Sequence, Structure, and Expression of a Wasp Venom Protein with a Negatively Charged Signal Peptide and a Novel Repeating Internal Structure

(Received for publication, January 22, 1992)

Davy Jones‡‡, Grzegorz Sawicki‡, and Mietek Wozniak‡
From the ‡Graduate Center for Toxicology and §School of Biological Sciences, University of Kentucky, Lexington, Kentucky 40506

An expression cDNA library prepared from mRNA from the venom gland of a parasitic wasp, Chelonus sp. near curvimaculatus, was screened with polyclonal antibodies against a 33-kDa venom protein from this wasp. Immunoreactive clones were sequenced, yielding a complete inferred sequence for a protein with an NH₂ terminus identical with that of the 33-kDa protein. The structure of the cDNA showed an apparent encoded signal peptide, which was unusual in possessing 2 glutamic acid residues juxtapositioned next to, or replacing, the conventional basic residues. The bulk of the mature protein sequence which follows the NH₂-terminal, 5000-kDa hydrophobic domain is composed of a dozen tandem repeats of a highly charged, approximately 14-residue sequence, except for the truncated eighth repeat which terminates in the only proline in this large domain. The primary structure is not closely related to any sequence in the GenBank data bank. Secondary structure analysis identified a turn occurring at or near each of 12 invariantly conserved serine residues. Further, the codon used for this serine was invariant, whereas other serines in the protein (including a nearly invariant serine 2 residues away) used various codons. Results of epitope mapping experiments supported a proposed tertiary structure in which the NH₂-terminal 5-kDa forms a hydrophobic core, overlain with the charged repeats. Northern analysis of poly(A) RNA from the venom gland of young adult female wasps showed expression of a single 1-kilobase transcript, for which there is no corresponding message in normal or parasitized host larvae. The remarkable structure of this protein and structural data on other wasp venom proteins suggest an evolutionary pattern in which some proteins critical for venom function evolve by internal tandem duplication, and which are secreted after biosynthesis by a different mechanism from that used for proteins with classical signal peptides containing basic residues.

Venom proteins comprise a group of biologically active molecules with interesting and useful properties and modes of action. Venoms from snakes have been most extensively investigated, and a number of individual proteins have been purified and characterized from snake venom. Examples of such proteins are phospholipases (1) and smaller polypeptides which are neurologically active. Venom proteins have been identified which interact with presynaptic targets (2) and post-synaptic targets such as the acetylcholine receptor (3). Toward an understanding of their molecular mechanisms of action, antigenicity, and of serological neutralization, several of such snake venom proteins have been characterized by x-ray crystallography and their epitopes mapped (4, 5). Mapping of neutralizing epitopes has been particularly useful in identifying functional groups which interact with the target site.

Arthropod venoms have also been of general interest as a source of medically important and pharmacologically useful proteins, such as honey bee melitin and apamin (6, 7). Phospholipases from the honey bee and other higher Hymenoptera have also been isolated (8). Additional characterization of biologically active components of hymenopteran venoms have produced the structures of kinins and mastoparans (9) and the cloning of a hornet antigen (10). Less well known are the venoms of the parasitic Hymenoptera, although there are many more species of such parasites, each contending with different hosts with different biochemical templates. Partial purification has been reported for paralyzing components of venom of several paralyzing wasps (11), but no sequencing, epitope mapping, or other molecular information is published. Detailed characterization of biologically active protein components of nonparalyzing venoms of parasitic wasps is not available, except for research which we have conducted on the endoparasitic wasp Chelonus sp. near curvimaculatus.

Recent studies have shown that a function of the nonparalyzing venom of some endoparasitic wasps is to suppress the immune response of the host caterpillar (12). We have identified and isolated a 33-kDa protein from the venom of C. sp. near curvimaculatus which is necessary for the survival of the larval endoparasite in its lepidopteran host, Trichoplusia ni (13). The NH₂-terminal sequence and cyanogen bromide fragmentation map of this protein are not obviously related to any of the previously reported venom proteins of any arthropod. However, conserved epitopes have been found on venom proteins of other Chelonus sp. and related genera (14). Furthermore, epitopes on this and other venom proteins in Chelonus are immunologically conserved in the venoms of a wide variety of higher predatory and social Hymenoptera, including venoms whose constituents possess strong allergic properties (15). This wasp and its close relatives use a wide variety of lepidopteran hosts, suggesting that this protein that is necessary for parasite survival is active against a biochemical
pathway fundamental and common to many insects.

In this report we describe the isolation and sequence of full-length cDNA for the protein, probing of tertiary structure through epitope mapping, and assessment of expression of the transcript in the wasp and its host.

**Materials and Methods**

**Insects—**Parasitic wasps (C. neotropicalis) were reared in culture with the host (T. nil) as described previously (16). Usually, "naive" wasps (which had not stung any host eggs), 1–2 days old, were used for dissection of venom glands. The venom glands were dissected in 7% sucrose containing 150 mM NaCl and 20 mM Tris-HCl buffer, pH 6.8 (TBS), then homogenized manually and centrifuged 20 min at 4 °C and 10,000 × g.

**Protein Purification—**The 33-kDa venom protein was purified as described elsewhere (13) using a Mono-Q anion-exchange column (Pharmacia LKB Biotechnology Inc.) and high performance liquid chromatography Gold System (Beckman). Extracts from 200–300 venom glands were used for each protein purification. After separation on the Mono-Q column, the pure fractions of the 33-kDa venom protein were concentrated about 10-fold with a Centricon device and stored in TBS solution at −70 °C.

**Chemical Fragmentation in CNBr—**Pure 33-kDa protein was diluted with 1 ml of water and concentrated in a Centricon device to 50 μl and then lyophilized in a Speedvac concentrator. Dry protein was dissolved in 70% trifluoroacetic acid and 0.1 mg of CNBr was added (17). After 24 h of incubation at room temperature under nitrogen, the sample was dried by Speedvac concentration and dissolved in sample buffer for electrophoresis in a polyacrylamide gel (10–20% concentration gradient of acrylamide) with sodium dodecyl sulfate (SDS-PAGE).

**Synthesis of Octapeptide and Conjugate—**An octapeptide corresponding to residues 1–8 of the NH2 terminus of the 33-kDa protein (IFSDDDVL) (13) was synthesized by Immuno-Dynamics (La Jolla, CA) according to the method of Margin and Marrifield (18). The octapeptide was conjugated to a succinyl residue from the NH2-terminal side for coupling to keyhole limpet hemocyanin, or BSA, by m-maleimidoobenzoyl-N-hydroxysuccinimide ester (19).

**Immunization—**Rabbit antiserum against octapeptide-keyhole limpet hemocyanin conjugate was raised by subcutaneous injection of 2 mg of conjugate emulsified with complete Freund’s adjuvant followed by four booster injections of 2 mg of conjugate with incomplete Freund’s adjuvant at monthly intervals.

**Gel Electrophoresis and Immunoblotting—**SDS-10% PAGE of the protein or its fragments, and then electrotransfer from the gel to nitrocellulose followed by immunoblotting, was performed as described (20). After electrotransfer, the nitrocellulose was incubated in blocking solution (20 mM Tris-HCl buffer, pH 7.4, 150 mM NaCl, 20% horse serum, and 0.1% Triton X-100), washed three times with the same buffers and one time with 20 mM Tris-HCl, pH 6.8, buffer, boiled with sample loading buffer, and then subjected to SDS-PAGE. After electrophoresis in SDS-PAGE, the proteins were transferred to nitrocellulose and probed with a mixture of two antisera: antiserum against native venom proteins and antiserum against denatured 33-kDa venom protein.

**Nucleic Acid Hybridizations—**Northern hybridization was performed on poly(A) RNA electrophoretically fractionated in denaturing agarose gels as described previously (25, 26). The recombinant plasmid containing the cDNA for the full sequence for the 33-kDa protein was radiolabeled with [32P]dCTP as described (27). Hybridizations were performed on 0.7% agarose gels. Northern blots were hybridized to [32P]cDNA probes obtained from inserts of at least 1 kb by PCR amplification of inserts from eluted phage plaques. Selected clones were then removed into Bluescript plasmid form and then sequenced using dinucleotide sequencing (28).

Isolation of cDNA Clones—Poly(A) RNA was extracted from several thousand venom glands from female wasps by the procedure of LeMeur et al. (21). Poly(A) RNA was selected by one pass through an oligo(dT)-cellulose matrix (22) and used as a template to make first strand cDNA with avian myeloblastosis virus reverse transcriptase and an oligo(dT) primer. A degenerate primer encoding the NH2-terminal sequence reported previously was used, along with an oligo(dT) primer, to amplify by PCR from the total first strand cDNA the specific CDNA encoding the protein, so as to obtain an initial estimate of the size of anticipated cDNA clones (primer coding sequence: AT(TCA) TT(CT) TCN TT(TC) GA(TC) GA(TC) (CT)T). The library was screened with polyclonal antibodies (14) against the 33-kDa protein, and positive clones were plaque purified and rescreened. Clones positive in rescreening were verified to have inserts of at least 1 kb by PCR amplification of inserts from eluted plaques. Selected clones were then removed into Bluescript plasmid form and sequenced by double-sanded dideoxy sequencing (23). Clones were verified as corresponding to the 33-kDa protein by comparison of the encoded sequence with the NH2-terminal sequence obtained by direct protein sequencing (13). The entire sequence was verified by obtaining the sequence for each region from several independent clones.

**In Vitro Transcription/Translation and Immunoprecipitation—**A clone possessing the full coding sequence for the 33-kDa protein, as well as an inserted signal peptide and 5'-untranslated sequences, was transcribed using T3 RNA polymerase as described by the supplier. The transcript was then translated in a cell-free rabbit reticulocyte lysate system as described (24). Radiolabeled (35S) methionine products were directly subjected to SDS-PAGE or further immunoprecipitated using the monoclonal antiserum in the same manner as for the mature, native 33-kDa protein. Extracts from 40 venom glands in 20 μl of TBS were diluted to 200 μl with TBS, 0.1% Triton X-100, centrifuged 10 min at 10,000 × g, raised to 200 μl, and incubated with antibody for 4 h at 4 °C. The antibodies used for incubation were 5 μl of preimmune serum, 5 μl of anti-NH2-terminal octapeptide, 1 μl of antiserum against all venom proteins, or a mixture of this latter antiserum with 1 μl of antiserum against denatured, purified 33-kDa protein. After incubation, 10 μg of protein A-Sepharose was added to the RIPA buffer and incubated an additional 1 h. The mixture was incubated again for 1.5 h. Finally, the pellet was washed three times with the same buffers and one time with 20 mM Tris-HCl, pH 6.8, buffer, boiled with sample loading buffer, and then subjected to SDS-PAGE. After electrophoresis in SDS-PAGE, the proteins were transferred to nitrocellulose and probed with a mixture of two antisera: antiserum against native venom proteins and antiserum against denatured 33-kDa venom protein.

**Protein Electrophoresis—**Mature proteins and primary translation products were subjected to SDS-PAGE as described by Lammle (29). Enzymes and Reagents—Commercial sources of enzymes and reagents were as follows: avian myeloblastosis virus reverse transcriptase, New England Biolabs; T4 DNA polymerase, New England Biolabs; RNase H, Promega; Taq polymerase, Perkin Elmer Cetus; T3 RNA polymerase, Boehringer Mannheim; rabbit reticulocyte lysate system (Promega) with [35S]methionine. The translation products were directly subjected to SDS-PAGE or first immunoprecipitated with polyclonal antibodies prior to SDS-PAGE.

1 The abbreviations used are: TBS, tris-buffered saline; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; kb, kilobases(s); BSA, bovine serum albumin; PCR, polymerase chain reaction.
protein. The sequence of the clone with the longest 5' end is shown in Fig. 1. The complete sequence of the 894-base cDNA shows an encoded protein sequence from position 106 to 165 which corresponds exactly to the NH₂-terminal protein sequence obtained through direct protein sequencing (13). A methionine codon at position 253 is situated so as to predict an NHz-terminal signal peptide of 20 amino acids and preceded by an acidic residue (lysine) immediately follows the lysine residue. These two glutamic acid codons were confirmed in sequencing of independent clones.

The theoretical isoelectric point of 4.55 for the encoded protein is exactly the result observed here (see below). The primary structure analysis of the mature protein is not closely related to any sequence in the GenBank data bank. The acidic nature of the protein is due mostly to residues which are in the carboxyl region, giving the carboxyl end a strongly polar characteristic, while the NH₂-terminal end is relatively much more hydrophobic.

The encoded protein possesses an interesting internal structure. After the NH₂-terminal 5,000-kDa region there abruptly begins a series of 12 tandem repeats of usually 14 residues (Fig. 2). Analysis of the most common functional group at each position (e.g., acidic, basic, hydrophobic, etc.) yields a consensus sequence of

\[ \text{A/V D/E X/V/L S G S Z D/E Q X K/R X S/T} \]

where Z indicates a hydrophobic group. Within this consensus sequence are several invariant or highly conserved residues. For example, positions 5 and 7 are serine residues conserved in 12, and 11 of the repeats, respectively, while glycine at position 6 is conserved in 10 repeats. Comparison of the sequences between repeats shows several distinct groupings. For example, repeats 2, 3, and 11 uniquely share the sequence AEXLSGSFD for the first 9 residues, while repeats 4 and 12 share the sequence VDEKSGSVG for this stretch. Repeats 1 and 9 can be grouped by their unique sharing of the sequence.

The protein predicted by the primary structure is not closely related to any sequence in the GenBank data bank. The acidic nature of the protein is due mostly to residues which are in the carboxyl region, giving the carboxyl end a strongly polar character, while the NH₂-terminal end is relatively much more hydrophobic.
truncated eighth repeat. Within each repeat is the invariant predict regions of the protein for α-helix predicted many of the same turns as the Garnier method. The proline at position 7.

VKQKRXT for positions 8–13. Clearly in a category by itself terminal 5-kDa segment is hydrophobic, with little α-helical structure, was encoded elsewhere variously as GGA, GGC, or GGT, it used various serine codons, the first serine of this motif was invariantly encoded as TCA within this conserved tripeptide motif.

Secondary Structure Analysis—The protein appears to consist of three major, identifiable regions (Fig. 2): a leading hydrophobic 5,000-kDa portion, a larger central portion composed of repeats 1–7, and a final region consisting of repeats 9–12. Examination of the inferred secondary structure of the leading hydrophobic domain suggests it is composed primarily of uncharged or weakly charged stretches, and sequences with β-sheet propensity, connected by regions predicted to have a strong turning character.

The second and third domains are dominated by a repeating structure of α-helices, or charged extended stretches separated by regions with a strong turning/coil component. These turns are located at or near the Ser-Gly-Ser motif containing the invariant serine (position 5, Fig. 2). The first of the two α-helices with the strongest amphipathic structure (32, 33) is located across positions 6–14 of the first repeat and extending to the first two positions of the second repeat. Such amphipathic helices are often on the surface of a protein, with the charged face of the helix oriented externally. This region is also predicted by the method of Hopp and Woods (34) to be strongly antigenic (see below).

Specificity of Antisera—Antisera generated against either total venom proteins, the 33-kDa protein, or the NH₂-terminal octapeptide possessed specificities corresponding exactly to the original antigens. Antibodies against native venom proteins interacted with all major proteins visualizable with silver staining even after denaturing SDS-PAGE, including the 33-kDa protein (Fig. 3, lanes A1 and B1). Antibodies generated against the SDS and heat-denatured 33-kDa protein yielded a strong signal only for the 33-kDa antigen (Fig. 3, lanes A2 and B2). Neither of these antisera interacted with the NH₂-terminal octapeptide-BSA conjugate (Fig. 3, lanes

![FIG. 2. Secondary structure predictions of the encoded mature 33-kDa protein.](image)

**FIG. 2.** Secondary structure predictions of the encoded mature 33-kDa protein. The region of the protein consisting of the 12 tandem repeats is indicated by the numbers on the left. The methods of Garnier et al. (51), in PC Gene software, were used to predict regions of the protein for α-helix (→→→), extended hydrophobic (←←←), and turn/coil (−−−) forms. The method of Novotny and Auffrey (52) was used to predict the locations of β-sheet structure (−−−−). The method of Chou and Fasman (53) predicted many of the same turns as the Garnier method. The consensus predictions of the methods employed are that the NH₂-terminal 5-kDa segment is hydrophobic, with little α-helical structure, while the hydrophilic repeats are characterized by extensive α-helical structure between periodic turns centered around the conserved Ser-Gly-Ser motif. Positions in the repeat at which at least half of the residues are of a single functional type are deemed to have a consensus residue or class, which is shown in the consensus motif provided underneath. The letter Z designates a hydrophobic residue. Positions 5, 6, and 7 are the most strongly conserved, with a consensus motif of SGS. Boldface letters indicate residues which are the same as the consensus sequence.

VKQKRXT for positions 8–13. Clearly in a category by itself is repeat 8, which is truncated and uniquely terminates in a proline at position 7.

**Codon Usage**—An intriguing conservation of codons occurs for the Ser-Gly-Ser motif that is invariant in all but the truncated eighth repeat. Within each repeat is the invariant octanucleotide sequence TCA-GGA-TC, with TCA-GGA-TCN encoding the conserved tripeptide. Although the second of the 2 serines, and other serines elsewhere in the protein, used various serine codons, the first serine of this motif was invariantly encoded as TCA. In addition, although glycine was encoded elsewhere variously as GGA, GGC, or GGT, it was strongly antigenic.

![FIG. 3. Interaction of total venom proteins with various antisera.](image)

**FIG. 3.** Interaction of total venom proteins with various antisera. Panel A, SDS-10% PAGE gel after silver staining. Lane 1, venom gland extract; lane 2, purified venom 33-kDa protein. Panel B, immunoblot of the venom gland extract probed with: lane 1, polyclonal antibodies against all (native) venom proteins; lane 2, polyclonal antibodies against denatured 33-kDa protein; lane 3, antibodies against NH₂-terminal peptide of 33-kDa protein. Panel C, NH₂-terminal octapeptide conjugated with BSA (all three lanes) and probed with the same set of antibodies as panel B, lanes 1–3, respectively. The NH₂-terminal antibodies react with the 33-kDa protein (lane B3) and the octapeptide-BSA conjugate (lane C3), while the other antibodies have epitopes elsewhere on the protein and not on the NH₂ terminus (lanes C1 and C2).
C1 and C2). However, the antibodies generated specifically against the NH$_2$-terminal octapeptide reacted strongly with the oligopeptide-BSA conjugate (Fig. 3, lane 3), as well as with the NH$_2$ terminus of heat- and SDS-denatured 33-kDa protein (Fig. 3, lane B3). These results demonstrate that while antibodies can be generated specifically against the octapeptide, this fragment of the protein is not antigenic under standard rabbit immunization conditions, using either native or denatured 33-kDa protein.

Mapping of Epitopes of NH$_2$-terminal Antiserum on CNBr Fragments—Toward immunological testing of the tertiary structure of the protein, the location of epitopes for each antiserum was determined. The protein was subjected to partial cyanogen bromide fragmentation, yielding fragments of 27–27.5 and 4.5–5 kDa and a portion of uncleaved protein. The location of intact epitopes for each antiserum were then determined by immunoblotting the fragments after denaturing SDS-PAGE. The NH$_2$-terminal octapeptide antibodies reacted with the free end of the denatured uncleaved protein and with the 5-kDa fragment (Fig. 4, lanes B1 and B2), which also confirmed the inferred MET residue encoded at position 253 of the cDNA (Fig. 1). In addition, the NH$_2$-terminal antiserum reacted with the positive control samples (the octapeptide, Fig. 4, lane B3; the octapeptide-BSA conjugate, Fig. 4, lane B4). Thus, the NH$_2$-terminal antiserum possesses an epitope(s) on the NH$_2$ terminus that is not duplicated elsewhere on the protein.

Mapping of Epitopes of Anti-total and Anti-33-kDa Protein Antisera—The 5-kDa NH$_2$-terminal fragment did not contain epitopes for antisera generated against either the native protein or the denatured protein, since a mixture of these two antisera did not react with this fragment (Fig. 4, C1). However, the antibodies against the native mature 33-kDa protein strongly cross-react with the denatured 27-kDa fragment (13). Consistent with the inferred localization of epitopes on the 27-kDa carboxyl fragment, and not on the NH$_2$ terminus for antisera against denatured anti-33-kDa protein, was the immunological selection with that antisera of a cDNA clone truncated so as to eliminate the first 10 residues (1 kDa) of the mature protein (not shown). Thus, there is at least one epitope on the protein for these antisera that is localized to the region of tandem repeats that comprises the 27 kDa end of the protein. This epitope also appears to be a linear epitope, comprised of a sequential group of amino acids (i.e. the epitope is stable to denaturing SDS-PAGE). The accessibility under native conditions of the epitopes located in the 27-kDa region was further tested by immunoprecipitation of native 33-kDa protein. As shown in Fig. 5, polyclonal antibodies which react with the 27-kDa fragment, but not with the 5-kDa fragment, were capable of immunoprecipitating the native protein. These data strongly support the hypothesis that at least one epitope is found on a tandem repeat which under native conditions is on the surface of the mature protein.

Mapping of Epitope on Tandem Repeat to Native Protein Surface—Further resolution of the location of surface epitopes was facilitated by immunodetection, under native conditions, of fusion proteins of two clones which were internally deleted from the 3rd-12th repeat. Since the NH$_2$-terminal 5-kDa domain has no epitopes for this antisera (above), the epitope on this fusion protein must be on one of, or across, the first two repeats of the mature protein.

Mapping of the NH$_2$ Terminus to the Protein Surface—While the bulk of the 5-kDa NH$_2$-terminal fragment is hydrophobic, other than several turns, the NH$_2$-terminal octamer has a strongly acidic charge (2 aspartic acids), and would be predicted to be at the surface of the protein. In support of this prediction, the native mature protein was immunoprecipitable with the NH$_2$-terminal antiserum, but not with control preimmune serum (Fig. 5).

Surface Accessibility of the Epitopes on Primary Translation Product—The accessibility of the NH$_2$-terminal octapeptide region and the surface epitope(s) of the tandem repeat(s) on the native primary translation product was also tested by immunoprecipitation. The cDNA encoding the full-length sequence was transcribed and translated in vitro, and the product was immunoprecipitable with either the NH$_2$-terminal antibody or the antibodies with epitopes located on the 27-kDa carboxyl fragment (not shown).

Northern Analysis—Poly(A) RNA from the venom gland of young adult female wasps and unstimmed hosts was probed with the clone encoding the full structural sequence. A single signal was observed corresponding only to venom gland mRNA of transcript size about 1 kb and no signal for samples from host tissues (not shown). This size compares favorably with the 894-base length of the cloned cDNA and the 1 kb size of the product of PCR amplification from mRNA. This result suggests that no other different sized transcripts in the venom gland have sequences closely related to that for the present Chelonus venom protein. Further, the absence of any signal in the lanes for poly(A) RNA from normal eggs or larvae strongly suggests that the venom protein is not closely
Unusual Signal Peptide—The encoded signal peptide itself is unusual. Usually, signal peptides have basic residues near each end which are necessary for anchoring in the cell membrane, and which are separated by an approximately 20-residue hydrophobic region, although on occasion the second basic residue may be missing or an acidic residue(s) placed on one end (36). In the signal peptide from the present venom protein, a glutamic acid residue occurs near the first basic amino acid, and the second basic amino acid has been replaced altogether by another glutamic acid. The protein processing (cleavage of signal peptide) and secretory pathways in hymenopteran venom glands may be highly specialized, as several of the few known signal peptides from venom gland proteins are also unusual. For example, the signal peptide for hornet antigen V has both basic residues replaced with glutamic acid residues (10). The site of predicted cleavage (von Heijne, 37) for secapin leaves a glutamic acid at position 29 of a 31-residue signal peptide (35). The signal peptide for honey bee melittin is the second basic residue (38, 39). It is interesting that in each of the known cases of occurrence of negatively charged residues in signal peptides of hymenopteran venom proteins, a glutamic acid residue is found.

Amino Acid Composition—The amino acid composition of the encoded sequence compares favorably with that found by direct amino acid analysis (13). The protein is very low in aromatic amino acids, and in fact has no tyrosine or tryptophan. It also has no histidine. Disulfide linkages are important to the structure of some small arthropod and vertebrate venoms (40). Interestingly, the 6 cysteines present are all in the hydrophobic domain that is apparently in the core of the protein.

Primary Structure and Evolution of the Protein—The protein has an intriguing primary structure which apparently encodes three regions or domains. The second and third domains have clearly arisen by repeated internal duplication of a core segment of approximately 14 residues. Sequence analysis also strongly suggests several subgroups of related repeats which may give clues as to the manner in which the pattern of repeats has developed. Intriguing, but unresolved at the present time, is the invariant retention of serine at position 5 of each repeat and the near invariant retention of glycine and serine at positions 6 and 7. Since none of the sequences of each subgroup of repeats is closely related to protein sequences obtained from GenBank, it is not possible to infer by analogy the function of this serine or of this repeat. It is of note that the homologous protein (adduced immunologically) in another species of this wasp genus has a size of approximately 35 kDa (15), and the venom of a species in a closely related genus has two homologous proteins (adduced immunologically) of approximately 43 and 45 kDa (14, 41). On the basis of the results presented here, it seems pertinent to determine whether these proteins will be found to be composed of a progressive increase in the number of tandem repeats.

Analysis of snake venom proteins has shown how one venom component (phospholipase A₂) can give rise to a related protein which has a function different than phospholipase A₂ (42). We propose that a similar principle is operating in the evolution of venom proteins in Hymenoptera. A venom protein prone to mutation through internal duplications and deletions will enable the parasite to keep pace with a host under selection pressure to evade the action of that protein. Further, as additional hosts and niches avail themselves as
potential substrates to the wasp over evolutionary time, highly
mutable venom proteins provide a mechanism by which forces 
of speciation can produce wasp lineages with venom com-
ponents functional against the given host or prey. 

Secondary and Tertiary Structure of the Protein—The 
inferences made about the secondary structure of each apparent
components functional against the given host or prey.

of speciation can produce wasp lineages with venom compo-
domains, together with immunological analyses and epitope
proteins (33), a reasonable conjecture is that much of the first
primarily of regions of hydrophobic sequence in extended
conformation which connect two 0-sheet structures, each of which
extended or a-helical, amphipathic conformation (36, 43).

The epitope-mapping experiments on the mature protein
confirm the location of surface epitopes on the region of the
tandem repeats and their absence in the 5-kDa NH2-terminal
portion, consistent with the proposed tertiary structure. The
placement of an epitope on the first or second repeat of the
delentum clone provides further support and is also consistent with
the Hopp and Woods method (34) predictions that repeat
1 is one of the most antigenic parts of the molecule. The very
NH2 terminus of the mature protein is exposed at the surface,
with the antibodies to it are capable of precipitating the
native protein. The presence of a signal peptide does not alter
the accessibility of the mature NH2 terminus to the antibodies,
nor does the presence of the signal peptide alter tertiary structure to
such an extent that the surface epitopes associated with the tandem repeats are subsequently lost. Future studies are clearly warranted to identify functional groups on the protein which interact with the target site or substrate.

Identity of the Protein—The Northern analysis shown here
detected no strong similarity of the encoded sequence for the
33-kDa protein to that for mRNA from normal host embryos
and larvae. Thus, the regulation of the host by the 33-kDa protein does not appear to be as injection of a protein resembling
a normal host protein.

The encoded sequence for the 33-kDa protein was also not
detectably related to any sequences in the genome of the
polydnavirus which replicates in the ovaries of the female wasp.2 Recently, in a study of a different family of parasitic wasps (Ichneumonidae), Webb and Summers (44) detected an immunological relationship between certain venom epi-
topes and epitopes on the polydnavirus nucleocapsid, and sequence similarity between certain venom and polydnavirus
mRNAs. In addition to the lack of hybridization between the
venom protein mRNA and the polydnavirus genome of the
Chelonus wasp (a braconid) which we have studied, we have not detected immunologically similar epitopes between the
33-kDa protein and the polydnavirus of the Chelonus wasp
(45).

Use of Venom Proteins as Experimental Probes—Arthropod
venoms have been useful tools in studies of various biochem-
ical questions in heterologous systems. Scorpion toxins have
been frequently tools of choice in purifying vertebrate (46) and
invertebrate (47) axonal sodium channels, while some
spider venoms can be used to block glutamate receptors or to
induce neurotransmitter release (2). Honey bee melittin has
been used to study cell membrane structure (6), and studies
on prepro melittin also contributed to early studies on protein
processing (48, 49). Already, crude preparations of paralyzing
wasp venoms have been used as probes to block the action of
arthropod venoms (50). With the identification and cloning of
this venom protein which participates in immunosuppres-
sion of the host, and with the features of this particular wasp-host
system, new areas of investigation are open on structure-
function relationships of one, and perhaps other, classes of
venom regulatory proteins that operate on embryonic tem-
plates. It is clear that wasp venom proteins offer exciting
leads on new mechanisms of protein secretion and of evolution
of natural toxins.

Acknowledgement—Appreciation is extended to Dr. Doug Melton,
Department of Biochemistry and Molecular Biology, Harvard Uni-
versity, in whose laboratory D. J. did the initial work while on sa-
batical leave.

REFERENCES
1. Ducancel, F., Guignery-Frelat, G., Bouchier, C., Menez, A., and Boulain,
Biochem. 171, 535
5. Boulain, J.-C., Menez, A., Coudere, J., Faure, G., Liaopoulos, P., and
258, 1781
J. Physiol. (Paris) 79, 254
Eur. J. Biochem. 48, 465
Press, London
Acad. Sci. U.S.A. 85, 895
J. Gen. Virol 69, 1027
Purification for Microsequencing (Matsudaira, P. T., ed) pp. 29-30, Aca-
demic Press, New York
85, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
21. LeNeur, M., Glaville, N., Mandel, J. L., Gerlinger, P., Palmier, R., and
U.S.A. 71, 3917
U.S.A. 74, 5463
Invert. Path. 47, 205
121, 193
37, 1

*J. Jones, G. Sawicki, and M. Wozniak, unpublished results.
Wasp Venom Protein with a Unique Repeating Structure