Márcia Maria da Silva Alves

Licenciada em Ciências da Saúde

hCES2: studies on the role of glycosylation in activity

Dissertação para obtenção do Grau de Mestre em

Genética Molecular e Biomedicina

Orientador: Ana Luísa Simplício, Ph.D, IBET/ITQB-UNL

Júri:

Presidente: Prof. Doutora Paula Gonçalves

Arguente: Prof. Doutora Isabel Sá Nogueira

Vogal: Prof. Doutora Ana Luísa Simplício

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro de 2013

Márcia Maria da Silva Alves

Licenciada em Ciências da Saúde

hCES2: studies on the role of glycosylation in activity

Dissertação para obtenção do Grau de Mestre em

Genética Molecular e Biomedicina

Orientador: Ana Luísa Simplício, Ph.D, IBET/ITQB-UNL

Júri:

Presidente: Prof. Doutora Paula Gonçalves

Arguente: Prof. Doutora Isabel Sá Nogueira

Vogal: Prof. Doutora Ana Luísa Simplício

“hCES2: studies on the role of glycosylation in activity.”

Agradecimentos

O desenvolvimento deste trabalho só foi possível graças ao contributo, directo ou indirecto, de várias pessoas, às quais gostaria de agradecer:

À Doutora Ana Luísa Simplício pela orientação, pela paciência e por me ter deixado pensar por mim própria. Não teria aprendido tanto durante este ano que passou sem a sua ajuda.

À Doutora Júlia Costa do Laboratório de Glicobiologia pela disponibilidade e ajuda.

À Doutora Joana Lamego por toda a ajuda no planeamento do trabalho e pelo pensamento crítico.

A todos os elementos do Laboratório de Tecnologia de Células animais por me acolherem e por me ensinarem o que quer que fosse preciso.

Aos colegas do Laboratório de Farmacocinética e Biofarmacêutica pelo bom ambiente.

Aos amigos, os mais antigos e os mais recentes, que nunca andam muito longe e que são sempre uma fonte de boa-disposição.

Aos meus pais, por tudo o que me proporcionaram, e em especial à minha mãe. À minha irmã, ao meu cunhado e ao meus padrinhos, pelo apoio e por acreditarem em mim.

Ao Pedro, este percurso não teria sido o mesmo se não estivesses ao meu lado.

À Fundação para Ciência e Tecnologia (FCT) pelo financiamento do projecto (PTDC/EBB-BIO/111530/2009).

Resumo

A carboxilesterase 2 humana (hCES2) é uma enzima com cada vez mais relevância devido ao seu papel na hidrólise de vários fármacos, pró-fármacos e xenobióticos. No entanto, a sua estrutura tridimensional ainda não foi resolvida, e muitas das suas propriedades estão ainda por esclarecer.

Neste estudo foi avaliado o impacto da glicosilação na actividade de uma variante de hCES2, hCES2-10xHis, através de vários métodos, como desglicosilação enzimática e inibição da glicosilação por inibidores e por diferentes mutações.

Chegou-se à conclusão que a glicosilação durante a síntese de hCES2 é necessária para manter a actividade da proteína, mas a presença de glicanos não é essencial após a proteína estar completamente maturada.

Palavras-chave: carboxilesterase, glicosilação, desglicosilação, PNGase F, tunicamicina, mutagénese dirigida.

Abstract

Human carboxylesterase 2 (hCES2) is becoming more relevant due to its role in the hydrolysis of several drugs, pro-drugs and xenobiotics. However, its three dimensional structuture is yet to be resolved, and many of its properties are still unclear.

In this work, the impact of glycosylation in the enzymatic activity of an hCES2 variant, hCES2-10xHis, was evaluated. To achieve this objective, different methods were used, such as enzymatic deglycosylation and glycosylation inhibition through specific inhibitors or plasmid mutation.

It was concluded that glycosylation is necessary during hCES2 synthesis to maintain protein activity, but the presence of the glycans is not essential after the protein is completely matured.

Keywords: carboxylesterase, glycosylation, deglycosylation, PNGase F, tunicamycin, site directed mutagenesis.

Table of Contents

Contents – Figure Index

Contents – Table Index

Abbreviations

%GC – Guanine Cytosine Content

10xHis – Ten Histidine Tag

4-MUB – 4-methylumbelliferone

4-MUBA – 4-methylumbelliferyl acetate

Amp – Ampicillin

Asn – Asparagine

Asn-X-Ser/Thr – Asparagine-Any Aminoacid-Serine/Threonine Aminoacid Sequence

ATCC – American Type Culture Collection

BHK – Baby Hamster Kidney Cells

Bp – Base Pairs

BSA – Bovine Serum Albumin

CES – Carboxylesterase

CES1 – Carboxylesterase 1

CES2 – Carboxylesterase 2

hCES2-10xHis – Human Carboxylesterase 2 with a C-terminal Ten-Histidine Tag

CES3 – Carboxylesterase 3

CHO – Chinese Hamster Ovary Cells

CMV – Cytomegalovirus

CO_2 – Carbon Dioxide

CPT-11 – Irinotecan

DMSO – Dimethyl Sulfoxide

DNA – Deoxyribonucleic Acid

Dol – Dolichol

***E.Coli* –** *Escherichia coli*

ECL – Enhanced Chemilluminescence

EDTA – Ethylenediaminetetraacetic Acid

Endo H – Endoglycosidade H

ER – Endoplasmic Reticulum

Fw – Forward

Glc – Glucose

GlcNAc – N-Acetylglucosamine

Gln – Glutamine

Glu – Glutamic Acid

HBSS – Hank's Balanced Salt Solution

hCES1 – Human Carboxylesterase 1

hCES2 – Human Carboxylesterase 2

HEK 293T – Human Embrionic Kidney 293T Cells

His – Histidine

HRP – Horse Raddish Peroxidase Conjugated Antibodies

HTEL – Histidine – Threonine – Glutamic Acid – Leucine Aminoacid Sequence

KDEL – Lysine – Aspartic Acid – Glutamic Acid – Leucine Aminoacid Sequence

kb – Kilo Base

kDa – Kilo Dalton

LB – Lysogeny Broth

Man – Mannose

MES – 2-(*N*-Morpholino)Ethanesulfonic Acid

NC – Negative Control

NSO – Mouse Myeloma Cells

P – Phosphate

PBS – Phosphate Buffered Saline

PC – Positive Control

pCi-neo hCES2-10xHis – pCi-neo Expression Vector With *hCES2-10xHis* Gene

PCR – Polymerase Chain Reaction

PEI – Poliethylamine

PNGase F – Peptide N-Glycosidase F

PVDF – Polyvinylidene Difluoride

RNA – Ribonucleic Acid

RNase B – Ribonuclease B

Rv – Reverse

SDS-Page – Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Ser – Serine

SN-38 – active metabolite of Irinotecan

SV40 – Simian Virus 40

TAE – Tris Acetate EDTA

TB – *Terrific Broth*

TBS – Tris Buffered Saline

Thr – Threonine

T_m – Melting Temperature

TM – Tunicamycin

UV – Ultraviolet Radiation

1 Introduction

1.1 Carboxylesterases

Carboxylesterases (CES) are a subset of esterases that belong to the α/β hydrolase family, which three dimensional structure consists of α helixes and β sheets connected through loops of variable length (Xie *et al*, 2002). These enzymes catalyze the hydrolysis of esters, thioesters and amides, freeing the respective carboxylic acid and alcohol (Lockeridge and Quinn, 2010). Ester bonds increase the lipophilicity and bioavailability of drugs, making them substrates for CES, among other esterases (Ross *et al*, 2012). Through hydrolysis, these exogenous chemicals are transformed into water soluble carboxylic acids and alcohols, promoting their excretion or activation of other functions (Laizure *et al*, 2013). Therefore, carboxylesterases are responsible for detoxifying exogenous substrates, like pesticides and environmental toxicants and are also involved in metabolizing drugs and/or activating pro-drugs (Satoh and Hosokawa, 1998).

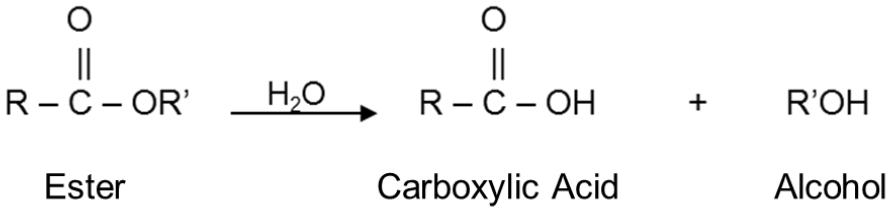


Figure 1.1: Ester Hydrolysis. CES catalyze the addition of a water molecule to an ester producing a carboxylic acid and an alcohol.

Multiple forms of CES are expressed in every mammal species, although there are differences in the tissues where each form is mainly expressed. In addition, some species like rodents express soluble forms that are therefore present in the plasma but others, like humans, only express intracellular forms (Bahar *et al*, 2012).

In general, CES enzymes are glycoproteins with one or more N-glycosylation sites, have four conserved cysteine residues and, since they are serine esterases, have a catalytic triad that consists of a Serine, a Glutamic Acid and a Histidine (Xie *et al*, 2002). Most substrates of carboxylesterases are also hydrolyzed by other esterases present in the organism. However, some CES have higher affinities for certain compounds, depending on their structure.

Three main human carboxylesterases have been described: CES1, CES2 and CES3. All of them are codified in chromosome 16 but have different numbers of exons, tissues of expression, localization within the cell and substrate preferences (Lockeridge and Quinn, 2010).

CES1 is expressed in a higher level in the liver, but is also present in the gastrointestinal tract, and is composed of a 180 kDa trimer with higher affinity to substrates with a small alcohol group and larger acyl groups, like clopidogrel, for example. CES2 can also be found in the liver in a lower expression level than CES1, but it is the most important CES present in the small intestine, where its activity is constant from the jejunum to the ileon, and is also present in the kidney and lung. It is reported to be present mostly in the monomeric form (60 kDa), however, recent reports have demonstrated that it can also be active in an oligomeric form (Lamego *et al*, 2013). It hydrolyzes preferentially substrates with a large alcohol group and smaller acyl group like irinotecan. The least studied human carboxylesterase, CES3, is also a 60 kDa monomer and is located in the liver and gastrointestinal tract, but to a lesser extent than the CES2 (Laizure *et al*, 2013; Lockeridge and Quinn, 2013).

Human CES enzymes have a signal peptide of approximately 17 to 22 aminoacid residues responsible for their targeting to the endoplasmic reticulum (ER) where they are retained due to a carboxy-terminal ER retention signal – the H-X-E-L (Histidine – X – Glutamic Acid – Leucine) consensus sequence that binds to the KDEL (Lysine – Aspartic Acid – Glutamic Acid – Leucine) receptor present in the luminal side of the ER, anchoring the protein (Satoh and Hosokawa, 2010). This explains why there are no CES present in human plasma, unlike what happens in most rodents (Satoh and Hosokawa, 2006; Bahar *et al*, 2012).

Human CES2 (hCES2), which is the main subject of the present work, is the major intestinal carboxylesterase although it can also be found in many other tissues and is overexpressed in some types of cancer, such as colon carcinoma (Xie *et al*, 2002).

The *CES2* gene is located in the human chromosome 16 and contains 12 exons (Lockeridge and Quinn, 2010). Its translation begins in the second ATG codon of the first exon, and the surrounding nucleotides are in agreement with the KOZAK consensus sequence, which is necessary for initiating protein translation. The N-terminal signal peptide targets the translated protein to the ER and the C-terminal tetrapeptide (H-T-E-L, in the case of hCES2) is responsible for anchoring the protein to the luminal side of the ER (Schwer *et al*, 1997). It has been shown that there are possible variants of hCES2 due to alternative splicing sites; however these did not show enzymatic activity for the tested substrates (Schiel *et al*, 2007).

Similarly to other carboxylesterases, the catalytic triad of hCES2 consists of Ser-Glu-His aminoacid residues, and the serine is included in a G-X-S-X-G motif, where X is any aminoacid (Schwer *et al*, 1997). The active site is located in the center of the molecule, providing the ideal hydrophobic environment for the hydrolysis of several hydrophobic substrates (Satoh and Hosokawa, 2010).

Several cardiovascular and chemotherapeutic drugs are mainly hydrolyzed by hCES2:

- Angiotensin receptor blockers

- Aspirin: when hydrolyzed by hCES2 into salicylic acid loses its anti-platelet activity and the anti-inflammatory effect is activated
- CPT-11 (irinotecan): chemotherapeutic agent with high affinity for hCES2. Hydrolysis leads to the SN-38 active metabolite, which is more cytotoxic and efficient.

It has been shown that patients respond differently to treatment with these drugs, due to genetic polymorphisms, diseases or drug interactions indicating that the level of expression of hCES2, and other carboxylesterases, can modulate the therapeutic response (Schiel *et al*, 2007; Ross *et al*, 2012; Laizure *et al*, 2013).

1.2 Glycosylation

Glycosylation occurs during protein synthesis in eukaryotic cells adding glycans to a peptide backbone, thus forming glycoproteins. The glycans have variable structures and are composed of sugar residues (Butler, 2006).

There are two main forms of glycosylation: O-glycosylation and N-glycosylation. The first involves the addition of a small sugar to a serine or threonine aminoacid and takes place within the Golgi apparatus. Despite the lack of consensus sequence, this glycosylation occurs in protein regions with a high proportion of serine, threonine and proline (Butler, 2006).

N-Glycosylation, the most common and most studied, is post-translational and takes place in the ER lumen. An N-glycan is added to an asparagine residue located in an Asparagine-X-Serine/Threonine (Asn-X-Ser/Thr) sequon. The sequon is only indicative of a potential N-glycosylation site, since not all asparagine residues from the sequons are glycosylated. If "X" (any aminoacid except proline) is acidic, the efficiency of glycosylation can decrease. Conformational limitations or the presence of disulfide bonds can also interfere with the accessibility of some sequons (Varki *et al*, 2009).

All the N-glycans share the same core saccharide ($Man_3GlcNAc_2$) and can be classified as high mannose, hybrid or complex type, according to how they are processed.

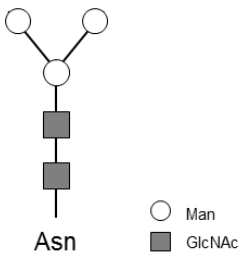


Figure 1.2: N-glycans core saccharide. $Man_3GlcNAc_2$. Man – Mannose, GlcNAc – N-acetylglucosamine. Adapted from Butler, 2006.

N-Glycan synthesis begins in the cytoplasmic side of the ER, where the GlcNAc-P-P-Dol precursor is formed. After the addition of mannose and N-acetylglucosamine (GlcNAc) residues to the precursor, the entire structure – $Man_5GlcNAc_2$-P-P-Dol – is translocated into the ER lumen. Inside the ER, mannose and glucose residues are added until the glycan is ready to be transferred to the newly synthesized peptide. The $Glc_3Man_9GlcNAc_2$-P-P-Dol is cleaved freeing the P-P-Dol, and the rest of the glycan is transferred to the asparagine residue from the Asn-X-Ser/Thr sequon. A series of reactions then trim the N-glycan, removing the glucose residues (Butler, 2006; Varki *et al*, 2009).

All of the processes that take part of N-glycosylation are catalyzed by specific enzymes. GlcNAc-1-phosphotransferase, for example, acts outside the ER and is responsible for the formation of GlcNAc-P-P-Dol. During the addition of sugar residues to form the glycan and its trimming glycosyltransferases and glucosidases are involved.

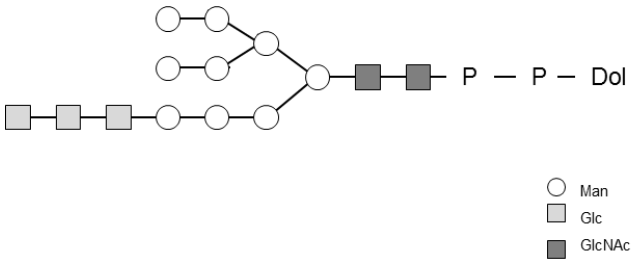


Figure 1.3: N-glycan precursor prior to the transference to the asparagine residue. $Glc_3Man_9GlcNAc_2$-P-P-Dol. Man – Mannose; Glc – Glucose; GlcNAc – N-acetylglucosamine; P – Phosphate; Dol - Dolichol. Adapted from Butler, 2006.

If the glycoprotein is retained in the ER, the N-glycans remain unalteratered and high mannose type. However, the glycoproteins that are directed to the Golgi are subjected to further processing: mannose residues are removed and other residues can be added like fucose, sialic acid, galactose and GlcNAc. This results in a variable number, as well as, longer branches, characteristic of complex and hybrid type N-glycans (Butler, 2006; Varki *et al*, 2009).

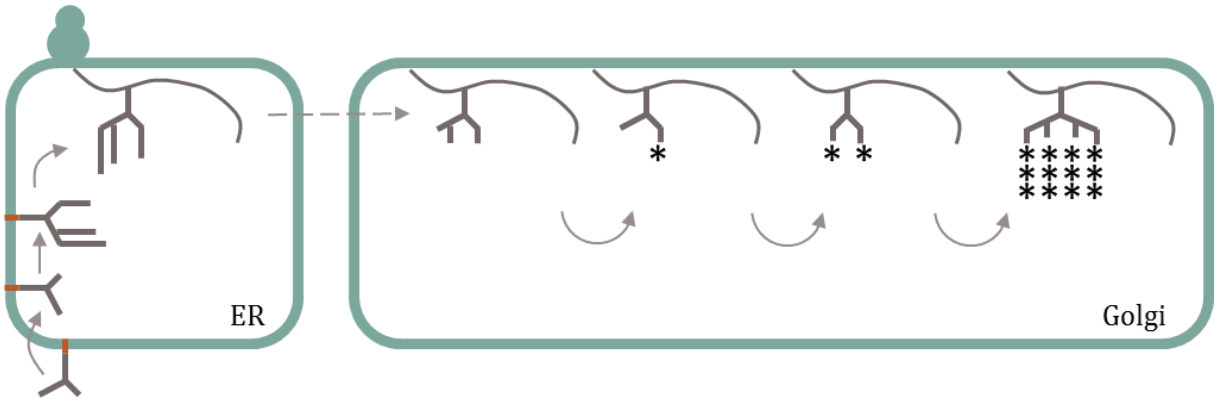


Figure 1.4: N-glycosylation process in the Endoplasmic Reticulum (ER) and Golgi. At the ER, the $Glc_3Man_9GlcNAc_2$ glycan (high mannose type) is attached to the asparagine residue of the glycosylation sequon of the translating protein. The glycoproteins that are directed to the Golgi are further processed, and their glycans become hybrid or of the complex type. Glc – Glucose, Man – Mannose, GlcNAc – N-acetylglucosamine.

New peptides begin to assume their final three dimensional conformations during protein synthesis. Once the peptide, in the case of a glycoprotein, is glycosylated its folding will be influenced by it – the glycans limit the conformational liberty of the peptide and maintain its structure and stability. Furthermore, N-glycosylation ensures that the peptide is correctly folded. Two chaperones in the ER bind to N-glycans of misfolded peptides, retaining them in the lumen until they are correctly folded. If the peptides remain misfolded they are translocated to the cytoplasm, deglycosylated and degraded in the proteasome. This mechanism improves folding efficiency, preventing aggregation and oligomerization (Varki *et al*, 2009).

hCES2 is described to have two potential N-glycosylation sites, identified by the presence of the Asn-X-Ser/Thr sequon at Asn103 and Asn267 (Satoh and Hosokawa, 2010). The purified enzyme has been deglycosylated with Endoglycosidase H (Endo H), which hydrolyses Asn-linked high mannose oligossacharides, resulting in a decrease in molecular mass (Pindel *et al*, 1997). This experiment showed that the protein was not processed in the Golgi, but failed to clarify the number of glycosylated sequons present.

The relevance of glycosylation in other mammalian CES has previously been studied. It was found that enzymatic deglycosylation did not alter the activity of two purified rat liver hydrolases, which have similar characteristics to human CES, although one of them adsorbed to plastic and glass in the absence of the glycans (Morgan *et al*, 1994). The expression of a recombinant rabbit liver carboxylesterase in different model systems yielded poor results: when expressed in *E.coli* the protein was inactive, probably due to the lack of post translational modifications in prokaryotes; and, although yeast cells are capable of glycosylation, production of active protein was achieved only if the signal peptide was altered because the signal peptide needs to be specific to the species in order to be recognized by the cell (Morton and Potter, 2000).

In an attempt to increase purification efficiency of recombinant hCES2, the human recombinant *CES2* gene was synthesized with a C-terminal 10 histidine tag (10xHis) and inserted in a pCI-neo expression vector – pCI-neo hCES2-10xHis. The 10xHis tag masked the H-T-E-L ER retention sequence, resulting in the secretion of the enzyme to the extracellular medium when HEK 293T suspension adapted cells were transfected. Analytical tests were performed using the hCES2-10xHis protein, which lead to the conclusion that it had a similar activity profile to recombinant hCES2 and that it was secreted through the classical pathway (Lamego *et al*, 2013).

To confirm processing in the Golgi, hCES2-10xHis was deglycosylated using two different enzymes: Endo H, which hydrolyses high mannose oligossacharydes, and PNGaseF, responsible for the hydrolysis of complex type glycans. Through Western Blot, a loss of molecular weight was only

visible when the protein was deglycosylated with PNGaseF – hCES2-10xHis suffers more glycan processing when it passes through the Golgi, making its glycans complex type, which are not sensitive to Endo H. This does not happen with the wildtype hCES2, since it remains in the ER instead of following the secretion pathway (Lamego et al, 2013).

1.3 In vitro protein expression

In vitro systems are used for protein expression and bioproduction of several therapeutic agents because they are more controllable and reproducible then systems based on laboratory animals (Hartung *et al*, 2002; Merten, 2006).

Animal cells play a crucial role in the synthesis of bioproducts, ensuring the post-translational modifications of proteins. The glycosylation profile, for example, must be consistent, although it can be affected by nutrient depletion, pH, temperature or oxygen in the cell culture (Butler, 2005; Merten, 2006).

Continual lineages, which consist of immortalized cell lines, have the advantages of being multiplied, expanded and cryopreserved. As long as culturing conditions are kept constant, the yielded results are reproducible, once the cell population is homogenous and stable. However, the cell cultures do not represent the in vivo environment because the cells retain little differentiation (Hartung *et al*, 2002).

The culture medium for cell growth generally consists of a defined base solution that includes salts, sugar and amino acids and that can be mixed with bovine serum (Hartung *et al*, 2002). Bovine serum is rich in hormones, growth factors and albumin, which protect cells against potentially hostile conditions. Despite its advantages, the serum is a variability source because its composition is inconstant and undefined, leading to inconsistent cell growth. Serum-free medium bypasses these problems and simplifies downstream purification processes, because less animal proteins are present in the cell culture (Butler, 2005).

Standard use mammalian cells include Chinese Hamster Ovary (CHO), NSO mouse myeloma, Baby Hamster Kidney (BHK) and Human Embryonic Kidney (HEK 293). In this work, suspension adapted HEK 293T cells were used for production of hCES2-10xHis (Cunha, 2011).

The HEK 293 cell line was isolated from human embryonic kidney cells that were transformed with human adenovirus type 5 DNA (Ad5) (Thomas and Smart, 2005).

For protein production, the cells are transfected using plasmid vectors containing cytomegalovirus (CMV) promoter and the entire protein synthesis machinery is sequestered, forcing the translation of genes incorporated in the plasmid (Thomas and Smart, 2005). A HEK 293 variant,

HEK 293T, expresses the T SV40 antigen. This modification makes HEK 293T highly transfectable if plasmids with SV40 promotor are used.

HEK 293T cells show epithelial morphology and are usually adherent, but can be adapted to suspension using serum-free medium.

Plasmid vectors, which originate from extra-chromosomal circular DNA, are routinely used as carriers for the genes of interest into the animal cells. These vectors contain structural elements crucial to DNA insertion and replication in different hosts. For example, the replication origin is responsible for replicating the vector within the bacterial host, while a promoter leads to the transcription and translation of the inserted gene in the expression system. The DNA codifying for the protein of interest is inserted in the multiple cloning site, which is a cluster of restriction enzyme recognition sites located downstream from the promoter. Other elements may be present, such as selective markers or polyhistidine tags (Wong, 2009).

If the vector is compatible with mammal cells it can be directly incorporated in the cell through several transfection techniques, like calcium phosphate co-precipitation, electroporation or with cationic polymers (Wong, 2009; Castilho *et al*, 2009). Once the cell is transfected it begins to express the gene product of interest.

Molecular cloning consists of the manipulation of DNA with the objective of, among other things, produce vectors such as plasmids. This manipulation is done using bacterial systems since they have high growth rates, use inexpensive substrates and have well characterized genetics (Walker and Rapley, 2009).

The pCI-neo Mammalian Expression Vector from Promega contains the human cytomegalovirus (CMV) enhancer/promoter region for the constitutive expression of cloned DNA in mammalian inserts, and SV40 enhancer/promoter that not only replicates DNA in bacterial cells but also increases transfectability in HEK 293T cells. Among other characteristics it also includes a neomycin phosphotransferase resistance gene and can be used either for transient or stable transfection (Wong, 2009).

2 Objectives

Although hCES1 structure is known since 2003 (Bencharit *et al*, 2003), the three dimensional structure of hCES2 hasn't yet been resolved. Its understanding would be particularly helpful for the development of inhibitors (for reduction of side effects derived from early activation of pro-drugs for example) or new substrates, leading to advances in the efficacy of treatment with several drugs and pro-drugs, making combined therapies a possibility to maximize the effects of the administered drugs. With this in mind, all contributions for a better understanding of hCES2 structure and function are extremely helpful.

The main objective in this work was to study the glycosylation of hCES2 and its impact on enzymatic activity. To achieve that goal, hCES2-10xHis, a variant of hCES2, and mutated forms lacking one or both glycans were produced. Four different methods (Figure 2.1) were used for the production of partially or completely deglycosylated enzymes:

- enzymatic deglycosylation of hCES2-10xHis;
- inhibition of N-glycosylation within the cell using inhibitors of the N-glycosylation pathway, like tunicamycin;
- avoiding targeting of the protein to the ER, and consequently its glycosylation, through deletion of the signal peptide;
- and production of mutated forms without the glycosylation sites through the substitution of the asparagine in Asn-X-Ser/Thr by a glutamine.

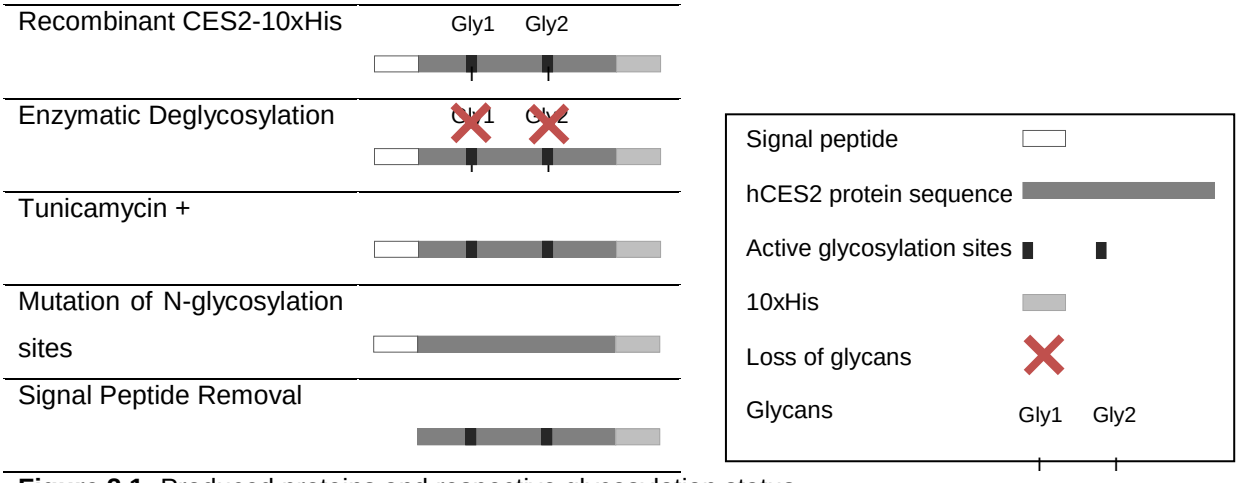


Figure 2.1: Produced proteins and respective glycosylation status.

The modified forms were tested for activity using usual hCES2 substrates.

Human CES2 is crucial in the metabolism of several xenobiotics. If the knowledge obtained from the study of hCES2-10xHis can be translated to human CES2 it will lead to a better understanding of the enzyme structure and functions. This could benefit some therapies that use its

substrates as drugs and pro-drugs. Moreover, the results of the present study would potentially be useful for devising methods for the crystallization of the human protein, which has eluded scientists for a decade.

3 Materials and Methods

3.1 Plasmid Production

3.1.1 Polymerase Chain Reaction (PCR)

PCRs were performed in a Biometra T3 Thermal Thermocycler (Germany), using Phusion High Fidelity Polymerase (Thermo Scientific, USA). Reaction buffers were either HF, recommended buffer with lower error rate, or GC, used for templates rich in guanines and cytosines, from the same brand. Maximum annealing temperature used was 71 ºC.

3.1.2 Site Directed Mutagenesis

Mutagenesis was performed in two consecutive reactions, using the GeneArt Mutagenesis Kit from Life Technologies (USA) that allows the insertion, deletion or substitution of up to 12 basepairs (bp) using reverse PCR.

Two sets of primers were designed according to the kit's specifications with the help of Vector NTI software (Life Technologies, USA), one for each glycosylation site. A first mutagenesis reaction was performed using pCI-neo hCES2-10xHis (map in appendix B-2) as template to mutate one glycosylation site. The new plasmid was extracted from the transformed bacteria and sequenced as described in section 3.1.6, and then it was used for the second mutagenesis reaction, resulting in a new plasmid with the two glycosylation sites mutated. (Diagram in Appendix C-1)

Table 3.1: Primers designed for the mutagenesis reactions. The first two pairs used different guidelines than the others. T_m: melting temperature; GC – Guanine Cytosine Content.

Primer Name	Sequence (5'-3')	T_m (ºC)	GC (%)
Glyco1 Fw	AGTTTCTTAGCCAGTTCCAGATGACCTTCCCTT	66.3	45.5
Glyco1 Rv	TCAAAGAATCGGTCAAGGTCTACTGGAAGGGAA		
Glyco2 Fw	ATCTCCACGGTGGTGGCCCAGCTGTCTGCCTGT	78	63.6
Glyco2 Rv	TAGAGGTGCCACCACCGGGTCGACAGACGGACA		
GS1 Fw	TAGCCAGTTCCAGATGACCTTCCCTTCCGACTCCATGTCT	76.1	52.3
GS1 Rv	GGAAGGTCATCTGGAACTGGCTAAGAAACTCTGACTCCACT	72.5	48.8
GS2 Fw	GGTGGTGGCCCAGCTGTCTGCCTGTGACCAAGTTG	78.7	62.9
GS2 Rv	ACACCGAGACAGCTGGGCCACCACCGTGGAGATGACA	81.4	62.2

3.1.3 Signal Peptide Removal

The gene of interest, *hCES2-10xHis* without the signal peptide, was amplified, purified from an agarose gel and digested with *Sal*I and *Not*I restriction enzymes (New England BioLabs, USA). The pCI-neo vector (Promega, USA) was linearized using the same enzymes, purified and dephosphorylated using Antarctic Phosphatase (New England BioLabs, USA). Next, the digested insert and dephosphorylated vector were ligated using two different methods: T4 Ligase and Quick Ligation Kit, both of which are manufactured by New England BioLabs, USA. Cells were transformed and plated.

Another approach was performed, cloning the gene of interest into a new pCI-neo (Promega, USA) vector using the InFusion HD Cloning Kit from Clontech, Japan. This technology is based on the InFusion enzyme that fuses linearized DNA fragments as long as they have an overlapped 15 bp sequence, which can be achieved through the correct designing of the PCR primers. Once the fragment has been amplified it is then fused with the linearized vector. (Diagram in Appendix C-2)

The PCR primers were designed according to the specifications of the kit with the help of Vector NTI software (Life Technologies, USA), and the reaction was set up using Phusion High Fidelity Polymerase (Thermo Scientific, USA) and the optimized annealing temperature.

Table 3.2: Primers designed for signal peptide removal. T_m: melting temperature; GC – Guanine Cytosine Content.

Primer Name	Sequence (5' – 3')	T_m (ºC)	GC (%)
Signal Fw	AATGTCGACTACAGCCACCATGCAGGACTCAGCCAGTCCC ATCCGGACCA	86.2	58
Signal Rv	AGTGGTAGTGGTGGTGATCCTCGAGCCGGCGGCTCAAAGG GAAA	84.1	59.1
SP Fw	TACCTCTAGAGTCGACGCCACCATGCAGGACTCAGCCAGT CCCAT	59	57.7
SP Rv	TAAAGGGAAGCGGCCGCCTAGTGGTGGTGA	54	65

3.1.4 DNA Purification

Whenever DNA needed to be purified, whether it was in solution or in agarose gel, the illustra GFX PCR DNA and Gel Band Purification Kit from GE Healthcare, UK, was used following manufacturer's instructions. This kit uses a chaotropic agent to dissolve agarose, extract DNA and denature proteins. The DNA then binds to the silica membrane inside the spin columns and is washed to remove any contaminants, until it can be eluted.

3.1.5 Bacterial Transformation

Unless otherwise stated, Stellar competent cells (Clontech by Takara Bio Company, Japan) were transformed using 5 ng of purified DNA following the protocol provided by the manufacturer and plated in Fast Media Amp Agar (InvivoGen, USA).

3.1.6 Plasmid Extraction, Quality Assessment and Sequencing

Transformed bacteria were grown overnight in Fast Media Growth Amp Agar (InvivoGen, USA) and selected colonies were transferred to Fast Media Amp TB (Invivogen, USA) also overnight. Plasmid DNA was extracted using either Genopure Plasmid Maxi Kit (Roche, Switzerland) or GenElute Plasmid Miniprep Kit (Sigma-Aldrich, USA): transformed bacteria are subjected to modified alkaline lysis, which results in plasmid DNA escaping to the supernatant while chromosomal DNA remains trapped inside the cell wall. The supernatant is then loaded onto a column where the DNA is adsorbed and, after the removal of contaminants through washing steps, is eluted. The obtained DNA is pure and free of RNA contamination.

To assess the quantity and quality of the plasmids, a sample was analyzed in a NanoDrop (Thermo Scientific, USA). The Abs260nm/Abs280nm ratio determined the quality of the plasmid, purity wise.

Enzymatic restriction assays confirmed that the plasmid was the expected size. The DNA and enzymes were incubated, in the correct proportions and buffers, at 37ºC for 1 hour (h). The assays were analyzed in 0.7 % (w/v) agarose gels (Lonza, Switzerland) with 1:20000 RedSafe (ChemBio, UK) incorporated, which is a safer alternative to Ethidium Bromide. The electrophoresis was performed in a horizontal device with buffer TAE (Tris-acetate EDTA) 1X at 100 volts (V) for 70 minutes (min), and the gels were visualized in the Gel Doc XR+ from Bio-Rad, USA. For the samples with the desired size and number of fragments, sequencing was performed at Macrogen, The Netherlands. Three primers were used for sequencing: two of them were universal and recognized sequences of the vector on either side of the gene; the third was gene specific and its respective sequence wasn't altered by any of the performed mutations.

Table 3.3: Primers used for sequencing the extracted clones. Universal primers were from Macrogen (The Netherlands).

Primer name	Sequence (5' – 3')	Hybridization site
T7 Universal	AATACGACTCACTATAG	T7 promoter, upstream to the gene
EBV RP Universal	GTGGTTTGTCCAAACTCATC	SV40 poly(a), downstream to the gene
CES2 Rv	ACATCAGCAGCGTTAACATTTTCTG	Gene specific

3.2 Cell Culture

3.2.1 Culturing Conditions

HEK 293T cells used, purchased from ATCC (CRL-11268), were adapted to suspension in-house, and cultured in 125 mL or 500 mL sterile glass Erlenmeyers (Schott AG, Germany) containing 20 mL or 80 mL of Freestyle 293 medium (Life Technologies, USA), respectively. The Freestyle 293 Expression Medium is serum free, animal origin free and chemically defined, having been developed for the culture and transfection of 293T cells in suspension. The medium was stored at 4 ºC, and protected from direct light.

The cells were sub-cultured two times a week when viable cell density was 2-3x10^6 cell/mL, using an inoculum of 0.3x10^6 cell/mL and for every 3-4 passages the cellular suspension was centrifuged for 10 min at 200 *g*; the pellet was resuspended in fresh medium. After carefully homogenizing the cellular suspension, the calculated volume of cells was added onto a sterile erlenmeyer containing the corresponding quantity of fresh medium, heated to 37 ºC. The cultures were then incubated in a humidified atmosphere of 8 % CO_2 at 37 ºC, using an orbital agitation of 130 rpm.

To freeze the cells, the cellular suspension was centrifuged for 10 min at 200 *g* and 4 ºC, and the pellet was resuspended in CryoStor CS10 cryopreservation medium (Stemcell Technologies, Canada) to obtain a concentration of 10x10^6 cell/mL. This suspension was aliquoted in cryovials and frozen at -80 ºC in a “Mr.Frosty” Cryo 1 ºC Freezing Container (Nalgene Nunc, USA), which ensured the gradual decrease of temperature at the rate of 1 ºC per min. To verify if the freezing was efficient, a vial of cells was thawed in a 37 ºC bath, added to fresh medium and centrifuged (200 *g*, 10 min, 4 ºC). The supernatant was quickly replaced with medium, to avoid damaging the cells due to DMSO toxicity of the freezing solution. The pellet was resuspended in the medium and transferred to a 125 mL Erlenmeyer.

3.2.2 Cell Viability and Concentration

To determine cell viability and concentration the method of Trypan Blue exclusion was used – while living cells are impermeable to the dye, dead cells incorporate it and are stained blue. The cellular suspension samples were diluted in 0.1 % Trypan Blue (Life Technologies, USA) solution prepared in phosphate buffered saline and the cells were counted twice on a Fuchs Rosenthal haemacytometer.

3.2.3 Transfection Assays

HEK 293T cells were transiently transfected with different expression vectors in 125 mL erlenmeyers containing 20 mL of cellular suspension. Poly(ethylenimine) (PEI) was used as the transfection agent, transporting DNA into the cell and releasing it into the cytoplasm.

The cells were subcultured three days prior to the transfection using an inoculum of $0.5x10^6$ cell/mL, and on the day of the transfection were diluted to $1x10^6$ cell/mL and incubated for 3-6 h. A transfection solution of 1mL containing fresh medium and DNA:PEI at a ratio of 1:3, in which the concentration of plasmid DNA was 5 µg/mL, was vortexed, incubated for 10 min at room temperature and added drop wise to the culture. This process had already been optimized in our group (Cunha, 2011; Lamego, 2013).

3.2.4 Tunicamycin

One cell culture transfected with pCIneo hCES2-10xHis was centrifuged 24 h post transfection (hpt) and the pellet was resuspended in Freestyle medium with 2 µg/mL (Oda *et al*, 2006) of Tunicamycin A1 (Sigma-Aldrich, USA), to inhibit protein N-glycosylation. The cultures were then incubated for 24 h (Powell, 2001). Samples were taken at different time points to assess cell viability and the supernatant was stored at -20 ºC.

3.3 Enzymatic Deglycosylation

Supernatant containing hCES2-10xHis protein was concentrated using Amicons Ultra-15, PLTK Ultracel-PL Membrane, 30 kDa (Merck Millipore, USA), in order to reduce the volumes needed. The increase in concentration was monitored through Bradford Assays.

The concentrated samples were deglycosylated using 50 mU of PNGaseF (ProZyme, USA) under native conditions, or 10 mU under denaturing conditions.

For native deglycosylation only the provided reaction buffer and enzyme were added to the sample, according to manufacturer's instructions. The denaturing deglycosylation required the addition of reaction buffer and denaturation solutions to the sample, which was then heated to 100 ºC for 5 min. After denaturation, detergent solution and enzyme were added.

For positive control RNaseB (Sigma-Aldrich, USA) was used, and reactions which lacked the PNGaseF enzyme were used as negative controls. Every reaction was incubated for an overnight period at 37 ºC.

ProQ Emerald 300 (Life Technologies, USA) was used to incubate polyacrilamide gels in order to detect the deglycosylated form of the protein. The stain reacts with periodate-oxydized carbohydrate groups, emitting a fluorescent green signal in glycoproteins that can be observed in a 300 nm transilluminator.

3.4 Analytical Assays

3.4.1 Protein Quantification

Quantification of total protein present in cell culture samples was performed using the Bradford Assay in a 96 well plate. An acidic red dye was used (Coomassie Brilliant Blue G250, from Bio-Rad USA), which binds primarily to basic and aromatic aminoacid residues and turns to blue in response to different protein concentrations.

A calibration curve using bovine serum albumin (BSA) was prepared with concentrations ranging from 5 to 500 µg/mL. The calibration curve and samples (usually 10 µL) were added to the plate in triplicate. Absorbance was read at 595 nm in a SpectraMax 340 (Molecular Devices, USA) spectrophotometer 10 min after the dye was added. Total protein concentration was determined using the calibration curve.

3.4.2 Enzyme Activity Assays

For the evaluation of enzymatic activity of the expressed enzymes the hydrolysis of 4-methylumbelliferyl acetate (4-MUBA) to 4-methylumbelliferone (4-MUB) was measured either by spectrophotometry or spectrofluorimetry. Every assay, including controls, was performed in triplicate, so the results presented are the average values obtained. Non enzymatic hydrolysis was subtracted from the reaction values.

The substrate stock solution was prepared by diluting 16.36 mg of 4-MUBA (Sigma-Aldrich, USA) in 5 mL of absolute ethanol (15 mM). This solution was diluted to 7.5 mM in a mixture of 90 mM KH_2PO_4 and 40 mM KCl (pH 7.3), which was used as reaction buffer.

Reaction buffer and samples were loaded into a 96 well plate placed on ice. The substrate was added to each assay for a final concentration of 5 mM in a total reaction volume of 250 µL, and the plate was read immediately after in a SpectraMax 340PC spectrophotometer (Molecular Devices, USA). Reading was performed in 30 second (s) intervals for 15 min at 37 ºC and a 350 nm wavelength.

When assaying the presence and activity of the expressed enzymes in different cell culture supernatants and a more sensitive detection method was needed, an FL800 Biotek Instruments Fluorescence Microplate reader (MTX Lab Systems, Inc., USA) was used. The substrate solution was prepared as mentioned above, but instead of absolute ethanol and reaction buffer, DMSO (Dimethyl Sulfoxide) and an HBSS/DMSO (Hank's Balanced Salt Solution, Life Technologies, USA) (mixture in a ratio of 1:1) were used respectively. A stock solution of 4-MUB (Sigma-Aldrich, USA) was also prepared in 5 mL DMSO and then diluted to different concentrations in HBSS/DMSO, which were then used to prepare the calibration curves. For each assay a final volume of 250 µL per well was used. Immediately before reading the substrate was added to the plate. Reaction occurred at 37 ºC, using excitation and emission filters at 350 and 450 nm for 15 min, in 50 s intervals.

3.4.3 SDS Page

For protein detection, electrophoresis was performed under native conditions. Ice cold ethanol was added to the samples in a ratio of 1:4 and incubated for 2 h at -20 ºC, precipitating the total protein. This step allows loading 50 µg of protein onto the gel using the smallest volume possible.

Total protein was resuspended in Sample Buffer and Reducing Agent and heated to 70 ºC for 20 min. Of each sample, 10 µL were loaded onto the wells of a NuPAGE BisTris 4-12 % acrylamide gel, using the X-Cell Surelock mini cell system and MES Running Buffer. Electrophoresis ran for 40 min at 200 V. All of the reagents and materials used were from Life Technologies, USA.

3.4.4 Western Blot

After electrophoresis the gel was transferred onto a PVDF or nitrocellulose membrane, using the Transblot Turbo Semi Dry Transfer System (Bio-Rad, USA) for 25 min at 25 V. The membrane was blocked in a solution 5 % (w/v) milk (skim milk for microbiology, Merck, U.S.A.) in 0.05 % (v/v) Tween (tween 20, Merck, U.S.A), Tris Buffered Saline (TBS, Sigma-Aldrich, U.S.A.) overnight.

The membrane was incubated for 1 h with the primary antibody goat anti-hCES2 (R&D Systems, USA) diluted 1:200, followed by 1 h in secondary antibody, HRP-mouse anti-goat, diluted 1:5000. The antibodies were diluted in the Tween TBS solution. In between incubations with the primary and secondary antibody the membrane was washed using Tween TBS.

Since the secondary antibody was horseradish peroxidase conjugated, the membrane was developed using a chemilluminescent detection method – ECL Amersham (GE Life Sciences, UK). Once the membrane was inside the ChemiDoc (Bio-Rad, USA), the two reagents were mixed and added to its surface. At least 3 pictures of each membrane were acquired at different exposure times.

4 Results and Discussion

4.1 Assessment of the importance of N-glycan residues through enzymatic deglycosylation

Native human CES2 has two potential N-glycosylation sites, but no O-glycosylation has been described since the protein is retained in the ER (Pindel *et al*, 1997). These glycosylation sites are characterized by an Asn-X-Ser/Thr sequon, in which X is any aminoacid, although acidic residues such as aspartate or glutamate may reduce the efficiency of glycosylation (Varki *et al*, 2009). In hCES2 these residues are located at Asn103 and Asn267 as previously mentioned.

The main objective in this work was to determine whether the glycan residues affect enzymatic activity of hCES2 using recombinant hCES2-10xHis as a model. The effect of glycosylation on rat liver hydrolases had already been tested (Morgan *et al*, 1994) with a slightly different approach.

For this experiment pCI-neo hCES2-10xHis plasmid, which had already been designed and used (Lamego *et al*, 2013), was extracted from bacteria. To verify the quality of the extracted DNA, a sample was used in a digestion reaction, using *Xho*I restriction enzyme. The reaction volume was loaded onto a 0.7% agarose gel, using non digested DNA as control. Figure 4.1 shows the result of the gel electrophoresis.

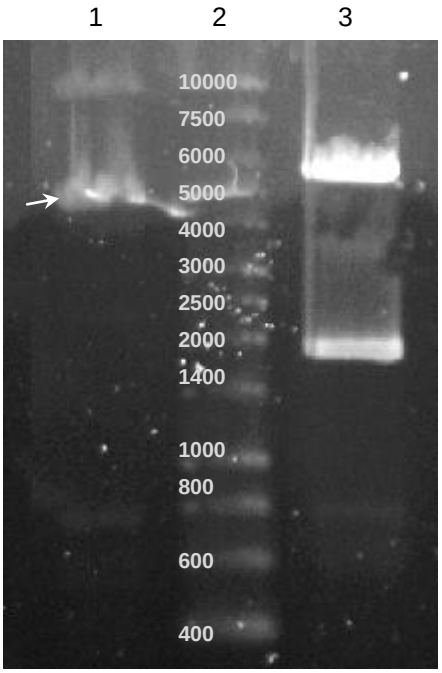


Figure 4.1: Gel electrophoresis of pCI-neo hCES2-10xHis in agarose gel after digestion with *Xho*I. Lane 1: non digested pCI-neo hCES2-10xHis, indicated by the white arrow; Lane 2: NZY III 1kb Ladder (masses indicated in the figure are in bp); Lane 3: pCI-neo hCES2-10xHis digested with *Xho*I.

For the control sample a single band in the 7.3 kb region was expected, but its migration pattern was slightly altered probably due to its non-linearized conformation – indicated by the white arrow in figure 4.1. The DNA sample digested with *Xho*I had two bands at 5.5 kb and 1.8 kb, as expected according to the restriction sites in the gene, indicating that the desired plasmid was present and had the correct size and digestion profile. A sample was also analyzed through spectrophotometry (NanoDrop), showing a plasmid concentration of 1287 ng/µL and an Abs_{260nm}/Abs_{280nm} ratio of 1.8. These values indicate that the DNA extraction was successful and that no protein contamination was present. Once the quality of the plasmid DNA had been established, HEK 293T cells were transiently transfected as previously described. After 96 h, the entire cell culture was centrifuged; the supernatant, containing the recombinant hCES2-10xHis, was collected and an enzymatic reaction assay was performed to ensure that the protein was present and active (Figure 4.2). The supernatant was then concentrated as described in the Methods section. The final total protein concentration was approximately 635 µg/mL., while the initial concentration was 89 µg/mL (Figure 4.3).

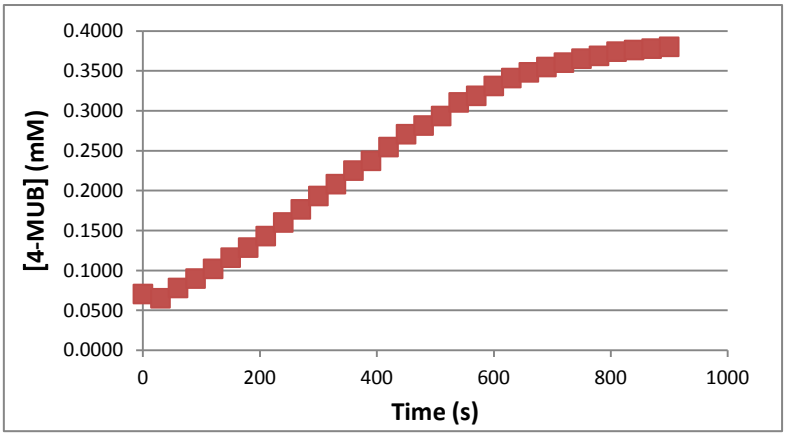


Figure 4.2: Enzymatic activity measured in the supernatant of the transfected cells using 4-MUBA as substrate. Product formation is indicative that the transfection was successful, since hCES2-10xHis is present and active. 4-MUBA – 4-methylumbelliferyl acetate; 4-MUB – 4-methylumbelliferone.

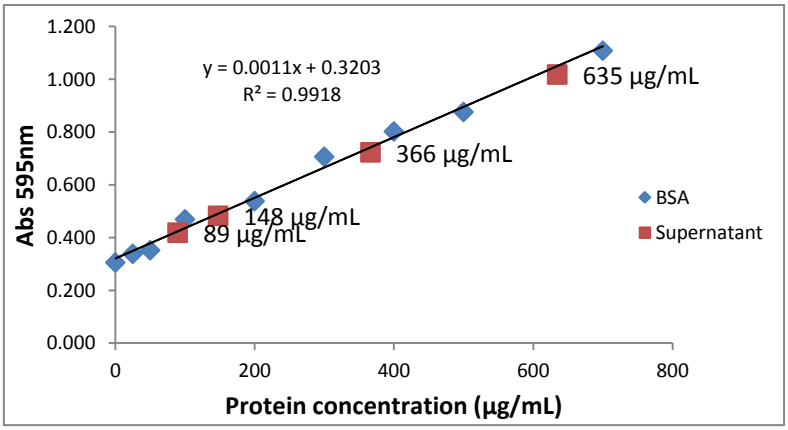


Figure 4.3: Bradford Assay for total protein quantification. Calibration curve defined using Bovine Serum Albumin. The supernatant samples were from four different points during the protein

concentration process, showing an increase in total protein concentration with the number of concentration steps.

The first experiment designed to evaluate the impact of glycosylation on enzyme activity consisted in removing the glycans from hCES2-10xHis through enzymatic deglycosylation, using PNGaseF. PNGaseF was chosen since the N-glycans present in this recombinant form of the protein are complex type instead of high mannose, as would happen in the wildtype (Lamego et al, 2013).

Deglycosylation reactions usually start with a protein denaturing step. However, since it was important to preserve protein integrity for the activity assays, four different reactions were prepared using the concentrated supernatant: two native reactions and two denaturing reactions, which served as controls. The enzyme was only added to one of each (PNGase+), while Reaction Buffer was added to the other (PNGase-) in the corresponding volume in order to maintain protein concentration. The reactions were then incubated overnight at 37 ºC. Immediately after the incubation period an enzymatic reaction assay was performed, using 5 µL of each native reaction volume.

To confirm the efficiency of the enzymatic deglycosylation every sample was subjected to SDS-Page/Western Blot, using anti hCES2 as primary antibody, as described in the Methods section. The results (Figure 4.4) show that there was a decrease in the molecular weight of the proteins incubated with PNGaseF (PNGase+)

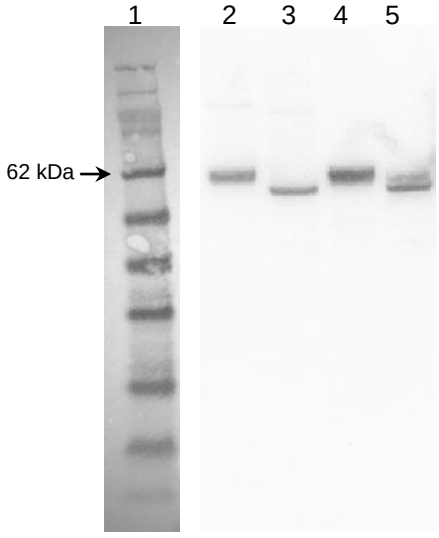


Figure 4.4: Western Blot of the enzymatic deglycosylation samples of hCES2-10xHis. Lane 1: SeeBlue Plus 2 Pre-stained Standard (Life Technologies, USA); Lane 2: denatured sample, PNGase-; Lane 3: denatured sample, PNGase+; Lane 4: native sample, PNGase-; Lane 5: native sample, PNGase+. There was a decrease in the protein molecular weight when incubated with PNGaseF.

There was a noticeable decrease in molecular weight of the samples that were deglycosylated. However, the protein in the native reaction appears to be only partially deglycosylated

– there is a slight smear in the band which shows that the native reaction wasn't as efficient as the denaturing one as expected, since the removal of N-glycans can be more difficult if the protein is in its native conformational state.

To verify if the protein remained active despite the loss of its glycans, an enzymatic activity assay was performed with both samples from the native reactions – PNGase+ and PNGase-. The kinectic curve obtained is represented in Figure 4.5.

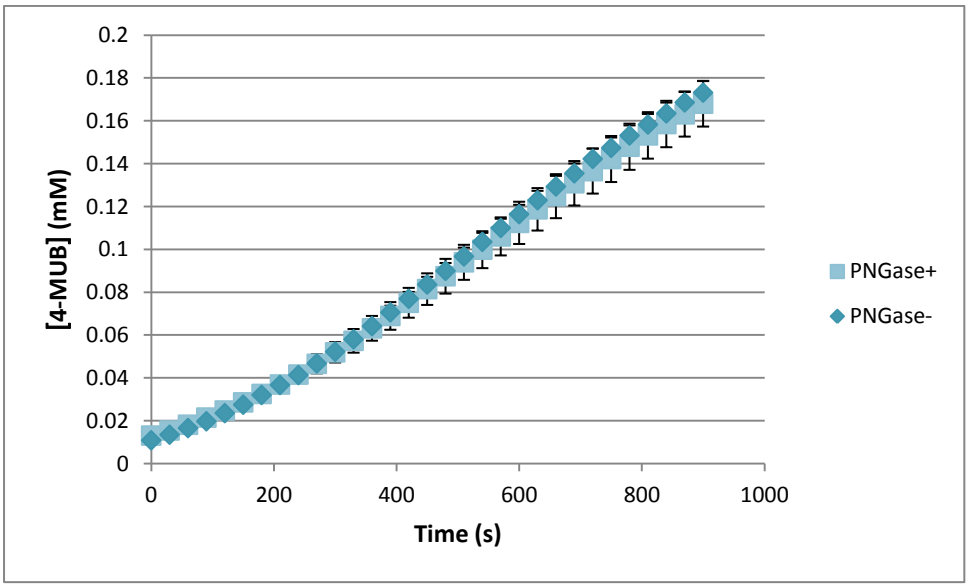


Figure 4.5: Enzyme kinectics of deglycosylated and non deglycosylated hCES2-10xHis, using 4-MUBA as substrate. Both samples analyzed were from native reactions in the presence (PNGase+) and absence (PNGase-) of the deglycosylating enzyme. Both reactions appear to have similar activity profiles. 4-MUBA – 4-methylumbelliferyl acetate; 4-MUB – 4-methylumbelliferone.

The previous figure (4.5) shows that both samples have similar enzymatic activity profiles, indicating that the N-glycans may not be essential for hCES2-10xHis activity. The slope of each curve was determined in the linear region, from 300 to 700 s. While the original sample (PNGaseF-) afforded a specific activity of 93.9 ± 3.7 pM s^{-1} μg^{-1}, the sample incubated with PNGase (PNGase+) afforded 88.9 ± 6.4 pM s^{-1} μg^{-1}. This appears to indicate that the presence of the N-glycans in the mature form of the protein (fully processed through the cell machinery) is not essential for protein activity. However the fact that the enzymatic deglycosylation was not complete as previously shown, could also be affecting this observation.

In order to evaluate whether more enzyme or a longer incubation time could improve the efficiency of the native reaction, the deglycosylation reactions were repeated. Incubation at 37 ºC lasted for 24 h, and at 18 h more enzyme or Reaction Buffer were added to each reaction. SDS-Page and Western Blot was performed with the new reactions, but the bands showed no obvious difference from the previous analysis (Figure 4.6).

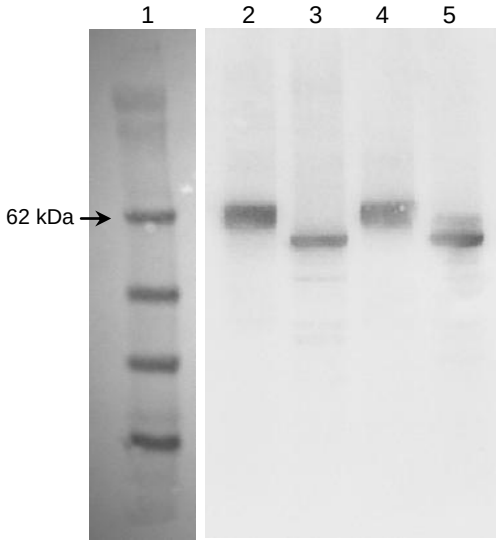


Figure 4.6: Western Blot of the enzymatic deglycosylation samples of hCES2-10xHis subjected to a longer incubation period and more deglycosylating enzyme. Lane 1: SeeBlue Plus 2 Pre-stained Standard (Life Technologies, USA); Lane 2: denatured sample, PNGase-; Lane 3: denatured sample, PNGase+; Lane 4: native sample, PNGase-; Lane 5: native sample, PNGase+. Although the protein under native conditions also showed a decrease in molecular weight when incubated with PNGaseF, some smear is visible.

The same results were obtained despite the increase of incubation time and amount of deglycosylating enzyme used, indicating that the remaining protein probably was not deglycosylated due to its native conformation limiting access of the enzyme to the glycosylation site. To verify whether hCES2-10xHis was completely deglycosylated in its native conformation, a different method of glycosylation detection was used.

Deglycosylation reactions were prepared using purified hCES2-10xHis available in the lab instead of the concentrated supernatant in order to allow glycoprotein staining without interference by other glycoproteins eventually present in the supernatant. Since the protein was purified, only 5 µg were loaded into each well for the SDS-Page. After electrophoresis, the gel was stained with ProQ Emerald 300 and the bands were detected using a 300 nm UV Transilluminator.

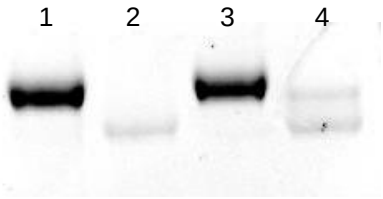


Figure 4.7: Acrylamide gel of the enzymatic deglycosylation stained with ProQ Emerald 300. Lane 1: denatured sample, PNGase-; Lane2: denatured sample, PNGase+; Lane 3: native sample, PNGase-; Lane 4: native sample, PNGase+

As expected in this method, the glycoproteins appear stained (lanes 1 and 3), while the deglycosylated proteins (lanes 2 and 4) only show faint bands (due to nonspecific binding), indicating that the protein lost its glycan residues. However, two bands are visible on lane 4, which could mean that either the deglycosylation was incomplete, although a more visible band was expected if this were the case, or that the protein started to degrade in the presence of the deglycosylation reagents, which may also have happened using the supernatant samples.

Since, from the stained gel, it was unclear whether the native deglycosylation of hCES2-10xHis was complete, new enzymatic assays were performed to evaluate if product formation was in the linear range. If it is established that there is no saturation, then conclusions can be made as to whether glycosylated and deglycosylated proteins are mixed in the sample or not. The same amount of sample was added to a 96 well plate in triplicate, as well as half and a fourth of that quantity – 2.3, 1.15 and 0.58 µg of total protein.

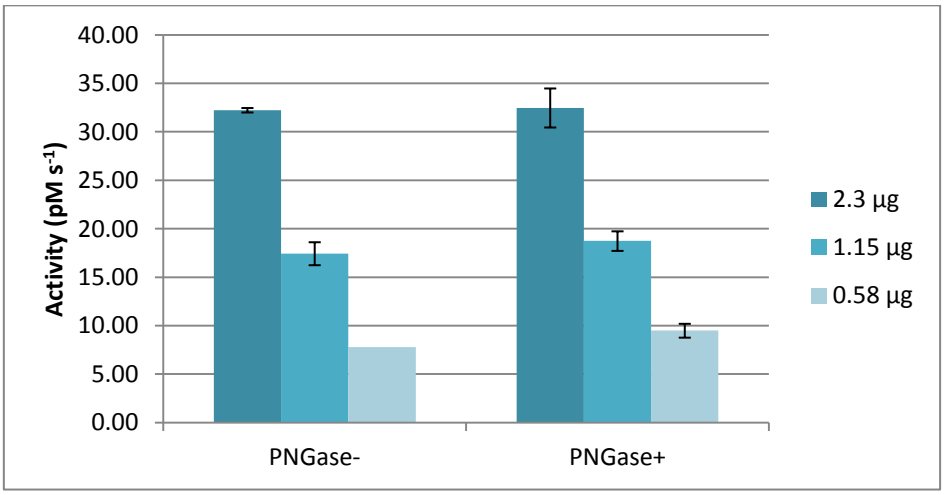


Figure 4.8: Graphic representation of the activity of each enzymatic reaction in the conversion of 4-MUBA, with different total protein quantities. A direct correlation between the amount of total protein used and the resulting activity is observed. 4-MUBA – 4-methylumbelliferyl acetate.

Figure 4.8 shows that there is a direct correlation between the amount of sample used in the enzymatic reactions and the resulting activity, indicating that the reaction was performed in the linear range. If hCES2-10xHis activity had saturated, a possible decrease in activity due to deglycosylated protein in the sample would not be proportional. As the activities of both samples (PNGase+ and PNGase-) are similar and in the linear range, it does appear that there are no glycosylated and deglycosylated proteins mixed in the same sample – the activity profile observed is solely due to one form of the protein.

The last enzymatic activity assay, along with the Western Blots and glycoprotein stain, indicate that the native deglycosylation was efficient and that hCES2-10xHis remains active despite the

removal of its N-glycans. This result is in agreement with the hypothesis raised after the deglycosylation of rat liver hydrolases (Morgan *et al*, 1994) that the glycan residues are not essential for enzyme activity.

However, with this experiment, it remains to be verified if glycosylation has an impact on activity by allowing correct protein maturation. This was investigated through glycosylation inhibition, as described in the next sections.

4.2 Inhibition of N-glycosylation during protein expression using Tunicamycin

Once it was verified that the N-glycans are not essential to hCES2-10xHis activity in a mature protein, the next step was to inhibit N-glycosylation. This inhibition should result in a non-glycosylated protein, which is expected to be inactive (Kroetz *et al*, 1993).

The inhibition of N-glycosylation in the cell was performed using tunacamycin. Tunicamycin is a mixture of homologue antibiotics that prevent the formation of the N-glycan precursor GlcNAc-P-P-Dol consequently inhibiting N-glycosylation during protein synthesis (Varki *et al*, 2009). Still, this does not avoid that the protein proceeds through the classical secretory pathway.

The assay was performed as described in the Methods section, and the cell cultures were labeled as to the presence (TM+) or absence (TM-) of the inhibitor in the medium. Since tunicamycin is stated as cytotoxic (Varki *et al*, 2009; Oda *et al*, 2006), cell viability was assessed. Increase in cell death between the two cultures was not significant, indicating that cell viability was not compromised by the addition of tunicamycin at a concentration of 2 µg/mL.

Table 4.1: Cell viabilities verified at different time points in the presence and absence of tunicamycin (2 µg/mL). An error rate of 15% was taken into account. hpt – hours post transfection; TM – hours post addition of tunicamicin.

	24 hpt	**4h TM**	**20h TM**	**24h TM**
TM-	85 ± 13%	84 ± 13%	79.5 ± 12%	78 ± 12%
TM+	88 ± 13%	73 ± 13%	80 ± 12%	84 ± 13%

Supernatant samples from 4, 20 and 24 h post addition of tunicamycin were quantified for protein and loaded into a SDS-Page acrylamide gel, using 25 µg of total protein per well. After the transfer the membrane was incubated with anti-hCE2 primary antibody.

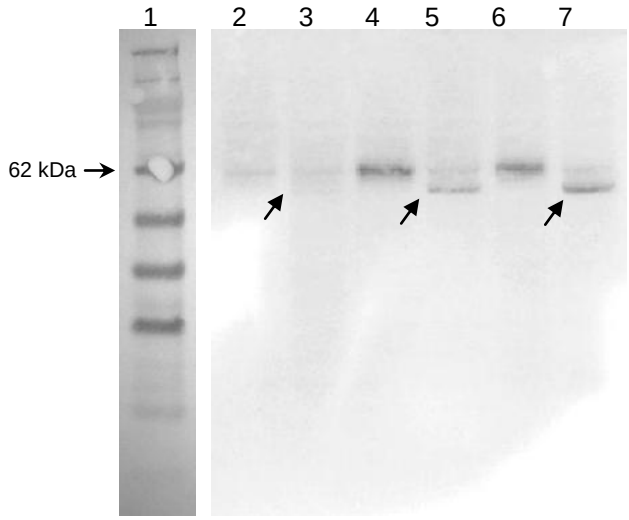


Figure 4.9: Western Blot of supernatant samples from the TM- and TM+ cell cultures. Lane 1: SeeBlue Plus 2 Pre-stained Standard (Life Technologies, USA); Lane 2: TM- 4 h; Lane 3: TM+ 4 h; Lane 4: TM- 20 h; Lane 5: TM+ 20 h; Lane 6: TM- 24 h; Lane 7: TM+ 24 h.TM+ samples are indicated in the membrane with an arrow. TM+ is indicative of supernatant samples from cell cultures incubated with tunicamycin, while TM- represents control samples.

Initially, as seen in Figure 4.9, both cultures show a faint band that corresponds to the expressed protein – hCES2-10xHis. With time, a shift is noticeable in the molecular weight of the protein between TM- and TM+ samples, especially at the 20 and 24 h time points (lanes 3-6). This shift could indicate that initially the TM+ culture was expressing glycosylated hCES2-10xHis, but that after the tunicamycin entered the cell and inhibited glycosylation, hCES2-10xHis stopped being glycosylated, leading to a decrease in molecular weight as expected, which wasn't visible in the TM- samples.

Knowing that the addition of tunicamycin to the cell culture led to a decrease in molecular weight, attributable to N-glycosylation inhibition, and that this procedure had no significant effect on cell viability, the same samples were used for enzymatic activity assays. To evaluate whether hCES2-10xHis synthesized in the presence of tunicamycin was active, activity assays were performed as described in the Methods section, using 150 µL of supernatant from the 24 hpt cell cultures and from 24 h post the addition of tunicamycin which equaled to 48 hpt.

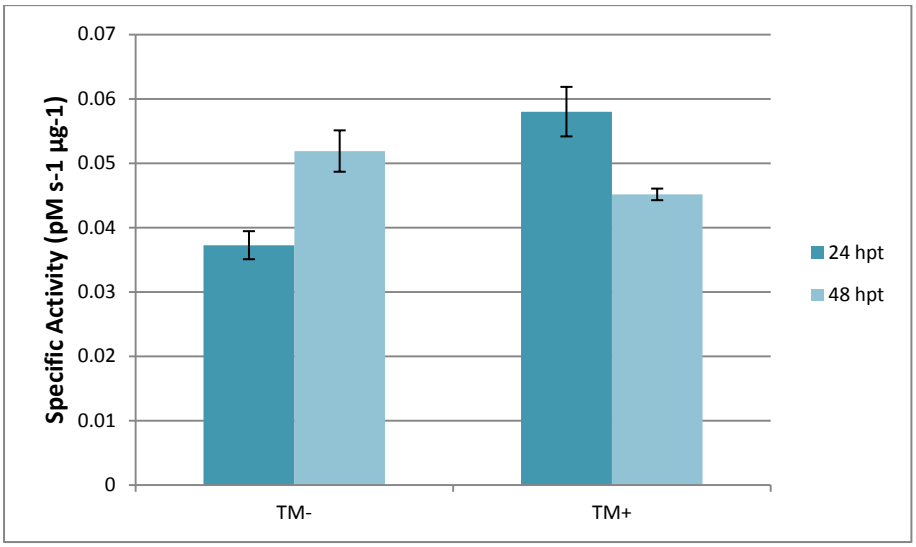


Figure 4.10: Graphic representation of the specific activity of hCES2-10xHis present in the supernatant of cell cultures treated or not with tunicamycin at two time points: 24 hpt and 48 hpt. TM+: cell culture treated with tunicamycin. TM-: cell culture not treated with tunicamycin. Results are average of three determinations. Error bars represent the standard deviation of the triplicates.

According to the results shown (Figure 4.10) a decrease in the specific activity is visible along time in the TM+ sample, while the specific activity of the culture with no tunicamycin increased. A more accentuated loss of activity was expected in TM+ supernatant: it is possible that inhibitor concentration and incubation time used were insufficient to affect its enzymatic activity – it has previously been stated that the efficiency of tunicamycin depends on those parameters (Kroetz *et al*, 1993). Another explanation could be the presence of glycosylated hCES2-10xHis synthesized between 0 and 4h post inhibitor addition. Further experiments are necessary with different tunicamycin concentrations and earlier addition in order to confirm or disprove this hypothesis.

4.3 Inhibition of N-glycosylation through mutation of the asparagine residues at the glycosylation sites

Site directed mutagenesis was performed to mutate the aminoacid residues at the potential N-glycosylation sites of hCES2-10xHis, in order to inhibit its N-glycosylation. It was expected that, if the N-glycans weren't able to bind to the new form of the protein, the folding might be incorrect and therefore the enzyme would be inactive. This experiment should, in theory, be complementary to the tunicamycin assay.

To avoid N-glycosylation of the protein, the asparagine residues, to which the N-glycans bind, had to be mutated. Firstly, the N-glycosylation sites – sequon Asn-X-Ser/Thr – present in hCES2-10xHis sequence were identified: there are two N-glycosylation sequons in the hCES2-10xHis protein sequence at Asn175 and Asn340.

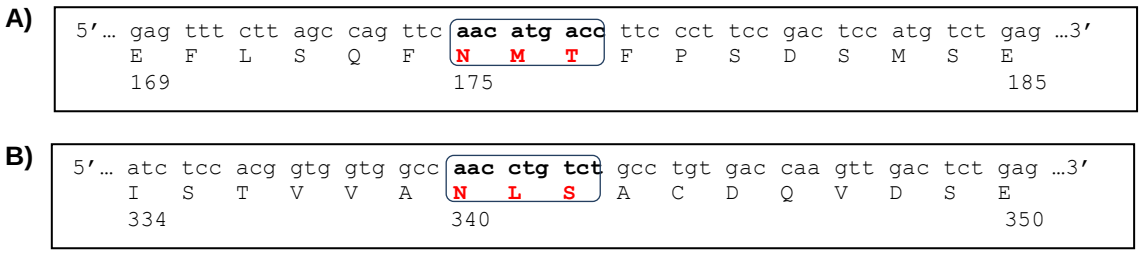


Figure 4.11: Potential N-glycosylation sites in the hCES2-10xHis codifying sequence and protein. A) First N-glycosylation site at Asn175; B) Second N-glycosylation site at Asn340. Full sequence available in Appendix A.

In order to inhibit glycosylation glutamine can be used to substitute asparagine since both residues are chemical and structurally similar: both are small and polar aminoacids with a neutral charge that exchange frequently, according to Dayhoff's mutation odds matrix (Taylor, 1986). Replacing one residue with the other should, in principle, destabilize protein structure and function as little as possible. There are two codons that codify for glutamine in the human genetic code, CAG and CAA, however the CAG codon has a higher usage frequency and was thus chosen (Kotlar and Lavner, 2006).

After identifying the glycosylation sites and target residues, mutagenic primers were designed. These primers have the same sequence as the gene, except for the mutated region. Once the PCR was set up, the primers amplified the entire template plasmid incorporating the desired mutation. According to the specifications from the Site Directed Mutagenesis kit, both primers of each pair would:

- Be approximately 30 to 45 nucleotides in length
- Have a centrally located mutation site
- Be 100% complementary with no overhangs

In addition to these guidelines others were taken into account, such as Guanine-Cytosine (GC) content between 40 and 60%, similar melting temperatures for each primer within a pair, and avoiding high melting temperatures and secondary structures (Sambrook and Russel, 2001). The primers were designed with the help of Vector NTI software (Life Technologies, USA).

Primer optimization of Glyco1 and Glyco2 pairs (sequences defined in the Methods section) started with a standard PCR using pCI-neo hCES2-10xHis as template and only varying the annealing temperature. A wide range of temperatures were tested for both pairs of primers, from 50 to 72 ºC, but no bands were visible when the products were analyzed in agarose gel electrophoresis (Figure 4.12) . To achieve a visible band other components of the reactions were modified, such as polymerase buffer, amount of DNA or polymerase, and number and duration of the program cycles.

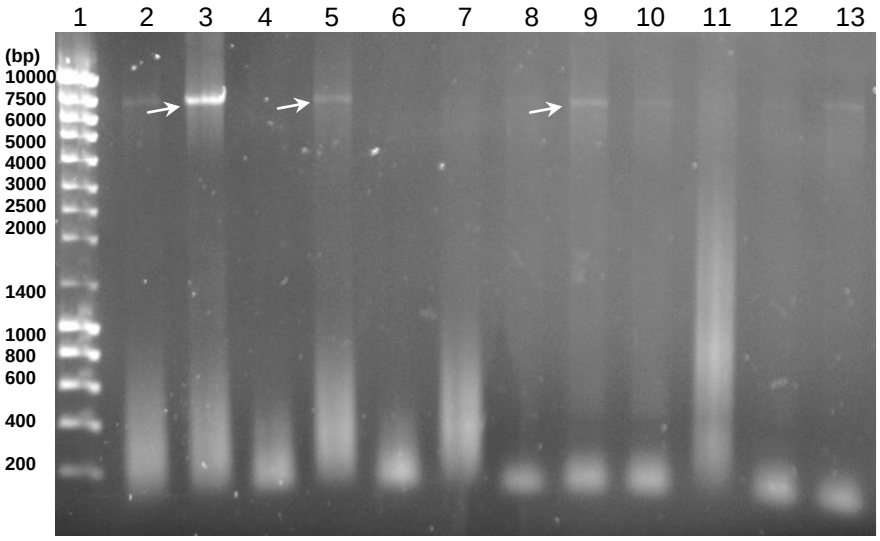


Figure 4.12: Agarose gel of primer optimization under different conditions. Lane 1: NZY III 1kb Ladder; Lane2:_Glyco1, Buffer HF, 65 ºC; Lane 3: Glyco1, Buffer GC, 65 ºC; Lane 4: Glyco1, Buffer HF, 68 ºC; Lane 5: Glyco1, Buffer GC, 68 ºC; Lane 6: Glyco1, Buffer HF, 71 ºC; Lane 7: Glyco1, Buffer GC, 71 ºC; Lane 8: Glyco2, Buffer HF, 65 ºC; Lane 9: Glyco2, Buffer GC, 65 ºC; Lane 10: Glyco2, Buffer HF, 68 ºC; Lane 11: Glyco2, Buffer GC, 68 ºC; Lane 12: Glyco2, Buffer HF, 71 ºC; Lane 13: Glyco2, Buffer GC, 71 ºC. White arrows indicate the bands corresponding to the intended product. Glyco1 – mutagenic primers for the first glycosylation site; Glyco1 – mutagenic primers for the second glycosylation site.

The best conditions found for the two pairs of primers were: Phusion GC buffer instead of HF, 20 ng of template DNA and 0.04 U/µL Phusion polymerase, with a 65 ºC annealing temperature and 35 PCR cycles, which corresponded to the conditions used for lanes 3, 5 and 9 (bands indicated with white arrow).

The protocol provided by the Site-Directed Mutagenesis System describes three different steps: (i) Methylation and Mutagenesis reactions, (ii) Recombination reaction and (iii) Transformation into DH5α-T1 *E.coli*. In order to assess the presence and quality of the mutated product, the DNA extracted from DH5α-T1 *E.coli* was digested with restriction enzymes, usually *Xho*I, and analyzed in an agarose gel after electrophoresis. When bands with the expected size and restriction pattern were present, the corresponding non-digested samples were sent to Macrogen for sequencing.

Initially the only changes made to the protocol were the optimized primer conditions and the addition of a DNA purifying step after the recombination reaction, but the product was not present. The DNA was digested with *Xho*I, but the bands visible in the agarose gel (Figure 4.13) weren't the expected size and didn't have the expected restriction profile, which should be similar to the pCI-neo hCES2-10xHis used as control; still a couple of samples were sequenced. The sequencing results showed that the *hCES2-10xHis* gene was not present.

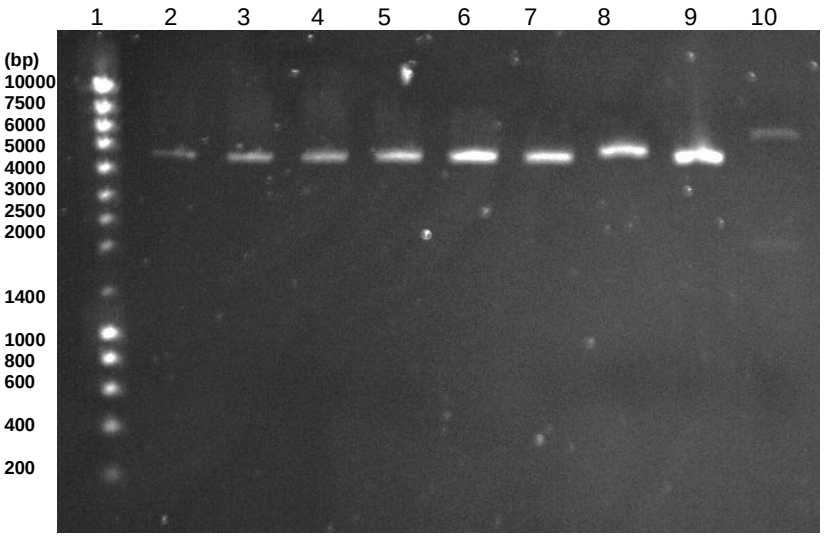


Figure 4.13: Gel electrophoresis of mutated plasmids digested with *Xho*I. Lane 1: NZY III 1kb Ladder; Lanes 2-9: mutated plasmid clones digested with *Xho*I; Lane 10: pCI-neo hCES2-10xHis digested with *Xho*I. The mutated plasmids do not correspond to the intended product, since they show a different restricition profile from the pCI-neo hCES2-10xHis plasmid.

For the next reaction new aliquots of template DNA and primers were prepared and the Phusion polymerase buffer was switched back to HF, which should have a lower error rate than GC. Two PCR programs were tested: the one that was optimized and the one described in the protocol. Despite these alterations the analyzed DNA did not correspond to the intended products.

Since no alteration to the protocol seemed to yield the desired results, the mutagenic primers were re-designed using slightly different specifications: instead of the sequence being 100% complementary, the new primers were designed to have overhangs of 10 to 20 nucleotides. This new approach resulted in the GS1 and GS2 pairs of primers, for the first and second glycosylation sites respectively, listed in the Methods section.

The optimization of the new primers used a standard PCR mix (with HF or GC buffer) and the recommended program. Three annealing temperatures were tested. While the first pair of primers showed the expected band after gel electrophoresis, the second pair did not. However, the same conditions were used for both pairs in future reactions: 62 ºC annealing temperature and HF Buffer.

For each pair of primers, to increase the probability that at least one of the reactions worked, two annealing temperatures were used in the mutagenesis reaction: 62 ºC (the optimized temperature), and 57 ºC (recommended temperature in the protocol). After the PCR, the reaction volume was purified and eluted in sterile H_2O. The recombination reaction was prepared without diluting the DNA and the provided DH5α-T1 cells were transformed. The plasmid DNA was extracted from the cells, digested with *Sma*I (Thermo Scientific, USA) restriction enzyme and analyzed by gel electrophoresis (Figure 4.14).

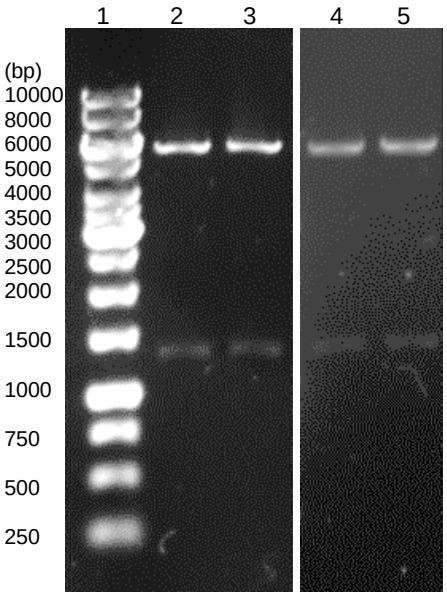


Figure 4.14: Gel electrophoresis of mutant clones digested with *Sma*I. Lane 1: Gene Ruler 1 kb DNA Ladder (Thermo Scientific, USA); Lane 2: mutagenesis product with GS1 primers, annealing temperature 57 ºC; Lane 3: mutagenesis product with GS1 primers, annealing temperature 62 ºC; Lane 4: mutagenesis product with GS2 primers, annealing temperature 57 ºC; Lane 5: mutagenesis product with GS2 primers, annealing temperature 62 ºC. All clones showed the expected restriction profile. GS1 – mutagenic primers for the first glycosylation site; GS2 – mutagenic primers for the second glycosylation site.

All samples showed the expected restriction pattern when analyzed in agarose gel – one band with 6 kb and the other with 1.3 kb. Two DNA samples of each reaction (M1 57 ºC, M1 62 ºC, M2 57 ºC and M2 62 ºC) were sequenced. All of the clones had the intended mutation, either the first or the second glycosylation site was mutated with the glutamine codon instead of asparagine, and no other mutations were present in the gene.

The second mutagenesis reaction was prepared, using the mutated pCI-neo hCES2-10xHis as template. If the template had the first glycosylation site mutated (pCI-neo hCES2-10xHis Glyco1) the primers used were for the second mutation site (GS2), so that the final plasmid had both glycosylation sites mutated. The same procedure was done for the clone with the second glycosylation site mutated (pCI-neo hCES2-10xHis Glyco2), in order to increase the probability of having one clone completely and correctly mutated. The same procedure from the first mutagenesis reaction was followed, and enzymatic restriction also used *Sma*I (Figure 4.15).

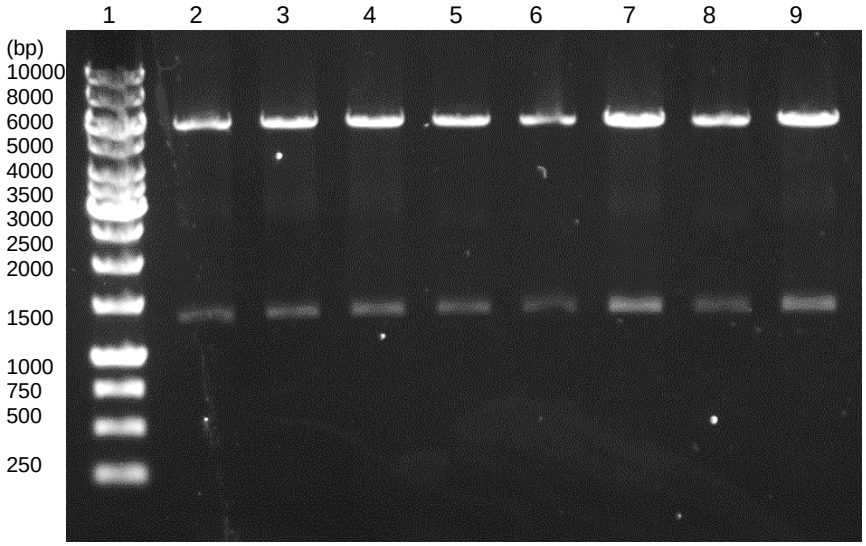


Figure 4.15: Gel electrophoresis of double mutant clones digested with *Sma*I. Lane 1: Gene Ruler 1 kb DNA Ladder (Thermo Scientific, USA); Lanes 2-5 : pCI-neo hCES2-10xHis Glyco1 + GS2 primers; Lanes 6-9: pCI-neo hCES2-10xHis Glyco2 + GS1 primers. All clones showed the expected restriction profile. GS1 – mutagenic primers for the first glycosylation site; GS2 – mutagenic primers for the second glycosylation site.

Since the restriction pattern was correct (6 kb and 1.3 kb bands), the samples were sequenced. Every clone had both glycosylation sites correctly mutated and no other mutations were detected.

This mutagenesis process lead to three new plasmids, all of which showed no other mutation in the gene except for the ones intended:

- pCI-neo hCES2-10xHis Glyco1, with the first glycosylation site mutated (Asn175 -> Gln175)
- pCI-neo hCES2-10xHis Glyco2, with the second glycosylation site mutated (Asn340 -> Gln340)
- pCI-neo hCES2-10xHis Glyco1+2, with both glycosylation sites mutated.

4.4 Inhibition of N-glycosylation through signal peptide removal

The signal peptide is responsible for targeting the protein to the endoplasmic reticulum. Once the protein is inserted in the ER, the signal peptide is cleaved by a specific protease, and other post translational modifications occur (Rapoport, 2007).

Taking this into consideration, one way of avoiding protein glycosylation is by deleting the signal peptide from the codifying sequence since the protein will not be targeted to the ER, bypassing the classical secretory pathway. With this procedure, it was therefore expected that the protein would remain in the cytoplasm (Schiel, 2009) and would not be glycosylated.

A signal peptide of 26 aminoacids in length has been described for human CES2 (Pindel et al, 1997, Schwer et al, 1997). This sequence starts at the second ATG codon from the first exon (Schiel, 2009).

```
5'…atg act gct cag tcc cgc tct cct acc aca ccc acc ttt ccc ggc cca agc cag cgc acc ccg
   M   T   A   Q   S   R   S   P   T   T   P   T   F   P   G   P   S   Q   R   T   P
   1                                                                               21

ctg act ccc tgc cca gtc caa act cca agg ctg ggc aag gca ctg atc cac tgc tgg aca gac
L   T   P   C   P   V   Q   T   P   R   L   G   K   A   L   I   H   C   W   T   D
22                                                                              42

ccg ggg cag cct ctg ggt gaa cag cag cgt gtc cgc cgg cag cga acc gag acc agc gag ccg
P   G   Q   P   L   G   E   Q   Q   R   V   R   R   Q   R   T   E   T   S   E   P
43                                                                              63

acc atg cgg ctg cac aga ctt cgt gcg cgg ctg agc gcg gtg gcc tgt ggg ctt ctg ctg ctt
T   M   R   L   H   R   L   R   A   R   L   S   A   V   A   C   G   L   L   L   L
64                                                                              84

ctt gtc cgg ggc cag ggc cag gac tca gcc agt ccc atc cgg acc aca cac acg ggg cag gtg
L   V   R   G   Q   G   Q   D   S   A   S   P   I   R   T   T   H   T   G   Q   V
85                                                                              105

ctg ggg agt ctt gtc cat gtg aag ggc gcc aat gcc ggg gtc caa acc ttc ctg gga att cca…3'
L   G   S   L   V   H   V   K   G   A   N   A   G   V   Q   T   F   L   G   I   P
106                                                                             126
```

Figure 4.16: hCES2-10xHis signal peptide of 26 aminoacid residues located downstream from the second methionine residue (in bold). Complete codifying and protein sequences of hCES2-10xHis in Appendix A.

In order to remove the signal peptide from recombinant hCES2-10xHis DNA, the gene was amplified from the end of the signal peptide to the stop codon and re-inserted in the pCI-neo vector.

The designed primers (sequence in the Methods section) had very specific characteristics. To ensure that the cloning would be successful, restriction enzyme sites were included so that after digestion of the amplicon its ends would match those of the linearized vector in which the same sites are located. The KOZAK sequence was added to warrant the correct reading frame upon protein synthesis.

- The forward primer (SignalFW) had a 5' non complementary sequence, which included a restriction site for *Sal*I (surrounded by other nucleotides), and the KOZAK sequence (GCCACCATG). The complementary sequence started downstream from the signal peptide.

- The reverse primer (SignalRV) was complementary do the plasmid sequence, beginning before the stop codon and ending after the *Not*I restriction site present in the vector.

Two PCR reactions were prepared with Phusion polymerase and both primers, but only one had pCI-neo hCES2-10xHis template. The optimized annealing temperature was 71 ºC. The reaction products were analyzed through gel agarose electrophoresis (Figure 4.17).

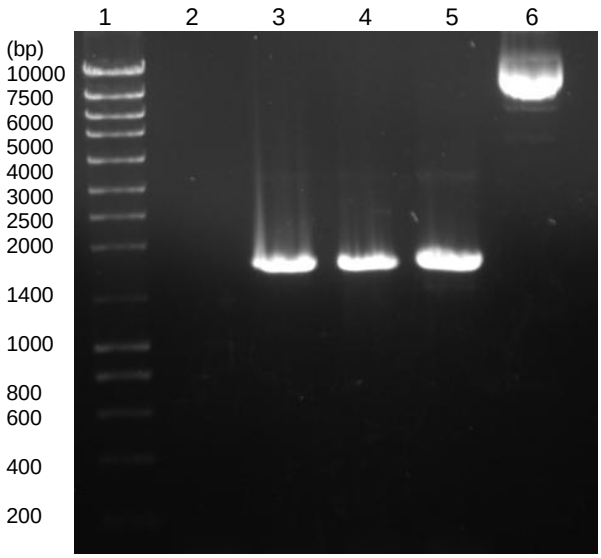


Figure 4.17: Agarose gel analysis of the PCR products. Lane 1: NZYTech III 1 kb Ladder; Lane 2: negative control, Lanes 3-5: reaction; Lane 5: pCI-neo hCES2-10xHis. Every product has the desired size.

The visible bands in the previous figure were the expected size, approximately 1.7 kb, but some smear was present. To ensure that no unspecific products would later be formed, the entire reaction volume was loaded into an agarose gel and purified. Simultaneously, pCI-neo vector was extracted from bacteria already present in the lab and its quality was assessed. Both the fragment and the destination vector were digested using *Sal*I and *Not*I restriction enzymes and afterwards purified.

Dephosphorylation and ligation reactions were performed, and bacteria were transformed. Plasmid DNA was extracted from bacteria and digested using *Bsr*GI and *Dra*III restriction enzymes (New England Biolabs, USA) that cut within the vector. Agarose gel electrophoresis is shown in the following figure.

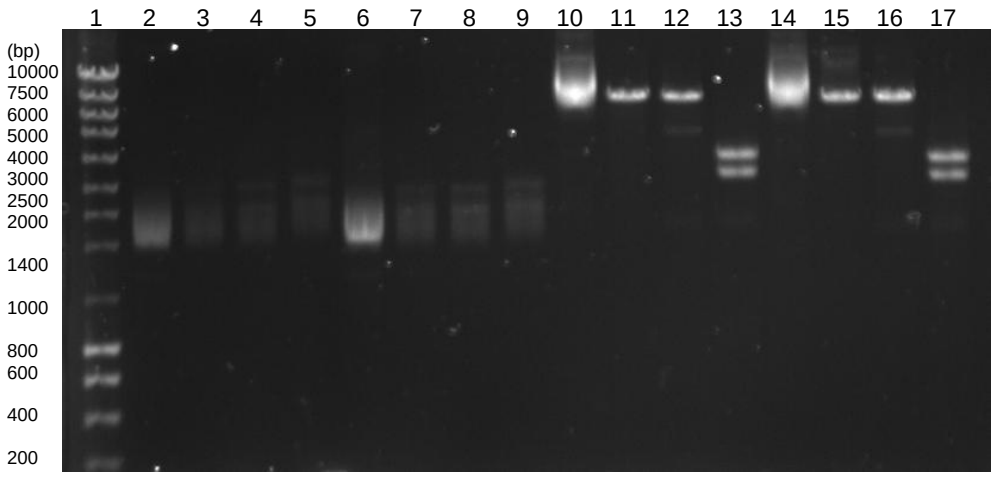


Figure 4.18: Digestion analysis of four different clones from the construction. Lane 1: NZYTech III 1 kb Ladder; Lane 2-5: clone A; Lane 6-9: clone B; Lane 10-13: clone C; Lane 14-17: Clone D. Each clone was digested with *Bsr*GI, *Dra*III or both, respectively.

Only the last two samples (C and D) showed the expected restriction pattern: two bands of 3.4 kb and 3.9 kb each. These samples were sent for sequencing, but the result did not match the correct sequence. It is possible that the restriction reactions performed during DNA preparation or the ligation reaction itself led to an incorrect sequence in the plasmid. The construction had to be repeated using a different method.

The In-Fusion HD Cloning Kit allows fast and directional cloning of DNA fragments into any vector by recognizing 15 bp overlaps at the fragments ends and fusing them, using the In-Fusion Enzyme. Since this cloning method requires specific primers so that the enzyme can work properly, new primers were designed according to the supplied guidelines:

- The 5' end of each primer must contain 15 homologue bases to 15 bases in one end of the destination vector
- The 3'end of each primer should be gene specific, be between 18 and 25 bp in length and have a maximum melting temperature of 65 ºC.

SP Fw TACCTCTAGAGTCGAC**GCCACCATG**CAGGACTCAGCCAGTCCCAT

Complementary to pCI-neo — complementary to the gene

KOZAK sequence

SP Rv TAAAGGGAAGCGGCCGCCTAGTGGTGGTGA

Complementary to pCI-neo — complementary to the gene

Figure 4.19: In Fusion primers diagram, showing the regions complementary to the insert and the vector, as well as the KOZAK sequence.

The SP primers were optimized using three different annealing temperatures: 50, 55 and 60 ºC. The expected fragment was present in every sample analyzed in gel electrophoresis. New PCR reactions were prepared using the 60 ºC annealing temperature. The reactions were loaded into an agarose gel and, after electrophoresis, band purification from the gel was performed.

The pCI-neo destination vector was digested using *Sal*I and *Not*I and purified. The In-Fusion reaction was set up according to the protocol using the purified fragment and vector. Stellar competent cells were transformed and plasmid DNA from four different colonies was extracted. Each clone was digested using *Sma*I and *Xho*I restriction enzymes. The *Sma*I enzyme cuts pCI-neo hCES2-10xHis twice in the gene, resulting in two bands of 6 kb and 1.3 kb; however, one of the restriction sites is lost with the deletion of the signal peptide, so only one 7.1 kb band would be expected in a successful construction. Both restriction sites for *Xho*I should be intact in the construction, but the lightest band would be smaller than in the template plasmid. (Figure 4.20)

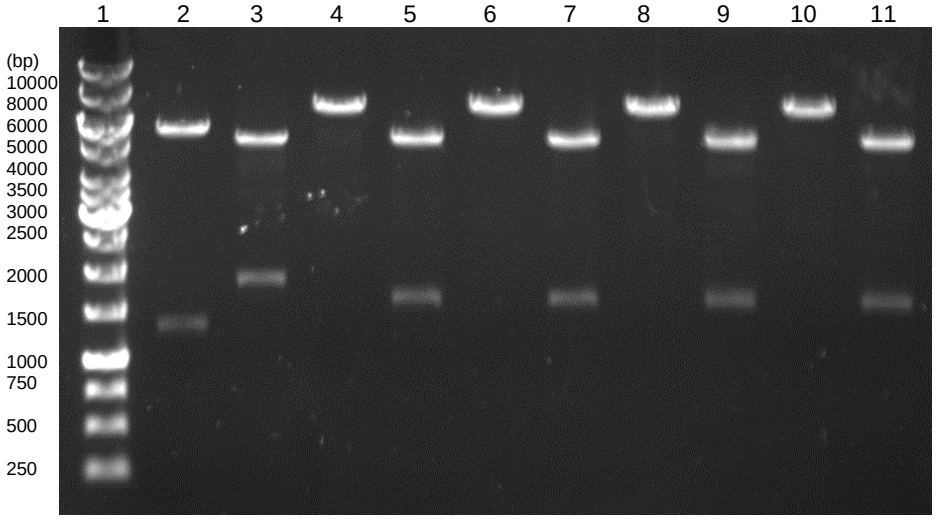


Figure 4.20: Agarose gel analysis of different clones digested with two enzymes. Lane 1: Gene Ruler 1 kb DNA Ladder; Lane 2: pCI-neo hCES2-10xHis + *Sma*I; Lane 3: pCI-neo hCES2-10xHis + *Xho*I; Lane 4: Clone 1 + *Sma*I; Lane 5: Clone 1 + *Xho*I; Lane 6: Clone 2 + *Sma*I; Lane 7: Clone 2 + *Xho*I; Lane 8: Clone 3 + *Sma*I; Lane 9: Clone 3 + *Xho*I; Lane 10: Clone 4 + *Sma*I; Lane 11: Clone 4 + *Xho*I.

Analysis of the agarose gel indicated that every clone had the expected restriction pattern. Since there were only four clones, all were sent for sequencing. According to the sequencing results, clones 1 and 3 were correctly constructed and presented no spontaneous mutations; clone 2 had a single base mutation that wasn't confirmed nor denied by the other sequencing products; and one of the primers didn't sequence the final segment of clone 4.

The new mutated plasmid, pCI-neo hCES2-10xHis w/oSP (map in Appendix B-3), lacked the signal peptide but the KOZAK sequence was successfully inserted, ensuring the correct reading frame for protein expression.

4.5 Mutant protein production and analysis

Once the plasmids had been designed and its quality established, HEK 293T cells were transiently transfected with them. This production of different mutants of hCES2-10xHis had the objective of studying the differences in the proteins, such as their molecular weight and activity.

A total of six erlenmeyers were cultured, and five were transfected with plasmid DNA:

- NC: negative control. Not transfected.
- PC: positive control. Transfected with pCI-neo hCES2-10xHis plasmid.
- w/oSP: transfected with pCI-neo hCES2-10xHis w/oSP. The protein should have no signal peptide, and could remain in the cytoplasm.
- G1: transfected with pCI-neo CES2 10xHis Glyco1. Secreted protein not glycosylated at the first site.
- G2: transfected with pCI-neo CES2 10xHis Glyco2. Secreted protein not glycosylated at the second site.
- G1+2: transfected with pCI-neo hCES2-10xHis Glyco1+2. Secreted protein not glycosylated.

The clones were selected according to the sequencing results and NanoDrop quality assessment. The chosen plasmids had no mutations other than the ones intended, and their concentration and ratios were also taken into consideration.

Sixty hours post-transfection each cell culture was centrifuged and stored. The same volume of each supernatant was used for precipitating total protein. SDS-Page and Western Blot were performed as previously described.

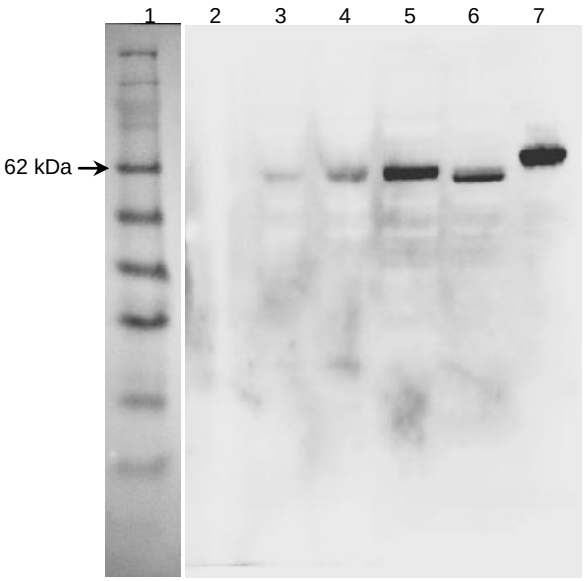


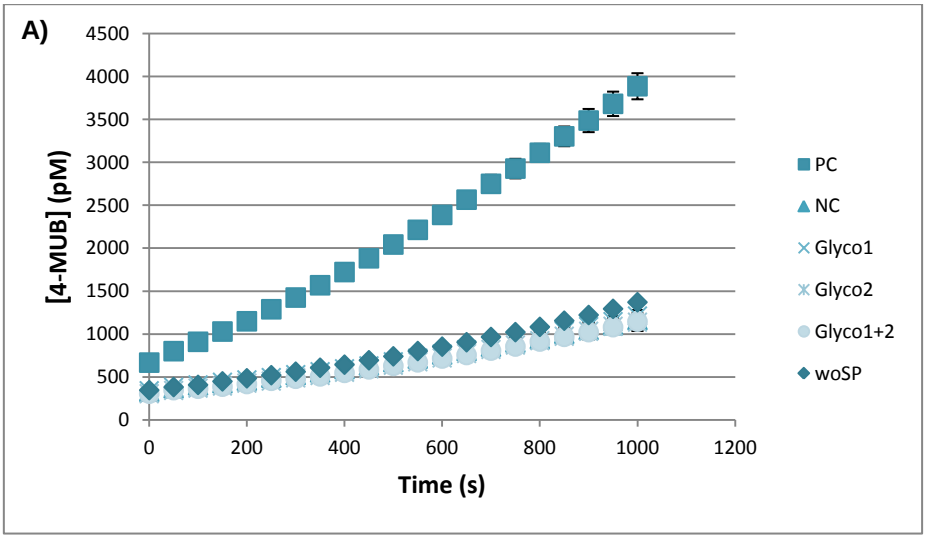
Figure 4.21: Western Blot of supernatant samples from the transfected cell cultures. Lane 1: SeeBlue Plus 2 Pre-stained Standard (Life Technologies, USA); Lane 2: negative control; Lane 3: hCES2-10xHis woSP; Lane 4: hCES2-10xHis Glyco1; Lane 5: hCES2-10xHis Glyco2; Lane 6: hCES2-10xHis Glyco 1+2; Lane 7: hCES2-10xHis.

The hCES2-10xHis band is visible in lane 7 and has the expected molecular weight. Since this is the only fully glycosylated protein, no other lane presented a band at the same weight. Lanes 4 and 5 that correspond to the mutant proteins with only one glycosylation site, showed a slight decrease in the molecular weight of the protein. This decrease is more accentuated in lane 6 as expected, since this protein should not be glycosylated at all. Lane 3 shows a very faint band, which can either mean that less protein was loaded into the gel, or that not all of the protein is being excreted from the cell. This was expected since, without the signal peptide, the protein shouldn't take part of the classical secretory pathway and should be retained in the cytoplasm; however, since transfection causes an increase in protein synthesis, perhaps excess protein is being excreted to the cell medium through the non canonic pathway. This band also has lower molecular weight than the positive control, which could indicate that it hasn't been glycosylated.

Total protein concentration in the supernatants was assessed by Bradford Assay. The glycosylation site mutants (hCES2-10xHis Glyco1, hCES2-10xHis Glyco2 and hCES2-10xHis Glyco1+2) were all in the same range of concentrations of the positive control, approximately 25 to 35 µg/mL. However, the hCES2-10xHis w/oSP mutant had a lower concentration of 16,6 µg/mL. This result could indicate that the protein expressed without the signal peptide isn't secreted as the other mutants, but instead remains in the cytoplasm to be degraded.

Human CES2 is usually described as having two potential N-glycosylation sites. In this experiment, hCES2-10xHis was mutated resulting in plasmids with only one or neither of the glycosylation sites. Figure 4.21 shows a sequential decrease in molecular weight of the mutated proteins: hCES2-10xHis Glyco1 and Glyco2 both have the same weight, although it has decreased from control hCES2-10xHis; and hCES2-10xHis Glyco1+2 has an even lower molecular weight. This gradual decrease accompanies the loss of the N-glycosylation sites, indicating that both sites are effectively glycosylated.

The enzymatic activity assay was performed using 150 µL of each supernatant. The 96 well plate was prepared and read in the fluorimeter as described in the materials and methods section.



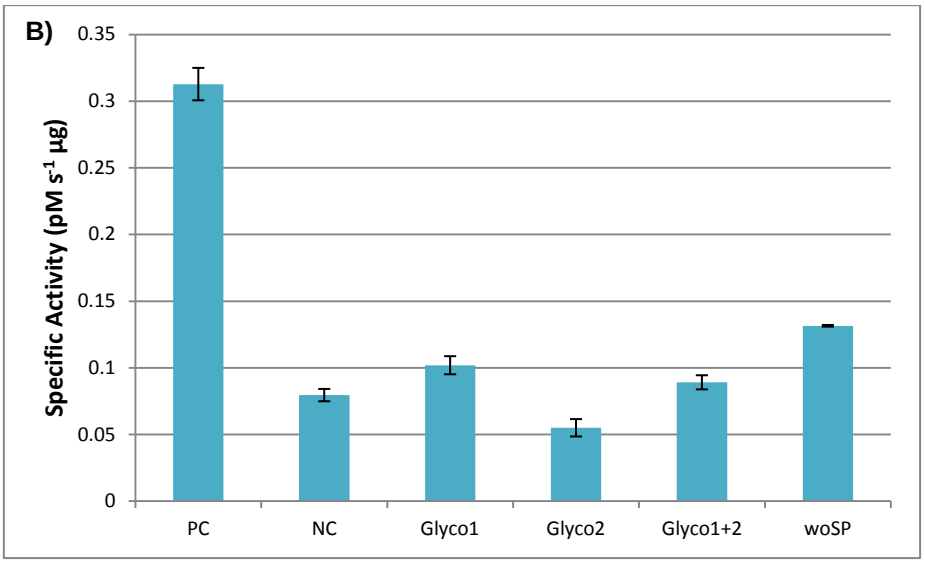


Figure 4.22: Activity evaluation of mutated and control hCES2-10xHis, using 4-MUBA as substrate. A) Product formation curve of each protein; B) Specific activity of each protein, calculated from curves plotted in panel A, from 300 to 700 seconds. 4-MUBA – 4-methylumbelliferyl acetate; 4-MUB – 4-methylumbelliferone.

Every mutant protein was expressed, as confirmed by the Western Blot in Figure 4.21. However, only hCES2-10xHis, the positive control, showed any activity in the presence of 4-MUBA substrate. These results confirm the initial hypothesis that N-glycans are necessary during protein expression – without the sugar residues hCES2-10xHis mutants probably did not achieve the correct conformation that would lead to enzymatic activity. However, there is also the possibility that the exchange of the asparagine residues for glutamine could have impacted protein conformation.

5 Final Remarks

The main objective of this work was to assess N-glycosylation influence in hCES2-10xHis protein activity as a surrogate of hCES2.

Enzymatic deglycosylation allowed studying the protein post-synthesis, after the post translational modifications. It was verified the deglycosylation was successful and that N-glycans are not required for the activity of fully synthesized hCES2-10xHis, as it was expected, although they may be somehow involved in its correct synthesis.

The addition of tunicamycin to a cell culture transfected with pCI-neo hCES2-10xHis resulted in a decrease in protein molecular weight. A decrease in specific activity of the supernatant along the culturing period was visible, indicating that the protein stopped being glycosylated and therefore lost activity. The culture in the absence of tunicamicyn had an increase in the specific activity consistent with an increase in protein concentration. However, only a slight difference in activity was found between samples from treated and untreated cell cultures at 48 hpt, although a more pronounced effect was expected. These cultures were from two independent transfections which, although they were both transfected with the same plasmid, could be responsible for some variability.

The removal of the signal peptide inhibited the processing of the protein through the classical secretory pathway and therefore avoided glycosylation in the ER. Considering that glycosylation is essential for correct conformation of the active site, an inactive form of the protein was expected in this assay. Since a significant difference in specific activity was observed between the control and the protein expressed without the signal peptide, two hypotheses can be considered: either the amount of hCES2-10xHis in the supernatant was too low, or the protein in fact lost activity because it was not matured in the ER and in the Golgi.

The mutated forms of the protein that lacked either one or both glycosylation sites showed the expected shift in molecular weight. This assay also confirmed that both N-glycosylation sites are indeed glycosylated in the non mutated protein, as the molecular weight of the protein decreased in the mutants. Enzymatic activity assays showed that the mutants did not have esterase activity larger than the negative control, as predicted – the inhibition of N-glycosylation lead to the loss of activity. This loss of activity can be explained by two hypotheses: ER and Golgi processing alone, in the absence of glycans, is not sufficient for the correct folding of hCES2-10xHis around the active site or that the mutation of one or two aminoacids provokes a larger effect in the three dimensional structure than anticipated.

These experiments, in general, may indicate that N-glycosylation is needed during hCES2-10xHis synthesis but not after its maturation.

6 Future Work

In the future, some of the assays should be revised and repeated, in order to validate the obtained results. The tunicamycin assay, for instance, should be repeated with larger inhibitor concentration and for a longer time.

If possible, the produced mutants should be purified to normalize the concentration of hCES2-10xHis present, instead of using total protein concentration as normalization factor.

Different methods should be developed in order to obtain a form of the protein that can lead to the crystallographic structure of hCES2.

7 Bibliography

Bahar, F.G., Ohura, K., Ogihara, T. and Imai, T. 2012. Species difference of esterase expression and hydrolase activity in plasma. J Pharm Sci 101(10):3979-88.

Bencharit, S., Morton, C.L., Xue, Y., Potter, P.M. and Redinbo, M.R. 2003. Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme. Nature Structural & Molecular Biology 10, 349.

Butler M. 2005. Animal cell cultures: recent achievements and perspectives in the production of biopharmaceuticals. Appl Microbiol Biotechnol 68(3):283-91.

Butler, M. 2006. Optimisation of the cellular metabolism of glycosylation for recombinant proteins produced by Mammalian cell systems. Cytotechnology 50(1-3):57-76.

Castilho, L., Moraes, A., Augusto, E. and Butler, M. 2008. Animal Cell Technology: from Biopharmaceuticals to Gene Therapy. 1st edition. Taylor & Francis.

Cunha, B. 2011. Expression of the human carboxylesterase 2 enzyme (CES2) in mammalian cells. Master Thesis. Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

Hartung, T., Balls, M., Bardouille, C., Blanck, O., Coecke, S., Gstraunthaler, G., Lewis, D.; ECVAM Good Cell Culture Practice Task Force. 2002. Good Cell Culture Practice. ECVAM Good Cell Culture Practice Task Force Report 1. Altern Lab Anim 30(4):407-14.

Kotlar, D. and Lavner, Y. 2006. The action of selection on codon bias in the human genome is realted to frequency, complexity, and chronology of amino acids. BMC Genomics 7:67.

Kroetz, D.L., McBride, O.W. and Gonzalez, F.J.. 1993. Glycosylation-dependent activity of baculovirus-expressed human liver carboxylesterases: cDNA cloning and characterization of two highly similar enzyme forms. Biochemistry 32(43):11606-17.

Laizure, S.C., Herring, V., Hu, Z., Witbrodt, K. and Parker, R.B. 2013. The role of human carboxylesterases in drug metabolism: have we overlooked their importance? Pharmacotherapy 33(2):210-22.

Lamego, J., Cunha, B., Peixoto, C., Sousa, M.F., Alves, P.M., Simplício, A.L., Coroadinha, A.S. 2013. Carboxylesterase 2 production and characterization in human cells: new insights into enzyme oligomerization and activity. Appl Microbiol Biotechnol 97(3):1161-73.

Lockridge, O. and Quinn, D.M.. 2010. Esterases. *In* Comprehensive Toxicology, 2nd Edition, pp 243-273, Elsevier.

Maley, F, Trimble, R.B., Tarentino, A.L. and Plummer, T.H. Jr. 1989. Characterization of glycoproteins

and their associated oligosaccharides through the use of endoglycosidases. Anal Biochem 180(2):195-204.

Merten, O.W. 2006. Introduction to animal cell culture technology-past, present and future. Cytotechnology 50(1-3):1-7.

Morgan, E.W., Yan, B., Greenway, D., Petersen, D.R., Parkinson, A. 1994. Purification and characterization of two rat liver microsomal carboxylesterases (hydrolase A and B). Arch Biochem Biophys 315(2):495-512.

Morton, C.L. and Potter, P.M. 2000. Comparison of *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris, Spodoptera frugiperda*, and COS7 cells for recombinant gene expression. Application to a rabbit liver carboxylesterase. Mol Biotechnol 16(3):193-202.

Oda, Y., Okada, T., Yoshida, H., Kaufman, R.J., Nagata, K. and Mori, K. 2006. Derln-2 and Derln-3 are regulated by the mammalian unfolded protein response and are required for ER associated degradation. JCB 172(3):383-393.

Pindel, E.V., Kedishvili, N.Y., Abraham, T.L., Brzezinski, M.R., Zhang, J., Dean, R.A. and Bosron, W.F. 1997. Purification and cloning of a broad substrate specificity human liver carboxylesterase that catalyzes the hydrolysis of cocaine and heroin. J Biol Chem 272(23):14769-75.

Powell, L.D. 2001. Inhibition of N-linked Glycosylation. Current Protocols in Immunology 8.14.1-8.14.9.

Rapoport, T.A. 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature 450(7170):663-669.

Ross, M.K., Borazjani, A., Wang, R., Crow, J.A. and Xie, S. 2012. Examination of the carboxylesterase phenotype in human liver. Arch Biochem Biophys 522(1):44-56.

Sambrook, J. and Russel, D.W. 2001. In vitro amplification of DNA by PCR. *In* Molecular Cloning: A Laboratory Manual, 3rd edition. Cold Spring Harbor, New York.

Satoh, T. and Hosokawa, M. 1998. The mammalian carboxylesterases: from molecules to functions. Annu Rev Pharmacol Toxicol 38:257-88.

Satoh, T. and Hosokawa, M. 2006. Structure, function and regulation of carboxylesterases. Chem Biol Interact 162(3):195-211.

Satoh, T. and Hosokawa, M. 2010. Carboxylesterases: structure, function and polymorphism in mammals. Journal of Pesticide Science 35 (3): 218-228.

Schiel, M.A., Green, S.L., Davis, W.I., Sanghani, P.C., Bosron, W.F. and Sanghani, S.P. 2007. Expression and characterization of a human carboxylesterase 2 splice variant. J Pharmacol Exp Ther 323(1):94-101.

Schiel, M.A. 2009. Human Carboxylesterase 2 splice variants: expression, activity, and role in the metabolism of irinotecan and capecitabine. Doctoral Thesis. Department of Biochemistry and Molecular Biology, University of Indiana.

Schwer, H., Langmann, T., Daig, R., Becker, A., Aslanidis, C. and Schmitz, G. 1997. Molecular cloning and characterization of a novel putative carboxylesterase, present in human intestine and liver. Biochem Biophys Res Commun 233(1):117-20.

Taylor, W.R. 1986. Classification of amino acid conservation. J Theor Biol 119:205-218.

Thomas, P. and Smart, T.G. 2005. HEK 293 cell line: a vehicle for the expression of recombinant proteins. J Pharmacol Toxicol Methods 51(3):187-200.

Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Wart, G.W. and Etzler, M.E. 2009. Essentials of Glycobiology, 2nd edition. Cold Spring Harbor Laboratory Press, New York.

Walker, J. and Rapley, R. 2009. Molecular Biology and Biotechnology. 5th edition. Royal Society of Chemistry, Cambridge.

Wong, D.W.S. 2006. The ABCs of Gene Cloning. Springer US.

Xie, M., Yang, D., Liu, L., Xue, B. and Yan, B. 2002. Human and rodent carboxylesterases: immunorelatedness, overlapping substrate specificity, differential sensitivity to serine enzyme inhibitors, and tumor-related expression. Drug Metab Dispos. 30(5):541-7.

Zschenker, O., Bähr, C., Hess, U.F., Ameis, D. 2005. Systematic mutagenesis of potentialglycosylation sites of lysosomal acid lipase. J Biochem 137(3):387-94.

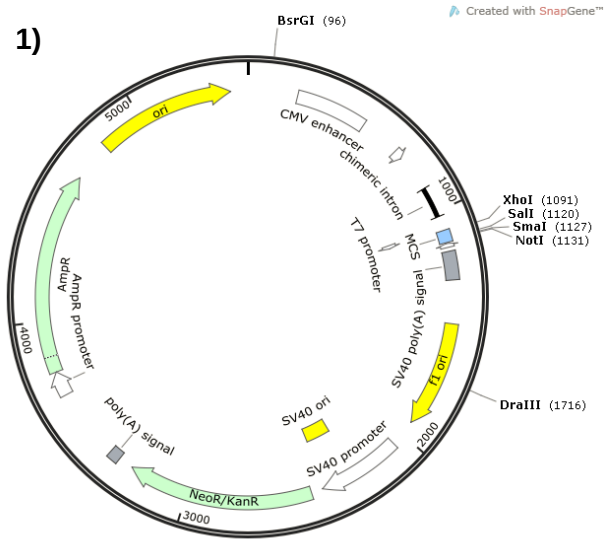
Appendixes

Appendix A – hCES2-10xHis codifying sequence and respective amino acid residues (Ensembl). Initiation codon is indicated by a box, and the signal peptide and N-glycosylation site residues are in bold.

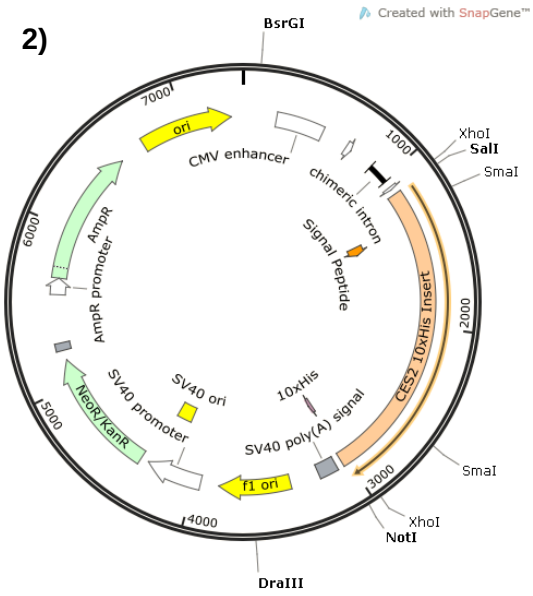
```
5'…atg act gct cag tcc cgc tct cct acc aca ccc acc ttt ccc ggc cca agc cag cgc acc ccg
   M   T   A   Q   S   R   S   P   T   T   P   T   F   P   G   P   S   Q   R   T   P
   1                                                                               21

  ctg act ccc tgc cca gtc caa act cca agg ctg ggc aag gca ctg atc cac tgc tgg aca gac
  L   T   P   C   P   V   Q   T   P   R   L   G   K   A   L   I   H   C   W   T   D
  22                                                                              42

  ccg ggg cag cct ctg ggt gaa cag cag cgt gtc cgc cgg cag cga acc gag acc agc gag ccg
  P   G   Q   P   L   G   E   Q   Q   R   V   R   R   Q   R   T   E   T   S   E   P
  43                                                                              63

  acc atg cgg ctg cac aga ctt cgt gcg cgg ctg agc gcg gtg gcc tgt ggg ctt ctg ctg ctt
  T   M   R   L   H   R   L   R   A   R   L   S   A   V   A   C   G   L   L   L   L
  64                                                                              84

  ctt gtc cgg ggc cag ggc cag gac tca gcc agt ccc atc cgg acc aca cac acg ggg cag gtg
  L   V   R   G   Q   G   Q   D   S   A   S   P   I   R   T   T   H   T   G   Q   V
  85                                                                             105

  ctg ggg agt ctt gtc cat gtg aag ggc gcc aat gcc ggg gtc caa acc ttc ctg gga att cca
  L   G   S   L   V   H   V   K   G   A   N   A   G   V   Q   T   F   L   G   I   P
  106                                                                            126

  ttt gcc aag cca cct cta ggt ccg ctg cga ttt gca ccc cct gag ccc cct gaa tct tgg agt
  F   A   K   P   P   L   G   P   L   R   F   A   P   P   E   P   P   E   S   W   S
  127                                                                            147

  ggt gtg agg gat gga acc acc cat ccg gcc atg tgt cta cag gac ctc acc gca gtg gag tca
  G   V   R   D   G   T   T   H   P   A   M   C   L   Q   D   L   T   A   V   E   S
  148                                                                            168

  gag ttt ctt agc cag ttc aac atg acc ttc cct tcc gac tcc atg tct gag gac tgc ctg tac
  E   F   L   S   Q   F   N   M   T   F   P   S   D   S   M   S   E   D   C   L   Y
  169                 175                                                        189

  ctc agc atc tac acg ccg gcc cat agc cat gaa ggc tct aac ctg ccg gtg atg gtg tgg atc
  L   S   I   Y   T   P   A   H   S   H   E   G   S   N   L   P   V   M   V   W   I
  190                                                                            210

  cac ggt ggt gcg ctt gtt ttt ggc atg gct tcc ttg tat gat ggt tcc atg ctg gct gcc ttg
  H   G   G   A   L   V   F   G   M   A   S   L   Y   D   G   S   M   L   A   A   L
  211                                                                            231

  gag aac gtg gtg gtg gtc atc atc cag tac cgc ctg ggt gtc ctg ggc ttc ttc agc act gga
  E   N   V   V   V   V   I   I   Q   Y   R   L   G   V   L   G   F   F   S   T   G
  232                                                                            252

  gac aag cac gca acc ggc aac tgg ggc tac ctg gac caa gtg gct gca cta cgc tgg gtc cag
  D   K   H   A   T   G   N   W   G   Y   L   D   Q   V   A   A   L   R   W   V   Q
  253                                                                            273

  cag aat atc gcc cac ttt gga ggc aac cct gac cgt gtc acc att ttt ggc gag tct gcg ggt
  Q   N   I   A   H   F   G   G   N   P   D   R   V   T   I   F   G   E   S   A   G
  274                                                                            294

  ggc acg agt gtg tct tcg ctt gtt gtg tcc ccc ata tcc caa gga ctc ttc cac gga gcc atc
  G   T   S   V   S   S   L   V   V   S   P   I   S   Q   G   L   F   H   G   A   I
  295                                                                            315
```

```
atg gag agt ggc gtg gcc ctc ctg ccc ggc ctc att gcc agc tca gct gat gtc atc tcc acg
M   E   S   G   V   A   L   L   P   G   L   I   A   S   S   A   D   V   I   S   T
316                                                                             336


gtg gtg gcc aac ctg tct gcc tgt gac caa gtt gac tct gag gcc ctg gtg ggc tgc ctg cgg
V   V   A   N   L   S   A   C   D   Q   V   D   S   E   A   L   V   G   C   L   R
337         340                                                                 357

ggc aag agt aaa gag gag att ctt gca att aac aag cct ttc aag atg atc ccc gga gtg gtg
G   K   S   K   E   E   I   L   A   I   N   K   P   F   K   M   I   P   G   V   V
358                                                                             378

gat ggg gtc ttc ctg ccc agg cac ccc cag gag ctg ctg gcc tct gcc gac ttt aag cct gtc
D   G   V   F   L   P   R   H   P   Q   E   L   L   A   S   A   D   F   K   P   V
379                                                                             399

cct agc att gtt ggt gtc aac aac aat gaa ttc ggc tgg ctc atc ccc aag gtc atg agg atc
P   S   I   V   G   V   N   N   N   E   F   G   W   L   I   P   K   V   M   R   I
400                                                                             420

tat gat acc cag aag gaa atg gac aga gag gcc tcc cag gct gct ctg cag aaa atg tta acg
Y   D   T   Q   K   E   M   D   R   E   A   S   Q   A   A   L   Q   K   M   L   T
421                                                                             441

ctg ctg atg ttg cct cct aca ttt ggt gac ctg ctg agg gag gag tac att ggg gac aat ggg
L   L   M   L   P   P   T   F   G   D   L   L   R   E   E   Y   I   G   D   N   G
442                                                                             462

gat ccc cag acc ctc caa gcg cag ttc cag gag atg atg gcg gac tcc atg ttt gtg atc cct
D   P   Q   T   L   Q   A   Q   F   Q   E   M   M   A   D   S   M   F   V   I   P
463                                                                             483

gca ctc caa gta gca cat ttt cag tgt tcc cgg gcc cct gtg tac ttc tac gag ttc cag cat
A   L   Q   V   A   H   F   Q   C   S   R   A   P   V   Y   F   Y   E   F   Q   H
484                                                                             504

cag ccc agc tgg ctc aag aac atc agg cca ccg cac atg aag gca gac cat ggt gat gag ctt
Q   P   S   W   L   K   N   I   R   P   P   H   M   K   A   D   H   G   D   E   L
505                                                                             525

cct ttt gtt ttc aga agt ttc ttt ggg ggc aac tac att aaa ttc act gag gaa gag gag cag
P   F   V   F   R   S   F   F   G   G   N   Y   I   K   F   T   E   E   E   E   Q
526                                                                             546

cta agc agg aag atg atg aag tac tgg gcc aac ttt gcg aga aat ggg aac ccc aat ggc gag
L   S   R   K   M   M   K   Y   W   A   N   F   A   R   N   G   N   P   N   G   E
547                                                                             567

ggt ctg cca cac tgg ccg ctg ttc gac cag gag gag caa tac ctg cag ctg aac cta cag cct
G   L   P   H   W   P   L   F   D   Q   E   E   Q   Y   L   Q   L   N   L   Q   P
568                                                                             588

gcg gtg ggc cgg gct ctg aag gcc cac agg ctc cag ttc tgg aag aag gcg ctg ccc caa aag
A   V   G   R   A   L   K   A   H   R   L   Q   F   W   K   K   A   L   P   Q   K
589                                                                             609

atc cag gag ctc gag gag cct gaa gag aga cac aca gag ctg agc gct cat cac cac cat cat
I   Q   E   L   E   E   P   E   E   R   H   T   E   L   S   A   H   H   H   H   H
610                                                                             630

cac cat cac cac cac tag …3'
H   H   H   H   H   *
631             636
```

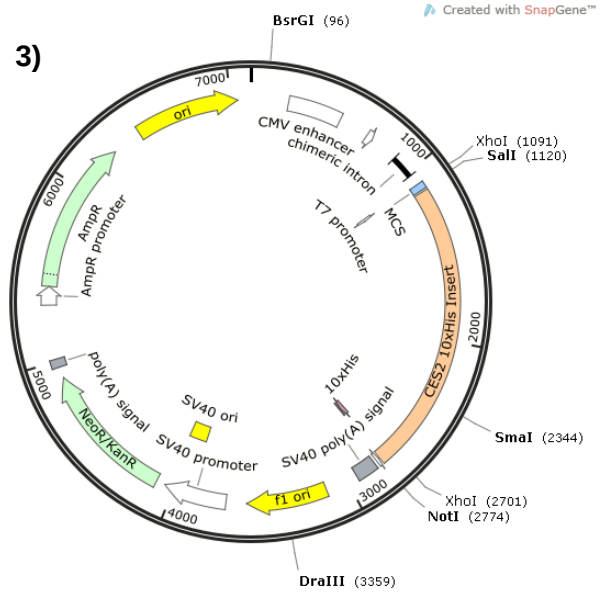
Appendix B – Maps pCI-neo (panel 1), pCI-neo hCES2-10xHis (panel 2) and pCI-neo hCES2-10xHis w/oSP (panel 3). Some structural elements are represented, as well as restriction sites for the enzymes used in the course of the work.



pCI-neo (5472 bp)

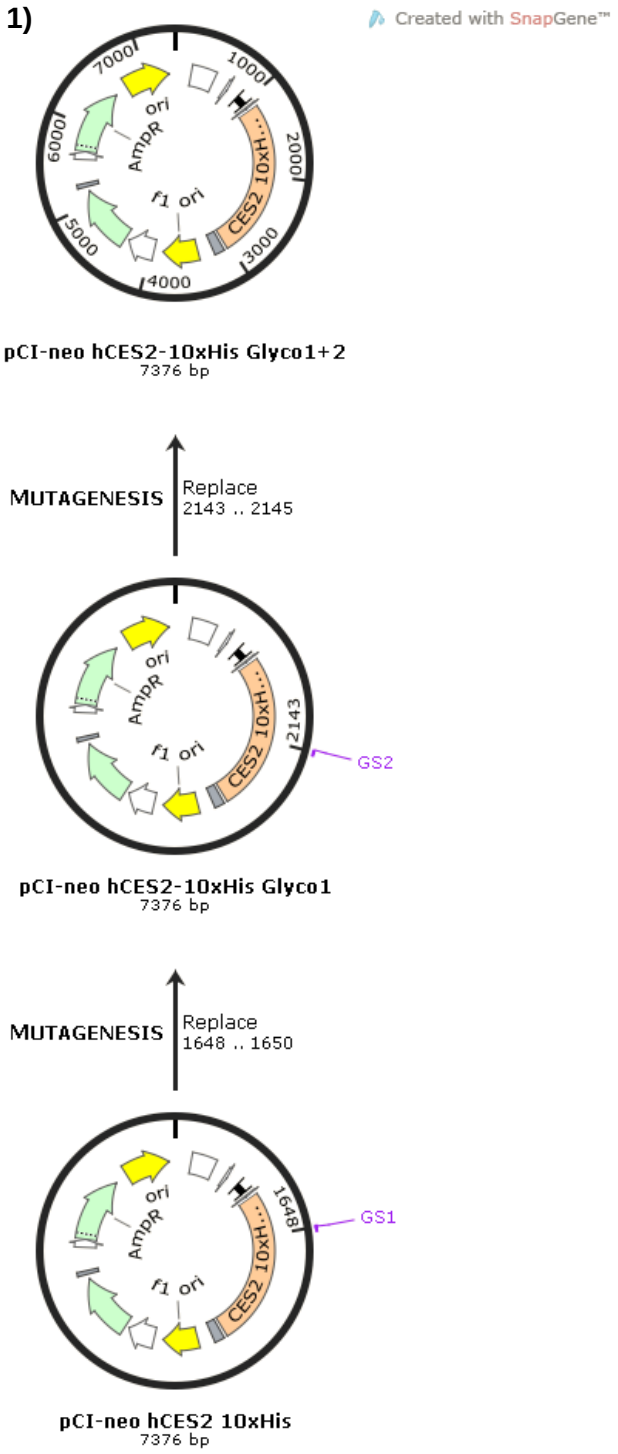


pCI-neo hCES2-10xHis (7376 bp)

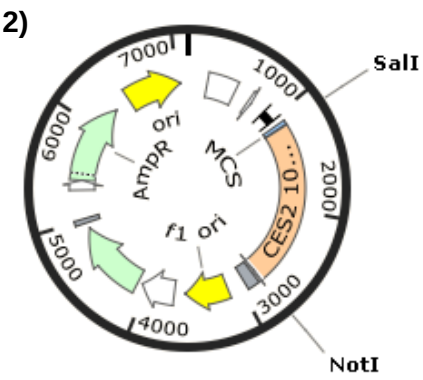


pCI-neo hCES2-10xHis w/oSP (7115 bp)

Appendix C – Diagrams of plasmid mutation. Panel 1: Site directed mutagenesis of the two glycosylation sites; Panel 2: removal of the signal peptide using In-Fusion Cloning.



2)



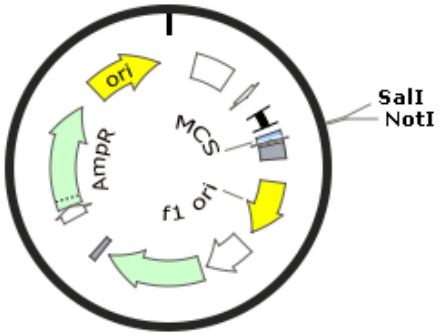
pCI-neo hCES2-10xHis woSP
7115 bp

Created with SnapGene™

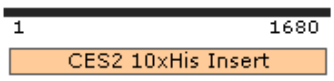
IN-FUSION CLONING

Replace
SalI (1120) – NotI (1131)

Overlap and insert
1 .. 1680

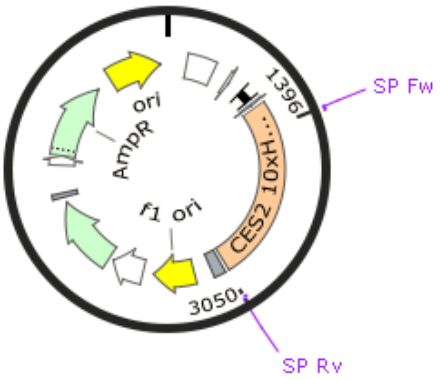


pCI-neo
5472 bp



Fragment
1680 bp

PCR Amplify
1396 .. 3050



pCI-neo hCES2 10xHis
7376 bp