Sequence similarity of a hornet (D. maculata) venom allergen phospholipase A₁ with mammalian lipases

Lyudmila Soldatova, Loucia Kochoumian and Te Piao King

Rockefeller University, New York, NY 10021, USA

Received 3 February 1993

We have determined the sequence of a venom allergen phospholipase A₁ from white-faced hornet (Dolichovespula maculata) by cDNA and protein sequencings. This protein of 300 amino acid residues (Dol m 1) has no sequence similarity with other known phospholipases. But it has sequence similarity with mammalian lipases; about 40% identity in overlaps of 123 residues. Tests suggest that hornet phospholipase has weak lipase activity. Hornet venom has 3 major allergens, and another hornet allergen antigen 5 (Dol m V) was previously found to have sequence similarity with a mammalian testis protein and a plant leaf protein.

Hornet venom allergen; Phospholipase A,

1. INTRODUCTION

One important group of insects which cause insect sting allergy of the immediate type is the vespids which include hornets, wasps and yellowjackets. The vespids have similar venom compositions, and their major venom allergens are hyaluronidase (43 kDa), phospholipase A₁ (37 kDa) and antigen 5 (23 kDa) which has as yet unknown biochemical function [1-3]. We have previously determined the sequence of homologous antigen 5s from white-faced hornet (Dolichovespula maculata) [4] and other wasps and yellowjackets [5]. Antigen 5 was found to have sequence similarities with pathogenesisrelated proteins from tobacco and tomato leaves and a set of homologous proteins from human, rat and mouse testis; about 30% identity in overlaps of 130 residues [5,6]. According to the accepted allergen nomenclature system [7], white-faced hornet phospholipase A₁ and antigen 5 are designated Dol m I and Dol m V, respectively.

Allergy of the immediate type is due to the presence of allergen-specific antibodies of IgE isotype. Antibody response to an immunogen/allergen depends on the genetic make-up of the host, the route and mode of immunization/sensitization and the nature of the immunogen/ allergen. To what extent an allergen determines the outcome of IgE response is not known (cf. [8,9]). In our continuing effort to understand what immunochemical properties of a protein contribute to its allergenicity, we have cloned and sequenced the second major allergen of hornet venom.

Correspondence address: L. Soldatova, Rockefeller University, New York, NY 10021, USA.

2. MATERIALS AND METHODS

2.1. Isolation and characterization of Dol m I and its CNBr peptides Dol m I was isolated from venom sac extracts of white-faced hornet (Vespa Laboratory, Spring Mills, PA) as described [3]. The protein (0.6 mg) was cleaved with CNBr (15 mg) in 75% HCO₂H (0.2 ml) at 25°C overnight. After cleavage the lyophilized mixture was separated on PepRPC column (Pharmacia, Piscataway, NJ) with a 2-propanol gradient of 0.1% per ml in 0.1% trifluoroacetic acid at a flow rate of 40 ml per hour. Selected fractions were rechromatographed under the same conditions after reduction and S-carboxymethylation [4]. The recovered peptides were characterized by Edman degradation on an Applied Biosystems gas phase sequencer.

2.2. Dol m I-specific cDNA

Total RNAs were isolated from the acid gland of white-faced hornet using the guanidine thiocyanate extraction procedure [4]. Dol m Ispecific cDNA was obtained from total RNAs by the procedure of Frohman [10,11] for rapid amplification of 3' or 5' cDNA ends

First strand cDNAs were prepared using the MeHgOH (Invitrogen, San Diego, CA) denatured total RNAs (6 µg) as the template and other reagents of a cDNA synthesis kit from Gibco-BRL (Gaithersburg, MD) and RNasin (Promega Biotech) in a total reaction volume of 37 μ l. For 5' RACE, the single strand cDNAs (from 6 μ g of total RNAs) were poly-dA tailed with terminal deoxynucleotidyl transferase (US Biochemical, Cleveland, OH). 3' or 5' RACE was carried out with GenAmp PCR reagent kit (Perkin-Elmer Cetus, Norwalk, CT) using AmpliTaq polymerase, and 3' RACE was also made with Vent polymerase (New England Biolabs, Bevery, MA). For first round PCR, 1/100 of the first strand cDNAs was used as a template. For the second round PCR, 1/1,000 of the first round PCR products was used as a template.

PCR products were examined by electrophoresis in 1.5% agarose gel with ethidium bromide staining and by Southern blot analysis. DNA was transferred to nitrocellulose membrane (Schleicher & Schuell, Keene, NH) and then was immobilized by UV cross-linking. Membranes were soaked for 2 h at 42°C in a prehybridization solution of 30% formamide, $6 \times SSPE$ [12], $5 \times Denhardt's$ solution [12], 100 μg/ml salmon sperm DNA, 0.1% SDS then hybridized overnight at 42°C with ³²P-labeled oligonucleotide probe (1 × 106 cpm per ml of

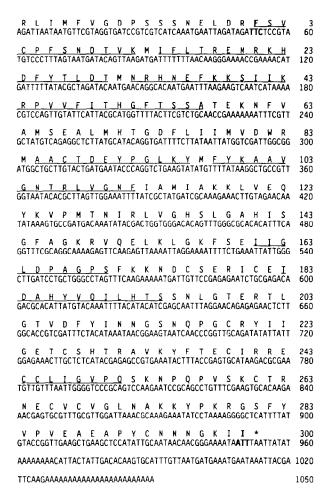


Fig. 1. cDNA and amino acid sequences of hornet phospholipase A₁ (*Dol m* 1). Nucleotide and amino acid positions are numbered on the right. Numbering of amino acid residues begins and ends at its N- and C-termini of phenylalanine and isoleucine, corresponding to nucleotide positions of 52–54 and 949–951 and the amino acid residues and nucleotides are shown in bold characters. The underlined amino acid residues were also established by Edman degradation of CNBr peptides. These sequence data are available from EMBL/Genbank/DDBJ under accession number X66869.

prehybridization solution). Post hybridization membranes were twice washed for 20 min at 60° C in a solution of 3 M tetramethylammonium chloride, 0.2% SDS and 0.05 M Tris-HCl, pH 8.0 [13]. Oligonucleotides of specific activity $5 \times 10^7 - 10^8$ cpm/ μ g were labeled with $[\gamma^{-32}$ P]ATP (New England Nuclear Corp.) in presence of T4 polynucleotide kinase (New England Biolabs). The labeling procedure as well as other molecular biology procedures were taken from [12].

PCR products contain single 3'-overhanging A-nucleotides [14] and were used directly for cloning into the pCR vector with compatible T-nucleotide overhangs (Invitrogen Corp., San Diego, CA). Plasmid DNAs were isolated from appropriate clones using the QIAGEN plasmid kit (QIAGEN, Chatworth, CA). DNA sequences were determined by the dideoxynucleotide chain-termination method [15] using alkaline-denatured plasmid DNAs and the Sequenase version 2.0 kit (US Biochemical, Cleveland, OH).

2.3. Phospholipase and lipase assays

Phospholipase activity was measured titrimetrically at 24 ± 1 °C and pH 8 in a volume of 1.5 ml with 0.2–2 μg of enzyme and 10% egg yolk as substrate in 0.2 N NaCl containing 0.5% Triton [2]. One unit

of activity is defined as the release of one μ mol of acid per min per mg of enzyme. Lipase activity was measured similarly using either a 2% solution of triacetin or a 2% emulsion of tributyrin (Sigma Biochemical, St. Louis, MO) as substrates in 0.2 N NaCl containing 0.5% Triton.

3. RESULTS

3.1. Partial amino acid sequence of Dol m I

These data were obtained from its CNBr peptides. The partial or complete sequences of 7 of these peptides correspond to residues 1–12, 14–30, 32–57, 85–96, 98–112, 161–170, 183–194 and 244–251 of the molecule shown in Fig. 1. The first five peptides correspond to the expected cleavage as it is in each case either preceded or terminated with a methionine residue. The last three peptides represent side products from acid cleavage of glutamyl peptide bonds. These partial amino acid sequence data were used for the design and synthesis of oligonucleotides 4 to 7 in Table I.

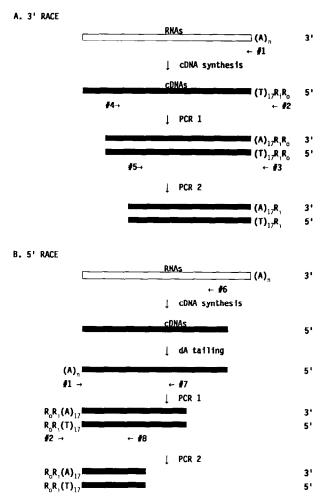


Fig. 2. Schematic diagram for rapid amplification of the 3' and 5' cDNA ends (RACE) of *Dol m* I. Open and solid bars represent RNA and DNA, respectively. The oligonucleotide primers are numbered, and their sequences are given in Table I.

3.2. cDNA sequence of Dol m I

cDNA encoding amino acid residues 22–300 and the 3'-untranslated region was amplified from venom RNAs by the RACE procedure as outlined in Fig. 2A. Single stranded venom cDNAs were synthesized from total RNAs using a dT primer with R₁ + R_o adaptor (oligonucleotide 1 in Table I). Double stranded *Dol m* I-specific cDNA was amplified from single stranded venom cDNAs by two successive rounds of PCR using the nested primers as indicated. Several PCR products were detected and a major band of about 1 kb (Fig. 3) appeared to be the expected product when tested on Southern blot by hybridization with oligonucleotide 6 (Table I). As shown in Fig. 3 the 1 kb band was only found when Taq polymerase was used and it was not found with Vent polymerase.

The PCR products which contain the 1 kb band were cloned directly into plasmids. After transformation into bacteria, plasmids from 3 colonies were selected and sequenced. Two colonies have the nucleotide sequence of 115 to 1,050 in Fig. 1. One of them differs from that shown: by the deletion of one adenine base at position 968, and by the insertion of an additional 99 nucleotides at position 1,027 (these data are not shown but are on deposit at EMBL/Genbank). A third colony differs from that shown at position 807 (C to T substitution;

both encoding serine) and at position 812 (A to G substitution; asparagine to serine change).

Using the cDNA data of Fig. 1, oligonucleotides 8 and 9 in Table I were synthesized for amplifying the cDNA region which is 5' of nucleotide 115 in Fig. 1. As shown in Fig. 2B, single stranded Dol m I-specific cDNA was synthesized from total RNAs using oligonucleotide 6 as the primer then poly-dA tailed with terminal deoxynucleotidyl transferase. Double-stranded Dol m I-specific cDNA was amplified from poly-dA tailed specific cDNA by two successive rounds of PCR with the indicated primers. Several products formed after the second round of amplification and two bands of about 0.32 and 0.25 kbp (Fig. 3) appeared to be the expected products when detected on Southern blot by hybridization with oligonucleotide 9 in Table I. Following cloning into a plasmid, the product of 0.32 kbp was established to contain the cDNA sequence from nucleotide 1 to 262 in Fig. 1.

The region preceding nucleotide position 52 in Fig. 1 encodes a leader sequence of 17 amino acid residues as the N-terminal amino acid residue of *Dol m* I was found on Edman degradation to begin at nucleotide position 52. The protein sequence suggests the presence of two possible glycosylation sites at residue 8 and 212. The site at residue 8 is probably glycosylated as re-

Table I

Oligonucleotides used as primers or probes for cloning hornet phospholipase

No.	Oligonucleotide*	Comment
1	AAG GAT CCG TCG ACA TCG ATA ATA CGA CTC ACT ATA GGG ATT T_{15}	(dT) ₁₇ R ₁ R ₀ primer for first strand cDNA synthesis of 3' RACE
2	AAG GAT CCG TCG ACA TC	R _o anti-sense primer for first round PCR of 3' RACE
3	GAC ATC GAT AAT ACG AC	R, anti-sense primer for second round PCR of 3' RACE
4	D° T V K M I ¹⁴ GAY ACI GTI AAR ATG AT	Sense primer for first round PCR of 3' RACE
5	K ²² H D F Y T ²⁷ AAR CAY GAY TTY TAY AC	Sense primer for second round PCR of 3' RACE
6	I ¹⁹⁰ Q V Y H A D ¹⁸⁴ AT YTG IAC RTA RTG IGC RTC	Hybridization probe of PCR produce of 3' RACE; or primer for first strand cDNA synthesis of 5' RACE
7	P ⁹² Y E D T C ⁸⁷ GG RTA YTC RTC IGT RCA	Anti-sense primer for first round PCR of 5' RACE
8	$ exttt{M}^{70}$ L A E $ exttt{S}^{66}$ G CAT AAG AGC CTC TGA C	Anti-sense primer for second round PCR of 5' RACE
9	${ m M}^{31}$ T D L T 27 T CAT TGT ATC TAG CGT A	Hybridization probe for PCR product of 5' RACE

^{*}R represents A or G, Y represents C or T, and I represents inosine.

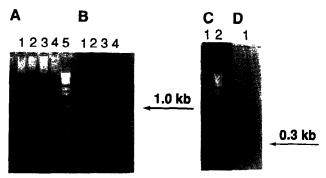


Fig. 3. 3' and 5' RACE of white-faced hornet phospholipase-specific cDNA. In panels (A) and (B) are shown, respectively, the agarose gel electrophoresis and Southern blot analysis products for 3' RACE. In lanes 1 and 3 are shown the products from first and second round PCR obtained with AmpliTaq DNA polymerase, in lanes 2 and 4 are shown similar products obtained with Vent polymerase; and in lane 5 is shown a 1 kb DNA Ladder (BRL). In panels (C) and (D) are shown similar results (as in panels (A) and (B)) for 5' RACE products (lane 1) obtained with AmpliTaq DNA polymerase; and in lane 2 (panel (C)) is shown 1 kb DNA Ladder. The arrows in panels (B) and (D) indicate the desired products. The hybridization probes are given in Table I.

peated attempts to identify this residue by Edman degradation gave negative results. This is also suggested by the difference in the molecular weight of 33,745 calculated from the sequence and that of about 37,000 estimated from SDS gel electrophoresis.

3.3. Lipase activity of hornet phospholipase

We reported previously [3] that vespid phospholipases catalyze a rapid hydrolysis of the acyl group at position 1 of synthetic phosphatidylcholines and a slow hydrolysis of the acyl group at position 2. Therefore vespid phospholipases have both A_1 and B types of phospholipase activities. The present finding on sequence similarity of hornet phospholipase with lipases prompted us to test for lipase activity.

The enzyme sample tested had about 280 units of phospholipase activity per mg when tested with egg yolk as a substrate as compared to the previously reported specific activity of 1,100 units per mg [3], and its low specific activity was due to inadvertent prolonged exposure to low pH. This sample had the following lipase activities of 13 and 33 (± 20%) units/mg with triacetin and tributyrin, respectively, as substrates. These data suggest hornet phospholipase has a weak lipase activity.

4. DISCUSSION

Sequence comparison by the FASTA method [11] showed that hornet phospholipase Dol m I has no similarity with other known phospholipases in the literature, but it has similarity with mammalian lipases. This is shown in Fig. 4 for lipoprotein lipases and hepatic lipases from human and mouse [16–19]. Human pancreatic lipase [20] has about the same degree of similarity with Dol m I as human hepatic lipase (data not shown).

There is about 40% identity in overlaps of 123 residues of mammalian lipases and *Dol m* I. The sequence region of lipases shown in Fig. 4 is highly conserved as similar sequences are found for a number of other mammalian and prokaryotic lipases and a *Drosophila* protein vitellogenin [21,22]. Thus these proteins also have significant sequence similarity with *Dol m* I.

The most strongly conserved region of all lipases is reported to be in the undecapeptide region of residue 153-163 of human lipoprotein lipase [21]. This region is believed to be of importance for lipase activity, and it is the region of highest identity of lipases and *Dol m* I. Tests with two synthetic triglycerides suggest the presence of a weak lipase activity for *Dol m* I.

All vespid allergic patients invariably have antibodies specific for both $Dol\ m$ I and V. Therefore we compared the sequences of these two proteins and they are found to share one similar octapeptide sequence; VNRHNQFR and LKRHNDFR at position 45–52 of $Dol\ m$ VA and B, respectively, and MNRHNEFK at position 31–38 of $Dol\ m$ I. However this octapeptide sequence is not in the sequence region where these allergens show similarity with other proteins.

There are several examples of sequence similarity of allergens with other proteins in our environment. They are birch pollen allergen $Bet\ v\ I$ with a pea disease resistance response protein gene [23], another birch pollen allergen and its homologs from timothy and mugwort pollens with human profilin [24], ragweed pollen allergens $Amb\ a\ I$ and II with proteins from tomato anther and pistil (cf. [25]), cat allergen $Fel\ d\ I$ with rabbit uteroglobin [26], mite allergen $Der\ p\ I$ with human cathepsin and other cysteine proteases [27], bee venom allergen phospholipase A_2 with human pancreatic enzyme and bee venom allergen melittin $Api\ m\ III$ with

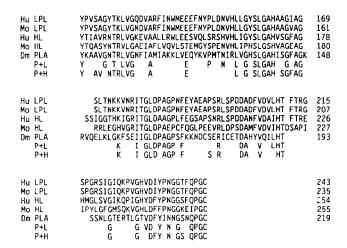


Fig. 4. Sequence similarity of *Dol m* I and mammalian lipases. Amino acid positions are numbered on the right. Abbreviations used: Hu, human; Mo, mouse; LPL, lipoprotein lipase; HL, hepatic lipase; Dm, white face hornet; and PLA, phospholipase. P + L and P + H indicate residues of hornet phospholipase which are identical to human lipoprotein or hepatic lipases, respectively.

human complement C9 (cf. [6]). However several other major allergens from mite [28] and ragweed and grass pollens [29–31] have no known sequence similarity with other proteins in our environment.

The accumulated sequence data makes it possible to analyze whether the T and/or B cell epitope of these allergens are located in the regions of sequence similarity with other proteins in our environment. Such analysis will provide the necessary proof whether sequence similarity plays a role in the immunogenicity of allergens in man.

Acknowledgements: We thank Mr. Miles Guralnick of Vespa Laboratories for gifts of hornets and hornet venom. Also we thank Donna Atherton, Sheenah Mische and the staff of the Rockefeller University Protein Sequencing Facility for amino acid sequence analysis and oligonucleotide synthesis.

REFERENCES

- [1] King, T.P., Sobotka, A., Alagon, A., Kochoumian, L. and Lichtenstein, L.M. (1978) Biochemistry 17, 5165-5174.
- [2] King, T.P., Kochoumian, L. and Joslyn, A. (1984) Arch. Biochem. Biophys. 230, 1–12.
- [3] King, T.P., Joslyn, A. and Kochoumian, L. (1985) J. Allergy Clin. Immunol. 75, 621–628.
- [4] Fang, K.S.F., Vitale, M., Fehlner, P. and King, T.P. (1988) Proc. Natl. Acad. Sci. USA 85, 895-899.
- [5] Lu, G., Villalba, M., Coscia, M., Hoffman, D. and King, T.P. (1993) J. Immunol. 150.
- [6] King, T.P., Moran, D.C., Wang, D.F., Kochoumian, L. and Chait, B. (1990) Protein Sequences and Data Analysis 3, 263 266.
- [7] Marsh, D.G., Goodfriend, L., King, T.P., Lowenstein, H. and Platts-Mills, T.A.E. (1987) J. Allergy Clin. Immunol. 80, 639-645.
- [8] Kapsenberg, M., Jansen, H., Bos, J. and Wierenga, E. (1992) Curr. Op. Immunol. 4, 788-793.
- [9] Vercelli, D. and Geha, R. (1992) Curr. Op. Immunol. 4, 794-797.
- [10] Frohman, M. (1990) Amplifications: A Forum for PCR Users 5, 11-15.
- [11] Frohman, M., Dush, M. and Martin, G. (1988) Proc. Natl. Acad. Sci. USA 85, 8998–9002.
- [12] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning, Vols. 1 and 2, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

- [13] Wood, W., Gitschier, J., Lasky, L. and Lawn, R. (1985) Proc. Natl. Acad. Sci. USA 82, 1585-1588.
- [14] Clark, J. (1988) Nucleic Acids Res. 16, 9677-9686.
- [15] Sanger, F., Nicklen, S. and Coulson, A. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
- [16] Wion, K., Kirchgessner, T., Lusis, A., Schotz, M. and Lawn, R. (1987) Science 235, 1638-1641.
- [17] Kirchgessner, T., Svenson, K., Lusis, A. and Schotz, M. (1987) J. Biol. Chem. 262, 8463-8466.
- [18] Martin, G., Busch, S., Meredith, G., Cardin, A., Blankenship, D., Mao, S., Rechtin, A., Woods, C., Racke, M., Schafer, M., Fitzgerald, M., Burke, D., Flanagan, M. and Jackson, R. (1988) J. Biol. Chem. 263, 10907-10914.
- [19] Oka, K., Nakano, T., Tkalcevic, G., Scow. R. and Brown, W. (1991) Biochim. Biophys. Acta 1089, 13-20.
- [20] Winkler, F., D'Arcy, A. and Hunziker, W. (1990) Nature 343, 771, 773
- [21] Persson, B., Bengtsson-Olivecrona, G., Enerback, S., Olivecrona, T. and Jornvall, H. (1989) Eur. J. Biochem. 179, 39-45.
- [22] Bownes, M., Shirras, A., Blair, M., Collins, J. and Coulson, A. (1988) Proc. Natl. Acad. Sci. USA 85, 1554-1557.
- [23] Breiteneder, H., Petternburger, K., Bito, A., Valenta, R., Kraft, D., Rumpold, H., Scheiner, O. and Breitenbach, M. (1989) EMBO J. 8, 1935–1938.
- [24] Valenta, R., Duchene, M., Ebner, C., Valent, P., Sillaber, C., Deviller, P., Ferreira, F., Tejkl, M., Edelmann, H., Kraft, D. and Scheiner, O. (1992) J. Exp. Med. 175, 377-385.
- [25] Griffith, I., Pollock, J., Klapper, D., Rogers, B. and Nault, A. (1991) Int. Arch. Allergy Appl. Immunol. 96, 296-304.
- [26] Morgenstern, J., Griffith, I., Brauer, A., Rogers, B., Bond, J., Chapman, M. and Kuo, M. (1991) Proc. Natl. Acad. Sci. USA 88, 9690-9694.
- [27] Chua, K.Y., Stewart, G.A. and Thomas, W.R. (1988) J. Exp. Med. 167, 175-182.
- [28] Chua, K.Y., Dilworth, R.J. and Thomas, W.R. (1990) Int. Arch. Allergy Appl. Immunol. 91, 124-129.
- [29] Griffith, I.J., Smith, P.M., Pollock, J., Thecrakulpisut, P., Avijioglu, A., Davies, S., Shough, T., Singh, M.B., Simpson, R.J., Ward, L.D. and Knox, R.B. (1991) FEBS Lett. 279, 210-215.
- [30] Silvanovich, A., Astwood, J., Zhang, L., Olsen, E., Kisil, F., Sehon, A., Mohapatra, S. and Hill, R. (1991) J. Biol. Chem. 266, 1204–1210.
- [31] Singh, M.B., Hough, T., Theerakulpisut, P., Avjioglu, A., Davies, S., Smith, P.M., Taylor, P., Simpson, R.J., Ward, L.D., McCluskey, J., Puy, R. and Knox, R. (1991) Proc. Natl. Acad. Sci. USA 88, 1384–1388.