Structural properties of trypsin from cold-adapted fish, arabesque greenling (*Pleurogrammus azonus*)

Gaku Kanno [a], Hideki Kishimura [a *], Seiichi Ando [b], Sappasith Klomklao [c], Sitthipong Nalinanon [d], Soottawat Benjakul [d], Byung-Soo Chun [e], Hiroki Saeki [a]

[a] *Laboratory of Marine Products and Food Science, Research Faculty of Fisheries Sciences, Hokkaido University, Hakodate, Hokkaido 041-8611, Japan*

[b] *Faculty of Fisheries, Kagoshima University, Shimoarata, Kagoshima 890-0056, Japan*

[c] *Department of Food Science and Technology, Faculty of Technology and Community Development, Thaksin University, Phattalung Campus, Phattalung 93110, Thailand*

[d] *Department of Food Technology, Faculty of Agro-Industry, Prince of Songkla University, Hat Yai, Songkhla 90112, Thailand*

[e] *Faculty of Food Science and Biotechnology, Pukyong National University, Busan 608-737, Republic of Korea*

[*]Corresponding author. Tel. /fax: +81 138 40 5519.

E-mail address: kishi@fish.hokudai.ac.jp (H. Kishimura).

Abstract

A cDNA clone encoding trypsin (AG-T) was isolated from the pyloric ceca of cold-adapted fish, arabesque greenling (*Pleurogrammus azonus*). The cDNA was composed of 892 bp with an open reading frame of 729 bp at nucleotide positions 25-753. Similar to all the known trypsin, the AG-T seemed to be synthesized as preproenzyme that contains a hydrophobic signal peptide, an activation pentapeptide and a mature trypsin of 222 amino acid residues. The AG-T also completely conserved the major structural features common to trypsin such as the catalytic triad (His57, Asp102 and Ser195), the obligatory Asp189 and twelve Cys residues. On the other hand, the AG-T possessed the deletion of Tyr151 and substitution of Pro152 for Gly in the autolysis loop when aligned with the sequence of tropical-zone fish and bovine trypsins. In addition, Val75 concerned in a combination with calcium ion was exchanged for Ala in the AG-T and the content of positively charged amino acid residues at the calcium-binding site of the AG-T was three times higher than those of tropical-zone fish trypsins. Moreover, the ratio between charged and hydrophobic amino acid residues in the *N*-terminal region of the AG-T was also higher than those of temperate-zone fish and tropical-zone fish trypsins. Such structural properties of the AG-T would contribute to its low thermostability.

Keywords Trypsin, Pyloric caecum, Arabesque greenling, *Pleurogrammus azonus*, Primary structure, Thermostability, Cold-adaptation

Introduction

Trypsin (EC 3.4.21.4), a major digestive enzyme, is a member of the large family of serine proteases. The enzyme specifically cleaves the peptide bond on the carboxyl side of Lys and Arg residues. It is characterized by a common catalytic mechanism involving the catalytic triad of three essential amino acid residues (His57, Asp102 and Ser195) and a substrate determinant residue (Asp189) [1]. Trypsin has been studied in a broad range of species from bacteria to humans because bovine trypsin was isolated and analyzed as the first proteolytic enzyme [1]. The enzyme is an excellent model to study structure-function relationships and many studies have been done to elucidate the structural properties of mammalian pancreatic trypsin [2-7]. On the other hand, fish viscera are a source of trypsin that has some unique properties of interest to both basic research and practical applications [8-13]. Especially, fish trypsin displays substantially higher *kcat*/*Km* value at low temperatures than their mammalian counterparts. For example, the catalytic efficiency of cod trypsin was found to be higher than that of bovine trypsin [14, 15], and that of salmon anionic trypsin was reported to be as much as 40-fold higher [16]. In addition, fish trypsin is more sensitive to inactivation by heat, low pH and autolysis than that of mesophilic analogues [14, 17]. These enzymatic properties of fish trypsin are interesting for several industrial applications, such as in certain food processing operations that require low processing temperatures. Indeed, Atlantic cod trypsin has already been used in industrial applications [18, 19] and medical applications [18].

Recently, we isolated and characterized trypsins from various species of marine fish [29-31]. Then, it was obtained that the strong positive correlation between habitat temperature of these marine fish and thermostability of the fish trypsins [30], and it was also noted that the trypsin from arabesque greenling (AG-T) showed remarkably low

thermostability [21]. The property of AG-T stimulated an interest in the potential of its commercial use. However, there is no structural data concerning the AG-T. Therefore, in this study, we investigated a primary structure of the AG-T.

Materials and methods

Materials

Live specimen of arabesque greenling (*Pleurogrammus azonus*) was purchased at the fish market in Hakodate, Hokkaido prefecture, Japan, in April 2008.

RT (reverse transcription)-PCR and cDNA sequencing

Total RNA was extracted from pyloric ceca of arabesque greenling with a TRIzol regent (Invitrogen, CA, USA). Poly $(A)^+$RNA was isolated from the total RNA using an Oligotex-dt30 (TaKaRa, Kyoto, Japan). The poly $(A)^+$RNA was reverse-transcribed by using a RT-RACE primer (5'-GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTT-3') and a SuperScript Ⅱ (Invitrogen, CA, USA). PCR was carried out using the first strand cDNA and an Amplitaq Gold (TaKaRa, Kyoto, Japan) under the following conditions: 95 ^{o}C for 9 min, 45 cycles of 94 ^{o}C for 15 s, 54 ^{o}C for 30 s, 72 ^{o}C for 60 s and 72 ^{o}C for 7 min. A forward primer (5'-ATCGTCGGAGGGTATGAGTG-3') and a reverse primer (5'-AGTCACCCTGGCAAGAGTCC-3') were designed on the basis of well-conserved partial amino acid sequences (Ile16-Cyc22 and Lys185-Ser192, respectively) in various vertebrate trypsins (Fig. 4). The PCR products were subcloned in a pDrive Cloning Vector

(QUIAGEN, Duesseldorf, Germany) and transformed into JM109 cell (Promega, WI, USA). Plasmid DNA was purified from the positive clone using a Wizard SV Gel and PCR Clean-Up System (Promega, WI, USA). The nucleotide sequence of the cDNA was determined using a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Ca, USA) and an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Ca, USA).

RACE (rapid amplification of cDNA ends)-PCR

Based on the partial nucleotide sequence determined by RT-PCR, the remaining sequences were analyzed by 3'- and 5'-RACE. The 3'-terminal cDNA fragments were amplified using the above first strand cDNA and an Amplitaq Gold (TaKaRa, Kyoto, Japan) under the following conditions: 95 ^{o}C for 9 min, 45 cycles of 94 ^{o}C for 15 s, 54 ^{o}C for 30 s, 72 ^{o}C for 60 s and 72 ^{o}C for 7 min. A forward primer (5'-TCTGCGCTGGATACCTGGAG-3') designed on the basis of the above partial nucleotide sequence and a reverse primer (5'-GGCCACGCGTCGACTAGTAC-3') designed on the basis of the RT-RACE primer.

5'-RACE was performed using a 5' Full RACE Core Set (TaKaRa, Kyoto, Japan). Total RNA was extracted from pyloric ceca of arabesque greenling with a TRIzol regent (Invitrogen, CA, USA). Poly (A)$^{+}$RNA was isolated from the total RNA using an Oligotex-dt30 (TaKaRa, Kyoto, Japan). The poly (A)$^{+}$RNA was reverse-transcribed by using a RT primer (5'-(p)TAGCTGGCCATG-3') and a AMV Reverse Transcriptase XL, and then the first strand cDNA was ringed with a T4 RNA Ligase. The first PCR was carried out using the ringed cDNA and a Premix Taq under the following conditions: 94 ^{o}C for 3 min, 25 cycles of 94 ^{o}C for 30 s, 65 ^{o}C for 30 s, 68 ^{o}C for 30 s. The second PCR was carried out using the amplified product of the first PCR and a Premix Taq under the following conditions: 27 cycles of 94 ^{o}C for 30 s, 65 ^{o}C for 30 s, 68 ^{o}C for 30 s. For the first PCR, a forward

primer (5’-TGTGCTCAGAGAGACAACCC-3’) and a reverse primer (5’-TGTAGCAGTGAGCAGCCAGAC-3’) designed on the basis of the above partial nucleotide sequence. For the second PCR, a forward primer (5’-GGTGTCTAGCGAAAGGTCTG-3’) and a reverse primer (5’-AGAAGTGGTAGCCGGAGTTC-3’) designed on the basis of the above partial nucleotide sequence.

The nucleotide sequences of the RACE-PCR products were determined by the same method in the RT-PCR.

Results and discussion

cDNA characteristics of the isolated clone

The isolated cDNA clone encoding the AG-T (Accession No.: AB441709 in DDBJ) was composed of 892 bp with an open reading frame of 729 bp from the ATG start codon at nucleotide position 25 through to the TAA stop codon at position 753 (Fig. 1). The 5’-noncoding region of AG-T appeared to contain the longest nucleotide sequence (24 bp) among the vertebrate sequence reported so far [32]. The poly (A) signals, as expected, occurred at 25 bp upstream from the first adenine of poly (A) track. The 3’-noncoding region of AG-T had 138 bp nucleotides, and it is much longer than those of other trypsin. The nucleotide sequence of AG-T encoded 242 amino acids starting from the first Met. It was suggested that the AG-T is synthesized as preproenzyme which contains hydrophobic signal peptide, activation pentapeptide and mature trypsin of 222 amino acid residues.

Signal peptide

As shown in Fig. 2, the signal peptide of AG-T was composed of fifteen amino acids and had a hydrophobic core containing seven contiguous hydrophobic residues. It was predicted that the cleavage between signal peptide and activation peptide of the AG-T occurs at the *C*-terminal of Ala residue (amino acid position 15) considering its structural feature. The amino acid sequence of the AG-T signal peptide was compared to those of frigid-zone fish (Atlantic cod, Antarctic fish, Atlantic salmon), temperate-zone fish (anchovy and flounder), tropical-zone fish (zebrafish and tilapia) and bovine trypsins. The signal peptide of bovine trypsin typically contains a hydrophobic core terminated by a helix-breaking Gly residue, and the hydrophobic core falls into two clusters interrupted by a less hydrophobic Ala residue at position 8 [33]. The signal peptide of zebrafish and tilapia (tropical-zone fish) showed the same structural composition as those of bovine trypsin. However, the signal peptides of frigid-zone and temperate-zone fish trypsins had seven contiguous hydrophobic residues common to the AG-T signal peptide.

Activation peptide

The AG-T activation peptide was pentapeptide composed by a poly-anionic cluster (two Glu residues followed by an Asp residue) and a Lys residue located at the *C*-terminal end (Fig. 3). It was suggested that the cleavage site between activation peptide and mature enzyme of the AG-T is at the *C*-terminal of the Lys residue (position 9) and that the enterokinase or trypsin itself of arabesque greenling recognize the poly-anionic cluster to cleave the activation peptide from the proenzyme [34]. The activation peptide of

mammalian trypsin is usually consisted by octapeptide containing a hydrophobic cluster of three amino acid residues followed by a poly-anionic cluster of four contiguous Asp residues. Equally to the AG-T, the sequences of activation peptides of frigid-zone and temperate-zone fish trypsins required the deletion of two amino acid residues (positions 1 and 2) in the hydrophobic cluster when aligned with those of tropical-zone fish and bovine trypsins. The number of amino acid in the poly-anionic cluster of other fish trypsins except for zebrafish was one residue shorter than bovine equivalent in accordance with the AG-T.

Mature trypsin

Common structural properties of AG-T to vertebrate trypsin

The predicted amino acid sequence of mature AG-T is showed in Fig. 4. The AG-T was composed of 222 amino acid residues, and its molecular weight was calculated at 24,028. The theoretical isoelectric point (pI) of AG-T was 5.9. The amino acid sequence identities of the AG-T to frigid-zone fish trypsins (Atlantic cod, 82 %; Antarctic fish, 85 %; Atlantic salmon, 85 %) and temperate-zone fish trypsins (anchovy, 81 %; flounder, 84 %) were higher than to tropical-zone fish trypsins (zebrafish, 65 %; tilapia, 69 %) and bovine trypsin (64 %). The AG-T possessed twelve Cys residues at the appropriate positions in bovine trypsin and completely conserved the catalytic triad (His57, Asp102 and Ser195) and the consensus sequence of GDSGG around the Ser195. The AG-T also preserved the structures for substrate specificity, i.e. S1 substrate-binding pocket (positions 189-195, 214-220 and 225-228), loop 1 (positions 184-188), loop 2 (positions 221-225) and Tyr172 residue.

Bovine trypsin has six disulfide-bridges (Cys15-Cys145, Cys33-Cys49,

Cys117-Cys218, Cys124-Cys191, Cys156-Cys170 and Cys181-Cys205) [35]. The pairing of half-Cys residues in the AG-T was not determined in this study. However, the AG-T may also have six disulfide-bridges, because it possesses twelve Cys residues at the appropriate positions in bovine trypsin. The catalytic activity of trypsin is due to the ability of His57 to transfer a proton from Asp102 to Ser195 [35], and the consensus repeat (GDSGG) is diagnostic of a serine protease [36]. In addition, the steric and electrostatic properties of the S1 pocket are need for substrate specificity [3]. Especially, the preference of trypsin for Lys and Arg residues results from the presence of Asp189 residue at the bottom of the S1 pocket. Moreover, Tyr172 residue interacts synergistically with the residues of the S1 pocket and two surface loops (loop 1 and 2) to determine substrate specificity [4]. In this study, it was demonstrated that the AG-T completely conserved these important structures for catalytic function of trypsin. Therefore, the AG-T has a common structure of vertebrate trypsin family, and its catalytic mechanism is essentially the same as those of them.

Structural properties for low thermostability of AG-T

The deletion of Tyr151 residue and substitution of Gly for Pro152 residue in the autolysis loop (positions 143-153) of the AG-T was existed when aligned with the sequences of bovine trypsin (Fig. 4). These structural differences were also found in frigid-zone and temperate-zone fish trypsins, but not in tropical-zone fish trypsin. Gable and Kasche [37] reported that the cleavage of single-chain bovine trypsin (β-trypsin) at Lys145 residue in the autolysis loop results α-trypsin which shows less thermal stability, and it is thought that the structure of autolysis loop is strongly relates to thermostability of trypsin. From the structural characteristics of anchovy trypsin, Ahsan et al. [32] considered that Gln192 residue located at the entrance of the S1 substrate-binding pocket in anchovy trypsin is much more

freedom with lack of the bulky ring of Tyr151 residue, and substitution of Gly for Pro152 residue in anchovy trypsin could result in a completely different orientation of the autolysis loop between anchovy trypsin and bovine trypsin.

Four amino acid residues (Glu70, Asn72, Glu77 and Glu80) were conserved in the calcium-binding loop (positions 68-80) of the AG-T, but Val75 was exchanged for Ala. Also, the AG-T had three positive charged amino acid residues (His71, Arg74 and Arg84) and three negative charged amino acid residues (Glu70, Glu77 and Glu80) at the region. Bovine trypsin has been found to require Ca^{2+} for thermal stability and resistance to degradation, and this stabilizing effect is accompanied by a conformational change in the trypsin molecule resulting in a more compact structure [38, 39]. The calcium-binding site of bovine trypsin is in the external loop, and five amino acid residues (Glu70, Asn72, Val75, Glu77 and Glu80) are concerned in a combination with calcium ion [39]. These residues were completely conserved in tropical-zone fish trypsins. However, Val75 in the AG-T was exchanged for Ala as well as other frigid-zone and temperate-zone fish trypsins (Atlantic cod, Asn72-His; Antarctic fish, Asn72-His; anchovy, Asn72-His and Val75-Gln; flounder, Asn72-Lys). In addition, it was obtained that the strong positive correlation between the charge at the calcium-binding region and the thermostability of fish trypsins. That is to say, at the calcium-binding region (positions 68-84), the ratio between positive and negative charged amino acid residues of frigid-zone fish trypsins (mean: 1.06) was highest followed by temperate-zone fish trypsins (0.75), tropical-zone fish trypsins (0.29) and bovine trypsin (0.00). It was interpreted that the combination of Ca^{2+} with the AG-T and frigid-zone fish trypsins could be weaker than temperate-zone fish, tropical-zone fish and bovine trypsins, because calcium ion repels positive charged amino acid residues at the calcium-binding region of them.

The percentage of charged amino acid residues (Lys, Arg, Aps, Glu and His) in the

AG-T (19 %) was same as those of other frigid-zone and temperate-zone fish trypsins (mean: 18 %) differing from tropical-zone fish trypsins (mean: 14 %) and bovine trypsin (13 %). Also, it was clarified that the correlation between the charge at the *N*-terminal region and the thermostability of fish trypsins. That is, at the *N*-terminal region (positions 20-50), the ratio between charged and hydrophobic amino acid residues (Trp, Phe, Leu, Ileu, Val, Tyr and Pro) of frigid-zone fish trypsins (mean: 0.73) was highest followed by temperate-zone fish trypsins (mean: 0.63), tropical-zone fish trypsins (mean: 0.27) and bovine trypsin (0.10). Genicot, et al. [40] reported that thermal stability and flexibility of fish trypsin were affected by its overall decrease in hydrophobicity and an increase in surface hydrophilicity as compared to mammalian counterparts. In the previous study, we aligned the *N*-terminal 20 amino acid sequences of several fish and mammalian trypsins and demonstrated that the charged amino acid contents of frigid-zone fish trypsins (mean: 19.2 %) including the AG-T were higher than those of temperate-zone fish trypsins (mean: 11.0 %), tropical-zone fish trypsins (mean: 12.1 %) and mammalian trypsins (mean: 5.0 %) [41].

We predicted the 3D-structures of the calcium-binding region and the *N*-terminal region of the AG-T (Fig. 5). As shown in Fig. 5, the positively charged amino acid residues were observed in the surface of the calcium-binding region of the AG-T, but that was not detected in bovine trypsin. Additionally, *N*-terminal region located at the surface of the AG-T molecule contained more charged amino acid residues than bovine trypsin.

From these results, we believe that the structural characteristics at the autolysis loop, calcium-binding region and *N*-terminal region of the AG-T would contribute its lower thermostability. We should produce the recombinant trypsin from cDNA and obtain the evidences supporting the structural properties. These are currently under investigation.

Conclusions

The AG-T conserved common structural properties of vertebrate trypsin family, and its catalytic mechanism would be essentially the same as those of vertebrate trypsins. On the other hand, the AG-T possessed the deletion of Tyr151 and substitution of Pro152 for Gly in the autolysis loop. In addition, the ratio between positive and negative charged amino acid residues at the calcium-binding region (positions 68-84) and the ratio between charged and hydrophobic amino acid residues at the *N*-terminal region (positions 20-50) of the AG-T and frigid-zone fish trypsins was highest followed by temperate-zone fish trypsins, tropical-zone fish trypsins and bovine trypsin. Such structural characteristics of the AG-T would contribute to its low thermostability.

Acknowledgements This research was partly supported by the grant from HOKUSUI Association. We also appreciate the support of the Core University Program on Fisheries Sciences founded by JSPS & KOSEF.

References

[1] W.R. Rypniewski, A. Perrakis, C.E. Vorgias, K.S. Wilson, Evolutionary divergence and conservation of trypsin, Protein Eng. 7 (1994) 57-64.

[2] R.M. Stroud, L.M. Kay, R.E. Dickerson, The structure of bovine trypsin: electron density maps of the inhibited enzyme at 5 Å and 2.7 Å resolution, J. Mol. Biol. 83 (1974) 185-208.

[3] L. Hedstrom, L. Szilagyi, W.J. Rutter, Converting trypsin to chymotrypsin: the role of surface roops, Science 255 (1992) 1249-1253.

[4] L. Hedstrom, J. Perona, W.J. Rutter, Converting trypsin to chymotrypsin: residue 172 is a substrate specificity determinant, Biochemistry 33 (1994) 8757-8763.

[5] L. Hedstrom, T. Lin, W. Fast, Hydrophobic interactions control zymogen activation in the trypsin family of serine proteases, Biochemistry 35 (1996) 4515-4523.

[6] A. Pasternak, X. Liu, T. Lin, L. Hedstrom, Activating a zymogen with out proteolytic processing: mutation of Lys15 and Asn194 activates trypsinogen, Biochemistry, 37 (1998) 16201-16210.

[7] E. Szabo, Z. Bocskei, G. Naray-Szabo, L. Graf, The three-dimensional structure of Asp189Ser trypsin provides evidence for an inherent structural plasticity of the proteases, Eur. J. Biochem. 263 (1999) 20-26.

[8] G. Jonsdottir, J.B. Bjarnason, A. Gudmundsdottir, Recombinant cold-adapted trypsin I from Atlantic cod – expression , purification, and identification, Protein Expr. Purif. 33 (2004) 110-122.

[9] A. Gudmundsdottir, H.M. Palsdottir, Atlantic cod trypsin: from basic research to practical applications, Mar. Biotech. 7 (2005) 77-88.

[10] M. Macouzet, B.K. Simpson, B.-H. Lee, Expression of a cold-adapted fish trypsin in

Pichia pastoris, FEMS Yeast Res. 5 (2005) 851-857.

[11] H.M. Palsdottir, A. Gudmundsdottir, Development of a qRT-PCR assay to determine the relative mRNA expression of two different trypsins in Atlantic cod (*Gadus morhua*), Comp. Biochem. Physiol. B 146 (2007) 26-34.

[12] H.M. Palsdottir, A. Gudmundsdottir, The novel trypsin Y from Atlantic cod (*Gadus morhua*) – isolation, purification and characterization, Food Chem. 111 (2008) 408-414.

[13] E. Toyota, D. Iyaguchi, H. Sekizaki, M. Tateyama, K.K.S. Ng, A structural comparison of three isoforms of anionic trypsin from chum salmon (*Oncorhynchus keta*), Biol. Crystal. D65 (2009) 713-723.

[14] B. Asgeirsson, J.W. Fox, J.B. Bjarnason, Purification and characterization of trypsin from the poikilotherm *Gadus morhua*, Eur. J. Biochem. 180 (1989) 85-94.

[15] B. Stefansson, L. Helgadottir, S. Olafsdottir, A. Gudmundsdottir, J.B. Bjarnason, Characterization of cold-adapted Atlantic cod (*Gadus morhua*) trypsin 1 – kinetic parameters, autolysis and thermal stability, Comp. Biochem. Physiol. B 155 (2010) 186-194.

[16] A.O. Smalas, E.S. Heimstad, A. Hordvic, N.P. Willassen, R. Male, Cold adaptation of enzymes: structural comparison between salmon and bovine trypsins, Proteins 20 (1994) 149-166.

[17] B.K. Simpson, N.F. Haard, Cold-adapted enzymes from fish, In D. Knorr (Ed.), Food Biotech. New York, Marcel Dekker, 1987, pp. 495-528.

[18] J.B. Bjarnason, Fish serine proteases and their pharmaceutical and cosmetic use, Patent PCT, WO 00/78332 A2, December 28, 2000.

[19] J.B. Bjarnason, B. Benediktsson, Protein hydrolysates produced with the use of marine proteases, Patent PCT, WO 01/28353 A2, April 26, 2001.

[20] H. Kishimura, K. Hayashi, Y. Miyashita, Y. Nonami, Characteristics of two trypsin isozymes from the viscera of Japanese anchovy (*Engraulis japonica*), J. Food Biochem. 29

(2005) 459-469.

[21] H. Kishimura, K. Hayashi, Y. Miyashita, Y. Nonami, Characteristics of trypsins from the viscera of true sardine (*Sardinops melanostictus*) and the pyloric ceca of arabesque greenling (*Pleuroprammus azonus*), Food Chem. 97 (2006) 65-70.

[22] H. Kishimura, Y. Tokuda, S. Klomklao, S. Benjakul, S. Ando, Enzymatic characteristics of trypsin from the pyloric ceca of spotted mackerel (*Scomber australasicus*), J. Food Biochem. 30 (2006) 466-477.

[23] H. Kishimura, Y. Tokuda, S. Klomklao, S. Benjakul, S. Ando, Comparative study on enzymatic characteristics of trypsins from the pyloric ceca of yellow tail (*Seriola quinqueradiata*) and brown hakeling (*Physiculus japonicus*), J. Food Biochem. 30 (2006) 521-534.

[24] S. Klomklao, S. Benjakul, W. Visessanguan, H. Kishimura, B.K. Simpson, H.Saeki, Trypsins from yellowfin tuna (*Thunnus albacores*) spleen: Purification and characterization, Comp. Biochem. Physiol. B 144 (2006) 47-56.

[25] S. Klomklao, S. Benjakul, W. Visessanguan, H. Kishimura, B.K. Simpson, Purification and characterization of trypsin from spleen of tongol tuna (*Thunnus tonggol*), J. Agric. Food Chem. 54 (2006) 5617-5622.

[26] H. Kishimura, Y. Tokuda, M. Yabe, S. Klomklao, S. Benjakul, S. Ando, Trypsins from the pyloric ceca of jacopever (*Sebastes schlegeli*) and elkhorn sculpin (*Alcichthys alcicornis*) : Isolation and characterization, Food Chem. 100 (2007) 1490-1495.

[27] S. Klomklao, S. Benjakul, W. Visessanguan, H. Kishimura, B.K. Simpson, Purification and characterization of trypsins from skipjack tuna (*Katsuwonus pelamis*) spleen, Food Chem. 100 (2007)1580-1589.

[28] S. Klomklao, S. Benjakul, W. Visessanguan, H. Kishimura, B.K. Simpson, Trypsin from the pyloric ceca of bluefish (*Pomatomus saltatrix*), Comp. Biochem. Physiol. B 148 (2007)

382-389.

[29] S. Klomklao, S. Benjakul, W. Visessanguan, H. Kishimura, B.K. Simpson, A 29 kDa protease from the digestive glands of Atlantic bonito (*Sarda sarda*): recovery and characterization, J. Agric. Food Chem. 55 (2007) 4548-4553.

[30] H. Kishimura, S. Klomklao, S. Benjakul, B.-S. Chun, Characteristics of trypsin from the pyloric ceca of walleye pollock (*Theragra chalcogramma*), Food Chem. 106 (2008) 194-199.

[31] T. Fuchise, H. Kishimura, H. Sekizaki, Y. Nonami, G. Kanno, S. Klomklao, S. Benjakul, B.-S. Chun, Purification and characteristics of cold-zone fish trypsin, Pacific cod (*Gadus macrocephalus*) and saffron cod (*Eleginus gracilis*), Food Chem. 116 (2009) 611-616.

[32] M.N. Ahsan, D. Funabara, S. Watabe, Molecular cloning and characterization of two isoforms of trypsinogen from anchovy pyloric ceca, Marine Biotech. 3 (2001) 80-90.

[33] M.E.E. Watson, Compilation of published signal sequences, Nuc. Acids Res. 12 (1984) 5145-5164.

[34] M.N. Louvard, A. Puigserver, On bovine and porcine anionic trypsinogens, Biochim. Biophys. Acta 371 (1974) 177-185.

[35] I.L. Huerou, C. Wicker, P. Guilloteau, R. Toullec, A. Puigserver, Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic trypsinogen: structural identity within the trypsin family, Eur. J. Biochem. 193 (1990) 767-773.

[36] M.M. Krem, T. Rose, E.D. Cera, The C-terminal sequence encodes function in serine proteases, J. Biol. Chem. 274 (1999) 28063-28066.

[37] D. Gable, V. Kasche, Autolysis of β-trypsin: influences of calcium ions and heat, Acta Chem. Scandinavia 27 (1973) 1971-1981.

[38] K.A. Walsh, Trypsinogens and trypsins of various species. Methods Enzymol. 19 (1970) 41-63.

[39] W. Bode, P. Schwager, The single calcium-binding site of crystalline β-trypsin, FEBS

LETTERS 56 (1975) 139-143.

[40] S. Genicot, F. Rentier-Delrue, D. Edwards, J. Vanbeeumen, C. Gerday, Trypsin and trypsinogen from Antarctic fish: molecular basis of cold adaptation, Biochim. Biophys. Acta 1298 (1996) 45-57.

[41] G. Kanno, T. Yamaguchi, H. Kishimura, E. Yamaha, H. Saeki, Purification and characteristics of trypsin from masu salmon (*Oncorhybchus masou*) cultured in fresh-water, Fish Physiol. Biochem. 36 (2009) 637-645.

[42] A. Gudmundsdottir, E. Gudmundsdottir, S. Oskarsson, J.B. Bjarnason, A.K. Eakin, C.S. Craik, Isolation and characterization of cDNAs from Atlantic cod encoding two different forms of trypsinogen, Eur. J. Biochem. 217 (1993) 1091-1097.

[43] R. Male, L.B. Lorens, A.O. Smalas, K.R. Torrissen, Molecular cloning and characterization of anionic and cationic variants of trypsin from Atlantic salmon, Eur. J. Biochem. 232 (1995) 677-685.

[44] B.S. Hartley, D.L. Kauffman, Corrections to the amino acid sequence of bovine chymotrypsinogen A, Biochem. J. 101 (1966) 229-231.

[45] M.C. Peitsch, Protein modeling by e-mail, Biol. Tech. 13 (1995) 658-660.

[46] M.C. Peitsch, ProMod and Swiss-Model: internet-based tools for automated comparative protein modeling, Biochem. Soc. Trans. 24 (1996) 274-279.

(Captions to figures)

Fig. 1. Nucleotide and deduced amino acid sequences of arabesque greenling trypsin. The deduced amino acid sequence and the residue numbers are shown above the codons. The ATG initiation codon and the TAA termination codon are asterisked. Putative polyadenylation signal is written in bold-italic.

Fig. 2. Comparison of the signal peptide of arabesque greenling trypsin with those of other fish and bovine trypsins. Arabesque greenling, arabesque greenling trypsin; Atlantic cod I, Atlantic cod, *Gadus morhua*, trypsin I [42]; Antarctic fish, Antarctic fish, *Paranotothenia magellanica*, trypsin [40]; Atlantic salmon I, Atlantic salmon, *Salmo salar*, trypsin I [43]; Anchovy I, anchovy, *Engraulis japonicus*, trypsin I [32]; Flounder I, flounder, *Paralichthys plivaceus*, trypsin I [Accession No.: AB029750 in GenBank]; Zebrafish, zebrafish, *Dnio rerio*, trypsin [Accession No.: AF541952 in DDBJ]; Tilapia, tilapia, *Oreochromis niloticus*, trypsin [Accession No.: AY510093 in DDBJ]; Bovine cationic, bovine cationic trypsin [Accession No.: BC134797 in DDBJ].

Fig. 3. Primary structures of the activation peptides of several fish and bovine trypsins. The names of sequences are the same in Fig. 2. Dashes indicate deletions introduced for maximizing the sequence similarity.

Fig. 4. Alignment of the deduced amino acid sequence of arabesque greenling mature trypsin with the sequences of other fish and bovine trypsins. The amino acids are numbered by the standard chymotrypsin numbering system [44]. The names of sequences are the same in Fig. 2. Dashes indicate deletions introduced for maximizing the sequence similarity. The residues of catalytic triad (His57, Asp102 and Ser195) and obligatory Asp189 are marked with asterisks.

Fig. 5. Electrostatic surface potentials of the predicted 3D-structure of AG-T. We built the 3D-structure of the AG-T on the basis of the primary structural information using 3-D JIGSAW server [http://bmm.cancerresearchuk.org/~3djigsaw/] and Swiss-PdbViewer DeepView v4.1 [47, 47], and the results were compared to that of bovine trypsin.

A: Calcium-binding region. Red color represents the negatively charged amino acid residues and amino acid residues which are concerned in a combination with calcium ion. Blue color represents the positively charged amino acid residues. Green color represents the other amino acid residues. A-1; AG-T, A-2; bovine trypsin [2].

B: N-terminal region. Red color represents charged amino acid residues. Blue color represents hydrophobic amino acid residues. Green color represents the other amino acid residues. B-1; AG-T, B-2; bovine trypsin [2].

Fig. 1

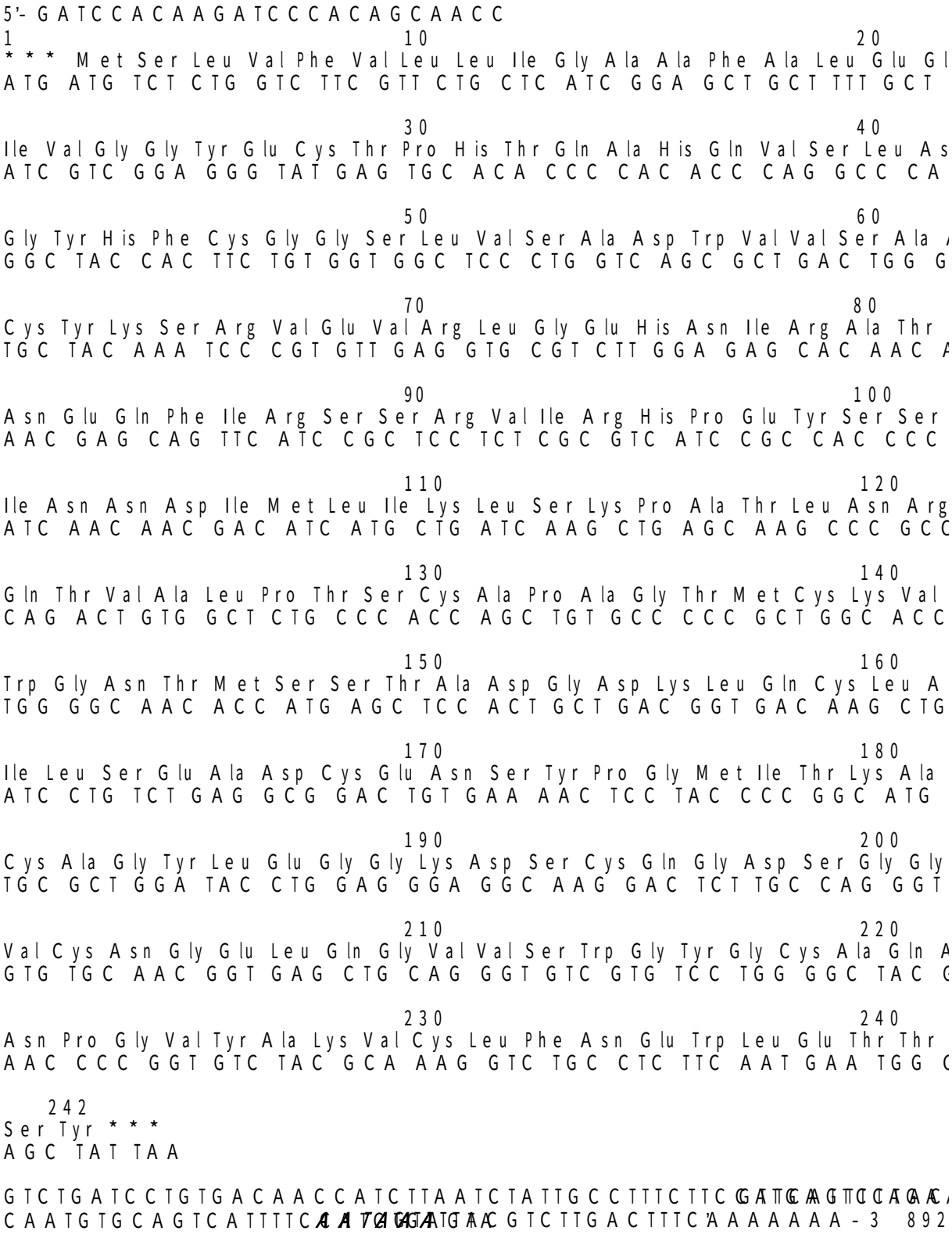

5'- G A T C C A C A A G A T C C C A C A G C A A C C
1 10 20
* * * Met Ser Leu Val Phe Val Leu Leu Ile Gly Ala Ala Phe Ala Leu Glu Gl
A T G A T G T C T C T G G T C T T C G T T C T G C T C A T C G G A G C T G C T T T G C T
30 40
Ile Val Gly Gly Tyr Glu Cys Thr Pro His Thr Gln Ala His Gln Val Ser Leu As
A T C G T C G G A G G G T A T G A G T G C A C A C C C C A C A C C C A G G C C C A
50 60
Gly Tyr His Phe Cys Gly Gly Ser Leu Val Ser Ala Asp Trp Val Val Ser Ala
G G C T A C C A C T T C T G T G G T G G C T C C C T G G T C A G C G C T G A C T G G G
70 80
Cys Tyr Lys Ser Arg Val Glu Val Arg Leu Gly Glu His Asn Ile Arg Ala Thr
T G C T A C A A A T C C C G T G T T G A G G T G C G T C T T G G A G A G C A C A A C A
90 100
Asn Glu Gln Phe Ile Arg Ser Ser Arg Val Ile Arg His Pro Glu Tyr Ser Ser
A A C G A G C A G T T C A T C C G C T C C T C T C G C G T C A T C C G C C A C C C C
110 120
Ile Asn Asn Asp Ile Met Leu Ile Lys Leu Ser Lys Pro Ala Thr Leu Asn Arg
A T C A A C A A C G A C A T C A T G C T G A T C A A G C T G A G C A A G C C C G C C
130 140
Gln Thr Val Ala Leu Pro Thr Ser Cys Ala Pro Ala Gly Thr Met Cys Lys Val
C A G A C T G T G G C T C T G C C C A C C A G C T G T G C C C C G C T G G C A C C
150 160
Trp Gly Asn Thr Met Ser Ser Thr Ala Asp Gly Asp Lys Leu Gln Cys Leu A
T G G G G C A A C A C C A T G A G C T C C A C T G C T G A C G G T G A C A A G C T G
170 180
Ile Leu Ser Glu Ala Asp Cys Glu Asn Ser Tyr Pro Gly Met Ile Thr Lys Ala
A T C C T G T C T G A G G C G G A C T G T G A A A A C T C C T A C C C C G G C A T G
190 200
Cys Ala Gly Tyr Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly
T G C G C T G G A T A C C T G G A G G G A G G C A A G G A C T C T T G C C A G G G T
210 220
Val Cys Asn Gly Glu Leu Gln Gly Val Val Ser Trp Gly Tyr Gly Cys Ala Gln A
G T G T G C A A C G G T G A G C T G C A G G G T G T C G T G T C C T G G G G C T A C G
230 240
Asn Pro Gly Val Tyr Ala Lys Val Cys Leu Phe Asn Glu Trp Leu Glu Thr Thr
A A C C C C G G T G T C T A C G C A A A G G T C T G C C T C T T C A A T G A A T G G C
242
Ser Tyr * * *
A G C T A T T A A
G T C T G A T C C T G T G A C A A C C A T C T T A A T C T A T T G C C T T T C T T C G A T G A G T C T T A A
C A A T G T G C A G T C A T T T T C A T A T G T A A C G T C T T G A C T T T C A A A A A A A A - 3 8 9 2

Fig. 2

	1				5				10				15	
Arabesque greenling	Met	Met	Ser	Leu	Val	Phe	Val	Leu	Leu	Ile	Gly	Ala	Al	
Atlantic cod I	Met	Lys	Ser	Leu	Ile	Phe	Val	Leu	Leu	Leu	Gly	Ala	Va	
Antarctic fish	Met	Arg	Ser	Leu	Val	Phe	Val	Leu	Leu	Ile	Gly	Ala	Al	
Atlantic salmon I	Met	Ile	Ser	Leu	Val	Phe	Val	Leu	Leu	Ile	Gly	Ala	Ala	
Anchovy I	Met	Arg	Pro	Leu	Val	Phe	Leu	Val	Leu	Leu	Gly	Ala	Al	
Flounder II	Met	Arg	Ser	Leu	Val	Phe	Val	Leu	Leu	Ile	Gly	Ala	Al	
Zebrafish	Met	Lys	Ala	Phe	Ile	Leu	Leu	Ala	Leu	Phe	Ala	Val	Al	
Tilapia	Met	Lys	Tyr	Phe	Ile	Leu	Leu	Ala	Leu	Phe	Ala	Ala	Al	
Bovine cationic	Met	Lys	Thr	Phe	Ilu	Phe	Leu	Ala	Leu	Leu	Gly	Ala	Al	

Fig. 3

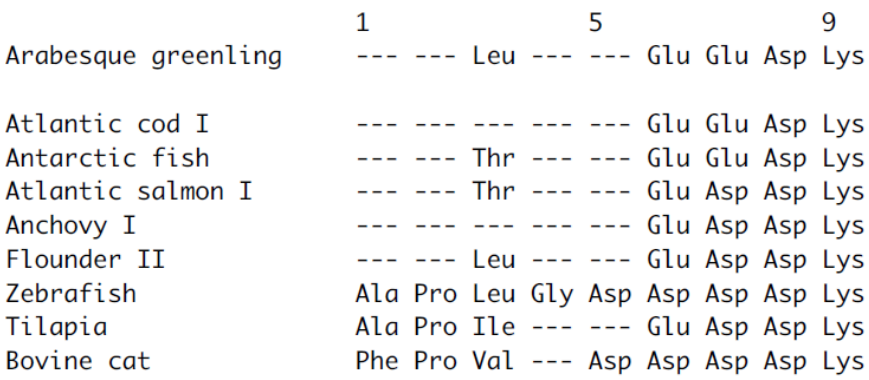

1 5 9
Arabesque greenling --- --- Leu --- --- Glu Glu Asp Lys
Atlantic cod I --- --- --- --- --- Glu Glu Asp Lys
Antarctic fish --- --- Thr --- --- Glu Glu Asp Lys
Atlantic salmon I --- --- Thr --- --- Glu Asp Asp Lys
Anchovy I --- --- --- --- --- Glu Asp Asp Lys
Flounder II --- --- Leu --- --- Glu Asp Asp Lys
Zebrafish Ala Pro Leu Gly Asp Asp Asp Asp Lys
Tilapia Ala Pro Ile --- --- Glu Asp Asp Lys
Bovine cat Phe Pro Val --- Asp Asp Asp Asp Lys

Fig. 4

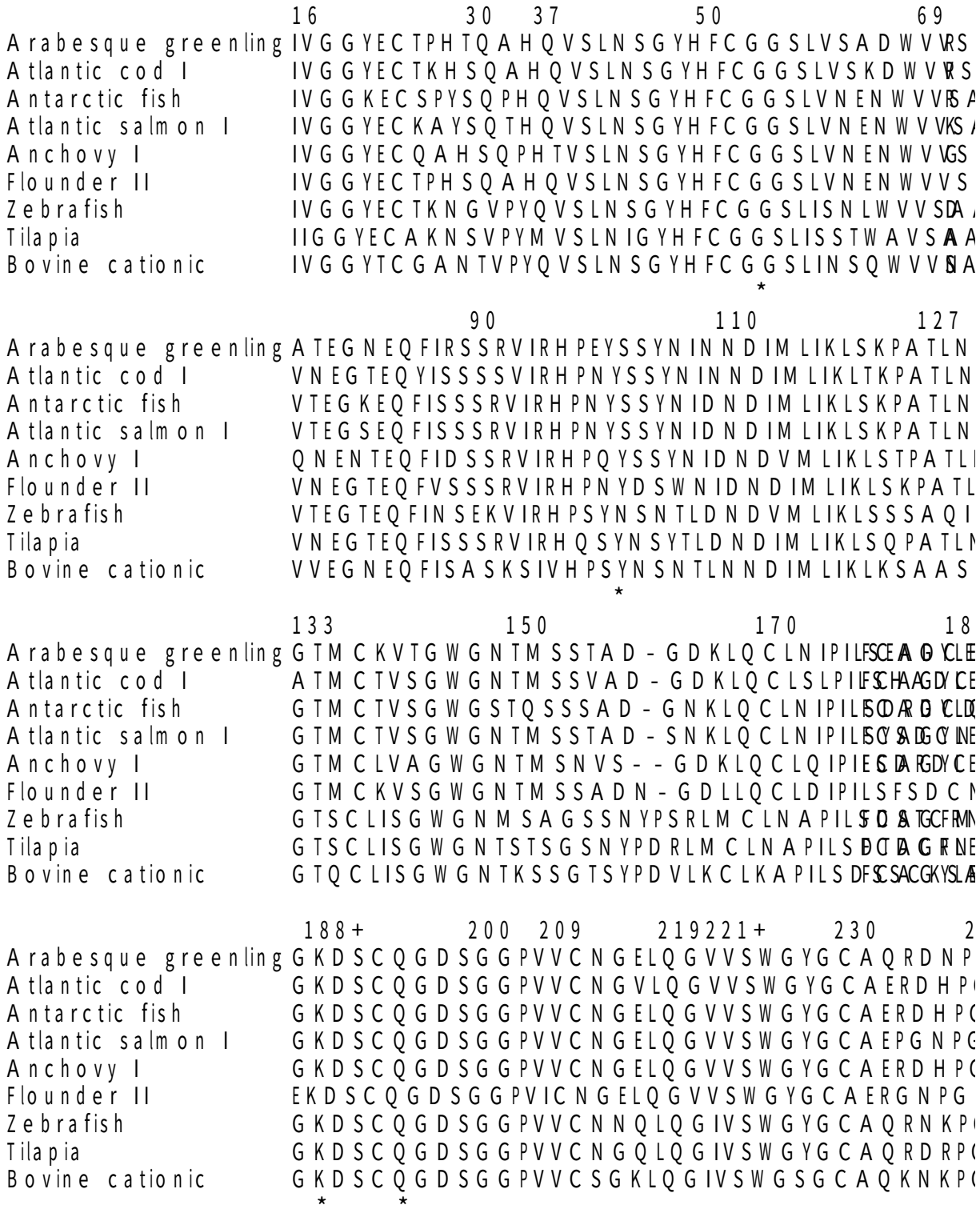
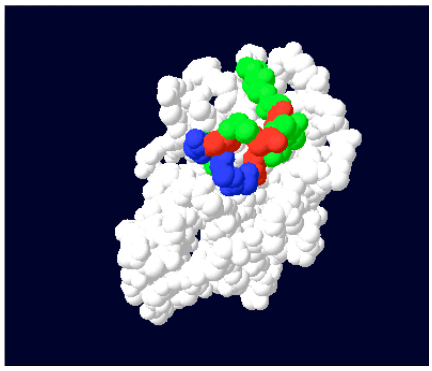
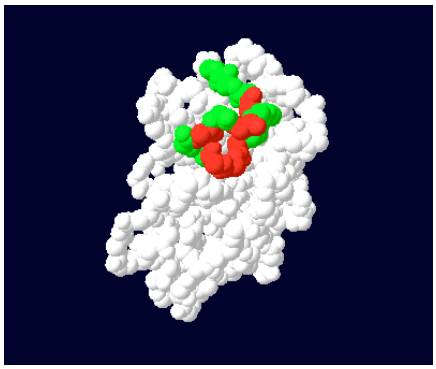

16 30 37 50 69
Arabesque greenling IVGGYECTPHTQAHQVSLNSGYHFCGGSLVSADWVVRS
Atlantic cod I IVGGYECTKHSQAHQVSLNSGYHFCGGSLVSKDWVVRS
Antarctic fish IVGGKECSPYSQPHQVSLNSGYHFCGGSLVNENWVVRS
Atlantic salmon I IVGGYECKAYSQTHQVSLNSGYHFCGGSLVNENWVVKS
Anchovy I IVGGYECQAHSQPHTVSLNSGYHFCGGSLVNENWVVGS
Flounder II IVGGYECTPHSQAHQVSLNSGYHFCGGSLVNENWVVS
Zebrafish IVGGYECTKNGVPYQVSLNSGYHFCGGSLISNLWVVSD
Tilapia IIGGYECAKNSVPYMVSLNIGYHFCGGSLISSTWAVSA
Bovine cationic IVGGYTCGANTVPYQVSLNSGYHFCGGSLINSQWVVS
90 110 127
Arabesque greenling ATEGNEQFIRSSRVIRHPEYSSYNINNDIMLIKLSKPATLN
Atlantic cod I VNEGTEQYISSSSVIRHPNYSSYNINNDIMLIKLTKPATLN
Antarctic fish VTEGKEQFISSSRVIRHPNYSSYNIDNDIMLIKLSKPATLN
Atlantic salmon I VTEGSEQFISSSRVIRHPNYSSYNIDNDIMLIKLSKPATLN
Anchovy I QNENTEQFIDSSRVIRHPQYSSYNIDNDVMLIKLSTPATL
Flounder II VNEGTEQFVSSSRVIRHPNYDSWNIDNDIMLIKLSKPATL
Zebrafish VTEGTEQFINSEKVIRHPSYNSNTLDNDVMLIKLSSSAQI
Tilapia VNEGTEQFISSSRVIRHQSYNSYTLDNDIMLIKLSQPATLN
Bovine cationic VVEGNEQFISASKSIVHPSYNSNTLNNDIMLIKLKSAAS
133 150 170 18
Arabesque greenling GTMCKVTGWGNTMSSTAD-GDKLQCLNIPILFSCEAQ
Atlantic cod I ATMCTVSGWGNTMSSVAD-GDKLQCLSLPILFSCHAAGDY
Antarctic fish GTMCTVSGWGSTQSSSAD-GNKLQCLNIPILSCDARD
Atlantic salmon I GTMCTVSGWGNTMSSTAD-SNKLQCLNIPILSCYADG
Anchovy I GTMCLVAGWGNTMSNVS--GDKLQCLQIPIESCDARGDY
Flounder II GTMCKVSGWGNTMSSADN-GDLLQCLDIPILSFSDC
Zebrafish GTSCLISGWGNMSAGSSNYPSRLMCLNAPILSDSATGCFM
Tilapia GTSCLISGWGNTSTSGSNYPDRLMCLNAPILSDSTAGCFN
Bovine cationic GTQCLISGWGNTKSSGTSYPDVLKCLKAPILSDFSCSACGKYSL
188+ 200 209 219221+ 230 2
Arabesque greenling GKDSCQGDSGGPVVCNGELQGVVSWGYGCAQRDNP
Atlantic cod I GKDSCQGDSGGPVVCNGVLQGVVSWGYGCAERDHP
Antarctic fish GKDSCQGDSGGPVVCNGELQGVVSWGYGCAERDHP
Atlantic salmon I GKDSCQGDSGGPVVCNGELQGVVSWGYGCAEPGNP
Anchovy I GKDSCQGDSGGPVVCNGELQGVVSWGYGCAERDHP
Flounder II EKDSCQGDSGGPVICNGELQGVVSWGYGCAERGNPG
Zebrafish GKDSCQGDSGGPVVCNNQLQGIVSWGYGCAQRNKP
Tilapia GKDSCQGDSGGPVVCNGQLQGIVSWGYGCAQRDRP
Bovine cationic GKDSCQGDSGGPVVCSGKLQGIVSWGSGCAQKNKP

Fig. 5

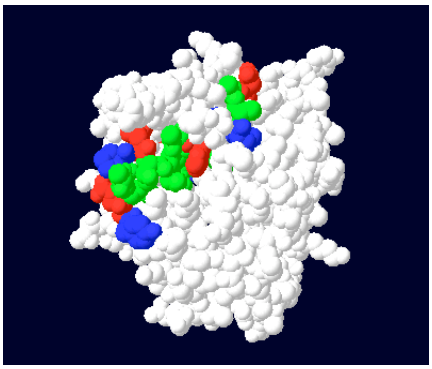
A-1



A-2



B-1



B-2