Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

X-ray crystallographic characterization and phasing of an NtrC homologue

László Sallai, Jörg Hendle† and Paul A. Tucker*

European Molecular Biology Laboratory
Hamburg Outstation, c/o DESY,
Notkestrasse 85, D-22603 Hamburg, Germany

† Present address: Structural GenomiX Inc.,
10505 Roselle Street, San Diego,
California 92121, USA.

Correspondence e-mail:
tucker@embl-hamburg.de

Received 3 June 2003
Accepted 7 July 2003

The ZraR (HydG) protein is a 441-amino-acid protein with three functional domains and is homologous to the general nitrogen-regulatory protein NtrC that regulates nitrogen assimilation in many bacteria. The AAA and DNA-binding domains (residues 141–441) of the ZraR protein from *Salmonella typhimurium* were crystallized using the sitting-drop vapour-diffusion method. X-ray diffraction data from the native crystal have been collected to 3.0 Å resolution. Initial phasing was successfully performed by the SIRAS method using derivativatized crystals soaked in 1 m*M* ethylmercuric phosphate. Preliminary structural analysis shows the presence of a hexamer in the asymmetric unit. Model building is in progress.

1. Introduction

Two-component regulatory systems can be characterized by phosphotransfer reactions between a sensor protein and a response regulator (Stock *et al.*, 2000). The σ^{54}-dependent (the σ factor being a component of the RNA polymerase complex) general nitrogen-regulatory proteins B and C (NtrB/NtrC) that control the nitrogen assimilation in many bacteria belong to this family and have been studied extensively. The NtrC (NR_I) protein comprises three domains: an N-terminal regulatory domain, an AAA (ATPase activity with various cellular functions) domain and a C-terminal DNA-binding domain. The N-terminal domain is phosphorylated by the sensor kinase NtrB (when glutamine levels fall) and transmits the signal to the central domain, which oligomerizes. Upon ATP hydrolysis, this oligomer is able to generate the activated open promoter complex (RNA polymerase–σ^{54}) and thus initiate transcription at the *glnA* enhancer, which regulates the expression of glutamine synthetase (Austin & Dixon, 1992; Wyman *et al.*, 1997). The structure of the N-terminal receiver domain has already been solved in both the inactive and active states (Volkman *et al.*, 1995; Kern *et al.*, 1999), as has that of the DNA-binding domain (Pelton *et al.*, 1999).

We have purified and tried to crystallize a number of full-length homologues and orthologues of NtrC. These included NtrCs from *Klebsiella pneumoniae*, *Azotobacter vinlandii* and *Escherichia coli*, the hydrogenase G (HydG or ZraR) protein from *E. coli* and *Salmonella typhimurium* and YfhA (a protein with unknown function) from *E. coli*. Only ZraR from *E.coli* gave crystals, which diffracted to 8 Å on a synchrotron beamline, but we were unable to improve the diffraction limit. For this reason, and in view of the shorter inter-domain linkers in ZraR relative to the other proteins, we concentrated our efforts on shorter constructs of ZraR orthologues.

ZraR from *S. typhimurium*, a homologue of NtrC, has been postulated to play a role in the regulation of hydrogenase expression (Stoker *et al.*, 1989; Chopra *et al.*, 1991); however, neither the physiological signal nor the molecular mechanism of its action have been established. Recent studies have shown that the ZraS/ZraR system is responsive to high Zn^{2+} concentrations and that it regulates the expression of the *zraP* gene, which codes for a periplasmic Zn^{2+}-binding protein (Leonhartsberger *et al.*, 2001). Here, we describe the overexpression, purification, crystallization and phasing of a construct comprised of the central and C-terminal domains of ZraR. The structure may help us understand the activation mechanism of the σ^{54}-driven transcription.

2. Materials and methods

2.1. Protein expression and purification

A PCR product containing the coding sequence of the central and C-terminal domains (residues 141–441) of the ZraR protein was cloned into the pET-28 vector. *E. coli* B834 (DE3) cells transformed with the plasmid were grown overnight at 310 K on LB agar plates containing 50 mg l^{-1} kanamycin. A single colony was picked to inoculate 2 l of LB medium containing antibiotic. Protein induction was carried out at OD $\simeq$ 0.6 by the addition of IPTG to a final concentration of 0.5 m*M*. After induction, growth continued for 3 h at 310 K. The SeMet-containing protein

was overexpressed in M9 media supplemented with SeMet overnight at 310 K.

The cell pellet was resuspended in degassed buffer *A* (500 m*M* NaCl, 50 m*M* NaH_2PO_4 pH 8.0, 5% glycerol) and cells were lysed by the addition of lysozyme followed by sonication on ice for 3 × 1 min. The extract was centrifuged at 20 000*g* for 20 min and the supernatant was loaded onto an Ni-affinity column and eluted with a linear gradient to degassed buffer *B* (buffer *A* plus 1 *M* imidazole). The sample was dialyzed against 500 m*M* NaCl, 10 m*M* Tris–HCl pH 8.0, 2 m*M* DTT and 2 m*M* $CaCl_2$ and digested with 10 U mg^{-1} bovine thrombin overnight at room temperature to remove the His tag. After treatment with thrombin, the sample was dialyzed against buffer *C* (50 m*M* NaCl, 20 m*M* CHES pH 9.5, 2 m*M* DTT, 0.1 m*M* EDTA) and applied onto a heparin column. The protein was eluted with a gradient to buffer *D* (buffer *C* plus 1.95 *M* NaCl). Finally, the buffer of the protein sample was exchanged to 50 m*M* NaCl, 10 m*M* CHES pH 9.5, 2 m*M* DTT and 0.1 m*M* EDTA, which was supplemented with 20% glycerol; the protein was divided into aliquots and frozen in liquid nitrogen when crystallization experiments were not to be performed immediately.

2.2. Crystallization and data collection

The protein was concentrated to 10–12 mg ml^{-1} and crystallized using sitting-drop vapour diffusion by mixing 1 µl of protein solution with 0.6 µl of reservoir solution. Initially, crystals were obtained at 293 K from 15% ethanol, 0.1 *M* HEPES pH 7.2, 200 m*M* $MgCl_2$, a condition from the Wizard II sparse-matrix screen (the sparse-matrix screens tried were Wizard I, Wizard II, Cryo I and Cryo II from Emerald Bio-Structures, and Crystal Screen and Crystal Screen 2 from Hampton Research), but conditions could not be optimized until it was realised that 2-propanol, which was probably an additive in the ethanol used, was essential for crystallization. Subsequently, crystals were obtained at 291 K in 100 m*M* HEPES pH 7.5, 10% 2-propanol and 200 m*M* NaCl. The crystallization conditions were further optimized, with the best crystals obtained using a well solution containing 100 m*M* HEPES pH 8.2, 9% 2-propanol, 3% methanol and 200 m*M* NaCl. The selenomethionine (SeMet) containing protein was crystallized in an N_2 atmosphere using a glove bag. For heavy-metal soaking experiments, the native crystals were transferred to 100 m*M* HEPES pH 8.2, 9% 2-propanol, 3% methanol and 200 m*M* NaCl containing 1 m*M* ethylmercuric phosphate for 5 min. The crystals were back-soaked in the cryobuffer (25% 2-propanol, 20% PEG 200, 2% glycerol and 100 m*M* HEPES pH 8.2) for 5 min. The crystals were then either flash-cooled in a nitrogen stream at 100 K or plunged into and stored in liquid nitrogen.

Data from the native and SeMet crystals at the peak and inflection-point wavelengths were collected at the ID-14-4 beamline at the ESRF in Grenoble using an ADSC Quantum-4 detector. The wavelengths were chosen on the basis of the X-ray absorption fluorescence spectrum. Data were collected from the mercury derivative at the X13 beamline at the EMBL in Hamburg using a MAR165 CCD detector. A summary of the data-collection statistics is given in Table 1. The data were processed and reduced using the programs *DENZO*, *SCALEPACK* (Otwinowski & Minor, 1997) and *TRUNCATE* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994).

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	Native	SeMet crystal		Hg derivative
		Peak	Edge	
Crystal data				
Space group	$P222_1$			
Unit-cell parameters				
a (Å)	107.44	109.43	109.72	106.88
b (Å)	114.74	112.00	111.97	113.94
c (Å)	187.26	187.84	187.84	186.46
Mosaicity (°)	0.73	0.57	0.68	0.49
Data collection				
Wavelength (Å)	0.9806	0.9795	0.9806	0.8020
Resolution range (Å)	30.0–3.0 (3.11–3.00)	30.0–3.2 (3.25–3.20)	30.0–3.4 (3.45–3.40)	20.0–4.0 (4.10–4.00)
Observed reflections	173549	142605	118851	86339
Unique reflections	46186	38771	32788	17045
R_{merge}† (%)	0.06 (0.44)	0.08 (0.35)	0.07 (0.35)	0.08 (0.26)
Completeness (%)	97.5 (87.8)	98.4 (90.1)	98.1 (88.5)	81.8 (58.5)
$I/\sigma(I)$	10.9 (2.6)	9.2 (3.4)	11.6 (3.3)	9.0 (4.5)

† $R_{merge} = \sum|I_i - \langle I_i\rangle|/\sum\langle I_i\rangle$, where I_i is the observed intensity and $\langle I_i\rangle$ is the average intensity over symmetry-equivalent intensities.

2.3. Phasing

SHELXD (Schneider & Sheldrick, 2002), *SnB* (Weeks & Miller, 1999), *CRUNCH* (de Graaf *et al.*, 2001) and *SOLVE* (Terwilliger, 2002) were used to search for the Hg and Se sites using SAS, SIRAS or MAD strategies. The self-rotation function was calculated with *POLARRFN* (Collaborative Computational Project, Number 4, 1994). The Hg positions were refined using *SOLVE*; the initial map was solvent-flattened and NCS averaged with *RESOLVE*. *SHARP* (La Fortelle & Bricogne, 1997) was used to find additional Se sites from SAS data at the peak wavelength with the initial phases from the mercury derivative.

3. Results and discussion

Crystals grew overnight to a maximum length of ∼300 µm; the average approximate dimensions were 150 × 80 × 50 µm (Fig. 1). The crystals belong to space group $P222_1$. The volume of the unit cell is 2.31×10^6 $Å^3$, which is consistent with a hexamer in the asymmetric unit (V_M = 2.91 $Å^3$ Da^{-1}) as confirmed by the self-rotation function. The best native crystal diffracted to 3.0 Å. The SeMet crystal initially diffracted (at the peak wavelength) to 3.2 Å but, owing to radiation damage, the inflection-point data were only processed to 3.4 Å. Probably as a consequence of the radiation damage, SAS and two-wavelength MAD strategies using different programs failed to give solutions for the selenium sites. The SeMet crystal was non-isomorphous with the native crystal (Table 1) and consequently the isomorphous

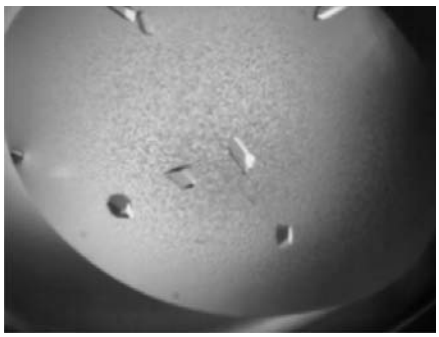

Figure 1
Native crystals of *S. typhimurium* ZraR central and DNA-binding domains. The typical dimensions are approximately 150 × 80 × 50 µm.

differences between the native and SeMet data sets also failed to yield the selenium positions. Soaking the native crystals in 1 m*M* ethylmercuric phosphate for 5 min gave Hg-derivative crystals that diffracted to 4.0 Å. All the six initial heavy-atom positions were found from the SAS data with *SHELXD*; the sites were refined with *SOLVE*. Data from the native crystal were used to calculate SIRAS phases with an overall figure of merit of 0.17. With help of the initial phases from the Hg SIRAS phasing, all the 36 Se sites were located using an SAS strategy with *SOLVE* and *SHARP*. The maps produced from *SOLVE* and *SHARP* were noisy and did not reveal any recognizable protein features. Surprisingly, phase combination from the two different sources did not improve the map quality, so only the SIRAS phases from the Hg sites were used to calculate an electron-density map. Solvent-flattening and sixfold NCS operators determined from the Se sites were applied to extend the phases from 4.0 to 3.0 Å, which yielded a map with recognisable features and with an overall figure of merit of 0.45. Model building is in progress.

We wish to thank Steffi Arzt at the ID-14-4 beamline at the ESRF in Grenoble for helping us to collect the native and MAD data sets. We would like to thank Martin Drummond and Ray Dixon (Department of Molecular Microbiology, John Innes Centre, Norwich, England) for providing us with proteins and especially for the plasmid pKH-420 of the *zraR* gene. Part of this work was supported by the European Union (EU-Biotechnology BIO4-CT-1997-2143).

References

Austin, S. & Dixon, R. (1992). *EMBO J.* **11**, 2219–2228.

Chopra, A. K., Peterson, L. W. & Prasad, R. (1991). *Biochim. Biophys. Acta*, **1129**, 115–118.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Graaf, R. A. G. de, Hilge, M., van der Plas, J. L. & Abrahams, J. P. (2001). *Acta Cryst.* D**57**, 1857–1862.

Kern, D., Volkman, B. F., Luginbuhl, P., Nohaile, M. J., Kustu, S. & Wemmer, D. E. (1999). *Nature (London)*, **402**, 894–898.

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Leonhartsberger, S., Huber, A., Lottspeich, F. & Böck, A. (2001). *J. Mol. Biol.* **307**, 93–105.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pelton, J. G., Kustu, S. & Wemmer, D. E. (1999). *J. Mol. Biol.* **292**, 1095–1110.

Schneider, T. R. & Sheldrick, G. M. (2002) *Acta Cryst.* D**58**, 1772–1779.

Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). *Annu. Rev. Biochem.* **69**, 183–215.

Stoker, K., Reijnders, W. N. M., Oltmann, L. F. & Stouthamer, A. H. (1989). *J. Bacteriol.* **171**, 4448–4456.

Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1937–1940.

Volkman, B. F., Nohaile, M. J., Amy, N. K., Kustu, S. & Wemmer, D. E. (1995). *Biochemistry*, **34**, 1413–1424.

Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.

Wyman, C., Rombel, I., North, A. K., Bustamante, C. & Kustu, S. (1997). *Science*, **275**, 1658–1661.