IPS Instituto Politécnico de Setúbal
Escola Superior de Tecnologia do Barreiro

JOÃO MIGUEL PINTO DOS PRAZERES

EXPRESSION OF RECOMBINANT HUMAN PROTEINS INVOLVED IN DRUG METABOLISM

Relatório de Dissertação do Mestrado em Engenharia Biológica e Química

ORIENTADOR
Professora Doutora Marta Campos Justino

SUPERVISOR
Doutor Gonçalo Campos Justino

Junho de 2021

JOÃO MIGUEL PINTO DOS PRAZERES

EXPRESSION OF RECOMBINANT HUMAN PROTEINS INVOLVED IN DRUG METABOLISM

JÚRI

Presidente: Professora Doutora Maria de Lurdes de Figueiredo Gameiro, ESTBarreiro/IPS

Supervisor: Doutor José Gonçalo Deira Duarte de Campos Justino, Centro de Química Estrutural, Instituto Superior Técnico, Universidade de Lisboa

Vogal: Doutora Cláudia Sofia Pires Godinho, BioData Consortium, Instituto Superior Técnico, Universidade de Lisboa

Junho de 2021

I've experiments to run; there is research to be done

On the people who are still alive”

Still Alive - Jonathan Coulton

ACKNOWLEDGEMENTS

I would like to take this opportunity to, first of all, thank both *Escola Superior de Tecnologia do Barreiro*, ESTB/IPS, and *Centro Estrutural de Quimica*, CQE/IST, for providing all of their facilities and human resources needed for this internship to go as smoothly as it did, even in these troubling quarantine times.

I want to thank Professor Matilde Marques (CQE/IST) for providing the resources necessary for this internship and both my supervisors, Professor Marta Justino (ESTB/IPS) and Doctor Gonçalo Justino (CQE), for all the support, both in lab work and in the writing of this dissertation, availability to enlighten any doubts that appeared along the way, and all-around guidance throughout the time I was able to work alongside them. I also want to thank Professor Pedro Pinheiro, Cátia Marques, and Pedro Rosado, for all of the lab experience they shared with me during my time at CQE, as well as the companionship. A final thank you has also to be granted to D. Rosa, the lab technician from the BERG group, for all her sympathy and willingness to help me whenever she could.

I can not thank my close friends and family enough for all the moral support they offered, especially during the first months of the internship when the future of all ours lives looked a bit uncertain at times.

It is also important to acknowledge that this work was funded by *Fundaçao para a Ciência e a Tecnologia* through projects *UIDB/00100/2020, SAICTPAC/0019/2015 and PTDC/QUI-QAN/32242/2017.*

RESUMO

Ao longo dos anos, a expressão heteróloga de proteínas recombinantes tem sido um campo muito estudado na biotecnologia, com diversas aplicações laboratoriais e industriais, desde a produção de proteínas para estudos de estrutura e função até à produção industrial, em larga escala, de proteínas como a insulina, que são utilizadas como agentes terapêuticos. A expressão heteróloga de proteínas está intrinsecamente ligada à tecnologia de DNA recombinante, que, ao permitir a manipulação dos genes envolvidos nos processos de expressão proteica, permite a obtenção de praticamente qualquer proteína desejada em quantidades apreciáveis.

O objectivo deste trabalho foi o estabelecimento de estirpes bacterianas capazes de expressar proteínas humanas envolvidas no metabolismo de fármacos, de modo a possibilitar o seu posterior uso em modelos *in vitro* de metabolismo de fármacos. Recorrendo à sub-clonagem de cDNA, procurou-se clonar duas isoformas do citocromo P450 (CYP2C8 e CYP3A4), a NADPH:citocromo P450 oxidorreductase (POR), e a UDP-glucuronosiltransferase 2B7 (UGT2B7).Enquanto que os CYP e a POR estão envolvidas no metabolismo de fase I de muitos fármacos, bem como outros xenobióticos e compostos endógenos, catalisando a introdução de grupos polares de modo a facilitar a sua posterior metabolização, a UGT2B7 actua a nível da fase II, catalisando a inserção de grupos glucuronosilo nos fármacos, facilitando a sua excreção.

Para realizar estas subclonagens foram aplicadas técnicas de clonagem tradicional usando amplificação de genes por PCR e sucessivas digestões enzimáticas, fazendo a ligação dos *inserts* em vetores de expressão com o auxílio de ligases. Em algumas das proteínas (CYP2C8, CYP3A4, e POR) foi aplicada também uma técnica de clonagem independente de ligação enzimática, recorrendo ao uso de *primers* e reações de PCR para amplificar os *inserts* de interesse e ligar estes a vetores específicos para este tipo de procedimento.

Foram obtidos clones prontos a expressar em *E. coli* C43(DE3) das proteínas POR e UGT2B7 e clones da CYP2C8 e CYP3A4 em *E.coli* DH5α. As proteínas POR e UGT2B7 foram expressas concluindo que a proteína UGT2B7 pode ser obtida purificando o sobrenadante da lise celular das células expressas com uma coluna de permuta iónica eluída com um buffer com 25 mM de imidazole. Não foi possível obter a proteína POR através do mesmo método de purificação visto que, apesar de ser expressa, esta fica presa nas membranas da célula.

PALAVRAS-CHAVE: Citocromo P450, glucuronosiltransferase, citocromo P450 oxidorredutase, subclonagem, expressão proteica.

ABSTRACT

Over the years, heterologous expression of recombinant proteins has been a very important field in biotechnology, with various laboratory and industrial applications, from protein expression for structural and functional studies to the large-scale industrial production of proteins, such as insulin, that are used as therapeutics. Protein heterologous expression is intrinsically connected to recombinant DNA technology, which, by allowing the manipulation of the genes involved in protein expression, allows obtaining virtually any protein in appreciable amounts.

This work aimed to establish bacterial strains able to express human proteins involved in drug metabolism to allow their use in subsequent *in vitro* studies of drug transformation. By resorting to cDNA sub-cloning, the specific goals were to clone two cytochrome P450 isoforms (CYP2C8 and CYP3A4), NADPH:cytochrome P450 oxidoreductase (POR), and the 2B7 isoform of the UDP-glucuronosyltransferase enzyme (UGT2B7). While CYP and POR are involved in the phase I metabolism of drugs, xenobiotics and endogenous compounds, catalysing the introduction of polar groups to facilitate further metabolism, UGT2B7 is responsible for introducing the large polar glucuronosyl moiety, facilitating the excretion of drugs.

To perform these subcloning tasks, traditional cloning techniques were applied using gene amplification by PCR, coupled to enzymatic digestions, ligating the inserts in the vectors with the aid of ligases. For some of the proteins (CYP2C8, CYP3A4, and POR), a ligation-independent technique was also applied, using primers and PCR reactions to amplify the inserts of interest and ligate these to specific vectors for this type of procedure.

Clones ready to express in *E. coli* C43 (DE3) of POR and UGT2B7 proteins were obtained, and clones of the CYP2C8 and CYP3A4 were transformed in *E. coli* DH5α. The POR and UGT2B7 proteins were expressed concluding that the UGT2B7 protein can be attained by purifying the cell lysis supernatant from the expressed cells with an ion-exchange column eluted with a 25 mM imidazole buffer. It was not possible to obtain the POR protein through the same purification method as this protein,although being expressed, remains attached to the cell membranes.

KEYWORDS: Cytochrome P450, glucuronosyltransferase, cytochrome P450 oxidoreductase, subcloning, protein expression.

GENERAL INDEX

FIGURE INDEX

TABLE INDEX

SYMBOLS AND ABBREVIATIONS

5-ALA	aminolevulinic acid
CAB	carbenicillin
CYP	cytochrome P450 (see Enzyme Glossary)
dNTP	deoxynucleotide triphosphate
e^-	electron
E.coli	*Escherichia coli*
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
GFP	green fluorescent protein
GOI	gene of interest
IPTG	isopropyl-β-*D*-thiogalactopyranoside
LB	Lysogeny Broth
LBAMP50	LB medium supplemented with 50 µg/µL of ampicillin
LBCAB50	LB medium supplemented with 50 µg/µL of carbenicillin
LBKAN30	LB medium supplemented with 30 µg/µL of kanamycin
LIC	ligation-independent cloning
MB H_2O	molecular biology-grade water
MCS	multiple cloning site
NADP	nicotinamide adenine dinucleotide phosphate
NADPH	dihydronicotinamide-adenine dinucleotide phosphate
$O.D._{600\ nm}$	optical density at 600 nm

PCR	polymerase chain reaction
P_{Fwd}	forward primer
POR	NADPH:cytochrome P450 oxidoreductase (see Enzyme Glossary)
P_{Rev}	reverse primer
RE	restriction enzyme
rhP	recombinant human proteins
RT	room temperature
SDS-PAGE	sodium dodecyl sulphate–polyacrylamide gel electrophoresis
SOC	super optimal broth with catabolite repression
SOCKAN30	SOC supplemented with kanamycin [30 µg/mL]
T	temperature
TAE	tris-acetate-EDTA
T_{Ann}	annealing temperature
T_m	melting temperature
UGT2B7	UGT glucuronosyltransferase 2B7 (see Enzyme Glossary)

ENZYME GLOSSARY

All information from this glossary was obtained using BRENDA – Enzyme Database as a source (Chang, Jeske et al. 2021).

Cytochrome P450 (CYP), E.C. 1.14.14.1

Trivial name	Cytochrome P450
Recommended name	Unspecific monoxygenase
Systematic name	substrate,NADPH-hemoprotein reductase:oxygen oxidoreductase (RH-hydroxylating or -epoxidizing)
CYP2C8	cytochrome P450 family 2 subfamily C member 8
CYP3A4	cytochrome P450 family 3 subfamily A member 4

Catalysed reaction:

RH + [reduced NADPH-hemoprotein reductase] + O_2 ⇌ ROH + [oxidized NADPH-hemoprotein reductase] + H_2O

CYPs are a group of P450 heme-thiolate proteins, acting on a wide range of substrates including many xenobiotics, steroids, fatty acids, vitamins, and prostaglandins; reactions catalysed include hydroxylation, epoxidation, *N*-oxidation, sulfoxidation, *N*-, *S*- and *O*-dealkylations, desulfation, deamination, and reduction of azo, nitro and *N*-oxide groups. Together with E.C. 1.6.2.4, NADPH: hemoprotein reductase, it forms a system in which two reducing equivalents are supplied by NADPH.E.C.

Cytochrome P450 oxidoreductase (POR), E.C. 1.6.2.4

Trivial name	NADPH-cytochrome P450 oxidoreductase
Recommended name	NADPH-hemoprotein reductase
Systematic name	NADPH:hemoprotein oxidoreductase

Reaction:

NADPH + H^+ + *n* oxidized hemoprotein ⇌ $NADP^+$ + *n* reduced hemoprotein

POR is a flavoprotein containing both FMN and FAD. This enzyme catalyses the transfer of electrons from NADPH, an obligatory two-electron donor, to microsomal P450 monooxygenases (e.g. E.C. 1.14.14.1, unspecific monooxygenase) by stabilizing the one-electron reduced form of the flavin cofactors FAD and FMN.

UDP-glucuronosyltransferase (UGT), E.C. 2.4.1.17

Trivial name	UDP glucuronosyltransferase
Recommended name	glucuronosyltransferase
Systematic name	UDP-glucuronate β-*D*-glucuronosyltransferase (acceptor-unspecific)

Reaction:

UDP-α-*D*-glucuronate + acceptor ⇌ UDP + acceptor β-*D*-glucuronoside

UGT enzymes accept a wide range of substrates, including phenols, alcohols, amines and fatty acids, catalysing the insertion of a glucuronosyl group in a hydroxyl group E.C. 2.4.1.42, E.C. 2.4.1.59, E.C. 2.4.1.61, E.C. 2.4.1.76, E.C. 2.4.1.77, E.C. 2.4.1.84, E.C. 2.4.1.107 and E.C. 2.4.1.108. A temporary nomenclature for the various forms, whose delineation is in a state of flux.

Endonucleases, Restriction enzymes, type II, E.C. 3.1.21.4

Trivial name	resctriction enzymes
Recommended name	type II site-specific deoxyribonuclease

Endonucleases catalyse the endonucleolytic cleavage of DNA to give specific double-stranded fragments with terminal 5'-phosphates. Endonucleases used in this work are listed below.

Endonuclease	Recognition sequence
*Nde*I (from *Neisseria denitrificans*)	$5'CA^{\downarrow}TATG$ $3'GTAT_{\uparrow}AC$
*Nhe*I (from *Neisseria mucosa heidelbergensis*)	$5'G^{\downarrow}CTAGC$ $3'CGATC_{\uparrow}G$
*Hind*III (from Haemophilus influenzae)	$5'A^{\downarrow}AGCTT$ $3'TTCGA_{\uparrow}A$
*Xho*I (from *Xanthomonas holcicola*)	$5'C^{\downarrow}TCGAG$ $3'GAGCT_{\uparrow}C$

1. INTRODUCTION

1.1. GENETICAL ENGINEERING AND SUBCLONING

Around half a century ago, the term “Genetic Engineering” was introduced to the scientific community to describe the most recent technology, recombinant DNA. Since its beginnings, this technology has developed vastly in a relatively short time frame evolving from cloning simple DNA sequences to be grown in simple life forms, like bacteria, to cloning entire genomes and introducing them into cells using a myriad of ways possible (Bodine, Nicholl 2008, Brown 2010). Under this umbrella term lies one of the central foci of this dissertation, the molecular biology technique of cloning.

While molecular cloning addresses recombinant DNA assembling and replication in host organisms (Watson, Richard M. Myers et al. 2007), subcloning consists of either extracting or creating copies of a target gene from an already constructed plasmid and inserting the gene of interest (GOI) into a new vector. This technique is often used to more easily obtain a protein of interest that is native to a complex organism, *e.g.* the human body, and produce it on a more laboratory accessible scale for further research. Although similar to the cloning technique, these two differ in that whilst in cloning the targeted gene is directly obtained from the genome of the native organism, subcloning works on moving the GOI from one plasmid to another that is more suitable for its intended goal, either gene storage or gene expression.

When performing subcloning, there are a variety of known and patented ways to obtain our final product, some of them being PCR based cloning, traditional digestion/ligation cloning, Gibson Assembly, and Golden Gate Assembly.

1.1.1. SUBCLONING

Traditional cloning is the original subcloning technique and consists of using restriction enzymes (RE), endonucleases, to excise the gene of interest from its plasmid and ligate it to a prepared vector of choice using a DNA ligase, such as T4 DNA Ligase (E.C. 6.5.1.1).

Due to the destructive nature of this process, it is only feasible to follow this route if the plasmid where the GOI stems is of no use for other applications since it is removed from the vector

backbone. When choosing which RE to use, it is important to analyse the plasmid maps of both the donor plasmid and accepting plasmid (vector backbone) carefully and see which restriction sites flank the GOI, being aware of the presence of these sites inside the gene itself, as that would lead to a destruction of the DNA sequence, and also to verify if the chosen REs cut the vector backbone in the desired area and will not interrupt any of the genes necessary for its replication or for the expression of the GOI (normally the restrictions sites fall in a region called multiple cloning site (MCS) of the vector). Using two different REs offers the advantage of always having an insert with non-compatible ends that will force unidirectional cloning when ligated with the backbone, eliminating the chance of having the insert in the wrong direction, and will also reduce the chance of having a recircularization reaction where the vector, instead of ligating with the insert, ligates back with itself producing transformants without the GOI in it (Nicholl 2008, Brown 2010).

Upon proper digestion of both the plasmid and insert, a purification step ensues obtaining these components as pure as possible to ensure a proper ligation and to avoid self-ligation reactions where the RE digestion is inverted. Most common purification methods include, but are not limited to, agarose gel electrophoresis and DNA adsorption via column membrane.

To complete this process, the insert and the vector are ligated in an enzymatic ligation performed by the aforementioned DNA ligase, resulting in a new plasmid ready to be transformed into its intended host or to be stored until further use. These hosts can vary from prokaryotic microorganisms, like bacteria, to more complex eukaryotic hosts, like animal cells.

This whole process can be visually summed up as described in Figure 1.

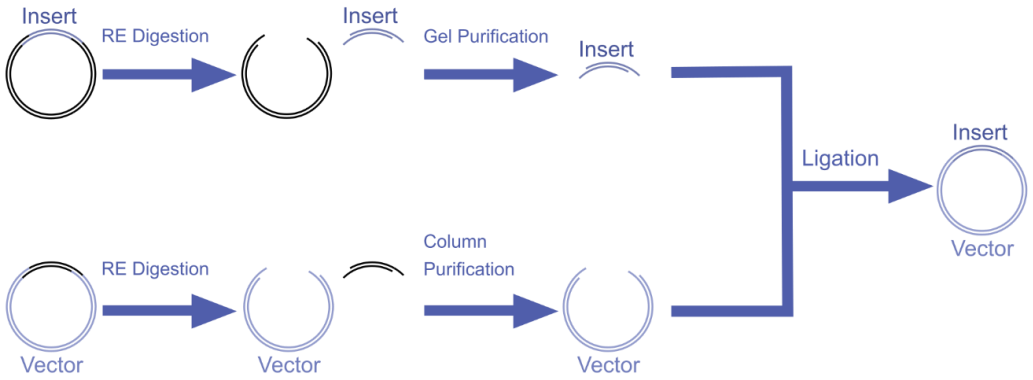


Figure 1: Traditional cloning workflow. In parallel, the insert and vector suffer the same RE digestion and go through a purification step after which an enzymatic ligation ensues to obtain the final plasmid.

1.1.1.1. PCR cloning

This cloning strategy enables the use of polymerase chain reaction (PCR) technology to create a large number of copies of the insert and/or the vector, avoiding the use of REs and the destruction of the template plasmids. Although very similar to the previously described technique (traditional cloning), PCR based cloning presents some unique advantages and disadvantages.

As the name implies, PCR utilizes the properties of the polymerase enzymes, *e.g. Taq* DNA Polymerase, *Pfu* DNA Polymerase (2'-deoxyribonucleoside-5'-triphosphate:DNA deoxynucleotidyltransferase (DNA-directed); E.C. 2.7.7.7 (McDonald 2019)) to construct an exponential number of copies of the targeted sequence. To delimit the region of the plasmid intended to be copied, a pair of oligonucleotide primers is used. These primers also give a starting point for the polymerase since these enzymes aren't capable of making *de novo* sequences. These primers are meant to anneal to each of the DNA strands ("Forward primer" (P_{Fwd}) is the oligonucleotide that anneals to the start codon of the protein to be expressed and "Reverse primer" (P_{Rev}) the one that anneals to the stop codon on the complementary strand) flanking the zone of interest.

Since the primers need to anneal to the sequence, rarely can the same pair of primers be used in two different sequences, which can be a disadvantage. Optimal primers are usually very short oligonucleotides (between 18-30 nt), have a G/C content of 40-60% (percentage of guanine and cytosine nucleotides present in the total sequence), and have a melting temperature (T_m) between 55 ºC and 75 ºC (BRCF, Brown 2010). Nowadays there are multiple software and online calculators that aid in primer syntheses, such as "PrimerSelect" (DNASTAR Version 7.0.0, Madison, WI, USA) and Benchling, making the process much easier and efficient.

One advantage that can be used while building the primer pair is the ability to introduce new restriction sites in the sequence to ensure the final product can be digested to yield sticky-ends, guaranteeing a unidirectional ligation. This implementation is made in primers that anneal to plasmids that either lack any restriction site relatively close to the interest sequence or have a restriction site close to it but this site doesn't correspond to desirable RE and can easily be changed to others more readily available (NEB,2020).

The PCR itself consists of an enzymatic reaction with three main stages in a repeating cycle to be repeated 25-35 times (denaturation, annealing, and extension) and two non-cyclical stages that can occur once each outside the cycle (initial denaturation and final extension) controlled with abrupt temperature (T) changes (Figure 2). A complete reaction would encompass all these stages in the following order (Scientific, Brown 2010):

- Initial denaturation – the reaction mixture is heated to *ca.* 95 ºC for 5 minutes. This ensures complete separation of the double-strand DNA (dsDNA) into single strands;
- Denaturation – the first cyclical step with the same goal as the previous step at the same T with a much shorter duration (*ca.* 10 seconds);

- Annealing – a T decrease occurs to the approximate primer T_m to let the primers properly anneal to each of the strands; depending on the DNA polymerase used, optimal annealing can be obtained at about the T_m or at about $T_m - 5$.
- Extension – T is then increased to *ca.* 72 ºC (optimal polymerase T), to start the synthesis of the complementary target sequence;
- Final extension – in this last reactional step, T is kept at 72 ºC for *ca.* 5 minutes to ensure that any incomplete ends get corrected;

With the insert now amplified, the rest of the process follows the traditional cloning protocol to purify the insert, digest it if needed to obtain sticky-ends, and final ligation.

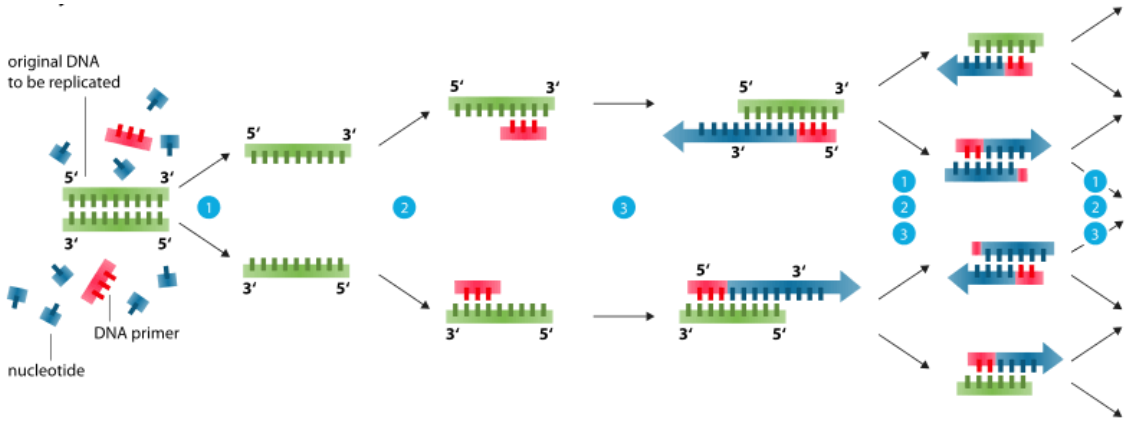


Figure 2: Main three stages of the polymerase chain reaction (PCR). Before these three, usually occurs a one-time initial denaturation with the same temperature as the cyclical denaturation (1) but with much greater duration; following the denaturation is the annealing period (2) and after that the sequence extension (3). After all the cycles finish, in similarity with the denaturation, a one-time final extension occurs with the same temperature as the cyclical extension but with a higher duration. Adapted from: "Schematic drawing of the PCR cycle" by Enzoklop, distributed under a CC-BY 2.0 license (available on Wikimedia Commons).

1.1.1.2. Ligation-independent cloning

Ligation-Independent cloning (LIC) is a variation of the PCR-based cloning method where the enzymatic ligation step between the insert and the vector is substituted by a T4 DNA polymerase mediated reaction. This process cuts down on the need of involving any kind of extra enzymatic reactions to prepare the insert/vector after the PCR (possible vector dephosphorylation, insert phosphorylation, or enzymatic digestions) (Aslanidis and de Jong 1990, Zerbs, Frank et al. 2009).

When amplifying the GOI, a specific overhang is synthesised into the PCR primers resulting in a blunt double-stranded PCR product with the overhang sequence plus the GOI. After the reaction, both the resulting insert and the linearized vector are mixed and treated with T4 DNA polymerase in the presence of only one deoxynucleotide triphosphate (dNTP). In the absence of sufficient dNTPs, the polymerase obtains a 3' → 5' exonuclease activity and begins digesting the insert and vector sequence at the same pace stopping at the first complementary nucleotide to the dNTP present in solution. This process creates complementary sticky ends

between the insert and vector that anneal simply via that complementarity, *i.e.*, the insert and vector anneal by establishing the hydrogen bonds between the two strands and not by having the nucleotides in the insert chain form a phosphodiester bond with the vector strand Figure 3. With the resulting mix, the polymerase is inactivated and the resulting plasmid is transformed into the intended competent host where the plasmid will be replicated producing plasmids where the insert and vector have the phosphodiester bond formed (Aslanidis and de Jong 1990, Zerbs, Frank et al. 2009).

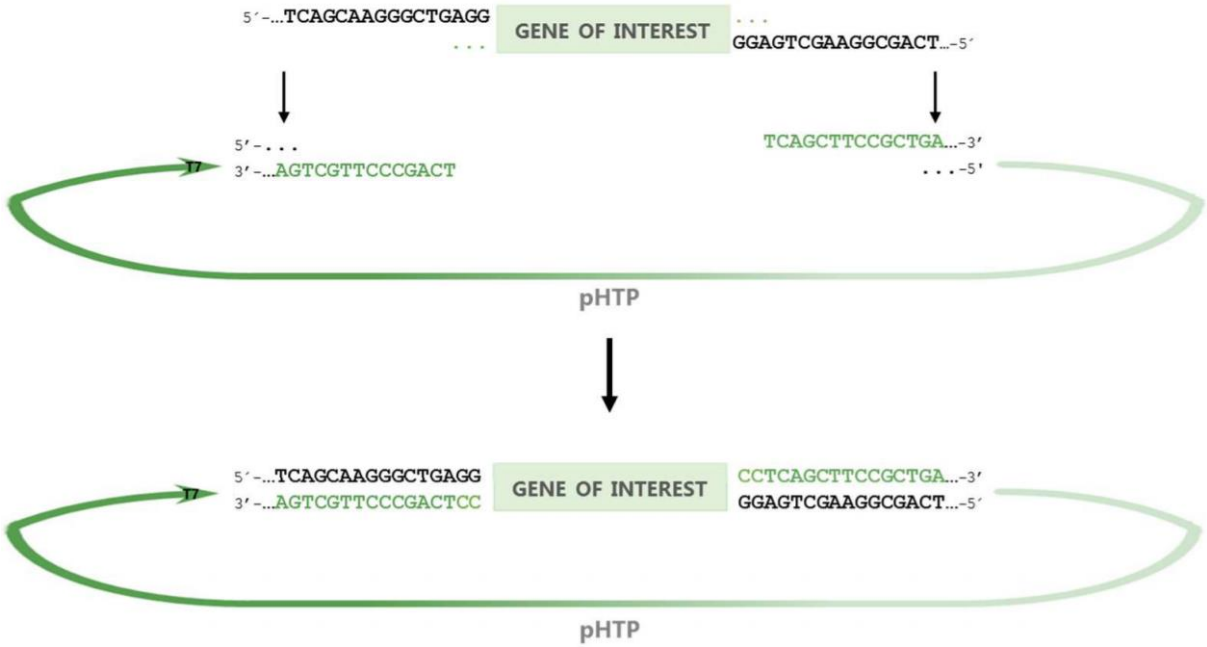


Figure 3: Representation of the Ligation-Independent cloning. After the GOI is amplified using the designed primers capable of introducing the needed overhangs, the insert and vector are treated with the enzyme mixture promoting the annealing of these two sequences. After this treatment, the vector is then transformed for further replication and formation of the phosphodiester bonds between nucleotides (NZYTech 2019).

1.2. HUMAN PROTEINS

From antibodies to insulin, human proteins have been a crucial point of research since the 1970s and have come a long way to better help the lives of countless people. With the lab synthesis of human proteins, it has also been made possible to study certain pathologies that would otherwise be impossible to study without recurring to the use of such a complex host organism (Wuest, Hou et al. 2011).

1.2.1. THE CYTOCHROME P450 SYSTEM

A highly researched group of recombinant human proteins are the human heme proteins, among which the superfamily of cytochromes P450 (CYP), E.C. 1.14.14.1. This protein family

is comprised of a vast number of monooxygenase enzymes with a high level of stereoselectivity and regioselectivity making them highly sought by the pharmaceutical industry in drug metabolism (Manikandan and Nagini 2018).

Since the first published work on CYPs in 1955 (Brodie, Axelrod et al. 1955), much has been researched involving these enzymes, including their mechanism of action. CYPs mediate monooxygenation reactions where one atom of oxygen is inserted into the substrate causing a water molecule to be formed in the process, this reaction behaves cyclically, as represented in Figure 4 (Manikandan and Nagini 2018).

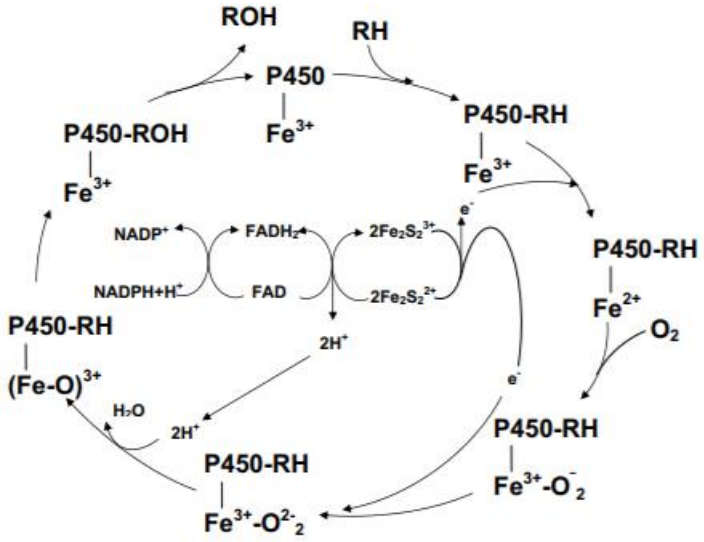


Figure 4: Overall view of the cytochrome P450 monoxygenation catalytic cycle. The RH substrate forms a complex with the P450 while the heme group receives the O_2. After several intermolecular transfers, the substrate receives the needed OH group to become hydrophilic and detaches from the enzyme. In parallel (in the centre) the POR enzyme cycle can be seen providing the needed electrons. Adapted from (Manikandan and Nagini 2018).

In this cycle, a cytochrome P450 oxidoreductase (POR) works in parallel transferring electrons (e-) from NADPH via the central diagram chain, also providing the two needed H^+ ions. From the formation of the CYP-substrate complex, the CYP heme group is reduced, and then an O_2 molecule is incorporated forming the stable complex Fe^{3+}-O_2^- that is further reduced to Fe^{3+}-O_2^{2-}. When this complex is reduced by the protons from the solvent, the bond between the oxygen atoms is broken forming an H_2O molecule and $(Fe\text{-}O)^{3+}$. This $(Fe\text{-}O)^{3+}$ receives an H atom from the substrate forming $FeOH^{3+}$, and the formed OH group is subsequently transferred to the substrate, resulting in the hydroxylated final product. The release of the product from the CYP-Substrate complex leaves a CYP with a regenerated heme group ready to catalyze further reactions (Isin and Guengerich 2007, Manikandan and Nagini 2018).

As more and more different CYPs were discovered, a need to organize and classify them emerged. The nomenclature that was settled uses the CYP abbreviation followed by an Arabic numeral representing its family (*e.g.* CYP1 CYP2, CYP3), a letter for its subfamily (*e.g.* CYP2A, CYP2B, CYP2C), and finally, a second number, to further distinguish isoenzymes (*e.g.* CYP2C1, CYP2C2, CYP2C3). When referring to mutated variants (isoforms) from one

enzyme, an asterisk can be added at the end followed by a number representing the mutation reference (*e.g.* CYP2C8*1, CYP2C8*2, CYP2C8*3) (McDonnell and Dang 2013).

While most studies of CYP structure have focused on the haem-containing active site, from a cloning point of view some insights about the overall 3D structure and cellular location of the proteins are required. CYPs are membrane-associated proteins, particularly to the endoplasmic reticulum. CYPs present a short N-terminal membrane anchoring segment, leading to the catalytic site to be located on the cytosol, with the N-terminal residues on the luminal side of the endoplasmic reticulum (Šrejber, Navrátilová et al. 2018)

1.2.1.1. Cytochrome P450 2C8

Officially known as cytochrome P450 family 2 subfamily C member 8 (CYP2C8), this enzyme comprises *ca.* 7% of the CYP content in the human liver and is responsible for the metabolization of *ca.* 5% of the drugs modified in phase I metabolism, *i.e.* montelukast (used in the treatment of chronic asthma), propacetamol (prodrug of paracetamol), and paclitaxel (used in the treatment of Kaposi's sarcoma and several types of cancer) (Desta and Flockhart 2017). To date, a total of 15 mutations have been registered with either no change (CYP2C8*10) (Hichiya, Tanaka-Kagawa et al. 2005) or a decrease (CYP2C8*4) (Gao, Liu et al. 2010) in enzymatic activity with *in vitro* experiments. Figure 5 depicts a representation of the structure of this protein.

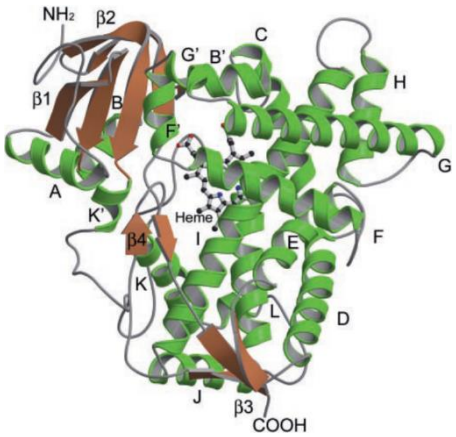


Figure 5: Representation of the CYP2C8 structure. Display of the main structures and the heme group in the centre (ball-and-stick model). Adapted from (Schoch, Yano et al. 2004).

1.2.1.2. Cytochrome P450 3A4

First described in 1988 (Gonzalez, Schmid et al. 1988), cytochrome P450 family 3 subfamily A member 4 (CYP3A4) (Figure 6) is primarily expressed in the liver (95% of the total

expression), with only a small part being expressed in the small intestine. CYP3A4 is responsible for metabolizing a large part (*ca.* 30%) of the drugs consumed by humans.

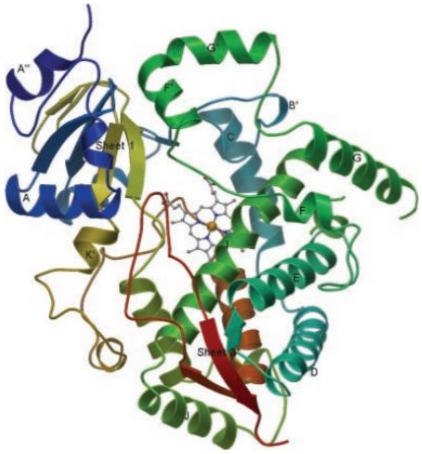


Figure 6: Structural representation of CYP3A4. Display of the main structures and the heme group in the centre (ball-and-stick model). Adapted from (Williams, Cosme et al. 2004).

1.2.1.3. Cytochrome P450 oxidoreductase

The cytochrome P450 oxidoreductase (POR) is an enzyme of extreme importance in all of the aforementioned CYP mechanisms as it mediates e^- transfer to the CYP itself. This enzyme is comprised of three different domains, one domain for flavin mononucleotides (FMN), one domain for flavin adenine dinucleotides (FAD) and NADPH, and a connecting domain between these two (Xia, Panda et al. 2011). POR is a membrane-bound enzyme, anchored to the endoplasmic reticulum and the nuclear envelope by a short N-terminal domain, akin to the anchoring mechanism of CYPs (Wang, Roberts et al. 1997).

As the name implies, this oxidoreductase will oxidize NADPH into $NADP^+$, supplying the two e^- needed to forward the cycle. There are 2 cycles in which the POR operates (in humans), but *in vivo,* most of the time only one is actively running in the liver where the semiquinone $FAD/FMNH^\bullet$ is predominantly found. This essential semiquinone is produced in yet another parallel priming reaction from its fully oxidized form FAD/FMNH. (Montellano 2015)

The main POR cycle (Figure 7) consists of transferring two hydrogen atoms from the NADPH to the FAD molecule, forming $NADP^+$ and $FADH_2/FMNH^\bullet$. Within the flavin complex, an intramolecular transfer forms $FADH^\bullet/FMNH_2$ that will further donate the first e^- to reduce the P450 heme group, ending up with a $FADH^\bullet/FMNH^\bullet$. After this first P450 reduction, another intramolecular transfer takes place to form the more stable $FAD/FMNH_2$, which will then reduce the P450 for the second and final time forming then the original $FAD/FMNH^\bullet$, bringing the cycle to a close (Montellano 2015).

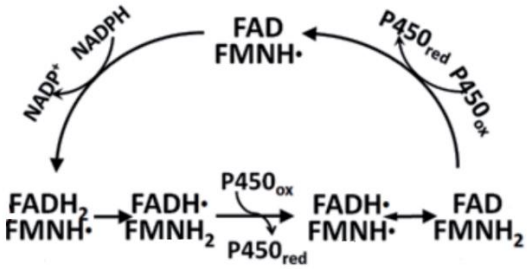


Figure 7: Schematic representation of the cytochrome P450 oxidoreductase (POR) reaction cycle. Oxidation of the NADPH into $NADP^+$ and the following intermolecular transfers to reduce the CYP enzyme. Adapted from (Montellano 2015).

Several human variants have been documented so far. Amongst them, the POR*13 (Q153R) showed a significant improvement in aiding CYP2D6 in the metabolization of certain common substrates (like dextromethorphan, a cough suppressor) by 98% when compared with its wild-type variant (Sandee, Morrissey et al. 2010).

1.2.1.4. UDP-glucuronosyltransferase 2B7

UDP-glucuronosyltransferase 2B7 (UGT2B7) is one of the enzymes that act on phase II of drug metabolism primarily on the liver and kidneys (Fujiwara, Yokoi et al. 2016) but also in smaller amounts in the gastrointestinal tract (Gregory, Lewinsky et al. 2004), and even on the brain (King, Rios et al. 1999).

UGT enzymes use the glucuronidation mechanism to conjugate hydrophilic glucuronic acid molecules (using UDP-glucuronic acid as a co-factor) into the xenobiotic metabolites resultant from phase I drug metabolism. This conjugation will result in a more water-soluble metabolite ready for excretion from the human organism. (Testa and Clement 2015). In Figure 8 it is depicted a structural representation of this enzyme.

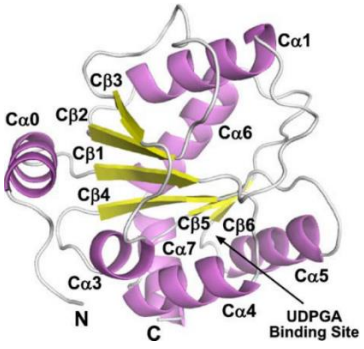


Figure 8: Structural representation of UGT2B7.Cartoon representation of the α helixes and β sheets as well as the UDPGA binding centre. Adapted from (Miley, Zielinska et al. 2007).

2. CYTOCHROME P450 2C8 CLONING

The first enzyme studied in this work was the 2C8 isoform of cytochrome P450 (CYP2C8), where both the traditional cloning and the LIC techniques were employed.

2.1. CYP2C8 TRADITIONAL CLONING

The CYP2C8 cDNA, cloned into a pCW vector (a gift from Joyce Goldstein; Addgene plasmid # 69604; http://n2t.net/addgene:69604; RRID: Addgene_69604, supplied as 25% glycerol stocks of transformed *E. coli* DH5α), translates into a polypeptide chain with 7 amino acid changes in the N-terminal region relative to the canonical protein sequence (UniProtKB ID P10635), introduced to optimize bacterial expression (vector map in Annex 1). The translated sequence of the pCW-CYP2C8 was checked for signal peptides in both eukaryotic and prokaryotic Gram-negative organisms using the SignalP-5.0 server (Almagro Armenteros, Tsirigos et al. 2019). The initial 24-aminoacid sequence has a 16 % probability of being a signal peptide in eukaryotic cells, but no signal peptide for Gram-negative organisms was found.

The pET-28a(+) vector (Novagen, USA) that was used for routine protein expression carries an N-terminal His_6 tag/thrombin/T7 tag configuration in addition to an optional C-terminal His_6 tag sequence (map in Annex 2). The T7 expression region is reversed on the circular map due to the sequence numbering from the pBR322 convention. This vector also features a T7 promoter, a multiple cloning site (MCS) and the *lacI* operon, allowing for protein expression mediated by the T7 RNA polymerase under control of the *lacI* repressor. The MCS of this vector (Figure 9) presents several restriction sites that can be used to drive the introduction of inserts using various approaches.

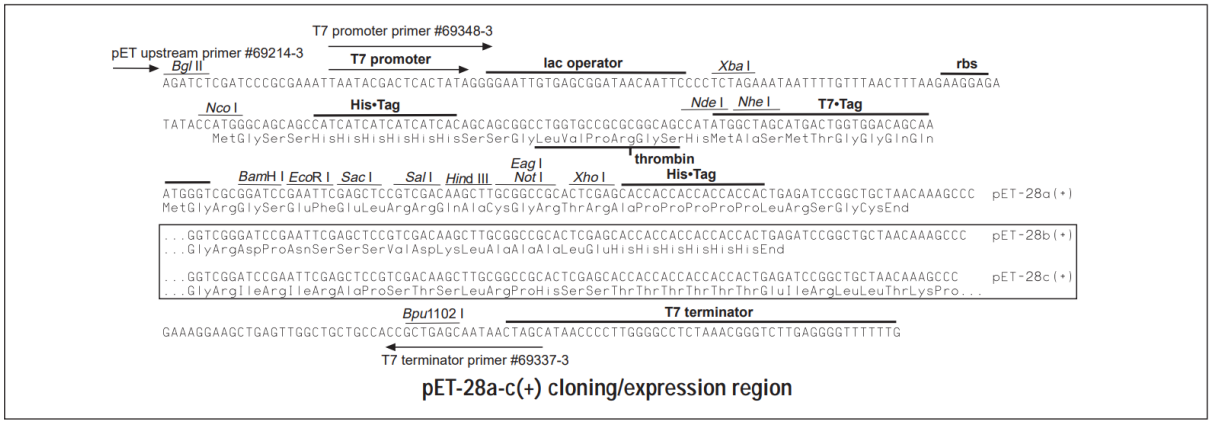


Figure 9: Organization of the pET-28a (+) cloning/expression region. The region flanked by the T7 promoter and the T7 terminator is shown, including restriction enzyme cut sites, tag-coding sequences and the thrombin cut site. Taken from the pET28a-c(+) vector maps (Novagen, reference TB074VM).

From the analysis of the pCW-2C8 plasmid sequence, and considering the protein sequence, a double digestion approach was chosen to sub-clone this gene. pCW-CYP2C8 was digested with *Nde*I and *Hind*III REs to excise the CYP2C8 cDNA, in which the initial methionine was encoded by the ATG sequence of the *Nde*I cut site, and then inserted into the pET28a(+) vector, thus inserting the N-terminal His_6 tag fusion in frame with the CYP2C8 cDNA. The whole sub-cloning process was simulated in Benchling to determine the expected fragment sizes. Digestion of the pCW-CYP2C8 vector with *Nde*I and *Hind*III generates an insert containing the CYP2C8 coding sequence at 1.5 kbp, while digestion of the pET28a(+) with the same RE pair produces a linearized fragment at 5.3 kbp. The ligation product will have a total 6.8 kbp size, from which PCR amplification with primers for the T7 promoter and T7 terminator generates a 1.8 kbp fragment.

To obtain both plasmids in sufficient quantities, the *E.* coli DH5α cells containing the pCW-CYP2C8 and pET-28a (+) plasmids were grown in Lysogeny Broth (LB) (MB145, NZYTech) with the respective selection antibiotics (LB with 30 µg/mL of kanamycin (MB020, NZYTech) for the pET-28a(+)-carrying cells and 50 µg/mL of ampicillin (MB021, NZYTech) for the pCW-CYP2C8-carrying cells). For each strain, two equal liquid growths were conducted with a volume of 10 mL each in 50 mL tubes, to guarantee a minimum aeration headspace of four times the culture volume (Sambrook, Russell et al. 2001).

In parallel to the liquid growth, two solid medium growths were also made for each strain. The solid medium growths consisted of LB Agar (MB118, NZYTech) plates with the respective antibiotics in the same concentrations of the liquid growths. Whilst the liquid growths served to obtain large quantities of cells, the solid plates served to qualitatively evaluate cell growth to more easily assess the presence of contaminations and the morphology of the colonies, and also to obtain, where necessary, isolated colonies. Both the liquid and solid growths were incubated at 37 ºC overnight (*ca.* 18h) with an agitation of 250 rpm (agitation for the liquids only).

With the liquid growth, and after the confirmation of correct colony growth via analysis of the plates, plasmids were extracted using a plasmid purification kit (illustra™ plasmidPrep Mini Spin Kit, GE Healthcare) following the provided protocol in (Annex 12). Extracted DNA was quantified using a nanodrop spectrophotometer (Spectrostar Nano, Biogen), by reading the absorbance at 260 nm and using a modified Lambert-Beer law to calculate the corresponding concentration (Equation 1), where c is the nucleic acid concentration (ng/µL); A_{260} is the absorbance at 260 nm; A_{340} is the absorbance at 340 nm (baseline normalization); S_c is the standard coefficient provided for the measuring equipment (50 µg/mL); and D is the dilution factor.

$$\mathrm{c} = (\mathrm{A}_{260} - \mathrm{A}_{340}) * \mathrm{S}_\mathrm{c} * \mathrm{D} \qquad (1)$$

The absorbance values at 230 nm and 280 nm were also analysed and used to calculate the 260/230 and 260/280 purity ratios through which the levels of chaotropic salts/detergents contamination and protein contamination can be determined, respectively.

After quantification, pCW-CYP2C8 was concentrated and quantified once more, results can be observed in Table 1 alongside respective final volumes obtained. With the liquid growths, stocks were also prepared to store samples at -80 ºC, by mixing 500 µL of 50% (v/v) sterile glycerol (in water) and 500 µL of each growth in a sterile cryovial.

Table 1: pCW-CYP2C8 and pET-28a(+) fractions obtained.Purity ratios and concentration for each fraction obtained via spectrophotometric analysis.

Plasmid	Volume (µL)	Abs_{260}/Abs_{230}	Abs_{260}/Abs_{280}	Concentration (ng/µL)
pCW-CYP2C8	150.0	0.3	2.1	19.7
	170.0	0.3	2.1	17.0
	70.0	0.2	2.0	36.5
	70.0	0.3	1.8	35.8
	550.0	1.3	1.8	55.8
pET-28a(+)	400.0	1.3	1.8	44.8
	400.0	1.2	1.9	41.0
	400.0	1.3	1.9	44.5
	300.0	0.8	1.8	11.1

With the plasmids ready, both the pET-28a(+) and the pCW-CYP2C8 were digested using both *Hind*III and *Nde*I REs (MB097 and MB101, NZYTech, manufacturers protocols in annexes Annex 13 and Annex 14, respectively) for 3 hours at 37 ºC with an inactivation step of 20 minutes at 65 ºC. The digestion reaction followed the protocols provided as a base, with few modifications, and its conditions are presented in Table 2. Digestion time was changed from

the recommended 15 min to 3 hours, following preliminary work. Although *Nde*I is extremely useful for restriction-based DNA work, because it contains the ATG start codon in its recognition sequence, and the pET expression vectors also contain the *Nde*I site in the MCS, *Nde*I generates a two-base overhang, giving rise to products with a lower melting temperature than REs that produces larger overhangs. For this reason, digestion time was increased to 3 hours, as it generally produced higher amounts of the desired digested DNA. Also, this lower melting temperature requires the ligation of *Nde*I-generated fragments to be performed with higher ligase concentrations over a longer ligation time (Chang, Ge et al. 2005).

Table 2: Reaction conditions for pET28a(+) and pCW-CYP2C8 enzymatic digestion

Reagents	**Initial Concentration**	**pET-28a(+)**	**pCW-CYP2C8**	**Final concentration / amount**
*Hind*III		7.00 µL	40.00 µL	-
*Nde*I		7.00 µL	40.00 µL	-
NZYSpeedyBuffer	10x	15.00 µL	100.00 µL	1x
pET-28a(+)	44.80 ng/µL	111.70 µL	-	5.16 µg
pCW-CYP2C8	55.80 ng/µL	-	532.00 µL	29.66 µg
MB H2O		9.30 µL	288.00 µL	-
TOTAL		150.00 µL	1000.00 µL	

To purify the CYP2C8 insert and verify the successful enzymatic digestion, a 1.5% (w/v) agarose gel was made with agarose (MB027, NZYTech) using Tris-acetate-EDTA (TAE) buffer to prepare the gel and run the electrophoresis at 80 V for 1h30m. Figure 10 shows the electrophoresis result for both digested pCW-CYP2C8 and, as control, undigested pCW-CYP2C8. The band at 1.5 kbp corresponds to the excised insert.

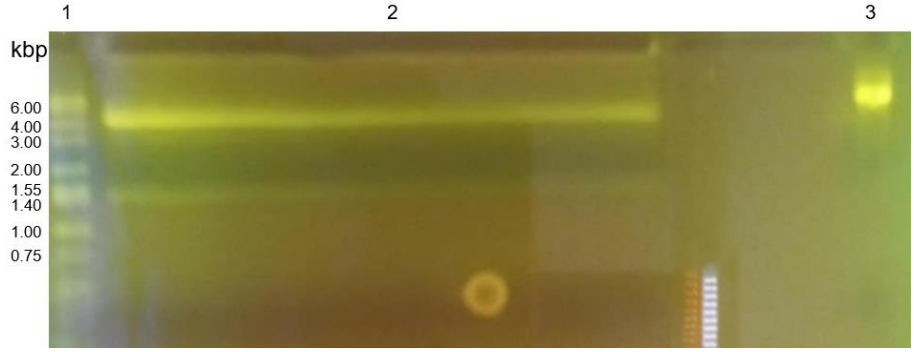


Figure 10: Agarose gel (1.5 %) of the pCW-CYP2C8 enzymatic digestion. The DNA ladder used (lane 1) was Fisher BioReagents™ Routine DNA Ladder (BP2582200, Fisher scientific) and on lane 2 all the digested pCW-CYP2C8 was loaded and a band at 1.5 kbp can be observed, as expected. For comparison, a sample of undigested pCW-CYP2C8 was loaded into lane 3.

With the insert identified in the gel, the corresponding band was cut out and, using a commercial DNA purification kit from agarose gel (AN0070, Canvax), the insert was obtained in a solution. This extracted insert was then quantified. The pET-28a (+) plasmid was also purified via a commercial DNA purification kit for PCR products (AN0063, Canvax) and quantified afterwards. In both kits, the protocols provided were followed without any alterations, and both can be found in Annex 15 and Annex 16, respectfully. The concentration of both sequences is presented in Table 3.

Table 3: Isolated CYP2C8 insert and digested pET-28a(+). Purity ratios and concentration for each fraction obtained via spectrophotometric analysis

DNA	Volume (µL)	260/280	Concentration (ng/µL)
CYP2C8	270.0	2.4	16.3
pET-28a(+)	140.0	1.9	25.8

From this purification, and with both the insert and vector fully prepared, the ligation step followed using T4 DNA Ligase (MB00703, NZYTech). To ensure the maximum use of the vectors supplied to the reaction, a vector:insert molar ratio of 1:10 was selected. Given the concentrations of the insert and vector, and using the protocol provided as a guide (Annex 17), Table 4 depicts the conditions at which the reaction was performed. This reaction took 18 hours (overnight) and occurred at 16 ºC.

Table 4: Ligation reaction conditions for CYP2C8 insertion into the pET28a(+) vector

Component	Concentration	Volume (µL)	Final concentration / amount
MB H2O		3.30	
Reaction buffer	10 x	2.00	1 x
pET-28a(+)	25.80	2.70	69.66 ng
CYP2C8	16.30	11.00	179.30 ng
T4 DNA Ligase	5.00 U/µL	1.00	0.25 U
TOTAL		20.00	

For the transformation step, the now ligated plasmid was inserted in chemically competent *E. coli* DH5α cells. This *E. coli* strain was made competent following a standard chemically competent cell protocol using calcium chloride. Briefly, *E. coli* cells were grown overnight in LB medium without antibiotics. Cells were harvested by centrifugation at 16 000 x *g* for 5 minutes at 4 ºC. The supernatant was discarded, and cells were resuspended in 750 µl of ice-cold $CaCl_2$ 0.1 M, followed by an incubation step in ice for 1h. Then, cells were harvested and resuspended in 50 µl of fresh ice cold $CaCl_2$ 0.1 M. Cells were stored at -80 ºC until used (Sambrook, Russell et al. 2001).

To insert the plasmid into the cells, 20 µL of the ligation reaction (*ca.* 50 ng of DNA) were pipetted into 50 µL of competent *E. coli* DH5α and left for 30 minutes on ice. This process was made on an ice bath at all times, using ice-cold tips and polypropylene tubes, to avoid bringing the competent cells to warmer temperatures that would lead to a fluidification of the bacterial cell membrane. Following the 30-minute incubation, the cell + plasmid mix was transferred to a preheated water/dry bath at 42 ºC for 45 seconds. This transfer has to be very quick from ice to heat to achieve the desired heat shock that will create the necessary fissures on the wall enabling the plasmid to enter the cell. When the 45 seconds have passed, the mix was put on ice again for 2 more minutes before adding 800 µL of Super Optimal Broth with Catabolite repression (SOC) and incubated at 37 ºC; 250 rpm for 45 minutes to let the cells fully regenerate and start expressing the antibiotic resistance without any added stress from antibiotic media. To collect the transformed cells a spindown was made for 20 seconds and cells were resuspended in 100 µL of SOC or LB medium. The cells were plated into LB agar plates with kanamycin 30 µg/mL (KAN30) and incubated at 37 ºC overnight. The resulting colonies were isolated into a new plate to be reincubated, and into liquid growths. The new grown colonies were used to perform a colony PCR to verify the presence of the plasmid whilst the liquid growth were used to make 25% glycerol stocks and plasmid extraction.

Colony PCR was performed in a standard manner using T7 forward (promoter) and reverse (terminator) primers (Table 5) to verify the presence of the desired insert in the T7-controlled expression region. The DNA sample was gathered by collecting, with a sterile loop, a visible mass of each colony into 50 µL of sterile H_2O, followed by intense vortexing and boiling for 10 minutes. With the samples boiled, they were stored in ice till needed for the PCR to avoid

polymerase activity. When needed, the samples were spun down and the supernatant used for the PCR.

Table 5: Primers for colony PCR for the T7 promoter and T7 terminator regions

Name	T_m (ºC)	Length (bp)	Sequence (5' → 3')
T7_PROMOTER	46.8	20.0	TAATACGACTCACTATAGGG
T7_TERMINATOR	52.9	19.0	GCTAGTTATTGCTCAGCGG

The preparation of the PCR reaction is resumed in Table 6 and was based on the protocol provided by the polymerase supplier(MB283, NZYTech, Annex 18), altering the final volume of the reaction to 30 µL. Enough reactions were prepared to have every sample in duplicate (four samples, one for each of the colonies that grew; one for a negative control where pET-28a(+) plasmid was used in place of the plasmid to evaluate the presence of pET-28a(+) without the insert; and one for a neutral control where 15 µL of H_2O where used in place of a sample to evaluate the presence of DNA contaminations in the reagents and primer self-dimerization that could lead to false PCR results) totalling twelve reactions plus one extra to minimize possible pipetting errors (Master Mix of 12+1).

Table 6: PCR reaction mixtures for T7 colony PCR assay. † The volume of the sample is not counted for the final volume of the PCR master mix, since the sample is only added after the mix is already prepared and divided into the tubes. ‡ Final volume of each reaction

Component	Initial concentration	Volumes (µL) for:		Final concentration/amount
		1 reaction	12+1 reactions	
MB H_2O	-	5.46	70.98	-
Reaction buffer	5 x	6.00	78.00	1 x
dNTPs	25.00 mM	0.24	3.12	0.20 mM
P_{Fw}	10.00 µM	1.50	19.50	0.50 µM
P_{Rv}	10.00 µM	1.50	19.50	0.50 µM
Colony sample †	-	(15.00)	(15.00 x 13)	-
Supreme NZYProof DNA Polymerase	2.50 U/µL	0.30	3.90	0.025 U/µL
TOTAL	-	15.00 (30.00)‡	195.00	-

The thermocycler program used for the reaction is presented in Table 7. The program used was based on the program supplied with the polymerase Annex 18, using an annealing temperature (T_{Ann}) calculated for the T7 primers and an extension time calculated according to supplier specifications (rounded up to ensure total sequence extension).

Following the PCR reaction, a 1.5% (w/v) agarose gel electrophoresis (Figure 11) was performed in the same conditions as the prior gel. Unlike the expected positive results, it was not possible to gather any conclusive result at all from the electrophoresis due to the overly smeared bands. Possible causes for this result may have stemmed from the 3'→5' exonuclease activity inherent to the Supreme NZYProof DNA Polymerase (MB283, NZYTech) used in the PCR. Although proofreading is a useful attribute to have on polymerase in cloning PCR, when too much polymerase is present or when the dNTPs get depleted/are in inadequate quantities the exonuclease activity can start degrading the DNA chains and the primers creating secondary artefacts that lead to smeared bands (Longley, Bennett et al. 1990).

Table 7: Thermocycler program for the T7 colony PCR assay.† Using the directions provided by the supplier of 30 s/kbp, and expecting a PCR product of 1.78 kbp, the estimated time would be *ca.* 54 seconds but this duration was rounded up to ensure the complete extension of all sequences

Cycles	Step	T (ºC)	t (s)
	Pre-heat	95	∞
1	Initial denaturation	96	240
35	Denaturation	96	30
	Annealing	55	30
	Extension	72	60†
1	Final Extension	72	600
	Hold	4	∞

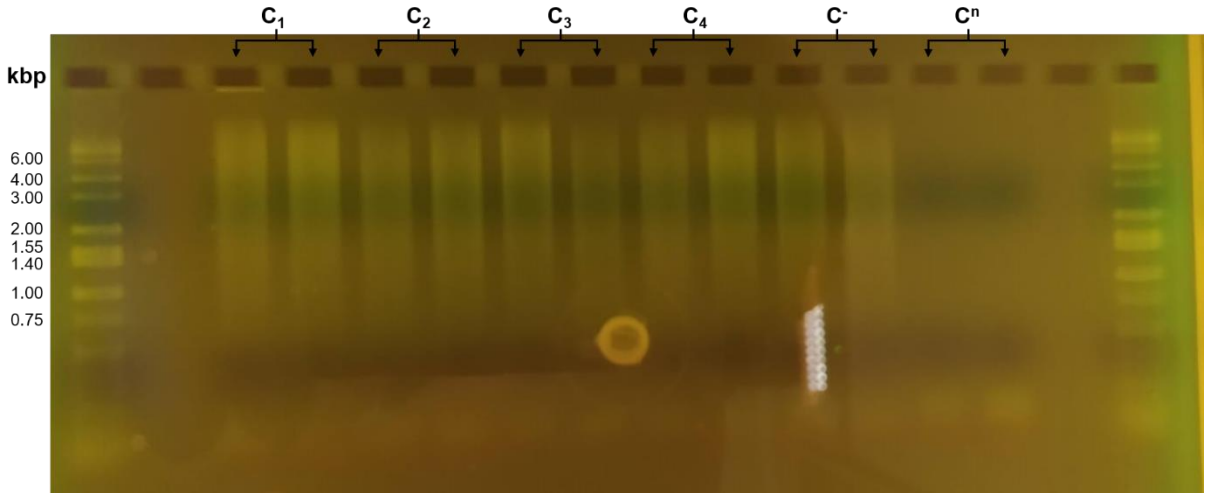


Figure 11: Agarose gel (1.5 %) electrophoresis of the T7 colony PCR products. C_1-C_4 lanes have the samples from the four colonies obtained; C^- has the negative control with pET-28a(+) instead of colony DNA; C^n has the DNA volume substituted for MB H_2O. No relevant result is observable.

Due to the low number of recombinant colonies obtained, and because no clear positive transformant was identified from the colony PCR, an alternative method of evaluating if the colonies were successfully transformed was to proceed to the overexpression of the CYP2C8 protein and then evaluate by protein electrophoresis. To overexpress this protein, the plasmids from the four colonies were extracted using the same commercial kit used before (Annex 12) and transformed into *E. coli* C43(DE3) chemically competent cells, generically, following the transformation protocol used above.

Once grown, each of the four different pET-CYP2C8 clone plasmid containing C43(DE3) cells were used to inoculate a 5 mL LBKAN30 pre-inoculum grown in standard conditions (37 ºC, 250 rpm, overnight) in 50 mL tubes. Using this pre-inoculum, three tubes with 15 mL of LB were inoculated to have a starting optical density at 600 nm ($O.D._{600\ nm}$) between 0.4 and 0.6 (equivalent to adding 375 µL to the fresh medium) for each of the colonies. With the medium inoculated, the cells were allowed to incubate for 6 hours at 37 ºC and then the expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG), a compound analogue to allolactose that doesn't get consumed overtime avoiding the need to constantly supplement the medium with the *lac* operon inducer (lactose (allolactose)). Alongside the IPTG another compound, 5-aminolevulinic acid (5-ALA), was also added as a supplement to boost the efficiency of the expression of heme production (Miura, Ito et al. 2015). Both the supplements were added to a final concentration of 1 mM and the expression took place for 3 hours at 250 rpm and at two different temperatures (one of the tubes was incubated at room temperature (RT) and the other at 37 ºC). With two out of the three inoculated tubes used for induction, the third was left not induced to serve as a point of reference of what proteins are normally expressed by the cells and what proteins would be only expressed due to the presence of the supplements.

To obtain the target protein, the growth was centrifuged for 10 minutes at 5150 g. The supernatant was discarded, the pellet weight was calculated (subtracting the empty tube

weight to the weight of the tube post centrifugation) and the pellet was kept frozen at -20 ºC overnight.

With the frozen pellet, cellular lysis was performed according to the NZY Bacterial Cell Lysis Buffer (MB178, NZYTech) protocol (Annex 19) with all the optional steps, and also using UltraCruz® Protease Inhibitor Cocktail Tablet, EDTA-free (sc-29131, Santa Cruz Biotechnology). The protein fraction was then analysed via sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) The gels used for this process were made with the reagents and corresponding volumes presented in Table 8.

Table 8: Preparation of the SDS-PAGE stacking and resolving gels. Volumes provided for 2 gels

Components	**4% Stacking Gel**	**10% Resolving Gel**
40% Acrylamide/Bis (1610144, Bio-Rad)	1.00 mL	5.00 mL
Tris-HCl [0.5 M], pH 6.8	2.52 mL	-
Tris-HCl [1 M], pH 8.8	-	5.00 mL
SDS 10% (w/v) (1610416, Bio-Rad)	0.10 mL	0.20 mL
MB H_2O	6.36 mL	9.70 mL
TEMED (1610800, Bio-Rad)	0.01 mL	0.01 mL
10 % APS	0.10 mL	0.10 mL
TOTAL	10.00 mL	20.00 mL

After polymerisation, the gels were loaded with the samples. Said samples were prepared by boiling 15 µL of each of the expressed cell colonies with 5 µL of loading buffer. These mixes were then loaded into the gels alongside the molecular weight marker mix (10 µL of Amersham™ ECL™ Rainbow™ Marker - Full range (RPN800E, Sigma-Aldrich) and 5 µL of loading buffer) and the electrophoresis ran at 100 V for 40 minutes (until all the samples reached the top of the resolving gel) and then the voltage was raised to 120 V for the remainder 50 minutes. The resulting gels can be observed in Figure 12.

It was expected to observe a relatively thick band around the 58 kDa (approximate size of the Human CYP2C8 protein) mark but unfortunately, no band outside of the regular bands seen in the non-induced cells was obtained. This lead to the conclusion that the CYP2C8 was not successfully cloned with traditional cloning. Without the confirmation from the colony PCR, it is not possible to confirm if the colonies grown had the intended insert cloned into them, as the pET-28a(+) could have self ligated and during the transformation inserted into the host cell enabling them to grow through the selection media without giving proper results.

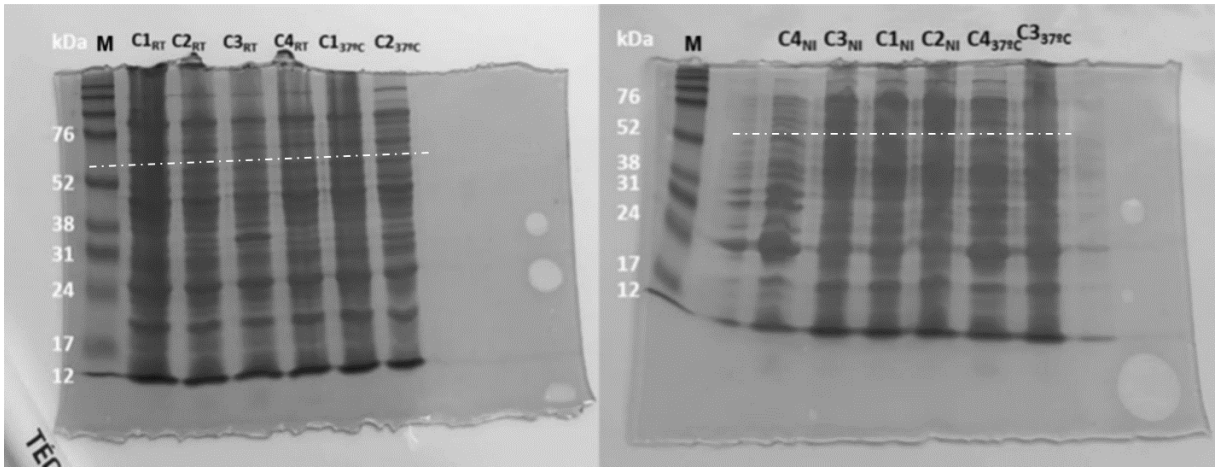


Figure 12: SDS-PAGE analysis of target protein expression induction in *E. coli* C43(DE3) cells transformed with pET-CYP2C8 plasmids.The first lane in each gell had the marker mix loaded into it whilst the remaining lanes had the sample mixes; $C1_{RT}$ - $C4_{RT}$ are the cells Induced at RT; $C1_{37\,ºC}$ – $C4_{37\,ºC}$ are the cells induced at 37 ºC; $C1_{NI}$ – $C4_{NI}$ are the non-induced cells.

As a final attempt with the traditional cloning, the entire cloning process described above was repeated with the introduction of vector dephosphorylation and insert phosphorylation step. To do so a new batch of the same commercial pCW-CYP2C8 in *E. coli* DH5α was grown in the same conditions as before and the plasmid was extracted the extraction resulted in a new fraction with a concentration of 74.69 ng/µL.

This new pCW-CYP2C8 fraction and the previously used pET-28a(+) fraction (44.79 ng/µL) were digested once again with the same *Hind*III and *Nde*I enzymes using the quantities listed in Table 9 for 3 hours at 37 ºC with an inactivation step of 20 minutes at 65 ºC.

Table 9: Reaction mixtures for the *Nde*I/*Hind*III double digestion of the pET28a(+) and pCW-CYP2C8 plasmids. † *Nde*I was used in excess due to the nature of the enzyme being less efficient than the *Hind*III

Reagents	**Initial Concentration**	**pET-28a(+)**	**pCW-CYP2C8**	**Final concentration/amount**
*Hind*III		4.00 µL	14.00 µL	-
*Nde*I		5.00 µL	16.00 µL	-
NZYSpeedyBuffer	10x	10.00 µL	30.00 µL	1x
pET-28a(+)	44.79 ng/µL	67.00 µL	-	3001.60 ng
pCW-CYP2C8	74.69 ng/µL	-	137.00 µL	10232.53 ng
MB H_2O		14.00 µL	108.00 µL	-
TOTAL		100.00 µL	300.00 µL	

To obtain the vector and insert it as pure of a form as possible, the same purification steps were followed as the ones from the previous attempt, resulting in a pET-28a(+) vector with 10.08 ng/µL in 30 µL and an insert with 33.78 ng/µL in 125 µL.

Using alkaline phosphatase (MB01801, NZYTech), the phosphate group at the 5' end of the pET-28a(+) vector was removed to avoid the possibility of recircularization. This reaction was performed following the provided conditions in the enzyme protocol, Annex 20, and with the volumes presented in Table 10.

Table 10: Reaction mixture for dephosphorylation of the pET-28a (+) digested vector

Reagents	**Initial concentration**	**Alkaline Phosphatase**	**Final concentration/amount**
MB H_2O	-	2.00 µL	-
Reaction buffer	10 x	4.00 µL	1.00 x
pET-28a(+)	10.08 ng/µL	30.00 µL	302.40 ng
Alkaline Phosphatase	5.00 U/µL	4.00 µL	0.5 U/µL
TOTAL	-	40.00 µL	-

Although enzymatic digestion of the insert already produces a 5' phosphate group, a portion of the digested CYP2C8 was further phosphorylated using the T4 Polynucleotide Kinase (403005, Bioron) to test if this phosphate group was being created as intended. For this reaction, the commercial protocol from a similar T4 polynucleotide kinase (M0201, New England Biolabs) was followed (Annex 21) with the volumes presented in Table 11.

Table 11: Reaction mixture for the phosphorylation of the digested CYP2C8 insert

Reagents	**Initial concentration**	**Volume (µL)**	**Final concentration / amount**
MB H_2O	-	1.00	-
Reaction buffer	10 x	4.00	1 x
ATP	10.00 mM	4.00	1.00 mM
CYP2C8	42.40 ng/µL	30.00	1272.00 ng
T4 Polynucleotide Kinase	20.00 U/µL	1.00	0.50 U/ µL
TOTAL	-	40.00	-

To ligate the treated vectors and inserts, the same ligase as before was once again used. These ligations were done in four different insert:vector ratios (1:1, 2:1, 10:1, and 20:1) always using dephosphorylated vector but using both the phosphorylated and not phosphorylated inserts. The conditions of these reactions followed the commercial protocol provided from the ligase (Annex 17) and the volumes used for each of the ratios are presented in Table 12 and Table 13.

Table 12: Ligation reaction of the phosphorylated CYP2C8 insert and the dephosphorylated pET-28a (+) vector. The reactions occurred with four different insert:vector ratios (1:1, 2:1, 10:1, and 20:1)

Reagents	**Initial conc**	**1:1**	**2:1**	**10:1**	**20:1**	**Final conc. / amount**
MB H_2O	-	12.25 µL	11.85 µL	8.68 µL	4.72 µL	-
Reaction buffer	10 x	2.00 µL	2.00 µL	2.00 µL	2.00 µL	1 x
pET-28a(+)	7.55 ng/µL	4.36 µL	4.36 µL	4.36 µL	4.36 µL	32.92 ng
CYP2C8	25.34 ng/µL	0.39 µL	0.79 µL	3.96 µL	7.92 µL	-
T4 DNA Ligase	5.00 U/µL	1.00 µL	1.00 µL	1.00 µL	1.00 µL	0.50 U/µL
TOTAL	-	20.00 µL	20.00 µL	20.00 µL	20.00 µL	-

Table 13: Reaction mixtures for the ligation reactions of the untreated CYP2C8 insert and the dephosphorylated pET-28a (+) vector. The reactions were performed with four different insert:vector ratios (1:1, 2:1, 10:1, and 20:1)

Reagents	**Initial conc.**	**1:1**	**2:1**	**10:1**	**20:1**	**Final conc./ amount**
MB H_2O	-	12.35 µL	12.05 µL	9.71 µL	6.78 µL	-
Reaction buffer	10 x	2.00 µL	2.00 µL	2.00 µL	2.00 µL	1 x
pET-28a (+)	7.55 ng/µL	4.36 µL	4.36 µL	4.36 µL	4.36 µL	32.92 ng
CYP2C8	33.78 ng/µL	0.29 µL	0.59 µL	2.93 µL	5.86 µL	-
T4 DNA Ligase	5.00 U/µL	1.00 µL	1.00 µL	1.00 µL	1.00 µL	0.50 U/µL
TOTAL	-	20.00 µL	20.00 µL	20.00 µL	20.00 µL	-

As the last step of the transformation, plasmids were transformed into *E. coli* C43(DE3) cells and plated into SOC plates supplemented with kanamycin (30 µg/mL) (SOCKAN30). The plates were incubated at 37 ºC for 18 hours, at which point the colonies were annotated and the plates were grown for 24 more hours following a colony recount. At this stage, the grown colonies from the 20:1 ratio group were picked and isolated into a gridded SOCKAN30 plate and grown for another 42 hours.

With these grown colonies a group-colony PCR was performed in which the colonies were grouped in groups of approximately 10 colonies each, totalling 10 group-colony PCRs for the cells with phosphorylated insert and 9 group-colony PCRs for the cells where the insert was not phosphorylated before ligation.

A standard electrophoresis analysis followed, in a 1.5% (w/v) agarose gel (Figure 13), showing no positive results neither for the colonies with a phosphorylated insert nor the colonies with the untreated insert. Although no colony group showed the expected band at the 1.785 kbp mark, groups 1-3 and 6-9 showed a band close to the 0.6 kbp mark. A probable cause for the appearance of these bands could be from improper primer annealing to the plasmids due to temperature, shorter extension period than the theoretically calculated.

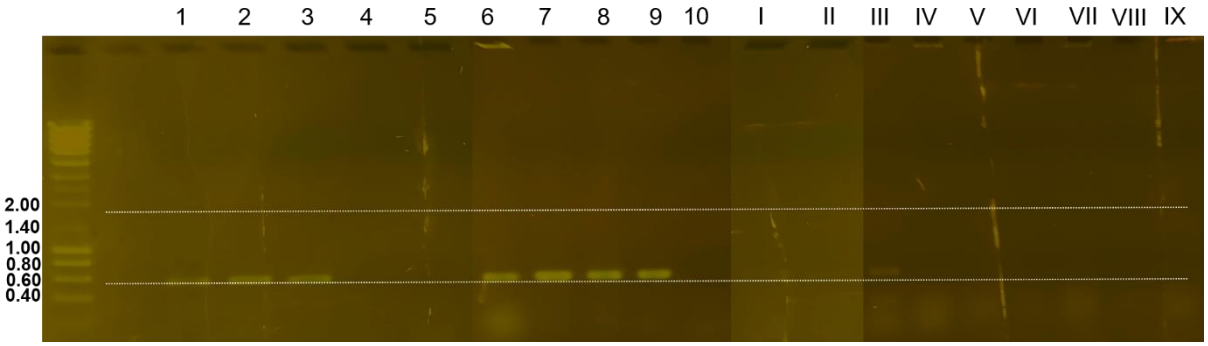


Figure 13: Agarose gel (1.5 %) electrophoresis of group-colony PCR of transformed pET-CYP2C8 cells. In Arabic numerals are the groups with the phosphorylated insert and in roman numerals the colonies with unphosphorylated inserts.

As no positive results were obtained, despite the different approaches tried, and repeated several times, a different approach was taken with a ligation-independent technique. This change in technique was prioritized as the relative complexity of the aforementioned process (dephosphorylation of the vector, phosphorylation of the inserts, mass colony PCR) could have easily led to errors at many stages and the growth times of the transformed colonies ended up taking more than twice the time as expected from this *E. coli* strain growth time (42h vs the usual 18h).

2.2. CYP2C8 LIGATION-INDEPENDENT CLONING

For this ligation-independent cloning (LIC), two receiving vectors were chosen, pHTP1 and pHTP9 (MB282 and MB324, NZYTech, Portugal), part of the NZYEasy Cloning & Expression kit. Like the pET28a(+) vector, both offer the capacity for protein expression using IPTG induction, and also have the possibility of introducing an N- and/or C-terminal His_6 tag, thus facilitating the purification of the protein. Also, pHTP9 plasmid leads to the expression of a fusion protein where the desired protein comes with an N-terminal Green Fluorescent Protein which can be proteolytically cleaved out if necessary. The vector maps and brochures from

each kit are presented in the annexe (pHTP1 - Annex 3; pHTP9 - Annex 4; Cloning kit - Annex 22).

The CYP2C8 insert was PCR amplified from the pCW-2C8 source plasmid (same source plasmid used in 2.1. CYP2C8 traditional cloning) using primers specifically designed for this purpose (Table 14). While the 3' region of each primer is inserted specific, the 5' region will anneal with the receiving plasmid. To promote primer-target annealing, a nucleotide was changed, as illustrated in Figure 14. These primers were designed using DNASTAR Lasergene PrimerSelect and the whole LIC process was simulated using Benchling. Primers designed for the NZYEasy Cloning&Expression kits are independent of the destination vector, and the same set of primers can be used for both pHTP1 and pHTP9 destination vectors.

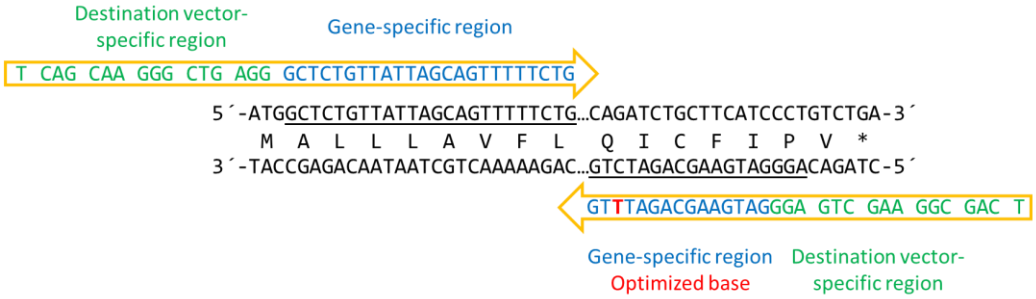


Figure 14: Ligation-independent primer design for CYP2C8 cloning. To promote the primer-target annealing, a cytosine nucleotide was changed to thymine.

Simulation of the cloning process with Benchling indicates that PCR amplification of the CYP2C8 coding sequence with the designed primers will produce a 1.5 kbp fragment. Upon insertion of this fragment in the destination vectors, the pHTP1-CYP2C8 construct will have a total 6.8 kbp size and the pHTP9-CYP2C8 construct will have a total size of 7.6 kbp. The presence of the insert in these constructs can be verified with primers that are specific for the T7 promoter and terminator or using pHTP specific primers. Using T7-annealing primers, amplification from the pHTP1-CYP2C8 construct will produce a 1.8 kbp fragment, while amplification from the pHTP9-CYP2C8 construct generates a 2.5 kbp fragment. Using pHTP specific primers, where the reverse primer anneals to the T7 terminator and the forward primer anneals just upstream of the ATG codon for the inserted fragment, PCR amplification from either the pTHP1-CYP2C8 or the pHTP9-CYP2C8 constructs produces a 1.8 kbp fragment.

Table 14: Primers used for CYP2C8 amplification via ligase-independent cloning. Required overhangs shaded in grey

Name	Length (bp)	Tm (ºC)	Sequence (5' →3')
2C8_EASY_FW_NCHIS	40.0	67.5	TCAGCAAGGGCTGAGGGCTCT GTTATTAGCAGTTTTTCTG
2C8_EASY_RV_NCHIS	31.0	65.3	TCAGCGGAAGCTGAGGGATGA AGCAGATTTG

Using the designed primers and the Supreme NZYProof 2x Green Master Mix (MB285, NZYTech) the T_{ann} optimization for the insert was carried out with the reaction components in Table 15; volumes are given for a single reaction.

The reactional mixture was then transferred to the thermocycler with a 95 ºC pre-heat step and reacted following the parameters in Table 16. An electrophoresis agarose gel 1.5% (w/v) was then run (Figure 15) to properly analyze the reaction results.

Table 15: Reaction components for the CYP2C8 PCR T_{ann} optimization with LIC primers

Components	Initial concentration	Volumes (µL)	Final concentration/ammount
MB H2O	-	5.83	-
2C8_EASY_FW_NCHIS	100.00 µM	0.06	0.50 µM
2C8_EASY_RV_NCHIS	100.00 µM	0.06	0.50 µM
pCW-CYP2C8	82.08 ng/µL	0.30	25.00 ng
Supreme NZYProof Green Master	2 x	6.25	1 x
TOTAL	-	12.50	-

Table 16: PCR program for the T_{ann} optimization with LIC primers.PCR was performed using the Supreme NZYProof DNA polymerase (Supreme NZYProof 2x Green Master Mix

Cycles	Step	T (ºC)	t (s)
1	Initial denaturation	96	240
35	Denaturation	96	30
	Annealing	[65,71]; Δ = 2	30
	Extension	72	60
1	Final Extension	72	600
	Hold	4	∞

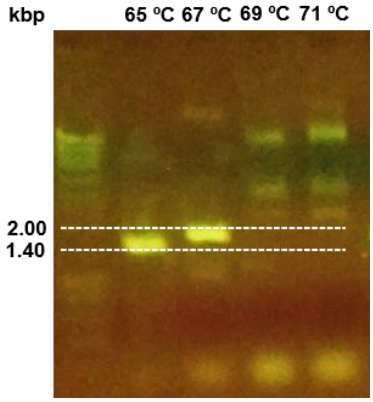


Figure 15: Agarose gel electrophoresis of the CYP2C8 PCR T_{ann} optimization from 65 ºC to 71 ºC.DNA marker used NZYDNA Ladder III (MB044, NZYTech).

Typically, to optimize the T_{ann} of a pair of primers a gradient of temperatures is selected from temperatures close to T_m – 5 ºC. In this case, the optimization reaction occurred alongside the T_{ann} optimization for all other genes in this study (not just the CYP2C8 insert) leading to the out of norm range presented. Non the less, both 65 ºC and 67 ºC presented results compatible with the theoretical result of a band with 1.493 kbp. From these temperatures, the amplification of the insert was then done following the same protocol of the optimization but using a T_{Ann} = 67 ºC in a bigger reaction volume (150 µL). The PCR resulting mixture was separated on a 1.5 % agarose gel and the relevant bands were cut and the DNA was purified as described earlier; the final insert concentration was 23.48 ng/µL.

The purified insert was used for the LIC reaction, following the manufacturer's guide (Annex 22 – NZYEasy Cloning & Expression kits I and IX protocol) and the reactional mix is described in Table 17; the thermocycler program used is presented in Table 18.

Table 17: LIC reaction for the ligation of the CYP2C8 insert with both the pHTP1 and pHTP9 vectors

Components	**Initial concentration**	**pHTP1 LIC**	**pHTP9 LIC**	**Final concentration/ammount**
MB H_2O	-	0.50 µL	0.50 µL	-
Reaction Buffer	10 x	1.00 µL	1.00 µL	1 x
CYP2C8	23.48 ng/µL	6.40 µL	6.40 µL	150.27 ng
pHTP1	-	1.00 µL	-	-
pHTP9	-	-	1.00 µL	-
NZYEasy enzyme mix	-	0.50 µL	0.50 µL	-
TOTAL	-	12.50 µL	62.50 µL	-

Table 18: Thermocycler program for the LIC CYP2C8 insertion into pHTP1 and pHTP9 vectors

Cycles	**Step**	**T (ºC)**	**t (s)**
1	LIC reaction	37	3600
		80	600
		30	600
1	Hold	4	∞

Following the LIC reaction, the plasmids were transformed similarly to the previous transformations. The only alteration to the provided protocol was the competent cells used (same *E. coli* DH5α used in 2.1. CYP2C8 traditional cloning instead of the advised NZY5α cells) and the volume of said cells and ligation product (instead of 100 µL of competent cells + 10 µL of product, 50 µL of cells were used with 5 µL of product). The transformed cells were then plated in LBKAN30 agar plates and incubated overnight at 37 ºC resulting in the colonies observed in Figure 16.

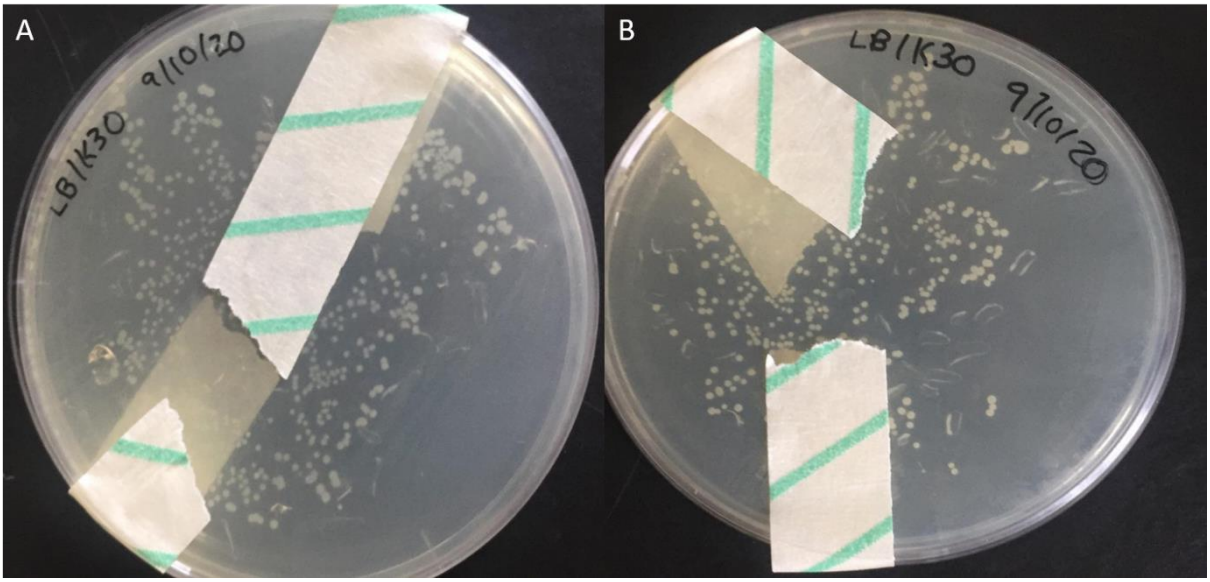


Figure 16: 24-hour growth of *E.coli* DH5α cells transformed with (A) pHTP1-CYP2C8 ligation product and (B) pHTP9-CYP2C8 ligation product.

Colonies were then isolated into new plates and regrown at the same parameters. From the isolated colonies, a colony PCR was performed to evaluate the transformation efficiency using specific primers (Table 19) and following the colony PCR protocol used in 2.1. CYP2C8 traditional cloning with the reaction components described in Table 6. The pHTP sequencing primers are designed so that the reverse primer (pHTPUni_RV) anneals at the T7 terminator region of the plasmid, and can be used for all pHTP vectors, while the forward primers (pHTP1_FW and pHTP9_FW) are vector-specific and anneal just upstream the 5' region of the inserted fragment, generating amplicons with sizes similar to those obtained by amplifying pET28a(+)-based constructs with T7 promoter and terminator specific primers.

Table 19: Primers used for colony PCR of transformed pHTP1-CYP2C8 and pHTP9-CYP2C8 cells

Name	T_m (ºC)	Length (bp)	Sequence (5'→3')
pHTP1_FW	57.2	28	GCGAAATTAATACGACTCACTATAGGGG
pHTP9_FW	59.7	25	GAATGAAAAACGCGACCACATGGTG
pHTPUni_RV	56.1	24	GGTTATGCTAGTTATTGCTCAGCG

When analysing the results of the PCR (Figure 17), near all samples had a nonspecific band of over 10 kbp and nonspecific smears under 0.2 kbp. These bands can be related to primer dimerization where the forward and reverse primers form a stable enough attachment at their 3' ends for the polymerase to then attach to this site and begin elongating them forming a new template DNA sequence to be further amplified in the future cycles. For the 10 kbp nonspecific

bands, a probable cause can be from genomic DNA contamination from other lab work occurring at the time. From all the colonies tested, the G7 and G8 colonies showed the presence of a band with a size of 1.00 kbp and 1.40 kbp, respectively, whilst none of the pHTP1-CYP2C8 colonies showed any specific band.

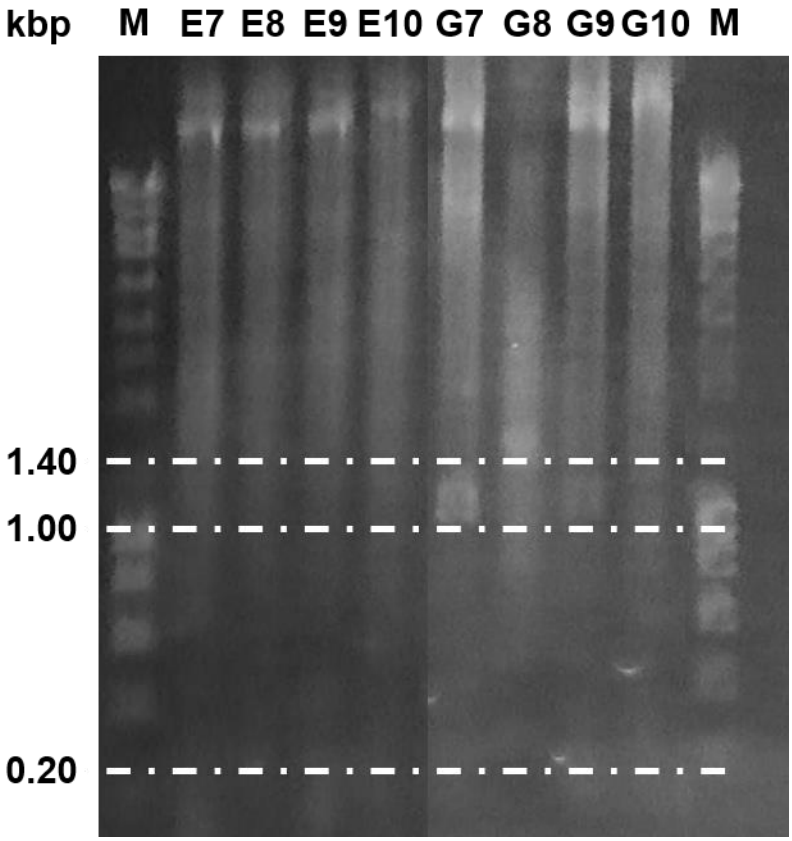


Figure 17: Agarose gel electrophoresis (1.5 %) of colony PCR analysis from the transformed *E. coli* DH5α cells with pHTP ligated products. Cells were transformed with the results of pHTP1-CYP2C8 (E7-E10, using pHTP1_FW and pHTPUni_RV primers) and pHTP9-CYP2C8 ligations (G7-G10, using pHTP9_FW and pHTPUni_RV primers).

Although not in the right size, the presence of these bands raised awareness that a correctly transformed plasmid could be present in these colonies. To verify if this was true, these clones were grown in LBKAN30 liquid medium overnight at 37 ºC to extract their plasmids, following Annex 12**,** and the PCR was repeated using these isolated plasmids instead of direct colony samples. From this PCR, the same three bands persisted in the samples but toke a shift in size showing bands with *ca*. 2.50 kbp, 7.50 kbp, and more than 10.00 kbp (Figure 18). The presence of the band *ca.* 2.50 kbp confirms that the extracted plasmids presented the intended CYP2C8 insert. The vector map for this newly constructed pHTP9-CYP2C8 plasmid is in Annex 5.

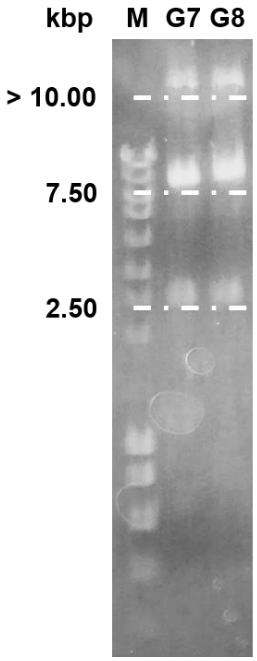


Figure 18: Agarose gel electrophoresis (1.5 %) of PCR amplification of the extracted plasmids from G7 and G8 colonies (pHTP9-CYP2C8) using the T7 primers.

3. CYTOCHROME P450 3A4 CLONING

Similarly to the rhCYP2C8, the process used to obtain the rhCYP3A4 also followed both traditional cloning as well as ligation-independent cloning.

3.1. CYP3A4 TRADITIONAL CLONING

The source plasmid used in the cloning of the CYP3A4 protein was the pCR4-TOPO-CYP3A4 (HsCD00341290, Harvard PlasmID Database, vector map in Annex 6), to obtain the GOI via PCR amplification, for subsequent insertion into the pET-28a(+) expression vector (Novagen, USA). This plasmid contains the complete coding sequence for the CYP3A4 protein. No clear definition of the signalling peptide could be found for this protein, neither in the literature nor using computational approaches. As such, primers were designed to PCR amplify the whole coding sequence, while simultaneously inserting *Nde*I and *Hind*III cut sites at the 5' and 3' regions, respectively; these enzymes were checked to be non-cutters of the coding sequence using NEB cutter 2.0.

From an analysis of the pCR4-TOPO-CYP3A4 plasmid, the amplification of the GOI via PCR was preferred, as opposed to the double enzymatic digestion approach used in 2.1. CYP2C8 traditional cloning, since this approach enables inserting new REs recognition sequences near the GOI. These recognition sequences need to be added to the insert for the GOI to be situated between the *lac* operon (enabling protein expression control) and the His_6 tag to facilitate purification after ligation to the pET-28a (+) vector. The primers used for the amplification are described in Table 20. PCR amplification from the pCR4-TOPO-CYP3A4 plasmid with these primers generates a 1.6 kbp fragment. Ligation of this fragment into a *Nde*I/*Xho*I-digested pET28a(+) vector produces a 6.9 kbp plasmid, from which the insert can be PCR amplified with primers for the T7 promoter and terminator regions, producing a 1.8 kbp fragment.

Table 20: Primers pair designed for the amplification of the GOI CYP3A4. The forward primer inserts a *Nde*I recognition sequence (underlined) upstream of the GOI while the reverse inserts an *Xho*I recognition sequence (underlined) downstream of the GOI

Name	T_m (ºC)	Length (bp)	Sequence (5' → 3')
CYP3A4_Fw_NdeI	62.4	31.0	ACAGTACATATGGCTCTCATC CCAGACTTGG
CYP3A4_Rev_XhoI	65.3	31.0	ACTCTCGAGAGGGCGAATTGA ATTTAGCGGC

E. coli containing the source plasmid for the CYP3A4 gene was grown in two 50 mL centrifuge tubes with 10 mL of LBAMP50 overnight at 37 ºC and agitation of 250 rpm. The plasmid was then extracted from 18 mL of the cell growth using the plasmid extraction kit (illustra™ plasmidPrep Mini Spin Kit, GE Healthcare) altering the volume of elution buffer used from 100 µL to 50 µL (Annex 12**)** resulting in a pCR4-TOPO-CYP3A4 solution with 280.4 ng/µL, a 260/280 ratio of 1.8 and a volume of 300 µL. The remaining cell growth was used for 25 % (v/v) glycerol stocks.

To optimize the best annealing temperature for the amplification primers, six PCRs were made using different annealing temperatures ranging from 67 ºC to 72 ºC. The polymerase used for the reactions was the One-Fusion high-speed high-fidelity Polymerase (S450, GeneON) and the master mix reagents are detailed in Table 21. The PCR program used is given in Table 22.

Table 21: Reaction mix for the PCR T_{ann} optimization of the CYP3A4 amplification primers

	Initial concentration	Volume (µL) for		Final concentration/amount
		1 reaction	6+1 reactions	
MB H_2O	-	11.57	70.98	-
Reaction buffer	2.5 x	10.00	70.00	1 x
dNTPs	10.00 mM	0.50	3.50	0.20 mM
CYP3A4_Fw_NdeI	10.00 µM	1.25	8.75	0.50 µM
CYP3A4_Rev_XhoI	10.00 µM	1.25	8.75	0.50 µM
pCR4-TOPO-CYP3A4	304.40 ng/µL	0.18	1.26	-
One-Fusion polymerase	2.00 U/µL	0.25	1.75	0.02 U/µL
TOTAL		25.00	175.00	

Table 22: PCR program for the T_{ann} optimization of the CYP3A4 amplification primers

Cycles	Step	T (ºC)	t (s)
1	Initial denaturation	98	300
35	Denaturation	98	10
	Annealing	[67,72]; Δ = 1	30
	Extension	72	50
1	Final Extension	72	180
	Hold	4	∞

The PCR reaction products were analysed via electrophoresis with a 1.5% (w/v) agarose gel (Figure 19) ran at standard conditions. Although all the samples presented a smear, it was possible to identify two distinct bands of greater intensity on the sample that had the T_{ann} at 72 ºC with the size of *ca.* 1.6 kbp and 2.0 kbp. The 1.6 kbp is the intended amplicon containing the CYP3A4 (theoretical size of 1.637 kbp) whilst the 2.0 kbp band can be ruled out as a nonspecific band derived from possible lower primer stringency.

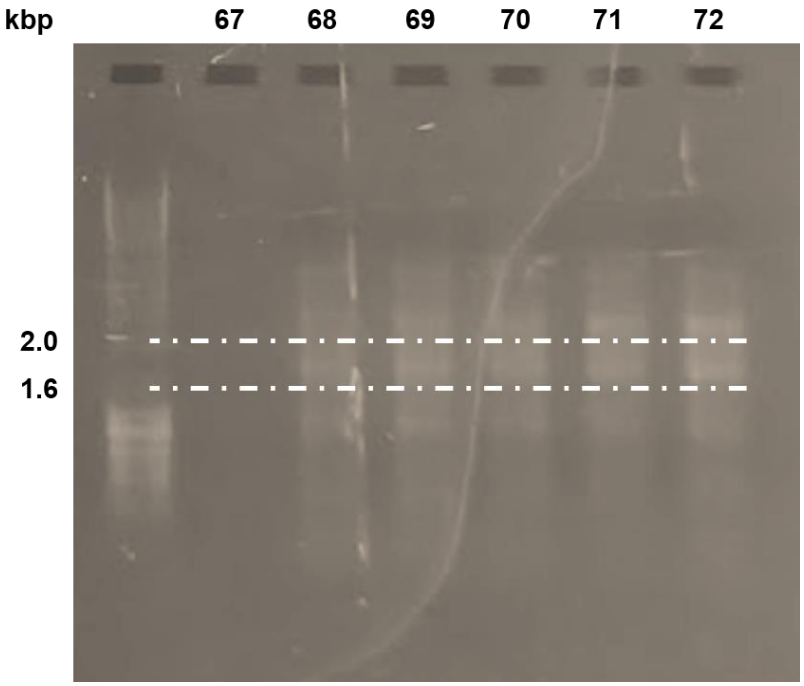


Figure 19: Agarose gel (1.5 %) electrophoresis of the CYP3A4 T_{ann} optimization PCR products.Six different annealing temperatures were used, from 67 ºC to 72 ºC. The presence of two bands was registered, one with a size of *ca.* 1.6 kbp (intended size) and a nonspecific with *ca.* 2.0 kbp.

A larger-scale amplification was done scaling the master mix used in the previous optimization from 6+1 reactions to 48+1 reactions. The PCR program used was the same described in Table 22 without the temperature gradient (substituted with a fixed T_{ann} = 72 ºC) and the master mix is described in Table 23.

Table 23: CYP3A4 PCR amplification

	Initial conc.	Volume (µL) for		Final conc./amount
		1 reaction	48+1 reactions	
MB H_2O	-	23.30	163.10	-
Reaction buffer	2.50 x	20.00	140.00	1.00x
dNTPs	10.00 mM	1.00	7.00	0.20 mM
CYP3A4_Fw_NdeI	10.00 µM	2.50	17.50	0.50 µM
CYP3A4_Rev_XhoI	10.00 µM	2.50	17.50	0.50 µM
pCR4-TOPO-CYP3A4	304.40 ng/µL	0.20	1.40	-
One-Fusion polymerase	2.00 U/µL	0.50	3.50	0.02 U/µL
TOTAL	-	50.00	350.00	-

An electrophoresis 1% (w/v) agarose gel was run (80 V) to purify the reaction products and obtain an isolated 1.6 kbp amplicon separated from the 2.0 kbp amplicon (Figure 20). The 1.6 kbp amplicon was then extracted and further purified with a commercial DNA purification kit from agarose gel (AN0070, Canvax) following the protocol provided without alterations (Annex 15). This extraction resulted in a CYP3A4 insert solution with 56.83 ng/µL in a volume of 300 µL (260/280 purity of 1.7).

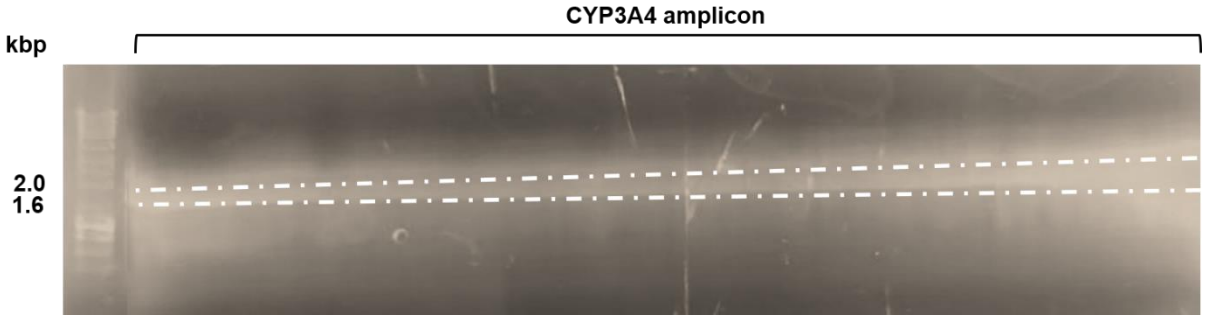


Figure 20: Agarose gel (1.5%) electrophoresis of the pCR4-TOPO-CYP3A4 insert PCR amplification. Two bands were present, the intended near 1.6 kbp and a nonspecific near 2.0 kbp.

Both the insert and vector were then digested using REs *Nde*I (Annex 14) and *Xho*I (Annex 23), with the reactions described in Table 24. These digestions (after being purified with a DNA commercial kit (AN0063, Canvax)) resulted in an insert solution with 42.40 ng/µL in 390 µL (260/280 purity of 1.7) and a vector solution with 9.14 ng/µL in 30 µL (260/280 purity of 1.8).

Table 24: Reaction mixtures for the enzymatic digestion of the CYP3A4 insert and PET28a(+) vector with *Nde*I and *Xho*I

Components	Initial concentration	Insert	Vector	Final concentration/amount
MB H2O	-	3.00 µL	13.00 µL	-
NZYSpeedyBuffer	10.00 x	38.00 µL	8.00 µL	1.00 x
CYP3A4	56.83 ng/ µL	300.00 µL	-	17049.00 ng
pET-28a(+)	44.80 ng/µL	-	50.00 µL	2240.00 ng
*Nde*I	10.00 U/µL	20.00 µL	5.00 µL	-
*Xho*I	10.00 U/µL	19.00 µL	4.00 µL	-
TOTAL	-	380.00 µL	80.00 µL	-

Similar to the CYP2C8 cloning, the vector went through alkaline phosphatase and some digested insert was also phosphorylated. These reactions utilized the same enzymes from 2.1-CYP2C8 traditional cloning and are described in Table 25 (alkaline phosphatase) and Table 26 (insert phosphorylation).

Table 25: Reaction mixture for the alkaline phosphatase-catalysed dephosphorylation of the pET-28a(+) *Nde*I/*Xho*I digested vector

Reagents	Initial concentration	Alkaline Phosphatase	Final concentration/amount
MB H_2O	-	2.00 µL	-
Reaction buffer	10 x	4.00 µL	1 x
pET-28a(+)	9.14 ng/µL	30.00 µL	274.20 ng
Alkaline Phosphatase	5 U/µL	4.00 µL	0.5 U/µL
TOTAL	-	40.00 µL	-

Table 26: Reaction mixture for the phosphorylation of the *Nde*I/*Xho*I digested CYP3A4 insert

Reagents	**Initial concentration**	**T4 polynucleotide kinase**	**Final concentration/amount**
MB H2O	-	7.40 µL	-
Reaction buffer	10 x	4.00 µL	1 x
ATP	10.00 mM	4.00 µL	1.00 mM
CYP3A4	42.40 ng/µL	23.60 µL	274.20 ng
T4 Polynucleotide Kinase	20.00 U/µL	1.00 µL	0.50 U/ µL
TOTAL	-	40.00 µL	-

With the vectors and inserts now properly treated, the ligation step followed using the same ligase used for previous ligations. The ligations took place in four different insert:vector ratios (1:1, 2:1, 10:1, and 20:1) using dephosphorylated vector and using both the phosphorylated and untreated inserts. The ligation reactions followed the commercial protocol from the ligase Annex 17 and the volumes used are depicted in Table 27 and Table 28.

Table 27: Ligation reaction mixture of the phosphorylated CYP3A4 inserts and the dephosphorylated pET-28a (+) vectors. The reactions occurred with four different insert:vector ratios (1:1, 2:1, 10:1, and 20:1)

Reagents	**Initial conc.**	**1:1**	**2:1**	**10:1**	**20:1**	**Final conc./amount**
MB H_2O	-	11.80 µL	11.40 µL	8.23 µL	4.27 µL	-
Reaction buffer	10 x	2.00 µL	2.00 µL	2.00 µL	2.00 µL	1 x
pET-28a (+)	6.85 ng/µL	4.81 µL	4.81 µL	4.81 µL	4.81 µL	32.95 ng
CYP3A4	40.00 ng/µL	0.39 µL	0.79 µL	3.96 µL	7.92 µL	-
T4 DNA Ligase	5.00 U/µL	1.00 µL	1.00 µL	1.00 µL	1.00 µL	0.50 U/µL
TOTAL	-	20.00 µL	20.00 µL	20.00 µL	20.00 µL	-

Table 28: Ligation reaction mixture of the untreated CYP3A4 inserts and the dephosphorylated pET-28a (+) vectors. The reactions occurred with four different insert:vector ratios (1:1, 2:1, 10:1, and 20:1)

Reagents	Initial conc.	1:1	2:1	10:1	20:1	Final conc./amount
MB H_2O	-	11.90 µL	11.60 µL	9.26 µL	6.33 µL	-
Reaction buffer	10 x	2.00 µL	2.00 µL	2.00 µL	2.00 µL	1 x
pET-28a (+)	6.85 ng/µL	4.81 µL	4.81 µL	4.81 µL	4.81 µL	32.95 ng
CYP3A4	42.40 ng/µL	0.29 µL	0.59 µL	2.93 µL	5.86 µL	-
T4 DNA Ligase	5.00 U/µL	1.00 µL	1.00 µL	1.00 µL	1.00 µL	0.50 U/µL
TOTAL	-	20.00 µL	20.00 µL	20.00 µL	20.00 µL	-

The ligations were then transformed into *E. coli* C43(DE3) cells following the protocol defined in 2.1. CYP2C8 traditional cloning plated into SOCKAN30 agar plates and incubated for 18 hours at 37 ºC. Once this 18 hour period elapsed the number of colonies grown was registered, the plates were incubated for another 24 hours and the colonies recounted after that second growth period. The colonies from the 20:1 insert:vector ratio group were picked and isolated into a new SOCKAN30 agar plate and grown for another 42 hours.

Once grown, colonies were grouped in groups of ten colonies each and a group-colony PCR was performed. A total of ten colony groups were made for the colonies that had the insert phosphorylated and nine colonies groups for the transformed cells with untreated inserts. The reactions were then analysed through a 1.5% (w/v) agarose gel electrophoresis (Figure 21) not showing any band with the expected size of *ca.* 1.8 kbp but showing instead a non-specific band with a size *ca.* 0.6 kbp on the groups 1-3 and 6 that could indicate the presence of the intended template but wrong annealing of the primers due to multiple factors.

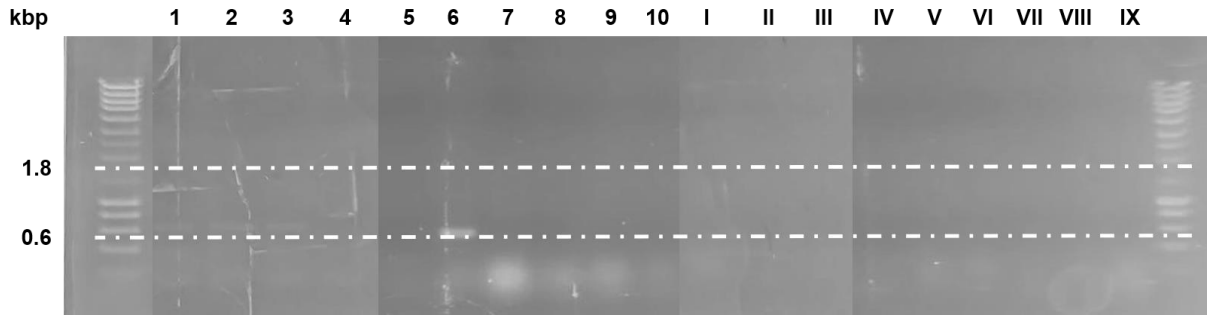


Figure 21: Group-colony PCR of the pET-CYP3A4-transformed *E. coli* C43(DE3) colonies using T7 primers. The samples represented in Arabic numerals correspond to the group-colony samples (10 colonies each) in which the CYP3A4 insert was phosphorylated before ligation; the samples in roman numerals correspond to the group-colony samples (10 colonies each) in which the CYP3A4 insert was untreated before ligation.

With the colony PCR returning negative results, a new approach to this protein cloning was taken using the ligation-independent technique. This technique had a much lower complexity relative to the specific steps this traditional cloning had involved that could have lead to an increase in methodology errors.

3.2. CYP3A4 LIGATION-INDEPENDENT CLONING

The ligation-independent cloning of the CYP3A4 protein closely followed the methodology used in the 2.2 CYP2C8 ligation-independent cloning as these two proteins were cloned in parallel to each other using the same vectors used for the CYP2C8, the pHTP1 and pHTP9 (MB282 and MB324, NZYTech, Portugal). The vector maps and protocols from the cloning kit are present in Annex 3**,** Annex 4**,** and Annex 22**.**

Following the same procedure from 3.1 CYP3A4 traditional cloning, the insert was amplified via PCR from the pCR4-TOPO-CYP3A4 using specific primers needed for this technique. The primers were designed using DNASTAR Lasergene PrimerSelect and following the design guidelines offered by the cloning kit of these vectors (Table 29).

Table 29: Amplification primers for the ligase-independent approach to CYP3A4 cloning

Name	Tm (ºC)	Length (bp)	Sequence (5' → 3')
3A4_Fw_NCHis	67.1	36.0	TCAGCAAGGGCTGAGGGCTCT CATCCCAGACTTGGCC
3A4_Rv_NCHis	65.4	34.0	TCAGCGGAAGCTGAGGGGCTC CACTTACCGTACC

To optimize the T_{ann} of this primer pair, a PCR with T_{ann} gradient was made using the Supreme NZYProof 2x Green Master Mix (MB285, NZYTech) polymerase. The master mix volumes are presented in Table 30 and the PCR program used was the same used in Table 16.

Table 30: Reaction mixtures for the PCR T_{ann} optimization of the primer pair used for the amplification of the CYP3A4 insert

Components	Initial conc.	Volume (µL) for		Final conc./amount
		1 reaction	4+1 reactions	
MB H2O	-	6.04	30.20	-
3A4_Fw_NCHis	100.00 µM	0.06	0.30	0.50 µM
3A4_Rv_NCHis	100.00 µM	0.06	0.30	0.50 µM
pCR4-TOPO.CYP3A4	304.42 ng/µL	0.08	0.40	25.00 ng
Supreme NZYProof Green Master Mix	2.00 x	6.25	31.25	1.00 x
TOTAL	-	12.50	62.50	-

The PCR reactions were then analysed in 1.5 % (w/v) agarose gel electrophoresis (80 V) showing the presence of the amplified insert with a band at *ca.* 1.5 kbp (expected amplicon size of 1.538 kbp) using T_{ann} of 65, 67, and 69 (Figure 22). To stay in conformity with the T_{ann} selected in 2.2 CYP2C8 ligation-independent cloning, the T_{ann} for this primer pair was set as 67 ºC.

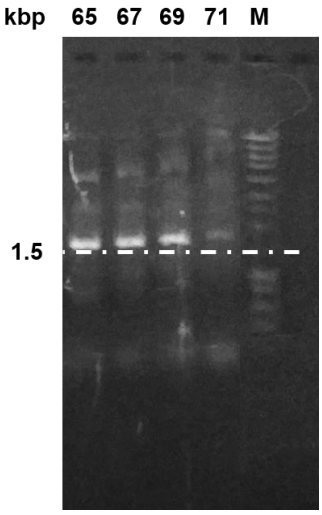


Figure 22: Agarose gel (1.5 %) electrophoresis of the PCR products for T_{ann} optimization of the primer pair used for the CYP3A4 insert amplification. All the samples went through the same reaction, changing the T_{ann} used (65 ºC, 67 ºC, 69 ºC, and 71 ºC).

With the T_{ann} optimized, the insert amplification followed. This amplification used the same reagents as

Table 30 but with a final reaction volume of 150 µL (corresponding to 12 reactions), and the same PCR program used for the T_{ann} optimization, changing only the temperature gradient to a static temperature of 67 ºC. To purify the insert, a 1.5% agarose gel electrophoresis was run and the isolated band cleaned with the same commercial kit used in prior clonings for this effect (Annex 15) obtaining an insert solution of 41.18 ng/µL.

The LIC reaction followed the provided protocol Annex 22 with the volumes depicted in Table 31 and with the same thermocycler program described in the LIC of the CYP2C8 protein (Table 18).

Table 31: LIC reaction mixtures for the ligation of the CYP3A4 insert with both the pHTP1 and pHTP9 vectors

Components	**Initial concentration**	**pHTP1 LIC**	**pHTP9 LIC**	**Final conc./amount**
MB H2O	-	0.50 µL	0.50 µL	-
Reaction Buffer	10.00 x	1.00 µL	1.00 µL	1.00 x
CYP3A4	41.18 ng/µL	3.70 µL	3.70 µL	152.37 ng
pHTP1	-	1.00 µL	-	-
pHTP9	-	-	1.00 µL	-
NZYEasy enzyme mix	-	0.50 µL	0.50 µL	-
TOTAL	-	12.50	62.50	-

The pHTP1-CYP3A4 and pHTP9-CYP3A4 plasmids were then transformed following the cloning kit transformation protocol with the same modifications made in the transformation of the pHTP1-CYP2C8 and pHTP9-CYP2C8 plasmids where the host cells used were the same *E. coli* DH5α used for all the prior transformations (as opposed to the NZY5α cells advised) and the volume of competent cells and ligation product where halved (from 100 µL + 10 µL to 50 µL + 5 µL, respectively). The transformation product was then plated in LBKAN30 agar plates (as commercially advised) and incubated at standard conditions (overnight at 37 ºC) until transformants grew (Figure 23).

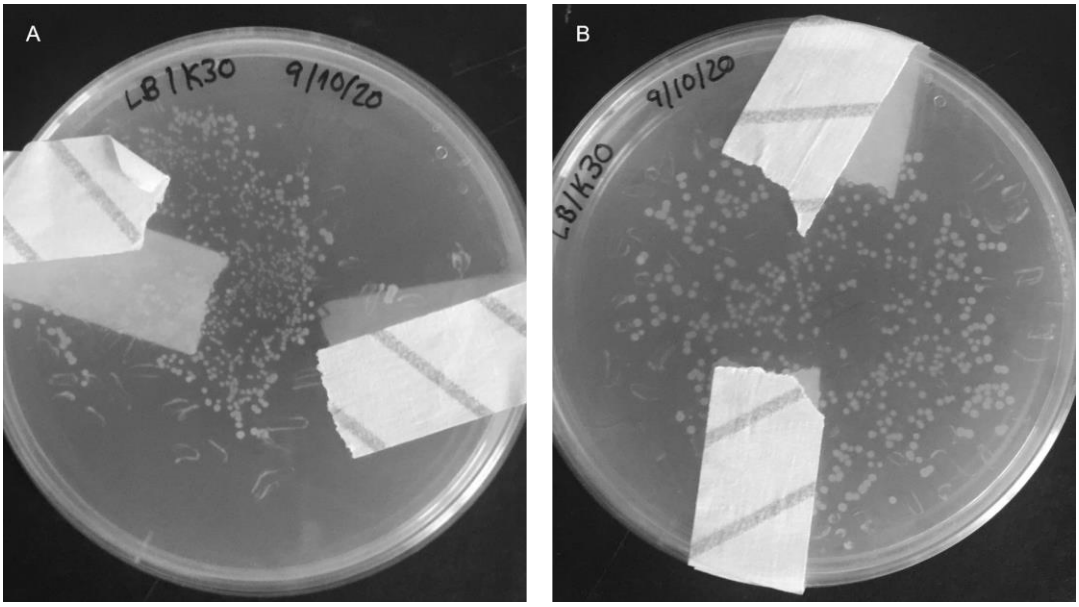


Figure 23: *E.* coli DH5α cells transformed with potential pHTP1-CYP3A4 (A) and pHTP9-CYP3A4 (B), via the LIC technique.

The transformed colonies were then isolated into a new LBKAN30 plate and grown at the same conditions as the first plates, from which a colony PCR was performed to ensure the presence of the intended CYP3A4 insert in the plasmids of these transformants. The colony PCR used the primers shown in Table 19 and following the reaction components of Table 6 and using the thermocycler program in Table 32. This PCR was analysed in a 1.5% (w/v) agarose electrophoresis gel (ran at 80 V) obtaining no amplification at the expected size of *ca.* 1.80 kbp, only appearing the same bands of *ca.* 0.20 kbp and over 10.00 kbp (Figure 24) similar to the colony PCR from section 2.2.

Table 32: Colony PCR thermocycler steps with corresponding temperatures and durations

Cycles	Step	T (ºC)	t (s)
1	Initial denaturation	95	300
35	Denaturation	94	30
	Annealing	57	30
	Extension	72	85
1	Final Extension	72	600
	Hold	4	∞

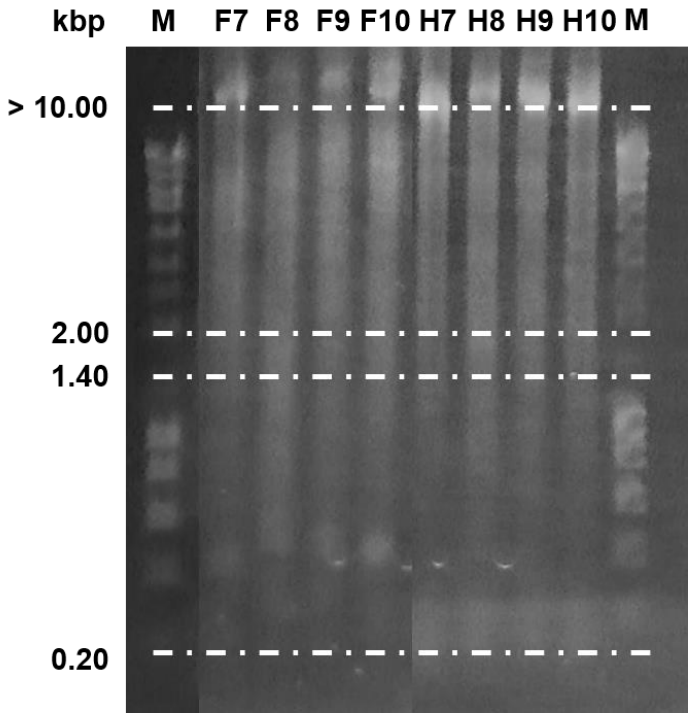


Figure 24: Agarose gel (1.5%) electrophoresis of the colony PCR products of the pHTP1-CYP3A4 (F7 – F10) and pHTP9-CYP3A4 (H7 – H10) isolated colonies.

As none of the colonies tested showed the expected amplicon, a new colony PCR was made using extracted plasmid from the grown colonies. The previously isolated pHTP9-CYP3A4 colonies were grown overnight at 37 ºC in 10 mL of LBKAN30 and extracted following Annex 12**.** The PCR reaction was repeated using the T7 primers (Table 5) and using 50 ng of each plasmid as a DNA template maintaining the remaining volumes and adjusting the final volume with MB H_2O accordingly.

From this new PCR, four main bands appeared when running the products through a 1% (w/v) agarose electrophoresis gel at 80 V (Figure 25). Three of these bands were nonspecific bands with their sizes close to 0.60 kbp, 2.00 kbp, and 4.00 kbp, and the fourth one had a size very close to 2.50 kbp which puts it in line with the theoretical size expected (2.519 kbp) from this PCR had the plasmids been successfully ligated. The vector map for the constructed pHTP9-CYP3A4 is given in Annex 7.

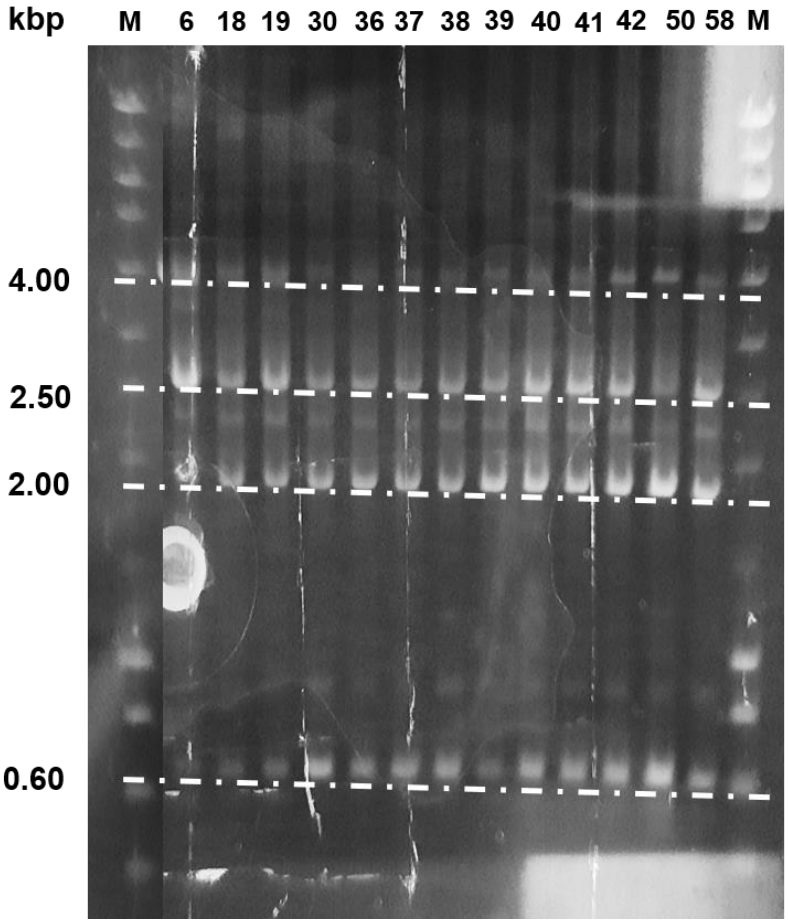


Figure 25: Agarose gel (1.5%) electrophoresis of the colony PCR of the extracted plasmids from the pHTP9-CYP3A4 transformants. The numbers above the samples are indicative of colony tracking IDs.

4. CYTOCHROME P450 OXIDOREDUCTASE CLONING AND EXPRESSION

This chapter presents the work performed to sub-clone the cytochrome P450 oxidoreductase gene to obtain an *E. coli* strain capable of over-expressing this protein.

4.1. CYTOCHROME P450 OXIDOREDUCTASE TRADITIONAL CLONING

The cloning of the human CYP reductase didn't stray too far from the previously presented traditional clonings. The original plasmid used to obtain the GOI was the pCMV-SPORT6-POR (HsCD00327178, Harvard PlasmID Database, map in Annex 8) and the final expression vector was once again the pET-28 a (+) (69864, Novagen). The source plasmid contains the complete coding sequence for the canonical isoform of the protein. A zero-cutter analysis of the coding sequence, performed with NEBCutter v2.0 (Vincze, Posfai et al. 2003), identified *Nde*I, *Nhe*I, *Hind*III and *Xho*I as zero-cutter for the coding sequence; as *Nde*I and *Nhe*I are present upstream in the pET28a(+) MCS, they were chosen for posterior work in the 5' region of the insert; in parallel, as *Hind*III and *Xho*I are located downstream in the MCS, they were chosen for work in the 3' region of the insert.

While a computational signal peptide search did not identify any signal peptides, POR is a protein that localizes to the endoplasmic reticulum during processing and is known to be well associated with the membrane (Lu, Junk et al. 1969, Pandey and Flück 2013). The UniProt database (UniProt 2020) entry for this protein (ID P16435) lists the N terminal region (residues 2 to 21) to be lumenal and indicates that residues 22 to 42 form a transmembrane helix; the remainder of the protein is cytoplasmatic. The TMHMM server (Sonnhammer, von Heijne et al. 1998) indicates that the sequence presents a likely TM helix formed by residues 20 to 42. Building on this information, different primer pairs were designed, corresponding to possible expression products, and also considering the possibility that the N-terminal residues (before the TM helix) may correspond to a signalling peptide. In version 1, residues 1 to 42 are lost, and NdeI and HindIII restriction sites are introduced. In version 2, the N-terminal region of the protein is also lost, and restriction sites are created for NheI and HindIII. In version 3, the TM region is maintained, and sites for NheI and HindIII are introduced. This strategy is shown in Figure 26. The primers used through the cloning process are presented in Table 33.

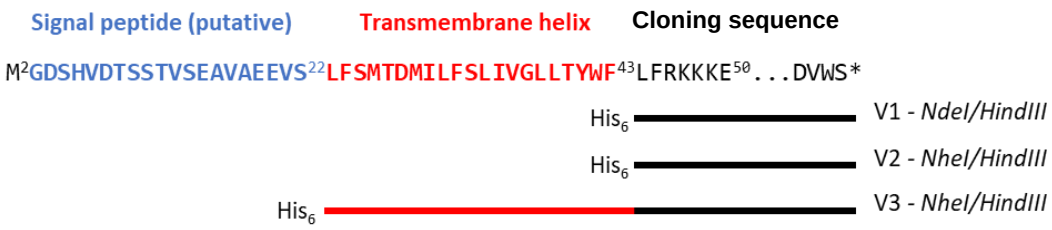


Figure 26: Resulting PCR amplicons from cytochrome P450 amplification using the different primer pairs designed. Both V1 and V2I primer pairs result in a protein without the initial transmembrane helix while V2II results in the main POR protein with the transmembrane helix included. All of the pairs exclude the initial signal peptide.

Table 33: List of primers designer for the POR cloning with respective T_m and length

	Name	Tm (ºC)	Length (bp)	Sequence (5' → 3')
V1	POR_Fw_NdeI	59.2	31.0	TCTCCTACATATGTGGTTCCT CTTCAGAAAG
	POR_Rev_HindIII	64.5	30.0	CCAAGCTTAAGCCAAACACA CCCAGGAGAC
V2	POR_Fw_NheI	62.3	31.0	TCTCGCTAGCTACTGGTTCCT CTTCAGAAAG
	POR_Rev_HindIII	64.5	30.0	CCAAGCTTAAGCCAAACACA CCCAGGAGAC
V3	POR_Chis_Rev_HindII	68.7	29.0	GCAGGAAGCTTGCTCCACA CGTCCAGGGA
	POR_TM_Rev_NheI	69.9	30.0	AACGCTAGCGACTCCCACGT GGACACCAGC

E. coli DH5α cells containing the pCMV-SPORT6-POR plasmid were grown in 30 mL of LB supplemented with carbenicillin (CAB) at a final concentration of 50 µg/mL (LCAB50) overnight at 250 RPM. The plasmid was then extracted from 27 mL using the same protocol used in earlier chapters and 25 % (v/v) glycerol stocks were made using the remainder 3 mL. The extraction resulted in a plasmid solution with a concentration of 26.33 ng/µL, a 260/280 purity of 1.8, and a volume of 850 µL.

Three different pairs of primers (Table 34) were used to optimize the amplification primer T_{ann} and two different PCR programs with different extension times (Table 35), totalling six different PCR parameters. Since the optimization occurs with a gradient of six temperatures, this brings the number of distinct reactions to thirty-six.

Table 34: pCMV-SPORT6-POR reaction mixtures for PCR T_{ann} optimization

	Volume (µL) per reaction	**Volume (µL) for 6+1 reactions:**		
		V1	**V2**	**V3**
MB H_2O	9.85	68.95	68.95	68.95
Buffer [2.5 x]	10.00	7.00	7.00	7.00
dNTPs [10 mM]	0.50	3.50	3.50	3.50
POR-FW-NDE1 [10 µM]		8.75	-	-
POR-RV-HIND3 [10 µM]		8.75	8.75	-
POR-FW-NHE1 [10 µM]	1.25 P_{Fw} / P_{Rv}	-	8.75	-
POR-CHIS-RV-HIND3 [10 µM]		-	-	8.75
POR+TM-NHE10 [10 µM]		-	-	8.75
pCMV-SPORT6-POR [26.33 ng/µL]	1.90	13.30	13.30	13.30
One-Fusion polymerase (2 U/µl) (S450, GeneOn)	0.25	1.75	1.75	1.75
TOTAL	20.00	175.00	175.00	175.00

Table 35: pCMV-SPORT6-POR gradient PCR programs. Between the two programs, the extension step varies

Cycles		**T (ºC)**	**t (s)**
1	Initial denaturation	98	300
35	Denaturation	98	10
	Annealing	[67,72]; Δ = 1	30
	Extension A	72	35
	Extension B	72	70
1	Final Extension	72	180
1	Hold	4	∞

From all the thirty-six reactions only the reaction using the extension time of 70 seconds at T_{ann} of 67 ºC showed an intense band with the wrong size (0.5 kbp instead of the expected 2 kbp), as shown in Figure 27.

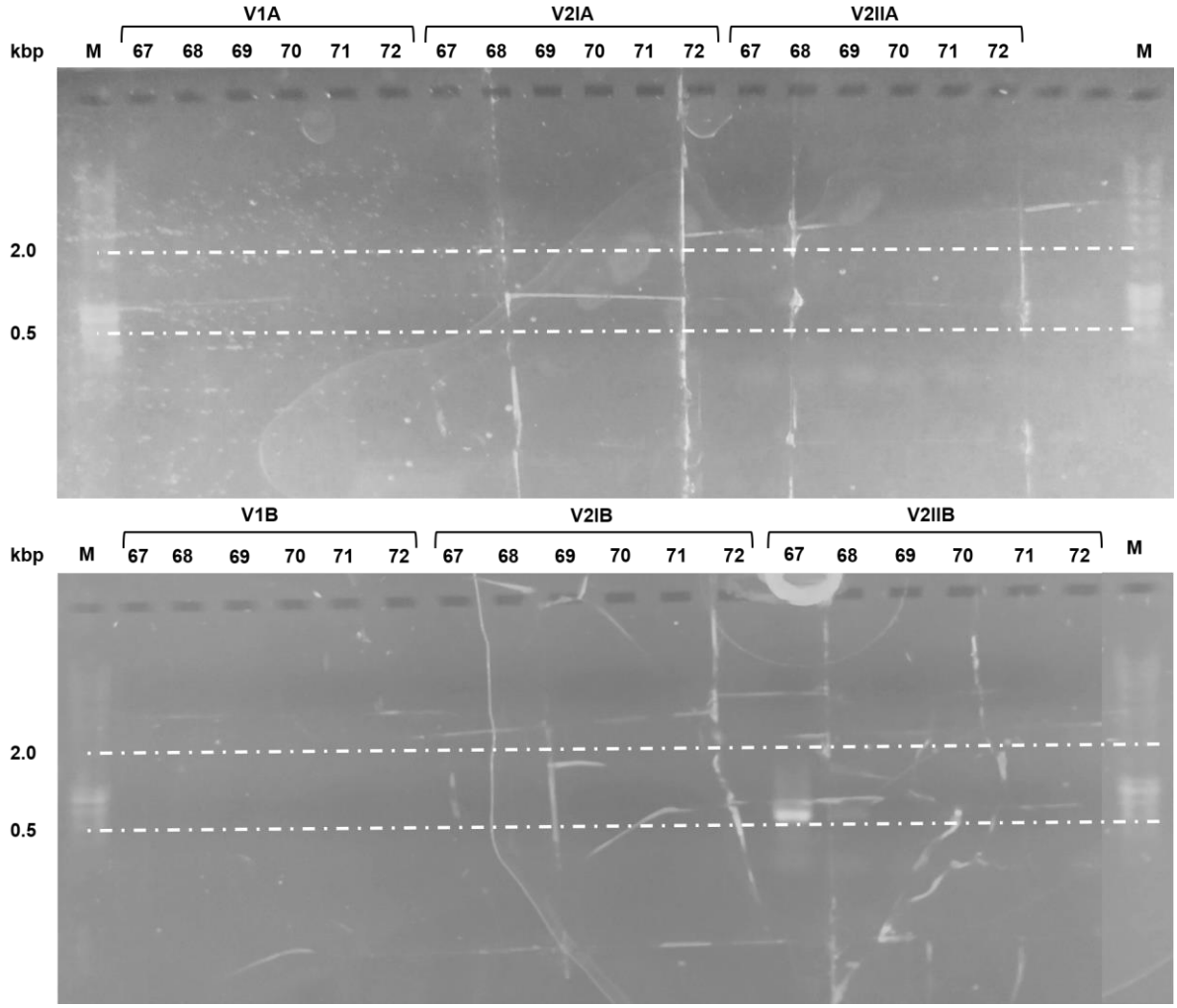


Figure 27: Agarose gel (1.5 %) electrophoresis of the POR T_{ann} primer optimization PCR products.The reactions are divided into two gels were the first gel corresponds to all the reactions with the PCR Extension A and the second one corresponds to the reactions that used the PCR Extension time B. in each of the gels, the groups with V1 correspond to the reactions with the primer pairs POR-FW-NDE1 + POR-RV-HIND3; the groups with V2 correspond to the reactions with the primer pairs POR-RV-HIND3 + POR-FW-NHE1; the groups with V3 correspond to the reactions with the primer pairs POR-CHIS-RV-HIND3 + POR+TM-NHE10.

In face of these results, some PCR optimization approaches were tested. The above PCR amplifications were repeated with the addition of 3 and 5 % (v/v) dimethylsulfoxide (DMSO) to the PCR mix. DMSO acts by eliminating secondary structure motifs in DNA, particularly important in high GC situations, and leads to a diminished annealing temperature (Chester and Marshak 1993). Neither the 3 % nor the 5 % reactions produced any relevant bands.

Similarly, magnesium ions can also play an important role in optimizing primer annealing as they contribute to the denaturation of the template strands. Also, most DNA polymerase buffers already contain $MgCl_2$ at *ca.* 1.5 mM, as magnesium is an essential cofactor for DNA

polymerase activity. The above reactions (at 0, 3 and 5 DMSO) had been performed with the standard manufacturer-supplied reaction buffer, at a final concentration of 1.5 mM, and were repeated with the addition of $MgCl_2$ to a final 2.0 and 2.5 mM concentration. Again, no relevant bands were obtained. Non-systematic PCR mixtures with simultaneous addition of up to 9 % DMSO and 4.0 mM $MgCl_2$ were also tested, to no avail.

As no bands could be obtained by PCR amplification of the POR coding sequence, a LIC-based approach, as used in previous chapters, was tested.

4.2. CYTOCHROME P450 OXIDOREDUCTASE LIGASE-INDEPENDENT CLONING

Similarly to the process described above, a LIC-based approach was used to try to sub-clone the POR gene into pET28a(+). Using the same NZYEasy Cloning&Expression kit, specific primers were designed to PCR-amplify the POR coding sequence with the appropriate overhangs for subsequent insert into pHTP vectors. The primer design approach is shown in Figure 28, and the primer information is given in Table 36.

Figure 28: Ligation-independent primer design for POR cloning.

Table 36: Primers used for POR amplification via ligase-independent cloning with the required overhangs shaded in grey

Name	Length (bp)	Tm (ºC)	Sequence (5' →3')
2C8_EASY_FW_NCHIS	36.0	>75.0	TCAGCAAGGGCTGAGGGGAGACTCCCAC GTGGACAC
2C8_EASY_RV_NCHIS	36.0	>75.0	TCAGCAAGGGCTGAGGGCTCCACACACG TCCAGGGA

Using these primers, a new set of PCR reaction mixtures were prepared using standard conditions, and also with added DMSO and magnesium chloride, as described above. Once

again, no relevant amplicons could be observed from these reactions by agarose gel electrophoresis.

Given the high T_m for these primers (above 75 ºC), two different PCR techniques were also tested. In the first approach, a two-temperature cycle was tested, where denaturation was immediately followed by the extension step, and again no results were obtained.

In a second test, a touchdown PCR (TD-PCR) approach was employed, where the annealing temperature is varied along the run – in the first cycle T_{Ann} was set at 78 ºC, and then it was lowered by 0.5 ºC every two steps until 70 ºC. This would allow primer annealing to occur at high temperatures in the initial cycles, although with the possibility for non-specific annealing, and in later cycles primers would anneal at the correct place to pre-formed products, allowing obtaining the desired product. However, no bands at the expected band size were observed.

4.3. CYTOCHROME P450 OXIDOREDUCTASE EXPRESSION

Given the difficulties in amplifying the POR coding sequence from the source plasmid, it was decided to outsource the construction of a pET28a(+) plasmid containing the POR coding sequence.

A pET28a(+)-TEV-POR vector was designed and ordered. This is a pET28a(+) derivative in which the thrombin site is replaced by a tobacco etch virus (TEV) protease site, that has a consensus cut site ENLYFQS, cutting between the Gln (Q) and Ser (S) residues. The inserted sequence contains 12 upstream (5'-CATATGATCAAC…) and 6 downstream (…CTCGAG-3') added nucleotides, encoding the *Nde*I and *Xho*I cut sites, respectively. This plasmid codes for a POR construct with an N-terminal His_6 tag followed by a TEV site just upstream the canonical POR sequence, and also a C-terminal His_6 tag, generating a 77.2 kDa molecular mass expression product. The map for this vector is present in Annex 9

The plasmid was transformed into expression-competent *E. coli* C43(DE3) cells via the calcium chloride method. Upon establishing this transformed strain, cells were inoculated from an overnight pre-inoculum into LBK30 media and grown for 6 hours at 37 ºC at 220 rpm. At this point, protein expression was induced by adding 1 mM IPTG (final concentration) to the growth and incubating for further 4 hours at 37 ºC with agitation.

Following incubation, cells were harvested by centrifugation (5 000 x *g* for 30 min at 4 ºC) and the pellets were frozen overnight to promote cell lysis. Pelleted cells were resuspended in NZY Bacterial Cell Lysis at 5 mL per gram of cell paste, supplemented with 2 µL of lysozyme (from a 50 mg/mL stock solution) and 2 µL of DNAseI (from a 2 mg/mL stock solution) per mL of buffer. The cell suspension was incubated at RT for 1 hand then centrifuged (15 000 x *g* for 15 minutes at 4 ºC); the supernatant contains the soluble protein fraction. The pellet of the previous centrifugation contains cell debris and inclusion bodies, that often appear when over-expressing recombinant proteins in bacteria.

To isolate the protein from the inclusion bodies, the pellets were ressuspended in 10 mM Tris buffer (pH 7.4, with 0.15 M NaCl) at 10 mL of buffer per L of the original culture and sonicated for 5 min. Upon vortexing, samples were centrifuged at 20 000 x g for 15 min at 4 ºC and the supernatant was collected. This washing procedure was repeated, and the supernatant was once again collected. The resulting pellet was resuspended in 1.5 mL of 10 mM Tris buffer (pH 8) supplemented with 0.1 M NaH_2PO4 and 8 M urea; the suspensions were incubated for 2 h at RT at 500 rpm to solubilize inclusion bodies and centrifuged for 10 min at 10 ºC at 20 000 x g. Both the supernatant and the pellet were collected.

The various fractions collected during this process were analysed by electrophoresis on 10 % SDS polyacrylamide gels, shown in Figure 29

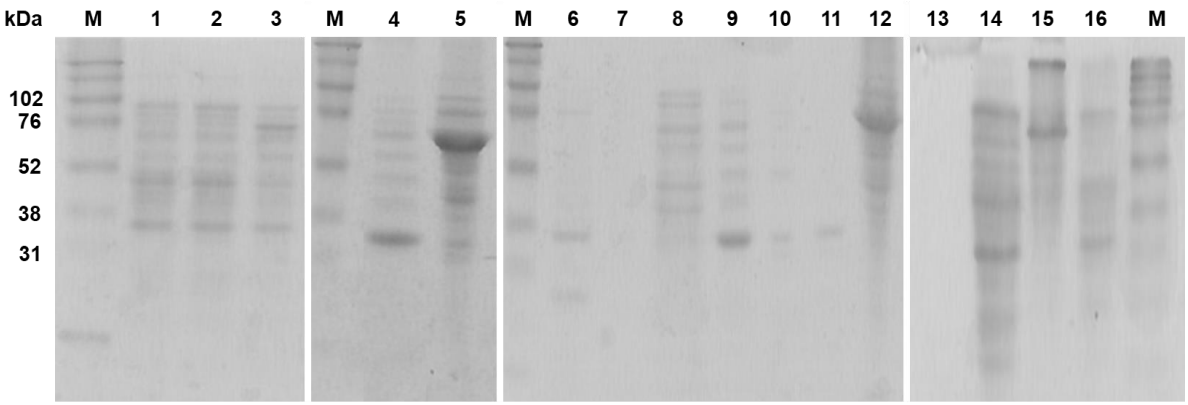


Figure 29: SDS-PAGE analysis (10 % gels) of IPTG-induced POR over-expression by *E. coli* C43(DE3) cells transformed with the pET28a(+)-TEV-POR plasmid. Lanes labelled M correspond to the molecular weight marker. 1 – whole cell sample was taken before IPTG induction; 2 – whole cell sample from control growth (not induced); 3 – whole cell sample taken 6 h after IPTG induction. Lanes 4 and 5 correspond to the supernatant and pellet, respectively, from the cell lysis from the induced growth. Lanes 6, 7 and 8 correspond to the supernatant of successive urea lysis steps performed on the pellet from the cell lysis step. Lane 9 is the same sample as lane 4. Lane 10 corresponds to the supernatant gathered after de centrifugation of the sample used in lane 5. Lanes 11 and 12 correspond to the final supernatant and pellet from the urea lysis steps, respectively. Lanes 13 and 15 correspond, respectively, to the supernatant and the pellet of a Tween-20 step performed on the pellet corresponding to lane 12. Lanes 14 and 16 correspond to the supernatant and pellet of the Tween-20 step performed on the pellet of the control growth.

Comparing lanes 1 to 3, it is clear that a new band at *ca.* 77 kDa appears on the IPTG induced culture, which corresponds to the expected molecular mass, and that is not present on the non-induced control growth nor the test culture before induction. Lane 4, which corresponds to the supernatant of cell lysis, does not show that band, which is instead located on the cell lysis pellet, suggesting that it is not a soluble protein. Lanes 6 to 8 correspond to successive urea lysis steps of possible inclusion bodies, but no band at 77 kDa was observed. Following this, the pellet was once again treated with DNAseI and lysozyme, and in lane 12 a clear band is observed, suggesting that the protein is still located in the overall pellet.

Being a membrane protein, it may be still attached to membrane debris, and so this pellet was incubated with 2.5 %(v/v) Tween-20 in 10 mL 10 mM Tris-HCl pH 7.4 supplemented with 0.15 M NaCl. This Tween-20 approach was also applied to the cell lysis pellet of the non-induced culture. Lanes 13 and 15 contain the supernatant and the pellet following the Tween-20 treatment of the pellet from the induced culture, and lanes 14 and 16 contain the supernatant and the pellet following the Tween-20 treatment of the cell lysis pellet from the non-induced culture. A band at 77 kDa is still observable in the pellet of the Tween-20 treated fraction from the induced culture, meaning that the protein remains membrane-bound or is then present in a non-renaturable conformation following all this treatment.

5. UDP-GLUCURONOSYLTRANSFERASE 2B7 CLONING AND EXPRESSION

5.1. UDP-GLUCURONOSYLTRANSFERASE 2B7 TRADITIONAL CLONING

Much like the CYP3A4 cloning, the UGT2B7 was also cloned using the PCR amplification technique. This gene was amplified (introducing the REs recognition sequences in the process, map in Annex 10) from the plasmid pDONR221-UGT2B7 (HsCD00045674, Harvard PlasmID Database) using the primers depicted in

Table 37. The insert sequence of the pDONR221-UGT2B7 plasmid corresponds entirely to the coding sequence for the human gene in its canonical form.

The translated sequence was checked for the presence of signal peptides using the SignalP5.0 server (Almagro Armenteros, Tsirigos et al. 2019)**.** In eukarya, the initial 23 amino acids are likely to be a signal peptide (Figure 30), and in Gram-negative bacteria, the same region is also likely to be a localization peptide (Figure 31). Moreover, as this protein can be found to be membrane-associated, an analysis of membrane-interacting domains was conducted using the TMHMM server (Sonnhammer, von Heijne et al. 1998). A likely transmembrane (TM) helix was found, comprising residues 495 to 517, with the C-terminal 518-529 oriented to the outside of the organelle (Figure 32). This agrees with the UniProt (UniProt 2020) database info (ID P16662), which lists a TM helix formed by residues 493 to 509. A zero-cutter analysis using the NEBCutter v2.0 server (Vincze, Posfai et al. 2003) identified *Nde*I, *Nhe*I, *Hind*III and *Xho*I enzymes as zero cutters for the coding sequence. *Nde*I and *Nhe*I sites are located on the pET28a(+) MCS site on the upstream region and *Hind*III and *Xho*I sites are located downstream in the MCS region; ultimately, *Nde*I and *Xho*I were selected for further cloning work.

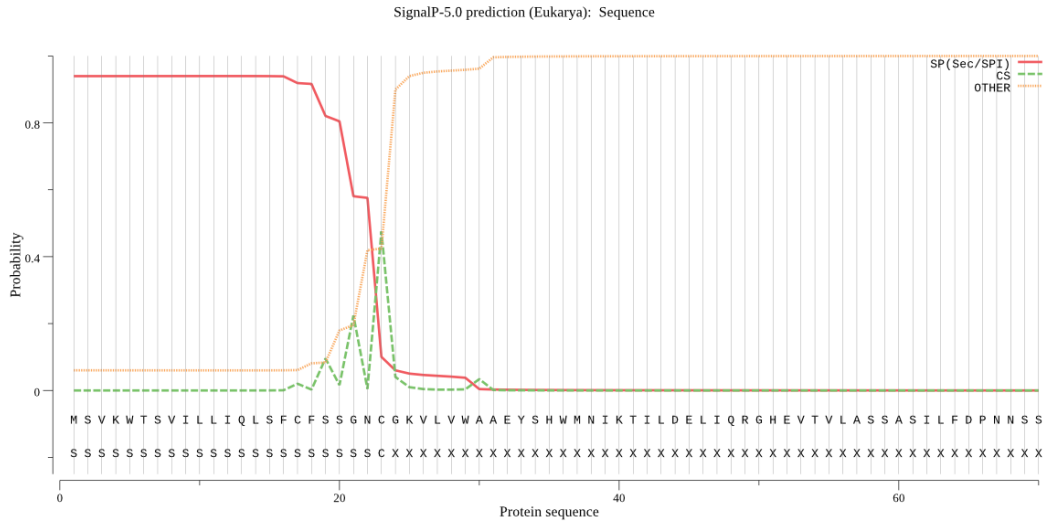


Figure 30: SignalP-5.0 prediction of signal peptides for Eukarya present in the UGT2B7 translated sequence. It is predicted that the first 23 amino acids have a high probability of being part of a signal peptide.

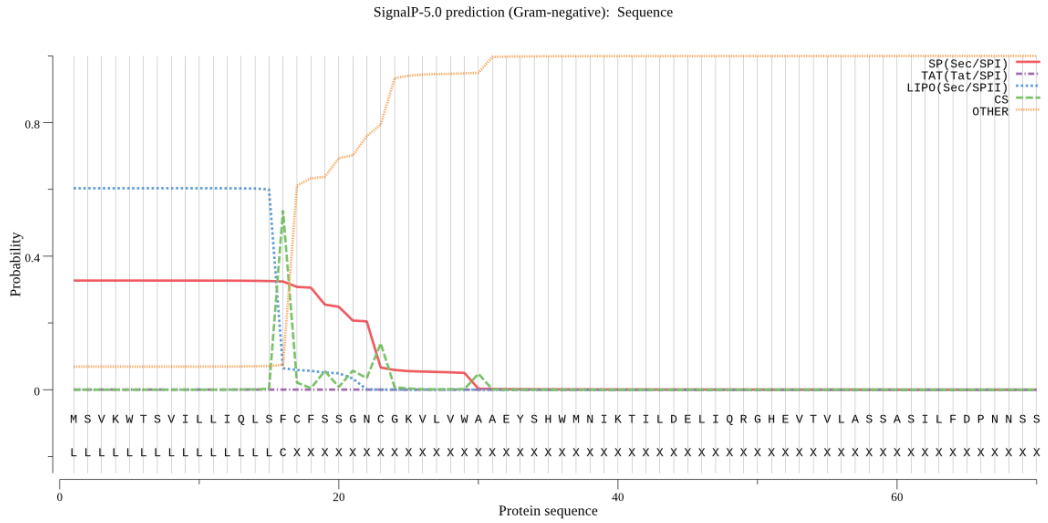


Figure 31: SignalP-5.0 prediction of signal peptides when expressing the UGT2B7 sequence in gram-negative bacteria. It is predicted that the first 16 amino acids are part of a localization peptide.

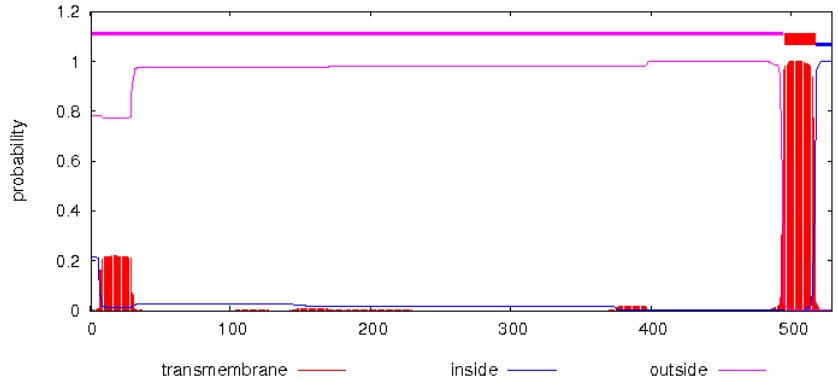


Figure 32: TM helix prediction via TMHMM Server v. 2.0 software. The amino acids from 495 to 517 encode a TM hélix with the C-terminal 518-529 oriented to the outside of the organelle.

With this information, various primer pairs were designed, as detailed in

Table 37, allowing for obtaining different amplification products. In version 1, the expression product will have an N truncated terminal (without the signal peptide), and the C-terminal will be intact, containing the TM region and the original stop codon. In version 2, the N-terminal region is intact, and a C-terminal His_6 tag is added, yielding a product likely to lose the N-terminal peptide (and the His_6 tag) and with the C-terminal purification tag. In version 3 the signal peptide will not be present in the final product, which carried His tags at both the N and C terminals. Version 4 is similar to version 3 but without the C terminal His_6 tag. Finally, version 5 leads to a product with an N terminal His_6 tag but no signal peptide, and a C-terminal region with an His6 tag and no TM helix. These alternative primer pairs were designed to account for the possible hydrolysis of the signal peptide in the final product, and also to add the possibility of having different C-terminal regions, as the presence of a terminal TM helix may introduce complications in the expression and purification steps. The design strategy is shown in Figure 33.

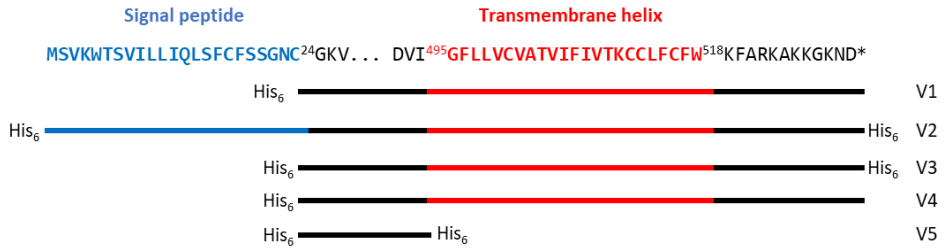


Figure 33: Primers designed for the cloning process of the UGT2B7. V1: No signal peptide (in the N-terminal region) and C-terminal intact. V2: N-terminal intact, C-terminal His_6 tag is added yielding a product likely to lose the N-terminal peptide (and the His_6 tag). V3: Removal of the signal peptide and presence of His_6 tag s at both C and N terminals. V4: equal to V3 but without C terminal His_6 tag. V5: No signal peptide or TM helix but the presence of His_6 tag at both terminals.

Table 37: Primer pairs designed for the UGT2B7 gene amplification

	Name	Tm (ºC)	Length (bp)	Sequence (5' → 3')
V1	Fw_Nde_v1	65.4	29.0	TCTGGGCATATGGGAAAGGTG CTGGTGTG
	Rev_Xho_v1	62.8	32.0	GACCTCGAGTTACCCAATCAC ATCCAAAGAGT
V2	Fw_Nde_v2	59.3	30.0	TTGCACTCATATGTCTGTGAA ATGGACTTC
	Rev_Xho_v2	58.0	33.0	GATATAACTCGAGATTTTTTC CCTTCTTTGCTT
V3	Fw_ Nde_v1	65.4	29.0	TCTGGGCATATGGGAAAGGTG CTGGTGTG
	Rev_Xho_v2	58.0	33.0	GATATAACTCGAGATTTTTT CCCTTCTTTGCTT
V4	Fw_Nde_v1	65.4	29.0	TCTGGGCATATGGGAAAGGTG CTGGTGTG
	Rev_Xho_v4	61.8	31.0	GCATCTCGAGCTTTCTTGCTG GAATAAACTG
V5	Fw_Nde_v2	59.3	30.0	TTGCACTCATATGTCTGTGAA ATGGACTTC
	Rev_Xho_v5	63.6	28.0	ACACACTCGAGCAGGAACCCA ATCACAT

Cloning work started with the primer pair for version 1, to obtain a recombinant protein without the N signal peptide and with a single N-terminal His tag. For the PCR amplification, the Supreme NZYProof DNA was used, with a final reaction volume of 50 µL, and 50 ng of template DNA was added. The rest of the component volumes were added according to the commercial protocol of the enzyme. As for the thermocycler program used (35 cycles), its step durations and temperatures are depicted in Table 38. To ensure proper amplification, the program was run with a T_{ann} gradient from 55 ºC to 64 ºC.

Table 38: Thermocycler PCR program. A temperature gradient was made with intervals of 1 ºC

Cycles	Step	T (ºC)	t (s)
	Pre-heat	95	∞
1	Initial denaturation	96	240
35	Denaturation	96	30
	Annealing	[59,64]; Δ = 1	30
	Extension	72	60 †
1	Final Extension	72	600
	Hold	4	∞

The reactions were then analysed via agarose gel electrophoresis (agarose gel of 1.5% (w/v)) in the same conditions as previous electrophoresis analysis. As observable from Figure 34, all T_{ann} (apart from T_{ann} = 61 ºC) had successful amplification, with the T_{ann} = 62 ºC showing the band most approximate to 1.609 kbp (estimated correct PCR product size). This reaction was then digested, alongside some pET-28a(+) plasmid, with *Nde*I and *Xho*I.

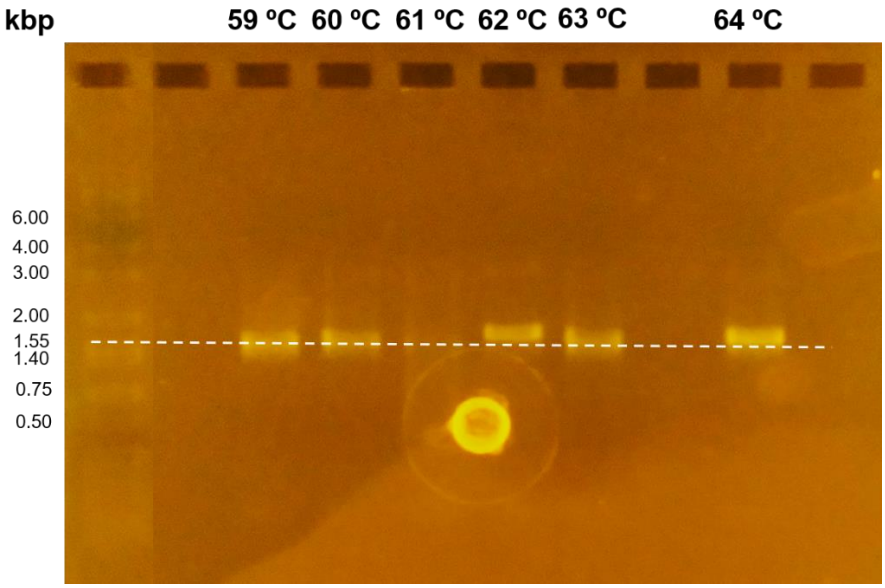


Figure 34: Agarose gel (1.5 %) electrophoresis of the PCR reactions for UGT2B7 insert amplification. The samples depicted are the different PCR reactions (different T_{ann}) with T_{ann}=62 ºC showing the most accurate result (closest to the expected 1.609 kbp).

The enzymatic digestions with the *Nde*I and *Xho*I (MB101 and 107, NZYTech) were performed following the commercial protocol provided, in which 16 µL of pET-28a(+) (*ca.* 716.608 ng) and 30 µL of UGT2B7 PCR amplification product (*ca.* 2917 ng) were used into a final reactional

volume of 20 µL, in the pET-28a(+) digestion, and 60 µL, in the UGT2B7 insert digestion. These digestions were then followed up with a quantification step, using the spectroscopy method previously mentioned, ending up with the characteristics presented in

Table 39.

Table 39: Digested UGT2B7 insert and pET-28a(+). Purity ratios and concentration for each fraction obtained via spectrophotometric analysis

DNA	Volume (µL)	260/280	Concentration (ng/µL)
UGT2B7	60.00	1.60	66.60
pET-28a(+)	20.00	1.60	52.70

With the now digested vector and insert, an enzymatic ligation step followed using the same T4 Ligase used in previous enzymatic ligation reactions and its commercial protocol was followed rigorously (Annex 17**)** using a ratio of *ca.* 1:10 (vector:insert), therefore adding 0.95 µL of the digested pET-28a(+) (*ca.* 50 ng) and 2.19 µLof the digested insert (*ca.* 145.455 ng). the reaction occurred over a period of 16 hours at a temperature of 16 ºC.

The pET-UGT2B7 was transformed into competent DH5α *E. coli* cells in the same manner described in section 2.1. CYP2C8 traditional cloning. These transformed cells were then plated into LBK30 agar plates and grown overnight at 37 ºC. After this growing period, the colonies were isolated into a new LBK30 plate and left to grow once more in the same conditions. In parallel, the loops used to isolate said colonies were also used to inoculate 2 mL of LBK30 media that underwent growth in an orbital incubator at 37 ªC, 250 rpm, and overnight. These liquid growths were used to make 25 % (v/v) glycerol stocks for future use.

Following the incubation period, a colony PCR was made of all the colonies that had grown both in the solid and liquid media which happen to be all of the 64 isolated colonies. The conditions were similar to the colony PCR in 2.1. CYP2C8 traditional cloning, with the reaction components represented in Table 40 and the PCR program used was the same described in Table 7 but using T_{ann} = 55 ºC and 60 seconds of extension time. From this PCR it was possible to affirm that out of the 64 isolated colonies, 12 of them appeared to present only the intended band (*ca.* 1.8 kbp) and another 25 presented the intended band plus a nonspecific one around 0.5 kbp (Figure 35). Amongst all the colonies, the #40 and #53 clones were selected for further transformation into *E. coli* C43(DE3) cells.

Table 40: Colony PCR master mix components. † The total volume of the master mix does not take into account the 10 µL of colony sample

Components	**Initial concentration**	**Volume (µL) for**		**Final concentration/amount**
		1 reaction	**64+2 reactions**	
MB H_2O		3.64	240.24	-
Reaction buffer	5.00 x	4.00	264.00	1.00 x
dNTPs	25.00 mM	0.16	10.56	0.20 mM
PFw	10.00 µM	1.00	66.00	0.50 µM
PRv	10.00 µM	1.00	66.00	0.50 µM
Colony sample	-	(10.00)	(10.00)	-
Supreme NZYProof DNA Polymerase	2.50 U/µL	0.20	13.20	0.025 U/µL
TOTAL	-	10.00 (20.00)	660.00 †	-

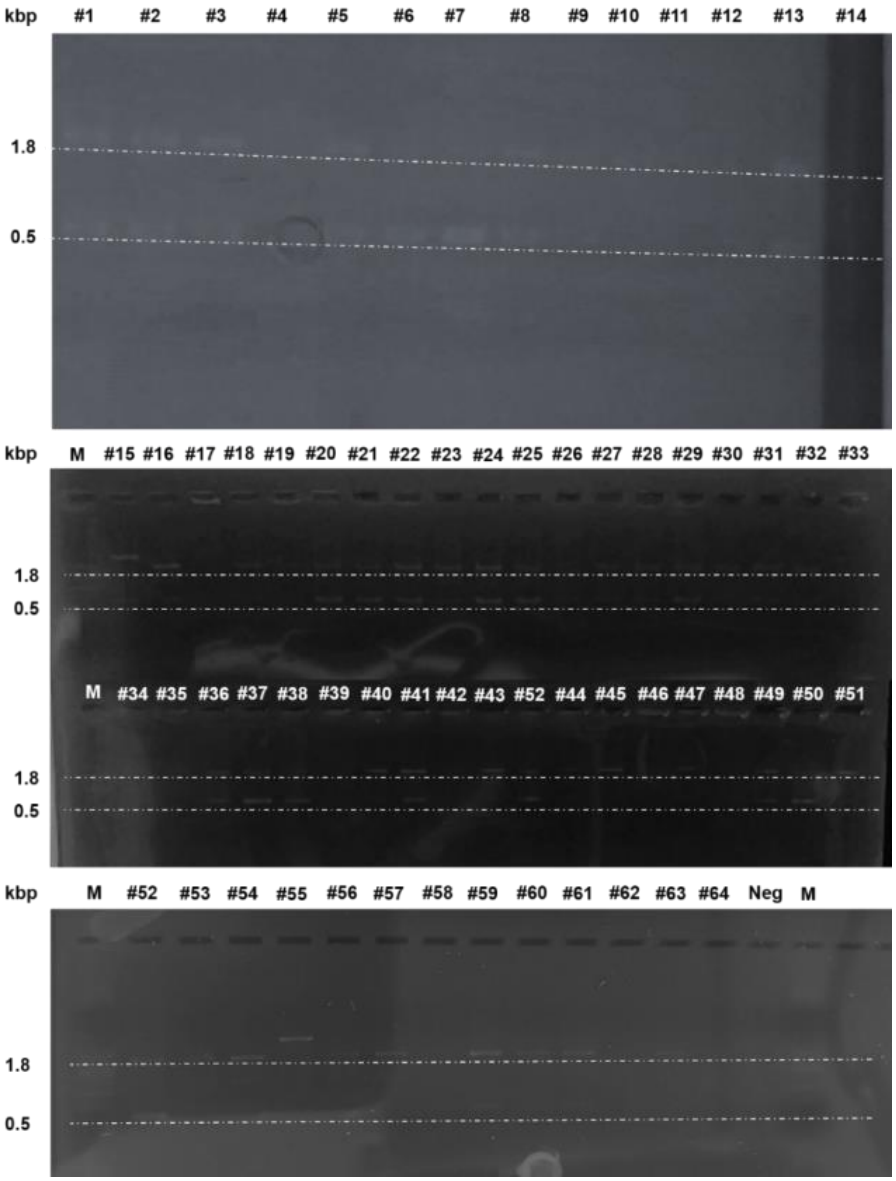


Figure 35: Agarose gel (1.5 %) electrophoresis of the colony PCR reactions from the isolated *E.* coli DH5α colonies from pET-UGT2B7 transformation. pET-28 a (+) was used as a negative control. In all gels, the DNA ladder used was the NZYDNA Ladder III (MB044, NZYTech).

To transform the plasmids of these clones into *E. coli* C43(DE3) cells, 30 mL of LBK30 were inoculated with each of the clones and grown overnight at 37 ºC and 250 rpm. The plasmids were then extracted using the same method described in the above chapters and 25% (v/v) glycerol stocks were made for future use. After quantification via spectrophotometry, the #40 plasmid solution had a concentration of 128.80 ng/µL ± 1.54 ng/µL and the #53 plasmid solution had a concentration of 112.01 ng/µL ± 0.28 ng/µL. Both of the plasmid solutions presented a 260/280 purity ratio of 1.8 and a final volume of 500 µL. C43(DE3) cells were then transformed following the same protocol used for the transformation into DH5α cells.

Once more a colony PCR followed suit to verify if the transformation had been successful using the program from the DH5α colony transformation PCR and the master mix reagents in Table 41. In this reaction, two colonies from each plasmid were tested and two reactions with dH_2O substituting the template DNA volumes were made to serve as a negative control totalling six PCR reactions.

Table 41: pET-UGT2B7 transformed into *E. coli* C43 colony PCR

	Volume for 1 reaction (µL)	**Volume for 6+1 reactions (µL)**
MB H_2O	5.10	35.70
Reaction buffer [5x]	6.00	42.00
dNTPs [25 mM]	0.60	4.20
Fwd primer [10 µM]	1.50	10.50
Rev primer [10 µM]	1.50	10.50
Colony sample	(15.00)	(15.00)
Supreme NZYProof DNA Polymerase [2.5 U/µL]	0.30	2.10
TOTAL	15.00 (30.00)	210.00

Observing the PCR results (Figure 36), two bands appeared in each of the four colonies tested, all with the same sizes (one band with *ca.* 1.8 kbp and another nonspecific band with 1.4 kbp), while no bands appeared in the negative controls. Both the #40 and #53 clones are then ready to start the UGT2B7 protein expression. The vector map for the pET-UGT2B7 is given in Annex 11.

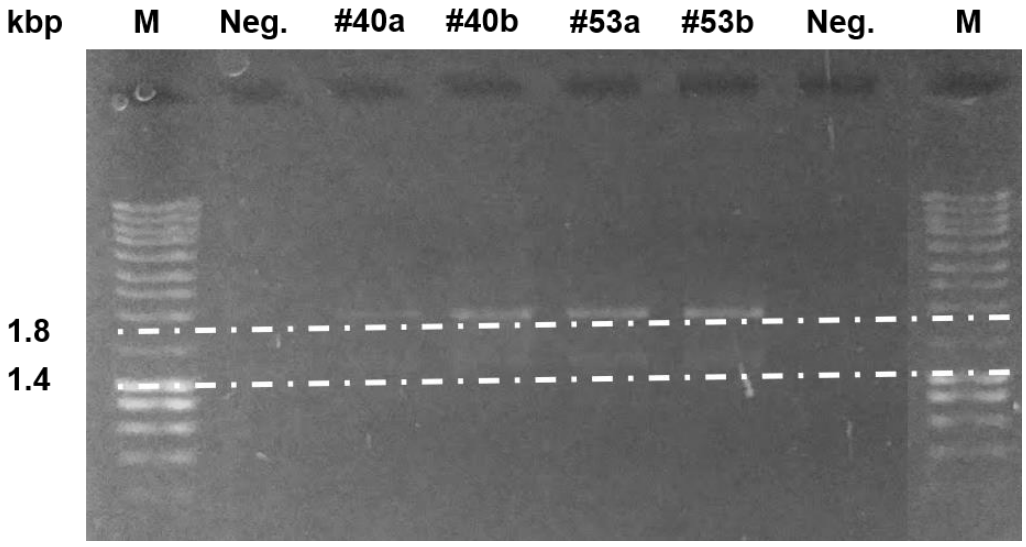


Figure 36: Agarose gel (1.5 %) electrophoresis of the colony PCR assay of the #43 and #50 plasmids from pET-UGT2B7-transformed *E.coli* C43(DE3) cells. "Neg." corresponds to the PCR reaction where the volume of plasmid was substituted with H_2O serving as a negative control.

5.2. UDP-GLUCURONOSYLTRANSFERASE 2B7 EXPRESSION

Clone 50 was used for an initial expression assay. Cells from an overnight pre-inoculum were used to inoculate an LBKAN30 media that was grown for 6 h at 37 ºC at 220 rpm. During this period, protein expression was induced by adding 0.5 mM IPTG (final concentration). Cells were harvested 4 h after centrifugation and resuspended in NZY Bacterial Cell Lysis at 5 mL per gram of cell paste, supplemented with 2 µL of lysozyme (from a 50 mg/mL stock solution) and 2 µL of DNAseI (from a 2 mg/mL stock solution) per mL of buffer. The cell suspension was incubated at RT for 1 hand then centrifuged (15 000 x *g* for 15 minutes at 4 ºC).

The crude supernatant was applied on a HisTrap FF Crude (GE LifeSciences, UK) column. The His_6 tag added during the cloning process generates a hexahistidine terminal peptide, which displays a high affinity towards nickel ions. HisTrap chromatography columns contain a nickel-loaded resin that allows binding of His_6-proteins, while most of the cellular products are not bound. Upon binding and washing, the retained proteins are easily eluted with an imidazole gradient, which outcompetes bound proteins.

Briefly, the column was washed with 5 column volumes of water and equilibrated with 5 volumes of binding buffer (20 mM sodium phosphate pH 7.4 with 0.5 M NaCl). The cell lysate was loaded on the column and it was washed with 6 column volumes of binding buffer. Protein elution was performed by increasing the imidazole concentration from 0 mM to 500 mM in step gradient and washing the column with 6 column volumes of elution buffer (20 mM sodium phosphate, 0.5 M NaCl, 500 mM imidazole) afterwards. Absorbance at 280 nm was recorded for the eluted fractions (5 mL each). The obtained chromatogram is shown in Figure 37: The purity of collected protein fractions was assessed by SDS-PAGE analysis, the results of which are shown in Figure 38.

A band at an elution volume of *ca*. 80 mL can be observed in the chromatogram, corresponding to a fraction eluted with 25 mM imidazole buffer. This corresponds to lane 8 in the acrylamide gel, where a band between 52 and 76 kDa can be observed, which may correspond to the UGT2B7 protein, at 64 kDa.

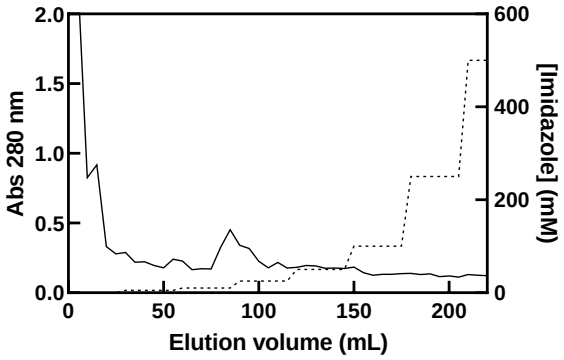


Figure 37: Chromatogram of UGT purification on a Ni-loaded HisTrap FF Crude (GE LifeSciences) column. The full line corresponds to the absorbance of each fraction recorded at 280 nm, and the dashed line represents the imidazole step gradient used for elution.

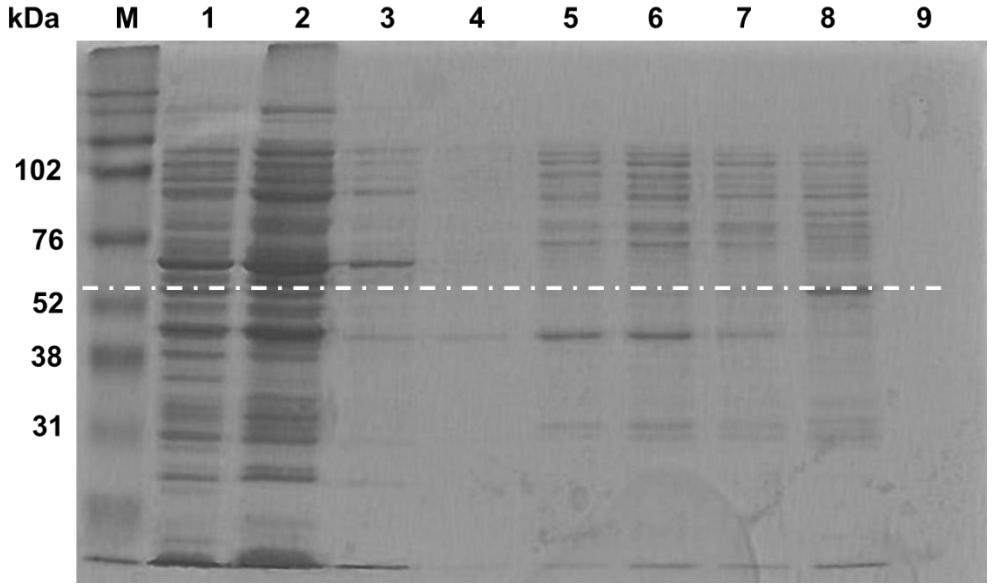


Figure 38: SDS-PAGE analysis (10 % resolving gel) of the elution process. Lane M corresponds to the molecular weight marker. A sample of the lysis pellet was added in lane 1 for comparison. Lane 2 corresponds to the injection fraction, lanes 3 and 4 to fractions eluted with the washing buffer, lanes 5 and 6 to fractions eluted with 10 mM imidazole buffer, lanes 7 and 8 to fractions eluted with 25 mM imidazole and lane 9 to a final 500 mM imidazole fraction.

6. CONCLUSION

In conclusion, although originally the primary objective of this dissertation was to obtain cloned *E. coli* cells expressing the desired cytochrome P450 and UDP-glucuronosyltransferase recombinant human proteins, a bigger focus was put on the cloning process instead. This change of focus happened due to the difficulty and complexity of the cloning process that was originally thought to take a smaller internship time that ended up taking leaving less time to focus on the expression of the proteins themselves.

Even though the traditional cloning techniques did not work, after using the LIC technique both the CYP2C8 and CYP3A4 proteins were successfully cloned by a ligase-independent method; pHTP9-CYP2C8 (#7 and #8) and pHTP9-CYP3A4 (#6 and #41) clones, corresponding to final GFP-CYP fusion proteins, will be used in future protein expressions to confirm and characterize the protein expressed.

Regarding POR, amplification of the coding insert by PCR could not be achieved. Four sets of primers were used, with varying concentrations of different additives; a two-step PCR approach and a touch-down PCR variant were also employed, but no amplification was obtained. This plasmid was then acquired from a supplier and was successfully transformed into expression-competent cells and protein expression via IPTG-induction afforded a definitive protein band that requires further purification work.

Last but not least, the UGT2B7 protein had a smooth and effective cloning process having obtained two fully working pET-UGT2B7 clones (#40 and #53). One of the clones was successfully expressed via IPTG induction and purified by a simple straightforward chromatography approach.

REFERENCES

Almagro Armenteros, J. J., K. D. Tsirigos, C. K. Sønderby, T. N. Petersen, O. Winther, S. Brunak, G. von Heijne and H. Nielsen (2019). "SignalP 5.0 improves signal peptide predictions using deep neural networks." Nat Biotechnol **37**(4): 420-423.

Aslanidis, C. and P. J. de Jong (1990). "Ligation-independent cloning of PCR products (LIC-PCR)." Nucleic Acids Research **18**(20): 6069-6074.

Bodine, D. M. "Genetic Engineering." Retrieved 03/09, 2020, from https://www.genome.gov/genetics-glossary/Genetic-Engineering.

BRCF. "How do I design my own primers?" Retrieved November, 2020, from https://brcf.medicine.umich.edu/cores/advanced-genomics/faqs/sanger-sequencing-faqs/how-do-i-design-my-own-primers/.

Brodie, B. B., J. Axelrod, J. R. Cooper, L. Gaudette, B. N. La Du, C. Mitoma and S. Udenfriend (1955). "Detoxication of drugs and other foreign compounds by liver microsomes." Science **121**(3147): 603-604.

Brown, T. A. (2010). Gene Cloning & DNA Analysis - An Introduction. UK, Wiley-Blackwell.

Chang, A., L. Jeske, S. Ulbrich, J. Hofmann, J. Koblitz, I. Schomburg, M. Neumann-Schaal, D. Jahn and D. Schomburg (2021). "BRENDA, the ELIXIR core data resource in 2021: new developments and updates." Nucleic acids research **49**(D1): D498-D508.

Chang, E., B. Ge, M. Lee, M. So and W.-b. Wang (2005). Investigation of the Ligation Efficiency of *Nde*I Digested Fragments.

Chester, N. and D. R. Marshak (1993). "Dimethyl sulfoxide-mediated primer Tm reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction." Anal Biochem **209**(2): 284-290.

Desta, Z. and D. A. Flockhart (2017). Chapter 18 - Pharmacogenetics of Drug Metabolism. Clinical and Translational Science (Second Edition). D. Robertson and G. H. Williams, Academic Press**:** 327-345.

Fujiwara, R., T. Yokoi and M. Nakajima (2016). "Structure and Protein-Protein Interactions of Human UDP-Glucuronosyltransferases." Front Pharmacol **7**: 388.

Gao, Y., D. Liu, H. Wang, J. Zhu and C. Chen (2010). "Functional characterization of five CYP2C8 variants and prediction of CYP2C8 genotype-dependent effects on in vitro and in vivo drug-drug interactions." Xenobiotica **40**(7): 467-475.

Gonzalez, F. J., B. J. Schmid, M. Umeno, O. W. McBride, J. P. Hardwick, U. A. Meyer, H. V. Gelboin and J. R. Idle (1988). "Human P450PCN1: sequence, chromosome localization, and

direct evidence through cDNA expression that P450PCN1 is nifedipine oxidase." DNA **7**(2): 79-86.

Gregory, P. A., R. H. Lewinsky, D. A. Gardner-Stephen and P. I. Mackenzie (2004). "Regulation of UDP glucuronosyltransferases in the gastrointestinal tract." Toxicol Appl Pharmacol **199**(3): 354-363.

Hichiya, H., T. Tanaka-Kagawa, A. Soyama, H. Jinno, S. Koyano, N. Katori, E. Matsushima, S. Uchiyama, H. Tokunaga, H. Kimura, N. Minami, M. Katoh, K. Sugai, Y. Goto, T. Tamura, N. Yamamoto, Y. Ohe, H. Kunitoh, H. Nokihara, T. Yoshida, H. Minami, N. Saijo, M. Ando, S. Ozawa, Y. Saito and J. Sawada (2005). "Functional characterization of five novel CYP2C8 variants, G171S, R186X, R186G, K247R, and K383N, found in a Japanese population." Drug Metab Dispos **33**(5): 630-636.

Isin, E. M. and F. P. Guengerich (2007). "Complex reactions catalyzed by cytochrome P450 enzymes." Biochim Biophys Acta **1770**(3): 314-329.

King, C. D., G. R. Rios, J. A. Assouline and T. R. Tephly (1999). "Expression of UDP-glucuronosyltransferases (UGTs) 2B7 and 1A6 in the human brain and identification of 5-hydroxytryptamine as a substrate." Arch Biochem Biophys **365**(1): 156-162.

Longley, M. J., S. E. Bennett and D. W. Mosbaugh (1990). "Characterization of the 5′ to 3′ exonuclease associated with Thermus aquaticus DNA polymerase." Nucleic Acids Research **18**(24): 7317-7322.

Lu, A. Y., K. W. Junk and M. J. Coon (1969). "Resolution of the cytochrome P-450-containing omega-hydroxylation system of liver microsomes into three components." J Biol Chem **244**(13): 3714-3721.

Manikandan, P. and S. Nagini (2018). "Cytochrome P450 Structure, Function and Clinical Significance: A Review." Curr Drug Targets **19**(1): 38-54.

McDonald, A. (2019). The Enzyme List: Class 2 - Transferases, International Union of Biochemistry and Molecular Biology (NC-IUBMB)**:** 423-424.

McDonnell, A. M. and C. H. Dang (2013). "Basic review of the cytochrome p450 system." J Adv Pract Oncol **4**(4): 263-268.

Miley, M. J., A. K. Zielinska, J. E. Keenan, S. M. Bratton, A. Radominska-Pandya and M. R. Redinbo (2007). "Crystal Structure of the Cofactor-Binding Domain of the Human Phase II Drug-Metabolism Enzyme UDP-Glucuronosyltransferase 2B7." Journal of Molecular Biology **369**(2): 498-511.

Miura, M., K. Ito, M. Hayashi, M. Nakajima, T. Tanaka and S.-i. Ogura (2015). "The Effect of 5-Aminolevulinic Acid on Cytochrome P450-Mediated Prodrug Activation." PloS one **10**(7): e0131793-e0131793.

Montellano, P. R. O. (2015). Cytochrome P450: Structure, mechanism, and biochemistry, fourth edition.

NEB. "PCR Cloning Method." Retrieved November, 2020, from https://international.neb.com/applications/cloning-and-synthetic-biology/pcr-cloning.

Nicholl, D. S. T. (2008). An Introduction to Genetic Engineering. New York, USA, Cambridge University Press.

NZYTech (2019). NZYEasy Cloning & Expression System. NZYTech.

Pandey, A. V. and C. E. Flück (2013). "NADPH P450 oxidoreductase: structure, function, and pathology of diseases." Pharmacol Ther **138**(2): 229-254.

Sambrook, J., D. W. Russell, C. A. Irwin, C. S. H. Laboratory, R. E. C. M. Fund and K. A. Janssen (2001). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press.

Sandee, D., K. Morrissey, V. Agrawal, H. K. Tam, M. A. Kramer, T. S. Tracy, K. M. Giacomini and W. L. Miller (2010). "Effects of genetic variants of human P450 oxidoreductase on catalysis by CYP2D6 in vitro." Pharmacogenet Genomics **20**(11): 677-686.

Schoch, G. A., J. K. Yano, M. R. Wester, K. J. Griffin, C. D. Stout and E. F. Johnson (2004). "Structure of human microsomal cytochrome P450 2C8. Evidence for a peripheral fatty acid binding site." J Biol Chem **279**(10): 9497-9503.

Scientific, T. "PCR Cycling Parameters–Six Key Considerations for Success." Retrieved November, 2020, from https://www.thermofisher.com/pt/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/pcr-education/pcr-reagents-enzymes/pcr-cycling-considerations.html.

Sonnhammer, E. L., G. von Heijne and A. Krogh (1998). "A hidden Markov model for predicting transmembrane helices in protein sequences." Proc Int Conf Intell Syst Mol Biol **6**: 175-182.

Šrejber, M., V. Navrátilová, M. Paloncýová, V. Bazgier, K. Berka, P. Anzenbacher and M. Otyepka (2018). "Membrane-attached mammalian cytochromes P450: An overview of the membrane's effects on structure, drug binding, and interactions with redox partners." Journal of Inorganic Biochemistry **183**: 117-136.

Testa, B. and B. Clement (2015). Chapter 24 - Biotransformation Reactions and their Enzymes. The Practice of Medicinal Chemistry (Fourth Edition). C. G. Wermuth, D. Aldous, P. Raboisson and D. Rognan. San Diego, Academic Press**:** 561-584.

UniProt (2020). "UniProt: the universal protein knowledgebase in 2021." Nucleic Acids Research **49**(D1): D480-D489.

Vincze, T., J. Posfai and R. J. Roberts (2003). "NEBcutter: A program to cleave DNA with restriction enzymes." Nucleic acids research **31**(13): 3688-3691.

Wang, M., D. L. Roberts, R. Paschke, T. M. Shea, B. S. Masters and J. J. Kim (1997). "Three-dimensional structure of NADPH-cytochrome P450 reductase: prototype for FMN- and FAD-containing enzymes." Proceedings of the National Academy of Sciences of the United States of America **94**(16): 8411-8416.

Watson, J. D., J. W. J. D. W. A. A. C. Richard M. Myers, W. J. D, U. R. M. Myers, R. M. Myers, A. A. Caudy and J. A. Witkowski (2007). Recombinant DNA: Genes and Genomes: A Short Course, W. H. Freeman.

Williams, P. A., J. Cosme, D. M. Vinkovic, A. Ward, H. C. Angove, P. J. Day, C. Vonrhein, I. J. Tickle and H. Jhoti (2004). "Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone." Science **305**(5684): 683-686.

Wuest, D. M., S. Hou and K. H. Lee (2011). Comprehensive Biotechnology.

Xia, C., S. P. Panda, C. C. Marohnic, P. Martásek, B. S. Masters and J.-J. P. Kim (2011). "Structural basis for human NADPH-cytochrome P450 oxidoreductase deficiency." **108**(33): 13486-13491.

Zerbs, S., A. M. Frank and F. R. Collart (2009). Chapter 12 Bacterial Systems for Production of Heterologous Proteins. Methods in Enzymology. R. R. Burgess and M. P. Deutscher, Academic Press. **463:** 149-168.

ANNEXES – PLASMID MAPS

Annex 1 – pCW-LIC-CYP2C8 map

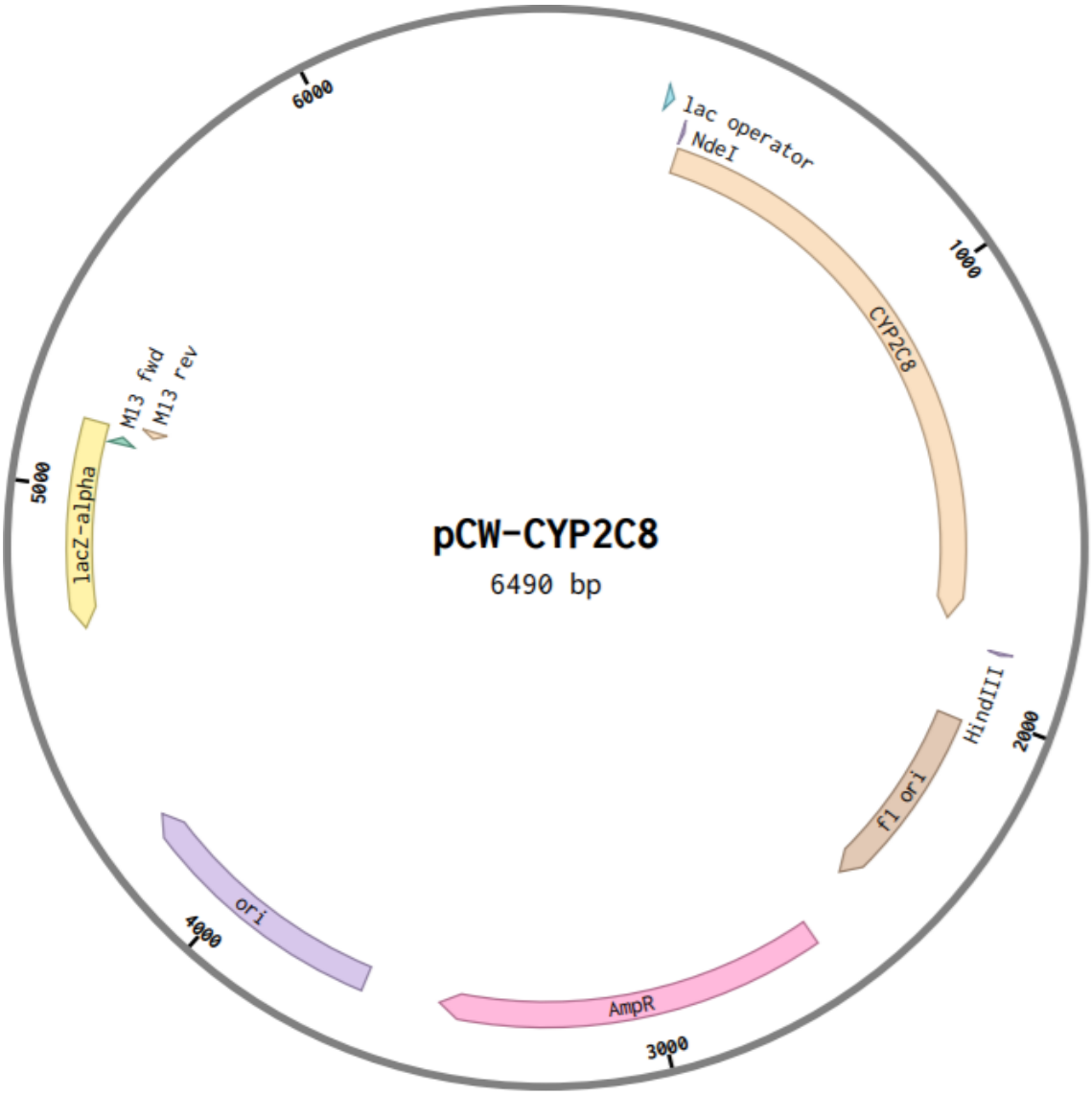


Figure A1 – Vector map for plasmid pCW-LIC-CYP2C8. Plasmid map source: Addgene.

Annex 2 – pET-28a(+) map

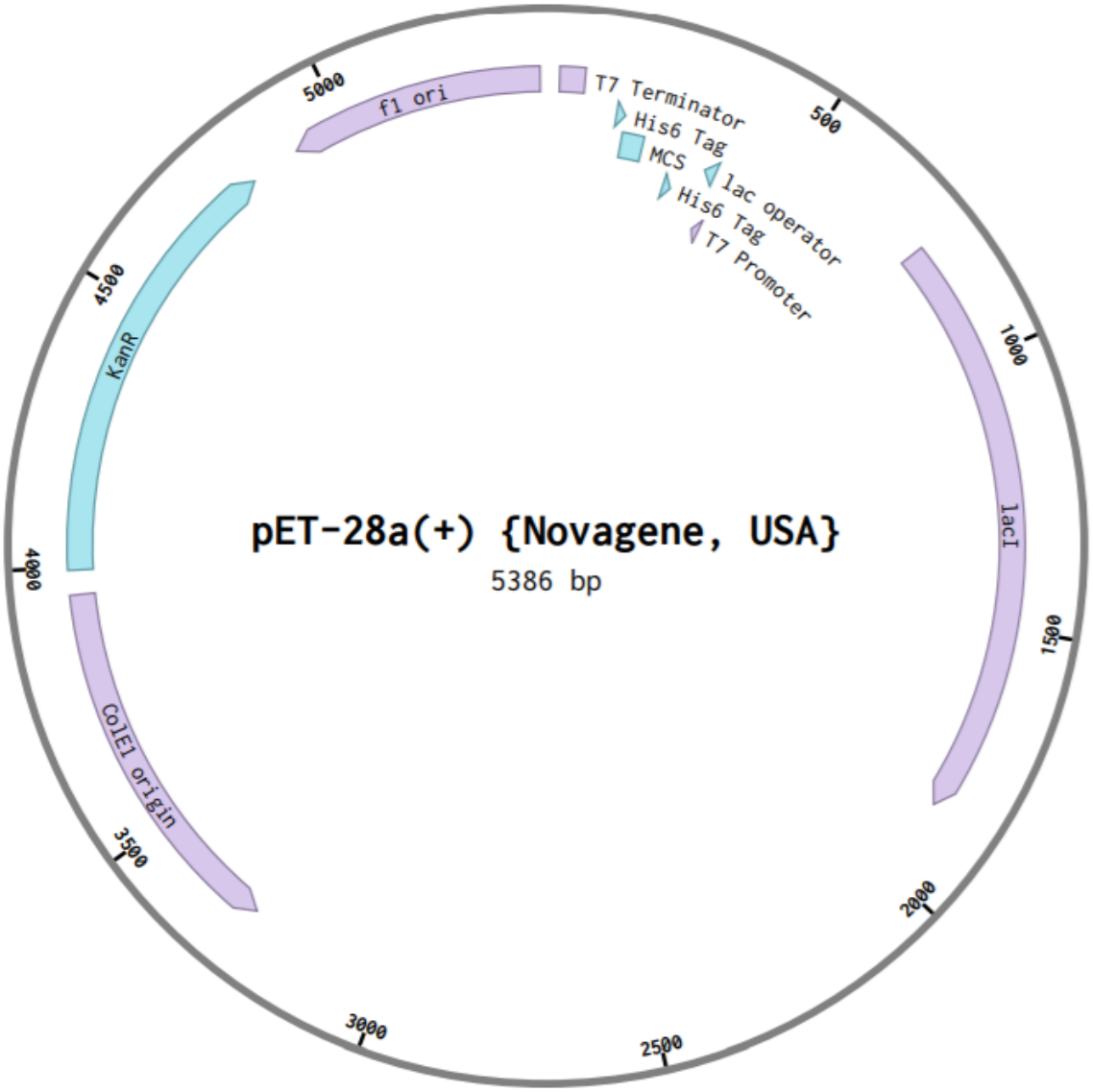


Figure A2 – Vector map for plasmid pET-28a(+). Plasmid map source: Novagen/MerckMillipore.

Annex 3 – pHTP1 map

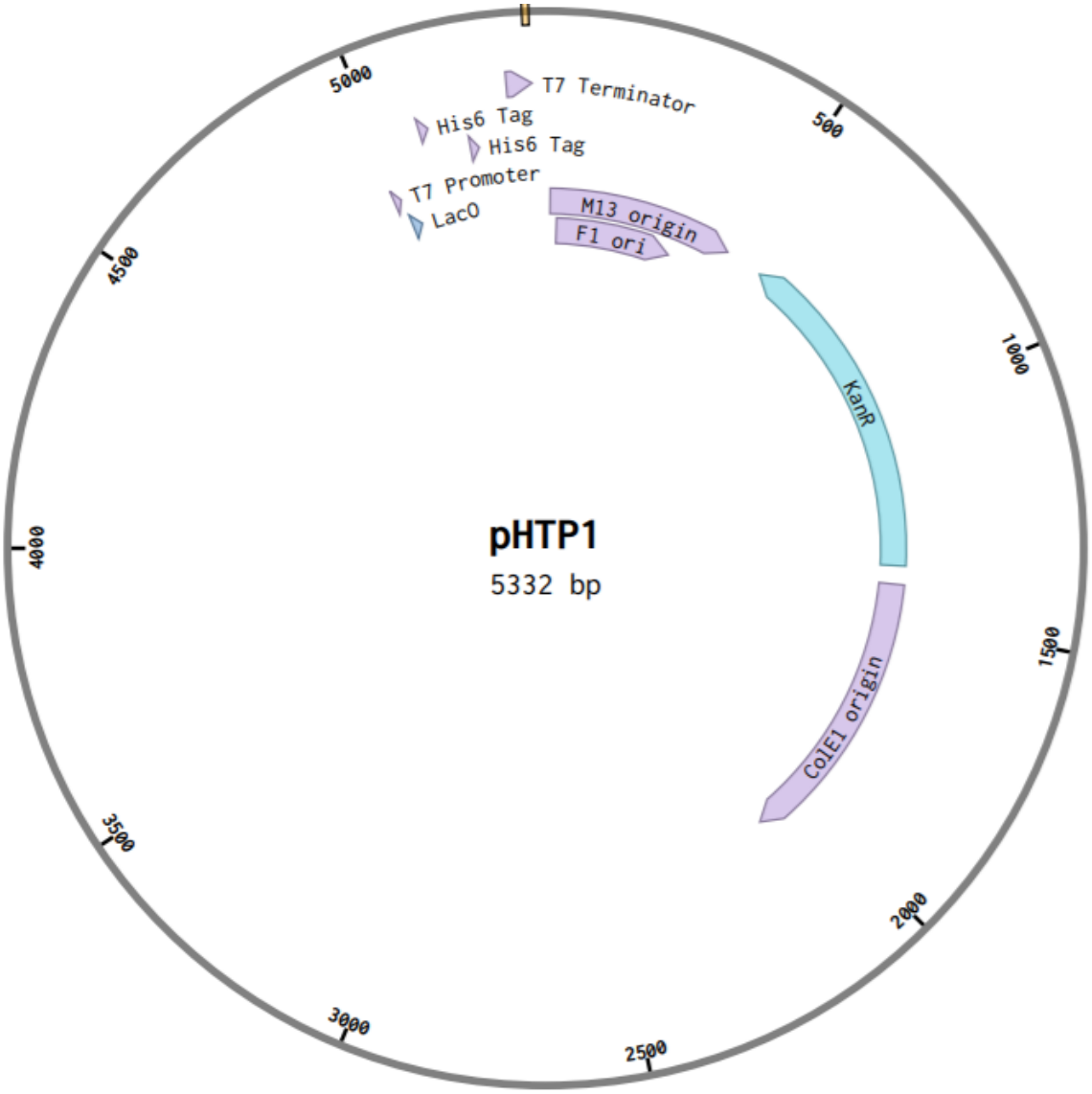


Figure A3 – Vector map for plasmid pHTP1. Plasmid map source: NZYTech

Annex 4– pHTP9 map

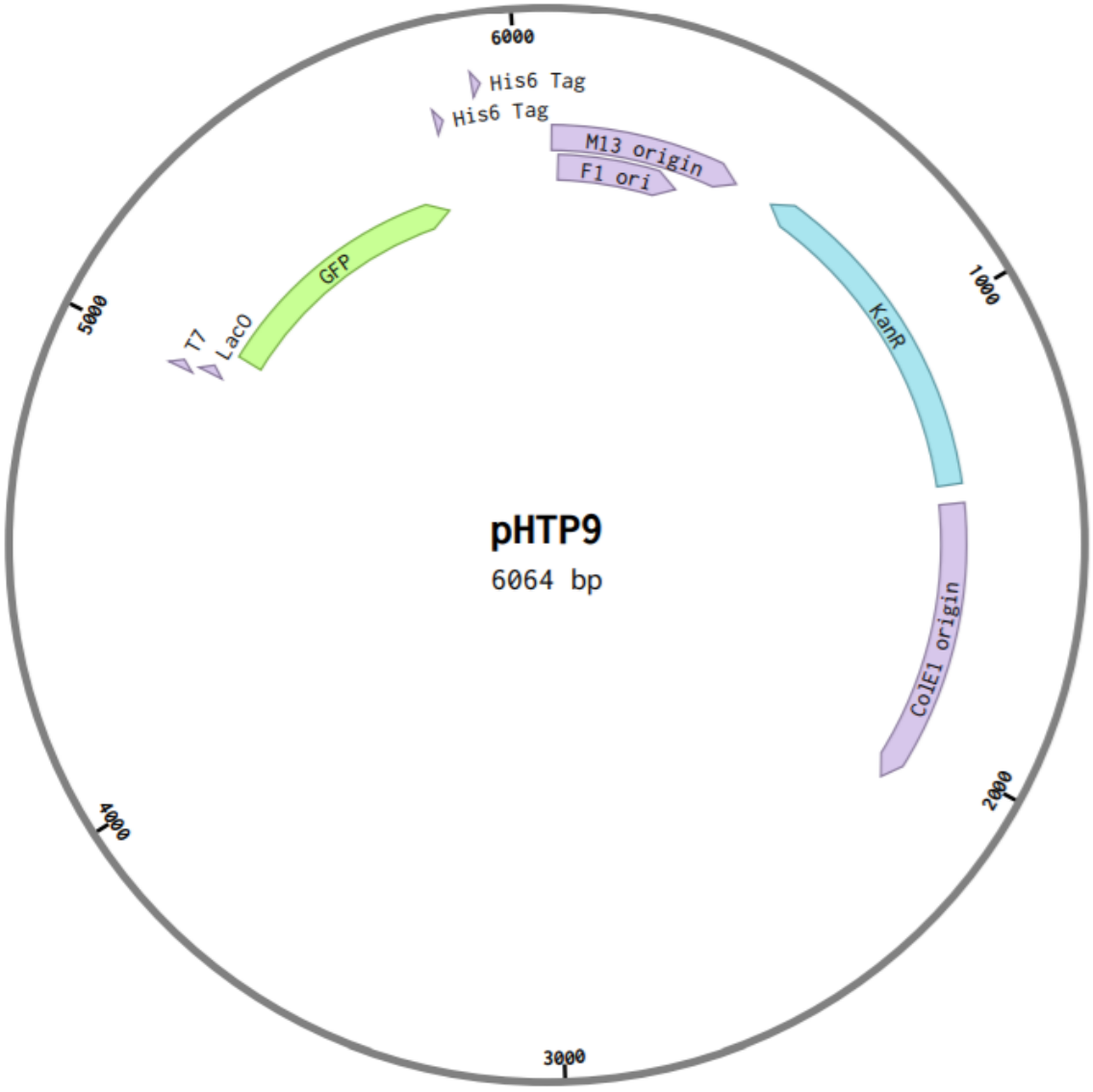


Figure A4 – Vector map for plasmid pHTP9. Plasmid map source: NZYTech

Annex 5 – pHTP9-CYP2C8 map

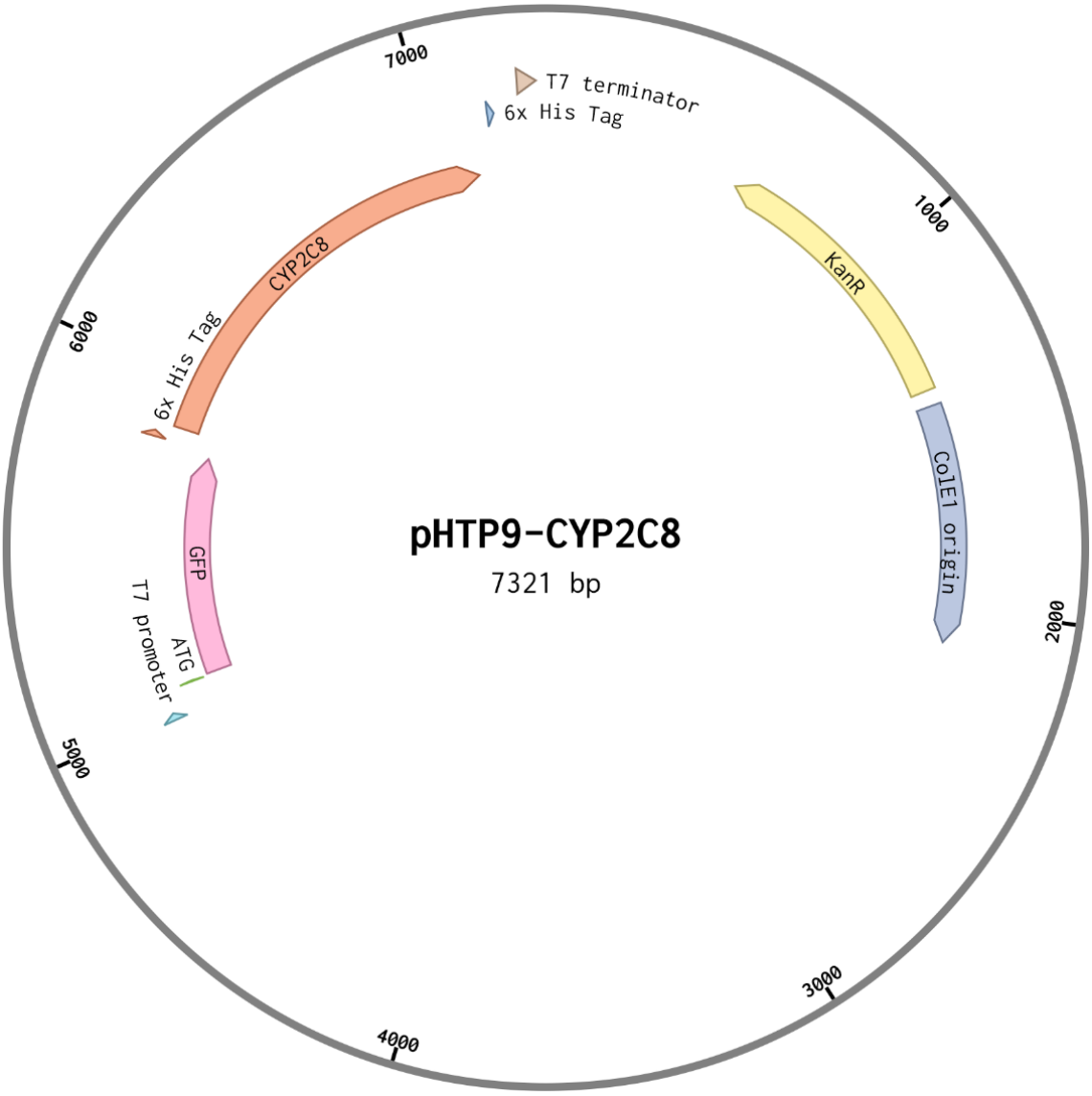


Figure A5 – Vector map for plasmid pHTP9-CYP2C8.

Annex 6 – pCR4-TOPO-3A4 map

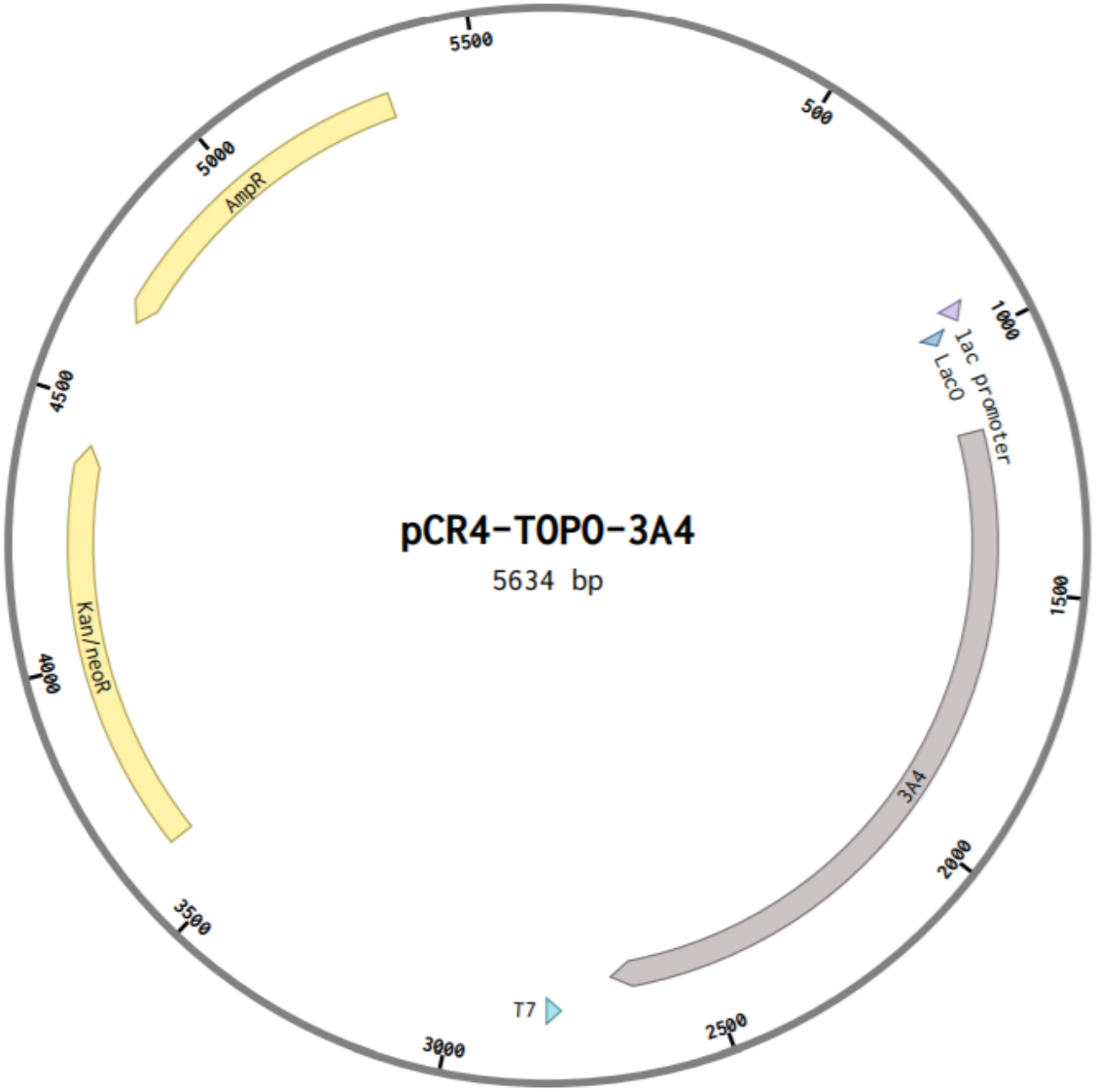


Figure A6 – Vector map for plasmid pCR4-TOPO-3A4. Plasmid map: Harvard PlasmID Database.

Annex 7 – pHTP9-CYP3A4 map

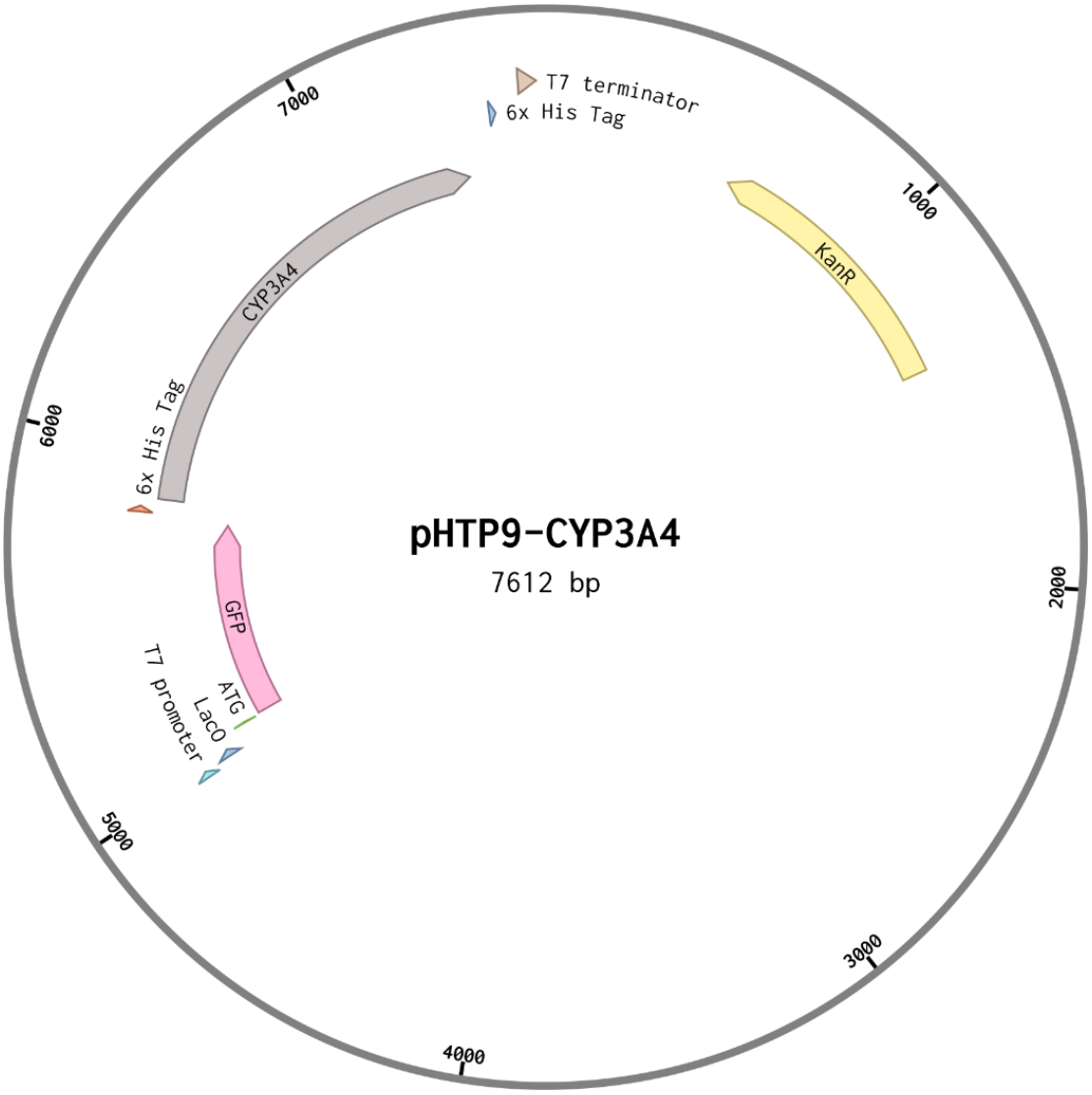


Figure A7 – Vector map for plasmid pHTP9-CYP3A4.

Annex 8 – pCMV-SPORT6-POR map

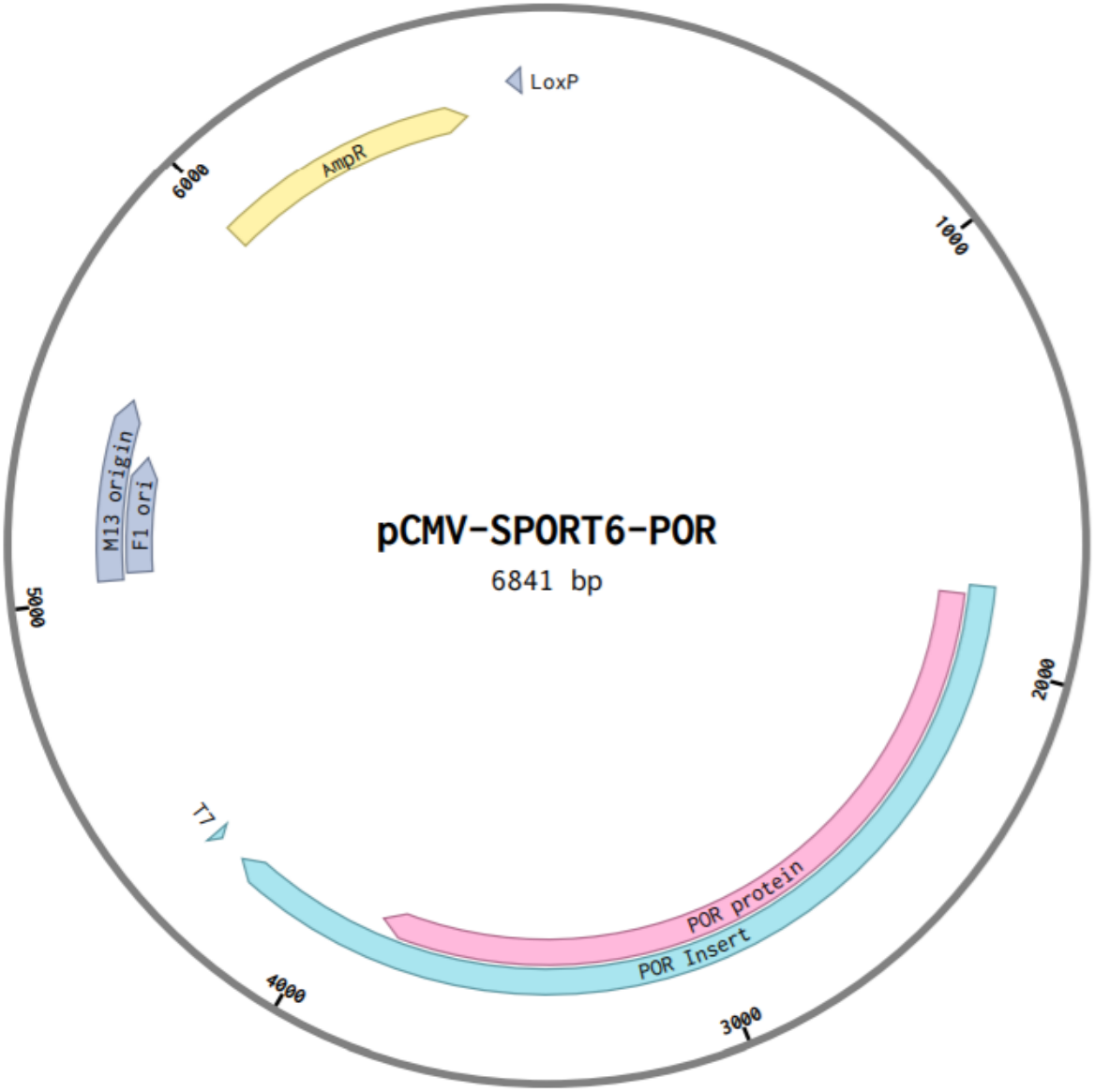


Figure A8 – Vector map for plasmid pCMV-SPORT6-POR. Plasmid map source: Harvard PlasmID Database.

Annex 9 – pET28a(+)-TEV-POR map

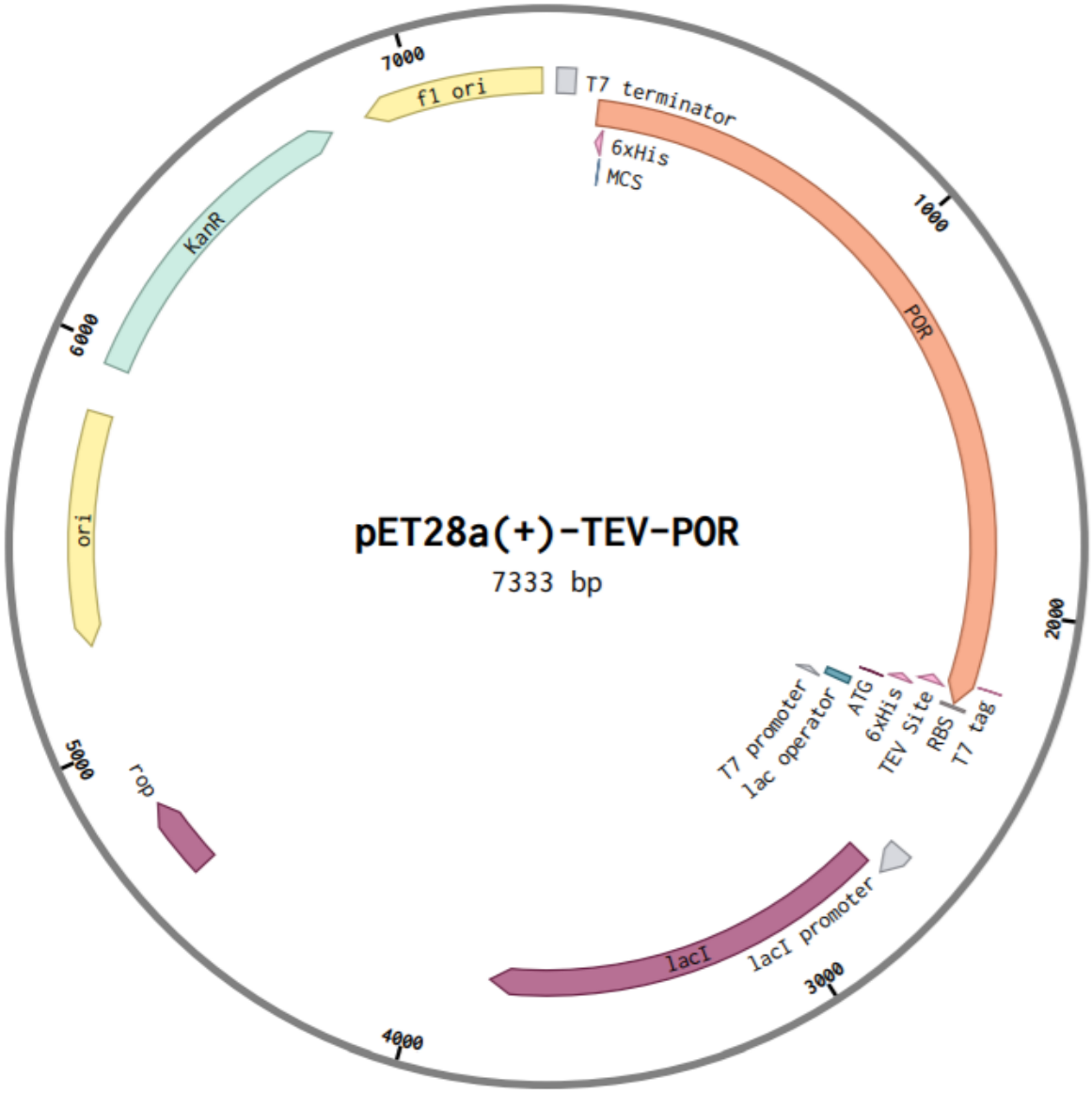


Figure A9 – Vector map for plasmid pET28a(+)-TEV-POR. Plasmid map source: GeneScript.

Annex 10 – pDONR221-UGT2B7 map

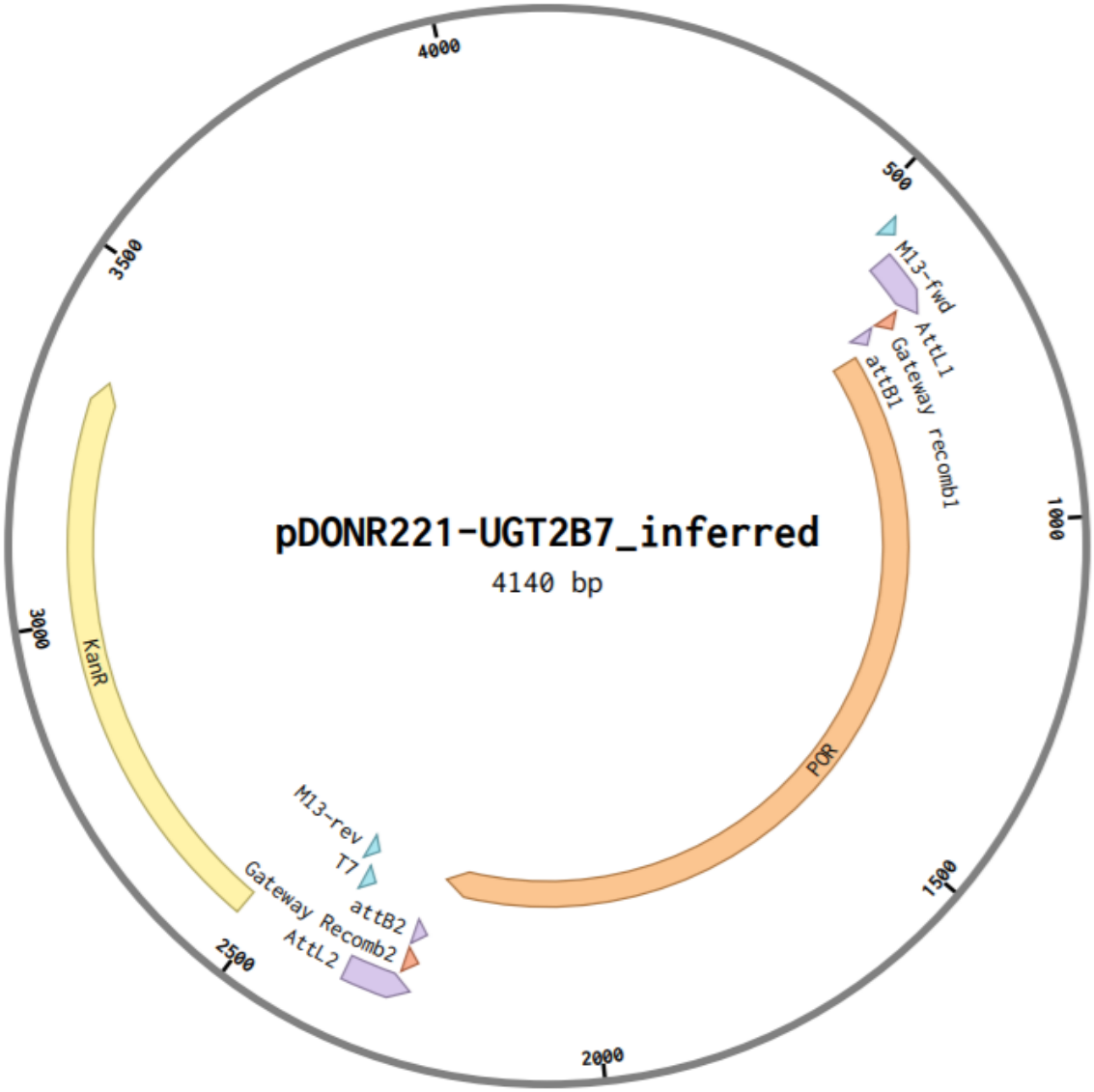


Figure A10 – Vector map for plasmid pDONR221-UGT2B7. Plasmid map source: Harvard PlasmID Database.

Annex 11 – pET-UGT2B7 map

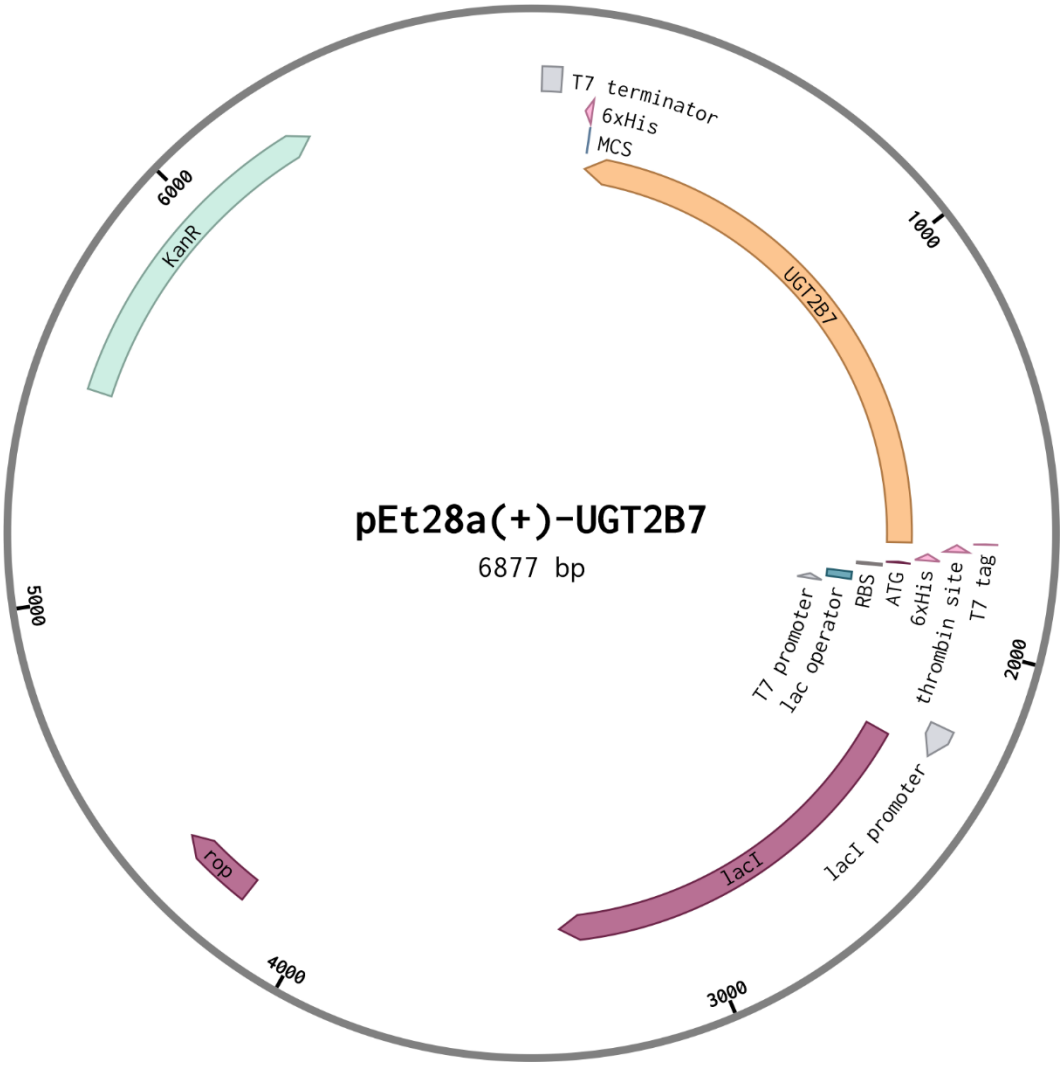
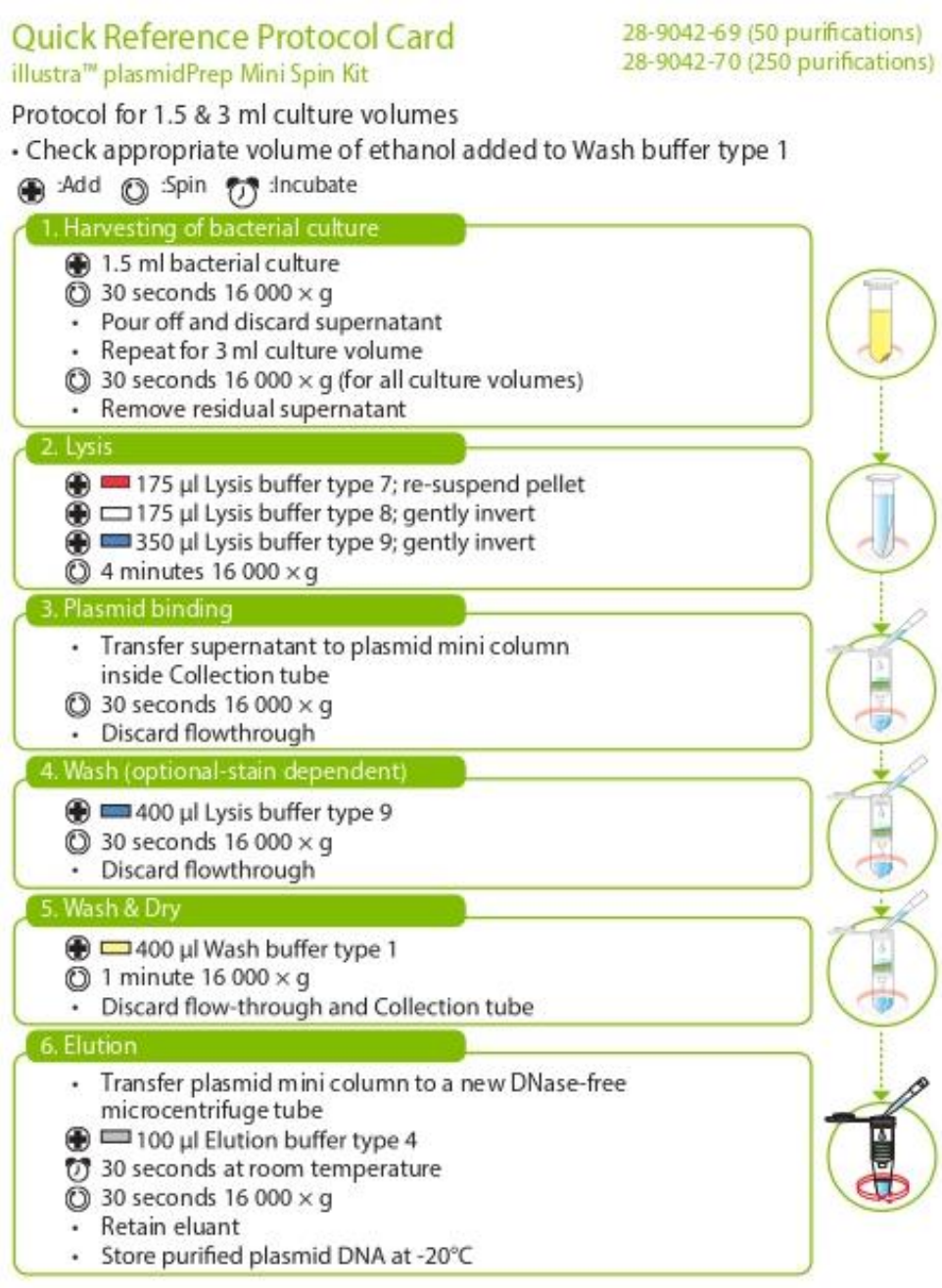


Figure A11 – Vector map for plasmid pET28(+)-UGT2B7.

ANNEXES – PROTOCOLS

Annex 12 – Plasmid extraction protocol



Quick Reference Protocol Card

illustra™ plasmidPrep Mini Spin Kit

28-9042-69 (50 purifications)
28-9042-70 (250 purifications)

Protocol for 1.5 & 3 ml culture volumes

- Check appropriate volume of ethanol added to Wash buffer type 1

:Add :Spin :Incubate

1. Harvesting of bacterial culture

- 1.5 ml bacterial culture
- 30 seconds 16 000 × g
- Pour off and discard supernatant
- Repeat for 3 ml culture volume
- 30 seconds 16 000 × g (for all culture volumes)
- Remove residual supernatant

2. Lysis

- 175 µl Lysis buffer type 7; re-suspend pellet
- 175 µl Lysis buffer type 8; gently invert
- 350 µl Lysis buffer type 9; gently invert
- 4 minutes 16 000 × g

3. Plasmid binding

- Transfer supernatant to plasmid mini column inside Collection tube
- 30 seconds 16 000 × g
- Discard flowthrough

4. Wash (optional-stain dependent)

- 400 µl Lysis buffer type 9
- 30 seconds 16 000 × g
- Discard flowthrough

5. Wash & Dry

- 400 µl Wash buffer type 1
- 1 minute 16 000 × g
- Discard flow-through and Collection tube

6. Elution

- Transfer plasmid mini column to a new DNase-free microcentrifuge tube
- 100 µl Elution buffer type 4
- 30 seconds at room temperature
- 30 seconds 16 000 × g
- Retain eluant
- Store purified plasmid DNA at -20°C

Annex 13 – Speedy *Hind*III nezymatic digestion protocol

nzytech

Speedy HindIII

Catalogue number:

MB09701, 500 reactions
MB09702, 2500 reactions

5'...A↓AGCTT...3'
3'...TTCGA↓A...5'

Description

NZYTech's Speedy restriction enzymes are a new generation of DNA modifying enzymes developed for rapid DNA digestion. All Speedy enzymes are 100% active in the NZYSpeedyBuffers and are able to digest DNA in 5-15 minutes. NZYSpeedyBuffers are universal and enable any combination of NZY Speedy restriction enzymes to work simultaneously. It's ideal for double or multiple digestions, eliminating the need for sequential digestions. NZYSpeedyBuffer Orange also allows ready gel load after digestion. Speedy HindIII activity is not affected by *dam* methylation, *dcm* methylation or CpG methylation. NZYTech's Speedy HindIII is a fast enzyme that performs its reaction usually in 5-15 minutes at 37 °C.

Concentration

One µL of enzyme can completely digest up to 1 µg of DNA in 5-15 min.

Reagents supplied with the enzyme

10× NZYSpeedyBuffer Colourless
10× NZYSpeedyBuffer Orange

Storage buffer

25 mM NaHepes, pH 7.5, 500 mM NaCl, 150 mM Imidazol, 2.5 mM $CaCl_2$, 50% (v/v) glycerol.

Recommended reaction conditions

Speedy HindIII	1 µL
10× NZYSpeedyBuffer	2 µL
Substrate DNA	≤ 1 µg
Sterilized ultrapure water	up to 20 µL

Incubation temperature

37 °C for 5-15 minutes.

Heat inactivation

65 °C for 20 minutes.

Quality control

Purity

Speedy HindIII has been determined to be >90% pure as judged by SDS-PAGE followed by BlueSafe staining (MB15201).

DNases assay

10 U of Speedy HindIII were incubated with 0.2-0.3 µg of pNZY28 for 14-16 hours at 37 °C. No nicking activity was observed following agarose gel electrophoresis.

Functional assay

Speedy HindIII was tested for performance in a digestion of 1 µg of a recombinant pNZY28 derivative using 1 µL of enzyme. The resulting digestion was visualized in an agarose gel.

Storage conditions

Speedy HindIII should be stored at -20 °C at all times. Do not keep it on ice while working at the bench.

Shipping conditions

Dry ice.

Product life

Please see the product labels.

V1901

Certificate of Analysis

Assay	Result
Protein purity	Pass
DNases assay	Pass
Functional assay	Pass

Approved by:

Patrícia Ponte
Senior Manager, Quality Systems

For research use only.

nzytech
genes & enzymes

Estrada do Paço do Lumiar, Campus do Lumiar - Edifício E, R/C, 1649-038 Lisboa, Portugal
Tel.: +351.213643514 Fax: +351.217151168
www.nzytech.com

Annex 14 – Speedy *Nde*I enzymatic digestion protocol

Speedy NdeI

Catalogue number:

MB10101, 50 reactions
MB10102, 250 reactions

5'...CA↓TATG...3'
3'...GTAT↓AC...5'

Description

NZYTech's Speedy restriction enzymes are a new generation of DNA modifying enzymes developed for rapid DNA digestion. All Speedy enzymes are 100% active in the NZYSpeedyBuffers and are able to digest DNA in 5-15 minutes. NZYSpeedyBuffers are universal and enable any combination of NZY Speedy restriction enzymes to work simultaneously. It's ideal for double or multiple digestions, eliminating the need for sequential digestions. NZYSpeedyBuffer Orange also allows ready gel load after digestion. Speedy NdeI activity is not affected by *dam* methylation, *dcm* methylation or CpG methylation. NZYTech's Speedy NdeI is a fast enzyme that performs its reaction usually in 5-15 minutes at 37 °C.

Concentration

One µL of enzyme can completely digest up to 1 µg of DNA in 5-15 min.

Reagents supplied with the enzyme

10× NZYSpeedyBuffer Colourless
10× NZYSpeedyBuffer Orange

Storage buffer

10 mM Tris-HCl pH 7.5, 50 mM KCl, 0.1 mM EDTA, 1 mM DTT, 200 µg/mL BSA, 50% (v/v) glycerol.

Recommended reaction conditions

Speedy NdeI	1 µL
10× NZYSpeedyBuffer	2 µL
Substrate DNA	≤ 1 µg
Sterilized ultrapure water	up to 20 µL

Incubation temperature

37 °C for 5-15 minutes.

Heat inactivation

65 °C for 20 minutes.

Quality control

Purity

Speedy NdeI has been determined to be >90% pure as judged by SDS-PAGE followed by BlueSafe staining (MB15201).

DNases assay

10 U of Speedy NdeI were incubated with 0.2-0.3 µg of pNZY28 for 14-16 hours at 37

°C. No nicking activity was observed following agarose gel electrophoresis.

Functional assay

Speedy NdeI was tested for performance in a digestion of 1 µg of a recombinant pNZY28 derivative using 1 µL of enzyme. The resulting digestion was visualized in an agarose gel.

Storage conditions

Speedy NdeI should be stored at -20 °C at all times. Do not keep it on ice while working at the bench.

Shipping conditions

Dry ice.

Product life

Please see the product labels.

V1901

Certificate of Analysis

Assay	Result
Protein purity	Pass
DNases assay	Pass
Functional assay	Pass

Approved by:

Patrícia Ponte
Senior Manager, Quality Systems

For research use only.

Estrada do Paço do Lumiar, Campus do Lumiar - Edifício E, R/C, 1649-038 Lisboa, Portugal
Tel.: +351.213643514 Fax: +351.217151168
www.nzytech.com

Annex 15 – Agarose gel DNA purification protocol

Data sheet

Clean-Easy Agarose Purification Kit

Cat. No: AN0070 (50 reactions)
Cat. No: AN0071(100 reactions)

Description

Clean-Easy Agarose Purification Kit provides a rapid and efficient method to extract DNA from agarose gels. It is based on the solubilisation and binding of DNA to a silica membrane in presence of chaotropic salts. Clean-Easy minispin columns contains an exclusive membrane that allows to bind an unique DNA fragment, previously excised from agarose gel.

Features

- **Simple** and **Just a few minutes** procedure.
- **Wide spectrum of size** fragments could be purified, (suitable since 100 bp up)
- **High Percentage of Recovery**, greater than 80% on 0.7-1% agarose. Recovery is lower in more concentrated agarose gels (50-60% on 2% agarose).
- DNA purified **Ready to use** in all molecular biology procedures.
- Suitable for any kind of agarose and gel buffer systems.

Applications

- Purification of DNA fragments (*obtained by PCR or digestion with restriction enzymes*) from agarose gels.
- The purified DNA can be used in all molecular biology applications.

Kit Components

Item	AN0070	AN0071
Clean-Easy minispin columns	50	100
Collection tubes (2 mL)	50	100
QG Buffer	60 ml	2X60 ml
PE Buffer*	11.25 ml	2X11.25 ml
EB Buffer	10ml	10ml

* Add ethanol (96%-100%) [not included] to PE Buffer prior to initial use. After ethanol has been added, mark the bottle to indicate that this step has been completed.

Storage:

Clean-Easy Agarose Purification Kit should be stored at room temperature (15–25°C) for up to 12 months without any reduction in performance.

Quality Certifications

Clean-Easy Agarose Purification Kit is tested in the purification of a 0.5 kb DNA fragment excised from 2% agarose gel. The purified band is analysed in agarose gel electrophoresis.

(Continued on reverse side)

Canvax Biotech, S.L. C/Astrónoma Cecilia Payne. Edif. Canvax. 14014 Córdoba, Spain.
☎ : +34 957 348 066
🖷 : +34 957 346 217
✉ : info@canvaxbiotech.com

canvax

canvax

DETAILED PROTOCOL

1. Using a clean, sharp razor blade or scalpel, excise the DNA ban from the agarose gel. Remove the extra agarose to reduce the size of gel slice. Place the gel slice in a 1.5 ml preweighted tube and weigh the gel slice (The maximum amount of gel slice per column is 400 mg).
2. Add 3 volumes of QG Buffer to 1 volume of gel. (For example, if the agarose gel slice is 100 mg, add 300 µl of QG buffer)

   For gels containing more than 2% agarose, add 6 volumes de QG Buffer per mg of gel.
3. Incubate at 50 ° C in a water bath for 10 min or until the gel slice has completely dissolved. During incubation at 50 °C, mix by vortexing or inverting the tubes every 1 minute. **Make sure the gel slice completely dissolved. For >2% gels, increase incubation time.**

   Important! *For fragments <500 bp and >4 kb, add 1 volume of isopropanol to the sample and mix (For example, if the agarose gel slice is 100 mg, add 100 µl isopropanol)*
4. Label the lid of a new minispin column placed in a 2 ml collection tube. Carefully apply the mix from previous step to the spin column and Centrifuge at 13000 rpm for 1 minute. For mixture volumes of more than 750 µl, load and centrifuge again using the same column.
5. Place the spin column in a new 2 ml collection tube, and discard the collection tube containing the filtrate. Add 700 µl of PE buffer to the minispin column and centrifuge at 13000 rpm for 1 minute.

   Remember! *Before using for the first time, add ethanol (96–100%) to PE Buffer as indicated on the bottle.*
6. Discard the flow-through and centrifuge again at 13000 rpm for 1 minute. This step is essential for removing trace buffer PE.
7. Transfer the column to a clean 1.5 ml microcentrifuge tube. Add 30 µl of Elution Buffer (EB) or H_2O (pH=7.0-8.5) to the center of the column membrane and incubate at room temperature for 1 minute. Centrifuge at 13000 rpm for 1 minute to elute and collect DNA.

*To increase the DNA yield you can warm the buffer EB/H2O to 65 °C

PRODUCT USE LIMITATION

This product is developed, designed and sold exclusively for research purposes and *in vitro* use only. The product was not tested for use in diagnostics or for drug development, and is not suitable for administration to humans or animals. Please refer to www.canvaxbiotech.com for the Material Safety Data Sheet of the product.

Canvax Biotech, S.L. C/Astrónoma Cecilia Payne. Edif. Canvax. 14014 Córdoba, Spain.
☎ : +34 957 348 066
🖷 : +34 957 346 217
✉ : info@canvaxbiotech.com

canvax

Annex 16 – PCR/enzymatic digestion product purification protocol

Data sheet

Clean-Easy PCR Purification Kit

Cat. No: AN0063-S (20 reactions)
Cat. No: AN0063 (50 reactions)
Cat. No: AN0064 (100 reactions)

Description

Clean-Easy PCR Purification kit provides a rapid and efficient method to purify DNA and remove contaminants from reaction mixtures (e.g. PCR, digestion or labeling reactions,). Clean-Easy minispin columns contain an exclusive membrane that allows DNA adsorption in presence of chaotropic salts and the removal of contaminants.

Features

- **Simple** and **Just a few minutes** procedure.
- **70-90% DNA recovery**.
- Suitable for DNA fragments as short as 75 bp.
- DNA purified **Ready to use** in all molecular biology procedures.

Applications

- Removal of proteins and salts from PCR, restriction digestion, dephosphorylation, ligation or labelling reactions.
- Changing of a restriction enzyme buffer.
- Re-purification of genomic DNA

Kit Components

ITEM	AN0063	AN0064
Clean-Easy minispin columns	50	100
Collection tubes (2 mL)	50	100
PB Buffer	25 ml	50 ml
PE Buffer *	11.25 ml	22.5 ml
EB Buffer	9 ml	9 ml

* Add the volume ethanol (96%-100%) specified [Not included] prior to initial use (see bottle label for volume). After ethanol has been added, mark the bottle to indicate that this step has been completed.

Store at RT.

Quality Certifications

Clean-Easy PCR Purification Kit is tested in the purification of a 0.6 kb DNA fragment from PCR mixture. The purified band is analysed in agarose gel electrophoresis.

(Continued on reverse side)

Canvax Biotech, S.L. C/Astrónoma Cecilia Payne. Edif. Canvax. 14014 Córdoba, Spain.
☎ :+34 957 348 066
🖷 :+34 957 346 217
✉ : info@canvaxbiotech.com

canvax

canvax

Assay procedure

1. Add 5 volumes of **Buffer PB** to one volume of PCR solution and mix thoroughly by pipette.
2. Label the lid of a new spin column placed in a 2 ml collection tube. Carefully apply the mix from step 1 to the spin column and Centrifuge at 13000 rpm for 1 minute.
3. Place the spin column in a new 2 ml collection tube, and discard the collection tube containing the filtrate.
4. Add 700 µl of **buffer PE** for Wash to the minispin column and centrifuge at 13000 rpm for 1 minute.
   Remember! *Before using it for the first time, add ethanol (96–100%) to the PE Buffer as indicated on the bottle.*
5. Discard the flow-through and centrifuge at 13000 rpm for 1 minute. *This step is essential for removing traces of PE buffer.*
6. Transfer the minispin column into a new, labeled 1.5 ml microcentrifuge tube.
7. Carefully open the minispin column and pipet 30 µl **Buffer EB** or H_2O (pH=7.0-8.5) directly onto the membrane. Close the cap and incubate for 1 min at room temperature, then centrifuge at 13000 rpm for 1 min to elute DNA. *To increase the DNA yield you can warm the buffer EB/H_2O to 65 °C before adding to the column.*

PRODUCT USE LIMITATION

This product is developed, designed and sold exclusively for research purposes and *in vitro* use only. The product was not tested for use in diagnostics or for drug development, and is not suitable for administration to humans or animals. Please refer to www.canvaxbiotech.com for the Material Safety Data Sheet of the product.

Canvax Biotech, S.L. C/Astrónoma Cecilia Payne. Edif. Canvax. 14014 Córdoba, Spain.
☎ : +34 957 348 066
🖷 : +34 957 346 217
✉ : info@canvaxbiotech.com

canvax

Annex 17 – T4 DNA ligase protocol

nzytech
genes & enzymes

T4 DNA Ligase

Catalogue number: MB00703, 500 U

Description

T4 DNA Ligase is an ultrapure recombinant enzyme purified from *Escherichia coli* and supplied with an optimized 10× Reaction Buffer. T4 DNA ligase catalyses the formation of a phosphodiester bond between juxtaposed 5'-phosphoryl and 3'-hydroxyl termini in duplex DNA. It repairs single-strand nicks in duplex DNA and will join both blunt and cohesive-end restriction fragments of duplex DNA or RNA. The enzyme requires ATP as cofactor.

Storage temperature

T4 DNA Ligase should be stored at -20 °C in a constant temperature freezer. The protein will remain stable till the expiry date if stored as specified.

Unit definition

One unit catalyses the exchange of 1 nmol of radiolabelled phosphate from pyrophosphate into Norit-absorbable material in 20 min at 37 °C under standard assay conditions.

Enzyme concentration: 5 U/µL

Inactivation

T4 DNA ligase is heat inactivated at 65 °C for 10 min.

System components and Reaction conditions

T4 DNA Ligase is provided with a dedicated highly optimized NZYTech reaction buffer, which contains ATP that is critical for this enzyme. Repeated freeze-thaw cycles will affect the stability of ATP, so we recommend making 10-20 µL aliquots of the reaction buffer and store at -20 °C. Vortex the reaction buffer solution thoroughly after thawing and prior to use.

The enzyme performs well at temperatures ranging from 16 °C to 25 °C. The optimal temperature for a ligation reaction is a balance between the enzyme's optimal temperature and the temperature required to ensure annealing of the DNA fragment ends, which can vary with the length and base composition of the terminal sequences.

Ligation Protocol

We recommend using a 1:3-10 molar ratio of vector:insert. To calculate optimal amounts of insert DNA in ligation reaction, see below:

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of insert}$$

Example: If using 50 ng of a vector plasmid with 3 kb, for a 1:10 molar ratio of vector:insert then you will require the following amount of a 500 bp insert:

$$\frac{50 \times 0.5}{3} \times 10 = 83 \text{ ng}$$

1. On ice, in a sterile, nuclease-free microcentrifuge tube, prepare a reaction mixture, combining the following components (volumes for a 20 µL reaction):

Component	Volume
10× Reaction buffer (provided)	2 µL
Vector DNA (20-50 ng)	x µL
Insert DNA (3-10 molar excess)	y µL
T4 DNA Ligase (5 U/µL)	1 µL
Nuclease-free water	up to 20 µL

2. Mix and centrifuge briefly to bring the contents to the bottom of the tube.

3. Incubate at 16-20 °C for 16 hours.

4. Use the ligation reaction to transform NZYTech competent cells.

Important notes

- It is extremely important not to change the ratio between the volume of T4 DNA Ligase and the reaction final volume to prevent decrease in efficiency of cloning reactions.
- For blunt-end ligations, use higher quantities of both vector and insert DNA.
- For sticky (cohesive)-end ligations, we recommend to heat both vector and insert DNA prior to the ligation.
- If the ligation mixture will be used for electroporation, a DNA purification step is recommended before the transformation. Use a spin column purification method (NZYGelpure, Cat. No. MB011) or chloroform extraction.

Quality control assays

Purity

Recombinant T4 DNA Ligase is >95% pure as judged by SDS polyacrylamide gel electrophoresis followed by BlueSafe (NZYTech, Cat. No. MB152) staining.

Nuclease assays

0.2-0.3 µg of pNZY28 plasmid DNA are incubated with 5 U of T4 DNA Ligase in 1× Reaction buffer for 14-16 hours at 37 °C. Following incubation, the DNA is visualized on a GreenSafe-stained agarose gel. There must be no visible nicking or cutting of the DNA.

Functional assay

Linearized pNZY28 plasmid (leaving either blunt-end or cohesive ends) is re-ligated with 5 U of T4 DNA Ligase. The DNA is then transformed into NZY5α competent cells that are plated on ampicillin plates. The re-ligation efficiency is determined by counting transformed bacterial colonies.

Related products

Product name	Cat. No.
NZY5α	MB004
NZYStar	MB005
Speedy Ligase	MB130
NZYGelpure	MB011

V2101

Annex 18 – Supreme NZYProof DNA polymerase protocol

nzytech
genes & enzymes

Supreme NZYProof DNA polymerase

Catalogue number: MB28301, 125 U
MB28302, 500 U
MB28303, 1000 U

Description

Supreme NZYProof DNA polymerase is an engineered highly accurate, fast and sensitive variant of NZYProof DNA polymerase displaying a hot-start like PCR capacity. This feature is achieved by a novel hot-start technology, which inhibits both polymerase and 3'→5' exonuclease activities and thus avoids extension of non-specifically annealed primers or primer-dimers, as well as the degradation of primers and template DNA during PCR reaction setup. Thus, Supreme NZYProof DNA polymerase generates higher specificity, sensitivity and yield during the accurate amplification of DNA. Supreme NZYProof DNA polymerase was also engineered for higher processivity, thus allowing fast PCR reactions of longer PCR products.

This highly robust version of NZYProof DNA polymerase is a broad range enzyme suitable for a variety of applications, including amplification of longer (≤10 kb) and difficult PCR products, as well as site-directed mutagenesis. Supreme NZYProof DNA polymerase possesses 3'→5' exonuclease proofreading activity that enables the polymerase to amplify DNA with increased accuracy. The error rate of Supreme NZYProof DNA polymerase is similar to that of *Pfu* and *Kod* DNA polymerases and significantly lower than the error rate of *Taq* DNA polymerases. It generates blunt-ended PCR products that are suitable for cloning with NZYTech's NZY-blunt PCR cloning kit (MB121).

Storage Conditions

Supreme NZYProof DNA polymerase should be stored at -20°C at all times in a constant temperature freezer. Keep it on freezer while perform PCR set up until use. Supreme NZYProof DNA polymerase will remain stable till the expiry date if stored as specified.

Unit definition

One unit is defined as the amount of enzyme required to catalyse the incorporation of 10 nmoles of dNTPs into acid insoluble material in 30 minutes at 72 °C.

Enzyme concentration 2.5 U/µL

Reaction buffer (5×)

The enzyme reaction buffer contains Mg^{2+} (1.5 mM at the final, 1×, reaction concentration). Presence of a sugar alcohol in its composition decreases freezing temperatures. The buffer remains stable at 4 °C up to one month.

Standard Protocol

The following standard protocol is provided to ensure successful PCR amplification using Supreme NZYProof DNA polymerase. Optimal reaction conditions (e.g. concentration of DNA Polymerase, primers, dNTPs and template DNA) may need to be optimized for long amplicons or difficult templates. It is strongly recommended to assemble all reaction components on ice and quickly transfer the reactions to a thermocycler preheated to the denaturing temperature to start the PCR.

1. Gently mix and briefly centrifuge all components after thawing. At room temperature, or on ice, in a sterile, nuclease-free microcentrifuge tube, prepare a mixture for the appropriate number of PCR reactions. Add water first and the remaining components in the order specified in the table below. It is strongly advisable that the enzyme is the last component to add to the reaction in order to minimize primer degradation due to the 3'→5' exonuclease activity. A single 50 µL reaction mixture should combine the following components:

Component	Volume	Final conc.
5× Reaction buffer	10 µL	1×
25 mM dNTPs	0.4 µL	0.2 mM
10 µM Forward primer	2.5 µL	0.5 µM
10 µM Reverse primer	2.5 µL	0.5 µM
Template DNA	X µL	(*)
Supreme NZYProof DNA Polymerase (2.5 U/µL)	0.5 µL	0.025 U/µL
Nuclease-free water	to 50 µL	-

(*) Typically, add DNA at a range of 1ng-0.5µg. See below general recommendations for DNA template.

2. Mix and quickly pulse the reactions.

3. Immediately initiate the PCR by transferring the PCR mixtures to the thermocycler with the block pre-heated to 95°C and following the below cycling parameters:

Cycle step	Temp.	Time	Cycles
Initial denaturation	96 °C	4 min	1
Denaturation	96 °C	30 s	25-35
Annealing	*	30 s	
Extension	72 °C	30 s/kb¥	
Final Extension	72 °C	5-10 min	1

* Annealing temperature should be optimized for each primer set based on the primer T_m; typically it should be T_m-5 °C.

¥ Use 40s/kb for PCR products >3 kb.

4. Analyse PCR products by agarose gel electrophoresis (0.7-1.2%, w/v) and visualize with GreenSafe (MB088) or any other means.

Primer Design

PCR primers generally range in length from 15–30 bases and are designed to flank the region of interest. Sequences longer than 30bp may improve PCR yield using Supreme NZYProof DNA polymerase since its 3'→5' exonuclease activity may degrade primers. In addition, to overcome primer degradation, the 3' termini of primers may be protected with phosphorothioate modifications. Primers should contain 40–60% GC, and avoid sequences that might produce internal secondary structure. The 3'-ends of the primers should not be complementary to avoid the production of primer-dimers. Primer-dimers unnecessarily remove primers from the reaction and result in an unwanted polymerase reaction that competes with the desired reaction. Avoid three G or C nucleotides in a row near the 3'-end of the primer as this may result in non-specific primer annealing increasing the synthesis of

undesirable reaction products. Ideally, both primers should have nearly identical melting temperatures (T_m), allowing their annealing with the denatured template DNA at roughly the same temperature.

DNA template

The optimal amount of starting material may vary depending on its quality and complexity. In general, we recommend using 10ng to 500ng of genomic DNA templates, although the enzyme is sensitive enough to amplify fragments from as little as 1ng of human gDNA, for example. Lower amounts of template may be used for amplification of less complex DNA (typically 1-50ng). When using a cDNA synthesis reaction as template do not exceed 10% of the final PCR reaction volume.

Enzyme concentration

In general, we recommend using 1.25 U of enzyme (0.5µL) in a 50 µL reaction. You may increase the volume of enzyme to a maximum of 2.5 U (1 µL) in a 50 µL reaction when amplifying abundant templates (>50 ng gDNA). Do not exceed this enzyme concentration in particular for longer PCR products (>5 kb). For convenience during PCR assembly, enzyme may be dilute in water (for example, dilute 1/10 in water to add 5 µl of diluted enzyme instead of 0.5 µl of undiluted preparation).

Quality control assays

Purity

Supreme NZYProof DNA polymerase purity is >90% as judged by SDS-PAGE followed by Coomassie Blue staining.

Genomic DNA contamination

Supreme NZYProof DNA polymerase must be free of any detectable DNA contamination as evaluated through PCR.

Nuclease assays

0.2-0.3 µg of pNZY28 plasmid DNA are incubated with 5 U of Supreme NZYProof DNA polymerase, in 1× reaction buffer, for 14-16 hours at 37 °C. Following incubation, the DNA is visualised on a GreenSafe-stained agarose gel. There must be no visible nicking or cutting of the nucleic acid. Similar tests are performed with reaction buffer.

Functional assay

Supreme NZYProof DNA polymerase is extensively tested for performance in a PCR reaction using 1.25 units of enzyme for the amplification of different-sized DNA fragments (1 and 2.5 kb) from human genomic DNA. The resulting PCR products are visualized as a single band in a GreenSafe stained agarose gel.

Troubleshooting

No product amplification or low yield
• Inadequate annealing temperature
The reaction mix composition may affect the melting properties of primers and DNA. Adjust the annealing temperature to accommodate the primer with the lowest melting temperature (5 ° to 10 °C lower than T_m).
• Presence of PCR inhibitors
Some DNA isolation procedures, particularly genomic DNA isolation, can result in the co-purification of PCR inhibitors. Reduce the volume of template DNA in reaction or dilute template DNA prior to adding to the reaction. Diluting samples even 1:10,000 has been shown to be effective in improving results, depending on initial DNA concentration.

V1902

Certificate of Analysis

Test	Result
Enzyme purity	Pass
Genomic DNA contamination	Pass
DNase contamination	Pass
Functional assay	Pass

Approved by:

Patrícia Ponte
Senior Manager, Quality Systems

For research use only.

nzytech
genes & enzymes

Estrada do Paço do Lumiar, Campus do Lumiar - Edifício E, R/C, 1649-038 Lisboa, Portugal Tel.: +351.213643514 Fax: +351.217151168
www.nzytech.com

Annex 19 – Cell lysis protocol

nzytech
genes & enzymes

NZY Bacterial Cell Lysis Buffer

Catalogue number:

MB17801, 250 mL
MB17802, 2 × 250 mL

Features

- Disrupt cells without mechanical procedures
- Ready-to-use formulation
- Ideal for high-throughput (HTP) methods
- DNase I and Lysozyme provided separately
- Outstanding yields of extracted proteins

Description

NZY Bacterial Cell Lysis Buffer is an innovative product developed at NZYTech for the gentle disruption of *Escherichia coli* cell wall, generating a homogeneous cell-free extract. It provides a rapid and cost-effective alternative to harsh chemicals or mechanical procedures, such as French Press or sonication, for releasing recombinant and native proteins without denaturation. This extraction reagent is a Tris-buffered formulation (pH 7.5) with lysozyme and DNase I provided separately. The addition of these enzymes allows a most efficient extraction. However, for some over-expressed proteins and particular *E. coli* strains the addition of lysozyme may not be required. Please consider that the presence of lysozyme and DNase I might interfere with some downstream applications (in these situations, do not add the enzymes).

Additional components, such as protease inhibitors, salts, reducing agents and chelating agents (not provided), may be added to the lysate obtained using the NZY Bacterial Cell Lysis Buffer, depending on the particular application. Note that chelating agents should not be added to the lysate in the presence of lysozyme and DNase I.

Shipping conditions

NZY Bacterial Cell Lysis Buffer, Lysozyme and DNase I are shipped at 4 °C to room temperature.

Storage conditions

Store NZY Bacterial Cell Lysis at room temperature and lysozyme and DNase I at -20 °C in a freezer without defrost cycles. Buffer and enzymes are stable until their expiration dates if stored and handled properly.

System Components

Component	MB17801	MB17802
NZY Bacterial Cell Lysis buffer	250 mL	5 x 250 mL
Lysozyme (at 50 mg/mL)	0.5 mL	2 x 0.5 mL
DNase I (at 2 mg/mL)	0.5 mL	2 x 0.5 mL

Protocol for extracting recombinant or native proteins from bacteria (*E. coli*)

1. Harvest cells from liquid culture by centrifugation at 2000-5000 ×*g* for 10 min at 4 °C.

2. Decant and allow the pellet to drain, removing as much supernatant as possible. Determine the weight of the pellet. **Note:** The protein extraction is typically more efficient if the pellets are frozen for at least 30 min.

3. Resuspend the cell pellet at room temperature by pipetting or gentle vortexing using 5 mL of NZY Bacterial Cell Lysis Buffer per gram of cell past.
Note: In high-throughput protocols, when using small cultures, typically use 1 mL of Buffer per 5-mL cultures.

4. Optional but highly recommended: Add 2 µL of lysozyme at 50 mg/mL and 2 µL of DNase I per 1 mL of NZY Bacterial Cell Lysis Buffer. Optional: Add EDTA-free protease inhibitors (not provided).

5. Incubate the cell suspension on a shaking platform or rotating mixer for 10-20 min at room temperature.

6. Remove insoluble cell debris by centrifugation at 15,000 ×*g* for 15 min at 4 °C.
Note: If desired, save the pellet for inclusion body purification.

Important notes

- NZY Bacterial Cell Lysis Buffer can be used on fresh or frozen cell pellets. Superior extraction efficiencies can be obtained by freezing the pellets before adding the Buffer.
- Extraction efficiency can be strain-dependent. The NZY Bacterial Cell Lysis Buffer can be used to disrupt the most common bacterial host strains and it is especially efficient with BL21 strain and its derivatives. The use of pLysS or pLysE hosts enhances the extraction procedure.
- Whole cell lysates from NZY Bacterial Cell Lysis Buffer are compatible with Bradford protein assay as well as with Lowry and biuret reagents.
- The resulting protein extract can be used in preparation of different purification methods, such as GST or His-tagged immobilized metal affinity chromatography (IMAC), or in other applications.

Quality control assays

Functional assay

NZY Bacterial Cell Lysis Buffer is tested functionally in a protocol for extracting recombinant proteins from *E. coli* strain BL21(DE3). The released recombinant proteins are purified through IMAC and separated through SDS-PAGE.

Data

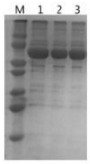


Figure 1. Comparing the efficiency of cell lysis using three different protein extraction methods. *E. coli* cells harvested from 5 mL of cultured media were lysed using three different procedures (Lane 1: NZY Bacterial Cell Lysis Buffer; Lane 2: Sonication; Lane 3: Competing product). The recombinant protein was purified through IMAC and separated through SDS-PAGE. M: Low Molecular Weight (LMW) Protein Marker (MB082).

Troubleshooting

The cell lysate is hazy
• Cell density is too high
Additional NZY Bacterial Cell Lysis Buffer can be added for high density cell cultures to solubilize remaining particulates. Cell lysate can be clarified by centrifugation at 16,000×*g* for 20 min.
• Incubation time is too short
Incubate the cell lysate for at least 20 min to ensure that all cell components are completely solubilized.
• Temperature is too low
Incubate lysate at room temperature (20-25 °C). When using frozen pellets extend the incubation period or quickly warm lysate to room temperature.

Viscosity of extract is high
• DNase I has low activity or is inactive Cell density is too high
Add more DNase I. Check if the enzyme is stored and handled properly.

Protein of interest is not solubilized
• Protein of interest is expressed in inclusion bodies
Adjust expression conditions. Ensure that lysozyme and DNase I are added during cell lysis.
• Cells are not completely lysed
Ensure cell lysis is performed for at least 20 min at room temperature to allow for complete lysis of the cells.

V1901

Annex 20 – Alkaline phosphatase protocol

nzytech
genes & enzymes

Alkaline Phosphatase (*E. coli*)

Catalogue number: MB01801, 200 U

Description

Alkaline Phosphatase (*E. coli*) non-specifically catalyses the removal of most phosphomonoester bonds from the 5' and 3' termini of DNA and RNA without degrading diphosphate or triphosphate linkages. The enzyme is suitable for removal of terminal monoesterified phosphates from deoxyribo-oligonucleotides. The enzyme is generally used for the removal of single phosphate groups from 5´-ends of linear vectors to prevent re-circularization during cloning or to dephosphorylate DNA prior to kinase labelling protocols.

Storage temperature

Alkaline Phosphatase should be stored at -20 °C in a constant temperature freezer. The protein will remain stable till the expiry date if stored as specified.

Unit definition

One unit liberates 1 µmol of p-nitrophenol per minute at 25 °C, pH 8.0 with p-nitrophenyl phosphate as the substrate.

Enzyme concentration: 0.5 U/µL

Inactivation

Alkaline Phosphatase (*E. coli*) is highly resistant to heat inactivation. Thus, alternative protocols should be considered when requiring removing the enzyme from reactions, such as DNA silica column purification or phenol/chloroform extraction.

System components and Reaction conditions

Alkaline Phosphatase (*E. coli*) is provided with a dedicated highly optimized NZYTech reaction buffer, which includes Zn^{2+} is essential for enzyme activity. The optimum reaction temperature is 37 °C.

Protocol for dephosphorylation of 5´-ends of DNA

1. Perform a typical dephosphorylation reaction using 1.0 pmol of DNA termini (1.0 µg of a 3.0 kb plasmid) and using between 0.1 to 1 U of Alkaline Phosphatase (higher amounts for 5´-recessed terminus) in 1× Reaction buffer (provided).

2. Incubate at 37 °C for 1 hour.

3. After the reaction is complete, perform a DNA purification step using NZYGelpure (NZYTech, Cat. No. MB011).

Quality Control Assays

Purity

Alkaline Phosphatase (*E. coli*) is >95% pure as judged by SDS polyacrylamide gel electrophoresis followed by BlueSafe staining (NZYTech, Cat. No. MB15201).

Nucleases assays

To test for DNase contamination, 0.2-0.3 µg of pNZY28 plasmid DNA are incubated with 1 U of Alkaline Phosphatase for 14-16 hours at 37 °C. To test for RNase contamination, 1 µg of RNA is incubated with 1 U of Alkaline Phosphatase for 1 hour at 37 °C. Following incubation, the nucleic acids are visualized on a GreenSafe-stained agarose gel. There must be no visible nicking or cutting of the nucleic acids.

Functional assay

Alkaline Phosphatase (*E. coli*) is assayed in a dephosphorylating reaction followed by ligation and transformation. The efficiency of enzyme reaction is evaluated by the number of bacterial colonies transformed with the cloning product.

Related products

Product name	Cat. No.
Water for Molecular Biology	MB111
T4 Polynucleotide Kinase	MB008
Klenow Fragment of DNA Polymerase I	MB00901
Speedy Ligase	MB130
NZY5α competent cells	MB004

V2101

Certificate of Analysis

Test	Result
Enzyme purity	Pass
Nucleases contamination	Pass
Functional assay	Pass

Approved by:

Patrícia Ponte
Senior Manager, Quality Systems

For research use only.

nzytech
genes & enzymes
Estrada do Paço do Lumiar, Campus do Lumiar - Edifício E, R/C, 1649-038 Lisboa, Portugal Tel.: +351.213643514 Fax: +351.217151168
www.nzytech.com

Annex 21 – Non-radioactive Phosphorylation with T4 PNK protocol

Non-radioactive Phosphorylation with T4 PNK or T4 PNK (3´ phosphatase minus)

Protocols.io also provides an interactive version of this protocol where you can discover and share optimizations with the research community. It is available for both the Non-Radioactive Phosphorylation with T4 PNK and the Non-Radioactive Phosphorylation with T4 PNK (3´ phosphatase minus).

1. Set-up the following reaction in a microcentrifuge tube on ice:

DNA	up to 300 pmol of 5´ termini
T4 PNK Reaction Buffer (10X)	5 µl
ATP (10 mM)	5 µl
T4 PNK	1 µl (10 units)
Nuclease-free Water	up to 50 µl

2. Incubate at 37°C for 30 minutes.

3. Heat inactivate by incubating at 65°C for 20 minutes.

Notes:

- ATP is not supplied
- 1X T4 DNA Ligase Buffer contains 1 mM ATP and can be substituted in non-radioactive phosphorylations (T4 Polynucleotide Kinase exhibits 100% activity in this buffer).

Annex 22 – NZYEasy Cloning & Expression kits I and IX protocol

nzytech

NZYEasy Cloning & Expression kits

Catalogue numbers:

MB282, NZYEasy Cloning & Expression kit I
MB319, NZYEasy Cloning & Expression kit II
MB320, NZYEasy Cloning & Expression kit III
MB321, NZYEasy Cloning & Expression kit IV
MB322, NZYEasy Cloning & Expression kit VII
MB323, NZYEasy Cloning & Expression kit VIII
MB324, NZYEasy Cloning & Expression kit IX
MB325, NZYEasy Cloning & Expression kit X
MB326, NZYEasy Cloning & Expression kit XI
MB327, NZYEasy Cloning & Expression kit XIII
MB328, NZYEasy Cloning & Expression kit XIV
MB329, NZYEasy Cloning & Expression kit XVI
MB330, NZYEasy Cloning & Expression kit XVII

Unit sizes: (available for each kit)

8 and 96 reactions

Description

NZYEasy Cloning & Expression kits were designed to allow directional cloning of any PCR-generated fragment or synthetic gene into a linearized pHTP *Escherichia coli* expression vector. Cloning proceeds in a single ligase-independent reaction mediated by the NZYEasy enzyme mix. Vector-complementary overhangs containing a specific sequence recognized by the NZYEasy enzyme are incorporated in the PCR product by using primers with appropriate 5' extensions. When you combine the insert thus generated with the linearized pHTP vector, also containing complementary overhangs, in the presence of NZYEasy enzyme mix, the two DNA molecules will anneal through base-pair complementation of the single-strand regions. The reaction occurs in a single-tube along three temperature-dependent steps. Circular recombinant vector containing the fragment of interest is obtained by transforming the annealed plasmid DNA into competent *E. coli* cells. The system allows achieving high cloning efficiency (80-100%) and does not require the use of DNA ligases. In addition, the insert does not require any preliminary treatment (e.g. restriction digestion, phosphorylation, or blunt-end polishing).

Cloning is performed using conventional *E. coli* strains. Once pHTP recombinant plasmid has been constructed and its sequence confirmed it should be used to transform λDE3 *E. coli* lysogens, such as BL21(DE3), for high levels of protein expression. NZYEasy Cloning & Expression kits have been successfully used in high-throughput (HTP) platforms for the efficient cloning and expression of a large number of genes at a scale compatible with the functional screen of hundreds to thousands of genes/proteins.

NZYTech provides a comprehensive portfolio of pHTP expression vectors, which include different fusion tags commonly used to enhance expression and/or solubility of recombinant proteins in *E. coli*, as well as fluorescent tags. pHTP1 vector (included in NZYEasy Cloning & Expression kit I, cat. No. MB281) contains two poly-histidine (6xHis) sequences (N- and C-terminal) which allow subsequent recombinant protein purification by immobilized metal ion affinity chromatography (IMAC). The other pHTP expression vectors were constructed by inserting fusion tags (see Table 1) into the pHTP1 backbone such that the fusion partner will be at the N-terminus of the recombinant protein.

Table 1. pHTP expression vectors

Vector	Fusion Protein	Kit cat. No.
pHTP1	*No fusion tag besides His_6 sequences*	MB282
pHTP2	Leader less disulfide-bond isomerase DsbC (**LLDsbC**)[1]	MB319
pHTP3	Mutant version of disulfide-bond isomerase Dsbc (**mutDsbC**)[2]	MB320
pHTP4	Disulfide-bond isomerase **DsbC**[1]	MB321
pHTP7	Disulfide oxidoreductase **DsbA**[2]	MB322
pHTP8	Thioredoxin (**Trx**)[3]	MB323
pHTP9	Green fluorescent protein (**GFP**)[4]	MB324
pHTP10	N-utilization substance A (**NusA**)[5]	MB325
pHTP11	Glutathione S-transferase (**GST**)[6]	MB326
pHTP13	Gb1 Domain of Protein G (**GB1**)[7]	MB327
pHTP14	Ketosteroid isomerase (**KSI**)[8]	MB328
pHTP16	*R. flavefaciens* cellulosomal protein (**cpA**) [A]	MB329
pHTP17	*R. flavefaciens* cellulosomal protein (**cpB**) [A]	MB330

[A] CpA and CpB are two recombinant cellulosomal proteins (Cps) that are highly expressed in E. coli. CpA is a carbohydrate-binding module, displaying affinity for ß-glycans (xyloglucan, glucomannan, galactomannan and barley ß-glucan).

References

1) Nozach, H. et al. 2013 *Microb. Cell Fact.* **12**(37):2-16
2) Collins-Racie, L.A. et al. 1995 *Biotechnol.* **13**(9):982-987
3) LaVallie, E.R. et al. 1993 *Biotechnol.* **11**(2):187-193
4) Prendergast, F.G & Mann, K.G. 1978 *Biochemistry* **17**(17):3448-53
5) Davis, G.D. et al. 1999 *Biotechnol. Bioeng.* **8**:*1668-1674*
6) Smith, D.B. & Johnson, K.S. 1988 *Gene* **67**(1):31-40
7) Huth, J.R. et al. 1997 *Protein Sci.* **6**:2359-64
8) Kuliopulos, A. & Walsh, C.T. 1994 *J. Am. Chem. Soc.* **116**:4599-4607

Storage temperature

Kit components may be stored at -20 °C or at -80 °C.

Kit components

Component	8 reactions	24 reactions	96 reactions
10x Reaction Buffer	8 µL	24 µL	96 µL
NZYEasy enzyme mix	4 µL	12 µL	48 µL
pHTP vector	8 µL	24 µL	96 µL
Positive control [1]	10 µL	10 µL	10 µL

[1] Positive Control: PCR fragment provided for 5 experiments.

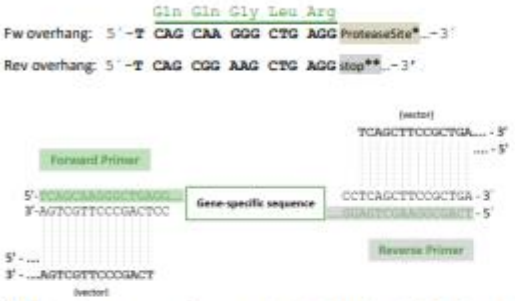
NZYEasy cloning protocol

Before you start using this protocol, please read carefully the NZYEasy Cloning Expression System User Guide available at the product resources tab of the product.

1. Preparing DNA inserts by PCR

1.1 Guidelines for Primers Design:

Besides gene-specific sequences, the following 16 bp overhangs must be included on 5'-ends of both forward (Fw) and reverse (Rev) primers, in order to provide the required vector-complementary single-strand terminals:



(*) If desired, a protease specific site can be included at the end of the forward overhang, just before the gene-specific sequence, in order to remove the N-terminal protein tag by a protease. For example, the cleavage recognition sequence of the Tobacco Etch Virus (TEV) protease is ENLYFQ↓(G/S).

(**) If a C-terminal His-tag is not desired, include an in-frame stop codon on the reverse primer (TTA, TCA or CTA). Omit the stop codon if you require both N- and C-terminal His-tags. For details please see the pHTP vector maps available on our website.

1.2 Guidelines for PCR amplification:

- We strongly recommend using a high-fidelity enzyme to reduce the error rate.
- When genes are isolated from plasmids with kanamycin resistance (same as the pHTP expression vectors), use 0.1-0.5 ng of template per 50 µL PCR reaction. Digestion with DpnI (NZYTech, cat. No. MB078) is recommended when high amounts of plasmid template are used.
- For optimal cloning efficiencies, spin-column purification of the PCR product using NZYGelpure kit (NZYTech, cat. No. MB011) is highly recommended. Gel-extraction of the desired band should be performed in case non-specific amplifications or primer-dimers are formed, thus enhancing cloning efficiencies

1.3 Sub-cloning of synthetic genes into pHTP1:

When transferring a synthetic gene to pHTP vectors, the following overhang regions are required upstream and downstream the gene-specific sequence:

2. Ligase-independent cloning reaction

We recommend using a vector:insert molar ratio of 1:5; Please use the table below to determine the optimal amount of PCR product to be used in a cloning reaction:

Fragment length (bp)[1]	Optimal DNA quantity for Cloning reaction (ng)
100	8.3
300	25.0
500	42.0
1000	83.0
2000	166.0
3000	249.0
4000	332.0

[1] ng of insert required = DNA fragment length (bp) × 0.083
e.g. 1348 bp gene = 1348 x 0.083 = 114.9 ng of DNA

2.1. On ice, in a sterile, nuclease-free microcentrifuge tube, prepare the following reaction mixture:

Component	Volume
Purified DNA fragment	x µL[2,3]
pHTP vector [3]	1 µL
10x Reaction Buffer	1 µL
NZYEasy enzyme mix	0.5 µL
Nuclease-free water	up to 10 µL

[2] Use a maximum of 7.5 µL of purified PCR insert when it is not possible to use the recommended optimal amount.
[3] Positive Control: PCR fragment of 500 bp is provided at 21.0 ng/µL (enough for 5 experiments). Please use 2 µL per reaction.
[3] pHTP vectors are provided in a ready-to-use form.

2.2. Mix the reactions by pipetting and spin to collect contents at the bottom of the tubes.

2.3. Perform the cloning reaction in a thermal cycler programmed with the following protocol:

Temperature (°C)	Time (min)
37	60
80	10
30	10
4	∞

2.4. Centrifuge briefly to collect the reaction components.

3. Transformation

3.1. Add 10 µL of ligation product directly into 100 µL NZY5α cells (NZYTech cat No. MB004) competent cells.
3.2. Place the mixture on ice for 30 min. Heat shock cells at 42 °C for 40 seconds. Place tube on ice for 2 minutes.
3.3. Add 900 µL of pre-warmed SOC media and incubate at 200 rpm at 37 °C for 1 hour.
3.4. Centrifuge at 5000 rpm for 1 min. Remove 900 µL of supernatant.
3.5. Re-suspend cells by gentle pipetting. Spread 100 µL of the cells onto the selection LB agar plates containing 50 µg/mL kanamycin.
3.6. Incubate inverted plates overnight at 37 °C.

Note: Significantly lower cloning efficiencies can result from using other E. coli strains than DH5α.

4. Screening for recombinant clones

Screening for recombinants can easily be achieved by colony-PCR, restriction analysis and/or sequencing. For colony PCR or sequencing, use the following pHTP vector-specific primers:

Vector	Forward primer (5'→ 3')
pHTP1	GCGAAATTAATACGACTCACTATAGGGG
pHTP2	CAATGGCACACTTGTTCCGGGTTAC
pHTP3	CAATGGCACACTTGTTCCGGGTTAC
pHTP4	CAATGGCACACTTGTTCCGGGTTAC
pHTP7	GAATCCGCAGGGTATGGATACCAGC
pHTP8	GTTCAAAAACGGTGAAGTGGCGGC
pHTP9	GAATGAAAAACGCGACCACATGGTG
pHTP10	GGCTGATATCGAAGGGTTGACCG
pHTP11	CTTGAAATCCAGCAAGTATATAGCATGG
pHTP13	GGAAAAAGTTTTCAAACAGTACGCTAAC
pHTP14	GCCCCGATTGACCATTTTCGTTTC
pHTP16	CCCACTTGCTGACGCTGTAGTAG
pHTP17	CATTCGTCATAGAAAAAGACCTGAAAG

Reverse primer (5'→ 3') common for all the pHTP expression vectors:
GGTTATGCTAGTTATTGCTCAGCG

Note: *After running on an agarose gel, the expected size of the insert amplified using the pHTP vector-specific primers will be incremented by extra 294 bp.*

pHTP vectors

Nucleotide sequence and properties of pHTP expression vectors are available for download at the product resources tab of the product.

Multiple fragment cloning protocol

NZYEasy Cloning & Expression System offers the possibility to clone multiple inserts simultaneously into one vector in a single reaction. Please read the Manual for Multiple Fragment Cloning using pHTP vectors available for download on the product page on our website.

Protein Expression & Purification

pHTP expression vectors are T7/lac promoter based-plasmids and can be used to transform competent *E. coli* cells expressing T7 RNA polymerase, such as BL21(DE3) cells. His-tagged recombinant proteins can be purified by immobilized metal-affinity chromatography (IMAC).

NZYEasy Cloning & Expression Systems

For more details, please read the NZYEasy Cloning & Expression System Manual available for download on the product page on our website.

Quality control assays

Purity
NZYEasy enzyme mix is >95% pure as judged by SDS polyacrylamide gel electrophoresis followed by BlueSafe (MB152) staining.

Nucleases assay
All components of the kits are tested for nucleases activities, using 0.2-0.3 µg of pNZY28 plasmid DNA. Following incubation at 37 °C for 14-16 hours, the DNA is visualized on a GreenSafe-stained agarose gel. There must be no visible nicking or cutting of the DNA.

Functional assay
All components of the kits are functionally tested in a ligase-independent cloning reaction, followed by a transformation assay. >90% of the recombinant plasmids must contain the appropriate insert.

V2001

Certificate of Analysis

Test	Result
Enzyme purity	Pass
Nucleases assay	Pass
Functional assay	Pass

Approved by:

Patrícia Ponte
Senior Manager, Quality Systems

For research use only.

nzytech
genes & enzymes
Estrada do Paço do Lumiar, Campus do Lumiar - Edifício E, R/C, 1649-038 Lisboa, Portugal Tel.: +351.213643514 Fax: +351.217151168
www.nzytech.com

Annex 23 – Speedy *Xho*I protocol

Speedy XhoI

Catalogue number:

MB10701, 200 reactions
MB10702, 1000 reactions

5'...C↓TCGAG...3'
3'...GAGCT↓C...5'

Description

NZYTech's Speedy restriction enzymes are a new generation of DNA modifying enzymes developed for rapid DNA digestion. All Speedy enzymes are 100% active in the NZYSpeedyBuffers and are able to digest DNA in 5-15 minutes. NZYSpeedyBuffers are universal and enable any combination of NZY Speedy restriction enzymes to work simultaneously. It's ideal for double or multiple digestions, eliminating the need for sequential digestions. NZYSpeedyBuffer Orange also allows ready gel load after digestion. Speedy XhoI activity is impaired by CpG methylation (CG methylase cleaves CT^{5m}CGAG very slowly). NZYTech's Speedy XhoI is a fast enzyme that performs its reaction usually in 5-15 minutes at 37 °C.

Concentration

One µL of enzyme can completely digest up to 1 µg of DNA in 5-15 min.

Reagents supplied with the enzyme

10× NZYSpeedyBuffer Colourless
10× NZYSpeedyBuffer Orange

Storage buffer

25 mM NaHepes, pH 7.5, 500 mM NaCl, 150 mM Imidazol, 2.5 mM $CaCl_2$, 50% (v/v) glycerol.

Recommended reaction conditions

Speedy XhoI	1 µL
10× NZYSpeedyBuffer	2 µL
Substrate DNA	≤ 1 µg
Sterilized ultrapure water	up to 20 µL

Incubation temperature

37 °C for 5-15 minutes.

Heat inactivation

65 °C for 20 minutes.

Quality control

Purity

Speedy XhoI has been determined to be >90% pure as judged by SDS-PAGE followed by BlueSafe staining (MB15201).

DNases assay

10 U of Speedy XhoI were incubated with 0.2-0.3 µg of pNZY28 for 14-16 hours at 37 °C. No nicking activity was observed following agarose gel electrophoresis.

Functional assay

Speedy XhoI was tested for performance in a digestion of 1 µg of a recombinant pNZY28 derivative using 1 µL of enzyme. The resulting digestion was visualized in an agarose gel.

Storage conditions

Speedy XhoI should be stored at -20 °C at all times. Do not keep it on ice while working at the bench.

Shipping conditions

Dry ice.

Product life

Please see the product labels.

V1901

Certificate of Analysis

Assay	Result
Protein purity	Pass
DNases assay	Pass
Functional assay	Pass

Approved by:

Patrícia Ponte
Senior Manager, Quality Systems

For research use only.

nzytech
genes & enzymes

Estrada do Paço do Lumiar, Campus do Lumiar - Edifício E, R/C, 1649-038 Lisboa, Portugal
Tel.: +351.213643514 Fax: +351.217151168
www.nzytech.com