Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray studies on the putative dTDP sugar epimerase NovW from the novobiocin biosynthetic cluster of *Streptomyces spheroides*

Piotr Jakimowicz,[a] Caren L. Freel Meyers,[b] Christopher T. Walsh,[b] Mark J. Buttner[a] and David M. Lawson[c]*

[a]Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, England, [b]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA, and [c]Department of Biological Chemistry, John Innes Centre, Norwich NR4 7UH, England

Correspondence e-mail: david.lawson@bbsrc.ac.uk

Received 20 May 2003
Accepted 9 June 2003

Crystals of recombinant NovW (subunit MW = 22 289 Da), a putative dTDP sugar epimerase from *Streptomyces spheroides*, were grown by vapour diffusion. The protein crystallizes in space group $P4_32_12$, with unit-cell parameters $a = b = 59.20$, $c = 109.23$ Å. Native data to a resolution of 2.0 Å were collected from a single crystal at 100 K on a rotating-anode X-ray generator. Preliminary analysis of these data indicated that the asymmetric unit corresponded to a monomer, whilst dynamic light scattering (DLS) suggested that NovW was a dimer in solution. NovW is involved in the biosynthesis of the aminocoumarin antibiotic novobiocin, which targets the bacterial enzyme DNA gyrase, and represents the first enzyme to be crystallized from this biosynthetic pathway.

1. Introduction

Streptomycetes are numerous and widespread soil-dwelling bacteria that represent rich sources of useful natural products (Hodgson, 2000). In particular, these bacteria are responsible for the production of over two-thirds of the naturally derived antibiotics in clinical use today, including tetracycline, chloramphenicol, erythromycin and vancomycin (Bentley *et al.*, 2002). However, emerging resistance to many current drugs means that we need to continually identify new targets and develop novel agents against them. Indeed, vancomycin was hailed as the last antibiotic effective against staphylococcal infections, including methicillin-resistant *Staphylococcus aureus* (MRSA). Nevertheless, vancomycin-resistant strains of MRSA are now beginning to appear (Pearson, 2002). Thus, keeping ahead of bacterial antibiotic resistance will be a continual challenge for the 21st century. To this end, we have begun to study the biosynthesis of the aminocoumarin antibiotic novobiocin (Fig. 1) that is produced by *Streptomyces spheroides* and *S. niveus*. The aminocoumarins are powerful inhibitors of DNA gyrase (Gellert *et al.*, 1976), which is an essential enzyme in bacteria and a validated drug target (Maxwell, 1997; Maxwell & Lawson, 2003). However, despite its potency against DNA gyrase, novobiocin has rather unsatisfactory pharmacological properties, displaying limited solubility in water, poor penetration into Gram-negative bacteria and moderate toxicity to humans. Nevertheless, it has seen limited clinical use in the treatment of some Gram-positive bacterial infections, including MRSA (Montecalvo *et al.*, 1995; Walsh *et al.*, 1993). Using a combination of structural and mechanistic approaches, we aim to understand the biosynthetic pathway for novobiocin in sufficient detail to facilitate the rational redesign of the product through the manipulation of the individual enzymes in the pathway. Ultimately, we hope to enable the production of novel derivatives of novobiocin that retain their ability to inhibit DNA gyrase but have enhanced pharmacokinetics.

Novobiocin is essentially comprised of three moieties that are produced in parallel branches of the biosynthetic pathway: a prenylated hydroxybenzoate (ring *A*), a substituted coumarin (ring *B*) and a noviose sugar (ring *C*). These rings become ultimately linked through peptide (*A*–*B*) and glycosidic (*B*–*C*) bonds

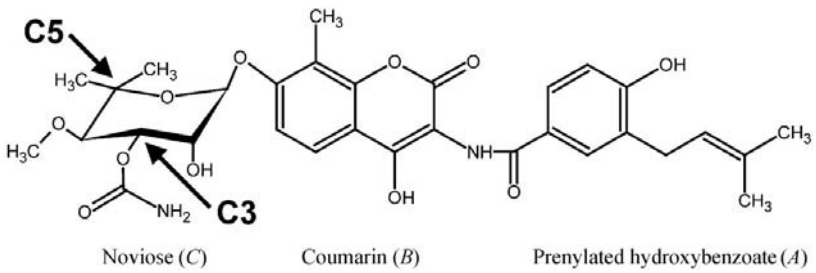

Figure 1
Structure of novobiocin. The C3 and C5 C atoms of the noviose sugar are indicated.

© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

(Fig. 1). The gene cluster encoding this biosynthetic pathway in *S. spheroides* has recently been cloned and was originally thought to contain 23 open reading frames (Steffensky *et al.*, 2000). However, comparisons with the closely related coumermycin A_1 (Wang *et al.*, 2000) and clorobiocin (Pojer *et al.*, 2002) gene clusters suggests that no more than 19 genes may be involved in the biosynthesis of novobiocin. The functions of many of the *nov* gene products have already been assigned by sequence homology to known proteins (Steffensky *et al.*, 2000). We report here the crystallization and preliminary X-ray analysis of NovW, a putative dTDP sugar epimerase, which is implicated in the biosynthesis of the noviose sugar. This represents the first enzyme from the *nov* gene cluster to be crystallized.

2. Materials and methods

2.1. Protein expression and purification

The *S. spheroides novW* gene was amplified by PCR using *S. spheroides* genomic DNA as a template. The amplified DNA was cloned into the vector pET-15b (Novagen) to give a plasmid encoding a polypeptide with a thrombin-cleavable N-terminal hexahistidine tag. This added a further 20 residues to the native protein (with the sequence MGSSHHHHHHSSGLVPRGSH), giving a total deduced molecular mass of 24 452 Da. This plasmid was transformed into *Escherichia coli* strain BL21 (DE3) Gold (Studier & Moffatt, 1986) and the cells were grown in 6 l of Luria–Bertani medium at 288 K for 72 h, without induction by isopropyl-β-D-thiogalactopyranoside. Harvested cells were resuspended in buffer [25 m*M* Tris–HCl pH 8.0, 400 m*M* NaCl, 10%(*v*/*v*) glycerol] and lysed by passage through a French press at 110 MPa. The cell lysate obtained by centrifugation at 13 000*g* for 20 min was loaded onto an Ni^{2+}-charged Hi-trap metal-chelation column and unbound proteins were removed by washing with 2 m*M* imidazole in 25 m*M* Tris–HCl pH 8.0, 400 m*M* NaCl, 10%(*v*/*v*) glycerol. His-tagged NovW bound to the column was eluted with a 5–500 m*M* imidazole gradient in 25 m*M* Tris–HCl pH 8.0, 400 m*M* NaCl, 10%(*v*/*v*) glycerol. The fractions containing NovW were pooled and dialysed against buffer containing 100 m*M* NaCl, 10%(*v*/*v*) glycerol and 1 m*M* EDTA in 50 m*M* Tris–HCl pH 8.0. A second dialysis was carried out in buffer containing 100 m*M* NaCl, 10%(*v*/*v*) glycerol and 1 m*M* dithiothreitol (DTT) in 50 m*M* Tris–HCl pH 8.0. The His tag was not cleaved. The purified protein was stored under liquid nitrogen in 30 µl aliquots and subsequently shipped on dry ice for crystallization trials.

Prior to crystallization, dynamic light scattering (DLS) was used to monitor the solution properties of the protein in different buffers. Small samples (30 µl) of purified NovW were buffer-exchanged and concentrated to around 2.5 mg ml^{-1} using Ultrafree 10 kDa cutoff concentrators (Millipore). The samples were then centrifuged through 0.1 µm Ultrafree filters (Millipore) to remove particulate material before introduction into a 12 µl microsampling cell. The latter was then inserted into a Dynapro-MSTC molecular-sizing instrument at 293 K (Protein Solutions Inc.). For each sample, at least 15 scattering measurements were taken and the data were analysed using the *DYNAMICS* software package (Protein Solutions Inc.). Once buffer conditions that gave rise to favourable solution properties had been obtained, a large sample (400 µl) was buffer-exchanged into this solution and concentrated to approximately 10 mg ml^{-1} for crystallization trials using a Centriprep-30 concentrator (Amicon).

2.2. Crystallization and X-ray diffraction analysis

Crystallization trials were performed by vapour diffusion in hanging drops using 24-well VDX plates (Hampton Research) at a constant temperature of 291 K. Drops consisted of 1 µl protein solution mixed with 1 µl well solution; the well volume was 1.0 ml. Initial crystallization conditions were established using an in-house screen (denoted MYXS) consisting of three different concentrations each of ammonium sulfate and PEG 8000 at four different pH values, giving 24 conditions (see http://www.jic.bbsrc.ac.uk/staff/david-lawson/research/myxs.htm for details). The successful conditions were subsequently optimized to include ethylene glycol. The crystals were then mounted directly from the mother liquor onto cryoloops (Hampton Research) and flash-cooled to 100 K in a stream of gaseous nitrogen produced by an X-Stream cryocooler (Rigaku-MSC). Diffraction data were collected in-house using a MAR 345 image-plate detector (X-ray Research) mounted on a Rigaku RU-H3RHB rotating-anode X-ray generator (operated at 50 kV and 100 mA) fitted with Osmic confocal optics and a copper target (Cu $K\alpha$; λ = 1.542 Å). X-ray data were processed using the *HKL* software package (Otwinowski & Minor, 1997). All other downstream data processing and statistical analysis was carried out using programs from the *CCP*4 software suite (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

NovW was overexpressed and purified with an approximate yield of 10 mg of protein from 1 l culture and was judged to be greater than 95% pure from SDS–PAGE analysis. After purification, the sample was at a concentration of approximately 2.4 mg ml^{-1} in 50 m*M* Tris–HCl buffer pH 8.0 containing 100 m*M* NaCl, 10%(*v*/*v*) glycerol and 1 m*M* DTT. DLS analysis before manipulation of the buffer components gave a monomodal distribution and a relatively low polydispersity (26.5%), although two aggregate peaks were present. From these results, the molecular weight was estimated to be 91 kDa after correcting for the glycerol present. This value lies between the figures expected for a trimer (73.4 kDa) and a tetramer (97.8 kDa) of His-tagged NovW, although this value was most likely skewed upwards by the presence of some aggregated material. Indeed, this sample could not be concentrated significantly by centrifugation, owing to precipitated protein clogging the membrane of the concentrator. An attempt to exchange the protein into a minimal buffer consisting of 1 m*M* DTT in 20 m*M* Tris–HCl buffer pH 7.5 also induced protein precipitation. However, the situation could be improved through the reintroduction of NaCl to a concentration of 0.5 *M*. Other factors that could promote precipitation include metal ion-mediated intermolecular interactions through the N-terminal His tags and intermolecular disulfide-bond formation (NovW contains five Cys residues). To counter these possible effects, fresh samples of purified NovW were independently exchanged into 0.5 *M* NaCl in 20 m*M* Tris–HCl buffer at pH 7.5 supplemented with relatively high concentrations (10 m*M*) of either EDTA (to chelate metal ions) or DTT (to keep thiols reduced). Although no noticeable enhancement of the solution properties of the sample was apparent with either supplement, neither did they appear to have any deleterious effects. Thus, they were both included in the final sample buffer, which comprised 0.5 *M* NaCl, 10 m*M* EDTA and 10 m*M* DTT in 20 m*M* Tris–HCl pH 7.5. In this buffer, NovW could readily be concentrated to 10 mg ml^{-1} for crystallization trials. DLS analysis at this stage gave a monomodal distribution and a polydispersity of 25.5%, with a very minor

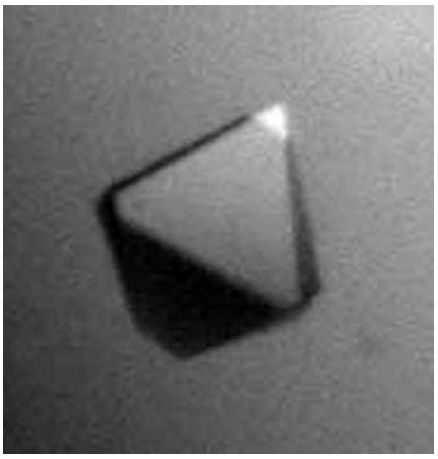

Figure 2
Single crystal of *S. spheroides* NovW, approximately 250 × 150 × 150 µm in size.

aggregate peak. From these results, the molecular weight was estimated to be 58 kDa, being relatively close to the value expected for a dimer (48.9 kDa).

Preliminary crystals of NovW were obtained overnight in the 11th condition of the MYXS screen, which consists of 25% saturated $(NH_4)_2SO_4$ in 0.1 *M* Tris–HCl pH 8.5. They were octahedral in morphology, with approximate dimensions of 150 × 100 × 100 µm. Larger crystals were subsequently obtained at the same precipitant concentration, but with 25% of the buffer volume replaced by ethylene glycol. These took up to 7 d to reach maximum dimensions of 250 × 150 × 150 µm (Fig. 2) and could be flash-cooled to 100 K without further cryoprotection.

X-ray data were collected from a single crystal to a maximum resolution of 2.0 Å. The symmetry was established as primitive tetragonal, with unit-cell parameters $a = b = 59.20$, $c = 109.23$ Å. A total of 144 × 1° oscillation images were recorded in a continuous sweep, yielding a data set that was 99.8% complete. Data-collection statistics are summarized in Table 1. After inspection of pseudo-precession plots, the space group was assigned as either $P4_12_12$ or $P4_32_12$. Analysis of the contents of the asymmetric unit based on a single protomer (24 452 Da) gave a crystal-packing parameter (V_M) of 2.0 Å^3 Da^{-1}, with a corresponding solvent content of 36.7% (Matthews, 1968). Given the relatively high sequence identity (44%) with RmlC from *Methanobacterium thermoautotrophicum*, a protein of known structure (Christendat *et al.*, 2000), molecular replacement could be used to solve the structure. The RmlC subunit from PDB entry 1ep0, converted to polyalanine and stripped of solvent molecules, was used as the search model in the program *AMoRe*. Both of the possible space groups were tested, which unambiguously identified $P4_32_12$ as the true space group, giving a correlation coefficient of 55.7% and an overall *R* factor[1] of 54.9% at 2.0 Å resolution after rigid-body refinement. As expected, a single subunit was located in the asymmetric unit with plausible crystal-packing interactions. In particular, an extensive interface was apparent across a twofold crystallographic axis that enables two symmetry-equivalent antiparallel β-sheets to extend from one subunit into the other. This is consistent with the quaternary interactions observed in the RmlC dimer and therefore validates this as the correct molecular-replacement solution. Moreover, this agrees with the DLS result and thus strongly suggests that the biologically relevant oligomer of NovW is a dimer. This preliminary model is currently being refined.

Table 1
Summary of X-ray data for NovW.

Values in parentheses are for the outer resolution shell.

Resolution range (Å)	30.0–2.0 (2.07–2.00)
Unique reflections	13753
Completeness (%)	99.8 (100.0)
Redundancy	10.7
R_{merge}†	0.074 (0.311)
$\langle I\rangle/\langle\sigma(I)\rangle$	26.6 (6.6)

† $R_{merge} = \sum(|I_j - \langle I_j\rangle|)/\sum\langle I_j\rangle$, where I_j is the intensity of an observation of reflection j and $\langle I_j\rangle$ is the average intensity for reflection j.

NovW is closely related in both sequence and structure to RmlC, a dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13) from the rhamnose biosynthetic pathway. Thus, reference to previous mechanistic and structural studies of RmlC (Christendat *et al.*, 2000; Dong *et al.*, 2003; Giraud *et al.*, 2000; Graninger *et al.*, 1999; Stern *et al.*, 1999) will yield much useful information concerning the location of the active centre and the identity of key catalytic residues and specificity determinants in NovW. Nevertheless, important questions remain to be answered concerning the details of the enzyme mechanism, in particular the order in which NovW and NovU (a putative C-methylase) act in TDP-L-noviose biosynthesis. There are two possible scenarios: (i) epimerization at both the C3 and C5 positions of the deoxy sugar prior to C-methylation at the 5-position by NovU or (ii) epimerization at C3 only following C-methylation at the 5-position by NovU (Fig. 1). We aim to address these unresolved issues through the structural analysis of ligand-bound complexes of NovW, coupled with complementary mechanistic and mutagenic studies.

[1] R factor = $\sum(|F_{obs} - F_{calc}|)/\sum|F_{obs}| \times 100$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively.

PJ, MJB and DML would like to thank the BBSRC for financial support. CLFM and CTW are funded by NIH grants F32 AI054007 and GM 49338, respectively. R. Chen is acknowledged for isolation of *S. spheroides* genomic DNA. We are also grateful to C. Stevenson and A. Maxwell for critically reading this manuscript.

References

Bentley, S. D. *et al.* (2002). *Nature (London)*, **417**, 141–147.
Christendat, D., Saridakis, V., Dharamsi, A., Bochkarev, A., Pai, E. F., Arrowsmith, C. H. & Edwards, A. M. (2000). *J. Biol. Chem.* **275**, 24608–24612.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dong, C., Major, L. L., Allen, A., Blankenfeldt, W., Maskell, D. & Naismith, J. H. (2003). *Structure*, **11**, 715–723.
Gellert, M., O'Dea, M. H., Itoh, T. & Tomizawa, J. (1976). *Proc. Natl Acad. Sci. USA*, **73**, 4474–4478.
Giraud, M. F., Leonard, G. A., Field, R. A., Berlind, C. & Naismith, J. H. (2000). *Nature Struct. Biol.* **7**, 398–402.
Graninger, M., Nidetzky, B., Heinrichs, D. E., Whitfield, C. & Messner, P. (1999). *J. Biol. Chem.* **274**, 25069–25077.
Hodgson, D. A. (2000). *Adv. Microb. Physiol.* **42**, 47–238.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Maxwell, A. (1997). *Trends Microbiol.* **5**, 102–109.
Maxwell, A. & Lawson, D. M. (2003). *Curr. Top. Med. Chem.* **3**, 283–303.
Montecalvo, M. A., Horowitz, H., Wormser, G. P., Seiter, K. & Carbonaro, C. A. (1995). *Antimicrob. Agents Chemother.* **39**, 794.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pearson, H. (2002). *Nature (London)*, **418**, 469.
Pojer, F., Li, S. M. & Heide, L. (2002). *Microbiology*, **148**, 3901–3911.
Steffensky, M., Muhlenweg, A., Wang, Z. X., Li, S. M. & Heide, L. (2000). *Antimicrob. Agents Chemother.* **44**, 1214–1222.
Stern, R. J., Lee, T. Y., Lee, T. J., Yan, W., Scherman, M. S., Vissa, V. D., Kim, S. K., Wanner, B. L. & McNeil, M. R. (1999). *Microbiology*, **145**, 663–671.
Studier, F. W. & Moffatt, B. A. (1986). *J. Mol. Biol.* **189**, 113–130.
Walsh, T. J., Standiford, H. C., Reboli, A. C., John, J. F., Mulligan, M. E., Ribner, B. S., Montgomerie, J. Z., Goetz, M. B., Mayhall, C. G., Rimland, D., Stevens, D. A., Hansen, S. L., Gerard, G. C. & Ragual, R. J. (1993). *Antimicrob. Agents Chemother.* **37**, 1334–1342.
Wang, Z. X., Li, S. M. & Heide, L. (2000). *Antimicrob. Agents Chemother.* **44**, 3040–3048.