

UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS BIOLÓGICAS Y DE LA SALUD

PROGRAMA DE DOCTORADO EN CIENCIAS

(QUÍMICO BIOLÓGICAS Y DE LA SALUD)

Molecular Characterization, Expression Analysis, Recombinant
Expression and Refolding of the Chicken-type and Goose-type
Lysozymes from Totoaba (*Totoaba macdonaldi*)



TESIS

Que para obtener el grado de

Doctorado en Ciencias
(Químico Biológicas y de la Salud)

Presenta

Elena Nohelí Moreno Córdova

Universidad de Sonora

Repositorio Institucional UNISON



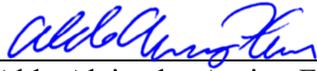
**"El saber de mis hijos
hará mi grandeza"**



Excepto si se señala otra cosa, la licencia del ítem se describe como openAccess

FORMA DE APROBACIÓN

Los miembros del Jurado Calificador designado para revisar el trabajo de **Elena Nohelí Moreno Córdova**, lo han encontrado satisfactorio y recomiendan que sea aceptado como requisito parcial para obtener el grado de Doctor en Ciencias (Químico Biológicas y de la Salud).



Dr. Aldo Alejandro Arvizu-Flores
Director Académico



Dr. Eduardo Ruiz-Bustos
Secretario



Dra. María Auxiliadora Islas-Osuna
Sinodal



Dr. Michael Frederick Criscitiello
Sinodal



Dr. Alonso Alexis López-Zavala
Sinodal

FUNDING

This research would not have been possible without the valuable financial support of the Collaborative Research Grant [Project number 2015-020] between the Texas A&M University and CONACyT “*Immunogenetic assessment of the critically endangered totoaba fish to enhance survival and repopulation after hatchery rearing*”.

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to my professors, mentors, academic institutions, laboratories and colleagues, whose valuable support and assistance have made this research possible.

To CONACyT, for the scholarship granted to carry out my Doctoral studies.

To Universidad de Sonora, my *Alma mater*, to Departamento de Ciencias Químico Biológicas y de la Salud, and Posgrado en Ciencias de la Salud, for their valuable academic support and their infrastructure that made possible the development and culmination of this research project.

To the University of Texas A&M (College of Veterinary Medicine and Biomedical Sciences), and especially to the Comparative Immunogenetics Laboratory (Criscitiellos's Lab), for their invaluable support in academic advising and with indispensable laboratory resources that made possible this research.

To Centro de Investigación en Alimentación y Desarrollo (CIAD), to Coordinación de Tecnología de Alimentos de Origen Vegetal, and especially to Laboratorio de Biología Molecular de Plantas, for opening their doors to carry out my doctoral stay, and for their valuable academic support.

To Centro Reprodutor de Especies Marinas del Estado de Sonora (CREMES) for their kind contribution in donating totoaba fish specimens to carry out this research and for their technical support.

I would like to express my deepest gratitude to Dr. Aldo A. Arvizu-Flores, my research supervisor and mentor, for giving me the opportunity to be part of his work group, for his valuable shared knowledge and experience, for his patient guidance and unrelenting support at each stage of this research, for his insightful suggestions and constructive advice, for his profound belief in my work, motivation, and his invaluable contributions to my professional and personal growth.

I would like to express my deepest appreciation to my committee, Dr. María A. Islas-Osuna, Dr. Michael F. Criscitiello, Dr. Alonso A. López-Zavala and Dr. Eduardo Ruiz-Bustos, for their patient guidance and unrelenting support, for their valuable advice and shared experience, for their extensive knowledge, insightful suggestions, useful critiques and helpful contributions that made this research possible. Thank you for your profound beliefs in my work, your encouragement, motivation and your invaluable contributions to my professional and personal growth.

I would like to express my very great appreciation to Dr. Carmen A. Contreras-Vergara for her valuable contribution and suggestions during the planning, development and analysis of RT-PCR, RACE and molecular cloning experiments. Her willingness to share her experience, insightful knowledge, and time so generously has been very much appreciated.

I'm deeply indebted to M. Sc. Mónica G. Reséndiz-Sandoval, for her invaluable collaboration and assistance in the planning, development and analysis of gene expression experiments. Her great technical expertise and willingness to share her insightful knowledge so generously have been deeply appreciated.

I'd like to extend my gratitude to Q. B. César Otero-León, for his valuable technical support in the laboratory, for his positive attitude and constant willingness to help in everything he can and more.

I gratefully acknowledge the assistance of Dr. Frans Tax (University of Arizona) for his support in the safekeeping of laboratory material essential to carry out this research.

I'd like to extend my sincere gratitude to Laboratorio de Inmunología (CIAD) and Dr. Jesús Hernández, for their collaboration and valuable support in the acquisition of laboratory tools, and loan of equipment essential to carry out this research work.

I'd like to extend my sincere gratitude to Laboratorio de Estructura Biomolecular (CIAD) and Dr. Rogerio Sotelo-Mundo, for their valuable advice, technical and academic support, and collaboration with laboratory material and infrastructure essential to accomplish this research project.

I'm grateful to Laboratorio de Inmunología y Biología Celular (UNISON), M. Sc. Lucila Rascón and my colleagues Dr. Gloria López, M. Sc. Thania Garzón and Dr. Efraín Alday, for their valuable technical and academic support, collaboration in the loan of laboratory equipment, their good advice and friendship.

I would like to extend my sincere gratitude to my colleagues from Laboratorio de Investigación en Alimentos (UNISON): cDr. Brenda Samaniego, Dr. María Moreno, Dr. Manuel Carretas, M. Sc. Luis Vázquez, and M. Sc. Brisa Ibarra, and from Laboratorio de Biología Molecular de Plantas (CIAD): Dr. Aldana Laino, M. Sc. Lucía Angulo, M. Sc. Martín Andrade, M. Sc. Nacxit Delgado, M. Sc. Andrés López and M. Sc. Kathya, whom I had the great pleasure of working with, for their unwavering academic and technical support, for their friendship and for making the long laboratory working hours nice.

The development and culmination of this research work would not have been possible without the invaluable support of my family and friends who accompanied me on this journey, to whom I wish to express my sincere gratitude.

I'd like to express my deepest gratitude to Elena Moreno, my Mother, for her unrelenting support and faithful company throughout all my professional training, for her example of strength, courage and perseverance, for her tender love, dedication and care for me, for her profound belief in me, for her motivation to fight for my dreams and her encouragement to not give up in the face of adversity or disappointment, for her prayers, advice and wisdom, and for being my faithful companion in sleepless nights and long working hours in the laboratory and at home.

I'm deeply grateful to my family, especially my aunts Carmen Moreno and Cande Moreno, and my cousins Claudia Gálvez, David Naranjo, Myriam Villalba and Jorge Villalba, for their unfailing emotional support, encouragement and motivation.

I'm very grateful to my friends Brenda Samaniego, María Moreno, Lucy Angulo, Leo Fox, Nitzia Arce, Ana Karenth López and Edgar Melgoza, who went through hard times with me, cheered me on, and celebrated each accomplishment, for their valuable academic and emotional support, shared knowledge, motivation and inspiration to give my best always and to never give up.

I'm deeply grateful to my friends Claudia Gálvez, Fernanda Amaya, Jessica Wong, Marysol Navarro, and my mentors Eliza Osuna, Javier Pérez and Georgina Hernández, for their unfailing support, prayers, good advice, encouragement and motivation, for going with me through hard times, and celebrate with me every accomplishment. Thank you for inspiring me to grow in wisdom, knowledge, virtue and grace, and to never give up on my dreams.

DEDICATION

To God, the Creator of life and the most extraordinary Designer, the One in whom all science, knowledge, and wisdom comes from and finds its purpose, the Scientist by excellence, the One who is great above all else. My deepest gratitude is to You. Thank you for being with me all throughout this long journey, for giving me the strength, grace, and wisdom needed to carry out this work, for guiding my every step so patiently, and for being the one who has always believed in me the most, for teaching me to overcome my fears and every obstacle, and for granting my biggest dreams. Thank you for your great faithfulness to me and to your promises, for your unfailing love and care, and for being the God who not only sustains life and the whole universe with his might but so personally sustains me. Without you, I could not have finished this race, so I'm honored to dedicate this achievement to you.

To my Mother, my greatest example and inspiration, and my strongest pillar. My deepest appreciation is to you. Thank you for your unfailing love and unrelenting care for me, for your unconditional support, timely encouragement, motivation, and endless patience. Thank you for being with me in every step and season of this long journey, for believing in me always, for teaching me to be brave and perseverant, for inspiring me to fight for my dreams and not give up, and most important, thank you for modeling me true faith in God and teach me to love him and trust in him. It was your tender love and prayers that raised me up when I got weary. The completion of my dissertation would not have been possible without your support and nurturing. This achievement is also yours.

“We all know that chance, fortune, fate or destiny – call it what you will, has played a considerable part in many of the great discoveries in science. We do not know how many, for all scientists who have hit on something new have not disclosed exactly how it happened. We do know, though, that in many cases it was a chance observation which took them into a track which eventually led to a real advance in knowledge or practice. This is especially true of the biological sciences for there we are dealing with living mechanisms about which there are enormous gaps in our knowledge”

Sir Alexander Fleming (1945)- From his speech at the Nobel Banquet

“And when after the long search, some new fragment of truth has been captured, it comes not as something discovered but as something revealed. The answer is so unexpected, and yet so simple and aesthetically satisfying, that it carries instant conviction, and that wonderful never-to-be forgotten moment when one says to oneself ‘Of course that’s it’”

Lawrence Bragg (1968) – The margins of Science: the white coated worker / Quoted by Louise N. Johnson (1998) – The early history of lysozyme

“We shall hear more about lysozyme”

Sir Alexander Fleming (1922)

“We cannot wrap our minds around the riches of God, the depth of his wisdom, and the marvel of his perfect knowledge. Its depths can never be measured! We cannot understand His decisions or explain the mysterious ways that He works! For all that exists originates in Him, comes through Him, and ends up in Him. So, for Him be the glory forever and ever”
(Romans 11: 33, 36)

“All things were created and exist through Him; every detail was crafted through His design and for His purposes. And He Himself existed and is before all things, and in Him all things hold together” (Colossians 1:16, 17)

INDEX

	Page
List of Tables	xii
List of Figures	xiii
Objectives	xvi
Abstract	xvii
Resumen	xviii
Introduction	1
Bibliographic Background	4
Overview of Fish Innate Immune System.....	4
Recognition of Non-self and Activation of the Innate Immune System.....	9
Lysozyme: Structure and Function.....	11
The Catalytic Mechanisms of C-type and G-type Lysozymes.....	18
The Antibacterial Activity of Lysozymes Independent of their Catalytic Function.....	22
Lysozymes in Teleost Fish.....	23
Totoaba: A Critically Endangered Species with Great Potential in Aquaculture.....	25
Chapter I	29
Materials and Methods	30
cDNA Cloning and Expression Profile of the C-type and G-type Lysozymes from Totoaba (<i>Totoaba macdonaldi</i>).....	30
Fish Animal Samples.....	30
cDNA synthesis from Extracted RNA.....	30
Cloning of the Full-length cDNA Sequences.....	31
Sequence Analysis and Molecular Modeling.....	35
Expression Analysis in Totoaba Tissues.....	36
Results and Discussion	37
Characterization of the C-type and G-type Lysozymes cDNA Sequences.....	37
Analysis of Conserved Structural Features in the C-type and G-type Lysozymes of Totoaba.....	40
Phylogenetic Analysis of the C-type and G-type Lysozymes from Totoaba.....	48

	Page
Expression Analysis of the C-type and G-type Lysozyme Genes in Different Tissue Samples.....	50
Conclusions	53
Chapter II	54
Materials and Methods	55
Purification and Refolding of the Recombinant C-type and G-type Lysozymes from Totoaba.....	55
Recombinant Expression	55
Inclusion Body Isolation and Solubilization	57
Protein Purification by Immobilized Metal Ion Affinity Chromatography.....	58
Refolding Assay Conditions	59
Lysozyme Activity Assays.....	62
Turbidimetric assay.....	62
Solid phase assay.....	62
Binding Affinity Experiments by Isothermal Titration Calorimetry...	62
Results and Discussion	64
Recombinant Expression and Purification of the C-type and G-type Lysozymes from Inclusion Bodies.....	64
<i>In vitro</i> Refolding of Denatured C-type and G-type Lysozymes.....	68
Conclusions	93
Recommendations	95
References	96
Appendixes	117

LIST OF TABLES

	Page
Background	
Table I. Summary of the general features of lysozymes in the animal kingdom.....	13
Table II. Current research on totoaba's immune response against bacterial pathogens and immunostimulants.....	28
Chapter I	
Table III. List of oligonucleotide primers used for RT-PCR, RACE PCR and qRT-PCR	32
Chapter II	
Table IV. Matrix design for refolding of TmLyzc under reducing conditions (GSH:GSSG).....	60
Table V. Matrix design for refolding of TmLyzg under reducing conditions (DTT)	60
Table VI. Base refolding formulation for TmLyzc refolding experiments.....	61
Table VII. Base refolding buffer formulation for TmLyzg refolding experiments.....	61
Table VIII. Experimental results of refolding TmLyzc.....	69
Table IX. Experimental results of refolding TmLyzg.....	72

LIST OF FIGURES

	Page
Background	
Figure 1. Teleost fish as models for the innate immune system evolution in vertebrates.....	5
Figure 2. Overview of the defense mechanisms involved in the innate immune system of teleost fish.....	6
Figure 3. Recognition of non-self and activation of the innate immune response.....	10
Figure 4. Structure of HEWL and binding cleft subsites.....	15
Figure 5. Structure of GEWL and binding cleft subsites.....	17
Figure 6. General mechanism for retaining glycosidases: the catalytic mechanism of HEWL (c-type lysozymes – retaining β -glycosidases).....	20
Figure 7. General mechanism for inverting glycosidases: the catalytic mechanism of GEWL (g-type lysozymes – inverting β -glycosidases).....	21
Figure 8. Totoaba facts and localization of the Gulf of California habitat.....	26
Chapter I	
Figure 9. General strategy diagram for the amplification of TmLyzg and TmLyzc partial cDNA sequences by RT-PCR and the amplification of 3' and 5' cDNA ends by RACE-PCR.....	34
Figure 10. Nucleotide and deduced amino acid sequence of <i>Totoaba macdonaldi</i> g-type lysozyme (TmLyzg).....	38

	Page
Figure 11. Nucleotide and deduced amino acid sequence of <i>Totoaba macdonaldi</i> c-type lysozyme (TmLyzc).....	39
Figure 12. Multiple sequence alignment and predicted secondary structures of g-type lysozymes from totoaba and other fish	41
Figure 13. Multiple sequence alignment and predicted secondary structures of c-type lysozymes from totoaba, other fish and higher vertebrates.....	43
Figure 14. Predicted tertiary structure of TmLyzg, showing the active site and conserved structural elements.....	45
Figure 15. Predicted tertiary structure of TmLyzc, showing the active site, and conserved structural elements.....	46
Figure 16. Cladogram of TmLyzg, TmLyzc and known orthologs in the vertebrate phylum.....	49
Figure 17. Tissue profile for the relative expression levels of the g-type and c-type lysozymes mRNA of totoaba.....	51
Chapter II	
Figure 18. Design of synthetic gene constructs of TmLyzg and TmLyzc for recombinant expression	56
Figure 19. SDS-PAGE analysis of the kinetics of recombinant expression of TmLyzg and TmLyzc in <i>E. coli</i> BL21 Gold (DE3).....	65
Figure 20. SDS-PAGE analysis for the purification of recombinant TmLyzg by IMAC chromatography at denaturing conditions using 8 M urea.....	66
Figure 21. SDS-PAGE analysis for the purification of recombinant TmLyzc by IMAC chromatography at denaturing conditions using 8 M urea.....	67

	Page
Figure 22. SDS-PAGE analysis of the renaturation of TmLyzc inclusion bodies using the Pierce Protein Folding kit system.....	70
Figure 23. SDS-PAGE analysis of the renaturation of TmLyzg inclusion bodies using the Pierce Protein Folding kit system.....	73
Figure 24. Lysozyme activity by solid phase assay of TmLyzg and TmLyzc refolding trials.....	74
Figure 25. SDS-PAGE analysis and solid phase lysozyme assay of refolded recombinant TmLyzc and TmLyzg.....	75
Figure 26. Proposed energy landscape of the folding process of hen-egg lysozyme	84
Figure 27. Proposed schematics of dominants folding pathways of hen-egg lysozyme studied by molecular dynamics simulations	89

OBJECTIVES

General Aim

To characterize the chicken-type and goose-type lysozymes from totoaba (*Totoaba macdonaldi*) at the molecular level and to produce them as functional recombinant proteins.

Particular Aims

1. To determine the full-length cDNA sequences of the genes that encode the c-type and g-type lysozymes from totoaba (*Totoaba macdonaldi*).
2. To evaluate the expression level of totoaba's lysozymes in tissues of healthy totoaba juveniles by qRT-PCR.
3. To produce totoaba's lysozymes as recombinant proteins by heterologous expression in *Escherichia coli*.
4. To evaluate conditions for *in vitro* refolding of the recombinant lysozymes from totoaba.

ABSTRACT

Lysozyme (EC 3.2.1.17) plays a critical role in the innate immune response against bacterial pathogens, hydrolyzing the peptidoglycan, which shapes their cell walls. C-type and g-type lysozymes exhibiting unique features are found in teleost fish species, suggesting lysozyme's critical role in these organisms' defense system. Totoaba (*Totoaba macdonaldi*) is an endemic and critically endangered species from the Gulf of California, Mexico. Totoaba's aquaculture demands further research towards the development of immunoprophylactic strategies and improve this species large-scale farming production. In this study, the c-type (TmLyzc) and g-type (TmLyzg) lysozymes genes from totoaba were cloned and characterized. The TmLyzc and TmLyzg full-length cDNA sequences were of 432 bp and 582 bp, encoding to polypeptides of 143 and 193 amino acids, respectively. The amino acid sequences shared high identities (90-60%) and close phylogenetic relationships with other fish and higher vertebrate lysozymes, as well as structural and functional domains typical of the lysozyme superfamily. Expression analysis by qRT-PCR showed that TmLyzc and TmLyzg were mainly expressed in the stomach, pyloric caeca, and heart. The findings suggest that lysozymes have a significant role in defense of totoaba against bacterial infections and also may be involved in digestion. In the second leg of this research, TmLyzc and TmLyzg were recombinantly expressed in *E. coli* BL21 (DE3) at 25 °C with 0.5 mM IPTG, isolated from inclusion bodies and subjected to *in vitro* refolding experiments using the Pierce® Protein Folding system. Refolding both lysozymes did not yield active enzymes, either because the proteins were partially folded or because fusion tags somehow hindered their active site. Although the tested refolding conditions were found promising, further optimization is required to recover active enzymes.

RESUMEN

La lisozima (EC 3.2.1.17) desempeña un papel crítico en la respuesta inmune innata contra patógenos bacterianos, catalizando la hidrólisis del peptidoglicano que da forma a sus paredes celulares. En peces teleósteos se encuentran lisozimas de tipo-c y tipo-g que presentan características únicas, lo que sugiere el importante papel de la lisozima en el sistema de defensa de estos organismos. La totoaba (*Totoaba macdonaldi*) es una especie endémica del Golfo de California, México, y en peligro de extinción. La acuicultura de la totoaba exige más investigación sobre su sistema inmunológico para desarrollar estrategias inmunoprolácticas y mejorar las condiciones de cultivo para su cría a gran escala. En este estudio, se clonaron y caracterizaron los genes que codifican a las lisozimas de tipo-c (TmLyzc) y tipo-g (TmLyzg) de totoaba. Las secuencias de ADNc de longitud completa de TmLyzc y TmLyzg fueron de 432 pb y 582 pb, que codifican polipéptidos de 143 y 193 aminoácidos, respectivamente. Las secuencias de aminoácidos mostraron gran identidad (90-60%) y estrechas relaciones filogenéticas con lisozimas de peces y vertebrados superiores, así como dominios estructurales y funcionales típicos de la superfamilia de lisozimas. El análisis de expresión por qRT-PCR reveló que TmLyzc y TmLyzg se expresan principalmente en el estómago, ciego pilórico y corazón. Estos hallazgos sugieren que las lisozimas tienen un papel importante en la defensa de la totoaba contra las infecciones bacterianas y también pueden estar involucradas en la digestión. En la segunda etapa de esta investigación, TmLyzc y TmLyzg se expresaron de forma recombinante en *E. coli* BL21 (DE3) a 25 °C con 0.5 mM IPTG, se obtuvieron como cuerpos de inclusión, se desnaturalizaron y se sometieron a experimentos de replegamiento *in vitro* utilizando el sistema Pierce® Protein Folding. El replegamiento de ambas lisozimas no produjo enzimas activas, debido posiblemente a que las proteínas se plegaron parcialmente o porque su sitio activo se vio obstaculizado de alguna manera por la presencia de las etiquetas de fusión. Aunque las condiciones de replegamiento probadas resultaron prometedoras, se requiere una mayor optimización para recuperar a las enzimas activas.

INTRODUCTION

Lysozyme is an iconic enzyme in the legacy of protein science since it was first isolated and named after Alexander Fleming in 1922. It was the first enzyme whose three-dimensional structure was solved using X-ray crystallography by David C. Phillips in 1965. It was also the first enzyme whose complete amino acid sequence was known (Canfield, 1963). All of this led to lysozyme being the first enzyme to have a suggested chemical mechanism to explain enzyme catalysis (Dahlquist *et al.*, 1969; Vocadlo *et al.*, 2001). Although it was later revisited, this discovery supported the importance of the relationship between protein structure and function. Moreover, recent findings have shown that lysozyme retains its antibacterial activity even when the amino acid residues critical for catalysis at the active site are replaced through site-directed mutagenesis. These results open new debates in lysozyme function and broadens the horizons to new research fields (Ibrahim, 1998; Ibrahim, Thomas, *et al.*, 2001; Ibrahim *et al.*, 2005, 2011).

From the biological point of view, lysozyme is a cornerstone in the innate immune system and one of the most essential molecules in protecting against bacterial pathogens. Lysozyme (EC 3.2.1.17) catalyzes the hydrolysis of the β -(1, 4) glycosidic bonds between N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG) present in the peptidoglycan layer of bacterial cell walls, causing the lysis of bacteria (Saurabh and Sahoo, 2008). It is also involved in other defense mechanisms and can also contribute to the resolution of inflammation (Saurabh and Sahoo, 2008; Ragland and Criss, 2017). Lysozyme is ubiquitous and distributed in diverse organisms, including animals, plants, fungi, bacteria, and phages. In the animal kingdom, three types of lysozymes are found, named c-type, g-type and i-type according where first discovered: chicken, goose or invertebrates, respectively (Callewaert and Michiels, 2010).

Lysozyme has been widely studied as a component of the innate immune system in fish. It is expressed in lymphoid tissues and also present in sites prone to bacterial invasion (Saurabh and Sahoo, 2008). In teleost fish species, c-type and g-type lysozymes genes have been identified and characterized (Irwin and Gong, 2003; Li *et al.*, 2021). Lysozyme

expression is an indicator of the proper functioning of the innate immune system and the stress level in farmed fish, suggesting an important role of this protein in these organisms' defense system (Hikima *et al.*, 2003).

In this work, I studied the c-type and g-type lysozymes expressed in totoaba (*Totoaba macdonaldi*) fish, which is a critically endangered species endemic to the central and northern Gulf of California, in Mexico. Totoaba is a demersal teleost fish and the largest member of the *Sciaenidae* family, also known as croakers. It can live up to 50 years and reach up to 2 m long and 100 kg of weight (Cisneros-Mata *et al.*, 1997). This species has great potential for aquaculture in northern Mexico for its high value. Formerly, its commercial fishery used to be one of the most prominent economic activities until it was banned in 1975 due to environmental protection by the U.S. and Mexican governments (Findley, 2010; Valenzuela-Quiñonez *et al.*, 2015). Nowadays, aquaculture represents a viable approach to recover totoaba's natural population, which is still facing technical challenges including disease control. In this sense, basic research on genes associated with the innate immune system of totoaba, such as lysozyme, will give us insight into this species' early defense mechanisms against pathogens relevant in aquaculture.

This doctoral dissertation project was divided into two parts: the first part describes the cloning and analysis of both lysozymes genes. The second part describes the recombinant expression and refolding of the lysozyme proteins. In the first part of this project, the full-length cDNA sequences of *Totoaba macdonaldi* c-type and g-type lysozyme were described. The characterized genes are homologous to other c-type or g-type lysozymes from other fish and higher vertebrate species. They also contain conserved structural and functional domains present in the β -glycosyl-hydrolases. Also, the expression of both genes in spleen, kidney, pyloric caeca, stomach, heart, and brain is constitutively at the mRNA level. Furthermore, their expression showed higher levels in specific tissues for each lysozyme. It reflects the importance of its function on innate immunity. In the second part, the c-type and g-type lysozymes of *Totoaba macdonaldi* were overexpressed in *E. coli*, isolated from inclusion bodies, and subjected to *in vitro* refolding experiments. However, refolding both lysozymes did not yield active enzymes,

either because the proteins were kinetically trapped in a folding pathway intermediate, or because their active sites were somehow hindered by the presence of fusion tags. Even though the tested refolding conditions were promising, further experiments are required to recover enzyme activity.

The findings described in this project suggest that the lysozymes genes cloned from totoaba play important roles in defense against bacterial infections and also may be involved in digestion. In the future, this knowledge will be crucial for establishing sustainable long-term cultures for the rearing of totoaba, aimed to improve farming conditions and develop diagnosis and immunoprophylactic strategies.

BIBLIOGRAPHIC BACKGROUND

Overview of Fish Innate Immune System

The immune response in vertebrate animals is divided into innate and adaptive systems, with the first as initial and non-specific response to an infection (Smith *et al.*, 2019), including physical barriers, cellular processes, and humoral components including proteins and enzymes (Riera Romo *et al.*, 2016). If after displaying an innate immune response, a pathogen is still standing, then the adaptive immune system is activated, generating a highly specific response, providing long-term immunity (Iwasaki and Medzhitov, 2010; Netea *et al.*, 2019). Innate immunity is ancient in origin, emerging > 600 million years ago approximately, while some components of the adaptive immunity are comparatively newer, from ~ 450 million years ago in the first jawed vertebrates (Gnathostomes) (Flajnik and Kasahara, 2010; Buchmann, 2014).

Though the mechanisms of innate immunity are well conserved and ancient than the adaptive system ones, this system is plastic and continuously evolving. In other words, the innate immune has been working constantly under natural selection to counteract and avoid infections from causing disease (Parham, 2003). Understanding the proper response of the vertebrate immune defense requires the analysis of the evolution of the first jawed vertebrates (Gnathostomata). In this sense, fish are used as model systems to reconstruct the evolutionary history and discover the conserved elements of the vertebrate immune defense, as they occupy a crucial role in the establishment of basic immune functions (Figure 1) (Sunyer, 2013; Nakanishi *et al.*, 2018; Smith *et al.*, 2019).

In fish, the innate immune system is considered as the main response attacking pathogens due to adaptive immune system limitations (Magnadóttir, 2006; Uribe *et al.*, 2011; Rebl and Goldammer, 2018). Innate immunity include three defense mechanisms, comprising physical barriers, cellular components and humoral responses (Magnadóttir, 2006). These defense mechanisms are summarized in Figure 2, with emphasis in teleost

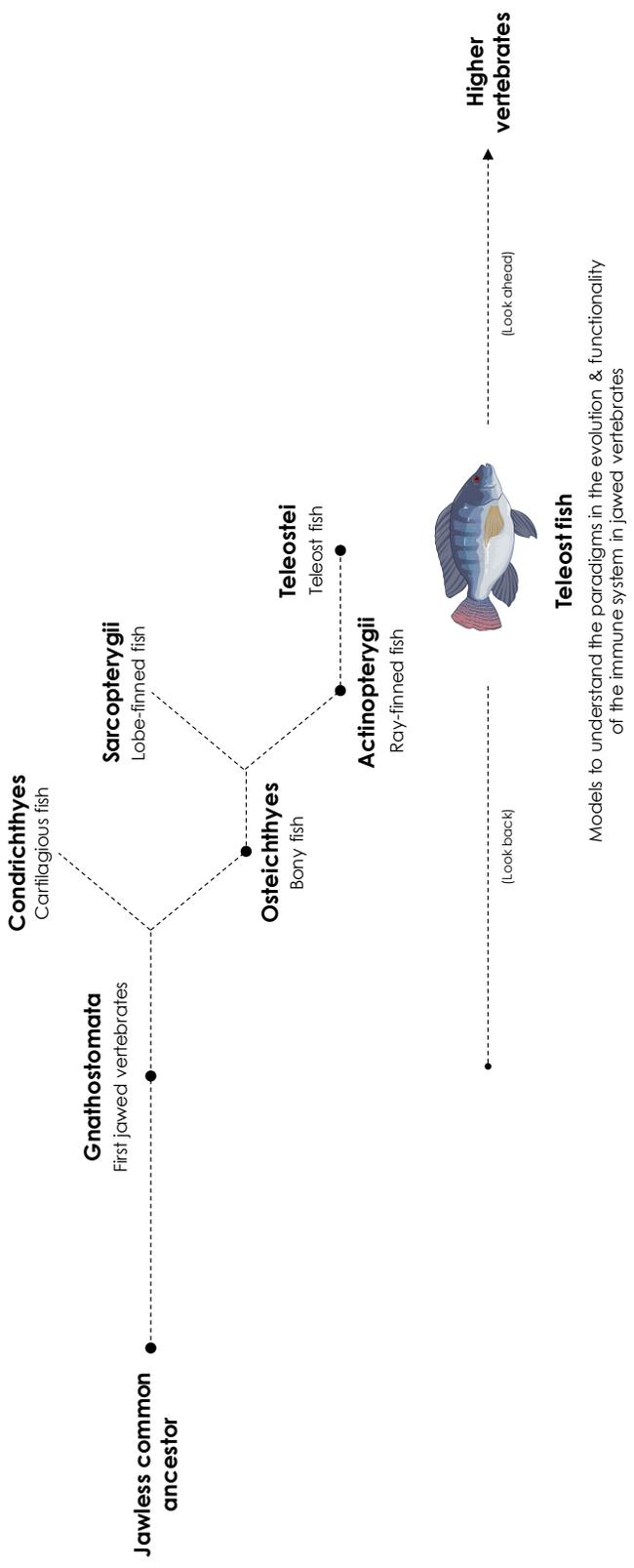


Figure 1. Teleost fish as models to understand the evolution and function of the immune system in vertebrates. The first jawed vertebrates (Gnathostomes) diverged from a jawless common ancestor with the lineage leading to other bony vertebrates. The key branches of Gnathostomes are Chondrichthyes and Osteichthyes. The major class of bony fishes are the Actinopterygii (ray-finned fish), of which 96% are from the infraclass Teleostei (teleost fish) (Near et al., 2012).

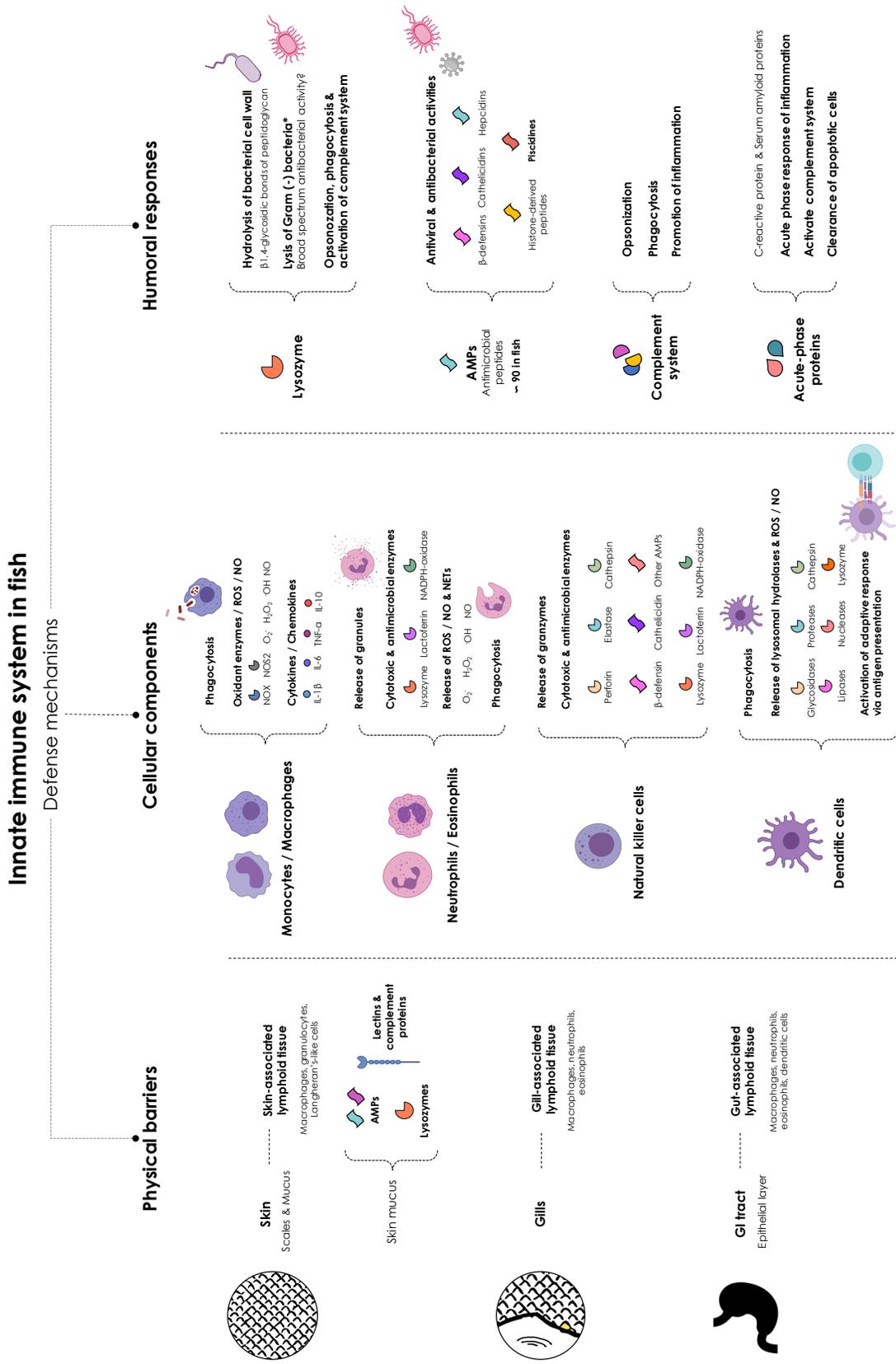


Figure 2. Overview of the defense mechanisms involved in the innate immune system of teleost fish.

fish innate immunity (as described from now on in the text). Fish inhabit an aquatic environment since early developmental stages and hence always exposed to potential pathogens. The earliest defense mechanisms in the innate immune system of these organisms are physical barriers, including the scales and mucus of the skin preventing pathogen entry, the gills, and the epithelial cells at the gastrointestinal tract (Magnadóttir, 2006).

The skin of teleost fish contains skin-associated lymphoid tissue (SALT) which are multiple immune cells like granulocytes, macrophages, Langerhans-like cells, secretory cells, and lymphocytes B and T (Ángeles Esteban, 2012; Xu *et al.*, 2013). Moreover, secreted skin mucus acts as a physical and chemical barrier, including lectins, lysozymes, complement proteins, and antimicrobial peptides that play an important role in neutralizing pathogens (Fast *et al.*, 2002; Xu *et al.*, 2013). The gills are another physical barrier, containing gill-associated lymphoid tissue (GIALT) that consist of both innate and adaptive immune cells like macrophages, neutrophils, eosinophils and lymphocytes (Haugarvoll *et al.*, 2008; Salinas, 2015). Furthermore, the gut-associated lymphoid tissue (GALT) in the epithelial layer of the gastrointestinal tract (especially in the posterior segment of the intestine) will respond if a pathogen is ingested. GALT contains innate and adaptive immune cells like macrophages, mast/eosinophilic granule cells, dendritic cells, and B and T cells (Salinas, 2015; Tafalla *et al.*, 2016).

The innate immune system, both humoral and cellular, is faced only if a pathogen breaks the physical barriers. The cellular components of the fish's innate immune defense include monocytes/macrophages, eosinophilic granule cells, neutrophils, natural killer (NK) cells, and dendritic cells (DC). The primary sites for leukocyte production are the anterior (head) kidney and thymus in bony fish (Smith *et al.*, 2019). The innate immune cell will recognize small conserved molecular motifs within the pathogen, resulting in cell activation. Once activated, the immune cells will participate in several responses that mainly depend on cell subtype to counteract the infection. These responses include phagocytosis and degradation of the pathogen, production of various cytokines and

activation of the adaptive immune system via antigen presentation, along with cytokine stimulation among others (Smith *et al.*, 2019).

Macrophages and neutrophils are among the first cells arriving in response to an infection. Macrophages play key roles during inflammation, pathogen infection, and tissue homeostasis. In teleost fish, macrophages can attack pathogens by phagocytosis, by producing reactive oxygen species (ROS) and nitric oxide (NO), and by releasing inflammatory cytokines and chemokines (Hodgkinson *et al.*, 2015; Grayfer *et al.*, 2018). Moreover, subpopulations of macrophages with different functionalities have been identified in teleost fish. These include the classically activated macrophages (M1), that produce ROS, NO, and pro-inflammatory cytokines (TNF- α and IL-1 β), and the alternative activated macrophages (M2), associated to immunosuppression, wound repair and production of anti-inflammatory cytokines (IL-10) (Hodgkinson *et al.*, 2015; Grayfer *et al.*, 2018). Furthermore, macrophages are one type of antigen-presenting cells for the adaptive immune system (Smith *et al.*, 2019). In addition, neutrophils play important roles to the innate defense against pathogens with strong responses to microbial infections through several intracellular and extracellular mechanisms (Havixbeck and Barreda, 2015). Neutrophils respond by releasing granules with cytotoxic and antimicrobial enzymes, producing ROS and NO, releasing of neutrophil extracellular traps, and by phagocytosis (Neumann *et al.*, 2001; Havixbeck and Barreda, 2015).

Furthermore, NK cells are the main players in cell-mediated cytotoxicity of the innate immune response to counteract intracellular pathogens and eliminate infected cells. Two types of NK cell are found in bony fish, named the non-specific cytotoxic cells and the NK-like cells (Fischer *et al.*, 2013). The cytotoxic activity of NK cells is mediated by granule exocytosis, in which proteins (perforin, granulysin and NK-lysin) form pores in the outer membrane of pathogen-infected cells, thus allowing the entry of granzymes into the cytoplasm leading to cell death by apoptosis (Fischer *et al.*, 2013). Finally, DC are specialized antigen-presenting cells linking the innate and adaptive immune responses. Their role is to process and present foreign antigens through the major histocompatibility complex (class I or II) at the cell surface to prime and stimulate T cells (Mellman and

Steinman, 2001; Savina and Amigorena, 2007; Soletto *et al.*, 2018). In teleost fish, some subsets of DC have been identified to undertake specific DC-activities such as phagocytic and strong T-cell activating capacities, however, more research in this area is needed (Bassity and Clark, 2012; Granja *et al.*, 2015; Soletto *et al.*, 2018).

Recognition of Non-self and Activation of the Innate Immune System

In general, the basis of the innate immune response is the recognition of pathogen-associated molecular patterns (PAMPs), which are well conserved composites of molecular signatures found in pathogenic microorganisms, through a system of pattern recognition receptors (PRRs) within the host immune cells (Mogensen, 2009). Examples of PAMPs include bacterial-derived lipopolysaccharide and peptidoglycan, viral RNA, bacterial DNA, or damage-associated molecular patterns which are released from stressed or damaged cells. In fish, the major groups of identified PRRs include Toll-like receptors (TLRs), nucleotide-binding oligomerization domain-like receptors (NLRs) and retinoic acid inducible gene I-like receptors (Li *et al.*, 2017; Sahoo, 2020). The PRRs are composed by a domain that recognizes the PAMP and connects to another domain which in turn binds with downstream signaling molecules (Mogensen, 2009; Zhang *et al.*, 2014). The recognition of PAMPs triggers signaling cascades of transcription factors related to the immune response, mainly the nuclear factor NF- κ B and interferon regulatory factors (IRF3). This leads to the production of proinflammatory cytokines and chemotactic cytokines (IL-1 β , IL-6, IL-8, IL-10, TNF α) and a battery of antimicrobial compounds, including oxidant enzymes involved in respiratory burst (NADPH-oxidase and nitric oxide synthase 2), antimicrobial peptides (β -defensins, cathelicidins, hepcidins, histone-derived peptides and fish-specific piscidins) and lytic enzymes (i.e., cathepsin, lactoferrin, lysozyme, etc.) (Figure 2 and Figure 3) (Magnadóttir, 2006; Magnadóttir, 2010). These molecules are part of the humoral innate response, produced by innate immune cells to counteract infections caused by pathogenic microorganisms through different

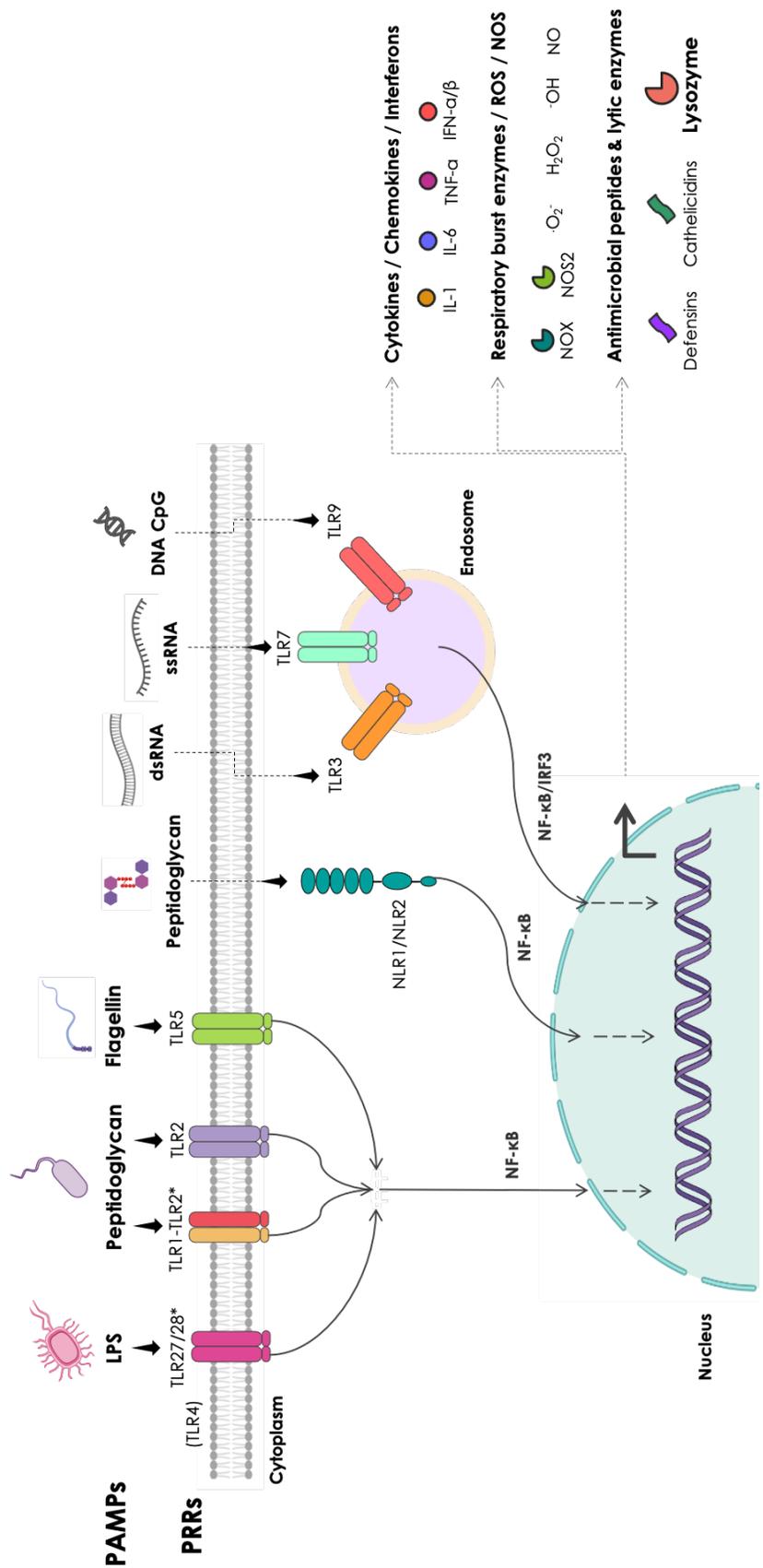


Figure 3. Recognition of non-self and activation of the innate immune response. The numbering of TLRs and their corresponding ligands are according to those reported in teleost fish species. TLR1-3, 5, 7-9 are structurally and functionally similar to mammalian counterparts; TLR6 and TLR10 mammalian homologues not found in teleosts; TLR5s, 18-20, 25-28 are teleost-specific (Smith *et al.*, 2019).

mechanisms, including inflammation, phagocytosis, and direct microbicidal effect. The humoral components in fish comprise the complement system, the acute phase proteins, antimicrobial peptides and lysozyme (Smith *et al.*, 2019). For further review, the reader is kindly referred to (Katzenback, 2015; Zou and Secombes, 2016; Bruce and Brown, 2017; Smith *et al.*, 2019; Semple and Dixon, 2020).

Lysozyme: Structure and Function

Lysozyme, described for the first time by Alexander Fleming (Fleming, 1932), is a cornerstone in the innate immune system. This enzyme is one of the most important molecules of the innate immune defense aiming to protect against bacterial pathogens (Ragland and Criss, 2017). Lysozyme (EC 3.2.1.17) is a bacteriolytic enzyme that catalyzes the hydrolysis of the β -(1, 4) glycosidic bonds between N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG), the main components of the peptidoglycan layer of bacterial cell wall (Saurabh and Sahoo, 2008). This protein is also involved in other defense mechanisms such as opsonization, phagocytosis, activation of the complement system, and stimulation of pro-inflammatory responses (Saurabh and Sahoo, 2008). More recent findings reveal that lysozyme can also contribute to the resolution of inflammation, acting as an important immunomodulatory agent (Ragland and Criss, 2017). Lysozyme is ubiquitously distributed among living organisms, including animals, plants, fungi, bacteria and even bacteriophages (Callewaert and Michiels, 2010). Three types of lysozymes are found in the animal kingdom: c-type (conventional or chicken-type), g-type (goose-type), and i-type (invertebrate-type) (Callewaert and Michiels, 2010).

In general, c-type and g-type lysozymes are basic proteins due to their high isoelectric point (pI) values (~ 9.53 or lower), while the pI of i-type lysozymes are quite variable. This might be attributed to the diverse biological functions of these lysozymes, including host defense and digestion (Callewaert and Michiels, 2010). The c-type and i-type lysozymes are much smaller than the typical g-type lysozymes, with molecular weights of

~11-15 kDa (c-type and i-type lysozymes) and ~20-22 kDa (g-type lysozymes), respectively (Callewaert and Michiels, 2010). The primary structure of c-type and i-type lysozymes consist of ~129 and ~123 amino acid residues (mature peptides), respectively, whereas g-type lysozymes are larger proteins composed of ~185 amino acids (Callewaert and Michiels, 2010). The presence of cysteine residues and disulfide bonds in c-type and i-type lysozymes is well preserved, while these characteristics vary among g-type lysozymes. C-type and i-type lysozymes contain four and seven intact disulfide bonds, respectively (8 and 14 cysteine residues, respectively), whereas g-type lysozymes show a striking difference in cysteine content and thus, the number of disulfide bridges. In birds and mammals, g-type lysozymes form two intramolecular disulfide bonds in the mature proteins, formed by four conserved cysteine residues (Irwin and Gong, 2003; Pooart *et al.*, 2004).

Conversely, the g-type lysozymes found in fish have variable number of cysteine residues. As in Japanese flounder and orange-spotted grouper where no cysteine residues are present (Hikima *et al.*, 2001; Yin *et al.*, 2003), whereas in carp and salmon there is only one cysteine residue (Savan *et al.*, 2003; Kyomuhendo *et al.*, 2007), or like zebrafish with two cysteines but with no potential to form a disulfide bond (Irwin and Gong, 2003). The absence of intramolecular disulfide bonds is a common feature among g-type lysozymes in fish (Irwin and Gong, 2003). Concerning its cellular localization, c-type and i-type lysozymes are recognized as extracellular proteins as they invariably possess a signal peptide sequence in their primary structure (Qasba *et al.*, 1997; Goto *et al.*, 2007; Callewaert and Michiels, 2010; Van Herreweghe and Michiels, 2012). Conversely, the presence of signal sequence within g-type lysozymes is variable. Most g-type lysozymes from vertebrates are known as extracellular proteins, with a clear exception from fish orthologues, in which no signal peptide is identified within its primary structure and are recognized as intracellular proteins (Irwin and Gong, 2003). This might have implications in a highly evolved and richly diversified innate immune system in teleost fish, with the capability to counteract both extracellular and intracellular pathogens by these enzymes (Lieschke and Trede, 2009; Li *et al.*, 2021). All these features are summarized in Table I.

Table I. Summary of the general features of lysozymes in the animal kingdom.

General characteristics	Lysozyme type		
	c-type	g-type	i-type
Amino acids (mature peptide)	~ 129	~ 185	~ 123
Molecular weight	~ 11-15 kDa	~ 20-22 kDa	~ 11-15 kDa
Isoelectric point	~ 9.53 or lower	~ 9.53 or lower	Variable
Cysteine residues	8	Variable 4 – Mammals & birds 0-2 – Fish	14
Disulfide bonds	4	Variable 2 – Mammals & birds 0 – Fish	7
Signal peptide	Yes	Variable Yes – Mammals & birds No – Fish	Yes
Cellular localization	Extracellular	Variable Extracellular – Mammals & birds Intracellular – Fish	Extracellular
Active site residues *	Glu ³⁵ , Asp ⁵²	Glu ⁷³ , Asp ^{97*}	Glu ¹⁸ , Asp ³⁰
Catalytic mechanism	Retaining β -glycosydases	Inverting β -glycosydases	Retaining β -glycosydases

*Active site residues correspond to hen-egg white lysozyme (c-type lysozyme), goose-egg white lysozyme (g-type lysozyme) and tapes japonica (i-type lysozyme).

Despite the fact that the different types of lysozymes present low identity in their primary structures (16-24%) and differ in their biochemical characteristics and catalytic mechanisms, they all share a very similar three-dimensional structure, which strongly suggest a common evolutionary origin (Callewaert and Michiels, 2010; Wohlkönig *et al.*, 2010). The overall three-dimensional structure of lysozymes is an α/β structure with a deep cleft that contains the active and substrate binding sites and separates the larger α -helical domain from the small β -sheet domain (Wohlkönig *et al.*, 2010). The crystal structures of hen-egg white (HEWL, c-type lysozyme) and goose-egg white (GEWL, g-type lysozyme) lysozymes are excellent representatives of the main structural elements present in the three-dimensional structures of lysozymes from vertebrates.

In HEWL (representative of c-type lysozymes), the α -helical domain (Lys1 to Asn39 and Ser86 to Leu 129) includes both the N- and C-terminal segments of the protein and contains four α -helices (α 1, Arg5 to Arg14; α 2, Leu25 to Ser36; α 3, Thr89 to Asp101; α 4, Val109 to Arg114) and a small 3^{10} helix (Val120 to Trp123). The β -domain (Thr40 to Leu84) contains a three-stranded antiparallel β -pleated sheet (β 1- β 3, Thr43 to Ser60) with a long loop (Arg61 to Ile78) and a single-turn 3^{10} helix (Pro79 to Leu84). Residues that line the cleft include the β -sheet residues Thr43, Asn44, Asn46, Asp52, and Leu56-Asn59, Glu35 in α 2-helix, the loop connecting helices α 3 and α 4 (residues Ile98 and Asp101-Asn103), residues Ala107-Ala110 in α 4 helix, and residues Trp62, Trp63, and Arg73. The native structure of HEWL is stabilized by four disulfide bonds: two of them in the α -domain, one in the β -domain, and another connecting the two domains (Figure 4) (Blake *et al.*, 1965; Artymiuk *et al.*, 1982; McKenzie and White, 1991). This general architecture is conserved in the crystal structure of rainbow trout c-type lysozyme (PDB code: 1LMN) (Karlsen *et al.*, 1995).

Concerning GEWL (representative of g-type lysozymes), the α -helical domain (Arg1 to His75 and Gly109 to Tyr185) covers both the N- and C-terminal segments of the protein and contains seven α -helices (α 1, Ser17 to Gly25; α 2, Val31 to Met45; α 3, Lys49 to Lys58; α 4, Pro63 to Glu73; α 5, Glu110 to Lys130; α 6, Lys136 to Asn148; α 7, Tyr169 to Gln184), and a small 3^{10} helix (Asn7 to Ile11). The β -domain (Trp84 to Gln104) comprises

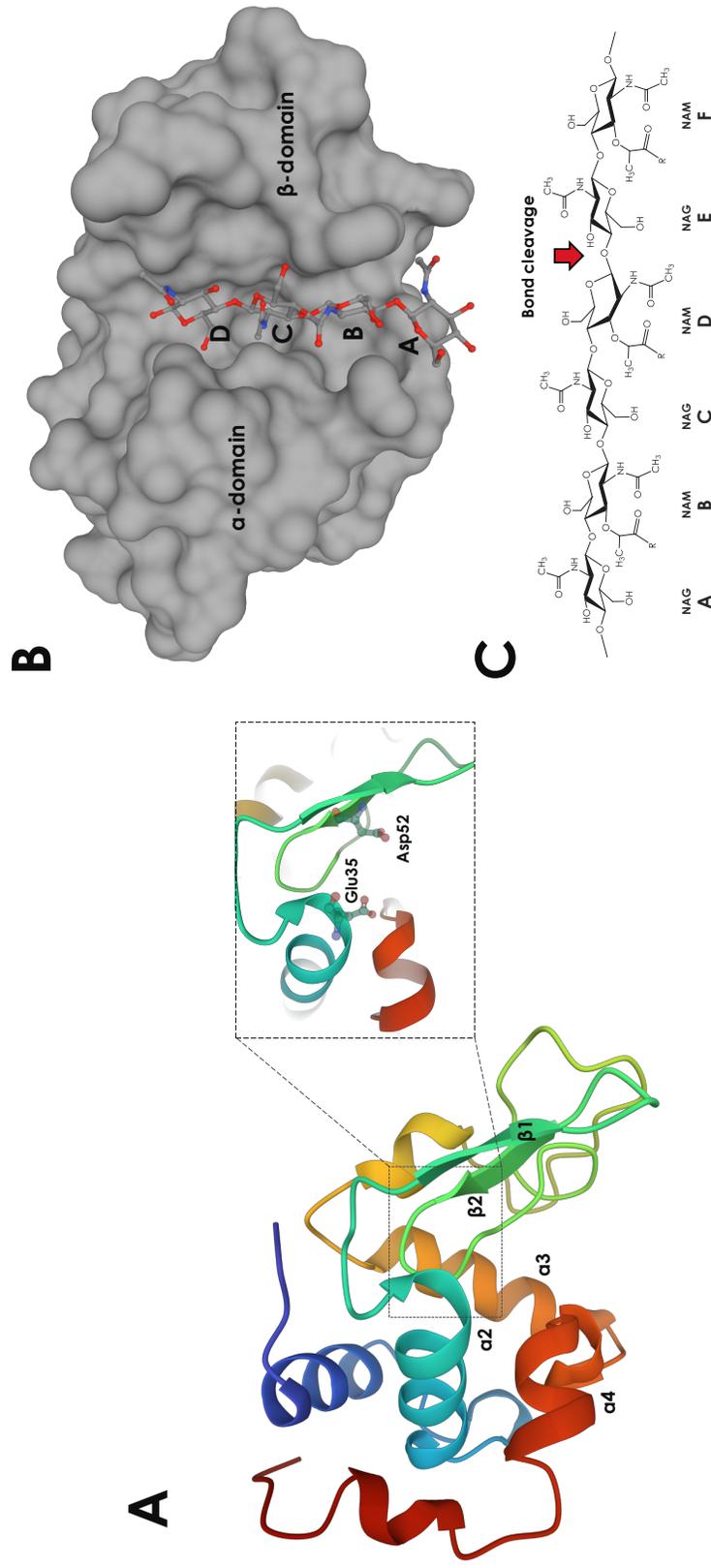


Figure 4. Three-dimensional structure of HEWL and binding cleft subsites. (A) Catalytic residues and main structural elements that shape the substrate binding site in HEWL. (B) Substrate binding cleft (subsites A-F) with bound tetrasaccharide (GlcNAc₄). (C) Natural substrate of lysozymes with marked A-F subsites (c-type lysozymes) and cleavage of the β -1,4 glycosidic bond between subsites D (NAM) and E (NAG). PDB codes for crystal structures of hen-egg lysozyme used for this illustration 1AKI and 1LZC.

a three-stranded antiparallel β -sheet (β 1, Gly90 to Gly92; β 2, Asn95 to Lys98; β 3, His101 to Gln104) and a short loop (Gly105 to Asn108) (Figure 5). The residues that comprise the binding cleft include the α 1-helix residues Pro23 and Glu24, Arg17 and Glu73 in α 4-helix, the loop connecting α 4-helix and β 1-sheet (His75, Lys78 and Gly85), the β -domain residues Asp86-Gly90, Gln85, Asp97, and Arg99-His101, the α 5-helix residue Ile119, Tyr147 and Asn148 in α 6-helix, residues in the loop connecting α 6 and α 7 helices (Ala51, Gly152, Thr165 and His166), and Tyr169 in α 7-helix. The native structure of GEWL is stabilized by two disulfide bonds: Cys18-Cys29, that connects the N-terminus with a loop between α 1 and α 2 helices, and Cys4-Cys60, that connects the N-terminal with the loop between α 3 and α 4 helices (Weaver *et al.*, 1995; Honda and Fukamizo, 1998; Pooart *et al.*, 2004; Hirakawa *et al.*, 2008; Kawamura *et al.*, 2008). These bonds are highly conserved in g-type lysozymes from birds and mammals, but are absent in their counterparts from fish species (Irwin and Gong, 2003; Pooart *et al.*, 2004). However, the general architecture of g-type lysozymes is conserved in the structure of Atlantic salmon (PDB: 4G9S) and Atlantic cod (PDB: 3GXX) lysozymes (Helland *et al.*, 2009; Leysen *et al.*, 2013).

C-type and g-type lysozymes present conserved structural features found in the members of the lysozyme superfamily. In particular, these conserved structural elements are α 2- α 4 helices in HEWL and the α 4- α 6 helices in GEWL, along with the three-stranded β -sheet, that altogether shape the substrate binding site in these lysozymes (Monzingo *et al.*, 1996; Wohlkönig *et al.*, 2010). As stated above, the natural substrate of lysozymes is the peptidoglycan cell wall of Gram-positive bacteria, which is composed of crosslinked oligosaccharides consisting of alternating NAM and NAG residues. Lysozymes cleave this substrate by first binding the polysaccharide in the enzyme's cleft (substrate binding site) located in the interface of the α - and β -domains. The crystal structures of HEWL and GEWL reveal that the substrate binding cleft can accommodate six sugar residues (GlcNAc₆ oligomer), named as subsites A to F in HEWL, and B to G in GEWL (Hirakawa *et al.*, 2006, 2008).

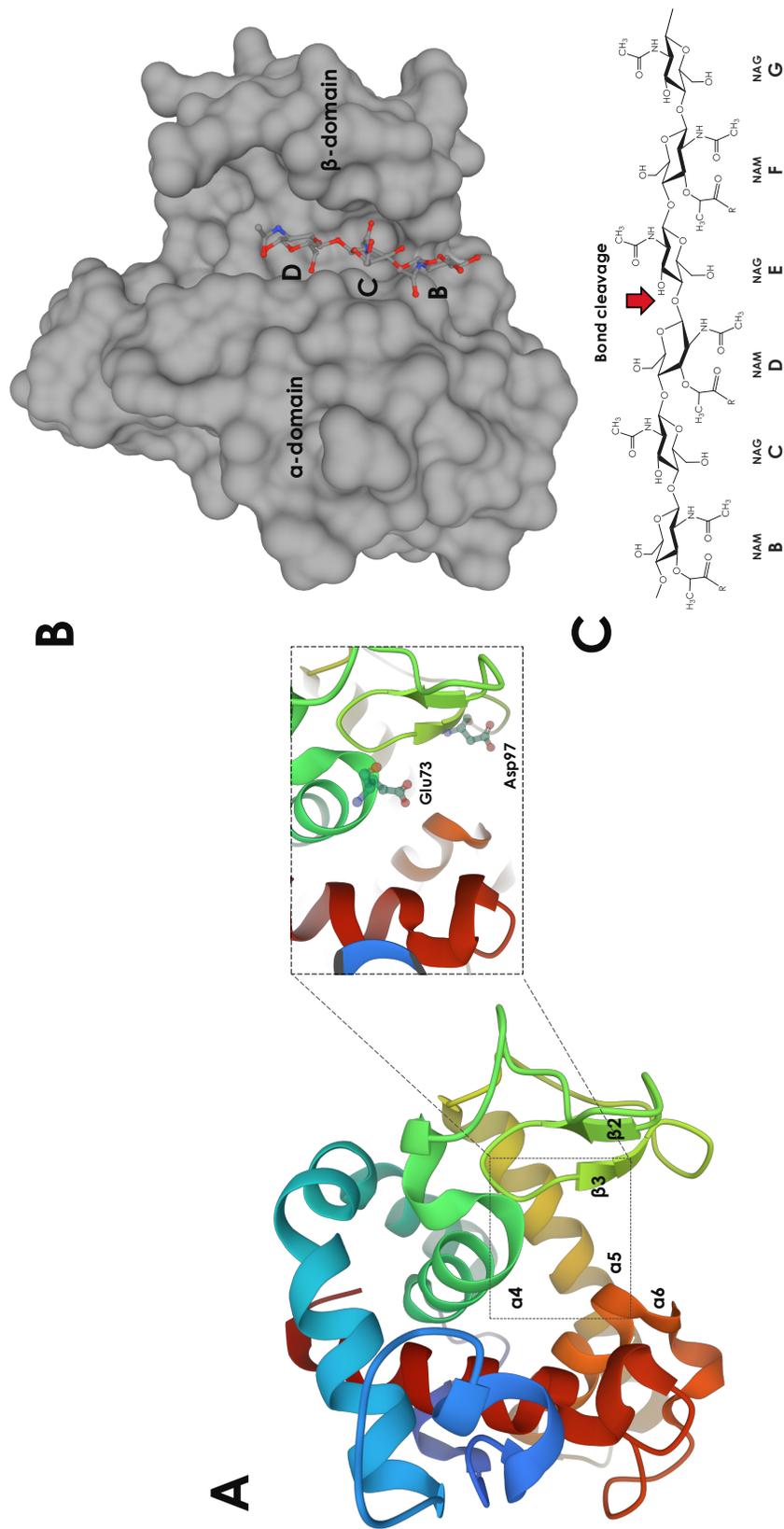


Figure 5. Three-dimensional structure of GEWL and binding cleft subsites. (A) Catalytic residues and main structural elements that shape the substrate binding site in GEWL. (B) Substrate binding cleft (subsites A-D) with bound trisaccharide (GlcNAc₃). (C) Natural substrate of lysozymes with marked B-G subsites (g-type lysozymes) and cleavage of the β -1,4 glycosidic bond between subsites D (NAM) and E (NAG). PDB codes for crystal structures of goose-egg lysozyme used for this illustration 153L and 154L.

The residues that constitute the A-F subsites in the binding cleft of HEWL structure are: A (Arg73, Asp101-Asn103), B (Asp101 and Asn103), C (Asn59, Trp62, Trp63, Asn103, Ala107), D (Glu35, Asn46, Asp52, Leu56, Gln57, Ala107, Val109), E (Glu35, Asn44-Asn46, Asp52, Gln57, Ala110), F (Lys33, Phe34, Glu35, Asn37, Arg45, Asn46, Arg114) (Hirakawa *et al.*, 2006). In GEWL, the residues that form the B-G binding subsites are: B (Ser100, His101, Ala151, Gly152), C (Asp97, Arg99, Ser100, His101, Tyr147, Asn148), D (Glu73, Arg87, Asn89, Gly90, Gln95, Asp97, Arg99, Tyr147, Asn148, Thr165, Tyr169), E (Arg72, Glu73, Gly85, Asp86, Arg87, Gly88, Asn89, Gly90, Gln95, Asn148, Thr165, His166), F (Glu24, Arg72, Glu73, His75, Asp86, Arg87, Gln95, His166), G (Pro23, Glu24, His75, Lys78, Asp86) (Hirakawa *et al.*, 2008). Furthermore, although the tridimensional structures of c-type and g-type lysozymes are quite similar, only the central α -helix (α 2 in HEWL and α 4 in GEWL) and the β -hairpin (β 2- β 3 strands in both lysozymes) are the invariant structural motifs found in the lysozyme superfamily, and constitute the core that shapes the enzymes' active site (Wohlkönig *et al.*, 2010). The central α -helix contains the catalytic Glu35 (HEWL) or Glu73 (GEWL), while the β -hairpin contains the second catalytic residues Asp52 (HEWL) or Asp97 (GEWL). In both types of lysozymes, glutamate invariably acts as a general acid catalyst whereas aspartate acts as a general base (Zechel and Withers, 2000; Vocadlo *et al.*, 2001).

The Catalytic Mechanisms of C-type and G-type Lysozymes

Concerning their catalytic activity and mechanism, lysozymes are a group of β -glycosidases that catalyze the hydrolysis of β -1,4 glycosidic bonds between NAM and NAG or oligomers of NAG. The hydrolysis of the glycosidic bond can occur with one of two possible stereochemical outcomes: (1) net inversion or (2) net retention of anomeric configuration, which are carried out by two different mechanisms. The c-type lysozymes (HEWL) are known as retaining β -glycosidases, which exert their catalytic activity through a double displacement (two-step) mechanism. Conversely, g-type lysozymes (GEWL) perform their catalytic activity by a single displacement (single-step) mechanism

and are known as inverting β -glycosidases (Honda and Fukamizo, 1998; Zechel and Withers, 2000; Vocadlo *et al.*, 2001).

In the catalytic mechanism exhibited by HEWL (c-type lysozymes in general), the binding of substrate to the enzyme's catalytic cleft induces substrate strain leading to a distorted conformation of the substrate. In this form, the NAM residue at the D subsite shifts to a half-chair conformation, thus bending the oligosaccharide between subsites D and E where glycosidic bond cleavage occurs (Figure 4). In the first step of catalysis, the carboxyl group of Glu35 functions as a general acid catalyst to protonate the glycosidic oxygen, while Asp52 acts as a nucleophile (general base) to form a covalent glycosyl-enzyme intermediate. This first part of the reaction necessarily results in the net inversion of the anomeric configuration. In the second step, the enzyme carboxylate is then itself displaced from the glycoside-enzyme intermediate by an incoming water molecule, which attacks at the anomeric center and displaces the sugar. This second part of the reaction results in a second inversion restoring the original anomeric configuration (Figure 6) (Zechel and Withers, 2000; Kirby, 2001; Vocadlo *et al.*, 2001).

Concerning the catalytic mechanism of GEWL (g-type lysozymes in general), substrate distortion after binding to the enzyme's catalytic cleft is also observed in the NAM unit located at subsite D, thereby hydrolysis of the glycosidic bond also takes place between NAM and NAG units that occupy D and E subsites, respectively (Figure 5). Unlike c-type lysozymes, catalysis of GEWL occurs in a single step in which Glu73 acts as a general acid catalyst, protonating the glycosidic oxygen, while Asp97 acts as a general base to assist the attack of a nucleophilic water molecule at the anomeric center, which ultimately leads to the displacement of the sugar moiety. This reaction involves an oxocarbenium ion-like transition state and results in the net inversion of the anomeric configuration (Figure 7) (Honda and Fukamizo, 1998; Kuroki *et al.*, 1999; Zechel and Withers, 2000; Hirakawa *et al.*, 2008).

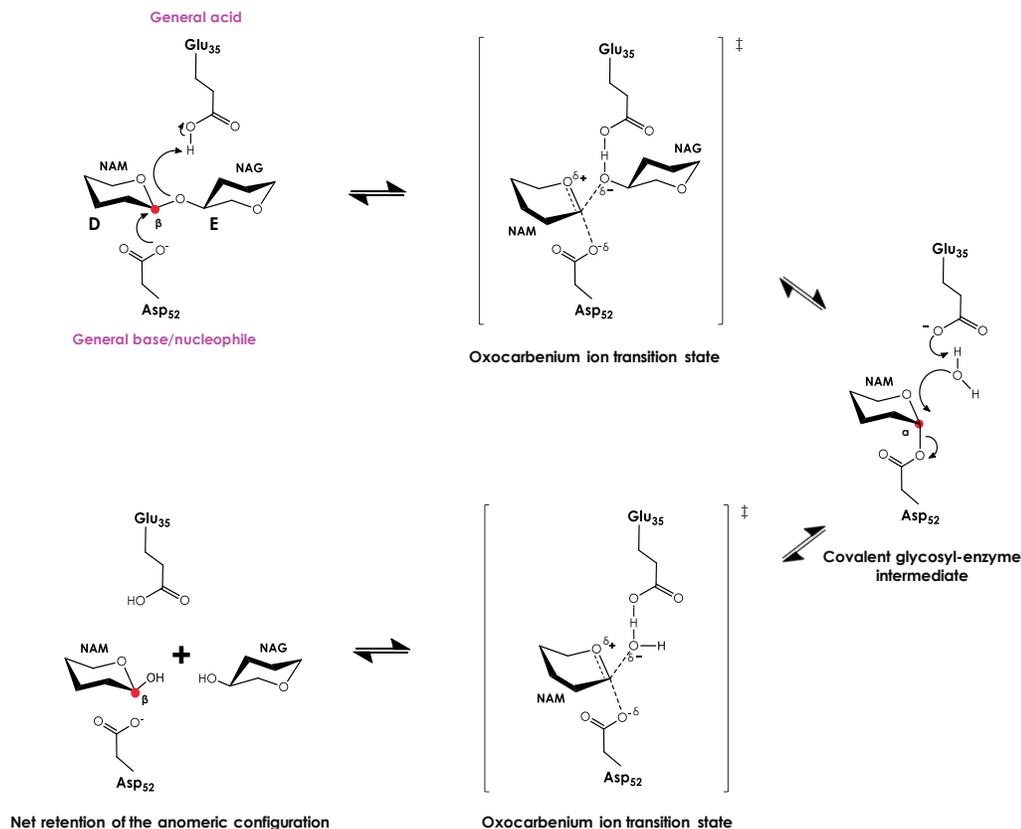


Figure 6. General mechanisms for retaining glycosidases: the catalytic mechanism of HEWL (c-type lysozymes – retaining β -glycosidases). Lysozymes exert their catalytic function through acid/base catalysis, in which the cleavage of the β -1,4 glycosidic bond between NAM (D subsite) and NAG (E subsite) residues requires substantial general acid catalytic assistance (Glu35 in HEWL), as well as a general base catalysis (Asp52 in HEWL) to assist the attack of a nucleophilic water molecule. In retaining glycosidases, the carboxyl groups of the acid/base catalysts are only 5.5 Å apart, consistent with a double-displacement mechanism involving a covalent glycosyl-enzyme intermediate. This two-step reaction occurs via transition states with substantial oxocarbenium ion character. Substrate distortion plays a critical role in assisting bond cleavage, as the active site of β -glycosidases seems to be structurally static but electronically dynamic. Conformational changes in the NAM residue positioned in the D subsite of HEWL permit an in-line attack of the enzymic nucleophile at the anomeric center, moves the substrate closer to the conformation of the oxocarbenium ion transition state, and places the glycosidic oxygen in an appropriate position for protonation by the general acid catalyst. Illustrations and information adapted from (Zechel and Withers, 2000).

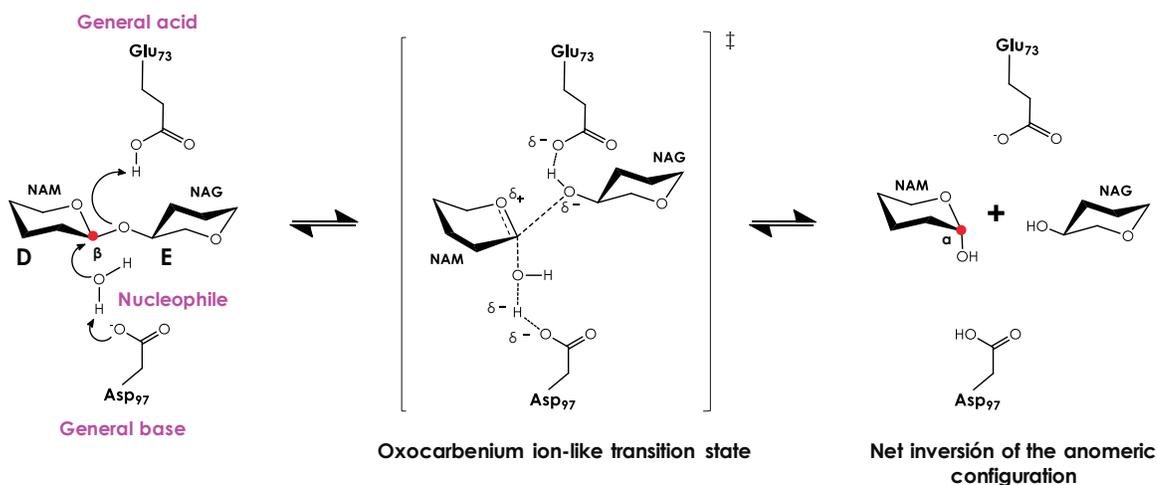


Figure 7. General mechanism for inverting glycosidases: the catalytic mechanism GEWL (g-type lysozymes – inverting β -glycosidases). Lysozymes exert their catalytic function through acid/base catalysis, in which the cleavage of the β -1,4 glycosidic bond between NAM (D subsite) and NAG (E subsite) residues requires substantial general acid catalytic assistance (Glu73 in GEWL), as well as a general base catalysis (Asp97 in GEWL) to assist the attack of a nucleophilic water molecule. In inverting glycosidases, the carboxyl groups of the acid/base catalysts are suitably placed ~ 10.5 Å apart, allowing enough space to bind the substrate and a water molecule. This reaction occurs via single-displacement mechanism, which involves an oxocarbenium ion-like transition state. Substrate distortion that occurs in NAM residue positioned in the D subsite of GEWL plays a critical role in assisting bond cleavage, as stated for HEWL. Illustrations and information adapted from (Zechel and Withers, 2000).

The Antibacterial Activity of Lysozymes Independent of their Catalytic Function

Studies on heat-denatured (partially or totally unfolded) and mutant lysozyme lacking muramidase activity have shown that the enzyme retains its inherent antibacterial activity against Gram-positive and even exhibits enhanced bactericidal action against Gram-negative bacteria. In this sense, the research has proposed that the antibacterial activity of lysozyme is independent of its catalytic activity but rather it might be associated with a structural phase transition, which would allow lysozyme to act as an antimicrobial peptide (Ibrahim *et al.*, 1996; Ibrahim, 1998; Düring *et al.*, 1999; Ibrahim, Matsuzaki, *et al.*, 2001). More recent studies have given insights into the structural elements in lysozyme that can be involved in its broad-spectrum antibacterial activity. In HEWL, a helix-loop-helix peptide (residues Glu87-Arg114) located at the upper lip of the active site cleft was microbicidal against both Gram-positive and Gram-negative bacteria (Ibrahim, Thomas, *et al.*, 2001). In particular, the N-terminal helix (residues Asp87-Ser100) of this α -helical hairpin peptide was specifically bactericidal to Gram-positive bacteria, such as *Micrococcus luteus*, *Bacillus subtilis*, *Staphylococcus aureus*, and *S. epidermidis*, whereas the C-terminal helix (residues Ala107-Arg114) exhibited broad-spectrum bactericidal effect against Gram-positive bacteria and Gram-negative bacteria (*E. coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *Salmonella typhimurium*). This study provides evidence that the α -helical hairpin peptide and its C-terminal helix domain can cross the outer membrane through a self-promoted uptake. This assembly promotes the formation of channels, damaging the inner membrane, killing Gram-negative bacteria (Ibrahim, Thomas, *et al.*, 2001). The α -helical hairpin structures are motifs commonly found in bactericidal and cytolytic pore-forming proteins (Engelman and Steitz, 1981; Srisailam *et al.*, 2000).

Furthermore, HEWL processed with pepsin generated peptides with bactericidal activity against several strains of Gram-positive (*S. aureus*, *S. epidermidis*, *B. subtilis* and *M. luteus*) and Gram-negative bacteria (*E. coli* K-12, *Salmonella enteritidis* and *Helicobacter pylori*) (Ibrahim *et al.*, 2005). They found that HEWL contains two cationic α -helical peptides (Lys1-Leu17 and Arg125-Leu129) within the bactericidal helix-loop-

helix domain (Lys1-Phe38), a peptide consisting of an α -helix (Asp18-Phe38), and two helix-sheet peptides (a helix-two-stranded β -sheet Asp18-Leu56/Arg114-Thr118, and a helix triple-stranded β -sheet Asp18-Trp62/Arg114-Thr118) that also exhibited bactericidal activity against *S. aureus* and *E. coli* in a dose-dependent manner and to different extents depending on their structure. These peptides can disrupt the outer membrane of Gram-negative bacteria via carpet-like and pore-forming mechanisms, through their multiple structural motifs (Ibrahim *et al.*, 2005).

Regarding GEWL, an α -helical peptide (Thr20-Tyr28) located at the outer surface of the active site (a part of the C-terminal α 1-helix and the loop that links α 1- α 2 helices) was obtained through enzymatic hydrolysis and showed broad spectrum antibacterial activity against *Vibrio cholerae* and *S. epidermidis* (Thammasirirak *et al.*, 2010). The results in this study suggest that this peptide binds to the outer membrane surface at bacterial cells wall, and then inserting into the lipid bilayer of the cytoplasmic membrane, which induces protrusion of the cell surface (Thammasirirak *et al.*, 2010). Similar results were also observed for peptides derived from human and mouse lysozymes (Nash *et al.*, 2006; Ibrahim *et al.*, 2011). All these findings strongly suggest that lysozymes possess non-enzymatic bactericidal/bacteriostatic domains in their primary structure that may be involved and can help explain the broad-spectrum antibacterial activity of these enzymes.

Lysozymes in Teleost Fish

Lysozyme is one of the most studied innate immune components in fish, which is highly expressed in lymphoid tissues and distributed in sites susceptible to bacterial invasion. Lysozymes are mainly found in neutrophils and monocytes, and to a lesser extent in macrophages of several tissues including kidney, liver, spleen, gills, gastrointestinal tract, and in skin mucus (Saurabh and Sahoo, 2008). In fish species, lysozyme expression is an indicator of the proper functioning of the innate immune system, suggesting the important role of this protein in the defense system of these organisms (Hikima *et al.*, 2003;

Callewaert and Michiels, 2010). In teleost fish species, c-type and g-type lysozymes genes have been identified and functionally characterized (Irwin and Gong, 2003; Irwin *et al.*, 2011; Li *et al.*, 2021). The antibacterial activity of c-type lysozymes has been described in several teleost species, such as Japanese flounder (*Paralichthys olivaceus*) (Minagawa *et al.*, 2001), rainbow trout (*Oncorhynchus mykiss*) (Grinde, 1989), taimen (*Hucho taimen*) (Li *et al.*, 2016), Asian seabass (*Lates calcarifer*) (Fu *et al.*, 2013), among others. Likewise, the activity of g-type lysozymes has also been described in Atlantic salmon (*Salmo salar*) (Kyomuhendo *et al.*, 2007), Atlantic cod (*Gadus morhua*) (Larsen *et al.*, 2009; Seppola *et al.*, 2016), Japanese flounder (*P. olivaceus*) (Hikima *et al.*, 2001), orange-spotted grouper (*Epinephelus coioides*) (Yin *et al.*, 2003), mandarin fish (*Siniperca chuatsi*) (Sun *et al.*, 2006), European seabass (*Dicentrarchus labrax*) (Buonocore *et al.*, 2014), large yellow croaker (*Larimichthys crocea*) (Zheng *et al.*, 2007), among others.

Various studies have also shown that the expression of lysozyme genes in fish significantly increase after an immunological challenge with bacterial pathogens (Saurabh and Sahoo, 2008; Zhao *et al.*, 2011; Wang *et al.*, 2013; Gao *et al.*, 2016). Furthermore, recombinant c-type and g-type lysozymes from various fish species exhibit broad-spectrum antibacterial activity against Gram-positive and Gram-negative pathogens, including those relevant in aquaculture or in marine environments. Antibacterial activity against Gram-positive pathogens includes *Streptococcus iniae*, *Corynebacterium glutamicum*, *Staphylococcus epidermidis*, *S. aureus*, *Listeria monocytogenes* and *Micrococcus lysodeikticus*. Antibacterial activity against Gram-negative pathogens includes *Vibrio parahaemolyticus*, *V. salmonicida*, *V. harveyi*, *V. anguillarum*, *V. alginolyticus*, *Photobacterium damsela*, *Aeromonas sobria*, *A. hydrophila*, and *Yersinia ruckeri* (Hikima *et al.*, 2001; Zheng *et al.*, 2007; Whang *et al.*, 2011; Fu *et al.*, 2013; Yu *et al.*, 2013; Wei *et al.*, 2014; Li *et al.*, 2016; Zhang *et al.*, 2018). In this sense, lysozyme plays an essential role in the innate defense against infectious diseases in fish. In addition to its important role in immune defense, lysozyme can function as an indicator of stress in

cultured fish, as its activity is known to vary according to the intensity, duration and type of induced stress in these organisms (Fevolden *et al.*, 2002; Saurabh and Sahoo, 2008).

Totoaba: A Critically Endangered Species with Great Potential in Aquaculture

Totoaba (*Totoaba macdonaldi*) is a demersal teleost fish and the largest member of the *Sciaenidae* family, commonly known as croakers or drums. It is a long-lived fish, which can live up to 50 years, and it can reach up to 2 m long and 100 kg of weight (Cisneros-Mata *et al.*, 1997). This species is endemic to the central and northern Gulf of California, Mexico, ranging from the Colorado River delta to Concepcion Bay at the west coast, and up to the El Fuerte River mouth on the east coast (Figure 8), where brackish waters are found (Valdez-Muñoz *et al.*, 2010). At the beginning of the 20th century, totoaba's commercial fishery was one of the most prominent economic activities in the Gulf of California due to its extremely high value in the Asian markets. It was until 1975, when totoaba's natural population was dramatically reduced due to habitat depletion, unregulated fishing, and poaching that this species was declared as threatened and since then is currently classified as critically endangered in the red list by the International Union for the Conservation of Nature (Guevara, 1990; Lercari and Chávez, 2007; Findley, 2010; Bobadilla *et al.*, 2011). Since 1975, totoaba's fishery was completely banned and in 1991 it was officially declared as threatened with extinction by the Mexican government. At present, totoaba is a protected species listed in the NOM-059-SEMARNAT-2010 (Environmental protection of native Mexican species of wild flora and fauna-Risk categories) and in the Convention on International Trade of Endangered Species of Wild Flora and Fauna (CITES, Appendix I) (Bobadilla *et al.*, 2011; Valenzuela-Quiñonez *et al.*, 2015).

However, aquaculture represents a feasible strategy to recover totoaba's natural population (True, Loera, *et al.*, 1997; Mata-Sotres *et al.*, 2015; Juarez *et al.*, 2016). Recently, the development of totoaba aquaculture started in the states of Baja California and Sonora, Mexico, with assisted breeding and fertilization programs which has allowed

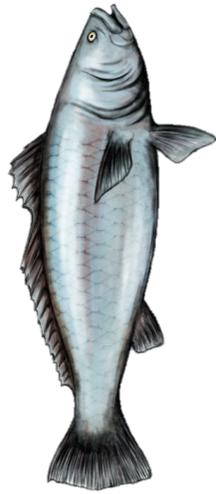


Phylogenetic family
 Sciaenidae family
 Known as drums or croakers

Habitat
 Endemic species from the Gulf of California, Mexico

Natural population
 Declined ~ 95% over the past 78 years (1942-2020)

State of conservation
 Critically Endangered
 Facing extremely high risk of extinction in the wild



Totoaba
(Totoaba macdonaldi)

Figure 8. Totoaba, a critically endangered species endemic from the Gulf of California, Mexico.

the successful management of this species from broodstock to juvenile stages. Therefore, totoaba is gaining interest, and efforts are being pursued in intensive aquaculture due to its rapid growth. Nevertheless, large-scale hatchery rearing of this new species still faces notable technical challenges, including those related to disease outbreak control (True, Silva-Loera, *et al.*, 1997; Reyes-Becerril *et al.*, 2016). It is known that farming conditions including high biomass density, low water quality and nutrient availability are factors that can cause chronic stress in fish. All these lead to immunosuppression and increases the fish vulnerability to the attack of pathogens, also compromising its lifecycle (Saurabh and Sahoo, 2008; Noga *et al.*, 2011; Kiron, 2012). Thus, the understanding of the totoaba protective immune response against pathogens is critical, in order to promote fish health and the development of immunoprophylactic strategies and improve farming conditions. However, the genetic, biochemical and immunological knowledge of this fish is still scarce (Reyes-Becerril *et al.*, 2016; González-Félix *et al.*, 2018).

Current research on totoaba's aquaculture has been focused on the effects of diet and supplementation in digestive system enzyme's, growth performance and physiological state of the fish (Galaviz *et al.*, 2015; López *et al.*, 2015; Satriyo *et al.*, 2017; Mata-Sotres *et al.*, 2018; Cabanillas-Gámez *et al.*, 2020; Lazo *et al.*, 2020). Conversely, very few studies have been conducted to investigate and describe totoaba's immune response (Reyes-Becerril *et al.*, 2016, 2018; González-Félix *et al.*, 2018; Angulo *et al.*, 2019). Major findings in totoaba's immune system research up to date are summarized in Table II.

In this sense, the growth of totoaba's aquaculture requires a further and wider knowledge of its immune system to develop and improve immunoprophylactic strategies, such as immunostimulation, for the control of diseases in this species. In this context, basic research on genes associated with the innate immune system of totoaba, such as lysozyme, will give us insight into the early defense mechanisms of this species against pathogens relevant in aquaculture. In the future, this knowledge will be crucial to improve farming conditions and to develop diagnosis and immunoprophylactic strategies for the establishment of sustainable long-term cultures of this species.

Table II. Current research on totoaba's immune response against bacterial pathogens and immunostimulants.

Study and major findings	References
<p>Leukocyte susceptibility and immune response against <i>Vibrio parahaemolyticus</i> in <i>Tototoaba macdonaldi</i></p> <ul style="list-style-type: none"> • IL-1β and IL-8 were up-regulated in spleen and intestine, and down-regulated in liver after the immunological challenge with <i>V. parahaemolyticus</i>. • The percentage of head kidney and thymus leukocytes changed (populations of lymphocytes increased and granulocytes/monocytes decreased) after exposure to <i>V. parahaemolyticus</i>. • The phagocytic activity and respiratory burst decreased in head kidney, spleen and thymus leukocytes. • The expression pattern of IL-1β and IL-8 was up-regulated in head kidney and spleen leukocytes exposed to <i>V. parahaemolyticus</i>. 	<p>(Reyes-Becerril <i>et al.</i>, 2016)</p>
<p>Caspase -1, -3, -8 and antioxidant enzyme genes are key molecular effectors following <i>Vibrio parahaemolyticus</i> and <i>Aeromonas veronii</i> infection in fish leukocytes</p> <ul style="list-style-type: none"> • The expression levels of casp-1, casp-3, casp-8, and antioxidant enzymes (catalase, glutathione peroxidase 1 and 4) from <i>T. macdonaldi</i> were strongly induced in head kidney, liver and spleen leukocytes after <i>V. parahaemolyticus</i> infection (peak at 24h post-infection in head kidney leukocytes; peak at 6h post-infection in spleen leukocytes). 	<p>(Reyes-Becerril <i>et al.</i>, 2018)</p>
<p>C-type lectin 17A and macrophage-expressed receptor genes are magnified by fungal β-glucan after <i>Vibrio parahaemolyticus</i> infection in <i>Totoaba macdonaldi</i> cells</p> <ul style="list-style-type: none"> • C-type lectin 17A, TLR2 and natural resistance-macrophage protein 2 genes were strongly up-regulated in head kidney leukocytes after immunostimulation with fungal β-glucan 197A and <i>V. parahaemolyticus</i>. 	<p>(Angulo <i>et al.</i>, 2019)</p>

CHAPTER I

Overview

In the first part of this research, the genes encoding for the c-type and g-type lysozymes expressed in totoaba were cloned and characterized. Totoaba's lysozymes were homologous to other lysozymes from their corresponding type from fish and higher vertebrates' species, presented conserved structural-functional signatures typical of the lysozyme superfamily and were constitutively expressed in all examined tissues. The findings suggest that lysozymes play a significant role in the innate defense of totoaba against bacterial infections and also may be involved in digestion.

The results of this part of the research were published as an original short communication article in an indexed journal under the following citation: Moreno-Córdova, E. N. *et al.* (2020). Molecular characterization and expression analysis of the chicken-type and goose-type lysozymes from totoaba (*Totoaba macdonaldi*). *Developmental & Comparative Immunology*, 113:103807.

MATERIALS AND METHODS

cDNA Cloning and Expression Profile of the C-type and G-type Lysozymes from *Totoaba macdonaldi*

Fish Animal Samples

Healthy totoaba juveniles (15 cm of average body size) were kindly provided by the Reproductive Center of Marine Species of the State of Sonora (CREMES), located in Kino Bay, Sonora, Mexico (28°49'22" N, 111°56'27" W). Fish were maintained in an indoor recirculating aquaculture system at 23 ± 1.0 °C. Tank conditions were kept at $35 \pm 0.5\%$ of salinity, oxygen concentration higher than 6 mg L^{-1} and a photoperiod cycle of 12 h light/12 h dark. The juveniles were euthanized, and the spleen, kidney, pyloric caeca, stomach, heart and brain tissues were surgically removed. Sampled tissues were immediately soaked in RNeasy[®] tissue collection solution (Invitrogen) and stored at -80 °C until RNA extraction.

cDNA Synthesis from Extracted RNA

Total RNA from spleen, kidney, pyloric caeca, stomach, heart and brain was isolated using Quick-RNA[™] Miniprep kit (Zymo Research), according to the manufacturer's specifications. Thereafter, first-strand cDNA was synthesized from spleen RNA using SuperScript[®] III first-strand synthesis system for RT-PCR (Invitrogen), according to the manufacturer's protocol. The resulting cDNA was stored at -20 °C and it was used to first amplify the partial-length cDNA sequences of totoaba g-type and c-type lysozymes. RNA samples for the rest of tissues were immediately processed to perform one-step qRT-PCR analysis as described below.

Cloning of the Full-length cDNA Sequences

Since the coding sequences of totoaba g-type (TmLyzg) and c-type (TmLyzc) lysozymes were unknown, their protein sequence had to be deduced from homologous genes that belong to the same protein family. Genetic information of the *Sciaenidae* family was only available for the large yellow croaker (*Larimichthys crocea*, Genome ID: 12197) and Mi-iuy croaker (*Miichthys miuiy*, Genome ID: 23982), but lysozyme genes have only been mapped in the *L. crocea* genome so far. Thus, degenerate primers for TmLyzg and TmLyzc were designed based on the available amino acid sequences deduced from *L. crocea* (GenBank ABR66917.1 and XP_019114159.1) and closely related teleost fish to amplify highly conserved regions in g-type and c-type lysozymes. The set of gene specific and degenerated primers used is listed in Table III. The partial cDNA sequence of each TmLyzg and TmLyzc was cloned by RT-PCR using a GoTaq[®] DNA polymerase (GoTaq[®] Green Master Mix, Promega) and their respective degenerate primer set (Figure 9). The cDNA amplification for each gene was performed in a DNA Engine[®] thermal cycler (Bio-Rad). The PCR program was an initial denaturation of 94 °C for 5 min, followed by 35 cycles of denaturation at 94 °C for 1 min, annealing at 50 °C (TmLyzg) or 52 °C (TmLyzc) for 1 min and extension at 72 °C for 45 s, and a final extension step of 72 °C for 5 min. PCR products were analyzed on a 1% agarose gel and the amplified DNA fragments were purified using QIAEX II gel extraction kit (Qiagen).

Purified DNA fragments were treated with a small amount (1 µL in a total reaction volume of 4 µL) of GoTaq[®] Green Master Mix at 72 °C for 10 min to allow the addition of 3' deoxyadenosine-overhang to DNA by the 5' → 3' exonuclease activity of GoTaq[®] DNA polymerase. PCR products were ligated into 25 ng of pCR2.1[®] vector (at 1:3 vector-insert molar ratio, in 10 µL total reaction volume) overnight at 4 °C using the TA Cloning[®] kit (Invitrogen). The resulting constructs were separately used to transform *Escherichia coli* TOP10 competent cells by heat shock at 42 °C for 1 min.

Table III. List of oligonucleotide primers used for RT-PCR, RACE and qRT-PCR.

Primer	Sequence ^a (5' → 3')	Description
Lyzg1Fw	ATGGGTTAYGGAAACATYATG	Isolation of totoaba g-type lysozyme cDNA (partial sequence)
Lyzg1Rv	CTGAGCTCTGGCAACRAC	
Lyzc1Fw	ATGAGGICTCTGGTGKTTT	Isolation of totoaba c-type lysozyme cDNA (partial sequence)
Lyzcm1Rv	ATCAGTCAGAAGCTSGCTGC	
TmLyzcFw3RACE	GGCTAATGGAATGGACGGCTACCACGG	Amplification of TmLyzc 3'-end (3'-RACE)
TmLyzcFw3NRACE	GCCAACTGGGTTTGTCTGTCCAAGTGG	
TmLyzcRv5RACE	GCACCACCAGCGGCTGTTGATCTGG	Amplification of TmLyzc 5'-end (5'-RACE)
TmLyzcRv5NRACE	CCACTTGGACAGACAAACCCAGTTGGC	
TmLyzgFw3RACE	CGCCTGGGGACTGATGCAGGTTGATG	Amplification of TmLyzg 3'-end (3'-RACE)
TmLyzgFw3NRACE	GAGAGGTGGACACACTGCAGAGGGCG	
TmLyzgRv5RACE	ATCCCAGTAGCTTGGCAGAGGTGTTCC	Amplification of TmLyzg 5'-end (5'-RACE)
TmLyzgRv5NRACE	CTCTGCAGTGTGTCCACCTCTCTCTGG	
TmLyzcFw LyzcRvqPCR	ATGAGGGCTCTGGTGTTT C CAGACAAACCCAGTTGGC	qRT-PCR TmLyzc
TmLyzgFw555 Lyzg1RvqPCR	ATGGGTTATGGAAACATCAAGAG GCCTGCATCTTCTCTGC	qRT-PCR TmLyzg
TmEF1aFw TmEF1aRv	CATTGTCAAACCTCATTCCACAG CACGGTCTGCCTCATGTC	Reference gene qRT-PCR Elongation factor-1 α from <i>T. macdonaldi</i> (Genbank KX524957.1)

^a IUPAC codes for degenerate bases: I (Inosine), K (G/T), M (C/A), R (A/G), S (G/C), W (T/A), Y (T/C).

Briefly, 2 μL of each ligation reaction were added to 50 μL of competent cells, mixed gently and then incubated for 30 min on ice. Heat shock was carried out at 42 $^{\circ}\text{C}$ for 1 min, then the cells were immediately incubated on ice for 5 min. Then, 250 μL of SOC medium were added and the cell suspension was incubated at 37 $^{\circ}\text{C}$ for 90 min with shaking at 225 rpm. After that time, 50-150 μL of transformed cells were cultured in LB plates containing ampicillin (100 $\mu\text{g mL}^{-1}$) and X-Gal (40 mg mL^{-1}) and incubated overnight at 37 $^{\circ}\text{C}$. Positive clones (observed as round white colonies) were used to isolate plasmid DNA by alkaline lysis (Green and Sambrook, 2016) and the purified recombinant plasmids were sequenced by Sanger sequencing at the USSDNA (IBT UNAM) using M13+ and T7+ universal primer sites.

A new first-strand cDNA for 5' and 3' RACE was synthesized from total spleen RNA using the SMARTer™ RACE cDNA amplification kit (Takara Bio, USA), following the manufacturer's protocol (Figure 9). The gene specific primers TmLyzcFw3RACE and TmLyzcRv5RACE, and UP (supplied by the kit) were used for the first PCR experiment for TmLyzc. Amplification conditions were conducted at 94 $^{\circ}\text{C}$ for 5 min, followed by 25 cycles of 94 $^{\circ}\text{C}$ for 30 s, 63 $^{\circ}\text{C}$ for 30 s and 72 $^{\circ}\text{C}$ for 3 min, and a final extension step of 72 $^{\circ}\text{C}$ for 5 min. Nested 3' and 5'-RACE PCR was carried out using TmLyzcFw3NRACE and TmLyzcRv5NRACE as gene specific primers and UP-short (supplied by the kit) under identical amplification conditions. RACE for TmLyzg was carried out using gene specific primers TmLyzgFw3RACE, TmLyzgRv5RACE and UP for first round amplification run by 25 cycles of 94 $^{\circ}\text{C}$ for 30 s, 65 $^{\circ}\text{C}$ for 30 s (3'-RACE) or 63 $^{\circ}\text{C}$ for 45 s (5'-RACE) and 72 $^{\circ}\text{C}$ for 3 min. Nested 5'-RACE was conducted using TmLyzgRv5NRACE and UP-short with the same amplification conditions. The resulting RACE products were analyzed by 1% agarose gel electrophoresis and purified by QIAEX II gel extraction kit (Qiagen). The purified products were refreshed and cloned separately into the pCR 2.1® vector as described above. Recombinant clones were sequenced by Sanger sequencing using M13+/- universal primers on vector at the USSDNA (IBT UNAM).

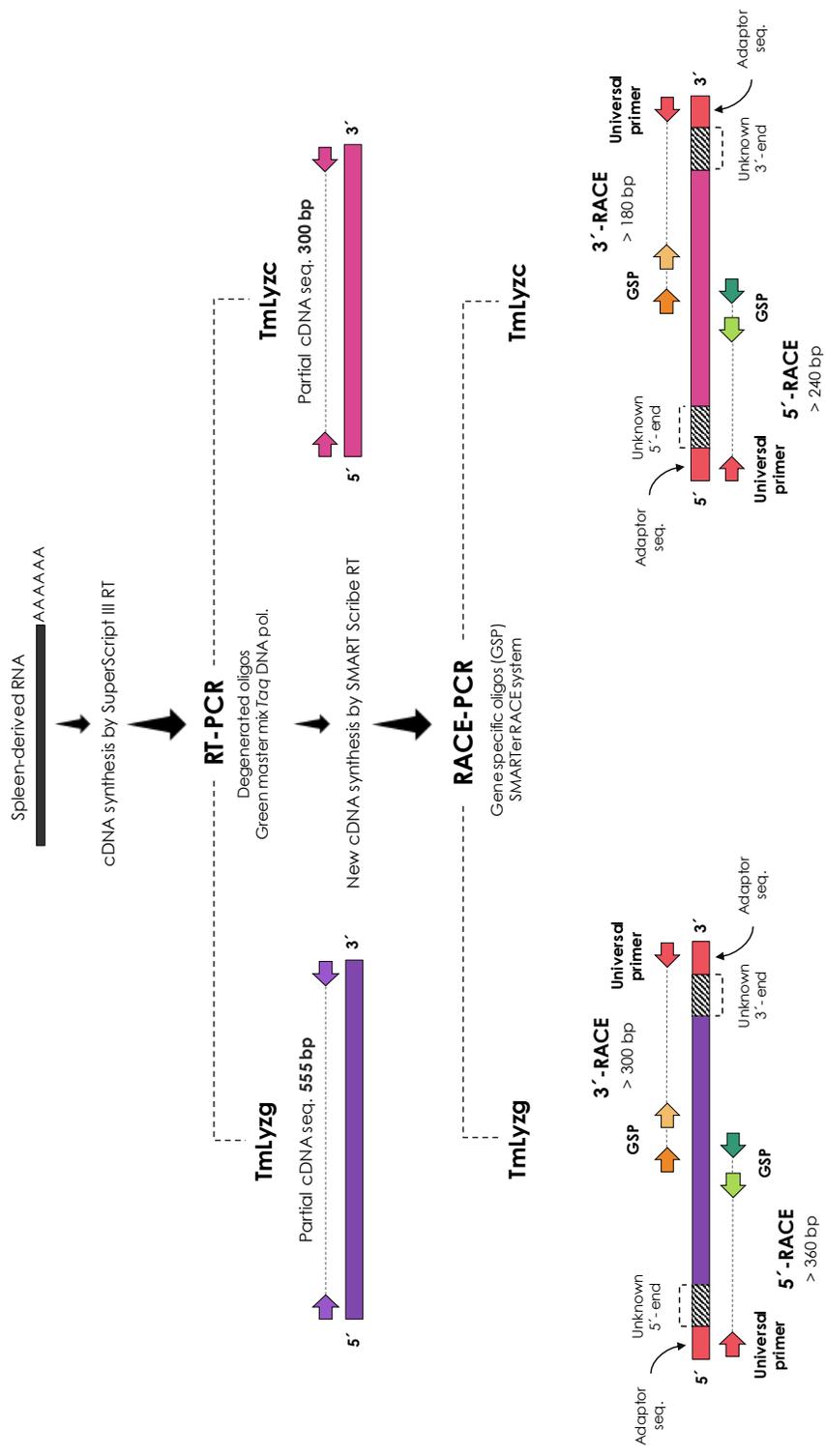


Figure 9. General strategy diagram for the amplification of TmLyzg and TmLyzc partial cDNA sequences by RT-PCR and the amplification of 3' and 5' cDNA ends by RACE-PCR.

Sequence Analysis of and Molecular Modeling

The cDNA sequences of TmLyzg and TmLyzc were analyzed using the ORF Finder tool to predict the open reading frame regions and translate the deduced amino acid sequences (<https://www.ncbi.nlm.nih.gov/orffinder/>). BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) search using the algorithms BLASTX and BLASTP was conducted to identify known protein sequences that are homologous to the deduced amino acid sequences of TmLyzg and TmLyzc. In order to infer the homologies, multiple amino acid sequence alignments of TmLyzg, TmLyzc and lysozymes from other teleost fish and other vertebrates were performed using the T-Coffee server (<http://tcoffee.crg.cat/apps/tcoffee/do:regular>) (Notredame *et al.*, 2000). Conserved domains in the TmLyzg and TmLyzc amino acid sequences were identified using the Conserved Domain Database by NCBI (<https://www.ncbi.nlm.nih.gov/cdd>) (Marchler-Bauer *et al.*, 2017).

Physicochemical parameters, such as molecular mass and theoretical isoelectric point, were computed using ProtParam tool on (<https://web.expasy.org/protparam/>) (Gasteiger *et al.*, 2005). In addition, the presence of signal peptide and the subcellular location of the lysozymes were predicted using the SignalP-5.0 server (<http://www.cbs.dtu.dk/services/SignalP/>) and YLoc tools (<http://abi.inf.uni-tuebingen.de/Services/YLoc/webloc.cgi>), respectively (Briesemeister *et al.*, 2010; Petersen *et al.*, 2011).

The evolutionary histories of TmLyzg and TmLyzc were inferred through phylogenetic analysis using MEGA X software, through the Maximum Likelihood method at 1000 bootstraps per analysis (Kumar *et al.*, 2018; Stecher *et al.*, 2020). Finally, the software MOE v2018.01 (Chemical Computing Group) was used to construct and refine a predicted homology model for the tridimensional structures of TmLyzg and TmLyzc, using the crystal structures of the g-type lysozyme from Atlantic cod (*Gadus morhua*) (PDB: 3GXX) (Helland *et al.*, 2009) and the c-type lysozyme from rainbow trout (*Oncorhynchus mykiss*) (PDB: 1LMN) (Karlsen *et al.*, 1995) as templates, respectively. A set of 25 independent models were made for each lysozyme sequence under the

CHARMM27 forcefield. The final models were optimized from the best-scoring intermediate model according to the Protein Geometry analysis function and after a default energy minimization procedure in MOE.

Expression Analysis in Totoaba Tissues

Expression profile of TmLyzg and TmLyzc genes in healthy tissues (spleen, kidney, intestine, stomach, pyloric caeca, heart and brain) from totoaba was evaluated by qPCR analysis using Brilliant III Ultra-Fast SYBR[®] Green qRT-PCR master mix (Agilent). The reactions were conducted on a StepOne[™] Real-Time PCR System thermal cycler (Thermo Fisher Scientific) under the following amplification conditions: 50°C for 10 min, 95°C for 3 min and 35 cycles of 94°C for 10 s and 55°C for 5 s. Primers used to amplify amplicons from each lysozyme gene are shown in Table III. Elongation Factor-1 α from totoaba was used to normalize gene expression in each sample using the C_T method ($2^{-\Delta C_t}$) (Livak and Schmittgen, 2001; Reyes-Becerril *et al.*, 2016). Three healthy totoaba individuals were used as biological replicates. All data were presented as relative mRNA expression (mean \pm standard deviation) with experimental triplicates (n = 3). Descriptive statistics were performed using GraphPad Prism version 8.4.1 for MacOS (GraphPad Software, San Diego California USA).

RESULTS AND DISCUSSION

Characterization of the C-type and G-type Lysozymes cDNA Sequences

The full-length cDNA sequences of TmLyzg and TmLyzc were obtained by RT-PCR and RACE techniques and deposited in Genbank under the accession numbers MT211628 and MT211627, respectively. For TmLyzg, we cloned a cDNA sequence of 711 bp in length, which includes an open reading frame (ORF) of 582 bp and a 3'-UTR of 129 bp with a typical polyadenylation signal (AATAAA) and a poly(A) tail (Figure 10). The deduced amino acid sequence for TmLyzg has 193 residues, with a calculated molecular mass of 21.5 kDa and a theoretical isoelectric point (pI) of 7.7. The subcellular location prediction indicated that TmLyzg might be an intracellular lysozyme. This is in accordance with the features exhibited by most g-type lysozymes from fish (Irwin and Gong, 2003). The TmLyzg amino acid sequence showed significant homology with other g-type lysozymes from various fish species. TmLyzg presented the highest identity of 92% and 83% with members of the *Sciaenidae* family large yellow croaker (*L. crocea*) and big head croaker (*Collichthys lucidus*), respectively (Zheng *et al.*, 2007). High identities were also found with other bony fishes such as 76% with mandarin fish (*Siniperca chuatsi*) (Sun *et al.*, 2006), 76% with European seabass (*Dicentrarchus labrax*) (Buonocore *et al.*, 2014), 76% with barred knifejaw (*Oplegnathus fasciatus*) (Whang *et al.*, 2011), 75% with turbot (*Scophthalmus maximus*) (Zhao *et al.*, 2011), 74% with black rockfish (*Sebastes schegelii*) (Nilojan *et al.*, 2017) and 64% with Atlantic cod (*Gadus morhua*) (Larsen *et al.*, 2009).

Regarding TmLyzc, the full-length cDNA sequence contains a 5'-UTR of 130 bp, a 3'-UTR of 280 bp with a 25 bp poly (A) tail and a consensus polyadenylation signal sequence AATAAA (Figure 11). The ORF of 432 bp for TmLyzc encodes a polypeptide of 143 amino acids with a calculated molecular mass of 16.12 kDa and a theoretical pI of 8.45. Moreover, this lysozyme was predicted to localize extracellularly, as a signal sequence of 15 residues was identified at the N-terminus of the deduced protein. This is consistent with the characteristics displayed by most of the c-type lysozymes from the vertebrate phylum, including fish, birds and mammals (Tullio *et al.*, 2015).

<i>atgggttatggaaacatcaagagggttcagactactggtgcgtcacagcaaacatctcag</i>	60
M G Y G N I K R V Q T T G A S Q Q T S Q	20
<i>caggacaaaactgggatactcaggtgtgagggcatcacatgcgatggcagaagaagatgca</i>	120
Q D K L G Y S G V R A S H A M A E E D A	40
<i>ggccgaatggaaaagtacagatgtaaaatcaacagtggttgacgtagatatgatatcgat</i>	180
G R M E K Y R C K I N S V G R R Y D I D	60
<i>ccagctctcatcgctgcaatcatctccagagaatctagggctggaaatgcactaactaat</i>	240
P A L I A A I I S R E S R A G N A L T N	80
<i>ggatggggagactatagcccagtgagaggacaatacaacgcctggggactgatgcaggtt</i>	300
G W G D Y S P V R G Q Y N A W G L M Q V	100
<i>gatgtcaatccagagagaggtggacacactgcagagggcgcatgggacagtgaggaacac</i>	360
D V N P E R G G H T A E G A W D S E E H	120
<i>ctctgccaaagctactgggatcttggtttatttcatcaaagtgattcgcaacaaatcttct</i>	420
L C Q A T G I L V Y F I K V I R N K F P	140
<i>ggctggagcacagaggagcagctgaaaggaggatagcagcatacaatatgggggatcga</i>	480
G W S T E E Q L K G G I A A Y N M G D R	160
<i>agtgtcgaagacaaagatgtggataagataacaacaggtagagactactccaatgatgtt</i>	540
S V E D K D V D K I T T G R D Y S N D V	180
<i>gttgccagagctcagtggtacaaaaacaataaaaaactattaacagctgaagctgtctgtt</i>	600
V A R A Q W Y K N N K N Y *	193
<i>acaaagatctcctacaaaatgacagaaacatttctgtaatgtgtgctgtgtcaaaacgta</i>	660
<i>tgctgaataaagaaaactgcaaagcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa</i>	711

Figure 10. Nucleotide and deduced amino acid sequence of *Totoaba macdonaldi* g-type lysozyme (TmLyzg). Coding sequence is depicted in italic letters. Deduced amino acid sequence is marked by gray box and stop codon is marked by an asterisk.

	atggggattc	-130
agagaaaat	atggagttagtggaaccagagcagccacagcaat	-120
aactataaa	aggagatttagcagcagctgtggaaacattcgatccagcagagagtgccatc	-60
<i>atgagggctctggtgtttctgctcttggttaactttggccagcgctaaaatctaccagcgc</i>		60
M R A L V F L L L V T L A S A K I Y Q R		20
<i>tgtgaatgggcccgaatcctgaaggctaattggaatggacggctaccacggttacagcctg</i>		120
C E W A R I L K A N G M D G Y H G Y S L		40
<i>gccaaactgggtttgtctgtccaagtgggagtcagactacaacaccagagccaccaaccac</i>		180
A N W V C L S K W E S D Y N T R A T N H		60
<i>aaactgatggatccactgactacggcatcttcagatcaacagccgctggtggtgcaac</i>		240
N T D G S T D Y G I F Q I N S R W W C N		80
<i>gatggccgcacctctacatcaaatgcatgccacatccagtgagccagcttctgactgat</i>		300
D G R T S T S N A C H I Q C S Q L L T D		100
<i>gatgtcagtggtggcgatcaactgtgccaaacgtgtcggttagggatcccagcggcatcgga</i>		360
D V S V A I N C A K R V V R D P S G I G		120
<i>gcctgggtggcctggcggcgtcactgcgagaaccgtgacctgagctcctatgtgtcagga</i>		420
A W V A W R R H C E N R D L S S Y V S G		140
<i>tgtggcctttaatcaacacatggaacaggacagtgatcagcataataaaattctccc</i>		480
C G L *		143
atcttcagctgaaggggattaaaaggtttctgctgttacacgcagttactaaaacagatt		540
attctcaaatacaagtgaattcagatctttactttcaatgtatttagttatacttcactt		600
gattttatagatgtagtgagtgaacagaaatgtgctgatttcaacttgaatctgttgca		660
ttgtgttcatttcaaacctgattaaaatggttgattttgtgttaaaaaaaaa		712

Figure 11. Nucleotide and deduced amino acid sequence of *Totoaba macdonaldi* c-type lysozyme (TmLyzc). Coding sequence is depicted in italic letters. Deduced amino acid sequence is marked by gray box and stop codon is marked by an asterisk.

The inferred amino acid sequence of TmLyzc showed high identity with c-type lysozymes from other fish species and even higher vertebrates. TmLyzc presented identity of 92% with large yellow croaker (Ao *et al.*, 2015), 91% with big head croaker, 85% with barramundi perch (*Lates calcarifer*), 80% with orange-spotted grouper (*Epinephelus coioides*) (Wei *et al.*, 2012), 77% with taimen (*Hucho taimen*) (Li *et al.*, 2016), 75% with rainbow trout (*Oncorhynchus mykiss*) (Dautigny *et al.*, 1991), 70% with mouse (*Mus musculus*), 64% with human (*Homo sapiens*) and 55% with chicken (*Gallus gallus*) protein sequences.

Analysis of Conserved Structural Features in the C-type and G-type Lysozymes of Totoaba

The amino acid sequence alignment of TmLyzg with homologous g-type lysozymes from teleost fish is shown in Figure 12. The analysis indicated that TmLyzg possesses a highly conserved goose egg white lysozyme (GEWL) domain (from Gln17 to Tyr193) (Monzingo *et al.*, 1996), that contains two conserved catalytic residues (Glu71 and Asp101), eleven *N*-acetyl-D-glucosamine binding sites residues (Glu71, Gln99, Val100, Asp101, Pro104, His109, Ile127, Phe131, Tyr155, Asn156, Gly158), and a GXXQ signature motif (Gly96, Leu97, Met98 and Gln99). Altogether, these residues are fundamental to maintain the three dimensional structure and the biological activity of g-type lysozymes (Callewaert and Michiels, 2010).

Furthermore, TmLyzg contained only one cysteine residue and no signal peptide sequence was predicted. These are common features in most of the known fish g-type lysozymes, but differ from mammalian and avian counterparts whose g-type lysozymes are secreted proteins (Irwin, 2014). Nevertheless, the presence of intracellular lysozymes in fish species might accomplish an important role in controlling cytosolic growth of intracellular bacteria, possibly mediated by their muramidase activity, bacterial membrane disruption capacity and aggregation of bacterial cells (Seppola *et al.*, 2016).

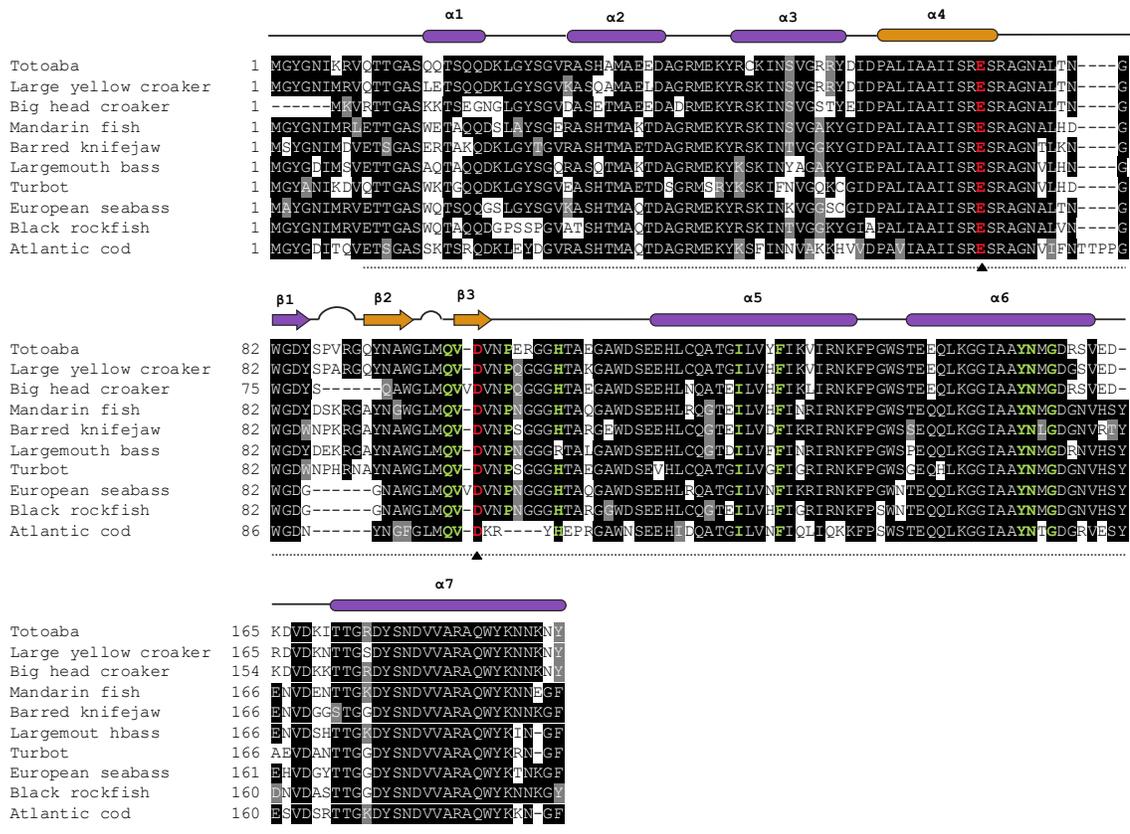


Figure 12. Multiple sequence alignment and predicted secondary structures of g-type lysozymes from totoaba and other fish. Identical amino acids are shaded in black, similar are shaded in grey and different are marked with white background. Dotted line highlights the GEWL domain. Red letters and triangles indicate active site residues and substrate binding sites are shown in green. Alpha-helices are depicted as cylinders, β -sheets as arrows, coils as solid lines and β -turns as solid line loops. Color palette is representative for structural components in the three-dimensional structure.

Figure 13 shows the amino acid sequence alignment of TmLyzc with homologous c-type lysozymes from vertebrate species, including fish, birds and mammals. The analysis revealed that TmLyzc presents a highly conserved α -lactalbumin/c-type lysozyme signature motif called LYZ1 (from Lys16 to Gly142) (all amino acids are numbered based on the full TmLyzc polypeptide sequence) (Nitta and Sugai, 1989; Qasba *et al.*, 1997). This region includes the two conserved catalytic residues Glu50 and Asp67, the residues that constitute the active site cleft (Asn61, Asp67, Gln72, Ile73, Asn74, Trp77, Trp78, Glu115, Ser117, Ala121, Trp122 and Val123) and a highly conserved signature sequence from Gly64 to Ser75. We found eight highly conserved cysteines (Cys21, Cys45, Cys79, Cys90, Cys94, Cys108, Cys129 and Cys141) in this motif that could be able to form disulfide bridges as predicted by the molecular model (below). Taken as a whole, these residues are essential to preserve the structure and function of the c-type lysozymes. These characteristics are consistent with those reported for c-type lysozymes from all vertebrate species, ranging from fish to mammals (Callewaert and Michiels, 2010). The presence of extracellular c-type lysozymes in fish indicate that these proteins may play important roles in the host defense against extracellular bacterial pathogens. Furthermore, some reports also suggest the implication of secreted lysozymes in the modulation of the immune response (Ragland and Criss, 2017). Secondary to their ability to kill bacteria, extracellular lysozymes can release immunomodulatory bacterial ligands, including peptidoglycan fragments, that in certain scenarios can further activate pro-inflammatory responses or either resolve inflammation (Ragland and Criss, 2017).

The predicted secondary structures for TmLyzg and TmLyzc amino acid sequence are also shown in Figures 12 and 13, respectively. In general, the results indicated that these proteins present structural features commonly found in the members of the lysozyme superfamily. More specifically, six distinct α -helices (α 1- α 6) and an antiparallel three-stranded β -sheet (β 1- β 3) were identified in both TmLyzg and TmLyzc. The conserved α 3- α 5 helices in TmLyzg and α 2- α 5 in TmLyzc, along with the three-stranded β -sheet, constitute the catalytic and substrate binding sites in these lysozymes. Altogether, these characteristics are considered the only constant elements of the secondary structure in all

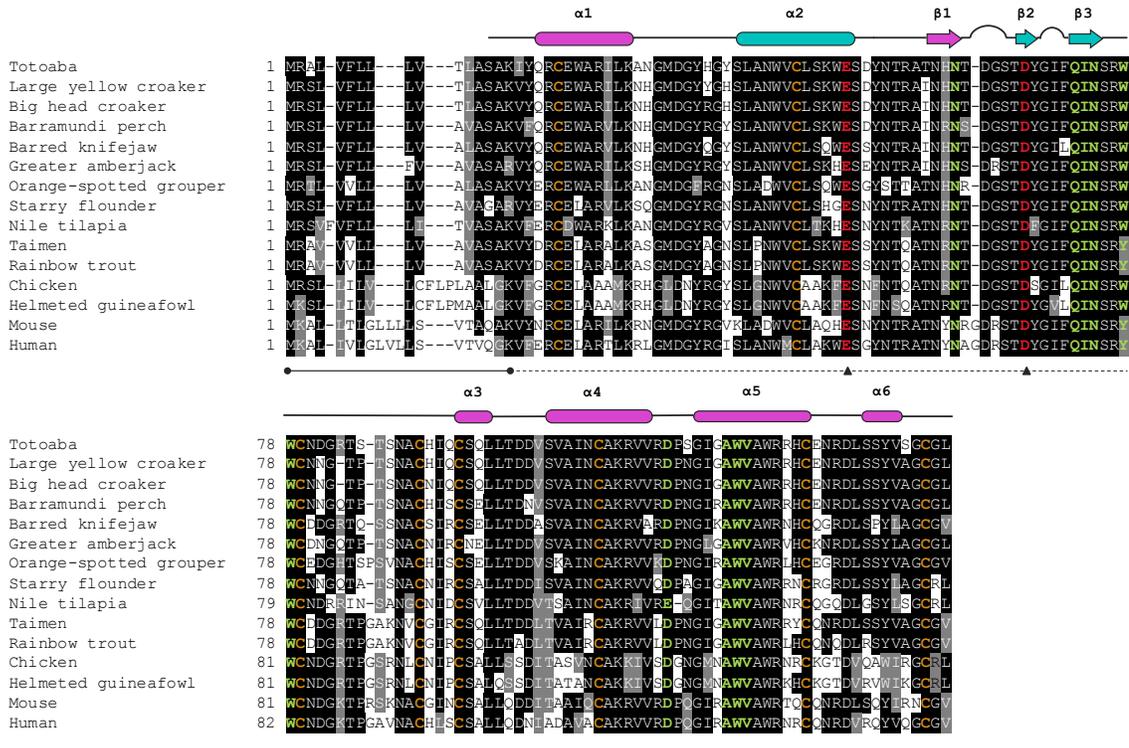
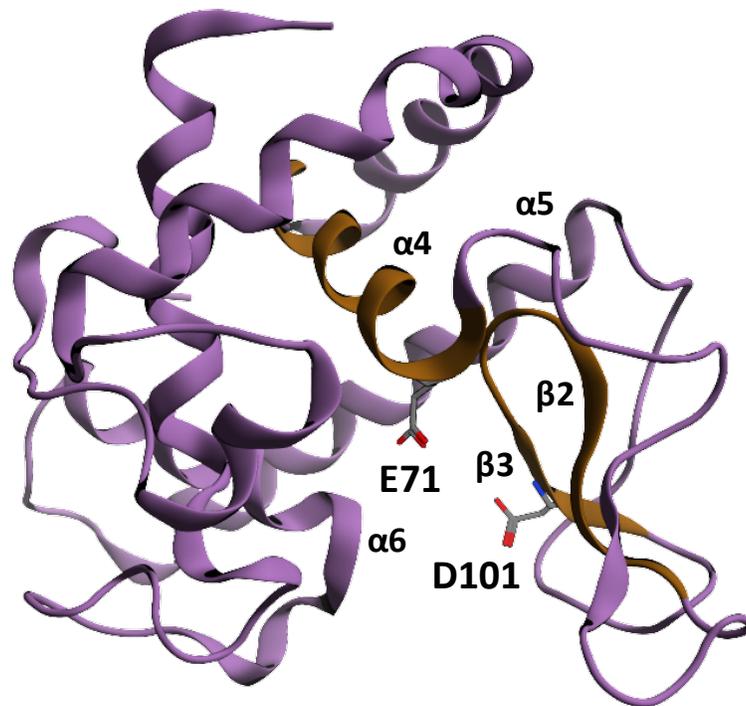


Figure 13. Multiple sequence alignment and predicted secondary structures of c-type lysozymes from totoaba, other fish and higher vertebrates. Identical amino acids are shaded in black, similar are shaded in grey and different are marked with white background. Dotted line highlights the LYZ1 domain in c-type lysozymes. Red letters and triangles indicate active site residues; substrate binding sites are shown in green and cysteine residues are shown in yellow; continuous line at the start of the sequence underlines predicted signal peptide. Alpha-helices are depicted as cylinders, β -sheets as arrows, coils as solid lines and β -turns as solid line loops. Color palette is representative for structural components in the three-dimensional structure.

members of the lysozyme superfamily, including g-type and c-type lysozymes (Monzingo *et al.*, 1996; Wohlkönig *et al.*, 2010).

In agreement with these characteristics, the predicted molecular models of TmLyzg (Figure 14) and TmLyzc (Figure 15) also showed a common lysozyme fold, which consists of a large α -helical domain separated from the small β -stranded domain by a wide substrate binding cleft (Karlsen *et al.*, 1995; Helland *et al.*, 2009; Wohlkönig *et al.*, 2010). In TmLyzg, the $\alpha 4$ and $\alpha 5$ helices contain hydrophobic residues important for substrate binding, and an $\phi\psi$ sequence signature (Tyr156 and Asn157) that is commonly found in g-type lysozymes at the end of $\alpha 5$ -helix. Regarding TmLyzc, key residues for substrate binding are distributed mainly in $\alpha 2$ - $\alpha 4$ helices as well as in $\beta 1$ - $\beta 3$ -sheets. In spite of presenting a quite similar tridimensional structure, only the central α -helix ($\alpha 4$ in TmLyzg and $\alpha 2$ in TmLyzc) and the β -hairpin ($\beta 2$ - $\beta 3$ strands in both lysozymes) are the invariant structural motifs of the lysozyme superfamily, and constitute the core that shapes the protein's active site (Wohlkönig *et al.*, 2010).

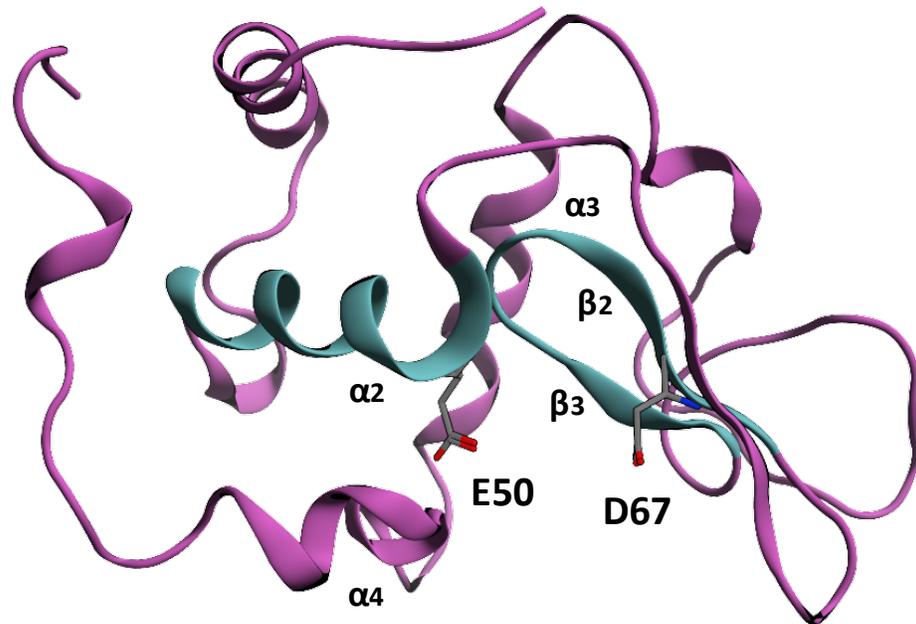
Particularly, the central α -helix contains the residues Glu71 in TmLyzg or Glu50 in TmLyzc, while the β -hairpin has the residues Asp101 in TmLyzg or Asp67 in TmLyzc that might play a crucial role in catalysis. These residues are identified as the counterparts of the well-defined catalytic amino acids of other g-type lysozymes such as those from goose egg (Glu73 and Asp97) (Weaver *et al.*, 1995) and Atlantic cod (Glu71 and Asp99) (Helland *et al.*, 2009), and other c-type lysozymes such as those from hen egg (Glu35 and Asp52) (Blake *et al.*, 1965; Artymiuk *et al.*, 1982) and rainbow trout (Glu50 and Asp67) (Karlsen *et al.*, 1995). Furthermore, these residues are fingerprints in g-type and c-type lysozymes and serve to differentiate between the catalytic mechanisms exhibited by each type. In both types of lysozymes, glutamate invariably acts as a general acid catalyst whereas aspartate acts as a general base (Zechel and Withers, 2000).



	$\alpha 4$		$\beta 2$	$\beta 3$
Totoaba	61	PAIIAAIISRES	92	YNAWGLMQVDVN
Cod	61	PAVIAAIIISRES	90	YNGFGLMQVDKR
Salmon	78	PAIIAGIISRES	103	GNGFGLMQVDKR
Goose	63	PAVIAGIISRES	88	GKGFGLMQVDKR

GXXQ
Signature motif

Figure 14. Predicted tertiary structure of TmLyzg, showing the active site, and conserved structural elements. GXXQ motif for TmLyzg in alignments fragments is highlighted in purple.



		$\alpha 2$		$\beta 2$	$\beta 3$
					
Totoaba	40	LANWVCLSKWES		66	TDY GIFQ INS
Rainbow trout	40	LPNWVCLSKWES		66	TDY GIFQ INS
Human	43	LANWMCLAKWES		70	TDY GIFQ INS
Chicken	43	LGNWVCAAKFES		69	TD S GILQ INS
					GXXQ
					Signature motif

Figure 15. Predicted tertiary structure of TmLyzc, showing the active site, and conserved structural elements. GXXQ motif for TmLyzc in alignments fragments is highlighted in pink.

In TmLyzc, Asp67 could act as a nucleophile, stabilizing the acyl-enzyme intermediate to accomplish the enzymatic reaction. This is a hallmark in all c-type lysozymes, which exert their catalytic activity through a double displacement mechanism and are known as retaining glycosidases (Zechel and Withers, 2000; Kirby, 2001). In TmLyzg, Asp101 could participate in catalysis by activating and positioning a catalytic water in order to facilitate the nucleophilic attack, as well as in stabilizing the oxocarbenium ion-like intermediate. This is a special feature of g-type lysozymes, as the unique members of the lysozyme superfamily which perform their catalytic activity by a single displacement mechanism and are known as inverting glycosidases (Zechel and Withers, 2000).

Moreover, the β -hairpin which comprises the β 2 and β 3 sheets, is recognized as the most conserved structural element (at the amino acid sequence level) among the lysozyme superfamily members, which displays a specific sequence, according to the type of lysozyme (Wohlkönig *et al.*, 2010). Based on this, we were able to differentiate between lysozyme families, according to the particular sequence motifs located in this region. The most distinctive signature motif is the GXXQ sequence, was identified as Gly96, Leu97, Met98 and Gln99 in TmLyzg (Figure 14), and as Gly69, Ile70, Phe71 and Gln72 in TmLyzc (Figure 15) (Pei and Grishin, 2005). These residues are part of the loop linking the β -sheets (β 2 and β 3) and are known to participate in substrate binding, according to the findings in the crystal structures of Atlantic cod *Gadus morhua*, rainbow trout *Oncorhynchus mykiss* and their counterparts from avian and mammalian species (Karlsen *et al.*, 1995; Helland *et al.*, 2009).

Amongst lysozymes families, these signatures show low sequence similarity. This reveals the high sequence plasticity in the β -hairpin region. For a given family, however, the high degree of conservation of a particular group of residues, located in a specific structural region, might be a requirement to adopt a specific structural conformation to accomplish a particular functional role (Pils *et al.*, 2005).

Phylogenetic Analysis of the C-type and G-type Lysozymes from Totoaba

A phylogenetic tree was constructed using the maximum likelihood method to analyze the evolutionary relationship of TmLyzg and TmLyzc with other vertebrate g-type and c-type lysozymes (Figure 16). The results indicated that lysozymes were clustered into two distinct clades, separating the orthologous g-type lysozymes from the c-type lysozymes. In the case of g-type lysozymes, the clade was split into two distinct branches or nodes, setting aside the orthologues from teleost fish from those of mammals and birds. TmLyzg was positioned within the teleost group and showed the closest relationship with the counterparts of large yellow croaker and big head croaker, suggesting a common recent ancestor between the three species, as they belong to the *Sciaenidae* family. The diverse evolutionary paths of g-type lysozymes among vertebrates is in accordance with previous reports (Irwin, 2014; Ko *et al.*, 2016). Particularly in fish species, the major contributors in the evolution of g-type lysozymes are the considerable insertions/deletions in their amino acid sequences, for example the absence of a signal peptide and the lack of disulfide bonds caused by replacement of cysteine residues (Pooart *et al.*, 2004; Irwin, 2014).

The c-type lysozymes clade, was divided into two major nodes, separating other mammal orthologues, from the fish and avian counterparts. Nevertheless, the evolutionary relationship between c-type lysozymes from distinct vertebrate species seem to be very close. In particular, TmLyzc was positioned within the teleost and avian group and showed the closest relationship with its counterparts from large yellow croaker and big head croaker. The close evolutionary path among c-type lysozymes from different species indicates a well preserved structural characteristics and functional roles during evolution (Nitta and Sugai, 1989). Fish species, including totoaba, clearly display the conservation of typical features of c-type lysozymes, which includes the presence of a signal peptide and disulfide bonds.

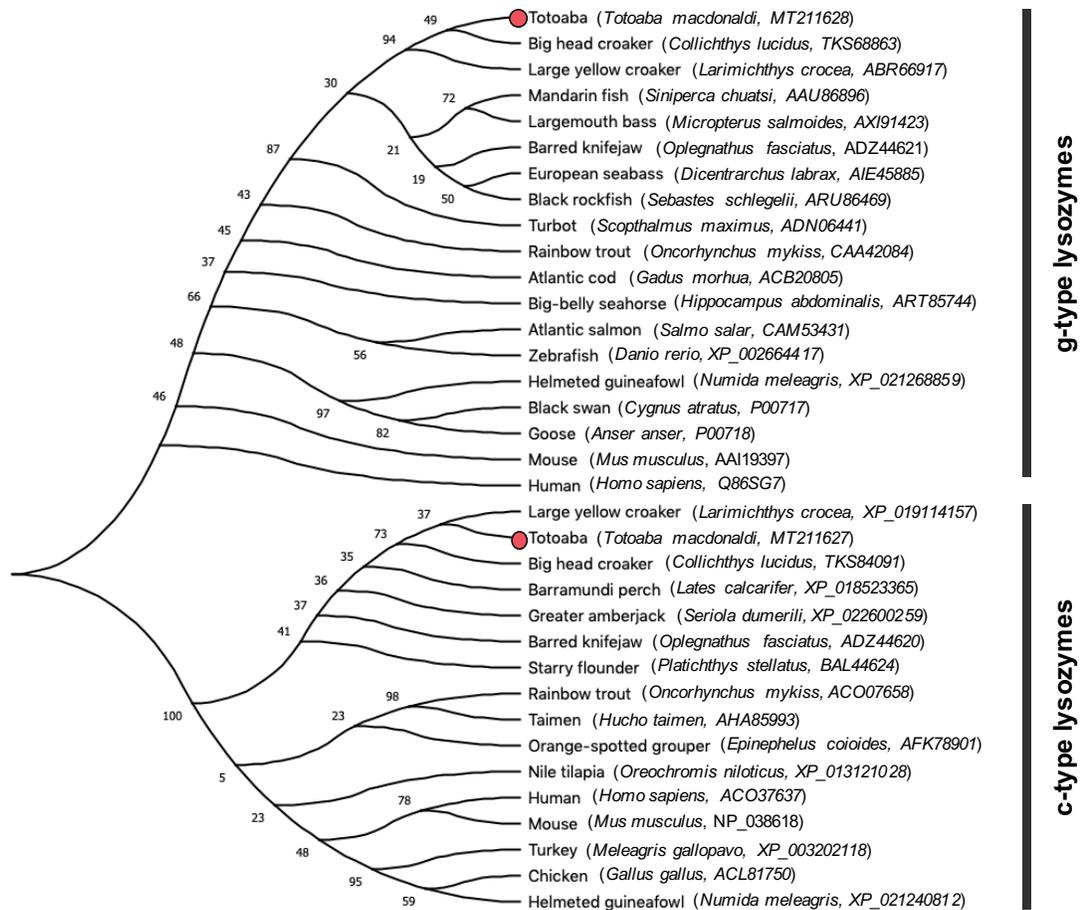


Figure 16. Cladogram of TmLyzg, TmLyzc and known orthologs in the vertebrate phylum. Genbank accession numbers of g-type and c-type lysozymes used for the analysis are depicted in italic letters.

It is generally accepted that proteins of the lysozyme superfamily have diverged from a common ancestor (Holm and Sander, 1994; Monzingo *et al.*, 1996). This inference is based on the fact that although their amino acid sequences appear to be unrelated, the overall structure between lysozymes' families is strikingly similar. In the hypothesis of divergent evolution, this means that the ancestral fold has been conserved across species and during evolution, while the complete sequences have diverged (Grütter *et al.*, 1983; Yoshikuni *et al.*, 2006). This evidence the high sequence plasticity displayed by these proteins in order to exert specific catalytic functions, but also their amazing ability to maintain an overall well-conserved fold to preserve its biological function.

Expression Analysis of the C-type and G-type Lysozyme Genes in Different Tissue Samples

TmLyzc and TmLyzg were ubiquitously expressed in all examined tissues, with predominant expression in stomach, pyloric caeca and heart (Figure 17). The highest expression level of TmLyzc was detected in stomach and pyloric caeca, whereas TmLyzg was mainly expressed in stomach and heart. C-type and g-type lysozymes from other teleosts such as Japanese flounder (Hikima *et al.*, 2001), orange-spotted grouper (Yin *et al.*, 2003), large yellow croaker (Zheng *et al.*, 2007), brill (Jiménez-Cantizano *et al.*, 2008), Atlantic cod (Larsen *et al.*, 2009), Asian seabass (Fu *et al.*, 2013), starry flounder (Kim and Nam, 2015) and Chinese black sleeper (Wei *et al.*, 2020) were also expressed in the gastrointestinal tract (stomach and intestine) and in heart. Nevertheless, in contrast to our results, the highest expression level of these lysozymes was observed in spleen and kidney. Lysozymes genes in fish are not expressed in a tissue restricted fashion, as it is with their counterparts in birds and mammals (Nakano and Graf, 1991; Irwin and Gong, 2003; Irwin *et al.*, 2011), but are expressed predominantly in hematopoietic organs including spleen, head kidney and liver, as well as tissues exposed to external environment that are the first interaction sites between host and bacteria such as gills, skin and gastrointestinal tract (Saurabh and Sahoo, 2008).

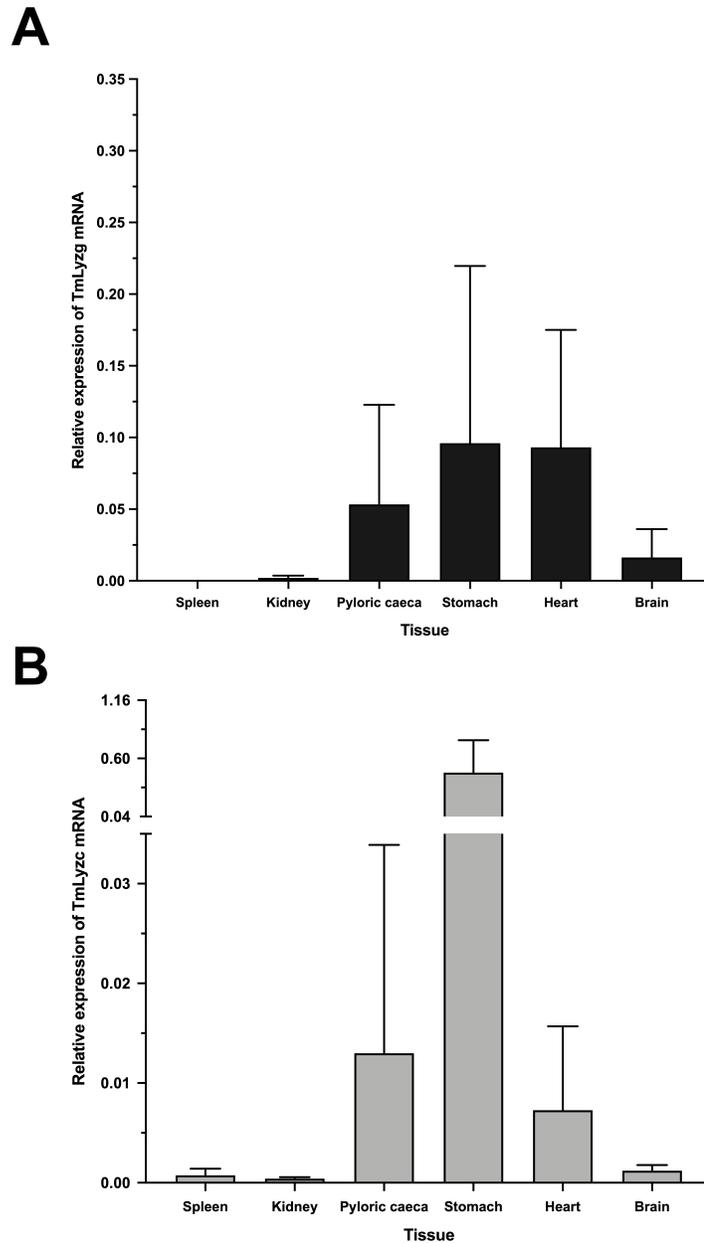


Figure 17. Tissue profile for the relative expression levels of the g-type (A) and c-type (B) lysozymes mRNA of totoaba. Total mRNA was isolated from tissues of three fish and individually processed to perform RT-qPCR with experimental triplicates (n = 3). Expression values were normalized to those of EF-1 α as housekeeping gene. Data are expressed as the mean of mRNA relative expression (means \pm S.D., n = 3).

In a previous study with reared Atlantic salmon, the expression levels of the immune response genes in the gut-associated lymphoid tissue was affected by diet and farming conditions (Løkka *et al.*, 2014). Lysozymes may prevent bacterial infection through the humoral innate immunity in the fish gastrointestinal tract by a direct bactericidal effect (Smith *et al.*, 2019). González-Félix *et al.* (2018) showed that the level of plasma lysozyme activity in totoaba can be affected depending on the fish diet (basal or probiotic diet), without analyzing the affected type of lysozyme activity nor the expression profile in the gastrointestinal tract. In the near future, it will become necessary to explore the effects in the expression and activity pattern of totoaba lysozymes under different immunological challenges, especially since lysozyme activity can also increase in other farmed fish species, serving as a possible marker of inflammatory response to diet components. Another possible function of lysozyme can relate to digestion, since high expression of lysozymes in the gastrointestinal tract suggest that function (Gao *et al.*, 2012). All this suggest that totoaba g-type and c-type lysozymes may play important roles in the defense against infectious diseases. However, different studies must be carried out, such as the temporal expression of lysozyme after immunological stimuli, in complement to their functional characterization at the protein level.

CONCLUSIONS

As the first part of this research project, the full-length cDNA sequences of *Totoaba macdonaldi* g-type lysozyme (TmLyzg) and c-type lysozyme (TmLyzc) were described. Both lysozymes were homologous to other g-type or c-type lysozymes from different fish and higher vertebrate species. They presented conserved structural and functional signatures typical of the lysozyme superfamily as part of glycosyl-hydrolases. TmLyzg and TmLyzc mRNAs' expression profiles revealed that both genes are constitutively expressed in all tissues examined. However, the expression level of each lysozyme gene is higher in specific tissues and reflects its importance on innate immunity. Although further experiments are needed to prove the expression of these lysozymes under an immune challenge, the lysozymes from totoaba may play essential roles in defense against bacterial infections and also may be involved in digestion. In the long term, this knowledge will also contribute to establishing improved immunoprophylactic strategies for the rearing of totoaba, including immune stimulation and vaccination.

CHAPTER II

Overview

In the second part of this research, the c-type and g-type lysozymes from totoaba were recombinantly expressed in *E. coli*, isolated from inclusion bodies, and subjected to *in vitro* refolding experiments. The tested refolding conditions were found promising; however, lysozyme activity could not be demonstrated probably because the proteins were kinetically trapped in a folding intermediate, or because the presence of fusion tags somehow hindered their active site. Further optimization of refolding conditions and removal of fusion tags is required to recover active enzymes.

MATERIALS AND METHODS

Purification and Refolding of Recombinant C-type and G-type Lysozymes from Totoaba

Recombinant Expression

The expression constructs pET-28a(+)-TmLyzg and pET-32a(+)-TmLyzc were synthesized at GenScript® (Piscataway, New Jersey, USA). TmLyzg was cloned at *Nde* I and *Xho* I cloning sites of pET-28a(+) vector, which carries a *Kan* coding sequence that provides resistance to kanamycin and a N-terminal 6xHis-Tag® sequence for protein purification. TmLyzc was cloned at *Kpn* I and *Xho* I restriction sites of pET-32a(+) vector, which contains a *bla* coding sequence that provides ampicillin resistance, a Trx-Tag™ as fusion protein and a 6xHis-Tag® sequence for protein purification at the amino-terminus (Figure 18). A cleavage sequence for PreScission Protease (GE Life Sciences) (Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro) was added at the N-terminus of both TmLyzg and TmLyzc synthetic genes to allow the removal of His-tag and Trx-tag from the expressed proteins.

The expression plasmids were used to transform *E. coli* BL21 Gold (DE3) competent cells by heat shock. Briefly, 100 ng of pET-28a(+)-TmLyzg or pET-32a(+)-TmLyzc were added to 50 µL of chemically competent bacterial cells, mixed carefully and then, the suspensions were incubated for 30 min on ice. Heat shock was carried out at 42 °C for 1 min and then the bacterial suspensions were immediately placed on ice for 5 min. Thereafter, 250 µL of SOC medium were added to the bacterial suspensions and the cultures were grown for 90 min at 37 °C with shaking at 225 rpm. After that time, 100-200 µL of bacterial suspensions were cultured in LB plates added with 100 µg mL⁻¹ of ampicillin (TmLyzc) or 50 µg mL⁻¹ of kanamycin (TmLyzg) and were incubated at 37 °C overnight. A single-isolated colony of each transformation reaction was used to prepare the starting bacterial inoculum for expression assays.



Figure 18. Design of synthetic gene constructs of TmLyzg (A) and TmLyzc (B) for recombinant expression.

Transformed bacteria were grown in 5 mL of LB broth, supplemented with 50 $\mu\text{g mL}^{-1}$ of kanamycin (TmLyzg) or 100 $\mu\text{g mL}^{-1}$ of ampicillin (TmLyzc) as corresponds, at 37 °C with shaking at 225 rpm for 16 h. This comprised the starting bacterial inoculum for each expression assay. Thereafter, a Fernbach flask containing 1 L of LB broth (supplemented with the corresponding antibiotic) was inoculated with 5 mL of the bacterial inoculum, and the cultures were grown at 37 °C with shaking at 200 rpm until the cell count reached 0.6 at OD₆₀₀. Overexpression of TmLyzg and TmLyzc was induced by the addition of 0.5 mM IPTG (isopropyl- β -D-thiogalactopyranoside) and the cells were grown at 25 °C for 18 h with shaking at 220 rpm. Bacterial cells were harvested by centrifugation at 7 000 x g for 20 min at 4 °C and then suspended in lysis buffer (50 mM Tris-HCl, 100 mM NaCl, 5 mM EDTA, 5 mM benzamidine, 1 mM DTT at pH 7.0), at a ratio of 4 mL per gram of the bacterial pellet. Thereafter, the cells were lysed by sonication with 6 pulses of 10 s at 40% of amplitude at intervals of 1 min, meanwhile, the sample was always kept on ice. The lysate was clarified by centrifugation at 8 000 x g for 20 min at 4 °C to separate soluble and insoluble phases, which were then analyzed by 12% SDS-PAGE.

Inclusion Body Isolation and Solubilization

TmLyzg and TmLyzc inclusion bodies were isolated as previously described by (Vazquez-Morado *et al.*, 2021). In brief, after cell lysis, the cellular precipitate was subjected to a series of 4 cycles to wash and extract lysozymes from inclusion bodies. In each cycle, the pellet was suspended in 4 mL of the corresponding buffer, dissolved by sonication with 6 pulses of 10 s at 40% of amplitude on an ice bath, and then centrifuged at 25 000 x g for 20 min at 4 °C. The resulting pellet was recovered in cycles 1-3, whereas supernatant (which contained solubilized inclusion bodies) was kept in cycle 4. The buffers used in each cycle are described below: Cycle 1 (Washing buffer 1: 50 mM Tris-HCl, 2% Triton X-100, 5 mM EDTA, 5 mM benzamidine, 1 mM DTT, pH 7.0), cycles 2-3 (Washing buffer 2: 50 mM Tris-HCl, 5 mM EDTA, 5 mM benzamidine, 1 mM DTT,

pH 7.0) and cycle 4 (Extraction buffer: 50 mM Tris-HCl, 8 M urea, 5 mM EDTA, 5 mM benzamidine, 1 mM DTT, pH 7.0). After solubilization of inclusion bodies (cycle 4), the suspensions containing TmLyzg and TmLyzc were incubated at 4 °C for 12 h with gently stirring. Then, the solutions were clarified by centrifugation at 25 000 x g for 20 min at 4 °C and the supernatant containing the soluble fraction was recovered (Figure 20 and 21, lane 4) . Protein concentration in the samples was estimated by A_{280} using the extinction coefficients $\epsilon_{A_{280}} = 42\,770\text{ M}^{-1}\text{ cm}^{-1}$ (TmLyzg) and $\epsilon_{A_{280}} = 62\,650\text{ M}^{-1}\text{ cm}^{-1}$ (TmLyzc), calculated from their amino acid sequence using ProtParam tool on (<https://web.expasy.org/protparam/>) (Gasteiger *et al.*, 2005).

Protein Purification by Immobilized Metal Ion Affinity Chromatography

Lysozymes were purified by immobilized metal ion affinity chromatography (IMAC) under denaturant conditions using an HisTrap™ FF column (1 mL) in an Äkta Prime Plus chromatographer (GE Life Sciences), according to the manufacturer's protocol. The column was equilibrated with 10 volumes of binding buffer A (50 mM Tris-HCl, 500 mM NaCl, 8 M urea, pH 7.4) and was loaded with 4 mL of lysozyme solution at $\sim 4\text{ mg mL}^{-1}$ (TmLyzg) and $\sim 6\text{ mg mL}^{-1}$ (TmLyzc). Lysozymes were eluted with a lineal gradient (1-100%) of elution buffer B (50 mM Tris-HCl, 500 mM NaCl, 8 M urea, 500 mM imidazole, pH 7.4) at a flow rate of 1 mL min^{-1} . Eluted fractions with increased absorbance at A_{280} were analyzed by 12% SDS-PAGE.

The purified sample fractions from each lysozyme were pooled and reduced with the addition of 5 mM DTT (final concentration) before proceeding to protein refolding. For TmLyzc, which is predicted to contain four disulfide bonds in its native state, DTT was removed just before conducting refolding experiments to minimize reformation of disulfide bonds. This was carried out using a protein desalting column PD-10 (GE Life Sciences) equilibrated in elution buffer A that did not contain the reducing agent, according to the manufacturer's protocol. TmLyzg was kept in the same elution buffer

and concentrated to 1 mg mL^{-1} by ultracentrifugation at 5 000 rpm and $4 \text{ }^{\circ}\text{C}$ using an Amicon[®] Ultra-15 10K centrifugal filter device (Merk Millipore).

Refolding Assay Conditions

Refolding of lysozymes was carried out using the Pierce[®] Protein Refolding kit (Thermo Scientific), according to the manufacturer's protocol. A primary screen of nine trials was performed to examine the effect of denaturant concentration (0-1.4 M guanidine and 0-0.8 M arginine) and reducing environment (reduced/oxidized glutathione or DTT) on refolding of TmLyzc and TmLyzg. The matrix set-up for TmLyzc and TmLyzg is described in detail in Table IV and Table V, respectively. For TmLyzc, 900 μL of each base refolding buffer were supplemented with 1 mM EDTA and varying ratios of reduced and oxidized glutathione ([GSH]:[GSSG]) as follows: 2 mM GSH:0.2 mM GSSG; 2 mM GSH:0.4 mM GSSG; 1 mM GSH:1 mM GSSG (final concentration), as determined by the matrix layout shown in Table VI. For TmLyzg, 900 μL of each base refolding buffer were supplemented with 1 mM EDTA and 5 mM DTT (final concentration) as described in Table VII. All solutions were equilibrated to $4 \text{ }^{\circ}\text{C}$ on ice. Then, denatured lysozymes were added to the refolding buffers to a final concentration of $50 \text{ } \mu\text{g mL}^{-1}$ in cycles of addition of 10 μL of protein solution (at 1 mg mL^{-1}), immediate vortex and incubation on ice for 2 min until 50 μL of lysozyme solution were added to each reaction. Refolding was allowed to proceed for 18 h at $4 \text{ }^{\circ}\text{C}$ and refolding yields were determined by measuring lysozyme activity according to the methods described below and by 12% SDS-PAGE.

Table IV. Matrix design for refolding of TmLyzc under reducing conditions (GSH:GSSG).

Tube #	Base Refolding Buffer (900 μL)	100 mM EDTA (μL)	200 mM GSH (μL)	100 mM GSSG (μL)	H₂O (μL)	Lysozyme [1mg/mL] (μL)*
1	1	10	10	2	28	50
2	2	10	10	4	26	50
3	3	10	5	10	25	50
4	4	10	10	4	26	50
5	5	10	5	10	25	50
6	6	10	10	2	20	50
7	7	10	5	10	25	50
8	8	10	10	2	28	50
9	9	10	10	4	26	50

*Denatured lysozyme in 8 M urea, 50 mM Tris, pH 7.4. Final protein concentration in the reaction was 50 μ g mL⁻¹.

Table V. Matrix design for refolding of TmLyzg under reducing conditions (DTT).

Tube #	Base Refolding Buffer (900 μL)	100 mM EDTA (μL)	500 mM DTT (μL)	H₂O (μL)	Lysozyme [1mg/mL] (μL)
1	1	10	10	30	50
2	2	10	10	30	50
3	3	10	10	30	50
4	4	10	10	30	50
5	5	10	10	30	50
6	6	10	10	30	50
7	7	10	10	30	50
8	8	10	10	30	50
9	9	10	10	30	50

*Denatured lysozyme in 8 M urea, 50 mM Tris, 5 mM DTT pH 7.4. Final protein concentration in the reaction was 50 μ g mL⁻¹.

Table VI. Base refolding formulation for TmLyzc refolding experiments.

Base Refolding Buffer*	Factor 1 Guanidine (M)**	Factor 2 L-Arginine (M)	Factor 3 GSH (mM)	Factor 4 GSSG (mM)
1	0	0	2	0.2
2	0	0.44	2	0.4
3	0	0.88	1	1
4	0.55	0	2	0.4
5	0.55	0.44	1	1
6	0.55	0.88	2	0.2
7	1.1	0	1	1
8	1.1	0.44	2	0.2
9	1.1	0.88	2	0.4

*Each base buffer contains the indicated denaturant concentrations as well as 55 mM Tris, 21 mM NaCl, 0.88 mM KCl, adjusted to pH 8.2. **Addition of solubilized protein sample supplied additional 0.4 M urea to each base refolding buffer.

Table VII. Base refolding formulation for TmLyzg refolding experiments.

Base Refolding Buffer*	Factor 1 Guanidine (M)**	Factor 2 L-Arginine (M)	Factor 3 DTT (mM)
1	0	0	5
2	0	0.44	5
3	0	0.88	5
4	0.55	0	5
5	0.55	0.44	5
6	0.55	0.88	5
7	1.1	0	5
8	1.1	0.44	5
9	1.1	0.88	5

*Each buffer contains the indicated denaturant concentrations as well as 55 mM Tris, 21 mM NaCl, 0.88 mM KCl, adjusted to pH 8.2. **Addition of solubilized protein sample supplied additional 0.4 M urea to each base refolding buffer.

Lysozyme Activity Assays

Turbidimetric assay. Lysozyme activity was determined by the turbidimetric assay according to (Shugar, 1952). A suspension of 0.15 mg mL^{-1} of *Micrococcus luteus* in 50 mM potassium phosphate buffer, pH 8.2, with an initial OD_{450} of ~ 0.7 was used as substrate. TmLyzc and TmLyzg refolding reactions at $50 \text{ } \mu\text{g mL}^{-1}$ (as described above) and concentrated TmLyzc and TmLyzg (from refolding reactions) at 0.8 mg mL^{-1} were used as the enzyme solution. The reaction was started by adding $100 \text{ } \mu\text{L}$ of enzyme solution were added to 2.5 mL of substrate suspension and lysozyme activity was measured as the decrease in optical density at 450 nm for 20-30 min at $25 \text{ } ^\circ\text{C}$. Hen egg white lysozyme (HEWL) at 0.8 mg mL^{-1} was used as the control reaction, whereas the substrate without HEWL was used as the correction blank. One unit of lysozyme is defined as a ΔA_{450} of 0.001 per minute at $25 \text{ } ^\circ\text{C}$ in a 2.6 mL reaction mixture (Shugar, 1952).

Solid phase assay. Lysozyme activity was also determined by solid phase assay, which is a modification of the turbidimetric method (de-la-Re-Vega *et al.*, 2004). In this approach, a suspension of 0.05% (w/v) of *M. luteus* in 50 mM potassium phosphate buffer, pH 8.2, was mixed with 1% agarose and poured in a Petri dish. Wells of $\sim 3 \text{ mm}$ were punched out from the solid media and were filled with $20 \text{ } \mu\text{L}$ of TmLyzc or TmLyzg (at $50 \text{ } \mu\text{g mL}^{-1}$ and 0.8 mg mL^{-1}), using HEWL (at 0.8 mg mL^{-1} and 0.1 mg mL^{-1}) as positive controls. The plates were incubated overnight at $37 \text{ } ^\circ\text{C}$ and lysozyme activity was observed as a cleared halo zone where the enzyme lysed *M. luteus* opaque substrate.

Binding Affinity Experiments by Isothermal Titration Calorimetry (ITC)

Refolding reactions where TmLyzg and TmLyzc were detected as presumably folded (soluble fraction) were pooled and extensively dialyzed against 50 mM potassium phosphate buffer, pH 8.2, for 16 h with two exchanges of buffer. The dialyzed protein solutions were clarified by centrifugation at $25\ 000 \times g$ for 20 min at $4 \text{ } ^\circ\text{C}$ and the soluble

fractions were concentrated to 0.6 mg mL^{-1} (TmLyzg) and 1 mg mL^{-1} (TmLyzc) by centrifugation at 5 000 rpm and $4 \text{ }^{\circ}\text{C}$ using an Amicon[®] Ultra-15 10K centrifugal filter device (Merk Millipore). Working solutions of $8 \text{ }\mu\text{M}$ TmLyzg, $11 \text{ }\mu\text{M}$ TmLyzc and $200 \text{ }\mu\text{M}$ Tri-N-acetylchitotriose (Toronto Research Chemicals, Canada) were prepared using dialysis buffer (50 mM potassium phosphate buffer, pH 8.2). HEWL ($10 \text{ }\mu\text{M}$) was prepared in the same buffer and used as control. ITC experiments were performed at $30 \text{ }^{\circ}\text{C}$ using a Microcal VP-ITC calorimeter (Malvern Panalytical). The binding reaction was monitored by recording the heat released upon small additions of Tri-N-acetylchitotriose to the protein solution. A total of 30 aliquots of titrant ($2 \text{ }\mu\text{L}$ each) were injected, while the stirring-syringe was kept rotating at 394 rpm. The heat of dilution of the saccharide was obtained by adding ligand to a buffer solution under identical conditions as with the protein sample. ITC data was analyzed using Origin software supplied with the calorimeter (García-Hernández *et al.*, 2003).

RESULTS AND DISCUSSION

Recombinant Expression and Purification of the C-type and G-type Lysozymes from Inclusion Bodies

The g-type and c-type lysozymes from totoaba were overexpressed as recombinant proteins in *E. coli*. TmLyzg was produced as the mature polypeptide plus poly-histidine tag and protease cleavage sequences at its N-terminus. In contrast, TmLyzc was produced as a fusion protein with thioredoxin (Trx-tag) plus poly-histidine tag and protease cleavage sequences at its N-terminal for protein purification and tag removal by proteolysis. TmLyzg and TmLyzc fusion proteins were expressed after 18 h of induction with 0.5 mM IPTG at 25 °C and were identified as 23.2 kDa (TmLyzg) and 27.9 kDa (TmLyzc) bands by analysis of the bacterial pellet on SDS-PAGE (Figure 19). The electrophoretic analysis of soluble cytosolic and insoluble fractions of the induced bacterial lysate indicated that TmLyzg and TmLyzc were expressed as inclusion bodies (Figure 19). Therefore, a protocol for inclusion body solubilization and refolding was carried out for both recombinant proteins.

TmLyzg (~23 kDa) and TmLyzc (~28 kDa) were the most abundant proteins in the urea-solubilized inclusion bodies, as shown in Figure 20 and Figure 21 respectively. However, several contaminant proteins were present in the samples. Thus, denatured TmLyzg and TmLyzc were subjected to purification by IMAC chromatography before performing the refolding experiments. Denatured lysozymes were eluted at ~250 mM (TmLyzg) and ~200 mM (TmLyzc) of imidazole, respectively, allowing the removal of many contaminant proteins. Although complete purification of lysozymes might require further chromatographic steps, TmLyzg and TmLyzc were observed as the most abundant proteins in the samples before refolding (Figures 20 and 21).

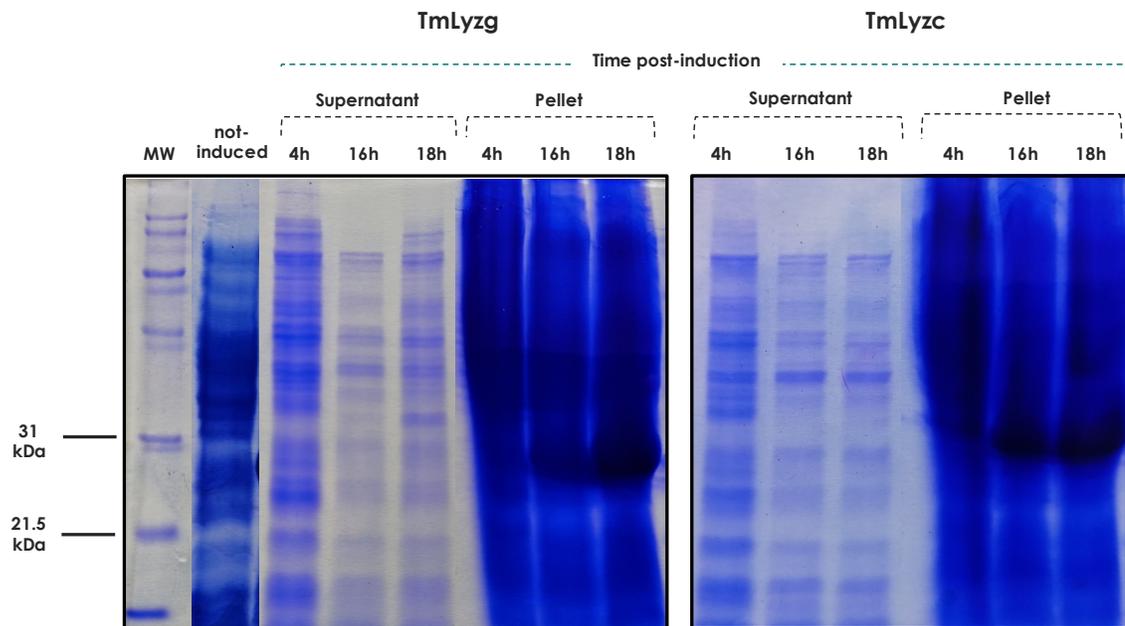


Figure 19. SDS-PAGE analysis of the kinetics of recombinant expression of TmLyzg and TmLyzc in *E. coli* BL21 Gold (DE3). Samples were collected at 4, 16 and 18 h after induction with 0.5 mM IPTG and soluble (supernatant) and insoluble (pellet) phases were examined. Uninduced bacterial lysate was used as expression control. Proteins were separated on 12% acrylamide gels and stained with Coomassie blue.

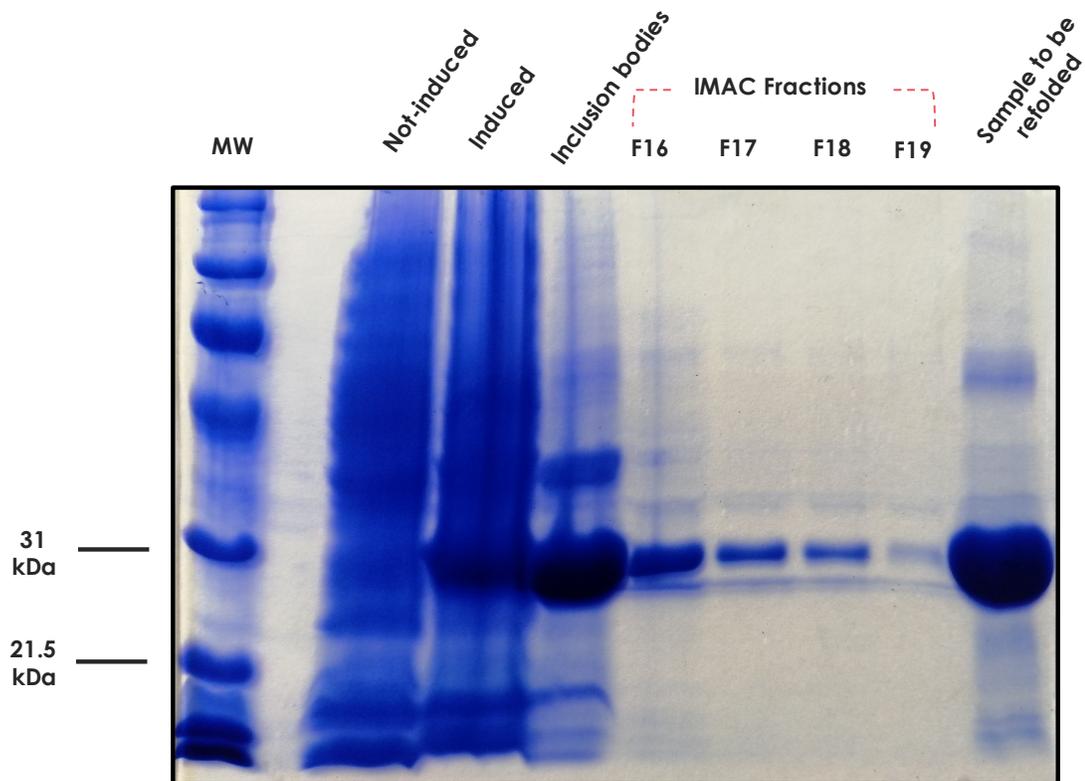


Figure 20. SDS-PAGE analysis for the purification of recombinant TmLyzg by IMAC chromatography at denaturing conditions using 8 M urea. Samples are as follows: lane 1, uninduced bacterial lysate; lane 2, bacterial lysate after induction with IPTG (18 h post-induction); lane 3, inclusion bodies in 8 M urea; lanes 4-7, chromatographic fractions of the eluted protein; lane 8, sample to be refolded (pool of chromatographic fractions). Proteins were separated on 12% acrylamide gels and stained with Coomassie blue.

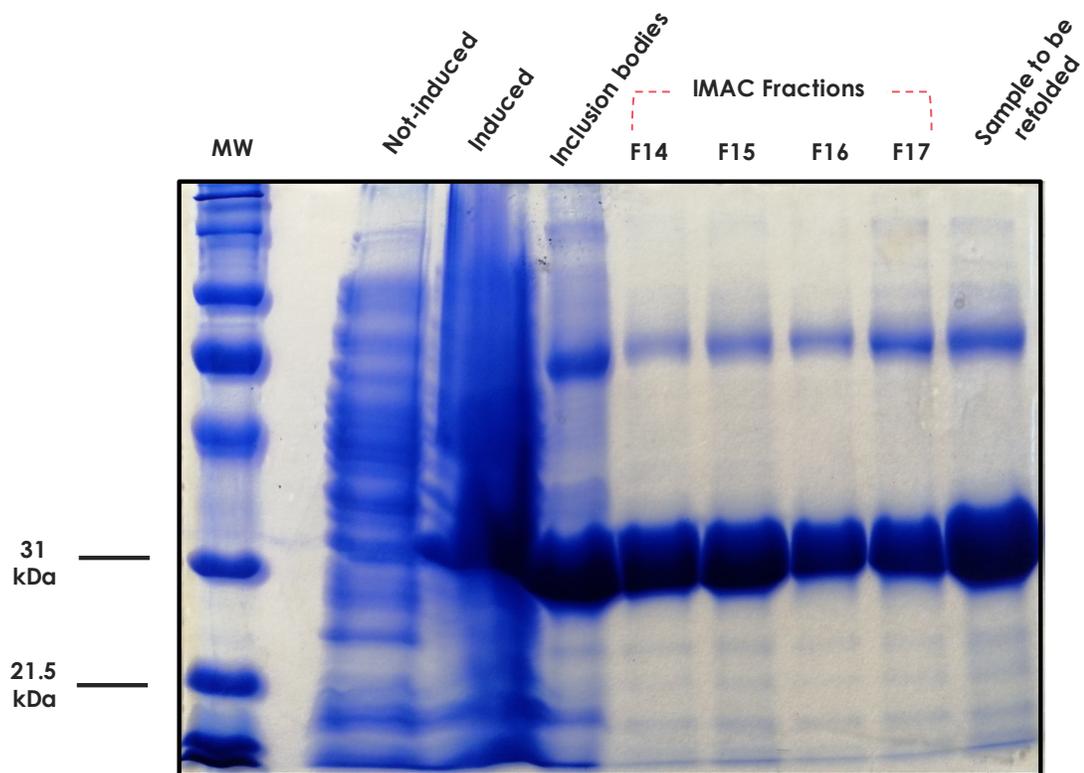


Figure 21. SDS-PAGE analysis for the purification of recombinant TmLyzc by IMAC chromatography at denaturing conditions using 8 M urea. Samples are as follows: lane 1, uninduced bacterial lysate; lane 2, bacterial lysate after induction with IPTG (18 h post-induction); lane 3, inclusion bodies in 8 M urea; lanes 4-7, chromatographic fractions of the eluted protein; lane 8, sample to be refolded (pool of chromatographic fractions). Proteins were separated on 12% acrylamide gels and stained with Coomassie blue.

Protein concentration was monitored during the inclusion body isolation and purification of lysozymes, and total yields of 13.5 mg and 20 mg (per g of bacterial pellet) were estimated for TmLyzg and TmLyzc, respectively. Since high levels of purity (~80-90%) of the denatured target proteins are not required for successful refolding, the eluted IMAC fractions where TmLyzg and TmLyzc were detected as the most prominent proteins were pooled and prepared (see Materials & Methods section) to carry out the refolding experiments.

***In vitro* Refolding of Denatured C-type and G-type Lysozymes**

The refolding of denatured TmLyzg and TmLyzc was assayed at 9 different folding conditions using the Pierce Protein Refolding kit. This system is based on a fractional factorial matrix design, where only the conditions and components having the most significant and general utility in refolding buffers are examined, with the possibility to adjust and optimize refolding conditions. In all refolding experiments, TmLyzc was treated as a disulfide-containing protein using GSH:GSSG to control the redox environment, whereas TmLyzg was treated as a non-disulfide-containing protein, maintaining a reducing environment during refolding with DTT.

The results for refolding denatured TmLyzc and TmLyzg, along with the summary of the examined folding conditions are reported in Table V and Table VI, respectively. The folding experiments of TmLyzc show that protein refolding was suppressed at the lowest denaturant concentrations present within the refolding matrix (Table VIII, trial 1). Moreover, protein precipitation was only observed in the absence of L-arginine at these conditions. However, soluble TmLyzc was found in refolding trials 2-9, as evidenced by electrophoretic analysis of the supernatants of the clarified reactions (Figure 22). A single band of ~27.9 kDa was observed and coincides with the expected electrophoretic pattern of TmLyzc. Indeed, it seems that protein refolding also enhanced protein purity.

Table VIII. Experimental results of refolding TmLyzc.

Trial/Base Refolding Buffer*	Guanidine (M)**	L-Arginine (M)	Redox Environment	Protein Precipitation	Relative Percent Recovery
1	0	0	2 mM GSH:0.2 mM GSSG	Yes	ND
2	0	0.4	2 mM GSH:0.4 mM GSSG	No	ND
3	0	0.8	1 mM GSH:1 mM GSSG	No	ND
4	0.55	0	2 mM GSH:0.4 mM GSSG	No	ND
5	0.55	0.4	1 mM GSH:1 mM GSSG	No	ND
6	0.55	0.8	2 mM GSH:0.2 mM GSSG	No	ND
7	1.1	0	1 mM GSH:1 mM GSSG	No	ND
8	1.1	0.4	2 mM GSH:0.2 mM GSSG	No	ND
9	1.1	0.8	2 mM GSH:0.4 mM GSSG	No	ND

*Each buffer contains the indicated denaturant and redox concentrations as well as 50 mM Tris, 18 mM NaCl, 8 mM KCl, 1 mM EDTA adjusted to pH 8.2.

** 0.4 M urea is provided with the added protein.

Relative percent of recovery as the percentage of lysozyme activity measured by turbidimetric assay.

ND, no lysozyme activity was detected

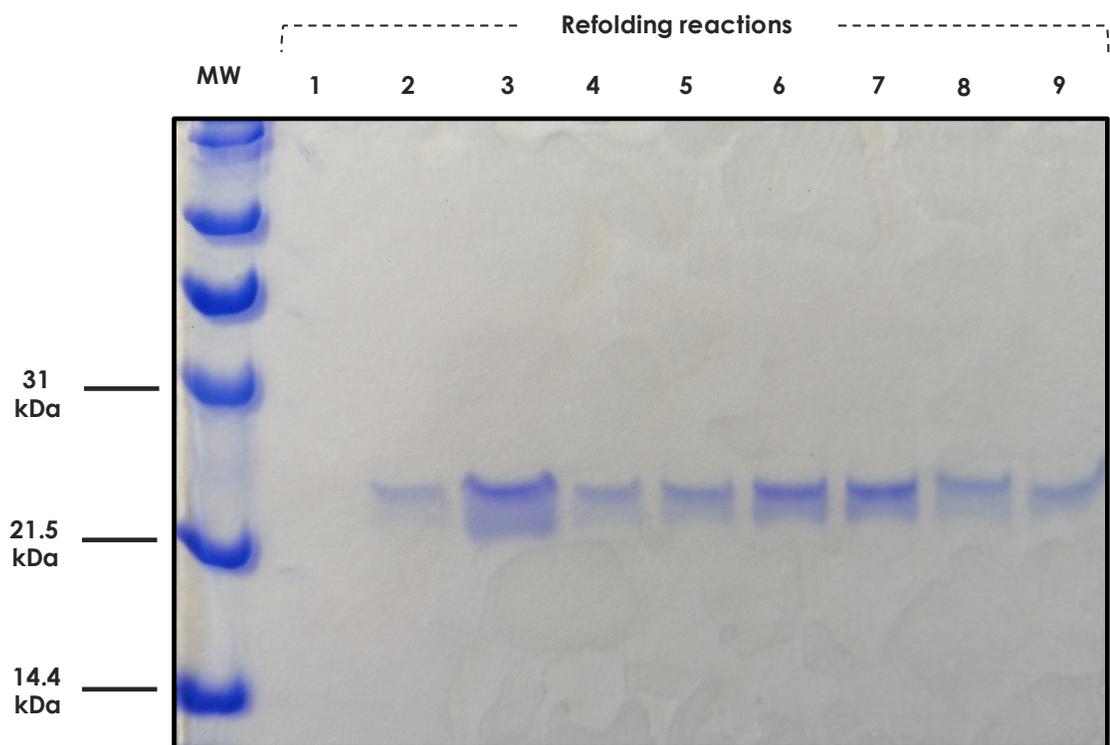


Figure 22. SDS-PAGE analysis of the renaturation of TmLyzc inclusion bodies using the Pierce Protein Folding kit system. Samples are the soluble fractions of refolding reactions and samples in lanes 1-9 represent the nine different refolding conditions tested in this experiment in sequential order. Proteins were separated on 12% acrylamide gels and stained with Coomassie blue.

Even though the soluble protein was detected in all these reactions, a slightly higher band density in SDS-PAGE was noticed in trials 3, 6, and 7 that may suggest better refolding conditions. On the other hand, TmLyzg presented a very similar behavior. The folding experiments of TmLyzg show that refolding was also suppressed at the lowest denaturant concentrations within the refolding matrix, as protein precipitation was observed in trial 1 (Table IX). Nevertheless, the soluble protein was found in all refolding trials (1-9) as shown by SDS-PAGE analysis (Figure 23). A prominent band of approximately 23 kDa was detected in the supernatants of the clarified refolding reactions, which is consistent with the predicted molecular weight of TmLyzg. In addition, a slightly more protein concentration was noticed in refolding trials 2, 3, 4, 5, and 8.

The percentage of recovery of the refolded TmLyzc and TmLyzg was evaluated by measuring lysozyme activity in each refolding trial. For this purpose, turbidimetric and solid-phase lysoplate assays were performed using *M. luteus* as substrate and hen-egg white lysozyme as a positive control. The results show that no lysozyme activity was detected in TmLyzc (Table V and Figure 24B) nor in TmLyzg (Table VI and Figure 24A), as compared with hen-egg white lysozyme as control (Figure 24C). These results may indicate that although TmLyzc and TmLyzg were presumably refolded (or at least were soluble) under different folding conditions, protein concentration in each activity assay was too low or refolding buffers were not suitable to measure lysozyme activity properly. To prove these speculations, the soluble fractions of TmLyzc and TmLyzg refolding reactions were pooled, dialyzed extensively against lysozyme activity buffer, clarified and then, concentrated by ultrafiltration. Thereafter, lysozyme activity was once more measured at 0.8 mg mL⁻¹ by turbidimetric and lysoplate assay. However, although the presence of soluble TmLyzc and TmLyzg was again evidenced by protein electrophoresis (Figure 25A), no lysozyme activity could be detected even at these conditions as compared to control (Figure 25B).

With the attempt to investigate if the reformation of native conformation of TmLyzc and TmLyzg was accomplished at these refolding conditions, binding of lysozymes with

Table IX. Experimental results of refolding TmLyzg.

Trial/Base Refolding Buffer*	Guanidine (M)**	L-Arginine (M)	Redox Environment	Protein Precipitation	Relative Percent Recovery
1	0	0	5 mM DTT	Yes	ND
2	0	0.4	5 mM DTT	No	ND
3	0	0.8	5 mM DTT	No	ND
4	0.55	0	5 mM DTT	No	ND
5	0.55	0.4	5 mM DTT	No	ND
6	0.55	0.8	5 mM DTT	No	ND
7	1.1	0	5 mM DTT	No	ND
8	1.1	0.4	5 mM DTT	No	ND
9	1.1	0.8	5 mM DTT	No	ND

*Each buffer contains the indicated denaturant and redox concentrations as well as 50 mM Tris, 18 mM NaCl, 8 mM KCl, 1 mM EDTA adjusted to pH 8.2.

** 0.4 M urea is provided with the added protein.

Relative percent of recovery as the percentage of lysozyme activity measured by turbidimetric assay.

ND, no lysozyme activity was detected

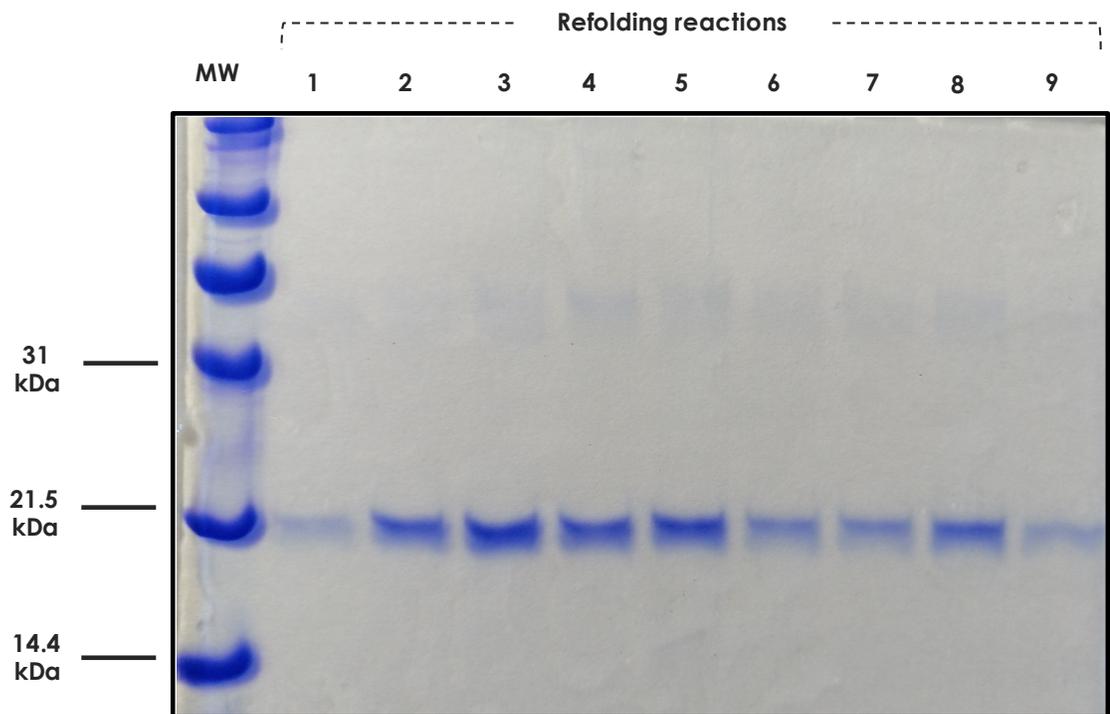


Figure 23. SDS-PAGE analysis of the renaturation of TmLyzg inclusion bodies using the Pierce Protein Folding kit system. Samples are the soluble fractions of refolding reactions and samples in lanes 1-9 represent the nine different refolding conditions tested in this experiment in sequential order. Proteins were separated on 12% acrylamide gels and stained with Coomassie blue.

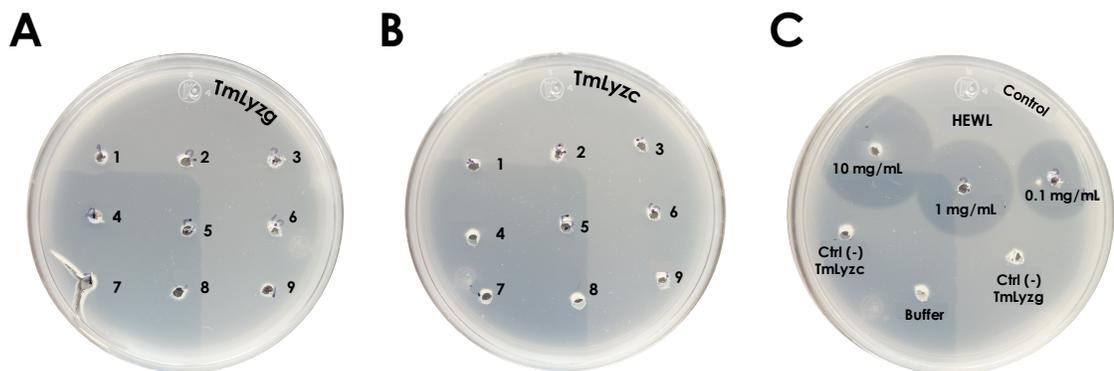


Figure 24. Lysozyme activity by solid phase assay of TmLyzg and TmLyzc refolding trials using *Micrococcus luteus* (0.05% w/v) on 1% agarose gel. A) Refolding trials (1-9) of TmLyzg, B) Refolding trials (1-9) of TmLyzc, and C) lysozyme activity controls: hen-egg lysozyme (HEWL) as positive control and solubilization buffers with and without DTT (TmLyzg and TmLyzc, respectively) in where the inclusion bodies of recombinant lysozymes were solubilized prior to refolding experiments. Wells of ~3 mm were filled with 20 μ L of TmLyzc and TmLyzg at 50 μ g mL⁻¹.

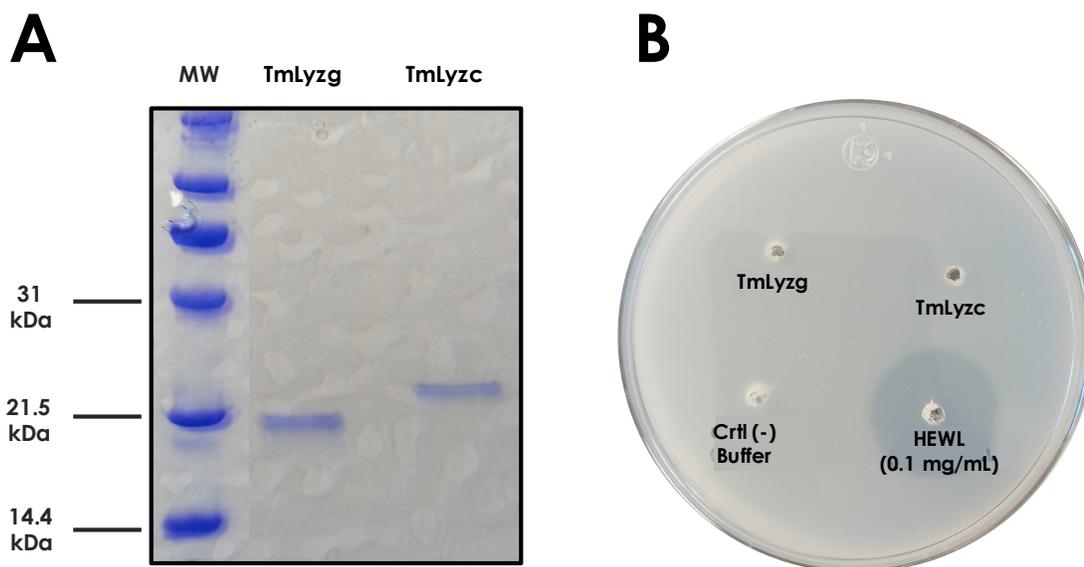


Figure 25. SDS-PAGE analysis and solid phase lysozyme assay of refolded recombinant TmLyzc and TmLyzg. In both cases, lysozymes from refolding trials (soluble fractions) were pooled, extensively dialyzed against 0.5 M phosphate buffer (pH 8.2), clarified and concentrated by ultrafiltration. In the lysoplate activity assay, wells of ~3 mm were filled with 20 μ L of TmLyzc and TmLyzg at 0.8 mg mL⁻¹, and HEWL was used as positive control.

the inhibitor triacetyl-chitotriose (used as substrate analog) was measured by isothermal titration calorimetry (ITC) (data not shown). The results indicated that the calorimetric signal was weak (not enough) and no clue of binding between triacetyl-chitotriose and TmLyzc, TmLyzg, or even hen-egg lysozyme was detected. This can be due to multiple factors, including low protein concentration in the reaction, inadequate buffer composition (pH, ionic strength) and inappropriate reaction conditions, thus ITC assays must be further optimized.

The production of fully functional recombinant proteins is an intricate process and the details involved in the process of protein folding are complex. *Escherichia coli* is a well-characterized and the most widely used prokaryotic host system for heterologous protein expression due to easy handling, inexpensive cultivation, and large scale of production (Ferrer-Miralles and Villaverde, 2013). However, the expression of eukaryotic proteins in a bacterial system has been always challenging, especially when these proteins contain disulfide bonds or when they are toxic to the host. Disulfide bonds are a common feature in eukaryotic proteins, and they are crucial for maintaining proper protein folding, stability, and activity (Wedemeyer *et al.*, 2000). In general, the cytoplasm of *E. coli* is not favorable for the expression of proteins containing disulfide bonds, since this environment is constantly kept reduced. Therefore, the formation and folding of disulfide-bond containing proteins in bacterial cytosol is unstable and normally results in the formation of inactive inclusion bodies (Villaverde and Mar Carrió, 2003; Landeta *et al.*, 2018). Moreover, when the protein that is being synthesized is toxic for the host, the bacteria protect itself by producing it as inactive inclusion bodies (Dumon-Seignovert *et al.*, 2004; Rosano and Ceccarelli, 2014). The production of recombinant proteins in the form of inclusion bodies may represent an additional advantage in protein purification and most recent research highlights their diverse applications in industrial and medical biotechnology (Singhvi *et al.*, 2020). However, additional *in vitro* refolding is required to obtain biologically functional recombinant proteins and it is well known that this process is often unpredictable and challenging (Yamaguchi and Miyazaki, 2014).

Bacterial expression of lysozyme is not a trivial task due to problems with protein folding. Recombinant TmLyzc and TmLyzg were expressed in the form of inclusion bodies and such behavior was also observed in other lysozymes when *E. coli* was used as the expression system. These include lysozymes from different species such as hen-egg (Schlörb *et al.*, 2005), equine (Časaitė *et al.*, 2009), canine (Koshiha *et al.*, 1999) and tobacco hornworm (*Manduca sexta*) (García-Orozco *et al.*, 2005) lysozymes. Amongst lysozymes from marine species, marine shrimp (de-la-Re-Vega *et al.*, 2004), taimen (Li *et al.*, 2016), redbtail shrimp (Cai *et al.*, 2019) and mud crab (Xie *et al.*, 2019) c-type lysozymes were also expressed as inclusion bodies. On the other hand, expression of soluble hen-egg and equine lysozymes was achieved when fused with the OmpA protein (Fischer *et al.*, 1993) and DsbA protein (Časaitė *et al.*, 2009), respectively. However, the fused proteins were active and the host cells were lysed, making the growth of cells hard to control.

A more innovative approach to produce soluble active lysozyme was reported by (Lamppa *et al.*, 2013) with human lysozyme, where *E. coli* was genetically engineered to co-produce lysozyme along with its Ivy inhibitor, DsbC isomerase and other chaperones/foldases. This system was successful, but its complexity demonstrates that producing lysozyme as a soluble recombinant protein, which may also be harmful to the *E. coli* host cells, is not trivial but multifactorial. Regarding lysozymes from fish species, the g-type lysozymes from mandarin fish (Sun *et al.*, 2006), orange-spotted grouper (Wei *et al.*, 2014), Indian major carp (Mohapatra *et al.*, 2019), and the c-type lysozymes from red-spotted grouper (Mai *et al.*, 2014), orange-spotted grouper (Wei *et al.*, 2012) and Asian seabass (Fu *et al.*, 2013) were expressed in *E. coli* as only a small proportion of soluble proteins, but enough to further characterize these enzymes. Moreover, some systems may be useful to express at least a small proportion of soluble lysozymes when these are fused with protein such as maltose-binding protein, glutathione-s-transferase, and thioredoxin (Trx) (Costa *et al.*, 2014). Nevertheless, our experience with TmLyzc and TmLyzg show that recombinant expression of these proteins is not always predictable but involves multifactorial details.

In the attempt to obtain fully functional recombinant TmLyzc and TmLyzg, the denatured proteins were subjected to *in vitro* refolding process. The results showed that both lysozymes were soluble, and consequently were refolded at different experimental conditions. However, lysozyme activity could not be detected nor demonstrated in any sample. These findings suggest, at least, two possible events that may have occurred with TmLyzc and TmLyzg throughout activity assays or during the refolding process. The first one and the most straightforward is that at least a proportion of these lysozymes were refolded correctly, but fusion tags at their N-terminus may be hindering lysozyme activity. Recombinant TmLyzc and TmLyzg expressed in *E. coli* cells have additional N-terminal amino acid residues required for the construction of the recombinant plasmid, purification and protease digestion (see recombinant plasmid maps in Materials & Methods section). It is precise to note that cleavage of these additional residues was not done prior to measuring lysozyme activity.

Several reports have indicated that additional N-terminal amino acid residues can influence protein stability, folding and activity. In hen-egg lysozyme, an additional methionine residue attached to its N-terminus was enough to depress refolding yield and decrease protein solubility (Mine *et al.*, 1997). Foreign residues such as methionine, serine and proline attached to the N-terminal of human lysozyme negatively affected the conformational stability of the protein (Takano *et al.*, 1999). This was probably by removal of intramolecular hydrogen bond networks between the protein and two ordered water molecules identified at the N-terminal region of lysozymes and their homologous proteins α -lactalbumins (Acharya *et al.*, 1989; Motoshima *et al.*, 1997; Ohmura *et al.*, 1997; Chrysina *et al.*, 2000). Moreover, structural changes of other parts far from the N-terminal due to extra amino acid residues at this region might also affect lysozyme stability (Goda *et al.*, 2000). Studies with bovine, goat and human α -lactalbumins have reported a similar behavior (Ishikawa *et al.*, 1998; Chaudhuri *et al.*, 1999; Makabe *et al.*, 2013). The presence of an extra methionine residue at the N-terminus contributes unfavorably to protein stability, increasing its unfolding rate probably by an excess of conformational entropy at the N-terminal region in the transition state of refolding of these enzymes

(Chaudhuri *et al.*, 1999). In fact, the destabilization of recombinant lysozymes has shown to be enthalpic in nature (Takano *et al.*, 1999), whereas the destabilization of recombinant α -lactalbumins is entropic (Ishikawa *et al.*, 1998). The translation of this way of behaving to TmLyzc and TmLyzg must require further investigation but it may represent a possible explanation to the results described above.

Besides, the effect of polyhistidine tags at the N-terminus on the native state stability of recombinant proteins has been reported in several studies. For example, polyhistidine tag in the N-terminal of the transcriptional repressor Arc from *Salmonella phage* had little effect on protein stability and folding (Milla and Sauer, 1994), whereas polyhistidine tag in the same region of the cold shock protein CspA from *E. coli* altered its folding behavior by interacting with the native portion of the protein or by adopting a defined structure (Reid *et al.*, 1998). The crystal structures of these proteins have shown that the structure around the N-terminal residue in CspA is rigid (Schindelin *et al.*, 1994), whereas that of the Arc repressor is flexible (Bonvin *et al.*, 1994). The structure around the N-terminal region of lysozymes and α -lactalbumins is rigid (Chaudhuri *et al.*, 1999). Therefore, these studies together with those of lysozymes and α -lactalbumins suggest that when the N-terminal region of a protein is rigid, the addition of extra residues at this vicinity destabilizes the native state of the protein. On the other hand, when the N-terminal structure is flexible, the extra residues do not interfere with the native state stability of the protein (Chaudhuri *et al.*, 1999). Moreover, (Časaitė *et al.*, 2009) reported that N-terminal polyhistidine-tagged equine lysozyme was inactive, but its activity was completely restored after removal of the extra histidine residues by protease digestion. Thus, alterations of lysozyme N-terminus can directly influence its activity. In this sense, removal of fusion-tags from TmLyzc and TmLyzg is needed to investigate the effect of these residues on lysozyme activity.

Another possible explanation for the lack of activity in the refolded recombinant lysozymes could be caused by enzyme inhibition. A competitive inhibition of hen lysozyme has been observed with imidazole and indole derivatives, that can form a

charge-transfer complex with tryptophane residues and acidic residues of lysozyme (Swan, 1972). Low resolution crystallographic analysis of hen lysozyme crystals diffused with histidine showed an increase in electron density at the active site, suggesting the interaction of imidazole ring with key active site residues (Swan, 1972). It is important to note that polyhistidine tag is commonly used to direct purification of recombinant proteins with IMAC. This strategy was used in this work for the purification process but the removal of the polyhistidine tag was not assayed. The presence of the polyhistidine tag in TmLyzc and TmLyzg could be obstructing the active site cleft for the substrate and retrieving inactive enzymes. On the other hand, it has been reported that hen lysozyme can form soluble oligomers during the initial phases of the aggregation process. These oligomeric forms are distinguished by the presence of misfolded species that have undergone partial unfolding of the α/β -structure characteristic of the native monomeric lysozyme (Frare *et al.*, 2009). Furthermore, some reports suggest that the use of tri-*N*-acetylchitotriose as ligand for lysozyme could suppress protein aggregation by rising the activation energy barrier for protein misfolding (Kumar *et al.*, 2009). The use of glycosidic ligands as additives during refolding could improve the recovery of active recombinant lysozymes in future works.

The second possibility and more complex to explore and explain has to be with the process of protein folding itself and the intricacy of pathways that a protein has to follow to acquire its full native conformation. It has been shown that the amino acid sequence at the primary structure level has chemical instructions to acquire the tridimensional structure of most proteins (Anfinsen, 1973), but the detailed process at the molecular level is still not well understood (Dill *et al.*, 2008; Dill and MacCallum, 2012). Fortunately, folding of hen egg-white lysozyme (HEWL) has been thoroughly studied by a combination of biophysical techniques and computational simulations (Miranker *et al.*, 1991; Radford *et al.*, 1992; Itzhaki *et al.*, 1994; Kiefhaber, 1995; Mallamace *et al.*, 2016). Thus, knowledge of its *in vitro* folding process gives us insight into the complex and diverse paths in the folding process of globular proteins and so will help us translate this knowledge to give a rational explanation of the refolding process of TmLyzc and TmLyzg.

HEWL is C-type lysozyme of 129 residues (Callewaert and Michiels, 2010). The overall three-dimensional structure HEWL has structural motifs commonly found in globular proteins and is a good representation of the general molecular architecture and folding of enzymes of the lysozyme superfamily. This protein is made up of two structural domains (Artymiuk *et al.*, 1982; Callewaert and Michiels, 2010; Wohlkönig *et al.*, 2010). The α -domain is rich in α -helices comprising from Lys1 to Asn39 and Ser86 to Leu129 (Lys1 to Asn39 and Ser80 to Leu128 in TmLyzc; Met1 to Asn76 and Ser117 to Tyr193 in TmLyzg), while the β -domain contains the β -sheets elements from residues Thr40 to Ser85 (Thr40 to Thr84 in TmLyzc; Thr79 to Thr110 in TmLyzg). Residue numbers correspond to the mature polypeptides of all three lysozymes and from predictions of secondary/tertiary structures in TmLyzc and TmLyzg (Moreno-Córdova *et al.*, 2020) (for details, the reader is referred to Results & Discussion section in Chapter I).

The α -domain contains three large α -helices ($\alpha 1$ - $\alpha 3$), a small helix of six residues ($\alpha 4$), and a short C-terminal 3^{10} helix, including both the N- and C- terminal segments of the protein. The β -domain has a triple-stranded antiparallel β -sheet ($\beta 1$ - $\beta 3$), a short 3^{10} helix, and a long loop. The native structure of HEWL is stabilized by four disulfide bonds, two at the α -domain, one at the β -domain, and another connecting both domains. As stated from the results in Chapter I, counterparts of cysteine residues are not found in the vast majority of g-type lysozymes from marine species, as it is the case of TmLyzg, which are predicted to be intracellular proteins (Moreno-Córdova *et al.*, 2020). However, it is precise to consider that the overall three-dimensional structure of lysozyme is strikingly similar and well preserved across all members of the lysozyme superfamily (Wohlkönig *et al.*, 2010). In this sense, general principles of hen lysozyme folding process could also be translated to other lysozymes such as TmLyzc and TmLyzg.

The folding kinetics of HEWL with its four disulfide bonds intact has been extensively studied by biophysical techniques and therefore, a detailed model for the *in vitro* folding process of this enzyme has been proposed (Dobson *et al.*, 1994, 1998; Radford *et al.*, 1995). This folding process is a complex mechanism, where the two

structural domains (α - and β -domains) that shape the native fold of lysozyme behave as distinct folding domains, although the formation of each domain is a highly cooperative process (Dobson *et al.*, 1994). Moreover, this folding process cannot be reduced to a simple sequential assembly process, since it involves parallel events that show distinct kinetic profiles (Matagne and Dobson, 1998).

The earliest detected events in hen lysozyme folding process include the formation of a substantial proportion of the native secondary structure and a considerable degree of hydrophobic collapse. The experimental findings reveal that at this very early stage of folding, an embryonic native-like structure in the α -domain of lysozyme is formed and this probably plays a significant role to direct the later stages of folding, acting as a template-like structure (Gladwin and Evans, 1996). Moreover, the hydrophobically collapsed state formed at the onset of this process displays characteristics of the equilibrium molten globule state, which is observed as an important intermediate in the general process of protein folding (Radford *et al.*, 1992; Judy and Kishore, 2019). Since the collapsed state is heterogeneous in nature, kinetically distinct populations of folding lysozyme molecules are generated: some molecules already display substantial native-like features after the initial rapid collapse, while others need first to reorganize (Radford *et al.*, 1992, 1995; Kotik *et al.*, 1995).

These populations of refolding molecules shape the native state of lysozyme via two different paths: 25 % of the protein folds through a fast track path, whereas 75 % fold on a slow track path (~75%) (Matagne *et al.*, 1997). The fast track in refolding occurs via an intermediate which has an extensive native-like structure in both the α - and β -domains, but which is not entirely native since the active site cleft ($\alpha\beta$ -domain interface) is not formed yet. This intermediate accumulates because it forms prior to the rate-limiting step, which may be associated with the reorganization and docking of the two partly structured α - and β -domains to form the active site and the fully native state of lysozyme (Matagne *et al.*, 1997). Conversely, in the slow folding track, the active site could be achieved cooperatively with the formation of a native-like β -domain and its integration into the

native lysozyme structure (Matagne *et al.*, 1997). However, the rate-determining step in the formation of the native state for this slowly folding population of molecules involves structural reorganization. This is likely to be associated with the formation and stabilization of structure within the β -domain, and from rearrangement of the interfacial region between the α - and β - domains (Matagne *et al.*, 1997; Matagne and Dobson, 1998).

The number of possible conformations that an extended polypeptide chain can adopt is enormously large, so that systematic search for the native (lowest energy) structure would require infinite length of time (Karplus, 1997). However, protein folding occurs many orders of magnitude more rapidly than predicted by a random search of conformational space (S̃ali *et al.*, 1994; Dill and Chan, 1997). This indicates that constraints on the conformational energy of the protein restricts the configurations visited in the folding process only to a limited number (Melkikh and Meijer, 2018). In addition, the nature of the interactions that develop within the protein structure serve to direct the search towards the native state by decreasing the protein's conformational energy at the cost of configurational entropy loss (Melkikh and Meijer, 2018). In this sense, to understand folding it is necessary to focus on how proteins search conformational space. Energy surfaces or landscapes enable to understand the protein folding process in the light of structural thermodynamics (Wolynes, 2015). This view of folding describes the process as a progressive organization of an ensemble of partially folded structures in parallel events, that explore certain region of the protein's conformational space to find the conformation with lowest energy (native state). Intermediates resulted from the presence of barriers that generate local minima in the energy surface, thus, fast and slow folding populations can simply result from the distributions of molecules encountering such energy minima (Dill and Chan, 1997).

A schematic energy surface has been proposed for hen lysozyme by (Dobson *et al.*, 1998), which is broadly consistent with the experimental findings of the folding process of this protein (Figure 26) (Miranker *et al.*, 1991; Chaffotte *et al.*, 1992; Radford *et al.*, 1992, 1995; Itzhaki *et al.*, 1994; Matagne and Dobson, 1998). This schematic surface is

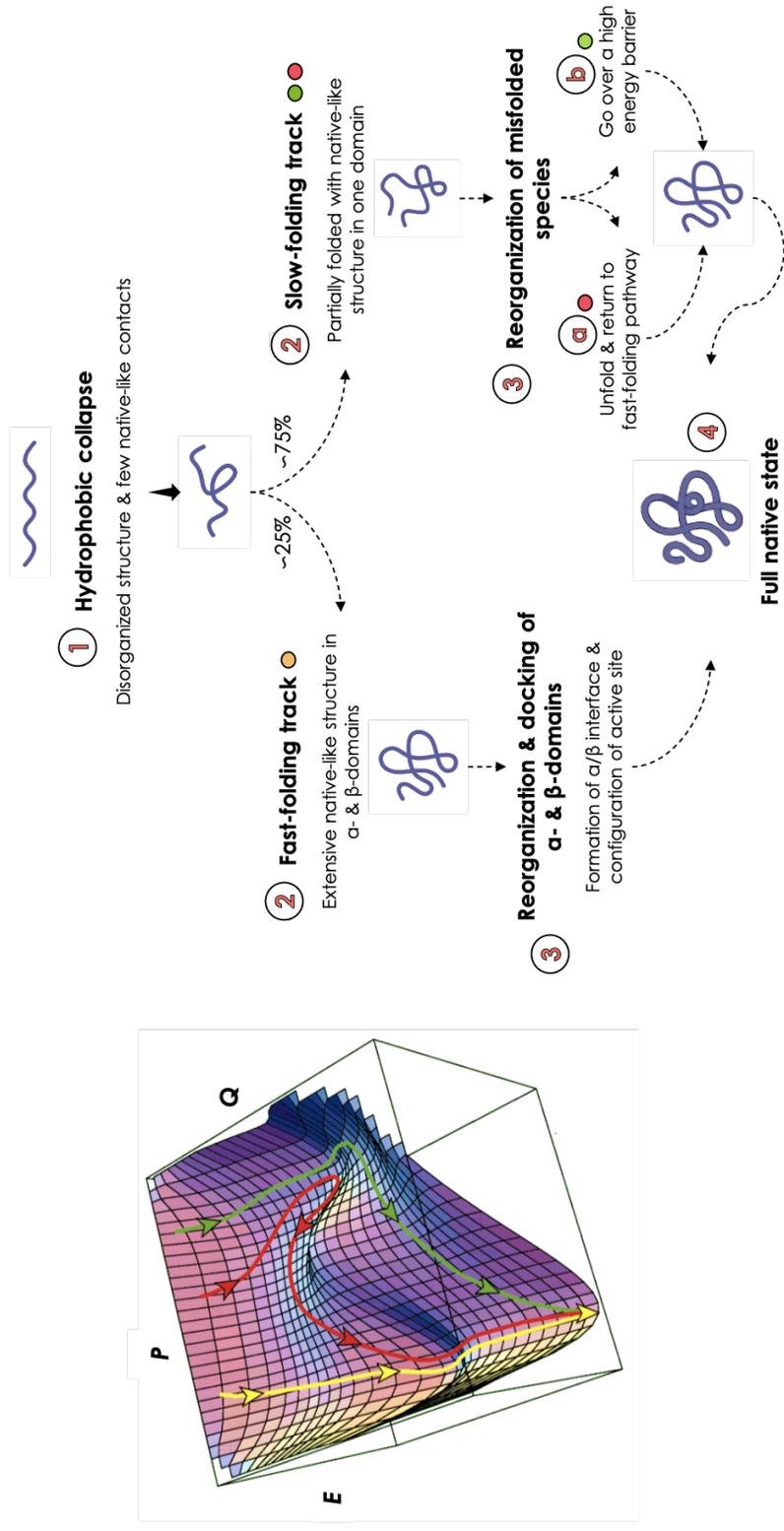


Figure 26. Proposed energy landscape of the folding process of hen-egg lysozyme. Possible folding trajectories are depicted as arrows: Yellow (fast-folding), green (slow-folding that crosses a high energy barrier) and red (slow-folding trajectory that returns to a less folded state and then follows the valley of the fast-folding trajectory). All folding trajectories aim to reach the lowest-energy full native folded state. This illustration is adapted from Matagne and Dobson (1998).

designed to illustrate the three major experimental features of the lysozyme folding reactions (at 20 °C): 1) The early collapse with very rapid formation of a native-like secondary structure, 2) the sampling of distinct intermediate species during folding (search of conformational space) and 3) the existence of complex kinetics arising from partitioning between fast and slow populations of refolding molecules (Dobson *et al.*, 1998). As depicted in Figure 26, this energy surface is based on the results of simulations using simple lattice models to describe protein folding (Dobson *et al.*, 1998). The E axis represents the averaged effective energy of folding, Q axis depicts the fraction of native contacts that are formed at a given stage of folding, and P axis is related to the configurational entropy of the system and describes the extent of conformational space accessible to the polypeptide chain (Dobson *et al.*, 1998).

The first stage of the folding process of hen lysozyme is the early collapse of the polypeptide chain, represented by the steep drop at the top of Figure 26 (the first segments of the yellow, red and green trajectories). This structure is a relatively disorganized globule and is characterized by the presence of few native-like contacts (low values of Q), in which the driving force is the simple burial of hydrophobic groups (Dill *et al.*, 1993; Rothwarf and Scheraga, 1996; Dobson *et al.*, 1998). From this highly heterogeneous collapsed state, native-like structure develops independently in the two structural domains, α - and β -domains, of lysozyme and hence intermediates are formed.

Along the slow folding track (red and green trajectories), a partially folded, and almost native structure relative to both domains, resulted when the α -domain folds in the fast folding path (Figure 26, deep minimum in the right-hand side). On the contrary, the presence of an intermediate state with an extensive native-like structure within both α - and β -domains is identified along the fast folding track (yellow trajectory in Figure 26, energy minimum on the left-hand side of the diagram). The two partly structured α - and β -domains join together in a slower step (representing a barrier), resulting in the formation of the fully functional native state of lysozyme. Some studies strongly suggest that the rate-limiting step in the folding process for the slow folding population of lysozyme

molecules involves the reorganization of misfolded species (Radford *et al.*, 1992; Guo and Thirumalai, 1995; Matagne *et al.*, 1998). These reorganization processes are likely to be trapped by energy barriers associated with interactions within the collapsed protein (Dobson *et al.*, 1998). Hence, in the slow folding track, the misfolded species must jump the energy barrier to fold in the native state (green trajectory), or to take another by unfolding their structure and taking the fast folding track explained before (red trajectory) (Dobson *et al.*, 1998; Matagne and Dobson, 1998). This schematic surface is aimed to make qualitatively congruent between the kinetic partitioning of folding pathways and the distinct intermediate structural species during lysozyme folding at a given set of conditions. It is important to know that this surface will be modified by changing the experimental folding conditions, and so the folding behavior as well (Matagne and Dobson, 1998).

More recently, with the advance of more sophisticated computational tools, molecular dynamics simulations have been used to predict and describe in more detail the events of protein folding kinetics. The findings in the work done by (Muttathukattil *et al.*, 2019) using coarse-grained protein models and molecular dynamics simulations (Liu *et al.*, 2011), exquisitely depict a more detailed picture of the kinetics behind hen lysozyme folding in the light of folding pathways, the structure of the intermediates along folding pathways and the role of disulfide bonds in the process of lysozyme folding. Along with the extensive research that has been conducted for the past years, this new knowledge will help us understand more about protein folding and to construct a rational explanation of our results on lysozyme refolding experiments

In agreement with previous reports (Radford *et al.*, 1992; Guo and Thirumalai, 1995; Kiefhaber, 1995; Matagne *et al.*, 1997), molecular dynamics simulations show that hen lysozyme folds through fast and slow folding pathways, exhibiting kinetic partitioning. Lysozyme folds through a fast folding pathway in 49% of folding trajectories (average folding time ~ 9.4 ms), whereas in 51% of trajectories the protein folds through a slow folding pathway (average folding time > 500 ms), where a partially folded intermediate

state is populated (Kiefhaber, 1995; Radford *et al.*, 1995; Muttathukattil *et al.*, 2019). The folding dynamics of this protein is stochastic, as the initial conformations in all folding trajectories is a heterogeneous mixture of unfolded molecules, therefore, the protein in these conformations can choose either the fast or the slow folding route (Kotik *et al.*, 1995; Matagne *et al.*, 1997). The simulations also broadly support the triangular folding mechanism proposed by (Wildegger and Kiefhaber, 1997), that includes an intermediate with a trapped kinetic state between the unfolded and native states.

Lysozyme folding initiates by forming compact structures in both the fast and slow folding pathways. In the rapid collapse, the formation of the β -domain and α -helices contributes to the compaction in protein dimensions (from $> 25 \text{ \AA}$ to $\sim 20 \text{ \AA}$). As previously described, this compacted state mainly lacks the tertiary contacts present in the α -domain and the interdomain contacts (interface between α - and β -domain). After the compaction, bifurcation in the folding pathways to fast and slow folding routes is observed. In the fast folding pathways, the protein continues to fold to its native structure with the formation of interdomain contacts (Segel *et al.*, 1999). However, in the slow folding pathways, the protein gets trapped in an intermediate state that is close to the native folded state. In both the fast and slow folding pathways, it is the β -domain that folds completely and more rapidly ($\sim 5 \text{ ms}$), reaching the native state. In contrast, the folding of the α -domain is the rate-determining step in the folding of lysozyme, contrary to what has been proposed by (Matagne *et al.*, 1997; Matagne and Dobson, 1998). In the fast folding trajectory, the α -domain folds just a few milliseconds slower than the β -domain (average time scale $\sim 15 \text{ ms}$), reaching the native state. However, in the slow folding trajectories, the α -domain is stuck in an intermediate state showing that the folding of this domain is responsible for the slow folding pathways in lysozyme folding (Muttathukattil *et al.*, 2019).

A close-up view of the folding process of hen lysozyme gives us insights into the major structural components involved and helps us formulate a coherent explanation of the determinants that lead this process. In the folding of the α -domain, the three large α -

helices α 1- α 3 present in hen lysozyme (α 1 Arg5 to Gly22, the central α 2-helix Leu25 to Ser36 that contains the catalytic Glu52, and α 3 Thr89 to Ala107) influence folding kinetics (Muttathukattil *et al.*, 2019). Counterparts in TmLyzc are α 1 Arg5 to Lys13, the central α 2-helix Leu25 to Ser36 that contains the catalytic Glu50, and α 4 Ser88 to Ala106 as depicted in Figure 13, Chapter I. In TmLyzg, the molecular architecture is a bit different compared to c-type lysozymes and so it might include at least the helices α 3 Arg47 to Tyr57, the central α 4-helix Pro61 to Ser72 (which contains the catalytic Glu71), and α 5 Glu119 to Lys138, as depicted in Figure 12, Chapter I. This structural comparison will be useful to translate the findings of hen lysozyme folding into, at least, a general picture of the folding process of other lysozymes such as TmLyzc and TmLyzg.

In both the fast and slow folding pathways of HEWL, the α -helices are formed first and then, the interactions between the secondary structural elements take place. For most part of the protein molecules going through the fast folding track, after the α 2-helix and α 3-helix (α 2 α 3) merged first, this ensemble contacts to α 1-helix at the same time the interdomain contacts between the α -domain and β -domains are formed, resulting in a fully folded native HEWL (Figure 27, fast trajectory 1). For the remaining protein molecules in the fast folding track, first α 1 α 2 is assembled, then the contacts that join α 2 α 3, α 1 α 3, as well as the interdomain contacts occur simultaneously (Figure 27, fast trajectory 2) (Muttathukattil *et al.*, 2019). In the fast folding trajectories, contacts in either α 2 α 3 or α 1 α 2 form first, followed by the simultaneous formation of the other two sets of interhelical contacts and $\alpha\beta$ interdomain contacts. The formation of the interhelical contacts through these pathways allow HEWL to fold straightforward and without the formation of kinetically trapped species (Muttathukattil *et al.*, 2019). In the slow folding pathway, α 1 α 2 contacts form first. This structure prevents the formation of α 2 α 3 and α 1 α 3 contacts because in this trajectory, the orientation of the helix α 1 is misaligned (in the opposite direction) relative to the helices α 2 and α 3 compared to the native folded state. For the protein trapped in this state to fold correctly, α 1 has to reorient by breaking

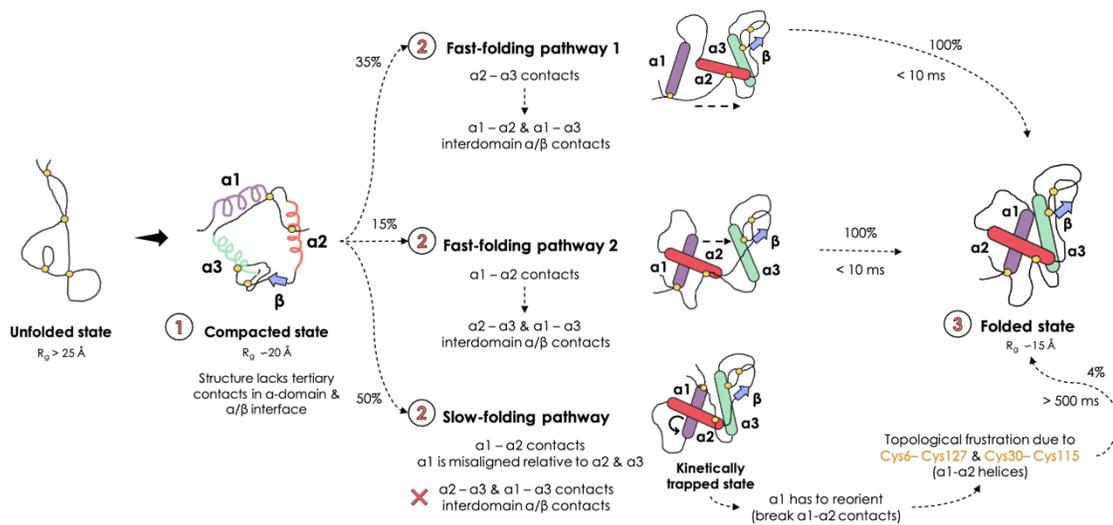


Figure 27. Proposed schematics of dominant folding pathways of hen-egg lysozyme studied by molecular dynamics simulations. The three α -helices, $\alpha 1 - \alpha 3$, that influence the folding kinetics are shown in purple, red, and green, respectively. The triple-stranded β -sheets that fold early are shown as one blue arrow. Folding is initiated from an unfolded state and the polypeptide compacts due to the folding of the β -domain and α -helices. From this compacted state, the folding bifurcates to three different pathways: 1) fast folding pathway with transient intermediates. In fast-folding pathway 1, $\alpha 2 \alpha 3$ contacts are formed initially, followed by $\alpha 1 \alpha 2$, $\alpha 1 \alpha 3$, and interdomain contacts. In fast-folding pathway 2, $\alpha 1 \alpha 2$ contacts are formed initially, and helices ($\alpha 1$ and $\alpha 2$) have the correct orientation relative to the native state. In the later stages, $\alpha 1 \alpha 3$, $\alpha 2 \alpha 3$ and interdomain contacts form, and the protein folds correctly to the native state. 2) slow-folding pathway with a kinetically trapped state. In this state, the orientation of $\alpha 1$ relative to $\alpha 2$ is in the opposite direction compared to the native state. The protein has to backtrack by breaking $\alpha 1 \alpha 2$ contacts to fold to the native state. This rearrangement is restricted by the presence of two disulfide bonds, Cys6-Cys127 and Cys30-Cys115, present in the α -domain ($\alpha 1 - \alpha 2$ helices). This illustration is adapted from Muttathukattil *et al.* (2019).

the $\alpha 1\alpha 2$ contacts, and this leads to a long-lived kinetic intermediate (Figure 27, slow trajectory) (Muttathukattil *et al.*, 2019). These observations strongly suggest the important role of the central $\alpha 2$ -helix, which is part of the substrate binding site and contains the catalytic Glu52, as one of the initiation sites that assist the configuration of the native state fold of hen lysozyme.

Furthermore, the simulations show that in the kinetically trapped intermediate (slow folding pathway), the β -domain is completely folded, and the α -domain is partially folded with $\alpha 2\alpha 3$ and $\alpha 1\alpha 3$ interhelical contacts, but some of the interdomain contacts are missing (Muttathukattil *et al.*, 2019). The kinetically trapped misfolded intermediate has to partially unfold before folding back to the native state, as previously inferred in some experiments (Rothwarf and Scheraga, 1996; Matagne *et al.*, 1998). In these slow folding trajectories, the contacts between the misaligned helices $\alpha 1\alpha 2$ (present in the misfolded intermediate state) need first to break down, allowing $\alpha 1$ to reorient to the correct alignment relative to $\alpha 2$, thereby the protein folds back to the native state. In this sense, the disruption of the $\alpha 1\alpha 2$ contacts are a primary bottleneck for the protein to escape from this trapped state. However, the percentage of molecules that can overcome this trapped state and completely fold in the slow folding trajectories is too low (4%) (Muttathukattil *et al.*, 2019). The models show that the population of the intermediate in the slow folding pathway is not caused by the formation of non-native interactions within lysozyme folding domains, as previously hypothesized (Rothwarf and Scheraga, 1996; Kazmirski and Daggett, 1998). Instead, it should be the topological frustration between $\alpha 1$ and $\alpha 2$ helices, with a possible role of disulfide bonds, which constrains the conformational space available for the protein and facilitates the population of the kinetically trapped intermediate state (Muttathukattil *et al.*, 2019).

The possible role of disulfide bonds in the population of the kinetic intermediate was investigated by mimicking the folding of lysozyme in the presence of reducing agents, such as dithiothreitol, which prevent the formation of disulfide bonds. This scenario also resembles the refolding conditions tested for TmLyzg, which contains no disulfide bonds.

At these conditions, no kinetically trapped intermediate states are observed and all trajectories reached the folded state at ~ 200 ms. This shows that in hen lysozyme conformations, when $\alpha 1$ is misaligned relative to $\alpha 2$, the topological constraints due to the disulfide bonds formation make the protein less flexible for structural interconversion, and these conformations result in kinetic traps in the folding pathways (Muttathukattil *et al.*, 2019). To escape from this conformation, local unfolding or backtracking of lysozyme $\alpha 1\alpha 2$ interhelical contacts is required. The simulations show that the two disulfide bonds Cys6-Cys127 and Cys30-Cys115, attached to $\alpha 1$ and $\alpha 2$ helices in the α -helical domain of hen lysozyme, are responsible for the kinetic partitioning in the folding pathways and explain the topological restrictions in the rearrangement of these helices in the folding process (Denton *et al.*, 1994; Guez *et al.*, 2002; Muttathukattil *et al.*, 2019). The population of a misfolded kinetic intermediate due to topological frustration is also observed in the folding process of green fluorescent protein (Andrews *et al.*, 2008) and cytochrome c (Weinkam *et al.*, 2010).

Finally, another important question to address in the folding process of lysozyme is whether protein folding guides disulfide bond formation or vice versa, as studies have shown evidence for the former but not the latter (Qin *et al.*, 2015). Simulations mimicking oxidative folding of hen lysozyme show that, in all folding pathways, the formation of disulfide bonds has a specific sequential order, and it is guided by protein folding (Muttathukattil *et al.*, 2019). The order of disulfide bond formation in hen lysozyme is Cys64-Cys80 (present in the β -domain), Cys76-Cys94 (located at the interface of the α - and β -domains), Cys30-Cys115, and Cys6-Cys127 (located in the α -domain) (Muttathukattil *et al.*, 2019). The first two disulfide bonds are short-ranged native contacts and are formed during the initial protein collapse, whereas the last two are long-ranged native contacts formed in the last stages of folding. The number of amino acids between the cysteine residues forming these bonds along the contour of the polypeptide chain is 15, 17, 84 and 120, respectively (Artymiuk *et al.*, 1982). The rate of disulfide bond formation depends on the protein contour loop length between the Cys residues forming the bond, and it follows the proximity rule proposed by (Camacho and Thirumalai, 1995).

If any of the disulfide bonds forms before its precursor in the order, then the protein does not fold to its native state. In the simulations of folding pathways, it was not observed any non-native disulfide bond formation, and this shows that protein folding guides disulfide bond formation in agreement with previous work (Qin *et al.*, 2015). Furthermore, ostrich-egg lysozyme (a g-type lysozyme, containing two disulfide bonds) is considered to require no disulfide bonds for folding and function (Kawamura *et al.*, 2008). The three α -helices ($\alpha 5$, $\alpha 7$ and $\alpha 8$), which are a structurally invariant core, may pack together in the early stage of the folding process and act as nucleation sites around which the structure can be formed. The disulfide bonds may confer stability after the protein reaches its final folded form but are not essential to protein folding and function (Kawamura *et al.*, 2008).

Taken as a whole, all these findings might help to describe the observed behavior of TmLyzc and TmLyzg during the *in vitro* refolding process in the light of misfolded or partially folded kinetic intermediates. The events involved in the folding process of hen lysozyme can resemble those that occurred in the refolding process of TmLyzc and TmLyzg, as they are homologous counterparts. In this sense, although protein folding seems to be still like a ‘black box’, the advances in the knowledge of the folding of small proteins like lysozyme can help to depict a general picture of the events and pathways involved in this process. In this light, the art of *in vitro* protein refolding resides in finding the precise experimental conditions that guide folding through the fast folding pathways. That is, to force the equilibrium of kinetic partitioning to folding intermediates that follow the fast folding trajectories that lead to successful folding.

CONCLUSIONS

As the second part of this research project, the recombinant g-type (TmLyzg) and c-type (TmLyzc) lysozymes of *Totoaba macdonaldi* were recombinantly overexpressed in *E. coli*, partially purified by IMAC chromatography and subjected to refolding experiments. Both lysozymes were expressed in the form of inclusion bodies, as it has also been observed for other c-type and g-type lysozymes from diverse organisms when using *E. coli* as the expression system. Furthermore, TmLyzg and TmLyzc were assayed under several refolding conditions under reducing (TmLyzg, no disulfide bonds) and non-reducing (TmLyzc, four disulfide bonds) conditions, according to the predicted secondary/tertiary structures of each lysozyme. Both lysozymes were soluble under some refolding conditions and thus, presumably refolded. Refolding yields could not be measured in terms of lysozyme activity, suggesting that TmLyzg and TmLyzc could be partially refolded, or their active sites were somehow hindered, and thus were inactive. However, even though some of the assayed refolding conditions seem promising, further optimization is required to obtain soluble-active lysozymes for biochemical and structural studies, and to investigate their broad-spectrum antibacterial activity.

The folding pathways that a protein can follow are complex. Studies with hen-egg lysozyme revealed that this protein presents two folding domains (α - and β -domains) that cooperate to configure the binding cleft, which is the interface between these domains. In addition, simulation studies also support a triangular folding mechanism of hen-lysozyme, in which a trapped kinetic intermediate can be found between the unfolded and native states. Moreover, lysozyme can follow fast and slow folding trajectories, each with considerable energy barriers to overcome to reach the protein's native fold. In this sense, if TmLyzc and TmLyzg were kinetically trapped in a misfolded or partially folded intermediate that prevented them achieving their native and active fold, further biophysical studies to monitor folding kinetics or protein stability should be conducted.

Despite the fact that lysozyme is a well-known model for protein folding and structure-activity relationships, the most represented studies come from human and avian sources. In the future, studies like those on lysozyme from fish species will expand the knowledge in protein structure and function that may impact aquaculture systems.

RECOMMENDATIONS

To follow up this research, I recommend:

- Remove fusion tags (His-tag and Trx-tag) from TmLyzg and TmLyzc by proteolytic digestion with PreScission protease, in order to re-assess lysozyme activity before and after refolding assays.
- Use a second chromatographic step to purify recombinant TmLyzg and TmLyzc using a substrate-affinity matrix of immobilized *M. luteus*, to ensure that only properly folded lysozymes are obtained to be used in further biochemical and structural studies.
- Use of biophysical techniques, such as Differential Scanning Calorimetry (DSC) and Circular Dichroism (CD), to monitor the overall structure of TmLyzg and TmLyzc during and after refolding.
- Use molecular exclusion chromatography to evaluate the oligomeric state of TmLyzg and TmLyzc during and after refolding.
- Use molecular dynamics simulation based on high precision molecular 3D models to predict folding trajectories and mechanisms of TmLyzg and TmLyzc.

REFERENCES

- Acharya, K. R., Stuart, D. I., Walker, N. P. C., Lewis, M. and Phillips, D. C. (1989). Refined structure of baboon α -lactalbumin at 1.7 Å resolution: Comparison with C-type lysozyme. *J. Mol. Biol.* 208(1): 99–127.
- Andrews, B. T., Gosavi, S., Finke, J. M., Onuchic, J. N. and Jennings, P. A. (2008). The dual-basin landscape in GFP folding. *Proc. Natl. Acad. Sci.* 105(34): 12283–12288.
- Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science.* 181(4096): 223–230.
- Ángeles Esteban, M. (2012). An Overview of the Immunological Defenses in Fish Skin. *ISRN Immunol.* Edited by A. Bensussan, P. Puccetti, E. Flaño, and J. D. Hayball. 2012: 853470.
- Angulo, C., Sanchez, V., Delgado, K. and Reyes-Becerril, M. (2019). C-type lectin 17A and macrophage-expressed receptor genes are magnified by fungal β -glucan after *Vibrio parahaemolyticus* infection in *Totoaba macdonaldi* cells. *Immunobiology.* 224(1): 102–109.
- Ao, J., Mu, Y., Xiang, L.-X., Fan, D., Feng, M., *et al.* (2015). Genome sequencing of the perciform fish *Larimichthys crocea* provides insights into molecular and genetic mechanisms of stress adaptation. *PLOS Genet.* 11(4): e1005118.
- Artymiuk, P. J., Blake, C. C. F., Rice, D. W. and Wilson, K. S. (1982). The structures of the monoclinic and orthorhombic forms of hen egg-white lysozyme at 6 Å resolution. *Acta Crystallogr. Sect. B.* 38(3): 778–783.
- Bassity, E. and Clark, T. G. (2012). Functional identification of dendritic cells in the teleost model, rainbow trout (*Oncorhynchus mykiss*). *PLoS One.* 7(3): e33196.
- Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. and Sarma, V. R. (1965). Structure of hen egg-white lysozyme: A three-dimensional fourier synthesis at 2 Å resolution. *Nature.* 206(4986): 757–761.
- Bobadilla, M., Alvarez-Borrego, S., Avila-Foucat, S., Lara-Valencia, F. and Espejel, I. (2011). Evolution of environmental policy instruments implemented for the protection of totoaba and the vaquita porpoise in the Upper Gulf of California. *Environ. Sci. Policy.* 14(8): 998–1007.

- Bonvin, A. M. J. J., Vis, H., Breg, J. N., Burgering, M. J. M., Boelens, R. and Kaptein, R. (1994). Nuclear magnetic resonance solution structure of the Arc repressor using relaxation matrix calculations. *J. Mol. Biol.* 236(1): 328–341.
- Briesemeister, S., Rahnenführer, J. and Kohlbacher, O. (2010). YLoc—an interpretable web server for predicting subcellular localization. *Nucleic Acids Res.* 38(suppl_2): W497–W502.
- Bruce, T. J. and Brown, M. L. (2017). A review of immune system components, cytokines, and immunostimulants in cultured finfish species. *Open J. Anim. Sci.* 7(03): 267.
- Buchmann, K. (2014). Evolution of innate immunity: Clues from invertebrates via fish to mammals. *Front. Immunol.* 459.
- Buonocore, F., Randelli, E., Trisolino, P., Facchiano, A., de Pascale, D. and Scapigliati, G. (2014). Molecular characterization, gene structure and antibacterial activity of a g-type lysozyme from the European sea bass (*Dicentrarchus labrax* L.). *Mol. Immunol.* 62(1): 10–18.
- Cabanillas-Gámez, M., Bardullas, U., Galaviz, M. A., Rodriguez, S., Rodriguez, V. M. and López, L. M. (2020). Tryptophan supplementation helps totoaba (*Totoaba macdonaldi*) juveniles to regain homeostasis in high-density culture conditions. *Fish Physiol. Biochem.* 46(2): 597–611.
- Cai, S., Zhang, Y., Wu, F., Wu, R., Yang, S., Li, Y. and Xu, Y. (2019). Identification and functional characterization of a c-type lysozyme from *Fenneropenaeus penicillatus*. *Fish Shellfish Immunol.* 88: 161–169.
- Callewaert, L. and Michiels, C. W. (2010). Lysozymes in the animal kingdom. *J. Biosci.* 35(1): 127–160.
- Camacho, C. J. and Thirumalai, D. (1995). Theoretical predictions of folding pathways by using the proximity rule, with applications to bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci.* 92(5): 1277–1281.
- Canfield, R. E. (1963). The amino acid sequence of egg white lysozyme. *J. Biol. Chem.* 238(8): 2698–2707.
- Časaitė, V., Bružytė, S., Bukauskas, V., Šetkus, A., Morozova-Roche, L. A. and Meškys, R. (2009). Expression and purification of active recombinant equine lysozyme in *Escherichia coli*. *Protein Eng. Des. Sel.* 22(11): 649–654.

- Chaffotte, A. F., Guillou, Y. and Goldberg, M. E. (1992). Kinetic resolution of peptide bond and side chain far-UV circular dichroism during the folding of hen egg white lysozyme. *Biochemistry*. 31(40): 9694–9702.
- Chaudhuri, T. K., Horii, K., Yoda, T., Arai, M., Nagata, S., *et al.* (1999). Effect of the extra N-terminal methionine residue on the stability and folding of recombinant α -Lactalbumin expressed in *Escherichia coli*. *J. Mol. Biol.* 285(3): 1179–1194.
- Chrysina, E. D., Brew, K. and Acharya, K. R. (2000). Crystal Structures of Apo- and Holo-bovine α -Lactalbumin at 2.2 Å resolution reveal an effect of calcium on inter-lobe interactions. *J. Biol. Chem.* 275(47): 37021–37029.
- Cisneros-Mata, M. A., Botsford, L. W. and Quinn, J. F. (1997). Projecting viability of *Totoaba macdonaldi*, a population with unknown age-dependent variability. *Ecol. Appl.* 7(3): 968–980.
- Costa, S., Almeida, A., Castro, A. and Domingues, L. (2014). Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system . *Front. Microbiol.* 63.
- Dahlquist, F. W., Rand-Meir, T. and Raftery, M. A. (1969). Application of secondary α -deuterium kinetic isotope effects to studies of enzyme catalysis: Glycoside hydrolysis by lysozyme and β -glucosidase. *Biochemistry*. 8(10): 4214–4221.
- Dautigny, A., Prager, E. M., Pham-Dinh, D., Jollès, J., Pakdel, F., Grinde, B. and Jollès, P. (1991). cDNA and amino acid sequences of rainbow trout (*Oncorhynchus mykiss*) lysozymes and their implications for the evolution of lysozyme and lactalbumin. *J. Mol. Evol.* 32(2): 187–198.
- de-la-Re-Vega, E., García-Orozco, K. D., Calderón-Arredondo, S. A., Romo-Figueroa, M. G., Islas-Osuna, M. A., Yepiz-Plascencia, G. M. and Sotelo-Mundo, R. R. (2004). Recombinant expression of marine shrimp lysozyme in *Escherichia coli*. *Electron. J. Biotechnol.* 12–13.
- Denton, M. E., Rothwarf, D. M. and Scheraga, H. A. (1994). Kinetics of folding of guanidine-denatured hen egg white lysozyme and carboxymethyl (Cys6,Cys127)-lysozyme: A stopped-flow absorbance and fluorescence study. *Biochemistry*. 33(37): 11225–11236.
- Dill, K. A. and Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4(1): 10–19.

- Dill, K. A., Fiebig, K. M. and Chan, H. S. (1993). Cooperativity in protein-folding kinetics. *Proc. Natl. Acad. Sci.* 90(5): 1942–1946.
- Dill, K. A. and MacCallum, J. L. (2012). The protein-folding problem, 50 years on. *Science*. 338(6110): 1042–1046.
- Dill, K. A., Ozkan, S. B., Shell, M. S. and Weikl, T. R. (2008). The protein folding problem. *Annu. Rev. Biophys.* 37(1): 289–316.
- Dobson, C. M., Evans, P. A. and Radford, S. E. (1994). Understanding how proteins fold: the lysozyme story so far. *Trends Biochem. Sci.* 19(1): 31–37.
- Dobson, C. M., Šali, A. and Karplus, M. (1998). Protein folding: A perspective from theory and experiment. *Angew. Chemie Int. Ed.* 37(7): 868–893.
- Dumon-Seignovert, L., Cariot, G. and Vuillard, L. (2004). The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr. Purif.* 37(1): 203–206.
- Düring, K., Porsch, P., Mahn, A., Brinkmann, O. and Gieffers, W. (1999). The non-enzymatic microbicidal activity of lysozymes. *FEBS Lett.* 449(2–3): 93–100.
- Engelman, D. M. and Steitz, T. A. (1981). The spontaneous insertion of proteins into and across membranes: The helical hairpin hypothesis. *Cell*. 23(2): 411–422.
- Fast, M. D., Sims, D. E., Burka, J. F., Mustafa, A. and Ross, N. W. (2002). Skin morphology and humoral non-specific defence parameters of mucus and plasma in rainbow trout, coho and Atlantic salmon. *Comp. Biochem. Physiol. Part A Mol. Integr. Physiol.* 132(3): 645–657.
- Ferrer-Miralles, N. and Villaverde, A. (2013). Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microb. Cell Fact.* 12(1): 113.
- Fevolden, S.-E., Røed, K. H. and Fjalestad, K. T. (2002). Selection response of cortisol and lysozyme in rainbow trout and correlation to growth. *Aquaculture*. 205(1): 61–75.
- Findley, L. (2010). *Totoaba macdonaldi*. *The IUCN Red List of Threatened Species 2010: e.T22003A9346099*. *Totoaba macdonaldi*. *IUCN Red List Threat. Species 2010 e.T22003A9346099*. Available at: <https://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T22003A9346099.en>. (Accessed: 29 January 2021).

- Fischer, B., Perry, B., Phillips, G., Sumner, I. and Goodenough, P. (1993). Physiological consequence of expression of soluble and active hen egg white lysozyme in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 39(4): 537–540.
- Fischer, U., Koppang, E. O. and Nakanishi, T. (2013). Teleost T and NK cell immunity. *Fish Shellfish Immunol.* 35(2): 197–206.
- Flajnik, M. F. and Kasahara, M. (2010). Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat. Rev. Genet.* 11(1): 47–59.
- Fleming, A. (1932). Lysozyme. *Proc. R. Soc. Med.* 26(2): 71–84.
- Frare, E., Mossuto, M. F., de Laureto, P. P., Tolin, S., Menzer, L., Dumoulin, M., Dobson, C. M. and Fontana, A. (2009). Characterization of oligomeric species on the aggregation pathway of human lysozyme. *J. Mol. Biol.* 387(1): 17–27.
- Fu, G. H., Bai, Z. Y., Xia, J. H., Liu, F., Liu, P. and Yue, G. H. (2013). Analysis of two lysozyme genes and antimicrobial functions of their recombinant proteins in Asian seabass. *PLoS One.* 8(11): e79743.
- Galaviz, M. A., López, L. M., García Gasca, A., Álvarez González, C. A., True, C. D. and Gisbert, E. (2015). Digestive system development and study of acid and alkaline protease digestive capacities using biochemical and molecular approaches in totoaba (*Totoaba macdonaldi*) larvae. *Fish Physiol. Biochem.* 41(5): 1117–1130.
- Gao, C., Fu, Q., Zhou, S., Song, L., Ren, Y., Dong, X., Su, B. and Li, C. (2016). The mucosal expression signatures of g-type lysozyme in turbot (*Scophthalmus maximus*) following bacterial challenge. *Fish Shellfish Immunol.* 54: 612–619.
- Gao, F., Qu, L., Yu, S., Ye, X., Tian, Y., Zhang, L., Bai, J. and Lu, M. (2012). Identification and expression analysis of three c-type lysozymes in *Oreochromis aureus*. *Fish Shellfish Immunol.* 32(5): 779–788.
- García-Hernández, E., Zubillaga, R. A., Chavelas-Adame, E. A., Vázquez-Contreras, E., Rojo-Domínguez, A. and Costas, M. (2003). Structural energetics of protein–carbohydrate interactions: Insights derived from the study of lysozyme binding to its natural saccharide inhibitors. *Protein Sci.* 12(1): 135–142.
- García-Orozco, K. D., López-Zavala, A. A., Puentes-Camacho, D., Calderón-de-la-Barca, A. M.

- and Sotelo-Mundo, R. R. (2005). Recombinant bacterial expression of the lysozyme from the tobacco-hornworm *Manduca sexta* with activity at low temperatures. *Biotechnol. Lett.* 27(15): 1075–1080.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. and Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server BT - The Proteomics Protocols Handbook. in Walker, J. M. (ed.). Totowa, NJ: Humana Press. 571–607.
- Gladwin, S. T. and Evans, P. A. (1996). Structure of very early protein folding intermediates: new insights through a variant of hydrogen exchange labelling. *Fold. Des.* 1(6): 407–417.
- Goda, S., Takano, K., Yutani, K., Yamagata, Y., Nagata, R., Akutsu, H., Maki, S. and Namba, K. (2000). Amyloid protofilament formation of hen egg lysozyme in highly concentrated ethanol solution. *Protein Sci.* 9(2): 369–375.
- González-Félix, M. L., Gatlin, D. M., Perla Urquidez-Bejarano, de la Reé-Rodríguez, C., Duarte-Rodríguez, L., *et al.* (2018). Effects of commercial dietary prebiotic and probiotic supplements on growth, innate immune responses, and intestinal microbiota and histology of *Totoaba macdonaldi*. *Aquaculture.* 49:239–251.
- Goto, T., Abe, Y., Kakuta, Y., Takeshita, K., Imoto, T. and Ueda, T. (2007). Crystal Structure of *Tapes japonica* lysozyme with substrate analogue: Structural basis of the catalytic mechanism and manifestation of its chitinase activity accompanied by quaternary structural change. *J. Biol. Chem.* 282(37): 27459–27467.
- Granja, A. G., Leal, E., Pignatelli, J., Castro, R., Abós, B., Kato, G., Fischer, U. and Tafalla, C. (2015). Identification of teleost skin CD8 α ⁺ dendritic-like cells, representing a potential common ancestor for mammalian cross-presenting dendritic cells. *J. Immunol.* 195(4): 1825 LP – 1837.
- Grayfer, L., Kerimoglu, B., Yaparla, A., Hodgkinson, J. W., Xie, J. and Belosevic, M. (2018). Mechanisms of fish macrophage antimicrobial immunity. *Front. Immunol.* 1105.
- Green, M. R. and Sambrook, J. (2016). Preparation of plasmid DNA by alkaline lysis with sodium dodecyl sulfate: minipreps. *Cold Spring Harb. Protoc.* 2016(10): pdb.prot093344.
- Grinde, B. (1989). Lysozyme from rainbow trout, *Salmo gairdneri* Richardson, as an antibacterial agent against fish pathogens. *J. Fish Dis.* 12(2): 95–104.
- Grütter, M. G., Weaver, L. H. and Matthews, B. W. (1983). Goose lysozyme structure: an

- evolutionary link between hen and bacteriophage lysozymes? *Nature*. 303(5920): 828–831.
- Guevara, J. C. B. (1990). The conservation of *Totoaba macdonaldi* (Gilbert), (Pisces: *Sciaenidae*), in the Gulf of California, Mexico. *J. Fish Biol.* 37(sA): 201–202.
- Guez, V., Roux, P., Navon, A. and Goldberg, M. E. (2002). Role of individual disulfide bonds in hen lysozyme early folding steps. *Protein Sci.* 11(5): 1136–1151.
- Guo, Z. and Thirumalai, D. (1995). Kinetics of protein folding: Nucleation mechanism, time scales, and pathways. *Biopolymers*. 36(1): 83–102.
- Haugarvoll, E., Bjerås, I., Nowak, B. F., Hordvik, I. and Koppang, E. O. (2008). Identification and characterization of a novel intraepithelial lymphoid tissue in the gills of Atlantic salmon. *J. Anat.* 213(2): 202–209.
- Havixbeck, J. J. and Barreda, D. R. (2015). Neutrophil development, migration, and function in teleost fish. *Biology (Basel)*. 4(4): 715–734.
- Helland, R., Larsen, R. L., Finstad, S., Kyomuhendo, P. and Larsen, A. N. (2009). Crystal structures of g-type lysozyme from Atlantic cod shed new light on substrate binding and the catalytic mechanism. *Cell. Mol. Life Sci.* 66(15): 2585–2598.
- Van Herreweghe, J. M. and Michiels, C. W. (2012). Invertebrate lysozymes: Diversity and distribution, molecular mechanism and in vivo function. *J. Biosci.* 37(2): 327–348.
- Hikima, J., Hirono, I. and Aoki, T. (2003). The lysozyme gene in fish - Aquatic genomics: Steps toward a great future. In: Shimizu, N., Aoki, T., Hirono, I., and Takashima, F. (eds). Tokyo: Springer Japan. 301–309.
- Hikima, J., Minagawa, S., Hirono, I. and Aoki, T. (2001). Molecular cloning, expression and evolution of the Japanese flounder goose-type lysozyme gene, and the lytic activity of its recombinant protein. *Biochim. Biophys. Acta - Gene Struct. Expr.* 1520(1): 35–44.
- Hirakawa, H., Kawahara, Y., Ochi, A., Muta, S., Kawamura, S., Torikata, T. and Kuhara, S. (2006). Construction of enzyme-substrate complexes between hen egg-white lysozyme and N-acetyl-D-glucosamine hexamer by systematic conformational search and molecular dynamics simulation. *J. Biochem.* 140(2): 221–227.
- Hirakawa, H., Ochi, A., Kawahara, Y., Kawamura, S., Torikata, T. and Kuhara, S. (2008).

Catalytic reaction mechanism of goose egg-white lysozyme by molecular modelling of enzyme–substrate complex. *J. Biochem.* 144(6): 753–761.

Hodgkinson, J. W., Grayfer, L. and Belosevic, M. (2015). Biology of bony fish macrophages. *Biology (Basel)*. 4(4): 881–906.

Holm, L. and Sander, C. (1994). Structural similarity of plant chitinase and lysozymes from animals and phage. *FEBS Lett.* 340(1–2): 129–132.

Honda, Y. and Fukamizo, T. (1998). Substrate binding subsites of chitinase from barley seeds and lysozyme from goose egg white. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* 1388(1): 53–65.

Ibrahim, H. R. (1998). On the novel catalytically-independent antimicrobial function of hen egg-white lysozyme: A conformation-dependent activity. *Food/Nahrung.* 42(03-04): 187–193.

Ibrahim, H. R., Higashiguchi, S., Juneja, L. R., Kim, M. and Yamamoto, T. (1996). A structural phase of heat-denatured lysozyme with novel antimicrobial action. *J. Agric. Food Chem.* 44(6): 1416–1423.

Ibrahim, H. R., Imazato, K. and Ono, H. (2011). Human lysozyme possesses novel antimicrobial peptides within its N-terminal domain that target bacterial respiration. *J. Agric. Food Chem.* 59(18): 10336–10345.

Ibrahim, H. R., Inazaki, D., Abdou, A., Aoki, T. and Kim, M. (2005). Processing of lysozyme at distinct loops by pepsin: A novel action for generating multiple antimicrobial peptide motifs in the newborn stomach. *Biochim. Biophys. Acta - Gen. Subj.* 1726(1): 102–114.

Ibrahim, H. R., Matsuzaki, T. and Aoki, T. (2001). Genetic evidence that antibacterial activity of lysozyme is independent of its catalytic function. *FEBS Lett.* 506(1): 27–32.

Ibrahim, H. R., Thomas, U. and Pellegrini, A. (2001). A helix-loop-helix peptide at the upper lip of the active site cleft of lysozyme confers potent antimicrobial activity with membrane permeabilization action. *J. Biol. Chem.* 276(47): 43767–43774.

Irwin, D. M. (2014). Evolution of the vertebrate goose-type lysozyme gene family. *BMC Evol. Biol.* 14(1): 188.

Irwin, D. M., Biegel, J. M. and Stewart, C.-B. (2011). Evolution of the mammalian lysozyme gene

family. *BMC Evol. Biol.* 11(1): 166.

Irwin, D. M. and Gong, Z. M. (2003). Molecular evolution of vertebrate goose-type lysozyme genes. *J. Mol. Evol.* 56(2): 234–242.

Ishikawa, N., Chiba, T., Chen, L. T., Shimizu, A., Ikeguchi, M. and Sugai, S. (1998). Remarkable destabilization of recombinant alpha-lactalbumin by an extraneous N-terminal methionyl residue. *Protein Eng. Des. Sel.* 11(5): 333–335.

Itzhaki, L. S., Evans, P. A., Dobson, C. M. and Radford, S. E. (1994). Tertiary interactions in the folding pathway of hen lysozyme: kinetic studies using fluorescent probes. *Biochemistry.* 33(17): 5212–5220.

Iwasaki, A. and Medzhitov, R. (2010). Regulation of adaptive immunity by the innate immune system. *Science.* 327(5963): 291–295.

Jiménez-Cantizano, R. M., Infante, C., Martín-Antonio, B., Ponce, M., Hachero, I., Navas, J. I. and Manchado, M. (2008). Molecular characterization, phylogeny, and expression of c-type and g-type lysozymes in brill (*Scophthalmus rhombus*). *Fish Shellfish Immunol.* 25(1): 57–65.

Juarez, L. M., Konietzko, P. A. and Schwarz, M. H. (2016). Totoaba aquaculture and conservation: Hope for an endangered fish from Mexico's Sea of Cortez. *World Aquac.* 47(4): 30–38.

Judy, E. and Kishore, N. (2019). A look back at the molten globule state of proteins: thermodynamic aspects. *Biophys. Rev.* 11(3): 365–375.

Karlsen, S., Eliassen, B. E., Hansen, L. K., Larsen, R. L., Riise, B. W., Smalås, A. O., Hough, E. and Grinde, B. (1995). Refined crystal structure of lysozyme from the rainbow trout (*Oncorhynchus mykiss*). *Acta Crystallogr. Sect. D Biol. Crystallogr.* 51(3): 354–367.

Karplus, M. (1997). The Levinthal paradox: yesterday and today. *Fold. Des.* 2: S69–S75.

Katzenback, B. A. (2015). Antimicrobial peptides as mediators of innate immunity in teleosts. *Biology.* 4(4):607-639.

Kawamura, S., Ohkuma, M., Chijiwa, Y., Kohno, D., Nakagawa, H., Hirakawa, H., Kuhara, S. and Torikata, T. (2008). Role of disulfide bonds in goose-type lysozyme. *FEBS J.* 275(11): 2818–2830.

Kazmirski, S. L. and Daggett, V. (1998). Non-native interactions in protein folding intermediates:

- molecular dynamics simulations of hen lysozyme¹ Edited by B. Honig. *J. Mol. Biol.* 284(3): 793–806.
- Kiefhaber, T. (1995). Kinetic traps in lysozyme folding. *Proc. Natl. Acad. Sci.* 92(20): 9029–9033.
- Kim, Y. K. and Nam, Y. K. (2015). Molecular characterization and expression pattern of c-type and g-type lysozyme isoforms in starry flounder *Platichthys stellate* infected with *Streptococcus parauberis*. *Fish. Sci.* 81(2): 353–363.
- Kirby, A. J. (2001). The lysozyme mechanism sorted — after 50 years. *Nat. Struct. Biol.* 8(9): 737–739.
- Kiron, V. (2012). Fish immune system and its nutritional modulation for preventive health care. *Anim. Feed Sci. Technol.* 173(1): 111–133.
- Koshiha, T., Hayashi, T., Miwako, I., Kumagai, I., Ikura, T., Kawano, K., Nitta, K. and Kuwajima, K. (1999). Expression of a synthetic gene encoding canine milk lysozyme in *Escherichia coli* and characterization of the expressed protein. *Protein Eng. Des. Sel.* 12(5): 429–435.
- Kotik, M., Radford, S. E. and Dobson, C. M. (1995). Comparison of the Refolding Hen Lysozyme from Dimethyl Sulfoxide and Guanidinium Chloride. *Biochemistry.* 34(5): 1714–1724.
- Kumar, S., Ravi, V. K. and Swaminathan, R. (2009). Suppression of lysozyme aggregation at alkaline pH by tri-N-acetylchitotriose. *Biochim. Biophys. Acta - Proteins Proteomics.* 1794(6): 913–920.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35(6): 1547–1549.
- Kuroki, R., Weaver, L. H. and Matthews, B. W. (1999). Structural basis of the conversion of T4 lysozyme into a transglycosidase by reengineering the active site. *Proc. Natl. Acad. Sci.* 96(16): 8949 – 8954.
- Kyomuhendo, P., Myrnes, B. and Nilsen, I. W. (2007). A cold-active salmon goose-type lysozyme with high heat tolerance. *Cell. Mol. Life Sci.* 64(21): 2841–2847.
- Lamppa, J. W., Tanyos, S. A. and Griswold, K. E. (2013). Engineering *Escherichia coli* for soluble expression and single step purification of active human lysozyme. *J. Biotechnol.* 164(1): 1–8.

- Landeta, C., Boyd, D. and Beckwith, J. (2018). Disulfide bond formation in prokaryotes. *Nat. Microbiol.* 3(3): 270–280.
- Larsen, A. N., Solstad, T., Svineng, G., Seppola, M. and Jørgensen, T. Ø. (2009). Molecular characterisation of a goose-type lysozyme gene in Atlantic cod (*Gadus morhua* L.). *Fish Shellfish Immunol.* 26(1): 122–132.
- Lazo, J. P., Fuentes-Quesada, J. P., Villareal-Rodarte, G., Viana, M. T. and Baron-Sevilla, B. (2020). The effect of dietary n-3 LC-PUFA levels on growth, survival, and feed utilization in juvenile *Totoaba macdonaldi*. *Aquaculture.* 525: 735350.
- Lercari, D. and Chávez, E. A. (2007). Possible causes related to historic stock depletion of the totoaba, *Totoaba macdonaldi* (Perciformes: *Sciaenidae*), endemic to the Gulf of California. *Fish. Res.* 86(2): 136–142.
- Leysen, S., Vanderkelen, L., Weeks, S. D., Michiels, C. W. and Strelkov, S. V. (2013). Structural basis of bacterial defense against g-type lysozyme-based innate immunity. *Cell. Mol. Life Sci.* 70(6): 1113–1122.
- Li, L., Cardoso, J. C. R., Félix, R. C., Mateus, A. P., Canário, A. V. M. and Power, D. M. (2021). Fish lysozyme gene family evolution and divergent function in early development. *Dev. Comp. Immunol.* 114: 103772.
- Li, S., Wang, D., Liu, H., Yin, J. and Lu, T. (2016). Expression and antimicrobial activity of c-type lysozyme in taimen (*Hucho taimen*, Pallas). *Dev. Comp. Immunol.* 63: 156–162.
- Li, Yajuan, Li, Yuelong, Cao, X., Jin, X. and Jin, T. (2017). Pattern recognition receptors in zebrafish provide functional and evolutionary insight into innate immune signaling pathways. *Cell. Mol. Immunol.* 14(1): 80–89.
- Lieschke, G. J. and Trede, N. S. (2009). Fish immunology. *Curr. Biol.* 19(16): R678–R682.
- Liu, Z., Reddy, G., O'Brien, E. P. and Thirumalai, D. (2011). Collapse kinetics and chevron plots from simulations of denaturant-dependent folding of globular proteins. *Proc. Natl. Acad. Sci.* 108(19): 7787–7792.
- Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative pcr and the $2^{-\Delta\Delta CT}$ method. *Methods.* 25(4): 402–408.

- Løkka, G., Austbø, L., Falk, K., Bromage, E., Fjellidal, P. G., Hansen, T., Hordvik, I. and Koppang, E. O. (2014). Immune parameters in the intestine of wild and reared unvaccinated and vaccinated Atlantic salmon (*Salmo salar* L.). *Dev. Comp. Immunol.* 47(1): 6–16.
- López, L. M., Flores-Ibarra, M., Bañuelos-Vargas, I., Galaviz, M. A. and True, C. D. (2015). Effect of fishmeal replacement by soy protein concentrate with taurine supplementation on growth performance, hematological and biochemical status, and liver histology of totoaba juveniles (*Totoaba macdonaldi*). *Fish Physiol. Biochem.* 41(4): 921–936.
- Magnadóttir, B. (2010). Immunological Control of Fish Diseases. *Mar. Biotechnol.* 12(4): 361–379.
- Magnadóttir, B. (2006). Innate immunity of fish (overview). *Fish Shellfish Immunol.* 20(2): 137–151.
- Mai, W., Liu, P., Chen, H. and Zhou, Y. (2014). Cloning and immune characterization of the c-type lysozyme gene in red-spotted grouper, *Epinephelus akaara*. *Fish Shellfish Immunol.* 36(1): 305–314.
- Makabe, K., Nakamura, T. and Kuwajima, K. (2013). Structural insights into the stability perturbations induced by N-terminal variation in human and goat α -lactalbumin. *Protein Eng. Des. Sel.* 26(2): 165–170.
- Mallamace, F., Corsaro, C., Mallamace, D., Vasi, S., Vasi, C., Baglioni, P., Buldyrev, S. V, Chen, S.-H. and Stanley, H. E. (2016). Energy landscape in protein folding and unfolding. *Proc. Natl. Acad. Sci.* 113(12): 3159 LP – 3163.
- Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., *et al.* (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* 45(D1): D200–D203.
- Mata-Sotres, J. A., P.Lazo, J. and Baron-Sevilla, B. (2015). Effect of age on weaning success in totoaba (*Totoaba macdonaldi*) larval culture. *Aquaculture.* 437: 292–296.
- Mata-Sotres, J. A., Tinajero-Chavez, A., Barreto-Curiel, F., Pares-Sierra, G., Del Rio-Zaragoza, O. B., Viana, M. T. and Rombenso, A. N. (2018). DHA (22:6n-3) supplementation is valuable in *Totoaba macdonaldi* fish oil-free feeds containing poultry by-product meal and beef tallow. *Aquaculture.* 497: 440–451.

- Matagne, A., Chung, E. W., Ball, L. J., Radford, S. E., Robinson, C. V and Dobson, C. M. (1998). The origin of the α -domain intermediate in the folding of hen lysozyme. *J. Mol. Biol.* 277(5): 997–1005.
- Matagne, A. and Dobson, C. M. (1998). The folding process of hen lysozyme: a perspective from the ‘new view’. *Cell. Mol. Life Sci. C.* 54(4): 363–371.
- Matagne, A., Radford, S. E. and Dobson, C. M. (1997). Fast and slow tracks in lysozyme folding: insight into the role of domains in the folding process. *J. Mol. Biol.* 267(5): 1068–1074.
- Mckenzie, H. A. and White, F. H. (1991). Lysozyme and α -Lactalbumin: Structure, function, and interrelationships. In: Anfinsen, C. B., Richards, F. M., Edsall, J. T., and Eisenberg, D. S. B. T.-A. in P. C. (eds). Academic Press. 173–315.
- Melkikh, A. V and Meijer, D. K. F. (2018). On a generalized Levinthal’s paradox: The role of long- and short range interactions in complex bio-molecular reactions, including protein and DNA folding. *Prog. Biophys. Mol. Biol.* 132: 57–79.
- Mellman, I. and Steinman, R. M. (2001). Dendritic cells: specialized and regulated antigen processing machines. *Cell.* 106(3): 255–258.
- Milla, M. E. and Sauer, R. T. (1994). p22 Arc repressor: Folding kinetics of a single-domain, dimeric protein. *Biochemistry.* 33(5): 1125–1133.
- Minagawa, S., Hikima, J., Hirono, I., Aoki, T. and Mori, H. (2001). Expression of Japanese flounder c-type lysozyme cDNA in insect cells. *Dev. Comp. Immunol.* 25(5): 439–445.
- Mine, S., Ueda, T., Hashimoto, Y. and Imoto, T. (1997). Improvement of the refolding yield and solubility of hen egg-white lysozyme by altering the Met residue attached to its N-terminus to Ser. *Protein Eng. Des. Sel.* 10(11): 1333–1338.
- Miranker, A., Radford, S. E., Karplus, M. and Dobson, C. M. (1991). Demonstration by NMR of folding domains in lysozyme. *Nature.* 349(6310): 633–636.
- Mogensen, T. H. (2009). Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin. Microbiol. Rev.* 22(2): 240 LP – 273.
- Mohapatra, A., Parida, S., Mohanty, J. and Sahoo, P. K. (2019). Identification and functional characterization of a g-type lysozyme gene of *Labeo rohita*, an Indian major carp species. *Dev.*

Comp. Immunol. 92: 87–98.

Monzingo, A. F., Marcotte, E. M., Hart, P. J. and Robertas, J. D. (1996). Chitinases, chitosanases, and lysozymes can be divided into procaryotic and eucaryotic families sharing a conserved core. *Nat. Struct. Biol.* 3(2): 133–140.

Moreno-Córdova, E. N., Islas-Osuna, M. A., Contreras-Vergara, C. A., López-Zavala, A. A., Ruiz-Bustos, E., Reséndiz-Sandoval, M. G., Castillo-Yañez, F. J., Criscitiello, M. F. and Arvizu-Flores, A. A. (2020). Molecular characterization and expression analysis of the chicken-type and goose-type lysozymes from totoaba (*Totoaba macdonaldi*). *Dev. Comp. Immunol.* 113: 103807.

Motoshima, H., Ueda, T., Masumoto, K., Hashimoto, Y., Chijiwa, Y. and Imoto, T. (1997). Influence of mutations of the N-cap residue, Gly4, on stability and structure of hen lysozyme. *J. Biochem.* 122(1): 25–31.

Muttathukattil, A. N., Singh, P. C. and Reddy, G. (2019). Role of disulfide bonds and topological frustration in the kinetic partitioning of lysozyme folding pathways. *J. Phys. Chem. B.* 123(15): 3232–3241.

Nakanishi, T., Hikima, J. and Yada, T. (2018). Osteichthyes: Immune Systems of Teleosts (*Actinopterygii*) BT - Advances in Comparative Immunology. In: Cooper, E. L. (ed.). Cham: Springer International Publishing. 687–749.

Nakano, T. and Graf, T. (1991). Goose-type lysozyme gene of the chicken: sequence, genomic organization and expression reveals major differences to chicken-type lysozyme gene. *Biochim. Biophys. Acta - Gene Struct. Expr.* 1090(2): 273–276.

Nash, J. A., Ballard, T. N. S., Weaver, T. E. and Akinbi, H. T. (2006). The peptidoglycan-degrading property of lysozyme is not required for bactericidal activity in vivo. *J. Immunol.* 177(1): 519 LP – 526.

Netea, M. G., Schlitzer, A., Placek, K., Joosten, L. A. B. and Schultze, J. L. (2019). Innate and adaptive immune memory: an evolutionary continuum in the host's response to pathogens. *Cell Host Microbe.* 25(1): 13–26.

Neumann, N. F., Stafford, J. L., Barreda, D., Ainsworth, A. J. and Belosevic, M. (2001). Antimicrobial mechanisms of fish phagocytes and their role in host defense. *Dev. Comp. Immunol.* 25(8): 807–825.

- Nilojan, J., Bathige, S. D. N. K., Kugapreethan, R., Thulasitha, W. S., Nam, B.-H. and Lee, J. (2017). Molecular, transcriptional and functional insights into duplicated goose-type lysozymes from *Sebastes schlegelii* and their potential immunological role. *Fish Shellfish Immunol.* 67: 66–77.
- Nitta, K. and Sugai, S. (1989). The evolution of lysozyme and α -lactalbumin. *Eur. J. Biochem.* 182(1): 111–118.
- Noga, E. J., Ullal, A. J., Corrales, J. and Fernandes, J. M. O. (2011). Application of antimicrobial polypeptide host defenses to aquaculture: Exploitation of downregulation and upregulation responses. *Comp. Biochem. Physiol. Part D Genomics Proteomics.* 6(1): 44–54.
- Notredame, C., Higgins, D. G. and Heringa, J. (2000). T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302(1): 205–217.
- Ohmura, T., Ueda, T., Motoshima, H., Tamura, T. and Imoto, T. (1997). Analysis of the stability of mutant lysozymes at position 15 using X-ray crystallography. *J. Biochem.* 122(3): 512–517.
- Parham, P. (2003). Innate immunity: The unsung heroes. *Nature.* 423(6935): 20.
- Pei, J. and Grishin, N. V. (2005). The P5 protein from bacteriophage phi-6 is a distant homolog of lytic transglycosylases. *Protein Sci.* 14(5): 1370–1374.
- Petersen, T. N., Brunak, S., von Heijne, G. and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods.* 8(10): 785–786.
- Pils, B., Copley, R. R. and Schultz, J. (2005). Variation in structural location and amino acid conservation of functional sites in protein domain families. *BMC Bioinformatics.* 6(1): 210.
- Pooart, J., Torikata, T. and Araki, T. (2004). The primary structure of a novel goose-type lysozyme from rhea egg white. *Biosci. Biotechnol. Biochem.* 68(1): 159–169.
- Qasba, P. K., Kumar, S. and Brew, K. (1997). Molecular divergence of lysozymes and α -Lactalbumin. *Crit. Rev. Biochem. Mol. Biol.* 32(4): 255–306.
- Qin, M., Wang, W. and Thirumalai, D. (2015). Protein folding guides disulfide bond formation. *Proc. Natl. Acad. Sci.* 112(36): 11241 LP – 11246.
- Radford, S. E., Dobson, C. M., Dobson, C. M. and Fersht, A. R. (1995). Insights into protein folding using physical techniques: studies of lysozyme and α -lactalbumin. *Philos. Trans. R. Soc.*

London. Ser. B Biol. Sci. 348(1323): 17–25.

Radford, S. E., Dobson, C. M. and Evans, P. A. (1992). The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature*. 358(6384): 302–307.

Ragland, S. A. and Criss, A. K. (2017). From bacterial killing to immune modulation: Recent insights into the functions of lysozyme. *PLOS Pathog.* 13(9): e1006512.

Rebl, A. and Goldammer, T. (2018). Under control: The innate immunity of fish from the inhibitors' perspective. *Fish Shellfish Immunol.* 77: 328–349.

Reid, K. L., Rodriguez, H. M., Hillier, B. J. and Gregoret, L. M. (1998). Stability and folding properties of a model β -sheet protein, *Escherichia coli* CspA. *Protein Sci.* 7(2): 470–479.

Reyes-Becerril, M., Alamillo, E., Sánchez-Torres, L., Ascencio-Valle, F., Perez-Urbiola, J. C. and Angulo, C. (2016). Leukocyte susceptibility and immune response against *Vibrio parahaemolyticus* in *Totoaba macdonaldi*. *Dev. Comp. Immunol.* 65: 258–267.

Reyes-Becerril, M., Sanchez, V., Delgado, K., Guerra, K., Velazquez, E., Ascencio, F. and Angulo, C. (2018). Caspase -1, -3, -8 and antioxidant enzyme genes are key molecular effectors following *Vibrio parahaemolyticus* and *Aeromonas veronii* infection in fish leukocytes. *Immunobiology.* 223(10): 562–576.

Riera Romo, M., Pérez-Martínez, D. and Castillo Ferrer, C. (2016). Innate immunity in vertebrates: an overview. *Immunology.* 148(2): 125–139.

Rosano, G. L. and Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiology.* 172.

Rothwarf, D. M. and Scheraga, H. A. (1996). Role of non-native aromatic and hydrophobic interactions in the folding of hen egg white lysozyme. *Biochemistry.* 35(43): 13797–13807.

S̃ali, A., Shakhnovich, E. and Karplus, M. (1994). How does a protein fold? *Nature.* 369(6477): 248–251.

Sahoo, B. R. (2020). Structure of fish Toll-like receptors (TLR) and NOD-like receptors (NLR). *Int. J. Biol. Macromol.* 161: 1602–1617.

Salinas, I. (2015). The mucosal immune system of teleost fish. *Biology.* 4(3):525-39.

Satriyo, T. B., Galaviz, M. A., Salze, G. and López, L. M. (2017). Assessment of dietary taurine

- essentiality on the physiological state of juvenile *Totoaba macdonaldi*. *Aquac. Res.* 48(11): 5677–5689.
- Saurabh, S. and Sahoo, P. K. (2008). Lysozyme: an important defence molecule of fish innate immune system. *Aquac. Res.* 39(3): 223–239.
- Savan, R., Aman, A. and Sakai, M. (2003). Molecular cloning of G type lysozyme cDNA in common carp (*Cyprinus carpio* L.). *Fish Shellfish Immunol.* 15(3): 263–268.
- Savina, A. and Amigorena, S. (2007). Phagocytosis and antigen presentation in dendritic cells. *Immunol. Rev.* 219(1): 143–156.
- Schindelin, H., Jiang, W., Inouye, M. and Heinemann, U. (1994). Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. *Proc. Natl. Acad. Sci.* 91(11): 5119–5123.
- Schlörb, C., Ackermann, K., Richter, C., Wirmer, J. and Schwalbe, H. (2005). Heterologous expression of hen egg white lysozyme and resonance assignment of tryptophan side chains in its non-native states. *J. Biomol. NMR.* 33(2): 95–104.
- Segel, D. J., Bachmann, A., Hofrichter, J., Hodgson, K. O., Doniach, S. and Kiefhaber, T. (1999). Characterization of transient intermediates in lysozyme folding with time-resolved small-angle X-ray scattering. *J. Mol. Biol.* 288(3): 489–499.
- Semple, S. L. and Dixon, B. (2020). Salmonid antibacterial immunity: an aquaculture perspective. *Biology.* 9(10):331.
- Seppola, M., Bakkemo, K. R., Mikkelsen, H., Myrnes, B., Helland, R., Irwin, D. M. and Nilsen, I. W. (2016). Multiple specialised goose-type lysozymes potentially compensate for an exceptional lack of chicken-type lysozymes in Atlantic cod. *Sci. Rep.* 6(1): 28318.
- Shugar, D. (1952). The measurement of lysozyme activity and the ultra-violet inactivation of lysozyme. *Biochim. Biophys. Acta.* 8: 302–309.
- Singhvi, P., Saneja, A., Srichandan, S. and Panda, A. K. (2020). Bacterial inclusion bodies: a treasure trove of bioactive proteins. *Trends Biotechnol.* 38(5): 474–486.
- Smith, N. C., Rise, M. L. and Christian, S. L. (2019). A comparison of the innate and adaptive immune systems in cartilaginous fish, ray-finned fish, and lobe-finned fish. *Front. Immunol.* 10:2292.

- Soletto, I., Fischer, U., Tafalla, C. and Granja, A. G. (2018). Identification of a potential common ancestor for mammalian cross-presenting dendritic cells in teleost respiratory surfaces. *Front. Immunol.* 9:59.
- Srisailam, S., Arunkumar, A. I., Wang, W., Yu, C. and Chen, H. M. (2000). Conformational study of a custom antibacterial peptide cecropin B1: implications of the lytic activity. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* 1479(1): 275–285.
- Stecher, G., Tamura, K. and Kumar, S. (2020). Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Mol. Biol. Evol.* 37(4): 1237–1239.
- Sun, B. J., Wang, G. L., Xie, H. X., Gao, Q. and Nie, P. (2006). Gene structure of goose-type lysozyme in the mandarin fish *Siniperca chuatsi* with analysis on the lytic activity of its recombinant in *Escherichia coli*. *Aquaculture.* 252(2): 106–113.
- Sunyer, J. O. (2013). Fishing for mammalian paradigms in the teleost immune system. *Nat. Immunol.* 14(4): 320–326.
- Swan, I. D. A. (1972). The inhibition of hen egg-white lysozyme by imidazole and indole derivatives. *J. Mol. Biol.* 65(1): 59–62.
- Tafalla, C., Leal, E., Yamaguchi, T. and Fischer, U. (2016). T cell immunity in the teleost digestive tract. *Dev. Comp. Immunol.* 64: 167–177.
- Takano, K., Tsuchimori, K., Yamagata, Y. and Yutani, K. (1999). Effect of foreign N-terminal residues on the conformational stability of human lysozyme. *Eur. J. Biochem.* 266(2): 675–682.
- Thammasirirak, S., Pukcothanung, Y., Preecharram, S., Daduang, S., Patramanon, R., Fukamizo, T. and Araki, T. (2010). Antimicrobial peptides derived from goose egg white lysozyme. *Comp. Biochem. Physiol. Part C Toxicol. Pharmacol.* 151(1): 84–91.
- True, C. D., Loera, A. S. and Castro, N. C. (1997). Technical Notes: Acquisition of Broodstock of *Totoaba macdonaldi*: Field Handling, Decompression, and Prophylaxis of an Endangered Species. *Progress. Fish-Culturist.* 59(3): 246–248.
- True, C. D., Silva-Loera, A. and Castro-Castro, N. (1997). Acquisition of broodstock of *totoaba macdonaldi*: field handling, decompression, and prophylaxis of an endangered species. *Progress. Fish Cult.* 246–248.

- Tullio, V., Spaccapelo, R. and Polimeni, M. (2015). Lysozymes in the animal kingdom. in *Hum. Mosq. Lysozymes*. Springer. 45–57.
- Uribe, C., Folch, H., Enríquez, R. and Moran, G. (2011). Innate and adaptive immunity in teleost fish: a review. *Vet. Med. (Praha)*. 56(10): 486–503.
- Valdez-Muñoz, C., Aragón-Noriega, E. A., Ortega-Rubio, A., Salinas-Zavala, C. A., Arreola-Lizárraga, J. A., Hernández-Vázquez, S. and Morales, L. F. B. (2010). Distribución y abundancia de juveniles de totoaba, *Totoaba macdonaldi* y la salinidad del hábitat de crianza. *Interciencia*. 35(2): 136–139.
- Valenzuela-Quiñonez, F., Arreguín-Sánchez, F., Salas-Márquez, S., García-De León, F. J., Garza, J. C., Román-Rodríguez, M. J. and De-Anda-Montañez, J. A. (2015). Critically endangered totoaba *Totoaba macdonaldi*: signs of recovery and potential threats after a population collapse. *Endanger. Species Res.* 29(1): 1–11.
- Vazquez-Morado, L. E., Robles-Zepeda, R. E., Ochoa-Leyva, A., Arvizu-Flores, A. A., Garibay-Escobar, A., Castillo-Yañez, F. and Lopez-zavala, A. A. (2021). Biochemical characterization and inhibition of thermolabile hemolysin from *Vibrio parahaemolyticus* by phenolic compounds. *PeerJ*. 9: e10506.
- Villaverde, A. and Mar Carrió, M. (2003). Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol. Lett.* 25(17): 1385–1395.
- Vocadlo, D. J., Davies, G. J., Laine, R. and Withers, S. G. (2001). Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature*. 412(6849): 835–838.
- Wang, R., Feng, J., Li, C., Liu, S., Zhang, Y. and Liu, Z. (2013). Four lysozymes (one c-type and three g-type) in catfish are drastically but differentially induced after bacterial infection. *Fish Shellfish Immunol.* 35(1): 136–145.
- Weaver, L. H., Grütter, M. G. and Matthews, B. W. (1995). The refined structures of goose lysozyme and its complex with a bound trisaccharide show that the ‘Goose-type’ lysozymes lack a catalytic aspartate residue. *J. Mol. Biol.* 245(1): 54–68.
- Wedemeyer, W. J., Welker, E., Narayan, M. and Scheraga, H. A. (2000). Disulfide bonds and protein folding. *Biochemistry*. 39(15): 4207–4216.
- Wei, K., Ding, Y., Yin, X., Zhang, J. and Shen, B. (2020). Molecular cloning, expression analyses

- and functional characterization of a goose-type lysozyme gene from *Bostrychus sinensis* (family: Eleotridae). *Fish Shellfish Immunol.* 96: 41–52.
- Wei, S., Huang, Y., Cai, J., Huang, X., Fu, J. and Qin, Q. (2012). Molecular cloning and characterization of c-type lysozyme gene in orange-spotted grouper, *Epinephelus coioides*. *Fish Shellfish Immunol.* 33(2): 186–196.
- Wei, S., Huang, Y., Huang, X., Cai, J., Wei, J., Li, P., Ouyang, Z. and Qin, Q. (2014). Molecular cloning and characterization of a new G-type lysozyme gene (Ec-lysG) in orange-spotted grouper, *Epinephelus coioides*. *Dev. Comp. Immunol.* 46(2): 401–412.
- Weinkam, P., Zimmermann, J., Romesberg, F. E. and Wolynes, P. G. (2010). The folding energy landscape and free energy excitations of cytochrome C. *Acc. Chem. Res.* 43(5): 652–660.
- Whang, I., Lee, Y., Lee, S., Oh, M.-J., Jung, S.-J., *et al.* (2011). Characterization and expression analysis of a goose-type lysozyme from the rock bream *Oplegnathus fasciatus*, and antimicrobial activity of its recombinant protein. *Fish Shellfish Immunol.* 30(2): 532–542.
- Wildegger, G. and Kiefhaber, T. (1997). Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate. *J. Mol. Biol.* 270(2): 294–304.
- Wohlkönig, A., Huet, J., Looze, Y. and Wintjens, R. (2010). Structural relationships in the lysozyme superfamily: significant evidence for glycoside hydrolase signature motifs. *PLoS One.* 5(11): e15388.
- Wolynes, P. G. (2015). Evolution, energy landscapes and the paradoxes of protein folding. *Biochimie.* 119: 218–230.
- Xie, J.-W., Cheng, C.-H., Ma, H.-L., Feng, J., Su, Y.-L., Deng, Y.-Q. and Guo, Z.-X. (2019). Molecular characterization, expression and antimicrobial activities of a c-type lysozyme from the mud crab, *Scylla paramamosain*. *Dev. Comp. Immunol.* 98: 54–64.
- Xu, Z., Parra, D., Gómez, D., Salinas, I., Zhang, Y.-A., *et al.* (2013). Teleost skin, an ancient mucosal surface that elicits gut-like immune responses. *Proc. Natl. Acad. Sci.* 110(32): 13097–13102.
- Yamaguchi, H. and Miyazaki, M. (2014). Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules.* 4(1): 235-51.

- Yin, Z.-X., He, J.-G., Deng, W.-X. and Chan, S.-M. (2003). Molecular cloning, expression of orange-spotted grouper goose-type lysozyme cDNA, and lytic activity of its recombinant protein. *Dis. Aquat. Organ.* 55(2): 117–123.
- Yoshikuni, Y., Ferrin, T. E. and Keasling, J. D. (2006). Designed divergent evolution of enzyme function. *Nature.* 440(7087): 1078–1082.
- Yu, L., Sun, B., Li, J. and Sun, L. (2013). Characterization of a c-type lysozyme of *Scophthalmus maximus*: Expression, activity, and antibacterial effect. *Fish Shellfish Immunol.* 34(1): 46–54.
- Zechel, D. L. and Withers, S. G. (2000). Glycosidase mechanisms: Anatomy of a finely tuned catalyst. *Acc. Chem. Res.* 33(1): 11–18.
- Zhang, J., Kong, X., Zhou, C., Li, L., Nie, G. and Li, X. (2014). Toll-like receptor recognition of bacteria in fish: Ligand specificity and signal pathways. *Fish Shellfish Immunol.* 41(2): 380–388.
- Zhang, Y., Yang, H., Song, W., Cui, D. and Wang, L. (2018). Identification and characterization of a novel goose-type and chicken-type lysozyme genes in Chinese rare minnow (*Gobiocypris rarus*) with potent antimicrobial activity. *Genes Genomics.* 40(6): 569–577.
- Zhao, L., Sun, J. and Sun, L. (2011). The g-type lysozyme of *Scophthalmus maximus* has a broad substrate spectrum and is involved in the immune response against bacterial infection. *Fish Shellfish Immunol.* 30(2): 630–637.
- Zheng, W., Tian, C. and Chen, X. (2007). Molecular characterization of goose-type lysozyme homologue of large yellow croaker and its involvement in immune response induced by trivalent bacterial vaccine as an acute-phase protein. *Immunol. Lett.* 113(2): 107–116.
- Zou, J. and Secombes, C. J. (2016). The function of fish cytokines. *Biology.* 5(2)23:1-35.

APPENDIXES

Scientific productivity derived from this research project (Published papers).



Contents lists available at ScienceDirect

Developmental and Comparative Immunology

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

Short communication

Molecular characterization and expression analysis of the chicken-type and goose-type lysozymes from totoaba (*Totoaba macdonaldi*)

Elena N. Moreno-Córdova^a, María A. Islas-Osuna^b, Carmen A. Contreras-Vergara^b,
 Alonso A. López-Zavala^a, Eduardo Ruiz-Bustos^a, Mónica G. Reséndiz-Sandoval^c,
 Francisco J. Castillo-Yañez^a, Michael F. Criscitiello^{d,e,**}, Aldo A. Arvizu-Flores^{a,*}

^a Departamento de Ciencias Químico Biológicas, Universidad de Sonora. Blvd. Rosales S/N, Centro. Hermosillo, SON, CP, 83000, Mexico

^b Departamento de Tecnología de Alimentos de Origen Vegetal, Centro de Investigación en Alimentación y Desarrollo, Gustavo Enrique Astiazarán Rosas, NO. 46. Hermosillo, SON, CP, 83304, Mexico

^c Laboratorio de Inmunología, Centro de Investigación en Alimentación y Desarrollo, Gustavo Enrique Astiazarán Rosas, NO. 46. Hermosillo, SON, CP, 83304, Mexico

^d Comparative Immunogenetics Laboratory, Department of Veterinary Pathobiology, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX, 77843, USA

^e Department of Microbial Pathogenesis and Immunology, College of Medicine, Texas A&M University. Bryan, TX, 77807, USA

ARTICLE INFO

Keywords:

Totoaba
 c-type lysozyme
 g-type lysozyme
 Lysozyme superfamily
 Innate immunity

ABSTRACT

Lysozymes play a key role in innate immune response to bacterial pathogens, catalyzing the hydrolysis of the peptidoglycan layer of bacterial cell walls. In this study, the genes encoding the c-type (TmLyzc) and g-type (TmLyzg) lysozymes from *Totoaba macdonaldi* were cloned and characterized. The cDNA sequences of TmLyzg and TmLyzc were 582 and 432 bp, encoding polypeptides of 193 and 143 amino acids, respectively. Amino acid sequences of these lysozymes shared high identity (60–90%) with their counterparts of other teleosts and showed conserved functional-structural signatures of the lysozyme superfamily. Phylogenetic analysis indicated a close relationship with their vertebrate homologues but distinct evolutionary paths for each lysozyme. Expression analysis by qRT-PCR revealed that TmLyzc was expressed in stomach and pyloric caeca, while TmLyzg was highly expressed in stomach and heart. These results suggest that both lysozymes play important roles in defense of totoaba against bacterial infections or as digestive enzyme.

1. Introduction

Totoaba (*Totoaba macdonaldi*) is a demersal teleost fish and one of the largest members of the *Sciaenidae* family, commonly known as croakers or drums. This species is endemic and only found in the central and northern Gulf of California, Mexico (Cisneros-Mata et al., 1997). Due to its extremely high value in Asian markets, totoaba was an important species for commercial fisheries until 1975, when its natural population was severely reduced due to unregulated fishing, by-catch and habitat loss. Since then, totoaba is declared to be a threatened species and is currently classified as critically endangered (Bobadilla et al., 2011). Aquaculture of totoaba represents a feasible strategy to recover the natural population (Mata-Sotres et al., 2015; True et al., 1997). Nevertheless, large-scale hatchery rearing of this species faces

notable technical difficulties, including those related to disease outbreak control. For this reason, understanding the immune response of totoaba against pathogens is critical to prevent diseases, promote fish health and develop immunoprophylactic strategies to improve farming conditions. However, knowledge of this fish at genetic, biochemical and immunological levels is lacking and demands further investigation (González-Félix et al., 2018; Reyes-Becerril et al., 2016).

Amongst the vast innate immune arsenal, lysozyme is one of the most important molecules that participate in protection against bacterial pathogens (Ragland and Criss, 2017). Lysozyme (EC 3.2.1.17) catalyzes the hydrolysis of β -(1,4)-glycosidic bonds between *N*-acetyl-muramic acid and *N*-acetyl-glucosamine present in the peptidoglycan layer of bacterial cell walls (Kirby, 2001; Sukhithasri et al., 2013). This enzyme is ubiquitously distributed in diverse organisms, including animals,

* Corresponding author. Departamento de Ciencias Químico Biológicas, Universidad de Sonora. Blvd. Rosales S/N, Centro. Hermosillo, SON, CP, 83000, Mexico.

** Corresponding author. Comparative Immunogenetics Laboratory, Department of Veterinary Pathobiology, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX, 77843, USA.

E-mail addresses: mcriscitiello@cvm.tamu.edu (M.F. Criscitiello), aldo.arvizu@unison.mx (A.A. Arvizu-Flores).

<https://doi.org/10.1016/j.dci.2020.103807>

Received 10 April 2020; Received in revised form 23 July 2020; Accepted 23 July 2020

Available online 29 July 2020

0145-305X/© 2020 Elsevier Ltd. All rights reserved.