Acta Cryst. (1998). D**54**, 684–686

Preliminary crystallographic investigations of recombinant GDP-4-keto-6-deoxy-D-mannose epimerase/reductase from *E. coli*

M. Tonetti,[a] M. Rizzi,[b,c] P. Vigevani,[c] L. Sturla,[a] A, Bisso,[a] A. De Flora[a] and M. Bolognesi[d]* at [a]*Istituto Policattedra di Chimica Biologica, Universita' di Genova, Viale Benedetto XV, 1. I-16132 Genova, Italy,* [b]*Dipartimento di Scienza e Tecnologia del Farmaco, Universita' di Torino, Via Pietro Giuria, 9. I-10125 Torino Italy,* [c]*Dipartimento di Genetica e Microbiologia, Universita' di Pavia, Via Abbiategrasso, 207. I-27100 Pavia Italy, and* [d]*Centro Biotecnologie Avanzate - IST and Dipartimento di Fisica - INFM Universita' di Genova, Largo Rosanna Benzi, 10. I-16132 Genova Italy. E-mail: bolognesi@genova.infn.it*

(*Received 3 October 1997; accepted 9 December 1997*)

Abstract

The GDP-4-keto-6-deoxy-D-mannose epimerase/reductase (GM_ER) isolated from *E. coli* has been overexpressed as a GST-fusion protein and purified to homogeneity. The enzyme, an $NADP^+$(H)-binding homodimer of 70 kDa, is responsible for the production of GDP-L-fucose. GM_ER shows significant structural homology to the human erythrocyte protein FX, which is involved in blood-group glycoconjugate biosynthesis, displaying 3,5 epimerase/reductase activity on GDP-4-keto-6-deoxy-D-mannose. GM_ER has been crystallized in a trigonal crystalline form, containing one molecule per asymmetric unit, suitable for high-resolution crystallographic investigations.

1. Introduction

Fucosylated glycolipids and glycoproteins have been recognised as cell-surface molecules playing fundamental roles in several physio-pathological processes, including blood group determination, tumorigenicity and neutrophil adhesion (Feizi & Childs, 1987). In bacterial cells, L-fucose is a component of capsular antigens and of lipopolysaccharides of Gram-negative microorganisms (Ørskov & Ørskov, 1992). The multi-step biosynthetic pathway converting GDP-D-mannose into GDP-L-fucose, the main substrate for fucosyltransferase enzymes involved in the modification of cell-surface structures, has been described in both prokaryotes and eukaryotes, and shown to be evolutionarily conserved. Nevertheless, the precise structures and mechanisms of the enzymes involved in GDP-L-fucose biosynthesis have received little attention so far.

A 68 kDa homodimeric $NADP^+$(H)-binding protein from human erythrocytes, known as FX, has recently been shown to display both 3,5 epimerase and reductase activities on GDP-4-keto-6-deoxy-D-mannose (Tonetti *et al.*, 1996). This is the product of GDP-D-mannose dehydratase, the first enzyme in the GDP-L-fucose biosynthetic pathway (Sturla *et al.*, 1997). FX displays a clear amino-acid sequence relationship with UDP-galactose 4-epimerase from *E. coli*, a homodimeric enzyme belonging to the superfamily of short-chain dehydrogenases (Jörnvall *et al.*, 1995), built up by a NAD^+(H) binding (Rossmann fold) *N*-terminal domain, coupled to a smaller *C*-terminal mixed α/β domain, responsible for substrate binding (Bauer *et al.*, 1991, 1992). Moreover, human FX displays very high amino-acid sequence homology with respect to the murine tumor antigen P35B [92.6% amino-acid identities (Camardella *et al.*, 1995)]. Lower, but significant, sequence relationships have also been found between FX and the nodulating protein *Nolk* from *Azorhizobium caulinodans* [which has recently been shown to be involved in GDP-L-fucose production; 26% amino-acid identities (Mergaert *et al.*, 1997)] and the *yefb* gene product protein from *E.coli* [involved in the metabolism of GDP-D-mannose and object of the present study; 27% amino-acid identities, see Fig. 1 (Tonetti *et al.*, 1996)]. Notably, the *E. coli yefb* gene is located in a cluster

```
GM_ER   MSKQ----RVFIAGHRGMVGSAIRRQLEQRGD---------------V---ELV---LRT
FX      MGEPQGSMRILVTGGSGLVGKAIQKVVADGAG---------------LPGEDWV---FVS
UG_ER   M-------RVLVTGGSGYIGSHTCVQLLQNGHDVIILDNLCNSKRSVLPVIERLGGKHPT
        *       *....*  * .*      . . .                .   . .     .

GM_ER   RDELNLLDSRAVHDFFASERIDQVYLAAAKVGGIVANNTYPADFIYQNMMIESNIIHAAH
FX      SKDADLTDTAQTRALFEKVQPTHVIHLAAMVGGLFRNIKYNLDFWRKNVHMNDNVLHSAF
UG_ER   FVEGDIRNEALMTEILHDHAIDTVIHFAG-LKAVGESVQKPLEYYDNNVNGTLRLISAMR
          . .. .      ..       *   *. . ..  .     ..   .*.     .. .

GM_ER   QNDVNKLLFLGSSCIYPKLAKQPMAESELLQGTLEPTNEPYAIAKIAGIKLCESYNRQYG
FX      EVGARKVVSCLSTCIFPDKTTYPIDETMIHNGPPHNSNFGYSYAKRMIDVQNRAYFQQYG
UG_ER   AANVKNFIFSSSATVYGDQPKIPYVES-FPTGTPQS---PYGKSKLMVEQILTDLQK--A
             . .   *. ..      *  *. .  *  .     *. .*           .  .

GM_ER   RDYRSVMPTNLYGPHDNFHP----SNSHVIPA-LLRRFHEATAQNAPDVVVWGS------
FX      CTFTAVIPTNVFGPHDNFNI----EDGHVLPG-LIHKVHLAKS-SGSALTVWGT------
UG_ER   QPDWSIALLRYFNPVGAHPSGDMGEDPQGIPNNLMPYIAQVAVGRRDSLAIFGNDYPTED
            ..     . * .         . . .*  *.             . . *

GM_ER   GTPMREFLHVDDMAAASIHVMELAHEVWLENTQPMLSHINVGTGVDCTIRDVAQTIAKVV
FX      GNPRRQFIYSLDLAQLFIWVLREYNEV-----EPII--LSVGEEDEVSIKEAAEAVVEAM
UG_ER   GTGVRDYIHVMDLADGHVVAMEKLAN------KPGVHIYNLGAGVGNSVLDVVNAFSKAC
        *   *...   *.*   .  .    .       * .   ..*     .. .  ..

GM_ER   GYKGRVVFDASKPDGTPRKLLDVTRLHQ-LGWYHEISLEAGLASTYQWFLENQDRFRG
FX      DFHGEVTFDTTKSDGQFKKTASNSKLRTYLPDFRFTPFKQAVKETCAWFTDNYEQARK
UG_ER   GKPVNYHFAPRREGDLPAYWADASKADRELNWRVTRTLDEMAQDTWHWQSRHPQGYPD
        .      *   . ..        ..    *       .      *  *   . .
```

Fig. 1. Amino-acid sequence alignments (Higgins & Sharp, 1988) of *E. coli* GDP-4-keto-6-deoxy-D-mannose epimerase/reductase (GM_ER), human GDP-4-keto-6-deoxy-D-mannose epimerase/reductase (FX) and *E. coli* UDP-galactose 4-epimerase (UG_ER).

© 1998 International Union of Crystallography
Printed in Great Britain – all rights reserved

ISSN 0907-4449 © 1998

containing other genes involved in GDP-D-mannose and GDP-L-fucose metabolism (Stevenson *et al.*, 1996).

In the context of our investigations on the biosynthesis of GDP-L-fucose, and on its role in determining cell surface molecular properties, we cloned and expressed *yefb* and confirmed that its product indeed displays GDP-4-keto-6-deoxy-D-mannose epimerase/reductase activity. The purified enzyme (GM_ER) has been crystallized in a form suitable for high-resolution X-ray crystallographic analysis.

2. Materials and methods

GM_ER was overexpressed in *E. coli* as a GST-fusion protein. The putative coding sequence was identified in *E. coli* K12 *wca* gene cluster, responsible for cholanic acid production (Stevenson *et al.*, 1996), on the basis of sequence homology of its gene product with the human FX protein, corresponding to the *wcaG* gene (formerly known as *yefb*) (GenBank acc. No. U38473). Gene specific primers (TibMolBiol, Genova, Italy), containing *Eco*RI and *Xho*I restriction sites in sense and antisense oligonucleotides, respectively, were used to amplify a 996 kb sequence corresponding to the *wcaG* gene, using purified *E. coli* genomic DNA as template. The PCR fragment so obtained was purified by agarose gel, digested with *Eco*RI and *Xho*I, and ligated in pGEX-6P-1 vector (Pharmacia, Milan, Italy). The construct obtained was used to transform *E. coli* strain BL21 and to express the recombinant GST-fusion protein as previously described (Sturla *et al.*, 1997). The GST-GM_ER fusion protein was purified from cell lysate supernatant using GSH-Sepharose 4B, and cleaved while bound to the matrix, using Prescission Protease (Pharmacia, Milan, Italy) at 277 K for 24 h. Purity of the recombinant GM_ER was determined by sodium dodecyl sulfate polyacrylamide gel electrophoreseis (SDS–PAGE) analysis (Laemmli, 1970) and size-exclusion chromatography (Tonetti *et al.*, 1996; Sturla *et al.*, 1997).

The enzymatic activity of the recombinant protein was confirmed by high-pressure liquid chromatography (HPLC) and thin-layer chromatography (TLC) analysis, following the conversion of the substrate GDP-4-keto-6-deoxy-D-mannose to GDP-L-fucose, as previously described (Tonetti *et al.*, 1996; Sturla *et al.*, 1997). Since the substrate is unstable, it was prepared immediately before the enzymatic assay using recombinant *E. coli* GDP-D-mannose 4,6 dehydratase (Sturla *et al.*, 1997). GM_ER enzymatic assay was performed in 50 m*M* Tris–HCl, pH 8.0, containing 0.1 m*M* substrate and 0.2 m*M* NADPH, at 310 K. At different times an aliquot was withdrawn, extracted and analysed for the GDP-L-fucose content.

The purified protein was employed in a sparse-matrix search for crystal growth (Hampton Research, Laguna Hills, CA, USA), at a concentration of 16 mg ml^{-1}, at 277 K. A crystalline precipitate was promptly identified under two different growth conditions: (i) at 2.0 *M* sodium formate, 0.1 *M* sodium acetate, pH = 4.6 (as elongated thin needles); and (ii) at 1.5 *M* lithium sulfate, 0.1 *M* Hepes buffer, pH = 7.5. Refinement of the physico-chemical parameters affecting crystallization led to the isolation of large bipyramidal crystals at 1.5 *M* lithium sulfate, 0.1 *M* Mes buffer, pH = 6.5, 293 K. The crystals grow to a size of approximately 0.5 × 0.3 × 0.3 mm, in about a week. No further search for growth conditions in sodium formate [(i) above] was conducted.

Table 1. *Data collection statistics*

Values in parentheses refer to the upper resolution shell (2.6–2.5 Å).

Resolution (Å)	2.5
Total observations	62779
Unique reflections	16715
Redundancy for all data	3.8 (2.4)
R_{merge} (I) for all data	0.039 (0.154)
Completeness for all data (%)	98.6 (92)
$I/\sigma(I)$ for all reflections	13.6 (5.0)

Diffraction data at room temperature were collected on an MSC R-axis IIc system, coupled to a Rigaku RU200 rotating-anode generator, run at 50 kV, 100 mA (Cu $K\alpha$ radiation). The diffraction limit of GM_ER crystals extends to approximately 2.3 Å resolution, on this X-ray source; a full native data set was collected from one crystal in the 30.0–2.5 Å resolution range (see Table 1). The *MOSFLM* program package (Leslie, 1996) was used in conjunction with the *CCP*4 suite for data reduction and analysis (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

About 20 mg of pure GM_ER could be obtained from 1 l of *E. coli* culture. In SDS–PAGE the protein migrated as a single band, with an apparent molecular mass of 35 kDa, in agreement with the expected molecular weight derived from the translated amino-acid sequence. The protein eluted as a single peak in size-exclusion chromatography; comparison of the retention time of GM_ER with those of known markers indicated a molecular mass of approximately 70 kDa, suggesting a homodimeric assembly, as observed for human FX (Morelli *et al.*, 1977). Enzymatic assays confirmed that the product of the *wcaG* gene is indeed an NADPH-dependent GDP-4-keto-6-deoxy-D-mannose epimerase/reductase, responsible for the last step in the synthetic pathway leading to GDP-L-fucose in bacterial cells. The specific activity of GM_ER was 140 µmol h^{-1} mg^{-1}.

Analysis of the diffraction data sets collected allowed to assign the GM_ER crystals to the trigonal space group $P3_121$ (or enanthiomorph), with unit-cell constants $a = b = 105.0$, $c = 75.6$ Å, $\gamma = 120°$. Assuming one GM_ER subunit of 35 kDa molecular mass per asymmetric unit, the lattice packing parameter for this crystalline form is 3.43 $Å^3$ Da^{-1} (64% solvent content) (Matthews, 1968).

Although GM_ER and UDP-galactose 4-epimerase, whose three-dimensional structure is known (Thoden *et al.*, 1996*a,b,c*; Thoden, Hegeman *et al.*, 1997; Thoden, Gulick *et al.*, 1997), catalyze very similar reactions and display low but significant amino-acid sequence relationship (23% identity; see Fig. 1), several molecular replacement attempts employing widely different experimental strategies (Brünger, 1992; Navaza, 1994) failed to provide a clear solution for the correct orientation/location of the unknown molecule in its unit cell. The crystal structure analysis of *E. coli* GDP-4-keto-6-deoxy-D-mannose epimerase/reductase will therefore be pursued through multiple isomorphous replacement techniques. A search for suitable heavy-atom derivatives is in progress.

Part of this work was supported by ASI grant ARS-96-114 to MB and by a MURST grant to ADF.

References

Bauer, A. J., Rayment, I., Frey, P. A. & Holden, H. M. (1991). *Proteins,* **9**, 135–142.

Bauer, A. J., Rayment, I., Frey, P. A. & Holden, H. M. (1992). *Proteins*, **12**, 372–381.

Brünger, A. T. (1992). *A system for X-ray crystallography and NMR: X-PLOR. Version* 3.1, Yale University Press, New Haven, CT, USA.

Camardella, L., Carratore, V., Ciardiello, A., Damonte, G., Benatti, U. & De Flora, A. (1995). *Blood*, **85**, 264–267.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–767.

Feizi, T. & Childs, R. A. (1987). *Biochemistry*, **245**, 1–11.

Higgins, D. G. & Sharp, P. M. (1988), *Gene*, **73**, 237–244.

Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzales-Duarte, R., Jeffery, J. & Ghosh, D. (1995). *Biochemistry*, **34**, 6003–6013.

Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.

Leslie, A. (1996). In *Crystallographic Computing*, pp. 110–125. Oxford University Press.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Mergaert, P., Van Montagu, M. & Holsters, M. (1997). *FEBS Lett.* **409**, 312–316.

Morelli, A., Benatti, U., Radin, L., Wrigley, N. G. & De Flora, A. (1977). *FEBS Lett.* **80**, 1–4.

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

Ørskov, F. & Ørskov, I. (1992). *Can. J. Microbiol.* **38**, 699–704.

Stevenson, G., Adrianopoulos, K., Hobbs, M. L. & Reeves, P. R. (1996). *J. Bacteriol.* **178**, 4885–4893.

Sturla, L., Bisso, A., Zanardi, D., Benatti, U., De Flora, A. & Tonetti, M. (1997). *FEBS Lett.* **412**, 126–130.

Thoden, J. B., Frey, P. A. & Holden, H. M. (1996*a*). *Biochemistry*, **35**, 2557–2566.

Thoden, J. B., Frey, P. A. & Holden, H. M. (1996*b*). *Biochemistry*, **35**, 5137–5144.

Thoden, J. B., Frey, P. A. & Holden, H. M. (1996*c*). *Protein Sci.* **5**, 2149–2161.

Thoden, J. B., Gulick, A. & Holden, H. M. (1997). *Biochemistry*, **36**, 10685–10695.

Thoden, J. B., Hegeman, A. D., Wesenberg, G., Chapeau, M. C., Frey, P. A. & Holden, H. M. (1997). *Biochemistry*, **36**, 6294–6304.

Tonetti, M., Sturla, L., Bisso, A., Benatti, U. & De Flora, A. (1996). *J. Biol. Chem.* **271**, 27274–27279.