Structural Analysis of an *Echinococcus granulosus* Actin-Fragmenting Protein by Small-Angle X-Ray Scattering Studies and Molecular Modeling

Eliana D. Grimm,* Rodrigo V. Portugal,† Mário de Oliveira Neto,† Nádia H. Martins,† Igor Polikarpov,† Arnaldo Zaha,*‡ and Henrique B. Ferreira*‡

*Laboratório de Biologia Molecular de Cestódeos, Centro de Biotecnologia, and ‡Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre RS, Brazil; and †Grupo de Cristalografia, Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos SP, Brazil

ABSTRACT The *Echinococcus granulosus* actin filament-fragmenting protein (EgAFFP) is a three domain member of the gelsolin family of proteins, which is antigenic to human hosts. These proteins, formed by three or six conserved domains, are involved in the dynamic rearrangements of the cytoskeleton, being responsible for severing and capping actin filaments and promoting nucleation of actin monomers. Various structures of six domain gelsolin-related proteins have been investigated, but little information on the structure of three domain members is available. In this work, the solution structure of the three domain EgAFFP has been investigated through small-angle x-ray scattering (SAXS) studies. EgAFFP exhibits an elongated molecular shape. The radius of gyration and the maximum dimension obtained by SAXS were, respectively, 2.52 ± 0.01 nm and 8.00 ± 1.00 nm, both in the absence and presence of Ca^{2+}. Two different molecular homology models were built for EgAFFP, but only one was validated through SAXS studies. The predicted structure for EgAFFP consists of three repeats of a central β-sheet sandwiched between one short and one long α-helix. Possible implications of the structure of EgAFFP upon actin binding are discussed.

INTRODUCTION

The larval stage of the cestode tapeworm *Echinococcus granulosus* is the causative agent of cystic hydatid disease or hydatidosis, recognized as one of the world's major zoonoses (1). This parasite requires two mammalian hosts for completion of its life cycle. Adult tapeworms develop in the small intestine of definitive hosts (domestic dogs and wild canids), whereas the metacestode or hydatid cyst usually develops in the liver or lungs of intermediate hosts (mainly in ungulates, and accidentally in humans). The pathological effect of the disease is caused by the pressure exerted by the hydatid cyst on the intermediate host's viscera. Within the cyst, protoscoleces are produced by asexual reproduction and develop into the adult worm when ingested by the definitive host.

The *E. granulosus* complex life cycle involves important changes in cell morphology and physiology (2). The molecular and cellular mechanisms involved in *E. granulosus* development are still largely unknown but are likely to require extensive cytoskeleton reorganization (3,4).

The actin cytoskeleton is a vital component of several key cellular and developmental processes in eukaryotes, such as motility, cytokinesis, cytoplasmic organization, and endocytosis (5). In cells, the assembly and disassembly of actin filaments, in addition to their organization into functional three-dimensional (3D) networks, are regulated by a variety of actin-binding proteins (6–10). Among these proteins, those from the gelsolin superfamily control actin organization by severing filaments, capping filament ends, and nucleating actin assembly (11).

The best-studied members of this protein family are severin (12–14) and fragmin (15,16) from *Dictyostelium discoideum* and *Physarum polycephalum*, respectively, and gelsolin (17,18) and villin (19,20) from higher organisms. A common feature of this family is the segmental organization into three (severin, fragmin) or six (gelsolin, villin) homologous domains that might have evolved from an ancestral one domain protein through a stepwise process, involving a gene triplication followed by an additional duplication event (21,17). The activities of these proteins are often modulated by signaling molecules, such as Ca^{2+} or phosphorylated phosphoinositides (22). Based on gelsolin (23,24), the most extensively studied member of the family, it is generally accepted that the second domain binds F-actin, whereas the first domain (and the fourth one, for six domain proteins) binds G-actin.

So far, only a few proteins of the gelsolin superfamily have solved 3D structures. A search in the database of protein structures indicates that of all known members of this family to date, gelsolin is the only protein that has its full-length (six domains) structure solved (25). Structures of other proteins, like villin (26) and severin (27,28), have been determined only for the first or second domains, usually bound to ligands (Ca^{2+} and/or actin). Domain comparisons between the known structures show that they all share a common fold built around a central five-stranded mixed

Submitted June 1, 2005, and accepted for publication December 12, 2005.

Address reprint requests to Dr. Henrique B. Ferreira, Laboratório de Biologia Molecular de Cestódeos, Centro de Biotecnologia, UFRGS, Caixa Postal 15005, Porto Alegre RS 91501-970, Brazil. Tel.: 55-51-33167768; Fax: 55-51-33167309; E-mail: henrique@cbiot.ufrgs.br.

0006-3495/06/05/3216/08 $2.00

doi: 10.1529/biophysj.105.067801

β-sheet, which is flanked by a long α-helix running parallel to the sheet and a short perpendicular running α-helix (25–29).

Our laboratory has previously cloned and functionally characterized a 42 kDa actin filament-fragmenting protein from *E. granulosus* (EgAFFP) (30). The recombinant EgAFFP protein is recognized by sera of ~69% of human hydatid disease patients (31) and, in vitro, was able to induce actin polymerization and sever actin filaments, confirming that it belongs to the gelsolin superfamily (30). According to sequence analysis, EgAFFP presents three repeated domains and is similar (36% identity) to the gelsolin NH_2-terminal half (G1–G3). The lack of structural data for full-length three domain members of the gelsolin superfamily, such as EgAFFP, represents an obstacle to the understanding of structure-function relationships of these smaller proteins, which are functionally equivalent to their six domain counterparts. Structural characterization of EgAFFP might help to understand how three domain members function and how they are regulated by calcium.

There is a relatively high degree of sequence identity between gelsolin and orthologous proteins, such as villin and severin, particularly in regions that are buried or forming secondary structure (25). Therefore, it has been suggested that the domain folds in gelsolin may be used to predict the geometries for the as of yet unsolved structures of the corresponding domains in other members of the gelsolin superfamily. Homology or comparative protein structure modeling allows the construction of a 3D model for a given protein sequence based on its similarity to one or more known structures. It is the most reliable method to predict the 3D structure of a protein, with an accuracy that can be comparable to that of a low resolution, experimentally determined structure (32). The accuracy of comparative modeling is correlated with sequence identity between the template structure and the modeled sequence (33). High accuracy models can be obtained when template structures with >50% sequence identity are available, medium accuracy with 30% ± 50% sequence identity, and low accuracy with <30% sequence identity.

Small-angle x-ray scattering (SAXS) is a well-established technique for obtaining structural information on macromolecules in solution under close to physiological conditions (34) and is particularly useful when high resolution structural information, such as that from NMR or x-ray crystallography, is not yet available (35) as is the case for EgAFFP. Previous work has shown that SAXS data can be conveniently used to confirm computer-modeled protein structures (36,37).

In this study, we constructed a homology model for EgAFFP both in the presence and in the absence of Ca^{2+}, based on its similarity with horse plasma gelsolin (Protein Data Bank (PDB) id: 1RGI (38) and 1D0N (25), respectively). We also obtained an ab initio low resolution model of EgAFFP through SAXS studies, which is in agreement with the structure of our homology model based on Ca^{2+}-free gelsolin (PDB id: 1D0N) but differs significantly from the homology model generated from Ca^{2+}-gelsolin bound to actin (PDB id: 1RGI). We also examined changes in EgAFFP conformation caused by Ca^{2+} and found no significant variations in the molecular envelope of the protein upon addition of Ca^{2+}. The functional implications of the proposed structural model upon actin binding are discussed.

MATERIALS AND METHODS

Protein expression and purification

Full-length EgAFFP (370 aa) (30) was expressed as a fusion with glutathione-S-transferase using the pGEX system (Amersham Biosciences, Uppsala, Sweden) and purified by affinity chromatography according to Smith and Johnson (39). The fusion protein was cleaved with thrombin (10 units/mg of matrix-bound protein) for 12 h at 4°C, and the recombinant EgAFFP was eluted in phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4) plus 3 mM dithiothreitol (DTT) and 1 mM phenylmethylsulfonyl fluoride (PMSF) and concentrated in Centricon YM-10 microconcentrators (Millipore, Bedford, MA) to 12 mg/ml. Protein purity was evaluated by SDS-PAGE (40), and concentration was determined using the Bio-Rad dye assay (Bio-Rad, Hercules, CA), with BSA as standard. Typically, 30 mg of recombinant protein were obtained from 1 liter of culture.

Pyrene actin preparation

Pyrene actin was prepared according to Cooper et al. (41) with minor modifications. Monomeric actin (G-actin) from bovine skeletal muscle (Sigma-Aldrich, St. Louis, MO) was diluted in buffer G (2 mM Tris-HCl, pH 8.0, 0.2 mM ATP, 0.5 mM DTT, 0.2 mM $CaCl_2$, 0.01% NaN_3) to a concentration of 1 mg/ml. G-actin was polymerized at room temperature for 1 h by adding KCl and $MgCl_2$ to 100 mM and 2 mM, respectively. *N*-(1-pyrene) iodoacetamide (Molecular Probes, Eugene, OR) freshly dissolved in dimethyl formamide was added from a 5 mM stock solution to the F-actin (final concentration of 160 μM) and incubated overnight at 4°C in the dark. The pyrene F-actin was collected by centrifugation at 163,000 × *g* for 90 min at 4°C, resuspended in 1 ml of buffer G, and dialyzed against the same buffer for 48 h with several changes of buffer to depolymerize the actin and remove free pyrenyl. The pyrene G-actin was then clarified at 163,000 × *g* for 90 min at 4°C and divided into 50 μl aliquots, which were frozen in liquid nitrogen and stored at −70°C. Before use, the pyrene G-actin was thawed on ice and then centrifuged at 16,000 × *g* for 15 min at 4°C to remove larger precipitates. After the absorbance at 344 nm and 290 nm was determined, the total actin concentration was calculated using mM actin = $(A_{290} - 0.127\ A_{344})/26.6\ mM^{-1}\ cm^{-1}$. The pyrene/actin ratio was calculated using an extinction coefficient of 22,000 $M^{-1}\ cm^{-1}$ for pyrene with pyrene/actin = $(A_{344}/22.0)$/mM actin.

Fluorescence spectroscopy

All fluorescence measurements with pyrene actin were performed at 20°C using an ISS K2 fluorimeter (ISS, Fluorescence, Analytical and Biomedical Instrumentation; Champaign, IL). The excitation and emission wavelength were set at 365 nm and 407 nm, respectively, with sample volumes of 300 μl. To avoid bleaching the fluorophore, pyrene actin was prepared and stored in the dark and, during fluorescence experiments, it was exposed to the light source only intermittently. The data were plotted as arbitrary relative fluorescence values.

For nucleation assays, actin (6 μM final concentration, 10% of which was pyrene labeled) was preincubated for 10 min with EgAFFP (250 nM). The polymerization was initiated by the addition of buffer F (buffer G with 100

mM KCl and 2 mM $MgCl_2$) from a 10× concentrated stock solution, and the increase in fluorescence was recorded as a function of time.

For severing assays, actin (8 μM, 25% pyrene labeled) was polymerized in buffer F for 15 min. Subsequently, the filaments were precapped with recombinant gelsolin (23) (40 nM final concentration) to avoid measuring nucleation activity. The mixture was stored overnight on ice. This solution was then diluted in buffer G to a final concentration of 400 nM in the absence or presence of EgAFFP (10 nM), and fluorescence was recorded.

Modeling of EgAFFP

The PDB database was searched for proteins homologous to EgAFFP with solved atomic structures using the BLAST program (42). The selected sequences were aligned with CLUSTALW (43), and model building by homology was performed with the MODELLER 6 program (44). For a given alignment, 20 model structures were built and evaluated with the PROCHECK software suite (45). Only the best-evaluated model was retained after the analysis.

SAXS measurements and data analysis

SAXS data were collected at the small-angle scattering beamline on the LNLS (National Synchrotron Light Laboratory, Campinas, Brazil) using a one-dimensional position-sensitive detector. EgAFFP samples at a concentration of 4, 8, and 12 mg/ml were prepared in phosphate-buffered saline (pH 7.4) in the absence or presence of 0.2 mM $CaCl_2$. EGTA was added to the Ca^{2+}-free samples to a final concentration of 1 mM. Samples were measured at a wavelength of 1.488 Å for a sample-detector distance of 1634.5 mm covering the momentum transfer range $0.21\ nm^{-1} < q < 2.35\ nm^{-1}$ ($q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle). The scattering curves of the protein solutions and the corresponding solvents were collected in 20 frames of 100 s each to monitor radiation damage and beam stability. The data were normalized to the intensity of the incident beam and corrected for detector response. The scattering of the buffer was subtracted, and the difference curves were scaled for concentration. The distance distribution functions p(r) and the radii of gyration R_g were evaluated by indirect Fourier transform using GNOM (46), which also corrected the smearing effect.

Ab initio molecular shape determination from SAXS measurements

The low resolution particle shape was restored using the ab initio procedure described by Petoukhov et al. (47), as implemented in the CREDO program. The known part of the high resolution structure was fixed, and missing portions, such as disordered loops and N-terminal residues, were then modeled to fit the experimental scattering data obtained from the intact particle. A simulated annealing procedure was used to minimize a scoring function that depends on the discrepancy between the experimental and calculated patterns and relevant penalty terms. The final shape restoration was performed using 436 residues and 378 waters assuming no molecular symmetry. The values for R_g, D_{max}, and the scattering curve simulated from the atomic coordinates of the EgAFFP homology models based both on Ca^{2+}-free and Ca^{2+}-bound gelsolin and the discrepancies between the homology model and experimental data were calculated using the CRYSOL program (48). The SUPCOMB program (49) was used to superpose the low resolution ab initio model to the structure of the homology model.

RESULTS

Functional integrity of recombinant EgAFFP

Recombinant EgAFFP functionality has been previously assessed and, compared to gelsolin, EgAFFP nucleates actin at a similar rate, whereas fragmentation is 20% less efficient (30). To assess recombinant EgAFFP functionality before its structural characterization, the pyrene actin fluorescence assay was used to indirectly determine its actin-severing and -nucleating activities by monitoring the polymerization/depolymerization rate of actin. The ability of recombinant EgAFFP to nucleate actin assembly (Fig. 1 *A*) and sever actin filaments (Fig. 1 *B*) was confirmed by its direct effect on the rate of increase and decrease, respectively, in the fluorescence of pyrene-labeled actin. These results indicated that the recombinant EgAFFP samples used for structural characterization retained these Ca^{2+}-dependent activities, which are typical of proteins of the gelsolin superfamily.

EgAFFP modeling

Sequence alignments with the database of structures revealed that the EgAFFP sequence has the highest sequence identity, 36% of identical residues and 58% of conserved residues, with the N-terminal half (amino acids 27–367) of horse plasma gelsolin (in the Ca^{2+}-free form, PDB id: 1D0N; (25)). Relatively high homology and the fact that 1D0N is the only full-length structure of the gelsolin superfamily available justified its use as a template for EgAFFP modeling. For modeling, we aligned the EgAFFP target sequence to that of amino acids 27–367 of the 1D0N template. The first 35

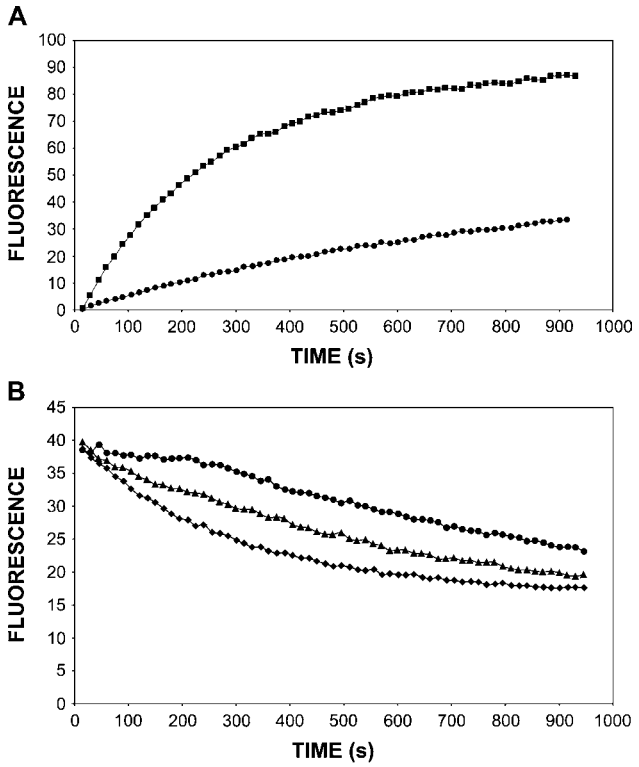

FIGURE 1 (*A*) Nucleating and (*B*) severing activity of recombinant EgAFFP. Assays were performed in the presence of Ca^{2+} as described in Materials and Methods. Nucleating activity of 60 nM EgAFFP (□). Severing activity of 4 nM (△) and 10 nM (◇) EgAFFP. Actin alone (○) was used as a control in both assays.

amino acids of the EgAFFP sequence, which would consist of two short α-helices, were eliminated, since they did not align with the template sequence. The obtained alignment was then used to generate a homology model (Fig. 2, *shaded*). As expected from the building procedure, the model resembles the N-terminal half of the 1D0N structure (domains G1–G3) and suggests that EgAFFP is composed of three independent domains (E1, Phe^{49}–Tyr^{143}; E2, Leu^{147}–His^{255}; and E3, Ser^{278}–Trp^{368}), as predicted by sequence analysis (30). Each domain adopts a similar geometric fold composed of a core of mixed β-sheet comprising five or six strands sandwiched between a long helix (helix 1) running roughly parallel to the strands in the sheet, and a short helix (helix 2) running approximately perpendicular to the strands. The sheets of E1 and E3 form a continuous 10-stranded β-sheet. E1 also contributes a single strand to the sheet of E2, forming a continuous seven-stranded β-sheet. The C-terminus of E1 immediately enters the N-terminus of E2, whereas there is an extended 23-residue loop connecting E2 and E3, which runs over the central sheet of E1. EgAFFP was also modeled in the presence of Ca^{2+}, based on the crystallographic structure of the Ca^{2+}-bound N-terminal half (G1–G3 domains) of gelsolin:actin complex (PDB id: 1RGI; (38)), from which the actin fragment was discarded. The obtained model (Fig. 2, *solid*) shows that the two halves of the β-sheet that runs continuously from E1 to E3 in the Ca^{2+}-free model are separated in the Ca^{2+}-bound model, and E1 no longer interacts with E3. In contrast, the E2–E3 linker shortens and brings E2 close to E3, whereas the first strand of E2 extends, forming a long linker between E1 and E2.

SAXS data analysis

Experimental scattering curves from EgAFFP are presented in Fig. 3, and the structural parameters derived from these curves are given in Table 1. Guinier plots (50) of the data obtained at 4 mg/ml exhibited linear profiles, indicating satisfactory monodispersity of EgAFFP both in the absence and in the presence of Ca^{2+} (Fig. 3 *A, inset*). The scattering

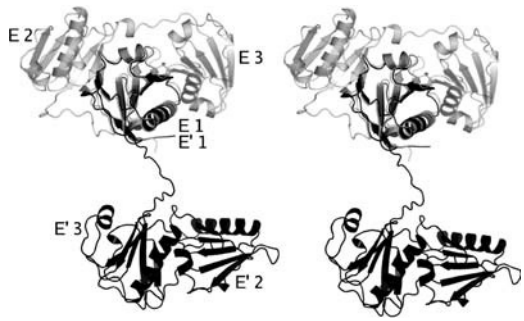

FIGURE 2 Stereo view of superposition of the EgAFFP homology model based on Ca^{2+}-bound gelsolin (PDB id 1RGI) (*solid*) and EgAFFP homology model based on Ca^{2+}-free gelsolin (PDB id 1D0N) (*shaded*). Domains E1, E2, and E3 are indicated in each model. The relative positions of both models were based on the superposition of the corresponding E1 domains.

TABLE 1 Structural parameters derived from SAXS data and models

	EgAFFP absence of Ca^{2+}			EgAFFP presence of Ca^{2+}		
Parameter/sample	Exp.*	Mod.†	DR‡	Exp.*	Mod.†	DR‡
D_{max} (nm)	8.00 ± 1.00	7.68	8.53	8.00 ± 1.00	9.02	8.47
R_g (nm)	2.53 ± 0.01	2.37	2.35	2.51 ± 0.01	2.71	2.29
Discrepancy χ	–	1.29	1.15	–	6.00	1.13
Resolution (nm)	2.67	–	2.67	2.67	–	2.67

*Exp., calculated from the experimental data.
†Mod., parameters obtained from the EgAFFP homology models.
‡DR, parameters of the experimental SAXS low resolution structure, averaged over 10 independent DR models.

curves obtained in the presence or in the absence of Ca^{2+} were virtually identical at all measured protein concentrations (Fig. 3 *A*). The radius of gyration (R_g), estimated using the Guinier approximation, was 2.53 ± 0.01 nm in the absence of Ca^{2+} and 2.51 ± 0.01 nm in the presence of Ca^{2+}. Since scattering curves at higher concentrations showed nonnegligible interference effects, we combined a low resolution part of the scattering curve obtained at 4 mg/ml (<0.5 nm^{-1}) with a high resolution part of the scattering curve measured at 12 mg/ml in a composed scattering curve, which was used in ab initio model reconstruction (Fig. 3 *B*). The pair distance distribution function p(r) for both protein samples is shown in Fig. 4 and was calculated from the GNOM fitted data. These curves also provided the EgAFFP maximum dimension (D_{max}), determined from the value of r for which the function p(r) drops to zero. The determined EgAFFP D_{max} was 8.0 ± 1.0 nm, both in the absence and in the presence of Ca^{2+}. The ratio of the radius of gyration to the maximum dimension indicates an elongated shape for EgAFPP in solution, also indicated by the shape of the p(r) function, typical of an asymmetric particle (51). To evaluate our computer-designed EgAFFP models, based on the structure of Ca^{2+}-free and Ca^{2+}-bound gelsolin domains G1–G3, the SAXS experimental curves were compared to the scattering intensity and the distance distribution p(r) function calculated from their atomic coordinates. The scattering intensity and the distance distribution p(r) curve calculated for the EgAFFP Ca^{2+}-free homology model (presented in Figs. 3 *B* and 4, respectively) are in good agreement with the SAXS experimental data, indicating that scattering data are fully compatible with this structural model. Conversely, both scattering intensity and distance distribution function for the EgAFFP Ca^{2+}-bound model grossly deviate from the scattering intensity curves and p(r) function obtained from SAXS experiments (Figs. 3 *B* and 4).

Ab initio molecular shape determination from SAXS experiments

The low resolution ab initio models of EgAFFP in the absence and presence of calcium were both built using the

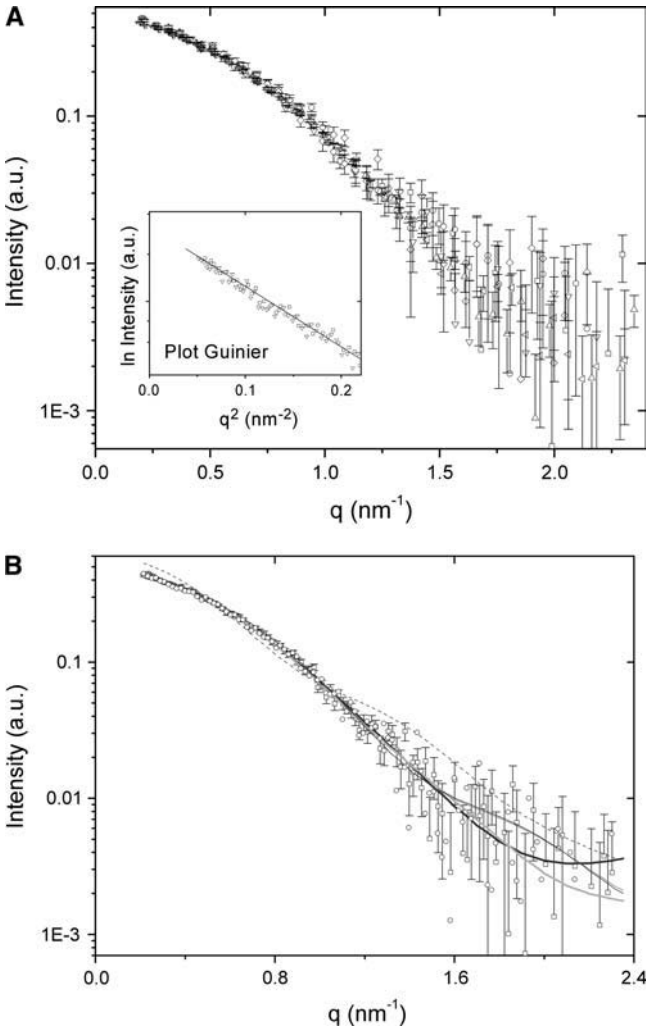

FIGURE 3 Experimental scattering curves for EgAFFP. (*A*) Experimental scattering curves corrected by the smearing effect of EgAFFP in the absence of Ca^{2+} at 4 mg/ml (□), 8 mg/ml (○), and 12 mg/ml (△) and in the presence of Ca^{2+} at 4 mg/ml (∇), 8 mg/ml (◇), and 12 mg/ml (◁). The Guinier plot with linear fit profiles is shown in the inset, allowing the estimation of $R_g = 2.53$ nm in the absence of Ca^{2+}, and $R_g = 2.51$ nm in the presence of Ca^{2+}. (*B*) Composed scattering curve of EgAFFP corrected by the smearing effect in the absence of Ca^{2+} (□), in the presence of Ca^{2+} (○), and scattering intensities from the DR models (*thick solid line*), from the homology model based on the N-terminal half of Ca^{2+}-free gelsolin (PDB id 1D0N, *thick shaded line*), from the homology model based on the N-terminal half of Ca^{2+}-bound gelsolin (PDB id 1RGI, *traced line*), from the crystallographic structure of the Ca^{2+}-bound C-terminal half of gelsolin (PDB id 1NPH, *thin shaded line*), and from the x-ray structure of the C-terminal half of gelsolin bound to Ca^{2+} and actin (PDB id 1H1V, *thin solid line*). The actin structure was extracted from the later model.

CREDO program from the composed x-ray scattering curves. The models were derived from experimental data assuming P1 symmetry. To verify the validity of the dummy residue (DR) models, several independent constructions were made from different starting conditions and, in all cases, very similar envelopes were obtained. Virtually identical low resolution models were also obtained when the GASBOR

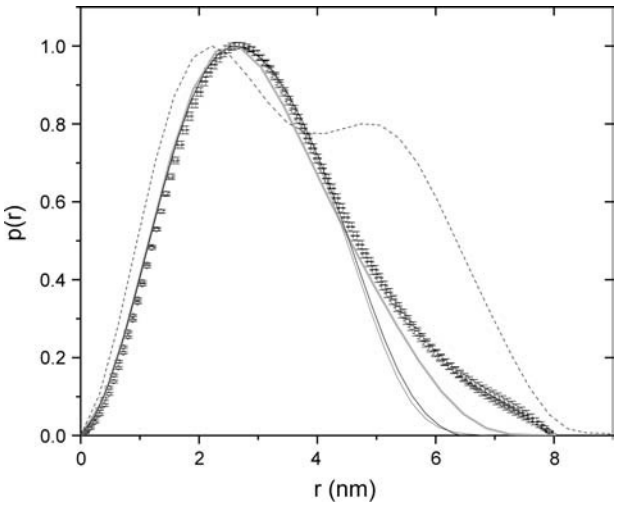

FIGURE 4 Comparison of distance distribution functions for EgAFFP. The p(*r*) curves obtained were calculated using the GNOM program, and the experimentally derived distributions for EgAFFP are given in the absence (□) and in the presence (○) of Ca^{2+}. Distributions were calculated for the homology model based on the N-terminal half of Ca^{2+}-free gelsolin (PDB id 1D0N, *thick shaded line*), for the homology model based on the N-terminal half of Ca^{2+}-bound gelsolin (PDB id 1RGI, *traced line*), for the crystallographic structure of the Ca^{2+}-bound C-terminal half of gelsolin (PDB id 1NPH, *thin shaded line*), and for the x-ray structure of the C-terminal half of gelsolin bound to Ca^{2+} and actin (PDB id 1H1V, *thin solid line*). The actin structure was extracted from the later model.

program (52) was used. The homology model of EgAFFP based on Ca^{2+}-free gelsolin, but not the model based on the Ca^{2+}/gelsolin:actin complex, could be unambiguously positioned inside the SAXS-derived molecular envelopes of EgAFFP both in the absence and in presence of Ca^{2+} (Fig. 5). The structural parameters calculated for the DR models and the Ca^{2+}-free EgAFFP homology model and the structural parameters obtained from experimental SAXS measurements (all presented in Table 1) show very good agreement, further supporting our computer-designed Ca^{2+}-free EgAFFP structure. Furthermore, molecular envelopes obtained from the SAXS curves measured for EgAFFP in the absence and presence of Ca^{2+} were virtually identical, as one would expect given the extremely high similarity of the initial scattering curves (Fig. 3). At the same time, structural parameters derived from the EgAFFP homology model based on the Ca^{2+}-bound gelsolin:actin complex are very different from the experimental data (Table 1). Moreover, fitting of the predicted x-ray scattering curve based on this model with the experimental scattering data (with and without Ca^{2+}) indicates gross differences between these curves (Fig. 3 *B*), which results in a discrepancy parameter $\chi = 6.0$. Given the fact that the structure of the N-terminal half of gelsolin (domains G1–G3) closely resembles its C-terminal half (domains G4–G6), we also compared the predicted x-ray scattering intensity curves and distance distribution for crystallographic models of G4–G6 domains of Ca^{2+}-bound gelsolin (PDB id: 1NPH (53)) and Ca^{2+}/actin:gelsolin

FIGURE 5 Stereo views showing the superposition of the Ca^{2+}-free EgAFFP homology model with the envelope obtained from the 10 DR models (*spheres*). Only DR models retrieved from the Ca^{2+}-free EgAFFP scattering curves are shown since these are virtually identical to the DR models based on Ca^{2+}-bound EgAFFP scattering curves. Middle and bottom images are rotated by 90° around the *y* and *x* axes, respectively, compared to the top image.

complex (PDB id: 1H1V (54)). Actin structure was removed from the later model. Both G4–G6 gelsolin structures agree with the experimental x-ray scattering curves significantly better than the EgAFFP model based on Ca^{2+}-bound G1–G3 domains of gelsolin (Fig. 3 *B*; $\chi = 1.77$ for 1NPH and $\chi = 1.78$ for 1H1V crystallographic models, respectively) but not as well as the EgAFFP model based on the Ca^{2+}-free G1–G3 gelsolin structure. Distance distribution functions for G4–G6 gelsolin domains also show smaller deviations from the experimentally derived curve as compared to the EgAFFP model based on Ca^{2+}-bound G1–G3 gelsolin domains but not when compared with the model based on the Ca^{2+}-free G1–G3 gelsolin structure (Fig. 4).

DISCUSSION

We have previously cloned and expressed a functional three domain EgAFFP and demonstrated that this protein has two actin-binding sites and is regulated by calcium (30). To understand how this protein functions and is regulated by calcium, we started its structural characterization. Our attempts to solve its structure by x-ray crystallography have not been successful so far due to problems with protein crystallization. Therefore, we have used the alternative approaches of SAXS studies and molecular homology modeling to obtain structural information on EgAFFP.

Overall SAXS results (R_g, D_{max}, and DR models) indicate that EgAFFP in solution both in the presence and in the absence of Ca^{2+} is an asymmetric elongated protein. This is in agreement with the homology model based on the Ca^{2+}-free gelsolin structure (PDB id 1D0N) in which domains E1 and E3 are gathered at one side of the protein, forming an intimate contact through a continuous β-sheet, whereas E2 is relatively apart from them, on the other side. Small differences in R_g and D_{max} between the SAXS measurements and the homology model are probably due to subtle variations in position and orientation of the three distinct domains and their connecting loops in the model.

Comparison of x-ray data obtained in the absence and in the presence of Ca^{2+} indicates that Ca^{2+} does not significantly change x-ray scattering curves, rendering practically identical radii of gyration, maximum dimensions, and overall shapes of the protein in both Ca^{2+}-free and Ca^{2+}-bound states. This is surprising, since the homologous six domain protein, gelsolin, is known to suffer drastic domain shifts upon interaction with Ca^{2+} alone (53–56) or with Ca^{2+} and actin (38). As a consequence, EgAFFP homology models based on these two different structural templates (Ca^{2+}-free gelsolin and Ca^{2+}/gelsolin:actin complex, domains G1–G3) have very different molecular shapes (Fig. 2). However, only one of them (the Ca^{2+}-free gelsolin-based model) is in agreement with the experimental SAXS data measured from the EgAFFP samples both in the absence and in the presence of Ca^{2+}, suggesting that addition of Ca^{2+} alone is not sufficient to promote conformational changes in EgAFFP similar to those observed in the Ca^{2+}-bound N-terminal half of gelsolin: actin complex.

EgAFFP is more similar to the N-terminal half (36% of identical residues and 58% of conserved residues) than to the C-terminal half of gelsolin (29% of identical residues and 43% of conserved residues). Therefore, the G1–G3 gelsolin structure (in its Ca^{2+}-free form) was used for homology modeling. The predicted x-ray data for G4–G6 gelsolin structures also agree with the experimental x-ray scattering curves but less consistently than the model based on the G1–G3 Ca^{2+}-free gelsolin structure.

In EgAFFP, at least one actin-binding site is known to be Ca^{2+} dependent (30), but apparently the mere binding of Ca^{2+} does not determine structural changes large enough to be detected by SAXS. It is possible that more significant domain shifts would depend upon actin binding to EgAFFP, characterizing a different behavior from that of gelsolin, in which calcium activation triggers large domain movements both in the N- and C-terminal halves (53–59). This issue remains to be experimentally addressed for EgAFFP.

In gelsolin, Ca^{2+} activation is mediated by type 1 and type 2 Ca^{2+}-binding sites (54). Type 1 Ca^{2+}-binding sites are shared between actin and gelsolin domains, allowing the prediction that those domains will bind G-actin in a Ca^{2+}-sensitive matter. Type 2 Ca^{2+}-binding sites are completely contained within gelsolin domains, facilitating disruption of interactions between domains to release latches and render the activated gelsolin able to bind actin (59). Results obtained through sequence alignment and the molecular homology model (data not shown) indicate that EgAFFP has one type 1 Ca^{2+}-binding site, located on the first domain (E1), and three type 2 Ca^{2+}-binding sites, one on each domain. Considering this is the same organization found in each half of gelsolin, the different EgAFFP behavior upon Ca^{2+} binding cannot be simply attributed to the nature of Ca^{2+}-binding sites.

It has been shown that an acidic pH could induce conformational changes in gelsolin similar to those predicted to occur upon Ca^{2+} binding (60). Therefore, we cannot discard, at least formally, the hypothesis that the conformational differences between Ca^{2+}-bound EgAFFP and Ca^{2+}-bound gelsolin are due, at least to some extent, to differences in pH conditions in which EgAFFP SAXS measurements (this work) and gelsolin crystallization (38) were performed. It would be possible that gelsolin crystallization at an acidic pH (4.7) could have induced the large conformational change observed in its N-terminal (G1–G3) domains.

In conclusion, based on the low resolution ab initio structures and the Ca^{2+}-free homology model, we can infer that EgAFFP is a compact protein very similar to the N-terminal half of Ca^{2+}-free gelsolin. In contrast to what has been described for six domain members of the gelsolin family, the addition of Ca^{2+} does not result in major conformational changes in EgAFFP, suggesting a distinct functioning mechanism for three domain members of the gelsolin family. Subtle structural changes of regulatory relevance upon Ca^{2+} binding are likely to occur, but their study will possibly require EgAFFP crystallization. It also remains to be clarified how actin-binding would affect EgAFFP conformation and how EgAFFP binds to and severs F-actin.

We thank Dr. Elizabeth Cortez-Herrera for the gift of purified gelsolin.

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico, Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul, and the Network for Research and Training in Parasitic Diseases at the Southern Cone of Latin America/Swedish International Development Agency. E.D.G. was a recipient of a fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico. R.V.P. was a recipient of a fellowship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. M.O.N. and N.H.M. were recipients of fellowships from Fundação de Amparo à Pesquisa do Estado de São Paulo.

REFERENCES

1. Thompson, R. C. A., and D. P. McManus. 2001. Aetiology: parasites and life-cycles. *In* WHO/OIE Manual on Echinococcosis in Humans and Animals: A Public Health Problem of Global Concern. J. Eckert, M. A. Hemmell, F.-X. Meslin, and Z. S. Pawlowski, editors. World Organization for Animal Health, Paris. 1–19.
2. Howell, M. J., and J. D. Smyth. 1995. Maintenance and cultivation of *Echinococcus* species in vivo and in vitro. *In Echinococcus* and Hydatid Disease. R. C. A. Thompson and A. J. Lymbery, editors. Allen & Unwin, London. 201–232.
3. Galindo, M., M. J. González, and N. Galanti. 2002. *Echinococcus granulosus* protoscolex formation in natural infections. *Biol. Res.* 35:365–371.
4. Martínez, C., R. Paredes, R. P. Stock, A. Saralegui, M. Andreu, C. Cabezón, R. Ehrlich, and N. Galanti. 2005. Cellular organization and appearance of differentiated structures in developing stages of the parasitic platyhelminth *Echinococcus granulosus*. *J. Cell. Biochem.* 94:327–335.
5. Pollard, T. D., L. Blanchoin, and R. D. Mullins. 2000. Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annu. Rev. Biophys. Biomol. Struct.* 29:545–576.
6. Pollard, T. D., and J. A. Cooper. 1986. Actin and actin-binding proteins. A critical evaluation of mechanisms and functions. *Annu. Rev. Biochem.* 55:987–1035.
7. Stossel, T. P. 1993. On the crawling of animal cells. *Science.* 260: 1086–1094.
8. Welch, M. D., D. A. Holtzman, and D. G. Drubin. 1994. The yeast actin cytoskeleton. *Curr. Opin. Cell Biol.* 6:110–119.
9. Schafer, D. A., and J. A. Cooper. 1995. Control of actin assembly at filament ends. *Annu. Rev. Cell Dev. Biol.* 11:497–518.
10. Sun, H. Q., K. Kwiatkowska, and H. L. Yin. 1995. Actin monomer binding proteins. *Curr. Opin. Cell Biol.* 7:102–110.
11. Yin, H. L. 1987. Gelsolin: calcium- and polyphosphoinositide-regulated actin-modulating protein. *Bioessays.* 7:176–179.
12. André, E., F. Lottspeich, M. Schleicher, and A. Noegel. 1988. Severin, gelsolin, and villin share a homologous sequence in regions presumed to contain F-actin severing domains. *J. Biol. Chem.* 263:722–727.
13. Eichinger, L., A. A. Noegel, and M. Schleicher. 1991. Domain structure in actin-binding proteins: expression and functional characterization of truncated severin. *J. Cell Biol.* 112:665–676.
14. Eichinger, L., and M. Schleicher. 1992. Characterization of actin- and lipid-binding domains in severin, a Ca^{2+}-dependent F-actin fragmenting protein. *Biochemistry.* 31:4779–4787.
15. Ampe, C., and J. Vandekerckhove. 1987. The F-actin capping proteins of *Physarum polycephalum*: cap42(a) is very similar, if not identical, to fragmin and is structurally and functionally very homologous to gelsolin; cap42(b) is *Physarum* actin. *EMBO J.* 6:4149–4157.
16. Gettemans, J., Y. De Ville, J. Vandekerckhove, and E. Waelkens. 1993. Purification and partial amino acid sequence of the actin-fragmin kinase from *Physarum polycephalum*. *Eur. J. Biochem.* 214: 111–119.
17. Way, M., and A. G. Weeds. 1988. Nucleotide sequence of pig plasma gelsolin. Comparison of protein sequence with human gelsolin and other actin severing proteins shows strong homologies and evidence for large internal repeats. *J. Mol. Biol.* 203:1127–1133.
18. Cunningham, C. C., T. P. Stossel, and D. J. Kwiatkowski. 1991. Enhanced motility in NIH 3T3 fibroblasts that overexpress gelsolin. *Science.* 251:1233–1236.

19. Janmey, P. A., and P. T. Matsudaira. 1988. Functional comparison of villin and gelsolin: effect of Ca^{++}, KCl and polyphosphoinositides. *J. Biol. Chem.* 263:16738–16743.

20. Friederich, F., K. Vancompernolle, C. Huet, M. Goethals, J. Finidori, J. Vandekerckhove, and D. Louvard. 1992. An actin binding site containing a conserved motif of charged amino acid residues is essential for morphogenic effect of villin. *Cell.* 70:81–92.

21. Kwiatkowski, D. J., T. P. Stossel, S. H. Orkin, J. E. Mole, H. Colten, and H. L. Yin. 1986. Plasma and cytoplasmic gelsolins are encoded by a single gene and contain a duplicated actin binding domain. *Nature.* 323:455–458.

22. Janmey, P. A. 1994. Phosphoinositides and calcium as regulators of cellular actin assembly and disassembly. *Annu. Rev. Physiol.* 56: 169–191.

23. Way, M., J. Gooch, B. Pope, and A. G. Weeds. 1989. Expression of human plasma gelsolin in *Escherichia coli* and dissection of actin binding sites by segmental deletion mutagenesis. *J. Cell Biol.* 109: 593–605.

24. Pope, B., M. Way, and A. G. Weeds. 1991. Two of the three actin-binding domains of gelsolin bind to the same subdomain of actin. *FEBS Lett.* 280:70–74.

25. Burtnick, L. D., E. K. Koepf, J. M. Grimes, E. Y. Jones, D. I. Stuart, P. J. McLaughlin, and R. C. Robinson. 1997. The crystal structure of plasma gelsolin: implications for actin severing, capping and nucleation. *Cell.* 90:661–670.

26. Markus, M. A., P. Matsudaira, and G. Wagner. 1997. Refined structure of villin-14T and a detailed comparison with other actin-severing proteins. *Protein Sci.* 6:1197–1209.

27. Schnuchel, A., R. Wiltscheck, L. Eichinger, M. Schleicher, and T. A. Holak. 1995. Structure of severin domain 2 in solution. *J. Mol. Biol.* 247:21–27.

28. Puius, Y. A., E. V. Fedorov, L. Eichinger, M. Schleicher, and S. C. Almo. 2000. Mapping the functional surface of domain 2 in the gelsolin superfamily. *Biochemistry.* 39:5322–5331.

29. McLaughlin, P. J., J. T. Gooch, H.-G. Mannherz, and A. G. Weeds. 1993. Structure of gelsolin segment 1-actin complex and the mechanism of filament severing. *Nature.* 364:685–692.

30. Cortez-Herrera, E., R. R. Yamamoto, J. J. S. Rodrigues, S. E. Farias, H. B. Ferreira, and A. Zaha. 2001. *Echinococcus granulosus:* cloning and functional in vitro characterization of an actin filament fragmenting protein. *Exp. Parasitol.* 97:215–225.

31. Virginio, V. G., A. Hernández, M. B. Rott, K. M. Monteiro, A. F. Zandonai, A. Nieto, A. Zaha, and H. B. Ferreira. 2003. A set of recombinant antigens from *Echinococcus granulosus* with potential for use in the immunodiagnosis of human cystic hydatid disease. *Clin. Exp. Immunol.* 132:309–315.

32. Marti-Renom, M. A., A. C. Stuart, A. Fiser, R. Sanchez, F. Melo, and A. Sali. 2000. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29:291–325.

33. Baker, D., and A. Sali. 2001. Protein structure prediction and structural genomics. *Science.* 294:93–96.

34. Trewhella, J. 1997. Insights into biomolecular function from small-angle scattering. *Curr. Opin. Struct. Biol.* 7:702–708.

35. Svergun, D. I., and M. H. J. Koch. 2003. Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* 66: 1735–1782.

36. Casadio, R., E. Polverini, P. Mariani, F. Spinozzi, F. Carsughi, A. Fontana, P. Polverino de Laureto, G. Matteucci, and C. M. Bergamini. 1999. The structural basis for the regulation of tissue transglutaminase by calcium ion. *Eur. J. Biochem.* 262:672–679.

37. Occhipinti, E., P. L. Martelli, F. Spinozzi, F. Corsi, C. Formantici, L. Molteni, H. Amenitsch, P. Mariani, P. Tortora, and R. Casadio. 2003. 3D structure of *Sulfolobus solfataricus* carboxypeptidase developed by molecular modeling is confirmed by site-directed mutagenesis and small angle x-ray scattering. *Biophys. J.* 85:1165–1175.

38. Burtnick, L. D., D. Urosev, E. Irobi, K. Narayan, and R. C. Robinson. 2004. Structure of the N-terminal half of gelsolin bound to actin: roles in severing, apoptosis and FAF. *EMBO J.* 23:2713–2722.

39. Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *E. coli* as fusions with gluthatione S-transferase. *Gene.* 67:31–40.

40. Laemmli, U. K. 1970. Cleavage of structured proteins during assembly of head of bacteriophage T4. *Nature.* 227:680–685.

41. Cooper, J. A., S. B. Walker, and T. D. Pollard. 1983. Pyrene actin: documentation of the validity of a sensitive assay for actin polymerization. *J. Muscle Res. Cell Motil.* 4:253–262.

42. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.

43. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.

44. Šali, A., and T. L. Blundell. 1993. Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815.

45. Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.

46. Svergun, D. I. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 25:495–503.

47. Petoukhov, M. V., N. A. J. Eady, K. A. Brown, and D. I. Svergun. 2002. Addition of missing loops and domains to protein models by x-ray solution scattering. *Biophys. J.* 83:3113–3125.

48. Svergun, D. I., C. Barberato, and M. H. Koch. 1995. CRYSOL: a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* 28:768–773.

49. Kozin, M. B., and D. I. Svergun. 2001. Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* 34:33–41.

50. Guinier, A., and G. Fournet. 1955. Small-Angle Scattering of X-Rays. John Wiley and Sons, New York.

51. Glatter, O., and O. Kratky. 1982. Small Angle X-Ray Scattering. Academic Press, London, New York.

52. Svergun, D. I., M. V. Petoukhov, and M. H. J. Koch. 2001. Determination of domain structure of proteins from x-ray solution scattering. *Biophys. J.* 80:2946–2953.

53. Kolappan, S., J. T. Gooch, A. G. Weeds, and P. J. McLaughlin. 2003. Gelsolin domains 4–6 in active, actin-free conformation identifies sites of regulatory calcium ions. *J. Mol. Biol.* 329:85–92.

54. Choe, H., L. D. Burtnick, M. Mejillano, H. L. Yin, R. C. Robinson, and S. Choe. 2002. The calcium activation of gelsolin: insights from the 3 Å structure of the G4–G6/actin complex. *J. Mol. Biol.* 324:691–702.

55. Pope, B. J., J. T. Gooch, and A. G. Weeds. 1997. Probing the effects of calcium on gelsolin. *Biochemistry.* 36:15848–15855.

56. Robinson, R. C., M. Mejillano, V. P. Le, L. D. Burtnick, H. L. Yin, and S. Choe. 1999. Domain movement in gelsolin: a calcium-activated switch. *Science.* 286:1939–1942.

57. Gremm, D., and A. Wegner. 2000. Gelsolin as a calcium-regulated actin filament-capping protein. *Eur. J. Biochem.* 267:4339–4345.

58. Lin, K.-M., M. Mejillano, and H. L. Yin. 2000. Ca^{2+} regulation of gelsolin by its C-terminal tail. *J. Biol. Chem.* 275:27746–27752.

59. Narayan, K., S. Chumnarnsilpa, H. Choe, E. Irobi, D. Urosev, U. Lindberg, C. E. Schutt, L. D. Burtnick, and R. C. Robinson. 2003. Activation in isolation: exposure of the actin-binding site in the C-terminal half of gelsolin does not require actin. *FEBS Lett.* 552:82–85.

60. Lamb, J. A., P. G. Allen, B. Y. Tuan, and P. A. Janmey. 1993. Modulation of gelsolin function. Activation at low pH overrides Ca^{2+} requirement. *J. Biol. Chem.* 268:8999–9004.