UNIVERSIDADE DA BEIRA INTERIOR
Ciências

Stabilization of COMT by experimental design approaches

Diana Margarida Tavares de Oliveira

Dissertação para obtenção do Grau de Mestre em
Biotecnologia
(2º ciclo de estudos)

Orientador: Professor Doutor Luís António Paulino Passarinha
Co-orientador: Mestre Augusto Quaresma Henriques Pedro

Covilhã, junho 2014

“Conquistas sem riscos são sonhos sem mérito.”

Augusto Cury

Acknowledgments

Em primeiro lugar quero deixar um profundo agradecimento ao Professor Doutor Luís Passarinha e ao Mestre Augusto Pedro não só pela orientação, mas pela confiança que demonstraram em mim para a realização deste trabalho. Foi um privilégio trabalhar neste grupo de investigação que contribui imenso para o enriquecimento da minha formação académica e científica.

Não posso deixar de agradecer a Filipa Correia, Fátima Santos e Luís Martins pelo apoio, companheirismo e ensinamentos transmitidos ao longo deste ano de trabalho.

Quero agradecer aos meus amigos (Diana Rodrigues, Joana Barros, Soraia Ferreira, Filipa Pires, Zélia Proença, Tiago Tavares, Sara Tavares e Ruben Coval) por me terem apoiado neste que foi a maior aventura da minha vida. Obrigada pelas palavras de conforto e "puxões de orelhas", obrigada por todos os momentos que vivemos juntos, guardo todos esses momentos no meu coração.

Por último, mas mais importantes, quero agradecer aos meus pais Rosa e Humberto e irmão Luís pelo apoio que me deram ao longo deste ano. Dedico este trabalho a vocês, vocês merecem tudo de bom. Tenho muito orgulho em vocês, são o meu pilar, a minha razão de viver. Desculpem todas as ausências. Espero um dia poder-vos recompensar por todo o esforço que fizeram para eu poder conquistar este meu sonho, vosso sonho.

Resumo

A catecol-*O*-metiltransferase (COMT) é uma enzima monomérica dependente de magnésio envolvida na inativação de substratos com estrutura catecólica como a dopamina, epinefrina e norepinefrina. A COMT apresenta-se sob duas formas, uma isoforma solúvel (SCOMT) cuja localização é o citoplasma e uma isoforma membranar (MBCOMT) que se encontra predominantemente na membrana do retículo endoplasmático rugoso. Ambas as isoformas são extremamente termolábeis, sendo que a MBCOMT apresenta maior afinidade (menor K_M) para todos os seus substratos, com exceção dos estrogénios com estrutura catecólica. Adicionalmente, tem sido descrito que estas enzimas possuem um papel importante em diversas patologias humanas tais como cancros induzidos por estrogénios, doenças cardiovasculares ou doenças neurodegenerativas. Atualmente, a terapêutica para a doença de Parkinson consiste numa tripla profilaxia incluindo levodopa, um inibidor da enzima descarboxilase dos aminoácidos aromáticos e um inibidor da COMT. Visto que ambas as isoformas são expressas em baixos níveis nos tecidos, para ser possível a sua caracterização estrutural ou funcional (que permita por exemplo, o desenvolvimentos de novos inibidores com potência e eficácia melhoradas), geralmente estas enzimas são expressas de forma recombinante utilizando um microrganismo adequado, tendo como objetivo final a obtenção de extratos proteicos em quantidade suficiente e pureza adequada que permita a realização dos estudos anteriormente mencionados.

Em particular, neste trabalho, pretendeu-se estabelecer uma formulação que permita o armazenamento da MBCOMT recombinante no seu estado nativo e, consequentemente, com atividade biológica obtida a partir de lisados de *Pichia pastoris*. Sendo assim, foi aplicado um desenho experimental (DOE) de forma a poder encontrar as condições ótimas de armazenamento da MBCOMT em termos da concentração dos estabilizadores (trealose, glicerol e cisteína), temperatura (-80, -20 e 4ºC) e tempo de armazenamento (24, 48 e 72h).

Com este trabalho pretendeu-se encontrar as condições ótimas de armazenamento que permitam manter a MBCOMT num estado nativo pelo máximo tempo possível sem comprometer a sua atividade biológica. Para alcançar este objetivo foi criada uma rede neuronal onde através da análise dos gráficos de resposta de superfície, foi-nos dado o ponto ótimo onde a percentagem de recuperação de atividade biológica da MBCOMT foi máxima (114.56 %). Assim sendo, uma formulação contendo 100 mM Trealose, 30 %(v/v) Glicerol e 150 mM Cisteína permitiu a estabilização da MBCOMT recombinante durante 72 horas a 4 ºC, como foi comprovado pelo ligeiro aumento na atividade biológica da MBCOMT às 72 horas, em comparação com os valores obtidos imediatamente após a lise celular (0 horas).

Em conclusão, pela primeira vez efetuou-se um estudo sistemático acerca das melhores condições para o armazenamento da MBCOMT recombinante, permitindo obter uma formulação que mantém a MBCOMT num estado nativo e biologicamente ativo durante 72 horas.

Palavras-chave

COMT, Proteína Membranar, *Pichia pastoris*, Estabilização proteica, Rede neuronal

Abstract

Catechol-*O*-methyltransferase (COMT) is monomeric enzyme magnesium dependent that is involved in the inactivation of catechol substrates such as dopamine, epinephrine and norepinephrine. The COMT enzyme is present as two isoforms, a soluble isoform (SCOMT) that is found in the cytoplasm and a membrane-bound isoform (MBCOMT) which is present mainly in the rough endoplasmatic reticulum membrane. Both isoforms are extremely thermolabile and, except for catechol estrogens, which possess similar K_M values for both isoforms, MBCOMT presents higher affinity (lower K_M) for the catechol substrates. In the last years, these enzymes have been described as being involved in several human diseases such as catechol-estrogen induced cancers, cardiovascular diseases and degenerative brain disorders. Actually, the most effective treatment for Parkinson's disease is a triple prophylaxis consisting of the dopamine replacement with levodopa together with an inhibitor of aromatic amino acid decarboxylase and an inhibitor of COMT.

Since both COMT isoforms are present in tissues at low concentrations, in order to accomplish its structural or functional characterization, these enzymes are usually expressed in a recombinant form using a suitable host, in which the final aim is the achievement of protein extracts with enough quantity and purity, enabling structural and functional studies.

Specifically, in this work, it was intended to establish a formulation that allows the storage of recombinant and biologically active MBCOMT in a native state from *Pichia pastoris* lysates. Therefore, it was applied a design of experiments (DOE) in order to obtain the optimum conditions for MBCOMT storage according to the stabilizers concentrations (trehalose, glycerol and cysteine), the storage temperature (-80 ºC, -20 ºC and 4 ºC) and the storage time (24, 48 and 72 hours).

The global aim of this work was to establish the optimum storage conditions that allow MBCOMT to maintain its biological activity in a native state for as long as possible without compromising its biological activity. Thus, in order to achieve this aim, an artificial neural network was created where by the analysis of the response surface graphics, a specific run (optimum point) was obtained where the percentage of the MBCOMT biological activity recovery was the highest (114.56 %). Therefore, a specific formulation with 100 mM trehalose, 30 % (v/v) glycerol and 150 mM cysteine allow MBCOMT stabilization during 72 hours at 4 ºC, as judged by the slightly improvement in MBCOMT biological activity at 72 hours, when compared with the MBCOMT biological activity obtained immediately after cell lysis (0 hours).

In conclusion, for the first time, a systematic study was carried out in order to evaluate the best conditions to storage recombinant MBCOMT, allowing to obtain a specific formulation that maintain MBCOMT in a native and biologically active state during 72 hours.

Keywords

COMT, Membrane Protein, *Pichia pastoris*, Protein Stabilization, Artificial Neural network.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ANN	Artificial neural network
AOX	Alcohol oxidase
AOX_1	Alcohol oxidase 1 gene
AOX_2	Alcohol oxidase 2 gene
BMGY	Buffered Glycerol-complex medium
BMMY	Buffered Methanol-complex medium
BSM	Basal salts medium
COMT	Catechol-*O*-methyltransferase
COOH	Carboxyl group
DOE	Experimental design
DTT	Dithiothreitol
EGTA	Glycol-bis(2-aminothylether)-N,N,N',N'-tetraacetic acid
FM22	Standard medium alternative to the BSM
HIS4	Selectable marker
HPLC	High-performance liquid chromatography
KH_2PO_4	Monopotassium phosphate
$KHPO_4$	Dipotassium phosphate
K_M	Michaelis constant
MBCOMT	Membrane-bound catechol-*O*-methyltransferase
Met	Methionine
$MgCl_2$	Magnesium chloride
mRNA	Messenger ribonucleic acid
Mut^+	Methanol utilization plus phenotype
Mut^s	Methanol utilization slow phenotype
NaH_2PO_4	Sodium acetate anydrous
NH_2	Amine group
MN	Metanephrine

OSA	Sodium octil sulfate
OD_{600}	Optical density 600 nm
P. pastoris	*Pichia pastoris*
PD	Parkinson's disease
PCR	Polymerase chain reaction
RSM	Response surface methodology
SAM	S-adenosly-*l*-methionine
SCOMT	Soluble form of catechol-*O*-methyltransferase
Tg	Glass transition temperature
Val	Valine
Vmax	Maximum velocity
YPDS	Yeast Extract Peptone Dextrose Medium
3-OMD	3-*O*-methyldopa
3,5-DNC	Inhibitor 3,5-dinitrocatechol

Chapter 1 – Introduction

1.1 Catechol-*O*-methyltransferase: an overview

1.1.1 COMT physiological role

Catechol-*O*-methytransferase (COMT; EC 2.1.1.6) is a monomeric and magnesium-dependent enzyme that catalyses the methylation of catechol substrates using S-adenosyl-*l*-methionine (SAM) as a methyl donor (see figure 1), being involved in the inactivation of catechols and catecholamines neurotransmitters like dopamine, epinephrine and norepinephrine [1].

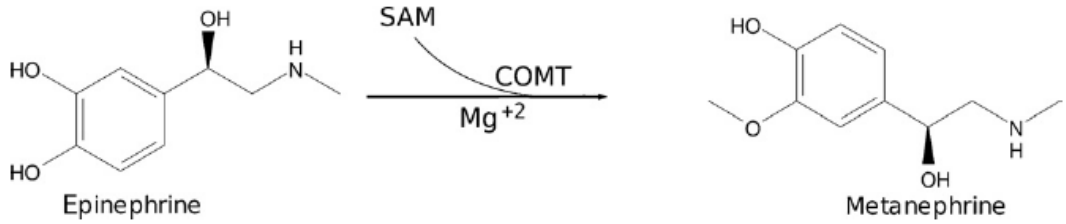

Figure 1 - Reaction catalyzed by COMT [1].

As a result of the reaction catalyzed by COMT, *O*-methylated compounds [2, 3] are obtained. In fact, the *O*-methylation reactions catalyzed by COMT are an important mechanism to detoxify biotransformation or inactivate toxic catechol compounds, which may be endogenous or exogenous to the organism [3, 4]. COMT substrates include a wide variety of catechols, including catecholamines, their hydroxilated metabolites, catecholestrogens, ascorbic acid, dietary phytochemicals and medicinal compounds [2]. Moreover, it has been reported an association between the human COMT gene and various diseases affecting the central nervous system such as neurodegenerative diseases [5, 6].

1.1.2 Soluble and membrane-bound COMT isoforms

COMT is present in prokaryotes and eukaryotes, namely bacteria, yeast, plants and animals [2]. In humans, the COMT gene is located on chromossome 22 band and is composed of six exons, whereas for example the rat COMT gene is located on chromosome 11 and is composed of five exons [2].

The COMT gene encodes two major transcripts, the long and short transcript. The long transcript can be translated into two major isoforms of COMT, the membrane-bound and soluble forms [6].

The MBCOMT (membrane-bound catechol-*O*-methyltransferase) is a membrane protein that is associated with the rough endoplasmatic reticulum membrane and contains 50 additional amino acids when compared with the soluble form, resulting in a total of 271 amino acids corresponding to a final molecular weight of 30 kDa (in humans). In fact, this extra peptide contains a strech of 21 (humans) hydrophobic amino acid residues that constitute the membrane anchor region [1, 6].

On the other hand, the human SCOMT (soluble catechol-*O*-methyltransferase) contains 221 amino acids residues and a molecular weight of 24.7 KDa (humans) [7] and is found mainly in the cytoplasm [5].

Although the intracellular localization of soluble and membrane-bound COMT isoforms differ [7], both enzymes share similar affinities for SAM, similar magnesium dependency, inhibition by calcium, and a similar optimum pH for the activity, but they can have significantly different affinities for substrates [2].

Moreover, MBCOMT acts in the synaptic dopaminergic neurotransmission termination when there is low physiologic concentrations of catecholamines [2], while SCOMT is responsible for the biological removal of catecholamines by acting as a detoxifying agent between the blood and other tissues [8]. Finally, while MBCOMT is the COMT isoform predominant in the brain, SCOMT is present in higher quantities in peripheral tissues such as the liver and kidney [9].

X-ray experiments revealed that the active site of SCOMT consists in a catalytic site, which binds one magnesium ion and the cathecol substrateas as well as the SAM-binding region [2]. The analysis of the X-ray structure and of kinetics experiments made it possible to establish that the methylation reaction follows a sequentially ordered mechanism with SAM binding first, followed by ion magnesium (Mg^{2+}) and, finally, the catechol substrate [2].

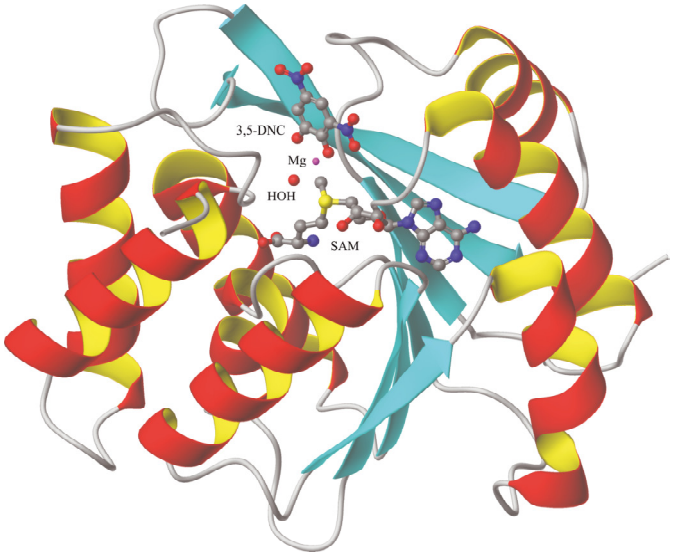

Figure 2 - Schematic representation of the three-dimensional structure of COMT. The SAM co-substrate, the inhibitor 3,5-dinitrocatechol (3,5-DNC), the magnesium ion, and coordinated water molecules water [2].

The magnesium ion is responsible for the binding of the substrate (or inhibitor) in the catalytic site and catalytically is important because while it is positively charged, the catechol substrate presents a negative charge [2].

Some immunohistochemical studies show that residues of COMT are present at low levels in postsynapic neurons instead of presynaptic dopaminergic neurons [10]. On the other hand, MBCOMT is present in presynaptic neurons and the catalytic site is oriented to the extracellular space [10]. Studies showed that dopamine must be carried out for the catalytic site so that it can be metabolized by COMT [6].

Another aspect to consider is the orientation of MBCOMT in the cell membrane, since its orientation is critical for synaptic mechanism and inactivation of dopamine. The C-terminal domain has a catalytic domain that is oriented out of the neuron and thus the extracellular dopamine is endocytosed and transported to be inactivated by MBCOMT. If the C-terminal domain is oriented directly into the extracellular space, dopamine is directly inactivated by MBCOMT [6].

1.1.3 COMT biological activity: Affinities of soluble and membrane-bound COMT for several substrates

Kinetic data showed that MBCOMT presents a higher affinity for the substrates, but lower reaction rate when compared with SCOMT, thus being the main responsible for the catecholamines metabolism in the brain [5, 7, 9].

As can be seen in Table 1, MBCOMT has a higher affinity for the substrates when compared to SCOMT isoform [11], suggesting that MBCOMT is the most important at a physiological level, i.e., it is more relevant to the metabolism of catecholamines *in vivo* [7]. The reasons for these differences remain to be determined. However, it can be hypothesized that additional interactions between the substrate and the extra N-terminal residues of MBCOMT (absent in SCOMT), or a change in the conformation of the active site, due to membrane interactions, could account for such differences [2].

According to Table 1, the kinetic parameter K_M varies according to the enzyme source tissue and organism as well as the substrate. Moreover, with exception of the catecholestrogens, MBCOMT has a higher affinity for all substrates (lower K_M) when compared with SCOMT. Finally, the K_M is greater in humans compared for example with rat tissues [2].

Table 1 - Affinities of Soluble and MBCOMT, from different sources, for several substrates [2].

	Enzyme source	Substrate	SCOMT	MBCOMT
			K_M (μM)	
Human	Brain	Dopamine	280	3,3
	Recombinant *E.coli*	Catechol	108	10
	Recombinat Sf9 cells	Dopamine	207*	15*
		Norepinephrine	369*	24*
		DBA	39*	30*
		Levedopa	613*	266*
Rat	Liver and Brain	Epinephrine	168-345	0.9-3
	Liver, Brain and Kidney	Norepinephrine	304-464	5.5-11.4
Pig	Brain	R-Salsolinol	156	43
Mouse	Liver	Epinephrine	242	12
Rabbit	Aorta	2-Hydroxyestradiol	0.27	0.15
		Isoproterenol	121	0.91

*Values for the 3-*O*-methylation.

1.1.4 COMT genetic polymorphisms

The polymorphisms of the COMT gene are involved in a variety of phenotypes and have associated COMT with several human diseases, namely Parkinson's disease (PD), Alzheimer's disease, breast cancer, obsessive-compulsive disorder, agressive and suicidal manisfestations of schizophrenia [1, 12]. The most studied polymorphism is the Val 108/158Met (respectively for soluble and MBCOMT) that is characterized by the substitution of a valine for a methionine (Val → Met) in the polypeptide chain [2, 13]. According to this polymorphism, in humans, the COMT activity is divided into three groups: a group that presents high activity ($COMT^{H/H}$), followed by a group with intermediate activity ($COMT^{H/L}$) and a group that display the lowest levels biological activity ($COMT^{L/L}$). The Met 158 variant of this polymorphism is associated with a decreased COMT activity and low thermal stability, thereby increasing dopamine levels, while the Val158 variant exhibits increased biological

activity [14], as shown in picture 3. In fact, the substitution Val158Met leads to a decrease in COMT enzymatic activity of approximately 40% in the brain [15].

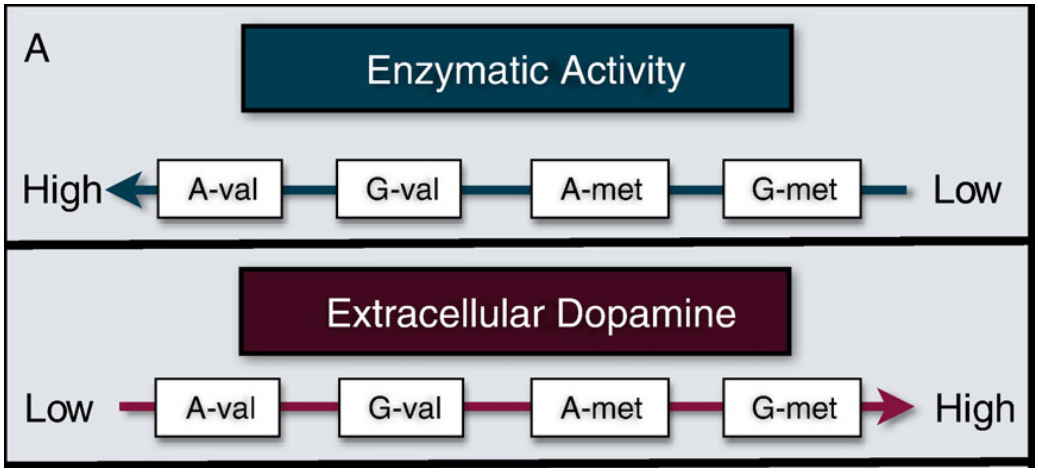

Figure 3 - Relationship between enzyme activity and dopamine levels in several genetic polymorphisms reported for COMT [15].

Interestingly, individuals who have the genotype met/met present COMT lower biologically active and, consequently, display higher levels of dopamine but they present a greater ability to memorize. The reverse occurs for individuals who possess the val/val genotype [9]. Other polymorphisms reported to the human COMT gene (described in table 2) have also been shown to affect mRNA levels, mRNA structure and have been associated with clinical phenotypes [1, 3].

Table 2- Table showing polymorphisms previously reported for the COMT gene [1].

S. No	Polymorphism	NCBI SNP id	COMT activity[a]
1.	Val108Met (G/A)	rs4680	Yes
2.	Ala22Ser (G/T)	rs6267	No
3.	His12His (C/T)	rs4633	No
4.	Leu86Leu (C/G) and (C/T)	rs4813	No
5.	Ala52Thr (G/A)	rs5031015	No

[a]COMT activity refers to if function of COMT is affected by the mutation.

1.1.5 COMT Implication in disease

In recent decades the COMT enzyme has been implicated in different human diseases such as PD, obessive-compulsive disorder, schizophrenia, depression and estrogen-related cancers [12, 15–17].

In particular, PD is one the most common neurodegenerative diseases [18], which causes deficiencies in motor and cognitive level resulting in the loss of quality of life [19]. Additionally, PD is characterized by tremor, rigidity, brandykinesia and postural instability, being one of the first symptoms the instability in movements [20].

The major symptom of PD results in progressive decrease of dopaminergic neurons with consequent decrease in dopamine levels [20, 21]. Actually, the most effective treatment for PD remains to be a triple prophylaxis consisting of levodopa, a peripheral amino acid decarboxilase inhibitor and a COMT inhibitor [2].

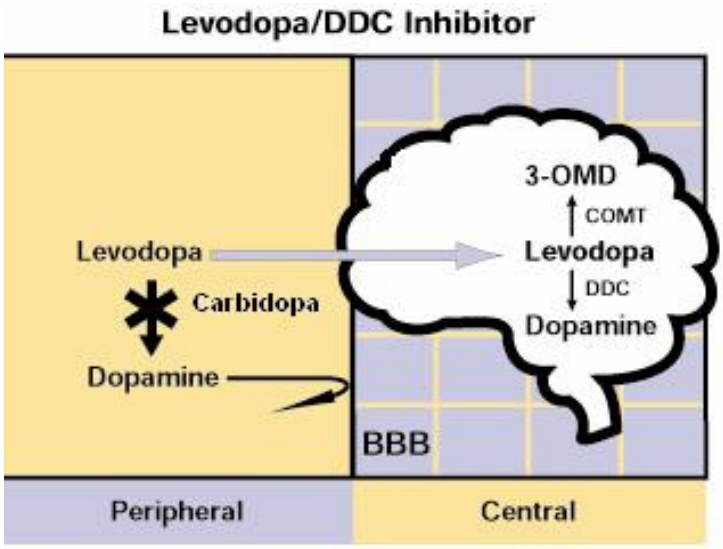

Figure 4: Mechanism of action of the triple prophylaxis therapy for PD [21].
(BBB – Blood-brain barrier; COMT – catechol-*O*-methyltransferase; DDC – Decarboxilase amino acid aromatic; 3-OMD – 3-*O*-methyldopa).

The mechanism of action of this therapy is depicted in figure 4 and involves the conversion of levodopa to dopamine in the target region of the brain, thus facilitating the action of amine groups in the receptor sites [2]. Then, COMT inhibitors prevent normal *O*-methylation of levodopa to its metabolite 3-*O*-methyldopa (3-OMD) limiting this metabolic pathway, thereby increasing the availability of the conversion of levodopa to dopamine in the brain [2]. However, the PD therapy will depend on the polymorphisms of the COMT gene involved in the metabolism of the gene encoding the drug target. Thus, the potential of gene/drug is related to COMT activation or inhibition [21].

COMT is a key enzyme in the modulation of dopamine levels in brain , being a good candidate for treating various neurodegenerative diseases and in particular Parkinson's disease [22]. In addition, kinetic studies showed that MBCOMT is the main responsible for the catecholamines catabolism in the brain and, for this reason, MBCOMT is an important therapeutic target for the treatment of PD [11].

In addition to COMT, other enzymes play an important role in PD therapy, such as monoamine oxidase, an enzyme that has two isoforms that are involved in the degradation of neurotransmitters such as norepinephrine, epinephrine, dopamine, and serotonin [23].

1.2 MBCOMT Biosynthesis

Membrane proteins are drug targets for a wide range of diseases and the achievement of appropriate samples for further research encourages the pharmaceutical industry for developing new and more effective drugs. In fact, as the majority of medical and pharmaceutically relevant membrane proteins, MBCOMT is present in tissues at low concentrations, which makes the screening of new heterologous expression systems in large-scale production-adapted cells a prerequisite for structural and functional studies [20]. During the last decades, several expression systems have been applied successfully to the biosynthesis of recombinant human MBCOMT.

1.2.1 Recombinant systems available for MBCOMT biosynthesis

Several systems have been employed for the recombinant expression of MBCOMT in a biologically active form. Concerning the prokaryotic systems, *Escherichia coli* SG13009 [24] and BL21 [25] were able to produce MBCOMT at relatively high levels. More recently, an entire flowsheet for MBCOMT recombinant expression and purification using *Brevibacillus choshinensis* as the host was successfully reported [5, 7, 11]. On the other hand, over the last years, several eukaryotic systems such as Sf9 insect cells [26] transfected human embryonic kidney fibroblast cell lines [20] and human HeLa and hamster BHK-cells [20] were also applied for MBCOMT biosynthesis.

1.2.2 *Pichia pastoris* as a recombinant microfactory

Pichia pastoris (*P. pastoris*) is a methylotrophic yeast that can be modified by genetic engineering techniques and, thereby, producing the heterologous protein of interest [27]. In fact, *P. pastoris* is capable of producing proteins at high levels, both intracellularly or extracellularly and its cell machinery can introduce diverse post-translational modifications such as glycosylation, dissulfide bridges formation and proteolytic processing [27]. With all

these changes, the *P. pastoris* can be useful to produce proteins with therapeutic and commercial interest [28].

The *P. pastoris* can use integrative or autonomous plasmids. In order to produce proteins in a recombinant way, the gene of the protein of interest included in the expression vector needs to be integrated into the *P. pastoris* genome, more precisely into the alcohol oxidase (AOX) locus, which has the AOX_1 and AOX_2 gene [28]. As demonstrated in figure 5, the expression vector containing the gene of interest can be integrated in the *P. pastoris* genome by gene replacement at a specific site of the *P. pastoris* genome [28].

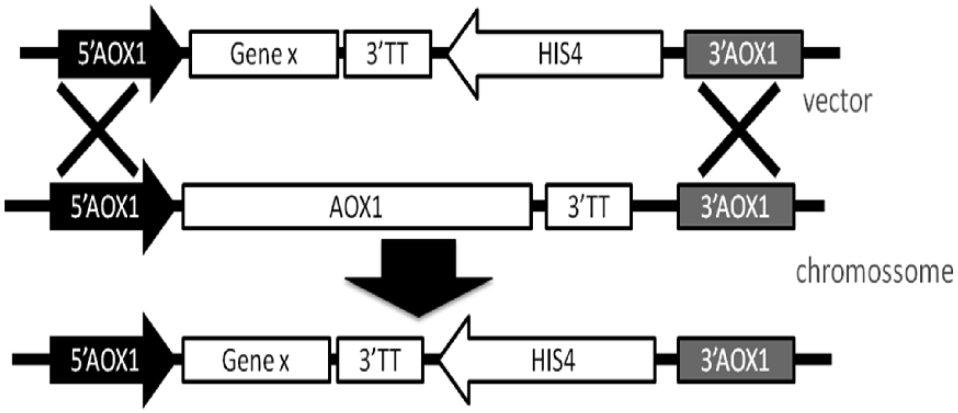

Figure 5 - Integration of the vector by gene replacement into the specific locus of the *Pichia pastoris* genome [28].

Alternatively, the vector can be linearized and oriented to integrate into the genome by non-disruptive insertion HIS4 (marker), thus generating strains that grow normally on methanol (Mut $^+$), as shown in Figure 6 [13, 14].

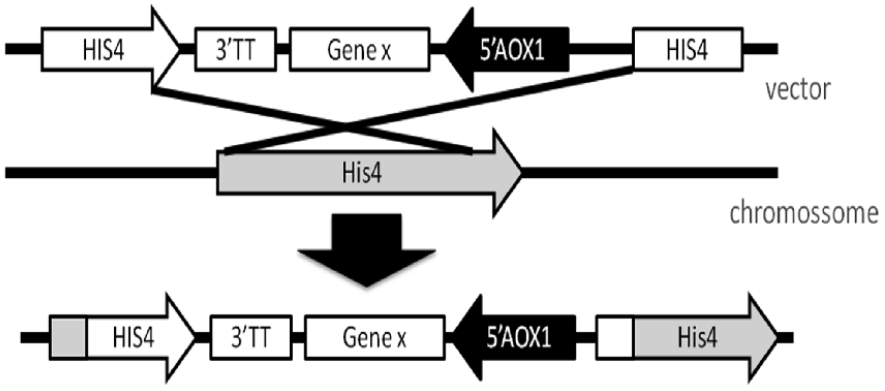

Figure 6 - Integration of the vector by gene insertion at hIS4 into the the *Pichia pastoris* genome [28].

Actually, there are many *P. pastoris* strains available for recombinant protein production (see table 3) [20]. Specifically, the use of protease-deficient strains such as SMD1168 or MC 100-3 may be extremely useful when the recombinant protein is secreted into the extracellular medium, since its proteolysis is avoided [27].

Table 3 – *P. pastoris* strains commonly applied for recombinant protein production. Adapted from [28].

Strain	Genotype	Phenotype
Y-11430	Wild type	-----
X-33	Wild type	-----
Auxotrophic strain		
GS115	his4	Mut^+, His^-
Protease-deficient strains		
KM71	Δaox 1::SARG4 his4 arg4	Mut^S, His^-
SMD1168	His4, pep4	Mut^+, His^-, pep^{4-}
SMD1165	His4, prb1	Mut^+, His^-, $peb1^-$
MC 100-3	Arg4his4aox1Δ::SARGAaox2Δ::Phis4	Mut^+, His^-

The *P. pastoris* is a yeast capable of metabolizing methanol since it presents a unique set of enzymes (alcohol oxidase, catalase and dihydroxyacetone synthase) (as shown in figure 7) that enable this yeast to convert methanol as the sole source of carbon and energy [27].

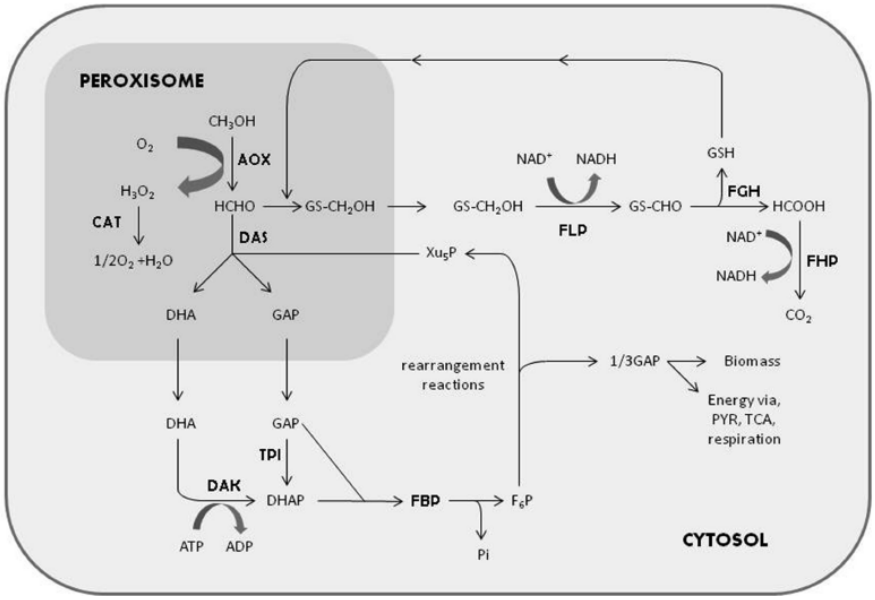

Figure 7 - Metabolic pathway that use methanol as a carbon source can be divided into three steps. **Step 1:** The methanol is oxidized to formaldehyde by alcohol oxidase. **Step 2:** Formaldehyde is rapidly oxidized to formic acid by formaldehyde dehydrogenase enzyme. **Step 3:** The formic acid is metabolized by dehydrogenase (formic acid is toxic in the main metabolic pathway that use methanol) [28].

The tight regulation and high level of AOX1 expression make AOX a strong promoter for the expression of heterologous proteins, being the most widely reported and utilized of all available promoters [27]. As a matter of fact, the alcohol oxidase 1 (AOX_1) is responsible for over 90% of the activity in the cell, while the alcohol oxidase 2 (AOX_2) accounts for less than 10% of the activity of the cell [30].

According to the functionality of AOX genes, there are *P. pastoris* strains with a Mut^+ (Methanol utilization plus) phenotype that contain both functionally AOX genes or Mut^S (methanol utilization slow) strains where there is a deletion in the AOX1 gene or in both genes, the AOX 1 and the AOX [28, 31]. Therefore, strains with a Mut^s phenotype present little capacity to metabolize methanol and, usually, the culture medium is supplemented with an additional carbon source such as sorbitol or trehalose in order to maintain productivity of the target bioproduct [28]. On the other hand, when the Mut^s phenotype is used, less methanol is consumed, sometimes with the same productivities as for the Mut^+ phenotype but with longer induction times [31].

As can be seen in figure 8, for the production of recombinant human growth hormone production in *P. pastoris*, the growth of Mut^+ strains in methanol is much higher when compared with strains with a Mut^s phenotype [29].

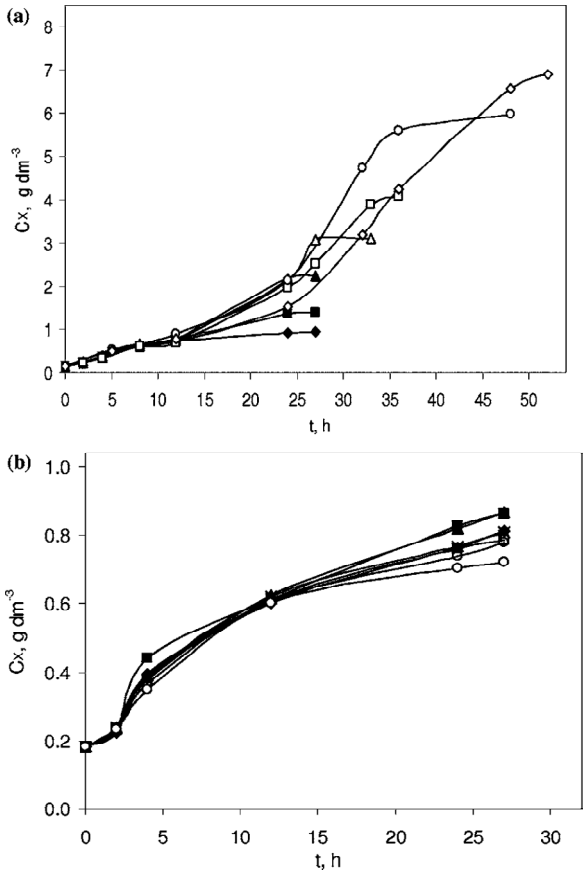

Figure 8 - Change in cell concentration with cultivation time and the concentration of methanol. (a) strain of *P. pastoris* hGH-Mut^{+}; (b) strain of *P. pastoris* initial hGH-Mut^{S} [32].

Although *P. pastoris* possess the capacity to metabolize methanol, high methanol levels can be toxic to the cell while low levels may be insufficient to initiate transcription [28], [33]. The metabolic pathway that use methanol may have some disadvantages such as the formation of hydrogen peroxide and toxic intermediates which together may affect cell viability and decrease the production of heterologous proteins [34]. Therefore, other promoters are available such as the GAP and the YPT1 that work in a constitutive strategy [27].

Finally, the *P. pastoris* cells allows the production of the recombinant proteins either intra or extracellularly whenever it contains the signal sequence with the information targeting the product to the secretory pathway [35]. Moreover, for the production of intracellular proteins is suggested to use Mut^{s} phenotype, as this will have fewer AOX protein, may be easily purified [36].

1.2.2.1 Bioprocess design and operating conditions for recombinant protein production

Expected yields of the *P. pastoris* expression system in the production of heterologous proteins have met new levels of demands, and so, bioprocess design including all the parameters inherent has a great influence on the final result and quality of the target protein [27]. In fact, regardless of which system to use, there are some parameters that must be controlled to maximize the production of the protein of interest.

The culture medium is the key to expressing high levels of protein in *P. pastoris* [37]. The composition of the culture medium have an effect on the production of the protein and on cell growth and viability [28]. In *P. pastoris*, complex media such as BMGY or BMMY consisting of biotin (micronutrient), phosphate buffer (used to control the pH), glycerol or methanol as carbon and energy sources are suitable for the expression at a small scale in shake-flasks [28]. Moreover, depending on the strain, the culture media may need to be supplemented with essential components for *P. pastoris* growth, such as histidine and arginine [28]. On the other hand, alternative defined media such as BSM (basal salts medium) or FM22 medium (Alternative formulation to the BSM) [28, 38]may be used in bioreactor cultures where ammonia is used as nitrogen source but also works to control the pH during fermentation [35]. As a matter of fact, in the formulations of the defined media, the chosen nitrogen source needs to be taken into account since its elimination may increase the action of proteases, thus reducing protein stability [30].

The culture media composition may have important implications in the overall productivity since the addition of specific components such as peptone can reduce or even prevent the proteolytic degradation the target protein [28]. Another way to prevent protease activity by enzymatic competition is through the addition of amino acids to the culture media that also promote *P. pastoris* growth [30].

Both the complex and defined media can be supplemented with macro and micronutrients (such as Fe, Mn, Cu and biotin) as this can increase cell productivity [38].

Another required parameter is the pH control. Monitoring of pH is essential for obtaining cell growth and productivity [39]. Also, the pH control is important for the protease activity [33]. Therefore, in order to control the activity of proteases and maximize protein production, the pH should be between 3.0 and 7.0 [18, 21]. The pH affects cell growth but also affects the stability of the protein and influences protease activity. If the pH is not suitable, undesired precipitation of some components of the culture media can occur [30].

The temperature is a critical parameter of operation [40] and the optimum temperature for growth *P. pastoris* is 30 ºC [33]. However, it had been reported that using lower temperatures than 30 °C until 20 ºC can benefit the production levels of recombinant proteins [33]. Finally, in order to maximize cell growth and protein production, it should be ensured intensive stirring at 250 rpm [41].

1.2.2.2 Advantages and disadvantages of *Pichia pastoris* for recombinant protein expression

The type and size of the target protein determines the most suitable expression system for the expression of the biomolecule of interest [28]. In fact, *P. pastoris* as an eukaryotic organism that is easily genetically manipulated, allowing to obtain high growth rates coupled to high protein levels has been extensively used in the production of eukaryotic membrane and soluble proteins [27, 33].

Furthermore, *P. pastoris* is able to secrete soluble proteins with a complete post-translational maturation, one advantage over the traditional expression system based on *Escherichia coli*. In addition, *P. pastoris* is readily distinguished from other yeasts such as in *S. cerevisiae* since the latter has a smaller capacity compared to the glycosylation *P. pastoris* [25].

In table 4, is presented a short summary of the advantages and disadvantages of using *P. pastoris* for the production of recombinant proteins compared to other expression systems.

Table 4 - Summary of the advantages and disadvantages of using *P. pastoris* as an expression host for recombinant protein production.

	Advantages	Disadvantages	Ref.
E . coli	- Allows a rapid growth rate and the production of proteins at high levels. - Low cost compared to other expression systems.	- Does not allow post-translational modifications because it is a prokaryotic organism.	[4, 7, 27, 43]
S. cerevisiae	- Show great similarity with the *P. pastoris*, having some applicability in the production of proteins.	- Have less glycosylation capacity compared with the P. pastoris. - Requires defined culture media.	[27, 44]
P. pastoris	- Easy to manipulate and cultivate. - It allows the expression of high levels of protein with post-translational modifications: - Processing of the final sequence - Disulfide bridge formation - Folding - Addition of certain lipids - Glycosylation - Methylation - Acetylation - Produce soluble and membrane proteins to their proper folding. - Secrete proteins, since the plasmid has used this signal sequence.	- Control of various operational parameters (pH, temperature, aeration) is required. - Risk of proteolysis. - *P. pastoris* instability due to the formation of intermediate compounds catabolism of methanol are toxic to cells (formaldehyde, hydrogen peroxide).	[27, 30, 33, 39, 40,42 43, 46]
Mammalian cells	- Are a good option in terms of glycosylation, as there are addition of N-glicolilneuramic, leading to development of immune responses.	- Requires very complex culture media. - Requires control of several operational parameters such as temperature, pH and aeration.	[28]

1.2.2.3 Strategies for the optimization of membrane protein expression in *Pichia pastoris*

Membrane proteins are very common as a drug target. Expression of membrane proteins is difficult and requires a balance between the translation and membrane folding. In the last decade the *P. pastoris* expression system started to be used in the expression of membrane proteins [44]. Indeed, one of the major factors that greatly influences the production levels of membrane proteins is the gene dosage used [45].

Currently, strategies to increase the production of membrane proteins in *P. pastoris* have been developed. The following table summarizes some strategies that have been applied in the production of membrane proteins in *P. pastoris* for improving the overall yield of the process.

Table 5 - Strategies applied in the production of membrane proteins in *P. pastoris* [45].

Protein	**Strategies used to improve production**
Human AQPs	– Use of mammalian Kozak's consensus for the ATG
G protein-coupled receptors	– Low culture temperature (20°C) – Supplementation of media with specific ligands – Addition of DMSO (chemical chaperone)
G protein-coupled receptors	– Optimization of culture scale, pH and temperature
G-protein coupled receptor adenosine A2A	– Activation of UPR through overexpression of Hac1p
Rat K^+ channel	– Elimination of glycosylation sites
Human μ-opioid receptor	– Enhanced translocation to the membrane through fusion to *S. cerevisiae* α-factor signal
Mammalian *Shaker* family Kv channel	– Use of detergents and lipids during purification and crystallization

1.3 Role of Protein Stabilizers

The majority of the enzymes, including both COMT isoforms, loses its activity rapidly even at physiological temperature [3]. Indeed, one of the major bottlenecks in working with proteins is that they are extremely sensitive to environmental conditions such as pH, temperature and the presence of salts [46]. Therefore, inappropriate storage conditions can lead to the formation of aggregates (as shown in figure 9) that may be irreversible and consequently lead to the loss of protein activity [47]. To prevent these problems, it is common the application of stabilizers during the hands-on in a laboratory platform [48, 49].

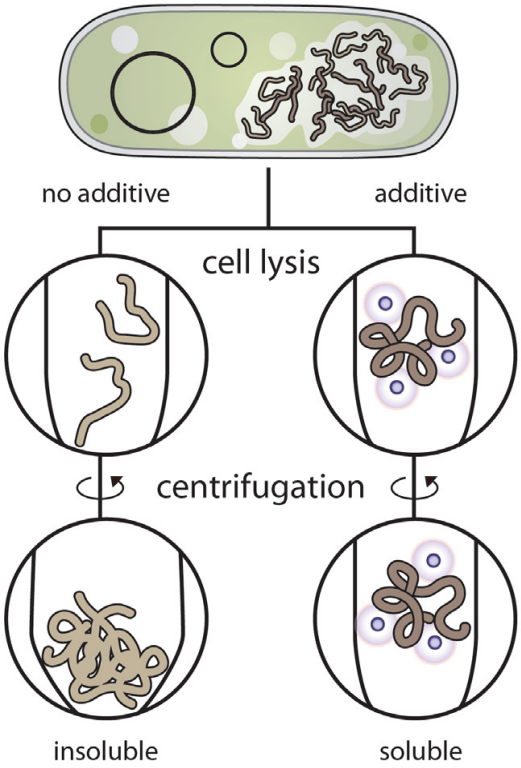

Figure 9 - Comparison of the structural stability of a protein with and without the addition of stabilizers [46].

The stabilizers can be classified into several categories according to their function. They can be amino acid, additives that reduces protein-protein interactions (chaotropic agents), additives that stabilize intra-molecular bonds (kosmotropic agents) and additives that affect protein stability such as reducing agents, sugars and polyols [46].

The trehalose and sucrose are sugars, acting as thermal stabilizer [50, 51]. The addition of these stabilizers to a protein will increase the glass transition temperature (T_g) [52] and, consequently, the sugars are widely used in cryoprotection [53]. Both sugars confer similar levels of cryoprotection [53]. The efficiency of sugars on stabilization of macromolecules and in the modulation of their function resides in the capacity that the

sugars possess to establish hydrogen bonds, thus stabilizing the tertiary structure of the target macromolecule [50, 52].

Specifically, trehalose (non-reducing sugar) maintains the stability of proteins, sometimes by promoting the folding and refolding of the target protein [46, 52] Furthermore, it has been described that trehalose maintains the stability of the protein acting at the level of the properties of water and thereby creating a bioprotective environment for the biomolecule [46, 52, 54]. Sucrose, on the other hand, presents a particularity with respect to trehalose, because it allows the preservation of proteins that are readily volatilized during desiccation and subsequent storage at higher temperatures [55, 56].

Dithiotreitol (DTT) is a reducing agent that is added to protein solutions to inhibit the formation of disulfide bridges that are unfavorable to the target protein [57]. The disulfide bridges are responsible for maintaining the thermodynamic stability of the protein, helping in protein folding [58].

The mechanism that living organisms present to avoid denaturation and controlling osmotic pressure is through the addition of osmolytes. One of the most common osmolytes used is the glycerol. The presence of glycerol in protein solutions will increase the temperature of denaturation allowing the storage of macromolecules during longer periods of time as compared to solutions without osmolytes [59].

Moreover, cysteine is a good stabilizer of proteins, since the presence of thiol groups in cysteine allows its binding to free groups of the protein (- NH_2 and-COOH), thus increasing the protein stability by preventing aggregation and denaturation [60].

Finally, it is critical the establishment of a commitment between the stability of the target protein and the concentration of stabilizers. Otherwise, as we can see in figure 10, a dramatically decrease in the target enzyme activity could be observed if the glycerol concentration isn't the most appropriate [61]. It is believed that this effect is observed for any type of stabilizer [61].

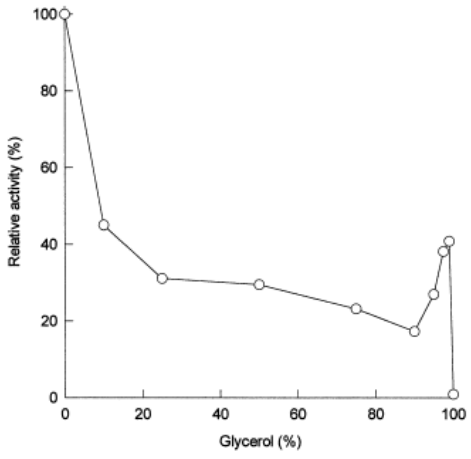

Figure 10 - Effect of glycerol concentration in protein activity [61].

Finally, the temperature and the storage time are also critical parameters for the correct storage of recombinant proteins. Usually, recombinant proteins are believed to remain in a native and stable form for 1 month at 4 ºC, 1 year in 25 – 50 % glycerol or several years frozen at – 20 or – 80 ºC [62]. However, these conditions apply to highly stable proteins where enzymes and antibodies may already significantly loose activity after 2 hours at 4 ºC [63] .

The optimization of physical and chemical parameters that influence protein storage by the classical method, i. e., changing one independent variable (trehalose, glycerol and cysteine concentrations, storage temperature and storage time) while fixing all other at a fixed level is extremely time consuming and expensive for a large number of variables. To overcome this difficulty, experiment factorial design and response surface methodology are an attractive strategy to optimize the stabilization conditions for recombinant proteins storage, minimizing the number of experiments.

1.4 The relevance of experiments in bioprocess optimization

Artificial neural networks (ANN) have been used in recent years to solve a variety of problems [64]. The ANN is a mathematical tool used in a training data set and to generalize this knowledge to predict a specific response.

The network is organized in three distinct layers: the input layer, the hidden layer(s) and the output layer [65].

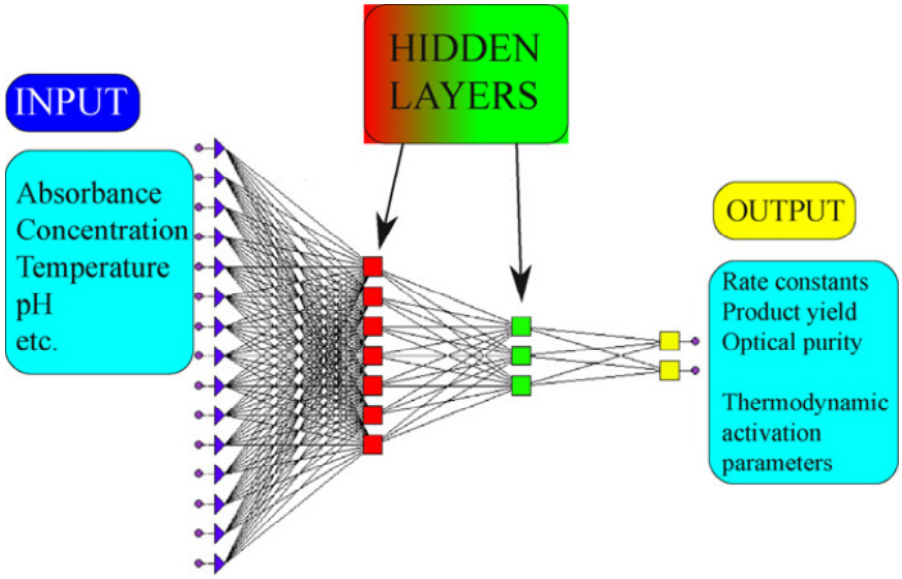

Figure 11 - Example of a typical structure for a neural network [66].

Typically, for the optimization process are considered two major steps: (A) step of screening, where after studying many factors, the most significant are chosen for the model and (B) optimization step, where the previously selected factors are studied in order to find the best conditions for method development [67].

The experimental design is a model based on statistics. It is an important tool in evaluating the interaction between independent variables [68]. The experimental design (DOE) and response surface methodology (RSM) is used in the optimization process, as can be seen in figure 12 [67].

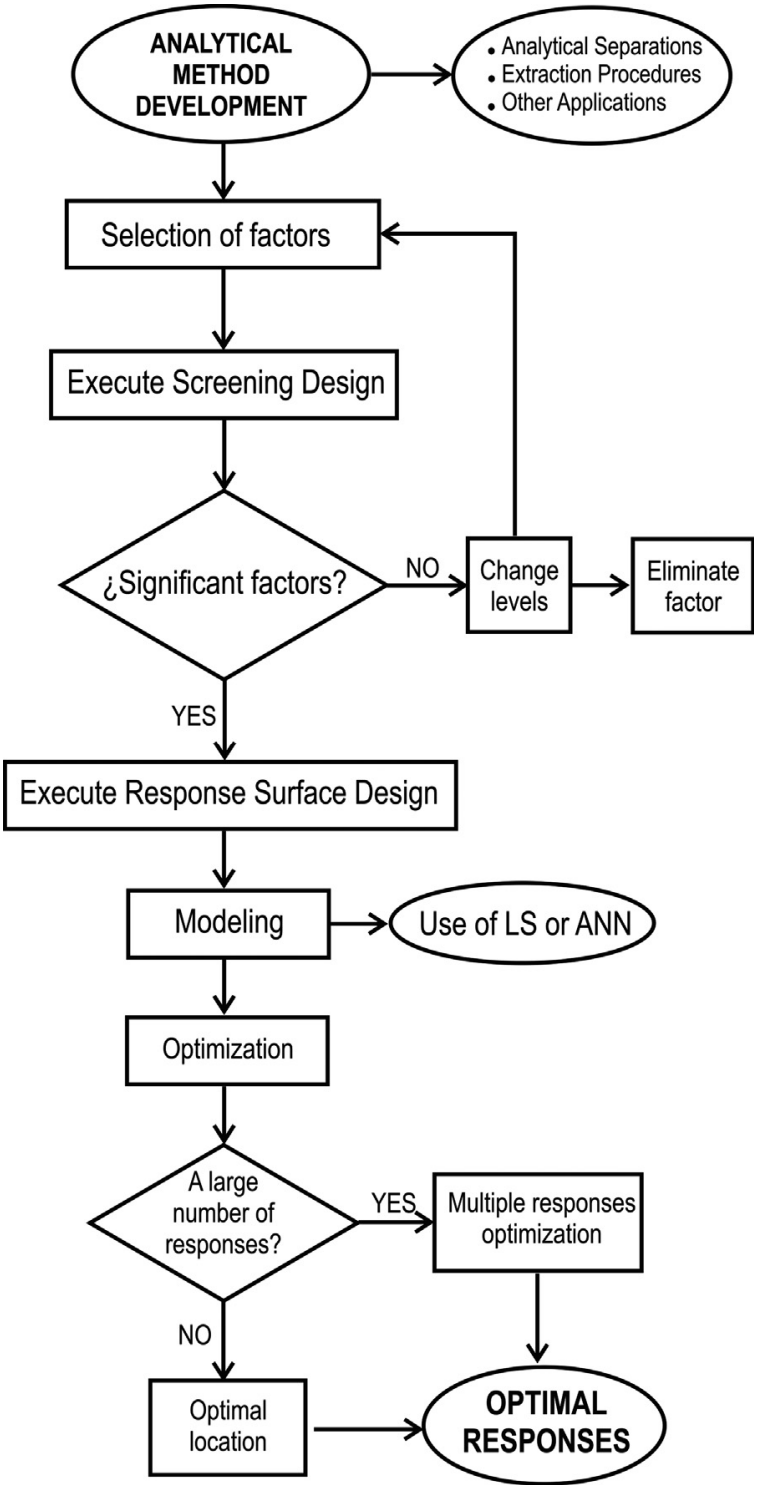

Figure 12 - Flowchart with DOE and RSM [67].

The response surface methodology is used to find the optimal conditions and is an efficient statistical technique for optimizing multiple variables with the minimum number of experiments. The response surface graphs are obtained through the interaction between the various factors under study [35]. So, RSM is a set of statistics and mathematical techniques

used to analyze several variables and determine those who are the most significant for the model [68]. For a model to be accepted, this has to be optimized. Basically, the optimization process can be summarized in three steps: statistical analysis of the variables under study, estimating the correlation coefficients (R^2) and adaptation of the model to the study [69].

Currently, there are several examples in the literature concerning the application of ANN in the initial stages of bioprocess development, either in the upstream [70] or in the downstream stage [71] but fewer examples were reported concerning the employment of experimental design in the optimization of the storage conditions of recombinant proteins [72] .

Chapter 2 - Aims

Due to the growing importance of MBCOMT for the pharmaceutical industry, particularly in the treatment of Parkinson's disease, this study aims to assess the stability of the protein so that it can maintain the native protein activity by the maximum time possible.

Therefore, in the development of this study we will:

- evaluate the most significant factors for protein activity such as: stabilizers (trealose, glycerol and cysteine) as well as temperature (-80, -20 and 4 ° C) and time (24, 48 and 72 hours). Through the interaction of the above factors we will identify / assess the conditions that can promote sustainable protein storage, without compromising the MBCOMT activity.

- develop and optimize a specific formulation containing redutor an cryogenic agents, at a specific temperature and storage time by the application of a artificial neural network.

Chapter 3 – Materials and methods

3.1 Materials

Ultrapure reagent water for HPLC was obtained with a Mili-Q system (Milipore/Waters). Zeocin was obtained from Invitrogen (Carlsbad, Ca, USA). Glass bleads, S-adenosyl-L-methionine (SAM), Ethylene glycol-bis(2-aminothylether)-N,N,N',N'-tetraacetic acid (EGTA), epinephrine, citric acid monohydrate, sodium octil sulfate (OSA), glucose, cysteine (L-), sucrose and trehalose were obtained from Sigma (St. Louis, MO). Yeast nitrogen base, yeast extract and dithiothreitol (DTT) were obtained from Himedia (Mumbai, India). Peptone was obtained from Becton, Dickinson and Company (Sparks, MD). Dipotassium phosphate (K_2HPO_4), perchloric acid and sodium acetate anhydrous (NaH_2PO_4) were obtained from Panreac (Barcelona, Spain). Monopotassium phosphate (KH_2PO_4) and magnesium chloride ($MgCl_2$) were obtained from Chem-Lab (Zedelgem, Belgium). Biotin was obtained from Roche (Basileia, Swiss). Methanol, glycerol and acetonitrile were obtained from VWR (Carnaxide, Portugal). Pierce BCA Protein Assay Kit and protease inhibitors (leupeptin, pesptatin and PMSF) were obtained from Thermo Scientific (Rockford, USA). Agar was obteined from Pronadisa (Basel, Switzerland).

3.2 Construction of the expression vector pPICZα-hMBCOMT

In this study, the hMBCOMT gene was PCR amplified from the plasmid pNCMO2-hMBCOMT [7] using specific primers containing a unique restriction site for Xho I (forward primer; 5' AACTCGAGAAAAGAATGCCGGAGGCCCCGCCT3'; reverse primer, 5' AACTCGAGTCAGGGCCCTGCTTCGCTGCCTG 3') and cloned into the plasmid pPICZα, previously digested with the same restriction enzyme. After confirming that the isolated positive clone contained the hMBCOMT gene [73], it was introduced into freshly made competent cells by electroporation. Then, in order to confirm the correct integration of the hMBCOMT gene into the *P. pastoris* host cell genome, both transformed and untransformed yeast DNA was extracted and analyzed by PCR using the AOX primers (AOX1 5' GACTGGTTCCAATTGACAAGC 3' and AOX1 5' CAAATGGCATTCTGACATCC) and one band corresponding to the size of our gene of interest cloned into pPICZα was obtained. Finally, the Mut phenotype of the *Pichia pastoris* X33 host strain transformed with the hMBCOMT cDNA was determined according to the manufacturer's instructions and it was confirmed to be Mut^+.

3.3 Recombinant hMBCOMT biosynthesis and recuperation

P. pastoris cells transformed with the expression construct were grown for 72 hours at 30ºC in YPD medium plates containing 200 µg/mL Zeocin. A single colony was precultived in 100 mL of BMGY in 500 mL shake-flasks. Cells were grown at 30ºC and 250 rpm to a cell density at 600 nm (OD_{600}) between 5-6. Subsequently, an aliquot was added to 100 mL of BMMY medium in 500 mL shake-flasks and the initial OD_{600} was fixed to 1.0 unit. After a 24 hour growth at 30ºC and 250 rpm, cells were harvested by centrifugation (1500xg, 10 min, 4ºC).

The cell pellet was ressuspended in 2 mL of an appropriate Buffer (150 mM NaCl, 50 mM Tris, 1 mM $MgCl_2$, pH 7.8), suplemented with protease inihibitors (5 µg/mL leupeptin, 0.7 µg/mL pesptatin and 10 µL/mL PMSF), disrupted by mechanical treatment with glass beads (seven cycles of 1 minute with 1 minute of interval on ice). The resultant supernatant was recovered after centrifugation (500xg, 5min, 4ºC).

3.4 Quantification of total protein using BCA micro-assay

Protein contents in samples was measured by the Pierce BCA Protein Assay Kit using Bovine serum albumin as standards (0.025-2.5 mg/mL), according to manufacturer´s instructions and the obtained concentration in lysates were adjusted to 2 mL.

3.5 hMBCOMT stability assays

The MBCOMT stability was assessed by evaluating the effect of several experimental conditions in the target enzyme activity levels. A total of 23 experiments were carried out combining several stabilizers with different concentrations, different temperatures and different storage periods. The stabilizers evaluated were trehalose, glycerol and cysteine with the respective concentrations ranging from 100-300 mM, 10-30 %v/v and 100-150 mM. The storage conditions were assessed by varying the time (24, 48 and 72h) and temperature (4, -20 and -80ºC) of the assays.

The hMBCOMT stability experiments consisted in addition of 500 µL stabilizing solution to 500 µL of protein lysates (with the total protein being normalized to 2.0 mg/mL) in order

to obtain the desired stabilizers concentration and stored according to the experimental design.

After experimental assays, hMBCOMT activity levels were evaluated by the measurement of metanephrine levels converted from epinephrine as previously described [11]. Briefly, samples were incubated in 5 mM sodium phosphate buffer (pH 7.8) containing 0.2 µm $MgCl_2$, 2 mM EGTA, 250 µm SAMe and 1 mM epinephrine in a total sample of 1 ml at 37ºC for 15 min. The reaction is stopped by the addition of 200 µL of 2 M perchloric acid.

3.6 Analytical methods: hMBCOMT enzymatic assays

The methylating efficiency of recombinant MBCOMT was evaluated by measuring the amount of metanephrine formed from epinephrine as previously described [11]. The samples were injected into an HPLC (High-Performance Liquid Chromatography) Agilent 1260 system (Agilent, Santa Clara, USA) equipped with autosampler and quaternary pump coupled to an ESA Coulochem III detector (Milford, MA, USA).

The samples are centrifuged at a speed of 6000 rpm for 10 min at a temperature of 4 °C and then are filtered with filters of pore size 0.22 µm, to remove some material that is precipitated.

After HPLC system with electrochemical detector being balanced with the following mobile phase: 0.12 M NaH_2PO_4, 0.03 M citric acid monohydrate, 2 mM sodium octyl sulfate and 9% acetonitrile, pH 3.0, the samples are injected on Zorbax 300SB C18 RP analytical column (250× 4.6 mm id 5 µm; Agilent, Santa Clara, USA) with a column temperature set point of 30 ºC and flow rate of 1mL/min.

The data are monitored by electrochemical detection using two analytical cells using an oxidation and reduction potencial (analytical cell #1: +410 mV; analytical cell #2: -350 mV. The sensitivity of the method used previously is 1 µA. To calculate the activity of the protein is used in the chromatograms obtained reduction potential (-350 mV) [74].

The activity of hMBCOMT is calculated from the peak area of metanephrine by integration of the characteristic peak metanephrine using a calibration curve. The units of total activity comes metanephrine produced in nmol/h while the units of specific activity comes metanephrine produced in nmol/h/mg protein. The amount of protein produced is previously determined by using the micro-assay BCA.

3.7 Artificial neural network

A feed-forward artificial neural network was applied to predict the activity recovery as function of the concentration of trehalose (100-300 mM), glycerol (10-30 % v/v) and cysteine (100-150 mM) and temperature of -80, -20 and 4ºC and time of 24, 48 and 72h. The ANN models were implemented in MATLAB™ using the Neural Network Toolbox. The ANN structure included an input layer with five neurons (one for each input variables), an output layer with one neuron (activity recovery) and one hidden layer with three neurons (5/3/1). Therefore, the resulting model contains a total of 23 parameters. The transfer functions of the input and output layers were linear functions 'purelin' and of the hidden layer were log-sigmoid functions 'logsig'. The output function can be represented mathematically as follows:

$$f(x,w) = w_{output} \cdot s(w_{hidden} \cdot x + bias_{hidden}) + bias_{output}$$

with, w_{hidden}, w_{output}, b_{hidden} and b_{output} the parameter matrices associated with connections between the nodes of the network, w is a vectored form of the previous parameters and s is the activation function:

$$s(t) = \frac{1}{1+e^{-(t)}}$$

with t representing:

$$t = w_{hidden} \cdot x + bias_{hidden}$$

The ANN structure was built using the "newff" function. ANN was trained with the Levenberg-Marquardt back-propagation function, up to 1000 epochs, using the "train" function. The learning rate and the momentum constant were set at 0.01 and 0.9. The remaining training parameters were set at the default values defined in MATLAB™. The input and output data were scaled according to the 'mapminmax' function. The well-known early stopping method was used, i.e. the mean squared error between measured and predicted number of column volumes of the validation dataset (25 % of overall dataset) was monitored during training. The training was stopped when the validation increases over a certain threshold. The ANN was trained until either the maximum number of epochs or the performance goal was reached.

Chapter 4 – Results and Discussion

P. pastoris has been sucessfully used in the production of various proteins [41]. The expression system based on *P. pastoris* enables the biosynthesis of proteins at high levels in media of easy preparation that allow the achievement of high cell densities [33, 58].

Specifically, while COMT is an important enzyme responsible for the inactivation of catechol substrates [75, 76], its stability as for other proteins is affected by physical-chemical parameters such as pH and temperature [77]. According to some authors, COMT is an extremely unstable enzyme that loses 50-70% of its activity in less than 24 hours at 4 °C [78]. However, proteins are thermodynamically stable in a certain temperature range

In this work, the stabilizers are added to the hMBCOMT in order to increase the thermal stability and help maintain its native conformation [49]. hMBCOMT stability was assessed by measuring the levels of enzyme activity in a HPLC system coupled to electrochemical detection. Stability was studied as a way to evaluate the action of stabilizers, verifying that these compounds decreased, kept or increased hMBCOMT biological activity. Therefore, after MBCOMT recombinant biosynthesis and recuperation from *Pichia pastoris* cells, we intended to improve the storage conditions of MBCOMT lysates in order to maintain its biological activity during longer periods that those that have been reported. Finally, in order to fulfill the global aim of this work, an artificial neural network was generated to give a specific formulation consisting of the best concentration of stabilizers as well as the best temperature and the period time where MBCOMT retains more biological activity.

4.1 Recombinant MBCOMT biosynthesis and recuperation from *Pichia pastoris* X33 cells

MBCOMT biosynthesis was developed using *Pichia pastoris* X33 cells under the control of the strongly methanol inducible AOX promoter in shake-flasks containing a semi-defined medium (BMGY) with 0.5 % (v/v) methanol. Subsequently, a sequential procedure was established as the main method for *P. pastoris* lysis: vortexing of the cells (combined with lysis buffer and glass beads) for 7 times during 1 minute with an interval of 1 minute on ice. Then, after a centrifugation at 500g, MBCOMT in an immunological and biological active form was isolated whose kinetic parameters are identical to its correspondent native enzyme (data not shown).

4.2 Preliminary results obtained from SCOMT stabilization assays

Previous stabilization studies carried out for SCOMT lysates obtained from *Escherichia coli* showed that a combination of stabilizers (DTT and glycerol) are effective in maintaining the activity of SCOMT [17]. In fact, preliminary results obtained for SCOMT were helpful in order to define which stabilizer at which concentration should be used in the development of a specific formulation that can maintain and/or improve MBCOMT stability in *P. pastoris* lysates. Therefore, the SCOMT stability was assessed after incubation at 4 ºC with the specified stabilizer concentrations during 12 hours and the activity recovery was determined by the ratio between the SCOMT specific activity at 12 hours without stabilizers and the SCOMT specific activity at 12 hours with the stabilizers at the reported concentration.

The first stabilizer to be evaluated was sucrose, whose results are depicted in figure 13. In fact, higher SCOMT specific activity was observed for sucrose concentrations within the range 250-350 mM as compared to the activity of the protein without the stabilizer addition. The same tendency was observed for the activity recovery.

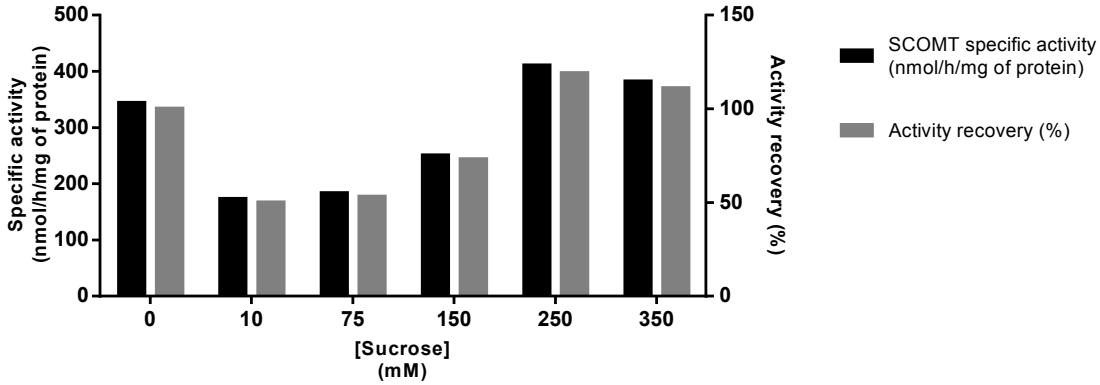

Figure 13 - Specific activity and percentage recovery of SCOMT activity in the presence and absence of sucrose during a 12 hours period (Activity recovery (%) is defined as the ratio between SCOMT specific activity at 12 hours without stabilizer and SCOMT specific activity at 12 hours with the specified concentration of stabilizer; each value represents the mean of 3 independent samples).

Another stabilizer in study is DTT, which prevents the formation of oligomers, thus helps stabilizing the protein during storage as it rearranges the protein disulfide groups during folding, thereby facilitating the conformation of the native protein [77–79]. However, the use of DTT can be somewhat controversial. In fact, while there are some authors who argue that the use of DTT in a concentration up to 4 mM prevent protein aggregation [81, 82], others argue that DTT can bind to the active site of the protein, affecting protein activity [57].

Furthermore, cysteine is extremely important in maintaining enzyme activity since it is involved in the formation of disulfide bridges which are essential for maintaining the native conformation of the protein [83]. The relationship between the DTT and cysteine concentrations with the SCOMT specific activity and the percentage of recovery is shown in figure 14.

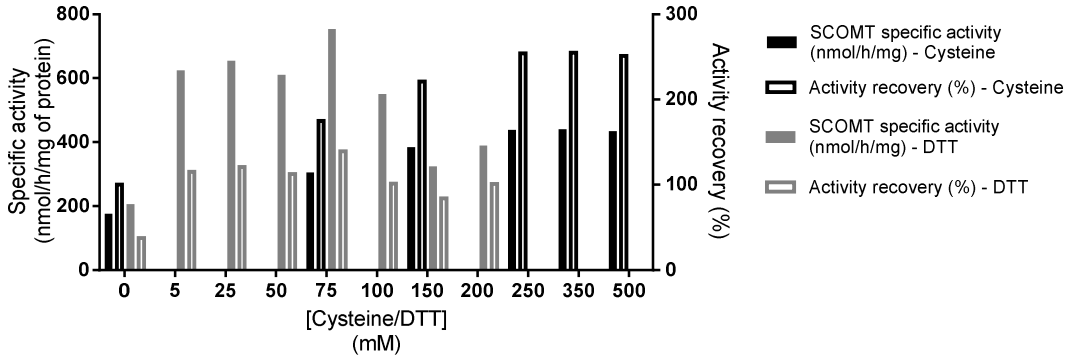

Figure 14 - Specific activity and percentage recovery of COMT activity in the presence and absence of DTT and cysteine during a 12 hours period (Activity recovery (%) is defined as the ratio between SCOMT specific activity at 12 hours without stabilizer and SCOMT specific activity at 12 hours with the specified concentration of stabilizer; each value represents the mean of 3 independent samples).

The analysis of the results in figure 14 allowed us to observe that DTT is more effective in maintaining SCOMT activity at concentrations ranging between 50 and 100 mM while for cysteine, higher values were detected between 200 and 350 mM.

Glycerol belongs to the group of polyols and is used as a cryoprotectant because it increases the melting temperature, thus protecting the macromolecule against thermal inactivation and helps keeping the properties of the native protein [83, 84]. The effect of the glycerol in the SCOMT specific activity was also analyzed (see figure 15) and it was found that the glycerol concentration where the highest percentage of SCOMT activity recovery was obtained at 20 % (v/v).

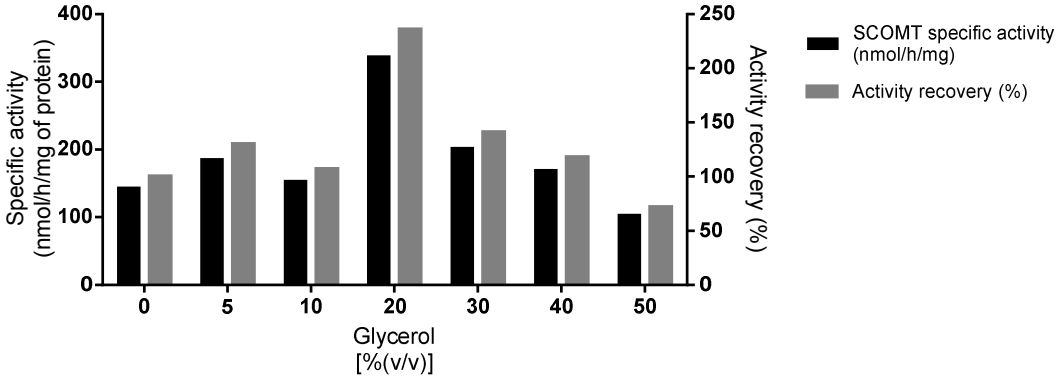

Figure 15 - Specific activity and percentage recovery of COMT activity in the presence and absence of glycerol during a 12 hours period (Activity recovery (%) is defined as the ratio between SCOMT specific activity at 12 hours without stabilizer and SCOMT specific activity at 12 hours with the specified concentration of stabilizer; each value represents the mean of 3 independent samples).In addition to sucrose, trehalose is also commonly studied as a stabilizer in pharmaceutical formulations. In fact, it has been demonstrated that trehalose 1 M increases the stability of lysozyme at 8 °C whereas sucrose enhances the RNase stability only at 4 ºC [85, 86]. Again, as demonstrated in figure 16, the effect of adding trehalose at several concentrations on the SCOMT specific activity was determined and it was found that higher percentage of SCOMT activity recoveries were found with a trehalose concentration between 75 and 250 mM.

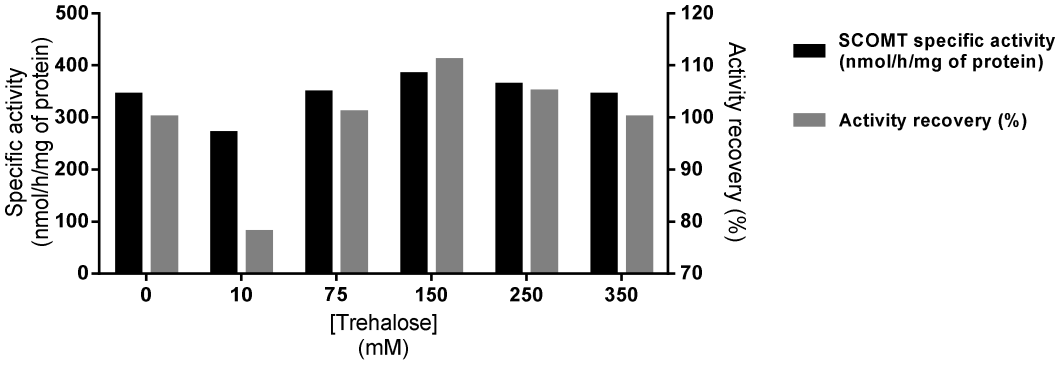

Figure 16 - Specific activity and percentage recovery of COMT activity in the presence and absence of trehalose during a 12 hours period (Activity recovery (%) is defined as the ratio between SCOMT specific activity at 12 hours without stabilizer and SCOMT specific activity at 12 hours with the specified concentration of stabilizer; each value represents the mean of 3 independent samples).

4.3 Choice of stabilizers

Following the preliminary results obtained for the SCOMT specific activity and percentage of recovery of SCOMT activity during 12 hours, the ranges of stabilizers concentrations that would be applied for MBCOMT were specified. Therefore, and based on the results obtained for SCOMT activity (please see figures 13 – 17), the following concentrations ranges were defined: 150 – 300 mM sucrose, 10 – 30 % (v/v) glycerol, 100 – 150 mM cysteine and 50 – 100 mM DTT.

The next step was to evaluate if there is any interference of the retention time of these compounds in the HPLC-reversed phase chromatographic runs designed to evaluate the formation of metanephrine and, consequently, that allows the determination of the MBCOMT specific activity. Typically the retention time of metanephrine (MN) was observed at around 8 minutes (figure 17). Therefore, a 300 mM sucrose solution was directly injected into the HPLC system in the same conditions as the samples obtained through the MBCOMT biological activity assays would be performed. The results are depicted in figure 17.

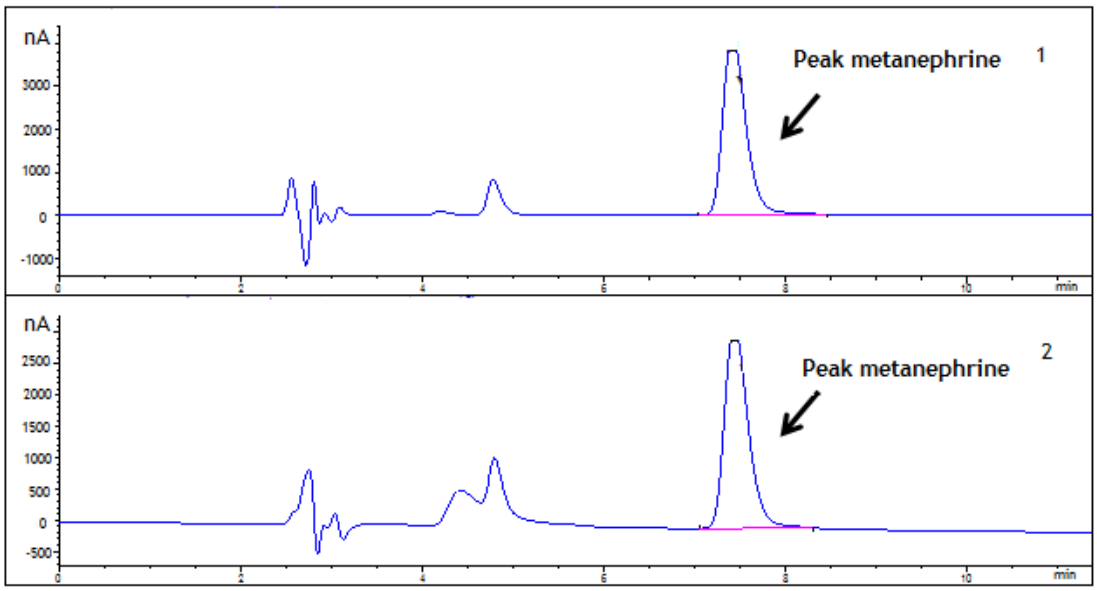

Figure 17 – HPLC chromatogram of a standard metanephrine with 15 nmol/mL. (1) Oxidation potential and (2) Reduction potential.

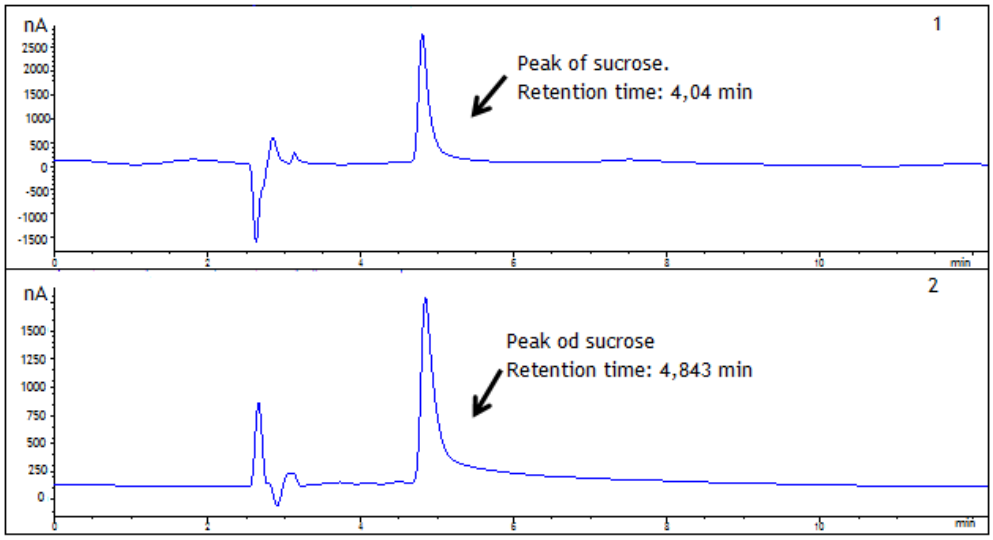

Figure 18 - HPLC chromatogram of a 350 mM sucrose solution. (1) Oxidation potential and (2) Reduction potential.

The analysis of the chromatogram in figure 18 allowed us to conclude that there is no interference of the sucrose solution retention time (retention time 4.8 min), compared with that from metanephrine (retention time around 8 min). Similarly, no interferences were detected when a cysteine solution (see figure 19) or a glycerol solution (see figure 20) were directly injected onto the HPLC system.

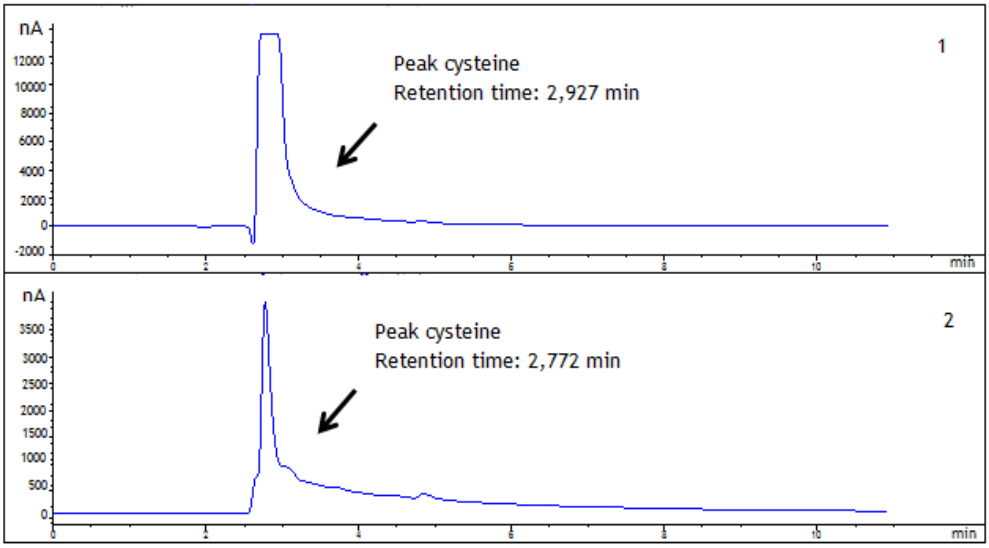

Figure 19 - HPLC chromatogram of a 300 mM cysteine solution. (1) Oxidation potential and (2) Reduction potential.

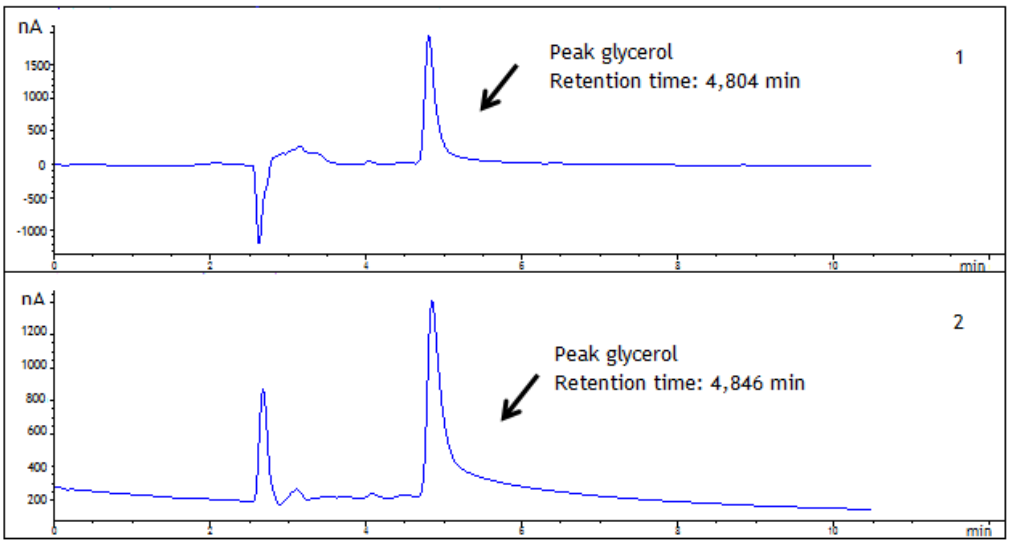

Figure 20 – HPLC chromatogram of a 30% (v/v) glycerol solution. (1) Oxidation potential and (2) Reduction potential.

The last compound to be analyzed in a chromatographic run was the DTT (see figure 21). In fact, after the injection of a 100 mM DTT solution in the HPLC system, it was found that the retention time of this solution interfere with the retention time of metanephrine. Therefore, as the peak obtained from the 100 mM DTT solution mask the peak obtained from the metanephrine, we decided to remove this stabilizer from the compounds present in the formulation since it wouldn't allow the quantification of metanephrine from the MBCOMT biological activity assays.

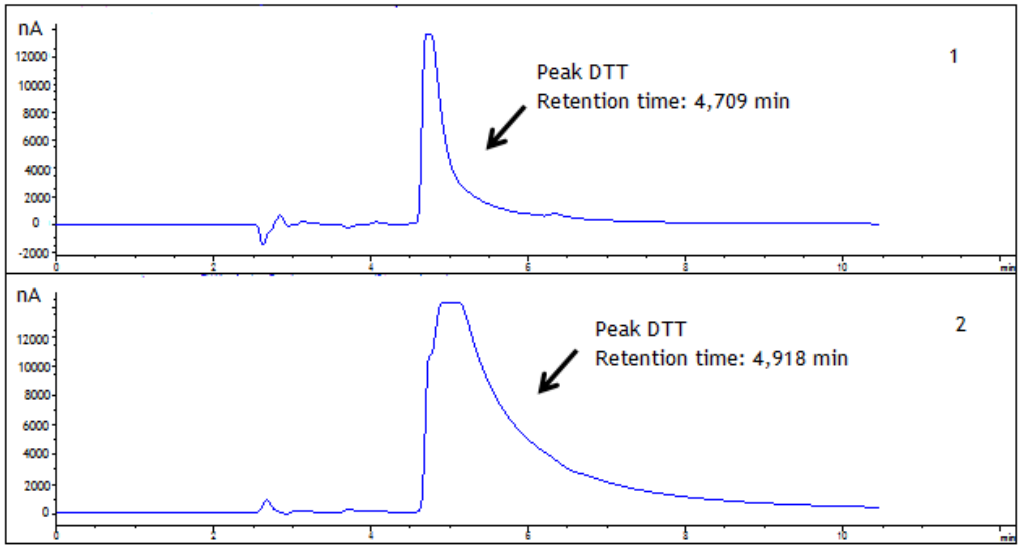

Figure 21 – HPLC chromatogram of a 100 mM DTT solution. (1) Oxidation potential and (2) Reduction potential.

However, apart from the fact that a 100 mM DTT solution interferes with the assessment of metanephrine, it is important the presence of DTT in the lysis buffer, even that it is present in lower concentrations. As a matter of fact, the complete removal of DTT from this buffer has an extremely negative effect on the initial MBCOMT biological activity, decreasing it to very low values. Therefore, since when MBCOMT is in a partially denatured state, matching with initial very low levels of biological activity, it is extremely difficult to promote MBCOMT renaturation into a highly biological active state. So, we chose to keep the DTT in the lysis buffer at a smaller concentration (10 mM) in order to avoid MBCOMT specific activity losses during the initial recuperation step immediately after the *P. pastoris* lysis. In fact, the DTT in these concentrations doesn't interfere with the metanephrine retention time, as demonstrated in figure 22.

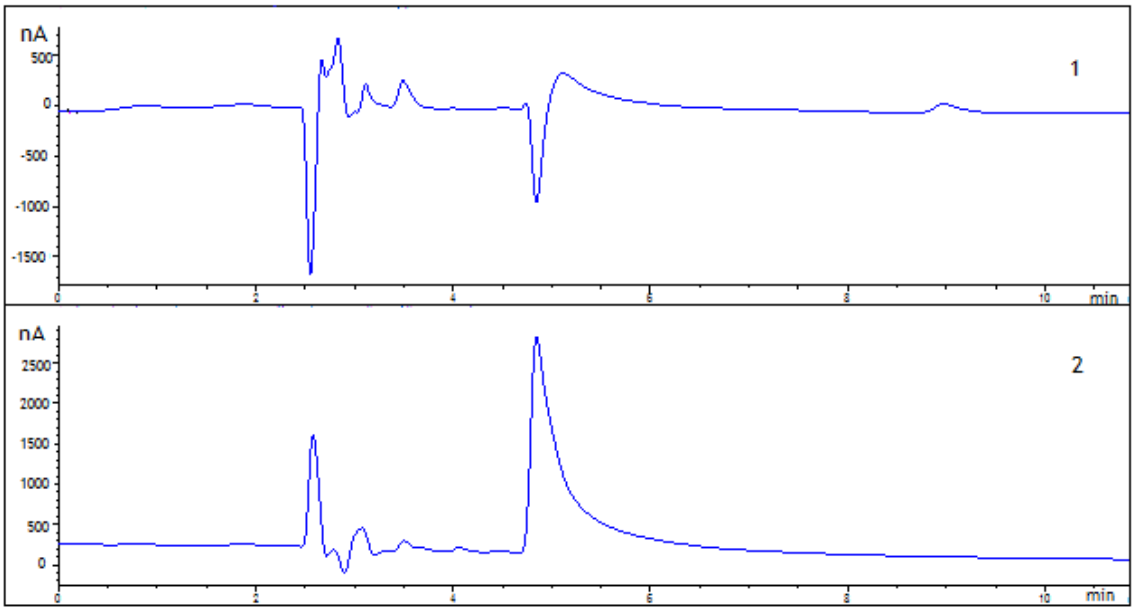

Figure 22 – HPLC chromatogram of lysis buffer (containing 10 mM DTT). (1) Oxidation potencial and (2) Reduction potencial.

.

Finally, we wanted to evaluate in a HPLC chromatographic run if the stabilizer solution composed of all stabilizers in study (previously analized separately) presented some interference concerning the retention time of metanephrine. As demonstrated in figure 23, the peak of the stabilizing solution partially overlaps the peak corresponding to metanephrine, thus compromising again the determinations of MBCOMT biological activity. Therefore, additional trials proved that this effect was due to the sucrose, when in solution with the other stabilizers.

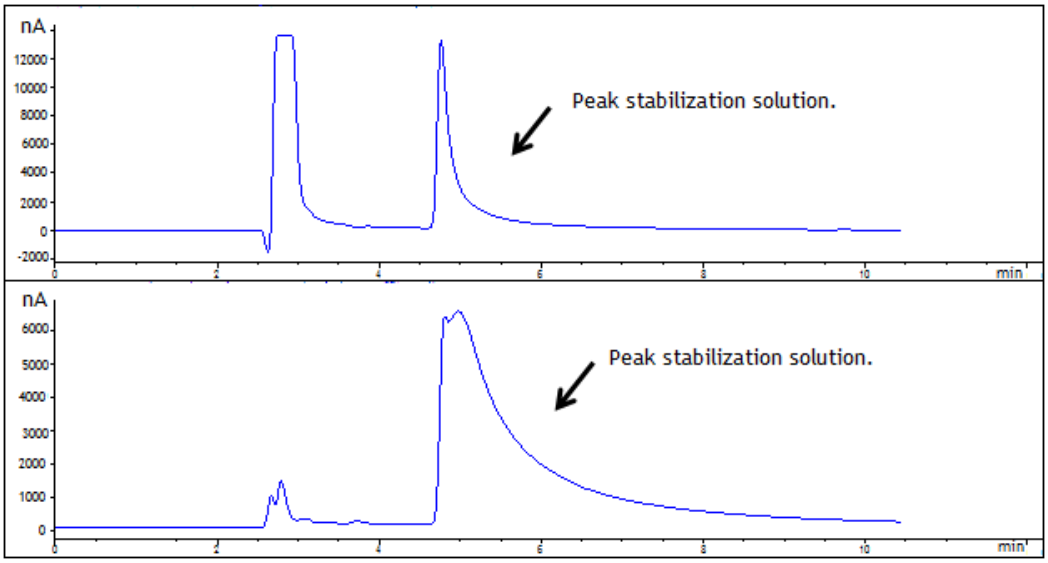

Figure 23 – HPLC chromatogram of the stabilizer solution containing sucrose, cysteine, DTT and glycerol. (1) Oxidation potential and (2) Reduction potential.

In the preliminary stabilization studies carried out for SCOMT, another sugar in study with similar functions than sucrose was trehalose. Therefore, as we did for the other stabilizers, a solution of 300 mM trehalose was injected in the HPLC (see figure 24) and none interferences with the metanephrine retention time were observed. Thus, the best option was to replaced sucrose to trehalose, within the range 100-300 mM, in the stabilizers formulations in order to perform the experimental design for achieve the optimal storage conditions.

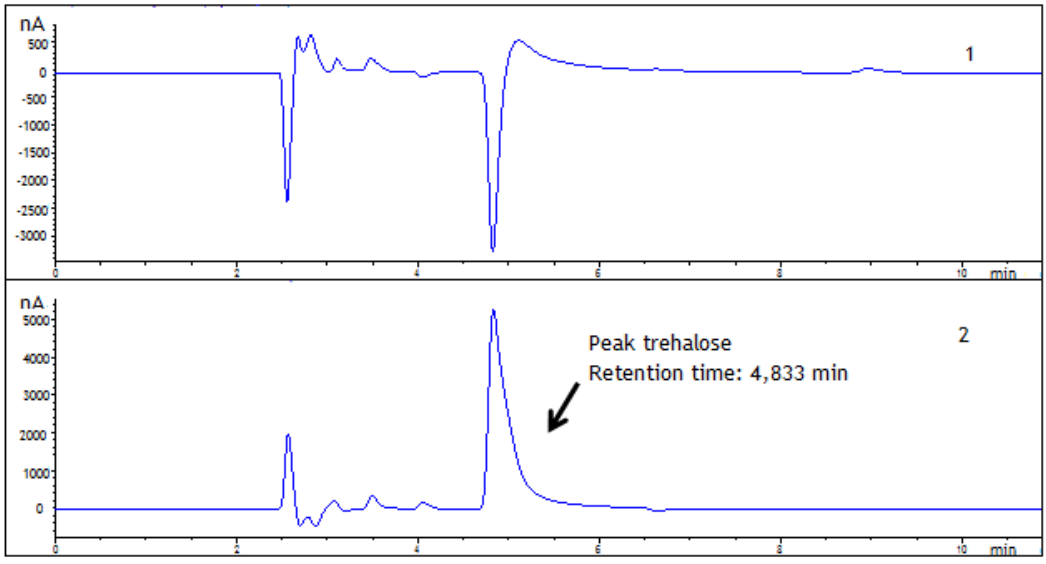

Figure 24 – HPLC chromatogram of a 300 mM trehalose solution. (1) Oxidation potential and (2) Reduction potential.

4.4 Experimental Design

The initial trials proved to be useful in determining either which range of stabilizers concentrations could be useful to maintain MBCOMT biological activity as to determine if they exert any interference in the chromatographic determination of metanephrine. After carrying out these assays, a computer program generated the experimental design based on the selected inputs. Indeed, three levels were defined for the stabilizers concentrations (specific for each compound) as well as for the storage time (24, 48 and 72 hours) and the storage temperature (4 ºC, - 20 ºC and – 80 ºC), as shown in Table 6. Nevertheless, although higher storage periods such as 1 month or 1 year had been reported for several proteins, these proteins are thought to be stable and don't lose their native conformation easily. Therefore, as MBCOMT is described to be a highly unstable and thermolabile enzyme, shorter periods of storage were considered in this work at the temperatures that are commonly reported for the storage of mostly proteins (4, - 20 and – 80 ºC).

Table 6 - Structure of the DOE.

Exp No	Exp Name	Run Order	Trehalose	Gycerol	Cysteine	Temperature	Time
1	N1	17	100	10	100	-80	24
2	N2	5	300	30	100	-80	24
3	N3	14	300	10	150	-80	24
4	N4	21	100	30	150	-80	24
5	N5	15	200	10	100	-20	24
6	N6	20	300	10	100	4	24
7	N7	4	100	30	100	4	24
8	N8	22	100	10	150	4	24
9	N9	23	300	30	150	4	24
10	N10	18	200	20	125	-80	48
11	N11	7	300	20	150	-20	48
12	N12	19	300	10	100	-80	72
13	N13	1	100	30	100	-80	72
14	N14	10	100	10	100	-80	72
15	N15	12	300	30	150	-80	72
16	N16	3	100	30	125	-20	72
17	N17	8	100	10	100	4	72
18	N18	6	300	30	100	4	72
19	N19	16	300	10	150	4	72
20	N20	9	100	30	150	4	72
21	N21	2	300	30	150	4	72
22	N22	11	300	30	150	4	72
23	N23	13	300	30	150	4	72

4.5 Artificial Neural Network

In order to achieve a specific formulation for maintaining MBCOMT stability, the concentration of each stabilizer was analyzed at three levels as well as the storage temperature (- 80 ºC, - 20 ºC and 4 ºC) and the storage time (24, 48 and 72 hours) by an artificial neural network. After completing the 23 experiments in which the inputs previously selected were varied, i.e., the concentration of stabilizers, the storage temperature and the storage time, an output was obtained: the percentage of recovery of MBCOMT specific activity. The percentage of MBCOMT biological activity recovery is calculated by dividing the specific activity after incubation in the specific conditions by the MBCOMT specific activity at 0 hours without stabilizers. In table 7 are depicted the results obtained for all the 23 experiments where the most relevant results are highlighted in bolt.

Table 7 - Results obtained after performing the DOE.

Exp No	Specific activity (nmol/h/mg) (0 hours)	Specific activity (nmol/h/mg) real(after incubation)	% MBCOMT bioactivity recovery
1	145,45	112,87	77,60
2	16,51	10,39	62,94
3	268,53	112,25	41,80
4	243,68	3,96	1,63
5	182,75	46,88	25,65
6	283,45	72,59	25,61
7	229,58	57,90	25,22
8	291,80	62,35	21,37
9	209,00	57,24	27,39
10	170,32	109,75	64,44
11	334,36	274,52	82,10
12	368,72	124,69	34,47
13	403,24	326,32	80,92
14	170,77	82,84	48,51
15	**97,73**	**107,39**	**109,89**
16	193,94	172,65	89,02
17	**175,28**	**188,72**	**107,67**
18	240,45	167,23	69,43
19	214,04	59,22	27,67
20	**69,61**	**79,74**	**114,56**
21	73,94	48,66	65,81
22	227,68	29,28	12,86
23	104,07	78,85	75,77

A generic analysis of the MBCOMT specific activity at 0 hours where higher activity levels were detected in almost all experiments proves the usefulness of using DTT 10 mM in the lysis buffer since it removal leads to a marked decrease in these values. In fact, DTT behaves as a reducing agent that through the inhibition of the formation of disulfide bridges that are unfavorable to the target protein might stabilizes MBCOMT.

Through the analysis of table 7, it is possible observe that the experiments 15 (300 mM trehalose, 30 % (v/v) glycerol, 150 mM cysteine at – 80 ºC during 72 hours), 17 (100 mM trehalose, 10 % (v/v) glycerol, 100 mM cysteine at + 4 ºC during 72 hours) and 20 (100 mM trehalose, 30 % (v/v) glycerol, 150 mM cysteine at + 4 ºC during 72 hours) gave a percentage of recovery of MBCOMT biological activity above 100% These data are extremely important, because through the optimal stabilizers concentrations, MBCOMT biological activity is maintained and, in some cases, increased for extended periods of storage.

In particular, from the experiment 15 where 109.9 % of MBCOMT activity recovery was obtained, it seems that all the three stabilizers in study are the major components that are responsible for maintaining MBCOMT biological activity at – 80 ºC. On the other hand, in the experiments 17 and 20 that were carried out at 4 ºC and 107.7 and 114.6 % of MBCOMT activity recovery was obtained, respectively, the requirements for the stabilizers concentrations doesn't seem to have a great impact. A typical chromatogram (run 20) from the determination of metanephrine in these biological extracts is shown in figure 25.

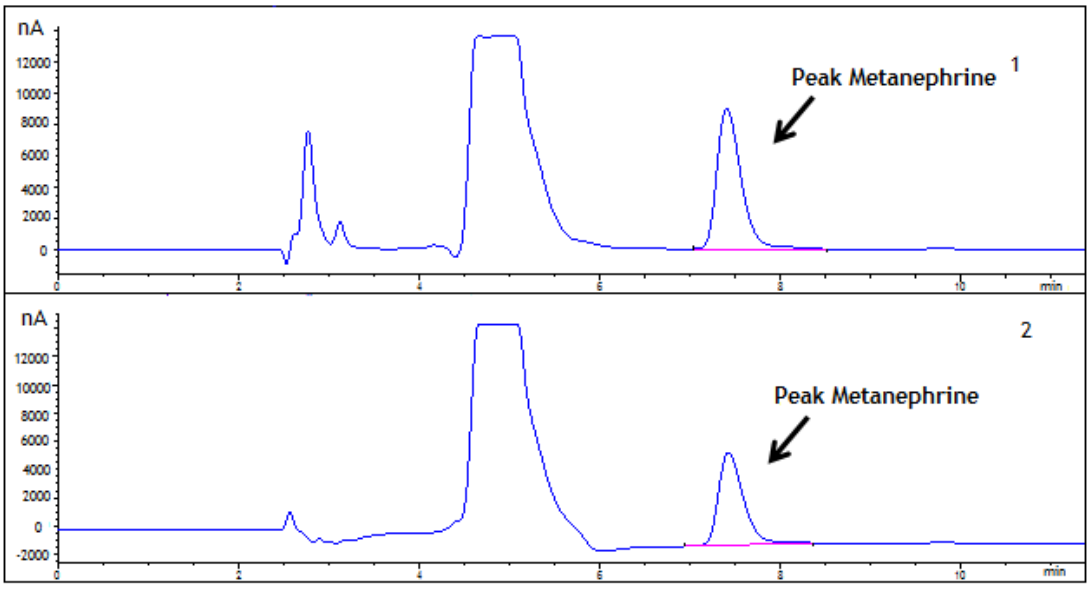

Figure 25 - HPLC chromatogram of MBCOMT extract sample. (1) Oxidation potential and (2) Reduction potential

Interestingly, trehalose is widely used in cryoprotection since it increases the glass transition temperature and, in this particular case, seems to be required in higher concentrations for MBCOMT storage at – 80 ºC. In fact, while at – 80 ºC, 300 mM trehalose

are needed to maintain MBCOMT biological activity, at 4 ºC, the best results were obtained with the lower trehalose concentration in study. Moreover, from the analysis of the 23 experiments, it is possible to observe that in experiments in which the storage time is shorter (24 hours, -80 ° C), the percent recovery of bioactivity is lower when compared for example with longer storage times (48 to 72 hours) for the same temperature.

Succinctly, from the analysis of the results in table 7, the optimum point was obtained at 4 ºC with 100 mM trehalose, 30 % (v/v) glycerol and 150 mM cysteine with an incubation period of 72 hours where 114.56 % of MBCOMT activity recovery was obtained.

Finally, an ANN preliminary model aiming the optimization of experimental conditions of activity recovery was developed by the structure described in the section of material and methods. The ANN model was trained using the experiments defined in Table 7.

The ANN model was developed using a stepwise process until the maximum activity recovery was achieved by applying the optimal conditions. In the present work, only one iteration was run. The model estimation for maximum activity recovery was 219.24 under the optimal conditions (160 mM trehalose, 30% (v/v) glycerol and 150 mM cysteine at -64.5ºC and over 72 hours) in the first iteration. This value is 92 % higher than those achieved under the best condition performed in the DoE step (see experiment 20 in Table 7). The obtained ANN model is almost unbiased because the slope and R^2 of the fitting between the measured and predicted output were close to 1 (1.0127 and 0.9952, respectively). Figure 26 shows the contour plot obtained from the ANN model for all combinations between the five operational conditions. The modeling results showed that the activity recovery is sensitive to the operational conditions. However, the activity recovery seems to be not extensively affected by cysteine concentration when it varied simultaneously with trehalose, glycerol and temperature under the optimal conditions of the remaining operational conditions. As a future work, new optimization iterations is required for model validation and/or fine tuning. In subsequent iterations, results interpretation regarding to the sensitivity of the operations conditions may vary.

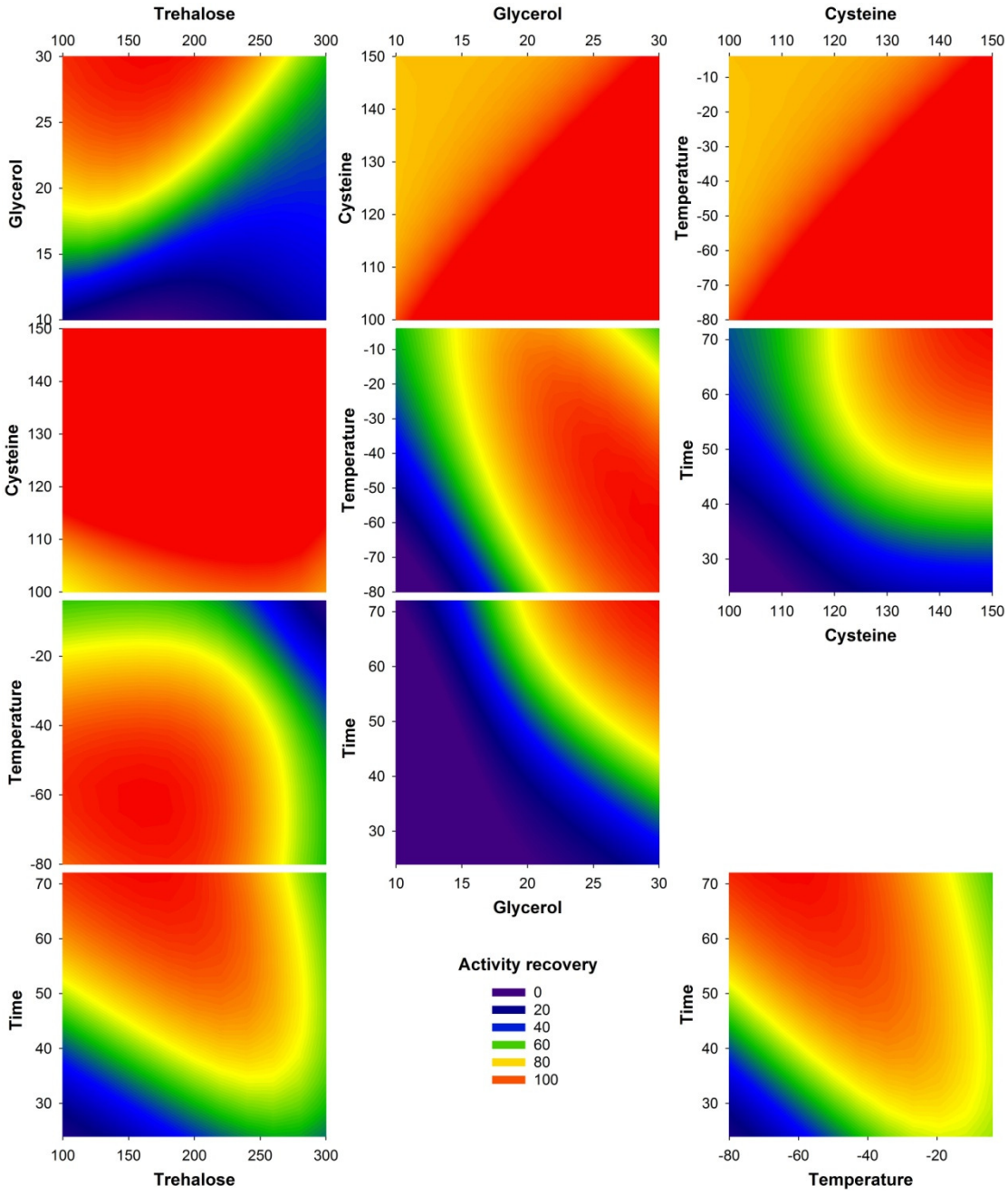

Figure 26 - Contour plots of MBCOMT specific activity (nmol/h/mg of protein), as a function of the stabilizers concentration (trehalose, glycerol and cysteine), storage temperature and storage time

Chapter 5 - Conclusions and Future Perspectives

Membrane proteins are drug targets for a wide range of diseases and the achievement of appropriate samples for further research encourages the pharmaceutical industry for developing new and more effective drugs. These samples are often obtained toward a strategy concerning the synthesis of the target protein on a large scale, after which an equal appropriated purification flowsheet has to be designed in order to isolate pure protein in higher quantities than those found in the protein's native source. Throughout this strategy, one major issue is to maintain the target recombinant protein in a native and biologically active state.

Specifically, COMT has been regarded as an important therapeutic target for PD. In fact, the improvement of the clinical efficacy of the currently available COMT inhibitors largely depends on structural and functional studies performed with recombinant forms of this enzyme. However, a major bottleneck in the expression of membrane proteins and, in particular, in MBCOMT expression from *P. pastoris* is the stability of the target protein. As MBCOMT is extremely thermolabile and losses it biological activity in few hours after its recuperation, it is critical the establishment of a strategy that allows MBCOMT to maintain its biological activity over time allowing performing the aforementioned studies. Hence, in this work and for the first time, it was attempted to create a specific formulation comprising several stabilizers (trehalose, glycerol and cysteine) that maintains MBCOMT in a biologically active state for more than 24 hours. Thus, after defining the stabilizers concentrations ranges that were more suitable in stabilizing MBCOMT (assessed by the activity recovery at the end of the incubation time) and were compatible with the HPLC chromatographic runs designed to evaluate the formation of metanephrine, a product that is correlated with the MBCOMT methylating efficiency, the 23 experiments belonging to the DOE were performed. Among these experiments, 3 conditions gave excellent results and allowed percentages of MBCOMT activity recovery higher than 100%. In fact, while two of these experiments were performed at + 4 ºC with distinct requirements in terms of the stabilizers concentrations, the one that was performed at -80 ºC had the highest stabilizers concentrations under study. Therefore, it seems that for lower storage temperature, there is an increased demand on the stabilizers concentrations required. In particular, for – 80 ºC, the use of higher glycerol concentrations seem to be important since it exerts an important effect in controlling the osmotic pressure, increasing the temperature of denaturation.

Overall, a new approach intended to evaluate MBCOMT lysates stability at short-term storage based on DOE approaches was successfully implemented. From this strategy, 3 specific formulations that gave similar results allow the storage of MBCOMT lysates with a slightly increase on MBCOMT biological activity after 72 hours, the highest period ever reported for this enzyme.

In the future, it would be interesting to evaluate the MBCOMT behavior at long-term storage where the results obtained here can be a good start point for those trials. Moreover, the assessment of MBCOMT stability on purified fractions is a critical issue that also needs to be addressed.

Chapter 6 – References

[1] N. Jatana, A. Sharma, and N. Latha, *J. Mol. Graph. Model.*, vol. 39, pp. 145–64, 2013.

[2] M. Jo and P. N. Palma, CNS Drug Review, vol. 13, no. 3, pp. 352–379, 2007.

[3] N. J. H. Cotton, B. Stoddard, and W. W. Parson, *J. Biol. Chem.*, vol. 279, no. 22, pp. 23710–8, May 2004.

[4] L. A Passarinha, M. J. Bonifácio, P. Soares-da-Silva, and J. A Queiroz, *J. Chromatogr. A*, vol. 1177, no. 2, pp. 287–96, 2008.

[5] F. M. Santos, A. Q. Pedro, R. F. Soares, R. Martins, M. J. Bonifácio, J. A. Queiroz, and L. A. Passarinha, "*J. Sep. Sci.*, vol. 36, no. 11, pp. 1693–702, 2013.

[6] J. Chen, J. Song, P. Yuan, Q. Tian, Y. Ji, R. Ren-Patterson, G. Liu, Y. Sei, and D. R. Weinberger, *J. Biol. Chem.*, vol. 286, no. 40, pp. 34752–60, Oct. 2011.

[7] A.Q. Pedro, M.J. Bonifácio, J.A. Queiroz, C.J. Maia and L.A Passarinha, J. Biotechnology, vol. 156 pp.141-6, 2011.

[8] T. T. Myöhänen, N. Schendzielorz, and P. T. Männistö, *J. Neurochem.*, vol. 113, no. 6, pp. 1632–43, 2010.

[9] M. Matsumoto, C. S. Weickert, M. Akil, B. K. Lipska, T. M. Hyde, M. M. Herman, J. E. Kleinman, and D. R. Weinberger, *Neuroscience*, vol. 116, no. 1, pp. 127–37, 2003.

[10] N. Schendzielorz, J.-P. Oinas, T. T. Myöhänen, I. Reenilä, A. Raasmaja, and P. T. Männistö, *PLoS One*, vol. 8, no. 4, pp. e61392, Jan. 2013.

[11] F. F. Correia, F. M. Santos, A. Q. Pedro, M. J. Bonifácio, J. A. Queiroz, and L. A. Passarinha, *J. Sep. Sci.*, vol. 37, no. 1–2, pp. 20–9, 2014.

[12] J. Chen, B. K. Lipska, N. Halim, Q. D. Ma, M. Matsumoto, S. Melhem, B. S. Kolachana, T. M. Hyde, M. M. Herman, J. Apud, M. F. Egan, J. E. Kleinman, and D. R. Weinberger, Am. J. Hum. Genet., vol. 75 pp. 807–821, 2004.

[13] K. Rutherford, E. Alphandéry, a McMillan, V. Daggett, and W. W. Parson, *Biochim. Biophys. Acta*, vol. 1784, no. 7–8, pp. 1098–105, 2008.

[14] A Cerasa, A Cherubini, A Quattrone, M. C. Gioia, P. Tarantino, G. Annesi, F. Assogna, C. Caltagirone, and G. Spalletta, "*Neuroscience*, vol. 167, no. 3, pp. 809–14, 2010.

[15] R. Honea, B. a Verchinski, L. Pezawas, B. S. Kolachana, J. H. Callicott, V. S. Mattay, D. R. Weinberger, and A. Meyer-Lindenberg, *Neuroimage*, vol. 45, no. 1, pp. 44–51, 2009.

[16] L. A. Passarinha, M. J. Bonifácio, and J. A. Queiroz, vol. 438, pp. 430–438, 2007.

[17] V. S. Nunes, M. J. Bonifácio, J. A Queiroz, and L. A Passarinha, *Biomed. Chromatogr.*, vol. 24, no. 8, pp. 858–62, 2010.

[18] L. Klingelhoefer and H. Reichmann, *Basal Ganglia*, vol. 4, no. 1, pp. 9–13, 2014.

[19] J. E. Pompeu, L. A Arduini, a R. Botelho, M. B. F. Fonseca, S. M. A A Pompeu, C. Torriani-Pasin, and J. E. Deutsch, *Physiotherapy*, vol. 100, no. 2, pp. 162–8, 2014.

[20] M. J. Bonifácio, J. S. Sutcliffe, L. Torrão, L. C. Wright, and P. Soares-da-Silva, *Neuropharmacology*, vol. 77, pp. 334–41, 2014.

[21] http://www.pharmainfo.net/reviews/parkinsons-disease , assessed 11 June 2014

[22] A. Hatzimanolis, S. Vitoratou, L. Mandelli, C. Vaiopoulos, F. a Nearchou, C. N. Stefanis, A. Serretti, and N. C. Stefanis, *J. Affect. Disord.*, vol. 148, no. 2–3, pp. 316–22, 2013.

[23] F. Cormier, J. Muellner, and J.-C. Corvol, *J. Neural Transm.*, vol. 120, no. 4, pp. 665–71, 2013.

[24] F. Vilbois, P. Caspers, M. da Prada, G. Lang, C. Karrer, H. W. Lahm, and a M. Cesura, *Eur. J. Biochem.*, vol. 222, no. 2, pp. 377–86, 1994.

[25] H.-W. Bai, J.-Y. Shim, J. Yu, and B. T. Zhu, "*Chem. Res. Toxicol.*, vol. 20, no. 10, pp. 1409–25, 2007.

[26] B. Bertocci, V. Miggiano, M. Da Prada, Z. Dembic, H. W. Lahm, and P. Malherbe, *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. 4, pp. 1416–20, Feb. 1991.

[27] G. P. L. Cereghino, J. L. Cereghino, C. Ilgen, and J. M. Cregg, Current Opinion in Biotechnoloy, vol. 13 pp. 329–332, 2002.

[28] A M. Gonçalves, A Q. Pedro, C. Maia, F. Sousa, J. A Queiroz, and L. A Passarinha, *J. Microbiol. Biotechnol.*, vol. 23, no. 5, pp. 587–601, 2013.

[29] J. M. Cregg, J. L. Cereghino, J. Shi, and D. R. Higgins, Molecular Biotechnology, vol. 16, pp. 23-38, 2000.

[30] G. Potvin, A. Ahmad, and Z. Zhang, *Biochem. Eng. J.*, vol. 64, pp. 91–105, 2012.

[31] J. L. Cereghino and J. M. Cregg, *FEMS Microbiol. Rev.*, vol. 24, no. 1, pp. 45–66, 2000.

[32] M. A. Orman, C. Pınar, and H. Ozdamar, vol. 255, pp. 245–255, 2009.

[33] M. Gao and Z. Shi, *Chinese J. Chem. Eng.*, vol. 21, no. 2, pp. 216–226, 2013.

[34] H. Wang, J. Li, L. Liu, X. Li, D. Jia, G. Du, J. Chen, and J. Song, *Bioresour. Technol.*, vol. 124, pp. 338–346, 2012.

[35] A Ghosalkar, V. Sahai, and a Srivastava, *Bioresour. Technol.*, vol. 99, no. 16, pp. 7906–10, Nov. 2008.

[36] K. Sreekrishna, R. G. Brankamp, K. E. Kropp, D. T. Blankenship, J. T. Tsay, P. L. Smith, J. D. Wierschke, a Subramaniam, and L. ABirkenberger, *Gene*, vol. 190, no. 1, pp. 55–62, 1997.

[37] J. Zheng, W. Zhao, N. Guo, F. Lin, J. Tian, L. Wu, and H. Zhou, *Bioresour. Technol.*, vol. 118, pp. 257–64, 2012.

[38] O. Cos, R. Ramón, J. L. Montesinos, and F. Valero, Microbial Cell Factories, vol.5, pp. 1–20, 2006.

[39] J. Ye, J. Ly, K. Watts, A. Hsu, A. Walker, K. McLaughlin, M. Berdichevsky, B. Prinz, D. Sean Kersey, M. d'Anjou, D. Pollard, and T. Potgieter, *Biotechnol. Prog.*, vol. 27, no. 6, pp. 1744–50, 2011.

[40] EasySelect™ Pichia Expression Kit, no. 25, 2010.

[41] P. Yu, J. Yang, and H. Gu, *Carbohydr. Polym.*, vol. 106, pp. 223–9, 2014.

[42] C. Rabert, D. Weinacker, A. Pessoa, and J. G. Farías, *Braz. J. Microbiol.*, vol. 44, no. 2, pp. 351–6, 2013.

[43] D. Weinacker, C. Rabert, A. B. Zepeda, C. A. Figueroa, A. Pessoa, J. G. Farías, D. D. T. Bioquímico-farmacêutica, F. D. C. Farmacêuticas, U. D. S. Paulo, and S. Paulo, vol. 1048, no. 2013, pp. 1043–1048, 2014.

[44] T. Vogl, G. G. Thallinger, G. Zellnig, D. Drew, J. M. Cregg, A. Glieder, and M. Freigassner, *N. Biotechnol.*, vol. 00, no. 00, Mar. 2014.

[45] A. Ramón and M. Marín, *Biotechnol. J.*, vol. 6, no. 6, pp. 700–6, 2011.

[46] D. J. Leibly, T. N. Nguyen, L. T. Kao, S. N. Hewitt, L. K. Barrett, and W. C. Van Voorhis, *PLoS One*, vol. 7, no. 12, pp. e52482, 2012.

[47] S. E. Bondos and A. Bicknell, *Anal. Biochem.*, vol. 316, no. 2, pp. 223–231, 2003.

[48] H. L. White and J. C. Wu, *Biochem. J.*, vol. 145, no. 2, pp. 135–43, 1975.

[49] S. Belluzo, V. Boeris, B. Farruggia, and G. Picó, *Int. J. Biol. Macromol.*, vol. 49, no. 5, pp. 936–41, Dec. 2011.

[50] Y. Wei, C. Li, L. Zhang, and X. Xu, *Biochim. Biophys. Acta*, vol. 1838, no. 7, pp. 1911–1920, 2014.

[51] X. C. Tang and M. J. Pikal, *Pharm. Res.*, vol. 22, no. 7, pp. 1167–75, 2005.

[52] D. Barreca, G. Laganà, S. Magazù, F. Migliardo, and E. Bellocco, *Chem. Phys.*, vol. 424, pp. 100–104, 2013.

[53] A. Chefson, J. Zhao, and K. Auclair, *J. Biotechnol.*, vol. 130, no. 4, pp. 436–40, 2007.

[54] E. K. Barbour, a Abdelnour, F. Jirjis, O. Faroon, and M. T. Farran, *Vaccine*, vol. 20, no. 17–18, pp. 2249–53, 2002.

[55] Ahmad Fazeli, Soyed Abbas Shojaosadati, Mohammad Reza Fazeli, Khrosrow klalifeh, Sholred Ariaeenejad, Ali Akkbor moosavi-Movahedi, *J. Biotechnol.*, vol. 163, no. 3, pp. 318–24, Feb. 2013.

[56] G. Bonacci, D. Ceschin, G. Fidelio, and M. A. Vides, J. Biochem. Biophys. Methods, vol. 46, pp. 95–105, 2000.

[57] M. C. Alliegro, *Anal. Biochem.*, vol. 282, no. 1, pp. 102–6, Jun. 2000.

[58] D. Fass, *Annu. Rev. Biophys.*, vol. 41, pp. 63–79, Jan. 2012.

[59] N. Penha-Silva, L. R. Arvelos, C. C. Cunha, T. a Aversi-Ferreira, L. F. Gouvêa-e-Silva, M. S. Garrote-Filho, C. J. Finotti, M. Bernardino-Neto, and F. G. de Freitas Reis, *Bioelectrochemistry*, vol. 73, no. 1, pp. 23–9, Jun. 2008.

[60] L. S. B. Upadhyay and N. Verma, *Process Biochem.*, 2014.

[61] G. R. Castro, Enzyme and Microbial Technology, vol. 27, pp. 143–150, 2000.

[62] http://sites.bio.indiana.edu/~ptotocol_files/protein_storage. pdf assessed 20 June 2014

[63] http://www.uni-leipzig.de/uspdu/docs/Protein%20guide_Storage_Working.pdf assessed 20 June 2014

[64] Y.-S. Kim and B.-J. Yum, *Eng. Appl. Artif. Intell.*, vol. 17, no. 3, pp. 249–263, 2004.

[65] F. M. Alam, K. R. McNaught, and T. J. Ringrose, *Simul. Model. Pract. Theory*, vol. 12, no. 7–8, pp. 559–578, 2004.

[66] F. Amato, J. L. González-Hernández, and J. Havel, "Artificial neural networks combined with experimental design: a 'soft' approach for chemical kinetics.," *Talanta*, vol. 93, pp. 72–8, 2012.

[67] L. Vera Candioti, M. M. De Zan, M. S. Cámara, and H. C. Goicoechea, *Talanta*, vol. 124, pp. 123–138, 2014.

[68] D. Ghosh and P. C. Hallenbeck, *Bioresour. Technol.*, vol. 101, no. 6, pp. 1820–5, 2010.

[69] V. Y. Yatsyshyn, D. V. Fedorovych, and A. A. Sibirny, *Biochem. Eng. J.*, vol. 49, no. 1, pp. 52–60, 2010.

[70] R. Silva, S. Ferreira, M. J. Bonifácio, J. M. L. Dias, J. A Queiroz, and L. A Passarinha, *J. Biotechnol.*, vol. 160, no. 3–4, pp. 161–8, 2012.

[71] A. Afonso, P. Pereira, J. a Queiroz, Â. Sousa, and F. Sousa, *J. Chromatogr. A*, vol. 1343, pp. 119–27, 2014.

[72] C. Karlsson and M. G. Karlsson, "*J. Histochem. Cytochem.*, vol. 59, no. 12, pp. 1113–21, 2011.

[73] J. Tenhunen, M. Salminen, K. Lundström, T. Kiviluoto, R. Savolainen, and I. Ulmanen, *Eur. J. Biochem.*, vol. 223, no. 3, pp. 1049–59, Aug. 1994.

[74] Pedro AQ, Soares RF, Oppolzer D, Santos FM, Gonçalves AM, Bonifacio MJ, Queiroz JA, Gallardo E and Passarinha LA, *Chromatogr. Sep. Tech.*, p. 8, 2014.

[75] T. Vogl, F. S. Hartner, and A. Glieder, *Curr. Opin. Biotechnol.*, vol. 24, no. 6, pp. 1094–101, 2013.

[76] D. Tsao, S. Liu, and N. V Dokholyan, *Chem. Phys. Lett.*, vol. 506, no. 4–6, pp. 135–138, 2011.

[77] C. Di Salvo, D. Barreca, G. Laganà, M. di Bella, E. Tellone, S. Ficarra, and E. Bellocco, *Int. J. Biol. Macromol.*, vol. 57, pp. 63–8, Jun. 2013.

[78] C. Tilgmann and I. Ulmanen, *J. Chromatogr. B. Biomed. Appl.*, vol. 684, no. 1–2, pp. 147–61,1996.

[79] B. Bouma, J. Westerman, N. Dekker, P. Gros, and K. W. Wirtz, *Biochim. Biophys. Acta*, vol. 1546, no. 1, pp. 216–25, 2001.

[80] T. Tomoyasu, A. Tabata, Y. Ishikawa, R. a Whiley, and H. Nagamune, *J. Biosci. Bioeng.*, vol. 115, no. 1, pp. 15–9, 2013.

[81] S. Ratnayake, I. H. K. Dias, E. Lattman, and H. R. Griffiths, *J. Proteomics*, vol. 92, pp. 160–70, 2013.

[82] F. Gonzalez-Fernandez, D. Sung, K. M. Haswell, A. Tsin, and D. Ghosh, *Exp. Eye Res.*, vol. 120, pp. 167–74, 2014.

[83] M. N. Kiritsi, E. G. Fragoulis, and D. C. Sideris, *FEBS J.*, vol. 279, no. 7, pp. 1318–26, 2012.

[84] M. Z. Kamal, S. Ahmad, and N. M. Rao, *Biophys. Chem.*, vol. 156, no. 1, pp. 68–71, 2011.