Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and X-ray diffraction analysis of the sensor domain of the HemAT aerotactic receptor

Wei Zhang[a] and George N. Phillips Jr[a,b]*

[a]Department of Biochemistry and Cell Biology, Rice University, Houston, TX 77005, USA, and [b]Department of Biochemistry and Computer Sciences, University of Wisconsin-Madison, Madison, WI 53706, USA

Correspondence e-mail: phillips@biochem.wisc.edu

Received 12 December 2002
Accepted 10 February 2003

HemAT is a 432-amino-acid protein with two structural domains identified from *Bacillus subtilis*. It is responsible for sensing oxygen and delivering the signal to the downstream signal transduction cascade through a two-component system [Hou *et al.* (2000), *Nature (London)*, **403**, 540–544]. The consequence of such events is to change the flagellar movement and alter the swimming behavior of bacteria. To elucidate the molecular mechanism of oxygen sensing, the sensor domain of HemAT from *B. subtilis* was cloned, expressed and crystallized. Multiple-wavelength anomalous dispersion (MAD) data were collected from the intrinsic anomalous scatterer, iron, using synchrotron radiation. Three-wavelength iron MAD data sets were collected to 2.8 Å resolution. The native data set was collected to 2.15 Å resolution. Initial crystallographic analysis revealed the crystals to belong to space group $P2_12_12_1$, with unit-cell parameters $a = 50.00$, $b = 80.12$, $c = 85.95$ Å. There is one dimer in the asymmetric unit, with 40% solvent content. Structure determination using MAD methods and model building are currently under way.

1. Introduction

Sensing changes in the surrounding environment and making corresponding responses are essential for the growth and survival of living organisms. Chemotaxis is a vital survival strategy which causes movement of bacterial cells in response to chemicals present in the nearby environment, such as sugars, amino acids *etc.* Attractants and repellents can make bacteria move toward or away from their sources, respectively. The bacterial chemotactic response is one of the most extensively studied signal transduction systems (Djordjevic & Stock, 1998; Manson *et al.*, 1998; Stock & Mowbray, 1995). Aerotaxis is one example of a chemotactic response and is how bacteria search for a favorable gas environment.

Most chemoreceptors are methyl-accepting chemotaxis proteins (MCP) that are methylated or demethylated during excitation and adaptation (Le Moual & Koshland, 1996). External stimulants can either bind directly to the sensor domains of chemoreceptors or causes effects on receptors indirectly through soluble periplasmic binding proteins (Sharff *et al.*, 1992; Yeh *et al.*, 1993). The chemoreceptors do not influence the flagellar rotation directly, but rather utilize a two-component regulatory system that is commonly used in signal transduction pathways in prokaryotes and eukaryotes (Djordjevic & Stock, 1998). Two-component systems are composed of a histidine kinase protein and a response regulator (Stock *et al.*, 2000). A signaling ligand or ligand-bound receptor can interact with the histidine kinase protein to induce its autophosphorylation at a conserved histidine residue. The phosphate on the phosphorylated histidine kinase can then be transferred to an aspartate residue of its specific partner, a response-regulator protein. Formation of the phosphoaspartate in the regulator protein will activate downstream events in the signaling pathway, resulting in initiation of the transcription of specific genes or in changing swimming patterns by interacting with flagellar motor proteins (Bunn & Poyton, 1996; Stock *et al.*, 2000).

Thus far, most chemoreceptors that have been identified indicate transmembrane regions in the amino-acid sequence (Grebe & Stock, 1998; Le Moual & Koshland, 1996). However, several soluble chemoreceptors have been reported, including aerotactic HemATs from *Bacillus subtilis* and *Halobacterium salinarum* and another soluble chemoreceptor Car, which senses arginine (Hou *et al.*, 2000; Storch *et al.*, 1999).

HemAT (heme-based aerotactic transducer) is a soluble heme-containing protein that has recently been identified from *B. subtilis* (Hou *et al.*, 2000). It is a 432 amino-acid protein with two structural domains, the sensing domain (residues 1–178) and the signaling domain (residues 190–432). The amino-terminal 100 or so residues show ~20% sequence similarity to other heme-containing proteins, in particular

myoglobin and rice hemoglobin. The carboxy-terminal sequence is homologous to the cytoplasmic domain of a variety of bacterial chemotactic receptors, including the aspartate receptor Tar and the serine receptor Tsr (Hou *et al.*, 2000). This sensor domain plays a critical role in the chemotaxis signal transduction pathway. The atomic structure of the HemAT sensor domain is needed to develop a sound molecular mechanism for the initial signal transduction.

2. Materials and methods

2.1. Protein expression and purification

The HemAT gene from *B. subtilis* was commercially cloned into a pET29b expression plasmid by PCR methods (ATG Inc.). In addition to the six histidine residues at the N-terminus of the HemAT sequence, a factor Xa cleavage site (Ile-Glu-Gly-Arg) sequence was inserted before the first residue, methionine, of the sensor domain HemAT (residues 1–178). *Escherichia coli* [BL21(DE3)] cells transformed with the pET29b-HemAT plasmid were grown overnight at 310 K on 2× YT agar plates with addition of 30 µg ml^{-1} kanamycin. A single colony was picked to inoculate 1 l 2× YT media containing antibiotics. Protein induction was carried out by the addition of IPTG to a final concentration of 1 m*M*. The growth temperature was adjusted to 301 K after induction and growth continued for 6–10 h.

The cell pellet was resuspended in buffer *A* (300 m*M* NaCl, 50 m*M* Tris–HCl pH 8.0, 10 m*M* imidazole) and was broken with a small amount of lysozyme and DNase I at 295 K. After centrifugation at 277 K, the clear supernatant was applied to a nickel-chelating column. After extensive washing with buffer *A*, the final protein product was eluted with buffer *A* containing 500 m*M* imidazole. Factor Xa enzyme (Novagen) digestion was carried out to remove the plasmid-encoded residues after the first step of purification. The time courses of the enzymatic digestion reaction for the sensor domain were monitored by SDS–PAGE. Two additional purification steps, anion-exchange and hydroxyapatite chromatography (resin purchased from Pharmacia and BioRad, respectively), were carried out after the enzymatic cleavage in order to obtain high-purity protein. Basically, after dialyzing against buffer *B* (20 m*M* NaCl, 50 m*M* Tris–HCl pH 8.5), the protein sample was loaded onto an anion-exchange column and eluted with a linear gradient of buffer *C* (1 *M* NaCl, 50 m*M* Tris–HCl pH 8.5). For the last purification step, the buffer of the protein sample was exchanged to 10 m*M* NaH_2PO_4 pH 6.8 and loaded onto a hydroxyapatite column. The final protein was eluted with a linear phosphate gradient to a final concentration of 500 m*M* NaH_2PO_4 pH 6.8.

Table 1
Statistics of MAD data collection.

The wavelengths used in the data collection and their corresponding resolutions are listed. Values in parentheses are for the highest resolution shell.

Wavelength (Å)	Resolution (Å)	Reflection	Completeness (%)	R_{sym} (%)
1.7416	2.85	8218 (649)	96.5 (78.8)	0.103 (0.223)
1.7400	2.90	7902 (682)	97.4 (87.1)	0.112 (0.333)
1.6712	3.10	6467 (552)	97.0 (85.8)	0.113 (0.217)
1.0000	2.15	18610 (1588)	96.1 (84.1)	0.067 (0.220)

2.2. Crystallization and data collection

The protein used in crystallization setups was prepared by converting purified HemAT to a cyano-liganded form using an excess of potassium cyanide. The sample was then passed through a small Sephadex G-25 column equilibrated with 10 m*M* phosphate buffer pH 7.0, 5 m*M* NaCl and 4% glycerol. The final protein concentration was 10–15 mg ml^{-1}. The sample was aliquoted and immediately frozen at 193 K. When performing crystallization experiments, the protein sample was thawed and centrifuged using a bench-top microcentrifuge at 14 000 rev min^{-1} for 15 min in a cold room. Crystallization conditions were initially screened using the hanging-drop vapor-diffusion method, employing a variety of precipitants at different temperatures.

Multiple-wavelength anomalous dispersion (MAD) data were collected at the iron edge using flash-cooling techniques at the APS BioCARS beamline on an ADSC Quantum-4 detector. The cryoprotectant consisted of 2 *M* sodium citrate, 20% glycerol and 1 *M* lithium sulfate. The X-ray fluorescence scan was taken near the iron edge (7.12496 keV). Although the absorption signal was weak, three data sets were collected at peak, inflection point and remote wavelengths. Additionally, another native data set was collected from the same crystal after MAD data collection. The native data were collected to 2.15 Å resolution. The collected data were processed on-site using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). A summary of the data-collection statistics is given in Table 1.

3. Results and discussion

3.1. Protein production and initial crystal characterization

B. subtilis HemAT sensor domains were heterologously expressed in *E. coli*. The cell-growth temperature was initially 310 K for this strain both before and after IPTG induction. However, we found that a large portion of the HemAT protein was in an insoluble form under these conditions. Lowering the growth temperature and prolonging the induction time were very helpful methods in obtaining soluble HemAT and yielded 20–30 mg of protein per litre of cell culture. No further effort was taken to solubilize the inclusion bodies, since large amounts of protein were in soluble form.

The nickel-affinity column facilitated the purification by providing relatively high purity protein in one step. After the enzymatic cleavage of the His tag, the purity of the HemAT protein was further improved by anion-exchange and hydroxyapatite chromatography. The sample purity at each purification step was monitored by SDS–PAGE (Fig. 1). The purity of the protein sample used for crystallization was examined by electrospray mass spectrometry and the measured mass was within 2 Da of the expected value (data not shown). Protein used for crystallization was oxidized to the more stable cyano-liganded form, as initial

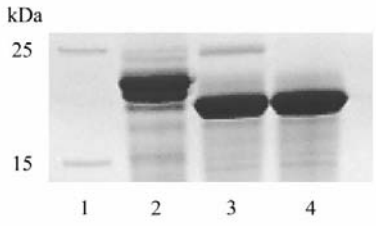

Figure 1
SDS–PAGE of purified *B. subtilis* HemAT sensor domain expressed in *E. coli*. Lane 1, protein standards, labelled on the left. Lane 2, elutant from nickel-chelating column. After breaking the cells and centrifugation, supernatant was loaded onto the nickel-chelating affinity column. Lane 3, after factor Xa enzymatic digestion. Lane 4, elutant from a hydroxyapatite column.

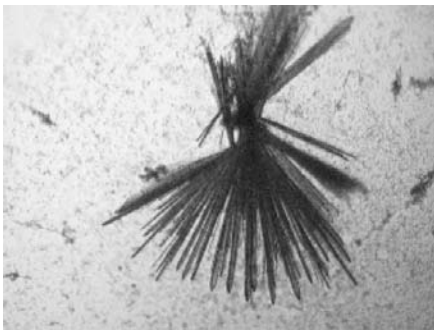

Figure 2
Cluster of reddish plate-like crystals of *B. subtilis* HemAT sensor domain. The typical dimensions of a single crystal are approximately 0.2 × 0.05 × 0.05 mm. Sodium citrate was used as a precipitant in the crystallization conditions.

trials using the unliganded form were not successful.

Crystals of the HemAT sensor domain were grown from 10–18% sodium citrate, 100 m*M* KH_2PO_4 pH 7.0 at room temperature. However, these crystals formed needle-like clusters and were unsuitable for X-ray diffraction studies. After testing crystallization conditions with detergent additives, improved crystals were obtained. Large and reproducible crystals were obtained using a 0.1%(*w*/*v*) concentration of *n*-octyl-β-D-glucoside (Sigma). The morphology was dramatically improved, but clusters were still formed. By breaking down the crystal clusters, single plate-like crystals were obtained. Crystals of HemAT showed a red color and had dimensions of about 0.2 × 0.05 × 0.05 mm for a single plate-shaped crystal (Fig. 2).

Solvent-content analysis of HemAT sensor-protein crystals showed that the Matthews coefficient would be 4.12 $Å^3\ Da^{-1}$ and the solvent content 69.9% if there was one subunit in the asymmetric unit. The solvent content would be 39.8% and the Matthews coefficient 2.06 $Å^3\ Da^{-1}$ for two subunits in the asymmetric unit. It is therefore reasonable to assume that two subunits exist in one asymmetric unit. The space group was determined to be $P2_12_12_1$ based on systematic absences, with unit-cell parameters a = 50.00, b = 80.1, c = 85.95 Å.

3.2. Iron MAD phasing

There is limited sequence similarity between the HemAT sensor domain and other members of the myoglobin family. However, all attempts to solve the sensor-domain structure using molecular replacement failed. The iron in the heme group was used as an anomalous scatterering candidate to collect a MAD data set. The X-ray absorption scan readily showed the Fe atom; MAD data were then collected at peak, inflection-point and remote wavelengths. The data were collected to 2.8 Å resolution (Table 1) near the iron edge. A native data set was also collected to a resolution of 2.15 Å from the same crystal after MAD data collection (1.00 Å wavelength).

The initial heavy-atom positions were found by *SOLVE* (Terwilliger & Berendzen, 1999) from the three-wavelength MAD data (peak, inflection point and remote). However, the map produced from *SOLVE* was noisy and did not reveal any recognizable protein features. The model built by *RESOLVE* based on this map gave only a few candidate helices. Another approach used was to find the heavy-atom sites using programs from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). Four data sets, including the native data, were processed and scaled using the programs *TRUNCATE* and *SCALEIT*. Direct methods were used to find the heavy atoms using *RANTAN*. The heavy-atom sites identified were refined with *MLPHARE*. Two heavy atoms were found in the asymmetric unit. It was found that the sites identified from *CCP*4 and from *SOLVE* were symmetry-related. There were two heavy-atom sites corresponding to the iron positions from two subunits in one asymmetric unit. We then found that several helices produced by *RESOLVE* bore non-crystallographic twofold symmetry. Along with homology model building based on other heme proteins, the map was improved after density modification taking the initial phases from *MLPHARE*. The protein chain has been traced from the map. Refinement is in progress.

The crystal structure of the sensor domain will help us to understand the molecular mechanism of receptor-mediated signaling transduction. It will provide accurate structural information on the composition of the ligand-binding pocket, where the first events occur in the chemotaxis signal transduction cascade.

We wish to thank staff Gary Navrotski, Keith Brister, Bill Desmarais and Michael Bolbat at the BioCARS beamline at the APS for helping us to collect the iron-edge MAD data and the native data set. This work was supported by the Robert A. Welch Foundation C-1142 (GNP), the W. M. Keck Center for Computational Biology and the Wisconsin Alumni Research Foundation.

References

Bunn, H. F. & Poyton, R. O. (1996). *Physiol. Rev.* **76**, 839–885.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Djordjevic, S. & Stock, A. M. (1998). *J. Struct. Biol.* **124**, 189–200.
Grebe, T. W. & Stock, J. (1998). *Curr. Biol.* **8**, R154–R157.
Hou, S., Larsen, R. W., Boudko, D., Riley, C. W., Karatan, E., Zimmer, M., Ordal, G. W. & Alam, M. (2000). *Nature (London)*, **403**, 540–544.
Le Moual, H. & Koshland, D. E. Jr (1996). *J. Mol. Biol.* **261**, 568–585.
Manson, M. D., Armitage, J. P., Hoch, J. A. & Macnab, R. M. (1998). *J. Bacteriol.* **180**, 1009–1022.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sharff, A. J., Rodseth, L. E., Spurlino, J. C. & Quiocho, F. A. (1992). *Biochemistry*, **31**, 10657–10663.
Stock, A. M. & Mowbray, S. L. (1995). *Curr. Opin. Struct. Biol.* **5**, 744–751.
Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). *Annu. Rev. Biochem.* **69**, 183–215.
Storch, K. F., Rudolph, J. & Oesterhelt, D. (1999). *EMBO J.* **18**, 1146–1158.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Yeh, J. I., Biemann, H. P., Pandit, J., Koshland, D. E. & Kim, S. H. (1993). *J. Biol. Chem.* **268**, 9787–9792.