

Review

The polyprotein lipid binding proteins of nematodes

Malcolm W. Kennedy *

Division of Infection and Immunity, Institute of Biomedical and Life Sciences, Joseph Black Building, University of Glasgow, Glasgow, UK

Received 23 September 1999; received in revised form 16 November 1999; accepted 8 December 1999

Abstract

The nematode polyprotein allergens/antigens (NPAs) are specific to nematodes, and are synthesised as tandemly repetitive polypeptides comprising 10 or more repeated units. The polyproteins are post-translationally cleaved at consensus sites to yield multiple copies of the approximately 15-kDa NPA units. These units can be highly diverse in their amino acid sequences, but absolutely conserved signature amino acid positions are identifiable. NPA units are helix-rich and possibly fold as four helix bundle proteins. The NPA units have relatively non-specific lipid binding activities, binding fatty acids and retinoids, with dissociation constants similar to those of lipid transport proteins of vertebrates. Fluorescence-based analysis has indicated that, like most lipid transport proteins, the ligand is taken into the binding site in its entirety, but the binding site environment is unusual. NPAs are synthesised in the gut of nematodes, and presumably act to distribute small lipids from the gut, via the pseudocoelomic fluid, to consuming tissues (muscles, gonads, etc.). In some species, one of the units has a histidine-rich extension peptide which binds haems and certain divalent metal ions. NPAs appear to be released by parasitic nematodes, and may thereby be involved in modification of the local inflammatory and immunological environment of the tissues they inhabit by delivering or sequestering pharmacologically active lipids – they are known to bind arachidonic acids and some of its metabolites, lysophospholipids, and retinoids. NPAs are the only known lipid binding protein made as polyproteins, and are exceptions to the rule that repetitive polyproteins are only produced by cells undergoing programmed cell death and producing specialist products. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Nematode polyprotein allergen; Nematode polyprotein antigen; Lipid binding; Fatty acid; Retinol; Polyprotein

1. Introduction: if all matter were swept away ...

Nematodes are a highly abundant and ecologically diverse group of multicellular, eucaryotic organisms. They are worm-like, but not related to the segmented annelid worms, flukes, tapeworms or other flatworms. The vast majority of nematode species are small, often microscopic, but large species (up to a

metre long) occur as parasites of vertebrates, such as ourselves. Their abundance can be illustrated by the following statistics: inter-tidal marine muddy sands, up to 5 million nematodes per square metre; arable land, up to 6 billion per acre; roots of a single potato, up to 40 000. An appreciation of their abundance and ubiquity led in 1915 to the comment [1]:

If all the matter in the universe except the nematodes were swept away, our world would still be dimly recognisable, and if, as disembodied spirits, we could investigate it, we should find its mountains, hills, vales, rivers, lakes and oceans represented by a film of nematodes ...

* Fax: +44-141-330-4600;
E-mail: malcolm.kennedy@bio.gla.ac.uk

In the case of parasitic nematodes, probably all species of plant and animal would be identifiable by their own particular species of nematodes. In some tropical areas, even the outline of entire human beings would be discernible through the presence of millions of larval forms in the skin. It is for their clinical and agricultural impact that nematodes are most infamous, but a free-living species, *Caenorhabditis elegans*, is the first multicellular organism to have its genome sequenced [2]. This exercise has provided us with many new insights into biology, not the least of which is the identification of many novel type of proteins. One of these is the subject of this review.

The polyprotein allergens/antigens of nematodes (NPAs) represent one of two novel classes of lipid binding protein which have recently been described from this abundant and diverse group of organisms [3–5]. NPAs are small, helix-rich proteins, and have no known structural counterparts in other phyla, including the animals and plants in which parasitic nematodes cause so much harm. The NPAs are therefore of intrinsic interest because of their novel structure and ligand binding properties, but also for their potential relevance to the success of nematode parasitism. They are also produced by free-living nematodes, so may represent the adaptation of an existing system for parasitism.

2. Polyproteins

The term polyprotein is used in the literature to denote two quite different types of entity. The first is exemplified by viral proteins and some vertebrate hormone proproteins in which a precursor protein is cleaved into polypeptides with different functions [6–10]. The second type is typified by certain proteins of vertebrate keratinocytes in which a large precursor containing tandemly repeated units is cleaved post-translationally into polypeptides which are presumed (in the absence of direct testing) to have similar or identical functions [11–15]. There will presumably also be cases of this type in which the different units have diverged such that they now serve different biochemical functions yet reflect their antecedents in primary, secondary and tertiary structures. The

NPAs fall into this tandemly repetitive category of polyproteins.

The advantage in producing proteins in repetitive arrays is presumably that it presents a highly efficient means by which to produce large quantities of a functional product [13]. This is argued for the polyproteins of terminally differentiated mammalian keratinocytes such as filaggrin, involucrin and loricin, which comprise peptide or polypeptide repeats that display sequence variations yet retain distinctive structural motifs [11–15]. The filaggrin gene appears not to have any introns in the coding region, which represents a further economy and presumably means that the (evolutionary) advantage of introns is outweighed by their cost in this situation. The necessity for these economies is postulated to be that a maturing epidermal granular cell requires to mobilise rapidly, and precisely regulate, a large pool of cell type-specific proteins in the face of nuclear function which is compromised by approaching senescence [13]. The NPAs, however, appear to break the rule that repetitive polyproteins with (apparently) intronless coding regions are confined to products of a senescing cell; the somatic cells of nematodes are thought not to undergo division once development has been completed, and apoptotic sacrifice of cells is unlikely once late stages of development are reached, although it does occur at early stages [16,17].

Approximately 58% of the proteins predicted from the genome of *C. elegans* encode proteins with no identifiable homologues in other groups of organisms [2]. Only two classes amongst these apparently nematode-specific proteins have known biochemical activities, the NPAs and an unrelated family of small (~20-kDa) lipid binding proteins [5,18].

3. Nematode proteins and allergens

Parasitic nematodes tend to elicit high levels of IgE antibody, which is the immunoglobulin class involved in type I hypersensitivity reactions, such as asthma and food allergies. This does not necessarily mean, though, that infected people will suffer from hypersensitivity responses due to the infections [19], but life-threatening reactions occur in certain cases, such as in infection with the large roundworm of

humans, *Ascaris lumbricoides* [20,21]. IgE antibody, however, also appears to associate with protective immune responses to worm parasites in humans [22–26], prompting increasing interest in the targets of IgE antibody (allergens by definition) because of their potential role in pathology and protective immunity. Most effort has concentrated on environmental (pollens, dust mites, hymenopteran venoms) and food allergens because of their role in allergic conditions occurring in developed countries. No common structural or biochemical feature of protein allergens has yet been identified which can explain their allergenicity, so other aspects of sensitisation to allergens are probably crucial to the initiation of hypersensitivity reactions. Possibilities include the site or route of exposure, co-exposure with irritants or infections, combined with the genetics of an allergic individual with respect to tendency to high IgE responses [27].

The concern about allergens arises from the increasing occurrence of asthma and other conditions in developed countries. The IgE antibody response involved in these conditions appears, however, to be a type of immune response which has evolved as protection against multicellular eucaryotic parasites, such as nematodes and other worms, and arthropods, such as ticks [28,29]. In the case of many nematode infections, it is now well established that responses dominated by the type 2 subset of helper T-cells (Th2) are crucial to immune elimination of the infections in experimental animals [28–31]. There is debate about whether worm infections may protect against or exacerbate allergies [27,32–33]. The frequency of allergies in areas endemic for worm infections is dramatically lower than in developed countries, leading to the idea that the immune system is unbalanced in the absence of worm infections, and allergies are the consequence. It is also argued that the removal of infections (such as tuberculosis) which provoke immune responses of a type counter-regulatory to Th2 responses is also a contributory factor to the allergy epidemic in developed countries [27,34].

There seem to be two distinct routes whereby a parasite product becomes the target of IgE responses. The first is self-allergenicity by which the proteins are intrinsically allergenic (however that may be) and are capable of direct induction of IgE responses (for example, [35]). So far, very few exam-

ples of this have been demonstrated satisfactorily. The second is potentiation of IgE responses. This is a well-described phenomenon in nematode infections by which any foreign proteins which are presented to the immune system at the time of infection will themselves become bystander targets of IgE antibody [36,37]. This appears to be the case for the NPAs, which are strong allergens in the context of worm infection [38–41], but do not appear to be intrinsically allergenic ([42], and unpublished).

4. Type specimen: the ABA-1 allergen of *Ascaris*

The first of the NPAs to be described was the 14.4-kDa ABA-1 allergen protein of the large roundworm of pigs (*Ascaris suum*) and humans (*Ascaris lumbricoides*). The latter parasite is estimated to infect about one-quarter of humanity, children being most at risk of injury by these large worms (each female worm ~8 g in weight, 200–400 mm long, 3–6 mm diameter) parasitising the intestine. The early developmental stages of *Ascaris* migrate through the lungs, and hypersensitivity reactions to this phase of the infection can be lethal [20,21]. ABA-1 was found to be secreted by the tissue-invasive stages of the parasites and it is also abundant within the adult worms [38,43–45,46]. The protein was found to be a potent allergen [38,40,41] and, curiously, only certain people are capable of responding to it [23,47,48]. This latter effect is due to the inheritance of responder or non-responder alleles of the class II region of the major histocompatibility complex in experimental animals [38,40–42,47,49], and this genetic control of the immune repertoire to ABA-1 is also likely to apply in humans. While work on ABA-1 was underway, other groups working on nematodes which cause elephantiasis in humans (*Brugia malayi*) had characterised a protein of similar size to ABA-1 which appeared to be close to the surface of the parasites [50,51]. It was from these parasites and the dog heartworm (*Dirofilaria immitis*) that cDNAs encoding NPAs were first cloned, and their sequences led to the realisation that the proteins are related to ABA-1 [50–52], whose N-terminal amino acid sequence had by then appeared [41]. Cloning of ABA-1 cDNA followed closely [53], and cDNA sequences encoding related proteins now appear regu-

larly in the databases from screening of cDNA expression libraries with antibody from infected humans or domestic animals (for example [54–57]).

5. Precursors and processing of nematode polypeptides

In the majority of cases, cDNAs encoding only fragments of the NPA proprotein have been isolated. This is possibly because the parent mRNAs are large and consequently are poorly represented intact in cDNA expression libraries. The isolation of central sequences may additionally be due to the fact that the encoded proteins tend to be lysine-rich, so cDNA synthesis may be initiated in adenosine-rich internal codons rather than the polyA tails intended. The effect may be exaggerated by the AT-richness of the genomes of many species of nematode.

Despite their fragmentary nature, the first NPA cDNAs to be obtained clearly encoded tandemly repetitive polypeptides [50–52]. Upon alignment with the N-terminal amino acid sequence of the ABA-1 protein, a consensus processing cleavage site was identified immediately N-terminal of the beginning of each repeating unit (Figs. 1 and 2). This cleavage site comprised the $^{Lys}/_{Arg}$ -Xaa- $^{Lys}/_{Arg}$ -Arg motif of processing endoproteinases [59], and, in the case of *Ascaris* and *C. elegans*, the site is commonly a tetrad of arginines. The tetrad is probably removed after

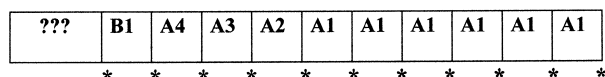
the cleavage event because mass spectroscopy of parasite-derived ABA-1 gives a mass of $14\,643.2 \pm 1.4$, which, within error, corresponds to the calculated mass (14 642.6) of a single unit of the ABA-1 array minus the arginine tetrad [60].

Biosynthetic labelling work on living nematodes had shown that the nascent polypeptides of NPAs are very large, comprising perhaps up to 20 repeats of the ~ 15 -kDa units, each unit encoding an ABA-

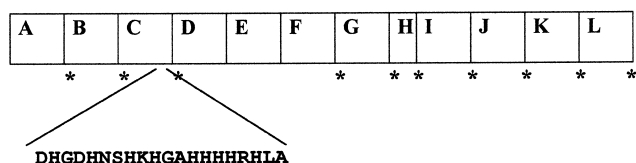
←

Fig. 1. Organisation of the polypeptide from three species of nematode. Only the central, repetitive, part of the polypeptides are illustrated, ignoring the short hydrophobic N-terminal (signal) and C-terminal extension peptides. The full structure of the NPA of *D. viviparus* is known from cDNA analysis [57], and the complete genomic sequence for the latter derives from the *C. elegans* genome project (cosmids VC5 and F27B10). In the case of the *A. suum* NPA, which is the next best understood, the 5' region is unresolved. For the NPAs of other species, only fragments of the central region are known. The positions of the consensus $^{Lys}/_{Arg}$ -Xaa- $^{Lys}/_{Arg}$ -Arg cleavage sites are marked (*). In some species, including *D. viviparus*, no consensus cleavage site is evident between some of the units, suggesting either that another type of enzyme is involved in processing or that double or triple units are produced without further processing – there is some evidence for the latter possibility in that intact multimers have been indicated from immunoblotting experiments using antisera raised against recombinant NPA [57,58]. The histidine-rich stretches which are found in *D. viviparus* and *C. elegans* are positioned as indicated. The positions of the small (40–396-bp) introns in the genomic DNA encoding the *C. elegans* NPA are indicated (V). For complete sequences, the numbering system for the units is alphabetical from the N-terminus. The *Ascaris* NPA units were named according to order of discovery and similarity, and should be renamed to conform to the convention for the other two when sequenced completely. In this NPA, the A1–A4 units are grouped because of their closely similar amino acid sequences, which are only $\sim 50\%$ similar to the single, abruptly different, B1 unit. All the A1 units have identical amino acid sequences, although with slight differences in the encoding DNA. The units within the NPA arrays of *D. viviparus* and *C. elegans* are quite different from one another (see Fig. 2), but those of *Ascaris* species and certain filarial nematodes (e.g. *Brugia malayi*) are more similar [51,53]. The (poly)proteins of *A. suum*, and *D. viviparus* were originally designated ABA-1 and DvA-1 [3,4,46,57], respectively, but under suggested rules for the naming of nematode genes and their products, the polypeptide should now be designated As-NPA, and Dv-NPA, and the units thereof given as suffixes, Dv-NPA-A, Dv-NPA-B, etc. (see Table 1). The genes encoding the polypeptides should be named *ce-npa-1*, *dv-npa-1* and *as-npa-1* for *C. elegans*, *D. viviparus* and *A. suum*, respectively.

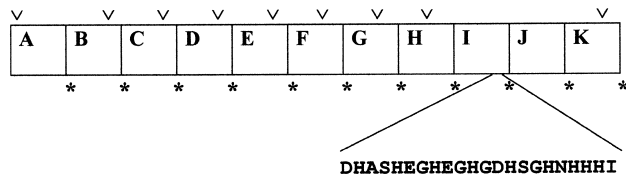
Ascaris suum polypeptide (ABA-1; As-NPA)



Dictyocaulus viviparus polypeptide (DvA-1; Dv-NPA)



Caenorhabditis elegans polypeptide (Ce-NPA)



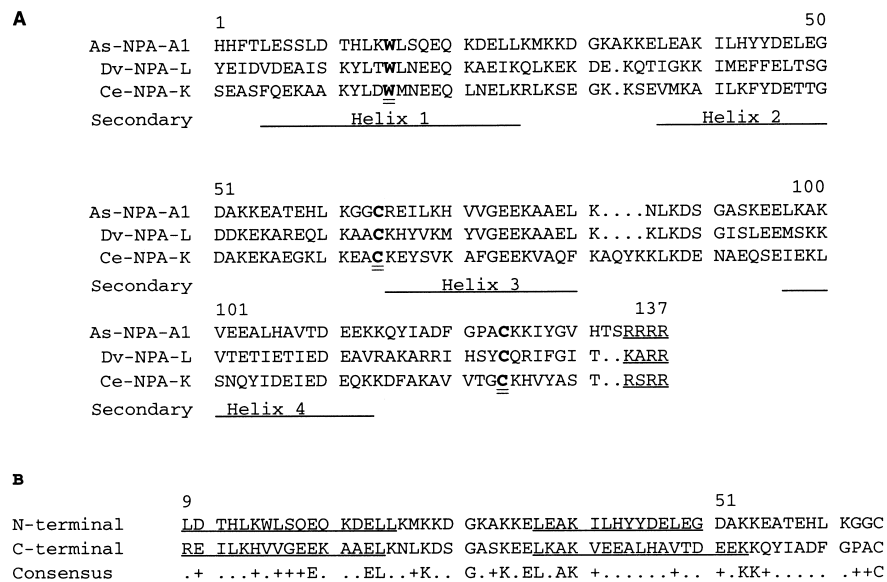


Fig. 2. (A) Alignment of the amino acid sequences of three NPA units to illustrate their diversity. The three selected are those encoding the most 3' repeat unit from the NPA array of three different species, *Ascaris suum* (As-NPA-A1, originally designated ABA-1A1), *Dictyocaulus viviparus* (Dv-NPA-L; DvA-1L) and *Caenorhabditis elegans* (Ce-NPA-K). The amino acids which are absolutely conserved in all NPA units from all species (with the single exception of the truncated unit H from *D. viviparus*) are indicated by double underlining. The terminal four amino acids which conform to the $^{Lys}_{Arg}$ -Xaa- $^{Lys}_{Arg}$ -Arg consensus cleavage site for processing enzymes [59], and which are probably removed after cleavage [60], are single underlined. The regions predicted to form α -helices are as indicated, and no β /extended structure has been predicted or detected by circular dichroism [3]. (B) Alignment of the N- and C-terminal halves of one unit (from *A. suum*) to illustrate the putative duplication which might have led to the modern ~ 15 -kDa units of the NPAs; a similar effect is observed for other NPA units. The consensus line shows where amino acids are identical (letters), or there are substitutions with amino acids of similar properties (+) using the groupings (single letter amino acid codes) GASTP, DEQN, C, VIML, KRH and FYW (grouped according to substitution likelihood; [61] and [62]). The predicted helical sections are underlined, and the numbering corresponds to that used for the alignment in A.

l-like protein. Thus, in a pulse-chase experiment, a very large product is observed initially, which is then progressively degraded to form a ladder of bands on SDS-PAGE, the rungs of which are approximately 15 kDa apart [51]. This process culminates in an abundant product of 15 kDa. Such a ladder is also observed in Western blots of parasite material separated on SDS-PAGE and probed with antibody from infected humans or animals, which is presumably catching polyprotein undergoing processing. In some cases, it appears that the polyproteins are not fully processed down to ~ 15 -kDa polypeptides, leaving products representing two or three units uncut [57,58]. cDNA analysis has shown that some units of the polyproteins of some species are indeed uninterrupted by consensus cleavage sites (Fig. 1, and [57,58]).

The size of an NPA array has been estimated from Southern blotting experiments to be between 10 and 20, or as many as 50, units, depending on the species

[50–52]. Complete cDNA sequence is available from one species, *Dictyocaulus viviparus* (a highly pathogenic pulmonary parasite of cattle) [57], and the *C. elegans* genome project has provided the genomic sequence from which the coding region of its NPA mRNA has been predicted (spread over cosmids VC5 and F27B10). The NPAs of these two species comprise 11 and 10 units, respectively, and in the only other species for which cDNA analysis has been pursued with vigour (*Ascaris*), only 10 contiguous units have been found, although the sequence is incomplete [63]. It might therefore be that the higher estimates of the number of units in an NPA may not survive closer examination.

6. Lipid binding

The biochemical function of the NPAs was first identified using protein purified from a parasite [3],

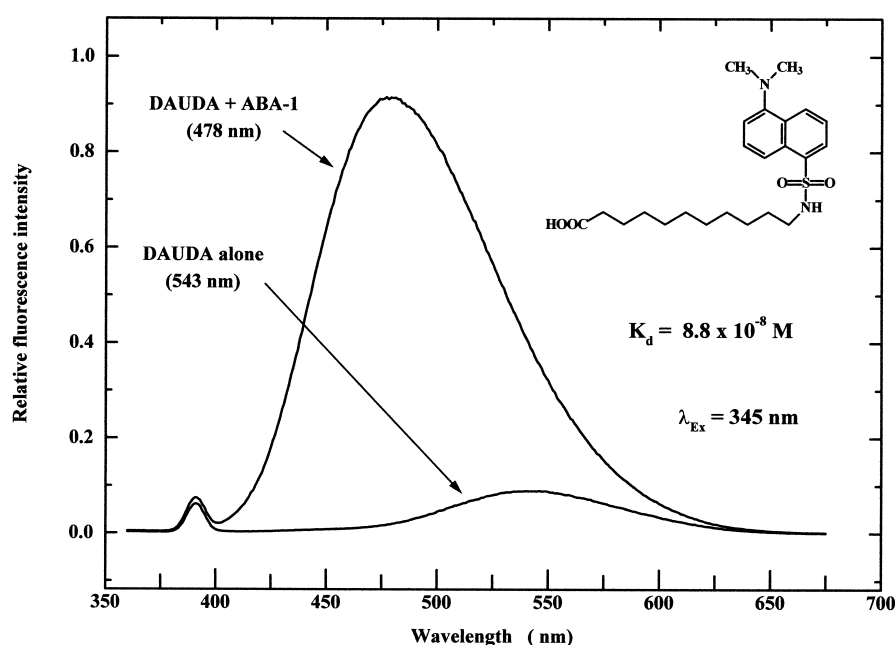


Fig. 3. Binding of fatty acids by the ABA-1 protein isolated directly from *Ascaris suum*. Fluorescence spectra of 11-((5-dimethylamino-naphthalene-1-sulphonyl)amino)undecanoic acid (DAUDA) in buffer alone and upon addition of ABA-1 protein to illustrate the dramatic increase in fluorescence intensity and the blue-shift in wavelength of peak emission. The change in emission spectrum is reversible by addition of/competition with non-fluorescent ligands. The structure of DAUDA is shown. (Data from [3]).

but only large nematodes can provide adequate protein for testing, and it would be difficult to obtain sufficient material from highly pathogenic species. Most biochemical and structural analysis has consequently been carried out on recombinant protein expressed in *E. coli*, using PCR to amplify DNA encoding single units of an NPA array from cDNA or genomic DNA [3,4,64,65].

Using fluorescence-based methods, all NPAs have been found to bind fatty acids and retinol in a 1:1 ligand:protein monomer ratio [3,4,63,64,65], and more conventional radio-isotope methods have confirmed this activity (K. Copeland, M.W. Kennedy and L. Fixter, unpublished). Dissociation constants estimated either in direct fluorescence titration experiments, or in competitive titration with natural, non-fluorescent, compounds are in the micromolar or sub-micromolar range. These values are similar to those obtained from other types of lipid transport proteins using similar methods [66,67].

The most useful fluorescent lipid in the study of NPAs has been 11-((5-dimethylaminonaphthalene-1-sulphonyl)amino)undecanoic acid (DAUDA), which bears the environment-sensitive dansyl fluorophore (Fig. 3). The extent of blue shift in fluorescence emis-

sion by this fluorophore is taken as an indicator of the degree of apolarity of the binding site and isolation of the bound ligand from solvent water [68]. In the case of NPAs, the degree of shift is unprecedented amongst fatty acid binding proteins, including liver FABP and serum albumin, which presumably means that the binding site of NPAs is highly unusual (Table 1). Indeed, the shift is almost as extreme as that occurring when the DAUDA probe is dissolved in organic solvents (Table 1 and [3]).

It is now known from X-ray crystallography, nuclear magnetic resonance and biochemical techniques that fatty acid binding proteins, with only a few exceptions [76], take their ligands entirely into their binding cavity, such that even the charged carboxylate is internalised [77–81]. For the NPAs, a similar behaviour has been inferred from the fact that the dansyl fluorophore undergoes a very similar blue shift in fluorescence emission whether it is attached at the α or at the ω /methyl carbon of a fatty acid [3,4,64]. Such full internalisation would be consistent with a role for the proteins in protecting their cargo from chemical degradation.

Binding retinol (vitamin A) by NPAs was also detected by fluorescence, and fatty acids compete for

Table 1
Comparative properties of NPAs and other small lipid binding proteins

Protein/solvent	New designation	Source organism	Natural/ recombinant	Fluorescence emission maxima (nm)		References
				DAUDA	DACA	
<i>Mammalian</i>						
Serum albumin		<i>Bos taurus</i>	natural	495	485	[3,69]
Serum albumin		<i>Homo sapiens</i>	natural	480		[69]
Intestinal-FABP		<i>Rattus norvegicus</i>	recombinant	492		[70]
Liver-FABP		<i>Rattus norvegicus</i>	recombinant	496		[71]
<i>NPAs</i>						
ABA-1	As-NPA	<i>Ascaris suum</i>	natural	475	483	[3]
ABA-1 unit A1	Al-NPA-A1	<i>Ascaris lumbricoides</i>	recombinant	475		[65]
ABA-1 unit A1	As-NPA-A1	<i>Ascaris suum</i>	recombinant	473	471	[63]
ABA-1 unit B	As-NPA-B1	<i>Ascaris suum</i>	recombinant	480	491	[63]
DvA-1 unit L	Dv-NPA-L	<i>Dictyocaulus viviparus</i>	recombinant	490	483	[4]
gp15/400	Bm-NPA	<i>Brugia malayi</i>	recombinant	478		[64]
Ce-1	Ce-NPA	<i>Caenorhabditis elegans</i>	recombinant	502		[72]
<i>Other nematode proteins</i>						
As-p18		<i>Ascaris suum</i>	recombinant	498	no binding	[73] and unpublished
Ov20		<i>Onchocerca volvulus</i>	recombinant	485	483	[5] and unpublished
<i>Tapeworm LBP</i>						
MeLBP		<i>Moniezia expansa</i>	natural	480	no binding	[74,75]
<i>Solvents</i>						
Water				543	543	[3] and unpublished
Ethanol				506	502	[3] and unpublished
Dimethylformamide				505	502	[3] and unpublished
Cyclohexane				475	469	[3] and unpublished

I-FABP, L-FABP and As-p18 all members of the FABP/P2/CRBP/CRABP family of β -barrel ~ 14 kDa, predominately intracellular, lipid transporter proteins. As-p18 is unusual in this family in that it has a hydrophobic leader sequence and is secreted [73]. Ov20 is a novel type of retinol and fatty acid binding protein which is unrelated to the NPAs, although, like the NPAs, it appears to be rich in α -helix structure and shows no evidence of β /extended structure [5]; family designation LBP-20 (lipid binding protein, ~ 20 Da). The 'DAUDA' and 'DACA' columns give the wavelength of peak emission of the dansylated fatty acids DAUDA or DACA upon binding to the proteins indicated. DAUDA binds in both the bilirubin (predominately) and fatty acid binding sites of serum albumin [69]. DAUDA, 11-((5-dimethylaminonaphthalene-1-sulphonyl)amino)undecanoic acid. DACA, dansyl-D,L- α -aminocaproic acid. The excitation wavelengths used were 345 or 350 nm. See legend to Fig. 1 for details of NPA (poly)protein and gene nomenclature.

binding, so either the retinol and fatty acid binding sites are the same, or they interact [3,4,63].

7. Biologically active lipids and nematocidal drugs

Most of the drugs commonly used against nematode infections in humans and animals are hydrophobic compounds which might need carrier proteins

to distribute them to their site of action within the target organisms. But, NPAs appear to bind none of the widely used anti-nematode drugs tested (Table 2).

Of more interest is the finding that NPAs bind several pharmacologically active lipids (Table 2), including arachidonic acid and some of its metabolites, lysophosphatidic acid, lyso-platelet activating factor, retinoic acid, and retinol itself. Given that NPAs appear to be released by living parasites [38,50,52,

Table 2
Ligand binding by NPAs

Binding	No binding
Oleic acid ($K_{app} = 0.72 \mu M$)	tocopherol
Retinol	tocopherol acetate
<i>cis</i> -Parinaric acid	β -carotene
<i>trans</i> -Parinaric acid	mebendazole ^b
Arachidonic acid	albendazole ^b
Lysophosphatidic acid ($K_{app} = 3.3 \mu M$)	thiabendazole ^b
Lysophosphatidyl ethanolamine	oxibendazole ^b
Lysophosphatidyl choline	piperazine ^b
Platelet-activating factor	tetramisole ^b
Lysoplatelet activating factor	pyrantel ^b
Leukotrienes B ₄ , C ₄ , D ₄ , E ₄ ^a	DEC ^b
Bilirubin ($K_{app} = 5.3 \mu M$)	levamisole ^b
DAUDA ($K_d = 1.37 \mu M$)	
DACA	

The binding studies were carried out either by competitive displacement of a bound fluorescent ligand (DAUDA), or by detecting changes in the fluorescence of natural ligands (e.g. retinol, parinaric acid). The dissociation constant (K_d) for protein:DAUDA interaction was calculated by fluorescence titration, and this value then used to estimate the apparent dissociation constant (K_{app}) for non-fluorescent compounds in competitive titration experiments [3]. DAUDA, 11-((5-dimethylaminonaphthalene-1-sulphonyl)amino)undecanoic acid; DACA, dansyl-D,L- α -aminocaproic acid; dansyl, dimethylaminonaphthalene-1-sulphonyl.

^aBinding with these compounds was detectable, but not readily quantifiable. Composite data are presented from natural and recombinant NPAs [3,4,63–65], but mostly based on work with recombinant ABA-1 variant A1 (AlABA-1A; Al-NPA-A1) protein from *Ascaris lumbricoides* [65].

^bThese compounds are widely used anti-nematode drugs.

57,82,83], the binding and potential sequestration of mediators of local inflammatory and immune reactions could pertain to the survival of the parasites in vertebrate tissues. Such compounds are also considered to influence the nature of the immune response systemically, so the secretion of NPAs by parasites could have wide implications. The relevance of this activity is further emphasised by the finding that some species of tissue-parasitic nematode absorb and modify host lipids for secretion as biologically active eicosanoids [84], for which they may use carrier proteins.

Because NPAs appear to have no structural counterparts in mammals, they may represent novel drug targets/nematocides. The relative non-specificity of their binding activities, however, would seem at this

stage to preclude the development of inhibitors/competitors which would not also affect mammalian proteins/systems. On the other hand, until their structure and ligand binding mechanisms are more fully understood, their potential as drug targets should not be ignored.

The binding of retinol may be of particular note because of the role this lipid has in tissue morphogenesis and repair. Sequestration of retinol, for instance, could be of significance to diseases such as river blindness which affects both the skin and the eyes. In the case of retinol, however, it is notable that nematodes produce, and in some cases secrete, another unusual type of binding protein which binds both retinol and fatty acids (the Ov20 protein, generically known as Ov-LBP-20; [5,18,85]).

8. Polyfunctional polyproteins?

A curious feature of the NPA of *D. viviparus*, which was not appreciated until the *C. elegans* genomic sequence became available, is the occurrence of a histidine-rich stretch of amino acids in the C-terminus of a single unit of the polyprotein array in each case. His-rich proteins are known to bind metals and/or haems, such as the nickel-binding protein of *E. coli* [86], the haem-polymerising protein of the malaria parasite *Plasmodium falciparum* [87], and the plasma His-rich protein of humans [88–90]. We have begun to examine the His-rich regions of NPAs using synthetic peptides and found that they indeed bind divalent cations (Ni, Cu, Zn, but not Ca) and haemin, and calorimetry studies have indicated that there are more than one binding site per peptide, with differing characteristics (M.W. Kennedy, M. Nutley, and A. Cooper, unpublished). The possibility is therefore that the His-rich peptides are involved in salvage or transport of haems or metals. This activity might be important to the biology of nematodes given, for instance, their possession of globins ('nemo-globins') for acquisition or sequestration of oxygen [91]; *Ascaris*, for instance, is notable in possessing a globin with the highest affinity for oxygen known for any such protein, presumably functioning to protect this microaerobic parasite from oxygen toxicity [92]. The breakdown of globin would release haem, which is itself toxic, and a means for its capture and safe

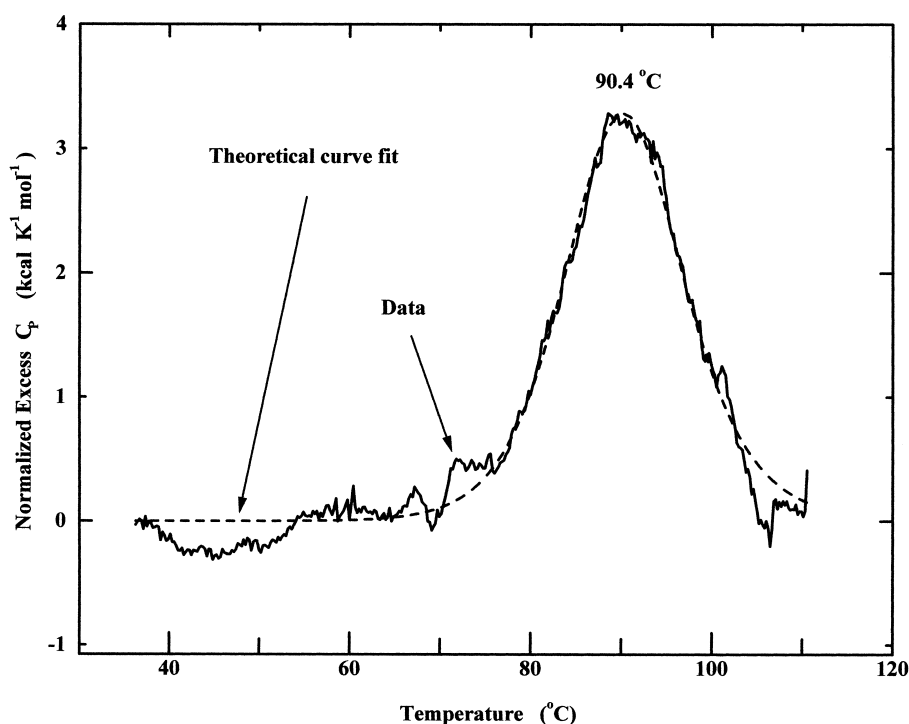


Fig. 4. Differential scanning calorimetry analysis of recombinant ABA-1 allergen from *Ascaris lumbricoides*, the large roundworm of humans. Protein solution (0.9 mg/ml, 2 ml) was scanned from 20 to 110°C yielding a single cooperative unfolding transition at 90.4°C. Reasonable numerical agreement between calorimetric ($\Delta H_{\text{cal}} = 58.0$ kcal/mol) and van 't Hoff enthalpies ($\Delta H_{\text{vH}} = 59.5$ kcal/mol) analysed assuming dimer concentration supports the idea that the protein unfolds cooperatively as a dimer [65]. Values for recombinant Dv-NPA-L protein (= DvA-1L) are similar [4].

disposal is important, which might be particularly so for blood-feeding nematodes [93]. Metal ion-binding histidine-rich proteins are also thought to be anti-microbial through their sequestration of essential ions for bacterial and fungal growth [94], so the His-rich segment of NPAs might act in anti-microbial defense of nematodes in addition to a known bactericidal protein present in nematodes [95].

9. Structural novelty

The binding site characteristics of NPA units marks them out as novel, but so too does their structural characteristics. At first site, their size and lipid binding activities would ally them to the beta-rich, predominately cytoplasmic, lipid binding proteins of the ~ 14 kDa FABP/P2/CRBP/CRABP family [96], but there are no sequence similarities between NPAs and these proteins. Moreover, secondary structure predictions and circular dichroism measurements in-

dicate a predominance of α -helix in NPAs, but no β -structure [3]. The predictions are that the proteins fall into four helical regions, so they may fold as four helix bundle proteins, such as is known for other invertebrate carrier proteins [97]. This therefore sets the NPAs apart from other proteins of their size and binding propensities, although plants also have small helix-rich fatty acid binding proteins (~ 9 kDa) [98] and the cytoplasmic acetyl-CoA binding protein is also small (~ 14 kDa) and helix-rich [99].

The predicted helices of the NPAs are all amphipathic, although there are certain highly conserved residues which are predicted to occur on faces of the helices contrary to their nature [3]. For example, the conserved Trp-15 is predicted to occur on a strongly hydrophilic face in ABA-1, yet Gln-20 appears on a hydrophobic face [3]. This should mean that Trp-15 is exposed to solvent and therefore exhibit a red-shifted fluorescence emission spectrum. Instead, Trp-15 has a strongly blue-shifted emission spectrum ($\lambda_{\text{max}} = 314$ nm) [3,4,63,64]. The Trps pre-

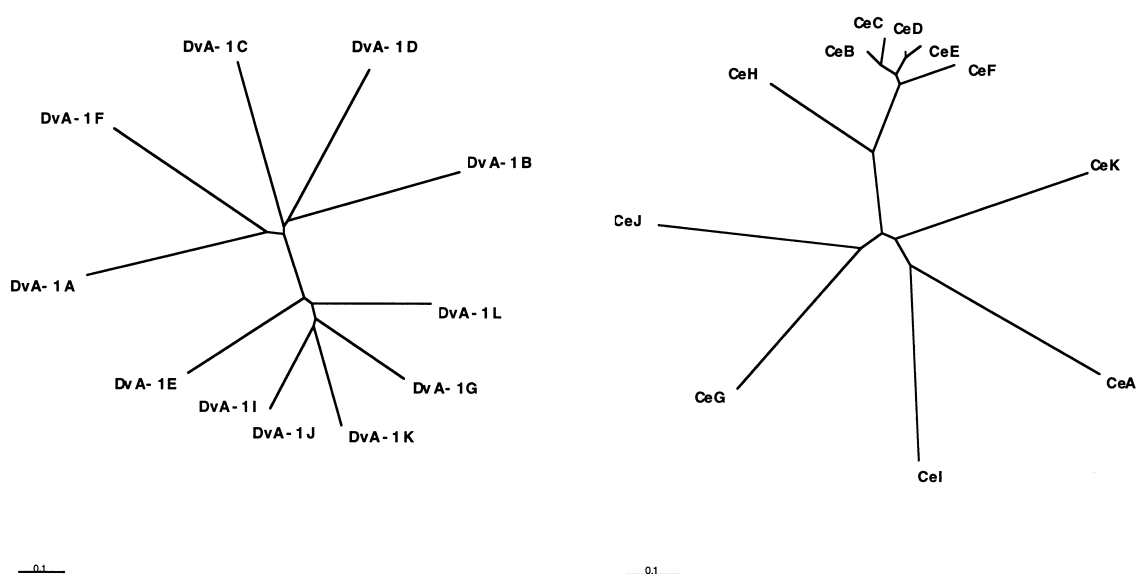


Fig. 5. Radial dendrograms to show the diversity and relationships between units in single NPA arrays. The dendrograms were constructed using ClustalW set for the neighbourhood joining method. In order to improve the search for relationships between the units, each was edited to remove the N-terminal sequences up to the conserved Trp (W15, Fig. 2), and beyond the second conserved Cys (C124, Fig. 2) because it is in these regions where the proteins appear to be most diverse and disrupt the analysis. This procedure also removes the His-rich regions. The truncated unit H from *D. viviparus* was also removed. The units towards the C-terminus of the polypeptide cluster together in the case of the *D. viviparus* NPA, but the opposite is the case for the *C. elegans* array. Those units clustering together may represent the products of more recent amplifications, with the more divergent representing the ancestral units. For the *C. elegans* NPA, therefore, it could be postulated that the more recent duplications/amplifications were centred around the B unit, but without involvement of the A unit. It might be of interest to note that the most divergent units in *C. elegans* are also those uninterrupted with introns, so the introns may be important in the amplification mechanism, or relics of it.

sent within or close to the binding sites of other proteins which bind fatty acids, such as serum albumin and lactoglobulin do not have such extreme blue shifts [3]. Mutagenesis of Trp-15 to Arg did not alter the ligand binding characteristics of recombinant ABA-1, although it did destabilise the protein in terms of increased susceptibility to denaturation with guanidine (L. McDermott, J. Moore, A. Cooper and M.W. Kennedy, unpublished, and [100]). Trp-15 is therefore unlikely to be in the ligand binding site, and may instead be either in an unusual environment within the folded protein, or involved in the interface of the dimers which the NPAs appear to form [3,4,46,63,64].

10. Quaternary structure

Differential scanning calorimetry (DSC) has been used to examine the stability of NPAs, but also to provide information on their quaternary structure

[3,4,65]. Using parasite-derived ABA-1 protein, it was initially found that the protein exists as a dimer (initially suspected from gel-permeation chromatography work [46]), and was observed to undergo a two-phase unfolding process [3]. Similar analysis of two different recombinant NPAs, however, showed only single transitions at very high temperatures ($T_m = 98$ or 90.4°C ; Fig. 4, [4,65]). The more complex transition curves using nematode-derived material may have been due to contamination with different types of proteins, and/or heterogeneity in the NPA proteins present. Recombinant material is probably more representative of the behaviour of a single species of NPA unit, and the data are consistent with the idea that the proteins unfold cooperatively as dimers. That is, the dimer association and the structures of the monomers are interdependent. In these experiments, only homodimers could form, but it is quite possible that heterodimers occur in vivo, and that this may influence their biochemical function and structures.

Dimerisation will increase the effective size of NPAs, which will presumably limit the protein's diffusability within, and loss from, the nematode. Of immunological importance, existence as a homodimer means that cross-linking of IgE receptors on mast cells will be facilitated, even where the number of epitopes on a monomer unit is limited. We have yet to establish, however, whether NPA proteins naturally exist as homodimers or as heterodimers with other members of the array from which they are derived, or whether dimerisation is necessary for their ligand binding activities.

It has previously been noted that the allergens of ascaridids are heat stable, retaining their allergenicity following cooking (nematodes commonly contaminate fish meat) [101] or even autoclaving [60]. This could be due to the IgE epitopes being resistant to heat denaturation (as, for example, would be simple linear peptide epitopes), and/or that the allergens are structurally stable at high temperatures or refold efficiently upon cooling. The calorimetry experiments on NPAs show that the latter is a likely explanation, so it may be that the NPAs are important in allergic reactions to ingested nematode materials [101].

11. Evolution

The sequence of a single unit of an NPA shows evidence of an ancient duplication event [53,63]. Thus, if one takes the first 60 amino acids or so of a given ~ 133 amino acid unit and aligns them with the remaining amino acids, then the two halves align easily (Fig. 2). This may mean that the progenitor of the NPAs was a small helical protein which was then duplicated to form the ~ 15 kDa NPA unit, and then this was in turn duplicated/amplified to produce the NPA array of today. Whether the small putative progenitor protein was itself a lipid binding protein is open to conjecture, but very small fatty acid binding proteins have been described from plants (~ 9 kDa) [98] and flatworms (~ 8 kDa) [74,75]. Also, we have noticed that a recombinant polypeptide representing the C-terminal half of one NPA unit binds DAUDA, albeit poorly [L. McDermott, J. Moore, A. Cooper and M.W. Kennedy, unpublished; [100]], so a very small lipid binding progenitor is conceivable.

If the ~ 15 -kDa NPA unit were derived from a

small precursor, then it is possible that some arrays will possess a relic of this. Within the DvA-1 NPA of *D. viviparus* there is a truncated unit of approximately the expected length (unit DvA-1H; Fig. 1). It might therefore be that this represents the progenitor, although it has yet to be functionally tested. The DvA-1 polyprotein is, however, the only case in which a truncated peptide has been found, and there is no evidence of such in the sequence from *C. elegans*. Further evidence for an ancient duplication events has arisen from sequence analysis of the NPA cDNAs of the dog heartworm *Dirofilaria immitis*, several units encoded by which have very similar amino acid sequences in their N-terminal halves, but divergent C-terminal halves [58].

If the NPA genes arose from a series of duplications to encode multiple 15-kDa units, then it might be expected that when one compares the sequences of each repeat unit of a single array, then a pattern of relationships might emerge. Such an analysis for DvA-1 shows that the amino acid sequences of the units encoded at the 3' end of the cDNA are much more closely related than are those at the 5' end (Fig. 5, and [57]). This suggests that the duplications occurred predominately at the 3' end of the gene, followed by divergence of the products of the multiplications. The opposite appears to be the case for *C. elegans*, for which the units at the 5' end of the gene are more closely related. So, the duplications may have occurred in independent lineages of nematode, or there have been successive deletions or additional gene conversion events since.

12. Internal sequence diversity

The reason for the dramatic differences between the amino acid sequences of NPA units could be that the amplification of the array is ancient, and that neutral substitutions have accumulated without changes in function, and/or that, following amplification, the separate units evolved independently to have modified functions. Testing the latter possibility would require the production of recombinant proteins representing every member of an array and characterising their individual biochemical activities. This heroic experiment has not yet been attempted, but a comparison of two quite different units in the

ABA-1 array of *Ascaris* has been carried out [63]. This revealed that there are essentially no differences in the binding functions of the two in terms of the lipids they bind and their dissociation constants, although subtle differences in the environment of their binding sites were detectable. Whether these slight differences have any biological significance remains to be seen.

The diversity within an array can be very high. In *D. viviparus* DvA-1, for instance, the individual units can differ by as much as 82% of their amino acids. This diversity is similar to that between the beta barrel FABPs from different tissues in humans, although their structures are virtually identical [81,96]. The predicted secondary structures of NPAs are also very similar, and their sequence diversity can be used to identify conserved features which are essential to their structure and function. Mutagenesis is then a straightforward means by which to examine conserved features, and our chemical and genetic mutagenesis study showed that alteration of the NPA signature amino acids (Trp-15, Gln-20, Cys-64, Cys-124) does indeed affect the binding function and/or stability of the proteins (L. McDermott, J. Moore, N.C. Price, S. Kelly, A. Cooper and M.W. Kennedy, unpublished; [100]).

A curious fact is that whilst some species of nematode have diverse arrays, others have arrays in which the units are very similar or even identical in amino acid sequence. The reason why some species have relatively uniform arrays and others are highly diverse is far from clear. It might be pertinent to note, however, that those which appear to have NPAs comprising similar units (genera *Ascaris*, *Brugia*) are allotted to the same phylogenetic clade (although *Loa* is an exception), whereas those with highly diverse arrays (genera *Dictyocaulus*, *Caenorhabditis*, *Haemonchus*, *Ostertagia*) are in a distantly-related clade [102]. It could therefore be that the differences reflect the ancestral condition of the progenitor of each clade. On the other hand, the similarity between the units in *Ascaris*, for example, is so close (at both protein and DNA levels) that it is likely that they are derivatives of a recent amplification event.

The species whose NPAs comprise very similar units have no free-living stages, whereas those with diverse arrays do. It could therefore be that an array

with a variety of products is of advantage to organisms which encounter fluctuating environmental conditions through their life cycles, but one form is optimal for a parasitic stage.

It must be pointed out, however, that complete information is available for too few species to permit robust generalisations. It may merely be that diverse units are yet to be found in those species in which only very similar units have been found so far.

13. Why polyproteins?

Nothing of what we know about NPAs to date gives us strong clues as to the special reason why they have evolved as polyproteins. Synthesis that way seems to be highly economical, but why, then, are more proteins not produced in this manner? ABA-1 is an abundant product of *Ascaris* and related nematodes, so the need for rapid, mass production seems logically related to production as polyproteins, but the NPAs are not so predominant in other, smaller, species. Also, whilst vertebrates produce polyproteins (e.g. the filaggrins), abundantly produced proteins (such as serum albumin) are not synthesised this way. If NPA/filaggrin-type polyproteins are confined, for some reason, to terminally differentiating, apoptotic cells in vertebrates, then why does this rule not apply in other metazoans? Is there some connection between synthesis as polyproteins and the fact that the encoding genes have no (or very small – see below) introns?

The fact that sequence-divergent units from the same nematode polyprotein array appear to have similar functional lipid-binding properties prompts two, possibly related, questions: (1) why are NPAs produced as polyproteins, and (2) why has the dramatic sequence diversity evolved when the end products have similar functional properties? Perhaps the answer to both questions arises from a requirement for large amounts of functionally active protein to be produced rapidly and efficiently, at high local concentrations at the site of synthesis or action. The polyprotein mechanism would achieve this, but the limited solubility of a structurally homogeneous protein would impose limits on the concentrations safely achievable without aggregation occurring [103]. The diversity amongst the units might overcome this

problem by raising the threshold concentration of protein which could be produced before significant aggregation occurs. In relation to the solubility argument, it may be relevant to note that different component units of polyprotein arrays show significant differences in predicted *pI* (ranging from 5.0 to 6.4), but will all be negatively charged at neutral pH (unpublished).

Alternatively, the explanation might be genetic. Following the amplification of DNA encoding the ancestral NPA unit, an array comprising tandemly repeated sequences might be at risk of disruption through homologous recombination. There would therefore be a selective advantage to diversifying the DNA, within functional constraints imposed on the encoded proteins. Instability of such a gene might, for instance, be the reason for the apparent difference in the sequences and number of repeat units in the flaggrins between individual humans, and apparent size heterogeneity of the products [104,105]. On the other hand, the amplification event(s) might merely be sufficiently ancient that the diversified NPA arrays we observe today result from a gradual divergence in the sequences of the different units in an array, but with conservation of function.

A final potential explanation for the evolution of internal diversity in the NPA units is that it might reflect some other additional function(s) that we have not yet identified. For example, although these molecules might be fairly indiscriminating in what they bind, the sites of delivery – what they deliver and where – may be more specific. The diversity in the active site and other features of the proteins might therefore reflect different mechanisms for transfer or release of their payload at different receptors or membrane locations.

14. Gene structure: truly intronless?

The idea that the genes encoding NPAs are intronless came from two observations. First, Southern blots of genomic DNA which had been partially digested with a restriction enzyme cutting only once within a repeat unit produced a ladder of bands [50–52,57]. The sizes of the DNA fragments in the ladder had increments which approximated to that

required to encode a single unit polypeptide, and no more. Secondly, genomic DNA amplified by PCR contained no introns [51,65,106]. Preliminary genomic analysis of the *Ascaris* NPA gene has indeed revealed no introns in the repetitive part of the protein, but a large (4-kbp) intron interrupts the region encoding the small extension peptide at the COOH terminus [107]. The genomic sequence from *C. elegans* provides more complete information, and reveals very small introns (~ 40 –396 bp; Fig. 1) interrupting the coding region for the most 5' repeat unit, and in some, but not all, of the subsequent units. It is feasible, then, that small introns may also exist in *npa* genes of other nematodes, but below the resolution of the Southern blotting experiments carried out to date, and that PCR products containing introns have been ignored.

Introns may have been involved in the amplification process of NPA units, and then subsequently lost. If so, then contiguous units with no introns may be the more ancient and likely to have accumulated changes in their sequences. In the case of *C. elegans*, this would match the fact that the most divergent units do not have introns (Figs. 1 and 5). On the other hand, as stated above, no introns have been found in the ABA-1 array of *Ascaris*, yet the units are very similar.

It might be, therefore, that other polyproteins which are thought to be encoded by genes which are intronless in the coding region, such as the flaggrins of human skin, also have undiscovered/small introns. Such introns may be remnants of the original amplification process, or are material to future amplification (or deletion) events in the evolution of this type of polyprotein.

15. Site of synthesis and biological function

The NPAs appear to have two broad biochemical functions. The first is the extracellular transport, solubilisation and chemical protection of small lipids (such as fatty acids and retinoids), as with our own serum albumin. This fits with their presence in the pseudocoelomic fluid of nematodes (the approximate equivalent of vertebrate blood and the haemolymph of insects), and immunoelectronmicroscopy studies have localised NPA to extracellular connective/struc-

tural tissues of the organisms [82]. They may therefore function to distribute small lipids from the gut to the musculature, gonads and other lipid-requiring tissues. In *Ascaris*, this may be a particularly important function because its pseudocoelomic fluid contains very high levels of fatty acids – equivalent to quantities which begin to appear in the plasma of humans after several days of starvation (K. Copeland, M.W. Kennedy and L. Fixter, unpublished). They may also be involved in the sequestration, transportation and salvage of haem compounds and metal ions, a function which might be important to blood-feeding species. Some haem-derived compounds appear to bind in the lipid binding site, but the His-rich peptides may also perform this function in addition to binding metal ions.

NPA mRNA has only been detected in the gut (of *Ascaris*) [65], and transfection work in *C. elegans* has indicated that the promoter is active only in the gut (I. Johnstone and M.W. Kennedy, unpublished). NPAs may therefore be synthesised in the gut, processed down to ~15-kDa units, loaded with lipids and released for distribution. They may then unload their cargo upon interaction with appropriate cell surface receptors, be briefly internalised and recycled as *apo* proteins into the pseudocoelomic fluid for continued shuttling, or merely transfer their cargo directly to cell membranes.

Their second major function would be confined to parasites, and be to alter the tissue environment to the parasites' advantage, either by delivering pharmacologically active lipids or sequestering the host's own lipid mediators. This would require the proteins to be secreted, but, to date, there is no evidence that specialised secretory cells or organs of nematodes produce and/or release NPAs. Release directly from the gut, however, is clearly a possibility.

16. Concluding remarks

NPAs are the only lipid binding proteins known to be produced as polyproteins. Their existence questions some of the explanations for the evolution of repetitive polyproteins in higher organisms, such as synthesis uniquely by apoptotic cells producing specialist products. A more general question is: why are more proteins not made this way? If there are good

reasons for this, then what is special about NPAs and filaggrins? Why do the encoding genes possess small and few introns?

Finally, if, as seems likely, the ~15-kDa units of NPAs derive from a duplication event (and even larger entities exist because of the absence of processing cleavage sites between some units in an array), then might our own serum albumin have arisen similarly? Could they even have a common ancestor?

Acknowledgements

I am greatly indebted to the many people who have contributed to the work and my/our thinking about these proteins, particularly Heather Spence, Joyce Moore, Collette Britton, Jacqui Christie, Eleanor Fraser, Yu Xia, Murray Selkirk, Andy Brass, Tim Lee, Angie McGibbon, Judi Allen, Rick Mairrels, Mark Blaxter, Larry McReynolds, Nick Price, Sharon Kelly, Alan Cooper and Lindsay McDermott. I am also grateful to Iain Johnstone for guiding me through the *C. elegans* genome database and for his experiments on the NPA promoter. Our work on the NPAs and other nematodes was mainly supported by the Wellcome Trust, and also by the Medical Research Council (UK).

References

- [1] N.A. Cobb, Nematodes and their Relationships, U.S. Department of Agriculture Yearbook, 1915, pp. 457–490.
- [2] The *C. elegans* Sequencing Consortium, *Science* 282 (1998) 2012–2018.
- [3] M.W. Kennedy, A. Brass, A.B. McCruden, N.C. Price, S.M. Kelly, A. Cooper, *Biochemistry* 34 (1995) 6700–6710.
- [4] M.W. Kennedy, C. Britton, N.C. Price, S.M. Kelly, A. Cooper, *J. Biol. Chem.* 270 (1995) 19277–19281.
- [5] M.W. Kennedy, L.H. Garside, L.E. Goodrick, L. McDermott, A. Brass, N.C. Price, S.M. Kelly, A. Cooper, J.E. Bradley, *J. Biol. Chem.* 272 (1997) 29442–29448.
- [6] J.M. Fisher, R.H. Scheller, *J. Biol. Chem.* 263 (1988) 16515–16518.
- [7] P. Neddermann, L. Tomei, C. Steinkuhler, P. Gallinari, A. Tramontano, R. DeFrancesco, *Biol. Chem.* 378 (1997) 469–476.
- [8] V.M. Vogt, *Curr. Top. Microbiol. Immunol.* 214 (1996) 95–131.
- [9] I.G. Maia, K. Seron, A.L. Haenni, F. Bernardi, *Plant Mol. Biol.* 32 (1996) 367–391.

- [10] C.J.A. Wallace, *Protein Sci.* 2 (1993) 697–705.
- [11] R.P. Presland, P.V. Haydock, P. Fleckman, W. Nirunskisiri, B.A. Dale, *J. Biol. Chem.* 267 (1992) 23772–23781.
- [12] P.V. Haydock, B.A. Dale, *DNA Cell Biol.* 9 (1990) 251–261.
- [13] J.A. Rothnagel, P.M. Steinert, *J. Biol. Chem.* 265 (1990) 1862–1865.
- [14] S.-Q. Gan, O.W. McBride, W.W. Idler, N. Markova, P.M. Steinert, *Biochemistry* 29 (1990) 9432–9440.
- [15] K. Yoneda, D. Hohl, O.W. McBride, M. Wang, K.U. Cehrs, W.W. Idler, P.M. Steinert, *J. Biol. Chem.* 267 (1992) 18060–18066.
- [16] W.B. Wood, in: W.B. Wood (Ed.), *The Nematode Caenorhabditis elegans*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988, pp. 1–16.
- [17] J. Soultson, in: W.B. Wood (Ed.), *The Nematode Caenorhabditis elegans*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988, pp. 123–155.
- [18] M. Blaxter, *Science* 282 (1998) 2041–2046.
- [19] J.E. Allen, R.M. Maizels, *Int. Arch. Allergy Immunol.* 109 (1996) 3–10.
- [20] A.P. Gelpi, A. Mustapha, *Am. J. Trop. Med. Hyg.* 16 (1967) 646–657.
- [21] A.P. Gelpi, A. Mustapha, *Am. J. Med.* 44 (1968) 377–389.
- [22] I. Hagel, N.R. Lynch, M.C. di Prisco, E. Rojas, M. Perez, N. Alvarez, *Clin. Exp. Immunol.* 94 (1993) 80–83.
- [23] C. McSharry, Y. Xia, C.V. Holland, M.W. Kennedy, *Infect. Immun.* 67 (1999) 484–489.
- [24] D.W. Dunne, A.E. Butterworth, A.J. Fulford, H.C. Kariuki, J.G. Langley, H. Ouma, A. Capron, R.J. Pierce, R.F. Sturrock, *Eur. J. Immunol.* 22 (1992) 1483–1494.
- [25] P. Hagan, U.J. Blumenthal, D. Dunn, A.J.G. Simpson, H.A. Wilkins, *Nature* 349 (1991) 243–245.
- [26] P. Rihet, C.E. Demeure, A. Burgois, A. Prata, A.J. Dessein, *Eur. J. Immunol.* 21 (1991) 2679–2686.
- [27] W.O.C.M. Cookson, M.F. Moffatt, *Science* 275 (1997) 41–42.
- [28] J.F. Urban, K.B. Madden, A. Svetic, A. Cheever, P.P. Trotta, W.C. Gause, I.M. Katona, F.D. Finkelman, *Immunol. Rev.* 127 (1992) 205–220.
- [29] A. Sher, R.L. Coffman, *Annu. Rev. Immunol.* 10 (1992) 385–409.
- [30] K.J. Else, F.D. Finkelman, C.R. Maliszewski, R.K. Grencis, *J. Exp. Med.* 179 (1994) 347–351.
- [31] R.K. Grencis, L. Hultner, K.J. Else, *Immunology* 74 (1991) 329–332.
- [32] N.R. Lynch, in: R. Moqbel (Ed.), *Allergy and Immunity to Helminth Infection: Common Mechanisms or Divergent Pathways?*, Taylor and Francis, London, 1992, pp. 51–62.
- [33] N.R. Lynch, M. Palenque, I. Hagel, M.C. DiPrisco, *Am. J. Respir. Crit. Care Med.* 156 (1997) 50–54.
- [34] T. Shirakawa, T. Enomoto, S.-I. Shimazu, J.M. Hopkin, *Science* 275 (1997) 77–79.
- [35] G.J. Waite, D.R. Mazzer, D.P. McManus, *Parasit. Immunol.* 19 (1997) 531–533.
- [36] E.E.E. Jarrett, S. Mackenzie, H. Bennich, *Nature* 282 (1980) 302–304.
- [37] E.E.E. Jarrett, H.R.P. Miller, *Prog. Allergy* 31 (1982) 178–233.
- [38] L.A. Tomlinson, J.F. Christie, E.M. Fraser, D. McLaughlin, A.E. McIntosh, M.W. Kennedy, *J. Immunol.* 143 (1989) 2349–2356.
- [39] J.E. Allen, R.A. Lawrence, R.M. Maizels, *Infect. Immun.* 63 (1995) 2892–2898.
- [40] M.W. Kennedy, E.M. Fraser, J.F. Christie, *Immunology* 72 (1991) 577–579.
- [41] J.F. Christie, B. Dunbar, I. Davidson, M.W. Kennedy, *Immunology* 69 (1990) 596–602.
- [42] J.F. Christie, E.M. Fraser, M.W. Kennedy, *Parasite Immunol.* 14 (1992) 59–73.
- [43] M.W. Kennedy, F. Qureshi, *Immunology* 58 (1986) 515–522.
- [44] M.W. Kennedy, F. Qureshi, E.M. Fraser, M. Haswell-Elkins, D.B. Elkins, H.V. Smith, *Clin. Exp. Immunol.* 75 (1989) 493–500.
- [45] M.W. Kennedy, F. Qureshi, M. Haswell-Elkins, D.B. Elkins, *Clin. Exp. Immunol.* 67 (1987) 20–30.
- [46] A.M. McGibbon, J.F. Christie, M.W. Kennedy, T.D.G. Lee, *Mol. Biochem. Parasitol.* 39 (1990) 163–172.
- [47] M.W. Kennedy, L.A. Tomlinson, E.M. Fraser, J.F. Christie, *Clin. Exp. Immunol.* 80 (1990) 219–224.
- [48] E.M. Fraser, J.F. Christie, M.W. Kennedy, *Int. Arch. Allergy Immunol.* 100 (1993) 283–286.
- [49] M.W. Kennedy, A.M.S. Gordon, L.A. Tomlinson, F. Qureshi, *Parasite Immunol.* 9 (1987) 269–273.
- [50] J. Culpepper, R.B. Grieve, L. Friedman, M. Mika-Grieve, G.R. Frank, B. Dale, *Mol. Biochem. Parasitol.* 54 (1992) 51–62.
- [51] S. Tweedie, W.A. Paxton, L. Ingram, R.M. Maizels, L.A. McReynolds, M.E. Selkirk, *Exp. Parasitol.* 76 (1993) 156–164.
- [52] C.B. Poole, A.G. Grandea, C.V. Maina, R.E. Jenkins, M.E. Selkirk, L.A. McReynolds, *Proc. Natl. Acad. Sci. USA* 89 (1992) 5986–5990.
- [53] H.J. Spence, J. Moore, A. Brass, M.W. Kennedy, *Mol. Biochem. Parasitol.* 57 (1993) 339–344.
- [54] D.C. de Graaf, L.J. Peelman, E. Claerebout, H. Hilderson, H.D. Schallig, J. Vercruysse, *Mol. Biochem. Parasitol.* 72 (1995) 239–241.
- [55] S. Yahiro, G. Cain, J.E. Butler, *Parasite Immunol.* 20 (1998) 351–357.
- [56] P.M. Ajuh, J.P. Akue, P. Boutin, S. Everaere, T.G. Egwang, *Exp. Parasitol.* 81 (1995) 145–153.
- [57] C. Britton, J. Moore, J.S. Gilleard, M.W. Kennedy, *Mol. Biochem. Parasitol.* 72 (1995) 77–88.
- [58] C.B. Poole, L.J. Hornstra, J.S. Benner, J.R. Fink, L.A. McReynolds, *Mol. Biochem. Parasitol.* 82 (1996) 51–65.
- [59] P.J. Barr, *Cell* 66 (1991) 1–3.
- [60] J.F. Christie, B. Dunbar, M.W. Kennedy, *Clin. Exp. Immunol.* 92 (1993) 125–132.
- [61] M.O. Dayhoff, R.M. Schwartz, B.C. Orcutt, in: *Atlas of Protein and Sequence Structure*, M.O. Dayhoff (Ed.), Vol. 5, Suppl. 3, National Biomedical Research Foundation, Washington, DC, 1978, pp. 345–362.

- [62] D. Bordo, P. Argos, *J. Mol. Biol.* 217 (1991) 721–729.
- [63] J. Moore, L. McDermott, N.C. Price, S.M. Kelly, A. Cooper, M.W. Kennedy, *Biochem. J.* 340 (1999) 337–343.
- [64] M.W. Kennedy, J.E. Allen, A.S. Wright, A.B. McCrudden, A. Cooper, *Mol. Biochem. Parasitol.* 71 (1995) 41–50.
- [65] Y. Xia, H.J. Spence, J. Moore, N. Heaney, L. McDermott, A. Cooper, D.G. Watson, B. Mei, R. Komuniecki, M.W. Kennedy, *Parasitology*, in press.
- [66] T.C.I. Wilkinson, D.C. Wilton, *Biochem. J.* 238 (1986) 419–424.
- [67] A.E.A. Thumser, C. Evans, A.F. Worrall, D.C. Wilton, *Biochem. J.* 297 (1994) 103–107.
- [68] R.B. Macgregor, G. Weber, *Nature* 319 (1986) 70–73.
- [69] D.C. Wilton, *Biochem. J.* 270 (1990) 163–166.
- [70] J.H. Veerkamp, H.T.B. van Moerkerk, C.F.M. Prinsen, T.H. van Kuppevelt, *Mol. Cell. Biochem.* 192 (1999) 137–142.
- [71] A.E. Thumser, D.C. Wilton, *Biochem. J.* 300 (1994) 827–833.
- [72] A. Cooper, D. Ma, L. McReynolds, M.W. Kennedy, *Worm Breeders Gazette* 14 (1) (1995) 76.
- [73] B. Mei, M.W. Kennedy, J. Beauchamp, P.R. Komuniecki, R. Komuniecki, *J. Biol. Chem.* 272 (1997) 9933–9941.
- [74] D. Janssen, J. Barrett, *Biochem. J.* 311 (1995) 49–57.
- [75] J. Barrett, N. Saghir, A. Timanova, K. Clarke, P.M. Brophy, *Eur. J. Biochem.* 250 (1997) 269–275.
- [76] J. Thompson, N. Winter, D. Terwey, J. Bratt, L. Banaszak, *J. Biol. Chem.* 272 (1997) 7140–7150.
- [77] G. Zanotti, G. Scapin, P. Spadon, J.H. Veerkamp, J.C. Sacchettini, *J. Biol. Chem.* 267 (1992) 18541–18550.
- [78] J. LaLonde, M.A. Levenson, J.J. Roe, D.A. Bernlohr, L.J. Banaszak, *J. Biol. Chem.* 269 (1994) 25339–25347.
- [79] Z. Zu, D.A. Bernlohr, L.J. Banaszak, *J. Biol. Chem.* 268 (1993) 7874–7884.
- [80] J.C. Sacchettini, J.I. Gordon, L.J. Banaszak, *J. Mol. Biol.* 208 (1989) 327–339.
- [81] N.R. Coe, D.A. Bernlohr, *Biochim. Biophys. Acta* 1391 (1998) 287–306.
- [82] M.E. Selkirk, W.F. Gregory, R.E. Jenkins, R.M. Maizels, *Parasitology* 107 (1993) 449–457.
- [83] M.W. Kennedy, F. Qureshi, *Immunology* 58 (1986) 515–522.
- [84] L.X. Liu, J.E. Buhlmann, P.F. Weller, *Am. J. Trop. Med. Hyg.* 46 (1992) 520–523.
- [85] T.I.M. Tree, A.J. Gillespie, K.J. Shepley, M.L. Blaxter, R.S. Tuan, J.E. Bradley, *Mol. Biochem. Parasitol.* 69 (1995) 185–195.
- [86] C. Wulffing, J. Lombardero, A. Pluckthun, *J. Biol. Chem.* 269 (1994) 2895–2901.
- [87] D.J. Sullivan, I.Y. Gluzman, D.E. Goldberg, *Science* 271 (1996) 219–222.
- [88] W.T. Morgan, *Biochemistry* 20 (1981) 1054–1061.
- [89] M. Kappel, W.T. Morgan, *Biochemistry* 24 (1985) 5919–5924.
- [90] T. Koide, D.C. Foster, S. Yoshitake, E.W. Davie, *Biochemistry* 25 (1986) 2220–2225.
- [91] M.L. Blaxter, *Parasitol. Today* 9 (1993) 353–360.
- [92] S. Huang, J. Huang, A.P. Kloek, D.E. Goldberg, J.M. Friedman, *J. Biol. Chem.* 271 (1996) 958–962.
- [93] M.F. Oliveira, J.R. Silva, M. Dansa-Petreski, W. de Souza, U. Lins, C.M.S. Braga, H. Masuda, P.L. Oliveira, *Nature* 400 (1999) 517–518.
- [94] H.J. Loomans, B.L. Hahn, Q.-Q. Li, S.H. Phadnis, P.G. Sohnle, *J. Infect. Dis.* 177 (1998) 812–814.
- [95] Y. Kato, S. Komatsu, *J. Biol. Chem.* 271 (1996) 30493–30498.
- [96] J.C. Sacchettini, J.I. Gordon, *J. Biol. Chem.* 268 (1993) 18399–18402.
- [97] S. Sheriff, W.A. Hendrickson, J.L. Smith, *J. Mol. Biol.* 197 (1987) 273–296.
- [98] M.H. Lerche, F.M. Poulsen, *Protein Sci.* 7 (1998) 2490–2498.
- [99] B.B. Kragelund, K.V. Andersen, J.C. Madsen, J. Knudsen, F.M. Poulsen, *J. Mol. Biol.* 230 (1993) 1260–1277.
- [100] L. McDermott, PhD thesis, University of Glasgow, 1999.
- [101] L. Audicana, M.T. Audicana, L. FernandezdeCorres, M.W. Kennedy, *Vet. Rec.* 140 (1997) 235.
- [102] M.L. Blaxter, P. Deley, J.R. Garey, L.X. Liu, P. Scheldeman, A. Vierstraete, J.R. Vanfleteren, L.Y. Mackey, M. Dorris, L.M. Frisse, J.T. Vida, W.K. Thomas, *Nature* 392 (1998) 71–75.
- [103] T. Arakawa, S.N. Timasheff, *Methods Enzymol.* 114 (1985) 49–77.
- [104] S.-Q. Gan, O.W. McBride, W.W. Idler, N. Markova, P.M. Steinert, *Biochemistry* 29 (1990) 9432–9440.
- [105] C.D. Thulin, J.A. Taylor, K.A. Walsh, *Protein Sci.* 5 (1996) 1157–1164.
- [106] W.A. Paxton, M. Yazdanbakhsh, A. Kurniawan, F. Partono, R.M. Maizels, M.E. Selkirk, *Infect. Immun.* 61 (1993) 2827–2833.
- [107] H.J. Spence. PhD thesis, University of Glasgow, 1994.