Proc. Nat. Acad. Sci. USA
Vol. 69, No. 11, pp. 3185-3188, November 1972

Structure of the Oxidized Form of a Flavodoxin at 2.5-Å Resolution: Resolution of the Phase Ambiguity by Anomalous Scattering

(protein structure/flavin mononucleotide/single isomorphous derivative/x-ray diffraction)

K. D. WATENPAUGH*, L. C. SIEKER*, L. H. JENSEN*, J. LEGALL†‡, AND M. DUBOURDIEU†

* Department of Biological Structure, University of Washington, Seattle, Wash. 98195; and † Laboratoire de Chimie Bacterienne, C.N.R.S., Marseilles, France

Communicated by Hans Neurath, August 21, 1972

ABSTRACT **Flavodoxin from *Desulfovibrio vulgaris* crystallizes in the oxidized form as well-formed, tetragonal bipyramids, space group $P4_32_12$, unit-cell parameters, $a = b = 51.6$ Å, $c = 139.6$ Å, 8 molecules per unit cell.**

The structure has been determined at 2.5-Å resolution with phases based on a single isomorphous derivative. The phase ambiguity of a single derivative was resolved by use of anomalous scattering from the single-site Sm^{+3}.

The molecule has a five-strand pleated sheet core with two long helices on either side of the sheet. The flavin mononucleotide lies mostly buried on one side of the molecule, but the methyl groups, one edge of the flavin, and part of the ribityl are exposed at the surface.

Originally, the name flavodoxin was given by Knight *et al.* (1) to a small flavoprotein found in extracts of *Clostridium pasteurianum* grown in iron-deficient medium. The protein was able to replace ferredoxin as an electron carrier (2, 3). After this discovery, other organisms have been shown to contain a similar flavoprotein: *Desulfovibrio gigas* (4), *Peptostreptococcus elsdenii* (5), *Desulfovibrio vulgaris* (6), *Clostridium* MP (7), *Rhodospirillum rubrum* (8), *Chlorella fusca* (9), and *Escherichia coli* (10).

Flavodoxins have been defined as low molecular weight proteins with one flavin mononucleotide prosthetic group per molecule and having the ability to function interchangeably with ferredoxin in photosynthetic $NADP^+$ reduction (11).

In *Desulfovibrio*, flavodoxin can replace ferredoxin in the reduction of sulfite and other sulfur compounds (4, 12, 13), and in the formation of hydrogen from pyruvate via the phosphoroclastic reaction (14).

When thiosulfate is reduced by molecular hydrogen in the presence of *Desulfovibrio* extracts that have been depleted of flavodoxin, ferredoxin, and cytochrome cc_3, any one of these three electron carriers can couple the reaction (13, 15). Since the three electron carriers are synthesized simultaneously by the cells, proteins with either flavin mononucleotide (flavodoxin), nonheme iron (ferredoxin), or heme iron (cytochrome cc_3) as prosthetic group are carriers in the same reaction.

Since these three electron carriers can be obtained in crystalline form, it is hoped that their structures can be established as a basis of better understanding their mechanism of action. Four of the five *Desulfovibrio* species, namely *D. gigas*, *D. sulfuricans*, *D. salexigens*, and *D. vulgaris*, have been examined and found to contain flavodoxin. Thus far, only *D. vulgaris*, strain Hildenborough, N.C.I.B. 8303, has given crystals suitable for x-ray studies. We report here the first crystallographic results on the structure of a *Desulfovibrio* flavodoxin. The following paper in this issue by Andersen *et al.* compares certain features of the flavodoxin from *Clostridium* MP with the flavodoxin reported here.

† Also at Department of Biochemistry, University of Georgia, Athens, Ga.

EXPERIMENTAL

Crystallization

Conditions for growing the bacterium *Desulfovibrio vulgaris* and the steps for the purification of flavodoxin have been reported (4, 6). It crystallizes readily from 3.5 M solution of $(NH_4)_2SO_4$ over a pH range of 6.0–10.0. Crystallization appears to be unaffected by the presence of different buffering reagents, and crystals have been obtained from solutions free of buffer or with acetate, phosphate, Tris·HCl, or glycinate. The crystals used for data collection were routinely grown from 3.5 M $(NH_4)_2SO_4$ buffered with 0.1 M Tris·HCl at pH 7.5 and protein concentrations in the range 0.6–1.0%. Crystals develop as well-formed, elongated tetragonal bipyramids, easily growing to sizes with the major dimension up to 1–2 mm.

Crystals of flavodoxin are tetragonal, space group $P4_12_12$ or $P4_32_12$. The unit-cell dimensions are $a = b = 51.6$ Å, $c = 139.6$ Å based on $\lambda_{CuK\alpha} = 1.5418$ Å. The unit-cell volume is 372,000 Å^3 with eight asymmetric units of 46,500 Å^3 per cell. If we assume a molecular weight of 16,000, the volume per dalton is 2.9, indicating rather high solvent content (16).

Heavy atom derivatives

All attempts to obtain heavy atom derivatives were made by soaking flavodoxin crystals in a solution 3.7 M in $(NH_4)_2SO_4$ buffered with 0.05 M Tris·HCl or phosphate at pH 7.5.

Four heavy-atom reagents, K_2PtCl_4, $UO_2(NO_3)_2$, $Yb_2(SO_4)_3$, and $Sm(NO_3)_3$, were found to bind sufficiently that intensity differences could be observed in the diffraction pattern. 0.01 M K_2PtCl_4–phosphate buffer, pH 7.5, gave good intensity changes, but the difference Patterson map indicated light substitution. Since K_2PtCl_4 disordered the crystals in higher concentrations, that derivative has not been pursued further. 4–10 mM $UO_2(NO_3)_2$ in Tris·HCl (pH 7.5) gave a highly substituted derivative, but thus far the UO_2^{+2} sites have not been located with certainty.

Both $Yb_2(SO_4)_3$ and $Sm(NO_3)_3$ gave identical single-site derivatives when crystals were soaked for 9–10 days in Tris·HCl buffer solution, 0.04 M in heavy-atom reagent. Since anomalous scattering for CuK_α radiation is much greater for

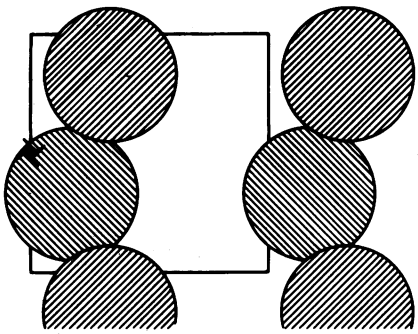

FIG. 1. Diagram showing packing of molecules in slice z = 0–35/140.

Sm^{+3} than for Yb^{+3}, the Sm^{3+} derivative was used in this work, and to this point only about 50 mg of the protein has been required.

Data collection

Data were collected on a four-circle, computer-controlled diffractometer with a modified version of the control program of Lenhert and Henry (17). Ni-filtered CuK_α radiation was used with a focal spot 0.4 × 10 mm and a take-off angle set at 6°. For both the native crystal and the Sm^{3+}-derivative crystal, data were collected in the same way by a five-step $\omega/2\theta$ scan across the top of each reflection. Each step was 0.08° in 2θ.

Backgrounds were not measured for the individual reflections. Instead, the background was determined as a function of 2θ by collecting counts at many points between reflections and assuming it to be independent of direction in space. The step scan used to collect data and the method of approximating the background was made necessary by the long c axis.

Friedel-related reflections were collected by alternately measuring 32 reflections at $+2\theta$, followed by 32 reflections at -2θ. The groups were staggered, however, so that the reflections of each Friedel pair were collected 16 reflections apart.

A group of three monitor reflections were collected every 200 reflections or approximately every 1.5 hr to check crystal alignment and equipment malfunction. Ten standard reflections were collected every 2000 reflections, and these were used to correct for crystal deterioration.

The five steps for each reflection were fit by least-squares to the Gaussian,

$$y_i = C_1 e^{-C_3(X_i - C_2)^2},$$

where C_1 = peak height, C_2 = peak position, C_3 = peak width, y_i = count for the ith step, and $x_i = 2\theta$ for the ith step. Both C_1 and C_2 were refined, but C_3 was fixed at a value determined from the standard reflections. Because of the absence of heavily absorbing material and the near ideal shape of the crystal, the absorption correction was small and was not applied to the native or to the Sm^{3+}-derivative data.

Over 14,000 reflections were collected to a $\sin\theta/\lambda$ of $0.2 (d$ = 2.5 Å) for one native and one derivative crystal. Total deterioration of the standard reflections was 29% for the native crystal and 25% for the Sm^{+3}-derivative crystal. For the 7151 Friedel pairs from the native crystal, the relative deviation from the mean defined by the equation,

$$D = \sum||F_{\text{mean}}| - |F(\pm)||/\sum|F_{\text{mean}}|$$

was 0.016. The set of native data was taken as $I_{hkl} = (1/2)(I_{hkl} + I_{\bar{h}\bar{k}\bar{l}})$.

The 1676 Friedel pairs of centrosymmetric reflections for the Sm^{+3}-derivative data were also averaged. For these reflections, D was 0.018.

Location of heavy atoms and phasing of the data

The native and the Sm^{+3}-derivative data were placed on the same relative scale by the method described by Singh and Ramaseshan (18). No difference of fall off in intensity or anomalous contribution between the two sets of data as a function of $\sin\theta/\lambda$ was observed, so that no correction for an over-all difference in temperature factor was necessary. The Sm^{+3} was located by a difference Patterson synthesis in which the coefficients F_M were calculated according to the expression given by Matthews (19), which takes into account the anomalous difference. The Patterson synthesis with F_M as coefficients was very clean, the only prominent feature being the large peaks from the single Sm^{+3} site. This site was refined against the F_M coefficients by full-matrix least-squares, giving a conventional R of 0.41.

A difference Fourier synthesis was calculated to check for any minor Sm^{+3} sites. This map confirmed the conclusion

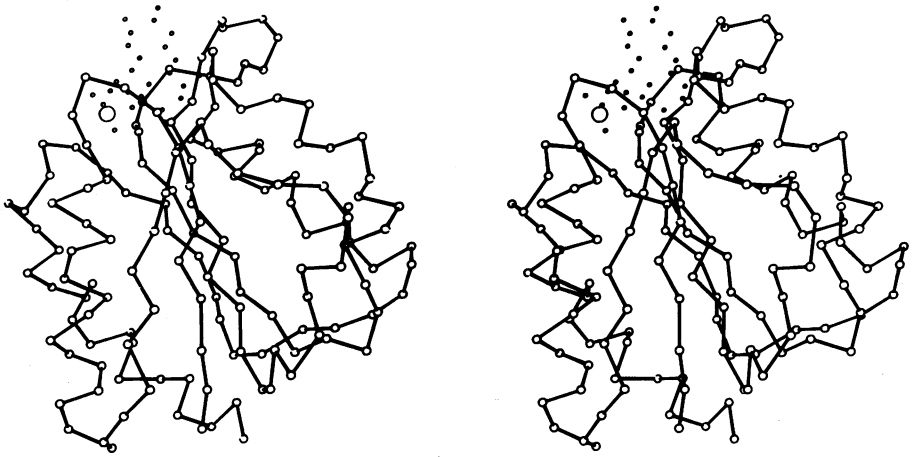

FIG. 2. Stereo view of molecule along pleated sheet and normal to c.

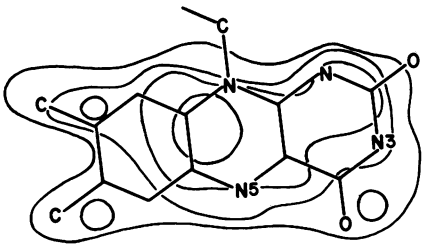

FIG. 3. Electron density in plane of flavin group.

drawn from the Patterson map that there was only one significant Sm^{+3} site. The refined parameters in space group $P4_12_12$ are

$$x = 0.230,\ y = 0.359,\ z = 0.019,\ B = 12.5\ \text{Å}^2.$$

Phase determination of the native crystal data was performed with the single-site Sm^{+3} derivative according to the method of Blow and Crick (20). Use of anomalous scattering to resolve the phase ambiguity of a single heavy-atom protein derivative was suggested by Blow and Rossman (21), and has been successful in phasing several proteins (22–24). Two possible sets of phases can be calculated for the native flavodoxin data when anomalous scattering is used to resolve the phase ambiguity, one in space group $P4_12_12$, the other in space group $P4_32_12$. Table 1 contains summary criteria for the phase determination.

A Fourier synthesis with observed native magnitudes weighted by their figures of merit and the "best" or centroid phases was calculated in space group $P4_32_12$. A long section of helix and a section of pleated sheet were readily apparent, indicating that the correct space group and enantiomorph had been chosen. A large residual peak appeared, however, at the Sm^{+3} position, suggesting an error either in the scale factor between the Sm^{+3} occupancy and the derivative magnitudes, or between the derivatives and the protein magnitudes. Accordingly, the scale factors of the protein and the derivative magnitudes were refined only with the centrosymmetric reflections by a least-squares fit of

$$f_{Sm^{+3}} = (k_{Sm^{+3}})(F_{Sm^{+3}}) - (k_{Fl})(F_{Fl})$$

where $f_{Sm^{+3}}$ = magnitude and phase of the Sm^{+3} scattering, $F_{Sm^{+3}}$ = magnitude and phase of the Sm^{+3} derivative data, F_{Fl} = magnitude and phase of the native data, and $k_{Sm^{+3}}$ and k_{Fl} = scale factors to be adjusted. This gave new scale factors $k_{Sm^{+3}} = 0.982$ and $k_{Fl} = 0.99$ relative to the previous values of 1.0.

The phase determination was again carried through, but with the new scales. The R factors did not change significantly, but the new electron-density map showed a large decrease in the peak at the Sm^{+3} site, approximately to the density of the main chain atoms.

Description of the structure

The molecules of flavodoxin are oblate spheroids of approximate dimensions 25 Å × 40 Å × 40 Å packed in the crystal lattice as shown in Fig. 1. They are arranged side by side in a zig-zag, rod-like fashion with adjacent parallel rods separated by large solvent channels. Sheets of rods are stacked, with the rods in successive layers alternately parallel to the x and the y axes.

The core of the molecule is a five-strand pleated sheet.

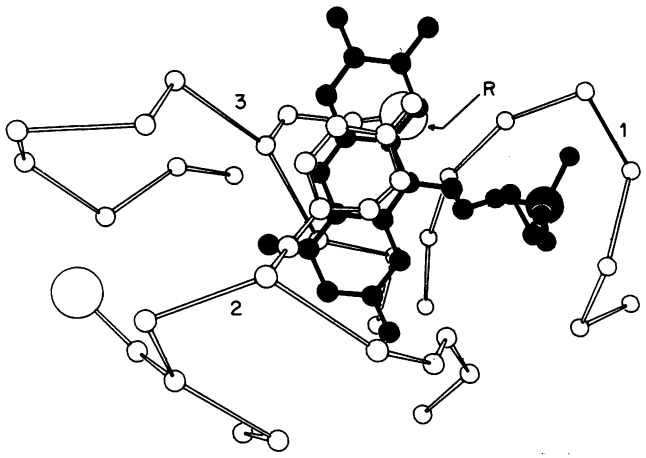

FIG. 4. Plot of main-chain C_α atoms and some side chains (*open circles and bonds*) and the flavin mononucleotide group (*solid circles and bonds*). Direction of view normal to plane of flavin. Three pieces of main chain are shown: a section looped around the phosphate, a section with what appears to be a Tyr on one side of the flavin, and a section with another large group labeled R on the opposite side of the flavin.

Four long helices, two on either side of the pleated sheet, are arranged as shown in Fig. 2. About one-third of the residues are involved in the pleated sheet, one-third in the helices, and one-third in extended chain or other configurations.

The flavin mononucleotide is distinguishable from the protein and lies mostly buried near the surface of the molecule, Fig. 2. Inspection of the electron-density map shows, however, that the methyl groups and one edge of the flavin, along with part of the ribityl group, are exposed at the surface.

The flavin appears to be planar, as expected, and Fig. 3 shows the electron density in a plane through it, with a skeleton model superposed. The group is tucked in between and parallel to what appears to be an aromatic group (probably

TABLE 1. *Summary criteria of phase determination*

R_1	R_2	R_K	R_C	$\langle m \rangle$	n
0.415	0.075	0.144	0.607	0.67	6354

$$R_1 = \sum_{hkl} \big| |F_H| - |f_H| \big| / \sum_{hkl} |F_H|$$

where F_H is the observed heavy-atom contribution calculated according to Matthews (17); f_H is the calculated structure factor of heavy atoms.

$$R_2 = \sum_{hkl} \big| |F_{PH}| - |F_P + f_H| \big| / \sum_{hkl} |F_{PH}|$$

where F_{PH} is the observed structure factor of derivative, F_P is the observed structure factor of native protein.

$$R_K = \sum_{hkl} \big| |F_{PH}| - |F_P + f_H| \big| / \sum_{hkl} |F_{PH}|$$

for data from centrosymmetric projections.

$$R_C = \sum_{hkl} \big| |F_{PH}| - |F_P + f_H| \big| / \sum_{hkl} \big| |F_{PH}| - |F_P| \big|$$

for data from centrosymmetric projections.

$\langle m \rangle$ = mean figure of merit.

n = number of observed reflections.

tyrosine) on one side and another large group (possibly aromatic) on the opposite side (Fig. 4). The two carbonyl oxygen atoms on the pyrimidine ring of the flavin are hydrogen bonded to the protein. N5 does not appear to be hydrogen bonded either to the protein or to any other part of the flavin mononucleotide (25).

The ribityl group is clearly visible connecting the flavin and the phosphate. In this region of the map, the electron density is sufficiently good that three bulges on the ribityl density appear to correspond to 02′, 03′, and 04′.

The phosphate position corresponds to the largest peak in the electron-density map. It is buried most deeply in the protein and is bent away from the flavin, rather than back toward it (26). Each of the three phosphate oxygen atoms (excluding the one bonded to C5′) appears to be hydrogen bonded to different R groups in the protein.

The sequence of this flavodoxin has not yet been completed, and most *R* groups cannot be identified with certainty from the electron-density map. It appears, however, that none of the cysteine residues are involved in disulfide bridges or in bonding to flavin mononucleotide.

We thank Professors J. C. Senez and E. H. Fischer through whom this collaborative effort was arranged. We are indebted to J. M. Stewart for the X-ray System programs and to C. K. Johnson for the ORTEP program. This work was supported by USPHS Grant AM-3288 from the National Institutes of Health and Grant GB-29618X from the National Science Foundation.

1. Knight, E., Jr., D'Eustachio, A. J. & Hardy, R. W. F. (1966) *Biochim. Biophys. Acta* **113,** 626–628.
2. Knight, E., Jr. & Hardy, R. W. F. (1966) *J. Biol. Chem.* **241,** 2752–2756.
3. Knight, E., Jr. & Hardy, R. W. F. (1967) *J. Biol. Chem.* **242,** 1370–1374.
4. LeGall, J. & Hatchikian, E. C. (1967) *C. R. H. Acad. Sci.* **264,** 2580–2583.
5. Mayhew, S. G. & Massey, V. (1969) *J. Biol. Chem.* **244,** 794–802.
6. Dubourdieu, M. & LeGall, J. (1970) *Biochem. Biophys. Res. Commun.* **38,** 965–972.
7. Mayhew, S. G. (1971) *Biochim. Biophys. Acta* **235,** 278–288.
8. Cusanovich, M. A. & Edmonson, D. E. (1971) *Biochem. Biophys. Res. Commun.* **45,** 327–335.
9. Zumft, W. G. & Spiller, H. (1971) *Biochem. Biophys. Res. Commun.* **45,** 112–118.
10. Vetter, H. & Knappe, J. (1971) *Z. Physiol. Chem.* **352,** 433–446.
11. Van Lin, B. & Bothe, H. (1972) *Arch. Mikrobiol.* **82,** 155–172.
12. Guarraia, L. J., Laishley, E. J., Forget, N. & Peck, H. D., Jr. (1968) *Bacteriol. Proc.* 133.
13. Hatchikian, E. C., LeGall, J., Bruschi, M. & Dubourdieu, M. (1972) *Biochim. Biophys. Acta* **258,** 701–708.
14. Hatchikian, E. C. & LeGall, J. (1970) *Ann. Inst. Pasteur Paris* **118,** 288–301.
15. Bruschi, M., LeGall, J., Hatchikian, E. C. & Dubourdieu, M. (1969) *Bull. Soc. Fr. Physiol. Veg.* **15,** 381–390.
16. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497.
17. Lenhert, P. G. & Henry, D. (1970) *Amer. Crystallogr. Ass., Abstracts of Winter Meeting,* Tulane University, New Orleans, p. 71.
18. Singh, A. K. & Ramaseshan, S. (1966) *Acta Crystallogr.* **21,** 279–280.
19. Matthews, B. W. (1966) *Acta Crystallogr.* **20,** 230–239.
20. Blow, D. M. & Crick, F. H. C. (1959) *Acta Crystallogr.* **12,** 794–802.
21. Blow, D. M. & Rossman, M. G. (1961) *Acta Crystallogr.* **14,** 1195–1202.
22. Herriott, J. R., Sieker, L. C., Jensen, L. H. & Lovenberg, W. (1970) *J. Mol. Biol.* **50,** 391–406.
23. Arnone, A., Brier, C. J., Cotton, F. A., Day, V. W., Hazen, E. E., Richardson, D. C., Richardson, J. S. and in part, Yonath, A. (1971) *J. Biol. Chem.* **246,** 2302–2316.
24. Sieker, L. C., Adman, A. T. & Jensen, L. H. (1972) *Nature* **235,** 40–42.
25. Edmondson, D. E. & Tollin, G. (1971) *Biochemistry* **10,** 133–145.
26. Ludwig, M. L., Andersen, R. D., Apgar, P. A., Burnett, R. M., LeQuesne, M. E. & Mayhew, R. M. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **35,** 369–385.