ELSEVIER

FEBS Letters 339 (1994) 291–296

FEBS 13650

FEBS LETTERS

Three-dimensional structure of a recombinant peroxidase from *Coprinus cinereus* at 2.6 Å resolution

Jens F.W. Petersen*, Anders Kadziola, Sine Larsen

Centre for Crystallographic Studies, Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark

Received 6 December 1993; revised version received 4 January 1994; accepted 4 January 1994

Abstract

The structure of a recombinant peroxidase from the ink cap, *Coprinus cinereus*, CiP, is reported to 2.6 Å resolution and refined to a *R*-value of 18.1%. The structure was solved by molecular replacement using the coordinates from a newly published ligninase structure, LiP. CiP crystallizes in space group $P2_12_12_1$ with two independent molecules in the asymmetric unit related by the vector $0.29\vec{b} + 0.5\vec{c}$. The two CiP molecules are structurally identical; each contains two Ca^{2+} ions in positions equivalent to those found in the LiP structure. Two *N*-acetylglucosamines and one mannose residue were fitted into the density adjacent to two of the three predicted glycosylation sites. The refinement also included 40 and 41 water molecules, respectively, in the two CiP molecules. The structure of CiP displays a folding very similar to that of LiP. The active sites are almost identical in the LiP and CiP structures. CiP has a much larger opening to the active site than LiP.

Key words: Peroxidase; *Coprinus cinereus*; Crystal structure; Recombinant

1. Introduction

Only one glycosylated peroxidase, CiP, has been isolated from the ink cap basidiomycete, *Coprinus cinereus* [1]. From the alignment of amino acid sequences Welinder showed that this peroxidase belongs to the plant peroxidase superfamily (class II) [2]. Baunsgaard et al. found that the cDNA sequence of CiP is identical to the amino acid sequence of *Coprinus macrorhizus* peroxidase, CMP. Both have a single chain of 343 amino acids and the N-terminal blocked by pyroglutamate [3]. It has been shown that the peroxidase isolated from *Arthomyces ramosus*, ARP, has enzymatic and structural properties [4] identical to those of the CiP and CMP peroxidases. The three peroxidases, CiP, CMP and ARP, appear to vary only in their degree of glycosylation.

CiP shows 40–45% sequence identity to the lignin degrading peroxidases isolated from the white-rot fungi, *Phanerochaete chrysosporium*, but CiP is unable to degrade lignin. In its enzymatic specificity CiP resembles the classical plant peroxidase from horseradish, (HRP-C), with which it only shares 18% sequence identity [2]. This makes CiP an attractive candidate for the study of the relationship between structure and function of peroxidases.

Yeast cytochrome *c* peroxidase, CCP, has long been the only example of a peroxidase structure known at high resolution [5], but recently two independently performed structure determinations have been published for the ligninase isozyme from *Phanerochaete chrysosporium*, LiP [6,7]. The two structure determinations were performed for crystals obtained with different precipitating agents in different space groups.

A preliminary X-ray crystallographic characterization of ARP has recently been published by Kunishima et al. [8]. The crystals of ARP are prepared with $(NH_4)_2SO_4$ as the precipitating agent and have one molecule per asymmetric unit. We have previously described the crystallization conditions and X-ray characterization of CiP crystals [9]. They were crystallized with PEG 6K and have two CiP molecules per asymmetric unit. We present here the three dimensional structure of CiP to a resolution of 2.6 Å and make a comparison of its structure with the structures of other known peroxidase.

2. Experimental

2.1. Precipitation and crystallographic characterization

Crystals were obtained with highly purified recombinant CiP as previously reported [9]. The crystal growth conditions were 18% w/v PEG 6K, 0.35 M $MgCl_2$, buffered at pH 7.0 with 0.1 M HEPES. CiP crystallizes in a primitive orthorhombic cell with unit cell dimensions a = 127.4 Å, b = 75.4 Å and c = 76.7 Å. Assuming 2 CiP molecules in the asymmetric unit V_m is 2.8 Å^3/Da according to Matthews [10]. The systematically absent reflections indicated $P2_12_12_1$ or $P2_12_12$ as possible space groups. The extinctions for 00ℓ: $\ell = 2n + 1$ could be a consequence of the observation that the $h0\ell$ reflections were absent or systematically very weak for $\ell = 2n + 1$. The apparent extinctions in the $h0\ell$ zone were interpreted as originating from pseudo translational symmetry relating the two crystallographically independent molecules. An inspection of

*Corresponding author. Fax: (45) 3532 0299.
E-mail: Jens@xray.ki.ku.dk

SSDI 0014-5793(94)00034-S

the native Patterson function confirmed this interpretation and showed that the two independent molecules are connected by the vector $0.29\bar{b} + 0.5\bar{c}$. In the space group $P2_12_12_1$ this pseudo translational symmetry gives rise to a non-crystallographic two-fold axis parallel to the c-axis. Alternatively it corresponds to the presence of a twofold screw axis parallel to the c-axis, if $P2_12_12$ were the correct space group.

2.2. Data collection and processing

CuKα X-ray data were collected on a Rigaku R-AXIS II imaging plate system. The X-ray source was a Rigaku Rotaflex RU 200 rotating anode generator operated at 50 kV and 180 mA, with a normal focus, a graphite monochromator and a 0.5 mm collimator. The area detector was set at 14 cm from the crystal with $2\theta = 0°$. Diffraction data were collected on a series of frames each comprising a 2.5° oscillation around the ω-axis. The exposure time was 30 min for each frame. The data were processed with software supplied with the instrument. The data set was 91.6% complete to 2.6 Å and the corresponding R_{sym} defined as

$$R_{sym} = \frac{\sum_{i,\bar{h}} |I_i(\bar{h}) - \langle I(\bar{h})\rangle|}{\sum_{i,\bar{h}} I_i(\bar{h})}$$

was 6.3% where i runs over all symmetry-related reflections.

2.3. Structure determination and refinement

The structure was solved by molecular replacement using the LiP structure as a search model. The coordinates for LiP were generously provided by Dr. Klaus Piontek, ETH, Zurich, Switzerland. The programs Merlot [11] and AMORE [12,13] were used. In each case only one prominent peak in the rotation function was seen, as expected from the pseudo symmetry that relates the two independent molecules. The translation search was conducted with AMORE in both space groups $P2_12_12$ and $P2_12_12_1$ in two steps with a monomer LiP as a search molecule. The top translational solution was fixed and a second monomer translational search was performed including the previously 'located' monomer. A solution where the two search molecules were related by approximately $0.29\underline{b} + 0.5\underline{c}$ was found in both $P2_12_12$ and $P2_12_12_1$. However, the corresponding crystal packing suffered from significant overlap in both space groups.

A one step translation search with a dimer (two 'correctly' oriented molecules related by the $0.29\bar{b} + 0.5\bar{c}$ vector) was more successful. A significant peak was found in the translation function in both space groups $P2_12_12_1$ and $P2_12_12$, but only in the space group $P2_12_12_1$ was the solution without any severe overlap in the crystal packing. As an additional check of this solution a rigid body refinement was performed with XPLOR [14] where the two molecules were treated independently. This resulted in shifts of up to 2 Å and 2°of the molecules relative to each other. As a check of the molecular replacement solution the iron atoms were removed from the haem groups prior to the refinement. In the resulting difference Fourier map strong peaks were found in the center of the haem groups.

After the rigid body refinement the two molecules were refined independently without any modifications using simulated annealing with XPLOR [14], including the data from 10 to 2.8 Å. The CiP model was built after this round according to the alignment of Baunsgaard et al. [3] and fitted to the density where it was possible. Turbo-Frodo [15] was used for the model building. Residues 9–343 were included in the model for both molecules. In the refinement of the structure each cycle of simulated annealing was followed by cycles of restrained individual *B*-factor refinement. Firstly one round was performed with data from 10–2.8 Å followed by 6 rounds of refinement and consecutive model fitting including data from 10–2.6 Å. Additional difference density appeared in the N-terminal region during the refinement and residue 8 could be fitted to this density in both molecules. Similarly the density close to Asn^{142} and Ser^{338} could be assigned to two *N*-acetylglucosamine (NAG) residues (Fig. 6) bonded to Asn^{142} and one mannose residue bonded to Ser^{338}. Two structurally bound calcium ions and one magnesium ion on the non-crystallographic twofold axis, that relate the two CiP molecules were included in the model during the refinement. Peaks in the density map were included as water molecules if they fulfilled the criteria: density in the F_o-F_c map above 2σ, density in the $2F_o$-F_c map above 1σ, chemically reasonable contact distances to potential hydrogen bond partners and a *B*-value after refinement less than 50 Å^2. This lead to the inclusion of 40 and 41 water molecules, respectively, in the two CiP molecules. The Fe atoms in the haem groups were refined without any restraints.

The final *R*-factor residual $\Sigma||F_o| - |F_c||/\Sigma|F_o|$ based on data between 10 and 2.6 Å above 1σ is 18.1%. Based on a 10% test data set the corresponding R_{free}-value [16] is 27.2%. The average *B*-value is 15.6 Å^2 and the rms differences of bond lengths and angles from ideal geometry are 0.016 Å and 3.5°, respectively.

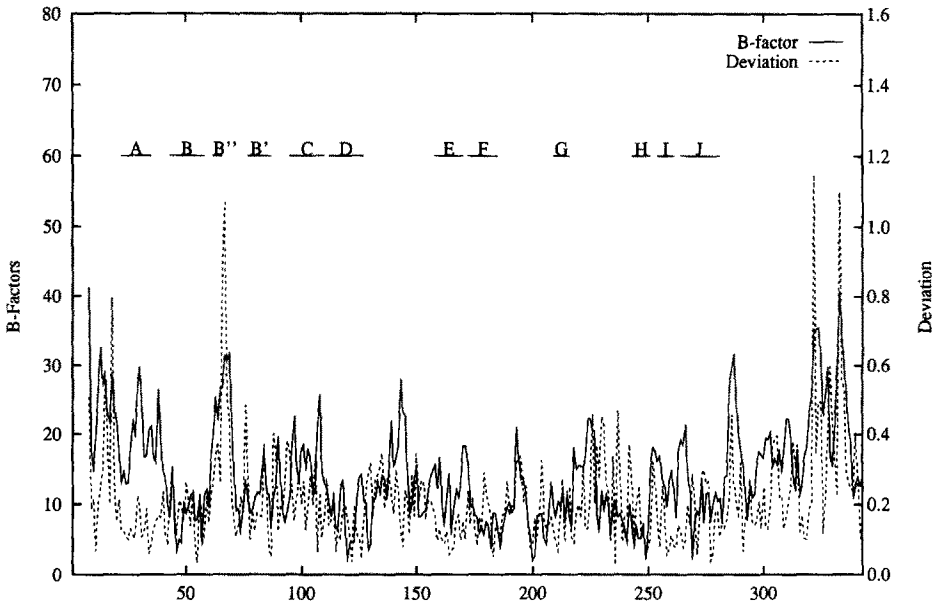

Fig. 1. The deviation (Å) of Cα's positions between the two independent molecules in an asymmetric unit after superposition. The average *B*-factor for Cα atoms in unit of Å^2 as a function of residue number is also shown. The letters indicate helical regions according to the definition of Kabsch and Sander [18]. The B'' helix, which is found in a loop region, was not identified in the LiP structure.

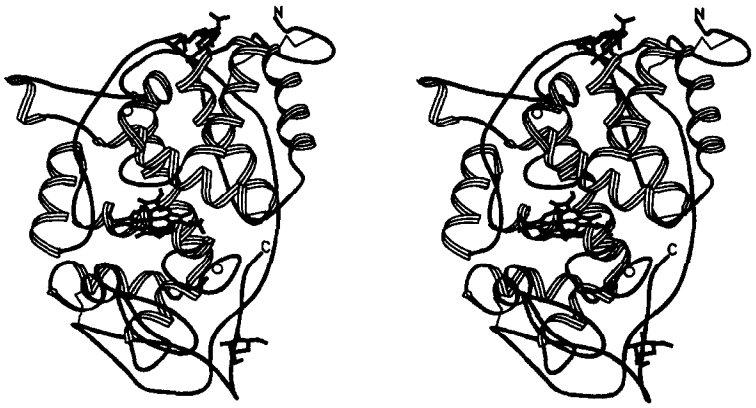

Fig. 2. The overall fold of the CiP molecule, the haem group, Ca^{2+} ions and bound carbohydrate are also shown. Ribbon drawing from Setor [19].

3. Results and discussion

3.1. Quality of the model

Each independent CiP molecule consists of residues 8–343, a haem group, two calcium ions, two covalently bound sugar moieties: two *N*-acetylglucosamines and one mannose. In addition the asymmetric unit contains 81 water molecules. We were unable to assign any density to the first N-terminal seven residues. It seems likely that this region exhibits conformational disorder due to the four glycines that constitute residues 4–7.

The presence of two molecules in the asymmetric unit provides an independent estimate of the accuracy of the fit to the density. After the two molecules have been superimposed the deviation for each Cα atom is shown in Fig. 1 as a function of residue number. The rms difference of 0.25 Å, is comparable to the average rms of 0.34 Å that Poulos et al. [7] found in the analysis of the two independent LiP molecules. The average thermal *B*-factors as a function of residue number are displayed in the same diagram. Regions with large *B*-factors are also regions with large deviations between the molecules. They include the N- and C-termini and residues 60–70 which are within a loop region. Only Gly residues are found in the 'forbidden' regions of the Ramachandran plot.

3.2. Overall description of the molecule

The overall folding of CiP is illustrated in Fig. 2. It is largely helical and has the same topology as CCP, therefore we employ the nomenclature used for the CCP and LiP helices [5,7]. The only difference in the helical topology between CiP and CCP is the two additional helices, B″ and B′, between B and C in CiP. Like CCP and LiP, CiP can be described as consisting of two domains named after the residues adjacent to the haem group. The distal domain contains the helices A, B, B′,B″, C, D and E, the helices F, J, H, I and G are part of the proximal domain. The J helix connects the two domains and continues in a large loop, which extends over the top of the distal domain and goes back to the proximal domain, where the C-terminus in CCP is located. Relative to CCP both CiP and LiP have an extension of the C-terminal end of about 50 residues; as shown in Fig. 2 this is a region where significant differences can be observed

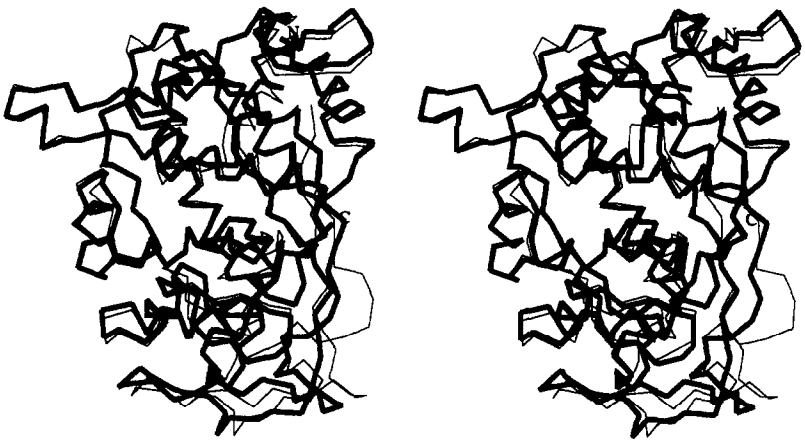

Fig. 3. The backbone of LiP (thin line) superimposed on the CiP backbone (thick line) drawn with TURBO FRODO [15].

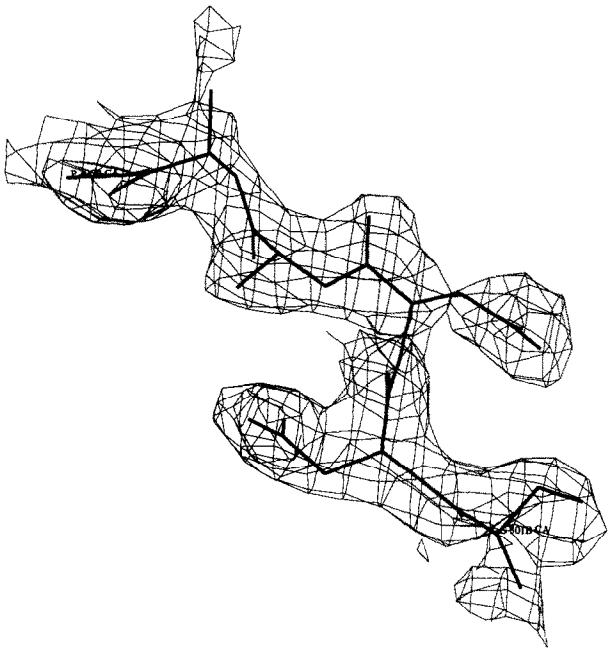

Fig. 4. Simulated annealed $2F_o$-F_c electron density omit map of the region around residue 303

between the CiP and LiP structures. In LiP the C-terminus extends along the distal domain, whereas in CiP it ends between the proximal and distal domain.

The folding of CiP is almost identical to the folding of LiP. The stereo pair in Fig. 3 illustrates a superimposition of their backbones, which shows that the folding of the CiP and LiP structures differ only in the loop and C-terminal regions. The rms difference between Cα atoms in CiP and LiP is 0.95 Å. Alignment of the CiP and LiP sequences shows two gaps in the CiP sequence relative to LiP [3]. The first gap is at residue number 170 and the second at residue 300. At residues 300–304 there are significant differences between the two structures. It has been found that this region is susceptible to proteolytic cleavage between residues 303 and 304 in CiP [4]. A simulated annealed omit map was calculated in this region (Fig. 4). All the density, accounting for both main and side chain, was clearly seen in this map.

3.3. Carbohydrate

The two independent molecules have an identical degree of glycosylation. One glycosylation site is at Asn^{142}, where the sequence Asn-Ser-Ser is found. This is the consensus sequence for a N-linked glycan. Kjalke et al. [4] found a N-linked glycan in CiP at Asn^{142}. A very clear F_o-F_c density appeared at this site as shown in Fig. 5. It was possible to build two NAGs into this density, and no additional density appeared after refinement. The average *B*-value of the last NAG residue is rather high (40 $Å^2$), indicating that if other carbohydrate residues are present they are disordered.

Baunsgaard et al. [3] have proposed that Thr^{331} and Ser^{338} should be O-linked to a single mannose residue in CMP. At these sites difference density appeared in both molecules. The density at Ser^{338} could be fitted with one mannose residue. No additional density appeared at this site after refinement, but the high average *B*-value (40 $Å^2$) could indicate that this sugar residue is slightly disordered. The density at Thr^{331} was very weak, so we did not attempt to fit any mannose residue into this density.

3.4. Ca^{2+}-binding sites

Like the LiP structure [6,7], CiP contains two structurally bound Ca^{2+} ions. One is found in the proximal and one in the distal domain as shown in Fig. 2. The ligands to the Ca^{2+} ion in the proximal site are eight oxygen atoms from main and side chains of the residues: Ser^{184}

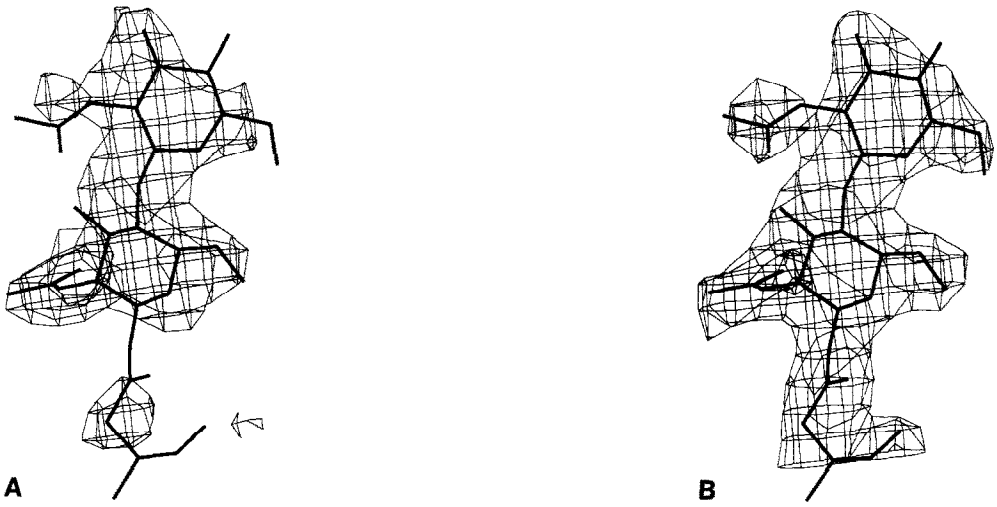

Fig. 5. (A) F_o-F_c electron density map at Asn^{142} before the two *N*-acetylglucosamines (NAG) were fitted into the model. (B) The $2F_o$-F_c electron density at Asn^{142} after two NAGs were built into the model.

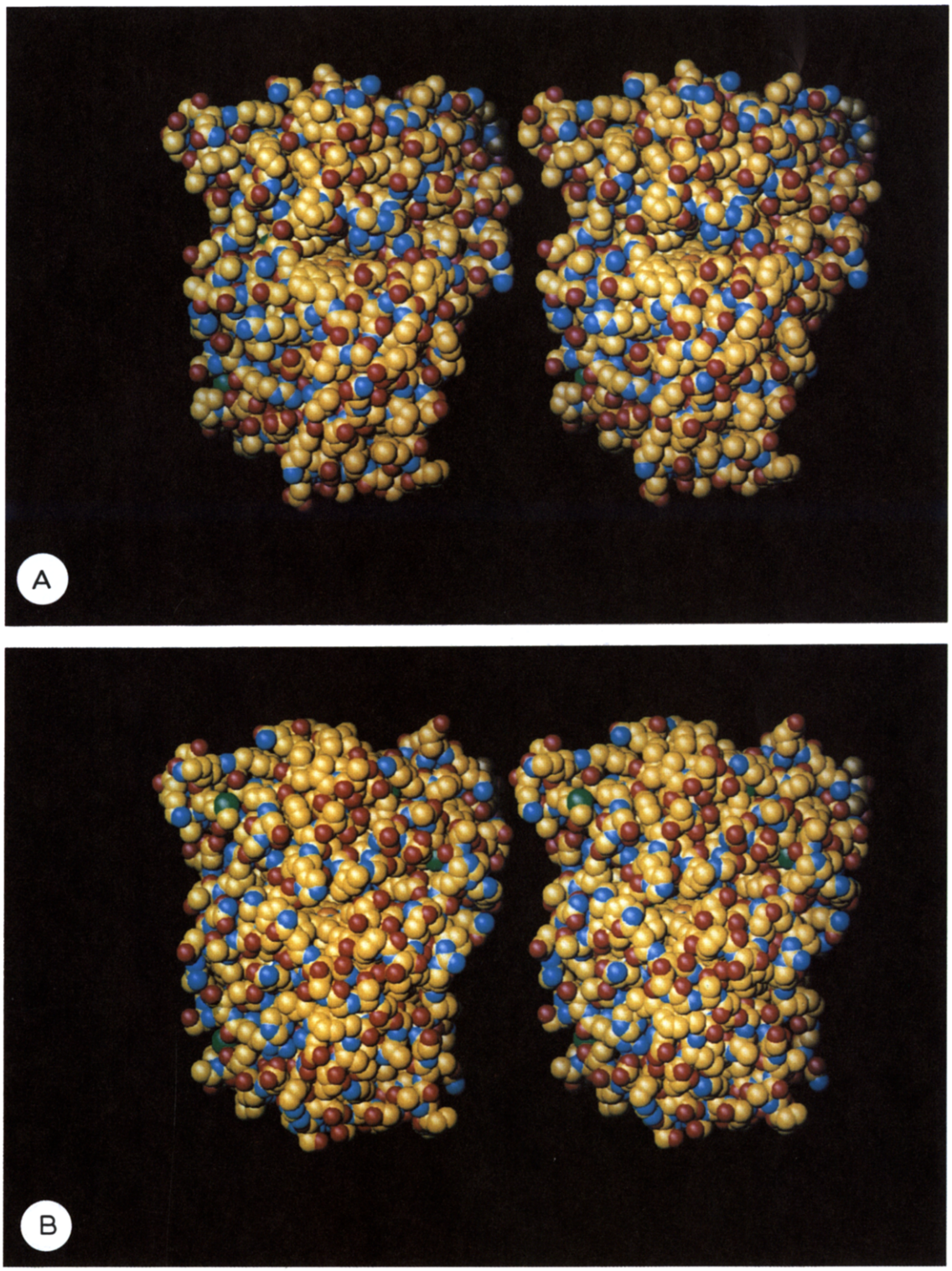

Fig. 6. Space filling models [15] of (A) CiP and (B) LiP.

(O + OD1), Asp^{201} (OD1 + OD2), Thr^{203} (O + OG1), Val^{206} (O) and Asp^{208} (OD1). These ligands are equivalent to those found for the proximal Ca^{2+} ion in the LiP structure [7]. The Ca–O distances in the CiP structure are between 2.2 and 2.8 Å, the equivalent distances in the LiP structure are in the range 2.3–2.9 Å. The distal Ca^{2+} sites are also similar in the CiP and LiP structures, Ca^{2+} is heptacoordinated with five ligands from the main and side chain residues: Asp^{56} (O + OD1), Gly^{74} (O), Asp^{76} (OD2) and Ser^{78} (O6) and two water molecules. Though the density corresponding to two water molecules is quite clear it was so weak that we did not include them in the

refinement at this stage. The distal Ca^{2+} ligand distances in CiP are between 2.1 and 2.9 Å. They vary more than those in LiP, which are between 2.3 and 2.6 Å.

3.5. Active site and haem environment

The difference in substrate specificity must be related to differences in the three dimensional structure of the two enzymes of LiP and CiP. The catalytic activity of peroxidases is associated with the haem group and conserved catalytic residues in the haem pocket. In CiP numbering scheme the catalytically important amino acids are: Arg^{51}, Phe^{54}, His^{55}, His^{183} and Asp^{245}, where His^{183} (proximal His) coordinates to the iron in the haem group. In addition Asn^{92} forms a bridge between His^{55} and main chain carbonyl O from Glu^{86} via hydrogen bonds. This bridge may also be important for the catalysis [7]. The active sites of CiP and LiP are remarkably similar with respect to the position, orientation and hydrogen bond pattern for all these residues. Fe appears to be pentacoordinated as it was observed in LiP. The distance to a possible sixth ligand, a water molecule is so long that we do not consider it to be coordinating.

CiP and LiP differ in their substrate specificity. Considering that the surroundings of the haem groups are similar, the difference in substrate specificity could be related to differences in their active site channels. Fig. 6 shows space filling models of CiP and LiP. The two molecules are oriented in an identical way, so a good view can be obtained of the opening to the active site. These photographs demonstrate clearly that there is a great difference between the CiP and LiP structures with respect to the opening into their active sites. As shown in Fig. 2 the two enzymes have virtually identical folding of the polypeptide backbone in this region.

CiP has a much larger opening to the active site than LiP. Considering that lignin (a large complex molecule) is a substrate for LiP but not for CiP, we find that this difference may support the view that LiP oxidizes lignin indirectly via small molecule mediators. The larger opening to the active site in CiP compared to LiP is caused exclusively by differences in the side chains: the His^{82} in LiP corresponding to Pro^{90} in CiP is turned so that it closes parts of the opening and the Gly's in CiP (93, 154, 156 and 190) are replaced with the bulkier residues in LiP (Ile^{85}, Glu^{146}, Phe^{148} and Asp^{183}). Through kinetic investigation and structural modifications of the haem group in horseradish peroxidase isozyme *c*, Ator et al. [17] provided evidence that different substrates interact with different sites of the haem edge. The large opening to the active site in CiP makes the haem group accessible to substrate binding at different sites of the haem edge, which may explain the broad substrate specificity of CiP.

Acknowledgements: We express our gratitude to Dr. Jeppe Tams for purification of the enzyme. All gene technology work was carried out at Novo Nordisk A/S and we thank cand. scient. Lone Baunsgaard and Dr. Jesper Vind for the cDNA clone and Dr. Ejner Bech Jensen for fermentation of transformed *Aspergillus*. This work was supported by the Danish Technology Council, Grant 133/001–90.0090. The Centre for Crystallographic Studies is supported by The Danish National Research Foundation.

References

[1] Morita, Y., Yamashita, H., Mikami, B., Iwamato, H., Aibara, S., Terada, M. and Minami, J. (1988) J Biochem. 103, 693–699.
[2] Welinder, K.G. (1992) Curr. Opin. Struct. Biol. 2, 388–393.
[3] Baunsgaard, L., Dalbøge, H., Houen, G., Rasmussen, E M. and Welinder, K.G. (1993) Eur. J. Biochem. 213, 605–611.
[4] Kjalke, M., Andersen, M.B., Schneider, P., Christensen, B., Schülein, M. and Welinder, K.G. (1992) Biochim. et Biophys. Acta 1120, 248–256.
[5] Finzel, B.C., Poulos, T.L. and Kraut, J. (1984) J. Biol. Chem. 259, 13027–13036.
[6] Piontek, K., Glumoff, T. and Winterhalter K. (1993) FEBS Lett. 315, 119–124.
[7] Poulos, T.L., Edwards, S.L., Wanishi, H. and Gold, M.H. (1993) J Biol. Chem. 268, 4429–4440.
[8] Kunishima, N., Fukuyama, K., Wakabayashi, S., Sumida, M., Takaya, M., Shibano, Y., Amachi, T. and Matsubara, H. (1992) Proteins 15, 216–220.
[9] Petersen, J.F.W., Tams, J.W., Vind, J., Svensson, A., Dalbøge, H., Welinder, K.G. and Larsen, S. (1993) J. Mol. Biol. 232, 989–991.
[10] Matthews, B.W. (1968) J. Mol. Biol. 33, 491–497.
[11] Fitzgerald, P.M.D. (1988) J. Appl. Crystallogr 21, 273–278
[12] Navaza, J. (1987) Acta Crystallogr. A43, 645–650.
[13] Navaza, J. (1990) Acta Crystallogr. A46, 619–620.
[14] Brünger, A.T. (1992) X-PLOR Version 3.0 (Yale University, New Haven).
[15] Roussel, A. and Cambillau, C. (1989) Turbo-Frodo, Biographics, LCCMB-CNRS, Bvd. Pierre Dranard, F-13916 Marseille Cedex 20.
[16] Brünger, A.T. (1992) Nature 355, 4722–475.
[17] Ator, M.A. and Ortiz de Montellano, P.R. (1987) J. Biol. Chem 262, 1542–1551.
[18] Kabsch, W. and Sander, C. (1983) Biopolymers 22, 2577–2637.
[19] Evans, S.V. (1993) J. Mol. Graphics 11, 134–138.