Structure of human erythrocyte catalase

Tzu-Ping Ko, Martin K. Safo, Faik N. Musayev, Martino L. Di Salvo, Changqing Wang, Shih-Hsiung Wu and Donald J. Abraham
Structure of human erythrocyte catalase

Catalase (E.C. 1.11.1.6) was purified from human erythrocytes and crystallized in three different forms: orthorhombic, hexagonal and tetragonal. The structure of the orthorhombic crystal form of human erythrocyte catalase (HEC), with space group $P2_12_12_1$ and unit-cell parameters $a = 84.9$, $b = 141.7$, $c = 232.5$ Å, was determined and refined with 2.75 Å resolution data. Non-crystallographic symmetry restraints were employed and the resulting $R$ value and $R_{	ext{free}}$ were 0.206 and 0.272, respectively. The overall structure and arrangement of HEC molecules in the orthorhombic unit cell were very similar to those of bovine liver catalase (BLC). However, no NADPH was observed in the HEC crystal and a water was bound to the active-site residue His75. Conserved lattice interactions suggested a common growth mechanism for the orthorhombic crystals of HEC and BLC.

1. Introduction

The presence of reactive oxygen species (ROS) in all aerobic organisms is harmful to DNA and other cellular components and causes disease and ageing if not immediately scavenged (Halliwell & Gutteridge, 1990). Hydrogen peroxide ($H_2O_2$) is one of the ROS that decomposes into free hydroxyl radicals, the most deleterious species of activated oxygen (Scandalias, 1997). Catalase (E.C. 1.11.1.6) catalyzes the reaction

$$2H_2O_2 \rightarrow 2H_2O + O_2$$

and is directly involved in eliminating oxidative stress. Loss of the catalase gene decreased the viability of fruit flies (Griswold et al., 1993) and reduced catalatic activity in Xeroderma pigmentosum cells could be directly related to impaired DNA repair (Quilliet et al., 1997), while stable overexpression of the catalase gene markedly attenuated the $H_2O_2$-induced toxic effect in immortalized neural cells (Mann et al., 1997). Catalase was shown to be effective in inhibiting the degeneration of neurons (Busciglio & Yankner, 1995). A growth-promoting factor derived from human erythrocytes with a wide target-cell spectrum was also identified as catalase (Takeuchi et al., 1995).

Human erythrocyte catalase (HEC) is a tetrameric protein of 244 kDa, comprising four identical subunits of 59.7 kDa plus four heme groups and four NADPH molecules (Bona et al., 1996; Sevinc et al., 1999). It is very similar to bovine liver catalase (BLC), which is the only known mammalian catalase that has had its three-dimensional structure determined (Murthy et al., 1981). The human gene of catalase has been cloned (Quan et al., 1986). Besides the 25-residue extension at the C-terminus, HEC differs from BLC in 43 amino-acid residues. In BLC, the four subunits are related by a 222 point-group molecular symmetry. Catalases from other sources, e.g. the Proteus mirabilis catalase (PMC; Gouet et al., 1995), catalase-A from Saccharomyces cerevisiae (SCC-A; Maté, Zamocky et al., 1999) and a larger catalase from Escherichia coli (HPII; Bravo et al., 1995), have a similar core structure and the same tetrameric organization as BLC. The tetrameric enzyme contains a central cavity and several channels that reach the active-site heme groups and facilitate the entry and exit of substrates and products (Fita & Rossmann, 1985; Gouet et al., 1996; Sevinc et al., 1999).

Because of their large molecular size, catalases have been exploited in a number of crystal-growth experiments (Sato et al., 1993; Malkin et al., 1995). HEC was observed as two-dimensional crystals by electron microscopy (Harris et al., 1993; Harris & Holzenburg, 1995), but three-dimensional crystals have not been reported until recently (Maté, Lombardia et al., 1999). In this paper, we describe the purification of HEC and its crystallization in new crystal forms. We have also determined the structure of HEC in an orthorhombic crystal by molecular replacement. This turned out to be very similar to the orthorhombic BLC crystals recently reported by Ko et al. (1999). Packing and intermolecular interactions explained the growth mechanism of the BLC crystals, which may also apply to the HEC crystals.
ammonium sulfate to 50% saturation. The pellet was dialyzed against a buffer containing 5 mM 2-mercaptoethanol, 0.2 mM EDTA, 10 mM potassium phosphate pH 6.5 and loaded onto a CM-Sephadex C-50 column equilibrated with the same buffer. The flow-through solution (part A) which had an A_{400}/A_{278} of 1.04 was found to contain the purest catalase, while most of the proteins were bound to the column, and fractions containing catalase (part B) were eluted with a 0–250 mM potassium phosphate gradient. The total yield was 110 mg, comprising 45 mg in part A and 65 mg in part B. To grow crystals, part A was used after concentration by ammonium sulfate precipitation and dialysis against 50 mM KCl, 1 mM 2-mercaptoethanol, 0.2 mM EDTA, 20 mM sodium acetate pH 5.2.

Crystallography refinement

Table 1

Data-collection and structure-refinement statistics.

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Unit-cell parameters (Å)</th>
<th>a = 84.9, b = 141.7, c = 232.5</th>
<th>Resolution (Å)</th>
<th>100–2.73 (2.82–2.73)</th>
<th>Number of observations</th>
<th>131236 (7325)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unique reflections</td>
<td>64674 (4406)</td>
<td>Completeness (%)</td>
<td>86.0 (67.0)</td>
<td>R_{max} (%)</td>
<td>10.9 (35.4)</td>
</tr>
<tr>
<td>Crystallographic refinement</td>
<td>Resolution range (Å)</td>
<td>40–2.75 (2.85–2.75)</td>
<td>Number of reflections (all data)</td>
<td>59969 (4498)</td>
<td>R value based on 92% data</td>
<td>23.9 (34.6)</td>
</tr>
<tr>
<td></td>
<td>With F &gt; 2σ cutoff</td>
<td>20.6 (29.9)</td>
<td>With F &gt; 2σ cutoff</td>
<td>31.6 (43.2)</td>
<td>R_{max} for 8% test data set</td>
<td>27.2 (39.5)</td>
</tr>
<tr>
<td></td>
<td>R.m.s.d. from ideal bond length (Å)</td>
<td>0.005</td>
<td>R.m.s.d. from ideal bond angle (°)</td>
<td>0.897</td>
<td>Dihedral angles</td>
<td>80.3</td>
</tr>
<tr>
<td></td>
<td>In most favored regions (%)</td>
<td>8.9</td>
<td>In additional allowed regions (%)</td>
<td>19.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Materials and methods

2.1. Purification, crystallization and data collection

DEAE–Sephadex A-50, CM–Sephadex C-50, 2-mercaptoethanol and PMSF were purchased from Sigma (St Louis, Missouri, USA). All other chemicals were from Fisher Scientific (Springfield, New Jersey, USA).

The procedure was a modification of Mori-kofer-Zwez et al. (1969). Outdated packed red blood cells were obtained from the Virginia Blood Bank. About 600 ml of red blood cells were mixed with two volumes of buffer containing 10 mM 2-mercaptoethanol, 10 mM EDTA, 10 mM 4-amino-caproic acid, 1 mM PMSF, 10 mM potassium phosphate pH 7.0 plus 10% (v/v) cold toluene. After homogenization and centrifugation to remove the upper fat layer, ammonium sulfate was added to the suspension to 50% saturation. The pellet was washed with 500 ml of buffer containing ammonium sulfate at 45% saturation, resuspended in a minimum volume of buffer at pH 7.2 and dialyzed against the same buffer. This sample was then loaded onto a DEAE–Sephadex A-50 column equilibrated with the pH 7.2 buffer. A green-brown band was observed to bind to the top of the column. After extensive washing with the buffer, the band was eluted using a 0–250 mM NaCl gradient. Fractions having an A_{400}/A_{278} ratio higher than 0.65 were pooled and the protein was precipitated by adding system and a rotating-anode generator (Rigaku RU-200) operating at 50 kV and 180 mA. Data were processed using the BIOTEX software (Molecular Structure Corporation, The Woodlands, Texas, USA) and the CCP4 suite (Collaborative Computational Project, Number 4, 1994). Statistics are listed in Table 1.

The second crystal form, having approximately hexagonal bipyramidal morphology, grew in drops equilibrated against a reservoir containing 5.5% PEG 8000, 15 mM MgSO_{4}, 70 mM Tris buffer at the pH value of 8.5. The crystals were of red-brown color and reached maximum dimensions of 0.4 × 0.3 × 0.2 mm. This crystal form belongs to a hexagonal space group but diffracted poorly; no attempt was made to determine the space group. The unit-cell parameters were a = b = 88.3, c = 257.1 Å, similar to the unit-cell dimensions reported by Maté, Lombarda et al. (1999) in their characterization of the HEC crystal. Their crystal belonged to the hexagonal space group P6_2_2_2; it is most likely that our crystal also belongs to this space group. A further search for crystallization conditions yielded a third crystal form. The reservoir contained 0.7 M potassium sodium tartrate and 0.1 M HEPES buffer pH 7.0. Protein concentration before mixing was 16 mg ml^{-1} and the initial drop volume was 3 μl. These crystals grew slowly over 4–6 months at room temperature and had well defined morphology. The final crystals had dimensions of 0.32 × 0.25 × 0.25 mm and diffracted X-rays to about 2.9 Å resolution. The crystals belonged to the space group I4_1, with unit-cell parameters a = b = 202.9, c = 152.4 Å.

2.2. Structure determination and refinement

The structure determination was carried out using the orthorhombic crystal of HEC and the BLC model (PDB entry 4blc; Ko et al., 1999). Because the space group and unit-cell dimensions of the BLC and HEC crystals were almost identical, the molecular packing and overall protein fold were presumed to be identical. Therefore, no rotational or translational search with the model was necessary. Direct application of the BLC model to the HEC crystal yielded a correlation coefficient of 0.74 and an R value of 0.30, calculated by the program AMoRe (Navaza, 1994) using data in the resolution range 8.0–4.0 Å. Subsequent refinement of the model employed the program X-PLOR (Brünger, 1992a). Model building was carried out on an SGI Octane using the
randomly selected 8% of reflections in the data set were excluded for cross-validation using \( R_{\text{free}} \) (Brünger, 1992b). The procedure of simulated annealing (Brünger et al., 1987, 1990) at high temperature was used; bulk-solvent correction (Jiang & Brünger, 1994) was also employed to incorporate the low-resolution data. Map calculations and density modifications included the entire data set. The NADPH molecules in BLC were not observed in the initial Fourier maps of the HEC crystal and were deleted. The HEC sequence was obtained from the SWISS-PROT database (accession number P04040). After substituting the 43 different amino-acid residues in BLC with those of HEC, preliminary refinement of the model gave an \( R \) value and an \( R_{\text{free}} \) of 0.30 and 0.37, respectively. This provided starting phase angles for subsequent density modification (DM) with solvent flattening, histogram matching and fourfold molecular averaging using the CCP4 suite (Collaborative Computational Project, Number 4, 1994). In DM, \( R_{\text{free}} \) improved from 0.362 to 0.304 and correlation coefficients between the subunit densities improved from 0.783–0.795 to 0.919–0.925. The map calculated using DM phases showed well defined density for the entire tetramer and additional densities for water molecules. Fig. 1 shows a DM map around one of the heme groups. However, in the N- and C-terminal regions the densities were still insufficient to allow tracing of the missing residues 1–3 and 503–527. The refinement proceeded with non-crystallographic symmetry (NCS) restraints. No improvement in the terminal regions was seen in the maps recalculated after refinement.

Figure 1
Stereodiagram and ribbon drawing of the heme group in HEC (subunit A). The residues Arg354 and Tyr358 on the proximal side and His75 and Asn148 on the distal side as well as the substrate water are shown. The electron-density map was calculated using DM phase angles and contoured at the 1.5σ level. The programs TOM, MOLSCRIPT and Raster3D were used in making this figure.

program O (Jones et al., 1990). The CCP4 suite (Collaborative Computational Project, Number 4, 1994) was used for density modification. The stereochemistry of the model was evaluated by PROCHECK (Laskowski et al., 1993). The programs TOM (Jones, 1982), MOLSCRIPT (Kraulis, 1991), Raster3D (Merritt & Murphy, 1994) and GRASP (Nicholls et al., 1991) were used to produce the figures.

Rigid-body refinement using the orthorhombic BLC model with all solvent molecules removed gave an initial \( R \) value of 0.38 for all data. In the following refinement, a

3. Results and discussion

3.1. Crystallographic refinement and structural comparison

The final model contained 1997 amino-acid residues in four subunits (named A, B, C and D), four heme groups and 393 waters. The polypeptide chain of each subunit had residues 4–502. An additional residue, Glu503, was included as an alanine in subunit B because its backbone was visible in the electron-density maps. With strong NCS and stereochemical restraints applied, the \( R \) value and \( R_{\text{free}} \) were 0.239 and 0.316, respectively, for all data in the resolution range 40–2.75 Å. For data with a \( 2\sigma \) cutoff, the corresponding values became 0.206 and 0.272, respectively. The average coordinate errors were between 0.30 and 0.35 Å, as estimated by a Luzzati plot (Luzzati, 1952). Other statistical values are listed in Table 1. The Ramachandran plot from PROCHECK showed that 80.3% of non-glycine and non-proline residues were in the most favored regions and only one residue, Ser217, in all four subunits was in a disallowed region. It had \( \varphi \) and \( \psi \) angles of 73 ± 5 and −59 ± 4°, respectively. Similar conformations were observed in the corresponding residues of other catalase models, e.g. Ser216 in BLC and Ile274 in HPII. The average temperature factor of 12.8 Å\(^2\) was lower than the value of about 20 Å\(^2\) estimated by a Wilson plot (Wilson, 1950). This was probably a consequence of the low resolution of the data set.

The overall protein fold of HEC is identical to BLC, with a core structure of an eight-stranded \( \beta \)-barrel surrounded by a number of \( \alpha \)-helices (Fita & Rossmann, 1985). The four subunits were organized around three perpendicularly intersecting dyad axes and express 222 point-group symmetry. The r.m.s.d. coordinate differences between the four subunits were 0.156 Å for the backbone atoms, 0.400 Å for the side chains and 0.125 Å for the heme groups, reflecting appropriate maintenance of the NCS restraints. The entire HEC and BLC tetramers could also be superimposed with an r.m.s.d. of 0.387 Å for 1996 equivalent C\(^\alpha\) atoms. The r.m.s.d. for the 7984 backbone atoms was 0.423 Å and was 1.006 Å for the 7740 equivalent atoms in the side chains. The four heme groups deviated by an r.m.s.d. of 0.724 Å or 0.41–0.53 Å when superimposed separately. Fig. 2 shows the two catalase models superposed. The 43 different amino-acid residues appear to be randomly distributed throughout the protein. Nevertheless, the \( \beta \) strand, which is involved in an association with its symmetry-
related equivalent, contains six amino-acid substitutions. The ligand to the heme iron, Tyr358 of HEC, as well as the essential His75 assumed virtually identical conformations to those observed in BLC. Other active-site residues, more than 30 in number, which constitute the binding pocket and proximal and distal sides of the heme group, are all conserved in HEC and BLC, except for the substitution of Ala157 in BLC by Pro158 in HEC. This residue is in contact with the vinyl group of the heme pyrrole ring I and has α-helical conformation. It is the first residue of the α3 helix and is well accommodated.

At 2.75 Å resolution, not much information could be extracted from the HEC model, even though it was refined with strong stereochemical and fourfold NCS restraints. The major difference between the two almost isomorphous orthorhombic crystal structures of HEC and BLC was the absence of bound NADPH molecules in HEC. The 19 amino-acid residues of BLC involved in contact with NADPH are also conserved in HEC, with the exception of Asp212 in BLC versus Asn213 in HEC. The acidic side chain of Asp212 in BLC forms hydrogen bonds with the 3′-phosphate of the NADPH molecule. Substitution by Asn213 in HEC neutralizes the side chain yet still maintains hydrogen-bonding capacity. Comparison with the BLC structure showed that the guanidinium group of Arg203 in HEC was rotated 60° away from the NADPH-binding position to form hydrogen bonds with the phenolic O atom of Tyr215.

The 76 backbone atoms of the 19 residues that line the NADPH pocket in HEC and BLC could be superimposed with r.m.s.d.s in the range 0.30–0.46 Å, depending on the subunits compared. The side chains deviated by an r.m.s.d. of 0.72–0.86 Å for 99 equivalent atoms. Because the differences are small, all of these residues in HEC can be readily adjusted to make proper interactions with the NADPH molecule as in BLC. Therefore, the binding site should be intact and functional. In addition, the side-chain amino group of Lys306 in HEC was close to the pyrophosphate group of NADPH in BLC, with distances of 1.7–2.4 Å. By forming salt bridge and/or hydrogen bonds with the pyrophosphate, this lysine should further strengthen the binding of NADPH to HEC. The corresponding residue in BLC is Gly305 and similar arrangement is not possible. In the orthorhombic BLC structure, the side chain of Phe197 had two conformations ‘A’ and ‘B’, which differ by a χ1 rotation of 120°; in the trigonal BLC crystal (PDB entry 7cat), it had the ‘B’ conformation; in the current HEC model, all equivalent Phe198 had the ‘A’ conformation. In either conformation, this side chain makes close contacts with the nicotinamide and adenine groups of the NADPH molecule; it is uncertain which conformation would be favored in binding NADPH.

In the structure of SCC-A, the NADP(H) site had only a partial occupancy of about 50% (Maté, Zamocky et al., 1999). The NADPH molecules in the orthorhombic BLC crystals had extremely high temperature factors (Ko et al., 1999). The results suggested weak binding of the dinucleotide to HEC.

3.2. Solvent model and lattice contacts

The orthorhombic HEC crystal had a specific volume (Matthews, 1968) of \(V_m = 2.87 \, \AA^3 \, \text{Da}^{-1}\), corresponding to a solvent content of 57%. There were 384 bound waters in the BLC crystal structure and 393 in the current HEC model. When the two models were superimposed, 38 pairs of equivalent waters were matched within 1.5 Å distance. If the matching criterion was 1.0 Å, the number of equivalents would be 19. These are few compared with the 99 equivalent waters in the trigonal and orthorhombic crystal structures of BLC (Fita et al., 1986; Ko et al., 1999). The 19 conserved waters showed a featureless pattern of distribution, except those bound to the O1A atoms of heme propionate and Thr361 OG1 atoms in all four subunits. In addition, a notable active-site water bound to the NE2 of His75 (at a distance of 2.59–2.85 Å) and possibly OD1 of Asn148 (at a distance of 3.21–3.89 Å) was also observed in each subunit. These ‘substrate’ waters had low temperature factors and strong electron densities in the Fourier maps (Fig. 1). The equivalent solvent molecules have been observed in a number of other catalases, including an azide ion in SCC-A (Maté, Zamocky et al., 1999) and a water in HP II (Bravo et al., 1995), but not in either crystal form of BLC. On the proximal side of the heme group, solvent channels with 3–4 bound water molecules were visible, but they appeared not to be directly involved in interaction with the active site. The channels were connected to the central cavity of the HEC tetramer, in which there were 16 waters.

Because the HEC crystal is almost identical or quasi-isomorphous to the BLC crystal, the interactions between the tetrameric catalase molecules should be very similar. This was verified by analyzing the crystal contacts. Like BLC, an HEC molecule was in contact with four symmetry-related neighbors in the crystal and the interactions could be divided into two types. The first type (interface I) is between molecules related by the 2l screw axis.
Table 2
Possible residues involved in crystal contacts.

<table>
<thead>
<tr>
<th>Residue 1 Atom</th>
<th>Residue 2 Atom</th>
<th>d (Å)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu424 NE2</td>
<td>Ser276 OG</td>
<td>3.44</td>
<td>Hydrogen bond (salt bridge).</td>
</tr>
<tr>
<td>Glu290 OE1</td>
<td>Lys243 NZ</td>
<td>&gt; 4</td>
<td>Could be mediated by water</td>
</tr>
<tr>
<td>Glu290 OE1</td>
<td>Lys106 NZ</td>
<td>&gt; 4</td>
<td>(salt bridge),</td>
</tr>
<tr>
<td>Glu428 OE2</td>
<td>Lys315 NZ</td>
<td>3.14</td>
<td>Salt bridge</td>
</tr>
<tr>
<td>Phe432 CE1</td>
<td>Thr341 CB</td>
<td>3.84</td>
<td>vdw contact</td>
</tr>
<tr>
<td>ProC46 CB</td>
<td>IleC103 CG2</td>
<td>3.98</td>
<td>vdw contact</td>
</tr>
<tr>
<td>ArgC47 NH1</td>
<td>GlnC104 O</td>
<td>2.8</td>
<td>Hydrogen bond</td>
</tr>
</tbody>
</table>

Interface II: molecules related by (i) x, y, z and (ii) \(-x, \frac{1}{2} + y, \frac{1}{2} - z\).

<table>
<thead>
<tr>
<th>Residue 1 Atom</th>
<th>Residue 2 Atom</th>
<th>d (Å)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asn449 O</td>
<td>Glu453 N</td>
<td>3.01</td>
<td>Hydrogen bond</td>
</tr>
<tr>
<td>Asn449 O</td>
<td>Arg456 NH2</td>
<td>3.26</td>
<td>Hydrogen bond</td>
</tr>
<tr>
<td>Asn452 ND2</td>
<td>Glu455 OE1</td>
<td>3.19</td>
<td>Hydrogen bond</td>
</tr>
<tr>
<td>Glu453 N</td>
<td>Asn449 O</td>
<td>3.46</td>
<td>Hydrogen bond</td>
</tr>
<tr>
<td>Glu455 OE1</td>
<td>Asn452 ND2</td>
<td>3.68</td>
<td>Hydrogen bond</td>
</tr>
<tr>
<td>Arg456 NH2</td>
<td>Asn449 O</td>
<td>2.68</td>
<td>Hydrogen bond</td>
</tr>
</tbody>
</table>

parallel to the b axis. It covered a surface area of 563 Å² on one molecule and 541 Å² on the other. There were at least 14 water molecules involved, which contributed an additional 380 Å² to the interface area. The second type (interface II) is between molecules related by the screw axis parallel to the a axis and, by coincidence, a non-crystallographic pseudo-dyad axis. The contact areas were 304 Å² and 287 Å² on the two protein molecules and 141 Å² on six waters. Interface I covered twice as much surface area as did interface II. Specific interactions could not be identified for sure owing to the low resolution and the NCS restraints of the refinement. Nevertheless, a few possible contacts are listed in Table 2. The salt bridges of Glu290–Lys106 and Glu428–Lys315 in HEC correspond to Glu289–Arg105 and Asp106–Arg107 in the orthorhombic space group as described by Ko et al. (1999), i.e. by formation of half unit-cell layers. Within a layer, molecules are associated through the stronger interface I; between adjacent layers, they interact through the weaker interface II.

We thank Professor Donatella Barra and Dr Bruno Maras for protein sequencing. This work was supported by the National Institutes of Health (Grant HL-32793).

References

short communications