Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray analysis of the *Entamoeba histolytica* α-actinin-2 rod domain

Barbara Addario,* Shenghua Huang, Uwe H. Sauer and Lars Backman

Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden

Correspondence e-mail: barbara.addario@chem.umu.se

Received 6 May 2011
Accepted 1 July 2011

α-Actinins form antiparallel homodimers that are able to cross-link actin filaments. The protein contains three domains: an N-terminal actin-binding domain followed by a central rod domain and a calmodulin-like EF-hand domain at the C-terminus. Here, crystallization of the rod domain of *Entamoeba histolytica* α-actinin-2 is reported; it crystallized in space group $P2_12_12_1$, with unit-cell parameters $a = 47.8$, $b = 79.1$, $c = 141.8$ Å. A Matthews coefficient V_M of 2.6 Å^3 Da^{-1} suggests that there are two molecules and 52.5% solvent content in the asymmetric unit. A complete native data set extending to a *d*-spacing of 2.8 Å was collected on beamline I911-2 at MAX-lab, Sweden.

1. Introduction

α-Actinins are ubiquitous actin-binding proteins that are present in all eukaryotes with the exception of plants and baker's yeast (*Saccharomyces cerevisiae*). This group of proteins are characterized by an N-terminal actin-binding domain, a C-terminal calmodulin-like EF-hand domain and an intervening rod domain (Blanchard *et al.*, 1989; Otey & Carpen, 2004). The evolutionarily highly conserved actin-binding domain comprises two calponin-homology domains (Bañuelos *et al.*, 1998; Virel & Backman, 2004). The EF-hand domain and in particular the rod domain show greater sequence variation among different species. In vertebrates, four isoforms have evolved: two calcium-insensitive muscle isoforms (α-actinin-2 and α-actinin-3) and two nonmuscle isoforms (α-actinin-1 and α-actinin-4) that show calcium-dependent actin binding. In invertebrates, only a single calcium-sensitive isoform is found. This might indicate that only a single calcium-dependent α-actinin existed before the invertebrate–vertebrate bifurcation.

α-Actinins are able to cross-link actin filaments into extended networks or bundles owing to their ability to form antiparallel dimers *via* the rod domain. It has also been suggested that the rod functions as a scaffold for several other proteins (Djinovic Carugo *et al.*, 2002; Otey & Carpen, 2004). In addition, α-actinins are important and crucial organizers of focal adhesion contacts and of the muscle Z-disc.

We have previously cloned, expressed and biochemically characterized the α-actinins (α-actinin-1 and α-actinin-2) of *Entamoeba histolytica* (Virel & Backman, 2006; Virel *et al.*, 2007), the aetiological agent of human amoebiasis (Stanley, 2003; Santi-Rocca *et al.*, 2009). Both isoforms are able to form dimers as well as to bind and cross-link actin filaments in a calcium-sensitive manner. However, the rod domains of the *E. histolytica* α-actinins are shorter compared with those of vertebrate isoforms. Analysis of multiple sequence alignments suggests that these isoforms contain one or two spectrin repeats, respectively, whereas in animals the rod domain comprises four spectrin repeats. Further, on analysing the domain content with the SUPERFAMILY (Gough *et al.*, 2001), Pfam (Finn *et al.*, 2010), SMART (Letunic *et al.*, 2009) and FISH (Tångrot *et al.*, 2006) servers, only SUPERFAMILY and FISH identified two spectrin repeats in α-actinin-2, albeit with a low probability. As noted previously (Virel & Backman, 2007), neither of these servers was able to detect a spectrin repeat in the rod of α-actinin-1 but rather suggested a coiled-coil structure.

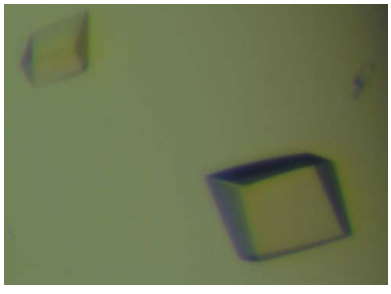

Table 1
Parameters and data-collection statistics for crystals of *E. histolytica* α-actinin-2.

Values in parentheses are for the highest resolution shell.

No. of crystals	1
Beamline	I911-2, MAX-lab, Sweden
Wavelength (Å)	1.03786
Detector	MAR CCD
Crystal-to-detector distance (mm)	169.8
Rotation range per image $\Delta\varphi$ (°)	0.5
Total rotation range (°)	180
Exposure time per image (s)	20
Resolution range (Å)	32.34–2.80 (2.95–2.80)
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 47.8, b = 79.1, c = 141.8
Mosaicity (°)	0.27
Total No. of measured intensities	98668 (14155)
Unique reflections	13831 (1949)
Multiplicity	7.1 (7.3)
Mean $I/\sigma(I)$	19.5 (6.3)
Completeness (%)	99.4 (97.9)
$R_{\rm merge}$† (%)	9.2 (34.8)
$R_{\rm meas}$ or $R_{\rm r.i.m.}$‡ (%)	10.0 (37.5)
Overall B factor from Wilson plot (Å^2)	54.8

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the weighted average intensity of all observations i of reflection hkl. ‡ $R_{\rm meas} = R_{\rm r.i.m.} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the weighted average intensity of all observations i of reflection hkl. N is the multiplicity.

E. histolytica α-actinins have been implicated in the infectious mechanism of the parasite. It has been suggested that one of the α-actinins or possibly both interact with the intracellular C-terminus of a Gal/GalNAc lectin that is required for adhesion (Blazquez *et al.*, 2007; Vargas *et al.*, 1996; Seigneur *et al.*, 2005). In addition, it is plausible that the rod domain may serve as an interaction platform for other proteins involved in the process.

Structural information about these two α-actinins would improve our understanding of the role they play in the life cycle of *E. histolytica*. Crystallization attempts using the full-length proteins have not yet yielded diffracting crystals. We therefore cloned and expressed the domains separately. In this study, we present the crystallization and preliminary diffraction data analysis of the rod domain of *E. histolytica* α-actinin-2.

2. Materials and methods

2.1. Cloning, expression and purification

The central rod-domain region containing the putative spectrin-like repeats (nucleotides 754–1431, corresponding to amino-acid residues Glu252–Ala477) was amplified by PCR using the full-length *E. histolytica* (strain HM-1:IMSS) α-actinin-2 gene as a template (accession No. XM_648191). For directional ligation into the expression vector, the forward primer 5′-TTT **GGA TCC** GAA GGT ATG GTT CAT GAT TAT G-3′ and reverse primer 5′-TTT **CTC GAG** CTA TTA TGC TTC ATT AAT TTG GG-3′, containing a *Bam*HI and a *Xho*I restriction site, respectively (shown in bold), were used. The amplified fragment was digested and cloned into the *Bam*HI and *Xho*I sites of a modified pET-19b vector, producing the plasmid pTEV-EH ROD. The vector encodes an N-terminal 10×His tag followed by a TEV-protease cleavage site upstream of the *Bam*HI site. The sequence of the cloned construct was confirmed by DNA sequencing (Eurofins MWG GmbH, Germany).

Competent *Escherichia coli* BL21 (DE3) cells were transformed by heat-shock with pTEV-EH ROD and grown for 16 h at 310 K on agar plates containing Luria–Bertani (LB) medium supplemented with 100 µg ml^{-1} carbenicillin. 10 ml LB medium containing 100 µg ml^{-1} carbenicillin was inoculated with a single colony and incubated in a shaker incubator (200 rev min^{-1}) at 310 K for 16 h. 2 ml of the liquid culture was used to inoculate 2000 ml LB medium containing 100 µg ml^{-1} carbenicillin until an optical density of 0.6–0.8 at 600 nm was reached. At this point the temperature was decreased to 296 K and protein expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside to a final concentration of 0.5 m*M*. After induction for 16 h at 296 K, the cells were harvested by centrifugation, resuspended in 40 ml NaPB (150 m*M* NaCl in 25 m*M* sodium phosphate buffer pH 7.6) and stored at 253 K until further processing.

For purification, frozen cells (∼40 ml) were first thawed on ice and Triton X-100 was added to 1% final concentration. Cells were lysed by sonication on ice and centrifuged at 37 000*g* for 15 min in a Beckman–Coulter JA-20 rotor. The clarified supernatant was loaded onto a 5 ml HiTrap Chelating HP column (GE Healthcare) charged with Ni^{2+} ions and extensively washed with NaPB containing 10 m*M* imidazole to elute unbound proteins. Bound proteins were eluted in a gradient ranging from 10 to 510 m*M* imidazole in NaPB. The His-tagged rod domain eluted at 163 m*M* imidazole. The eluate (15 ml) was incubated at 277 K for 16 h with 6×His-tagged tobacco etch virus (TEV) protease (produced from a plasmid kindly provided by Dr David S. Waugh) and simultaneously dialyzed against 1000 ml 50 m*M* Tris–HCl pH 7.6 for 1 h and then against 2000 ml of the same buffer overnight to remove imidazole. The released 10×His tag as well as the 6×His-tagged TEV protease was removed by passing the dialysate over a second 5 ml Ni^{2+}-charged HiTrap Chelating HP column (GE Healthcare, Sweden). An additional ion-exchange chromatography step on a HiTrap Q column (GE Healthcare, Sweden) was used to remove most of the remaining impurities. The recombinant rod domain was eluted in a salt gradient from 0 to 1 *M* NaCl in 50 m*M* Tris–HCl pH 7.6. Finally, 16 ml of the protein was dialysed against 1000 ml 50 m*M* Tris–HCl pH 7.6 for 1 h and then against 2000 ml of the same buffer overnight at 277 K. After dialysis, the protein solution was concentrated using an Ultracel 3K centrifugal filter with a cutoff of 3 kDa (Millipore, Ireland). The protein concentration was determined from the absorbance at 280 nm using a molar extinction coefficient ε of 13 535 M^{-1} cm^{-1}, which was calculated from the amino-acid sequence (using *ProtParam* from the ExPASy proteomics server). The purity of the protein was determined under denaturing conditions using SDS–PAGE (Laemmli, 1970); we estimated the purity to be greater than 95–97% from the stained gel.

2.2. Crystallization and diffraction data collection

Initial screening for crystallization conditions was performed in 96-well sitting-drop plates (MRC–Wilden, UK) using Crystal Screen and Crystal Screen 2 (Hampton Research, USA) and a Mosquito liquid-handling robot (TTP LabTech, UK). For this, the 48 first reagents of Crystal Screen as well as all reagents of Crystal Screen 2 were transferred to a deep-well plate. Equal volumes (100 nl) of protein in 50 m*M* Tris–HCl pH 7.6 and precipitant solution were then pipetted into the sitting-drop wells and sealed with Crystal Clear tape (Hampton Research). Initially, small crystals were obtained from condition No. 17 of Crystal Screen (0.2 *M* Li_2SO_4, 0.1 *M* Tris–HCl pH 8.5, 30% PEG 4000).

These conditions were manually optimized by screening about 200 drops containing different buffers with varying pH and molecular-weight PEGs, as well as different alcohols such as 2-propanol, 1,6-hexanediol, β-mercaptoethanol, 2-methyl-2,4-pentanediol and dimethyl sulfoxide. Optimization was performed using 24-well Linbro plates and the hanging-drop vapour-diffusion method. During the

optimization process, three-part droplets were produced by pipetting together 2 µl protein solution (in 50 m*M* Tris–HCl pH 7.6), 2 µl precipitant solution containing different molecular-weight PEGs and 0.5–2 µl alcohol and were equilibrated over 500 µl reservoir solution (precipitant without alcohol).

2.3. Diffraction data collection and processing

The crystals were harvested into CryoLoops (Hampton Research, USA), immediately vitrified in liquid nitrogen without any additional cryoprotecting agents and stored under liquid nitrogen until further use. X-ray diffraction data were collected at 100 K on beamline I911-2 at the MAX-lab II synchrotron, Lund, Sweden. Rotating the crystals by 0.5° per exposure, a data set consisting of 360 diffraction images was collected from a single crystal. Intensity data were indexed, integrated and processed with *XDS* (Kabsch, 2010). Unmerged intensities from *XDS* were read into *POINTLESS* (Evans, 2011), which confirmed the Laue group and transformed the intensity data into *CCP*4 format. *SCALA* (Evans, 2011) was used to scale and merge the data and *TRUNCATE* (French & Wilson, 1978) was used to transform the intensities to structure factors. The data-collection and processing statistics are listed in Table 1. The Matthews coefficient (Matthews, 1968) was calculated using *MATTHEWS_COEF* (Kantardjieff & Rupp, 2003) from the *CCP*4 suite (Winn *et al.*, 2011).

3. Results and discussion

The correspondence of the cloned and the original genomic nucleotide sequences of the *E. histolytica* α-actinin-2 rod domain was confirmed by DNA sequencing. Owing to the choice of cloning vector, the nucleotide sequence of the translated protein harboured codons for a ten-His tag and a TEV protease recognition site upstream of the first rod-domain codon. Cleavage with TEV protease produced a 228-residue protein with an additional Gly-Ser in front of Glu252, the first native residue of the rod domain, resulting in a molecular weight of 25 817 Da.

The purification of the rod domain of *E. histolytica* α-actinin-2 yielded about 10 mg pure protein per litre of bacterial cell culture. As judged from the elution position after gel filtration, it was apparent that the rod domain eluted as a dimer (data not shown). We confirmed that the α-helical content of the purified protein, as determined by circular-dichroism spectroscopy, agreed well with the expected content, indicating that the rod domain was correctly folded (Addario & Backman, 2010).

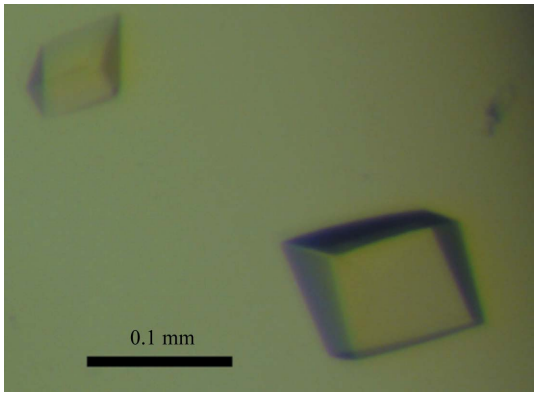

Figure 1
Typical crystal of the *E. histolytica* α-actinin-2 rod domain. The crystal was ∼0.1 × 0.1 × 0.03 mm in size.

For crystallization, pure protein was concentrated to 1.6 mg ml^{-1} by ultrafiltration; the protein precipitated at higher concentrations.

The best crystals for diffraction data collection were obtained from conditions consisting of a 1:1:1 mixture of 2 µl protein solution, 2 µl precipitation buffer (0.1 *M* Tris–HCl pH 7.5, 50 m*M* Li_2SO_4, 20% PEG 1450) and 2 µl 5% 2-propanol equilibrated over 500 µl well solution (0.1 *M* Tris–HCl pH 7.5, 50 m*M* Li_2SO_4, 35% PEG 1450). The crystals grew at 291 K in less than a week to dimensions of about 0.1 × 0.1 × 0.03 mm (Fig. 1).

Indexing the diffraction data revealed a unit cell with parameters $a = 47.8$, $b = 79.1$, $c = 141.8$ Å in space group $P2_12_12_1$. Further analysis of the asymmetric unit suggested two molecules and 52.5% solvent content, with a Matthews coefficient V_M of 2.6 Å^3 Da^{-1}.

Attempts to use molecular replacement for structure solution using *CNS* (Brünger *et al.*, 1998; Brunger, 2007), *MOLREP* (from the *CCP*4 suite; Winn *et al.*, 2011), *AMoRe* (Navaza, 1994; Trapani & Navaza, 2008) and *EPMR* (Kissinger *et al.*, 1999) were unsuccessful. As structural templates, we used individual repeats of PDB entries 1hci (Ylänne *et al.*, 2001), 1quu (Djinovic-Carugo *et al.*, 1999) and 1u5p (Kusunoki *et al.*, 2004) as polyalanine or polyserine models with or without removed loop regions. The lack of success was most likely due to the low search mass of the structural templates and the low sequence identity of the two spectrin repeats of the rod domain to proteins of known crystal structure. The first spectrin repeat, amino acids 252–364, shows 30% sequence identity to amino acids 763–878 of *Dictyostelium* α-actinin (PDB entry 1g8x; Kliche *et al.*, 2001) and the second rod repeat, amino acids 375–477, shows 23% sequence identity to residues 368–472 of the rod domain of human α-actinin-2 (PDB entry 1hci; Ylänne *et al.*, 2001). We have therefore initiated the crystallization of selenomethionine-labelled protein in the hope of determining the structure using anomalous diffraction methods.

This work was supported by grants from Carl Tryggers Stiftelse to UHS and LB.

References

Addario, B. & Backman, L. (2010). *Cell. Mol. Biol. Lett.* **15**, 665–678.
Bañuelos, S., Saraste, M. & Djinović Carugo, K. (1998). *Structure*, **6**, 1419–1431.
Blanchard, A., Ohanian, V. & Critchley, D. (1989). *J. Muscle Res. Cell Motil.* **10**, 280–289.
Blazquez, S., Rigothier, M.-C., Huerre, M. & Guillén, N. (2007). *Int. J. Parasitol.* **37**, 425–433.
Brunger, A. T. (2007). *Nature Protoc.* **2**, 2728–2733.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., 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.
Djinovic Carugo, K., Gautel, M., Ylänne, J. & Young, P. (2002). *FEBS Lett.* **513**, 119–123.
Djinovic-Carugo, K., Young, P., Gautel, M. & Saraste, M. (1999). *Cell*, **98**, 537–546.
Evans, P. R. (2011). *Acta Cryst.* D**67**, 282–292.
Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R. & Bateman, A. (2010). *Nucleic Acids Res.* **38**, D211–D222.
French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.
Gough, J., Karplus, K., Hughey, R. & Chothia, C. (2001). *J. Mol. Biol.* **313**, 903–919.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
Kliche, W., Fujita-Becker, S., Kollmar, M., Manstein, D. J. & Kull, F. J. (2001). *EMBO J.* **20**, 40–46.

Kusunoki, H., Minasov, G., Macdonald, R. I. & Mondragón, A. (2004). *J. Mol. Biol.* **344**, 495–511.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Letunic, I., Doerks, T. & Bork, P. (2009). *Nucleic Acids Res.* **37**, D229–D232.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otey, C. A. & Carpen, O. (2004). *Cell Motil. Cytoskeleton*, **58**, 104–111.
Santi-Rocca, J., Rigothier, M.-C. & Guillén, N. (2009). *Clin. Microbiol. Rev.* **22**, 65–75.
Seigneur, M., Mounier, J., Prevost, M.-C. & Guillén, N. (2005). *Cell. Microbiol.* **7**, 569–579.
Stanley, S. L. (2003). *Lancet*, **361**, 1025–1034.
Tångrot, J., Wang, L., Kågström, B. & Sauer, U. H. (2006). *Nucleic Acids Res.* **34**, W10–W14.
Trapani, S. & Navaza, J. (2008). *Acta Cryst.* D**64**, 11–16.
Vargas, M., Sansonetti, P. & Guillén, N. (1996). *Mol. Microbiol.* **22**, 849–857.
Virel, A., Addario, B. & Backman, L. (2007). *Mol. Biochem. Parasitol.* **154**, 82–89.
Virel, A. & Backman, L. (2004). *Mol. Biol. Evol.* **21**, 1024–1031.
Virel, A. & Backman, L. (2006). *Mol. Biochem. Parasitol.* **145**, 11–17.
Virel, A. & Backman, L. (2007). *Mol. Biol. Evol.* **24**, 2254–2265.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Ylänne, J., Scheffzek, K., Young, P. & Saraste, M. (2001). *Structure*, **9**, 597–604.