Max-Planck Institut für Biochemie
Abteilung Strukturforschung

Biosynthesis and function of biological pteridines

Structural studies on two molybdopterin containing Aldehyde oxido-reductases, from *Desulfovibrio desulfuricans* ATCC 27774 and from *Desulfovibrio gigas*, and the GTP cyclohydrolase I on *E. coli*, responsible for the first step of the tetrahydropterin biosynthesis.

Jorge Manuel Baeta Simões Rebelo

Vollständiger Abdruck der von der Fakultät für Chemie der
Technische Universität München zur Erlangung
des akademischen Grades eines
Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Dr. Dr. Adelbert Bacher
Prüfer der Dissertation:	1. apl.-Prof. Dr. Dr. h. c. Robert Huber
	2. Univ.-Prof. Dr. Johannes Buchner

Die Dissertation wurde am 7.6.2004 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 28.7.2004 angenommen.

This work, although a valid scientific work, is nevertheless far from deserving an important scientific prize, but is a very important step on my life! Unfortunately this long step is under my personal aspirations, but is simply what was possible and nothing more. I will take the responsibility and the consequences for this damage.

I will mainly address thanks to my supervisors, Professors Maria João Romão and Robert Huber and also to Professor Adelbert Bacher, those who made this work possible. I also thank some coleagues and some good friends that supported me. I am grateful to everyone that at some stage helped a little bit to build it and even to those who not doing anything at all, did not disturbed it.

Besides the practical knowledge that I could take from it as a "young" scientist, the most important to me during these long years is what I learned about human society and human nature.

I will dedicate this humble work, to two persons who always supported me unconditionally, and those are the persons I love most in this world, my parents Maria Helena and Manuel.

This work was financially and exclusively supported by grants PRAXIS/BM/12730/97 and PRAXISXXI/BD/21493/99 from the Fundação para a Ciência e Tecnologia / Ministério da Ciência e Tecnologia (Portugal) and Fundo Social Europeu (III Quadro comunitário de apoio).

Summary

Pteridine derivatives take part in different biochemical processes inside organisms. Two general classes of enzymes are directly related with these compounds, enzymes that participate in their biosynthesis, and enzymes that use these substances as active cofactors for their enzymatic activity.

To the first class belongs the GTP cyclohydrolase I that catalyses the hydrolytic release of formate from GTP followed by cyclization to dihydroneopterin triphosphate. Found as homodecamers in bacteria and animals, they contain one catalytic zinc ion per subunit. Replacement of Cys 110, Cys 181, His 112 or His 113 of the enzyme from *Escherichia coli* by serine affords catalytically inactive mutant proteins with reduced capacity to bind zinc. These mutant proteins are unable to convert GTP or the committed reaction intermediate, 2-amino-5-formylamino-6-(β-ribosylamino)-4(3*H*)-pyrimidinone 5'-triphosphate, to dihydroneopterin triphosphate. On the present work the crystal structures of GTP complexes of the His113Ser, His112Ser and Cys181Ser mutant proteins determined at resolutions of 2.5, 2.8 and 3.2 Å, respectively, revealed the conformation of substrate GTP in the active site cavity. The carboxylic group of the highly conserved residue Glu 152 anchors the substrate GTP, by hydrogen bonding to N-3 and to the position 2 amino-group. Several basic amino acid residues interact with the triphosphate moiety of the substrate. Mutant and wild type structural comparison indicates a direct coordination of the catalytic zinc with the Cys 110, Cys 181, His 113 amino-acid residues and indirect coordination over a complexed water molecule to His 112. In close analogy to zinc proteases, the zinc-coordinated water molecule is suggested to attack C-8 of the substrate affording a zinc-bound 8*R* hydrate of GTP. Opening of the hydrated imidazole ring affords a formamide derivative, which remains coordinated to zinc. The subsequent hydrolysis of the formamide motif has an absolute requirement for zinc ion catalysis. The hydrolysis of the formamide bond shows close mechanistic similarity with peptide hydrolysis by zinc proteases.

To the second class belong the molybdopterin containing enzymes of the xanthine oxidase (molybdenum hydroxylase) family of enzymes. In the present work, two sulphate reducing bacterium, aldehyde oxidoreductases from *Desulfovibrio gigas* (MOP) and *Desulfovibrio (D.) desulfuricans* ATCC 27774 (MOD) where studied.

MOPs single peptide chain with 907 amino-acid residues model, also containing one molybdopterin cytosine dinucleotide (MCD) cofactor and two [2Fe-2S] iron-sulphur clusters was studied using almost atomic resolution data. Improved cryo-cooled crystals were measured using a synchrotron x-ray source and yield about 2% reduction on room temperature cell constant sizes. A total of 233755 unique reflections under the space group $P6_122$ where measured at 1.28 Å resolution. Throughout the entire refinement the full gradient least squares method was used, leading to a final R factor of 14.5 and R_{free} factor of 19.3. The model contains 8146 non-hydrogen atoms described by anisotropic displacement parameters with an observations/parameters ratio of 4.4. Three Cl^- and two Mg^{2+} ions from the crystallization solution were clearly identified, that with the exception of one Cl^- which is buried and 8 Å distant from the Mo atom, are close to the molecular surface and may contribute to crystal packing. The overall structure has not changed in comparison to the lower resolution model, apart from local corrections that included some loop adjustments and 17 alternate side-chain or main-chain conformations. Estimated errors of bond distances obtained by blocked least squares matrix inversion, enabled a more detailed analysis of the three redox-centres. For the MCD cofactor the resulting geometric parameters confirmed its reduction state as a tetrahydropterin. At the Mo centre estimated corrections calculated for the Fourier ripples artifact are very small when compared to the experimental associated errors, supporting the suggestion that the fifth ligand is a water molecule rather than a hydroxide. Concerning the two iron-sulphur centres, asymmetry in the Fe-S distances as well as differences in the pattern of NH-S hydrogen bonding interactions was observed, which influence the electron distribution upon reduction and cause non equivalence of the individual Fe atoms in each cluster.

The aldehyde oxidoreductase from *Desulfovibrio (D.) desulfuricans* ATCC 27774 (MOD) has similar substrate specificity as the homologous enzyme from *D. gigas* (MOP) and the primary sequences from both enzymes show 68% identity. The crystallographic structure was solved at 2.8 Å resolution on the space group $P6_122$ with unit cell dimensions of a=b=156.4 Å and c=177.1 Å, by Patterson Search Techniques using the coordinates of the *D. gigas* enzyme. The overall fold of the *D. desulfuricans* enzyme is very similar to MOP and the few differences are mapped to exposed regions of the molecule. This is reflected in the electrostatic potential surfaces of both homologous enzymes, one exception being the surface potential in a region identifiable as the putative docking site of the physiological electron acceptor. Two important mutations are located in the pocket bearing a chain of catalytically relevant water molecules. Other essential features of the

MOP structure such as residues of the active site cavity are basically conserved in MOD. The comparison made allowed confirming and establishing features which are essential for their function, namely conserved residues in the active site, catalytically relevant water molecules and recognition of the physiological electron acceptor docking site.

Parts of this work where published on the following scientific articles:

Rebelo J., Macieira S., Dias J.M., Huber R., Ascenso C., Rusnak F., Moura J.J.G., Moura I. and Romão M.J. (2000) Gene sequence and crystal structure of the Aldehyde oxido-reductase from *Desulfovibrio desulfuricans* ATCC 27774. *J. Mol. Biol.* **297**, 135-146

J.M.Rebelo, J.M. Dias, R. Huber, J.J.G.Moura and M.J.Romão (2001) Structure refinement and analysis of the Aldehyde Oxidoreductase from Desulfovibrio gigas (MOP) at 1.28Å. *J. Biol. Inorg. Chem.* **6**:791-800.

Rebelo J, Auerbach G., Bader G., Bracher A., Nar H., Hösl C., Schramek N., Kaiser J., Bacher A., Huber R. and Fischer M. (2003). Biosynthesis of pteridines. Reaction mechanism of GTP cyclohydrolase I. *J. Mol. Biol.* 326 (2):503-16.

Zusammenfassung

Pteridin Derivate sind Bestandteil grundlegend unterschiedlicher biochemischer Prozesse. Zwei allgemeine Enzymenkategorien stehen direkt mit diesen Derivaten in Verbindung: Enzyme, die direkt an ihrer Biosynthese beteiligt sind, und Enzyme, die diese Substanzen als Kofaktoren bei ihrer enzymatischen Tätigkeit benutzen. Zur ersten Kategorie gehört die GTP Cyclohydrolase I, die die hydrolytische Freisetzung von Formiat aus GTP katalysiert, gefolgt von der Zyklisierung zum Dihydroneopterin Triphosphat. Das Enzym ist bei Bakterien und Eukarioten als Homodekamer zu finden. Diese Homodekamere enthalten ein katalytisches Zink pro Untereinheit. Austausch von Cys110, Cys181, His112 oder His113 durch zielgerichtete Mutagenenese des Enzyms von *Escherischia coli* gegen Serin, erzeugt katalytisch inaktive Proteine mit der verringerter Zinkbeladung. Diese durch Mutation erzeugten Proteinvarianten sind nicht imstande GTP oder das Reaktions-intermediat 2-amino-5-formylamino-6-(β-ribosylamino)-4(3*H*)-pyrimidinon 5'-triphosphat zum Dihydroneopterin Triphosphat umzuwandeln. In der vorliegenden Arbeit wurden die Kristallstrukturen der GTP-Protein Komplexe der Mutanten His113Ser, His112Ser und Cys181Ser mit Auflösungen von 2.5, 2.8 und 3.2 Å aufgeklärt. Dadurch könnte die Lage des GTP im Aktiven Zentrum bestimmt werden. Die Karboxygruppe der in hohem Grade konservierten Aminosäure Glu 152 bindet das Substrat GTP durch Wasserstoffbrückebindungen zu N-3 und der Aminogruppe Position 2 des GTP. Der Phosphatrest des GTP wird mit Ausnahme von Arg139, von konservierten Aminosäuren koordiniert. Durch strukturellen Vergleich der Mutanten mit dem Wildtyp-Enzym lässt sich eine Koordination des katalytischen Zinks durch die Aminosäuren Cys110, Cys181 und His113, und eine indirekte Koordination über ein komplexiertes Wassermolekül zur Aminosäure His112 nachweisen. In der nahen Analogie zu Zinkproteasen wird in dieser Arbeit vorgeschlagen, dass das zinkkoordinierte Wassermolekül ein nukleophilen Angriff auf das C-8-Atom des Substrates durchführt. Dabei entsteht das mit dem Zink verbundene Hydrat 8R von GTP. Bei der Öffnung des hydratisierten Imidazolringes wird das Formamidintermediat 2-amino-5-formylamino-6-(ribosyl-amino)-4(3H)-pyrimidone 5'-triphosphate gebildet, welches an Zink koordiniert bleibt. Die darauf folgende Hydrolyse des Formamidintermediats verlangt zwingend eine katalytische Beteiligung des Zinks. Auffallig dabei ist die mechanistische Ähnlichkeit mit der Peptidhydrolyse durch Zinkproteasen.

In der vorliegenden Arbeit wurden die Aldehyd-oxidoreductasen von *Desulfovibrio gigas* (MOP) und *Desulfovibrio (D.) desulfuricans* ATCC 27774 (MOD) studiert. Das monomer von MOP hat eine Länge von 907 Aminosäuren, ein Molybdopterin Cytosin-Dinucleotid (MCD) als Kofaktor und zwei Eisen-Schwefel [2Fe2S] Cluster. Die Kristallstruktur dieses Proteins konnte mit nahezu atomarer Auflösung (1.28 Å) studiert werden. Kristalle hoher Qualität, Niedrigtemperaturmessungen und die Verwendung eines Synchrotrons als Röntgenstrahlungsquelle ergaben eine 2 prozentige Verkleinerung der Zellenkonstanten im Vergleich mit vorherigen Strukturen. Bei 1.28 Å Auflösung wurden 233755 einzigartige Reflexionen gemessen. Während des refinements wurde die Methode volle Steigung kleinstes Quadrat verwendet. Das führte zu einem abschließenden R Faktor von 14.5 % und R_{free} Faktor von 19.3 %. Das Modell enthält 8146 Nichtwasserstoffatome, die durch anisotrope Distanzparameter mit einem Beobachtung/berechnungs Parameterverhältnis von 4.4 beschrieben wurden. Die Kristallisationslösung enthält $MgCl_2$. Die Kristallstruktur zeigt Zwei Mg und drei Cl Ionen, eines der Cl Ionen befindet sich in der Nähe des aktiven Zentrums im Abstand von 8Å von dem Mo Atom, die anderen Ionen befinden sich nah an der molekularen Oberfläche und können zur Kristallverpackung beitragen. Die gesamte Struktur hat sich im Vergleich zum Modell niedrigerer Auflösung, abgesehen von lokalen Korrekturen, anpassungen von loops und 17 geringfügigen Konformationsänderungen nicht geändert. Die Fehlerberechnung der Bindungsabstände wurde mit der Methode der blockierten kleinsten Quadratmatrixumstellung durchgeführt. Das ermöglichte eine ausführliche Analyse der drei Redoxzentren. Für den MCD Kofaktor bestätigten die resultierenden geometrischen Parameter seinen Zustand als Tetrahydropterin. Das Fourier Kräuselungkunstprodukt ware bei den Kristallstrukturen im Vergleich zu experimentell bedingten störungen vernachlässigbar klein. Dies erlaubt und unterstüzt den Vorschlag von Wasser statt Hydroxid als fünften Liganden des Mo Ions. Asymmetrien in den zwei Eisen-Schwefel Zentren hinsichtlich der Fe-S Abständen sowie Unterschiede bezüglich des Musters der NH-S Bindungsinteraktionen, könnten die Elektronenverteilung nach Oxidation beeinflussen.

Die Aldehydeoxydoreduktase von *Desulfovibrio desulfuricans* ATCC 27774 hat ähnliche Substratanforderungen wie das übereinstimmende Enzym von *Desulfovibrio gigas*. Die Aminosäurensequenzen der beiden Enzyme zeigen 68 % Identität. Die Kristallstruktur wurde bei 2.8 Å Aufgelost. Die Raumgruppe $P6_122$ hat Zelldimensionen von a=b=156.4 Å und c=177.1 Å. Die Kristallstruktur wurde nach der Methode von Patterson mit den

Koordinaten des *Desulfovibrio gigas* Enzyms gelöst. Die gesamte Faltung des *Desulfovibrio desulfuricans* ATCC 27774 ist sehr ähnlich. Die wenigen unterschieden manifestieren sich im der Veränderung der elektrostatischen Oberflächen von beiden Enzymen. Dies erlaubt es, eine bestimmte Region an der Oberfläche der Proteine als mutmaßliche Bindestelle des physiologischen Elektronenakzeptors zu identifizieren. Außerdem befinden sich in der Nähe des Katalytischen Zentrums, in einer Bindungstasche für eine Kette katalytisch relevanter Wassermoleküle, zwei wichtige Mutationen im Vergleich zur MOP-Struktur

Andere wesentliche Eigenschaften der MOP Struktur, wie Aminosäuren des aktiven Zentruns, werden im Allgemeinen in MOD konserviert. Der Vergleich beider Enzyme erlaubte die Bestätigung der für ihre Funktion wesentlichen Eigenschaften.

Veröffentlichungen, die Teil dieser Arbeite beinhalten:

Rebelo J., Macieira S., Dias J.M., Huber R., Ascenso C., Rusnak F., Moura J.J.G., Moura I. and Romão M.J. (2000) Gene sequence and crystal structure of the Aldehyde oxido-reductase from *Desulfovibrio desulfuricans* ATCC 27774. *J. Mol. Biol.* **297**, 135-146

J.M.Rebelo, J.M. Dias, R. Huber, J.J.G.Moura and M.J.Romão (2001) Structure refinement and analysis of the Aldehyde Oxidoreductase from Desulfovibrio gigas (MOP) at 1.28Å. *J. Biol. Inorg. Chem.* **6**:791-800.

Rebelo J, Auerbach G., Bader G., Bracher A., Nar H., Hösl C., Schramek N., Kaiser J., Bacher A., Huber R. and Fischer M. (2003). Biosynthesis of pteridines. Reaction mechanism of GTP cyclohydrolase I. *J. Mol. Biol.* 326 (2):503-16.

I Introduction

1 Scope of this work

This work is an overall overview of the importance of pteridines in biological systems. Both types of enzymes studied here have completely different crystallographic structures and are involved in diverse biochemical processes, but are important examples of biosynthesis and functionality of pteridines, that otherwise would be completely unrelated.

The first task that was defined at the start of this work was to verify and confirm the proposed enzymatic mechanism of the pteridine-containing enzyme, the aldehyde oxidoreductase, belonging to the xanthine oxidase family of Molybdenum containing enzymes (XO) (also known as Molybdenum hydroxylase family of enzymes - MH). These studies included a high-resolution description of the cofactors on an oxidised desulfurated form of the enzyme. Furthermore, to localize the possible position of interaction of the unknown electron acceptor responsible for the regeneration of an active state on this enzyme, the crystallographic structure was also determined on a different but very close organism, and then a comparison with the crystallographic high-resolution refined model of the first structure was carried out.

A second task was assigned to the present work, namely, the establishment, and possible final clarification of the detailed mechanism on the biosynthesis of 7,8-dihydroneopterin 3´-triphosphate, as catalysed by the GTP cyclohydrolase I in animals and micro-organisms excluding plants. This included the study of the crystallographic structure of several completely inactive single amino acid mutants, and the description of the structural state of the active site, including a description of the substrate GTP molecule positioned to start the reaction.

2 Protein Crystallography

2.1 Crystallization and X-ray diffraction

Concentrated protein solutions purified to a high degree of homogeneity and bio-chemical activity, are required for the use of relatively well-established crystallization techniques. By far, the most used crystallization technique is the vapour diffusion, operated on small volume containers, where a small drop of concentrated protein solution is mixed with a crystallization solution. The protein/crystallization solution, as a homogeneous mixture obtained on the drop, is left to equilibrate slowly, by vapour diffusion, with a separated larger volume of this crystallization solution (for an example see Materials and Methods later in the text). After a variable number of crystallization solutions have been tested (screening), it is expected that at least in one of these, small protein crystals have been formed. Sets of conditions that have shown to produce crystals on several different proteins are many times used as starting point on the search of successful crystallization conditions. When this systematic approach fails, other more scientific screening strategies were made to test a number of variables, being statistically significant they provide useful information and important clues to explore. These screening tests perform variations not only over crystallization solutions, but also over a spectrum of physical-chemical relevant factors like pH, concentrations of precipitating agents or working volumes. Several substances either organic like polyetilenglycol (PEG), or inorganic like ammonium sulphate emerged as important crystallization agents. Still uncertainty is very high during this phase of structure determination, nevertheless is thought that almost hundred percent of the soluble “good behaved” proteins are possible to crystallize.

Once a safe indication of the presence of protein crystals on one of the tested hypothesis is ensured, a step directed to the improvement of the quality or/and of the size of these first grown crystals is necessary in most of the cases. This development, although a guided step, is many times more effort demanding then obtaining the starting crystallization conditions themselves.

Single crystals with a minimum size between 100 and 200 μm, and good dimensioned on the three spatial directions are normally enough for successful X-ray diffraction tests and eventual characterization. This first diffraction characterization permits normally the determination of the crystal internal symmetry (space group), the measurement of lattice cell dimensions and inter-axial angles, and an overall assumption of the crystal quality by determining the diffraction resolution obtained as well as other data quality statistics.

At this stage there is still no information available on the protein atomic structure. Although the diffraction observations contain a significant amount of information something more is needed. The observed measurements give information about diffraction intensities, but according to the quantum mechanics, more specifically, with the dynamic theory of diffraction, we are in presence of waves, and waves have an amplitude and a phase. The observed intensities are the squares of the amplitudes of the sum of all the time independent waves produced on the interaction of the x-ray main beam with the atomic electrons from the crystal due to coherent scattering (Thomson scattering), obviously not excluding other less important phenomena present like Compton scattering or absorption, among others.

Each spot experimentally observed is called structure-factor $F(\underline{\mathbf{r}}^*)$ and corresponds to constructive interference between the scattered waves. $\underline{\mathbf{r}}^*$ (1) is a space vector on a reciprocal space (an abstracted physically inexistent space) that represents the difference of directions between a determined direction $\underline{\mathbf{s}}$ and the direction from the direct beam $\underline{\mathbf{s}_0}$, weighted by the reciprocal of the wavelength λ of the x-ray radiation.

$$\underline{\mathbf{r}}^* = \lambda^{-1} (\underline{\mathbf{s}} - \underline{\mathbf{s}_0}) \qquad (1)$$

On a continuum of electronic distribution inside the crystal $\rho(\underline{\mathbf{r}})$, the mathematic expression for $F(\underline{\mathbf{r}}^*)$ is given by equations (2) and (3).

$$\mathbf{F}(\underline{\mathbf{r}}^*) = \int_v \rho(\underline{\mathbf{r}}) \exp(2\pi i \underline{\mathbf{r}}^* . \underline{\mathbf{r}}) \, d\underline{\mathbf{r}} \qquad (2)$$

$$= \int_v \rho(\underline{\mathbf{r}}) \exp(\alpha_{\underline{\mathbf{r}}} i) \, d\mathbf{v} \qquad (3)$$

$$= T[\rho(\underline{\mathbf{r}})] \qquad (4)$$

$$= |\mathbf{F}(\underline{\mathbf{r}}^*)| \exp(\alpha_{\mathbf{F}} i) \qquad (5)$$

That is, the contribution of all differential volume scattered amplitudes summed over the complete volume of the crystal (2) and (3), or, the Fourier transform operator (4) over the electronic 3-dimensional density function over the volume of the crystal. **F** is a complex number (5) that also represents a wave, where the module $|\mathbf{F}(\underline{\mathbf{r}}^*)|$ corresponds to the square root of an observed intensity (diffraction spot) and $\alpha_{\mathbf{F}}$ a phase that cannot be observed.

Obviously, as the crystal is rotated over an axis the directions from the vectors **r** are altered and the observed structure factors will also change, corresponding to new conditions of constructive interference.

Introducing a discrimination of point scattering centres (atoms) the expression for **F** becomes simpler (6), and relatively easy to calculate using the fast-Fourier-transform (FFT).

$$\mathbf{F}(h,k,l) = V^{-1} \Sigma_x \Sigma_y \Sigma_z \, A_{x,y,z} \exp \{ (2\pi i) (xh + yk + zl) \} \qquad (6)$$

x,y,z - positions of the scattering centres

h,k,l - Miller indexes of the scattering planes

In the sampling process, if on equation (1), **r*** is replaced by the h,k,l indexes, and projections of the ($\mathbf{s} - \mathbf{s_0}$) vectors over the unit cell directions (periodicity of the dispersion centres over the crystal) are carried out, we obtain 3 equations called von Laue conditions (7).

a . $(\mathbf{s} - \mathbf{s_0}) = h \lambda$ (7a)

b . $(\mathbf{s} - \mathbf{s_0}) = k \lambda$ (7b)

c . $(\mathbf{s} - \mathbf{s_0}) = l \lambda$ (7c)

2.2 Solving the phase problem

The determination of starting phases that enable the construction of a model is achieved mainly by three different ways, multiple isomorphous replacement (MIR), multiple anomalous dispersion (MAD) and molecular replacement (MR). Each of these techniques is used on a different context, but mixed solutions are also sometimes necessary. A full description of the concepts and the theoretical ground of both MIR and MAD techniques can be found on "*Fundamentals of crystallography*" (1992) by Giacovazzo *et al.*

2.2.1 Molecular Replacement (MR)

Identity and homology between amino-acid sequences of proteins correspond on many cases to secondary or even tertiary structural similarities. This opens the possibility of creating relatively simple initial models for proteins, using already known folding motives belonging to other sequence related proteins whose crystallographic structure is already known. The number of unique folding motives is relatively small as can be seen by a search on the SCOP database (Murzin *et al.*, 1995). This is the ground for the molecular replacement method. This method is particularly efficient for solving structures of proteins common to different organisms that in most cases have extensive amino-acid sequence homology.

After an initial model is selected, a search is made on the three-dimensional space and is normally carried out in two steps, namely a rotation and a translation. The first step, the rotational step, searches for the correct orientation of the starting model, to which corresponds a transformation matrix **[M]**. Then it is necessary to find the correct position of the model on the a.u. (asymmetric unit) by carrying a translation search. The final translation is represented by a vector $\underline{\mathbf{t}}$. The two-step search produces a solution on the form of a composed transformation, where the position vector of each atom at the end $\underline{\mathbf{X}}'$ is related with the position vector of the starting point $\underline{\mathbf{X}}$, according to the following expression:

$$\underline{\mathbf{X}}' = [\mathbf{M}]\,\underline{\mathbf{X}} + \underline{\mathbf{t}} \qquad (8)$$

The minimization of the difference between observed and calculated structure factors, respectively F_{obs} and F_{calc}, is the criteria guiding the search process. This is not done directly, but by the use of the Patterson function $P(\underline{\mathbf{u}})$, which is defined as the self-convolution of the electron density $\rho(\underline{\mathbf{r}})$:

$$P(\underline{\mathbf{u}}) = \rho(\underline{\mathbf{r}}) * \rho(-\underline{\mathbf{r}}) = \int_v \rho(\underline{\mathbf{r}})\, \rho(\underline{\mathbf{u}}+\underline{\mathbf{r}})\, d\underline{\mathbf{r}} \qquad (9)$$

According to the convolution theorem:

$$T[P(\underline{\mathbf{u}})] = T[\rho(\underline{\mathbf{u}})] \,.\, T[\rho(-\underline{\mathbf{u}})] \qquad (10)$$

Using (4) and the rules of complex analysis:

$$\begin{aligned} &= [|\mathbf{F}(\underline{\mathbf{r}}^*)| \exp(i\theta)].[|\mathbf{F}(\underline{\mathbf{r}}^*)| \exp(-i\theta)] \\ &= |\mathbf{F}(\underline{\mathbf{r}}^*)|^2 \end{aligned} \qquad (11)$$

Then:

$$P(\underline{\mathbf{u}}) = T^{-1}[|\mathbf{F}(\underline{\mathbf{r}}^*)|^2]$$

Or, relating to a finite set of structure factors:

$$P(\underline{\mathbf{u}}) = V^{-1} \sum_h |\, F_h \,|^2 \, \mathrm{Cos}(2\pi\, \underline{\mathbf{h}} \,.\, \underline{\mathbf{u}}) \qquad (12)$$

The Patterson function presents two types of maximum, those arising from intra-molecular vectors, or self-vectors and those from inter-molecular vectors, or cross vectors. The rotation of the starting model, is to be analysed using a rotation function defined by:

$$R(\mathbf{[M]}) = \int_v P_{crystal}\, (\underline{\mathbf{u}})\, P_{mol}\, (\mathbf{[M]}\underline{\mathbf{u}})\, d\underline{\mathbf{u}} \qquad (13)$$

This represents the superimposition of Patterson function of the rotated molecule and the Patterson function of the crystal. Other rotation functions can be defined based on different criteria like the sum of two Patterson, for some examples see Buerger (1959) *In* "Vector space", or Huber (1969) *In* "Crystallographic Computing Proc.".

The maxima occurring on this function corresponding to self-vectors are the possible solutions for the reorientation of the starting model. At this stage, the number of maxima found should already give some indication on the symmetry inside the crystal.
The maxima of the rotation function, that stand for the possible correct orientations of the molecule are then explored, in order to find the location inside the unit cell. The intra-molecular vectors are irrelevant at this stage, and only inter-molecular or cross-vectors have a role. The determination of the translation vector is performed by comparison of two Patterson maps, but now the target consists on maximization of superposition of different classes of peaks. Several translation functions can be used. An example for two molecules, the set of cross-vectors of the calculated Patterson from molecule 1 to molecule 2 is given by the following function:

$$P_{12}(\mathbf{u}) = \int_v \rho_1(\mathbf{r})\, \rho_2(\mathbf{u}+\mathbf{r})\, d\mathbf{r} \qquad (14)$$

ρ_1 and ρ_2 represent electron densities of the two molecules.
Defining a vector $\mathbf{t} = \mathbf{s}_2 - \mathbf{s}_1$ (translation vector), a possible translation function can be defined as:

$$T(t) = \int_v P_{obs}(\mathbf{u}).\, P_{12}(\mathbf{u},\mathbf{t})\, d\mathbf{u} \qquad (15)$$

Which will have a maximum when the two Patterson functions superimpose. On the reciprocal space the equation takes the form:

$$T(t) = \sum_n I_{obs}(\mathbf{h})\, F_1(\mathbf{h})\, F_1^*(\mathbf{hA}) \exp(-2\pi i \mathbf{ht}) \qquad (16)$$

Being $F_1(\mathbf{h})$ and $F_1^*(\mathbf{hA})$ the Fourier transforms for molecule 1 and for the same molecule after a symmetry operation **A** is performed. An advantage of this function is that the calculations do not have to be extended to the complete unit cell. A simple example of a two-fold screw axis parallel to z-axis will yield a vector (2x,2y,1/2) which will be confined to a plane z=1/2.

2.3 Structure refinement

Structure solution is understood as the initial approach to the solution of the problem of creating a rough model that agrees with the measured experimental data. Still, to be a valid solution the model has to be close enough to the solution so that no new experimental information has to be added to it. To the process of introduction of corrections or additional small model parts that are missing - model building/rebuilding in order to fit electronic density maps, followed by automated refinement on computers using mathematical minimization methods is called refinement. It is carried out in an iterative way, and should converge to a final and accurate solution.

Automated refinement methods are mainly based on the Least Squares minimization method. For that, a least squares function is defined, S_1, that is the weighted sum of the squares of the differences between the observed and the calculated structure factors:

$$S_1 = \Sigma_i \, w_i \, (F_{i(obs)} - F_{i(calc)})^2 \qquad (17)$$

This function is to be minimized in function of each of its model parameters, included on the expression defining $F_{i(calc)}$ that directly reflect the initial model.

The Least Squares method requires a large system over-determination, that is, the system has to have enough information in order to be refined. The ratio of observations to variables has to be high, ideally about 10 as occurs on small molecules crystallography, where the resolution obtained on x-ray diffraction is very high, to 0.7 Å or more. Protein crystals are intrinsically much less ordered, 2.0 Å is already considered a good result. To overcome this problem the number of parameters to be determined is decreased, using constrains, or the number of observations is artificially increased, using restrains, or both can be applied simultaneously.

When the geometry of a group of atoms is accurately known, and there are reasons to believe that the environment will not modify significantly this geometry, the entire group can be treated as a rigid entity. This is called constrained or rigid body refinement, and is always used on the first refinement of a structure solved by the molecular replacement method.

The artificial increase of the number of observations is normally made by the introduction of geometrical conditions. These conditions include properties observed in proteins as they are built of segmented sequences of connected aminoacids, for instance typical bond

distances or bond angles. An accurate set of these conditions, $d_{j(\text{“ideal”})}$, was determined by Engh and Huber (1991), now universally used. In addition to the initial S_1 function, a S_2 function containing the information about these restrains is also minimized:

$$S_2 = \Sigma_j \, w_j \, (d_{j(\text{ideal})} - d_{j(\text{calc})})^2 \qquad (18)$$

Although a set of new equations is introduced, the number of parameters to be determined does not change since $d_{j(\text{calc})}$ is a function of the coordinates. The weight w_j reflects the uncertainty of the restraint and is the reciprocal value of the standard deviation associated with the measurement of the specific geometrical condition. Other conditions reflecting other geometrical parameters or isotropy/anisotropy of thermal parameters, occupancy, or even non-crystallographic symmetry can be used.

A second strategy of automated refinement, particularly efficient at the initial stages of the refinement due to a larger radius of convergence is the refinement by Molecular Dynamics. This usually ensures a faster convergence of the refinement, due to the fact that the molecular model is “shacked” to adopt a better starting position to be followed by a more fine refinement performed by the Least Squares method. The method consists on solving the classical Newton equation of motion taking into account the effect of the medium and the approximations used to calculate the total energy. A set of Langevin equations (accounting for dynamical effects):

$$m_i \, . d^2x_i(t)/dt^2 = - \, \text{grad}_x \, E_{tot} + f_i(t) - m_i b_i \, . \, dx_i(t)/dt \qquad (19)$$

That states that the forces applied on each atom (left side) are originated from the gradient of the total potential energy (first term on the right), a random function dependent on the temperature ($f_i(t)$) and a damping function related with friction (third term on the right). The simulation starts with an initial set of coordinates. To each atom is assigned a velocity at random from a Maxwell distribution corresponding to the temperature selected. The equation is then integrated at a given temperature for a given time. New coordinates arise from the previous step and once again new velocities are assigned at a certain temperature to perform another step. A special case of this type of refinement called simulated annealing, changes the temperature in each step, from room temperature to higher temperature and back, or just starting at a higher temperature and cooling down. These calculations are performed for small periods of time, normally pico-seconds. The higher

temperatures used are very far from the normal biological conditions, however this helps the model to escape local minima, increasing largely the convergence radius.

Refinement of a model cannot proceed indefinitely and stopping criteria are based mainly on two things, the so called crystallographic R factor acquiring for the adjustment between the experimental and the calculated structure factors, and the accordance of the model with the calculated difference electronic density maps. A third indicator called R_{free} factor is as important as the other two because it validates the quality of the refinement. This last indicator consists in exactly the same computation formula as used for the R factor but on an independent set of observations selected at the beginning of the refinement. The R_{free} factor cannot increase during the refinement and its final value should be near the R factor or the refinement was biased and was consequently driving the model to a wrong result.

Further insight over the crystallographic theory, can be found on "*Fundamentals of crystallography"* by Giacovazzo *et al.*, (1992).

3 Biological pteridines

Pteridines are a fundamental group of natural heterocyclic compounds, which are structurally related by a pyrazinopyrimidine moiety and serve as cofactors on several different enzyme-catalysed reactions. Biochemical active pteridine compounds include vitamins, folic acid, riboflavin, and cofactors for redox reactions. Particularly relevant while related with the enzymes studied here, are:

- Tetrahydrofolate (H_4Folate) and tetrahydromethanopterin (H_4MPT) are used as mediators for the transfer of one-carbon fragments. One-carbon chemistry includes a variety of reactions that involve the movement of single carbon atoms from one molecule to another and is especially important in amino acid metabolism and nucleotide biosynthesis. These processes include methylation reactions, which use mainly S-adenosylmethionine as carbon donor, or carboxylase reactions involving the substrate incorporation of carbon dioxide, or even reactions involving derivatives of tetrahydrofolate.

  On the tetrahydrofolate molecule the bound carbon can be in the formic acid oxidation state ($HCOO^-$), formaldehyde oxidation state (CH_2=O) or methanol

oxidation state (CH_3OH), and is carried on either the 5-position or the 10-position nitrogen of the folic acid structure (figure 1).

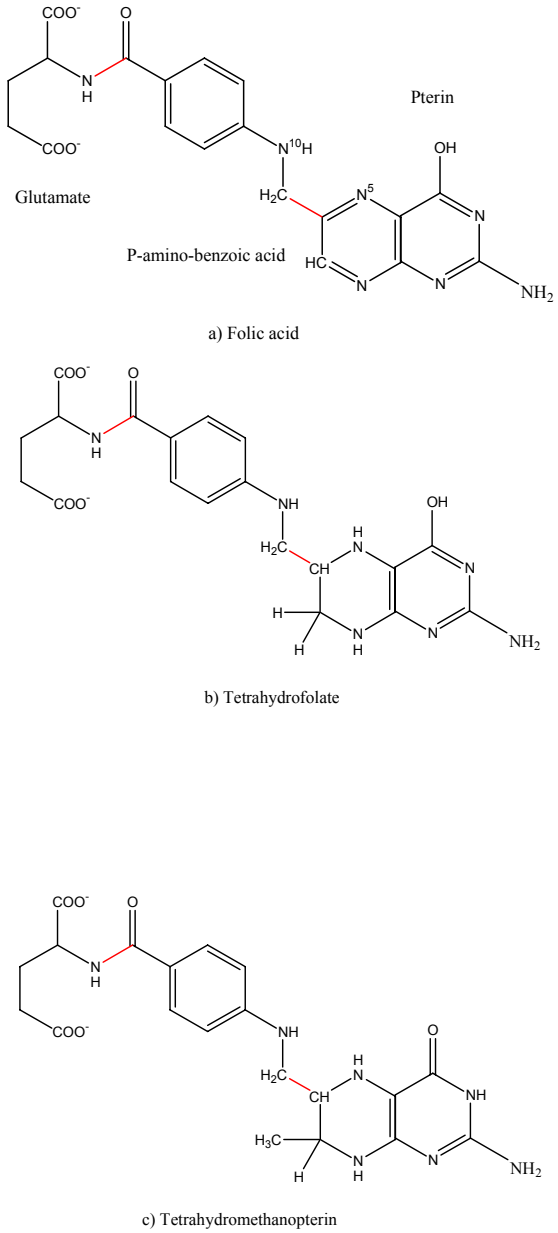


Figure 1 - Folic acid, tetrahydrofolate and tetrahydromethanopterin.

- Tetahydrobiopterin (BH_4) (figure 2) plays a key role in a vast number of biological processes and pathological states associated with neurotransmitter formation, vaso-relaxation, and immune response. BH_4 is functionally active as an electron donor for the hydroxylation of aromatic amino acids, that is, as the natural cofactor

for phenylalanine-4-hydroxylase (PAH), tyrosine-3-hydroxylase (TH), and tryptophan-5-hydroxylase (TPH); the latter two are key enzymes in the biosynthesis of biogenic amines (Scriver *et al.*, 1994) like catecholamines (Nichol *et al.*, 1985) (Tayhe and Marletta, 1989) (Kwon *et al.*, 1989). BH_4 also serves as the cofactor for nitric oxide synthase (NOS) (for the formation of nitric oxide from arginine) (Marletta 1993) and glyceryl-ether monooxygenase (Kaufman *et al.* 1990), and is used as a modulator in signal transduction. The function of these reactions derives from the ability of BH_4 to react with molecular oxygen to form an active oxygen intermediate that can hydroxylate substrates. In the hydroxylation process, the co-enzyme loses two electrons and is regenerated *in vivo* in an NADH-dependent reaction. Although BH_4 was found to be absolutely essential for nitric oxide synthase activity, the exact function in different forms of the enzyme and the mechanism of action are as yet not clear.

Abnormalities in the biosynthesis or regeneration pathways of BH_4 have been demonstrated in some diseases affecting the central nervous system such as hyperphenylalaninemia, atypical phenylketonuria or hereditary progressive dystonia, among other neurological disorders (Blau *et al.*, 1996), or even vascular diseases like artherosclerosis.

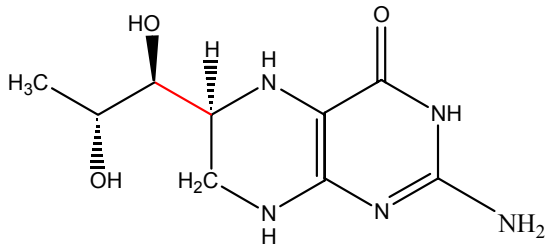


Figure 2 – Tetrahydrobiopterin (BH_4)

- Molybdopterins are cofactors for a large number of redox enzymes. Molybdenum-containing enzymes are a wide redox-active class of enzymes that mediate the metabolism of nitrogen, carbon and sulfur (Pilato and Stiefel, 1999). Molybdenum has been found in two distinct forms originating a classification in two different groups: a mixed [Fe-Mo_S] metallic cluster, which occurs in the group of nitrogenases (performs nitrogen fixation)(Stiefel, 1997), and a Mo atom associated with one or two pyranopterin-ene-1,2-dithiolate (molybdopterin) cofactors occurring in the oxotransferases group

(also designated as oxomolybdenum or mononuclear molybdenum group of enzymes). Molybdopterin cofactors are structurally active as tricyclic-pterin systems bound to molybdenum or tungsten via the dithiolene side chain (and termed as Moco cofactors). On prokaryotic sources the tricyclic pyrano-pterin system is usually bound to a nucleotide of guanine, cytosine, adenine or hypoxantine (figure 3).

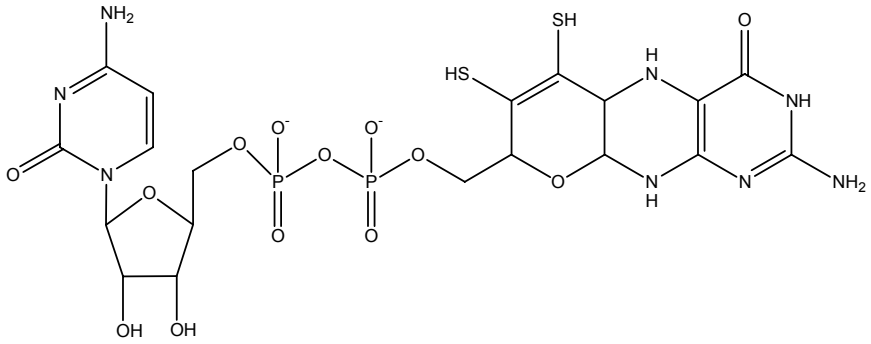


Figure 3. – The MCD cofactor: Molybdopterin Cytosine Dinucleotide.

Moco containing enzymes catalyse oxo-hydroxy two-electron transfer reactions, are functionally subdivided into five families on the basis of sequence homology, similar spectroscopic properties, catalysed reaction and environment of the active site (Hille, 1996). Examples of crystal structures from members of all these families were reported in the last six years: for the DMSO reductase family, a nitrate reductase (Dias *et at.*, 1999) and three DMSO reductase structures (McAlpine *et al.*, 1997) (Schindelin *et al.*, 1996) (Schneider *et al.*, 1996), for the sulfite oxidase family, the structure of sulfite oxidase (Kisker *et al.*, 1997) and finally the tungsten aldehyde ferredoxin oxidoreductase (AfOR) (Chan *et al.*, 1996) and formaldehyde ferredoxin oxidoreductase (FfOR) (Hu *et al.*, 1999), which constitutes a distinct family of enzymes – the Aldehyde Oxidoreductase family. To the family of the presently reported enzymes, the xanthine oxidase/molybdenum hydroxylase family of Mo enzymes, the first structurally elucidated representative was the aldehyde oxidoreductase from *Desulfovibrio gigas*, MOP (Romao *et al.*, 1995), followed by the structure of the CO dehydrogenase (Dobbek *et al.*, 1999). Recently the crystal structures of the bovine milk xanthine oxidase (Enroth *et al.*, 2000) and xanthine dehydrogenase from *Rhodobacter* capsulatus (Truglio *et al.*, 2002) were also described. This family

also comprise other members like formate dehydrogenase (FDH), quinoline-2-oxidoreductase, isoquinoline-1-oxidoreductase and nicotinic acid hydroxylase.

As an important example of a molybdopterin enzyme, the xanthine oxidase (XO) is a highly versatile flavoprotein enzyme, present among species from bacteria to human. This enzyme catalyses the oxidative hydroxylation of purine substrates and at the same time generates reactive oxygen species, either superoxide anion radical or hydrogen peroxide.

Either too small or too large concentrations of XO in human serum are likely to produce problems. The stimulation of the production of XO may provide higher protection against infections, as XO is an important generator of reactive nitrogen and oxygen species like peroxynitrite that work as antibacterial substances. On the other side the inhibition of this enzyme reduces the amount of uric acid in human serum, which is known to be a very important factor contributing for the occurrence of gout. The members of the xanthine oxidase family also include the enzymes xanthine dehydrogenase, aldehyde oxidoreductase, aldehyde oxidase and CO dehydrogenase. This family is characterized by the presence of a $Mo^{VI}OS$ nucleus and a mechanistically active water/hydroxide molecule coordinated to Mo and one molybdopterin cofactor (moco) non-covalentely bound to the polypeptide chain. These enzymes are present in both eukaryotic and prokaryotic organisms showing different substrate specificities. These are structurally complex enzymes, using up to five different types of cofactors to transfer electrons from or to their active sites located at the moco. While some bacterial enzymes like MOP only use two types of [2Fe-2S] clusters in addition to the moco, most of the other family members, like XO or CODH contain an additional FAD-containing domain or subunit.

3.1 GTP cyclohydrolase I enzyme

The formation of NH_2TP dihydroneopterin 3´-triphosphate is the first step in the biosynthesis of tetrahydrobiopterin (BH_4) (Nichol *et al.*, 1985) on animals and tetrahydrofolate (Brown and Williamson, 1987) (Green *et al.*, 1996) in plants and microorganisms with the exception of methanogenic bacteria, which utilize tetrahydromethanopterin (Eisenreich and Bacher, 1994). For the tetrahydrobiopterin (BH_4) biosynthesis, GTP-cyclohydrolase I, 6-pyruvoyltetrahydropterin synthase and sepiapterin reductase are catalytically responsible for the sequential formation of NH_2TP

dihydroneopterin 3´-triphosphate, PPH_4 6-pyruvoyltetrahydropterin and tetrahydrobiopterin (BH_4), respectively.

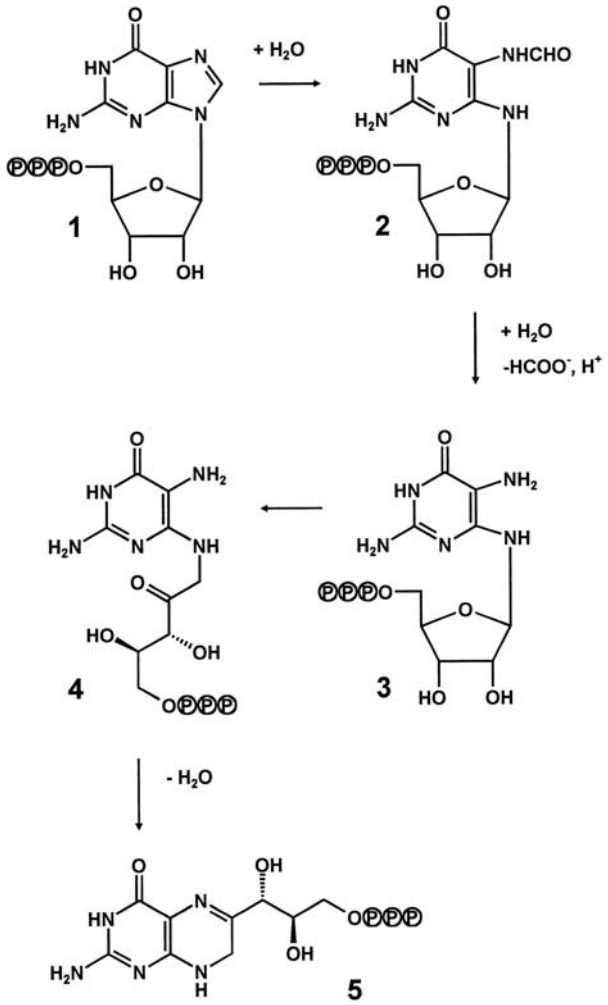


Figure 4 - Main steps on the formation of dihydroneopterin triphosphate (reactions catalysed by GTP Cyclohydrolase I).

The complex reaction sequence catalysed by GTP cyclohydrolase I is supposed to involve the hydrolytic opening of the imidazole ring of GTP under formation of a formamidopyrimidine type intermediate (Compound **2**), (Shiota *et al.*, 1969, 1967) (Schramek *et al.*, 2001a, 2001b) (Bracher *et al.*, 1998) the removal of the purine carbon atom 8 of GTP as formate affording the triaminopyrimidine derivative (Compound **3**), an Amadori rearrangement of the ribose moiety, and finally closure of the dihydropyrazine ring by intramolecular condensation affording dihydroneopterin triphosphate (Compound **5**). By this reaction sequence involving the hydrolytic cleavage of two carbon-nitrogen

bonds and one carbon-oxygen bond, carbon atoms 1' and 2' of the ribose side chain of GTP become part of the pteridine ring system of the product (Figure 4) (Weigand *et al.*, 1961) (Burg and Brown, 1968). Genetic defects of GTP cyclohydrolase I are proved to be responsible for severe neurological impairement (Thony *et.al.*, 2000) (Nagatsu and Ichinose, 1997) (Bandmann *et al.*, 1996a, 1996b).

The crystal structure of wild type *Escherichia coli* GTP cyclohydrolase I has been determined at a resolution of 2.6 Å (Nar *et at.*, 1995a, 1995b). The homodecameric protein has a mass of 247 kDa and the shape of a torus with dimensions of 65 Å x 100 Å. The D_5 symmetric structure comprises a β barrel composed of twenty anti-parallel strands surrounding a core of α-helical segments (Figure 5).

Recently, it was shown that the active sites of human and bacterial GTP cyclohydrolase I contain essential zinc ions that had escaped detection in earlier crystallographic studies (Auerbach *et al.*, 2000). The metal ion is supposed to act as a Lewis acid enabling the hydrolytic opening of the imidazole ring of GTP via activation of an attacking water molecule.

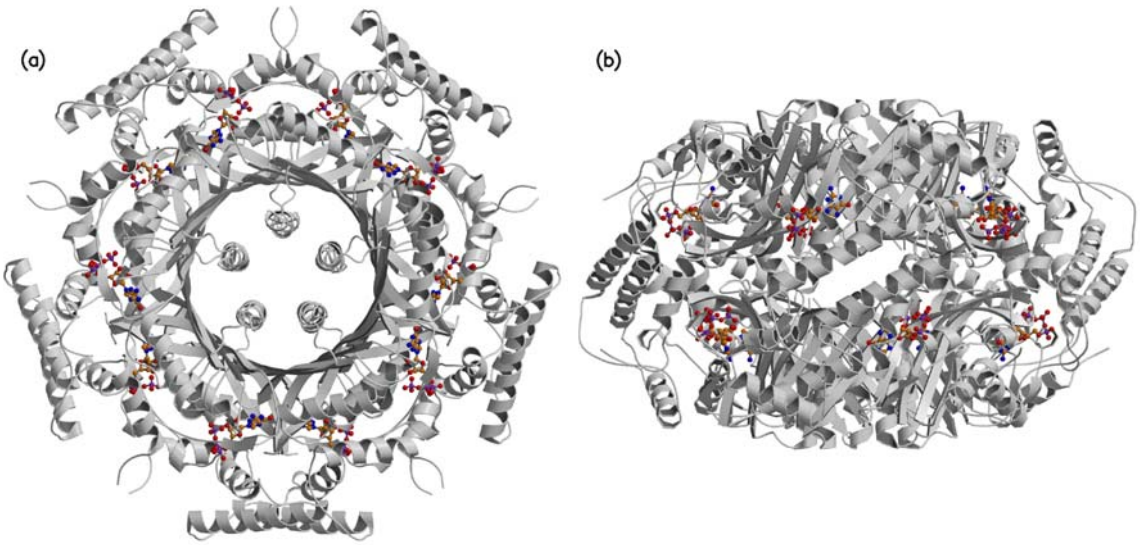


Figure 5 – D_5 homodecamer of the GTP Cyclohydrolase I from *E. coli*. The GTP substrates are represented as coloured ball-and-stick at their observed places. a) Top view. b) Side view. Figure prepared using MolScript (Kraulis, 1991) and Raster3D (Merrit & Murphy, 1994).

3.2 Molybdopterin containing enzyme Aldehyde oxidoreductase (AOR), as a member of the Xanthine oxidase family.

Like all the members of the xanthine oxidase family, the aldehyde oxidoreductase enzymes act as an oxo donor (by hydroxylation) to a substrate in a two-step one-electron transfer reaction, with water as the source of the oxygen atom that is incorporated into the substrate:

$$RH + H_2O \longrightarrow ROH + 2e^- + 2H^+$$

Where RH is an aldehyde for AOR or a purine aromatic heterocycle for the XO.

By this oxidation process aldehydes are transformed into carboxylic acids, releasing two electrons. The electron transfer from the substrate to the electron acceptors is in general made through Fe-S centers and flavin moieties, but in this case only two [2Fe2S] units were identified as being involved. There is very little specificity on the aldehydes that react with these enzymes, however it is known that glyceraldehyde is a substrate of the *D. gigas* enzyme (MOP), and therefore its physiological role may be linked to the degradation of polyglucose, which is accumulated by *D. gigas* under normal grouth conditions (Stams *et al.*, 1983).

The enzymatic mechanism was proposed on the basis of the 1.8 Å crystallographic studies on different functional forms of the enzyme (Huber *et al.*, 1996). A bound inhibitory isopropanol in the inner compartment of the substrate-binding tunnel is a model for the Michaelis complex of the reaction with aldehydes (H-C=O, R). The reaction was proposed to proceed by transfer of the molybdenum-bound water as OH after proton transfer to Glu869 to the carbonyl carbon of the substrate concomitant with hydride transfer to the sulfido group to generate [MoIV, =O, -SH, ---(O-C=O, -R)]. A Glu869 transient binding to the molybdenum may facilitate dissociation of the final carboxylic acid product (Figure 6). Regeneration of the metal bound water is then possible to happen due to the proximity of an internal chain of water molecules. In the organisms where it is present the AOR enzyme was shown to be part of an electron-transfer chain comprising several proteins, which is capable of linking the oxidation of aldehydes to the reduction of protons.

Several known features deepely relate the structure and consequently the mechanism of action from the AOR and the mechanism of hydroxylation performed by XO. Among these are the conservation of amino-acid sequences, particularly on segments engaged on

binding the molybdopterin cofactor and both Fe/S centres, also strong resemblance on the structural and spectroscopic properties of these common metal centres, and finally the widely demonstrated capacity of XO of using aldehydes as substrates.

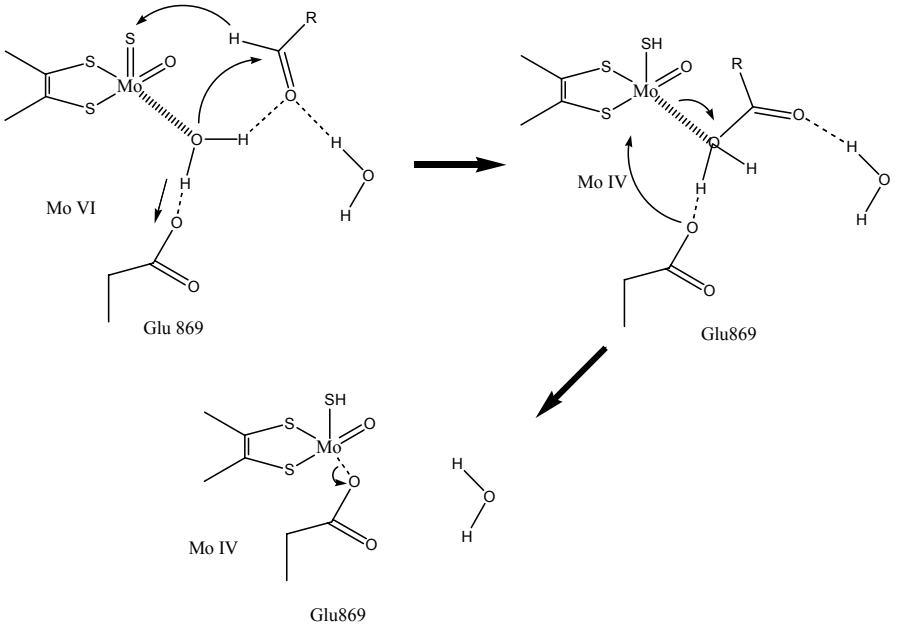


Figure 6 – Reaction mechanism of AOR

The latest crystallographic results from the XO demonstrate that the hydroxylation reaction mechanism is similar to the one previously proposed for MOP, as the reaction is initiated by a water derived hydration of a neighbouring Glu residue, then followed by a base-assisted nucleophilic attack of the MoVI-OH group to the C8 position of the xanthine substrate. This last process is assisted by a hydride transfer from the substrate to the Mo=S group leading to a reduced LMoIVO(SH)OR species, where OR represents the product coordinated to the molybdenum. Two one-electron steps re-oxidize the Mo and release the uric acid product (Figure 7).

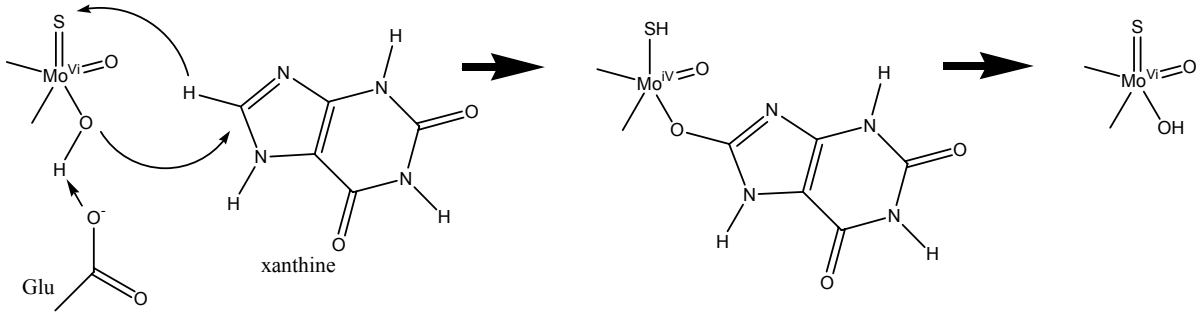


Figure 7 – Reaction mechanism of XO/XDH

II Materials and Methods

A. GTP cyclohydrolase I from *Escherischia coli*

II.A.Methods.1 Molecular Biology Methods

II.A.Methods.1.1 Transformation in chemically competent cells

The expression vector pNC0113 was used for hyprexpression of the GTP cyclohydrolase I point mutants from *E. coli*. The competent *E. coli* strain M15 [pREP4] was utilized for transformation purposes. The working genotype was: lac, ara, gal, mtl, $recA^+$, uvr^+, [pREP4: lacl, kan^r] (Zamenhof *et al*., 1972).

A heat chock transformation inserted the gene containing plasmid into the chemically competent *E. coli* cells. Transformed selection was achieved by overnight growth at 37ºC on a 2% agarose plate containing Luria Bertani (LB) medium plus ampicillin (170mg/l) and kanamycin (20 mg/l).

II.A.Methods.2 Protein Chemistry Methods

II.A.Methods.2.1 Growth of Cells overexpressing GTP cyclohydrolase I

Previously reported recombinant *E. coli* strains for hyperexpression of mutants of GTP cyclohydrolase I were grown and harvested as previously described (Nar *et al*., 1995a). Normal cell growth was carried until an optical density of 0.6 at 220 nm was achieved, followed by induction by 500 mg/l of IPTG, and overnight growth at 32ºC. No supplement of zinc was added to the growth solution.

II.A.Methods.2.2 Purification protocol

Cells suspended in 25 mM Tris hydrochloride, pH 8.0, containing 20 mM KCl and 3 mM sodium azide were lysed by ultrasonic treatment. The suspension was centrifuged, and the supernatant was heated to 60 °C for 10 min. After centrifugation, the supernatant was dialysed against 25 mM Tris hydrochloride pH 8.2, containing 20 mM potassium chloride, 3 mM sodium azide, and 24 mM GTP. The solution was applied to a column of Sepharose Q (Pharmacia; 3.8 x 11 cm), which was developed with a gradient of 100 to 600 mM potassium chloride. Fractions were combined and applied to a Superdex-200 column (from Pharmacia), which was developed with 25 mM Tris hydrochloride, pH 8.2, containing 100 mM KCl, and 3 mM sodium azide.

II.A.Methods.2.3 Enzyme Assays

GTP cyclohydrolase I activity was determined as described (Bracher *et al*., 1998 and 1999).

II.A.Methods.2.4 Determination of zinc content

To a solution of enzyme (3 mg/ml) in 50 mM Tris hydrochloride pH 8.0 buffer, hydrochloric acid was added to a final acid concentration of 1 M. This mixture was heated to 96 °C for 2 h. Zinc was then determined by atomic absorption spectrometry using a Unicam 919 spectrometer (Unicam, Cambridge, UK).

II.A.Methods.2.5 Protein concentration measurement

GTP cyclohydrolase I concentrations were measured by the Bradford method. Four standard probes and one measure probe were obtained by adding 200 µl from BioRad protein assay solution and 798 µl of water plus 2µl of 1 mg/ml, 2mg/ml, 3mg/ml and 4mg/ml for each one of the standard probes respectively and 1µl of protein sample plus 1µl of water for the measurement probe.

II.A.Methods.3 Crystallography Methods

II.A.Methods.3.1 Protein crystallization

Crystallization experiments were performed at room temperature employing the vapour diffusion technique. Crystals of GTP cyclohydrolase I mutant proteins were obtained under the following conditions. (a) Mutant His112Ser*/unsoaked crystals: 0.1 M sodium MES, pH 6.0, containing 0.2 M sodium acetate and 3 mM sodium azide. (b) Mutants His112Ser and Cys181Ser: 0.1 M MOPS, pH 7.0, with 10 % (w/v) PEG 6K and 0.1 M ammonium sulphate. (c) Mutant His113Ser: 0.1 M Tris hydrochloride, pH 8.5, containing 0.2 M ammonium dihydrogen phosphate and 50 % MPD. The orthorhombic crystals corresponding to the two first crystallization conditions (conditions (a) and (b)) belong to the space group $C222_1$ with 15 monomers in the asymmetric unit and lattice constant a = 315.9 Å, b = 220.6 Å, c = 131.4 Å for mutant His112Ser* (under cryo conditions), a = 224.46 Å, b =313.26 Å, c = 130.18 Å for mutant His112Ser (under cryo conditions) and a = 223.14 Å, b = 317.8 Å, c = 132.1 Å for mutant Cys181Ser at room temperature. The crystals of mutant His113Ser obtained with MPD as precipitant (condition (c)) belong to the space group $P4_32_12$ with 10 monomers in the asymmetric unit and lattice constants a = b = 123.9 Å, c = 388.6 Å. The cell angles are $\alpha = \beta = \gamma = 90^{\circ}$ for all cases.

II.A.Methods.3.2 X-ray diffraction data collection and processing

Crystals were soaked for 30 min in harvesting buffer supplemented with GTP to a final concentration of 10 mM GTP. After soaking, the crystals were immediately frozen in a nitrogen stream using glycerol or MPD (mutant His113Ser) as cryoprotectant.

The data set of the complex of Cys181Ser with GTP was collected to a limit resolution of 3.2 Å from a single crystal at room temperature on a Mar Research imaging plate system installed on a Rigaku rotating anode generator (λ = 1.54 Å; 0.5 ^{o} frames). The other data sets were collected from single frozen crystals at the BW6 beam-line of DESY (Hamburg) (λ = 1.05 Å; 0.3 ^{o} frames) at a temperature of 95 K using a CCD detector.

Indexation and integration of the crystal diffraction images from the mutants His112Ser* and GTP-Cys181Ser was made using the MOSFLM package (Leslie 1991), and from the mutants GTP- His112Ser and GTP-His113Ser using the DENZO/SCALEPACK package (Otwinowski and Minor, 1997). Further processing, was carried by the CCP4 suite of programs (Bailey, 1994). A total of 661,363 measured reflection intensities for the His112Ser* merged on to 238,566 unique reflections with R_{merge} on intensities of 8.5 %. This set represents 90.5 % of the possible data to a resolution of 2.1 Å. For the His113Ser

complex, 100128 unique merged reflections (R_{mefge} of 4.7 %) were obtained from a total of 1005735 collected measurements, representing 96.9 % completeness of the data. The His112Ser complex yield 105682 unique reflections from 2178312 intensity reflections measured, characterized by R_{merge} of 5.1 % and completeness of 97.4 % for the total data to 2.8 Å resolution. Finally, the Cys181Ser complex 1260271 intensity measurements resulted on 70145 unique reflections complete to 98.6 % with overall R_{merge} of 6.3 % for all the data to a resolution of 3.2 Å (Table 1).

II.A.Methods.3.3 Structure solution and refinement

The model of the wild type enzyme (Nar *et al.*, 1995a, 1995b) was used for refinement of the mutant complex structures after molecular replacement with Molrep (Vagin and Teplyakov, 1997), followed by rigid body refinement with CNS (Brünger, 1990,1998). The electron density of His112Ser* was modified applying non-crystallographic symmetry averaging with the program MAIN (Turk, 1992). Model building was done with O (Jones, Bergdoll & Kjeldgaard, 1990) and refinement with CNS. The 15 monomers of His112Ser*, His112Ser and Cys181Ser and the 10 monomers of His113Ser per asymmetric unit were strongly constrained by non-crystallographic symmetry during initial refinement and released at the medium to final stages. All models refined to satisfying overall quality, with R-factors of $R_{His112Ser*}$ = 20.0 %, $R_{His113Ser}$ = 26.2 %, $R_{His112Ser}$ = 21.8 % and $R_{Cys181Ser}$ = 18.5 % obtained at 3σ cut-off level. Other important refinement information is also presented in Table 1.

II.A.Methods.4 Software

Data processing was made using DENZO-SCALEPACK (Otwinowski & Minor, 1997) or MOSFLM (CCP4, 1994), followed by scaling and/or formatting software from the CCP4 suite of programs (CCP4, 1994). Molecular Replacement calculations were done with MOLREP (CCP4, 1994), map calculations and refinement were carried out with X-PLOR (Brünger, 1990, 1998) using the Engh & Huber (1991) force field. Model building was done with O (Jones et *al.,* 1990). Sequence alignment and homology and identity calculations were made with CLUSTALW (Higgins *et al.*, 1994) and sequence alignment figure was prepared with AlScript (Barton, 1993). The validation tool for geometric quality assessment was PROCHECK (Laskowski *et al.*, 1993). All model pictures where drawn with either BobScript (Esnouf, 1997) or MolScript (Kraulis, 1991) and Raster3D (Merrit & Murphy, 1994).

Data set	**His112Ser /GTP**	**His113Ser /GTP**	**Cys181Ser /GTP**
Data collection			
Source[1]	DESY	DESY	RA
Space group	$C222_1$	$P4_32_12$	$C222_1$
Observations, No.	2,178,312	1,005,735	1,260,271
Unique reflections, No.	105,682	100,128	70,145
R_{merge}, %[2]	5.1	4.7	6.3
Resolution, Å	2.8	2.55	3.2
Data completeness, %	97.4	96.9	98.6
Refinement			
No. of atoms in refinement	26,509	17,827	26,475
Resolution limits, Å	10.0 – 2.8	20.0 – 2.5	11 – 3.2
R factor, %[3]	21.8	26.2	18.5
R_{free}, %[3,4]	27.2	29.3	22.8
r.m.s.d.[5] bonds, Å	0.009	0.013	0.008
r.m.s.d. angles, degrees	1.51	1.94	1.47
B factor (averaged, all monomers), Å^2	55.0	56.3	43.5
r.m.s.d., bonded B factor, Å^2 (main chain)	1.2	1.5	1.4

Table 1: Crystallographic Data. [1]RA, rotation anode source; DESY, Synchrotron Hamburg, Germany. [2]R_{merge}: $R_s=\Sigma(I-\langle I\rangle)/\Sigma(I)$, where I is the measured intensity and $\langle I\rangle$ is the averaged value; the summation is over all measurements. [3]Observations above 3σ cut off. [4]Free R factor calculated by setting aside 10 % of the reflections before refinement. [5]Root mean square deviations from ideal values.

B. The Aldehyde Oxidoreductase from *Desulfovibrio gigas* (MOP)

II.B.Methods.1 Protein Chemistry Methods

II.B.Methods.1.1 Cell growth and protein purification

Aerobic purification was carried out following the strategy described by Moura *et* al. (1976, 1978).

II.B.Methods.2 Crystallography Methods

II.B.Methods.2.1 Protein crystallization

Crystallisation was achieved by vapour diffusion on sitting drops using a mixture of purified protein in 10 mM Tris/HCl buffer (pH 7.5) and a crystallizing solution of 30% (v/v) isopropanol as precipitant and 0.2M $MgCl_2$ as additive in 0.2M HEPES (pH 7.6), following the procedure described by Romão *et* al. (1993). The presence of $MgCl_2$ (or $CaCl_2$) was essential for crystallization and contributes to the high stability of these crystals (see below). Crystal nucleation and growth up to a size of about 0.6x0.4x0.4 mm takes about three weeks at 4ºC. A cryo-protectant harvesting buffer prepared with the crystallization solution enriched in polyethylene glycol 4000 (PEG 4000) at 30% (w/v) was added to the crystals which were thus stable over one month.

II.B.Methods.2.2 X-ray diffraction and data processing

For the data collection of one single crystal cryo-cooled at 100 K, two independent sets of frames were collected, one corresponding to the lower and the other to the higher resolution shells. The measurements were made on a MAR 300 imaging plate x-ray detector at the BW6 beam-line of DESY. The diffraction data were indexed, recorded, integrated and processed using DENZO, scaled and merged with SCALEPACK (Otwinowski & Minor, 1997). Structure factors were then derived from intensities using TRUNCATE from the CCP4 package (1994), and a starting X-PLOR formatted file was generated by MTZ2VARIOUS also from CCP4. The total set of 452521 reflections was processed with anomalous data included, yielding 233755 unique reflections with an overall R_{sym} 3.6% (43.6% in the outer shell). Important statistics are summarized in Table 2.

Temperature (º K)	100
Wavelength (Å)	1.0004
Scan rate (º min^{-1})	0.5
θ Range (º)	0.8
Resolution range (Å)	24.4 – 1.28
Space group	$P6_122$
Cell constants (Å)	141.78 160.87
Total number of measured reflections	452521
Nº of independent reflections (all)	233755
(4 σ cutoff)	155842
Completeness (%) (overall)	97.8
(last shell: 1.33-1.28 Å)	92.3
<I> / σ (<I>) (overall)	17.7
(last shell)	1.6
Rsym (%) (overall)	3.6
(last shell)	43.6
Rmerge(%) (overall)	9.1
(last shell)	64.9
Solvent content (%)	54
Overall temperature factor ($Å^2$)	
Wilson plot (3.84-1.28 Å)	16.2

Table 2. Data collection statistics

The new cell constants of the cryo-cooled crystals decreased to about 2% in comparison to the previous crystals which were analysed at –17 ºC, reflecting a tighter packing, keeping 54% of solvent content as calculated by the Matthews coefficient (Matthews, 1968).

II.B.Methods.2.3 Structure Solution and Refinement

An initial randomly chosen subset of 3% of the total observations was isolated for cross validation by the R_{free} factor (Brünger, 1992). Refinement started with X-PLOR 3.1 (Brünger, 1990, 1998), first including only the data below 1.8 Å of resolution. The first atomic model used was taken from an earlier 1.8 Å refinement (Huber *et al.*, 1996). In this first partial model redox cofactors, alcohol and water molecules were excluded. After 40 cycles of rigid body refinement, the R and R_{free} factors decreased from 53.6 and 51.9 to 35.2 and 39.5, respectively. At this stage, the quality of the electron density maps was good enough to correctly place all cofactors as well as some ordered internal solvent water molecules. The MCD cofactor and both iron sulphur clusters were included in the atomic model for the subsequent refinements. The geometric and force field parameter sets, implemented by Engh and Huber (Engh & Huber, 1991), were applied. As the

cofactors were included on the model, it was clear that the MCD would also require guiding target values to avoid distortion (see below). Rigid body refinement was followed by 150 cycles of conjugated gradient Powell minimisation (Brünger, 1990, 1998), which led to an R_{free} and an R-value of 34.7 and 31.1 respectively. Additional 25 cycles of individual restrained isotropic temperature factor refinement decreased the R_{free} and R to 33.9 and 29.9 respectively. Higher resolution data were stepwise included from 1.8 Å to 1.45 Å and finally to 1.28 Å. At each step some cycles of automated refinement were applied and followed by visual map inspection and model rebuilding in TURBO-FRODO (Cambillau, 1997). In each step of visual inspection, more solvent molecules were included when their density was above 1.85 $Å.e^{-3}$ (3σ) on Fo-Fc electron density maps, and when they were within H-bonding distance to acceptor groups. Then an automated bulk solvent model was activated, the X-PLOR bulk solvent correction (Brünger, 1990, 1998), which allowed all measured low-resolution data to be included, and improved significantly the quality of the electron density maps. At this stage the R factor decreased from 23.7 to 20.7 and the R_{free} from 27.6 to 24.4 on a 2σ cut-off electron density level.

Refinement was continued with SHELX97-2 (Sheldrick & Schneider, 1997) and introducing anisotropic ADP´s (atomic displacement parameters) described by the symmetric Uij second rank tensor for all non-hydrogen atoms. This change in the refinement protocol was made because SHELX shows advantages in relation to X-PLOR on high-resolution refinements. It employs conventional Fourier structure factor summations rather than FFT structure factor summations and allows the anisotropic ADP refinement as well as the ability to refine all positional and thermal displacement parameters simultaneously. Beyond keeping the Engh and Huber geometric restraints (Engh & Huber, 1991), new types of restraints related to Uij refinement were also necessary. The first new set of restraints working as rigid bond Uij restraints, forced bonded atoms to have the same vibration along the direction of the bond (DELU restraint of SHELXH (Sheldrick & Schneider, 1997)). A second set of restraints limited the degree of anisotropy (ISOR restraint (Sheldrick & Schneider, 1997)). The last set was used to avoid sharp variations between adjacent atoms (SIMU restraint (Sheldrick & Schneider, 1997)).

A refinement algorithm with CGLS (Conjugated Gradient Least Squares) (Tronrud, 1992) and diffuse solvent modelling by the Babinet´s principle was used throughout the rest of the subsequent refinement (Table 3). Single atom anisotropic refinement, alternate conformation modelling, water occupancy rough refinement by setting the occupancy of

some waters to 0.5 on several disordered solvent zones, and finally the automatic placement of non-refined ridding H atoms on calculated positions lead to a final R of 14.8 and R_{free} of 19.4. A summary of the most important refinement statistics is presented on Table 3 and final model parameters are summarized in Table 4.

Working set (95%)	226751	
Independent cross validation set (3%)	7004	
XPLOR (CGLS)*	R	R_{free}
Resolution range	24.4	1.80
-After molecular replacement	53.6	51.9
-Rigid body refinement	35.2	39.5
-Cofactor inclusion, atomic co-ordinate refinement	31.1	34.7
-B-fact refinement	29.9	33.9
-Model corrections	25.1	29.3
Resolution range	24.4	1.5
-Side chain corrections, water inclusion	24.5	28.7
Resolution range	24.4	1.28
-Water inclusion, rotamers correction	23.7	27.6
-Bulk solvent correction	20.7	24.4
SHELXL (CGLS) **	R	R_{free}
-Isotropic refinement	21.1	25.2
-Anisotropic refinement	17.3	22.1
-Double conformation modeling corrections, more solvent	15.6	20.6
-End of refinement (4 σ cutoff)	14.5	19.3

*2 sigma cutoff;
**2 sigma cutoff estimated by linear interpolation between 0 and 4 sigma cutoff.

Table 3. Refinement protocol

Non-H atoms		8146
907 amino acids		6844
1 MCD		44
2 [2Fe-2S]		8
Waters		1243
Mg		2
Cl		3
Isopropanol		2
Refinement parameters/statistics (SHELX [ref])		
Anisotropic temperature U_{ij} restraints *		
	rmsd	target σ
DELU	0.012	0.030
ISOR	0.108	0.200
SIMU	0.053	0.274
Restrained goodness of fit		1.4
Total nº of reflections		233755
Nº of restraints		87737
Nº of parameters		73378
(observations+restraints)/parameters		4.4
Observations/restraitns		3.6
Structure quality statistics		
Estimated average atomic Co-ordinate error (Å)**		0.10
Average B_{eq}-factor (Å^2)	overall	24.7
	protein	21.7
	cofactors	16.0
	solvent	38.9
	rms	Target σ
Bond distances (Å)	0.013	0.020
Bond angle distances (Å)	0.029	0.040
Bonded B-factors	3.3	-
Ramachandran (%)		
	Allowed	91.6
	Generously allowed	8.0
	Disallowed	0.4
Overall quality measure	(g-factor)	-0.11
RSC outliers (%)		2.6
Peptide flips (%)		3.0
R-factor	(4σcutoff)	14.5
	(all data)	15.9
Free R-factor	(4σ cutoff)	19.3
	(7004 reflections)	21.4

* According to definition described on the text

** Average error estimated from Luzzatti plot

Table 4. Final model parameters and statistics

II.B.Methods.2.4 MCD cofactor target values driven by CSD search

The MCD cofactor structure as analysed below was obtained by almost atomic resolution crystallographic refinement. However during this refinement some structural information was required at early stages and used as restraint. There are no known small molecule structures with related or similar structure, excluding the associated cytosine dinucleotide. To overcome this problem, a total of 44 bond types searches corresponding to all the unique MCD internal bonds were executed on the Cambridge Structural Database (CSD) (Allen *et al.*, 1979). Each of these searches produced a histogram of bond distances found among all structures available. The bond distance search pattern for each bond between pairs of atoms within the cofactor also included information about the environments of these pairs of atoms, such as, to which atoms the named pair of atoms are bonded to, and the number and types of bonds this pair of atoms establishes. For all bonds many occurrences could be detected, and several histograms reached the stop limit of 600 occurrences imposed to end the search. Following histogram analysis, a standard deviation and a mean value were calculated for each histogram. The mean values were calculated and were subsequently introduced as initial bond distance target values for each of the atomic bonds on the MCD moiety, but as refinement progressed these restraints were decreased and completely released in the final calculations on which the reported MCD geometry is based.

II.B.Methods.3 Software

Data processing was made using DENZO-SCALEPACK (Otwinowski & Minor, 1997) formatting software from the CCP4 suite of programs (CCP4, 1994). Molecular Replacement calculations, map calculations and refinement were carried out with X-PLOR (Brünger 1990, 1998) and using the Engh & Huber (1991) force field. Model building was done with TURBO-FRODO (Cambillau *et al.*, 1997). The validation tool for geometric quality assessment was PROCHECK (Laskowski *et al.*, 1993). All model pictures where draw with either on TURBO-FRODO, or MolScrip (Kraulis, 1991) followed by Raster3D (Merrit & Murphy, 1994).

C. The Aldehyde Oxidoreductase from *Desulfovibrio desulfuricans* 27775 (MOD)

II.C.Methods.1 Molecular biology methods

II.C.Methods.1.1 Sequencing of the gene encoding MOD.

The gene sequece was kindly provided by Dr S. Macieira when developing the following protocol: Oligonucleotide primers were synthesized based on the amino acid sequence of the N-terminus of MOD and a highly conserved region of molybdopterin enzymes, and were used in a PCR reaction to amplify a DNA fragment representing about 250 bp near the 5'-end of the MOD gene. The amino acid sequence ALLVDV, obtained by Edman degradative sequencing of the N-terminus, yielded one of the primers, lucky1: 5'-GCCCTGCTGGTNGACGT-3'. A multiple sequence alignment of aldehyde oxidoreductases, aldehyde oxidases, and xanthine oxidases (Hille, 1996), identified a region approximately 100 amino acids from the N-terminus that is absolutely conserved in this family of enzymes. The conserved amino acids, QCGFC, were used to design a second primer, lucky2: 5'-GCAGAAGCCGCATTG-3' (complementary strand). PCR amplification followed by automated DNA sequencing of the PCR product yielded a DNA sequence that coded for an open reading frame highly homologous to the *D. gigas* MOP gene (Thoenes *et al.*, 1994).

This fragment served as a guide to obtain additional DNA sequence of the MOD gene using two approaches. In one approach, nested primers based on the sequence obtained above were synthesized and used according to manufacturer's directions provided in the Universal Genome Walker Kit (Clontech, Palo Alto, CA) to generate two additional PCR fragments encompassing 578 bp of the 5'-end of the MOD gene including 67 bases of the 5'-untranslated region. In a parallel approach, the DNA sequence obtained in the first step was used to design a homologous forward primer, MOD1:5'-GCGTGGATCCAGCACGGCGCGGCACAATGCGGCTTCTGC-3', and X-ray derived sequence DGPFGA, was used for designing a reverse primer, MOD864R:5'-GCICCRAAIGGICCRTC-3'. A DNA fragment spanning to the 3'-end of the MOD gene, representing 2319 bp, was amplified by PCR using Taq DNA polymerase (MBI Fermentas).

This product was direct-sequenced on both forward and reverse directions by primer walking, using an automated DNA sequencer (model 373, Applied Biosystems, Forster City, USA) and the PRISM ready reaction dye deoxy terminator cycle sequencing kit

(Applied Biosystems, Forster City, USA). From the obtained DNA sequences, the 'inside-out' primers LASTFWD: 5'-CCCTCTATCAAGATGATTCC-3', and MOD1REV: 5'-GCAGAAGCCGCATTGTGCCGCGCCGTGCTGGATCCACGC-3' were constructed for Inverse PCR DNA amplification (Triglia *et al.*, 1988) (Ochman *et al.,* 1988) (Silver *et al.*, 1989).

A restriction enzyme map of the known MOD gene sequence was produced with the program MAP of the GCG package (Genetics Computer Group, 1994). The restriction enzyme NdeI (New England Biolabs) was chosen for digestion of an aliquot of genomic DNA (5 µg) from *D. desulfuricans* ATCC 27774 in a volume of 20 µl. After overnight incubation the enzyme was heat inactivated for 20 min at 70 °C. Intramolecular ligation of the restriction fragments was performed using T4 DNA ligase (New England Biolabs) for 16 h at 16 °C in a total volume of 100 µl. Using these DNA circles as templates, the primers LASTFWD and MOD1REV, and Taq DNA polymerase, a PCR product of about 1300 bp was obtained. The PCR product was directly sequenced giving the remaining sequence of the MOD gene. Assembly of the overlapping sequence fragments allowed reconstitution of an open reading frame of 2721 bp, coding for a 907 amino acid polypeptide.

II.C.Methods.2 Protein Chemistry Methods

II.C.Methods.2.1 Protein purification

The protein was purified by a procedure similar to the one used for the *D. gigas* enzyme (Barata *et al*., 1993) with some modifications and it was dialyzed against 10mM TRIS-HCl, pH=7.6.

II.C.Methods.3 Crystallography methods

II.C.Methods.3.1 Protein crystallisation

For protein crystallisation, the vapour diffusion method was used at room temperature and the crystallisation solution contained 1.8 M ammonium sulphate, 0.1 M MES (pH 6.0) and 10% dioxane. This last additive was essential for obtaining larger and better diffracting crystals. Hanging drops were prepared with 4 μl of reservoir buffer mixed with 4 μl of a 15 mg/ml protein solution in 10 mM Tris-HCl, pH=7.6. Single crystals appear within three to four weeks and are red-brownish hexagonal prisms.

II.C.Methods.3.2 X-ray diffraction and data collection

X-ray-diffraction intensity data was collected with a conventional X-ray source, an Enraf-Nonius rotating anode generator (Cu-K_α) operated at 4.5 kW and equipped with a Mar Research Imaging Plate detector. Data were collected at room temperature from a single crystal mounted in a sealed capillary, leading to a maximum resolution of 2.5 Å, from which only 2.8 Å resolution had acceptable quality.

II.C.Methods.3.3 Crystallographic data processing

The complete set of the detector images were indexed, and followed by parameter refinement in DENZO. The SCALEPACK package ended data processing by scaling and merging the data yielding a final file with a complete set of data (Otwinowski & Minor, 1997). Data were truncated and Friedel mates were merged (CCP4, 1994) and a startup file with X-PLOR format was finally available.

II.C.Methods.3.4 Structure solution

The MOD structure was solved by molecular replacement techniques using the program package AMoRe (Navaza, 1994). The model of the closely related aldehyde oxidoreductase from *D. gigas* (MOP) in the oxidized form was used as a search model, after omitting cofactors and solvent molecules. Using data in the resolution range between 15.0 and 3.5 Å a unique and clear solution was obtained with an R factor of 41.6% and a correlation coefficient of 0.475. After rigid body refinement the R factor dropped to 36.7% with a correlation coefficient of 0.591. The final values for the rotation and

translation functions were α=20.1º, β=19.4º, γ=50.3º (in Euler angles) and t_x=0.1942, t_y=0.4149, t_z=0.4434 (fractional co-ordinates), respectively.

II.C.Methods.3.5 Model building and refinement

The refinement of the correctly positioned molecule was initiated, in the first cycles, still with the unmodified MOP sequence and was carried out with X-PLOR 3.81 (Brünger, 1990, 1998) using a least-squares refinement algorithm with FFT (Agarwal, 1978) and the parameter set of Engh & Huber (1991). The total recorded range of data, 24.9 Å to 2.8 Å with a 2σ cut-off level were employed. In the beginning only the first 100 amino acid residues from the N-terminus were known, which led to a fall on the R-factor from 36.7 % to 26.8 % and on the free R-factor from 36.0 % to 32.6 %. Each cycle of refinement involved map inspection and correction followed by 100 steps of positional, and 50 steps of individual B-factor automated refinement. The parameter file for the molybdopterin cofactor was taken from the optimised geometry of the cofactors from the 1.28 Å resolution refined model of MOP (see Results). Improved phases enabled the calculation of clearer electron density maps, which contributed to the design of new PCR primers for completing the gene sequence. When more than half of the sequence was known, bulk solvent correction was applied reducing the R and free R factors by about 3%.

	All (last shell)
Resolution (Å)	24.4-2.8 (2.9-2.8)
Total no. of measured reflections	190328
No. of unique reflections	31573
Completeness (%)	98.2 (98.0)
R_{merge} * (%)	16.9 (56.1)
$I/\sigma_{(I)}$	6.7 (2.4)

*Rmerge *(I)* = Σ (| *I(k)* - <*I*> |) / Σ *I(k),* summation over all measurements of individual diffraction intensity values *I(k)*. <*I*> represent the corresponding mean values.

Table 5. Data collection and processing statistics

Throughout the refinement, the R-free (Brünger, 1992) was calculated by using a random independent set with 3% of all collected data omitted from refinement. The rebuilding and adjustments of the model, such as sequence replacement, to the displayed electron

density maps of this structure were performed using the graphical program TURBO (Cambillau *et al.*, 1997). Data collection statistics and the progress of the refinement of the MOD model are summarised on Tables 5 and 6 respectively.

	R	Rfree
-Rigid body refinement with		
100 N-terminus amino acid residues known	26.8	32.6
-Sequenced until amino acid residue 344		
and chain adjustments	25.8	32.5
-Sequenced until amino acid residue 614		
Insertions not included	25.0	31.3
-Bulk solvent modelling		
Inclusion of two insertions on primary sequence	21.6	28.3
-Sequenced until amino acid residue 713		
Conformational corrections and solvent inclusion	18.0	25.1
-Sequenced until amino acid residue 864		
and solvent inclusion	16.4	23.6
-Primary sequence completed		
	16.4	22.4

Table 6. Stepwise R-factor and free R-factor improvement (%)

II.C.Methods.4 Software

Rough data processing was made using DENZO-SCALEPACK (Otwinowski & Minor, 1997), and the CCP4 suite of programs (CCP4, 1994). Molecular Replacement calculations were done with AmoRe (Navaza, 1994), map calculations and refinement were carried out with X-PLOR (Brünger, 1990, 1998) and using the Engh & Huber (1991) force field. Model building was done with TURBO-FRODO (Cambilau *et al.*, 1997). Sequence alignment and homology and identity calculations were made with CLUSTALW (Higgins *et al.*, 1994) and sequence alignment figure was prepared with ALSCRIPT (Barton, 1993). The validation tool for geometric quality assessment was PROCHECK (Laskowski *et al.*, 1993). Hydrogen bond analysis of the cofactors was made using HBLPLUS (McDonald *et al.*, 1994). All model pictures where draw with

either TURBO-FRODO (Cambilau *et al.*, 1997) or MolScript (Kraulis, 1991) and Raster3D (Merrit & Murphy, 1994).

III Results and Discussion

III-A - GTP cyclohydrolase I – The complex first step of tetrahydropterin biosynthesis in *E. coli*.

III.A Results

III.A.Results.1 Background

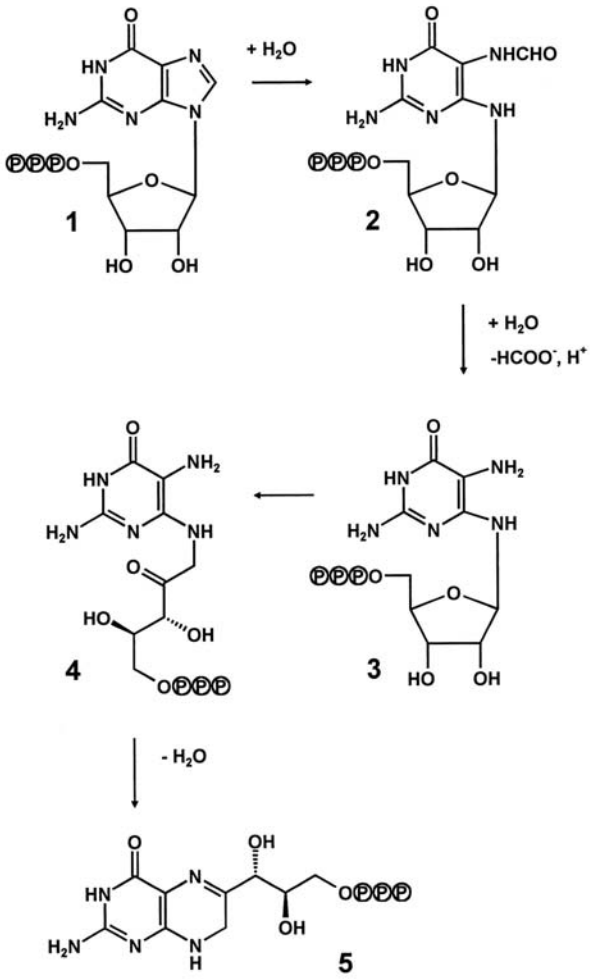


Figure 4. Hypothetical mechanism of the reaction catalysed by GTP cyclohydrolase I.

Earlier X-ray diffraction studies on a complex of wild type GTP cyclohydrolase I with the substrate analogue, 2'-deoxy GTP, had shown that each of the 10 active sites of GTP cyclohydrolase I is located at the interface of three adjacent subunits (subsequently labelled A, B, and D) (Nar *et al.*, 1995a). However, the electron density had not been sufficiently clear in order to analyse the interaction between protein and bound ligand in detail. Since several purified mutant proteins failed to produce dihydroneopterin 3'-triphosphate (Compound **5**) from GTP (Table 7), some of these where considered as good candidates for a crystallographic analysis of the interactions between the GTP substrate and the active site residues (Nar *et al.*, 1995a). For the purpose of this study, single amino acid replacement mutant proteins carrying serine residues instead of Cys 110, Cys 181, His 112 or His 113 were purified from recombinant *E. coli* strains as described under Methods.

Mutation	**Zn content**	**Activity** nmol mg^{-1} min^{-1}	
		Substrate	
	mol/mol	GTP	compound **2**
none	0.9	91	88
H112S	< 0.2	< 0.1	< 0.1
H113S	< 0.2	< 0.1	< 0.1
C110S	< 0.2	< 0.1	< 0.1
C181S	< 0.2	< 0.1	< 0.1

Table 7: Properties of GTP cyclohydrolase I wild type and mutants under study

Zinc content and dihydroneopterin triphosphate formation from the natural substrate GTP and 2-amino-5-formylamino-6-ribofuranosylamino-4(*3H*)-pyrimidinone triphosphate (compound **2**). The contents of bound zinc and GTP are referenced to the estimated number of GTP binding sites. Enzymatic activity is given as nmol dihydroneopterin triphosphate formed per minute with 1 mg of protein.

Recently, 2-amino-5-formylamino-(6β)-ribosylamino-4(3*H*)-pyrimidinone 5'-triphosphate (Compound **2**) has been identified as an early intermediate of the reaction catalyzed by GTP cyclohydrolase I. This intermediate was prepared by the catalytic action of the His179Ala mutant of GTP cyclohydrolase I as described elsewhere and was used as substrate for the mutants under study (Bracher *et al.*, 1999). As shown in Table 7, all mutants studied were unable to convert this reaction intermediate to dihydroneopterin 3'-triphosphate (Compound **5**).

Zinc was measured in the enzyme preparations by atomic absorption spectrometry. Whereas the wild type protein contained about 0.9 zinc ions per subunit, the zinc concentration of the recombinant mutant proteins was below the level of detection (Table 7).

III.A.Results.2 Crystallographic analysis

The His112Ser mutant was crystallized using the experimental conditions, which have been reported earlier for the wild type protein. A complete data set of a single crystal was determined at a temperature of 95 K to a resolution of 2.1 Å. The structure was determined using model phases of the wild type enzyme (Nar *et al.*, 1995b) and was refined to a resolution of 2.1 Å with an R-factor of 20.0 %.

The inspection of the electron density map showed that the active site contained non-protein density that could be interpreted as a nucleoside triphosphate, although none had been added during purification or crystallization. Subsequent biochemical analysis revealed that the chromatographic purification of the Cys110Ser mutant and the other mutants described in this paper affords the proteins in complex with a mixture of nucleotide triphosphates as well as some other low molecular weight phosphoric acid esters. In a typical preparation of the His112Ser mutant, the amounts of GTP and ATP present were estimated by HPLC to yield about 17% GTP and 23% ATP.

In light of this unexpected finding, all crystals were soaked with GTP prior to data collection in the following experiments. The Cys181Ser and His112Ser mutants could be crystallized under the experimental conditions used for the wild type enzyme, albeit at an increased concentration of precipitant. Quality crystals were obtained at room temperature on a sitting drop vapour diffusion set-up, by mixing equal amounts of 0.1M Mops (pH 7.0), 10% (w/v) polyethylene glycol 6000 and 0.1 M ammonium sulphate solution (crystallization solution) and 6 mg/ml of purified protein solution under 20 mM tris-HCl

buffer (pH 8.0). The His113Ser mutant enzyme was also crystallized on a sitting drop vapour diffusion set-up, again at room temperature but under new chemical conditions, namely, 0.1 M Tris hydrochloride (pH 8.5), containing 0.2 M ammonium dihydrogen phosphate and 50 % MPD, to which and equal amount of 6 mg/ml purified protein solution was added.

Complete data sets to resolutions of 2.5, 2.8 and 3.2 Å were collected from single crystals of His113Ser, His112Ser and Cys181Ser mutant protein that had been soaked for 30 min with 10 mM GTP. The His113Ser and His112Ser mutant crystal diffraction sets were measured at a temperature of 95 K using synchrotron radiation. The diffraction data set of the Cys181Ser mutant was obtained at a temperature of 277 K with a Cu rotating anode x-ray source. Structures were determined using model phases of the wild type enzyme (Nar *et al.*, 1995b) and refined to their limit resolutions (2.5, 2.8 and 3.2 Å) with R-factors of 26.2 %, 21.8 % and 18.5 % respectively (Table 1 in Materials and Methods).

Figure 8. Amino-acid sequence and secundary structure from the wild type GTP-Cyclohydrolase I from *E. coli*. Red – single site mutation with reduced catalytic activity; Green – conserved residues, strictly conserved with additional box; Yellow ▲ - Residue closer than 6 Å to the GTP; Green ▲ – Residue closer than 4 Å to the GTP.

A superposition of the structure of His112Ser mutant in complex with the nucleoside triphosphate mixture at 2.1 Å resolution with the wild type structure reveals no significant structural changes; r.m.s. deviations were less than 0.4 Å for all monomers. However, based on this first high-resolution structure of GTP cyclohydrolase I, the secondary structure of the protein can be described in more detail. Two β-strands, β1A and β2B, and two α-helices, α5B and α6A, were additionally assigned (Nar *et al.*, 1995b).

The analysis of the main chain atom temperature factors of the wild type and mutant structures indicated that the most flexible region of the monomer is the loop between Gly 18 and Asp 31, connecting the α-helices α1 and α2. High main chain B-factors were also observed at the chain termini. On the other hand, the active site is within a rather rigid region of the GTP cyclohydrolase I structure including the four-stranded antiparallel β-sheet.

Amino acid sequences have been reported for GTP cyclohydrolase I from a variety of prokaryotic and eukaryotic organisms (Katzenmeier *et al.*, 1991) (Maier *et al.*, 1995). The evolution of the C-terminal domain containing the regions relevant for oligomerisation and enzyme catalysis appears to have proceeded in a relatively conservative way. With two exeptions, Lys68 and Ile132, every residue within a distance of 4 Å around the bound GTP is conserved (Figure 8). Three adjacent subunits participate in each respective catalytic site. Specifically, ^{D}Arg 65 (superscripts A to D mark the subunits A to D), ^{A}Cys110, ^{A}His 112, ^{A}His 113, ^{B}Tyr 125 (contact distance 6 Å, via water molecule W2_A), ^{B}Ser 135, ^{B}Lys 136, ^{B}Arg 139, ^{A}Gln 151, ^{A}Glu 152, ^{A}His 179, ^{A}Cys 181, ^{A}Arg 185 and ^{A}Thr 196 are involved either in substrate binding and/or catalytic mechanism. Correlating the conserved sequence regions with the three-dimensional structure reveals

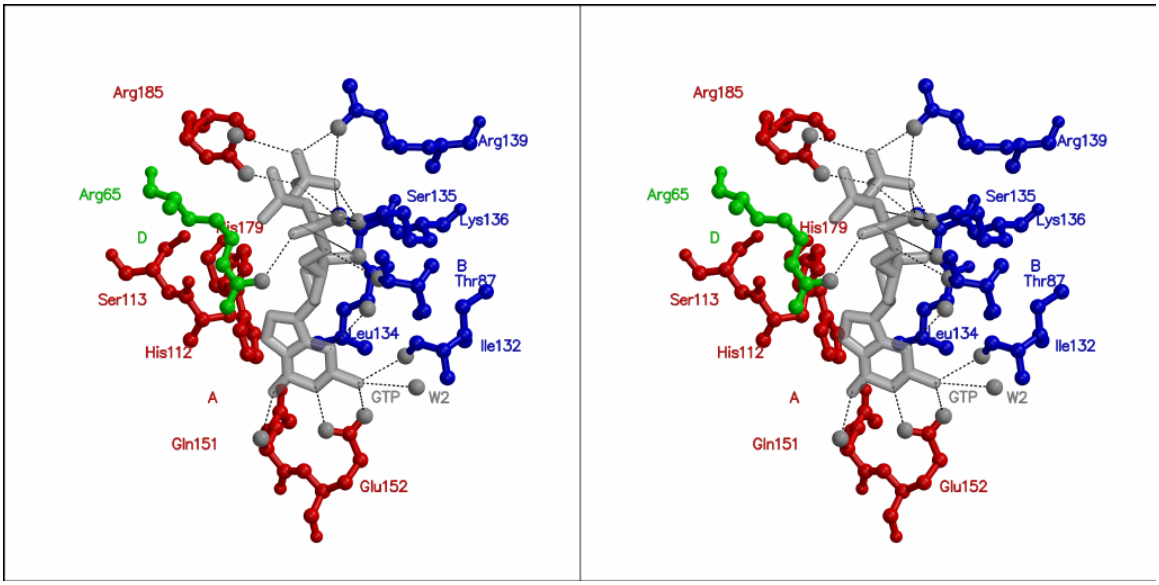


Figure 9. Stereo diagram from the active site of the *E. coli* GTP cyclohydrolase I His113Ser mutant in complex with the substrate GTP. The GTP molecule (shown as transparent wire model representation) is embedded in a large hydrogen bond network (dashed lines) within the active site. Amino acids are shown as ball-and-stick models coloured accorded to the subunit they belong, A – red, B – blue and D – green. The figure

was created using MolScript (Kraulis, 1991) and Raster3D (Merrit & Murphy, 1994).

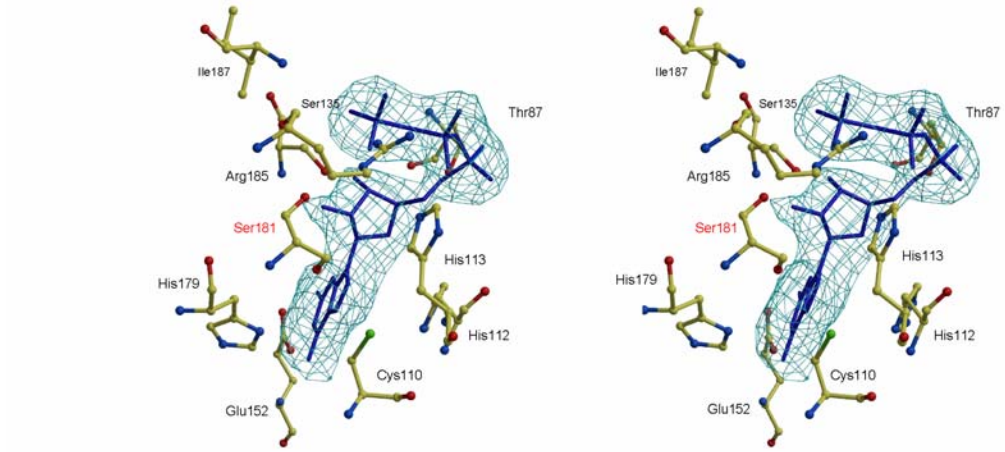
that the region from ^{D}Ser 58 to ^{D}Met 69 narrows the entrance to the active site pocket. The loop between the β-strands β1 and β2, ^{A}Cys 110 to ^{A}Val 115 forms the active site loop and the conserved N-terminal region of helix 5, ^{A}Arg 148 to ^{A}Thr 155, provides the purine anchor Glu 152. The sequence part ^{B}Val 131 to ^{B}Arg 139 sides both the ribose and the triphosphate, ^{A}Ile 175 to ^{A}His 179 the nucleobase and ^{A}Cys 181 to ^{A}Arg 188 the triphosphate moiety (Figure 9).

III.A.Results.3 Substrate (GTP) contacts with the protein matrix

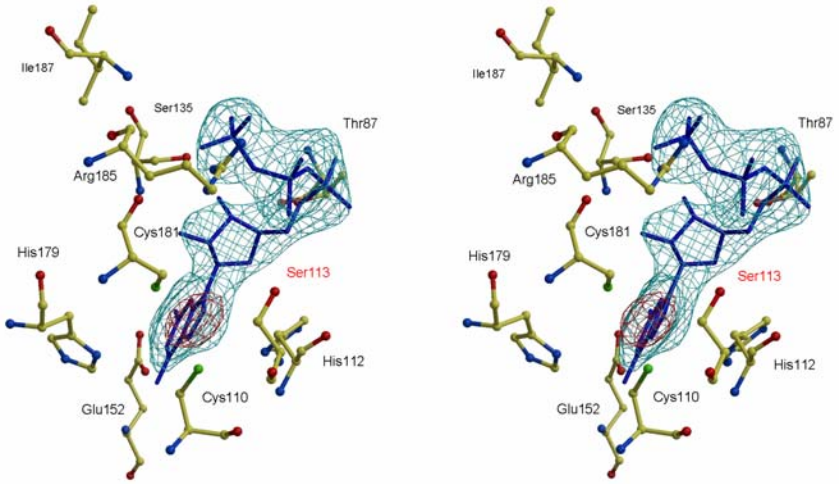
Parts of the electron densities of the mutant proteins under study can be unequivocally attributed to bound substrate. The well-defined F_{obs}–F_{calc} electron difference density maps clearly show the intact GTP molecule (Figure 10), thus confirming the complete catalytic inactivity of the mutants under study. In all structures, the GTP molecule is bound to ^{A}Glu 152 via the ring N1 nitrogen atom and the position 2 amino group of the 2-amino-4-oxopyrimidine moiety of the substrate, thus confirming the binding mode found for the 2'-deoxy GTP complex (Bracher *et al.*, 1998). Supporting this anchor, the amide hydrogen of the preceding peptide bond between ^{A}Val 150 and ^{A}Gln 151 forms a hydrogen bond to the 6-oxo group of the purine (Figure 9).

Furthermore, a hydrogen bond is formed between the position 2 amino group of the guanine moiety and the carbonyl oxygen of residue ^{B}Ile 132. Whereas the peptide bond of ^{A}Gln 151 is involved in anchoring the guanidino moiety of the substrate, the sidechain of ^{A}Gln 151 is hydrogen bonded to both ^{A}His 179 and ^{A}Thr 196. The anchor residue, ^{A}Glu 152, is located at the strictly conserved N-terminal region of helix α5.

It appears plausible that this hydrogen bond system remains in place throughout a substantial part of the reaction sequence, where it would initially contribute to the formation of the Michaelis complex and subsequently continue to serve as an anchor for the reaction intermediates (Figure 9).



a)



b)

Figure 10. Stereo diagrams of the F_{obs}-F_{calc} electron density maps immediately after rigid body refinement. Wired molecular model of GTP (dark blue) is superimposed on the GTP assignable electron density maps. Protein amino-acid residues are represented using ball-and-stick models.

(a) Mutant Cys181Ser with 2.0 σ contour level F_{obs}-F_{calc} electron density maps, coloured light blue. (b) Mutant His113Ser model with mutation. The F_{obs}-F_{calc} electron density map contoured at 2.5 σ level is coloured as light blue. Additionally red coloured F_{obs}-F_{calc} electron density map contoured at 4.5 σ level shows the presence of an undetermined ligand occupying the wild type zinc position (see also figure 9). Picture was elaborated using BobScript (Esnouf, 1997) and Raster 3D (Merrit & Murphy, 1994).

While the guanidine moiety of GTP binds to the negatively charged part of the active site region, the triphosphate moiety is bound by a cluster of basic residues provided by the

helices $^{D}\alpha 3$, $^{B}\alpha 4$ and $^{A}\alpha 5B$ (Figure 8). The charge compensation by these residues may be the reason why metal-assisted binding of the triphosphate group is not employed in GTP cyclohydrolase I. The strands $^{A}\beta 2B$ and $^{A}\beta 3$, which are N-terminally adjacent to these α helices ($^{B}\alpha 4$ and $^{A}\alpha 5B$ respectively), form the widest part of the active site pocket by lining both the base and the ribose of the GTP molecule. The enzyme substrate interactions observed in the structure are summarised in Table 8.

Guanine base (^{A}GTP)	**Protein**	**Triphosphate (^{A}GTP)**	**Protein**
N1	^{A}Glu152 Oε1	O1A	^{B}Lys136 Nz
N2	^{A}Glu152 Oε2	O2A	^{D}Arg65 NH1
	^{B}Ile132 O	O1G	^{B}Arg139 NH1
O6	^{A}Gln151 N		^{A}Arg185 NH1
N3	^{B}Leu134 N	O2G	^{B}Arg139 NH1
Ribose (^{A}GTP)	**Protein**		^{B}Ser135 Oγ
			^{B}Lys136 Nz
O2’	^{B}Ser135 N	O3G	^{A}Arg185 NH2
O3’	^{B}Ser135 Oγ		^{B}Ser135 Oγ
O5’	^{B}Thr87 Oγ1		

Table 8: Hydrogen bonds established between the active site residues and the bound GTP substrate molecule, as observed on the crystallographic structure of the GTP-cyclohydrolase I His113Ser mutant. Superscript marks the molecular subunits.

In the difference electron density maps of the three complex mutant protein structures, the orientation of the hydroxyl groups in the GTP carbohydrate unit is clearly defined and thus the conformation of the ribose moiety can be unambiguously assigned to C-2' *endo.* The 2' and 3' hydroxy groups of GTP form hydrogen bonds to the ^{B}Ser 135 main chain

amino group and to ^{B}Ser 135 Oγ respectively. ^{B}Ser 135 is furthermore hydrogen-bonded to the γ-phosphate of GTP. Additionally, a nearby water molecule, W1_A (His113Ser), connects the 2'-hydroxyl group to the γ-phosphate.

The triphosphate chain of GTP was found in a bent conformation, which is almost identical to the conformation in the deoxy-GTP complex of the zinc-depleted wild type enzyme (Nar *et al.*, 1995a). Small GTP conformation differences were found in different mutants in the phosphate Pβ to O3 region, a solvent exposed region where no fixed hydrogen bonds were assigned. The position of the γ-phosphate is fixed by interactions with the surrounding side chains of residues ^{B}Ser135, Lys^{B}136, Arg^{B}139 and Arg^{B}185 and virtually coincides with a bound inorganic phosphate in the wild type structure.

The resolution of the His113Ser mutant structure allowed the location of the active site water molecules and the study of their interactions with substrate and enzyme. A water molecule, W2_A, extends the hydrogen bond network at the guanidino moiety, thus contributing to the tight anchoring of the substrate. This is in line with the strict conservation of the residues Tyr125 and Glu152, which interact with this water molecule. As already suggested by the complex structure with the substrate analog, deoxyGTP (Nar *et al.*, 1995a), there is some free space on both sides of the imidazole portion of the nucleobase that might be occupied by solvent during catalysis. The water W3_A is held in position by hydrogen bonds to Gln149, Glu111, and His112 and lies in close proximity to the nucleotide base N7. The water molecules W1_A, W4_A, and W5_A are hydrogen-bonded to the γ-phosphate. No water molecules were found around the exposed α-phosphate and β-phosphate.

III.A.Results.4 Mutants configuration/conformation description

Mutation analysis had indicated earlier that the amino acid residues Cys 110 and Cys 181 are both indispensable for catalytic activity. The active wild type enzyme has both these amino acid residues bound to the essential zinc ion (Nar *et al.*, 1995b)(Auerbach *et al.*, 2000). Replacement of either one of these by any of several amino acid residues yielded completely inactive protein. Moreover, no zinc was detected by atomic absorption spectromerty in both these mutants (table 7). Finally, the inspection of the electron density maps for the present zinc free Cys181Ser mutant model confirmed the complete absence of zinc in this mutant, as no residual electron density was observed on the F_{obs}-F_{calc} density maps (figure10(a)).

Residues His 112 and His 113 have also been shown to be of major importance for the mechanism of this enzyme. The mutants His112Ser and His113Ser are devoid of catalytic activity (table 7). The zinc free crystal structure models of His112Ser and His113Ser, revealed important positive residual electron densities in the F_{obs}-F_{calc} electron density maps. These densities occur on the region where the zinc metal ion is observed in the active wild type form of the enzyme. Due to the fact that zinc was not detected by atomic absorption spectrometry (table 7), a fully occupied disulphide group formed between Cys 110 and Cys 181 on the His113Ser mutant model was calculated but could not explain this residual electronic density. With both His/Ser mutants, the position and coordination of the residual electron density led us to presume the presence of another metallic positive ion replacing zinc. As shown in figure 10(b) for the mutant His113Ser, the detected F_{obs}-F_{calc} electron density above 4.5σ cut-off level corresponding to the replaced metal is within the range of coordination of both Cys 110 and Cys 181 residues. For the mutant His112Ser, the weaker F_{obs}-F_{calc} residual electron density detected on some of the active sites of the a.u. (above 3.0σ) is shifted in the direction of His 179 by about 3 Å to establish coordination to the imidazole side chain via Nδ1.

In the structure of the wild type enzyme (Auerbach *et al.*, 2000) the imidazole sidechain of His112 forms a hydrogen bond to the zinc-bound water of approximate 2.6 Å, thus positioning it properly for nucleophilic attack on the substrate (Figure 11). A residue acting as a hydrogen bond acceptor for a metal-activated hydroxide is a common feature among zinc metalloenzymes (Christianson & Cox, 1999). In the Cys181Ser mutant the same orientation is observed for ^{A}His112. In the complex structure of deoxyGTP and Zn-depleted GTP cyclohydrolase I, the imidazole ring plane of His112 was found rotated by about 90 degrees in hydrogen-bond distance to the furanosyl ring oxygen (Nar *et al.*, 1995a). In the structure of the His112Ser mutant, ^{A}Ser112 forms a hydrogen bond with the side chain from the adjacent residue ^{A}His113, also engaged in GTP binding by acting as hydrogen bond donor to the β-phosphate.

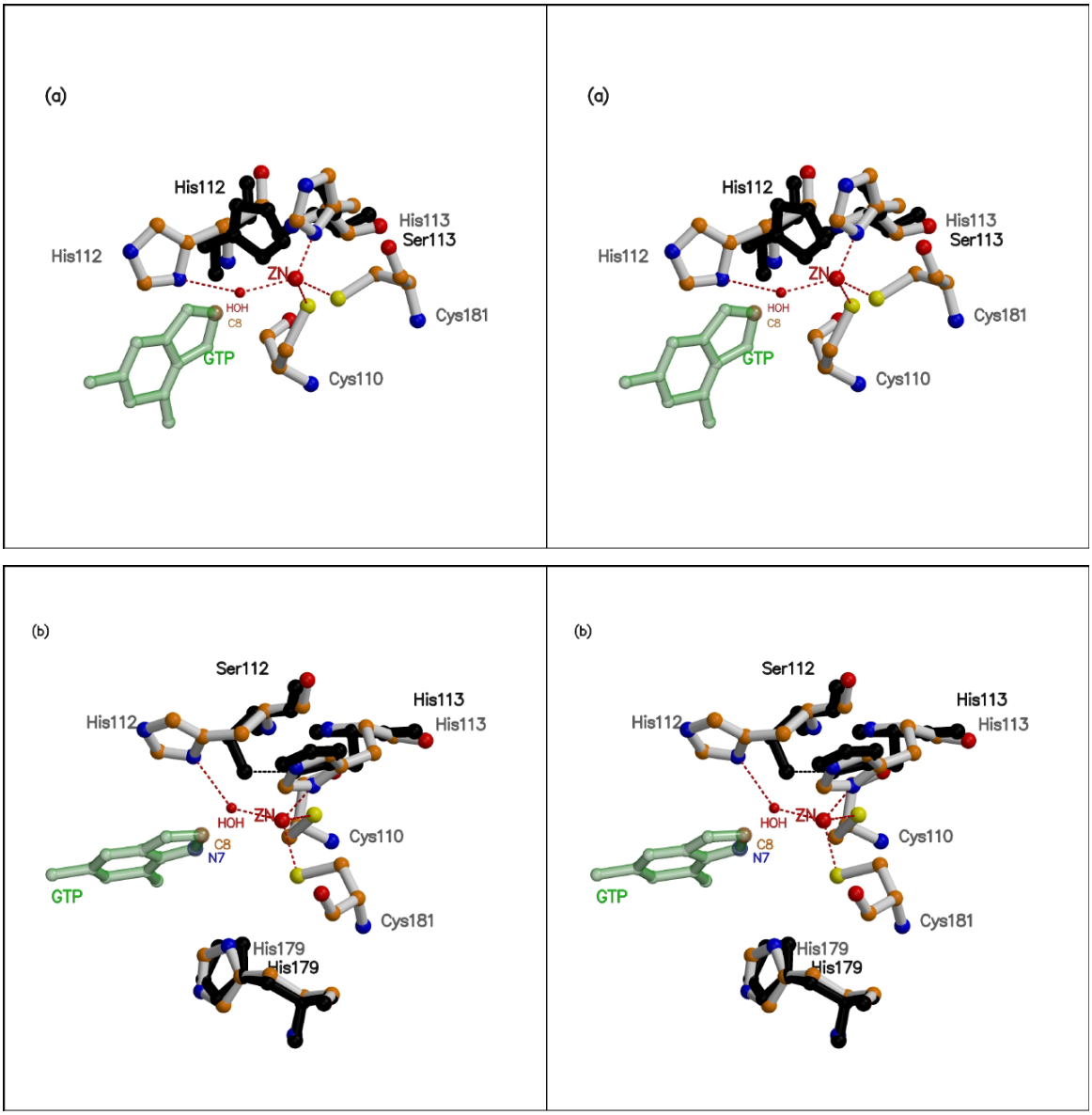


Figure 11. Orientation of the GTP substrate related to the zinc cofactor on the active site of GTP cyclohydrolase I. The peptide backbones from the wild type GTP cyclohydrolase I and from the GTP complex of His113ser (a) and His112Ser (b) mutants were aligned. The resulting locations of the His112, His113 side chains, of wild type zinc and the mechanistically important zinc-bound water molecule are superimposed on the binding pocket. The wild type residues are printed in grey. The zinc ion, the catalytic active water molecule and all the bonds are coloured red. Mutated structures and the hydrogen bonds within are coloured in black. Figure created using MolScript (Kraulis, 1991) and Raster3D (Merritt & Murphy, 1994).

When the structures of the zinc bound wild type enzyme and the His113Ser mutant are superimposed (Figure 11a), a short distance (below Van der Waals contact distance) is found between the zinc-bound water molecule and C-8 of GTP, suggesting a compressed

state favouring nucleophilic attack of the oxygen on C-8. The product of this initial reaction could assume a more relaxed conformation.

Another structurally characterized enzyme with a zinc cofactor bound by two cysteine residues, one histidine residue and one water molecule is cytidine deaminase, which catalyses the conversion of cytidine to uracil (Betts *et al.*, 1994) (Xiang *et al.*, 1995) (Xiang *et al.*, 1996). It has been demonstrated that cytidine deaminase binds transition state analogues formed *in situ* by hydration of 5-fluorozebularine and zebularine with very high affinity. In these complexes, the zinc-bound water molecule is replaced by the substrate analogue hydrate, a state resembling the transition state of the GTP imidazole ring opening step in the GTP cyclohydrolase I reaction.

III.A Discussion

III.A.Discussion.1 The reaction mechanism

GTP cyclohydrolase I catalyses a complex reaction sequence. The formamide intermediate, Compound **2**, has been confirmed by spectroscopic analysis (Bracher *et al.*, 1998). It has also been shown that the carbohydrate rearrangement involves the introduction of a proton from solvent into the pro-7*R* position of dihydroneopterin triphosphate (Bracher *et al.*, 1998).

Whereas the enzyme reaction had been known for more than three decades, the presence of an essential zinc ion at the active site of the enzyme had escaped detection. Moreover, auto-oxidation of the zinc-depleted *E. coli* enzyme had resulted in the formation of an artifactual disulfide bond between cysteine residues 110 and 181, which were no longer engaged in zinc chelation (Nar *et al.*, 1995b). Disulfide bond formation in the absence of zinc was also reported for cytidine deaminase, whose metal coordination resembles GTP cyclohydrolase I (Mejlhede & Neuhard, 2000). The presence of the metal cation at the active site was ultimately discovered during the crystallographic analysis of the human protein, which had been purified without the use of chelating agents (Auerbach *et al.*, 2000).

We have now shown that the replacement of Cys 110, Cys 181, His 112 or His 113 by serine affords mutant proteins with reduced metal binding capability. Each of the mutants was unable to convert GTP to dihydroneopterin 3'-triphosphate, even after addition of $ZnCl_2$. Notably, each of these mutant proteins contained nucleotide triphosphate GTP that was not removed during purification. These findings leave no doubt that all four amino acid residues are required for the chelation of active zinc with significant affinity. However, Nε of His 112 is 4.5 Å apart from the zinc ion and is unlikely to interact directly with the metal ion. The loss of activity in this mutant may result from a lack of fixation of the zinc-bound water molecule by its short hydrogen bond to His112 and in consequence lead to the loss of the reactive ligand of the zinc. On the other side, the hydrogen bond from the His113 imidazole ring to the mutated His112Ser effectively prevents zinc coordination (Figure 11b). On the His113Ser mutant, the absence of His 113 imidazole ring allows an alternate side-chain conformation for residue His 112 (Figure 11a).

The mutants under study are also unable to catalyse the release of formate from GTP and from the formamide type intermediate, Compound **2**. It follows that a correctly positioned

zinc ion is not only required for the opening of the imidazole ring but also for the subsequent hydrolysis of the formamide motif. This reaction is highly reminiscent of the hydrolysis of peptide bonds by zinc proteases (Christianson & Lipscomb, 1988) (Lipscomb & Sträter, 1996).

The data suggest that zinc acts as a Lewis acid activating water molecules (i) for the hydration of the imidazole ring of GTP, which precedes the ring opening and (ii) for the hydration of the formyl group of Compound **2**. The residue His 112 appears to be required to properly orientate the zinc-activated water for nucleophilic attack on the substrate and to accept the proton released in this step (Figure 12). As corroborated by the orientation of His 112 in the complex structure of zinc-depleted GTP cyclohydrolase I with deoxy-GTP, the resulting imidazolium group of His 112 could rotate and transfer the proton to the ribose ring oxygen of the GTP hydrate. This would facilitate the opening of the imidazole ring in the substrate by intermediary formation of a Schiff base as previously proposed (Nar *et al.*, 1995a). In analogy to the mechanism of zinc proteases, the coordination number of zinc could then be increased to five through complexation of an additional water molecule which would attack the zinc-complexed formyl group of the intermediate **9** conducive to a tetrahedral transition state (intermediate **10**) for the abstraction of formate (Hangauer *et al.*, 1984) (Christianson *et al.*, 1987) (Holden & Matthews, 1988). The imidazole sidechain of the adjacent residue His 179 could support the cleavage of the formamide bond by acid catalysis (Figure 12). This role is supported by the reversible conversion of GTP to 2-amino-5-formylamino-6-(ribosylamino)-4(3*H*)-pyrimidinone 5'-triphosphate catalysed by His 179 mutants of GTP cyclohydrolase I.

Single turnover kinetic studies showed that the velocity of the ring opening reaction exceeds that of product formation by approximately one order of magnitude (Schrameck *et al.*, 2001a) (Bracher *et al.*, 2001) (Schrameck *et al.*, 2001b). More recent studies suggest that the hydrolysis of the formamide bond is even more rapid than the ring opening step (Schrameck *et al.*, 2002) (Schrameck *et al.*, 2001b). Thus, the hydrolytic reactions involving zinc ion catalysis are rapid by comparison with the isomerisation of the carbohydrate moiety.

Comparison of the zinc-containing and zinc-free wild type enzymes and the zinc-deficient mutants indicates that the topology of the active site is relatively rigid and apparently independent of possible thiol oxidation or metallic coordination. This view is also supported by the crystallographic temperature factors, which are low for the active site area. The structural comparison also suggests that the conformation of the substrate GTP

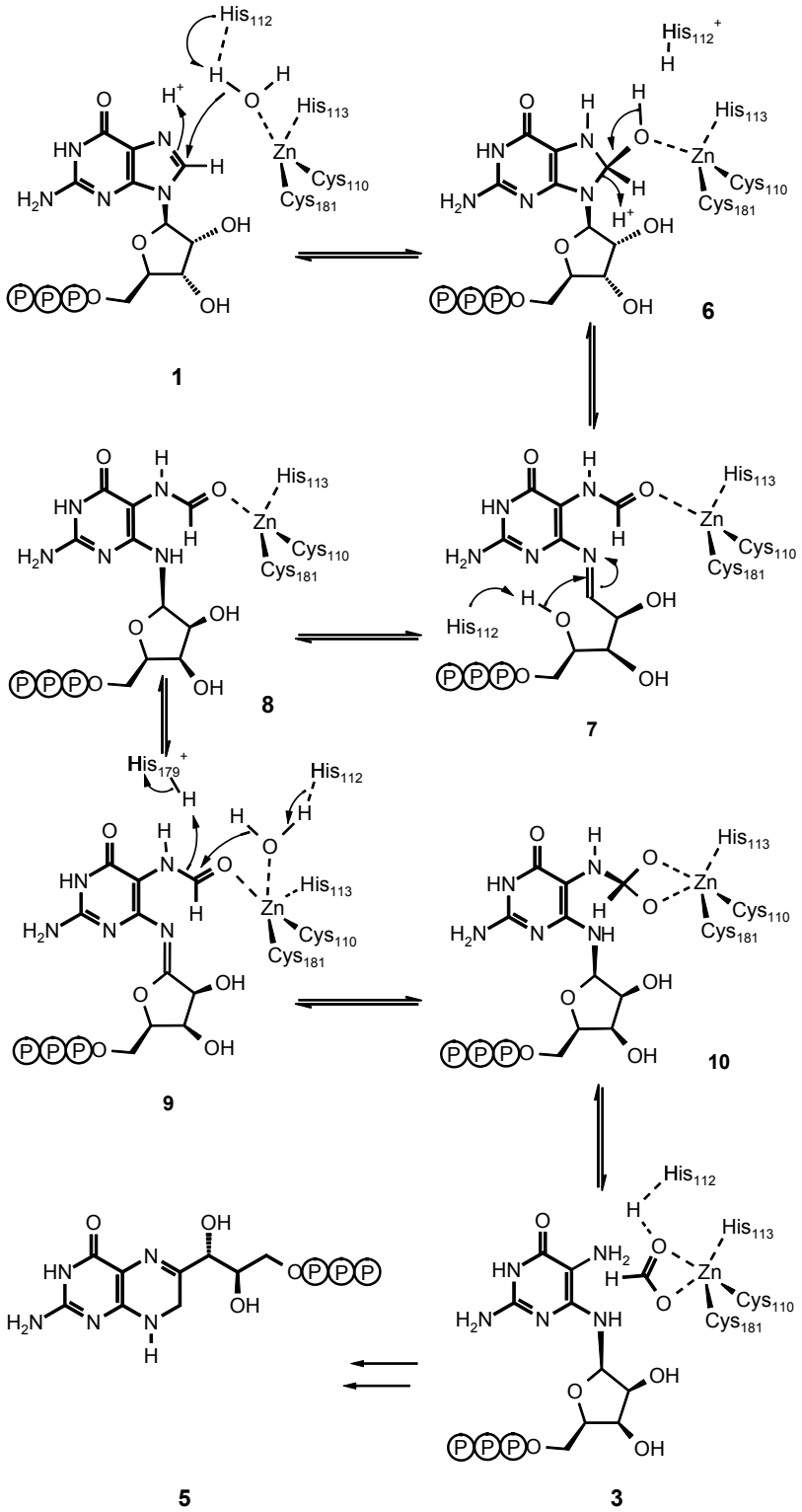


Figure 12. Detail from the hypothetical reaction mechanism for GTP cyclohydrolase I. This mechanism comprises the following steps: Nucleophilic attack by a zinc complexed water molecule to the position C8 from the imidazole ring of the nucleotide base (from 1 to 6); Hydrolitic opening of the imidazole ring (from 6 to 8) with transient formation of a Schiff base (7); Hydration from the formyl group of 2-amino-5-formylamino-6-(ribosyl-amino)-4(3H)-pyrimidone 5'-triphosphate (from 8 to 9); Formamide hydrated tetrahedral state supported by a penta-coordinated zinc (10); Release of formate (from 10 to 3); Amadori rearrangement of the ribose ring and subsequent closure of the dihydropyrazine ring by intramolecular condensation (from 3 to 5).

is substantially constrained by the presence of the zinc ion. A putative deformation of the guanine moiety could further assist the water addition mediated by the zinc ion. The substrate GTP is most probably fixed on both ends during the enzymatic reaction. It is likely to be flexible only on the phosphate Pα, ribose and imidazole ring regions.

III - B/C/D - Molybdopterin containing enzymes belonging to the Xanthine oxidase family (Molybdenum hydroxylases).

B. The Aldehyde Oxidoreductase from *Desulfovibrio gigas*.

III.B Results

III.B.Results.1 Crystallisation, data collection and refinement.

One hexagonal bipyramid crystal with the longest axis measuring almost 1 mm was obtained following the procedure described by Romão et al. (1993). This crystal was frozen at 100 K on a loop embedded on crystallization based solution containing PEG4000 (harvesting solution), and was used to obtain a data set almost 100% complete up to 1.28 Å of resolution, on the BW6 X-ray beamline from the DESY synchrotron. An X-PLOR (Brünger 1990) formatted data file containing 233755 unique reflections was obtained after data processing with Denzo/Scalepack (Otwinowski & Minor 1997) and several software utilities belonging to the CCP4 software package (CCP4, 1994) as can be seen on table 2 (c.f. Methods).

With a 3% randomly chosen subset of reflections selected (7004 reflections) for validation (R_{free}), the single polypeptide chain from the previous model refined to a limit resolution of 1.8 Å was used for rigid body refinement including only the observations up to this resolution. After the Molecular Replacement / Rigid Body Refinement step has been performed by XPLOR, the computed R and R_{fre} values where 35.2% and 39.5% respectively. The following inclusion of the cofactors, the introduction of solvent modulation among other corrections and extending of the resolution of the observations to its higher value (1.28 Å), resulted on quality indicators R and R_{free} of 20.7 and 24.4 respectively.

The resolution of 1.28 Å is high enough to permit an anisotropic refinement of the individual atomic temperature factors (B factors). On the present case the ratio between the total number of unique reflections to the number of parameters for a non isotropic single atom model (x,y,z,B) was 7.2. When including the bond and angle restrains, this ratio jumps to 9.9, justifying the extension of the refinement to an anisotropic temperature modulation for single atoms (x,y,z,U_{ij}). This refinement was then performed under SHELXL (Sheldrik & Schneider, 1997) achieving a final quality of refinement, R- factor

under 15% at 4σ cutoff, showing a good agreement with the expected for this level of resolution (Wilson et *al.* 1998). Further details on refinement progression are described on table 3 under Methods.

III.B.Results.2 Overall quality of the final model

The quality of the model now presented was analysed with PROCHECK (Laskowski et al., 1993) yielding an overall structure quality measure g-factor of –0.11, close to the mean g-factor value for structures of the same nominal resolution (Wilson et *al.* 1998). The Ramachandran plot has 91.6 % of the residues in the most favoured regions. Only two residues other than glycine or proline are on disallowed positions, Tyr142 (Φ=69º Ψ=-37º), Leu254 (Φ=62º Ψ=-68º). Both residues are very clear in density. Tyr 142 is a conserved residue in the xanthine oxidase family and is located on the FeS_I domain on a loop that connects helix III to helix IV. Its hydroxyl group establishes hydrogen bond interactions with the amide from Cys137, a ligand from one of the iron atoms, and to the carbonyl group of Pro 839, thus defining the loop conformation of helix III to helix IV. Leu254, the second outlier, is also very clear in density and is part of the hydrophobic patch, which surrounds the entrance to the tunnel leading to the active site.

The final r.m.s.d. values for bond distances and angle distances are 0.013 Å and 0.029 Å respectively (Table 4) (c.f. Materials and Methods). All other stereochemical parameters analysed are within the expected standard uncertainty ranges for a set of very high-resolution (0.92 to 1.20 Å) structures available (Wilson et *al.* 1998), except for few outliers that occur on low ordered regions of the molecule.

The final bonds and angles error estimates were calculated from a LS (least squares) blocked matrix inversion for the cofactors. In this blocked matrix least-square minimization only the positional parameters included in the defined block were refined. The block contained exclusively atoms belonging to the cofactors, while all protein atoms were held at fixed refined positions.

These estimated bond distance error values are summarized on Table 4 (c.f. Materials and Methods) and on Figures 16, 17 and 18 for all bonds within the cofactor molecules. They are one order of magnitude higher than usual for small molecules (Wilson et *al.* 1998).

III.B.Results.3 Fourier termination effects on the Molybdenum site

The quality of the final model may be restricted by Fourier ripples at resolutions above 1Å. This may shift the positions of atoms on the vicinity of strongly scattering atoms, usually metal ions, which is of particular interest in metal-proteins (Schindelin *et* al., 1997). Possible Fourier ripple effects were estimated mainly to determine changes in the Mo-H_2O/OH bond distances.

Starting with the final refined atomic model, two Fc maps where calculated on the vicinity of the Mo atom, with data truncated at 1.28 Å and 0.5 Å respectively. The 0.5 Å map is virtually free from the Fourier ripples and the comparison between the two maps allows the estimation of a first order correction to be applied on each one of the Mo ligands positions. The differences between the two maps are minimal and clearly within the experimental error, when calculated with the observed B_{eq} factor of 17.5 $Å^2$ for the Mo atom. Lowering the value for the B_{eq} of Mo to 5 $Å^2$ the effect of this artefact is enhanced, but even here calculated corrections were small (Mo-OH differs by 0.08 Å)

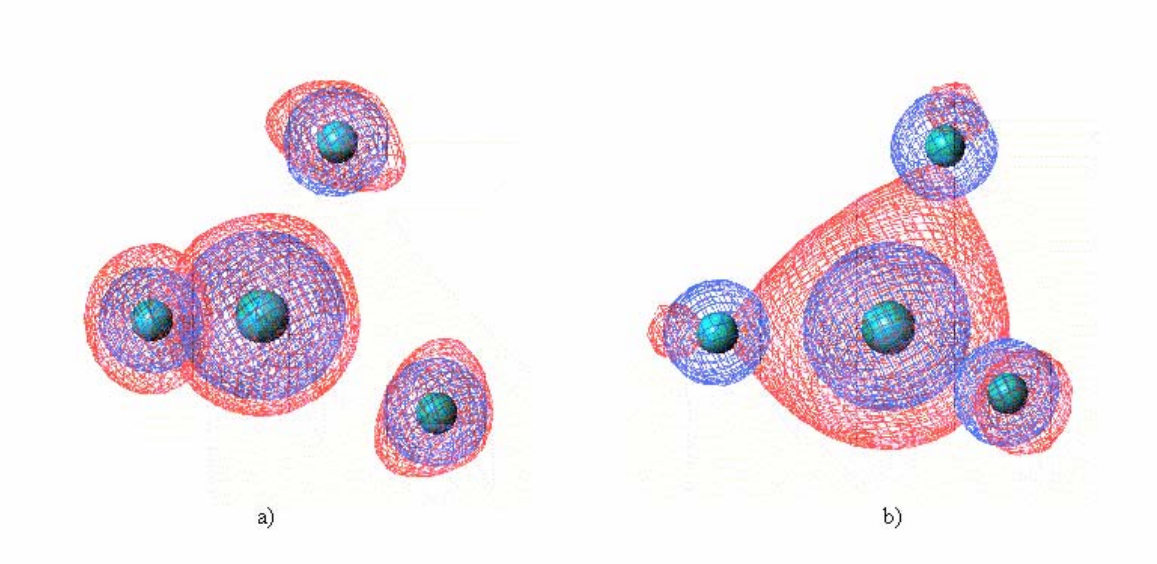


Figure 13. Calculated electron density maps at 2.0 σ level for the Mo –Oxo1, -Oxo2, -OH/H_2O, starting from the refined MOP model with a B_{eq} factor of 5 $Å^2$at resolutions: a) 0.5 (blue) and 1.28 Å (red). The OH/H_2O molecule has a new electron density maximum less than 0.1 Å from the refined maximum of the 1.28 Å map; b) 0.5 (blue) and 2.0 Å (red). Noticeable shifts on all the positions of all Mo ligands are observed.

(Figure 13). At 1.1 Å, the electron density maps can be considered free from these series termination errors (Dobbek and Huber, 2002), at least for a scattering centre consisting of

a single Molybdenum atom (atomic number: 42). For this reason the problem is immediately overruled for the two other metallic [2Fe-2S] centres (Fe atomic number: 26). Still and since we are dealing with synchrotron radiation, ionisation effects frequently occur, which may change the oxidation states of the enzymes. Particularly sensitive to these phenomena are metallic centres containing the Fe ion. Consequently and although working at atomic resolution, a careful interpretation of the information is required.

III.B Discussion

III.B.Discussion.1 Refined MOP model

MOP is a globular shaped protein with a 75 Å average diameter, with a single polypeptide chain of 907 amino acid residues folded as a four-domain structure. Each of the first two domains, named FeS_II (distal) (plant-type ferredoxin (Fd) fold) and FeS_I (proximal) (4-helical fold, new for Fd), enclose the two [2Fe-2S] clusters, while the larger following domains Mo 1 and Mo 2, anchor and stabilize the extended conformation of the pyranopterin cofactor (Romão *et* al. 1995) (Figure 14). The current

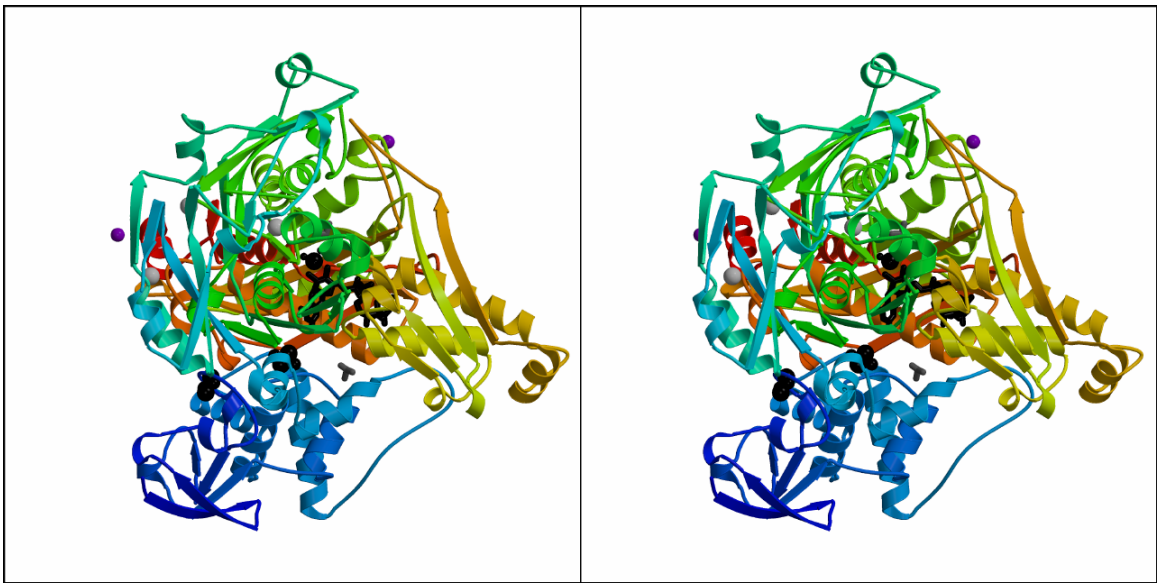


Figure 14 - Stereo view of the protein folding represented in a colour gradient from the N-terminus (dark blue) to the C-terminus (red). The [2Fe-2S] and MCD cofactors as well as the two isopropanol molecules are shown in black. Cl^- and Mg^{2+} ions are drawn as purple spheres, those with larger radius corresponding to Cl^- and those with smaller radius to Mg^{2+} (Picture build with Molscript (Kraulis 1991) and Raster3D (Merrit & Murphy 1994)).

model is very well ordered especially on hydrophobic and core regions of the molecule as indicated by an overall B-factor of 24.7 $Å^2$ and by an estimated average error on the co-ordinates of 0.10 Å (Luzzati, 1952). A few external and solvent exposed loops suffer from a variable amount of disorder**.** The exposed loop connecting domains FeS_I and

Mo1, between residues 168 and 172, is the most disordered region of the molecule and the site of insertion of the FAD domain present in the larger XO structure.

The r.m.s.d. between 904 C_α atoms of the two models (the former model refined at 1.8 Å (Huber *et* al. 1996) and the current one at 1.28 Å) is 0.26 Å. The improvement of the new model is reflected particularly in much clearer electron density maps for most of the molecule especially in poorly ordered surface residues. Alternate conformations for 17 side-chains and 1 for a main chain were clearly identified, having a shared associated occupancy ranging from 0.2-0.8 to 0.5-0.5 as described on Table 9. Conformational flexibility affects several exposed side-chains where the electron densities vanish to a variable degree.

Residue	Occupancy 1	Occupancy 2	Residue	Occupancy 1	Occupancy 2
Ile 2	0.64	0.36	Leu 410	0.53	0.47
Ala 18	0.52	0.48	Glu 561	0.78	0.22
Gln 44	0.26	0.74	Cys 661	0.74	0.26
Met 233	0.52	0.48	Asp 736	0.69	0.31
Ile 240	0.52	0.48	Thr 742	0.54	0.46
Ile 322	0.48	0.52	Asn 744	0.42	0.68
Asp 379	0.68	0.32	Val 834	0.49	0.51
Leu 396	0.56	0.44	Leu 872	0.63	0.37
			Lys 828-CO	0.58	0.42

Table 9. Relative occupancy of the alternate conformations.

Two isopropanol molecules from the crystallising solution are present on the model, IPP1 is in the second co-ordination sphere of the molybdenum acting as an inhibitor, in agreement with our earlier studies, which provided the basis for the enzymatic mechanism when taking a Michaelis complex as a model (Huber *et* al. 1996). A second isopropanol molecule, IPP2 is on the external surface of the protein molecule, on its first hydrating shell, and is H-bounded to water molecules.

III.B.Discussion.2 Mg^{2+} and Cl^- ions in the crystal packing

As described in the experimental part the salts $MgCl_2$ or $CaCl_2$ were essential for the crystallization of MOP and the same crystal form was always obtained either when using different combinations of alcohols or when slightly changing the pH.

At later stages of refinement, Cl^- and Mg^{2+} ions from the crystallisation solution were clearly identified on Fo-Fc electron density maps over 4 σ levels (2.43 e.$Å^{-3}$). As depicted on Figure 14, three Cl^- and two Mg^{2+} ions were included in the final model and, with the exception of Cl1, which is rather buried, all others are close to the molecular surface, some favouring and stabilizing crystal contacts.

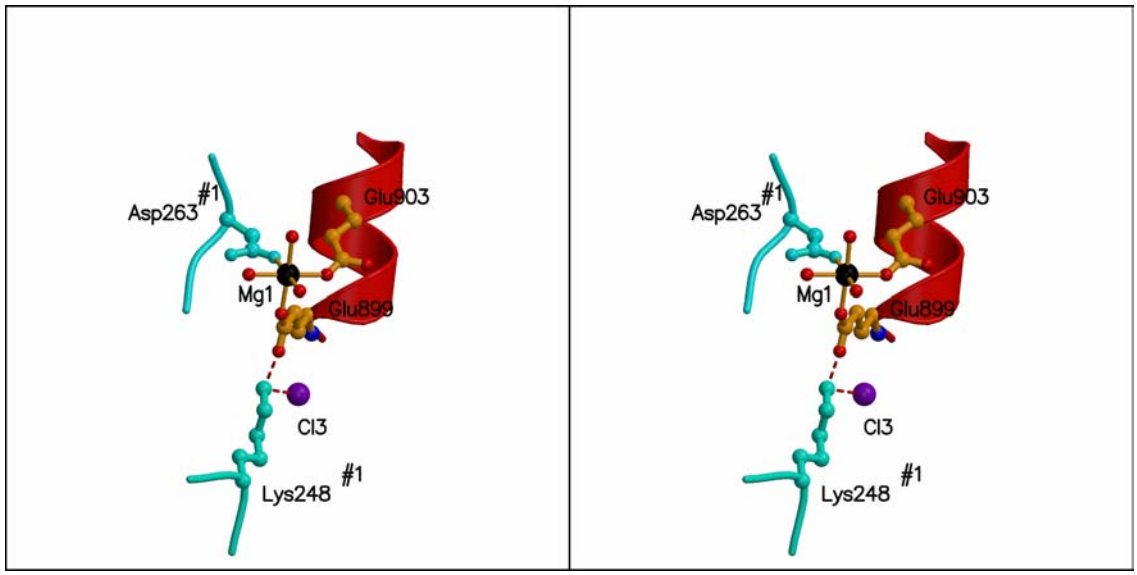


Figure 15 – Stereo representation of the crystal packing contacts around one of the Mg^{2+} ions, Mg1 (black). It shows perfect octahedral co-ordination geometry to 3 water molecules, to Glu899 Oε2 and Glu903 Oε1 from the C-terminus helix (red) and to Asp263 Oε1 from a symmetry mate #1 (blue). Glu899 Oε1 establishes a salt bridge to Lys248 Nε, which is 2.99 Å from Cl3 (Figure prepared with MolScript (Kraulis, 1991) and Raster3D (Merrit & Murphy 1994)).

One of the Cl^- ions, Cl1, is located in an internal pocket, filled with 4 water molecules, which form a water chain directed towards the isopropanol molecule of the active site. The distance from Cl1 to the Mo atom is 8.2 Å. It is stabilized by the vicinity of a positively charged residue, Arg 460 only 3.4 Å away. This Arg 460 is conserved in the

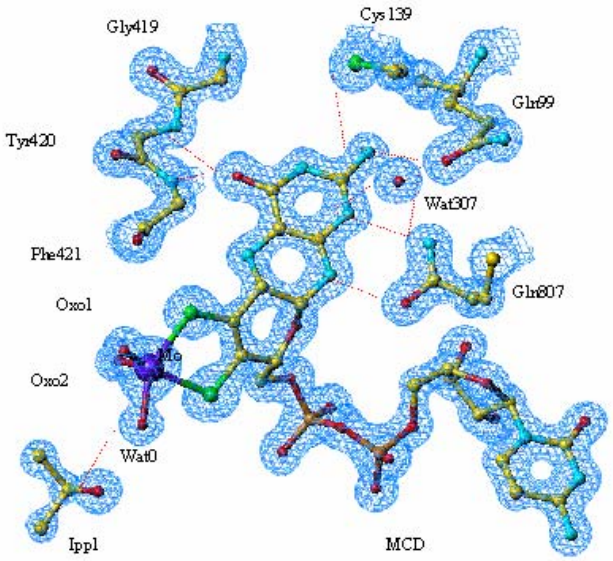
XO family. In the homologous enzyme MOD, where the crystallization conditions were different the Cl1 is replaced by a water molecule.

As depicted on Figure 15, in the first hydration shell, Mg1 and Cl3 make protein symmetry contacts that constitute a strong intermolecular contact and contribute to the molecular packing. Mg1 makes two ionic contacts to Glu899Oε2 and to Glu903Oε1 and to the symmetry related #1 Asp263Oε1. Three water molecules complete the coordination sphere of Mg1, which displays a perfect octahedral geometry. The carboxylate of Glu899 makes a salt bridge with Lys248#1 from a symmetry mate (the distance Glu 899 Oε1---Nε Lys248 is 2.74 Å). Chloride Cl3, which is only 2.99Å away from Nε of Lys248, completes a network of electrostatic interactions at the crystal contact possible due to the presence of Mg^{2+} and Cl^- ions.

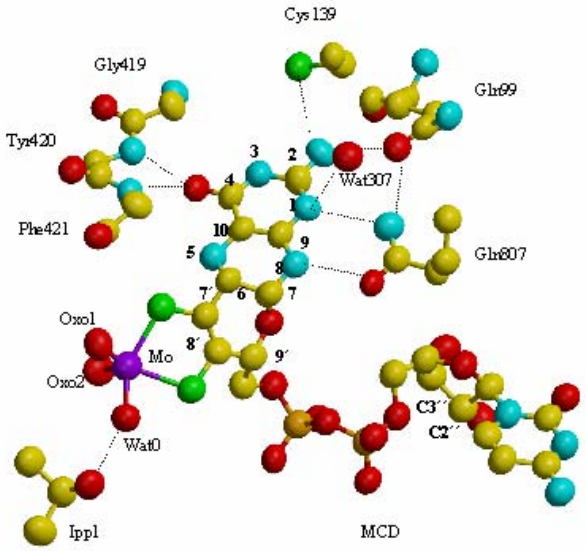
Other salt ions were recognised on the surface such as an additional Cl^- ion, Cl2 found near the C-terminus and relatively close to positively charged residues, Lys900 (5.4 Å away) and Arg 893 (7.1 Å away). A second Mg^{2+} ion is also located at the molecular surface, far from crystal contacts. It defines a perfect octahedron with a cluster of water molecules as ligands. This cluster is found on a rather hydrophilic cavity, filled with 15 well ordered water molecules which establish a network of hydrogen bonds among themselves and also to polar residues, which coat this exposed pocket, Asn511, Asp572, Arg569, Lys565 and Asp474.

III.B.Discussion.3 The Mo active site

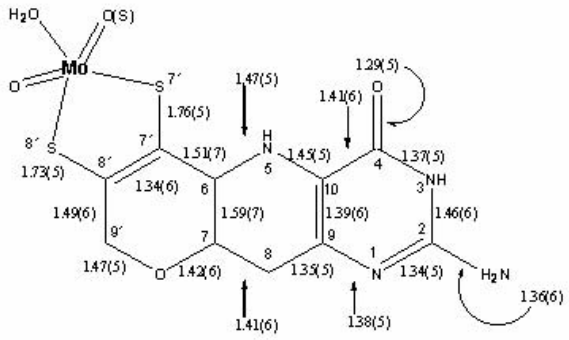
Mo is present in the oxidised form and is penta co-ordinated by two dithiolene sulphurs and three oxygen ligands. The cis-dithiolene group and the molybdenum atom form the equatorial plane of the metal ligands, with covalent bonds of 2.41(2) to S7′ and 2.50(2) to S8′. The molybdenum atom is further coordinated at a distance of 1.77(4) Å to the apical oxo ligand, 1.79(4) Å to the basal oxo ligand and 2.02(4) Å to the water molecule. These values confirm the 1.8 Å resolution results for the desulfo-oxidised form of the enzyme (Huber *et* al., 1996), interpreted with two oxo and one water ligand.



a)



b)



	Oxo1	Wat0	Oxo2	MCD-S7´	MCD-S8´
Mo	1.74(4)	2.00(4)	1.79(4)	2.41(2)	2.50(2)

c)

Figure 16 – Representation of the MCD cofactor and main contacts of the pyranopterin moiety with the surrounding protein residues.

a) The final atomic model with 1.6 σ (0.97.eÅ^{-3}) 2Fo-Fc electron density map superimposed.

b) Representation of the anisotropic atomic model, with the ellipsoids contoured to enclose a probability of 35%.

c) Schematic representation of Moco with bond distances including errors in parenthesis. (Figures a) and b) prepared with TURBO-FRODO (Cambillau *et al.*, 1997), MolScript (Kraulis, 1991) and Raster3D (Merrit & Murphy 1994)).

III.B.Discussion.4 The MCD cofactor

The chemical nature of the pterin cofactor had been suggested by Rajagopalan and co-workers (1991) with a pterin ring substituted at position 6 with a phosphorylated dihydropterin side chain containing a cis-dithiolene bond. This model was confirmed and refined with the crystal structures of the W AOR from *P. furiosus* (Chan *et al.*, 1996) and MOP (Romão *et* al., 1995). In MOP, as in almost all bacterial enzymes, the tricyclic pyranopterin cofactor occurs as a dinucleotide. The MCD cofactor is buried 7 Å deep inside the protein, at the interface between Mo1 and Mo2 with a 15 Å deep hydrophobic tunnel allowing the access of substrate molecules to the catalytic site.

The MCD is subdivided in three parts: The cytosine dinucleotide which has the pyrimidine base in anti conformation and the D-ribofuranose sugar twisted with C2´´ endo and C3´´ exo; The pyrophosphate which assumes different conformations in known enzyme structures; The tricyclic pyranopterin-ene-1,2-dithiolate, formed by a bicyclic pterin fused to a pyran ring, were the pterin system and pyran ring enclose an angle of about 28º. The pyrazine part of the system is twisted, with C6 exo and C7 endo being 0.4 Å and 0.3 Å apart from the reference plan respectively, and the pyran ring is in envelope conformation with O9´ endo.

The pyranopterin system is in the tetrahydropterin form. Protonation at C6 and C7 is in agreement with a tetrahedral geometry with bond angles ranging from 107º to 114º. Furthermore, the C6-C7 single bond is 1.59(6) Å, approximately 0.20 Å longer than the C9-C10 double bond. N5 and N8 are probably also protonated, as the N5-C6 and N5-C10 distances and the N8-C9 distance are equal or larger than the database derived target values for bonds with protonated N atoms (Figure 16).

The pterin ring is stabilised on one side by Phe421 and Tyr420 via the carbonyl O4 atom pair of hydrogen bonds, whilst on the other side H-bonds are made from the amino group NH_2 to the Sγ of Cys139 and to the carbonyl Oε of Gln99. The N1 edge is H-bonded to Nδ1 amine side-chain group of Gln807 and to a water molecule. The last H-bond interaction from the tricyclic pterin belongs to the pyrazine ring from the N8 edge to the Oε1 of Gln807 (Figure 16).

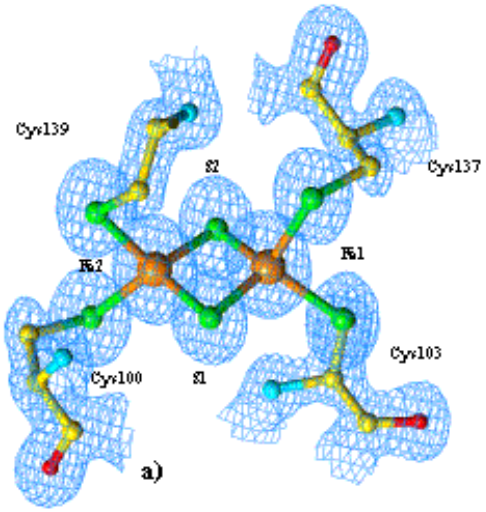
III.B.Discussion.5 The [2Fe-2S] clusters I and II

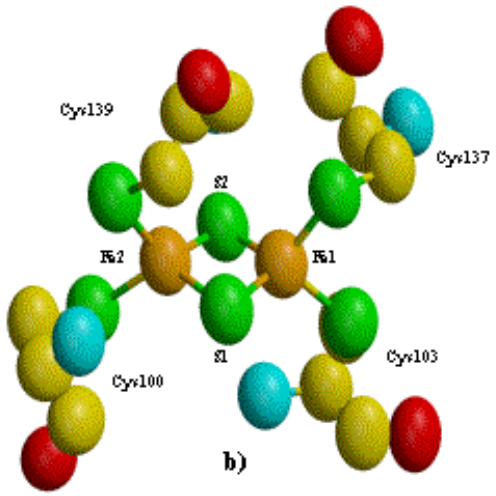
Two inorganic one electron redox [2Fe-2S] centres contact the MCD cofactor and their role consists in transferring electrons out of the buried MCD to the external surface, and from them to an electron acceptor. As the reaction mechanism consists of a coupled pair of reductive and oxidative half-reactions, characterized by reduction of Mo(VI) and oxidation of Mo(IV) (Huber *et* al., 1996), molybdenum has to be restored to the resting Mo(VI) state through a two step sequence of one-electron transfers by the iron-sulphur clusters, generating the intermediate Mo(V) state.

The Fe/S centre proximal to Mo (centre I) (4-helical fold domain), is buried ca. 15 Å from the nearest surface with a distance of 15 Å to the active site (Fe1…Mo). The distal Fe/S cluster (centre II), has a typical plant-type Fd coordination motive, is near the molecular surface, and 12.2 Å (FeS-I, Fe2---Fe1, FeS-II) from cluster I. The clusters were recently assigned for MOP (Caldeira *et al.*, 2000) as well as for the homologous enzyme from *D. alaskensis* (Andrade *et al.*, 2000) on the basis of the anisotropy of their EPR signals (**g**-tensor). Accordingly, the centre of type I with the lower anisotropy is the proximal Fe/S centre, while the distal, plant-type Fe/S cluster, with a higher anisotropy is FeS_II. The redox potentials are –260mV and -285mV for centers I and II respectively (Moura *et* al., 1978).

Both [2Fe-2S] clusters are depicted in Figures 17 and 18, superimposed to the final electron density map, contoured at 1.8 σ, as well as the anisotropy model. The FeS_I cluster is surrounded by seven N-H…S hydrogen bonds to cysteinyl S^{γ} and to inorganic S atoms, and one N-H…O to a water molecule, while cluster II establishes a total of nine N-H…S bonds plus one N-H…O to an internal water molecule (see Table 10). In both clusters the pattern of N-H…S^{γ}Cys bonds is rather asymmetrical, which may explain the differences in the Fe-S^{γ} bond lengths (2.27(1)≤Fe-S^{γ} ≤2.37(1)). Cluster Fe/S_II, with a plant-type Fd fold exhibits an NH-S hydrogen bonding pattern very similar to the one found in the [2Fe-2S] cluster from plant-type Fds, although only for one of the Fe atoms. For example, FdI from the blue alga *Aphanothece sacrum* (Tsukihara *et* al., 1990), the pattern around the Fe atom bound to Cys 39 and Cys 44 (Cys 40 and Cys 45 in MOP) is identical to what is found around Fe2 in MOP (Table 10). The only exception is the interaction with the O^{γ} of Ser 38 in *A. sacrum* FdI, which in MOP is replaced by a NH-O bond to a water molecule. In FdI, the second iron atom has no NH-S bonds in its environment, while in MOP, Fe2 cysteinyl sulphur atoms make 2 NH-S bonds.

[2Fe-2S] I

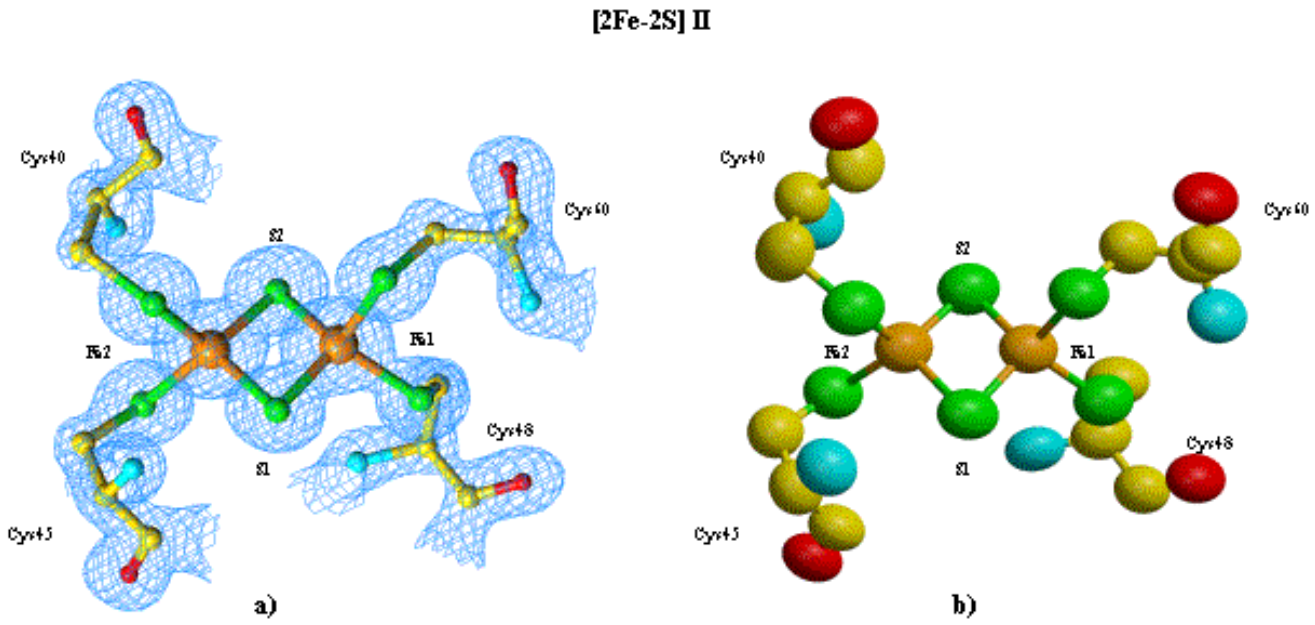




	Cys103-S$^{\gamma}$	Cys137-S$^{\gamma}$	S1	S2		Cys100-S$^{\gamma}$	Cys139-S$^{\gamma}$	S1	S2
Fe1	2.29(1)	2.36(1)	2.23(2)	2.20(2)	**Fe2**	2.27(1)	2.33(1)	2.27(2)	2.24(2)

Figure 17 – The atomic model of the proximal FeS_I cluster, from the 4-helical domain and with the cysteinyl co-ordination Cys100-X_2-Cys103-X_n-Cys137-X_2-Cys139.

a) The final atomic model with the 1.6 σ (0.97 eÅ^{-3}) 2Fo-Fc electron density map superimposed. **b)** Anisotropic atomic model with the ellipsoids contoured for a 35% probability distribution. Bond distances include errors in parenthesis. (Figures prepared with TURBO-FRODO (Cambillau et *al.* 1997), MolScript (Kraulis, 1991) and Raster3D (Merritt & Murphy, 1994).



	Cys60-Sγ	Cys48-Sγ	S1	S2
Fe1	2.29(1)	2.33(1)	2.21(2)	2.20(2)

	Cys40-Sγ	Cys45-Sγ	S1	S2
Fe2	2.37(1)	2.31(1)	2.27(2)	2.26(2)

Figure 18. – The atomic model of the distal FeS_II cluster, with the plant type ferredoxin fold and co-ordination motif Cys40-X_4-Cys45-X_2-Cys48-X_n-Cys60.

a) The final atomic model with the 2Fo-Fc electron density maps drawn at 1.6 σ (0.97 eÅ^{-3}) superimposed.

b) Representation of the anisotropic atomic model, with the ellipsoids contoured to enclose a probability of 35%.

c) Schematic representation of Moco with bond distances including errors in parenthesis. (Figures a) and b) prepared with TURBO-FRODO (Cambillau et *al.* 1997), Molscript (Kraulis, 1991) and Raster3D (Merrit & Murphy, 1994)).

Acceptor	Donor	Distance (Å)
[2Fe-2S] I (new 4-helical fold)		
S^{γ} Cys 100	HN Phe 102	3.36
S^{γ} Cys 139	H_2N MCD	3.14
S^{γ} Cys 137	HN Tyr 140	3.38
	HN Cys 139	3.52
S1	HN Gly 101	3.48
	HN Glu 99	3.54
S2	HN Arg 138	3.39
	H_2O Wat 105	3.38
[2Fe-2S] II (plant-type Fd fold)		
S^{γ} Cys 40	HN Gln 44	3.33
	HN Gln 42	3.18
S^{γ} Cys 45	HN Ala 47	3.37
	H_2O Wat 200	3.63
S^{γ} Cys 48	HN Cys 60	3.40
S^{γ} Cys 60	HN Gly 43	3.49
S2	HN Gly 39	3.24
	HN Glu 41	3.42
S1	HN Cys 45	3.33
	HN Gly 46	3.25

Table 10. NH-S and OH-S hydrogen bonds established by [2Fe-2S] I and [2Fe-2S] II centres to neighbouring residues.

This asymmetry found in both Fe/S clusters may explain why one of the two iron atoms in [2Fe-2S] clusters is more easily reduced. Studies carried out with several ferredoxins, showed a correlation of the number of NH-S bonds around each Fe center with the redox potential, suggesting an increase of 75-80 mV per single NH-S bond [41,42]. In cluster II, Fe2 is stabilized by a total of four of NH-S^{γ} bonds, two more then around Fe1. On this basis one might expect that in cluster II, Fe2 would be more easily reduced than Fe1.

C. The Aldehyde Oxidoreductase from *Desulfovibrio desulfuricans* ATCC 27774.

III.C Results

III.C.Results.1 Primary sequence determination.

To obtain the amino acid sequence, several attempts were made for cloning the gene using the PCR (Polymerase Chain Reaction) technique. Based on the MOD N-terminal sequence obtained by Edman sequencing (Edman & Begg, 1967), and on a highly conserved region found in molybdopterin enzymes, primers lucky1 and lucky2 (cf. Materials and Methods) were synthesized and a DNA fragment of about 250 base pairs (bp) was amplified. Nested primers based on this sequence were used to generate two additional PCR fragments encompassing 578 bp of the 5'-end of the MOD gene, including 67 bases of the 5'-untranslated region. To extend the gene sequence towards the C-terminus, information obtained from the crystallographic data was used for oligonucleotide construction. After the initial MOD model was refined with X-PLOR, a careful inspection of the electron density map clearly revealed regions where the amino acid sequence was conserved in relation to MOP. The peptide 859-DGPFGA-864, at the C-terminus, was selected to construct a degenerate oligonucleotide, MOD864R, which together with the homologous primer MOD1 was used to amplify a DNA fragment of 2319 bp. Information on the regions flanking this sequence was obtained by Inverse PCR (Triglia *et al.*, 1988) (Ochman *et al.,* 1988) (Silver *et al.*, 1989), using primers LASTFWD and MOD1Rev (Materials and Methods). The resulting 1300 bp product was used to determine the remaining sequence of the MOD gene.

Assembly of the overlapping sequence fragments allowed the reconstitution of an open reading frame of 2721 bp, coding for a 907 amino acid polypeptide chain with a deduced molecular mass of 97964.3 Da. This value is in good agreement with electrophoresis data.

III.C.Results.2 Crystallisation and crystal characterisation

Crystals of MOD were obtained by vapor diffusion, using ammonium sulfate at pH 6.0 as precipitant and dioxane as additive. These crystals belong to the space group $P6_122$, with unit cell dimensions a =b =156.4 Å and c = 177.1 Å. Diffraction was obtained to beyond 2.5 Å resolution. A full data set was measured using X-rays from a rotating anode generator operating at 4.9 kW, yielding an overall completeness of 98.2% to a maximum

resolution of 2.8 Å.The solvent content is approximately 60% assuming one monomer of 98 kDa in the asymmetric unit (Mathews, 1968). Since the protein is known to be homodimeric in solution, the two subunits must be related by a crystallographic dyad. Other important data collection statistics are summarised in Table 5 (c.f. Materials and Methods).

III.C.Results.3 Structure solution and refinement

The MOD structure was solved using the closely related MOP molecule as a search model. The high homology between both structures was clear upon inspection of the first electron density maps, when still no primary sequence was available. Even with data to only 2.8 Å resolution it was possible to carefully analyse the electron density maps and locate the regions of the primary sequence conserved in respect to the MOP protein. In the first stages of the refinement a X-ray derived sequence was proposed for several segments highly homologous to the MOP sequence and dubious regions were left as polyalanines. The refinement and primary sequence determination were carried out in parallel and the refinement advanced as progress was made on the knowledge of the primary sequence, which allowed insertion of new blocks of corrected amino-acid sequence (Table 6 in Methods).

The side chains of eleven exposed amino acid residues - eight lysine residues (Lys 132, 168, 232, 248, 301, 478, 606 and 752), two glutamates (Glu 572 and 612) and Arg 475 - which were not defined in the electron density, were omitted from the final refinement and phase calculation. Seven other amino acid side chains that were partially defined Glu 163, Glu 607, Lys 246, Lys 603, Lys 801, Arg 604 and Arg 608 were set to occupancy of 0.5. The detailed progress of the refinement as well as statistical data of the refined MOD model is summarised in Tables 5 and 6 respectively (c.f. Methods).

The final overall geometric quality of the model was determined with PROCHECK (Laskowski *et al.*, 1993) with 88.4% of all residues found in most favoured regions of the Ramachandran plot, 11.1% in the additional allowed regions and 0.3% in the generously allowed regions. As found in the MOP protein, only two residues, Tyr 142 (Φ=69º Ψ=-37º) and Leu 254 (Φ=62º Ψ=-68º) are on disallowed regions but in both cases the electron density is very well defined. Tyr 142 (in MOP and in MOD) is placed on the N-terminus of helix IV from domain Fe/S_proximal and its hydroxyl group establishes hydrogen bond interactions with amide from Cys137 and the carbonyl group of Pro 839.

Tyr 142 is totally conserved within the xanthine oxidase family of enzymes (Figure 19). There are two X-Pro peptide bonds that adopt a *cis* conformation very clear in the electron density maps, Phe 408-Pro 409, in a new inserted loop and Leu 895-Pro 896, which is also present in the MOP structure. The MOD crystal structure has an estimated average atomic positional error of 0.2 Å (Luzzati *et al.*, 1952).

III.C Discussion

III.C.Discussion.1 Crystal structure description of MOD

As expected, the MOD enzyme has the same overall globular structure as MOP (Romão *et al.*, 1995) with 27 % α-helix and 21 % β-sheet as defined by PROCHECK (Laskovski *et al*, 1993) (Figure 19) and the same domain organisation (c.f. III.D.Discussion.1 for further details). Domain I (Fe/S_distal - residues 1-76), has a fold typical of a plant-type ferredoxin (Rypniewski *et al.,* 1991, Fukuyama *et al,* 1995) with a five-stranded β half-barrel plus an α-helix orthogonal to the strands while domain II (Fe/S_proximal - residues 84-156) is a four-helical bundle with the [2Fe-2S] centre bound at the N-terminus of the two central and larger helices (Figure 20). These domains I and II are involved in the binding of the distal and proximal iron-sulphur redox centres respectively (classified with respect to their relative proximity to the molybdenum site), and are followed by a 37 residues connecting peptide (residues 158-195) leading to contiguous third domain. Domain III (Mo1, residues 196-581) is rather elongated (ca 75 Å long and 28 Å wide) and consists of two sub-domains: a larger N-terminal part folded as a seven-stranded incomplete β-barrel with one α-helix in its central cavity and two additional helices which flank the barrel and are exposed to solvent. The smaller C-terminal subdomain of Mo1 consists of a five-stranded mixed parallel/anti-parallel β-sheet flanked by two helices approximately parallel to the strands direction. Domain IV (Mo2, residues 582 to 907) is also organized in two sub-domains each with a similar basic fold consisting of a four-stranded β-sheet bended around a pair of helices. Both sub-domains resemble two large wings extending over 80 Å and with the cofactor MCD lying at their intersection. The molybdenum catalytic site is located at the extensive interface defined by these two domains Mo1 and Mo2.

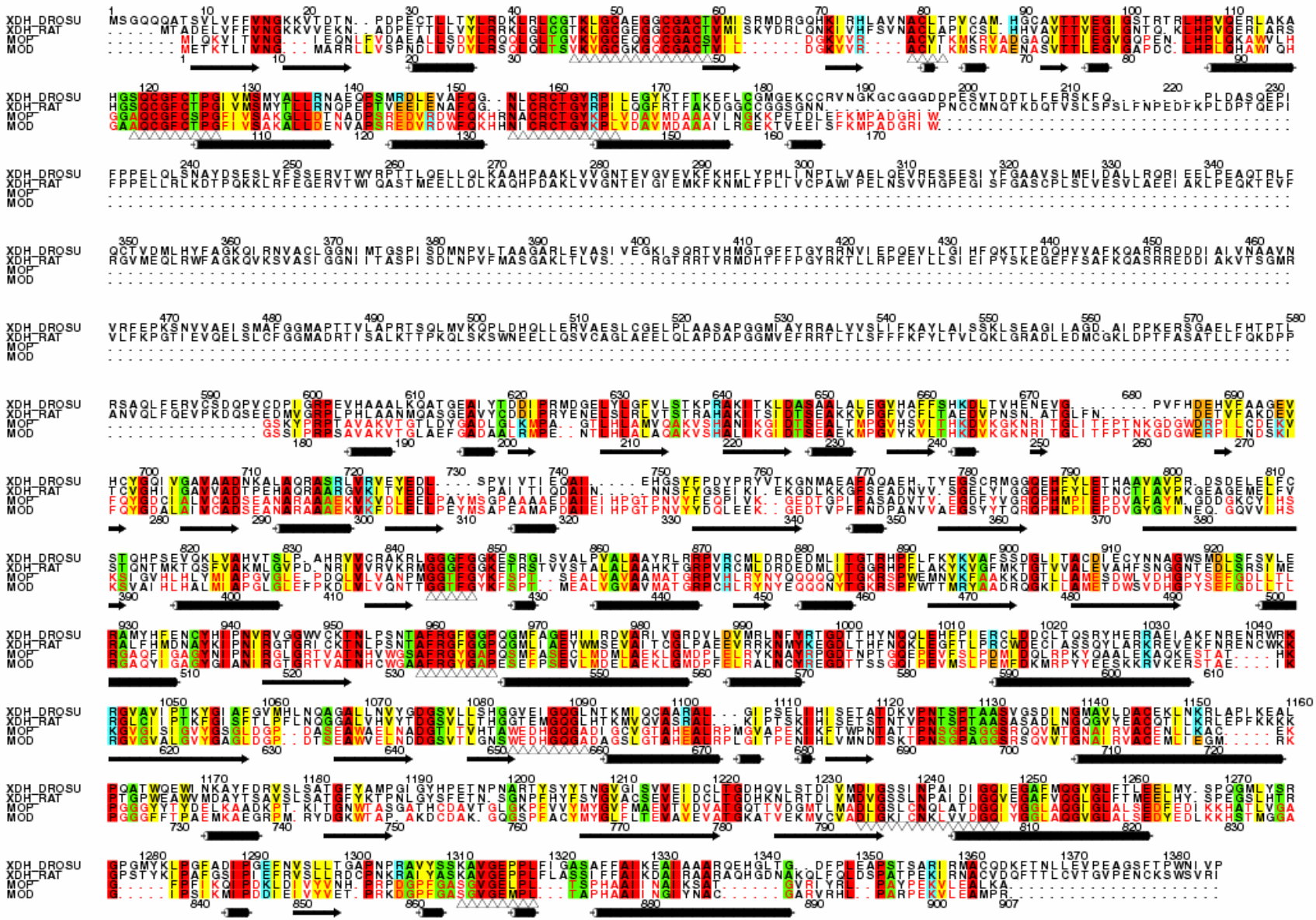

XDH_DROSU
XDH_RAT
MOP
MOD

Figure 19 - Amino acid sequence alignment of MOD with MOP and Xanthine Dehydrogenases. XDH_DROSU: xanthine dehydrogenase from *Drosophila subobscura*, Swiss-Prot accession number P91711; XDH_RAT: xanthine dehydrogenase from *Rattus norvegicus*, Swiss-Prot accession number P22985; MOP: aldehyde oxidoreductase from *Desulfovibrio gigas*, Swiss-Prot accession number Q46509; MOD: aldehyde oxidoreductase from *Desulfovibrio desulfuricans* ATCC 27774. Alignments were performed with programs CLUSTALW and ALSCRIPT, within the GCG package. Invariant residues are coloured red, conserved hydrophobic residues yellow, conserved small neutral residues green and conserved charged residues brown (negatively charged) and blue (positively charged). The secondary structural elements of β-strands (→) and α-helices (–■–) and the segments contacting the iron-sulphur clusters, the molybdopterin (Δ), and the cytosine (∇), are marked.

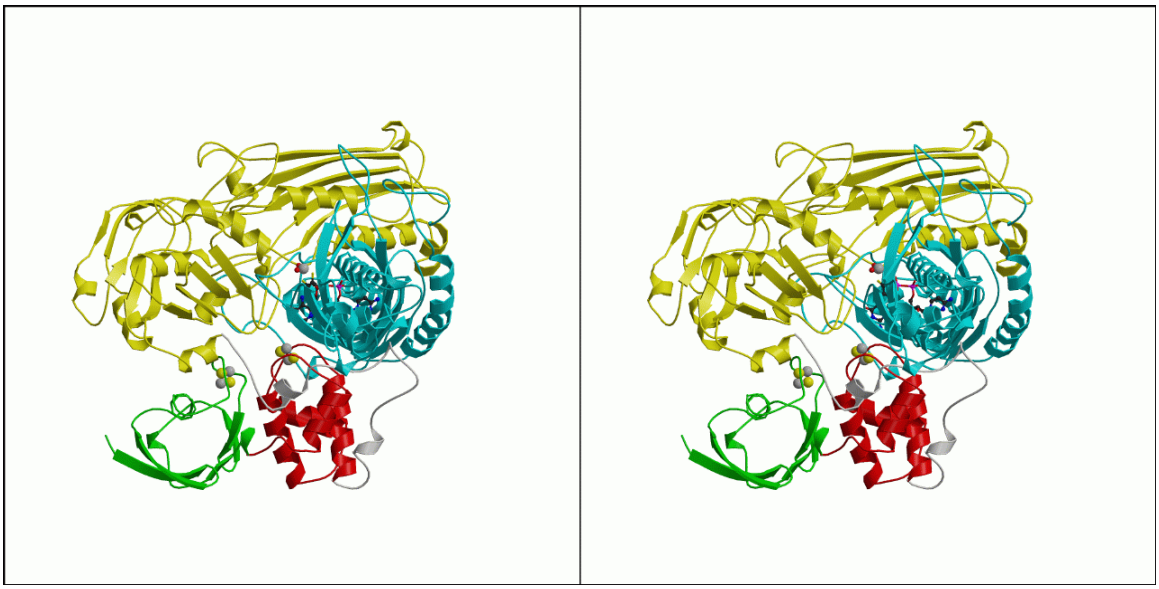


Figure 20 - Stereo plot of the molecular structure of MOD with the four independent domains represented in different colors and cofactors shown as colored spheres. I- Fe/S_distal - green, residues 1-76; II- Fe/S_proximal - red, residues 84-156; connecting peptide-white, residues 158-195; III- Mo1- yellow, residues 196-581; IV- Mo2- blue, residues 582-907. Figure prepared with MolScript (Kraulis, 1991) and Raster3D (Merrit & Murphy, 1994).

III.C.Discussion.2 MCD cofactor and the active site

The Molybdopterin Cytosine Dinucleotide (MCD) structure was parameterised using the model from the very high-resolution structure of MOP. In the refined structure, all cofactors had very clear density and the individual atoms have comparable temperature factors with the exception of the molybdenum atom and its three oxygen ligands that presented very high temperature factors for full occupancy. When the individual occupancies were refined to lower values, the resulting model of the Mo coordination sphere showed B-factors (35.3-42.2 $Å^2$) near the upper range observed for the atoms of the pterinic moiety of the cofactor (16.2-35.8 $Å^2$). The final electron density maps in the active site were of much better quality with the refined occupancy. This observation is in agreement with the known experimental evidence that enzymes from the xanthine oxidase group are often found with a certain fraction in the ´demolybdo´ form (Moura & Barata, 1994). The thermal parameters found for Moco is not evenly distributed and the largest values are found for the atoms of the pterinic moiety (25-36$Å^2$). This may reflect some disorder of this part of the cofactor, present in its demolybdenum form. Series termination errors that are undoubtedly present at this level of resolution (c.f. III.B.Results.3. for further details) also contribute in extent to reduce the quality and intensity of the observed density. In the refined structure, the molybdenum atom has been modelled in a square pyramidal geometry with the two sulphur atoms of the dithiolene unit, one water and one oxo ligand occupying the base and one additional oxo group on the apex (Figure 21).

On the MCD cofactor binding region, structural differences are minimal when compared to the MOP case, and only two residue changes were identified, Gln 700 and Ser 797, which have been replaced in MOD by Ser 702 (and a water molecule) and Lys 797 respectively. These mutations do not affect the hydrogen-bonding pattern, which stabilises the MCD cofactor in the protein matrix (Figure 22).

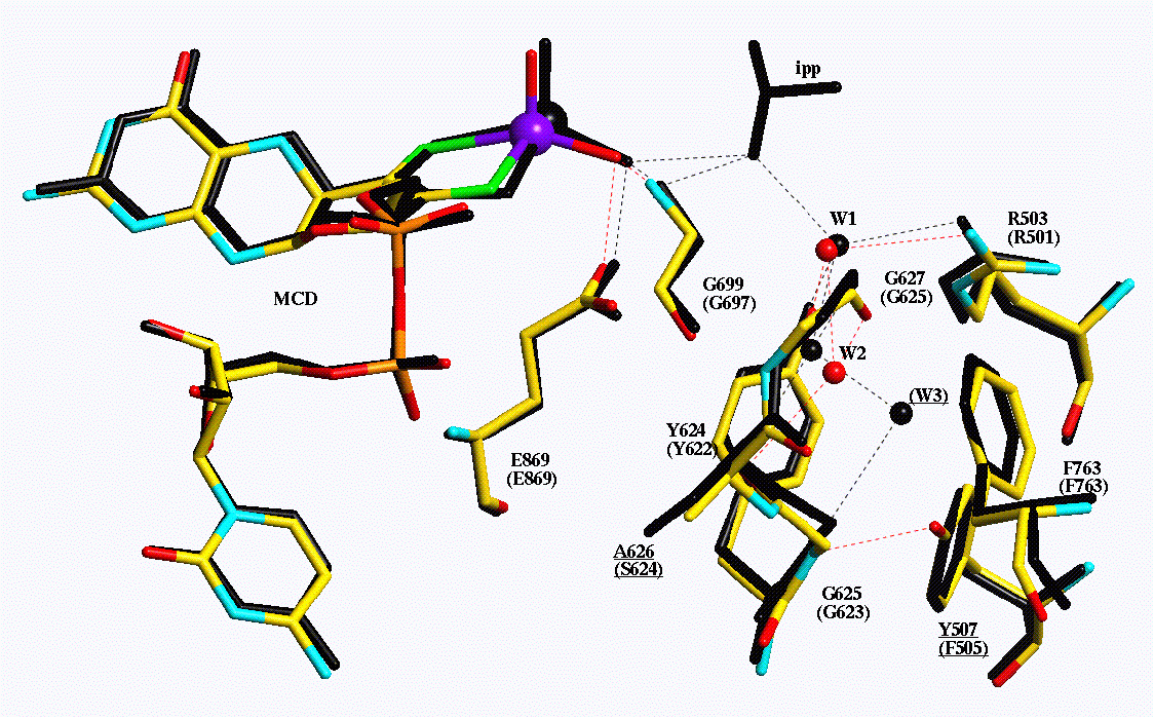


Figure 21 – Close-up of the active site residues of MOD (in TURBO-FRODO color code) superimposed with those of MOP (in black). Two water molecules are present on the hydrophobic pocket of MOD rather than the three present in the MOP structure. The isopropanol inhibitor molecule is not present in the Mod crystallizing solution and is therefore absent in this structure. Figure prepared with TURBO-FRODO (Cambillau *et al.*, 1997).

Amino acid residues in the close vicinity of the molybdenum catalytic site are also essentially conserved. Some of these residues were important for the proposal of a reaction mechanism based on the presence of an inhibitor bound form (Huber *et al.*, 1996) (Romão *et al.*, 1997). The Glu 869, which is very close to molybdenum site and is totally conserved in XDH, is involved in the catalytic cycle. It activates, as a general

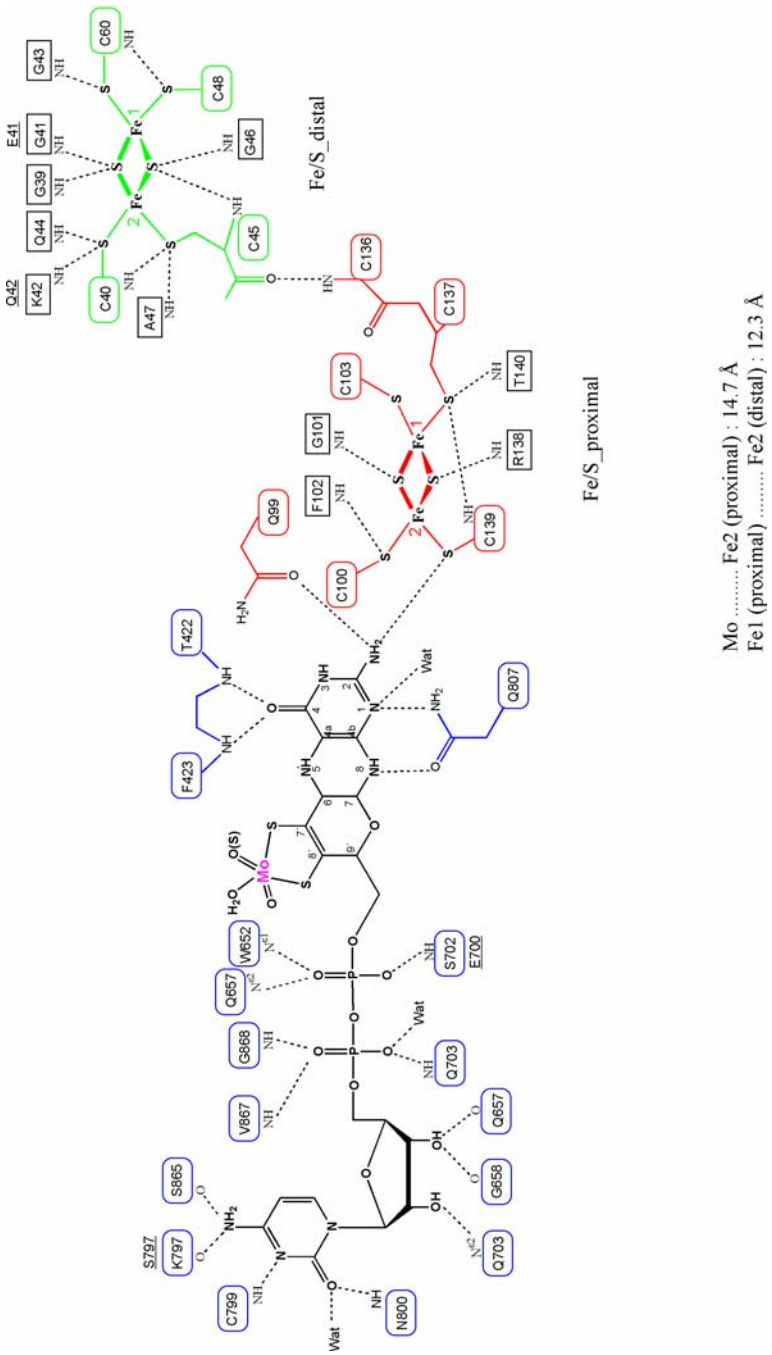


Figure 22 - Schematic representation of hydrogen bonded contacts established between the MOD residues and the molybdopterin cytosine dinucleotide and the two [2Fe-2S] centres. Residues are coloured according to their location in one of the four domains, Fe/S_distal (green), Fe/S_proximal (red), Mo1 (yellow) or Mo2 (blue). Contacts to a few buried water molecules are also included. NH--S hydrogen bonds surrounding both Fe/S clusters are also marked. The residues that are different in the MOP structure (E 41, Q 42, E 700 and S 797) are specifically underlined.

base, the co-ordinated water molecule and probably binds transiently to the metal when the carboxylic acid product is released, before the vacant co-ordination position is reoccupied by one water molecule from the chain of interior water molecules. The isopropanol inhibitor molecule present in MOP is absent in the MOD structure, as expected, since it was not present in the crystallisation solution. However, a disordered water molecule most probably replaces it, which is seen in density but poorly defined and was therefore not included in the final model.

Another important aspect of the active site cavity is the presence of the above-mentioned chain of buried water molecules, proposed as catalytically relevant. In the MOD structure only two such water molecules were found instead of three as in MOP. The innermost water molecule, W 3, present on MOP in a rather hydrophobic environment, is not seen in the MOD electron density maps. This may be due either to the low resolution of the data or to the two mutations present in the pocket surrounding the water molecules. The Phe 505 and Ser 624 amino-acid residues in MOP, are replaced respectively by Tyr 507 and Ala 626 in MOD (Figure 22). These mutations change the hydrogen-bonding pattern of the loop Tyr 624-Gly 627. While in MOP, water W 3 is stabilised by one H bond to water W 2 and to amide from Gly 623, in MOD, NH from Gly 625 makes a favourable interaction to the OH from Tyr 507 (NH---OH 2.8 Å). Additionally, on the MOP structure the Ser 624 amino-acid residue that stabilises a loop conformation via the interaction with the NH from Gly 866 (~3.0 Å) is in the MOD structure replaced by Ala 626.

On the basis of point mutations of the *Aspergillus nidulans* xanthine dehydrogenase (XDH), it was possible to define residues essential for substrate positioning (Glatigny *et al*., 1998). Some of these mutations implied partial loss of function of XDH and corresponded to residues in the vicinity of the MOP/MOD substrate-binding pocket. Particularly interesting is Arg 503 (Arg 501 in MOP) whose mutation in *Aspergillus nidulans* XDH (Arg 911 to Gln or Gly) changes the hydroxylation position of 2-hydroxypurine from C8 to C6. This Arg is totally conserved within the XDH family of enzymes and was proposed to be essential to position purine substrates in the Mo center, in agreement with our earlier proposals based on the crystallographic studies of MOP (Huber *et al*., 1996).

III.C.Discussion.3 The two [2Fe-2S] redox centres

The two redox active iron-sulphur centres represent distinct forms, which are structurally related but very differently bound to the protein. For the distal cluster, a sequence Cys 40-X1-X2-Gly-Gln-Cys 45-Gly-Ala-Cys 48- (as found in plant-type ferredoxins) traps the centre, which is closed by Cys 60, placed 11 residues ahead. For the proximal cluster a totally different cysteine pattern is involved, as one of the iron atoms binds on a Cys 100-Gly-Phe-Cys 103 motif. The cluster is enclosed by a similar pattern located 34 residues ahead, Cys 137-Arg-Cys 139. As a consequence of the different cysteine-chelating motif both clusters exhibit different NH--S bonding patterns (Figure 22). These NH--S bonds are involved in the accommodation of both Fe/S centres inside the protein and modulate the redox potential as found in other types of iron clusters (Backes *et al.*, 1991)(Sheridan *et al.*, 1981).

The two Fe/S clusters are about 12.3 Å far apart (Fe2 atom of FeS_distal to Fe1 atom of Fe/S_proximal). The Fe2 atom from the proximal iron-sulphur centre is 14.7 Å away from the molybdenum atom and in contact with the pterin via residue Cys 139. The side chains of Val 38 and Arg 58 at the surface partially shield the distal cluster from the solvent. Not surprisingly both clusters maintain their co-ordination as in MOP, in spite of a few mutations on the close vicinity of the metals. Residues Ser 104, Ala 136 and Glu 41, Gln 42 and Ser 49 have been replaced by residues Thr 104, Ile 136 and Gly 41, Lys 42 and Thr 49 for the proximal and the distal clusters respectively. The most conspicuous mutation is on residue 41 (Glu to Gly) which makes one NH--S contact to a sulphur from the centre. The side chain of Glu 41 is an exposed one. In residues 104 and 49, where a Thr in MOD replaces a Ser in MOP, the same hydrogen-bonding pattern is maintained in both structures. The proximal cluster establishes six NH--S hydrogen bonds. A direct NH--S contact between the S^{γ} atom of Cys 139 and the NH_2 group at C2 of the pterin is the most probable electron transfer pathway (Romão *et al.*, 1995). The distal, iron-sulphur centre has a different NH--S hydrogen bonding pattern being stabilised by a total of ten such interactions.

D. Further comparison between MOP and MOD and the Mo hydroxylases.

III.D Discussion

III.D.Discussion.1 Protein structure

The structure determination of the aldehyde oxidoreductase from *Desulfovibrio desulfuricans* ATCC 27774 (MOD) was carried out at a resolution of 2.8 Å which is lower than obtained for the homologous protein from *D. gigas* (MOP) (1.8Å reported by Huber *et al.*, 1996 and 1.28 Å on the present work). When comparing both structures, non-conserved residues are homogeneously spread over the polypeptide chain. There are two single residue insertions and two single residue deletions placed on exposed loops of the structure. The two insertions on the MOP sequence occur between residues 354 and 355 (355 in MOD) and between residues 406 and 407 (408 in MOD), and two deletions between residues 750 and 751 (749 in MOP) and 757 and 758 (757 in MOP).

When we superimpose 879 common Cα atoms of MOD and MOP, the average Cα distance is 0.59 Å and the r.m.s. deviation is 0.39Å which is not surprising due to the high degree of primary sequence identity (68%) between both enzymes (Figure 19). The largest differences, C_α deviations greater than 2 Å, occur essentially at the C terminus and on a few exposed loop regions (residues 66-67; 165-166; 170-172 (disordered loop); 319-320; 349-352 (1 amino acid insertion in MOD); 642-644; 683-685; 722-730; 750-751). The core and active site of the molecule is basically conserved and practically identical in both structures. Also the region surrounding the entrance to the funnel like cavity, which leads to the active site is remarkably conserved. Important non-polar residues at the half length of the access tunnel, Phe 425, Phe 494, Leu 497 and Leu 626 which we have described as "swinging doors" (Romão *et al.,* 1995) are also conserved in MOD (Phe 427, Phe 496, Leu 499 and Leu 628 respectively).

The active site topology as well as the conformation assumed by the moco cofactor superimpose extremely well, not only between MOP and MOD as already referred (figure 21), but also with XO/XDH and CODH. The difference to be found on the CODH active site is the presence of Cu ion supported by a cysteine that belongs to a second sphere of Mo coordination via μ-sulfido, thereby forming a [CuSMo] cluster. These findings have obviously deep consequences, leading to different enzymatic mechanism.

MOD and MOP primary sequences are homologous to eukaryotic xanthine dehydrogenases as well as to carbon monoxide dehydrogenase (CODH) both containing the flavin domain absent in the aldehyde oxidoreductases. The calculation of the electrostatic potentials for MOD and MOP (Figure 23) shows that there are differences evenly spread on the molecular surfaces, but that the upper right quadrant is dominated by a cluster of positively charged residues in both structures. This same region corresponds roughly to the docking site of the flavin domain of CODH from *Oligotropha (O.) carboxidovorans* (Dobbek, et al., 1999) and bovine milk xanthine dehydrogenase (Enroth, et al., 2000) and may mark the docking site of the physiological electron acceptor in MOP and MOD.

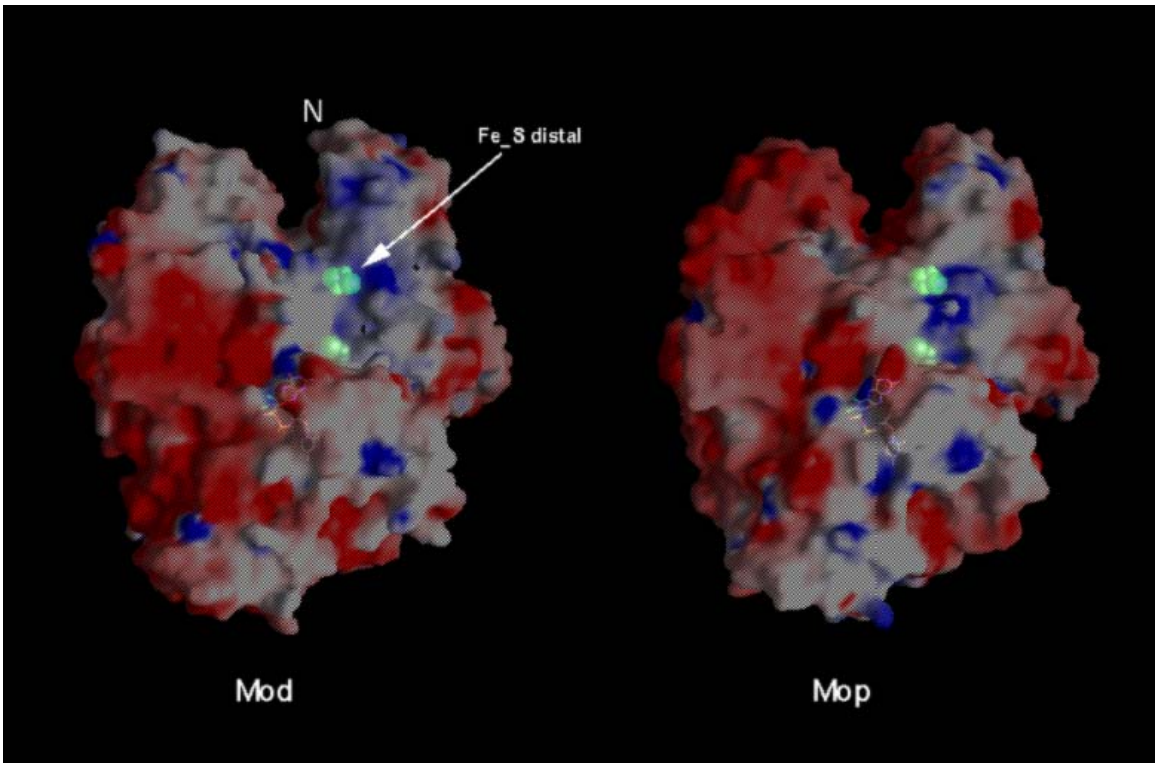


Figure 23 – Electrostatic surface potential for MOP and Mod calculated with GRASP (Nicholls *et al.*, 1991). These molecular surfaces show the electrostatic potential coloured from $-10k_BT$ (red) to $10\ k_BT$ (blue) (k_B is the Boltzman constant and T the absolute temperature). The arrow points to the distal, solvent exposed, [2Fe-2S] centre.

III.D.Discussion.2 The dimerization

As found in several other molybdenum hydroxylases (MH) (Hille, 1996), MOD is isolated as a homo-dimer although the subunits are functionally independent and, as in MOP, both subunits are related by a crystallographic dyad (Figure 24). In both XO/XDH and CODH the dimerization also occurs in solution and both are crystallized on the dimeric form enclosed on a single asymmetric unit (a.u.).

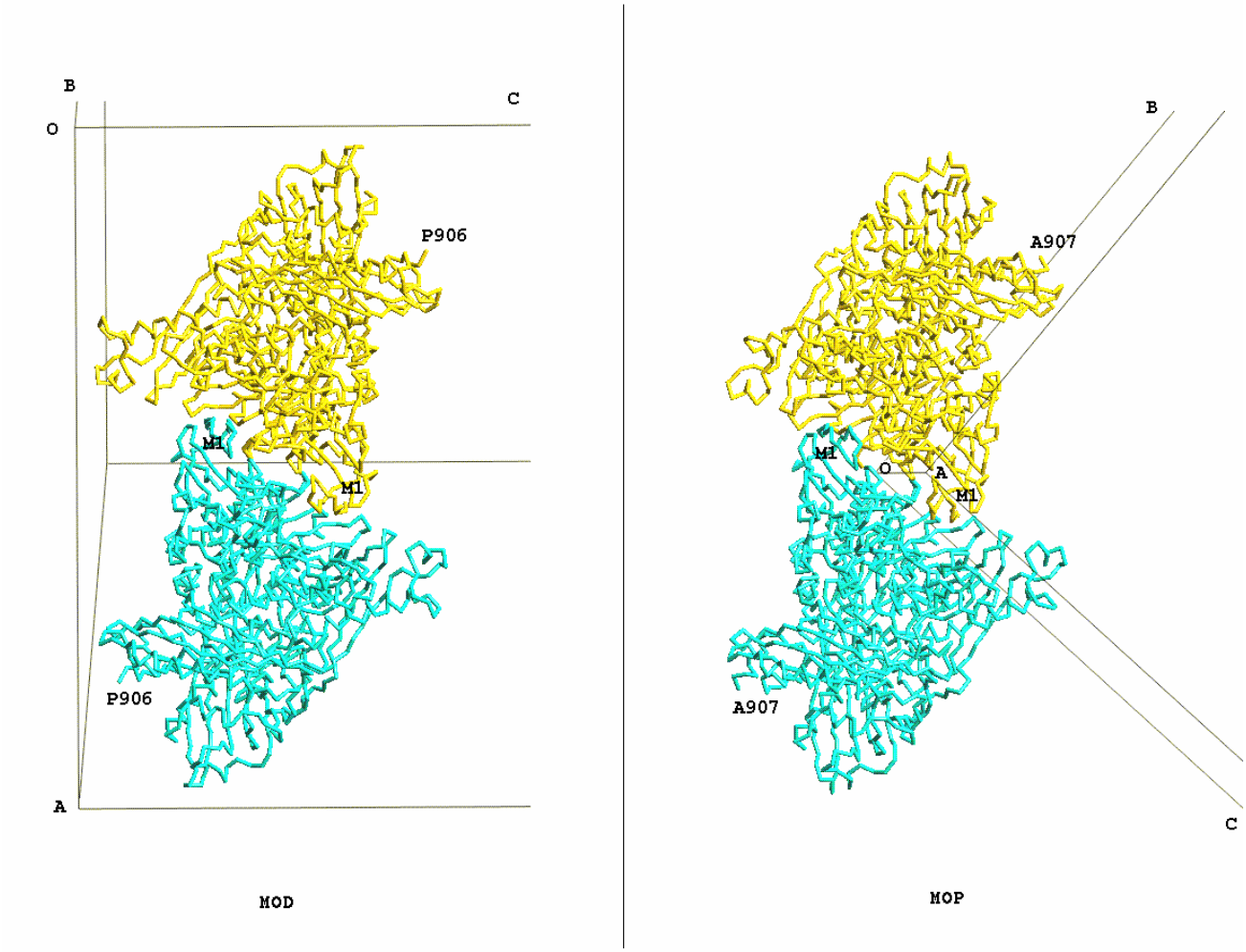


Figure 24 – (a) Packing of one crystallographically related Mod dimer in the $P6_122$ unit cell and (b) packing of one MOP dimer in the corresponding unit cell viewed in the same orientation as the Mod dimer. Figure prepared with TURBO-FRODO (Cambillau *et al.*, 1997).

In spite of the fact that both molecules, MOP and MOD crystallize under totally different conditions (30% isopropanol, 0.2M $MgCl_2$ at pH 7.2 (Romão *et al.*, 1993) vs. ammonium sulfate, 2.5M at pH 6, respectively), both yield hexagonal crystals (hexagonal bi-pyramides vs. hexagonal prisms, respectively) belonging to the same space group $P6_122$. The cell parameters are however different and, due to different packing effects, the location of the crystallographic α_2 dimer is different in both hexagonal lattices.

In the dimer, the redox cofactors are relatively far apart in both subunits with the shortest distance of 39.3 Å between the nearest iron atoms (from centre Fe/S_distal and its symmetry partner). The molecular dyad runs along the contact of both Fe/S_distal domains through Gly 32, Leu31, such that residues from Fe/S_distal (I) and Mo1 (III) domains contact symmetry partners from domains Fe/S_distal* and Mo1* (Figure 24).

subunit 1	**subunit 2 ***
Met 1	Leu 192*
Thr 3 (Gln 3)	Thr 377*
Met 11 (Ile 11)	Leu 192* (Thr 192*)
	Ala 187*
Arg 14 (Asn 14)	Asn 417*
Leu 16 (Phe 16)	Met 203*
	Tyr 377*
Met 203	Leu 16*
Glu 205 (Ala 205)	Ser 289*
Tyr 377	Thr 3* (Gln 3*)
Glu 380 (Asp 380)	Pro 19* (Ala 19*)

Table 11. Contacting residues at the subunit interface of the MOD dimer [(a)]

(a) for comparison purposes, corresponding residues in the MOP dimer which are different from MOD are shown in parenthesis.

* crystallographically related

These symmetrical contacts are mainly of a hydrophobic nature as seen in Table 11 where contacts through hydrogen bonding interactions between both subunits are summarized. An interesting feature observed only in the MOD dimer interface is the presence of Cys 84 (Asn 84 in MOP) in close vicinity to its symmetry mate (distance S-S* 3.1 Å) but not forming a disulfide bridge.

In the related enzyme carbon monoxide dehydrogenase from *Oligotropha carboxidovorans* (Dobbek, et al., 1999), the dimer axis is different, but the redox cofactors are also placed far apart in both subunits. This indicates that for all these enzymes of the Xanthine Oxidase family the two subunits are functionally independent.

IV References

Agarwal, R.C. (1978). A new least-squares refinement technique based on the Fast Fourier Transfor. Algorithm. *Acta Cryst*. **A34**, 791-809.

Allen F H, Bellard S, Brice M D, Cartwright B A, Doubleday A, Higgs H, Hummelink T, Hummelink-Peters B G, Kennard O, Motherwell W D S, Rodgers J L & Watson D G (1979). The Cambridge Crystallographic Data Centre: computer-based search, retrieval, analysis and display of information *Acta Cryst*. **B35**, 2331-2339.

Andrade S L A, Brondino C, Feio M J, Moura I and Moura, J J G (2000). Aldehyde oxidoreductase activity in Desulfovibrio alaskensis NCIMB 13491 EPR assignment of the proximal [2Fe-2S] cluster to the Mo site. *Eur J Biochem*. **267**, 2054-2061.

Auerbach, G., Herrmann, A., Bracher, A., Bader, G., Gütlich, M., Fischer, M., Neukamm, M., Garrido-Franco, M., Richardson, J., Nar, H., Huber, R. & Bacher, A. (2000). Zinc plays a Key Role in GTP Cyclohydrolase I. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13567-13572.

Backes G., Mino Y., Löhr T.M., Meyer T.E., Cusanovich M.A., Sweeney W.V., Adam E.T.,Sanders-Löhr J. (1991). The environment of [4Fe-4S] clusters in Ferredoxins and High-Potential Iron Proteins. New information from X-ray Crystallography and Raman Spectrometry. *J.Am.Chem.Soc.* **113**:2055-2064.

Bailey, S. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst.* **D50**, 760-763.

Bandmann O., Daniel S., Mardsen C.D., Wood N.W. & Harding A.E. (1996a) The GTP-cyclohydrolase I gene in atypical parkinsonian patients: a clinico-genetic study. *J. Neurol. Sci.*. **141**, 27-32.

Bandmann O., Nygaard T.G., Surtees R., Mardsen C.D., Wood N.W. & Harding A.E. (1996b). Dopa-responsive dystonia in British patients: new mutations of the GTP-

cyclohydrolase I gene and evidence for genetic heterogeneity. *Hum. Mol. Gen.* **5**, 403-406.

Barata B. A. S., J. Legall, J. J. G. Moura (1993). Aldehyde oxidoreductase activity in *Desulfovibrio gigas: In vitro* reconstitution of an electron-transfer chain from aldehydes to the production of molecular hydrogen. *Biochemistry* **32**, 11559-11568.

Barton, G.J. (1993). AlScript, a tool to format multiple sequence alignments. *Protein Engineering* **6**, 37-40.

Betts, L., Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W., Jr. (1994). Cytidine deaminase. The 2.3 crystal structure of an enzyme: transition-state analog complex. *J. Mol. Biol.* **235**, 635-656.

Blau, N., Barnes, I. & Dhondt, J.L.J. (1996) International database of tetrahydrobiopterin deficiencies. *J. Inherit. Metab. Dis.* **19**, 8-14.

Bracher, A., Eisenreich, W., Schramek, N., Ritz, H., Götze, E., Herrmann, A., Gütlich, M. & Bacher, A. (1998). Biosynthesis of pteridines. NMR studies on the reaction mechanisms of GTP cyclohydrolase I, pyruvoyltetrahydropterin synthase, and sepiapterin reductase. *J. Biol. Chem.* **273**, 28132-28141.

Bracher, A., Fischer, M., Eisenreich, W., Ritz, H., Schramek, N., Boyle, P., Gentili, P., Huber, R., Nar, H., Auerbach, G. & Bacher, A. (1999). Histidine 179 mutants of GTP cyclohydrolase I catalyze the formation of 2-amino-5-formylamino-6-ribofuranosylamino-4(3H)-pyrimidinone triphosphate. *J. Biol. Chem.* **274**, 16727-16735.

Bracher, A., Schramek, N. & Bacher, A. (2001). Biosynthesis of pteridines. Stopped flow kinetic analysis of GTP cyclohydrolase I. *Biochemistry* **40**, 7896-7902.

Brown, G. M. & Williamson, J.M. (1987). Biosynthesis of folic acid, riboflavin, thiamine, and panthotenic acid. In escherichia coli and Salmonella typhimurium: *Celular and Molecular Biology* (Neidhardt F.C., ed.), vol. 1, pp.521-538, American Society for Microbiology, Whashington, DC.

Brünger A.T., Adams P., Clore G., DeLano W., Gros B., Grosse-Kunstleve R., Jiang J-S., Kuszewski J., Nilges M., Pannu N., Read R., Rice L., Simonson T. & Warren G. (1998). Crystallography & NMR system: A new software for macromolecular structure determination. *Acta cryst.* **D54**, 905-921.

Brünger A.T. X-PLOR *: A System for Crystallography and NMR.*, (version 3.851, manual (Yale Univ. Press, New Haven, CT, 1990)).

Brünger A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472-475.

Buerger, M.J. (1959). Vector space and its application in vector structure investigation. John Willey, New York.

Burg, A. W. & Brown, G. M. (1968). The biosynthesis of folic acid. VIII. Purification and properties of the enzyme that catalyzes the production of formate from carbon atom 8 of guanosine triphosphate. *J. Biol. Chem.* **243**, 2349-2358.

Caldeira J., Belle V., Asso M., Guigliarrelli B., Moura I., Moura J. J. G. and Bertrand P. (2000). Analysis of the electron paramagnetic resonance properties of the [2Fe-2S]1+ centers in molybdenum enzymes of the xanthine oxidase family: assignment of signals I and II. *Biochemistry* **39**, 2700-7.

Cambillau, C., Roussel, A., Inisan, A., G. & Knoops-Mouthay, K., (1997). *TURBO-FRODO*. Version OpenGL.1. CNRS/Université Aix-Marseille II, Marseille, France.

CCP4, Collaborative Computational Project number 4, (1994). The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst.* **D 50**, 760-763.

Chan M., Mukund S., Kletzin A., Adams M., Rees D. (1996). Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* **267**: 1463-1469.

Christianson N. and Lipscomb W. N. (1988). Structural aspects of zinc proteases mechanisms. *Mol. Struct. Energ.* **9**, 1-25.

Christianson, D. W. & Cox, J. D. (1999). Catalysis by metal-activated hydroxide in zinc and manganese metalloenzymes. *Annu. Rev. Biochem.* **68**, 33-57.

Christianson, D. W., David, P. R. & Lipscomb, W. N. (1987). Mechanism of carboxypeptidase A: hydration of a ketonic substrate analogue. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 1512-1515.

Dias, J. M., Than, M. E., Humm, A., Huber, R., Bourenkov, G.P., Bartunik, H. D., Bursakov, S., Calvete, J., Caldeira, J., Carneiro, C., Moura, J.J.G., Moura, I. & Romão, M.J. (1999). Crystal structure of the first dissimilatory nitrate reductase at 1.9Å solved by MAD methods. *Structure* **7**, 65-79.

Dobbek H., Gremer L., Meyer O. & Huber R. (1999). Crystal structure and mechanism of the CO dehydrogenase, a molybdo iron-sulfur flavoprotein containing S-selanylcysteine. *Proc. Natl. Acad. Sci. USA,* **96**, 8884-8889.

Dobbek H. and Huber R. (2002) Metal Ions in Biological Systems, Vol. **39,** Molybdenum and Tungsten. Their Roles in Biological Processes**.** Ed. A. Sigel, H. Sigel, 539-570.

Edman P., Begg G.(1967). A protein sequenator. *Eur.J. Biochem.* **1**:80-91.

Eisenreich, W. & Bacher, A. (1994). Biosynthesis of methanopterin in *Methanobacterium thermoautotrophicum. Pteridines*, **5**, 8-17.

Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for x-ray protein structure refinement. *Acta Cryst.,* **A47**, 392-400.

Enroth C., Eger B.T., Okamoto K., Nishino T., Pai E.F. (2000). Crystal structures of bovine milk xanthine dehydrogenase and xanthine oxidase: structure-based mechanism of conversion. *Proc.Natl.Acad.Sci.* USA **97**:10723-10728.

Esnouf R. (1997). An extensively modified version of MolScript that includes greatly enhanced colour capabilities. *J.Mol.Graph*. **15**, 112-113.

Fukuyama K., Ueki N., Nakamura H., Tsukihara T., Matsubara H. (1995). Tertiary structure of [2Fe-2S] Ferredoxin from *Spirulina platensis* refined at 2.5 Å resolution: Structural comparison of plant-type ferredoxins and an electrostatic potential analysis. *J.Biochem (Tokyo)* **117**, 1017-1123.

Genetics Computer Group (1994) Program Manual for the Wisconsin Package, version 8, 575 Science Drive, Madison, Wisconsin, USA.

Giacovazzo C., Monaco H.L., Viterbo D., Scordari F., Gilli G., Zanotti G., Catti M. (1992). *Fundamentals of crystallography.* Oxford University Press.

Glatigny A., Hof P., Romão M.J, Huber R., Scazzocchio C. (1998). Altered specificity mutations define residues essential for substrate positioning in xanthine dehydrogenase. *J. Mol. Biol.*, **278**, 431-438.

Green, J.M., Nichols, B.P., and Matthews, R.G. (1996) Folate biosynthesis, reduction, and polyglutamation. In *Escherichia coli and Salmonella*, **Vol. 1**. Neidhardt, F.C. (ed.). Washington, DC: American Society for Microbiology Press, pp. 665-673.

Hangauer, D. G., Monzingo, A. F. & Matthews, B. W. (1984). An interactive computer graphics study of thermolysin-catalyzed peptide cleavage and inhibition by N-carboxy-methyl dipeptides. *Biochemistry* **23**, 5730-5741.

Hille, R. (1996). The Mononuclear Molybdenum enzymes. *Chem. Rev.* **96**, 2757-2816.

Higgins D., Thompson J., Gibson T.Thompson J.D., Higgins D.G., Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. **22**:4673-4680.

Holden, H. M. & Matthews, B. W. (1988). The binding of L-valyl-L-tryptophan to crystalline thermolysin illustrates the mode of interaction of a product of peptide hydrolysis. *J. Biol. Chem.* **263**, 3256-3260.

Hu Y., Faham S., Roy R., Adams M. W. and Rees D.C. (1999). Formaldehyde ferredoxin oxidoreductase from Pyrococcus furiosus: the 1.85 A resolution crystal structure and its mechanistic implications. *J.Mol.Biol.*, **286**, 899-914.

Huber R., Hof P., Duarte R. O., Moura J.J.G., Moura I., Liu M-Y., LeGall J., Hille R., Archer M & Romão M.J. (1996). A structure-based catalytic mechanism for the xanthine oxidase family of molybdenum enzymes. *Proc. Natl. Acad. Sci*. USA **93**, 8846-8851.

Huber, R. (1969) *In* "Crystallographic Computing Proc.", p.96 (ed. Ahmed, F.R.) Munksgaard, Copenhagen.

Jones, T. A., Bergdoll, M. & Kjeldgaard, M. (1990). O: a macromolecule modelling environment. In *Crystallography and Modelling Methods in Molecular Design* (Bugg, C. & Ealick, S., eds.), pp. 189-199. Springer Verlag, New York.

Katzenmeier, G., Schmid, C., Kellermann, J., Lottspeich, F. & Bacher, A. (1991). Biosynthesis of tetrahydrofolate. Sequence of GTP cyclohydrolase I from *Escherichia coli*. *Biol. Chem. Hoppe Seyler* **372**, 991-997.

Kaufman S, Pollock RJ, Summer GK, Das AK, Hajra AK. (1990). Dependence of an alkyl glycol-ether monooxygenase activity upon tetrahydropterins. *Biochim. Biophys. Acta*. **1040**:19-27.

Kisker C., Schindelin H., Pacheco A., Wehbi W.A., Garrett R.M., Rajagopaln K.V., Enemark J.H., Rees D.C. (1997). Crystal structure of sulfite oxidase and identification of mutations responsible for sulfite oxidase deficiency. *Cell* **91**:973-983.

Kraulis, P.J. (1991). MolScript: a program to produce both detailde and schematic plots for protein structures. *J. Appl. Crystallogr.* **24**, 946-950.

Kwon, N.S., Nathan, C.F. & Stuherl, D.J. (1989). Reduced biopterin as a cofactor in the generation of nitrogen oxides by murine macrophages. *J. Biol. Chem.* **264**, 20496-20501.

Laskowski, R.A., McArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK - a program to check stereochemical quality of protein structures. *J. Appl. Crystallogr*. **26**, 283-291.

Leslie, A. G. W. (1991). MOSFLM: Macromolecular Data Processing. In *Crystallographic Computing* (Moras, V. D., Podjarny, A. D. & Thierry, J. C., eds.), pp. 27-38. Oxford University Press, Oxford.

Lipscomb W.N. and Sträter N. (1996) Recent advances in zinc enzymology. *Chem. Rev.* **96**, 2375-2433.

Luzzati, P.V. (1952). Traitement statístique des erreurs dans la déterminacion des structures cristalines. *Acta Cryst.* ***D*** **5**, 802-810.

Maier, J., Witter, K., Gütlich, M., Ziegler, I., Werner, T. & Ninnemann, H. (1995). Homology cloning of GTP-cyclohydrolase I from various unrelated eukaryotes by reverse-transcription polymerase chain reaction using a general set of degenerate primers. *Biochem. Biophys. Res. Commun.* **212**, 705-711.

Marletta M.A. (1993). Nitric oxide synthase structure and mechanism. *J. Biol. Chem*. **268**:12231-12234.

Mathews, B.W. (1968) Solvent contents of protein crystals. *J. Mol. Biol.* **33**, 491-497.

McAlpine, A.S., McEwan, A.G., Shaw, A. L. and Bailey, S. (1997) Molybdenum active centre of DMSO redutase from Rhodobacter capsulatus: crystal structure of the oxidised enzyme at 1.8A resolution and the dithionite-reduced enzyme at 2.8A resolution. *J.B.I.C.* **2,** 690-701.

McDonald, I.K & Thornton, J.M. (1994). Satisfying Hydrogen Bonding Potential in Proteins. *J. Mol. Biol*. **238**, 777-793.

Mejlhede, N. & Neuhard, J. (2000). The role of zinc in Bacillus subtilis cytidine deaminase. *Biochemistry* **39**, 7984-7989.

Merritt, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0., a program for photorealistic molecular graphics. *Acta Cryst.,* **D50**, 869-873.

Moura J J G, Xavier A V, Brusci M,LeGall J, Hall D and Carmmack R, (1976). A molybdenum-containing iron-sulphur protein from Desulphovibrio gigas. *Biochim Biophys Res Commun* **72**, 782-789.

Moura J.J.G., Barata B.A. (1994), Aldehyde oxidoreductases and other molybdenum-containing enzymes. *Methods Enzymol.* **243**, 24-42.

Moura J.J.G., Xavier A.V., Carmmack R., Hall D.O., Buschi M. & LeGall J. (1978). Oxidation-Reduction studies of the Mo(2Fe-2S) protein from *Desulfovibrio gigas*. *Biochem. J.* **173**, 419-425.

Murzin, A.G., Brenner, S.E., Hubbard, T., Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536-540.

Nagatsu T. & Ichinose, H. (1997). GTP cyclohydrolase I gene, dystonia, juvenile parkinsonism, and Parkinson's disease. *J. Neural. Transm. Suppl.* **49**, 203-209.

Nar, H., Huber, R., Auerbach, G., Fischer, M., Hösl, C., Ritz, H., Bracher, A., Meining, W., Eberhardt, S. & Bacher, A. (1995a). Active site topology and reaction mechanism of the GTP cyclohydrolase I. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 12120-12125.

Nar, H., Huber, R., Meining, W., Schmid, C., Weinkauf, S. & Bacher, A. (1995b). Atomic structure of GTP cyclohydrolase I. *Structure* **3**, 459-466.

Navaza, J (1994).AMoRe: An automated package for molecular replacement. *Acta Cryst.* **A50**, 157-163.

Nichol, C.A., Smith, G.K. & Duch, D.S. (1985). Biosynthesis and metabolism of tetrahydropterins and molybdopterin. *Annu. Rev. Biochem.* **54**, 729-764.

Nicholls A., Sharp K. & Honig B. (1991). Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 134-138.

Ochman, H., Gerber, A. S., Hartl, D. L.(1988). Genetic applications of an inverse polymerase chain reaction. *Genetics*, **120**, 621-623.

Otwinowski Z & Minor W, (1997). Processing X-ray diffraction data collected in oscillation mode. *Methods in Enzymol* **276**, 307-326.

Pilato R. S. and Stifel E.I., Molybdenum and tungsten enzymes. In: J.Reedijk and E. Buowman, Editors, *Bioinorganic Catalysis* (2nd ed.),, Dekker, NY, USA (1999), pp. 81-152.

Rajagopalan K. V. (1991). Novel aspects of the Biochemistry of the molybdenum cofactor. *Adv Enzym* **64** : 215-290.

Romão M.J., Archer M., Moura I., Moura J.J.G., LeGall J., Engh R., Schneider M., Hof P. & Huber R. (1995). Crystal Structure of the Xanthine Oxidase-Related Aldehyde Oxido-Reductase from *D. gigas*. *Science* **270**, 1170-1176.

Romão M.J., Barata B.A.S., Archer M., Lobeck K., Moura I., Carrondo M.A., LeGall J., Lottspeich F., Huber R., Moura J.J.G. (1993). Subunit composition, crystallization and preliminary crystallographic studies of *Desulfovibrio gigas* aldehyde oxidoreductase containing molybdenum and [2Fe-2S] centres. *Eur. J. Biochem.* **215**, 729-732.

Romão M.J., Rösch N., Knäblein, J & Huber R. (1997). The Molybdenum site in the Xanthine Oxidase-related aldehyde oxidoreductase from *Desulfovibrio gigas* and a catalytic mechanism for this class of enzymes. *JBIC*. **2**, 782-785.

Rypniewski W. R., Breiter D.R., Benning M.M., Wesenberg G., Oh B.H., Markley G.L., Rayment I., Holden H.M.(1991).Crystallization and structure determination to 2.5Å resolution of the oxidized [2Fe-2S] ferredoxin isolated from Anabaena 7120. *Biochemistry* **30**, 4126-4131.

Schindelin H., Kisker C. and Rees, D.C. (1997) The molybdenumcofactor: a crystallographic perspective. *J. Biol. Inorg. Chem.* **2**, 773-781.

Schindelin H., Kisker C., Hilton J., Rajagopalan K.V., Rees D.C. (1996). Crystal structure of DMSO reductase: redox-linked changes in molybdopterin coordination. *Science* **272**:1615-1621.

Schneider F., Löwe J., Huber R., Schindelin H., Kisker C., Knäblein J. (1996). Crystal structure of dimethyl sulfoxide reductase from Rhodobacter capsulatus at 1.88 Å resolution. *J. Mol. Biol.* **263**:53-69.

Schramek N., Bracher A., Bacher A. (2001a). Biosynthesis of riboflavin. Single turnover kinetic analysis of GTP cyclohydrolase I. *J. Biol. Chem.* **276**, 44157-62.

Schramek N., Bracher A., Bacher A. (2001b) Ring opening is not rate limiting in the GTP Cyclohydrolase I reaction. *J. Biol. Chem.* **242**, 2622-2626.

Schramek N., Bracher A., Fischer M., Auerbach G., Nar H., Huber R., Bacher A. (2002). Reaction mechanism of GTP cyclohydrolase I: Single turnover experiments using a kinetically competent reaction intermediate. *J. Mol. Biol.* **316**, 829-37.

Scriver CR, Eisensmith RC, Woo SLC, Kaufman S. (1994). The hyperphenylalaninemias of man and mouse. *Annu Rev Genet* **28**:141-165.

Sheldrick G & Schneider T (1997) SHELXL: high resolution refinement. *Methods in Enzymol* **277**, 319-343.

Sheridan R.P., Allen L.C., Carter C.W. Jr (1981). Coupling between oxidation state and hydrogen bond conformation in high potential iron-sulphur protein. *J. Biol. Chem.* **256**:5052-5057.

Shiota, T., Baugh, C. M. & Myrick, J. (1969). The assignment of structure to the formamidopyrimidine nucleoside triphosphate precursor of pteridines. *Biochim. Biophys. Acta* **192**, 205-210.

Shiota, T., Palumbo, M. P. & Tsai, L. (1967). A chemically prepared formamidopyrimidine derivative of guanosine triphosphate as a possible intermediate in pteridine biosynthesis. *J. Biol. Chem.* **242**, 1961-1969.

Silver, J., Keerikatte,V. (1989). Novel use of the polymerase chain reaction to amplify celular DNA adjacent to an integrated provirus. *Journal of Virology*, **63**,1924-1928.

Stams F.J.M., Veenhuis M., Weenk G.H. and Hansen T.A. (1983). Occurrence of polyglucose as a storage polymer in *Desulfovibrio* species and *Desulfobulbus propionicus*. *Arch. Microbiol.*, **136**, 54-59.

Stiefel E I (1997). Chemical Keys to Molybdenum Enzymes. *J Chem Soc*, Dalton Trans 3915-3923.

Tayeh, M.A. & Marletta, M.A. (1989) Macrophage oxidation of L-arginine to nitric oxide, nitrite, and nitrate. Tetrahydrobiopterin is required as a cofactor. *J. Biol. Chem.* **264**, 19654-19658.

Thoenes U., Flores O.L., Neves A., Devreese B., Van Beeumen J.J., Huber R., Romão M.J., LeGall J., Moura J., Rodrigues-Pousada C. (1994). Molecular cloning and sequence analysis of the gene of the molybdenum-containing aldehyde oxido-reductase of *Desulfovibrio gigas*. The deduced amino acid sequence shows similarity to xanthine dehydrogenase. *Eur. J. Biochem*. **220(3)**, 901-910.

Thony, B., Auerbach, G. & Blau, N. (2000). Tetrahydrobiopterin biosynthesis, regeneration and functions. *Biochem. J.* **347**, 1-16.

Triglia, T., Peterson, M.G., Kemp, D.J., (1988). A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucleic Acids Research*, **16**, 8186.

Tronrud D E (1992). Conjugate-direction minimization: an improved method for the refinement of macromolecules. *Acta Cryst.* **A48**, 912-916.

Truglio, J.J., Theis, K., Leimkühler, S., Rappa, R., Rajagopalan, K.V., and Kisker, C. (2002). Crystal structure of the active and alloxanthine-inhibited forms of xanthine dehydrogenase from *Rhodobacter capsulatus*. *Structure* **10**, 115–125.

Tsukihara T, Fukuyama K, Mizushima M, Harioka T, Kusunoki M, Katsube Y, Hase T and Matsubara H (1990). Structure of the [2Fe-2S] ferredoxin I from the blue-green alga Aphanothece sacrum at 2.2 A resolution. *J. Mol. Biol.* **216**, 399-410.

Turk, D. (1992). Development and usage of a macromolecular graphics program. Ph.D., Technical University Munich, Germany.

Vagin, A., Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022-1025.

Weygand, F., Simon, H., Dahms, G., Waldschmidt, M., Schliep, H. J. & Wacker, H. (1961). Über die Biogenese des Leucopterins. *Angew. Chem.* **73**, 402-407.

Wilson K. S., Butterworth S., Dauter Z., Lamzin V.S., Walsh M., Wodak S., Pontius J., Richelle J., Vaguine A., Sander C., Hooft R. W. W., Vriend G., Thornton J. M., Laskowski R. A., MacArthur M. W., Dodson E. J., Murshudow C., Oldfield T. J., Kaptein R., Rullman J. A. C. (1998). Who checks the checkers? Four validation tools applied to eight atomic resolution structures. *J. Mol. Biol*. **276**, 417-436.

Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W., Jr. (1995). Transition-state selectivity for a single hydroxyl group during catalysis by cytidine deaminase. *Biochemistry* **34**, 4516-4523.

Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W., Jr. (1996). Cytidine deaminase complexed to 3-deazacytidine: a "valence buffer" in zinc enzyme catalysis. *Biochemistry* **35**, 1335-1341.

Zamenhof P.J., Villarejo M. (1972). Construction and properties of *Escherichia coli* strains exhibiting α-complementation of β-galoctosidase fragments in vivo. *J. Bacteriol*. **110**, 171-178.

V Abbreviations

[CuSMo]	Copper-Sulphur-Molybdenum cluster
[FeS]	Iron-Sulphur cluster
Å	Angstron (10^{-10} m)
a u	Asymmetric unit
ADP	Atomic displacement parameters (temperature factor)
AfOR	Aldehyde ferredoxin oxidoreductase
AOR	Aldehyde oxidoreductase
ATP	Adenosine Triphosphate
B, B_{eq}	Isotropic atomic displacement parameter (temperature factor)
BH_4	(6r)-5,6,7,8-Tetrahydro-L-biopterin
CCD	Charge coupling device
CCP4	Collaborative computational project 4
CGLS	Conjugated gradient least squares
CNS	Crystallography and NMR system
CO	Carbon monoxide
CODH	Carbon monoxide dehydrogenase
CSD	Cambridge structural database
D.	*Desulfovibrio*
DESY	Deutsches Elektronen-Synchrotron
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
FAD	Flavin-adenine dinucleotide
FDH	Formate Dehydrogenase
FfOR	Formaldehyde ferredoxin oxidoreductase
FFT	Fast Fourier Transform
FR	Fourier ripples
GTP	Guanine triphosphate dinucleotide
H_4Folate	Tetrhydrofolate
H_4MPT	Tetrahydromethanopterin
HCl	Hydrochloride
HPLC	High pressue Liquid chromatography
IPP	Isoprpanol
IPTG	Isopropyl-beta-D-thiogalctopyranoside
KCl	PotassiumChloride
LB	Luria Bertani
LS	Least squares
M15[pREP4]	Cell strain M15 including plasmid pREP4
MAD	Multiwavelength anomalous dispersion
MCD	Molybdopterin cytosine dinucleotide
MGD	Molybdopterin guanine dinucleotide
MH	Molybdenum hydroxilases
MIR	Multiple isomorphous replacement
Mo	Molybdenum ion
Mo1 (/2)	Molybdenum binding domain 1 (/2)
Moco	Pyranopterin-ene-1,2-dithiolate cofactor or molybdopterin cofactor
MOD	Aldehyde oxidoreductase from *Desulfovibrius desulfuricans* ATCC 27774
MOP	Aldehyde oxidoreductase from *Desulfovibrio gigas*
MOPS	3-morpholinopropanesulfonic acid
MPD	2-Methyl-2,4-pentanediol
MPT	Monophosphate form of molybdopterin
MR	Molecular replacement
NADH	Nicotinamide adenin dinucleotide
NH_2TP	7,8-dihydroneopterin 3´-triphosphate
NOS	Nitric-oxide synthase
PAH	Phenylalanin-4-hydroxylase
PCR	Polymerase chain reaction
PEG	Polyethileneglycol
PPH_4	6-pyruvoyl-5,6,7,8-tetrahydropterin
SCOP	Structural classification of proteins
TH	Tyrosine-3-hydroxylase
TPH	Tryptophan-5-hydroxylase
U_{ij}	2^{nd} rank Tensor of atomic thermal/displacement parameters
XDH	Xanthine dehydrogenase
XO	xanthine oxidase