Crystal structure of human uroporphyrinogen decarboxylase

Frank G. Whitby, John D. Phillips¹, James P. Kushner¹ and Christopher P. Hill²

Department of Biochemistry and ¹Department of Medicine, University of Utah School of Medicine, 50 N. Medical Drive, Salt Lake City, UT 84132, USA
²Corresponding 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 (β/α)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 (Kₐ = 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 enzymatic steps of the heme biosynthetic pathway (Figure 1) and its isomer uroporphyrinogen I, to produce coproporphyrinogen III (see Figure 1) and its isomer uroporphyrinogen I, to produce coproporphyrinogen III (Romeo et al., 1986). Purified recombinant human URO-D has been shown to be a homodimer with a Kₐ 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. 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 (β/α)-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); Rhodobacter capsulatus (U16796); Bacillus subtilis (M97208); Mycobacterium leprae (G699194); Caulobacter crescentus (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 (β/α)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.

![Image of URO-D crystal structure](image)

**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 4.0σ, and shown in yellow. 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.

**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 P3121 (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% (Rfree = 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% (Rfree = 28.9%) with good stereochemistry (r.m.s.d. bonds = 0.010 Å, r.m.s.d. angles = 2.049°). 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

<table>
<thead>
<tr>
<th>Wavelength (Å)</th>
<th>ScMet URO-Dt</th>
<th>URO-Dt</th>
<th>cutURO-Dt</th>
</tr>
</thead>
<tbody>
<tr>
<td>λ1</td>
<td>0.9795</td>
<td>1.54</td>
<td>0.98</td>
</tr>
<tr>
<td>λ2</td>
<td>0.9795</td>
<td>2.03</td>
<td>1.60</td>
</tr>
<tr>
<td>λ3</td>
<td>1.0688</td>
<td>2.20</td>
<td>1.60–1.63</td>
</tr>
<tr>
<td>λ4</td>
<td>0.9252</td>
<td>2.00</td>
<td>1.80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observed reflections</th>
<th>272 178</th>
<th>305 819</th>
<th>156 721</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unique reflections</td>
<td>29 817</td>
<td>60 487</td>
<td>42 721</td>
</tr>
<tr>
<td>dmin (Å)</td>
<td>2.03</td>
<td>2.00</td>
<td>1.60</td>
</tr>
<tr>
<td>High resolution shell</td>
<td>2.03–2.06</td>
<td>2.20–2.24</td>
<td>1.60–1.63</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>96.8 (65.4)</td>
<td>92.3 (67.3)</td>
<td>98.6 (85.2)</td>
</tr>
<tr>
<td>Rsym (%)</td>
<td>3.8 (8.4)</td>
<td>3.0 (5.3)</td>
<td>3.9 (17.9)</td>
</tr>
<tr>
<td>Average I/σ(I)</td>
<td>&gt;20 (9)</td>
<td>18 (13)</td>
<td>19 (2.1)</td>
</tr>
<tr>
<td>Mosaicity (°)</td>
<td>0.439</td>
<td>0.440</td>
<td>0.437</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rsym = 100*Σ</th>
<th>I–&lt;I&gt;</th>
<th>/ΣI.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crystals grown from MPD solution. Data collected at SSRL beamline 1–5 using Fuji image plates and a BAS2000 off-line scanner. Space group P3121, unit cell dimensions a = b = 103.3 Å, c = 74.3 Å.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crystals grown from MPD solution. Data collected from a rotating-anode X-ray source and an RAXIS-IV image plate detector. Space group P3121, unit cell dimensions a = b = 103.3 Å, c = 74.3 Å.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crystals grown from citrate solution. Data collected at SSRL beamline 9–1 using a MAR image plate detector. Space group P3121, unit cell dimensions a = b = 103.0 Å, c = 73.7 Å.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Values in parentheses refer to the highest resolution shell.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 (β/α)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 (β/α)₈ motif (Sergeev and Lee, 1994). Strands and helices of the (β/α)₈-barrel are named S1–S8.

### Table II. Refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>cutURO-D (citrate)</th>
<th>URO-Dt (MPD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range (Å)</td>
<td>20.0–1.80</td>
<td>20.0–1.60</td>
</tr>
<tr>
<td>No. of protein atoms</td>
<td>4254</td>
<td>4193</td>
</tr>
<tr>
<td>No. of solvent atoms</td>
<td>282</td>
<td>267</td>
</tr>
<tr>
<td>R-factor (%)</td>
<td>18.4</td>
<td>24.2</td>
</tr>
<tr>
<td>R_free (%)</td>
<td>23.3</td>
<td>28.9</td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.018</td>
<td>0.010</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>2.049</td>
<td>1.703</td>
</tr>
<tr>
<td>&lt;B&gt; (Å²) main chain</td>
<td>16.67</td>
<td>22.41</td>
</tr>
<tr>
<td>&lt;B&gt; (Å²) side chains</td>
<td>20.20</td>
<td>25.87</td>
</tr>
<tr>
<td>&lt;B&gt; (Å²) water molecules</td>
<td>26.96</td>
<td>35.39</td>
</tr>
<tr>
<td>&lt;B&gt; (Å²) β-ME</td>
<td>59.80</td>
<td>e</td>
</tr>
<tr>
<td>No. of ϕ/ψ angles (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>most favored</td>
<td>92.4</td>
<td></td>
</tr>
<tr>
<td>additional</td>
<td>6.3</td>
<td></td>
</tr>
</tbody>
</table>

*Non-hydrogen atoms only.

bR-factor = 100*Σ(|F(obs)|−|F(calc)|)/Σ|F(obs)|. All data were used in the resolution range indicated, without application of a cut based upon the estimated standard deviation.

R_free = R-factor for a selected subset (2.5–3.5%) of the reflections which were not included in prior refinement calculations.

Stereochemistry was assessed with PROCHECK (Laskowski et al., 1993).

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.](image-url)
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 (β/α)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
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 $\sim 0.1 \, \mu M$ (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 ($<B> = 16.9 \, \text{Å}^2$) 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 ($B < 22.0 \, \text{Å}^2$ 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.
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 sulphydryl-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 over-expressed 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 CaCl2 at room temperature for 24 h.
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 uncleaved 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-D 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 5.6–5.6. 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 offline 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 P31 21 with cell dimensions that range from α = 103.0–103.3 Å and c = 73.7–74.3 Å. There is one molecule in the asymmetric unit, and the Matthews’ coefficient (V_m) of 2.7 Å³/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 offline 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 Blow, 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-Dt was built using the program O (Jones et al., 1991) and refined against the SeURO-Dt data with XPLOR (Brünger, 1992). 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 Å²) 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 of the C-terminal 31 residues of URO-Dt. 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-D crystal grown in MPD. 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 18.2% and R_free is 24.2% for all data to 1.80 Å resolution with good stereochemistry (see Table II).
References


Received February 12, 1998; accepted March 4, 1998