

Acta Crystallographica Section D

**Biological  
Crystallography**

ISSN 0907-4449

Editors: **E. N. Baker** and **Z. Dauter**

## **Structure of the molybdenum-cofactor biosynthesis protein MoaB of *Escherichia coli***

**Gerd Bader, Mariola Gomez-Ortiz, Christoph Haussmann, Adelbert Bacher,  
Robert Huber and Markus Fischer**

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site provided that this cover page is retained. Republication of this article or its storage in electronic databases or the like is not permitted without prior permission in writing from the IUCr.

Gerd Bader,<sup>a\*</sup> Mariola Gomez-Ortiz,<sup>a</sup> Christoph Haussmann,<sup>b</sup> Adelbert Bacher,<sup>b</sup> Robert Huber<sup>a</sup> and Markus Fischer<sup>b\*</sup>

<sup>a</sup>Abteilung Strukturforschung, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany, and <sup>b</sup>Lehrstuhl für Organische Chemie und Biochemie, Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany

Correspondence e-mail:  
gerd.bader@vie.boehringer-ingenheim.com,  
markus.fischer@ch.tum.de

## Structure of the molybdenum-cofactor biosynthesis protein MoaB of *Escherichia coli*

The *moaABC* operon of *Escherichia coli* is involved in early steps of the biosynthesis of the molybdenum-binding cofactor molybdopterin, but the precise functions of the cognate proteins are not known. The crystal structure of the MoaB protein from *E. coli* was determined by multiple anomalous dispersion at 2.1 Å resolution and refined to an *R* factor of 20.4% ( $R_{\text{free}} = 25.0\%$ ). The protein is a 32-symmetric hexamer, with the monomers consisting of a central  $\beta$ -sheet flanked by helices on both sides. The overall fold of the monomer is similar to those of the MogA protein of *E. coli*, the G-domains of rat and human gephyrin and the G-domains of Cnx1 protein from *A. thaliana*, all of which are involved in the insertion of an unknown molybdenum species into molybdopterin to form the molybdenum cofactor. Furthermore, the MoaB protein shows significant sequence similarity to the cinnamon protein from *Drosophila melanogaster*. In addition to other functions, all these proteins are involved in the biosynthesis of the molybdenum cofactor and have been shown to bind molybdopterin. The close structural homology to MogA and the gephyrin and Cnx1 domains suggests that MoaB may bind a hitherto unidentified pterin compound, possibly an intermediate in molybdopterin biosynthesis.

Received 15 December 2003

Accepted 25 March 2004

**PDB Reference:** MoaB, 1r2k,  
r1r2ksf.

### 1. Introduction

Molybdoenzymes play essential roles in the carbon, sulfur and nitrogen cycles in most organisms. In all molybdoenzymes known to date, with the sole exception of nitrogenases, the catalytically active centre is formed by the molybdenum cofactor (MoCo). It consists of a mononuclear molybdenum ion bound to the dithiolene moiety of a tricyclic pyranopterin derivative termed molybdopterin (MPT). While in eukaryotes only MPT has been observed as the organic part of the molybdenum cofactor, additional variability is achieved in bacteria through the attachment of a second nucleotide (GMP, AMP, IMP or CMP) to the phosphate group of MPT (Rajagopalan & Johnson, 1992). *Escherichia coli* has been shown to contain the molybdopterin guanidine dinucleotide (MGD; Eaves *et al.*, 1997). In the past few years, several structures of proteins containing molybdopterin have been determined (Chan *et al.*, 1995; Kisker, Schindelin, Pacheco *et al.*, 1997; Romao *et al.*, 1995), each of them representing one of the currently known families of MoCo-containing enzymes (Kisker, Schindelin & Rees, 1997). In mammals, MoCo is required for the activity of xanthine oxidase, aldehyde oxidase and sulfite oxidase. Deficiency in MoCo biosynthesis results in a fatal neurological disorder similar to the isolated form of sulfite oxidase deficiency and in early childhood death (Scriver *et al.*, 1988). Recently, the first genes involved in human molybdopterin biosynthesis and mutations within these genes

**Table 1**  
Bacterial strains and plasmids used in this study.

Strain/plasmid	Relevant characteristics	Source
<i>E. coli</i> strains		
XL1-Blue	recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac[F', proAB, lacI <sup>q</sup> ZΔM15, Tn10(tet <sup>r</sup> )]	Bullock <i>et al.</i> (1987)
M15[pREP4]	lac, ara, gal, mtl, recA <sup>+</sup> , uvr <sup>+</sup> , Str <sup>R</sup> , (pREP4:Kan <sup>R</sup> , lacI)	Stüber <i>et al.</i> (1990)
Expression plasmids		
pNCO113	<i>E. coli</i> expression vector	Bullock <i>et al.</i> (1987)
pNCO-EC-moaB	Expression vector for the overexpression of EC-moaB in <i>E. coli</i>	This study

**Table 2**  
Oligonucleotides used for the construction of the EC-moaB expression plasmid (restriction sites are in bold).

Designation	Sequence (5' to 3')
ECMOB-1	GAGGAGAAAATTACTATGATGAGTCAGGTAAGCACTGAATTTATC
ECMOB-2	TATTATGGATCCTTATTCTTCAAATGTGGATGGAATTAC
BSEcoRI	ATAATAGA <b>AATTC</b> ATTAAAGAGGAGAAATTAAC TATG

leading to molybdenum-cofactor deficiency have been identified (Reiss, Christensen *et al.*, 1998; Reiss, Cohen *et al.*, 1998).

The pathways leading to the formation of the molybdenum cofactor and its insertion into the corresponding apoenzymes to form active molybdoenzymes are not yet fully understood. Considerable progress towards an understanding of the reactions has been achieved through the use of mutants which are chlorate-resistant owing to mutations in the molybdoenzymes. These mutants, formerly referred to as *chl* mutants and now called *mo* mutants, helped to identify the genes involved in molybdenum-cofactor formation (Shanmugam *et al.*, 1992). The gene products of at least five operons are involved: *moa*, *mob*, *mod*, *moe* and *mog* proteins. The role of the respective gene-product transcripts in molybdopterin biosynthesis can be divided in two steps. In the second, presently quite well understood process, molybdopterin synthase, the heterodimer of MoaD and MoaE, catalyzes the conversion of precursor Z to molybdopterin by insertion of the dithiolene sulfurs into the pterin side chain. The MoaD moiety of molybdopterin synthase is in turn resulfurated by the molybdopterin synthase sulfurylase MoeB. The first step, the formation of precursor Z, is less well characterized. GTP (or another guanine derivative) serves as a substrate for MoaABC (Pitterle *et al.*, 1993; Rieder *et al.*, 1998; Rivers *et al.*, 1993) in a reaction that clearly differs from the pterin formation by GTP cyclohydrolase I (Rebelo *et al.*, 2003; Wuebbens & Rajagopalan, 1995), although for the yeast *Pichia canadensis* different findings have been reported (Irby & Adair, 1994). MoaA presumably catalyses the first step; it has been shown to be an Fe–S protein and has been proposed to act as an oxidoreductase (Menendez *et al.*, 1996). The crystal structure of the *E. coli* MoaC protein has been reported (Wuebbens *et al.*, 2000) and the residues necessary for catalytic activity have been identified, but the reaction

performed by the enzyme remained unclear. The role of the MoaB protein is presently not understood. To date, no MPT-deficient mutants with a mutation in the *moaB* gene have been identified, which may simply reflect the need to examine a larger number of mutants (Rajagopalan, 1997; Rivers *et al.*, 1993). MoaB protein shows significant homology to MogA protein, another protein of the molybdenum-cofactor biosynthesis, as well as to several eukaryotic proteins, *i.e.* gephyrin (*Rattus norvegicus*), cinnamon (*Drosophila melanogaster*) and Cnx1 (*Arabidopsis thaliana*).

Gephyrin is responsible for the postsynaptic anchoring of inhibitory glycine receptors to the cytoskeleton (Kirsch *et al.*, 1993) and the postsynaptic localization of major GABA<sub>A</sub>-receptor subtypes (Essrich *et al.*, 1998). Furthermore, gephyrin, as well as cinnamon and Cnx1, are involved in MPT biosynthesis (Schwarz *et al.*, 1997; Stallmeyer *et al.*, 1999; Wittle *et al.*, 1999). All these proteins consist of one domain with a high homology to *E. coli* MoaB and MogA and a second domain homologous to MoeA, another *E. coli* protein that is involved in MPT biosynthesis, even though the arrangement of the domains is not identical.

## 2. Materials and methods

### 2.1. Materials

Restriction enzymes and Vent DNA polymerase were purchased from New England Biolabs, Schwalbach, Germany. T4 DNA ligase was from Gibco BRL, Eggenstein, Germany. Taq polymerase was from Finnzyme, Espoo, Finland. DNA fragments were purified with the QIAquick PCR Purification Kit from Qiagen, Hilden, Germany. Oligonucleotides were custom-synthesized by MWG Biotech, Ebersberg, Germany.

### 2.2. Strains and plasmids

The bacterial strains and plasmids used in this study are summarized in Table 1.

### 2.3. Construction of an hyperexpression plasmid

PCR amplification was performed using chromosomal *E. coli* DNA as template and the oligonucleotides ECMOB-1 and ECMOB-2 as primers (Table 2). The amplificate served as template for a second PCR with the oligonucleotides BSEcoRI and ECMOB-2 as primers. The final 558 bp amplificate was digested with *EcoRI* and *BamHI* and ligated into the plasmid pNCO113 which had been treated with the same restriction enzymes. The resulting plasmid designated pNCO-EC-moaB was transformed into *E. coli* XL1-Blue cells using previously described procedures (Bullock *et al.*, 1987). Transformants were selected on LB agar plates supplemented with ampicillin (170 mg l<sup>-1</sup>). The plasmid was reisolated and transformed into *E. coli* M15[pREP4] cells (Stüber *et al.*, 1990) carrying the pREP4 repressor plasmid for the overexpression

of *lac* repressor protein. Kanamycin ( $20 \text{ mg l}^{-1}$ ) and ampicillin ( $170 \text{ mg l}^{-1}$ ) were added to secure the maintenance of both plasmids in the host strain.

## 2.4. Biosynthetic substitution of methionine by its analogue selenomethionine

For the incorporation of selenomethionine, the expression plasmid pNCO-EC-moaB was transformed into *E. coli* B834 cells, which are auxotrophic for methionine. The recombinant strain was grown in New Minimal Medium (NMM) under methionine-limited conditions with an excess of the analogue in the culture medium. *E. coli* B834 cells harbouring pNCO-EC-moaB were grown at 310 K to an  $\text{OD}_{600}$  of 0.5 in NMM containing  $7.5 \text{ mM}$  ammonium sulfate,  $8.5 \text{ mM}$  NaCl,  $55 \text{ mM}$  potassium dihydrogenphosphate,  $100 \text{ mM}$  potassium hydrogenphosphate,  $1 \text{ mM}$  magnesium sulfate and  $20 \text{ mM}$  glucose. The medium was supplemented with  $\text{Ca}^{2+}$ ,  $\text{Fe}^{2+}$  ( $1 \text{ mg l}^{-1}$ ), the trace elements  $\text{Co}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{MoO}_4^{2-}$  ( $1 \text{ } \mu\text{g l}^{-1}$ ), the vitamins thiamine and biotin ( $10 \text{ mg l}^{-1}$ ) and a mixture of all amino acids without methionine ( $50 \text{ mg l}^{-1}$ ).

For the preculture,  $50 \text{ mg l}^{-1}$  L-methionine was added to the medium. Cells were harvested and transferred into new medium containing  $100 \text{ mg l}^{-1}$  D,L-selenomethionine. The bacterial strain was grown under the selective pressure of  $150 \text{ mg l}^{-1}$  ampicillin. The expression of protein under the control of the T5 promoter was induced with  $2 \text{ mM}$  isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG) after the methionine supply had been exhausted. Purification of SeMet-MoaB was performed as described below.

## 2.5. Protein purification

The recombinant *E. coli* strain M15[pRep4]-pNCO-EC-moaB was grown in LB medium containing  $20 \text{ mg l}^{-1}$  kanamycin and  $170 \text{ mg l}^{-1}$  ampicillin. At an optical density (600 nm) of 0.8, IPTG was added to a final concentration of  $2 \text{ mM}$ . Incubation was continued for 18 h. The cells were harvested by centrifugation, washed twice with  $0.9\%$  NaCl and stored at 253 K.  $5 \text{ g}$  frozen cell mass was thawed in  $35 \text{ ml}$   $50 \text{ mM}$  Tris-HCl pH 7.8 containing  $1 \text{ mM}$  DTT and  $0.02\%$  sodium azide (buffer A). The suspension was cooled on ice and subjected to ultrasonic treatment (Branson-Sonifier B-12A, Branson SONIC Power Company, Dunbury, CT, USA). The suspension was centrifuged (Sorvall SS34 rotor,  $15\,000 \text{ rev min}^{-1}$ , 15 min, 277 K) and the supernatant was applied to a column of Q Sepharose FF (Pharmacia Biotech, Freiburg, Germany) equilibrated with buffer A. The column was developed using a gradient of  $0$ – $1 \text{ M}$  KCl in buffer A. MoaB protein was eluted at  $400 \text{ mM}$  KCl. Fractions containing the MoaB protein were collected and the solution was applied to a column of HA Macrorep  $40 \text{ } \mu\text{m}$  (Biorad) which had been equilibrated with buffer A. The column was developed using buffer A. The enzyme was not retarded under these conditions. The MoaB fractions were collected and concentrated by ultrafiltration (Macrosept  $10 \text{ kDa}$ ), yielding a solution containing  $1.5 \text{ mg ml}^{-1}$  protein.  $3 \text{ ml}$  aliquots of this solution were applied to a gel-filtration column (Superdex 200,

$2.6 \times 60 \text{ cm}$ , Pharmacia Biotech, Freiburg, Germany), which was then developed using buffer A. Selenomethionine incorporation was checked by electrospray mass spectrometry.

The native molecular weight was estimated using a Pharmacia FPLC system equipped with a Superdex 200 ( $1.6 \times 60 \text{ cm}$ , Pharmacia Biotech, Freiburg, Germany) column. The elution buffer was  $50 \text{ mM}$  Tris-HCl pH 7.8,  $100 \text{ mM}$  KCl and  $1 \text{ mM}$   $\text{MgCl}_2$ . The column was calibrated using the following standard proteins: RNase A ( $13.7 \text{ kDa}$ ), chymotrypsinogen A ( $25.7 \text{ kDa}$ ), ovalbumin ( $43.0 \text{ kDa}$ ), BSA ( $67.0 \text{ kDa}$ ), aldolase ( $156 \text{ kDa}$ ), catalase ( $232 \text{ kDa}$ ) and apoferritin ( $440 \text{ kDa}$ ).

## 2.6. Estimation of protein concentration

Protein concentration was estimated by a modified Bradford procedure (Read & Northcote, 1981).

## 2.7. DNA and protein sequencing

DNA was sequenced by the automated dideoxynucleotide method (Sanger *et al.*, 1977) using a Prism 377 sequencer from Applied Biosystems (Foster City, CA, USA). Protein was sequenced by the automated Edman method using a 471A sequencer from Perkin-Elmer (Weiterstadt, Germany).

## 2.8. Polyacrylamide gel electrophoresis

SDS-PAGE was performed as described by Laemmli (1970). Molecular-weight standards were supplied by Sigma (Munich, Germany).

## 2.9. Crystallization

Crystals were prepared by the sitting-drop vapour-diffusion method using  $5 \text{ } \mu\text{l}$   $14.5 \text{ mg ml}^{-1}$  MoaB protein solution in  $25 \text{ mM}$  Tris-HCl pH 8.0 containing  $100 \text{ mM}$  NaCl and  $1 \text{ } \mu\text{l}$  reservoir solution. Droplets were equilibrated against a reservoir solution containing  $0.6 \text{ M}$   $(\text{NH}_4)_2\text{SO}_4$ ,  $1.25 \text{ M}$   $\text{Li}_2\text{SO}_4$  and  $125 \text{ mM}$  citrate pH 5.6. Crystals grew within 4 d at room temperature to dimensions of  $200 \times 300 \times 900 \text{ } \mu\text{m}$  and belonged to space group *P*321 (unit-cell parameters  $a = b = 69.68$ ,  $c = 125.82 \text{ } \text{Å}$ ), with two MoaB subunits in the asymmetric unit.

## 2.10. Data collection

Crystals were transferred to cryobuffer [ $0.6 \text{ M}$   $(\text{NH}_4)_2\text{SO}_4$ ,  $1.25 \text{ M}$   $\text{Li}_2\text{SO}_4$ ,  $125 \text{ mM}$  citrate pH 5.6,  $25\%$  (*w/v*) sucrose] and flash-cooled at liquid-nitrogen temperature. MAD diffraction data were collected at beamline BW6, DESY, Hamburg using a  $135 \text{ mm}$  MAR CCD detector (MAR Research, Hamburg). Wavelengths were as follows:  $\lambda_1 = 0.9792$ ,  $\lambda_2 = 0.97976$ ,  $\lambda_3 = 1.20 \text{ } \text{Å}$ . Data sets were processed and scaled using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

## 2.11. Structure solution and refinement

The positions of the Se atoms were determined using *SOLVE* (Terwilliger & Berendzen, 1999) and further refined using *SHARP* (de La Fortelle & Bricogne, 1997). Solvent flattening was performed with *SOLOMON* (Abrahams &

**Table 3**  
Refinement statistics.

Values in parentheses are for the highest resolution shell.

	Peak	Edge	Remote
Data-collection statistics			
Wavelength (Å)	0.97920	0.97976	1.20
Resolution (Å)	2.1 (2.17–2.10)	2.1 (2.17–2.10)	2.1 (2.17–2.10)
Total data	276109	276995	275848
Unique data	21651	21615	21683
Completeness (%)	99.3 (95.2)	98.1 (93.5)	98.3 (93.7)
$R_{\text{sym}}^{\dagger}$ (%)	4.3 (16.3)	4.5 (19.7)	4.5 (25.9)
Phasing statistics			
Mean FOM			0.47
$R_{\text{cullis}}^{\ddagger}$	0.35	0.37	—
Refinement statistics			
$R_{\text{cryst}}^{\S}$ (%)	—	—	20.45
$R_{\text{free}}^{\P}$ (%)	—	—	25.05
R.m.s.d. bonds (Å)	—	—	0.010
R.m.s.d. angles (°)	—	—	1.47
Ramachandran plot quality			
Most favoured	—	—	89.9
Allowed	—	—	9.9
Generously favoured	—	—	0.3

$\dagger R_{\text{sym}} = \frac{\sum_{hkl} \sum_i (I(hkl)_i) - \langle I(hkl) \rangle / \sum_{hkl} \sum_i \langle I(hkl)_i \rangle}{\sum_{hkl} \sum_i |F_{\text{calc}}(hkl)| \pm |F_{\text{p}}(hkl)|} - \frac{F_{\text{calc}}(hkl) / \sum_{hkl} |F_{\text{p}}(hkl)| \pm F_{\text{p}}(hkl)}{\sum_{hkl} |F_{\text{obs}}(hkl)|}$ .  $\ddagger R_{\text{cullis}} = \frac{\sum_{hkl} |F_{\text{p}}(hkl)| \pm |F_{\text{calc}}(hkl)|}{\sum_{hkl} |F_{\text{obs}}(hkl)|}$ .  $\S R_{\text{cryst}} = \frac{\sum_{hkl} |F_{\text{obs}}(hkl)| - |F_{\text{calc}}(hkl)|}{\sum_{hkl} |F_{\text{obs}}(hkl)|}$ .  $\P$  As for  $R_{\text{cryst}}$  with 5% of the reflections excluded from refinement calculation.

Leslie, 1996). The resulting electron-density map allowed the tracing of the complete model. Multiple rounds of model building using *MAIN* (Turk, 1992) and refinement using *CNS* (Brünger *et al.*, 1998) resulted in a final *R* factor of 20.4% and an  $R_{\text{free}}$  of 25.0% (calculated from a test set containing 5% of all reflections). All refinement was performed using the maximum-likelihood target as implemented in *CNS*. The final model contains two MoaB subunits with 2597 protein atoms, 150 water molecules and two sulfate ions.

In the Ramachandran plot for the 300 non-glycine and non-proline residues as calculated with *PROCHECK* (Laskowski *et al.*, 1993), 89.9% of the residues are in the most favoured regions, 9.9% in additionally allowed regions and 0.3% in generously allowed regions. The oligomer interfaces were analysed using <http://www.biochem.ucl.ac.uk/bsm/PP/server> (Jones & Thornton, 1995). Structure comparisons were performed using *TOP* from the *CCP4* suite (Collaborative Computational Project, Number 4, 1994).

### 3. Results and discussion

#### 3.1. Cloning and purification

The *moaB* gene was cloned from *E. coli* and the protein was labelled with selenomethionine by expression in a methionine-auxotrophic *E. coli* strain. MoaB protein was purified in a three-step procedure. The native hexameric structure of the protein in solution was determined by gel-filtration chromatography calibrated with standard proteins.

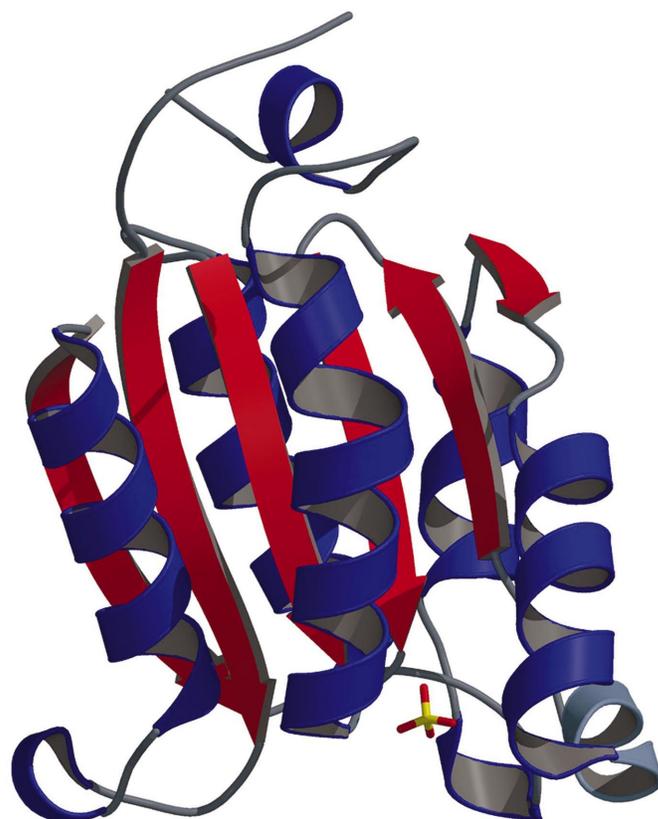
#### 3.2. Structure solution and refinement

Recombinant MoaB protein was crystallized from a mixture of ammonium sulfate and lithium sulfate in space group *P321*,

with two monomers in the asymmetric unit that were related by twofold non-crystallographic symmetry. The crystal symmetry operations form the observed hexamer with  $D_{32}$  symmetry. MAD data sets to 2.1 Å resolution were collected from a crystal of the selenomethionine-substituted protein using synchrotron radiation at three wavelengths. The selenium sites were located using automated Patterson methods. MAD phasing with four selenium sites and subsequent solvent flattening resulted in a figure of merit of 0.71 and an electron-density map that was traceable throughout the model. The structure was refined to a final  $R_{\text{cryst}}$  of 20.4% ( $R_{\text{free}} = 25.0$ ). The r.m.s.d. values are 0.010 Å for bond lengths and 1.47° for bond angles. 89.9% of the residues are in the most favoured regions of the Ramachandran plot (Ramachandran & Sasi-sekharan, 1968). The two monomers in the asymmetric unit can be superimposed with a r.m.s.d. of 0.27 Å (for  $C^{\alpha}$  atoms). Data-collection, phasing and refinement statistics are summarized in Table 3. Despite the high homology of MoaB to the proteins mentioned below solution of the structure by molecular replacement was not attempted, as the homologous structures had not yet become accessible at the time of structure solution.

#### 3.3. Overall structure

MoaB protein has an  $\alpha/\beta/\alpha$  structure and is composed of a central predominantly parallel six-stranded  $\beta$ -sheet ( $\beta 1$ – $\beta 6$ ) surrounded by two  $\alpha$ -helices on one side and four  $\alpha$ -helices



**Figure 1**  
Ribbon representation of a MoaB monomer.

and a  $3_{10}$ -helix on the other side. The only antiparallel strand of the  $\beta$ -sheet,  $\beta 5$ , is located near the edge of the sheet. Helix  $\alpha 4$  (residues 99–113) and the  $3_{10}$ -helix form one long quasi-continuous but strongly kinked helix. On the opposite side of the  $\beta$ -sheet, residues 139–154 form  $\alpha 7$ , a continuous helix that is bent and somewhat unwound at Ile150 (Fig. 1).

The crystal structure confirms that MoaB is a hexamer, or more exactly a dimer of trimers, with 32 symmetry (Fig. 2). The trimer motif is formed by residues located in helix  $\alpha 6$ , the following  $3_{10}$ -helix, helix  $\alpha 4$  and residues Pro160–His166, including the beginning of helix  $\alpha 6$ . Approximately  $1480 \text{ \AA}^2$  of accessible surface area is buried upon trimerization, which corresponds to 18% of the surface of each monomer. The hexamer is made up of two trimers mainly through the interaction of helix  $\alpha 4$  with the reversely oriented helix  $\alpha 4$  of another subunit and the residues Pro92 and Leu93 in the loop connecting  $\alpha 4$  and  $\beta 3$ . Here, only  $970 \text{ \AA}^2$  (corresponding to 11.8% of the monomer surface) becomes buried upon oligomerization, supporting the notion that MoaB is a dimer of trimers.

### 3.4. Comparison to MogA and the gephyrin and CNX1 G-domains

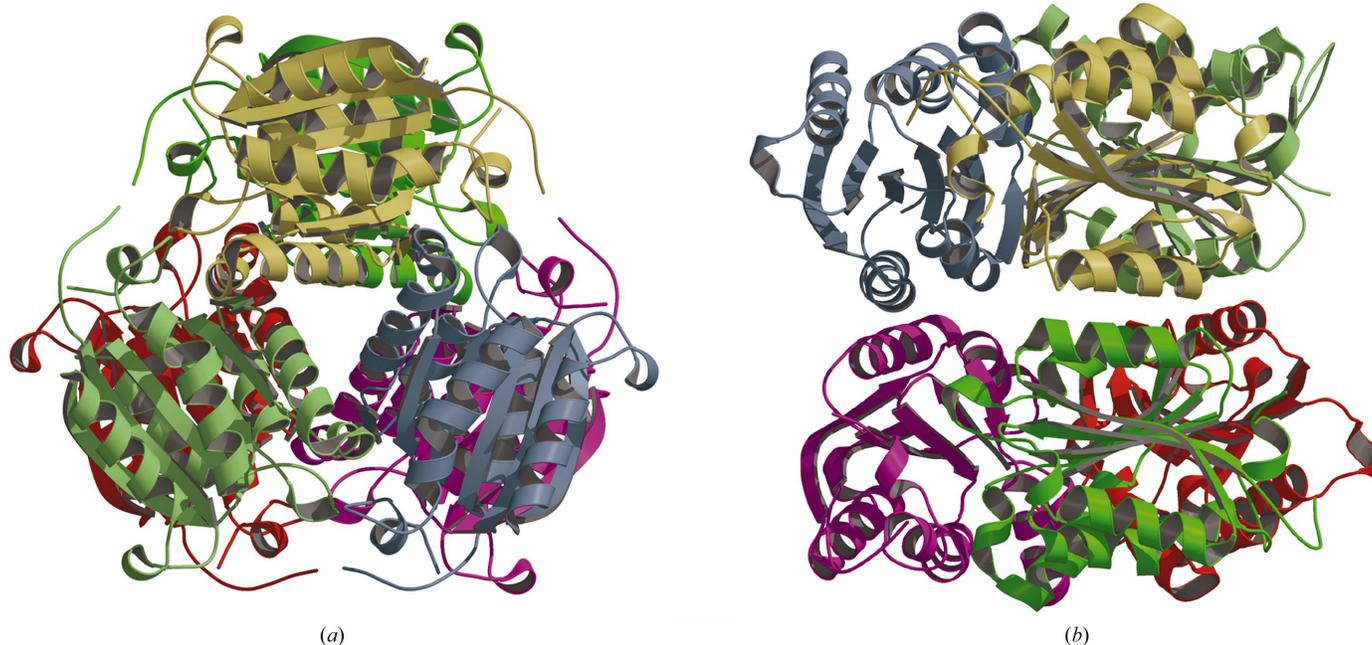
Several proteins show sequence similarity to MoaB, the best known of which are MogA, which is 27.4% identical and 36.9% similar, gephyrin from *R. norvegicus* (30.4% identical, 41.8% similar in the corresponding domains) and Cnx1 (*A. thaliana*; 35.9% identical and 45.7% similar). Others include cinnamon (*D. melanogaster*; 30.3% identical and 41.4% similar). Fig. 3 shows a structure-based sequence alignment of MoaB with MogA and the G-domain of gephyrin and Cnx1.

The molybdopterin-cofactor biosynthesis protein signature 1 (Prosite accession No. PS01078; <http://www.expasy.ch/prosite/>), VVLITGGTG in the case of MoaB, is conserved in all sequences (Fig. 3).

MogA has been proposed to act as a molybdochelataase, incorporating molybdenum into molybdopterin, owing to the observation that *mogA* mutants can be rescued by high concentrations of molybdate (Joshi *et al.*, 1996). Furthermore, MogA has been shown to bind molybdopterin (Pitterle & Rajagopalan, 1989). The crystal structure of *E. coli* MogA and accompanying biochemical investigations identified the residues that are important for molybdopterin binding (Liu *et al.*, 2000) and catalytic activity, but failed to demonstrate molybdate binding; the exact function of the protein could not yet be elucidated.

Gephyrin, a 93 kDa protein, has been shown to play a crucial role in the clustering of the glycine receptor at the postsynapse by anchoring it to the subsynaptic cytoskeleton (Kirsch & Betz, 1995). In addition to this function, gephyrin has been shown to be involved in the biosynthesis of the molybdenum cofactor. The structures of the N-terminal G-domain of rat gephyrin (Sola *et al.*, 2001) and its human orthologue, which is identical in sequence (Schwarz *et al.*, 2001), have recently been solved and show only minor differences in the position of the C-terminus.

As can be seen in Fig. 4 (overlay and topology), the structure and topology are MoaB is remarkably similar to those of its homologues. Topologically, the only significant difference between the monomeric structures of MoaB and MogA is the insertion of the two  $\beta$ -strands before the C-terminal helix of MoaB. The two structures can be superimposed with an r.m.s.d. of  $1.7 \text{ \AA}$  for 147  $C^\alpha$  atoms. The N-terminal parts of the two structures show an almost identical structure, whereas the



**Figure 2** Hexameric assembly of *E. coli* MoaB viewed (a) along the threefold crystallographic symmetry axis and (b) perpendicular to the threefold axis.

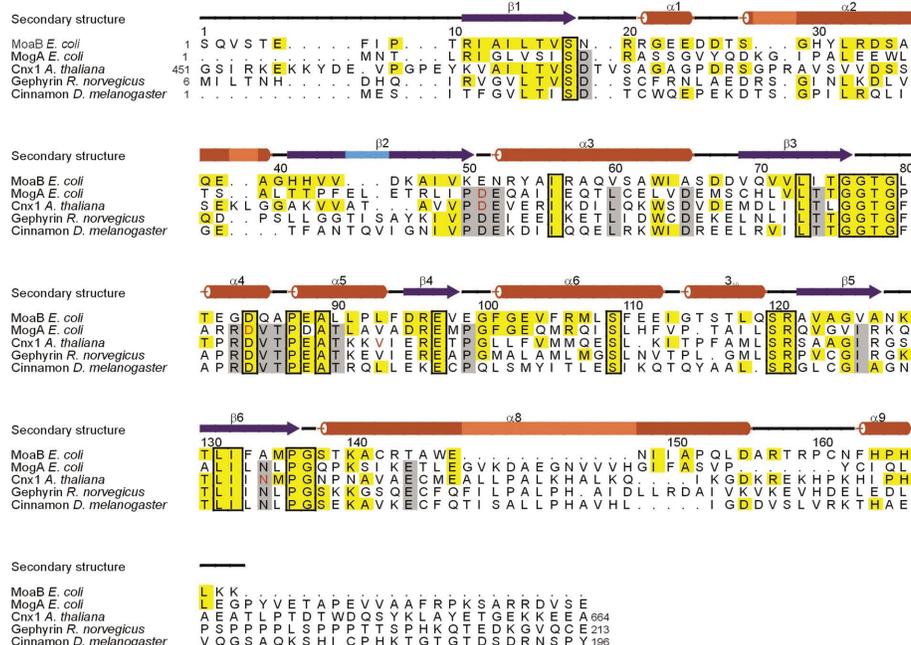
C-terminal parts differ significantly. In the MoaB structure,  $\alpha 7$  is a long bent helix that is also present in MogA, with the difference that 12 residues are inserted in MogA at the position of the kink, where they form a  $\beta$ -sheet protruding perpendicularly from the helix. Closer to the C-terminus, the structures become quite different, but interestingly the C-terminus of MogA lies in almost the same position as the MoaB N-terminus. The structures of the gephyrin and Cnx1 G-domains are even more homologous to MoaB in that they also lack the inserted  $\beta$ -strands in the C-terminal helix.

Both proteins form trimers in a similar way; most residues that make up the trimer contacts are either conserved or conservatively substituted, namely residues Leu80, Val96, Phe101, Gly114 and Leu118. This also holds true for the other homologous proteins (Fig. 3). Contrary to MogA, gephyrin and Cnx1, which form a trimer in solution, MoaB forms a hexamer or rather a dimer of trimers. On first sight, the parts of MoaB forming the hexamer contact seem to be identical to the corresponding parts of MogA, but a closer inspection reveals some definite differences in the residues directly

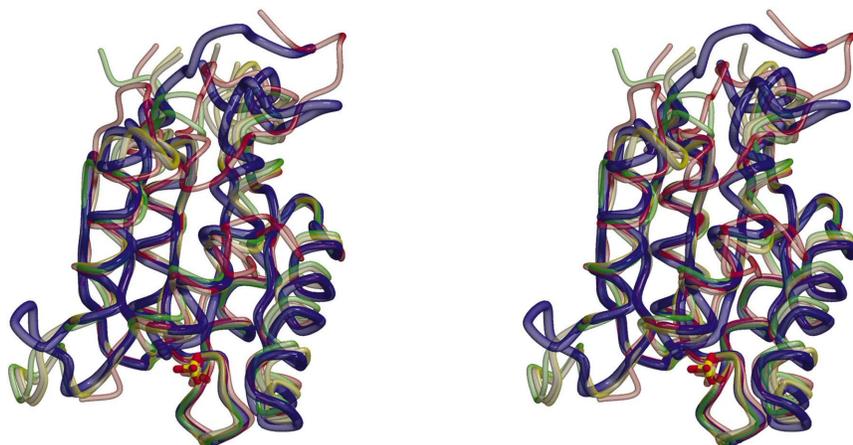
involved in protein–protein interaction. Tyr54 is replaced by Ala52, which prevents the formation of a hydrogen bond, Ala58 is replaced by the significantly larger Gln56 (Ala65→Asp63) and, probably most importantly, Ser66 is replaced by Glu64, which would, owing to slight changes in the overall structure, point directly into the potential interface and collide with Gln56, obstructing possible hexamer formation. Furthermore, the side chain of MogA Lys125 might interfere with the contact surface.

In the MogA structure, two residues that are strictly conserved in the homologous proteins (Fig. 3), Asp49 and Asp82, have been shown to be of functional importance (Liu *et al.*, 2000). Mutation of either of these residues to alanine led to a protein that was unable to complement a mutant *E. coli* strain in which the chromosomal copy of MoaA had been disrupted. Furthermore, this mutation inhibited an apo nitrate reductase reconstitution assay, which was ascribed to closer molybdopterin binding in these mutants. The corresponding residues in MoaB are Glu51 and Asp84, respectively. In the MogA structure, a sulfate molecule has been modelled into the density close to these residues, which are located near the conserved TXGGTG-motif. Additional density features connected to this sulfate molecule have been attributed to possible residual molybdopterin binding. Interestingly, in the MoaB structure a bound sulfate molecule has been found at an almost identical position.

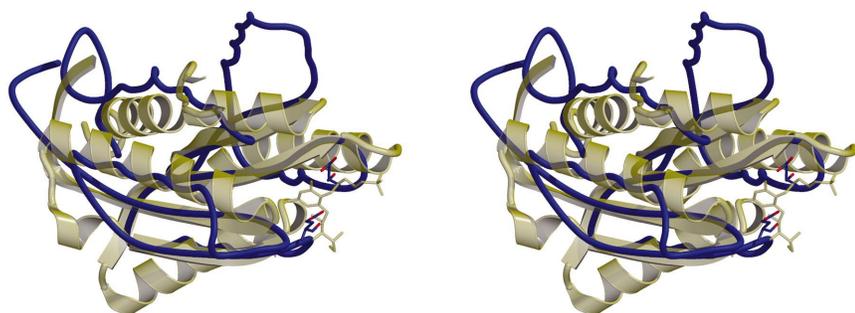
A recent biochemical investigation (Kuper *et al.*, 2000) identified on the basis of chlorate resistance some residues that are indispensable for the function of the *A. thaliana* Cnx1 G-domain, which is homologous to MoaB/MogA, and again the corre-



**Figure 3** Secondary-structure assignment of MoaB of *E. coli* and sequence alignment with homologous proteins. The numbering above the alignment corresponds to MoaB of *E. coli*. Secondary-structure elements ( $\beta$ -strands, blue;  $\alpha$ -helices, red) found in *E. coli* MoaB are shown above the sequences. Residues found to be crucial for function are marked in red. Identical residues in all shown sequences are shown in black boxes on a yellow background; conserved residues in most of the sequences are shown on a yellow background. The figure was drawn with *ALSCRIPT* (Barton, 1993).



**Figure 4** Overlay of all MoaB homologues: MoaB (blue), MogA (red), Cnx1 G-domain (green), rat gephyrin (brown) and human gephyrin (grey). The sulfate ions conserved in MoaB and MogA are depicted in yellow and red, respectively.



**Figure 5**

Overlay of MoaB from *E. coli* (yellow) with lumazine synthase from *S. cerevisiae* (blue). The intermediate analogue is depicted in yellow. The amino acids Glu51 and Asp84 above and below the analogue are shown in dark grey.

sponding residues in MoaB and MogA are located at identical positions. The G-domain has been shown to bind MPT, to be responsible for the insertion of molybdenum into MPT and to stabilize the formed molybdenum cofactor. The mutations Asp515Asn and Asp515His (Glu51 in MoaB, Asp49 in MogA) resulted in unaltered MPT binding, pointing to a function distinct from MPT binding, possibly the insertion of molybdenum into molybdopterin, which is in agreement with the findings in the case of MogA (Liu *et al.*, 2000). Val557Gly (Leu93 in MoaB, Val91 in MogA) and Asn597Leu (Ala134 in MoaB, Asn131 in MogA) showed a decrease in molybdopterin binding and a loss of MoCo stabilization. As the latter two residues seem to mediate interactions between secondary-structure elements, the mutants were suggested to influence the formation of the substrate-binding pocket. Unfortunately, the role of Asp548, which corresponds to the catalytically active Asp82 of MogA (Asp84 in MoaB), has not been investigated in this study.

### 3.5. Structural homologues

A search for structural homologues with the program DALI (Holm & Sander, 1995) revealed, aside from the close homology to *E. coli* MogA, a large number of distantly related structures (302 structures with a *Z* score greater than 3). The most remarkable of these homologues are the lumazine synthases (Braden *et al.*, 2000; Holm & Sander, 1995; Ladenstein *et al.*, 1988; Ritsert *et al.*, 1995). Lumazine synthase from *Brucella abortus* (Braden *et al.*, 2000) has a *Z* score of 8.7 and 116 C $\alpha$  atoms can be overlaid with an r.m.s.d. of 2.7 Å<sup>2</sup>; that of *Saccharomyces cerevisiae* has a *Z* score of 6.8 (the highest scoring protein not involved in molybdopterin biosynthesis). In this case, 114 C $\alpha$  atoms can be overlaid with an r.m.s.d. of 3 Å<sup>2</sup>. Lumazine synthase catalyzes the formation of 6,7-dimethyl-8-ribityllumazine, a compound that contains a pterin moiety and is thus very similar to molybdopterin or precursor Z. An overlay of *E. coli* MoaB with the crystal structure of *S. cerevisiae* lumazine synthase in complex with a transition-state intermediate analogue (Meining *et al.*, 2000) places the pyrimidine ring of the analogue exactly between MoaB Glu51 and Asp84 in the presumed binding pocket (Fig. 5). The pterin ring of dimethylribityllumazine would come to lie in the same position.

### 3.6. Concluding remarks

Even though the exact biochemical function of MoaB has not yet been elucidated, several features implying a possible function of MoaB can be found. The localization of the gene in the *moa* operon and the close homology in sequence and structure to the corresponding enzymes of other species clearly indicate involvement in molybdopterin metabolism. The presence of sequence signatures characteristic of molybdopterin biosynthesis and sequence homologies to folate-binding proteins also point in the same direction. MoaB mutants

with a deficiency in molybdopterin have not been reported. Hence, it is unknown whether or not the *moaB* gene is essential for molybdopterin biosynthesis, but the fact that no *mo* mutants with a defective MoaB have been identified thus far suggests that it is not. In the light of the previously mentioned facts, one might speculate that MoaB transiently binds a possibly unstable intermediate of molybdopterin biosynthesis.

We thank Dr Hans D. Bartunik and Gleb P. Bourenkov for assistance during crystallographic data collection at the BW6 beamline at DESY (Hamburg) and Dr Andreas Giessauf for help with protein purification. We also thank Dr I. Haase for a critical review of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie and by the European Community Grants ERB CHRX CT93-0243 and ERB FMRX-CT98-0204.

### References

- Abrahams, J. & Leslie, A. (1996). *Acta Cryst.* **D52**, 30–42.  
 Barton, G. J. (1993). *Protein Eng.* **6**, 37–40.  
 Braden, B. C., Vekilovsky, C. A., Cauerhff, A. A., Polikarpov, I. & Goldbaum, F. A. (2000). *J. Mol. Biol.* **297**, 1031–1036.  
 Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.  
 Bullock, W. O., Fernandez, J. M. & Short, J. M. (1987). *Biotechniques*, **5**, 376–379.  
 Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. & Rees, D. C. (1995). *Science*, **267**, 1463–1469.  
 Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.  
 Eaves, D. J., Palmer, T. & Boxer, D. H. (1997). *Eur. J. Biochem.* **246**, 690–697.  
 Essrich, C., Lorez, M., Benson, J. A., Fritschy, J. M. & Luscher, B. (1998). *Nature Neurosci.* **1**, 563–571.  
 Holm, L. & Sander, C. (1995). *Trends Biochem. Sci.* **20**, 478–480.  
 Irby, R. B. & Adair, W. L. Jr (1994). *J. Biol. Chem.* **269**, 23981–23987.  
 Jones, S. & Thornton, J. M. (1995). *Prog. Biophys. Mol. Biol.* **63**, 31–65.  
 Joshi, M. S., Johnson, J. L. & Rajagopalan, K. V. (1996). *J. Bacteriol.* **178**, 4310–4312.  
 Kirsch, J. & Betz, H. (1995). *J. Neurosci.* **15**, 4148–4156.  
 Kirsch, J., Wolters, I., Triller, A. & Betz, H. (1993). *Nature (London)*, **366**, 745–748.

- Kisker, C., Schindelin, H., Pacheco, A., Wehbi, W., Garrett, R., Rajagopalan, K., Enemark, J. & Rees, D. (1997). *Cell*, **91**, 973–983.
- Kisker, C., Schindelin, H. & Rees, D. C. (1997). *Annu. Rev. Biochem.* **66**, 233–267.
- Kuper, J., Palmer, T., Mendel, R. R. & Schwarz, G. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 6475–6480.
- Ladenstein, R., Schneider, M., Huber, R., Bartunik, H. D., Wilson, K., Schott, K. & Bacher, A. (1988). *J. Mol. Biol.* **203**, 1045–1070.
- Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
- La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
- Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). *J. Mol. Biol.* **231**, 1049–1067.
- Liu, M. T. W., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. (2000). *J. Biol. Chem.* **275**, 1814–1822.
- Meining, W., Mortl, S., Fischer, M., Cushman, M., Bacher, A. & Ladenstein, R. (2000). *J. Mol. Biol.* **299**, 181–197.
- Menendez, C., Siebert, D. & Brandsch, R. (1996). *FEBS Lett.* **391**, 101–103.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pitterle, D. M., Johnson, J. L. & Rajagopalan, K. V. (1993). *J. Biol. Chem.* **268**, 13506–13509.
- Pitterle, D. M. & Rajagopalan, K. V. (1989). *J. Bacteriol.* **171**, 3373–3378.
- Rajagopalan, K. V. (1997). *Biochem. Soc. Trans.* **25**, 757–761.
- Rajagopalan, K. V. & Johnson, J. L. (1992). *J. Biol. Chem.* **267**, 10199–10202.
- Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–438.
- Read, S. M. & Northcote, D. H. (1981). *Anal. Biochem.* **116**, 53–64.
- Rebelo, J., Auerbach, G., Bader, G., Bracher, A., Nar, H., Hosl, C., Schramek, N., Kaiser, J., Bacher, A., Huber, R. & Fischer, M. (2003). *J. Mol. Biol.* **326**, 503–516.
- Reiss, J., Christensen, E., Kurlemann, G., Zobot, M. T. & Dorche, C. (1998). *Hum. Genet.* **103**, 639–644.
- Reiss, J., Cohen, N., Dorche, C., Mandel, H., Mendel, R. R., Stallmeyer, B., Zobot, M. T. & Dierks, T. (1998). *Nature Genet.* **20**, 51–53.
- Rieder, C., Eisenreich, W., O'Brien, J., Richter, G., Gotze, E., Boyle, P., Blanchard, S., Bacher, A. & Simon, H. (1998). *Eur. J. Biochem.* **255**, 24–36.
- Ritsert, K., Huber, R., Turk, D., Ladenstein, R., Schmidt-Base, K. & Bacher, A. (1995). *J. Mol. Biol.* **253**, 151–167.
- Rivers, S. L., McNairn, E., Blasco, F., Giordano, G. & Boxer, D. H. (1993). *Mol. Microbiol.* **8**, 1071–1081.
- Romao, M. J., Archer, M., Moura, I., Moura, J. J., LeGall, J., Engh, R., Schneider, M., Hof, P. & Huber, R. (1995). *Science*, **270**, 1170–1176.
- Sanger, F., Niklen, S. & Coulson, A. R. (1977). *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Schwarz, G., Boxer, D. H. & Mendel, R. R. (1997). *J. Biol. Chem.* **272**, 26811–26814.
- Schwarz, G., Schrader, N., Mendel, R. R., Hecht, H.-J. & Schindelin, H. (2001). *J. Mol. Biol.* **312**, 405–418.
- Scriven, C. R., Kaufman, S. & Woo, S. L. (1988). *Annu. Rev. Genet.* **22**, 301–321.
- Shanmugam, K. T., Stewart, V., Gunsalus, R. P., Boxer, D. H., Cole, J. A., Chippaux, M., DeMoss, J. A., Giordano, G., Lin, E. C. & Rajagopalan, K. V. (1992). *Mol. Microbiol.* **6**, 3452–3454.
- Sola, M., Kneussel, M., Heck, I. S., Betz, H. & Weissenhorn, W. (2001). *J. Biol. Chem.* **276**, 25294–25301.
- Stallmeyer, B., Schwarz, G., Schulze, J., Nerlich, A., Reiss, J., Kirsch, J. & Mendel, R. R. (1999). *Proc. Natl. Acad. Sci. USA*, **96**, 1333–1338.
- Stüber, D., Matile, H. & Garotta, G. (1990). *Immunological Methods IV*, edited by I. Lefkovits & P. Pernis, pp. 121–125. New York: Academic Press.
- Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* **D55**, 849–861.
- Turk, D. (1992). PhD thesis. Technische Universität München, Germany.
- Wittle, A. E., Kamdar, K. P. & Finnerty, V. (1999). *Mol. Gen. Genet.* **261**, 672–680.
- Wuebbens, M. M., Liu, M. T., Rajagopalan, K. & Schindelin, H. (2000). *Structure*, **8**, 709–718.
- Wuebbens, M. M. & Rajagopalan, K. V. (1995). *J. Biol. Chem.* **270**, 1082–1087.