Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of neural haemoglobin from the nemertean worm *Cerebratulus lacteus*

Alessandra Pesce,[a] Marco Nardini,[a] Sylvia Dewilde,[b] Paolo Ascenzi,[a,c] Austen F. Riggs,[d] Kiyoshi Yamauchi,[e] Eva Geuens,[b] Luc Moens[b] and Martino Bolognesi[a]*

[a]Department of Physics – INFM, Advanced Biotechnology Centre, University of Genova, Largo Rosanna Benzi 10, I-16146 Genova, Italy, [b]Department of Biochemistry, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium, [c]Department of Biology, University 'Roma Tre', Viale Guglielmo Marconi 446, I-00146 Roma, Italy, [d]Section of Neurobiology, School of Biological Sciences, University of Texas, CO920 Austin, Texas TX 78712-1064, USA, and [e]Department of Biology and Geoscience, Faculty of Science, Shizuoka University, 836 Oya, Shizuoka 422-8529, Japan

Correspondence e-mail: bolognes@fisica.unige.it

Received 18 July 2001
Accepted 26 September 2001

The nemertean worm *Cerebratulus lacteus* neural tissue haemoglobin (109 amino acids, the shortest known haemoglobin) has been overexpressed in *Escherichia coli*, purified and crystallized. A highly redundant native data set has been collected at the Cu $K\alpha$ wavelength to 2.05 Å resolution. The crystals belong to the orthorhombic $P2_12_12_1$ space group, with unit-cell parameters a = 42.5, b = 43.1, c = 60.2 Å and one molecule per asymmetric unit. The anomalous difference Patterson map clearly reveals the position of the haem Fe atom, thus paving the way for MAD/SAD structure determination.

1. Introduction

Haemoglobins (Hbs) occur in all the kingdoms of living organisms and display a variety of functions. In addition to oxygen transport and storage, several novel Hb functions have emerged, including NO metabolism and dehaloperoxidase activity (Ascenzi *et al.*, 2001; Weber & Vinogradov, 2001; Witting *et al.*, 2001). Despite the large variability in their primary and quaternary structures, Hbs display a well conserved tertiary structure (the 'globin fold'), based on 140–160 amino acids, typically characterized by a three-on-three α-helical sandwich (Perutz, 1979; Bolognesi *et al.*, 1997; Riggs, 1998). However, considerable editing of the 'classical' globin fold has recently been characterized. In particular, small haemoproteins, known as 'truncated Hbs', have been identified in unicellular organisms, in a unicellular alga and in some higher plants (Moens *et al.*, 1996). They consist of 115–130 amino acids arranged in a two-on-two α-helical sandwich, which can be recognized as a subset of the classical globin fold (Pesce *et al.*, 2000; Milani *et al.*, 2001).

The nemertean worm *C. lacteus* hosts a blood Hb, so far uncharacterized. A second Hb is found in the body-wall tissues and a third is expressed in the neural ganglia and lateral nerve cords (Vandergon *et al.*, 1998). The latter two Hbs are composed of 109 residues, thus being the shortest stable Hbs known. Alignment of *C. lacteus* Hb (CerHb) amino-acid sequences with those of (in)vertebrate haemoglobins indicates substantial structural perturbation of the *A*, *B* and *H* helices. Moreover, lack of typical structural determinants, such as Gly-Gly motifs, suggests that CerHb may deviate not only from the classical globin fold, but also from the recently discovered truncated Hb fold (Pesce *et al.*, 2000).

The neural CerHb has been suggested to store O_2, extending neuronal activity in an anoxic environment (Vandergon *et al.*, 1998). Remarkably, neural specific Hbs occur in a number of invertebrates (Dewilde *et al.*, 1996) and have very recently also been discovered in mouse and human brain (Burmester *et al.*, 2000). However, their physiological role is still open to debate.

Here, we describe the expression, purification, crystallization and preliminary X-ray diffraction analysis of the neural Hb of *C. lacteus* as the first step in the elucidation of its atomic structure.

2. Materials and methods

2.1. Cloning, expression and purification of CerHb

A neural Hb was identified and sequenced in *C. lacteus* as previously described by Vandergon *et al.* (1998). A synthetic *C. lacteus* cDNA was constructed using the codon frequency of *E. coli* from 15 oligonucleotides according to the method of Ikehara *et al.* (1984). The synthetic globin gene was designed to be 335 bp long, with a *Nde*I site at the 5′ end and a *Bam*HI site at the 3′ end. It was ligated into the corresponding restrictions site of the pET3a expression vector. The correctness of the construct was verified by dideoxy sequencing. The expression plasmid was subsequently transformed in *E. coli* strain BL21(DE3)pLysS and expressed in TB medium as described previously (Dewilde *et al.*, 1998; Das *et al.*, 2000).

CerHb was purified from the total lysate by low-speed (10 000*g*) and high-speed (105 000*g*) centrifugation, ammonium sulfate precipitation (40 and 90% saturated), DEAE fast-flow chromatography and finally by two cycles of gel filtration on a Sephacryl S200 high-resolution column (Pharmacia). The purity of the fractions was monitored by SDS–PAGE; 36 mg of pure CerHb were obtained per litre of medium.

2.2. CerHb crystallization

All the crystallization experiments were carried out at 277 K using vapour-diffusion techniques (hanging drop) at a protein concentration of 27 mg ml^{-1}. A preliminary screening varying the concentration of ammonium sulfate as precipitant at neutral pH and a subsequent pH screen at the optimal precipitant concentration were carried out. The best condition, corresponding to 2.6 *M* ammonium sulfate, 50 m*M* sodium acetate pH 5.5, resulted in bunches of elongated prismatic crystals with typical dimensions of 0.07 × 0.07 × 0.4 mm (Fig. 1) after one week of equilibration against the reservoir solution. The crystals were stored in 2.9 *M* ammonium sulfate, 50 m*M* sodium acetate pH 5.5 at 277 K and transferred to the same solution supplemented with 15%(*v*/*v*) glycerol immediately prior to data collection at 100 K.

2.3. Data collection and processing

A X-ray native data set from a CerHb crystal was collected at 100 K using a MAR Research 345 imaging-plate detector coupled to a Rigaku RU-H3R rotating-anode generator, using monochromatic Cu $K\alpha$ radiation, to a resolution limit of 2.05 Å. High redundancy and completeness of the data set were planned in order to optimize accuracy in measurement of the haem Fe-atom anomalous signal. After one week, at the end of the data collection, the crystal did not display signs of decay.

Figure 1
Crystals of the recombinant neural haemoglobin from the nemertean worm *C. lacteus*.

Diffraction data were processed using *DENZO*, *SCALEPACK* (Otwinowski & Minor, 1997) and programs from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

Inspection of the diffraction pattern and systematic absences allowed the assignment of the CerHb crystals to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 42.5$, $b = 43.1$, $c = 60.2$ Å (Table 1). Evaluation of the crystal packing parameter (Matthews, 1968) indicated that the lattice can accommodate one molecule per asymmetric unit (V_M = 2.3 Å^3 Da^{-1}), with an estimated solvent content of 47%.

Sequence-similarity searches did not highlight any significant structural homology (*i.e.* amino-acid identity > 15%) with Hbs of known three-dimensional structure. A molecular-replacement solution was extensively sought (*EPMR* program; Kissinger *et al.*, 1999) using search models derived from (in)vertebrate and truncated Hbs. However, the correct orientation and translation of the template molecules in the CerHb unit cell was not possible. Therefore, the crystallographic analysis of CerHb will be based on the multiple isomorphous replacement or multiple anomalous diffraction methods based on the haem Fe-atom absorption edge.

The high-redundancy data set collected in-house was used in order to locate the

Table 1
Data-collection statistics for CerHb.

Outer shell statistics (2.10–2.05 Å) are shown in parentheses.

Wavelength (Å)	1.5418
Resolution limit (Å)	2.05
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 42.5, b = 43.1, c = 60.2
Total reflections	65904
Unique reflections	7388
Completeness (%)	99.3 (98.5)
Redundancy	9 (10)
R_{merge} (%)	5.6 (10.9)
Average $I/\sigma(I)$	29 (14)
R_{anom} (%)	3.1
R_{anom}^{theor}† (%)	2.3

† Calculated according to Smith (1991).

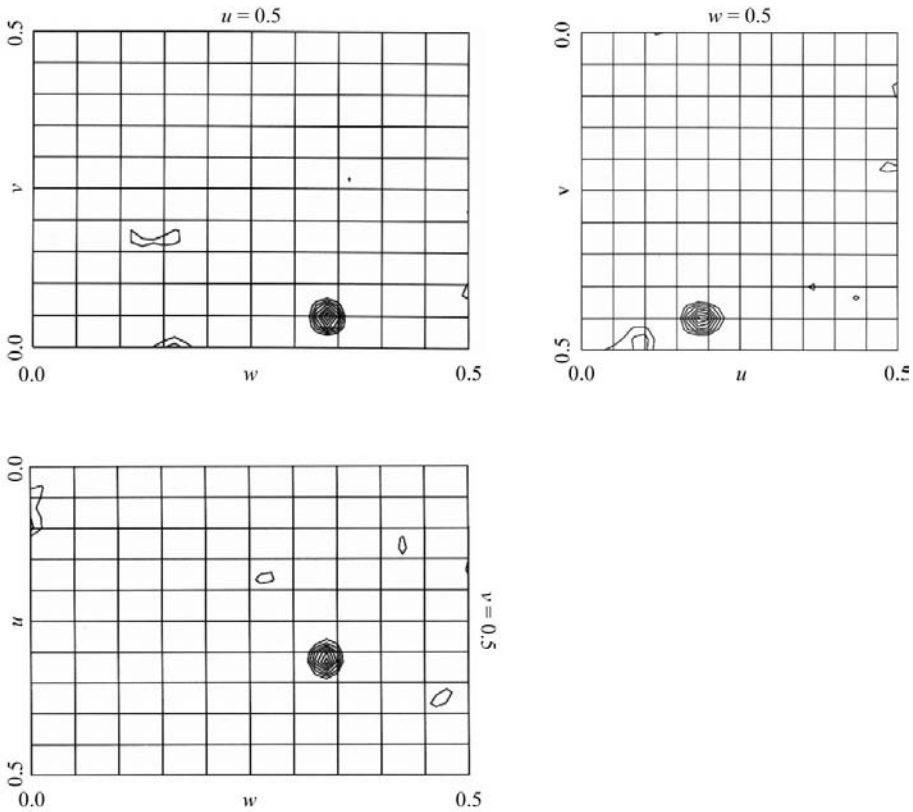

Figure 2
The three Harker sections (at $u = 0.5$, $v = 0.5$ and $w = 0.5$) of the anomalous difference Patterson map, showing the haem Fe-atom peak, calculated in the 30–2.05 Å resolution range. Contours are drawn at 2σ levels, starting from 3σ.

haem Fe-atom site, considering that at the Cu $K\alpha$ wavelength the Fe anomalous contribution to the scattering factor (f'') is 3.2 e^- (http://www.bmsc.washington.edu/scatter/AS_index.html; see Table 1). An anomalous difference Patterson map calculated in the resolution range 30–2.05 Å with the program *PHASES* (Furey & Swaminathan, 1997) clearly showed coherent 15–17σ peaks in the three Harker sections at u = 0.5, v = 0.5 and w = 0.5 (Fig. 2). Remarkably, the anomalous Patterson peaks could be clearly recognized even using data at resolutions as low as 4.0 Å, although in this case the peak height drops to 5σ.

CerHb structure solution using SAD/MAD methods based on the haem Fe-atom position is currently in progress. An electron-density map calculated with SAD phases at 2.05 Å resolution using the program *SOLVE* (Terwilliger & Berendzen, 1999) and solvent flattened with *DM* (Cowtan, 1994) shows interpretable features, particularly in the haem-pocket region.

This work was supported by grants from the Italian Space Agency (IR/167/01) and from the National Research Council (CNR) Projects Functional Genomics and Biotechnology to MB and by a grant from the Fund for Scientific Research-Flanders (project No. G.0069.98) to LM. SD is a post-doctoral fellow of the latter fund.

References

Ascenzi, P., Salvati, L. & Brunori, M. (2001). In the press.
Bolognesi, M., Bordo, D., Rizzi, M., Tarricone, C. & Ascenzi, P. (1997). *Prog. Biophys. Mol. Biol.* **68**, 29–68.
Burmester, T., Weich, B., Reinhardt, S. & Hankeln, T. (2000). *Nature (London)*, **407**, 529–523.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50,** 760–763.
Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
Das, T. K., Weber, R. E., Dewilde, S., Wittenberg, J. B., Wittenberg, B. A., Yamauchi, K., Van Hauwaert, M. L., Moens, L. & Rousseau, D. L. (2000). *Biochemistry*, **39**, 14330–14340.
Dewilde, S., Blaxter, M., Van Hauwaert, M.-L., Vanfleteren, J., Esmans, E., Marden, M., Griffon, N. & Moens, L. (1996). *J. Biol. Chem.* **271**, 19865–19870.
Dewilde, S., Blaxter, M., Van Hauwaert, M.-L., Van Houte, K., Pesce, A., Griffon, N., Kiger, L., Marden, M. C., Vermeire, S., Vanfleteren, J., Esmans, E. & Moens, L. (1998). *J. Biol. Chem.* **273**, 32467–32474.
Furey, W. & Swaminathan, S. (1997). *Methods Enzymol.* **277**, 590–620.
Ikehara, M., Ohtsuka, E., Tokunaga, T., Taniyama, Y., Iwai, S., Kitano, K., Miyamoto, S., Ohgi, T., Sakuragawa, Y., Fujiyama, K., Ikari, T., Kobayashi, M., Miyake, T., Shibahara, S., Ono, A., Ueda, T., Tanaka, T., Baba, H., Miki, T., Sakurai, A., Oishi, T., Chisaka, O. & Matsubara, K. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 5956–5960.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Milani, M., Pesce, A., Ouellet, Y., Ascenzi, P., Guertin, M. & Bolognesi, M. (2001). *EMBO J.* **20**, 3902–3909.
Moens, L., Vanfleteren, J., Van de Peer, Y., Peeters, K., Kapp, O., Czeluzniak, J., Goodman, M., Blaxter, M. & Vinogradov, S. (1996). *Mol. Biol. Evol.* **13**, 324–333.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perutz, M. F. (1979). *Annu. Rev. Biochem.* **48**, 327–386.
Pesce, A., Couture, M., Dewilde, S., Guertin, M., Yamauchi, K., Ascenzi, P., Moens, L & Bolognesi, M. (2000). *EMBO J.* **19**, 2424–2434.
Riggs, A. F. (1998). *J. Exp. Biol.* **201**, 1073–1084.
Smith, J. L. (1991). *Curr. Opin. Struct. Biol.* **1**, 1002–1011.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Vandergon, T. L., Riggs, C. K., Gorr, T. A., Colacino, J. M. & Riggs, A. F. (1998). *J. Biol. Chem.* **273**, 16998–17011.
Weber, R. & Vinogradov, S. N. (2001). *Phys. Rev.* **81**, 569–628.
Witting, P. K., Douglas, D. J. & Mauk, A. G. (2001). *J. Biol. Chem.* **276**, 3991–3998.