Developmental and Comparative Immunology 148 (2023) 104920

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/devcompimm

Two granzyme A/K homologs in *Zebra mbuna* have different specificities, one classical tryptase and one with chymase activity

Erdem Aybay [a], Mamoun Elkhalifa [a], Srinivas Akula [a,b], Sara Wernersson [b], Lars Hellman [a,*]

[a] Department of Cell and Molecular Biology, Uppsala University, The Biomedical Center, Box 596, SE-751 24, Uppsala, Sweden
[b] Department of Anatomy, Physiology and Biochemistry, Swedish University of Agricultural Sciences, Box 7011, SE-75007, Uppsala, Sweden

ARTICLE INFO

Keywords:
Cytotoxic T cells
NK cell
Granzyme A
Apoptosis
Caspase
Chymase

ABSTRACT

Granzymes A and K are two highly homologous serine proteases expressed by mammalian cytotoxic T cells (CTLs) and natural killer (NK) cells. The locus encoding these two proteases is the first of the hematopoietic serine protease loci to appear during vertebrate evolution. This locus is found in all jawed vertebrates including the cartilaginous fishes. Granzyme A is the most abundant of the different granzymes expressed by CTLs and NK cells and its potential function has been studied extensively for many years. However, no clear conclusions concerning its primary role in the immune defense has been obtained. In all mammals, there are only one copy each of granzyme A and K, whereas additional copies are found in both cartilaginous and ray finned fishes. In cichlids two of these copies seem to encode new members of the granzyme A/K family. These two new members appear to have changed primary specificity and to be pure chymases based on the amino acids in their active site substrate binding pockets. Interestingly, one of these gene copies is located in the middle of the granzyme A/K locus, while the other copy is present in another locus, the met-ase locus. We here present a detailed characterization of the extended cleavage specificity of one of these non-classical granzymes, a *Zebra mbuna* granzyme positioned in the granzyme A/K locus. This enzyme, named granzyme A2, showed a high preference for tyrosine in the P1 position of substrates, thereby being a strict chymase. We have also characterized one of the classical granzyme A/Ks of the *Zebra mbuna*, granzyme A1, which is a tryptase with preference for arginine in the P1 position of substrates. Based on their extended specificities, the two granzymes showed major similarities, but also some differences in preferred amino acids in positions surrounding the cleavable amino acid. Fish lack one of the hematopoietic serine protease loci of mammals, the chymase locus, where one of the major mast cell enzymes is located. An interesting question is now if cichlids have by compensatory mechanisms generated a mast cell chymase from another locus, and if similar chymotryptic enzymes have appeared also in other fish species.

1. Introduction

Several of the hematopoietic cells store massive amounts of serine proteases within their cytoplasmic granules for rapid release upon activation. These hematopoietic serine proteases do all belong to the large family of chymotrypsin-related serine proteases (Akula et al., 2015). These proteases, which are stored in their active form within granules, are expressed primarily by mast cells, basophils, neutrophils, cytotoxic T cells (CTLs) and natural killer cells (NK-cells), but not in B-cells, macrophages, dendritic cells and very small amounts in eosinophils. Several of the chymotrypsin-related serine proteases are expressed by activated CTLs and NK cells where they take part in the killing of virus infected cells. Due to their granule storage they have been named granzymes. Human CTLs and NK cells express five such granzymes, granzymes A, K, B, H and M (Hellman and Thorpe, 2014), of which granzyme B is the most well characterized enzyme (Heibein et al., 1999; MacDonald et al., 1999; Metkar et al., 2008). Granzyme B is one of the key components in the killing of cells infected by intracellular parasites and it initiates both caspase dependent and independent activation of cellular apoptosis mechanisms (Heibein et al., 1999; MacDonald et al., 1999; Metkar et al., 2008). The entry of granzyme B into a target cell during the fusion of granules with the target cell membrane is dependent of perforin, a pore-forming molecule related in structure to the final component in the membrane attack complex of the complement

Abbreviations: Gzm, granzyme; CTL, cytotoxic T cell; NK-cell, Natural killer cell.

* Corresponding author.
E-mail address: Lars.Hellman@icm.uu.se (L. Hellman).

https://doi.org/10.1016/j.dci.2023.104920
Received 30 June 2023; Received in revised form 8 August 2023; Accepted 15 August 2023
Available online 18 August 2023

cascade, the C9 component (Shinkai et al., 1988; Kondos et al., 2010). The function of the other human granzymes is less well characterized. Both human granzymes A and K are tryptases with preference for arginine (Arg) in the P1 position of a substrate. However, they differ markedly in their extended specificity (Aybay et al., 2023). The difference is primarily located N-terminally of the cleavage site, where granzyme A prefers smaller amino acids like glycine (Gly) and valine (Val), whereas granzyme K prefers large aromatic amino acids in the P2, P3 and P4 positions of a substrate (Aybay et al., 2023).

Out of the different loci encoding hematopoietic serine proteases, the locus encoding granzymes A and K was the first to appear during vertebrate evolution (Akula et al., 2015). This locus is found in all jawed vertebrate species from cartilaginous fishes to humans, and has thereby been around for more than 450 million years of vertebrate evolution (Akula et al., 2015). However, neither this nor any of the other loci encoding hematopoietic serine proteases are found in lamprey and hagfishes, two jawless fishes, indicating that they first appear with the jawed vertebrates. We have started to investigate the presence and specificity of enzymes related to the mammalian hematopoietic serine proteases including the granzyme A/K locus in different fish species, to obtain more detailed information on the evolution of these enzymes and their possibly conserved functions in fish immunity (Wernersson et al., 2006; Thorpe et al., 2016). We here present the analysis of two granzyme A/K related enzymes from a cichlid, the *Zebra mbuna*. One of these enzymes was found to be a classical tryptase with primary specificity for arginine and the second enzyme had changed specificity to now being a tyrosine specific chymase.

2. Materials and methods

2.1. Phylogenetic analyses

The phylogenetic analysis was performed essentially as described in a previous publication using the same strategy and sequences (Akula et al., 2015).

2.2. Production and purification of recombinant Zebra mbuna granzymes A2 and A1

The sequences of *Zebra mbuna* granzyme A1 and A2 were retrieved from the NCBI database under the accession numbers GzmA1 (XM_004546954) (with DGG triplet) and GzmA2 (XM_004546953) (with SGA triplet). The sequences were ordered as designer genes from GenScript (Piscataway, NJ, USA) and inserted into the mammalian expression vector pCEP-Pu2, followed by sequence verification. The construct contains a signal sequence, an N-terminal His_6-tag and an enterokinase site (Signal sequence-HHHHHHDDDDK-active protease). The His_6-tag and the enterokinase site were inserted for purification and activation purposes. The vector was transfected into the human embryonic kidney cell line HEK293-EBNA for expression of the recombinant enzymes. After purification on IMAC Ni^{2+} agarose (Qiagen, Hilden Germany), the enzymes were activated by the addition of 1 μl enterokinase (Roche, Mannheim, Germany) into 90 μl recombinant protein. The samples were mixed, followed by a 37 °C incubation for 5 h to activate the proteases. The purity and activation were determined by separation on 4–12% pre-cast SDS-PAGE gels (Invitrogen, Carlsbad, CA, USA). 2.5 μl of 4 x sample buffer, containing sodium dodecyl sulfate (SDS), and 0.5 μl β-mercaptoethanol were added to 10 μl protein followed by 85 °C heating and SDS-PAGE gel electrophoresis. Overnight staining of gels in colloidal Coomassie staining solution were followed by de-staining by 30% methanol in water and several washes in water to enable the visualization of the protein bands.

2.3. Determination of cleavage specificity by phage display

A library of T7 phages was designed to contain random nonameric peptides followed by a His_6-tag at the C-terminus of the capsid protein. This library, containing approximately 50 million independent phages, was used to determine the cleavage specificities of the two *Zebra mbuna* enzymes. Two hundred and fifty μl nickel–nitriloacetic acid (Ni–NTA) beads was used to immobilize the phages based on the interaction with His_6-tags by mixing and gentle agitation at 4 °C for 1 h. One and a half ml of washing buffer (1M NaCl, 0.1% Tween-20 in PBS, pH 7.2) was then added to remove unbound phages. This was repeated 10 times to ensure proper washing. The beads were then washed two times in 1.5 ml PBS and suspended in PBS in a total volume of 500 μl. Approximately 100 ng of recombinant protease was then added to the Eppendorf tubes followed by gentle agitation at 37 °C for approximately 16 h (overnight). One tube without protease was set as control. Phages containing peptide sequences cleavable by the protease were released from Ni–NTA beads and by centrifugation, the released phages were separated from the still uncleaved and thereby bead-bound phages, enabling the recovery of the cleaved phages from the supernatant. Fifteen μl Ni–NTA beads were added to the supernatant in order to remove phages that still had His tags but had been released by other means than cleavage ensuring that the His_6-tags had been removed on all the phages in the supernatant. One hundred μl of 100 mM imidazole was then used to elute the phages from the remaining Ni–NTA beads to determine the number of phages initially bound. In order to determine the number of phages detached from the Ni–NTA beads by the protease, a series of dilutions of the supernatant were plated onto LA-Amp plates together with 3 ml of 0.6% top agarose, 300 μl of *Escherichia coli* (BLT 5615) and 100 μl of 100 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Ten ml of *Escherichia coli* (BLT 5615) was prepared to amplify the remaining phages for the next round of selection. To this culture, 100 μl of a 100 mM solution of IPTG was added to induce the expression of T7 phage capsid protein. The bacteria were lysed after approximately 2 h of gentle agitation at 37 °C, and the lysate was centrifuged to remove cell debris before storage at 4 °C until the next round of selection. After 5–7 rounds of selection, 100 plaques were picked by glass Pasteur pipettes from the LA-Amp plates and the plugs were transferred into Eppendorf tubes containing phage extraction buffer (100 mM NaCl and 6 mM $MgSO_4$ in 20 mM Tris-HCl, pH 8.0). Vigorous shaking for 30 min at 4 °C was used for extracting the phages from the agarose. Polymerase chain reaction (PCR) was then used to amplify the region of phage DNA encoding the random nonamer region. The quality and quantity of the amplified DNA was determined by gel electrophoresis and the 96 samples with best DNA quality were sent in a microtiter plate for sequencing to Eurofins (Ebersberg, Germany).

2.4. Generation of a consensus sequence from sequenced phage inserts

The results from the sequence analysis were translated into amino acid sequences by CLC Sequence viewer 8. The amino acid sequences were then aligned by hand looking for sequence patterns and based on the results of previous studies of similar enzymes. Amino acids were classified based on the particular characteristics as follows: aromatic amino acid (Phe, Tyr, Trp); negatively charged amino acid (Asp, Glu); positively charged amino acid (Lys, Arg); small aliphatic amino acid (Gly, Ala); larger aliphatic amino acid (Val, Leu, Ile, Pro) and other amino acid (Ser, Thr, His, Asn, Gln, Cys, Met). The peptide bond between P1 and P1′position represents the scissile bond, based on the nomenclature by Schechter and Berger to designate the substrate cleavage region (Schechter and Berger, 1967).

The sequences form the phage display were also added to a file in the Web-logo program to generate a graphical presentation of the selectivity of the two *Zebra mbuna* enzymes (https://weblogo.berkeley.edu/logo.cgi).

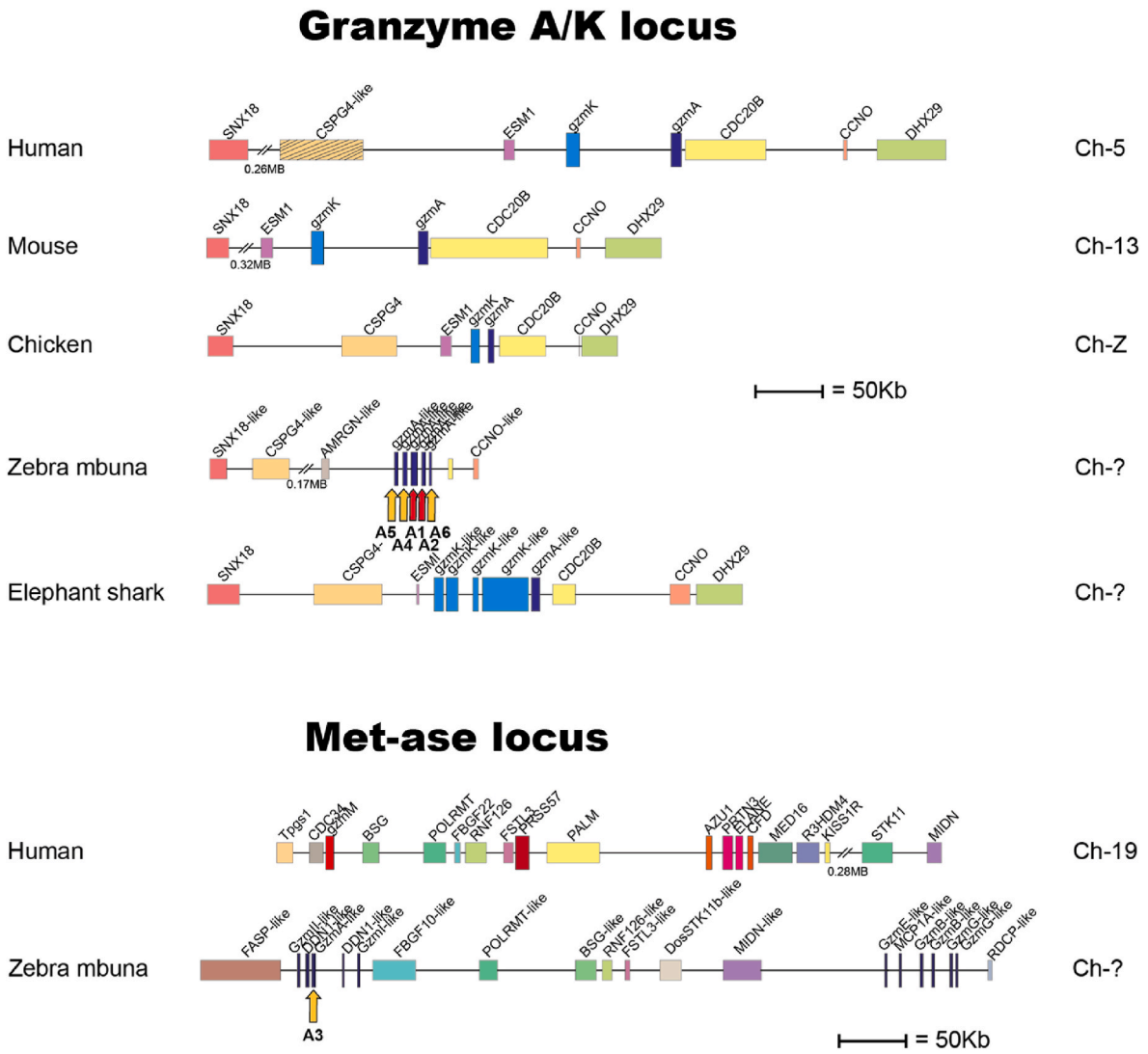

Fig. 1. The gene loci encoding granzymes A and K. In almost all jawed vertebrates the granzymes A and K are encoded from the granzyme A/K locus, also named the T cell tryptase locus. However, in a few fish species, primarily cichlids, a closely related gene is also present in the met-ase locus, a locus encoding other hematopoietic serine proteases, including granzyme M (GzmM), complement factor D (CFD), neutrophil elastase (ELANE), proteinase 3 (PRTN3), the inactive azurocidin (AZU1) and neutrophil serine protease 4 (NSP-4), which in the gene map is named PRSS57. Genes are colored according to the classification. Granzyme A-related genes are dark blue and the granzyme K-related genes are blue. The serine protease genes within the human met-ase locus are shown in orange or red colors. The protease genes are in double hight compared to the surrounding non-protease genes for an easier identification. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

2.5. Generation of recombinant substrates for the analysis of the cleavage specificity

To verify the result from the phage display we used a new type of recombinant substrates. The vector pET21 was modified to carry two copies of the *E. coli* thioredoxin (Trx) gene for bacterial expression. A His_6-tag was inserted in the C-terminal end of the second copy of the Trx for purification with Ni–NTA beads. Substrate sequences were inserted between the two thioredoxin molecules by the use of two complementary oligonucleotides. The oligonucleotides were ordered to match the sticky ends of two unique restriction sites (BamHI and SalI) followed by ligation and cloning. Individual clones were picked and sequenced to ensure the correct linker sequence between the two Trx molecules inserted by the oligos. The verified plasmids were then transformed into the *E. coli Rosetta Gami* strain for protein expression (Novagen, Merck, Darmstadt, Germany). The transformed strains were grown overnight. In the morning, 10 ml of the culture was added to 90 ml of LB-Amp in a new E-flask and incubated at 37 °C for 1 h until the OD (600 nm) reached 0.5. The culture was then induced by the addition of 1 ml IPTG to a final concentration of 1 mM and incubated at 37 °C for an additional 3 h under vigorous shaking. The bacteria were centrifuged at 3000 rpm for 12 min at 4 °C, resuspended in 25 ml PBS containing 0.05% Tween followed by an additional centrifugation whereafter the pellet was dissolved in 2 ml PBS. Sonication was then used to open the cells to release their cellular contents. Cooling on ice after every 30 s of sonication was important for avoiding overheating. Sonication 5 times for 30 s each time were performed to ensure efficient breaking of the cell wall. The debris was removed by a centrifugation at 13,000 rpm for 3 min under 4 °C and the supernatant was transferred into new tubes. 250 µl Ni–NTA slurry (50:50; Qiagen, Hilden, Germany) was added and the samples were incubated for 45 min in cold room under slow rotation. The Ni–NTA beads were then transferred into a 2 ml syringe with a glass filter in the bottom of the syringe. The beads were washed 3 times with 1 ml, 2 ml and 2 ml of washing buffer (PBS, 0.05% Tween 20, 10 mM Imidazole and 1 M NaCl) and the protein was then eluted in six fractions with 100 µl (for the first fraction) and 200 µl (for the following fractions)

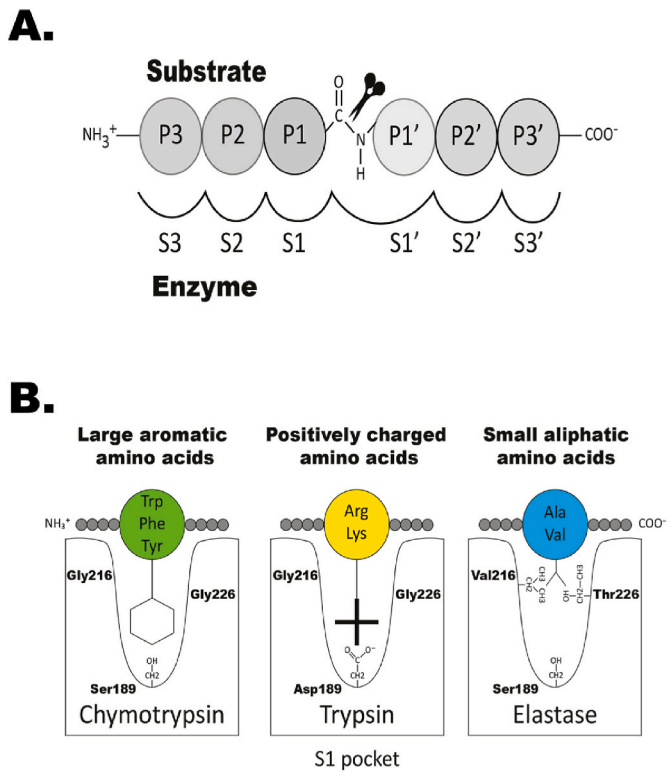

Fig. 2. Nomenclature of the amino acids surrounding the cleavage site and the amino acids forming the active site pocket. In panel **A,** we show the nomenclature of the amino acids surrounding the cleaved peptide bond. The amino acids N-terminal from the cleaved bond are termed P1 (where cleavage occurs, depicted by scissors), P2, P3 etc. Amino acids C-terminal of the cleaved bond are termed P1' (adjacent to P1), P2′, P3′ etc. The corresponding interacting subsites in the enzyme are denoted with S. In panel **B**, we show the three amino acids forming the active site pocket (S1 pocket). These three residues correspond to positions 189, 216 and 226 in bovine pancreatic chymotrypsinogen and have been found to determining the primary specificity of the enzyme as either chymotrypsin-, trypsin- or elastase-like specificity (Schechter and Berger, 1967). The preferred amino acids in the P1 position of the corresponding substrates are illustrated in green, yellow and blue, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

of elution buffer (PBS, 0.05% Tween 20, 100 mM imidazole). Each eluted fraction was analyzed on SDS bis-Tris 4–12% PAGE gel by mixing 10 μl of protein sample with 2.5 μl 4x sample buffer and 0.5 μl β-mercaptoethanol followed by heating for 8 min at 85 °C. The protein concentration for each fraction was estimated by comparing with a control sample for which the concentration was already known. The fractions were diluted into similar concentration as the control sample and 10 μl of the dilutions were mixed with 2 μl of the *Zebra mbuna* enzyme (based on the activity of the enzyme) leaving at room temperature for the reaction at time points 0, 15, 45 and 150 min respectively. Sample buffer was used to stop the reaction and 0.5 μl of β-mercaptoethanol was then added to each sample before heating for 8 min at 85 °C. The results from the cleavage reactions were analyzed on 4–12% pre-cast SDS-PAGE gels (Invitrogen) and, according to previous description, the gels were stained overnight in colloidal Coomassie staining solution followed by de-staining in 30% methanol in water one time and for several times in water.

3. Results

3.1. Genomic and phylogenetic analysis of granzyme A/K like enzymes from the Zebra mbuna, a cichlid

From an initial screening of different fish species concerning granzyme A/K genes cichlids turned out to be the most interesting species. In the *Zebra mbuna*, we found several genes encoding classical enzymes located in the granzyme A/K locus. However, a related enzyme is also present in a locus encoding other hematopoietic serine proteases, the met-ase locus. This is a locus harboring several of the neutrophil proteases in mammals, including the neutrophil elastase, proteinase 3, the inactive azurocidin and neutrophil serine protease 4 (NSP-4) (Fig. 1). The nomenclature of the cleavage site and the corresponding region of the enzyme is shown in Fig. 2A and the three amino acids forming the active site pocket is shown in Fig. 2B.

To look closer at the sequence similarity between the different human and *Zebra mbuna* granzyme A/K enzymes we performed a sequence alignment (Fig. 3). The six *Zebra mbuna* enzymes are more closely related to each other than to the two human enzymes, and human granzyme K is the enzyme showing the lowest degree of sequence identity to the other enzymes (Fig. 3Aand C). In Fig. 3B there is a separate alignment of only the *Zebra mbuna* enzymes to focus on the difference between these six enzymes. We can from the alignment see that all six *Zebra mbuna* enzymes show more than 70% identity and that granzymes A4 and A5 show more than 90% identity (Fig. 3B and C).

By looking at the three amino acid residues forming the substrate binding cleft, corresponding to the positions 189, 216 and 226 in bovine chymotrypsinogen, we found that two of the six granzyme A/K genes in the cichlid *Zebra mbuna* had a triplet very different from the other four members (Fig. 4B). All the other members had the classical DGG triplet (aspartic acid, glycin, glycin) indicating tryptase specificity similar to all the mammalian granzyme A/K genes (Fig. 4B) (Schechter and Berger, 1967). The two members with a different triplet were one of the genes in the classical A/K locus and the related gene found in the met-ase locus. Both of these genes had SGA triplets (serine, glycine and alanine), which indicated chymotryptic primary specificity of these enzymes (Figs. 2B and 4B). To look closer at both primary and extended cleavage specificities of the cichlid enzymes we therefore decided to determine the specificity of *Zebra mbuna* granzyme A1, one of the classical granzyme A/K members with a DGG triplet, and granzyme A2, one of the members with a SGA triplet, both encoded from the classical granzyme A/K locus.

3.2. Purification and activation of recombinant Zebra mbuna granzymes A2 and A1

The *Zebra mbuna* granzymes A1 and A2 were produced as recombinant enzymes in the mammalian cell line HEK293-EBNA. These recombinant enzymes contain an N-terminal His_6-tag to enable purification from the conditioned media by the use of Ni-NTA agarose and an enterokinase site to be able to remove the His tag and obtain an enzymatically active protease. Following purification, both enzymes were activated by the cleavage with enterokinase, resulting in a reduction by 1.5–2 kDa in size of the protein as confirmed by SDS-PAGE analysis (Fig. 5).

3.3. Determination of the extended cleavage specificity by substrate phage display

A library of T7 phages was used to determine the extended cleavage specificities of the two *Zebra mbuna* enzymes. The phage library used contains approximately 50 million individual phage clones and each clone expresses a unique random sequence of 9 amino acids. Selection of the cleavage by the *Zebra mbuna* granzyme A2 was performed for 7 rounds and the result showed approximately 620 times more phages compared to the PBS negative control. The phage display of *Zebra mbuna*

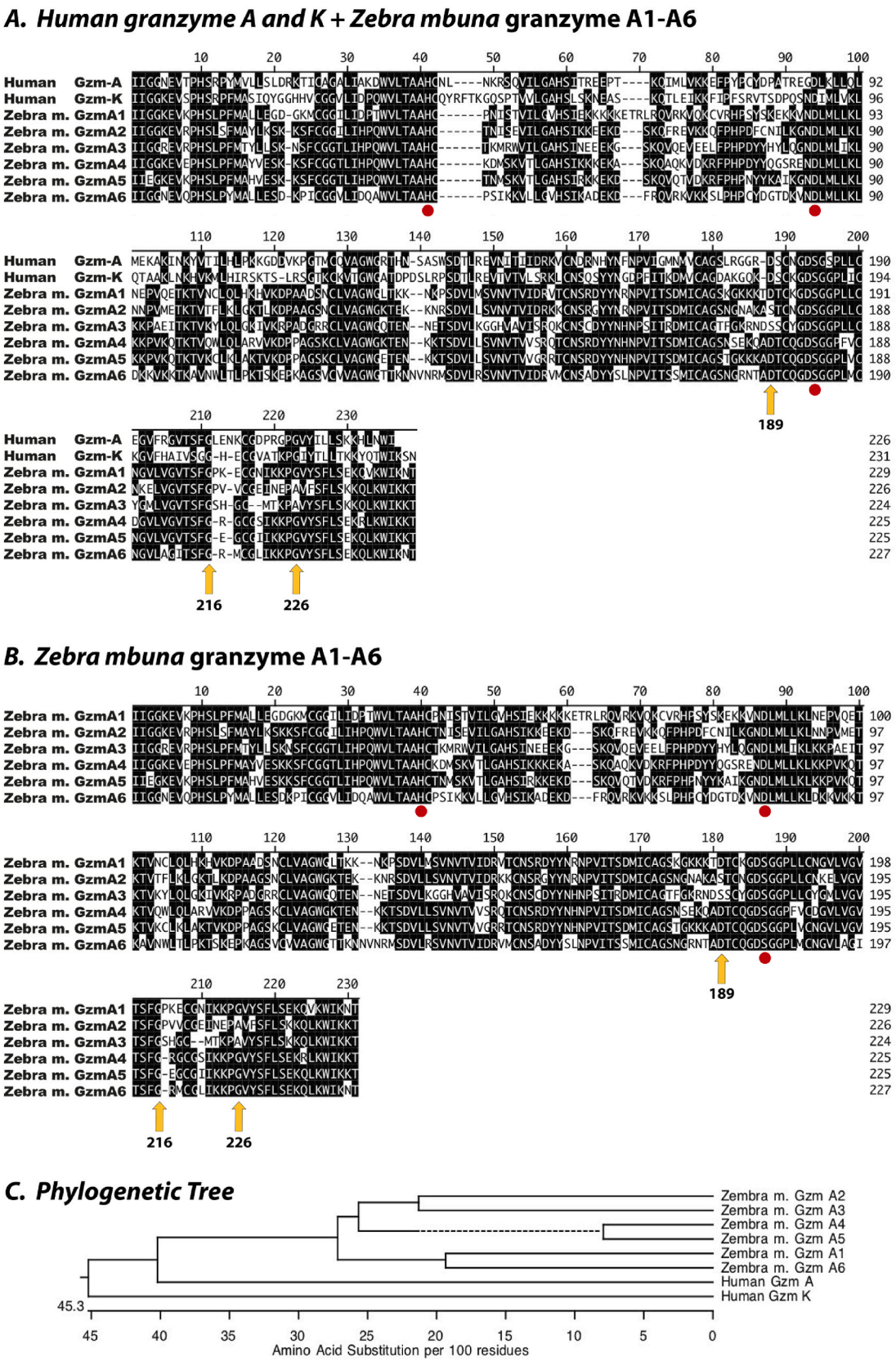

Fig. 3. Sequence alignment of *Zebra mbuna* and human granzymes A and K. DNA Star Megalign program and Clustal W algorithm were used to create the alignments (panel **A** and **B**) and phylogenetic tree (panel **C**) of the *Zebra mbuna* granzymes. Human granzymes A and K were used as reference granzymes in panel **A** and in the phylogenetic tree in panel **C**. The positions of the three residues of the catalytic triad His-Asp-Ser are marked by red dots. The three amino acids which are predicted as the substrate pocket in number 189, 216 and 226 are indicated by yellow arrows in panels **A** and **B**. Gzm: granzyme. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

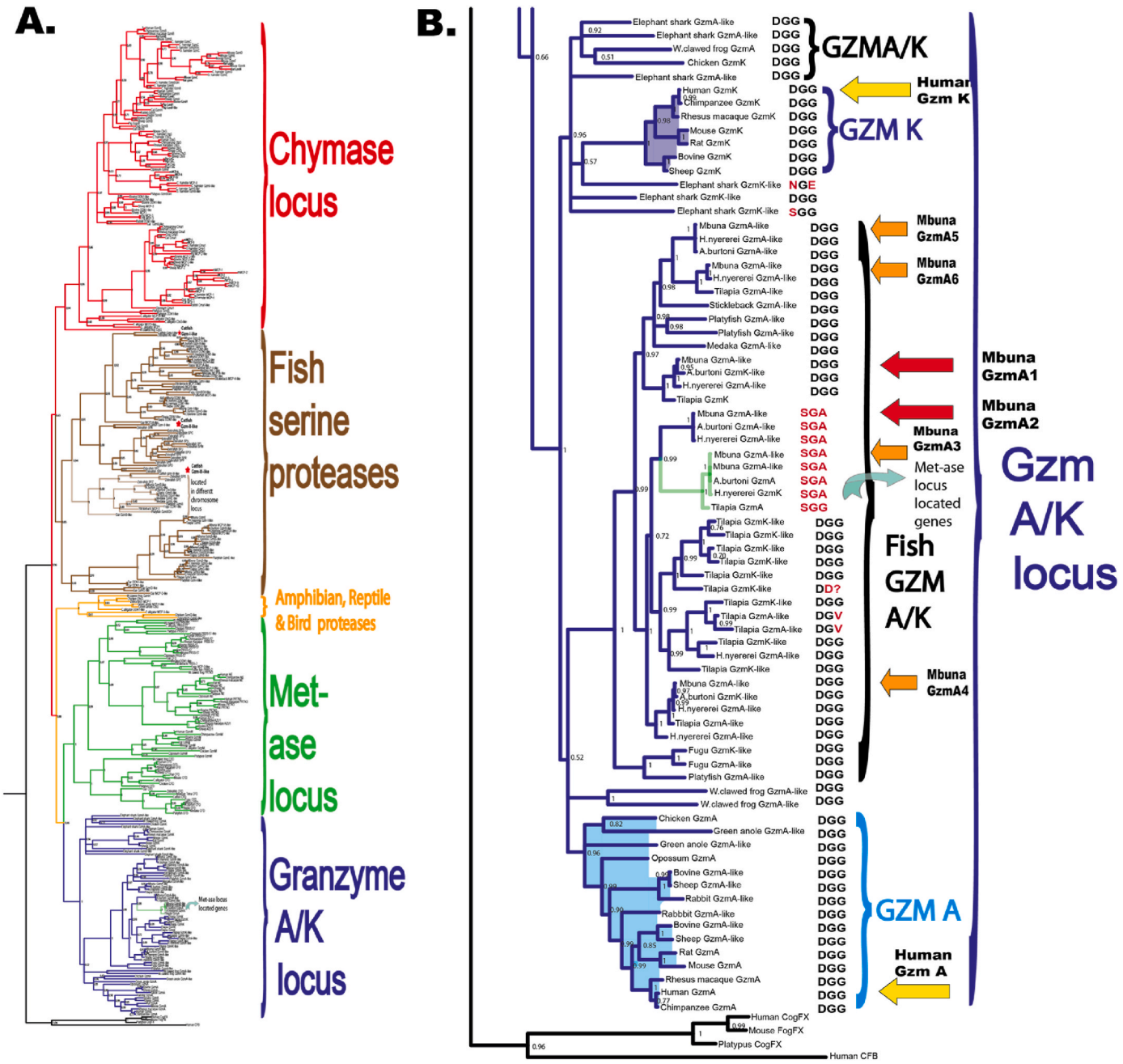

Fig. 4. A phylogenetic tree of selected granzyme A/K locus genes. A phylogenetic tree was constructed based on the relationships among different hematopoietic serine protease genes located in five different loci of a number of vertebrates. The phylogenetic tree was constructed using both MrBase analysis program and Maximum-likelihood algorithm using the manual standard protocol (Ronquist et al., 2012) and was drawn in FigTree 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). In panel **A,** we show the phylogenetic tree of genes from all five loci forming individual branches for the chymase locus (red), the met-ase locus (green), the major fish serine protease locus (brown), a locus found in amphibians, reptiles and birds (orange) and the granzyme A/K locus (dark blue). In panel **B,** we show the granzyme A/K locus including the enzymes of major interest for this study, with those analyzed by phage display marked by red arrows. The other *Zebra mbuna* granzyme A/K genes are marked by orange arrows and human granzymes A and K by yellow arrows. Note also the small subfamily of cichlid granzyme A/K genes that have become chymases and have the triplet SGA (red). Five of these genes, representing different cichlid species, are present in their met-ase loci (light green branches and light blue arrow). Gzm: granzyme. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

granzyme A1 was performed over 6 rounds and the result showed approximately 100 times more phages compared to the PBS control. One hundred and ten individual phage colonies were picked from each of these two selections and a region of approximately 300 bp including the coding region for the 9 amino acid random region was amplified by PCR. Ninety-six of the PCR products from each of them with the most distinct PCR bands were sent for sequencing. The sequencing results were translated into amino acid sequence and the 9 random amino acid of each sequence was aligned by hand based on common sequence characteristics (Fig. 6). As reference and for comparison we have also added the results from a recent study of human granzymes A and K by phage display, using the same T7 phage library as for the two *Zebra mbuna* enzymes (Fig. 6) (Aybay et al., 2023). The frequency of each amino acid in different positions for both *Zebra mbuna* enzymes were plotted into a Web-logo figure based on the alignment in Fig. 6 A (Fig. 6C and D).

The *Zebra mbuna* granzyme A1 showed a strong preference for the

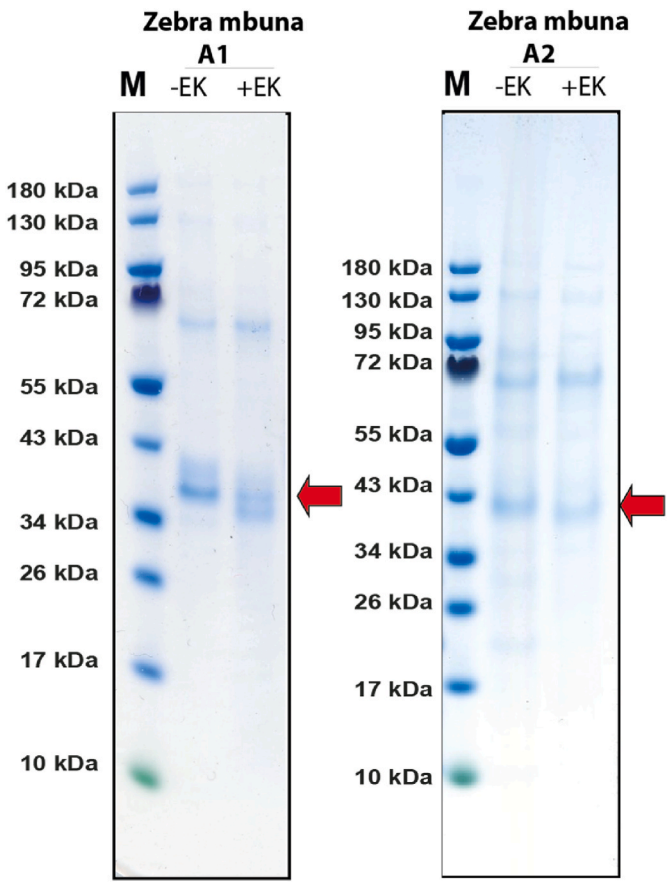

Fig. 5. SDS-PAGE gel of *Zebra mbuna* granzyme A1 and A2. The *Zebra mbuna* granzymes were produced as inactive enzymes containing an N-terminal His_6 tag and an enterokinase site. The enzymes were produced in the human cell line HEK293-EBNA with the episomal vector pCEP-Pu2. Enterokinase (EK) was used to cleave off the N-terminal tail to activate the enzymes. The inactive enzymes with the N-terminal purification tag and the active enzymes were analyzed by separation on SDS-PAGE gel and visualized with Coomassie Brilliant Blue staining. PAGE Ruler was used as marker. M: marker; +EK: with entrokinase; -EK: without entrokinase. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

positively charged amino acid arginine (Arg) in the P1 position, similar to both human granzymes A and K (Fig. 6A and B). In fact, all the 89 clones with high quality sequences had an Arg in the P1 position and none of them had a lysine (Lys) in this position (Fig. 6A). In contrast we found that the *Zebra mbuna* granzyme A2 had a strong preference for the aromatic amino acid tyrosine (Tyr), as demonstrated by the Tyr found in the P1 position of all the 76 clones with high quality sequences (Fig. 6A).

Concerning the extended specificities, we found marked differences in some of the positions between the two enzymes (Fig. 6A). Granzyme A1 showed a high preference for aromatic amino acids, such as Phe and Trp, in the P2 position just N-terminally of the cleavage site, similar to what we have seen for human granzyme K, whereas granzyme A2 preferred small amino acids such as Gly and Ala in this position (Fig. 6A and B). In general, granzyme A2 also preferred smaller aliphatic amino acids N terminally of the cleavage site similar to human granzyme A (Fig. 6A and B). The *Zebra mbuna* enzymes also differed in the P3 position where granzyme A1 preferred negatively charged amino acids such as Glu and Asp, whereas granzyme A2 preferred Val (Fig. 6A). In the P4 position, both enzymes preferred Val (Fig. 6A). Both granzymes preferred Ser in the P1′position just C terminally of the cleavage site and there were also similarities in the P3′and P4' positions where both enzymes preferred Val, Leu, Ala and Ser (Fig. 6A). However, a minor difference was seen in the P2′position where granzyme A2 preferred small amino acids such as Gly and Ala, whereas A1 preferred slightly larger aliphatic amino acids such as Leu and Val (Fig. 6A). Together, these findings show that there are marked differences between granzyme A1 and A2 in a few positions but also major similarities in other positions. It is important to note that the His residues, sometimes frequently occurring in positions P3', P4′and P5', should be interpreted with caution in these analyses as they in most cases originate from the histidine tag of the modified phage coat protein.

3.4. Verifying the consensus sequence by the use of recombinant protein substrates

To verify the sequences obtained from the phage display analysis and to obtain quantitative estimates of the importance of amino acids at and around the cleavage site, we used an inhouse established type of recombinant protein substrates. The dominating sequences obtained for the P2 and P1 positions in the phage display analysis and several variants of these sequences were designed and produced as recombinant substrates in a two-thioredoxin (Trx) system. We have produced almost 400 such 2xTrx substrates however only a few of the most relevant existing substrates showed any detectable cleavage with the *Zebra mbuna* enzymes, why most of the substrates for the validation of the two *Zebra mbuna* enzymes had to be newly produced. Double stranded oligonucleotides encoding the consensus substrates and a number of variants of these substrates were designed, ordered and ligated into the 2xTrx substrate vector. A His_6-tag was added in the C-terminal end of the second Trx molecule to facilitate purification (Fig. 7A). The vectors carrying the target sequences were transferred into *E. coli Rosetta gami* for expression and purification. These purified 2xTrx proteins were then used to analyze the specificities of the two *Zebra mbuna* enzymes (Figs. 7 and 8).

The analysis of the *Zebra mbuna* granzyme A1 showed that Ser was favored over Asp by a factor of 2–3 in the P1′position. Unexpectedly, we found that Phe was favored over Gly, and Tyr over Leu, in the P5 and P2′positions, respectively (Fig. 7C). This result indicates that the sequence FVEFRSYSL is actually a better consensus compared to the GVEFRSLSL that was the result from the phage display. This is the first time we have seen a sequence tested in the 2xTrx system that is better than the consensus from the phage display. Of the six additional variants tested in panels D and E, four of them were not cleaved at all, one in panel D and three in panel E (Fig. 7D and E). Two of these six sequences, the WGWRSLVL and VYWRSLVL in panel D, were 15–20 times less efficiently cleaved compared to the consensus, indicating an overall very high selectivity of this enzyme.

The analysis of *Zebra mbuna* granzyme A2 confirmed the high specificity for Tyr (Y) in the P1 position (Fig. 8A). Compared to the strong cleavage seen with Tyr in the P1 position, a 20–50 times lower rate of cleavage was obtained with a phenylalanine (Phe, F) in this position (Fig. 8A). In contrast, we found no cleavage of a substrate with a Leu (L) in the P1 position, and an almost undetectable cleavage of a substrate with a tryptophan (Trp, W) in this position (Fig. 8A). We also found that Gly (G) in the P2 position was clearly favored over larger amino acids such as the aromatic amino acid Trp and Leu, and that Ser (S) in the P1′position was favored over Val (V) and Asp (D) (Fig. 8B, C and D). The consensus cleavage site for *Zebra mbuna* granzyme A2 thereby show several characteristics similar to the mouse and rat mucosal mast cell chymases mMCP-1 and rMCP-2, and the rat vascular chymase (Akula et al., 2021).

4. Discussion

The detailed analyses of the primary and extended specificity of the *Zebra mbuna* granzyme A1 and A2 confirmed our initial predictions concerning granzyme A-like enzymes in this fish species. Based on the three amino acid positions forming the substrate binding pocket in these enzymes (Fig. 2B), we predicted that two of the *Zebra mbuna* granzyme

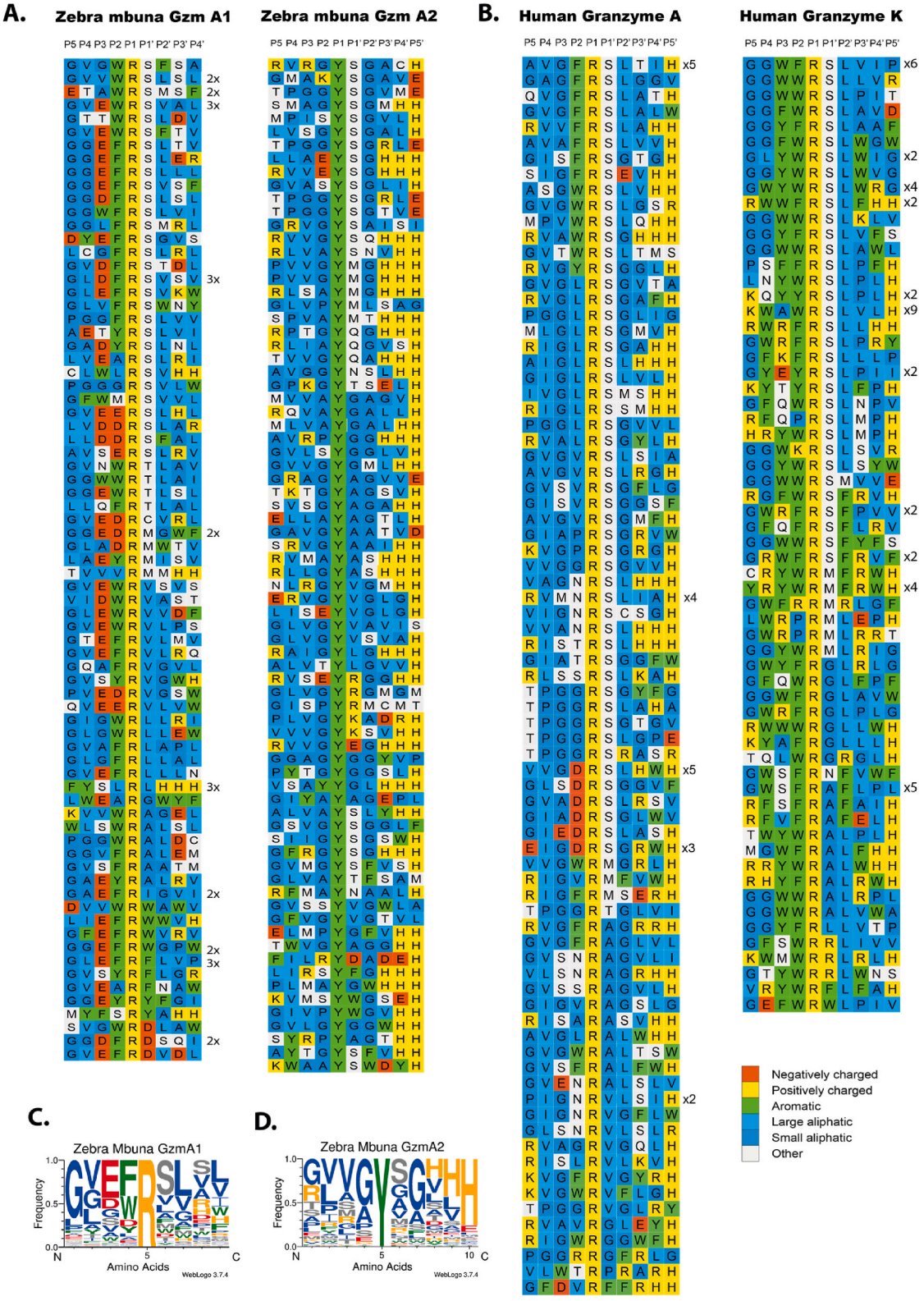

Fig. 6. Sequences from phage display analyses of *Zebra mbuna* granzymes A1 and A2 in comparison with human granzyme A and K. Panel **A** shows phage display sequences generated after cleavage by *Zebra mbuna* granzymes A1 and A2 after 7 and 6 rounds of selection, respectively. After the last round of selection, the released phages were collected for sequencing. The general amino acid sequences are P-G-G-?-?-?-?-?-?-?-?-?-H-H-H-H-H-H and the "?" means randomized amino acids. The amino acid sequences were aligned into a P5–P4′ consensus. The cleavage occurs between positions P1 and P1'. The classification of amino acids is marked by color, based on the side chain properties as indicated in the right bottom corner of the figure. Gzm: granzyme. In panel **B**, for comparison and as reference samples, we show the phage display results from a recent analysis of human granzymes A and K (Aybay et al., 2023). In panels **C** and **D,** we have included Web-logo presentations of the target selectivity of the two *Zebra mbuna* enzymes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

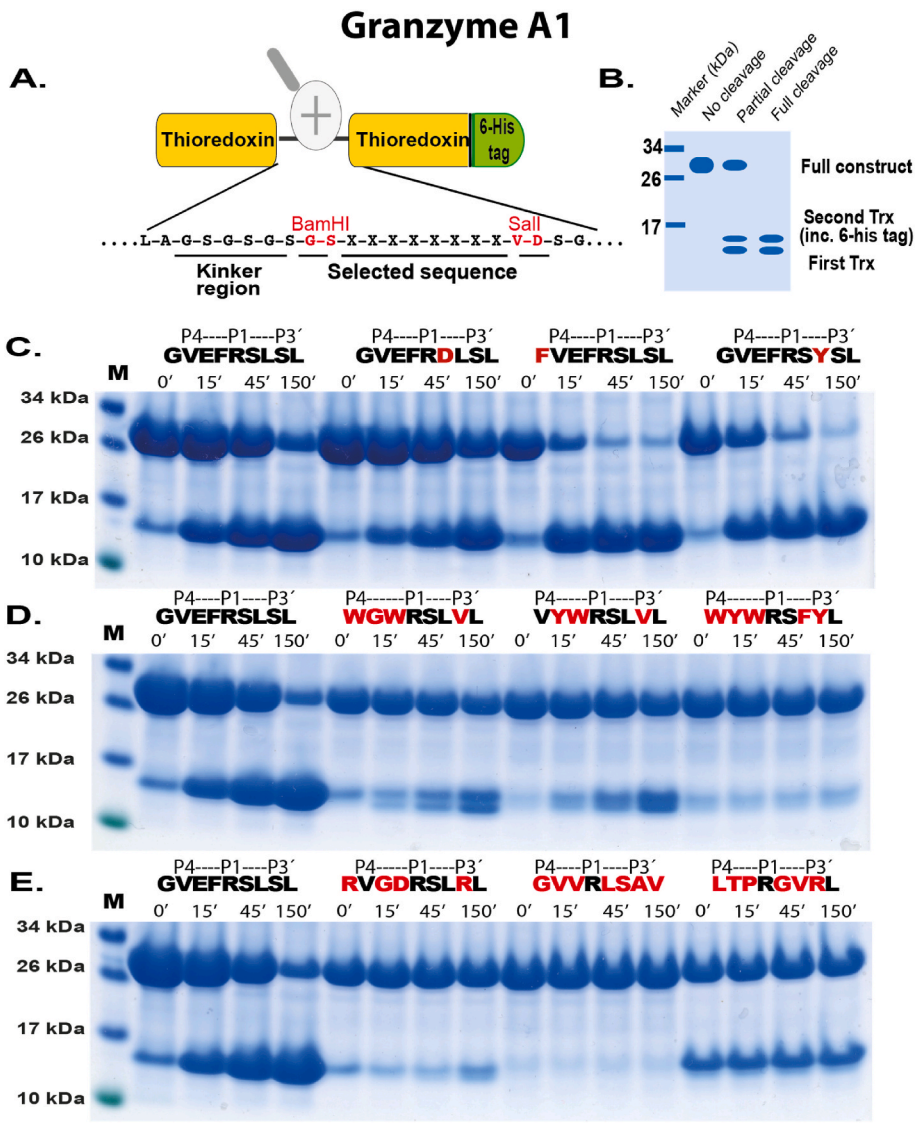

Fig. 7. Analysis of the cleavage specificity of *Zebra mbuna* granzyme A1 by the use of recombinant protein substrates. In panel **A**, the overall structure of the recombinant protein substrates used for the analysis of the cleavage by *Zebra mbuna* granzyme A1 is depicted. The sequences analyzed were positioned between two thioredoxin molecules with a His_6-tag attached to the C terminal of the second Trx molecule. Two unique restriction sites (Bam HI and Sal I) were used for the insertion of the target sequences between the two thioredoxin molecules. A schematic illustration of hypothetical result from this type of analysis is depicted in panel **B** explaining various possibilities of the cleavage patterns. The cleavage of substrates by *Zebra mbuna* granzyme A1 is shown in panels **C**, **D** and **E**. The sequences and cleavage times (minutes) are indicated above the lanes of different samples. The bands at approximately 30 kDa represent uncleaved substrates and those between 10 kDa and 17 kDa represent cleaved substrates. The difference in size of cleaved bands was caused by the His_6-tag of one of the Trx molecules.

A/K homologs, granzymes A2 and A3, had changed their primary specificity from the traditional tryptic specificity to chymotryptic specificity and that granzyme A1 had a tryptic activity similar to human granzymes A and K (Fig. 4B). In agreement with this, phage display followed by recombinant substrate analyses clearly demonstrated that granzyme A2 exhibited a strong preference for Tyr in the P1 position, thereby confirming that it is a chymase having chymotrypsin-like primary specificity, and that granzyme A1 was a traditional tryptase with primary specificity for Arg. Interestingly, we found that the extended specificity of the A2 enzyme was very similar to that of human granzyme A (Fig. 6) (Aybay et al., 2023). This finding suggests that this was the original extended specificity of granzyme A/K enzymes and that granzyme K has drifted away from this extended specificity, most likely to obtain new functions different from those performed by the granzyme A and its homologs (Fig. 6) (Aybay et al., 2023). Notably, the mammalian granzyme Ks form a separate subfamily indicating that this enzyme is under an evolutionary selective pressure (Fig. 4B).

The alignment presented in Fig. 3 shows that all six *Zebra mbuna* granzyme A/K enzymes are very homologous, with over 70% identity between the most divergent members and over 90% identity between granzymes A4 and A5, in their sequences indicating homogenization, most likely by gene conversion. All of these enzymes are more similar to each other than to human granzymes A and K (Fig. 3A). An alternative explanation for this situation may be recent internal duplications within the locus. However, this would explain the high homology of the internal genes but not of the granzymes A5 and A6 which are the ends of the locus. Moreover, the high sequence similarity between the gene located in the met-ase locus and the genes within the granzyme A/K locus is striking, and it indicates that this gene transfer has occurred relatively recently or that gene conversion can occur over larger distances than previously expected (Figs. 1 and 3). We used dot plot analyses to study the sequences surrounding the gene in the met-ase locus and the homologous genes in the granzyme A/K locus, and we detected homology only over the exons and not in introns or in regions upstream or downstream of the gene. This result leaves us without any clues as to how the gene has been copied and inserted from the granzyme A/K locus to the met-ase locus. However, the clustering of the two chymases, granzyme A2 and A3, in the phylogenetic tree in Fig. 3C suggests that the A2 gene was copied and transferred to the met-ase locus. A transfer of one of the tryptase genes is less likely as it would have involved a secondary change of the A3 gene from a tryptase to a chymase (Figs. 1 and 3). One remarkable feature of this gene family, also shared by other enzyme families, is the relatively few mutations needed for a total change of specificity of an enzyme. The *Zebra mbuna* granzyme A/K genes are highly homologous but have completely different primary as well as extended specificities. A similar phenomenon has occurred in the

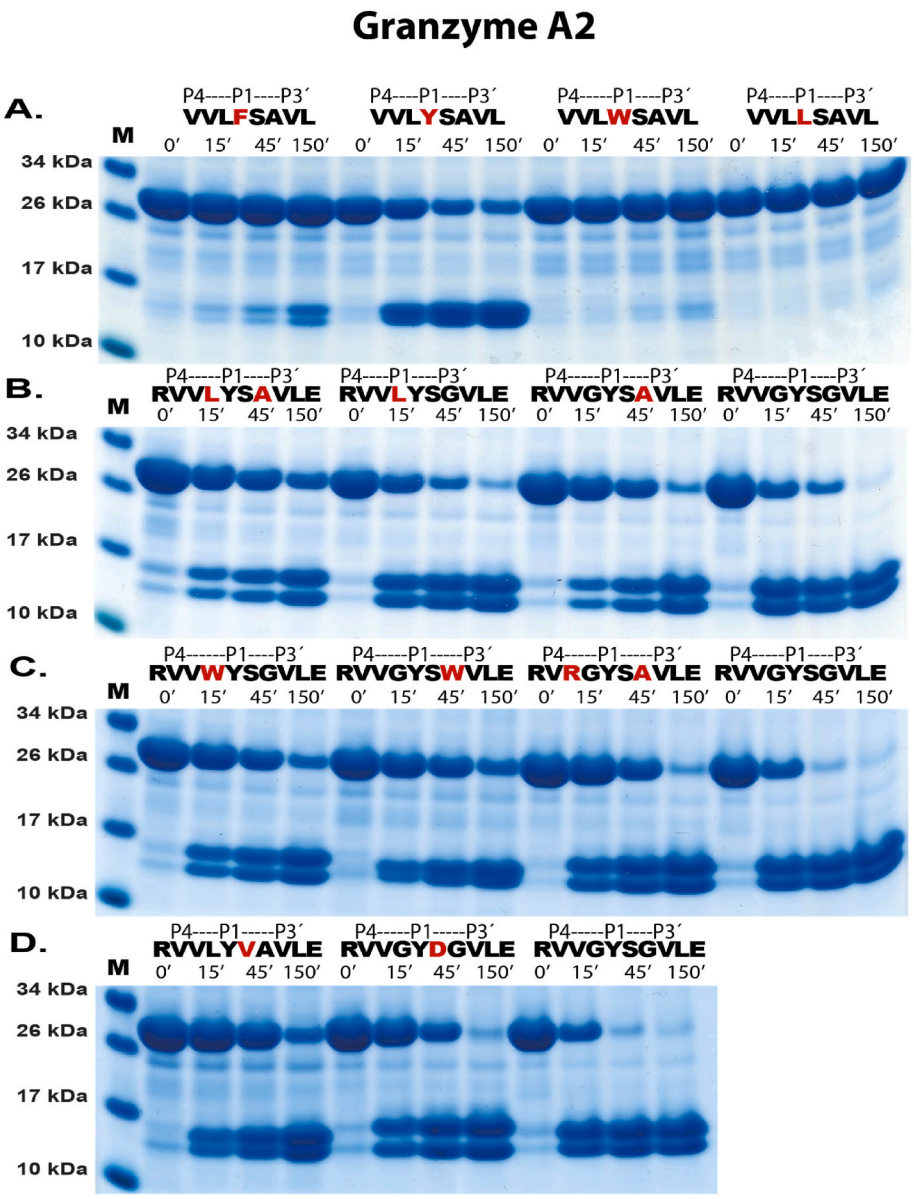

Fig. 8. Analysis of the cleavage specificity of *Zebra mbuna* granzyme A2 by the use of recombinant protein substrates. The cleavage of substrates by *Zebra mbuna* granzyme A2 is shown in panels **A**, **B**, **C** and **D**. The sequences and cleavage times (minutes) are indicated above the lanes of different samples. The bands at approximately 30 kDa represent un-cleaved substrates and those between 10 kDa and 17 kDa represent cleaved substrates. The difference in size of cleaved bands was caused by the His_6-tag of one of the Trx molecules.

chymase locus in cows and sheep where novel genes within the locus have obtained very different specificities, including asp-ase, tryptase and chymase specificities (Fu et al., 2021). These novel genes have most likely emerged from gene duplications of very similar genes followed by relatively few point mutations. Moreover, these novel ruminant enzymes have also changed tissue localization and are no longer immune proteases but instead they are expressed in the duodenum and has therefore been named duodenases (Fu et al., 2021).

As mentioned in the abstract, there are still unsolved questions concerning the biological functions of both granzyme A and K. Granzyme A was first thought to be involved in apoptosis induction of target cells (Lieberman, 2010). This was later questioned using more physiological concentrations of the enzyme. Evidence was instead presented for the role of this enzyme in the induction of a number of inflammatory cytokines and chemokines by monocytes (Metkar et al., 2008; Wensink et al., 2016; van Eck et al., 2017). Doubts on both of these functions are presented in a recent article on human granzymes A and K indicating that major questions remain on the primary targets and the major functions of these highly conserved enzymes in immunity (Aybay et al., 2023). Another intriguing question concerns the new functions adopted by the two new cichlid enzymes, granzyme A2 and A3, as they changed from being classical tryptases to now being chymases. We have no information about the tissue distribution of these enzymes and their potential in vivo targets. However, we found that both the primary and the extended specificity of one of these enzymes, granzyme A2, were relatively similar to the mammalian mast cell chymases and primarily the mucosal mast cell chymases, and the rat vascular chymase (Akula et al., 2021). We may therefore speculate that this enzyme has similar functions and expression pattern as these enzymes. However, this and other interesting questions need to wait for the production of specific antibodies that can detect the presence of this enzyme in different tissues and cells. A major challenge for production of specific antibodies could be the large similarities in primary sequence between granzyme A2 and the other members of the granzyme A/K family, which is why peptide antisera may be the most reliable way to obtain such reagent. It is interesting to note that fish lack the chymase locus, encoding the classical mast cell chymases. Chymases with a very similar specificity have been conserved in most tetrapods from reptiles to humans (Akula et al., 2021). The obvious question is then if fish have generated similar enzymes from other loci harboring hematopoietic serine proteases in fish. One possibility is that the new enzymes, granzyme A2 and A3 in *Zebra mbuna,* are examples of such a process and that they have adopted

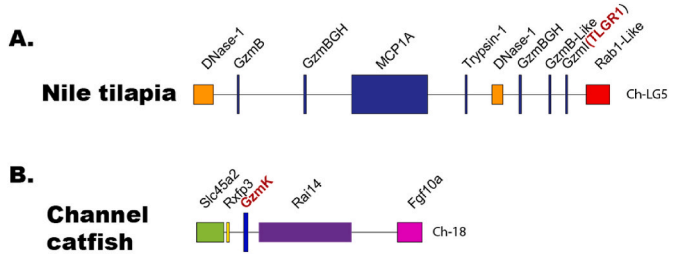

Fig. 9. An updated map of the fish specific hematopoietic serine protease locus of the Nile tilapia, and the location of a characterized catfish granzyme A/K gene. An in-scale picture of the fish specific hematopoietic serine protease locus from Nile tilapia. The protease genes are illustrated in double hight compared to the surrounding non-protease genes for an easier identification. The updated version of the locus has now several additional serine protease genes compared to the first published version of this locus indicating that the early versions of sequenced genomes should be handled with care as the initial information is often incomplete (Akula et al., 2015). The chymotryptic protease gene, named TLGR1, that is expressed in tilapia cytotoxic cells is marked in red bold text in panel A (Praveen et al., 2006a). Panel B shows the location of the catfish granzyme A/K gene named granzyme K in the present update of the catfish genome. This protease was initially named CFGR1 in the previous publications describing its isolation and cleavage specificity (Praveen et al., 2004, 2006b). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

functions similar to mast cell chymase and cathepsin G. We have so far only found these new granzyme A/K variants in cichlids but we cannot exclude that enzymes with chymase activity originating from other loci have also emerged in other fish species. To explore this possibility, future studies of hematopoietic serine proteases in different fish species are warranted.

Interestingly, an mRNA for a chymotryptic enzyme has been isolated from cytotoxic cells of another cichlid, the Nile tilapia (*Oreochromis niloticus*) (Praveen et al., 2006a). The gene for this enzyme, named GzmI, or previously TLGR1, is however not found in the granzyme A/K locus but instead in a fish specific locus (Fig. 9A) (Praveen et al., 2006a). This fish specific locus contains together with the fish met-ase locus the majority of the serine proteases that are found in brown branch of the phylogenetic tree in Fig. 4A (Akula et al., 2015). Chymotryptic activity seems thereby to be present in fish cytotoxic cells but may originate from an enzyme that clusters with non-granzyme A/K like fish hematopoietic serine proteases. Major updates on many fish genomes have recently resulted in a more correct picture of the various loci encoding hematopoietic serine proteases in fish. We therefore include the updated version of the fish specific locus encoding the Nile tilapia chymotryptic enzyme, named GzmI, or previously TLGR1 (Fig. 9A). Moreover, a channel catfish granzyme A/K homologue has been shown to exhibit tryptase activity, similar to what is observed in mammalian cytotoxic T cells and NK cells, indicating that granzyme A/K homologs with tryptase activity are present in fish (Praveen et al., 2004, 2006b). The gene for this enzyme, which cluster with granzyme A/Ks in a phylogenetic tree, is located in a completely new location not matching any of the previously identified loci for hematopoietic serine proteases in fish (Fig. 9B). This indicates that more changes in the mapping of the different hematopoietic serine protease genes are to be expected with coming updates of fish genomes.

We have previously analyzed an enzyme expressed by channel catfish NK-like cells, Catfish granzyme-like 1 (Thorpe et al., 2016). This enzyme was found to be a very specific met-ase, which indicates that fishes have homologs of mammalian granzyme M. Granzyme M has been shown to be expressed by both NK cells and cytotoxic T cells in humans (Aybay et al., 2023; Thorpe et al., 2016). We actually also detected mRNA for granzyme M in human CD4-positive T cells indicating that these met-ases could have a broader expression pattern and may not only be restricted to NK cells (Aybay et al., 2023). The catfish granzyme-like 1 enzyme may have a function in apoptosis induction of target cells by cleavage and activation of catfish caspase 6 (Thorpe et al., 2016). However, direct evidence for this function is still lacking due to major difficulties in producing catfish caspase 6 as recombinant protein.

A detailed study of different enzyme activities in various immune cell populations in the marine gilthead seabream and the European sea bass using chromogenic substrates has detected trypsin-like, chymotrypsin and met-ase activities in fish leukocytes and also low levels asp-ase activity in these cells (Chaves-Pozo et al., 2019). The enzymes directly responsible for these activities have not yet been identified, although the granzyme A homologue most likely was responsible for the tryptase activity (Chaves-Pozo et al., 2019). The findings of this study nicely demonstrated that similar enzyme specificities are present in fish immune cells as in the corresponding mammalian cells. Interestingly, our analyses of the catfish granzyme-like 1 and a few related, still unpublished enzymes, suggest that fish hematopoietic serine proteases generally show a higher extended specificity than their mammalian counterparts indicating that the analysis of their activity with chromogenic substrates may involve a risk of being undetected (Thorpe et al., 2016). Detailed analysis of a larger panel of hematopoietic serine proteases from various fish species with high resolution techniques, including phage display, is therefore needed to decipher their complexity in cleavage specificities, their in vivo targets and their functions. Hence, we can obtain a better picture of similarities and differences in the repertoire of hematopoietic serine proteases between fish and mammals. Such a detailed picture would be of major interest both from an evolutionary standpoint and for our understanding of what key functions are essential for a well-functioning immune system in all vertebrates.

Data availability

All data of importance for the manuscript is available in the manuscript.

Acknowledgments

The study was supported by grants from Knut and Alice Wallenberg Foundation (KAW 2017-0022).

References

Akula, S., Thorpe, M., Boinapally, V., Hellman, L., 2015. Granule associated serine proteases of hematopoietic cells - an analysis of their appearance and diversification during vertebrate evolution. PLoS One 10 (11), e0143091.
Akula, S., Fu, Z., Wernersson, S., Hellman, L., 2021. The evolutionary history of the chymase locus -a locus encoding several of the major hematopoietic serine proteases. Int. J. Mol. Sci. 22 (20).
Aybay, E., Ryu, J., Fu, Z., Akula, S., Enriquez, E.M., Hallgren, J., et al., 2023. Extended cleavage specificities of human granzymes A and K, two closely related enzymes with conserved but still poorly defined functions in T and NK cell-mediated immunity. Front. Immunol. 14, 1211295.
Chaves-Pozo, E., Valero, Y., Lozano, M.T., Rodriguez-Cerezo, P., Miao, L., Campo, V., et al., 2019. Fish granzyme A shows a greater role than granzyme B in fish innate cell-mediated cytotoxicity. Front. Immunol. 10, 2579.
Fu, Z., Akula, S., Qiao, C., Ryu, J., Chahal, G., de Garavilla, L., et al., 2021. Duodenases are a small subfamily of ruminant intestinal serine proteases that have undergone a remarkable diversification in cleavage specificity. PLoS One 16 (5), e0252624.
Heibein, J.A., Barry, M., Motyka, B., Bleackley, R.C., 1999. Granzyme B-induced loss of mitochondrial inner membrane potential (Delta Psi m) and cytochrome c release are caspase independent. J. Immunol. 163 (9), 4683–4693.
Hellman, L., Thorpe, M., 2014. Granule proteases of hematopoietic cells, a family of versatile inflammatory mediators - an update on their cleavage specificity, in vivo substrates, and evolution. Biol. Chem. 395 (1), 15–49.
Kondos, S.C., Hatfaludi, T., Voskoboinik, I., Trapani, J.A., Law, R.H., Whisstock, J.C., et al., 2010. The structure and function of mammalian membrane-attack complex/perforin-like proteins. Tissue Antigens 76 (5), 341–351.
Lieberman, J., 2010. Granzyme A activates another way to die. Immunol. Rev. 235 (1), 93–104.
MacDonald, G., Shi, L., Vande Velde, C., Lieberman, J., Greenberg, A.H., 1999. Mitochondria-dependent and -independent regulation of Granzyme B-induced apoptosis. J. Exp. Med. 189 (1), 131–144.

Metkar, S.S., Menaa, C., Pardo, J., Wang, B., Wallich, R., Freudenberg, M., et al., 2008. Human and mouse granzyme A induce a proinflammatory cytokine response. Immunity 29 (5), 720–733.

Praveen, K., Evans, D.L., Jaso-Friedmann, L., 2004. Evidence for the existence of granzyme-like serine proteases in teleost cytotoxic cells. J. Mol. Evol. 58 (4), 449–459.

Praveen, K., Leary 3rd, J.H., Evans, D.L., Jaso-Friedmann, L., 2006a. Molecular characterization and expression of a granzyme of an ectothermic vertebrate with chymase-like activity expressed in the cytotoxic cells of Nile tilapia (Oreochromis niloticus). Immunogenetics 58 (1), 41–55.

Praveen, K., Leary 3rd, J.H., Evans, D.L., Jaso-Friedmann, L., 2006b. Nonspecific cytotoxic cells of teleosts are armed with multiple granzymes and other components of the granule exocytosis pathway. Mol. Immunol. 43 (8), 1152–1162.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., et al., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61 (3), 539–542.

Schechter, I., Berger, A., 1967. On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. 27 (2), 157–162.

Shinkai, Y., Takio, K., Okumura, K., 1988. Homology of perforin to the ninth component of complement (C9). Nature 334 (6182), 525–527.

Thorpe, M., Akula, S., Hellman, L., 2016. Channel catfish granzyme-like I is a highly specific serine protease with metase activity that is expressed by fish NK-like cells. Dev. Comp. Immunol. 63, 84–95.

van Eck, J.A., Shan, L., Meeldijk, J., Hack, C.E., Bovenschen, N., 2017. A novel proinflammatory role for granzyme A. Cell Death Dis. 8 (2), e2630.

Wensink, A.C., Kok, H.M., Meeldijk, J., Fermie, J., Froelich, C.J., Hack, C.E., et al., 2016. Granzymes A and K differentially potentiate LPS-induced cytokine response. Cell Death Dis. 2, 16084.

Wernersson, S., Reimer, J.M., Poorafshar, M., Karlson, U., Wermenstam, N., Bengten, E., et al., 2006. Granzyme-like sequences in bony fish shed light on the emergence of hematopoietic serine proteases during vertebrate evolution. Dev. Comp. Immunol. 30 (10), 901–918.