

Alexandra S. Solovyova · Nicola Meenan
Lindsay McDermott · Antonio Garofalo
Jannette E. Bradley · Malcolm W. Kennedy
Olwyn Byron

The polyprotein and FAR lipid binding proteins of nematodes: shape and monomer/dimer states in ligand-free and bound forms

Received: 11 June 2002 / Revised: 13 February 2003 / Accepted: 13 February 2003 / Published online: 12 April 2003
© EBSA 2003

Abstract Nematodes produce two classes of small, helix-rich fatty acid- and retinol-binding proteins whose structures and in vivo functions remain to be elucidated. These are the polyprotein allergens (NPA) and the FAR proteins. The solution properties of recombinant forms of these proteins from parasitic [*Ascaris suum* (As) and *Onchocerca volvulus* (Ov)] and free-living [*Caenorhabditis elegans* (Ce)] nematodes have been examined. Analytical ultracentrifugation (AUC) showed that, contrary to previous findings, the rAs-NPA-1A polyprotein unit (~15 kDa) is a monomer, and this stoichiometry is unaltered by ligand (oleic acid) binding. The rOv-FAR-1 and rCe-FAR-5 proteins differ in that the former forms a tight dimer and the latter a monomer, and these oligomeric states are also unaffected by ligand binding or protein concentration. Sedimentation equilibrium

experiments showed that the partial specific volume \bar{v} of the unliganded proteins agree well with the value calculated from amino acid composition extrapolated to experimental temperature, and was unaffected upon ligand binding. Data from small-angle X-ray scattering (SAXS) indicated that both of the monomeric proteins rAs-NPA-1A and rCe-FAR-5 are globular, although slightly elongated and flattened. These data are in good agreement with shapes predicted from sedimentation velocity experiments and hydrodynamic bead modelling. On the basis of functional and secondary structural homology with the ligand-binding domain of the retinoic acid receptor RXR α , de novo atomic resolution structures for rAs-NPA-1A and rCe-FAR-5 have been constructed which are consistent with the SAXS and hydrodynamic data.

Presented at the Conference for Advances in Analytical Ultracentrifugation and Hydrodynamics, 8–11 June 2002, Grenoble, France

A. S. Solovyova (✉) · O. Byron
Division of Infection and Immunity,
Institute of Biomedical and Life Sciences,
Joseph Black Building, University of Glasgow,
Glasgow, G12 8QQ, UK
E-mail: alexandra.solovyova@bio.gla.ac.uk
Fax: +44-141-3304600

A. S. Solovyova
Institute for Problems of Cryobiology and Cryomedicine,
National Academy of Science of the Ukraine,
23 Pereyaslavskaya Str., 61015 Kharkov, Ukraine

N. Meenan · L. McDermott
Department of Chemistry, Joseph Black Building,
University of Glasgow, Glasgow, G12 8QQ, UK

L. McDermott · M. W. Kennedy
Division of Environmental and Evolutionary Biology,
Institute of Biomedical and Life Sciences,
Graham Kerr Building, University of Glasgow,
Glasgow, G12 8QQ, UK

A. Garofalo · J. E. Bradley
School of Life and Environmental Science, University of
Nottingham, University Park, Nottingham, NG7 2RD, UK

Keywords Analytical ultracentrifugation · Hydrodynamic bead modelling · Nematode fatty acid binding proteins · Protein-ligand interaction · Small-angle X-ray scattering

Introduction

Vertebrates and other animal groups are known to produce two types of small, soluble, lipid-binding proteins. The first of these (with a monomer mass of ~14 kDa) is the FABP/P2/CRBP/CRABP family of intracellular fatty acid- and retinoid-binding proteins that function to transport small, insoluble or oxidation-sensitive lipids between sites within cells (Coe and Bernlohr 1998; Storch and Thumser 2000). The other is the lipocalins (~20 kDa) that transport a number of ligand types in blood or secretions, some of which may interact with other proteins in circulation or with cell surface receptors (Flower 1996). Both of these classes of proteins are rich in beta structure, with only one or two short helical stretches. Smaller (~7–9 kDa) fatty acid

storage/transporter proteins have been reported from plants and parasitic flatworms; these are helix-rich, with no apparent beta content (Barrett et al. 1997; Janssen and Barrett 1995; Lerche and Poulsen 1998).

Two new classes of small, lipid-binding protein have recently been described from nematodes (roundworms), and may be confined to this group of medically, economically and ecologically important and abundant organisms (Kennedy 2001). They originally received attention for their presence in the secretions of parasitic nematodes, and therefore for their potential roles in the pathology, acquired resistance and immunomodulation in the infections. Moreover, their ligand-binding specificities could be relevant to modification of the host tissue occupied by the parasites. Both classes are also notable for the strong apolarity of their binding sites, as attested to by the unprecedented degree of blue shift they cause in the fluorescence emission of the environment-sensitive dansyl fluorophore (Kennedy et al. 1995a, 1995b, 1997; MacGregor and Weber 1986).

The first of these new classes is the nematode polyprotein allergens/antigens (NPA), which are synthesized as large precursor polypeptides comprising about 11 repeat units that are posttranslationally cleaved into multiple copies of the ~14.4 kDa functional proteins (Kennedy 2000). The polyprotein array in one species of parasite (*Ascaris*) has units that are very similar to one another (Xia et al. 2000), but other species produce arrays with units whose amino acid sequences differ dramatically (e.g. *Dictyocaulus viviparus*, *Caenorhabditis elegans*) (Blaxter 1998; Britton et al. 1995). Such repetitive polyproteins appear to be rare in nature, but one (filaggrin) is produced in abundance by terminally differentiating keratinocytes of mammals (Rothenagel and Steinert 1990). One of our interests, therefore, is to use the NPAs to understand what is special about repetitive polyproteins: what is the advantage of synthesis as repetitive polyproteins, what is the danger of not doing so, why are other proteins not made this way? An understanding of their structures (tertiary and quaternary) and protein-protein interactions could begin to answer these questions.

The other class of nematode lipid binding proteins is the FARs (fatty acid and retinol binding), which are unrelated to the NPAs but have similar binding properties (Garofalo et al. 2002; Kennedy et al. 1997; Prior et al. 2001). They are larger (~20 kDa), helix-rich and are, like the NPAs, secreted by parasites into their host's tissues, plant or animal (Garofalo et al. 2002; Prior et al. 2001). They have similar lipid binding properties to the NPAs, but presumably play a different role in the biology of, and parasitism by, nematodes. The first FAR protein to be investigated was Ov-FAR-1 (formerly Ov20) of the river blindness parasite *Onchocerca volvulus* (Kennedy et al. 1997). Only one type of FAR protein has been identified in any species of parasitic nematode, but the genome of *C. elegans* appears to encode eight FAR-like proteins (Garofalo et al. 2003). The latter are under investigation in anticipation that others will be found in parasites and

because *C. elegans* provides a more tractable model for investigating the biological role of the FARs.

In this study we have examined one NPA (As-NPA-1A), representing a single unit from the array encoding the ABA-1 allergen of *Ascaris suum*, and two FARs, one from the *O. volvulus* parasite (Ov-FAR-1), and the other from the free-living *C. elegans* (Ce-FAR-1). The FAR proteins selected are representative of the two major subfamilies of FAR proteins (Garofalo et al. 2003). We have examined their properties in solution by analytical ultracentrifugation and by small-angle X-ray scattering and report here on their protein-protein self-interactions, how these are unaffected by ligand binding, and make predictions on their shape based on low-resolution modelling.

Materials and methods

Recombinant As-NPA-1A protein expression and purification

Recombinant As-NPA-1A (henceforth rAs-NPA-1A) was derived from the A1 type unit of the ABA-1 polypeptide of *Ascaris suum*. The protein was over-expressed in *Escherichia coli* and purified according to a previously published protocol (McDermott et al. 2001). The purified protein migrated as a single band with the expected molecular weight on an SDS-PAGE gel (data not shown).

Recombinant Ov-FAR-1 protein expression and purification

Recombinant Ov-FAR-1 (henceforth rOv-FAR-1) was over-expressed in *E. coli* and purified according to a previously published protocol (Kennedy et al. 1997). The purified protein migrated as a single band with the expected molecular weight on an SDS-PAGE gel (data not shown).

Recombinant Ce-FAR-5 protein expression and purification

A fragment of Ce-Far-5 (from *C. elegans*), corresponding to the entire coding region except for the predicted signal sequence, was amplified from plasmid stocks with primers, cloned into pET30 Xa/LIC (Novagen) using standard protocols and then sequenced. The recombinant pET30 Xa/LIC plasmids were maintained in *E. coli* NovaBlue cells (Novagen) and transformed into *E. coli* BLR(DE3) cells (Novagen) for expression of the recombinant fusion protein. The expression and purification of the protein was performed in accordance with the kit manufacturer's instructions for high yield. rCe-Far-5 was released from the glutathione S-transferase moiety by cleavage with thrombin. The purified proteins were exhaustively dialysed at 4 °C against phosphate buffered saline (171 mM NaCl, 3.35 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.2) and subsequently passed down a 2 mL Extracti-Gel D column (Pierce) to remove any residual detergent. The purity of the protein was confirmed by standard SDS-PAGE analysis on a tris/glycine mini protean 10–20% (w/v) polyacrylamide gel (Bio-Rad).

Protein concentration determination

The concentrations of the protein solutions were determined spectrophotometrically using a Shimadzu UV-160A visible recording spectrophotometer and 1 cm path-length quartz cuvettes. Molar extinction coefficients (at 280 nm) for the proteins were

estimated from their tyrosine and tryptophan content (Gill and von Hippel 1989): 10,810 M⁻¹ cm⁻¹ for rAs-NPA-1A, 5120 M⁻¹ cm⁻¹ for rOv-FAR-1, 6200 M⁻¹ cm⁻¹ for rCe-FAR-5.

Sample preparation

rAs-NPA-1A and rCe-FAR-5 were characterized in sodium phosphate buffer [141 mM NaCl, 2.7 mM KCl, 10.1 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.3 (PBS)]. The buffer for rOv-FAR-1 was 20 mM tris, 0.1 mM EDTA, pH 8.0, because initial biophysical experiments were performed on an rOv-FAR-1 sample that had been prepared for other studies requiring that the protein be in a buffer at pH 8. Sedimentation velocity studies confirm that exchange of this protein into sodium phosphate buffer has no effect on the sedimentation velocity profile of rOv-FAR-1 (data not shown).

Binding experiments with DAUDA (Molecular Probes Europe, Leiden, The Netherlands) were conducted at a molar ratio of protein:DAUDA=1:2 to ensure that all protein had bound DAUDA. DAUDA stock solutions (2.3 mM and 115 mM) in ethanol (EtOH) were added to the protein stock solution and diluted with buffer to the required protein concentration. In parallel, a control mixture of protein + EtOH was prepared to ascertain the effect of EtOH on the protein. The amount of EtOH in the final protein solution did not exceed 0.7% (v/v). Protein solutions were dialysed extensively before use in all experiments. In DAUDA binding experiments the concentration of DAUDA in the dialysing buffer was sufficient to ensure the equilibration of the DAUDA chemical potential on both sides of the dialysis membrane. The molecular weight cut-off of the dialysis membrane was 7 kDa, so the DAUDA could easily diffuse. Thus, the residual absorbance of free DAUDA in solution at 280 nm was matched.

The protein samples for density variation AUC experiments were prepared as follows: solvents containing 10–100% (v/v) D₂O were prepared by mixing appropriate volumes of the buffer made up in 100% D₂O with buffer made up in 100% H₂O. Concentrated protein stock solutions were then diluted to ~30 μM with the different density solvents. The addition of each component (the protein, 100% D₂O buffer and 100% H₂O buffer) was performed using an analytical balance (Mettler Toledo, Switzerland). Diluted protein solutions were then dialysed against 250 volumes of the corresponding D₂O:H₂O buffer for 15 h at 4 °C. The density of the dialysate was measured using an Anton Paar (Graz, Austria) precision densimeter.

Analytical ultracentrifugation

Sedimentation velocity (SV) and sedimentation equilibrium (SE) experiments were performed in a Beckman Coulter (Palo Alto, USA) Optima XL-I analytical ultracentrifuge in double-sector cells using absorbance optics at 280 nm for protein and at 335 nm to observe the protein-DAUDA complex. Interference detection was also used. This was especially important for rOv-FAR-1 and rCe-FAR-5 since these proteins have no tryptophan residues. Data were recorded at a rotor temperature of 4 °C. The partial specific volume \bar{v} of the proteins studied was calculated from their amino acid composition and corrected for temperature using Eq. (1) (Durchschlag 1986):

$$\bar{v}_T = \bar{v}_{298.15} + 4.25 \times 10^{-4}(T - 298.15) \quad (1)$$

The partial specific volumes and weight average molecular weights of the protein-DAUDA complexes were determined experimentally via SE using the density variation method (Edelstein and Schachman 1973), in which \bar{v} is obtained from the concentration dependence of the buoyant mass (M_b) on the solvent density (ρ) and the mass comes from extrapolation of the data to zero solvent density (see e.g. Fig. 5). SE scans were analysed with Beckman XL-A/XL-I software based on Microcal ORIGIN 6.0.

Equilibrium in SE experiments was attained after 33 h. The speeds of rotation were selected so that the value for the parameter σ [the reduced apparent molecular weight (Yphantis 1964; Yphantis and Waugh 1956)] was between 2 and 4 for each plausible oligomeric species. Thus, the speeds of rotation were 30,000 rpm and 35,000 rpm for rAs-NPA-1A, 24,000 rpm and 31,000 rpm for rOv-FAR-1 and 26,000 and 32,000 rpm for rCe-FAR-5.

SV experiments were carried out at 45,000 rpm. Experimental data were analysed with SEDFIT software (Schuck 2000), which performs direct fitting of the velocity profiles. Also, size-distribution analysis [$c(s)$ versus s] was performed using SEDFIT to describe the heterogeneity of the material moving in the AUC cell.

Small-angle X-ray scattering

Small-angle X-ray scattering (SAXS) experiments were carried out on Beamline BM26B at the European Synchrotron Radiation Facility (Grenoble, France) at an electron energy of 17 keV. Camera lengths of 4 m and 1.5 m were used to obtain small-angle and wide-angle data, respectively. The concentration of rAs-NPA-1A was 4.2 mg mL⁻¹ for the 4 m camera length and 8.5 mg mL⁻¹ for the shorter camera length. Samples of rCe-FAR-5 were run at the same concentration (6.5 mg mL⁻¹) for both camera lengths. For the X-ray wavelength of $\lambda = 0.728$ Å used, the momentum transfer interval of $0.01 < s < 0.5$ Å⁻¹ was covered, where $s = 4\pi \sin \theta / \lambda$ and 2θ is the scattering angle. Experiments were performed at 5 °C. The data were normalized to the intensity of the incident beam and corrected for the detector response. The scattering of the background (empty cell) and buffer were subtracted. SAXS data were acquired as a series of 15 s frames. The data in the first frame were compared with those in the last to check for the presence of aggregates induced by radiation damage. These would be clearly seen as an increase in the low-angle data as time progressed. This was not observed.

Bioinformatics-based structure prediction

High-resolution models for rAs-NPA-1A and rCe-FAR-5 were constructed on the basis of their functional and secondary structural homology with the ligand-binding domain (LBD) of the retinoic acid receptor RXR α (Bourguet et al. 1995). The secondary structures of rAs-NPA-1A and rCe-FAR-5 were predicted using the consensus of several algorithms run by the JPRED server (Cuff and Barton 1999; Cuff et al. 1998) (<http://www.compbio.dundee.ac.uk/Software/JPred/jpred.html>). This information, together with the predicted surface accessibility (also from JPRED), through-space separation of key residues [extrapolated from the atomic structure of the RXR α LBD, protein data bank (PDB) accession code 1LBD] and a multiple sequence alignment [generated using the program MULTALIN (Corpet 1988) (<http://prodes.toulouse.inra.fr/multalin/multalin.html>)] of each of the two target protein sequences, together with 21 (rAs-NPS-1A) or 20 (rCe-FAR-5) homologous sequences identified using the program PSI-BLAST (Altschul et al. 1997) (<http://www.ncbi.nlm.nih.gov/BLAST/>), was used as input for the distance geometry-based de novo protein modelling program DRAGON (Aszodi et al. 1995; Aszodi and Taylor 1994; Taylor and Aszodi 1994). DRAGON was run 20 times for each protein; the resultant structures were superimposed to check for reproducibility of the predicted structure and ranked according to bond score, non-bond score, restraint score and secondary structure score. The polypeptide backbone and side-chains of the highest ranking structure were built onto the predicted alpha carbon and side-chain centroid positions using the biopolymer module of the program InsightII (Molecular Simulations, now Accelrys, Cambridge, UK). The sedimentation coefficients and radii of gyration of these predicted structures were calculated using the program HYDROPRO (García de la Torre et al. 2000).

Results

Stoichiometry of the nematode proteins in unliganded form

Nematode fatty acid-binding proteins were examined for the presence of oligomeric species in solution via SV size-distribution analysis. The resultant distribution curves for the three proteins, each of which comprises a single narrow peak, are shown in Fig. 1A. Thus, rAs-NPA-1A, rCe-FAR-5 and rOv-FAR-1 are monodisperse in solution. The rAs-NPA-1A and rCe-FAR-5 peaks are centered about sedimentation coefficients of 1.1 S and 1.2 S, respectively, which would be typical sedimentation coefficients for globular proteins with masses equal to those of the proteins in monomeric form under the experimental conditions used, while the rOv-FAR-1 peak (centred at 1.8 S) corresponds to the mass of the

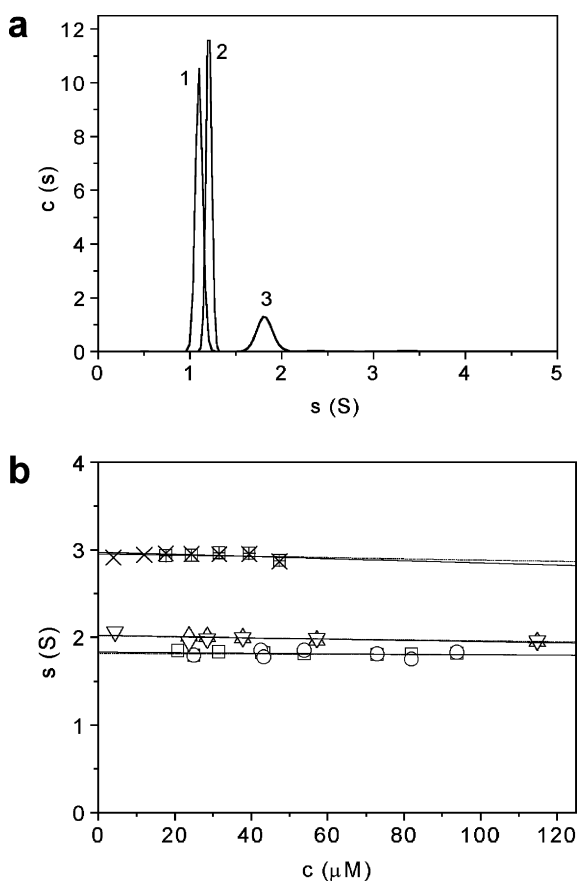


Fig. 1a, b Sedimentation velocity results for unliganded rAs-NPA-1A, rCe-FAR-5 and rOv-FAR-1; rotor speed = 45,000 rpm, temperature = 4 °C. **a** Size-distribution analysis for (1) rAs-NPA-1A (85 μ M), (2) rCe-FAR-5 (115 μ M) and (3) rOv-FAR-1. Data were recorded at 280 nm. **b** Concentration dependence of the sedimentation coefficient for rAs-NPA-1A (squares, absorbance at 280 nm; circles, interference), rCe-FAR-5 (up triangles, absorbance at 280 nm; down triangles, interference) and rOv-FAR-1 (plusses in squares, absorbance at 280 nm; crosses, interference). Data were fitted with least-squares best lines to obtain the values of $s_{20,w}^0$ reported in Table 1

dimeric form of the protein. The area beneath the peaks depends on the sample concentration and extinction coefficient (for absorbance data) or refractive index (for interference data) of the proteins. The rOv-FAR-1 peak has the lowest area since this protein has a very low extinction coefficient (it contains no tryptophan residues). Also, the concentration of rOv-FAR-1 was significantly lower than the other two proteins. The preliminary size-distribution analysis was confirmed by direct fitting of the SV profiles with a model for single non-interacting species in all three cases. The extrapolation of the $s(c)$ dependence to zero concentration is illustrated in Fig. 1B. Numerical values of the sedimentation coefficient agree well for both absorbance and interference data. The almost complete concentration independence for the data in Fig. 1B is consistent with the ideal monodisperse behaviour observed in SE studies (below).

On the basis of earlier differential scanning calorimetry data (Kennedy 2000; Kennedy et al. 1995a), rAs-NPA-1A was thought to be a dimer in solution. We sought to check this finding via sedimentation equilibrium and also to ascertain the association state of rCe-FAR-5 and rOv-FAR-1. Two rotor speeds were chosen, appropriate for observing a monomer-dimer equilibrium for all three proteins. The equilibrium solute distributions were firstly fitted with a model for a single ideal molecular species in order to obtain the weight average molecular masses (Fig. 2). Comparison

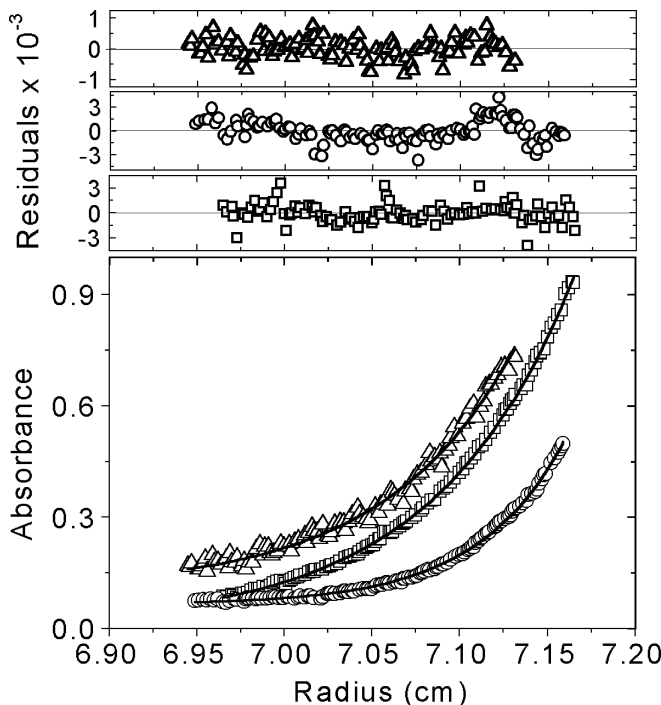


Fig. 2 An example of sedimentation equilibrium absorbance data (280 nm) for rAs-NPA-1A (39 μ M) at 30,000 rpm (squares), rCe-FAR-5 (67 μ M) at 26,000 rpm (circles) and rOv-FAR-1 (48 μ M) at 24,000 rpm (triangles). All data were fitted with a single species model. The residuals of the fit are shown in the upper panels

Table 1 Summary of AUC data. The results are distributed between the columns headed 280 nm or 335 nm depending on the wavelength at which the data were recorded

	$s_{20,w}^0$ (S)		M (kDa)		\bar{v} (mL g ⁻¹)		
	280 nm	335 nm	Calculated	280 nm	335 nm	280 nm	335 nm
	Calculated	Calculated	Calculated	Calculated	Calculated	Calculated	Calculated
rAs-NPA-1A	1.83 ± 0.01 ^a		15.160	15.1 ± 1.5 ^b	0.732	0.73 ± 0.15 ^b	
rAs-NPA-1A-DAUDA	1.85 ± 0.02 ^a	1.87 ± 0.04 ^a	15.629 ^c	15.8 ± 0.5 ^b	0.744	0.71 ± 0.06 ^b	0.71 ± 0.12 ^b
rAs-NPA-1A-EtOH	1.76 ± 0.02 ^a		15.629 ^c	13.6 ± 1.5 ^b		0.66 ± 0.21 ^b	
rCe-FAR-5	2.02 ± 0.01 ^a		22.658	20.5 ± 0.3 ^d	0.737		
rOv-FAR-1	2.95 ± 0.03 ^a		20.945	40.0 ± 0.7 ^d	0.730		
rOv-FAR-1-DAUDA	2.88 ± 0.10 ^a	2.88 ± 0.20 ^a	21.414 ^c	49.0 ± 7.5 ^b	0.739	0.75 ± 0.30 ^b	0.74 ± 0.48 ^b
rOv-FAR-1-EtOH	2.96 ± 0.10 ^a		21.414 ^c	40.1 ± 1.1 ^d			

^aDetermined via SV experiment; errors represent the quality of the linear extrapolation to infinite dilution of the sedimentation coefficient as a function of solute concentration

^bDetermined via SE density variation experiment; errors represent the quality of the linear extrapolation to zero density of the buoyant mass as a function of solvent density

^cAssuming 1:1 binding

^dDetermined via SE experiment; errors represent the quality of the weighted fit of the self-association model used by the Beckman XL-A XL-I Microcal ORIGIN 6.0 software

of these masses with those calculated for the three proteins from their amino acid compositions (Table 1) suggested that rAs-NPA-1A and rCe-FAR-5 were monomeric in solution while rOv-FAR-1 was a dimer. Therefore, the equilibrium data were also fitted with a monomer-dimer model but the fit with the single species model was more appropriate.

The partial specific volume \bar{v} of rAs-NPA-1A was measured by SE density variation. The results are in excellent agreement with the calculated value extrapolated to experimental temperature according to Eq. 1. It was important to take into account the temperature dependence of \bar{v} , since implausible molecular masses were obtained from the SE data when values for \bar{v} uncorrected for temperature were used. The errors incurred by failure to correct \bar{v} are particularly significant for low molecular weight proteins. The apparently large errors in the values of \bar{v} obtained by this method are the result of the combined errors in the slope and intercept of the experimental data extrapolation. In summary, SE data confirmed the results from the SV experiments and no dissociation processes were observed; all nematode proteins studied are single species: monomers in the case of rAs-NPA-1A and rCe-FAR-5, whereas rOv-FAR-1 is a dimer. All numerical results from AUC experiments are shown in Table 1.

Stoichiometry of the nematode proteins in ligand-bound form

The AUC study of the ligand-bound form of these proteins concentrated on rAs-NPA-1A and rOv-FAR-1 owing to their different oligomeric states (monomer and dimer, respectively). DAUDA (a fluorescent fatty acid analogue), used in these studies as a ligand for each of the nematode proteins, has a higher solubility in alcohol than the fatty acids naturally bound by the proteins. Thus we were able to minimize the quantity of ethanol in the experimental protein solutions. Nevertheless, protein-ethanol control samples in the absence of DAUDA were examined by AUC to check for the effect of the ethanol on the proteins. Since DAUDA absorbs significantly at 335 nm (Haughland 1996), we were able to use this wavelength to observe the sedimentation velocity/equilibrium of liganded protein while the absorbance at 280 nm and interference signal were used to measure the distribution of all the protein in the system. There are some complications arising in detection at 280 nm since DAUDA itself has some residual absorbance at this wavelength (Fig. 3). Therefore, it was critical to obtain an optical baseline before the SE data could be fitted correctly. Normally any mismatch between the sample and reference channel baselines in SE experiments can be accounted for by overspeeding of the rotor; however, these proteins have low molecular masses and so form equilibrium distributions even at higher rotor speeds. Instead, matching baselines were achieved by extensive

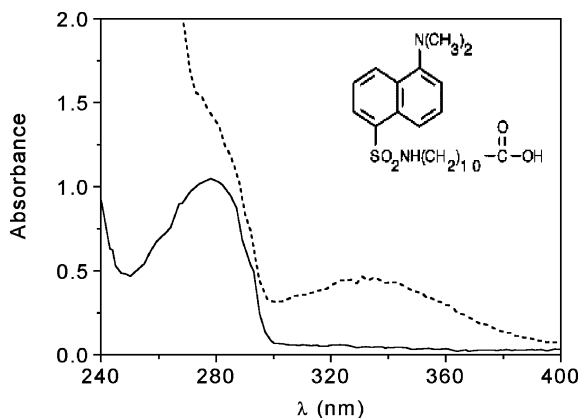


Fig. 3 Absorbance spectra for unliganded rAs-NPA-1A (full line) and a complex of rAs-NPA-1A-DAUDA (dashed line). The structural formula of DAUDA is shown in the insert

dialysis of the samples to equilibrate the amounts of DAUDA in the protein sample and solvent. The SE data treatment was thus performed using floating baselines for rAs-NPA-1A, whereas for the more massive protein (dimeric rOv-FAR-1) the floated baseline agreed well with the experimental baseline scan obtained at 45,000 rpm.

The size-distribution analysis plot for rAs-NPA-1A and rOv-FAR-1 complexed with DAUDA shows the peaks centred at the same sedimentation coefficients as for the unbound proteins (Fig. 4A) for both the 280 nm and 335 nm data sets. The area under the protein-EtOH control size-distribution peaks is lower than that for the corresponding protein-DAUDA samples because DAUDA contributes to absorbance at 280 nm. Again, this effect is especially apparent in the case of rOv-FAR-1 since it has a low extinction coefficient at 280 nm. The area under the peak for data obtained at 335 nm is smaller than for peaks from the 280 nm data since the DAUDA molar extinction coefficient is approximately half the usual protein extinction coefficient at 280 nm [$\epsilon_{335} = 4400 \text{ M}^{-1} \text{ cm}^{-1}$ (Haughland 1996)].

The concentration dependence of s for liganded rAs-NPA-1A is shown in Fig. 4B for two DAUDA stock concentrations (2.3 mM and 115 mM). When the lower stock concentration of DAUDA was used, it was necessary to add a higher volume of ligand solution to the protein sample. This resulted in a high final concentration of EtOH in the protein solution (up to 7% v/v). Under these conditions the sedimentation coefficient (observed at both 280 nm and 335 nm) decreases significantly at higher protein concentrations. Comparison of these data with the single data point for a rAs-NPA-1A control sample (in solvent of the same high EtOH concentration) indicates that it is the EtOH, and not the binding of DAUDA, that so significantly modifies the hydrodynamic properties of the protein. Interestingly, because the data points in Fig. 4B are identical for data acquired at 280 nm and 335 nm, rAs-NPA-1A remains active (i.e. able to bind DAUDA) even in the presence of a high EtOH concentration, and even though the shape

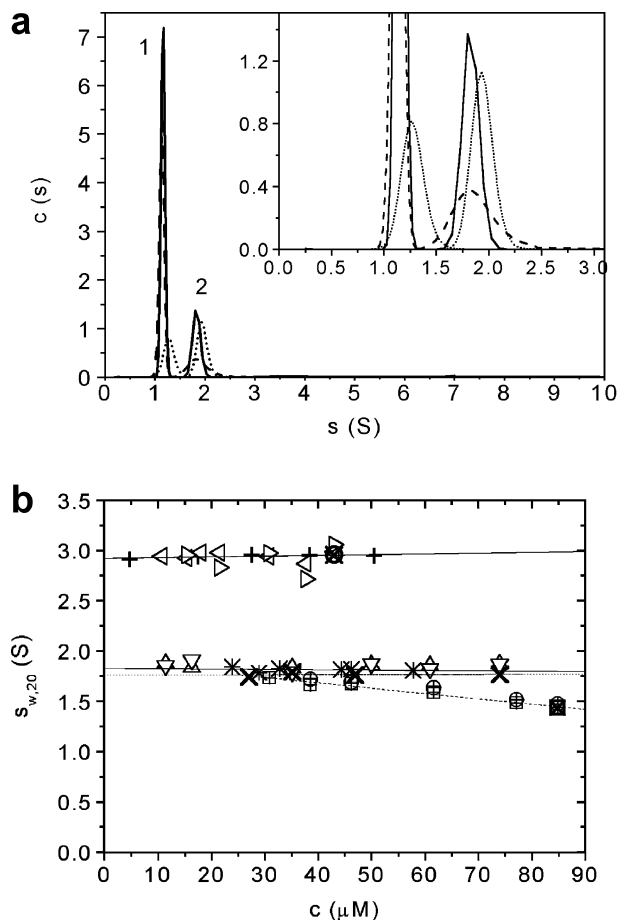


Fig. 4 **a** Size distribution analysis of sedimentation velocity data for (1) rAs-NPA-1A (73 μM) and (2) rOv-FAR-1 (43 μM); rotor speed = 45,000 rpm, temperature = 4 $^{\circ}\text{C}$: absorbance at 280 nm (full line), 335 nm (dotted line); EtOH (0.7% v/v) control, absorbance at 280 nm (dashed line). Rescaled data for low-intensity peaks are presented in the insert. **b** The concentration dependence of the sedimentation coefficient for rAs-NPA-1A-DAUDA in high-[EtOH] buffer [absorbance at 280 nm (pluses in squares), 335 nm (pluses in circles)], the control sample of rAs-NPA-1A in high-[EtOH] buffer [absorbance at 280 nm (crosses in squares)], rAs-NPA-1A-DAUDA in low-[EtOH] buffer [absorbance at 280 nm (up triangles), 335 nm (down triangles)], the control sample of rAs-NPA-1A in low-[EtOH] buffer [absorbance at 280 nm (crosses)], rOv-FAR-1-DAUDA in low-[EtOH] buffer [interference (left triangles), absorbance at 335 nm (right triangles)] and the control sample of rOv-FAR-1 in low-[EtOH] buffer (crosses in circles). The results for the unliganded proteins (from Fig. 1b) are shown for comparison: rAs-NPA-1A (asterisks) and rOv-FAR-1 (pluses)

of the protein is likely to have been altered by this solvent composition. In order then to observe the real effect of DAUDA, its stock concentration was increased to 115 mM so that the final concentration of EtOH in the protein solutions only reached a maximum of 0.7% (v/v). Under these conditions, the sedimentation coefficients (and thus solution conformation) of rAs-NPA-1A and rOv-FAR-1 do not change significantly upon binding DAUDA (Fig. 4B). Comparison of 0.7% EtOH control data with data for the unliganded proteins acquired in EtOH-free buffer confirms that the proteins are not sensitive to this low level of organic solvent.

The SE data for rAs-NPA-1A/rOv-FAR-1-DAUDA were fitted with a single species model. An increase in molecular mass consistent with the binding of one molecule of DAUDA per protein molecule was observed (Table 1). This is in accordance with the 1:1 stoichiometry determined earlier for this interaction (Kennedy 2000; Kennedy et al. 1995a, 1997; McDermott et al. 2001; Xia et al. 2000). However, the sensitivity of these SE experiments is not sufficient to rule out a lower affinity higher stoichiometric interaction which could result in the same weight average molecular mass.

The results of the SE density variation experiment for the DAUDA complexes of rAs-NPA-1A and rOv-FAR-1 are shown in Fig. 5. These data were collected at speeds of 35,000 rpm, which is appropriate for rAs-NPA-1A in monomeric form and 24,000 rpm for dimeric rOv-FAR-1. The \bar{v} value for rAs-NPA-1A-DAUDA determined in the D₂O/H₂O SE density variation experiment is the same as that calculated for the complex (0.744 mg mL⁻¹, Table 1). In this calculation, since we could not find an exact literature value for the DAUDA \bar{v} , we assumed that it was similar to that for natural fatty acids (1.14 mL g⁻¹) (Duel 1951). At the concentrations of DAUDA used in these studies, micelles were not formed since their presence would have been apparent as a significant, floating species in the sedimentation velocity and equilibrium experiments performed on solutions of protein in the presence of DAUDA, and this was not observed. A decreased \bar{v} for rAs-NPA-1A was obtained by density variation SE in the presence of EtOH (0.66 mL g⁻¹, Table 1); however, within the large error limits of the technique this decrease is not significant. Similarly, the increase in \bar{v} of rAs-NPA-1A to 0.71 mL g⁻¹ upon binding of DAUDA is not truly significant. The solvent density variation SE experiment on the rOv-FAR-1-DAUDA complex has shown the expected increase in \bar{v} for the protein-fatty acid complex (Table 1). The accu-

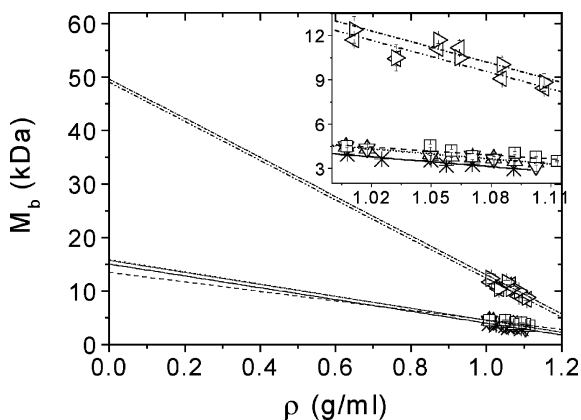


Fig. 5 Results of the sedimentation equilibrium density variation experiment. The dependence of the buoyant mass on solvent density for rAs-NPA-1A (rotor speed = 35,000 rpm): unliganded rAs-NPA-1A in the absence (*asterisks*) and presence (*squares*) of 0.7% (v/v) ethanol; rAs-NPA-1A-DAUDA [absorbance at 280 nm (*up triangles*) and 335 nm (*down triangles*)] and for rOv-FAR-1-DAUDA (rotor speed = 24,000 rpm): absorbance at 280 nm (*left triangles*) and 335 nm (*right triangles*)

racy of this experiment was also limited by the slight variation in protein concentration at each solvent density. However, given the minimal concentration dependence of s for rOv-FAR-1, the mass is also unlikely to depend significantly upon concentration.

Shape modelling from SAXS experiments

Indirect transformation of the reciprocal space scattering data [$I(Q)$ versus Q , Fig. 6A and B] into a real space $P(r)$ versus r function (Fig. 6C) was performed using the program GNOM (Svergun 1992) for SAXS data. The analysis provides an estimate of the maximum dimension of the particle (D_{\max}). As $P(r)$ corresponds to the distribution of distances r between any two scattering elements within a particle, it also offers an alternative calculation of the radius of gyration R_g . Low-resolution models were restored from the experimental data using the programs DAMMIN (Svergun 1999), to determine a dummy atom model (DAM) and GASBOR (Svergun et al. 2001), which generates a dummy residue model (DRM). The details of these methods are described elsewhere (Svergun 1999; Svergun et al. 2001). Parameters determined from the scattering curves are listed in Table 2.

The shapes from GASBOR and Dammin were slightly different but in general represented the same kind of structures and had similar overall dimensions. The rAs-NPA-1A DAM has dimensions of 46×30×32 Å and a resolution of 21 Å, while the DRM from the same data has dimensions of 46×26×28 Å approximately and a resolution of 19 Å. For rCe-Far-5 we have found the following dimensions: for the DAM, 52×31×34 Å, and for the DRM, 53×32×33 Å. Thus, the structures resolved by SAXS experiments are globular although slightly elongated and slightly flattened (Fig. 8).

To confirm that the rAs-NPA-1A and rCe-FAR-5 DAMs were consistent with the hydrodynamic data obtained from SV studies, we calculated their sedimentation coefficients using the program HYDROPRO (García de la Torre et al. 2000), which calculates hydrodynamic parameters for shell models that it constructs from user input files of model coordinates. In order to ensure that the primary hydrodynamic model was filled with overlapping spheres, we selected a radius for the atomic elements (the AER) greater than the radius of the dummy atoms, which are hexagonally packed in the DAM and thus do not overlap and do not properly fill the model volume (see Ackerman et al. 2003; Scott et al. 2002 for earlier examples of this approach). To determine the amount of bound water for each protein DAM, the anhydrous protein volume ($V_A = M\bar{v}/N_A$, where M is the protein mass and N_A is Avogadro's number) was subtracted from the Porod volume (the excluded volume of the hydrated protein) of the DAM. The DAM packing radius was 2.0 Å; therefore an AER of 2.6 Å was used for the HYDROPRO model. The range of shell (or mini) bead sizes was 0.9–2.0 Å. Five iterations for extrapolation to the hydration shell were used. These parameters yielded

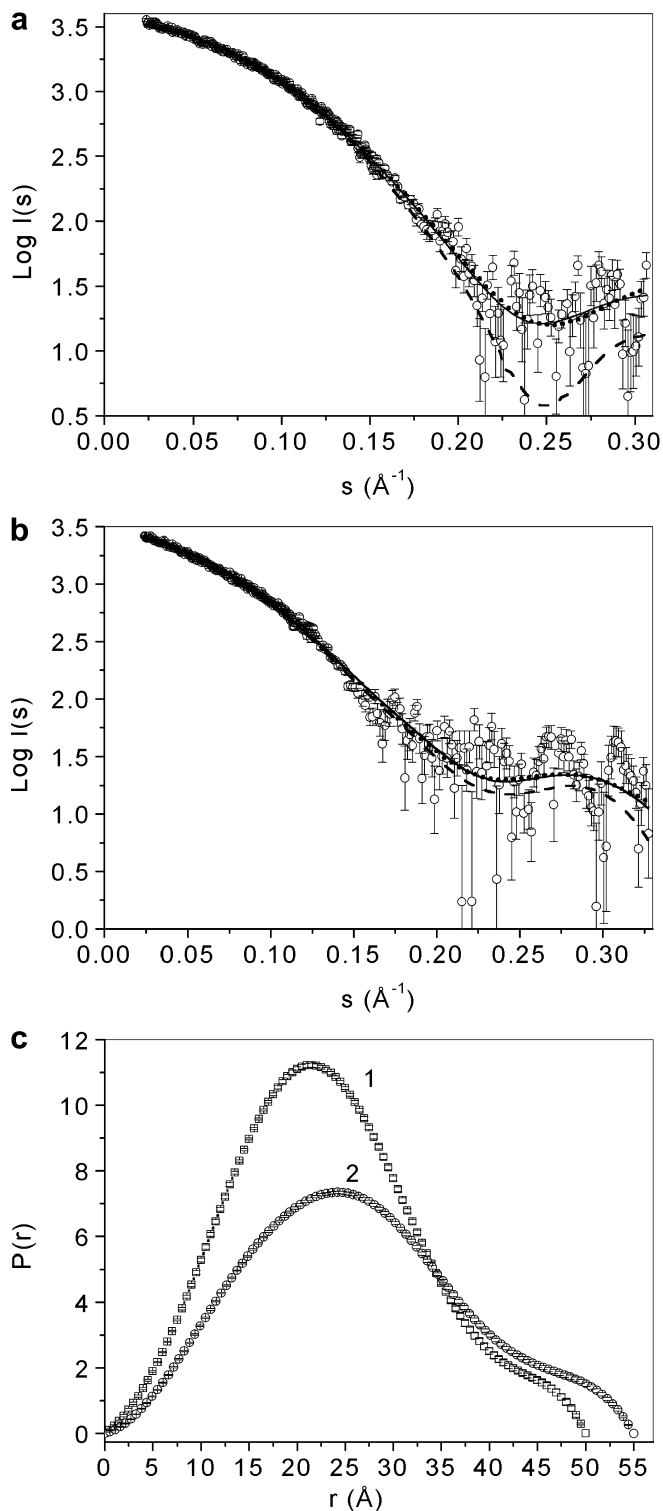


Fig. 6 Results of the SAXS experiment for (a) rAs-NPA-1A and (b) rCe-FAR-5. Scattering curve (circles), GNOM fit (full line), Dammin fit (dashed line) and GASBOR fit (dotted line). c The distance distribution function for (1) rAs-NPA-1A and (2) rCe-FAR-5 obtained by fitting the experimental data in a and b with GNOM (Svergun 1992)

models with effective hydrodynamic volumes that were in good agreement with those of the DAMs. For example, the effective hydrodynamic volume of the rAs-NPA-1A model was 40.1 nm³, while the volume of the anhydrous protein (calculated from its mass and partial specific volume) is 18.5 nm³. Thus the effective hydration is 0.87 g water/g protein (García de la Torre 2001). This is significantly higher than the hydration of the DAM (0.54 g water/g protein) and consequently gives a calculated sedimentation coefficient (1.50 S) somewhat lower than the experimental value (1.82 S). Following a similar procedure (using an AER of 3.0 Å and a range of mini-bead sizes from 1.0 to 2.5 Å) we obtained $s = 1.99$ S for the rCe-Far-5 DAM (compared with an experimental value of 2.02 S). This model has an effective hydration of 0.33 g water/g protein (compared with the hydration of the DAM, 0.23 g water/g protein). Thus the model for rCe-FAR-5 restored from SAXS data is completely consistent with the hydrodynamic data. However, this is not so for the rAs-NPA-1A model.

Structural models of rAs-NPA-1A and rCe-FAR-5 based upon functional and secondary structural homology with the ligand-binding domain of the retinoic acid receptor, RXR α

At present there are no high-resolution structures for fatty acid-binding proteins from nematodes. This precludes prediction of an atomic structure for As-NPA-1A and Ce-FAR-5 using conventional homology modelling programs such as Swiss-Model (Guex et al. 1999; Peitsch 1996). Our initial attempts to construct high-resolution homology models for rAs-NPA-1A based on its nearest primary sequence homologues and earlier predictions about its α -helical content (Kennedy et al. 1995a) proved unsatisfactory: the resultant models generated by DRAGON (Aszódi et al. 1995; Aszódi and Taylor 1994; Taylor and Aszódi 1994) contained non-protein-like features and had overall shapes at odds with the hydrodynamic data measured. Instead, we based our models upon homology of a different sort. Recognizing that JPRED (Cuff and Barton 1999; Cuff et al. 1998) predicted seven α -helices in rAs-NPA-1A and 10 in rCe-FAR-5 and that the ligand-binding domain (LBD) of the retinoic acid receptor RXR α (Bourguet et al. 1995) also comprises 10 α -helices, we elected to use the fold of the RXR α LBD as a template for folding rAs-NPA-1A and rCe-FAR-5. The alignment of the secondary structure of RXR α LBD with those predicted for rAs-NPA-1A and rCe-FAR-5 (Fig. 7) was obtained by introducing gaps in the MULTALIN (Corpet 1988) alignment of the pairs so that the helices were brought into alignment. Corresponding helices (e.g. helix 1 of RXR α LBD and rCe-FAR-5) tend to have different lengths; they were positioned relative to each other on the basis of obtaining maximal sequence homology without introduction of gaps along the extent of the helix pair. Asterisks in Fig. 7 identify the RXR α LBD



Fig. 7 Alignments of the helices in the ligand-binding domain of RXR α with the predicted helices in the recombinant forms of As-NPA-1A and Ce-FAR-5 (including extra N-terminal amino acids encoded by the expression vector or the 6 \times His affinity tag). *Asterisks* indicate the RXR α LBD residues used to define the relative positions of the corresponding rAs-NPA-1A and rCe-FAR-5 helix extremities in one of the DRAGON input files

residues used to define the positions of the corresponding rAs-NPA-1A and rCe-FAR-5 helix extremities in DRAGON. An additional constraint was introduced into the As-NPA-1A model so that the two cysteines it contains were maintained at a separation appropriate for disulfide bridging (McDermott et al. 2001).

The predicted spatial structures for rAs-NPA-1A and rCe-FAR-5 are superimposed on the dummy residue models (DRMs) restored from SAXS data using the program GASBOR (Svergun et al. 2001) for rAs-NPA-1A and rCe-FAR-5 in Fig. 8. Superimposition was achieved with the program SUPCOMB (Kozin and Svergun 2001). Examination of the log file generated by

the ranking process for the DRAGON models reveals that those ranked lower score less because an increasing number of α -carbons make too many bonds with other α -carbons. Given the uncertainties in this entire process, we view this as a relatively trivial violation. Therefore, lower ranking DRAGON models were also assessed for their agreement with the DRMs: the second ranked rCe-FAR-5 DRAGON model fits the DRM slightly better than the first; this is not so for rAs-NPA-1A.

Hydrodynamic quantities for the predicted models (the cysteine-bridged model for As-NPA-1A and the second-ranked Ce-FAR-1 model) and the DAMs are listed in Table 2. The predicted structures are a little more compact than the DRM/DAMs for both proteins, more so in the case of rCe-FAR-5. We ascribe this to the restrictions placed upon the DRAGON fold by the predicted surface accessible residues. This increased compactness is reflected in the decreased radius of gyration calculated by HYDROPRO for both the crude DRAGON model and the backbone/side-chain model for both proteins compared with the measured value (listed in the DAM column). The agreement between the sedimentation coefficients is variable: for rAs-NPA-1A the DAM and the crude DRAGON model agree well, while the backbone/side-chain model is in better agreement with the experimentally determined value. In the case of rCe-FAR-5, all models agree with each other and with the experimentally determined value. Therefore the structures modelled by DRAGON are consistent with the envelopes described by the DAMs, which in turn are consistent with the hydrodynamic data, to within reasonable limits.

Discussion

We report on the quaternary arrangements of one form of the NPA polyproteins and two forms of FAR

Table 2 Hydrodynamic parameters calculated for the dummy atom models (DAMs) of rAs-NPA-1A and rCe-FAR-5 compared with experimental values and values calculated for the high-resolution predicted structures

	rAs-NPA-1A				rCe-FAR-5			
	DAM	HYDROPRO			DAM	HYDROPRO		
		DAM	DRAGON model ^d	Side-chain model ^d		DAM	DRAGON model ^e	Side-chain model ^e
AER (Å) ^a	2.0	2.6	5.0	3.2	2.3	3.0	5.0	3.5
Volume (nm ³) ^b	32.4	40.1	33.7	23.0	36.7	40.6	50.9	34.8
Volume of water (nm ³)	13.7	22.8	15.16	4.5	8.6	12.5	22.8	6.7
$s_{20,w}^0$ (S) ^c		1.47	1.53	1.67		1.99	2.00	2.20
δ (g water/g protein)	0.54	0.87	0.60	0.18	0.23	0.33	0.61	0.18
R_g (Å)	17.6	18.5	17.8	16.4	20.0	20.6	19.0	17.0

^aAtomic element radius: the radius of bead used to construct the primary hydrodynamic particle in HYDROPRO (Garcia de la Torre et al. 2000)

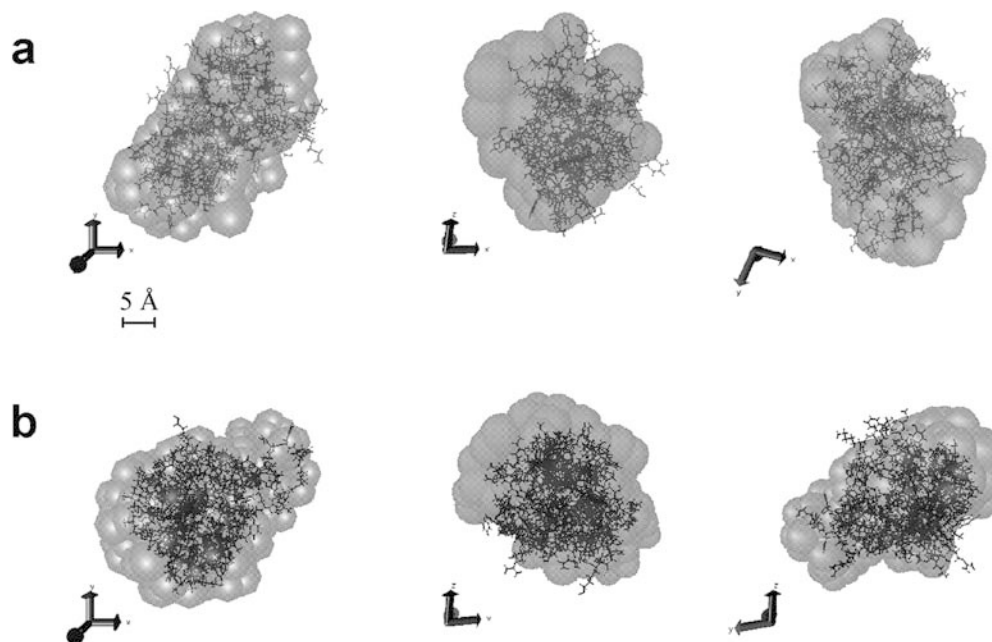
^bHydrated volume: it is necessary to increase the radius of spheres in the DAM above the packing radius in order to calculate hydrodynamic parameters using HYDROPRO (see e.g. Ackerman et al. 2003); thus the hydration increases

^cFor rAs-NPA-1A the anhydrous volume is 18.5 nm³ and the experimental sedimentation coefficient ($s_{20,w}^0$) is 1.83 S. For rCe-FAR-5 the anhydrous volume is 28.1 nm³ and the experimental sedimentation coefficient is 2.02 S

^dFor the cysteine-bridged model

^eFor the second-ranked DRAGON model

Fig. 8 Orthogonal views of dummy residue models for (a) rAs-NPA-1A and (b) rCe-FAR-5 superimposed upon the high-resolution predicted structures



protein from nematodes, show that this is unaltered by ligand (fatty acid) binding, and present low-resolution structural models based on a combination of hydrodynamics and SAXS. A recombinant form of the As-NPA-1A polyprotein unit of *A. suum* is shown to exist as a monomer in solution, as was the recombinant form of the Ce-FAR-5 protein of *C. elegans*, but not the recombinant Ov-FAR-1 counterpart of *O. volvulus*, which instead exists as a tight dimer.

Our previous calorimetry studies of a recombinant NPA protein (Xia et al. 2000), and material purified directly from the parasite *A. suum* (Kennedy et al. 1995a), indicated that the protein possibly formed a dimer, based on the ratio between ΔH_{cal} and ΔH_{vH} derived from analysis of the thermal transition curve, which gave an unusually high T_m (≥ 90 °C). The hydrodynamics data are, however, the more definitive, providing robust evidence for a monomer state, and none for higher order arrangements. Moreover, the new data were collected at protein concentrations of a similar magnitude to those used in the calorimetry studies, so the disparity is not due to monomer-dimer equilibrium effects. In addition to being released by living parasites, the protein is also present at high concentrations in the pseudocoelomic fluid (analogous to our blood or the haemolymph of insects) of the worms (Christie et al. 1990; Kennedy and Qureshi 1986; Kennedy et al. 1989). Such a small protein would likely be lost through an excretory system (and thereby found in “secretions”), as it would be in vertebrates, unless it formed part of a larger entity, as does human plasma retinol-binding protein with transthyretin (Flower 1996). No evidence has, however, arisen to date that the NPAs associate with other proteins in vivo. In the present study we analysed a homogeneous preparation of recombinant As-NPA-1A, and it remains possible that this unit of the polyprotein might associate

with units of different amino acid sequence. In the case of the As-NPA array, all but one of the 10 units known are very similar, and so are unlikely to associate, but one (As-NPA-1B) possesses only 49% amino acid identity with the A unit, although it has very similar lipid-binding properties (Moore et al. 1999). In species such as the cattle lungworm *D. viviparus* and *C. elegans*, the arrays contain units that differ dramatically in sequence (Blaxter 1998; Britton et al. 1995), so there is more scope for heterodimerization, and this awaits testing.

All NPA units described so far possess a Trp in an absolutely conserved position, and As-NPA-1A possesses a single Trp. The fluorescence emission spectrum of this Trp (Trp15) is unusual in being strongly blue-shifted (to ~ 315 nm) (Kennedy et al. 1995a; McDermott et al. 2001), indicative of removal from solvent water to an unusual degree. However, our original computer-based predictions for the disposition of Trp15 placed it on the solvent-exposed face of an amphipathic helix (Kennedy et al. 1995a), which should result in a red-shifted emission spectrum (Eftink and Ghiron 1976, 1984). In the present study, despite being identified in the DRAGON input files as a buried residue, the tryptophan is nonetheless also located on the surface of the final folded protein model, although now situated at the break between helices (at the end of helix 1). Moreover, site-directed mutagenesis of Trp15 to an Arg resulted in no detectable change in the ligand-binding properties of the protein (although it did compromise its resistance to denaturation by guanidine hydrochloride), and ligand binding made no difference to its emission spectrum (Kennedy et al. 1995a; McDermott et al. 2001). So Trp15 is probably not involved in the binding site. One possible explanation for these data was that Trp15 is involved in a protein-protein interface in a homodimer, but our new analysis refutes this possibility also.

Nematodes appear only to possess a single gene for the NPAs per haploid genome (Blaxter 1998), but recent analysis has identified eight genes encoding FAR-like proteins in the *C. elegans* genome (Garofalo et al. 2003). These fall into three groups according to sequence similarity, designated groups A, B and C. The Ov-FAR-1 protein examined here falls into group A, as do all other FAR proteins described from parasites to date, and Ce-FAR-5 falls into group B. All of the group A and B proteins have similar fatty acid- and retinol-binding activities, and Ce-FAR-5 is unusual in that it causes an unprecedented blue shift in the fluorescence emission spectrum of DAUDA (Garofalo et al. 2003).

The two FAR proteins examined here differ in that rOv-FAR-1 (group A) was found to form a tight homodimer, whereas rCe-FAR-5 (group B) existed only as a monomer. We cannot explain this different behaviour, and there is no structural information from any FAR protein yet available to shed light. The only difference between the two primary structures that may be pertinent is that Ov-FAR-1 exhibits a high probability of forming coiled coils (Kennedy et al. 1997), whereas Ce-FAR-5 does not (unpublished), according to predictions made by the Coils (Berger et al. 1995) and Paircoil (Lupas et al. 1991) programs. Coiled-coil regions of proteins can be involved in protein-protein interactions, but, equally, some coiled-coil proteins exhibit no such interactions (Burkhard et al. 2001).

What is clear, however, is that neither NPA nor FAR proteins require the formation of higher order complexes in order to bind ligands, since both As-NPA-1A and Ce-FAR-1 bind ligands as monomers, and ligand binding does not affect their monomeric status. In terms of their role in parasitism, an NPA secreted by a parasite will diffuse away in the form of a small ~14.4 kDa protein, whereas Ov-FAR-1 will be a larger ~40 kDa entity. Whether such differences are biologically meaningful in the performance of the two proteins remains to be seen.

Acknowledgements A.S.S. is supported by a grant from the Wellcome Trust to O.B., M.W.K. and others (05606/Z/99/Z), and we are also indebted to the Wellcome Trust for their support of this project through other grants to M.W.K. The work was also supported by an ESRF beam time award (LS-1872) and by the Universities of Salford and Nottingham. Special thanks to the team of Dr. Wim Bras (ESRF, Grenoble, particularly Dr. Marc Malfois) for assistance in the SAXS experiments and Marcelo Nöllmann for help in SAXS data treatment.

References

- Ackerman CJ, Harnett MM, Harnett W, Svergun DI, Byron O (2003) 19 Å solution structure of the filarial nematode immunomodulatory protein, ES-62. *Biophys J* 84:489–500
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Aszodi A, Taylor WR (1994) Folding polypeptide alpha-carbon backbones by distance geometry methods. *Biopolymers* 34:498–505
- Aszodi A, Gradwell MJ, Taylor WR (1995) Global fold determination from a small number of distance restraints. *J Mol Biol* 251:308–326
- Barrett J, Saghir N, Timanova A, Clarke K, Brophy PM (1997) Characterisation and properties of an intracellular lipid-binding protein from the tapeworm *Moniezia expansa*. *Eur J Biochem* 250:269–275
- Berger B, Wilson DB, Wolf E, Tonchev T, Milla M, Kim PS (1995) Predicting coiled coils by use of pairwise residue correlations. *Proc Natl Acad Sci USA* 92:8259–8263
- Blaxter M (1998) *Caenorhabditis elegans* is a nematode. *Science* 282:2041–2046
- Bourguet W, Ruff M, Chambon P, Gronemeyer H, Moras D (1995) Crystal structure of the ligand binding domain of the human nuclear receptor RXR α . *Nature* 375:377–382
- Britton C, Moore C, Gilleard JS, Kennedy MW (1995) Extensive diversity in repeat unit sequences of the cDNA encoding the polyprotein antigen/allergen from the bovine lungworm *Dictyocaulus viviparus*. *Mol Biochem Parasitol* 72:77–88
- Burkhard P, Stetefeld J, Strelkov SV (2001) Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol* 11:82–88
- Christie JF, Dunbar B, Davidson I, Kennedy MW (1990) N-terminal amino acid sequence identity between a major allergen of *Ascaris lumbricoides* and *Ascaris suum*, and MHC-restricted IgE responses to it. *Immunology* 69:596–602
- Coe NR, Bernlohr DA (1998) Physiological properties and functions of intracellular fatty acid-binding proteins. *Biochim Biophys Acta* 1391:287–306
- Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 16:10881–10890
- Cuff JA, Barton GJ (1999) Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins Struct Funct Genet* 34:508–519
- Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics* 14:892–893
- Duel HJ (1951) *The lipids: their chemistry and biochemistry*. Interscience, New York
- Durchschlag H (1986) Specific volumes of biological macromolecules and some other molecules of biological interest. In: Hinz H-J (ed) *Thermodynamic data for biochemistry and biotechnology*. Springer, Berlin Heidelberg New York, pp 45–128
- Edelstein SJ, Schachman HK (1973) Measurement of partial specific volume by sedimentation equilibrium in H₂O-D₂O solutions. *Methods Enzymol* 27:82–98
- Eftink MR, Ghiron CA (1976) Exposure of tryptophanyl residues in proteins: quantitative determination by fluorescence quenching studies. *Biochemistry* 15:672–679
- Eftink MR, Ghiron CA (1984) Indole fluorescence quenching studies on proteins and model systems: use of the efficient quencher succinimide. *Biochemistry* 23:3891–3899
- Flower DR (1996) The lipocalin protein family: structure and function. *Biochem J* 318:1–14
- García de la Torre J (2001) Hydration from hydrodynamics. General considerations and applications of bead modelling to globular proteins. *Biophys Chem* 93:159–170
- García de la Torre J, Huertas ML, Carrasco B (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J* 78:719–730
- Garofalo A, Kläger SL, Rowlinson MC, Nirmalan N, Klion A, Allen JE, Kennedy MW, Bradley JE (2002) The FAR proteins of filarial nematodes: secretion, glycosylation and lipid binding characteristics. *Mol Biochem Parasitol* 122:161–170
- Garofalo A, Rowlinson M-C, Ngwa A, Hughes JM, Kelly SM, Price NC, Cooper A, Watson DG, Kennedy MW, Bradley JE (2003) The FAR protein family of the nematode *Caenorhabditis elegans*. Differential lipid binding properties, structural characteristics and developmental regulation. *J Biol Chem* 278:8065–8074

- Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino-acid sequence data. *Anal Biochem* 182:319–326
- Guex N, Diemand A, Peitsch MC (1999) Protein modelling for all. *Trends Biol Sci* 24:364–367
- Haughland RP (1996) Handbook of fluorescent probes and research chemicals, 6th edn. Molecular Probes, Eugene, Ore., USA
- Janssen D, Barrett J (1995) A novel lipid-binding protein from the cestode *Moniezia expansa*. *Biochem J* 311:49–57
- Kennedy MW (2000) The polyprotein lipid binding proteins of nematodes. *Biochim Biophys Acta* 1476:149–164
- Kennedy MW (2001) Structurally novel lipid-binding proteins. In: Kennedy MW, Harnett W (eds) Parasitic nematodes. Molecular biology, biochemistry, and immunology. CABI, Wallingford, UK, pp 309–330
- Kennedy MW, Qureshi F (1986) Stage-specific secreted antigens of the parasitic larval stages of the nematode *Ascaris*. *Immunology* 58:515–522
- Kennedy MW, Qureshi F, Fraser EM, Haswell-Elkins MR, Smith HV (1989) Antigenic relationships between the surface-exposed, secreted and somatic materials of the nematode parasites *Ascaris lumbricoides*, *Ascaris suum*, and *Toxocara canis*. *Clin Exp Immunol* 75:493–500
- Kennedy MW, Brass A, McCrudden AB, Price NC, Kelly SM, Cooper A (1995a) The ABA-1 allergen of the parasitic nematode *Ascaris suum*: fatty acid and retinoid binding function and structural characterisation. *Biochemistry* 34:6700–6710
- Kennedy MW, Britton C, Price NC, Kelly SM, Cooper A (1995b) The DvA-1 polyprotein of the parasitic nematode *Dictyocaulus viviparus*: a small helix-rich lipid binding protein. *J Biol Chem* 270:19277–19281
- Kennedy MW, Garside LH, Goodrick LE, McDermott L, Brass A, Price NC, Kelly SM, Cooper A, Bradley JE (1997) The Ov20 protein of the parasitic nematode *Onchocerca volvulus*. A structurally novel class of small helix-rich retinol-binding proteins. *J Biol Chem* 272:29442–29448
- Kozin MB, Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* 34:33–41
- Lerche MH, Poulsen FM (1998) Solution structure of barley lipid transfer protein complexed with palmitate. Two different binding modes of palmitate in the homologous maize and barley nonspecific lipid transfer proteins. *Protein Sci* 7:2490–2498
- Lupas A, Van Dyke M, Stock J (1991) Predicting coiled coils from protein sequences. *Science* 252:1162–1164
- MacGregor RB, Weber G (1986) Estimation of the polarity of the protein interior by optical spectroscopy. *Nature* 319:70–73
- McDermott L, Moore J, Brass A, Price NC, Kelly SM, Cooper A, Kennedy MW (2001) Mutagenic and chemical modification of the ABA-1 allergen of the nematode *Ascaris*: consequences for structure and lipid-binding properties. *Biochemistry* 41:9918–9926
- Moore J, McDermott L, Price NC, Kelly SM, Cooper A, Kennedy MW (1999) Sequence-divergent units of the ABA-1 polyprotein array of the nematode *Ascaris suum* have similar fatty-acid- and retinol-binding properties but different binding-site environments. *Biochem J* 340:337–343
- Peitsch MC (1996) ProMod and Swiss-Model: internet-based tools for automated comparative protein modelling. *Biochem Soc Trans* 24:274–279
- Prior A, Jones JT, Blok VC, Beauchamp J, McDermott L, Cooper A, Kennedy MW (2001) A surface-associated retinol- and fatty acid-binding protein (Gp-FAR-1) from the potato cyst nematode *Globodera pallida*: lipid binding activities, structural analysis and expression pattern. *Biochem J* 356:387–394
- Rothnagel JA, Steinert PM (1990) The repeating structure of the units for mouse filaggrin and a comparison of the repeating units. *J Biol Chem* 265:1862–1865
- Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys J* 78:1606–1619
- Scott DJ, Grossmann JG, Tame JRH, Byron O, Wilson KS, Otto BR (2002) Low resolution structure of the apo form of *Escherichia coli* haemoglobin protease Hbp. *J Mol Biol* 315:1179–1187
- Storch J, Thumser AE (2000) The fatty acid transport function of fatty acid-binding proteins. *Biochim Biophys Acta* 1486:28–44
- Svergun DI (1992) Determination of the regularisation parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25:495–503
- Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76:2879–2886
- Svergun DI, Petoukhov MV, Koch MHJ (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* 80:2946–2953
- Taylor WR, Aszodi A (1994) Building protein folds using distance geometry: towards a general modelling and prediction method. In: Merz JM Jr, LeGrand SM (eds) The protein folding problem and tertiary structure prediction. Birkhäuser, Boston, pp 165–192
- Xia Y, Spence HJ, Moore J, Heaney N, McDermott L, Cooper A, Watson DG, Mei B, Komuniecki R, Kennedy MW (2000) The ABA-1 allergen of *Ascaris lumbricoides*: sequence polymorphism, stage and tissue-specific expression, lipid binding function, and protein biophysical properties. *Parasitology* 120:211–224
- Yphantis DA (1964) Equilibrium ultracentrifugation of dilute solutions. *Biochemistry* 3:297–317
- Yphantis DA, Waugh DF (1956) Ultracentrifugal characterisation by direct measurement of activity. I. Theoretical. *J Phys Chem* 60:623–635