THE AMINO ACID SEQUENCE OF PHOSPHOLIPASE A

Dawie P. Botes and Cornelis C. Viljoen
From the National Chemical Research Laboratory, Council for Scientific and Industrial Research, Pretoria, Republic of South Africa

SUMMARY

Bitis gabonica phospholipase A contains 118 amino acid residues in a single chain and is cross-linked by six disulfide bridges. Cyanogen bromide cleavage, and tryptic and chymotryptic digests provided the necessary overlapping peptides allowing the derivation of the complete amino acid sequence. In its primary structure the enzyme is more closely related to porcine pancreatic phospholipase A than to the enzyme from the honey bee Apis mellifica.

EXPERIMENTAL PROCEDURE

Materials—Phospholipase A was purified from B. gabonica venom, and the reduced and S-carboxymethylated derivative prepared as described (13). Trypsin was obtained from Serva Laboratories, Cape Town, South Africa, as a twice crystallized, diphenylcarbamyl chloride-treated, salt-free preparation (Batch 336A). \( \alpha \)-Chymotrypsin (three times crystallized, Batch DC1 6100-1) was obtained from Worthington. Subtilisin, T 3081, was a product of Mann Research Laboratories. Phenylisothiocyanate and cyanogen bromide were supplied by Fluka. Trifluoroacetic acid was purchased from Merck. DEAE-cellulose and Sephadex were described in the preceding paper (13), and all other reagents were of analytical grade.

Digestion with Proteolytic Enzymes—Digestions with trypsin and chymotrypsin were carried out at 37\(^{\circ}\)C for 2 hours in 2% \( \text{NH}_4 \text{HCO}_3 \) and an enzyme to substrate ratio of 1:100 (w/w). Digestion of trypic peptide T6 with subtilisin was carried out at 25\(^{\circ}\) for 21 hours, using the same solvent and enzyme to substrate ratio.

Cleavage by Cyanogen Bromide—The procedure of Gross and Witkop (25) was applied under conditions similar to those used by Steers et al. (26). After 24 hours at 20\(^{\circ}\)C the reagents were removed under reduced pressure, the peptides dissolved in dilute ammonia and separated on Sephadex G-50.

Fractionation of Peptides—Tryptic and chymotryptic peptides were separated by chromatography on DEAE-cellulose.

Phospholipase A (EC 3.1.1.4), the heat-stable esterolytic enzyme acting on 3-sn-phosphoglycerides, liberating specifically the fatty acid esterified at the glycerol C-2 position, has been successfully purified from pancreatic tissues (1), snake venoms (2-13), and also bee venoms (14, 15). Whereas the pancreatic enzymes from pig (16), human (17), and rat (18) have been found to occur as an enzymatically inactive zymogen, no evidence of a precursor has been found in either snake or bee venom. During the activation of the porcine zymogen by trypsin, a heptapeptide which constitutes the NH\(_2\)-terminal sequence of the porcine enzyme and showed it to consist of 123 amino acid residues in a single polypeptide chain, cross-linked by six disulfide bridges.

The phospholipase A from bee venom was first distinguished from the hemolytic peptide melittin by Neumann and Habermann (21). Habermann and Reiz (22) described a purification procedure for the enzyme which was improved on by Shipolini et al. (15). Shipolini and co-workers (23) published the primary structure of the bee venom enzyme, a 129-amino acid residue single polypeptide chain containing covalently bound carbohydrates, and four disulfide bridges.

In spite of the large number of snake venom phospholipases A that have been obtained in pure form, primary structural data on enzymes derived from these venoms are notably scanty. Sanejima et al. (24) reported the amino acid composition of a phospholipase A from Agkistrodon halys blomhoffii. These authors showed the enzyme to consist of 126 amino acid residues with seven disulfide bonds, but positioned only the first 17 amino acid residues from the NH\(_2\)-terminus and showed half-cystine to be COOH-terminal.

In a separate communication the purification of a phospholipase A from Bitis gabonica venom is described (13). The enzyme was shown to consist of 118 amino acid residues in a single polypeptide chain with six disulfide bridges. The present communication describes the complete amino acid sequence of the gaboon adder enzyme.

Some of the data are presented as a miniprint supplement immediately following this paper. Material published in miniprint form can be easily read with the aid of a large-field reading glass of a type readily available at most opticians. For the convenience of those who prefer to obtain supplementary material in the form of a microfiche or full size photocopy, these same data are available as JBC Document No. 73M-1392. Orders for supplementary material should specify the title, authors, and reference to this paper and the JBC Document Number, the form desired, microfiche or full sized photocopy of 33 pages, and the number of copies desired. Orders should be addressed to The Journal of Biological Chemistry, 9600 Rockville Pike, Bethesda, Maryland 20814, and must be accompanied by a remittance to the order of the Journal in the amount of $2.50 for microfiche or $3.30 for photocopy.
gradient from 0.02 to 0.6 M NH₄HCO₃ was applied in both cases. The eluate was continuously monitored at 280 nm.

**Paper Chromatography and Electrophoresis of Peptides**—The solvent systems for paper chromatography were Solvent I, butan-1-ol-acetic acid-water (40:6:16, v/v) and Solvent II, butan-1-ol-pyridine-acetic acid-water (15:10:3:12, v/v). Electrophoresis was performed on a Gilson high voltage electrophorator, model D, with Varsol as cooling medium. The buffer systems were Buffer A, pyridine-acetic acid-water, pH 4.5 (10:14:576, v/v), and Buffer B, acetic acid-formic acid-water, pH 1.9 (100:2:95:576, v/v). The electrode vessels contained the above buffers, while the papers used for the electrophoretic separation were moistened with 1:10 and 1:5 (v/v) dilutions, respectively, of the above buffers. Whatman No. 3MM paper was used in all cases. Peptides were located on paper with the collidine-ninhydrin reagent (29). Tryptophan-containing peptides were revealed by the Ehrlich reagent (50).

**Amino Acid Composition and Sequence Analyses of Peptides**—Peptides were hydrolyzed and amino acid analyses performed as described in the preceding paper (13). The NH₂-terminal sequence of the RSCM-phospholipase A was determined by Edman degradation (31) with a Beckman Sequencer, model 890B, using the Quadrol protein program, or in the case of the larger peptides, the volatile buffer peptide program. 3-PHENYL-2-thiohydantoins were identified by gas chromatography with SP-400 support (32) and by thin layer chromatography (33-35). The ninhydrin reagent of the cGANOCN bromide peptides, together with that of phospholipase A, was used for the electrophoretic separation were moistened with 1:10 and 1:5 (v/v) dilutions, respectively, of the above buffers. Whatman No. 3MM paper was used in all cases. Peptides were located on paper with the collidine-ninhydrin reagent (29). Tryptophan-containing peptides were revealed by the Ehrlich reagent (50).

**RESULTS**

**Isolation and Amino Acid Composition of CNBr Peptides**—The separation of the cyanogen bromide peptides by gel filtration through Sephadex G-50 is presented in Fig. 1. Fractions I and II constituted aggregated material and a mixture of undegraded and partially degraded material, respectively, and no further purification of these fractions was attempted. Peptides CN3 and CN4 required no further purification. CN1 and CN2 eluted in one fraction and were separated by paper electrophoresis at pH 1.9. Table I reports the amino acid compositions of the cyanogen bromide peptides, together with that of phospholipase A.

**Sequences of RSCM-Phospholipase A and Peptides**—The application of the Edman procedure, either through the use of the automatic sequencer or by manual manipulation, and direct identification of the released derivitized amino acid, provided sufficient data to arrive at a complete linear structure of the phospholipase A. Both the amino acid compositions of tryptic and chymotryptic peptides and the experimental data that led to the complete sequence are detailed in the supplement to this communication.

**Alignment of Peptides and Derivation of Complete Sequence**—Sequence studies on RSCM-phospholipase A, together with the amino acid compositions of peptides CN1 and CN2, allowed the latter two cyanogen bromide peptides to be positioned NH₂-terminal, in that order. Peptide CN4 contained no homoserine or the lactone and was therefore placed COOH-terminal, allowing CN3 to be positioned by difference. Except for a small segment toward the COOH-terminal end of this peptide, the sequence of CN3 could be derived from sequence studies on the intact peptide. The sequence of this particular segment was clarified by Edman degradation on T4 and C7a, peptide derived from a tryptic and chymotryptic digest, respectively, of RSCM-phospholipase A.

Sequence analyses on CN4 as well as on secondary peptides derived from CN4 allowed the sequence of this peptide to be deduced. Due to the complexity of both the tryptic and chymotryptic digests, not all the peptides derived from the latter two digests were identified. Nevertheless, peptides providing the necessary overlaps and small enough to confirm, by their amino acid compositions, sequence analyses performed on the larger fragments, were obtained from the tryptic and chymotryptic digests.

**DISCUSSION**

The complete amino acid sequence of a phospholipase A from the venom of the gaboon adder, *B. gabonica*, is reported in Fig. 2. The nonenzymatic cleavage of the protein by cyanogen bromide at the 3 methionine residues yielded the expected four peptides. Although the close proximity of 2 of the 3 methionine residues to the NH₂-terminus limited the information to that part of the molecule, the position of the third methionyl was more favorable and gave rise to the required large fragments. Tryptic and chymotryptic cleavage produced peptides amenable to sequential degradation and also, together with the cyanogen bromide fragments, provided overlaps allowing the complete sequence to be deduced.

Residues 67 to 74 proved to be particularly refractory to conventional sequential degradation due to the particularly strong tendency of glutamyl residue 67 to convert to pyroglutamic acid, when it became NH₂-terminal, rendering the peptide inaccessible to further cleavage. This difficulty, however, could be bridged by limiting the acid cleavage, during the preceding cycle, to 1 min. Although this resulted in incomplete cleavage of the isoleucyl residue at position 66, its minimized the exposure of the newly formed NH₂-terminal glutamyl residue to anhydrous acid and hence also reduced conversion to pyroglutamic acid. Acid cleavage during the subsequent step released not only glutamine but also the residual isoleucine at position 66. A somewhat prolonged anhydrous acid exposure, however, ensured complete cyclization of the glutamyl residue, NH₂-terminal after removal of the residual isoleucine, leaving a peptide with a free NH₂-terminal, starting at position 68. This allowed sequential degradation without the complications of part of the peptide lagging 1 residue behind in the Edman procedure.
Aspartic acid, rather than asparagine, was identified at residue 68. However, since a glycy1 residue is found in position 69, the possibility of the aspartyl residue at position 68 having arisen from deamidation of an asparaginyl, had to be considered. The very ready deamidation of asparaginyl residues whenever in the sequence asparaginyl-glycyl is now well recognized (38-41). Fortuitously, the sequence Asn-Gly is readily distinguished from Asp-Gly by the sensitivity of the former and stability of the latter toward hydroxylamine. Bornstein (42) attributed this selectivity to the ability of hydroxylamine to cleave the cyclic imides formed from asparaginyl-glycyl bonds. Support for the suggestion that asparaginyl-glycyl bonds are cleaved in preference to aspartyl-glycyl bonds comes from the finding that asparaginyl-glycyl at residues 24 to 25 and 90 to 91 in rat skin collagen is not of a phospholipase A from a mammalian source has been published by De Haas and co-workers (20). The latter authors noted that, for the porcine pancreatic enzyme, all 6 glycine residues present are found in the NH2-terminal region, between positions 22 and 42. Also, the basic residues showed a tendency to cluster toward the COOH-terminal, where 5 out of the 9 lysine residues are located between positions 114 and 128. In accordance with the idea that the highly charged COOH-terminal region of the protein chain will probably be located at the surface of the molecule, trypsin in low concentrations preferentially released a pentapeptide between residues 116 and 120, leaving the rest of the molecule intact.

The second phospholipase A to be sequenced was that from the common European honey bee Apis mellifera (23). In a communication describing the purification and some properties of the bee venom enzyme, Shipolini et al. (23) accounted for only four disulfide bonds in their proposed sequence. No overlap had been established between residues 76 and 77, and although the sequence is reported with some confidence, these authors submitted the formal possibility of an incomplete sequence, in that a fragment lying between residues 76 and 77 might have escaped detection.

The present communication is the first report of a complete sequence of a phospholipase A derived from a snake venom. Homology between the snake venom and pancreatic enzymes is far more pronounced than that between any of these and the bee venom enzyme. Proper alignment of the sequences, such as to make it phylogenetically more meaningful, will have to await the completion of more structures, both mammalian and reptilian. At this stage, however, it is possible to point out some of the more outstanding similarities between the different structures (Fig. 3).

Considering first of all the mammalian and snake venom enzymes, similarity from the NH2 terminus to residue 9, and again from residue 16 to 41 is quite pertinent. The high frequency of occurrence of glycyl residues in this region of the molecule is a common feature of the two enzymes. The region between

---

**Table I**

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Phospholipase A</th>
<th>CN1</th>
<th>CN2</th>
<th>CN3</th>
<th>CN4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td>7.83 (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histidine</td>
<td>1.96 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arginine</td>
<td>3.07 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carboxymethylcysteine</td>
<td>11.47 (12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>15.60 (13)</td>
<td>2.07 (2)</td>
<td>1.12 (1)</td>
<td>4.14 (4)</td>
<td>11.02 (12)</td>
</tr>
<tr>
<td>Threonine</td>
<td>7.68 (8)</td>
<td>0.92 (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serine</td>
<td>6.00 (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>10.15 (10)</td>
<td>1.12 (1)</td>
<td></td>
<td></td>
<td>8.31 (8)</td>
</tr>
<tr>
<td>Proline</td>
<td>1.90 (2)</td>
<td></td>
<td></td>
<td>1.12 (1)</td>
<td>1.09 (1)</td>
</tr>
<tr>
<td>Glycine</td>
<td>12.77 (13)</td>
<td>1.18 (1)</td>
<td></td>
<td></td>
<td>5.07 (5)</td>
</tr>
<tr>
<td>Alanine</td>
<td>4.10 (4)</td>
<td></td>
<td></td>
<td>1.01 (1)</td>
<td>3.02 (3)</td>
</tr>
<tr>
<td>Valine</td>
<td>2.97 (3)</td>
<td></td>
<td>1.73 (2)</td>
<td></td>
<td>0.74 (1)</td>
</tr>
<tr>
<td>Methionine</td>
<td>2.47 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homoserine</td>
<td>0.84 (1)</td>
<td>0.88 (1)</td>
<td></td>
<td>0.91 (1)</td>
<td></td>
</tr>
<tr>
<td>Isoleucine</td>
<td>5.74 (6)</td>
<td></td>
<td></td>
<td>1.72 (2)</td>
<td>2.09 (3)</td>
</tr>
<tr>
<td>Leucine</td>
<td>1.95 (2)</td>
<td>0.96 (1)</td>
<td></td>
<td></td>
<td>1.03 (1)</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>5.82 (10)</td>
<td></td>
<td></td>
<td>4.74 (5)</td>
<td>4.79 (5)</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>5.02 (5)</td>
<td>0.92 (1)</td>
<td></td>
<td>1.93 (2)</td>
<td>1.93 (2)</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>2.18 (2)</td>
<td></td>
<td>(1)</td>
<td>(1)</td>
<td>(1)</td>
</tr>
<tr>
<td>Total</td>
<td>118</td>
<td>8</td>
<td>4</td>
<td>40</td>
<td>66</td>
</tr>
</tbody>
</table>

* Extrapolated to zero hydrolysis time.
* 72-hour hydrolysis.
* Spectrophotometrically determined.
* Ehrlich-positive.
* Values in the CN-peptides were determined by summation of homoserine and its lactone.
FIG. 2. Amino acid sequence of phospholipase A.

positions 10 and 15 is not only highly variant, but also requires insertions in the pancreatic sequences or deletions in the snake venom structure. In this region of the mammalian enzyme is present the first cysteinyl residue. Noteworthy is the fact that the next 3 cysteinyl residues in the linear structures of both enzymes are completely in alignment. The partial sequence of the phospholipase A from Agkistrodon halys blomhoffi (24), in common with the B. gabonica enzyme, is void of half-cystine residues up to position 17.

Since the mammalian enzyme is derived from a precursor by enzymic release of a heptapeptide, the significance of the proximity of a disulfide cross-link to the NH₂-terminal end might be in stabilizing the zymogen or in restricting the activating cleavage to the first 7 residues.

Apart from noncovalent stabilizing effects such as hydrogen and hydrophobic bonding, the disulfide bridge is a major factor in maintaining and stabilizing secondary and tertiary structures. One would therefore expect the cysteinyl residues to be conserved. However, whenever folding of the chain, as determined by the primary structure, no longer necessitates a disulfide bridge at a particular region of the molecule, the pressures for maintaining invariancy of that particular cysteinyl residue would also be diminished. Provided that the basic structural requirements for biological activity are maintained, cross-linking to a higher, or lesser degree, should be tolerable. This seems to apply for phospholipases A from different biological sources.

In the snake venom primary structure the sequence cysteinyl-cysteinyl is found at residues 41 to 42 and again at residues 47 to 48. A similar situation is found for the bee enzyme, the first cysteinyl, however, being replaced by a methionyl residue in the bee enzyme. Nonetheless, this region in the bee enzyme structure shows a considerably higher degree of homology with the snake venom enzyme than any part of the rest of the molecule, where homology is by no means obvious. A similar structural feature is absent from the pancreatic enzyme. However, a cysteinyl residue corresponding to either residue 47 or 48 in the snake venom phospholipase A is present in the pancreatic enzyme. A clear-cut choice between aligning this particular cysteinyl with either residue 47 or 48 in the snake venom phospholipase A is present in the pancreatic enzyme. A clear-cut choice between aligning this particular cysteinyl with either residue 47 or 48 in the snake venom enzyme will have to await the positioning of the disulfide links in the snake enzyme. Although it is tempting to align the Cys-Tyr sequences in the corresponding parts of the two enzymes, a Cys-Cys, Cys-Tyr alignment seems equally feasible at this stage, since a Cys-Tyr substitution requires only a single base substitution in the coding triplet and would furthermore allow an identical number of residues in the linear structures between this cysteinyl residue and the preceding one.

Homology is again evident in the region 89 to 93 where both the porcine and snake enzymes have a Ala-Ala-Ile-Cys-Phe sequence. The fact that both enzymes terminate in a cysteinyl residue seems purely coincidental. The sequence data, although fragmentary in the sense that homology is inferred more from a constancy in the number of residues between cysteinyl residues in this region of the molecule rather than invariance in the chem-
Table and diagram: Alignment of phospholipases A from *Bitis gabonica* venom, porcine pancreas, and *Apis mellifica* venom.

**Fig. 3.**
ical sense, suggested homology between Cys 112 in the snake enzyme and the COOH terminal cysteinyl in the porcine enzyme. Some support for this alignment is the absence of a highly positive region toward the COOH-terminal end of the snake enzyme, a noticeable feature of the pancreatic phospholipase.

From an inspection of the three linear sequences, the phospholipase A structure can tolerate a considerable degree of variancy and still retain its biological activity. The primary structures provide no hint as to the reason for one of the most interesting properties of this water-soluble enzyme, i.e. its high activity toward substrates which give aggregated structures in water. Whether the uneven distribution of certain residues over the chain is significantly reflected in the three-dimensional pattern, or whether apolar regions, distal in the linear structure, form a particularly hydrophobic pocket to accommodate the hydrophobic substrate in the active conformation of the enzyme can not be judged from the sequences alone.

Acknowledgments—The authors wish to thank Dr. T. Haylett and Mrs. A. Ruele for the gas chromatographic analyses, and Miss C. M. Potgieter for quantitative amino acid analyses.

REFERENCES
The sequence of CN-3 was established up to residue 43, and this part of the molecule was resolved by sequencing peptides T-2, P-3, and T-4.

No sequence studies were performed on this peptide.

Sequence studies performed on this peptide.

No sequence studies performed.

Sequence studies performed on this peptide.
Dawie P. Botes and Cornelis C. Viljoen

*Bitis gabonica* Venom: THE AMINO ACID SEQUENCE OF PHOSPHOLIPASE A


Access the most updated version of this article at http://www.jbc.org/content/249/12/3827

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/249/12/3827.full.html#ref-list-1