Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural genomics of *Caenorhabditis elegans*: structure of the BAG domain

J. Symersky,[a] Y. Zhang,[a] N. Schormann,[a] S. Li,[a] R. Bunzel,[a] P. Pruett,[a] C.-H. Luan[a] and M. Luo[a,b]*

[a]Southeast Collaboratory for Structural Genomics, Center for Biophysical Sciences and Engineering, University of Alabama at Birmingham, Birmingham, AL 35294, USA, and [b]Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA

Correspondence e-mail: mingluo@uab.edu

Received 21 June 2004
Accepted 17 July 2004

PDB Reference: BAG-1, 1t7s, r1t7ssf.

Binding of the BAG domain to the eukaryotic chaperone heat-shock protein (Hsp70) promotes ATP-dependent release of the protein substrate from Hsp70. Although the murine and human BAG domains have been shown to form an antiparallel three-helix bundle, the *Caenorhabditis elegans* BAG domain is formed by two antiparallel helices, while the third helix is extended away and stabilized by crystal-packing interactions. A small β-sheet between helices 2 and 3 interferes with formation of the intramolecular three-helix bundle. However, intermolecular three-helix bundles are observed throughout the crystal packing and suggest that stable functional dimers and tetramers can be formed in solution. The structure may represent a new folding type of the BAG domain.

1. Indroduction

BAG (Bcl2-associated athanogene) family proteins are relatively conserved in eukaryotes. BAG-1 was first identified by virtue of its interaction with the anti-apoptotic protein Bcl2 (Takayama *et al.*, 1995). The BAG family is characterized by the BAG domain, which mediates direct interaction with the Hsp70 ATPase domain (Takayama *et al.*, 1997; Zeiner *et al.*, 1997) and directly interacts with a number of other proteins, *e.g.* protein kinase Raf-1 (Wang *et al.*, 1996), growth-factor receptors (Bardelli *et al.*, 1996) and the retinoic acid receptor (Liu *et al.*, 1998). The minimal BAG domain of BAG-1 was identified by limited proteolysis of BAG-1M and shown to have the same affinity for Hsc70 (Sondermann *et al.*, 2001).

Binding and release of a protein substrate by the Hsp70 molecular chaperone is ATP-dependent. In addition, it also depends on Hsp40 cochaperone. ATP-bound Hsp70 has a low affinity for substrate, whereas the ADP-bound form has a high affinity. Hsp70 binds ATP with high affinity and slowly hydrolyzes it to ADP. Substrate binding to Hsp70–ATP stimulates ATP hydrolysis, resulting in a more stable complex of Hsp70–ADP with bound substrate. Release of substrate from the complex depends on the exchange of bound ADP for ATP. Binding of the BAG domain to the Hsp70 ATPase domain promotes ADP release from Hsp70 by disrupting the nucleotide-binding cleft. The BAG domain induces a relatively small conformational change in the ATPase that is incompatible with tight nucleotide binding. Interestingly, the BAG domain itself remains relatively distant from the bound nucleotide. Thus, the BAG proteins function as nucleotide-exchange factors for Hsp70 by stabilizing the nucleotide-free state of the ATPase. Because of the excess of ATP over ADP and BAG proteins in the eukaryotic cytosol, ATP will enter the nucleotide-binding pocket and displace bound BAG protein (Sondermann *et al.*, 2001). In prokaryotes, the same

nucleotide-exchange effect is mediated by GrpE. Structurally unrelated to the BAG domain, GrpE binds to the ATPase domain of DnaK, the bacterial homolog of Hsp70 (Harrison *et al.*, 1997). Unlike the BAG protein, which binds as a monomer, GrpE functions as a dimer. Remarkably, both BAG and GrpE elicit the same conformational change in the corresponding ATPase domain, but do not cross-react with the other ATPase domain.

The murine (Briknarova *et al.*, 2001) and human BAG domains have been shown to form an antiparallel three-helix bundle and a crystal structure of human BAG domain bound to Hsp70 ATPase domain has been described (Sondermann *et al.*, 2001). We have determined the crystal structure of the BAG domain (ceBAG) as part of the structural genomics initiative on *Caenorhabditis elegans* (Norwell & Machalek, 2000) and we report a novel BAG-domain fold compared with the previous structures.

2. Experimental

The full-length BAG-1 protein (GenBank accession No. AAD16125) was expressed as described by Finley *et al.* (2004). A spontaneous degradation of the protein led to identification of the BAG-1 domain. The corresponding cDNA for residues 74–210 of the BAG-1 protein was subcloned into pET28b vector (Novagen) with a histidine tag at the N-terminus and transformed into *Escherichia coli* strain B834(DE3). Expression was carried out at 310 K in M9 media supplied with selenomethionine. The protein was isolated on an Ni column, treated with thrombin to remove the histidine tag and loaded onto an S75 gel-filtration column (Amersham Biosciences). The pooled fractions were concentrated to 7.7 mg ml^{-1} in a final buffer consisting of 25 m*M* NaCl, 10 m*M* HEPES pH 7.5 and 1 m*M* DTT. Incorporation of selenomethionine was confirmed by electrospray ionization mass spectrometry.

A robotic screening for crystallization conditions was performed on HONEYBEE (Cartesian) with commercial kits (Hampton Research, Emerald BioStructures). Diffraction-grade crystals were grown by the vapor-diffusion method in manually set-up hanging drops. The reservoir solution contained 1.6 *M* ammonium sulfate, 0.1 *M* MES pH 5.9 and 5%(*v*/*v*) dioxane. Crystallization drops were made of 1.7 µl protein stock solution and 1 µl reservoir solution. For data collection at 100 K, the crystals were transferred to 3 *M* sodium malonate pH 6 and then flash-cooled in liquid nitrogen (Holyoak *et al.*, 2003). Based on the diffraction data, the crystals are orthorhombic, space group *I*222 or $I2_12_12_1$, with unit-cell parameters a = 75.518, b = 86.164, c = 125.49 Å. Assuming the presence of two protein molecules in the asymmetric unit, the Matthews coefficient is 3.32 Å^3 Da^{-1} and the solvent content is 62.6%.

The diffraction data were collected at the Advanced Photon Source (APS) beamline 22ID at an incident wavelength of 0.97924 Å, which corresponded to the peak of the fluorescence spectrum. The imaginary scattering component f'' was derived to be 5.65 e. The diffraction images were acquired in oscillation mode using a MAR225 CCD area detector. The crystal-to-detector distance was 200 mm and the oscillation angle was 1° around the ω axis. A total of 180 consecutive images were collected and processed as an anomalous data set to 2.8 Å resolution using *HKL*2000 (Otwinowski & Minor, 1997). The statistics are shown in Table 1.

The structure was solved by the single-wavelength anomalous diffraction (SAD) method (Dauter *et al.*, 2002) in *SOLVE*/*RESOLVE* (Terwilliger & Berendzen, 1999) as implemented in the Southeast Collaboratory for Structural Genomics (SECSG) Pipeline (https://genome.rcr.uga.edu/). The solution was only attainable in space group *I*222. Out of ten expected positions of Se atoms, eight were found by anomalous difference Patterson and difference Fourier methods. Based on the selenium positions, SAD phases were

Table 1
Data-collection and refinement statistics.

Values in parentheses refer to the last resolution shell.

Data collection	
Resolution range (Å)	50–2.8 (2.9–2.8)
R_{sym} (%)	6.8 (31.4)
Completeness (%)	99.9 (100.0)
Average $I/\sigma(I)$	12.3 (3.4)
Average redundancy	3.9 (3.9)
Refinement	
No. reflections	19761 (1971)
Resolution range (Å)	50–2.8 (2.98–2.8)
R factor (%)	22.4 (33.2)
R_{free} (%)	29.8 (40.1)
No. amino-acid residues	258
No. protein atoms	2108
No. water sites	50
Wilson B factor (Å^2)	64.7
Average B factor (Å^2)	57.7
R.m.s.d. bond lengths (Å)	0.007
R.m.s.d. bond angles (°)	1.3
Ramachandran plot regions	
Most favored (%)	91.9
Additional (%)	7.6
Disallowed (%)	0.4

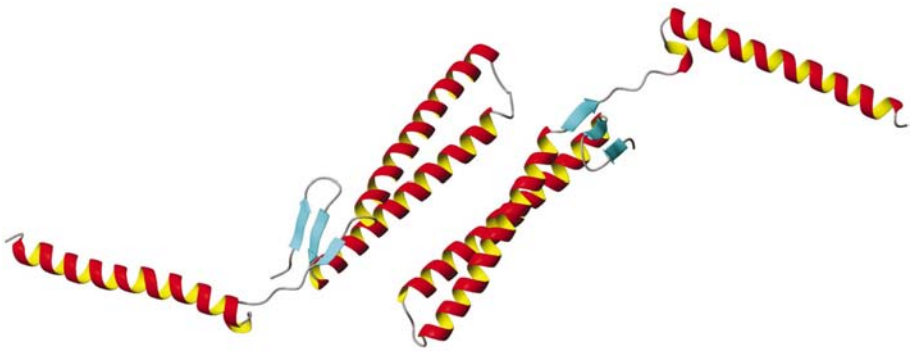

Figure 1
Ribbon representation of the two chains from the asymmetric unit of the crystal structure of the BAG domain from *C. elegans*. Rendered using *MOLMOL* (Koradi *et al.*, 1996).

calculated and further improved by *RESOLVE*. An electron-density map generated from these phases unambiguously showed two polypeptide chains compatible with the amino-acid sequence of the protein. Approximately 50% of the amino acids of each chain were built automatically in *RESOLVE*. The model was completed by manual building to the original SAD map using the programs *O* (Jones *et al.*, 1993) and *CHAIN* (Sack, 1988). Refinement of the model was performed in *CNS* v.1.1 (Brünger *et al.*, 1998). Although related by an approximate non-crystallographic twofold, both chains were refined independently. Additional manual building using $2F_o - F_c$ maps improved the positions of side chains and terminal residues and assisted in the localization of 50 water molecules. The crystallographic *R* factor converged at 0.224 and the R_{free} at 0.298 using all anomalous data with no cutoffs. The model statistics are shown in Table 1.

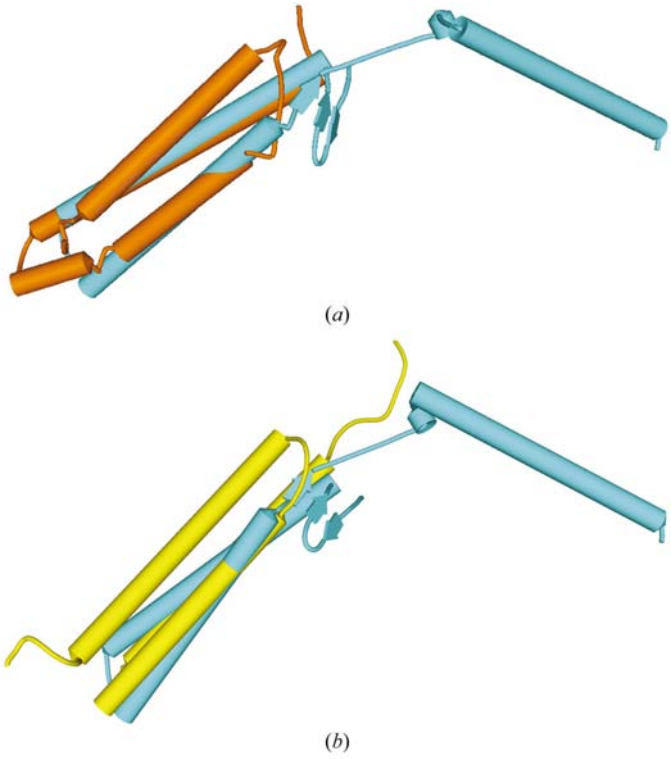

Figure 2
Superimpositions of the BAG domain from *C. elegans* chain *A* (cyan), with (*a*) the human BAG domain (gold) and (*b*) murine BAG domain (yellow) through the corresponding helices 1 and 2. The helices are represented by cylinders and the β-strands by arrows. The superimpositions were made using the program *TOPP* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The figures were prepared in *Viewer Lite* 5.0 (Accelrys).

3. Results and discussion

Both polypeptide chains were surprisingly well traceable in the original SAD map and refined well without breaks in the $2F_o - F_c$ map. Each chain includes residues 74–202 of the BAG-1 protein from *C. elegans* (GenBank accession No. AAD16125). Weaker electron density persisted in connecting loop regions *A*122–*A*125 and *B*120–*B*125 and residue Leu124*A* was refined on the edge of a disallowed region of the Ramachandran plot. The electron density for C-terminal residues 203–210 was completely missing in both chains. Because of the lower resolution and/or possible disorder, a number of side chains, especially Lys and Arg, were poorly defined and a relatively small number of ordered water molecules could be included in the model. This may also contribute to the higher *R* factors. The root-mean-square deviations of the model from target values (Engh & Huber, 1991), as well as the Ramachandran plot, met general standards (Table 1). It is noteworthy that to our knowledge ceBAG is the first crystal structure of the BAG domain as such. The human BAG crystal structure is in a complex with Hsp70 ATPase domain, whereas the murine BAG is a solution NMR structure.

The ceBAG molecule is largely helical as expected from previous studies on human and murine analogs (Fig. 1). At the N-terminus the chain folds into a short β-hairpin and proceeds to a pair of long antiparallel helices 1 and 2 which bring it back to the β-hairpin. At this switch point a short β-strand is formed and lined up with the β-hairpin into a small three-stranded β-sheet. The chain proceeds further from the β-sheet in an extended loop and finally folds into a long isolated helix 3, which gives the molecule an overall L-shape. Compared with human and murine BAG, which fold into an anti-parallel three-helix bundle with no β-strand (Fig. 2), ceBAG apparently possesses distinct features in terms of secondary structure and folding.

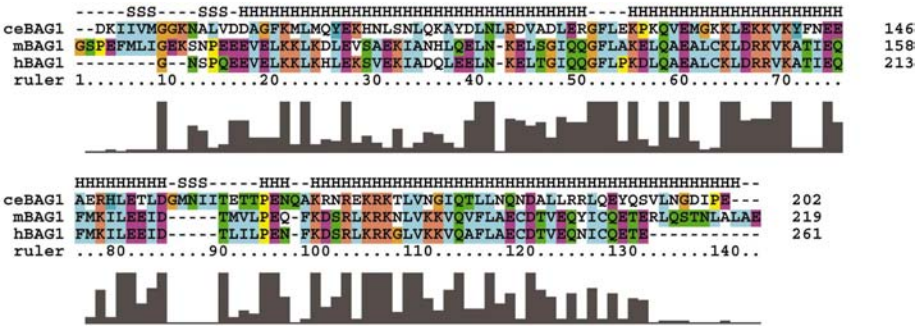

Figure 3
Sequence alignment of the BAG domains from *C. elegans* (ceBAG), murine (mBAG) and human (hBAG) with overlaid secondary structure of the *C. elegans* BAG domain (S, β-strand; H, helix). The color coding reflects different chemical types of amino-acid residues. The higher the bar under the alignment, the higher is conservation at the given amino-acid position. This figure was prepared using *ClustalX* (Jeanmougin *et al.*, 1998).

Fig. 3 shows a sequence alignment of the human, murine and *C. elegans* BAG domains. A closer inspection shows that the ceBAG sequence is less homologous to human and murine sequences (22 and 25% identity) than they are to each other (70% identity). Moreover, ceBAG has a unique insertion of five amino acids, Gly, Met, Asn, Ile, Ile, in the second half of the

domain. This insertion corresponds to the third β-strand which lines up with the β-hairpin forming the β-sheet.

Both ceBAG chains superimpose well on themselves using their C^{α} atoms with a root-mean-square deviation (r.m.s.d.) of 0.66 Å. In addition, ceBAG superimposes reasonably well with human and murine BAG through their corresponding helices 1 and 2 (Fig. 2). The r.m.s.d.s are less than 2 Å for the matching subsets of C^{α} atoms. Remarkably, when the human BAG domain is superimposed with the ceBAG crystal structure through helices 1 and 2, the third helix of a ceBAG symmetry-related molecule matches the third helix of the human BAG (Fig. 4). Thus, instead of forming an intramolecular three-helix bundle, ceBAG forms intermolecular helical bundles throughout the crystal packing and helix 3 maintains principally the same relationship with helices 1 and 2 as in human and murine BAG domains. The intermolecular interactions between helix 3 and helices 1 and 2 are apparently more productive, in terms of buried molecular surface, than interactions between ceBAG molecules in the asymmetric unit. The dimeric assembly formed in the asymmetric unit is therefore improbable in solution. However, both monomeric and dimeric forms were implicated by the gel-filtration profiles during protein purification. It would be very likely that a dimer is formed in solution in a similar way as in the crystal packing (Fig. 4). This dimer consists of two three-helix bundles and a small β-sheet core, as the β-sheets from each ceBAG molecule pack against each other. In the crystal packing, two such dimers form a tetramer which may also be found in solution. Here, the dimers interact mainly through the β-strands, which are lined up into two six-stranded β-sheets that pack aginst each other. Four three-helix bundles extend from the β-sheet core in a tetrahedral fashion (see also http://qps.ebi.ac.uk/).

In a modeling session, we attempted to refold the ceBAG molecule into a three-helical bundle by manipulating the φ, ψ main-chain angles in the connecting region between helices 2 and 3. The objective was to align helix 3 with the corresponding helix of the human BAG molecule superimposed with ceBAG through helices 1 and 2. The helices could not be aligned unless the third β-strand was unfolded into a coiled structure.

We have shown that the considerable differences in the ceBAG sequence compared with those of human and murine result in a new secondary structure and folding, as observed in the crystal structure. While the basic folding motif consisting of the two antiparallel helices 1 and 2 is essentially conserved, the unique insertion of five amino acids downstream in the sequence, along with the β-hairpin at the N-terminus, result in a β-sheet which prevents helix 3 from proper alignment with helices 1 and 2 to form an intramolecular three-helix bundle as observed in human and murine BAG domains. The crystal structure of ceBAG thus presents a new folding type of the BAG domain. In solution, helix 3 may have more conformational freedom. However, the β-sheet is expected to be stable and keep the domain in a relatively open state with an overall L- or V-shape. The domains are likely to associate into dimers or even tetramers. Ultimately, only a solution NMR structure of ceBAG can verify whether the new fold is inherent to the ceBAG domain or whether it was formed casually by the crystal packing.

Sondermann *et al.* (2001) have shown that in the crystal structure the human BAG binds to Hsp70 ATPase primarily through bundled helices 2 and 3. This would not be possible with the helices split as is the case in ceBAG. On the other hand, GrpE binds differently to DnaK ATPase, which is structurally homologous to Hsp70 ATPase and elicits the same conformational change (Harrison *et al.*, 1997). It is possible that ceBAG binds to ATPase in a similar mode as GrpE, *i.e.* as a more extended helical molecule stabilized in a dimer. If ceBAG forms the same dimer in solution as observed in the crystal packing (Fig. 4), then it can interact with Hsp70 ATPase similarly to human BAG through helices 2 and 3, each from a different monomer. Mechanistically, it is even conceivable that the tetramer may interact with the ATPase domain through one of the three-helix bundles.

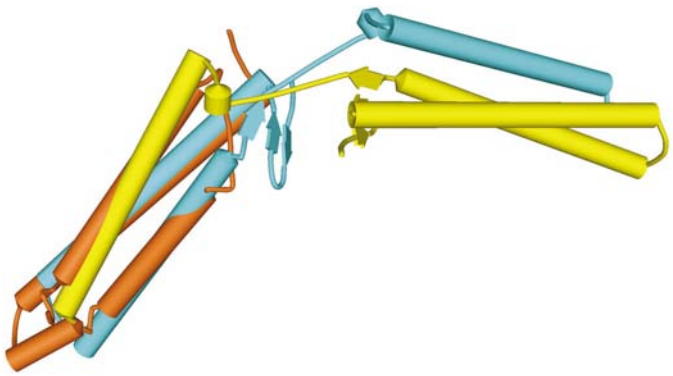

Figure 4
A superimposition of the human BAG domain (gold) with chain *A* of the *C. elegans* BAG domain (cyan) through the corresponding helices 1 and 2. A symmetry-related molecule of *C. elegans* BAG domain is shown in yellow. Helix 3 of the symmetry-related molecule matches helix 3 of the superimposed human BAG domain. Rendered in *Viewer Lite* 5.0 (Accelrys).

Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at http://www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. This work was supported by NIGMS grant IP50-M62407.

References

Bardelli, A., Longati, P., Albero, D., Goruppi, S., Schneider, C., Ponzetto, C. & Comoglio, P. M. (1996). *EMBO J.* **15**, 6205–6212.
Briknarova, K., Takayama, S., Brive, L., Havert, M. L., Knee, D. A., Velasco, J., Homma, S., Cabezas, E., Stuart, J., Hoyt, D. W., Satterthwait, A. C., Llinas, M., Reed, J. C. & Ely, K. R. (2001). *Nature Struct. Biol.* **8**, 349–352.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, G. M., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M.,

Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dauter, Z., Dauter, M. & Dodson, E. (2002). *Acta Cryst.* D**58**, 494–506.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Finley, J. B., Qiu, S.-H., Luan, C.-H. & Luo, M. (2004). *Protein Expr. Purif.* **34**, 49–55.
Harrison, C. J., Hayer-Hartl, M., Liberto, M. D., Hartl, F. & Kuriyan, J. (1997). *Science*, **276**, 431–435.
Holyoak, T., Fenn, T. D., Wilson, M. A., Moulin, A. G., Ringe, D. & Petsko, G. A. (2003). *Acta Cryst.* D**59**, 2356–2358.
Jeanmougin, F. Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. (1998). *Trends Biochem. Sci.* **23**, 403–405.
Jones, T. A., Bergdoll, M. & Kjeldgaard, M. (1993). *Crystallographic Computing and Modeling Methods in Molecular Design.* New York: Springer.
Koradi, R., Billeter, M. & Wuthrich, K. (1996). *J. Mol. Graph.* **14**, 51–55.
Liu, R., Takayama, S., Zheng, Y., Froesch, B., Chen, G.-Q., Zhang, X., Reed, C. & Zhang, X.-K. (1998). *J. Biol. Chem.* **273**, 16985–16992.
Norwell, J. C. & Machalek, A. Z. (2000). *Nature Struct. Biol.* **7** (*Suppl.*), 931.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sack, J. S. (1988). *J. Mol. Graph.* **6**, 224–225.
Sondermann, H., Scheufler, C., Schneider, C., Hohfeld, J., Hartl, F.-U. & Moarefi, I. (2001). *Science*, **291**, 1553–1557.
Takayama, S., Bimston, D. N., Matsuzawa, S.-I., Freeman, B. C., Aime-Sempe, C., Xie, Z., Morimoto, R. I. & Reed, J. C. (1997). *EMBO J.* **16**, 4887–4896.
Takayama, S., Sato, T., Krajewski, S., Kochel, K., Irie, S., Millan, J. A. & Reed, J. C. (1995). *Cell*, **80**, 279–284.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Wang, H.-G., Takayama, S., Rapp, U. R. & Reed, J. C. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 7063–7068.
Zeiner, M., Gebauer, M. & Gehring, U. (1997). *EMBO J.* **16**, 5483–5490.