

Diversity in the structures and ligand-binding sites of nematode fatty acid and retinol-binding proteins revealed by Na-FAR-1 from *Necator americanus*

M. Florencia Rey-Burusco*, Marina Ibáñez-Shimabukuro*, Mads Gabrielsen†, Gisela R. Franchini*, Andrew J. Roe†, Kate Griffiths‡, Bin Zhan§, Alan Cooper||, Malcolm W. Kennedy‡¶¹, Betina Córscico* and Brian O. Smith¶¹

*Instituto de Investigaciones Bioquímicas de La Plata, CONICET-UNLP, Facultad de Ciencias Médicas, calles 60 y 120, 1900-La Plata, Argentina

†Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow G12 8QQ, U.K.

‡Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow G12 8QQ, U.K.

§National School of Tropical Medicine and Department of Pediatrics, Section of Tropical Medicine, Baylor College of Medicine, Houston, TX 20031, U.S.A.

||School of Chemistry, University of Glasgow, Glasgow G12 8QQ, U.K.

¶Institute of Molecular, Cell & Systems Biology, University of Glasgow, Glasgow G12 8QQ, U.K.

Fatty acid and retinol-binding proteins (FARs) comprise a family of unusual α -helix rich lipid-binding proteins found exclusively in nematodes. They are secreted into host tissues by parasites of plants, animals and humans. The structure of a FAR protein from the free-living nematode *Caenorhabditis elegans* is available, but this protein [C. *elegans* FAR-7 (Ce-FAR-7)] is from a subfamily of FARs that does not appear to be important at the host/parasite interface. We have therefore examined [*Necator americanus* FAR-1 (Na-FAR-1)] from the blood-feeding intestinal parasite of humans, *N. americanus*. The 3D structure of Na-FAR-1 in its ligand-free and ligand-bound forms, determined by NMR (nuclear magnetic resonance) spectroscopy and X-ray crystallography respectively, reveals an α -helical fold similar to Ce-FAR-7, but Na-FAR-1 possesses a larger and more complex internal ligand-binding cavity and an additional C-terminal α -helix. Titration

of *apo*-Na-FAR-1 with oleic acid, analysed by NMR chemical shift perturbation, reveals that at least four distinct protein–ligand complexes can be formed. Na-FAR-1 and possibly other FARs may have a wider repertoire for hydrophobic ligand binding, as confirmed in the present study by our finding that a range of neutral and polar lipids co-purify with the bacterially expressed recombinant protein. Finally, we show by immunohistochemistry that Na-FAR-1 is present in adult worms with a tissue distribution indicative of possible roles in nutrient acquisition by the parasite and in reproduction in the male.

Key words: fatty acid-binding protein, *Necator americanus*, nematode, nuclear magnetic resonance (NMR), parasite, protein structure, retinol-binding protein, X-ray.

INTRODUCTION

Lipids such as acids (FAs) and retinoids are relatively insoluble in water, can be susceptible to oxidation and are potentially damaging to membranes in their free form. Consequently, they are usually transported within proteins or protein-mediated lipid aggregates. Examples of vertebrate FA and retinoid transporter proteins include serum albumins (~64 kDa), which bind a range of compounds, lipocalins (~20 kDa), which are found in many secretions and the proteins of the cytoplasmic FA-binding/cellular retinol-binding/cellular retinoic acid-binding protein (FABP/CRBP/CRABP) family (~14 kDa) that are confined to the cytoplasm, where they may bind FAs, retinol or retinoic acid, depending on the isoform. These proteins are variously involved in transport and storage of lipids

and also in the delivery of small signalling lipids to their destinations [1].

Nematodes, both free-living and parasitic to plants and animals, exhibit several additional types of lipid-binding proteins that are not found in other phyla. Prominent examples include the nematode polyprotein antigens (NPAs), such as ABA-1, whose structure has previously been solved [2]. These are synthesized as large polypeptide precursors that are post-translationally processed down to multiple copies of small lipid-binding proteins of ~14 kDa [3–7]. Other examples include the nemFABPs (nematode FABPs), which are similar to the intracellular FABP/CRBP/CRABP family of proteins, but in nematodes have structural modifications found in no other group of animals and are not confined to the cytosol [8,9]. The subject of the present paper is an unusual class of lipid-binding protein, the FA and retinol-binding proteins (FARs). These occur in several

Abbreviations: Ace-FAR-1, *Ancylostoma ceylanicum* fatty acid- and retinol-binding protein-1; Ac-FAR-1, *Ancylostoma caninum* fatty acid- and retinol-binding protein-1; Ac-SPI, serine protease inhibitor from *Ancylostoma caninum*; Bm-FAR, *Brugia malayi* fatty acid- and retinol-binding protein; Ce-FAR-7, *Caenorhabditis elegans* fatty acid- and retinol-binding protein-7; CRABP, cellular retinoic acid-binding protein family; CRBP, cellular retinol-binding protein; DAG, diacylglycerol; DAUDA, 11-(dansylamino)undecanoic acid; FA, fatty acid; FABP, fatty acid-binding protein; FAR, fatty acid- and retinol-binding protein; GC, gas chromatography; Gp-FAR-1, *Globodera pallida* fatty acid- and retinol-binding protein-1; Hp-FAR-1, *Heligmosomoides polygyrus* fatty acid- and retinol-binding protein-1; HPLC, high performance liquid chromatography; HSQC, heteronuclear single-quantum correlation spectroscopy; L3, third (infective) larval stage of a nematode; Mj-FAR-1, *Meloidogyne javanica* fatty acid- and retinol-binding protein-1; Na-FAR-1, *Necator americanus* fatty acid- and retinol-binding protein-1; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; NPA, nematode polyprotein antigen; Oo-FAR-1, *Ostertagia ostertagi* fatty acid- and retinol binding protein-1; Ov-FAR-1, *Onchocerca volvulus* fatty acid- and retinol-binding protein-1; PBS, phosphate buffered saline; PDB, protein structure database; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PL, phospholipid; RDC, residual dipolar coupling; RMSD, root-mean-square deviation; TLC, thin layer chromatography.

¹ Correspondence may be addressed to either of these authors (email Brian.Smith@glasgow.ac.uk or malcolm.kennedy@glasgow.ac.uk).

The structures of Na-FAR-1 determined by NMR and X-ray have been deposited in the PDB under accession codes 4UET and 4XCP respectively.

isoforms of ~20 kDa, eight having been found encoded within the genome of the free-living species *Caenorhabditis elegans*, each of which binds FAs and retinol to varying extents [10]. There are as yet uncharted numbers of FAR isoforms in parasitic nematodes, but FARs have drawn attention because they are secreted by both plant- and animal-parasitic species and their encoding transcripts are relatively abundant [11–18]. They have also proven useful for serodiagnosis, have shown promise in experimental vaccines [19,20] and have been proposed to facilitate infection by manipulating host lipid-mediated defences [21,22]. At least one FAR has been shown to bind an anthelmintic drug, these drugs being typically hydrophobic and so may require carrier proteins to conduct them to their site of action within parasites [23]. The definitive role of FARs in parasitism is not known, but their presence in secretions of the worms, coupled with their ligand-binding propensities, suggests roles for them either in acquiring lipids from the hosts, or in delivering or in sequestering signalling lipids and so modifying the tissues the parasites occupy or the immune responses against them.

The structure of one FAR protein, *C. elegans* FAR-7 (Ce-FAR-7), has been solved by X-ray crystallography, revealing a helix-rich structure that is unlike any type of lipid-binding protein previously described [24]. The mRNA for Ce-FAR-7 does not encode a secretory signal peptide and its amino acid sequence indicates that it is in a different subfamily of FARs from those that are secreted from the synthesizing cell. We set out to confirm the expression pattern of a secreted FAR protein from a parasite and characterize its structure and ligand-binding characteristics. The protein Na-FAR-1 derives from the blood feeding intestinal hookworm of humans, *N. americanus*. This parasite and the other hookworm of humans, *Ancylostoma duodenale*, together infect over 300 million people worldwide [25,26], causing considerable morbidity, together with adverse social and economic consequences. We have determined the structure of Na-FAR-1 by both X-ray crystallography and NMR spectroscopy in solution. We find that it has a similar overall fold to Ce-FAR-7, indicating that FARs are structurally conserved despite considerable sequence diversity. But, in addition to the structural differences between Ce-FAR-7 and Na-FAR-1, we find that their ligand-binding sites differ significantly in position and form.

EXPERIMENTAL

Protein expression and purification

Recombinant Na-FAR-1 (rNa-FAR-1) was expressed in BL21 (λ DE3) *Escherichia coli* cells as described [27]. For native crystallographic studies, Na-FAR-1 was purified to homogeneity, as previously described [28], from cells grown in LB media. Selenomethionine-labelled protein was purified from B834 cells grown in M9 minimal medium supplemented with a cocktail of free amino acids (each $0.5 \text{ g} \cdot \text{l}^{-1}$) and selenomethionine ($50 \text{ mg} \cdot \text{l}^{-1}$; Generon).

For NMR studies, samples of unlabelled, ^{15}N -labelled and $^{13}\text{C}^{15}\text{N}$ -labelled protein were purified by nickel-affinity, size exclusion and reverse-phase chromatographies, as described [27], from cells grown in M9 minimal medium containing $^{15}\text{NH}_4\text{Cl}$, [$^{13}\text{C}_6$]-glucose or their unlabelled equivalents.

Western blotting and immunolocalization of Na-FAR-1

Antiserum prepared against recombinant Na-FAR-1 was raised in three rabbits by subcutaneous injection with 0.7 mg of purified recombinant Na-FAR-1 in Freund's complete adjuvant.

Antiserum was tested by ELISA and Western blot analysis against the recombinant protein. To analyse the expression of Na-FAR-1 in the worm, soluble extracts of adult *N. americanus*, as well as recombinant proteins (each 100 ng) of homologues from *Ancylostoma caninum* FAR-1 (Ac-FAR-1); *Brugia malayi* FARs (Bm-FAR-1 and Bm-FAR-2) and the unrelated protein, recombinant Ac-SPI (serine protease inhibitor from *A. caninum*) were separated on a 4%–20% gradient SDS precast polyacrylamide gel (Invitrogen) and subsequently electrotransferred on to a PVDF membrane (Millipore). Rabbit anti-Na-FAR-1 serum was diluted 1:5000 into PBS (phosphate buffered saline), pH 7.4, with 0.05% Tween-20 and incubated with transferred membrane for 2 h. Horseradish peroxidase-conjugated goat anti rabbit-IgG was used as secondary antibody and ECL was used to develop the reaction (GE).

To determine the tissue-specific localization of Na-FAR-1, adult *N. americanus* worms were prepared as previously described [29]. Briefly, adult worms were collected from the intestines of hamsters infected with *N. americanus* L3 [third (infective) larval stage of a nematode] for 45 days and fixed with 10% formalin. The fixed worms were sectioned and mounted on glass slides. The non-specific binding sites on worm sections were blocked with 5% FBS in PBS for 1 h. The rabbit anti-Na-FAR-1 serum was applied (1:500 dilution) to each tissue section and incubated for 2 h at room temperature in a humidified chamber. Pre-immune rabbit serum at the same dilution was used as a negative control. Sections were washed six times for 5 min each in PBS and probed with anti-rabbit Cy3-conjugated IgG (Rockland). Sections were viewed under a Nikon TE-2000 Inverted fluorescence microscope using a 550 nm excitation filter block and emission at 565 nm.

Crystallization, data collection, processing and structure solution

We have shown previously that Na-FAR-1 crystallizes in two crystal forms, one of which (form 2) shows significant twinning [28]. Here, in order to obtain phasing information, selenomethionine-substituted protein was purified and crystallized, selecting only the cubic crystal form 1. Crystals were frozen in a stream of cool nitrogen gas (100 K) and brought to the Diamond Light Source, station I04 (DLS) for X-ray diffraction data collection.

Data were collected at 0.7° increments per image, for a total of 200 images [wavelength 0.9793 \AA ($1 \text{ \AA} = 0.1 \text{ nm}$)] and processed by the automatic processing routines *fast_dp*, which utilized XDS [30], POINTLESS and SCALA [31]. The structure was solved using the SAS protocol of Auto-Rickshaw [32]. The input diffraction data were prepared and converted for use in Auto-Rickshaw, using programs of the CCP4 suite [33]. F_A values were calculated using the program SHELXC [34]. Based on an initial analysis of the data, the maximum resolution for sub-structure determination and initial phase calculation was set to 2.14 \AA based on the scaling statistics and the increase in R_{meas} in the highest resolution bin. All the four heavy atoms requested were found using the program SHELXD [35]. The correct handedness for the substructure was determined using the programs ABS [36] and SHELXE [34]. Initial phases were calculated after density modification using the program SHELXE. 83.9% of the model was built using the program ARP/wARP [37]. Despite a solvent content above 70%, only one copy of the protein was observed in the asymmetric unit. The model was completed by hand using COOT [38] and iterative rounds of BUSTER [39]. Ligands and water were added using COOT. Based on its prevalence in the lipids co-purifying with recombinant Na-FAR-1, palmitate was fitted into unoccupied electron density within the binding pocket.

Further electron density that may represent additional ligand molecules was observed, but as the density was incomplete and not at full occupancy, it was left unmodelled. The geometry of the finished models was validated using Molprobity [40].

Preparation of NMR samples

Apo-Na-FAR-1 samples were delipidated by reverse phase-HPLC (high performance liquid chromatography) as previously described [27] using a C₈-silica stationary phase and a water/acetonitrile gradient in the presence of 0.1 % trifluoroacetic acid as the mobile phase. Samples were lyophilized, reconstituted and concentrated to approximately 0.6 mM in 20 mM sodium phosphate, pH 7.20. ²H₂O was added to a final concentration of 5 % (v/v). The sample used for residual dipolar coupling (RDC) measurements was prepared by partial alignment of the uniformly ¹³C/¹⁵N-enriched Na-FAR-1 in a solution of magnetically aligned filamentous Pf1 bacteriophage [41] (ASLA biotech). The degree of alignment was evaluated by measuring the ²H quadrupolar splitting in the HDO resonance. After testing several conditions, a NaCl concentration of 300 mM with 9 mg·ml⁻¹ of bacteriophage and 300 μM protein was selected.

NMR spectroscopy and analysis

All spectra were recorded at 311 K on a Bruker AVANCE 600 MHz spectrometer equipped with TCI cryoprobe. Resonances were assigned, as described previously [27]. All spectra were processed in AZARA (Wayne Boucher, <http://www.bio.cam.ac.uk/azara>). Maximum entropy reconstruction [42] was used to enhance resolution of the indirect dimensions of 3D experiments. Spectra were analysed with CCPNmr analysis software [43]. Frequency-based methods were employed to measure ¹D_{NH} [44] and ¹D_{CαHα} [45] couplings from in-phase/antiphase ¹⁵N-HSQC (heteronuclear single-quantum correlation spectroscopy) spectra. Distance restraints for structure calculations were derived from 3D ¹⁵N-NOESY-HSQC [46] and 3D ¹³C-NOESY-HSQC spectra, each recorded with 100 ms mixing time.

Structure calculations from NMR-based restraints

Distance restraints were derived from NOESY (nuclear Overhauser effect spectroscopy) cross-peaks with the initial mapping from normalized intensity to distance and grouped in distance bins. NOE distance restraints were incorporated in restrained MD calculations using the ambiguous distance restraints formalism [47] using ARIA 2.3 [48] and CNS [49]. Loose backbone dihedral restraints for regions of regular secondary structure predicted based on secondary chemical shifts by DANGLE [50] were incorporated during the high temperature phases of the simulations but omitted during the final cooling phase. RDC and hydrogen bond restraints were then introduced. The average RDC alignment tensor was estimated from the ensemble calculated using only NOEs with PALES [51] and used to incorporate the RDC restraints via the SANI potential [52] in square-well mode. The 20 structures that best satisfy the experimental restraints were chosen from 100 structures generated in the final iteration and refined in explicit water [53]. The quality of these structures was analysed using PROCHECK_nmr [54] and their co-ordinates deposited in the Protein Data Bank under accession code 4UET. Structure figures were generated using PyMOL (<http://www.PyMOL.org>).

NMR relaxation measurements

¹⁵N-relaxation time constants, *T*₁ and *T*₂ were assessed using the method of Kay [55–57] at a field strength of 14.1 T. Relaxation delays for assessment of *T*₁ were 101, 601, 1001 and 1401 ms whereas those for *T*₂ were 17, 34, 68, 102 and 136 ms. Selected time points in each series were repeated in order to estimate the inherent error in calculation of cross-peak intensities. Relaxation times *T*₁ and *T*₂ were calculated using non-linear least squares fitting. Collection of ¹⁵N-HSQC-heteronuclear NOE experiments with and without saturation allowed extraction of ¹H,¹⁵N NOE values. Both saturation and reference experiments were repeated for the purpose of error estimation.

Ligand-binding analysis by chemical shift perturbation

In order to examine Na-FAR-1's ligand binding properties by NMR, the chemical shift changes induced in ¹⁵N HSQC spectra of the protein by the presence of increasing amounts of ligand were analysed. Double-labelled ¹³C-¹⁵N recombinant Na-FAR-1 (0.4 mM) was titrated by sequential addition of small volumes of unlabelled sodium oleate (Sigma–Aldrich) from a 125 mM stock solution in water (pH ~9). Triple resonance experiments were recorded at selected points of the titration in order to assign the displaced cross-peaks. The titration started by addition of 0.5 molar equivalent of the ligand and was continued until turbidity due to the presence of insoluble oleic acid was observed.

Lipid extraction and analysis

Total lipids were extracted according to the methodology described by [58] and [59], with minor modifications. For unstripped Na-FAR-1 (no RP-HPLC purification), lipids were extracted from 15 mg of protein and compared with control *apo*-Na-FAR-1 (RP-HPLC purified protein). For comparison, approximately 3 ml of culture of *E. coli* BL21 (λDE3) cells were lysed by sonication. Each sample was mixed with 15 ml of CHCl₃–CH₃OH (2:1) and vigorously shaken for 15 min in an ice bath. The homogenate was washed with 250 μl of 2.9 % (w/v) NaCl solution. After agitation, the phases were separated by centrifugation and the upper, aqueous phase discarded. The lower phase containing lipids was recovered and dried under a stream of N₂ gas, re-dissolved in CHCl₃ and stored at –20 °C under N₂ gas until analysis.

Lipid classes bound to Na-FAR-1 were analysed by TLC (thin layer chromatography) on silica gel plates Si250 (J.T.Baker) with the methodology and solvent systems described by [59]. Lipid samples obtained from *holo*-Na-FAR-1 and *E. Coli* lysates controls and standards were spotted on TLC plates (20 × 20 cm) previously activated at 100 °C for 30 min and developed with methyl acetate–isopropanol–chloroform–methanol–0.25 % KCl (25:25:25:10:9, by volume) for polar lipids and hexane/diethyl-ether/acetic acid (80:20:1, by volume) for neutral lipids.

Non-esterified FA and phospholipid (PL) TLC spots were scraped and extracted from the silica with pure chloroform for FAs and chloroform–methanol–water for PLs respectively. FA composition was analysed by GC (gas chromatography)–MS of their methyl ester derivatives, prepared with BF₃–methanol according to the method of Morrison and Smith [60] as described previously [61]. The individual FA methyl ester peaks were identified by comparison of their retention times with those of standards and by their mass spectra.

Fluorescence spectroscopy

Fluorescence experiments were performed with a Fluorolog-3 Spectrofluorometer (Horiba-Jobin Yvon). Buffer alone was

used to correct for Raman and background scattering. RP-HPLC delipidated Na-FAR-1 was employed in all ligand-binding experiments.

The ligand-binding capacity of Na-FAR-1 was investigated with the fluorescent ligands 11-(dansylamino)undecanoic acid (DAUDA) and retinol. Stock solutions (10 mM) were prepared in ethanol and then diluted in PBS for use in the assays. Retinol solutions were diluted in ethanol and added directly to the cuvette to minimize degradation.

The FA chain length preference of rNa-FAR-1 was tested by displacement of the fluorescent ligand DAUDA as described [13]. Binding of non-fluorescent ligands was detected by a reversal of the wavelength shift and a decrease in fluorescence emission intensity on equal additions of test ligands to a DAUDA-rNa-FAR-1 complex recorded at the peak fluorescence emission wavelength of DAUDA in the protein (470 nm). The concentration of Na-FAR-1 in the cuvette was 1.5 μ M. DAUDA ethanol stock solution was diluted 1:10000 in PBS for use in the assays at 1 μ M. Stock solutions of all the non-fluorescent competitors were made to approximately 10 mM in ethanol, then diluted in PBS for use in the assays.

RESULTS AND DISCUSSION

Na-FAR-1

Na-FAR-1 had been identified in a gene survey of the *N. americanus* transcriptome under the name *N. americanus* LBP-20 (sequence ID NAC00128) [62]. LBP-20 proteins are currently named FAR due to their capacity for binding FAs and retinol and Na-FAR-1 was renamed accordingly [63]. More recently, the genome of *N. americanus* has been published and at least six FAR proteins have been recognized (Figure 1; Supplementary Figure S1), including Na-FAR-1, under the sequence name NECAME_14208 [64]. Na-FAR-1 cDNA predicts a 19364.57 Da protein with a 14 amino acid secretion signal peptide predicted by SignalP [65]. The post-translational removal of the leader peptide would yield a mature protein of 155 amino acids with a molecular mass of 17082.49 Da.

A phylogenetic tree was constructed to show the relationship between Na-FAR-1 and FARs from other species (Figure 1b). The tree was constructed with amino acid sequences omitting any leader peptides identified by SignalP. The sequences included those of *C. elegans* (Ce-FAR-1–8), the human parasitic nematodes *Onchocerca volvulus* FAR-1 (Ov-FAR-1), *B. malayi* (Bm-FAR-1), animal parasitic nematodes *A. caninum* (Ac-FAR-1 and Ac-FAR-2), *Ancylostoma ceylanicum* FAR-1 (Ace-FAR-1), *Ostertagia ostertagi* FAR-1 (Oo-FAR-1) and *Heligmosomoides polygyrus* FAR-1 (Hp-FAR-1) and the plant parasitic nematodes *Globodera pallida* FAR-1 (Gp-FAR-1) and *Meloidogyne javanica* FAR-1 (Mj-FAR-1). FAR transcript levels in some of these parasites are notably high, particularly in the parasitic stages of life cycles [11–18,62,66]. Figure 1 shows that most FAR proteins known from parasites group together, including those from *N. americanus*, and a subset of FARs from *C. elegans* (Ce-FAR-1, 2 and 6). As noted above, Ce-FAR-7, the only FAR protein whose structure was previously known, falls in a subfamily of FARs distant from all of the FAR proteins from parasites [10] (Figure 1b).

Na-FAR-1 localization within the parasite

Na-FAR-1-encoding mRNA had been detected in adult and larval L4 *N. americanus* life stages [62] and analysis of differences in gene expression between infective larvae (iL3) and adult parasitic stages shows that Na-FAR-1 is mainly expressed in the adult,

Table 1 Data collection, phasing and structure refinement statistics for *holo*-Na-FAR-1 determined by X-ray crystallography (PDB:4XCP)

Numbers in brackets indicate values in the highest resolution bin (2.27–2.14 Å). Abbreviation: PDB, protein structure database.

Data collection	
Space group	P432
Unit cell	$a = b = c = 121.123$
Wavelength (Å)	0.9793
Resolution (Å)	29.38–2.14
Observed reflections	502169 (17556)
Unique reflections	17353 (1212)
Completeness	99.91 (97.0)
Multiplicity	28.9 (14.5)
R_{meas} (%)	14.2 (78.5)
R_{pim} (%)	2.6 (20.3)
R_{anom} (%)	5.2 (17.5)
$\langle I/\sigma(I) \rangle$	25.5 (4.2)
Wilson B (Å) ²	30.11
Number of Selenomethionines found	4
Refinement	
R_{work} (%)	20.5
R_{free} (%)	22.4
Protein residues/atoms	155 (1213)
Water molecules	139
Palmitic acid	1
RMSD bond lengths (Å)/angles (°)	0.01/0.99
Average isotropic thermal parameters (Å ²)	
Main chain/side chain atoms	25.06/33.69
Water molecules	38.09
Palmitic acid	38.11
Ramachandran analysis	
% Favoured regions	98.05
% Outliers	0.65

blood-feeding stage [64]. Immunohistochemistry revealed that Na-FAR-1 protein occurs in the cloacal aperture and copulatory bursa of male worms (Figures 1c and 1d) and in the intestine (Figures 1f and 1g). The presence of Na-FAR-1 in the parasite's intestinal cells may indicate involvement in transportation and storage of lipids derived from host blood and its location in copulatory bursa of male worms suggests its function may additionally be related to reproduction. Na-FAR-1 was detected in the intestinal cells but not the reproductive structures of females (Figure 1e).

3D structure of Na-FAR-1

The 3D structure of Na-FAR-1 was determined by both protein X-ray crystallography and solution state NMR spectroscopy. Na-FAR-1, which had not been subjected to reverse phase HPLC purification to strip out co-purifying ligands, crystallized as previously reported and diffracted to 2.14 Å [28]. Molecular replacement using the Ce-FAR-7 structure proved unsuccessful and we therefore obtained the crystal structure of Na-FAR-1 in complex with co-purifying ligands, hereafter referred to as *holo*-Na-FAR-1, by anomalous dispersion from crystals of Se-Met-labelled protein (Table 1). In contrast, NMR spectra of unstripped Na-FAR-1 in solution, like those of other FAR proteins we have tested, were characterized by broad signal peaks indicative of multiple conformations and/or conformational exchange. However, stripped Na-FAR-1 gave good solution NMR spectra and the structure of *apo*-Na-FAR-1 was determined from a total of 7289 NOE-derived distance restraints, 316 dihedral angle restraints and restraints derived from 177 RDCs (101 ¹D_{NH} and

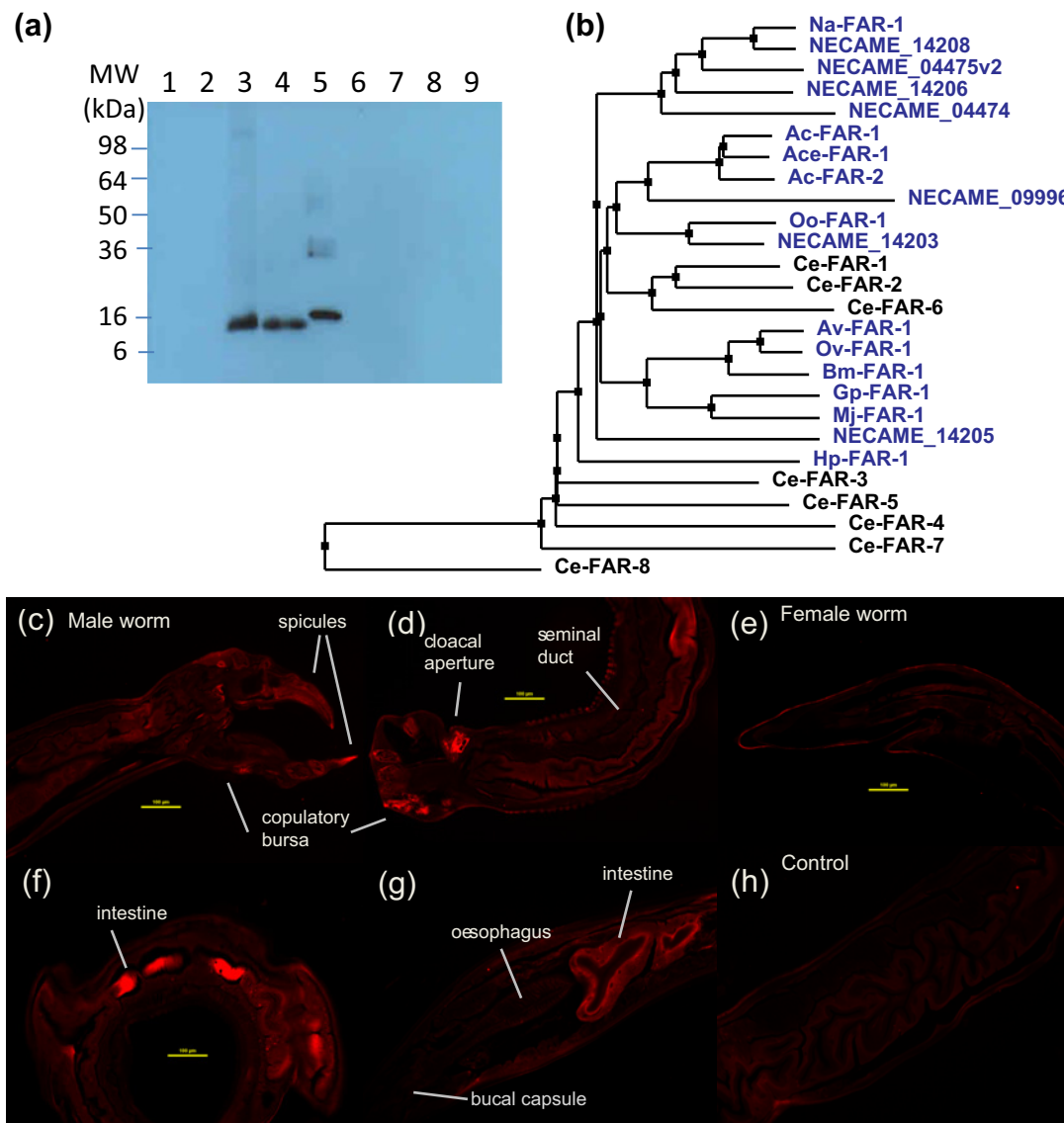


Figure 1 FAR protein relationships and Na-FAR-1 expression

(a) Western blot with anti-Na-FAR-1 serum that specifically recognizes native Na-FAR-1 in *N. americanus* extracts (lane 3, 0.5 μg) and excreted/secreted (ES) products (lane 4, 0.5 μg) at an approximate M_r of 14 kDa, but not in L3 extracts (lane 1) and L3 ES products (lane 2) at the same loading, indicating the specific expression of Na-FAR-1 in adult stage as a secreted protein. The antiserum also recognized the recombinant Na-FAR-1 at 16 kDa (with His-tag, lane 5, 20 ng). There was no cross-reaction with FAR homologues from dog hookworm *A. caninum* (Ac-FAR-1, lane 6); *B. malayi* (Bm-FAR-1/lane 7, Bm-FAR-2/lane 8) and non-relevant recombinant protein Ac-SPI (lane 9) loaded at the same amount (20 ng). (b) Neighbour joining tree of Na-FAR-1 and other nematode FAR proteins. The tree was generated in *jalview* 2.8 [69] using the BLOSUM 62 matrix from a *T-coffee* WS sequence alignment [70] with FAR protein amino acid sequences of: *N. americanus* recently identified FAR proteins (NECAME_09996, NECAME_04475, NECAME_04474, NECAME_14206, NECAME_14205 and NECAME_14203), the free living nematode *C. elegans* (Ce-FAR-1 to Ce-FAR-8), the human parasitic nematodes *O. volvulus* (Ov-FAR-1), *B. malayi* (Bm-FAR-1), animal parasitic nematodes *A. caninum* (Ac-FAR-1 and Ac-FAR-2), *A. ceylanicum* (Ace-FAR-1), *O. ostertagi* (Oo-FAR-1) and *H. polygyrus* (Hp-FAR-1) and the plant parasitic nematodes *G. pallida* (Gp-FAR-1) and *M. javanica* (Mj-FAR-1). All the parasite proteins are coloured blue. See Supplementary Figure S1 for the multiple sequence alignment from which the tree was constructed. (c–g) Localization of Na-FAR-1 within adult male (c and d) and female (e–g) worms. Indirect immunofluorescence localization with rabbit anti-Na-FAR-1 serum stains the intestinal cells of adult *N. americanus* worms. Na-FAR-1 was also detected on the copulatory bursa and cloacal aperture of male worms. (h) Control carried out using pre-immune serum. Scale bars represent 100 μm.

76 $^1D_{CaHa}$) observed in a sample that had been partially aligned in Pf1 filamentous bacteriophage (Table 2).

The overall folds of the Na-FAR-1 structures in complex with ligands (*holo*) and without ligand (*apo*) are similar (Figure 2a–c). Superposition of the *holo* and *apo* structures gives a 1.814 Å co-ordinate RMSD (root-mean-square deviation) for all heavy atoms (1.579 Å for main chain heavy atoms only). Na-FAR-1 presents a wedge-shaped structure with two larger faces each of approximately 40 by 30 Å in area and of ~17 Å in width at the wide end of the wedge. The fold is organized into 11 helices of various lengths, defining an internal cavity (Figure 3). The N-

terminal helices, 3₁₀ (residues 3–5), α2 (8–11) and α3 (16–24) are co-planar with the C-terminal helices α9 (126–135), α10 (137–144) and α11 (147–152) and together with one of the longest helices, α6 (57–75), form one large face of the wedge. The other two long helices, α7 (79–99) and α8 (107–122), are co-planar with α5 (45–55) and form the other large face. These two faces enclose an internal cavity that is closed at the thick end of the wedge by α4 (29–37), which is almost perpendicular to the long axis of the molecule. Hydrophobic residues are located almost exclusively on the inward-facing sides of the helices with their side chains pointing towards the internal cavity. Most of the polar and charged

Table 2 Structural statistics for the ensemble of 20 apo-Na-FAR-1 structures determined by NMR spectroscopy

NOE distance restraints	
Total NOE	7097
Ambiguous	3443
Unambiguous	3654
Intra-residue	1548
Inter-residue	2106
Sequential ($ i - j = 1$)	833
Short range ($1 < i - j \leq 4$)	756
Long range ($ i - j \geq 5$)	517
Violations $> 0.5 \text{ \AA}$	0.3
Violations $> 0.3 \text{ \AA}$	1.6
Distance restraints RMSD (\AA)	0.016
Coordinate RMSD (\AA)	
Backbone heavy atoms	0.597
All heavy atoms	1.012
Parameter RMSD from idealized geometry (mean \pm S.D.)	
Bond lengths (\AA)	$0.00326 \pm 1.4 \times 10^{-4}$
Bond angles ($^\circ$)	0.595 ± 0.018
Impropers ($^\circ$)	1.57 ± 0.07
RDC restraints	
$^1\text{D}_{\text{NH}}$	102
$^1\text{D}_{\text{C}\alpha\text{H}\alpha}$	76
RDC Q factor	0.164
Dihedral angles	312
Ramachandran statistics (%)	
Most favoured	92.9
Additionally allowed	6.0
Generously allowed	0.8
Disallowed	0.3

residues are on the external surface of the protein generating a predominately hydrophilic surface. The notable exceptions are Ser⁸⁸, Lys⁹⁶ and Tyr¹⁰⁰ whose side chains are located within the cavity, whereas His⁶⁷ and Arg⁹³ stand on either sides of the largest opening to the cavity.

Despite their similarities, the structures of the *apo* and *holo* forms of Na-FAR-1 reveal significant differences that reflect structural alterations upon ligand binding. The internal cavity is much larger for the *holo* protein, achieved by a global outward displacement of the surrounding helices by a little over 1 \AA compared with the *apo* form, with the biggest changes seen at the C-terminal end of helix $\alpha 7$ (Figure 2a). For the *apo* structure, the central cavity has a volume of approximately 1220 \AA^3 accessible to a 1.4 \AA radius probe (equivalent to a water molecule) or 940 \AA^3 accessible to a 1.925 \AA probe (equivalent to a CH₂ group). In the *holo* form, the cavity is more than twice as big, reaching 2570 and 2170 \AA^3 for probes of 1.4 and 1.925 \AA respectively.

The main part of the *apo*-Na-FAR-1 cavity is lined by the side chains of residues Leu¹³, Met¹⁴, Pro¹⁵, Ala¹⁸, Phe²¹, Leu²², Leu³³, Phe³⁷, Thr⁴⁵, Val⁶³, Leu⁶⁶, His⁶⁷, Val⁷⁰, Ile⁸⁹, Ala⁹², Arg⁹³, Lys⁹⁶, Tyr¹⁰⁰ and Leu¹³⁹, most of which are apolar, with polar groups contributed by the side chain of Tyr¹⁰⁰ and backbone carbonyl of Leu¹³ and the positively charged Lys⁹⁶. His⁶⁷ and Arg⁹³ are located at the entrance to the cavity. A further, less accessible volume is located beyond a pinch point between Ala⁹² and Leu¹³⁹ and is lined by the predominately hydrophobic side chains of Tyr¹⁰, Ser⁸⁸, Val¹¹⁴,

Val¹¹⁷ and Phe¹⁴³. The cavity has a single opening located on one side of the wedge between $\alpha 6$ and $\alpha 7$ and the $\alpha 4$ – $\alpha 5$ loop that is surrounded by charged residues. In the *holo* structure the expanded cavity is lined by the same residues as in the *apo* structure, but with 21 additional, predominantly apolar, side chains contributing to the surface of the expanded cavity (Phe¹, Ile⁶, Leu²⁵, Lys³⁰, Lys³⁴, Val³⁶, Ser⁴⁸, Ile⁴⁹, Leu⁵², Met⁶⁹, Phe⁸⁴, Ala⁸⁵, Ile⁹¹, Ile⁹⁵, Tyr⁹⁹, Leu¹¹⁰, Ser¹¹³, Phe¹³², Phe¹³⁶, Val¹⁴² and Ala¹⁴⁸). In addition to the opening seen in the *apo* form, the cavity is also accessible to solvent via an opening between the $\alpha 2$ – $\alpha 3$ and $\alpha 9$ – $\alpha 10$ turns and $\alpha 6$. It is therefore possible that more than one region could be involved in ligand entrance or exit and the loop between $\alpha 7$ and $\alpha 8$, which exhibits a higher mobility than the average, could be a candidate region to participate in those processes. Within the cavity of the *holo* form, clear electron density was observed that is well fitted by a single FA molecule, containing 16 carbon atoms (Figure 2d). Since the predominant FA found in *E. coli* lipids is palmitate, we fitted this to the electron density. The palmitate appears to be curled into a C-shape around the side chain of Lys⁹⁶ and contacts 11 other residues (Phe¹, Ile⁶, Met¹⁴, Phe²¹, Leu²², Leu²⁵, Leu³³, Leu⁶⁶, Ile⁹⁵, Tyr⁹⁹ and Tyr¹⁰⁰). Additional electron density, possibly indicative of several other bound ligand molecules, is seen in the cavity of Na-FAR-1, but in light of the heterogeneous mixture of lipids associated with the protein and the poor quality of the density, we have not modelled them.

Comparison of the structures of Na-FAR-1 and Ce-FAR-7

Ce-FAR-7 has previously been noted to lie in a distinct sub-group of FARs because it possesses two cysteines that may form a disulfide bond, has distinct lipid-binding characteristics, lacks a secretory leader peptide and because of the intron arrangement of its encoding gene [10,24]. The X-ray crystallographic structure of Ce-FAR-7 [24] displays similarities but several distinct differences to Na-FAR-1 that could relate to differences in their ligand binding and biological properties.

A structure-based sequence alignment (Figure 3a) reveals sequence identity between the two proteins at 32 positions (~21%). Among these are: hydrophobic residues, Leu⁷⁷ and Ala⁸¹ (Leu⁷⁸ and Ala⁸² in Ce-FAR-7) that help to determine the angle between $\alpha 6$ and $\alpha 7$; Leu³³, Leu⁵⁹ (Leu³³ and Leu⁶⁰ in Ce-FAR-7) together with Leu⁵², which is replaced by another hydrophobic residue Val⁵³, which are involved in maintaining the orientation between $\alpha 6$ and helices $\alpha 4$ and $\alpha 5$; a salt bridge between Glu³⁵ and Lys⁵⁵ (Glu³⁵ and Lys³⁶ in Ce-FAR-7) that helps to maintain the relative orientation of $\alpha 4$ and $\alpha 5$; and another salt bridge between Lys⁷³ (Arg⁷⁴ Ce-FAR-7) and Glu¹⁷ that helps to maintain the relative orientation of $\alpha 6$ and $\alpha 3$. Na-FAR-1 has an insertion of four residues (107–110) that constitute an additional turn of helix at the N-terminal end of $\alpha 8$. Conserved prolines Pro⁷, Pro¹⁵ and Pro⁷⁹ initiate helices $\alpha 2$, $\alpha 3$ and $\alpha 7$, corresponding to Pro⁷, Pro¹⁵ and Pro⁸⁰ in Ce-FAR-7, as they seem likely to, in other FARs. Ce-FAR-7 has an additional residue in the loop between helices $\alpha 4$ and $\alpha 5$ and it is 14 residues shorter at the C-terminal region, thereby lacking helices $\alpha 10$ and $\alpha 11$. In Ce-FAR-7 $\alpha 1$ comprises a full turn of α -helix, whereas Na-FAR-1 has a 3_{10} helix at this position. A comparison of the structures gives an RMSD of 1.711 \AA (for a total of 636 atoms aligned) with the *apo* form and 1.646 \AA (608 atoms aligned) with the *holo* form. The region comprising the $\alpha 4$ – $\alpha 5$ joining loop, which might constitute part of the entrance to the cavity, was not modelled in the structure of Ce-FAR-7 due to insufficient electron density. It is likely that the lack of good definition is indicative of dynamic processes occurring in the $\alpha 4$ – $\alpha 5$ loop, as we found by NMR for Na-FAR-1 (Supplementary Figure S2).

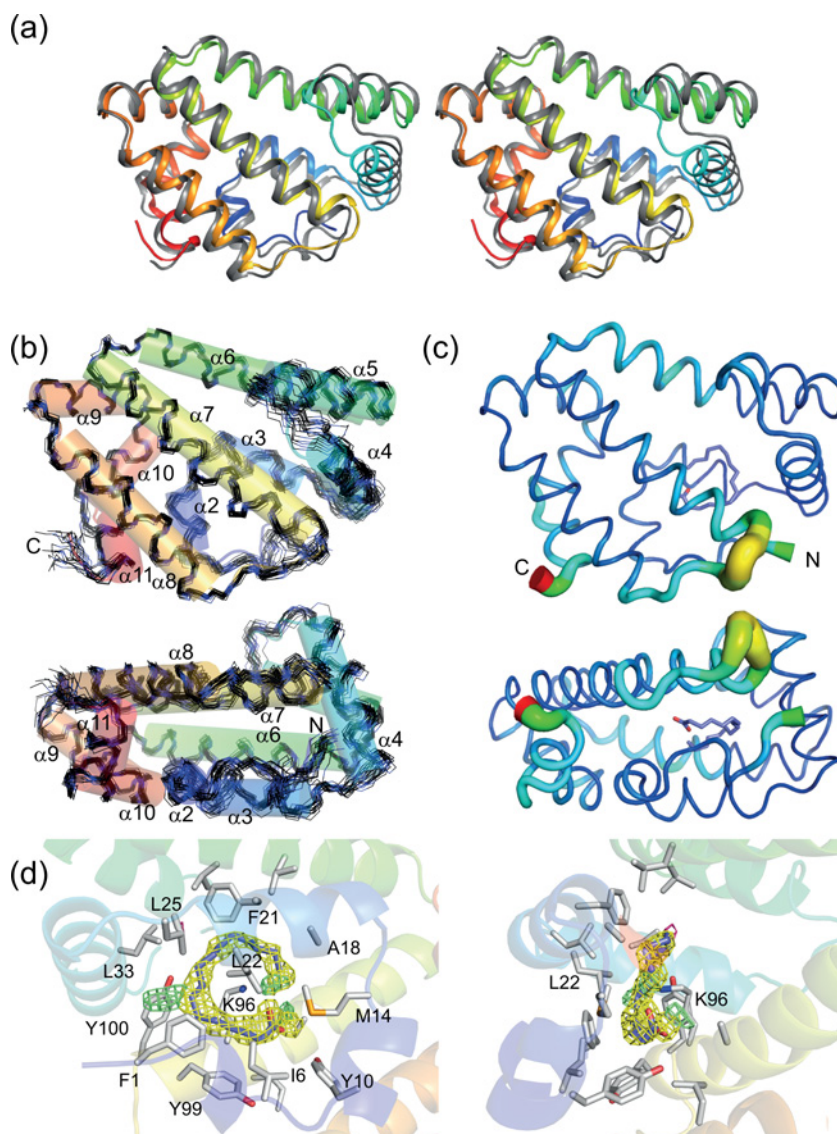


Figure 2 The structure of Na-FAR-1

(a) Stereoscopic view (wall-eyed) of the superimposed *apo*- and *holo*-Na-FAR-1 structures shown in cartoon form. The *apo* structure determined by NMR is coloured from blue at the N-terminus through to red at the C-terminus. The *holo* form determined by X-ray crystallography is shown in grey. (b) The ensemble of the 20 lowest energy structures of *apo*-Na-FAR-1 shown in two orientations related by a 90° rotation. (c) The *holo*-Na-FAR-1 structure shown in b-factor putty representation with the palmitate shown as sticks. (d) Expanded view of the bound palmitate and its environment with 2Fo-Fc electron density within 4 Å of the ligand displayed as a mesh at 1σ and shaded yellow and Fo-Fc difference density displayed at ±3σ (green and red).

The estimated cavity for Ce-FAR-7 calculated with CASTp has a much smaller size than for both forms of Na-FAR-1, with a volume of 687.7 Å³ accessible to a probe equivalent to a water molecule and just to 390.0 Å³ for a probe equivalent to a CH₂ group. In Ce-FAR-7, helices α2 and α3 approach α7 and α8 more closely, reducing the volume of the cavity and resulting in two distinct hydrophobic pockets. The majority of residues that line the Na-FAR-1 cavity and the Ce-FAR-7 P1 and P2 pockets are conserved or conservatively substituted. The exceptions are that the Na-FAR-1 cavity includes the side chains of more polar amino acids, such as Ser⁸⁸, Arg⁹³ and Lys⁹⁶, which are found in positions occupied by Leu⁸⁹, Thr⁹⁴ and Leu⁹⁷ respectively, in Ce-FAR-7.

Ligand binding

In fluorescence-based ligand-binding assays, Na-FAR-1 bound the fluorescent FA analogue DAUDA and the naturally fluorescent

lipid retinol (Figures 4d and 4e). The degree of blue shift in DAUDA fluorescence emission (from 543 nm in buffer to 480 nm) indicates Na-FAR-1 has a highly apolar binding site, as shown for other members of the FAR protein family [10]. It is also interesting to note that Ce-FAR-7 has shown no binding capacity for DAUDA [10] and binding to retinol showed no saturation [24]. The preference of Na-FAR-1 for FAs was investigated by the addition of FAs with different chain lengths to preformed protein–DAUDA complexes (Supplementary Figure S3), showing that DAUDA displacement was greatest with saturated FAs in the range C14:0–C19:0.

In order to determine the lipid classes bound to Na-FAR-1 in a cellular environment, a Folch extraction of the lipidic fraction of Na-FAR-1 purified from *E. coli* and subsequent TLC analyses were performed. The bacterial cytoplasm is not the natural environment of Na-FAR-1, but this approach could contribute to the understanding of its binding preferences. The

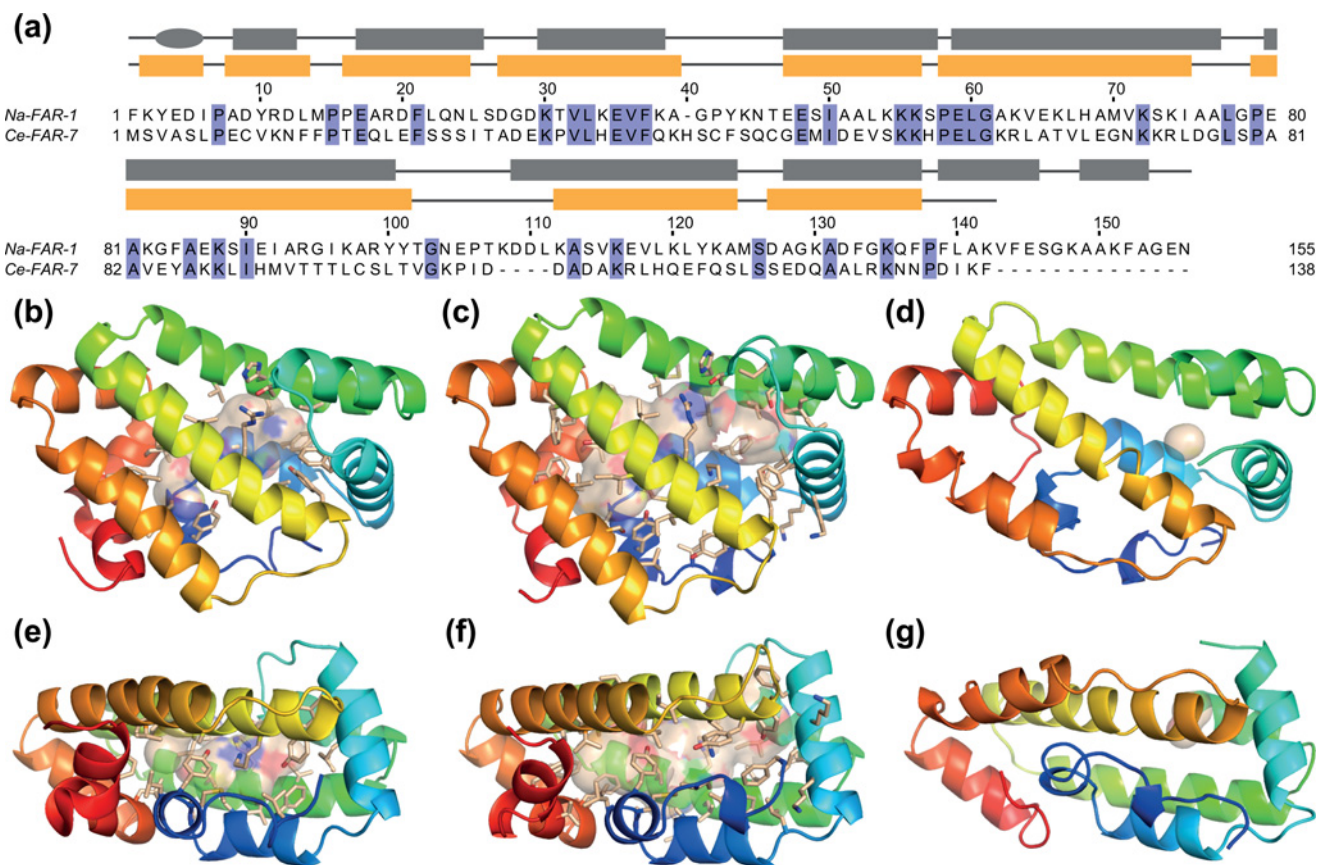


Figure 3 Na-FAR-1 cavities and comparison with Ce-FAR-7

(a) Sequence alignment of Na-FAR-1 and Ce-FAR-7. Secondary structure elements are indicated in boxes (α -helix) and lines (loops) coloured grey for Na-FAR-1 and yellow Ce-FAR-7. Comparison of Na-FAR-1 *apo* (b and e) and *holo* (c and f) forms with Ce-FAR-7 (d and g) shown in cartoon representation and coloured from blue (N-terminus) to red (C-terminus) viewed from two orientations related by a 90° rotation about the horizontal axis. Internal cavities accessible to a probe of 1.925 Å, equivalent to a methylene group are shown as transparent surfaces with the surrounding sidechains shown as sticks. (PDB accession codes Na-FAR-1 *apo* 4UET and *holo* 4XCP; Ce-FAR-7, 29WY)

lipid fraction bound to Na-FAR-1 exhibits heterogeneous lipid content (Figure 4a). Previous studies have shown that other members of the FARs bind retinol and FAs [10,13,15,19,21]. In addition, the present analysis reveals that the protein binds not only to FAs but also a broader range of lipid classes such as PLs.

The qualitative analysis conducted suggests the presence of polar lipids such as phosphatidylethanolamine (PE) and phosphatidylglycerol (PG). Esterification and analysis of the non-esterified FA and PL-derived FA methyl esters by GC-MS (Figures 4b and 4c) showed that in the lipids carried by Na-FAR-1, palmitate (16:0) was the most common FA. Examination of total *E. coli* lipids showed a similar FA composition.

Protein–ligand interactions were further investigated by titration of Na-FAR-1 with sodium oleate followed using NMR. The first step of the titration was made by adding of 0.5 molar equivalents of sodium oleate (Figure 5), at which point two cross-peaks originating from each shifted amide group were detected, one remaining at the original shift that corresponds to the *apo* protein conformation and another cross-peak at a different shift resulting from the protein conformation in the presence of one ligand bound. At subsequent stages in the titration, further chemical shift changes were observed, in each case resulting in the appearance of a new cross-peak and disappearance of the original. The ligand binding process therefore exhibited slow exchange behaviour through the addition of 1, 2 and 3 molar equivalents of oleate, which would suggest that the protein

binds three ligands with high affinity. Beyond this, additional chemical shift changes were observed in the fast exchange regime from a 1:4 protein–ligand ratio where only one cross peak was seen for each amide group. The titration was carried on to a maximum protein–ligand ratio of 1:10 where it was limited by the precipitation of excess, unbound ligand. These results suggest that the protein binds up to at least four oleate molecules, three of them with high affinity in the sub micromolar range and a fourth with lower affinity. Substantial chemical shift changes were observed for residues throughout the protein, consistent both with ligand binding affecting the local environment of residues within the helices surrounding the cavity and with the changes in conformation of residues in the loops required to allow the structure to expand.

CONCLUSION

The FAR proteins are unique to nematodes and we in the present study demonstrate the first structure of a member of this family from a parasitic species. FARs are present in the secretory products of parasites, which is apparent both for those that reside in the lumen of the gut and those that penetrate and reside in the host tissue of animals and plants [22,67]. This family of proteins comprises subfamilies that differ in their ligand-binding characteristics and in whether or not they are secreted from

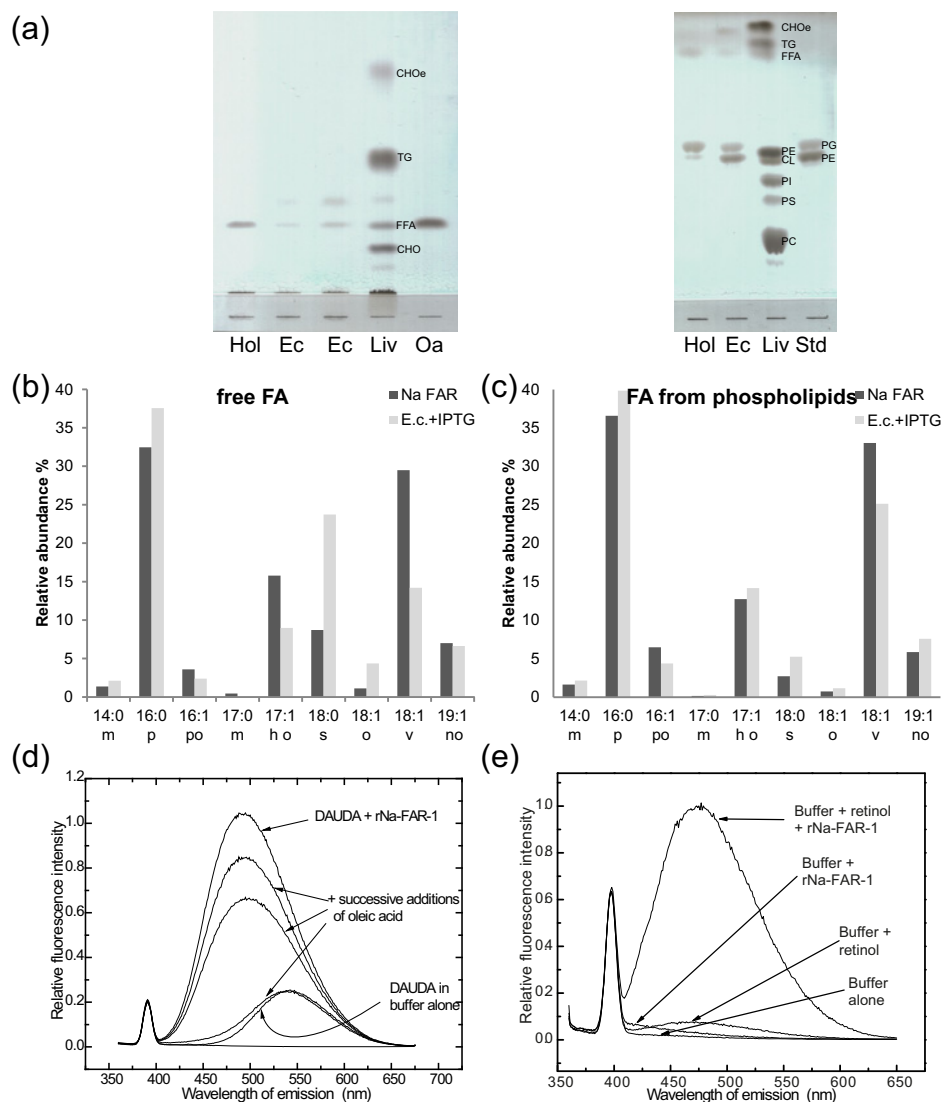


Figure 4 Lipid binding by recombinant Na-FAR-1

(a) Lipid fractions bound to bacterially-expressed Na-FAR-1 detected by TLC. Lipids were extracted from the protein and fractions were analysed by TLC in conditions for resolving separately neutral (left panel) and polar (right panel) lipid classes. Standards and samples applied to TLC plates were: Hol, Na-FAR-1 purified without an HPLC step; Ec, extract from whole *E. coli* cells; LIV, standard mix of lipids from rat liver homogenate; STD, *E. coli* whole extract. CHO, cholesterol; TG, triglycerides; CHOe, cholesterol esters; PC, phosphatidylcholine; PS, phosphatidylserine; PI, phosphatidylinositol; CL, cardiolipin. (b) GC-MS analysis of non-esterified FAs isolated from lipids associated with Na-FAR-1 purified from *E. coli* (dark grey) or found in *E. coli* extracts (light grey). FAs detected were: myristic (14:0), palmitic (16:0), palmitoleic (16:1), margaric (17:0), heptadecenoic (17:1), stearic (18:0), oleic (18:1 Δ 9), vaccenic (18:1 Δ 11) and nonadecenoic acid (19:1). (c) FAs found in PLs isolated from purified Na-FAR-1 (dark grey) or found in *E. coli* extracts (light grey). (d) Fluorescence emission spectra of 1 μ M DAUDA excited at 345 nm in buffer, on the addition of 1 μ M Na-FAR-1 and after successive additions of oleic acid (81 nM, 810 nM and 8.1 μ M). (e) Fluorescence emission spectra of 5 μ l of 0.15 mM retinol in ethanol added to the fluorescence cuvette and excited at 350 nm and on the addition of 1.5 μ M Na-FAR-1, compared with 1.5 μ M protein in buffer and buffer alone.

the cells that synthesize them. Transcripts encoding Na-FAR-1 are relatively abundant in the parasite from which it derives and the encoded polypeptide bears a secretory signal sequence. It falls into a new subfamily of the FARs that has not been found so far in any other group of nematodes, although the inventory of FARs is only comprehensive for those species for which genomic information is essentially complete and annotated (e.g. *C. elegans*).

It has been noted that all FAR protein sequences exhibit a consensus casein kinase II-type sequence that occurs in approximately the same place, despite the considerable degree of

sequence diversity among FARs [19,24,67,68]. This argues that this site may have some significance in, for instance, altering the ligand-binding activity of the proteins or modifying interaction with other cellular factors such as cell surface receptors (host or parasite). There is currently no evidence that this site is phosphorylated naturally, though experiments with Ce-FAR-7 indicate that modification of this site does alter ligand binding. The previous study found that direct phosphorylation rendered the protein unstable in some unknown way, but when phosphorylation was mimicked by substitution of the presumed target threonine by an aspartic acid, the affinity of Ce-FAR-7 for retinol increased,

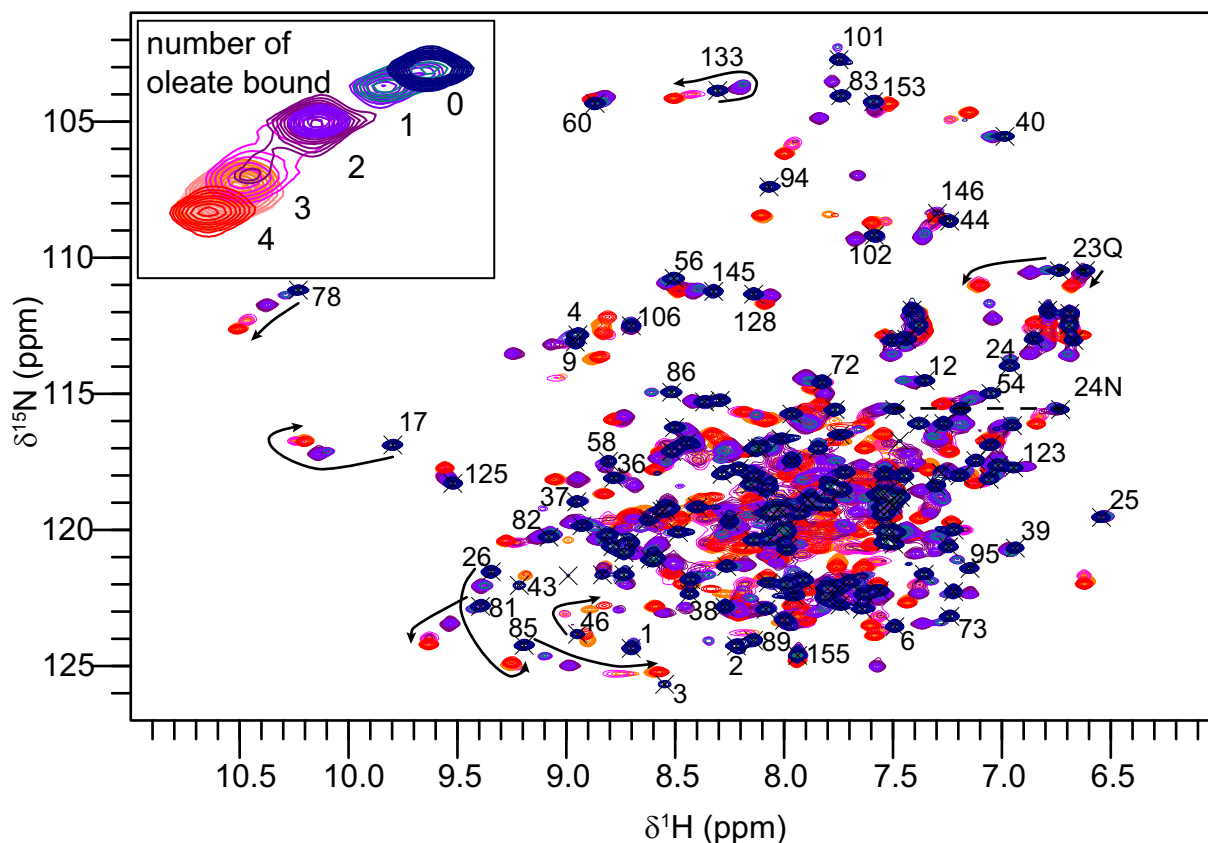


Figure 5 Titration of Na-FAR-1 with sodium oleate followed by NMR spectroscopy

^{15}N HSQC spectra of 0.4 mM Na-FAR-1 with successive additions of sodium oleate. The spectrum of the apo protein is shown in dark blue, with spectra coloured through purple to red at the final protein–ligand ratio of 1:10. A subset of the backbone and side chain (lettered) assignments are shown and the trajectories of selected residues indicated with arrows. The inset shows an expanded view of the cross-peaks from Gly⁷⁸ backbone amide with our interpretation of the protein–ligand stoichiometry responsible for each distinct cross-peak position.

although there was no change for FA. In both Ce-FAR-7 and Na-FAR-1, the amino acid side chain thought to be phosphorylated lies in a short unstructured region between two helices at one end of the molecule. In neither protein is the side chain of this residue oriented towards the internal cavity and nor in Ce-FAR-7 it is more distant from the internal surface than in Na-FAR-1. It is therefore not immediately clear how phosphorylation could affect the proteins' ligand-binding characteristics unless structural changes result.

The differences in the size and shape of the internal cavities of Ce-FAR-7 and Na-FAR-1 are probably indicative of differences in their ligand selectivity. Both proteins bind retinol and FAs, but whether these compounds are relevant *in vivo* and whether other classes of ligand (such as PLs here found to be bound by Na-FAR-1) are important, remains to be established. The fact that FARs are commonly found in the secretions of parasitic nematodes bolsters confidence that they play an important role in parasitism. Parasites need to acquire nutrients from their hosts, but also need to defend themselves against immune defence reactions and some parasites induce modifications to the tissues they occupy. If FARs do bind signalling lipids, then it is conceivable that they are released by parasites in order to subvert or modify the hosts' tissue and defence reactions. If so, then FARs would seem to be logical targets for vaccines, particularly against parasites that are embedded in host tissues. Development of drugs against them would need to take into account the possibility of interference with the hosts' proteins

and enzymes that transport or modify FAs and retinoids. But, it is entirely possible that we have yet to discover the true range of the ligand-binding propensities of FARs in nematodes. An understanding of their structures and how they vary between those secreted by different species of parasites could illuminate their roles in parasitism and suggest possible targets for therapeutic interventions.

AUTHOR CONTRIBUTION

Florencia Rey-Burusco, Mads Gabrielsen, Kate Griffiths, Gisela Franchini Bin Zhan, Alan Cooper, Malcolm Kennedy, Betina Córscico and Brian Smith performed the experiments. Florencia Rey-Burusco, Marina Ibáñez-Shimabukuro, Mads Gabrielsen, Gisela Franchini Andrew Roe, Kate Griffiths, Bin Zhan, Alan Cooper, Malcolm Kennedy, Betina Córscico and Brian Smith analysed the data. Florencia Rey-Burusco, Betina Córscico, Malcolm Kennedy and Brian Smith wrote the paper. All authors have seen, made appropriate modifications to and approved of the manuscript.

ACKNOWLEDGEMENTS

We gratefully acknowledge the Diamond Light Source (MX1229) for access to beamline I04. We thank Karl Burgess and ThermoScientific for the use of their LipidSearch software. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We declare that no competing interests exist.

FUNDING

This work was supported by the Wellcome Trust [grant number 083625 (to M.W.K., A.C., B.O.S. and B.C.)]; the National Research Council of Argentina (to M.F.R.-B., M.I.-S., G.R.F. and B.C.O.); the European Molecular Biology Organization short term fellowship [grant number ASTF 12-2012 (to M.F.R.-B.)]; and the Biotechnology and Biological Sciences Research Council [grant numbers (BB/G011389/1 (to A.J.R. and M.G.)).

REFERENCES

- Storch, J. and Corsico, B. (2008) The emerging functions and mechanisms of mammalian fatty acid-binding proteins. *Annu. Rev. Nutr.* **28**, 73–95 [CrossRef PubMed](#)
- Meenan, N.A., Ball, G., Bromek, K., Uhrin, D., Cooper, A., Kennedy, M.W. and Smith, B.O. (2011) Solution structure of a repeated unit of the ABA-1 nematode polyprotein allergen of ascaris reveals a novel fold and two discrete lipid-binding sites. *PLoS Negl. Trop. Dis.* **5**, e1040 [CrossRef PubMed](#)
- Kennedy, M.W. (2000) The polyprotein lipid binding proteins of nematodes. *Biochim. Biophys. Acta* **1476**, 149–164 [CrossRef PubMed](#)
- Kennedy, M.W., Allen, J.E., Wright, A.S., McCrudden, A.B. and Cooper, A. (1995) The Gp15/400 polyprotein antigen of *Brugia malayi* binds fatty-acids and retinoids. *Mol. Biochem. Parasitol.* **71**, 41–50 [CrossRef PubMed](#)
- Kennedy, M.W., Brass, A., McCrudden, A.B., Price, N.C., Kelly, S.M. and Cooper, A. (1995) The ABA-1 allergen of the parasitic nematode *Ascaris suum* - fatty-acid and retinoid-binding function and structural characterization. *Biochemistry* **34**, 6700–6710 [CrossRef PubMed](#)
- Kennedy, M.W., Britton, C., Price, N.C., Kelly, S.M. and Cooper, A. (1995) The DVA-1 polyprotein of the parasitic nematode *Dictyocaulus viviparus* - a small helix-rich lipid-binding protein. *J. Biol. Chem.* **270**, 19277–19281 [CrossRef PubMed](#)
- Moore, J., McDermott, L., Price, N.C., Kelly, S.M., Cooper, A. and Kennedy, M.W. (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 [CrossRef PubMed](#)
- Mei, B., Kennedy, M.W., Beauchamp, J., Komuniecki, P.R. and Komuniecki, R. (1997) Secretion of a novel, developmentally regulated fatty acid-binding protein into the perivitelline fluid of the parasitic nematode, *Ascaris suum*. *J. Biol. Chem.* **272**, 9933–9941 [CrossRef PubMed](#)
- Plenefisch, J., Xiao, H., Mei, B.S., Geng, J.M., Komuniecki, P.R. and Komuniecki, R. (2000) Secretion of a novel class of iFABPs in nematodes: coordinate use of the *Ascaris/Caenorhabditis* model systems. *Mol. Biochem. Parasitol.* **105**, 223–236 [CrossRef PubMed](#)
- Garofalo, A., Rowlinson, M.C., Amambua, N.A., Hughes, J.M., Kelly, S.M., Price, N.C., Cooper, A., Watson, D.G., Kennedy, M.W. and Bradley, J.E. (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 [CrossRef PubMed](#)
- Jones, J.T., Kumar, A., Pylpenko, L.A., Thirugnanasambandam, A., Castelli, L., Chapman, S., Cock, P.J., Grenier, E., Lilley, C.J., Phillips, M.S. and Blok, V.C. (2009) Identification and functional characterization of effectors in expressed sequence tags from various life cycle stages of the potato cyst nematode *Globodera pallida*. *Mol. Plant. Pathol.* **10**, 815–828 [CrossRef PubMed](#)
- Popeijus, M., Blok, V.C., Cardle, L., Bakker, E., Phillips, M.S., Helder, J., Smart, G. and Jones, J.T. (2000) Analysis of genes expressed in second stage juveniles of the potato cyst nematodes *Globodera rostochiensis* and *G. pallida* using the expressed sequence tag approach. *Nematology* **2**, 567–574 [CrossRef](#)
- Prior, A., Jones, J.T., Blok, V.C., Beauchamp, J., McDermott, L., Cooper, A. and Kennedy, M.W. (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 [CrossRef PubMed](#)
- Ranjit, N., Jones, M.K., Stenzel, D.J., Gasser, R.B. and Loukas, A. (2006) A survey of the intestinal transcriptomes of the hookworms, *Necator americanus* and *Ancylostoma caninum*, using tissues isolated by laser microdissection microscopy. *Int. J. Parasitol.* **36**, 701–710 [CrossRef PubMed](#)
- Basavaraju, S.V., Zhan, B., Kennedy, M.W., Liu, Y., Hawdon, J. and Hotez, P.J. (2003) Ac-FAR-1, a 20 kDa fatty acid- and retinol-binding protein secreted by adult *Ancylostoma caninum* hookworms: gene transcription pattern, ligand binding properties and structural characterisation. *Mol. Biochem. Parasitol.* **126**, 63–71 [CrossRef PubMed](#)
- McCarter, J.P., Mitreva, M.D., Martin, J., Dante, M., Wylie, T., Rao, U., Pape, D., Bowers, Y., Theising, B., Murphy, C.V. et al. (2003) Analysis and functional classification of transcripts from the nematode *Meloidogyne incognita*. *Genome Biol.* **4**, R26 [CrossRef PubMed](#)
- Vanholme, B., Mitreva, M., Van Criekinge, W., Logghe, M., Bird, D., McCarter, J.P. and Gheysen, G. (2006) Detection of putative secreted proteins in the plant-parasitic nematode *Heterodera schachtii*. *Parasitol. Res.* **98**, 414–424 [CrossRef PubMed](#)
- Ramanathan, R., Varma, S., Ribeiro, J.M., Myers, T.G., Nolan, T.J., Abraham, D., Lok, J.B. and Nutman, T.B. (2011) Microarray-based analysis of differential gene expression between infective and noninfective larvae of *Strongyloides stercoralis*. *PLoS Negl. Trop. Dis.* **5**, e1039 [CrossRef PubMed](#)
- Fairfax, K.C., Vermeire, J.J., Harrison, L.M., Bungiro, R.D., Grant, W., Husain, S.Z. and Cappello, M. (2009) Characterisation of a fatty acid and retinol binding protein orthologue from the hookworm *Ancylostoma ceylanicum*. *Int. J. Parasitol.* **39**, 1561–1571 [CrossRef PubMed](#)
- Burbeho, P.D., Leahy, H.P., Iadarola, M.J. and Nutman, T.B. (2009) A four-antigen mixture for rapid assessment of *Onchocerca volvulus* infection. *PLoS Negl. Trop. Dis.* **3**, e438 [CrossRef PubMed](#)
- Kennedy, M.W., Garside, L.H., Goodrick, L.E., McDermott, L., Brass, A., Price, N.C., Kelly, S.M., Cooper, A. and Bradley, J.E. (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 [CrossRef PubMed](#)
- Iberkleid, I., Vieira, P., de Almeida Engler, J., Frester, K., Spiegel, Y. and Horowitz, S.B. (2013) Fatty acid- and retinol-binding protein, Mj-FAR-1 induces tomato host susceptibility to root-knot nematodes. *PLoS One* **8**, e64586 [CrossRef PubMed](#)
- Lal, P.G. and James, E.R. (1996) Onchocerca retinol- and ivermectin-binding protein activity. *Parasitology* **112** Pt 2, 221–225 [CrossRef PubMed](#)
- Jordanova, R., Groves, M.R., Kostova, E., Woltersdorf, C., Liebau, E. and Tucker, P.A. (2009) Fatty acid- and retinoid-binding proteins have distinct binding pockets for the two types of cargo. *J. Biol. Chem.* **284**, 35818–35826 [CrossRef PubMed](#)
- Hotez, P.J., Alvarado, M., Basanez, M.-G., Bolliger, I., Bourne, R., Boussinesq, M., Brooker, S.J., Brown, A.S., Buckle, G., Budke, C. M. et al. (2014) The Global Burden of Disease Study 2010: interpretation and implications for the neglected tropical diseases. *PLoS Negl. Trop. Dis.* **8**, e2865 [CrossRef PubMed](#)
- Pullan, R.L., Smith, J.L., Jasrasaria, R. and Brooker, S.J. (2014) Global numbers of infection and disease burden of soil transmitted helminth infections in 2010. *Parasit. Vectors* **7**, 37 [CrossRef PubMed](#)
- Rey-Burusco, M.F., Ibanez-Shimabukuro, M., Cooper, A., Kennedy, M.W., Corsico, B. and Smith, B.O. (2014) ¹H, ¹³C and ¹⁵N chemical shift assignments of Na-FAR-1, a helix-rich fatty acid and retinol binding protein of the parasitic nematode *Necator americanus*. *Biomol. NMR Assign.* **8**, 19–21 [CrossRef PubMed](#)
- Gabrielsen, M., Rey-Burusco, M.F., Griffiths, K., Roe, A.J., Cooper, A., Smith, B.O., Kennedy, M.W. and Corsico, B. (2012) Two crystal forms of a helix-rich fatty acid- and retinol-binding protein, Na-FAR-1, from the parasitic nematode *Necator americanus*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **68**, 835–838 [CrossRef PubMed](#)
- Williamson, A.L., Brindley, P.J., Abbenante, G., Prociw, P., Berry, C., Girdwood, K., Pritchard, D.I., Fairlie, D.P., Hotez, P.J., Dalton, J.P. and Loukas, A. (2002) Cleavage of hemoglobin by hookworm cathepsin D aspartic proteases and its potential contribution to host specificity. *FASEB J* **16**, 1458–1460 [PubMed](#)
- Kabsch, W. (2010) Xds. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 [CrossRef](#)
- Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 72–82 [CrossRef](#)
- Pajankar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S. and Tucker, P.A. (2005) Auto-Rickshaw - an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61**, 760–763
- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A. et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 [CrossRef](#)
- Sheldrick, G.M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 479–485 [CrossRef](#)
- Schneider, T.R. and Sheldrick, G.M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58**, 1772–1779 [CrossRef](#)
- Hao, Q. (2004) ABS: a program to determine absolute configuration and evaluate anomalous scatterer substructure. *J. Appl. Crystallogr.* **37**, 498–499 [CrossRef](#)
- Perrakis, A., Morris, R. and Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6**, 458–463 [CrossRef PubMed](#)
- Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 [CrossRef](#)
- Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Vonrhein, C. and Womack, T.O. (2009) In BUSTER version 2.8.0. Global Phasing Ltd, Cambridge

- 40 Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, III, W.B., Snoeyink, J., Richardson, J.S. and Richardson, D.C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375–W383 [CrossRef PubMed](#)
- 41 Hansen, M.R., Mueller, L. and Pardi, A. (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5**, 1065–1074 [CrossRef PubMed](#)
- 42 Laue, E.D., Mayger, M.R., Skilling, J. and Staunton, J. (1986) Reconstruction of phase-sensitive two-dimensional NMR spectra by maximum entropy. *J. Magn. Reson.* **68**, 14–29
- 43 Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J. and Laue, E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 [CrossRef PubMed](#)
- 44 Ottiger, M., Delaglio, F. and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **131**, 373–378 [CrossRef PubMed](#)
- 45 Ball, G., Meenan, N., Bromek, K., Smith, B.O., Bella, J. and Uhrin, D. (2006) Measurement of one-bond $^{13}\text{C}\alpha\text{-}^1\text{H}\alpha$ residual dipolar coupling constants in proteins by selective manipulation of $\text{C}\alpha\text{H}\alpha$ spins. *J. Magn. Reson.* **180**, 127–136 [CrossRef PubMed](#)
- 46 Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989) Overcoming the overlap problem in the assignment of ^1H NMR spectra of larger proteins by use of three-dimensional heteronuclear $^1\text{H}\text{-}^{15}\text{N}$ Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. *Biochemistry* **28**, 6150–6156 [CrossRef PubMed](#)
- 47 Nilges, M. (1995) Calculation of protein structures with ambiguous distance restraints - automated assignment of ambiguous NOE crosspeaks and disulfide connectivities. *J. Mol. Biol.* **245**, 645–660 [CrossRef PubMed](#)
- 48 Habeck, M., Rieping, W., Linge, J.P. and Nilges, M. (2004) NOE assignment with ARIA 2.0: the nuts and bolts. *Method Mol. Biol.* **278**, 379–402
- 49 Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S. et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921
- 50 Cheung, M.S., Maguire, M.L., Stevens, T.J. and Broadhurst, R.W. (2010) DANGLE: a Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J. Magn. Reson.* **202**, 223–233 [CrossRef PubMed](#)
- 51 Zweckstetter, M. and Bax, A. (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR. *J. Am. Chem. Soc.* **122**, 3791–3792 [CrossRef](#)
- 52 Tjandra, N., Omichinski, J.G., Gronenborn, A.M., Clore, G.M. and Bax, A. (1997) Use of dipolar H-1-N-15 and H-1-C-13 couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat. Struct. Biol.* **4**, 732–738 [CrossRef PubMed](#)
- 53 Linge, J.P., Williams, M.A., Spronk, C.A., Bonvin, A.M.J.J. and Nilges, M. (2003) Refinement of protein structures in explicit solvent. *Proteins* **50**, 496–506 [CrossRef PubMed](#)
- 54 Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 [CrossRef PubMed](#)
- 55 Farrow, N.A., Zhang, O.W., Formankay, J.D. and Kay, L.E. (1994) A heteronuclear correlation experiment for simultaneous determination of N-15 longitudinal decay and chemical-exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4**, 727–734 [CrossRef PubMed](#)
- 56 Kay, L.E., Torchia, D.A. and Bax, A. (1989) Backbone dynamics of proteins as studied by N-15 inverse detected heteronuclear Nmr-spectroscopy - application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 [CrossRef PubMed](#)
- 57 Kay, L.E., Xu, G.Y. and Yamazaki, T. (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H₂O saturation. *J. Magn. Reson. A* **109**, 129–133 [CrossRef](#)
- 58 Folch, J., Lees, M. and Sloane Stanley, G.H. (1957) A simple method for the isolation and purification of total lipides from animal tissues. *J. Biol. Chem.* **226**, 497–509 [PubMed](#)
- 59 Obal, G., Ramos, A.L., Silva, V., Lima, A., Batthyany, C., Bessio, M.I., Ferreira, F., Salinas, G. and Ferreira, A.M. (2012) Characterisation of the native lipid moiety of *Echinococcus granulosus* antigen B. *PLoS Negl. Trop. Dis.* **6**, e1642 [CrossRef PubMed](#)
- 60 Morrison, W.R. and Smith, L.M. (1964) Preparation of fatty acid methyl esters and dimethylacetals from lipids with boron fluoride-methanol. *J. Lipid. Res.* **5**, 600–608 [PubMed](#)
- 61 Maté, S.M., Brenner, R.R. and Ves-Losada, A. (2004) Phosphatidyl choline fatty acid remodeling in the hepatic cell nuclei. *Prostaglandins Leukot. Essent. Fatty Acids* **70**, 49–57 [CrossRef PubMed](#)
- 62 Daub, J., Loukas, A., Pritchard, D.I. and Blaxter, M. (2000) A survey of genes expressed in adults of the human hookworm, *Necator americanus*. *Parasitology* **120** Pt 2, 171–184 [CrossRef PubMed](#)
- 63 Bradley, J.E., Nirmalan, N., Kläger, S.L., Faulkner, H. and Kennedy, M.W. (2001) River blindness: a role for parasite retinoid-binding proteins in the generation of pathology? *Trends Parasitol.* **17**, 471–475 [CrossRef PubMed](#)
- 64 Tang, Y.T., Gao, X., Rosa, B.A., Abubucker, S., Hallsworth-Pepin, K., Martin, J., Tyagi, R., Heizer, E., Zhang, X., Bhonagiri-Palsikar, V. et al. (2014) Genome of the human hookworm *Necator americanus*. *Nat. Genet.* **46**, 261–269 [CrossRef PubMed](#)
- 65 Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**, 1–6 [CrossRef PubMed](#)
- 66 Cantacessi, C., Mitreva, M., Jex, A.R., Young, N.D., Campbell, B.E., Hall, R.S., Doyle, M.A., Ralph, S.A., Rabelo, E. M., Ranganathan, S. et al. (2010) Massively parallel sequencing and analysis of the *Necator americanus* transcriptome. *PLoS Negl. Trop. Dis.* **4**, e684 [CrossRef PubMed](#)
- 67 Garofalo, A., Klager, S.L., Rowlinson, M.C., Nirmalan, N., Klion, A., Allen, J.E., Kennedy, M.W. and Bradley, J.E. (2002) The FAR proteins of filarial nematodes: secretion, glycosylation and lipid binding characteristics. *Mol. Biochem. Parasitol.* **122**, 161–170 [CrossRef PubMed](#)
- 68 Garofalo, A., Kennedy, M.W. and Bradley, J.E. (2003) The FAR proteins of parasitic nematodes: their possible involvement in the pathogenesis of infection and the use of *Caenorhabditis elegans* as a model system to evaluate their function. *Med. Microbiol. Immunol.* **192**, 47–52 [PubMed](#)
- 69 Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. and Barton, G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 [CrossRef PubMed](#)
- 70 Notredame, C., Higgins, D.G. and Heringa, J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 [CrossRef PubMed](#)

Received 16 January 2015/12 August 2015; accepted 27 August 2015
 Accepted Manuscript online 28 August 2015, doi:10.1042/BJ20150068