Universidade do Minho
Escola de Engenharia

André Filipe Fernandes Costa

Fusion proteins as innovative hair cosmetic products

novembro de 2021

Fusion proteins as innovative hair cosmetic products
André Costa
UMinho | 2021

Universidade do Minho
Escola de Engenharia

André Filipe Fernandes Costa

Fusion proteins as innovative hair cosmetic products

Dissertação de Mestrado
Mestrado em Biotecnologia

Trabalho efetuado sob a orientação do
Doutor Artur Jorge Araújo Magalhães Ribeiro
e da
Doutora Carla Manuela Pereira Marinho da Silva

novembro de 2021

AGRADECIMENTOS

No culminar de mais uma etapa, que foi sem sombra de dúvida um enorme crescimento tanto a nível científico como a nível profissional resta-me expressar os meus mais sinceros agradecimentos a todos que tornaram possível a sua conclusão.

Em primeiro lugar quero agradecer aos meus orientadores Doutor Artur Ribeiro e Doutora Carla Silva por todo o apoio, motivação e partilha de conhecimento constante. Agradecer ainda por toda a ajuda e paciência nas horas menos boas. Ao longo deste processo vincaram profundamente a vossa condição de excelentes mentores!

Não posso também deixar de agradecer ao Professor Artur Cavaco-Paulo por me receber no seu grupo de trabalho e por me ter dado a oportunidade de realizar a minha dissertação de mestrado numa área que é tão do meu interesse.

Um agradecimento também para todos os meus colegas do LBBN por todo o companheirismo, ajuda constante e por todos os concelhos. Um obrigado especial à Catarina e à Filipa por estarem sempre disponíveis para esclarecer as dúvidas que surgissem e por toda a ajuda dada.

Tenho também que agradecer ao cenas, ao Tisco, à crew, à Inês, à Alexandra e às Tias pelo suporte constante. A vossa amizade e a vossa alegria foram essenciais para a ultrapassar as frustrações do dia a dia! Quero vos também agradecer todos os momentos de alegria, convívio e mesmo de estupidez e por todas as vivências que partilhamos juntos. Espero que se sigam muitas meltdowns coletivos, discusões parvas e acima de tudo vivências incríveis.

Por último, mas decididamente não menos importante, um OBRIGADO aos meus pais e à minha irmã, que sem eles este percurso não teria sido possível. Quero agradecer todo o amor, incentivo e orgulho que demonstram constantemente ter em mim e por estarem sempre presentes! Obrigado por terem acreditado em mim!

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

RESUMO

O cabelo tem uma importância indiscutível na imagem corporal e na sociedade atual. É um sistema complexo composto maioritariamente por queratinas. Com a crescente procura por novos produtos capilares, a indústria cosmética tem-se focado na procura de novas alternativas às formulações químicas atualmente utilizadas. Proteínas e formulações à base de proteínas são vistas como excelentes alternativas. De entre as proteínas, as proteínas de fusão têm despertado particular interesse tendo sido recentemente exploradas no desenvolvimento de formulações cosméticas capilares.

A expressão de proteínas é uma etapa essencial no desenvolvimento de formulações de base proteica. Neste trabalho foram exploradas sete proteínas baseadas na proteína XZX (XZX–1, XZX–2, XZX–3, XZX–4, XZX–5, XZX–6 e XZX–7) como novos agentes cosméticos. Para a expressão das proteínas XZX (1 a 6) foram testadas cinco condições de cultura. Para a expressão da proteína XZX–7 foi realizada uma pré-seleção, para determinar o melhor meio de cultura comercial. E usando o software DesignExpert® foram realizadas três otimizações para encontrar a melhor mistura do meio de cultura (comercial, subprodutos industriais e aditivos) e as melhores condições de pH e temperatura para expressar a proteína XZX–7. Esta proteína foi produzida usando as condições previamente otimizadas e caracterizada. A capacidade da XZX–7 de se ligar a cabelo Asiático virgem e o seu potencial para o encaracolar foi avaliada.

As proteínas XZX (2 a 6) foram expressas em meio LB + IPTG durante 24 horas a 37 °C. No entanto é necessária uma otimização das condições de expressão ou uma mudança do hospedeiro de expressão, a fim de tornar a produção destas proteínas um processo com viabilidade económica.

Após otimização do meio de cultura e das condições de crescimento determinadas pelo desenho experimental, determinou-se que as condições para maximizar a expressão da proteína XZX–7 eram: 10 g/L TB-AIM, 30g/L CSL, 0.934 g/L NaCl, pH 7, 37 °C. Com esta otimização foi duplicado o rendimento de proteína para uma DO_{600} = 0,1 de 0,0012 mg para o meio comercial TB-AIM para 0,0025 mg para o meio otimizado. A proteína XZX–7 demonstrou capacidade de se ligar ao cabelo asiático com uma percentagem de ligação de 17,15% em água e de 33,03% numa formulação etanólica. Em relação à capacidade de encaracolamento da XZX–7, a condição de 20 mg/mL em formulação etanólica com pré-tratamento com ureia demonstrou a melhor eficiência de encaracolamento após seis lavagens com champô.

Palavras-chave: Cabelo; Desenho Experimental; Produtos cosméticos para os cabelos; Proteínas de fusão

ABSTRACT

Hair has an unquestionable importance in body image and in today's society. It is a complex system made up mostly of keratins. With the increasing demand for new hair care products, the cosmetic industry has focused on finding new alternatives to the chemical formulations currently used. Protein and protein-based formulations are seen as excellent alternatives. Among proteins, fusion proteins aroused particular interest and have recently been explored in the development of hair care cosmetic formulations.

Protein expression is an essential step in the development of protein-based formulations. In this work, seven proteins based on the XZX protein (XZX–1, XZX–2, XZX–3, XZX–4, XZX–5, XZX–6 and XZX–7) were explored as new cosmetic agents. For the expression of XZX proteins (1 to 6) five culture conditions were tested. For the expression of XZX–7 protein, a screening was performed to determine the best commercial culture medium. Using DesignExpert® software, three optimizations were performed to find the best mixture of culture medium (commercial media, industrial by-products and additives) and the best pH and temperature conditions to express the XZX–7 protein. This protein was produced using the conditions previously optimized and characterized. The ability of XZX–7 to bind to virgin Asian hair and its potential for curling was evaluated.

XZX proteins (2 to 6) were expressed in LB medium + IPTG for 24 hours at 37°C. However, an optimization of expressing conditions or a change in the expression host is necessary, in order to make the production of these proteins a process economical viable.

After optimizing the culture medium and growth conditions determined by the experimental design, it was determined that the conditions for maximizing the expression of XZX-7 protein were: 10 g/L TB-AIM, 30g/L CSL, 0.934 g/ L NaCl, pH 7, 37°C. With this optimization, the protein yield for an OD600 = 0.1 was doubled from 0.0012 mg for the commercial TB-AIM medium to 0.0025 mg for the optimized medium. The XZX-7 protein demonstrated the ability to bind to Asian hair with a binding percentage of 17.15% in water and 33.03% in an ethanolic formulation. Regarding the curling capacity of XZX-7, the condition of 20 mg/mL in an ethanolic formulation with urea pre-treatment demonstrated the best curling efficiency after six shampoo washes.

Keywords: Experimental Design; Fusion Proteins; Hair; Hair cosmetic products

LIST OF CONTENT

LIST OF ABBREVIATIONS

18-MEA	18-methyl eicosonoic acid
Ala	Alanine
Arg	Arginine
CMC	Cell membrane complex
CSL	Corn step liquor
Cys	Cysteine
E. coli	*Escherichia coli*
Gly	Glycine
IF	Intermediate filaments
IPTG	Isopropyl β-D-1-thioXZXactopyranoside
KAP	Keratin associated protein
LB	Lysogeny broth
LB-AIM	Auto induction Lysogeny Broth
OD	Optic density
PI	Isoelectric point
SB-AIM	Auto-induction Super Broth
SDS-PAGE	Sodium dodecyl sulfate- polyacrylamide gel electrophoresis
TB-AIM	Auto induction Terrific Broth
TEM	Transmission electron microscopy
UHSP	Ultra-high sulfur protein

LIST OF FIGURES

LIST OF TABLES

1.STATE OF THE ART

1.1. Hair

The hair represents an important component of body image with unquestionable relevance in society (Oliver *et al.*, 2019). Although the enormous aesthetic significance in our culture, the hair also presents a protection function (Yu *et al.*, 2017). Taking into consideration is great relevance, the haircare industry has developed several products to protect the hair or to change some of its properties. These properties can range from color, shape, volume, shine and hydration degree (Cruz *et al.*, 2017).

1.1.1.Hair anatomy

Hair is a complex system with an individual chemical and physical behavior totally adapted to fulfil its protective and insulating functions (Buffoli *et al.*, 2014; Malinauskyte *et al.*, 2021). The hair fibers are mainly composed by dead and fully keratinized epithelial cells, organized in an intricate structure, composed by several morphological components, which act as one (Buffoli *et al.*, 2014). The human hair has a diameter of 50 to 100 μm, and is characterized by three main regions: the cuticle, the cortex and the medulla (**Figure 1**) (Buffoli *et al.*, 2014; Yu *et al.*, 2017).

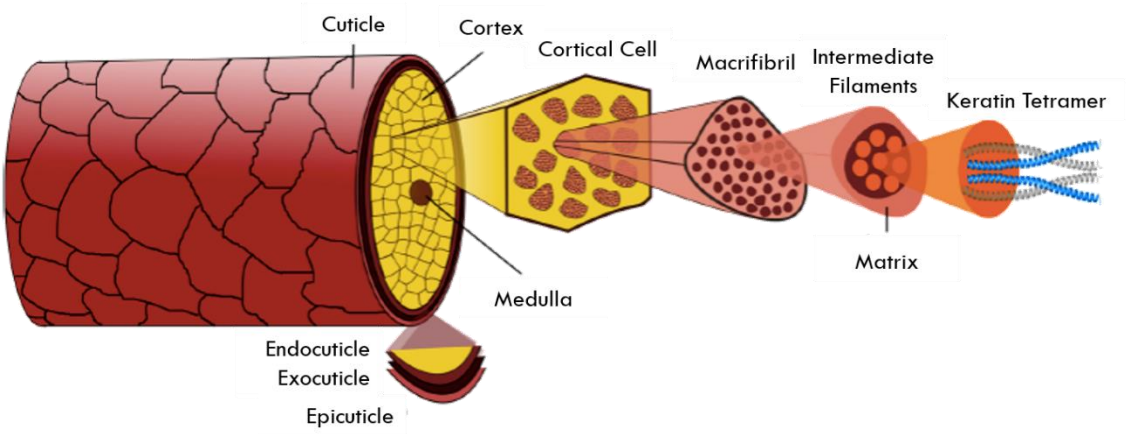

Figure 1 – Scheme of hairs' transversal cut, showing the different regions: medulla, cortex and cuticle, and the constitution of a cortical cell. Adapted of Cruz *et al.* (2016)

The cuticle is the most external layer of the hair and consists in overlapping cuticle cells with a root-to-tip orientation **(Figure 2-A)**. This layer is characterized to be highly resistant to chemical and physical damage. The cuticle cells have a thin protein membrane, the epicuticle, that is covered by a monolayer of covalently bound fatty acids, mainly 18-methyl eicosanoic acid (18-MEA). The

hydrophobicity of hair is related with the presence of 18-MEA and its removal may expose the hair to damage by increasing its surface hydrophilicity (Kamath and Weigmann, 1982; Dawber, 1996; Gavazzoni Dias, 2015).

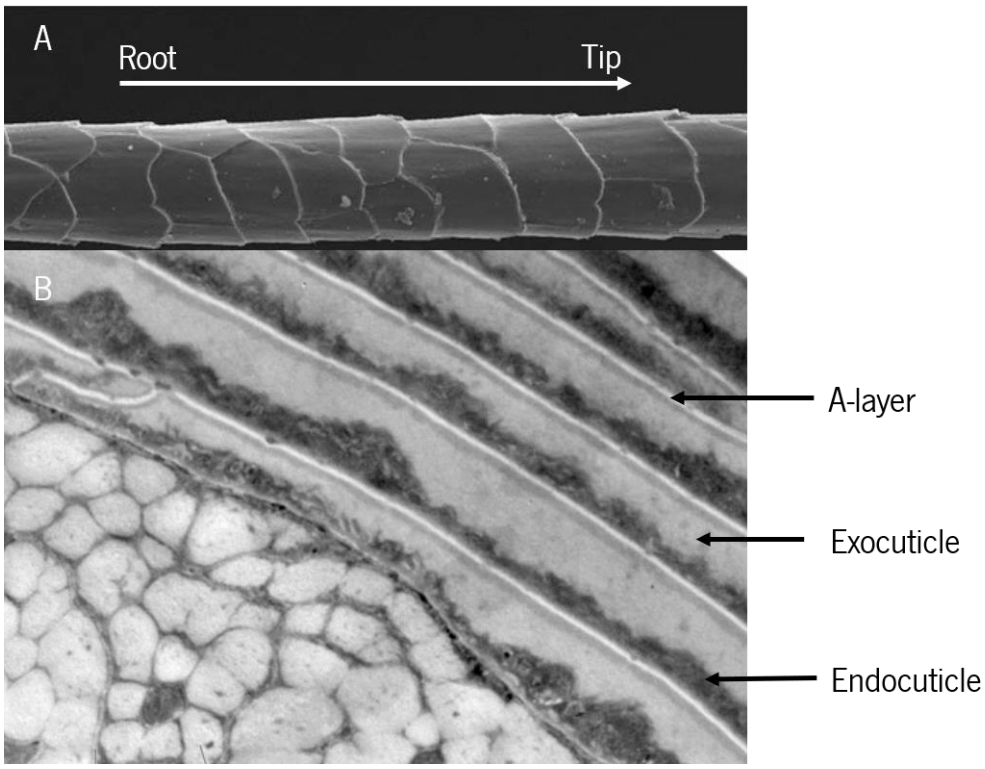

Figure 2 – A) Scanning Electron Microscopy micrographs of a hair fiber showing the overlapping scale structure. **B)** Transmission electron microscopy (TEM) of cross section of human scalp hair showing multiple overlapping cuticle layers. Adapted of Plowman, J.E., Duane 2018

The hair cuticle is also constituted by three other layers: the A-layer, the exocuticle and the endocuticle (**Figure 2-B**). These layers present a high content in cysteine (Cys) residues and are rich in heavily cross-linked keratin proteins. The A-layer is approximately 100 nm thick and has the highest Cys content. The high Cys and cross-linked proteins content turn this layer strongly resistant against degradation. Moreover, recent studies have identified several ultra-high sulfur proteins in the A-layer (1 in every 2.7 residues is Cys) (Swift, 1999; Bringans *et al.*, 2007; Plowman, J.E., Duane, 2018).

The exocuticle, also known as B-layer, varies in thickness and presents an amorphous appearance **(Figure 2-B)**. The sulfur content of the exocuticle is significantly lower when compared with the A-layer. Moreover, the exocuticle is easily dissolved when treated with urea in conjugation with the reductant tris(2-carboxyethyl)phosphine, suggesting the absence of isopeptide bonds (Swift, 1997, 1999; Bringans *et al.*, 2007; Plowman, J.E., Duane, 2018).

The endocuticle has an amorphous appearance and is located beneath the exocuticle **(Figure 2-B)**. This layer varies in thickness, and it is characterized by a low sulfur content and high level of acidic and basic amino acids. Based on its general appearance, the endocuticle appears to have been derived from developing cell cytoplasm and cytoplasmic components (Hallegot and Corcuff, 1993; Swift, 1999; Plowman, J.E., Duane, 2018).

Another important component of hair is the Cell Membrane Complex (CMC), which is located between the cuticle cells, the cortical cells and between the cuticle and the cortical cells. The CMC is composed by cell membranes and adhesive material, with the function to bind the different components of hair. The CMC contains a low proportion of sulfur-containing amino acids including Cys when compared to the other hair components and together with the endocuticle is sometimes referred as the non-keratinous region of the hair. The CMC composition differs between the cuticle and the cortical regions. The CMC is very vulnerable to chemical treatments such as bleaching, dyeing, hair straightening and perm procedures (Robbins 2016; Gavazzoni Dias 2015).

The cortex represents the main component of human hair and is constituted by elongated fusiform cells containing proteins, mainly keratin and keratin associated proteins (KAPs), and melanin granules (**Figure 1**). The cortical cells are composed by macrofibrils (approximately 90 %), nuclear debris and pigment granules (approximately 10 %) (Akkermans and Warren, 2004; Buffoli *et al.*, 2014).

The macrofibrils are composed by protein filaments (intermediate filaments (IFs)), which are constituted by keratin and CMC. The macrofibrils can present three different architectures: orthocortical, paracortical and mesocortical (**Figure 3**). The orthocortical cells consist of a straight IF surrounded by six IFs that are gently coiled around the core, where the next ring of filaments is coiled around the firsts (**Figure 3-B**). The paracortical architecture consist in a generally spaced IFs which are roughly aligned parallel along to the fiber axis (**Figure 3-A)**. The mesocortical architecture contain sizable regions in which the IFs are highly aligned and packed laterally into a tight hexagonal array (**Figure 3-C**). In human hair a single cortical cell can contain macrofibrils of different architectures, mainly orthocortical and paracortical (Orwin *et al.*, 1984; Bryson *et al.*, 2009).

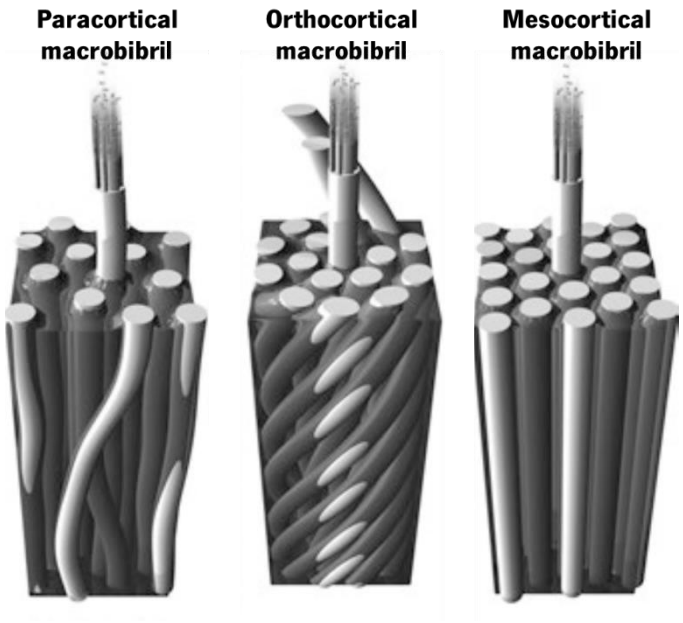

Figure 3 – Representation of the type of macrofibrils. Adapted of Plowman, J.E., Duane 2018.

The medulla is the core of the hair shaft (**Figure 1**) and it is not always present. The medulla is formed by loosely packed cells with variable thickness (5 to 10 μm), and its function is still unclear (Cruz *et al.*, 2016; Oliver *et al.*, 2019). It can be involved in the splitting of hair, since it is a weaker area ideal for the propagation of cracks along the hair fiber (Kamath and Weigmann, 1982; McMichael, 2007).

1.1.2.Hair cycle

The hair follicle is a multifaceted epithelial structure surrounded by an outer hair root sheath, that have as function protecting the growing hair, and an inner hair root sheath, that follow the hair shaft to the opening of the sebaceous gland (**Figure 4**).

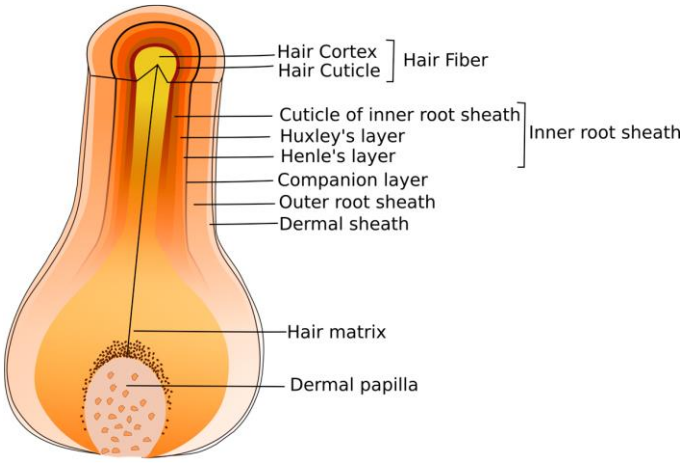

Figure 4 – Basic structure of the hair follicle adapted from Cruz *et al.*, 2016

Typically the hair grow 10-15 mm per month, this growth is cyclic and consist in three main phases: anagen, catagen and telogen (**Figure 5**) (Lai-Cheong and McGrath, 2009; Cruz *et al.*, 2016). The anagen phase (2 to 10 years) consist in the active growth of the hair. Cells in the bottom section of the hair divide quickly during this phase, whereas matrix cells move outward (Grymowicz *et al.*, 2020). The catagen phase (around 3 weeks) is a transient phase, where the hair stops growing but the cellular activity continues in the germinal center of the hair follicle. During this stage, the hair shaft loses its connections to the papillae and contracts (Cruz *et al.*, 2016; Grymowicz *et al.*, 2020). The telogen phase (3 – 4 mouths) is the resting and final phase of the cycle. This phase is defined as the regression of the matrix and retraction of the papilla to a place near the bulge. During this stage, there is no considerable proliferation or apoptosis (Grymowicz *et al.*, 2020).

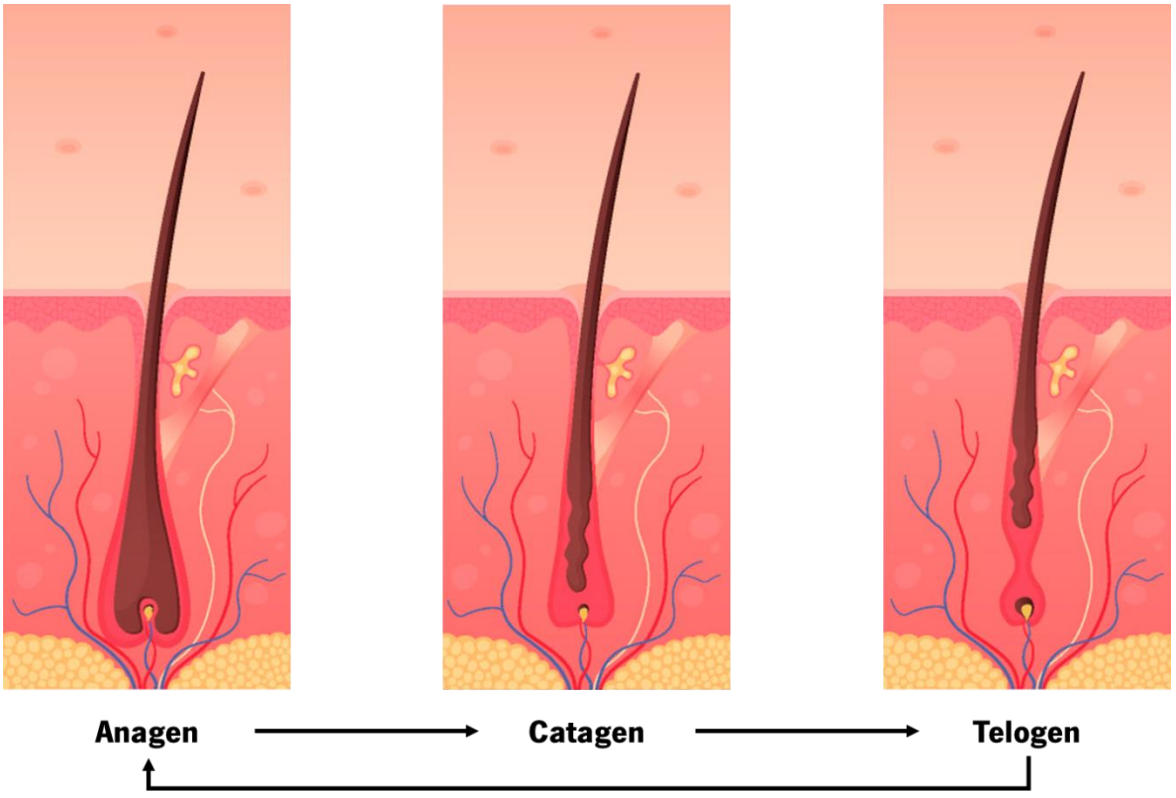

Figure 5 – Hair cycle adapted from Grymowicz *et al.*, 2020

1.1.3.Hair Proteins

Proteins constitute 90% of hair´s mass and 85% of those proteins are keratins and KAPs (**Figure 6**). The keratins constitute the IFs and the KAPs the matrix surrounding the IFs. The hair tensile strength and its flexibility is associated with the keratin and KAPs, respectively, present in hair (Cruz *et al.*, 2013; Malinauskyte *et al.*, 2021).

Keratin proteins refers to a group of insoluble proteins organized as intermediate filaments that form the bulk of cytoplasmic epithelia and epidermal appendage structures like hair, horns, nails, wool and feathers (Tinoco *et al.*, 2021b).

Considering the cysteine content and the number of dissulphide bonds keratins can be classified as high sulfur proteins and low sulfur proteins (**Figure 6**), or as soft or hard keratins. The soft keratins have a low content of dissulphide bonds and the hard keratins have a high number of dissulphide bonds (Basit *et al.*, 2018; Plowman, J.E., Duane, 2018).

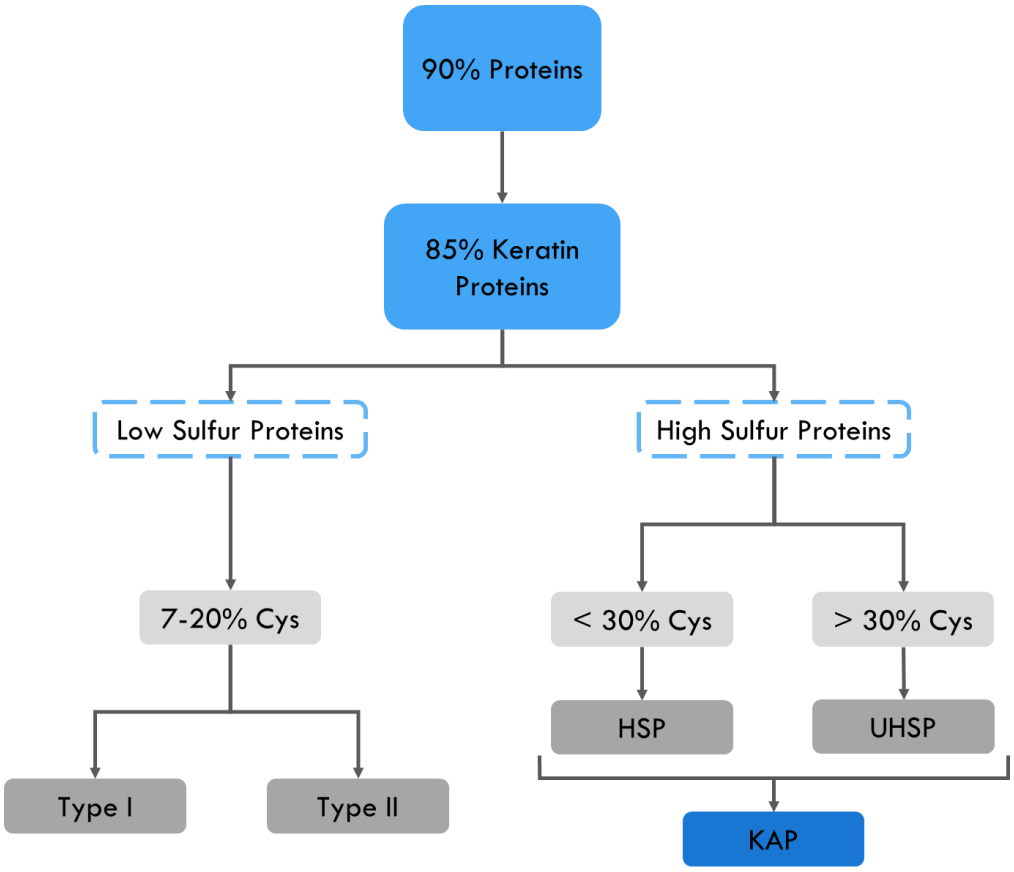

Figure 6 – Quantification of keratins in hair and their classification. Where HSP is high sulfur proteins, UHSP is ultra-high sulfur proteins and KAP is keratin associated protein.

Regarding the low sulfur proteins, they present 7 to 20% of Cys residues and are further divided in Type I and Type II. Type I and Type II keratins are designated by K and a number (Table 2) (Powell and Rogers, 1997; Schweizer *et al.*, 2006; Basit *et al.*, 2018; Plowman, J.E., Duane, 2018).

Type I keratins have an acid nature, and their size varies between 403-471 amino acids. Type II keratins are bigger with a length that varies between 479-507 amino acids, and they have a neutral/basic behavior. In **Table 1** are presented the Type I and Type II human hair's keratins, alongside their localization on the hair fiber (Schweizer *et al.*, 2006).

Table 1 – Human hair's Type I and Type II keratins and their location in the hair shaft (Schweizer *et al.*, 2006)

Name	Location	Type
K31	Cortex	Type I
K32	Cortex	
K33a	Cortex	
K33b	Cortex	
K34	Cortex	
K35	Cortex and Cuticle	
K36	Cortex	
K38	Cortex	
K39	Cortex and Cuticle	
K40	Cortex and Cuticle	
K81	Cortex	Type II
K82	Cuticle	
K83	Cortex	
K85	Cortex and Cuticle	
K86	Cortex	

The high sulfur proteins are divided according to their Cys content: the high sulfur proteins (HSP), with less than 30% of Cys, and the ultra-high sulfur proteins (UHSP), with more than 30% of Cys. With the evolution of the analytical methods, this group of proteins was classified as Keratin Associated Proteins (KAPs). A nomenclature was also stablished to classify the KAPs, KAPm.n, where m corresponds to a family or unique protein and n denotes a variant. Some KAPs present in human hair are categorized in **Table 2** (Gong et al., 2012; Plowman, J.E., Duane, 2018).

Table 2 – Known human hair keratin associated proteins (Gong *et al.*, 2012)

Class	Subfamily	Variants	Location
HSP	KAP1	4	Cortex
	KAP2	5	Cortex
	KAP3	3	Cortex
	KAP10	11	Cortex
	KAP11	1	Cortex
	KAP12	4	Cuticle
	KAP13	4	Cortex/Cuticle
	KAP15	1	Cortex/Cuticle
	KAP16	1	
	KAP23	1	Cortex/Cuticle
	KAP24	1	Cuticle
	KAP25	-	
	KAP26	1	Cuticle
	KAP27	1	
UHSP	KAP4	27	Cortex
	KAP5	4	Cuticle
	KAP9	7	Cortex
	KAP17	-	Cuticle

1.1.4. Pigments

The pigmentation of hair, skin, and eyes in animals is mainly a manifestation of the presence of melanin. In the hair and skin are two types of melanin (**Figure 7**) , the eumelanin that gives a black to brown coloration, and the pheomelanin that gives a yellow to reddish brown coloration (Ito and Wakamatsu, 2011; Cao *et al.*, 2021). The combination of these two types of melanin gives the hair its natural color.

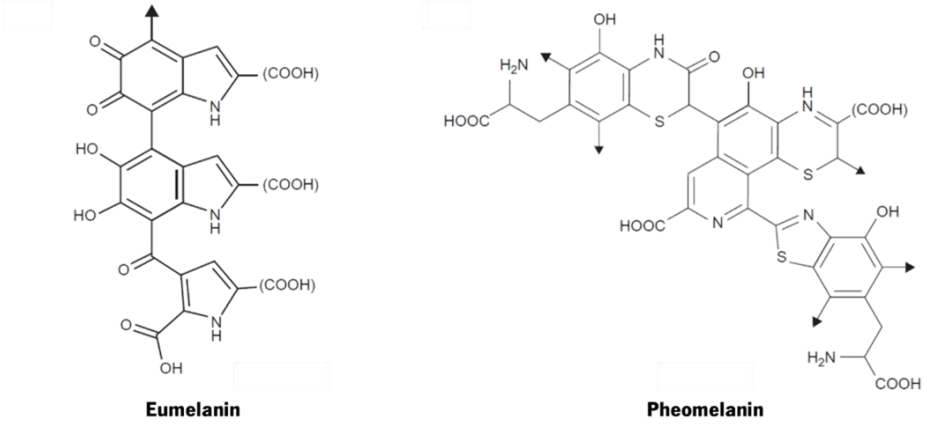

Figure 7 – Structure of Eumelanin and Pheomelanin, the two types of melanin present in hair. Adapted from Litwack, 2018.

1.1.5. Lipids

The human hair is composed by 1-9% of lipids (dry weight). The lipids are not located in a specific region of the hair, being scattered all over the fiber, and are mainly: cholesterol esters, free fatty acids, cholesterol, ceramides and cholesterol sulphate. They can be classified according to their origin, sebaceous glands or hair matrix, as exogeneous and endogenous respectively (Cruz *et al.*, 2013). One of the most important lipids in the hair is 18-MEA. The hydrophobicity of hair is related with the presence of this lipid and its removal may expose the hair to damage by increasing its surface hydrophilicity (Gavazzoni Dias, 2015). The lipids from hair can change the keratin fibers structure leading to changes in the texture of hair (Cruz *et al.*, 2013). The lipids also have a role in the dynamics of water in the fiber (Coderch *et al.*, 2019; Oliver *et al.*, 2019).

1.1.6. Hair ethnicity

Human hair can be classified according to its ethnic origin (Asian, Caucasian and African), color and degree of curliness. (Araújo *et al.*, 2010; Cruz *et al.*, 2016; Leerunyakul and Suchonwanit, 2020). The main differences between the three hair ethnicities are mainly related with the shape, appearance and the hair shaft diameter (**Table 3**). More recently, some differences were also found on the cysteine content and the type and amount of lipids present in the hair cortex of each hair ethnicity. This influences the degree of hair moisture and the tendency of a fiber to break, with the African hair presenting the lowest moisture content and the highest tendency to break (Cruz *et al.*, 2013). Moreover, it was also found that the hair follicles of African hair present less elastic fibers anchoring the hair follicles to the dermis, when compared with the other groups (Franbourg *et al.*, 2003; Richards *et al.*, 2003).

Table 3 – General characteristics of Asian, Caucasian and African hair according to Leerunyakul and Suchonwanit, 2020

	Asian	Caucasian	African
Hair ellipticity	Round	Ovoid	Flat and elliptical
Hair shape	Straight	Straight/Wavy/Curly	Curly
Cross-sectional area	4804 ± 159 μm^2	3857 ± 132 μm^2	4274 ± 215 μm^2
Lipid content	Polar lipid and free fatty acid	Polar lipid and free fatty acid	Apolar lipid
Hair pigmentation	Mainly eumelanin	Mainly eumelanin (in brown and blond hair); mainly pheomelanin (in red hair)	Mainly eumelanin

1.2. Proteins in Hair-Care Industry

With the increasing demand for new hair products, the hair care industry is becoming more interested in new alternatives to the current chemical formulations (Tinoco *et al.*, 2021b). Proteins and protein-based formulations are seen as excellent alternatives since they can be obtained using eco-friendly processes while exploring the functions of peptides and proteins (Tinoco *et al.*, 2018, 2021b). The development of novel protein-based ecofriendly hair formulations is thus viewed as a good alternative for mitigating or even avoiding the detrimental impact of some cosmetic chemicals, resulting in the development of cost-effective and sustainable solutions for the haircare sector (Tinoco *et al.*, 2021a). Proteins have been investigated by the hair cosmetic industry in recent years for the development of various protein-based cosmetic formulations (**Figure 8**) (Tinoco *et al.*, 2021a).

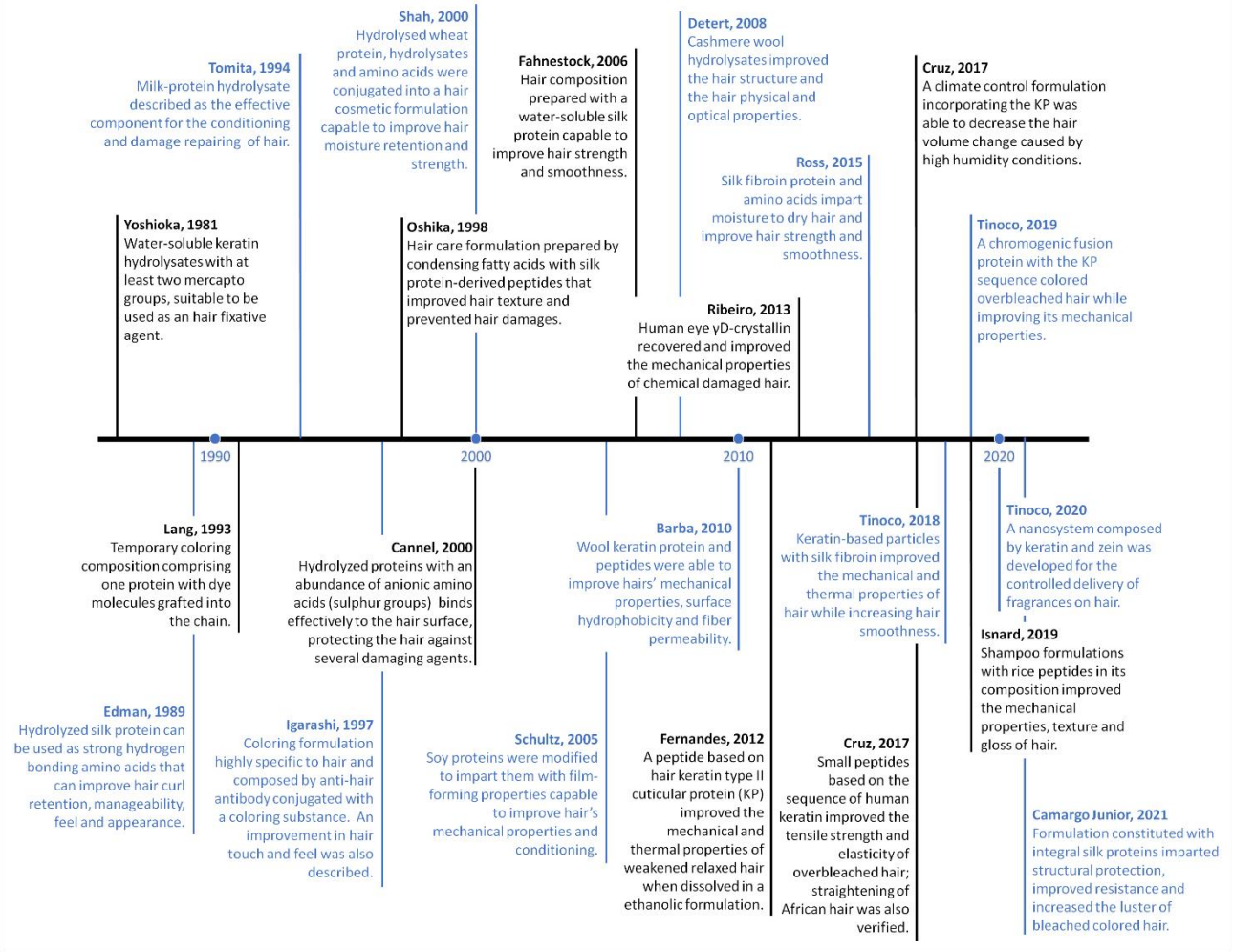

Figure 8 – A chronology of biotechnological applications of proteins/peptides in hair cosmetics over the years. Adapted from Tinoco *et al.*, 2021a

Based on their amphoteric and buffering properties, proteins have been added to hair care cosmetic formulations as conditioning agents, preventing and reducing the environmental damage (UV rays, brush and detangle the hair, etc.) caused to the hair fibers. Proteins also are very useful ingredients to generating an environment for healthy skin and hair, due of their ability to bind water with the horny layer skin and its annexes (Secchi, 2008; Tinoco *et al.*, 2018).

Hair damage caused by the use of chemical treatments, such as chemical straightening and perms, can be avoided, minimized and repaired using proteins and/or peptides.(Cruz *et al.*, 2017). Proteins and peptides can be used to increase hair's properties like hydration, brightness, softness, mechanical and the thermal properties. For example, the silk fibroin, a biocompatible and non-toxic natural polymer with a great water binding and absorbing capacity, have been explored to increase the moisturizing effect of hair formulations (Aarti V *et al.*, 2005; Villa *et al.*, 2013; Tinoco *et al.*, 2018).

Some proteins, like keratin, collagen, soy and silk fibroin, have been used in hair care cosmetic products due to their ability to form films on the hair surface (Benson Edward R. (Durham, 2012). The film-forming capacity of these proteins can be explored for example on the protection of hair during coloration procedures. Since these proteins have a low solubility, they are usually used in the form of proteins hydrolysates. Nevertheless, Fahnestock Stephen R. (2006) described a haircare formulation with water soluble silk proteins, where the protein was chemically modified to increase its solubility (Fahnestock Stephen R. (Wilmington, 2006).

Protein hydrolysates from cashmere wool have been used in hair-fixing formulations, to circumvent problems associated with the anionic or amphoteric polymers used in these formulations. These polymers tend to leave the hair brittle and sometimes with shedding particles (Detert Marion (Pommernweg 22, 2008). Protein hydrolysates from wheat, milk, soya and collagen with an abundance of anionic amino acids and, in particular, sulfur-containing amino acids, have also been used in haircare formulations. These proteins provide excellent finishing effects and protect the hair against environmental, chemical and grooming-associated damage (Sun Ziming (Fountain Valley, 2002).

Proteins and proteins hydrolysates could also contain at least one fatty chain. These chains can be added naturally or by chemical modification, such as grafting and quaternization. These proteins can be used in haircare formulations to improve the binding of ceramides to and/or in the hair fiber; provide stable aqueous dispersions and improve cosmetic properties as hold of the hairstyle (Cauwet-martin Daniele (Paris, 2000).

Cosmetic agents have been covalently linked to proteins and protein hydrolysates. The attachment of pigments to antibodies have been described has an approach to develop a new class of hair dyes (Huang Xueying (Hockessin, 2007). Horikoshi (1996) and Igarashi (1997) described a dye covalently attached to an anti-keratin antibody as a new hair coloring agent. Similarly Kizawa and co-workers also described a coloring agent composed by an antibody that recognizes the surface layer of hair coupled to a colored latex particle (Huang Xueying (Hockessin, 2007).

A fusion protein is a protein created by genetic engineering technologies, joining two or more genes that originally coded separate proteins. This process results in a protein with the functional properties derived from each of the original proteins (Schmidt, 2013).

The ability to combine several functions within the same sequence makes the fusion proteins excellent candidates for advanced biotechnological applications. Several fusion proteins have been reported in different areas such as pharmaceutical industry and enzymatic production (Schmidt, 2013).

Despite their great potential, there are few examples on the application of fusion proteins in hair cosmetic applications.

A fusion protein, constituted by a body surface binding peptide and a pigment-binding peptide, was developed to dye the hair and other body parts, such as skin and teeth. The patent describe a diblock peptide-based hair coloring agent having the general structure $[(HBP)_m\text{-}(PBP)_n]_x$, wherein HBP is the body surface binding peptide, and PBP is a pigment-binding peptide and m, n and x an independent coefficient that range from 1 to 10. In this invention is also described a triblock peptide-base hair coloring reagent having the general structure $[[(HBP)_m\text{-}S_q]_x\text{-}[(PBP)_n\text{-}S_r]_z]_Y$ wherein S is a molecular spacer, z and y is a coefficient that range from 1 to 10 and q and r is a coefficient that range from 1 to 5 (O'brien John P. (Oxford, 2010).

Some growth factors have been explored to stimulate hair growth. In this perspective, a fusion protein constituted by oleosin and the growth factor rhFGF9 was tested in mice. Fibroblast growth factor 9 (FGF9) is a heparin-binding growth factor, secreted by both mesothelial and epithelial cells, which participates in hair follicle regeneration. The human FGF9 (rhFGF9) was fused to the oleosin gene, and the olesin-rhFGF9 fusion protein was produced at large scale in *Arabidopsis thaliana* seeds and *Carthamus tinctorius L.* (safflower) seed. The fusion protein promoted the hair growth upon application, supporting its use on the development of hair cosmetic formulations. Interestingly the protein also showed potential to be used for wound healing (Yang *et al.*, 2015; Cai *et al.*, 2018).

The haircare business has created a variety of products to alter hair shape (Cruz *et al.*, 2017). However, the frequent chemical treatments have a harmful effect on the top layer of the cuticle, diminishing hair hydrophobicity and cysteine content. These processes damage the fibers, causing keratin structure loss and reduced hair mechanical properties, resulting in hair that is more dry and brittle, with less color and luster (Azizova Marina (New Canaan, 2016). Hair form is principally determined by disulfide bonds, salt connections, and hydrogen bonds. Because the normal form of hair is highly dependent on the orientation of disulfide bonds, their rupture and reorientation are required to achieve a long-term effect of the new hair shape (Azizova Marina (New Canaan, 2016; Tinoco *et al.*, 2021a).

Many publications (Dimotakis Emmanuel (Oradell, 2013; Azizova Marina (New Canaan, 2016; Song *et al.*, 2016; Tinoco *et al.*, 2021a) have identified amino acids and polyamino acids as active ingredients in cosmetic formulations to modify hair structure. A polylysine-containing formulation was able to keep the hair straight for at least 12 washing cycles, with much less damage to the fibers than chemical treatments (Azizova Marina (New Canaan, 2016). The polylysine produced a layer over the hair fibers, functioning as a barrier to vapor and moisture and preventing additional harm to the straightened hair.

Through changes in the protein structure, this compound was able to permanently straighten the hair, while the film offered some mechanical hold and protection from ambient humidity. This behavior was also observe by others authors (Dimotakis Emmanuel (Oradell, 2013).

Specific amino acids and peptides can also be utilized to replace some harmful substances used in aesthetic procedures. Cysteines, for example, can operate as reductants during the perming process, and when conjugated with polycarboxylic acids, it is feasible to get a hair-perming effect with less hair strength loss. This perming effect is dependent on the length of the polycarboxylic acids as well as the presence of extra carboxyl and hydroxyl groups (Song *et al.*, 2016). Small peptides based on the sequences of human keratins and keratin-associated proteins was studied as alternatives to the harsh chemicals used in hair straightening (Cruz *et al.*, 2017). These peptides straightened African hair by breaking down hair-protein disulfide links and forming new connections between the peptide thiol group and the hair proteins (Tinoco *et al.*, 2021a).

The capacity of proteins and peptides to form films over hair fibers is an important aspect during protein and peptide-driven hair perming. A novel perming technique investigated the *in situ* oligomerization of the tripeptide KCL catalyzed by α-chymotrypsin and the formation of a coating on the hair surface. Although α-chymotrypsin is frequently employed to hydrolyze proteins, it also works in the opposite way, resulting in the creation of peptide bonds. This process needs appropriate reaction conditions such as alkaline pH, temperature, and greater enzyme and substrate concentrations (Qin *et al.*, 2013; Yazawa and Numata, 2014; Tinoco *et al.*, 2021a). Because of its greater yields and lesser utilization of harmful chemicals, this protease may be employed in the polymerization of peptides and has the potential to replace the chemical procedures usually used for peptide oligomerization. Although this technology decreased perming time while maintaining the perming efficiency comparable to that seen in chemical formulations (Martins *et al.*, 2019; Tinoco *et al.*, 2021a). When a keratin-concentrate mixture including high- and low- molecular-weight keratins was applied to hair, a coating over the hair surface was also formed. Low-molecular-weight keratins may enter the hair cortex, but high-molecular-weight keratins produced a robust film over the hair surface, increasing style effects (Tinoco *et al.*, 2021a).

1.3. Protein production

1.3.1. *E. coli* as a biofactory

The use of proteins in biotechnological applications relies in their production and isolation from other cellular components. *Escherichia coli (E. coli)* is one of the hosts usually selected for the

recombinant expression of proteins. The cost-effective production of proteins in *E. coli* is associated with their fast growth rate, easiness to manipulate and ability to grow at different temperatures and using several nutrient sources contributing for cost-effective production of the proteins (Rosano *et al.*, 2019). For example, around 30% of approved therapeutic proteins are currently being produced using *E. coli* as a host biofactory(Baeshen *et al.*, 2015).

To produce recombinant/heterologous proteins the coding sequence (DNA) of interesse can be inserted into an expression vector and transformed in *E. coli*, using the bacteria as a microbial cell factory (**Figure 9**) (Baeshen *et al.*, 2015).

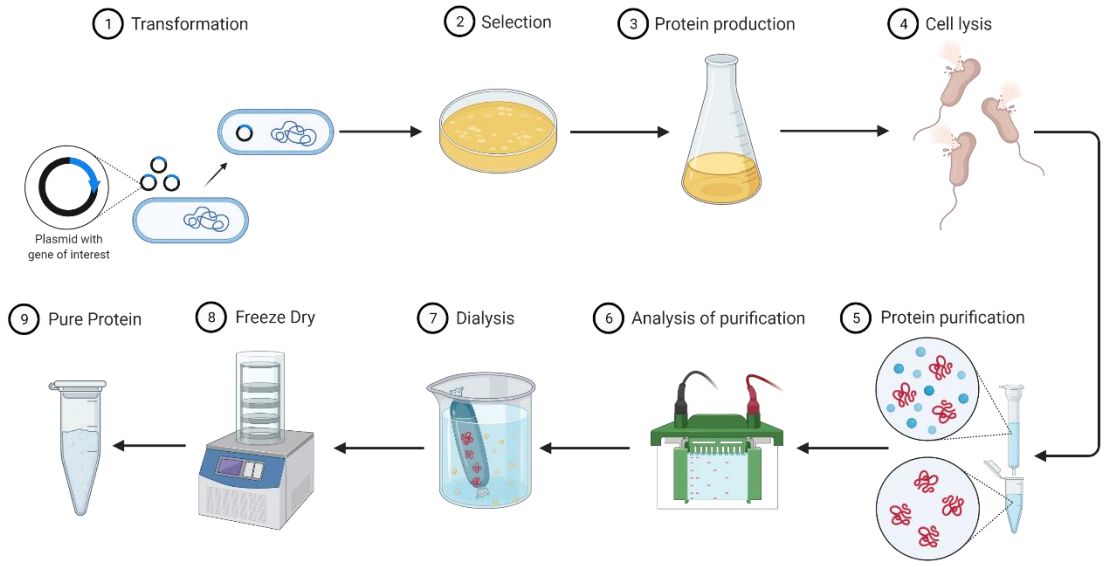

Figure 9 – General scheme of production of recombinant/heterologous proteins in *E. coli*.

With the biotechnological revolution on the 20th century several *E. coli* lines were tested for their characteristic for the heterologous expression of protein. The B line, especially the derivative BL21(DE3), given its features rapidly disseminated among the main biotechnological platforms for the expression of recombinant proteins. (Rosano *et al.*, 2019)

The BL21(DE3) strain is the go too strain for protein production. It carries a copy of phage T7 RNA polymerase (T7RNAP) controlled by the *lac* promoter (Rosano *et al.*, 2019). The T7 RNA allows higher translate rate of the interest sequence than the RNA polymerase naturally present in *E. coli* (Sousa, 2013). The sequence of interest is cloned next to the T7 promoter in the expression vector, and the expression of the heterologous protein begins upon the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG), that unlocks the T7 RNA polymerase (Sousa, 2013; Rosano *et al.*, 2019). This system allows the user full control of the induction protein synthesis with high selectivity and activity. With time, several

improvements were made, especially at the level of rare codon usage and dissulphide bonds. Nevertheless, the system based on the T7 RNAP and its promoter remained virtually unchanged. Making the BL21(DE3) strain, as aforementioned, the preferred host for protein production (Rosano *et al.*, 2019).

Altering the conditions of the culture is the simplest way to impact the growth of *E. coli* and directly alter the yield of recombinant proteins. The temperature of culture is one of the most common parameters altered in order to achieve better yields. Although LB medium is the standard to growth, it is far from being the best medium for protein expression. Richer broths, like Terrific Broth, and autoinduction media have gained popularity and there are many successful examples of proteins using these media in literature (OBP::Tat, OBP::Pep-1, OBP::pVEC (Gonçalves *et al.*, 2018), KP-UM (Tinoco *et al.*, 2019b), KP-Cryst Wt and KP-Cryst Mut (Tinoco *et al.*, 2019a)). The autoinduction medium consist of two or more carbon sources, glucose and lactose. Glucose is the preferred carbon source, and its normally depleted in the exponential phase, then lactose begins to be consumed, and activates the protein production system based on the lac promoter (Rosano *et al.*, 2019).

1.3.2. Industrial byproducts as a subtract to protein production

The culture medium is one of the most important factors when optimizing the expression of heterologous protein in *E. coli* (Ye *et al.*, 2010). Industry-by-products such whey, corn steep liquor (CSL) and molasse can be used as fermentation media for the production of heterologous proteins. These by-products beside being rich in sugars, minerals, amino acidic and other organic compounds can be obtained at low costs (Ye *et al.*, 2010; Pais *et al.*, 2014; Gudiña *et al.*, 2015b; Chen *et al.*, 2018).

Whey is the major by-product of cheese and casein production, representing 80 to 90% of the volume of transformed milk. The increasing worldwide cheese market alone generates $145x10^6$ ton of liquid whey every year. Half of the word`s cheese whey is discarded representing an important environmental problem (Pais *et al.*, 2014). Whey powder is the dried and concentrated form of whey and cost 20 to 401 cents per kg (Sar *et al.*, 2017). Cheese whey have 4.5% (w/v) lactose, 0.8% (w/v) protein, 1.0% (w/v) salts and 0.8% (w/v) lactic acid (Pais *et al.*, 2014; Sar *et al.*, 2017).

Molasses, a major by-product of sugar refinery, consists primarily of glucose, sucrose, fructose, minerals, amino acids, vitamins and nitrogen compounds.(He *et al.*, 2015) It can be used as a carbon source for fermentation in *E.coli* (Ye *et al.*, 2010). To date, molasses has been used to successfully produce L-lactic acid, succinic acid and benzaldehyde lyase (He *et al.*, 2015).

Corn steep liquor (CSL) is a low-cost agro-industrial by-product of starch production from the wet milling of corn (Ye *et al.*, 2010; Gudiña *et al.*, 2015b). It is rich in vitamins, peptides and amino acids

and normally is used as an inexpensive nitrogen source (Rivas *et al.*, 2004; Gudiña *et al.*, 2015a; Chen *et al.*, 2018). CSL has been used as a fermentative subtract in the production of ethanol by *Zymomonas mobilis* (Silveira *et al.*, 2001) and *Pichia stipites* (Amartey and Jeffries, 1994), and succinic acid by *Anaerobiospirillum succiniciproducens* (Lee *et al.*, 2000). And in the production of biosurfacans in *Bacillus subtilis* (Rivas *et al.*, 2004; Gudiña *et al.*, 2015b).

2.GOALS

Proteins are gaining interest as friendly alternatives to the chemicals commonly used in hair cosmetic industry. Particularly, the development and production of novel fusion proteins as new hair cosmetic products is a relatively unexplored field. In light of this, the potential of several proteins based on XZX protein for the development of novel hair cosmetic products should be investigated.

The focus of this work is the design, expression, optimization, purification and application of different XZX-based proteins. The proteins will be expressed in *Escherichia coli* (*E. coli*), and the optimization of the protein expression will be done using a combination of different culture media and expression temperatures. Using and experimental design approach, the optimization will explore the use of industrial by-products as subtract for fermentation combined with additives, in order to lower the final protein's production costs. After purification, the proteins XZX-based proteins will be characterized and applied to virgin Asian hair. The ability of these proteins to bind to the hair and to change its shape will be assessed. One of the main goals in hair cosmetic industry is to develop products to change the shape efficiently the shape of Asian hair. This type of hair, due to its properties, is very hard to be changed regarding its shape. This work aims to develop a protein-based green hair cosmetic product, with proven environmental and personal advantages, to replace the chemical-based products currently used.

3. MATERIALS AND METHODS

3.1. Optimization and production of XZX-based Proteins

3.1.1. Competent cells

In order to prepare competent cells, *E. coli* BL21(DE3) cells were inoculated in Lysogeny broth (LB) medium (Grisp, Portugal) and grown at 37 °C with constant agitation (170 RPM) overnight. The culture was diluted 1:100 in LB medium and grew until an OD_{600} of 0.6. The cells, previously incubated in ice for 10 min, were centrifuged at 2500 RPM, for 10 min at 4°C, and were gently resuspended in TB-buffer (10 mM HEPES, 15 mM $CaCl_2$, 250 mM KCl, 55 mM $MnCl_2$, pH 6.7). DMSO was added to the cells suspension to a final percentage of 7%, and the cells were incubated in ice for an additional 10 min. Aliquots of 100 µL were prepared and immediately frozen in liquid azote, and kept at -80°C until further use.

3.1.2. Transformation

E. coli BL21(DE3) competent cells were transformed with the pET-28 a (+) vectors (GenScript(USA).) harboring the genes coding for the XZX-based proteins. After adding the vectors (1 µL), the cells were incubated in ice for 1h. Then a thermal shock at 42 °C for 2 min followed by ice for 2 min was applied. LB medium pre-heated at 37 °C was added and the cells were incubated for 1h at 37 °C with constant agitation (170 RPM). The cells where then platted in LB-agar plates supplemented with kanamycin (0.05 mg/mL).

The theorical molecular weight and isoelectric point of each recombinant proteins coding by the aforementioned vectors are present in **Table 4**.

Table 4 – Theorical molecular weight of each recombinant protein.

Protein	XZX-1	XZX-2	XZX-3	XZX-4	XZX-5	XZX-6	XZX-7
Molecular weight (kDa)	30.35	20.70	20.74	20.18	20.10	19.89	22.75

3.1.3. Protein Expression Screening for the determination of the best commercial medium

The effect of different commercial media on protein expression was assessed at two different temperatures, 25 and 37 °C, for the XZX (1-6) proteins and at 37°C for the XZX–7 protein. The media tested were: Lysogeny Broth with IPTG (LB+IPTG), auto induction Lysogeny Broth (LB-AIM), auto induction

Terrific Broth (TB-AIM), TB-AIM supplemented with 1% glucose and auto induction Super Broth (SB-AIM) (**Table 5**). The TB-AIM and SB-AIM were only tested for the XZX–7 production. All the commercial media were purchased from Grisp, Portugal.

Table 5 – Composition of Lysogeny Broth (LB), auto induction Lysogeny Broth (LB-AIM), auto induction Terrific Broth (TB-AIM) and auto-induction Super Broth (SB-AIM) mediums

	LB	**LB-AIM**	**TB-AIM**	**SC-AIM**
Tryptone (g/L)	10.00	10.00	12.00	35.00
Yeast extract (g/L)	5.00	5.00	24.00	20.00
NaCl (g/L)	5.00	_____	_____	_____
MgSO4 (g/L)	_____	0.14	0.15	0.15
(NH4)2SO4 (g/L)	_____	3.30	3.30	3.30
KH2PO4 (g/L)	_____	6.80	6.50	6.80
Na2HPO4 (g/L)	_____	7.10	7.10	7.10
Glucose (g/L)	_____	0.50	.50	0.50
Lactose (g/L)	_____	2.000	2.00	2.00

A pre-inoculum of four transformants of each protein was prepared by growing the transformants in LB medium supplemented with kanamycin (0.05 mg/mL) overnight at 37 °C, 170 rpm. The OD_{600} of each transformant suspension was adjusted to 0.1 in the commercial culture media to be tested, using the pre-inoculum grown overnight. The volume of pre-inoculum necessary to prepare the inoculum was calculated using **Equation 1**.

$$\mathbf{OD_i \times V_i = OD_f \times V_f} \qquad \textbf{Equation 1}$$

Where OD_i is the OD_{600} of pre-inoculum; V_i the volume of pre-inoculum necessary to prepare the inoculum; OD_f is the final OD_{600} = 0.1; and V_f is the culture volume of the inoculum.

The cells inoculated in LB-AIM, TB-AIM and SB-AIM supplemented with kanamycin (0.05 mg/mL) were grown at different temperatures (25 and 37 °C for the XZX (1-6)proteins and at 37°C for the XZX–7 protein), for 24h (Tinoco *et al.*, 2019b). For IPTG induction, the transformants were grown at 37 °C until an OD_{600} equal to 1, and protein expression was induced by the addition of IPTG (0.1 mM) to the culture media. After induction, the cells were grown for 4 and 24 h at 37°C and 170 rpm. Protein

expression was evaluated by SDS-PAGE. Using the ImageJ software, the amount of protein expressed by an OD_{600} of 0.1 was calculated.

3.1.4. Sodium Dodecyl Sulfate- Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), running gel of 12.5% or 15% and staking gel of 4%, were performed to evaluate the expression of XZX-based proteins. The samples were prepared with 2x Sample Loading Buffers, followed a denaturation step at 100°C for 5 min. Molecular weight markers (Grisp), GRS Unstained Protein Marker for quantifications and GRS Protein Marker Blue for the rest, and the samples were loaded onto the gel and separated by electrophoresis (20 mA per gel). The gel was then stained with a Coomassie solution (methanol 50% (V/V), acetic acid 10% (V/V), Coomassie Brilliant Blue G-250 2.5g/L) for 40 min followed by distaining solution (methanol 15% (V/V), acetic acid 10% (V/V)) for 1.5 hours. Protein expression was analyzed and quantified using the ImageJ software all the quantifications were calculated in milligrams to a OD_{600} equal to 0.1.

3.1.5. Experimental design to determine the best conditions for XZX–7 protein expression

3.1.5.1. Best mixture of industrial by-products

The use of whey, CSL and molasses as induction media for protein expression was evaluated recurring to an experimental design with the help of DesignExpert® software (StatEase®(USA)).

The matrix was designed to a concentration between 10 and 40 g/L for TB-AIM, and between 0 and 30 g /L for whey, CLS and molasses, in a culture media concentration of 40 g/L. DesignExpert® software was used to design the matrix to evaluate the effect on protein expression of the different culture media mixtures (**Table 6**).

Table 6 – Concentrations of the different industrial by products and TB-AIM, given by DesignExpert®, in order to test the influence of each component on the expression of XZX–7 protein

Run	**TB-AIM** (g/L)	**Wey**(g/L)	**Corn Syrup** (g/L)	**Molasses** (g/L)
1	13.75	3.75	3.75	18.75
2	10	15	0	15
3	25	15	0	0
4	17.5	7.5	7.5	7.5
5	13.75	3.75	18.75	3.75
6	10	0	0	30
7	10	0	30	0
8	40	0	0	0
9	13.75	18.75	3.75	3.75
10	10	0	30	0
11	10	0	0	30
12	25	15	0	0
13	10	15	15	0
14	40	0	0	0
15	28.75	3.75	3.75	3.75
16	10	30	0	0
17	10	0	15	15
18	25	0	15	0
19	25	0	0	15
20	10	30	0	0

An *E. coli* BL21(DE3) transformant expressing the XZX–7 protein was incubated overnight in LB medium supplemented with kanamycin at 37°C and 170 rpm. The culture media mixtures were supplemented with kanamycin (0.05 mg/mL), inoculated to a starting OD_{600} of 0.1**,** and were further incubated at 37°C and 170 rpm for 24 hours. The final OD_{600} was recorded at the end of protein expression. XZX–7 yield was evaluated by SDS-PAGE (section 3.1.4) with all the samples normalized to an OD_{600} of 0.1. The final OD_{600} and the protein yield determined for each condition were analyzed using the DesignExpert® and three culture media mixtures were selected. The expression of XZX–7 protein using these mixtures was confirmed by SDS-PAGE (section 3.1.4).

3.1.5.2. Best temperature and pH

To further optimize the production of XZX–7 protein several temperatures and pH values were tested: 20°C, 25°C, 30°C, 37°C, 40°C and pH 6, 7 and 8.

XZX–7 producing *E. coli* cells were incubated overnight in LB medium supplemented with kanamycin at 170 rpm. The best mixture of culture media (selected in 3.1.5.1.) was adjusted to the three pH values (pH = 6, 7 or 8). The culture media was inoculated with the XZX–7 producing *E. coli* cells to an OD_{600} of 0.1 according to **Equation 1.** The cells were incubated at 20 °C, 25 °C, 30 °C, 37 °C and 40 °C and 170 rpm for 24 hours. The XZX–7 yield was evaluated by SDS-PAGE and determined using ImageJ software.

3.1.5.3. Best additives

To maximize the production of XZX–7 protein, several additives (NaCl, glycerol, glucose and sorbitol) were added to the best culture media mixture and their effect on protein expression was evaluated by an experimental design using the DesignExpert® software (StatEase®(USA)).

The matrix was design to a concentration between 0 and 20 g/L for NaCl, glycerol and glucose, and between 0 and 10 g/L for sorbitol. DesignExpert® software was used to design the matrix to evaluate the effect on protein expression of the different additives (**Table 7**).

Table 7 – Concentrations of the different additives, given by DesignExpert®, to test their influence on XZX–7 protein expression

Run	**NaCl** (g/L)	**Glycerol** (g/L)	**Glucose** (g/L)	**Sorbitol** (g/L)
1	10	10	10	5
2	10	10	10	5
3	0	20	0	0
4	10	10	10	5
5	0	0	0	10
6	10	10	10	5
7	10	0	10	5
8	10	10	10	5
9	10	10	10	10
10	20	0	0	0
11	10	10	10	5

Table 7 – Concentrations of the different additives, given by DesignExpert®, to test their influence on XZX–7 protein expression (**continued**)

Run	**NaCl** (g/L)	**Glycerol** (g/L)	**Glucose** (g/L)	**Sorbitol** (g/L)
12	10	10	10	5
13	10	10	10	5
14	10	10	10	5
15	20	10	10	5
16	10	10	0	5
17	20	20	0	10
18	10	10	10	5
19	10	10	10	5
20	10	20	10	5
21	10	10	10	5
22	0	20	0	10
23	10	10	10	5
24	20	20	20	0
25	20	0	20	0
26	10	10	10	5
27	10	10	10	5
28	10	10	10	0
29	10	10	10	5
30	10	10	10	5
31	10	10	20	5
32	10	10	10	5
33	20	20	20	10
34	0	20	20	0
35	0	10	10	5
36	0	0	20	0
37	20	20	0	0
38	10	10	10	5
39	0	0	20	10
40	10	10	10	5

Table 7 – Concentrations of the different additives, given by DesignExpert®, to test their influence on XZX–7 protein expression (**continued**)

Run	NaCl (g/L)	Glycerol (g/L)	Glucose (g/L)	Sorbitol (g/L)
41	10	10	10	5
42	0	0	0	0
43	10	10	10	5
44	20	0	0	10
45	0	20	20	10
46	20	0	20	10
47	10	10	10	5
48	10	10	10	5

The different combinations of additives determined by the matrix were added to the culture media mixture previously optimized (3.1.4.1 and 3.1.4.2 sections).

XZX–7 producing *E. coli* cells were incubated overnight in LB medium supplemented with kanamycin at 37°C, 170 rpm. The pre-optimized culture media with kanamycin was supplemented with the additives, and the mixtures were inoculated to an initial OD_{600} of 0.1 and incubated at 37°C and 170 rpm for 24 hours. The final OD_{600} was measured and the protein yield was evaluated by SDS-PAGE with the samples normalized to an OD_{600} of 0.1 for better comparison between the conditions. The final OD_{600} and the protein yield determined for each condition were analyzed using the DesignExpert® and three optimal culture media mixtures were selected. The expression of XZX–7 protein using the optimal culture media mixtures with additives was confirmed by SDS-PAGE.

The addition of ethanol can improve the production of recombinant proteins in *E. coli.* To study the effect of ethanol on the expression of XZX–7 protein three ethanol concentrations (0, 1, 2 and 3%) were added to optimal culture media, previously optimized. The effect of ethanol concentrations on protein yield was evaluated has described in 3.5.1.2. For this experiment the DesignExpert® was not applied (Chhetri *et al.*, 2015).

3.1.6. Growth kinetics of XZX–7 producing *E. coli* in the optimized media conditions

A calibration curve was obtained by measuring the OD_{600} and the respective dry weight (OD_{600} vs dry weigh) of XZX–7 expressing *E. coli* cells. In order to obtain the calibration curve, an overnight culture of

XZX–7 expressing *E. coli* was used to prepared solutions with growing OD values. These solutions were filtered and the dry weight was determined.

To determine the growth kinetics in the optimized culture media conditions, each media was inoculated to a starting OD_{600} of 0.1 and the absorbance was measured during 8 hours. The specific growth rate (μ) and the doubling time were calculated.

$$\mu = \frac{dX}{dt} \qquad \textbf{Equation 2}$$

$$t_d = \frac{\ln 2}{\mu} \qquad \textbf{Equation 3}$$

Where μ is the specific growth rate, dX is the variation of biomass, dt is the variation of time and t_d is the doubling time

3.1.7.Production of XZX-based Proteins

The XZX–7 protein was expressed using the optimized medium (10 g/L TB-AIM, 30g/L CSL, 0.934 g/L NaCl, pH 7). For the expression of the XZX–1 to 6 proteins, *E. coli* cells previously selected regarding the best protein yield, were incubated overnight in LB medium supplemented with kanamycin at 37°C with constant agitation (170 rpm). LB medium (0.5 L) supplemented with kanamycin was inoculated to an OD_{600} of 0.1 using the pre-inoculum, and the cells were incubated at 37 °C and 170 rpm until reaching an OD_{600} between 0.6 and 1. Afterwards the expression of the XZX–1 to 6 proteins was induced with IPTG (0.1mM) and the cells were incubated for an additional 24 hours at 37 °C and 170 rpm.

After fermentation the cells were harvested by centrifugation (4000 RPM, 4°C, 10 min) and sonicated (30 min with 3 ON, 9 OFF cycles). The total fractions were then separated into soluble and insoluble fractions by a centrifugation step (10000 rpm, 4°C, 40 min).

The insoluble fraction was resuspended in phosphate buffer (20 mM NaH_2PO_4, 500 mM NaCl, pH 7.4) and the protein of interest was purify by using Nickel magnetic beads with specificity to the His-tag sequence present in the N-terminal of the proteins. The purification process was controlled by SDS-PAGE (section 3.1.4) and the purified proteins solutions were dialyzed against distilled water for 4 days. After dialyze the proteins were freeze dried for 4 days.

3.1.8.Characterization of XZX-7 protein

3.1.8.1. Fourier Transform Infrared Spectroscopy (FTIR)

Fourier Transform Infrared Spectroscopy (FTIR) was used to characterized chemical structure of XZX-7 protein. The spectra were collected using a Bruker Alpha II (Massachusetts, USA) using Opus 8.22.28 software. Samples were placed directly on the crystal, and spectra were collected between 400 and 4000 cm^{-1} wavenumbers at a resolution of 2 cm^{-1}.

Convoluted FTIR curves corresponding to the Amide I spectral interval of [1600; 1720] cm1 were analyzed in the "Feat Peaks (Pro)" procedure of the "Peak Analyzer" menu in OriginPro software, v.8.5.0 (OriginLab Corporation, USA). Prior to fitting, no smoothing was applied to each convoluted curve. Following the baseline subtraction, a multiple pass fit was performed. A Second Derivative built-in-method was utilized to identify discrete starting peak spectral locations for fitting. Based on the provided peak assignments, the secondary conformational data arising from individual peaks were derived (KONG and YU, 2007; Ribeiro *et al.*, 2016).

3.1.8.2. Differential scanning calorimetry (DSC)

Differential scanning calorimetry (DSC) measurements were performed using stainless steel capsules in the temperature range of 20–250°C (heating rate: 5°C/min, sample weight: 2-3 mg) on a power-compensated DSC equipment (DSC 6000 Perkin Elmer) with a nitrogen flow of 20 mL/min.

3.2. Hair treatment with the XZX–7 protein

3.2.1. Binding assay to virgin Asian hair

A preliminary assay to assess the ability of XZX–7 protein to bind to hair was performed. Approximately 100 mg of XZX–7 inclusion bodies were suspended in 5 mL of deionized water or in 5 mL of an ethanolic formulation (phosphate buffer 0.1 M, 1.5 % (V/V) propylene glycol, 0.5 % Benzyl alcohol, 10% ethanol). Hair meshes with 150 mg of virgin Asian hair were incubated for one hour with the suspension of the inclusion bodies, previously prepared in deionized water and in the ethanolic formulation. After running the samples in a polyacrylamide gel the ImageJ software was used to determine the binding efficiency of XZX–7 protein to the hair fibers according to **Equation 4**.

$$\textbf{Binding efficiency}\ (\%) = \frac{[XZX\text{-}7]_I - [XZX\text{-}7]_F}{[XZX\text{-}7]_I} * 100 \qquad \textbf{Equation 4}$$

Where $[XZX\text{-}7]_I$ is the concentration of XZX-7 prior the incubation with hair and $[XZX\text{-}7]_F$ is the concentration of XZX-7 after incubation with hair

3.2.2. Hair perming with XZX–7 protein

The ability of the XZX–7 protein to perm virgin Asian hair was assessed with and without a pre-treatment step. Half of the virgin Asian hair bundles were pretreated prior application of XZX–7 protein to enhance the protein binding to the hair. Hair meshes of the same length and weight (mg) were first incubated in 2 mL of a 2 M urea solution, pH 9.5, for 20 minutes before being rinsed and dried with a towel. After, the hair meshes were incubated for 20 minutes in 2 mL of an ethanolic formulation (0.1 M phosphate buffer (pH 5) with 1.5 % (v/v) propylene glycol, 0.5 % (v/v) benzyl alcohol, and 10% (v/v) ethanol). Afterwards the hair meshes were gently dried with a towel, and 2 mL of a XZX–7 inclusion bodies solution (20 or 1 mg/mL in water or in ethanolic formulation) were applied to the hair for 20 minutes at room temperature. Hair without treatment was used as control. A commercial kit for curling (FarmaVita® Life the Perm) (water, ammonium thioglycolate, ammonium bicarbonate, ethanolamine, PEG-40 hydrogenated castor oil, cocamidopropyl betaine, ammonium hydroxide, polyquaternium-7, simethicone, styrene/vp copolymer and tetrasodium EDTA) and a neutralizing solution (FarmaVita® Universal Neutralizer) (water, SDS, hydrogen peroxide, simethicone, phosphoric acid and oxyquinoline sulfate)) was also used as control of the perming ability of XZX–7 protein. A BaByliss (BaByliss ®, Paris) of 19 mm previously heated at 200 °C, was used for 50 s to curl the treated hair meshes. To compare the initial results, all of the hair meshes were curled around the BaByliss the same number of times and in the same direction(Tinoco *et al.*, 2021c) (**Figure 10**).

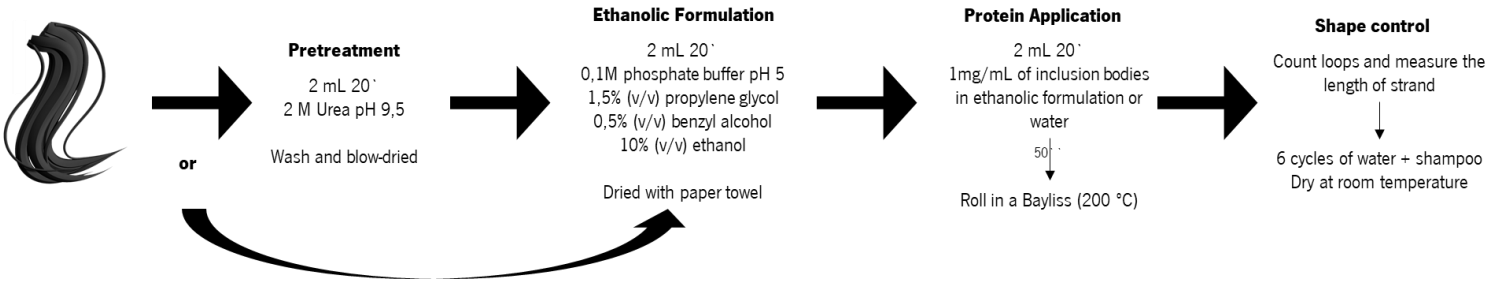

Figure 10 – Scheme of the protocol to assess the perming efficiency of XZX.

The perming efficiency of curled Asian hair was assessed following BaByliss application and throughout six washing cycles using tap water and a commercial shampoo (Pantene® Basic). The washing cycles were carried out without modifying the pH of the washing solution. The meshes were air-dried at room temperature after washing, and the perming efficiency was measured after the meshes were fully dry. The perming efficiency was calculated using the **Equation 5**.

$$\textbf{Perming efficiency (\%)} = \frac{\frac{\text{number of loops after perm}}{\text{fiber length afeter perm}}}{\left(\frac{\text{number of loops before perm}}{\text{fiber length before perm}}\right)} \times 100 \qquad \textbf{Equation 5}$$

4. RESULTS AND DISCUSSION

4.1. Optimization and production of XZX-based proteins

4.1.1. Protein expression screening for the determination of the best commercial medium

4.1.1.1. Screening of XZX–1, XZX–2, XZX–3, XZX–4, XZX–5 and XZX–6 proteins

In **Table 8** is presented the effect of the tested commercial culture media (TB-AIM and LB) in combination with different temperatures (25 °C and 37 °C) and additives (glucose and IPTG) on the expression of XZX–1, XZX–2, XZX–3, XZX–4, XZX–5 and XZX–6 proteins).

Table 8 – Culture conditions where the recombinant proteins were and were not expressed.

Protein / Conditions	XZX–1	XZX–2	XZX–3	XZX–4	XZX–5	XZX–6
TB-AIM (37ºC|w/glucose)	×	✓	×	×	×	✓
TB-AIM (37ºC|w/o glucose)	×	✓	×	×	×	×
TB-AIM (25ºC|w/o glucose)	×	×	×	×	×	×
LB+IPTG (4h)	×	×	×	×	×	×
LB+IPTG (24h)	×	✓	✓	✓	✓	✓

With the exception of XZX–1protein, all the XZX-based proteins tested in this experiment were expressed in at least one of the culture media conditions evaluated. XZX–1has a high repetitive sequence with several cysteine residues. Since *E. coli* is a simple organism, their cytoplasmic redox potential does not encourage the formation of intracellular disulfide bonds, hindering the expression of proteins with a high amount of cysteine residues. This results in the absence or in very low expression levels and in proteins with reduced solubility and consequent formation of inclusion bodies (Berkmen, 2012; Gutiérrez-González *et al.*, 2019).

The XZX–2 protein was successful expressed in TB-AIM medium with and without 1% glucose at 37°C (**Figure 11- A and B**). However, this protein expressed better in TB-AIM without 1% glucose at 37°C (0.00029 mg for a OD_{600} equal to 0.1) than in TB-AIM with glucose (0.00022 mg). This can be explained by the presence of glucose on the culture medium. Glucose is a simple and easily metabolized

monosaccharide, and its presence on the culture media retard the intake and metabolization of lactose (disaccharide). Since lactose is the induction factor on the TB-AIM medium, a delay on its consumption might delay the expression of the recombinant protein, which can favor the expression of proteins with a repetitive sequence and a high number of cysteine residues (Fox and Blommel, 2009; Zhang *et al.*, 2017). XZX–2was also expressed in LB medium with a 24 hour period of induction with IPTG (**Figure 11 C**). When the temperature was decreased to 25 °C the expression of the XZX–2 protein was affected and no expression was detected for any of the conditions tested (data not shown).

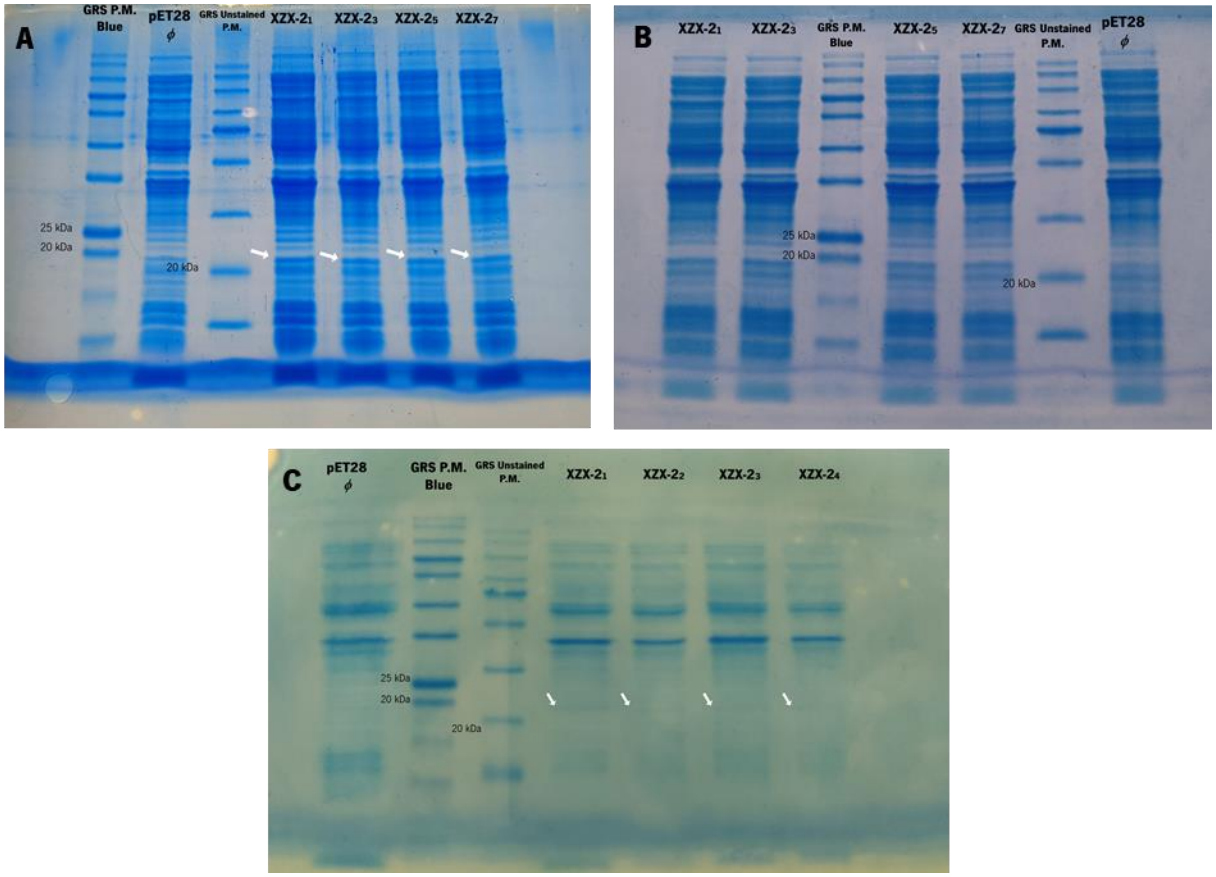

Figure 11 – SDS-PAGE electrophoresis (12%) of XZX–2expression in different culture media conditions: **A –** TB-AIM medium with 1% glucose at 37 °C; **B –** TB-AIM without 1% glucose at 37 °C; **C –** LB medium with 24 hour IPTG induction period at 37 °C. The XZX-based proteins are highlighted with an arrow.

The XZX–3 and XZX–4 proteins were expressed exclusively in LB medium with a 24 hour IPTG induction period (**Figure 12 A and B**). The inability of *E. coli* cells to metabolized IPTG, allowing longer induction times might help the expression of recombinant proteins with repetitive sequences and high percentage of cysteine residues (Fox and Blommel, 2009; Zhang *et al.*, 2017).

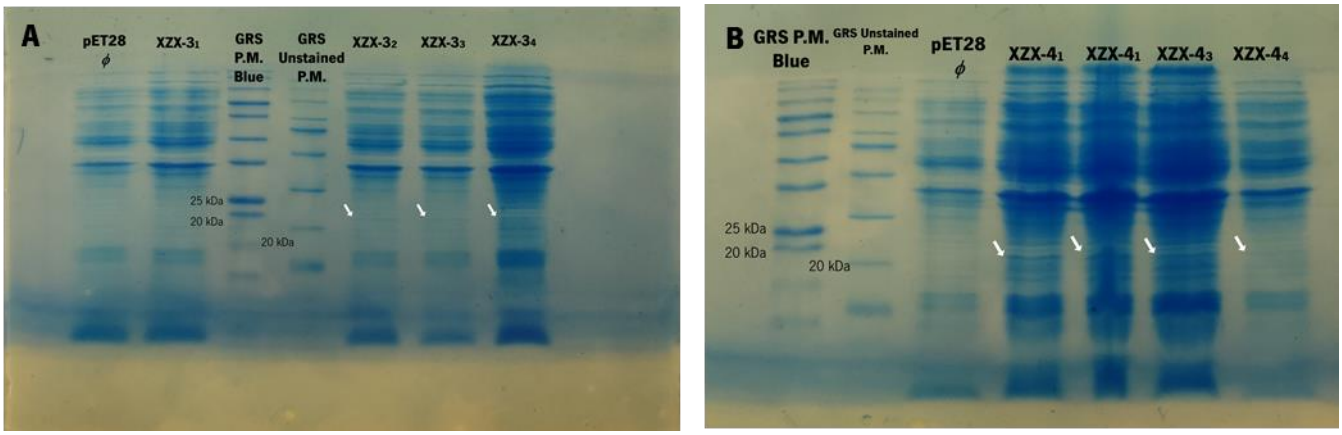

Figure 12 – SDS-PAGE electrophoresis (12%) of XZX–3 (**A**) and XZX–4 (**B**) expression in LB medium with a 24-hour IPTG induction period at 37°C. The XZX-based proteins are highlighted with an arrow.

The XZX–5 and the XZX–6 proteins were expressed in LB medium with a 24 hour of IPTG induction (**Figure 13 and Figure 14 A**). Moreover, the XZX–6protein was also expressed in TB-AIM with 1% glucose (**Figure 14 B**). Although the amino acidic of the two proteins is very similar, the single alteration in the aminoacidic sequence can alter the metabolic burden of *E. coli* cells leading to a reduction or inhibition of protein expression (Rahmen *et al.*, 2015).

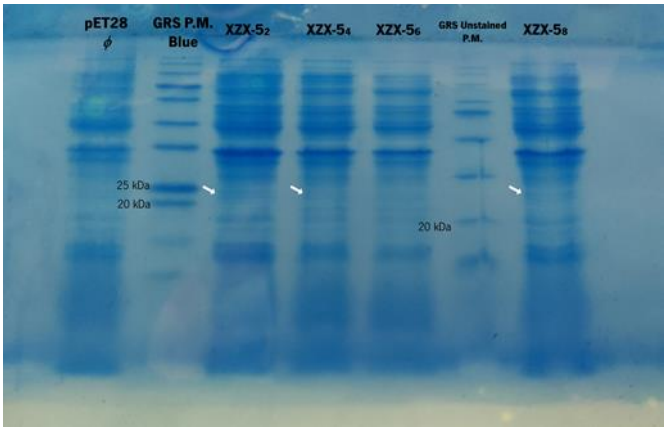

Figure 13 – SDS-PAGE electrophoresis (12%) of XZX-5 expression in LB medium with a 24 hour IPTG induction period at 37 °C. The XZX-based proteins are highlighted with an arrow.

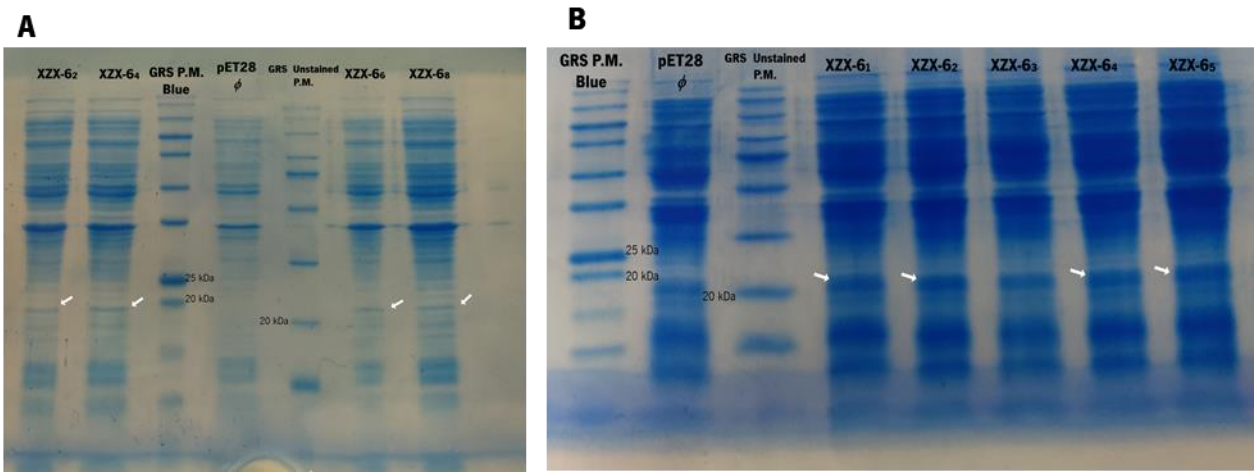

Figure 14 – SDS-PAGE electrophoresis (12%) of XZX-6 expression in: **A** - LB medium with a 2 -hour IPTG induction period at 37 °C; **B** - TB-AIM medium with 1% glucose at 37 °C. The XZX-based proteins are highlighted with an arrow.

4.1.1.2. XZX–7 protein expression

The expression of XZX–7 protein was evaluated in four different commercial media, at 37°C and 170 rpm: TB-AIM, SB-AIM, LB-AIM and LB with IPTG induction (4h) (**Figure 15 and Table 9**).

XZX–7 protein was successfully expressed in all the tested media, with the highest yield (0,0012 mg) being obtained for the TB-AIM. This medium was selected for further studies regarding the optimization of XZX–7 expression using industrial by-products (whey, CLS and molasses), additives (glucose, glycerol, sorbitol and NaCl) and testing several temperatures (20, 25, 30, 37 and 40 °C) and culture medium pH (6, 7 and 8).

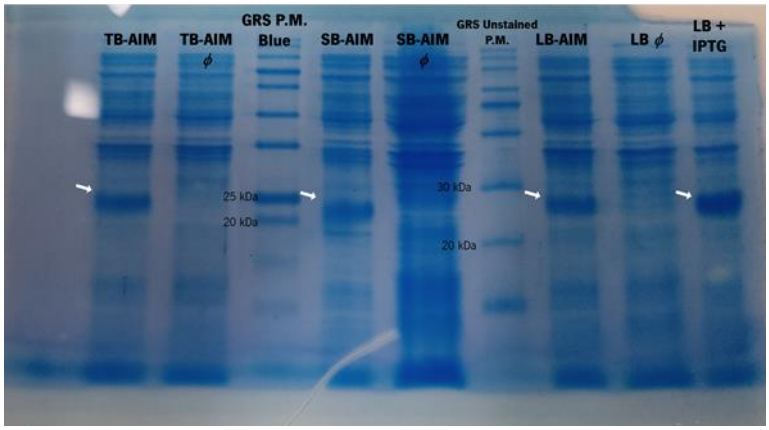

Figure 15 – SDS-PAGE electrophoresis (15%) of XZX–7 expression in TB-AIM, SB-AIM, LB-AIM and LB+IPTG media at 37°C. The XZX protein is highlighted with an arrow.

Table 9 – Yield (mg for an OD600 equal to 0.1) of XZX–7 protein expression for each medium tested

Medium	**Yield** (mg for OD_{600}=0.1)
TB-AIM	0.0012
SB-AIM	0.0001
LB-AIM	0.0007
LB + IPTG	0.0011

4.1.2. Experimental design to determine the best conditions for XZX–7 protein expression

4.1.2.1. Best mixture of industrial by-products

The final yield and the production costs are key parameters when developing industrial, medical, and cosmetic applications based on recombinant proteins. One of the goals of this work was to maximize the expression of XZX–7 protein by replacing commercial culture media by industrial by-products. Due to difficulties observed during the expression screening of the XZX–1, XZX–2, XZX–3, XZX–4, XZX–5 and XZX–6 proteins the optimization of the expression conditions were focused on the XZX–7 protein.

For this optimization, the commercial medium selected was the TB-AIM medium (highest protein yield comparing to the other commercial media). Different mixtures of culture media (TB-AIM + Industrial by-products or only industrial by-products) were tested according to the matrix (**Table 6**) obtained with the DesignExpert® software. After expression, the final OD_{600} was recorded and the expression yield was evaluated by SDS-PAGE using the ImageJ software and the reference band of the GRS unstained protein marker (**Figure 16 and Table 10**).

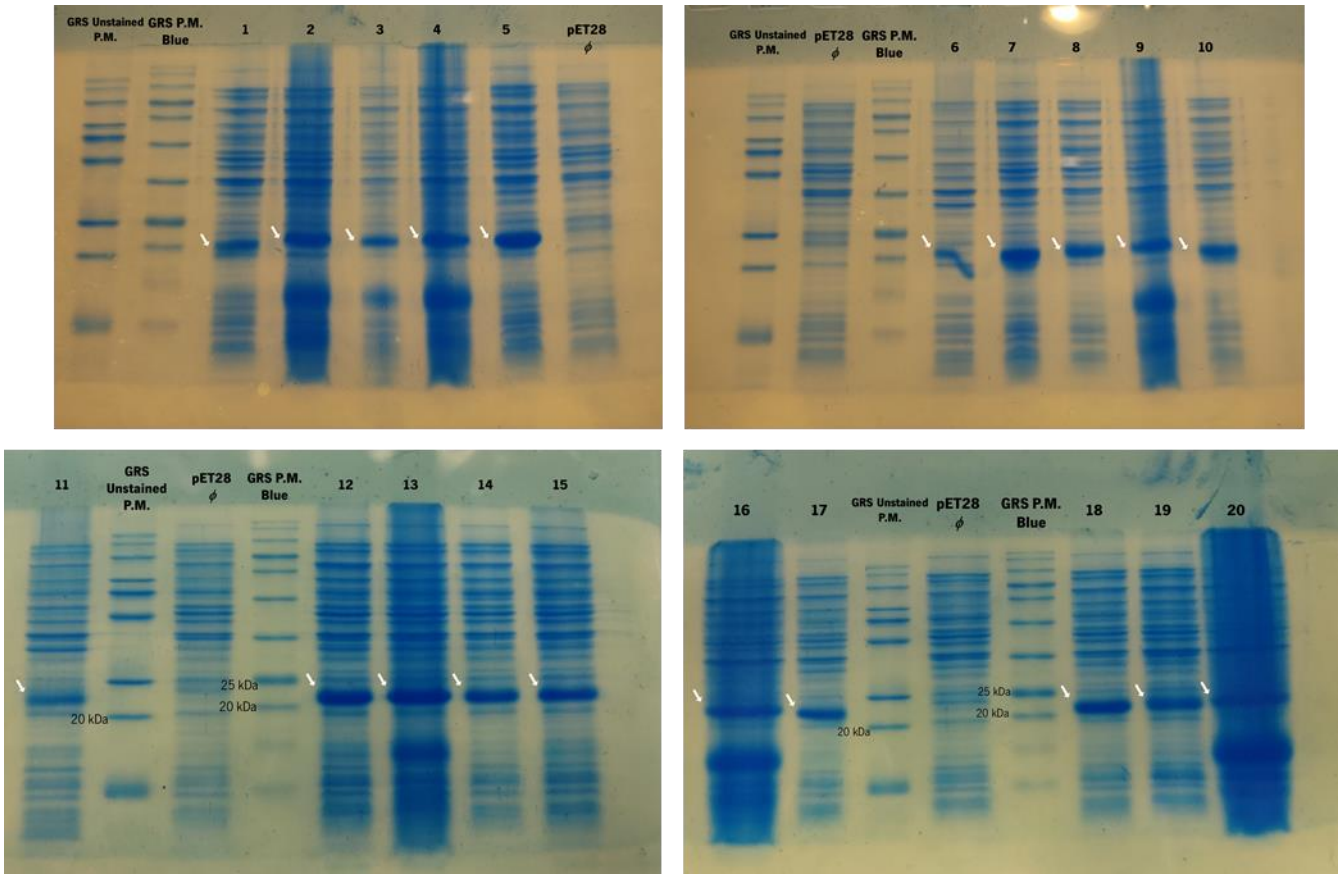

Figure 16 – SDS-PAGE electrophoresis (15%) of the expression conditions (culture media mixtures) dictated by the DesignExpert® matrix to evaluate the effect of industrial-by products in XZX–7expression. The composition of each condition is described in **Table 6**. The XZX–7 protein is highlighted with an arrow.

Table 10 – Composition of the different mixtures of culture media and their XZX expression yield (mg for a OD_{600} equal to 0.1)

Mixture	[TB-AIM] (g/L)	[Wey] (g/L)	[Syrup] (g/L)	[Molasses] (g/L)	XZX–7 Yield (mg for OD_{600}=0.1)
1	13.75	3.75	3.75	18.75	**0.0014**
2	10	15	0	15	**0.0023**
3	25	15	0	0	**0.0008**
4	17.5	7.5	7.5	7.5	**0.0021**
5	13.75	3.75	18.75	3.75	**0.0020**
6	10	0	0	30	**0.0008**
7	10	0	30	0	**0.0032**
8	40	0	0	0	**0.0013**
9	13.75	18.75	3.75	3.75	**0.0017**

Table 10 – Composition of the different mixtures of culture media and their XZX expression yield (mg for a OD600 equal to 0.1) (**Continued**)

Mixture	[TB-AIM] (g/L)	[Wey] (g/L)	[Syrup] (g/L)	[Molasses] (g/L)	XZX–7 Yield (mg for OD_{600}=0.1)
10	10	0	30	0	**0.0017**
11	10	0	0	30	**0.0011**
12	25	15	0	0	**0.0021**
13	10	15	15	0	**0.0025**
14	40	0	0	0	**0.0016**
15	28.75	3.75	3.75	3.75	**0.0017**
16	10	30	0	0	**0.0022**
17	10	0	15	15	**0.0012**
18	25	0	15	0	**0.0021**
19	25	0	0	15	**0.0021**
20	10	30	0	0	**0.0017**

The final OD_{600} and the amount of XZX–7 protein produced by an OD_{600} = 0.1 obtained for each condition was analyzed by the DesignExpert® software and three ideal mixtures and their desirability were selected (S1, S2 and S3), in order to maximize the final biomass and the protein yield (**Table 11**). The desirability is the objective function based on the parameters and DesignExpert® calculates their significance. This function can have values between 0, when the objective is out of the limits of the experiment, and 1, when the objective is achieved.

Table 11 – Composition of the three best culture media mixtures determined by the DesignExpert® software to maximize the expression of XZX–7 protein

	[TB-AIM]	[Wey]	[CSL]	[Molasses]	Desirability
S1	10	0	30	0	0.636
S2	10	15	15	0	0.450
S3	20	10	10	0	0.497

All the three solution offered by DesignExpert® contain CSL in the composition of the culture media and do not contain molasses. Molasses, as previously described is composed primarily of glucose and sucrose, the presence of glucose retard the intake and metabolization of lactose, the induction factor on

the TB-AIM medium (Fox and Blommel, 2009; Zhang *et al.*, 2017). CSL is an inexpensive nitrogen source, being able to have 5 g/L of protein (Gudiña *et al.*, 2015b), supplying the building blocks to protein production.

The three solutions of culture media mixtures determined by the DesignExpert® software were evaluated by growing the *E. coli* expressing the XZX–7 protein in the media for 24 h at 37 °C. The samples were analyzed by SDS-PAGE (12%) (**Figure 17 A**) and the amount of protein expressed was determined using the ImageJ software (**Figure 17 B**). After analysis of the results, the culture media composition was 10 g/L TB-AIM with 30 g/L CSL with a final yield of 0.0021 mg.

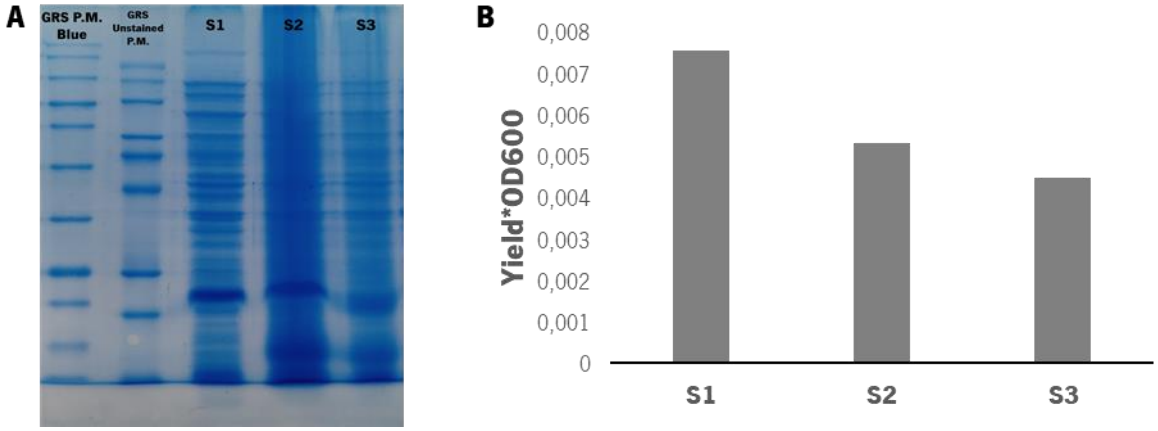

Figure 17 – Analysis of XZX–7 protein expression in the culture media mixtures determined by the DesignExpert® software: **A – SDS-PAGE electrophoresis (15%)** of the confirmation screening of XZX–7 protein; **B –** Expression yield of XZX-7 protein. The Yield*OD600 is obtain by multiplying the final OD_{600} with the yield (mg for a OD600=0.1)

Furthermore, the DesignExpert® software reports the influence of each compound in the responses evaluated the protein yield for an OD_{600} = 0.1 and the final OD_{600} (**Equation 3 and 4, Figure 18 and Figure 19**). For protein yield, the response fits a linear model, and all factors (TB-AIM (A), whey (B), CSL (C) and molasses (D)) have a positive impact, being the CSL the main contributing factor (**Equation 6**).

$$\textbf{Yield}=0.002A+0.002B+0.002C+0.001D \qquad \textbf{Equation 6}$$

Where A stands for TB-AIM, B for whey, C for CSL and D for molasses

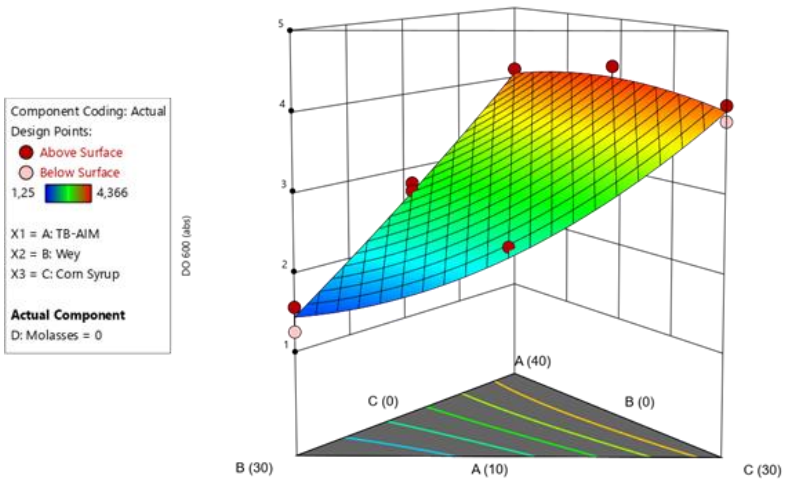

Figure 18 – 3D Surface graph of the response XZX–7 Yield and the impact of each factor (TB-AIM (A), whey (B) and CSL (C)), produced by DesignExpert® software.

The final OD_{600} response follows a quadratic fit, with the contribution of the isolated factors and with the combination of these. Besides the positive effect of TB-AIM (A), whey (B), CSL (C) and molasses (D), there is a negative impact of the combinations: TB-AIM with whey (AB), TB-AIM with molasses (AD), whey with CSL (BC) and whey with molasses (BD) (**Equation 7**).

$$\mathbf{OD_{600}} = 4.0\ 6A + 1.44\ B + 4.01\ C + 3.44\ D - 0.28\ AB - 0.73\ AC - 3.11\ AD - 2.13\ BC - 1.49\ BD + 0.\ 46CD \qquad \textbf{Equation 7},$$

Where A stands for TB-AIM, B for whey, C for CSL and D for molasses

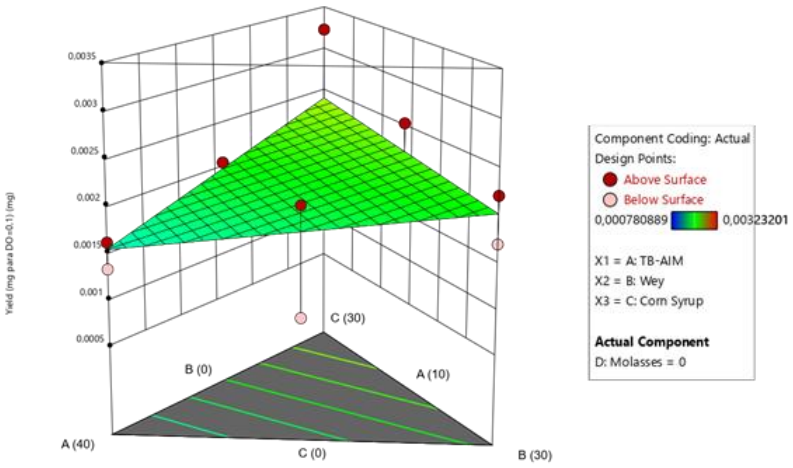

Figure 19 – 3D Surface graph of the response OD_{600} and the impact of each pf the factors study (TB-AIM (A), whey (B) and CSL (C)) produced by DesignExpert® software.

4.1.2.2. Best temperature and pH

To further optimize the production of XZX–7 protein, the effect of several growth temperatures (20, 25, 30, 37 and 40 °C) and pH (6, 7 and 8) were tested. To analyze the impact of this factors the DesignExpert® software was used. The final OD600 was measured and the amount of XZX–7 produced were evaluated by SDS-PAGE (**Figure 20**) using the reference band of the GRS unstained protein marker for ImageJ analysis (**Table 12**).

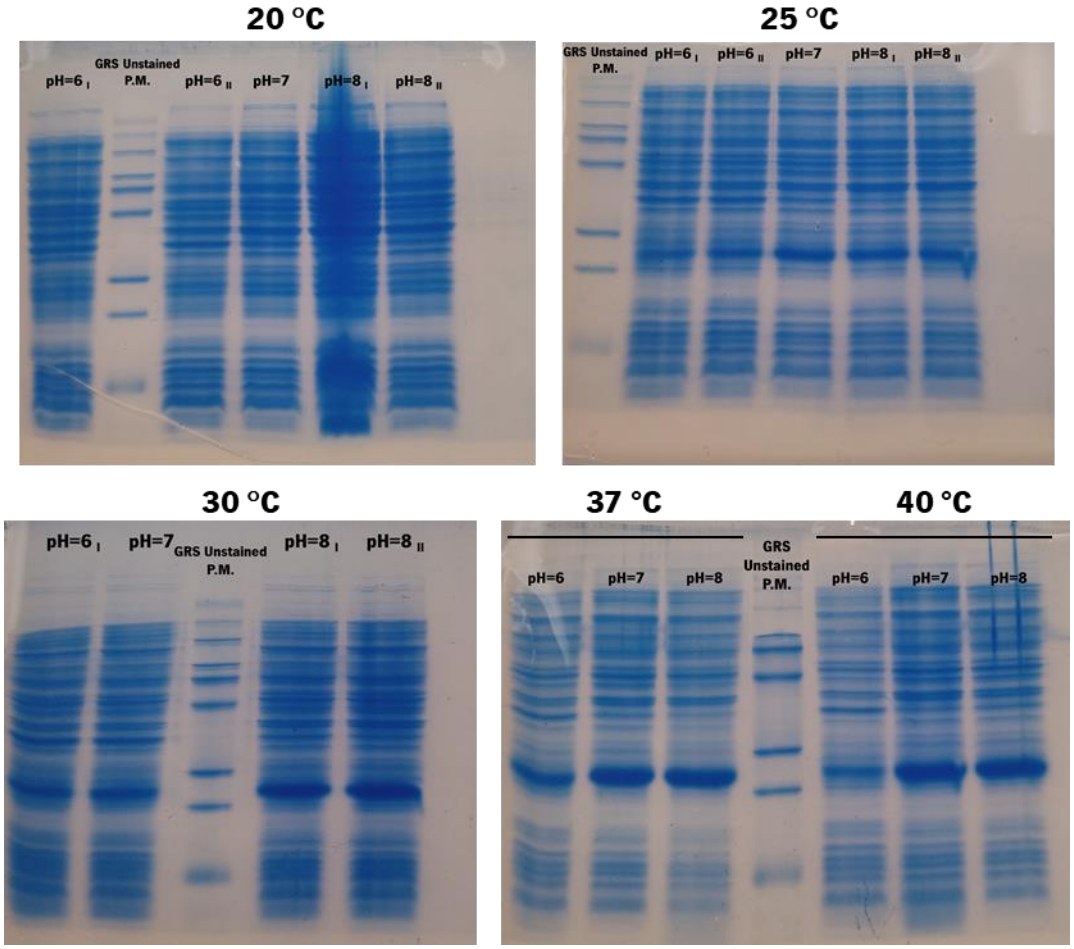

Figure 20 – SDS-PAGE electrophoresis (15%) of XZX–7 expression screenings to evaluate the best expression temperature (20, 25, 30 37 and 40°C) and pH (6, 7 and 8).

Table 12 – Different temperature (°C) and pH tested and their XZX–7 expression yields (mg for OD600=0.1)

Temperature (°C)	**pH**	**XZX–7 Yield** (mg for OD_{600}=0.1)
20	6	**0.0001**
20	6	**0.0002**
20	7	**0.0002**
20	8	**0.0001**
20	8	**0.0002**
25	6	**0.0005**
25	6	**0.0006**
25	7	**0.0011**
25	8	**0.0009**
25	8	**0.0006**
30	6	**0.0004**
30	6	**0.0006**
30	7	**0.0010**
30	8	**0.0008**
37	6	**0.0009**
37	7	**0.0034**
37	8	**0.0032**
40	6	**0.0012**
40	7	**0.0021**
40	8	**0.0025**

The results (yield and final OD_{600}) obtained for each condition were analyzed by the DesignExpert® software and the best combination of temperatures and culture medium pH, in order to maximize the final OD_{600} and the amount of protein, where found (**Table 13**).

Table 13 – Combinations of temperature and pH of culture medium, and their desirability, calculated by DesignExpert® in order to maximize the yield of XZX–7 protein

	Temperature	pH	Desirability
S1	37	7	0.809
S2	37	8	0.506
S3	40	7	0.276

The three solutions (combination of temperature and pH) were evaluated by a new expression screening and the results were analyzed by SDS-PAGE (**Figure 21 A**) and the amount of protein (Yield*$DO_{600\ final}$) for each condition was calculated (**Figure 21 B**).

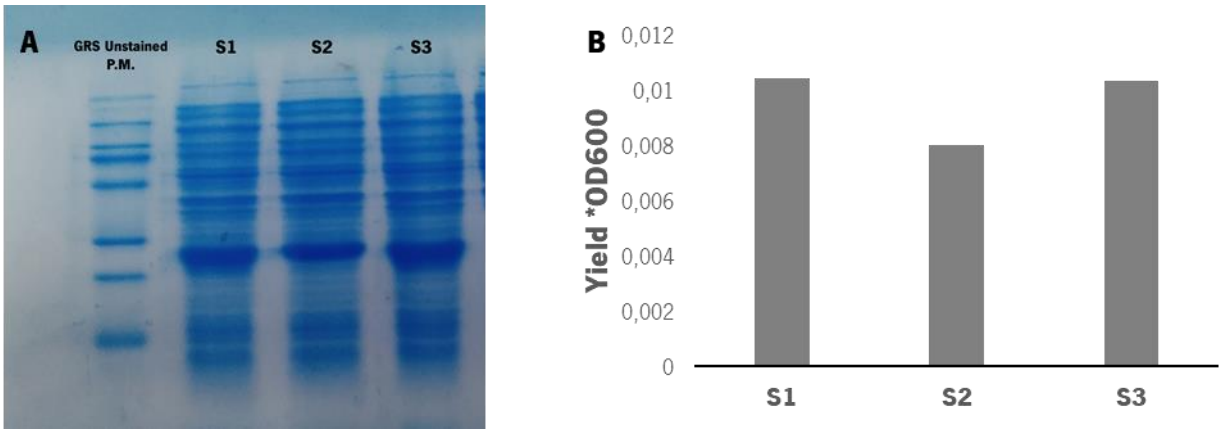

Figure 21 – Analysis of XZX–7 protein expression in the culture medium pH and temperature combinations determined by the DesignExpert® software: **A – SDS-PAGE electrophoresis (15%)** of the confirmation screening of XZX–7 protein; **B –** Expression yield of XZX–7 protein. The Yield*OD600 is obtain by multiplying the final OD600 with the yield.

The best combination of temperature and culture medium pH was 37°C and a pH 7 with a yield of 0.0022 mg. As temperature and pH are categorical factors the DesignExpert®, software was not able to produce the equations that correlate the different factors with the responses.

4.1.2.3. Best additives

After determining the best mixture of culture medium and industrial by-products, and the best combination of growth temperature and culture medium pH, the effect of different additives (glucose, glycerol, sorbitol and NaCl) on the expression of XZX–7 protein was analyzed. Several expression screenings were performed according to the matrix determined by the DesignExpert® software. The final OD_{600} was measured and the amount of XZX–7 protein was evaluated (**Figure 22 and Table 14)**

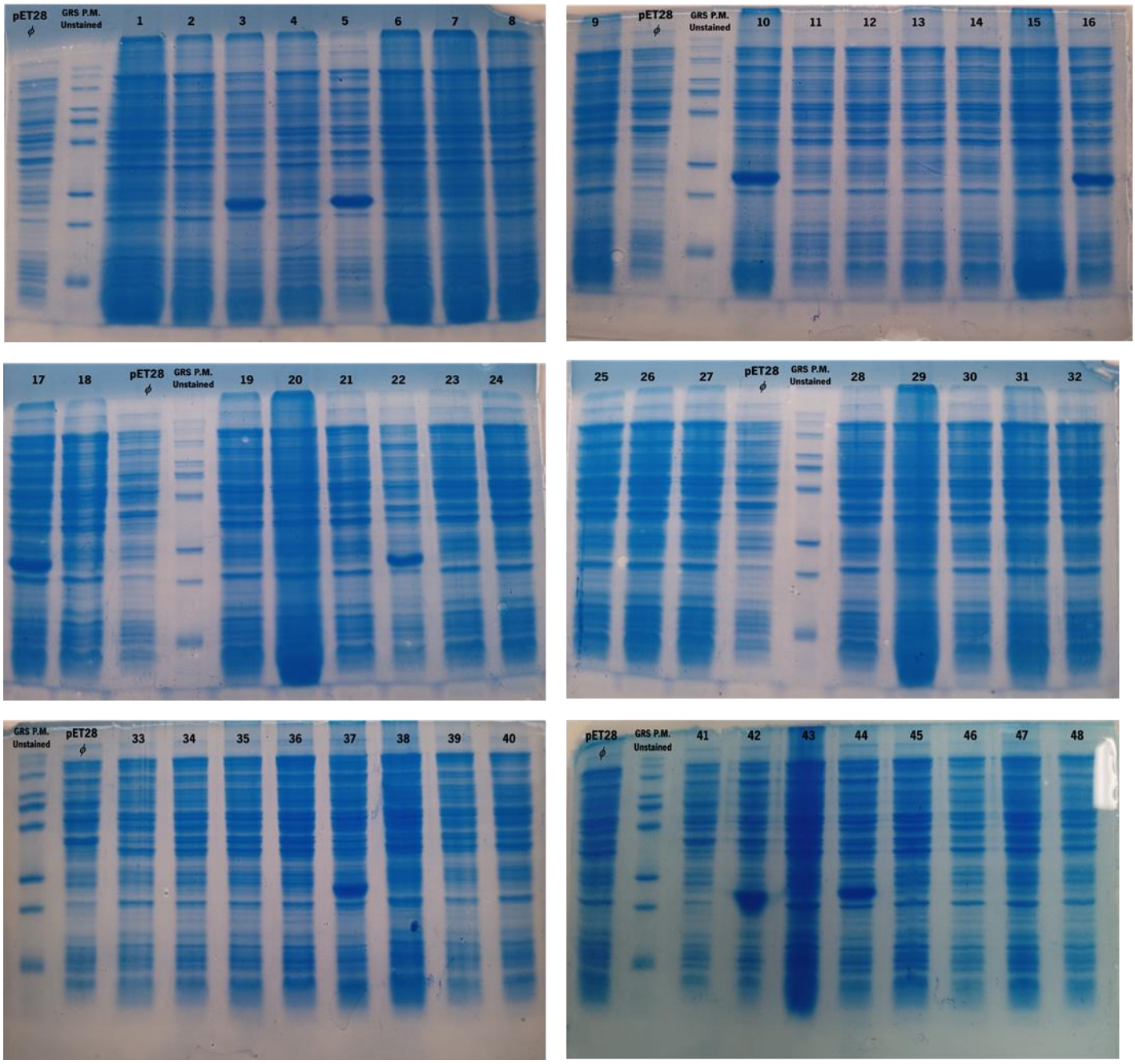

Figure 22 – SDS-PAGE electrophoresis (15%) of XZX–7 expression screenings to evaluate the best additives (NaCl, Glycerol, Glucose and Sorbitol).

Table 14 – XZX–7 Yields (mg for a OD_{600} equal to 0.1) of the different mixtures of additives (showing in detail in the **section 3.1.5.3.**)

Run	XZX–7 Yield (mg for OD_{600}=0.1)	Run	XZX–7 Yield (mg for OD_{600}=0.1)
1	0.0005	**25**	0.0004
2	0..0002	**26**	0.0006
3	0.0018	**27**	0.0005
4	0.0003	**28**	0.0014
5	0.0019	**29**	0.0010
6	0.0005	**30**	0.0004
7	0.0006	**31**	0.0008
8	0.0005	**32**	0.0006
9	0.0013	**33**	0.0000
10	0.0023	**34**	0.0000
11	0.0002	**35**	0.0001
12	0.0004	**36**	0.0002
13	0.0002	**37**	0.0019
14	0.0004	**38**	0.0007
15	0.0015	**39**	0.0000
16	0.0021	**40**	0.0003
17	0.0026	**41**	0.0000
18	0.0005	**42**	0.0026
19	0.0004	**43**	0.0001
20	0.0011	**44**	0.0019
21	0.0004	**45**	0.0000
22	0.0018	**46**	0.0000
23	0.0005	**47**	0.0000
24	0.0005	**48**	0.0000

In order to maximize the final OD_{600} and the protein yield the DesignExpert® software determined five solutions and their desirability (**Table 15**).

Table 15 – The best additives and their desirability, calculated by DesignExpert® in order to maximized XZX–7 production.

	[NaCl]	[Glycerol]	[Glucose]	[Sorbitol]	Desirability
S1	0	0	0	0	0.61
S2	0.365	0	0	0	0.606
S3	0.934	0	0	0	0.598
S4	1.889	0	0	0	0.583
S5	0	11.5	0	0	0.273

A new screening was performed to determine which of the five solutions rendered the highest protein yield **Figure 23**). Analyzing the results the additives that rendered the highest protein yield (0.0024 mg) was the NaCl with a concentration of 0.934 g/L.

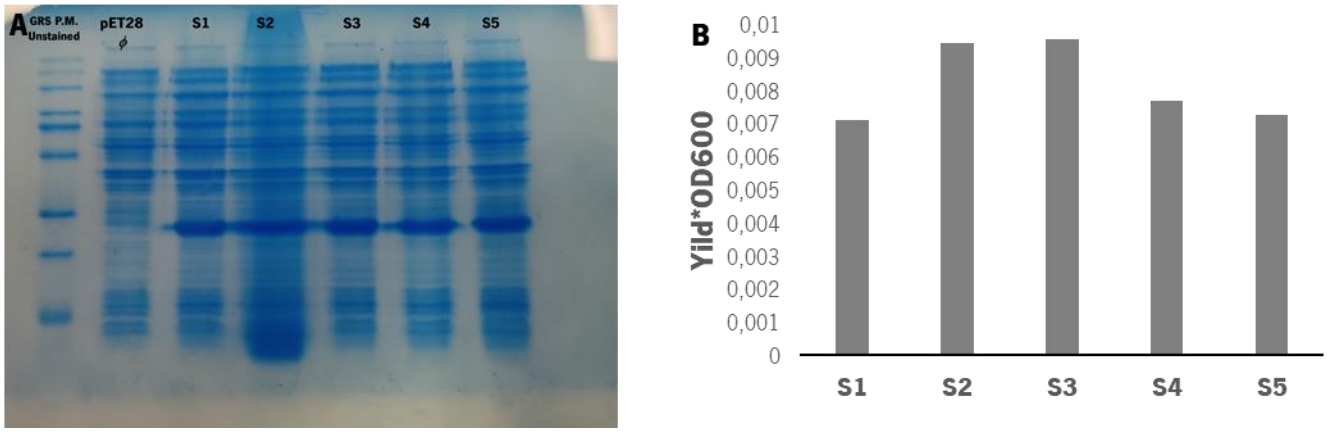

Figure 23 – Analysis of XZX–7 protein expression in the culture medium and additives combinations determined by the DesignExpert® software: **A – SDS-PAGE electrophoresis (15%)** of the confirmation screening of XZX–7 protein; **B –** Expression yield of XZX–7 protein. The Yield*OD600 is obtain by multiplying the final OD600 with the yield

Furthermore, the DesignExpert® reported the influence of each additive in the different responses (yield for a OD_{600} of 0.1 and the final OD_{600}) addressed (**Equation 8 and 9**).

$$\textbf{Yield} = 0.0005 + 0.0001\,A - 0.0009\,C - 0.0001\,D + 0.0001\,AB + 0.0001\,BD - 0.0001\,A^2 - 0.0001\,B^2 + 0.0005\,C^2 + 0.0004\,D^2 \quad \textbf{Equation 8},$$

$$\textbf{OD}_{600} = 2.2 - 0.36\,A - 0.2\,B - 0.37\,C - 0.06 + 0.02\,AB + 0.03\,AC + 0.06\,AD + 0.19\,BC\ 0.01\,BD + 0.21\,CD + 0.14\,A^2 + 0.01B^2 + 0.28\,C^2 + 0.01\,D^2 \quad \textbf{Equation 9}$$

Where A stand for NaCl, B for glycerol, C for glucose and D for sorbitol

For the response final OD_{600}, that fit a quadratic model, all the combined factors and the quadratic ones have a positive impact and all the single (NaCl (A), Glycerol (B), Glucose (C) and Sorbitol (D)) factors have a negative impact. The response yield is more complex, not having a clear distinction on the factors that have a positive and negative impact.

The addiction of ethanol can increase the expression of recombinant proteins in *E. coli* strains (Chhetri *et al.*, 2015). After optimizing the culture medium, the growth temperature, the culture medium pH and the additives, the effect of ethanol concentration (0, 1, 2 and 3%) on XZX–7 protein expression was evaluated (**Figure 24**).

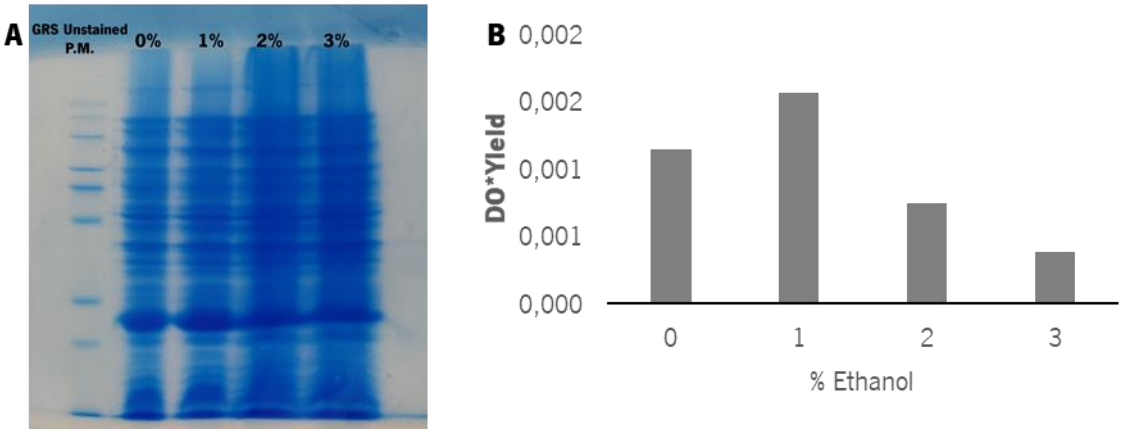

Figure 24 – Analysis of XZX–7 protein expression with different ethanol percentages (0, 1, 2, 3): **A – SDS-PAGE electrophoresis (15%)** of the screening of XZX–7 protein; **B –** Expression yield of XZX–7 protein. The Yield*OD600 is obtain by multiplying the final OD600 with the yield (mg for a OD600=0.1) and 10 (in order to obtain the total amount of recombinant protein (mg) expressed

Generally, the tested ethanol concentrations did not affected the cellular growth and 1% ethanol had a positive effect on the expression of the XZX–7 protein, with a final yield of 0.0025 mg. Ethanol is an amphipathic substance that may significantly alter the cellular environment by altering membrane fluidity, membrane transport, membrane lipid composition, and membrane protein assembly (Dombek and Ingram, 1984). These modifications may have an effect on membrane-associated phenomena such as DNA replication, resulting to an increase in DNA synthesis (Basu and Poddar, 1994). It is proposed that increasing DNA synthesis leads to gene amplification, which may increase the production of inducible proteins in ethanol-treated cells (Chhetri *et al.*, 2015).

4.1.2.4. Growth kinetics of XZX–7 producing E. coli in the optimized media conditions

After transforming all the data in dry weight (calibration curves in **supplementary information 1**), a growth curve for each medium was drawn (**Figure 25**). At the end of 8 hours, it is possible to observe that the dry weight in TB-AIM medium is the lowest and that the medium composed by 10 g/L TB-AIM, 30 g/L CSL and 0.934 g/L NaCl is the highest (**Figure 25**).

The specific growth rate (μ) and the doubling time were calculated (**Table 16**) for each medium. Comparing the specific growth rate (μ) of the TB-AIM (1.4 ± 0.2 h^{-1})medium with the described in the literature for TB medium (μ=0.97 (Romano *et al.*, 2009)) the higher value of TB-AIM can be explained by the higher amount of carbon sources in TB-AIM.

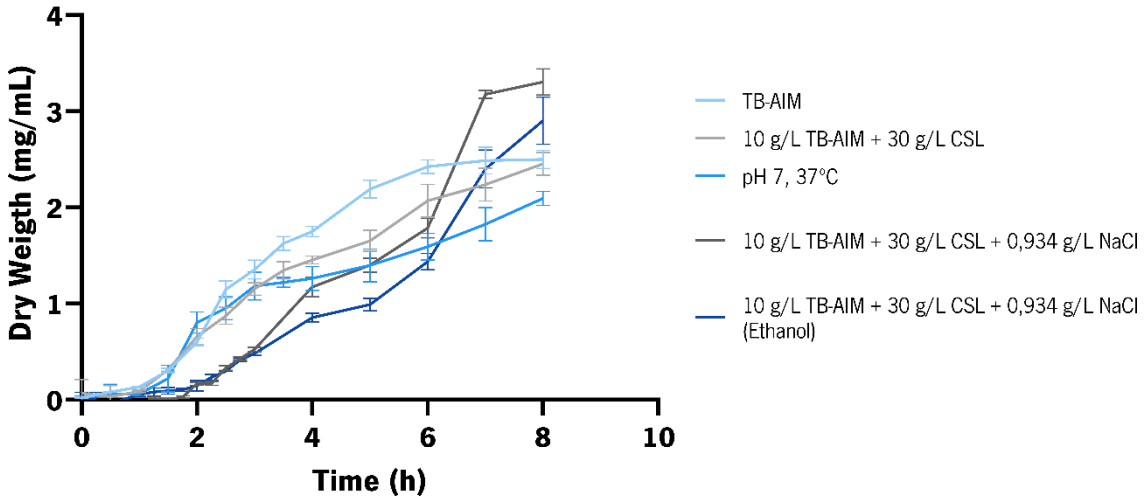

Figure 25 – Growth curve of XZX–7 producing *E. coli* cells in the commercial medium (TB-AIM) and in each of the optimized media mixtures and conditions: 1 - 10 g/L TB-AIM + 30 g/L CSL; 2 - 10 g/L TB-AIM + 30 g/L CSL at pH7 and 37°C; 3 - 10 g/L TB-AIM + 30 g/L CSL+0.934 g/L NaCl; 4 - 10 g/L TB-AIM + 30 g/L CSL+0.934 g/L NaCl+ 1% ethanol).

Table 16 – Specific growth rates (μ) and doubling times of XZX–7 producing *E. coli* cells in the optimized culture media conditions

	Specific Growth Rate (μ) (h^{-1})	Doubling Time (h)
TB-AIM	1.4 ± 0.2	0.5 ± 0.1
10 g/L TB-AIM+30g/L CSL	2.00 ± 0.6	0.3 ± 0.2
10 g/L TB-AIM+30g/L CSL, pH 7	2.57 ± 0.02	0.270 ± 0.006
10 g/L TB-AIM+30g/L CSL+0.931g/L NaCl, pH 7	0.8 ± 0.1	0.8 ± 0.1
10 g/L TB-AIM+30g/L CSL+0.931g/L NaCl + 1% ethanol, pH 7	0.7 ± 0.4	1.1 ± 0.4

4.1.3. Production and purification of XZX-based proteins

The XZX-based proteins (XZX–2, XZX–3 and XZX–5) which presented promising expression patterns, were produced in 0.5L of LB minimum with a IPTG induction period of 24 hours. After sonication, the samples were ran in a SDS-PAGE to evaluate if the proteins were in the soluble or insoluble fraction (**Figure 26**).

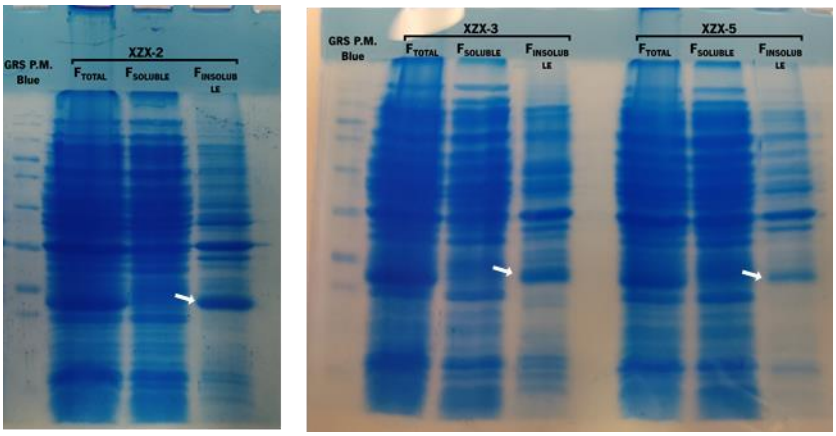

Figure 26 – SDS-PAGE electrophoresis (12%) of total, soluble and insoluble fraction of XZX–2, XZX–3 and XZX–5producing *E. coli* cells. The XZX–based proteins are highlighted with an arrow.

The three XZX–based proteins were present on the insoluble fraction. These proteins are highly repetitive and have a high number of cysteine residues on their amino acidic sequence. These factors might promote the formation of inclusion bodes during protein expression. In order to purify the proteins, the insoluble fraction was incubated with nickel magnetic beads and the proteins of interest were eluted with increasing concentrations of imidazole. The ability of this approach to purify proteins present on the insoluble fraction was evaluated by SDS-PAGE (**Figure 27**).

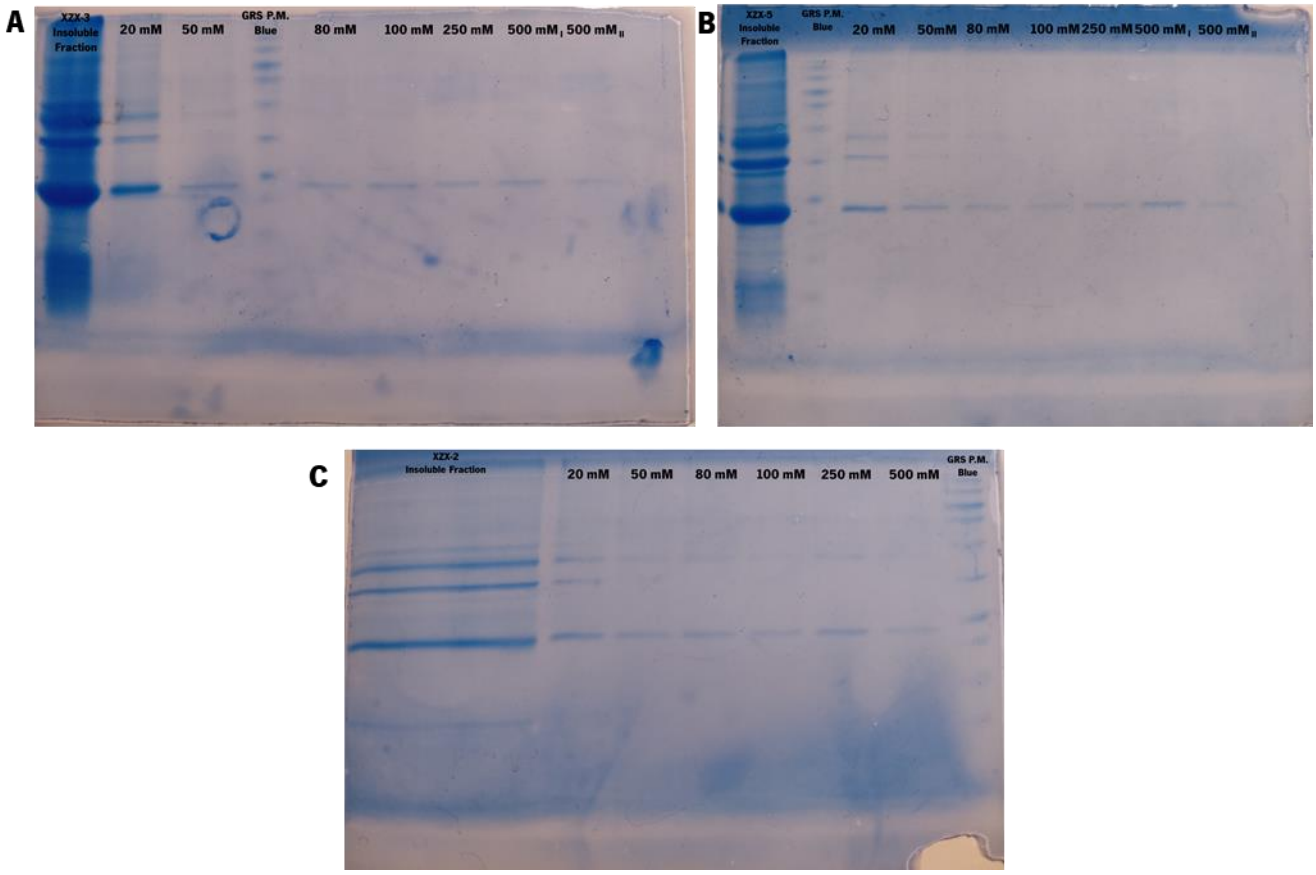

Figure 27 – SDS-PAGE electrophoresis (12%) of the insoluble fractions of XZX–3 (**A**), XZX–5 (**B**), and XZX–2 (**C**) purified with nickel magnetic beads and eluted with increasing imidazole concentrations.

The incubation of the insoluble fractions with the nickel magnetic beads allowed the purification of the XZX–2, XZX–3 and XZX–5 proteins with a good level of protein purity in the first elution fraction (20 mM of imidazole). As imidazole concentration increased the protein purity also increased.

The XZX–7 protein was expressed in the conditions determined by the experimental design (10 g/L TB-AIM, 30 g/L CSL, 0.93 g/L NaCl, pH 7, at 37°C). To access if XZX–7 protein was in the soluble or in the insoluble fraction the samples were ran in an SDS-PAGE electrophoresis (Figure 23).

As for the XZX–2, XZX–3 and XZX–5 proteins, the XZX–7 protein was expressed in the insoluble fraction (**Figure 28 A**). The same approach used for XZX–2, XZX–3 and XZX–5 proteins was applied to purify XZX–7 protein (**Figure 28 B**). The SDS-PAGE electrophoresis revealed that the XZX–7 protein had a great level of purity in the first elution fraction and was almost pure in the 50 mM imidazole fraction.

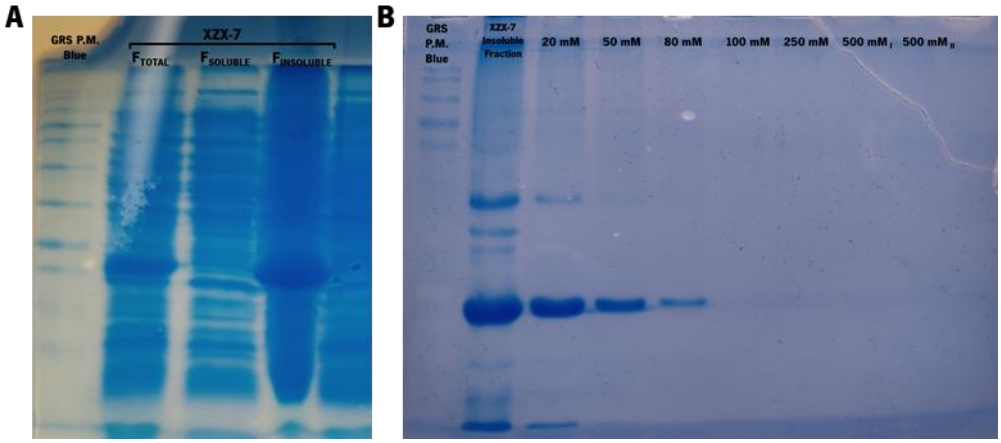

Figure 28 – SDS-PAGE electrophoresis (15%) of: **A –** total, soluble and insoluble fraction of XZX–7 producing *E. coli* cells; **B –** XZX–7 purified with nickel magnetic beads and eluted with increasing imidazole concentrations.

4.2. XZX–7 Protein Characterization

4.2.1. Fourier-transform Infrared Spectroscopy (FTIR)

In the FTIR spectrum of XZX–7 (**Figure 29 A**) is possible to note the characteristics peaks of proteins, corresponding to the amide I and amide II (C= O bonds) at 1661 and 1476 cm^{-1} respectively. Some other characteristic peaks were observed for these proteins (Tinoco *et al.*, 2021b). The peaks at 670, 1055 and 3140 cm^{-1}, which correspond to C–S bonds present in cystine, C–O stretching vibration and to N–H bonds, respectively, all these peaks were expected considering the amino acid sequence of XZX–7 (not shown) (Tinoco *et al.*, 2021b).

After deconvolution of the amide I peak is possible to infer the secondary structure of XZX–7 (KONG and YU, 2007). The deconvolution of the amide I peak of XZX–7 (**Table 17**) show 3 peaks corresponding to 3 different structures, a random one a β-turn and a β-sheet (KONG and YU, 2007).

Table 17 – Resulting discrete peaks, their respective contribution to the overall FTIR-derived curves and the corresponding structural assignments of XZX–7 protein. Structural assignment was performed according to KONG and YU, 2007

Deconvoluted data		Deconvoluted data assignment			
Peak center (cm^{-1})	**Area occupied (%)**	**Helical**	**β-sheet**	**β-turn**	**Random**
1646	24.3				✓
1668	21.7			✓	
1695	0.6		✓		

4.2.2. Differential Scanning Calorimetry (DSC)

The DSC analyze (**Figure *29* B**) revealed two endothermic transitions at 137.7 and 150.3°C with a ΔH of 30.73 and 146.13 J/g. The first transition is most likely associated with the glass transition while the second transition its most likely to be associated the solid-liquid transition (melting).

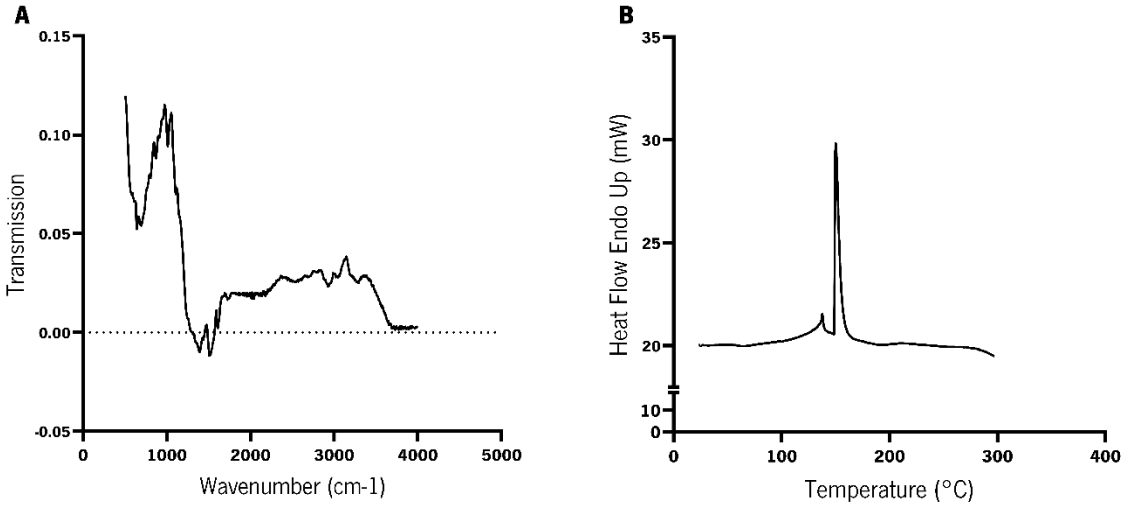

Figure 29 – A: FTIR spectrum of XZX–7 protein; **B:** DCS curve of XZX–7 with a T_g= 137.7 °C and T_m =150.3 °C.

4.3. Hair Treatment with XZX–7 Protein

The development of protein-based hair cosmetic formulations relays on the proteins' ability to bind to the hair fibers. To study the binding potential of XZX–7 protein towards hair, the XZX–7 inclusions bodies were first resuspended in deionized water or in an ethanolic formulation (phosphate buffer 0.1 M, 1.5 % (V/V) propylene glycol, 0.5 % Benzyl alcohol, 10% ethanol). Afterwards, the protein solution was applied to virgin Asian hair and incubated for 1 hour at room temperature. The amount of XZX–7 protein bound to the hair, measured in percentage, was analyzed in an SDS-PAGE electrophoresis by comparing the band intensity before and after application (**Figure 30**).

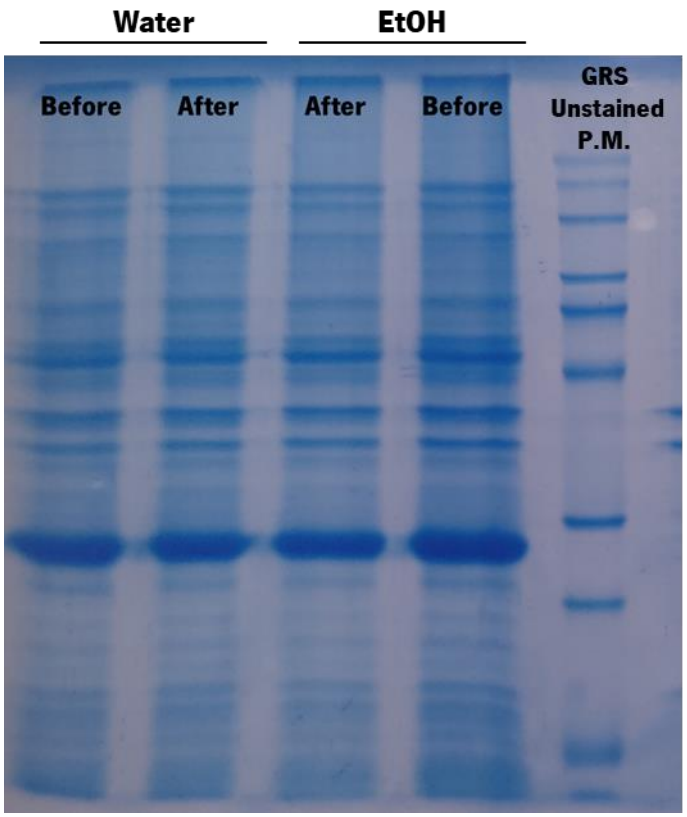

Figure 30 – SDS-PAGE electrophoresis (12%) analysis of XZX–7 protein binding assay: **Water** $_{Before}$ is the sample of inclusion bodies resuspended in water prior to incubation with hair, **Water** $_{After}$ is after incubation with hair; **EtOH** $_{Before}$ is the sample of inclusion bodies resuspended in ethanolic formulation prior to incubation with hair, **EtOH** $_{After}$ is after incubation with hair.

Regardless of the formulation used for the inclusion bodies resuspension, the XZX–7 protein was able to bind to the virgin Asian hair, with a binding efficiency of 17.15% and 33.03% in water and ethanolic formulation, respectively. Since the global charge of the XZX–7 protein is positive in both formulations (pH water = 7 and pH ethanolic formulation = 5), the differences observed in the binding of XZX–7 protein to hair are related with the effect of the ethanolic formulation on hair. This ethanolic formulation promoted hair swelling and exposed the hair shaft to the protein, leading to a higher binding efficiency (Tinoco *et al.*, 2019b).

After confirming the binding ability of XZX–7 protein towards virgin Asian hair, the perming potential of the protein was assessed. The virgin hair meshes were treated with 2 M of urea and/or ethanolic formulation prior to incubation with two concentrations of XZX–7 protein (1 and 20 mg/mL) applied in deionized water or in ethanolic formulation. The perming efficiency was evaluated after six washing cycles with a commercial shampoo (**Figure 31**).

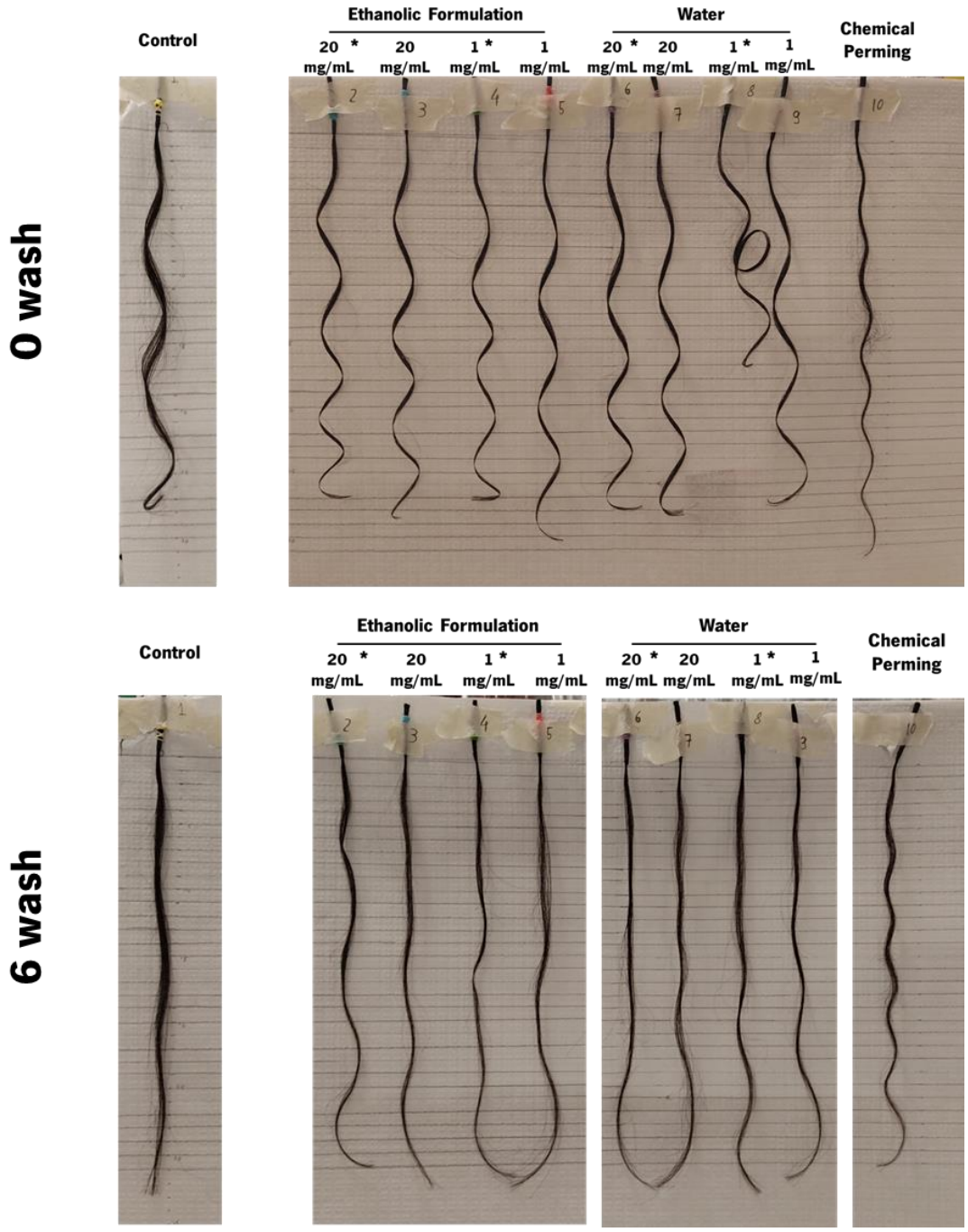

Figure 31 – Perming potential of XZX–7 protein before and after six washing **cycles** with a commercial shampoo. The hair tresses were treated with 20 mg/mL and 1 mg/mL of XZX–7 inclusion bodies resuspended in water or in ethanolic formulation and with (*) and without pre-treatment with 2 M urea. Hair treated with a commercial formulation for perming ("Chemical"), and without any treatment ("Control") were used as controls of proteins' capacity to curl virgin Asian hair.

Using **Equation 5** the perming efficiency was calculated, considering the number of loops and the length of the control without treatment (**Figure 32**).

The condition of 20 mg/mL in water without urea pre-treatment showed a perming efficiency similar to the control before washing with the commercial shampoo. Despite this result, this condition showed a good performance throughout washing, being able to maintain the shape of the hair along the

washing cycles, with a final perming efficiency of 84%. The condition of 1 mg/mL in ethanolic formulation without pre-treatment presented the lowest perming efficiency at 0 washes (90%), and the lowest perming efficiency after 6 washes (55%). The condition that showed the best perming efficiency (159%) before washing was the 1 mg/mL of XZX–7 protein in water with urea pre-treatment. This condition showed better results than the chemical control before washing (150%) and displayed the second best performance after six washing cycles (83%).

When the hair was pre-treated with urea and 20 mg/mL of XZX–7 inclusion bodies resuspended in ethanolic formulation, a perming efficiency of 121% and 87% was measured before washing and after the washing cycles, respectively. Besides the chemical treatment, this was the best condition in maintaining the shape of the hair along the washing cycles.

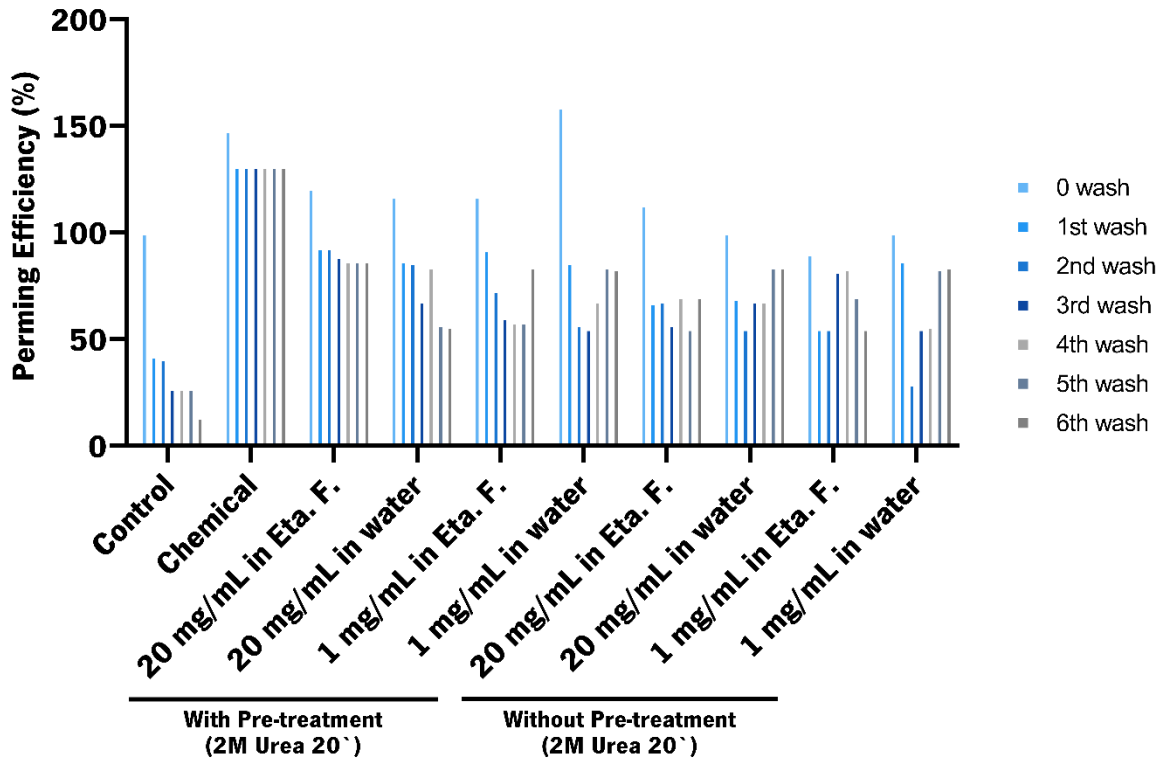

Figure 32 – Perming efficiency of virgin Asian hair treated with XZX–7 protein throughout six washing cycles with a commercial shampoo. The hair tresses were treated with 20 mg/mL and 1 mg/mL of inclusion bodies resuspended in water or in ethanolic formulation, and with and without pre-treatment of 2 M urea. Hair treated with a commercial formulation for perming ("Chemical"), and without any treatment ("Control") were used as controls for the proteins' capacity to curl the Asian hair.

5. CONCLUSION AND FUTURE PERSPECTIVES

This work aimed the development of hair cosmetic products based in fusion proteins. Several proteins were design based in the XZX protein sequence, and their expression conditions were optimized.

With exception of XZX–1, the other XZX–proteins (2-6) were expressed in LB medium with a 24 hour period of induction with IPTG at 37 °C and 170 rpm. Other conditions also resulted in the expression of some these proteins. XZX–2, XZX–3 and XZX–5 proteins were expressed in a in a larger scale in LB medium with a 24 hour period of induction with IPTG at 37 °C and 170 rpm. After fraction analyzes, it was observed that the three proteins were expressed in the insoluble fraction. The XZX proteins present in the insoluble fractions were purified using magnetic nickel beads with satisfactory levels of purity.

XZX–7 protein was expressed in all the commercial media tested, however the best yield was obtained for TB-AIM medium. To further enhance and reduce the cost of XZX–7 expression, a series of optimizations using an experimental design approach were tested. The first optimization aimed the replacing of commercial media by industrial by products such as corn syrup, molasses and whey. After analysis using the DesignExpert® software the best culture mixture of commercial media and industrial by-products was 10 g/L TB-AIM + 30 g/L CSL. The second optimization focused on the culture media pH and growth temperature. The results obtained showed that the optimal conditions were pH 7 and 37 °C. The third optimization evaluated the effect of additives on XZX–7 expression. Several additives were tested; however, only NaCl showed a positive effect on protein expression in a final concentration of 0.934 g/L (NaCl). The XZX–7 expression was further optimized by adding 1% ethanol to the culture medium. The final yield of all the optimizations was 0.0025 mg of XZX–7 per 0.1 unity of OD_{600}. Like for the other XZX-based proteins, the XZX–7 was also expressed on the insoluble fraction.

After purification the XZX–7 protein was physico-chemically characterized. The deconvolution of amide I band from the FTIR spectrum revealed that XZX–7 have three different arrangements in terms of secondary structure: random, β-turn and β-sheet. The DSC study revealed two endothermic transitions at 137.7 and 150.3°C, corresponding to the glass transition and the melting transition, respectively.

The ability of XZX–7 protein to bind to hair was evaluated by incubation of XZX–7 inclusion bodies in water and in an ethanolic formulation, with Asian virgin hair for one hour at room temperature. The percentage of XZX–7 bound to hair was dependent on the solution where the inclusion bodies were resuspended. When the XZX–7 protein was resuspended in water the binding ability of the protein was 17.15%, however when the protein was resuspended in the ethanolic formulation a higher binding was

observed (33.03%). This difference was related on the effect of ethanolic formulation on hair structure and properties. This formulation swells the hair and exposes the hair cuticles and cortex to the protein, increasing the binding of XZX–7 to hair. Regarding to the perming potential of XZX–7 the condition 20 mg/mL in ethanolic formulation with 2M urea treatment was the best in maintaining the hair shape throughout six washing cycles. .

In summary, from the seven XZX-based proteins tested during this project, it was found at least one culture condition which allowed the expression of six of them. However, the expression patterns were differentiated according to the protein tested. For XZX–2, XZX–3, XZX–4, XZX–5 and XZX–6 it is necessary to optimize the culture conditions in the future. Considering that, these proteins have a repetitive sequence and a high percentage of cysteine residues in their sequence, an eukaryotic host could be tested.

Regarding the XZX–7 protein, despite its expression was optimized during this work; other conditions should be tested to increase the expression levels. A scale-up to reactors could be tested in future. It is also necessary to explore the solubilization of XZX–7 protein in the future to exploit its full potential as a new eco-friendly hair cosmetic agent.

BIBLIOGRAPHY

Aarti V, D. *et al.* (2005) 'Moisturizing efficiency of silk protein hydrolysate: Silk fibroin', *Indian Journal of Biotechnology*, 4, pp. 115–121.

Akkermans, R. L. C. and Warren, P. B. (2004) 'Multiscale modelling of human hair', *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*. Edited by M. Al–Ghoul, J. P. Boon, and P. V. Coveney, 362(1821), pp. 1783–1793. doi: 10.1098/rsta.2004.1395.

Amartey, S. and Jeffries, T. W. (1994) 'Comparison of corn steep liquor with other nutrients in the fermentation of D-Xylose by Pichia stipitis CBS 6054', *Biotechnology Letters*, 16(2), pp. 211–214. doi: 10.1007/BF01021673.

Araújo, R. *et al.* (2010) 'Biology of Human Hair: Know Your Hair to Control It', in, pp. 121–143. doi: 10.1007/10_2010_88.

Azizova Marina (New Canaan, C. T. U. A. E. A. (Putnam V. N. Y. U. T. R. (Trumbull C. T. U. C. A. (Littleton M. A. U. (2016) 'Hair treatment composition with naturally - derived peptide identical to human hair'.

Baeshen, M. N. *et al.* (2015) 'Production of Biopharmaceuticals in E. coli: Current Scenario and Future Perspectives', *Journal of Microbiology and Biotechnology*, 25(7), pp. 953–962. doi: 10.4014/jmb.1412.12079.

Basit, A. *et al.* (2018) 'Health improvement of human hair and their reshaping using recombinant keratin K31', *Biotechnology Reports*, 20, p. e00288. doi: 10.1016/j.btre.2018.e00288.

Basu, T. and Poddar, R. K. (1994) 'Effect of ethanol onEscherichia coli cells. Enhancement of DNA synthesis due to ethanol treatment', *Folia Microbiologica*, 39(1), pp. 3–6. doi: 10.1007/BF02814520.

Benson Edward R. (Durham, N. C. U. F. S. R. (Wilmington D. E. U. H. P. (Cary N. C. U. O. J. P. (Oxford P. A. U. W. H. (Kennett S. P. A. U. (2012) 'Hair binding peptides and peptide-based hair reagents for personal care'. Available at: ,8273337.

Berkmen, M. (2012) 'Production of disulfide-bonded proteins in Escherichia coli', *Protein Expression and Purification*, 82(1), pp. 240–251. doi: 10.1016/j.pep.2011.10.009.

Bringans, S. D. *et al.* (2007) 'Characterization of the exocuticle a-layer proteins of wool', *Experimental Dermatology*, 16(11), pp. 951–960. doi: 10.1111/j.1600-0625.2007.00610.x.

Bryson, W. G. *et al.* (2009) 'Cortical cell types and intermediate filament arrangements correlate with fiber curvature in Japanese human hair', *Journal of Structural Biology*, 166(1), pp. 46–58. doi: 10.1016/j.jsb.2008.12.006.

Buffoli, B. *et al.* (2014) 'The human hair: from anatomy to physiology', *International Journal of Dermatology*, 53(3), pp. 331–341. doi: 10.1111/ijd.12362.

Cai, J. *et al.* (2018) 'Oil body bound oleosin-rhFGF9 fusion protein expressed in safflower (Carthamus tinctorius L.) stimulates hair growth and wound healing in mice', *BMC Biotechnology*, 18(1), p. 51. doi: 10.1186/s12896-018-0433-2.

Cao, W. *et al.* (2021) 'Unraveling the Structure and Function of Melanin through Synthesis', *Journal of the American Chemical Society*, 143(7), pp. 2622–2637. doi: 10.1021/jacs.0c12322.

Cauwet-martin Daniele (Paris, F. D. C. (Le C. F. (2000) 'Cosmetic compositions containing a lipid ceramide compound and a peptide having a fatty chain, and their uses'. Available at: ,6039962.

Chen, J. *et al.* (2018) 'Metabolic engineering of Escherichia coli for the synthesis of polyhydroxyalkanoates using acetate as a main carbon source', *Microbial Cell Factories*, 17(1), p. 102. doi: 10.1186/s12934-018-0949-0.

Chhetri, G., Kalita, P. and Tripathi, T. (2015) 'An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli', *MethodsX*, 2, pp. 385–391. doi: 10.1016/j.mex.2015.09.005.

Coderch, L. *et al.* (2019) 'External lipid function in ethnic hairs', *Journal of Cosmetic Dermatology*, 18(6), pp. 1912–1920. doi: 10.1111/jocd.12899.

Cruz, C. *et al.* (2016) 'Human Hair and the Impact of Cosmetic Procedures: A Review on Cleansing and Shape-Modulating Cosmetics', *Cosmetics*, 3(3), p. 26. doi: 10.3390/cosmetics3030026.

Cruz, C. F. *et al.* (2013) 'Keratins and lipids in ethnic hair', *International Journal of Cosmetic Science*, 35(3), pp. 244–249. doi: 10.1111/ics.12035.

Cruz, C. F. *et al.* (2017) 'Changing the shape of hair with keratin peptides', *RSC Advances*, 7(81), pp. 51581–51592. doi: 10.1039/C7RA10461H.

Dawber, R. (1996) 'Hair: Its structure and response to cosmetic preparations', *Clinics in Dermatology*, 14(1), pp. 105–112. doi: 10.1016/0738-081X(95)00117-X.

Detert Marion (Pommernweg 22, 22455 Hamburg DE) Dingler Christian (Hagenbeckallee 2c 22527 Hamburg DE) Pilzner Anke (Steendammswish 50 22459 Hamburg DE) Sass Viola (Am Kamp 10 25488 Holm DE) (2008) 'Hair styling preparations with special protein hydrolysates'. Available at: ,EP1878423.

Dimotakis Emmanuel (Oradell, N. J. U. S. J. M. (South O. N. J. U. B. H. S. (Piscataway N. J. U. S. J. (Mamaroneck N. Y. U. (2013) 'HAIR COSMETIC AND STYLING COMPOSITIONS BASED ON MALEIC ACID COPOLYMERS AND POLYAMINES'.

Dombek, K. M. and Ingram, L. O. (1984) 'Effects of ethanol on the Escherichia coli plasma membrane', *Journal of Bacteriology*, 157(1), pp. 233–239. doi: 10.1128/jb.157.1.233-239.1984.

Fahnestock Stephen R. (Wilmington, D. E. U. S. T. M. (Randolph N. J. U. (2006) 'Water-soluble silk proteins in compositions for skin care, hair care or hair coloring'. Available at: ,7060260.

Fox, B. G. and Blommel, P. G. (2009) 'Autoinduction of Protein Expression', *Current Protocols in Protein Science*, 56(1). doi: 10.1002/0471140864.ps0523s56.

Franbourg, A. *et al.* (2003) 'Current research on ethnic hair', *Journal of the American Academy of Dermatology*. Mosby Inc., 48(6), pp. S115–S119. doi: 10.1067/mjd.2003.277.

Gavazzoni Dias, M. F. (2015) 'Hair cosmetics: An overview', *International Journal of Trichology*, 7(1), p. 2. doi: 10.4103/0974-7753.153450.

Gonçalves, F. *et al.* (2018) 'OBP fused with cell-penetrating peptides promotes liposomal transduction', *Colloids and Surfaces B: Biointerfaces*, 161, pp. 645–653. doi: 10.1016/j.colsurfb.2017.11.026.

Gong, H. *et al.* (2012) 'An updated nomenclature for keratin-associated proteins (KAPs).', *International journal of biological sciences*, 8(2), pp. 258–64. doi: 10.7150/ijbs.3278.

Grymowicz, M. *et al.* (2020) 'Hormonal Effects on Hair Follicles.', *International journal of molecular*

sciences, 21(15). doi: 10.3390/ijms21155342.

Gudiña, E. J. *et al.* (2015a) 'Bioconversion of agro-industrial by-products in rhamnolipids toward applications in enhanced oil recovery and bioremediation', *Bioresource Technology*, 177, pp. 87–93. doi: 10.1016/j.biortech.2014.11.069.

Gudiña, E. J. *et al.* (2015b) 'Biosurfactant production by Bacillus subtilis using corn steep liquor as culture medium', *Frontiers in Microbiology*, 6. doi: 10.3389/fmicb.2015.00059.

Gutiérrez-González, M. *et al.* (2019) 'Optimization of culture conditions for the expression of three different insoluble proteins in Escherichia coli', *Scientific Reports*, 9(1), p. 16850. doi: 10.1038/s41598-019-53200-7.

Hallegot, P. and Corcuff, P. (1993) 'High-spatial-resolution maps of sulphur from human hair sections: an EELS study.', *Journal of microscopy*, 172(Pt 2), pp. 131–6. doi: 10.1111/j.1365-2818.1993.tb03404.x.

He, X. *et al.* (2015) 'Enhanced l-lysine production from pretreated beet molasses by engineered Escherichia coli in fed-batch fermentation', *Bioprocess and Biosystems Engineering*, 38(8), pp. 1615–1622. doi: 10.1007/s00449-015-1403-x.

Horikoshi Toshio, H. J. M. H. I. S. U. T. (1996) 'Hair pretreating agent, hair pretreating cosmetic and method for dyeing hair'. Available at: ,JPH08104614.

Huang Xueying (Hockessin, D. E. U. W. H. (Kennett S. P. A. U. W. Y. (Wallingford P. A. U. (2007) 'Peptide-based conditioners and colorants for hair, skin, and nails'. Available at: ,7220405.

Igarashi Shigeru (Odawara, J. U. T. (Odawara J. H. J. (Yokohama J. H. K. (Odawara J. U. H. (Odawara J. M. U. (Naka-gun J. S. K. (Odawara J. M. H. (Odawara J. H. T. (Chigasaki J. (1997) 'Hair coloring composition comprising anti-hair antibodies immobilized on coloring materials, and hair coloring methods'. Available at: ,5597386.

Ito, S. and Wakamatsu, K. (2011) 'Human hair melanins: what we have learned and have not learned from mouse coat color pigmentation', *Pigment Cell & Melanoma Research*, 24(1), pp. 63–74. doi: 10.1111/j.1755-148X.2010.00755.x.

Kamath, Y. K. and Weigmann, H.-D. (1982) 'Fractography of human hair', *Journal of Applied Polymer*

Science, 27(10), pp. 3809–3833. doi: 10.1002/app.1982.070271016.

KONG, J. and YU, S. (2007) 'Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures', *Acta Biochimica et Biophysica Sinica*, 39(8), pp. 549–559. doi: 10.1111/j.1745-7270.2007.00320.x.

Lai-Cheong, J. E. and McGrath, J. A. (2009) 'Structure and function of skin, hair and nails', *Medicine*, 37(5), pp. 223–226. doi: https://doi.org/10.1016/j.mpmed.2009.03.002.

Lee, P. C. *et al.* (2000) 'Fermentative production of succinic acid from glucose and corn steep liquor byAnaerobiospirillum succiniciproducens', *Biotechnology and Bioprocess Engineering*, 5(5), pp. 379–381. doi: 10.1007/BF02942216.

Leerunyakul, K. and Suchonwanit, P. (2020) 'Asian Hair: A Review of Structures, Properties, and Distinctive Disorders', *Clinical, Cosmetic and Investigational Dermatology*, Volume 13, pp. 309–318. doi: 10.2147/CCID.S247390.

Litwack, G. (2018) 'Metabolism of Amino Acids', in *Human Biochemistry*. Elsevier, pp. 359–394. doi: 10.1016/B978-0-12-383864-3.00013-2.

Malinauskyte, E. *et al.* (2021) 'Penetration of different molecular weight hydrolysed keratins into hair fibres and their effects on the physical properties of textured hair', *International Journal of Cosmetic Science*, 43(1), pp. 26–37. doi: 10.1111/ics.12663.

Martins, M., Silva, C. and Cavaco-Paulo, A. (2019) 'α-Chymotrypsin catalysed oligopeptide synthesis for hair modelling', *Journal of Cleaner Production*, 237, p. 117743. doi: 10.1016/j.jclepro.2019.117743.

McMichael, A. J. (2007) 'Hair Breakage in Normal and Weathered Hair: Focus on the Black Patient', *Journal of Investigative Dermatology Symposium Proceedings*, 12(2), pp. 6–9. doi: 10.1038/sj.jidsymp.5650047.

O'brien John P. (Oxford, P. A. U. W. H. (Kennett S. P. A. U. W. Y. (Wallingford P. A. U. (2010) 'Peptide-based body surface coloring reagents'. Available at: ,20100311641.

Oliver, M. A. *et al.* (2019) 'Lipid loses and barrier function modifications of the brown-to-white hair transition', *Skin Research and Technology*, 25(4), pp. 517–525. doi: 10.1111/srt.12681.

Orwin, D. F. G., Woods, J. L. and Ranford, S. L. (1984) 'Cortical cell types and their distribution in wool fibres', *Australian journal of biological sciences*, v. 37.

Pais, J. *et al.* (2014) 'Improvement on the yield of polyhydroxyalkanotes production from cheese whey by a recombinant Escherichia coli strain using the proton suicide methodology.', *Enzyme and microbial technology*, 55, pp. 151–8. doi: 10.1016/j.enzmictec.2013.11.004.

Plowman, J.E., Duane, P. H. (2018) *The Hair Fibre: Proteins, Structure and Development*. Edited by J. E. Plowman, D. P. Harland, and S. Deb-Choudhury. Singapore: Springer Singapore (Advances in Experimental Medicine and Biology). doi: 10.1007/978-981-10-8195-8.

Powell, B. C. and Rogers, G. E. (1997) 'The role of keratin proteins and their genes in the growth, structure and properties of hair', in *Formation and Structure of Human Hair*. Basel: Birkhäuser Basel, pp. 59–148. doi: 10.1007/978-3-0348-9223-0_3.

Qin, X. *et al.* (2013) 'Enzyme-triggered hydrogelation via self-assembly of alternating peptides', *Chemical Communications*, 49(42), p. 4839. doi: 10.1039/c3cc41794h.

Rahmen, N. *et al.* (2015) 'Exchange of single amino acids at different positions of a recombinant protein affects metabolic burden in Escherichia coli', *Microbial Cell Factories*, 14(1), p. 10. doi: 10.1186/s12934-015-0191-y.

Ribeiro, A. *et al.* (2016) 'BSA/HSA ratio modulates the properties of Ca2+-induced cold gelation scaffolds', *International Journal of Biological Macromolecules*, 89, pp. 535–544. doi: 10.1016/j.ijbiomac.2016.05.012.

Richards, G. M., Oresajo, C. O. and Halder, R. M. (2003) 'Structure and function of ethnic skin and hair', *Dermatologic Clinics*, 21(4), pp. 595–600. doi: 10.1016/S0733-8635(03)00081-0.

Rivas, B. *et al.* (2004) 'Development of culture media containing spent yeast cells of Debaryomyces hansenii and corn steep liquor for lactic acid production with Lactobacillus rhamnosus', *International Journal of Food Microbiology*, 97(1), pp. 93–98. doi: 10.1016/j.ijfoodmicro.2004.05.006.

Robbins, C. R. (2012) *Chemical and Physical Behavior of Human Hair*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-25611-0.

Romano, D. *et al.* (2009) 'Optimization of human d-amino acid oxidase expression in Escherichia coli',

Protein Expression and Purification, 68(1), pp. 72–78. doi: 10.1016/j.pep.2009.05.013.

Rosano, G. L., Morales, E. S. and Ceccarelli, E. A. (2019) 'New tools for recombinant protein production in Escherichia coli : A 5-year update', *Protein Science*, 28(8), pp. 1412–1422. doi: 10.1002/pro.3668.

Sar, T., Stark, B. C. and Yesilcimen Akbas, M. (2017) 'Effective ethanol production from whey powder through immobilized E. coli expressing Vitreoscilla hemoglobin', *Bioengineered*, 8(2), pp. 171–181. doi: 10.1080/21655979.2016.1218581.

Schmidt, S. R. (2013) 'Fusion Proteins: Applications and Challenges', in *Fusion Protein Technologies for Biopharmaceuticals*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 1–24. doi: 10.1002/9781118354599.ch1.

Schweizer, J. *et al.* (2006) 'New consensus nomenclature for mammalian keratins', *Journal of Cell Biology*, 174(2), pp. 169–174. doi: 10.1083/jcb.200603161.

Secchi, G. (2008) 'Role of protein in cosmetics', *Clinics in Dermatology*, 26(4), pp. 321–325. doi: 10.1016/j.clindermatol.2008.04.004.

Silveira, M. M. *et al.* (2001) 'Production of glucose-fructose oxidoreductase and ethanol by Zymomonas mobilis ATCC 29191 in medium containing corn steep liquor as a source of vitamins', *Applied Microbiology and Biotechnology*, 55(4), pp. 442–445. doi: 10.1007/s002530000569.

Song, K., Xu, H. and Yang, Y. (2016) 'Effects of chemical structures of polycarboxylic acids on molecular and performance manipulation of hair keratin Kaili', *RSC Advances*. Royal Society of Chemistry, 6, pp. 58594–58603. doi: 10.1039/C6RA08797C.

Sousa, R. (2013) 'T7 RNA Polymerase', in *Encyclopedia of Biological Chemistry*. Elsevier, pp. 355–359. doi: 10.1016/B978-0-12-378630-2.00267-X.

Sun Ziming (Fountain Valley, C. P. J. W. (San J. C. C. T. D. (Coto D. C. C. N. C. (Garden G. C. (2002) 'Composition and method for hair and scalp treatment'. Available at: ,6455058.

Swift, J. A. (1997) 'Morphology and histochemistry of human hair', in *Formation and Structure of Human Hair*. Basel: Birkhäuser Basel, pp. 149–175. doi: 10.1007/978-3-0348-9223-0_4.

Swift, J. A. (1999) 'Human hair cuticle : Biologically conspired to the owner's advantage', *Journal of*

Cosmetic Science, 50(1), pp. 23–47.

Tinoco, A. *et al.* (2018) 'Keratin-based particles for protection and restoration of hair properties', *International Journal of Cosmetic Science*. Blackwell Publishing Ltd, 40(4), pp. 408–419. doi: 10.1111/ics.12483.

Tinoco, A. *et al.* (2019a) 'Crystallin Fusion Proteins Improve the Thermal Properties of Hair', *Frontiers in Bioengineering and Biotechnology*, 7, p. 298. doi: 10.3389/fbioe.2019.00298.

Tinoco, A. *et al.* (2019b) 'Fusion proteins with chromogenic and keratin binding modules', *Scientific Reports*, 9(1), p. 14044. doi: 10.1038/s41598-019-50283-0.

Tinoco, A. *et al.* (2021a) 'Biotechnology of functional proteins and peptides for hair cosmetic formulations', *Trends in Biotechnology*. doi: 10.1016/j.tibtech.2021.09.010.

Tinoco, A. *et al.* (2021b) 'Keratin:Zein particles as vehicles for fragrance release on hair', *Industrial Crops and Products*, 159, p. 113067. doi: 10.1016/j.indcrop.2020.113067.

Tinoco, A. *et al.* (2021c) 'Proteins as Hair Styling Agents', *Applied Sciences*, 11(9), p. 4245. doi: 10.3390/app11094245.

Villa, A. L. V. *et al.* (2013) 'Feather keratin hydrolysates obtained from microbial keratinases: effect on hair fiber', *BMC Biotechnology*, 13(1), p. 15. doi: 10.1186/1472-6750-13-15.

Yang, J. *et al.* (2015) 'Expression of biologically recombinant human acidic fibroblast growth factor in Arabidopsis thaliana seeds via oleosin fusion technology', *Gene*, 566(1), pp. 89–94. doi: 10.1016/j.gene.2015.04.036.

Yazawa, K. and Numata, K. (2014) 'Recent Advances in Chemoenzymatic Peptide Syntheses', *Molecules*, 19(9), pp. 13755–13774. doi: 10.3390/molecules190913755.

Ye, Q. *et al.* (2010) 'High-level production of heterologous proteins using untreated cane molasses and corn steep liquor in Escherichia coli medium', *Applied Microbiology and Biotechnology*, 87(2), pp. 517–525. doi: 10.1007/s00253-010-2536-0.

Yu, Y. *et al.* (2017) 'Structure and mechanical behavior of human hair', *Materials Science and Engineering: C*, 73, pp. 152–163. doi: 10.1016/j.msec.2016.12.008.

Zhang, X. *et al.* (2017) 'Production of small cysteine-rich effector proteins in Escherichia coli for structural and functional studies', *Molecular Plant Pathology*, 18(1), pp. 141–151. doi: 10.1111/mpp.12385.

SUPPLEMENTARY INFORMATION

1. Calibration curves (OD600 vs Dry weight)

In this annex are presented the calibration curves used to quantify the dry weight (mg/mL) of biomass in order to study the growth kinetics.

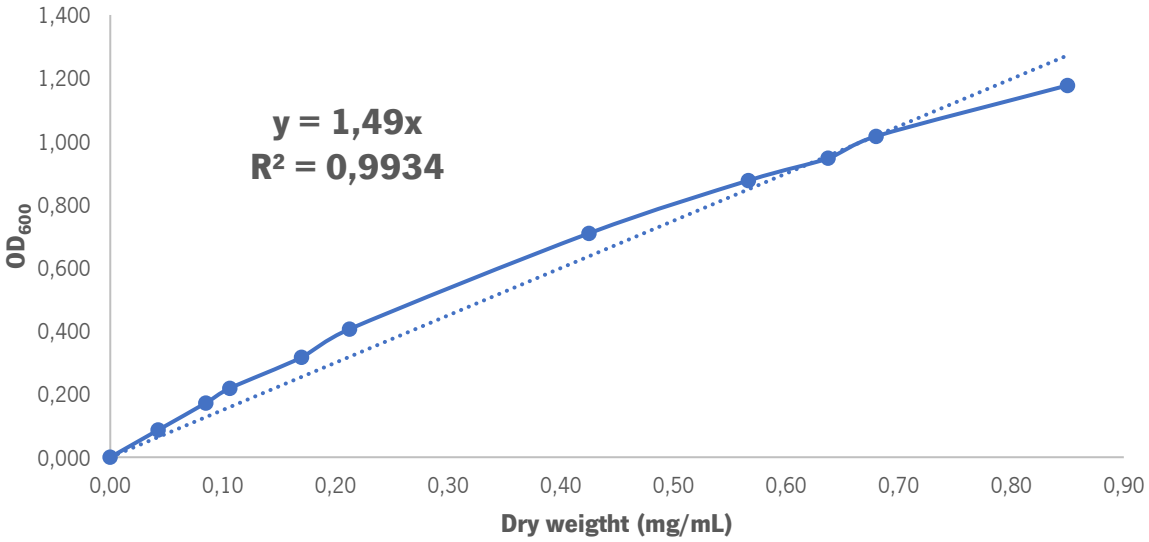

Figure SI 1 – Calibration curve of biomasses in reader 1. The equation OD600=1.49 ± 0.09 [biomasses]

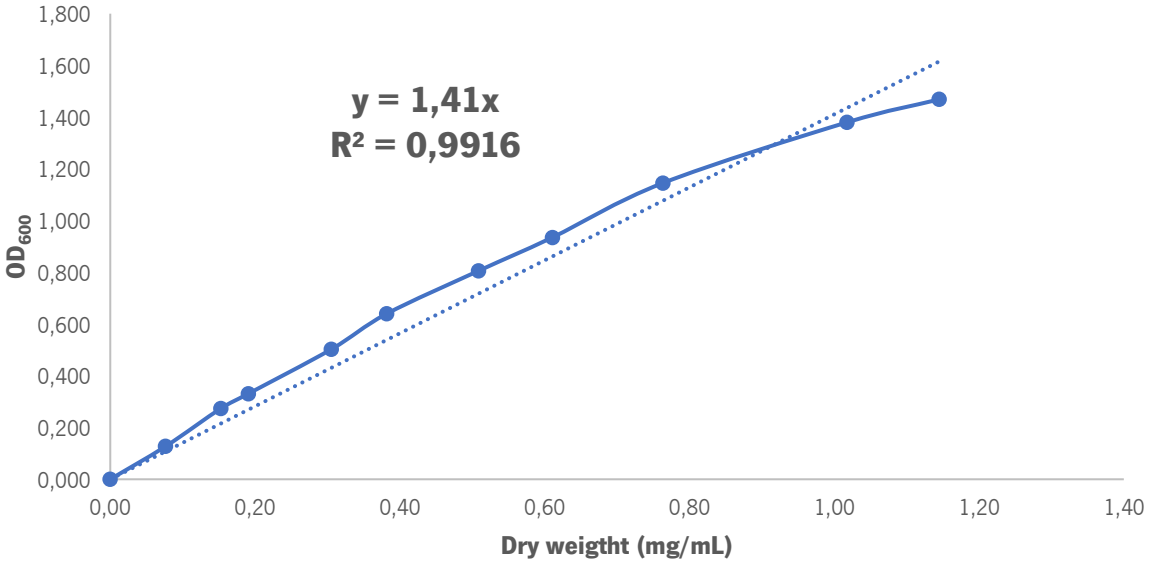

Figure SI 2 – Calibration curve of biomasses in reader 2. The equation OD600=1.41 ± 0.09 [biomasses]