unusual properties of C*c*P and its status as the benchmark peroxidase, and at the same time provided a very convenient comparative framework that was important for our understanding of substrate binding and oxidation across the entire peroxidase family. The development of recombinant bacterial expression systems for pea cytosolic (rpAPX)[67] and soybean cytosolic[68] (rsAPX) APXs, together with new structural information,[42] provided the first major step towards this overall aim and made structure/function studies on APX feasible for the first time.

The first crystal structure to appear for an APX was that for the recombinant pea cytosolic enzyme (rpAPX), which was published by Patterson and Poulos in 1995.[42] Subsequently, a crystal structure for the closely related recombinant soybean cytosolic APX enzyme (rsAPX) appeared[43] (rsAPX has 91% sequence identity to rpAPX). The overall structure of the rsAPX enzyme is shown in Fig. 4 and was very similar to that rpAPX (r.m.s. deviation between Cα positions for the 249 residues in the A-chain of rpAPX and the refined rsAPX structure is 0.443 Å); both APX structures are similar to that of C*c*P.[69] The active site of rsAPX is shown in Fig. 2(b) and, as expected from the high sequence identity, is essentially identical to that of C*c*P. These structural similarities were discussed when the crystal structure of rpAPX first appeared[42] and are not repeated here.

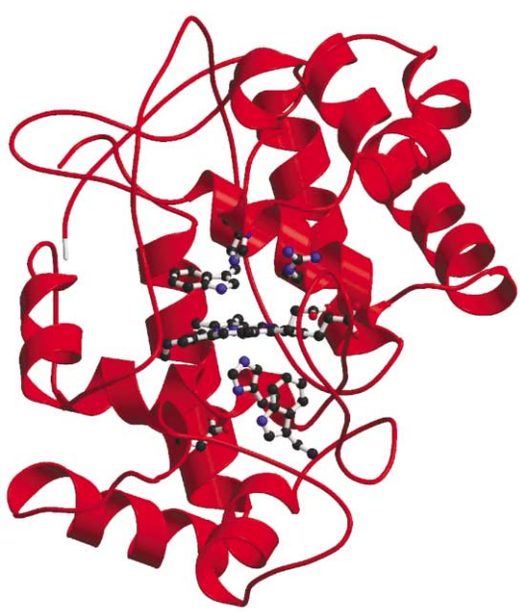


Fig. 4 The overall structure of rsAPX,[43] showing the haem and the active site residues.[105,107]

With its high sequence identity to C*c*P[60] and an almost identical three-dimensional structure,[42] APX was fully expected to duplicate the properties of the anomalous Compound I intermediate of C*c*P, to reconfirm the working hypothesis (above) and to revalidate C*c*P as the benchmark for other peroxidases. There was a very keen interest, therefore, in defining the nature of the Compound I intermediate in APX. This had not been attempted before the crystal structure was published, most probably due to low quantities of workable enzyme and the experimental difficulties associated with detection of the transient Compound I intermediate. Quite unexpectedly, however, APX was found not to duplicate the properties of C*c*P. Hence, despite containing exactly the same Trp residue (Trp179, Fig. 2(b)) that is used in C*c*P (Trp191, Fig. 2(a)) in almost exactly the same structural environment,[42] it was unambiguously demonstrated using EPR spectroscopy[70] that Compound I of rpAPX uses a porphyrin π-cation radical, Scheme 1. Spectroscopic work[71–75] from various laboratories has confirmed this original assignment for several other APXs, including rsAPX,[75] and there is complete agreement on this point.

A refinement to the original hypothesis was clearly necessary. It was proposed[70] that a potassium ion, located within 8 Å of Trp179 in rpAPX but not present in C*c*P (the residues ligating to the K^+ ion in APX – Thr164, Thr180, Asn182 and Asp187 – are replaced with Ala176, Gly192, Ala194 and Thr199, respectively in C*c*P,[65] Fig. 5), might destabilize radical cation formation at Trp179 on electrostatic grounds. Theoretical data were subsequently published in support of this hypothesis.[76] Although electrostatics may be influential, there are several reasons why we believe that this hypothesis is unlikely to provide a complete rationalisation of the results. First, we have shown[43] that the crystal structure of rsAPX contains no K^+ ion at the equivalent site and yet rsAPX also uses a porphyrin π-cation radical and not Trp179.[75] Second, we have evidence[77] that suggests that (slow) protein radical formation – most likely Trp radical formation – is possible in rpAPX if no substrate is present. Both observations argue against the K^+-destabilisation hypothesis. Finally, *ab initio* calculations[78,79] do not support a role for K^+ in defining Compound I structure. †

There was an additional complication. Site-specific removal of Trp179 in APX (Trp179Phe variant) has only a marginal effect on the ability of the enzyme to oxidise ascorbate,[85] in direct conflict with the C*c*P data which clearly indicate that Trp191 is absolutely essential for oxidation of cytochrome *c*.[86] No hypothesis had been advanced to explain this.

Development of a new hypothesis

We had suggested[64] that the key to defining these important functional differences might lie in understanding the conspicuously different substrate binding properties of the C*c*P and APX enzymes. In other words, that the properties of Compound I and the substrate binding specificity might be linked. To address this, we clearly needed information on the exact nature of the APX/ascorbate complex.

The first information on the substrate binding interaction came from NMR-derived distance restraints,[87] which indicated that the substrate was bound between 9.0 and 11.2 Å from the haem iron (NMR had been previously used in the study of substrate binding interactions in HRP, reviewed in ref. 46). These distances were consistent with two possible binding locations for the substrate: one close to close to the 6-propion-

† We note, however, that it has been possible to engineer a K^+-site, analogous to that found in APX, into C*c*P through a series of iterative mutations.[80–83] These experiments have shown that the new metal site destabilises protein radical formation at Trp191 but porphyrin π-cation formation was not observed instead. The corresponding experiment (removal of the K^+-site) has not been possible in APX because this leads to destabilisation of the active site structure.[84]

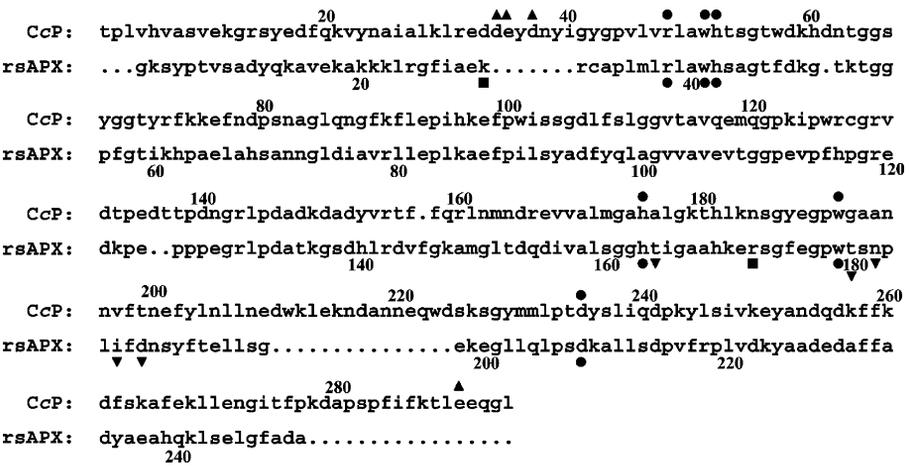


Fig. 5 Structural alignment of rsAPX and C*c*P (adapted from Veitch and Smith[46]). The numbering is as for the pea cytosolic APX enzyme.[60] The rpAPX sequence is identical to rsAPX except at positions 10 (Pro in rAPX), 16 (Ile), 21 (Arg), 31 (Lys), 36 (Ile), 49 (Ser), 50 (Lys), 63 (Gln), 69 (Gly), 84 (Ile), 86 (Glu), 87 (Gln), 91 (Val), 107 (Ile), 152 (Ser), 196 (Thr), 200 (Asp), 213 (Thr), 215 (Ser), 222 (Glu), 230 (Val), 240 (Leu) and 249 (Glu). Active site residues, Fig. 2, (●), the residues involved in binding of the potassium site in rpAPX (▼), the residues involved in binding of ascorbate in APX (■) and the residues involved in binding of cytochrome *c* in C*c*P (▲) are indicated. The additional loops in C*c*P and the C-terminal truncation in APX can be seen.

ate (γ-meso) and the other close to the δ-meso position of the haem, Fig. 1. (There is other (kinetic) evidence that APX may utilize a second binding site for ascorbate under certain conditions,[75] but the location of this second site is not known. C*c*P is also known to have more than one binding site for cytochrome *c*.[88–95]) Later, chemical modification and site-directed mutagenesis experiments[96,97] were used to show that Cys32 was close to the ascorbate site[96] and that the positively charged Arg172 residue (Arg172), located close to Cys32, was important in binding of the (anionic) ascorbate molecule.[97] These data implicated an ascorbate interaction close to Cys32 and the haem propionate groups, which was close to the γ-meso position originally implicated by NMR[87] but was not consistent with the substrate binding behaviour of most other peroxidases, which were known to use the δ-haem edge.[46,50–52,98]

In 2003, the crystal structure of the rsAPX enzyme in complex was ascorbate was published.[43] This structure confirmed the original predictions from NMR[87] and was consistent with the mutagenesis and chemical modification work.[96,97] In the absence of substrate, there are several ordered water molecules directly or indirectly hydrogen bonded to the side-chains of Arg172 and the haem 6-propionate group, Fig. 6(a), which are replaced by ascorbate in the rsAPX/ascorbate complex, Fig. 6(b). The ascorbate is bound to enzyme through hydrogen bonds between the 2-OH and 3-OH groups of the substrate, Scheme 2, and Arg172, and between the 2-OH group of the substrate and the (deprotonated) haem 6-propionate. The side chain of Lys30 swings in from solvent to provide additional hydrogen bonding stabilization (6-OH group of the substrate); this was not predicted from any functional work. Cys32 is close to the substrate binding site, but has no direct interaction with the substrate, which is consistent with the modest (≈3-fold) effect on ascorbate activity in the Cys32Ser variant.[96] The shortest distance from the substrate (oxygen of the 2-OH group) to the haem iron is 11.2 Å, in excellent agreement with the original NMR data[87] that predicted that the substrate binds between 9.0 and 11.2 Å from the iron.

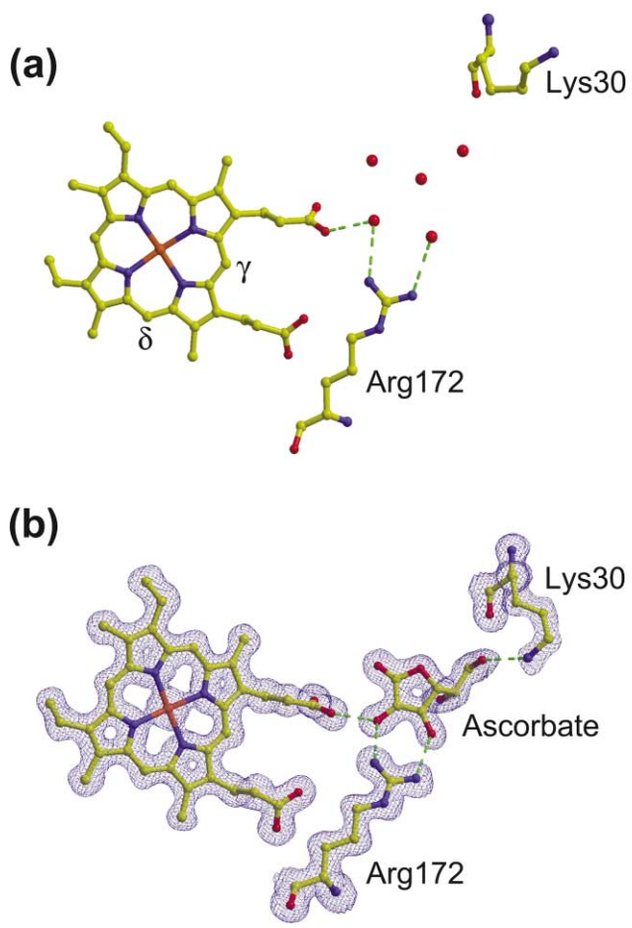


Fig. 6 (a) The structure of rsAPX showing the solvent water molecules bound in the ascorbate-binding site.[104,105] The γ-meso and δ-meso positions of the haem, Fig. 1, are indicated. (b) The rsAPX/ascorbate complex, showing refined electron density (blue) and the bound ascorbate. Hydrogen bonds are indicated (dotted lines).

Functional correlations

The rsAPX/ascorbate structure provides us with much more than just a glimpse of the ascorbate binding site. It has brought the more general question of substrate binding in other peroxidases – and how this is linked to functional diversity – into much sharper focus and has cast new light on the functional differences between the C*c*P and APX enzymes that is of relevance to our understanding of substrate binding and electron delivery in other haem enzymes. The main findings are most easily visualised with reference to Figs. 5–8.

Substrate binding

Despite high sequence identity, Fig. 5, and very similar three-dimensional structures,[29,42,43] the structural architecture of the APX and C*c*P enzymes is very subtly different and it is this that

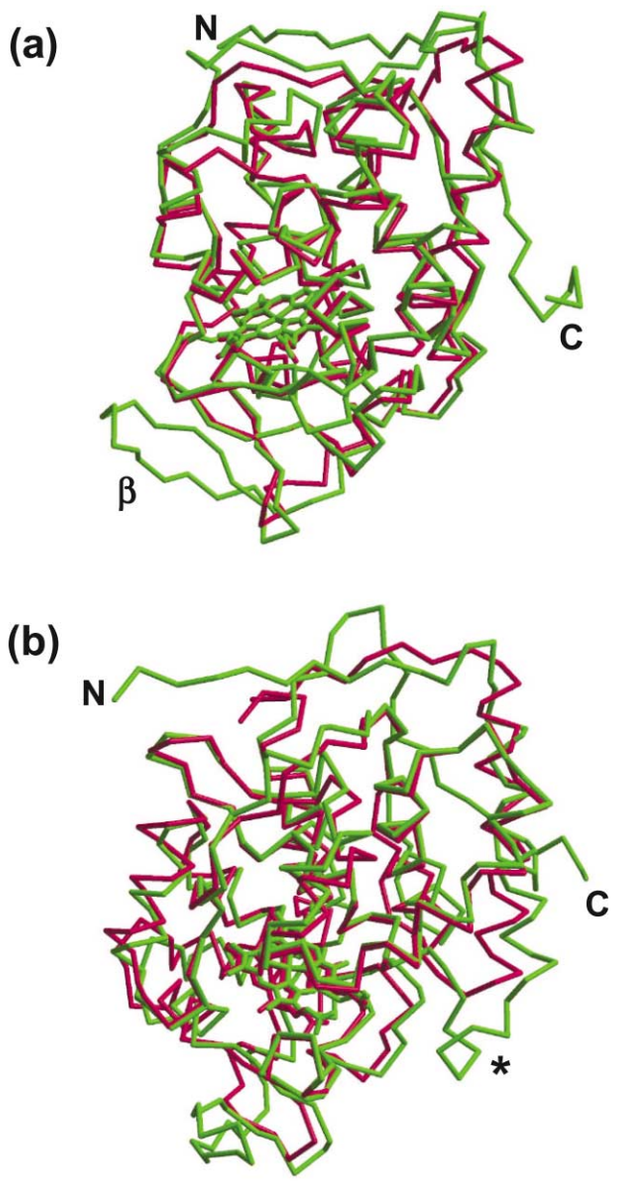


Fig. 7 Superposition of the C*c*P (green) and rsAPX (red) structures.[104,105] (a) Diagram showing the additional C-terminal tail (labelled **C**) and β-sheet structure (labelled **β**) in C*c*P; the N-terminus of C*c*P is also indicated (labelled **N**). (b) Diagram showing the additional loop (residues 34–41) in C*c*P (labelled *) that replaces Arg31 in rsAPX.

defines the substrate specificity. Hence, the key features required for substrate binding in APX (Arg172, Lys 30, Fig. 6) are not present in C*c*P (replaced by Asn184 and Asp33, Fig. 5). Similarly, the residues[47] required for substrate binding in C*c*P (Asp34, Glu35, Asp37, Glu290) are not present in APX because the loop containing residues 34–41 in C*c*P is completely missing in APX (replaced with Arg31) and because Glu290 in C*c*P is missing in APX because of a C-terminal truncation, Figs. 5 and 7. Superposition of the C*c*P and APX structures in this loop region is especially revealing and shows that the additional loop in C*c*P and the side chain of Tyr36 together prevent binding of ascorbate at the equivalent site, Fig. 8. We have proposed[43] that these differences account for the low activity of C*c*P towards ascorbate[99] and the low activity of APX towards cytochrome *c*.[63]

It is clear, therefore, that the additional loop in C*c*P (residues 34–41) and the residues in it (which were already known to be involved in binding of cytochrome *c*[47]), along with the C-terminal truncation in APX, are critical in defining substrate specificity. These differences are very easily visualised from a simple structural alignment, Fig. 5, and were even noted when

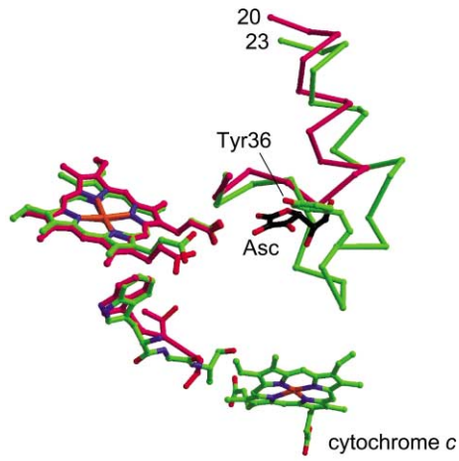


Fig. 8 Structural alignment of the C*c*P/cytochrome *c*[47] (residues 23–45, green) and APX/ascorbate[43] (residues 20–35, red) structures[104,105] showing the additional 8-residue loop in C*c*P (residues 34–41) that that replaces Arg31 in rsAPX and obscures the ascorbate binding site in rsAPX.[104,105] The side chain of Tyr36 in C*c*P is also shown (green) close to the bound ascorbate (Asc, black/red). The proposed pathway[47] for electron transfer from the haem in cytochrome *c* to C*c*P is shown (Ala193 to Trp191) superposed with the equivalent residues in rsAPX (Trp179-Ser181).

the crystal structure of rpAPX[42] first appeared. (C*c*P also contains additional β-sheet structure, residues 210–224,[42] that is missing in APX, Figs. 5 and 7(a), although we do not know whether this is functionally important.) Indeed, with the benefit of hindsight, is now much easier to predict that the loop may be important in fine-tuning substrate specificity. In practice, of course, our understanding of substrate specificity is not sufficiently sophisticated for accurate predictions of this kind to be made.

Compound I structure

As mentioned above, one of main incentives for studying APX was to establish whether it used the residue analogous to Trp191 in C*c*P during catalysis. APX was fully expected to utilize Trp179, but this has been clearly shown not to be the case.[42,71–75] The rsAPX/ascorbate structure provides some rationalization as to why not. For APX, there is direct coupling of the substrate to the haem, through the haem 6-propionate (γ-haem edge) – the substrate is 'hard-wired' directly to the haem, Fig. 6(b). Reduction of Compound I in APX therefore most sensibly occurs through a porphyrin π-cation intermediate, completely by-passing the need for involvement of Trp179. This would explain the unexpected observation that mutagenic replacement of Trp179 (Trp179Phe variant) does not affect catalytic oxidation of ascorbate by APX.[85] This contrasts with the C*c*P/cytochrome *c* interaction,[47] which has cytochrome *c* binding in an orientation that involves Trp191 as part of an electron transfer pathway to the haem in Compound I, Fig. 8. Accordingly, Trp191 of C*c*P would be expected to be essential for catalytic activity, which is indeed observed.[86]

For APX at least, it appears that it is the substrate binding orientation that defines the structure of the catalytic Compound I intermediate. If this is the case, a further question arises: is it the substrate, in the correct orientation, that precludes formation of a protein radical in Compound I of APX? There is no complete answer to this question, but this hypothesis would be consistent with experiments on rpAPX[77] in which protein radical formation has been observed in the absence of substrate (migration of the radical from the porphyrin to the the protein has also been observed in HRP[100,101]).

Electron delivery

The rsAPX/ascorbate structure also raises important questions in terms of our existing views on electron transfer in the haem

peroxidases, in which substrate binding and oxidation at the δ-haem edge, and not the γ-haem edge as seen in APX, Fig. 6, has been widely assumed.[98] In this context, we note that manganese peroxidase (a class II enzyme)[31] and NO synthase[102,103] both use the same γ-haem edge site. It remains to be seen whether that this represents a more widespread mechanism for electron delivery in other haem peroxidases or, indeed, haem enzymes more generally.

Summary

The role of the protein environment in controlling Compound I formation in the haem peroxidases has been studied extensively and a consistent mechanistic picture has emerged. Our understanding of the precise structural features that control substrate binding, and hence functional diversity, in the haem peroxidases is less well understood. In the case of APX, the sequence and structural similarity with that most famous of peroxidases, C*c*P, provides a particularly useful comparative framework that has raised questions that are relevant across the haem peroxidase family. A consensus view has not yet been reached, but the emergence of new crystallographic information for various peroxidase–substrate complexes is providing some answers and posing new questions that can subsequently be tested.

Acknowledgements

This work has been supported by the BBSRC (project grant 91/ B11469 to E. L. R.), the Wellcome Trust (grant 063688, fellowship to E. L. R.) and the EPSRC and The Institute of Applied Catalysis (studentship to K. H. S.).

References

1 C. A. I. Hutchison, S. Phillips, M. H. Edgell, S. Gillam, P. Jahnke and M. Smith, *J. Biol. Chem.*, 1978, **253**, 6551.
2 A. Razin, T. Hirose, K. Itakura and A. D. Riggs, *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 4268.
3 P. R. Ortiz de Montellano, *Acc. Chem. Res.*, 1987, **20**, 289.
4 T. L. Poulos, *Adv. Inorg. Biochem.*, 1987, **7**, 1.
5 J. H. Dawson, *Science*, 1988, **240**, 433.
6 J. Everse, K. E. Everse and M. B. Grisham, *Peroxidases in Chemistry and Biology*, CRC Press, Boca Raton, FL, 1991, vol. I and II.
7 A. G. Mauk, *Struct. Bonding (Berlin)*, 1991, **75**, 131.
8 T. L. Poulos and R. E. Fenna, in *Metal Ions in Biological Systems: Metalloenzymes Involving Amino-Acid and Related Radicals*, ed. H. Sigel and A. Sigel, Marcel Dekker, New York, 1994, p. 25.
9 B. A. Springer, S. G. Sligar, J. S. Olson and G. N. Phillips, *Chem. Rev.*, 1994, **94**, 699.
10 A. M. English and G. Tsaprailis, *Adv. Inorg. Chem.*, 1995, **43**, 79.
11 D. B. Goodin, *J. Biol. Inorg. Chem.*, 1996, **1**, 360.
12 J. E. Erman, *J. Biochem. Mol. Biol.*, 1998, **31**, 307.
13 A. T. Smith and N. C. Veitch, *Curr. Opin. Chem. Biol.*, 1998, **2**, 269.
14 H. B. Dunford, *Heme Peroxidases*, John Wiley, Chichester, 1999.
15 S. G. Sligar, *Essays Biochem.*, 1999, **34**, 71.
16 G. H. Loew and D. L. Harris, *Chem. Rev.*, 2000, **100**, 407.
17 S.-I. Ozaki, T. Matsui, M. P. Roach and Y. Watanabe, *Coord. Chem. Rev.*, 2000, **198**, 39.
18 S.-I. Ozaki, M. P. Roach, T. Matsui and Y. Watanabe, *Acc. Chem. Res.*, 2001, **34**, 818.
19 E. L. Raven and A. G. Mauk, *Adv. Inorg. Chem.*, 2001, **51**, 1.
20 A. G. Sykes and A. G. Mauk *Heme-Fe proteins*, Academic Press, London, 2001, vol. 51.
21 J. E. Erman and L. B. Vitello, *Biochim. Biophys. Acta*, 2002, **1597**, 193.
22 D. B. Goodin, A. G. Mauk and S. M., *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 1295.
23 I. S. Isaac and J. H. Dawson, *Essays Biochem.*, 1999, **34**, 51.
24 J. Everse, K. E. Everse and M. B. Grishham *Peroxidases in Chemistry and Biology*, CRC Press, Boca Raton, FL, 1991, vol. 1 and 2.
25 R. Timkovich and L. L. Bondoc, *Adv. Biophys. Chem.*, 1990, **1**, 203.
26 C. Colas and P. R. Ortiz de Montellano, *Chem. Rev.*, 2003, **103**, 2305.
27 K. G. Welinder, *Curr. Opin. Struct. Biol.*, 1992, **2**, 388.
28 T. L. Poulos, S. T. Freer, R. A. Alden, N. Xuong, S. L. Edwards, R. C. Hamlin and J. Kraut, *J. Biol. Chem.*, 1978, **253**, 3730.
29 T. L. Poulos, S. T. Freer, R. A. Alden, S. L. Edwards, U. Skogland, K. Takio, B. Eriksson, N. Xuong, T. Yonetani and J. Kraut, *J. Biol. Chem.*, 1980, **255**, 575.
30 M. Gajhede, D. J. Schuller, A. Henricksen, A. T. Smith and T. L. Poulos, *Nat. Struct. Biol.*, 1997, **4**, 1032.
31 M. Sundaramoorthy, K. Kishi, M. H. Gold and T. L. Poulos, *J. Biol. Chem.*, 1994, **269**, 32759.
32 J. Zeng and R. E. Fenna, *J. Mol. Biol.*, 1992, **226**, 185.
33 T. J. Fiedle, C. A. Davey and R. E. Fenna, *J. Biol. Chem.*, 2000, **275**, 11964.
34 A. Henricsen, K. G. Welinder and M. Gajhede, *J. Biol. Chem.*, 1998, **273**, 2241.
35 N. Kunishima, F. Amada, K. Fukuyama, M. Kawamoto, T. Matsunaga and H. Matsubara, *FEBS Lett.*, 1996, **378**, 291.
36 J. F. W. Petersen, A. Kadziola and S. Larsen, *FEBS Lett.*, 1994, **339**, 291.
37 S. L. Edwards, R. Raag, H. Wariishi, M. H. Gold and T. L. Poulos, *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 750.
38 K. Piontek, T. Glumoff and K. H. Winterhalter, *FEBS Lett.*, 1993, **315**, 119.
39 T. L. Poulos, S. L. Edwards, H. Wariishi and M. H. Gold, *J. Biol. Chem.*, 1993, **268**, 4429.
40 T. Choinowski, W. Blodig, K. H. Winterhalter and K. Piontek, *J. Mol. Biol.*, 1999, **286**, 809.
41 D. J. Schuller, N. Ban, R. B. Van Huystee, A. McPherson and T. L. Poulos, *Structure*, 1996, **4**, 311.
42 W. R. Patterson and T. L. Poulos, *Biochemistry*, 1995, **34**, 4331.
43 K. H. Sharp, M. Mewies, P. C. E. Moody and E. L. Raven, *Nat. Struct. Biol.*, 2003, **10**, 303.
44 Y. Yamada, T. Fujiwara, T. Sato, N. Igarashi and N. Tanaka, *Nat. Struct. Biol.*, 2002, **9**, 691.
45 M. Sundaramoorthy, J. Terner and T. L. Poulos, *Structure*, 1995, **3**, 1367.
46 N. C. Veitch and A. T. Smith, *Adv. Inorg. Chem.*, 2001, **51**, 107.
47 H. Pelletier and J. Kraut, *Science*, 1992, **258**, 1748.
48 C. A. Davey and R. E. Fenna, *Biochemistry*, 1996, **35**, 10967.
49 K. Fukuyama, K. Sato, H. Itakura, S. Takahashi and T. Hosoya, *J. Biol. Chem.*, 1997, **272**, 5752.
50 H. Itakura, Y. Oda and K. Fukuyama, *FEBS Lett.*, 1997, **412**, 107.
51 A. Henriksen, D. J. Schuller, K. Meno, K. G. Welinder, A. T. Smith and M. Gajhede, *Biochemistry*, 1998, **37**, 8054.
52 A. Henriksen, A. T. Smith and M. Gajhede, *J. Biol. Chem.*, 1999, **274**, 35005.
53 K. Meno, S. Jennings, A. T. Smith, A. Henricksen and M. Gajhede, *Acta Crystallogr., Sect. D*, 2002, **58**, 1803.
54 L. A. Fishel, J. E. Villafranca, J. M. Mauro and J. Kraut, *Biochemistry*, 1987, **26**, 351.
55 J. E. Erman, L. B. Vitello, J. M. Mauro and J. Kraut, *Biochemistry*, 1989, **28**, 7992.
56 C. P. Scholes, Y. Liu, L. A. Fishel, M. F. Farnum, J. M. Mauro and J. Kraut, *Isr. J. Chem.*, 1989, **29**, 85.
57 M. Sivaraja, D. B. Goodin, M. Smith and B. M. Hoffman, *Science*, 1989, **245**, 738.
58 A. L. Houseman, P. E. Doan, D. B. Goodin and B. M. Hoffman, *Biochemistry*, 1993, **32**, 4430.
59 J. E. Huyett, P. E. Doan, R. Gurbiel, A. L. P. Houseman, M. Sivaraja, D. B. Goodin and B. M. Hoffman, *J. Am. Chem. Soc.*, 1995, **117**, 9033.
60 R. Mittler and B. A. Zilinskas, *FEBS Lett.*, 1991, **289**, 257.
61 D. Groden and E. Beck, *Biochim. Biophys. Acta*, 1979, **546**, 426.
62 G. J. Kelly and E. Latzko, *Naturewissenschaften*, 1979, **66**, 617.
63 D. A. Dalton, in *Peroxidases in Chemistry and Biology*, ed. J. Everse, K. E. Everse and M. B. Grisham, CRC Press, Boca Raton, FL, 1991, p. 139.
64 E. L. Raven, in *Subcellular Biochemistry: Enzyme Catalysed Electron and Radical Transfer*, ed. A. Holzenberg and N. S. Scrutton, Kluwer Academic/Plenum Publishers, New York, 2000, p. 318.
65 E. L. Raven, *Nat. Prod. Rep.*, 2003, **20**, 367.
66 H. A. Heering, M. A. K. Jansen, R. N. F. Thorneley and G. Smulevich, *Biochemistry*, 2001, **40**, 10360.
67 W. R. Patterson and T. L. Poulos, *J. Biol. Chem.*, 1994, **269**, 17020.
68 D. A. Dalton, L. Diaz del Castillo, M. L. Kahn, S. L. Joyner and J. M. Chatfield, *Arch. Biochem. Biophys.*, 1996, **328**, 1.
69 B. C. Finzel, T. L. Poulos and J. Kraut, *J. Biol. Chem.*, 1984, **259**, 13027.
70 W. R. Patterson, T. L. Poulos and D. B. Goodin, *Biochemistry*, 1995, **34**, 4342.
71 L. A. Marquez, M. Quitoriano, B. A. Zilinskas and H. B. Dunford, *FEBS Lett.*, 1996, **389**, 153.

72 M. Kvaratskhelia, C. Winkel and R. N. F. Thorneley, *Plant Physiol.*, 1997, **114**, 1237.
73 M. Kvaratskhelia, C. Winkel, M. T. Naldrett and R. N. F. Thorneley, *J. Plant Physiol.*, 1999, **154**, 273.
74 L. Lad, M. Mewies, J. Basran, N. S. Scrutton and E. L. Raven, *Eur. J. Biochem.*, 2002, **369**, 3182.
75 L. Lad, M. Mewies and E. L. Raven, *Biochemistry*, 2002, **41**, 13774.
76 G. M. Jensen, *J. Phys. Chem.*, 1998, **102**, 8221.
77 A. N. P. Hiner, J. I. Martínez, M. B. Arnao, M. Acosta, D. D. Turner, E. L. Raven and J. N. Rodríguez-López, *Eur. J. Biochem.*, 2001, **268**, 3091.
78 G. Naray-Szabo, *J. Biol. Inorg. Chem.*, 1997, **2**, 135.
79 D. K. Menyhard and G. Naray-Szabo, *J. Phys. Chem.*, 1999, **103**, 227.
80 C. A. Bonagura, M. Sundaramoorthy, H. S. Pappa, W. R. Patterson and T. L. Poulos, *Biochemistry*, 1996, **35**, 6107.
81 C. A. Bonagura, M. Sundaramoorthy, B. Bhaskar and P. T. L, *Biochemistry*, 1999, **38**, 5538.
82 B. Bhaskar, C. A. Bonagura, H. Li and T. L. Poulos, *Biochemistry*, 2002, **41**, 2684.
83 C. A. Bonagura, B. Bhaskar, M. Sundaramoorthy and T. L. Poulos, *J. Biol. Chem.*, 1999, **274**, 37827.
84 J. Cheek, D. Mandelman, T. L. Poulos and J. H. Dawson, *J. Biol. Inorg. Chem.*, 1999, **4**, 64.
85 H. Pappa, W. R. Patterson and T. L. Poulos, *J. Biol. Inorg. Chem.*, 1996, **1**, 61.
86 J. M. Mauro, L. A. Fishel, J. T. Hazzard, T. E. Meyer, G. Tollin, M. A. Cusanovich and J. Kraut, *Biochemistry*, 1988, **27**, 6243.
87 A. P. Hill, S. Modi, M. J. Sutcliffe, D. D. Turner, D. J. Gilfoyle, A. T. Smith, B. M. Tam and E. Lloyd, *Eur. J. Biochem.*, 1997, **248**, 347.
88 S. H. Northrup, J. O. Boles and J. C. L. Reynolds, *Science*, 1988, **241**, 67.
89 E. D. A. Stemp and B. M. Hoffman, *Biochemistry*, 1993, **32**, 10848.
90 M. R. Mauk, J. C. Ferrer and A. G. Mauk, *Biochemistry*, 1994, **33**, 12609.
91 J. S. Zhou, J. M. Nocek, M. L. DeVan and B. M. Hoffman, *Science*, 1995, **269**, 204.
92 H. Mei, K. Wang, S. McKee, X. Wang, J. L. Waldner, G. J. Pielak, B. Durham and F. Millett, *Biochemistry*, 1996, **35**, 15800.
93 J. S. Zhou, S. T. Tran, G. McLendon and B. M. Hoffman, *J. Am. Chem. Soc.*, 1997, **119**, 269.
94 V. W. Leesch, J. Bujons, A. G. Mauk and B. M. Hoffman, *Biochemistry*, 2000, **39**, 10132.
95 H. Mei, L. Geren, M. A. Miller, B. Durham and F. Millett, *Biochemistry*, 2002, **41**, 3968.
96 D. Mandelman, J. Jamal and T. L. Poulos, *Biochemistry*, 1998, **37**, 17610.
97 E. H. Bursey and T. L. Poulos, *Biochemistry*, 2000, **39**, 7374.
98 M. A. Ator and P. R. Ortiz de Montellano, *J. Biol. Chem.*, 1987, **262**, 1542.
99 T. Yonetani and G. S. Ray, *J. Biol. Chem.*, 1965, **240**, 4503.
100 V. P. Miller, D. B. Goodin, A. E. Friedman, C. Hartmann and P. R. Ortiz de Montellano, *J. Biol. Chem.*, 1995, **270**, 18413.
101 A. Morimoto, M. Tanaka, S. Takahashi, K. Ishimori, H. Hori and I Morishima, *J. Biol. Chem.*, 1998, **273**, 14753.
102 T. L. Poulos, H. Li, C. S. Raman and D. J. Schuller, *Adv. Inorg. Chem.*, 2001, **51**, 243.
103 C.-C. Wei, B. R. Crane and D. J. Stuehr, *Chem. Rev.*, 2003, **103**, 2365.
104 D. McRee, *J. Mol. Graph.*, 1992, **10**, 44.
105 E. A. Merrit and M. E. P. Murphy, *Acta Crystallogr., Sect. D*, 1994, **50**, 869.
106 D. M. Bryan, S. D. Dell, R. Kumar, M. J. Clarke, V. Rodriguez, M. Sherban and J. Charkoudian, *J. Am. Chem. Soc.*, 1988, **110**, 1498.
107 P. J. Kraulis, *J. Appl. Crystallogr.*, 1991, **24**, 946.