UNIVERSIDADE DA BEIRA INTERIOR
Ciências

Biosynthesis, isolation and kinetic characterization of recombinant human catechol-*O*-methyltransferase from *Pichia pastoris* strains

Augusto Quaresma Henriques Pedro

Tese para a obtenção do grau de Doutor em
Bioquímica
(3º ciclo de estudos)

Orientador: Prof. Doutor Luís António Paulino Passarinha
Co-orientador: Prof. Doutor João António de Sampaio Rodrigues Queiroz

Covilhã, abril de 2016

To my family,

"If you can trust yourself
when all men doubt you,
…
If you can meet with triumph and disaster
and treat those two impostors just the same,
…
If you can talk with crowds and keep your virtue
Or walk with Kings – nor lose the common touch,
…
Yours is the Earth and everything that's in it,
And – which is more – you'll be a Man, my son!"

Rudyard Kipling

This work was financed by the Portuguese Foundation for Science and Technology (SFRH/BD/81222/2011) under the program QREN - POPH - Type 4.1 – Advanced Training, cofounded by the European Social Fund and by national funds from the MCTES. Moreover, this work was also supported in the framework of a research and development project (EXPL/BBB478/0960/2012) and strategic programs (PEst-C/SAU/UI0709/2011 and Pest-OE/SAU/UI0709/2014) co-founded by the operational program factors of competitiveness – COMPETE (FCOMP-01-0124-FEDER-041068 – EXPL/QEQ-MED/1068/2013) and co-founded by FEDER funds (FCOMP-01-0124-FEDER-027560) of COMPETE.

Final version printed in April 2016

Acknowledgments

First and foremost, I'm deeply grateful to my supervisors, Professor Luís Passarinha and Professor João Queiroz for all the careful guidance throughout this thesis. Yours valuable scientific knowledge, dedication, motivation, support, professionalism and teachings during the last years were major determinants for successfully conducting this work. Also, I'm thankful for all the friendship and availability. It was a huge privilege working with you.

I'm deeply grateful to Fundação para a Ciência e Tecnologia (FCT) for the financial support through the PhD fellowship (SFRH/BD/81222/2011).

I'm thankful to all the people from University of Beira Interior and those involved in Health Sciences Research Centre which cooperated in the development of this work. In particular, a special acknowledgment to the laboratory technicians Margarida Carrilho and Sofia Duarte that really cooperated for the successful development of this work.

To all my teachers, specially Professor Eugenia Gallardo and Cláudio Maia, I'm grateful for their help and support, their scientific expertise and kind collaboration throughout this thesis.

To Professor Fani, I'm grateful for all the friendship, teachings, help and support.

To all my colleagues, specially PhD Ângela Sousa, Susana Ferreira, Filomena Silva, Rita Martins, Ana Martinho and Msc Luís Martins, David Oppolzer, Eduarda Coutinho, Fátima Santos and Margarida Gonçalves, I would like to acknowledge all the teachings, friendship, support and kind collaboration during the course of this thesis.

To Patrícia, thank you not only by all the friendship and love but also for the companionship and collaboration in the good and bad times. I'm sure that you own a piece of the success of this work.

To my father, for all you mean to me; my brother, my fellow always forever; my uncles, always there for me; my grandmother and my uncle, I will always be grateful for your love, friendship and support.

Resumo

A catecol-*O*-metiltransferase (COMT; EC 2.1.1.6) é uma enzima dependente de magnésio que cataliza a reação de metilação de diferentes substratos com estrutura catecólica, nomeadamente catecolaminas, catecóis xenobióticos e catecolestrogénios. Em humanos, esta enzima encontra-se presente sob duas isoformas, uma solúvel (SCOMT) e uma membranar (MBCOMT), esta última associada à membrana do retículo endoplasmático rugoso.

Após os primeiros estudos realizados com estas enzimas, estabeleceu-se que estas poderiam estar potencialmente envolvidas em diversas patologias humanas. Em particular, chegou-se à conclusão que a sua inibição origina grandes benefícios na doença de Parkinson onde leva a um aumento da biodisponibilidade da levodopa, administrada para colmatar a falta de dopamina causada pela degeneração dos neurónios dopaminérgicos. Assim, o desenvolvimento de moléculas com capacidade para inibir a atividade biológica da COMT com potência e seletividade aperfeiçoadas poderia melhorar o prognóstico destes doentes. O processo para o desenvolvimento destas moléculas geralmente inclui a realização de estudos estruturais baseados em estruturas previamente determinadas que dependem da obtenção de elevadas quantidades de proteínas purificadas.

A determinação de estruturas tridimensionais de proteínas é geralmente dificultada por vários obstáculos, principalmente devido ao fato de se encontrarem em quantidades reduzidas nas suas fontes naturais e à dificuldade em obter cristais cuja Difração de Raios-X produza resultados positivos. Além disso, uma vez que as proteínas membranares se encontram naturalmente embebidas numa bicamada lipídica, a determinação da sua estrutura enfrenta desafios adicionais.

Visto que do ponto de vista laboratorial é praticamente impossível purificar a maioria destes alvos proteicos diretamente a partir das suas fontes naturais, em biologia estrutural, geralmente utiliza-se a produção recombinante heteróloga de proteínas em hospedeiros procariotas ou eucariotas. Adicionalmente, para isolar as proteínas-alvo heterólogas dos restantes contaminantes do hospedeiro, é necessário o desenvolvimento e implementação de estratégias de purificação, principalmente através de procedimentos cromatográficos. Ao longo deste processo, também é relevante que a estratégia desenvolvida permita que as proteínas mantenham o seu estado nativo e, consequentemente, biologicamente ativo. Desta forma, é importante evitar que a proteína se acumule num estado desenrolado e não funcional durante a sua biossíntese ou, por outro lado, evitar a sua agregação após a sua recuperação e purificação. Assim sendo, o principal objetivo deste trabalho é o desenvolvimento de uma abordagem integrada, simples e eficiente para a biossíntese e purificação de ambas as isoformas da COMT numa forma biologicamente ativa, passíveis de posteriores aplicações em estudos estruturais.

Não existe um único hospedeiro que reúna todas as características desejadas para a biossíntese recombinante de proteínas. No entanto, encontra-se descrito que a *Pichia pastoris* é capaz de realizar muitas modificações pós-tradução necessárias para o correto enrolamento de algumas proteínas e obtêm-se elevadas densidades celulares em meios relativamente pouco dispendiosos. Assim, tendo em conta estas vantagens, este hospedeiro foi selecionado neste trabalho para a expressão de ambas as isoformas da COMT. Por outro lado, procurou-se explorar a elevada seletividade geralmente associada à cromatografia de afinidade, tendo-se por isso definido este tipo de cromatografia como o principal passo de isolamento e purificação da COMT.

A determinação da atividade enzimática da COMT apresenta elevada importância do ponto de vista científico em todas as investigações relacionadas com a COMT recombinante. Desta forma, tendo em vista a quantificação da metanefrina produzida em ensaios enzimáticos da COMT, desenvolveu-se e implementou-se um método analítico baseado em cromatografia líquida de elevada performance com deteção coulométrica. Além do tempo de corrida de cada ensaio ser inferior aos reportados anteriormente, as melhorias na razão sinal/ruído atingidas com a deteção coulométrica permitiram detetar e quantificar quantidades inferiores do analito em questão, levando a um aumento da sensibilidade deste método em comparação com métodos alternativos descritos na literatura.

De seguida, desenvolveu-se uma estratégia integrada para a biossíntese da SCOMT em *Pichia pastoris* e subsequente purificação por cromatografia de afinidade com iões metálicos imobilizados. Através da aplicação com sucesso desta estratégia, frações de SCOMT com elevada pureza foram obtidas. Estas frações foram ainda avaliadas por ionização por dessorção a laser assistida por matriz com deteção em dois analisadores do tipo tempo de vôo (em inglês, MALDI-TOF/TOF) e verificou-se que apresentavam uma estrutura primária idêntica à correspondente SCOMT nativa, indicando que foi corretamente processada pela maquinaria intracelular do hospedeiro. Também se determinou a sensibilidade (IC_{50} - Concentração de composto necessária para inibir 50% da atividade enzimática máxima da enzima-alvo) da SCOMT para ser inibida por dois inibidores disponíveis comercialmente - entacapone e 3,5-dinitrocatecol - e obtiveram-se valores da mesma ordem de grandeza dos que foram reportados anteriormente. Em conjunto, estes resultados vêm corroborar que a proteína é idêntica à sua correspondente nativa. Perante o sucesso atingido com esta estratégia, como a expressão heteróloga de proteínas membranares é geralmente mais complexa do que a expressão de proteínas solúveis e visto que existem menos dados na literatura relativamente à MBCOMT quando comparada com a correspondente isoforma solúvel, as etapas subsequentes na dissertação tiveram como foco principal a MBCOMT.

No que diz respeito à expressão recombinante em *Pichia pastoris* da MBCOMT, desenvolveram-se e implementaram-se protocolos para a sua produção em pequena escala ou, alternativamente, em mini-biorreatores. Inicialmente, estudaram-se os perfis de

produção da MBCOMT em erlenmeyer por duas estirpes de *Pichia pastoris* com fenótipos opostos e em condições experimentais distintas. Efetivamente, verificou-se que os níveis de produção da proteína-alvo foram superiores quando a indução foi realizada apenas com metanol, por contraposição aos ensaios onde se utilizou metanol juntamente com glicerol ou sorbitol em modo de alimentação mista. Apesar da performance das duas estirpes em sintetizar a MBCOMT ter sido semelhante, selecionou-se para ensaios posteriores a estirpe X33 (Mut^+), em detrimento da KM71H (Mut^S). De fato, como a estirpe KM71H é mais influenciada pela concentração de indutor utilizada, a eventual aplicação de uma concentração de metanol inadequada compromete os níveis de produção, o que não se verifica com a estirpe X33.

No entanto, os resultados mais promissores foram obtidos em mini-biorreatores onde a fase de indução foi otimizada e modelada através de uma rede neuronal artificial. Dos fatores otimizados, destaca-se a adição de dimetilsulfóxido, um agente que atua como um chaperone químico, melhorando a produção de proteínas recombinantes membranares. Além deste bioprocesso não interferir com a viabilidade celular, permitiu a obtenção de elevados níveis de MBCOMT enzimaticamente ativa.

Após a implementação de um processo eficiente para a produção recombinante da MBCOMT, desenvolveram-se estratégias para a sua purificação. Inicialmente, utilizou-se cromatografia de afinidade com arginina imobilizada para o isolamento da MBCOMT a partir de lisados de *Pichia pastoris*. Durante este processo, avaliou-se o efeito do pH, temperatura e a concentração da amostra injetada na adsorção e grau de purificação da enzima-alvo. Assim, a estratégia otimizada foi obtida a pH 7 e conduzida a 20ºC com uma amostra inicial mais concentrada, a partir da qual a MBCOMT foi recuperada num estado intermédio de purificação. Finalmente, desenvolveu-se uma metodologia para a purificação de uma forma recombinante da MBCOMT com um tag de 6 histidinas no terminal carboxilo por cromatografia de afinidade com iões metálicos imobilizados. Apesar deste trabalho ainda não estar concluído, o seu progresso permitiu avaliar o efeito do tag na atividade biológica da MBCOMT e verificou-se que a introdução de um local de clivagem para uma protease leva à perda total da atividade biológica, o que não se observou quando a enzima contém apenas o tag. Finalmente, testaram-se vários detergentes para a solubilização da MBCOMT e verificou-se que o dodecilmaltosídeo foi o mais adequado. De seguida, efetuaram-se vários ensaios para avaliar o comportamento cromatográfico da proteína-alvo com 6 histidinas em diferentes condições experimentais, nomeadamente com diferentes tampões cromatográficos e diferentes iões imobilizados na matriz, nomeadamente Níquel e Zinco. No entanto, para obter uma separação efetiva da enzima-alvo dos demais contaminantes, serão necessários testar gradientes por passos com diferentes concentrações de imidazol ou, alternativamente, testar outros métodos de eluição como a diminuição do pH e a competição com histidina.

O progresso atingido com a elaboração deste trabalho permitiu de certa forma atingir os requisitos altamente exigentes necessários para a realização de diversas técnicas biofísicas, particularmente no que diz respeito à fase de produção e estabilização da COMT. De fato, com as estratégias anteriormente reportadas, foram obtidas quantidades moderadas a elevadas de ambas as isoformas da COMT num estado cataliticamente ativo. A estratégia aqui reportada para a isoforma solúvel permite a obtenção da proteína num estado de elevada pureza que quando associada a um passo de cromatografia de exclusão molecular poderá efetivamente cumprir os requisitos necessários para realizar estudos estruturais ou estudos de bio-interação com ligandos ou inibidores já comercializados.

Em sentido inverso, as estratégias reportadas neste trabalho com sucesso para a purificação parcial da MBCOMT necessitam de ser melhoradas, particularmente para a cromatografia de afinidade com iões metálicos imobilizados. Visto que este tipo de cromatografia é geralmente caracterizado por apresentar elevada seletividade, é exequível que a realização de ensaios adicionais poderá levar à obtenção de frações com elevada pureza. Desta forma, com a elaboração destes ensaios adicionais, poder-se-á obter proteína-alvo em condições que cumpram os requisitos necessários para realizar estudos estruturais e de bio-interação utilizando a apoenzima ou a enzima complexada com diferentes ligandos tais como cofatores ou inibidores já disponíveis. Estes estudos poderão ser conduzidos por ressonância magnética nuclear, calorimetria de titulação isotérmica ou até experiências cristalográficas.

Relativamente às perspetivas futuras, poderá vir a revelar-se útil desenvolver sistemas de expressão sem células, principalmente para a isoforma membranar da COMT visto que se poderá acoplar a sua reconstituição em sistemas não micelares tais como nanodiscos ou lipossomas, que poderá contribuir para a sua estabilização. Adicionalmente, tendo em vista a obtenção de frações purificadas com estas enzimas numa forma nativa, isto é, sem introdução de tags de histidinas por exemplo, poderá explorar-se a cromatografia de imunoafinidade. Através da imobilização de um anticorpo que reconheça um epítopo comum a ambas as isoformas, esta matriz poderá demonstrar elevada versatilidade e ser aplicada tanto para a purificação da COMT solúvel como da membranar.

Palavras-chave

Doença de Parkinson; Proteínas Recombinantes; Catecol-*O*-metiltransferase; *Pichia pastoris*; Biorreator; Cromatografia de Afinidade; Purificação de Proteínas.

Abstract

Catechol-*O*-methyltransferase (COMT; EC 2.1.1.6) is a magnesium-dependent enzyme that catalyzes the methylation reaction of different catecholic substrates such as catecholamines, xenobiotic catechols and catecholestrogens. Following the initial characterizations of these enzymes, it was described that they are potentially involved in diverse human disorders. Specifically, as its inhibition has proven to be of great interest in neurologic disorders such as Parkinson's disease, developing inhibitor molecules with increased potency and selectivity may improve the outcome of these patients. These molecules are usually accomplished using structure-based drug design studies that rely on the attainment of highly purified protein quantities. Indeed, challenges in the determination of protein structures are mainly associated with their low natural abundance coupled with the difficulty of obtaining crystals amenable to X-Ray diffraction. In particular, as membrane proteins are naturally embedded in the lipid bilayer, the determination of their structure faces additional difficulties.

As it is unrealistic to purify all of these targets from their natural sources, structural biology of proteins usually focus on the recombinant heterologous expression of these proteins onto an expression host. In addition, to isolate the target proteins from the other major host contaminants, equally appropriated purification strategies need to be designed and implemented, mostly using chromatographic procedures. Throughout this entire process, is also important that the developed strategy is able to keep the proteins in a stable and functional active form, thus avoiding its misfolding during biosynthesis and aggregation after its recovery and isolation in the downstream processing. Therefore, the main scope of this work is the development of a straightforward approach that allows the biosynthesis, isolation and purification of recombinant human COMT isoforms in a biologically active form for further application in structural studies or to evaluate their role as potential therapeutic proteins. Specifically, although no single host can provide all the desired properties for recombinant protein biosynthesis, *Pichia pastoris* is able to perform many post-translational modifications and is cultivated at high cell-densities in moderately cheap media. Therefore, in this work, it was selected for expression of COMT enzymes. On the other hand, the high selectivity often provided by affinity chromatography prompted us to employ it as the main isolation and purification step.

The determination of COMT enzymatic activity is greatly important in COMT recombinant research, either to assess COMT activity from recombinant lysates or purified fractions, for detergent-solubilized or unsolubilized samples and for both isoforms. Therefore, a faster and more sensitive analytical method based on HPLC coupled with coulometric detection was developed for quantifying metanephrine in these assays. Then, an integrated strategy for recombinant soluble catechol-*O*-methyltransferase (SCOMT) biosynthesis onto *P. pastoris* and

purification using immobilized-metal affinity chromatography was implemented where highly purified fractions of this target enzyme were obtained.

On the other hand, as heterologous membrane protein overexpression is usually more challenging than soluble proteins and less reports are available in the literature with recombinant human membrane-bound catechol-*O*-methyltransferase (MBCOMT) than COMT soluble isoform, our work were mostly focused on MBCOMT. Here, we established protocols for MBCOMT expression in *Pichia pastoris* methanol-induced cultures in baffled shake-flasks and mini-bioreactors. In particular, the optimization of the induction phase using artificial neural networks in mini-biorreactors allowed achieving high levels of biologically active MBCOMT. Then, arginine-affinity chromatography was successfuly applied for the direct capture of MBCOMT from *Pichia pastoris* lysates and it was recovered in a moderate purified form. Finally, the ongoing work is related to the purification of a hexa-histidine tagged form of MBCOMT using immobilized-metal affinity chromatography. Indeed, despite significant achievements were made concerning the construction of a tagged form of MBCOMT solubilized with an appropriated detergent in a biologically active form, additional stepwise gradients are required to effectively separate MBCOMT from the other contaminants.

In conclusion, the progress achieved with this work meets the highly demanding requirements of biophysical techniques, mainly regarding the upstream stage as well as COMT stabilization where moderate to high quantities of catalitically active enzymes were obtained. In particular, coupling the strategy here reported for SCOMT with a final polishing step will probably allow performing structural or bio-interaction studies with this enzyme. Nonetheless, the strategies here described successfully for partial MBCOMT purification need to be improved, especially for immobilized-metal affinity chromatography once it is considered to be highly selective and, thus, it is feasible that after succesful optimization procedures, fractions with high purity will be obtained. Therefore, the strategies here reported with the intensification and optimization of some procedures would possible permit performing structural and bio-interaction studies using the apo-enzymes or complexed with different ligands (cofactors or inhibitors) by Nuclear Magnetic Ressonance, Isothermal Titration Calorimetry or even using Crystallographic experiments.

Keywords

Parkinson's Disease; Recombinant Proteins; Catecol-*O*-methyltransferase; *Pichia pastoris*; Bioreactor; Affinity Chromatography; Protein Purification.

Thesis Overview

This thesis is structured in four main chapters. The **first chapter** is divided into three sections. The first section presents an explanation concerning the importance and interest of the enzyme catechol-*O*-methyltransferase and its potential pathophysiological role in several human diseases, especially Parkinson's disease. Also in this section, is highlighted the importance of developing systems for the recombinant expression of enzymes, particularly the catechol-*O*-methyltransferase as well as the implementation of chromatographic strategies in order to obtain pure samples in high quantity for structural and functional studies. This section results in a general introduction to the main issues/challenges presented in sections 2 and 3. The second section of the first chapter deals with the implementation and development of biosynthesis strategies in order to circumvent common problems associated to membrane protein structure determination (Paper 1 - From gene to structure: biosynthesis strategies to break the bottlenecks associated to membrane protein structure determination). Finally, the last section of chapter 1 adress the biosystem *Pichia pastoris* as a recombinant microfactory for the production of membrane proteins. In particular, are discussed *Pichia pastoris* morphological and biomolecular characteristics, biosynthesis strategies for protein production in *Pichia pastoris* and typical fermentation flowsheets usually applied for membrane proteins.

The **second chapter** includes the global as well as the intermediate aims established for the implementation and development of this study. Accordingly, the **third chapter** includes the presentation and discussion of the results obtained during this research work in the form of original research papers organized as follows:

Paper II – An improved HPLC method for quantification of metanephrine with coulometric detection.

Paper III – Biosynthesis and Purification of histidine-tagged human soluble catechol-*O*-methyltransferase.

Paper IV – Evaluation of Mut^S and Mut^+ *Pichia pastoris* strains for membrane-bound catechol-*O*-methyltransferase biosynthesis.

Paper V – An artificial neural network for membrane-bound catechol-*O*-methyltransferase biosynthesis with *Pichia pastoris* methanol-induced cultures.

Paper VI – Purification of membrane-bound catechol-*O*-methyltransferase by arginine-affinity chromatography.

Paper VII – Purification of histidine-tagged membrane-bound catechol-*O*-methyltransferase immobilized-metal affinity chromatography.

The **fourth chapter** summarizes the concluding remarks obtained during this research work, mostly concerning the application of *Pichia pastoris* for the small-scale and bioreactor production of recombinant catechol-*O*-methyltransferase isoforms as well as the novel chromatographic methods implemented in this work for its purification. Also, the impact of the findings here reported in developing structural and functional studies using both catechol-*O*-methyltransferase isoforms will also be addressed. Finally, it will be also presented the future trends of this thesis and additional research work beyond what has been performed here.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

3-MT	3-Methoxytyramine
3-OMD	3-*O*-methyldopa
AADC	Aromatic Amino Acid Decarboxylase
ANN	Artificial Neural Network
AOX	Alcohol Oxidase
BBB	Blood-Brain Barrier
BSM	Basal Salts Medium
COMT	Catechol-*O*-methyltransferase
DDM	n-dodecyl-β-$_D$-maltopyranoside
DNC	3,5-dinitrocatechol
DOPAC	Dihydroxyphenylacetic acid
EAE	Experimental Autoimmune Encephalomyelitis
E. coli	*Escherichia coli*
GPCR	G-Protein Coupled Receptor
GRAS	Generally Recognized As Safe
IC_{50}	Half maximal inhibitory concentration
IMAC	Immobilized-Metal Affinity Chromatography
IPTG	Isopropylthiogalactosidase
L-DOPA	Levodopa
LB	Lysogeny broth
MALDI-TOF/TOF	Matrix Assisted Laser Desorption/Ionization with tandem Time Of Flight
MAO	Monoamine oxidase
MBCOMT	Membrane-bound Catechol-*O*-methyltransferase
MBCOMT-His_6	Hexa-histidine tagged membrane-bound catechol-*O*-methyltransferase
MBCOMT-TEV-His_6	Hexa-histidine tagged membrane-bound catechol-*O*-methyltransferase with a linker for *Tobacco Etch Virus*
MES	2-(N-morpholino)-ethanesulfonic acid
Met	Methionine
MP	Membrane Protein
MPD	(4S)-2-methyl-2,4-pentanediol
PD	Parkinson's Disease
P. pastoris	*Pichia pastoris*
SAH	S-adenosyl-$_L$-homocysteine
SAM	S-adenosyl-$_L$-methionine
SCOMT	Soluble Catechol-*O*-methyltransferase
SCOMT-His_6	Hexa-histidine tagged soluble catechol-*O*-methyltransferase
Sf9	*Spodoptera frugiperda*
TEV	*Tobacco Etch Virus*
Val	Valine

List of Scientific Publications

Papers related with this Thesis

I. From gene to structure: biosynthesis strategies to break the bottlenecks associated to membrane protein structure determination

A. Q. Pedro, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

Submitted for publication (2015)

II. An improved HPLC method for quantification of metanephrine with coulometric detection

A. Q. Pedro, R. F. Soares, D. Oppolzer, F. M. Santos, L. A. Rocha, A. M. Gonçalves, M. J. Bonifacio, J. A. Queiroz, E. Gallardo, L. A. Passarinha

Journal of Chromatography and Separation Techniques (2014) 5 (2): 217-224

III. Biosynthesis and Purification of histidine-tagged human soluble catechol-*O*-methyltransferase

A. Q. Pedro, F. F. Correia, F. M. Santos, G. Espírito-Santo, A. M. Gonçalves, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

Submitted for publication (2015)

IV. Evaluation of Mut^S and Mut$^+$ *Pichia pastoris* strains for membrane-bound catechol-*O*-methyltransferase

A. Q. Pedro, D. Oppolzer, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

Applied Biochemistry and Biotechnology (2015) 175 (8): 3840-3855

V. An artificial neural network for membrane-bound catechol-*O*-methyltransferase biosynthesis with *Pichia pastoris* methanol-induced cultures

A. Q. Pedro, L. M. Martins, J. M. Dias, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

Microbial Cell Factories (2015) 14: 113-127

VI. Purification of membrane-bound catechol-*O*-methyltransferase by arginine-affinity chromatography

A. Q. Pedro, P. Pereira, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

Chromatographia (2015) 78 (21): 1339-1348

VII. Purification of histidine-tagged membrane-bound catechol-*O*-methyltransferase by immobilized-metal affinity chromatography

A. Q. Pedro, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

Ongoing work

Papers not related with this Thesis

I. A novel prokaryotic expression system for biosynthesis of recombinant human membrane-bound catechol-*O*-methyltransferase

A. Q. Pedro, M. J. Bonifacio, J. A. Queiroz, C. J. Maia, L. A. Passarinha

Journal of Biotechnology (2011) 156 (2): 141-146

II. Performance of hydrophobic interaction ligands for human membrane-bound catechol-*O*-methyltransferase purification

F. M. Santos, A. Q. Pedro, R. F. Soares, R. Martins, M. J. Bonifacio, J.A. Queiroz, L. A. Passarinha

Journal of Separation Science (2013) 36 (11): 1693-1702

III. *Pichia pastoris*: a recombinant microfactory for antibodies and human membrane proteins

A. M. Gonçalves, A. Q. Pedro, C. Maia, F. Sousa, J. A. Queiroz, L. A. Passarinha

Journal of Microbiology and Biotechnology (2013) 23 (5): 587-601

IV. Recovery of biological active catechol-*O*-methyltransferase isoforms from Q-sepharose

F. F. Correia, F. M. Santos, A.Q. Pedro, M. J. Bonifácio, J. A. Queiroz, L.A. Passarinha

Journal of Separation Science (2014) 37 (1-2): 20-29

V. Trends in protein-based biosensor assemblies for drug screening and pharmaceutical kinetic studies

A. M. Gonçalves, A.Q. Pedro, F. M. Santos, L. M. Martins, C. J. Maia, J.A. Queiroz, L.A. Passarinha

Molecules (2014) 19 (8): 12461-12485

VI. Development of fed-batch profiles for efficient biosynthesis of catechol-*O*-methyltransferase

G. M. Espírito Santo, A.Q. Pedro, D. Oppolzer, M. J. Bonifacio, J.A. Queiroz, F. Silva, L.A. Passarinha

Biotechnology Reports (2014) 3: 34-41

VII. Enhanced biosynthesis of plasmid DNA from *Escherichia coli* VH33 using Box-Behnken design associated to aromatic amino acids pathway

L. M. Martins, A. Q. Pedro, D. Oppolzer, F. Sousa, J. A. Queiroz, L. A. Passarinha

Biochemical Engineering Journal (2015) 15: 117-126

VIII. Advances in time-course extracellular production of human pre-miR-29b from *Rhodovulum sulfidophilum*

P. Pereira, A. Q. Pedro, J. Tomás, C. J. Maia, J. A. Queiroz, A. Figueiras, F. Sousa

Submitted for publication (2015)

List of Scientific Communications

Poster Communications related with this Thesis

I. Strategies to improve recombinant human membrane protein overexpression for structural, functional and clinical studies

A. Q. Pedro, M. J. Bonifacio, C. J. Maia, J. A. Queiroz, L. A. Passarinha

V Jornadas sobre Tecnologia e Saúde 2012. Guarda, Portugal

II. Evaluation of a Mut^S and Mut$^+$ *Pichia pastoris* strains for membrane-bound COMT biosynthesis

A. Q. Pedro, D. Oppolzer, M. J. Bonifacio, C. J. Maia, J. A. Queiroz, L. A. Passarinha

Portuguese Congress of Microbiology and Biotechnology – Microbiotec 2013. Aveiro, Portugal

III. Evaluation of a Mut^S and Mut$^+$ *Pichia pastoris* strains for membrane-bound COMT biosynthesis

A. Q. Pedro, D. Oppolzer, M. J. Bonifacio, C. J. Maia, J. A. Queiroz, L. A. Passarinha

Modern biophysical methods for protein-ligand interactions – EMBO practical course 2013. Oulu, Finland

IV. Biosynthesis, Purification and Biointeraction of human SCOMT with Parkinson's disease inhibitors

F. F. Correia, A. Q. Pedro, D. Oliveira, J. A. Queiroz, L. A. Passarinha

16th European Congress on Biotechnology (ECB) 2014. Edinburgh, Scotland

V. Membrane-bound catechol-*O*-methyltransferase biosynthesis in *Pichia pastoris*: optimization of the induction phase using artificial neural networks

A. Q. Pedro, L. M. Martins, J. M. L. Dias, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

Protein & Antibody Engineering Summit (PEGS) Europe 2014. Lisboa, Portugal

VI. An artificial neural network for membrane-bound catechol-*O*-methyltransferase biosynthesis in mini-bioreactors with *Pichia pastoris* methanol-induced cultures networks

A. Q. Pedro, L. M. Martins, J. M. L. Dias, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

8th Conference on Recombinant Protein Production (RPP) 2015. Mallorca, Spain

Oral Communications not related with this Thesis

I. Advances in time-course biosynthesis and solubilization of MBCOMT from *Brevibacillus* cells

A. Q. Pedro, M. J. Bonifacio, J. A. Queiroz, C. J. Maia, L. A. Passarinha

VI Anual CICS Symposium 2011. Covilhã, Portugal

II. Kinetics of a recombinant form of human membrane-bound Catechol-*O*-methyltransferase

Filipa Sousa, Maria João Bonifácio, Augusto Q. Pedro, Luís A. Passarinha, Lyndon Wright, Patrício Soares-da-Silva

XLII Reunião da Sociedade Portuguesa de Farmacologia 2012. Lisboa, Portugal

III. A novel prokaryotic expression system for biosynthesis of recombinant human membrane-bound catechol-*O*-methyltransferase

A. Q. Pedro, F. Sousa, M. J. Bonifacio, C. J. Maia, J. A. Queiroz, L. A. Passarinha

ESBES (9th European Symposium on Biochemical Engineering Science) 2012. Istanbul, Turkey

IV. Evaluation of human membrane-bound catechol-*O*-methyltransferase purification by hydrophobic interaction chromatography

F. M. Santos, A. Q. Pedro, R. Martins, C. J. Maia, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

ISPPP (32nd International Symposium on the Separation of Proteins, Peptides and Polynucleotides) 2012. Istanbul, Turkey

V. Purification of recombinant human membrane COMT by ionic-exchange chromatography

F. Correia, F. Santos, A. Q. Pedro, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

VIII Anual CICS Symposium 2013. Covilhã, Portugal

VI. Biosynthesis, purification and brain-targeted delivery of pre-miR-29b biopharmaceuticals for Alzheimer's disease therapy

P. Pereira, A. Q. Pedro, J. Tomás, C. Cruz, A. Sousa, C. Maia, J. A. Queiroz, A. Figueiras, F. Sousa

X Annual CICS Symposium 2015. Covilhã, Portugal

Poster Communications not related with this Thesis

I. Biosynthesis of human membrane-bound catechol-*O*-methyltransferase optimization using Plackett-Burman and central composite rotatable design

R. Soares, S. Ferreira, A. Q. Pedro, A. M. Gonçalves, C. J. Maia, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

ESBES (9th European Symposium on Biochemical Engineering Science) 2012. Istanbul, Turkey

Chapter 1

Section 1 - Introduction

1. General aspects on the catechol-*O*-methyltransferase enzymes

Catechol-*O*-methyltransferase (COMT; EC 2.1.1.6) is a magnesium-dependent enzyme that catalyzes the methylation reaction whereby a methyl group from S-adenosyl-$_L$-methionine (SAM) is transferred to one of the catecholic hydroxyls, according to Figure 1 [Bonifacio *et al.*, 2007; Ma *et al.*, 2014]. As reaction products, are obtained the *O*-methylated catechol and S-adenosyl-$_L$-homocysteine (SAH) [Bonifacio *et al.*, 2007].

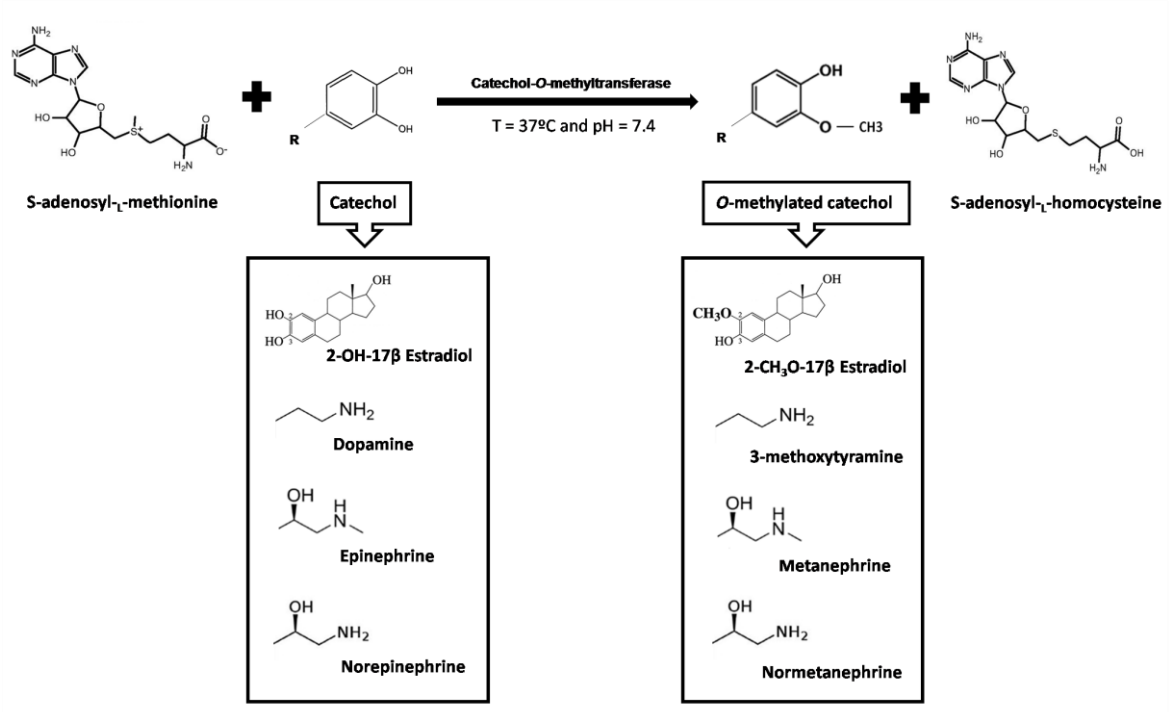


Figure 1 -Typical *O*-methylation reaction of different catechol substrates by COMT.

The methylation reaction by COMT occurs via a sequentially ordered mechanism where SAM firstly binds to the enzyme, followed by the Mg^{2+} ion and the substrate [Ma *et al.*, 2014]. This order is necessary because the binding of one changes the shape of the binding site so that the next can bind [Thomas and Potter, 2013]. Typically, the substrates of COMT (Please see Figure 1) in mammals include catecholestrogens and their metabolites, catecholamines with hormonal and neurotransmission activities such as dopamine, norepinephrine, epinephrine, ascorbic acids, some indolic intermediates of melanin metabolism and xenobiotic catechols like carcinogenic catechol-containing flavonoids [Bonifacio *et al.*, 2007].

COMT exists as two molecular forms, a soluble (SCOMT) and a membrane-bound (MBCOMT) isoform, both encoded by the same gene located on chromosome 22 band q11.2 [Müller, 2015].

In humans, SCOMT is a 221 amino acid protein with 24.7 KDa while MBCOMT has 30 KDa since it contains an additional stretch of 50 amino acid, the signal sequence for membrane anchoring [Bonifacio *et al.*, 2007; Ma *et al.*, 2014]. MBCOMT is anchored to the rough endoplasmic reticulum membrane and is characterized as an integral membrane protein (MP) with the catalytic portion of the enzyme oriented toward the cytoplasmic side of the membrane [Bonifacio *et al.*, 2007]. On the other hand, SCOMT is thought to be located in the cytosol [Bonifacio *et al.*, 2007], although it has also been reported in the nucleus of rat transfected COS-7 cells [Lundtröm *et al.*, 1995] as well as in mammary epithelial cells in the presence of increased levels of catecholestrogens [Weisz *et al.*, 2000].

1.1. The catechol-*O*-methyltransferase Valine/Methionine 108/158 polymorphism

Two codominant alleles in exon 4 of the COMT gene influence the amino acid structure [Valine (Val) or Methionine (Met)] at codon 108/158, respectively for soluble and MBCOMT [Bonifacio *et al.*, 2007; Hosak, 2007]. As a matter of fact, COMT enzymatic activity is genetically polymorphic with a trimodal activity distribution (High in Val/Val, intermediate in Val/Met and low in Met/Met genotypes) [Bonifacio *et al.*, 2007; Hosak, 2007]. The Met 108/158 variant is associated with low enzymatic activity and decreased thermal stability [Bonifacio *et al.*, 2007]. It was also reported that the Val158Met polymorphism appears to have pleiotropic effects on human behavior [Hosak, 2007] and recently it was demonstrated that the COMT inhibitor tolcapone-mediated cytotoxicity is influenced by the Val/Met genotypes [Chen *et al.*, 2011]. Recently, it was demonstrated that the Val/Met polymorphism influence plasma total homocysteine levels [Tunbridge *et al.*, 2008]. Indeed, it was observed that COMT Val158 individuals had significantly higher plasma total homocysteine levels than Met158 homozygotes [Tunbridge *et al.*, 2008]. Another study revealed that this polymorphism is associated with incident cardiovascular disease and this association was modified by randomized allocation to aspirin, a commonly used cardiovascular prevention agent [Hall *et al.*, 2014].

In several breast cancer *in vitro* models (MCF-7, MCF-10A and MCF-10F), it has been demonstrated that COMT plays an important role in the methylation of catechol estrogens, by blocking their oxidative metabolism to reactive quinones and the associated formation of reactive oxygen species [Yager, 2015]. Although with inconsistent results, the low activity polymorphism in COMT has been shown to cause a small but significant increased risk for developing breast cancer [Yager, 2012; Yager, 2015].

1.2. Soluble vs membrane-bound COMT: relative distribution and proposed function

COMT gene codes for two separates enzymes, soluble and MBCOMT that share the basic kinetic mechanism (Ca^{2+} inhibition, Mg^{2+} requirement, pH optimum, similar K_m value for SAM and recognition by SCOMT antiserum) [Myöhänen and Männistö, 2010]. As depicted in Table 1, in most human and rat tissues, the levels of SCOMT greatly exceed the levels of MBCOMT, except for the human brain where it only represents 30% of the total COMT [Bonifacio *et al.*, 2007].

An important difference between COMT isoforms is that with exception of catechol estrogens, which possess similar K_m values for both isoforms, all other catechol substrates including catecholamines have lower K_m values for MBCOMT [Roth, 1992]. Interestingly, the substrate concentration seems to influence the kinetic parameters of both isoforms since at low concentrations of catecholamines, *O*-methylation by the low K_m membrane-bound form of COMT would predominate, and only when this enzyme becomes saturated with substrate, does the contribution of the high K_m soluble form of COMT becomes significant [Roth, 1992].

Table 1 – Relative quantification of soluble and MBCOMT proteins in rat and human tissues expressed as % of total COMT in the immunoblot (adapted from [Bonifacio *et al.*, 2007]).

Rat Tissues		
Tissue	**SCOMT**	**MBCOMT**
Liver	93	7
Kidney	75	25
Heart	79	21
Cerebellum	86	14
Telencephalon	69	31
Human Tissues		
Tissue	**SCOMT**	**MBCOMT**
Liver	85	15
Kidney	77	23
Adrenal	74	26
Duodenum	89	11
Brain	30	70
HeLa-Cells	35	65
MCF7-Cells	92	8

In general, the two forms are proposed to have at least partially distinct roles: MBCOMT is believed to be primarily involved in the termination of dopaminergic and noradrenergic synaptic neurotransmission when there are physiologically relevant low concentrations of catecholamines [Myöhänen and Männistö, 2010], i. e., MBCOMT presents higher affinity but lower reaction velocity for catecholamines than SCOMT [Reenilä and Männistö, 2001]. On the other hand,

SCOMT is thought to be mainly responsible for the elimination of biologically active or toxic, particularly exogenous catechols, thus acting as an enzymatic destoxifying barrier between the blood and other tissues [Ma *et al.*, 2014]. Although quite controversial, it was hypothesized that MBCOMT is the predominant enzyme form at dopamine concentrations below 10 μM and at noradrenaline concentrations below 300 μM [Myöhänen and Männistö, 2010; Roth, 1992].

More recently, it was demonstrated that MBCOMT is located in the cell body and in axons and dendrites of rat cortical neurons [Chen *et al.*, 2011]. In addition, the analysis of MBCOMT orientation with computer simulations, flow cytometry and a cell surface enzyme assay revealed that the C-terminal catalytic domain of MBCOMT is in the extracellular space, which suggests that it can inactivate synaptic and extrasynaptic dopamine on the surface of presynaptic and postsynaptic neurons [Chen *et al.*, 2011].

1.3. Catechol-*O*-methyltransferase pathophysiological role in human diseases

During the last decades, COMT has been implicated in diverse human diseases including certain types of cancer [Yager, 2012; Yager, 2015; Wu *et al.*, 2015], cardiovascular diseases [Hall *et al.*, 2014; Voutilainen *et al.*, 2007] or neurologic disorders (Parkinson's disease, schizophrenia) [Apud and Weinberger, 2007; Bonifacio *et al.*, 2007; Philippu *et al.*, 1981], among others.

1.3.1. Catechol-*O*-methyltransferase inhibition in Parkinson's disease

Parkinson's disease (PD) is the most common chronic neurodegenerative disease that affects movement behavior [Müller, 2015] with characteristic symptoms such as tremor, rigidity, bradykinesia and postural instability [Bonifacio *et al.*, 2007]. PD is characterized by dopaminergic neuronal loss in the substantia nigra and by striatal dopamine loss with accumulation of the protein α-synuclein [Müller, 2015]. Initially, the therapy applied to PD was the dopamine replacement using levodopa (L-DOPA) together with an Aromatic Amino Acid Decarboxylase (AADC) inhibitor such as carbidopa [Bonifacio *et al.*, 2007]. However, the efficacy of this therapy decreases over time and most patients develop fluctuating responses and dyskinesias [Bonifacio *et al.*, 2007]. Therefore, the chronic L-DOPA/AADC inhibitor application with concomitant inhibition of COMT and monoamine oxidase (MAO) (inhibitors, such as selegiline and rasagiline) is suggested as standard L-DOPA application in PD patients who need L-DOPA, if they will tolerate it, as depicted in Figure 2 [Bonifacio *et al.*, 2007; Müller 2015].

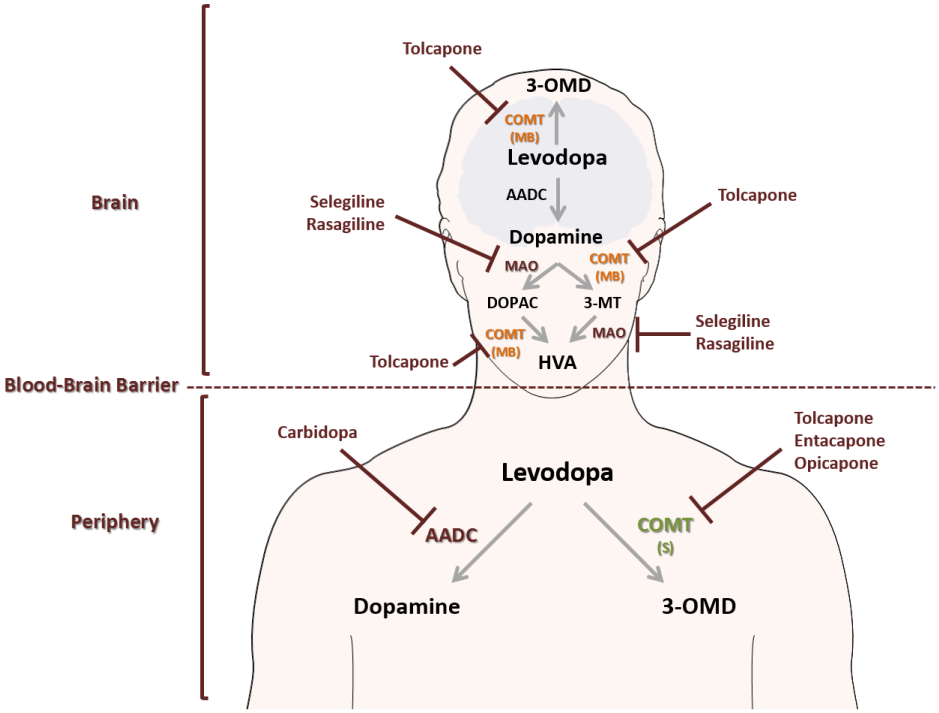


Figure 2 – Sites of action of PD drugs.

Abbreviations: 3-MT, 3-methoxytyramine; 3-OMD, 3-*O*-methyldopa; AADC, aromatic amino acid decarboxylase; COMT, catechol-O-methyltransferase; DOPAC, Dihydroxyphenylacetic acid; MAO, Monoamine oxidase. According to their relative distribution [Bonifácio *et al.*, 2007], in brackets are shown the most abundant COMT isoform in each location (S for SCOMT and MB for MBCOMT).

In fact, COMT inhibitors are well established for the treatment of wearing-off phenomena in PD patients [Müller, 2015]. Recently, it has been described that new pharmacological therapies consisting in using central nervous system-penetrant COMT inhibitors such as tolcapone can also be applied for the treatment of prefrontal cortex dysfunction in schizophrenia patients [Apud and Weinberger, 2007].

The first-generation of COMT inhibitor molecules were typically competitive substrates [Bonifacio *et al.*, 2007]. However, since these molecules were extremely toxic and ineffective *in vivo*, a new class of di-substituted catechols (such as tolcapone, entacapone and nitecapone) was described whose enhanced potency is due to the substitution with electron-withrawing groups at an ortho position to a hydroxyl group of the catechol moiety [Bonifacio *et al.*, 2007]. Typically, these compounds show potency in a low nanomolar range and while tolcapone was described as centrally active, the entacapone and nitecapone were mainly peripheral [Palma *et al.*, 2013]. More recently, the strictly peripheral BIA 9-1067 (Opicapone) inhibitor was developed by the Portuguese Pharmaceutical Company Bial – Portela & C^{a}, S.A. (Portugal) [Kiss *et al.*, 2010]. This compound was found to be extremely potent with a low propensity to cause toxicity, presenting the highest duration of action ever reported, leading to stable and sustained plasma levodopa levels over a 24 h period [Kiss *et al.*, 2010; Palma *et al.*, 2013].

Recently, it was shown that tolcapone induces cell death via the mechanism of apoptosis, and its cytotoxicity is dependent on dosage and correlated with COMT Val/Met genotypes in human lymphoblastoid cells, suggesting that MBCOMT specific inhibitors can be developed and that tolcapone may be less hazardous at low doses and in specific genetic backgrounds [Chen *et al.*, 2011]. It has also been suggested that the central nervous system levels of the endogenous neurotransmitter noradrenaline are decreased in multiple sclerosis patients as well as in its commonly used animal model Experimental Autoimmune Encephalomyelitis (EAE) [Polak *et al.*, 2014]. In fact, mice immunized with myelin oligodendrocyte glycoprotein peptide, the blood-brain barrier (BBB)-permeable COMT inhibitor 3,5-dinitrocatechol (DNC) reduced clinical signs, while entacapone, a non BBB-permeable inhibitor, had no effect [Polak *et al.*, 2014]. These results suggest that centrally active COMT inhibitors could provide benefit to multiple sclerosis patients [Polak *et al.*, 2014].

1.3.2. Catechol-*O*-methyltransferase as a therapeutic protein?

Estrogen catechol metabolites are carcinogenic and it is probable that the mechanism involves both their estrogenicity and oxidative metabolism to genotoxic quinones [Yager, 2012]. As COMT is a gatekeeper phase II enzyme that *O*-methylates the catechols, it blocks their estrogenicity and further metabolism [Yager, 2012]. In fact, several studies reported that COMT can reduce the potential of DNA damage and increase the concentration of 2-methoxy-estradiol, an antiproliferative metabolite in human breast cancer and lung cancer [Lakhani *et al.*, 2003; Michnovicz *et al.*, 1986; Wu *et al.*, 2015], although opposite results were obtained for breast cancer [Bergman-Jungestrom and Wingren, 2001]. Specifically, it has been reported that polymorphisms in COMT that affect its activity and expression have been extensively studied as risk factors for breast cancer [Yager, 2015]. However, as the penetrance of the phenotype encoded by these polymorphisms is low and their impact on breast cancer risk in different populations is not consistent, future investigations into the role of estrogen catechols are required in which samples should be collected for the measurement of specific biomarkers of metabolites hypothesized to be on the causative pathway [Yager, 2012].

Another study conducted with *in vitro* and *in vivo* models of colorectal cancer showed that COMT has tumor-supressive functions [Wu *et al.*, 2015]. Specifically, *in vitro* experiments showed that COMT inhibited cell proliferation by regulating p-Akt, PTEN (phosphatase and tensin homologue) and inhibited G1 to S phase transition by regulating p53, p27 and cyclin D1 and inhibited invasion by regulating E-cadherin [Wu *et al.*, 2015]. On the other hand, *in vivo* studies showed decrease tumor growth in COMT overexpressing cell line [Wu *et al.*, 2015].

1.4. Recombinant human catechol-*O*-methyltransferase biosynthesis and purification

Since the discovery and preliminary characterization of COMT, there was interest in developing biotechnological strategies that allow its biosynthesis and/or isolation for further studies. During the last decades, several eukaryotic and prokaryotic systems including *Escherichia coli* (*E. coli*) reported efficiently the production of both COMT isoforms at different scales with different production levels. Indeed, though it is difficult to perform a proper comparison between the different reports due to the lack of data, this information was gathered and summarized in Table 2. In general, the firstly reported systems for recombinant SCOMT production were important to obtain preliminary biochemical characterizations and kinetic data. However, it was the *E. coli*-based system [Cotton *et al.*, 2004] that allowed the target enzyme production in sufficient amounts for crystallographic studies that ended with the resolution of SCOMT Val/Met variants crystal structures. On the other hand, it was reported that *E. coli* wasn't able to produce the MBCOMT since it was considered to be toxic to the host cells [Lundtröm *et al.*, 1995]. Nevertheless, although enzymatically active MBCOMT was obtained from *E. coli* [Lundtröm *et al.*, 1992] and other prokaryotic and eukaryotic hosts, less work had been performed with this isoform and its structure hasn't been solved yet.

Table 2 – Characteristics and production levels for recombinant soluble and membrane-bound human COMT biosynthesis using different expression systems.

Abbreviations – 108M – Variant methionine; 108V – Variant valine; *E. coli* – *Escherichia coli*; IPTG – Isopropylthiogalactosidase; LB – Lysogeny Broth; NR – Not Reported; Sf9 – *Spodoptera frugiperda*.

SCOMT				
Host strain	**Culture medium/Induction regime**	**Scale**	**Production level**	**Reference**
E. coli SG 13009	LB/IPTG	Shake-flask	180 units/mg of protein (cytosolic fraction)	Malherbe *et al.*, 1992
Sf9 insect cells	TNM-FH + 10% foetal calf serum	1 L spinner flasks	NR	Tilgmann *et al.*, 1992
Human HeLa and Hamster BHK-cells	Minimum essential medium	35 or 100 mm dishes	NR	Ulmanen *et al.*, 1997
E. coli BL21	YT medium/IPTG	8L flask culture	22000 U (1.7 U/mg protein)	Lundtröm *et al.*, 1992
	Defined medium/IPTG	4L Bioreactor/Fed-batch	NR	Tomson *et al.*, 1995
	NR/IPTG	NR	Up to 15 mg/L of culture (Purified)	Cotton *et al.*, 2004
	LB/IPTG	NR	NR	Bai *et al.*, 2007
	Semi-defined medium	3.5 L Bioreactor/Fed-batech	571.8 U/L	Passarinha *et al.*, 2009
	SOB/IPTG	1L Shake flask	132.9 nmol/h	Silva *et al.*, 2012
	Semi-defined/IPTG		183.73 nmol/h	

MBCOMT				
Host strain	**Culture medium/Induction regime**	**Scale**	**Production level**	**Reference**
Human embryonic kidney fibroblast cells (293 cells)	Minimal essential medium + 8% fetal calf serum	100-mm dishes	200 U[1]	Bertocci *et al.*, 1991
	Minimal essential medium + 8% fetal calf serum	NR	33 units/mg protein – cytosolic fraction	Malherbe *et al.*, 1992
			114 units/mg protein – membrane fraction	
E. coli SG 13009	LB/IPTG	Shake-flask	57 units/mg protein (cytosolic fraction)	Malherbe *et al.*, 1992
Sf9 insect cells	TNM-FH + 10% fetal calf serum	1 L spinner flasks	NR	Tilgmann *et al.*, 1992
	NR	NR	NR	Robinson *et al.*, 2012
Human HeLa and Hamster BHK-cells	Minimum essential medium	35 or 100 mm dishes	NR	Ulmanen *et al.*, 1997
E. coli SG BL21	LB/IPTG	Shake-flask	NR	Bai *et al.*, 2007
Brevibacillus choshinensis SP3	Semi-defined (2SY and MT)/Constitutive	Shake-flask (62.5 mL)	50.5 nmol/h/mg of protein	Pedro *et al*, 2011

[1]One unit of COMT activity was defined as the quantity of enzyme producing 1 nmol of guaiacol per hour at 37ºC.

Over the years, from hydrophobic to ionic chromatography, through affinity chromatography, diverse methods were reported for human COMT isolation and purification from several biological and recombinant sources. In Table 3, we review the most important chromatographic methodologies previously reported for recombinant soluble and membrane bound COMT isolation and purification.

Table 3 -Chromatographic methodologies previously reported for recombinant human soluble and membrane-bound COMT isolation and purification.

Abbreviations: AS – Ammonium Sulfate; NR – Not Reported; SC – Sodium Citrate.

SCOMT					
Chromatographic methodology	**Stationary phase**	**Specific activity** (nmol/h/mg of protein)	**Purification factor** (fold)	**Bioactivity recovery** (%)	**Reference**
2 step: affinity chromatography + anion-exchange chromatography	Glutathione-sepharose 4B + HR5/5	NR	NR	NR	Tilgmann *et al.*, 1996
Immobilized-metal affinity chromatography	Talon	Purified 108V ≈ 10.5[1]	NR	NR	Cotton *et al.*, 2004
		Purified 108M ≈ 9.5[1]			
2 step: Immobilized-metal affinity chromatography + Gel Filtration	Talon + XK-16/100 Superdex 75	NR	NR	NR	Rutherford *et al.*, 2008
Hydrophobic Interaction Chromatography	Butyl-sepharose 4FF	1461 ± 30	3.9	23	Passarinha *et al.*, 2007
	Epoxy-sepharose CL-6B	NR	NR	NR	Passarinha *et al.*, 2007
	Octyl-sepharose 6FF 0.025M AS – 0M AS (room temperature)	677 ± 10	1.5	22	Passarinha *et al.*, 2007 Nunes *et al.*, 2010
	Octyl-sepharose 6FF 0.6M AS + 0.025M SC – (5ºC) 0M AS + 0M SC	1688 ± 11	5.9	13	Nunes *et al.*, 2010
	Phenyl-sepharose 6FF	NR	NR	NR	Passarinha *et al.*, 2007
2 step: Hydrophobic Interaction Chromatography + Gel Filtration	Butyl-sepharose 4FF	1688	1.8	14	Passarinha *et al.*, 2008
	Superose™ 12	5500	5.9	1	
Amino acid Affinity Chromatography	L-histidine L-aspartate L-methionine L-leucine L-arginine L-Glutamine	NR	NR	NR	Costa *et al.*, 2011
Anion-exchange chromatography	Q-sepharose	250	3.6	107	Correia *et al.*, 2014
	Q-sepharose HL 16/10 column	11000 U (87 U/mg protein)	51	50	Lundtröm *et al.*, 1992

MBCOMT					
Chromatographic methodology	**Stationary phase**	**Specific activity** (nmol/h/mg of protein)	**Purification factor** (fold)	**Bioactivity recovery** (%)	**Reference**
Immobilized-metal affinity chromatography	His trap FF	NR	NR	NR	Robinson *et al.*, 2012
Hydrophobic Interaction chromatography	Butyl-sepharose 4FF	NR	NR	NR	Santos *et al.*, 2013
	Epoxy-sepharose CL-6B	NR	NR	NR	
	Octyl-sepharose 6FF	NR	NR	NR	
Anion-exchange chromatography	Q-sepharose	496	7	67	Correia *et al.*, 2014
	Q-sepharose (with 0.5% triton X-100)	331	4.3	91	

[1]Enzymatic activities are expressed as μM methylated products formed/15 minutes/μM protein;

1.5. Analytical methods in catechol-*O*-methyltransferase enzymatic assays

As interest in COMT has increased, more reliable, sensitive and rapid analytical COMT assays are needed, either for measuring the activity of recombinant proteins, testing of *in vitro* efficacy of new COMT inhibitor candidates or for the determination of structure-activity relationships [Pihlavisto and Reenilä, 2002]. In a COMT enzymatic assay, three factors should be considered: 1) the detection method, which usually quantifies the reaction product derived from the 2) substrate, which has to be properly metabolized by the 3) enzyme source, the actual sample [Pihlavisto and Reenilä, 2002]. Also, the incubation conditions, including the required supplements, pH of the incubation mixture, temperature, incubation time and concentration of the enzyme during reaction have also to be optimized [Guldberg and Marsden, 1975; Pihlavisto and Reenilä, 2002].

Several analytical methods were described for COMT enzymatic assays. Some of these methods involve the direct quantification of the substrate or the product and don't include a separation step between the *O*-methylated products and the substrate [Pihlavisto and Reenilä 2002] as for example the colorimetric detection of hydroxylamine complexed with ferric chloride in a catecholamine COMT-catalyzed methylation [Abdel-Latif, 1969]. During the last decades, more reliable and sensitive methods were developed. The first methods to separate an *O*-methylated metabolite from the catecholic compound were the liquid-liquid extraction with organic solvents [Pihlavisto and Reenilä, 2002] where the products were posteriorly assessed using fluorometric [Axelrod and Tomchick, 1958] or radiochemical detection [Axelrod and Cohn, 1971]. Although

initially paper or thin-layer chromatography [Axelrod and Tomchick, 1958] were applied as a separation method to identify the *O*-methylated products, more reliable and efficient chromatographic methods were implemented and developed. Indeed, gas chromatography have been applied for example for the separation and quantification of *O*-methylated catechols of 3,4-dihydroxybenzoic acid using flame ionization detection [Koh *et al.*, 1991] or *O*-methylated catechols of catechol estrogens with mass spectrometry [Dawling *et al.*, 2001].

In spite of the methods described above, High-Performance Liquid Chromatography (HPLC) is the most common technique used for COMT activity measurements. The general characteristics as well as information regarding the detection limits and sensitivity of HPLC coupled with different types of detection is presented in Table 4.

Table 4 – HPLC -based analytical methods commonly applied as separation and quantification technique in COMT enzymatic activity assays (adapted from Pihlavisto and Reenilä, 2002).

Liquid chromatography separation techniques in COMT enzymatic assay					
Detection mode	**Characteristics**	**Sensitivity**	**Detection limit** (substrate: 3,4-dihydroxybenzoic acid)	**Substrates**	**Reference**
UV	Poor selectivity	Moderate	1.8 pmol/injection	3,4-dihydroxybenzoic acid	Pennings and Van Kempen, 1979 Brevitt and Tan, 1997
Fluorescence	Metabolites must exhibit native fluorescence or be derivatized with fluorogenic reagents;	Can be high	0.2 pmol/injection	Catecholamines, 3,4-dihydroxybenzoic acid	Okada *et al.*, 1981 Jorga *et al.*, 1998
Radiochemical	Regioisomeric products may be separated from each other; restricted use due to the use of radioactive material;	High	0.45 pmol 3H-labeled; 0.04 pmol 14C-labeled	Dopamine, L-DOPA, 2-hydroxyestradiol	Nissinen, 1985 Lautala *et al.*, 2001
Electrochemical	Can work in amperometric or coulometric mode with different sensitivities;	High	0.5 pmol/injection	catecholamines, L-DOPA, 3,4-dihydroxybenzoic acid	Tuomainen *et al.*, 1996 Passarinha *et al.*, 2006

In particular, HPLC coupled with electrochemical detection are commonly applied to the detection of catechols and phenolic hydroxyls of the *O*-methylated products since these are easily oxidized [Pihlavisto and Reenilä, 2002]. These compounds are separated by reversed-phase HPLC systems with ion-pairing reagents and detected by their reversible oxidation by amperometric [Passarinha *et al.*, 2006] or coulometric working electrodes [Achili *et al.*, 1985;

Sabbioni *et al.*, 2004]. In coulometric detectors, to reduce the noise in the analytical cell, a guard cell is installed prior the analytical cell [Pihlavisto and Reenilä, 2002]. Therefore, by using two analytical cells set at different potentials, interfering compounds may be oxidized or reduced before the cell that is recording, thereby improving the selectivity [Pihlavisto and Reenilä, 2002]. Specifically, in coulometric detection, all of the analyte in the column effluent is oxidized or reduced at the surface of an electrode at constant potential, contrasting with amperometric detection where only part of the analyte is converted [Matson *et al.*, 1984].

1.6. Catechol-*O*-methyltransferase: three-dimensional structure

Actually, nine structures of human SCOMT were deposited in Protein Data Bank, either in the apo form [Ehler *et al.*, 2014] or in complex with several substrates and/or inhibitors [Rutherford *et al.*, 2008]. In Table 5, is presented a short summary of the different SCOMT structures already reported in Protein Data Bank as well as different characteristics of each X-ray experiments.

Table 5 - Summary of the 3D structures previously reported for human SCOMT.

Abbreviations: 43G - 1-(biphenyl-3-yl)-3-hydroxypyridin-4(1H)-one; 43J – 2-(biphenyl-3-yl)-5-hydroxy-3-methylpyrimidin-4(3H)-one; 43H – [1-(biphenyl-3-yl)-5-hydroxy-4-oxo-1,4-dihydropyridin-3-yl]boronic acid; Cl^- - Chloride ion; DNC – 3,5-dinitrocatechol; MES – 2-(N-morpholino)-ethanesulfonic acid; Mg^{2+} - Magnesium ion; MPD – (4S)-2-methyl-2,4-pentanediol; NA – Not available; Na^+ - Sodium ion; SAM – S-adenosyl-$_L$-methionine.

Protein data bank ID	Polymorphic variant	Ligands	X-ray diffraction resolution	Reference
3BWY	Methionine 108	DNC; Mg^{2+}; SAM; MPD	1.30 Å	Rutherford *et al.*, 2008
3BWM	Valine 108	DNC; Mg^{2+}; SAM; MPD	1.98 Å	
3A7E	NA	DNC; Mg^{2+}; SAM	2.80 Å	Tsuji E. 2010 Not Published
4PYK	Valine 108	Cl^-; Mg^{2+}; Na^+	2.22 Å	Ehler *et al.*, 2014
4PYJ	Valine 108	Cl^-; Mg^{2+}; Na^+	1.90 Å	
4PYI	Valine 108	Na^+	1.35 Å	
4XUE	Valine 108	Mg^{2+}; 43J; SAM	2.30 Å	Harrison *et al.*, 2015
4XUD	Valine 108	Mg^{2+}; 43H; MES; SAM	2.4 0Å	
4XUC	Valine 108	Mg^{2+}; 43G; MES; SAM	1.80 Å	

Briefly, the structure of SCOMT is one of two sets of α-helices (α1- α5) sandwiching a seven-stranded β-sheet core (arranged in an order of 3214576) [Rutherford *et al.*, 2008; Ma *et al.*, 2014]. In the β-sheet, strand 7 is antiparallel to the others [Ma *et al.*, 2014].

The SAM binding pocket has a base formed by the C-terminal ends of β–sheets a-d with the walls of the pocket formed by helix 6 and the loops between helix 2 and helix 3, β-sheet a and helix 4, and β-sheet d and helix 7 [Thomas and Potter, 2013]. The SAM methyl group is directed toward the substrate binding site and the catechol oxygen to be methylated [Thomas and Potter, 2013]. The methyl group attached to the methionine sulfur atom in SAM is oriented toward the substrate binding site and specifically towards the catechol oxygen atom to be methylated [Ma *et al.*, 2014]. The substrate-binding site is a shallow pocket defined by M40, L198, W143 and "gatekeeper" residues W38 and P174 [Ma *et al.*, 2014]. All these residues are hydrophobic, suggesting that Van der Waals contacts are the main forces for ligand binding [Ma *et al.*, 2014]. The magnesium ion is held in place by coordination with three acidic residues, D141, D169 and E199 the latter two of which are oriented by interaction with K46 [Thomas and Potter, 2013]. The oxygen acting as the methyl recipient is activated by K44 which acts as a general base in the methylation [Thomas and Potter, 2013]. In a direct bimolecular transfer the SAM methyl group is transferred from the Sulphur to the catechol hydroxyl oxygen through an S_N2-like transition state [Thomas and Potter, 2013].

1.7. References

Abdel-Latif AA. Reaction of catecholamines with hydroxylamine and its application to the assay of catechol-*O*-methyltransferase. Anal Biochem 1969; 29 (3): 468-75.

Achili G, Perego C, Ponzio F. Application of the dual-cell coulometric detector: a method for assaying monoamines and their metabolites. Anal Biochem 1985; 148 (1): 1-9.

Apud JA, Weinberger DR. Treatment of cognitive deficits associated with schizophrenia: potential role of catechol-*O*-methyltransferase inhibitors. CNS Drugs 2007; 21 (7): 535-57.

Axelrod J, Cohn CK. Methyltransferase enzymes in red blood cells. J Pharmacol Exp Ther 1971; 176 (3): 650-54.

Axelrod J, Tomchick R. Enzymatic *O*-methylation of epinephrine and other catechols. J Biol Chem 1958; 233 (3): 702-5.

Bai HW, Shim JY, Yu J, Zhu BT. Biochemical and molecular modeling studies of the *O*-methylation of various endogenous and exogenous catechol substrates catalyzed by recombinant human soluble and membrane-bound catechol-*O*-methyltransferase. Chem Res Toxicol 2007; 20: 1409-25.

Bergman-Jungestrom M, Wingren S. Catechol-*O*-methyltransferase (COMT) gene polymorphism and breast cancer risk in young women. Br J Cancer 2001; 85: 859-62.

Bertocci B, Miggiano V, Da Prada M, Dembic Z, Lahm HW, Malherbe P. Human catechol-*O*-methyltransferase: Cloning and expression of the membrane-associated form. PNAS 1991; 88: 1416-20.

Bonifacio MJ, Palma PN, Almeida L, Soares-da-Silva P. Catechol-*O*-methyltransferase and its inhibitors in Parkinson's disease. CNS Drug Rev 2007; 13 (3): 252-79.

Brevitt SE, Tan EW. Synthesis and *in vitro* evaluation of two progressive series of bifunctional polyhydroxybenzamide catechol-*O*-methyltransferase inhibitors. J Med Chem 1997; 40 (13): 2035-9.

Chen J, Song J, Yuan P, Tian Q, Ji Y, Ren-Patterson R, Liu G, Sei Y, Weinberger DR. Orientation and cellular distribution of membrane-bound catechol-*O*-methyltransferase in cortical neurons. Implications for drug development. J Biol Chem 2011; 286 (40): 34752-60.

Correia FF, Santos FM, Pedro AQ, Bonifacio MJ, Queiroz JA, Passarinha LA. Recovery of biological active catechol-*O*-methyltransferase isoforms from Q-sepharose. J Sep Sci 2014; 37: 20-9.

Costa SR, Bonifacio MJ, Queiroz JA, Passarinha LA. Analysis of hSCOMT adsorption in bioaffinity chromatography with immobilized amino acids: The influence of pH and ionic strength. J Chromatogr B Analyt Technol Biomed Life Sci 2011; 879: 1704-06.

Cotton NJH, Stoddard B, Parson WW. Oxidative inhibition of human soluble catechol-*O*-methyltransferase. J Biol Chem 2004; 279 (22): 23710-18.

Dawling S, Roodi N, Mernaugh RL, Wang X, Parl FF. Catechol-*O*-methyltransferase (COMT)-mediated metabolism of catechol estrogens: comparison of wild-type and variant COMT isoforms. Cancer Res 2001; 15 (61): 6716-22.

Ehler A, Benz J, Schlatter D, Rudolph MG. Mapping the conformational space accessible to catechol-*O*-methyltransferase. Acta Crystallogr D Biol Crystallogr 2014; 70 (Pt 8): 2163-74.

Guldberg HC, Marsden CA. Catechol-*O*-methyltransferase: pharmacological aspects and physiological role. Pharmacol Rev 1975; 27 (2): 135-206.

Hall KT, Nelson CP, Davis RB, Buring JE, Kirsch I, Mittleman MA, Loscalzo J, Samani NJ, Ridker PM, Kaptchuk TJ, Chasman DI. Polymorphisms in Catechol-*O*-methyltransferase modify treatment effects of aspirin on risk of cardiovascular disease. Arterioscler Thromb Vasc Biol 2014; 34: 2160-67.

Harrison ST, Poslusney MS, Mulhearn JJ, Zhao Z, Kett NR, Schubert JW, Melamed JY, Allison TJ, Patel SB, Sanders JM, Sharma S, Smith RF, Hall DL, Robinson RG, Sachs NA, Hutson PH, Wolkenberg SE, Barrow JC. Synthesis and evaluation of heterocyclic catechol mimics as inhibitors of catechol-*O*-methyltransferase (COMT). ACS Med Chem Lett 2015; 6 (3): 318-23.

Hosak L. Role of the COMT gene Val158Met polymorphism in mental disorders: a review. Eur Psychiatry 2007; 22 (5): 276-81.

Jorga KM, Fotteler B, Heizmann P, Zürcher G. Pharmacokinetics and pharmacodynamics after oral and intravenous administration of tolcapone, a novel adjunct to Parkinson's disease therapy. Eur J Clin Pharmacol 1998; 54 (5): 443-7.

Kiss LE, Ferreira HS, Torrão L, Bonifacio MJ, Palma PN, Soares-da-Silva P, Learmonth DA. Discovery of a long-acting, peripherally selective inhibitor of catechol-*O*-methyltransferase. J Med Chem 2010; 53 (8): 3396-411.

Koh S, Uravama K, Kawai S, Takayama Y. Application of a metal capillary column in gas chromatographic determination of catechol-*O*-methyltransferase activity. J Chromatogr 1991; 19 (549): 434-39.

Lakhani NJ, Sarkar MA, Venitz J, Figg WD. 2-Methoxyestradiol, a promising anticancer agent. Pharmacotherapy 2003; 23 (2): 165-72.

Lautala P, Ulmanen I, Taskinen J. Molecular mechanisms controlling the rate and specificity of catechol *O*-methylation by human soluble catechol-*O*-methyltransferase. Mol Pharmacol 2001; 59 (2): 393-402.

Lundtröm K, Tilgmann C, Peränen J, Kalkinnen N, Ulmanen I. Expression of enzymatically active rat liver and human placental catechol-*O*-methyltransferase in *Escherichia coli*; purification and partial characterization of the enzyme. Biochim Biophys Acta 1992; 1129: 149-54.

Lundtröm K, Tenhunen J, Tilgmann C, Karhunen T, Panula P, Ulmanen I. Cloning, expression and structure of catechol-*O*-methyltransferase. Biochim Biophys Acta 1995; 1251 (1): 1-10.

Ma Z, Liu H, Wu B. Structure-based drug design of catechol-*O*-methyltransferase inhibitors for CNS inhibitors. Brit J Clin Pharmacol 2014; 77 (3): 410-20.

Malherbe P, Bertocci B, Caspers P, Zürcher G, Da Prada M. Expression of functional membrane-bound and soluble catechol-*O*-methyltransferase in *Escherichia coli* and a mammalian cell line.J Neurochem 1992; 58 (5): 1782-89.

Matson WR, Langlais P, Volicer L, Gamache PH, Bird E, Mark KA. N-Electrode three-dimensional liquid-chromatography with electrochemical detection for determination of neurotransmitters. Clin Chem 1984; 30 (9): 1477-88.

Michnovicz JJ, Hershcopf RJ, Naganuma H, Bradlow HL, Fishman J. Increased 2-hydroxylation of estradiol as a possible mechanism for the anti-estrogenic effect of cigarette smoking. N Engl J Med 1986; 315 (21): 1305-9.

Müller T. Catechol-*O*-methyltransferase inhibitors in Parkinson's disease. Drugs 2015; 75 (2): 157-74.

Myöhänen TT, Männistö PT. Distribution and functions of catechol-*O*-methyltransferase proteins: do recent findings change the picture? Int Rev Neurobiol 2010; 95: 29-47.

Nissinen E. Determination of catechol-*O*-methyltransferase activity in brain tissue by high-performance liquid chromatography with on-line radiochemical detection. Anal Biochem 1985; 144 (1): 247-52.

Nunes VS, Bonifacio MJ, Queiroz JA, Passarinha LA. Assessment of COMT isolation by HIC using a dual system and low temperature. Biomed Chromatogr 2010; 24 (8): 858-62.

Okada Y, Zaitsu K, Ohtsubo K, Nohta H, Ohkura Y. Fluorimetric assay for catechol-O-methyltransferase. Chem Pharm Bull (Tokyo) 1981; 29 (6): 1670-3.

Palma PN, Kiss LE, Soares-da-Silva P. Catechol-*O*-methyltransferase inhibitors: Present problems and relevance of the new ones. Emerging Drugs and Targets for Parkinson's disease. Cambridge, The Royal Society of Chemistry 2013; 4: 83-109.

Passarinha LA, Bonifacio MJ, Queiroz JA. The effect of temperature on the analysis of metanephrine for catechol-*O*-methyltransferase activity assay by HPLC with electrochemical detection. Biomed Chromatogr 2006; 20 (9): 937-44.

Passarinha LA, Bonifacio MJ, Queiroz JA. Comparative study on the interaction of recombinant human soluble catechol-*O*-methyltransferase on some hydrophobic adsorbents. Biomed Chromatogr 2007; 21: 430-38.

Passarinha LA, Bonifacio MJ, Soares-da-Silva P, Queiroz JA. A new approach on the purification of recombinant human soluble catechol-*O*-methyltransferase from an *Escherichia coli* extract using hydrophobic interaction chromatography. J Chromatogr A 2008; 1177: 287-96.

Passarinha LA, Bonifacio MJ, Queiroz JA. Application of a fed-batch bioprocess for the heterologous production of hSCOMT in *Escherichia coli*. J Microbiol Biotechnol 2009; 19 (9): 972-81.

Pedro AQ, Bonifacio MJ, Queiroz JA, Maia CJ, Passarinha LA. A novel prokaryotic expression system for biosynthesis of recombinant human membrane-bound catechol-*O*-methyltransferase. J Biotechnol 2011; 156: 141-6.

Pennings EJ, Van Kempen GM. Assay of catechol-*O*-methyltransferase by determination of the *m*- and *o*-*O*-methylated products using high-performance liquid chromatography. Anal Biochem 1979; 98 (2): 452-4.

Philippu G, Hoo JJ, Milech U, Argarwall DP, Schrappe O, Goedde HW. Catechol-*O*-methyltransferase of erythrocytes in patients with endogenous psychoses. Psychiatry Res 1981; 4 (2): 139-46.

Pihlavisto P, Reenilä I. Separation methods for catechol-*O*-methyltransferase activity assay: physiological and pathophysiological relevance. J Chromatogr B Analyt Technol Biomed Life Sci 2002; 781 (1-2): 359-72.

Polak PE, Lin SX, Pelligrino D, Feinstein DL. The blood-brain barrier-permeable catechol-*O*-methyltransferase inhibitor dinitrocatechol suppresses experimental autoimmune encephalomyelitis. J Neuroimmunol 2014; 276: 135-41.

Reenilä I, Männistö PT. Catecholamine metabolism in the brain by membrane-bound and soluble catechol-*O*-methyltransferase (COMT) estimated by enzyme kinetic values. Med Hypotheses 2001; 57 (5): 628-32.

Robinson RG, Smith SM, Wolkenberg SE, Kandebo M, Yao L, Gibson CR, Harrison ST, Polsky-Fisher S, Barrow JC, Manley PJ, Mulhearn JJ, Nanda KK, Schubert JW, Trotter BW, Zhao Z, Sanders JM, Smith RF, McLoughlin D, Sharma S, Hall DL, Walker TL, Kershner JL, Bhandari N, Hutson PH, Sachs NA. Characterization of non-nitrocatechol pan and isoform specific catechol-*O*-methyltransferase inhibitors and substrates. ACS Chem Neurosci 2012; 3: 129-40.

Roth JA. Membrane-bound catechol-*O*-methyltransferase: a reevaluation of its role in the O-methylation of the catecholamine neurotransmitters. Rev Physiol Biochem Pharmacol 1992; 120: 1-29.

Rutherford K, Trong IL, Stenkamp RE, Parson WW. Crystal structures of human 108V and 108M catechol-*O*-methyltransferase. J Mol Biol 2008; 380: 120-30.

Sabbioni C, Saracino MA, Mandrioli R, Pinzauti S, Furlanetto S, Gerra G, Raggi MA. Simultaneous liquid chromatographic analysis of catecholamines and 4-hydroxy-3-methoxyphenylethylene glycol in human plasma: Comparison of amperometric and coulometric detection. J Chromatogr A 2004; 1032 (1-2): 65-71.

Santos FM, Pedro AQ, Soares RF, Martins R, Bonifácio MJ, Queiroz JA, Passarinha LA. Performance of hydrophobic interaction ligands for human membrane-bound catechol-*O*-methyltransferase purification. J Sep Sci 2013; 36: 1693-1702.

Silva R, Ferreira S, Bonifacio MJ, Dias JML, Queiroz JA, Passarinha LA. Optimization of fermentation conditions for the production of human soluble catechol-*O*-methyltransferase by *Escherichia coli* using artificial neural network. J Biotechnol 2012; 160: 161-68.

Thomas MP, Potter BVL. The structural biology of oestrogen metabolism. J Steroid Biochem Mol Biol 2013; 137: 27-49.

Tilgmann C, Melen K, Lundstrom K, Jalanko A, Julkunen I, Kalkinnen N, Ulmanen I. Expression of recombinant soluble and membrane-bound catechol-*O*-methyltransferase in eukaryotic cells and identification of the respective enzymes in rat brain. Eur J Biochem 1992; 207: 813-21.

Tilgmann C, Ulmanen I. Purification methods of mammalian catechol-*O*-methyltransferases. J Chromatogr B 1996; 684: 147-61.

Tomson K, Paalme T, Laakso PS, Vilu R. Automatic laboratory-scale fed-batch procedure for production of recombinant proteins using inducible expression systems of *Escherichia coli*. Biotechnol Tech 1995; 9 (11): 793-8.

Tunbridge EM, Harrison PJ, Warden DR, Johnston C, Refsum H, Smith AD. Polymorphisms in the catechol-*O*-methyltransferase (COMT) gene influence plasma total homocysteine levels. Am J Med Genet B Neuropsychiatr Genet 2008; 147 B (6): 996-99.

Tuomainen P, Reenillä I, Mannistö PT. Validation of assay of catechol-*O*-methyltransferase activity in human erythrocytes. J Pharm Biomed Anal 1996; 14 (5): 515-23

Ulmanen I, Peränen J, Tenhunen J, Tilgmann C, Karhunen T, Panula P, Bernasconi L, Aubry J, Lundtröm K. Expression and intracellular localization of catechol-*O*-methyltransferase in transfected mammalian cells. Eur J Biochem 1997; 243: 452-59.

Voutilainen S, Tuomainen T, Korhonen M, Mursu J, Virtanen JK, Happonen P, Alfthan G, Erlund I, North KE, Mosher MJ, Kauhanen J, Tiihonen J, Kaplan GA, Salonen JT. Functional COMT Val158Met polymorphism, risk of Acute Coronary events and serum homocysteine: The Kuopio Ischaemic heart disease risk factor study, PLoS ONE 2007; 1: 181-87.

Weisz J, Fritz-Wolz G, Gestl S, Clawson GA, Crevelling CR, Liehr JG, Dabbs D. Nuclear localization of catechol-*O*-methyltransferase in neoplastic and noneoplastic mammary epithelial cells. Am J Pathol 2000; 156 (6): 1841-48.

Wu W, Wu Q, Hong X, Xiong G, Xiao Y, Zhou J, Wang W, Wu H, Zhou L, Song W, Dai H, Qiu H, Zhao Y. Catechol-*O*-methyltransferase inhibits colorectal cancer cell proliferation and invasion. Arch Med Res 2015; 46 (1): 17-23.

Yager JD. Catechol-*O*-methyltransferase: characteristics, polymorphisms and role in breast cancer. Drug Discov Today Dis Mech 2012; 9 (1-2): 41-46.

Yager JD. Mechanisms of estrogen carcinogenesis: The role of E2/E1 – Quinone metabolites suggests new approaches to preventive intervention – A review. Steroids 2015; 99 (Pt A): 56-60.

Paper I

2. From gene to structure: biosynthesis strategies to break the bottlenecks associated to membrane protein structure determination

A. Q. Pedro, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

(*Submitted for publication*)

Short description: This paper is a review article that summarizes the most recent and important achievements concerning different strategies previously reported that lead to an increase in the production yield of recombinant MP. Initially, a comparative overview of the MP biogenesis in eukaryotic and prokaryotic organisms is presented. Then, genetic and molecular biology based tools that are commonly applied to increase MP expression are also discussed as well as general guidelines to choose the right host. Afterwards, we summarize specific strategies that have been used to increase MP biosynthesis in different hosts such as *E. coli*, *Pichia pastoris* (*P. pastoris*) and mammalian cell lines and it is also addressed MP expression in preformed lipid-bilayers using cell-free expression systems. Finally, a case study on optimizing MBCOMT in a biologically active form is presented and discussed.

From gene to structure: Biosynthesis strategies to break the bottlenecks associated to membrane protein structure determination

Augusto Q Pedro[1], Maria J. Bonifacio[2], João A Queiroz[1], Luís A Passarinha[1*]

[1]*CICS-UBI – Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, 6201-001 Covilhã, Portugal*

[2]*Bial – Departamento de Investigação e Desenvolvimento, 4745-457 São Mamede do Coronado, Portugal*

*Corresponding author – E-mail: lpassarinha@fcsaude.ubi.pt - Phone: +351 275 329 069; Health Sciences Research Centre (CICS-UBI); Faculty of Health Sciences, University of Beira Interior, Av. Infante D. Henrique, 6200-506, Covilhã, Portugal

Outline

Abstract

Membrane proteins constitute 20-30% of all proteins encoded by the genome of various organisms and perform a wide range of essential biological functions. However, despite they represent the largest class of protein drug targets, few high-resolution 3D structures have been obtained yet. Membrane protein biogenesis is more complex than that of the soluble proteins and its recombinant biosynthesis has been a major drawback, thus blocking their further structural characterization. Indeed, the major limitation in structure determination of MP is the low yield achieved in recombinant expression, usually coupled to low functionality, pinpointing the optimization target in recombinant MP research. Recently, many efforts that have been put into the upstream stage of membrane protein bioprocesses began to succeed, allowing the growth of the number of membrane protein solved structures. Therefore, this review covers the most recent solutions and technical advances for common bottlenecks encountered in membrane protein expression using prokaryotic and eukaryotic organisms. Also, to guide researchers on their own MP targets, we also highlight skillful strategies to effectively increase the expression and/or folding of specific MP but readily adaptable to other targets.

Keywords: Production; Membrane Protein; Host; Biogenesis; Codon usage; Optimization; Cell-free expression; catechol-*O*-methyltransferase.

List of abbreviations: AQP - Aquaporin; ß-OG - n-octyl-ß-$_D$-glucopyranoside; BiP – Binding Protein; CF – Cell-free; CHAPS - 3-[(3-Cholamidopropyl)dimethylammonio]-1-propansulfonate; CHO - Chinese Hamster Ovary; COMT – catechol-*O*-methyltransferase; D-CF – Detergent-based cell-free; DPC - Dodecylphosphocholine; *E. coli* – *Escherichia coli*; ER - Endoplasmic Reticulum ; GRAS – Generally recognized as safe; IMAC - Immobilized-Metal Affinity Chromatography; JCat - Java Codon Adaptation Tool; L-CF – Lipid-based cell-free; LPPG - Lysopalmitoyl phosphatidylglycerol; LPS – Lipopolysaccharide; MBCOMT – Membrane-bound catechol-*O*-methyltransferase; mRNA – messenger Ribonucleic acid; MP – Membrane Protein; P-CF – Precipitate cell-free; PCN – Plasmid Copy Number; *P. pastoris* - *Pichia pastoris*; PTM – Post-translational modifications; qPCR – quantitative Polymerase Chain Reaction; RNC – Ribosome Nascent Chain; SAM – S-adenosyl-$_L$-methionine; SCOMT – Soluble catechol-*O*-methyltransferase; SDS - Sodium Dodecyl Sulfate; Sec – Secretory; SERT - Serotonin transporter; SRP – Signal recognition Particle; tRNA – transfer Ribonucleic Acid.

1. Introduction

Membrane proteins (MP) constitute 20-30% of all proteins encoded by the genome of various organisms [Pedro *et al.*, 2011] and perform a wide range of essential biological functions, thus, representing the largest class of protein drug targets [Bernaudat *et al.*, 2011]. A vast majority of these MP still have no assigned function and only a few unique high-resolution 3D structures have been obtained for transmembrane proteins so far [Bernaudat *et al.*, 2011]. Up until now, 555 unique MP structures from different species have been deposited in the MP of Known Structure Database (http://blanco.biomol.uci.edu/mpstruct/), whereas 2443 nonunique structures of MP (2139 alpha- and 304 beta folded) are listed in the MP Data Bank (http://pdbtm.enzim.hu/) [Kozma *et al.*, 2013; Lantez *et al.*, 2015]. The increasing understanding of MP insertion, folding and degradation means that MP overexpression can be more rationalized, both at the level of the host and the synthesized MP [Henrich *et al.*, 2015]. However, MP are very diverse in their structure and physico-chemical properties and thus it is unfeasible predict whether a protein of interest will express well, be easy to purify, be active or crystallize in any given experimental flowsheet [Bernaudat *et al.* 2011].

Indeed, challenges in MP structure determination come from their low natural abundance, their frequent toxicity when overexpressed in heterologous expression systems, the difficulty in purifying stable functional proteins and, finally, the difficulty in obtaining well-diffracting crystals [Massey-Gendel *et al.*, 2009; Gul *et al.*, 2014; Lantez *et al.*, 2015]. The technical challenges posed by the hydrophobic nature coupled with the low natural abundance and intrinsic instability of MP makes it difficult obtaining sufficient amounts of MP for functional and structural studies [Bernaudat *et al.*, 2011]. As it is impractical to purify MP from the poorly abundance natural sources, recombinant heterologous expression systems with the development of protein synthesis technologies have circumvented this problem and different hosts are usually applied, namely *Escherichia coli* (*E. coli*), *Pichia pastoris* (*P. pastoris*), mammalian or insect cells, among others [Zheng *et al.*, 2014]. However, the production of MP in cellular systems is besieged by several problems due to their hydrophobic nature which often causes misfolding, protein aggregation and cytotoxicity, resulting in poor yields of stable proteins [Rajesh *et al.*, 2011], pinpointing the optimization target in the field of MP structural biology. Protein overexpression involves the synergy of three key elements: a gene, a vector and an expression host [Bernaudat *et al.*, 2011]. Therefore, here we review the most recent and important achievements made in the upstream stage for increasing the recombinant expression of correctly folded and functionally active recombinant MP using prokaryotic or eukaryotic hosts. The MP biogenesis will be briefly scrutinized and generic guidelines to aid in choosing the right host that meets specific needs will be addressed. Also, it will be discussed the role that genetics and molecular biology tools play in increasing the expression levels of active MP and the applicability of Cell-Free (CF) expression systems for MP production in pre-formed lipid-bilayers. Finally, a

case-study on the "difficult to express" membrane-bound catechol-*O*-methyltransferase (MBCOMT) will be discussed.

2. Membrane protein biosynthesis for structural studies

Membrane proteins are very diverse in their structure and physic-chemical properties and thus it is unfeasible predict whether a protein of interest will express well, be corrected-inserted in the host system membranes, be easy to purify, be active or crystallize by a specific experimental flowsheet [Bernaudat *et al.*, 2011]. However, specific strategies (optimization of the codon usage, strain engineering adapted to MP production, addition of chemical or co-expression with biological chaperones, among others) often reported to a single MP may have a broad applicability and its use may be extended to other MP. Therefore, according to Figure 1, the main aim of this review is to gather information regarding the successful optimization of the biosynthesis of recombinant MP, discussing its general applicability and provide a practical guide with general strategies for the optimization of the upstream stage.

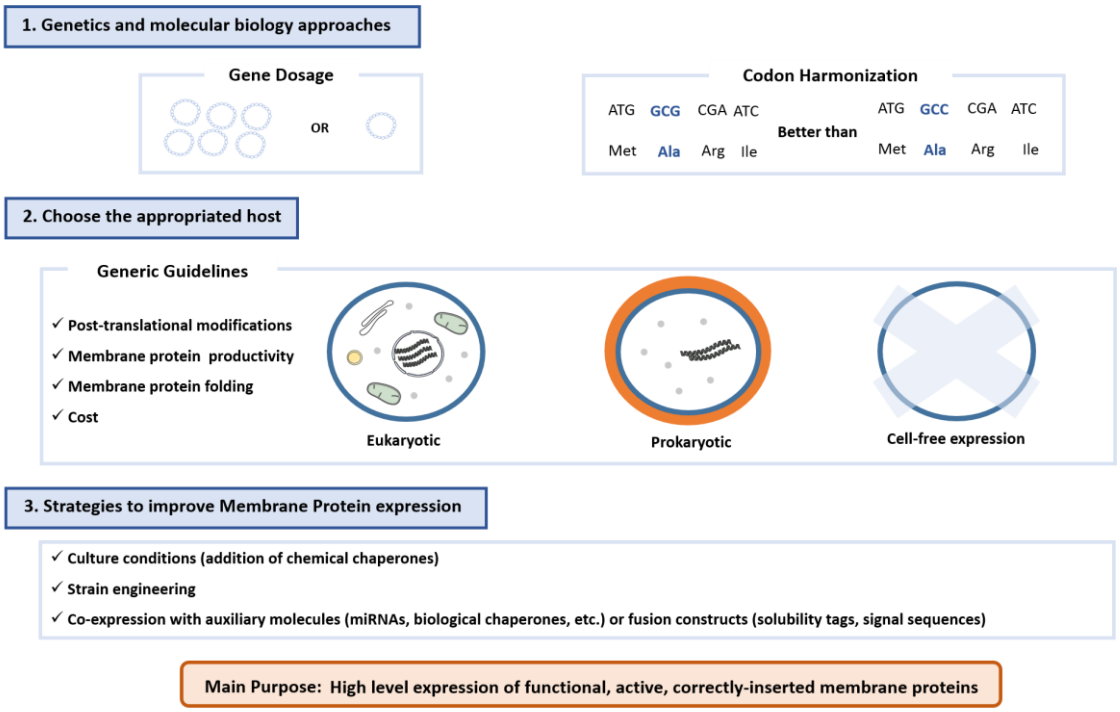


Figure 1 - Overview of the topics included in this review amenable to optimization and, thus, relevant for obtaining a successful strategy for recombinant MP biosynthesis.

2.1 Membrane protein biogenesis: eukaryotic vs prokaryotic organisms

The MP biogenesis is more complex than water-soluble proteins and requires machinery to target the proteins and to insert them into the membrane [Marredy *et al.*, 2011]. Recombinant MP biosynthesis and its biogenesis are closely associated once MP must be produced in a folded form, correctly inserted in the membranes. In order to elucidate this, a simplified overview and a schematic representation of the main paths involved in membrane protein targeting and insertion into prokaryotic (cytoplasmic) or eukaryotic (peroxisomal, mitochondrial and endoplasmic reticulum) membranes are presented in Figures 2 and 3. In bacteria, MP are overexpressed in the cytoplasmic membrane whereas in eukaryotes they are typically overexpressed in the endoplasmic reticulum (ER) membrane, from where they are transported to the plasma membrane in vesicles [Freigassner *et al.*, 2009; Wagner *et al.*, 2006].

In general, the following steps are crucial for a successful translocation and integration of proteins into the membrane in a functional form: 1) identification of the protein to be translocated; 2) discrimination between the protein to be translocated into the membrane and proteins to be secreted; 3) integration and exportation of the proteins into or across the lipid bilayer; and 4) functional folding of the membrane integrated protein without disrupting the membrane integrity [Marredy *et al.*, 2011]. Overall, heterologous overexpression of MP can be hampered by different synthesis, targeting, insertion and folding characteristics in host [Wagner *et al.*, 2006]. As for instance, some of the charged residues in the yeast Sec61 translocon, important for proper functioning, are not conserved in the prokaryotic Sec translocon, what can possible cause problems in the heterologous expression of MP [Goder *et al.*, 2004; Wagner *et al.*, 2006]. The different protein folding capacities, glycosylation patterns and membranes lipid composition between the organism from which the target membrane protein originates and the heterologous system are also important factors that affect the overexpression of MP [Freigassner *et al.*, 2009; Wagner *et al.*, 2006].

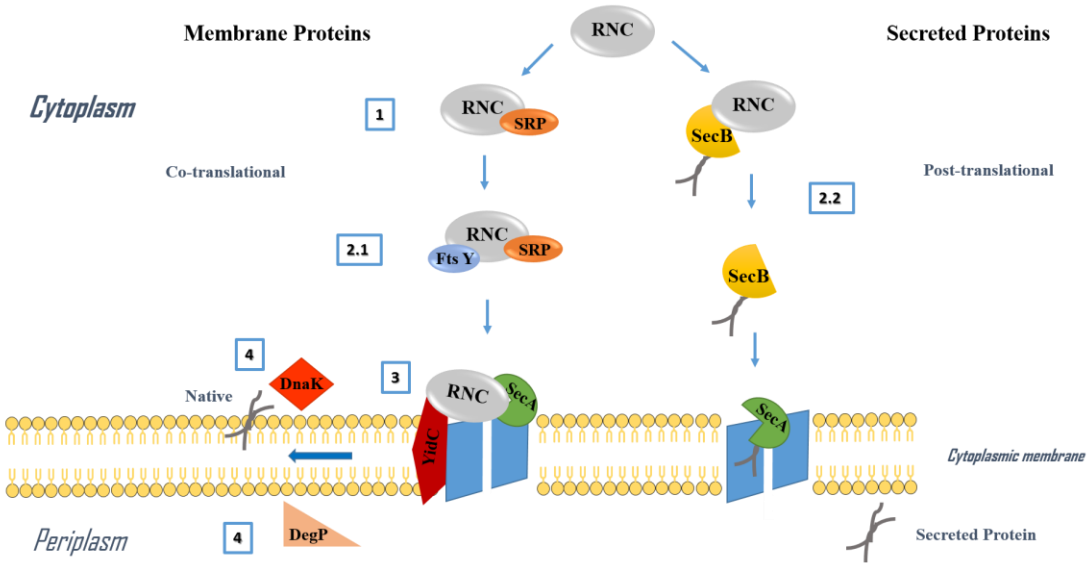


Figure 2 - Simplified overview of membrane protein biogenesis in prokaryotic organisms (adapted from [Freigassner *et al.*, 2009; Marredy *et al.*, 2011; McMorran *et al.*, 2014; Wagner *et al.*, 2006]).

1 - Signal Recognition Particle (SRP) identifies the Membrane Proteins (MP) that need to be inserted in the membrane via the Sec pathway; **2** - Ribosome-membrane protein nascent chain complexes (RNCs) are recognized by SRP; the transfer of a nascent polypeptide into the Sec pore requires energy and is mediated by SecA ATPase; these complexes are targeted **(2.1)** in a co-translational manner to the bacterial cytoplasmic membrane until it binds to the membrane receptor FtsY (for integral MP) or **(2.2)** the translocation may be post-translational (for secreted or outer MP that contain a positively charged N-terminal sequence) in which proteins are transported to the membrane once their synthesis is complete and are kept unfolded by SecB; **3** - YidC has been proposed to mediate the transfer of transmembrane segments from the Sec translocon into the lipid bilayer; YidC may also play a role in MP folding; **4** - DnaK and DegP are cytoplasmic and periplasmic chaperones, respectively. In particular, DnaK may be involved in targetting and folding of specific MP and it has been described that is upregulated in response to inclusion body formation. On the other hand, DegP can act as a chaperone or protease and the switching between these two activities seems to be dependent on structural changes promoted by changes in temperature. The fate of DegP-bound substrates lies in their ability to adopt their native structure within the cavity, as only unfolded substrates can be degraded.

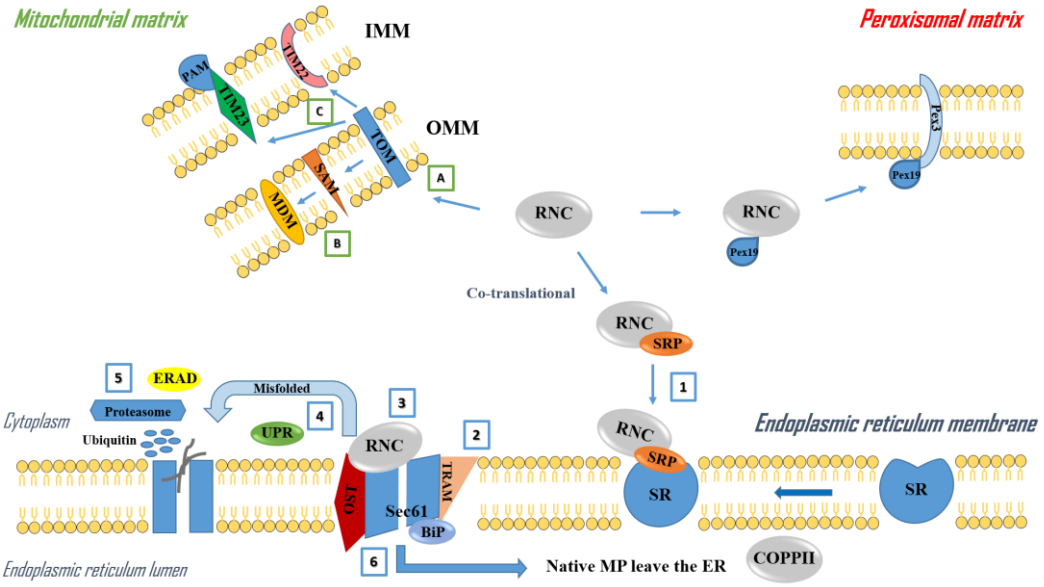


Figure 3 - Simplified overview of membrane protein biogenesis in eukaryotic mitochondrial, peroxisomal and reticulum endoplasmic membranes (adapted from [Becker *et al.*, 2008; Freigassner *et al.*, 2009; Marredy *et al.*, 2011; Wagner *et al.*, 2006]).

Mitochondria: A - Mitochondrial Membrane proteins (MP) pass through the outer mitochondrial membrane (OMM) via the translocase of the outer membrane (TOM) complex; **B** - Proteins of the OMM are transported to the SAM (Sorting and Assembly Machinery) and MDM (Mitochondrial Distribution and Morphology) complexes; **C** - Proteins are targeted to the inner mitochondrial membrane (IMM) by the Translocase of the Inner Mitochondrial membrane complexes TIM22 and TIM23.

Peroxissome: Peroxisomal MP use the Pex19/Pex3-mediated pathway (class I proteins).

Endoplasmic reticulum: 1 - Ribosome-membrane protein nascent chain complexes (RNC) are targeted in a co-translational fashion to the endoplasmic reticulum (ER) membrane via the SRP pathway (comprises the signal recognition particle and its receptor, SR) that unlike *Escherichia coli* SRP, can pause translation; **2** - RNC docks at the Sec61 translocon and after dissociation of SRP from RNC, translation resumes; **3** - Hsp70 chaperone Binding Protein (BiP) supports the translocation of polypeptides into the ER and folding of luminal protein domains; TRAM might assist the transfer of transmembrane segments from the Sec61 translocon into the lipid bilayer and MP folding; then, proteins can be glycosylated by oligossacharyltransferase (OST); **4** - The increased chaperone requirement often triggers the Unfolded Protein Response (UPR), an intracellular signaling pathway that acts to relieve stress phenomena's; **5** - Proteins that fail to reach their native fold are targeted to the proteasome for degradation, through dislocation to the cytoplasm and protein ubiquitination - ERAD process; **6** - Native MP leave the ER from exit sites by COPII-mediated trafficking.

2.2 Genetics and molecular biology: tools to improve membrane protein expression

Gathering structural information about MP targets requires higher quantities of properly folded and correctly inserted MP. To accomplish this purpose, scientists often drive their optimization efforts to the upstream stage where extensive optimization procedures are usually carried out with success. However, prior the upstream stage, optimization experiments on a genetic level can be carried out and promote an increase of more than fivefold in the yield of the desired MP [Nordén *et al.*, 2011]. If on one hand, the genetic code itself can be optimized to suit the selected host for expression, sometimes a higher copy number of the target gene may increase the levels of the target MP obtained in particular hosts such as *P. pastoris*.

2.2.1 Handling the target gene copy number

Plasmid-based expression is preferred in prokaryotic organisms once the gene dose, which depends on Plasmid Copy Number (PCN), is higher than when the recombinant gene is integrated into the host's chromosome [Palomares *et al.*, 2004]. In these systems, as PCN increases, the metabolic load increases and, consequently, the growth rate decreases and faster-growing plasmid-free cells eventually overtake the culture [Palomares *et al.*, 2004]. Also, high PCN values may drive high protein production rates, which can result in protein aggregation and deficient Post-Translational Modifications (PTM) [Palomares *et al.*, 2004]. In fact, although chromosomal integration of the gene of interest in *E. coli* [Olson *et al.*, 1998] is a powerful alternative for overcoming problems of segregational instability in plasmid-based systems, this usually results in a lower production rates due to the low copy number of the recombinant gene [Palomares *et al.*, 2004]. Nevertheless, plasmid-based are still the most popular *E. coli* systems and the accurate determination of PCN can be easily determined in *E. coli* cultures using qPCR (Quantitative Polymerase-Chain Reaction) procedures [Lee *et al.*, 2006; Martins *et al.*, 2015].
On the other hand, the copy number of the inserted gene can have a significant influence on the expression characteristics of recombinant protein production in mammalian cell lines [Bandaranayake and Almo, 2014]. In general, despite higher gene copy numbers are directly correlated with higher expression levels, when individual clones are compared, this correlation breaks down and, in some applications, stable expression from single or low gene copy number is preferred [Mueller *et al.*, 2003]. The isolation of high copy number clones in Chinese Hamster Ovary (CHO) cells may be accomplished based on the co-transfection of the dihydrofolate reductase gene together with the gene of interest and the clone selection on methotrexate [Mueller *et al.*, 2003] or otherwise using the glutamine synthetase system and selection in methionine sulfoximine [Bandaranayake and Almo, 2014]. Finally, in order to establish a correlation between PCN and protein yields in human cell lines, PCN can also be accurately determined in these hosts using qPCR-based methods [Fliedl *et al.*, 2015].

In *P. pastoris* expression systems, a well-established method for increasing recombinant protein titer is to increase the number of genes in order to upregulate the transcription and translation pathways of the desired gene [Aw and Polizzi, 2013]. The methods for generating multi-copy clones as well as the implications of the high gene dosage present in *P. pastoris* recombinant strains were well reviewed by Aw and Polizzi (2013). Actually, different qPCR-based methods were reported to determine the recombinant gene dosage for all constructs based on pPICZ vectors [Nordén *et al.*, 2011] using SYBR Green or TaqMan [Abad *et al.*, 2010]. In particular, one of the best results ever reported for MP was those published by Nordén and coworkers (2011) with the expression of aquaporins (AQP) from mammals (HsAQP5 and HsAQP8) or plants (SoPIP1;2 and AtSIP1;1) in *P. pastoris*. The authors observed that the heterologous expression levels of all four aquaporin isoforms strongly respond to an increase in recombinant gene dosage, independently of the amount of protein expressed from a single gene copy clone [Nordén *et al.*, 2011]. However, the results are more mixed and in many cases the observed titer is below what would be predicted from the number of genes inserted [Aw and Polizzi, 2013]. Regardless of the mixed success of multi-copy clones, many groups continue to apply this methodology as a key strategy for increasing heterologous protein yield [Aw and Polizzi, 2013].

Actually, using a higher gene dosage may or may not represent a direct increase in the protein titer and depends on the expression host. For prokaryotic hosts such as *E. coli*, we believe that the plasmid-based expression systems are preferred over the chromosomal integration and despite an increase in PCN may conduct to an increase of protein titer, care should be taken concerning the segregational instability and the quality of the target protein should be evaluated [Palomares *et al.*, 2004]. For CHO cell lines, although the procedure for obtaining high gene copy clones is laborious and time-consuming, it often leads to an increase in protein yield [Mueller *et al.*, 2003]. Finally, the effect of increasing the target gene number is best characterized in *P. pastoris*-based recombinant systems where different methods for generating multi-copy clones as well as qPCR-based methods to determine the absolute number of genes are now available. In these systems, the production levels of the target protein are usually highly increased for high copy number clones, although a direct correlation is not always observed [Nordén *et al.*, 2011].

2.2.2 Codon Harmonization

Universal genetic codes are degenerated with 61 codons specifying 20 amino acids, thus creating synonymous codons for a single amino acid [Hu *et al.*, 2013]. These synonymous codons have been shown to affect protein properties in a given organism [Hu *et al.*, 2013] and, consequently, are not used equally in protein-coding sequences, once the nucleotide sequence of a gene contains information beyond the amino acid sequence of the protein it encodes [Klumpp *et al.*, 2012]. This additional information is contained in the usage patterns of synonymous codons that is applied with a bias towards a set of species-specific preferred codons [Klumpp *et al.*, 2012]. Extracting this additional information from sequence data is a key challenge for systems biology

and numerous studies have linked the pattern of codon usage in gene sequences to various properties of the proteins such as their abundance in the cell, their domain structure, folding kinetics, and cost of misfolding [Klumpp *et al*. 2012]. On the mechanistic level, codon usage is known to affect the kinetics of translation as individual codons are translated at different rates [Klumpp *et al*. 2012]. The differences in translation rates of synonymous codons are believed to result mostly from distinct intracellular concentrations of the corresponding tRNA (transfer ribonucleic acid) species, but also small discrepancies in the intrinsic kinetics have been demonstrated [Klumpp *et al.*, 2012]. Codon usage bias varies significantly between different organisms and attempts at producing proteins in heterologous cells often result in poor synthesis or formation of insoluble aggregates [Hu *et al.*, 2013]. By mimicking *E. coli* codon usage bias, synonymous substitutions of foreign genes ("codon harmonization") improved protein synthesis despite the varying success and unpredictability of this approach [Hu *et al.*, 2013]. Synonymous codons also influence the function of proteins since naturally occurring silent mutants can affect protein folding, alter substrate recognitions, as well as triggering several diseases. These concerns suggest that synonymous codons might control the folding of nascent polypeptides emerging from ribosome by regulating polypeptide elongation rates [Hu *et al.*, 2013].

The bottleneck at protein translation has been recognized as an important issue in the design of heterologous gene for recombinant expression [Chung *et al.*, 2013]. Differences in codon usage bias between the expression host and the target recombinant gene may lead to low translation rates and, consequently, to diminished levels of a particular heterologous proteins [Welch *et al.*, 2011]. As a result of random mutation and selection pressure, different organisms may have evolved to apply the synonymous codons with disparate frequencies [Chung *et al.*, 2013]. Accordingly, in the expression of a foreign gene in a particular host organism, the differences in codon bias can impair the protein translation machinery in a manner whereby the host is unable to efficiently translate the rare codons that may occur frequently in the recombinant gene [Rosano *et al.*, 2014]. As such, coding sequence re-design via codon optimization has been practically employed to adapt the foreign gene for efficient heterologous expression [Chung *et al*. 2013; Welch *et al.*, 2011]. Specifically, there are many on-line applications such as the OPTIMIZER (http://genomes.urv.es/OPTIMIZER/) [Puigbò *et al*. 2007], the Java Codon Adaptation Tool (JCat) (http://www.jcat.de/) [Grote *et al.*, 2005] or the codon optimization tool from Integrated DNA technologies (https://eu.idtdna.com/CodonOpt) that optimize the codon usage of a gene to increase its expression levels. Adjustments using the aforementioned tools on the codon usage patterns in the target gene may have important implications in the quality and titer of the recombinant protein expressed.

2.3 Economics vs complexity: guidelines to choose the right host

There are no universal systems for the recombinant expression of specific genes [Sorensen *et al.*, 2010]. The most common expression systems currently available for MP overexpression are microbial (bacteria or yeasts) or higher eukaryotes (insect or mammalian cells) (reviewed in [Bernadaut *et al.*, 2011; Midgett *et al.*, 2007; Wagner *et al.*, 2006]). In Table 1 we present a comparative overview of the advantages and weaknesses of different kinds of expression hosts and the emerging technology of CF systems.

Although different factors must be taken into account when selecting the host, this choice remains empirical, being extremely difficult to predict levels of biologically active, properly folded and membrane-inserted MP in each system. The reasons why some MP are overexpressed easily whilst others are expressed poorly are not fully understood [Andréll *et al.*, 2013]. Nevertheless, it is clear that this problem is not proportional to the number of transmembrane α-helices or the size of the protein, but is related to the "complexity" of the MP, i.e., how difficult is to fold into a functional state [Andréll *et al.*, 2013]. In bacteria, MP are overexpressed in the cytoplasmic membrane, whereas in eukaryotes they are typically overexpressed in the ER membrane [Wagner *et al.*, 2006]. Therefore, the MP are targeted, inserted and folded by different ways [Wagner *et al.*, 2006]. Moreover, it has also been described that the successful overproduction of some MP was shown to be linked to the avoidance of stress responses in the host cell [Grisshamer *et al.*, 2006]. In addition, a eukaryotic MP that resides in the plasma membrane must be first targeted as a nascent polypeptide to a protein-conducting channel (translocon) in the ER membrane [Grisshamer *et al.*, 2006]. Thus, the topology of each MP influences the choice of an appropriated host for its biosynthesis [Grisshamer *et al.*, 2006]. It is likely that the level of functional MP expression is dictated by a complex interplay of factors that probably include the following: the amount of mRNA (messenger Ribonucleic Acid) synthesized and its stability, the secondary structure of the mRNA and the presence of translational pause sites, folding of the nascent polypeptide chain in the ribosome and translocon, the efficiency of insertion into the membrane, the role of PTM in the folding process, e. g. N-glycosilation, and the requirement for molecular chaperones to facilitate folding [Andréll *et al.*, 2013]. Heterologous overexpression of MP can be hampered by different synthesis, targeting, insertion and folding characteristics in host [Wagner *et al.*, 2006]. Indeed, the different protein folding capacities, glycosylation patterns and membranes lipid composition between the organism from which the target membrane protein originates and the heterologous system are also important factors that affect the overexpression of MP [Wagner *et al.*, 2006]. Once the protein has been inserted and assembled correctly into the membrane, specific lipid requirements may be important to keep its activity. In general, the bacterial cell membranes are devoid of sterols and derivatives, polyunsaturated fatty acid chains and sphingolipids [Freigassner *et al.*, 2009]. Specifically, the inner membrane and the inner leaflet of the outer membrane of *E. coli* are mainly composed by phosphatidylethanolamine followed by phosphatidylglycerol and few

cardiolipin while the outer leaflet of the outer membrane is highly enriched in Lipopolysaccharide (LPS) [McMorran *et al.*, 2014]. On the other hand, *Lactoccocus* lactis contains a single membrane with a high fraction of glycolipids while plasma membranes from yeasts are composed of phospholipids, sterols (ergosterol) and sphingolipids (inositol) [Marredy *et al.*, 2011]. Finally, the cholesterol present in mammalian cell lines membranes have proved to be essential for the functionality of certain MP, as analyzed in detail below [Freigassner *et al.*, 2009].

As a rule of thumb, the choice of which expression system to use for a new target is usually dictated initially by which systems are already in use in the laboratory or in adjacent laboratories [Andréll *et al.*, 2013]. This increases the speed at which results are obtained, because user knowledge can greatly improve the yield from a particular expression system, especially when using insect cells or mammalian cells where the health of the cells before production is vital to obtain good yields [Andréll *et al.*, 2013]. In general, the majority of the expression hosts described in Table 1 have been described as Generally Recognized As safe Microorganisms (GRAS) and, therefore, if we intend to give a therapeutic use to the target recombinant protein, mammalian cell lines may be the most appropriated host but others can be used such as *P. pastoris* or *Lactococcus Lactis*. On the other hand, when the time and the cost of the process are critical parameters, then the choice should fall in microbes such as yeast or bacteria since the implementation and maintenance of insect and mammalian cell cultures requires more time, are more labor intensive and extremely expensive. In particular for mammalian cell lines, transgenes can be expressed by two ways, stable or transiently [He *et al.*, 2014]. Stable gene expression in is initially slower and more technically challenging when compared to large-scale transient expression [Chaudhary *et al.*, 2012]. However, once a clonal cell line is generated, long-term overexpression from stably transfected cells can be much more easy and consistent, since the purification of large quantities of supercoiled plasmid DNA required for transient-expression is not required [Chaudhary *et al.*, 2012].

Specifically, the serotonin transporter (SERT) presents several unusual characteristics [Andréll *et al.*, 2013], namely: it presents two N-glycosylation sites in extracellular loop 2 that are essential for efficient folding of the protein [Tate and Blakely, 1994]; SERT appears to require the molecular chaperone calnexin for efficient folding [Tate *et al.*, 1999]; and it requires cholesterol, probably for stabilizing the folding state of the transporter [Andréll *et al.*, 2013]. Indeed, due to the above mentioned reasons it failed to express in other systems than mammalian cell lines. Therefore, this particular case-study highlights the fact that uncommon characteristics of a particular MP must require the employment of more complex hosts such as mammalian cell lines. Finally, a summary reporting the weaknesses and strengths of the most common expression systems and the emerging technology of CF expression is shown in Table 2

Table 1 - Major advantages, limitations and general characteristics of different hosts suitable for recombinant membrane protein expression.

Host	*Advantages*	*Drawbacks*	*Other characteristics*	*References*
Escherichia coli Gram-negative bacterium	Inexpensive; Rapid generation of expression plasmids; Fast growth; Easy scale up; Simple culture requirements.	Endotoxin; Inclusion body formation; Inefficient protein secretion; Many MP do not fold properly; Lack of efficient PTM.	Specific strains (e.g. Lemo21) or introduction of solubility tags may improve MP expression.	[Bernaudat *et al.*, 2011] [Midgett *et al.* 2007]
Pichia pastoris Methylotrophic yeast; GRAS organism	Efficient protein secretion with low levels of endogenous proteins. Capable of performing many PTM; Low cost of culture media; Industry-scale fermentation.	Glycosylation pattern different from mammalian; Intracellular recovery of large amount of cells may require specific equipment (French-press); High oxygen demand.	Improved glyco-engineered strains obtained using the GlycoSwitch® technology; Wide range of genetic tools, plasmids, strains and promoters available; The preference for the respiratory growth allow to be cultivated at high cell densities.	[Gonçalves *et al.*, 2013] [Laukens *et al.*, 2015] [Pedro *et al.*, 2015a]
Brevibacillus choshinensis Gram-positive bacterium Safe microorganism	Low protease activity; Protein levels highly increased when using inducible promoters from *Brevibacillus megaterium*.	Few examples of successfully expressed MP; Low levels with P2 constitutive promoter; As low levels of endogenous proteins are secreted, more suitable for secreted proteins.	Does not produce lipopolysaccharide (LPS); Able to export heterologous proteins directly into the extracellular medium.	[Pedro *et al.*, 2011] [D'Urzo *et al.*, 2013]
Lactococcus lactis Food-grade lactic acid bacteria; GRAS organism	Inclusion bodies have not been observed so far; Has only a single membrane.	Highly overproduced and poorly expressed MP both may result in severe growth defects and accumulation of misfolded MP, evoking a general stress response; Has not been employed widely as other hosts.	Expression of heterologous genes under the control of the nisin-inducible promoter P_{nisA}.	[Bernaudat *et al.*, 2011] [Noreen *et al.*, 2011] [Marreddy and Pinto *et al.*, 2011]
Insect cells Baculovirus-infected cells	More native environment than yeast; More compatible with eukaryotic MP because of similar codon usage rules than *E. coli* or *P. pastoris*; Well-established protocols; Good secretion.	Cost; Non-native glycosylation and lipid environment; Cell lysis; Some of the PTM are not identical to those found in mammalian; Long production time; Relative high media costs.	Used for MP expression as a compromise between bacterial and mammalian systems.	[Bernaudat *et al.*, 2013] [Midgett *et al.*, 2007]
Mammalian cells Stable integration and transient transfection	Proper folding; Stable/transient folding; Native lipid environment and post-translational pathways.	High media costs; Slow growth rates; Low expression; Viral infection; Cost; Higher technical requirements.	For particular targets, may be the only expression system able to express a given MP in a functional and properly folded state.	[Midgett *et al.*, 2007] [Andréll *et al.*, 2013]
Cell-Free expression	Short time reaction; Manipulation of reaction conditions allow to control conveniently the PTM; Plasmid or DNA can be directly used for protein expression; Special proteins can be expressed with a composition of non-natural amino acids.	High costs, Low protein production rates; Insufficiency of PTM is a bottleneck to obtain complex proteins in a functional form.	May be based in prokaryotic or eukaryotic CF systems; MP may be produced co-translationally in artificial membrane environments.	[Rajesh *et al.*, 2011] [Proverbio *et al.*, 2013] [Zheng *et al.*, 2014]

Table 2 – Comparative overview of the strengths and weaknesses of different types of expression systems actually employed for recombinant MP expression.

	Escherichia coli	*Pichia pastoris*	Mammalian cell lines	Cell-Free
Cost	−/−	−/−	+/+	+/+
Ease of manipulation	+/+	+/+	−/−	−/+
User-knowledge importance	−/−	−/−	+/+	+/−
Glycosilation pattern	−/−	[1]+/−	+/+	[2]+/+
Other PTM	−/−	+/+	+/+	[2]+/+
Protein productivity	+/−	+/+	−/−	−/−
Scalability	+/+	+/+	+/−	−/−

Legend: **−/−** Low/absent; **+/−** Medium; **+/+** High; [1]Using the GlycoSwitch® technology [Laukens *et al.*, 2015]; [2]Depends on the cell-free system.

2.3.1 Upstream-based strategies to improve membrane protein expression levels and/or folding in *Escherichia coli*, *Pichia pastoris* and mammalian cell lines

The development of recombinant heterologous expression systems such as *E. coli*, *P. pastoris or* mammalian cell lines coupled with the development of protein synthesis technologies allowed to overcome inherent problems associated with the purification of MP from poorly abundance natural sources [Zheng *et al*. 2014]. Although the rapidly increasing on understanding MP biogenesis, diverse improvements on MP expression are still achieved on a "trial and error" basis. Indeed, a bioprocess design to obtain an active, functional and correctly folded MP depends on a complex interplay between different factors such as the heterologous gene itself, the expression vector and appropriated biosynthesis conditions for the expression host. However, often the optimization of one or a synergy of these factors allows a high increase in the production levels and/or in the functionality of the target protein. Indeed, although specific strategies are reported using different expression systems for a specific MP, these approaches can often be applied to other strains in the same expression system or to other MP. Therefore, in Table 3 we review recent and important achievements in this field for *E. coli*, *P. pastoris* and mammalian cell lines.

Table 3 – Summary of different strategies applied for improving the heterologous expression of MP in diverse hosts, namely *E. coli* and *P. pastoris* and mammalian cell lines.

Escherichia coli								
Target membrane protein	**Strain**	**X-ray Structure[4]**	**Strategy**	**Mechanism**	**Adaptable to others:**		**Observations**	**Reference**
					Membrane proteins	**Strains**		
E. coli anti-transporter Na^+/H^+ Nha A	Lemo21 (DE3)	3.45Å Hunte *et al.*, 2005	Construction of a BL21 (DE3) derivative strain designated Lemo21 (DE3) that is transformed with a plasmid harboring the gene encoding T7 lysozyme, an inhibitor of the T7 RNA polymerase, under the control of the well-titratable rhamnose promoter.	Modulation of T7 RNA polymerase can have a tremendous impact on the amount of correctly folded and inserted MP in the cytoplasmic membrane relative to the non-inserted and aggregated form; At the optimal rhamnose concentrations, the amount of correctly inserted MP is highly enhanced.	Yes	No	Screening at 30ºC is sufficient to optimize MP expression; Codon optimization did not improve MP expression; Compatible with autoinduction. Scale reported: 1 in 2.5 L baffled shake-flasks.	Schlegel *et al.*, 2012
Microbacterium liquefaciens sodium-hydantoin transporter Mhp1		2.85 Å Weyand *et al.*, 2008						
Severe acute respiratory syndrome coronavirus membrane protein	BL21 (DE3)	2.2 Å (Chains C and F) Liu *et al.*, 2010	Expression of the target MP in *E. coli* fused with SUMO tag.	SUMO chaperoning properties facilitate proper protein folding, which enhances the solubility and biological activity of the purified MP.	Yes	Yes	After purification of the fusion construct by AC, the SUMO can be cleaved with SUMO protease 1, generating the native MP. Scale reported: 0.5 L.	Zuo *et al.*, 2005
Human 5-lipoxygenase-activating protein		4.00Å Ferguson *et al.*, 2007						
Rat neurotensin receptor	MG1655	2.8 Å White *et al.*, 2012	Construction of new prokaryotic expression vectors using the rhaB promoter (pRHA-67, pRHA-113, pRHA-109, and pMPX-66) inducible by L-rhamnose and tightly regulated by D-glucose.	The tightly regulated pRHA promoter is almost completely repressed until induced (with L-rhamnose) and, therefore, is suitable for the expression of toxic proteins.	Yes	Yes	Through variations in the plasmid copy number, additional regulatory control is achieved to fine-tune the expression of toxic proteins. Scale: 3 mL.	Giacalone *et al.*, 2006
ToxR-PhoA from *Vibrio cholera* ToxR		NF						
Thermotoga maritima YidC	MC1061-ffh77	NF	Design of a novel protein fusion partner (P8CBD) to utilize the co-translational SRP pathway in order to target heterologous proteins to the *E. coli* inner membrane.	MP targeting to *E. coli* inner membrane using the SRP pathway results in reduced cytoplasmic toxicity. Thereby, the over-expression of target MP is significantly increased as it is shown that the P8CBD fusion partner utilize the SRP membrane targeting pathway.	Yes[1]	Yes (Probably)	P8CBD features: epitopes for immuno-detection; FLAG and enterokinase nucleotide sequences; poly-histidine sequences may be incorporated (preferably in C-terminal).	Luo *et al.*, 2009
Yeast Oxa1p		NF						

Human GPCR Y4	MC1061	NF	Development of a method that makes use of a small additional RNA sequence upstream to the RNA sequence of the target MP (transcriptional fusions) and results in the production of a bicistronic mRNA. It was observed improvements in the quantity/quality of the produced material for several MP compatible with structural studies.	The translation of the upstream RNA sequence was not essential for increased expression but rather the sequence itself had a large impact on protein yields, suggesting that alternative folding of the transcript was responsible for the observed effect.	Yes	Yes	Does not require protease treatment and subsequent removal of the fusion protein.	Marino *et al.*, 2015
Channel MscL from *Mycobacterium tuberculosis*		3.5Å Steinbacher *et al.*, 2007						
E. coli ZraS (histidine kinase)	BW25113 (tig^+)	1.61Å Celikel *et al.*, 2012	Explore the feasibility of reprogramming the chaperone pathways to improve the biosynthesis of MP in *E. coli*, focusing on eliminating TF/SRP competition through TF inactivation, improving targeting and/or delivery to the inner membrane through SRP overexpression and promoting insertion and folding in the lipid bilayer by YidC overproduction.	Because TF and SRP bind nonexclusively to the same location of the ribosome, SRP should have unimpeded access to the nascent transmembrane segment of MP in tig^+ strains (TF deficient), guaranteeing efficient capture of MP and proper targeting to the SRP pathway.	Yes	Yes	YidC coexpression is beneficial to the production of polytopic proteins while higher levels of SRP have the opposite effect.	Nannenga *et al.*, 2011
Haloterrigena turkmenica deltarhodopsin (HtdR) and *Natronobacterium pharonis* sensory rhodopsin II (pSRII)		2.7Å Zhang *et al.*, 2013						
Glutamate transporter (GltP) from *E. coli*	BW25113B	NF	A new strategy was developed that involves fusion in tandem of green fluorescent protein and the erythromycin resistance protein (23S ribosomal RNA adenine N-6 methyltransferase, ErmC) to the C-terminus of a target MP. New strains are generated in which the partial removal of the transcriptional silencing mechanism changes the levels of proteins essential for the functional overexpression of MP.	With the increase of the erythromycin concentration, a number of evolved strains were obtained. Four (NG2, NG3, NG5 and NG6) were fully characterized and all carried a mutation in the hns gene, whose product is involved in genome organization and transcriptional silencing. The degree of expression of MP correlates with the severity of the hns mutation, but cells in which hns was deleted showed an intermediate expression performance.	Yes	No	Fluorescence intensity of GFP is used to report the folding state of the target protein while ErmC is used to select clones that present increased expression.	Gul *et al.*, 2014
Branched-chain amino acid permease (BcaP) from *L. Lactis*		NF						
Putative P-type cation transporter (PacL) from *L. lactis*		NF						

Alpha-helical inner membrane protein Rv1337 from *Mycobacterium tuberculosis*	TOP10	NF	Design and selection of mutant strains in which the coding sequence of the target MP is fused to a C-terminal selectable marker, so that the production and survival of cells on selective media is linked to expression of the target MP.	The double-selecting mechanism allows to discriminate between cells expressing high levels of target MP and those expressing basal protein concentrations.	Yes	NR	Curing of selected mutants is achieved by *in vivo* digestion with a rare-cutting endonuclease, (the homing endonuclease I-CreI).	Massey-Gendel *et al.*, 2009
Human selenoprotein K	BL21 (DE3)	NF	Expression of the target protein fused with maltose-binding partner; This tag is subsequently cleaved off the target MP	Selenoprotein K solubilization is enhanced after fusion with the maltose-binding partner.	Yes	Yes	The fusion partner in selenoprotein K can be cleaved off in the presence of several detergents.	Liu *et al.*, 2012
Human tetraspan vesivle protein /TVP) Synaptogyrin 1	Lemo21 (DE3)	NF	Optimization of the N-terminal sequence of the target gene coupled with the usage of Lemo21 (DE3) strain.	Codon usage optimization of the N-terminal guarantee an efficient translation start; the balanced transcription/translation rate provided by Lemo strain avoid the saturation of the cell insertion machinery - Sec translocon.	Yes	No	The protein was soluble and stable in several mild detergents and was correctly folded.	Löw *et al.*, 2012

Pichia pastoris

Target membrane protein	Strain	X-Ray structure[4]	Strategy	Mechanism	Adaptable to others:		Observations	Reference
					Membrane proteins	Strains		
Human adenosine A2A GPCR[2]	GS115 (his4)	2.6Å Jaakola *et al.*, 2008	Overexpression of Hac1p in *P. pastoris* generally increases heterologous proteins production, although it needs to be evaluated in an isolate perspective.	UPR upregulates factors that restore ER homeostasis upon protein folding stress. Hac1p overexpression in *Pichia* strongly upregulates the KAR2 (UPR) - responsive gene, leading to the expansion of the intracelular membranes.	Yes	Yes	Inducible Hac1p expression is more effective than constitutive expression. Alpha-mating factor prepro signal processing of a GPCR was more efficient with Hac1p overexpression, resulting in a significant homogeneity.	Guerfal *et al.*, 2010

Human Membrane-bound catechol-*O*-methyltransferase	X33	NF	Addition of chemical chaperone dimethylsulfoxide (6% and 2.5 % (v/v) for MBCOMT and GPCR, respectively) to the culture medium.	Not fully understood; may involve the up-regulation of genes involved in membrane lipid components and, consequently, the induction of membrane proliferation.	Yes	Yes	As dimethylsulfoxide possess antioxidant properties, the reduction of protein oxidation may also explain the increase in protein titer/quality.	Pedro *et al.*, 2015b
20 mammalian GPCR[2]	SMD1163	----------						André *et al.*, 2006
Human μ-opioid receptor[2]	SMD 1163	2.8Å Manglik *et al.*, 2012	Fusion of the target gene to Sacharomycces cerevisiae α-factor signal.	Protein translocation to the membrane is enhanced.	Yes	Yes	----------	Sarramegna *et al.*, 2005
Human Aquaporins[2]	X33	----------	Use of mammalian Kozak's consensus for the ATG.	Guanine (G) at position +4 in highly expressed genes gives rise to alanine (GCN) and glycine (GGN), small amino acids that are suitable for an efficient cleavage of the initiator methionine from the nascent polypeptide chain.	Yes	Yes	Apparently mimicking the yeast consensus sequence (ATGTCT) has a negative influence in the expression level.	Oberg *et al.*, 2009
Human Aquaporin 5	X33	2Å Horsefield *et al.*, 2008	Generation of multi-copy clones with several copy of the target gene.	An increase in the recombinant gene dosage correlates with an increased protein titer.	Yes	Yes	----------	Nordén *et al.*, 2011
Human P-glycoprotein	KM71H	NF	Codon usage optimization based on highly expressed *P. pastoris* genes.	Translational efficiency is maximized and, consequently, the expression levels of the target protein are increased.	Yes	Yes	P-glycoprotein sequence was adjusted taking into account relative codon frequencies for each amino acid, the GC content and controlling for mRNA instabilities.	Bai *et al.*, 2011
Human α3 and Β1 subunits of Na,K-ATPase	Derived from CBS7435 *Δhis4Δku70* or SMD1168	NF	Engineering of a *P. pastoris* strain in the sterol pathway towards the synthesis of cholesterol instead of ergosterol to foster the functional expressional of membrane proteins.	Humanized cholesterol-producing *P. pastoris* strain afford a good environment for stability of the target proteins, improving its expression levels and co-localization on the plasma membrane.	Yes	Yes	Especially important for expression of MP which function depends on cholesterol. Works better with alcohol oxidase than constitutive promoters.	Hirz *et al.*, 2013

Mammalian Cell Lines								
Target membrane protein	**Mammalian cell line**	**X-Ray structure[4]**	**Strategy**	**Mechanism**	**Adaptable to others:**		**Observations**	**Reference**
					Membrane proteins[1]	**Cell lines[3]**		
Human epidermal growth factor receptor	CHO	NF	Co-expression of the anti-apoptosis gene Bcl x_L.	Transient production of MP is improved by decreasing the stress caused by the transient transfection process; Unfolded protein response and cell death are diminished.	Yes	Unknown	For sustained production, multiple CHO Bcl x_L clonal isolates showed significant expression of the target receptors whereas CHO clonal isolates did not show any recognizable expression.	Ohsfeldt *et al.*, 2012
Human fibroblast growth receptor		2.3Å Ibrahimi *et al.*, 2001						
Rat serotonin receptor subtype 2c (5HT2c)	HEK-293T	NF	Development of a procedure using fluorescence-activated cell sorting for the direct selection of stable mammalian cell lines that express target proteins with high yield.	Target protein is co-expressed with GFP; after repeated rounds of cell sorting, a highly expressing cell line is developed based on GFP fluorescence; as GFP and the target protein are derived from the same mRNA, GFP levels are correlated with the target protein levels.	Yes	Yes (probably)	After the termination codon of the protein of interest is an IRES that enables translation of GFP to be initiated from an internal site of the bicistronic mRNA transcript, allowing the production of two separate proteins, GFP and the target protein.	Mancia *et al.*, 2004
Human Rh C glycoproteins	HEK-293 GnTI$^-$	2.1 Å Gruswitz *et al.*, 2010	Implementation of a protocol for toxic MP expression employing a strain lacking GnTI and use of a promoter that allows high-density cell cultures.	Lack of GnTI restricts *N*-linked glycans to a homogeneous Man5-GlcNac2 greatly facilitates their enzymatic removal via endo- and exoglycosidases; Tetracycline-inducible promoter allows the establishment of high-density cell cultures when cytotoxic proteins are expressed.	No	Yes	After identification of highly expressing cell lines, suspension cell cultures are easily scaled using spinner flasks or cell-bag bioreactors.	Chaudhary *et al.*, 2012
Human ABCG2	HEK-293; HeLa; MCF7	2.2Å Oldham and Chen, 2011	Use of recombinant baculoviruses containing mammalian cell-active expression cassetes (BacMam) with a cytomegalovirus promoter.	Transduction is highly enhanced using this system and different cell lines may be screened by simply adding a viral inoculum.	Yes	Yes	Addition of butyric acid (histone deacetylase inhibitor) increase protein expression in transduced cells; transduction efficiency varies among different cell lines.	Shukla *et al.*, 2012

Human sodium/iodide symporter (NIS)	HeLa; Hep3B; U87MG; TPC-1; FRO; B-CPAP	NF	Optimization of the coding sequence of human NIS gene by replacing NIS DNA codons from wild type to new codons that present the highest usage in human gene translation.	The increase in the codon adaptation index (0.97 and 0.79 for the non-optimized) associated with a more efficient translation renders higher levels of NIS and higher uptake of ^{125}I.	Yes	Yes	RNA instability motifs, very high (>80%) or very low (<30%) GC content regions and cis-acting sequence motifs were also removed.	Kim *et al.*, 2015

[1]By replacement of the native N-terminal signal or targeting sequence. These examples are especially encouraging for the over-expression of α-helical MP with an N-out topology such as GPCR. [2]Reviewed in [Ramón and Marín, 2011]. [3]These strategies are considered to be adaptable to other membrane proteins when more than one type of membrane proteins are reported and are probably adaptable when the reported strategy isn't MP-specific. [4]According to UniProt (http://www.uniprot.org/). **Abbreviations:** ABC - ATP-binding cassette; CHO - Chinese Hamster Ovary; ER - Endoplasmic Reticulum; GFP - Green Fluorescent Protein; $GnTI^-$ - Lacks *N*-acetylglucosaminyltransferase I; GPCR – G-Protein Coupled Receptor; HEK-293 - Human Embryonic kidney 293; HEK-293 - T antigen-transformed Human Embryonic Kidney 293; IRES - Internal Ribosome Entry Site; MBCOMT - Membrane-bound catechol-*O*-methyltransferase; MP – Membrane Protein; NF - Not found; NIS - Sodium-iodide symporter; NR - Not Reported; *P. pastoris* – *Pichia pastoris;* SRP - Signal Recognition Particle; SUMO - Small Ubiquitin-Related Modifier; TF – Trigger Factor; UPR – Unfolded Protein Response;

2.4 Expressing encapsulated membrane proteins in preformed lipid-bilayers using cell-free expression systems

Toxicity, limited membrane space for their functional folding as well as inefficient transport and membrane insertion mechanisms are the major problems associated with cellular-based expression systems [Henrich *et al.*, 2015]. As an alternative to these conventional *in vivo* expression systems, CF systems have been growingly applied for MP production [Carlson *et al.*, 2012; Roos *et al.*, 2012]. However, CF are still limited to low protein production rate, expensive reagent costs and short reaction durations of protein synthesis [Zheng *et al.*, 2014]. This technology makes use of cellular extracts/lysates containing the protein synthesis machinery and related elements to direct protein synthesis from added DNA or mRNA templates [He *et al.*, 2011]. Most common CF protein synthesis systems are derived either from prokaryotic crude cell extracts such as *E. coli* cell extracts or from eukaryotes, namely rabbit reticulocyte lysates, wheat germ extracts and insect cell lysates [Zheng *et al.*, 2014]. In particular, the most popular CF systems for MP expression are based on *E. coli* and wheat germ extracts [Rajesh *et al.*, 2011]. Whilst the *E. coli* method involves a coupled transcription/translation system where protein synthesis can be initiated from a plasmid DNA or PCR product, the wheat germ system uses decoupled translation and requires mRNA to initiate protein synthesis [Rajesh *et al.*, 2011]. The optimization of CF expression is relatively simple since the cytotoxic effects are eliminated, the absence of an enclosed membrane system facilitates the addition of ligands, cofactors or protease inhibitors and, therefore, the low MP yield is only associated with inefficient translation [Henrich *et al.*, 2015]. In fact, the magnesium ion concentration for each DNA template should always be optimized and attention should be taken regarding the first few codons of the mRNA since suboptimal nucleotide sequences could restrict their access by ribosomal RNA subunits [Henrich *et al.*, 2015]. An overview of the different procedures based on the screening of expression environments suitable to support folding and stability of the target MP are depicted in Table 4.

Table 4 – Summary of the different expression modes currently available for the CF expression of MP [Carlson *et al.*, 2012; Henrich *et al.*, 2015, Proverbio *et al.*, 2013; Rajesh *et al.*, 2011].

Cell-Free expression systems			
CF mode	**Co-translational events**	**Post-translational events**	**Advantages/Drawbacks**
Precipitate forming CF (P-CF)	Synthesized MP precipitates after translation.	Post-translational solubilization in detergents.	The yield is higher in the absence of hydrophobic compounds. MP may retain their folded structures without excessive refolding procedures.
Detergent-based CF (D-CF)	MP produced co-translationally through supplementation with appropriated detergents.	Reconstitution into artificial membranes by classical *in vitro* techniques.	Synthesis rate of MP is often drastically reduced, compared to P-CF. Different MP topologies may require different artificial bilayers. As lipids are stabilizing or even structural MP elements, fully functional MP production may only be obtained using L-CF.
Lipid-based CF (L-CF)	Co-translational reconstitution through addition of lipids as preformed nanodiscs or liposomes.	None.	

The open nature of CF systems provides an artificial membrane-mimicking systems composed of detergents, liposomes or other hydrophobic compounds such as fluorinated surfactants or amphipols that could be adjusted basis on the synthesized MP [Roos *et al.*, 2012; Henrich *et al.*, 2015]. Indeed, conventional CF of a MP results in insoluble precipitates that must be solubilized using detergents [Henrich *et al.*, 2015; Rajesh *et al.*, 2011]. This approach has been almost exclusively applied with success for β-barrel folds using the following detergents: lysopalmitoyl phosphatidylglycerol (LPPG), dodecylphosphocholine (DPC) and sodium dodecyl sulfate (SDS) [Rajesh *et al.*, 2011]. On the other hand, the addition of detergents to the CF reaction enables the direct expression of folded proteins as the close proximity of detergent micelles to the translation apparatus aids the direct solubilization of the synthesized MP [Henrich *et al.*, 2015]. The most successful classes of detergents are long-chain polyoxyethylene-ethers such as Brij derivatives and steroid glycosides such as digitonin, while the application of n-octyl-β-$_D$-glucopyranoside (β-OG) or 3-[(3-Cholamidopropyl)dimethylammonio]-1-propansulfonate (CHAPS) is avoided, once they inhibit the transcription/translation machinery [Rajesh *et al.*, 2011; Schwarz *et al.*, 2008]. A major limitation of using detergents to solubilize MP is that the absence of bound lipids can

compromise the stability of MP and, therefore, non-micellar membrane-mimicking systems have been developed [Carlson *et al.*, 2012; He *et al.*, 2011; Rajesh *et al.*, 2011]. In particular, MP from CF reactions may be expressed into a lipid bilayer following two approaches: 1) detergent-solubilized MP are reconstituted into liposomes by freeze-thaw or by the removal of detergents; 2) direct CF synthesis in the presence of unilamellar liposomes [Rajesh *et al.*, 2011]. However, it has been described that preformed liposomes often tend to rearrange and to precipitate during the expression reaction resulting in high sample heterogeneity, thus limiting its utility for structural studies [Roos *et al.*, 2012]. As an alternative, nanodiscs based on lipoprotein particles formed by membrane scaffold proteins can provide stable membranes of defined sizes [Carlson *et al.*, 2012; Roos *et al.*, 2012]. The nanodisc size seem to be important for an efficient formation of the complexes and the lipid composition of the membrane is important to obtain active MP, as previously described for the translocase MraY from *E. coli* [Roos *et al.*, 2012]. Moreover, different chemical chaperones such as polyethileneglycol derivatives or amino acids can be added into CF expression reactions acting as stabilizers [Kai *et al.*, 2013]. In general, P-CF mode is the fastest approach and usually routinely employed for the first level of MP expression protocol development to tune MP production to the desired levels and may be applied to complex proteins even with 12 transmembrane segments [Proverbio *et al.*, 2014]. On the other hand, proteomicelles formation in D-CF are usually applied when P-CF does not result in functionally folded protein since protein aggregation is avoided [Schwarz *et al.*, 2008]. Although P-CF and D-CF are feasible and can be efficient, the L-CF mode where proteins are targeted to artificial membrane environments usually increase MP stability, although problems with efficient translocation of the synthesized MP may occur like those observed in living cells [Schwarz *et al.*, 2008].

Finally, the native integrity of the target MP largely depends upon the choice of the CF system and, in general, a eukaryotic MP should be produced the intracellular machinery from a eukaryotic system and vice-versa [He *et al.*, 2011].

2.5 Case study on optimizing membrane-bound catechol-*O*-methyltransferase biosynthesis and purification in a biologically active form

Catechol-*O*-methyltransferase (COMT; EC 2.1.1.6) is a magnesium-dependent enzyme that catalyzes the methylation reaction whereby a methyl group from the S-adenosyl-$_L$-methionine (SAM) is transferred to one of the catecholic hydroxyls [Bonifacio *et al.*, 2007; Ma *et al.*, 2014]. As reaction products, are obtained the *O*-methylated catechol and S-adenosyl-$_L$-homocysteine [Bonifacio *et al.*, 2007]. In humans, COMT appears as two molecular forms, a soluble (SCOMT) and a membrane-bound (MBCOMT) isoforms [Bonifacio *et al.*, 2007]. Specifically, while SCOMT is located in the cytosol, MBCOMT is anchored to the rough ER membrane and is characterized as an integral MP with the catalytic portion of the enzyme

oriented toward the cytoplasmic side of the membrane [Bonifacio *et al.*, 2007]. In general, COMT major physiological role is the elimination of biologically active or toxic catechols and the affinity of both isoforms depends on the substrate. For instance, MBCOMT affinity for catecholamines is 10 to 100 fold higher than that of SCOMT [Bonifacio *et al.*, 2007; Ma *et al.*, 2014]. The two forms are proposed to have at least partially distinct roles: MBCOMT is believed to be primarily involved in the termination of dopaminergic and noradrenergic synaptic neurotransmission when there are physiologically relevant low concentrations of catecholamines [Myöhänen and Männistö, 2010]. So, MBCOMT presents higher affinity but lower reaction velocity for catecholamines than SCOMT [Reenilä and Männistö, 2001]. Also, a polymorphism at codon 108/158 (for soluble and MBCOMT, respectively) that involves a substitution of a valine to a methionine (Met) in the polypeptide chain influences COMT enzymatic activity where the Met variant is associated with low enzymatic activity and decreased thermal stability [Bonifacio *et al.*, 2007; Hosak, 2007]. During the last decades, COMT has been implicated in diverse human diseases including certain types of cancer [Yager, 2012; Wu *et al.*, 2015], cardiovascular diseases [Voutilainen *et al.*, 2007] or neurologic disorders (Parkinson's disease, schizophrenia) [Apud and Weinberger, 2007; Bonifacio *et al.*, 2007], among others. In particular, the best documented is the relevant role that COMT plays in Parkinson's disease, a disorder that is characterized by dopaminergic neuronal loss in the substantia nigra and by striatal dopamine loss with accumulation of the protein α-synuclein [Müller, 2015]. Initially, the therapy applied to Parkinson's disease (Please see Figure 4) was the dopamine replacement using levodopa together with an aromatic amino acid decarboxylase inhibitor such as carbidopa [Bonifacio *et al.*, 2007]. However, the efficacy of this therapy decreases over time and most patients develop fluctuating responses and dyskinesias [Bonifacio *et al.*, 2007]. Therefore, the chronic levodopa/aromatic amino acid decarboxylase inhibitor application with concomitant inhibition of COMT and monoamine oxidase (such as selegiline and rasagiline) is suggested as standard levodopa application in Parkinson's disease patients who need levodopa, if they will tolerate it [Bonifacio *et al.*, 2007; Müller 2015].

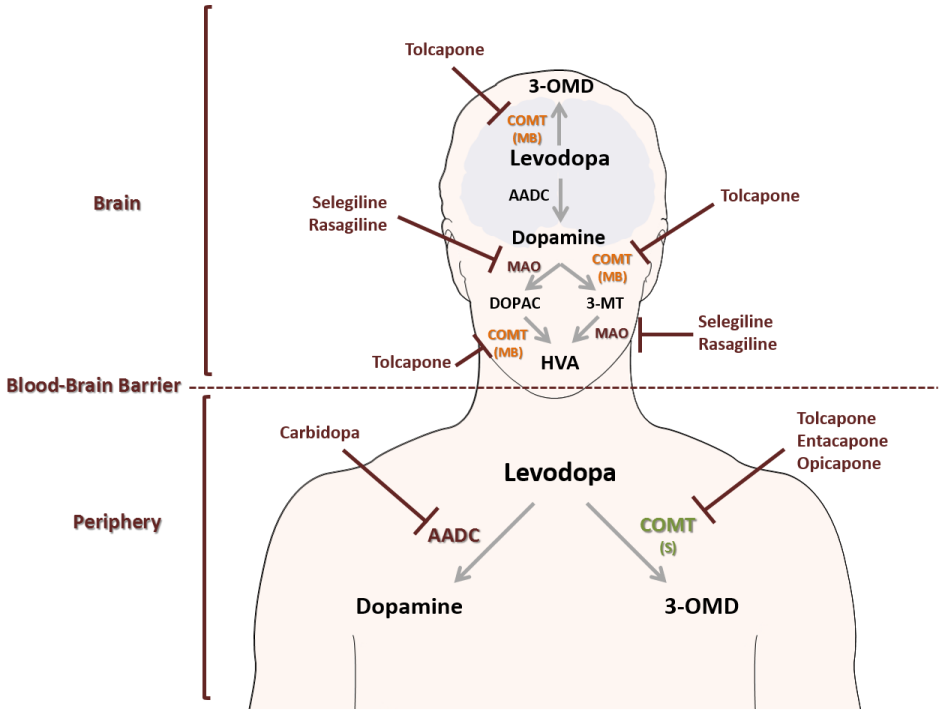


Figure 4 - Sites of action of PD drugs.

Abbreviations: 3-MT, 3-methoxytyramine; 3-OMD, 3-*O*-methyldopa; AADC, aromatic amino acid decarboxylase; COMT, catechol-O-methyltransferase; DOPAC, Dihydroxyphenylacetic acid; MAO, Monoamine oxidase. According to their relative distribution [Bonifácio *et al.*, 2007], in brackets are shown the most abundant COMT isoform in each location (S for SCOMT and MB for MBCOMT).

On the other hand, COMT is regarded a gatekeeper phase II enzyme that *O*-methylates the catechols, thus blocking their estrogenicity and further metabolism [Yager, 2012]. In fact, several studies reported that COMT can reduce the potential of DNA damage and increase the concentration of 2-methoxy-estradiol, an antiproliferative metabolite in human breast cancer and lung cancer [Michnovicz *et al.*, 1986; Wu *et al.*, 2015], although opposite results were obtained for breast cancer by other research group [Bergman-Jungestrom and Wingren, 2001]. Indeed, COMT is an important pharmacological target and its inhibition leads to an improved outcome in Parkinson's disease therapy, mostly via the soluble isoform. On the other hand, it has been described that new pharmacological therapies consisting in using central nervous system-penetrant COMT inhibitors may be important in other diseases than Parkinson's disease such as schizophrenia [Apud *et al.*, 2007]. In fact, conducting crystallization studies of MBCOMT in complex with different ligands similarly to what was previously performed for rat SCOMT [Bonifacio *et al.*, 2002; Rodrigues *et al.*, 2005] may aid in the design of new molecules isoform-specific with improved selectivity for MBCOMT. Actually, as far as we know, although several crystal structures of human and rat SCOMT in complex with different molecules are available on Protein Data Bank (http://www.rcsb.org/pdb/home/home.do), no structures of human and rat MBCOMT were yet solved or reported in the Membrane Proteins of Known

Structure Database (http://blanco.biomol.uci.edu/mpstruc/) or in the Membrane Protein Data Bank (http://pdbtm.enzim.hu/). In fact, attaining high quantities of pure recombinant human MBCOMT has been the major stumbling block to completion of crystallization experiments, thus compromising the development of new inhibitor molecules. As recombinant MBCOMT research has been intensified during the last years, an overview of the most important strategies recently reported for MBCOMT biosynthesis and purification is presented in Figure 5.

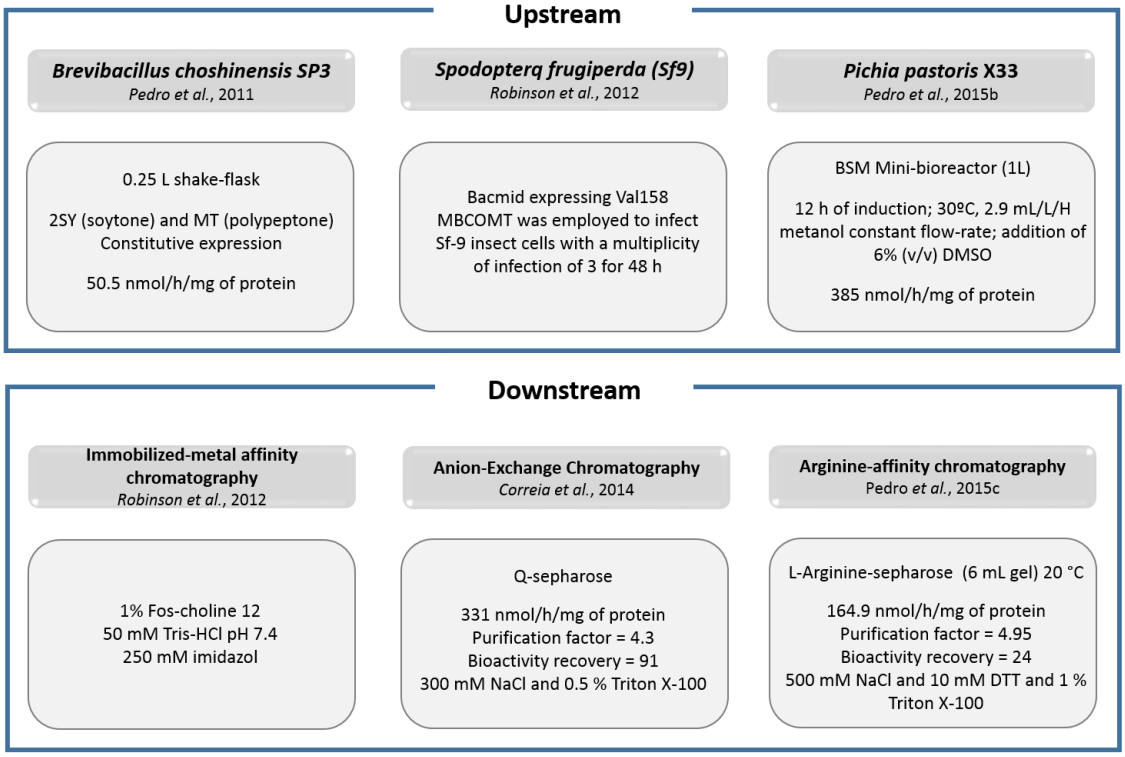


Figure 5 - Most recent and important strategies for MBCOMT biosynthesis (upstream) and purification (downstream) and general characteristics.

Abbreviations: BSM - Basal Salts Medium.

Up until now, several expression systems have been developed for efficient recombinant human biosynthesis. Recently, two eukaryotic-based expression systems using Sf-9 insect cells [Robinson *et al.*, 2012] or *P. pastoris* X33 cells [Pedro *et al.*, 2015b] were reported for MBCOMT expression. The optimization of the induction phase parameters (induction temperature, methanol flow-rate and dimethylsulfoxide concentrations) in *P. pastoris* mini-bioreactors cultures allowed an improvement of 6.4 fold in MBCOMT specific activity over the best results previously reported in shake-flasks [Pedro *et al.*, 2015a]. Directly coupled to the upstream stage, an alternative strategy to further increase functional and active MBCOMT biosynthesis may involve its recombinant production through CF expression systems. These systems have been described as particularly suitable for the expression of difficult proteins,

as they were previously reported as one of the most efficient systems for overexpression of human endothelin receptors and other GPCRs [Proverbio *et al.*, 2013]. Indeed, we believe it will be worthwhile to try to overexpress MBCOMT using a eukaryotic CF expression system based on wheat germ or insect cells and insert it in a co-translational form into pre-formed lipid bilayers such as liposomes or nanodiscs. Moreover, it should be explored not only the stabilizer potential of different sugars and polyols [Correia *et al.*, 2014] during the CF expression of MBCOMT but also the effect of SAM as well as the magnesium ion, the two cofactors that are essential to obtain catalytically active COMT [Bonifacio *et al.*, 2007].

Actually, there are four major reports concerning the chromatographic purification of MBCOMT by hydrophobic adsorbents [Santos *et al.*, 2013], using Q-sepharose as an anion-exchanger [Correia *et al.* 2014] or using affinity chromatography, either by Immobilized-Metal Affinity Chromatography (IMAC) [Robinson *et al.*, 2012] or arginine-affinity chromatography [Pedro *et al.*, 2015c]. Specifically, while the highest MBCOMT purity degree was obtained through IMAC, the use of tag sequences often interfere with further experiments or in the general properties of the target enzyme. Therefore, there is still room for the improvement concerning the design of a suitable chromatographic strategy for MBCOMT purification in a native, untagged form. Although uncommon, despite successful optimizations of the upstream stage, some MP are still in low abundance, requiring the application of highly specific and sensitive protocols for its enrichment. This is usually accomplished in immunoaffinity purification where a target antibody (called ligand) is immobilized into a matrix to form a solid absorbent [Zheng *et al.*, 2014]. Again, in contrast with the CF expression of MBCOMT, we believe that coupled with the optimization of the MBCOMT biosynthesis in mini-bioreactors using *P. pastoris* methanol-induced cultures [Pedro *et al.*, 2015b], it will be worthwhile to integrate this strategy with the development of a chromatographic matrix with an immobilized antibody anti-COMT.

The wide research conducted with recombinant MBCOMT allowed a deep characterization of its biochemical properties, catalytical mechanisms and the determination of its kinetic constants not only for different substrates but also for a wide range of inhibitor molecules already commercially available (3,5-dinitrocatechol, tolcapone, entacapone or opicapone). However, to consider its potential use as a therapeutic protein or to develop isoform-specific inhibitor molecules, its downstream processing will need to be improved and the suggestions previously reported may play an important role in this process.

3. Conclusions and Future Perspectives

Efforts to determine MP structure have been largely impaired by its recombinant expression, despite over the last years the progress in understanding MP biogenesis has triggered new insights in MP research. In some cases, proteins are inserted in the membranes at high levels while others are barely detected and as MP present diverse intrinsic properties, it becomes difficult or even impossible to predict which conditions are the most suitable to express a specific MP target. Actually, despite the increasing understanding of MP biogenesis, the development of upstream strategies are still conducted in some cases on a "trial and error" basis or alternatively using established protocols reported for other targets. Actually, *E. coli*, *P. pastoris* and mammalian cell lines are well-characterized and broadly reported microfactories for MP biosynthesis even though they are fairly different. *E. coli* is widely used due to the ease of manipulation, low cost, the immense availability of genetic tools and protocols and its capability to grow easily to high-cell densities. However, *E. coli* is unable to perform some of the PTM modifications that are usually required for the correct processing and insertion of MP. *Pichia pastoris* shares some of advantages of *E. coli* but as a eukaryote, is able to perform in a better extent some of the PTM performed by higher eukaryotes although the different membrane lipid compositions and the distinct glycosylation pattern are still cause for concerns. Mammalian cell lines are undoubtedly the best expression system to obtain properly folded and correctly inserted MP but it is way expensive and the complexity associated with cell culture and maintenance still limits its application as a widely used host. In general, there is no such a perfect host for MP overexpression and its choice remains largely empirical depending, among others, on the availability of specific equipment in each laboratory, the type of protein and the available budget. An alternative to *in vivo* systems that was rediscovered recently is the CF expression in which MP may be produced co-translationally with different kinds of artificial membrane environments. In fact, while these systems are particularly helpful for producing "difficult to express proteins" in which living cells have failed, there are specific aspects that should be taken into account to design a successful approach: origin of the target MP and type of CF. Also, since the co-translational expression in artificial membrane environments often leads to a decrease in yield, it must be considered only when there is no other option. The membrane-mimicking environment is also important in CF systems once some MP are active in nanodiscs but not with liposomes and the type of lipids required seem to be MP-dependent. Much research have been carried out with recombinant COMT and important achievements have been made over the years concerning either the development of appropriated strategies for its biosynthesis or the implementation of selective chromatographic methodologies for MBCOMT isolation. We consider MBCOMT a "difficult to express membrane protein" and this fact has encouraged the researchers worldwide to continuously search for the optimal conditions for its expression, purification and stabilization with tremendous achievements during the last three decades. Actually, we believe that MBCOMT CF expression in a co-translational form inserted into mimics of the lipid

bilayer may increase the production levels of functional and active MBCOMT. Also, the development of an immunoaffinity chromatographic methodology may aid in development of a biotechnological platform to obtain pure MBCOMT in a native active form.

Overall, the incredible achievements made in the last decades in understanding and overcoming the bottlenecks MP overexpression enabled the structural characterization of diverse MP. However, there is still room for improvement and new achievements will unveil new structures, allowing to develop new drugs for important pharmacological targets that are intimately linked to a large range of harmful human diseases.

4. References

Abad S, Kitz K, Hormann A, Schreiner U, Hartner FS, Glieder A. Real-time PCR-based determination of gene copy numbers in *Pichia pastoris*. Biotechnol J 2010; 5 (4): 413-20.

André N, Cherouati N, Prual C, Steffan T, Zeder-Lutz G, Magnin T, Pattus F, Michel H, Wagner R, Reinhart C. Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. Protein Sci 2006; 15: 1115-26.

Andréll J, Tate CG. Overexpression of membrane proteins in mammalian cells for structural studies. Mol Membr Biol 2013; 30 (1): 52-63.

Apud JA, Weinberger DR. Treatment of cognitive deficits associated with schizophrenia: potential role of catechol-*O*-methyltransferase inhibitors. CNS Drugs 2007; 21 (7): 535-57.

Aw R, Polizzi KM. Can too many copies spoil the broth? Microb Cell Fact 2013; 12: 128-37.

Bai J, Swartz DJ, Protasevich II, Brouillette CG, Harrel PM, Hildebrandt E, Gasser B, Mattanovich D, Ward A, Chang G, Urbatsch IL. A gene optimization strategy that enhances production of fully functional P-glycoprotein in *Pichia pastoris*. PLoS ONE 2011; 6 (8): 22577-92.

Bandaranayake AD, Almo SC. Recent advances in mammalian protein production. FEBS Lett 2014; 588 (2): 253-60.

Becker T, Vögtle FN, Stojanovski D, Meisinger C. Sorting and assembly of mitochondrial outer membrane proteins. Biochim Biophys Acta 2008; 1777 (7-8): 557-63.

Bergman-Jungestrom M, Wingren S. Catechol-*O*-methyltransferase (COMT) gene polymorphism and breast cancer risk in young women. Br J Cancer 2001; 85: 859-62.

Bernaudat F, Frelet-Barrand A, Pochon N, Dementin S, Hivin P, Boutigny S, Rioux JB, Salvi D, Seigneurin-Berny D, Richaud P, Joyard J, Pignol D, Sabaty M, Desnos T, Pebay-Peyroula E, Darrouzet E, Rolland N. Heterologous expression of membrane proteins: choosing the appropriate host. PLoS ONE 2011; 6 (12): 29191-208.

Bonifacio MJ, Archer M, Rodrigues ML, Matias PM, Learmonth DA, Carrondo MA, Soares-Da-Silva P. Kinetics and crystal structure of catechol-*O*-methyltransferase complex with co-substrate and a novel inhibitor with potential therapeutic application. Mol Pharmacol 2002; 62 (4): 795-805.

Bonifacio MJ, Palma PN, Almeida L, Soares-da-Silva P. Catechol-*O*-methyltransferase and its inhibitors in Parkinson's disease. CNS Drug Rev 2007; 13 (3): 352-79.

Carlson ED, Gan R, Hodgman CE, Jewett MC. Cell-free synthesis: Applications come of age. Biotech Adv 2012; 30 (5): 1185-94.

Celikel R, Veldore VH, Matthews I, Devine KM, Varughese KI. ATP forms a stable complex with the essential histidine kinase WalK (YycG) domain. Acta Crystallogr D Biol Crystallogr 2012; 68: 839-45.

Chaudhary S, Pak JE, Gruswitz F, Sharma V, Stroud RM. Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293S cells. Nat Protoc 2012; 7 (3): 453-66.

Chung BKS, Yusufi FNK, Mariati, Yang Y, Lee DY. Enhanced expression of codon optimized interferon gamma in CHO cells. J Biotechnol 2013; 167: 326-33.

Correia FF, Santos FM, Pedro AQ, Bonifácio MJ, Queiroz JA, Passarinha LA. Recovery of biological active catechol-*O*-methyltransferase isoforms from Q-sepharose. J Sep Sci 2014; 37: 20-9.

D'Urzo N, Martinelli M, Nenci C, Brettoni C, Telford JL, Maione D. High-level intracellular expression of heterologous proteins in *Brevibacillus choshinensis* SP3 under the control of xylose inducible promoter. Microb Cell Fact 2013; 12: 12-23.

Ferguson AD, McKeever BM, Xu S, Wisniewski D, Miller DK, Yamin TT, Spencer RH, Chu L, Ujjainwalla F, Cunningham BR, Evans JF, Becker JW. Crystal structure of inhibitor-bound human 5-lipoxygenase-activating protein. Science 2007; 317: 510-12.

Fliedl L, Kast F, Grillari J, Wieser M, Grillari-Voglauer R. Optimization of a quantitative PCR based method for plasmid copy number determination in human cell lines. N Biotechnol 2015; 32 (6): 716-19.

Freigassner M, Pichler H, Glieder A. Tuning microbial hosts for membrane protein production. Microb Cell Fact 2009; 8: 69-91.

Giacalone MJ, Gentile AM, Lovitt BT, Berkley NL, Gunderson CW, Surber MW. Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. Biotechniques 2006; 40: 355-64.

Goder V, Junne T, Spiess M. Sec61p contributes to signal sequence orientation according to the positive-inside rule. Mol Biol Cell 2004; 15: 1470-78.

Gonçalves AM, Pedro AQ, Maia C, Sousa F, Queiroz JA, Passarinha LA. *Pichia pastoris*: A recombinant microfactory for antibodies and human membrane proteins. J Microbiol Biotechnol 2013; 23 (5): 587-601.

Grisshammer R. Understanding recombinant expression of membrane proteins. Curr Opin Biotech 2006; 17: 337-40.

Grote A, Hiller K, Scheer M, Munch R, Nortemann B, Hempel DH, Jahn D. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acid Res 2005; 33: 526-31.

Gruswitz F, Chaudhary S, Ho JD, Schlessinger A, Pezeshki B, Ho CM, Sali A, Westhoff CM,Stroud RM. Function of human Rh based on structure of RhCG at 2.1A. PNAS 2010; 107: 9638-43.

Guerfal M, Ryckaert S, Jacobs PP, Ameloot P, Van Craenenbroeck K, Derycke R, Callewaert N. The HAC1 gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins. Microb Cell Fact 2010; 9 (49): 2859-71.

Gul N, Linares DM, Ho FY, Poolman B. Evolved *Escherichia coli* strains for amplified, functional expression of membrane proteins. J Mol Biol 2014; 426 (1): 136-49.

He M, He Y, Luo Q Wang M. From DNA to protein: No living cells required. Process Biochem 2011; 46: 615-20.

He Y, Wang K, Yan N. The recombinant expression systems for structure determination of eukaryotic membrane proteins. Protein Cell 2014; 5 (9): 658-72.

Henrich E, Hein C, Dotsch V, Bernhard F. Membrane protein production in *Escherichia coli* cell-free lysates. Febs Lett 2015; 589: 1713-22.

Hirz M, Richter G, Leitner E, Wriessnegger T, Pichler H. A novel cholesterol-producing *Pichia pastoris* strain is an ideal host for functional expression of human Na,K-ATPase α3β1 isoform. Appl Microbiol Biotechnol 2013; 97: 9465-78.

Horsefield R, Nordén K, Fellert M, Backmark A, Törnroth-Horsefield S, Terwisscha van Scheltinga AC, Kvassman J, Kjellbom P, Johanson U, Neutze R. High-resolution x-ray structure of human aquaporin 5. PNAS 2008; 105: 13327-32.

Hosak L. Role of the COMT gene Val158Met polymorphism in mental disorders: a review. Eur Psychiatry 2007; 22 (5): 276-81.

Hu S, Wang M, Cai G, He M. Genetic co-guided Protein synthesis and folding in *Escherichia coli*. J Biol Chem 2013; 288 (43): 30855-61.

Hunte C, Screpanti M, Venturi M, Rimon A, Padan E, Michel H. Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH. Nature 2005; 534: 1197-1202.

Ibrahimi OA, Eliseenkova AV, Plotnikov AN, Yu K, Ornitz DM, Mohammadi M. Structural basis for fibroblast growth factor receptor 2 activation in Apert syndrome. PNAS 2001; 98: 7182-87.
Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EYT, Lane JR, Ijzerman AP, Stevens RC. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science 2008; 322: 1211-17.

Kai L, Dotsch V, Kaldenhoff R, Bernhard F. Artificial environments for the co-translational stabilization of cell-free expressed proteins. PLoS ONE 2013; 8 (2): 56637-46.

Kim Y, Youn H, Na J, Hong K, Kang KK, Lee DS, Chung J. Codon-optimized human sodium iodide symporter (opt-hNIS) as a sensitive reporter and efficient therapeutic gene. Theranostics 2015; 5 (1): 86-96.

Klumpp S, Dong J, Hwa T. On Ribosome Load, Codon Bias and Protein abundance. PLoS ONE 2012; 7 (11): 48542-54.

Kozma D, Simon I, Tusnádi GE. PDBTM: Protein Data Bank of transmembrane proteins after 8 years. Nucleic Acids Res 2013; 41 (Database issue): D524-9.

Lantez V, Nikolaidis I, Rechenmann M, Vernet T, Noirclerc-Savoye M. Rapid automated detergent screening for the solubilization and purification of membrane proteins and complexes. Eng Life Sci 2015; 15: 39-50.

Laukens B, De Wachter C, Callewaert. Engineering the *Pichia pastoris* N-Glycosylation pathway using the GlycoSwitch technology. Methods Mol Biol 2015; 1321: 103-22.

Lee C, Kim J, Shin SG, Hwang S. Absolute and relative quantification of plasmid copy number in *Escherichia coli*. J Biotechnol 2006; 123: 273-80.

Liu J, Sun Y, Qi J, Chu F, Wu H, Gao F, Li T, Yan J, Gao GF. The membrane protein of severe acute respiratory syndrome coronavirus acts as a dominant immunogen revealed by a clustering region of novel functionally and structurally defined cytotoxic T-lymphocyte epitopes. J Infect Dis 2010; 202: 1171-80.

Liu J, Srinivasan P, Pham DN, Rozovsky S. Expression and purification of the membrane enzyme selenoprotein K. Protein Expr Purif 2012; 86 (1): 27-34.

Löw C, Jegerschöld C, Kovermann M, Moberg P, Nordlund P. Optimisation of over-expression in *E. coli* and biophysical characterization of human membrane protein synaptogyrin 1. PLoS ONE 2012; 7 (6): 38244-57.

Luo J, Choulet J, Samuelson JC. Rational design of a fusion partner for membrane protein expression in *E. coli*. Protein Sci 2009; 18: 1735-44.

Ma Z, Liu H, Wu B. Structure-based drug design of catechol-*O*-methyltransferase inhibitors for CNS inhibitors. Brit J Clin Pharmacol 2013; 77 (3): 410-20.

Mancia F, Patel SD, Rajala MW, Scherer PE, Nemes A, Schieren I, Hendrickson WA, Shapiro L. Optimization of protein production in mammalian cells with a coexpressed fluorescent marker. Structure 2004; 12: 1355-60.

Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK, Granier S. Crystal structure of the μ-opioid receptor bound to a morphinan antagonist. Nature 2012; 485: 321-6.

Marreddy RKR, Geertsma ER, Poolman B. Recombinant membrane protein production: past, present and future. Supramolecular Structure and function. Springer Netherlands 2011. Heidelberg. 41-74.

Marreddy RKR, Pinto JPC, Wolters JC, Geertsma ER, Fusetti F, Permentier HP, Kuipers OP, Kok J, Poolman B. The response of *Lactococcus lactis* to membrane protein production. PLoS ONE 2011; 6 (8): 24060-75.

Marino J, Hohl M, Seeger MA, Zerbe O, Geertsma ER. Bicistronic mRNA to enhance membrane protein overexpression. J Mol Biol 2015; 427 (4): 943-54.

Martins LM, Pedro AQ, Oppolzer D, Sousa F, Queiroz JA, Passarinha LA. Enhanced biosynthesis of plasmid DNA from *Escherichia coli* VH33 using Box-Behnken design associated to aromatic amino acids pathway. Biochem Eng J 2015; 98: 117-26.

Massey-Gendel E, Zhao A, Boulting G, Kim HY, Balamotis MA, Seligman LM, Nakamoto RK, Bowie JU. Genetic selection system for improving recombinant membrane protein expression in *E. coli*. Protein Sci 2009; 18: 372-83.

McMorran LM, Brockwell DJ, Radford SE. Mechanistic studies of the biogenesis and folding of outer membrane proteins in vitro and in vivo: What have we learned to date? Arch Biochem Biophys 2014; 564: 265-80.

Michnovicz JJ, Hershcopf RJ, Naganuma H, Bradlow HL, Fishman J. Increased 2-hydroxylation of estradiol as a possible mechanism for the anti-estrogenic effect of cigarette smoking. N Engl J Med 1986; 315 (21): 1305-9.

Midgett CR, Madden DR. Breaking the bottleneck: Eukaryotic membrane protein expression for high-resolution structural studies. J Struct Biol 2007; 160: 265-74.

Mueller PP, Wirth D, Unsinger J, Hauser H. Genetic Approaches to recombinant protein production in Mammalian cells. Handbook of Industrial Cell Culture- Mammalian, Microbial and Plant cells. Humana Press 2013. New York. 21-49.

Müller T. Catechol-*O*-methyltransferase inhibitors in Parkinson's disease. Drugs 2015, 75 (2): 157-74.

Myöhänen TT, Männistö PT. Distribution and functions of catechol-*O*-methyltransferase proteins: do recent findings change the picture? Int Rev Neurobiol 2010; 95: 29-47.

Nannenga BL, Baneyx F. Reprogramming chaperone pathways to improve membrane protein expression in *Escherichia coli*. Protein Sci 2011; 20: 1411-20.

Nordén K, Agemark M, Danielson JAH, Alexandersson E, Kjellborn P, Johanson U. Increasing gene dosage greatly enhances recombinant expression of aquaporins in *Pichia pastoris*. BMC Biotechnol 2011; 11: 47-59.

Noreen N, Hooi WY, Baradaran A, Rosfarizan M, Sieo CC, Rosli MI, Yusoff K, Raha AR. *Lactococcus lactis* M4, a potential host for the expression of heterologous proteins. Microb Cell Fact 2011; 10: 28-38.

Oberg F, Ekvall M, Nyblom M, Backmark A, Neutze R, Hedfalk K. Insight into factors directing high production of eukaryotic membrane proteins; production of 13 human AQPs in *Pichia pastoris*. Mol Membr Biol 2009; 26 (4): 215-27.

Ohsfeldt E, Huang S, Baycin-Hizal D, Kristoffersen L, Le TT, Li E, Hristova K, Betenbaugh MJ. Increased expression of the integral membrane proteins EGFR and FGFR3 in anti-apoptotic Chinese hamster ovary cell lines. Biotechnol Appl Biochem 2012; 59: 155-62.

Oldham ML, Chen J. Snapshots of the maltose transporter during ATP hydrolisis. PNAS 2011; 108: 15152-56.

Olson P, Zhang Y, Olsen D, Owens A, Cohen P, Nguyen K, Ye J, Bass S, Mascarenhas D. High-level expression of eukaryotic polypeptides from bacterial chromossomes. Protein Express Purif 1998; 14: 160-66.

Palomares AL, Estrada-Mondaca S, Ramírez OT. Production of Recombinant Proteins. Challenges and solutions. Methods Mol Biol 2004; 267: 15-52.

Pedro AQ, Bonifácio MJ, Queiroz JA, Maia CJ, Passarinha LA. A novel prokaryotic expression system for biosynthesis of recombinant human membrane-bound catechol-*O*-methyltransferase. J Biotechnol 2011; 156: 141-46.

Pedro AQ, Oppolzer D, Bonifácio MJ, Maia CJ, Queiroz JA, Passarinha LA. Evaluation of Mut^S and Mut^+ *Pichia pastoris* strains for membrane-bound catechol-*O*-methyltransferase biosynthesis. Appl Biochem Biotechnol 2015a; 175 (8): 3840-55.

Pedro AQ, Martins LM, Dias JML, Bonifacio MJ, Queiroz JA, Passarinha LA. An artificial neural network for membrane-bound catechol-*O*-methyltransferase biosynthesis with *Pichia pastoris* methanol-induced cultures. Microb Cell Fact 2015b; 14: 113-27.

Pedro AQ, Pereira P, Bonifácio MJ, Queiroz JA, Passarinha LA. Purification of membrane-bound catechol-*O*-methyltransferase by arginine-affinity chromatography. Chromatographia 2015c; DOI 10.1007/s10337-015-2970-3.

Proverbio D, Roos C, Beyermann M, Orbán E, Dotsch V, Bernhard F. Functional properties of cell-free expressed human endothelin A and endothelin B receptors in artificial membrane environments. Biochim Biophys Acta 2013; 1828: 2182-92.

Proverbio D, Henrich E, Orbán E, Dötsch V, Bernhard F. Chapter II: Membrane protein quality control in cell-free expression systems: tools, strategies and case studies in: Membrane Protein Production for structural analysis. 2014. Springer New York. 45-71.

Puigbò P, Guzmán E, Romeu A, Garcia-Vallvé S. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acid Res 2007; 35 (2): 126-31.

Rajesh S, Knowles T, Overduin M. Production of membrane proteins without cells or detergents. New Biotech 2011; 28 (3): 250-54.

Ramón A, Marín M. Advances in the production of membrane proteins in *Pichia pastoris*. Biotechnol J 2011; 6: 700-6.

Reenilä I, Männistö PT. Catecholamine metabolism in the brain by membrane-bound and soluble catechol-*O*-methyltransferase (COMT) estimated by enzyme kinetic values. Med Hypotheses 2001; 57 (5): 628-32.

Robinson RG, Smith SM, Wolkenberg SE, Kandebo M, Yao L, Gibson CR, Harrison ST, Polsky-Fisher S, Barrow JC, Manley PJ, Mulhearn JJ, Nanda KK, Schubert JW, Trotter BW, Zhao Z, Sanders JM, Smith RF, McLoughlin D, Sharma S, Hall DL, Walker TL, Kershner JL, Bhandari N,

Hutson PH, Sachs NA. Characterization of non-nitrocatechol pan and isoform specific catechol-*O*-methyltransferase inhibitors and substrates. ACS Chem Neurosci 2012; 3: 129-40.
Rodrigues ML, Bonifacio MJ, Soares-Da-Silva P, Carrondo MA and Archer M. Crystallization and preliminar X-ray diffraction studies if a catechol-*O*-methyltransferase/inhibitor complex. Acta Crystallogr Sect F Struct Biol Cryst Commun 2005; 61 (1): 118-20.

Roos C, Zocher M, Muller D, Munch D, Schneider T, Sahl HG, Scholz F, Wachtveitl J, Ma Y, Proverbio D, Henrich E, Dotsch V, Bernhard F. Characterization of co-translationally formed nanodiscs complexes with small multidrug transporters, proteorhodopsin and with the *E. coli* MraY translocase. Biochim Biophis Acta 2012; 1818: 3098-3106.

Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 2014; 5 (172): 1-17.

Santos FM, Pedro AQ, Soares RF, Martins R, Bonifácio MJ, Queiroz JA, Passarinha LA. Performance of hydrophobic interaction ligands for human membrane-bound catechol-*O*-methyltransferase purification. J Sep Sci 2013; 36: 1693-1702.

Sarramegna V, Muller I, Mousseau G, Froment C, Monsarrat B, Milon A, Talmont F. Solubilization, purification and mass spectrometry analysis of the human mu-opioid receptor expressed in *Pichia pastoris*. Protein Expr Purif 2005; 43 (2): 85-93.

Schlegel S, Lofblom J, Lee C, Hjelm A, Klepsch M, Strous M, Drew D, Slotboom DJ, de Gier JW. Optimizing membrane protein overexpression in the *Escherichia coli* Lemo21 (DE3). J Mol Biol 2012; 423: 648-59.

Schwartz D, Dötsch V, Bernhard F. Production of membrane proteins using cell-free expression systems. Proteomics 2008; 8: 3933-46.

Shukla S, Schwartz C, Kapoor K, Kouanda A, Ambudkar SV. Use of baculovirus BacMam vectors for expression of ABC drug transporters in mammalian cells. Drug Metab Dispos 2012; 40 (2): 304-12.

Sørensen HP. Towards universal systems for recombinant gene expression. Microb Cell Fact 2010; 9: 27-28.

Steinbacher S, Bass R, Strop P, Rees DC. Structures of the prokaryotic mechanosensitive channels MscL and MscS. Curr Top Membr 2007; 58.

Tate CG, Blakely RD. The effect of N-linked glycosilation on activity of the Na(+) and Cl(-)-dependent serotonin transporter expressed using recombinant baculovirus in insect cells. J Biol Chem 1994; 269: 26303-10.

Tate CG, Whiteley E, Betenbaugh MJ. Molecular chaperones stimulate the functional expression of the cocaine-sensitive serotonin transporter. J Biol Chem 1999; 274: 17551-58.

Voutilainen S, Tuomainen T, Korhonen M, Mursu J, Virtanen JK, Happonen P, Alfthan G, Erlund I, North KE, Mosher MJ, Kauhanen J, Tiihonen J, Kaplan GA, Salonen JT. Functional COMT Val158Met polymorphism, risk of Acute Coronary events and serum homocysteine: The Kuopio Ischaemic heart disease risk factor study, PLoS ONE 2007; 1: 181-87.

Wagner S, Bader ML, Drew D, de Gier J. Rationalizing membrane protein overexpression. Trends Biotechnol 2006; 24 (8): 364-71.

Welch M, Villalobos A, Gustafsson C, Minshull J. Designing genes for successful protein expression. Methods Enzymol 2009; 498: 43-66.

Weyand S, Shimamura T, Yajima S, Suzuki S, Mirza O, Krusong K, Carpenter EP, Rutherford NG, Hadden JM, O'Reilly J, Ma P, Saidijam M, Patching SG, Hope RJ, Norbertczak HT, Roach PC, Iwata S, Henderson PJ, Cameron AD. Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. Science 2008; 322: 709-13.

White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, Gvozdenovic-Jeremic J, Shah P, Shiloach J, Tate CG, Grisshammer R. Structure of the agonist-bound neurotensin receptor. Nature 2012; 490: 508-13.

Wu W, Wu Q, Hong X, Xiong G, Xiao Y, Zhou J, Wang W, Wu H, Zhou L, Song W, Dai H, Qiu H, Zhao Y. Catechol-*O*-methyltransferase inhibits colorectal cancer cell proliferation and invasion. Arch Med Res 2015; 46 (1): 17-23.

Xiao S, Chen YC, Betenbaugh MJ, Martins SE, Shiloach J. MiRNA mimic screen for improved expression of functional neurotensin receptor from HEK 293 cells. Biotechnol Bioeng 2015; DOI 10.1002/bit.25567.

Yager JD. Catehcol-*O*-methyltransferase: characteristics, polymorphisms and role in breast cancer. Drug Discov Today Dis Mech 2012; 9 (1-2): 41-6.

Zheng X, Dong S, Zheng J, Li D, Li F, Luo Z. Expression, stabilization and purification of membrane proteins via diverse protein synthesis systems and detergents involving cell-free associated with self-assembly peptide surfactants. Biotech Adv 2014; 32: 564-74.

Zhu BT, Liehr JG. Inhibition of catechol-*O*-methyltransferase-catalyzed *O*-methylation of 2- and 4-hydroxyetsradiol by catecholamine: implications for the mechanism of estrogen-induced carcinogenesis. Arch Biochem Biophys 1993; 304 (1): 248-56.

Zhu BT. On the mechanism of homocysteine pathophysiology and pathogenesis: a unifying hypothesis. Histol Histopathol 2002; 17 (4): 1283-91.

Zuo X, Li S, Hall J, Mattern MR, Tran H, Shoo J, Tan R, Weiss SR, Butt TR. Enhanced expression and purification of membrane proteins by SUMO fusion in *Escherichia coli*. J Struct Funct Genomics 2005; 6: 103-11.

3. Recombinant protein biosynthesis in *Pichia pastoris*

From:

Pichia pastoris: a Recombinant microfactory for antibodies and human membrane proteins

A. M. Gonçalves, A. Q. Pedro, C. Maia, F. Sousa, J. A. Queiroz, L. A. Passarinha

Journal of Microbiology and Biotechnology. 2013. 23 (5): 587-601

3. Recombinant protein biosynthesis in *Pichia pastoris*

During the last few decades, it has become evident that the compatibility of the yeast biochemical environment with the ability to process and translate the RNA transcript, along with its capacity to modify a translated protein, are relevant requirements for selecting this host cell for protein expression in several pharmaceutical and clinical applications. In particular, *P. pastoris* is used as an industrial host for recombinant protein and metabolite production, showing a powerful capacity to meet required biomolecular target production levels in high-throughput assays for functional genomics and drug screening.

3.1. *Pichia pastoris* as a recombinant microfactory

As a eukaryotic fungal organism, *Pichia* is easy to manipulate and cultivate. Initially used by Philips Petroleum Company for the production of single cellular proteins, this organism is nowadays a nonconventional yeast in which proteins have been successfully expressed in laboratory and large scale fermentation procedures to produce recombinant proteins [Cereghino and Cregg, 1999; Wegner, 1990]. Its unique features, which we will discuss below, combined with the possibility of designing suitable flowsheets for each specific protein, make *Pichia* a microfactory to be reckoned with. Being an organism that is easy to manipulate, *P. pastoris* can reach high levels of expression and produce complex biomolecules that need to undergo post-translational modification. Since *Pichia* has no native plasmids, the expression of foreign genes is achieved by chromosomal integration, using integrative plasmids or autonomous vectors. The integration of a foreign gene expression cassette can occur via gene insertion or gene replacement. For optimal genetic stability, cleavage of a *P. pastoris* vector within a sequence shared by the host genome stimulates homologous recombination events that efficiently target integration of the vector into a specific genomic locus, Alcohol Oxidase (AOX) 1 [Cregg and Maden, 1987]. These results in the site-specific eviction of the AOX1 structural gene, affording a strain that grows slowly in metanol (Mut^S). Alternatively, the vector can be linearized and targeted to integrate into the genome by non-disruptive insertion into HIS4, generating strains that grow normally in methanol (Mut^+) [Daly and Hearn, 2005; Domíngez *et al.*, 1998]. An understanding of the phenotype is important in deciding upon the culture conditions to be used. *Pichia* is a yeast capable of metabolizing methanol. A unique set of enzymes, alcohol oxidase, catalase, and dihydroxyacetone synthase, locked in peroxisomes, enable the yeast to convert methanol as a source of carbon and energy [Faber *et al.*, 1995]. The AOX enzyme is very important since it is responsible for the initial step, the oxidation of methanol to formaldehyde. There are two genes encoding AOX, aox1 and aox2, but the former is responsible for the vast majority of alcohol oxidase activity in the cell. Its regulation is similar to that of gal1 in *Sacharomyces cerevisiae*, involving two mechanisms: repression/derepression and induction. However, unlike gal1 regulation, the absence of a repressing carbon source such as glucose in the medium does not result in substantial

transcription of aox1. However, methanol is essential to the induction of high levels of transcription. 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 (Table 6) [Cereghino *et al.*, 2001; Faber *et al.*, 1995].

Table 6 – *Pichia pastoris* alternative promoters to AOX1.

Promoter	Features
GAP	Does not require methanol for induction, nor is it necessary to shift cultures from one carbon source to another, making strain growth more straightforward. Is expressed constitutively, so is not a good choice for expression of proteins that may be toxic to the cell.
FLD1	Can be induced by methanol as the sole carbon source (with ammonium sulfate as a nitrogen source), or Methylamine as the sole nitrogen source (with glucose as a carbon source). Is repressed with glucose and ammonium sulfate. Offers the flexibility of inducing high levels of expression using either methanol or methylamine, an inexpensive nontoxic nitrogen source.
PEX8	Moderate expression levels are desirable. Expressed at low but significant level on glucose. Induced modestly by methanol.
YPT7	Moderate expression levels are desirable. Provides a low but constitutive level of expression in media containing glucose, methanol, or mannitol as carbon sources.

Fundamentally, five promoters are used for recombinant MP production in *P. pastoris*. Of these, three are inducible by methanol (AOX1, FLD1, and PEX8), and the other two are constitutive (GAP and YPT1) [Zhang *et al.*, 2009]. Most strains grow on methanol at the wild-type rate (Mut^+); however, cells with the Mut^S and Mut^- phenotypes, because of the deletion in one or both AOX genes, requires less methanol to induce expression, which is an advantage in large-scale fermentations, where large quantities of methanol are considered a significant hazard [Macauley-Patrick *et al.*, 2005].

3.2. Bioprocess design: medium formulation and operation conditions

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 to the bioprocess has a great influence on the final result and quality of the target protein. Hence, in this section, we present an overview of the current strategies used for designing and optimizing a bioprocess using *P. pastoris* for MP and antibody heterologous overexpression. The culture medium composition exerts an effect on heterologous protein production in yeast through changes in cell growth and viability [Li *et*

al., 2007]. In general, all *P. pastoris* strains grow on defined medium supplemented with specific components for each strain. For instance, whereas the wildtype X-33 *P. pastoris* strain grows in minimal medium, strains GS115 and KM71H only grow on minimal medium supplemented with histidine and arginine, respectively [Li *et al.*, 2007].

For high cell density growth and induction of *P. pastoris* strains, the culture medium composition is similar to those described by Philips Petroleum Company in the 1970s [Cereghino and Cregg, 2000]. Essentially, the medium consists of biotin, ammonium hydroxide as the nitrogen source (also used to control the pH), glycerol or methanol as, respectively, the carbon and energy source, basal salts medium (BSM), and trace elements (zinc chloride, ferrous sulfate heptahydrate, and biotin being the most abundant constituents) [Celik and Calik, 2012; Cereghino and Cregg, 2000]. Alternative formulations to the standard BSM were proposed by Stratton and coworkers (the so-called FM22) [Stratton *et al.*, 1998] as well as d'Anjou and Daugulis [D'Anjou and Daugulis, 2000]. However, common problems associated with these formulations were the precipitation of one or more salts of the BSM during media preparation, changing the effective concentration of the dissolved minerals in the medium as well as turbidity, thus compromising cell density assessments [Cereghino and Cregg, 2000]. Nevertheless, a solution to this problem was presented by Oehler [1998] and Brady [2001]. Oehler and collaborators presented an alternative medium in which sodium hexametaphosphate, a non-precipitate-forming compound, replaced phosphoric acid as a source of phosphate [Oehler *et al.*, 1998], and Brady proposed the reduction of the basal salts concentrations to one quarter of that recommended [Brady *et al.*, 2001] (see Table 7 for the detailed composition of BSM preparations [Brady *et al.*, 2001; Invitrogen Co. 2005]). This alteration has proven to be effective since no adverse effects on cell growth rate, biomass yield, or protein expression levels have been detected.

Table 7 – Fermentation basal salts medium (BSM) proposed by Invitrogen Corporation

BSM component	Invitrogen Corporation	[Brady *et al.*, 2001]
Phosphoric acid 85%	2.67% (v/v)	----------
Calcium sulfate	0.93 g/L	0.23 g/L
Potassium sulfate	18.2 g/L	4.55 g/L
Magnesium sulfate heptahydrate	14.9 g/L	3.73 g/L
Potassium hydroxide	4.13 g/L	1.03 g/L
Glycerol	40 g/L	40 g/L
Sodium hexametaphosphate	----------	6.5 g/L

Regarding the production of heterologous proteins by *P. pastoris* in bioreactors, it is desirable to achieve high cell density cultures and, consequently, to attain high levels of recombinant protein. In fact, several strategies in fed-batch mode are available for protein production in *P. pastoris* strains in bioreactors under the control of AOX1 [Celik and Calik, 2012]. All of these strategies comprise three or four stages. First, the strains are batch cultured in a defined medium containing glycerol (or another repressing carbon source) to generate biomass but repress heterologous gene expression [Celik and Calik, 2012; Cereghino *et al.*, 2002; Cregg *et al.*, 2000]. Usually, in this stage, the feed rate is set to a growth limited level to avoid glycerol accumulation where the feed time will depend on the desired cell density. The second stage is a fed-batch transition phase in which more glycerol is added to the culture but at a growth-limiting rate in order to increase the biomass even more and to derepress the cells for induction [Cereghino *et al.*, 2002; Cregg *et al.*, 2000]. The last stage is the addition of methanol to the culture at a slow rate to induce protein expression [Cereghino *et al.*, 2002; Cregg *et al.*, 2000]. In a four-stage process, an additional stage of batch-methanol addition is employed between the second and third stages (called transition phase), to prepare the cells prior to fed-batch operation [Celik and Calik, 2012].

As discussed in detail above, there are two phenotypically distinct strains, Mut^S and Mut^+, that differ in their capacity to use methanol as a carbon source. For Mut^+ strains, when we switch the medium composition to methanol, we must observe a dissolved oxygen spike to ensure that all of the glycerol of the fed-batch phase is consumed before initiating the transition phase [White *et al.*, 1995]. On the other hand, in the induction phase for Mut^S strains, an excess of methanol not exceeding 0.3% is recommended [Zhang *et al.*, 2000]. Typically, for Mut^+ *P. pastoris* strains, the methanol concentration in the medium used for induction ranges between 0.5% and 1% (v/v) and it is an important factor that requires optimization for each target protein since it affects cell growth as well as protein levels [Daly and Hearn, 2005]. Moreover, alternatives to the common co-feeding strategy with glycerol/methanol have also been reported. Çelik and collaborators [2009] reported that the addition of sorbitol as a co-substrate at the induction phase of methanol fed-batch fermentation by a *P. pastoris* Mut^+ strain increased the levels of recombinant human erythropoietin with no adverse effects on AOX activity. Specifically for Mut^S phenotype strains, it may be necessary to use alternative nonrepressing carbon sources such as sorbitol, mannitol, alanine, or trehalose even after induction [Çelik and Calik, 2009. Three strategies are commonly employed to keep the methanol concentration at optimal limits. In the first one, the methanol feeding rate is controlled according to the concentration in the culture medium, which is determined by gas chromatography [Li *et al.*, 2007]. Another strategy consists of controlling the dissolved oxygen content, since the methanol feeding rate increases with higher levels of dissolved oxygen [Li *et al.*, 2007]. The third strategy is related to the specific growth rate during the induction phase, but the kinetic growth model needs to be set up well [Li *et al.*, 2007]. In any event, the strategy has to be adapted when protein production is under the control of another

promoter such as GAP in which the genes are constitutively expressed. In these cases, the strains are cultured in a medium with an appropriate carbon source (glycerol, methanol, oleic acid, glucose, among others) by a fed-batch process [Zhang *et al.*, 2009].

Monitoring the methanol concentration is of critical importance in a *P. pastoris* process since high levels of methanol can be toxic to the cells and low levels may not be sufficient to initiate transcription [Macauley-Patrick *et al.*, 2005] Chromatographic methods such as gas chromatography and high performance liquid chromatography are the most common off-line methanol monitoring methods used, despite being expensive and time consuming. On the other hand, the on-line methods are generally based on the liquid-gas equilibrium and monitor methanol in the broth by analyzing the fermenter exhaust gas. Moreover, biomass is an important parameter that has to be controlled during the bioprocess. Hohenblum and coworkers [2003] revealed for the first time the usefulness of flow cytometry as a tool for the analysis and optimization of recombinant protein production processes in *P. pastoris*. They applied *P. pastoris* strain GS115 to human trypsinogen biosynthesis in a fed-batch mode and evaluated the viability as well as the product that remained associated with the cell wall with propidium iodide and immunofluorescent staining, respectively. Interestingly, the authors claim that the viability of the culture dropped in an early phase of the fed batch and finally fell to 65% at the end of the fermentation. Concerning product localization, the authors described that after starting the methanol feed, first the cells accumulate product, and then, nearly 20 h later, they begin to release product into the supernatant at a higher rate. In addition, the authors concluded that the dead cells retain the product they contain as long as they are not lysed; however, they will not contribute to newly synthesized product [Hohenblum *et al.*, 2003].

3.3. Typical fermentation flowsheets for human membrane proteins

The methylotrophic yeast *P. pastoris* has been used to produce crystallization-grade proteins for several MPs, from which high-resolution 3D structures have been determined [Bornet *et al.*, 2012; Ramon and Marin, 2011]. In fact, *P. pastoris* is able to produce proteins with all kinds of membrane spanning topologies, including enzymes, aquaporins, and ion channels [Bornet *et al.*, 2012]. As highlighted previously by Ramón and Marin [Ramon and Marin, 2011], the production of functional, properly folded and inserted MPs is often achieved by using alternative approaches that are almost always specific for each target MP. The use of the mammalian Kozak sequence, optimization of codon usage, co-expression of MPs with the Hac1p, N-terminal fusion of signal sequences to the MP, lower temperature cultivation, and the addition of chemical chaperones such as dimethylsulfoxide are some of the successful strategies that have been employed to improve MP expression in *P. pastoris* [70]. Moreover, a set of guidelines and instructions with several tips to overexpress MP in the *P. pastoris* system using G-Protein Coupled Receptor (GPCR) as a model was recently published by Bornet [2012].

Pichia pastoris is able to produce proteins both intracellularly and secreted into the culture medium. When the latter option is preferred, the presence of a signal sequence is required to target the protein to the secretory pathway [Cregg *et al.*, 2000]. A variety of secretion signal sequences including the *Sacharomyces cerevisiae* α-factor prepropeptide [Cregg *et al.*, 2000; Damasceno *et al.*, 2004], acid phosphatase signal sequence [Cregg *et al.*, 2000], SUC2 gene signal sequence [Paifer *et al.*, 1994], 128 kDa pGKL killer protein [Kato *et al.*, 2001], and hydrophobins HFBI and HFBII from *Trichoderma reesei* [Kottmeier *et al.*, 2011] have been applied to protein secretion in *P. pastoris*, but the Sacharomyces cerevisiae α-factor prepropeptide has been the most common and successful strategy [Cregg *et al.*, 2000].

However, most of the fused signal sequences do not allow the secretion of the target MP into the culture medium as happens for soluble proteins, but, in some cases, it may exert a positive effect on MP yield, as previously shown for the μ-opioid receptor [Ramon and Marin, 2011; Sarramegna *et al.*, 2005].

The most common methods used for *P. pastoris* cell lysis employ glass beads, microfluidizers, and a French press [Alisio and Mueckler, 2010; Asada *et al.*, 2011; Invitrogen Co 2005], which are the more suitable choices for large amount of cells. Usually, MPs are produced intracellularly, in which case the membranes need to be isolated. This is accomplished by performing an ultracentrifugation step, typically at 200.000 ×g for 60 min at 4ºC or 30 min at 4ºC after cell lysis [Alisio and Mueckler, 2010; Asada *et al.*, 2011; Bornet *et al.*, 2012].

Construction of the target MP with an affinity tag allows purification by affinity chromatography [Alisio and Mueckler, 2010; Asada *et al.*, 2011; De Rivoyre *et al.*, 1996]. In fact, the application of affinity chromatography as the main tool for MP capture from total crude extract or total membrane fraction remains one of the most common and successfully employed techniques for the first step of MP purification. However, molecular exclusion and ion-exchange chromatographies have also been successfully employed [Craveiro *et al.*, 2006].

The main aim of developing an expression system based on *P. pastoris* strains for overexpression of an MP is to obtain recombinant MP of high quantity and purity that allows structural and functional studies. To date and according to the Membrane Protein Data Bank, 49 structures of membrane proteins have been solved using *P. pastoris* as the host for MP production [Raman *et al.*, 2006], making this microfactory a good option for performing heterologous protein expression for structural studies of MP [Bornet *et al.*, 2012].

3.4. References

Asada HT, Uemura T, Yurugi-Kobayashi M, Shiroishi T, Shimamura H, Tsujimoto H, Ito K, Sugawara T, Nakane T, Nomura N, Murata T, Haga T, Iwata S, Kobayashi T. Evaluation of the *Pichia pastoris* expression system for the production of GPCRs for structural analysis. Microb Cell Fact 2011; 10: 24-37.

Alisio A, Mueckler M. Purification and characterization of mammalian glucose transporters expressed in *Pichia pastoris*. Protein Expr Purif 2010; 70: 81-7.

Bornet O, Alkhalfioui F, Logez C, Wagner R. Overexpression of membrane proteins using *Pichia pastoris*. Curr Prot Protein Sci 2012; 29 (29.2).

Brady CP, Shimp RL, Miles AP, Whitmore M, Stowers AW. High-level production and purification of P30P2MSP1, an important vaccine antigen for malaria, expressed in the methylotrophic yeast *Pichia pastoris*. Protein Expr Purif 2001; 23: 468-75.

Celik E, Calik P. Production of recombinant proteins by yeast cells. Biotechnol Adv 2012; 30 (5): 1108-18.

Celik E, Calik P, Oliver SG. Fed-batch methanol feeding strategy for recombinant protein production by *Pichia pastoris* in the presence of co-substrate sorbitol. Yeast 2009; 26 (9): 473-84.

Cereghino GP, Cregg JM. Applications of yeast in biotechnology: Protein production and genetic analysis. Curr Opin Biotechnol 1999; 10: 422-27.

Cereghino GP, Cereghino JL, Sunga A. New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of *Pichia pastoris*. Gene 2001; 263: 159-69.

Cereghino GP, Cereghino JL, Ilgen C, Cregg JM. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. Curr Opin Biotech 2002; 13: 329-32.

Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol Rev 2000; 24: 45-66.

Craveiro RB, Ramalho JD, Chagas JR, Wang PHM, Casarini DE, Pesquero JL, Araújo RC, Pesquero JB. High expression of human carboxypeptidase M in *Pichia pastoris*: Purification and partial characterization. Braz J Med Biol Res 2006; 39 (2): 211-17.

Cregg JM, Madden KR. Development of yeast transformation systems and construction of methanol-utilizationdefective mutants of *Pichia pastoris* by gene disruption. Biol Res Ind Yeast 1987; 2: 1-18.

Cregg JM, Cereghino JL, Jianying S, Higgins D. Recombinant protein expression in *Pichia pastoris*. Mol Biotechnol 2000; 16: 23-52.

Daly R, Hearn MTW. Expression of heterologous proteins in *Pichia pastoris*: A useful experimental tool in protein engineering and production. J Mol Recognit 2005; 18: 119-38.

D'Anjou MC, Daugulis AJ. Mixed-feed exponential feeding for fed-batch culture of recombinant methylotrophic yeast. Biotechnol Lett 2000; 22: 341-46.

Damasceno LM, Pla I, Chang HJ, Cohen L, Ritter G, Old LJ, Batt CA. An optimized fermentation process for high-level production of single-chain Fv antibody fragment in Pichia pastoris. Protein Expr Purif 2004; 37: 18-26.

Damasceno LMC, Huang J, Batt C. Protein secretion in *Pichia pastoris* and advances in protein production. Appl Microbiol Biotechnol 2012; 93: 31-9.

De Rivoyre D, Bonino F, Ruel L, Bidet M, Thérond P, Mus-Veteau I. Human receptor Smoothened, a mediator of Hedgehog signalling, expressed in its native conformation in yeast. FEBS Lett 1996; 579: 1529-33.

Domíngez A, Fermiñán E, Sánchez M, González FJ, Pérez-Campo FM, García S, Herrero AB, San Vicente A, Cabello J, Prado M, Iglesias FJ, Choupina A, Burguillo FJ, Fernaández-Lago L, López MC. Non-conventional yeast as hosts for heterologous protein production. Int Microbiol 1998; 1 (2): 131-42.

Faber K, Harder W, Ab G, Veenhuis M. Review: Methylotrophic yeast as factories for the production of foreign proteins. Yeast 1995; 11: 1331-44.

Hohenblum H, Borth N, Mattanovich D. Assessing viability and cell-associated product of recombinant protein producing *Pichia pastoris* with flow cytometry. J Biotechnol 2003; 102: 281-90.

Invitrogen Co. (Assessed September 13, 2015) *Pichia* Fermentation Process Guidelines. http://www.invitrogen.com/ 2005; Invitrogen Co., San Diego, CA, USA.

Kato S, Ishibashi M, Tatsuda D, Tokunaga H, Tokunaga M. Efficient expression, purification and characterization of mouse salivary a-amylase secreted from methylotrophic yeast, *Pichia pastoris*. Yeast 2001; 18: 643-55.

Kottmeier K, Ostermann K, Bley T, Rödel G. Hydrophobin signal sequence mediates efficient secretion of recombinant proteins in *Pichia pastoris*. Appl Microbiol Biotechnol 2011; 91: 133-41.

Li P, Anumanthan A, Gao XG, Ilangovan K, Suzara VV, Düzgüne N, Renugopalakrishnan V. Expression of recombinant proteins in *Pichia pastoris*. Appl Biochem Biotechnol 2007; 142: 105-24.

Macauley-Patrick S, Fazenda LM, McNeil B, Harvey LM. Heterologous protein production using the *Pichia pastoris* expression system. Yeast 2005; 22: 249-70.

Oehler R, Lesnicki G, Galleno M. High cell density fermentation of *Pichia pastoris* using nonphosphate precipitate forming sodium hexametaphosphate as a phosphate source. Current topics in gene expression annual meeting 1998; SanDiego, CA, USA.

Paifer E, Margolles E, Cremata J, Montesino R, Herera L, Delgado JM. Efficient expression and secretion of recombinant alpha amylase in *Pichia pastoris* using two different signal sequences. Yeast 1994; 10: 1415-19.

Raman P, Cherezov V, Caffrey M. The membrane protein data bank. Cell Mol Life Sci 2006; 63: 36-51.

Ramon A, Marin M. Advances in the production of membrane proteins in *Pichia pastoris*. Biotechnol J 2011; 6: 700-6.

Sarramegna V, Muller I, Mousseau G, Froment C, Monsarrat B, Milon A, Talmont F. Solubilization, purification, and mass spectrometry analysis of the human mu-opioid receptor expressed in *Pichia pastoris*. Protein Expr Purif 2005; 43: 85-93.

Stratton J, Chiruvolu V, Meagher M. High cell density fermentation. Biotech Adv 1998; 103: 107-20.

Wegner G. Emerging applications of the methylotrophic yeast. FEMS Microbiol Rev 1990; 7: 279-83.

White CE, Hunter MJ, Meininger DP, White LR, Komives EA. Large-scale expression, purification and characterization of small fragments of thrombomodulin - the role of sixth domain and methionine 388. Protein Eng Des Sel 1995; 8: 1177-87.

Zhang W, Inan M, Meagher MM. Fermentation strategies for recombinant protein expression in the methylotrophic yeast *Pichia pastoris*. Biotechnol Bioprocess Eng 2000; 5: 275-87.

Zhang A, Luo J, Zhang T, Pan Y, Tan Y, Fu C, Tu F. Recent advances on the GAP promoter derived expression system of *Pichia pastoris*. Mol Biol Rep 2009; 36: 1611-19.

Zhang J, Mizuno K, Murata Y, Koide H, Murakami M, Ihara K, Kouyama T. Crystal structure of deltarhodopsin-3 from *Haloterrigena thermotolerans*. Proteins 2013; 81: 1585-92.

Chapter 2

Global aims

The main scope of this work is the development of a straightforward approach that allows the biosynthesis, isolation and purification of recombinant human COMT isoforms in a biologically active form for further application in structural studies or to evaluate their role as potential therapeutic proteins. In fact, the development of a biotechnological platform that allows obtaining high quantities of pure COMT isoforms will allow the development of structural studies and it is clear the health benefits that may come from these studies concerning several human diseases, including neurologic disorders. On the other hand, the possible use of COMT as a recombinant therapeutic protein in some types of cancer will also depend on the successfully application of the aforementioned strategy. In general, the production and purification of membrane proteins are more challenging than soluble proteins since they are extracted from the lipid bilayer, their natural environment. Therefore, these differences should be taken into account when designing and implementing efficient and adequate strategies for their biosynthesis and purification. In this work, the methylotrophic yeast *P. pastoris* was selected as the expression host since it presents diverse advantages: is a GRAS (Generally Recognized As Safe) organism, high protein productivities are usually achieved in cheap media and it is capable of performing many post-translational modifications. Then, in order to separate the target proteins from the host cell proteins, affinity-chromatography strategies will be applied as the main isolation step.

To fulfill the main aim of this thesis and to efficiently guide the experimental work developed throughout this project, intermediate tasks were set:

1. Development, implementation and validation of an analytical method based on the use of reversed-phase HPLC with coulometric detection for the quantification of metanephrine, an *O*-methylated catechol product in COMT enzymatic assays with higher sensitivity and low chromatographic runs.
2. Cloning separately the human soluble and membrane-bound COMT wild-type genes (polymorphism Val108/158 for soluble and membrane-bound COMT, respectively) in the expression vector pPICZα and transformation of several *P. pastoris* strains with the different constructs.
3. Establishment of an appropriated flowsheet for biosynthesis of both COMT isoforms at a small-scale onto baffled shake-flasks or at a medium-scale using mini-biorreactors with *P. pastoris* methanol-induced cultures.
4. Integration of the innovative biosynthesis strategy with an equally suitable recovery step, depending on the preferential location of each isoform. Study and implementation of novel affinity chromatographic strategies for the purification of recombinant COMT from *P. pastoris*.

Chapter 3

Paper II

An improved HPLC method for quantification of metanephrine with coulometric detection

A. Q. Pedro, R. F. Soares, D. Oppolzer, F. M. Santos, L. A. Rocha, A. M. Gonçalves, M. J. Bonifácio, J. A. Queiroz, E. Gallardo, L. A. Passarinha

Journal of Chromatography and Separation Techniques. 2014. 5 (2): 217-224

Short description: This article is an original research that reports an improved reversed-phase HPLC method with coulometric detection for the quantification of metanephrine, an *O*-methylated product in COMT enzymatic assays. This method presents several advantages over other previously published, namely the shorter chromatographic runs, higher sensitivity and lower detection limit for metanephrine. Finally, the method reported in this paper is widely applied in the subsequent papers, whenever COMT enzymatic activity is evaluated.

Chromatography
Separation Techniques

Pedro et al., J Chromatograph Separat Techniq 2014, 5:2
http://dx.doi.org/10.4172/2157-7064.1000217

Research Article Open Access

An Improved HPLC Method for Quantification of Metanephrine with Coulometric Detection

Pedro AQ[1#], Soares RF[1#], Oppolzer D[1], Santos FM[1], Rocha LA[1], Gonçalves AM[1], Bonifacio MJ[2], Queiroz JA[1], Gallardo E[1] and Passarinha LA[1*]

[1]*CICS-UBI-Centro de Investigacaoem Ciencias da Saude, Universidade da Beira Interior, 6201-506 Covilha, Portugal*
[2]*Departamento de Investigacao e Desenvolvimento, BIAL 4745-457 Sao Mamedeo Coronado, Portugal*
[#]*Authors contributed equally to this work*

Abstract

A rapid and straightforward analytical method, based on the use of RP-HPLC with coulometric detection, was developed and validated for the quantification of metanephrine, an O-methylated product in catechol-O-methyltransferase enzymatic assays. The isocratic separation was achieved on a reverse column with a mobile phase consisting of 0.1 M sodium dihydrogen phosphate, 0.024 M citric acid monohydrate, 0.5 mM sodium octyl sulphate and 9% acetonitrile (%v/v). The method was found to be linear between 0.25 and 15 nmol/mL with a determination coefficient of 0.9997 for metanephrine. Intra-and interday precision and accuracy were in conformity with the criteria accepted in bioanalytical method validation and the LOD and LLOQ were 0.25 nmol/mL. The main focus of the developed method is the lower LLOQ achieved that can have important implications in laboratory research for COMT activity determinations, in particular for the methionine 108/158 variant obtained either from native or recombinant extracts. Another major advantage of the present method is the shorter run times on automated chromatographic systems that allow the analysis of several samples in a short time.

In addition, metanephrine was stable in the samples for at least 24 h at room temperature, for at least 24 h in HPLC system injector and for at least three freeze/thaw cycles. The developed method demonstrated higher sensitivity, precision, accuracy, stability, and linearity when compared with the methods previously described. Finally, a catechol-O-methyltransferase activity assay, resulting in an O-methylated reaction product, was used to evaluate the method applicability.

Keywords: HPLC; Coulometric detection; Metanephrine; Catechol-O-methyltransferase.

Introduction

Normetanephrine (NMN) and metanephrine (MN), O-methylated metabolites of norepinephrine and epinephrine respectively, are produced by the actions of catechol-O-methyltransferase (COMT) (EC 2.1.1.6), an enzyme largely confined to extraneuronal tissues [1], that needs accurate and selective measurements, not only for clinical diagnosis but also for pathological studies of several diseases. In fact, as the catechol-metabolizing system that comprises COMT has a potential pathophysiological and pathogenic significance in several disorders [2,3], it becomes important to the determination of COMT biological activity and determination of metanephrine with lower detection limits and improved sensitivity.

MN quantification, after chemical or enzymatic hydrolysis of the conjugated forms, is still important for diagnosing neural crest tumors, particularly in differentiating between pheochromocytoma and hypertension [4]. Several analytical methods, in particular, HPLC, have been proposed for the analysis of catecholamines and their metabolites (O-methylated reaction products) in biological fluids. Several detecting techniques are depicted in bibliography, namely MS [5], UV spectrophotometry [6], Fluorometry [7] , RIA [8] and chemical luminescence [9].

Liquid chromatography coupled with electrochemical detection (LC-ECD) has provided a new tool to evaluate the levels of these compounds in urine [10], plasma [11] and COMT assays [12] and is considered a reliable technique for catecholamine assay [13]. Typically, catecholamines and their metabolites can be separated by RP HPLC systems with ion-pairing reagents or by ion-exchange HPLC and detected by their reversible oxidation by amperometric [11,12] or coulometric carbon-based working electrodes [14,15]. Analysis of electroactive compounds, like catecholamines, in biological samples/extracts by HPLC with electrochemical detectors in coulometric mode has gained more interest in the last years. As a matter of fact, ECD has been improved since the appearance of the flow-through electrodes (coulometric detectors) that can react with near of 100% of the electroactive components of the analyte. Recently, developed coulometric sensors provide selectivity, identification, and resolution of compounds when used as detectors in HPLC systems. Coulometric detectors claim a better performance for sensitivity and selectivity than the classic amperometric detector, making the coulometric detection an advanced tool capable of addressing the analytical complexity of biological samples/extracts. In fact, coulometric detection in HPLC ECD, a technique in which all of the analyte in the column effluent is oxidized or reduced at the surface of an electrode at constant potential, offer certain advantages over amperometric detection, in which only a few percent of the analyteis converted [16]. A major concernin amperometric HPLC ECD is the decrease of electrode response in time, which is often attributed to the reduction of the active area by adsorption phenomena. On the other hand, when the electrode area of a coulometric detector is large enough, the conversion efficiency may still be 100% despite some loss of active area, so that the response is unaffected [16]. Coulometric detection has been used in detection of

***Corresponding author:** Passarinha LA, Health Sciences Research Centre - University of Beira Interior (CICS-UBI), Av. Infante D. Henrique, 6200-506, Covilha, Portugal, Tel: +351 275 329 069; Fax: +351 275 329 099; E-mail: lpassariha@fcsaude.ubi.pt

Received March 12, 2014; **Accepted** April 14, 2014; **Published** April 18, 2014

Citation: Pedro AQ, Soares RF, Oppolzer D, Santos FM, Rocha LA, et al. (2014) An Improved HPLC Method for Quantification of Metanephrine with Coulometric Detection. J Chromatograph Separat Techniq 5: 217. doi:10.4172/2157-7064.1000217

clinical important substances in Plasma [17], Saliva [18], Urine [19], including catecholamine detection [20].

The sensitivity limitations imposed by the large areas of coulometric (>99% efficient) detectors and the theoretical advantages of 100% detection compared to the 2-5% detection achieved in amperometric sensors, have been recognized with detection limits inside the femtogram (10-15) region [15]. In coulometric detection, 100% oxidation efficiency at one electrode of the analyte and their reduction at a second electrode allow a great increase in selectivity and sensitivity for compounds with reversible oxidation over non-reversible compounds, such as MN, a typical O-methylated product of COMT in the presence of epinephrine. Furthermore, COMT has been described as an important drug target in Parkinson's disease [2], emphasizing the importance of the development of methods with improved LLOQ that allow the quantification of lower MN levels and, consequently, the determination of low enzyme activity levels.

In this work, we describe a novel method using HPLC with coulometric detection, with adequate sensitivity for reliable measurement of MN in biological bacteria lysate extracts for the first time, by calibration procedures, including intra- and interday precision and stability. Moreover, this method was fully validated in a wide concentration range with lower limit of quantification. We evaluated the method applicability through in vitro membrane COMT activity assays, using epinephrine as substrate, resulting in an O-methylated reaction product such as MN. The method reported here can be applied to clinical MN assessment on biological fluids (plasma, urine) after suitable extraction procedures due to the high sensitivity and speed shown, associated to a greater selectivity of the coulometric detectors.

Experimental

Reagents and standards

Analytical standards of MN (DL-Metanephrine Hydrochloride), citric acid monohydrate and 1-Octanesulfonic acid were purchased from Sigma Aldrich (Steinheim, Germany). Acetonitrile (HPLC grade) and sodium dihydrogen phosphate were obtained from Fisher Scientific (Leicestershire, UK) and USB Corporation (Ohio, USA), respectively. Perchloric acid was purchased from Panreac (Barcelona, Spain). Deionized water was obtained from a Millipore purification system in our research laboratory. Stock solution at 25 nmol/mL of MN was prepared in 0.2 M perchloric acid. Standard solutions were obtained by diluting stock solution with the same solution of perchloric acid. All these buffers were stored and protected from light at 4°C and were stable for at least three months.

Instrumentation

Chromatographic analysis was performed using a HPLC model Agilent 1260 system (Agilent, Santa Clara, California, USA) equipped with an autosampler and quaternary pump coupled to an ESA Coulochem III (Milford, Massachussets, USA) coulometric detector. Chromatographic separation was achieved on an analytical column Zorbax 300SB C_{18} RP analytical column (250×4.6 mm i.d. 5 µm) (Agilent, Santa Clara, California, USA). The mobile phase (0.1 M sodium dihydrogen phosphate, 0.024 M citric acid monohydrate, 0.5 mM sodium octyl sulphate and 9% acetonitrile, v/v), pH 2.9, was filtered under vacuum (0.2 µM hydrophilic polypropylene filter), degassed in ultrasonic bath before use. Under this procedure the mobile phase was maintained stable during 2 weeks. Column effluent was monitored with an electrochemical detector in the coulometric mode, which was equipped with a 5011 high sensitivity dual electrode analytical cell (electrodes I and II) using a procedure of oxidation/reduction (analytical cell #1: +410 mV; analytical cell #2: -350 mV). The high surface area of electrodes I and II results in a 100% reaction of the electroactive compound (MN). The method sensitivity was set at 1 µA and the flow rate applied was 1mL/min. Column temperature was optimized to 30°C. The chromatograms were obtained by monitoring the reduction signal of the working electrode II. The retention time was around 8.8 min for MN.

Standards preparation

A stock solution of MN (25 nmol/mL) was prepared by dissolving the appropriate salt in 0.2 M perchloric acid. The initial solution was diluted in 0.2 M perchloric acid in order to obtain MN standard solutions at several concentrations (0.25 to 15 nmol/mL). The standard samples were agitated by rotation/inversion movements for 2 min and injected into the HPLC-ECD system according to the experimental conditions.

Validation procedure

The procedure was validated in terms of selectivity, linearity, intra- and inter-day precision, accuracy, and stability. Calibration data was generated by spiking samples and the calibration curve was established between 0.25 and 15 nmol/mL (eight calibrators evenly distributed). Five calibration curves were prepared, and the criteria for acceptance included a R^2 value of at least 0.99, and the calibrators accuracy within a ± 15% interval, except at the lower LLOQ (LLOQ), for which ± 20% was accepted. The LLOQ was defined as the lowest amount of analyte that presented a signal-to-noise ratio of at least 5 and could be measured with adequate precision and accuracy (coefficient of variation of less than 20% and an accuracy of ± 20%). Intraday precision was characterized in terms of RSD (%) by analyzing sets of 5 MN samples at five different concentrations (0.25, 2, 4, 8 and 12 nmol/mL) over a 5-day period. Interday precision was assessed at eight concentrations (0.25, 0.5, 1, 2, 4, 8, 12, 15 nmol/mL) over a 5-day period. Accuracy was evaluated in terms of mean relative error between the measured and the spiked concentrations for the calibrators and also in the intra- and interday precision assays; the limits of acceptable variability were set at 15% for all concentrations, except at the LLOQ, for which 20% was accepted. Processed sample-stability, short-term stability and freeze/thaw stability were studied (n=3) at two concentration levels (3 and 10 nmol/mL).

MN samples were subjected to different storage conditions, and the obtained results were compared to those achieved after analysis of freshly prepared samples. The compound was considered stable under the tested conditions if the coefficients of variation between the two sets of samples were less than 15%.

Membrane-bound catechol-O-methyltransferase expression and enzymatic assay

Recombinant human membrane-bound catechol-O-methyltransferase (hMBCOMT) biosynthesis was carried out according to the procedure previously described by Pedro et al. [3]. Cells were grown in a mineral salt medium supplemented with soytone 1% (w/v) for 50 hrs at 30°C and 120 rpm. The experiments of activity were designed to evaluate the methylation efficiency of recombinant hMBCOMT, by measuring the amount of MN, using epinephrine as substrate, as previously described by Passarinha et al. [12]. In hMBCOMT activity assay, a 2000 µg/mL aliquot of the membrane extract after suitable solubilization, was incubated in 5 mM sodium phosphate buffer (pH 7.8) containing 0.2 µM $MgCl_2$, 2 mM EGTA, 250

µM SAM e and 1 mM epinephrine in a total sample volume of 1 mL. Reactions were carried out at 37°C during 15 min and were stopped by incubation in ice following the addition of 2M 200 µl perchloric acid. The supernatants were centrifuged, filtered and subsequently injected into the HPLC system.

Results and Discussion

The methodology presented in this work concerns the development and validation of an HPLC method with coulometric detection for assessment of MN, allowing the determination of the biological activity of recombinant COMT as well as native COMT extracted from animal tissues or cell lines. This method was validated using authentic samples obtained from a bioprocess intended to synthesize the membrane-bound isoform of COMT [3]. On the other hand, in humans, the activity of COMT can be distributed into three classes with high, intermediate and low activity groups in which such difference is correlated with a functional polymorphism at codon 108/158 (SCOMT/MBCOMT) involving a methionine/valine substitution in the polypeptide chain [2]. While the metionine 108/158 variant is associated with low enzymatic activity [2], the development of HPLC methods with improved LLOQ and sensitivity capable of measuring lower quantities of MN seems to be of great importance not only to monitor bioprocesses by measuring the biological activity of the target recombinant COMT but also to measure the biological activity of low activity native COMT extracted from animal tissues or cell lines in different pathophysiological and pathogenic conditions.

Chromatographic methodology

The chromatographic conditions were chosen to allow selectivity on the basis of the ionic and hydrophobic characteristics of MN. Chromatographic parameters such as temperature and mobile phase (data not shown) were investigated in order to obtain a higher detection of MN within an acceptable time span. Cell potentials were optimized according to preliminary experiments (data not shown) and ESA Biosciences, Inc indications. The C18 stationary phase has been successfully employed for the separation and quantification of catecholamines and its metabolites, as previously shown by Passarinha et al. [12], Hollenbach and collaborators [21] and Unceta [22]. Therefore and based on these results, the C_{18} was applied as the stationary phase in this work.

In order to obtain better electrochemical oxidation efficiency, the larger electrode surface area coulometric cell, compared to the amperometric, results in a 99% reaction of our electroactive compound-MN. While the electrochemical behavior of MN is reversible at carbon electrodes, we used the oxidation/reduction mode in the coulometric detector because it plays an important role in improving both selectivity and sensitivity of our analysis. The electrode 1 (E1-oxidation channel) was applied only to modify the molecules and reduce contaminants in the sample before the detection on electrode 2 (E2-reduction channel). In the E1, MN was oxidized generating oxidation products which were reduced at E2. In this way, monitoring of the channel E2 current eliminates a great deal of the interferences and improves the signal/noise ratio, emphasizing the chromatographic response for MN (see Figures 1 and 2).

Method Validation

The method was validated in a 5-day validation protocol. The validation parameters included linearity and LLOQ, intra-and interday precision and accuracy, and stability were performed according to the

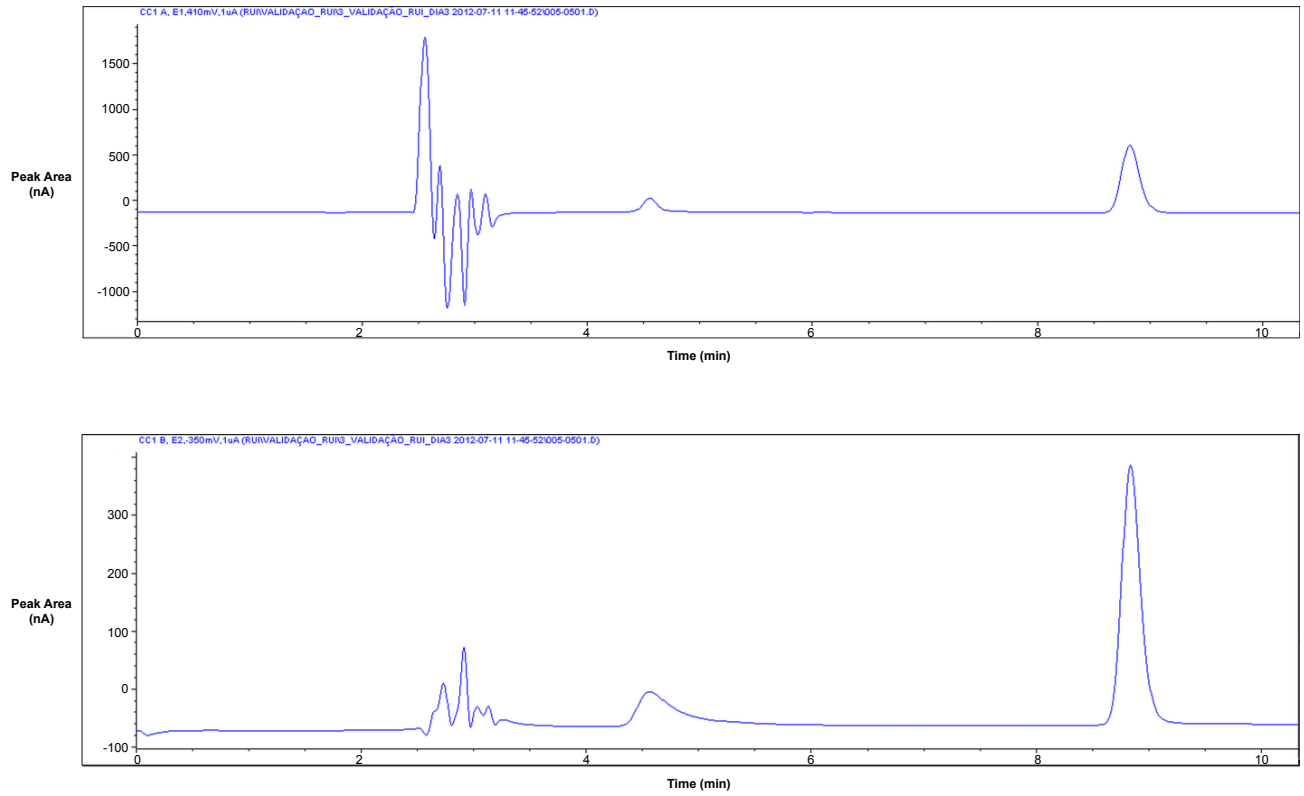


Figure 1: Chromatogram of MN at 2 nmol*mL^{-1}; (a) E1-Oxidation, (b) E2-Reduction (down); retention time (MN) of 8.873 min.

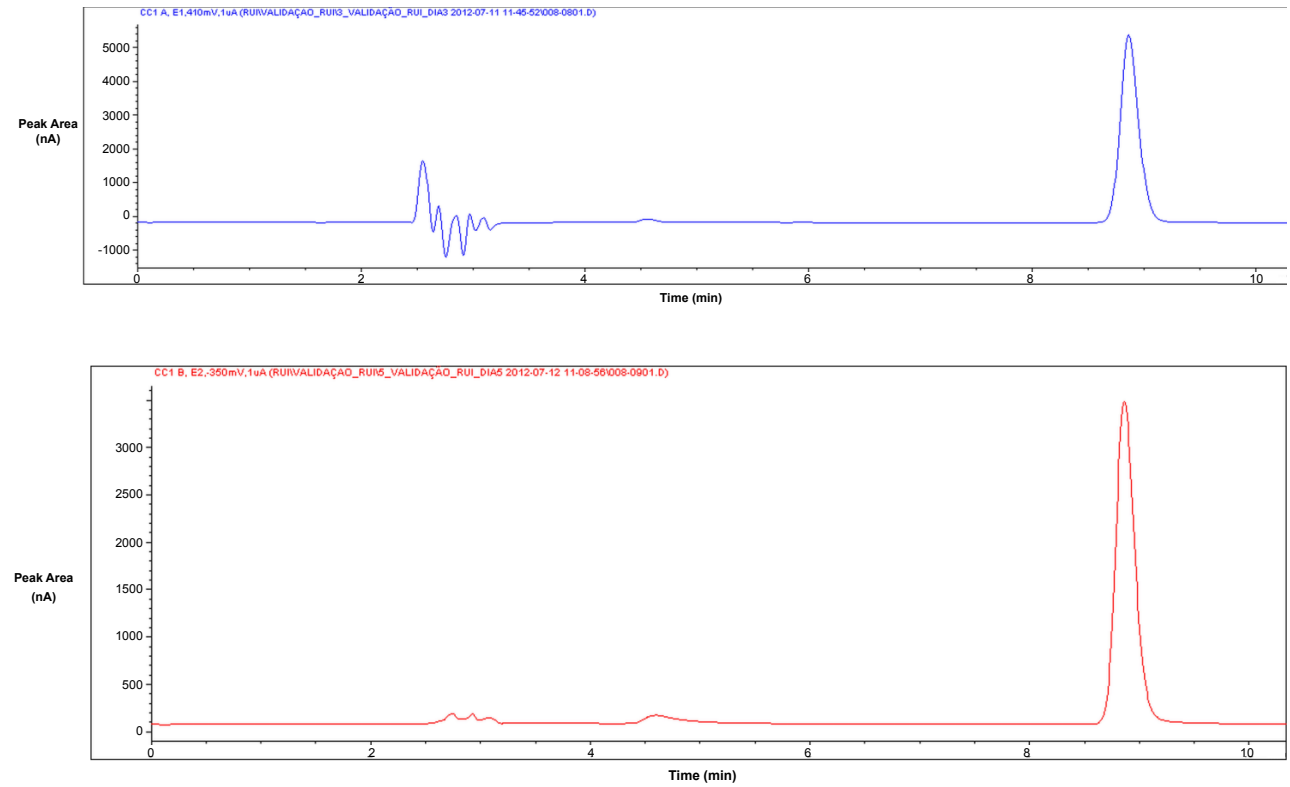


Figure 2: Chromatogram of MN at 15 nmol*mL^{-1}; (a) E1-Oxidation, (b) E2-Reduction (down); retention time (MN) of 8.873 min.

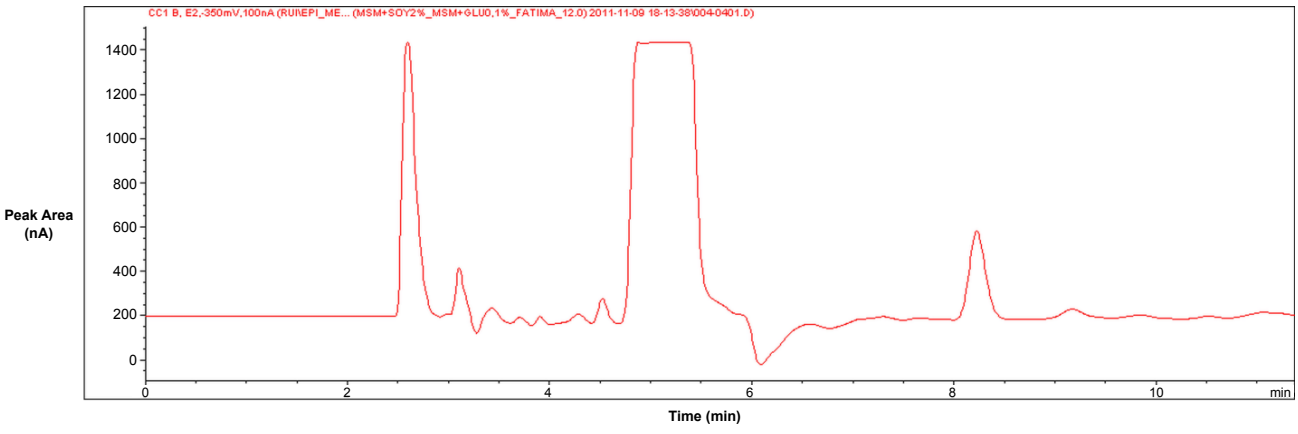


Figure 3: Chromatogram of a biological extract sample (Retention time (MN) of 8.220 min).

guidelines principles of the Food and Drug Administration [23] and International Conference on Harmonization [24].

As shown in Figure 4, the HPLC chromatogram of a fermentation extract sample demonstrates that the method provides an excellent resolution and selectivity between the compounds of interest, allowing the samples to be directly injected without pretreatment. In addition, the different components of the bacteria extract sample were injected as a "blank analysis" in order to evaluate possible interferents at the respective retention times of the analyte but none was observed.

The linearity of the method for MN quantification was established between 0.25 and 15 nmol/mL with eight evenly distributed calibrators and the following values were obtained (mean values ± SD): 2943.4 ± 24.2664 for slope (m), 126.3828 ± 76.0743 for intercept (b) and a R^2 of 0.9997 ± 0.0001. Each calibration level was in quintuplicate and determination coefficients higher than 0.99 were obtained. The calibration curves were obtained by plotting the peak area ratio between the analyte versus analyte concentration. The calculated concentration of each calibrator had to be within a ± 15% interval of target except for the LLOQ, where ± 20% was accepted. In addition, to each calibration curve, two control samples were prepared in triplicate (n=15) at the concentration 3 and 10 n nmol/mL, the low quality control (LQC)

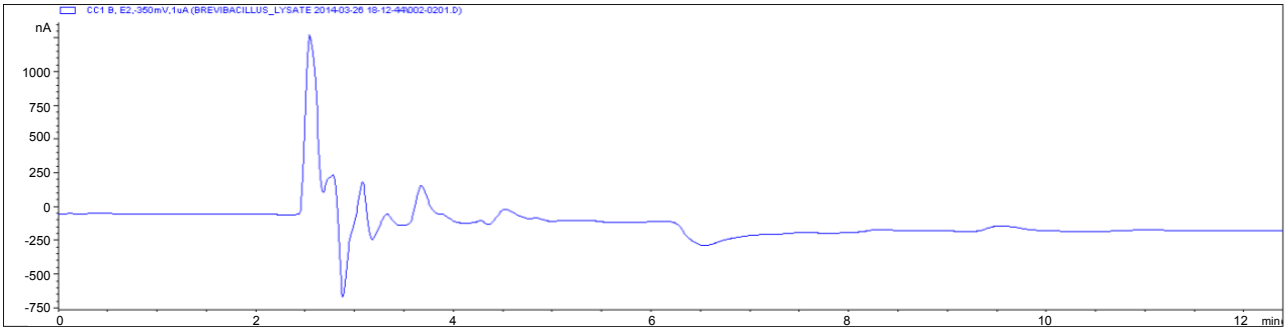


Figure 4: Representative chromatogram of a *Brevibacillus choshinensis* lysate extract sample (blank).

	Measured concentration		CV (%)		RE (%)	
Spiked	**Intraday (n=6)**	**Interday (n=5)**	**Intraday (n=6)**	**Interday (n=5)**	**Intraday (n=6)**	**Interday (n=5)**
0.25	0.28 ± 0.01	0.24 ± 0.02	4.36	8.78	11.55	-4.13
0.5	-	0.51 ± 0.02	-	4.52	-	1.42
1	-	1.01 ± 0.04	-	4.16	-	1.33
2	1.98 ± 0.02	1.98 ± 0.04	1.03	1.96	-0.85	-0.87
4	3.94 ± 0.03	3.94 ± 0.07	0.74	1.67	-1.57	-1.56
8	7.94 ± 0.03	8.10 ± 0.07	0.33	0.87	-0.79	1.22
12	11.81 ± 0.05	12.03 ± 0.14	0.47	1.18	-1.60	0.28
15	-	14.94 ± 0.13	-	0.88	-	-0.40

All concentrations in nmol*mL^{-1}; CV, coefficient of variation; RE, relative error [(measured concentration−spiked concentration)/spiked concentration×100].

Table 1: Intra-day and Inter-day precision and accuracy (n=5).

and high quality control (HQC) respectively. The LLOQ was defined as the smallest concentration of analyte (MN) that could be measured reproducibly and accurately (coefficient of variation less than 20% and calculated concentration within a ±20% interval from the target level) and was established at 0.25 nmol/mL. The LOD, was not systematically evaluated, and was defined to be the same as the LLOQ for practical reasons.

On the other hand, since the coulometric detector presents a high sensitivity, this parameter was set at 1 μA, allowing the detection of the MN standards with higher concentration without channel saturation and not compromising the detection of low standards. Moreover, following this methodology, a retention time around 8.8 min for MN was obtained, faster and more stable hands-on sample than reported by Gamache et al. [25] and Lenders [11] using coulometric detection and better precision/accuracy and lower retention time than reported by Passarinha et al. [12] and Parker and collaborators [26] with amperometric detection. In addition, our method showed a better efficiency and sensitivity, being able to detect MN concentrations of 0.25 nmol/mL with an excellent signal-to-noise ratio, better than reported by amperometric detectors [12], making the coulometric detection an advantageous method to MN quantification.

Other parameters evaluated for this method were the intraday, interday and intermediate precisions. In what concerns to the intraday precision and accuracy (relative error, %) for MN, it was determined by analysis of six independent replicates at five concentrations across the dynamic range of the assay: 0.25; 2; 4; 8 and 12 nmol/mL. The obtained CVs were typically below 5% at all concentrations, while relative errors were within a ±12% interval (see Table 1).

In addition, the interday precision and trueness were evaluated at eight concentrations within a 5-day period. The analysis of the interday precision and trueness yielded CVs generally lower than 9% at all concentration levels, while trueness was within a ±5% interval (see Table 1).

Another parameter evaluated in this work was the intermediate precision (combined intra- and interday) that was determined using the QC samples (LQC and HQC) that were prepared and analysed simultaneously with the calibration curves on 5 different days (15 measurements for each concentration). The CVs were typically below 2% at all concentrations, while accuracy was within ±6% of the nominal concentration.

Finally, the stability of metanephrine was assessed using samples spiked at the above-mentioned QC concentrations (n=3). In order to study the stability of the processed samples, the samples were left standing at room temperature and in the HPLC autosampler for 24 hrs. Those samples were compared to freshly prepared samples, and the obtained coefficients of variation were less than 10% for all compounds, meaning that MN are stable in the samples for at least 24 h at room temperature. Furthermore, the freeze/thaw stability was evaluated as follows: the MN QC samples were spiked at the intended concentrations and stored at −20°C for 1hrs and after this period, the samples were thawed single-handedly at room temperature. Following the third freeze/thaw cycle, the samples were analyzed and the obtained peak areas were compared to those obtained by analysis of freshly prepared samples. The MN was found to be stable for at least 3 freeze/thaw cycles (the obtained CVs were less than 10% for all compounds) with all values lying within ±15% and, hence, were considered acceptable (see Table 2). Also, the target analyte shows stability at room and injector temperature (4°C) up to 24 hrs, which is in accordance with previous findings [27]. Extended storage is possible at -20°C, because at least 3 freeze-thaw cycles had no influence on MN detection.

Method application with biological samples

After validation, the described procedure was applied in three

	Stability at room temperature			Stability of injector			Freeze stability		
Spiked	Measured	CV (%)	RE (%)	Measured	CV (%)	RE (%)	Measured	CV (%)	RE (%)
3	2.63 ± 0.28	9.83	-12.49	2.72 ± 0.13	2.72	-9.45	2.63 ± 0.26	9.81	-12.48
10	9.45 ± 0.93	9.89	-5.53	9.70 ± 0.58	9.70	-3.01	9.51 ± 0.84	8.81	-4.86

All concentrations in nmol*mL^{-1}; CV, coefficient of variation; RE, relative error [(measured concentration-spiked concentration/spiked concentration)×100]

Table 2: Stability data (n=3).

biological samples obtained from a MB-COMT enzymatic activity assay. A typical chromatogram of one of these biological samples is shown in Figure 3. Moreover, as MN is susceptible to oxidation due to the presence of the catechol structure, the presence of a reducing agent in standard solutions (perchloric acid 0.2 M) usually guarantees a higher stability of the analyte avoiding the oxidative degradation products of MN, which strongly influence the accuracy of the analysis [12]. Therefore, as can be seen in the chromatogram of a biological sample (see Figure 3), is possible to observe extra peaks corresponding to perchloric acid and epinephrine (used as substrate in MB-COMT enzymatic assay). However, these peaks don't interfere with MN (reaction product of MB-COMT enzymatic assay) in terms of magnitude and shape of peaks. In addition, all the compounds of the sample, namely MN, give neat and resolved chromatographic peaks as well as a very good selectivity to our interest compound. The mean concentration of MN present in the 3 biological samples was 1.499 ± 0.0135 nmol/mL. Finally, in order to evaluate possible interferents that could overlapping at the respective retention time of MN, a blank (bacteria lysate extract deproteinized with percloric acid) was injected into the HPLC system with the same operatory conditions (Figure 4). As we can observed for the retention time range of 8 to 9 min none components from the bacteria extract sample can be visualized, demonstrating that the analytical method develop is appropriated for this biological matrices.

Conclusions

A simple procedure employing a HPLC system with coulometric detector was developed and fully validated for the qualitative and quantitative determination of MNs from MB-COMT recombinant *Brevibacillu schosinensis* extracts. The proposed HPLC method with ECD, which employs an innovative coulometric detection in MB-COMT enzymatic assays, is suitable for the analysis of the main endogenous catecholamines, namely epinephrine and its main metabolite MN.

The optimization of the chromatographic conditions led to a more rapid, efficient and sensitive method, allowing the clear determination of MNs in MB-COMT enzymatic assays. The HPLC-ECD method developed in this work, in comparison with previously reported methods, namely the metanephrine detection quantification by HPLC coupled with amperometric detection [12], presents shorter chromatographic runs (10 minutes in contrast with 20 minutes). In addition, this method also presents a very good resolution to detect MN and the greater sensitivity associated with the coulometric detector, in contrast to other existing methods. Finally, another major advantage of this method is the improvement on the signal/noise ratio achieved that allowed us to lower the LLOQ to metanephrine from 1 nmol/mL [12] to 0.25 nmol/mL^{-1}, what can be of extremely importance in determining low COMT activity levels.

The method sensitivity, linearity, limits and intra- and interday precision and accuracy were adequate, allowing MN detection even at very low concentrations. Therefore, this procedure may be useful both in research and in routine analysis, more specifically in the quantification of low metanephrine concentrations associated with the COMT biological activity obtained either from recombinant or native sources. In conclusion, the accurate determination of lower metanephrine levels within a short time analysis provided by this method may have important implications in COMT research and, consequently, in the development of new and more effective COMT inhibitors for application in Parkinson's disease.

Acknowledgements

A.Q. Pedro acknowledges a doctoral fellowship (SFRH/BD/81222/2011) from Fundaçãopara a Ciência eTecnologia within the scope of QREN-POPH-Advanced Formation programs co-funded by Fundo Social Europeu and MEC. This work was partially funded by Fundação da Ciência e Tecnologia I.P. (PIDDAC) and Fundo Europeu de Desenvolvimento Regional-FEDER funds through Programa Operacional Factores de Competitividade (POFC)-COMPETE: FCOMP-01-0124-FEDER-027563 and by National Funds through FCT-Fundaçãopara a Ciência eTecnologia within the scope of Project "EXPL/BBB478/BQB/0960/2012". Also, Santos F.M. and Gonçalves A.M. acknowledges a fellowship from the project "Technologies for purification and controlled release of biopharmaceuticals to be applied in age-related disease" (CENTRO-07-ST24_FEDER-002014), funded by ProgramaOperacional Regional do Centro2007-2013 QREN (Programa "Mais Centro"). The authors also acknowledge the MscFilipa Correia, Luís Martins and Guilherme Espirito Santo for technical support on HPLC-ECD.

References

1. Chan E, Ho P (2000) High performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometric method for the analysis of catecholamines and metanephrines in human urine. Rapid Commun Mass Spectrom 14: 1959-1964.
2. Bonifacio MJ, Soares-da-Silva P (2007) Catechol-O-methyltransferase and its inhibitors in Parkinson's disease. CNS Drug Rev 13: 352-379.
3. Pedro AQ, Bonifacio MJ, Queiroz JA, Maia CJ, Passarinha LA (2011) A novel prokaryotic expression system for biosynthesis of recombinant human membrane-bound catechol-O-methyltransferase. J Biotechnol 156: 141-146.
4. Rosano T, Swift T, Hayes L (1991) Advances in catecholamine and metabolite measurements for diagnosis of pheochromocytoma. Clin Chem 37: 1854-1867.
5. Middelkoop C, Dekker G, Kraayenbrink A, Popp-Snijders C (1993) Platelet-poor plasma serotonin in normal and preeclamptic pregnancy. Clin Chem 39: 1675-1678.
6. Camanas RMV, Mallols JMS, Lapasió JRT, Ramis-Ramos G (1995) Analysis of pharmaceutical preparations containing catecholamines by micellar liquid chromatography with spectrophotometric detection. Analyst 120: 1767-1772.
7. Lizuka H, Ishige T, Ohta Y, Yajima T (1999) Simultaneous determination of 5-hydroxyndoles and catecholamines by HPLC with fluorometric precolumn derivatization. Adv Exp Med Biol 467: 821-826.
8. Engbaek F, Voldby B (1982) Radioimmunoassay of serotonin (5-hydroxytryptamine) in cerebrospinal fluid, plasma and serum. Clin Chem 28: 624-628.
9. Takezawa K, Tsunoda M, Murayama K, Santa T, Imai, K (2000) Full papers–Separation science-automatic semi-microcolumn liquid chromatographic determination of catecholamines in rat plasma utilizing peroxy oxalate chemiluminescence reaction. Analyst 125: 293-296.
10. Volin P (1992) Determination of urinary normetanephrine, metanephrine and 3-methoxytyramine by high-performance liquid chromatography with electrochemical detection: comparison between automated column-switching and manual dual-column sample purification methods. J Chromatogr B 578: 165-174.
11. Lenders J, Eisenhofer G, Armando I, Keiser H, Goldstein D, et al. (1993) Determination of metanephrines in plasma by liquid chromatography with electrochemical detection. Clin Chem 39: 97-103.
12. Passarinha LA, Bonifacio MJ, Queiroz JA (2006) The effect of temperature on

the analysis of metanephrine for catechol-O-methyltransferase activity assay by HPLC with electrochemical detection. Biomed Chromatogr 20: 937-944.

13. Kagedal B, Goldstein DS (1988) Catecholamines and their metabolites. J Chromatogr 429: 177-233.
14. Achili G, Perego C, Ponzio F (1985) Application of the dual-cell coulometric detector: a method for assaying monoamines and their metabolites. Anal Biochem 148: 1-9.
15. Matson W, Langlais P, Volicer L, Gamache P, Bird E, et al. (1984) n-Electrode three-dimensional liquid chromatography with electrochemical detection for determination of neurotransmitters. Clin Chem 30: 1477-1488.
16. Frei RW, Zech K (1988) Selective sample handling and detection in high-performance liquid chromatography part A. Journal of High Resolution Chromatography 11: 613.
17. Saracino M, Raggi M (2010) Analysis of soy isoflavone plasma levels using HPLC with coulometric detection in postmenopausal women. J Pharmaceut Biomed 53: 682-687.
18. Saracino MA, de Palma A, Boncompagni G, Raggi MA (2010) Analysis of risperidone and its metabolite in plasma and saliva by LC with coulometric detection and a novel MEPS procedure. Talanta 81: 1547-1553.
19. Marklund M, Landberg R, Aman P, Kamal-Eldin A (2011) Comparison of gas chromatography-mass spectrometry and high-performance liquid chromatography with coulometric electrode array detection for determination of alkylresorcinol metabolites in human urine. J Chromatogr B 879: 647-651.
20. Sabbioni C, Saracino MA, Mandrioli R, Pinzauti S, Furlanetto S, et al. (2004) Simultaneous liquid chromatographic analysis of catecholamines and 4-hydroxy-3-methoxyphenylethylene glycol in human plasma: Comparison of amperometric and coulometric detection. J Chromatogr A 1032: 65-71.
21. Hollenbach E, Schulz C, Lehnert H (1998) Rapid and sensitive determination of catecholamines and the metabolyte 3-methoxy-3-hydroxyphen-ethyleneglycol using HPLC following novel extraction procedures. Life Sci 63: 737-750.
22. Unceta N, Rodriguez E, de Balugera ZG, Sampedro C, Goicolea MA, et al. (2001) Determination of Catecholamines and their metabolites in human plasma using liquid chromatography with coulometric multi-electrode cell-desing detection. Anal Chim Acta 444: 211-221.
23. http://www.fda.gov/downloads/drugs/guidancecompilance/regulatoryinformation/guidances/ucm070107.pdf.
24. http://www.ich.org/lob/media/media417.pdf
25. Gamache P, Kingery M, Acworth I (1993) Urinary metanephrine and normetanephrine determined without extraction by using liquid chromatography and coulometric array detection. Clin Chem 39: 1825-1830.
26. Parker NC, Levtzow C, Wright P, Woodard LL, Chapman JF (1986) Uniform chromatographic conditions for quantifying urinary catecholamines, metanephrines, vanillylmandelic acid, 5-hydroxyindoleacetic acid, by liquid chromatography with electrochemical detection. Clin Chem 32: 1473-1476.
27. de Jong WHA, Graham KS, Van der Molen JC, Links TP, Morris MR, et al. (2007) Plasma free metanephrine measurement using automated online solid-phase extraction HPLC tandem mass spectrometry. Clin Chem 53: 1684-1693.

Citation: Pedro AQ, Soares RF, Oppolzer D, Santos FM, Rocha LA, et al. (2014) An Improved HPLC Method for Quantification of Metanephrine with Coulometric Detection. J Chromatograph Separat Techniq 5: 217. doi:10.4172/2157-7064.1000217

Paper III

Biosynthesis and Purification of histidine-tagged-human soluble catechol-*O*-methyltransferase

A. Q. Pedro, F. F. Correia, F. M. Santos, G. Espírito-Santo, A. M. Gonçalves, M. J. Bonifácio, J. A. Queiroz, L. A. Passarinha

(Submitted for publication)

Short description: In the present paper is introduced a novel system for biosynthesis and purification of recombinant hexa-histidine tagged Val 108 human soluble catechol-*O*-methyltransferase (SCOMT-His_6). SCOMT-His_6 biosynthesis was carried out with *P. pastoris* X33 methanol-induced cultures and its expression levels were monitored during the fermentation period. Moreover, the target protein was purified using one-step Immobilized-Metal Affinity Chromatography (IMAC). Finally, SCOMT-His_6 sensitivity to be inhibited by nitrocatecholic inhibitors was evaluated and a MALDI-TOF/TOF (Matrix Assisted Laser Desorption/Ionization with tandem Time Of Flight) analysis was performed to confirm the correct processing of the target enzyme, i.e., if it contains the correct primary sequence.

Biosynthesis and purification of histidine-tagged human soluble catechol-*O*-methyltransferase

Augusto Q. Pedro[1], Filipa F. Correia[1], Fátima M. Santos1, Guilherme Espírito-Santo1, Ana M. Gonçalves1, Maria J. Bonifácio[2], João A. Queiroz[1], Luís A. Passarinha[1]*

[1]*CICS-UBI – Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, 6201-001 Covilhã, Portugal*

[2]*Bial – Departamento de Investigação e Desenvolvimento, 4745-457 São Mamede do Coronado, Portugal*

*Corresponding author – E-mail: lpassarinha@fcsaude.ubi.pt - Phone: +351 275 329 069; Health Sciences Research Centre (CICS-UBI); Faculty of Health Sciences, University of Beira Interior, Av. Infante D. Henrique, 6200-506, Covilhã, Portugal

Abstract

Background: Catechol-*O*-methyltransferase (COMT, EC 2.1.1.6) has been implicated in several human diseases including Parkinson's disease and the most appropriated therapy depends on the efficacy of the COMT inhibitors applied. Therefore, more effective drugs for COMT inhibition are important, and the development of such inhibitors requires research with recombinant COMT. **Results:** The time-course production of biologically active hexa-histidine-tagged COMT in methanol-induced *Pichia pastoris* cultures was evaluated and immobilized-metal affinity chromatography was applied as the main capture step for the purification of the target enzyme that proved to be extremely efficient and selective.. The best strategy allowed recovering soluble COMT at 300 mM in a highly purified fraction with a purification fold of 81 and a bioactivity recovery of 57.35 %. Using this strategy, a concentration of 3.68 mg/L of shake-flask culture of highly purified recombinant COMT was obtained. Finally, the MALDI-TOF/TOF analysis of the purified recombinant SCOMT demonstrated that it is correctly processed since no modifications in the primary sequence were observed. **Conclusion:** A new strategy was developed and implemented for the biosynthesis and purification of biologically active soluble COMT in a highly purified fraction with the ability to be inhibited by commercial Parkinson's disease inhibitors: 3,5-dinitrocatechol and entacapone.

Keywords: Catechol-*O*-methyltransferase; Immobilized-metal affinity chromatography; Chromatography; Purification; *Pichia pastoris*.

Nomenclature: AOX – alcohol Oxidase; COMT – catechol-*O*-methyltransferase; *E. coli* – *Escherichia coli*; IMAC – immobilized-metal affinity chromatography; Mut – methanol utilization phenotype; *P. pastoris* – *Pichia pastoris*; SAM – S-adenosyl-$_L$-methyonine; SCOMT – soluble catechol-*O*-methyltransferase; SCOMT-His_6 – hexahistidine-tagged soluble catechol-*O*-methyltransferase; 3,5-DNC – 3,5-dinitrocatechol.

1. Introduction

Catechol-*O*-methyltransferase (COMT, EC 2.1.1.6) is a magnesium-dependent enzyme that catalyzes the methylation of catechol substrates using S-adenosyl-$_L$-methionine (SAM) as a methyl donor [1]. In addition, the *O*-methylated catechol and S-adenosyl-$_L$-homocysteine are the products obtained from the reaction catalyzed by COMT [1]. In humans, COMT appears as two molecular forms, a membrane-bound isoform and a soluble isoform (SCOMT) comprising 221 amino acids with a corresponding molecular weight of 24.7 KDa that is found mainly in the cytoplasm [1]. In most human tissues, the levels of SCOMT greatly exceed the levels of membrane-bound COMT and, despite SCOMT present lower affinity for the majority of substrates (higher K_M), it shows a great catalytic number (V_{Max}) [2]. Therefore, SCOMT is responsible for the elimination of biologically active or toxic catechols [3]. During the last decades, COMT has been implicated in several human diseases such as cardiovascular diseases [4], estrogen-induced cancers [5] and neurologic disorders [1]. Specifically, the best documented is the important role that COMT plays in Parkinson's disease whose most effective treatment remains the dopamine replacement therapy with levodopa together with an inhibitor of aromatic amino acid descarboxylase and a COMT inhibitor [1]. Therefore, it becomes clear the importance of developing new and more effective drugs for COMT inhibition in which biophysical studies between recombinant forms of this enzyme and already available molecules can play an important role for this therapy. However, major requirements associated with these studies are the purity and quantity of the target enzyme, thus emphasizing the importance of developing new systems capable of delivery protein in higher quantities in a highly purified form.

The yeasts are an attractive eukaryotic host for the production of heterologous proteins due to their fast growth rates coupled with high cell densities and the well-developed variety of genetic tools [6]. Initially described by Philips Petroleum Company for the production of single cell protein, *Pichia pastoris (P. pastoris)* has demonstrated its most powerful success as a large scale (fermentation) recombinant production tool [7]. Moreover, *P. pastoris* as a yeast, is a single-cell microorganism that is easy to manipulate and cultivate and as a eukaryote, is capable of many posttranslational modifications performed by higher eukaryotic cells such as proteolytic processing, folding, disulfide bond formation and glycosylation [7]. During the last decade, several chromatographic approaches have been employed for the purification of SCOMT from complex lysates. In fact, our research group previously reported different strategies for the isolation and purification of SCOMT either by employing hydrophobic resins [8 – 10], an anion-exchange resin [11] or immobilized amino acids [12]. In fact, the best results were obtained for the two-step SCOMT purification using a butyl-sepharose resin followed by a gel filtration step [10]. With this strategy, SCOMT is recovered in a highly purified state (purification fold of 5.9 and a bioactivity recovery of 1%) but the application of ammonium sulfate has proven to have a deleterious effect on SCOMT biological

activity [9], thus compromising SCOMT stability and further structural studies. Immobilized-Metal Affinity Chromatography (IMAC) has emerged over the last decade as one of the most widely used methods at the laboratory scale to isolate proteins from fermentation broths, mammalian cell culture supernatants or other biological sources [13]. Over the years, several proteins with different topologies had been successfully purified using IMAC as the main chromatography step such as the human growth hormone [13], immunoglobulins [14], the coagulation factor VIII [15] or membrane-bound proteins such as the dehydrogenase/reductase SDR family member 7 [16]. The IMAC is a separation technique in which metal ions entrapped on solid chromatographic supports serve as affinity ligands for several proteins [17] , usually histidine-tagged fusion proteins [18]. In IMAC, adsorption of proteins is based on the interaction between an immobilized metal ion [the most frequently used are Cu (II), Ni (II), Zn (II), Co (II) and Fe (III)] and electron donor groups located on the surface of the proteins, mainly histidine residues [18, 19], often introduced into a target protein as an N- or C- terminal peptide "tag" [13]. In this work, for the first time, recombinant SCOMT biosynthesis was carried out in *P. pastoris* with a hexa-histidine tag in its C-terminal. Moreover, after a suitable cell lysis stage employing glass beads, the lysate was recovered and directly injected onto an IMAC resin.

2. Materials and Methods

2.1. Materials

Ultrapure reagent-grade water was obtained with a Mili-Q system (Milipore/Waters). The easy select expression kit for expression of recombinant proteins using pPICZα in *P. pastoris* and zeocin were obtained from Invitrogen (Carlsbad, CA). Yeast nitrogen base (YNB), glucose, agar, yeast extract, peptone, dithiotreitol, SAM, epinephrine (bitartrate salt), deoxyribonuclease (DNase), protease inhibitor cocktail, cysteine, sucrose, glycerol, DL-metanephrine hydrochloride, citric acid monohydrate, glass beads (500 µm), 3,5-dinitrocatechol and entacapone were purchased from SigmaChemical Co. (St. Louis, MO).

The NZYcolour Protein Marker II used for estimation of subunit molecular weight was purchased from NZYTech (Lisboa, Portugal). Anti-rabbit IgG alkaline phosphate secondary antibody and the silver nitrate solution kit were purchased from GE Healthcare Biosciences (Uppsala, Sweden) while monoclonal rabbit COMT antibody purchased by Abcam (Cambridge, England). All chemicals used were of analytical grade, commercially available, and used without further purification.

2.2. Strains, plasmids and media

Escherichia coli (*E. coli*) TOP10F' was used for DNA manipulations. *E. coli* transformants were selected on low-salt luria-bertani plates with 25 µg/mL Zeocin. *P. pastoris* X-33 was used for fusion gene expression. The following media supplemented with Zeocin 200 µg/mL were employed in *Pichia* cells fermentations: YPD medium (1% yeast extract, 2% peptone and 2% glucose), YPDS medium (YPD medium supplemented with 1M Sorbitol), BMGH (100 mM potassium phosphate buffer [pH 6.0], 1.34% yeast nitrogen base, 4 $*10^{-4}$ g/L biotin and 1% glycerol) and BMMH (100 mM potassium phosphate buffer [pH 6.0], 1.34% yeast nitrogen base, 4 $*10^{-4}$ g/L biotin and 0,5% methanol). *P. pastoris* transformants were selected on YPDS plates with 200 µg/mL Zeocin.

2.3. Construction of expression vector pICZαA-SCOMT-His_6

Easy select expression kit for expression of recombinant proteins using pPICZα in *P. pastoris* (Invitrogen, Carlsbad, CA) was used for the expression of human SCOMT with the α-mating factor secretion signal and a hexahistidine in its carboxyl-terminal. This process was carried out according to manufacturer's instructions. Briefly, the DNA fragment coding for SCOMT was obtained from the pET101/D-hSCOMT plasmid [10] by PCR using specific primers for cloning (forward primer, 5' AAC TCG AGA AAA GAA TGG GTG ACA CCA AGG AGC AG 3' and reverse primer, 5' AAC TCG AGT CAG TGA TGG TGA TGG TGA TGG GGC CCT GCT TCG CTG CCT G 3') in which the reverse primer was designed in order to introduce a hexahistidine tag in SCOMT carboxyl-terminal. The PCR conditions were as follows: denaturation at 95ºC for 5min, followed by 30 cycles at 95ºC for 30 s, 60ºC for 30 s and 72ºC for 1 min, and a final elongation step at 72ºC for 5min. The amplified DNA was purified by low melting agarose gel

electrophoresis, digested with Xho I and cloned into the vector pPICZα (previously digested with Xho I) by T4 DNA ligase. This construct was transformed into *E. coli* TOP10F' cells, grown overnight at 37ºC in plates with low salt luria bertani-agar medium containing zeocin (25 µg/mL) and colonies were screened for the presence of the construct pICZαA-hSCOMT_His_6. Therefore, some colonies were inoculated in 2.0 mL of low salt luria bertani medium and grown at 37ºC and 250 rpm overnight. From these cultures, highly purified plasmids were prepared using NzyMiniprep (Nzytech, Lisboa, Portugal) and were then subjected to DNA sequence analysis to confirm the identity of the amplicon, orientation and frame. Since the sequence was confirmed to correspond to human SCOMT gene plus the six histidines in its carboxyl-terminal [20] and was in frame with the α-mating factor secretion signal, the cloned plasmid was digested with Sac I and introduced into freshly made *P. pastoris* X33 competent cells by electroporation. After confirming that the X33 integrants presented a methanol utilization phenotype plus (Mut^+) phenotype, the stable occurrence of the expression cassette was verified in the colonies genomic DNA by PCR using AOX1 promoter and terminator specific primers (AOX1 5' GACTGGTTCCAATTGACAAGC 3' and AOX1 5' CAAATGGCATTCTGACATCC).

2.4. Recombinant SCOMT production and recuperation

Recombinant SCOMT production was performed using *P. pastoris* X33 cells containing the expression construct pICZα A-hSCOMT_His_6 according to the following protocol: cells containing the expression construct were grown at 30ºC in YPD plates and then a single colony was inoculated in 100.0 mL of BMGH medium in 500.0 mL shake flasks. Subsequently, cells were grown at 30ºC and 250 rpm overnight when the cell density at 600 nm (OD_{600}) typically reached 6.0. Afterwards, since the inoculation volume was fixed to achieve an initial OD_{600} of 1.0, an aliquot of the fermentation in the medium BMGH was collected and centrifuged at room temperature during 5 minutes. After centrifuging the cells and ensure that all glycerol was removed, the cells were ressuspended in the induction medium and added to 500.0 mL shake-flasks to a total volume of 100.0 mL. The fermentations were carried out during 120 hours at 30ºC and 250 rpm and were supplemented with methanol at a final concentration of 1% every 24 hours. Then, the cells were harvested by centrifugation (1500 x g, 10 min, 4ºC) and stored frozen at –20ºC until use. *P. pastoris* lysis was accomplished though the application of a sequential procedure involving glass beads consisting of 7 cycles of vortexing during 1 minute with 1 minute of interval on ice. Therefore, cell suspensions were lysed in equilibrium buffer (150 mM NaCl, 10 mM DTT, 50 mM Tris, 1 mM $MgCl_2$, pH 8.0) containing protease inhibitor cocktail at a ratio of 1:2:2 (1 g cells, 2 mL lysis buffer and 2 g glass beads) and, after the lysis process were completed, the mixture was centrifuged (500g, 5 minutes, 4ºC). Finally, the supernatant was removed and the obtained pellet was ressuspended in the chromatographic binding buffer (500 mM NaCl, 50 mM Tris and 1 mM $MgCl_2$ at pH 7.8) and DNase I was added (250 µg/ml).

2.5. Immobilized Metal Affinity Chromatography

Chromatographic experiments were performed in an ÄKTA Avant system with UNICORN 6.1 software (GE Healthcare, Uppsala, Sweden). The chromatographic experiments were performed on a HisTrap™ FF crude (5mL), a prepacked IMAC stationary phase with nickel ions (GE Healthcare, Uppsala, Sweden). The chromatographic runs were performed at 20ºC. All buffers pumped into the system were prepared with Mili-Q system water, filtered through a 0.20 µm pore size membrane (Schleicher Schuell, Dassell, Germany) and degassed ultrasonically. Preliminary experiments were conducted in order to assess the most suitable imidazole concentration required for SCOMT elution. Unless otherwise stated, the column was initially equilibrated with 500 mM NaCl, 50 mM Tris and 1 mM $MgCl_2$ at pH 7.8. Aliquots of ressuspended pellet in equilibration buffer (diluted 1:2 with equilibrium buffer), were applied onto the column using a 500 µL at a flow-rate of 0.5 mL/min. After elution of the unretained species, imidazole concentration was increased to 50 mM at 1mL/min in a stepwise mode for 5 column volume (CVs). Subsequently, it was applied a stepwise mode at 70 mM of imidazole concentration for 5 CVs, followed by a step at 300 mM of imidazole for 5 CVs. Finally, the imidazole concentration in the elution buffer was increased to 500 mM for 5 CVs.

In all chromatographic runs, the conductivity, pH, pressure, temperature as well as absorbance at 280 nm were continuously monitored. Fractions were pooled according to the chromatograms obtained, collected in tubes containing a stabilizing solution (Final concentrations: 10 mM DTT, 140 mM of sucrose, 150 mM of cysteine and 10% of glycerol), concentrated and desalted with Vivaspin concentrators (10.000 MWCO) and conserved at 4ºC until further analysis. The protein content in samples were measured by Pierce BCA Protein Assay Kit (Thermo Scientific, USA) using Bovine Serum Albumin (BSA) as standard (25-2000 µg/mL) and calibration control samples according to the manufacturer's instructions.

Finally, the experiments of activity were designed to evaluate the methylating efficiency of recombinant SCOMT [21], by measuring the amount of metanephrine formed from epinephrine in the different fractions obtained in the different IMAC assays. In SCOMT enzymatic assay, an aliquot of 150 µg/mL of the SCOMT ressuspended pellets and purified extracts was incubated in 5 mM sodium phosphate buffer (pH 7.8) containing 0.2 µM $MgCl_2$, 2 mM EGTA, 250 µM SAM and 1 mM epinephrine in a total sample volume of 1 mL. Reactions were carried out at 37ºC during 15 minutes and were stopped by incubation in ice following the addition of 2M 200 µL perchloric acid. The samples were centrifuged at 6000 rpm during 10 minutes, the supernatants were filtered and were subsequently injected into the HPLC system.

2.6 SCOMT enzymatic assay

The methylating efficiency of recombinant SCOMT was evaluated by measuring the amount of metanephrine formed from epinephrine as previously described [9, 21]. After processing the samples, the metanephrine levels in the different samples obtained in the IMAC assays were

determined using HPLC with coulometric detection, as previously described [21]. Briefly, the chromatographic analysis was performed using a HPLC Agilent 1260 system (Agilent, Santa Clara, USA) equipped with an autosampler and quaternary pump coupled to an ESA Coulchem III detector (Milford, MA, USA). Chromatographic separation was achieved on an analytical column Zorbax 300SB C18 RP analytical column (250x4.6 mm i.d. 5 µm) (Agilent, Santa Clara, California, USA). The chromatographic method was developed using as mobile phase [0.1M NaH2PO4, 0.024M citric acid monohydrate, 0.5mM sodium octyl sulfate and 9% (v/v) acetonitrile] and the column effluent was monitored with an electrochemical detector in the coulometric mode, which was equipped with a 5011 high sensitivity dual electrode analytical cell (electrodes I and II) using a procedure of oxidation/reduction (analytical cell #1: +410mV; analytical cell #2: -350 mV). The method sensitivity was set at 1 µA, the flow-rate applied was 1 mL/min and the column temperature was optimized to 30ºC. The chromatograms were obtained by monitoring the reduction signal of the working electrode were metanephrine retention time was around 8.8 minutes. Finally, the metanephrine content in samples was measured using metanephrine standards (1 – 15 nmol/mL) as a calibration control.

2.7. MALDI-TOF/TOF analysis of recombinant purified SCOMT-His_6

To further characterize the purified recombinant SCOMT-His_6, matrix assisted laser desorption/ionization time-of-flight (MALDI-TOF/TOF) mass spectrometry analysis were performed. For N-terminal analysis and protein identification, purified recombinant SCOMT-His_6 was digested overnight using a trypsin/protein ratio of 1:30, followed by desalting with C18 Zip Tips. For MALDI analysis, the obtained tryptic peptides were mixed with α-cyano-4-hydroxycinnamic acid matrix (5 mg/mL in 50:50 acetonitrile: 0.3% TFA) and spotted on a MALDI sample plate. Samples were analyzed on 4800 plus MALDI-TOF/TOF analyzer (Applied Biosystems), equipped with a 355 nm laser. All spots were initially analyzed in positive MS mode in the range 800 to 4000 m/z by averaging 1500 laser spots. The eight more intense MS ions per spot that satisfied the precursor criteria (200 ppm fraction-to-fraction precursor exclusion, S/N ratio >25) were selected for subsequent MS/MS analysis. All MS/MS data were acquired using 1 keV collision energy with a total of 1500 laser shots per spectrum. Peak lists were export to an MGF file using the function Peaks to Mascot 4000 *Series Explorer*™ Software (Applied Biosystems). Protein identification were performed using Paragon algorithm from ProteinPilot™ Software, in which peak list were searched against the *Homo sapiens* UniprotKB canonical & isoform reviewed database (47,869 entries) downloaded from Swiss-Prot at 1 July 2015 with an unused score of at least 1.3 (95% confidence interval).

2.8. COMT inhibition by 3,5-dinitrocatechol and entacapone

The experiments with the aim of studying the effect of COMT inhibitors were conducted according to what was previously described by Vieira-Coelho and collaborators [22]. These trials consisted in measuring the percentage of inhibition of SCOMT activity when compared to the control (without inhibitor). These experiments were similar from those designed to

simply evaluate the SCOMT biological activity of the fractions obtained from IMAC with one minor modification: the reaction mixture containing the SCOMT purified fraction, SAM, $MgCl_2$, phosphate buffer and EGTA was preincubated for 20 minutes with increasing concentrations of the selected compounds (3,5-dinitrocatechol and entacapone) before the addition of the substrate epinephrine. All data analysis was performed using Prism 6 (GraphPad Software Inc. San Diego, CA).

2.9. Electrophoresis, Western-blot and Dot-blot analysis

Reducing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out according to the method of Laemmli [23] and as previously described [9] with a minor modification, instead of using coomassie blue staining, the gels were stained using the silver nitrate solution kit according to the manufacturer's instructions (GE Healthcare, Uppsala, Sweden). Immunologically active SCOMT was detected either by western-blot or dot-blot analysis as previously described [11].

3. Results and discussion

During the last years, COMT has been widely studied due its harmful role in several human diseases such as Parkinson's disease whose most effective treatment is the dopamine replacement therapy with levodopa together with an inhibitor of aromatic amino acid decarboxylase and a COMT inhibitor [1]. As a matter of fact, an improvement in Parkinson's therapy may come from the design of new COMT inhibitor molecules with increased selectivity and efficacy which can be achieved by the study of the interaction between COMT purified extracts and it´s inhibitors. Therefore, in order to obtain recombinant biologically active protein in high quantities, necessary to perform several structural and functional studies, it is imperative to develop a biotechnological platform capable of producing high levels of COMT with a high purity degree. Therefore, in this work, the major bottlenecks associated with previous published works concerning SCOMT production and purification are overcome and a new integrated strategy is reported for hexahistidine tagged-COMT (SCOMT-His_6) biosynthesis in *P. pastoris* and subsequently purification using immobilized metal affinity chromatography.

3.1. SCOMT-His_6 biosynthesis and recuperation

P. pastoris has the ability to produce proteins with therapeutic and commercial interest in concentrations ranging from miligrams to grams per liter [24]. In particular, as methylotrophic yeast, *P. pastoris* possess the strongest, most regulated promoter alcohol oxidase (AOX) where high levels of transcription are induced by methanol [24]. Previously, our research group reported several strategies for the intracellular production of SCOMT using *E. coli* [10, 25]. Therefore, in this work, an expression vector that contains the α-mating factor secretion signal was used in order to target the heterologous protein to the *P. pastoris* secretory pathway. Specifically, this expression vector contains a vector copy of the α-mating factor secretion signal prior to the cloning site [26] that is still part of the open reading frame and, therefore, this signal sequence is translated along with the target gene. Then, after translation, the pro-region cleavage site is recognized by the yeast kex2 protease and the signal sequence removed, resulting in the release of the mature, fully processed protein [26]. The cleavage process removes the signal peptides of pre-proteins in the late Golgi, which facilitates the subsequent entry of mature proteins toward the secretion vesicles and thus represents a key step in the yeast secretion pathways [27]. Therefore, the forward primer (5' AAC TCG AGA AAA GAA TGG GTG ACA CCA AGG AGC AG 3') designed to amplify the SCOMT gene contained a Xho I restriction site (CTC GAG) and the Kex2 cleavage site (AAA AGA) immediately before the first nucleotides of the target gene [20]. Therefore, we ensure that after translation and cleavage of the signal sequence onto the kex2 site, the matured target protein is released with the native N-terminus.

Recombinant SCOMT-His_6 biosynthesis was carried out under the control of AOX promoter in *P. pastoris* X33 cells. After cloning the SCOMT gene plus the sequence encoding the six

histidines in the SCOMT carboxyl-terminal in pICZα A, the target expression vector was introduced into freshly made X33 competent cells by electroporation and the Mut phenotype of the selected colony was confirmed to be Mut^+. In fact, the Mut^+ *P. pastoris* X33 strain was selected in this work since previous studies carried out by our research group showed that this strain is more suitable for expressing membrane-bound COMT than Mut^S KM71H [28].

Concerning the upstream stage, SCOMT-His_6 biosynthesis was initiated at 30ºC and 250 rpm in a semi-defined medium (BMGH) with glycerol as the main carbon source. Then, after achieving an OD_{600} of 6, the fermentations were initiated by shifting the carbon source from glycerol to methanol (0.5% (v/v)) that acts as inducer. In addition, to maintain the induction over the entire fermentation period without causing cell toxicity, whenever the fermentation took more than 24 hours, 1% (v/v) methanol was added to the cultures. Initially, since the vector used in this work comprises the α-factor mating signal sequence, required to target the protein to the secretory pathway [24], SCOMT-His_6 biosynthesis was monitored in the extracellular medium from *P. pastoris* fermentations. Indeed, one of the major advantages in *P. pastoris* bioprocesses is that the recovery and purification steps are simplified due to the low levels of endogenous proteins that are secreted into the extracellular medium [24]. Therefore, samples were taken at different times of fermentation, concentrated and desalted in vivaspin concentrators (10000 MCW) and were subsequently analyzed by dot-blot. As depicted in Figure 1, the dot-blot analysis of SCOMT production during the fermentation time indicates that at 120 hours are achieved the highest SCOMT levels. However, in the extracellular medium, biologically active SCOMT was only present in at the initial stages of the fermentation period and with low levels. In fact, although the initial pH of the culture medium is 6.0, during the course of *P. pastoris* fermentation, the accumulation of acidic metabolites leads to a decrease in pH to values ranging from 3.0 to 4.0. Therefore, to ascertain if there is a secreted SCOMT inhibitor or if the low pH of the culture medium together with the operational conditions such as temperature and stirring lead to losses in SCOMT biological activity, an experiment was carried out where the medium was spiked with biologically active SCOMT. Then, SCOMT activity was measured at the beginning of the fermentation (0 hours) and after 24 hours of the fermentation and it was found that only 60% of the initial biological activity is recovered, indicating that SCOMT from the extracellular medium tend to lose its biological activity. Therefore, we believe that the low pH, the stirring, the temperature and the possible presence of a secreted SCOMT inhibitor can lead to SCOMT denaturation, once this protein is regarded as being extremely unstable in several conditions [9, 12].

Subsequently, in an attempt to establish the preferential location of the biologically active target protein after *P. pastoris* lysis, samples from a 120 hour fermentation were taken every 24 hour and analyzed either by dot-blot or by the determination of SCOMT specific activity. As depicted in Figure 1, for the cell lysis supernatant and the ressuspended pellet, despite the levels of immunologically active SCOMT increase from 24 to 120 hours of fermentation,

the profile of SCOMT biological activity is quite different. It was verified that SCOMT specific activity reaches a peak at 24 and 48 hours of fermentation for the cell lysis supernatant and the ressuspended pellet, respectively, after which its levels decrease over time until no biologically active SCOMT was detected at 120 hours of fermentation. Nevertheless, the highest levels were detected for the ressuspended pellet.

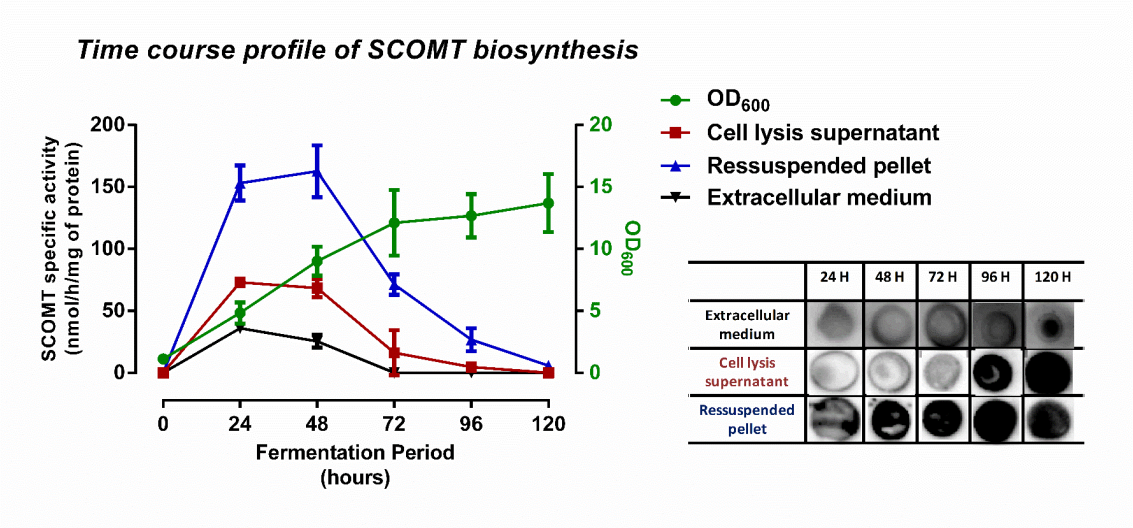


Figure 1 --Growth profile of *Pichia pastoris* harboring the recombinant plasmid pICZαA_SCOMT_His_6 at 250 rpm and 30ºC in BMMH medium. SCOMT-His6 biosynthesis was monitored through the determination of the target protein biological activity in the three major fractions obtained during this bioprocess (extracellular medium, cell lysis supernatant and ressuspended pellet). An inset of the dot-blot analysis over different periods of fermentation for each fraction is also shown. Each point represents the mean of three independent experiments.

As a matter of fact, polypeptide aggregation during overexpression could result from the accumulation of high concentrations of folding intermediates in spontaneous folding pathways or from inefficient recognition or processing of polypeptide substrates by molecular chaperones [29]. In addition, it has been described that the production of recombinant proteins at higher temperatures may lead to the exposure of more hydrophobic surfaces during peptide folding and favor hydrophobic interactions, thus predispose proteins to aggregation [30]. However, after *P. pastoris* lysis, the target protein may also be co-aggregating with membrane components, as it was previously described for some tagged-proteins [31]. Nevertheless, despite the unusual location after cell lysis, SCOMT-His_6 is recovered in a highly biologically active form and, based on the obtained results, the ressuspended pellet was chosen for the chromatographic trials and the fermentation period was fixed in 48 hours.

3.2. SCOMT-His_6 purification

IMAC has emerged over the past decade as one of the most widely used methods at the laboratory scale for the isolation of target proteins from fermentation broths, culture supernatants and other biological sources, demonstrating high selectivity [13]. In general, the biomolecules are retained in IMAC using equilibrium buffer without imidazole or low concentrations and the elution is usually achieved by increasing the imidazole concentration [19, 32].

Preliminary studies were carried out in order to study the chromatographic behavior of SCOMT-His_6 interacting with the IMAC Ni-sepharose matrix. Therefore, several buffers were studied and it was found that the formulation containing 500 mM NaCl, 50 mM Tris and 1 mM $MgCl_2$ at pH 7.8 was the most suitable since SCOMT is quite stable in this buffer and the inclusion of 500 mM NaCl allowed the removal of some host contaminants that bind the matrix through ionic interactions. In addition, an increasing linear gradient of elution buffer containing 500 mM imidazole was tested in order to assess the imidazole concentration necessary to elute our target protein. Indeed, it is possible to observe that the target protein interacts with the matrix, being almost all eluted in Peak III with an imidazole concentration close to 150 mM, as shown by the dot-blot analysis in the different peaks depicted in Figure 2A. Then, after these initial trials we concluded that, it wouldn't give the desired selectivity needed to separate SCOMT from major host contaminants. In addition, based on the fact that immunologically active SCOMT was also detected in the fraction that eluted at 0 mM imidazole (see Figure 2A), we also decided to lower the flow-rate during the injection of the sample from 1 mL/min to 0.5 mL/min since it can improve the binding of tagged proteins to the IMAC matrix [33]. Therefore, after decreasing the flow-rate from 1.0 mL/min to 0.5 mL/min during the injection of the sample, we found that the target protein is totally retained in the column (data not shown).

In IMAC, through the competition with nickel ions, the imidazole is responsible for eluting the proteins and when present at low concentrations in the binding buffer, it may prevent the binding of host proteins with exposed histidines, allowing the removal of contaminants in the flowthrough during the injection of the sample [34]. In fact, maintaining a lower flow-rate during the injection, a new strategy comprising the injection of the sample with binding buffer containing 10 mM imidazole, followed by four stepwise gradients of 70 mM, 100 mM, 150 mM and 500 mM imidazole was tested. However, the presence of 10 mM imidazole during the injection of the sample leads to the elution of a high quantity of the target protein in the flow through. Moreover, with this strategy the remaining target protein that was bound to the matrix, eluted with 70 mM imidazole in a more contaminated fraction, despite when the linear gradient was applied (see Figure 2A) SCOMT eluted with an imidazole concentration near 150 mM imidazole. Taking into account these results, we needed to improve the binding of the target protein to the matrix and, simultaneously, increase the purity of the target fraction where preferentially SCOMT accumulates. Therefore, the imidazole concentration

was lowered from 10 to 5 mM in the binding buffer, followed by two stepwise gradients of 50 and 70 mM imidazole, designed to remove contaminants that presented some affinity to the matrix and two final stepwise gradients, 300 mM and 500 mM imidazole. These results are shown in Figure 2B and it is possible to observe that the majority of the contaminants are eluted in the flowthrough with 5 mM imidazole while the dot-blot analysis showed that none SCOMT was eluted in this fraction. Then, with 50 mM (peaks II and III) and 70 mM imidazole (peak IV), the host protein contaminants that bind to the matrix are eluted without significant losses of the target protein, as demonstrated in Figure 2B. Subsequently, a stepwise gradient of 300 mM imidazole was applied where the target protein was eluted in a highly purified fraction, followed by a stepwise gradient with 500 mM imidazole where almost no proteins were eluted.

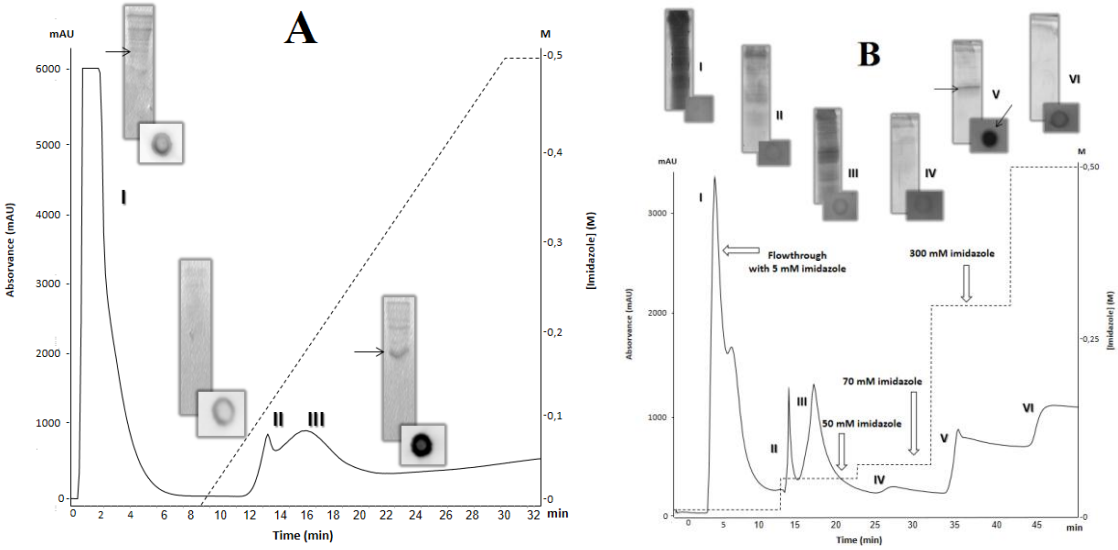


Figure 2 - Chromatographic profiles of the different strategies evaluated for the recombinant human SCOMT_His_6 purification from *P. pastoris* lysates on a HisTrap FF crude resin (5 mL column volume) with nickel ions immobilized and respective SDS-PAGE/Dot blotting analysis of collected fractions. Black line represents absorbance at 280 nm, black dashed line the imidazole concentrations in mobile phase. SCOMT_His_6 (25 kDa) position in the SDS-PAGE gels in active fractions is represented by an arrow. (A) Adsorption performed at 0 mM imidazole (1mL/min), followed by an increasing linear gradient from 0 to 500 mM of imidazole and a final step at 500 mM imidazole in 150 mM NaCl, 50 mM Tris and 1 mM $MgCl_2$ buffer at pH 7.8 and 1mL/min. (B) Adsorption was carried out with 5 mM imidazole (0.5 mL/min), followed by four different steps at 50 mM, 70 mM, 300 mM and 500 mM imidazole in 500 mM, 50 mM Tris and 1 mM $MgCl_2$ buffer at pH 7.8 with a flow rate at 1.0 mL/min. SDS-PAGE (silver nitrate staining) and dot-blot analysis are depicted for each peak.

Along with dot-blot, western-blot analysis was carried out in the initial sample as well as in the fraction obtained at 300 mM imidazol to confirm if immunologically active SCOMT was present in a monomeric form. Indeed, as shown in Figure 3, it is possible to conclude that immunologically active and monomeric SCOMT with the correct molecular weight (25 KDa) is present either in the initial sample (lane II), as in the purified fraction (lane I). However, in the initial sample, it was also detected a band with a molecular weight close to 35 KDa that may correspond to SCOMT with the alpha mating factor still attached, as it was previously described for the production of the Necator americanus secretory protein in *P. pastoris* [35]. This may be due to the inefficient cleavage of the alpha mating factor by the Kex2 protease, leading to an increase of 7KDa in the size of the target protein [35] what is consistent to what was obtained in this work. In addition, it seems that the inefficient cleavage of the signal sequence by Kex2 is associated with the target gene copy number [35]. Nevertheless, in the purified fraction (Figure 3, lane I), only a band with a molecular weight of nearly 25 KDa is observed, indicating that the fractions containing biologically active and purified SCOMT correspond to the SCOMT with the native N-terminus.

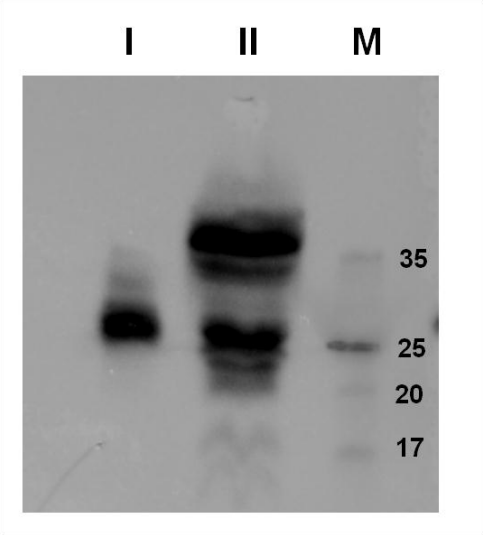


Figure 3 - Western-blot analysis of recombinant human SCOMT in the main fraction obtained during its purification using IMAC (Lane I –; IMAC Fraction obtained with 300 mM imidazol; Lane II – Ressuspended pellet; M – Molecular weight marker).

Then, we also wanted to evaluate if the target protein obtained in the purified fraction was in a biologically active form. Therefore, as demonstrated in Table 1, the fraction obtained at 300 mM imidazole showed a value of 121.1 nmol/h/mg of protein for SCOMT specific activity, a value that is much higher than that obtained for the sample that was initially injected onto the column (1.5 nmol/h/mg of protein for the ressuspended pellet). In fact, a purification fold of 81 was obtained with this process along with a bioactivity recovery of 57.4%, which are the best results ever obtained for the SCOMT chromatographic purification from recombinant lysates. Nonetheless, the application of IMAC as the main capture step for the purification of SCOMT-His_6 from recombinant lysates proved to be more efficient and

selective than the strategies previously applied using either hydrophobic interaction chromatography [10] or ion-exchange chromatography [11] where purifications fold of 1.8 and 3.6 were obtained, respectively. Moreover, even when a polishing step is applied after the hydrophobic interaction chromatography [10], the purification fold is lower (5.9) than that obtained with this work (81) with a single chromatographic step. Finally, after the application of the IMAC chromatographic step, highly purified SCOMT-His_6 was obtained in a concentration of 3.68 mg/L onto a shake-flask culture, a value that is much higher than those reported before (0.4 mg/L of *E. coli* shake flask culture) [10].

Table 1 - Purification of recombinant SCOMT_His_6 from a *P. pastoris* cell extract (results are the mean of 5 independent samples, N = 5).

Stage	Protein (mg/mL)	Protein recovery (%)	Specific Activity (nmol/h/mg of protein)	Total activity (nmol/h)	Bioactivity recovery (%)	Purification (fold)
Initial sample - Ressuspended pellet	21.139	100	1.5 ± 0.23	30.9 ± 0.634	100	1
IMAC – 300 mM imidazol	0.368	0.87	121.1 ± 5.1	17.74 ± 0.34	57.4	81

3.3 MALDI-TOF/TOF analysis of purified SCOMT-His_6

MALDI-TOF/TOF provides a simple and direct means to unequivocally confirm identity of recombinant proteins based on tryptic peptide analysis. Thus, MALDI-TOF/TOF analysis were performed in this work to complement the characterization of recombinant SCOMT-His_6 and to confirm that it was correctly processed. The recombinant protein was successfully identified as the soluble isoform of catechol *O*-methyltransferase (sp|P21964-2|COMT_HUMAN) with an unsed score of 11.87 and 45.25% (Supplementary material 1). The Figure 4 represents the MS spectrum obtained from SCOMT tryptic peptides analysis by MALDI-TOF/TOF, where the peaks corresponding to SCOMT-His_6 peptides that allowed its identification are highlighted in bold, as well as the respective identified sequences. However, based on these results it was not possible to find the tryptic peptide correspondent to the N-terminus of SCOMT-His_6. In general, the detection of N-terminal and C-terminal sequences is difficult, even further if the peptide correspondent to N-terminal sequence (MGDTK) is a low mass peptide of about 551.2494 Da. Moreover, the search engine applied for protein/peptide identification does not identify peptides with less than 5 amino acids. So, peptide mass tool from ExPASy Bioinformatics Resource Portal (http://web.expasy.org/peptide_mass/) was applied in order to theoretically calculate the

corresponding N-terminal fragment of unprocessed SCOMT sequence (Supplementary material 2). Theoretically, N-terminal fragment of unprocessed SCOMT corresponds to a high mass peptide of 7639.9624 Da, which would be easily identified in an MS spectrum. The spectrum obtained by MS analysis (Supplementary material 3) clearly confirms the absence of peptide corresponding to N-terminal unprocessed sequence, indicating the expressed and purified recombinant SCOMT was correctly processed.

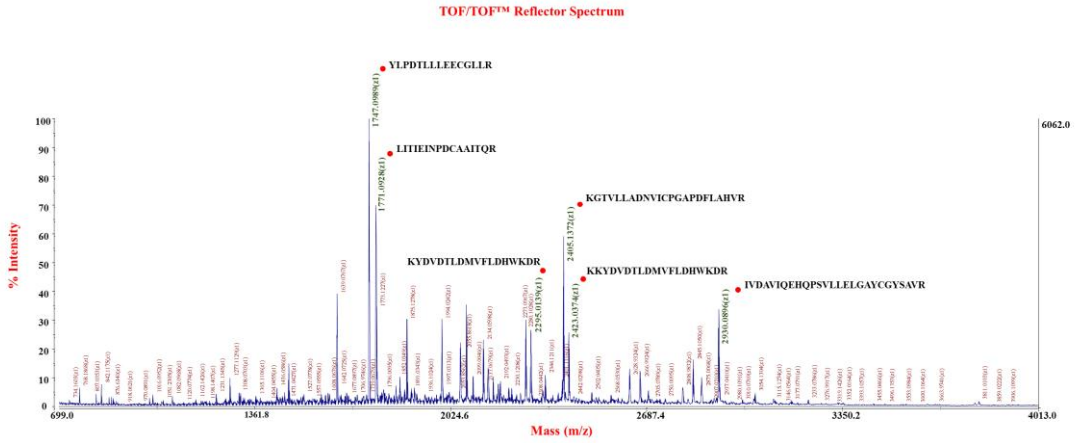


Figure 4 - Mass spectrometry spectrum obtained from recombinant sCOMT tryptic peptides analysis by MALDI-TOF/TOF. The peaks corresponding to sCOMT tryptic peptides that allowed its identification are highlighted in bold, as well as the respective identified sequences.

3.4 SCOMT inhibition

In general, there are two main classes of COMT inhibitors, the "first" and the "second" generation COMT inhibitors [1, 36]. While the first generation are generally characterized by weak or nonselectivity activity with dissociation constants in the micromolar range [1], the second generation COMT inhibitors, are tight-binding inhibitors with dissociation constants within the nanomolar range [1, 36]. These "second-generation" inhibitors, commonly called nitrocatecholic inhibitors include the tolcapone, nitecapone, entacapone, opicapone and 3,5-dinitrocatechol (3,5-DNC) [1, 36 – 38] In this work, the sensitivity of IMAC-purified recombinant SCOMT to be inhibited by 3,5-DNC and Entacapone was assessed (see Figure 5) and the IC_{50} for these compounds determined.

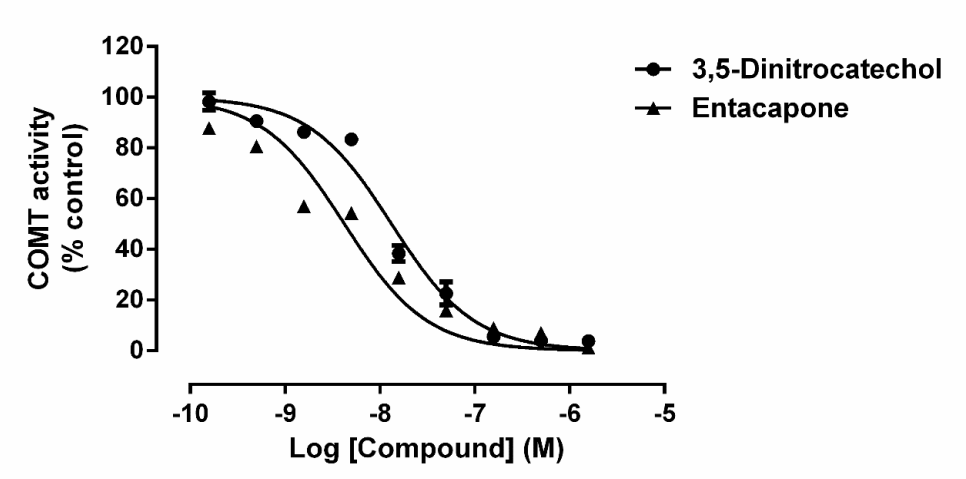


Figure 5 - Effect of increasing concentrations of 3,5-DNC (closed circles) and Entacapone (closed triangles) on the biological activity of recombinant SCOMT purified by IMAC using a fixed amount of SCOMT (150 µg/mL). COMT activity is shown as % of control (without inhibition). Each point represents the mean of two experiments while vertical lines show S. D.

Specifically, as shown in Table 2, the IC_{50}s for 3,5-DNC and Entacapone were, respectively, 13.26 and 4.224 nM. In general, differences in the experimental procedures employed to determine the IC_{50} of nitrocatechol-type inhibitors preclude an accurate and strictly comparison between the IC_{50} values obtained from different sources [38]. Nevertheless, these values are in the same order of magnitude than values previously reported in the literature using native SCOMT extracted from rat tissues (see Table 2).

Table 2 - IC_{50} values for inhibition by 3,5-DNC and Entacapone of SCOMT activity purified from IMAC obtained in this work and comparison with previous values reported in the literature. Values are means with 95% confidence values (n = 2).

Sample	Compound	IC_{50} (nM)	Reference
Recombinant purified SCOMT	3,5-DNC	13.26 (10.70 to 16.44)	This work
SCOMT from rat brain	3,5-DNC	28 (15 to 52)	[22]
SCOMT from rat liver		356 (156 to 811)	
Recombinant purified SCOMT	Entacapone	4.224 (2.949 to 6.050)	This work
SCOMT isolated from rat duodenum	Entacapone	10.0	[39]
SCOMT isolated from rat erythrocytes		20.0	
SCOMT isolated from rat liver		160.0	

4. Conclusion

COMT has been implicated in several human diseases including Parkinson's disease in which the most appropriated therapy depends on the efficacy of the COMT inhibitors applied. Studies intended to develop new and more effective drugs are usually dependent on the achievement of samples with high purity and quantity. Most of the previously reported procedures for the isolation of SCOMT are time-consuming, involving multiple steps where the loss of the enzyme activity is significant in the earlier stages of the process. Therefore, in this work a new process was implemented for the synthesis of SCOMT-His_6 in *P. pastoris*, followed by a purification strategy involving IMAC as the main capture step. Nonetheless, while IMAC proved to be extremely efficient and selective for the direct capture of hexahistidine tagged SCOMT from recombinant *P. pastoris* lysates, a purification fold of 81 and a bioactivity recovery of 57.4% were obtained, the best results ever reported for this protein. Moreover, after the application of the IMAC step, SCOMT-His_6 was obtained at a concentration of 3.68 mg/L from shake-flask culture. Also, the MALDI-TOF/TOF analysis of the recombinant form of SCOMT produced in this work allow us to conclude that the target purified protein is well processed during its biosynthesis in *P. pastoris* showing a native primary sequence. Finally, the sensitivity of the purified protein fractions to be inhibited by 3,5-DNC and Entacapone was evaluated and it was found that the determined IC_{50} values are in agreement with previous values reported in the literature. In conclusion, a new strategy was developed and implemented for the biosynthesis and purification of biologically active and highly purified SCOMT.

5. Acknowledgements

This research was supported by University of Beira Interior—Health Sciences Research Centre (CICS) and FCT (Portuguese Foundation for Sciences and Technology) by the project "EXPL/BBB478/BQB/0960/2012" and COMPETE: FCOMP-01-0124-FEDER-027563. F. F. Correia and G. Espírito-Santo acknowledge a fellowship from the project "EXPL/BBB478/BQB/0960/2012" while A.Q. Pedro acknowledges a doctoral fellowship (SFRH/BD/81222/2011) from FCT and A. M. Gonçalves a fellowship from the project "Technologies for purification and controlled release of biopharmaceuticals to be applied in age-related disease" (CENTRO-07-ST24_FEDER-002014), financed by the program 2007-2013 QREN. The authors also acknowledge the program COMPETE, the FCT project (Pest-C/SAU/UI0709/2011).

6. References

[1] Bonifacio MJ, Palma PN, Almeida L, Soares-da-Silva P, Catechol-*O*-methyltransferase and its inhibitors in Parkinson's disease, CNS Drug Rev 13 (3) 352-379 (2007).

[2] Lotta T, Vidgren J, Tilgmann C, Ulmanen I, Melen K, Julkunen I, Taskinen J, Kinetics of human soluble and membrane-bound catechol *O*-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme, Biochemistry 34 (13): 4202-4210 (1995).

[3] Myohanen TT and Mannisto PT, Distribution and functions of catechol-*O*-methyltransferase proteins: do recent findings change the picture? Int Rev Neurobiol 95: 29-47 (2010).

[4] Zhu BT, On the mechanism of homocysteine pathophysiology and pathogenesis: a unifying hypothesis, Histol Histopathol 17 (4): 1283-1291 (2002).

[5] Zhu BT and Liehr JG, Inhibition of the catechol-*O*-methyltransferase-catalyzed O-methylation of 2- and 4-hydroxyestradiol by catecholamine: implications for the mechanism of estrogen-induced carcinogenesis, Arch Biochem Biophys 304 (1): 248-256 (1993).

[6] Junge F, Schneider B, Reckel S, Schwarz D, Dotsch V, Bernhard F, Large-scale production of functional membrane proteins, Cell Mol Life Sci 65 (11): 1729-1755 (2008).

[7] Cregg JM, Cereghino JL, Shi J, Higgins DR, Recombinant protein expression in *Pichia pastoris*, Mol Biotechnol 16 (1): 23-52 (2000).

[8] Nunes VS, Bonifacio MJ, Queiroz JA, Passarinha LA, Assessment of COMT isolation by HIC using a dual salt system and low temperature, Biomed Chromatogr 24 (8): 858-862 (2010).

[9] Passarinha LA, Bonifacio MJ, Queiroz JA, Comparative study on the interaction of recombinant human soluble catechol-*O*-methyltransferase on some hydrophobic adsorbents, Biomed Chromatogr 21 (4): 430-438 (2007).

[10] Passarinha LA, Bonifacio MJ, Soares-da-Silva P, Queiroz JA, A new approach on the purification of recombinant human soluble catechol-*O*-methyltransferase from an Escherichia coli extract using hydrophobic interaction chromatography, J Chromatogr A 1177 (2): 287-296 (2008).

[11] Correia FF, Santos FM, Pedro AQ, Bonifacio MJ, Queiroz JA, Passarinha LA, Recovery of biological active catechol-*O*-methyltransferase isoforms from Q-sepharose, J Sep Sci 37 (1-2): 20-29 (2014).

[12] Costa SR, Bonifacio MJ, Queiroz JA, Passarinha LA, Analysis of hSCOMT adsorption in bioaffinity chromatography with immobilized amino acids: the influence of pH and ionic strength, J Chromatogr B Analyt Technol Biomed Life Sci 879 (19): 1704-1706 (2011).

[13] Mooney JT, Fredericks DP, Zhang C, Christensen T, Jespergaard C, Schiodt CB, Hearn MT, Purification of a recombinant human growth hormone by an integrated IMAC procedure, Protein Expr Purif 94: 85-94 (2014).

[14] Boden V, Winzerling JJ, Vijayalakshmi M, Porath J, Rapid one-step purification of goat immunoglobulins by immobilized metal ion affinity chromatography, J Immunol Methods 181 (2): 225-232 (1995).

[15] Rodrigues ES, Verinaud CI, Oliveira DS, Raw I, Lopes AP, Martins EA, Cheng E, Purification of coagulation factor VIII by immobilized metal affinity chromatography, Biotechnol Appl Biochem (2014) doi: 10.1002/bab.1276.

[16] Skarka A, Skarydova L, Stambergova H, Wsol V, Purification and reconstitution of human membrane-bound DHRS7 (SDR34C1) from Sf9 cells, Protein Expr Purif 95: 44-49 (2014).

[17] Gaberc-Porekar V, Menart V, Perspectives of immobilized-metal affinity chromatography, J Biochem Biophys Methods 49 (1-3): 335-360 (2001).

[18] Cheung RCJ, Wong JH, Ng TB, Immobilized metal ion affinity chromatography: a review on its applications, Appl Microbiol Biotechnol 96 (6): 1411-1420 (2012).

[19] Fanou-Ayi L, Vijayalakshmi M, Metal-chelate affinity chromatography as a separation tool, Ann N Y Acad Sci 413: 300-306 (1983).

[20] Tenhunen J, Salminen M, Lundstrom K, Kiviluoto T, Savolainen R, Ulmanen I, Genomic organization of the human catechol-*O*-methyltransferase gene and its expression from two distinct promoters, Eur J Biochem 223 (3): 1049-1059 (1994).

[21] Pedro AQ, Soares RF, Oppolzer D, Santos FM, Rocha LA, Gonçalves AM, Bonifacio MJ, Queiroz JA, Gallardo E, Passarinha LA, An improved HPLC method for quantification of metanephrine with coulometric detection, J Chromatograph Separat Techniq 5 (217) (2014).

[22] Vieira-Coelho MA, Soares-da-Silva P, Effects of tolcapone upon soluble and membrane-bound brain and liver catechol-*O*-methyltransferase, Brain Res 821: 69-78 (1999).

[23] Laemmli UK, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (5259) 680-685 (1970).

[24] Goncalves AM, Pedro AQ, Maia C, Sousa F, Queiroz JA, Passarinha LA, *Pichia pastoris*: a recombinant microfactory for antibodies and human membrane proteins. J Microbiol Biotechnol 23 (5): 587-601 (2013).

[25] Silva R, Ferreira S, Bonifacio MJ, Dias JML, Queiroz JA, Passarinha LA, Optimization of fermentation conditions for the production of human soluble catechol-*O*-methyltransferase by *Escherichia coli* using artificial neural network, J Biotechnol 160: 161-168 (2012).

[26] Daly R and Hearn MTW. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production, J Mol Recognit 18: 119-138 (2005).

[27] Yang S, Kuang Y, Li H, Liu Y, Hui X, Li P, Jiang Z, Zhou Y, Wang Y, Xu A, Liu P, Wu D. Enhanced production of recombinant secretory proteins in *Pichia pastoris* by optimizing Kex2 P1' site, PLoS One 8 (9): e75347 (2013).

[28] Pedro AQ, Oppolzer D, Bonifacio MJ, Maia CJ, Queiroz JA, Passarinha LA, Evaluation of Mut^S and Mut^+ *Pichia pastoris* strains for membrane-bound and Catechol-*O*-methyltransferase biosynthesis. Appl Biochem Biotechnol 175: 3840-3855 (2015).

[29] Zhang YB, Howitt J, McCorkle S, Lawrence P, Springer K, Freimuth P, Protein aggregation during overexpression limited by peptide extensions with large net negative charge, Protein Expr Purif 36 (2): 207-216 (2014).

[30] Li P, Anumanthan A, Gao XG, Ilangovan K, Suzara VV, Duzgunes N, Renugopalakrishnan V, Expression of recombinant proteins in *Pichia pastoris*, Appl Biochem Biotechnol 142 (2): 105-124 (2007).

[31] Terpe K, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, Appl Microbiol Biotechnol 60 (5): 523-533 (2003).

[32] Fazen CH, Kahkoska AR, Doyle RP, Expression and purification of human PYY(3-36) in Escherichia coli using a His-tagged small ubiquitin-like modifier fusion. Protein Expr Purif 85 (1) (2012) 51-59.

[33] Nieba L, Nieba-Axmann SE, Persson A, Hamalainen M, Edebratt F, Hansson A, Lidholm J, Magnusson K, Karlsson AF, Pluckthun A, BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. Anal Biochem 252 (2): 217-228 (1997).

[34] Bornhorst JA, Falke JJ, Purification of proteins using polyhistidine affinity tags. Methods Enzymol 326: 245-254 (2000).

[35] Inan M, Aryasomayajula D, Sinha J, Meagher MM. Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. Biotechnol Bioeng 93 (4): 771-778 (2006).

[36] Bonifati V, Meco G, New, Selective Catechol-*O*-methyltransferase Inhibitors as Therapeutic Agents in Parkinson's Disease. Pharmacol Ther 81 (1): 1-36 (1999).

[37] Kiss LE, Ferreira HS, Torrão L, Bonifacio MJ, Palma PN, Soares-da-Silva P, Learmonth DA, Discovery of a long-acting peripherally selective inhibitor of catechol-*O*-methyltransferase. J Med Chem, 53 (8): 3396-3411 (2010).

[38] Mannisto PT, Kaakkola S, Catechol-*O*-methyltransferase (COMT): Biochemistry, molecular biology, pharmacology, and clinical efficacy of the new selective COMT inhibitors. Pharmacol Rev 51 (4): 594-628 (1999).

[39] Nissinen E, Lindén IB, Schultz E, Pohto P, Biochemical and pharmacological properties of a peripherally acting catechol-*O*-methyltransferase inhibitor Entacapone. Naunyn Schmiedebergs Arch Pharmacol 346: 262-266 (1992).

7. Supplementary Material

7.1 Supplementary material

Sequence Coverage

Canonical Sequence coverage

Protein Sequence Coverage - Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2

MPEAPPLLLAAVLLGLVLLVVLLLLLRHWGWGLCLIGWNEFILQPIHNLLMGDTKEQRILNHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGK**IVDAVIQEHQPSVLLELGAYCGYSAVR**MARLLSPGAR**LITIEINPDCAAITQR**MVDFAGVKDKVTLVVGASQDIIPQ
LK**KKYDVDTLDMVFLDHWKDRYLPDTLLLEECGLLRKGTVLLADNVICPGAPDFLAHVR**GSSCFECTHYQSFLEYREVVDGLEKAIYKGPGSEAGP

Isoform Soluble of Catechol O-methyltransferase Sequence coverage

Protein Sequence Coverage - Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT

MGDTKEQRILNHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGK**IVDAVIQEHQPSVLLELGAYCGYSAVR**MARLLSPGAR**LITIEINPDCAAITQR**MVDFAGVKDKVTLVVGASQDIIPQLK**KKYDVDTLDMVFLDHWKDRYLPDTLLLEECGLLRKGTVLLADNVICPG**
APDFLAHVRGSSCFECTHYQSFLEYREVVDGLEKAIYKGPGSEAGP

Human database_Protein Summary

Proteins detected								
Unused	**Total**	**%Cov**	**%Cov(50)**	**%Cov(95)**	**Accession**	**Name**	**Species**	**Peptides(95%)**
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2	HUMAN	9
0	11.87	45,25	45,25	45,25	sp|P21964-2|COMT_HUMAN	Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	HUMAN	9

Human database_Peptide Summary

Peptide Summary

Unused	Total	% Cov	% Cov(50)	% Cov(95)	Accessions	Names	Contrib	Conf	Sequence	Modifications	Cleavages	dMass	Prec MW	Prec m/z	Theor MW	Theor m/z	Theor z	Sc	Spectrum
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	2	99.0000009536743	GTVLLADNVICPGAPDFLAHVR	Ala->Ser@6		0,1820	2293,3706	2294,3780	2293,1890	2294,1960	1	11	66,1,1,77500,75655
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	2	99.0000009536743	IVDAVIQEHQPSVLLELGAYCGYSAVR			-0,1409	2929,3596	2930,3670	2929,5007	2930,5081	1	7	20,1,1,77841,76115
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	2	99.0000009536743	KGTVLLADNVICPGAPDFLAHVR		missed K-G@1	-0,1112	2405,1777	2406,1850	2405,2888	2406,2961	1	7	21,1,1,77842,76126
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	2	99.0000009536743	KKYDVDTLDMVFLDHWKDR		missed K-K@1; missed K-Y@2; missed K-D@17	-0,1130	2423,0818	2424,0890	2423,1943	2424,2017	1	8	1,1,1,77809,75966
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	2	99.0000009536743	YLPDTLLLEECGLLR			-0,0676	1746,8547	1747,8620	1746,9226	1747,9299	1	8	17,1,1,77838,76095
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	1.2291479110717 8	94.0699994564056	KYDVDTLDMVFLDHWKDR		missed K-Y@1; missed K-D@16	-0,0979	2295,0017	2296,0090	2295,0994	2296,1067	1	6	1,1,1,77809,75962
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0.6382721066647491	99.0000009536743	LITIEINPDCAAITQR			-0,8431	1769,0917	1770,0990	1769,9346	1770,9419	1	9	62,1,1,77543,75506

11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	GTVLLADNVICPGAPDFLAHVR	Ala->Ser@6		0,1820	2293,3706	2294,3780	2293,1890	2294,1960	1	11	67,1,1,77500,75655
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	GTVLLADNVICPGAPDFLAHVR	Ala->Ser@6		0,1820	2293,3706	2294,3780	2293,1890	2294,1960	1	11	68,1,1,77500,75655
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	GTVLLADNVICPGAPDFLAHVR			-0,0943	2277,0996	2278,1070	2277,1941	2278,2012	1	7	17,1,1,77838,76092
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	GTVLLADNVICPGAPDFLAHVR			-0,0794	2277,1147	2278,1220	2277,1941	2278,2012	1	6	16,1,1,77837,76084
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	GTVLLADNVICPGAPDFLAHVR			-0,1077	2277,0867	2278,0940	2277,1941	2278,2012	1	6	19,1,1,77840,76108
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	97.9900002479553	GTVLLADNVICPGAPDFLAHVR			-0,1009	2277,0928	2278,1000	2277,1941	2278,2012	1	6	21,1,1,77842,76123
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	57.0299983024597	GTVLLADNVICPGAPDFLAHVR			-0,0919	2277,1016	2278,1090	2277,1941	2278,2012	1	5	18,1,1,77839,76100
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	IVDAVIQEHQPSVLLELGAYCGYSAVR			-0,1042	2929,3967	2930,4040	2929,5007	2930,5081	1	6	16,1,1,77837,76081
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2;	0	94.5900022983551	IVDAVIQEHQPSVLLELGAYCGYSAVR			-0,1274	2929,3738	2930,3810	2929,5007	2930,5081	1	6	18,1,1,77839,76097

					2|COMT_HUMAN	Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT												
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	KGTVLLADNVICPGAPDFLAHVR	Ala->Ser@7	missed K-G@1	0,0629	2421,3467	2422,3540	2421,2839	2422,2910	1	9	66,1,1,77500,75660
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	KGTVLLADNVICPGAPDFLAHVR	Ala->Ser@7	missed K-G@1	0,0629	2421,3467	2422,3540	2421,2839	2422,2910	1	9	67,1,1,77500,75660
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	KGTVLLADNVICPGAPDFLAHVR	Ala->Ser@7	missed K-G@1	0,0629	2421,3467	2422,3540	2421,2839	2422,2910	1	9	68,1,1,77500,75660
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	KGTVLLADNVICPGAPDFLAHVR	Leu->Gln@6	missed K-G@1	-0,2827	2419,9807	2420,9880	2420,2634	2421,2708	1	9	63,1,1,77499,75650
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	KGTVLLADNVICPGAPDFLAHVR	Leu->Gln@6	missed K-G@1	-0,2827	2419,9807	2420,9880	2420,2634	2421,2708	1	9	64,1,1,77499,75650
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	KGTVLLADNVICPGAPDFLAHVR	Leu->Gln@6	missed K-G@1	-0,2827	2419,9807	2420,9880	2420,2634	2421,2708	1	9	65,1,1,77499,75650
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.0000009536743	KGTVLLADNVICPGAPDFLAHVR		missed K-G@1	-0,1126	2405,1768	2406,1840	2405,2888	2406,2961	1	6	19,1,1,77840,76110
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	97.7800011634827	KGTVLLADNVICPGAPDFLAHVR		missed K-G@1	-0,1141	2405,1746	2406,1820	2405,2888	2406,2961	1	6	20,1,1,77841,76118

11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	84.68000292778 02	KGTVLLADNVICPGAPDFLAHVR	Leu->Met@6	missed K-G@1	0,1656	2423,4106	2424,4180	2423,2454	2424,2527	1	7	61,1,1,77542,75493
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	43.18999946117 4	KGTVLLADNVICPGAPDFLAHVR	Lys->Phe@1	missed K-G@1	0,5289	2424,7917	2425,7990	2424,2625	2425,2698	1	7	60,1,1,77541,75483
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	81.91999793052 67	KKYDVDTLDMVFLDHWKDR		missed K-K@1; missed K-Y@2; missed K-D@17	-0,1113	2423,0828	2424,0900	2423,1943	2424,2017	1	6	2,1,1,77810,75974
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	99.00000095367 43	LITIEINPDCAAITQR			-0,0795	1769,8547	1770,8620	1769,9346	1770,9419	1	8	20,1,1,77841,76117
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	71.38001306533 8	LITIEINPDCAAITQR			-0,0701	1769,8646	1770,8720	1769,9346	1770,9419	1	6	17,1,1,77838,76093
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	63.92999887466 43	LITIEINPDCAAITQR			-0,0598	1769,8746	1770,8820	1769,9346	1770,9419	1	6	16,1,1,77837,76085
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	63.83000016212 46	LITIEINPDCAAITQR			-0,0672	1769,8677	1770,8750	1769,9346	1770,9419	1	6	18,1,1,77839,76101
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	48.53999912731 88	LITIEINPDCAAITQR			-0,0765	1769,8577	1770,8650	1769,9346	1770,9419	1	6	19,1,1,77840,76109
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2;	0	48.10999929904 94	LITIEINPDCAAITQR			-0,0714	1769,8627	1770,8700	1769,9346	1770,9419	1	6	21,1,1,77842,76125

					2|COMT_HUMAN	Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT												
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	31.5499991178513	LITIEINPDCAAITQR			-0,0531	1769,8817	1770,8890	1769,9346	1770,9419	1	5	15,1,1,77836,76077
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	24.3100002408028	LITIEINPDCAAITQR			-0,0446	1769,8896	1770,8970	1769,9346	1770,9419	1	5	14,1,1,77835,76069
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	97.759997844696	YLPDTLLLEECGLLR			-0,0459	1746,8767	1747,8840	1746,9226	1747,9299	1	7	13,1,1,77834,76063
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	97.6999998092651	YLPDTLLLEECGLLR			-0,0578	1746,8646	1747,8720	1746,9226	1747,9299	1	7	16,1,1,77837,76087
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	96.670001745224	YLPDTLLLEECGLLR			-0,0722	1746,8507	1747,8580	1746,9226	1747,9299	1	7	21,1,1,77842,76127
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	95.3599989414215	YLPDTLLLEECGLLR			-0,0766	1746,8457	1747,8530	1746,9226	1747,9299	1	7	20,1,1,77841,76119
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	78.4500002861023	YLPDTLLLEECGLLR			-0,0670	1746,8557	1747,8630	1746,9226	1747,9299	1	6	18,1,1,77839,76103
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	64.5699977874756	YLPDTLLLEECGLLR			-0,0770	1746,8457	1747,8530	1746,9226	1747,9299	1	6	19,1,1,77840,76111

11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	64.410039505	YLPDTLLLEECGLLR			-0,0892	1746,8336	1747,8410	1746,9226	1747,9299	1	6	1,1,1,7780 9,75963
11.87	11.87	36,90	36,90	36,90	sp|P21964|COMT_HUMAN; sp|P21964-2|COMT_HUMAN	Catechol O-methyltransferase OS=Homo sapiens GN=COMT PE=1 SV=2; Isoform Soluble of Catechol O-methyltransferase OS=Homo sapiens GN=COMT	0	55.8499991893768	YLPDTLLLEECGLLR			-0,0874	1746,8357	1747,8430	1746,9226	1747,9299	1	6	2,1,1,7781 0,75972

7.2. Supplementary Material 2

In Silico digestion from unprocessed and processed recombinant sCOMT

sCOMT sequence:

Obtained from UniProtKB [http://www.uniprot.org/uniprot/P21964#sequences].

MGDTKEQRILNHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGKIVDAVIQEHQPSVLLELGAYCG
YSAVRMARLLSPGARLITIEINPDCAAITQRMVDFAGVKDKVTLVVGASQDIIPQLKKKYDVDTLDMVFL
DHWKDRYLPDTLLLEECGLLRKGTVLLADNVICPGAPDFLAHVRGSSCFECTHYQSFLEYREVVDGLEKA
IYKGPGSEAGP

Unprocessed sCOMT sequence:

Calculated based on α-factor signal sequence from expression vector.

**MRFPSIFTAVLFAASSALAAPVNTTTENETAQIPAEAVIGTSVLEGDFDVAVVLPFSNSTNNGLLFINTT
IASIAAKEEGVS**MGDTKEQRILNHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGKIVDAVIQEHQ
PSVLLELGAYCGYSAVRMARLLSPGARLITIEINPDCAAITQRMVDFAGVKDKVTLVVGASQDIIPQLKK
KYDVDTLDMVFLDHWKDRYLPDTLLLEECGLLRKGTVLLADNVICPGAPDFLAHVRGSSCFECTHYQSFL
EYREVVDGLEKAIYKGPGSEAGP

PeptideMass

The entered sequence is:

```
        10         20         30         40         50         60
MGDTKEQRIL NHVLQHAEPG NAQSVLEAID TYCEQKEWAM NVGDKKGKIV DAVIQEHQPS

        70         80         90        100        110        120
VLLELGAYCG YSAVRMARLL SPGARLITIE INPDCAAITQ RMVDFAGVKD KVTLVVGASQ

       130        140        150        160        170        180
DIIPQLKKKY DVDTLDMVFL DHWKDRYLPD TLLLEECGLL RKGTVLLADN VICPGAPDFL

       190        200        210        220
AHVRGSSCFE CTHYQSFLEY REVVDGLEKA IYKGPGSEAG P
```

The selected enzyme is: Trypsin

Maximum number of missed cleavages (MC): 0

All cysteines in reduced form.

Methionines have not been oxidized.

Displaying peptides with a mass bigger than 500 Dalton.

Using monoisotopic masses of the occurring amino acid residues and giving peptide masses as $[M+H]^+$.

The peptide masses from your sequence are:

[Theoretical pI: 5.15 / Mw (average mass): 24449.08 / Mw (monoisotopic mass): 24433.38]

mass	position	#MC	modifications	peptide sequence
3120.5418	9-36	0		ILNHVLQHAEPGNAQSVLEA IDTYCEQK
2930.5080	49-75	0		IVDAVIQEHQPSVLLELGAY CGYSAVR
2278.2012	163-184	0		GTVLLADNVICPGAPDFLAH VR
2056.8528	185-201	0		GSSCFECTHYQSFLEYR
1896.8836	130-144	0		YDVDTLDMVFLDHWK
1770.9418	86-101	0		LITIEINPDCAAITQR
1747.9298	147-161	0		YLPDTLLLEECGLLR
1680.9894	112-127	0		VTLVVGASQDIIPQLK
1049.4720	37-45	0		EWAMNVGDK
888.4672	202-209	0		EVVDGLEK
866.4440	102-109	0		MVDFAGVK
713.4304	79-85	0		LLSPGAR
671.2995	214-221	0		GPGSEAGP
551.2494	1-5	0		MGDTK

91.0% of sequence covered (you may modify the input parameters to display also peptides < 500 Da or > 100000000000 Da):

```
        10         20         30         40         50         60
MGDTKeqrIL NHVLQHAEPG NAQSVLEAID TYCEQKEWAM NVGDKkgkIV DAVIQEHQPS

        70         80         90        100        110        120
VLLELGAYCG YSAVRmarLL SPGARLITIE INPDCAAITQ RMVDFAGVKd kVTLVVGASQ

       130        140        150        160        170        180
DIIPQLKkkY DVDTLDMVFL DHWKdrYLPD TLLLEECGLL RkGTVLLADN VICPGAPDFL
```

```
       190        200        210        220
AHVRGSSCFE CTHYQSFLEY REVVDGLEKa iykGPGSEAG P
```

Display the list of masses in raw text format to be exported into an external application

PeptideMass

The entered sequence is:

```
        10         20         30         40         50         60
MRFPSIFTAV LFAASSALAA PVNTTTENET AQIPAEAVIG TSVLEGDFDV AVVLPFSNST

        70         80         90        100        110        120
NNGLLFINTT IASIAAKEEG VSMGDTKEQR ILNHVLQHAE PGNAQSVLEA IDTYCEQKEW

       130        140        150        160        170        180
AMNVGDKKGK IVDAVIQEHQ PSVLLELGAY CGYSAVRMAR LLSPGARLIT IEINPDCAAI

       190        200        210        220        230        240
TQRMVDFAGV KDKVTLVVGA SQDIIPQLKK KYDVDTLDMV FLDHWKDRYL PDTLLLEECG

       250        260        270        280        290        300
LLRKGTVLLA DNVICPGAPD FLAHVRGSSC FECTHYQSFL EYREVVDGLE KAIYKGPGSE


AGP
```

The selected enzyme is: Trypsin

Maximum number of missed cleavages (MC): 0

All cysteines in reduced form.

Methionines have not been oxidized.

Displaying peptides with a mass bigger than 500 Dalton.

Using monoisotopic masses of the occurring amino acid residues and giving peptide masses as $[M+H]^+$.

The peptide masses from your sequence are:

[Theoretical pI: 4.80 / Mw (average mass): 32863.56 / Mw (monoisotopic mass): 32842.67]

mass	position	#MC	modifications	peptide sequence
7639.9624	3-77	0		FPSIFTAVLFAASSALAAPV NTTTENETAQIPAEAVIGTS VLEGDFDVAVVLPFSNSTNN GLLFINTTIASIAAK
3120.5418	91-118	0		ILNHVLQHAEPGNAQSVLEA IDTYCEQK
2930.5080	131-157	0		IVDAVIQEHQPSVLLELGAY CGYSAVR
2278.2012	245-266	0		GTVLLADNVICPGAPDFLAH VR
2056.8528	267-283	0		GSSCFECTHYQSFLEYR
1896.8836	212-226	0		YDVDTLDMVFLDHWK
1770.9418	168-183	0		LITIEINPDCAAITQR
1747.9298	229-243	0		YLPDTLLLEECGLLR
1680.9894	194-209	0		VTLVVGASQDIIPQLK
1052.4564	78-87	0		EEGVSMGDTK
1049.4720	119-127	0		EWAMNVGDK
888.4672	284-291	0		EVVDGLEK
866.4440	184-191	0		MVDFAGVK
713.4304	161-167	0		LLSPGAR
671.2995	296-303	0		GPGSEAGP

92.7% of sequence covered (you may modify the input parameters to display also peptides < 500 Da or > 100000000000 Da):

```
        10         20         30         40         50         60
mrFPSIFTAV LFAASSALAA PVNTTTENET AQIPAEAVIG TSVLEGDFDV AVVLPFSNST

        70         80         90        100        110        120
NNGLLFINTT IASIAAKEEG VSMGDTKeqr ILNHVLQHAE PGNAQSVLEA IDTYCEQKEW

       130        140        150        160        170        180
AMNVGDKkgk IVDAVIQEHQ PSVLLELGAY CGYSAVRmar LLSPGARLIT IEINPDCAAI

       190        200        210        220        230        240
TQRMVDFAGV KdkVTLVVGA SQDIIPQLKk kYDVDTLDMV FLDHWKdrYL PDTLLLEECG

       250        260        270        280        290        300
LLRkGTVLLA DNVICPGAPD FLAHVRGSSC FECTHYQSFL EYREVVDGLE KaiykGPGSE


AGP
```

Display the list of masses in raw text format to be exported into an external application

7.3. Supplementary Material 3

Figure 1 – Mass spectrometry spectrum obtained from recombinant sCOMT tryptic peptides analysis by MALDI-TOF/TOF (positive MS mode in the range 800 to 8000 m/z, 1500 laser shots)

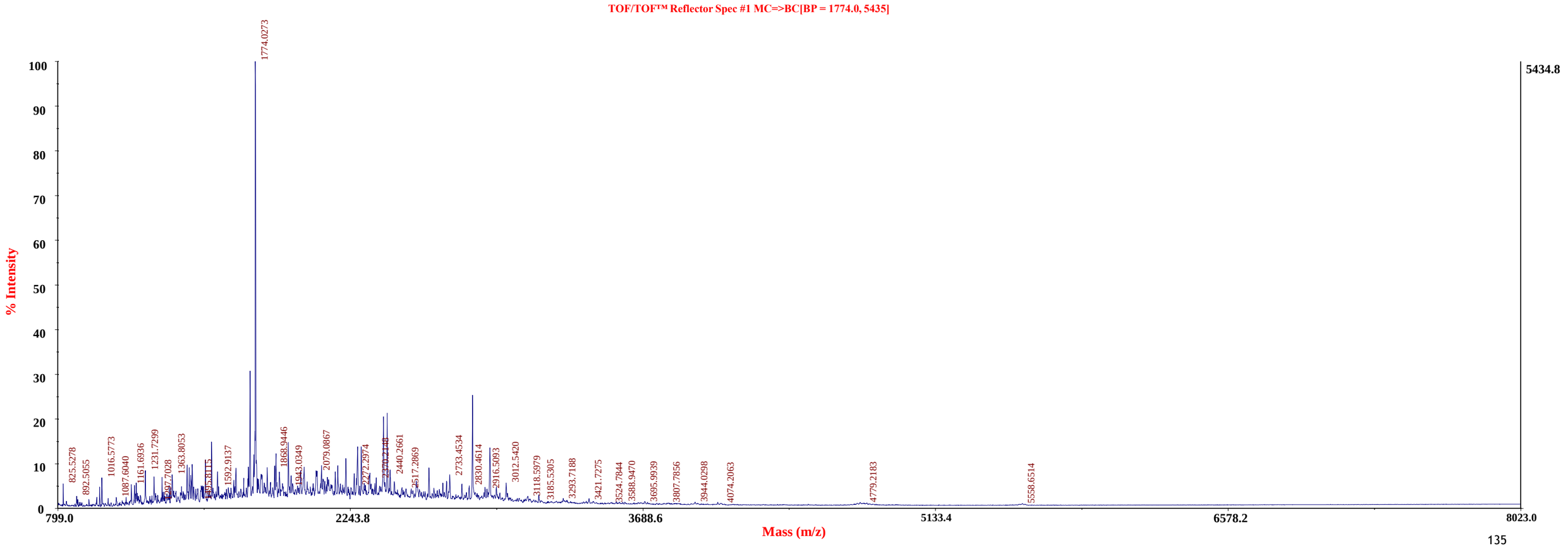
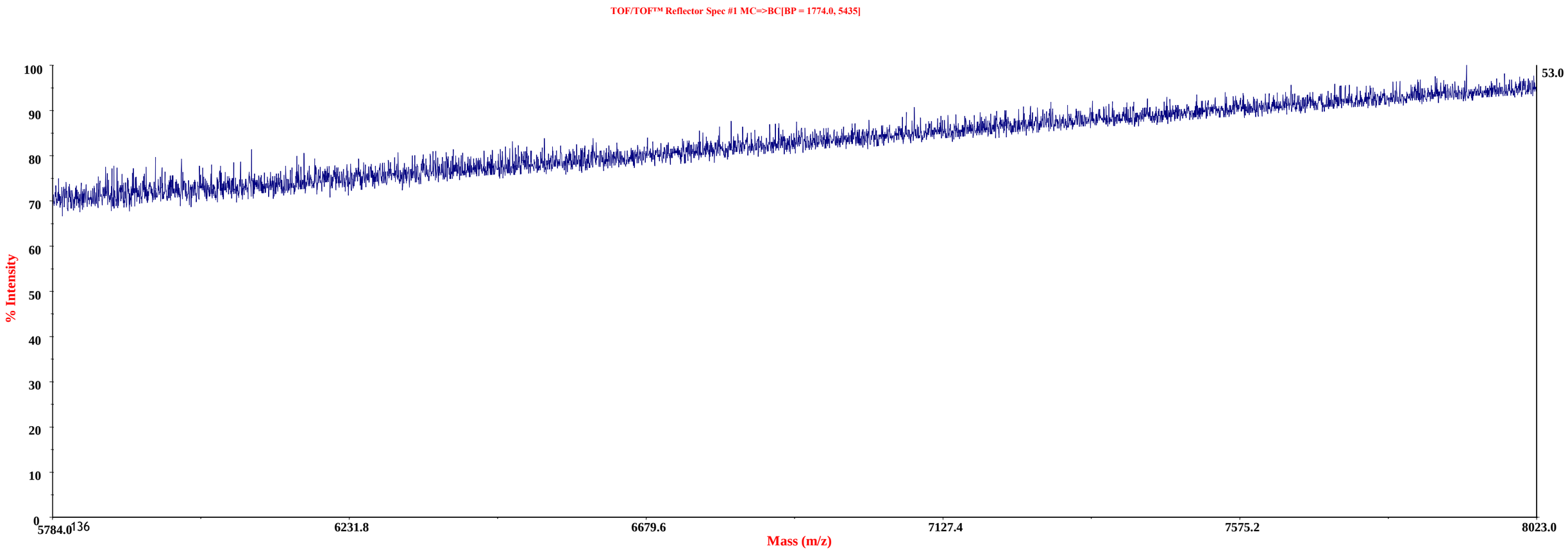


Figure 2 – Mass spectrometry spectrum obtained from recombinant sCOMT tryptic peptides analysis by MALDI-TOF/TOF (positive MS mode in the range 800 to 8000 m/z, 1500 laser shots). Zoom in of spectrum of Figure 1 in the mass range from 5700 to 8000.



Paper IV

Evaluation of Mut^S and Mut$^+$ *Pichia pastoris* strains for membrane-bound catechol-*O*-methyltransferase biosynthesis

A. Q. Pedro, D. Oppolzer, M. J. Bonifácio, J. A. Queiroz, L. A. Passarinha

Applied Biochemistry and Biotechnology. 2015. 175 (8): 3840-3855

Short description: Previously, we developed a successful strategy for production and purification of SCOMT-His$_6$ (Paper III). Here, our focus is to report a global strategy for the biosynthesis of recombinant human MBCOMT in its native form (Val 158) on a small-scale with *P. pastoris* methanol-induced cultures. Therefore, the influence of *P. pastoris* phenotype as well as the feed strategy (methanol or mixed-feed with glycerol or sorbitol) on the target protein production levels was determined. Also, in order to promote the recovery of biologically active MBCOMT, different *Pichia pastoris* lysis methods were tested. Finally, the kinetic parameters K_m and V_{Max} were determined in the different fractions obtained after MBCOMT compartmentalization.

Appl Biochem Biotechnol (2015) 175:3840–3855
DOI 10.1007/s12010-015-1551-0

Evaluation of Mut^S and Mut$^+$ *Pichia pastoris* Strains for Membrane-Bound Catechol-*O*-Methyltransferase Biosynthesis

A. Q. Pedro · D. Oppolzer · M. J. Bonifácio · C. J. Maia · J. A. Queiroz · L. A. Passarinha

Received: 14 April 2014 / Accepted: 15 February 2015 /
Published online: 25 February 2015

Abstract Catechol-*O*-methyltransferase (COMT, EC 2.1.1.6) is an enzyme that catalyzes the methylation of catechol substrates, and while structural and functional studies of its membrane-bound isoform (MBCOMT) are still hampered by low recombinant production, *Pichia pastoris* has been described as an attractive host for the production of correctly folded and inserted membrane proteins. Hence, in this work, MBCOMT biosynthesis was developed using *P. pastoris* X33 and KM71H cells in shake flasks containing a semidefined medium with different methanol concentrations. Moreover, after *P. pastoris* glass beads lysis, biologically and immunologically active hMBCOMT was found mainly in the solubilized membrane fraction whose kinetic parameters were identical to its correspondent native enzyme. In addition, mixed feeds of methanol and glycerol or sorbitol were also employed, and its levels quantified using liquid chromatography coupled to refractive index detection. Overall, for the first time, two *P. pastoris* strains with opposite phenotypes were applied for MBCOMT biosynthesis under the control of the strongly methanol-inducible alcohol oxidase (AOX) promoter. Moreover, this eukaryotic system seems to be a promising approach to deliver MBCOMT in high quantities from fermentor cultures with a lower cost-benefit due to the cheaper cultivation media coupled with the higher titers tipically achieved in biorreactors, when compared with previously reported mammallian cell cultures.

Keywords COMT · *Pichia pastoris* · Membrane protein · Methanol feeding · Alcohol oxidase promoter · Fermentation

Abbreviations

AOX	Alcohol oxidase
COMT	Catechol-*O*-methyltransferase
E. coli	*Escherichia coli*

A. Q. Pedro · D. Oppolzer · C. J. Maia · J. A. Queiroz · L. A. Passarinha (✉)
CICS-UBI–Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, Av. Infante D. Henrique, 6201-001 Covilhã, Portugal
e-mail: lapp@ubi.pt

M. J. Bonifácio
Bial–Departamento de Investigação e Desenvolvimento, 4745-457 São Mamede do Coronado, Portugal

DNA	Deoxyribonucleic acid
gDNA	Genomic DNA
HPLC	High-peformance liquid chromatography
LLOQ	Lower limit of quantification
MBCOMT	Membrane-bound catechol-*O*-methyltransferase
MP	Membrane protein
Mut	Methanol utilization
Mut^S	Methanol utilization slow
Mut^+	Methanol utilization plus
OD_{600}	Optical density 600 nm
PCR	Polymerase chain reaction
P. pastoris	*Pichia pastoris*
RID	Refractive index detection
SAM	*S*-adenosyl-*L*-methionine
SCOMT	Soluble catechol-*O*-methyltransferase
SDS-PAGE	Reducing sodium dodecyl sulfate polyacrylamide

Introduction

Catechol-*O*-methyltransferase (COMT, EC 2.1.1.6) is a magnesium-dependent enzyme that catalyzes the methylation of catechol substrates using *S*-adenosyl-*L*-methionine (SAM) as a methyl donor and yielding, as reaction products, the *O*-methylated catechol and S-adenolsyl-*L*-homocysteine [1]. In humans, COMT appears as two molecular forms, a soluble isoform (SCOMT) that is found mainly in the cytoplasm and a membrane-bound (MBCOMT) isoform that is an integral membrane protein with the catalytic portion of the enzyme oriented toward the cytoplasmic side of the membrane [1]. In general, the major physiological role of COMT is the elimination of biologically active or toxic catechols [1]. Moreover, since the affinity of MBCOMT for catecholamines is 10- to 100-fold higher than that of SCOMT [1], MBCOMT is believed to be more important for the metabolism of catecholamines in vivo [2]. Furthermore, COMT has been implicated in several human diseases such as pre-eclampsia [3], pain [4], estrogen-induced cancers [5], cardiovascular diseases [6], and a number of psychiatric conditions [7]. However, the best documented is the important role that COMT plays in Parkinson's disease whose most effective treatment remains the dopamine replacement therapy with levodopa together with an inhibitor of aromatic amino acid descarboxylase and a COMT inhibitor [1]. Recombinant MBCOMT expression has been reached using prokaryotic hosts such as *Escherichia coli* (*E. coli*) SG13009 [8, 9] and BL21 [10]. More recently, a flow sheet for recombinant MBCOMT expression [11] and purification using either hydrophobic matrices [12] or Q-sepharose as an anion exchanger [13] was developed using *Brevibacillus choshinensis* as the host. Beyond these systems, eukaryotic hosts such as Sf9 insect cells [14], transfected human embryonic kidney fibroblast cell lines [15], human HeLa, and hamster BHK cells [16] have also been employed for recombinant MBCOMT expression. Nevertheless, neither of the aforementioned systems has proven to be capable of producing crystallization-grade human MBCOMT protein for structural studies. On the other hand, yeasts are an attractive eukaryotic host for the production of heterologous proteins owing to their fast growth rates coupled with high cell densities and the well-developed variety of genetic tools [17]. Initially described by Philips Petroleum Company for the production of single cell protein, *Pichia pastoris* (*P. pastoris*) has demonstrated its most powerful success as a large-scale (fermentation) recombinant production tool [18]. Moreover, *P. pastoris*, as a

yeast, is a single-cell microorganism that is easy to manipulate and cultivate and as an eukaryote, is capable of many posttranslational modifications performed by higher eukaryotic cells such as proteolytic processing, folding, disulfide bond formation, and glycosylation [18]. In fact, this system stands out as a successful expression host for the production of a number of functional, properly folded, and inserted membrane proteins (MP) that have not been produced successfully in other systems [19, 20]. Among *P. pastoris*, more remarkable features are the promoter derived from the alcohol oxidase I (AOX 1) gene strictly regulated and induced by methanol that is uniquely suited for the controlled expression of foreign genes [18]. Moreover, *P. pastoris* has two alcohol oxidase genes, AOX 1 and AOX 2, of which AOX 1 is much more strongly transcribed than AOX 2 [21]. Actually, methanol-induced *P. pastoris* cultures grown in shake flask cultures result in expression levels of approximately 5 % of total soluble proteins [22]. However, this level is increased to more than 30 % of total proteins in cells grown in bioreactor cultures applying growth-limiting rates of methanol [22]. As a matter of fact, this can be explained by the preferential use of the *P. pastoris* system for respiratory growth, which makes easier its cultivation at high-cell densities relative to fermentation yeasts such as *Saccharomyces cerevisiae* [23]. The methanol levels in the culture medium must be carefully controlled since *P. pastoris* strains are sensitive to high residual methanol concentrations and sudden accumulation of methanol may result in loss of AOX activity and cell death [24] while low levels of methanol may not be enough to initiate the AOX transcription [25]. Also, keeping constant the methanol concentration during the induction phase has positive effects on the production of foreign proteins [26]. Therefore, several strategies have been employed to monitor the methanol concentrations in *P. pastoris* bioprocesses that can be divided into two classes, the off-line and on-line techniques [25]. The off-line techniques are based on gas chromatography [27], high-performance liquid chromatography (HPLC) [28], or enzymatic reactions [29] and require a pretreatment of the sample while some on-line techniques are based on automatization of the most typical off-line techniques in order to achieve an autonomous system [25]. Along with the wild-type methanol utilization plus (Mut^+) *P. pastoris* strains (e.g., X33 or GS115) that contain two functional AOX genes, strains with a methanol utilization slow (Mut^S) phenotype can be generated. In what concerns to Mut^S strains, due to their slow utilization of methanol, a mixed feed combining glycerol and methanol is commonly employed in the fermentation induction phase where glycerol acts as an efficient substrate for cell growth and target protein production while methanol works as an inducer [30]. On the other hand, for Mut^+ strains, due to efficient methanol utilization, a typical fed-batch fermentation strategy is to feed methanol alone as inducer and energy source in the induction phase [30]. Furthermore, the application of glycerol mixed feeding strategies to Mut^+ strains similar to those applied for Mut^S strains yielded different results since in some cases, the expression levels of the target protein are increased [31] while in other studies, strong inhibition of the protein biosynthesis is observed [32]. Moreover, sorbitol is a nonrepressing carbon source for AOX1 promoter where its accumulation during the induction phase does not seem to affect the expression levels of recombinant proteins [33]. Indeed, mixed feeds of sorbitol/methanol have also been applied successfully for both Mut^+ and Mut^S *P. pastoris* strains [33]. Here, we present a new process for hMBCOMT biosynthesis and recovery from *P. pastoris* cells. In fact, recombinant hMBCOMT production was accomplished using the pPICZα vector in *P. pastoris* X-33 and KM71H cells where after employing a simple lysis protocol through glass beads vortexing, biologicallly active hMBCOMT preferentially accumulates in the membrane fraction. Also, a method employing an HPLC coupled to a refractive index detection (RID) for the simultaneous detection and quantification of methanol, glycerol, and sorbitol in *Pichia* culture broths was successfully implemented and validated. Based on this method, the methanol consumption was measured over the fermentation time as well as

the glycerol and sorbitol levels in mixed feed experiments. Finally, a full kinetic characterization to epinephrine in the two major fractions containing biologically active COMT was performed. In general, for the first time, the ability of a Mut^+ and a Mut^S *P. pastoris* strains for the production of recombinant human MBCOMT, measured as hMBCOMT-specific activity (nmol/h/mg of protein), was thoroughly investigated under different culture conditions.

Materials and Methods

Materials

Ultrapure reagent-grade water was obtained with a Mili-Q system (Milipore/Waters). Yeast extract, lysozyme, dithiotreitol, SAM, epinephrine (bitartrate salt) were obtained from Sigma Chemical Co. (St. Louis, MO). *P. pastoris* strains X-33 and KM71H as well as Zeocin were purchased from Invitrogen (Carlsbad, CA). The Nzycolour protein marker II used for estimation of subunit masses was bought from Nzytech (Lisbon, Portugal). Polyclonal rabbit anti-COMT antibody affinity purified was produced in BIAL using purified recombinant rat COMT. All other chemicals were of analytical grade commercially available and used without further purification.

Strains, Plasmids, and Media

E. coli TOP10F′ was used for deoxyribonucleic acid (DNA) manipulations. *E. coli* transformants were selected on low-salt Luria-Bertani plates with 25 μg/mL Zeocin. *P. pastoris* X-33 and KM71H were used for fusion gene expression. The following media supplemented with Zeocin 200 μg/mL were employed in *Pichia* cell fermentations: YPD medium (1 % yeast extract, 2 % peptone, and 2 % glucose), YPDS medium (YPD medium supplemented with 1 M sorbitol), BMGH (100 mM potassium phosphate buffer [pH 6.0], 1.34 % yeast nitrogen base, 4 $*10^{-4}$ g/L biotin, and 1 % glycerol), and BMMH (100 mM potassium phosphate buffer [pH 6.0], 1.34 % yeast nitrogen base, 4 $*10^{-4}$ g/L biotin, and 0.5 % methanol). *P. pastoris* transformants were selected on YPDS plates with 200 μg/mL Zeocin.

Construction of the Expression Vector pPICZα-hMBCOMT

Easy select expression kit for expression of recombinant proteins using pPICZα in *P. pastoris* (Invitrogen, Carlsbad, CA) was used for the expression of human MBCOMT in its native form, and the process was carried out according to manufacturer's instructions. Briefly, the DNA fragment coding for MBCOMT was obtained from the pNCMO2_MBCOMT expression vector [11] previously constructed by our research group by polymerase chain reaction (PCR) using specific primers for cloning (forward primer; 5′ AACTCGAGAAAAGAATGC CGGAGGCCCCGCCT 3′; reverse primer, 5′ AACTCGAGTCAGGGCCCTGCTTCGCTGC CTG 3′). PCR was conducted as follows: denaturation at 95 °C for 5 min, followed by 30 cycles at 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min, and a final elongation step at 72 °C for 5 min. The amplified DNA was purified by low melting agarose gel electrophoresis, digested with Xho I, and cloned into the vector pPICZα (previously digested with Xho I) by T4 DNA ligase. This construct was transformed into *E. coli* TOP10F′ cells, grown overnight at 37 °C in plates with low-salt Luria-Bertani agar medium containing zeocin (25 μg/mL), and

colonies were screened for the presence of the construct pPICZα-hMBCOMT. Therefore, some colonies were inoculated in 2.0 mL of low-salt Luria-Bertani medium and grown at 37 °C and 250 rpm overnight. From these cultures, highly purified plasmids were prepared using Wizard SV Plus SV Minipreps and were then subjected to DNA sequence analysis to confirm the identity of the amplicon, orientation, and frame. This was confirmed to correspond to human MBCOMT gene [34].

Pichia pastoris Transformation and Selection for Positive Clones

P. pastoris competent cells were prepared according to the manufacturer's instructions (Invitrogen, Carlsbad, CA) and used freshly. Briefly, the plasmid pPICZα-hMBCOMT was linearized with Sac I and introduced into freshly made *P. pastoris* X-33 and KM71H cells by electroporation (2.5 kV, 25 μF, 1000 Ω) using an Eppendorf 2510. Immediately after pulsing, cells were suspended in 1-mL ice-cold 1 M sorbitol, plated in YPDS containing zeocin at a final concentration 200 μg/mL, and grown for 4 days at 30 °C. Subsequently, high-level expression transformants were screened from these plates. In what concerns to the transformants obtained from X33 cells, these were tested in order to confirm the methanol utilization (mut) phenotype according to the manufacturer's instructions (Invitrogen, Carlsbad, CA). After confirming that the X33 integrants presented a Mut^+ phenotype, genomic DNA (gDNA) was extracted from each colony of X33 and KM71H transformants using the Wizard SV Genomic DNA Purification System supplemented with zymolyase. Finally, the stable occurrence of the expression cassette was verified in the colonies gDNA by PCR using AOX1 promoter and terminator-specific primers (AOX1 5′ GACTGGTTCCAATTGACAAGC 3′ and AOX1 5′ CAAATGGCATTCTGACATCC).

Yeast Fermentation

Unless otherwise stated, recombinant hMBCOMT was carried out according to the following protocol. Cells containing the expression construct were grown at 30 °C in YPD plates. A single colony was inoculated in 50.0 mL of BMGH medium in 250-mL shake flasks. Cells were grown at 30 °C and 250 rpm overnight when the cell density at 600 nm (OD_{600}) typically reached 6.0. Subsequently, since the inoculation volume was fixed to achieve an initial OD_{600} of 1.0, an aliquot of the fermentation in the medium BMGH was collected and centrifuged at room temperature during 5 min. After centrifuging the cells and ensure that all glycerol was removed, the cells were resuspended in the induction medium and added to 500.0-mL shake flasks to a total volume of 100.0 mL.

The fermentations were carried out during 120 h at 30 °C and 250 rpm and were supplemented with methanol at a final concentration of 1 % every 24 h. Finally, the cells were harvested by centrifugation (1500×*g*, 10 min, 4 °C) and stored frozen at −20.0 °C until use.

Protein Recovery and Subcellular Fractionation

Cell suspensions were lysed in lysis buffer (150 mM NaCl, 10 mM DTT, 50 mM Tris, 1 mM $MgCl_2$, pH 8.0) and freshly made protease inhibitors (1 mM PMSF, 5.0 μg/mL leupeptin, and 0.7 μg/mL pepstatin A) containing an equal volume of acid-washed glass beads (500 μm, Sigma). The mixture was vortexed 7 times for 1 min with an interval of 1 min on ice and centrifuged at 500 g (4 °C) for 5 min to remove cell debris and glass beads. Subsequently, the supernatant (defined as S500g) was collected, DNase (Sigma) (1 mg/mL) added, and it was

centrifuged at 16000*g* (4 °C) for 30 min. The 16,000*g* supernatant (defined as S16000g) was collected, and the 16,000*g* pellet (defined as P16000g) was solubilized in solubilization buffer (lysis buffer plus 1 % Triton X-100) at 4 °C until its full solubilization. The cellular distribution of hMBCOMT in *P. pastoris* X-33 and KM71H cells was determined by monitoring the target enzyme activity in the several fractions. Finally, beyond the glass beads method, the freeze-thaw lysis through six consecutive freeze (in liquid nitrogen at −120 °C)/thaw (at 42 °C) cycles as well as the sonication in ice (100 W, 90 % amplitude with 0.5-s interval) were also applied in order to assess the most suitable *P. pastoris* lysis method.

SDS-PAGE and Western Blot Analysis

Reducing sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel electrophoresis was performed according to the method of Laemmli [35]. Moreover, the SDS-PAGE and Western blot analysis were carried out as previously described [36]. Finally, the protein content in samples was measured by the Pierce BCA Protein Assay Kit (Thermo Scientific, USA), using bovine serum albumin as the standards (0.025–2.0 mg/mL), according to manufacturer's instructions.

HPLC Analytical Methods

Methanol, Glycerol, and Sorbitol Determinations

The chromatographic analysis was performed using an HPLC model Agilent 1260 (Agilent, Santa Clara, CA, USA) equipped with an autosampler and quaternary pump coupled to a 1260 Infinity Refractive Index Detector (Agilent, Santa Clara, CA, USA). The chromatographic separation was achieved on a cation-exchange analytical column Agilent Hi-Plex H (300× 7.7 mm i. d.; 8 μm) [28] acquired from Specanalítica (Lisbon, Portugal). The analysis was performed at 65 °C with a flow rate of 0.6 mL/min using isocratic elution with 0.005 M H_2SO_4 [28]. The mobile phase was filtered prior to the analysis under vacuum using a 0.2-μm pore nylon membrane and degassed for 15 min in an ultrasonic bath. The samples were centrifuged at 6000 rpm for 10 min, and the supernatant was filtered through a 0.22-μm cellulose-acetate filter.

MBCOMT Enzymatic Assay

The methylating efficiency of recombinant MBCOMT was evaluated by measuring the amount of metanephrine formed from epinephrine as previously described [37]. To determine the recombinant MBCOMT kinetics parameters K_M and V_{max}, aliquots of either the S500g supernatant or the P16000g solubilized membrane (previously optimized 1.0 mg of total protein per mL) obtained after glass beads lysis were added to increasing concentrations of epinephrine (20–400 μM) (maintaining SAMe concentration on 250 μM) for 15 min at 37 °C [11]. The reaction was stopped with 2 M of percloric acid and before the HPLC with coulometric detection; the samples were processed as described elsewhere [37]. The chromatographic analysis was peformed using an HPLC model Agilent 1260 system (Agilent, Santa Clara, CA, USA) and as previously described [13]. Both presented chromatographic systems were controlled by Chemstation software supplied by Agilent Technologies (Waldbronn, Germany). All data analysis was performed using Prism 6 (GraphPad Sofware Inc., San Diego, CA).

Results and Discussion

Recombinant hMBCOMT has been produced using either prokaryotic or eukaryotic hosts. Since hMBCOMT exists in cells at low concentrations and neither of the aforementioned systems has had complete success, the development of a system capable of deliver hMBCOMT in quantity and quality enough to proceed further structural studies remains an essential hurdle to overcome. The methylotrophic yeast *P. pastoris* has been used as an industrial host for recombinant protein and metabolite production, showing a powerful capacity to meet the required biomolecular target production levels in high-throughput assays for functional genomics and drug screening [19]. As a matter of fact, proteins with all kinds of membrane spanning topologies, including enzymes [38], aquaporins [39], and ion channels [40], have been successfully expressed using *P. pastoris* [20]. Therefore, in this work and for the first time, we successfully applied a Mut^+ (X33) and a Mut^S (KM71H) *P. pastoris* strains for recombinant hMBCOMT biosynthesis. Moreover, the ability of these two strains to produce hMBCOMT correctly inserted in the membranes and, consequently, in a properly folded state was evaluated, as judged by the kinetic parameter K_M assessments in the different fractions obtained during the bioprocess. Finally, aspects such as the relevance of the subcellular fractionation to monitor hMBCOMT preferential acumulation in *Pichia* cells, time course experiments in order to evaluate the growth profile of X33 and KM71H cells, as well as the methanol, glycerol, and sorbitol consumption during mixed feed fermentations are also addressed in this work.

pPICZα-hMBCOMT Expression Vector Construction

In this study, the hMBCOMT gene was PCR amplified from the plasmid pNCMO2-hMBCOMT [11] using specific primers containing a unique restriction site for Xho I and cloned into the plasmid pPICZα, previously digested with the same restriction enzyme. After confirming that the isolated positive clone contained the hMBCOMT gene [34], 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 were extracted and analyzed by PCR using the AOX primers. For both *P. pastoris* strains in study, the PCR analysis in the transformed gDNA yielded one band corresponding to the size of our gene of interest cloned into pPICZα (data not shown). Finally, the Mut phenotype of the *P. pastoris* X33 host strain (there is no need to screen the KM71H cells since all present a Mut^S phenotype) transformed with the hMBCOMT cDNA was determined according to the manufacturer's instructions, and it was confirmed to be Mut^+.

Bioprocess Monitoring: Time Course hMBCOMT-Specific Activity, *P. pastoris* Optical Density, and Methanol Concentrations

In *P. pastoris* bioprocesses that include the AOX promoter, the most commonly used carbon sources are methanol, glycerol, sorbitol, and mannitol in which the methanol is used not only as carbon source but also as an inducer [41]. Actually, the methanol concentration in culture broths is one of the most important parameters in these bioprocesses and, therefore, needs to be carefully monitored. As a matter of fact, as methanol is used as inducer, low levels may not be enough to initiate the AOX transcription while high levels can be toxic to the cells, making it important to keep constant the methanol concentration during the induction phase [25]. Similarly, it is also important to monitor the glycerol and sorbitol levels in *Pichia* culture broths during mixed feed fermentations since for example, an excess of glycerol represses the AOX 1 promoter, which results in lower productivity [41]. The use of mixed feeds in

P. pastoris fermentations with another carbon source in addition to methanol is commonly applied to increase the cell concentration and productivity, as well as to decrease the induction time [41]. In fact, mixed feeds employing glycerol have been applied for recombinant protein production in *P. pastoris* Mut^+ and Mut^S [41] while sorbitol has also been used as cosubstrate in Mut^+ [33] and Mut^S [42] *P. pastoris* bioprocesses.

Hence, in this work, we implemented and validated an analytical method for simultaneous detection and quantification of methanol, glycerol, and sorbitol using HPLC-RID in *P. pastoris* fermentation broths. The method was fully validated according to internationally accepted guidelines from the Food and Drug Administration [43]. As shown in Fig. 1, the HPLC chromatogram of a fermentation extract sample spiked with sorbitol, glycerol, and methanol demonstrates that the method provides an excellent resolution and selectivity between the compounds of interest, allowing the samples to be directly injected without pretreatment. In addition, components of the fermentation broth were injected as a "blank" analysis in order to evaluate possible interferents at the respective retention times of the analytes; however, none was observed (data not shown). Linearity was established between 0.01 and 4 % (*v*/*v*) for glycerol, between 0.01 and 5 % (*v*/*v*) for methanol and between 0.00625 and 0.2 M for sorbitol, with determination coefficients (R^2) higher than 0.99 for all compounds. The lower limits of quantification (LLOQ) were 0.01 % (*v*/*v*) for glycerol and methanol and 0.00625 M for sorbitol. Intraday and interday precision ranged from 1 to 10 %, while accuracy was within a ±15 % interval for all compounds. As a matter of fact, this analytical method has proven to be of great interest in *P. pastoris* bioprocesses because not only it allows the screening of the most appropriated methanol concentrations as it also allows to monitor glycerol or sorbitol concentrations, usually employed in mixed feed *P. pastoris* fermentations. In order to evaluate the methanol consumption during the fermentation stage, preliminary assays were conducted for both strains in BMMH (methanol 0.5 %) medium, supplemented each 24 h with pulses of 1 % methanol. As we can see in Fig. 2a, for both strains, even with the increase in the optical density at OD_{600} observed for the X33 strain at 96 and 120 h, the methanol concentration increases from 24 (after the beginning of the 1 % methanol pulses) to 120 h of fermentation. Therefore, to avoid a possible inhibition in the growth of both strains due to the accumulation of the methanol in the culture broth, we decided to lower both the initial methanol concentration as well as the methanol levels used in the pulses to 0.25 % (*v*/*v*). When the methanol concentration is lowered to 0.25 % (see Fig. 2b), higher OD_{600} values were obtained for the X33 strain, and the methanol is almost all consumed. On the other hand, the profile obtained for the KM71H strain in these

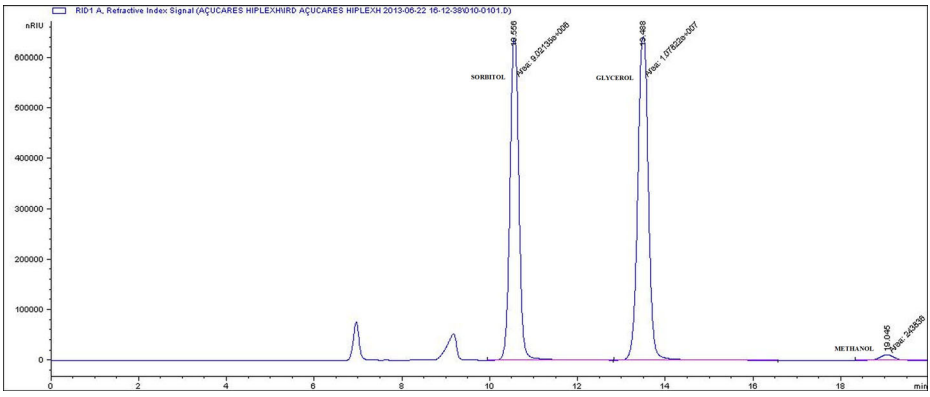


Fig. 1 Chromatogram of a spiked fermentation broth extract sample (retention time of sorbitol, glycerol, and methanol is 10.5, 13.5, and 18.9 min, respectively)

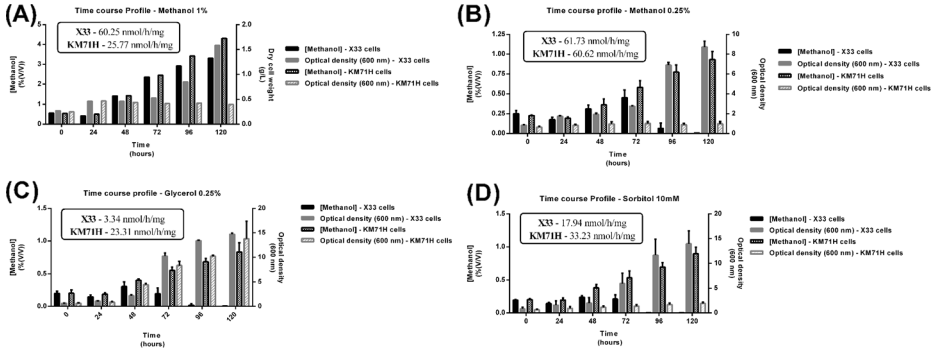


Fig. 2 Assessment of the growth profile of *Pichia pastoris* X33 and KM71H harboring the plasmid pICZα-MBCOMT as well as the methanol consumption at different incubation periods and different medium formulations, keeping constant the culture conditions (30 °C and 250 rpm). The hMBCOMT biological activity levels (nmol/h/mg of protein) at the end of each fermentation are depicted in each graph. **a** *Pichia pastoris* X33 and KM71H in BMGH medium (methanol 0.5 %) supplemented with 1 % methanol every 24 h; **b** *Pichia pastoris* X33 and KM71H in BMGH medium (methanol 0.25 %) supplemented with 0.25 % methanol every 24 h; **c** *Pichia pastoris* X33 and KM71H in BMGH medium (methanol 0.25 %) supplemented with 0.25 % methanol and 0.25 % glycerol every 24 h; **d** *Pichia pastoris* X33 and KM71H in BMGH medium (methanol 0.25 %) supplemented with 0.25 % methanol and 10 mM sorbitol every 24 h; samples were taken before each pulse. The values represent the mean of three independent samples. *Vertical lines* show SD

conditions is similar to those obtained with a concentration of methanol of 1 % (*v*/*v*). As mentioned before, the main difference between the *P. pastoris* X33 and the KM71H strains is the fact that the KM71H presents a Mut^S phenotype caused by the loss of AOX activity encoded by the AOX1 gene. Therefore, this strain only codes the enzyme AOX 2 and, consequently, presents a slow growth phenotype in methanol medium [44]. This fact helps to explain the huge differences between the two strains since the strain KM71H consumes less methanol, and therefore, the OD_{600} values are much lower when compared with the X33 strain. Following the evaluation of the growth profile of the X33 and KM71H strains and maintaining the methanol concentration in 0.25 % (*v*/*v*), we also evaluated the effect of the addition of glycerol (0.25 % (*v*/*v*)) or sorbitol (10 mM) in the pulses along with methanol 0.25 %. Therefore, assays (shown in Fig. 2c) were carried out with both strains in which the main difference between those described in Fig. 2b was that besides the addition of methanol 0.25 % every 24 h, we also added glycerol at a final concentration of 0.25 %. In these conditions, higher cell growth was achieved for both strains, especially for the KM71H strain where an OD_{600} of near 14 was obtained, compared to the 0.98 obtained when 0.25 % (*v*/*v*) methanol was fed alone. Along with the methanol, the glycerol levels were also quantified, and the following values in % (*v*/*v*) were obtained: 0.109± 0.024 at 48 h and 0 at 72, 96, and 120 h of fermentation for the X33 strain while for the KM71H strain, all the glycerol was consumed. Additionally, Fig. 2b, c depicted the hMBCOMT biological activity levels obtained at 120 h of fermentation for each condition and strain where it is possible to observe that lower levels for hMBCOMT biological activity were detected for both strains in the mixed feeds of methanol/glycerol when compared with fermentations where methanol was fed alone. This can be explained by the fact that the glycerol can be causing some repression on the AOX promoter, probably due to the conversion of glycerol in ethanol in the induction phase [44]. Another strategy concerning the use of mixed feeds of methanol [0.25 % (*v*/*v*)] and sorbitol (10 mM) (see Fig. 2d) was also employed for both strains in study. In this case, the sorbitol levels (in M) at 48, 72, 96, and 120 h of fermentation for the X33 strain were, respectively, 0.00910±0.00061, 0.01650±0.00370, 0.01550±0.00434, and 0.00797±0.00030 while for the KM71H at the same periods of fermentation were 0.00961±0.00160, 0.01959±

0.00230, 0.02910±0.00400, and 0.04100±0.00370. Despite that the sorbitol quantification in the culture broths showed that higher quantities of sorbitol accumulated in the culture medium, in contrast with glycerol, sorbitol accumulation during the induction phase does not affect the expression level of the target protein [45]. However, as shown in Fig. 2d, lower MBCOMT biological activity levels were determined in this strategy when compared where methanol was fed alone (see Fig. 2b). Indeed, sorbitol has been described as a nonrepressing carbon source but presents the disadvantage that the maximum specific growth rate is too low and since the fermentation time was maintained in 120 h, probably an induction time above 120 h will improve the overall yield of the mixed feed employing sorbitol as a cosubstrate.

hMBCOMT Recovery

The recovery of intracellular proteins requires the disruption of the *P. pastoris* rigid cell wall, which is composed of multiple layers of cross-linked β-1,3-glucan, chitin, and glycosilated mannoproteins [46]. Actually, common methods usually applied to *P. pastoris* lysis are the microfluidizers [47], a french press [48] or glass beads [49]. In this work, a sequential procedure was established as the main method for *Pichia* strains lysis: vortexing of the cells (combined with lysis buffer and glass beads) for 7 times during 1 min with an interval of 1 min on ice. In an attempt to improve the yield recovery (measured as the total protein released) of this method as well as the quantity of the target protein (assessed as the hMBCOMT activity) obtained after lysis, the freeze-thaw lysis (six consecutive freeze (liquid nitrogen)/thaw (42 °C) cycles) and sonication (short bursts of acoustic waves with a power input up to 100 W until the cell suspension reached a limit temperature of 12 °C) were also evaluated. Therefore, these methods were applied for *P. pastoris* X33 lysis alone or associated, and the efficiency of each method was evaluated in the several fractions obtained during the bioprocess and measured as the hMBCOMT biological activity (nmol/h/mg of protein), whose results are shown in Table 1. In general, only through the application of glass beads was possible to disrupt the *P. pastoris* rigid cell wall since higher enzyme activities were detected, in comparison with the other methods employed. In addition, the combination of this method with the freeze-thaw lysis or the sonication did not improve the yield of the recovery step. Therefore, from the methods applied, the use of glass beads proved to be the most effective method in the disruption of the *P. pastoris* cell wall, although it is moderately effective and reproducible. Nevertheless, it is performed very quickly following a simple protocol, what can be advantageous for maintaining stability and, consequently, the biological activity of the target enzyme.

The subcellular distribution enables stepwise separation and extraction of cytoplasmic and membrane-bound proteins in *P. pastoris* cells, allowing a protein localization assessment and enzyme enrichment from specific cellular compartments. Initial Western blot trials conducted for *P. pastoris* X33 cells revealed that immunologically active hMBCOMT correctly inserted in the membranes in a monomeric and properly folded form was detected in the 16,000*g* pellet membrane fraction (see Fig. 3, lane 2). On the other hand, in the 16,000*g* supernatant (see Fig. 3, lane 1), immunologically active hMBCOMT was detected as dimmers, as judging by the comparison with the molecular weight marker. After the preliminary results obtained by the Western blot analysis, through the determination of hMBCOMT biological activity in the different fractions, it is possible to evaluate in which fraction biologically active hMBCOMT preferentially accumulates. So, after two consecutive centrifugations at 500*g* and 16,000*g*, the hMBCOMT biological activity was determined in each fraction for each strain, and the results are depicted in Fig. 4. Indeed, after performing the cell lysis and the first centrifugation step at 500*g*, biologically active hMBCOMT was recovered in the supernatant (see Fig. 4, supernatant S500*g*) for both strains. Following this first centrifugation, the supernatant was recovered, and

Table 1 Efficiency of different disruption methods including freeze-thaw, sonication, and glass beads for *P. pastoris* X33 lysis

Lysis method	Fraction	Total protein (mg)	Protein recovery (%)	Specific activity (nmol/h/mg)
Glass beads	S500g	19.4	100	60.25
	S16000g	10.7	55.2	65.8
	P16000g	3.04	15.7	128.2
Glass beads+freeze-thaw	S500g	29.2	100	8.25
	S16000g	19.9	68.2	1.91
	P16000g	7.73	26.5	0
Freeze-thaw+glass beads	S500g	33	100	6.1
	S16000g	15.2	46.1	1.1
	P16000g	7.4	22.4	11.3
Freeze-thaw	S500g	12.2	100	3,2
	S16000g	9.7	79,5	1,1
	P16000g	2.4	19,7	13,5
Sonication	S500g	———	———	0
	S16000g	———	———	0
	P16000g	———	———	0
Sonication+freeze-thaw	S500g	———	———	0
	S16000g	———	———	0
	P16000g	———	———	0
Freeze-thaw+sonication	S500g	68.6	100	22.86
	S16000g	39.7	57.9	18.41
	P16000g	19.8	28.9	1
Sonication+glass beads	S500g	66.6	100	14.2
	S16000g	40.5	60.8	4
	P16000g	9.4	12	0.5
Glass beads+sonication	S500g	67.83	100	3.8
	S16000g	———	———	0
	P16000g	4.5	6.6	5.74

a second centrifugation at 16,000*g* was performed, enabling an additional fractionation of the samples. As a matter of fact and as depicted in Fig. 4, biologically active hMBCOMT was mostly detected in the 16,000*g* membrane pellet solubilized in solubilization buffer, a fraction where MBCOMT seems to appear as a monomer in a properly folded form, correctly inserted in the membranes, as previously shown in the Western blot analysis (see Fig. 4, lane II). In general, contrary to other findings by Orman and coworkers for the human growth hormone [41] and Pla and collaborators for the A33 scFv antibody fragment [50], the *P. pastoris* strain X33 seems to be most suitable for hMBCOMT expression when compared with the KM71H strain as it can be seen by the assessment of the target enzyme activity in the several fractions obtained.

hMBCOMT Kinetic Characterization

The acquired stability after *P. pastoris* lysis stage allows an intrinsic kinetic characterization of hMBCOMT. Therefore, the kinetic parameters K_M and V_{Max} of the *O*-methylation reaction of epinephrine were determined to evaluate the functional properties of the recombinant protein

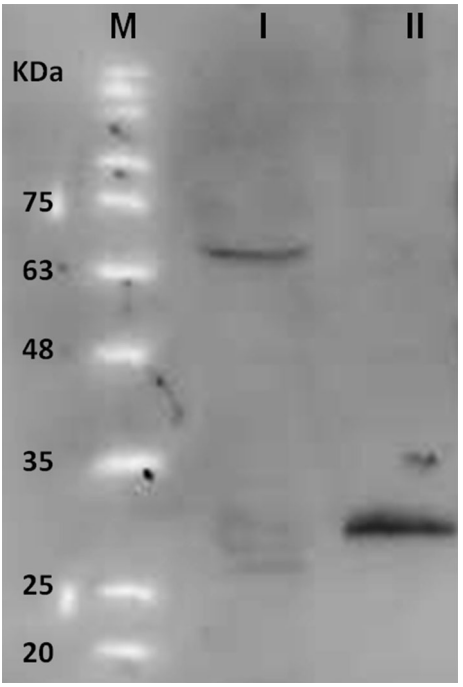


Fig. 3 Western blot analysis of recombinant hMBCOMT in the two major fractions obtained during a *Pichia pastoris* X33 bioprocess production, the 16,000*g* supernatant (*lane I*) and the 16,000*g* pellet solubilized in a specific solubilization buffer (*lane II*) (*M* molecular weight marker)

produced in this work using both *P. pastoris* strains (X33 and KM71H) as well as from the major fractions obtained during this bioprocess (500*g* supernatant and 16,000*g* pellet solubilized without detergent). As expected, the incubation of the enzyme preparations with

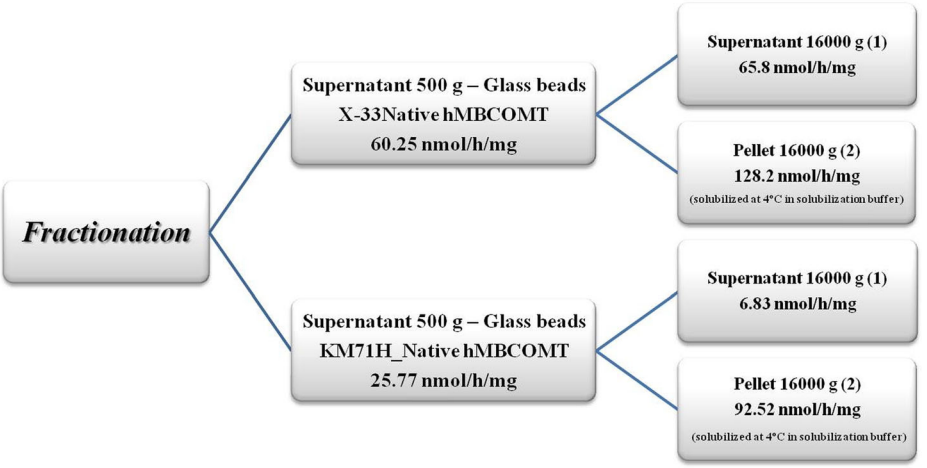


Fig. 4 Diagram showing the subcellular distribution (after two consecutive centrifugations at 500*g* and 16,000*g*) of hMBCOMT for both X33 and KM71H *P. pastoris* strains measured through the determination of enzyme-specific activity (nmol/h/mg of protein) in each fraction

increasing concentrations of epinephrine resulted in a concentration-dependent formation of metanephrine. In general, the affinity values obtained in this work in the different fractions are in agreement with previous results reported in the literature, as we can see in Table 2. Moreover, lower affinity values were obtained in the solubilized membrane fractions (defined as P16000*g* in Table 2), which indicates that MBCOMT inserted in the membranes is in a more properly folded form, thus presenting more affinity to epinephrine, as expected. Since it is well documented that the detergent influences the determination of COMT K_M values [51], the kinetic characterization in the solubilized 16,000*g* pellet fraction was performed without the addition of detergent unlike the determination of the enzyme activity in the subcellular distribution. Therefore, the V_{Max} values obtained in this fraction are lower than those reported in Fig. 4 since the detergents play an important role in the extraction of MP from cellular membranes [52], thus improving biologically active hMBCOMT extraction.

Conclusions

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. In fact, while structural and functional studies of hMBCOMT are still hampered by its low recombinant production, a new process intended to synthesize recombinant human MBCOMT using a eukaryotic host was sucessfully developed in this work. The implemented HPLC-RID procedure to monitor the carbon source in *Pichia* broths proved to be of great interest since it allows the simultaneous quantification of methanol, glycerol, and sorbitol in a single run. In general, the implementation of mixed feed strategies of methanol and glycerol or sorbitol led to a decrease in hMBCOMT biological activity, probably because the AOX promoter is repressed under the conditions tested. On the other hand, since the maximum specific growth rate in sorbitol mixed feeds is too low, probably the extension of the *P. pastoris* fermentation time above 120 h may possibly increase the overall productivity of this strategy. In addition, a higher decrease in the hMBCOMT biological activity levels in the mixed feed fermentations was observed for the X33 strain when compared with the fermentations where methanol was fed alone because the AOX1 gene is more likely to be repressed in these conditions than the AOX2, the only functional AOX gene in KM71H strain. Regarding the subcellular fractionation, biologically and immunologically active MBCOMT correctly inserted in the membranes was

Table 2 Affinity values (kinetic parameters K_M and V_{Max} for epinephrine) of the recombinant hMBCOMT produced in this work from different protein extracts obtained during the bioprocess and comparison with previous results reported in the literature

Kinetic Parameter	(Bai and collaborators) [10]	(Pedro and coworkers) [11]	This work - S500g		This work - P16000g	
			X33 cells	KM71H cells	X33 cells	KM71H cells
K_M (μM)	28.8	50.24	55.33±2.347	28.51±2.812	34.05±1.264	26.53±2.303
V_{Max} (nmol/h/mg of protein)	2778	48.07	66.08±7.497	54.14±1.493	46.56±0.4883	56.68±1.533

COMT activity is shown as the rate of formation of metanephrine (nmol/h/mg of protein) vs concentration of epinephrine (μM)

mainly recovered in a membrane fraction after a centrifugation step at 16,000*g*. Finally, the determination of the affinity values of the hMBCOMT produced in this work allows us to conclude that these recombinant forms are kinetically identical to its correspondent native enzyme. Overall, the influence of the *P. pastoris* strain phenotype to synthesize human MBCOMT was analyzed, and it was found that strain Mut^+ X33 gave better results than Mut^S KM71H. Indeed, this system seems to be a promising approach to deliver MBCOMT at high productivities for several pharmacological trials. Furthermore, the developed tools implemented in this work to monitor *P. pastoris* experiments can also play an important role in a further scale-up of this bioprocess.

Acknowledgments A.Q. Pedro acknowledges a doctoral fellowship (SFRH/BD/81222/2011) from Fundação para a Ciência e Tecnologia and D. Oppolzer acknowledges a fellowship (CENTRO-07-ST24_FEDER-002014-TPCR-2-004) from Programa "Mais Centro" within the scope of QREN–POPH–Advanced Formation programs cofunded by Fundo Social Europeu and MEC. This work was partially funded by Fundação para a Ciência e Tecnologia I.P. (PIDDAC) and Fundo Europeu de Desenvolvimento Regional-FEDER funds through Programa Operacional Factores de Competitividade (POFC)–COMPETE: FCOMP-01-0124-FEDER-027563 and by National Funds through FCT–Fundação para a Ciência e Tecnologia within the scope of Project "EXPL/BBB478/BQB/0960/2012."

Conflict of Interest All authors declare they do not have any conflict of interest.

Compliance with Ethical Standards In this work, no studies involving human participants or animals were carried out.

References

1. Bonifacio, M. J., Palma, P. N., Almeida, L., & Soares-da-Silva, P. (2007). Catechol-*O*-methyltransferase and its inhibitors in Parkinson's disease. *CNS Drug Reviews, 13*(3), 352–379.
2. Reenila, I., & Mannisto, P. T. (2001). Catecholamine metabolism in the brain by membrane-bound and soluble catechol-*O*-methyltransferase (COMT) estimated by enzyme kinetic values. *Medical Hypotheses, 57*(5), 628–632.
3. Kanasaki, K., Palmsten, K., Sugimoto, H., Ahmad, S., Hamano, Y., Xie, L., Parry, S., Augustin, H. G., Gattone, V. H., Folkman, J., Strauss, J. F., & Kalluri, R. (2008). Deficiency in catechol-*O*-methyltransferase and 2-methoxyoestradiol is associated with pre-eclampsia. *Nature, 453*(7198), 1117–1121.
4. Papaleo, F. J., Crawley, N., Song, J., Lipska, B. K., Pickel, J., Weinberger, D. R., & Chen, J. (2008). Genetic dissection of the role of catechol-*O*-methyltransferase in cognition and stress reactivity in mice. *Journal of Neuroscience, 28*(35), 8709–8723.
5. Zhu, B. T., & Liehr, J. G. (1993). Inhibition of the catechol-*O*-methyltransferase-catalyzed *O*-methylation of 2- and 4-hydroxyestradiol by catecholamine: implications for the mechanism of estrogen-induced carcinogenesis. *Archives of Biochemistry and Biophysics, 304*(1), 248–256.
6. Zhu, B. T. (2002). On the mechanism of homocysteine pathophysiology and pathogenesis: a unifying hypothesis. *Histology and Histopathology, 17*(4), 1283–1291.
7. Lachman, H. M., Papolos, D. F., Saito, T., Yu, Y. M., Szumlanski, C. L., & Weinshilboum, R. M. (1997). Human catechol-*O*-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics, 6*(3), 243–250.
8. Malherbe, P., Bertocci, B., Caspers, P., Zurcher, G., & Da Prada, M. (1992). Expression of functional membrane-bound and soluble catechol-*O*-methyltransferase in *Escherichia coli* and a mammalian cell line. *Journal of Neurochemistry, 58*(5), 1782–1789.
9. Vilbois, F., Caspers, P., da Prada, M., Lang, G., Karrer, C., Lahm, H. W., & Cesura, A. M. (1994). Mass spectrometric analysis of human soluble catechol-*O*-methyltransferase expressed in *Escherichia coli*. Identification of a product of ribossomal frameshifting and of reactive cysteines involved in S-adenosyl-*L*-methionine binding. *European Journal of Biochemistry, 222*(2), 377–386.

10. Bai, H. W., Shim, J. Y., Yu, J., & Zhu, B. T. (2007). Biochemical and molecular modeling studies of the *O*-methylation of various endogenous and exogenous catechol substrates catalyzed by recombinant human soluble and membrane-bound catehcol-*O*-methyltransferase. *Chemical Research in Toxicology, 20*(10), 1409–1425.
11. Pedro, A. Q., Bonifacio, M. J., Queiroz, J. A., Maia, C. J., & Passarinha, L. A. (2011). A novel prokaryotic expression system for biosynthesis of recombinant human membrane-bound catechol-*O*-methyltranferase. *Journal of Biotechnology, 156*(2), 141–146.
12. Santos, F. M., Pedro, A. Q., Soares, R. F., Martins, R., Bonifacio, M. J., Queiroz, J. A., & Passarinha, L. A. (2013). Performance of hydrophobic interaction ligands for human membrane-bound catechol-*O*-methyl-transferase purification. *Journal of Separation Science, 36*(11), 1693–1702.
13. Correia, F. F., Santos, F. M., Pedro, A. Q., Bonifácio, M. J., Queiroz, J. A., & Passarinha, L. A. (2013). Recovery of biological active catechol-*O*-methyltransferase isoforms from Q-sepharose. *Journal of Separation Science, 37*(1–2), 20–29.
14. Tilgmann, C., Melen, K., Lundstrom, K., Jalanko, A., Julkunen, I., Kalkkinen, N., & Ulmanen, I. (1992). Expression of recombinant soluble and membrane-bound catechol-*O*-methyltransferase in eukaryotic cells and identification of the respective enzymes in rat brain. *European Journal of Biochemistry, 207*(2), 813–821.
15. Bertocci, B., Miggiano, V., Da Prada, M., Dembic, Z., Lahm, H. W., & Malherbe, P. (1991). Human catechol-*O*-methyltransferase: cloning and expression of the membrane-associated form. *Proceedings of the National Academy of Sciences of the United States of America, 88*(4), 1416–1420.
16. Ulmanen, I., Peranen, J., Tenhunen, J., Tilgmann, C., Karhunen, T., Panula, P., Bernasconi, L., Aubry, J. P., & Lundstrom, K. (1997). Expression and intracellular localization of catechol-*O*-methyltransferase in transfected mammalian cells. *European Journal of Biochemistry, 243*(1–2), 452–459.
17. Junge, F. B., Schneider, B., Reckel, S., Schwarz, V., Dotsch, V., & Bernhard, F. (2008). Large-scale production of functional membrane proteins. *Cellular and Molecular Life Sciences, 65*(11), 1729–1755.
18. Cregg, J. M., Cereghino, J. L., Shi, J., & Higgins, D. R. (2000). Recombinant protein expression in *Pichia pastoris. Molecular Biotechnology, 16*(1), 23–52.
19. Goncalves, A. M., Pedro, A. Q., Maia, C., Sousa, F., Queiroz, J. A., & Passarinha, L. A. (2013). *Pichia pastoris*, a recombinant microfactory for antibodies and human membrane proteins. *Journal of Microbiology and Biotechnology, 23*(5), 587–601.
20. Ramon, A., & Marin, M. (2011). Advances in the production of human membrane proteins in *Pichia pastoris. Biotechnology Journal, 6*(6), 700–706.
21. Vogl, T., & Glieder, A. (2013). Regulation of *Pichia pastoris* promoters and its consequences for protein production. *New Biotechnology, 30*(4), 385–404.
22. Burrowes, O. J., Diamond, G., & Lee, T. C. (2005). Recombinant expression of pleurocidin cDNA using the *Pichia pastoris* expression system. *Journal of Biomedicine & Biotechnology, 4*, 374–384.
23. Villatte, F., Hussein, A. S., Bachmann, T. T., & Schmid, R. D. (2001). Expression level of heterologous proteins in *Pichia pastoris* is influenced by flask design. *Applied Microbiology and Biotechnology, 55*(4), 463–465.
24. Jahic, M., Veide, A., Charoenrat, T., Teeri, T., & Enfors, S. O. (2006). Process technology for production and recovery of heterologous proteins with *Pichia pastoris. Biotechnology Progress, 22*(6), 1465–1473.
25. Cos, O., Ramon, R., Montesinos, J. L., & Valero, F. (2006). Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: a review. *Microbial Cell Factories, 5*, 17–37.
26. Chiruvolu, V., Cregg, J. M., & Meagher, M. M. (1997). Recombinant protein production in an alcohol oxidase-defective strain of *Pichia pastoris* in fed-batch fermentations. *Enzyme Microbiology Technology, 21*, 277–283.
27. Pontes, H., Pinho, P., Casal, S., Carmo, H., Santos, A., Magalhaes, T., Remião, F., Carvalho, F., & Lourdes Bastos, M. (2009). Determination of ethanol, metanol, acetone and acetaldehyde by a simple capillary direct injection gas chromatographic method in human whole blood, vitreous humour and urine and in cell culture medium. *Journal of Chromatographic Science, 47*(4), 272–278.
28. Parpinello, G., & Versari, A. (2000). A simple high-performance liquid chromatography method for the analysis of glucose, glycerol and methanol in a bioprocess. *Journal of Chromatographic Science, 38*(6), 259–261.
29. Almuzara, C., Cos, O., Baeza, M., Gabriel, D., & Valero, F. (2002). Methanol determination in *Pichia pastoris* cultures by flow injection analysis. *Biotechnology Letters, 24*, 413–417.
30. Zhang, W., Hywood Potter, K. J., Plantz, B. A., Schlegel, V. L., Smith, L. A., & Meagher, M. M. (2003). Pichia pastoris fermentation with mixed-feeds of glycerol and methanol: growth kinetics and production improvement. *Journal of Industrial Microbiology & Biotechnology, 30*(4), 210–215.
31. Katakura, Y., Zhang, W., Zhuang, G., Omasa, T., Kishimoto, M., Goto, Y., & Suga, K.-I. (1998). Effect of the methanol concentration on the production of human beta(2)-glycoprotein I domain V by a recombinant

Pichia pastoris: a simple system for the control of methanol concentration using a semiconductor gas sensor. *Journal of Fermentation and Bioengineering, 86*, 482–487.
32. Hellwig, S., Emde, F., Raven, N. P., Henke, M., Van Der Logt, P., & Fischer, R. (2001). Analysis of single-chain antibody production in *Pichia pastoris* using on-line methanol control in fed-batch and mixed-feed fermentations. *Biotechnology and Bioengineering, 74*, 344–352.
33. Celik, E., Calik, P., & Oliver, S. G. (2009). fed-batch methanol feeding strategy for recombinant protein production by Pichia pastoris in the presence of co-substrate sorbitol. *Yeast, 26*(9), 473–484.
34. Tenhunen, J., Salminen, M., Lundstrom, K., Kiviluoto, T., Savolainen, R., & Ulmanen, I. (1994). Genomic organization of the human catechol-*O*-methyltransferase gene and its expression from two distinct promoters. *European Journal of Biochemistry, 223*(3), 1049–1059.
35. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, 227*(5259), 680–685.
36. Passarinha, L. A., Bonifacio, M. J., & Queiroz, J. A. (2007). Comparative study on the interaction of recombinant human soluble catechol-*O*-methyltransferase on some hydrophobic adsorbents. *Biomedical Chromatography, 21*(4), 430–438.
37. Passarinha, L. A., Bonifacio, M. J., & Queiroz, J. A. (2006). The effect of temperature on the analysis of metanephrine for catechol-*O*-methyltransferase activity assay with electrochemical detection. *Biomedical Chromatography, 20*(9), 937–944.
38. Goodrick, J. C., Xu, M., Finnegan, R., Schilling, B. M., Schiavi, S., Hoppe, H., & Wan, N. C. (2001). High-level expression and stabilization of recombinant human chitinase produced in a continuous constitutive *Pichia pastoris* expression system. *Biotechnology and Bioengineering, 74*(6), 492–497.
39. Oberg, F., Ekvall, M., Nyblom, M., Backmark, A., Neutze, R., & Hedfalk, K. (2009). Insight into factors directing high production of eukaryotic membrane proteins; production of 13 human AQPs in *Pichia pastoris*. *Molecular Membrane Biology, 26*, 215–227.
40. Long, S. B., Tao, X., Campbell, E. B., & Mackinnon, R. (2007). Atomic structure of a voltage-dependent K^+ channel in a lipid membrane-like environment. *Nature, 450*, 376–382.
41. Orman, M. A., Çalik, P., & Ozdamar, H. O. (2009). The influence of carbon sources on recombinant-human-growth hormone production by *Pichia pastoris* is dependent on phenotype: a comparison of Mut^S and Mut^+ strains. *Biotechnology and Applied Biochemistry, 52*, 245–255.
42. Horstkotte, B., Arnau, C., Valero, F., Elsholz, O., & Cerdà, V. (2008). Monitoring of sorbitol in *Pichia pastoris* cultivation applying sequential injection analysis. *Biochemical Engineering Journal, 42*, 77–83.
43. Food and Drug administration - *Guidance for Industry – bioanalytical method validation.*, U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM) (2001), Available from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070107.pdf. Assessed 20/03/2014.
44. Zhang, P., Zhang, W., Zhou, X., Bai, P., Cregg, J. M., & Zhang, Y. (2010). Catabolite repression of Aox in *Pichia pastoris* is dependent on hexose transporter PpHxt1 and pexophagy. *Applied and Environmental Microbiology, 76*(18), 6108–6118.
45. Sreekrishna, K., Brankamp, R. G., Kropp, K. E., Blankenship, D. T., Tsay, J. T., Smith, P. L., Wierschke, J. D., Subramaniam, A., & Birkenberger, L. A. (1997). Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene, 190*, 55–62.
46. Kim, S. I., Wu, Y., Kim, K. L., Kim, G. J., & Shin, H. J. (2013). Improved cell disruption of Pichia pastoris utilizing aminopropyl magnesium phyllosilicate (AMP) clay. *World Journal of Microbiology and Biotechnology, 29*(6), 1129–1132.
47. Giersing, B., Miur, a. K., Shimp, R., Wang, J., Zhou, H., Orcutt, A., Stowers, A., Saul, A., Miller, L. H., Long, C., & Singh, S. (2005). Posttranslational modification of recombinant *Plasmodium falciparum* apical membrane antigen 1: impact on functional immune responses to malaria vaccine candidate. *Infection and Immunity, 73*(7), 3963–3970.
48. Abad, S., Nahalka, J., Bergler, G., Arnold, S. A., Speight, R., Fotheringam, I., Nidetzsky, B., & Glieder, A. (2010). Stepwise engineering of a *Pichia pastoris* D-amino acid oxidase whole cell catalyst. *Microbial Cell Factories, 9*, 24–36.
49. Daniels, M. J., Wood, M. R., & Yeager, M. (2006). *In vivo* functional assay of a recombinant aquaporin in *Pichia pastoris*. *Applied Environmental Microbiology, 72*(2), 1507–1514.
50. Pla, I. A., Damasceno, L. M., Vannelli, T., Ritter, G., Batt, C. A., & Shuler, M. L. (2006). Evaluation of Mut^S and Mut^+ *Pichia pastoris* phenotypes for high level extracellular scFv expression under feedback control of the methanol concentration. *Biotechnology Progress, 22*, 881–888.
51. Bonifacio, M. J., & Soares-da-Silva, P. (2003). Purification of membrane-bound catechol-*O*-methyltransferase. *Methods in Molecular Biology, 228*, 231–238.
52. Prive, G. C. (2007). Detergents for the stabilization and crystallization of membrane proteins. *Methods, 41*(4), 388–397.

Paper V

An artificial neural network for membrane-bound catechol-*O*-methyltransferase biosynthesis with *Pichia pastoris* methanol-induced cultures

A. Q. Pedro, L. M. Martins, J. M. Dias, M. J. Bonifácio, J. A. Queiroz, L. A. Passarinha

Microbial Cell Factories. 2015. 14: 113-127

Short description: Taking into account the work developed in paper IV, here we report the expression of MBCOMT at medium-scale in mini-bioreactors with *P. pastoris* methanol-induced cultures. Initially, a first standard strategy for MBCOMT bioreactor biosynthesis was developed following the optimization of several fermentation parameters. Then, in order to maximize MBCOMT biological activity, several independent variables (methanol concentration, dimethylsulfoxide concentration and temperature) associated with the methanol induction phase were optimized using Artificial Neural Networks (ANN).

Pedro *et al. Microb Cell Fact (2015) 14:113*
DOI 10.1186/s12934-015-0304-7

MICROBIAL CELL FACTORIES

RESEARCH Open Access

An artificial neural network for membrane-bound catechol-*O*-methyltransferase biosynthesis with *Pichia pastoris* methanol-induced cultures

Augusto Q Pedro[1], Luís M Martins[1], João M L Dias[2], Maria J Bonifácio[3], João A Queiroz[1] and Luís A Passarinha[1*]

Abstract

Background: Membrane proteins are important drug targets in many human diseases and gathering structural information regarding these proteins encourages the pharmaceutical industry to develop new molecules using structure-based drug design studies. Specifically, membrane-bound catechol-*O*-methyltransferase (MBCOMT) is an integral membrane protein that catalyzes the methylation of catechol substrates and has been linked to several diseases such as Parkinson's disease and Schizophrenia. Thereby, improvements in the clinical outcome of the therapy to these diseases may come from structure-based drug design where reaching MBCOMT samples in milligram quantities are crucial for acquiring structural information regarding this target protein. Therefore, the main aim of this work was to optimize the temperature, dimethylsulfoxide (DMSO) concentration and the methanol flow-rate for the biosynthesis of recombinant MBCOMT by *Pichia pastoris* bioreactor methanol-induced cultures using artificial neural networks (ANN).

Results: The optimization trials intended to evaluate MBCOMT expression by *P. pastoris* bioreactor cultures led to the development of a first standard strategy for MBCOMT bioreactor biosynthesis with a batch growth on glycerol until the dissolved oxygen spike, 3 h of glycerol feeding and 12 h of methanol induction. The ANN modeling of the aforementioned fermentation parameters predicted a maximum MBCOMT specific activity of 384.8 nmol/h/mg of protein at 30°C, 2.9 mL/L/H methanol constant flow-rate and with the addition of 6% (v/v) DMSO with almost 90% of healthy cells at the end of the induction phase. These results allowed an improvement of MBCOMT specific activity of 6.4-fold in comparison to that from the small-scale biosynthesis in baffled shake-flasks.

Conclusions: The ANN model was able to describe the effects of temperature, DMSO concentration and methanol flow-rate on MBCOMT specific activity, as shown by the good fitness between predicted and observed values. This experimental procedure highlights the potential role of chemical chaperones such as DMSO in improving yields of recombinant membrane proteins with a different topology than G-coupled receptors. Finally, the proposed ANN shows that the manipulation of classic fermentation parameters coupled with the addition of specific molecules can open and reinforce new perspectives in the optimization of *P. pastoris* bioprocesses for membrane proteins biosynthesis.

Keywords: Catechol-*O*-methyltransferase, Artificial neural network, Bioreactor, *Pichia pastoris*, DMSO, Alcohol oxidase, Membrane protein

*Correspondence: lpassarinha@fcsaude.ubi.pt
[1] CICS-UBI, Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, Avenida Infante D. Henrique, 6201-001 Covilhã, Portugal
Full list of author information is available at the end of the article

Background

Membrane proteins (MP) are central to many cellular processes: they are involved in the uptake and export of diverse charged and uncharged molecules, as well as mediating the interaction of cells with their environment [1]. As a consequence, they are of prime importance as drug targets to the pharmaceutical industry [1]. Catechol-*O*-methyltransferase (COMT, EC 2.1.1.6) is a magnesium-dependent enzyme that catalyzes the methylation of catechol substrates using *S*-adenosyl-L-methionine (SAM) as a methyl donor and yielding, as reaction products, the *O*-methylated catechol and *S*-adenosyl-L-homocysteine [2]. In humans, COMT appears as two molecular forms, a soluble and a membrane-bound isoform (MBCOMT) that is found mainly associated with the rough endoplasmic reticulum membrane [2]. Specifically, SCOMT is a nonglycosylated protein containing 221 amino acid residues and a molecular weight of 24.7 kDa while MBCOMT has an additional peptide in its amino terminal of 50 amino acid residues and a molecular weight of 30 kDa [2].This extra peptide contains a stretch of 21 hydrophobic amino acid residues that constitute the membrane anchor region [2]. In fact, MBCOMT is an integral membrane protein with the catalytic portion of the enzyme oriented toward the cytoplasmic side of the membrane [2]. Recently, MBCOMT has gained a major importance as therapeutic target due to its high abundance in human brain and its higher affinity for catechol substrates when compared to soluble isoform [2]. During the last decades, COMT has been implicated in several diseases such as cardiovascular diseases [3], estrogen-induced cancers [4] and neurologic disorders [2]. Specifically, the best documented is the important role that COMT plays in Parkinson's disease whose most effective treatment remains the dopamine replacement therapy with levodopa together with an inhibitor of aromatic amino acid decarboxylase and a COMT inhibitor [2]. Therefore, it becomes clear the importance of developing new and more effective drugs for COMT inhibition in which structure-based drug design can play an important role in this process. However, in order to structurally and functionally characterize a MP, a stable active sample is required, meaning the requirement for a regular supply of milligram quantities of purified MP [1]. The foremost requirements associated with the majority of biophysical techniques emphasize the importance of developing new systems capable of delivery biologically active MBCOMT in higher quantities from high cell-density cultures. Around the mid of the twentieth century, bacteria and filamentous fungi have taken over the lead role in the development of bioprocesses [5]. However, novel developments of recombinant protein production, metabolic engineering and systems biology open a range of new applications of yeasts in the upstream stage of a bioprocess [5]. In fact, over the last two decades, the methylotrophic *Pichia pastoris* (*P. pastoris*) has been established as one of the most frequently used expression systems for recombinant protein production [6]. The benefits of this system include growth up to high cell densities quantity on defined minimal medium, high expression level of heterologous proteins, typical eukaryotic post-translational modifications, efficient secretion of extracellular proteins and the presence of the efficient methanol-inducible promoter from alcohol oxidase I gene (AOX) [7, 8]. Moreover, the *P. pastoris* preference for respiratory rather than fermentative metabolism, even at high cell density processes, prevents the accumulation of secondary metabolites such as ethanol and acetic acid [7]. Finally, following the recognition of *P. pastoris* as a GRAS organism by FDA in 2006 [6], the importance of this host as a platform for biopharmaceuticals production is highlighted. Upon the design of a bioprocess for recombinant protein production in *P. pastoris* under the control of the AOX promoter, a key step is the optimization of the induction phase since it will directly impact on the yield of the process [9]. Over the past few years, many efforts have allowed relevant advances in the development of *P. pastoris* for the production of MP where significant achievements were made in order to improve yield and proper folding of these target proteins [10]. Specifically, chemical chaperones such as dimethylsulfoxide (DMSO) have been shown to increase the expression of different G protein-coupled receptors such as the human neuromedin U subtype II receptor [11], the human adenosine A_{2A} receptor or the human β2-adrenergic receptor [12], mostly due to the up-regulation of the expression of genes involved in membrane lipid components [10, 13]. In addition, it was also reported that lowering the culture temperature from 30 to 20°C also leads to an improvement of the expression of MP, possibly because it slows down protein production, not overloading the translocation machinery, protein processing or intracellular trafficking [13]. Finally, while the methanol feeding strategy is one of the most important factors for maximizing heterologous protein expression, the methanol induction phase may also depend on other operational conditions (temperature, pH and culture medium), phenotype and specific characteristics of the heterologous protein produced [14]. In general, the traditional optimization method, commonly called "one factor/variable at a time", consists in varying one factor while keeping the others constant [15, 16] and is extremely time-consuming requiring a large number of experiments [15]. In alternative, statistical experimental designs have been widely used and they can be applied at distinct phases of an optimization process, either for screening experiments or for searching

for the optimal conditions for targeted response(s) [17]. Overall, response surface methodology (RSM), which includes factorial design and regression analysis, seeks to identify and optimize significant factors to maximize the response [18]. On the other hand, artificial neural networks (ANN) allow estimating relationships between one or more input and one or more output (also called responses) [16]. In general, ANNs are greater and more accurate modeling techniques when compared with RSM since they can cope with nonlinearities among the factor in the prediction of a given response [18]. Indeed, ANNs coupled with design of experiments have been successfully applied in diverse areas such as the optimization of the culture conditions [16, 18], pharmaceutics [19] or chromatography [15, 20].

The main aim of this work was to optimize the induction phase for recombinant MBCOMT production by *P. pastoris* X33 Mut^+ cultures in bioreactor applying central composite design (CCD) and ANNs.

Results and discussion

The structural and functional characterization of a MP depends on the production of a sufficient amount of active protein, meaning a regular supply of milligram quantities of the target enzyme [1]. Therefore, to fulfill this requirement, in this work and for the first time the biosynthesis of MBCOMT by *P. pastoris* bioreactor cultures is reported. Initially, in order to select the most appropriated *P. pastoris* strain for MBCOMT biosynthesis, trials at a small-scale in baffled shake-flasks were carried out. Then, a three-stage bioprocess for the biosynthesis of the target protein by *P. pastoris* bioreactor cultures was implemented and the lengths of the glycerol fed-batch and the methanol induction phases were optimized.

Moreover, after selecting a set of independent variables associated with the methanol induction phase that greatly influence the levels of the MBCOMT, ANN modeling was carried out in order to maximize the biological activity of the target protein. The massic and volumetric productivities were not incorporated as an output since the values of those parameters are in strictly dependence on MBCOMT biological activity [18]. Also, the biomass levels were evaluated in all assays performed in this work but were not considered in the optimization and validation procedures as an output, since higher biomass levels not always lead to higher mass productivities of the target protein.

Small-scale MBCOMT biosynthesis in *P. pastoris*

Membrane-bound catechol-*O*-methyltransferase biosynthesis was initially carried out in shake-flasks containing BMGH medium using a Mut^+ (X33) and a Mut^S (KM71H) *P. pastoris* strains [21]. Sometimes, an increase in the number of the heterologous gene can possibly lead to an increase in transcription and translation rate of the desired gene [22]. In fact, although opposite results had already been published, there are several examples including the mouse epidermal growth factor or miniproinsulin in which higher target gene copy numbers lead to higher titers for *P. pastoris* bioprocesses driven by AOX1 promoter [22]. Therefore, upon the transformation procedure with the target recombinant plasmid, clones from both strains in study were isolated from plates containing high zeocin concentrations (2 mg/mL). Following the isolation of these clones from both strains, it was determined the target gene copy number that was integrated in each strain. Therefore, using the method previously reported by Nordén and collaborators [23] that takes advantage of the fact that part of the plasmid pPICZ α, namely the AOX1 TT region is incorporated in the *P. pastoris* genome together with the gene to be expressed. In particular, for the X33 strain, the primer efficiencies were 1.88 and 1.87, respectively for the AOX1 TT and AOX2 PROM primer pairs. Similarly, for the KM71H strain, the primer efficiencies were 1.91 and 1.94, respectively, for the AOX1 TT and AOX2 PROM primer pairs. Finally, according the equation described in the "Methods", the target gene copy number introduced in each recombinant strain was determined and it was found that X33-PICZα-MBCOMT had nine copies of the target plasmid while the KM71H-PICZα-MBCOMT had ten copies. In fact, Nordén and coworkers [23] reported with the aquaporins that colonies isolated from 0.5 mg/mL zeocin could harbor from 4 to 15 plasmids while from 1 mg/mL, as many as 17 heterologous DNA sequences can be incorporated. Therefore, although the isolation of clones from plates containing higher antibiotic concentrations doesn't exclude completely the occurrence of false positives, the values here reported (9 and 10 copies for the X33 and KM71H strains, respectively) are in the same order of magnitude. Then, small-scale fermentation trials were carried out using 0.5% (v/v) methanol and higher biomass levels were detected for the X33 strain (OD_{600} = 7.5) when compared with those obtained for the KM71H strain (OD_{600} = 1.8). Similarly, the target enzyme recovered from the X33 strain presented higher biological activity (60.25 nmol/h/mg) in comparison to KM71H cells (25.77 nmol/h/mg of protein) [21]. On the other hand, when the methanol concentration is lowered from 1 to 0.25% (v/v), similar values for MBCOMT biological activity were obtained for the X33 (61.73 nmol/h/mg of protein) and the KM71H (60.62 nmol/h/mg of protein) strains [21]. Specifically, we believe that the observed differences in these two strains concerning their performance in MBCOMT biosynthesis seem to

be associated with the methanol concentration used for induction and not for example with the target gene copy number inserted in the genome since it is similar.

The value previously reported [21] with both *P. pastoris* strains for MBCOMT biological activity is higher than those previously reported by our research group using *Brevibacillus choshinensis* as the expression system (48.07 nmol/h/mg of protein) [24]. In general, for intracellular expression, it has been reported that it is preferable use Mut^S over Mut^+ *P. pastoris* strains because of increased specific yield of heterologous protein [25]. However, as previously reported by Maurer and collaborators, the volumetric productivity QP is the most plausible target for optimization in fed-batch processes [26]. Therefore, since the main aim of this work was to maximize MBCOMT expression irrespective the biomass levels, *P. pastoris* Mut^+ X33 was chosen for further bioreactor trials since regardless the methanol concentration used, the expression levels of the target protein were the highest obtained and they didn't significantly change when different methanol concentrations are applied.

MBCOMT biosynthesis from methanol-induced *Pichia pastoris* bioreactor cultures

Membrane-bound catechol-*O*-methyltransferase biosynthesis was carried out in mini-bioreactors (working volume 0.25 L) in modified basal salts medium (BSM) containing 4.35 mL/L trace metal solution (SMT) [27] and the pH was adjusted to 4.7 in order to minimize precipitation and, consequently, undesired operational problems such as starvation of nutrients and optical densities measurement interferences [14]. *P. pastoris* cultivations in bioreactor were initiated with a glycerol batch phase (30 g/L glycerol) that ends when glycerol was depleted, indicated by a sharp increase in the dissolved oxygen (DO) [14]. After this stage, a fed-batch growth on glycerol [50% (v/v) at 18.54 mL/L/H] during different periods was employed, followed by the methanol induction phase where *P. pastoris* was cultivated on a methanol fed-batch mode. In order to promote the derepression of the AOX promoter prior to induction, 1 h before starting the induction phase, methanol was added to the reaction vessel at the flow-rate later employed in the methanol fed-batch phase.

Preliminary trials were carried out in order to analyze the optimal period of the glycerol fed-batch phase as well as the optimal duration of the methanol induction phase that maximize MBCOMT expression. Therefore, keeping constant the methanol flow-rate (3.6 mL/L/H) in the induction phase, assays with 3, 5 or 7 h glycerol fed-batch phase were carried out. Methanol induction phase was maintained during 60 h and samples were collected with an interval of 2 h until 12 h and then every 12 h to follow the MBCOMT expression profile. As depicted in Fig. 1, the highest MBCOMT biological activity levels were detected when a 3 h period was applied in the glycerol fed-batch phase. In addition, concerning to the methanol induction phase, MBCOMT achieved a maximum expression of 121.0 nmol/h/mg of protein at 12 h of induction, what led us to assume a 3 h glycerol fed-batch period and a 12 h induction period for further experiments. In fact, a shorter induction period can be greatly advantageous over other previously reported strategies [27, 28] where induction usually takes more than 48 h, being more time-consuming and laborious. Moreover, the shorter induction period allows terminating the fermentation before a decrease in the cell's physiological activity is observed [29].

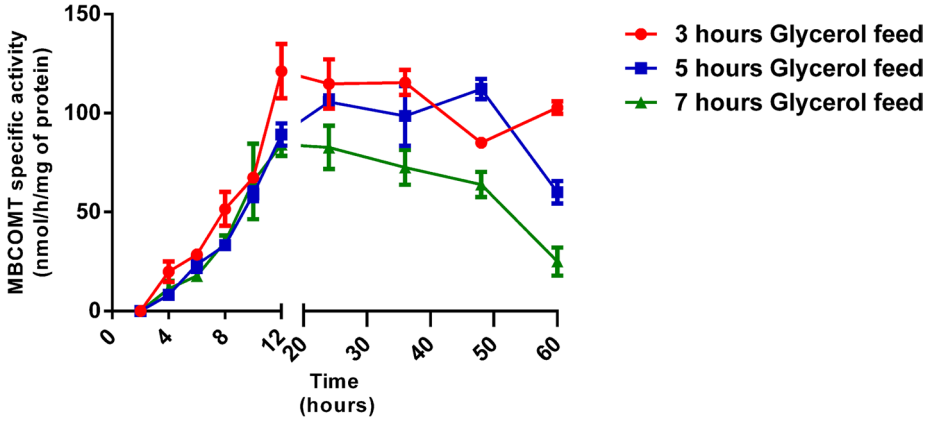


Fig. 1 Typical time profile of MBCOMT specific activity (nmol/h/mg of protein) obtained by *P. pastoris* bioreactor cultures using different periods of the glycerol fed-batch phase with a methanol constant feed flow-rate at 3.6 mL/L/H (each value represents the mean of three independent samples).

Following these findings, we evaluated if the expression of the target protein was significantly affected by the methanol constant flow-rate as well as the addition of the chemical chaperone DMSO that has been described to increase the expression levels of some MP [11–13, 30, 31]. Therefore, keeping constant the operational parameters previously optimized, distinct assays were performed: with different methanol constant flow rates at 2, 3.6 and 5.2 mL/L/H while others were performed maintaining the methanol flow-rate at 3.6 mL/L/H and changing the DMSO concentration [2.5, 5 and 7.5% (v/v)] in the culture according to what previously described [11–13, 30]. As demonstrated in Fig. 2a, for the lowest methanol constant flow-rate (2 mL/L/H), a highest MBCOMT expression level of 158 nmol/h/mg were obtained, contrasting with 120 and 107 nmol/h/mg for 3.6 and 5.2 mL/L/H, respectively. Also, the methanol and the biomass levels at distinct stages of the induction phase were quantified in these assays, as depicted in Fig. 2b and Table 1, respectively. In general, for the different methanol flow-rates applied, the methanol levels increase from 0 to 6 h and then they decrease until the end of the induction phase. At the early stage of the induction phase, methanol doesn't seem to be consumed in a large extent since *P. pastoris* cells may be going through a transition period where they stop consuming glycerol and start to oxidize methanol. Nevertheless, it is possible to observe that for methanol constant-flow rates of 3.6 and 5.2 mL/L/H, the concentration of methanol in the culture broth is higher (near 10 and 12.5 g/L respectively) at 6 h of induction when compared with the lowest flow-rate employed (1 g/L). Therefore, it is feasible to assume that using a lower flow rate (2 mL/L/H) allows the establishment of an appropriated balance between activation of the AOX promoter and, consequently, production of the target enzyme and accumulation of methanol in the culture medium that can be responsible for the undesired toxicity, as it may be happening for 3.6 and 5.2 mL/L/H [14]. Moreover, an optimal ratio of methanol to cell concentration should be applied [32], otherwise high methanol feeding rates stress the cell machinery and negatively affect the process performance [32, 33].

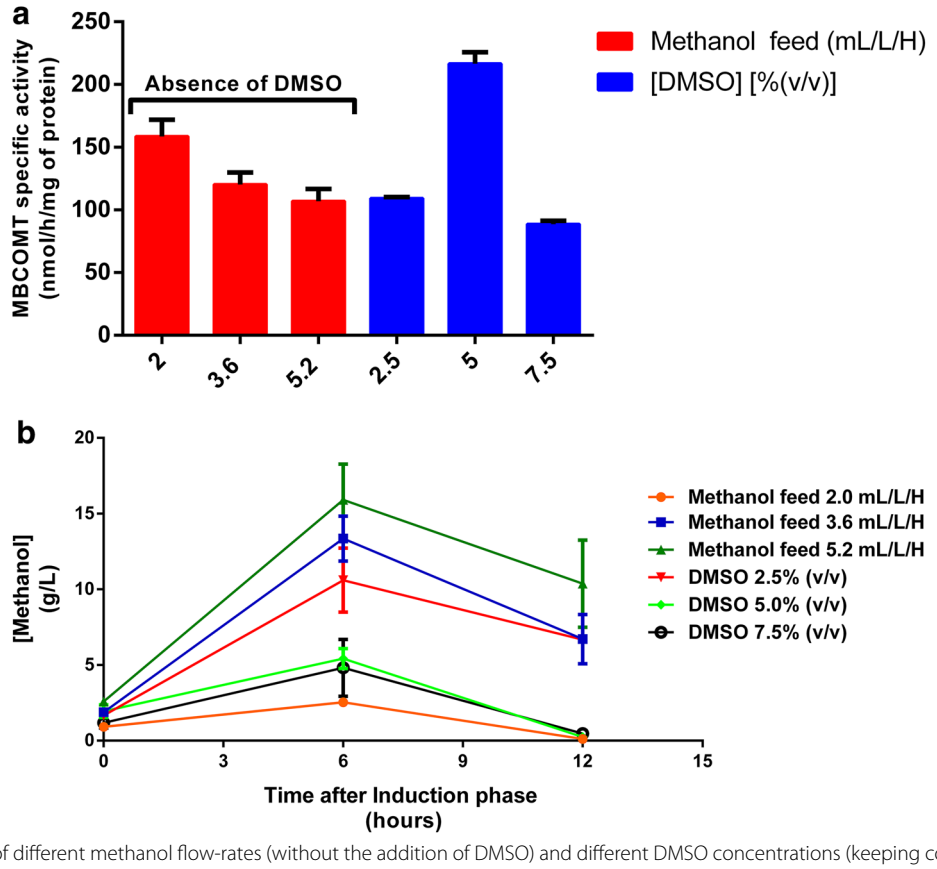


Fig. 2 **a** Analysis of different methanol flow-rates (without the addition of DMSO) and different DMSO concentrations (keeping constant the methanol flow-rate at 3.6 mL/L/H) on MBCOMT specific activity (nmol/h/mg of protein) obtained by *P. pastoris* bioreactor cultures. **b** Time course analysis of the methanol levels in the above mentioned assays measured by HPLC-RID. In both experiments, a 3-h period of the glycerol fed-batch was applied and induction was conducted during 12 h (each value represents the mean of three independent samples).

Table 1 Time course profile of the biomass levels (measured as $OD_{600\ nm}$) obtained in the trials where the methanol constant feed flow-rate (2, 3.6 and 5.2 mL/L/H) and the DMSO levels added to the culture were changed, in accordance with the results shown in Fig. 2b

Time after induction phase (h)	Optical density measurements at 600 nm					
	Methanol constant feed flow-rate			DMSO concentration		
	2 mL/L/H	3.6 mL/L/H	5.2 mL/L/H	2.5% (v/v)	5% (v/v)	7.5% (v/v)
3	111.75 ± 1.23	105.19 ± 5.57	116.75 ± 4.42	112.75 ± 4.95	113.44 ± 4.33	104.88 ± 2.47
9	110.32 ± 2.38	106.88 ± 7.95	110.5 ± 3.36	110.19 ± 2.21	114.43 ± 1.17	113.44 ± 0.27
15	111.31 ± 4.68	111.5 ± 9.02	117.38 ± 2.47	116.31 ± 2.21	132.00 ± 7.07	115.06 ± 3.62

On the other hand, when different DMSO concentrations were added to the *P. pastoris* cultures, the highest MBCOMT biosynthesis of 216 nmol/h/mg was detected for 5% (v/v), which represents an increase of 1.8-fold when compared with the control (without DMSO). Again, the methanol levels were also quantified in these trials and interestingly, its time course profile with the addition of 5% (v/v) DMSO conducted with 3.6 mL/L/H of methanol resembles the profile previously obtained for the 2 mL/L/H methanol flow rate and not the 3.6 mL/L/H. Following these results, it is reasonable to think that adjusting the DMSO concentration to the cell needs, the methanol is more efficiently used what, in a last analysis, leads to an increase in the biosynthesis of the target protein.

The addition of 5% (v/v) DMSO proved to have a positive effect on the expression of this particular MP, has been demonstrated previously for G protein-coupled receptors by other authors [11–13, 30, 31]. Although the mechanism by which DMSO increases MP expression is not yet fully understood, Murata and collaborators showed that DMSO induces membrane proliferation through the increase of the phospholipid content within *Saccharomyces cerevisiae* cells [34]. On the other hand, it was also reported that DMSO possess antioxidant properties, preventing protein oxidation (increase in protein carbonyl content and decrease in free thiol content) in rat brain homogenates induced by ferrous chloride/hydrogen peroxide [35]. Therefore, it is likely that the benefits of using DMSO on the expression of membrane proteins can be associated with the induction of membrane proliferation or with the reduction of protein oxidation or a combination of both. Moreover, despite the optimal temperature for growth and production of proteins in *P. pastoris* is 30°C [14], some authors claim that working at lower temperatures (until 20 to 25°C) may improve the target protein biosynthesis [36], lower cell lysis [37] and decrease the proteolytic activity [38]. Therefore, in this work, the temperature was also included as an independent process parameter to optimize MBCOMT biosynthesis from *P. pastoris* and the ranges (20, 25 and 30°C) were selected according to what has been reported in the literature [14, 37].

According to the results reported in this section and the synergy observed between methanol flow rate and DMSO concentration in the culture broth, the most appropriated ranges of the independent variables selected for performing the experimental design were defined, as shown in Table 2. Finally, a summary of the optimized conditions for the expression of MBCOMT from *P. pastoris* bioreactor methanol-induced cultures is presented in Fig. 3 where the ranges of the independent variables selected for the ANN modeling are presented as well as the major experimental conditions selected.

Experimental design and artificial neural network modeling

A set of 17 experiments defined by CCD for optimization of the induction phase for maximizing MBCOMT biosynthesis in *P. pastoris* culture are listed in Tables 2 and 3. In general, lower MBCOMT biological activity levels were detected when the input variables defined in CCD were at the lowest levels. Specifically, MBCOMT biosynthesis is maximized at higher methanol constant-flow rate concentrations and when the concentration of DMSO added is higher. On the other hand, an increased in the induction temperature coupled to an increase in the other input variables also lead to an increase in biologically active MBCOMT expression. According to the ANN modeling results in calibration dataset (DoE experiments) (Table 3), the predicted maximum for MBCOMT

Table 2 Coded levels used for temperature, methanol constant feed flow-rate and DMSO in the CCD

Input variables	Coded levels		
	−1	0	1
Temperature (°C)	20	25	30
Methanol constant feed rate (mL/h/L)	1	2	3
DMSO [% (v/v)]	4	5	6

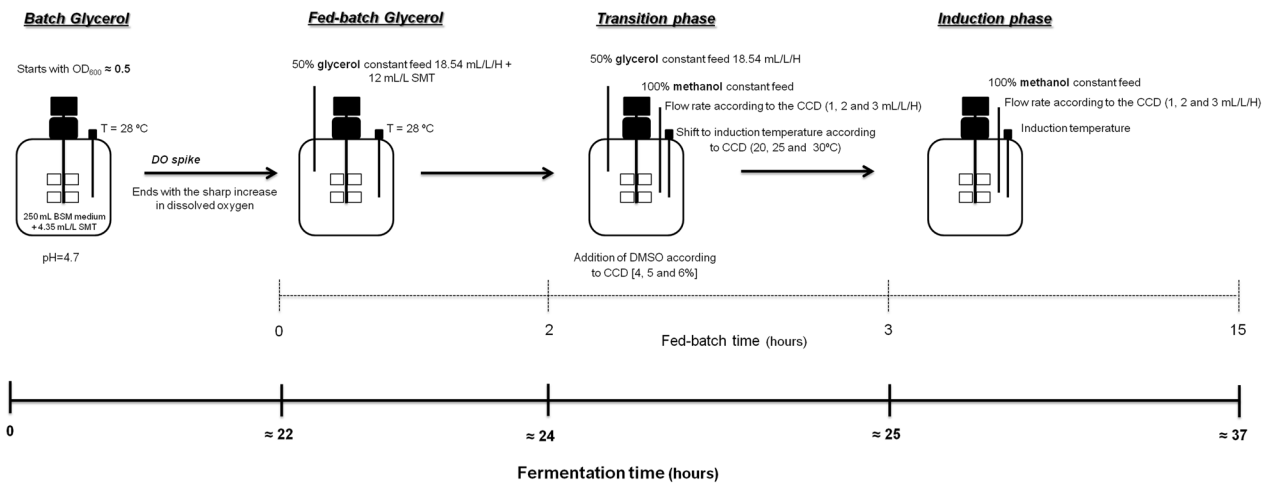


Fig. 3 Structure of the optimized four-stage bioprocess implemented in this work for recombinant MBCOMT biosynthesis by *P. pastoris* bioreactor cultures.

specific activity (384.8 nmol/h/mg of protein) was achieved at 30°C, 2.9 mL/L/H methanol constant flow-rate and with the addition of 6% (v/v) DMSO. In general, as previously demonstrated for others MP [11–13, 30, 31], the addition of DMSO to the culture proved to have a positive effect on MBCOMT expression since over the model optimization the maximum target protein specific activity is achieved at higher DMSO concentrations. In addition, the output seems to be maximized when the methanol constant flow-rate and the induction temperature are close to the upper values defined in the CCD. This can be explained by the increase in the biomass levels (data not shown) caused by the increase in the temperature and, since there is more methanol that is being oxidized by the AOX promoter, the supply of inducer needs to be higher in order to maintain induction. An ANN model was developed in order to optimize the induction phase for maximizing MBCOMT biosynthesis from *P. pastoris* bioreactor cultures. The model was calibrated with the experiments defined in Table 3.

Modeling of the methanol induction phase using artificial neural network

The ANN model was applied for the optimization of the induction phase for MBCOMT biosynthesis in *P. pastoris* bioreactor cultures using a stepwise process until the maximum MBCOMT biological activity was achieved. Four iterations were required to achieve the maximum MBCOMT specific activity (384.8 nmol/h/mg of protein) under the optimal conditions [30°C, 2.9 mL/L/H methanol constant flow-rate and 6% (v/v) DMSO] and to validate the model with new experiments. In the end, an improvement of 1.53-fold over the best conditions performed in the DoE step (see experiment 15, Table 3) was achieved while an improvement of 6.4-fold over the small-scale biosynthesis in baffled shake-flasks was achieved.

The obtained ANN model is mostly unbiased because the slope and R^2 of the fitting between the measured and predicted output were close to 1 (0.9064 and 0.97161, respectively) (see Fig. 4). In Fig. 5 are depicted the contour plots obtained from the ANN model for two combinations between the three operational conditions in study. The modeling results showed that the MBCOMT specific activity is sensitive to the operational conditions. The ANN parameters for the final validation model are presented in Additional file 1.

Bioprocess monitoring at the optimal conditions estimated by the ANN model

At the optimal conditions estimated by the ANN model [30°C, 2.9 mL/L/H methanol constant flow-rate and 6% (v/v) DMSO], the carbon source levels as well as the biomass levels and the number of viable/depolarized cells were analyzed, as depicted in Fig. 6. In what concerns to the *P. pastoris* growth, a marked increase in OD_{600} was detected between the end of the batch phase and the fed-batch growth of glycerol and it keeps increasing until the end of the induction phase with a value near 123 units of OD_{600}. The methanol and glycerol levels were quantified using a HPLC with refractive index detection and it was verified that the glycerol concentration also increases during the fed-batch glycerol phase, despite the higher accumulation of biomass during this stage. On the other hand, a low consumption of methanol was verified between the second and the third hours of the glycerol

Table 3 List of experiments performed for MBCOMT biosynthesis from *P. pastoris* bioreactor methanol-induced cultures based on CCD and ANN modeling

Experiment number (ANN model iterations)	Input variables level			Output	
	Methanol constant feed flow-rate (mL/L/h)	Induction temperature (°C)	DMSO concentration [%(v/v)]	Observed	Predicted
DoE					
1	1	20	4	126.1	122.2
2	3	20	4	163.9	139.3
3	1	30	4	47.4	97.0
4	3	30	4	188.0	187.1
5	1	20	6	138.0	143.8
6	3	20	6	130.6	139.4
7	1	30	6	151.4	97.5
8	3	30	6	(153.9)	358.1
9	1	25	5	105.3	116.5
10	3	25	5	137.9	134.2
11	2	20	5	115.2	136.8
12	2	30	5	101.1	105.5
13	2	25	4	183.9	197.2
14	2	25	6	222.6	218.5
15	2	25	5	252.5	243.3
16	2	25	5	243.8	243.3
17	2	25	5	230.3	243.3
I					
18	1	22.5	6	364.3	343.1
19	1	22.5	6	364.6	343.1
20	1	22.5	6	357.6	343.1
II					
21	2.9	30	6	390.6	383.1
22	2.9	30	6	391.5	383.1
III					
23	3	30	6	377.1	358.1
24	3	30	6	377.4	358.1
IV					
25	2.5	30	6	263.0	258.9
26	2.5	30	6	283.7	258.9
Final validation					
27	2.9	30	6	–	384.8

The predicted values of MBCOMT specific activity (nmol/h/mg of protein) are those obtained in the last optimization iteration. Observed outputs in parentheses represent the model outliers.

fed-batch phase since we consider that the consumption of glycerol is preferred over the methanol. On the other hand, at the end of the induction phase, almost no methanol was detected since *P. pastoris* cells are oxidizing it all, what can be indicating that the AOX promoter is highly active. Finally, the flow cytometry analysis led us to conclude that the changes introduced at the second hour of the glycerol fed-batch phase (namely the shift to the induction temperature, the addition of DMSO and the initiation of the methanol flow-rate) did not altered significantly the number of viable cells (94.8% compared to 95.4%) in culture. Furthermore, at the end of the induction phase, approximately 90% of viable cells were obtained, a value that is acceptable and is in accordance with *P. pastoris* bioprocesses that include the AOX promoter with a 12 h induction period [39].

To our best knowledge, this is the first systematic study where the interaction between two commonly studied operational parameters (induction temperature and methanol flow rate) and the addition of chemical

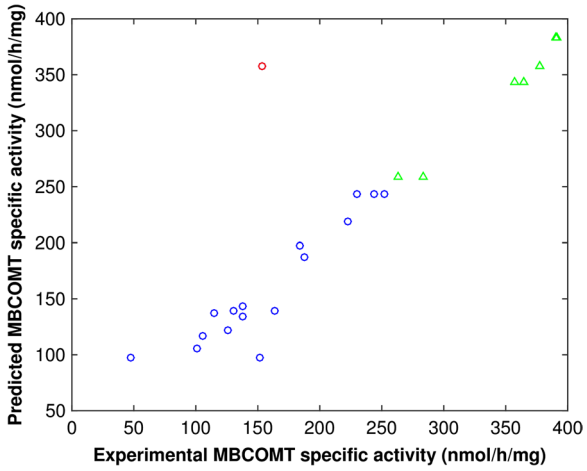


Fig. 4 ANN modeling results of MBCOMT specific activity for the last optimization steps. *Blue circle*, *red circles* and *green triangles* represent data from the CCD, outliers and from model optimization.

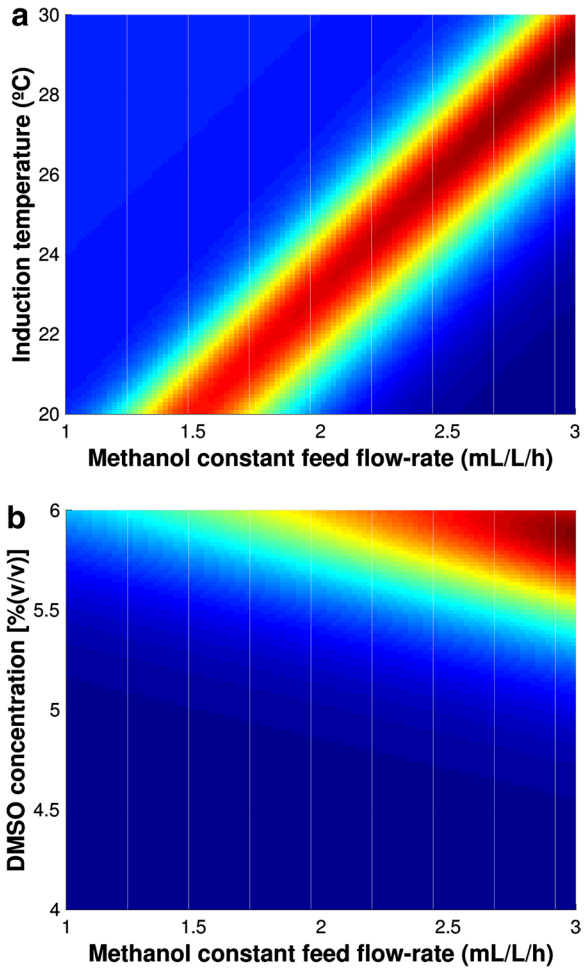


Fig. 5 Contour plots of MBCOMT specific activity as function of induction temperature, methanol constant flow-rate and DMSO concentration: **a** modeling results of MBCOMT specific activity as function of DMSO concentration and methanol constant flow-rate for the last optimization step. **b** Modeling results of MBCOMT specific activity as function of induction temperature and methanol constant flow-rate for the last optimization step.

chaperones (specifically the DMSO) are successfully reported to optimize MP expression by *P. pastoris* bioprocesses using ANN modeling.

Conclusions

Membrane-bound catechol-*O*-methyltransferase biosynthesis in a highly biological active form was successfully attained for the first time by *P. pastoris* bioreactor cultures under the control of the AOX promoter. The ANN model was able to describe the effects of temperature, DMSO concentration and methanol flow-rate on MBCOMT specific activity, as shown by the good fitness between the predicted and measured values. At the optimal conditions estimated by the ANN model [30°C, 2.9 mL/L/H methanol constant flow-rate and 6% (v/v) DMSO], a 1.58-fold increase was obtained for MBCOMT specific activity (384.8 nmol/h/mg of protein) over the highest value achieved in the experimental design while an improvement of 6.4-fold was found over the small-scale biosynthesis in baffled shake-flasks. Furthermore, in these conditions, almost 90% of viable cells were obtained at the end of the induction phase, indicating that the implemented experimental strategy allowed maintaining the viability of *P. pastoris* cultures. This experimental procedure highlighted the potential of chemical chaperones such as DMSO play to improve the yield of recombinant membrane proteins with a different topology than G-coupled receptors. In addition, this is the first systematic study where the interaction between two commonly studied operational parameters (induction temperature and methanol flow rate) and the addition of chemical chaperones (specifically the DMSO) were successfully reported for the optimization of *P. pastoris* bioprocesses using ANN models. The experimental strategy developed in this work shows that the manipulation of fermentation conditions coupled with the addition of specific molecules can open new perspectives in the optimization of *Pichia pastoris* bioprocesses for recombinant membrane protein biosynthesis.

Methods

Materials, strains and media

The easy select expression kit for expression of recombinant proteins using pPICZα in *P. pastoris* and zeocin were obtained from Invitrogen (Carlsbad, CA, USA). Bis-(1,3-dibutylbarbituric acid) trimethine oxonol was acquired from Molecular Probes® (Part of Life

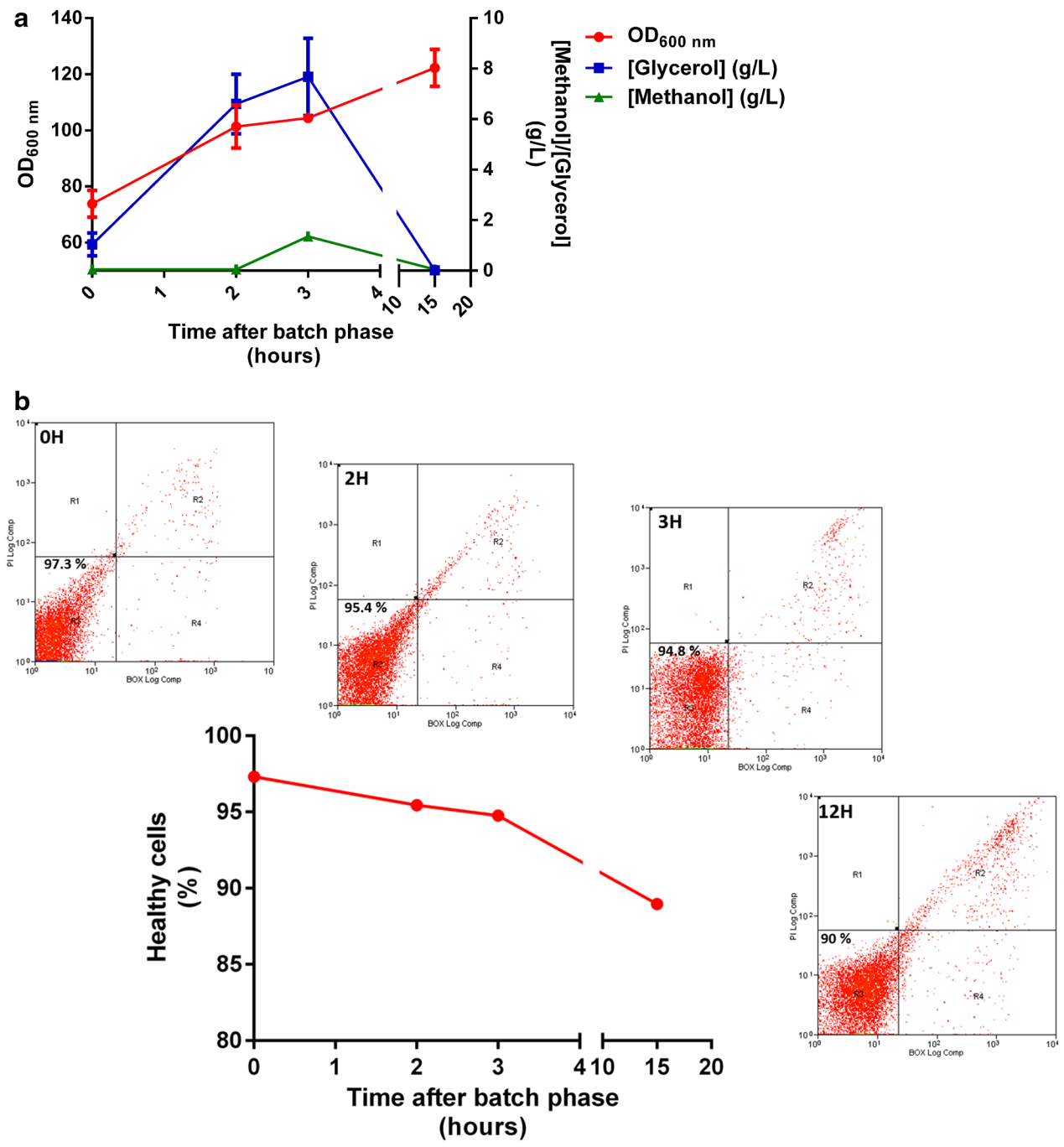


Fig. 6 Time course analysis of biomass levels, carbon sources concentrations and number of healthy *P. pastoris* cells at different stages of the optimal point estimated by the ANN model [30°C, 2.9 mL/L/H methanol constant flow-rate and 6% (v/v) DMSO]. **a** Biomass levels measured by spectrophotometric determination at 600 nm and methanol and glycerol levels measurements by HPLC with RID; (each value represents the mean of three independent samples). **b** Dot plots of green fluorescence of cells (BOX, x-axis) plotted against red fluorescence (PI, y-axis) obtained with cell samples taken at different stages of the optimum point retrieved from the ANN modelling. Three main subpopulations of cells can be distinguished corresponding to: healthy cells, no staining; cells with depolarized membranes, stained with BOX; and cells with permeabilized membranes, stained with PI. A total of 10,000 events were collected for these analysis. The variation on the percentage of healthy cells at different stages of the bioprocess is depicted in the *graph*. Each experiment was conducted in duplicate.

technologies; Carlsbad, CA, USA). Yeast nitrogen base (YNB), dithiothreitol, *S*-(5′-adenosyl)-L-methionine, epinephrine (bitartrate salt), deoxyribonuclease (DNase), protease inhibitor cocktail, DL-metanephrine hydrochloride, glass beads (500 µm) and propidium iodide were purchased from Sigma-Aldrich (St. Louis, MO, USA). All chemicals used were of analytical grade, commercially available, and used without further purification.

E. coli TOP10F' was used for DNA manipulations. *E. coli* transformants were selected on low-salt Luria–Bertani plates with 25 µg/mL Zeocin. *P. pastoris* X-33 and KM71H was used for fusion gene expression. YPD and YPDS media [40] were used for routine manipulation of *Pichia* cells. *P. pastoris* transformants were selected on YPDS plates with 200 µg/mL Zeocin. Small-scale fermentations were carried out in BMGH and BMMH media [40]. *P. pastoris* bioreactor cultures were carried out in modified basal salts medium (BSM) [27] with 200 µg/mL zeocin and supplemented with trace metal solution (SMT) [27].

Small-scale MBCOMT biosynthesis in *Pichia pastoris*

Easy select expression kit for expression of recombinant proteins using pPICZα in *P. pastoris* X33 cells (Invitrogen, Carlsbad, CA, USA) was used for the expression of human MBCOMT in its native form and the process was carried out according to manufacturer's instructions. Specifically, as the correct membrane protein targeting to the membrane is usually enhanced when secretion signals are used [41], the pPICZα expression vector was employed for expressing MBCOMT expression as it contains the alpha mating factor from *Saccharomyces cerevisiae*. For more details about the construction of the expression vector, please refer to Additional file 2. Subsequently, before conducting the initial trials for MBCOMT biosynthesis at a small-scale, the recombinant plasmid was sequenced in order to confirm the presence of the full sequence of the MBCOMT protein. In fact, after the analysis of the obtained results (Please refer to Additional file 3) concerning the sequencing analysis, it was possible to conclude that the recombinant plasmid contains the full sequence of the MBCOMT protein.

Recombinant hMBCOMT biosynthesis at a small-scale was carried out according to the following protocol [21]: cells containing the expression construct were grown at 30°C in YPD plates. A single colony was inoculated in 50 mL of BMGH medium in 250 mL shake flasks. Cells were grown at 30°C and 250 rpm overnight when the OD_{600} typically reached 6.0. Subsequently, since the inoculation volume was fixed to achieve an initial OD_{600} of 1, an aliquot of the fermentation in the medium BMGH was collected and centrifuged at room temperature during 5 min. After centrifuging the cells and ensuring that all glycerol was removed, the cells were resuspended in the induction medium and added to 500 mL shake-flasks to a total volume of 100 mL. The fermentations were carried out during 120 h at 30°C and 250 rpm, the cells were harvested by centrifugation (1,500×*g*, 10 min, 4°C) and stored frozen at −20°C until use.

Fed-batch *Pichia pastoris* bioreactor cultures

A single colony was used to inoculate a 100 mL BMGH seed culture in 500 mL shake-flasks and it was grown overnight at 250 rpm and 30°C. This culture was grown to an OD_{600} of 6 and used to inoculate 250 mL of modified basal salts medium (BSM) [26] containing 4.35 mL/L SMT [27] and 200 µg/mL zeocin in a 0.75 L (total volume) bioreactor (Infors HT, Switzerland). The bioreactors were operated with strictly controlled parameters including pH, temperature, airflow, agitation and dissolved oxygen. The pH was set at 4.7 and the DO set point was 20%. The temperature was 28°C in the batch phase while the pH was set at 4.7 during the entire assay and maintained by the addition of 12.5% (v/v) ammonium hydroxide and 0.75 M sulfuric acid. Foaming was controlled manually by the addition of the antifoam agent antifoam A (Sigma-Aldrich, St. Louis, MO, USA). The dissolved oxygen concentration was maintained at 20% by automatic adjustment of the airflow (maximum gas flow-rate used was 2 vvm) and the agitation rate (maximum agitation rate was fixed in 950 rpm). Preliminary trials were carried out in order to determine the best strategy for the biosynthesis of MBCOMT from *P. pastoris*. Therefore, unless otherwise stated, the optimized strategy (see Fig. 3) consisted of a glycerol batch phase that was carried out at 28°C until all glycerol had been consumed, indicated by a DO spike to 45%. Then, a glycerol fed-batch phase was initiated with a constant feed rate of 18.54 mL/L of 50% (v/v) glycerol containing 12 mL/L of SMT during 3 h. After 2 h elapsed, a transition phase was initiated through the addition of a 100% methanol at a constant feed rate, the temperature was changed for the induction temperature and the DMSO was added to the reaction vessel. The constant methanol feed rate, the temperature and the DMSO concentration were defined according to the experimental design. Then, after 3 h elapsed, the induction phase was maintained during additional 12 h using methanol as sole carbon and energy source. The whole system was controlled by IRIS software (Infors HT, Switzerland) and, in particular, the addition of feed medium was achieved using peristaltic pumps that were automatically controlled through a feeding profile previously programmed.

MBCOMT recuperation

Cell suspensions were broken down using a lysis buffer (150 mM NaCl, 10 mM DTT, 50 mM Tris, 1 mM $MgCl_2$,

pH 8.0) and freshly made protease inhibitors (1 mM PMSF, 5.0 μg/mL leupeptin and 0.7 μg/mL pepstatin A) containing an equal volume of acid-washed glass beads (500 μm, Sigma-Aldrich, St. Louis, MO, USA). The mixture was vortexed seven times for 1 min with an interval of 1 min on ice and centrifuged at 500*g* (4°C) for 5 min to remove cell debris and glass beads. Finally, the supernatant was collected, DNase (Sigma-Aldrich, St. Louis, MO, USA) (1 mg/mL) was added and the MBCOMT specific activity was determined (see "Determination of copy number by qPCR" for details).

Experimental design

A CCD with three levels and three factors was employed for the experimental design. The factors and levels for the optimization of MBCOMT specific activity were conditions associated with the fermentation process, namely, the temperature (20, 25 and 30°C), the 100% (v/v) methanol constant feed rate (1, 2 and 3 mL/h/L of culture) and the DMSO concentration [4, 5 and 6% (v/v)]. Table 2 lists the fermentation conditions parameters used in the experimental design and in model development and optimization by ANN.

Artificial neural network

A feed-forward artificial neural network was applied to predict the MBCOMT specific activity as function of the fermentation conditions (temperature, methanol constant flow-rate and DMSO concentration). The ANN models were implemented in MATLAB™ using the Neural Network Toolbox. The ANN structure included an input layer with three neurons (one for each input variables), an output layer with one neuron (MBCOMT specific activity) and one hidden layer with two neurons (3/2/1). Therefore, the resulting model contains a total of 11 parameters. The transfer functions of the input and output layers, the mathematically representation of the output function and the ANN structure were described elsewhere [18]. The ANN structure was built using the "newff" function. ANN was trained with the Levenberg–Marquardt back-propagation function, up to 1,000 epochs, using the "train" function. The learning rate and the ratio to increase learning rate were set at 0.01 and 1.05, respectively.

Flow cytometry assays

Cellular viability was assessed during the fermentation runs. Samples were collected at specific periods and analyzed by flow cytometry following the protocol described by Hyka and co-authors [39]. Briefly, the samples OD_{600} was measured, a dilution with PBS buffer was prepared to obtain a final OD_{600} of 0.1 and appropriated volumes of PI and BOX were added in order to attain final concentrations of 10 and 2 mg/L, respectively. The samples were incubated for 15 min at room temperature in the dark, centrifuged for 10 min at 1,500 rpm, resuspended in PBS and sonicated in the "hotspot" during 1 min. The samples were analyzed on a BD Biosciences FACSCalibur (Becton–Dickinson GmbH, Heidelberg, Germany), acquisition was performed with CellQuest™ Pro Software Light scatter measurements and fluorescence was collected in two optical channels, FL1 (515–545 nm, BOX) and FL4 (>670 nm, PI). Threshold was set on SSC to exclude noise, other particles and debris while sample acquisition was operated at flow rate of no more than 300 events per second and a total of 10,000 cells were gated and analyzed in each sample. Data analysis was performed using FCS Express Version 3 Research Edition (De Novo Software™, Los Angeles, CA, USA). The samples were incubated 30 min at 70°C to provide positive staining controls, thereby allowing detection of dead cells and were incubated 2 min at 60°C in order to provide the identification of three subpopulations.

HPLC analytical methods

The methylating efficiency of recombinant MBCOMT was evaluated by measuring the amount of metanephrine using epinephrine as substrate and as previously described [42]. Briefly, the MBCOMT lysates were incubated at 37°C for 15 min, using epinephrine as substrate and the reaction was stopped with 2 M of perchloric acid. Then, after processing the samples [42], the metanephrine levels in the samples were determined using HPLC with electrochemical detection in a coulometric mode, as previously described [43]. On the other hand, the levels of glycerol and methanol in the culture broth were quantified using a HPLC coupled to a 1260 Infinity Refractive Index Detector (Agilent, Santa Clara, CA, USA), according to what was previously described [21]. The chromatographic separation was achieved on a cation-exchange analytical column Agilent Hi-Plex H (300 × 7.7 mm i. d.; 8 μm) and the analysis was performed at 65°C with a flow rate of 0.6 mL/min using isocratic elution with 0.005 M H_2SO_4. The samples were centrifuged at 6,000 rpm for 10 min and the supernatant was filtered prior the injection through a 0.22 μm cellulose-acetate filter.

Determination of copy number by qPCR

The recombinant gene dosage present in the plasmid pPICZα-hMBCOMT introduced into the strains X33 and KM71H was determined according to the method reported by Nordén and collaborators [23]. Initially, gDNA was extracted from untransformed colonies of X33 and KM71H *P. pastoris* strains as well as from the X33 and KM71H transformants using the Wizard SV Genomic DNA Purification System (Promega, Madison,

USA) supplemented with zymolyase. Briefly, for internal standardization, a primer pair—PpAOX2_Prom_FW and PpAOX2_Prom_RV (5′-GACTCTGATGAGGGGCACAT-3′ and 5′-TTGGAAACTCCCAACTGTCC-3′, respectively)—was used that amplifies a stretch of the AOX2 promoter sequence, which is present as one copy in *P. pastoris* genome [23]. Then, in order to determine the number of recombinant gene sequences, it was designed another primer pair—PpAOX1_TT_FW and Pp_AOX1_TT_RV (5′-TGGGCACTTACGAGAAGACC-3′ and 5′-GCAAATGGCATTCTGACATC-3′, respectively)—that is directed towards the 3′TT sequence of the AOX1 gene, which is also present in the pPICZ and also in the pPICZ α plasmids and is integrated together with the gene of interest [23]. The mean efficiency (E) of the two primer pairs was determined according to the serial dilution method using gDNA extracted from both untransformed strains, starting from 100 ng. For each reaction, 10 ng of template were used and the thermal cycler was programmed to perform an initial incubation step at 95°C during 10 min and then 40 cycles of: 15 s at 95°C, 30 s at 60°C, 30 s at 72°C. According to what was previously described by Nordén and collaborators [23], the average copy number was calculated with the following equation:

$$\begin{aligned} Ravg &= \frac{E^{-\Delta\Delta Ct\ sample}}{E^{-\Delta\Delta Ct\ references}} \\ &= \frac{E^{-\Delta Ct\ sample}}{E^{-\Delta Ct\ references}} \\ &= \frac{E^{-(Ct\ A\ sample - Ct\ B\ sample)}}{E^{-(Ct\ A\ references - Ct\ B\ references)}} \end{aligned}$$

where Ravg is the average copy number, E the mean primer efficiency, Ct the critical take off cycle, sample the clone in study, reference the strain used (X33 or KM71H), A the AOX1-TT, B the AOX2 promoter. Finally, in order to obtain the MBCOMT copy number, the AOX1 TT copy number was subtracted by 1 to compensate for the endogenous AOX1 TT sequence.

Additional files

Additional file 1: ANN parameters employed for the final validation model.

Additional file 2: Detailed description of the construction of the expression vector pPICZα-hMBCOMT.

Additional file 3: Sequencing data of the recombinant expression vector pPICZα-hMBCOMT.

Abbreviations

ANN: artificial neural network; AOX: alcohol oxidase; BSM: basal salts medium; CCD: central composite design; COMT: catechol-*O*-methyltransferase; DO: dissolved oxygen; DMSO: dimethylsulfoxide; MBCOMT: membrane-bound catechol-*O*-methyltransferase; MP: membrane protein; OD_{600}: optical density 600 nm; PI: propidium iodide; *P. pastoris*: *Pichia pastoris*; SAM: *S*-adenosyl-L-methionine.

Authors' contributions

AQP carried out all the experimental procedures and wrote the manuscript. AQP, LMM, JAQ and LAP designed the study. LMM helped to perform the experimental procedures. JMLD carried out the ANN modeling. JMLD, MJB, JAQ and LAP contributed to drafting the manuscript. JAQ and LAP were, respectively, co-supervisor and supervisor of the project and were responsible for revising the manuscript. All authors read and approved the final manuscript.

Author details

[1] CICS-UBI, Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, Avenida Infante D. Henrique, 6201-001 Covilhã, Portugal. [2] Department of Biochemistry, Cambridge System Biology Centre, University of Cambridge, Sanger Building, 80 Tennis Court Road, Cambridge CB2 1GA, UK. [3] Departamento de Investigação e Desenvolvimento, Bial, 4745-457 São Mamede do Coronado, Portugal.

Acknowledgements

This research was supported by University of Beira Interior-Health Sciences Research Centre (CICS) and FCT (Portuguese Foundation for Sciences and Technology) by the project "EXPL/BBB478/BQB/0960/2012" and COMPETE: FCOMP-01-0124-FEDER-027563. A.Q. Pedro acknowledges a doctoral fellowship (SFRH/BD/81222/2011) from FCT and L. M. Martins a fellowship from the project PTDC/EBB-BIO/114320/2009. The authors also acknowledge the program COMPETE, the FCT project (Pest-C/SAU/UI0709/2011). Finally, the authors would like to acknowledge Filomena Silva and Carlos Gaspar for the valuable help with the flow cytometry assays.

Compliance with ethical guidelines

Competing interests

The authors declare that they have no competing interests.

Received: 11 February 2015 Accepted: 25 July 2015
Published online: 07 August 2015

References

1. Bawa Z, Bland CE, Bonander N, Nora N, Cartwright SP, Clare M et al (2011) Understanding the yeast host cell response to recombinant membrane protein production. Biochem Soc Trans 39(3):719–723
2. Bonifacio MJ, Palma PN, Almeida L, Soares-da-Silva P (2007) Catechol-*O*-methyltransferase and its inhibitors in Parkinson's diseaese. CNS Drug Rev 13(3):352–379
3. Zhu BT (2002) On the mechanism of homocysteine pathophysiology and pathogenesis: a unifying hypothesis. Histol Histopatol 17(4):1283–1291
4. Zhu BT, Liehr JG (1993) Inhibition of catechol-*O*-methyltransferase catalyzed *O*-methylation of 2- and 4-hydroxyestradiol by catecholamine: implications for the mechanism of estrogen-induced carcinogenesis. Arch Biochem Biophys 304(1):248–256
5. Mattanovich D, Sauer M, Gasser B (2014) Yeast biotechnology: teaching the old dog new tricks. Microb Cell Fact 13:34–39
6. Vogl T, Hartner FS, Glieder A (2013) New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. Curr Opin Biotechnol 24(6):1094–1101
7. Noseda DG, Blasco M, Recupero M, Galvagno MA (2014) Bioprocess and downstream optimization of recombinant bovine chymosin B in *Pichia* (*Komagataella*) *pastoris* under methanol-inducible AOX1 promoter. Protein Expr Purif. doi:10.1016/j.pep.2014.09.014
8. Darby RA, Cartwright SP, Dilworth MV, Bill RM (2012) Which yeast species shall I choose? *Saccharomyces cerevisiae* versus *Pichia pastoris* (review). Methods Mol Biol 866:11–23

9. Batra J, Beri D, Mishra S (2014) Response surface methodology based optimization of beta-glucosidase production from *Pichia pastoris*. Appl Biochem Biotechnol 172(1):380–393
10. Ramon A, Marin M (2011) Advances in the production of membrane proteins in *Pichia pastoris*. Biotechnol J 6(6):700–706
11. Shukla AK, Haase W, Reinhart C, Michel H (2007) Heterologous expression and comparative characterization of the human neuromedin U subtype II receptor using the methylotrophic yeast *Pichia pastoris* and mammalian cells. Int J Biochem Cell B 39:931–942
12. Asada H, Uemura T, Yurugi-Kobayashi T, Shiroishi M, Shimamura T, Tsujimoto H et al (2011) Evaluation of the *Pichia pastoris* expression system for the production of GPCRs for structural analysis. Microb Cell Fact 10:24–37
13. Andre N, Cherouati N, Prual C, Steffan T, Zeder-Lutz G, Magnin T et al (2006) Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. Protein Sci 15(5):1115–1126
14. Cos O, Ramon R, Montesinos JL, Valero F (2006) Operational strategies, monitoring and controlo f heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: a review. Microb Cell Fact 5:17–37
15. Afonso A, Pereira P, Queiroz JA, Sousa A, Sousa F (2014) Purification of Pre-miR 29 by a new *O*-phospho-L-tyrosine affinity chromatographic strategy optimized using design of experiments. J Chromatogr A 1343:119–127
16. Giordano PC, Martinez HD, Iglesias AA, Beccaria AJ, Goicochea HC (2010) Application of response surface methodology and artificial neural networks for optimization of recombinant Oryza sativa non-symbiotic hemoglobin 1 production by *Escherichia coli* in medium containing byproduct glycerol. Bioresour Technol 101(19):7537–7544
17. Abdel-Fattah YR, Soliman NA, Yousef SM, El-Helow ER (2012) Application of experimental designs to optimize medium composition for production of thermostable lipase/esterase by *Geobacillus thermodenitrificans* AZ1. Genet Eng Biotechnol J 10(2):193–200
18. Silva R, Ferreira S, Bonifacio MJ, Dias JM, Queiroz JA, Passarinha LA (2012) Optimization of fermentation conditions for the production of human soluble catechol-*O*-methyltransferase by *Escherichia coli* using artificial neural networks. J Biotechnol 160(3–4):161–168
19. Leonardi D, Salomon CJ, Lamas MC, Olivieri AC (2009) Development of novel formulations for Chagas disease: optimization of benznidazole chitosan microparticles based on artificial neural networks. Int J Pharm 367(1–2):140–147
20. Almeida AM, Queiroz JA, Sousa F, Sousa A (2015) Optimization of supercoiled HPV-16 E6/E7 plasmid DNA purification with arginine monolith using design of experiments. J Chromatogr B 978–979:145–150
21. Pedro AQ, Oppolzer D, Bonifacio MJ, Maia CJ, Queiroz JA, Passarinha LA (2015) Evaluation of Mut^S and Mut^+ *Pichia pastoris* strains for membrane-bound catechol-*O*-methyltransferase biosynthesis. Appl Biochem Biotechnol 175(8):3840–3855
22. Aw R, Polizzi KM (2013) Can too many copies spoil the broth? Microb Cell Fact 12:128–137
23. Nordén K, Agemark M, Danielson JAH, Alexandersson E, Kjelibom P, Johanson U (2011) Increasing gene dosage greatly enhances recombinant expression of aquaporins in *Pichia pastoris*. BMC Biotechnol 11:47–59
24. Pedro AQ, Bonifacio MJ, Queiroz JA, Maia CJ, Passarinha LA (2011) A novel prokaryotic expression system for biosynthesis of membrane-bound catechol-*O*-methyltransferase. J Biotechnol 156(2):141–146
25. Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F (2005) Combined effect of the methanol utilization (Mut) phenotype and gene dosage on recombinant protein production in *Pichia pastoris* fed-batch cultures. J Biotechnol 116:321–335
26. Maurer M, Kuhleitner M, Gasser B, Mattanovich D (2006) Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with *Pichia pastoris*. Microb Cell Fact 5:37–47
27. Ferrara MA, Severino BMN, Mansure JJ, Martins AS, Oliveira EMM, Siani AC et al (2006) Asparaginase production by a recombinant *Pichia pastoris* strain harbouring Saccharomyces cerevisiae ASP3 gene. Enzyme Microb Technol 39(7):1457–1463
28. Holmes WJ, Darby RA, Wilks MD, Smith R, Bill RM (2009) Developing a scalable model of recombinant protein yield from *Pichia pastoris*: the influence of culture conditions, biomass and induction regime. Microb Cell Fact 8:35–49
29. Pla IA, Damasceno LM, Vannelli T, Ritter G, Batt CA, Shuler ML (2006) Evaluation of Mut^+ and Mut^S *Pichia pastoris* phenotypes for high level extracellular scFv expression under feedback control of the methanol concentration. Biotechnol Prog 3:881–888
30. Xia H, Liu L, Reinhart C, Michel H (2008) Heterologous expression of human Neuromedin U receptor 1 and its subsequent solubilization and purification. Biochim Biophys Acta 1778:2203–2209
31. Salunkhe S, Soorapaneni S, Prasad KS, Raiker VA, Padmanabhan S (2010) Strategies to maximize expression of rightly processed human interferon α2b in *Pichia pastoris*. Protein Expr Purif 71:139–146
32. Wang Y, Wang Z, Du G, Hua Z, Liu L, Li J et al (2009) Enhancement of alkaline polygalacturonate lyase production in recombinant *Pichia pastoris* according to the ratio of methanol to cell concentration. Bioresour Technol 100:1343–1349
33. Anasontzis GE, Salazar Penã M, Spadiut O, Brumer H, Olsson L (2014) Effects of temperature and glycerol and methanol-feeding profiles on the production of recombinant galactose oxidase in *Pichia pastoris*. Biotechnol Prog 30(3):728–735
34. Murata Y, Watanabe T, Sato M, Momose Y, Nakahara T, Oka S et al (2003) Dimethyl sulfoxide exposure facilitates phospholipid biosynthesis and cellular membrane proliferation in yeast cells. J Biol Chem 278(35):33185–33193
35. Sanmartín-Suárez C, Soto-Otero R, Sánchez-Sellero I, Méndez-Álvarez E (2011) Antioxidant properties of dimethylsulfoxide and its viability as a solvent in the evaluation of neuroprotective antioxidants. J Pharm Toxicol Methods 63:209–215
36. Hong F, Meinander NQ, Jonsson LJ (2002) Fermentation strategies for improved heterologous expression of laccase in *Pichia pastoris*. Biotechnol Bioeng 79:438–449
37. Wu JM, Wang SY, Fu WC (2012) Lower temperature cultures enlarge the effects of *Vitreoscilla* hemoglobin expression on recombinant *Pichia pastoris*. Int J Mol Sci 13:13212–13226
38. Jahic M, Wallberg F, Bollok M, Garcia P, Enfors SO (2003) Temperature limited fed-bacth technique to control of proteolysis in *Pichia pastoris* bioreactor cultures. Microb Cell Fact 2:6–17
39. Hyka P, Zullig T, Ruth C, Looser V, Meier C, Klein J et al (2010) Combined use of fluorescent dyes and flow cytometry to quantify the physiological state of *Pichia pastoris* during the production of heterologous proteins in high-cell-density fed-batch cultures. Appl Environ Microbiol 76(13):4486–4496
40. *Pichia* Fermentation Process Guidelines, Version B 053002 (2002) Invitrogen Company (Carlsbad, CA, USA). http://www.invitrogen.com. Assessed 11 Jan 2015
41. Byrne B (2015) *Pichia pastoris* as an expression host for membrane protein structural biology. Curr Opin Struct Biol 32:9–17
42. Passarinha LA, Bonifacio MJ, Queiroz JA (2007) Comparative study on the interaction of recombinant human soluble catechol-*O*-methyltransferase on some hydrophobic adsorbents. Biomed Chromatogr 21(4):430–438
43. Pedro AQ, Soares RF, Oppolzer D, Santos FM, Rocha LA, Gonçalves AM et al (2014) An improved HPLC method for quantification of metanephrine with coulometric detection. J Chromatogrph Sep Tech 5(2):17–24

Additional Files

Additional file 1 - ANN parameters employed for the final validation model

4th Iteraction				
W_{hidden}	Temperature (ºC)	Methanol constant feed rate (mL/H/L)	DMSO [%(v/v)]	Bias
	30	2.9	6	-
Hidden Node (1)	5,768	-7,191	1,164	-0,100
Hidden Node (2)	4,138	-5,026	1,226	0,508

	Hidden Node (1)	Hidden Node (2)	Bias
W_{output}	-4,661	4,908	-0,712

Additional file 2 - Detailed description of the construction of the expression vector pPICZα-hMBCOMT

Briefly, the DNA fragment coding for MBCOMT was obtained from the pNCMO2_MBCOMT expression vector [23]. Previously constructed by our research group by PCR using specific primers for cloning (forward primer; 5' AACTCGAGAAAAGAATGCCGGAGGCCCCGCCT 3'; reverse primer, 5' AACTCGAGTCAGGGCCCTGCTTCGCTGCCTG 3'). PCR was conducted as follows: denaturation at 95ºC for 5 minutes, followed by 30 cycles at 95ºC for 30 seconds, 60ºC for 30 seconds and 72ºC for one minute, and a final elongation step at 72ºC for 5 minutes. The amplified DNA was purified by low melting agarose gel electrophoresis, digested with Xho I and cloned into the vector pPICZα (previously digested with Xho I) by T4 DNA ligase. This construct was transformed into *E. coli* TOP10F' cells, grown overnight at 37ºC in plates with low salt luria bertani-agar medium containing zeocin (25 μg/mL) and colonies were screened for the presence of the construct pPICZα-hMBCOMT. Therefore, some colonies were inoculated in 2.0 mL of low salt luria bertani medium and grown at 37ºC and 250 rpm overnight. From these cultures, highly purified plasmids were prepared using Wizard SV Plus SV Minipreps and were then subjected to DNA sequence analysis to confirm the identity of the amplicon, orientation and frame. This was confirmed to correspond to human MBCOMT gene [23]. Then, the target plasmid was introduced into freshly made *P. pastoris* X33 and KM71H competent cells by electroporation according to the manufacturer's instructions. After plating the resultant mixture in YPDS plates and following the incubation at 30ºC during 4 days, high level expression transformants were screened and further tested to confirm the methanol utilization phenotype. Finally, the stable occurrence of the expression cassette was verified in the colonies gDNA by PCR using AOX1 promoter and terminator specific primers (AOX1 5' GACTGGTTCCAATTGACAAGC 3' and AOX1 5' CAAATGGCATTCTGACATCC).

Additional file 3 - Sequencing data of the recombinant expression vector pPICZα-hMBCOMT

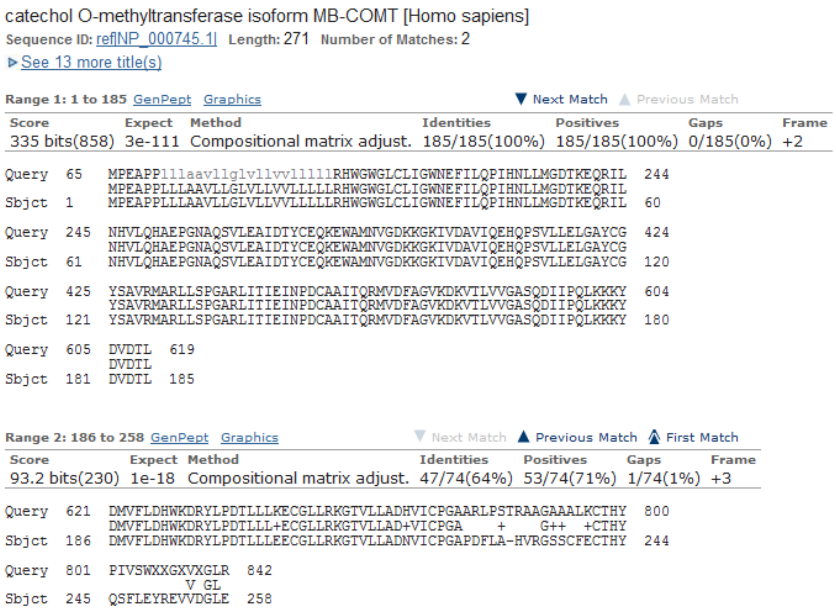
1. Recombinant plasmid sequencing using the 5' AOX1 forward sequencing primer (5'GACTGGTTCCAATTGACAAGC3')

a. Raw data obtained:

>A01_PN2+Primer_fw_27.ab1 1100 0 1100 ABI NNNNNNNCGACTTTTACGACACT
TGAGAGATCAAAAAACAACTAATTATTCGAAACGATGAGATTTCCTTCAATTTTTACTGCTGTTTTATTCGCAGC
ATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCAT
CGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATT
GTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAATGCCGGAGGC
CCCGCCTCTGCTGTTGGCAGCTGTGTTGCTGGGCCTGGTGCTGCTGGTGGTGCTGCTGCTGCTTCTGAGGCACTG
GGGCTGGGGCCTGTGCCTTATCGGCTGGAACGAGTTCATCCTGCAGCCCATCCACAACCTGCTCATGGGTGACAC
CAAGGAGCAGCGCATCCTGAACCACGTGCTGCAGCATGCGGAGCCCGGGAACGCACAGAGCGTGCTGGAGGCCAT
TGACACCTACTGCGAGCAGAAGGAGTGGGCCATGAACGTGGGCGACAAGAAAGGCAAGATCGTGGACGCCGTGAT
TCAGGAGCACCAGCCCTCCGTGCTGCTGGAGCTGGGGGCCTACTGTGGCTACTCAGCTGTGCGCATGGCCCGCCT
GCTGTCACCAGGGGCGAGGCTGATCACCATCGAGATCAACCCCGACTGTGCCGCCATCACCCAGCGGATGGTGGA
TTTCGCTGGCGTGAAGGACAAGGTCACCCTTGTGGTTGGAGCGTCCCAGGACATCATCCCCCAGCTGAAGAAGAA
GTATGATGTGGACACACTGGGACATGGTCTTCCTCGACCACTGGAAGGACCGGTACCTGCCGGACACGCTTCTCT
TGAAGGAATGTGGCCTGCTGCGGAAGGGGACAGTGCTACTGGCTGACCACGTGATCTGCCCAGGTGCCGCCAGAC
TTCCTAGCACACGTGCCGCGGGAGCAGCTGCTTTGAAGTGCACACACTACCCAATCGTTTCCTGGNATNNNGGAG
NGGTGNACGGCCTGAGANGCATCTNAA

Specifically, the Xho I restriction site is highlighted at yellow (CTCGAG) while the yeast endopeptidase kex2 recognition site (AAAAGA) and the initiation codon (ATG) for the hMBCOMT gene are highlighted, respectively, in red and green.

b. Basic local assignment search tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi) of the raw data:



catechol O-methyltransferase isoform MB-COMT [Homo sapiens]
Sequence ID: ref|NP_000745.1| Length: 271 Number of Matches: 2
▶ See 13 more title(s)

Range 1: 1 to 185 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
335 bits(858)	3e-111	Compositional matrix adjust.	185/185(100%)	185/185(100%)	0/185(0%)	+2

```
Query  65   MPEAPPlllaavllglvllvvllllRHWGWGLCLIGWNEFILQPIHNLLMGDTKEQRIL  244
            MPEAPPLLLAAVLLGLVLLVVLLLLLRHWGWGLCLIGWNEFILQPIHNLLMGDTKEQRIL
Sbjct  1    MPEAPPLLLAAVLLGLVLLVVLLLLLRHWGWGLCLIGWNEFILQPIHNLLMGDTKEQRIL  60

Query  245  NHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGKIVDAVIQEHQPSVLLELGAYCG  424
            NHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGKIVDAVIQEHQPSVLLELGAYCG
Sbjct  61   NHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGKIVDAVIQEHQPSVLLELGAYCG  120

Query  425  YSAVRMARLLSPGARLITIEINPDCAAITQRMVDFAGVKDKVTLVVGASQDIIPQLKKKY  604
            YSAVRMARLLSPGARLITIEINPDCAAITQRMVDFAGVKDKVTLVVGASQDIIPQLKKKY
Sbjct  121  YSAVRMARLLSPGARLITIEINPDCAAITQRMVDFAGVKDKVTLVVGASQDIIPQLKKKY  180

Query  605  DVDTL  619
            DVDTL
Sbjct  181  DVDTL  185
```

Range 2: 186 to 258 GenPept Graphics ▼ Next Match ▲ Previous Match ▲ First Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
93.2 bits(230)	1e-18	Compositional matrix adjust.	47/74(64%)	53/74(71%)	1/74(1%)	+3

```
Query  621  DMVFLDHWKDRYLPDTLLLKECGLLRKGTVLLADHVICPGAARLPSTRAAGAAALKCTHY  800
            DMVFLDHWKDRYLPDTLLL+ECGLLRKGTVLLAD+VICPGA    +    G++  +CTHY
Sbjct  186  DMVFLDHWKDRYLPDTLLLEECGLLRKGTVLLADNVICPGAPDFLA-HVRGSSCFECTHY  244

Query  801  PIVSWXXGXVXGLR  842
                    V GL
Sbjct  245  QSFLEYREVVDGLE  258
```

2. Recombinant plasmid sequencing using the 3' AOX1 reverse sequencing primer (5' GCAAATGGCATTCTGACATCC 3')

a. Raw data obtained:

>B02_PN2+Primer_rv_27.ab1 768 0 768 ABI ANNNNNNNNNANNCAGACCGGTCTT
CTCGTAAGTGCCCAACTTGAACTGAGGAACAGTCATGTCTAAGGCTACAAACTCAATGATGATGATGATGATGGT
CGACGGCGCTATTCAGATCCTCTTCTGAGATGAGTTTTTGTTCTAGAAAGCTGGCGGCCGCCGCGGCTCGAGTCA
GGGCCCTGCTTCGCTGCCTGGGCCCTTGTAGATGGCCTTCTCCAGGCCGTCCACCACCTCCCTGTATTCCAGGAA
CGATTGGTAGTGTGTGCACTCAAAGCAGCTGCTCCCGCGCACGTGTGCTAGGAAGTCTGGCGCACCTGGGCAGAT
CACGTTGTCAGCCAGTAGCACTGTCCCCTTCCGCAGCAGGCCACATTCCTCCAAGAGAAGCGTGTCCGGCAGGTA
CCGGTCCTTCCAGTGGTCGAGGAAGACCATGTCCAGTGTGTCCACATCATACTTCTTCTTCAGCTGGGGGATGAT
GTCCTGGGACGCTCCAACCACAAGGGTGACCTTGTCCTTCACGCCAGCGAAATCCACCATCCGCTGGGTGATGGC
GGCACAGTCGGGGTTGATCTCGATGGTGATCAGCCTCGCCCCTGGTGACAGCAGGCGGGCCATGCGCACAGCTGA
GTAGCCACAGTAGGCCCCCAGCTCCAGCAGCACGGAGGGCTGGTGCTCCTGAATCACGGCGTCCACGATCTTGCC
CTTTCTTGTCGCCCACGTTCATGGCCCACTCCTTCTGCTCCGCAGAATGTGTTCAATGGCCCTCCAGC

Specifically, the Xho I restriction site is highlighted at yellow (CTCGAG) while the stop codon is highlighted in blue.

b. Basic local assignment search tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi) of the raw data:

catechol O-methyltransferase isoform MB-COMT [Homo sapiens]
Sequence ID: ref|NP_000745.1| Length: 271 Number of Matches: 1
▶ See 13 more title(s)

Range 1: 95 to 271 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
363 bits(933)	5e-124	Compositional matrix adjust.	176/177(99%)	177/177(100%)	0/177(0%)	-3

```
Query  534  RKGKIVDAVIQEHQPSVLLELGAYCGYSAVRMARLLSPGARLITIEINPDCAAITQRMVD  355
            +KGKIVDAVIQEHQPSVLLELGAYCGYSAVRMARLLSPGARLITIEINPDCAAITQRMVD
Sbjct  95   KKGKIVDAVIQEHQPSVLLELGAYCGYSAVRMARLLSPGARLITIEINPDCAAITQRMVD  154

Query  354  FAGVKDKVTLVVGASQDIIPQLKKKYDVDTLDMVFLDHWKDRYLPDTLLLEECGLLRKGT  175
            FAGVKDKVTLVVGASQDIIPQLKKKYDVDTLDMVFLDHWKDRYLPDTLLLEECGLLRKGT
Sbjct  155  FAGVKDKVTLVVGASQDIIPQLKKKYDVDTLDMVFLDHWKDRYLPDTLLLEECGLLRKGT  214

Query  174  VLLADNVICPGAPDFLAHVRGSSCFECTHYQSFLEYREVVDGLEKAIYKGPGSEAGP  4
            VLLADNVICPGAPDFLAHVRGSSCFECTHYQSFLEYREVVDGLEKAIYKGPGSEAGP
Sbjct  215  VLLADNVICPGAPDFLAHVRGSSCFECTHYQSFLEYREVVDGLEKAIYKGPGSEAGP  271
```

3. Comments

The recombinant plasmid was sequenced using the AOX forward and reverse sequencing primers. In general, after performing the blast of the raw data obtained, it was observed that the sequence cloned in the recombinant plasmid matches with the nucleotides sequence that codes for human MBCOMT. Moreover, despite some mismatches were obtained at the end of each sequencing reaction, they are attributed to the sequencing reaction itself. Therefore, when the sequencing raw data obtained using the forward and the reverse primers, one can conclude that the cloned sequence obtained in this recombinant plamid codes the human MBCOMT protein.

Paper VI

Purification of membrane-bound catechol-*O*-methyltransferase by arginine-affinity chromatography

A. Q. Pedro, P. Pereira, M. J. Bonifácio, J. A. Queiroz, L. A. Passarinha

Chromatographia. 2015. 78 (21): 1339-1348

Short description: Following the establishment of appropriated strategies for MBCOMT biosynthesis and recovery from *P. pastoris* methanol-induced cultures (papers IV and V), here we describe an affinity methodology for the purification of the target enzyme from crude lysates. Indeed, this work reports the application of several immobilized amino acids as affinity ligands for MBCOMT purification using conditions that mainly promote either hydrophobic or electrostatic interactions. Preliminary results indicated that L-arginine-sepharose provided the adequate selectivity for MBCOMT purification, mostly using sodium chloride. Finally, after selecting the most appropriated stationary phase, the chromatographic methodology was refined through optimization of the buffers pH, the amount of sample injected and temperature.

Chromatographia
DOI 10.1007/s10337-015-2970-3

ORIGINAL

Purification of Membrane-Bound Catechol-*O*-Methyltransferase by Arginine-Affinity Chromatography

A. Q. Pedro[1] · P. Pereira[1] · M. J. Bonifácio[2] · J. A. Queiroz[1] · L. A. Passarinha[1]

Received: 16 April 2015 / Revised: 3 September 2015 / Accepted: 14 September 2015

Abstract Affinity chromatography strategies using amino acids as immobilized ligands have been successfully applied for the purification of different biomolecules from complex mixtures. Therefore, in this work, several supports with immobilized amino acids were applied for the purification of membrane-bound catechol-*O*-methyltransferase (MBCOMT) from *Pichia pastoris* lysates and it was verified that L-arginine provided the required selectivity for MBCOMT isolation. The optimization of the binding and elution buffers composition allowed the recovery of purified MBCOMT in a biological and immunologically active state from the arginine support. Additional optimization experiments varying the mobile phase pH, temperature and the concentration of the injected sample were carried out and an improvement of MBCOMT adsorption and purity was observed. Indeed, the optimized conditions for MBCOMT isolation and purification consisted in: loading of 4 mg of total protein onto the column previously equilibrated at 20 °C where the target enzyme was recovered in a purified fraction using 500 mM NaCl, 10 mM DTT and 0.5 % (v/v) Triton X-100 in 10 mM Tris buffer (pH 7) with a total bioactivity recovery of 24 ± 2.2 % and a purification fold of 4.95 ± 0.23, a value that is consistent with the best values ever reported for MBCOMT. Moreover, the L-arginine support demonstrated the ability to bind the target protein in a wide range of pH values (above and below the p*I* of the target protein) and the MBCOMT elution occurs in a single peak pattern. Finally, the strategy here reported can aid in the implementation of crystallization studies with MBCOMT in complex with clinically relevant inhibitors since it is obtained in a purified form with biological activity. In conclusion, a novel affinity chromatography strategy was developed and implemented for recombinant MBCOMT purification in a highly immunological and biologically active state.

Keywords Affinity chromatography · Arginine · Catechol-*O*-methyltransferase · Membrane protein · Protein purification

Abbreviations

BMGH	Buffered minimal glycerol
BMMH	Buffered minimal methanol
COMT	Catechol-*O*-methyltransferase
DTT	Dithiothreitol
MBCOMT	Membrane-bound catechol-*O*-methyltransferase
P. pastoris	*Pichia pastoris*
YNB	Yeast nitrogen base
YPD	Yeast extract peptone dextrose

Introduction

Catechol-*O*-methyltransferase (COMT, EC 2.1.1.6) is a magnesium-dependent enzyme that catalyzes the methylation of catechol substrates and exists as two molecular forms, a soluble and a membrane-bound (MBCOMT) isoform [1]. Furthermore, COMT has been implicated in

✉ L. A. Passarinha
lpassarinha@fcsaude.ubi.pt

[1] CICS-UBI-Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, 6201-001 Covilhã, Portugal

[2] Bial-Departamento de Investigação e Desenvolvimento, 4745-457 São Mamede do Coronado, Portugal

Published online: 05 October 2015

several human diseases [1–3] and the best documented is the important role that COMT plays in Parkinson's disease whose most effective treatment remains the dopamine replacement therapy with levodopa together with an inhibitor of aromatic amino acid decarboxylase and a COMT inhibitor [1]. In particular, our research group had previously published some studies concerning recombinant MBCOMT biosynthesis and kinetic characterization [4] as well as MBCOMT isolation and purification procedures [5, 6]. However, there is still room for improvement concerning the design of a chromatographic strategy for the purification of this target enzyme with higher yield and purity. Specifically, concerning the application of hydrophobic interaction chromatography for MBCOMT purification, several hydrophobic adsorbents (butyl, octyl and epoxy-sepharose) were tested using sodium phosphate to promote protein adsorption and an increasing detergent gradient until 1 % (v/v) of Triton X-100 to promote its elution [6]. Interestingly, from the ligands in study, octyl was found to be the most selective for MBCOMT purification over butyl and epoxy, probably due to its higher hydrophobicity and long chain length, despite MBCOMT is recovered under mild conditions [6]. In particular, it was observed a multiple peak pattern for MBCOMT elution that may be related to amino acids that are responsible for the interaction with the stationary phase [6]. Specifically, if the target protein binds to the matrix with the highly hydrophobic amino acids present in the MBCOMT membrane anchor region, higher detergent concentrations will be required for elution than other less hydrophobic MBCOMT regions [6]. Nevertheless, despite MBCOMT was isolated from a complex lysate, it was recovered without biological activity (unpublished data).

On the other hand, MBCOMT was also purified through anion-exchange chromatography using Q-sepharose and it was verified that the presence of 0.5 % triton X-100 led to an improvement in MBCOMT adsorption and isolation, probably because it increases MBCOMT solubility [5]. Thus, with 0.5 % Triton X-100 in the chromatographic buffers, MBCOMT was recovered with NaCl 300 mM with a purification factor of 4.3 [5].

In general, despite membrane proteins play crucial roles in many fundamental cell processes, there are few membrane proteins with known 3D structures [7] since its biosynthesis and purification are more challenging than soluble proteins due to the fact that membrane proteins are being manipulated outside their natural lipid environment [8]. Therefore, a promising and alternative solution to this problem may be the application of novel affinity chromatography strategies as the main step for membrane protein isolation and purification. Among different separation methods, affinity chromatography is a unique technique that allows the purification of proteins based on biological functions rather than individual physical or chemical properties [9]. The arginine-affinity chromatography has been regarded as a powerful technique with great applicability in the purification of several biomolecules [10–13]. In affinity chromatography, the separation is based on a reversible non-covalent interaction between the target biomolecule and its biospecific ligand [14], thus exploiting the principle of biomolecular recognition [11]. Moreover, this specific interaction that is occurring between the ligand and the target molecule can be the result of several interactions, namely electrostatic interactions, hydrophobic interactions, van der Waals forces and/or hydrogen bonding [10, 12]. In fact, the highly specific character of the aforementioned interactions constitute a major progress in affinity chromatography since, in a single step, offers many advantages when compared with other less selective and time-consuming multi-step procedures [14]. In particular, the development of affinity-based methodologies using amino acids as specific ligands was first introduced by Vijayalakshmi et al. (1989) for the purification of immunoglobulins and proteins [15] and has proved to be extremely successful. Specifically, arginine-affinity chromatography has been successfully applied for the purification of plasmid DNA [12, 14] RNA [11, 16] and proteins [17–19]. Accordingly, the use of amino acids as immobilized ligands for plasmid DNA purification was based on atomic studies from molecular modeling in which it was observed that the amino acids would preferentially interact with specific nucleic acid bases [14] while molecular recognition studies performed on RNA–protein interfaces yielded similar results for amino acids–RNA interactions [11]. On the other hand, in affinity chromatography with immobilized amino acids for proteins purification, the degree of interaction depends more specifically on the protein surface amino acids accessibility, where the three-dimensional structure can be extremely important, affecting the rate and extent of these interactions [10]. In this work and for the first time, arginine-affinity chromatography was applied for MBCOMT purification from *Pichia pastoris* (*P. pastoris*) lysates. Moreover, since MBCOMT biological activity is crucial for its function, aspects concerning the optimization of binding and elution buffers for maintaining MBCOMT biological activity during the chromatographic assay were addressed. Finally, the effect of mobile phase pH on MBCOMT adsorption considering its pI (isoelectric point) as well as the effect of the temperature and the concentration of the initial sample on MBCOMT adsorption and purity was thoroughly inspected.

Materials and Methods

Materials

Ultrapure reagent-grade water was obtained with a Mili-Q system (Milipore/Waters). The EasySelect *Pichia* expression kit for expression of recombinant proteins using pPICZ and pPICZα in *P. pastoris* and Zeocin were obtained from Invitrogen (Carlsbad, CA, USA). Triton X-100, Yeast Nitrogen Base (YNB), glucose, agar, yeast extract, NaCl, biotin, peptone, dithiothreitol (DTT), SAM, epinephrine (bitartrate salt), deoxyribonuclease, protease inhibitor cocktail, cysteine, sucrose, glycerol, DL-metanephrine hydrochloride, citric acid monohydrate and glass beads (500 µm) were purchased from SigmaChemical Co. (St. Louis, MO, USA). L-Histidine agarose and L-arginine-sepharose 4B gel were obtained from GE Healthcare Biosciences (Uppsala, Sweden) while L-methionine Agarose was obtained from SigmaChemical Co. (St. Louis, MO, USA). The NZY-colour Protein Marker II used for estimation of subunit molecular weight was purchased from NZYtech (Lisboa, Portugal). Anti-rabbit IgG alkaline phosphate secondary antibody and the silver nitrate solution kit were purchased from GE Healthcare Biosciences (Uppsala, Sweden) while monoclonal rabbit COMT antibody purchased by Abcam (Cambridge, England). All chemicals used were of analytical grade, commercially available, and used without further purification.

Recombinant hMBCOMT Biosynthesis and Recuperation

Recombinant hMBCOMT biosynthesis was carried out according to a previously developed protocol [20]. Briefly, *P. pastoris* cells transformed with the expression construct were grown for 72 h at 30 °C in yeast extract peptone dextrose (YPD) medium plates containing 200 µg/mL Zeocin. A single colony was used to inoculate 100 mL of Buffered Minimal Glycerol medium (BMGH) [100 mM potassium phosphate (pH 6.0), 1.34 % YNB, 4×10^{-5} % biotin and glycerol 1 %] in 250 mL shake-flasks. Cells were grown at 30 °C and 250 rpm to a cell density at 600 nm (OD_{600}) between 5 and 6. Subsequently, an aliquot was added to 100 mL of buffered minimal methanol (BMMH) medium (BMGH medium with methanol 0.5 % instead of glycerol) in 500 mL shake-flasks and the initial OD_{600} was fixed to 1.0 unit. After a 24 h growth at 30 °C and 250 rpm, cells were harvested by centrifugation (1500×*g*, 10 min, 4 °C). The cell pellet was resuspended in 2 mL of an appropriate buffer (150 mM NaCl, 50 mM Tris, 1 mM $MgCl_2$, pH 7.8), supplemented with protease inhibitor cocktail from Sigma-Chemical and disrupted by mechanical treatment with glass beads (seven cycles of 1 min with 1 min of interval on ice). The resultant supernatant was recovered after centrifugation (500×*g*, 5 min, 4 °C) and directly injected onto the column.

Affinity Chromatography

Affinity chromatography experiments were performed in an ÄKTA Avant system with UNICORN 6 software (GE Healthcare, Uppsala, Sweden). Unless otherwise stated, a 10 mm diameter × 20 mm long (about 2 mL) column was packed either with L-arginine-sepharose 4B gel, L-histidine agarose or L-methionine-agarose. All solutions were filtered through a 0.20 µm pore size membrane. Chromatographic runs were performed at 20 or 4 °C where the water-jacketed column was connected to a circulating water bath to maintain the appropriated temperature. Unless otherwise stated, the column was first equilibrated with 10 mM Tris buffer (pH 7.0), 10 mM DTT and 0.5 % (v/v) Triton X-100. The *P. pastoris* lysates (approximately 400 µg of total protein) were applied onto the column using a 200 µL loop at a flow rate of 0.5 mL/min. After the elution of unbound species, the ionic strength of the buffer was increased to 0.5 M of NaCl and then to 3 M NaCl in 10 mM Tris buffer (pH 7.0), 10 mM DTT and 0.5 % (v/v) Triton X-100. The absorbance of the eluate was continuously monitored at 280 nm. Fractions were pooled according to the chromatograms obtained and collected in tubes containing a stabilizing solution [final concentrations: sucrose 20 mM, glycerol 4 % (v/v) and DTT 5 mM], and were further concentrated and desalted with Vivaspin concentrators (10,000 MWCO).

MBCOMT Enzymatic Assay

The methylating efficiency of recombinant MBCOMT was evaluated by measuring the amount of metanephrine formed from epinephrine as previously described with minor modifications [21]. Briefly, the MBCOMT lysates and purified extracts were incubated at 37 °C for 15 min, using epinephrine as substrate and the reaction was stopped with 2 M of perchloric acid. Then, after processing the samples [21], the metanephrine levels in the samples were determined using HPLC with coulometric detection, as previously described [22].

SDS-PAGE and Dot-Blot Analysis

The purity of MBCOMT samples were analyzed by reducing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) according to the method of Laemmli [23] and as previously described [4]. Finally, immunologically active MBCOMT samples were identified after performing dot-blot analysis according to what was previously described [6].

Results and Discussion

Selection of the Amino Acid Stationary Phase for MBCOMT Isolation

In the present study, we wanted to evaluate the applicability of immobilized amino acids-affinity chromatography as the main step for MBCOMT purification from *P. pastoris* lysates. Preliminary assays were carried out to study the chromatographic behavior of these matrices interacting with MBCOMT according the salt type in the following conditions: a decreasing stepwise gradient of ammonium sulfate to promote mainly hydrophobic interactions and an increasing stepwise gradient of sodium chloride to favor mostly electrostatic interactions. The results obtained with sodium chloride are depicted in Fig. 1a and it is possible to observe that MBCOMT binding was only achieved using L-arginine as stationary phase. On the other hand and according to Fig. 1b, when the three columns were first equilibrated with 2 M ammonium sulfate, no binding was verified in any stationary phase. Nevertheless, the high concentrations of ammonium sulfate commonly used in these assays have a negative impact on COMT biological activity [24], a major drawback when developing suitable MBCOMT purification strategies. Therefore, after evaluating these three supports (L-histidine agarose, L-arginine-sepharose 4B gel and L-methionine agarose), we observed that L-arginine gel demonstrated high selectivity for the MBCOMT purification from *P. pastoris* lysates, mostly using an increasing gradient of sodium chloride. In fact, no interactions were established between any proteins from the *P. pastoris* lysates and L-histidine or L-methionine supports in the conditions used. Therefore, after these preliminary results, the L-arginine support was chosen for further assays in conditions that promote mainly electrostatic interactions.

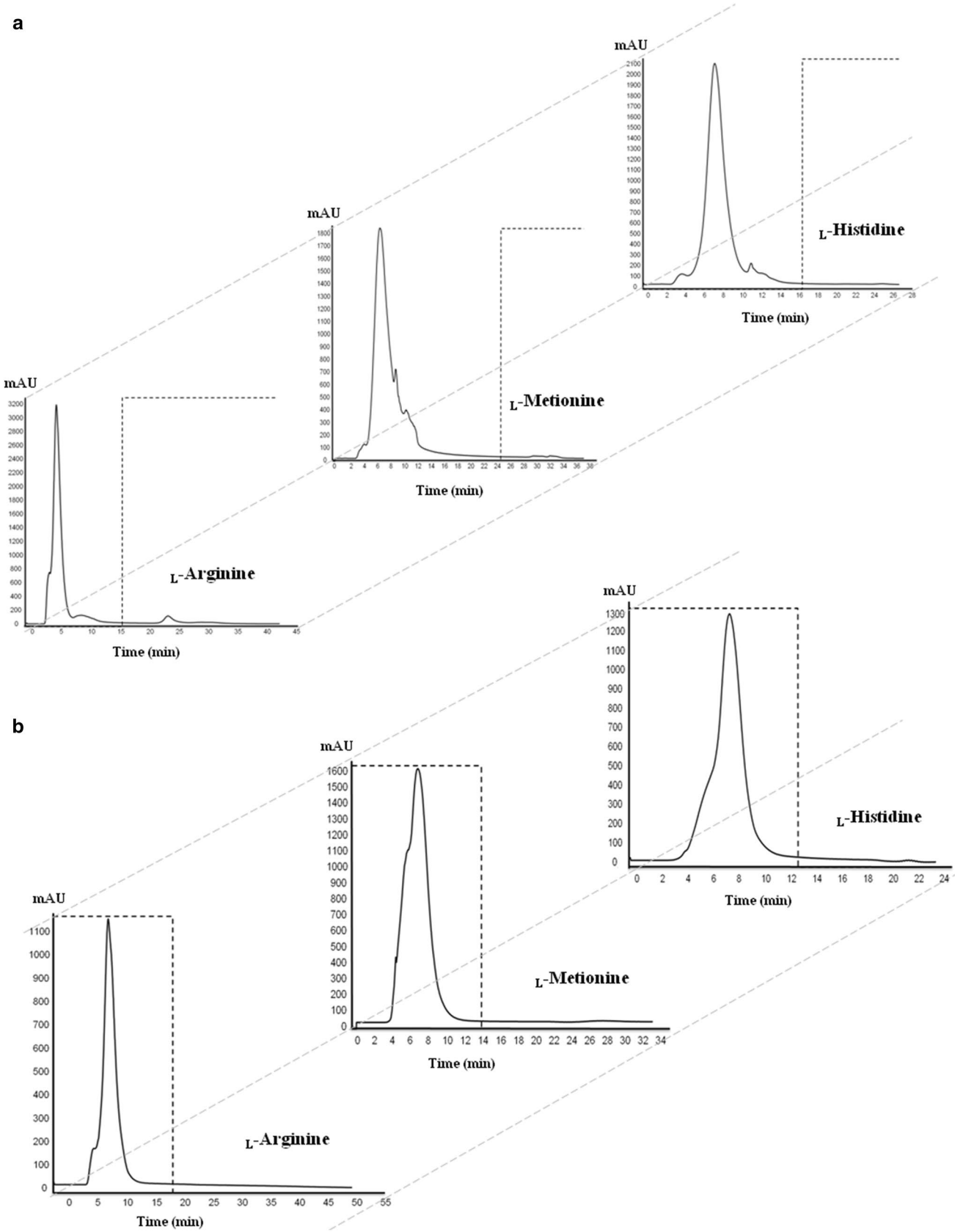
Effect of Binding and Elution Buffers Composition and pH on MBCOMT Biological Activity and Adsorption onto L-Arginine

Following the preliminary results, several binding/elution studies were performed through linear (from 0 to 1 M NaCl) and stepwise (from 0 to 3 M NaCl) gradients and it was determined that the stepwise elution is more suitable to obtain MBCOMT separation, once this strategy leads to a greater selectivity between the target biomolecule and contaminants. Therefore, after the injection of a sample containing 800 µg of total protein and binding with 10 mM Tris (pH 7) at 20 °C, MBCOMT elution was achieved, with the application of a stepwise gradient increasing the NaCl concentration to 500 mM and 3 M in 10 mM Tris (pH 7). In Fig. 2a, is depicted the typical chromatogram as well as the dot-blot and SDS-PAGE analysis of the several samples

Fig. 1 Preliminary chromatographic profiles of recombinant MBCOMT isolation from a *P. pastoris* lysate in L-arginine, L-histidine and L-methionine using different binding/elution conditions. **a** Elution performed at 0.5 mL/min by increasing NaCl stepwise gradient from 0 to 1 M in 10 mM tris buffer (pH8). **b** Elution performed at 0.5 mL/min by decreasing ammonium sulfate concentration from 2 to 0 M in 10 mM tris buffer (pH8)

obtained during the chromatographic run where it is possible to observe that immunologically active MBCOMT is mainly eluted with 500 mM NaCl with a substantial reduction of contaminants. Moreover, a regeneration step is performed with a stepwise gradient at 3 M NaCl to remove non-protein contaminants and assure that the binding capacity in the next trials is not compromised. Since it is crucial to recover MBCOMT in a biologically active state, through changes in binding/elution buffers composition, we attempted to improve the recovery of MBCOMT biological activity in the purified fractions. Therefore, the final optimized elution buffer composition was 3 M NaCl, 0.5 % (v/v) Triton X-100 and 10 mM DTT in 10 mM tris (pH 7) and it was verified that MBCOMT was recovered in a highly biologically active state while the balance between retained and unretained species was not significantly affected. In addition, a major bottleneck raised from the preliminary chromatographic studies intended to obtain COMT in a purified form identified this enzyme as highly unstable since it was described as being extremely labile, rapidly losing its biological activity during recovery and storage [6, 24, 25]. Therefore, in this work, the MBCOMT purified fractions were collected in a protein stabilizing solution comprising sucrose, DTT and glycerol that helped MBCOMT to maintain its biological activity [5]. Finally, the MBCOMT biological activity levels are depicted in Table 1 and, with this strategy, a purification fold of 1.69 ± 0.11 and a bioactivity recovery of 61.4 ± 3.8 % were obtained.

After this characterization of MBCOMT behavior in L-arginine, it was evaluated the pH effect on MBCOMT retention and stability. Indeed, as the arginine pKa is 12.1 [26] and the MBCOMT pI is 6.2 [5], at pH values higher than approximately 7.2, MBCOMT is negatively charged while at pH values lower than 5.2, MBCOMT is positively charged, despite the ligand at the pHs in study is always positively charged. Thus, several chromatographic experiments were designed to analyze how the pH ranging from 4 to 8 affects the retention of MBCOMT and other host proteins. The experiments at pHs 4, 5 and 6 were carried out with the buffer system sodium acetate-acetic acid in the conditions previously described. In fact, at pH 4 and 5, MBCOMT is not biologically active and at pHs 6, 7 and 8, the protein is recovered in a biologically active form, despite MBCOMT overall net charge at these pH is different. Based on these results, we consider that along with


a
mAU
L-Arginine
Time (min)
mAU
L-Metionine
Time (min)
mAU
L-Histidine
Time (min)
b
mAU
L-Arginine
Time (min)
mAU
L-Metionine
Time (min)
mAU
L-Histidine
Time (min)

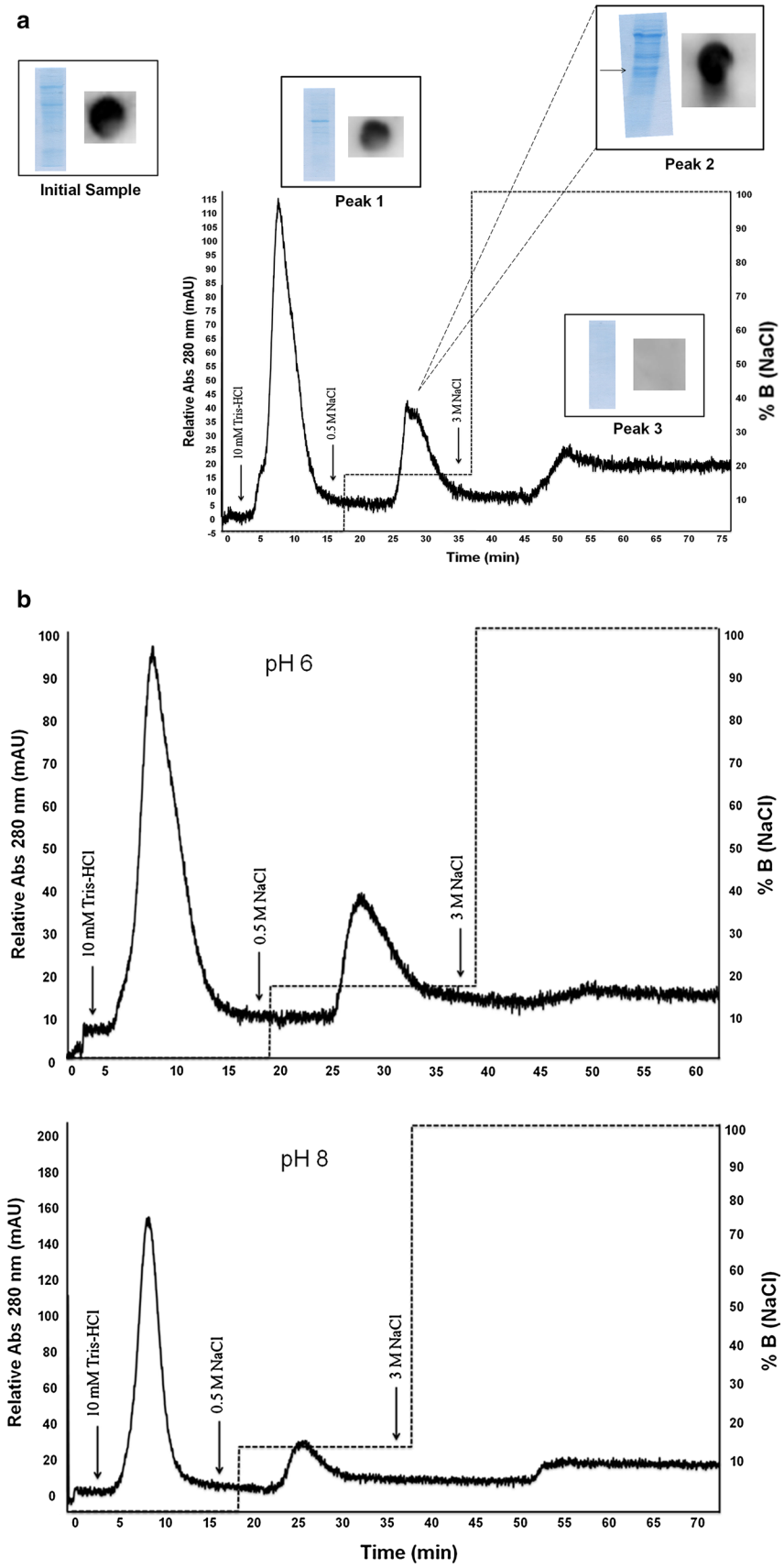


Fig. 2 Typical chromatographic profiles of recombinant MBCOMT isolation from a *P. pastoris* lysate by arginine-Sepharose chromatography at different pHs. **a** Elution was performed at 0.5 mL/min by increasing NaCl stepwise gradient from 0 M to 500 mM and, subsequently, to 3 M in 10 mM tris buffer (pH 7), 10 mM DTT and 0.5 % (v/v) Triton X-100, as represented by the *arrows*. For each stepwise gradient, the dot-blot and SDS-PAGE analysis of the different samples are represented. **b** Elution was performed at 0.5 mL/min by increasing NaCl stepwise gradient from 0 M to 500 mM and, subsequently, to 3 M in 10 mM tris buffer (pH 8) or sodium acetate buffer (pH 6), 10 mM DTT and 0.5 % (v/v) Triton X-100, as represented by the *arrows*

Table 1 Recombinant MBCOMT bioactivity recovery and purification fold obtained in this work for the arginine-sepharose matrix and comparison with previous results obtained for the target protein fraction using Q-sepharose

Chromatographic matrix	Fraction	Total activity (nmol/h)	Specific activity (nmol/h/mg of protein)	Purification fold	Bioactivity recovery (%)
L-Arginine-sepharose (2 mL gel) 20 °C, pH 7	Initial sample	67.3 ± 2.4	110.8 ± 3.6	1	100
	500 mM NaCl—800 µg initial sample	41.3 ± 3.2	187.3 ± 6.5	1.69 ± 0.11	61.4 ± 3.8
L-Arginine-sepharose (2 mL gel) 20 °C, pH 6	Initial sample	27.6 ± 2.3	40 ± 4.5	1	100
	500 mM NaCl—800 µg initial sample	16 ± 3	47.5 ± 5.7	1.19 ± 0.21	58 ± 4.5
L-Arginine-sepharose (2 mL gel) 20 °C, pH 8	Initial sample	50.0 ± 7.2	66.7 ± 8.9	1	100
	500 mM NaCl—800 µg initial sample	12.4 ± 4.2	65.5 ± 6.7	0.98 ± 0.08	24.8 ± 1.9
L-Arginine-sepharose (6 mL gel) 4 °C	Initial sample	1350.0 ± 46	337.7 ± 3.9	1	100
	500 mM NaCl—800 µg initial sample	783.0 ± 27	341.1 ± 9.7	1.01 ± 0.07	58 ± 4.5
L-Arginine-sepharose (6 mL gel) 4 °C	Initial sample	2219 ± 72	234.1 ± 24	1	100
	500 mM NaCl—4 mg initial sample	510.4 ± 33	533.7 ± 44	2.28 ± 0.12	23 ± 3.2
L-Arginine-sepharose (6 mL gel) 20 °C	Initial sample	604.2 ± 44	33.3 ± 4.2	1	100
	500 mM NaCl—4 mg initial sample	145.4 ± 22	164.9 ± 7.6	4.95 ± 0.23	24 ± 2.2
Q-Sepharose 20 °C [5]	300 mM NaCl with 0.5 % (v/v) TX-100 in buffers	181.0	331.0	4.3	91

electrostatic interactions, there are other non-covalent interactions involved in MBCOMT adsorption onto the matrix such as multiple non-covalent interactions, namely, hydrogen bonds, hydrogen π interactions and water mediated bonds, a fact that is characteristic from affinity chromatography. The representative chromatograms of thee assays at the pHs where MBCOMT was recovered biologically active are depicted in Fig. 2 and the values for the purification fold as well as for the bioactivity recoveries are depicted in Table 1. Indeed, we found that at pH 7, the highest purification factor was achieved (1.69 ± 0.11, 1.19 ± 0.21 and 0.98 ± 0.08 at pH 7, 6 and 8, respectively) since more contaminants were eluted with 10 mM Tris. Therefore, as MBCOMT is eluted with 500 mM NaCl in a more purified fraction at pH 7, we selected this pH for further studies. Moreover, as the buffers pH is increased from 6 to 8, along with a decrease in the purification factor, it was verified that the bioactivity recoveries are similar at pH 6 and 7 (58 % ± 4.5 and 61.4 % ± 3.8, respectively) while at pH 8 it rapidly decreases to 24.8 % ± 1.9. In fact, since COMT is regarded as being stable in this pH [10], we believe at pH 8 a considerable amount of target enzyme elutes with 10 mM tris, leading to a considerable decrease in the bioactivity recovery. In general, as it was suggested for arginine and particular oligonucleotide bases from plasmid DNA [14] and RNA [16], we believe that the affinity interaction promoted by the arginine support is responsible for the specific recognition of particular MBCOMT amino acid residues involving multiple non-covalent interactions.

Effect of Temperature and Concentration of the Injected Sample on the Purity of the MBCOMT Samples Recovered in L-Arginine

The optimization of the binding and elution buffers composition as well as the evaluation of the effect of pH on MBCOMT adsorption led to a strategy where a purification fold of 1.69 ± 0.11 was obtained, a value that is acceptable but could be improved. Therefore, since the previous experiments were carried out at 20 °C, the effect of changing the temperature to 4 °C on the MBCOMT biological activity and adsorption was evaluated, maintaining the same buffer composition and pH. Therefore, these experiments were performed with a sample containing 800 µg total protein and, as depicted in Table 1, it was verified that although the levels of bioactivity recovery were similar from those obtained at 20 °C (58 ± 4.5 % compared to 61.4 ± 3.8 % obtained at 20 °C), the purification fold was slightly lower (1.01 ± 0.07 compared to 1.69 ± 0.11). Unlike the hydrophobic interactions that are weakened with the decrease in the temperature [27], the electrostatic interactions are not affected by the temperature, notwithstanding it often affects the structure of a protein and, therefore, the interaction of the protein with the matrix. Therefore, the decrease

of the temperature led to a decrease in the purification fold since less target protein binded to the matrix, what can be explained by temperature-induced modifications on MBCOMT conformation, preventing it from binding to the L-arginine support. On the other hand, at higher temperatures, the mobile phase viscosity is reduced and thus improves the mass transfer [28] and since we are dealing with large molecules, the lower purification factor obtained at 4 °C can result from some limitations to mass transport.

Along with the temperature, the concentration of total protein in the initial lysate was also evaluated. Therefore, while in the aforementioned experiments, 800 µg of total protein were injected using 2 mL of L-arginine matrix, novel experiments were set up where 4 mg of total protein were injected in 6 mL of matrix. These experiments were carried out at 4 and 20 °C and the results are depicted in Table 1. In general, despite the bioactivity recoveries are lower in these conditions (23 ± 3.2 % at 4 °C and 24 ± 2.2 % at 20 °C), indicating that probably some target protein is eluting in the binding step, higher purifications fold were obtained (2.28 ± 0.12 at 4 °C and 4.95 ± 0.23 at 20 °C) since the fraction obtained with 500 mM NaCl were enriched in the target protein.

Comparison Between the MBCOMT Purification onto the L-Arginine Support and Previously Applied Chromatographic Techniques

The MBCOMT biological activity levels for the different strategies reported in this work as well as previously reports are shown in Table 1. In general, for the temperatures in study, MBCOMT adsorption over the host proteins increases with the temperature, thus increasing the purification fold (1.69 ± 0.11 at 20 °C and 1.01 ± 0.07 at 4 °C when the total protein in the initial sample was 800 µg) of the process. On the other hand, when the concentration of the initial sample was increased from 800 µg to 4 mg, the purifications fold of the process increased at both temperatures in study (see Table 1). In addition, the increase in the concentration of the initial sample leads to a decrease in the bioactivity recoveries of the process, lowering from 61.4 ± 3.8 to 24 ± 2.2 % at 20 °C. In conclusion, when the temperature of the experiment is 20 °C, we achieve a better purity in the target fractions and the bioactivity recoveries can be improved when the concentration of the initial sample is lower. On the other hand, when the purity of the target protein is a critical issue, the initial sample concentration should be higher since the purification fold is higher. The values obtained in this work are consistent with previous results obtained for MBCOMT purification using anion-exchange chromatography [5], as shown in Table 1. However, these values are excellent when compared to other strategies previously reported for soluble COMT purification using immobilized amino acids [10] or for MBCOMT purification through hydrophobic adsorbents [6] where none COMT biological activity was detected after the chromatographic process. Another advantage of this method over traditional strategies (hydrophobic interaction chromatography [6] or anion-exchange chromatography [5]) previously applied to MBCOMT purification is that MBCOMT elution occur in a single peak while in the aforementioned strategies, MBCOMT elution usually occur in a multiple peak pattern, leading to losses in the yield of the chromatographic process. Specifically, contrary to other chromatographic strategies previously reported [5, 6] where hydrophobic and electrostatic interactions are mainly involved in MBCOMT adsorption to the matrix, we believe that the affinity interaction promoted by the arginine support involves multiple non-covalent interactions, what can be greatly advantageous for the purification of other recombinant proteins. Finally, conducting crystallization studies of MBCOMT in complex with clinically and pharmacologically relevant inhibitors similarly to what was previously performed for SCOMT has proven to be valuable in the design of new molecules with improved COMT selectivity [29, 30]. However, the implementation of these studies relies on the development and implementation of efficient chromatographic strategies where batches of pure protein are required. Therefore, the strategy here reported can help to break the bottlenecks associated with the implementation of structural studies with MBCOMT since it can aid in obtaining the target protein with higher purity in a biologically active form.

Conclusions

Overall, a new affinity chromatographic process was described for the isolation of biologically active MBCOMT based on the specific molecular recognition between particular MBCOMT amino acid residues and the arginine support. The careful optimization of the binding and elution buffers composition carried out in this work allowed the recovery of MBCOMT in a biologically active state. In addition, the L-arginine support showed the ability to bind the target protein in a wide range of pH values (above and below the pI of the target protein), what can be greatly advantageous since the pH where the target protein is more stable can be chosen as the mobile phase pH without significant losses in the selectivity of the chromatographic process. The temperature effect on MBCOMT adsorption and purity was also evaluated and it was found that at 20 °C the purity of the isolated MBCOMT fractions was higher, probably because at 4 °C, temperature-induced modifications on MBCOMT conformation may be preventing MBCOMT adsorption to the matrix. In addition, as the mobile phase

viscosity is lower at 20 °C, the mass transfer is improved and since we are working with large molecules, the lower purification factor obtained at 4 °C can result from some limitations to mass transport. The concentration of the sample that is injected onto the column also seems to have a profound effect on MBCOMT adsorption and the purity of the isolated MBCOMT fractions increases with the increase in the concentration of the initial sample. Although the purification factor obtained with this strategy was similar from the best previously reported for MBCOMT isolation, its elution occurs in a single peak pattern, contrasting with the multiple peak pattern observed for hydrophobic interaction and anion-exchange chromatography that leads to losses in the yield of the target protein. Finally, we believe that the affinity interaction promoted by the L-arginine support and MBCOMT involves multiple non-covalent interactions, what can trigger new possibilities in the chromatographic isolation of other recombinant proteins in which low purity and low recovery yields are usually obtained with the traditional chromatographic strategies. In conclusion, a novel affinity chromatography strategy was developed and implemented for recombinant MBCOMT purification in a highly immunological and biologically active state.

Acknowledgments This research was supported by University of Beira Interior—Health Sciences Research Centre (CICS) and FCT (Portuguese Foundation for Sciences and Technology) by the project "EXPL/BBB478/BQB/0960/2012" and COMPETE: FCOMP-01-0124-FEDER-027563. A. Q. Pedro and P. Pereira acknowledge a doctoral fellowship (SFRH/BD/81222/2011 and SFRH/BD/81914/2011) from Fundação para a Ciência e Tecnologia. The authors also acknowledge the program COMPETE, the FCT project (Pest-C/SAU/UI0709/2011).

Compliance with Ethical Standards

The authors declare no commercial or financial conflict of interests. In this work, no research involving human participants or animals was carried out.

References

1. Bonifacio MJ, Palma PN, Almeida L, Soares-da-Silva P (2007) Catechol-*O*-methyltransferase and its inhibitors in Parkinson's disease. CNS Drug Rev 13:352–379
2. Karege F, Bovier P, Gaillard JM, Tissot R (1987) The decrease of erythrocyte catechol-*O*-methyltransferase activity in depressed patients and its diagnostic significance. Acta Psychiatr Scand 76:303–308
3. Yager JD, Liehr JG (1996) Molecular mechanisms of estrogen carcinogenesis. Annu Rev Pharmacol Toxicol 36:203–232
4. Pedro AQ, Bonifácio MJ, Queiroz JA, Maia CJ, Passarinha LA (2011) A novel prokaryotic expression system for biosynthesis of recombinant human membrane-bound catechol-*O*-methyltransferase. J Biotechnol 156:141–146
5. Correia FF, Santos FM, Pedro AQ, Bonifácio MJ, Queiroz JA, Passarinha LA (2014) Recovery of biological active catechol-*O*-methyltransferase isoforms from Q-sepharose. J Sep Sci 37:20–29
6. Santos FM, Pedro AQ, Soares RF, Martins R, Bonifácio MJ, Queiroz JA, Passarinha LA (2013) Performance of hydrophobic interaction ligands for human membrane-bound catechol-*O*-methyltransferase purification. J Sep Sci 36:1693–1702
7. Smith SM (2011) Strategies for the purification of membrane proteins. Methods Mol Biol 681:485–496
8. Seddon AM, Curnow P, Booth PJ (2004) Membrane proteins, lipids and detergents: not just a soap opera. Biochim Biophys Acta 1666:105–117
9. Zou H, Luo Q, Zhou D (2011) Affinity membrane chromatography for the analysis and purification of proteins. J Biochem Biophys Methods 49:199–240
10. Costa SR, Bonifácio MJ, Queiroz JA, Passarinha LA (2011) Analysis of hSCOMT adsorption in bioaffinity chromatography with immobilized amino acids: the influence of pH and ionic strength. J Chromatogr B Analyt Technol Biomed Life Sci 879:1704–1706
11. Martins R, Queiroz JA, Sousa F (2013) New approach in RNA quantification using arginine-affinity chromatography: potential application in eukaryotic and chemically synthesized RNA. Anal Bioanal Chem 405:8849–8858
12. Sousa F, Prazeres DMF, Queiroz JA (2009) Improvement of transfection efficiency by using supercoiled plasmid DNA purified with arginine affinity chromatography. J Gene Med 11:79–88
13. Sousa F, Prazeres DMF, Queiroz JA (2008) Affinity chromatography approaches to overcome the challenges of purifying plasmid DNA. Trends Biotechnol 26:518–525
14. Sousa A, Sousa F, Queiroz JA (2009) Selectivity of arginine chromatography in promoting different interactions using synthetic oligonucleotides as model. J Sep Sci 32:1665–1672
15. Vijayalakshmi MA (1996) Histidine ligand affinity chromatography. Mol Biotechnol 6:347–357
16. Pereira P, Sousa A, Queiroz J, Correia I, Figueiras A, Sousa F (2014) Purification of pre-miR-29 by arginine-affinity chromatography. J Chromatogr B Analyt Technol Biomed Life Sci 951:16–23
17. Patel A, O'Hara M, Callaway JE, Green D, Martin J, Nishikawa AH (1990) Affinity purification of tissue plasminogen activator using transition-state analogues. J Chromatogr 510:83–93
18. Suzuki T, Takahashi H (1974) Purification of prekallikrein with arginine-agarose. Methods Enzymol 34:432–435
19. Sakuragawa N, Takahashi K, Ashizawa T (1977) Isolation of prothrombin. Acta Med Biol 25:119–125
20. Pedro AQ, Oppolzer D, Bonifácio MJ, Maia CJ, Queiroz JA, Passarinha LA (2015) Evaluation of Mut^S and Mut^+ *Pichia pastoris* strains for membrane-bound and catechol-*O*-methyltransferase biosynthesis. Appl Biochem Biotechnol 175:3840–3855
21. Vieira-Coelho MA, Soares-da-Silva P (1999) Effects of tolcapone upon soluble and membrane-bound brain and liver catechol-*O*-methyltransferase. Brain Res 821(1):69–78
22. Pedro AQ, Soares RF, Oppolzer D, Santos FM, Rocha LA, Gonçalves AM, Bonifácio MJ, Queiroz JA, Gallardo E, Passarinha LA (2014) An improved HPLC method for quantification of metanephrine with coulometric detection. J Chromatograph Sep Tech 5:217
23. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685
24. Passarinha LA, Bonifacio MJ, Queiroz JA (2007) Comparative study on the interaction of recombinant human soluble catechol-*O*-methyltransferase on some hydrophobic adsorbents. Biomed Chromatogr 21:430–438
25. Cotton NJ, Stoddard B, Parson WW (2004) Oxidative inhibition of human soluble catechol-*O*-methyltransferase. J Biol Chem 279(22):23710–23718
26. Lee D, Lee J, Seok C (2013) What stabilizes close arginine pairing in proteins? Phys Chem Chem Phys 15:5844–5853

27. Queiroz JA, Tomaz CT, Cabral JMS (2001) Hydrophobic interaction chromatography of proteins. J Biotechnol 87:143–159
28. Ahuja S (2003) Chromatography and separation science. Academic Press, San Diego
29. Bonifacio MJ, Archer M, Rodrigues ML, Matias PM, Learmonth DA, Carrondo MA, Soares-Da-Silva (2002) Kinetics and crystal structure of catechol-*O*-methyltransferase complex with co-substrate and a novel inhibitor with potential therapeutic application. Mol Pharmacol 62(4):795–805
30. Rodrigues ML, Bonifacio MJ, Soares-Da-Silva P, Carrondo MA, Archer M (2005) Crystallization and preliminar X-ray diffraction studies if a catechol-*O*-methyltransferase/inhibitor complex. Acta Crystallogr Sect F Struct Biol Cryst Commun 61(1):118–120

Paper VII

Purification of histidine-tagged membrane-bound catechol-*O*-methyltransferase by immobilized-metal affinity chromatography

A. Q. Pedro, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha

(Ongoing work)

Short description: The arginine-affinity methodology reported in paper VI for MBCOMT purification was developed and implemented successfully, although MBCOMT was recovered with purity that is not compatible with the majority of biophysical techniques. Therefore, here we present some results on the purification of hexa-histidine tagged MBCOMT (MBCOMT-His_6) using IMAC. Initially, a screening of different detergents was carried out in order to evaluate which one most favors MBCOMT-His_6 extraction from *P. pastoris* membranes with higher biological activity. Then, several chromatographic buffers were tested in order to maximize MBCOMT binding onto the stationary phase in a biologically active form. Finally, the specificity of the stationary phase containing different immobilized ions to selectively bind MBCOMT-His_6 over other host contaminants was also evaluated.

Purification of histidine-tagged membrane-bound catechol-*O*-methyltransferase by Immobilized-metal affinity chromatography

Pedro A.Q.[1], Bonifacio M.J.[2], Queiroz J.A.[1], Passarinha L.A.[1]*

[1]*CICS-UBI – Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, 6201-001 Covilhã, Portugal*

[2]*Bial – Departamento de Investigação e Desenvolvimento, 4745-457 São Mamede do Coronado, Portugal*

*Corresponding author – E-mail: lpassarinha@fcsaude.ubi.pt - Phone: +351 275 329 069; Health Sciences Research Centre (CICS-UBI); Faculty of Health Sciences, University of Beira Interior, Av. Infante D. Henrique, 6200-506, Covilhã, Portugal

Abstract

Catechol-*O*-methyltransferase is an enzyme that exists as two molecular forms, a soluble and a membrane-bound isoform (hMBCOMT) and catalyzes the metabolism of catecholic substrates, including catecholestrogens and catecholamines. As this enzyme is involved in several human diseases, an improvement in the outcome of these diseases may come from structural studies between recombinant hMBCOMT in complex with inhibitor molecules or other ligands. However, these highly demanding techniques depend upon the achievement of highly quantities of purified and homogenous enzymes, thus highlighting the need of developing more accurate and selective methods for its biosynthesis and purification. In this work, human recombinant hMBCOMT was expressed in *Pichia pastoris* methanol-induced cultures using two different constructs, one with a hexa-histidine tag (hMBCOMT-His_6) and another with a linker for TEV protease prior the hexa-histidine tag. Despite immunologically active hMBCOMT was obtained from both constructs, only the hMBCOMT-His_6 was produced in a catalytically active form (27.3 ± 3.4 nmol/h/mg of protein) and thereby, this construct was chosen for further experiments. Therefore, after conducting a screening study with different detergents, we found that DDM is the most suitable detergent for extracting hMBCOMT-His_6 from *Pichia pastoris* membranes, owing to its favorable properties, namely its mildness due to the long alkyl chain length and the large micelles-forming capacity. First results obtained for hMBCOMT purification using Immobilized-Metal Affinity Chromatography showed that the majority of the target protein is lost during injection but it is minimized when the experiments were conducted in the hepes buffer system (Hepes 50 mM, NaCl 500 mM, $MgCl_2$ 1 mM, DDM 0.03% at pH 7.8) and with the ion nickel immobilized in the stationary phase where almost 50% of the target enzyme is recovered.

Keywords: Catechol-*O*-methyltransferase; Immobilized-metal affinity chromatography; Membrane protein; Purification; *Pichia pastoris*.

Abbreviations: $C_{12}E_8$ - Octaethylene glycol monododecyl ether; CMC - Critical Micellar Concentration COMT – catechol-*O*-methyltransferase; CYMAL-5 - 5-cyclohexyl-1-pentyl-β-$_D$-maltosyde; DDM - n-dodecyl-β-$_D$-maltopyranoside; DM - n-decyl-ß-$_D$-maltopyranoside; *E. coli* – *Escherichia coli*; FOS12 - FOS-choline 12; IMAC – Immobilized-Metal Affinity chromatography; LDAO - n-dodecyl-N,N-dimethylamine-N-oxide; hMBCOMT – human membrane-bound catechol-*O*-methyltransferase; MBCOMT-TEV-His_6 - human hexahistidine-tagged membrane-bound catechol-*O*-methyltransferase with a linker for TEV; hMBCOMT-His_6 - human hexahistidine-tagged membrane-bound catechol-*O*-methyltransferase; Mut – methanol utilization phenotype; MP – Membrane Protein; OG - octyl- β-$_D$-glucopyranoside; *P. pastoris* – *Pichia pastoris*; SAM – S-adenosyl-$_L$-methyonine; SDS-PAGE - Reducing sodium dodecyl sulphate-polyacrylamide gel electrophoresis; TEV - Tobacco Etch Virus.

1. Introduction

Catechol-*O*-methyltransferase (COMT; EC 2.1.1.6) is a magnesium-dependent enzyme that catalyzes the methylation reaction whereby a methyl group from S-adenosyl-$_L$-methionine (SAM) is transferred to one of the catecholic hydroxyls [Bonifacio *et al.*, 2007; Ma *et al.*, 2014]. As reaction products, are obtained the *O*-methylated catechol and S-adenosyl-*L*-homocysteine [Bonifacio *et al.*, 2007]. COMT exists as two molecular forms, a soluble and a membrane-bound (hMBCOMT) isoform, both encoded by the same gene located on chromosome 22 band q11.2 [Müller, 2015]. In humans, SCOMT is a 221 amino acid protein with 24.7 KDa while hMBCOMT has 30 KDa since it contains an additional stretch of 50 amino acid, the signal sequence for membrane anchoring [Bonifacio *et al.*, 2007; Ma *et al.*, 2014]. Typically, the substrates of COMT in mammals include catecholestrogens and their metabolites, catecholamines with hormonal and neurotransmission activities such as dopamine, norepinephrine, epinephrine, ascorbic acids, some indolic intermediates of melanin metabolism and xenobiotic catechols like carcinogenic catechol-containing flavonoids [Bonifacio *et al.*, 2007]. In most human and rat tissues, the levels of SCOMT greatly exceed the levels of hMBCOMT, except for the human brain where it only represents 30% of the total COMT [Bonifacio *et al.*, 2007]. Specifically, hMBCOMT is believed to be primarily involved in the termination of dopaminergic and noradrenergic synaptic neurotransmission when there are physiologically relevant low concentrations of catecholamines, indicating that this isoform is more important for catecholamine metabolism in vivo [Myöhänen and Männistö, 2010]. During the last decades, COMT has been implicated in diverse human diseases including certain types of cancer [Yager, 2015], cardiovascular diseases [Voutilainen *et al.*, 2007] or neurologic disorders (Parkinson's disease, schizophrenia and alzheimer's disease) [Apud and Weinberger, 2007; Bonifacio *et al.*, 2007; Serretti and Olgiati, 2012]. Specifically, the best documented is the important role that COMT plays in Parkinson's disease whose most effective treatment remains the dopamine replacement therapy with levodopa together with an inhibitor of aromatic amino acid decarboxylase and a COMT inhibitor [Bonifacio *et al.*, 2007]. An improvement in the outcome of the aforementioned diseases may come from structural studies between recombinant hMBCOMT in complex with inhibitor molecules or other ligands such as SAM or the magnesium ion. Typically, structure-based drug design for the development of new pharmacological molecules with improved selectivity and potency depend upon the attainment of highly purified protein quantities. Up until now, our research group reported diverse chromatographic strategies for hMBCOMT isolation and purification [Correia *et al.*, 2014; Santos *et al.*, 2013]. However, there is still room for improvement concerning the design of a chromatographic strategy for the purification of this enzyme with purity compatible with the majority of biophysical techniques for structural studies.

The isolation of stable and functional proteins constitutes a major bottleneck in the Membrane Protein (MP) production pipeline, since MP are embedded in the lipid bilayer of the

cell wall and have to be extracted from their native environment [Lantez *et al.*, 2015]. This is almost always accomplished using a detergent at a concentration above its Critical Micellar Concentration (CMC) [Lantez *et al.*, 2015]. Detergents interact with lipids and MP, thus leading to the dissolution of the membrane and solubilization of proteins. But each protein is a specific case, and the type of detergent greatly affect the success of solubilizing the target protein from its native environment in a biologically active form [Lantez *et al.*, 2015]. Initially proposed by Porath and coworkers [Porath *et al.*, 1975], Immobilized-Metal Affinity Chromatography (IMAC) is a separation technique based on the affinity of transition metals ions like zinc, copper, cobalt and nickel toward cysteine, histidine and tryptophan in aqueous solutions [Cheung *et al.*, 2012]. The metal ions are immobilized within a column through the use of chelating agents like iminodiacetic acid, nitriloacetic acid, carboxymethylated-aspartic acid and $_{L}$-glutamic acid [Hage *et al.*, 2012]. The affinity of a protein with the IMAC resin relies on the metal ions that take part in coordination where the affinity of Cu^{2+} and Ni^{2+} are usually the highest and the ion immobilized in the matrix is usually dictated by the biomolecular application [Cheung *et al.*, 2012]. For instance, trivalent ions are usually applied for the purification of phosphoproteins, divalent cations are commonly applied with histidine-tagged fusion proteins while Ca^{2+} was specifically applied for *biflorus* seed lectin [Cheung *et al.*, 2012]. Actually, IMAC is mainly applied from a bench to an industrial scale for the purification of proteins containing histidine residues, often introduced into a target protein as a N- or C- terminal peptide "tag" [Cheung *et al.*, 2012; Mooney *et al.*, 2014]. Over the years, several proteins with different topologies had been successfully purified using IMAC as the main chromatography step such as the human growth hormone [Mooney *et al.*, 2014], immunoglobulins [Boden *et al.*, 1995] or membrane-bound proteins such as the dehydrogenase/reductase SDR family member 7 [Skarka *et al.*, 2014]. In this work, recombinant hMBCOMT biosynthesis was carried out in *Pichia pastoris* (*P. pastoris*) with two constructs, one with a hexa-histidine tag in its carboxy-terminal and another with a linker for Tobacco Etch Virus (TEV) protease prior the histidine tag. Then, several detergents were screened in order to evaluate which detergent is more suitable for hMBCOMT extraction and solubilization from *P. pastoris* membranes with higher biological activity. Finally, hMBCOMT adsorption profiles in IMAC stationary phases charged with different ions and using different chromatographic buffers combinations were evaluated.

2. Materials and Methods

2.1. Materials

Ultrapure reagent-grade water was obtained with a Mili-Q system (Milipore/Waters). HisTrap™ FF crude was obtained from GE Healthcare Biosciences (Uppsala, Sweden). The easy select expression kit for expression of recombinant proteins using pPICZα in *P. pastoris* and zeocin were obtained from Invitrogen (Carlsbad, CA, USA). Yeast nitrogen base (YNB), glucose, agar, yeast extract, peptone, dithiotreitol, SAM, epinephrine (bitartrate salt), deoxyribonuclease (DNase), protease inhibitor cocktail, cysteine, sucrose, glycerol, DL-metanephrine hydrochloride, citric acid monohydrate, glass beads (500 µm), 3,5-dinitrocatechol and entacapone were purchased from SigmaChemical Co. (St. Louis, MO, USA). The NZYcolour Protein Marker II used for estimation of subunit molecular weight was purchased from NZYTech (Lisboa, Portugal). Anti-rabbit IgG alkaline phosphate secondary antibody and the silver nitrate solution kit were purchased from GE Healthcare Biosciences (Uppsala, Sweden), while monoclonal rabbit COMT antibody acquired by Abcam (Cambridge, England). All chemicals used were of analytical grade, commercially available, and applied without further purification.

2.2. Strains, plasmids and media

Escherichia coli (*E. coli*) TOP10F' was used for DNA manipulations. *E. coli* transformants were selected on low-salt luria-bertani plates with 25 µg/mL Zeocin. *P. pastoris* X-33 was used for fusion gene expression. The following media supplemented with Zeocin 200 µg/mL were employed in *Pichia* cells fermentations: YPD medium (1% yeast extract, 2% peptone and 2% glucose), YPDS medium (YPD medium supplemented with 1M Sorbitol), BMGH (100 mM potassium phosphate buffer at pH 6.0, 1.34% yeast nitrogen base, $4x10^{-4}$ g/L biotin and 1% glycerol) and BMMH (100 mM potassium phosphate buffer at pH 6.0, 1.34% yeast nitrogen base, $4x10^{-4}$ g/L biotin and 0.5% methanol). The *P. pastoris* transformants were selected on YPDS plates with 200 µg/mL Zeocin as a selective marker.

2.3. Construction of the expression vectors pICZαA-hMBCOMT-His_6 and pICZαA-hMBCOMT-TEV-His_6

Easy select expression kit for the biosynthesis of recombinant proteins using pPICZα in *P. pastoris* (Invitrogen, Carlsbad, CA) was used for the expression of human MBCOMT using two different constructs with the α-mating factor secretion signal; one construct with a hexahistidine in its carboxyl-terminal and a second construct with a linker for TEV protease followed by a hexahistidine in its carboxyl-terminal. This process was carried out according to manufacturer's instructions. Briefly, the DNA fragment coding for hMBCOMT was obtained from the pICZαA-hMBCOMT plasmid [Pedro *et al.*, 2015] by PCR using specific primers for cloning (a forward primer common to the two constructs - 5' AACTC GAGAAGAGAATGCCGGAGGCC CCGCCT 3'; and a reverse primer for hMBCOMT-His_6 - 5' AA CTC

GAG TCA GTG ATG GTG ATG GTG ATG GGG CCC TGC TTCGCTGCCTG 3'; and another reverse primer for hMBCOMT-TEV-His_6 5' AACTCGAGAAGAGAAT GCATCACCAT CACCATCACGAAA ACTTGTACTTCCAGGGTCCGGAGGCCCCGCCT 3') in which the reverse primers were designed in order to introduce a hexahistidine tag (hMBCOMT-His_6) and a linker for TEV protease before the hexahistidine tag (hMBCOMT-TEV-His_6), both in the carboxyl-terminal. The PCR conditions were as follows: denaturation at 95ºC for 5min, followed by 30 cycles at 95ºC for 30 s, 60ºC for 30 s and 72ºC for 1 min, and a final elongation step at 72ºC for 5min. The amplified DNAs was purified by low melting agarose gel electrophoresis, digested with Xho I and cloned into the vector pPICZα (previously digested with Xho I) by T4 DNA ligase. These constructs were transformed separately into *E. coli* TOP10F' cells, grown overnight at 37ºC in plates with low salt luria bertani-agar medium containing zeocin (25 μg/mL) and colonies were screened for the presence of the constructs pICZαA-hMBCOMT-His_6 and pICZαA-hMBCOMT-TEV-His_6. Therefore, some colonies were inoculated in 2.0 mL of low salt luria bertani medium and grown overnight at 37ºC and 250 rpm. From these cultures, highly purified plasmids were prepared using NzyMiniprep (Nzytech, Lisboa, Portugal) and were then subjected to DNA sequence analysis to confirm the identity of the amplicon, orientation and frame. Since the sequence was confirmed to correspond to human hMBCOMT gene plus the six histidines in its carboxyl-terminal or the human hMBCOMT gene plus the TEV protease linker and the six histidines [Tenhunen *et al.*, 1994] and were in frame with the α-mating factor secretion signal. The cloned plasmids were digested with Sac I and introduced into freshly made *P. pastoris* X33 competent cells by electroporation. After confirming that the X33 integrants presented a methanol phenotype plus (Mut^+), the stable occurrence of the expression cassette was verified in the colonies genomic DNA by PCR using AOX1 promoter and terminator specific primers (AOX1 5' GACTGGTTCCAATTGACAAGC 3' and AOX1 5' CAAATGGCATTCTGACATCC 3').

2.4. Recombinant hMBCOMT-His_6 and hMBCOMT-TEV-His_6 production and recuperation

Recombinant hMBCOMT biosynthesis was performed separately using *P. pastoris* X33 cells either containing the expression construct pICZα A-hMBCOMT-His_6 or pICZα A-hMBCOMT-TEV-His_6 according to the protocol previously reported [Pedro *et al.*, 2015]. Briefly, cells containing the expression construct were grown at 30ºC in YPD plates and then a single colony was inoculated in 100.0 mL of BMGH medium in 500.0 mL shake flasks. Subsequently, cells were grown at 30ºC and 250 rpm overnight when the cell density at 600 nm (OD_{600}) typically reached 6. Afterwards, since the inoculation volume was fixed to achieve an initial OD_{600} of 1, an aliquot of the BMGH fermentation medium was collected and centrifuged at room temperature during 5 minutes. After centrifuging the cells and ensure that all glycerol was removed, the pellet were ressuspended in the induction medium and added to 500.0 mL shake-flasks to a total volume of 100 mL. The fermentations were carried out during 120 hours at 30ºC and 250 rpm and were supplemented with methanol at a final concentration of 1% every 24 hours. Then, the cells were harvested by centrifugation (1500 x g, 10 min, 4ºC)

and stored frozen at –20ºC until use. Subsequently, *P. pastoris* lysis was accomplished through the application of a sequential procedure involving glass beads consisting of 7 cycles of vortexing during 1 minute with 1 minute of interval on ice. Therefore, cell suspensions were lysed in the chromatographic equilibrium buffer (please see section 2.5 – Immobilized-Metal Affinity Chromatography) containing protease inhibitor cocktail at a ratio of 1:2:2 (1 g cells, 2 mL lysis buffer and 2 g glass beads) and, after the lysis process were completed, the mixture was centrifuged (500g, 5 minutes, 4ºC) and the supernatant recovered.

2.5 Detergent screening for hMBCOMT-His_6 solubilization

The hMBCOMT solubilization studies were performed in the supernatant obtained after *P. pastoris* lysis and centrifugation at 500 g. The total protein concentration was adjusted to 5 mg/mL. Solubilization was achieved by mixing 900 µL of a *P. pastoris* lysate with 100 µL of a 10% detergent solution, giving a final detergent concentration of 1%. Different detergents were tested: n-dodecyl-β-$_D$-maltopyranoside (DDM), n-decyl-β-$_D$-maltopyranoside (DM), n-dodecyl-N,N-dimethylamine-N-oxide (LDAO), FOS-choline 12 (FOS12), Octaethylene glycol monododecyl ether ($C_{12}E_8$), 5-cyclohexyl-1-pentyl-β-$_D$-maltosyde (CYMAL-5) and n-octyl-β-$_D$-glucopyranoside (OG). After 3 hours incubation at 4ºC with mild agitation, the samples were centrifuged during 10 min either at 1500xg or 15000xg. Then, the supernatants were collected, the pellets ressuspended in the same solubilization buffer and the hMBCOMT specific activity (please see section 2.6 – hMBCOMT enzymatic assay – for further details) was determined in both fractions.

2.6. Immobilized-metal affinity chromatography

Chromatographic experiments were performed in an ÄKTA Avant system with UNICORN 6.1 software (GE Healthcare, Uppsala, Sweden). The chromatographic experiments were performed using the following stationary phase: HisTrap™ FF crude (5mL) packed with nickel or zinc ions (GE Healthcare, Uppsala, Sweden). The chromatographic runs were performed at 20ºC. All buffers pumped into the system were prepared with Mili-Q system water, filtered through a 0.20 µm pore size membrane (Schleicher Schuell, Dassell, Germany) and degassed ultrasonically. Preliminary experiments were conducted in order to assess which chromatographic buffers promote higher hMBCOMT retention onto the stationary phase and maintains enzyme stability. These experiments were conducted using different buffer-systems: 20 mM sodium phosphate, Tris 100 mM or Hepes 50 mM, all at pH 7.8 and supplemented with 1 mM $MgCl_2$, DDM 0.03% and NaCl 500 mM. On the other hand, for each experiment, the elution buffer composition was identical to the binding buffer but supplemented with imidazol 500 mM. Then, several experiments were performed in order to evaluate the specificity and selectivity of the stationary phase charged with different ions to the target enzyme. Therefore, the HisTrap™ FF crude (GE Healthcare, Uppsala, Sweden) was charged with nickel ($NiCl_2$ 0.2 M) or zinc ions ($ZnCl_2$ 0.2 M). All procedures including the charging of each stationary phase as well as the regeneration steps were carried out with

each manufacturer's instructions (https://www.thermofisher.com/order/catalog/product/K174001 assessed 21/09/2015 at 01:12 h GMT).

Unless otherwise stated, the column was initially equilibrated with 50 mM Hepes, 500 mM NaCl, and 1 mM $MgCl_2$ at pH 7.8. Aliquots of the lysis supernatant in equilibration buffer (500 µL) were applied onto the column at a flow-rate of 0.5 mL/min. After elution of the unretained species, imidazol concentration was increased by a stepwise mode to 500 mM for the HisTrap™ FF crude at 1mL/min. In all chromatographic runs, the conductivity, pH, pressure, temperature as well as absorbance at 280 nm were continuously monitored. Fractions were pooled according to the chromatograms obtained, collected in tubes containing a stabilizing solution (Final concentrations: 10 mM DTT, 140 mM of sucrose, 150 mM of cysteine and 10% of glycerol) [Correia *et al.*, 2014], concentrated and desalted with Vivaspin concentrators (10.000 MWCO) and conserved at 4ºC until further analysis. The protein content in samples were measured by Pierce BCA Protein Assay Kit (Thermo Scientific, USA) using Bovine Serum Albumin (BSA) as standard (25-2000 µg/mL) and calibration control samples according to the manufacturer's instructions.

Finally, each sample was subjected to the determination of hMBCOMT biological activity [Pedro *et al.*, 2014; Vieira-Coelho and Soares-da-Silva 1999]. Immunologically active hMBCOMT was detected using dot-blot [Santos *et al.*, 2013] and the purity of the samples detected by reducing sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

2.7. hMBCOMT enzymatic assay

The methylating efficiency of recombinant hMBCOMT was evaluated by measuring the amount of metanephrine formed from epinephrine as previously described with specific modifications [Pedro *et al.*, 2011; Vieira-Coelho and Soares-da-Silva 1999]. Briefly, the hMBCOMT lysates and purified extracts were incubated at 37ºC for 15 minutes, using epinephrine as substrate and the reaction was stopped with 2 M of perchloric acid. Then, after processing the samples [Vieira-Coelho and Soares-da-Silva 1999], the metanephrine levels in the samples were determined using a HPLC with coulometric detection, as previously described [Pedro *et al.*, 2014].

2.8. Electrophoresis, Western-blot and Dot-blot analysis

The purity of hMBCOMT samples were analyzed by SDS-PAGE according to the method of Laemmli [Laemmli, 1970] and as previously described [Pedro *et al.*, 2011]. Finally, immunologically active hMBCOMT samples were identified after performing western-blot analysis according to what was previously described [Pedro *et al.*, 2011].

3. Results and Discussion

Up until now, nine crystal structures of human recombinant SCOMT were deposited in the protein data bank but none corresponds to the membrane-bound isoform. Indeed to our best knowledge, the only membrane-bound methyltransferase crystal structure ever reported was isoprenylcysteine careboxyl methyltransferase from Methanosarcina acetivorans [Yang *et al.*, 2011].

The COMT enzyme has been implicated in diverse human diseases including certain types of cancer [Yager, 2015], cardiovascular diseases [Voutilainen *et al.*, 2007] or neurologic disorders (Parkinson's disease, schizophrenia) [Apud and Weinberger, 2007; Bonifacio *et al.*, 2007]. Specifically, the development of more selective and isoform-specific inhibitors can improve the outcome of several diseases, including schizophrenia, mostly because it is a promising avenue for treatment of cognitive deficits [Apud and Weinberger, 2007; Harrison *et al.*, 2015]. The majority of structure-based drug design depend upon the attainment of highly purified protein quantities. However, the isolation of stable and functional proteins still constitutes a major bottleneck in the MP production pipeline, since MP are embedded in the lipid bilayer of the cell wall and have to be extracted from their native environment [Lantez *et al.*, 2015]. Therefore, in order to circumvent these challenges, new integrated approaches must be developed where appropriated MP solubilization conditions need to be screened and highly selective chromatographic methodologies should be designed in order to obtain functional pure MP, such as hMBCOMT.

3.1. Histidine tagged-hMBCOMT biosynthesis in *Pichia Pastoris*

The hMBCOMT was recombinantly expressed in *P. pastoris* methanol-induced cultures using two different constructs, one encoding hMBCOMT and a hexahistidine tag and another that encodes hMBCOMT with a linker for TEV protease before the hexahistidine tag, both in the carboxyl terminal. After a typical fermentation run and the application of an appropriated lysis step using glass beads as previously described [Pedro *et al.*, 2015], a western-blot analysis was performed to verify the presence of hMBCOMT from both constructs. As depicted in Figure 1, immunologically active hMBCOMT was detected when it was expressed in *P. pastoris* with a hexa-histidine tag (Lane I, Figure 1) and with a linker for TEV protease prior the tag (Lane II, Figure 1). However, high molecular weight bands that may correspond to dimeric and trimeric MBCOMT forms were detected for both constructs, indicating that the tag may be causing aggregation of the target enzyme, contrary to what was previously reported for untagged MBCOMT [Pedro *et al.*, 2011; Pedro *et al.*, 2015].

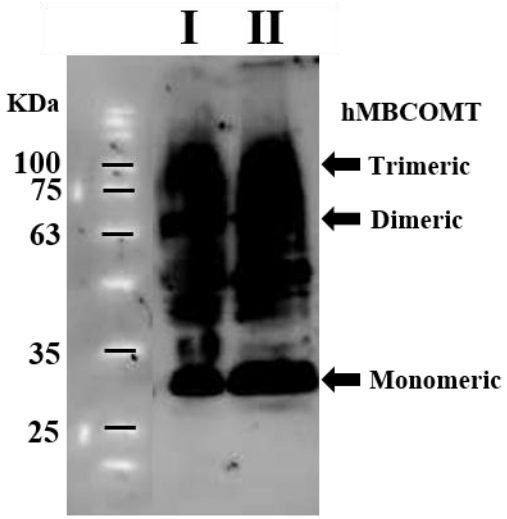


Figure 1 – Western-blot analysis of recombinant hMBCOMT-His_6 (Lane I) and hMBCOMT-TEV-His_6 (Lane II) obtained from typical methanol-induced *P. pastoris* cultures. Immunologically active hMBCOMT is indicated by the arrow.

On the other hand, only the construct hMBCOMT-His_6 was produced in a catalytically active form (27.3 ± 3.4 nmol/h/mg of protein). Specifically, the values obtained in this work for hMBCOMT-His_6 are lower than those previously reported for native MBCOMT using *Brevibacillus choshinensis* as the host (50.5 nmol/h/mg of protein) [Pedro *et al.*, 2011] or by *P. pastoris* X33 and KM71H strains (61.73 and 60.62 nmol/h/mg of protein, respectively) [Pedro *et al.*, 2015], demonstrating that the tag may slightly affect the overall folding of the target enzyme and, consequently, its biological activity. Indeed, the structure of the tag, its position, sequence, and length, can influence production of a protein on several levels: expression rate, accessibility for binding to the IMAC ligand, protein three-dimensional structure, protein crystal formation, and although to a minor extent solubility and activity [Block *et al.*, 2009]. Specifically, it was described that introduction of histidine tags at the carboxy-terminal may be detrimental to the target enzyme activity [Jones *et al.*, 1995], as it was previously demonstrated for $_L$-lactate dehydrogenase [Halliwell *et al.*, 2001].

The construct hMBCOMT-TEV-His_6 didn't show any biological activity, even after the cleavage of the tag with TEV protease. Indeed, as the tag is placed in the carboxy-terminal, cleavage by TEV protease will leave 3 or 4 amino acids behind the protease recognition sites [Jones *et al.*, 1995], what can still compromise MBCOMT biological activity. Therefore, based on results obtained, the construct hMBCOMT-His_6 was selected for further chromatographic experiments.

3.2. Detergent screening for hMBCOMT solubilization

Detergents are amphipatic molecules, consisting of a polar head group and a hydrophobic chain that can solubilize MP by mimicking the natural lipid bilayer environment normally inhabited by the protein [Seddon *et al.*, 2004]. Therefore, in this work, we performed a detergent screening in order to find the optimal solubilization detergent for the extraction and selective enrichment of biologically active hMBCOMT from *P. pastoris* Mut^+ lysates. Detergents are classified according to their structure and can be divided into four major classes: ionic which contain a head group with a net charge that can be either cationic or anionic and a hydrophobic chain or steroidal backbone; bile acid salts that are ionic but their backbone consists of rigid steroidal groups; non-ionic that contain uncharged hydrophilic head groups of either polyoxyethilene or glycosidic groups and zwitterionic that combine the properties of ionic and non-ionic detergents [Privé, 2007; Seddon *et al.*, 2004]. The type of detergents used in this work as well as structure characteristics are summarized in Table 1. In general, although affinity chromatography based on nickel chelates and histidine tags is usually not affected by detergents, the use of ionic detergents is not recommended [Arnold and Linke, 2008]. In addition, we previously reported that ionic detergents affect negatively hMBCOMT biological activity [Pedro *et al.*, 2011] and therefore, ionic detergents and bile acid salts were not considered in this study.

In general, all detergents were mixed to *P. pastoris* membranes to a final concentration of 1% and with a total protein concentration of 5 mg/mL. The yield, measured as hMBCOMT-His_6 specific activity (nmol/h/mg of protein), was defined as the main control parameter for the screening experiments. After the incubation period, the mixtures were centrifuged at 1500g or 15000g, the supernatant collected and defined as S1500g/S15000g while the pellets were ressuspended in the respective solubilization buffer and defined as P1500g/P15000g. A control experiment in which no detergent was added was also performed. Preliminary experiments were carried out to evaluate which detergents are useful to solubilize hMBCOMT-His_6 in a biologically active form. So, it was found that with LDAO, OG or FOS12 the enzyme doesn't retain any biological activity, but in a solubilization step performed with DDM, DM, CYMAL-5 and $C_{12}E_8$, hMBCOMT-His_6 was pre concentrated with activity (data not shown).

Table 1 – Type and general characteristics of the detergents applied for hMBCOMT-His_6 solubilization screening.

[1]Data obtained from GE Healthcare Biosciences (Uppsala, Sweden): http://www.gelifesciences.com/webapp/wcs/stores/servlet/catalog/en/GELifeSciences-pt/service-and-support/discovery-matters-archive/. Assessed 2 September 2015 15:15 H GMT. [2]Approximate critical micelle concentrations in water at 20ºC. [3]Data obtained from Anatrace (Dussel, USA): https://www.anatrace.com/. Assessed 2 September 2015 at 15:02 h GMT.

Detergent[1]	Common abbreviation[1]	Detergent type[1]	CMC (%)[1, 2]	Aggregation number (H_2O)[3]
n-dodecyl-β-$_D$-maltopyranoside	DDM	Non-ionic	0.009	78-149
n-decyl-β-$_D$-maltopyranoside	DM	Non-ionic	0.09	69
n-dodecyl-N,N-dimethylamine-N-oxide	LDAO	Zwitterionic	0.02	76
FOS-choline 12	FOS12	Zwitterionic	0.05	54
Octaethylene glycol monododecyl ether	$C_{12}E_8$	Non-ionic	0.005	90-120
5-cyclohexyl-1-pentyl-ß-$_D$-maltosyde	CYMAL-5	Non-ionic	0.12	47
octyl-β-$_D$-glucopyranoside	OG	Non-ionic	0.53	47

In general, despite the non-ionic detergents are usually applied for MP solubilization in a biologically active form, short chain non-ionic detergents such as the OG can often lead to deactivation of the protein [Seddon *et al.*, 2004], thus explaining why hMBCOMT-His_6 doesn't retain any biological activity when solubilized in OG. Likewise, hMBCOMT-His_6 wasn't solubilized in a biologically active form using either LDAO or FOS12, two zwitterionic detergents that are generally more deactivating than non-ionic detergents [Seddon *et al.*, 2004]. In fact, LDAO is a relatively harsh detergent and it was estimated that only 20% of MP are resistant to denaturing effects of this detergent [Privé, 2007].

Following the initial experiments, it was evaluated the performance of each detergent to extract biologically active hMBCOMT-His_6 from *P. pastoris* membranes, measured as the ratio between the target enzyme specific activity (nmol/h/mg of protein) in the supernatant by the sum of hMBCOMT-His_6 specific activity in the control (supernatant plus pellet), as demonstrated in Figure 2A.

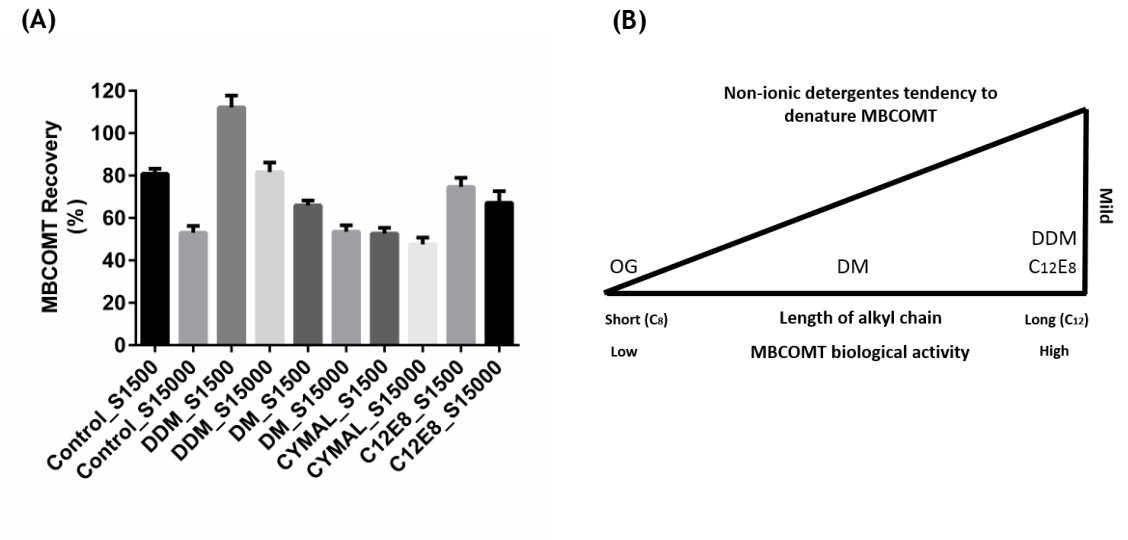


Figure 2 – (A) Performance of different detergents in extracting biologically active hMBCOMT-His_6 from *P. pastoris* membranes after a centrifugation step at 1500g or 15000g. The percentage of recovery is defined as the ratio between hMBCOMT specific activity (nmnol/h/mg of protein) for each detergent in the supernatant and the sum of the specific activity in the supernatant and the pellet of the control. **(B)** Tendency of different non-ionic detergents tested to denature hMBCOMT as function of the length of the alkyl chain.

Abbreviations- C12E8: Octaethylene glycol monododecyl ether; CYMAL-5: 5-cyclohexyl-1-pentyl-β-D-maltosyde; DDM: n-dodecyl-β-D-maltopyranoside; DM: n-decyl-β-D-maltopyranoside; hMBCOMT – human Membrane-bound catechol-*O*-methyltransferase; OG: octyl-β-D-glucopyranoside.

A common feature to all detergents in study was the fact that when a higher centrifugation velocity was applied (15000g), the percentage of hMBCOMT-His_6 recovery was lower in the supernatant, indicating that the target enzyme is lost in the pellet. According to Figure 2A, hMBCOMT extraction is greatly enhanced when the membranes are treated with DDM, following by $C_{12}E_8$, DM and CYMAL-5. According to Privé (2007), zwitterionic detergents that contain more charged groups tend to be harsh in solubilizing MP while mild detergents are the ones which contain large sizes of the head group and long length of the alkyl chain. In fact, we found that with exception of CYMAL-5 whose structure is more complex, the longer the length of the alkyl chain, greater the ability of each detergent to extract functional active hMBCOMT-His_6 from *P. Pastoris* lysates, as shown in Figure 2B. Unlike the short chain non-ionic detergents such as OG, MP are readily solubilized in other alkylglucosides such as DM or DDM in a functional active form [Seddon *et al.*, 2004]. Nevertheless, although DM properties are similar from those of DDM, it is more enzymatic deactivating than DDM (MBCOMT recoveries = 66 and 112%, respectively), probably due to shorter length of the alkyl chain (C_{10} and C_{12} respectively for DM and DDM) [Privé, 2007; Seddon, 2004]. DDM is one of the gentler detergents and has very favorable properties for maintaining the functionality of more-aggregation-prone MP in solution [Privé, 2007; Sonoda *et al.*, 2010], what can explain it

successful application in this work as hMBCOMT is thought to be highly unstable [Correia *et al.*, 2014].

On the other hand, $C_{12}E_8$ is much less denaturing than C_8E_4 or OG [Le Maire *et al.*, 2000; Privé, 2007] and according to the results obtained in this work (MBCOMT recovery = 75%), was the second best detergent to extract MBCOMT-His_6, next to DDM (MBCOMT recovery = 112%). Finally, CYMAL-5 (MBCOMT recovery = 53%) was moderately effective in extracting hMBCOMT-His_6. In addition, although the aggregation number of each detergent - the number of detergent monomers in a micelle - is influenced by the buffer additives where it is solubilized [Arnold and Linke, 2001] and thus may differ from those reported in Table 1, a positive correlation was found in this work between higher detergent aggregation numbers and higher biologically active hMBCOMT-His_6 extraction efficiencies.

A major drawback of using DDM is that it forms large micelles leading to a large protein-detergent complex which could compromise several structural studies [Privé, 2007]. Therefore, although DDM is the most appropriated detergent for hMBCOMT-His_6 extraction and it should be applied at the early stages of the bioprocess (solubilization and chromatographic procedures), it may not be compatible with crystallization experiments since one specific detergent is, in most cases, not the best choice for all demands in the whole process [Arnold and Linke, 2001]. As a consequence, it may be required that the detergent would be exchanged and this is accomplished by several ways. In particular, detergent exchange or removal have been achieved successfully in membrane protein bioprocesses using specific procedures comprising biobeads, dialysis, hydrophobic adsorption, affinity chromatography, ion-exchange chromatography or gel filtration [Arnold and Linke, 2001; Seddon, 2004].

3.3. hMBCOMT-His_6 adsorption in immobilized-metal affinity supports

To evaluate the hMBCOMT-His_6 adsorption onto the HisTrap™ FF crude loaded with nickel, solubilized *P. pastoris* membranes were injected using 20 mM sodium phosphate, NaCl 500 mM, 1 mM $MgCl_2$ at pH 7.8 as binding buffer, while the elution buffer contained 500 mM imidazole. In addition, according to the experiments reported in the previous section, 0.03% DDM was included in all chromatographic buffers to maintain the target enzyme in a solubilized state. Initially, a linear gradient was performed where the sample was injected with the binding buffer and then the imidazole concentration was increased from 0 to 500 mM during 30 minutes, as depicted in Figure 3. Indeed, it was observed that the elution of the retained proteins occurred in three peaks, corresponding approximately to 67.5, 210 and 500 mM of imidazole. Then, each peak was concentrated and desalted separately and after the determination of hMBCOMT biological activity, the following recoveries (calculated in each peak where 100% is the sum of hMBCOMT specific activity at 0 mM imidazole and the remaining peaks) were obtained: 5.91, 8.34 and 10.4 % for 67.5, 210 and 500 mM imidazole, respectively.

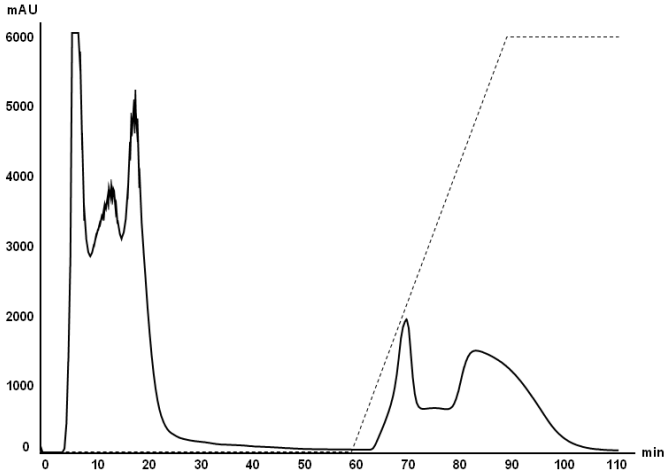


Figure 3 – Typical chromatographic profile of recombinant hMBCOMT-His_6 isolation from a *P. pastoris* solubilized-membranes by HisTrap™ FF crude loaded with Nickel in 20 mM sodium phosphate, 500 mM NaCl, 1 mM $MgCl_2$, DDM 0.03% at pH 7.8. Elution was performed with the binding buffer supplemented with 500 mM imidazole with a linear gradient from 0 to 500 mM imidazole during 30 min. Flow-rate and temperature were kept constant at 1 mL/min and 20ºC, respectively.

Looking at these results, it was verified that the majority of the target protein doesn't bind to the nickel matrix and the linear gradient applied wouldn't lead to the desired selectivity to effectively separate the different proteins in the crude extract. Therefore, a new strategy was elaborated in the previously described experimental conditions but with three stepwise gradients: 0, 50 and 500 mM of imidazole. In fact, although it was observed that hMBCOMT recovery at 500 mM increased from 10.4 to 32%, a large enzyme percentage eluted during the binding conditions. In order to improve the binding of hMBCOMT-His_6 to the matrix, the flow-rate was decreased to 0.5 mL/min during the injection of the sample once it often increases the binding of tagged proteins to the IMAC matrix [Nieba *et al.*, 1997]. Also, other chromatographic buffers than sodium phosphate were tested using the three stepwise strategy (0, 50 and 500 mM of imidazole), namely Tris 100 mM, and Hepes 50 mM, pH 7.8, both supplemented with NaCl 500 mM, 1 mM $MgCl_2$, DDM 0.03%. Using these buffers and lowering the flow-rate to 0.5 mL/min, hMBCOMT recoveries at 500 mM were 8.7% and 49.97 % for the Tris and Hepes buffers respectively, indicating that the Hepes buffer is more suitable for hMBCOMT adsorption onto the nickel-charged IMAC stationary phase. According to these results, it is likely that the target enzyme is more stable in this buffer but also that the pH may actually affecting hMBCOMT adsorption onto the matrix. Also in this specific case, after the injection of the sample, lower changes in the pH were observed. Following these results, we evaluated the effect of the immobilized ion on the target enzyme adsorption. Therefore, using the Hepes buffer system and maintaining the three stepwise gradient strategy, along

with nickel, we evaluated the performance of zinc to purify hMBCOMT in IMAC. As depicted in Figure 4, we observed that although the chromatographic profiles are similar for both nickel and zinc, the hMBCOMT recovery is much higher in nickel (49.97%) than with zinc, where only 6.76% of activity recuperation was obtained at 500 mM of imidazole. These observations are in accordance to what was previously reported [Cheung *et al.*, 2012; Gaberc-Porekar *et al.*, 2001], since the ions loading capacity for IMAC in an increasing order is as follows: cobalt>zinc>nickel>copper. In addition, it would be interesting to evaluate copper, which is usually the ion with the highest binding capacity.

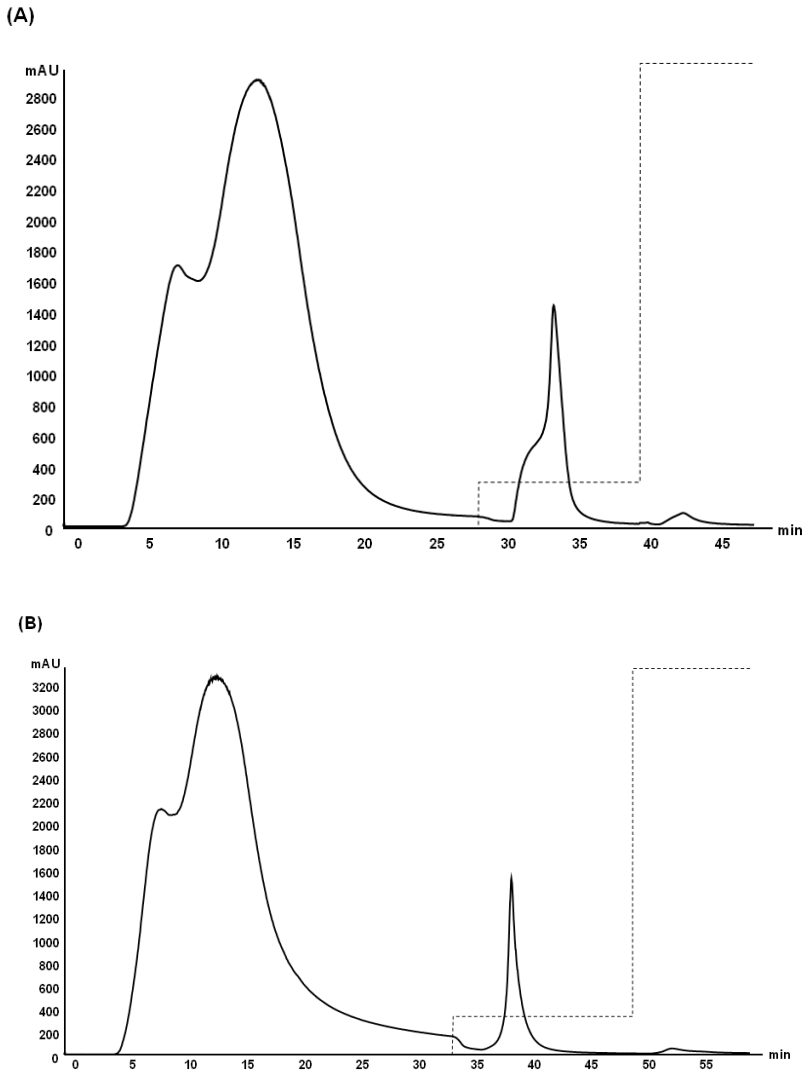


Figure 4 - Chromatographic profiles of recombinant hMBCOMT-His_6 isolation from a *P. pastoris* solubilized-membranes by HisTrap™ FF crude loaded with Nickel (A) or Zinc (B) in Hepes 50 mM, 500 mM NaCl, 1 mM $MgCl_2$, DDM 0.03% pH 7.8. Elution was performed with the binding buffer supplemented with 500 mM imidazole with three stepwise gradients: 0, 50 and 500 mM. Flow-rate was maintained constant at 0.5 mL/min during injection and 1 mL/min during elution. Temperature was kept constant at 20ºC.

On the other hand, the use of longer tag sequences or the use of a linker can be helpful to allow binding of the protein to the IMAC resin [Block *et al.*, 2009]. As we previously reported, using a linker for the TEV protease before the histidine tag greatly affects hMBCOMT biological activity and, consequently, is unappropriated. However, it would be worthwhile to evaluate the effect of using a longer histidine tag (10 or 12 histidines) on hMBCOMT adsorption onto the best IMAC matrix.

4. Conclusions

Hexa-histidine tagged-hMBCOMT was successfully expressed in *Pichia pastoris*. Specifically, despite including a linker for TEV protease prior the hexa-histidine tag may be favorable to promote the adsorption of detergent-solubilized MP in IMAC resins, it seems that affects the overall folding of hMBCOMT since no catalytically active enzyme was obtained in these conditions. From the detergent screening performed in this work, we observed that DDM was the most gentle and mild and thus the most suitable for hMBCOMT-His_6 extraction in a biologically active form from *P. pastoris* membranes.

The results obtained for hMBCOMT isolation and purification using IMAC showed that the majority of the target protein is lost during injection but it is minimized when using the hepes buffer system and the ion nickel immobilized in the stationary phase. It would be interesting to express hMBCOMT with a longer histidine tag than 6 histidine residues and to evaluate if it improves hMBCOMT affinity to the IMAC resin without affecting its biological activity. Following these achievements, additional stepwise gradients with different imidazole concentrations need to be tested in order to improve the selectivity of this process.

5. Acknowledgments

This research was supported by University of Beira Interior — Health Sciences Research Centre (CICS) and FCT (Portuguese Foundation for Sciences and Technology) by the project "EXPL/BBB478/BQB/0960/2012" and COMPETE: FCOMP-010124-FEDER-027563. A. Q. Pedro acknowledge a doctoral fellowship (SFRH/BD/81222/2011) from Fundação para a Ciência e Tecnologia. The authors also acknowledge the program COMPETE, the FCT project (Pest-C/SAU/ UI0709/2011).

6. References

Apud JA, Weinberger DR. Treatment of cognitive deficits associated with schizophrenia: potential role of catechol-O-methyltransferase inhibitors. CNS Drugs 2007; 21 (7): 535-57.

Arnold T, Linke D. The use of detergents to purify membrane proteins. Curr Protoc Protein Sci 2008; 4 (4.8.1-4.8.30).

Block H, Maertens B, Spriestersbach A, Brinker N, Kubicek J, Fabis R, Labahn J, Schäfer F. Immobilized-metal affinity chromatography (IMAC): a review. Methods Enzymol 2009; 463: 439-73.

Boden V, Winzerling JJ, Vijayalakshmi M, Porath J. Rapid one-step purification of goat immunoglobulins by immobilized metal ion affinity chromatography. J Immunol Methods 1995; 181 (2): 225-32.

Bonifacio MJ, Palma PN, Almeida L, Soares-da-Silva P. Catechol-*O*-methyltransferase and its inhibitors in Parkinson's disease. CNS Drug Rev 2007; 13 (3): 352-79.

Cheung RCF, Wong JH, Bun Ng T. Immobilized metal ion affinity chromatography: a review on its applications. Appl Microbiol Biotechnol 2012; 96: 1411-20.

Correia FF, Santos FM, Pedro AQ, Bonifacio MJ, Queiroz JA, Passarinha LA. Recovery of biological active catechol-*O*-methyltransferase isoforms from Q-sepharose. J Sep Sci 2014; 37 (1-2): 20-9.

Gaberc-Porekar V, Menart V. Perspectives of immobilized-metal affinity chromatography. J Biochem Biophys Methods 2001; 49: 335-60.

Hage DS, Anguizola JA, Bi C, Li R, Matsuda R, Papastavros E, Pfaunmiller E, Vargas J, Zheng X. Pharmaceutical and biomedical applications of affinity chromatography: Recent trends and developments. J Pharm Biomed Anal 2012; 69: 93-105.

Halliwell CM, Morgan G, Ou C, Cass AEG. Introduction of a (Poly)histidine tag in $_{L}$-Lactate dehydrogenase produces a mixture of active and inactive molecules. Anal Biochem 2001; 295: 257-61.

Harrison ST, Poslusney MS, Mulhearn JJ, Zhao Z, Kett NR, Schubert JW, Melamed JY, Allison TJ, Patel SB, Sanders JM, Sharma S, Smith RF, Hall DL, Robinson RG, Sachs NA, Hutson PH, Wolkenberg SE, Barrow JC. Synthesis and evaluation of heterocyclic catechol mimics as inhibitors of catechol-*O*-methyltransferase (COMT). ACS Med Chem Lett 2015; 6 (3): 318-23.

Lantez V, Nikolaidis I, Rechenmann M, Vernet T, Noirclerc-Savoye M. Rapid automated detergent screening for the solubilization and purification of membrane proteins and complexes. Eng Life Sci 2015; 15: 39-50.

Jones C, Patel A, Griffin S, Martin J, Young P, O'Donnell K, Silverman C, Porter T, Chaiken I. Current trends in molecular recognition and bioseparation. J Chromatogr A 1995; 707: 3-22.

Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970; 227: 680-85.

Le Maire M, Champeil P, Moller JV. Interaction of membrane proteins and lipids with solubilizing detergents. Biochim Biophys Acta 2000; 1508: 86-111.

Ma Z, Liu H, Wu B. Structure-based drug design of catechol-*O*-methyltransferase inhibitors for CNS inhibitors. Brit J Clin Pharmacol 2013; 77 (3): 410-20.

Mooney JT, Fredericks DP, Zhang C, Christensen T, Jespergaard C, Schiodt CB, Hearn MT. Purification of a recombinant human growth hormone by an integrated IMAC procedure. Protein Expr Purif 2014; 94: 85-94.

Müller T. Catechol-*O*-methyltransferase inhibitors in Parkinson's disease. Drugs 2015; 75 (2): 157-74.

Myöhänen TT, Männistö PT. Distribution and functions of catechol-*O*-methyltransferase proteins: do recent findings change the picture? Int Rev Neurobiol 2010; 95: 29-47.

Nieba L, Nieba-Axmann SE, Persson A, Hamalainen M, Edebratt F, Hansson A, Lidholm J, Magnusson K, Karlsson AF, Pluckthun A. BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. Anal Biochem 1997; 252 (2): 217-28.

Pedro AQ, Bonifácio MJ, Queiroz JA, Maia CJ, Passarinha LA. A novel prokaryotic expression system for biosynthesis of recombinant human membrane-bound catechol-*O*-methyltransferase. J Biotechnol 2011; 156: 141-46.

Pedro AQ, Soares RF, Oppolzer D, Santos FM, Rocha LA, Gonçalves AM, Bonifácio MJ, Queiroz JA, Gallardo E and Passarinha LA. An Improved HPLC Method for Quantification of Metanephrine with Coulometric Detection. J Chromatograph Separat Techniq 2014; 5:217.

Pedro AQ, Oppolzer D, Bonifácio MJ, Maia CJ, Queiroz JA, Passarinha LA. Evaluation of Mut^S and Mut^+ *Pichia pastoris* strains for membrane-bound and catechol-*O*-methyltransferase biosynthesis. Appl Biochem Biotechnol 2015; 175: 3840-55.

Porath J, Carlsson J, Olsson I, Belfrage G. Metal chelate affinity chromatography, a new approach to protein fractionation. Nature 1975; 258: 598-99.

Privé GG. Detergents for the stabilization and crystallization of membrane proteins. Methods 2007; 41: 388-97.

Santos FM, Pedro AQ, Soares RF, Martins R, Bonifácio MJ, Queiroz JA, Passarinha LA. Performance of hydrophobic interaction ligands for human membrane-bound catechol-*O*-methyltransferase purification. J Sep Sci 2013; 36: 1693-1702.

Seddon AM, Curnow P, Booth PJ. Membrane proteins, lipids and detergents: not just a soap opera. Biochim Biophys Acta 2004; 1666: 105-17.

Serretti A, Olgiati P. Catechol-O-methyltransferase and Alzheimer's disease: a review of biological and genetic findings. CNS Neurol Disord Drug targets 2012; 11 (3): 299-305.

Skarka A, Skarydova L, Stambergova H, Wsol V. Purification and reconstitution of human membrane-bound DHRS7 (SDR34C1) from Sf9 cells. Protein Expr Purif 2014; 95: 44-9.

Sonoda Y, Cameron A, Newstead S, Omote H, Moriyama Y, Kasahara M, Iwata S, Drew D. Tricks of the trade used to accelerate high-resolution structure determination of membrane proteins. FEBS Lett 2010; 584: 2539-47.

Vieira-Coelho MA, Soares-da-Silva P. Effects of tolcapone upon soluble and membrane-bound brain and liver catechol-*O*-methyltransferase. Brain Res 1999; 821 (1): 69-78.

Tenhunen J, Salminen M, Lundstrom K, Kiviluoto T, Savolainen R, Ulmanen I. Genomic organization of the human catechol-*O*-methyltransferase gene and its expression from two distinct promoters. Eur J Biochem 1994; 223 (3): 1049-59.

Voutilainen S, Tuomainen T, Korhonen M, Mursu J, Virtanen JK, Happonen P, Alfthan G, Erlund I, North KE, Mosher MJ, Kauhanen J, Tiihonen J, Kaplan GA, Salonen JT. Functional COMT Val158Met polymorphism, risk of Acute Coronary events and serum homocysteine: The Kuopio Ischaemic heart disease risk factor study. PLoS ONE 2007; 1: 181-87.

Yager JD. Mechanisms of estrogen carcinogenesis: The role of E2/E1 – Quinone metabolites suggests new approaches to preventive intervention – A review. Steroids 2015; 99 (Pt A): 56-60.

Yang J, Kulkami K, Manolaridis I, Zhang Z, Dodd RB, Mas-Droux C. Mechanism of isoprenylcysteine careboxyl methylation from the crystal structure of the integral membrane methyltransferase ICMT. Mol Cell 2011; 44: 997-1004.

Chapter 4

Concluding remarks

This thesis established the general hypothesis of expressing both soluble and membrane-bound COMT isoforms by *P. pastoris* methanol-induced cultures and employing affinity chromatography to specifically and accurately isolate COMT from *P. pastoris* lysates. Following the first biochemical characterizations of COMT enzymes and identification of their role in several human diseases, there was interest in developing more rapid and accurate analytical methods for measuring COMT activity. Initially, we implemented and validated an analytical method based on the use of reversed-phase HPLC coupled with coulometric detection for quantifying metanephrine, an *O*-methylated product in COMT enzymatic assays. Specifically, as reported in paper II, the improvement in the signal/noise ratio achieved with this method allowed us to lower the low limit of quantification of metanephrine with shorter chromatographic runs. This method may be valuable not only for determining endogenous metanephrine but in COMT recombinant research as it was deeply employed throughout the entire work developed in this thesis.

The harmful role of COMT in several human neurological disorders is well characterized and the efficacy of the therapy strongly depends on the selectivity and accuracy of the COMT inhibitors applied. Indeed, developing such inhibitors requires research with recombinant COMT, thus highlighting the need of developing bioprocess strategies that meet the highly demanding requirements of the majority of biophysical techniques. Therefore, here we implemented and developed an alternative upstream strategy for expressing a recombinant hexa-histidine tagged form of SCOMT (SCOMT-His_6) using *P. pastoris* coupled with one step purification employing IMAC. As reported in paper III, the time course profile biosynthesis of SCOMT-His_6 after *P. pastoris* lysis and subsequent fractionation demonstrated that highly biologically active SCOMT was found in the ressuspended pellet after 48 hours of fermentation. On the other hand, IMAC proved to be extremely efficient and selective for the direct capture of SCOMT-His_6 and the MALDI-TOF/TOF analysis allowed us to conclude that the recombinant form produced is well processed by *P. pastoris* intracellular machinery. Finally, the sensitivity of COMT to be inhibited by 3,5-DNC and entacapone was determined and the IC_{50}s (half-maximal inhibitory concentration) values in the same order of magnitude than those previously reported were obtained. The approach described successfully in this work demonstrates that the recombinant form of SCOMT produced in *P. pastoris* and purified using IMAC is obtained in a high purity degree with its native primary sequence and thus, may be potentially applied for structure-activity relationship studies.

Following the development of an appropriated strategy for biosynthesis and purification of SCOMT, our focus was the membrane-bound isoform. The benefits accomplished in PD therapy with the pharmacological inhibition of COMT is mainly achieved via soluble isoform while

recently it has been described that MBCOMT may be more important in other neurologic disorders than PD. Consequently, developing more selective and MBCOMT-specific inhibitors may improve the outcome of specific neurologic disorders, thus highlighting the need of developing new bioprocess strategies with recombinant MBCOMT. Initially, we reported recombinant MBCOMT biosynthesis by shake-flask *P. pastoris* methanol-induced cultures with opposite phenotypes, according to paper IV. Indeed, the performance of *P. pastoris* strains X33 (Mut^+) and KM71H (Mut^S) for expressing MBCOMT was similar employing a specific concentration of a feed composed uniquely by methanol, contrary to what was found by mixed-feeds. The determination of the affinity values to epinephrine in the different fractions obtained after *P. pastoris* subcellular fractionation allowed us to conclude that the recombinant form produced in this work is kinetically identical to its correspondent native enzyme. Overall, an upstream stage was sucessfully developed for MBCOMT biosynthesis by *P. pastoris* methanol-induced cultures. As it was observed that the methanol concentration greatly influenced the expression levels in the KM71H Mut^S strain and to avoid unsuccessfull optimization issues, the Mut^+ X33 strain was selected for scaling-up this bioprocess.

The ANN modelling of the methanol induction phase was conducted to optimize MBCOMT biosynthesis in mini-bioreactors using *P. pastoris* X33 methanol-induced cultures under the control of alcohol oxidase promoter, as demonstrated in paper V. The ANN model was able to describe the effects of temperature, dimethylsulfoxide concentration and methanol flow-rate on MBCOMT specific activity, as shown by the good fitness between the predicted and measured values. At the optimal conditions estimated by the ANN model, a 1.58-fold increase was obtained for MBCOMT specific activity over the highest value achieved in the experimental design while an improvement of 6.4-fold was found over the small-scale biosynthesis in baffled shake-flasks. In addition, this strategy doesn't impair *P. pastoris* viability once approximately 90% of viable cells were obtained at the end of the induction phase. The experimental strategy developed in this work shows that the manipulation of fermentation conditions coupled with the addition of specific molecules can open new perspectives in the optimization of *P. pastoris* bioprocesses for recombinant MP biosynthesis.

The success of an integrated strategy for obtaining recombinant MP for structural and functional studies relies not only in the production and recovery steps but also in developing equally appropriated downstream strategies, mostly using chromatographic procedures. Therefore, the successfull report of biosynthesis strategies for recombinant MBCOMT at a small and a medium-scale prompted us to apply affinity chromatography protocols for MBCOMT isolation and purification. As shown in paper VI, a new affinity chromatographic process was described for the isolation of MBCOMT based on the specific molecular recognition between particular MBCOMT amino acid residues and a commecial arginine-matrix. This procedure was performed using conditions that mainly promote electrostatic interactions and the supplementation of the chromatographic buffers with triton X-100 and dithiothreitol allowed MBCOMT isolation in a catalitically active form. Moreover, it was also

evaluated the effect of the temperature, pH and concentration of the injected sample in MBCOMT adsorption and selectivity of the chromatographic process. In particular, despite the $_L$-arginine support was able to bind the target protein in a wide range of pH values (above and below the MBCOMT isoelectric point), the highest purification factors were obtained at pH 7 and with the injection of a more concentrated lysate sample. In addition, to avoid mass transport limitations and consequently to increase the overall yield, the procedure was carried out at 20 ºC. Finally, we believe that the affinity interaction promoted by the $_L$-arginine support and MBCOMT involves multiple non-covalent interactions, what can trigger new possibilities in the chromatographic isolation of other recombinant proteins in which low purity and low recovery yields are usually obtained with the traditional chromatographic strategies. In general, although the purification factor obtained with this method is similar from those previously reported, MBCOMT elution occurs in a single peak pattern, thereby improving the yield of the process. Nevertheless, the purity achieved with this method is not as high as it is required for the highly demanding techniques used in structural biology of MP.

Therefore, in order to explore the high selectivity usually achieved with IMAC, paper VII describes the construction of new *P. pastoris* expression plasmids encoding the target protein with a hexa-histidine tag (MBCOMT-His_6) in the carboxy-terminal or alternatively, with a cleavage site for Tobacco-Etch Virus (TEV) prior the hexa-histidine tag (MBCOMT-TEV-His_6). Initial studies were designed to evaluate the influence of the histidine tag or the histidine tag plus the TEV linker on MBCOMT enzymatic activity and it was observed that MBCOMT-His_6 was produced with lower biological activity than the native, non-tagged form, indicating that the tag slightly affects MBCOMT overall folding. On the other hand, MBCOMT-TEV-His_6 was produced in a catalitically inactive form and even after cleavage of the tag, non-native amino acids will be present behind the protease recognition sites, thus compromising the enzyme activity. From the detergent screening performed, it was observed that n-dodecyl-β-$_D$-maltopyranoside (DDM) was the most gentle and mild and thus the most suitable for hMBCOMT-His_6 extraction in a biologically active form from *P. pastoris* membranes. The results obtained for hMBCOMT-His_6 isolation and purification using IMAC showed that the majority of the target protein is lost during injection but it is minimized when using the hepes buffer system and the ion nickel immobilized in the stationary phase. Following these achievements, more research is needed to obtain fractions of MBCOMT-His_6 in a more purified fraction using additional stepwise gradients with different imidazole concentrations or other elution methods to improve the selectivity of this strategy.

More than three decades ago, COMT was described as an important pharmacological target in PD and as it was described that COMT inhibition improved the outcome in PD patients, there was interest in developing COMT inhibitor molecules. Therefore, as COMT recombinant research was intensified, major progresses were achieved, which led to the resolution of the first structures of SCOMT, either from rat (1994) or human recombinantly expressed in *E. coli* (2008). Throughout this entire process, the major achievements were made with SCOMT over

MBCOMT once it was reported that the benefits in PD may come mainly from the inhibition of the soluble isoform.

In general, the integrated strategy implemented for recombinant SCOMT in our laboratory will be important once when coupled to a final polishing step will allow performing structural or bio-interaction studies.

On the other hand, the interest of the scientific community in MBCOMT is more recent as relatively new achievements implicated this isoform in other diseases than PD. Indeed, unveiling the epitope binding domain of the target inhibitor molecules responsible for interaction with MBCOMT could have major implications in the future design of more selective and isoform specific inhibitors. However, these biophysical techniques require high quantities of purified enzyme in a biologically active state, emphasizing the importance of developing new strategies with recombinant MBCOMT research. In a general way, the progress achieved with this work meets those highly demanding requirements, mainly regarding the upstream stage as well as COMT stabilization where moderately to high quantities of catalitically active enzyme were obtained. Nonetheless, the strategies successfully described here for partial MBCOMT purification need to be improved, especially for IMAC once it is considered to be highly selective and, thus it is feasible that after succesfull optimization procedures, fractions with high purity will be obtained.

Future perspectives

The progress achieved with this thesis allows to discloser diverse important findings regarding the biosynthesis and purification procedures of human recombinant COMT enzymes. However, despite all the efforts conducted throughout this thesis, further experiments are still required to emphasize and increase the significance of the findings reported.

Overall, the optimized upstream stage using *P. pastoris* as the host for both COMT isoforms (108/158 Valine variant) delivered moderately to high quantitities of the target enzymes. However, improved purification procedures, mostly for MBCOMT, will be required to perform structural studies. Concerning the purification of MBCOMT using IMAC matrices, additional stepwise gradients with different imidazole concentrations needed to be tested to obtain the target enzyme in a highly purified fraction. Also, it would be interesting to express MBCOMT with a longer histidine tag than 6 histidine residues (8 or 10) and evaluating if it improves MBCOMT adsorption to the IMAC resin without affecting its biological activity.

COMT role in human disorders can be seen from two different angles; if on one hand, it is an important pharmacological target and thus are required COMT inhibitor molecules; on the other hand, it was also reported that the inability of COMT - associated for example with the polymorphic methionine variant - to inactivate specific substrates such as catecholestrogens can also be associated with the progression of some cancers. Therefore, to target this last strategy, chromatographic strategies for COMT purification in a native, untagged form with high purity are required. A promising method may be the application of immunoaffinity matrices where the immobilization of an antibody with the same epitope for soluble and MBCOMT can render a chromatographic matrix feasible to be applied irrespective the COMT isoform. In addition, an alternative strategy to MBCOMT biosynthesis may involve its recombinant production through eukaryotic cell-free expression systems based on wheat germ or insect cells and insert it in a co-translational form into pre-formed lipid bilayers such as liposomes or nanodiscs. The functional polymorphism at codon 108/158 (for soluble and MBCOMT, respectively) may also influence the genesis and/or progression of different diseases. Consequently, it will be interesting to develop equally appropriated strategies for biosynthesis and purification of COMT methionine variants 108/158 and evaluate if the binding mechanism of the COMT inhibitor molecules is different.

Finally, we believe the improvements that may be achieved with the additional work here reported for both COMT enzymes will allow obtaining high quantities of purified enzymes. Therefore, it would be possible performing structural and bio-interaction studies using the apo-enzymes or complexed with different ligands (cofactors or inhibitors) by nuclear magnetic ressonance, isothermal titration calorimetry or even using crystallographic experiments