Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction studies of the transcriptional regulator TraR bound to its cofactor and to a specific DNA sequence

Alessandro Vannini,[a] Cinzia Volpari,[a] Cesare Gargioli,[a] Ester Muraglia,[b] Raffaele De Francesco,[b] Petra Neddermann[a] and Stefania Di Marco[a]*

[a]Department of Biochemistry, IRBM P. Angeletti, 00040 Pomezia, Rome, Italy, and [b]Department of Chemistry, IRBM P. Angeletti, 00040 Pomezia, Rome, Italy

Correspondence e-mail: stefania_dimarco@merck.com

Received 18 April 2002
Accepted 5 June 2002

TraR is an *Agrobacterium tumefaciens* transcriptional regulator which binds the pheromone *N*-3-oxooctanoyl-L-homoserine lactone (AAI) in response to the bacterial population density. The TraR–AAI complex dimerizes and interacts with a specific 18-base-pair DNA sequence (TraBox), activating promoters containing this site. TraR was overexpressed and purified from *Escherichia coli*. Crystals of the ternary complex, in which dimeric TraR–AAI is bound to the TraBox sequence, have been obtained by the vapour-diffusion method. The crystals belong to space group $P2_12_12_1$, with unit-cell parameters a = 66.99, b = 94.67, c = 209.66 Å, with two $(\text{TraR–AAI})_2$–TraBox complexes in the asymmetric unit. A three-wavelength MAD data set for the seleno-L-methionine-substituted form has been collected to a resolution of 3 Å. 20 of the 24 crystallographically independent selenium sites were located as part of the MAD-phasing process.

1. Introduction

Quorum sensing, or the control of gene expression in response to cell density, is used by both Gram-negative and Gram-positive bacteria to regulate a variety of physiological functions. In all cases, quorum sensing involves the production and detection of extracellular signalling molecules called autoinducers. In this strategy, in response to the bacterial population density, a diffusible molecule (acting as a 'signal') can bind to and activate a transcriptional regulator, which then binds to a recognition site close to the promoters of the target genes, regulating their expression (reviewed in Fuqua & Winans, 1994).

TraR is an *Agrobacterium tumefaciens* transcriptional regulator involved in the conjugal transfer of the Ti plasmid to the host cells and belongs to the family of LuxR-type proteins. Genes required for conjugation (*tra* genes) are positively regulated by the TraR protein, which is activated by the pheromone *N*-3-oxooctanoyl-L-homoserine lactone, sometimes referred to as *Agrobacterium* autoinducer (AAI; Piper *et al.*, 1993; Piper & Farrand, 2000).

TraR is a 26.7 kDa protein (234 amino acids) which is only expressed in a soluble form in the presence of AAI and is present as a dimer in solution. The AAI is tightly bound to TraR with a stoichiometry of one molecule of pheromone to one molecule of protomer, as assayed by amino-acid analysis (AAI is quantified by conversion to homoserine; Zhu & Winans, 1999) and by mass spectrometry (Orsatti *et al.*, 2002). Strong evidence suggests that AAI is critical for the folding of nascent TraR protein into its mature tertiary structure (Zhu & Winans, 2001). In contrast, another study suggests the possible role of AAI as a signal which drives the dimerization of TraR and its release from membranes into the cytoplasm (Qin *et al.*, 2000). TraR is thought to contain an amino-terminal module that binds AAI, which also mediates multimerization, and a carboxyl-terminal module that binds particular recognition DNA sites of 18 base pairs, referred to as TraBox, near target promoters. This module is also supposed to interact with bacterial RNA polymerase (Fuqua *et al.*, 1996; Zhu & Winans, 1999).

We have purified TraR in the presence of its cofactor AAI from *Escherichia coli*. We have also crystallized the ternary complex $(\text{TraR–AAI})_2$–TraBox. As at present no structure of a LuxR-type protein has been solved and no relevant sequence homology has been detected with other proteins, we applied the multi-wavelength anomalous diffraction (MAD) method (Hendrickson, 1991) using Se atoms as anomalous scatterers to obtain initial phases.

2. Experimental and results

2.1. Protein production

AAI (Fig. 1; molecular mass 241.29 Da; $C_{12}H_{19}NO_4$) was synthesized from L-homoserine lactone hydrobromide and the acyl Meldrum's acid intermediate of Oikawa according to Zhang *et al.* (1993).

TraR (SWISS-PROT P33905; plasmid pJZ358 kindly provided by S. C. Winans,

Cornell University, USA) was overexpressed in *E. coli* BL21-SI cells (Gibco, Life Technologies) using a T7 expression system. A total of 6 l of cells were grown at 310 K in LB medium deprived of NaCl until the $OD_{600\ nm}$ was 0.9. The medium was then supplemented with 2 µ*M* AAI and expression was induced for 20 h by addition of 0.3 *M* NaCl. Cells were resuspended in 300 ml of lysis buffer [50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 1 m*M* ethylenediaminetetraacetic acid (EDTA), 2 m*M* dithiothreitol (DTT), 10% glycerol and 1 µ*M* AAI] and the soluble fraction was loaded onto an ion-exchange column (Heparin) pre-equilibrated with lysis buffer. The protein was eluted with a 400 ml linear gradient of NaCl from 0.3 to 1 *M* and loaded onto a second ion-exchange column (Resource S) pre-equilibrated with the same buffer but containing 0.1 *M* NaCl. The TraR–AAI complex was eluted with a 300 ml linear gradient of NaCl from 0.1 to 0.45 *M* and loaded onto a gel-filtration column (Superdex G-200) pre-equilibrated with the same buffer but containing 0.25 *M* NaCl. The purified protein was concentrated to 1.1 m*M* by ultrafiltration. TraR concentrations were determined by quantitative amino-acid analysis and by UV spectroscopy (ε_{280} = 37 440 M^{-1} cm^{-1}). The TraR–AAI complex (50 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 1 m*M* EDTA, 2 m*M* DTT, 10% glycerol, 1 µ*M* AAI) was then stored at 193 K.

The pure protein, at 0.2 m*M*, elutes as a dimer from the gel-filtration column. This result was confirmed by dynamic light scattering (DLS) performed at 297 K in globular mode settings using a DynaPro-801 molecular-sizing instrument with temperature control. The TraR–AAI complex is monodisperse with an hydrodynamic radius of 3.47 nm, corresponding to a molecular mass of 59.1 kDa (a dimer). This complex binds to a TraBox-containing oligonucleotide with an estimated K_d of 4 n*M*, as determined by gel-retardation assay.

Selenomethionine-labelled TraR was overexpressed in the methionine-auxotrophic *E. coli* strain B834-DE3 (Novagen) in M9 medium with 40 mg ml^{-1} of all amino acids except methionine and 60 mg ml^{-1} seleno-L-methionine and was grown at 310 K until $OD_{600\ nm}$ = 0.7–0.8. The medium was then supplemented with 2 µ*M* AAI and protein expression was induced for 36 h at 296 K by the addition of 0.4 m*M* isopropyl β-D-thiogalactoside. The labelled protein was purified using the same procedure as the native protein. Selenomethionine incorporation was confirmed by ESI mass spectrometry (LCQ quadrupole ion trap mass spectrometer; Finnigan MAT). Native TraR had MW = 26 731 (1) Da, compared with a molecular weight of 27 013 (1) Da for the selenomethionine derivative, indicating 100% incorporation of six Se atoms. There are no evidence of post-translational modifications in either the native or SeMet-derivatized protein, as assayed by mass spectrometry.

2.2. DNA syntheses

The HPLC-purified oligonucleotide 5′-ATG-TGC-AGA-TCT-GCA-CAT-3′ (referred to as TraBox, Oswell Oligonucleotides) was resuspended to 2 m*M* and, using its perfect dyad symmetry, annealed to give a 18 bp double-stranded DNA. A brominated oligonucleotide 5′-ATG-TGC-ABA-TCT-GCA-CAT-3′ (where B is 8-bromo-dG and referred to as TraBoxBr) was also used in order to include a Br atom as an alternative scatterer to the Se atom.

2.3. Crystallization

Solutions, each containing one of the following complexes, $(TraR–AAI)_2$–TraBox, $[TraR(SeMet)–AAI]_2$–TraBox, $[TraR(SeMet)–AAI]_2$–TraBoxBr and $(TraR–AAI)_2$–TraBoxBr, hereafter referred to as native, SeMet, SeMet/2Br and 2Br, respectively, were prepared by mixing equimolar amounts of protein dimer and double-stranded DNA and immediately used for crystallization. Trials were conducted at 297 K in 24-well plates using the hanging-drop vapour-diffusion method (McPherson, 1999). In all experiments, 2–3 µl of $(TraR–AAI)_2$–DNA complex at concentrations ranging from 0.1 to 0.22 m*M* were mixed with an equal volume of reservoir solution and equilibrated against 0.6 ml of reservoir solution. Initial screenings were carried out using the crystallization reagent kits Natrix, Crystal Screen I, Crystal Screen II and Crystal Screen Lite (Hampton Research). Needle-like crystals were obtained in several conditions for all four different complexes. After extensive optimization, the best crystallization conditions were 50 m*M* MES buffer pH 6.2, 200 m*M* calcium acetate, 3–5%(*w*/*v*) PEG 8000, 3 m*M* DTT using a protein concentration of 0.21 m*M*. Large rod-shaped crystals of all four complexes (native, SeMet, SeMet/2Br and 2Br) were obtained under the same conditions mentioned above after 4–6 d at 297 K and had dimensions of up to 1.4 × 0.25 × 0.12 mm (Fig. 2*a*). Prior to data collection, crystals were transferred into a stabilizing solution with the same composition of the mother liquor, but with a slight increase of PEG 8000 to 10%(*w*/*v*). Crystals were then flash-cooled for data collection at 100 K

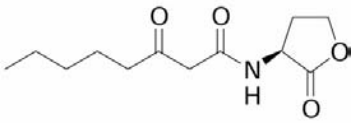
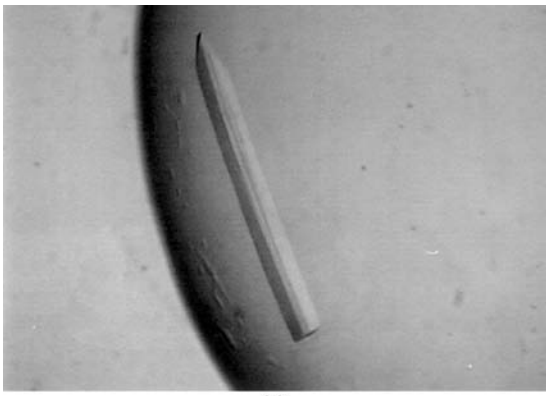
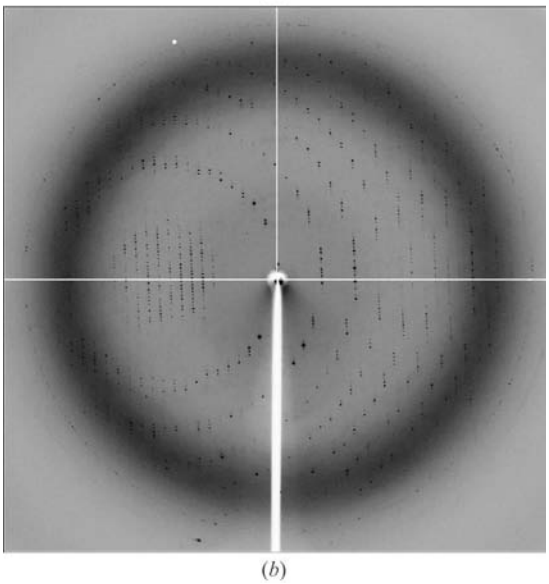


Figure 1
TraR cofactor, the pheromone *N*-3-oxooctanoyl-L-homoserine lactone (AAI).



(*a*)



(*b*)

Figure 2
(*a*) Typical crystal of $(TraR–AAI)_2$–TraBox complex. (*b*) Diffraction image from SeMet-derivatized crystals.

Table 1
Summary of X-ray diffraction data.

Values in parentheses are for the highest resolution bin.

	Se MAD†			Monochromatic‡	
	Inflection	Peak	Remote	SeMet	SeMet/2Br
Wavelength (Å)	0.979349	0.979186	0.914997	0.934	0.934
Resolution (Å)	50–3.0 (3.16–3.0)	50–3.0 (3.16–3.0)	50–3.0 (3.16–3.0)	30–2.79 (2.9–2.79)	50–3.06 (3.24–3.06)
R_{sym}§ (%)	9.0 (56.9)	8.5 (47.9)	9.5 (67.2)	6.0 (47.9)	9.0 (38.5)
$\langle I/\sigma(I)\rangle$	21.1 (5.2)	21.8 (5.8)	20.6 (4.3)	17.7 (4.7)	17.4 (3.1)
No. observations	274775	278470	275075	271940	288989
No. unique observations	27075	27075	27136	37181	30384
Completeness (%)	98.8 (98.1)	98.8 (98.2)	98.5 (98.6)	99.2 (100)	99.9 (99.7)

† Data collected at ESRF beamline ID29 using an ADSC Q210 detector with a crystal-to-detector distance of 280 mm and an oscillation range of 1°. ‡ Data collected at ESRF beamline ID14-H1 using an ADSC Q4R detector with crystal-to-detector distances of 260 and 280 mm and oscillation ranges of 0.8 and 1.0° for SeMet and SeMet/2Br crystals, respectively. § $R_{sym} = \sum_h \sum_i |\langle I_h\rangle - I_i(h)/\sum_h \sum_i I_i(h)$, where $I_i(h)$ is the mean intensity of the N reflections.

after stepwise transfer from the stabilizing solution to the final cryosolution [50 m*M* MES buffer pH 6.2, 200 m*M* calcium acetate, 25%(*w*/*v*) PEG 8000, 25% glycerol and 3 m*M* DTT] by gradually equilibrating them with increasing concentrations of PEG and glycerol.

2.4. Data collection

All data were collected at 100 K using synchrotron radiation at the European Synchrotron Radiation Facility (ESRF), Grenoble. MAD data were collected at beamline ID29 using an ADSC Quantum 210 charge-coupled device (CCD) detector. A crystal scan was performed to determine the wavelengths for data collection. Data to 3.0 Å resolution (Fig. 2*b*) were collected from a randomly oriented crystal at the Se absorption peak (12.6620 keV), at the absorption edge (12.6599 keV) and at a remote reference energy (13.5502 keV). Monochromatic data were collected on beamline ID14-H1 using an ADSC Q4R CCD detector on an SeMet crystal at 2.79 Å resolution and on an SeMet/2Br crystal at 3.06 Å resolution. Data were processed with *MOSFLM* (Leslie *et al.*, 1986) and scaled with *SCALA* (Evans, 1993). Crystals belong to space group $P2_12_12_1$, with unit-cell parameters a = 66.99, b = 94.67, c = 209.66 Å. Data-collection statistics are summarized in Table 1. Assumption of the presence of two (TraR–AAI)$_2$–TraBox complexes in the asymmetric unit gave a reasonable Matthews coefficient (Matthews, 1968) of 2.54 Å^3 Da^{-1} and a solvent content of 49.6% (assuming a molecular density of 1.30 g cm^{-3}). Therefore, 24 crystallographically independent selenium sites had to be located as part of the MAD-phasing process (six SeMet per protein monomer, including the starting Met).

2.5. Substructure solution and initial phases

The selenium substructure was determined with the *SnB* v2.1 package (Weeks & Miller, 1999). The peak-wavelength anomalous differences were processed with the *DREAR* suite (Blessing *et al.*, 1996) to generate difference normalized structure factors (diff*E* values). The largest 600 diff*E* values were used in *SnB* v2.1 to generate 6000 triplet invariants. Each trial was processed for 36 cycles of dual-space refinement. All these trials were carried out using an electron-density grid size of 1.0 Å and a minimum inter-peak distance of 3.0 Å. After 1000 trials had been processed, one trial was identified that had a significantly lower minimal function value (R_{min} = 0.308) than the others (R_{min} = 0.48–0.697). In all, 20 of the 24 independent selenium positions were located. These were confirmed in *CNX* (Accelrys, Pharmacopeia Inc.) by calculating and checking the anomalous difference Patterson maps (15–4.5 Å) from the experimental data and from the known locations of the selenium sites. Heavy-atom refinement and phase calculations were carried out at 3.0 Å resolution using the peak and the edge (reference) data in *CNX*, followed by solvent flattening with *CNX*. The figure of merit before and after solvent flattening was 0.477 and 0.915, respectively. The electron-density map, even without using any form of NCS averaging, shows interpretable regions of both protein and DNA. Model building and refinement are in progress.

We thank Bill Shepard and Andy Thompson for assistance in the MAD data collection at beamline ID29, and Elspeth Gordon and Joanne McCarthy for assistance in the monochromatic data collection at beamline ID14-H1, ESRF, Grenoble.

References

Blessing, R. H., Guo, D. Y. & Langs, D. A. (1996). *Acta Cryst.* D**52**, 257–266.
Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.
Fuqua, W. C & Winans, S. C. (1994). *J. Bacteriol.* **176**, 2796–2806.
Fuqua, C., Winans, S. C. & Greenberg, E. P. (1996). *Annu. Rev. Microbiol.* **50**, 727–751.
Hendrickson, W. H. (1991). *Science*, **254**, 51–58.
Leslie, A. G. W., Brick, P. & Wonacott, A. T. (1986). *Daresbury Lab. Inf. Quart. Protein Crystallogr.* **18**, 33–39.
McPherson, A. (1999). *Crystallization of Biological Macromolecules.* New York: Cold Spring Harbor Laboratory Press.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Orsatti, L., Di Marco, S., Volpari, C., Vannini, A., Neddermann, P. & Bonelli, F. (2002). In the press.
Piper, K. R., Beck von Bodman, S. & Farrand, S. K. (1993). *Nature (London)*, **362**, 448–450.
Piper, K. R. & Farrand, S. K. (2000). *J. Bacteriol.* **182**, 1080–1088.
Qin, Y., Luo, Z. Q., Smyth, A. J., Gao, P., Beck von Bodman, S. & Farrand, S. K. (2000). *EMBO J.* **19**, 5212–5221.
Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.
Zhang, L., Murphy, P. J., Kerr, A. & Tate, M. E. (1993). *Nature (London)*, **362**, 446–448.
Zhu, J. & Winans, S. C. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 4832–4837.
Zhu, J. & Winans, S. C. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 1507–1512.