Lilla Roszik

The BiP Chaperone from the Endoplasmic Reticulum of Zebrafish (*Danio rerio*) and Golden Hamster (*Mesocricetus auratus*)

UNIVERSIDADE DO ALGARVE

Faculdade de Ciências e Tecnologia

2025

Lilla Roszik

The BiP Chaperone from the Endoplasmic Reticulum of Zebrafish (*Danio rerio*) and Golden Hamster (*Mesocricetus auratus*)

Mestrado em Biologia Marinha

Supervisor:

Eduardo Pinho Melo

UNIVERSIDADE DO ALGARVE

Faculdade de Ciências e Tecnologia

2025

Abstract

The Hsp70 family of molecular chaperones plays a crucial role in maintaining cellular proteostasis by promoting protein folding, transport, and degradation. One of its component proteins, the binding immunoglobulin protein (BiP), is an essential endoplasmic reticulum chaperone required for proper protein folding and stress response.

The primary objective of this work was to clone the open reading frame of the BiP gene from zebrafish (*Danio rerio*), a model organism that is frequently used to study the molecular and genetic causes of human disorders. More than 70% of the genes between humans and zebrafish are similar, including many orthologs connected to neurodegenerative illnesses like Alzheimer's.

For the sake of further functional investigations, the open reading frame of the zebrafish BiP gene was amplified and inserted into a bacterial expression vector. From zebrafish, the total RNA of caudal cell lines was extracted and reverse transcribed to create complementary DNA. The target DNA sequence was amplified by polymerase chain reaction (PCR) using specific primers for BiP. The golden hamster BiP ORF was removed from the backbone pQE10, and the amplified product was cloned into the vector. After transforming the recombinant plasmid into *Escherichia coli* DH5α for propagation and isolation, the insertion of the open reading frame of zebrafish BiP was sent to sequencing. Unfortunately, *E coli* recombines the zebrafish BiP ORF, contrary to the golden hamster ORF previously cloned into the same vector. Different *E coli* strains deficient in DNA recombination should be tested soon.
This work aimed to establish the foundation for further investigations into the activity and biochemical characteristics of zebrafish BiP and its interactions with co-chaperones and client proteins in comparison with the mammalian BiP from golden hamster. At the longer term, we want to understand how the chaperone machinery of the endoplasmic reticulum has evolved in resolving protein aggregates associated with neurodegenerative diseases.

Keywords: Hsp70 family, BiP (Binding Immunoglobulin Protein), Zebrafish (*Danio rerio*), Golden hamster (*Mesocricetus auratus*), protein misfolding, neurodegenerative diseases

Resumo

As células possuem um conjunto de proteínas, designadas chaperones moleculares, absolutamente essenciais para assegurar a síntese de proteínas funcionais, responder a condições de stress que afetam a integridade do proteoma e direcionar proteínas danificadas para os mecanismos de reciclagem celular como o lisossoma e o proteassoma. A família de chaperones moleculares Hsp70 é das mais relevantes, desempenhando um papel crucial na manutenção da homeostase do proteoma celular (também designada proteostase). Um dos membros dessa família, o chaperone BiP (binding immunoglobulin protein), é essencial no retículo endoplasmático, a chamada via secretória das células, onde cerca de 20% do proteoma é sintetizado. Tem um papel fundamental no "folding" das proteínas no retículo endoplasmático e na resposta ao stress celular resultante do "misfolding" e agregação proteica que está associada à morte das células neuronais (neurodegeneração) em várias patologias como por exemplo a doença de Alzheimer. Em 2022, descobriu-se que a BiP, para além da função clássica de chaperone envolvido na catálise do folding de proteínas, possui a capacidade de desagregar proteínas no retículo endoplasmático (Melo et al., 2022). O estudo da capacidade da BiP para resolver agregados proteicos e impedir a sua acumulação no reticulo endoplasmático é importante para perceber os mecanismos celulares que previnem a neurodegeneração e, eventualmente, desenvolver estratégias terapêuticas para tratamento das doenças neurodegenerativas.

O objetivo principal deste trabalho consistiu na clonagem da sequência codificante (designada "Open Reading Frame" – ORF) da BiP do peixe zebra (*Danio rerio*), num vetor adequado para expressão da proteína na bateria *Escherichia coli*. O peixe zebra é um organismo modelo, frequentemente usado para estudar as causas genéticas e moleculares de doenças humanas. Mais de 70% dos genes humanos e do peixe zebra são similares, incluindo muitos genes ortologos relacionados com doenças neurodegenerativas. A ORF da BiP do peixe zebra foi amplificada e inserida num vetor para expressão da proteína na bactéria *E. coli.* A partir do ácido ribonucleico (RNA) total extraído de uma linha celular contínua isolada da cauda do peixe zebra, sintetizou-se a cadeia de ácido desoxirribonucleico complementar (cDNA) usando o enzima transcriptase reversa e "primers" com 18 nucleótidos timina e 6 nucleótidos aleatórios ("random hexamers"). O cDNA foi depois usado como molde, juntamente com primers específicos para a sequência alvo, para amplificar a ORF da BiP do peixe zebra pela reação de PCR ("polymerase chain reaction"). Nesta reação de PCR usou-se uma DNA polimerase à

prova de erro e, após otimização da temperatura de hibridação dos primers, conseguiu-se amplificar a ORF da BiP do peixe zebra. Para obter maiores quantidades desta ORF, usou-se o produto desta primeira reação de amplificação, após diluição adequada, como molde para nova reação de PCR, um procedimento frequentemente designado por "nested" PCR. Antes de avançar para a clonagem da ORF da BiP do peixe zebra num vetor adequado, mandou-se sequenciar o produto de PCR para confirmar a sequência alvo. Pela análise das sequenciações obtidas para as duas cadeias complementares de DNA, em comparação com a sequência esperada (gene ID 378848 e sequencia referência NM_213058.1 do National Center for Biotechnology Information), confirmou-se que era efetivamente a ORF da BiP do peixe zebra. Confirmou-se 92% da sequência e verificou-se a presença de 9 mutações silenciosas e uma mutação que substitui o aminoácido prolina por uma leucina, na posição 374 da BiP do peixe zebra. Para clonar esta ORF no vetor pQE10 desenhado para expressão de proteínas heterologas em *E. coli* e com resistência ao antibiótico ampicilina, primeiro digeriu-se o vetor com os enzimas de restrição BamHI e AflII para remover a ORF da BiP de hamster sírio, previamente clonada no vetor pQE10. Depois, a ORF da BiP do peixe zebra amplificada por PCR, foi também duplamente digerida com os mesmos enzimas de restrição cujas sequências de reconhecimento e corte estavam incluídas nos primers usados na amplificação. Por fim, procedeu-se à ligação do ORF ao vetor, usando o enzima DNA ligase T4. Após transformação de células de *E. coli* DH5α competentes, obtiveram-se colónias que foram crescidas em meio líquido para isolamento do plasmídeo e posterior sequenciação. Inesperadamente, depois de sequenciar várias colónias, verificou-se que ocorre recombinação da ORF pelas células de *E. coli* DH5α. Esta recombinação resulta da sequência da ORF da BiP do peixe zebra, visto que não ocorria quando a ORF inserida no vetor era a da BiP de hamster sírio. Para solucionar este problema, vai usar-se uma estirpe de *E. coli* alternativa à DH5α, designada Stbl3 e recomendada para reduzir a frequência de recombinação homologa.

Neste trabalho, efetuou-se ainda a expressão da ORF da BiP de hamster sírio previamente clonada no vetor pQE10 usando células de *E. coli* BL21(DE3) adequadas para expressão de proteínas heterólogas. Após expressão, a BiP de hamster sírio foi purificada por cromatografia de afinidade por metal imobilizado (IMAC – "Immobilized metal affinity chromatography) e o grau de pureza após purificação analisado por electroforese em gel de poliacrilamida desnaturante (SDS_Page – "Sodium dodecyl sulfate polyacrylamide gel electrophoresis").

A clonagem da ORF da BiP do peixe zebra tinha como objetivo estudar a função e o mecanismo da BiP de dois vertebrados evolutivamente distantes, o peixe zebra e o hamster sírio. Pretende-se no futuro, perceber se a evolução ao longo da linhagem dos vertebrados conduziu a uma BiP

mais eficiente na sua atividade de hidrólise de ATP (ATPase), na sua capacidade de catalisar o enrolamento correto de outras proteínas, na sua interação com outros co-chaperones importantes na manutenção da proteostase celular e, principalmente, na sua capacidade de resolver agregados proteicos resultantes do "misfolding" das proteínas. A longo prazo, pretende-se entender como é que a maquinaria de chaperones do retículo endoplasmático evoluiu para impedir/resolver a acumulação de agregados proteicos associados com doenças neurodegenerativas.

Palavras chave: família de proteínas Hsp70, chaperone BiP ("Binding Immunoglobulin Protein"), peixe zebra (*Danio rerio*), hamster sírio (*Mesocricetus auratus*), "misfolding" proteico, doenças neurodegenerativas

Declaração de autoria de trabalho

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

Declaração de direitos de autor

 Faro, 30 Janeiro de 2025

Acknowledgments

Above all, I would like first to express my most sincere gratitude to my supervisor, Professor Eduardo Pinho Melo, without whose invaluable guidance, expertise, encouragement, this journey would not have been possible. It has been my pleasure to learn so much under his mentorship, and I am grateful for the opportunity to have his support in my academic growth.

To Fernando, the best lab partner I could have ever had—thank you for your infinite patience and for always being willing to take the time to explain something to me when I needed it. Your support and your company have made every difficult moment in the lab easier and more enjoyable.

A very big thank you to my parents and my sister for the unconditional love, support, and faith they have given me, which have been my source of strength. I would like to take the opportunity to thank everyone who has ever loved me and always stood by my side no matter the distance or the obstacles.

Also want to say thank you to my boyfriend, Zoli, who is steadily on my side, and supporting me not just emotionally, but in the growth of my professional carrier.

To my best friends from Hungary- Fanni and Tamás, - I owe my success to your constant encouragement and support. Also thank you Bálint and Liza, my great friends from my home. All of your friendship has been a source of motivation and comfort to me, even when the miles between us. I am always grateful to have you by my side and believing in me.

I would like to express my deepest gratitude to my trainer and friend, Nagy Ferenc Gábor, for his unwavering support. Without his selfless help and constant presence in my life, completing this thesis would not have been possible.

Last but not least, my friends from the University of Algarve, thank you for all the unforgettable moments we have spent together and for the unwavering support you have given me during my studies. The amazing time we spent together, made this journey not only rewarding but also enjoyable.

I also want to dedicate this acknowledgment to Pirulito, my dog, and my loyal companion. I got Pirulito from Praia de Faro when I began my university studies and he has been with me

every day since. He has been a source of great joy and comfort and has been with me once daily throughout what has been the hardest time of this journey. Thank you, Pirulito, for being my source of unconditional love and always reminding me to find happiness in the simplest of things and for being my inspiration.

Table of contents

List of figures

1. Introduction

2.9.Recombinant DNA technology

Following the cloning of a complete bacterial gene (Chang & Cohen, 1974), the development of heterologous recombinant gene expression technologies was marked by the finding that DNA from one species can be cloned and replicated in another. In the late 1960s, Hamilton Smith and Daniel Nathans identified the first restriction endonuclease (Smith, 1979). The discovery of the enzyme DNA ligase in 1967 by Martin Gellert made it possible to combine DNA molecules from various sources to create recombinant DNA molecules by covalently joining their ends with phosphodiester linkages (Bedows, 1977). The groundbreaking discovery of PCR in the 1980s, which amplified DNA fragments *in vitro*, paved the way for the creation of complex cloning techniques.
Another significant development in molecular biology during the past ten years has been the routine synthesis of large DNA fragments of a particular sequence, which has become widely available and reasonably priced (Morrow et al., 1974).

2.10. Zebrafish as a Model Organism for human diseases

About 450 million years ago, the human (tetrapod) lineage split off from the zebrafish (teleost-bony fish) evolutionary lineage (Kumar and Hedges, 1998). Since separating from the tetrapod lineage, teleosts seem to have experienced another cycle of genome duplication, which was followed by the loss of many of the duplicated genes (Catchen et al., 2011). Zebrafish (*Danio rerio*) are small, resilient freshwater fish originally from India. Initially, they served as a model organism for research on the development of vertebrates. Zebrafish allow for the simple and efficient manipulation of many genes at physiologically appropriate levels (Newman et al., 2012). Zebrafish are vertebrates and so more relevant to comprehending human biology than invertebrate models like *Drosophila melanogaster* and *Caenorhabditis elegans*.
In the past ten years, the zebrafish model has been used more and more to study a wide range of human illnesses (Lieschke and Currie, 2007). Zebrafish is widely regarded, as an excellent model organism in molecular biology for various reasons, among them is the fact that they have genetic similarity to humans. Zebrafish share approximately 70% of their genes with humans, and about 84% of genes associated with human diseases have orthologues in zebrafish (Crouzier et al., 2021). Because of their transparency, high fertility, relatively low cost of

maintenance, and the possibility to manipulate their genome, zebrafish are an important model organism for studying human diseases and genetic functions. Zebrafish embryos are practically transparent, which enables live observation and modification of the embryonic development. Within 24-48 hrs post fertilization, they mimic many organs and systems, making it a fast model for developmental biology studies. Research employing zebrafish has uncovered distinctive features of genes that have been challenging to discover in rodent models, including genes orthologous to those altered in familial Alzheimer's disease. Therefore, zebrafish has become an increasingly common model for studying Alzheimer's dementia, complementing research employing other species to help us fully comprehend the condition (Newman et al., 2014). Indeed, genes orthologous to human genes known to play a role in AD are found in zebrafish. While the *APPA* and *appb* genes are "co-orthologs" of human amyloid precursor protein (APP) (Musa et al., 2001), the *psen1* (Leimer et al., 1999) and *psen2* (Groth et al., 2002) genes are orthologs of human *PSEN1* and *PSEN2*, respectively.

Autosomal-dominant Early Onset Alzheimer Disease (ADEOAD), whose frequency is estimated to be 5.3 per 100,000 persons at risk, is caused by mutations and duplications in amyloid protein precursor (APP), presenilin-1 (PSEN1), and presenilin-2 (PSEN2) (Sherrington et al., 1995; Rogaev et al., 1995; Levy-Lahad et al., 1995; Sherrington et al., 1996). A percentage of early-onset Alzheimer disease (eoAD) has been linked to mutations in three genes: presenilin 1 (PSEN1), amyloid precursor protein (APP), and presenilin 2 (PSEN2) (Krüger et al., 2012). According to the Alzforum database, PSEN1 is the most often implicated gene, with 221 mutations identified as harmful (Campion et al., 1999). With 32 pathogenic variants recorded, APP is the second most often affected gene. Moreover, 19 distinct PSEN2 pathogenic mutations have been documented. The amyloid-β precursor protein, which is encoded by APP, is processed by the β-secretase and γ-secretase complex to produce the amyloid β (Aβ) peptide, a crucial step in the pathophysiology of AD. A series of processes leading to AD are indeed triggered by the Aβ peptide's aggregation in the brain's parenchyma. When it aggregates in the cerebro meningeal arteries, it causes cerebral amyloid angiopathy (CAA), which is often linked to AD and causes white matter abnormalities and recurrent hemorrhagic strokes. Presenilins, the catalytic subunit of the γ-secretase complex, are encoded by PSEN1 and PSEN2 (Haass et al., 2012; Campion et al., 2016).

The increased aggregation of the Aβ peptide in the brain's parenchyma is believed to result from AD-EOAD causative mutations through one of two mechanisms: either a more

aggregation-prone form of the Aβ peptide or an increase in the production of all Aβ species (such as APP duplications or APP mutations around the β cleavage site).

2.11. Chaperone Proteins

Correct cellular function depends on protein homeostasis, also known as proteostasis, which is the balance of protein synthesis, folding, trafficking, assembly, and degradation. Cells have evolved a variety of mechanisms to regulate this process under stress (Zuiderweg et al., 2017). A ubiquitous class of cellular proteins known as molecular chaperones facilitates other polypeptides' proper folding and in certain situations, their assembly into oligomeric structures, although they are not components of those final structures. Although known chaperones lack the steric information necessary for protein folding, they do prevent ineffective folding and assembly pathways that could otherwise serve as dead-end kinetic traps and result in the wrong structures (Ellis, 1990).

Chaperones are similar in that they can identify and bind foreign proteins, which stops them from aggregating randomly (Beissinger & Buchner, 1998). One of the most common groups of chaperones, the 70-kDa heat shock protein (Hsp70) family of molecular chaperones is highly conserved across all organisms. The Hsp70 family of proteins regulates all aspects of cellular proteostasis, including the folding of developing protein chains, the import of proteins into organelles, the retrieval of proteins from aggregation, and the formation of multiprotein complexes. When a nucleoside binds to an adenosine-5′-triphosphate (ATP)-dependent chaperone (such as GroEL, Hsp70, or Hsp90), the chaperone may undergo significant conformational changes that enable it to switch between high- and low-affinity states for the substrate proteins (Ellis, 1990). Through improving cell viability and aiding in the repair of protein damage, these chaperones increase organismal survival and longevity in the face of proteotoxic stress. Extracellular Hsp70s play multiple cytoprotective and immunomodulatory functions. The last role can be chaperonins or stimulators of innate immune responses, or to facilitate the cross-presentation of immunogenic peptides through major histocompatibility complex (MHC) antigens. All of the physiological actions of Hsp70s depend on their ATP-regulated ability to bind with exposed hydrophobic surfaces of proteins. During the folding of polypeptides, ATP hydrolysis and the exchange of adenosine diphosphate (ADP) with ATP are essential processes for substrate binding and Hsp70 release.

Hsp70 co-chaperones are a group of proteins that bind to different Hsp70 subdomains and

subsequently alter the chaperone's function (Radons, 2016). The chaperone machinery of the endoplasmic reticulum (ER), a site responsible for the manufacturing of about one-third of the cell's protein repertoire (including disease determinants, e.g. APP) is relevant to understand how cells deal with proteotoxicity resulting from protein misfolding and aggregation of secretory proteins. As newly generated proteins are translocated into the endoplasmic reticulum, the chaperone binding immunoglobulin protein (BiP), which is an Hsp70 molecular chaperone situated in the ER lumen, binds them and keeps them in a state suitable for further folding and oligomerization. BiP also participates in the retrograde transport of abnormal proteins intended for proteasomal degradation across the ER membrane and is a crucial part of the translocation process (Multhoff et al., 1995; Multhoff & Hightower, 1996; Hantschel et al., 2000; Gehrmann et al., 2005).

1.1.1. The Hsp70 family

Hsp70 is a nanomachine that aids in protein folding by utilizing conformational changes driven by ATP hydrolysis. The common domain structure of Hsp70s, (Flaherty et al. 1990) is composed of a 44-kDa N-terminal nucleotide-binding domain (NBD) that binds and hydrolyzes ATP, a middle domain with protease-sensitive sites, and a 28-kDa C-terminal substrate-binding domain (SBD) that binds extended polypeptides (Figure 1.1). The ATP-binding pocket is surrounded by four subdomains (IA, IB, IIA, and IIB) that make up the nucleotide-binding domain (Flaherty et al. 1990). The SBD is separated into a C-terminal α-helical subdomain (SBDα), which functions as a flexible lid, and an N-terminal β-sheet (SBDβ) domain (Zhang, Hong, & et al., 2023).

Stable misfolded or aggregated proteins can be recognized by Hsp70s thanks to their unfoldase activity, which also allows the proteins to unfold and spontaneously refold into natively refoldable species (Sharma et al. 2010). Together with dimeric co-chaperone Hsp40 and nucleotide exchange factors (NEFs), Hsp70 recognizes stable misfolded polypeptides. These proteins undergo repeated cycles of binding, ATP-dependent unfolding, and spontaneous refolding to become native proteins. The N-terminal nucleotide-binding domain of Hsp70 is a V-shaped structure made up of two subdomains (lobes) that envelope the ATP-binding site. The NBD and the SBD interact allostery in the function of Hsp70. The SBD undergoes conformational changes driven by the NBD, while the NBD interface receives substrate-

induced conformational changes from the SBD (Zhuravleva & Gierasch, 2015; Taylor et al., 2018).

The accessibility of hydrophobic regions on proteins, which are typically hidden in the protein core, indicates a misfolded state on client proteins. There are two possible conformations for the SBD: open and closed.ATP interaction to the NBD causes the SBD open conformation, which has a poor affinity for client protein binding. The closed conformation is driven by ATP hydrolysis; structural changes start at the NBD and are transferred through the linker to cause the SBD base and lid subdomains to converge and bind to the client protein with high affinity. In the open conformation, the flexible and conserved linker makes contact with the NBD, and in the closed conformation, it separates from it.

J-protein (Hsp40) co-chaperones raise the basal ATPase activity of Hsp70 proteins, which are typically modest. In addition to binding client proteins to Hsp70, J-proteins also promote ATPase activity and effective client protein locking (Mayer & Bukau, 2005; Kampinga & Craig, 2010).

Another group of co-chaperones called the nucleotide exchange factors (NEFs) are responsible for ADP dissociation, which is necessary for the client protein to be released. Therefore, NEFs promote ADP/ATP exchange, inducing the open conformation, while J-proteins cause ATP hydrolysis, and enhance the closed conformation.

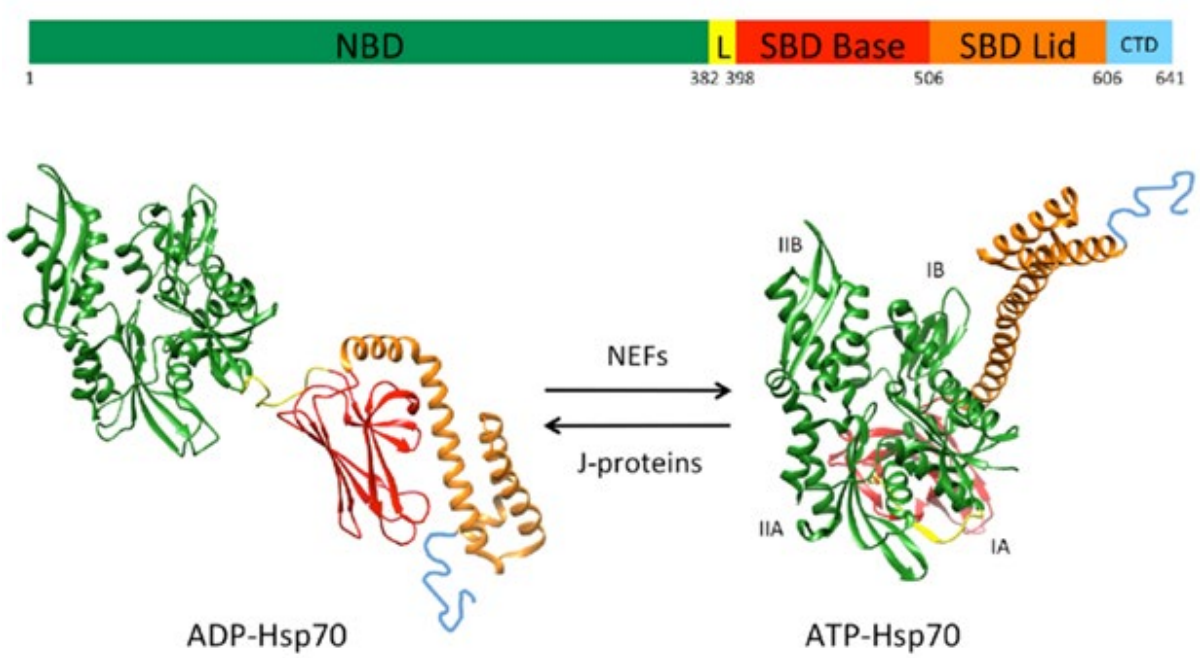


Figure 1.1 Primary and tertiary structure of Hsp70 showing the nucleotide-binding domain (NBD), linker (L), substrate-binding domain (SBD) including the lid (SBD lid) and the C-terminal domain (CTD). Hsp70 atomic structures in the closed (left structure PDB 2KHO) and open (right structure PDB 4JNE) conformations. NBD subdomains are referred to as IB, IIB, IA, and IIA. (Fernández-Fernández & Valpuesta, 2018). The role of nucleotide exchange factors (NEFs) and J-proteins co-chaperones on the equilibrium between the ADP- and ATP bound state is shown by the arrows.

Hsp70s are powerful anti-apoptotic proteins that prevent apoptosis on several levels. Hsp70 inhibits the assembly of the death-inducing signaling complex (DISC) (Guo et al. 2005b) and, on the other hand, prevents mitochondrial translocation and activation of Bax, which stops mitochondrial membrane permeabilization and the release of pro-apoptotic factors (Stankiewicz et al., 2005; Er et al., 2006).

1.1.1.1.The binding immunoglobulin protein BiP

BiP was first discovered on its own as the glucose-regulated protein Grp78 and as the immunoglobulin heavy chain binding protein. When it was established that BiP is the Hsp70 molecular chaperone family member that is found in the ER, its role as a broad specificity

molecular chaperone became evident. BiP is essential to every organism's biology, as evidenced by its great conservation across species. BiP functions as a molecular chaperone that aids in protein folding, import into the ER, protein elimination through endoplasmic reticulum associated degradation (ERAD), and Ca^{2+} homeostasis regulation. When proteins are misfolded, under glycosylated, or unassembled and their transport out of the ER is inhibited, BiP binds to these proteins temporarily. Furthermore, BiP inhibits apoptosis and promotes autophagy by activating the UPR pathway. BiP's links to cancer, cardiovascular illness, neurological disease, metabolic disease, autoimmune disease, and infectious disease demonstrate how frequently ER stress triggers these processes (Pouyssegur et al., 1977; Munro & Pelham, 1986).

The interaction of BiP's functions in autophagy, apoptosis, and unfolded protein response (UPR) needs more research to comprehend how BiP contributes to cell survival in response to ER stress and other illness situations. Drugs that target BiP have shown encouraging results in recent trials, especially in improving some cancer treatments and possible concurrent heart failure. These studies can be broadened to find novel chemical inhibitors and inducers of BiP to treat other pertinent disorders (Wang et al., 2017).

Unlike other Hsp70 proteins, BiP does not interact with native polypeptides. By identifying unfolded polypeptides and preventing intermolecular aggregation, BiP promotes the folding and assembly of newly synthesized proteins by keeping them in a state that allows for further folding and oligomerization. BiP three-dimensional structure has two major domains, just like other Hsp70 proteins: a C-terminal substrate-binding domain and an N-terminal domain containing the ATPase catalytic site (Figure 1.2.). The communication between these domains controls the duration and affinity of polypeptide binding (McKay, 1993). There is minimal interaction between the two domains in the nucleotide-free (apo) and ADP-bound states (Buchberger et al., 1995; Swain et al., 2007; Buchberger et al., 2008; Bertelsen et al., 2009). Similarly, to the isolated SBD, this state exhibits strong affinity for polypeptide substrates and delayed binding and release rates (Flynn et al., 1989; Schmid et al., 1994). While the resulting affinity is two to three orders of magnitude lower, the two domains are strongly linked in the ATP-bound state, leading to significantly faster kinetics in both binding and release of polypeptide substrates (Schmid et al., 1994).

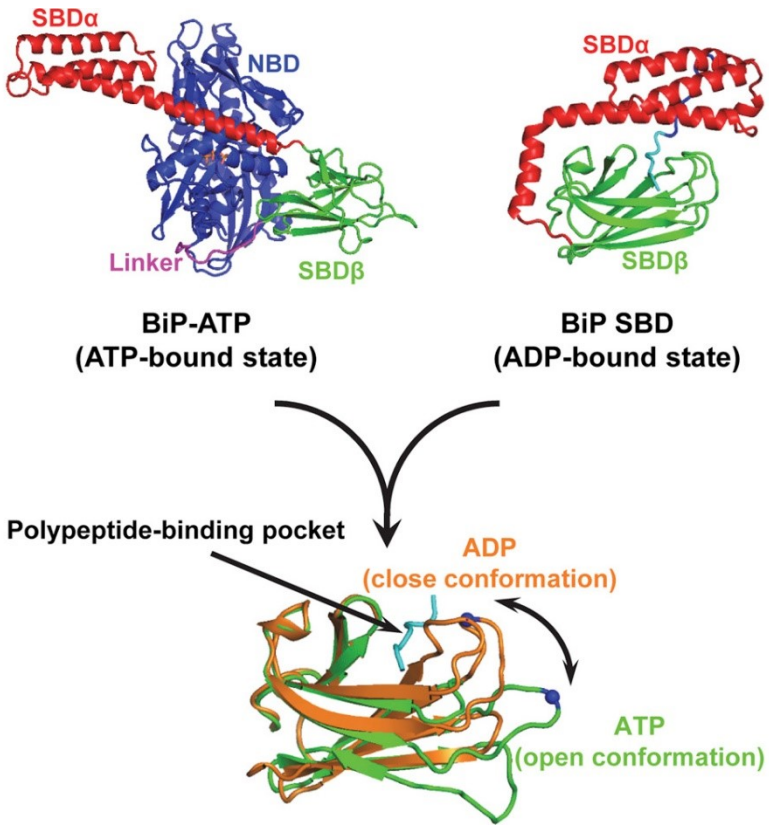


Figure 1.2 Close and Allosteric Opening of the Polypeptide-Binding Site in human Chaperone BiP (Yang et al., 2015)

The C-terminal domain's affinity for polypeptides is determined by the presence of ATP (low affinity but fast binding) or ADP (high affinity and sluggish exchange) in the nucleotide binding domain. The rate at which ATP is hydrolyzed by the N-terminal ATPase domain is accelerated upon polypeptide binding to the C-terminal domain. The nucleotide-binding domain (NBD) and substrate-binding domain (SBD) take on an open conformation when bound to ATP, which promotes substrate release. On the other hand, the polypeptide-binding pocket's close conformation is induced by the ADP-bound state, stabilizing the substrate association (Figure 1.2). As referred, effective chaperone function of BiP depends on the ATP-induced allosteric coupling (Mayer and Bukau, 2005). Therefore, cycles of binding and release of an unfolded polypeptide chaperoned by BiP can occur; the length of these cycles is determined by the rate at which BiP undergoes adenine nucleotide exchange of ATP for ADP and ATP hydrolysis. The polypeptide will interact with the chaperone until it no longer exhibits any BiP-binding motifs, but each time it is released, it will have the chance to advance down the folding pathway (McKay, 1993).

The nucleotide is bound in a deep cleft between the two lobes that make up the ATPase domain, which is made up of four subdomains. It is easy to map the amino acid sequence of the N-terminal domain of BiP onto the backbone of the Hsp70 structure. This indicates that the two proteins have extremely comparable conformations. The amino acid sequence of BiP shares 66% identity and 81% similarity with that of the cytosolic Hsp70 protein. The new fold found in the C-terminal substrate binding domain of *E coli* DnaK is probably shared by all members of the Hsp70 protein family, including BiP. It consists of a compact β-sandwich followed by helical elements. The beta sandwich is arranged in two sheets with four antiparallel beta strands in each, and the peptide is bound in an extended conformation in a channel formed by loops that connect strands of the sandwich. Binding of ATP to the N-terminal domain causes the lid to open, while hydrolysis of ATP to ADP results in a conformational change that causes the lid to close, encapsulating the peptide in the channel (Zhu et al., 1996).

BiP's capacity to distinguish between correctly folded and unfolded structures and to identify a broad range of polypeptides with no discernible sequence similarity is essential to its function as a generic chaperone in the ER lumen. Short synthetic peptides can interact with BiP *in vitro*, stimulating its ATPase activity and changing its oligomeric state upon engagement (Flynn et al., 1989; Blond-Elguindi et al., 1993a, 1993b; Fourie et al., 1994; Knarr et al., 1995; Chevalier et al., 1998; Carlsson & Lazarides, 1983; Leno & Ledford, 1990; Ledford & Leno, 1994). BiP, experiences a quick, reversible, and deactivating post-translational change. To match the supply of active BiP to the need for protein folding within the ER lumen, the post-translational covalent alteration works in tandem with the slower, traditional UPR. Now, it is known that this change is AMPylation of BiP's threonine 518, which was previously thought to be ADP-ribosylation and is catalyzed by the Ficdomain–containing protein (FICD) (Preissler et al., 2015; Perera & Ron, 2023; Freiden, Gaut, & Hendershot, 1992; Hendershot, Ting, & Lee, 1988; Leustek, Toledo, Brot, & Weissbach, 1991; Carlsson & Lazarides, 1983). In complexes with unfolded or unassembled polypeptides, only unaltered, monomeric BiP molecules are detected. Therefore, it would seem that post-translational modification of BiP creates a storage pool of BiP that may be called upon to return to the active form when necessary (Hendershot, Ting, & Lee, 1988; Freiden, Gaut, & Hendershot, 1992).

1.1.1.2.Role of chaperones in Neurodegenerative diseases

Aging-related impaired proteostasis is linked to a number of human illnesses. Highly ordered fibrillar-type protein aggregates are deposited into inclusion bodies in Parkinson's disease and other neurodegenerative diseases. The cellular protein quality-control machinery finds it challenging to handle these amyloid fibril aggregates due to their high stability (Kampinga & Bergink, 2016; Wentink et al., 2019).

Molecular chaperones are crucial for maintaining proteostasis; specifically, the chaperone heat shock protein 70 family is crucial for the folding, disaggregation, and degradation of proteins (Zuiderweg et al., 2017).ER stress has been implicated in the pathophysiology of neurodegenerative diseases in several studies (Ni & Lee, 2007). Huntington's disease and cerebral infarction are two conditions linked to cerebellar ataxia in which BiP is crucial. BiP has been demonstrated to bind to huntingtin aggregates in the Huntington's disease model and inhibit huntingtin aggregation at elevated doses (Muchowski et al, 2000). Consequently, cells that share both reduced BiP availability and mutant huntingtin exhibit hypersensitivity to ER stress (Lajoie & Snapp, 2011).

By repairing misfolded proteins, BiP reduces neuronal cell death brought on by ER stress in Marinesco-Sjögren syndrome, which is typified by cerebellar ataxia (Zhao et al., 2005; Anttonen et al., 2005). BiP expression is downregulated in cells with mutations that cause spinocerebellar ataxia type 17 (Lee et al., 2009). Through BiP-induced neuroprotective pathways, 2-deoxy-D-glucose therapy has been proposed to lessen brain injury in a mouse model of Parkinson's disease. By reprogramming ER stress signaling pathways to prevent apoptotic cell death, BiP overexpression was also reported to reduce α-synuclein cytotoxicity in rat and SH-SY5Y cell models (Gorbatyuk & Gorbatyuk, 2013).

The most common type of dementia is Alzheimer's disease (AD). According to the World Alzheimer Report (2010), there were an estimated 36 million people with AD in 2010. In 2024, this number rose to 55 million people. By 2030, that number is expected to rise to 66 million. Alzheimer's disease is characterized by gradual memory loss. However, despair, delusions, hallucinations, violent behavior, and deterioration of speech and motor abilities are also characteristics of AD (Voisin & Vellas, 2009). Despite these obvious behavioral changes, it might be challenging to accurately identify AD in its early stages. The causes and processes of Alzheimer's disease have been widely studied using rodent models. We still don't fully grasp the molecular processes that cause neurodegeneration, even after

years of rigorous research using these models. Alzheimer's disease has been firmly linked to BiP chaperone dysfunction, among other chaperones (Booth et al., 2015b; Ni & Lee, 2007). BiP's chaperone function may save neurons already loaded with β-amyloid protein, a significant contributor to extracellular senile plaques in Alzheimer's disease (Zhao et al., 2005; Anttonen et al., 2005). BiP inhibitors may therefore work better if administered early in Alzheimer's disease (Booth et al., 2015a).

As shown in Figure 1.1, Hsp70 function, including that of BiP, depends on co-chaperones from the J-domain protein (JDP) and nucleotide exchange factor (NEF) families, which guarantee substrate specificity, control ATPase cycle progression, and govern interactions with other chaperone networks and degradation pathways. It is yet unknown, nevertheless, how different co-chaperones allow the Hsp70 machinery to carry out particular tasks and, in turn, why particular Hsp70 activities call for particular co-chaperone partners. Recently it was shown that produced fibrils of the presynaptic protein α-synuclein, which is linked to Parkinson's disease, can be dissolved *in vitro* by the human Hsp70 chaperone and its co-chaperones DNAJB1 and Hsp110 (Figure 1.3) (Duennwald, Echeverria, & Shorter, 2012; Gao et al., 2015). DNAJB1 directs Hsp70 to fibrils and uses multivalent contacts to detect the oligomeric form of α-synuclein. Instead of the amyloid core itself, Hsp70 and DNAJB1 engage with the fibril via the exposed, flexible amino and carboxy termini of α-synuclein. By enabling the loading of many Hsp70 molecules in a densely packed configuration at the fibril surface—perfect for the production of "entropic pulling" forces—DNAJB1 and Hsp110 work in concert to significantly speed up α-synuclein disaggregation. Beyond the traditional substrate targeting and recycling roles that are ascribed to these Hsp70 co-chaperones, DNAJB1 and Hsp110 work together to facilitate amyloid disaggregation, which actively and crucially contributes to the amyloid substrate's remodeling. Molecular understandings of the necessary conditions for amyloid disintegration could serve as the foundation for novel neurodegenerative treatment approaches (Mayer & Gierasch, 2019; Rosenzweig et al., 2019). By speeding up ADP dissociation, which leads to ATP binding and Hsp70 interconversion to an open conformation, the NEF Hsp110 encourages Hsp70 recycling. High disaggregation efficiencies require thus the presence of a NEF (Bracher & Verghese, 2015).

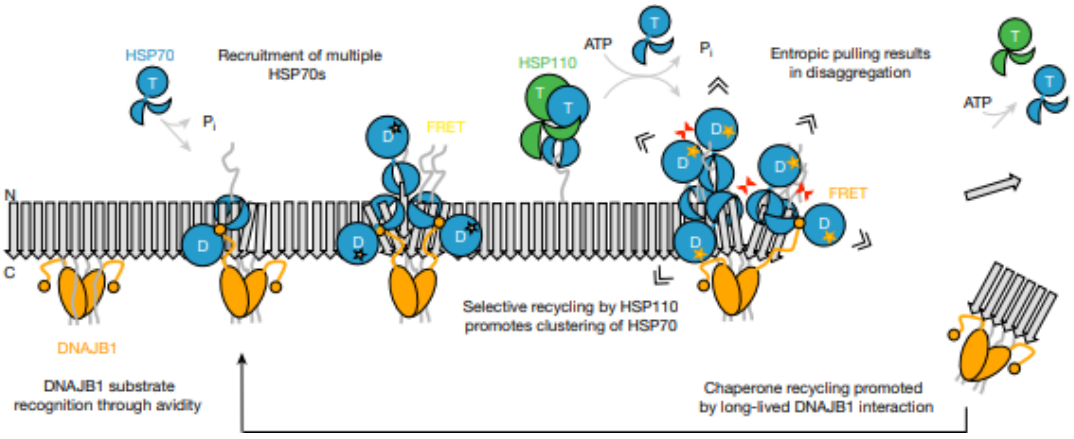


Figure 1.3 Amyloid disaggregation model using the Hsp70 chaperone apparatus. DNAJB1 uses multivalent contacts with the flexible C-terminal tails to specifically identify oligomeric α-synuclein. DNAJB1 uses several recruiting cycles to encourage the high-density loading of Hsp70 on the amyloid surface. Hsp110 preferentially dissociates Hsp70 that have not been grouped by DNAJB1 because of its large molecular mass. In order to provide entropic pulling forces that destabilize the amyloid fibril and cause fibril fragmentation and/or monomer extraction, this activity biases the positioning of Hsp70 to the most productive sites. Throughout multiple Hsp70 chaperoning cycles, DNAJB1 can stay connected and sense the substrate's oligomeric state. Either "T" for ATP or "D" for ADP denotes the nucleotide status of Hsp70 and Hsp110.

Several observations strongly support an entropic-pulling mechanism of amyloid disaggregation activity as described in Figure 1.3:

(1) chaperone engages the fibrillar substrate by exposing disordered regions near the amyloid surface;

(2) disaggregation necessitates Hsp70 crowding on the fibrils;

(3) crowding has an energetic cost, which is partially covered by the affinity of JDP-mediated recruitment of Hsp70 to the fibrils;

(4) Hsp110-specific activity encourages the selective reorganization of HSP70 proteins that are loaded ineffectively, favoring Hsp70 binding at high density where entropic-pulling effects are strongest;

(5) the high molecular mass of the Hsp110 family is a crucial characteristic, confirming that excluded volume effects play a significant role in the disaggregation mechanism (De Los Rios et al., 2006; Sousa et al., 2016; Assenza et al., 2019).

The spread of protein aggregation linked to dementia may be slowed by chaperone-mediated amyloid disaggregation. However, potentially hazardous oligomeric species and fibril

fragments that could act as a template for more protein aggregation could also be produced during the disaggregation process (Pieri et al., 2012; Gao et al., 2015).

The objective of this work was to clone the ORF of BiP from zebrafish into an appropriate vector for expression in *E. coli* to compare its capability to disaggregate proteins with mammalian BiP from golden hamster. By comparing the activity of BiP to disaggregate proteins from two distant vertebrate clades, warm-blooded mammals and cold-blooded fish, we aim to understand how the chaperone machinery from the endoplasmic reticulum has evolved to maintain proteostasis and prevent neurodegeneration.

2. Materials & Methods

2.1 Preparing the competent *E. coli* DH5α cells

The TB Buffer (100 mL) must be prepared sterile, using a 0.2 μm filter before beginning the entire procedure of preparing competent *E. coli* DH5α cells.
The components are listed below:

- 10 mM PIPES
- 15 mM $CaCl_2$
- 250 mM KCl
- 55 mM $MnCl_2$ (only add at the end after adjusting the pH)

The pH was adjusted to 6.7 with KOH.

The protocol to prepare the competent *E. coli* DH5 α cells was the following:

On the 1st Day: The *E.coli* strain needed to be grabbed from the -80 ºC Ark. After plating, let it grow overnight at 37 ºC. Inoculating a colony in 5 mL of LB took place on the 2nd Day and let it grow overnight at 37 ºC (Pre-inoculum). On the 3rd Day: In the afternoon, inoculate 250 mL of LB (in 1 L Erlenmeyer flask) with 1 mL of pre-inoculum and let it grow overnight at 18 ºC. The optical density (OD) was measured at 600 nm in a sample diluted 1:10 with water. On Day 4: Keep measuring the OD at 600 nm (sample diluted 1:10 with water), until an OD of 0.4 was obtained.

In the afternoon, the following steps were taken:

1. The Erlenmeyer flask was kept on ice for 10 minutes.
2. The culture was transferred to 5 sterile falcons (50 mL each) and centrifuged at 2500 g, at 4 ºC, for 10 minutes. (Pre-cool the centrifuge)
3. The pellet of each falcon was resuspended in 16 mL of cold TB Buffer and kept on ice for 10 minutes.
4. Centrifuged the falcons at 2500 g, at 4 ºC, for 10 minutes. (Pre-cool the centrifuge)
5. The pellet was resuspended in each falcon with 4 mL of cold TB Buffer and 280 μL of 7% DMSO was gently added. Falcons then were kept on ice for 10 minutes.
6. Aliquots of 100 uL were made in sterile Eppendorf and flash-freeze in liquid nitrogen or using dry ice.

7. The competent cells needed to be stored at -80 ºC.

2.2. Cloning of the open reading frame (ORF) of zebrafish BiP

Figure 2.1 below shows the vector map containing the ORF of golden hamster BiP where the ORF of zebrafish BiP will be cloned replacing the hamster BiP.

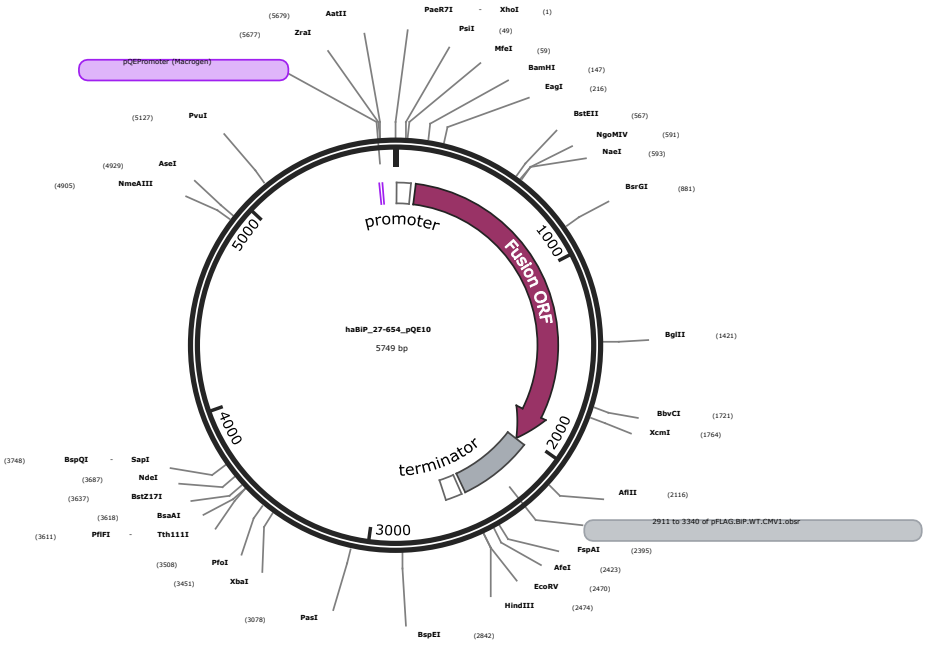


Figure 2.1. Shows the map for the vector haBiP_27-654_pQE10 that we have used as backbone after removing hamster BiP as insert to be replaced by zebrafish BiP. We planned the cloning between BamH1 and AflII that keeps the His tag on the vector for affinity purification.

Amplifying the open reading frame (ORF), and putting it into an appropriate vector for expression are the processes involved in cloning the zebrafish BiP ORF. The first step was to identify the zebrafish BiP (Uniprot Q7SZD3) ORF sequence using the database National Center for Biotechnology Information (NCBI Gene ID 378848). It is very important to make sure to have the correct full-length ORF and note any sequence-specific details like restriction sites or untranslated regions (UTRs). Confirmation of the right ORF was done by translation and comparison with the amino acid sequence from Uniprot reference.

2.3. Extraction of total RNA (tRNA)

Our main objective was to clone the zebrafish BiP ORF, using as original template the totalRNA extracted from Zebrafish. Total RNA was extracted from a caudal zebrafish cell line (SJD:1 from ATCC company) using the E.Z.N.A.® Total RNA Kit I.

The key of the E.Z.N.A.® RNA family of products is that it combines the speed of mini-column spin technology with the reversible binding qualities of the HiBind Matrix, allowing for the processing of one or more samples at once. Phenol/chloroform extractions are less specific, and laborious processes like precipitation with isopropanol or LiCl or CsCl gradient ultracentrifugation are unnecessary.

The Total RNA Kit by E.Z.N.A.® I is capable of purifying up to 100 µg of total RNA from tissue or cultured eukaryotic cells. Typically, a single experiment can process 5–30 mg of tissue or 1 x 10^6–1 x 10^7 eukaryotic cells. Denaturing circumstances that render RNases inactive cause cells or tissue to lyse. Following the homogenization procedure, samples are placed on the HiBind RNA spin column. After a few fast wash steps, cellular debris and other impurities are successfully removed. Finally, high-quality RNA elutes in water treated with Diethylpyrocarbonate (DEPC)-treated and sterile filtered water, which inactivates RNase enzymes in water by covalent modifications of the histidine residues. This kit is used to isolate total RNA larger than 200 nt.

2.4. Synthesis of cDNA

The NZY first-strand cDNA synthesis kit from NZYtech (MB17001) was used to synthesize the cDNA from total RNA. The NZY First-Strand cDNA Synthesis Kit by NZYtech contains all the components required for the synthesis of first-strand cDNA, except for the template RNA, and the very active NZY Reverse Transcriptase. Random hexamers and Oligo $(dT)_{18}$ primers are supplied in separate tubes. The NZY First-Strand cDNA Synthesis Kit was designed to boost sensitivity in subsequent applications and yield large amounts of full-length cDNA products. Ten pg to five µg of total RNA can be used as starting material. NZY Reverse Transcriptase (RNase H minus) and a potent NZY Ribonuclease Inhibitor are both included in this mixture to shield RNA from ribonuclease contamination. After the first-strand cDNA synthesis, RNase H (from *E. coli*) is supplied in a different tube to specifically break down the

RNA template in cDNA-RNA hybrids and facilitate PCR primers to anneal to the cDNA in following PCR assays.

To synthesize cDNA, the following components were mixed into autoclaved PCR tubes.

1. 20 µl of NZYRT 2x master mix
2. 4 µl NZYRT enzyme mix
3. 4 µl 400 ng/ µl RNA
4. 12 µl DEPC H_2O

We mixed the tubes gently, span down and then incubated at 25 °C for 10 minutes. Then raised the temperature to 50 °C and incubated for 30 minutes. Final heating to 85 °C for 5 minutes, inactivated the reaction. After inactivation, the tubes were chilled on ice, and 1 µl of NZY RNase was added to each tube and incubated for 20 minutes at 37 °C. The cDNA was stored at – 20 °C.

2.5. PCR to amplify the ORF of zebrafish BiP

The cDNA, non-diluted and diluted 1:10 and 1:100 in autoclaved water, was used as template for the PCR to amplify the ORF of zebrafish BiP. We made positive and control tubes as well. In the positive tubes we poured 32.5 µl of autoclaved water, 10 µl 5x reaction buffer, 1 µl of 10 mM dNTPs, 2.5 µl of 10 µM forward primer, 2.5 µl of 10 µM reverse primer, 1 µl of the cDNA template (1, 1:10 and 1:100 dilutions) and 0.5 µl of supreme DNA polymerase enzyme lastly to avoid non-specific random amplification. For the negative control instead of the template, an extra 1 µl of autoclaved water was added into the PCR tube.
The following primers were used to amplify ORF of zebrafish BiP using cDNA as template:
FW primer (Bam HI)
5′- ttaggatcc t gtt ggg aca gtg att ggg atc –3′

Rev primer (AflII/BspTI)
5′- gta cat cttaag cta cag ctc gtc ctt ctc ttc -3′

Adding 4-6 extra nucleotides upstream of restriction sites can enhance enzymatic efficiency.

The PCR cycles were set for the following parameters:
Starting the amplification with HOT START DNA polymerase with a temperature of 96 °C for 4 minutes. Then, 30 cycles were established, each consisting of denaturation at 96 °C for 30 seconds, annealing at three different temperatures (52°C, 55°C, 58°C) for 30 seconds, and

extending at 72°C for 60 seconds. A final step of extension was carried out at 72°C for 10 minutes. After the PCR reaction, 5 µl of 10x green buffer were added and samples were run in an agarose gel 1%.

The preparation of the agarose gel was as follows: 0.5 g agarose powder, 5 ml of TAE Buffer 10x and 45 ml of autoclaved water. The recipe for 1 liter of the 10 x TAE Buffer: 48.5 g tris, 11.4 mL of glacial acetic acid and 20 ml of 0.5 M EDTA (pH 8.0). We heated the agarose mixture in a microwave, until it was boiling, then we cooled it down, and added Green Safe Premium from NZYtech. When GreenSafe stain binds to DNA or RNA, it emitted green fluorescence. It exhibits one strong excitation peak at 490 nm and two secondary fluorescence excitation peaks at about 270 and 290 nm. The samples then were loaded into the agarose gel and the gel ran for 2 hours at 50 VOLT to confirm the PCR product size.

Extracting the amplicon from agarose gel

The amplicon of the right size was extracted from the gel using the kit (NZYGelPure) from NZYtech. The NZYGelpure kit is intended for direct purification of PCR products as well as DNA purification from TAE/TBE agarose gels. DNA fragments ranging in size from 50 bp to 20 kb can be purified using the kit. The NZYGelpure purification kit's silica-gel-based membrane may selectively adsorb up to 20 µg of DNA fragments. Impurities such as soluble agarose, nucleotides, oligos (less than 30 mers), primer dimers, enzymes, mineral oil, and others are washed away because they are unable to attach to the membrane. Following elution from the column, DNA fragments are ready for use in subsequent procedures without additional processing.

Procedure for DNA Purification from Agarose Gels:

All centrifugations should be carried out at room temperature in a table-top micro centrifuge at >12000 x g

1. The DNA fragment from the gel was excised with a clean, sharp scalpel. After weighting the gel slice, it was transferred to a 1.5 mL microcentrifuge tube.
2. 300 µL of Binding Buffer was added for each 100 mg of gel weight.
3. Incubation at 55-60 °C for 5-10 minutes and the tubes were occasionally shake until agarose was completely dissolved.

4. After making sure that the colour of the mixture was yellow, the mixture was loaded into theNZYtech spin column, which was placed into a Collection tube (2 mL). Centrifugation was the next step for 30 s to 1 minute, then flow-through was discarded.
5. 500 µL of Wash Buffer was added then centrifuged for 30s to 1 minute. Flow-through was discarded.
6. 600 µL of Wash Buffer was added, then centrifuged for 30s to 1 minute. The flow-through was discarded.
7. To dry NZYtech spin membrane of residual ethanol, centrifugation for 1 minute was essential.
8. The NZYtech spin column was then placed into a clean 1.5 mL microcentrifuge tube. 50 µL of ultrapure water was added to the centre of the column and incubated at room temperature for 1 minute. The DNA was eluted by centrifugation for 1 minute.
9. Store the purified DNA at -20 °C

2.6 Nested PCR to amplify the ORF of zebrafish BiP

To obtain a significant amount of the ORF of zebrafish BiP and confirm the identity of the band extracted from the PCR that used cDNA as template, a nested PCR was carried out using the band extracted in section 2.5 as template. This product was then diluted 1:10 and 1:100 to be used as a template in the nested PCR.

The PCR cycles were the following:

1. Hot start 96 °C, 4 minutes
2. Denaturation 96 °C, 30 seconds
3. Annealing 58 °C, 30 seconds
4. Extension 72 °C, 60 seconds
5. Final extension 72 °C, 10 minutes

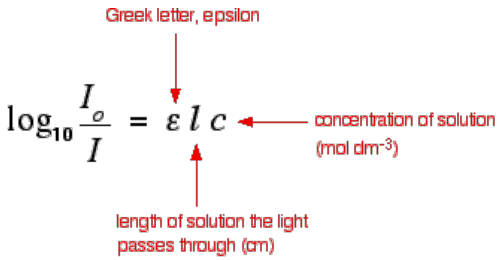
   30 cycles from step 2 to 4.

The amplicons with the right size were extracted from the gel as described in section **2.5.**

2.7. Double digestion of the backbone and insert

2.7.1. Double digestion of Bip vector

The backbone (plasmid haBiP_27-564_pQE10, Figure 2.1) and the ORF of zebrafish BiP used as insert were double digested with BamHI and Bspt restriction enzymes as described:
We digested the hamster BiP vector with BspTI from Thermo, and BamHI from NZYtech. Firstly, a single digestion was carried out using 1 µl of BamH1, 1 µl of Buffer Tango, 0.5 µl of the Bip vector and 7.5 µl of water. For the next single digestion 1 µl of BamH1, 1 µl of Speedy BamH1 Buffer, 0.5 µl of the Bip vector and 7.5 µl of water was used. Two other single digestions with Bspt1 with buffer Tango and buffer O were performed.
Double digestion with both enzymes Bspt1 and BamH1 and buffer Tango was also performed. For this mixture 1µl of each enzyme, 4 µl of Tango Buffer, 5 µl of BiP vector and 9 µl of water was used. After putting together the mixtures, the tubes were spun down and incubated at 37 °C for 5 hours. The inactivation required 20 minutes at 65 °C. Following these steps, green buffer was added to each tube and an agarose gel was run. For the single digestions, a band of 5.7 kbp was expected, for the double digestions, two bands, one with 1.97 kbp, and the other one with 3.8 kbp were expected. After testing the restriction endonucleases and the buffers, we double digested the BiP vector to get the backbone for the ligation.
To purify the double digested backbone, the right band from the gel were cut, and the NZY GELPURE cleaning kit was used to extract it.
Using NanoDrop One, the concentration for the double digested backbone was measured. First a tiny volume of the sample (1–2 µL) is placed directly on a measurement pedestal. A second pedestal (or arm) closes to form a liquid column, enabling the light to pass through the sample. A xenon flash lamp covers a wide wavelength range (from 190 to 850 nm) and acts as a light source. A photodetector measures the intensity of the light that passes through the sample. The Beer-Lambert Law enables the concentration (C) to be determined:



The constant ϵ is called molar absorptivity or molar extinction coefficient and is a measure of the probability of the electronic transition. For nucleic acids, the absorbance at 260 nm (A260) is used to calculate the concentration, as DNA and RNA absorb UV light strongly at this wavelength due to the presence of nucleotide bases. The measured concentration for the double digested backbone was 12 ng/µl.

2.7.2. Double digestion of the PCR product

To create the insert, the PCR product needed to be double digested. We have used the nested PCR product obtained at 58 °C with the template diluted 1:10 (section 2.6).
BspT1 and BamH1 in 2x Tango buffer were used for double digestion, the tubes were gently mixed, spun down, incubated for 16 hours at 37 °C, then inactivated for 20 minutes at 80 °C. Following the double digestion, the PCR product needed to be cleaned with the NZYGELPURE kit, described in section **2.5.** Then the concentration of the cleaned PCR product was also measured with NanoDrop One as 23 ng/µl.

2.8 Ligation and transformation of *E. coli* DH5α

2.8.1. Ligation

By using DNA ligase, restriction endonuclease-treated DNA fragments can be covalently recombined. The enzyme synthesizes two phosphodiester bonds using ATP. As soon as we had our backbone and insert double digested, we could carry on and start the ligation. 1 µl of Speedy ligase from NZYTech, 5 µl of 4x Speedy Buffer, 3 µl of the backbone, 10 µl of the insert and 1 µl of water was used for the ligation. The whole procedure was prepared on ice. Before using the speedy buffer, since it was very viscous, the buffer needed to be vortexed. The vector and the insert needed to be previously heated at a temperature of 95 °C, for 4 minutes. After adding the buffer, insert the vector into the autoclaved water, the tubes were gently mixed, then spun down. The mixtures were incubated at room temperature for 15 minutes, and then everything was set to perform the transformation.

2.8.2. Transformation

We have used competent DH5α cells for the transformation, which were stored at -80 °C until use. We added the plasmid to the competent cells, and then chilled the tubes for 20 minutes on ice. Then a step called, “heat shock”, followed this step for 45 seconds, at 42 °C. The tubes then were chilled on ice for another 5 minutes. Afterwards 0.7 ml of LB with no antibiotics was added. The tubes were incubated at 37 °C for 1 hour, and then placed into the centrifuge, to spin down the cells. Some media was removed, and then the cells were re-suspended, and plated on antibiotics containing agar plates, and incubated overnight at 37 °C for the colonies to grow. The next morning the colonies were picked up, grown in liquid media with antibiotic and the NZYMiniprep kit was used to extract the plasmid from the positive ligation mix colonies.

2.9. Minipreps of the construct for zebrafish BiP

The NZYMiniprep kit consists of an alkaline lysis of bacterial cells and DNA adsorption onto silica in the presence of high salt. The silica gel preferentially adsorbed the plasmid DNA while washing away other contaminants, such as proteins, salts, nucleotides, and oligos (less than 40 mers).
For the whole procedure all centrifugations were carried out at room temperature in a table-top microcentrifuge at >12000 g.
The steps to apply the NZYMiniprep kit were the following:

1. Cultivate and harvest bacterial cells

Pellet 1-5 mL of an *E. coli* LB culture with antibiotic for 30 s. Discard the supernatant.

2. Cell Lysis

Re-suspend the cell pellet in 250 µL Buffer A1 by vigorous vortexing. After this step, we add 250 µL of Buffer A2 and mixed it. Incubate for 4 minutes at room temperature. Then added 300 µL Buffer A3.

3. Lysate Clarification

Centrifuge for 5-10 min at room temperature, depending on initial culture volume.

4. DNA Binding

After placing NZYtech spin column in a 2 mL collecting tube, load the supernatant from step 3 onto the column. After centrifugation for 1 min at 11,000 g flow-through was discarded.

5. Wash Silica Membrane

After adding 500 µL of Buffer AY onto the column, centrifuge it for 1 min, flow-through was discarded.

Then 600 µL of ethanol containing Buffer A4 was added onto the column, then centrifuged for 1 min, and flow-through was discarded.

6. Dry Silica Membrane

The NZYtech spin column was Re-inserted into the empty 2 mL collecting tube and centrifuged for 2 min.

7. Elute Highly Pure DNA

The dried NZYtech spin column was placed into a clean 1.5 mL microcentrifuge tube and in order to get a higher concentrated miniprep, only 30 µL of autoclaved water was added into the column. After 1 min of incubation at room temperature, the eppendorf was centrifuged for 1 min. The purified DNA needed to be stored at -20 °C.

Afterwards all the concentrations were measured with NanoDrop One, and the samples were sent to sequencing.

The concentrations were the following:

ZFBip-1 = 64 ng/ µl
ZFBip-2 = 75 ng/ µl
ZFBip-3 = 56 ng/ µl
ZFBip-4 = 44 ng/ µl

The constructs were sent to STABvida for sequencing using the universal primers:

Universal primers:

Fw pQEPromoter: 5′-CCGAAAAGTGCCACCTG-3′
Rv pQE-R: 5′-GTTCTGAGGTCATTACTGG-3′

2.10. Expression and purification of golden hamster (*Mesocricetus auratus*) BiP (haBiP)

2.10.1Transformation of *E. coli* BL21 (DE3)

Tubes containing 100 µL of BL21 (DE3) competent *E. coli* cells prepared as described in section 2.1 were thawed on ice. Then, 1 µL of plasmid (haBiP_27-654_pQE10, see section 2.2) containing between 1 and 100 ng of DNA was added to the cell suspension, and the tube was kept on ice for 20 minutes. The transformation was performed as mentioned in section **2.8.2.** The cells were then incubated at 37 ºC for 60 minutes, with shaking at 300 rpm. After the incubation period, the cell suspension was centrifuged at 2300 *g* at room temperature. 450 µL of the supernatant was discarded and the cells were resuspended in the remaining 250 µL, before being plated in plates containing LB agar medium supplemented with 50 $\mu g.mL^{-1}$ of ampicillin and left to incubate overnight at 37°C.

2.10.2 Hamster BiP (haBiP) Expression

A single colony from previously transformed BL21(DE3) competent *E. coli* cells was used to inoculate 50 mL of LB medium containing ampicillin, grown overnight at 37 °C with an agitation of 150 rpm. Large-scale growth cultures were initiated by the addition of 5 mL of the starter culture into 500 mL of LB medium containing ampicillin and were allowed to grow at 37ºC until an OD600 of 0.6 to 0.8 was achieved. The expression of haBiP was induced by the addition of IPTG to a final concentration of 500 µM. The cultures were incubated at 18ºC overnight with 100 rpm agitation. The cells were then centrifuged at 6000 *g* at a temperature

of 4ºC, resuspended in HK buffer (50 mM HEPES, 75 mM KCl, pH = 7.4) with 40 mM Imidazole and stored at -20ºC.

2.10.3 Hamster BiP Purification

Resuspended cells were thawed on ice and lysed by sonication with an amplitude of 60%, a pulse of 15 seconds on and 20 seconds off, in the presence of DNase, RNase, and proteases inhibitors. The lysed cells were then centrifuged at 12,000 *g* at 4ºC for one hour. The supernatant was collected and filtered using a 0.45 µm filter before being loaded onto a 5 mL HisTrap HP column (Cytiva). The flow-through was collected with 50 mL of buffer HK with 40 mM Imidazole (pH = 7.4). The protein was then eluted by a gradient from HK buffer with 40 mM imidazole to HK with 500 mM imidazole over a volume of 50 mL. The eluted protein was then dialyzed against HK buffer (pH = 7.4) using a Snakeskin membrane (Thermofisher) with a molecular weight cutoff of 10000 Da. The concentration of BiP was assessed through nanodrop (ε at 280 nm for BiP is 29005 $M^{-1}cm^{-1}$) reaching a concentration of 257 µM, and the purity of the eluted protein was evaluated by a 12% sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-Page).

SDS-Page

The gel was prepared as described in the table below:

Table 1. Volumes of the different reagents used to prepare an SDS-Page

	Stacking 4 % (mL)	**Resolving 12% (mL)**
40 % Acrylamide/Bis Solution 29:1	0.69	7.5
0.5 M Tris-HCl pH 6.8 + 0.4% SDS	0.68	-
1.5 M Tris-HCl pH 8.8 +0.4 % SDS	-	6.25
Water	4.44	11.15
10 % APS	0.025	0.1
TEMED	0.005	0.01
Total	6.8	25

3. Results and Discussion

3.1. Cloning the ORF of zebrafish BiP

3.1.1. PCR to amplify the ORF of zebrafish BiP

The cDNA was synthesized from total RNA as described in section 2.4. The ORF of zebrafish BiP was obtained at NCBI (gene ID 378848 and NCBI Reference Sequence: NM_213058.1):

1 at gcggttgctt tgcctgtttt tgctggtggc cggcagcgtg
43 tttgccgaag aggacgataa gaaggagagt gttgggacag tgattgggat cgacctcggg
103 accacatact cctgtgttgg agtctacaag aatggccgtg ttgagattat tgccaatgac
163 cagggaaacc gcatcactcc gtcatacgtg gcctttacca ctgaaggaga gcggctcatc
223 ggagatgctg cgaagaacca gctcacatcc aaccctgaaa acactgtgtt tgacgccaag
283 aggctgatcg gacgcacatg gggcgactct tctgtgcagc aggacatcaa atacttcccc
343 tttaaggtga tcgagaagaa aaacaagcct cacatccagc tggacatcgg ctctggtcag
403 atgaagacgt ttgcaccgga ggaaatttcc gccatggttt tgaccaagat gaaggaaacc
463 gcagaggctt atctgggaaa gaaggtcact catgctgtgg tcaccgttcc tgcttatttc
523 aacgatgctc agcgtcaggc cactaaagat gctggaacca ttgctgggct gaatgtcatg
583 aggatcatca atgagcctac ggcggctgcc attgcatacg gtctggacaa gagggacgga
643 gagaaaaaca tcctggtgtt cgatctgggt ggtggcacct ttgacgtgtc tctgctgacc
703 atcgataacg gcgtgtttga agtggtggcc acaaacggag acactcacct gggcggagaa
763 gacttcgacc agcgcgtcat ggagcacttc atcaagctgt acaagaagaa gacgggcaaa
823 gacgtgcgca aagacaaccg cgccgtgcag aagctgcgca gagaggtgga gaaggctaag
883 agagcgctgt ctgcccagca tcaggcccgc atcgagatcg agtccttctt tgaaggagaa
943 gatttctctg agactctcac cagagccaag tttgaggagc tcaacatgga cttgttccgc
1003 tccactatga agccggttca gaaggttctg gaggactctg acccgaagaa gccagatatc
1063 gatgagatcg tgctggtcgg cggctccact cgtatcccga agatccagca gctggtgaag
1123 gagttcttca acggaaaaga gccgtccaga ggaatcaacc ctgacgaggc cgtggcgtac
1183 ggagctgctg tccaggctgg agtcctgtcc ggagaggagg agaccggtga tctggttctt
1243 ctggacgtgt gtccgctgac tctgggcatt gagactgttg gaggagtgat gaccaaactc
1303 attcccagaa acactgttgt tcccaccaag aaatcccaga tcttctccac tgcttccgac
1363 aaccagccca ccgtcactat caaagtttat gagggcgagc gtcccctgac caaagacaac

1423 catctgctgg gcacctttga cctgacaggc atccctccag cacctcgtgg tgtcccacag
1483 atcgaggtaa ctttcgagat cgacgtcaac ggcatcctgc gcgtcaccgc cgaagacaaa
1543 ggcaccggaa acaaaaacaa gatcaccatt accaacgacc agaaccggct gacccctgag
1603 gacatcgaga gaatggtgaa cgaagccgag agattcgctg atgaggacaa gaaactgaag
1663 gagagaatcg acagccgcaa tgaattggag agctacgcct attccctgaa gaaccagatc
1723 ggggataaag agaaattagg cggaaagtta tcctctgaag acaaggaggc catcgagaag
1783 gcagtggagg agaagatcga gtggctggag gcgcatcagg acgccgatct ggaggaattc
1843 caggccaaaa agaaggagct ggaggaggtg gtgcagccca tcgtcagcaa actgtacggc
1903 agtgcgggag gaccaccgcc tgaagaggcc gaagagaagg acgagctgta g

The first 16 codons (48 nucleotides) code for a peptide signal that directs the newly synthesized protein to the endoplasmic reticulum (also called secretory pathway) and therefore are useless to express the protein in *E coli,* which has no secretory pathway. The next 8 codons (24 nucleotides) code for a highly charged sequence (GluGluAspAspLysLysGluSer) that was not included as part of the golden hamster BiP previously cloned in pQE10 vector to be expressed in *E coli* (a kind gift from David Ron, University of Cambridge). Therefore, since we want to compare the zebrafish BiP with mammalian BiP from golden hamster to analyse how BiP has evolved in its chaperone function, we have decided to start the amplification at the codons gtt ggg aca..... coding for the residues ValGlyThr.

The PCR to amplify the ORF of zebrafish BiP starting on Val25 was carried out using the appropriate primers (see section 2.5) and cDNA diluted 1:10 and 1:100 as a template. An agarose gel was run to check the size of the bands (Figure 3.1).

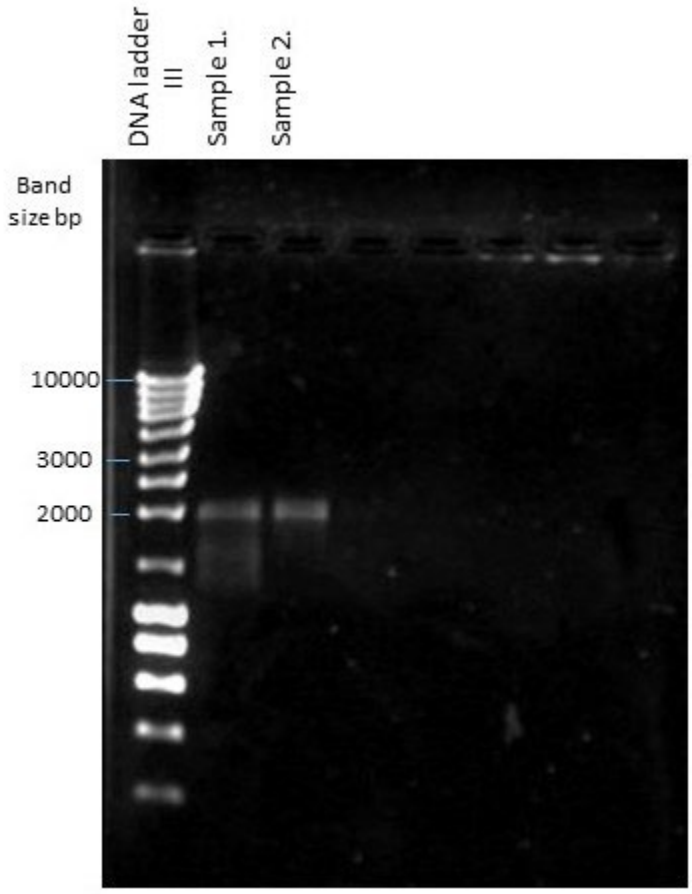


Figure 3.1 Agarose gel (1%) showing the amplicons from the PCR using cDNA as template using . Sample 1 is the dilution of 1:10 of the cDNA used as template, Sample 2 is the dilution of 1:100 of the cDNA used as template We were expecting a band with a size of 1881 bp.

An amplicon of 1881 bp is expected after discounting the first 72 nucleotides from the ORF of zebrafish shown above and indeed a band of around 2000 bp was observed in the gel indicating that amplification was successful.

3.1.2. Nested PCR to amplify the ORF of zebrafish BiP

To obtain a significant amount of the ORF of zebrafish BiP and confirm the identity of the band extracted from the PCR that used cDNA as template, a nested PCR was carried out using the bands extracted from the gel shown in figure 3.1 The gel in Figure 3.2 shows the amplicons from the nested PCR with a strong band around 2000 bp as expected. To get a more concentrated insert, the nested PCR reaction was repeated using as template the band extracted from the gel after dilution 1:100.

The amplicons with the right size were extracted from the gel and sent to sequencing to confirm the identity of the amplicon**.**

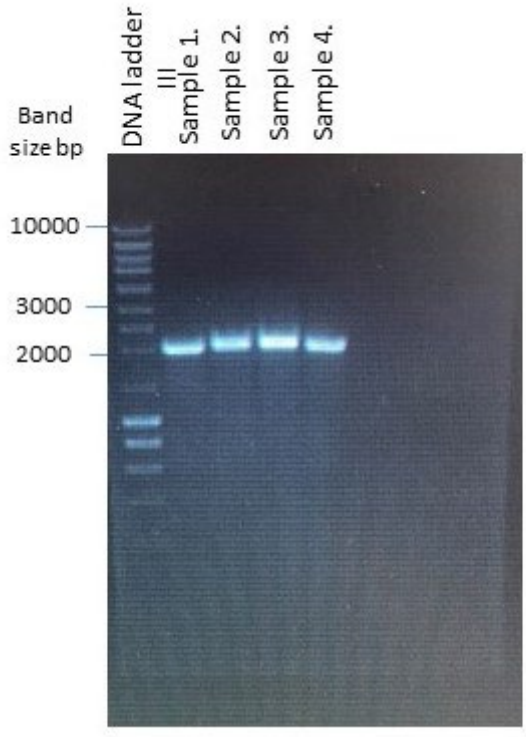


Figure 3.2 Nested PCR product where Sample1, Sample 2 are PCR products using cDNA template diluted 1:10; Sample 3, Sample 4 are PCR products using cDNA template diluted 1:100.

The sequence of the PCR product obtained with the forward primer used for amplification is compared with the expected zebrafish ORF from Val 25:

Alignment of Sequence_1: [ORF zebrafish from Val25.xdna] with Sequence_2: [PCR product FW.xdna]

Similarity: 1104/1107 (99.73 %)

```
Seq_1   1     Gttgggacagtgattgggatcgacctcgggaccacatactcctgtgttggagtctacaag  60
                                                          ||||||||||||||
Seq_2   1     ----------------------------------------------TTGGAGTCTACAAG  14

Seq_1   61    aatggccgtgttgagattattgccaatgaccagggaaaccgcatcactccgtcatacgtg  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2   15    AATGGCCGTGTTGAGATTATTGCCAATGACCAGGGAAACCGCATCACTCCGTCATACGTA  74
```

```
Seq_1  121   gcctttaccactgaaggagagcggctcatcggagatgctgcgaagaaccagctcacatcc  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  75    GCCTTTACCACTGAAGGAGAGCGGCTCATCGGAGATGCTGCGAAGAACCAGCTCACATCC  134

Seq_1  181   aaccctgaaaacactgtgtttgacgccaagaggctgatcggacgcacatggggcgactct  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  135   AACCCTGAAAACACTGTGTTTGACGCCAAGAGGCTGATCGGACGCACATGGGGCGACTCT  194
Seq_1  241   tctgtgcagcaggacatcaaatacttccccttttaaggtgatcgagaagaaaaacaagcct  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  195   TCTGTGCAGCAGGACATCAAATACTTCCCCTTTAAGGTGATCGAGAAGAAAAACAAGCCT  254

Seq_1  301   cacatccagctggacatcggctctggtcagatgaagacgtttgcaccggaggaaatttcc  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  255   CACATCCAGCTGGACATCGGCTCTGGTCAGATGAAGACGTTTGCACCGGAGGAAATTTCC  314

Seq_1  361   gccatggttttgaccaagatgaaggaaaccgcagaggcttatctgggaaagaaggtcact  420
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Seq_2  315   GCCATGGTTTTGACCAAGATGAAGGAAACCGCAGAGGCTTACCTGGGAAAGAAGGTCACT  374

Seq_1  421   catgctgtggtcaccgttcctgcttatttcaacgatgctcagcgtcaggccactaaagat  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  375   CATGCTGTGGTCACCGTTCCTGCTTATTTCAACGATGCTCAGCGTCAGGCCACTAAAGAT  434

Seq_1  481   gctggaaccattgctgggctgaatgtcatgaggatcatcaatgagcctacggcggctgcc  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  435   GCTGGAACCATTGCTGGGCTGAATGTCATGAGGATCATCAATGAGCCTACGGCGGCTGCC  494

Seq_1  541   attgcatacggtctggacaagagggacggagagaaaaacatcctggtgttcgatctgggt  600
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  495   ATCGCATACGGTCTGGACAAGAGGGACGGAGAGAAAAACATCCTGGTGTTCGATCTGGGT  554

Seq_1  601   ggtggcacctttgacgtgtctctgctgaccatcgataacggcgtgtttgaagtggtggcc  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  555   GGTGGCACCTTTGACGTGTCTCTGCTGACCATCGATAACGGCGTGTTTGAAGTGGTGGCC  614
```

```
Seq_1  661   acaaacggagacactcacctgggcggagaagacttcgaccagcgcgtcatggagcacttc  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  615   ACAAACGGAGACACTCACCTGGGCGGAGAAGACTTCGACCAGCGCGTCATGGAGCACTTC  674

Seq_1  721   atcaagctgtacaagaagaagacgggcaaagacgtgcgcaaagacaaccgcgccgtgcag  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  675   ATCAAGCTGTACAAGAAGAAGACGGGCAAAGACGTGCGCAAAGACAACCGCGCCGTGCAG  734

Seq_1  781   aagctgcgcagagaggtggagaaggctaagagagcgctgtctgcccagcatcaggcccgc  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  735   AAGCTGCGCAGAGAGGTGGAGAAGGCTAAGAGAGCGCTGTCTGCCCAGCATCAGGCCCGC  794

Seq_1  841   atcgagatcgagtccttctttgaaggagaagatttctctgagactctcaccagagccaag  900
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Seq_2  795   ATCGAGATCGAGTCCTTCTTTGAGGGAGAAGATTTCTCTGAGACTCTCACCAGAGCCAAG  854

Seq_1  901   tttgaggagctcaacatggacttgttccgctccactatgaagccggttcagaaggttctg  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  855   TTTGAGGAGCTCAACATGGACTTGTTCCGCTCCACTATGAAGCCGGTTCAGAAGGTTCTG  914

Seq_1  961   gaggactctgaccccgaagaagccagatatcgatgagatcgtgctggtcggcggctccact  1020
             |||||||||||||| ||||||||||||||| |||||||||||| ||||||||||||||||||
Seq_2  915   GAGGACTCTGACCTGAAGAAGCCAGATATTGATGAGATCGTGTTGGTCGGCGGCTCCACT  974

Seq_1  1021  cgtatcccgaagatccagcagctggtgaaggagttcttcaacggaaaagagccgtccaga  1080
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Seq_2  975   CGTATCCCGAAGATCCAGCAGCTGGTGAAGGAGTTCTTCAATGGAAAAGAGCCGTCCAGA  1034

Seq_1  1081  ggaatcaaccctgacgaggccgtggcgtacggagctgctgtccaggctggagtcctgtcc  1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  1035  GGAATCAACCCTGACGAGGCCGTGGCGTACGGAGCTGCTGTCCAGGCTGGAGTCCTGTCC  1094

Seq_1  1141  ggagaggaggagaccggtgatctggttcttctggacgtgtgtccgctgactctgggcatt  1200
             |||
Seq_2  1095  GGA----------------------------------------------------------  1097
```

There are 5 silent mutations (see codons in green) but there is one missense mutation from a Proline to a Leucine (see codon in red). The sequence was checked until nucleotide 1095 (see PDF file with the chromatogram in annex B), so translation was checked from residue 41 to 381 in 627.

Below is the comparison between zebrafish ORF from Val 25 and PCR product sequenced from reverse primer (reverse complementary sequence from nucleotide sequence):

Alignment of Sequence_1: [ORF zebrafish from Val25.xdna] with Sequence_2: [PCR product Rv-rev-comp.xdna]

Similarity: 1146/1148 (99.83 %)

```
Seq_1  661   acaaacggagacactcacctgggcggagaagacttcgaccagcgcgtcatggagcacttc  720
                                                      ||||||||||||||||
Seq_2  2     ---------------------------------------------GTCATGGAGCACTTC  16

Seq_1  721   atcaagctgtacaagaagaagacgggcaaagacgtgcgcaaagacaaccgcgccgtgcag  780
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  17    ATCAAGCTTTACAAGAAGAAGACGGGCAAAGACGTGCGCAAAGACAACCGCGCCGTGCAG  76
Seq_1  781   aagctgcgcagagaggtggagaaggctaagagagcgctgtctgcccagcatcaggcccgc  840
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Seq_2  77    AAGCTGCGCAGAGAG-TGGAGAAGGCTAAGAGAGCGCTGTCTGCCCAGCATCAGGCCCGC  135

Seq_1  841   atcgagatcgagtccttctttgaaggagaagatttctctgagactctcaccagagccaag  900
             ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Seq_2  136   ATCGAGATCGAGTCCTTCTTTGAGGGAGAAGATTTCTCTGAGACTCTCACCAGAGCCAAG  195

Seq_1  901   tttgaggagctcaacatggacttgttccgctccactatgaagccggttcagaaggttctg  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  196   TTTGAGGAGCTCAACATGGACTTGTTCCGCTCCACTATGAAGCCGGTTCAGAAGGTTCTG  255

Seq_1  961   gaggactctgaccccgaagaagccagatatcgatgagatcgtgctggtcggcggctccact  1020
             |||||||||||||| ||||||||||||||| |||||||||||| ||||||||||||||||||
Seq_2  256   GAGGACTCTGACCTGAAGAAGCCAGATATTGATGAGATCGTGTTGGTCGGCGGCTCCACT  315
```

```
Seq_1  1021  cgtatcccgaagatccagcagctggtgaaggagttcttcaacggaaaagagccgtccaga  1080
             ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Seq_2  316   CGTATCCCGAAGATCCAGCAGCTGGTGAAGGAGTTCTTCAATGGAAAAGAGCCGTCCAGA  375

Seq_1  1081  ggaatcaaccctgacgaggccgtggcgtacggagctgctgtccaggctggagtcctgtcc  1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  376   GGAATCAACCCTGACGAGGCCGTGGCGTACGGAGCTGCTGTCCAGGCTGGAGTCCTGTCC  435

Seq_1  1141  ggagaggaggagaccggtgatctggttcttctggacgtgtgtccgctgactctgggcatt  1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  436   GGAGAGGAGGAGACCGGTGATCTGGTTCTTCTGGACGTGTGTCCGCTGACTCTGGGCATT  495

Seq_1  1201  gagactgttggaggagtgatgaccaaactcattcccagaaacactgttgttcccaccaag  1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  496   GAGACTGTTGGAGGAGTGATGACCAAACTCATTCCCAGAAACACTGTTGTTCCCACCAAG  555

Seq_1  1261  aaatcccagatcttctccactgcttccgacaaccagcccaccgtcactatcaaagtttat  1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  556   AAATCCCAGATCTTCTCCACTGCTTCCGACAACCAGCCCACCGTCACTATCAAAGTTTAT  615

Seq_1  1321  gagggcgagcgtcccctgaccaaagacaaccatctgctgggcacctttgacctgacaggc  1380
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Seq_2  616   GAGGGCGAGCGTCCCCTGACCAAAGACAACCATCTGCTGGGCACCTTTGACCTGACTGGC  675

Seq_1  1381  atccctccagcacctcgtggtgtcccacagatcgaggtaactttcgagatcgacgtcaac  1440
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Seq_2  676   ATCCCTCCAGCACCTCGTGGCGTCCCACAGATCGAGGTAACTTTCGAGATCGACGTCAAC  735

Seq_1  1441  ggcatcctgcgcgtcaccgccgaagacaaaggcaccggaaacaaaaacaagatcaccatt  1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  736   GGCATCCTGCGCGTCACCGCCGAAGACAAAGGCACCGGAAACAAAAACAAGATCACCATT  795

Seq_1  1501  accaacgaccagaaccggctgacccctgaggacatcgagagaatggtgaacgaagccgag  1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  796   ACCAACGACCAGAACCGGCTGACCCCTGAGGACATCGAGAGAATGGTGAACGAAGCCGAG  855
```

```
Seq_1  1561  agattcgctgatgaggacaagaaactgaaggagagaatcgacagccgcaatgaattggag  1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  856   AGATTCGCTGATGAGGACAAGAAACTGAAGGAGAGAATCGACAGCCGCAATGAATTGGAG  915

Seq_1  1621  agctacgcctattccctgaagaaccagatcggggataaagagaaattaggcggaaagtta  1680
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Seq_2  916   AGCTACGCCTATTCCCTGAAGAACCAGATCGGGGATAAAGAGAAATTAGGTGGAAAGTTA  975

Seq_1  1681  tcctctgaagacaaggaggccatcgagaaggcagtggaggagaagatcgagtggctggag  1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  976   TCCTCTGAAGACAAGGAGGCCATCGAGAAGGCAGTGGAGGAGAAGATCGAGTGGCTGGAG  1035

Seq_1  1741  gcgcatcaggacgccgatctggaggaattccaggccaaaaagaaggagctggaggaggtg  1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  1036  GCGCATCAGGACGCCGATCTGGAGGAATTCCAGGCCAAAAAGAAGGAGCTGGAGGAGGTG  1095

Seq_1  1801  gtgcagcccatcgtcagcaaactgtacggcagtgcgggaggaccaccgcctgaagaggcc  1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  1096  GTGCAGCCCATCGTCAGCAAACTGTACGGCAGTGCGGGAGGACCACCGCC----------  1145

Seq_1  1861  gaagagaaggacgagctgtag  1881

Seq_2  1146  ----------------------  1145
```

There are 9 silent mutations (see codons in green) and the same missense mutation from a Proline to a Leucine at position 374 (see codon in red) found in the forward sequencing. The G deletion in nucleotide 92 of the reverse sequence was observed in the chromatogram as a shoulder. The sequence of the PCR product was checked from nucleotide 705 until nucleotide 1849 based on the reverse sequence (see PDF with the chromatogram), so translation was checked from residue 235 to 616.

Joining the information from the forward and reverse sequence, the PCR product was checked from residues 41 to 616 in 627 residues for the full length protein. The beginning and the end of the sequence of the PCR product could not be verified because we used the same primers used in the PCR reaction. Anyway, 92% of the sequence of the PCR product was checked and

only the missense mutation from Pro to Leu at position 334 was found. This is a very strong indication that we have succeeded in the amplification of the ORF of zebrafish BiP.

3.1.2 Backbone

The backbone (haBiP_27-654_pQE10, see Material and Methods, section 2.2) and the ORF of zebrafish BiP used as insert were double digested with BamHI and Bspt1 restriction enzymes as described in section **2.7**

To test the restriction enzymes and respective buffers, four single digestions and one double digestion were performed (Figure 3.3).

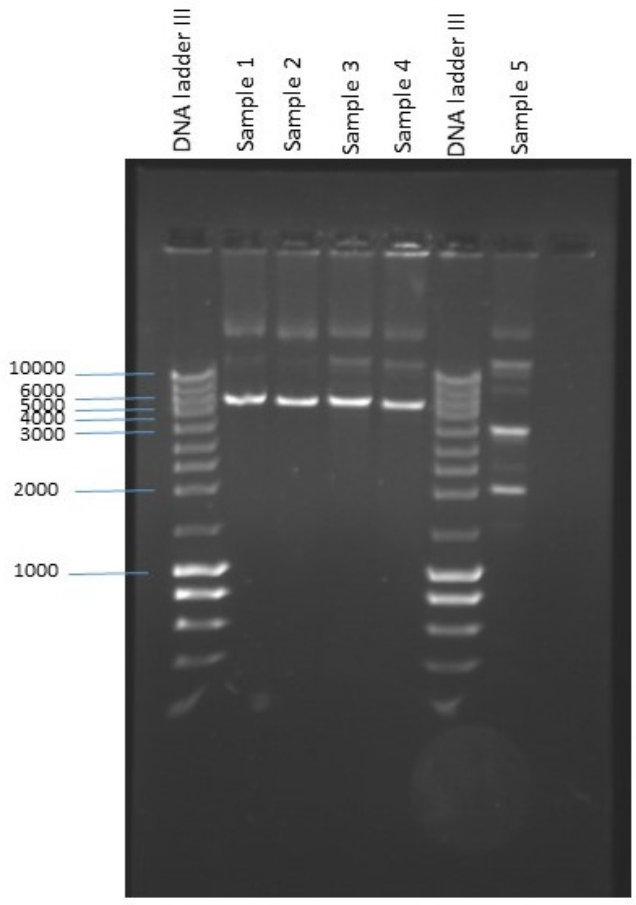


Figure 3.3 Four single, and one double digestion of the pQE10 vector to be used as backbone. Sample 1. Bspt1+ Buffer Tango ; Sample 2. Bspt1+ Buffer O; Sample 3. BamH1+Buffer Tango; Sample 4. BamH1+ Buffer BamH1+; Sample 5. Bspt1 + BamH1 + Buffer Tango. The

size of single digested band shown in samples 1, 2, 3 and 4 is 5749 bp, for sample 5, which is the double digested vector, the size of the first band is 3779 and band two is 1970 bp.

After testing the restriction endonucleases and the respective buffers, the haBiP_27-654_pQE10 vector was double digested with BamH1 and Bspt1 using buffer Tango to remove haBiP. The results are shown in Figure 3.4

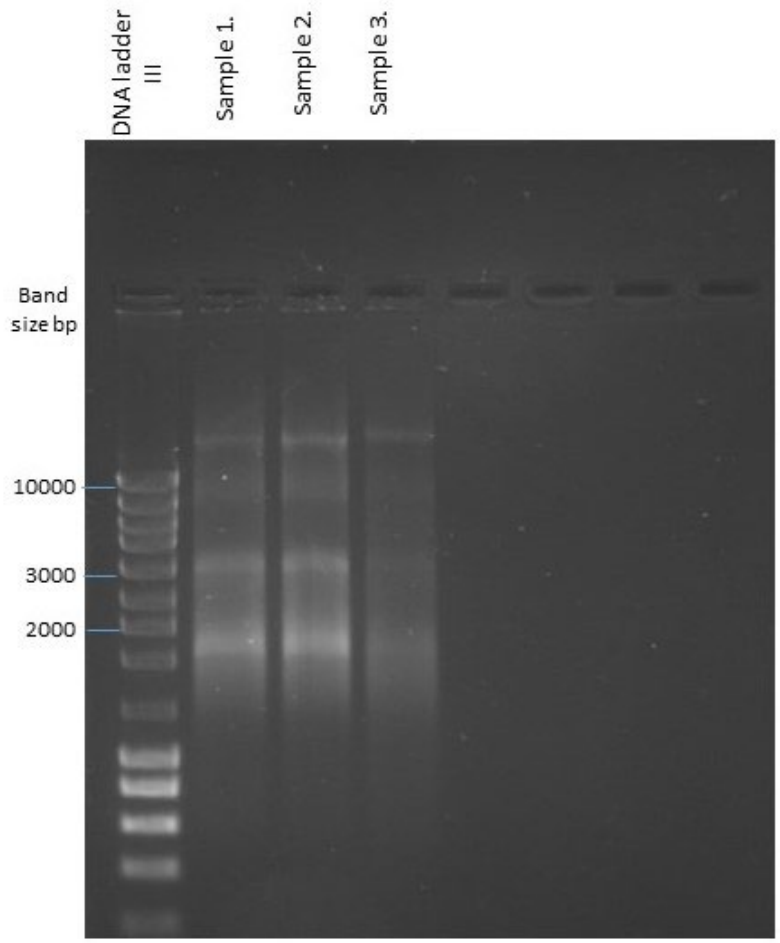


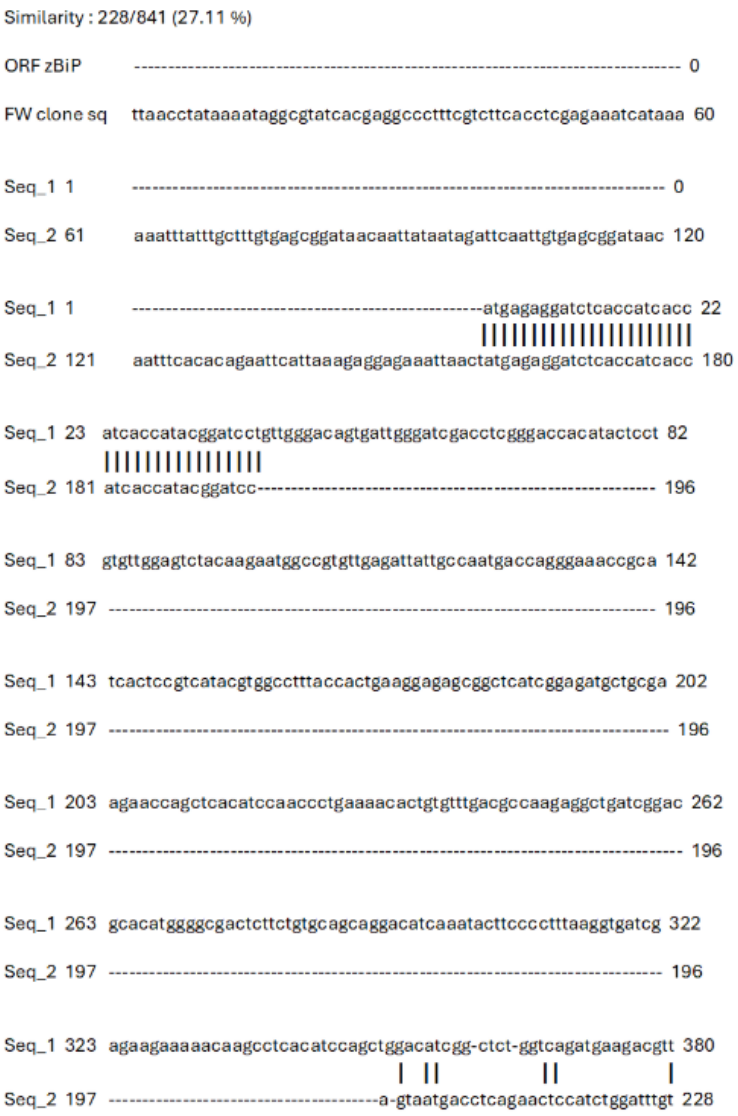
Figure 3.4 Double digested backbone with the restriction enzymes BamH1 and Bspt1 in buffer Tango. Sample 1., Sample 2 and Sample 3 are replicates. As referred in figure 3.3, expected sizes after double digestion are 3779 bp for the backbone and 1970 bp for the insert to be removed.

3.1.3 Ligation and Transformation

The ligation of the ORF of zebrafish BiP after double digestion into the backbone pQE10 after removing hamster BiP was carried out using a ratio insert: backbone 10:1 (v/v). To allow us to identify background colonies that arise from self-ligated backbone, a negative control without insert was carried out. After transformation of competent cells with the ligation mixture and the negative control that contains only the backbone, the cells were plated on agar plates with ampicillin.

Ideally, the negative control should have very few or no colonies, indicating minimal background from the backbone. There were a few colonies in the plates with the full ligation mixture, and minipreps from those colonies were prepared and sent to sequencing using the universal forward primer pQEPromoter as described in section 2.9. Below is a comparison of one miniprep sequence with the sequence expected for the ORF of zebrafish BiP.

Comparing ORF from zebrafish with sequences from pQE10 after ligation of ORF of zebrafish BIP



As can be seen from the sequence comparison above, after the BamHI restriction site (GGATCC) used to ligate the ORF of zebrafish BiP into the pQE10 vector there is no correspondence between the two sequences (see also the PDF file in annex A showing the comparison between the forward sequence from the miniprep and the ORF of zebrafish BiP

from Val 25). As the sequence of the PCR product was verified to be the right sequence (see section 3.1.2), we can conclude that *E coli* cells have recombined the ORF of zebrafish BiP after transformation. This result was unexpected, but E coli can indeed recombine certain DNA sequences (Persky & Lovett, 2008).

3.2.Expression and Purification of hamster BiP

3.2.1. IMAC Purification (Immobilized Metal Affinity Chromatography)

The ORF of hamster BiP cloned into vector haBiP_27-654_pQE10 was expressed in LB medium with ampicillin overnight at 18 °C as described in material and methods (section 2.10). Hamster BiP in this vector contains a polyhistidine tag (His-tag) with six histidine residues at the N-terminal. IMAC is a chromatographic method used to purify proteins containing an His-tag. The resin used for IMAC contains immobilized metal ions, such as Ni^{2+} (nickel), which bind specifically to the histidine residues in the His-tagged proteins. The chromatogram for the purification of hamster BiP is shown in Figure 3.5.

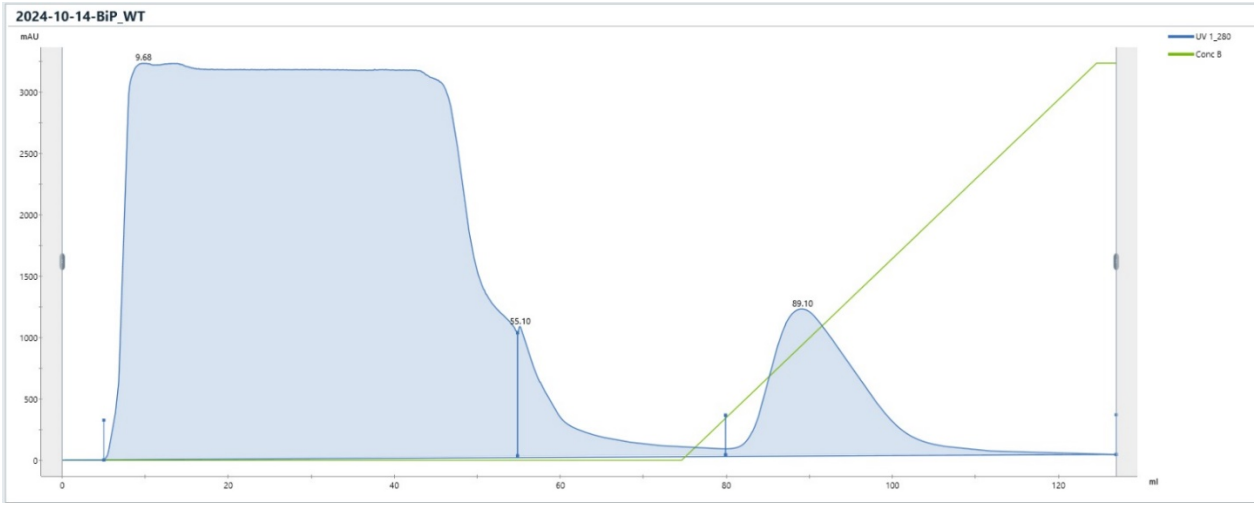


Figure 3.5 Chromatogram for the purification of golden hamster BiP showing (starting at 9.68 mL) the flow-through, which is the first large peak at the beginning of the run and the elution peak corresponding (at 89.10 mL) to the elution of BiP under increasing imidazole concentration.

The flow-through (starting at 9.68 mL) is expected to contain all the proteins from *E. coli* that do not bind to the nickel ions on the column because they do not have an His-tag. The elution peak (peak at 89.10 mL) should correspond to the BiP protein eluted during the gradient that increases the concentration of imidazole (green line in figure 3.5) which competes with the His-tag for the binding to the nickel ions in the column.

To remove the Imidazole, the eluted fractions containing Bip were subjected to dialysis against TR Buffer (50mM Tris-HCl, 200 mM NaCl) overnight using a SnakeSkin dialysis membrane (Thermofisher) with a cutoff of 10,000 Da.

.

3.2.2. SDS page

The purity of the eluted BiP was analyzed by a 12% SDS-PAGE (Figure 3.6)

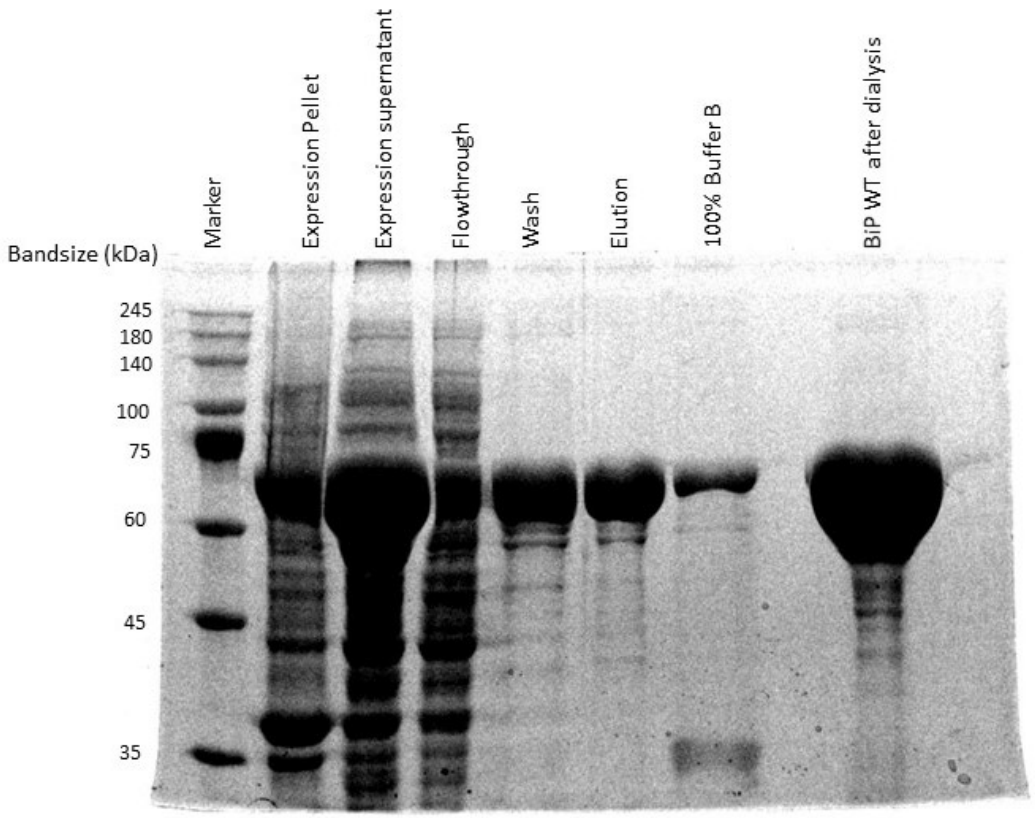


Figure 3.6 SDS Page showing the different fractions from the purification of hamster BiP by IMAC.

The molecular weight of hamster BiP is 71043 Da and the extinction coefficient at 280 nm is 29005 $M^{-1}cm^{-1}$ as calculated by ProtParam from Expasy (https://web.expasy.org/cgi-bin/protparam/protparam). A fraction of BiP is insoluble when expressed in *E coli* as revealed by the strong band at the right size in the lane identified as expression pellet. However, a major fraction is soluble in the supernatant after cell disruption, lane expression supernatant. In the flowthrough, proteins that have no affinity or very little affinity to the nickel ion can be found. Based on the analysis of the wash fraction, the column should be saturated and some weakly bound BiP was washed away previously to the elution gradient. BiP was eluted during the gradient to 500 mM imidazole (lane elution). At 500 mM imidazole (100% buffer B after the gradient) there is still a bit of BiP eluted from the nickel along with a faint band of a significantly smaller protein, which might be a truncated version of BiP based on the strong affinity to nickel ions. After dialysis to remove imidazole, the fraction of eluted BiP was concentrated to around 120 µM (measured based on the extinction coefficient at 280 nm) and analyzed on the gel to reveal a degree of purity over 95% (lane BiP WT after dialysis).

Conclusions

With an expected amplicon size of 1881 bp, the study successfully amplified the BiP ORF from zebrafish using reverse transcriptase and PCR, as validated by nested PCR. 99.73% similarity with the predicted zebrafish ORF was confirmed by sequencing the PCR product, which revealed one missense mutation (proline to leucine at location 334). After the golden hamster BiP sequence was removed, the pQE10 vector was used as the backbone. The backbone preparation was prepared by double digestion with BamHI and BsptI, and the zebrafish BiP ORF was ligated into the vector with a 10:1 insert-to-backbone ratio. Colony screening after transformation into *E. coli* DH5α showed some background from the self-ligated backbone, but minipreps from colonies in the positive plates were sent for sequencing. The zebrafish BiP sequence was recombined by *E. coli*, indicating the necessity of using a recombination-deficient *E. coli* strain for successfully cloning the ORF of zebrafish BiP.

Additionally, golden hamster BiP was successfully expressed in *E. coli* BL21 (DE3) and purified by Immobilized Metal Affinity Chromatography (IMAC). A distinct elution peak in

the chromatogram analyzed by SDS-page showed that the His-tagged BiP had effectively bound to the $Ni2^{+}$ column and was eluted by imidazole with a purity degree over 95%.

These results lay the groundwork for comparing the roles of BiP in zebrafish and mammals, namely in the areas of protein disaggregation and chaperone-mediated neuroprotection. Research on the evolutionary adaptations of BiP in vertebrates will be possible if the recombination in *E. coli* is solved.

In the future, we may be able to find novel strategies to treat proteostasis imbalances by comprehending how BiP interacts with co-chaperones and substrates. Modifying BiP activity may help treat disorders brought on by protein aggregation, which has potential implications for drug discovery. To compare with mammalian orthologs and investigate the precise biochemical and cellular pathways regulated by zebrafish BiP, more research is necessary.
This study provides a solid basis for expanding our knowledge of chaperone biology and devising new approaches to treat neurodegenerative diseases.

References

Aliyu, H.,, Habibu, Anke Neumann, A., & Ochsenreither, K. (2024). Enzymes in molecular biotechnology. In K.-E. Jaeger (Ed.), *Introduction to enzyme technology* (pp. 431–467). Springer International Publishing. https://doi.org/10.1007/978-3-031-42999-6_20

Anttonen, A. K., Mahjneh, I., Hämäläinen, R. H., Lagier-Tourenne, C., Kopra, O., Waris, L., Anttonen, M., Joensuu, T., Kalimo, H., Paetau, A., Tranebjaerg, L., Chaigne, D., Koenig, M., Eeg-Olofsson, O., Udd, B., Somer, M., Somer, H., & Lehesjoki, A. E. (2005). The gene disrupted in Marinesco-Sjögren syndrome encodes SIL1, an HSPA5 cochaperone. *Nature Genetics, 37*(12), 1309–1311. https://doi.org/10.1038/ng1677

Assenza, S., Sassi, A. S., Kellner, R., Schuler, B., De Los Rios, P., & Barducci, A. (2019). Efficient conversion of chemical energy into mechanical work by Hsp70 chaperones. *eLife, 8*, e48491. https://doi.org/10.7554/eLife.48491

Bedows, E. D. (1977). Studies on prokaryotic and eukaryotic polynucleotide ligases (Publication No. 2142/68078) [Doctoral dissertation, University of Illinois at Urbana-Champaign]. Illinois Digital Environment for Access to Learning and Scholarship (IDEALS). http://hdl.handle.net/2142/68078

Beissinger, M., & Buchner, J. (1998). How chaperones fold proteins. *Biological Chemistry, 379*(3), 245–259. https://doi.org/10.1515/BC.1998.033

Bertelsen, E. B., Chang, L., Gestwicki, J. E., & Zuiderweg, E. R. (2009). Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proceedings of the National Academy of Sciences of the United States of America, 106*(21), 8471–8476. https://doi.org/10.1073/pnas.0903503106

Blond-Elguindi, S., Cwirla, S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. R., Sambrook, J. F., & Gething, M. J. (1993). Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. *Cell, 75*(4), 717–728. https://doi.org/10.1016/0092-8674(93)90492-9

Blond-Elguindi, S., Fourie, A. M., Sambrook, J. F., & Gething, M. J. (1993). Peptide-dependent stimulation of the ATPase activity of the molecular chaperone BiP is the result of conversion of oligomers to active monomers. *Journal of Biological Chemistry, 268*(17), 12730–12735. https://doi.org/10.1016/S0021-9258(18)31449-2

Booth, L., Roberts, J. L., Cash, D. R., Tavallai, S., Jean, S., Fidanza, A., Cruz-Luna, T., Siembiba, P., Cycon, K. A., Cornelissen, C. N., & Dent, P. (2015a). GRP78/BiP/HSPA5/DnaK is a universal therapeutic target for human disease. *Journal of Cellular Physiology, 230*(7), 1661–1676. https://doi.org/10.1002/jcp.24919

Booth, L., Roberts, J. L., & Dent, P. (2015b). HSPA5/DnaK may be a useful target for human disease therapies. *DNA and Cell Biology, 34*(3), 153–158. https://doi.org/10.1089/dna.2015.2808

Bracher, A., & Verghese, J. (2015). The nucleotide exchange factors of Hsp70 molecular chaperones. *Frontiers in Molecular Biosciences, 2*, 10. https://doi.org/10.3389/fmolb.2015.00010

Buchberger, A., Theyssen, H., Schröder, H., McCarty, J. S., Virgallita, G., Milkereit, P., Reinstein, J., & Bukau, B. (1995). Nucleotide-induced conformational changes in the ATPase and substrate binding domains of the DnaK chaperone provide evidence for interdomain communication. *Journal of Biological Chemistry, 270*(28), 16903–16910. https://doi.org/10.1074/jbc.270.28.16903

Campion, D., Pottier, C., Nicolas, G., Le Guennec, K., & Rovelet-Lecrux, A. (2016). Alzheimer disease: Modeling an Aβ-centered biological network. *Molecular Psychiatry, 21*(7), 861–871. https://doi.org/10.1038/mp.2016.38

Campion, D., Dumanchin, C., Hannequin, D., Dubois, B., Belliard, S., Puel, M., Thomas-Anterion, C., Michon, A., Martin, C., Charbonnier, F., Raux, G., Camuzat, A., Penet, C., Mesnage, V., Martinez, M., Clerget-Darpoux, F., Brice, A., & Frebourg, T. (1999). Early-onset autosomal dominant Alzheimer disease: Prevalence, genetic heterogeneity, and mutation spectrum. *American Journal of Human Genetics, 65*(3), 664–670. https://doi.org/10.1086/302553

Carlsson, L., & Lazarides, E. (1983). ADP-ribosylation of the Mr 83,000 stress-inducible and glucose-regulated protein in avian and mammalian cells: Modulation by heat shock and glucose starvation. *Proceedings of the National Academy of Sciences of the United States of America, 80*(15), 4664–4668. https://doi.org/10.1073/pnas.80.15.4664

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes|Genomes|Genetics, 1*(3), 171–182. https://doi.org/10.1534/g3.111.000240

Chang, A. C., & Cohen, S. N. (1974). Genome construction between bacterial species in vitro: Replication and expression of Staphylococcus plasmid genes in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America, 71*(4), 1030–1034. https://doi.org/10.1073/pnas.71.4.1030

Chevalier, M., King, L., Wang, C., Gething, M. J., Elguindi, E., & Blond, S. Y. (1998). Substrate binding induces depolymerization of the C-terminal peptide binding domain of murine GRP78/BiP. *Journal of Biological Chemistry, 273*(41), 26827–26835. https://doi.org/10.1074/jbc.273.41.26827

Crouzier, L., Richard, E. M., Sourbron, J., Lagae, L., Maurice, T., & Delprat, B. (2021). Use of zebrafish models to boost research in rare genetic diseases. *International Journal of Molecular Sciences, 22*(24), 13356. https://doi.org/10.3390/ijms222413356

De Los Rios, P., Ben-Zvi, A., Slutsky, O., Azem, A., & Goloubinoff, P. (2006). Hsp70 chaperones accelerate protein translocation and the unfolding of stable protein aggregates by entropic pulling. *Proceedings of the National Academy of Sciences of the United States of America, 103*(16), 6166–6171. https://doi.org/10.1073/pnas.0510496103

Duennwald, M. L., Echeverria, A., & Shorter, J. (2012). Small heat shock proteins potentiate amyloid dissolution by protein disaggregases from yeast and humans. *PLoS Biology, 10*(6), e1001346. https://doi.org/10.1371/journal.pbio.1001346

Ellis, R. J. (1990). The molecular chaperone concept. *Seminars in Cell Biology, 1*(1), 1–9.

Er, E., Oliver, L., Cartron, P.-F., Juin, P., Manon, S., & Vallette, F. M. (2006). Mitochondria as the target of the pro-apoptotic protein Bax. Biochimica et Biophysica Acta (BBA) - *Bioenergetics, 1757*(9–10), 1301–1311. https://doi.org/10.1016/j.bbabio.2006.05.032

Fernández-Fernández, M. R., & Valpuesta, J. M. (2018). Hsp70 chaperone: A master player in protein homeostasis. *F1000Research, 7, F1000 Faculty Rev-1497*. https://doi.org/10.12688/f1000research.15528.1

Flaherty, K. M., DeLuca-Flaherty, C., & McKay, D. B. (1990). Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature, 346*(6285), 623–628. https://doi.org/10.1038/346623a0

Flynn, G. C., Chappell, T. G., & Rothman, J. E. (1989). Peptide binding and release by proteins implicated as catalysts of protein assembly. *Science, 245*(3859), 385–390. https://doi.org/10.1126/science.2773615

Fourie, A. M., Sambrook, J. F., & Gething, M. J. (1994). Common and divergent peptide binding specificities of hsp70 molecular chaperones. *Journal of Biological Chemistry, 269*(41), 30470–30478. https://doi.org/10.1074/jbc.269.41.30470

Freiden, P. J., Gaut, J. R., & Hendershot, L. M. (1992). Interconversion of three differentially modified and assembled forms of BiP. *EMBO Journal, 11*(1), 63–70. https://doi.org/10.1002/j.1460-2075.1992.tb05028.x

Gao, X., Carroni, M., Nussbaum-Krammer, C., Mogk, A., Nillegoda, N. B., Szlachcic, A., Guilbride, D. L., Saibil, H. R., Mayer, M. P., & Bukau, B. (2015). Human Hsp70 disaggregase reverses Parkinson's-linked α-synuclein amyloid fibrils. *Molecular Cell, 59*(5), 781–793. https://doi.org/10.1016/j.molcel.2015.07.012

Gehrmann, M., Marienhagen, J., Eichholtz-Wirth, H., Fritz, E., Ellwart, J., Jäättelä, M., Zilch, T., & Multhoff, G. (2005). Dual function of membrane-bound heat shock protein 70 (Hsp70), Bag-4, and Hsp40: Protection against radiation-induced effects and target structure for natural killer cells. *Cell Death & Differentiation, 12*(1), 38–51. https://doi.org/10.1038/sj.cdd.4401510

Gething, M.-J. (1999). Role and regulation of the ER chaperone BiP. *Seminars in Cell & Developmental Biology, 10*(5), 465–472. https://doi.org/10.1006/scdb.1999.0318

Gorbatyuk, M. S., & Gorbatyuk, O. S. (2013). The molecular chaperone GRP78/BiP as a therapeutic target for neurodegenerative disorders: A mini review. *Journal of Genetic Syndromes & Gene Therapy, 4*(2), 128. https://doi.org/10.4172/2157-7412.1000128

Groth, C., Nornes, S., McCarty, R., Tamme, R., & Lardelli, M. (2002). Identification of a second presenilin gene in zebrafish with similarity to the human Alzheimer's disease gene presenilin2. *Development Genes and Evolution, 212*(10), 486–490. https://doi.org/10.1007/s00427-002-0269-5

Guerreiro, R. J., Baquero, M., Blesa, R., Boada, M., Brás, J. M., Bullido, M. J., Calado, A., Crook, R., Ferreira, C., Frank, A., Gómez-Isla, T., Hernández, I., Lleó, A., Machado, A., Martínez-Lage, P., Masdeu, J., Molina-Porcel, L., Molinuevo, J. L., Pastor, P., Pérez-Tur, J., Relvas, R., Oliveira, C. R., Ribeiro, M. H., Rogaeva, E., Sa, A., Samaranch, L., Sánchez-Valle, R., Santana, I., Tàrraga, L., Valdivieso, F., Singleton, A., Hardy, J., & Clarimón, J. (2010). Genetic screening of Alzheimer's disease genes in Iberian and African samples yields novel mutations in presenilins and APP. *Neurobiology of Aging, 31*(5), 725–731. https://doi.org/10.1016/j.neurobiolaging.2008.06.012

Guo, F., Rocha, K., Bali, P., Pranpat, M., Fiskus, W., Boyapalle, S., Kumaraswamy, S., Balasis, M., Greedy, B., Armitage, E. S., Lawrence, N., & Bhalla, K. (2005). Abrogation of heat shock protein 70 induction as a strategy to increase antileukemia activity of heat shock protein 90 inhibitor 17-allylamino-demethoxy geldanamycin. *Cancer Research, 65*(22), 10536–10544. https://doi.org/10.1158/0008-5472.CAN-05-1799

Haass, C., Kaether, C., Thinakaran, G., & Sisodia, S. (2012). Trafficking and proteolytic processing of APP. *Cold Spring Harbor Perspectives in Medicine, 2*(5), a006270. https://doi.org/10.1101/cshperspect.a006270

Hantschel, M., Pfister, K., Jordan, A., Scholz, R., Andreesen, R., Schmitz, G., Schmetzer, H., Hiddemann, W., & Multhoff, G. (2000). Hsp70 plasma membrane expression on primary tumor biopsy material and bone marrow of leukemic patients. *Cell Stress & Chaperones, 5*(5), 438–442. https://doi.org/10.1379/1466-1268(2000)005<0438:hpmeop>2.0.co;2.

Hendershot, L. M., Ting, J., & Lee, A. S. (1988). Identity of the immunoglobulin heavy-chain-binding protein with the 78,000-dalton glucose-regulated protein and the role of

posttranslational modifications in its binding function. *Molecular and Cellular Biology, 8*(10), 4250–4256. https://doi.org/10.1128/mcb.8.10.4250-4256.1988

Kampinga, H. H., & Bergink, S. (2016). Heat shock proteins as potential targets for protective strategies in neurodegeneration. *Lancet Neurology, 15*(7), 748–759. https://doi.org/10.1016/S1474-4422(16)00099-5

Kampinga, H. H., & Craig, E. A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology, 11*(8), 579–592. https://doi.org/10.1038/nrm2941

Knarr, G., Gething, M. J., Modrow, S., & Buchner, J. (1995). BiP-binding sequences in antibodies. *Journal of Biological Chemistry, 270*(47), 27589–27594. https://doi.org/10.1074/jbc.270.47.27589

Krüger, J., Moilanen, V., Majamaa, K., & Remes, A. M. (2012). Molecular genetic analysis of the APP, PSEN1, and PSEN2 genes in Finnish patients with early-onset Alzheimer disease and frontotemporal lobar degeneration. *Alzheimer Disease & Associated Disorders, 26*(3), 272–276. https://doi.org/10.1097/WAD.0b013e318231e6c7

Kumar, S., & Hedges, S. (1998). A molecular timescale for vertebrate evolution. *Nature, 392*, 917–920. https://doi.org/10.1038/31927

Lajoie, P., & Snapp, E. L. (2011). Changes in BiP availability reveal hypersensitivity to acute endoplasmic reticulum stress in cells expressing mutant huntingtin. *Journal of Cell Science, 124*(19), 3332-3343. https://doi.org/10.1242/jcs.087510

Ledford, B. E., & Leno, G. H. (1994). ADP-ribosylation of the molecular chaperone GRP78/BiP. *Molecular and Cellular Biochemistry, 138*(1), 141-148. https://doi.org/10.1007/BF00926449

Lee, L. C., Chen, C. M., Chen, F. L., Lin, P. Y., Hsiao, Y. C., Wang, P. R., Su, M. T., Hsieh-Li, H. M., Hwang, J. C., Wu, C. H., Lee, G. C., Singh, S., Lin, Y., Hsieh, S. Y., Lee-Chen, G. J., & Lin, J. Y. (2009). Altered expression of HSPA5, HSPA8, and PARK7 in spinocerebellar ataxia type 17 identified by 2-dimensional fluorescence difference in gel

electrophoresis. *Clinical Chemistry and Laboratory Medicine, 400*(1-2), 56-62. https://doi.org/10.1016/j.cca.2008.10.013

Leimer, U., Lun, K., Romig, H., Walter, J., Grünberg, J., Brand, M., & Haass, C. (1999). Zebrafish (Danio rerio) presenilin promotes aberrant amyloid β-peptide production and requires a critical aspartate residue for its function in amyloidogenesis. *Biochemistry, 38*(41), 13602–13609. https://doi.org/10.1021/bi991453n

Leno, G. H., & Ledford, B. E. (1990). Reversible ADP-ribosylation of the 78 kDa glucose-regulated protein. *FEBS Letters, 276*(1), 29–33. https://doi.org/10.1016/0014-5793(90)80967-K

Leustek, T., Toledo, H., Brot, N., & Weissbach, H. (1991). Calcium-dependent autophosphorylation of the glucose-regulated protein, Grp78. *Archives of Biochemistry and Biophysics, 289*(1), 256–261. https://doi.org/10.1016/0003-9861(91)90469-Y

Levy-Lahad, E., Wijsman, E. M., Nemens, E., Anderson, L., Goddard, K. A., Weber, J. L., Bird, T. D., & Schellenberg, G. D. (1995). A familial Alzheimer's disease locus on chromosome 1. *Science, 269*(5226), 970–973. https://doi.org/10.1126/science.7638621

Lieschke, G. J., & Currie, P. D. (2007). Animal models of human disease: Zebrafish swim into view. *Nature Reviews Genetics, 8*(5), 353–367. https://doi.org/10.1038/nrg2091

Mayer, M. P., & Gierasch, L. M. (2019). Recent advances in the structural and mechanistic aspects of Hsp70 molecular chaperones. *Journal of Biological Chemistry, 294*(6), 2085–2097. https://doi.org/10.1074/jbc.REV118.002810

Mayer, M. P., & Bukau, B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences, 62*(7), 670-684. https://doi.org/10.1007/s00018-005-5171-5

McKay, D. B. (1993). Structure and mechanism of 70-kDa heat-shock-related proteins. *Advances in Protein Chemistry, 44*, 67–98. https://doi.org/10.1016/s0065-3233(08)60564-1

Melo, E.P., T. Konno, I. Farace, M.A. Awadelkareen, L.R. Skov, F. Teodoro, T.P. Sancho, A.W. Paton, J.C. Paton, M. Fares, P.M.R. Paulo, X. Zhang, E. Avezov (2022) Stress-induced protein disaggregation in the endoplasmic reticulum catalysed by BiP. Nature Commun. 13: 2501. https://doi.org/10.1038/s41467-022-30238-2

Morrow, J. F., Cohen, S. N., Chang, A. C., Boyer, H. W., Goodman, H. M., & Helling, R. B. (1974). Replication and transcription of eukaryotic DNA in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America, 71*(5), 1743–1747. https://doi.org/10.1073/pnas.71.5.1743

Muchowski, P. J., Schaffar, G., Sittler, A., Wanker, E. E., Hayer-Hartl, M. K., Hartl, F. U. (2000). Hsp70 and Hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. *Proceedings of the National Academy of Sciences of the United States of America, 97*(14), 7841-7846. https://doi.org/10.1073/pnas.140202897

Multhoff, G., & Hightower, L. E. (2011). Distinguishing integral and receptor-bound heat shock protein 70 (Hsp70) on the cell surface by Hsp70-specific antibodies. *Cell Stress and Chaperones, 16*(3), 251-255. https://doi.org/10.1007/s12192-011-0317-4

Multhoff, G., Botzler, C., Wiesnet, M., Müller, E., Meier, T., Wilmanns, W., & Issels, R. D. (1995). A stress-inducible 72-kDa heat-shock protein (HSP72) is expressed on the surface of human tumor cells, but not on normal cells. *International Journal of Cancer, 61*(2), 272-279. https://doi.org/10.1002/ijc.2910610222

Musa, A., Lehrach, H., & Russo, V. E. (2001). Distinct expression patterns of two zebrafish homologues of the human APP gene during embryonic development. *Developmental Genes and Evolution, 211*(8), 563-567. https://doi.org/10.1007/s00427-001-0189-9

Newman, M., Ebrahimie, E., & Lardelli, M. (2014). Using the zebrafish model for Alzheimer's disease research. *Frontiers in Genetics, 5*, 189. https://doi.org/10.3389/fgene.2014.00189

Lajoie, P., & Snapp, E. L. (2011). Changes in BiP availability reveal hypersensitivity to acute endoplasmic reticulum stress in cells expressing mutant huntingtin. *Journal of Cell Science, 124*(Pt 19), 3332–3343. https://doi.org/10.1242/jcs.087510

Nornes, S., Newman, M., Wells, S., Verdile, G., Martins, R. N., & Lardelli, M. (2009). Independent and cooperative action of Psen2 with Psen1 in zebrafish embryos. *Experimental Cell Research, 315*(16), 2791–2801. https://doi.org/10.1016/j.yexcr.2009.06.023

Perera, L. A., & Ron, D. (2023). AMPylation and endoplasmic reticulum protein folding homeostasis. *Cold Spring Harbor Perspectives in Biology, 15*(3), a041265. https://doi.org/10.1101/cshperspect.a041265

Persky, N. S., & Lovett, S. T. (2008). Mechanisms of recombination: lessons from E. coli. Critical reviews in biochemistry and molecular biology, 43(6), 347–370. https://doi.org/10.1080/10409230802485358

Pieri, L., Madiona, K., Bousset, L., & Melki, R. (2012). Fibrillar α-synuclein and huntingtin exon 1 assemblies are toxic to the cells. *Biophysical Journal, 102*(12), 2894–2905. https://doi.org/10.1016/j.bpj.2012.04.050

Pouysségur, J., Shiu, R. P., & Pastan, I. (1977). Induction of two transformation-sensitive membrane polypeptides in normal fibroblasts by a block in glycoprotein synthesis or glucose deprivation. *Cell, 11*(4), 941–947. https://doi.org/10.1016/0092-8674(77)90305-1

Preissler, S., Rato, C., Chen, R., Antrobus, R., Ding, S., Fearnley, I. M., & Ron, D. (2015). AMPylation matches BiP activity to client protein load in the endoplasmic reticulum. eLife, 4, e12621. https://doi.org/10.7554/eLife.12621

Radons, J. (2016). The human HSP70 family of chaperones: Where do we stand? *Cell Stress and Chaperones, 21*(3), 379–404. https://doi.org/10.1007/s12192-016-0676-6

Rogaev, E. I., Sherrington, R., Rogaeva, E. A., Levesque, G., Ikeda, M., Liang, Y., Chi, H., Lin, C., Holman, K., Tsuda, T., et al. (1995). Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature, 376*(6543), 775-778. https://doi.org/10.1038/376775a0

Rosenzweig, R., Nillegoda, N. B., Mayer, M. P., & Bukau, B. (2019). The Hsp70 chaperone network. *Nature Reviews Molecular Cell Biology, 20*(11), 665-680. https://doi.org/10.1038/s41580-019-0133-3

Rüdiger, S., Germeroth, L., Schneider-Mergener, J., & Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *The EMBO Journal, 16*(7), 1501-1507. https://doi.org/10.1093/emboj/16.7.1501

Schmid, D., Baici, A., Gehring, H., & Christen, P. (1994). Kinetics of molecular chaperone action. *Science, 263*(5149), 971-973. https://doi.org/10.1126/science.8310296

Sharma, S. K., De los Rios, P., Christen, P., Lustig, A., & Goloubinoff, P. (2010). The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. *Nature Chemical Biology, 6*(12), 914-920. https://doi.org/10.1038/nchembio.455

Sherrington, R., Froelich, S., Sorbi, S., Campion, D., Chi, H., Rogaeva, E. A., Levesque, G., Rogaev, E. I., Lin, C., Liang, Y., Ikeda, M., Mar, L., Brice, A., Agid, Y., Percy, M. E., Clerget-Darpoux, F., Piacentini, S., Marcon, G., Nacmias, B., Amaducci, L., Frebourg, T., Lannfelt, L., Rommens, J. M., & St George-Hyslop, P. H. (1996). Alzheimer's disease associated with mutations in presenilin 2 is rare and variably penetrant. *Human Molecular Genetics, 5*(7), 985-988. https://doi.org/10.1093/hmg/5.7.985

Sherrington, R., Rogaev, E. I., Liang, Y., Rogaeva, E. A., Levesque, G., Ikeda, M., Chi, H., Lin, C., Li, G., Holman, K., Tsuda, T., Mar, L., Foncin, J. F., Bruni, A. C., Montesi, M. P., Sorbi, S., Rainero, I., Pinessi, L., Nee, L., Chumakov, I., Pollen, D., Brookes, A., Sanseau, P., Polinsky, R. J., Wasco, W., Da Silva, H. A., Haines, J. L., Perkicak-Vance, M. A., Tanzi, R. E., Roses, A. D., Fraser, P. E., Rommens, J. M., & St George-Hyslop, P. H. (1995). Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature, 375*(6534), 754-760. https://doi.org/10.1038/375754a0

Smith H. O. (1979). Nucleotide sequence specificity of restriction endonucleases. Science (New York, N.Y.), 205(4405), 455–462. https://doi.org/10.1126/science.377492

Sousa, R., Liao, H. S., Cuéllar, J., Jin, S., Valpuesta, J. M., Jin, A. J., & Lafer, E. M. (2016). Clathrin-coat disassembly illuminates the mechanisms of Hsp70 force generation. *Nature Structural & Molecular Biology, 23*(9), 821-829. https://doi.org/10.1038/nsmb.3272

Stankiewicz, A. R., Lachapelle, G., Foo, C. P., Radicioni, S. M., & Mosser, D. D. (2005). Hsp70 inhibits heat-induced apoptosis upstream of mitochondria by preventing Bax translocation. *Journal of Biological Chemistry, 280*(46), 38729-38739. https://doi.org/10.1074/jbc.M509497200

Swain, J. F., Dinler, G., Sivendran, R., Montgomery, D. L., Stotz, M., & Gierasch, L. M. (2007). Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Molecular Cell, 26*(1), 27-39. https://doi.org/10.1016/j.molcel.2007.02.020

Taylor, I. R., Ahmad, A., Wu, T., Nordhues, B. A., Bhullar, A., Gestwicki, J. E., & Zuiderweg, E. R. P. (2018). The disorderly conduct of Hsc70 and its interaction with the Alzheimer's-related Tau protein. *Journal of Biological Chemistry, 293*(27), 10796-10809. https://doi.org/10.1074/jbc.RA118.002234

Voisin, T., & Vellas, B. (2009). Diagnosis and treatment of patients with severe Alzheimer's disease. *Drugs & Aging, 26*(2), 135-144. https://doi.org/10.2165/0002512-200926020-00005

Wang, J. L., Liem, D., & Ping, P. (2017). HSPA5 gene encoding Hsp70 chaperone BiP in the endoplasmic reticulum. *Gene, 618*, 14-23. https://doi.org/10.1016/j.gene.2017.03.005

Wentink, A. S., Nillegoda, N. B., Feufel, J., Ubartaitė, G., Schneider, C. P., De Los Rios, P., Hennig, J., Barducci, A., & Bukau, B. (2020). Molecular dissection of amyloid disaggregation by human HSP70. Nature, 587(7834), 483–488. https://doi.org/10.1038/s41586-020-2904-6

Wentink, A., Nussbaum-Krammer, C., & Bukau, B. (2019). Modulation of amyloid states by molecular chaperones. *Cold Spring Harbor Perspectives in Biology, 11*(7), a033969. https://doi.org/10.1101/cshperspect.a033969

Yang, J., Nune, M., Zong, Y., Zhou, L., & Liu, Q. (2015). Close and allosteric opening of the polypeptide-binding site in a human Hsp70 chaperone BiP. *Structure, 23*(12), 2191–2203. https://doi.org/10.1016/j.str.2015.10.012

Zhang, H., Hu, H., Wu, S., & Perrett, S. (2023). Effect of evolution of the C-terminal region on chaperone activity of Hsp70. *Protein Science, 32*(1), e4549. https://doi.org/10.1002/pro.4549

Zhao, F. G., Wang, Y. H., Yang, J. F., Ma, Q. L., Tang, Z., Dong, X. M., & Chan, P. (2005). Association between acyl-coenzyme A: cholesterol acyltransferase gene and risk for Alzheimer's disease in Chinese. *Neuroscience Letters, 388*(1), 17-20. https://doi.org/10.1016/j.neulet.2005.06.020

Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., & Hendrickson, W. A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science, 272*(5260), 1606–1614. https://doi.org/10.1126/science.272.5260.1606

Zhuravleva, A., & Gierasch, L. M. (2015). Substrate-binding domain conformational dynamics mediate Hsp70 allostery. *Proceedings of the National Academy of Sciences U S A, 112*(22), E2865–E2873. https://doi.org/10.1073/pnas.1506897112

Zuiderweg, E. R., Hightower, L. E., & Gestwicki, J. E. (2017). The remarkable multivalency of the Hsp70 chaperones. *Cell Stress & Chaperones, 22*(2), 173–189. https://doi.org/10.1007/s12192-017-0776-y

ANNEXES

Annex A - Fw sequence of miniprep from ligation colony

```
Alignment of Sequence_1:  [ORF zfBiP_V25.xdna] with  Sequence_2: [Sequence Window #8]

Similarity : 220/963 (22.85 %)

Seq_1  1     Gttgggacagtgattgggatcgacctcgggaccacatactcctgtgttggagtctacaag  60

Seq_2  1     ------------------------------------------------------------  0


Seq_1  61    aatggccgtgttgagattattgccaatgaccagggaaaccgcatcactccgtcatacgtg  120
                                                                        |
Seq_2  1     ------------------------------------------------------TTA  3


Seq_1  121   gcctttaccactgaaggagagcggctcatcggagatgctgcgaagaaccagctcacatcc  180
               |  ||  |   | |   | |       |       || |        |  ||| |
Seq_2  4     A-CC-TATAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCACCTCGAGAAATCATAAAA  61


Seq_1  181   aaccctgaaaacactgtgtttgacgccaagaggctgatcggacgcacatggggcgactct  240
              ||                  ||   |||       |                | |     |
Seq_2  62    AATTTATTTGCTTTGTGAGCGGATAACAATTTATAATAGATTCAATTGTGAGCGGATAACA  121


Seq_1  241   tctgtgcagcaggacatcaaatacttccccttaaggtgatcgagaagaaaaacaagcct  300
               |   ||           ||         ||||        ||    |      ||
Seq_2  122   ATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGAGGATCTCACCATCACCA  181


Seq_1  301   cacatccagctggacatcggctctggtcagatgaagacgtttgcaccggaggaaatttcc  360
                |      ||              |      |          || |   |
Seq_2  182   TCACCATACGGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGG  241


Seq_1  361   gccatggttttgaccaagatgaaggaaaccgcagaggcttatctgggaaagaaggtcact  420
               |                  |  |    |  ||     |  | || |
Seq_2  242   TTGCCGCCGGGCGTTTTTTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAG  301


Seq_1  421   catgctgtggtcaccgttcctgcttatttcaacgatgctcagcgtcaggccactaaagat  480
              |      | | |          | |     |||   |  |    |     |    ||
Seq_2  302   CTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAAT  361


Seq_1  481   gctggaaccattgctgggctgaatgtcatgaggatcatcaatgagcctacggcggctgcc  540
              |      |        |||      | ||     |     |     |     |
Seq_2  362   GGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGA  421


Seq_1  541   attgcatacggtctggacaagagggacggagagaaaaacatcctggtgttcgatctgggt  600
                | |  |      |        |       | |      |   |       |
Seq_2  422   CCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTT  481


Seq_1  601   ggtggcacctttgacgtgtctctgctgaccatcgataacggcgtgtttgaagtggtggcc  660
                |  | ||   |   |   |||   ||      |    | |     |  | ||
Seq_2  482   ATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGG  541


Seq_1  661   acaaacggagacactcacctgggcggagaagacttcgaccagcgcgtcatggagcacttc  720
               |              | | |  |     |   |   |  || |    |   | |
Seq_2  542   CAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCC  601


Seq_1  721   atcaagctgtacaagaagaagacgggcaaagacgtgcgcaaagacaaccgcgccgtgcag  780
              || |       ||          ||   |  |||   |     |      |||
Seq_2  602   ATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGT  661


Seq_1  781   aagctgcgcagagaggtggagaaggctaagagagcgctgtctgcccagcatcaggcccgc  840
                        ||| |          || ||    ||  || |    |||
Seq_2  662   TTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTA  721


Seq_1  841   atcgagatcgagtccttctttgaaggagaagatttctctgagactctcaccagagccaag  900
              |  |   |    |    | | |              |                 | |
Seq_2  722   AAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTT  781


Seq_1  901   tttgaggagctcaacatggacttgttccgctccactatgaagccggttcagaaggttctg  960
              ||     |          | |      ||  |       |      |     ||
Seq_2  782   TTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAAT  841


Seq_1  961   gaggactctgacccgaagaagccagatatcgatgagatcgtgctggtcggcggctccact  1020
```

```
                     |  |||  | |        |     |    |   | ||  |    |   |
Seq_2  842   ATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTT  901

Seq_1  1021  cgtatcccgaagatccagcagctggtgaaggagttcttcaacggaaaagagccgtccaga  1080
                ||  |    ||   |              |   | |||| | |  | |  |    |
Seq_2  902   GTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGC  961

Seq_1  1081  ggaatcaaccctgacgaggccgtggcgtacggagctgctgtccaggctggagtcctgtcc  1140
               |
Seq_2  962   AG----------------------------------------------------------  963

Seq_1  1141  ggagaggaggagaccggtgatctggttcttctggacgtgtgtccgctgactctgggcatt  1200

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1201  gagactgttggaggagtgatgaccaaactcattcccagaaacactgttgttcccaccaag  1260

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1261  aaatcccagatcttctccactgcttccgacaaccagcccaccgtcactatcaaagtttat  1320

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1321  gagggcgagcgtcccctgaccaaagacaaccatctgctgggcacctttgacctgacaggc  1380

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1381  atccctccagcacctcgtggtgtcccacagatcgaggtaactttcgagatcgacgtcaac  1440

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1441  ggcatcctgcgcgtcaccgccgaagacaaaggcaccggaaacaaaaacaagatcaccatt  1500

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1501  accaacgaccagaaccggctgacccctgaggacatcgagagaatggtgaacgaagccgag  1560

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1561  agattcgctgatgaggacaagaaactgaaggagagaatcgacagccgcaatgaattggag  1620

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1621  agctacgcctattccctgaagaaccagatcggggataaagagaaattaggcggaaagtta  1680

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1681  tcctctgaagacaaggaggccatcgagaaggcagtggaggagaagatcgagtggctggag  1740

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1741  gcgcatcaggacgccgatctggaggaattccaggccaaaaagaaggagctggaggaggtg  1800

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1801  gtgcagcccatcgtcagcaaactgtacggcagtgcgggaggaccaccgcctgaagaggcc  1860

Seq_2  964   ----------------------------------------------------------  963

Seq_1  1861  gaagagaaggacgagctgtag  1881

Seq_2  964   --------------------  963
```

Annex B – Chromatogram