Molecular Cloning and Sequence Analysis of Aggretin, a Collagen-like Platelet Aggregation Inducer

Ching-Hu Chung,* Lo-Chun Au,†^{,†} and Tur-Fu Huang^{‡,1}

*Department of Pharmacology, College of Medicine, National Taiwan University, Taiwan; †Department of Medical Research, Veterans General Hospital-Taipei, Taiwan; and ‡Graduate Institute of Medical Biotechnology, National Yang-Ming University, Taipei, Taiwan

Received September 2, 1999

A cDNA library derived from the Malayan-pit-viper (Calloselasma rhodostoma) venom gland was constructed in the phagemid vector. Using the information of the N-terminal amino acid sequences of two subunits of aggretin, synthetic mixed-base oligonucleotides were employed as a screening probe for colony hybridization. Separate cDNA clones encoding for the α and β chains of aggretin were isolated and sequenced. The results revealed that mature α and β chains contain 136 and 123 amino acid residues. respectively. Aggretin subunits show high degrees of identity with respective subunits (50–60% for α , 49– 58% for β) of C-type lectin-like snake venoms. The identity to rattlesnake lectin is relatively lower (i.e., 39 and 30%). All cysteine residues in each chain of aggretin are well conserved and located at the positions corresponding to those of C-type lectins. Thus, three intracatenary disulfide bridges and an interchain disulfide bond between Cys83(α) and Cys75(β) may be allocated. This is the first report regarding the entire sequence of venom GPIa/IIa agonist. According to the alignment of amino acid sequences, hypervariable regions among these C-type lectin-like proteins were revealed. These hypervariable regions are proposed to be the counterparts directly interacting with different receptors or different domains of a receptor on the surface of platelet. © 1999 Academic Press

Snake venoms affect blood coagulation and platelet function in a complex manner (1). The most pronounced anti-platelet constituents are Arg-Gly-Asp containing trigramin-like peptides, which have been identified as specific fibrinogen receptor antagonists and named as disintegrins (2-4). Several noncoagulant, non-enzymatic proteins that cause platelet aggregation have been purified from different snake venoms (5–7). Aggretin, a potent platelet-aggregating protein purified from *Calloselasma rhodostoma* venom, is a 29 kDa heterodimeric protein consisting of two subunits with molecular masses of 18 kDa and 15 kDa, linked by disulfide bridge (5). It elicited platelet aggregation in both human platelet-rich plasma and washed platelet suspension. It activates platelet through a Ca^{2+} -dependent mechanism and is devoid of enzyme activity. It may activate platelets through the binding of a collagen-like receptor (i.e. glycoprotein Ia/IIa) on the platelet membrane (5).

Rattlesnake lectin (RSL) from Crotalus atrox venom was determined as a C-type lectin and it consists of a carbohydrate-recognition domain (CRD) (8). RSL is composed of covalently linked and formed as a homodimer molecule. Many distinct proteins isolated from snake venoms consist of two heterogeneous CRDlike domain linked by an interchain S-S bond. For example, the sequences of glycoprotein Ib agonist or antagonist revealed high homology with those of C-type lectins (9), including botrocetin isolated from venom of Bothrops jararaca (10), bitiscetin from Bitis arietans (11), habu IX/X-bp from Trimeresurus fla*vovoridis* (12), and echicetin from *Echis carinatus* (13, 14). A collagen-like platelet aggregating protein, convulxin, from venom of Crotalus durissus terrificus binds to a specific platelet receptor GPVI. Convulxin consists of two subunits, α and β , and the two chains were homologous to those of the CRD of C-type lectins (15). From the available data concerning the N-terminal sequences of α and β -chain of aggretin, we found that they share a high degree of homology as compared with the respective chain of these GPIb-binding proteins, even though aggretin acts on a distinct target, GPIa/IIa rather than GPIb (5).

This study was supported by grants from National Science Council of the Republic of China (NSC87-2314-B-002-301).

¹ To whom correspondence and reprint requests should be addressed at Department of Pharmacology, College of Medicine, National Taiwan University, No. 1, Sec. 1, Jen-Ai Rd., Taipei, Taiwan. Fax: 886-2-23417930. E-mail: turfu@ccms.ntu.edu.tw.

In this study, we isolated cDNA clones of aggretin α and β subunits. The amino-acid sequences were deduced after the determination of nucleotide sequence of each clone. The polypeptide sequences were found to align well with those of CRD containing GPIb-binding proteins, although they may act in different way to induce platelet activation or to inhibit the interaction between von Willebrand factor and GPIb (16, 17).

MATERIALS AND METHODS

Determination of N-terminal amino acid sequence of aggretin. Aggretin was purified from crude venom of *Calloselasma rhodostoma* according to the method described previously (5). Aggretin (10 μ g) was dissociated by boiling in protein denaturing buffer containing β -mercaptoethanol and then electrophoresed into 12% SDS-PAGE. Two bands (subunits), α , 18 kDa and β 15 kDa were revealed after electro-blotted to PVDF membrane according to the method of Matsudarin *et al.* (18), and stained with Coomassie blue. The protein of each band was subjected to determination of N-terminal amino acid sequencing by using an Applied Biosystems model 477A pulsed liquid-phase sequencer equipped with an on-line 120 PTH amino acid analyzer. The N-terminal sequence of 24 residues for the α subunit and N-terminal sequence of 18 residues for the β subunit were determined (Fig. 1).

Construction of venom gland cDNA library. The cDNA library of the Calloselasma rhodostoma venom gland was constructed previously (19). A set of cDNAs derived from a poly(A)⁺ RNA preparation of the venom glands of Calloselasma rhodostoma were cloned into the Uni-ZAP XR vector (20) by inserting the cDNAs into the EcoRI and XhoI sites of the Uni-ZAP XR phage DNA. For convenience, the Uni-ZAP XR vectors was excised *in vivo* to generate a pBluescript SK-phagemid cDNA library by co-transfecting the Uni-ZAP XR phage and an R408 helper phage (21) into Escherichia coli XL1-Blue host cells according to the manufacturer's instructions (Stratagene, La Jolla, CA, USA).

Primer design. A 20meric mixed-base oligonucleotide probe, based on the N-terminal amino-acid sequence of α subunit was synthesized (Fig. 1). For the screening of β subunit, a 20meric oligonucleotide probe was synthesized according to the conserved sequence, mamushigin's coding region (22), of GPIb-binding protein (Fig. 1).

Isolating of cDNA clones of aggretin. E. coli XL1-Blue was infected with the rescued phagemid. The infected bacteria were plated out on to Luris broth/ampicillin plates. Ampicillin-resistant colonies were transferred on to nitrocellulose membrane and prepared for colony hybridization (23). The 20meric oligonucleotide probe (see Fig. 1) was ³²P-end-labeled with T4 polynucleotide kinase and $(\gamma^{-32}P)ATP$ (24). The membrane replicas were pre-hybridized at 65°C for 3 h in hybridization buffer (0.45 M sodium chloride, 0.045 M sodium citrate, 0.02 M potassium phosphate, pH 6.8, 0.1% (w/v), 0.1% (w/v) ficoll and 0.02% (w/v) polyvinylpyrrolidone) plus 100 μ g/ml sheared and denatured herring-sperm DNA. The membrane replicas were then hybridized at 49°C for 8 h in 30 ml of hybridization buffer containing 5 pmol of the labeled oligonucleotide probe. Subsequent washes of the hybridized membrane filters were carried out in 6 imesSSC (1 \times SSC is 0.15 M NaCl/0.015 M sodium citrate) and 0.1% (w/v) SDS as follows: four times at 4°C for 5 min each, once at room temperature (5 min) and once at 47°C (2 min). The positive clones were detected by autoradiography, and were further purified by the isolation of single cell colonies.

DNA sequencing and nucleotide analysis. DNA sequencing was performed by the PCR/dideoxyl method using a ABI PRISM BigDye

Aggretin α chain N-terminal



Aggretin B chain N-terminal



FIG. 1. The N-terminal amino acid sequences of aggretin α/β chains (bold letters) and the sequences of the oligonucleotides used as hybridization probes. Py, pyrimidines (T or C); Pu, purines (A or G).

Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, USA) and an automatic sequencer (ABI PRISM 377-96 DNA Sequencer, Perkin-Elmer, CA, USA). The two primers used, T3 and T7, are located in the flanking regions of cDNA.

T3 primer: 5' AATTAACCCTCACTAAAGGG 3'

T7 primer: 3' CGGGATATCACTCAGCATAATG 5'

RESULTS AND DISCUSSION

Screening of a cDNA Library

In order to design the oligonucleotide probes for screening the cDNA library, N-terminal sequences of α and β polypeptide chain were first determined (see Fig. 1). A 20meric mixed-based oligonucleotide probe with 64 combinations was synthesized for the isolation of α -subunit. On the other hand, we used defined-sequence probe for β subunit, which was referred to C-type lectin conserved mamushigin's cDNA codons (22), since too many combinations existed in mixed-base probe. Six positive clones of each subunit were isolated separately after colony hybridization. All clones had been sequenced. The nucleotide sequences for clones of the same group shared an identical sequence although they were different in length (different clones).

Analysis of cDNA Clones

The longest cDNA sequences and the open reading frames deduced from them are shown in Fig. 2. The results indicate that these cDNA do encode α and β subunits, since N-terminal sequences of α and β subunit shown in Fig. 1 appear in open reading frames of these two clones. Mature α and β subunits are composed of 136 and 123 amino-acid residues, respectively.

(A) aggretin α chain

- 37 gatcagcattgctaccaggcattcaatgaacaaaaaacctgggatgaggcaggaagttc 96 D Q H C Y Q A F N E O K T W D E A E K F
- 97 tgcagggcgcaggagaatggtgcgcatctggcctctatcgaaagcaatggagaagcagaa 156 C R A Q E N G A H L A S I E S N G E A D
- 217 ggactgaggggtcaaaacaaagaacagcaatgcagctggggtggagggatggctccagt 276 G L R A Q N K E Q Q C S S E W S D G S S
- 277 gtcagttatgagaacttgattgatctacatacgaaaaagtgtggtggcgctggaaaaactg 336 V S Y E N L I D L H T K K C G A L E K L
- 397 etectgecatattaaagatecagttgtgtgaagtetggagaaggaaggeececeacecaa 456 L L P Y *
- 457 tececeaccececcacctgecgaatetetgtagetggatetggtttgetgeteetga 516 517 tgggccagaaggtecaataaattetgectageatgaaaaa.....

(B) aggretin β chain

- -102 ttgtggccgagggacagggacggaaggacaggacggacgcattcatcttcgtgagcttc -43 M G R F I F V S F
 -42 ggcttgctggtcgtgttcctctccctgagtggtactggagctgattgtccctctggttgg 18 G L L V V F L S L S G T G A D C P S G W
- 19 teeteetatgaagggeattgetaeaageeetteaatgaaeegaaaaaetgggeegatgea 78 S S Y E G H C Y K P F N E P K N W A D A
- 139 gaagcagattttgtggtcaagttgaccagaccaaggttgaaagccaatttagtctggatg 198 E A D F V V K L T R P R L K A N L V W M
- 199 ggactgagcaatatctggcacggatgcaactggcagtggagtgatggcgcccaggctcaac 258 G L S N I W H G C N W Q W S D G A R L N
- 256 tacaaagactggcaggaacaatctgaatgtctcgcattcaggggagttcatacagaatgg 318 Y K D W Q E Q S E C L A F R G V H T E W
- 319 ttaaatatggactgcagcagtacttgctctttcgtctgcaagttcaaggcatagtctgaa 378 L N M D C S S T C S F V C K F K A

FIG. 2. The nucleotide sequences (top line) and the deduced amino acid sequences (bottom line) of the α - and β -subunits of aggretin. A +1 indicates the first amino acid residues of the mature subunits. The polyadenylation signal sequence AATAAA is underlined.

The peptide leading each N-terminal sequence is supposed to be the signal peptide or pre-peptide. We could not find a clone containing both α and β subunit. Besides, the α and β cDNAs shown in Fig. 2 are equipped with their own stop codons and polyadenylation-signal sequences. The results indicated that α/β heterodimer, although linked by a disulfide bond, are encoded by two genes. It is unlike the case of insulin which are originally synthesized by a single-chain precursor followed by folding, disulfide-bridge formation and cleavage to generate heterodimer (25). However, α and β subunit genes share 71% similarity in nucleotide sequence. They may originally duplicate from an ancestral gene and then evolved in diversely.

Amino Acid Sequence Homology

On comparison of the Swiss-Prot database, two groups of proteins, C-type lectins and C-type lectin-like proteins from snake venoms are found to show significant similarities with α and β subunits of aggretin (Fig. 3). The percentages of sequence similarity of these proteins are summarized in Table 1. The mature subunits of aggretin showed a high degree of similarity (38.5-59.7% identity for α subunit and 30.1-57.7%identity for β subunit) with those of C-type lectin-like proteins. Aggretin α chain contains seven cysteine residues, and β chain contains eight cysteine residues. All cysteine residues in each subunits of aggretin are well conserved and located at the positions corresponding to C-type lectins (9). According to the structural model of the CRD proposed by Spiess (26), six of these cysteines are involved in the intracatenary disulfide bridges (i.e. Cys 5-16, Cys 33-131, Cys 106-123 of the α subunit; and Cys 2-13, Cys 30-119, Cys 96-111 of the β subunit). According to the information of RSL and botrocetin, besides these six intracatenary disulfide bridges, a disulfide bridges between the two subunits has been shown to associate cysteine residues around position 80 (botrocetin α Cys80- β Cys75 and RSL 86-86) (8, 14). Therefore, it is likely that the interchain disulfide bond of aggretin may similarly exist around position 80 (Cys-83 of the α chain and Cys-75 of the β chain). It is noted that aggretin possesses an additional cysteine residue (Cys 115) in the β chain which is probably a candidate for disulfide linking. It is not known what a role the extra cysteine residue plays.

From alignment of the C-type lectin sequences, there was no specific fragment could be detected in the polypeptide sequence involved in the interaction between the C-type lectin and its receptor. Members of GPIb binding proteins show a high degree of similarity with C-type lectin, and most snake venom GPIb binding proteins do not induce direct platelet aggregation. However, one GPIb binding protein, AL-B from Trimeresurus albolabris does induce direct platelet aggregation (27). This is an open question because both categories of snake venom GPIb binding proteins having similar structural features exhibit different biological activities. From the result of the full sequences of α and β subunits of aggretin, we found that they share a high degree of homologous sequence as compared with the respective chain of these GPIb binding proteins. Therefore, the detailed study of their structureactivity in eliciting platelet responses for aggretin and these GPIb binding proteins will allow us to reveal which polypeptide sequences are likely to interact with receptor(s) on the platelet membrane.

A potent platelet-aggregating protein, convulxin, from the venom of *Crotalus durissus terrificus* has been recently cloned and fully sequenced (15). It

| | | | | | * | | | | | | | | | * | | 2 | 20 | | | | | | | | | * | | | | | 40 | | | | | | | | | | | | | | | 6 | 0 | | | | | | | | | | |
|---|--|---|---|---|--|---|---|--|--|---|---|---|--|--|---|---|---|--|---|---|---|---|---------------------------------|---|---|-------------------------|--|--|--|--|--------------------------------------|--|-------------------|--------------------------------------|---|-----------------------|--|---|---|---|---|---|--|---|---|--|--|--|--|--------------------------------------|--|--|--|---|---|--|---|
| aggretin a | 1 | (| L | ΕI | э с | D | FG | W | S | ΡY | ζD | 0 | н | CΥ | 0 | A | FN | ΕE | QK | Т | W | DE | A | Εŀ | F | CI | RΑ | 0 | ΕN | łG | A I | нι | A | S | ιE | S | - 1 | N G | βE | A C |) F | V | sν | VL | 1 | S (| Эk | D | Е | L | A I |) E | Ð | Y | - V | - 70 | 0 |
| convulxin a | 1 | (| ìΕ | Н | - C | Р | S D | w | Y | ΥŊ | í D | Ô. | н | C Y | R | I | ΕN | E | EN | 1 N | W | ЕΕ | A | ΕV | / F | C ' | тκ | ò | A K | G | A 1 | нL | v | S | ιк | S | - 4 | ٩к | E | A |) F | v | ΑV | vм | v | T (| ŝъ | εī | E | E | s | FS | H | - | - v | 6 | 8 |
| AL-B a | i | | | <u> </u> | ъċ | P | S D | w w | s | S F | FК | ò | γd | cγ | 0 | 1 | FΚ | 0 | L B | т | w | εс | A | ΕF | EF | Ċ N | ИD | ò | VK | G | A | нт | v | s | I E | s | . 1 | ίR | E | A | V F | v | A | 0 | ÷. | s | FΝ | ÷ν | К | т | ΤI | cγ | D | - | . v | 6 | 6 |
| hotrocetin a | i | | | 1 | ъ с | P | s G | w | s | \$ 1 | ΖĒ | à | N (| \mathbf{x} | ĸ | F | ΕO | òò | κn | f N | w | АΓ | - A | E F | F | C : | S E | ò. | AK | Ğ | G | нĩ | v | ŝ | ιĸ | ĩ | vš | k K | 3 | ĸ |) F | v | бι | Ϋ́ | v | т | 2 1 | | 0 | ç | s i | ĥ | v | | . A | 6 | 7 |
| monuchigin g | ÷ | FI | n n | s i | л с | P | s n | w | s | SN | JG | R | E (| ~ v | ĸ | 1 | ΕÔ | ñ | K N | 1 K | w | ΔГ | Δ. | E I | E F | È. | тЕ | à | R 1 | | A 1 | ні | v | s | 1 5 | ŝ | | ιп | - E | Δ. | A F | v | | х м. | i | è | E N | | ř | v | T I | 5.0 | | - | - 14 V | 61 | à |
| mannusingin o. | ÷. | | , р | 01 | 5 0 | T | 00 | | r e | Б X | , G Z D | 6 | и с и с | \tilde{v} | 0 | T I | ED | X | 1 12 | - т | w | пв | · | C L | · · | ~ I | I L | Ř | wr | 56 | 6 | ur | Ϋ́. | e . | 1 0 | 0 | | | v | - | 1 I C C | Ň | | 2 141 | - | с : | | | n. D | v | | | | - - | - v - v | 4 | <i>.</i> |
| echicetin a | - | | 0 | Ŷ | 20 | т. т | 50 | 1 11 | | г л г л | (15 7 E | 6 | u o | ~ v | ž | , L | | ~ | VV | | w | | | E I | | 2 | т ю | Š | ** L A 1/ | | | | . v | 3 | | 0 | - 1 | | · | ~ · | с г с г | Ň | | 2 I. | ÷. | 5 | ~ r | | r v | ь Б | 51 | 1 I 1 I | | | × v | 0 | 2 |
| nanu iX/X-op α | - | | | | J C | L | 30 | - W | | 51 | | 0 | n (| | Γ. | ? | г L С N | | 1 1 | • 1 • • | WV NU | с L г г | | E 1 E 1 | | C . | i E | Υ. | A B R B | | A | п L | · · | 3 | | 3 | | | | A 1 | | Ŷ. | A (| 2 | | 1 1 | , r | 111 | ĸ | ĸ | | 1 | . Y | - | - 1 | 0 | 3 |
| RSL | 1 | | | NI | N C | P | ιu | F W | L | РЛ | a N | 0 | L | Y | к | 1 : | r N | ιŲ | Γĸ | . 1 | w | εı | A | εN | 1 P | C I | КΚ | Y | K. F | 'G | C. | нц | A | S | rн | к | - 1 | é G | iΕ | 8 | LE | I | AI | : Y | 1 | S I | 2 Y | ' H | K. | G | Q. | 5.5 | - | - | - V | 6 |) |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | 80 | • | | | | | | | _ | | | | | | | _ | | | 100 | | _ | | | | Ι. | | | | | | | | | | 120 | | * | | | | | | | * | | | | | | |
| aggretin α | 71 | WI | G | L. | R A | Q | ΝК | E | Q | QC |) s | S | ΕV | N S | Ð | G | s s | v | S Y | - | - | | E | NI | . 1 | D : | LΗ | Т | Κĸ | - | - 1 | СG | A | L. | е к | L | τe | 3 F | R | Κ | γv | N | Υ' | C C | E | QN | Λŀ | ΙA | F | V | C 1 | < L | L | P ' | ŕ | | 136 |
| convulxin α | 69 | S I | G | L | R V | Q | и к | . E | К | QC | S | Т | ΚV | N S | D | G | s s | V | SY | - | ^ | | D | NI | . L | D | ιY | 1 | ΤĶ | - 1 | - 1 | C S | L | LI | ĸκ | Е | T (| j F | R | K١ | V F | v | A S | S C | 1 | GI | K I | Р | F | V | CI | ς F | P | P (| ŚС | | 135 |
| AL-Bα | 67 | WJ | G | L | sν | v | ΝK | G | Q | QC | S | S | ΕV | N S | D | G | s s | v | SΥ | - | - | | Е | NI | . V | K | ΡL | S | Κķ | | - 1 | CF | V | LI | ĸκ | G | ΤI | ΞF | R | K١ | VF | Ν | V / | ۱C | E | QI | Κŀ | ιL | F | М | C I | ΚF | L | R I | ? R | | 133 |
| botrocetin α | 68 | W 1 | G | L | RΥ | Е | ΝК | E | к | QQ | S | S | ΕV | N Ş | D | Ģ | S S | V | SΥ | (- | - | | E | N١ | / V | ΕI | RТ | v | ΚĶ | - 1 | - 1 | CF | A | Ľ | ΕK | D | L (| ĴΡ | V | Ľ١ | V I | Ν | L١ | (C | А | QI | K N | I P | F | V | C I | < S | Р | P : | > | | 132 |
| maniushigin α | 70 | W I | G | Ľ | ΓV | Q | ΝE | Ε | Q | QC | CΚ | S | RV | ₿ S | D | R. | S S | V | SΥ | (- | - | • • | Ε | NI | . V | ΚI | ΡN | S | Κk | ι. | - 1 | CF | V | LI | ΚE | Y | ΕC | 3 S | R | K١ | VF | Ν | ٧Y | ' C | G | QI | κı | ' A | F | М | C I | ΚF | L | R | P R | | 136 |
| echicetin a | 69 | W I | G | L | R D | R | sк | R | Ε | Q (| G | н | LΝ | ΝT | D | N | S F | v | ΗΥ | ′ - | - | | Е | н١ | / V | P 1 | Ρ- | - | Τķ | - 1 | - 1 | CF | V | L | ER | Q | ΤĒ | ΕF | R | K١ | V I | А | ٧N | I C | E | FΙ | ΚF | P | F | v · | C I | < A | ĸ | 1 | P R | | 133 |
| hahu IX/X-bp α | 67 | W | G | L | RΥ | Q | GК | v | к | Q (| C N | S | ΕV | V S | D | G | s s | V | SΥ | 1 - | - | | E | ΝV | 7.1 | E, | ΑE | s | κт | | - 4 | сL | G | L. | εĸ | E | ΤΙ | ЪF | R | К١ | vν | Ν | L Y | ′ C | G | 0 0 | o N | ΙP | F | v · | C I | ΞA | | | | | 129 |
| RSL | 67 | W I | G | L | R D | 2 | - к | K | К | DF | = s | W | ΕV | ΝT | D | R : | sс | Т | DΥ | L | T | wε | ĸ | N (|) P | DJ | ΗY | 0 | N K | ΞE | E (| сν | Е | Ľ | v s | L | тс | ĴΥ | R | L | νN | D | ٥١ | / c | Е | ŝı | λī | A | F | L. | c (| ъc | к | F | | | 135 |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | • | | | | | | | | | | | | | 2(|) | | | | | | | | | | | | | 4 | 40 | | | | | | | | | | | | | | | 6 | D | | | | | | | | | |
| aggretin β | 1 | | D | * C 1 | P S | G | ws | s | Y | ЕC | ЭH | * C | Yŀ | < P | F | N | 20 E P |) 'К | NV | / A | D | АE | R | , F (| к | L | ŞΡ | к | ня | а н | 4 L | 40 V S | F | Q | S A | - | ЕВ | ΞA | D | F | vv | ĸ | LI | R | Р | R | 6 L k | D : A | N | L | v | V M | 1 G | L | S N | 70 |) |
| aggretin β convulxín β | 1 | GE | D | * c : c | PS PS | G H | w s w s | S S | Y Y | E C D F | ЭН ХҮ | * c c | Y H Y H | < P < V | F | N I K (| 20 E P Q E | о к к | N V T V | / A / A | DD | A E A E | RK | , F (F (| к | LO | ŞР | К Т | H S G S | н | 4 L' L' | 40 V S V S | F | Q : H : | S A S T | | ЕН | e a e v | D / D | F | v v v v | KK | L T MI 1 | R | P | R | 6 L K L K | D A S | N T | L F | V V F V | V M V T | i G G | L | S N N N | 70 |) |
| aggretin β convulxin β AL-B β | 1 1 1 | GE | D C D | * C C | PS PS | G H D | ws ws | S S S | Y Y Y | E C D F D I | G H R Y V Y | * C C C | Y H Y H Y H | < P < V R V | F F F | N I K (| 20 E P Q E E K |) K M | N V T V N V | / A / A / E | D D D | A E A E A E | RK | F (F (F (| ск Т Т | L (Q (| а С Н С Н | К Т Т | H S G S D S | н н н | 4 L' L' | 40 VS VS VS | F F | Q : H : D : | SA ST SS | | E I E I E I | E A E V E V | . D 7 D 7 D | F | V V V V V A | K K S | LT MI1 K1 | R | P Q P | R I S I V I | 6 1 k 1 k 1 k | D A S S H | N T D | L F L | V V F V V V | V N V I V I | 1 G G G | | S N N N G S | 70 72 70 |) 2 |
| aggretin β convulxin β AL-B β botrocetin β | 1 1 1 | GE | | * C C C | PS PS PP | G H D | ws ws ws | S S S | Y Y Y Y | E C D F D I E C | GH XY XY | * C C C C | Y H Y H Y H Y H | <pre>< P < V < V < V < F </pre> | FFFF | N I K Q K | 2(E P Q E E K E W |) K K K | N V T V N V H V | / A / A / E / D | D D D D | A E A E A E | RK | F (F (F (F (| к т т | | Q P Q H Q H Q O | K T T | H S G S G A | H H H H | 4 L' L' L' | 40 VS VS VS | P F F | Q H D | SA ST SS SK | | E I E I E I | E A E V E A | 0 D / D / D / D | F | V V V V V A | K K S | L 1 M 1 K 1 | R | P Q P E | R I S I V I | 6 L K L K L K | D A A A A A A A A A A A A A A A A A A A | N T D | L F L | V N F N V N | V N V I V I V I | 1 G G G | | S N N N G S D | 70 71 70 70 |)) |
| aggretin β convulxin β AL-B β botrocetin β manuchicin β | 1 1 1 1 | G I | | * C C C C C C C | PS PS PS | G H D D | WS WS WS | S S S | Y Y Y Y | E C D F D L E C | 5 H 2 Y 5 H 5 H | * CCCCC | Y H Y H Y H Y H Y H | <pre>< P < V < V < F </pre> | FFFF | N I Q I K I | 20 E P Q E K E W |) K K K M | N V T V N V H V T V | / A / A / E / D | | A 8 A 8 A 8 A 8 | RKEK | F (F (F (F (| K C T T T | | Q P Q H Q H Q Q | K T T | H S G S G A | | 4 L' L' L' | 40 VS VS VS | P F F F | Q III | SA ST SSK | | E I E I E I E I | E A V E A V E A V | | F F F F | VV VV VA | K S S | L 1 M 1 K 1 L 1 | RHESW | P Q P E | R I S I M I | 6 L K L K L K | D A S A S A S A S A S A S A S A S A S A | NTDDD | L F L V | V N F N V N V N | V N V I V I V I | 1 G G G G | | S N N N G S D | 70 71 70 70 70 | 222 |
| aggretin β convulxin β AL-B β botrocetin β manushigin β cabinatin β | 1 1 1 1 1 | G I | | * 00000 | P S P S P P S P S | G H D D D | ws ws ws ws | S S S S V | Y Y Y Y Y Y | E C D F D I E C E C | 5 H 2 Y 5 H 5 H 5 H | * CCCCCC | Y H Y H Y H Y H Y H | <pre>< P < V < V < F < V </pre> | FFFFF | N I Q I Q I K Q I | 20 E P E E E E E E E E |) K K M M | N V T V N V H V T V | VA VA VA VA | | A 8 A 8 A 8 A 8 A 8 A 8 A 8 | RKK | F (F (F (F (F (| | | | K T T K | HSGDSGA | | 4 L' L' L' | 10 VS VS VS VS | F F F F F | Q H D H H | SA ST SS SS SS | | EI EI EI EI | | | FFFF | V V V V V A V R V V | K S S S | L 1 M 1 K 1 L 1 M 1 | RHESW | P Q P E P | R S V I | 6 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | D A A A A A A A A A A A A A A A A A A A | N T D D F | L F L F | | | 1 G G G G G G G | | S N N N G S D N N D | 70 71 70 70 70 70 |))) |
| aggretin β convulxin β AL-B β botrocetin β mamushigin β echicetin β | 1 1 1 1 1 | G I | | * 0000000 | P S P P S P P S P S P S P S P S P S P S | G H D D D D | ws ws ws ws | S S S V | Y Y Y Y Y Y | ECF | 5 H Y Y H H F Y H H H F Y H H H F Y H H H F Y H H H H | * 0000000 | Y H Y H Y H Y H Y H | < P < V < F < V < F < V | FFFFFF | NKQKQK | 20 EPEKEE EEKEE | K K K K M K M | N V T V H V T V N V | AAEDEAA | | A E E A E E A E E A E E A E E A E E A E E A E E E A E | RKKEKK | F (F (F (F (F (| KTTT | | Q P Q H Q Q Q R Q R Q R | K T T K K | HSGDS | | | 40 VS VS VS VS | F F F F F F F | Q H Q H R | SA ST SS SS SS SS | | E I E I E I E I K I | E E E E E E E E | | F F F F F | V V V V V A V R V V A I | K K S S S S | L 1 M 1 K 1 L 1 L 2 | RHESWF | P Q P E P | R S M I M | 6 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | D A A A A A A A A A A A A A A A A A A A | NTDDDE | L F L V F L | | | 1 G G G G G G G | | | 70 72 70 70 70 70 70 |) 2))) |
| aggretin β convulxin β AL-B β botrocetin β manushigin β echicetin β hahu 1X/X-bp β | 1 1 1 1 1 1 | GI | | * 0000000 | P S P S P S P S P S P S P S P S P S P S | G H D D D D D D | ws ws ws ws | S S S S V S | Y Y Y Y Y Y Y | EDFLECC | 6 H Y Y H 6 H 7 H 7 | * 00000000 | Y H Y H Y H Y H Y H | <pre>< P < V < V</pre> | FFFFFF | N K Q K S K S K S K S K S K S K S K S K S | 20 E P E E E E E E P I |) K K K K K K K K K K K K K K K K K K K | N V T N V T N V T N V N V N V | A A E D E A A E | | AEEEEEE | RKKEKKN | F () F () F () F () F () F () | C K T T T T T T T | | | K T T K K A | HSSDGEDG | | | io VS VS VS VS VS | FFFFFF | Q H D Q H R Q H | S A S T S S S S S S S S S S S S | | E I E I E I E I E I E I | E A E V E V E V E V E V | | F F F F F F | V V V V V V V V V V V V V V | K K S S S K | L 1 M 1 K 1 L 1 L 1 L 1 | RHESWFF | P Q P E P Q | R S V I M T | 6 L K L K L K L K L K | D A A A A A A A A A A A A A A A A A A A | N T D D E S | L F L V F L I | V V V V V V V V V V V V V V V V V V V | | 1 G G G G G G I G | L : A 1 L : L : L : L : | | 70 72 70 70 70 70 70 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 |
| aggretin β convulxin β AL-B β botrocetin β mamushigin β echicetin β hahu 1X/X-bp β RSL | 1 1 1 1 1 1 1 1 | G I | | * C C C C C C C C C C C C | P S P S P P S P P S P L P L | G H D D D D D D | WS WS WS WS WS | S S S V S P | Y Y Y Y Y Y M | E C D I E C E C E C N C | 3 H 2 Y 3 H 3 H 3 H 3 H 3 H 3 H 3 H 3 H 3 H 3 H | * 00000000 | Y H Y H Y H Y H Y H Y H | < P < V < V < F < V < P < I | F F F F F F F F | N Q K S N | 20 E P E E E E E E E E E E E E E E E E |) K K M M K K | N V T V N V N V N V T V | / A / E / D / E / A / E | D D D D D D D D D | A 8 A 8 A 8 A 8 A 8 A 8 A 8 A 8 A 8 A 8 | R K K K N | F (F (F (F (F (F (| K T T T T R | | Q P H Q Q V H K | K T K K P | H S G S G A S C C C C C C C C C C | | 4 L ' L ' L ' L ' L ' | 40 VS VS VS VS VS | P F F F F F F F F | Q H D Q H R Q H | SA ST SS SS SS SS SS SS SS SS SS SS SS SS | - - - - G | E I E I E I E I E I E I E I | E A E V E V E V E V E A E A | | F - F - F - F - F - F - F - F - F - F - | VV VA VR VV AE | K S S S K Y | L 1 M 1 L 1 L 1 L 1 L 1 L 1 L 1 | R H S S W F S D | P Q P E P P Q Y | R S M I M T | 6 L K L K L K L K L K L K C C | D A A A A A A A A A A A A A A A A A A A | N T D D E S E | L F L V F L I N | V \ F \ V \ V \ V \ V \ V \ V \ | VN VI VI VI VX VX | 1 G G G G G G G G G G | L : A 1 L : L : L : L : L : L : L : | 5 N S D N S N D S N D S N D S N D S N D S N D S N D S N D S N D S N D S N D S N D S N D S N D S N D S N D S N D | 70 71 70 70 70 70 70 70 70 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 |
| aggretin β convulxin β AL-B β botrocetin β mamushigin β cehicetin β hahu 1X/X-bp β RSL | 1 1 1 1 1 1 1 1 | G I | | * C C C C C C C C | PSPSPSPSPSLPS PSPLPSLPSL | G H D D D D D | WS WS WS WS WS WS | S S S V P | Y Y Y Y Y M | E (D F E (E (N (| 5 H 2 Y 5 H 5 H 5 H 5 H 5 H 5 L | * | Y H Y H Y H Y H Y H Y H Y H | < P < V < F < V < P < I | FFFFFF | N Q K Q K S N | 2(E P E E K E E R E P L |) K K M M K K | N V T V H V T V N V T V | / A / E / D / A / A / E | D D D D D D D D | A 8 A 8 A 8 A 8 A 8 A 8 A 8 A 8 A 8 | R K K K N M | F (F (F (F (F (F (| K T T T T R | | 2 P 2 H 2 Q 2 R 2 R 2 R 2 H 2 K | K T K K P | H S G S G A G C G C G C | 6 H 6 H 6 H 6 H 6 H 6 H 6 H | | 40 V S V S V S V S V S V S A S | P F F F F F F F | Q H D H R H H | SA SSK SSK SSS SS SS S S S S S S S S S S | | E I E I E I E I E I E I E I | E A E V E A E A E A E A E A E A E A | D / D / D / D / D / D / D / D / D / D / | F | √ V √ A √ V √ V √ V √ V ∧ E | K S S S K Y | L 1 M 1 K 1 L 1 L 1 L 1 L 1 | T R E S W F S D | P Q P E P P Q Y | R S M I H I | 6 L K K L K K L K K K K K K K | D A A A A A A A A A A A A A A A A A A A | N T D D E S E | L F L F L I N | V V V V V V V V V V V V | VN VI VI VI VN | 1 G G G G G G G G G | L : A 1 L : L : L : L : L : L : | 5 N 5 D 5 D 5 N 5 N 5 N | 70 71 70 70 70 70 70 70 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 |
| aggretin β convulxin β AL-B β botrocetin β mamushigin β echicetin β hahu 1X/X-bp β RSL | 1 1 1 1 1 1 1 1 | G I | D D D D N D N | * C C C C C C C C | PSPSPS | G H D D D D D | WS WS WS WS WS WS | S S S V S P 80 | Y Y Y Y Y M | E C D F E C E C N C | 3 H 3 Y 3 H 3 H 3 H 3 H 3 H 3 L | * 0000000 | Y H Y H Y H Y H Y H Y H | <pre>< P < V < I </pre> | F F F F F F F | N I Q I K I S I N I | 20 E P E E E E E E E E C E C |) K K M K K K | N V T V N V T V N V T V | / A / E / D / A / A / E | D D D D D D D D | A 8 A 8 A 8 A 8 A 8 A 8 A 8 A 8 A 8 A 8 | R K K K N M | F (F (F (F (F (| C K C T C T C T C T C T C T C T C T C T C T | L Q Q E Q K Q K * | Q P Q H Q Q Q R Q R Q R Q R V K | K T K K P | H S G S G A G C G C G C | 5 H 5 H 5 H 5 H 6 H 7 H | | 40 VS VS VS VS VS | P F F F F F F F | Q H Q H R H | SA SSK SSK SSS SS RY | - - - G | E I E I E I E I E I E I E I | E A V E V E V E V E A V E A V E A V | D / D / D / D / D / D / D / D / D / D / | F F F F F I | / V / A / R / V / I / V / V / A E | K S S S K Y | L 1 M 1 K 1 L 1 L 2 L 2 L 2 | TR H S W F S D | P Q P E P P Q Y 120 | R : S : M : M : M : H : H : | 6 L K L K K L K K C K C | D A A A A A A A A A A A A A A A A A A A | N T D D E S E | L F L V F L I N | V \ F \ V \ V \ V \ F \ V \ | VN VI VI VI VN VI | 1 G G G G G G G G G G | L : A 1 L : L : L : L : L : | 5 N 5 D 5 N 0 5 N | 70 72 70 70 70 70 70 70 70 | 2 2 2 2 2 2 2 2 2 2 2 |
| aggretin β convulxin β AL-B β botrocetin β mamushigin β cehicetin β hahu 1X/X-bp β RSL aggretin β | 1 1 1 1 1 1 1 71 | G I N I V | D D D D N D N V H | * CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC | PSPPS PSPPS PSPL PS SN | G H D D D D W | WS WS WS WS WS WL | S S S S S S S S S S S S S S S S S S S | Y Y Y Y Y M D | E C D I E C E C N C G A | 5 H | * CCCCCCCC | Y H Y H Y H Y H Y H Y H Y H | K P K K K K K K | F F F F F F F F F | N Q K K K N | 20 EPQEEWE EEVE EEPQ E |) K K K M K K | N V T V H V T V T V | / A / E / D / A / A / E / A | D D D D D D D D | A E E E A E E E A E E E A E E E A E E E E A E E E E E A E | R K E K K N M | F (F (F (F (F (F (F (Q) | C K C T C T C T C T C T C T C T C T C T C T | L Q Q E Q K Q K V I | QP QH QQ QR QR QR QR QR QR LA | K T K K P | H S G S G A G C G C G C O O O O O O O O O O | G H G H G H G H G H G H G V | | 40 VS VS VS VS VS AS | 9 F F F F F F F | Q H Q H R I H | SA SSS SSS SSS RY WL | - - - G | EI EI EI EI KI ES MI | E A E V E V E A E A E A E A E A E A E A E A | | F F F F F I S | VV VA VR VV AE TC | K S S S K Y S | L 1 M 1 L 1 L 7 L 7 F V | R H S S D F V F S D F V F S D | P Q P E P P Q Y 120 K | R S M I H I F I | 6 LK LK LK LK K K K | D A A A A A A A A A A A A A A A A A A A | N D D E S E | L F L V F L I N | V \ V \ V \ V \ V \ V \ V \ | VN VI VI VX VX | 1 G G G G G G G G G | L : A 1 L : L : L : L : L : | | 70 72 70 70 70 70 70 70 | 2 2 2 2 2 2 2 2 2 2 2 2 |
| aggretin β convulxin β AL-B β botrocetin β manuskigin β cohicetin β hahu 1X/X-bp β RSL aggretin β convulxin β | 1 1 1 1 1 1 1 1 71 73 | G H 1 1 V 1 V | D D D D N N N V H V N | * CCCCCCC GK | PSPPS PSPPS PSPL PS SN | G H D D D D W W | WS WS WS WS WL QW | S S S V 80 S S S S S S S S S S S S S S S S S S | Y Y Y Y Y M D D | E C D I D I E C E C N C G A G A | GH XYY GH GY GH GL K | * C C C C C C C C C C C C C C C C C C C | Y H Y H Y H Y H Y H Y H Y H Y H Y H Y H | K P K V R V F K K K K | F F F F F F F F F F F F F F F F F F F | N Q | 20 E P E E E E E E E E E E E E E E E E E E |) K K K K K K | N V T V H V T V T V | / A / E / A / A / A / A / E | | A E E E E E E E E E E E E E E E E E E E | R K K E K K N M | F () F () F () F () F () F () F () F () | C K C T T T T T T T T T T T T T T T T T T T | L Q Q E Q K Q K + C L I | Q P Q H Q Q Q R V K A L L | K T K K P F S | H S G S G A G C G C G C R T | G H G H G H G H G H G H G H G H G H G H | | 40 VS VS VS VS VS VS VS VS VS VS VS | PFFFFFF | | SA SSK SSK SSS SSS RY WL | G N S | EII EII EII KI EI MI | EAU EV EV EV EV EV EV EV EV EV EV C | | F F F F I S D | VV VA VR VV A E T Y | K S S S K Y S S | L 1 M 1 K 1 L 7 L 7 F 1 | TR TE S T W F V C | Р Q Р Р Р Р Р Р Р Р Р Р Р Р Р | R S M I H I F F | 6 ENNELNER ENNER ENNER ENNER ENNER | D A A A A A A A A A A A A A A A A A A A | N T D E S E 123 125 | L F L F L I N | V \ F \ V \ V \ F \ V \ | VN VI VI VX VX VX | 1 G G G G G G G G G G | L : A 1 L : L : L : L : L : L : | 5 N 5 S D 5 S D 5 S N 5 S N 5 S N | 70 72 70 70 70 70 70 70 70 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 |
| aggretin β convulxin β AL-B β botrocetin β mamushigin β cehicetin β habu $1X/X - bp \beta$ RSL aggretin β convulxin β AL-B β | 1 1 1 1 1 1 1 71 73 71 | G H I I V V N | D D D D N N N V H V N V H | * C C C C C C C C C C C C C C C C C C C | PSPPSPSPSPSPSPSPSPSSSSSSSSSSSSSSSSSSSS | G D D D D D W W L | WS WS WS WS WS WZ QW QW | S S S V S 80 S S S S S S S S S S S S S S S S S | Y Y Y Y Y M D D D | ECF DIC ECC GA GA GA | JH VY JH JH JH JH JH JH JH JH JH JH JH JH JH | * CCCCCCCC LPL | YYYYYYY NEXY | KVV KRVV KVV KVV KVV V | F F F F F F F F F F F F F F F F F F F | N I Q I K I V V W I W | 20 E P E E E E E E E E E E E E E E E E E E |) KM KM KK | NV TV HV TV TV | / A / E / A / E / A / E / A | D D D D D D D D | A 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 | R K K K K N M | F () F () | C K T T T T T T T T T T T T T T T T T T | LQQEQKQK * CCC | QP QH QQ QR QR VH V L I T | K T K K P F S S | H S G S G S G C G C G C G C R R S | GH GH GH GH GH GH GF T | | 40 VS VS VS VS VS VS VS VS VS VS VS VS VS | FFFFFFF | | SA SSS SSS SSS VL WL | G N S T | EII EII EII EII EII EII EII EII EII EII | EAVERV EAVERV EAVERV EAVERV EAVERV EAVERV | | F F F F I S D R | VV VV VR VV VV VV V V V V V V V V V V V | K S S S K Y S S P | LI MI KI MI FV FV | TREST TEST TEST TEST TEST TEST TEST TEST | Р | R : S : M : H : F : F : | ERRELLERC APP | | N T D D E S E 123 125 123 | L F L V F L I N | V V V V V V V V V V V V V V V V V V V | VN VI VI VI VN | 1 G G G G G G G G G G | L : A 1 L : L : L : L : L : | 5 N 1 S 2 S 1 N 2 S 2 S 1 S 2 S 1 S 2 S 1 S 1 S 2 S 1 S 1 S 2 S 1 S 1 S 2 S 1 | 70 70 70 70 70 70 70 70 | 2 2 2 2 2 2 2 2 2 2 2 2 |
| aggretin β convulxin β AL-B β botrocetin β manushigin β ecbicetin β hahu 1X/X-bp β RSL aggretin β convulxin β AL-B β botrocetin β | 1 1 1 1 1 1 1 1 1 71 73 71 | G I 1 V 1 V V N V V | DCDDDN DDN N VN VN VN VN | * CCCCCCC GKAK | PSSPPSPSPSS * NNKCCK | G H D D D D D W W L F | WSS WSS WSS WUS QW QW EW | S S S S S S S S S S S S S S S S S S S | Y Y Y Y Y M D D D D | EOF DDU ECCE CO GDU GDU GDU GDU GDU GDU | JHY JHHJJH JH JH JH JH JH JH JH JH JH JH JH | * CCCCCCCC PLF | YYYYYYY NEKY | K P V V V V V V V V V V V V V V V V V V | F F F F F F F F F F F F F F F F F F F | N Q | 20 E P E K E E P L E E E K E E P L E E E A L | K K K K K | N V N V N V N V T V T V | / A / E / A / E / A / E / A | | A 8 8 8 8 4 A 8 8 8 4 A 8 8 8 4 A 8 8 8 4 4 4 8 8 4 4 4 8 8 4 4 4 4 | RKKEKKNM | FF() FF() FF() FF() FF() FF() FF() FF() | C K T T T T T T T T T T T T T T T T T T | | QPHQQQVHK AITA | K T K K A P S S S | H S S S S S S S S S S S S S S S S S S S | GH GH GH GH GH CH VFT T | | 40 VS VSS VSS VSS AS EQ N N K | PFFFFFF | | SA SSS SSS SSS WL WL WL | G N S T 1 | EIEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE | | | F F F F F F F R R | VV VV VV VV VV VV VV VV VV VV VV VV VV | K S S S S K Y S S P N | LI MI MI MI FN FN FN | THESWFFD + CCCC | Р Р | R : V : M : M : H : F : F : F : | ELLELES CEAAA | | N T D D E S E 123 125 123 125 | L F L V F L I N | V V V V V V V V V V V V V V V V V V V | VN VI VI VN VN | 1 G G G G G G G G G G G G | L : A 1 L : L : L : L : L : L : | 5 N 3 S 5 D 5 D 5 D 5 D 5 D 7 D | 70 72 70 70 70 70 70 70 70 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 |
| aggretin β convulxin β AL-B β botrocetin β manushigin β cehicetin β habu 1X/X-bp β RSL aggretin β convulxin β AL-B β botrocetin β manushigin β | 1 1 1 1 1 1 1 71 73 71 71 71 | G I 1 V 1 V V V 1 V V V 1 V | DCDDNDN DNN VNNVN | CCCCCCC GKAKE | PSSPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPS | G H D D D D D W W L F V | WSS WSS WSSL QW EW | S S S S S S S S S S S S S S S S S S S | Y Y Y Y Y Y M D D D D D D | EDFLECCO GDLCCC GDLCCC GDLCCC GDLCCC GDLCCC GDLCCC GDLCCC GDLCCC GDLCCC GDLCCC GDLCCCC GDLCCCC GDLCCCC GDLCCCCC GDLCCCCCC GDLCCCCCCCCCC | GHYYGHYGG AKEEA | * CCCCCCCC CCCCCC FLFL | YYYYYYY NEKDY | くていてん くちょう くちょう しょうしょう しょうしょう しょうしん しょうしん しょうしん しょうしん しょうしん しょうしん しょうしん しょうしょう しんしょう しんしょ しんしょ | FFFFFF FFFF DEADA | N I Q I I Q I I Q I I Q I I Q I I Q I I Q I I Q I I Q I I Q I I Q I Q I I Q I | 20 PEEKEEPU PEEKEEPU PEEKERPU PEEKEEPU PEEKEEPU |) K K K K K K | N V V T V V T V V N V T V N V T V | / A / E / A / E / A / A / E / A / E / A | | A 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 | RKKEKKNM | FFFFFFF QEEEE | KTTTTMTR EEEEEE | LQQEQKQK * CCCCC | QPHQQQQVHK AITAA | K T T K K A P F S S S A | HSSAGEDGC DGSAGC GC RKFT | CHHHHH VFTTT | | 10 VSS VVSS VVSS TRQEK NNC | PFFFFFF | | SA SS SS SS SS V L W L W L W L | G N S T I S | EIEEEE EEEEE MIEE | | | F F F F F F F F I S D R R R | VV VA VV VV VV VV VV VV VV VV VV VV VV V | K K S S S K Y S S P N N | LI MI LI FN FN FN V | THEST FED + CCCCC | Р | R : S : M : M : F : F : F : F : | ELLELES C A A A A A A A A A A A A A A A A A A | | N T D D E S E 123 125 123 125 123 | L F L V F L I N | V \ V \ V \ V \ V \ V \ V \ | V N V I V I V I V N V I | 1 G G G G G G G G G G | L : A 1 L : L : L : L : L : | 5 N N 5 5 D N 5 5 N 0 | 70 72 70 70 70 70 70 70 72 | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 |
| aggretin β convulxin β AL-B β botrocetin β mamushigin β cchicetin β hahu 1X/X-bp β RSL aggretin β convulxin β AL-B β botrocetin β mamushigin β echicetin β | 1 1 1 1 1 1 1 71 71 71 71 71 | G F 1 V 1 V V V V V V V V V V V V V V V V V | DCDDNDN VNNVNVNVNVNV | * CCCCCCC GKAKED | PSPPPSLSL * COKRONY | GHDDDDWWLFVW | WSS WSS WSSL QW WSSL QW WEEW | SS SS SV 80 7 S 7 7 7 7 7 7 7 7 7 | Y Y Y Y Y Y M D D D D D D D | EOFLECCO ATT GGTGGA GGTGGA | GAYYHGGHL RKEERO | * CCCCCCCC LPLFLL | HHHHHHH VYYYYY NEXDY | くくマネネくくく イイイ | FFFFFFF DEADAA | N I Q I I Q I I Q I I Q I I Q I I Q I I Q I I Q I I Q I I Q I I Q | 20 EPEKWEEPL EEAL DHSLTN |) K K K K I I I I I I I I I I I I I I I | N V V V V V V V V V V V V V V V V V V V | / A / E / A / E / A / E / A / E / A | | A E E E E E E E E E E E E E E E E E E E | RKKEKKNM | FFFFFFF QEEEEE | KTTTTTR EEEEEH | | QHHQQQVHK AITAA | K T T K K A P S S S A A | HSSSASCOCORTSFI | | | 10 VVSSVVSS TNNNQ NNNN | PFFFFFF | | SA SSK SSS SSS WLL WL WWL | G N S T I S R | EFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF | | | F F F F F F F I S D R R R G | VV VV VV VV VV VV VV VV VV VV VV VV VV | K K S S S K Y S S P N N Y | LI MI LI LI FI FI FI FI FI | THESWEED COCCO | Р Р Р Р Р Р Р Р | R S V I M I F I F C F C | ELLEFK KEQQP | | N T D D E S E 123 125 123 125 123 123 | L F L V F L I N | V V V V V V V V V V V V V V V V V V V | V N V I V I V I V N V N | 1 G G G G G G G G G G G | L : A 1 L : L : L : L : L : | 5 N N 5 5 D N D 5 5 N D 7 5 7 D | 70 72 70 70 70 70 70 70 70 | 9 2 3 3 3 3 3 3 3 3 3 3 2 |
| aggretin β convulxin β AL-B β botrocetin β mamushigin β cehicetin β hahu 1X/X-bp β RSL aggretin β convulxin β AL-B β botrocetin β mamushigin β echicetin β | 1 1 1 1 1 1 1 1 1 1 7 1 7 1 7 1 7 1 7 1 | G F 1 V 1 V V V 1 V 1 V V V V V V V V V | DCDDNDN VNNVNVVNVVNVVNVVNV | * CCCCCCC GKAKEDO | PSPPPSL NNKRMYN | GHDDDDWWLFVWW | WSSSSWWSS WWSSSWW WWSSSWW WWSSSWW WWSSSWW WWSSSWW WWSSSWW | SS SS SS 80 SS S S S S S S S S S S S S S | | EDEEEEN GIIGA | GAYYHHGGHC AKEERAM | * CCCCCCC CCCCCC FLFLL | YYYYYYY NEKDSDR | くくマテマンター ドイングイン | FFFFFFF DEADAA | N Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q | 20 E P E E E E E E E E E E E E E E E E E E | M K K | | / A / E / A / E / A / E / A / E / A / E / A | | A E E E E E E E E E E E E E E E E E E E | RKKEKKNM | FFFFFFF QEEEEE | KTTTTMTR EEEEEHY | | QHHQQQVHK AITAAVIAV | K T K K A P S S S A A F | HGSSASCOORTSFI | CHHHH VFTTTT | | 10 VVSSVVSS EQEKQQK | FFFFFFF | | SATSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS | G NST1SR | EEEEEKEES MARIRE | | | F F F F F F F F F F F F F F F F F F F | VVVAVVVA VVVA TTYK TTYK | K K S S S S K Y S S P N N Y O | LII MIII MIII FFVV FVV | THESWEFD CCCCCCC | Р Q P E P P Q Y 120 К К К E K K F | RSV MIIMIPFC | ELLEFK KEQQEA | | N T D D E S E 123 125 123 125 123 123 123 | L F L V F L I N | V \ V \ V \ V \ V \ V \ V \ V \ | VN VI VI VI VN VI | 1 G G G G G G G G G G G G | | 5 N 5 5 D 5 5 N 5 5 N 5 5 N | 70 70 70 70 70 70 70 | 020000000000000000000000000000000000000 |

FIG. 3. Comparison of the amino acid sequences deduced from aggretin in cDNAs with those of snake venom C-type lectin and group VII C-type lectins. The two aggretin chains are aligned corresponding with the two chains of the following snake venom protein: convulxin from *Crotalus durissus terrificus* venom (15), mamushigin from *Agkistrodon halys blomhoffii venom* (22), alboagreggin-B (AL-B) from *Trimeresurus albolabris* venom (17), botrocetin from *Bothrops jararaca* venom (9), echicetin from *Echis carinatus* venom (13, 14), habu IX/X-bp from *Trimeresurus flavoviridis* venom (12), and Rattlesnake lectin (RSL) from *Crotalus atrox* venom (8). Asterisks indicate the position of well conserved cysteine residues. Gaps (–) are inserted to obtain the maximum similarity.

consists of two subunits, α and β , and encodes sequences of 135 amino acids for the α chain and 125 amino acids for the β chain, existing as hexameric structure (28). The convulxin α/β subunits were homologous to each other and to those of the CRD of C-type lectins. Convulxin has been found to bind to a specific platelet receptor, GPVI, which is also one of the collagen receptors (29, 30). The mature subunits of aggretin showed a high degree of similarity (56.4% identity for α chain and 56.1% identity for β chain) with the mature subunits of convulxin (15). Aggretin is a collagen-like platelet-aggregation protein and it may bind to GPIa/IIa (5). Now, we are trying to define the biological function of aggretin α/β chains

| Seg | uence identity among Aggretin an | a Other Venom Peptic | % of sequer with a | ce identity gretin | | |
|-----------------------------|---|----------------------|-----------------------|-----------------------|--|--|
| Class | Category | Venom peptide | α chain | β chain | | |
| C-type lectin-like peptides | Collagen-like peptide | Aggretin | _ | _ | | |
| | 0 11 | Convulxin | 56.4 | 56.1 | | |
| | GPIb antagonists | Habu IX/X-bp | 59.7 | 57.7 | | |
| | 0 | Mamushigin | 50.7 | 53.7 | | |
| | | Echicetin | 49.6 | 49.6 | | |
| | GPIb agonists | AL-B | 56.4 | 53.7 | | |
| | 5 | Botrocetin | 52.3 | 48.8 | | |
| C-type lectin | Platelet-aggregation inducer ^a | RSL | 38.5 | 30.1 | | |

 TABLE 1

 equence Identity among Aggretin and Other Venom Peptides

^a The mechanism in molecular level is unclear.

and to re-examine the possibility that aggretin may bind to other receptor besides GPIa/IIa. The elucidation of the essential binding motif of this molecule in eliciting platelet aggregation will highlight the identity of collagen-receptor and the signal transduction involved at molecular level.

REFERENCES

- 1. Huang, T. F. (1998) Cell. Mol. Life Sci. 54, 527-540.
- 2. Teng, C. M., and Huang, T. F. (1991) Platelets 2, 77-78.
- Gould, R. J., Polokoff, M. A., Friedman, P. A., Huang, T. F., Holt, J. C., Cook, J. J., and Niewiarowski, S. (1990) *Proc. Soc. Exp. Biol. Med.* 195, 168–171.
- Yen, C. H., Peng, H. C., Yin, J. B., and Huang, T. F. (1998) Biochim. Biophys. Acta 1425, 493–504.
- Huang, T. F., Liu, C. Z., and Yang, S. H. (1995) *Biochem. J.* 309, 1021–1027.
- Teng, C. M., Ko, F. N., Tsai, I. H., Hung, M. L., and Huang, T. F. (1993) *Thromb. Haemost.* 69, 286–292.
- Teng, C. M., Hung, M. L., Huang, T. F., and Ouyang, C. (1989) Biochem. Biophys. Acta 992, 258–264.
- Hirabayashi, J., Kusunoki, T., and Kasai, K. (1991) J. Biol. Chem. 266, 2320–2326.
- 9. Drickamer, K. (1993) Prog. Nucleic Acid Res. Mol. Biol. 45, 207-232.
- Usami, Y., Fujimura, Y., Suzuki, M., Ozeki, Y., Nishio, K., Fukui, H., and Titani, K. (1993) *Proc. Natl. Acad. Sci. USA* 90, 928–932.
- Matsui, T., Hamako, J., Suzuki, M., Hayashi, N., Ito, M., Makita, K., Fujimura, Y., Ozeki, Y., and Titani, K. (1997) *Res. Commun. Cell Mol. Biol.* 1, 271–284.
- Matsuzaki, R., Yoshihara, E., Yamada, M., Shima, K., Atoda, H., and Morita, T. (1996) *Biochem. Biophys. Res. Commun.* 220, 382–387.

- Polgar, J., Magnenat, E. M., Peitsch, M. C., Wells, T. N. C., Saqi, M. S. A., and Clemetson, K. J. (1997) *Biochem. J.* 323, 533–537.
- 14. Peng, M., Holt, J. C., and Niewiarowski, S. (1994) Biochem. Biophys. Res. Commun. 205, 68-72.
- 15. Leduc, M., and Bon, C. (1998) Biochem. J. 333, 389-393.
- Chang, M. C., Lin, H. K., Peng, H. C., and Huang, T. F. (1998) Blood 91, 1582–1589.
- Usami, Y., Suzuki, M., Yoshida, E., Sukurai, Y., Hirano, K., Kawasaki, T., Fujimura, Y., and Titani, K. (1996) *Biochem. Biophys. Res. Commun.* **219**, 727–733.
- 18. Matsudaria, P. (1987) J. Biol. Chem. 262, 10035-10038.
- Au, L. C., Lin, S. B., Chou, J. S., Ten, G. W., Chang, K. J., and Shih, C. M. (1993) *Biochem. J.* 294, 387–390.
- 20. Greener, A. (1990) Strategies in Molecular Biology 3, 5-6.
- 21. Russel, M., Kidd, S., and Kelly, M. R. (1986) Gene 45, 333-338.
- Sakurai, Y., Fujimura, Y., Kokubo, T., Imamura, K., Kawasaki, T., Handa, M., Suzuki, M., Matsui, T., and Titani, K. (1998) *Thromb. Haemost.* 79, 1199–1207.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *in* Molecular Cloning, A Laboratory Manual, 2nd edn., Chapter 1, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- 24. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol.* 65, 507–521.
- Chance, R. E., Ellis, R. M., and Bromer, W. W. (1968) Science 161, 165–167.
- 26. Spiess, M. (1990) Biochemistry 29, 10009-10018.
- Peng, M., Lu, W., and Kirby, E. P. (1991) *Biochemistry* 30, 11529–11536.
- 28. Marlas, G. (1985) Biochimie 67, 1231-1239.
- Jandrot-Perrus, M., Lagrue, A. H., Okuma, M., and Bon, C. (1997) J. Biol. Chem. 272, 27035–27041.
- Polgar, J., Clemetson, J. M., Kehrel, B. E., Wiedermann, M., Magnenat, E. M., Wells, T. N. C., and Clemetson, K. J. (1997) *J. Biol. Chem.* 272, 13576–13583.