*Acta Cryst.* (1995). D**51**, 1041–1051

# Refined Structures at 2 and 2.2 Å Resolution of Two Forms of the H-protein, a Lipoamide-Containing Protein of the Glycine Decarboxylase Complex

BY SERGE PARES AND CLAUDINE COHEN-ADDAD*

*Institut de Biologie Structurale Jean-Pierre Ebel, Centre National de la Recherche Scientifique/Commissariat l'Energie Atomique, 41 rue des Martyrs, F38027 Grenoble CEDEX 1, France*

LARRY C. SIEKER

*Department of Biological Structure, SM 20, University of Washington, Seattle, WA 98195, USA, and Institut de Biologie Structurale Jean-Pierre Ebel, Centre National de la Recherche Scientifique/Commissariat l'Energie Atomique, 41 rue des Martyrs, F38027 Grenoble CEDEX 1, France*

AND MICHEL NEUBURGER AND ROLAND DOUCE

*Physiologie Cellulaire Végétale, Unité de Recherche associée Centre national de la Recherche Scientifique, No. 576, Département de Biologie Moléculaire et Structurale/Commissariat à l'Energie Atomique, 17 rue des Martyrs, F38054 Grenoble CEDEX 9, France*

(*Received* 17 *January* 1995; *accepted* 15 *May* 1995)

## Abstract

H-protein, a 14 kDa lipoic acid-containing protein is a component of the glycine decarboxylase complex. This complex which consists of four protein components (P-, H-, T- and L-protein) catalyzes the oxidative decarboxylation of glycine. The mechanistic heart of the complex is provided by the lipoic acid attached to a lysine residue of the H-protein. It undergoes a cycle of transformations, *i.e.* reductive methylamination, methylamine transfer, and electron transfer. We present details of the crystal structures of the H-protein, in its two forms, H-Pro$_{Ox}$ with oxidized lipoamide and H-Pro$_{Met}$ with methylamine-loaded lipoamide. X-ray diffraction data were collected from crystals of H-Pro$_{Ox}$ to 2 Å and H-Pro$_{Met}$ to 2.2 Å resolution. The final *R*-factor value for the H-Pro$_{Ox}$ is 18.5% for data with $F > 2\sigma$ in the range of 8.0–2.0 Å resolution. The refinement confirmed our previous model, refined to 2.6 Å, of a $\beta$-fold sandwich structure with two $\beta$-sheets. The lipoamide arm attached to Lys63, located in the loop of a hairpin conformation, is clearly visible at the surface of the protein. The H-Pro$_{Met}$ has been crystallized in orthorhombic and monoclinic forms and the structures were solved by molecular replacement, starting from the H-Pro$_{Ox}$ model. The orthorhombic structure has been refined with a final *R*-factor value of 18.5% for data with $F > 2\sigma$ in the range of 8.0–2.2 Å resolution. The structure of the monoclinic form has been refined with a final *R*-factor value of 17.5% for data with $F > 2\sigma$ in the range of 15.0–3.0 Å. In these two structures which have similar packing, the protein conformation is identical to the conformation found in the H-Pro$_{Ox}$. The main change lies in the position of the lipoamide group which has moved significantly when loaded with methylamine. In this case the methylamine-lipoamide group is tucked into a cleft at the surface of the protein where it is stabilized by hydrogen bonds and hydrophobic contacts. Thus, it is totally protected and not free to move in aqueous solvent. In addition, the H-protein presents some sequence and structural analogies with other lipoate- and biotin-containing proteins and also with proteins of the phosphoenolpyruvate:sugar phosphotransferase system.

## 1. Introduction

Lipoate (1,2 dithiolane-3-valeric acid) is the prosthetic group of the H-protein of the glycine decarboxylase complex which catalyzes the oxidative decarboxylation of glycine. It is also the prosthetic group of core enzymes involved in large complexes which catalyze the decarboxylation of 2-oxo-acids (Reed, 1974; Mattevi, de Kok & Perham, 1992).

Glycine decarboxylase consists of four proteins named P, H, T and L. The P-protein catalyzes the decarboxylation of the glycine molecule and reductive transfer of the resultant methylamine group to an S atom of the lipoamide. Then the lipoamide-methylamine arm moves to the active site of the T-protein which catalyzes the release of the amino N atom as $NH_3$ and the transfer of the methylene group to tetrahydrofolate to form methylenetetrahydrofolate. Finally, the L-protein (dihydrolipoamide dehydrogenase) reoxidizes the reduced dihydrolipoamide, thereby completing the catalytic cycle of the H-protein (Bourguignon, Neuburger & Douce, 1988; Douce & Neuburger, 1989) (Fig. 1). During the course of the catalytic process, the

* Author for correspondence.




ISSN 0907-4449 

H-protein plays a central role as the lipoyl group interacts with the three other components. The lipoate is joined to the $\varepsilon$-amino group of a specific lysine side chain (Lys63) by an amide linkage. This attachment provides a 'flexible arm' permitting the dithiolane ring of the lipoyl moiety to interact with the different catalytic sites.

The role of the H-protein is in some way similar to the lipoyl-lysine domain of the core enzyme, E2, of the pyruvate dehydrogenase complex (Mattevi *et al.*, 1992). In this case the dithiolane ring of the lipoamide is reductively acetylated. Then the acetyl group is transferred to CoA and the lipoamide is reoxidized by a lipoamide dehydrogenase.

In order to understand how the H-protein recognizes the different protein components of the complex (P-, T- and L-proteins) and the mechanism by which the lipoamide interacts with different catalytic domains, we have crystallized the H-protein from pea leaf mitochondria in its oxidized and methylamine loaded forms. The latter structure was previously found to be surprisingly stable (Neuburger, Jourdain & Douce, 1991) indicating that, once bound to the lipoamide, the methylamine group is protected against hydrolysis.

The structure of the oxidized H-protein ($H\text{-}Pro_{Ox}$) was previously solved using multiple isomorphous replacement methods and has been refined to 2.6 Å (Pares, Cohen-Addad, Sieker, Neuburger & Douce, 1994). The structure of the methylamine loaded H-protein ($H\text{-}Pro_{Met}$) has been solved by molecular replacement. It provides new insight on the stabilization of the methylamine group by the movement of the lipoate-lysine arm (Cohen-Addad, Pares, Sieker, Neuburger & Douce, 1995). In particular, it has been emphasized that the methylamine group is locked into a very stable configuration preventing a nucleophilic attack by water molecules, which explains the stabilization of the $H\text{-}Pro_{Met}$ itself. We report here detailed descriptions of the structure of $H\text{-}Pro_{Ox}$ refined to 2 Å and of the structure of $H\text{-}Pro_{Met}$ refined to 2.2 Å. Two crystal forms of $H\text{-}Pro_{Met}$ have been obtained and a comparison between them is given in the present work. The structure of the H-protein presents another example of the barrel–sandwich hybrid fold that has been recently defined by Chothia & Murzin (1993) and a comparison of its structure with other analogous proteins is also discussed.

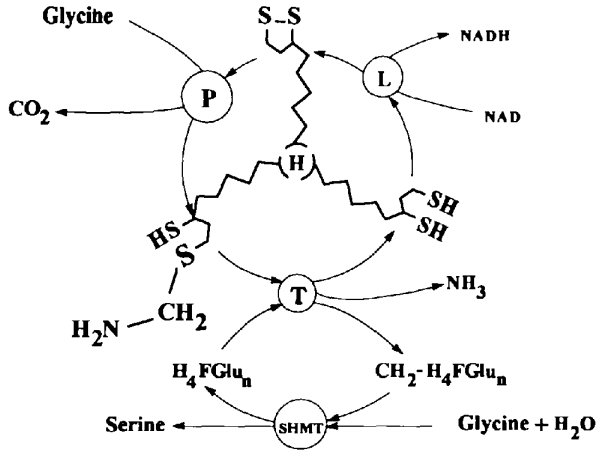


Fig. 1. Scheme outlining the reactions involved in oxidative decarboxylation and deamination of glycine in plant mitochondrial matrix. P, H, T and L are the protein components of the glycine decarboxylase system. The serine hydroxymethyltransferase (SHMT) is involved in the recycling of methylenetetrahydrofolate-polyglutamate ($CH_2H_4FGlu_n$; $n$ = number of glutamate residues) into $H_4FGlu_n$.

Table 1. *Data-collection statistics for H-Pro$_{Ox}$ and H-Pro$_{Met}$*

| | $H\text{-}Pro_{Ox}$ | $H\text{-}Pro_{Met}$ | |
|---|---|---|---|
| | Trigonal | Orthorhombic | Monoclinic |
| Resolution (Å) | 2.0 | 2.2 | 3.0 |
| No. of reflections | 87067 | 14236 | 8927 |
| No. of unique reflections | 17031 | 5684 | 4417 |
| $R_{sym}$* (%) | 8.1 | 5.4 | 5.9 |
| Completeness (%, $F > 2\sigma$) | 92 | 91 | 85 |
| Completeness (%, higher shell, $F > 2\sigma$) | 80 (2.1–2.0) | 73 (2.3–2.2) | 60 (3.14–3.0) |

* $R_{sym} = \sum |I_j(h) - \langle I(h)\rangle| / \sum I_j(h)$ summed over all measured reflections. $I_j$ represent the intensities of individual measurements $I$ and the corresponding mean values.

## 2. Experimental

### 2.1. *Crystallization and X-ray data collection*

H-protein was extracted and purified from pea leaf mitochondria (*Pisum sativum* L., var. Douce Provence) as previously described (Bourguignon *et al.*, 1988). Crystals of the oxidized form $H\text{-}Pro_{Ox}$ were obtained by vapour diffusion at 281 K in a solution containing 2 *M* $(NH_4)_2SO_4$ buffered at pH 5.2 with 0.1 *M* Tris maleate (Sieker, Cohen-Addad, Neuburger & Douce, 1991). The crystals belong to the space group $P3_121$ ($a = 57.14$, $c = 137.11$ Å), with two independent protein molecules in the asymmetric unit and 45% solvent content.

The methylamine loaded form $H\text{-}Pro_{Met}$ was purified according to Neuburger *et al.* (1991) and crystallized by vapour diffusion in two forms. An orthorhombic form was obtained in the same crystallization conditions as $H\text{-}Pro_{Ox}$. Under similar conditions, a monoclinic form was also obtained. The two corresponding space groups are $P2_12_12$ ($a$ = 61.34, $b$ = 55.38, $c$ = 33.77 Å, one protein molecule in the asymmetric unit and 39% solvent content) and $P2_1$ ($a$ = 55.70, $b$ = 68.30, $c$ = 33.90 Å, $\beta$ = 101°, two independent protein molecules in the asymmetric unit and 45% solvent content).

All X-ray data sets were collected on a FAST Enraf–Nonius area detector (Enraf–Nonius FR 571 rotating-anode generator, copper radiation, graphite monochromator). Diffraction images were processed with the program *MADNES* (Messerschmidt & Pflugrath, 1987).

Table 2. *Summary of the refinements of H-Pro$_{Ox}$ and H-Pro$_{Met}$*

| | H-Pro$_{Ox}$ Trigonal | | H-Pro$_{Met}$ Monoclinic | | |
|---|---|---|---|---|---|
| | Molecule *A* | Molecule *B* | Orthorhombic | Molecule *A* | Molecule *B* |
| *R* factor* (%) (No. of reflections) | 18.5 (16839) | | 18.4 (5457) | 17.5 (3450) | |
| Resolution range (Å) | 8.0–2.0 | | 8.0–2.2 | 15.0–3.0 | |
| No. of water molecules | 211 | | 146 | 0 | |
| R.m.s. deviations from ideal values | | | | | |
| Bonds (Å) | 0.012 | 0.011 | 0.014 | 0.016 | 0.017 |
| Angles ( ) | 2.82 | 2.77 | 2.90 | 3.73 | 3.61 |
| Dihedrals ( ) | 25.9 | 25.8 | 26.1 | 26.7 | 27.2 |
| Average *B* factors (Å$^2$) | | | | | |
| Main chain | 20.6 | 20.2 | 17.1 | 13.1 | 10.5 |
| Side chain | 26.6 | 28.0 | 18.8 | 17.6 | 13.8 |
| Lipoamide arm | 40.0 | 36.6 | 17.0 | 25.0 | 12.0 |
| Water molecules | 32.4 | | 22.2 | | |

* $R$ factor $= \sum |F_o - F_c| / \sum F_o$ where $F_o$ and $F_c$ are the observed and calculated structure-factor amplitudes.

For H-Pro$_{Ox}$, two X-ray data sets were collected at 2.2 and 2.0 Å, scaled and merged using the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994). For H-Pro$_{Met}$, data sets were collected on the orthorhombic crystal form to 2.2 Å resolution and on the monoclinic crystal form to 3 Å resolution. Results are summarized in Table 1.

## 2.2. *Refinement*

The starting model for the refinement at 2.0 Å was taken from the refined atomic coordinates at 2.6 Å resolution (Pares *et al.*, 1994) where the lipoate moiety attached to residue 63 was not included in the calculation. No water molecules were included. The refinement was carried out using *X-PLOR* (Brünger, 1993) with reflections in the resolution range 8.0–2.0 Å. The model fitting was carried out on an Evans and Sutherland ESV-10 molecular graphics systems, using the program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991; Jones & Kjeldgaard, 1993). The initial model gave an *R* factor of 31%. A simulated-annealing refinement was applied and the model was slow-cooled from the initial temperature of 4000 to 300 K reducing the temperature by 25 K after every 50 dynamic steps where each step corresponded to 0.5 fs. The simulated annealing was followed by an energy minimization and an individual *B*-factor refinement, leading to an *R* factor of 28.8%. At this stage, after the lipoate group was relocated on an $F_o - F_c$ map, some residues located in loops were not well fitted into the electron density (residues 90–99 and 112–122). They have been relocated using $2F_o - F_c$ and $F_o - F_c$ omit maps and a new refinement led to an *R* factor of 24.6%. Water molecules were located by considering density peaks higher than $2\sigma$ in the $F_o - F_c$ map and added to the model refinement. During that refinement, water molecules that had a *B* factor higher than 40 Å$^2$ and an occupancy factor less than 0.25 were rejected. The final *R* factor was 18.5% in the range 8–2 Å ($F > 2\sigma$) (Table 2).

The crystal structure of H-Pro$_{Met}$ (orthorhombic form) was solved by molecular-replacement techniques using the atomic coordinates of H-Pro$_{Ox}$ at 2.0 Å (Cohen-Addad *et al.*, 1995), and the *AMoRe* program package (Navaza, 1994). As for H-Pro$_{Ox}$, the lipoate group attached to residue 63 was not included in the calculations. With data between 8.0 and 3.0 Å a solution was found with a correlation of 60.6% and an *R* factor of 41.2%. After a rigid-body refinement, a correlation of 67.9 and an *R* factor of 37.4% were obtained. Initial refinement was performed with *X-PLOR*, using reflections in the resolution range 8.0–2.2 Å. An $F_o - F_c$ Fourier map was calculated and showed clearly the electron density of the lipoate arm with the methylamine group located in a cleft at the surface of the H-protein and attached to the distal S atom S8 (numbering of C and S atoms in the lipoamide group shown in Fig. 2 follows nomenclature used by Mattevi, Obmolova, Kalk, Teplyakov & Hol, 1993). After refinement of atomic positions and individual temperature factors, the *R* factor decreased to 23.5%. Extra electron density contiguous to the free sulfur of the lipoate group was interpreted as a molecule of $\beta$-mercaptoethanol linked by a disulfide bond to the proximal S atom S6 since $\beta$-mercaptoethanol was used to prevent the oxidation of the S atoms during the course of H-protein purification. Water molecules were then added, as described previously and rejected when the *B* values were higher than 40 Å$^2$ and the occupancy factor less than 0.5. The final *R* value was 18.4% for data with $F > 2\sigma$ in the range 8.0–2.2 Å (Table 2).

The H-Pro$_{Met}$ in the monoclinic form was also solved by molecular-replacement techniques using reflections in the resolution range 15.0–3.0 Å. A solution was found with a correlation of 60.9 and an *R* factor of 39.7%.

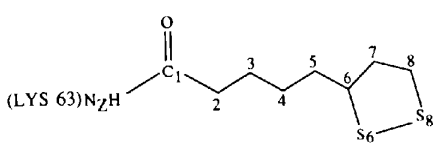


Fig. 2. The lipoyl moiety attached to the lysine. The oxidized form is represented. The atom numbering follows previous nomenclature (Mattevi, Obmolova, Kalk, Teplyakov & Hol, 1993).

After rigid-body refinement, a correlation of 69.9% and an $R$ factor of 35.3% were obtained. Initial refinement with *X-PLOR*, using reflections in the 15–3.0 Å range, was performed without any lipoate group included in the calculations. An $R$ factor of 20% was obtained and an $F_o - F_c$ map revealed clearly the positions of the lipoate group for each of the two molecules of the asymmetric unit and these groups were then included in the refinement. The final $R$ factor was 18.5% for data with $F > 2\sigma$ in the range 15.0–3.0 Å (Table 2). In contrast to the orthorhombic form, no density corresponding to a $\beta$-mercaptoethanol molecule was found in the monoclinic crystal form.

## 3. Results and discussion

### 3.1. *Quality of the structures*

3.1.1. *Model of H-Pro$_{Ox}$.* Statistics of the refinement are shown in Table 2. The average $B$-factor values for each residue of the two molecules of the asymmetric unit are shown in Fig. 3. They are nearly the same, for the two molecules of the asymmetric unit. The high $B$ values (36.6 and 40 Å$^2$) observed for the lipoamide arms reflect the flexibility of this long chain at the surface of the protein. A Ramachandran plot (Ramachandran & Sasisekharan, 1968) is shown in Fig. 4. All the residues are in the allowed regions, with 90% being within the most energetically favourable regions (Morris, MacArthur, Hutchinson & Thornton, 1992). The real-space coefficients calculated with the program *O* (Jones & Kjeldgaard, 1993) estimate how well the molecules fit the density for each residue (Jones *et al.*, 1991). The fit of each residue to the density is shown in Fig. 5. For both molecules, the average values of the correlation coefficients for main- and side-chain atoms are not less than 0.90 except for flexible parts (amino and carboxyl termini, lipoamide groups). The high correlation coefficients observed for the main chains give a good representation of the continuous electron density observed throughout both molecules of the asymmetric unit.

The mean positional error of atoms, calculated from a Luzzati plot (not shown) (Luzzati, 1952) is between 0.2 and 0.25 Å

3.1.2. *Model of H-Pro$_{Mel}$ in orthorhombic and monoclinic crystal forms.* Statistics of the refinements for both forms, orthorhombic and monoclinic, are given in Table 2. In the orthorhombic crystal form, the average $B$-factor value for both protein and the lipoamide group is 17.0 Å$^2$ (Table 2). This suggests that the lipoamide arm is strongly stabilized once it is loaded with a methylamine group which interacts with the protein. The $B$-factor values for each residue are shown in Fig. 6. A

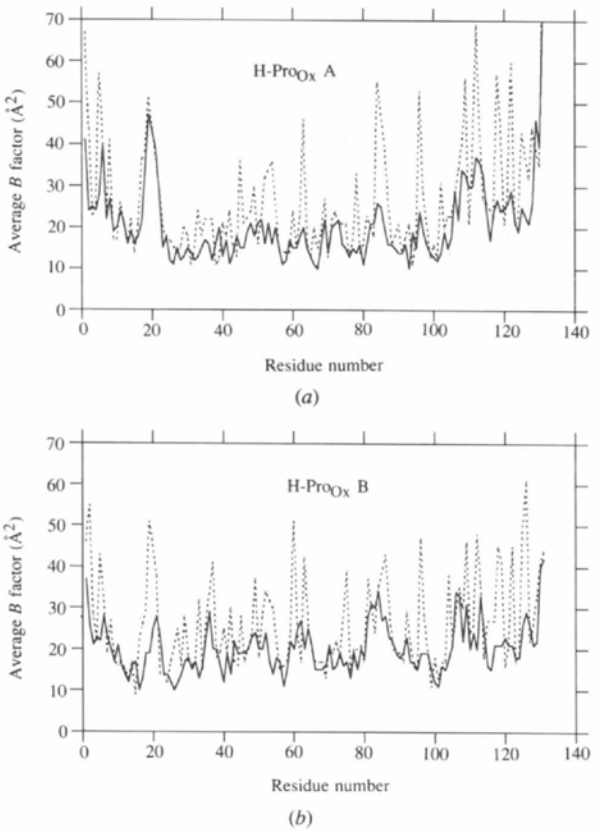


Fig. 3. H-Pro$_{Ox}$. Plot of the average $B$-factor values for the residues of the two independent molecules in the asymmetric unit. (*a*) molecule *A*, (*b*) molecule *B*. Values are averaged within each residue and plotted against the residue number. Full line: backbone, dotted line: side chains.

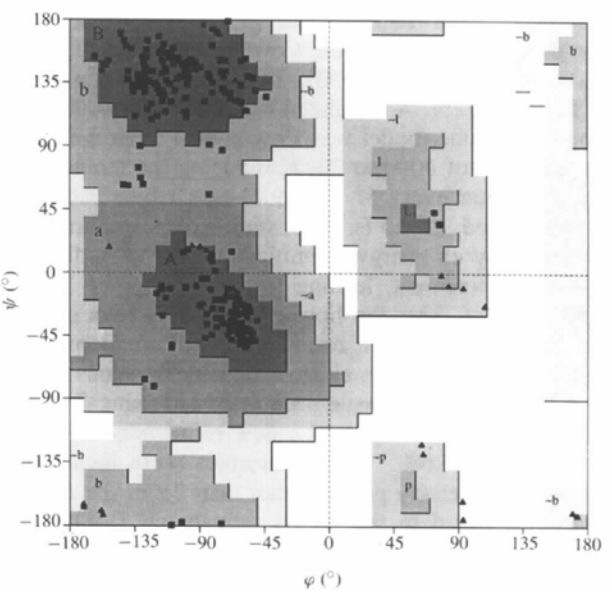


Fig. 4. H-Pro$_{Ox}$. Ramachandran plot for the $\varphi$, $\psi$ torsional angles. The values for the residues of the two molecules of the asymmetric unit are plotted. The figure is produced using the program *PROCHECK* (Morris, MacArthur, Hutchinson & Thornton, 1992). Triangles denote glycine residues, squares all others.

Ramachandran plot is shown in Fig. 7. All residues are in the allowed regions, with 89.5% in the most energetically favourable regions. The plot of the real-space fit of each residue to the density is shown in Fig. 8. The average values for main- and side-chain atoms are not less than 0.90 including the lipoamide arm. The mean positional error of atoms calculated from a Luzzati plot is between 0.2 and 0.25 Å.

In the monoclinic crystal form of H-Pro$_{Met}$, the average $B$ values of the lipoamide group in the two molecules of the asymmetric unit are lower than in the H-Pro$_{Ox}$ structure (Table 2). This confirms the stabilization of the lipoamide when loaded with methylamine group. In a Ramachandran plot (not shown here) all residues fall within the expected boundaries. A higher resolution study of the monoclinic crystal form is needed in order to obtain a more accurate description of this structure.

### 3.2. *Overall structures, description and comparison*

3.2.1. *The overall structure.* The refinement of the H-Pro$_{Ox}$ structure at 2.0 Å confirmed the initial model at 2.6 Å (Pares *et al.*, 1994). As previously described,

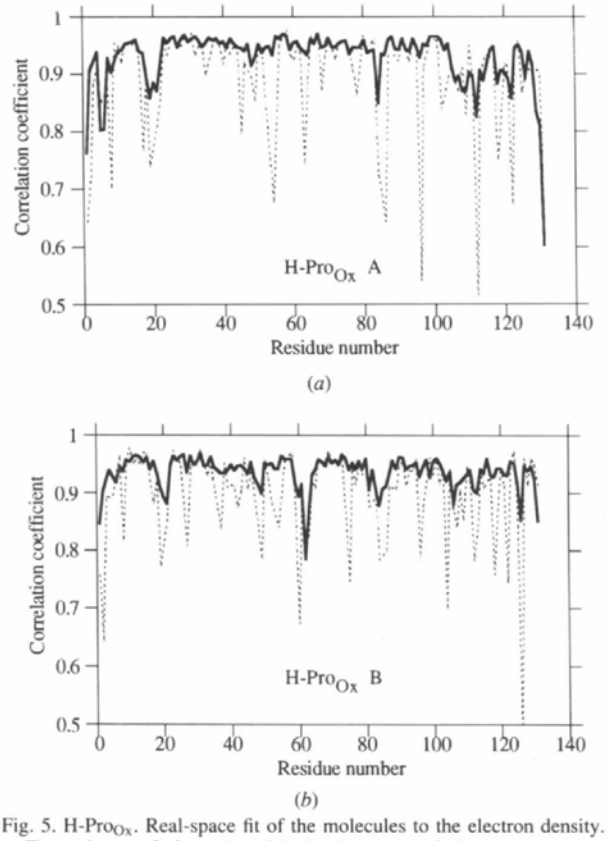


Fig. 5. H-Pro$_{Ox}$. Real-space fit of the molecules to the electron density. The real-space fit for each residue is given by the following expression (Jones *et al.*, 1991) $\sum|\rho_{obs} - \rho_{calc}|/\sum|\rho_{obs} + \rho_{calc}|$. where $\rho_{obs}$ is the observed density ($2F_o - F_c$ map) and $\rho_{calc}$ is the calculated density around the residue. Parameter values used as implemented in *O* (Jones & Kjelgaard, 1993). (*a*) molecule *A*, (*b*) molecule *B*. Full line, backbone; dotted line, side chains.

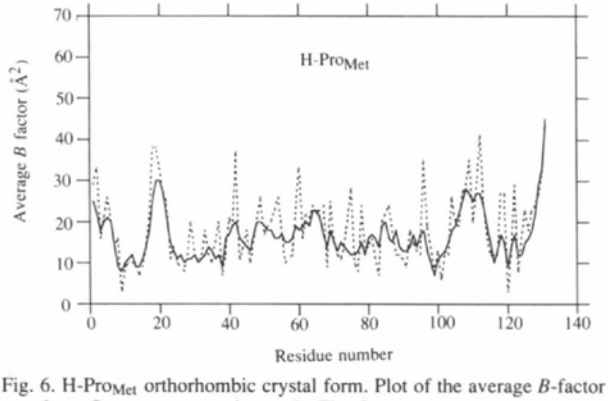


Fig. 6. H-Pro$_{Met}$ orthorhombic crystal form. Plot of the average $B$-factor values. Same representation as in Fig. 3.

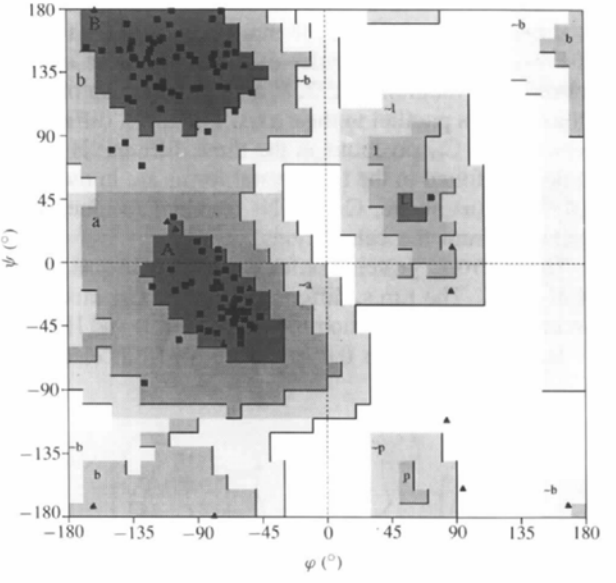


Fig. 7. H-Pro$_{Met}$ orthorhombic crystal form. Ramachandran plot for the $\varphi$, $\psi$ torsional angles. Same representation as in Fig. 4.

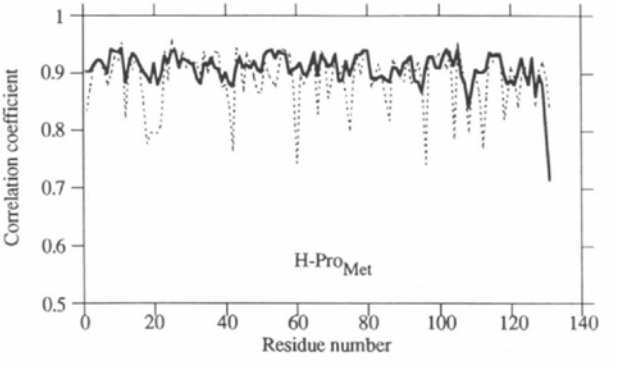


Fig. 8. H-Pro$_{Met}$ orthorhombic crystal form. Real-space fit of the molecule to the electron density. Full line: backbone, dotted line: side chains. Real-space fit and parameter values used in program *O* are the same as defined in Fig. 5.

the core of the H-protein, shown in Fig. 9(*a*), consists mainly of two antiparallel $\beta$-sheets, one with six $\beta$-strands and the other with three $\beta$-strands, forming a hybrid barrel–sandwich structure as previously described by Chothia & Murzin (1993). The lipoamide moiety attached to Lys63 located at the extremity of a hairpin loop is clearly visible at the surface of the protein. The two molecules of the asymmetric unit are related by a rotation of 140° and a translation of 2.8 Å. The two molecules are essentially identical. The r.m.s. differences in their $C_\alpha$ positions are equal to 0.3 Å (C- and N-terminal residues which have high *B* values were excluded from the calculations). The largest difference in the backbone is observed in the hairpin loop (1.3 Å) and this can be explained by the rather high flexibility of the lysine-lipoate group attached there. In fact the lipoamide moieties show r.m.s. differences of 1.6 Å.

For the H-$Pro_{Met}$ protein, the conformation of the protein molecule is identical in the two space groups and is shown in Fig. 9(*b*). In the monoclinic crystal form of H-$Pro_{Met}$, the two molecules of the asymmetric unit are related by a rotation of 177.9° and a translation of 1.4 Å along an axis parallel to the *c* axis. The r.m.s. differences between the $C_\alpha$ positions in the three different H-$Pro_{Met}$ molecules found in the two crystal forms are in the range 0.4–0.5 Å (as above, C- and N- terminal residues were excluded from the calculations).

The H-$Pro_{Met}$ protein conformation is identical to that of H-$Pro_{Ox}$. The r.m.s. differences on the $C_\alpha$ atoms between H-$Pro_{Met}$ (orthorhombic crystal form) and H-$Pro_{Ox}$ molecules is equal to 0.4 Å. As pointed out elsewhere

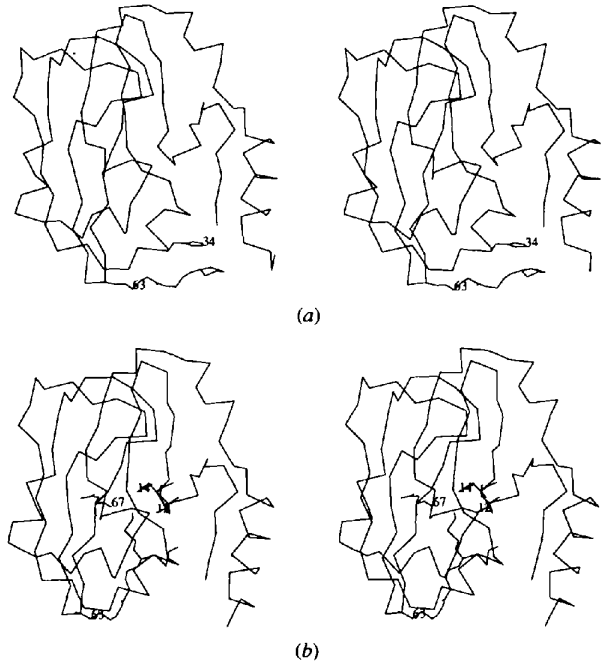


Fig. 9. Stereo view of the structure of the H-protein and location of the lipoamide arm. (*a*) H-$Pro_{Ox}$, (*b*) H-$Pro_{Met}$.

Table 3. *Summary of secondary-structure elements in H-protein*

The secondary structures have been assigned according to the hydrogen-bond pattern defined by Kabsch & Sander (1983).

| $\alpha$-Helices | $3_{10}$-Helix | Antiparallel $\beta$-strands | |
|---|---|---|---|
| *H*1 29–36 | | 1st $\beta$-sheet | *B*1 8–10 |
| *H*2 81–86 | | | *B*2 13–19 |
| *H*3 88–93 | | | *B*3 22–28 |
| *H*5 117–129 | | | *B*5 48–51 |
| | *H*4 107–114 | | *B*8 73–79 |
| | | | *B*9 100–106 |
| | | 2nd $\beta$-sheet | *B*4 37–43 |
| | | | *B*6 54–61 |
| | | | *B*7 64–70 |

Table 4. *H-$Pro_{Ox}$ contacts between the lipoamide groups and the protein molecules*

*A* and *B* refer to the two independent molecules in the asymmetric unit. Distances less than 3.5 Å are listed.

| | Protein molecules | | |
|---|---|---|---|
| | Residue | | Distance (Å) |
| Lipoamide molecule *A* | | | |
| S6 | His34 (*A*) | CD2 | 3.18 |
| S6 | His34 (*A*) | CB | 3.49 |
| S8 | His34 (*A*) | CD2 | 3.16 |
| C5 | Ser21 (*B*) | OG' | 3.48 |
| C7 | Ser21 (*B*) | OG' | 3.34 |
| Lipoamide molecule *B* | | | |
| NZ | His24 (*B*) | O | 2.94 |
| C1 | His34 (*B*) | O | 3.48 |
| C8 | Ser21 (*A*) | OG'' | 3.27 |

Symmetry operators: (i) $-x, y-x, -z+\frac{1}{3}$; (ii) $y, x, z$.

(Cohen-Addad *et al.*, 1995), the main difference between H-$Pro_{Ox}$ and H-$Pro_{Met}$ is the location of the methylamine-lipoate group which has moved significantly to bind into a cleft at the surface of the protein (Fig. 9*b*).

In the structure, 38% of the residues are in $\beta$-strand configurations and 27% of them form helices. The secondary elements are listed in Table 3 and a plot of the hydrogen-bond connections in the main chain is shown in Fig. 10. The two $\beta$-sheets of the structure contain the $\beta$-strands *B*1, *B*2, *B*3, *B*5, *B*8 and *B*9 for the first sheet, and $\beta$-strands *B*4, *B*6, *B*7 for the second one, respectively. The $\beta$-strands *B*6 and *B*7 are joined by a short hairpin which carries the prosthetic lipoamide arm. This hairpin can be classified in the 2:4 group as defined by Sibanda, Blundell & Thornton (1989). Besides the hydrogen bonds within each $\beta$-sheet, several hydrogen bonds between $\beta$-strands and the helical parts of the structure contribute to the stabilization of the molecule. These bonds involve residues His13, interacting with residues Tyr120 and Glu127, Val38 interacting with Asn91 and Ser61 interacting with Leu35. In addition, hydrophobic contacts (less than 4 Å) between the two sheets stabilize the barrel–sandwich structure. These contacts are established between the following residues: Ile27 with Val68, Leu43 with Val79, Phe56 with Val50

and Ile101, Met100 with Val41, Val59 and Val79. The structure contains 14 glycines which represent 11% of the molecule.

3.2.2. *Lipoamide arm interactions.* In the two molecules of the asymmetric unit of H-Pro$_{Ox}$, the lipoamide group has very few contacts with the protein residues (listed in Table 4) They involve mainly His34, the dithiolane ring being parallel to the His34 ring in molecule *A* and perpendicular to this residue in molecule *B* due to packing differences.

In both forms of the H-Pro$_{Met}$, orthorhombic and monoclinic, the methylamine-lipoate arm has moved dramatically so as to be tucked into a cleft which is delineated on one side by the $\beta$-sandwich structure and on the other side by the helix *H*1. It includes residues 12–14, 27–36 and 59–67. Our results indicate for the first time that the methylamine group is linked

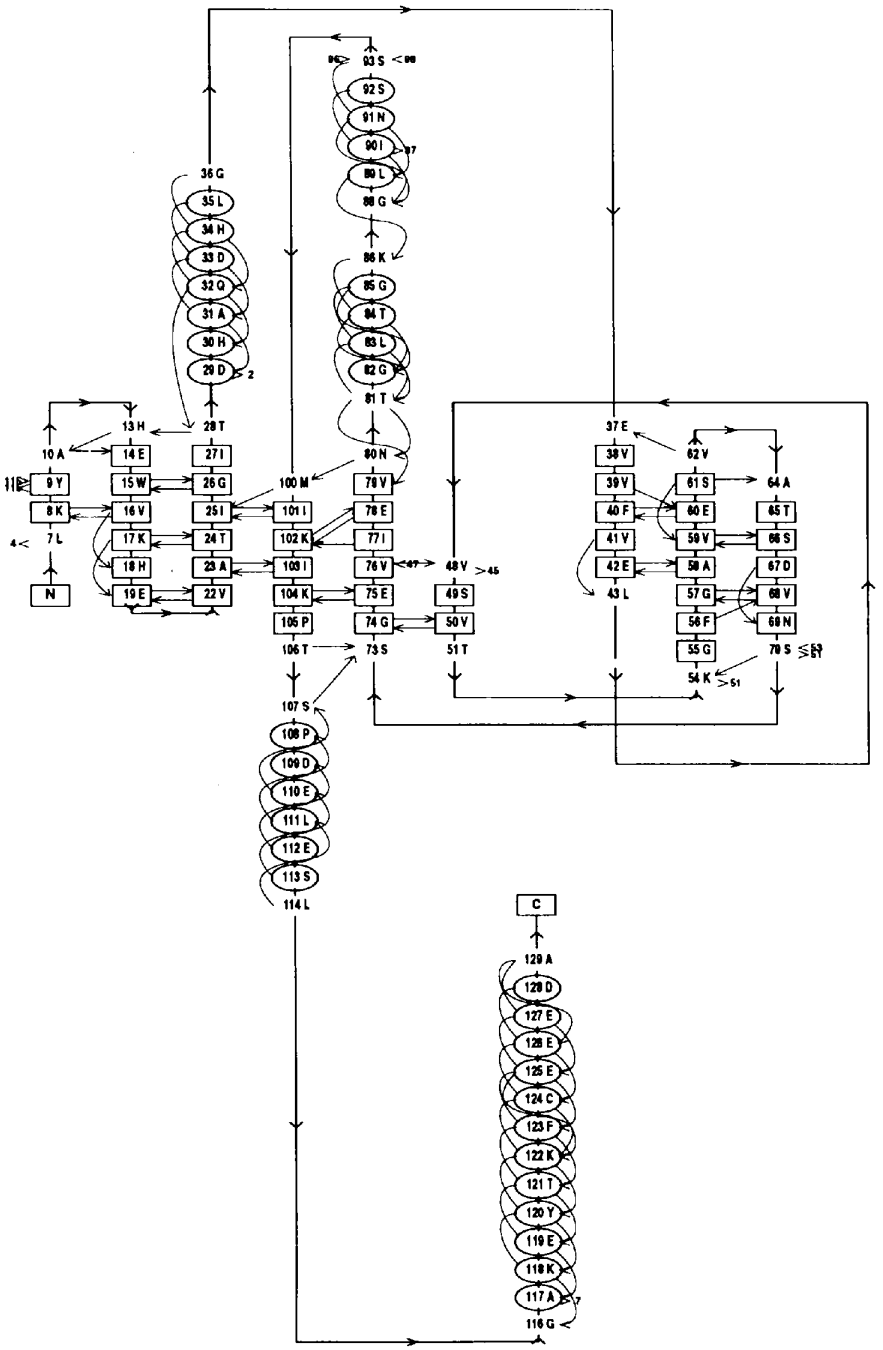


Fig. 10. H-protein. Hydrogen-bond connections in the main chain, drawn with the program *HERA* (Hutchinson & Thornton, 1990).

to the distal S atom S8 of the dithiolane ring. Its $NH_3^+$ moiety is tightly fixed by an ionic bond to the carboxyl group of Glu14, and two hydrogen bonds to carbonyl groups of Ser12 and Asp67 (Fig. 12). Hydrophobic contacts are established between the aliphatic chain of the lipoamide arm and some hydrophobic residues as shown in Fig. 12. In addition, the $N_\varepsilon$ atom of the lipoamide arm forms a hydrogen bond with the carbonyl group of His34 contributing to the stabilization of the long extended chain of C atoms. There is no interaction between the lipoate moiety and symmetrically related protein molecules. Under these conditions, the unstable lipoate-methylamine group which is very sensitive to nucleophilic attack by water molecules, is totally protected. This is in agreement with calculation of the solvent-accessible area for the methylamine group performed by the method of Lee & Richards (1971) with *X-PLOR*. With a probe radius of 1.4 Å used in the calculations, a very low value of 4 $Å^2$ was obtained for the methylamine group. This protective cleft is essentially identical in the H-$Pro_{Ox}$ molecule. In this case, the residues bound to the methylamine-lipoate moiety through hydrogen bonds are free to form hydrogen bonds with solvent molecules. The Connolly surface around the cavity is represented in Fig. 11.

In the orthorhombic form, a $\beta$-mercaptoethanol molecule is linked to the proximal S atom S6 of the lipoamide arm and forms a hydrogen bond with the carboxyl group of Glu127 (Fig. 12). As we have mentioned above, $\beta$-mercaptoethanol was used to prevent the oxidation of the S atoms. Due to the low redox potential of the free SH group of the lipoate, the oxidized form of $\beta$-mercaptoethanol (dimeric form) may have become fixed to the proximal SH group of lipoamide during one of the purification steps of the methylamine-loaded H-protein. In the monoclinic form, where no $\beta$-mercaptoethanol molecule is observed, the free S atom of the lipoate arm is in a reduced form. As the structure of the protein and the position of the lipoamide are similar in the two forms, we can conclude that the presence of the $\beta$-mercaptoethanol molecule has no effect on the protein structure and on the interaction of the lipoamide with the protein. Moreover, the observation of the same configuration in two different crystal forms is in favour of a stable structure which should be the same in solution.

The stabilization of the lipoamide group bound into a cleft of the protein casts doubts on the hypothetical movement of the lipoamide-methylamine chain in the aqueous solvent acting as a 'swinging arm' conveying the reactive dithiolane ring from one catalytic center to another in the enzyme complex (Cohen-Addad *et al.*, 1995). Furthermore we might suppose that during the methylamine transfer, the T-protein closely associates with the H-protein at the cleft region. In fact, preliminary small-angle X-ray scattering studies indicate that, in solution, the T- and the H-protein interact in a 1/1 ratio (Ober, personal communication).

3.2.3. *Crystal packing in monoclinic and orthorhombic forms of H-*$Pro_{Met}$. In both monoclinic and orthorhombic crystal forms of the H-$Pro_{Met}$, the packing in each unit cell is similar. In the monoclinic structure, the two independent molecules of the asymmetric unit are related by a pseudo twofold axis parallel to the *c* axis and the packing of the four molecules of the unit cell is very

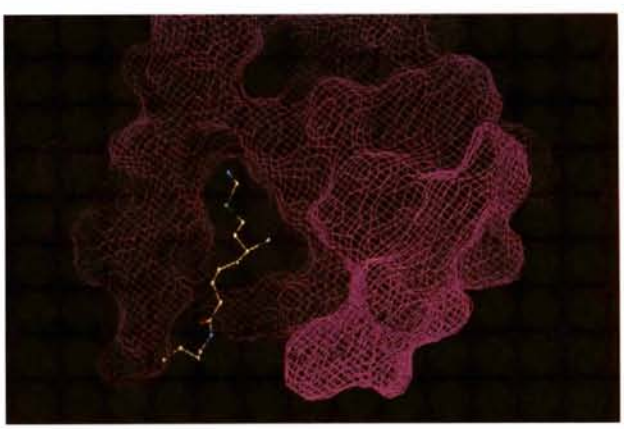

Fig. 11. Connolly surface of the cavity containing the lipoamide group. This cavity is essentially identical in H-$Pro_{Met}$ and H-$Pro_{Ox}$. The figure was generated with the program *MS* (Connolly, 1983) and *TURBO-FRODO* (Roussel & Cambiau, 1989).

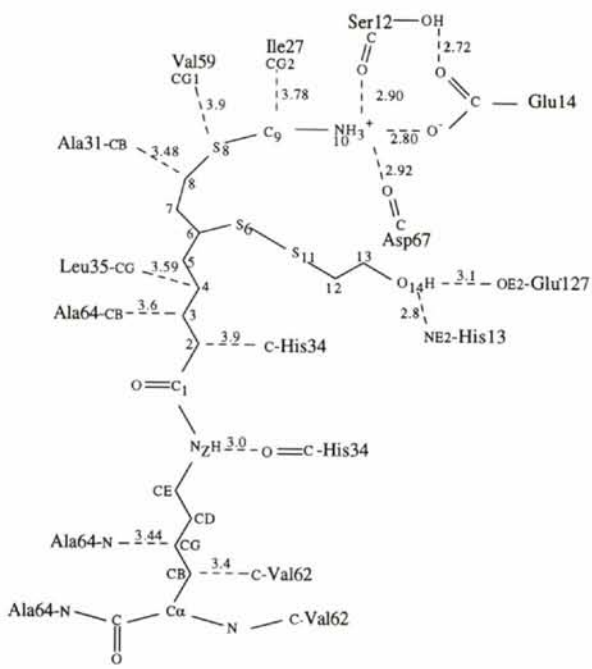


Fig. 12. H-$Pro_{Met}$. Hydrogen bonds and van der Waals interactions between the methylamine-lipoamide group and the protein residues. Distances less than 3.2 Å for hydrogen bond and less than 4 Å for van der Waals contacts are shown and represented by dashed lines. The numbering in the lipoamide group is defined in Fig. 2.

similar to that observed for the four molecules of the orthorhombic structure. Fig. 13 shows the similarity of the packing around the twofold axis parallel to *c* of the orthorhombic form and the corresponding pseudo twofold axis of the monoclinic form.

3.2.4. *Structural analogies between the H-protein and other proteins.* The folding of the $\beta$-portion observed in the crystal structure of the H-protein looks like a hybrid of a simple barrel and sandwich folds described previously by Chothia & Murzin (1993) for the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase complex (PDH) (Dardel, Davis, Laue & Perham, 1993) and for two non lipoylated proteins of the phosphoenolpyruvate:glucose phosphotransferase (PTS) system, the *Escherichia coli* phosphocarrier protein $III^{glc}$ (Worthylake *et al.*, 1991) and the IIA domain of the glucose permease of *Bacillus subtilis* (Liao *et al.*, 1991). As mentioned by Murzin (1994), the $\beta$-sheets are twisted and coiled as in barrel folds, but are packed in a sandwich-like manner.

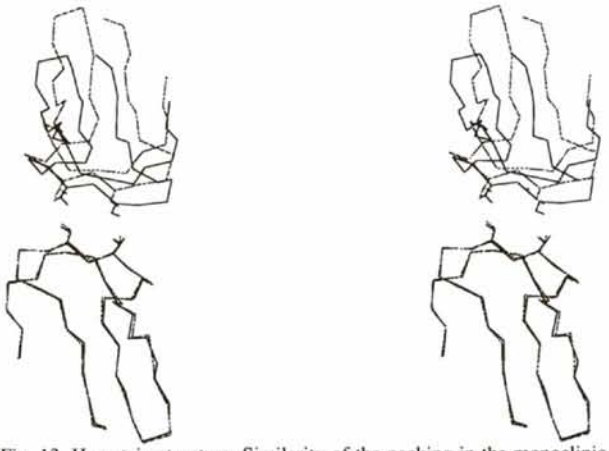

Fig. 13. H-protein structure. Similarity of the packing in the monoclinic and orthorhombic crystal structures of H-$Pro_{Met}$. The stereoview shows the optimized superposition of the orthorhombic asymmetric unit onto the monoclinic asymmetric unit. The view is along the twofold axis parallel to *c* in the orthorhombic structure and along the pseudo twofold axis parallel to the *c* axis in the monoclinic structure. Residues 38–80 are shown.

The similarities of the structures are shown in Fig. 14. In the antiparallel $\beta$-strands, which form the core of the proteins, the positions of the hydrophobic residues (represented in yellow on Fig. 14) are similar. The diagrams of Fig. 14 show also that three glycine residues (residues 36, 47 and 74 in the H-pea protein) are similarly located in the main chains. In the proteins of the PTS system where the catalytic hairpin loop is missing, the comparison of the topologies involves two different fragments: one fragment in the C-terminal side and another one in the N-terminal part which corresponds respectively to the N-terminal part and to the C-terminal part of the lipoate- or biotin-containing proteins. The superpositions of the structures of the lipoate domain from PDH and of the phosphocarrier protein $III^{glc}$ with the H-protein structure are shown in Fig. 15. Although

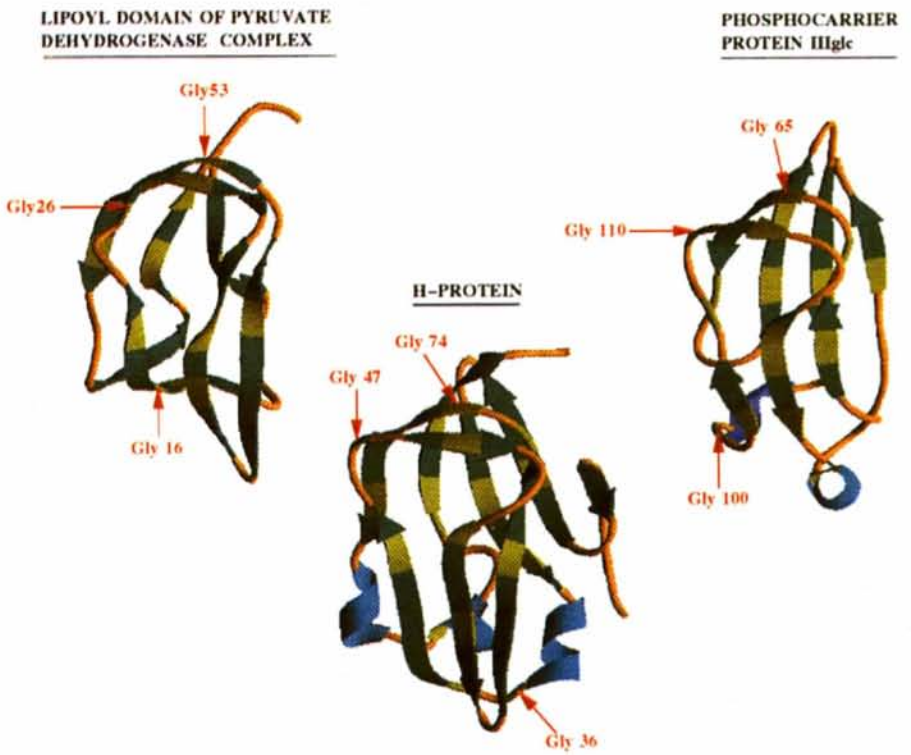


Fig. 14. Similarities between the $\beta$-fold hybrid sandwich observed in H protein and other protein known structures. Diagrams drawn with the program *Ribbons* (Carson, 1987). Conserved glycine are marked. Similarities in the hydrophobic residue positions in the core (represented in yellow colour) can be observed. (*a*) Structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase complex determined by NMR studies (Dardel *et al.*, 1993). (*b*) H-protein, present work; the $\beta$-sandwich domain is only represented and the C-terminal part beyond the residue 104 is not shown. (*c*) Proteins from the phosphotransferase system (PTS); the structure of the fragment (residues 58–126) of the *Escherichia coli* phosphocarrier protein $EIII^{glc}$ represented here (Worthylake *et al.*, 1991) is very similar to that of the IIa domain of the glucose permease of *Bacillus subtilis* not shown (Liao *et al.*, 1991). In the PTS system proteins the catalytic hairpin loop is missing (no lipoic acid of biotin cofactor).

the protein are different, the overall conformation of the sandwich is maintained.

On the other hand, a comparison between the lipoate domain from PDH and the biotinyl domain of yeast pyruvate carboxylase predicted three-dimensional structure analogies between these lipoyl and biotinyl domains (Brocklehurst & Perham, 1993).

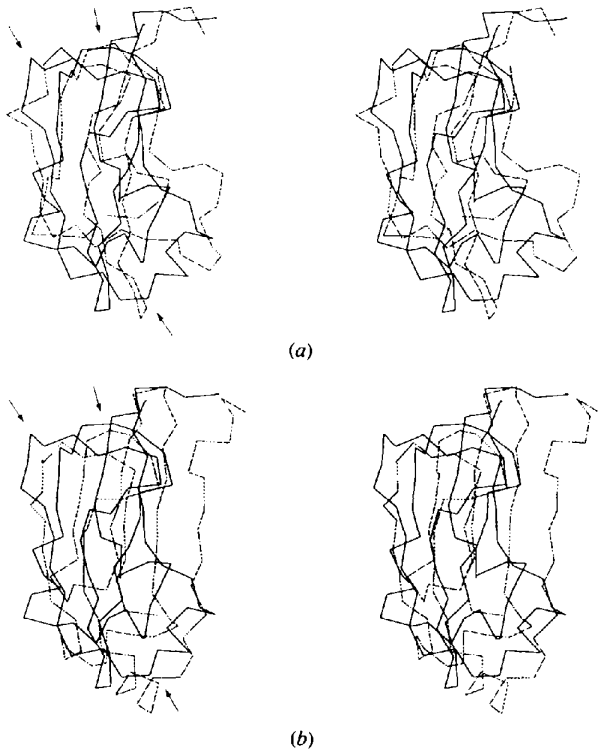


Fig. 15. Superposition of the three structures shown in Fig. 14. (*a*) The structure of the H-protein (residues 24–104, full line) is superposed to the structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase complex (residues 3–76, dotted line). (*b*) The structure of the H-protein is superposed to the structure of the *Escherichia coli* phosphocarrier protein EIII$^{glc}$ (residues 59–123, dotted line). Arrows indicate the positions of the three conserved glycines.

An alignment of the primary structures based on these analogies and using hydrophobic residues of the core and the three glycines was performed for the four proteins with known three-dimensional structures and applied to 52 other known sequences (Fig. 16). The percentage of homology show that the hydrophobic character of the residues used in the alignment is well conserved.

The three glycines are also very well conserved. The $\varphi$ and $\psi$ angles of the glycine residue have large conformational variability allowing two of them to be in equivalent positions (the glycines equivalent to Gly47 and Gly74 of H-protein). Such an observation suggests that these glycine residues are important for the overall protein conformation. In support of this suggestion, Fujiwara, Okamura-Ikeda & Motokawa, (1991) indicated that the replacement of Gly74 by an asparagine in bovine H-protein hinders the binding of lipoate to the H-protein. According to these authors such a mutation might induce a large conformational change of the protein preventing the recognition of the H-protein by the ligase catalyzing the lipoylation. Recently, Leon-Del-Rio & Gravels (1994) reported that, in human propionyl-CoA carboxylase, a biotin-containing protein, the mutations of the glycines equivalent to Gly47 and Gly74 of H-protein are critical for the enzyme biotinylation, and that is probably related to an important change in the protein conformation. The results in this report show that the folding of the proteins, crucial for the enzymatic activity, is sensitive to the glycine positions.

From the sequence and structure comparisons we can assume that the topology in lipoate- and biotin-containing proteins follows the $\beta$-barrel–sandwich hybrid fold described by Chothia & Murzin (1993). Site-directed mutagenesis studies support the hypothesis that this folding is necessary to the activity of the proteins.

In the case of the PTS system proteins where the same $\beta$-barrel–sandwich hybrid topology has been observed

First fragment

```
                <----------------------------------------------------------------------------->
H-pea           24  T I G I T D H A Q D H L - - G E V V F V E - L P E P G V S V T K G K G F G A V E S V K A T S D  67
B.stea. E2p      3  E F K L P D I G E G I H E - G E - I V K W - F V K P G D E V N E D D V L C E V Q N D K A V V E  46
EIII-glc        86  E L F V H F G I D T V E L K G E G F K R I - A E E - G Q R V K V G D T V I E F D L P           125
EIIABC-glc      79  E I L I H F G I D T V S L K G E G F T S F - V S E - G D R V E P G Q K L L E V D L D           118
Homology (%)            100                 90    100       95      95    100         95    100
```

Second fragment

```
                <-------------------------------------------------------------------------->
H-pea           68  V N S P I S G E V I E V - - N T G L T G K P G L I N S S P Y E D G W M I K I K               104
B.stea. E2P     47  I P S P V K G K V L E I L V P E G T V A V T G - - - - - - - - - Q T L I T L D               76
EIII-glc        59  M V A P V D G T I G K I F E T N H - - - - - - - - - - - - - - - - A F S I E               80
EIIABC-glc      52  V V S P V R G K I L N V F P T K H - - - - - - - - - - - - - - - - A I G L Q               73
Homology (%)        100         100 100   98                                        98  100
```

Fig. 16. Alignment of amino-acid sequences of lipoate- and biotin-containing proteins and of phosphotransferase system proteins. The alignment is shown for the four proteins with three-dimensional known structures. The alignment was also performed for 52 know sequences (not shown) of lipoate- and biotin-containing proteins and proteins of the PTS system. H-pea, H-protein of the glycine decarboxylase complex from pea leaves. (PDB code 1HPC, Swiss-prot entry: GCSH_PEA). B. stea. E2p, lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase E2p (PDB code 1LAB, Swiss-prot entry, ODP2_BACST). EIII-glc, *Escherichia coli* phosphocarrier protein III$^{glc}$ (PDB code 1F3G, Swiss-prot entry PTGA_ECOLI). EIIABC-glc, IIa domain of the glucose permease of *Bacillus subtilis* (PDB code, 1GPR, Swiss-prot entry PGTA_BACSU). The percentage of homology is given for the hydrophobic character of the residues and the glycine used in the alignment of the 52 sequences. The hydrophobic character of residue 25 (H-protein) is not conserved in the biotin-containing proteins.

for two of them, the sequence analogy shows that this structure could be expected in all these proteins.

In conclusion these initial results from the investigation of the H-protein of the glycine decarboxylase complex clearly indicate that the lipoamide arm is not free to move in aqueous solvent when loaded with methylamine group. The question remains whether the arm is released from the cleft when the methylamine moiety is released from the lipoamide group, *i.e.* when the T-protein interacts with the loaded H-protein. Structural studies are undertaken to determine the molecular structure of the other components of the complex and of their associations with the H-protein in order to improve our knowledge of the mechanism of the enzymatic process. The crystallization of the reduced form of the H-protein and of the association of the T-protein with the H-protein are in progress.*

We thank Dr O. Dideberg and Dr V. Villeret for helpful discussions. This work was supported by a grant from Centre National de la Recherche Scientifique (Interface Chimie/Biologie). This is publication No. 266 of the Institut de Biologie Structurale Jean-Pierre Ebel.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory [References: 1HPC (H-Pro$_{Ox}$) and 1HTP (H-Pro$_{Met}$)]. Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0428).

## References


BOURGUIGNON, J., NEUBURGER, M. & DOUCE, R. (1988). *Biochem. J.* **255**, 169–178

BROCKLEHURST, S. M. & PERHAM, R. N. (1993). *Protein Sci.* **2**, 626–639.

BRÜNGER, A. (1993). *X-PLOR Manual. Version* 3.1. Yale Univ., New Haven, CT, USA.

CARSON, M. (1987). *J. Mol. Graphics*, **5**, 103–106.

CHOTHIA, C. & MURZIN, A. G. (1993). *Curr. Biol. Struct.* **1**, 217–222

COHEN-ADDAD, C., PARES, S., SIEKER, L., NEUBURGER, M. & DOUCE, R. (1995). *Nature Struct. Biol.* **1**, 63–68.

CONNOLLY, M. L. (1983). *J. Appl. Cryst.* **16**, 548–558.

DARDEL, F., DAVIS, A. L., LAUE, E. D. & PERHAM, R. N. (1993). *J. Mol. Biol.* **229**, 1037–1048.

DOUCE, R. & NEUBURGER, M. (1989). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 371–414.

FUJIWARA, K., OKAMURA-IKEDA, K. & MOTOKAWA, Y. (1991) *FEBS Lett.* **293**, 115–118

HUTCHINSON, E. G. & THORNTON, J. M. (1990). *Proteins Struct. Funct. Genet.* **8**, 203–212.

JONES, T. A. & KJELDGAARD, M. (1993). *O – The Manual, Version* 5.9. Uppsala, Sweden.

JONES, T. A., ZOU, J.-Y., COWAN, S. W. & KJELDGAARD, M. (1991). *Acta Cryst.* A**47**, 110–119.

KABSCH, W. & SANDER, C. (1983). *Biopolymers*, **22**, 2577–2637.

LEE, B. K. & RICHARDS, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.

LEON-DEL-RIO, A. & GRAVELS, R. A. (1994). *J. Biol. Chem.* **269**, 22964–22968.

LIAO, D. I., KAPADIA, G., REDDY, P., SAIER, M. H. JR, REIZER, J. & HERZBERG, O. (1991). *Biochemistry*, **30**, 9583–9594.

LUZZATI, V. (1952). *Acta Cryst.* **5**, 802–810.

MATTEVI, A., DE KOK, A. & PERHAM, R. N. (1992). *Curr. Op. Struct. Biol.* **2**, 877–887.

MATTEVI, A., OBMOLOVA, G., KALK, K. H., TEPLYAKOV, A. & HOL, W. G. J (1993). *Biochemistry*, **32**, 3887–3901.

MESSERSCHMIDT, A. & PFLUGRATH, J. W. (1987). *J. Appl. Cryst.* **20**, 306–315.

MORRIS, A. L., MACARTHUR, M. W., HUTCHINSON, G. E. & THORNTON, J. M. (1992). *Proteins Struct. Funct. Genet.* **12**, 345–364

MURZIN, A. G. (1994). *Curr. Op. Struct. Biol.* **4**, 441–449.

NAVAZA, J. (1994). *Acta Cryst.* A**50**, 157–163.

NEUBURGER, M., JOURDAIN, A. & DOUCE, R. (1991). *Biochem. J.* **278**, 765–769.

PARES, S., COHEN-ADDAD, C., SIEKER, L., NEUBURGER, M. & DOUCE, R. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 4850–4853.

RAMACHANDRAN, G. N. & SASISEKHARAN, V. (1968). *Adv. Protein Chem.* **23**, 283–437.

REED, L. J. (1974). *Acc. Chem. Res.* **7**, 40–46.

ROUSSEL, A. & CAMBILLAU, C. (1989). *Silicon Graphics Geometry Partner Directory*, pp. 77–78. Mountain View CA: Silicon Graphics.

SIBANDA, B. L., BLUNDELL, T. L. & THORNTON, J. M. (1989). *J. Mol. Biol.* **206**, 759–777.

SIEKER, L., COHEN-ADDAD, C., NEUBURGER, M. & DOUCE, R. (1991). *J. Mol. Biol.* **220**, 223–224.

WORTHYLAKE, D., MEADOW, N. D., ROSEMAN, S., LIAO, D. I., HERZBERG, O. & REMINGTON, S. J. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 10382–10386.