FILIPPINI, G., GRAMACCIOLI, C. M., SIMONETTA, M. & SUFFRITTI, G. B. (1973). *J. Chem. Phys.* **59**, 5088–5101.
FILIPPINI, G., GRAMACCIOLI, C. M., SIMONETTA, M. & SUFFRITTI, G. B. (1974). *Acta Cryst.* A**30**, 189–196.
GOURSOT, P., GIRDHAR, H. L. & WESTRUM, E. F. (1968). *C. R. Acad. Sci. Ser. C*, **266**, 949–950.
GRAMACCIOLI, C. M., FILIPPINI, G., SIMONETTA, M., RAMDAS, S., PARKINSON, G. M. & THOMAS, J. M. (1980). *J. Chem. Soc. Faraday Trans.* 2, **76**, 1336–1346.
HIRSHFELD, F. L., HOPE, H. & RABINOVICH, D. (1990). In preparation.
JANKOWIAK, R., BÄSSLER, H. & KUTOGLU, A. (1985). *J. Phys. Chem.* **89**, 5705–5709.
JINDAL, V. K., KALUS, J., BOKHENDOV, E. L., CHAPLOT, S. L., DORNER, B., NATKANIEC, I., PAWLEY, G. S. & SHEKA, E. F. (1982). *J. Phys. C*, 7283–7294.
LEHMANN, M. S. & PAWLEY, G. S. (1972). *Acta Chem. Scand.* **26**, 1996–2004.
MASON, R. (1964). *Acta Cryst.* **17**, 547–555.
MATHIESON, A. MCL., ROBERTSON, J. M. & SINCLAIR, V. C. (1950). *Acta Cryst.* **3**, 245–250.
MIRSKY, K. & COHEN, M. (1978). *Chem. Phys. Lett.* **54**, 40–41.
NEERLAND, G., CYVIN, B. N., BRUNVOLL, J., CYVIN, S. J. & KLAEBOE, P. (1980). *Z. Naturforsch Teil A*, **35**, 1390–1394.
NICOL, M., VERNON, M. & WOO, J. T. (1975). *J. Chem. Phys.* **63**, 1992–1999.
PARKINSON, G. M., GORINGE, M. J., RAMDAS, S., WILLIAMS, J. O. & THOMAS, J. M. (1978). *J. Chem. Soc. Chem. Commun.* pp. 134–135.
PAULING, L. (1960). *The Nature of the Chemical Bond*, 3rd ed. Ithaca: Cornell Univ. Press.
PAULING, L. (1980). *Acta Cryst.* B**36**, 1898–1901.
PAWLEY, G. S. (1967). *Phys. Status Solidi*, **20**, 347–360.
PONOMAREV, V. I. & SHILOV, G. V. (1983). *Sov. Phys. Crystallogr.* **28**, 397–399.
PUFALL, R. & KALUS, J. (1988). *Acta Cryst.* A**44**, 1059–1065.
RAMDAS, S., PARKINSON, G. M., THOMAS, J. M., GRAMACCIOLI, C. M., FILIPPINI, G., SIMONETTA, M. & GORINGE, M. J. (1980). *Nature (London)*, **284**, 153–154.
ROBERTSON, J. M. (1933). *Proc. R. Soc. London Ser. A*, **140**, 79–98.
ROSENFIELD, R. E. JR, TRUEBLOOD, K. N. & DUNITZ, J. D. (1978). *Acta Cryst.* A**34**, 828–829.
ROSS, R. G., ANDERSSON, P. & BÄCKSTRÖM, G. (1979). *Mol. Phys.* **38**, 527–533.
RYZHENKOV, A. I., KOZHIN, V. M. & MYASNIKOVA, R. M. (1969). *Sov. Phys. Crystallogr.* **13**, 896–898.
SEILER, P., SCHWEIZER, W. B. & DUNITZ, J. D. (1984). *Acta Cryst.* B**40**, 319–327.
SINCLAIR, V. C., ROBERTSON, J. M. & MATHIESON, A. MCL. (1950). *Acta Cryst.* **3**, 251–256.
SPARKS, R. A. (1958). PhD Thesis, Univ. of California.
STEWART, J. M., KRUGER, G. J., AMMON, H. L., DICKINSON, C. & HALL, S. R. (1972). The *XRAY*72 system – version of June 1972. Tech. Rep. TR-192. Computer Science Center, Univ. of Maryland, College Park, Maryland, USA.
SUZUKI, M., YOKOYAMA, T. & ITO, M. (1968). *Spectrochim. Acta*, **24A**, 1091–1107.
TRUEBLOOD, K. N. (1978). *Acta Cryst.* A**34**, 950–954.
VOVELLE, F., CHEDIN, M.-P. & DUMAS, G. G. (1978). *Mol. Cryst. Liq. Cryst.* **48**, 261–271.

Acta Cryst. (1990). B**46**, 806–823

Structure Determination of Quinoprotein Methylamine Dehydrogenase from *Thiobacillus versutus*

BY F. M. D. VELLIEUX, K. H. KALK, J. DRENTH AND W. G. J. HOL

BIOSON Research Institute, Department of Chemistry, University of Groningen, Nijenborgh 16, 9747 *AG Groningen, The Netherlands*

(*Received* 2 *March* 1990; *accepted* 31 *May* 1990)

Abstract

The crystal structure of quinoprotein methylamine dehydrogenase from *Thiobacillus versutus* (EC 1.4.99.3, $M_r = 123\,500$) has been solved to 2·25 Å resolution. The crystals of space group $P3_121$ ($a = b = 129{\cdot}8$, $c = 104{\cdot}3$ Å) contain half a tetrameric enzyme molecule in the asymmetric unit, with a solvent content of *ca* 70%. The procedure used to solve this structure involved multiple isomorphous-replacement phasing, complemented by phase extension using solvent flattening, and phase combination with partial-model phases. The use of solvent flattening was essential to generate good quality electron density maps into which initial models were built. These partial models were refined using molecular-dynamics procedures. Refined model phases were then combined with solvent-flattening phases to generate improved electron density distributions. In the absence of an amino-acid sequence for this enzyme, the current 2·25 Å resolution electron density map was interpreted to provide a model for the complete molecule. The crystallographic *R* factor for this model, which lacks any water molecules, is 28·6% for data between 6·0 and 2·25 Å resolution.

Introduction

Bacterial methylamine dehydrogenase (abbreviated as MADH, EC 1.4.99.3) is a periplasmic enzyme which contains a pyrroloquinoline quinone (PQQ)

0108-7681/90/060806-18$03.00

related cofactor (de Beer, Duine, Frank & Large, 1980; van der Meer, Jongejan & Duine, 1987). Thus, MADH is a member of the recently discovered class of oxidoreductase enzymes known as quinoproteins (Duine, Frank & Verwiel 1980). Similar cofactors have been detected in numerous bacterial enzymes (for a review, see Duine, 1989), and have been discovered more recently in several mammalian enzymes, including human placental lysyl oxidase (van der Meer & Duine, 1986). Also of interest was the discovery of PQQ as the cofactor of soybean lipoxygenase (van der Meer & Duine, 1988). In view of these recent developments, it seems that quinoproteins are widely distributed among most, if not all, living organisms.

On the basis of the type of linkage of the quinone cofactor to the enzyme, there appear to be two distinct classes of quinoproteins: first, those containing a non-covalently bound prosthetic group, an example of which is bacterial glucose dehydrogenase (Duine, Frank & van Zeeland, 1979), and secondly, those where the quinone cofactor is covalently bound to the protein. It is interesting to note that all mammalian quinoproteins investigated so far contain such a covalently bound cofactor. Bearing this in mind, bacterial MADH could be an interesting model for studying the precise mode of binding of its PQQ-related cofactor, since the latter is also covalently bound to the protein (Ishii, Hase, Fukumori, Matsubara & Tobari, 1983).

A number of methylamine dehydrogenases have been isolated and characterized from different bacterial strains. All are very similar in their physicochemical and enzymatic properties (Vellieux *et al.*, 1986). They catalyze the oxidative deamination of primary alkylamines to the corresponding aldehydes (Eady & Large, 1968) according to:

$$R\mathrm{CH_2NH_3^+} + \mathrm{H_2O} \rightarrow R\mathrm{CHO} + \mathrm{NH_4^+} + 2\mathrm{H^+} + 2\mathrm{e^-}$$

All MADH's characterized so far are tetrameric and composed of two identical heavy (H) and two identical light (L) subunits. Each light subunit contains one quinone cofactor which is covalently linked to two side chains of the subunit (Ishii *et al.*, 1983).

We investigated the three-dimensional structure of methylamine dehydrogenase from *Thiobacillus versutus* in order to shed light on the exact nature of its quinone cofactor (de Beer *et al.*, 1980; Kenney & McIntire, 1983; McIntire & Stults, 1986), as well as to understand the reaction mechanism of this class of enzymes. Molecular mass determination (Vellieux *et al.*, 1986) gave $M_r = 123\,500$ for the complete MADH molecule, with the H and L subunits having M_r's of 47 500 and 12 900. At present, neither the amino-acid sequence of the heavy subunit, nor that of the light subunit are known, although investigations aimed at determining the complete sequence of this enzyme *via* cDNA are in progress (J. A. Duine & J. Frank, personal communication). The *N*-terminal amino-acid sequence of the L subunit of the *T. versutus* enzyme (F. Huitema, J. A. Duine, H. J. Bak & J. J. Beintema, unpublished results) shows a high degree of homology with the known sequence of the L subunit of the *Pseudomonas AM*1 enzyme (Ishii *et al.*, 1983).

Our crystallographic investigations have been carried out with MADH crystals containing the cofactor in a free-radical semiquinone form (de Beer *et al.*, 1980; Vellieux *et al.*, 1986). The crystal structure determination relied on the successful combination of: (i) multiple isomorphous replacement; (ii) phase improvement and addition by solvent-flattening procedures; and (iii) partial-structure phasing. An account of the three-dimensional structure of methylamine dehydrogenase from *T. versutus* has been published elsewhere (Vellieux *et al.*, 1989), and in this paper we wish to emphasize the structure determination procedure that was used. A few aspects of the structure will also be discussed.

Materials and methods

Crystallization

Large single crystals of MADH were grown from hanging-drop experiments as described before (Vellieux *et al.*, 1986, 1989). These crystals diffract to beyond 1·9 Å resolution and have space group $P3_121$ with $a = b = 129{\cdot}8$, $c = 104{\cdot}3$ Å. The asymmetric unit contains half an MADH tetramer, *i.e.* one (HL) dimer. The crystallographic twofold axes running along *a*, *b* and *i* relate pairs of (HL) dimers to form the complete $(HL)_2$ tetramers. The solvent content of the crystals is quite high, *ca* 70%.

Data collection and processing

For data collection, single MADH crystals were aligned with their crystallographic threefold screw axis parallel to the oscillation axis. A total of five data sets (Table 1) were collected by oscillation photography (Arndt, Champness, Phizackerley & Wonacott, 1973). An initial native data set was recorded using radiation from an Elliott GX 21 rotating-anode X-ray source. The remaining four data sets were collected using radiation from the X-31 synchrotron beamline of the DESY-EMBL outstation in Hamburg (FRG). Table 1 gives particulars of data collection for the two native and the three heavy-atom derivative data sets used for the structure determination.

All films were digitized with a Scandig microdensitometer using a 50 μm raster size and were converted to a 100 μm raster size using the program *SCANPCK*. This leads to better processing statistics

Table 1. *Summary of data collection*

Data set	Crystal* soaking characteristics	No. of crystals used	Cooling	X-ray source	Oscillation range per film pack (°)	No. of film packs collected	Max. resolution (Å)
Native *A*	–	7	None	Rotating anode	1·2–1·5	110	2·32
Native *B*	–	7	273 K	Synchrotron	1	143	2·21
K_2PtI_6	72 h in saturated reagent	3	None	Synchrotron	1·4–1·5	72	2·36
$Pt(en)Cl_2$	48 h in 5 m*M* reagent	2	None	Synchrotron	1·5–1·75	50	2·76
$UO_2(C_2H_3O_2)_2$	24 h in 2 m*M* reagent	2	None	Synchrotron	1·25–1·5	72	2·50

* The mother liquor surrounding the crystals was 42% saturated ammonium sulfate in 0·1 *M* sodium acetate, pH 5·0.

Table 2. *Data-processing characteristics*

$R_{sym} = [(\Sigma^N|F_i - \bar{F}|)/N\bar{F}] \times 100$. $R_{ano} = [(\Sigma|\bar{F}^+ - \bar{F}^-|)/\Sigma 2\bar{F}] \times 100$, for heavy-atom data only.

Data set	No. of film packs used	R_{sym} (%)	$R_{sym\ F^+}$ (%)	$R_{sym\ F^-}$ (%)	R_{ano} (%)	No. of unique reflections	Useful resolution (Å)
Native *A*	72	10·02	–	–	–	32 211	2·8
Native *B*	139	8·46	–	–	–	49 688	2·25
K_2PtI_6	46	–	7·14	6·03	5·01	33 427	2·6
$Pt(en)Cl_2$	40	–	6·51	6·70	6·03	24 533	2·9
$UO_2(C_2H_3O_2)_2$	71	–	8·11	8·28	5·69	27 988	2·5

than when films are scanned directly with a 100 μm raster size (W. P. J. Gaykema, unpublished results). The native data set *A*, and data for the two platinum derivatives were processed using the film-evaluation program *OSCIL* (Schwager, Bartels & Jones, 1975), with film non-linearity response correction carried out according to Wierenga (1978). The remaining two data sets were processed using the profile-fitting film-evaluation program *OSC* (Rossmann, 1979). Fully recorded reflections present in more than one film pack were used for the calculation of scale and temperature factors according to Hamilton, Rollett & Sparks (1965). During scaling, the native data set *A* was added as the reference 'film pack' for the $Pt(en)Cl_2$ derivative. Isomorphous amplitudes for the heavy-atom derivatives were calculated from merged amplitudes for the separately processed Friedel pairs (Petsko, 1976). Statistics for the processing of all five data sets can be found in Table 2.

Preliminary isomorphous-replacement phasing

Each heavy-atom derivative data set was scaled to the native data set *A* using a local-scaling procedure (Matthews & Czerwinski, 1975), as implemented in the Groningen *BIOMOL* protein structure determination package. With these data, several difference Patterson maps were calculated, and were interpreted with a vector-search procedure. The main heavy-atom sites for the platinum iodide and for the uranyl acetate derivatives could be derived independently. These heavy-atom models were completed by difference Fourier methods, which were also used to locate heavy-atom positions in the $Pt(en)Cl_2$ derivative. Phase calculation and refinement was carried out jointly with isomorphous and anomalous differences using the program *PHARE* (provided by Dr G. Bricogne). At this stage the ambiguity concerning space-group assignment ($P3_121$ or its enantiomorph $P3_221$) was solved by means of cross-difference Fourier calculations (Matthews, 1966) as is shown in Table 3.

Native protein phases were calculated to a resolution of 4·5 Å to generate an initial electron density distribution. This resolution limit was chosen since the phasing power of the heavy-atom derivatives dropped to values smaller than one at any higher resolution (Table 4). Also, electron density maps calculated to resolutions higher than 4·5 Å were noisier and far less continuous in the protein regions than the 4·5 Å map. This 4·5 Å electron density distribution showed the large solvent regions with a clear boundary between solvent and protein. The separation of the protein into subunit regions was also obvious at this stage, and secondary-structure elements could already be located in the map.

Phase extension to 3·35 Å *by solvent flattening*

A program suite to perform solvent-envelope calculations and solvent flattening (Wang, 1982) was kindly provided by Professor B. C. Wang. The programs were adapted to run on a Convex C1-XP computer, and were modified to expand their possibilities. During solvent flattening, a new solvent envelope was calculated at every cycle using the reciprocal-space algorithm of Leslie (1987). We used an integration radius of 10·0 Å, with the fractional solvent content set at a conservative value of 0·6. Great care was taken to ensure that no protein regions were truncated by the solvent masks. This was particularly important in the early stages of this procedure, when an incomplete phasing set was used (see Fig. 1*a*).

Table 3. *Space-group assignment by cross-difference Fourier calculations (Matthews, 1966)*

Phases were calculated after refinement *versus* both isomorphous and anomalous differences. The peak heights listed are those of the highest peak in the map, which corresponds in each case to a heavy-atom site.

Amplitudes	Phases from	Peak heights $P3_121$	$P3_221$
$F_{K_2PtI_6} - F_{NAT}$	$UO_2(C_2H_3O_2)_2$	35·01	19·96
$F_{Pt(en)Cl_2} - F_{NAT}$	$UO_2(C_2H_3O_2)_2$	35·66	18·52
$F_{UO_2(C_2H_3O_2)_2} - F_{NAT}$	K_2PtI_6	25·82	14·68

Table 4. *MIRAS phase calculation to* 4·5 Å

The heavy-atom model at this stage of the structure determination comprised two sites for the uranyl acetate derivative, three sites for the platinum iodide derivative and five sites for the $Pt(en)Cl_2$ derivative. $R_{CULLIS} = (\Sigma||F_{PH}| - |F_P + F_{H\,calc}||/(\Sigma||F_{PH}| - |F_P|)$ for centric reflections only.

Resolution (Å)	18·95	12·99	9·88	7·97	6·68	5·75	5·05	4·50	Total
$\langle m \rangle$	0·79	0·82	0·84	0·80	0·81	0·79	0·74	0·66	0·75
$UO_2(C_2H_3O_2)_2$									
F_c/E_j	1·76	2·08	1·78	1·90	2·04	1·94	1·47	1·07	1·56
R_{CULLIS}	0·61	0·47	0·53	0·49	0·56	0·52	0·67	0·82	0·61
K_2PtI_6									
F_c/E_j	1·48	1·50	1·65	1·68	1·77	1·69	1·45	1·14	1·45
R_{CULLIS}	0·55	0·58	0·47	0·44	0·55	0·59	0·79	0·87	0·62
$Pt(en)Cl_2$									
F_c/E_j	1·27	1·56	1·92	1·86	2·25	2·23	1·80	1·44	1·76
R_{CULLIS}	0·72	0·66	0·53	0·56	0·51	0·47	0·57	0·81	0·60

Phase combination with the original phases, which had been obtained by multiple isomorphous replacement supplemented by anomalous-scattering information (MIRAS), was not carried out at any stage. In other words, the MIRAS phases were used only once for the calculation of the initial 4·5 Å map, and were simply discarded afterwards. In contrast to the original solvent-flattening procedure (Wang, 1982), no density truncation was carried out within the protein regions. Also, during application of the solvent masks, the density at every grid point within the solvent zone was simply replaced by the average solvent density. The 4·5 Å phases were first improved with three cycles of solvent flattening at constant resolution: during this process, the 4949 reflections with an initial MIRAS phase were used. The average accumulated phase shift from the heavy-atom phases was 35·3°, the average Sim weight (Sim, 1959, 1960) was 0·80, and the *R* factor between map inversion and observed structure-factor amplitudes was 19·0%.

Next, we started using the phases calculated from map inversion with the 1098 reflections lacking a heavy-atom phase (Fig. 1*a*). This was carried out at 4·5 Å resolution for three cycles, giving an average accumulated phase shift of 36·4°, an average Sim weight of 0·76, and an *R*-factor value of 16·5%. Subsequently, both amplitudes and phases calculated from map inversion were used for the 267 unrecorded reflections to 4·5 Å resolution. After another two cycles of solvent flattening at that resolution, the average accumulated phase shift for the initial 4949 reflections was 37·2°, and the *R* factor had dropped to 14·5%.

Following our experience with hemocyanin (Gaykema, Hol, Vereijken, Soeter, Bak & Beintema, 1984; Gaykema, Volbeda & Hol, 1986) we then carried out phase extension gradually, increasing the resolution in thin shells of about half a reciprocal-lattice point thickness at a time. After each phase-extension step, phase refinement was carried out at constant resolution for one or two solvent-flattening cycles. A total of 16 phase-extension steps were required from 4·5 to 3·35 Å resolution. During this procedure, calculated amplitudes and phases were always used in place of the few unrecorded reflections, so that electron-density maps were calculated with fully complete data.

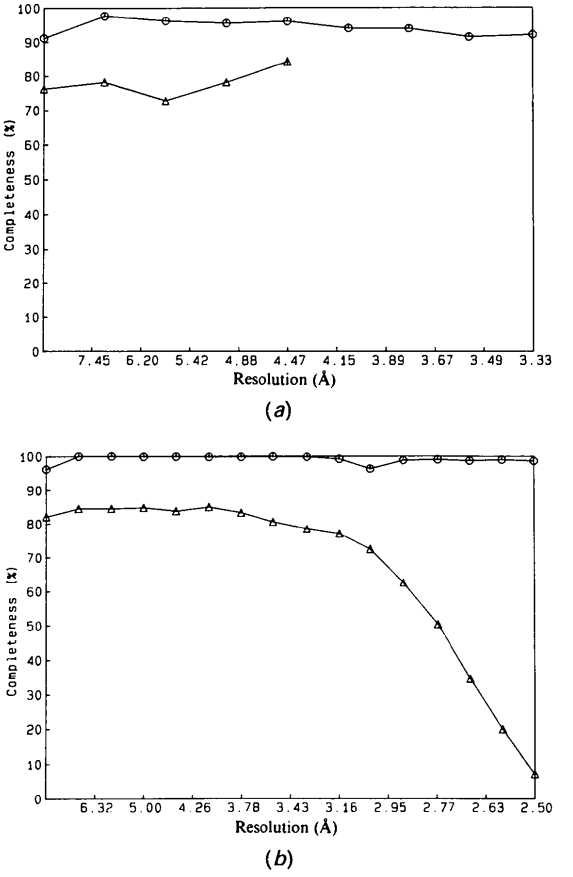


Fig. 1. Completeness of the data used for the isomorphous-replacement procedure. Circles represent completeness of the measured structure-factor amplitudes, and triangles indicate that of the isomorphous-replacement phases. The latter are less complete than the former because certain native reflections have been measured, for which no equivalent heavy-atom derivative measurements are available. Also, the very low intensity reflections have been discarded during phase refinement and phase calculation. (*a*) Preliminary phasing using the native data set *A*. (*b*) Phasing to 2·5 Å resolution using the combined native data set.

The overall effect of this phase-extension procedure can be summarized (Table 5) as follows: in the initial solvent-flattening cycle at 4·5 Å resolution, the mean phase change from multiple isomorphous-replacement phases was 27·8°, with a mean Sim weight of 0·69. In the last cycle, at 3·35 Å resolution, the average phase change from the previous cycle for all reflections had dropped to 1·7°, and the mean Sim weight was 0·86. A total of 10 102 new phases were introduced as the result of this procedure.

Using an electron density distribution calculated to a resolution of 3·5 Å with these solvent-flattening phases, it was possible to follow the polypeptide chain in the main body of the H subunit (339 residues) in the absence of any amino-acid sequence information.

Isomorphous-replacement phasing to 2·5 Å *resolution*

When the second, high-resolution native data set became available, a combined native data set was obtained by merging the two individual native data sets *A* and *B*. This was achieved with a local-scaling procedure in which deviating reflections from either data set were removed. A high *R* merge value (15·3% on *F*) was obtained owing to the differences in the dynamic range of the individual data sets, which were processed with two different program suites. Even so, better overall statistics were obtained when scaling the heavy-atom-derivative data sets on the combined native set than with either individual native data set (data not shown). Also, a new difference Patterson map calculated for the $Pt(en)Cl_2$ derivative with this combined native data could now be interpreted by vector search.

Using the latest solvent-flattening phases, each heavy-atom derivative was reexamined by means of difference Fourier maps calculated to a resolution of 3·5 Å. This led to the incorporation of one additional site for the uranyl acetate derivative. The multiple isomorphous-replacement phasing procedure was repeated to 2·5 Å resolution using all three heavy-atom derivatives, giving an overall mean figure of merit of 0·75. The final refined heavy-atom parameters are given in Table 6. 'Best' MIRAS phases were used with the combined native structure-factor amplitudes to calculate a figure-of-merit-weighted (Blow & Crick, 1959) 2·5 Å resolution electron density distribution.

Phase completion to 2·5 Å

The new 2·5 Å resolution electron density map was quite noisy, as a result of the lack of completeness of the MIRAS phases, in particular in the resolution range 3·5–2·5 Å (Fig. 1*b*), adding to the effects of phase errors. In order to improve and complete our phase set, and hence to obtain an improved 2·5 Å resolution electron-density distribution, we used a procedure similar to that described in *Phase extension to* 3·35 Å *by solvent flattening*. We also attempted to improve the resolution of the envelope

Table 5. *Statistics concerning phase extension from* 4·5 *to* 3·35 Å

$\langle m \rangle$ is the average Sim weight (Sim, 1959, 1960), except for the original unflattened map (cycle 0) where it refers to the heavy-atom figure of merit (Blow & Crick, 1959).

Cycle	Resolution limits (Å)	No. of reflections phased	$\langle m \rangle$	Accumulated phase shift* (°)	*R* factor
0	35·0–4·5	4967	0·75	0	-
3	–4·5	4967	0·80	35·3	0·190
6	–4·5	6083	0·76	36·4	0·165
8	–4·5	6314	0·76	37·2	0·145
11	–4·4	6746	0·78	37·4	0·147
13	–4·3	7230	0·78	37·6	0·156
15	–4·18	7835	0·78	37·8	0·161
17	–4·1	8303	0·79	37·8	0·154
20	–4·05	8611	0·80	38·1	0·132
22	–3·98	9053	0·80	38·3	0·141
24	–3·90	9628	0·78	38·5	0·150
26	–3·85	9997	0·79	38·6	0·144
29	–3·78	10554	0·80	38·8	0·142
31	–3·70	11229	0·77	38·9	0·158
33	–3·61	12079	0·84	39·0	0·148
35	–3·57	12478	0·85	39·1	0·129
38	–3·50	13243	0·85	39·3	0·122
40	–3·45	13815	0·86	39·4	0·123
42	–3·40	14427	0·86	39·5	0·122
45	–3·35	15051	0·86	39·6	0·116

* Only for data to 4·5 Å.

Table 6. *Results using isomorphous-replacement phasing to* 2·5 Å *resolution*

Statistics of the MIRAS phase calculation

Resolution (Å)	13·33	8·24	5·96	4·76	3·84	3·26	2·86	2·50	Total
$\langle m \rangle$	0·81	0·78	0·77	0·70	0·62	0·60	0·59	0·50	0·75
K_2PtI_6									
F_c/E_i	2·01	2·17	2·35	1·68	1·19	0·78	0·63	1·00	1·23
R_{CULLIS}	0·58	0·47	0·51	0·64	0·79	0·92	0·96	0·74	0·71
$Pt(en)Cl_2$									
F_c/E_i	2·51	2·36	2·14	1·38	1·02	0·83	0·75	-	1·22
R_{CULLIS}	0·45	0·53	0·54	0·61	0·88	0·93	1·29	-	0·69
$UO_2(C_2H_3O_2)_2$									
F_c/E_i	2·55	2·21	2·39	1·83	1·38	1·54	1·73	1·42	1·68
R_{CULLIS}	0·30	0·48	0·52	0·58	0·80	0·73	0·69	0·54	0·63

Heavy-atom parameters after refinement

Site No.	O_I	O_A	*x*	*y*	*z*
$UO_2(C_2H_3O_2)_2$					
1	0·16145	0·23485	0·44393	0·81178	0·08116
2	0·18985	0·33925	0·38234	0·92604	0·10592
3	0·05796	0·16295	0·37653	0·93486	0·06828
K_2PtI_6					
1	0·30774	0·34217	0·26671	0·45605	0·02380
2	0·20025	0·26591	0·48561	0·34642	0·07150
3	0·00468	0·08473	0·52307	0·12122	0·05180
$Pt(en)Cl_2$					
1	0·22867	0·29467	0·56107	0·11947	0·03874
2	0·25863	0·33240	0·48029	0·34299	0·07106
3	0·14155	0·23440	0·61806	0·52094	0·01637
4	0·21595	0·25569	0·23553	0·46780	0·01995
5	0·08868	0·12862	0·26697	0·45629	0·02310

Notes: R_{CULLIS} is as defined in Table 5; O_I, isomorphous occupancy; O_A, anomalous occupancy; *x*, *y* and *z* are expressed in fractional coordinates; anisotropic temperature factors were also refined for each heavy-atom site.

by gradually increasing the estimated solvent content from an initial value of 0·58 to a value of 0·63, while the integration radius used to calculate the solvent masks was decreased from a starting value of 10·0 Å to a final value of 8·0 Å in steps of 0·5 Å. This might result in a more convoluted, and hopefully more accurate envelope. As previously, phase combination was not carried out, so that the MIRAS phases were used only once for the calculation of the initial 2·5 Å map.

The initial 21 553 phases to 2·5 Å resolution (61% of all possible reflections to that resolution) were first improved by ten cycles of solvent flattening. Next, map-inversion phases were combined with the measured amplitudes for those reflections lacking an initial MIRAS phase (Fig. 1*b*), while both calculated amplitudes and phases were used in place of the missing reflections. This was carried out for the 2191 reflections to 3·5 Å resolution in the next two cycles. Phase extension to generate the 11 730 missing phases between 3·5 and 2·5 Å (53% of all possible reflections in that resolution range, Fig. 1*b*) was then gradually carried out in 28 solvent-flattening cycles with 15 extension steps. Finally, another 20 solvent-flattening cycles were performed with all data to 2·5 Å resolution.

During this procedure, the mean Sim weight (Sim, 1959, 1960) rose from a value of 0·63 for the initial solvent-flattening cycle to 0·84 for the final cycle (Table 7), and the *R* factor dropped from an initial value of 38·2% in the first cycle to a value of 11·4% in the final cycle. At the end of the procedure, the mean phase change from the initial MIRAS phases was 62·5°. The 2·5 Å map, calculated with all data to that resolution (35 474 reflections), was a great improvement on any previous map: the chain tracing of the H subunit was completed, and long stretches could be followed in the L subunit.

Table 7. *Phase completion from* 3·5 *to* 2·5 Å

Symbols are as explained in Table 5.

Cycle	Resolution* limit	Total No. of reflections	⟨*m*⟩	Accumulated‡ phase shift (°)	*R* factor
0	[2·5]†	21553	0·75	0	–
10	–	21553	0·81	58·9	0·176
12	3·5	23744	0·78	59·1	0·183
14	3·4	23994	0·79	59·3	0·168
16	3·3	24266	0·80	59·7	0·161
18	3·2	24617	0·80	60·1	0·154
20	3·1	25012	0·80	60·4	0·149
22	3·0	25600	0·80	60·7	0·146
24	2·94	26080	0·80	60·9	0·142
26	2·85	26962	0·79	61·1	0·144
28	2·77	28053	0·79	61·3	0·144
30	2·71	29169	0·79	61·4	0·143
32	2·67	30029	0·79	61·6	0·140
34	2·64	30761	0·80	61·6	0·137
36	2·60	31892	0·79	61·8	0·141
38	2·57	32890	0·79	61·9	0·141
40	2·53	34305	0·79	62·0	0·144
60	2·50	35474	0·84	62·5	0·114

* The resolution limit shown corresponds to the limit to which calculated phases (and amplitudes) were used for those reflections which did not receive an isomorphous-replacement phase.

† This number refers to the resolution limit for the MIRAS phase calculation. The corresponding reflections were used throughout the phase-completion procedure.

‡ Phase shifts were computed for all reflections with an initial MIRAS phase (to 2·5 Å resolution).

Partial model phasing

Although the electron density was of good quality, a full chain tracing in the L subunit was still not feasible; in particular, several forks were seen in the density. These were thought to correspond to the disulfide bridges which interconnect that subunit (Ishii *et al.*, 1983). In order to solve the remaining ambiguities, the method of partial-structure phasing (Rice, 1981) was applied.

Three successive models for the polypeptide chain of MADH were refined using molecular-dynamics procedures (Brünger, Kuriyan & Karplus, 1987). Each of these models was subjected to a refinement procedure (Table 8), which consisted of:

(i) Several steps of steepest-descent energy minimization, including low-resolution X-ray terms (EMX). This was carried out to reduce the highest energies of the starting models, as has been described by Fujinaga, Gros & van Gunsteren (1989).

(ii) Several hundred 0·002 ps steps of low-temperature molecular-dynamics refinement (MDX) using the *GROMOS* package (Fujinaga *et al.*, 1989; van Gunsteren & Berendsen, 1987). Following the experience of Gros, Fujinaga, Dijkstra, Kalk & Hol (1989) and Gros, Betzel, Dauter, Wilson & Hol (1989), this was performed using low-resolution X-ray terms initially, and the resolution limit was increased to a final value of 2·25 Å during this procedure.

(iii) Several steps of steepest-descent energy minimization, using 2·25 Å resolution X-ray terms. These steps were carried out to ensure that the refined polypeptide models were at the local energy minimum closest to the point reached after MDX.

(iv) A few cycles of individual temperature-factor refinement using the *FFT* program of Agarwal (1978).

The resulting partial-model phases were then combined (Hendrickson & Lattmann, 1970; Rossmann & Blow, 1961) with the 2·5 Å solvent-flattening phases. In order to overcome the domination of the resulting combined phases by the solvent-flattening phases (Table 9), the Hendrickson–Lattmann coefficients (Hendrickson & Lattmann, 1970) for the latter were divided by two before phase combination. These combined phases were used to generate improved 2·25 Å electron density maps, in which the polypeptide models were 'manually' rebuilt on a computer-graphics system using the program *FRODO* (Jones,

Table 8. *Overview of the partial structure refinement procedure*

Round	Cycle	Model refined	No. of atoms	Refinement stage	Resolution limits (Å)	No. of steps*	Temp. (K)	*R* factor	$\langle F_o - F_c \rangle$	Weight†	R.m.s. coordinate shift from previous cycle (Å)	R.m.s. coordinate shift from starting model (Å)
1	1	*A*	2703	EMX	10·0–3·0	100	–	0·397	557	300	–	0·38
	2	(H subunit		MDX	10·0–3·0	500	300	0·359	502	250	0·81	0·95
	3	only)		MDX	10·0–2·25	200	300	0·388	397	250	0·37	0·95
	4			EMX	10·0–2·25	50	–	0·376	367	200	0·10	0·95
	5			BF	8·0–2·25	6	–	0·360	349	–	–	–
2	1	*B*	3211	EMX	10·0–2·8	100	–	0·331	424	250	–	0·33
	2	(H subunit		MDX	10·0–2·8	400	300	0·301	384	200	0·59	0·74
	3	+ L subunit		MDX	10·0–2·25	200	300	0·344	325	190	0·32	0·75
	4	in 4 segments)		EMX	10·0–2·25	50	–	0·328	308	150	0·10	0·75
	5			BF	8·0–2·25	4	–	0·304	287	–	–	–
3	1	*C*	3385	EMX	10·0–3·5	100	–	0·284	442	450	–	0·26
	2	(HL dimer)		MDX	10·0–3·5	200	320	0·282	443	440	0·60	0·69
	3			MDX	10·0–3·0	200	320	0·301	417	440	0·50	0·62
	4			MDX	10·0–2·75	200	320	0·307	383	400	0·38	0·63
	5			EMX	8·0–2·7	50	–	0·297	355	375	0·10	0·62
	6			BF	8·0–2·7	3	–	0·269	324	–	–	–
	7			MDX	10·0–2·5	200	320	0·293	318	350	0·21	0·61
	8			EMX	10·0–2·5	50	–	0·289	313	325	0·09	0·60
	9			MDX	10·0–2·25	200	320	0·322	299	310	0·18	0·61
	10			EMX	8·0–2·25	50	–	0·311	284	295	0·09	0·60
	11			BF	6·0–2·25	4	–	0·286	260	–	–	–

Notes: EMX, energy minimization with X-ray constraints using *GROMOS*; MDX, low-temperature molecular-dynamics refinement using *GROMOS*: BF, individual isotropic temperature-factor refinement. During the refinement procedure, the weights given to the X-ray terms were adjusted according to the model being refined: with the two incomplete models (models *A* and *B*), these weights were set equal to half the average r.m.s. discrepancy between observed and calculated structure-factor amplitudes. With a more complete molecular model (model *C*), these weights were relaxed, and the resolution was increased very gradually, thus allowing for a more efficient search to be carried out (Gros, Fujinaga *et al.*, 1989). Models *A* and *B* were refined *versus* the combined native data set, whereas for model *C* we used the native data set *B*.

* One step for a molecular-dynamics run represents 2 fs.

† For the calculation of the X-ray energy term, the weight (w) is applied in the form $1/w^2 \sum/hkl(|F_o| - k|F_c|)^2$.

Table 9. *Phase-combination statistics: mean phase difference (°) from combined phase*

The first two colums correspond to figures obtained with simple phase combination; the next two columns correspond to the data obtained after dividing the Hendrickson–Lattmann coefficients (Hendrickson & Lattmann, 1970) for the solvent-flattening phases by two before phase combination.

Resolution (Å)	$\langle \Delta\varphi \rangle$ from solvent-flattening phases	$\langle \Delta\varphi \rangle$ from model *A* phases	$\langle \Delta\varphi \rangle$ from solvent-flattening phases	$\langle \Delta\varphi \rangle$ from model *A* phases
100·0–9·67	8·70	53·86	16·34	49·94
9·67–6·84	24·30	43·03	32·38	34·69
6·84–5·59	26·50	43·78	35·86	34·75
5·59–4·84	23·75	42·89	34·07	33·34
4·84–4·33	22·21	44·32	32·38	35·05
4·33–3·95	24·97	47·12	37·26	35·91
3·95–3·66	29·70	49·83	42·60	37·48
3·66–3·42	30·08	53·12	43·80	39·89
3·42–3·23	28·45	59·91	45·35	44·93
3·23–3·06	26·17	65·42	43·02	51·41
3·06–2·92	21·24	71·67	36·66	60·05
2·92–2·80	16·08	75·23	29·28	66·08
2·80–2·69	11·79	79·89	23·62	72·49
2·69–2·59	14·21	80·20	27·63	71·39
2·59–2·50	34·11	68·89	53·78	51·40

1985). This step involved the establishment and updating of an 'X-ray sequence'.

The three polypeptide models which were refined were obtained as follows:

Model *A*: Approximate Cα coordinates for the H subunit only were measured from electron density distributions plotted on mini-maps and were used to generate a polyalanine model with the *DGLOOP* option of the crystallographic display program *WHATIF* (Vriend, 1990). The latest 2·5 Å resolution solvent-flattened electron density map was displayed on an Evans and Sutherland PS 390 graphics system using the *FRODO* program (Jones, 1985). The polypeptide model was manually rebuilt in this map to give a good fit in the electron density. This step

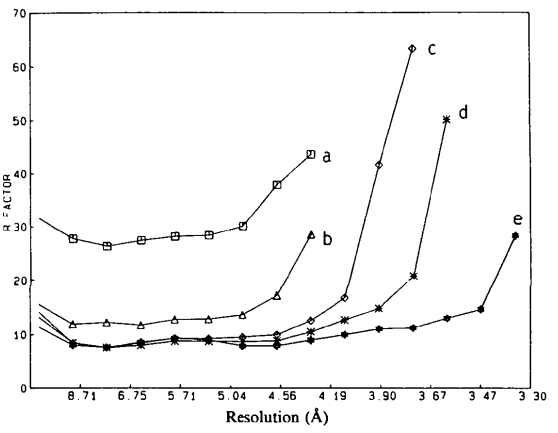


Fig. 2. Results of phase extension to 3·35 Å by solvent flattening: agreement between observed and map-inversion amplitudes as judged by the *R* factor ($\sum|F_{obs} - F_{inv}|/\sum F_{obs} \times 100$). (*a*) After the initial solvent flattening cycle, at 4·5 Å resolution. (*b*) After the final cycle at 4·5 Å. (*c*) After phase extension to 3·9 Å. (*d*) After phase extension to 3·61 Å. (*e*) At the end of the phase-extension procedure.

involved the determination of an X-ray sequence for the H subunit of MADH. This was achieved using the *SAM* option of *FRODO* by mutating the individual residues of the polyalanine model to those residues which were judged to fit best the electron density. The model at that stage comprised 385 residues in one segment.

Model *B*: In the 2·25 Å resolution map calculated with combined model *A*/solvent-flattening phases, the small subunit was partially traced in four segments. At this stage we established a partial preliminary X-ray sequence for that subunit, and the X-ray sequence for the H subunit was updated to improve the fit of the postulated side chains into the electron density. This model comprised 385 residues for the H subunit, and a total of 103 residues in the four segments corresponding to the L subunit.

Model *C*: The chain tracing of the L subunit was completed in the 2·25 Å resolution map calculated with combined model *B*/solvent-flattening phases. This stage involved an important model rebuilding session for the L subunit, where connections between the four segments of the previous model were established. The direction of the polypeptide chain in one of these segments was also reversed at this stage. This model, which was used for the final cycle of molecular-dynamics refinement, comprised 370 residues for the H subunit and 122 residues for the L subunit, and gave an *R*-factor value of 28·6%.

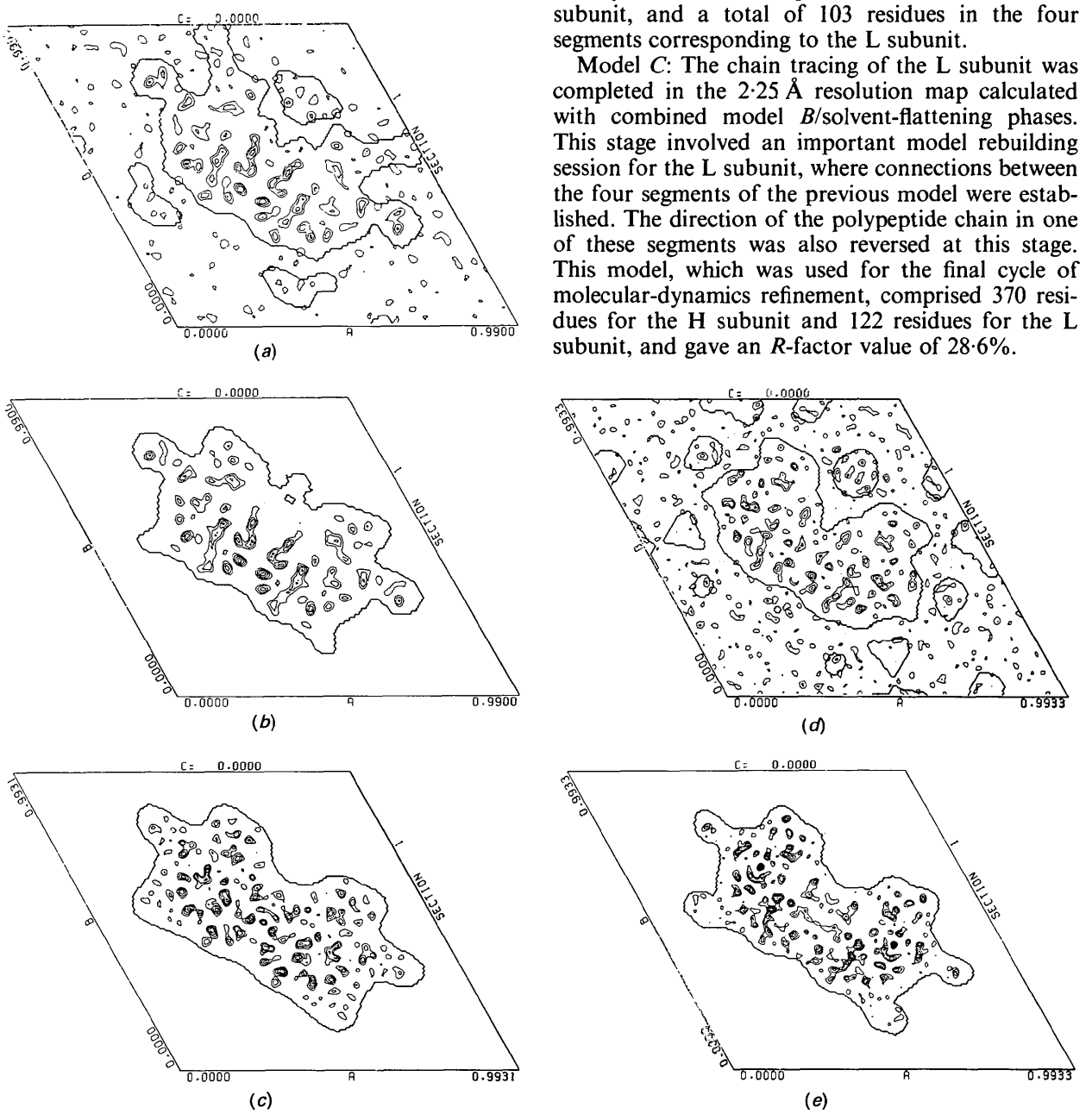


Fig. 3. Results of phase extension by solvent flattening illustrated by the section $z = 0$ of the figure-of-merit-weighted native maps with the corresponding section of the solvent mask at different stages of the procedure. (*a*) Initial 4·5 Å resolution MIRAS map. (*b*) After phasing of all measured native reflections at cycle 4, 4·5 Å resolution. (*c*) After phase extension to 3·35 Å. The map shown is that used for model building (3·5 Å data). (*d*) MIRAS map calculated at 2·5 Å resolution using the combined native data set. The phasing set is *ca* 80% complete to 3·5 Å, but only 7·1% complete between 2·6 and 2·5 Å. (*e*) After extension to 2·5 Å. The electron densities are contoured every σ level, starting at 1σ.

The 2·25 Å resolution electron density distribution, calculated with combined phases obtained with the refined model *C* was used for a final session of model rebuilding for the L subunit. During this session, four residues were mutated to the residue type deemed to fit best the electron density, and one deletion of a residue was made. The model which was used to study the polypeptide fold and the

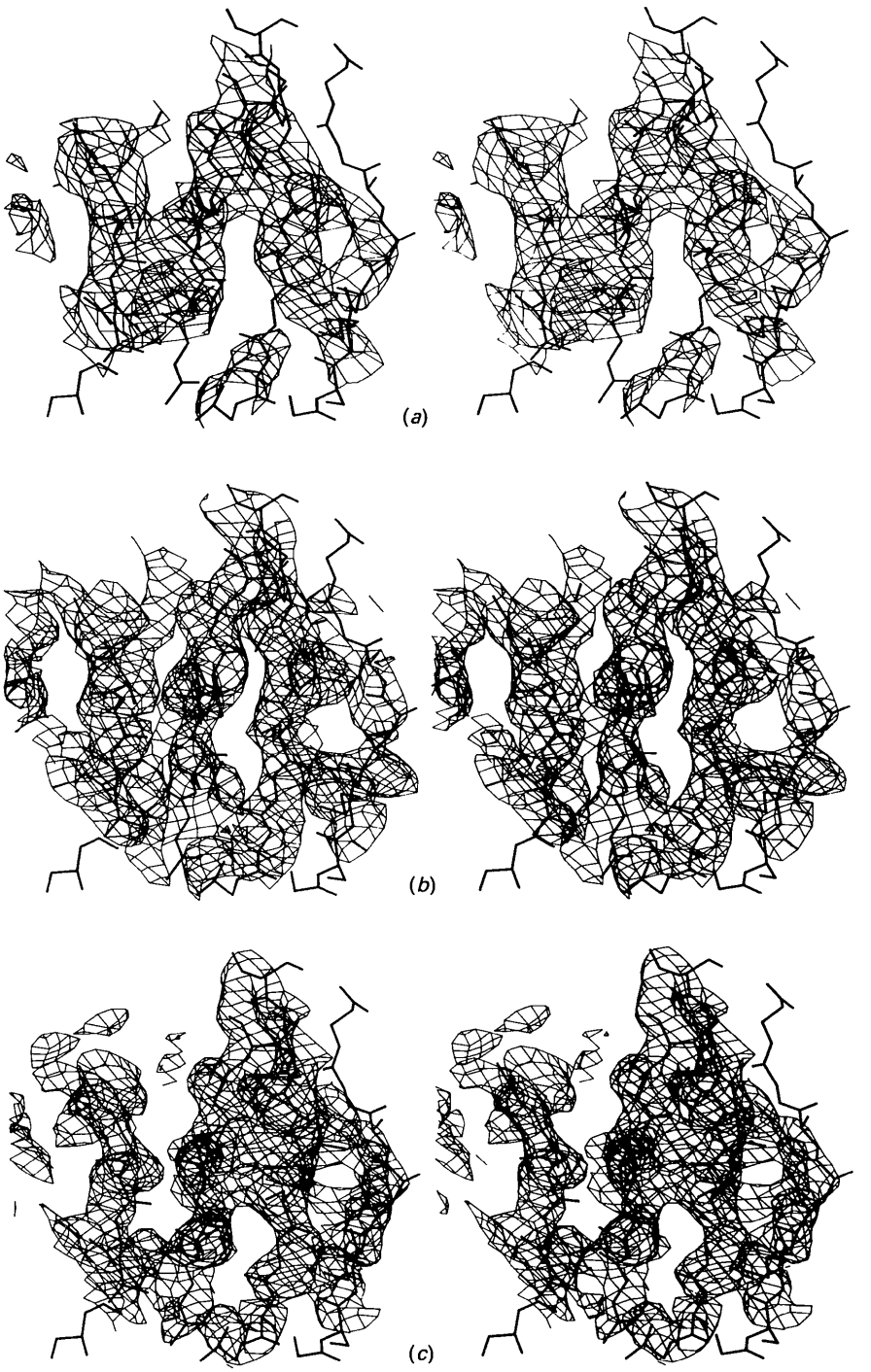


Fig. 4. Electron density corresponding to a four-stranded antiparallel β-sheet (thin lines) with the current model shown in thick lines. (*a*) Initial MIRAS map, 4·5 Å resolution. (*b*) 3·5 Å resolution density obtained after phase extension to 3·35 Å. (*c*) Electron density calculated with MIRAS phases to 2·5 Å resolution.

active-site region of the enzyme therefore consisted of 370 residues in the H subunit, and 121 residues in the L subunit.

Results and discussion

Phase improvement and extension by solvent flattening

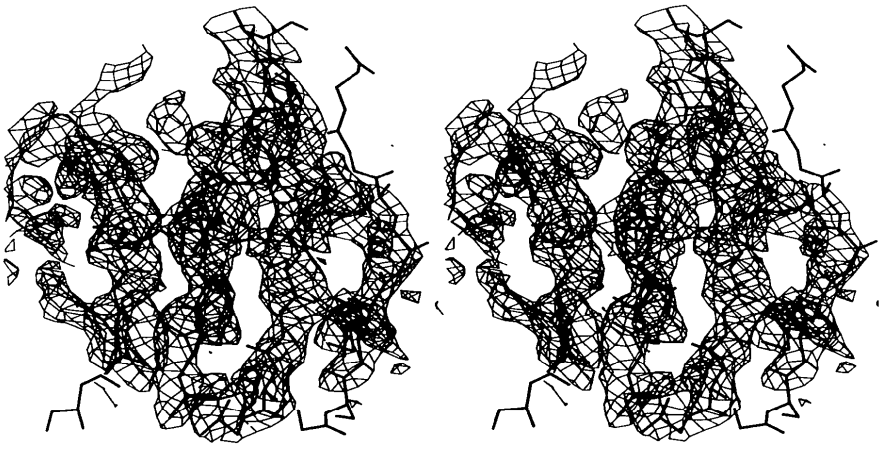
Our attempts to use solvent flattening alone to improve phases and extend the resolution gave results similar to those obtained using both molecular averaging and solvent flattening (see *e.g.* Gaykema *et al.*, 1984, 1986; Luo *et al.*, 1987; Filman, Syed, Chow, Macadam, Minor & Hogle, 1989). Convergence was attained rapidly since no phase combination was performed, and plots of map-inversion *R* factors *versus* resolution (Fig. 2) showed a typical pattern: a good agreement is seen between observed and calculated amplitudes, except at the current resolution limit.

As had been observed previously by other workers (Rayment, 1983; Holden, Rypniewski, Law & Rayment, 1987), attempts to truncate or attenuate negative electron densities within the protein regions led to discontinuities in the resulting electron density maps. Also, our attempts to combine phases obtained from map inversion with the original multiple isomorphous-replacement phases led to similar results. This can be attributed to two main reasons: first, by combining phase information from two sources which are not independent (solvent-flattening phases *versus* isomorphous-replacement phase information), we are probably overestimating the quality of the resulting phases. Secondly, phase combination between quite reliable solvent-flattening phases, as obtained from a successful procedure, and less reliable heavy-atom phases will shift the resulting combined phases back towards the latter, resulting in electron density maps of lower quality. That the procedure was successful can be seen from the resulting electron density maps:

(i) Superimposed sections of the solvent mask and of the corresponding electron density at various stages of the procedure (Fig. 3) show a gradual improvement of mask definition. It must be noted that the isolated islands of density within the solvent which are wrongly attributed to be part of the protein only disappear after phasing of all observed reflections, *i.e.* when nearly complete reflection sets are used. The level of detail within the protein zone is seen to increase with increasing resolution.

(ii) During the phase-extension procedure, the density corresponding to secondary-structure elements (Fig. 4) is seen to improve considerably. In the example shown, spurious electron density linking individual β-strands disappears, and gaps along the main chain are being filled. This results in maps where the tracing of the main chain is quite straightforward, even in the absence of sequence information.

(iii) A beneficial effect is also obtained at the level of individual side chains, as is shown on Fig. 5. There, a dramatic improvement in the density for a side chain occurs: with MIRAS phases at 4·5 Å resolution there is no density for the side-chain atoms. Already at 3·5 Å the density allows recognition of an isoleucine residue (F. Huitema, J. A. Duine, H. J. Bak & J. J. Beintema, unpublished results). Therefore, in the absence of sequence information, the resulting electron density map could be used to establish a preliminary X-ray sequence, which was used in the next stage of the structure determination.



(*d*)

Fig. 4 (*cont.*) (*d*) 2·5 Å map obtained after completion of the solvent-flattening procedure. Electron densities are contoured at the 1σ level.

Refinement and partial model phasing

As was noted by other workers before us (see *e.g.* Brünger, 1988; Gros, Fujinaga *et al.*, 1989; Gros, Betzel *et al.*, 1989) molecular-dynamics refinement procedures allow large shifts of atomic coordinates to occur without any manual intervention. In the course of the refinment of the three successive MADH models, we observed atomic shifts of up to 4·9 Å, while entire residues were shifted by up to 3·3 Å as judged from the largest coordinate differences seen for Cα atoms.

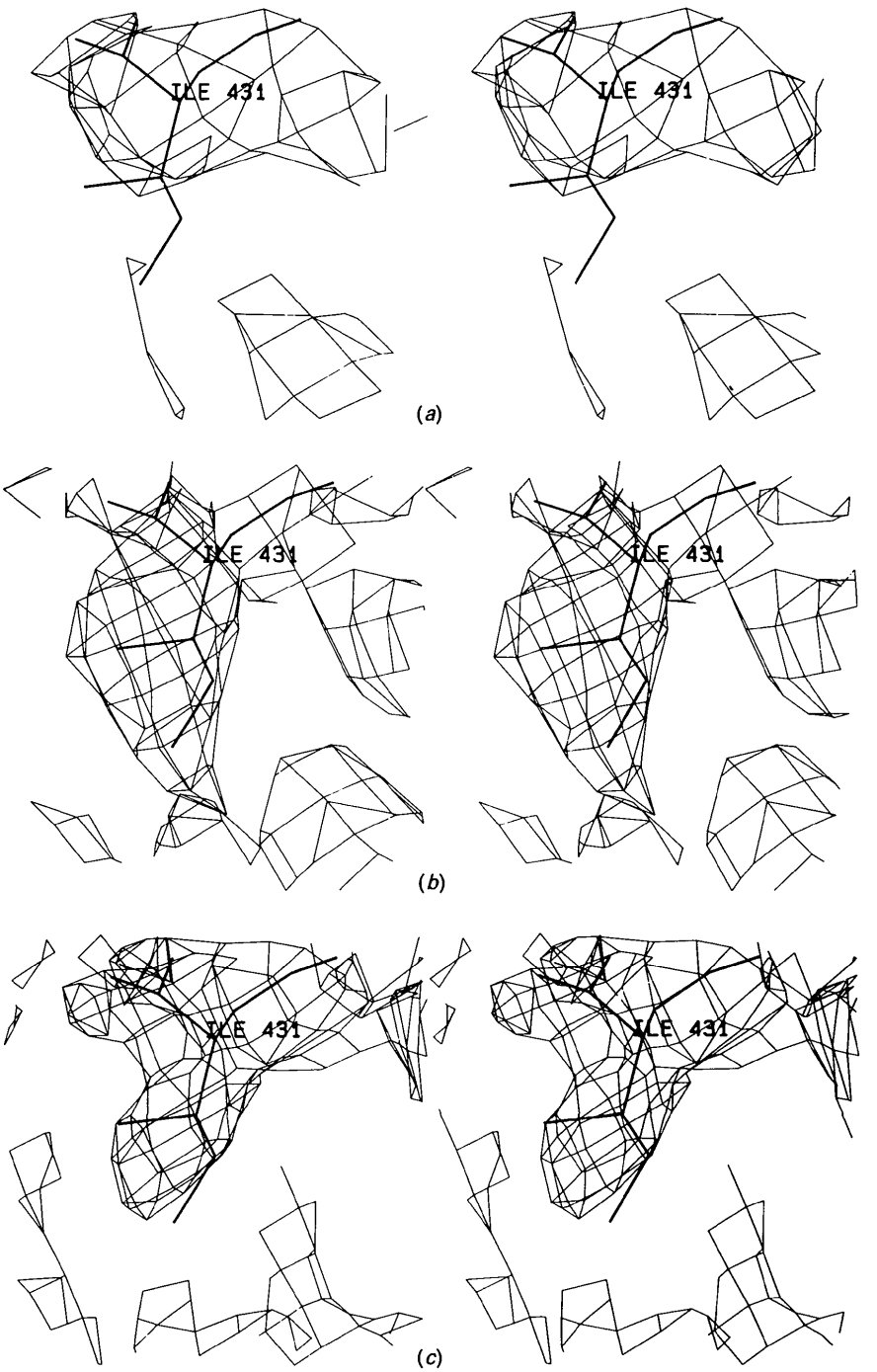


Fig. 5. Electron density corresponding to an isoleucine side chain at different stages of the solvent-flattening procedure: (*a*), (*b*) and (*c*) are as indicated for Fig. 4.

It must be noted that different situations will, in principle, arise if the model to be refined is obtained from a molecular-replacement solution to the phase problem, or if it is obtained from an isomorphous-replacement procedure: in the former case quite large shifts of residues and even domains are expected, for example as the result of insertions and deletions in the structure (Gros, Fujinaga *et al.*, 1989; Gros, Betzel *et al.*, 1989). On the other hand, with good heavy-atom or solvent-flattening phases the polypeptide model should be essentially correct, in particular for main-chain atoms, so that somewhat smaller shifts are expected. With MADH the situation was complicated by the lack of an amino-acid sequence, so that we depended essentially on the improvement of the electron density following refinement for the update of our 'X-ray sequence'.

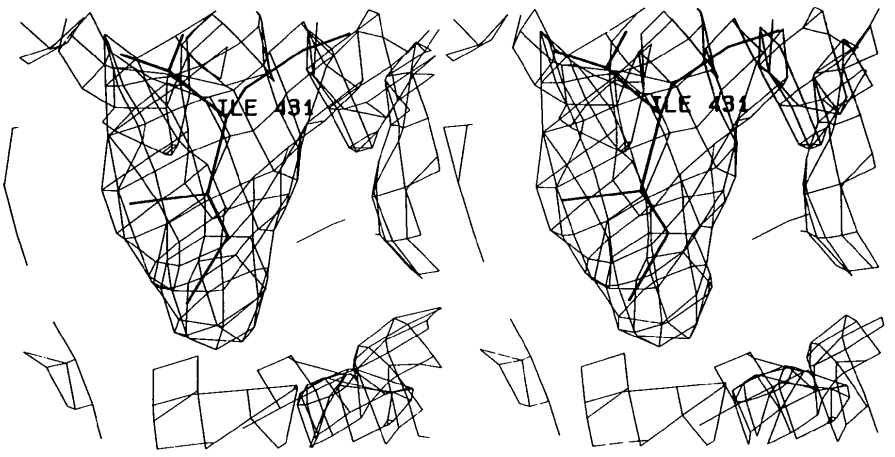
As an example, Fig. 6 shows a region of the map where side-chain assignments could not be made correctly: in the starting model *B*, the central residue had been left as alanine. During refinement (Fig. 6*a*), the atoms of this model are pulled towards the unexplained side-chain density. In the resulting 2·25 Å resolution map, calculated with combined phases (Fig. 6*b*), the electron density has sufficiently improved to allow the recognition of a tyrosine side chain, which is connected to the Cα atom of the central residue of Fig. 6.

Since all three models still contained errors after completion of the molecular-dynamics positional refinement, this was followed by individual isotropic temperature-factor refinement (Table 8). This step was introduced in the refinement procedure as a means of reducing the effects of the remaining model errors on the resulting electron density, *i.e.* in order to reduce model bias in the maps. For this reason, the plot of the temperature factors *versus* the sequence number for the refined model *C* is not very smooth, with most of the spikes in the plot caused by these errors. Apart from this use of temperature factors as an 'error sink', the *B* factors so derived still have some physical meaning, since the average temperature factor for the solvent-accessible residues in the MADH tetramer is 28·2 Å^2, whereas that for the solvent-inaccessible residues is smaller, 20·5 Å^2.

Motion of peptide planes during refinement

An argument which was proposed in favour of the use of molecular-dynamics refinement techniques *versus* conventional least-squares refinement is that large errors in the starting model can be corrected during the refinement procedure without manual intervention by the investigator if refinement is carried out using the molecular-dynamics approach (Brünger, Karplus & Petsko, 1989). A good criterion to judge this is to study the molecular corrections involving high-energy-barrier transitions that have occurred during molecular-dynamics refinement, since these are impossible to obtain with classical refinement.

An example of such a molecular correction is the flipping of peptide planes, which has been analyzed at the various stages of the refinement procedure (Table 10). As is expected, the model improves during the course of the refinement, and consequently the number of peptide units which have to be corrected decreases and the average peptide angular shift diminishes. However, as can be seen in Table 10, there are still numerous peptide planes which have flipped in the last round of refinement of the complete model *C*. When no sequence information is available, the errors made in the establishment of the X-ray sequence, such as the definition of a small



(*d*)

Fig. 5 (*cont.*) (*d*) 2·5 Å map obtained after completion of the solvent-flattening procedure. No model phases were used in these maps.

residue type where a large side chain is required (or inversely), will create additional motions in the model being refined, since the algorithm will try to pull atoms into unused electron density (see also Fig. 6). Thus, a number of these peptide flips will be caused by such errors, in addition to the motions expected from erroneous starting models which were fitted to regions of weak electron density (see *e.g.* Brünger *et al.*, 1989).

The refined model

The quality of a macromolecular model after refinement is usually estimated, among other things, from the correctness of its stereochemistry as represented by a Ramachandran plot (Ramakrishnan & Ramachandran, 1965). As is expected for a structure containing a total of 37 β-strands and little α-helical structure (Vellieux *et al.*, 1989) this φ–ϕ plot for MADH shows most residues in the region corresponding to β-conformation. 78 out of a total of 491 residues are found in the unallowed regions of conformational space. Although the distinction between glycine and non-glycine residues is not fully valid when the amino-acid sequence is unknown, it should be noted that 25 out of these 78 unallowed Ramachandran entries correspond to glycine residues in our X-ray sequence.

Obviously quite a large number of deviating residues will be found in a partially refined structure, probably still containing a few gross errors and lacking any solvent molecules, in which an X-ray

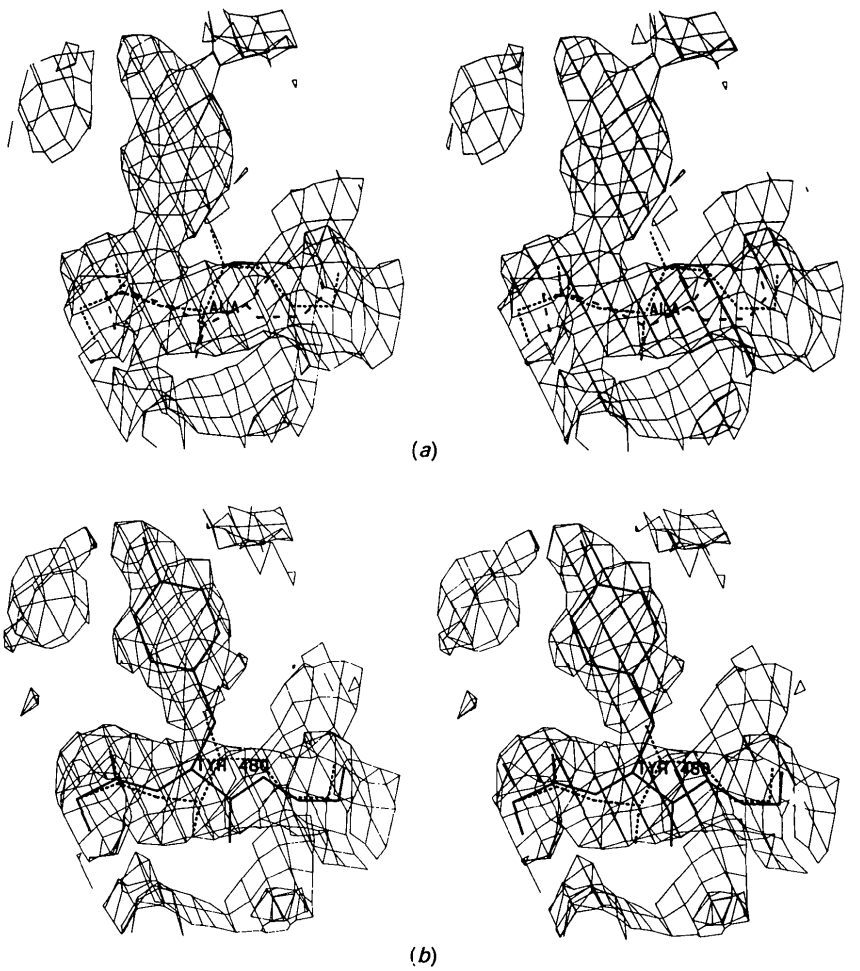


Fig. 6. Improvement of the electron density during refinement of model *C* allowing recognition of a tyrosine side chain. (*a*) Initial model (broken lines) and refined model (dotted lines) together with the 2·25 Å resolution map calculated from combined solvent-flattening/refined-model *B* phases. Even though density for a large aromatic side chain is present, it cannot be connected in a stereochemically reasonable manner to the 'left' residue of the figure. To avoid model bias these residues were introduced in the atomic model of MADH as alanines. (*b*) Refined model (dotted line) and new model after rebuilding (thick lines) shown with the 2·25 Å resolution map obtained from combined solvent-flattening/refined-model *C* phases.

Table 10. *Analysis of peptide plane conformation of the energy-minimized models*

Values are given for before and after molecular-dynamics refinement carried out as described in Table 8. A peptide plane is considered to have flipped during the refinement if the angular shift for the plane defined by the C_n, O_n and N_{n+1} atoms is larger than 90°.

Model	No. of peptide planes	Average angular shift (°)	No. of peptides having flipped	Largest angular shift observed (°)
A	383	39·0	23	162·4
B	484	30·9	13	158·5
C	490	23·5	16	131·3

Table 11. *Root-mean-square deviations from ideal geometry after molecular-dynamics refinement*

Calculations were carried out using the program *TNT* (Tronrud, Monzingo & Matthews, 1987).

Refined model	Bond lengths (Å)	Bond angles (°)	Torsion angles (°)	Trigonal non-planarity (Å)	Non-planarity (Å)	Bad contacts (Å)
A	0·031	6·369	34·317	0·044	0·126	0·375
B	0·038	7·926	35·104	0·059	0·140	0·293
C	0·021	3·642	31·047	0·027	0·076	0·212

sequence had to be used (see *e.g.* Anderson, Stenkamp & Steitz, 1978). It is therefore expected that the stereochemistry of the model will improve once the refinement has been completed using the correct amino-acid sequence, which is currently being determined (J. A. Duine & J. Frank, personal communication).

A plot of the σ_A parameter (Read, 1986) for the refined model *C* yielded an estimated coordinate error of 0·45 Å, which is a further indication that the refinement process is not yet complete. The r.m.s. deviations from ideal values that describe the molecular geometry for the refined MADH models (Table 11) show the typical decrease towards these ideal values as the model under refinement improves. Even though the current refined model is not at the end of the refinement process, these r.m.s. values compare well with the values obtained for similarly sized proteins refined to similar resolutions (see *e.g.* Henrick, Collyer & Blow, 1989).

In the absence of the correct amino-acid sequence, a detailed description of subunit contacts, geometries of side chains *etc.* cannot be given. In particular, the exact description of the active-site centre with its neighbouring residues involved in the reaction catalyzed by the enzyme will have to await the completion of the sequencing studies. However, in addition to the overall structure of MADH and its topology which have already been described elsewhere (Vellieux *et al.*, 1989), two points deserve further attention in this paper and are described in the following paragraphs.

Disulfide bridges of the L subunit of MADH

The final solvent-flattened 2·5 Å electron density map, as well as the subsequent phase-combined electron density maps all showed, in the L subunit region, six peaks of high density which interconnected the polypeptide chain and could be ascribed to disulfide bridges. This observation was in agreement with the results of sequence studies of the L subunit of the MADH from *Pseudomonas AM*1 (Ishii *et al.*, 1983), which indicated the presence in that subunit of 12 cysteine residues, proposed to be involved in disulfide bridges. The location of the so-deduced cysteinyl residues in our X-ray sequence (Vellieux *et al.*, 1989) was in excellent agreement with that found in the *Pseudomonas* enzyme. This high disulfide content in a subunit of M_r 12900 probably helps to stabilize the fold of the polypeptide chain, thus explaining the heat stability observed for that subunit (Shirai, Matsumoto & Tobari, 1978). It seems therefore justified to pay particular attention to the geometry of all six disulfide bridges located in the L subunit.

The results of this analysis, shown in Table 12, indicate that four out of the six disulfides found in the L subunit adopt classical conformations (Richardson, 1981): disulfides 1, 5 and 6 form left-handed spirals, whereas disulfide 4 is found in a *trans–gauche–trans* conformation. The remaining two disulfide bridges are found in unusual conformations: that between residues 29 and 61 of our X-ray sequence (Vellieux *et al.*, 1989) is found in a left-handed hook conformation. It is noteworthy that such a conformation has been proposed for a potential disulfide site between residues 154 and 166 of subtilisin Carlsberg (Hazes & Dijkstra, 1988). The last remaining disulfide, involving residues 36 and 119, clearly deviates: the total energy of the disulfide bond is 83·32 kJ mol^{-1}, which is due to the very unusual value of the χ_3 angle (*ca* 175°). The most likely possibility to account for this abnormal disulfide bond is the presence of an error (such as the deletion of one residue) in the X-ray sequence in the neighbourhood of the disulfide, thus constraining that bond to an unlikely conformation.

X-ray sequence and topology of the H subunit of MADH

Although the main elements of the secondary structure of the H subunit of MADH, together with its topological and possible evolutionary relationship to the neuraminidase subunit (Varghese, Laver & Colman, 1983) have already been described elsewhere (Vellieux *et al.*, 1989), it seems appropriate to describe here the X-ray sequence of that subunit as well as its relationship to its topology.

Table 12. *Conformation of the six disulfide bridges in the L subunit*

Calculations were carried out using the program *SSBOND* (Hazes & Dijkstra, 1988).

Disulfide	Residue No.*	Distances (Å) Cα—Cα	S—S	Torsion angles (°) χ_1	χ_2	χ_3	$\chi_{2'}$	$\chi_{1'}$	Energy (kJ mol^{-1})
1	23–88	5·56	2·03	−63·3	−57·8	−84·9	−64·7	−50·4	3·35
2	29–61	5·88	2·02	78·5	−78·9	−96·4	−36·4	−63·4	17·38
3	36–119	5·13	1·98	−88·1	−30·4	174·4	−91·9	−152·8	83·32
4	38–86	5·96	1·99	−61·1	−155·7	−87·3	141·3	−73·4	23·15
5	46–77	5·72	2·01	−79·9	−46·0	−90·4	−80·6	−56·3	6·70
6	78–108	6·19	2·00	−49·5	−100·5	−100·5	−66·3	−83·2	24·79

* The numbering scheme is that used in the X-ray sequence of the L subunit (Vellieux *et al.*, 1989).

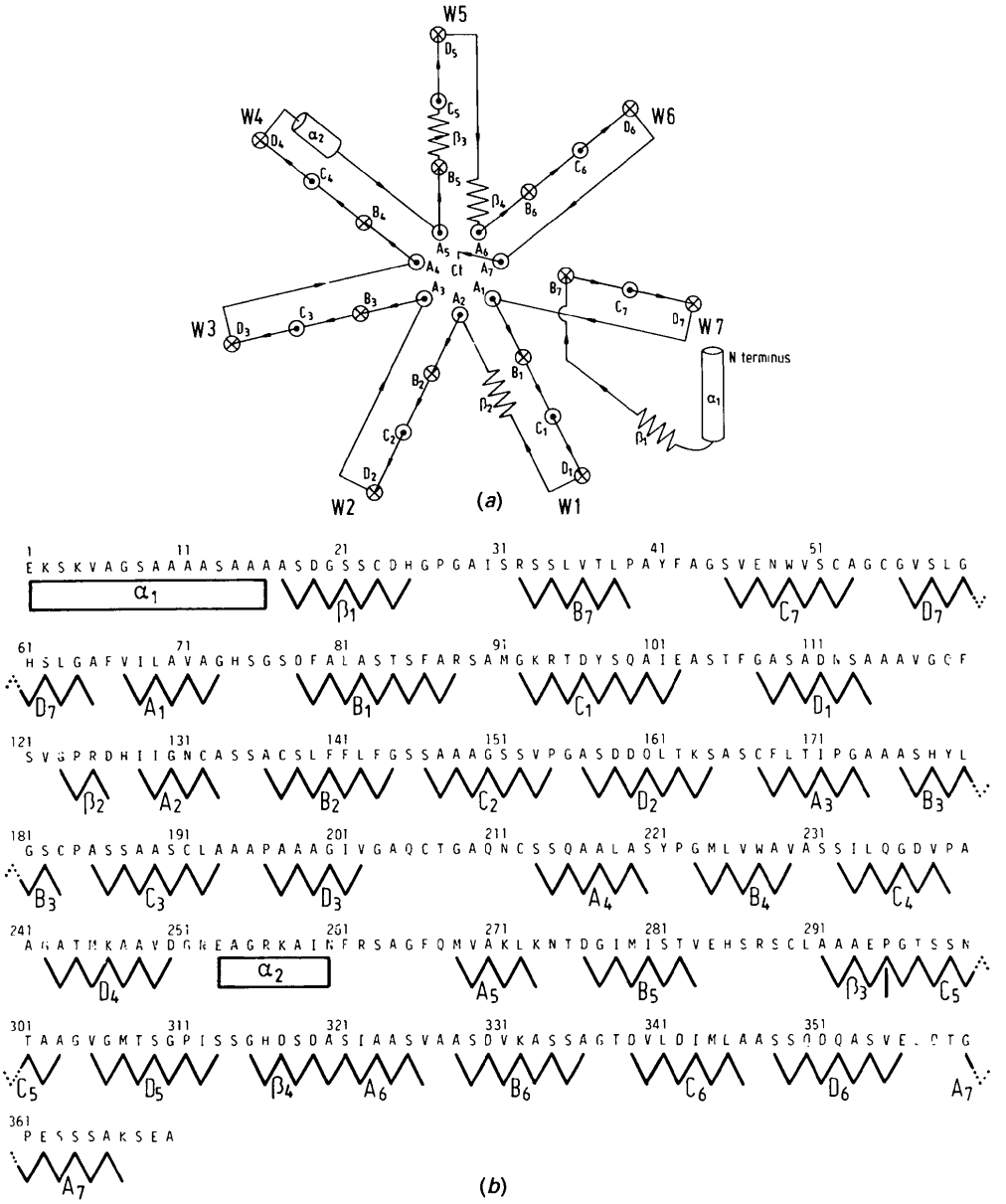


Fig. 7. Topology and X-ray sequence of the heavy subunit of MADH. (*a*) Schematic representation of the topology of the H subunit (adapted from Vellieux *et al.*, 1989) showing the secondary structure elements found in the subunit. The main structural feature of this subunit is the 'spinning wheel' formed by seven four-stranded antiparallel β-sheets (or W's), which has circular sevenfold symmetry. (*b*) The X-ray sequence as obtained from the interpretation of the electron density. The secondary structure elements are indicated in bold lines below the sequence: zigzags indicate residues found in β-sheets, and boxes indicate the α-helices. Note that 256 out of 370 residues, *i.e.* 70% of the residues found in the subunit are part of β structure.

This X-ray sequence, as derived from the 2·25 Å resolution map obtained after refinement of model *B*, is shown in Fig. 7, together with the elements of secondary structure found in the subunit. Since the side-chain assignment was carried out solely on the basis of the electron density, all residues which have been attributed an aromatic or long side chain will probably not differ too much from the correct sequence. On the other hand, short side-chain residues such as alanine, serine and proline were used either if the density indicated such a residue, or if the interpretation of the side-chain density proved difficult. Keeping this in mind, it is interesting to note that the *N*-terminal helix, α_1, which is fully accessible to the solvent in the MADH crystals, contains many of these short side-chain residues in our X-ray sequence.

Using this X-ray sequence, we were able to locate the chemically sequenced 32 *N*-terminal residues of an *N*-terminally truncated H subunit, which had been obtained by proteolytic cleavage of that subunit at a single site (F. Huitema, J. A. Duine, H. J. Bak & J. J. Beintema, unpublished results). Table 13 shows the agreement observed between both sequences: a total of 14 residues had been correctly identified from the electron density alone. With only one exception, large residues were either correctly identified, or differ by one or two atoms from the correct residue type. The identification of the smallest residues also did not seem to be a problem. However, most of the errors were made with residues having medium to short side chains. For example, none of the prolines had been correctly identified from the electron density. From the sequence alignment results, the proposed proteolytic cleavage site is located in a surface loop which is fully accessible to the solvent in the H subunit (Fig. 8), and would therefore agree well with its susceptibility to proteolytic attack.

Table 13. *Comparison of chemical and X-ray sequence information*

The amino-acid composition shown is that calculated from a 32 residue peptide whose sequence has been determined by chemical methods (F. Huitema, J. A. Duine, H. J. Bak & J. J. Beintema, unpublished results). Xxx is a residue whose identity was not established.

Amino acids observed in the 'chemical' sequence	Corresponding residues in the X-ray sequence
4 Asx	2 Asx, Glx, Ala
2 Thr	2 Thr
2 Glx	Glx, Ser
4 Pro	3 Ala, Gly
2 Gly	2 Gly
2 Ala	Ala, Val
4 Val	Val, 2 Ser, Ala
2 Ile	Asx, Ser
2 Leu	Ala, Gly
1 Tyr	Tyr
3 Phe	2 Phe, Ile
1 Lys	Lys
2 Arg	Arg, Gln
1 Xxx	Ser

Concluding remarks

The crystal structure of quinoprotein methylamine dehydrogenase has been determined to a resolution of 2·25 Å by using a combination of three crystallographic methods, namely multiple isomorphous replacement, solvent flattening and partial-model phasing.

Solvent flattening was used to generate phases for the 13580 reflections lacking heavy-atom phases, and to provide both amplitudes and phases for the 321 unrecorded reflections to 2·5 Å resolution (*ca* 38 and 1% of all possible reflections to that resolution, respectively). This procedure was essential for the success of the structure determination, since it led to good quality electron density maps, calculated with complete data, into which a preliminary molecular model was built. From these maps, useful sequence information was also obtained in the form of an X-ray sequence, which was established owing to the lack of an amino-acid sequence for this enzyme.

Molecular-dynamics refinement of three successive molecular models of increasing completeness eventually led, by a procedure of phase combination of refined model phases with solvent-flattening phases, to an improved 2·25 Å resolution electron density distribution. This was interpreted to provide a complete model for the enzyme. This current model, which has an *R* factor of 28·6% for data between 6·0 and 2·25 Å resolution, agrees with all currently available chemical and biochemical information.

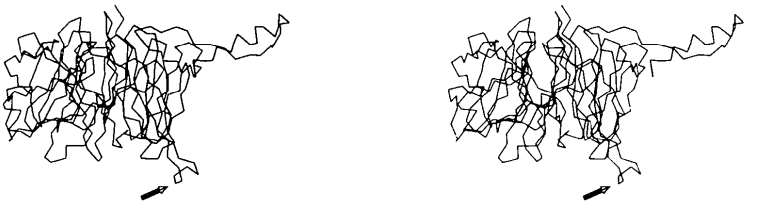


Fig. 8. The proposed proteolytic cleavage site in the H subunit of MADH. This figure shows the Cα tracing of a single H subunit, with the arrow pointing at the scissile bond located in a fully exposed surface loop.

We wish to thank Professor J. A. Duine, Drs J. Frank, Jzn. and F. Huitema (Department of Microbiology and Enzymology, Delft University of Technology) as well as Professor J. J. Beintema (Biochemical Laboratory, University of Groningen) for advice and for communicating sequence data prior to publication. We are grateful to Dr B. W. Dijkstra, who implemented the reciprocal space envelope calculation algorithm in the *BIOMOL* software, and to Dr M. Fujinaga (now at EMBL, Grenoble) for his help and advice during the early stages of refinement. The assistance of Drs K. S. Wilson, K. Petratos and Ch. Betzel (EMBL outstation, Hamburg) during data collection at the DESY synchrotron source is gratefully acknowledged.

This research was carried out with support from the Netherlands Foundation for Chemical Research (SON) and the Netherlands Organization for Scientific Research (NWO).

Note added in proof: This paper is also presented as part of the requirements for the award of a 'doctoraat' (PhD) degree for one of us (FMDV). Owing to the lack of an amino-acid sequence, the refinement process of the MADH model has not yet been completed. For this reason, only the Cα coordinates of this model will, for the time being, be deposited with the Protein Data Bank.* Regarding the exact nature of the quinone cofactor of MADH, recent reports have indicated that the covalently bound quinone cofactors found in two different quinoproteins, namely mammalian serum amine oxidase and fungal galactose oxidase, are in fact made up of a modified protein side chain (Janes *et al.*, 1990; Whittaker & Whittkaer, 1990).

* Cα coordinates have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1MAD), and are available in machine-readable form from the Protein Data Bank at Brookhaven or one of the affiliated centres at Melbourne or Osaka. The data have also been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 37035 (as microfiche). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

References

AGARWAL, R. C. (1978). *Acta Cryst.* A**34**, 791–809.
ANDERSON, C. M., STENKAMP, R. E. & STEITZ, T. A. (1978). *J. Mol. Biol.* **123**, 15–33.
ARNDT, U. W., CHAMPNESS, J. N., PHIZACKERLEY, R. P. & WONACOTT, A. J. (1973). *J. Appl. Cryst.* **6**, 457–463.
BEER, R. DE, DUINE, J. A., FRANK, J. & LARGE, P. J. (1980). *Biochim. Biophys. Acta*, **622**, 370–374.
BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794–802.
BRÜNGER, A. T. (1988). *J. Mol. Biol.* **203**, 803–816.
BRÜNGER, A. T., KARPLUS, M. & PETSKO, G. A. (1989). *Acta Cryst.* A**45**, 50–61.
BRÜNGER, A. T., KURIYAN, J. & KARPLUS, M. (1987). *Science*, **235**, 458–460.
DUINE, J. A. (1989). *PQQ and Quinoproteins*, edited by J. A. JONGEJAN & J. A. DUINE, pp. 351–360. Dordrecht: Kluwer Academic Publishers.
DUINE, J. A., FRANK, J. & VAN ZEELAND, J. K. (1979). *FEBS Lett.* **108**, 443–446.
DUINE, J. A., FRANK, J. & VERWIEL, P. E. J. (1980). *Eur. J. Biochem.* **108**, 187–192.
EADY, R. R. & LARGE, P. J. (1968). *Biochem. J.* **106**, 245–255.
FILMAN, D. J., SYED, R., CHOW, M., MACADAM, A. J., MINOR, P. D. & HOGLE, J. M. (1989). *EMBO J.* **8**, 1567–1579.
FUJINAGA, M., GROS, P. & VAN GUNSTEREN, W. F. (1989). *J. Appl. Cryst.* **22**, 1–8.
GAYKEMA, W. P. J., HOL, W. G. J., VEREIJKEN, J. M., SOETER, N. M., BAK, H. J. & BEINTEMA, J. J. (1984). *Nature (London)*, **309**, 23–29.
GAYKEMA, W. P. J., VOLBEDA, A. & HOL, W. G. J. (1986). *J. Mol. Biol.* **187**, 255–275.
GROS, P., BETZEL, CH., DAUTER, Z., WILSON, K. S. & HOL, W. G. J. (1989). *J. Mol. Biol.* **210**, 347–367.
GROS, P., FUJINAGA, M., DIJKSTRA, B. W., KALK, K. H. & HOL, W. G. J. (1989). *Acta Cryst.* B**45**, 488–499.
GUNSTEREN, W. F. VAN & BERENDSEN, H. J. C. (1987). *GROMOS. Groningen Molecular Simulation Library.* BIOMOS bv, Groningen, The Netherlands.
HAMILTON, W. C., ROLLETT, J. S. & SPARKS, R. A. (1965). *Acta Cryst.* **18**, 129–130.
HAZES, B. & DIJKSTRA, B. W. (1988). *Protein Eng.* **2**, 119–125.
HENDRICKSON, W. A. & LATTMANN, E. E. (1970). *Acta Cryst.* B**26**, 136–143.
HENRICK, K., COLLYER, C. A. & BLOW, D. M. (1989). *J. Mol. Biol.* **208**, 129–157.
HOLDEN, H. M., RYPNIEWSKI, W. R., LAW, J. H. & RAYMENT, I. (1987). *EMBO J.* **6**, 1565–1570.
ISHII, Y., HASE, T., FUKUMORI, Y., MATSUBARA, H. & TOBARI, J. (1983). *J. Biochem. (Tokyo)*, **93**, 107–119.
JANES, S. M., MU, D., WEMMER, D., SMITH, A. J., KAUR, S., MALTBY, D., BURLINGAME, A. L. & KLINMAN, J. P. (1990). *Science*, **248**, 981–987.
JONES, T. A. (1985). *Methods Enzymol.* **115**, 157–171.
KENNEY, W. C. & MCINTIRE, W. (1983). *Biochemistry*, **22**, 3858–3868.
LESLIE, A. (1987). *Acta Cryst.* A**43**, 134–136.
LUO, M., VRIEND, G., KAMER, G., MINOR, I., ARNOLD, E., ROSSMANN, M. G., BOEGE, U., SCRABA, D. G., DUKE, G. M. & PALMENBERG, A. C. (1987). *Science*, **235**, 182–191.
MCINTIRE, W. S. & STULTS, J. T. (1986). *Biochem. Biophys. Res. Commun.* **141**, 562–568.
MATTHEWS, B. W. (1966). *Acta Cryst.* **20**, 230–239.
MATTHEWS, B. W. & CZERWINSKI, E. W. (1975). *Acta Cryst.* A**31**, 480–487.
MEER, R. A. VAN DER & DUINE, J. A. (1986). *Biochem. J.* **239**, 789–791.
MEER, R. A. VAN DER & DUINE, J. A. (1988). *FEBS Lett.* **235**, 194–200.
MEER, R. A. VAN DER, JONGEJAN, J. A. & DUINE, J. A. (1987). *FEBS Lett.* **221**, 299–304.
PETSKO, G. A. (1976). *Acta Cryst.* A**32**, 473–476.
RAMAKRISHNAN, C. & RAMACHANDRAN, G. N. (1965). *Biophys. J.* **5**, 909–933.
RAYMENT, I. (1983). *Acta Cryst.* A**39**, 102–116.
READ, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
RICE, D. W. (1981). *Acta Cryst.* A**37**, 491–500.
RICHARDSON, J. S. (1981). *Adv. Protein Chem.* **34**, 167–339.
ROSSMANN, M. G. (1979). *J. Appl. Cryst.* **12**, 225–238.
ROSSMANN, M. G. & BLOW, D. M. (1961). *Acta Cryst.* **14**, 641–647.
SCHWAGER, P., BARTELS, K. & JONES, T. A. (1975). *J. Appl. Cryst.* **8**, 275–280.
SHIRAI, S., MATSUMOTO, T. & TOBARI, J. (1978). *J. Biochem. (Tokyo)*, **83**, 1599–1607.

SIM, G. A. (1959). *Acta Cryst.* **12**, 813–815.
SIM, G. A. (1960). *Acta Cryst.* **13**, 511–512.
TRONRUD, D. E., MONZINGO, A. F. & MATTHEWS, B. W. (1987). *Acta Cryst.* A**43**, 489–501.
VARGHESE, J. N., LAVER, W. G. & COLMAN, P. W. (1983). *Nature (London)*, **303**, 35–40.
VELLIEUX, F. M. D., FRANK, J., SWARTE, M. B. A., GROENDIJK, H., DUINE, J. A., DRENTH, J. & HOL, W. G. J. (1986). *Eur. J. Biochem.* **154**, 383–386.
VELLIEUX, F. M. D., HUITEMA, F., GROENDIJK, H., KALK, K. H., FRANK, J., JONGEJAN, J. A., DUINE, J. A., PETRATOS, K., DRENTH, J. & HOL, W. G. J. (1989). *EMBO J.* **8**, 2171–2178.
VRIEND, G. (1990). *J. Mol. Graphics*, **8**, 52–56.
WANG, B. C. (1982). *Methods Enzymol.* **115**, 90–112.
WHITTAKER, M. M. & WHITTAKER, J. W. (1990). *J. Biol. Chem.* **265**, 9610–9613.
WIERENGA, R. K. (1978). PhD Thesis, Univ. of Groningen, The Netherlands, pp. 60–92.

Acta Cryst. (1990). B**46**, 823–826

Polymorphism of 1-Chloro-2,4-dinitrobenzene: Structures of α and β Forms

BY A. WILKINS,* R. W. H. SMALL AND J. T. GLEGHORN

The Chemistry Department, The University, Lancaster, England

(*Received* 15 *February* 1990; *accepted* 25 *June* 1990)

Abstract

$C_6H_3ClN_2O_4$, $M_r = 202{\cdot}55$, λ(Mo $K\alpha$) = 0·7107 Å. α form: orthorhombic, $Pccn$, $a = 8{\cdot}937\ (5)$, $b = 11{\cdot}035\ (5)$, $c = 15{\cdot}844\ (5)$ Å, $V = 1562{\cdot}5$ Å^3, $F(000) = 816$, $Z = 8$, $D_x = 1{\cdot}721$ Mg m^{-3}, $\mu = 0{\cdot}407$ mm^{-1}, $T = 290$ K. β form: monoclinic, $P2_1/c$, $a = 8{\cdot}91\ (1)$, $b = 6{\cdot}870\ (5)$, $c = 12{\cdot}98\ (1)$ Å, $\beta = 96{\cdot}15\ (5)°$, $V = 790{\cdot}0$ Å^3, $F(000) = 408$, $Z = 4$, $D_x = 1{\cdot}701$ Mg m^{-3}, $\mu = 0{\cdot}403$ mm^{-1}, $T = 290$ K. For the α form, $R = 0{\cdot}045$ for 1474 observed reflexions; for the β form, $R = 0{\cdot}051$ for 976 observed reflexions. Different molecular conformations are found: in the α form the p-nitro group is rotated 41·8 (4)° relative to the benzene ring, in the β form the rotation is 68·8 (6)°. The p-nitro groups are rotated 11·1 (4) and 13·1 (6)°, in opposite directions. Molecular-energy calculations reveal negligible differences between the two conformations.

Introduction

This investigation is part of a series of studies on compounds which exhibit polymorphism. Two orthorhombic polymorphs of the title compound, with melting points 323 and 316 K, were reported by Groth (1917) who also referred to a more labile third form obtainable from the melt. The three polymorphs, designed α, β and γ, have been characterized by melting points 325, 316 and 300 K (*Dictionary of Organic Compounds*, 1965). The crystal structure of the α form of the title compound has previously been determined by Watson (1960) and Gopalakrishna (1969), both reporting cell dimensions, space group and R values of 0·12 but giving neither atomic coordinates nor bond distances and angles. More recently, however, an accurate determination at 110 K has been reported (Takazawa, Ohba & Saito, 1989).

Experimental

A commercial sample recrystallized from ethanol gave suitable crystals of the α form (melting point 325 K). Upon supercooling the melt to room temperature between a coverslip and microscope slide and pressing hard with a needle, slow polycrystalline growth was initiated. The melting point of this material was 316 K; when small fragments were seeded into a saturated ethanolic solution, single crystals formed. These, assumed to be of the β form, were rapidly removed from solution and dried; if they were left in the solution rapid transformation to the α form occurred. Transition of the β form in the solid state was found to occur when contact was made with crystals of the α form, a cloudy growth emanating from the point of contact. These transformed crystals were found to be the polycrystalline α form. No success was obtained in preparing the γ form from the melt. Space groups and cell dimensions were obtained from Weissenberg photographs taken with Cu $K\alpha$ radiation ($\lambda = 1{\cdot}5418$ Å). More accurate cell dimensions were obtained from 20 measured setting angles on a Stadi-2 two-circle diffractometer in the range $20 < 2\theta < 40°$ using graphite-monochromatized Mo $K\alpha$ radiation. The same instrument was used for intensity measurements, variable ω scan, $2\theta'$ fixed, stationary background count. Crystals were sealed in Lindemann glass capillaries. Separate standards for each layer

* Present address: Logica Energy and Industry Systems Ltd, Regal House, Duke St, Stockport, England.

0108-7681/90/060823-04$03.00