Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, purification and crystallization of full-length human annexin 2

Jacqueline T. Tran, Anja Rosengarth and Hartmut Luecke*

Department of Molecular Biology and Biochemistry, UCI Program in Macromolecular Structure, 3205 McGaugh Hall, University of California, Irvine, CA 92797-3900, USA

Correspondence e-mail: hudel@uci.edu

Received 27 March 2002
Accepted 23 July 2002

Annexin 2, a Ca^{2+}/phospholipid-binding protein, is involved in many biological processes, including membrane aggregation and the modulation of fibrinolytic activity. Here, the expression and purification of recombinant full-length human annexin 2 is reported, as well as crystals obtained by sitting-drop and hanging-drop vapor diffusion at 277 K. A condition consisting of 18% PEG 8000, 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* calcium acetate yielded long needles that diffracted to 3.20 Å. Another condition, consisting of 2.5 *M* NaCl, 0.1 *M* acetate pH 4.5, 0.2 *M* Li_2SO_4, gave crystals with unit-cell parameters $a = 48.36$, $b = 62.86$, $c = 119.11$ Å that diffracted to 1.52 Å. Both crystals belong to the orthorhombic $P2_12_12_1$ space group. The high-resolution 1.52 Å data set was collected at ALS beamline 5.0.2 and is 93.0% complete, with an R_{sym} of 4.5%. The structure of full-length annexin 2 will provide insight into how its N-terminal domain contributes to its functional role in a variety of biological processes.

1. Introduction

Annexins are a multigene family of proteins found in the majority of eukaryotic cells. These proteins possess a conserved core domain, typically comprised of four subdomains, which is responsible for calcium-dependent phospholipid binding (Hawkins *et al.*, 2000). In contrast, the N-terminal domain, which is highly variable in sequence and length, is believed to determine individual annexin function (Pepinsky *et al.*, 1988). Despite a wealth of experimental evidence that links annexins to a plethora of biological processes, definitive physiological functions have yet to be established.

Annexin 2, which exists as a monomer (p36) or a heterotetramer [$(p36)_2(p11)_2$] (Waisman, 1995), is known to aggregate and cause fusion of biological membranes and is involved in regulated membrane-trafficking events such as exocytosis, endocytosis and cell–cell adhesion (Drust & Creutz, 1988). More recently, annexin 2 has been found on the surface of endothelial cells as a component of the fibrinolysis machinery, where it is thought to co-localize tissue plasminogen activator (t-PA) and plasminogen, stimulating the t-PA-mediated proteolysis of plasminogen to plasmin (Kang *et al.*, 1999). The existing experimental evidence suggests two conflicting mechanisms in which either (i) the heterotetramer is the key physiological receptor for plasminogen on the extracellular surface of endothelial cells (Fitzpatrick *et al.*, 2000) or (ii) the annexin 2 monomer binds *via* its N-terminus to the C-terminal tail of t-PA, thereby activating plasminogen cleavage (Hajjar & Krishnan, 1999). In order to learn more about the role of the N-terminal domain of annexin 2 in plasminolysis and membrane aggregation, crystal structures of the full-length annexin 2 in the presence and absence of Ca^{2+} are essential.

The structure of a proteolytically truncated form of annexin 2 (lacking the first 30 residues) has been published previously (Burger *et al.*, 1996), but contained no structural information on the N-terminal domain or its interaction with the core domain. A crystal structure of the annexin 2 N-terminal peptide–p11 complex established that dimerization of p11 is essential for the heterotetrameric form of annexin 2 (Rety *et al.*, 1999). However, the structure did not reveal possible interactions of the annexin 2 N-terminal domain with its core domain.

Using a recently developed purification protocol that prevents proteolysis of the protease-sensitive N-terminal domain, we reported the structure of full-length annexin 1 in the absence of calcium (Rosengarth *et al.*, 2001). Surprisingly, the structure showed that the amphipathic N-terminal domain inserts into the third core-domain repeat, where it causes helix *D* of this repeat to unwind into a flap that covers the buried N-terminal domain. The full-length annexin 1 structure led us to propose a new model for annexin-mediated vesicle aggregation that is initiated by calcium-dependent binding of the core to the bilayer, followed by secondary binding of the N-

terminal domain to adjacent bilayers (Rosengarth *et al.*, 2001). The structure of the proteolytically truncated form of annexin 1 (lacking the first 32 residues) did not indicate such a rearrangement (Weng *et al.*, 1993). Hence, even though the structure of the N-terminally truncated form of annexin 2 is known (Burger *et al.*, 1996), the high-resolution structure of the full-length protein is required to gain information about its N-terminal domain and the interaction between the N-terminal domain and the core domain.

2. Experimental procedures and results

2.1. Subcloning of annexin 2

The original pCMV5-BX-AII-E967 (5.75 kbp) construct was kindly provided by Dr Volker Gerke of the Westfälische Wilhelms Universität. The construct consisted of the human annexin 2 gene (1.05 kbp) cloned into the *Bgl*II and *Xba*I restriction sites in the pCMV5 vector. The gene was subcloned into a prokaryotic expression vector, since the pCMV-based vectors are specific for mammalian expression. The construct was manipulated *via* mutagenesis to (i) eliminate the internal *Hin*dIII and *Eco*RI restriction sites in the gene sequence, (ii) insert a *Bam*HI restriction site upstream of the start codon and (iii) mutate base 246 from a cytosine to an adenine. The protein expressed from this mutated construct contains a glutamate (E) in place of the alanine (A) at amino-acid residue 66 (the residue number is inclusive of the *N*-formyl methionine). The A66E site mutation reconstructed the epitope for monoclonal antibody H28 and did not alter the biochemical properties of annexin 2 (Thiel *et al.*, 1992).

The human annexin 2 gene was amplified from the pCMV5-BX-AII-E967 construct *via* standard PCR with the Pwo DNA Polymerase kit (Roche Molecular Biochemicals) and the Biometra T Gradient thermocycler (reaction conditions: 334 K annealing temperature, 30 s annealing time, 30 cycles, 2 m*M* $MgSO_4$, 100 µl final volume). The sequence of the forward primer was 5′-TCTACTGTTCACGAAATCCTGTGC-AAGCTCAGC-3′ and the sequence of the reverse primer was 5′-GTTAGCTG-GAAGCTTGGTGAGCACCAT-3′ (Integrated DNA Technologies, Inc.). The resulting annexin 2 PCR product contained an engineered *Hin*dIII restriction site at the 3′ terminus and was consequently subjected to restriction digest with *Hin*dIII (New England Biolabs, Inc.) for subsequent cloning into the pSE420 vector (Invitrogen, Inc.). Purifications were performed between different enzymatic reactions with a PCR purification kit (Qiagen, Inc.).

The 4.6 kbp pSE420 expression vector (Invitrogen, Inc.) was linearized with the *Nco*I restriction endonuclease (New England Biolabs, Inc.). The cohesive ends of the linearized vector were filled-in with the Klenow Fragment (New England Biolabs, Inc.) to re-establish a start codon. The vector was then subjected to a restriction digest with *Hin*dIII (New England Biolabs, Inc.), which resulted in a linearized pSE420 vector with blunt (5′) and cohesive (3′) cloning sites. Finally, the vector was dephosphorylated with CIP (Roche Molecular Biochemicals). Gel purifications were performed between different enzymatic reactions with a purification kit (Qiagen, Inc.).

The final annexin 2 insert and pSE420 vector were ligated with T4 DNA Ligase (Promega). Plasmid DNA of the colonies produced from the transformation of the ligation products into DH5α cells was purified with the standard alkaline lysis miniprep protocol. The purified plasmid DNA was then subjected to restriction mapping: potentially positive constructs were identified on the basis that (i) they were uncleavable by *Nco*I, (ii) *Hin*dIII digests should result in one fragment of 5.3 kbp and (iii) *Nde*I digests should produce fragments of 3.3 and 2.0 kbp. Eight tentative positive clones were identified.

2.2. Annexin 2 protein expression

Cultures of the eight tentative positive clones were grown (5 ml, 310 K, 250 rev min^{-1}) until the optical density reached a value of 0.8 at 600 nm, followed by induction (1 m*M* IPTG) and a 4 h incubation. SDS–PAGE analysis showed that of eight tentative positive clones, only one expressed a protein with the correct target size of 38 kDa (Fig. 1*a*) and was designated A2-A66E. To test for annexin functionality, a lipid-binding assay (Isas *et al.*, 2000) was performed using crude cell extracts of pSE420, A12-E105A (an annexin

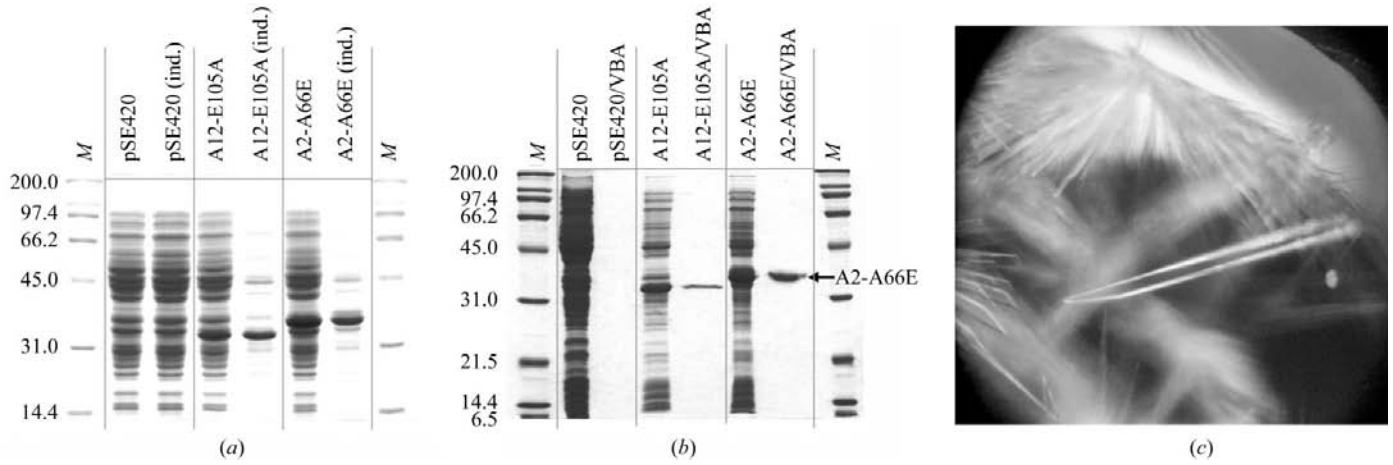

Figure 1
(*a*) A 12% SDS–PAGE gel of whole cell extracts for all lanes shows the expression characteristics of the putative annexin 2 clone. The molecular-weight markers are labeled *M*, the negative control (vector alone) is pSE420, the positive control is annexin 12 (A12-E105A) and our construct is A2-A66E. For each pair, the left lane is uninduced expression and the right lane is induced with IPTG. There is ample expression of clone A2-A66E. (*b*) To check whether the putative full-length annexin 2 clone was expressing the target protein, we performed a vesicle-binding assay. Again, as in (*a*), we have a negative control (pSE420 vector alone), a positive control (A12-E105A, annexin 12) and our construct A2-A66E. For each pair, the left lane is the soluble fraction post cell-lysis and the right lane is post calcium-induced vesicle-binding. It is clear from these results that the construct is indeed expressing A66E annexin 2 as indicated by the arrow. (*c*) Crystals of full-length annexin 2 crystal form *A*. The long needle in the center is over 1 mm in length and about 50 µm wide.

12 mutant) and A2-A66E cultures that were induced with 1 m*M* IPTG. Results of the assay showed the A2-A66E and A12-E105A proteins exhibited the typical annexin characteristic of binding phospholipid vesicles in the presence of Ca^{2+}. In the absence of Ca^{2+} no binding could be detected (Fig. 1*b*). No native *Escherichia coli* proteins have been shown to possess this Ca^{2+}-dependent phospholipid-binding property.

2.3. Annexin 2 protein purification

Large amounts of the A2-A66E protein were expressed in DH5α cells and purified using ion-exchange chromatography based on a modified version of the protocol that yielded full-length annexin 1 (Rosengarth *et al.*, 2001). Following induction, the cells were harvested by centrifugation at 6000*g* for 15 min at 277 K and resuspended in buffer *A* (50 m*M* Tris–HCl pH 8.5, 300 m*M* NaCl, 10 m*M* $MgCl_2$, 2 m*M* EGTA, 1 m*M* PMSF, 0.5 m*M* benzamidine, 10 μ*M* E-64 and 10 μg ml^{-1} aprotinin) at a ratio of 3.0 ml buffer *A* per gram of wet pellet. The sample was subsequently lysed twice with a Spectronic Unicam French press. The lysate was twice centrifuged at 39 000*g* for 30 min at 277 K. The supernatant was dialyzed against buffer *B* (10 m*M* Tris–HCl pH 8.5, 10 m*M* NaCl, 1 m*M* EGTA) and applied to a DEAE Sepharose Fast Flow (Amersham Pharmacia) column (80 ml bed volume), previously equilibrated with buffer *B*, at a flow rate of 300 ml h^{-1}. The column was washed with 250 ml buffer *B* and the A2-A66E protein eluted from the column in the flowthrough. The first 100 ml of the flowthrough was discarded and the remaining sample was then dialyzed against buffer *C* (50 m*M* MES–NaOH pH 6.4, 50 m*M* NaCl and 1 m*M* EGTA) and applied to a HiPrep 16/10 SP XL (Amersham Pharmacia) column (20 ml bed volume), previously equilibrated with buffer *C*, at a flow rate of 300 ml h^{-1}. SDS–PAGE analysis revealed that the A2-A66E protein eluted at 250 m*M* of a NaCl step gradient composed of 50 m*M* steps. The fractions containing the A2-A66E protein were pooled and concentrated *via* ultrafiltration (Amicon 8050 and Centricons). DNA sequencing and MALDI–TOF spectrometry confirmed the presence of the N-terminal domain and the molecular mass of the full-length protein (Fig. 2). The mass of the protein was determined to be 38 567.7 Da. The calculated mass for annexin 2 starting at position 2 is 38 529.9 Da. The difference in mass is most likely to arise from acetylation of the N-terminal serine residue, which has also been observed for other prokaryotic proteins (Polevoda & Sherman, 2002). This hypothesis is based on attempts to sequence the N-terminus, which showed it to be blocked. However, cyanogen bromide cleavage analysis confirmed the protein was indeed annexin 2 (Dr Agnes Henschen-Edman, personal communication).

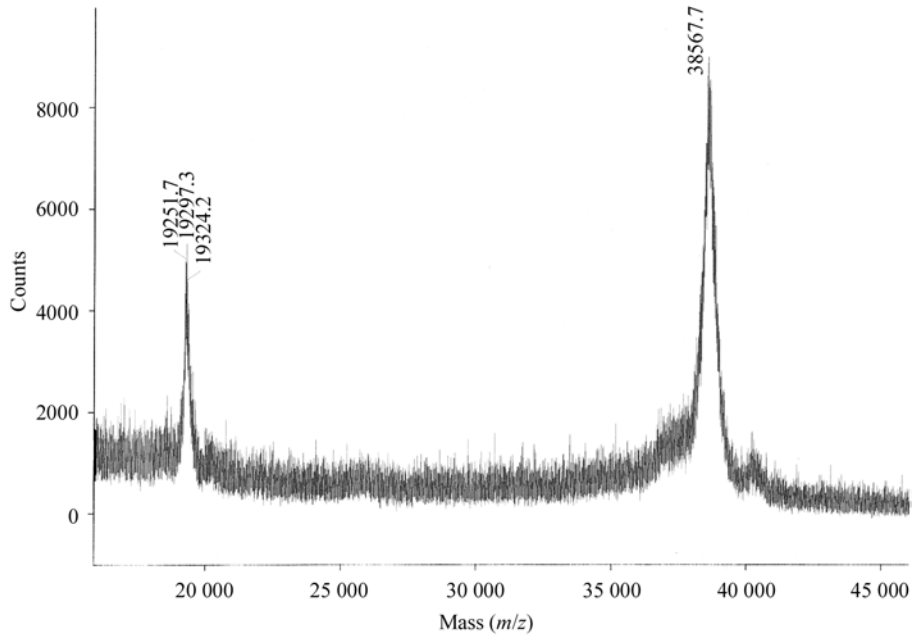

Figure 2
MALDI mass spectrometry of purified full-length annexin 2 (A2-A66E) shows a peak at 38 567.7 Da. A2-A66E has a predicted mass of 38 661.1 Da (DNA Star). Cleavage of the f-methionine would reduce the mass to 38 529.9 Da and acetylation of the following serine residue would increase the predicted mass to 38 571.9 Da. This leaves a difference of 4.2 Da from the final predicted mass of the A2-A66E protein.

2.4. Annexin 2 protein crystallization

Working at a protein stock concentration of 9.73 mg ml^{-1} in 50 m*M* MES–NaOH pH 6.4, 250 m*M* NaCl, 1 m*M* EGTA, we used various commercial kits to crystallize annexin 2. Drops were made with a volume ratio of protein solution to reservoir solution of 1:1. Crystals were obtained from two conditions. Condition *A*, using the sitting-drop vapor-diffusion method and a total drop volume of 6.0 µl (Hampton Research, Crystal Screen solution 46: 18% PEG 8000, 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* calcium acetate), yielded long thin plates ($1 \times 0.1 \times 0.05$ mm after 5–6 weeks at 277 K) that belong to space group $P2_12_12_1$, with unit-cell parameters $a = 60.97$, $b = 98.33$, $c = 176.06$ Å. With a unit-cell volume of 1 055 551 Å^3, the Matthews coefficient (Matthews, 1968) indicates the possibility of a dimer (V_M = 3.41 Å^3 Da^{-1}; predicted solvent content 64%) or a trimer (V_M = 2.28 Å^3 Da^{-1}; predicted solvent content 46%). Condition *B*, using the hanging-drop vapor-diffusion method and a total drop volume of 5.0 µl (Emerald Biostructures, Wizard II solution 38: 2.5 *M* NaCl, 0.1 *M* acetate pH 4.5, 0.2 *M* Li_2SO_4), yielded orthorhombic crystals ($0.5 \times 0.3 \times 0.15$ mm after 3–4 d at 277 K) that belong to space group $P2_12_12_1$, with unit-cell parameters $a = 48.36$, $b = 62.86$, $c = 119.11$ Å and a unit-cell volume of 361 334 Å^3. The Matthews coefficient indicates that only one monomer is present in the asymmetric unit (V_M = 2.34 Å^3 Da^{-1}; predicted solvent content 47%). The space group and unit-cell parameters of the condition *B* crystals are similar to the space group and unit-cell parameters published for the truncated annexin 2 structure (Burger *et al.*, 1996), where the crystal packing leaves ample space for a 30-residue N-terminal domain near the site of the truncation.

2.5. Data collection and preliminary data analysis

The crystals from condition *A* diffracted to 3.2 Å using a rotating-anode generator (Cu $K\alpha$ target at λ = 1.5418 Å) and an R-AXIS IV image-plate detector. A total of 180 frames of 0.5° oscillation range were collected and processed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Crystals obtained from condition *B* diffracted to 1.52 Å at the Advanced Light Source beamline 5.0.2 (λ = 0.97 Å) with a 2×2 array CCD detector (Area Detector

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Condition *A*	Condition *B*
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å, °)	a = 60.97, b = 98.33, c = 176.06	a = 48.36, b = 62.86, c = 119.11
Resolution (Å)	99.0–3.20 (3.20–3.26)	99.0–1.52 (1.52–1.55)
Reflections measured (all/unique)	165026/18032	296762/52751
Mosaicity (°)	0.55	0.32
R_{sym}† (%)	11.7 (48.1)	4.5 (60.3)
Completeness (%)	96.6 (97.0)	93.0 (83.7)
$I/\sigma(I)$, overall	13.0 (2.82)	29.1 (2.64)

† $R_{sym} = [\sum_h \sum_i |I(h,i) - \langle I(h)\rangle|/\sum_h \sum_i |I(h,i)] \times 100$, where $I(h, i)$ is the intensity value of the ith measurement of h and $\langle h\rangle$ is the corresponding mean value of h for all I measurements of h. The summation is over all measurements.

Systems Corporation, San Diego) at 100 K. Prior to being flash-cooled, the crystals were soaked in mother liquor containing 20% glycerol for 2–3 s for cryoprotection, with minimal deterioration during freezing. A total of 180 frames of 1° oscillation range were collected and processed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). All data-collection statistics are summarized in Table 1. Molecular replacement using the truncated form of annexin 2 (309 residues) as a search model will be used to solve the structure of the full-length annexin 2 (339 residues) in this high-resolution data set. Molecular replacement might not provide enough phasing power to render unbiased electron density for the N-terminal domain, as we experienced in the case of full-length annexin 1 (Rosengarth *et al.*, 2001). In this case, we will resort to heavy-atom phasing techniques. We are hopeful the structure of full-length annexin 2 will provide information on its N-terminal domain and a framework for the structural analysis of annexin 2/t-PA binding.

This work was supported by the NIH (grant GM56445). We would like to thank Jean-Philippe Cartailler for advice on experimental design and critical reading of the manuscript.

References

Burger, A., Berendes, R., Liemann, S., Benz, J., Hofmann, A., Gottig, P., Huber, R., Gerke, V., Thiel, C., Romisch, J. & Weber, K. (1996). *J. Mol. Biol.* **257**, 839–847.

Drust, D. S. & Creutz, C. E. (1988). *Nature (London)*, **331**, 88–91.

Fitzpatrick, S. L., Kassam, G., Choi, K. S., Kang, H. M., Fogg, D. K. & Waisman, D. M. (2000). *Biochemistry*, **39**, 1021–1028.

Hajjar, K. A. & Krishnan, S. (1999). *Trends Cardiovasc. Med.* **9**, 128–138.

Hawkins, T. E., Merrifield, C. J. & Moss, S. E. (2000). *Cell Biochem. Biophys.* **33**, 275–296.

Isas, J. M., Cartailler, J. P., Sokolov, Y., Patel, D. R., Langen, R., Luecke, H., Hall, J. E. & Haigler, H. T. (2000). *Biochemistry*, **39**, 3015–3022.

Kang, H. M., Choi, K. S., Kassam, G., Fitzpatrick, S. L., Kwon, M. & Waisman, D. M. (1999). *Trends Cardiovasc. Med.* **9**, 92–102.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pepinsky, R. B., Tizard, R., Mattaliano, R. J., Sinclair, L. K., Miller, G. T., Browning, J. L., Chow, E. P., Burne, C., Huang, K. S., Pratt, D., Wachter, L., Hession, C., Frey, A. Z. & Wallner, B. P. (1988). *J. Biol. Chem.* **263**, 10799–10811.

Polevoda, B. & Sherman, F. (2002). *Genome Biol.* **3**, reviews0006.1–0006.6.

Rety, S., Sopkova, J., Renouard, M., Osterloh, D., Gerke, V., Tabaries, S., Russo-Marie, F. & Lewit-Bentley, A. (1999). *Nature Struct. Biol.* **6**, 89–95.

Rosengarth, A., Gerke, V. & Luecke, H. (2001). *J. Mol. Biol.* **306**, 489–498.

Thiel, C., Osborn, M. & Gerke, V. (1992). *J. Cell Sci.* **103**, 733–742.

Waisman, D. M. (1995). *Mol. Cell Biochem.* **149–150**, 301–322.

Weng, X., Luecke, H., Song, I. S., Kang, D. S., Kim, S. H. & Huber, R. (1993). *Protein Sci.* **2**, 448–458.