

MOLBIO 01908

Short Communication

A cDNA encoding repeating units of the ABA-1 allergen of *Ascaris*

Heather J. Spence^a, Joyce Moore^a, Andrew Brass^b and Malcolm W. Kennedy^a

^aWellcome Laboratories for Experimental Parasitology, University of Glasgow, Glasgow, UK; and ^bDepartment of Biochemistry and Molecular Biology, University of Manchester, Manchester, UK

(Received 22 September 1992; accepted 20 October 1992)

Key words: *Ascaris suum*; Nematode; Allergen; ABA-1

The ABA-1 allergen of parasitic nematodes of the genus *Ascaris* is the most abundant protein in the pseudocoelomic fluid of the adult parasites and is also released by the tissue-penetrating larval stages of the parasite during culture in vitro [1,2]. Infected humans are known to respond heterogeneously to the allergen [3], and the antibody responses of experimentally infected rodents, both IgG and IgE, are under the control of the class II region of the major histocompatibility complex [4]. Therefore, IgE-mediated hypersensitivity responses to ABA-1 may contribute to the pathology of ascariasis in those individuals bearing the appropriate immune responder genes, particularly during the pulmonary migration phase.

ABA-1 homologues are known to occur in a wide spectrum of parasitic nematodes, although allergenicity has only been established for ABA-1. Immunologically cross-reactive components have been found in other ascaridids infective to humans [5,6] and recent cDNA clones encoding polypeptides whose

only known similarity is with ABA-1 have been described in filarial parasites (refs. 7,8; and M.E. Selkirk, personal communication). It is important, therefore, to understand the physiological role of such an abundant and widespread component of parasitic nematodes and any distinctive features that lend it to allergenicity and thereby pathogenicity.

With this in mind, a cDNA expression library was constructed in λ Zap II (Stratagene) using mRNA prepared from *A. suum* infective larvae hatched and cultured for 24 h [1], and screened with polyclonal rabbit antiserum to gel-excised ABA-1 protein [2]. This led to the isolation of a clone (pHS10) with an insert of 1255 bp containing part of a gene, here designated as *aba-1*. To enable sequencing of the entire insert, exonuclease III (Promega) deletions were constructed (Fig. 1A) and the inserts sequenced by Sanger's dideoxy method [9] using Sequenase 2.0 and TAQuenase 2.0 (United States Biochemical). The assembled sequence (Fig. 1B) suggests that there is an open reading frame of 1185 bp followed by a 3' untranslated region. An unusual feature of the insert is that it comprises one truncated and two complete repeat sequences of 399 bp; the complete repeats beginning at positions 286 and 685 of the cDNA. The 5' incomplete repeat would appear to start at position 120 with respect to a complete repeat. The repeats are identical

Correspondence address: Malcolm Kennedy, Wellcome Laboratories for Experimental Parasitology, University of Glasgow, Bearsden, Glasgow G61 1QH, UK. Tel.: 041-339-8855 Ext. 5819; Fax: 041-942-7215.

Note: Nucleotide sequence data reported in this paper have been submitted to the GenBank™ data base with the accession number L03211.

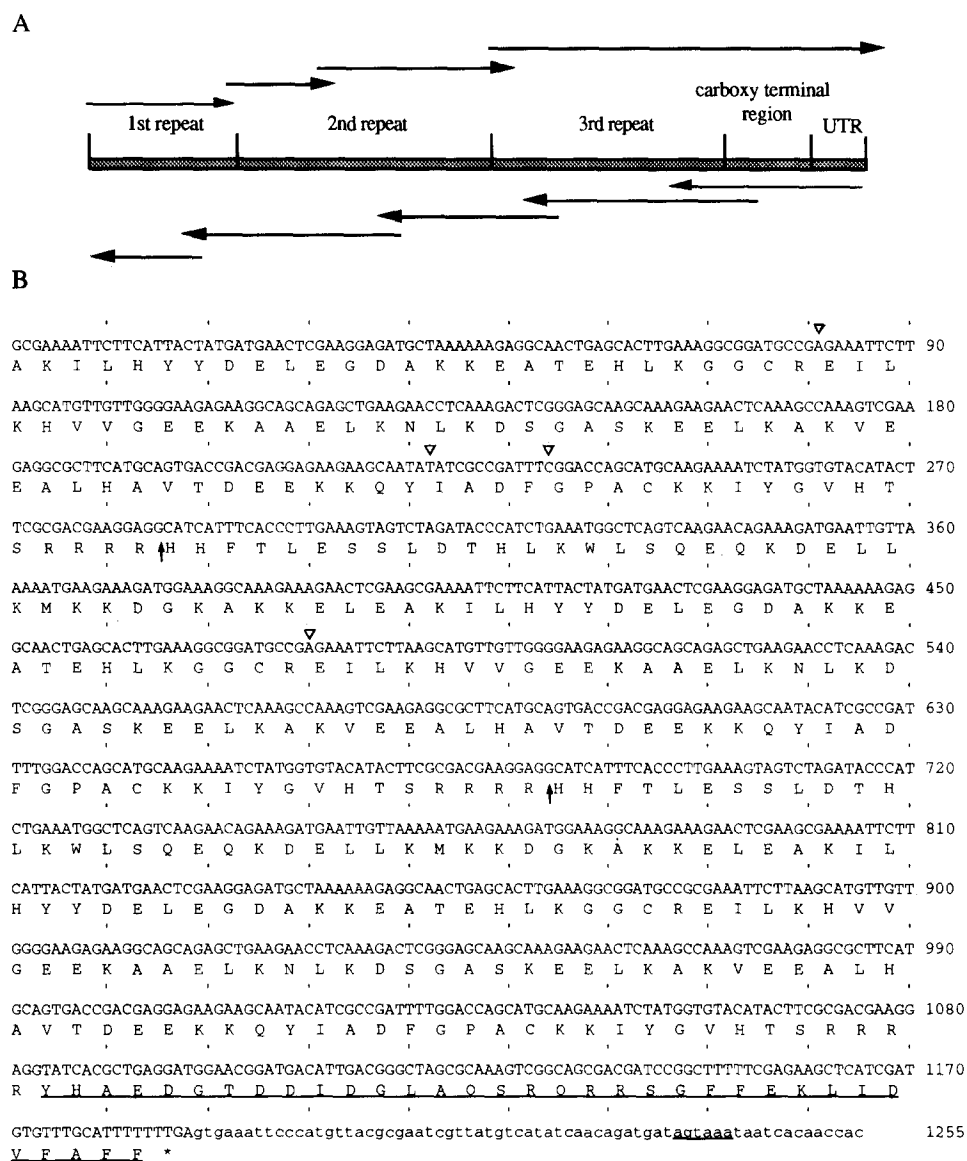


Fig. 1. (A) Diagram of the pH10 insert encoding part of the *aba-1* gene with arrows indicating the regions sequenced from exonuclease III deletions, or directly from the original plasmid. (B) DNA sequence of the entire pH10 insert and predicted amino acid sequence of the ABA-1 protein. The junction between repeat units is indicated by the arrows. The carboxy terminal region is underlined and the stop codon is indicated by an asterisk. Lower case letters indicate the putative 3' untranslated region and a putative polyadenylation signal is underlined. Inverted open triangles indicate base differences between the first and second repeats and the most 3' repeat.

apart from 3 conservative base pair differences between the first (positions 81, 222 and 234) and last repeat and one between the second (position 480) and the last. However, positions 81 and 480 are identical between repeats 1 and 2. Following the repeats there is a further 102

bp of open reading frame which bear no resemblance to the foregoing sequence, before the putative stop codon. Although the polyadenylation signal AAUAAA found in vertebrates [10] and nematodes [11] is not present in the 3' untranslated region of the gene, there is

the sequence AGUAAA at position 1248–1253 (underlined in Fig. 1B) which has been shown to act as a polyadenylation signal in vertebrates [10].

The predicted amino acid sequence from the beginning of each putative repeat matches perfectly with our previously reported N-terminal amino acid sequence obtained by direct sequencing of native ABA-1 to 41 residues [12]. The only sequence similar to this and the amino acid sequence predicted from a complete repeat within the cDNA found in the GenEMBL or SwissProt databases [13] (releases 67.0 and 22, respectively) was a cDNA clone encoding a 15-kDa antigen of *Dirofilaria immitis* [7,8]. A similar cDNA has also been cloned from *Brugia malayi* (M.E. Selkirk, personal communication). These two organisms are filarial nematode parasites distantly related to *Ascaris*. The cDNA from *D. immitis* also comprises tandem repeats and contains, like the *Ascaris* gene, slight differences between repeats. An alignment of the putative amino acid sequence of ABA-1 with one of the *D. immitis* repeats (Di22-16) is shown in Fig. 2A.

The last 4 amino acids (RRRR) of each repeat in the *Ascaris* gene match a consensus cleavage site of endoproteinases of the subtilisin family, a common motif for which is

K/R.X.K/R.R [14]. It is conceivable, therefore, that ABA-1 is produced as a large polyprotein which is subsequently cleaved to single units. Such a process has been inferred for the post-translational processing of the *D. immitis* 15-kDa antigen from biosynthetic labelling experiments on intact parasites [8]. We have previously found [3] that radio-iodinated somatic material from *Ascaris* immunoprecipitated with a polyclonal antibody to ABA-1 produces a ladder of bands similar to that found for the *D. immitis* antigen. This ladder extended to 6 units of ABA-1, and our recent unpublished work on probing Western blots of somatic material with the same antibody showed that the ladder ascends to a component of high molecular weight, as has been observed for *D. immitis*. It is likely, therefore, that the whole *aba-1* gene is of a similar large size to that of the latter parasite.

The slight differences between the DNA sequence of the repeats in the *Ascaris* gene do not result in changes in the amino acid sequence of the predicted translation product from each repeat unit. We have, however, preliminary sequence data on other cDNAs, encoding ABA-1, which have differences in base sequence which do alter the predicted amino acid sequence. These cDNAs appear to have primed in A-rich areas within the coding

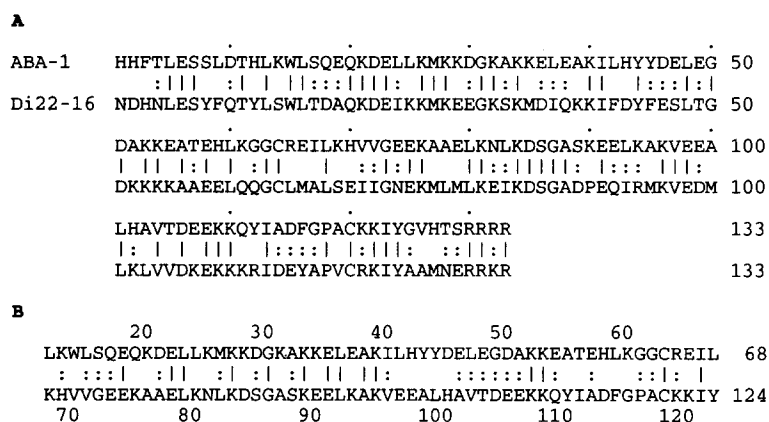


Fig. 2. (A) Alignment of the amino acid sequences predicted from a repeat unit in the *aba-1* gene with that of the Di22-16 repeat from *D. immitis*. Positions with identical amino acids are indicated (I) and those with conservative substitutions as defined by Dayhoff [16] are indicated (:). The amino acid sequences are 45.1% identical and 72.9% similar when conservative changes are allowed. (B) Possible internal duplication within the ABA-1 protein. An alignment from positions 13–68 and 69–124 of the ABA-1 protein giving 25% identity and 60.7% similarity between the sequences of the two putative domains.

region of the mRNA rather than the poly(A) tail, and could either represent copies in the array which are 5' to those present in pHs10, or encoded by another locus.

A single repeat unit within *aba-1* predicts a protein of 15257 Da, which is larger than previous estimates of the size of the native molecule [15]. We have consistently found the latter to be of M_r 14400 in gradient SDS-PAGE [1,2] and have unpublished mass spectrometry data indicating that ABA-1 is heterogeneous, but with a major component of 14643.2 ± 1.4 Da. This would be approximately the mass of a unit if the terminal arginines were trimmed after cleavage of the ABA-1 units during post-translational processing.

We have no information on the secondary or tertiary structure of ABA-1, but there appears from the primary sequence to be an indication of the existence of 2 domains. Alignment of the amino acid sequences from positions 13–68 with 69–124 in any repeat (Fig. 2B) shows 25% identity and 60.7% similarity when conservative changes are allowed. Of particular interest are the conserved cysteine, glycine and lysine residues at positions 28, 31, 40, 54, and 64 since these are also conserved in a similar alignment in the Di22-16 protein of *D. immitis* (not shown), suggesting that these residues may play an important structural or functional role in the ABA-1 protein. It is possible, therefore, that the ancestral *aba-1* gene was formed from a duplication event, the duplicated domains having since diverged whilst maintaining similarity at crucial residues.

The putative 34 amino acid carboxy-terminal region, only present in the last repeat, has two features worthy of note. Firstly, the terminal 12 amino acids would form a substantially more hydrophobic region than is found anywhere else in the protein. Secondly, there is a second consensus endoprotease site (RQRR) at positions 17–20 of the terminal region, cleavage at which would remove the most hydrophobic region. The function of this region, if any, remains unknown. It is conceivable, however, that it is involved in the

intracellular processing of the polypeptide. It is unlikely to be a membrane anchor because the hydrophobic region is shorter than is conventional for such anchors, it has no terminus comprising charged residues to act as a cytoplasmic holdfast, and charged residues appear within the hydrophobic region itself. Finally, while it has some superficial resemblances to signals for GPI linkages, it would be unusual in its content of positively charged residues [17]. The terminal region of the cDNA encoding ABA-1 has no equivalent in any nematode cDNA yet described, and may have important implication for our understanding of the post-translational processing of the products of this unusual type of gene.

Acknowledgements

This work was supported by grants from the Wellcome Trust and the Medical Research Council to MWK. Thanks also go to Drs. Lorna Proudfoot and Mike Ferguson for advice on the carboxy terminal tail and to Anne McIntosh and David McLaughlin for technical help.

References

- 1 Kennedy, M.W. and Qureshi, F. (1986) Stage-specific secreted antigens of the parasitic larval stages of the nematode *Ascaris*. *Immunology* 58, 515–522.
- 2 Kennedy, M.W., Qureshi, F., Haswell-Elkins, M. and Elkins, D.B. (1987) Homology and heterology between the secreted antigens of the parasitic larval stages of *Ascaris lumbricoides* and *Ascaris suum*. *Clin. Exp. Immunol.* 67, 20–30.
- 3 Kennedy, M.W., Tomlinson, L.A., Fraser, E.M. and Christie, J.F. (1990) The specificity of the antibody response to internal antigens of *Ascaris*: heterogeneity in infected humans, and MHC (H-2) control of the repertoire in mice. *Clin. Exp. Immunol.* 80, 219–224.
- 4 Kennedy, M.W., Fraser, E.M. and Christie, J.F. (1991) MHC class II (I-A) region control of the IgE antibody repertoire to the ABA-1 allergen of the nematode *Ascaris*. *Immunology* 72, 577–579.
- 5 Kennedy, M.W., Tierney, J., Ye, P., McMonagle, F.A., Macintosh, A., McLaughlin, D. and Smith, J.W. (1988) The secreted and somatic antigens of the third stage larva of *Anisakis simplex*, and antigenic relationship with *Ascaris suum*, *Ascaris lumbricoides*, and *Toxocara canis*. *Mol. Biochem. Parasitol.* 31, 35–46.
- 6 Kennedy, M.W., Qureshi, F., Fraser, E.M., Haswell-

- Elkins, M., Elkins, D.B. and Smith, H.V. (1989) Antigenic relationships between the surface-exposed, secreted, and somatic materials of the nematode parasites *Ascaris lumbricoides*, *Ascaris suum*, and *Toxocara canis*. Clin. Exp. Immunol. 75, 493–500.
- 7 Culpepper, J., Grieve, R.B., Friedman, L., Mika-Grieve, M., Frank, G.R. and Dale, B. (1992) Molecular characterization of a *Dirofilaria immitis* cDNA encoding a highly immunoreactive antigen. Mol. Biochem. Parasitol. 54, 51–62.
 - 8 Poole, C.B., Grandea, A.G., Maina, C.V., Jenkins, R.E., Selkirk, M.E. and McReynolds, L.A. (1992) Cloning of a cuticular antigen that contains multiple tandem repeats from the filarial parasite *Dirofilaria immitis*. Proc. Natl. Acad. Sci. USA 89, 5986–5990.
 - 9 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.
 - 10 Sheets, M.D., Ogg, S.C. and Wickens, M.P. (1990) Point mutations in AAUAAA and the poly-(A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. Nucleic Acids Res. 18, 5799–5805.
 - 11 Wheelock, M.J., Komuniecki, R., Duran, E. and Johnson, K.R. (1991) Characterization of cDNA clones for the beta subunit of pyruvate dehydrogenase from *Ascaris suum*. Mol. Biochem. Parasitol. 45, 9–18.
 - 12 Christie, J.F., Dunbar, B., Davidson, I. and Kennedy, M.W. (1990) N-terminal amino acid sequence identity between a major allergen of *Ascaris lumbricoides* and *Ascaris suum*, and MHC-restricted IgE responses to it. Immunology 69, 596–602.
 - 13 Devereux, J., Haeberli, P. and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12, 387–395.
 - 14 Barr, P.J. (1991) Mammalian subtilisins: the long-sought dibasic processing endoproteases. Cell 66, 1–3.
 - 15 McGibbon, A.M., Christie, J.F., Kennedy, M.W. and Lee, T.D.G. (1990) Identification of the major *Ascaris* allergen and its purification to homogeneity by high-performance liquid chromatography. Mol. Biochem. Parasitol. 39, 163–172.
 - 16 Dayhoff, M.O., Eck, R.V. and Park, C.M. (1972) A model of evolutionary change in proteins. In: Atlas of Protein Sequence and Structure, Vol. 5 (Dayhoff, M.O., ed.), pp. 89–99, Natl. Biomed. Res. Found. (USA).
 - 17 Ferguson, M.A.J. and Williams, A.F. (1988) Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. Annu. Rev. Biochem. 57, 285–320.