The EMBO Journal Vol.17 No.9 pp.2463–2471, 1998

Crystal structure of human uroporphyrinogen decarboxylase

Frank G.Whitby, John D.Phillips[1], James P.Kushner[1] and Christopher P.Hill[2]

Department of Biochemistry and [1]Department of Medicine, University of Utah School of Medicine, 50 N.Medical Drive, Salt Lake City, UT 84132, USA

[2]Corresponding author

F.G.Whitby and J.D.Phillips contributed equally to this work

Uroporphyrinogen decarboxylase (URO-D) catalyzes the fifth step in the heme biosynthetic pathway, converting uroporphyrinogen to coproporphyrinogen by decarboxylating the four acetate side chains of the substrate. This activity is essential in all organisms, and subnormal activity of URO-D leads to the most common form of porphyria in humans, porphyria cutanea tarda (PCT). We have determined the crystal structure of recombinant human URO-D at 1.60 Å resolution. The 40.8 kDa protein is comprised of a single domain containing a $(\beta/\alpha)_8$-barrel with a deep active site cleft formed by loops at the C-terminal ends of the barrel strands. Many conserved residues cluster at this cleft, including the invariant side chains of Arg37, Arg41 and His339, which probably function in substrate binding, and Asp86, Tyr164 and Ser219, which may function in either binding or catalysis. URO-D is a dimer in solution (K_d = 0.1 μM), and this dimer also appears to be formed in the crystal. Assembly of the dimer juxtaposes the active site clefts of the monomers, suggesting a functionally important interaction between the catalytic centers.

Keywords: coproporphyrinogen/heme/PCT/porphyria cutanea tarda/porphyrin

Introduction

The synthesis of heme is an essential process, and the enzymatic steps of the heme biosynthetic pathway have been highly conserved throughout evolution (Wyckoff and Kushner, 1994). In eukaryotes, heme is produced in eight steps, with the fifth step being catalyzed by the cytosolic enzyme uroporphyrinogen decarboxylase (URO-D; EC 4.1.1.37). URO-D sequentially decarboxylates the four acetate side chains of both uroporphyrinogen III (see Figure 1) and its isomer uroporphyrinogen I, to produce coproporphyrinogen (I or III) (Wyckoff and Kushner, 1994). Only coproporphyrinogen III is a substrate for the next enzyme in the pathway, coproporphyrinogen oxidase.

URO-D has been purified from many species, ranging from bacteria to humans. SDS–PAGE indicates that the enzyme has an apparent monomer mol. wt of ~40 000 Da in most species. The cDNA sequence of the human gene predicts a protein with a mol. wt of 40 831 Da (Romeo *et al*., 1986). Purified recombinant human URO-D has been shown to be a homodimer with a K_d of 0.1 μM (Phillips *et al*., 1997), and URO-D has also been shown to be dimeric in chicken (Seki *et al*., 1986). All studies to date indicate that there is no requirement for cofactors or prosthetic groups for enzymatic activity (Kappas *et al*., 1995). The mechanism of decarboxylation is not known, although it has been proposed that the protonated pyrrole ring of the porphyrinogen functions as an electron sink, promoting electron withdrawal in a manner analogous to the pyridine ring of pyridoxyl phosphate (Bernard and Akhtar, 1979). At physiologic substrate concentrations, the decarboxylation reactions are reported to occur in an ordered fashion, starting with the acetate on the asymmetric ring of uroporphyrinogen III (Jackson *et al*., 1976; Luo and Lim, 1993). Under conditions of substrate excess, decarboxylation occurs in a random fashion (Luo and Lim, 1993).

Alignment of 13 URO-D amino acid sequences, including both prokaryotic and eukaryotic sources, shows an overall 10% identity and 33% similarity (see Figure 2) (Romeo *et al*., 1986; Nishimura *et al*., 1993). The region of highest homology is found near the N-terminus, where eight out of 10 residues are invariant. Additionally, there are several other small regions of high homology scattered throughout the protein. Although models involving one or more URO-D active sites have been suggested (de Verneuil *et al*., 1980; Chelstowska *et al*., 1992), the absence of repetitive motifs in the URO-D sequence implies the presence of a single active site, as do stereospecific labeling experiments which show that the chirality of all four acetate α-carbons is the same and is conserved during the decarboxylation reactions (Bernard and Akhtar, 1979).

Subnormal activity of URO-D is responsible for porphyria cutanea tarda (PCT), the most prevalent of the human porphyrias. PCT is characterized clinically by hyperpigmentation, a severe photosensitive dermatosis and hypertrichosis. The photosensitivity is mediated by the presence of uroporphyrin and partially decarboxylated

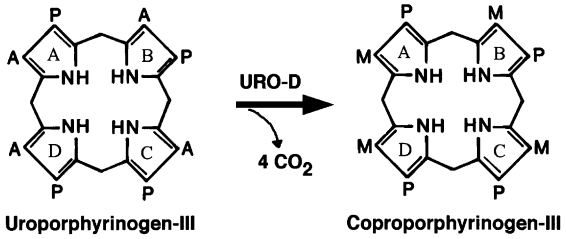

Fig. 1. Enzymatic conversion of uroporphyrinogen III to coproporphyrinogen III. A = acetate, P = propionate, M = methyl. The acetate side chains of uroporphyrinogen III are decarboxylated to methyl groups with the liberation of four molecules of carbon dioxide. In the type-I isomer, the acetate and propionate side chains of the D ring are ordered as on the other pyrrole rings, and the molecule is symmetrical.

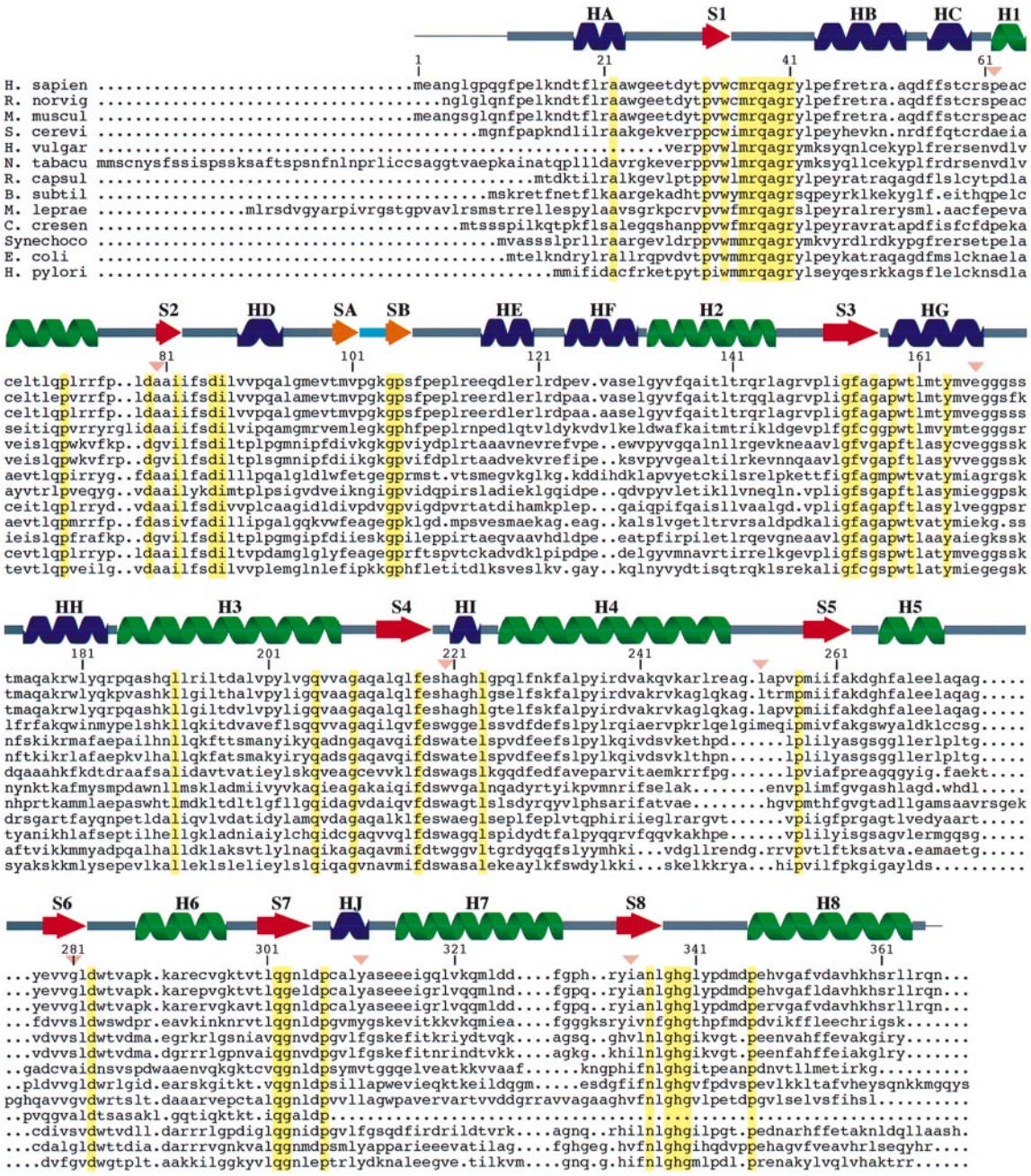

Fig. 2. Alignment of URO-D protein sequences. Thirteen sequences were aligned using the PILEUP routine of GCG. Residue numbering refers to the human sequence. Every twentieth residue is indicated. Invariant residues are highlighted in yellow. Secondary structure elements of human URO-D are indicated above, with helices and strands of the $(\beta/\alpha)_8$-barrel colored green and red, and other helices and strands colored blue and orange. A flexible segment that may cover the active site following ligand binding is colored cyan. Disordered residues that have been omitted from the model are indicated with a thin line. Mutated residues identified in humans with F-PCT are indicated by a red triangle. GenBank database entries shown are: *Homo sapiens* (U30787); *Rattus norvegicus* (Y00350); *Mus musculus* (J.D.Phillips, L.K.Jackson and J.P.Kushner, unpublished); *Saccharomyces cerevisiae* (Z49209); *Hordeum vulgare* (partial sequence X82832); *Nicotiana tabacum* (X82833); *Rhdobacter capsulatus* (U16796); *Bacillus subtilis* (M97208); *Mycobacterium leprae* (G699194); *Caulobacter cresentus* (partial sequence) (U13364); *Synochoccus* sp. strain PCC 7942 (Z11705); *Escherichia coli* (D12624); *Heliobacter pylori* (AE000574).

intermediate porphyrins in the skin and plasma. Biochemical findings include accumulation of uroporphyrin in the liver and excretion of large amounts of uroporphyrin and heptacarboxylic porphyrin in the urine (Wyckoff and Kushner, 1994; Kappas *et al*., 1995). The disease is classified as either familial (F-PCT) or sporadic (S-PCT). F-PCT is transmitted as an autosomal dominant trait caused by heterozygosity for mutations affecting the URO-D gene (Kushner *et al*., 1976). In F-PCT, URO-D activity is approximately half normal in all tissues (Wyckoff and Kushner, 1994). In S-PCT, the URO-D defect appears to be restricted to the liver (Elder *et al*., 1978), and no mutations have been identified in the URO-D gene (Garey *et al*., 1993). Lack of evidence for tissue-specific processing of the transcript generated from the single URO-D gene suggests that there is a liver-specific inhibitor of the enzyme in these cases. Clinical expression of both F-PCT and S-PCT has been associated with alcohol abuse, estrogen ingestion,

exposure to polyhalogenated hydrocarbons and hepatitis C infection (Kappas *et al*., 1995).

We report here the crystal structure of recombinant human URO-D. The protein adopts a distorted $(\beta/\alpha)_8$-barrel fold and contains a distinctive deep cleft that appears to provide the enzyme active site. URO-D forms a dimer in the crystal that probably represents the dimer formed in solution. This dimer places two active site clefts adjacent to each other in an arrangement that is probably required for enzyme function.

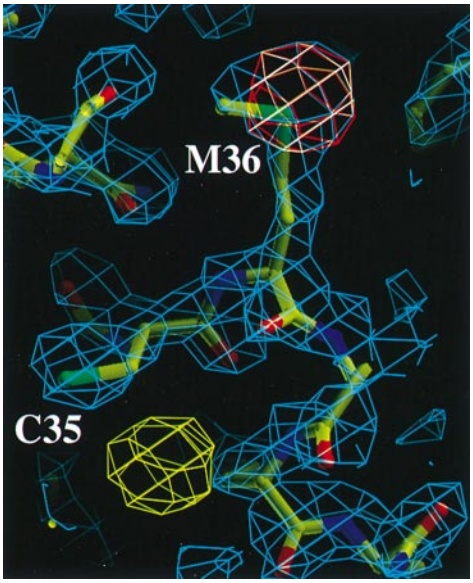

Fig. 3. Experimental MAD electron density map showing the current model. The experimental map, contoured at 1.4×r.m.s.d., is shown in cyan. An isomorphous difference Fourier map calculated from experimental phases and data collected from a previously identified mercury derivative is contoured at 3.0σ and shown in yellow. The position of the mercury atom, bound to Cys35, had been determined prior to MAD phasing, although this derivative possessed only poor phasing power and was not used in the structure determination. The position of the Met36 side chain is confirmed by an anomalous difference Fourier map, contoured in red at 4.0×r.m.s.d., that was calculated from Bijvoet differences and MAD protein phases –90°.

Results and discussion

Structure determination

Recombinant human URO-D with a 21 residue N-terminal extension containing 10 histidine residues (URO-Dt) was expressed in *Escherichia coli* and purified as described in Materials and methods. Crystals of URO-Dt were grown from a solution containing 24% 2-methyl-2,4-pentanediol (MPD) in space group $P3_121$ ($a = b = 103.3$ Å, $c = 74.3$ Å) with one molecule per asymmetric unit and a solvent content of ~54%. Selenomethionine-substituted URO-Dt (SeURO-Dt) was also prepared and yielded isomorphous crystals under the same conditions. The structure was determined at 2.0 Å resolution from the SeURO-Dt crystals by the method of multiple wavelength anomalous dispersion (MAD) (see Figure 3 and Table I). We also determined the structure of cutURO-D, which was prepared from URO-Dt by proteolytic removal of the 21 residue N-terminal histidine tag and the first 10 residues of authentic URO-D. CutURO-D did not crystallize in MPD but crystallized isomorphously from a solution containing 1.7 M citrate.

The model was refined against data to 1.80 Å resolution from the cutURO-D crystal grown in citrate to an *R*-value of 18.4% (R_{free} = 23.3%) with good stereochemistry (r.m.s.d. bonds = 0.018 Å, r.m.s.d. angles = 2.049°). The current model for this crystal includes all but the last residue of the cutURO-D sequence, i.e. residues 11–366, in addition to 282 water molecules and one molecule of β-mercaptoethanol (β-ME). Residues with ill-defined electron density (100–105) and a number of poorly defined side chains were included in the model with high *B*-factor.

This model was then refined against all data to 1.60 Å resolution from the URO-Dt crystal grown in MPD to an *R*-value of 24.2% (R_{free} = 28.9%) with good stereochemistry (r.m.s.d. bonds = 0.010 Å, r.m.s.d. angles = 1.703°). The structure is essentially identical to that of the citrate crystal described above. The histidine tag (residues –21 to –1), the authentic residues 1–10 and the C-terminal two residues (366–367) are entirely disordered. The model includes 355 residues and 267 water molecules.

Table I. Data collection statistics

	SeMet URO-Dt[a]				URO-Dt[b]	cutURO-D[c]
	λ1	λ2	λ3	λ4		
Wavelength (Å)	0.9798	0.9795	1.0688	0.9252	1.54	0.98
Observed reflections	272 178	185 889	93 663	187 246	305 819	156 721
Unique reflections	29 817	31 182	23 447	31 207	60 487	42 721
dmin (Å)	2.03	2.00	2.20	2.00	1.60	1.80
High resolution shell	2.03–2.06	2.00–2.03	2.20–2.24	2.00–2.03	1.60–1.63	1.80–1.83
Completeness (%)	96.8 (65.4)[d]	95.0 (58.9)	92.3 (67.3)	96.7 (93.2)	98.6 (85.2)	96.2 (91.7)
R_{sym}[e] (%)	3.8 (8.4)	3.8 (8.9)	3.0 (5.3)	3.7 (10.0)	3.9 (17.9)	4.8 (17.6)
Average I/σ(I)	>20 (9)	>20 (8)	18 (13)	19 (7)	19 (2.1)	15 (3.5)
Mosaicity (°)	0.439	0.449	0.440	0.446	0.437	0.530

[a]Crystals grown from MPD solution. Data collected at SSRL beamline 1–5 using Fuji image plates and a BAS2000 off-line scanner. Space group $P3_121$, unit cell dimensions $a = b = 103.3$ Å, $c = 74.1$ Å.
[b]Crystals grown from MPD solution. Data collected with a rotating-anode X-ray source and an RAXIS-IV image plate detector. Space group $P3_121$, unit cell dimensions $a = b = 103.3$ Å, $c = 74.3$ Å.
[c]Crystals grown from citrate solution. Data collected at SSRL, beamline 9–1 using a MAR image plate detector. Space group $P3_121$, unit cell dimensions $a = b = 103.0$ Å, $c = 73.7$ Å.
[d]Values in parentheses refer to the highest resolution shell.
[e]$R_{sym} = 100*\Sigma|I-<I>|/\Sigma I$.

As for the citrate crystal, residues with ill-defined electron density (100–105) and a number of poorly defined side chains were included in the model with high *B*-factor. Refinement statistics are given in Table II.

Structure of URO-D

URO-D is comprised of a single domain that includes a distorted $(\beta/\alpha)_8$-barrel and has overall dimensions of ~40×45×65 Å (Figure 4). Secondary structure assignments (Figure 5) were used as defined by the program PROMOTIF (Hutchinson and Thornton, 1996), with the exception that we also define residues 32–34 as a strand (S1) since this segment approximates the β conformation and closes the $(\beta/\alpha)_8$ motif (Sergeev and Lee, 1994). Strands and helices of the $(\beta/\alpha)_8$-barrel are named S1–S8 and H1–H8 respectively. All other helices have been given the alphabetical assignment HA–HJ, and the two short strands located within the loop following barrel strand S2 are named SA and SB. With the exception of a single hydrogen bond, the main chain hydrogen bonding network normally seen between the strands of a β-sheet is missing between strands S1 and S2, but is compensated for by two side chain interactions (see Figure 6). Helices HE, HI and HJ adopt the 3_{10} conformation, as do the last three residues of helix HB and the first three residues of helices HF and H5. As is typical of $(\beta/\alpha)_8$ structures, all of the helix–sheet turns at the N-terminal end of the barrel are comprised of only a few residues, while larger segments are found at the C-terminal end of the barrel strands (Murzin *et al*., 1994). The loops between the C-terminal end of each barrel strand and the subsequent barrel helix are numbered L1–L8. The smallest of these, L5, is composed of just four residues, while the largest, L2, contains three helices and two strands in a total of 48 residues.

The central residues of $(\beta/\alpha)_8$-barrels typically pack in three or four layers, each containing four residues from alternating strands of the barrel (Lesk *et al*., 1989). The core of the URO-D β-barrel is comprised of four distinct layers that contain predominantly hydrophobic residues. The first layer, at the C-terminal end of the β-barrel, is comprised of Ile82, Gln215, Gly281 and Asn336, and also includes part of Trp34 which is primarily part of the second layer. The second layer is comprised of Trp34, Ile152, Ile259 and Gln302, and the third layer of Ala80, Ala213, Val279 and Ile334. The fourth layer is formed by Pro32, Pro171, Pro257 and Thr300, which create an inward-pointing rim that encircles Phe19 and Ala23 which are part of helix HA. This helix contacts the barrel through a hydrophobic surface, and the side chain of Phe19 protrudes deeply into the core of the barrel, where it is also surrounded by the residues of the third layer.

A search for structurally related proteins with the DALI server (Holm and Sander, 1997) pairs URO-D with a large number of $(\beta/\alpha)_8$-containing proteins of broad func-

Table II. Refinement statistics

	cutURO-D (citrate)	URO-Dt (MPD)
Resolution range (Å)	20.0 – 1.80	20.0–1.60
No. of protein atoms[a]	4254	4193
No. of solvent atoms[a]	282	267
R-factor (%)[b]	18.4	24.2
R_{free} (%)[c]	23.3	28.9
Bond lengths (Å)[d]	0.018	0.010
Bond angles (°)[e]	2.049	1.703
<*B*> (Å^2) main chain	16.67	22.41
<*B*> (Å^2) side chains	20.20	25.87
<*B*> (Å^2) water molecules	26.96	35.39
<*B*> (Å^2) β-ME	59.80	e
No. of ø/ψ angles (%)[d]		
most favored	92.4	
additional	6.3	

[a]Non-hydrogen atoms only.
[b]*R*-factor = $100*\Sigma(||F_p(obs)|-|F_p(calc)||)/\Sigma|F_p(obs)|$. All data were used in the resolution range indicated, without application of a cut based upon the estimated standard deviation.
[c]R_{free} = *R*-factor for a selected subset (2.5–3.5%) of the reflections which were not included in prior refinement calculations.
[d]Stereochemistry was assessed with PROCHECK (Laskowski *et al*., 1993).
[e]One β-ME molecule has been positioned in the cutURO-D (citrate) structure, but not in the URO-Dt (MPD) model.

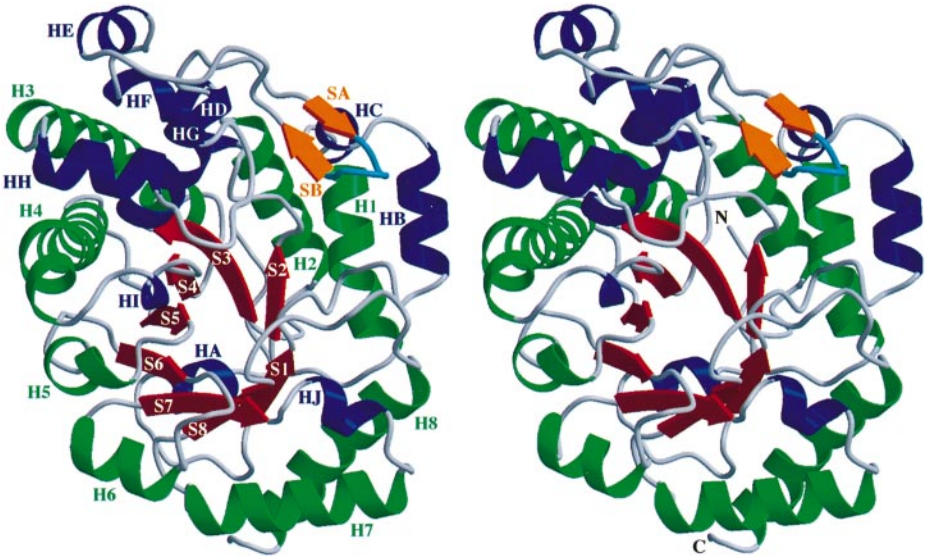

Fig. 4. Ribbon representation of the URO-D structure. This stereoview is approximately along the axis of the β-barrel, looking directly at the active site cleft. Secondary structural elements are labeled. Chain termini are indicated with an N and a C. The color code is the same as Figure 2.

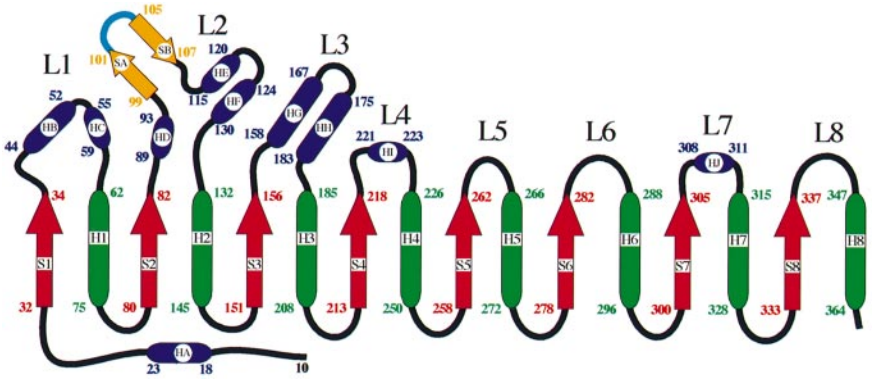

Fig. 5. Topology diagram showing secondary structure elements of human URO-D. Helices are shown as oval-shaped bars and strands are shown as arrows. The first and last residue of each secondary structural element is numbered. The color code is the same as Figure 2.

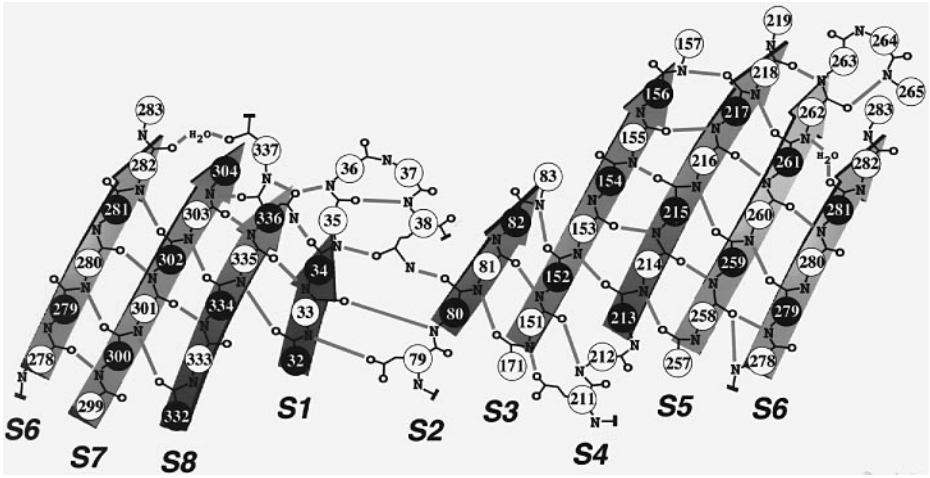

Fig. 6. Schematic representation of the strands in the β-barrel. β-Strands are shown as broad arrows and numbered from the most N-terminal (S1) to the most C-terminal (S8). Residues that point toward the barrel core are shown as white circles, and those that point out of the barrel are shown as black circles. Hydrogen bonds are represented by thin lines. The distortion in the barrel structure between strands S1 and S2 is compensated by the hydrogen bond network involving Gln38 and Asp79.

tionality, with closest structural similarity to trimethylamine dehydrogenase (PDB identifier, 2tmd) (Lim *et al*., 1986). Like many $(\beta/\alpha)_8$ proteins, the cross-sectional ratio of the barrel is nearly 1.0 (low ellipticity) and the diameter of the β-barrel itself is ~14 Å (Lesk *et al*., 1989; Murzin *et al*., 1994). However, further efforts at structural alignments of URO-D to structurally related proteins did not yield any insights regarding possible mechanisms.

Active site cleft

A deep cleft that appears to provide the enzyme active site is formed by loops L1, L2, L3, L4 and L8, at the C-terminal end of the β-barrel. The size of this cleft (~15×15×7 Å) appears to be suitable for insertion of a porphyrinogen into the active site, with most of the substrate shielded from solvent, consistent with the creation of an isolated environment in which catalysis can take place (see Figure 7). Consistent with binding the highly negatively charged substrate, the active site cleft contains a number of positively charged side chains (Arg37, Arg41, Arg47, Arg50, Lys263, His220, His223 and His339). The observation of a single active site cleft contradicts an earlier speculation that URO-D contains multiple active sites (de Verneuil *et al*., 1980) and is consistent with the observation that all decarboxylation reactions catalyzed by URO-D follow the same stereochemistry (Bernard and Akhtar, 1975, 1979).

The location of the proposed catalytic center is consistent with the observation that invariant residues are concentrated around the active site cleft (see Figure 7). Many of the 37 invariant residues (see Figure 2) perform obvious structural roles. For example, Ala22, Pro32, Trp34, Leu191 and Leu337 all pack in the hydrophobic core, while Asn336, Gln38 and Gln302 side chains participate in buried hydrogen bonding interactions.

Invariant residues that present groups capable of forming hydrogen bonding interactions into the active site cleft are excellent candidates for direct participation in substrate binding and catalysis. Thus, three invariant residues with positively charged side chains (Arg37, Arg41 and His339) may provide stabilizing interactions with the acetate and/or propionate carboxylate groups of the substrates. The invariant side chains of Tyr164 and Asp86 point into the active site cleft and are candidates for participation in catalysis. The presence of the invariant carboxylate side chain of Asp86 in the active site is particularly suggestive, as the initial substrate has eight negatively charged carboxylates and no positively charged groups. Asp86 is a good candidate for a role in either donating a proton or electrostatic destabilization of the charged substrate. We

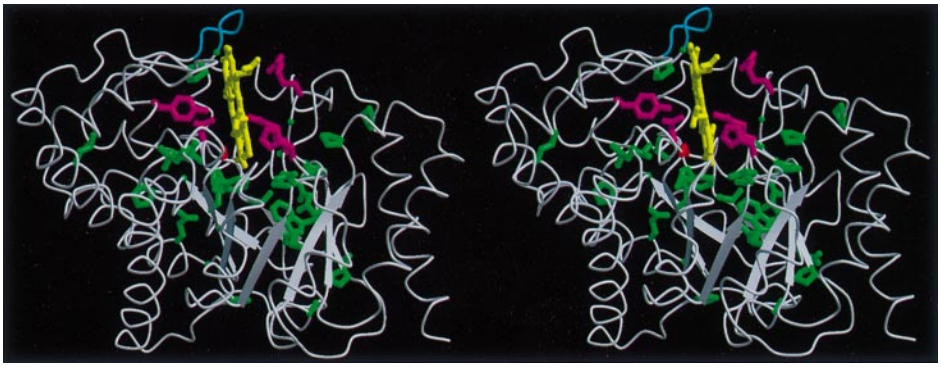

Fig. 7. Conserved residues of URO-D and the active site cleft. The URO-D backbone is shown in white, with the flexible segment of residues 100–105 colored cyan. The side chains of all invariant residues (see Figure 2) are shown, with glycine residues represented by a single sphere. Arg37, Arg41, Asp86, Tyr164 and His339 are colored magenta, with other invariant residues colored green. Ser110, which is invariant except for the *E.coli* sequence where it is a threonine, is shown in red. A uroporphyrinogen model has been roughly docked in the active site cleft to convey an impression of scale. This simple docking is not intended to imply any details beyond the location and approximate dimensions of the active site cleft. The only severe clashes in the docked conformation are with the Arg37 side chain, which can reasonably be modeled in several alternative conformations.

also propose that the side chain of Ser219, which projects into the active site cleft, may function to hydrogen-bond the substrate. This residue is conserved in all sequences, except for *E.coli*, which has a threonine at this position (see Figure 2).

The active site cleft also contains 10 solvent-exposed hydrophobic side chains (Met36, Phe46, Phe55, Ile82, Phe84, Ile87, Leu88, Phe154, Phe217 and Phe261) that are highly conserved or invariant and are candidates for participation in substrate binding. Binding of the substrate acetate group in a hydrophobic environment at the bottom of the active site cleft may contribute to catalysis by destabilization of the charged substrate with respect to the carbon dioxide product.

A flexible segment of loop L2 (residues 100–105) is located at the top of the putative active site cleft. We describe this segment as being flexible on the basis of high *B*-factors and broken electron density. The conformation adopted by this segment in the crystal structure appears partially to impede access to the active site cleft. We speculate that mobility of this segment may be of functional importance, with an open conformation accommodating substrate access and a closed conformation providing specific interactions that orient substrate and maintain a defined environment away from bulk solvent. An invariant Gly105–Pro106 pair is located at one end of the flexible segment, suggesting that the conformation(s) of this sequence is important for URO-D function.

URO-D as a dimer

URO-D has been shown by equilibrium sedimentation to dimerize with a K_d of ~0.1 μM (Phillips *et al.*, 1997). The solution dimer is apparently formed in the crystal by operation of a crystallographic 2-fold axis on the single molecule of the asymmetric unit (see Figure 8). The crystallographic dimer interface, which is formed by helix HH of loop L3 and all of the smaller loops (L4, L5, L6, L7 and L8), is remarkably flat and extensive, with a total of 2387 Å^2 of solvent-accessible surface area apparently buried upon dimerization. The interface is largely hydrophilic, and buries 27 ordered water molecules ($<B>$ = 16.9 Å^2) to create an extensive hydrogen bonding network. This hydrophilic character is consistent with the moderate dimerization affinity. The extensive nature of this interface indicates that it is more likely to be physiologically relevant than the three other considerably smaller contacts seen in the crystal, which bury 1006, 718 and 101 Å^2 of solvent-accessible surface area respectively. Furthermore, the 1006 Å^2 crystal contact is the result of translational symmetry and is thus not suitable for formation of a discrete dimer. The most powerful argument for relevance of the crystallographic dimer is that one edge of the monomer active site cleft is closed by the neighboring molecule. It is even possible that some residues, in particular Ser172, make direct contact with substrate bound to the neighboring molecule, although we note that Ser172 is not a conserved residue. At the least, however, dimer formation apparently serves to make a deeper cleft that is more protected from solvent.

The active site cleft of one monomer is adjacent to that of its neighbor in the dimer. This creates a single extended cleft that is partly divided by Ser172, which approaches within 7 Å of its symmetry-related residue in the dimer at about mid-depth of the cleft. This cleft is large enough to accommodate two substrate molecules in close proximity, or to allow reaction intermediates to shuttle between monomers. Ser172 and neighboring residues are well defined and have low *B*-factors (B <22.0 Å^2 for main chain atoms), suggesting that this dividing structure is not inherently flexible. Furthermore, helix HH makes buttressing interactions that would have to be disrupted in order to open the constriction caused by Ser172. In order for partially decarboxylated porphyrinogens to shuttle between monomer active sites, they would have to partially dissociate from the active site cleft. Thus, a deep active site allows for specific catalysis, and a less deep channel may allow for efficient transfer between active sites of intermediates generated during the stepwise decarboxylation of uroporphyrinogen to coproporphyrinogen. It is possible that porphyrinogen intermediates must at least partially exit the deep active site cleft in order to allow release of the carbon dioxide product. Transfer to an adjacent active site may provide an effective mechanism for this process. An alternate possibility is that the two adjacent active site clefts collaborate in a single decarboxylation reaction of one substrate molecule.

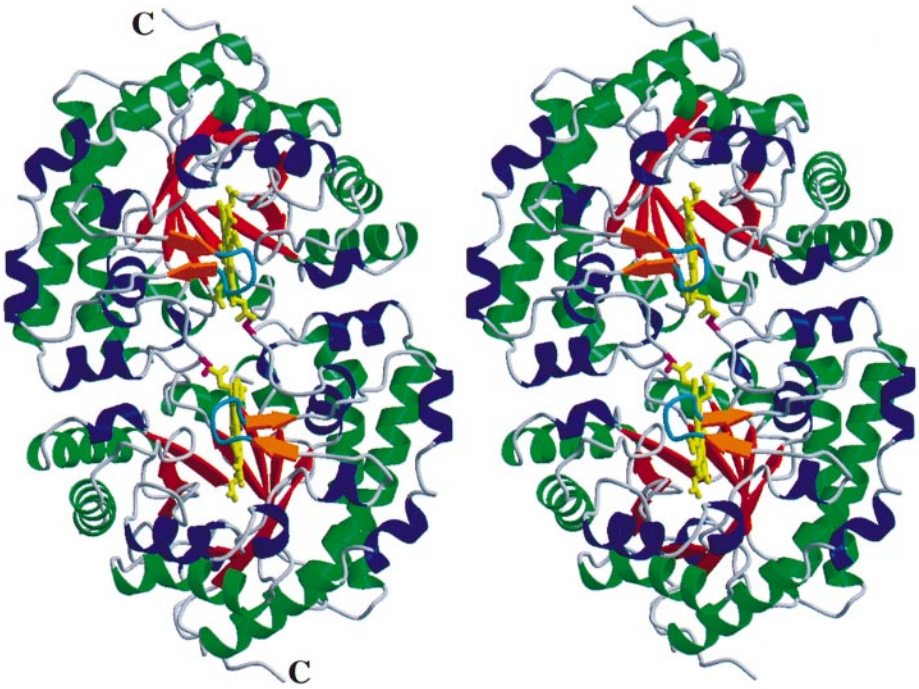

Fig. 8. Stereoview of the URO-D dimer. The color code is the same as Figure 2. In addition, the side chain of Ser172 is shown in magenta and the modeled uroporphyrinogen ligand is shown in yellow. The view direction is along the dimer 2-fold axis. In this view, the upper monomer is tilted ~45° from that of Figure 3 and rotated ~45° about the view direction.

Correlation with mutagenesis studies

URO-D has been the subject of mutagenesis studies targeting residues hypothesized to function in catalysis. One or more of the six cysteine residues in human URO-D were proposed to function in catalysis because sulfhydryl-modifying reagents, such as *N*-ethylmaleimide (NEM), inhibit enzymatic activity (Elder *et al*., 1978; Kawanishi *et al*., 1983). However, site-directed mutation of each of the six cysteine residues to serine indicated that no single thiol was critically involved in the catalytic process (Wyckoff *et al*., 1996). These data are explained by inspection of the structure. Cys35, Cys59 and Cys65 are buried, and modification with a bulky maleimide group is likely to disrupt the structure, but the conservative change to serine is tolerated. Cys66 and Cys294 are at the molecular surface, while Cys308 is on the surface of the monomer and buried in the dimer interface.

Histidine residues have been implicated in catalysis as URO-D activity is inhibited by ethoxyformic acid (diethylpyrocarbonate, DEPC) (Kawanishi *et al*., 1983; Koopman and Battle, 1987). Three invariant histidine residues were hypothesized to perform key functional roles, but URO-D with each of these residues changed to asparagine retained significant enzymatic activity (Wyckoff *et al*., 1996). These observations seem reasonable in light of the structure, since all three of the histidine side chains are on the molecular surface. One of the three histidine mutants described above, His339→Asn, retains normal activity against the initial 8-COOH substrate, but is only weakly active for additional decarboxylations of 7-, 6- and 5-COOH intermediates. This residue is located at the opening to the active site cleft and thus may function in the orientation of partially decarboxylated substrates in the active site.

Mutations causing a decrease in URO-D activity have been identified in patients with F-PCT. Most of the mutated residues appear to perform important structural roles away from the active site cleft or dimer interface (see Table III). Two clinical mutants, Glu167→Lys and His220→Pro, are located near the active site cleft. The Glu167 side chain is buried and its carboxylate forms hydrogen bonds that appear to stabilize the conformation of helix HH at the dimer interface and possibly also residues 170–172 at the active site cleft. Mutation of this residue may disrupt the active site geometry and/or dimerization. The moderate reduction in activity for the His220→Pro mutant indicates that this imidazole side chain, which faces the active site cleft, does not play a critical role in catalysis. The only clinical mutant located at the dimer interface is Tyr311, which retains 60% of wild-type activity upon mutation to cysteine. This does not invalidate the hypothesis that dimer formation is important for activity, however, since Tyr311 is not extensively close packed and the cysteine mutant can probably maintain equivalent contacts across the dimer interface.

Materials and methods

Protein chemistry

Recombinant histidine-tagged human URO-D (URO-Dt) was overexpressed in *E.coli* and purified with nickel-chelate column chromatography as described (Phillips *et al*., 1997). Selenomethionine-substituted URO-Dt (SeURO-Dt) was produced from the gal–, met– auxotroph B834(DE3) of the BL21 strain of *E.coli*. Cells were grown as described by Johnston *et al.* (1997). Purification of SeURO-Dt was essentially identical to that of URO-Dt (Phillips *et al*., 1997), although the yield was ~10-fold lower.

The histidine tag of recombinant URO-Dt was cleaved by incubating the purified protein at 6 mg/ml in 50 mM Tris–HCl pH 7.5, 1 mM β-ME, 10% glycerol and 1 mM $CaCl_2$ at room temperature for 24 h

Table III. Naturally occurring human URO-D mutations

Mutation	Enzymatic activity % WT	Structural effects
G281E[a]	(decreased *in vivo*)	buried in core of $(\beta/\alpha)_8$-barrel
G281V[b]	(decreased *in vivo*)	buried in core of $(\beta/\alpha)_8$-barrel
E167K[c]	(decreased *in vivo*)	stabilizes conformation of a turn at the active-site cleft and helix HH at the dimer interface
H220P[d]	40	in active site cleft
I334T[e]	15	buried in core of $(\beta/\alpha)_8$-barrel
L253Q[e]	5	caps a cluster of buried hydrophobic residues
A80G[e]	30	buried in core of $(\beta/\alpha)_8$-barrel
P62L[f]	5	at end of helix HC and start of helix HI. In van der Waals contact with helix H2
Y311C[f]	58	at dimer interface

[a]de Verneuil *et al*. (1984, 1988); [b]Garey *et al*. (1989); [c]Romana *et al*. (1991); [d]Meguro *et al*. (1994); [e]McManus *et al*. (1996); [f]Moran-Jimenez *et al*. (1996).

with 1 mg of factor Xa per 1000 mg of URO-Dt. The cleaved protein (cutURO-D) was then purified further by size-exclusion chromatography on a Superdex-75 sizing column (16/60, Pharmacia) in 150 mM NaCl, 100 mM Tris–HCl, pH 7.5. The cutURO-D was then separated from uncut URO-Dt by collecting the flow-through of a nickel-chelate column. The cutURO-D protein was shown to be >99% pure by SDS–PAGE. Enzymatic activity was assayed by the method of Straka and Kushner (Straka *et al*., 1982). The activity of cutURO-D was the same as that of native URO-D and URO-Dt.

N-terminal sequencing of fresh, purified URO-Dt indicated that the first five residues were GHHHH, as expected from the sequence of the histidine tag, without the N-terminal methionine. The mass of URO-Dt was 43 184 ± 6.8 Da, as determined by electrospray mass spectrometry (calculated mass = 43 177 Da). Following cleavage with factor Xa, the first five residues of cutURO-D (FPELK), corresponded to residues 11–15 of the authentic sequence (residues 32–36 of the URO-Dt sequence). This indicates that URO-D is highly susceptible to proteolysis at residue 10 of the authentic sequence.

Crystallization

SeURO-Dt, URO-Dt and cutURO-D were crystallized in sitting drops at 4°C in either MPD, as described (Phillips *et al*., 1997), or in citrate buffer. For crystallization in MPD, 5 µl of a SeURO-Dt or URO-Dt solution containing protein at 6 mg/ml, 50 mM Tris–HCl pH 7.5, 1 mM β-ME and 10% glycerol was mixed with 2 µl of the reservoir solution containing 16–24% MPD, 100 mM MES, pH 5.6–6.5. For crystallization in citrate, 5 µl of a cutURO-D or URO-Dt solution containing protein at 6 mg/ml, 50 mM Tris–HCl pH 7.5, 1 mM β-ME and 10% glycerol was mixed with 2 µl of the reservoir solution containing 1.2–1.7 M citrate, pH 5.8–6.2. Crystals grown under all of these conditions were isomorphous and extremely radiation sensitive unless cryo-cooled.

X-ray data collection

Data were collected using Fuji image plates and an off-line scanner on beamline 1–5 at the Stanford Synchrotron Radiation Laboratory (SSRL), an RAXIS-IV area detector on a rotating anode X-ray source, or with a MAR image-plate area detector on beamline 9–1 at SSRL (see Table I). Data integration and scaling was performed with the programs DENZO and SCALEPACK (Otwinowski, 1993). The space group is $P3_121$ with cell dimensions that range from a = 103.0–103.3 Å and c = 73.7–74.3 Å. There is one molecule in the asymmetric unit, and the Matthew's coefficient (V_m) of 2.7 Å^3/Da indicates a solvent content of ~54% (Matthews, 1968).

All data were collected from crystals maintained at 100 K. In preparation for data collection, crystals grown in MPD were removed from the crystallization drop, briefly washed in the reservoir buffer containing 24% MPD, suspended in a small rayon loop and plunged rapidly into liquid nitrogen. Crystals grown in citrate were removed from the crystallization drop, washed for 30 s in a buffer identical in composition to the reservoir solution with the addition of 5% glycerol, and cooled by plunging into liquid nitrogen.

Data for MAD phase determination were collected from a single SeURO-Dt crystal using Fuji image plates and a BAS2000 off-line scanner on beamline 1–5 at SSRL. Four wavelengths were chosen from the fluorescence spectrum (see Table I). Data were collected at each wavelength in contiguous 20° sweeps, with each sweep followed by a 'Friedel flip' to collect the equivalent data related by rotation of 180° about the phi drive. These data are of high quality to the edge of the detector (2.0 Å). Data from each wavelength were indexed according to the same crystal orientation matrix, and integrated and scaled independently for the four data sets. Scaled data sets from each of the four wavelengths were then scaled together from 20.0 to 2.0 Å resolution.

Structure determination and refinement

The eight ordered seleno-methionine positions in SeURO-Dt, from a total of 11 methionine residues, were determined from difference Patterson and Fourier maps using the package XtalView (McRee, 1992). Selenium positions were refined in PHASES (Furey, 1990), treating the four wavelength MAD experiment as a special case of multiple isomorphous replacement (Ramakrishnan and Biou, 1997). The mean figure of merit calculated by PHASES was 0.61.

The resulting electron density map was of high quality (see Figure 3). Solvent flattening was performed with PHASES (Furey, 1990) to a mean figure of merit of 0.89 but did not obviously improve the already very clear electron density map. A model consisting of the majority of URO-D was built using the program O (Jones *et al*., 1991) and refined against the SeURO-Dt data with XPLOR (Brünger, 1996). Rigid-body, simulated annealing, positional and *B*-factor refinement gave an *R*-value of 28% and R_{free} of 32% (Brünger, 1992) against all data between 8.0 and 2.0 Å resolution. The model was then rebuilt manually and refined seven times, including a bulk solvent correction, against all data from 20.0 to 1.80 Å resolution from the cutURO-D crystal grown in citrate.

Towards the end of refinement, the only significant unexplained electron density was a well-defined sickle-shaped region of strong density lying close to a 2-fold rotation axis. This density is well modeled by an ordered β-ME molecule, which is present at 1 mM in the crystallization buffer. The relatively high *B*-factor for the β-ME molecule (60 Å^2) may reflect partial occupancy. Orientation of the β-ME molecule is indicated by the strongest peak in a difference electron density map for the sulfur atom. The final model includes 282 water molecules, one β-ME molecule and 356 of the 367 residues of URO-D. The *R*-value for this model is 18.4% and the R_{free} is 23.3% for all data to 1.80 Å resolution with good stereochemistry (see Table II).

This model was then refined against all data of the native URO-Dt crystal grown in MPD. The final model includes 267 water molecules and 355 of the 367 residues of URO-D (355 of the 388 residues of URO-Dt). The β-ME molecule was not included in this model because the corresponding electron density was poorly defined. In this case, the N-terminal 31 residues are entirely disordered, as are the C-terminal two residues. The current *R*-value for this model is 24.2% and the R_{free} is 28.9% for all data from 20.0 to 1.60 Å resolution (see Table II). Coordinate and diffraction data will be deposited with the Brookhaven Protein Database and are available from the authors.

Acknowledgements

We thank Henry Bellamy, Anand Kolatkar and J.Randolph Knowlton for assistance with data collection; Laurie Jackson and Dan Smith for assistance with protein purification, and Robert Shackmann for N-terminal sequencing. This work was supported in part by a grant from the Lucille P.Markey Charitable Trust and National Institutes of Health grants RO1GM856775, RO1DK20503, MO1RR00064 and P50DK49219.

References

Barnard,G.F. and Akhtar,M. (1975) Stereochemistry of porphyrinogen carboxy-lyase reaction in haem biosynthesis. *J. Chem. Soc. Chem. Commun*., **10**, 494–496.

Barnard,G.F. and Akhtar,M. (1979) Stereochemical and mechanistic studies on the decarboxylation of uroporphyrinogen III in haem biosynthesis. *J. Chem. Soc. Perkin Trans. I*, **13**, 2354–2360.

Brünger,A.T. (1992) Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.

Brünger,A.T. (1996) *X-PLOR Version 3.843: A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.

Chelstowska,A., Zoladek,T., Garey,J., Kushner,J., Rytka,J. and Labbe-Bois,R. (1992) Identification of amino acid changes affecting yeast uroporphyrinogen decarboxylase activity by sequence analysis of *hem12* mutant alleles. *Biochem. J*., **288**, 753–757.

de Verneuil,H., Grandchamp,B. and Nordmann,Y. (1980) Some kinetic properties of human red cell uroporphyrinogen decarboxylase. *Biochim. Biophys. Acta*, **611**, 174–186.

de Verneuil,H., Grandchamp,B., Foubert,C., Weil,D., N′Guyen,V.C., Gross,M.S., Sassa,S. and Nordmann,Y. (1984) Assignment of the gene for uroporphyrinogen decarboxylase to human chromosome 1 by somatic cell hybridization and specific enzyme immunoassay. *Hum. Genet*., **66**, 202–205.

de Verneuil,H., Hansen,J., Picat,C., Grandchamp,B., Kushner,J., Roberts,A., Elder,G. and Nordmann,Y. (1988) Prevalence of the 281 (Gly→Glu) mutation in hepatoerythropoietic porphyria and porphyria cutanea tarda. *Hum. Genet*., **78**, 101–102.

Elder,G.H., Lee,G.B. and Tovey,J.A. (1978) Decreased activity of hepatic uroporphyrinogen decarboxylase in sporadic porphyria cutanea tarda. *N. Engl. J. Med*., **299**, 274–278.

Furey,W.S.S. (1990) PHASES: A Program Package for the Processing and Analysis of Diffraction Data from Macromolecules.

Garey,J.R., Hansen,J.L., Harrison,L.M., Kennedy,J.B. and Kushner,J.P. (1989) A point mutation in the coding region of uroporphyrinogen decarboxylase associated with familial porphyria cutanea tarda. *Blood*, **73**, 892–895.

Garey,J.R., Franklin,K.F., Brown,D.A., Harrison,L.M., Metcalf,K.M. and Kushner,J.P. (1993) Analysis of uroporphyrinogen decarboxylase complementary DNAs in sporadic porphyria cutanea tarda. *Gastroenterology*, **105**, 165–169.

Holm,L. and Sander,C. (1997) Dali/FSSP classification of three-dimensional protein folds. *Nucleic Acids Res*., **25**, 231–234.

Hutchinson,E.G. and Thornton,J.M. (1996) PROMOTIF—a program to identify and analyze structural motifs in proteins. *Protein Sci*., **5**, 212–220.

Jackson,A.H., Sancovich,H.A., Ferrmola,A.M., Evans,N., Games,D.E., Matlin,S.A., Elder,G.H. and Smith,S.G. (1976) Macrocyclic intermediates in the biosynthesis of porphyrins. *Philos. Trans. R. Soc. Lond*., **273**, 191–206.

Johnston,S.C., Larsen,C.N., Cook,W.J., Wilkinson,K.D. and Hill,C.P. (1997) Crystal structure of a deubiquitinating enzyme (human UCH-L3) at 1.8 Å resolution. *EMBO J*., **16**, 3787-3796.

Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and location of errors in these models. *Acta Crystallogr*., **A47**, 110–119.

Kappas,A., Sassa,S., Galbraith,R.A. and Nordmann,Y. (1995) The porphyrias. In Scriver,C.R., Beaudet,A.L., Sly,W.S. and Valle,D. (eds), *The Metabolic and Molecular Basis of Inherited Disease*. McGraw-Hill, New York, pp. 2103–2160.

Kawanishi,S., Seki,Y. and Sano,S. (1983) Uroporphyrinogen decarboxylase. Purification, properties, and inhibition by polychlorinated biphenyl isomers. *J. Biol. Chem*., **258**, 4285–4292.

Koopman,G.E. and Battle,A.d.C. (1987) Biosythesis of porphyrins in *Rhodoseudomonas palustris*. VI. The effect of metals, thiols, and other reagents on the activity of uroporphyrinogen decarboxylase. *Int. J. Biochem*., **19**, 373–377.

Kushner,J.P., Barbuto,A.J. and Lee,G.R. (1976) An inherited enzymatic defect in porphyria cutanea tarda: decreased uroporphyrinogen decarboxylase activity. *J. Clin. Invest*., **58**, 1089–1097.

Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*., **26**, 283–291.

Lesk,A.M., Branden,C.-I. and Chothia,C. (1989) Structural principles of α/β barrel proteins: the packing of the interior of the sheet. *Protein Struct. Funct. Genet*., **5**, 139–148.

Lim,L.W., Shamala,N., Mathews,F.S., Steenkamp,D.J., Hamlin,R. and Xuong,N.H. (1986) Three-dimensional structure of the iron–sulfur flavoprotein trimethylamine dehydrogenase at 2.4 Å resolution. *J. Biol. Chem*., **261**, 15140–15146.

Luo,J. and Lim,C.K. (1993) Order of uroporphyrinogen III decarboxylation on incubation of porphobilinogen and uroporphyrinogen III with erythrocyte uroporphyrinogen decarboxylase. *Biochem. J*., **289**, 529–532.

Matthews,B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol*., **33**, 491–497.

McManus,J.F., Begley,C.G., Sassa,S. and Ratnaike,S. (1996) Five new mutations in the uroporphyrinogen decarboxylase gene identified in families with cutaneous porphyria. *Blood*, **88**, 3589–3600.

McRee,D.E. (1992) A visual protein crystallographic software system for X11/XView. *J. Mol. Graphics*, **10**, 44–46.

Meguro,K. *et al*. (1994) Molecular defects of uroporphyrinogen decarboxylase in a patient with mild hepatoerythropoietic porphyria. *J. Invest. Dermatol*., **102**, 681–685.

Moran-Jimenez,M.J., Ged,C., Romana,M., Enriquez-De-Salamanca,R., Taieb,A., Topi,G., D'Alessandro,L. and de-Verneuil,H. (1996) Uroporphyrinogen decarboxylase: complete human gene sequence and molecular study of three families with hepatoerythropoietic porphyria. *Am. J. Hum. Genet*., **58**, 712–721.

Murzin,A.G., Lesk,A.M. and Chothia,C. (1994) Principles determining the structure of β-sheet barrels in proteins. *J. Mol. Biol*., **236**, 1382–1400.

Nishimura,K., Nakayashiki,T. and Inokuchi,H. (1993) Cloning and sequencing of the *hemE* gene encoding uroporphyrinogen III decarboxylase (UPD) from *Escherichia coli* K-12. *Gene*, **133**, 109–113.

Otwinowski,Z. (1993) Oscillation data reduction program. In Sawyer,L. Isaacs,N. and Bailey,S. (eds), SERC Daresbury Laboratory, Warrington UK, pp. 56–62.

Phillips,J.D., Whitby,F.G., Kushner,J.P. and Hill,C.P. (1997) Characterization and crystallization of human uroporphyrinogen decarboxylase. *Protein Sci*., **6**, 1343–1346.

Ramakrishnan,V. and Biou,V. (1997) Treatment of multiwavelength anomalous diffraction as a special case of multiple isomorphous replacement. *Methods Enzymol*., **276**, 538–557.

Romana,M., Grandchamp,B., Dubart,A., Amselem,S., Chabret,C., Nordmann,Y., Goossens,M. and Romeo,P.H. (1991) Identification of a new mutation responsible for hepatoerythropoietic porphyria. *Eur. J. Clin. Invest*., **21**, 225–229.

Romeo,P.-H., Raich,N., Dubart,A., Beaupain,D., Pryor,M., Kushner,J., Cohen-Solal,M. and Goossens,M. (1986) Molecular cloning and nucleotide sequence of a complete human uroporphyrinogen decarboxylase cDNA. *J. Biol. Chem*., **261**, 9825–9831.

Seki,Y., Kawanishi,S. and Sano,S. (1986) Uroporphyrinogen decarboxylase purification from chicken erythrocytes. *Methods Enzymol*., **123**, 415–421.

Sergeev,Y. and Lee,B. (1994) Alignment of beta-barrels in $(\beta/\alpha)_8$ proteins using hydrogen-bonding pattern. *J. Mol. Biol*., **244**, 168–182.

Straka,J.G., Kushner,J.P. and Pryor,M.A. (1982) Uroporphyrinogen decarboxylase: a method for measuring enzymatic activity. *Enzyme*, **28**, 170–185.

Wyckoff,E.E. and Kushner,J.P. (1994) Heme biosynthesis, the porphyrias, and the liver. In Arias,I.M., Boyer,J.L., Fausto,N., Jakoby,W.B., Schachter,D.A. and ShafritzD.A. (eds), *The Liver: Biology and Pathobiology*. Raven Press, Ltd, New York, pp. 505–527.

Wyckoff,E.E., Phillips,J.D., Sowa,A.M., Franklin,M.R. and Kushner,J.P. (1996) Mutational analysis of human uroporphyrinogen decarboxylase. *Biochim. Biophys. Acta*, **1298**, 294–304.

Received February 12, 1998; accepted March 4, 1998