The crystal structure of the *Escherichia coli* lipocalin Blc suggests a possible role in phospholipid binding

Valérie Campanacci[a,1], Didier Nurizzo[b,1], Silvia Spinelli[a], Christel Valencia[a,2], Mariella Tegoni[a], Christian Cambillau[a,*]

[a] *Architecture et Fonction des Macromolécules Biologiques, UMR 6098, CNRS and Universités Aix-Marseille I and II, 31 chemin J. Aiguier, F-13402 Marseille Cedex 20, France*
[b] *European Synchrotron Radiation Facility, ID23, P.O. Box 220, F-38043 Grenoble Cedex, France*

Received 19 January 2004; revised 19 February 2004; accepted 19 February 2004

First published online 4 March 2004

Edited by Hans Eklund

Abstract Lipocalins form a large multifunctional family of small proteins (15–25 kDa) first discovered in eukaryotes. More recently, several types of bacterial lipocalins have been reported, among which Blc from *Escherichia coli* is an outer membrane lipoprotein. As part of our structural genomics effort on proteins from *E. coli*, we have expressed, crystallized and solved the structure of Blc at 1.8 Å resolution using remote SAD with xenon. The structure of Blc, the first of a bacterial lipocalin, exhibits a classical fold formed by a β-barrel and a α-helix similar to that of the moth bilin binding protein. Its empty and open cavity, however, is too narrow to accommodate bilin, while the alkyl chains of two fatty acids or of a phospholipid could be readily modeled inside the cavity. Blc was reported to be expressed under stress conditions such as starvation or high osmolarity, during which the cell envelope suffers and requires maintenance. These data, together with our structural interpretation, suggest a role for Blc in storage or transport of lipids necessary for membrane repair or maintenance.

Key words: Lipocalin; Blc; Xenon; Remote SAD; Structural genomics; Lipid transport

1. Introduction

Lipocalins pertain to a large family of proteins which was identified initially in eukaryotes [1] and more recently in Gram-negative bacteria [2,3]. Lipocalins are considered to be lipophilic carriers (hence their name), but their functions (often putative) have been reported as very diverse [3,4] (see also papers in *Biochim. Biophys. Acta* Vol. 1482). Only a few lipocalins have a well defined and studied function. Among them, the retinol binding protein was the first whose three-dimensional structure has been solved [5]. Other lipocalins have been proposed to carry pheromones [6–8] or odorant molecules [9,10]. Outlier members of the lipocalin family, identified in the salivary gland of ticks, sequester histamine produced by the host [11]. Recently, the iron binding function of siderocalins, a new class of lipocalins, has been assigned by virtue of the crystal structure [12].

The lipocalin fold is very well conserved: it features an eight-stranded β-barrel and a C-terminal domain with an α-helix [13]. Decorations, such as an extra strand, may complete the structure. The number of disulfide bridges can vary from one to three in mammalian lipocalins, with the exception of bovine odorant binding protein which does not possess cysteines and forms a peculiar domain-swapped dimer [10]. The amino acid sequences of lipocalins, however, are very poorly conserved, and only a few signatures have been described, among which the most general is at the N-terminus and consists of a GXW motif [1,13]. Sequence comparisons and evolution analysis of lipocalins have led to a classification in which ~14 clades have been identified [14]. Clade 1 (the root clade) includes mainly bacterial lipocalins, clade 2 those from insects, whereas other eukaryotic and mammalian lipocalins are found in the higher clades (3–14).

Most bacterial lipocalins are lipoproteins which contain a type 2 signal peptide allowing export to the periplasm [2,3]. In the larger of the two identified groups, one or two cysteines are observed and no disulfide bridge is found. The first cysteine, located just after the signal peptide, is attached to the membrane anchor lipid forming a *N*-acyl-*S*-*sn*-1,2-diacylglycerylcysteine moiety at the N-terminus [2,3]. A smaller subgroup of bacterial lipocalins contains a disulfide bridge and its members have diverse predicted localizations in the cytoplasm, periplasm and the inner and outer membranes [2,3].

The paradigm for bacterial lipocalins is provided by the *blc* gene from *Escherichia coli*, the first bacterial lipocalin that has been identified [2]. Blc is an outer membrane bound protein, facing the periplasmic space. It is expressed during the steady phase, under stress conditions such as starvation or high osmolarity, during which the cell envelope suffers and requires maintenance [3]. A putative function of Blc might therefore be to store or carry lipids. Interestingly, the two close sequence neighbors of Blc in insects or mammals, namely Lazarillo [15] and ApoD [16], are believed to interact with lipids (Fig. 1A). It has been proposed that ApoD may participate in maintenance and repair in the central and peripheral nervous systems. ApoD could transport a ligand from one cell to another within an organ, scavenge a ligand within an organ for transport to the blood or could transport a ligand from the blood to specific cells within a tissue [17].

The three-dimensional structure of eukaryotic lipocalins is well documented for members of clades 2–14, but no structure

*Corresponding author. Fax: (33)-491-16 45 36.
E-mail address: cambillau@afmb.cnrs-mrs.fr (C. Cambillau).

[1] These authors contributed equally to this work.
[2] Present address: Institut Gilbert Laustriat, IFR85, 74 route du Rhin, P.O. Box 60024, F-67401 Illkirch Cedex, France.

0014-5793/04/$30.00
doi:10.1016/S0014-5793(04)00199-1

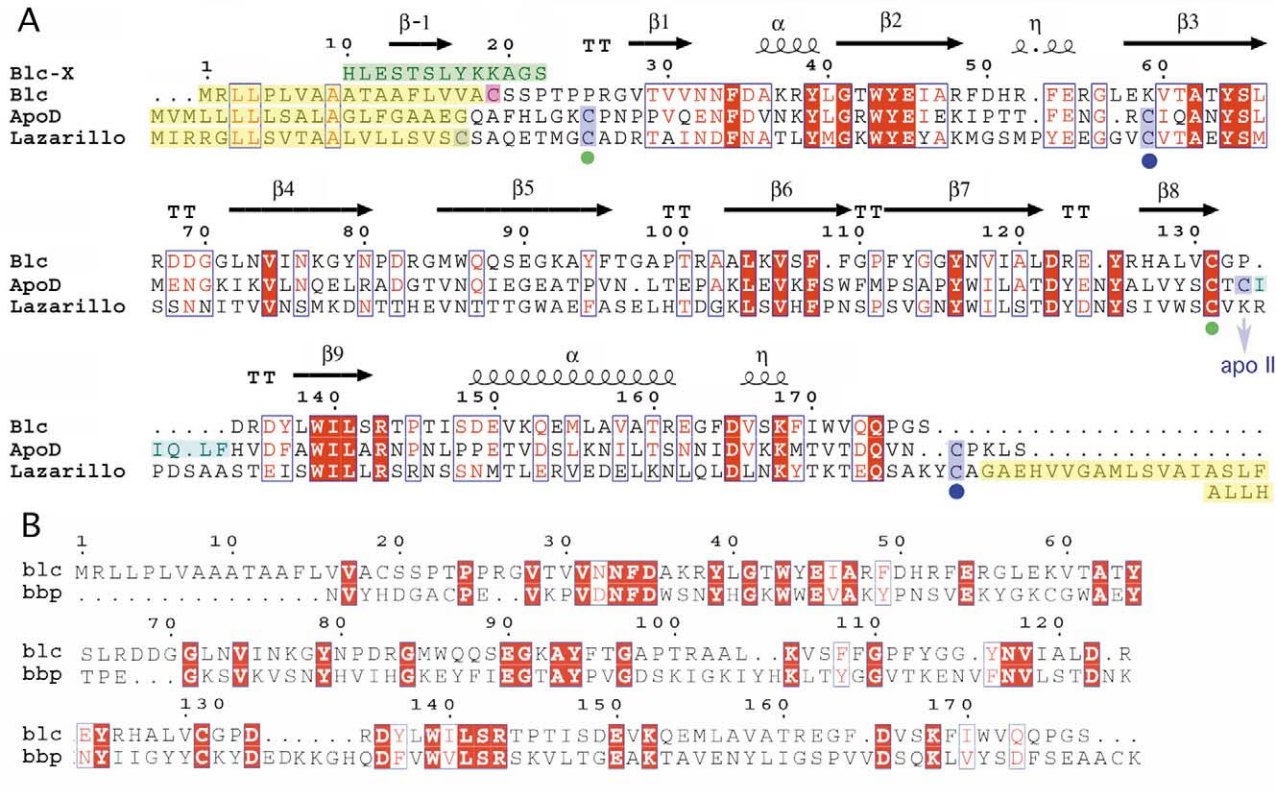

Fig. 1. A: Sequence alignment of Blc with human ApoD and *Schistocerca americana* Lazarillo. Identical residues are in white on a red background. The attb1 sequence at the N-terminus of the expressed protein is (Blc-X) boxed in green, and the hydrophobic loop of ApoD is boxed in light blue. The cysteines of ApoD and Lazarillo forming two disulfide bridges are identified by green and blue dots. The peptides cleaved at the N- and C-termini (leading to the mature proteins) are boxed in yellow. B: Sequence alignment of Blc with BBP [21].

of a true bacterial lipocalin was available. Recently, however, *E. coli* YodA protein was shown to possess a domain with a β-barrel fold resembling that of lipocalins [18]. In the course of a structural genomics program dealing with bacterial proteins, we have expressed, crystallized and characterized Blc. Its structure, solved by remote SAD methods with xenon [19,20], has been refined to 1.8 Å resolution. It reveals a fold similar to that of the moth bilin binding protein (BBP) [21] as expected from the 22% sequence identity between the two proteins (Fig. 1B). The elongated and open cavity present in Blc is narrower than the cavity of BBP. In consequence, large ligands such as bilins or hemes do not fit the cavity. Instead, the cavity could bind two fatty acid chains placed side to side.

2. Materials and methods

2.1. Expression and purification

Subcloning and expression strategies used for our *E. coli* targets, including the *blc* gene, have already been described elsewhere [22,23]. The *blc* gene segment designed for this study was almost complete, starting three residues after the cysteine residue used to anchor Blc to the membrane, thus yielding a putative soluble, well behaving protein. This shorter *blc* gene segment was amplified by polymerase chain reaction from *E. coli* K12 genomic DNA and was subcloned in the pDest17 vector by recombination [24] using the Gateway™ technology (Invitrogen). Expression was carried out using the Tuner(DE3)pLysS *E. coli* strain. The protein was soluble and was purified on a nickel column followed by gel filtration on a Superdex 200 column in HEPES 5 mM, NaCl 150 mM, pH 7.5.

2.2. Crystallization

Crystals of the protein were grown at 20°C using sitting drops and the nanodrop technology [23] followed by optimization [25]. A volume of 100 nl of a protein solution at 6–8 mg/ml in HEPES 5 mM, NaCl 150 mM, pH 7.5, was mixed with 100 nl of precipitant solution consisting of sodium citrate 800–900 mM, sodium borate 50 mM, pH 7.0–7.5. Crystals grew within 1–3 days and proved to belong to the orthorhombic space group $P2_12_12_1$ ($a = 58.0$ Å, $b = 80.8$ Å, $c = 89.0$ Å) with two molecules per asymmetric unit (Vm = 2.6 Å^3/Da; 59% solvent).

2.3. Data collection and processing

A native crystal was transferred into a cryo-solution made of the crystallization buffer supplemented with 25% glycerol and pressurized at 25 bar with xenon for 5 min using the Oxford Cryosystems XCell. Xenon-derivatized data were collected on a single flash-cooled crystal on beamline ID14-EH4 at the European Synchrotron Radiation Facility (ESRF) in Grenoble (France) using an ADSC Quantum 4 CCD detector. A total of 553 images were collected using a 1° oscillation range per image at 1.0332 Å wavelength. The native data were also collected on beamline ID14-EH4 at 0.9393 Å wavelength. Both data sets were processed with MOSFLM and scaling was performed using SCALA from the CCP4 [26] suite (Table 1).

2.4. Structure solution

Xenon positions were determined using SHELXD [27] with data in the resolution range from 20 Å to 2.3 Å. Two sites were found. Position refinement, phase calculation and density modification were performed using SHARP [28]. Phase extension using the isomorphous native data set was performed with DM [29]. ARP-wARP [30] version 6.0 integrated in CCP4i was run to build automatically the two molecules in the asymmetric unit. This final solution was assigned to its sequence using the X-Autofit option in Quanta (Accelerys, San Diego, CA, USA) and refined using REFMAC [31].

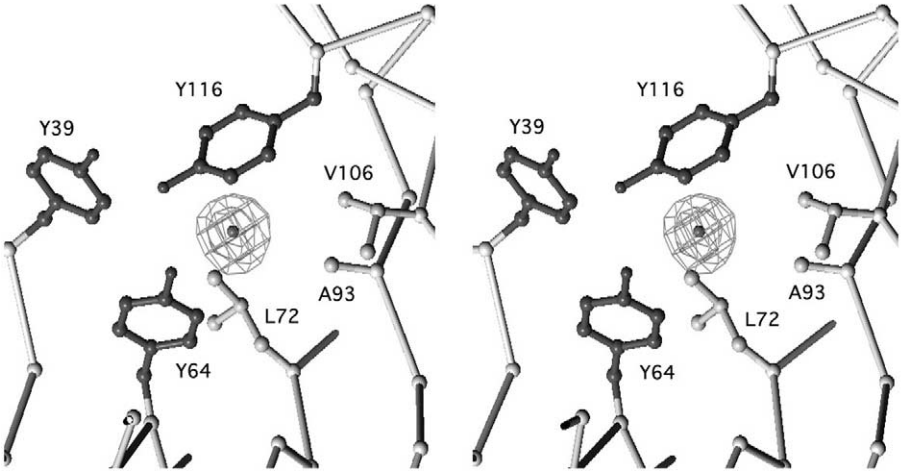

Fig. 2. Stereo view of the xenon binding site of Blc. The F_o-F_c electron density map is contoured at threshold levels of 5 and 8 σ (picture made with Turbo-Frodo [34]).

3. Results and discussion

3.1. Structure solution and xenon site

Although Blc shares 25% identity with BBP (Fig. 1B), its structure could not be solved by molecular replacement using AMoRe and the structure of BBP as a starting model. Since lipocalins are known to harbor a hydrophobic cavity, we decided to try high pressure soaking with xenon. Xenon has been frequently and successfully used as a MIR/SIR derivative [19] but not as an anomalous scatterer due to the high Xe L1 absorption edge (2.27 Å). We decided however to try a remote SAD approach and collected data at 1.03 Å resolution, providing 3.5 anomalous electrons. Such an approach has been described recently [20]. The structure could be readily solved (see Section 2) and the maps reveal the presence of a unique well ordered xenon per molecule, in a hydrophobic environment formed by Tyr 39, Tyr 64, Leu 72, Ala 93, Val 106 and Tyr 116, in a small pocket adjacent to the main cavity (Fig. 2).

3.2. The overall structure

The Blc structure has a typical lipocalin fold consisting of a β-barrel with eight antiparallel strands (strands 2–9, Fig. 3) and an α-helix at the C-terminus. An extra β-strand (strand 1, Fig. 3) does not belong to the β-barrel and weakly contacts strand 7 in a quasi-parallel fashion. A stretch of 13 residues, originating from the attb1 Gateway pDest17 vector recombination sequence [24], and a histidine from the 6×His tag, form an extended structure with a β-strand at its center (strand −1, Fig. 3), antiparallel to strand 1, and located opposite to the β-barrel open cavity.

Besides the N-terminal cysteine, which allows Blc to bind to the membrane through a lipid anchor, the single Cys 131 is buried 9 Å below the protein surface in a tight and hydrophobic environment: Pro 24, Asn 117 Cβ, Leu 119 and Phe 164. No particular function can be easily assigned to this residue. In the two eukaryotic lipocalins similar to Blc, Lazarillo and ApoD, the equivalent cysteine forms a disulfide bridge with another cysteine at the N-terminus (Fig. 1). Attachment to the membrane involves a GPI anchor in the case of Lazarillo [15] and binding of human ApoD [17] to ApoAII is performed through a disulfide bridge (see below) (Fig. 1) [3].

Table 1
Blc data collection and refinement

Data collection	Native
Detector type	ADSC Quantum 4R CCD
Radiation source	ESRF (ID14-EH4)
Wavelength (Å)	0.9393
Temperature (K)	100
Space group	$P2_12_12_1$
Cell (Å)	$a=58.0$, $b=80.8$, $c=88.95$
Resolution (Å)[a]	47.0–1.75 (1.84–1.75)
Completeness[a] (%)	98.1 (92.8)
R_{sym} [a] (%)	9.7 (36.9)
Multiplicity/$I/\sigma I$[a]	5.4 (2.7)/14.5 (1.8)
Refinement	
Resolution (Å)	47.0–1.75
Unique reflections	41 968
No. residues/water	333/306
R/R_{free}	0.169/0.204
Mean B value (Å^2)	19.9

[a]Values in parentheses are for the highest resolution shell (1.84–1.75 Å).

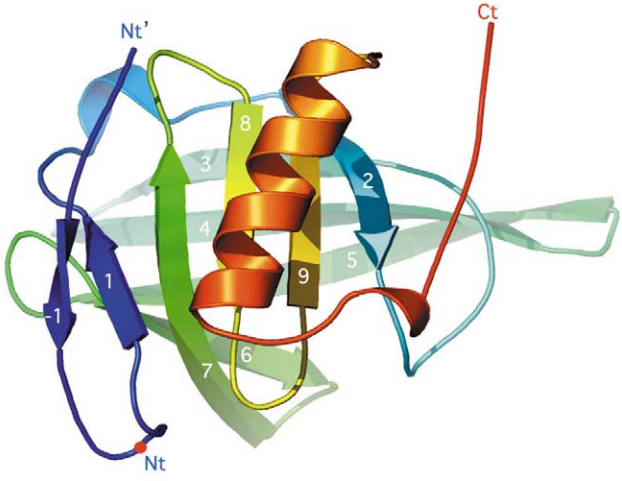

Fig. 3. Ribbon view of Blc, rainbow colored from N-terminus (blue) to C-terminus (red). The β-strands are numbered 1–9, plus the extra strand at the N-terminus (−1). The extra 12 residues at the N-terminus, from the attb1 sequence, are localized between the Nt′ and Nt labels. Blc itself starts at the red dot level (Nt) (picture made with PyMOL [35]).

Fig. 4. Superimposed ribbon view of Blc (blue) and BBP (magenta). Some strands, discussed in the text, have been numbered (picture made with PyMOL [35]).

When superimposing Blc to its closest three-dimensional neighbor, BBP [21], a rmsd value of 2.2 Å (on 160 Cα atoms) is obtained. Deviations occur at the helix, which is tilted differently, in the loops and along some strands (Fig. 4): strand 104–115, forming part of the wall of the cavity, has a position completely different in Blc and BBP. Strands 8 and 9, flanking the loop 132–137 in Blc, are much longer in BBP and form an extended structure of eight residues (Fig. 4). The loop between these strands is hydrophobic in ApoD, and has been postulated to mediate the interaction with HDL. It has been shown that a Cys residue at the beginning of the loop in human ApoD forms a disulfide bridge with ApoAII associated with HDL (Fig. 1) [3]. This cysteine does not exist in ApoDs from other species, and the hydrophobic loop anchors ApoD to the membrane.

In addition to the extra attb1 segment, Blc β-strand 1 also interacts with strand 7 and follows the same path as an extended structure in BBP. The most striking difference between the two lipocalins is in the N-terminus, which is ~10 residues longer in BBP than in Blc. At the very end of BBP, the N-terminal β-strand 1 interacts in an antiparallel manner with strand 9 and extends the size of the open cavity of BBP.

3.3. The cavity and its putative ligands

The calyx of Blc forms an open cavity ~18 Å deep by 12×7 Å. The cavity opening is located opposite to the real N-terminus of Blc (Fig. 3), the site of membrane attachment,

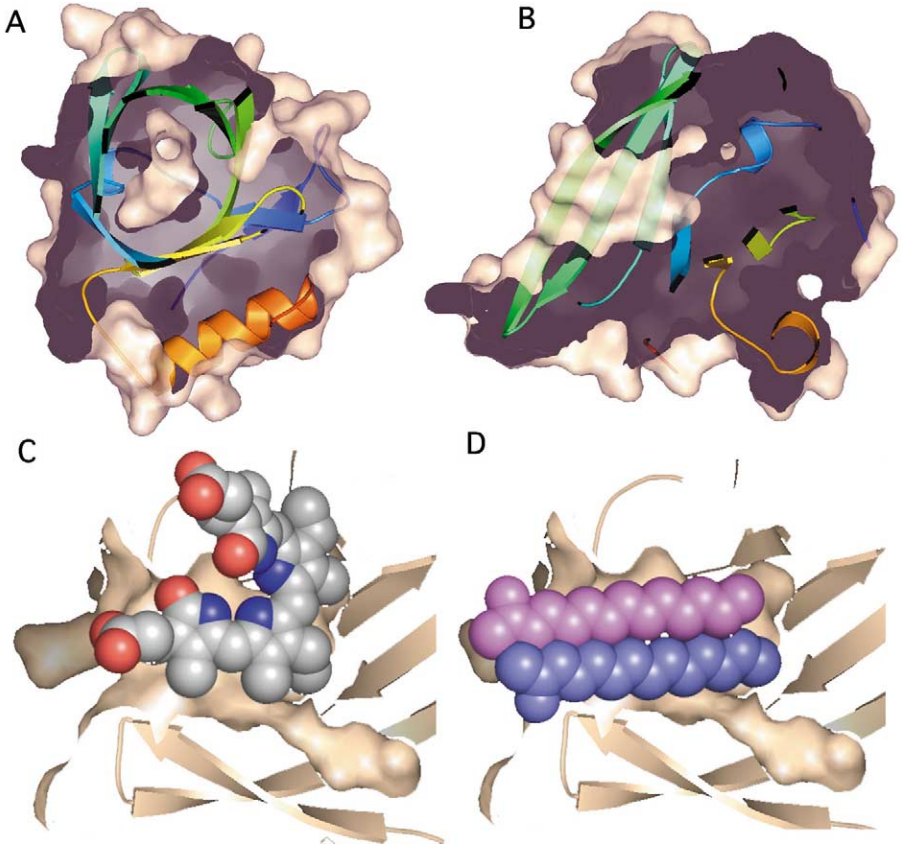

Fig. 5. Molecular surface of Blc cut at the level of the cavity. The β-strands are rainbow colored from the N-terminus (blue) to the C-terminus (red). A: View parallel to the β-barrel axis. B: View perpendicular to the β-barrel axis. C: A model of the bilin from the BBP structure [21] bound in the internal crevice. Note that only half of the bilin can fit in the cavity. D: A model with two fatty acids bound in the cavity. The alkyl chains fit the cavity walls well, indicating that ligands such as fatty acids or phospholipids might be transported or stored by Blc (pictures made with PyMOL [35]).

and thus faces the bulk periplasmic fluid. Adjacent to this main cavity, a smaller one forms the xenon binding site (Fig. 5A,B). The main cavity is less opened and narrower than the BBP calyx. After structural superimposition, bilin, the ligand of BBP [21], could not be fitted in the Blc cavity because of a large number of steric clashes with the Blc core (Fig. 5C).

In contrast with the cavity walls of retinol binding protein [5] or odorant binding protein [10], which are formed almost exclusively of hydrophobic residues, those of Blc are formed by hydrophobic, semi-polar and charged residues. The largest part is that of aromatic and aliphatic residues: Phe 53, Leu 57, Ala 62, Tyr 64, Val 74, Val 106, Phe 108, Tyr 116, Val 130, Trp 139 and Leu 141. In addition, two polar non-charged residues complete the wall, Asn 67 and Ser 89. These two residues protrude in the cavity. More surprisingly, two charged residues line the cavity wall, Glu 45 and Glu 54. Their carboxylic groups are in close contact (2.46 Å), which indicates that one of them is protonated. Glu 45 establishes an ionic bond with Arg 46, while Glu 54 is hydrogen-bonded to His 51. The two glutamic acids and the arginine are conserved in the sequence of BBP [21] and are in structurally equivalent positions compared to Blc.

Residues pertaining to the three classes, hydrophobic, semi-polar and polar, form the mouth of the cavity. Both positively and negatively charged residues are found: two negatively charged residues, Asp 82 and Glu 90, and four positively charged residues, arginines 52, 55 and 83, and Lys 92. The latter residues would be well suited for interacting with the phosphate moiety of a loaded phospholipid. Alternatively, they might also interact with phospholipids from the outer membrane.

As mentioned before, the binding cavity of Blc is long, flat and relatively narrow (Fig. 5A,B). It is not large enough to accommodate molecules such as bilins (Fig. 5C) or hemes, but it has a size sufficient to accommodate elongated chains or linear polycyclic structures such as anthracene derivatives. We have easily modeled two side-to-side fatty acids in the cavity (Fig. 5D). With an alkyl chain length of 14 carbons, both carboxylate moieties emerge in the solvent, near the Blc surface. Similarly, the two fatty acid chains of a phospholipid could fit the cavity, and its phosphate could interact with the positively charged residues around the cavity entrance, thus enhancing the binding affinity.

4. Conclusion and functional clues

Blc is a membrane-bound protein expressed in stress conditions which affect the membrane. The cavity of Blc is capable of binding up to two fatty acids. Other lipocalins such as cellular fatty acid binding proteins (FABP) [32] are lipocalins involved in fatty acid transport for membrane maintenance in, for example, heart and skeletal muscles [33]. A unique characteristic of Blc compared to FABPs, however, is its attachment to the outer membrane. This prevents Blc from traveling freely through the periplasmic space. However, Blc could still move laterally along the outer membrane and thus be capable of lipid transport. Such a two-dimensional diffusion may be more efficient and fast than free diffusion in the periplasm.

Acknowledgements: The ESRF is greatly acknowledged for beam time allocation. We thank the structural genomics team of the AFMB laboratory for technical assistance. This study was supported by the French Ministry of Industry (Grant ASG) and is a collaboration with the IGS laboratory and the Aventis company.

References

[1] Flower, D.R., North, A.C. and Sansom, C.E. (2000) Biochim. Biophys. Acta 1482, 9–24.
[2] Bishop, R.E., Penfold, S.S., Frost, L.S., Höltje, J.-V. and Weiner, J.H. (1995) J. Biol. Chem. 270, 23097–23103.
[3] Bishop, R.E. (2000) Biochim. Biophys. Acta 1482, 73–83.
[4] Åkerstrom, B., Flower, D.R. and Salier, J.-P. (2000) Biochim. Biophys. Acta 1482, 1–8.
[5] Newcomer, M.E., Jones, T.A., Aqvist, J., Sundelin, J., Eriksson, U., Rask, L. and Peterson, P.A. (1984) EMBO J. 3, 1451–1454.
[6] Cavaggioni, A., Findlay, J.B. and Tirindelli, R. (1990) Comp. Biochem. Physiol. B 96, 513–520.
[7] Singer, A.G. (1991) J. Steroid Biochem. Mol. Biol. 39, 627–632.
[8] Vincent, F., Lobel, D., Brown, K., Spinelli, S., Grote, P., Breer, H., Cambillau, C. and Tegoni, M. (2001) J. Mol. Biol. 305, 459–469.
[9] Pelosi, P., Baldaccini, N.E. and Pisanelli, A.M. (1982) Biochem. J. 201, 245–248.
[10] Ramoni, R., Vincent, F., Grolli, S., Conti, V., Malosse, C., Boyer, F.D., Nagnan-Le Meillour, P., Spinelli, S., Cambillau, C. and Tegoni, M. (2001) J. Biol. Chem. 276, 7150–7155.
[11] Paesen, G.C., Adams, P.L., Harlos, K., Nuttall, P.A. and Stuart, D.I. (1999) Mol. Cell 3, 661–671.
[12] Goetz, D.H., Holmes, M.A., Borregaard, N., Bluhm, M.E., Raymond, K.N. and Strong, R.K. (2002) Mol. Cell 10, 1033–1043.
[13] Flower, D.R. (2000) Biochim. Biophys. Acta 1482, 46–56.
[14] Gutiérrez, G., Ganfornina, M.D. and Sànchez, D. (2000) Biochim. Biophys. Acta 1482, 35–45.
[15] Ganfornina, M.D., Sànchez, D. and Bastiani, M.J. (1995) Development 121, 135–147.
[16] Davis, R.A. and Vance, J.E. (1996) in: Nex Comprehensive Biochemistry (Vance, D.E. and Vance, J.F., Eds.), Vol. 31, Biochemistry of Lipids, Lipoproteins and Membranes, pp. 473–493, Elsevier Science, Amsterdam.
[17] Rassart, E., Bedirian, A., Do Carmo, S., Guinard, O., Sirois, J., Terrisse, L. and Milne, R. (2000) Biochim. Biophys. Acta 1482, 185–198.
[18] David, G., Blondeau, K., Schiltz, M., Penel, S. and Lewit-Bentley, A. (2003) J. Biol. Chem. 278, 43728–43735.
[19] Prangé, T., Schiltz, M., Pernot, L., Colloc'h, N., Longhi, S., Bourguet, W. and Fourme, R. (1998) Proteins 30, 61–73.
[20] Olczak, A., Cianci, M., Hao, Q., Rizkallah, P.J., Raftery, J. and Helliwell, J.R. (2003) Acta Crystallogr. A 59, 327–334.
[21] Huber, R., Schneider, M., Mayr, I., Muller, R., Deutzmann, R., Suter, F., Zuber, H., Falk, H. and Kayser, H. (1987) J. Mol. Biol. 198, 499–513.
[22] Vincentelli, R., Bignon, C., Gruez, A., Sulzenbacher, G., Canaan, S., Tegoni, M., Campanacci, V. and Cambillau, C. (2003) Acc. Chem. Res. 36, 165–172.
[23] Sulzenbacher, G., Gruez, A., Roig-Zamboni, V., Spinelli, S., Valencia, C., Pagot, F., Vincentelli, R., Bignon, C., Salomoni, A., Grisel, S., Maurin, D., Huyghe, C., Johansson, K., Grassick, A., Roussel, A., Bourne, Y., Perrier, S., Miallau, L., Cantau, P., Blanc, E., Genevois, M., Grossi, A., Zenatti, A., Campanacci, V. and Cambillau, C. (2002) Acta Crystallogr. D 58, 2109–2115.
[24] Walhout, A.J., Temple, G.F., Brasch, M.A., Hartley, J.L., Lorson, M.A., van den Heuvel, S. and Vidal, M. (2000) Methods Enzymol. 328, 575–592.
[25] Lartigue, A., Gruez, A., Spinelli, S., Briand, L., Pernollet, J.-C., Tegoni, M. and Cambillau, C. (2003) Acta Crystallogr. D 59, 919–921.
[26] Collaborative Computational Project Number 4 (1994) Acta Crystallogr. D 50, 760–766.
[27] Schneider, T.R. and Sheldrick, G.M. Substructure solution with SHELXD, (2002) Acta Crystallogr. D 58, 1772–1779.
[28] de la Fortelle, E. and Bricogne, G. (1997) Methods Enzymol. 276, 472–494.
[29] Cowtan, K. (1994) Joint CCP4 ESF-EACBM Newsl. Protein Crystallogr. 31, 34–38.

[30] Perrakis, A., Morris, R. and Lamzin, V.S. (1999) Nat. Struct. Biol. 6, 458–463.
[31] Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Acta Crystallogr. D 53, 240–255.
[32] Thompson, J., Reese-Wagoner, A. and Banaszak, L. (1999) Biochim. Biophys. Acta 1441, 117–130.
[33] Luiken, J.J., Schaap, F.G., van Nieuwenhoven, F.A., van der Vusse, G.J., Bonen, A. and Glatz, J.F. (1999) Lipids 34 (Suppl.), S169–S175.
[34] Roussel, A. and Cambillau, C. (1991) in: Silicon Graphics Geometry Partners Directory, Vol. 81, pp. 77–78, Silicon Graphics, Mountain View, CA.
[35] DeLano, W.L., The PyMOL Molecular Graphics System, DeLano Scientific LLC, San Carlos, CA (http://www.pymol.org).