TABLE I
Crystallographic data and refinement

Values in parentheses refer to the high-resolution shell.

	CPO crystal form			
	C	C	I	II
Crystal				
Space group	$P4_12_12$	$P4_12_12$	$P4_32_12$	C2
Cell dimensions (Å)	a = 71.8, c = 118.1	a = 70.9, c = 114.5	a = 86.8, c = 207.8	a = 233.7, b = 65.4, c = 166.5, β = 108.1°
mol/asymmetric unit	1	1	2	6
Crystallographic data				
Wavelength (Å)	2.29	1.10	1.10	1.5418
Data collection[a]	RAXIS-Cr	NSLS-X8C	NSLS-X25	RAXIS-Cu
Data processing[b]	d*TREK	DENZO	DENZO	DENZO
Resolution (Å)	38.5–2.50 (2.59–2.50)	30.0–1.90 (1.97–1.90)	30.0–2.00 (2.07–2.00)	30.0–2.40 (2.55–2.40)
No. reflections measured	222,966	661,152	486,505	478,664
No. unique reflections	20,453	23,022	53,840	87,606
Complete (%)	99.9 (99.3)	97.0 (95.6)	98.4 (97.5)	93.9 (90.7)
$\langle I/\sigma I\rangle$	24.3 (4.2)	>20 (6.0)	19 (3.2)	15 (2.5)
Mosaicity	0.40°	0.58°	0.47°	0.83°
R_{sym} (%)[c]	0.057 (0.383)	0.072 (0.580)	0.057 (0.583)	0.077 (0.309)
Refinement statistics				
Resolution (Å)		30.0–1.90 (1.97–1.90)	30.0–2.00 (2.05–2.00)	30.0–2.40 (2.46–2.40)
R_{cryst} (%)[d]		0.228 (0.251)	0.207 (0.258)	0.208 (0.248)
R_{free}(%)[e]		0.231 (0.290)	0.254 (0.301)	0.282 (0.350)
Protein residues		8–265[f]	4–328[g]	3–328[g]
No. solvent molecules		96	418	639[h]
r.m.s.d. bonds (Å)		0.011	0.016	0.021
r.m.s.d. angles		1.441°	1.507°	1.832°
ϕ/ψ angles[i]				
Most favored (%)		91.3	94.0	90.1
Additional allowed (%)		7.1	6.0	9.6
⟨B⟩ protein (Å^2)		29.8	42.7	47.2
⟨B⟩ main chain (Å^2)		29.2	42.4	47.1
⟨B⟩ water (Å^2)		40.3	43.4	38.9

[a] Data were collected on Rigaku RAXIS-IV detectors mounted on rotating anode sources with chromium (RAXIS-Cr) and copper (RAXIS-Cu) targets. Other data were collected on ADSC Quantum-4 CCD area detectors at beamlines X8C and X25 at the Brookhaven National Laboratory National Synchrotron Light Source (NSLS).

[b] Data processing was performed using DENZO/SCALEPACK (44) or d*TREK (45).

[c] $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the intensity of an individual measurement and $\langle I\rangle$ is the corresponding mean value.

[d] $R_{cryst} = \Sigma||F_o - F_c||/\Sigma|F_o|$, where $|F_o|$ is the observed and $|F_c|$ is the calculated structure factor amplitude.

[e] R_{free} is the same as R_{cryst}, calculated with a randomly selected test set of reflections (5% of the total) that were never used in refinement calculations.

[f] 219 residues were modeled. Disordered regions of the protein that could not be modeled are postitions 42–50, 92–106, and 192–206.

[g] The full-length protein has been modeled as a single chain in all molecules in the asymmetric unit. For form II, residues 91–110 have only diffuse density and were assigned zero occupancies.

[h] This includes one sulfate ion.

[i] Values are for non-Gly and non-Pro residues only.

crystal and from a form I crystal at the National Synchrotron Light Source. Data were collected from a single form II crystal using an imaging plate area detector on a rotating anode x-ray source with a copper anode. Crystallographic statistics are given in Table I.

Structure Determination and Refinement—The form C crystal structure was determined by sulfur single wavelength anomalous diffraction phasing. SOLVE (23) was used to locate seven sulfur positions using 3.3-Å data and to calculate phases based on this solution. The M value was 0.27, and the overall Z score was 19. Density modification was performed with RESOLVE (24) to produce an electron density map that allowed building of several strands and helices. Phases from this refined partial model were used to calculate an anomalous difference Fourier map that confirmed the sulfur substructure and found an eighth site. SOLVE and RESOLVE were rerun, fixing the eight sulfur positions, to give a slightly improved map (see Fig. 2). Refinement and rebuilding using 1.9-Å data collected at National Synchrotron Light Source beamline X8C resulted in a model of good stereochemistry and agreement with x-ray terms (see Table I).

The form I and II crystal structures were determined by molecular replacement using the program MOLREP (25) with the CCP4i interface (26). The refined form C model was used to determine the form I structure, which has two molecules in the asymmetric unit. The refined form I model was subsequently used to determine the form II structure, which has six molecules in the asymmetric unit. In both cases, model building with the program O (27) and refinement with REFMAC5 (28) were straightforward (see Table I).

RESULTS AND DISCUSSION

Structure Determination—Crystals of histidine-tagged yeast odCPO/Hem13p were grown in three different space groups. One of these crystal forms is a truncated protein that underwent limited proteolysis during crystallization to yield a protein that started at Asp^6, as indicated by N-terminal sequencing of protein from a washed crystal. SDS-PAGE analysis gave a molecular mass of ~30 kDa, suggesting that protein in these crystals was also cleaved at the C terminus to remove the last ~60–70 residues. We refer to these crystals as form C (cleaved). Protein from the other two crystal forms (forms I and II) migrated similarly to full-length odCPO/Hem13p on SDS-polyacrylamide gel. N-terminal sequencing indicated that form I crystals were processed at the N terminus to start at Ala^3. This analysis was not performed on form II crystals.

We were unable to prepare crystals of selenomethionine-substituted odCPO/Hem13p that diffracted with sufficient strength to allow reliable phase determination. Structure determination was therefore approached using intrinsic anomalous scattering from the eight sulfur atoms in the truncated molecule of form C crystals. This crystal form offered the advantage of a relatively high fraction of sulfur atoms. Data were