Pure & Appl. Chem., Vol. 66, No. 4, pp. 715–720.
Printed in Great Britain.

Darwinian evolution of *Trimeresurus flavoviridis* venom gland phospholipase A_2 isozymes

Kin-ichi Nakashima*, Tomohisa Ogawa*, Naoko Oda*, Yasuyuki Shimohigashi*, Masahira Hattori†, Yoshiyuki Sakaki†, Hiroshi Kihara‡, and Motonori Ohno*

*Department of Chemistry, Faculty of Science, Kyushu University, Fukuoka 812, Japan; †Institute of Medical Science, University of Tokyo, Tokyo 108, Japan; ‡Biotechnology Research Laboratories, Takara Shuzo Co. Ltd., Shiga 520-21, Japan

Abstract: As a step toward understanding the structure and function relationships of phospholipases A_2 (PLA_2s), we isolated and sequenced five cDNAs encoding *Trimeresurus flavoviridis* (Habu snake) venom PLA_2 isozymes. Comparison of the nucleotide sequences of these cDNAs revealed that the homologies of the 5' and 3' untranslated regions were much higher than that of the protein-coding regions and that the base substitution rates at the first, second, and third position of codons are similar in the coding region. To gain further insight into this novel findings, six *T. flavoviridis* venom gland PLA_2 isozyme genes were isolated and sequenced. They were found to consist of four exons and three introns. The numbers of nucleotide substitutions per site (K_N) for introns are approximately one-fourth of the numbers of nucleotide substituions per synonymous site (K_S) for the protein-coding regions, indicating that the introns are unusually conserved. The fact that the numbers of nucleotide substitutions per nonsynonymous site (K_A) are close to or larger than K_S values for relevant pairs of genes revealed that Darwinian-type accelerated substitutions have occurred in the protein-coding regions of exons. This is compatible with the presence of PLA_2 species with diverse physiological activities in the venom.

Phospholipase A_2 (PLA_2, EC 3.1.1.4) catalyzes the hydrolysis of the 2-acyl ester linkage of 3-*sn*-phosphoglycerides with the requirement of Ca^{2+}. To date, the primary structures of more than 90 PLA_2s have been determined (1). Aspartate-49 (numbered accoding to the aligned numbering of PLA_2 enzymes from various sources) constitutes a part of the Ca^{2+} binding site (2). The PLA_2s are classified into two groups, Asp-49-PLA_2 with high activity and Lys-49-PLA_2 with extremely low activity (3). We isolated and sequenced Asp-49-PLA_2 (4, 5) and two Lys-49-PLA_2s called basic proteins I and II (6, 7) from *Trimeresurus flavoviridis* (Habu snake, crotalidae) venom. These PLA_2s consist of 122 amino acid residues and are structurally homologous to one another. Especially, basic protein I is identical to basic protein II except that aspartate at position 58 of the former is replaced by asparagine in the latter. These *T. flavoviridis* PLA_2 isozymes provide a useful system for understanding the structure-function relationships of PLA_2s. In addition, a question arises as to why PLA_2 isozymes with low lipolitic activity were produced in the venom. Here we report the cloning and the sequence analysis of cDNAs and genomic DNAs encoding venom gland *T. flavoviridis* PLA_2 isozymes. Comparison and analysis of the nucleotide sequences of five cDNAs and six genes indicated that *T. flavoviridis* PLA_2 isozyme genes have evolved so as to bring about accelerated amino acid substitutions in the protein-coding regions except for the signal peptide-coding domains. Such accelerated substitutions appear to be adaptive and consistent with the fact that *T. flavoviridis* venom contains PLA_2 species with diverse physiological activities.

Novel Structures of cDNAs Encoding *T. flavoviridis* PLA_2 Isozymes.

From *T. flavoviridis* venom gland cDNA library, five cDNAs encoding Asp-49-PLA_2, basic protein I and II, Thr-37-PLA_2, and PLX'-PLA_2 were isolated and sequenced. The nucleotide substitutions among five PLA_2 isozyme cDNAs are schematically shown in Fig. 1. The 5' and 3' untranslated regions (UTRs)

are much more homologous than the coding regions; nucleotide identities are 98% (59/60) for the 5' UTRs, 67% (281/414) for the coding regions, and 89% (97/109) for the 3' UTRs.. Such greater sequence homology in the UTRs than in the coding regions has not been known in cDNAs for other isozyme systems so far reported. Usually, less homology has been observed in the UTRs as in various isoforms of the G-protein α-subunit family (8) and the protein kinase C family (9). From the nucleotide sequences of cDNAs, it is noted that the signal sequences consist of 16 amino acid residues and that the domains coding for these sequences are exceptionally highly conserved when compared with other portions of the coding regions. The signal sequences fulfill the structures generally required for the signal peptides, that is, the presence of methionine and a positively charged residue at 1 and 2 positions followed by a stretch of hydrophobic amino acid residues. The high conservation may suggest that the signal peptides share with common roles in membrane penetration and susceptibility to a common signal peptidase. Examination of the nucleotide sequences of the coding regions of cDNAs revealed that substitutions have occurred in the first, second, and third positions of the triplet codons at similar rates; 45/140 (32.1%), 42/140 (30.0%), 40/140 (28.6%), respectively. This seems unusual compared with the other cases in which the silent (third) sites of codons are more variable than the first and second positions, where changes cause amino acid changes (10).

The unique features of cDNAs for *T. flavoviridis* PLA_2 isozymes prompted us to analyze the organization and structures of the PLA_2 family in the genome.

Structures of DNAs Encoding *T. flavoviridis* PLA_2 Isozymes.

Genomic Southern blot analysis of *T. flavoviridis* liver and venom gland DNAs digested by restriction enzymes gave several bands for each digest when the *Acc* I-*Pst* I fragment (the coding region) of *T. flavoviridis* Asp-49-PLA_2 cDNA was used as hybridization probe. Both venom gland and liver DNAs gave the identical profile. These results suggested that PLA_2 isozyme genes form a multigene family.

Genomic DNA library constructed with liver DNAs was screened with nearly full length *T. flavovirids* Asp-49-PLA_2 cDNA which contains the 5' and 3' UTRs highly conserved for *T. flavoviridis* PLA_2 isozyme cDNAs. One hundred and two positive clones were isolated from approximately 2 x 10^6 plaques. Eleven clones seleced on the basis of signal intensity were digested with *Eco* RI and blotted with the *Acc* I-

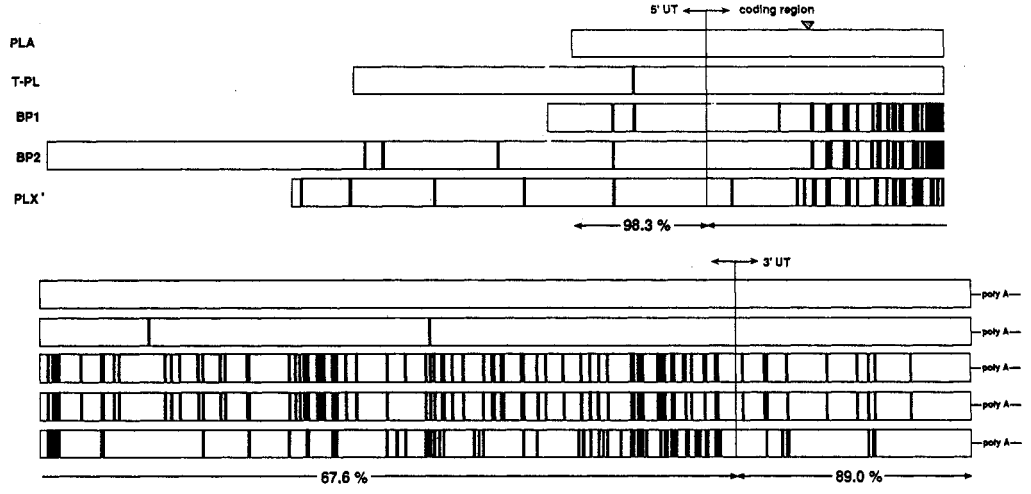

FIG. 1. Schematic representation of base substitutions among five *T. flavoviridis* PLA_2 isozyme cDNAs. Dark lines indicate the substituted bases when Asp-49-PLA_2 (PLA) cDNA is taken as a reference. T-PL, cDNA for Asp-49-PLA_2 with threonine at position 37; BP1, cDNA for basic protein I; BP2, cDNA for basic protein II; PLX', cDNA for an isozyme of Asp-49-PLA_2.

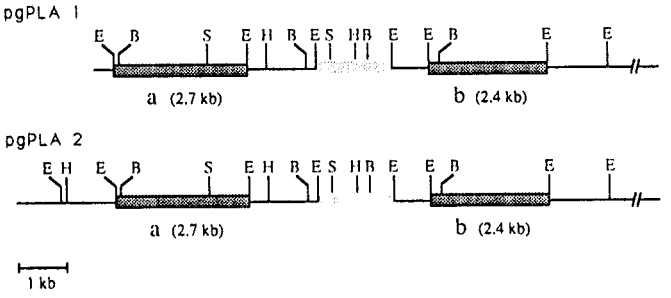

FIG. 2. Structures of phage clones pgPLA 1 and pgPLA 2 both encoding *T. flavoviridis* Asp-49-PLA_2 isozymes. The sites for restriction enzymes are designated as follows: B, *Bam*H I; E, *Eco*R I; H, *Hind* III; S, *Sac* I. *Eco*R I sites have not been detected in the dotted boxes. The fragments that hybridized with the *Acc* I-*Pst* I fragment (the coding region) of Asp-49-PLA_2 cDNA are shown by shaded boxes.

Pst I fragment (the coding region) of Asp-49-PLA_2 cDNA after electrophoresis. The two clones named pgPLA 1 and pgPLA2, which were 14 and 15.6 kb long, respectively, gave both 2.7 and 2.4 kb fragments which hybridized with this probe. The restriction maps of pgPLA 1 and pgPLA 2 are shown in Fig. 2. The four fragments that hybridized with the *Acc* I-*Pst* I fragment (the coding region) of Asp-49-PLA_2 cDNA were designated pgPLA 1a (2.7 kb), pgPLA 1b (2.4 kb), pgPLA 2a (2.7 b), and pgPLA 2b (2.4 kb). The first two and last two fragments are arranged in tandem, respectively, as shown in Fig. 2. Two out of the remaining 100 clones hybridized with the *EcoR* I-*Pst* I fragment (the coding region) of Lys-49-PLA_2 cDNAs but not with the *Acc* I-*Pst* I fragment (the coding region) of Asp-49-PLA_2 cDNA. One of them was found to encode basic protein I and named BP-I. The other encoded basic protein II and named BP-II. The nucleotide sequences of pgPLA 1a, pgPLA 1b, pgPLA 2a, pgPLA 2b, BP-I, and BP-II were determined by the dideoxy chain-termination method after subcloning into pUC13 or pUC19. The protein-coding regions were determined by matching their sequences with those of Asp-49-PLA_2 cDNA and Lys-49-PLA_2 cDNAs.

The transcription initiation sites were determined for pgPLA 1a and pgPLA 2a by primer extension

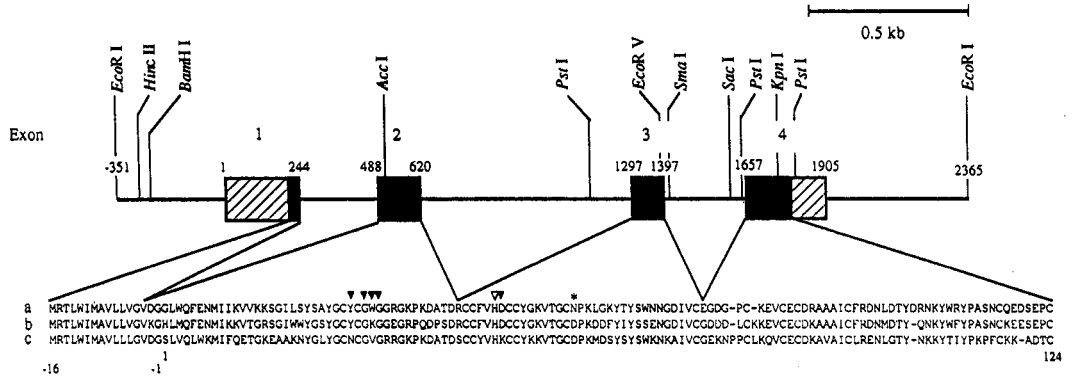

FIG. 3. Schematic representation of a common structure of six *T. flavoviridis* PLA_2 isozyme genes, pgPLA 1a, pgPLA 2a, pgPLA 1b, pgPLA 2b, BP-I, and BP-II, and the amino acid sequences corresponding to the coding regions of pgPLA 2a (a), pgPLA 1b (b), and BP-I (c). Four exons are indicated by boxes and the UTRs hatched. The numbers of the nucleotide positions represent those for pgPLA 1a and pgPLA 2a. Four coding regions and their corresponding amino acid sequences are correlated by lines. The residues involved in the Ca^{2+}-binding site are marked by ▼, those in catalytic site by ▽, and the residue at position 58 by *.

analysis using total venom gland RNAs as template (data not shown). The sites were identical for both genes and assigned to the adenosine residue located at 204 nucleotides upstream from the translation initiation codon (ATG), and the TATA-like sequence (CATAAA) was found 34 nucleotides upstream from the transcription initiation site. All genes spanned about 1.9 kb, contained four exons and three introns, and encoded proteins of 138 amino acid residues, including the signal sequence of 16 amino acid residues (Fig. 3). The nucleotide sequence for combined exons of pgPLA 2a was in accord with that of *T. flavoviridis* Asp-49-PLA$_2$ cDNA although the latter is 144 nucleotides shorter in the 5' UTR. Five base substitutions were noted between pgPLA 1a and pgPLA 2a but they carry the same mature protein. Genes pgPLA 1b and pgPLA 2b with five substitutions encoded an isozyme of Asp-49-PLA$_2$ (Fig. 3).

Analysis of Nucleotide Substitutions in PLA$_2$ Isozyme Genes Indicates Accelerated Evolution.

The nucleotide sequences and the corresponding amino acid sequences for the first introns and the second exons of six PLA$_2$ isozyme genes are shown in Fig. 4 as a representative. It is easily recognized that the second exon is much more variable than the first intron. This is also true of other introns and the protein-coding regions of other exons except for the signal sequence-coding domain of the first exon. For example, sectional homologies between pgPLA 1a and BP-I are 98.5% for the 5' UTR and 97.5% for the signal peptide-coding domain (first exon), 93.8% (first intron). 67.7% (second exon), 92.8% (second intron), 82.1% (third exon), 96.9% (third intron), and 73.2% for the protein-coding region and 91.5% for the 3'UTR (fourth exon). It is generally known that the evolution rates of introns are much greater than those

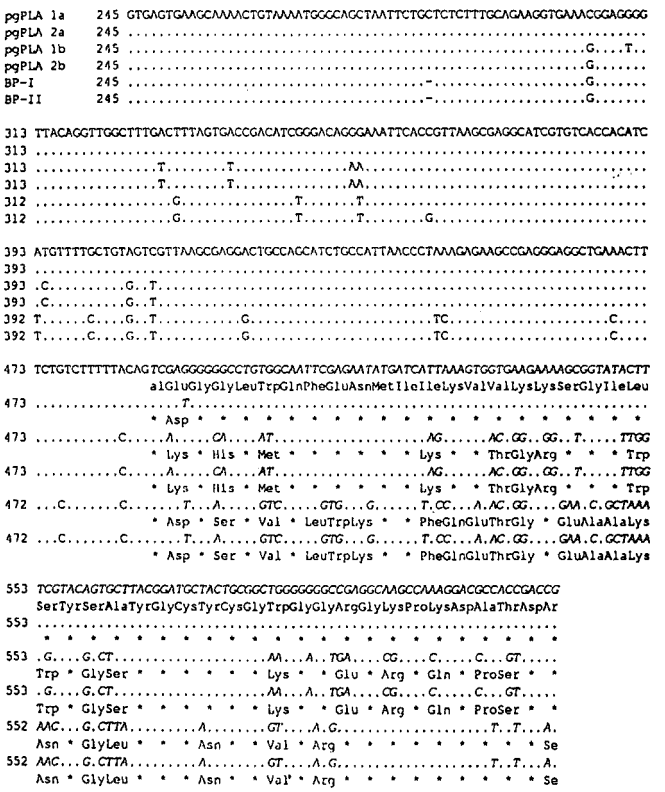

FIG. 4. The nucleotide sequences of the 1st introns and the 2nd exons of *T. flavoviridis* PLA$_2$ isozyme genes, pgPLA 1a, pgPLA 2a, pgPLA 1b, pgPLA 2b, BP-I, and BP-II.

of the protein-coding regions (or exons) (10), so that the structures of *T. flavoviridis* PLA_2 isozyme genes seem to be anomalous.

The evolutionary significance involved in the nucleotide sequences of *T. flavoviridis* PLA_2 isozyme genes was analyzed by computing K_N, K_S, and K_A values for all the relevant pairs of PLA_2 isozyme genes. Table 1 shows the values for three pairs of the genes. The data show several characteristics. First, similarly as in the 5' and 3' noncoding regions, K_N values for introns are approximately one-fourth of K_S values for all the pairs of the genes compared. This indicates that the introns are unusually conserved as compared to the protein-coding regions. The high homology of introns may suggest that there are some functionally important constraints in the introns. However, the regions corresponding to introns of precursor RNAs involved no significant secondary structure when analyzed by the method of Zuker and Steigler (11) (data not shown). Thus, it may be reasonable to consider that the introns have no significant functional role. The absence of apparent role in introns suggested that the protein-coding regions of exons have evolved at greater substitution rates than introns. Secondly, K_A/K_S values are much greater than those reported for other isoprotein genes (10, 12). Although synonymous sites are known to be much more variable than nonsynonymous sites because of much less functional constraint in the former (10, 12), this is not the case in the protein-coding regions of *T. flavoviridis* PLA_2 isozyme genes. The K_A/K_S values of the coding regions for pairs of pgPLA 1a and pgPLA 1b or pgPLA 2b and of pgPLA 1a and BP-I or BP-II are close to 1, and those for pairs of pgPLA 1b and BP-I or BP-II are about 1.8. Such high degrees of substitutions in nonsynonymous sites suggest that *T. flavoviridis* PLA_2 isozyme genes have evolved so as to bring about accelerated amino acid substitutions.

Diverse Physiological Activities of *T. flavoviridis* PLA2 Isozymes.

In terms of lipolytic activity, basic proteins I and II (Lys-49-PLA_2s), the products derived from BP-I and BP-II genes, respectively, are only 1.5-1.7% as active as Asp-49-PLA_2 (6, 7). Low lipolytic activity of basic proteins I and II is thought to be due in a large part to the presence of lysine in place of the Asp-49 of Asp-49-PLA_2. The Asp-49 is located in the Ca^{2+} binding site of Asp-49-PLA_2 (2). The Lys-49 in basic proteins I and II is unable to chelate Ca^{2+} by itself. However, necrosis-inducing activity of basic proteins I and II were 2 times more active than Asp-49-PLA_2 (13). This was ascribed to greater depolarization effects

TABLE 1. K_N/K_S and K_A/K_S values between pairs of *T. flavoviridis* PLA_2 isozyme genes.

	pgPLA 1a *vs.* pgPLA 1b			pgPLA 1a vs. BP-I			pgPLA 1b *vs.* BP-I		
	K_N	K_S	K_N/K_S	K_N	K_S	K_N/K_S	K_N	K_S	K_N/K_S
5'-Flanking region	0.051	0.224	0.228	0.035	0.253	0.138	0.029	0.183	0.158
5'-UTR	0.020	0.224	0.089	0.015	0.253	0.059	0.025	0.183	0.137
Intron 1	0.042	0.224	0.186	0.064	0.253	0.253	0.069	0.183	0.377
Intron 2	0.084	0.224	0.375	0.076	0.253	0.300	0.058	0.183	0.317
Intron 3	0.032	0.224	0.143	0.032	0.253	0.126	0.032	0.183	0.175
All introns	0.063	0.224	0.281	0.063	0.253	0.249	0.054	0.183	0.295
3'-UTR	0.059	0.224	0.263	0.059	0.253	0.233	0.065	0.183	0.355
3'-Flanking region	0.060	0.224	0.268	0.023	0.253	0.091	0.060	0.183	0.328
All noncoding regions	0.057	0.224	0.254	0.055	0.253	0.217	0.050	0.183	0.273
	K_A	K_S	K_A/K_S	K_A	K_S	K_A/K_S	K_A	K_S	K_A/K_S
Coding region of exon 1	0.000	0.116	0.000	0.039	0.000	-------	0.039	0.113	0.341
Exon 2	0.296	0.452	0.654	0.441	0.379	1.16	0.525	0.286	1.84
Exon 3	0.166	0.081	2.05	0.218	0.203	1.07	0.251	0.136	1.84
Coding region of exon 4	0.101	0.174	0.579	0.279	0.316	0.855	0.279	0.166	1.68
All coding regions	0.168	0.224	0.752	0.296	0.253	1.17	0.321	0.183	1.75

The K_S values used for estimation of and K_N/K_S for the noncoding regions are those for the complete coding resions of the corresponding pairs.

for muscle cell membranes of basic proteins I and II than Asp-49-PLA_2 (13). In guinea pig ileum contraction assay, Asp-49-PLA_2 and basic protein II elicited strong contraction at doses 10 to 100 times less than that for basic protein I (14). Ability of proteins to contract ileum appears to depend on whether the residue at position 58 is charged or noncharged (15). Asp-49-PLA_2 and basic protein II have asparagine while basic protein I has aspartate at this position. It is assumed that the residue at position 58 is involved in the site that specifically interacts with organized phospholipid matrices. These observations suggest that *T. flavoviridis* PLA_2 isozymes can manifest diverse physiological activities. It is likely that the genes encoding basic proteins I and II were produced from the gene encoding Asp-49-PLA_2 through base substitutions including converting the Asp-49 codon to a lysine codon. Such substitutions, otherwise, could afford potentiality for generating proteins with new physiological activity. The presence of PLA_2 isozymes with diverse physiological activities in venom must have a strong selective advantage to disrupt the physiological integrity of animals for catching prey or for defense against predators. Thus, it could be inferred that the nucleotide substitutions in *T. flavoviridis* PLA_2 genes to acquire PLA_2 species effective for such purposes have occurred by positive Darwinian selection.

The venom PLA_2s described here are regarded as proteins which interact functionally with diverse foreign substances. It may be assumed that such proteins each manifesting particular biochemical and physiological activity tend to evolve under adaptive pressure.

ACKNOWLEDGEMENTS

The authors thank Drs. T. Miyata and H. Hayashida, Department of Biophysics, Faculty of Science, Kyoto University, for computing K_N, K_S, and K_A values. This work was supported in part by Grants-in-Aid for Scientific Research (nos. 03453165, 03554021, and 05780441) from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. SWISS-PROT Protein Sequence Data Bank (1993, Feb.)
2. Dijkstra, B. W., Kalk, K. H., Hol, W. G. J., & Drenth, J. *J. Mol. Biol.* **147**, 97 (1981)
3. Maraganore, J. M., Mcrutka, G., Cho, W., Welches, W., Kezdy, F. J. & Heinrikson, R. L. J. Biol. Chem., **259**, 13839 (1984)
4. Tanaka, S., Mohri, H., Kihara, H., & Ohno, M. *J. Biochem.* (Tokyo) **99**, 281 (1986)
5. Oda, N., Ogawa, T., Ohno, M., Sasaki, H., Sakaki, Y., & Kihara, H. *J. Biochem.* (Tokyo) **108**, 816 (1990)
6. Yoshizumi, K., Liu, S.-Y., Miyata, T., Saita, S., Ohno, M., Iwanaga, S., & Kihara, H. *Toxicon* **28**, 43 (1990)
7. Liu, S.-Y., Yoshizumi, K., Oda, N., Ohno, M., Tokunaga, F., Iwanaga, S., & Kihara, H. *J. Biochem.* (Tokyo) **107**, 400 (1990)
8. Matsuoka, M., Itoh, H., Kozasa, T., & Kaziro, Y. *Proc. Natl. Acad. Sci. USA* **85**, 5384 (1988)
9. Ohno, S., Kawasaki, H., Imajoh, S., Suzuki, K., Inagaki, M., Yokokura, H., Sakoh, T., & Hidaka, H. *Nature* **325**, 161 (1987)
10. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge University Press, Cambridge, 1983).
11. Zuker, M., & Steigler, P. *Nucl. Acids Res.* **9**, 133 (1981)
12. Nei, M. *Molecular Evolutionary Genetics* (Columbia University Press, Irvington-on-Hudson, NY, 1987).
13. Kihara, H., Uchikawa, R., Hattori, S., & Ohno, M. *Biochem. Int.* **28**, 895 (1992)
14. Matsumoto, H., Shimohigashi, Y., & Ohno, M. unpublished results.
15. Fukagawa, T., Matsumoto, H., Shimohigashi, Y., Ogawa, T., Oda, N., Chang, C.-C., & M. Ohno. *Toxicon* **30**, 1331 (1992)