2010
2011

GENEESKUNDE

master in de biomedische wetenschappen: klinische moleculaire wetenschappen

Masterproef

In vivo site-specific modification of proteins with artificial click functionalized amino acids

Promotor :
dr. Ties STEEN REDEKER
Prof. dr. Peter ADRIAENSENS

Copromotor :
Prof. dr. Wanda GUEDENS

Birgit Baré

Masterproef voorgedragen tot het bekomen van de graad van master in de biomedische wetenschappen , afstudeerrichting klinische moleculaire wetenschappen

Maastricht University

universiteit hasselt
UNIVERSITEIT VAN DE TOEKOMST

De transnationale Universiteit Limburg is een uniek samenwerkingsverband van twee universiteiten in twee landen: de Universiteit Hasselt en Maastricht University

universiteit hasselt
UNIVERSITEIT VAN DE TOEKOMST

Maastricht University

Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek
Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt

2010
2011

GENEESKUNDE

master in de biomedische wetenschappen: klinische moleculaire wetenschappen

Masterproef

In vivo site-specific modification of proteins with artificial click functionalized amino acids

Promotor :
dr. Ties STEEN REDEKER
Prof. dr. Peter ADRIAENSENS

Copromotor :
Prof. dr. Wanda GUEDENS

Birgit Baré

Masterproef voorgedragen tot het bekomen van de graad van master in de biomedische wetenschappen , afstudeerrichting klinische moleculaire wetenschappen

Maastricht University

universiteit hasselt
UNIVERSITEIT VAN DE TOEKOMST

Table of Contents

Acknowledgements

A very important chapter in my life is about to end. The past five years have gone by so fast. I still remember the first day at the university as if it was yesterday. And here I am, at a few steps from graduation. All this would not have been possible without the help and support of my promoter, colleagues, friends and family. Therefore, I would like to take this opportunity to thank all of these people.

In the first place, I would like to thank my promoter, Dr. Steen Redeker Erik, whose door was always open so I could drop in anytime with my questions and problems. Erik, even in the weekends or late in the evening, you made the effort to answer my mails and read the draft versions of my thesis. Thank you very much! I would also like to thank the two PhD students, Drs. Reekmans Gunter and Drs. Cortens David for their help. Gunter, no matter how busy you were with your own students, you were always ready to help me figuring out what went wrong when my experiments failed. David, thank you for reading the draft versions of my thesis and helping me in the lab. Also my second promoter, Prof. Dr. Adriaensens Peter, and my co-promoter, Prof. Dr. Guedens Wanda, deserve a word of appreciation for giving me the opportunity to do my internship at their lab, as well as my second examiner, Prof. Dr. Michiels Luc, for reading my thesis.

Next, I would like to thank Brecht for the pleasant times we had in the lab and all of my fellow students for all their support and friendship. Not everything was always a bed of roses during the past five years. Besides the pleasant moments, there were also times of frustration, disappointment and stress before the exams. At those moments, I could always go to them and pour my hart out.

And last but definitely not least, I would like to thank the people closest to me: my parents, my sister Evelien, my best friend Sophie and my boyfriend Dries. They always motivated me, believed in me and offered me a listening ear when I needed it most. Mama, papa, I would also like to thank you for giving me the financial support to complete my studies. I really appreciate all the things you do for me!

List of abbreviations

aaRS	aminoacyl tRNA synthetase
Abs	antibodies
AMP	adenosinemonofosfaat
ATP	adenosinetriphosphate
bp	base pairs
$CaCl_2$	calcium chloride
DMSO	dimethylsulfoxide
dNTP	deoxyribonucleotide triphosphate
E. coli	*Escherichia coli*
ELISA	enzyme-Linked Immuno Sorbent Assay
IPTG	isopropyl β-D-1-thiogalactopyranoside
kb	kilo base pairs
LB	lysogeny broth
Leu	leucine
LiAc	lithium acetate
$MgCl_2$	magnesium chloride
Nb	nanobody
ODCase	orotidine-5'-monophosphate decarboxylase
PCR	polymerase chain reaction
PP_i	pyrophosphate
rpm	revolutions per minute
SD-URA	synthetic defined broth minimal media, which contains all essential components except uracil
S. cerevisiae	*Saccharomyces cerevisiae*
ssDNA	single stranded DNA
TAE	buffer solution containing tris base, acetic acid and ethylenediaminetetraacetic acid
Tm	primer melting temperature
$tRNA_{CUA}$	amber suppressor tRNA
Trp	tryptophan
TyrRS	tyrosyl tRNA synthetase
VHH	antigen-binding fragment of the camel antibody
X-gal	bromo-chloro-indolyl-galactopyranoside
YPD	yeast peptone dextrose

Abstract

Numerous biotechnological applications require the **immobilization of proteins on a solid substrate**. At present, protein immobilization is often established by weak and/or non-specific interactions. A disadvantage is that these interactions or not stable and/or that this results in proteins immobilized with a random orientation on the solid substrate. A possible consequence is that the biological activity of the biofunctionalized material decreases. An interesting technique to obtain site-specific and oriented coupling of proteins is **'click' chemistry**, of which the azide/alkyne cycloaddition is a popular example. It is hypothesized that the sensitivity of a biosensor will significantly increase when the proteins are site-specifically immobilized. An improvement of sensor-sensitivity would be beneficial in several sectors, including healthcare, scientific research, the food industry the environmental sector.

In order for the azide/alkyne cycloaddition to occur, one of the 'click' functionalities, an alkyne or an azide, has to be introduced into the proteins. A promising strategy to incorporate a 'click' functionality into proteins is by creating a **'click' functionalized amino acid**, which is built into the proteins during their translation. In order to incorporate this artificial amino acid, the **genetic repertoire of the host organism must be expanded with a codon that uniquely specifies this artificial amino acid**. Because there are three different stop codons in the genetic code, it is possible to reassign one of these codons and add a 21st amino acid to the organisms genetic repertoire. In this project was chosen to reassign the **'amber' stop codon**. For this, an **orthogonal aaRS/tRNA pair** is needed.

The main goal of this Senior internship was **to amplify the genetically encoded *E. coli* amber suppressor TyrRS/tRNA$_{CUA}$ pair** using PCR techniques and **to clone the amplified genes in the appropriate vector**. In a later stage, a library of mutant TyrRSs will be created by site-directed mutagenesis, from which a synthetase will be selected that specifically recognizes the artificial 'click' functionalized amino acid instead of tyrosine. The selected TyrRS/tRNA$_{CUA}$ pair will then be used as an orthogonal pair in the yeast *S. cerevisiae* to introduce the 'click' functionality in response to the amber codon strategically put into the genetic code of the protein to be coupled. The target protein in this project is the BCII10 nanobody. A second goal of the practical training was therefore **to amplify the BCII10 nanobody** using PCR techniques and **clone it in the pTEF-MF yeast expression vector.** Both constructs were successfully created and will be used in the further steps of the modification of the nanobodies.

Samenvatting

Talloze biotechnologische applicaties vereisen de **immobilisatie van eiwitten op een substraatoppervlak**. Tot de dag van vandaag wordt deze immobilisatie vaak verwezenlijkt met behulp van zwakke en/of niet-specifieke interacties. Het nadeel hiervan is dat deze bindingen niet stabiel zijn en/of dat deze interacties resulteren in een willekeurige oriëntatie van het eiwit op het substraatoppervlak, met als gevolg een verminderde biologische activiteit van het substraat. De **'click' chemie**, een relatief nieuwe techniek waarvan de azide/alkyn cycloadditie een goed voorbeeld is, maakt het mogelijk om een georiënteerde koppeling te bekomen. De hypothese is dat eiwitten met behulp van 'click' chemie stabiel en met een optimale oriëntatie geïmmobiliseerd kunnen worden op het substraat, en dat wanneer deze immobilisatie toegepast wordt op biosensoren de gevoeligheid van de biosensoren significant zal toenemen. Deze gevoelige biosensoren zullen een meerwaarde zijn voor verschillende sectoren, gaande van de gezondheidszorg en wetenschappelijk onderzoek tot de voedselindustrie en de milieu sector.

Om een eiwit met behulp van de azide/alkyn cycloadditie te immobiliseren, moet er eerst een 'click' groep, een azide of een alkyn, in het eiwit geïntroduceerd worden. Om de 'click' groep in een eiwit te introduceren kan gebruik gemaakt worden van een **'click' gefunctionaliseerd aminozuur**, dat tijdens de translatie in het eiwit wordt geïncorporeerd. Hiervoor moet het **genetische repertoire van het gast organisme uitgebreid** worden **met een codon, specifiek voor dat aminozuur**. De genetische code bevat drie verschillende stop codons, wat het mogelijk maakt om één van deze codons toe te wijzen aan het artificieel aminozuur. In dit project werd gekozen voor het **'amber' codon**, dat in gisten het minst frequent gebruikt wordt als stopcodon. Om het amber codon toe te wijzen aan het 'click' aminozuur is er een bijhorend **orthogonaal aaRS/tRNA paar** nodig.

Een eerste doel van deze senior stage was het **amplificeren van het genetisch gecodeerde *E. coli* amber suppressor TyrRS/tRNA$_{CUA}$ paar** met behulp van PCR-technieken en het **kloneren van de geamplificeerde genen in de pESC-TRP vector**. In een later stadium zal een bibliotheek van mutante TyrRSs aangemaakt worden, waaruit een synthetase geselecteerd zal worden dat specifiek het 'click' aminozuur herkent in plaats van tyrosine. Het geselecteerde TyrRS/tRNA$_{CUA}$ paar zal vervolgens als orthogonaal paar gebruikt worden in de giststam *S. cerevisiae* om de 'click' groep in het eiwit te brengen op de plaats van het amber codon, dat op een welbepaalde plaats geïntroduceerd werd in het eiwit. Het doeleiwit is het BCII10 nanobody. Een tweede doel van de stage was de **amplificatie van het BCII10 nanobody-gen** met behulp van PCR-technieken en het **kloneren van het BCII10-gen in de pTEF-MF gist expressievector**. Beide constructen werden met succes aangemaakt en gebruikt in verdere stappen in de amplificatie van de nanobodies.

INTRODUCTION

Biosensors are gaining interest since recent technical advances allow the development of high-quality sensors that can be used directly by the physician or even by laymen. They allow measurements on much shorter time scales than traditional detection methods such as mass spectrometry, ELISA and chromatography. Current biosensors show high levels of selectivity and sensitivity, approaching those of the traditional methods. However, there is still room for improvement [1].

A biosensor consists of a biological recognition element that is connected with a signal transducer directly, or through a connective layer. Together, they communicate the presence of a specific analyte in the sample by sending a quantitative or semi-quantitative signal. In order to have an optimal detection of the target molecule, it is important that there is a highly efficient and stable coupling between the transducer and the recognition molecules that preserves the native state of the biomolecule and immobilizes the molecules with optimal orientation for the molecule-target interaction [2-4].

A biomolecule frequently used in bio-applications is the antibody. Mammalian antibodies (Abs) (Fig. 1) are relatively large and complex, Y-shaped molecules made up of two light and two heavy chains, linked together by intermolecular disulfide bonds. The light and heavy chains are subdivided into regions with variable and constant sequences. The light chains exist of one constant (CL) and one variable (VL) domain, whereas the heavy chain contains three constant (CH1, CH2 and CH3) and one variable (VH) domain. In terms of function, two molecular parts can be distinguished. A first part is the antigen-binding fragment or Fab. A second part, the Fc fragment, is responsible for other biological effector functions [5]. Immunoassays based on polyclonal and monoclonal antibodies have been around for more than 30 years and are still among the most important diagnostic tool used in clinical an research areas. Despite the antibody's numerous successes, it also has its limitations. Their large size and complexity make them hard to manipulate. Several attempts were made to create smaller antibodies, consisting for example solely of the Fab-fragment or the variable fragment. However, these constructs were often not stable, still relatively big or had a lowered affinity [1, 6] In the beginning of the 1990s, the discovery by Hamers-Casterman et al. of a unique class of antibodies circulating in the blood of camelidae, opened up new perspectives [7]. Apart from the conventional antibodies, a considerable fraction of heavy-chain only antibodies were found in the serum of camelidae (Fig. 1). These antibodies lack the CH1 domain of conventional antibodies and the complete light-chain. The antigen-binding fragment is contained in one single domain, the VHH. The VHH is the smallest available intact antigen-binding fragment with a molecular weight of approximately 15 kDa. Because of its small size (~2,5 nm diameter and ~4 nm height) it is also referred to as 'nanobody'. The nanobodies harbour the full antigen-binding capacity of the original heavy-chain antibody that has evolved to be fully functional in the absence of light chains [8, 9]. Nanobodies have several advantages compared to the antigen-binding entities of regular antibodies. For example, the nanobodies are only a tenth of the size of conventional antibodies, which gives them access to clefts, grooves and buried epitopes on the target surface. Secondly, nanobodies are resistant to

stringent conditions; they are able to endure a broad range of temperatures and pH values. Furthermore, a very important advantage is that, in contrast to the conventional antibodies, the camel antibodies can be efficiently produced in bacteria and in yeast. Because of its single-domain character coded by one single gene, no complicated assemblage of the different antigen-binding domains is needed [6, 10]. These qualities and others make nanobodies a good model protein for site-specific modification and immobilization on a solid substrate. A representative example of a nanobody is the BCII10 nanobody, which is active against the β-lactamase BCII. β-lactamases are enzymes produced by nosocomial bacteria to escape the lethal action of β-lactam antibiotics [11,12]. The BCII10 nanobody is among other things used in this project because of its high stability and because it is possible to modify its C-terminus without interfering with its structure [13].

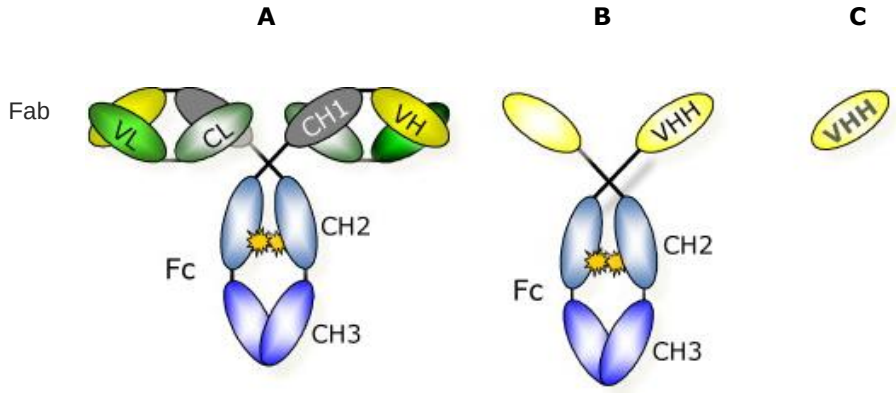

Fig. 1. Composition of a classical antibody (A), a heavy chain antibody (B) and a single domain antigen-binding entity derived from a heavy-chain antibody, the VHH or nanobody (C). The Fab or antigen-binding fragment in the heavy chain antibody is reduced to a single variable domain, the VHH, which harbours the full antigen-binding capacity [6].

In contrast to the progress that has been made in the immobilization of DNA structures, the immobilization of proteins is still in its infancy. This is explained by the heterogeneity of the protein and the close relation between its structure and function.
Generally, proteins are immobilized on a solid substrate via non-specific interactions. For example, non-covalent interactions such as adsorption onto hydrophobic surfaces and electrostatic interactions with charged surfaces are frequently used immobilization methods for the reason of their simplicity. However, non-covalent reactions are characterized by a low stability [14]. For more stable attachment, the formation of covalent bonds is required.

Covalent bonds can be realised by making use of the proteins' naturally available groups, which are coupled to a chemically activated surface. These can for example be lysine side chains that react with an amine-reactive surface [15-17]. Although this results in a strong covalent binding, the used groups are naturally occurring in the proteins and thus most of the time not unique. This means that it is not predictable which of these groups will bind the surface, and this random nature of

attachment can cause steric hindrance of the active site or denaturation of the protein due to strain from multiple attachment sites [18].

A promising method to overcome these problems and obtain site-specific coupling with a well-defined orientation of the protein, is to modify the proteins with a bioorthogonal group at a specific, well-defined position with 'new' chemistry. A popular example of a site-specific covalent reaction is the azide/alkyne cycloaddition, which is categorized as 'click' chemistry (Fig. 2) [19]. 'Click' chemistry is defined as a group of reactions that are easy to perform, are efficient and stereospecific, have readily available starting materials and reagents, have very high yields and require no or minimal purification. In the presence of the catalyst copper (I), the azide and the alkyne interact and form a triazole. Using the azide/alkyne cycloaddition has two important advantages: firstly, both alkyne and azide are bioorthogonal, which means that they do not occur naturally in proteins. This guarantees that the 'click' group is unique in the protein and consequently the only group capable of reacting with the surface. Secondly, the reaction can be accomplished in physiological conditions, making it ideal for the coupling of proteins in aqueous systems, for example for the use of a biosensor as an implant [20-22]. In order to immobilize proteins to a solid substrate using this 'click' reaction, either an alkyne or an azide has to be introduced into the protein.

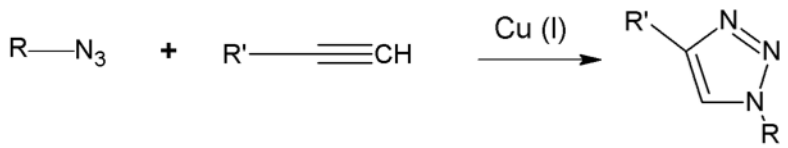

Fig. 2. Copper (I) catalyzed azide-alkyne cycloaddition (CuAAC)

A possible method to introduce one of the 'click' functionalities into the protein is via post-translational modification. Ideally, the group addressed by post-translational modification is unique in the protein and located at an optimal position for immobilization, because it is difficult to direct the modification to one single, specific group. As a consequence, this technique is not applicable for most proteins. A recent developed technique, the intein system, offers a solution to this specificity problem. Inteins are internal segments of precursor proteins that are able to excise themselves and ligate the two flanking regions through a native peptide bond, by the formation of an intermediate thioester [23]. A number of mutant inteins have been designed that can only complete the splicing reaction at one site of the intein and in the presence of thiols [24-25]. When the protein is cloned in a vector containing such an intein, a reactive C- or N-terminal thioester is created, depending on the type of intein. Via that thioester, the protein can then be directly coupled to a cysteine-functionalyzed bioorthogonal group. This system allows modifying the protein very specifically at a well-defined position, the C- or the N-terminus. However, when the active site of the protein is located nearby the C- or N-terminus, this is not the ultimate technique.

An innovative strategy to introduce a unique bioorthogonal group into proteins that overcomes the previous mentioned problems, is to make an artificial amino acid, for example a 'click' functionalized amino acid, which is built into the protein during the translation process [26]. The advantage of introducing a functional group using this technique is that the location of the group can be chosen very accurately at the level of the genetic code, enabling the selection of a position

where the introduced group will not interfere with the protein's structure or activity. This way, the protein can be site-specifically immobilized to the solid substrate with an optimal orientation.

The major objective of our research group is to establish site-specific immobilization of nanobodies on a solid substrate, using the azide/alkyne cycloaddition. In order to incorporate the artificial 'click' functionalized amino acid, the genetic repertoire of the host organism should be expanded with a codon that uniquely specifies this artificial amino acid, and a matching aminoacyl-tRNA synthetase (aaRS)/tRNA pair [26-28].

The genetic code is degenerate. Twenty amino acids are encoded by sixty-one codons, including one start codon. The remaining three codons serve as stop signals. This means that there is redundancy of several codons, making it possible to reassign for example the amber stop codon (TAG) to an artificial amino acid. The process of introducing an amino acid in response to the nonsense amber codon is called 'amber suppression'. Some organisms, such as *E. coli,* have a $tRNA_{CUA}$ that recognizes amber mutations and inserts a tyrosine in response to the mutation to prevent premature ending of translation. By changing the amino acid preference of the *E. coli* $tRNA_{CUA}$ corresponding aaRS to an artificial amino acid, the aaRS/$tRNA_{CUA}$ pair can be used to incorporate the artificial amino acid in response to an amber codon. When this amber codon is strategically put into the sequence of the target protein, site-specific modification will be possible [26-28]. A schematic representation of the incorporation of an artificial amino acid is shown in Fig. 3.

The expression of the modified proteins will be carried out in the yeast strain *Saccharomyces cerevisiae* MaV203 (Invitrogen). *S. cerevisiae* is an ideal organism for the production of proteins because they have high yield and they are easy to use. Since the translational machinery is well conserved in eukaryotes, it is believed that the whole system can be copied to higher eukaryotic organisms for the modification of more complex proteins [29-30]. The yeast strain MaV203 contains deletions in the endogenous GAL4 and GAL80 genes. It also contains three GAL4-inducible reporter genes, more specifically lacZ, HIS3 and URA3. MaV203 is also auxotrophic for leucine (Leu) and tryptophan (Trp), making this strain ideal for the construction and screening of libraries [31].
It is known from literature that the *E. coli* tyrosyl-tRNA synthetase (EcTyrRS)/$tRNA_{CUA}$ pair can be used as an orthogonal pair in the yeast *S. cerevisiae*. The EcTyrRS/$tRNA_{CUA}$ pair suppresses amber codons with tyrosine, the *E. coli* $tRNA_{CUA}$ is not charged by endogenous aminoacyl tRNA synthetases in the yeast cytosol nor does the EcTyrRS charge any yeast tRNAs [32-33]. In order to use the EcTyrRS/$tRNA_{CUA}$ pair as an orthogonal pair in *S. cerevisiae* for introducing an artificial amino acid, the amino acid specificity of EcTyrRS must be redirected from tyrosine to an artificial 'click' functionalized amino acid. In order to alter the amino acid specificity, a library of mutant EcTyrRSs will be constructed. Since synthetases show a strong sequence similarity, previous studies indicated five amino acids of the EcTyrRS, based on the crystal structure of the homologous synthetase from *Bacillus Stearothermophilus*, more specific Tyr37, Asn126, Asp182, Phe183 and Leu186, to be important for the specific amino acid binding [32].

In this project, based on the structure of *E. coli* TyrRS, five amino acids that are important for the binding of tyrosine will be defined. These amino acids will be the targets for site directed mutagenesis to create a mutant library that will be screened for 'click' amino acid selectivity. The starting products of the reaction include a double-stranded DNA template, in our case a pEcTyrRS/$tRNA_{CUA}$ vector construct containing the EcTyrRS/$tRNA_{CUA}$ pair and ADH1 promoter (Fig. 4), and the degenerate primers containing the desired mutations. The primers are designed in a way that every possible amino acid could be introduced at each of the selected sites important for the binding of the amino acid. In this way, a mutant collection containing 20 (amino acids) raised to the power of 5 (sites of mutation) or $3{,}2x10^6$ different mutants will be created. All the primers are designed to bind the same strand of the template DNA. Otherwise, there would be preferential binding of complementary strands of mutagenic primers to each other, limiting the representation of mutants in the library.

In order to select a mutant synthetase that incorporates the artificial amino acid with high fidelity, the library will be screened by both positive and negative selection methods. Apart from the pEcTyrRS/$tRNA_{CUA}$ vector construct (Fig. 4A), also a pGAL4/ADH1 vector construct (Fig. 4B) containing the ADH1 promoter and the GAL4 gene containing two amber mutations will be cloned into the yeast cells, which is necessary for the screening of the library (Fig. 4). The pGAL4/ADH1 vector has a Leu selection marker. After the screening of the library, leucine will be added to the medium. In response, the yeast cells will eject he pGAL4/ADH1 vector. For a first, positive selection, the cells will be allowed to grow in a selective medium containing the artificial amino acid and lacking uracil. Clones expressing an EcTyrRS that recognizes either the artificial amino acid or the natural tyrosine will suppress the amber codon and produce GAL4. GAL4 will enable the activation of URA3, a gene coding for the enzyme ODCase, which is necessary for the production of uracil. The production of uracil will lead to cell survival. In a second step, there will be selected against the EcTyrRS clones that introduce the natural amino acid tyrosine in response to the amber codon, leaving only the wanted clones that recognize the artificial amino acid. For this negative selection, the cells that survived the first selection step will be allowed to grow in a medium containing uracil and 5-Fluoroorotic acid (5-FOA), in the absence of the artificial amino acid. EcTyrRS clones that are able to introduce natural amino acids in response to the amber mutation in GAL4, will enable the activation of URA3. The enzyme ODCase encoded for by URA3, converts 5-FOA into a toxic substance, causing death of the cells. Both positive and negative selection steps will be repeated a couple of times, leaving only the clones of EcTyrRS that introduce the 'click' functionalized amino acid in response to an amber codon. After the screening of the library, the BCII10 nanobody modified with an amber mutation, cloned in the pTEF-MF/BCII10 expression vector (Fig. 4C), will be introduced in the yeast cells that survived the selection steps. The pTEF-MF expression vector contains a strong TEF1 promoter, which drives constitutive expression of the nanobody, cloned in the multiple cloning site of the vector, and a leader sequence of mating factor alpha to ensure secretion of the protein into the medium. At the C-terminus, the nanobody sequence is followed by a polyhistine-tag. This tag is an amino acid motif of six histidine residues, which will be attached to the protein upon expression. The His-tag has an affinity for metals, which makes it possible to purify the expressed proteins on an Immobilized Metal ion Affinity Chromatography (IMAC) column.

The primary goal of this senior practical training was to create the vector constructs necessary for the construction and screening of the library. For the first construct, the *E. coli* $tRNA_{CUA}$ and TyrRS and the yeast ADH1 promoter were amplified. The amplified genes were then cloned, one by one, in the pESC-TRP vector, resulting in the pEcTyrRS/$tRNA_{CUA}$ vector (Fig. 4A). The second construct that was made, was the pTEF-MF/BCII10 (Fig. 4C). The BCII10 nanobody was amplified from the pHEN6 vector and cloned in the pTEF-MF yeast expression vector.

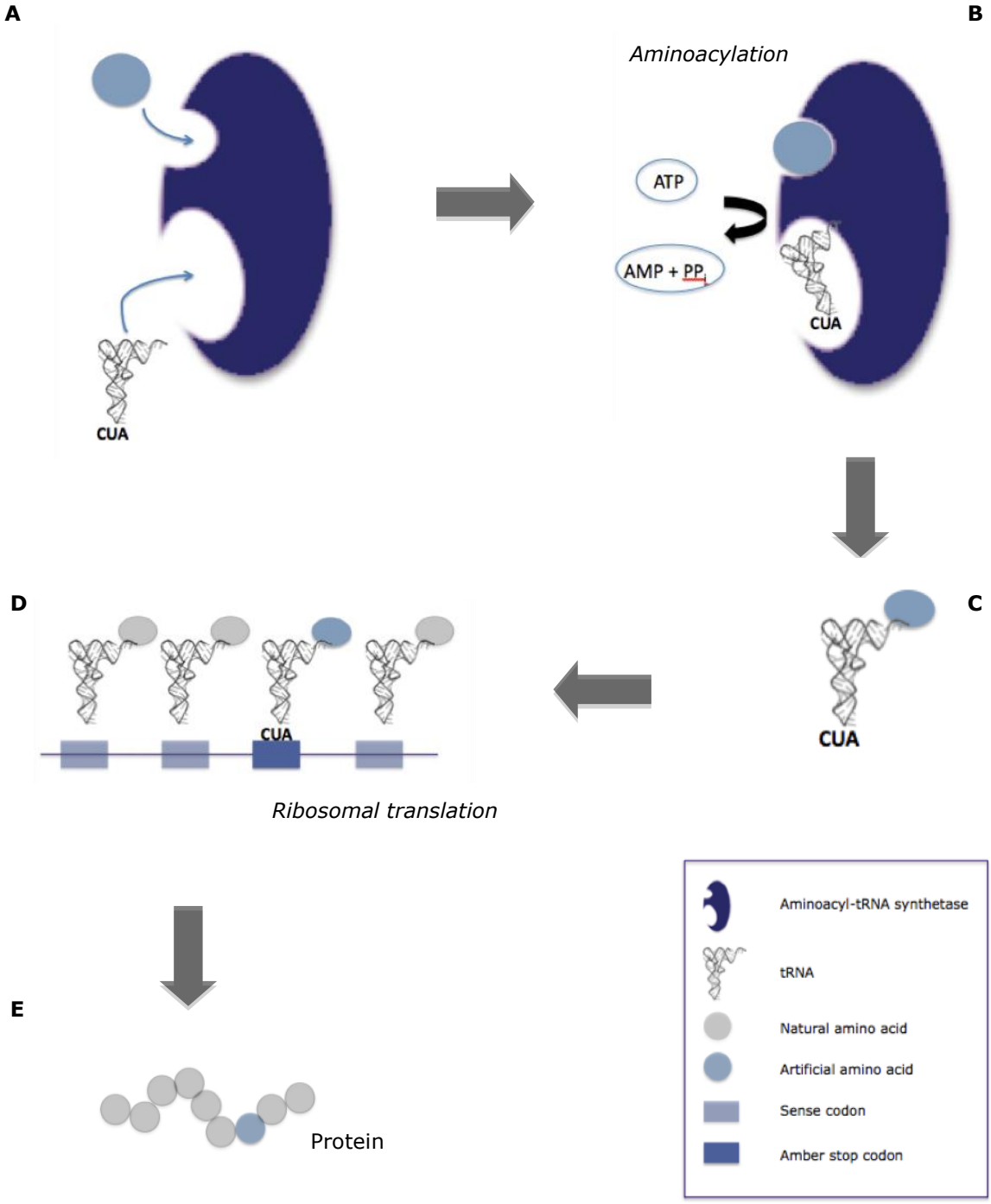

Fig. 3. Charging of tRNA molecules by the enzyme aminoacyl-tRNA synthetase and incorporation of an unnatural amino acid in a protein during ribosomal translation in response to the amber stop codon.
The enzyme aminoacyl-tRNA synthetase (aaRS) recognizes the amino acid and corresponding tRNA. The mutated *E. coli* tyrosyl-tRNA synthetase that recognizes the artificial amino acid and the *E. coli* $tRNA_{CUA}$ was chosen as an example (A). The enzyme catalyzes the coupling of the aa to the tRNA via the hydrolysis of ATP (B), resulting in an aa-tRNA (C). This process is called aminoacylation. During protein translation, the tRNA anti-codon pairs with its complementary codon in the DNA sequence and amino acids are added to the growing polypeptide chain one by one. The artificial amino acid is added in response to the amber codon, strategically introduced into the sequence of the protein (D,E). Figure adapted from [35].

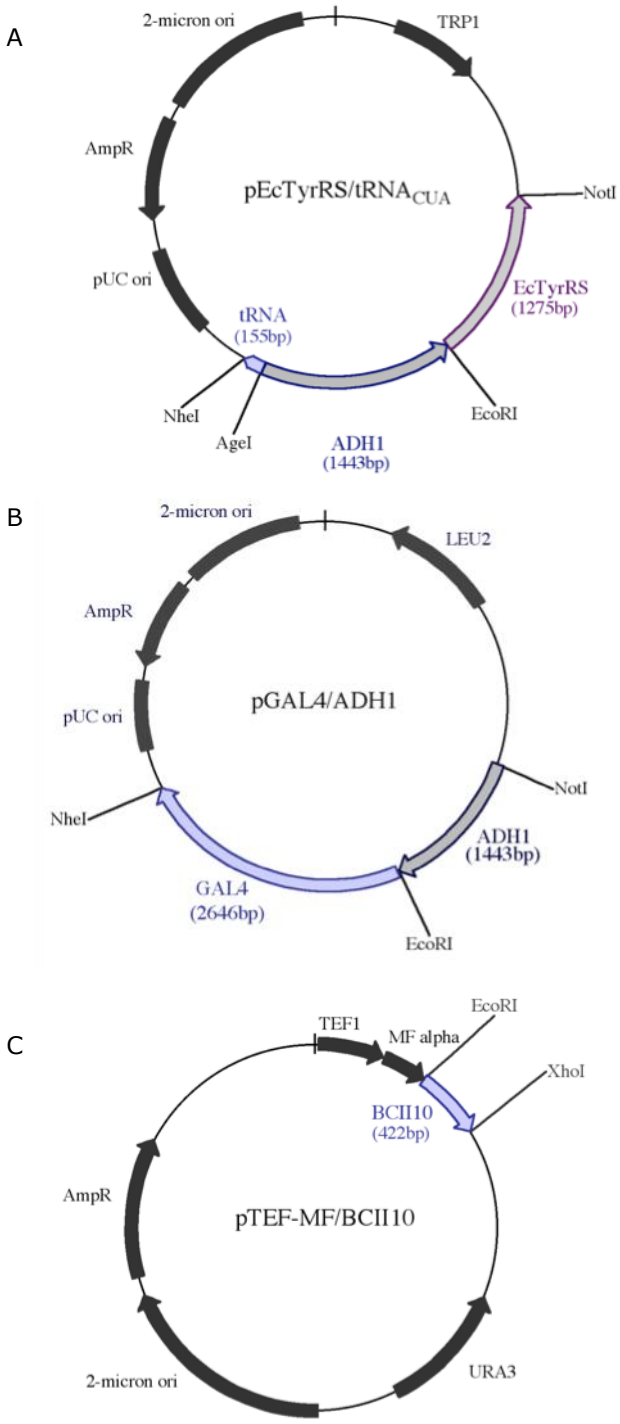

Fig. 4. Vector constructs. A) The EcTyrRS/tRNA$_{CUA}$ pair and an ADH1 promoter will be cloned in the plasmid pESC-TRP, for the formation of the pEcTyrRS/tRNA$_{CUA}$. Starting from this plasmid, a library of modified *E. coli* TyrRSs will be created, using the QuikChange Lightning Site-Directed Mutagenesis Kit (Stratagene). B) For the GAL4-based screening of the library, GAL4 containing an amber mutation and ADH1 will be cloned in the plasmid pESC-Leu, for the formation of pGAL4/ADH1. C) The BCII10 nanobody will be cloned in the plasmid pTEF-MF, which is a yeast expression vector. The resulting pTEF-MF/BCII10 vector will be transformed in the yeast cells containing the *E. coli* TyrRS selected from the library. The orthogonal EcTyrRS/tRNA$_{CUA}$ pair will then be used for the expression of the 'click' functionalized amino acid.

METHODS

1 Media, cells and plasmids

$tRNA_{CUA}$ and TyrRS were amplified from *E. coli* BL21 (DE3) pLys. PCR was performed on a single colony. This strain contains a chloramphenicol resistance gene. Cells were plated on Lysogeny Broth (LB) agar plates containing chloramphenicol (34 µg/ml, final concentration) (Supplement I). The full length ADH1 promoter was amplified from the plasmid pMA424 (kindly provided by Prof. Dr. Stanley Fields, Howard Hughes Medical Institute, USA). The BCII10 nanobody was amplified from the pHen6 vector (kindly provided by Prof. Dr. Serge Muyldermans, VIB/VUB).

All genes amplified by PCR were cloned in a pCR2.1 high-copy vector using the commercial TA Cloning Kit (Invitrogen) and transformed in *E. coli* Top10F' chemically competent cells. The pCR2.1 vector has an ampicillin resistance gene and the F' episome of the Top10 F' cells contains a tetracycline resistance gene. The transformed Top10F' cells were plated on LB agar plates containing both ampicillin (100µg/ml final concentration) and tetracycline (12,5 µg/ml final concentration). 40µl of both X-gal (40mg/ml) and IPTG (100mM) were added to the plates in order to select successful ligations with blue-white screening. The pCR2.1 vector contains the α fragment of the *lac*Z gene, which is part of the *lac* operon and responsible for the expression of β-galactosidase. The F' episome of the Top10F' cells, carries the $lacI^q$ repressor, which inhibits the production of β-galactosidase in the absence of lactose. The addition of IPTG, a lactose metabolite, induces the activation of *lac*Z and subsequent expression of β-galactosidase. β-galactosidase cleaves the colourless X-gal, yielding galactose and 5-bromo-4-chloro-3-hydroxyindole. The latter is then oxidized into 5,5'-dibromo-4,4'-dichloro-indigo, an insoluble blue product. However, the multiple cloning site of the vector is located in the middle of the *lac*Zα gene. When an insert is present in the multiple cloning site of the pCR2.1 vector, disturbing the *lac*Zα gene, the expression of β-galactosidase is inhibited and as a consequence, X-gal is not cleaved. Colonies containing an insert will appear white, whereas colonies without the insert will appear blue. Liquid cultures were made in LB-medium containing ampicillin (100µg/ml final concentration) and tetracycline (12,5 µg/ml final concentration).

2 Polymerase Chain Reaction

E. coli $tRNA_{CUA}$ and TyrRS, the ADH1 promoter and the NbBCII10 were amplified by PCR. Each reaction mixture contained 5µl of 10x PCR buffer (Invitrogen), 4µl of dNTPs (2,5mM each, TaKaRa), 1µl forward primer (10µM, Eurogentec), 1µl of the reverse primer (10µM, Eurogentec), 0,25µl Taq polymerase (Fermentas, 5U/µl), $MgCl_2$, DNA and MilliQ to a final volume of 50µl. The primers were designed to introduce unique restriction sites before and after the gene, which would be used later for the cloning. The primers used for the different reactions are displayed in table 1. The concentration of $MgCl_2$ varied for the different reactions. When an *E. coli* BL21 (DE3) pLys colony was used as DNA source, 5µl of a colony dissolved in 50µl MilliQ was added to the reaction mixture. Plasmid DNA was diluted in 100µl MilliQ, of which 1µl was added to the reaction mixture. The DNA source used for the amplification of the different genes is displayed in table 1. Two

negative controls were included. A first negative control was the reaction mixture without DNA. A second negative control was the DNA with 10x PCR buffer and MilliQ.

In the first step, denaturation of the DNA took place at 95°C for 5 minutes. Next, a cycle of denaturation (30" at 95°C) annealing (30") and elongation (72°C) was repeated 30 times. The reaction was ended with a last elongation step of 5'. The annealing temperature was chosen based on the melting temperature (Tm) of the primers (Table 1). The elongation time was based on the length of the fragment to be amplified. All reactions were carried out in the Mycycler Thermal Cycler (Biorad). The exact reaction mixtures and PCR-programs of the individual reactions are included in the supplementary information (Supplement 2). The sequences of the amplified genes are included in Supplement 3.

The amplified DNA was visualized by agarose gel electrophoresis. An agarose gel (UltraPure™, Invitrogen) was prepared with 1xTAE buffer (40mM Tris, 20mM acetic acid, 1mM EDTA, pH 8,4) and 5 µl GelRed™ (Biotium) of a 10000X solution in DMSO was added for the visualization of the DNA fragments with UV light. The percentage of agarose was based on the length of the DNA fragment. Sample buffer 6x (30% glycerol, 0,25% bromophenol blue, MilliQ) was added to the samples before they were loaded on the agarose gel. After running the gels for 60 minutes at 120 V in 1x TAE buffer, the DNA fragments were visualised with UV-light.

Table 1. Information on the construction of the pEcTyrRS/tRNA$_{CUA}$ and pTEF-MF/BCII10 vectors. The different genes that were cloned in the destination vectors are presented, together with the DNA source from which they were amplified, the primers used for their amplification with their melting temperatures and the restriction sites introduced into the amplified genes (indicated in bold in the primer sequences).

DESTINATION VECTOR	GENE	DNA SOURCE	PRIMER SEQUENCE		TM (°C)	RESTRICTION SITE
pEcTyrRS/tRNA$_{CUA}$	tRNA$_{CUA}$	E. coli DE3 pLys colony	Forward	5' GGGGGG**ACCGGT**AAGCTTCCCGATAAGGGAGCAG GCCAGTAAAAGCATTACCCCGTGGTGGGGTTCCCGA 3'	74,9	AgeI
			Reverse	5' GGCGGC**GCTAGC**GGGAAGTTCAGGGACT TTTGAAAAAAATGGTGGTGGGGGAAGGAT 3'	70,4	NheI
	TyrRS	E. coli DE3 pLys colony	Forward	5' TCATAACGA**GAATTC**ATGGCAAGCAGTAACTTG 3'	55,6	EcoRI
			Reverse	5' TTACTACGT**GCGGCCGC**T TATTTCCAGCAAATCAGAC 3'	61,6	NotI
	ADH1 promoter	pMA424 vector	Forward	5' GGGGGG**ACCGGT**CGGGATCGAAGAAA TGATGGTAAATGAAATAGGAAATCAAGG 3'	67,1	AgeI
			Reverse	5' GGGGGG**GAATTC**AGTTGATTGTATGCTTGG TATAGCTTGAAATATTGTGCAGAAAAAGAAAC 3'	64,9	EcoRI
pTEF-MF/BCII10	NbBCII10	pHEN6	Forward	5' GGC**GAATTC**ATGGCCCAGGTGC 3'	55,3	EcoRI
			Reverse	5' GGG**CTCGAG**TTAGTGATGGTGATGGTG 3'	57,7	XhoI

3 Preparation of plasmid DNA

3.1 TA cloning

The genes amplified by PCR were cloned in the high-copy vector pCR2.1 using the commercial TA Cloning Kit, according to the manufacturer's protocol (Invitrogen). The TA Cloning Kit provides a quick, one-step cloning strategy for the direct insertion of a PCR product into a plasmid vector. Taq polymerase adds a single deoxyadenosine (A) to the 3' ends of the PCR products during the amplification. The linearized vector supplied in the kit has single 3' deoxythymidine (T) residues. This allows PCR inserts to ligate efficiently with the vector.

3.2 Transformation in Top10 F' chemically competent *E. coli* cells

3.2.1 Preparation of Top10 F' chemically competent E. coli cells using $CaCl_2$

Top10F' cells from a glycerol stock were plated on LB-plates and incubated overnight at 37°C. A colony of the overnight plate was inoculated in 3ml LB-medium. The cells were allowed to grow overnight in an orbital shaker, set at 220 rpm and 37°C. 1ml from the overnight culture was transferred into 100ml LB-medium and grown to a final OD_{600} value of 0,4. The culture was divided in two falcon tubes (50ml), placed on ice for 20 minutes and the cells were harvested by centrifuging at 3000xg and 4°C for 10 minutes. The cells were kept on ice for the rest of the procedure. The pellet was resuspended in 30ml of cold $CaCl_2$ (0,1M) containing 15% glycerol. The competent cells were aliquoted and stored at -70°C until further use.

3.2.2 Transformation

8µl of the ligation product resulting from the TA cloning was added to 100µl of the competent cells and incubated on ice for 45 minutes. A first heat-shock was given to the cells by placing the reaction mixture in a water bath of 42°C for 90 seconds. The cells were shortly chilled on ice and subsequently the second heat-shock was given by the addition of 900µl of LB-medium, preheated at 37°C. The transformed Top10F' cells were allowed to grow in an ortbital shaker, set at 200 rpm and 37°C. After 1 hour, the cells were harvested by centrifuging at 3000xg for 5 minutes, resuspended in 50µl of the supernatant and plated on preheated LB-plates. The cells were allowed to grow overnight at 37°C.

3.3 Isolation of plasmid DNA

An overnight culture of the transformed Top10F' cells was started by inoculating 3ml LB-medium with a single white colony picked from the overnight agar plate and incubated overnight in an orbital shaker set at 220 rpm and 37°C. The plasmid DNA was isolated from the culture using the commercial QIAprep Spin Miniprep Kit according to the manufacturer's protocol.

3.4 Restriction analysis

The restriction reaction mixture with approximately 0,5µg plasmid DNA (2µl), 2µl 10X buffer (Fermentas), 1µl EcoRI (Fermentas) and 15µl water, was incubated in a water bath set at 37°C, for 1 hour. The restriction digestion was visualized by agarose gel electrophoresis as described earlier in 2 (Polymerase Chain Reaction). Plasmids that showed a restriction fragment of the expected length were sent for sequencing (Macrogen).

4 Site-Directed Mutagenesis of the tRNA anticodon

A site-directed mutagenesis was carried out using the QuickChange Lightning Site-Directed Mutagenesis kit (Stratagene, Agilent Technologies) according to the manufacturer's protocol. This kit enables the randomization of key amino acids in one single thermal cycling reaction by the use of degenerate primers. The pCR2.1 vector with tyrosyl-tRNA insert was used as template DNA. The forward and reverse primers are displayed in Table 2.

Table 2. Forward and reverse primers used for the site-directed mutagenesis of tRNA

Forward primer	5' CCAAAGGGAGCAGACTCTAAATCTGCCGTCATC 3'	Tm: 60,6°C
Reverse primer	5' GATGACGGCAGATTTAGAGTCTGCTCCCTTTGG 3'	Tm: 60,6°C

5 Construction of the pEcTyrRS/tRNA$_{CUA}$ and pTEF-MF/BCII10 vectors

5.1 Ethanol precipitation of tRNA and pEcTyrRS minipreps

The pCR2.1 vector containing tRNA and the pESC-TRP vector were transformed in *E. coli* Top10 F' cells by adding 1μl (±0,3μg) of the miniprep to 100μl competent cells. The transformation was carried out as described earlier in *3.2.2* and new minipreps were made as described earlier in 3.3. Of both tRNA and pESC-TRP, 4 minipreps were added together. 300mM NaAc, 10mM $MgCl_2$ and two volumes of ice-cold 100% ethanol were added to the DNA and incubated at -70°C for 1 hour. The precipiptated DNA was harvested by centrifuging 15' at 16000xg and 4°C. The pellet was washed in 500μl ice-cold 100% ethanol and the DNA was harvested again by centrifuging 15' at 16000xg and 4°C. After the DNA was air-dried, it was resuspended in 20μl MQ. The concentrations of both tRNA$_{CUA}$ and the plasmid pESC-TRP after ethanol precipitation were determined with a NanoDrop spectrophotometer (NanoDrop ® ND-1000).

5.2 Restriction digestion of pCR2.1/insert and the destination vector

The *E. coli* tRNA$_{CUA}$ and TyrRS, the *S. cerevisiae* ADH1 promoter and the NbBCII10 genes were excised from the pCR2.1 vector by restriction digestion, as described earlier for the restriction analysis after cloning in the pCR2.1 vector (3.4). The destination vector was opened using the same enzymes. The different restriction enzymes that were used to isolate the inserts and to open the vector are displayed in Table 1. The detailed composition of the reaction mixtures is included in Supplement IV.

5.3 Purification of inserts and destination vectors

After the restriction digestion, the inserts and the open vectors were separated by agarose gel electrophoresis. After the electrophoresis (60', 200V), the DNA fragments were excised from the agarose gel with a scalpel. Subsequently, the DNA was purified using the QIAquick Gel Extraction Kit according to the manufacturer's protocol. The DNA was eluted in 50μl water.

5.4 Cloning of the inserts in the destination vector

A reaction mixture was made with 10x T4 Ligase Buffer (Fermentas), T4 ligase (5u/μl, Fermentas), 20-100ng linear vector DNA, insert DNA and water. Different molar ratio's of insert:vector were

tested. The exact composition of the individual reaction mixtures is included in Supplement V. After vortexing, these mixtures were incubated overnight at 22°C.

5.5 Transformation and DNA isolation

The resulting pESC-constructs were transformed in Top10F' competent cells. The cells were allowed to grow overnight and subsequently, the plasmid DNA was isolated using the commercial QIAprep Spin Miniprep Kit according to the manufacturer's protocol (Qiagen). The presence of the insert was verified by restriction analysis and by sequencing, as described earlier for the pCR2.1 constructs (*3.2.2,* 3.3 and 3.4*)*.

RESULTS AND DISCUSSION

I. Construction of the pEcTyrRS/$tRNA_{CUA}$ vector

1 Amplification and cloning of *E. coli* TyrRS

1.1 Optimization of the PCR

In order to amplify the *E. coli* TyrRS, a PCR was performed on a single colony of *E. coli* BL21 (DE3) pLys. To obtain optimal PCR conditions, a gradient was applied for both the annealing temperature (35-55°C) and the $MgCl_2$ concentration (2, 4, 6 and 8mM). The PCR reaction mixtures were analysed by agarose gel electrophoresis, from which could be concluded that 2mM was the optimal $MgCl_2$ concentration (results not shown). The annealing temperature was further optimized in a second gradient PCR, with the annealing temperature ranging from 45-70°C. After the PCR, the reaction mixtures were analysed by agarose gel electrophoresis. *E. coli* TyrRS has a length of 1275bp. With the extra nucleotides added by the primers and the extra adenosine added by Taq polymerase included, a fragment of 1308bp was expected. Primer annealing appeared to be the most efficient at 54,3°C and 49,9°C, as indicated by the two clear bands at the expected height on the agarose gel (Fig. 5, respectively lanes 5 and 6). The reaction mixture from lane 6 (primer annealing temperature 49,9°C) was used in the TA cloning reaction. The resulting pCR2.1/EcTyrRS vector was transformed in Top10 F' chemically competent *E. coli* cells.

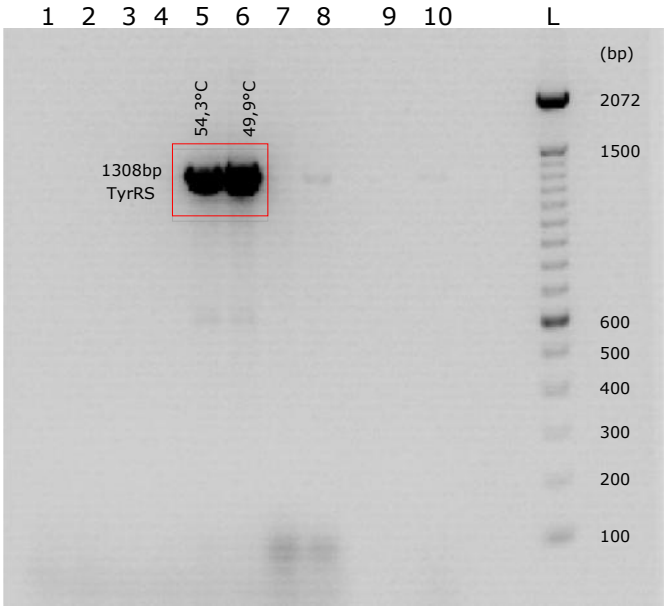

Fig. 5. Amplification of TyrRS from *E. coli* BL21 (DE3) pLys. A temperature gradient was applied to optimize the PCR. 1=70,0°C; 2=68,1°C; 3=65,0°C; 4=60,3°C; 5=54,3°C; 6=49,9°C; 7=46,8°C; 8=45,0 °C; 9= negative control I (mastermix without DNA); 10= negative control II (mastermix without primers and Taq polymerase); L=100bp ladder (Invitrogen). 49,9°C appeared to be the optimal annealing temperature for the amplification of *E. coli* TyrRS. (1% agarose)

1.2 Restriction analyses and sequencing

The *E. coli* cells, transformed with pCR2.1/EcTyrRS, were allowed to grow overnight. The next day, the plasmid DNA was isolated and the insert was analysed by restriction digestion. The pCR2.1

vector contains two EcoRI restriction sites, flanking the insert. After restriction digestion with restriction enzyme EcoRI, two fragments were expected for clones containing the *E. coli* TyrRS: one fragment of approximately 3,9kb (pCR2.1) and another fragment of 1323bp (TyrRS). In total, 10 clones were tested; clones 3, 4, 5, 7, 8 and 9 showed fragments of the expected length (Fig 6). In order to verify these results, the clones were sequenced and the resulting sequences were aligned with the sequence of *E. coli* TyrRS (Fig. 7). Clones 3, 8 and 9 contained a few point mutations. Therefore, clones 4, 5 and 7 were used in further cloning reactions.

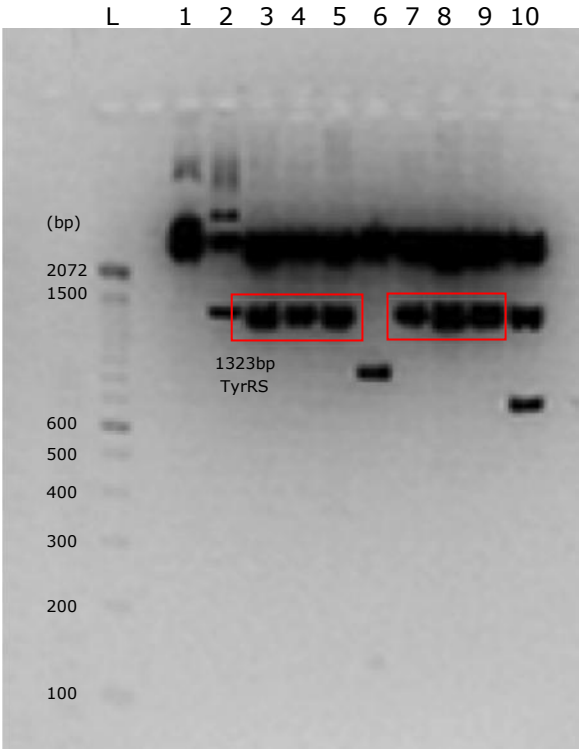

Fig. 6. Restriction digestion of pCR2.1/EcTyrRS with restriction enzyme EcoRI. 1-10= clones 1-10; L= 100bp ladder (Invitrogen). Fragments of the expected length are present in lanes 3, 4, 5, 7, 8 and 9. (1% agarose)

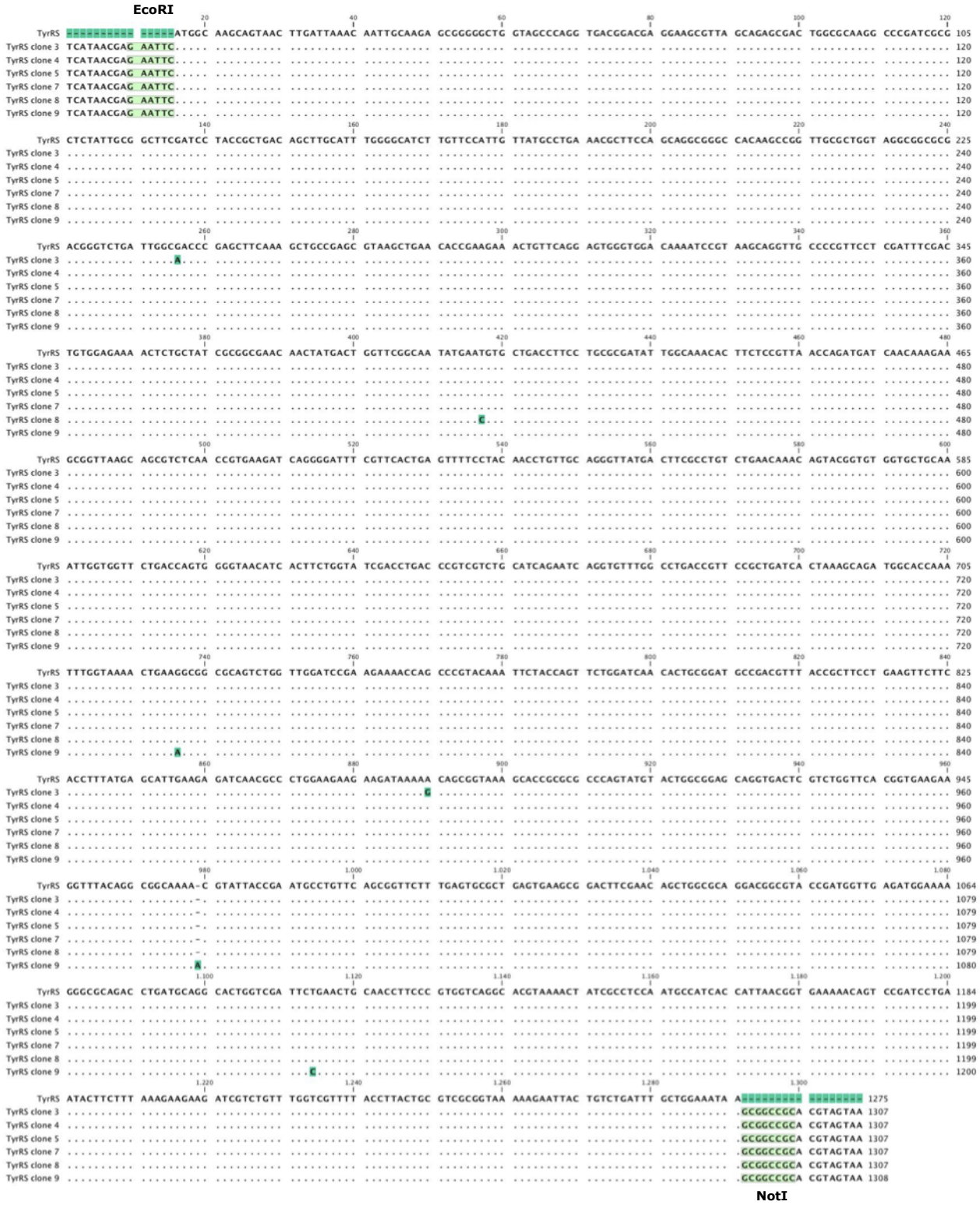

Fig. 7. Alignment of *E.coli* BL21 TyrRS sequence with the sequences of the different TyrRS clones, amplified by PCR. Clones 3, 8 and 9 contain one or more point mutations. Clones 4, 5 and 7 show complete sequence similarity and were used for cloning in the pESC-TRP vector. Mutations are indicated in blue, the restriction sites (EcoRI and NotI) are indicated in green.

1.3 Cloning in the pESC-TRP vector

In order to create the pEcTyrRS/tRNA$_{CUA}$ construct, clone 7 of the amplified *E. coli* TyrRS was cloned in the plasmid pESC-TRP. The *E. coli* TyrRS was first excised from the pCR2.1 vector by

restriction digestion with restriction enzymes EcoRI and NotI. The pESC-TRP vector was cut with the same enzymes. Restriction digestion of pESC-TRP should generate two fragments: a first fragment of approximately 6,5kb (representing the open vector) and a second fragment of approximately 22bp, which would not be visible on an agarose gel. Digestion of pCR2.1/EcTyrRS should result in a fragment of 1323bp (the TyrRS insert) and a fragment of approximately 3.9kb (the pCR2.1 vector). After restriction digestion, the reaction mixtures were analysed by agarose gel electrophoresis. Bands appeared at the expected height on the agarose gel (Fig. 8), indicating that the restriction digestion was successful. The TyrRS and the open pESC-TRP vector were excised from the agarose gel and after purification of the DNA, the TyrRS gene was cloned into the pESC-TRP vector, resulting in the pEcTyrRS vector.

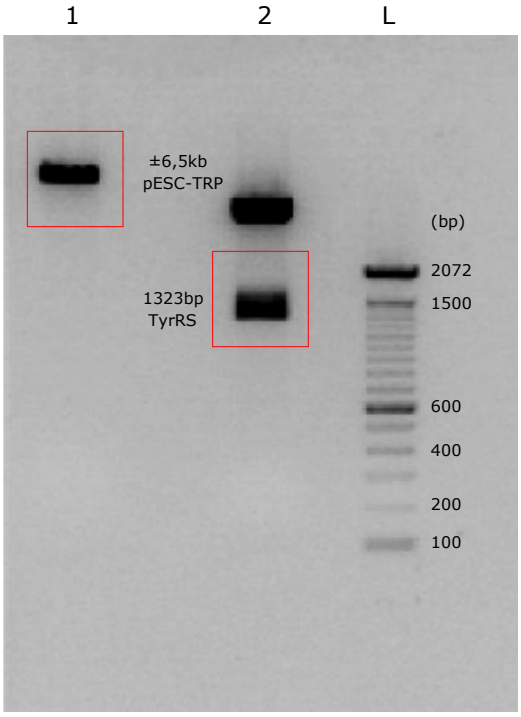

Fig. 8. Restriction digestion of the pESC-TRP vector and pCR2.1/EcTyrRS with restriction enzymes EcoRI and NotI. 1= opened pESC-TRP vector (6,5kb) indicated with a red frame; 2= the highest band represents the opened pCR2.1 vector, the lowest band, indicated with a red frame, represents the TyrRS insert (1323bp); L= 100bp ladder (Invitrogen). (1% agarose)

The pEcTyrRS vector was transformed into competent Top10 F' *E. coli* cells, which were allowed to grow overnight. Plasmid DNA from four colonies was isolated and sent for sequencing. All clones contained the *E. coli* TyrRS insert without mutations (data not shown).

2 Amplification and cloning of *E. coli* $tRNA_{CUA}$

2.1 Optimization of the PCR

In order to amplify the *E. coli* $tRNA_{CUA}$, a PCR was performed on a single colony of *E. coli* BL21 (DE3) pLys. To obtain optimal PCR conditions, a gradient was applied for both the annealing temperature (45-70°C) and the $MgCl_2$ concentration (2, 4, 6 and 8mM). The PCR reaction mixtures were analysed by agarose gel electrophoresis, from which could be concluded that 4mM was the

optimal $MgCl_2$ concentration (results not shown). The annealing temperature was further optimized in a second gradient PCR, with the annealing temperature ranging from 55°C to 75°C. The reaction mixtures of the second PCR were analysed by agarose gel electrophoresis. *E. coli* $tRNA_{CUA}$ has a length of 156bp. With the extra nucleotides added by the primers and the extra adenosine added by Taq polymerase taken into account, a fragment of 181bp was expected. The density of the $tRNA_{CUA}$ band on the agarose gel increased with decreasing primer annealing temperature, peaking at 62,5°C. When the primer annealing temperature was lowered, the density decreased again (Fig. 9). The extra bands in lane 1 are probably the product of nonspecific amplification and the bands smaller than 100bp are most likely primer dimers. The $tRNA_{CUA}$, amplified with the most optimal conditions (primer annealing at 62,5°C and 4mM $MgCl_2$) was cloned in the pCR2.1 vector. pCR2.1/EctRNA_{CUA} was transformed in Top10 F' competent *E. coli* cells.

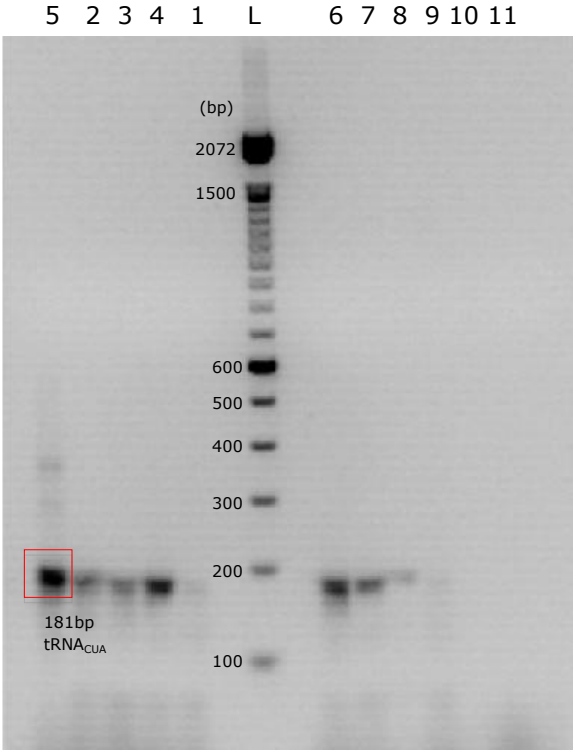

Fig. 9. Amplification of $tRNA_{CUA}$ from an BL21 (DE3) pLys *E. coli* colony. A temperature gradient was applied to optimize the PCR. L=100bp ladder (Invitrogen); 1=75,0°C; 2=73,5°C; 3=71,0°C; 4=67,2°C; 5=62,5°C; 6=58,9°C; 7=56,5°C; 8= 55,0°C; 9= no sample loaded, the visible bands are probably the result of leakage from one of the adjacent lanes; 10= negative control I (mastermix without DNA); 11= negative control II (mastermix without primers and Taq polymerase. 62,5°C was the optimal annealing temperature. Reaction mixture 5 and 1 were switched during the loading of the gel (2% agarose).

2.2 Restriction analysis and sequencing

The *E. coli* cells, transformed with pCR2.1/EctRNA_{CUA}, were allowed to grow overnight. The next day, the plasmid DNA was isolated and the insert was analysed by restriction digestion. After restriction digestion with restriction enzyme EcoRI, two fragments were expected for clones containing the *E. coli* $tRNA_{CUA}$: one fragment of approximately 3,9kb (pCR2.1) and another fragment of 196bp ($tRNA_{CUA}$). In total, 10 clones were tested; clones 5, 9 and 10 showed fragments of the expected lengths (Fig 10). In order to verify these results, the clones were sequenced and resulting sequences were aligned with the sequence of *E. coli* $tRNA_{CUA}$ (Fig 11). All clones showed one nucleotide differing from the sequence of *E. coli* $tRNA_{CUA}$. In order to locate the

difference, the secondary structure of the amplified tRNA was modelled using the online tool 'tRNAscan-SE Search Server' (Lowe Lab) and compared to the secondary structure of the *E. coli* $tRNA_{CUA}$. The tRNA structures are shown in Fig. 12.

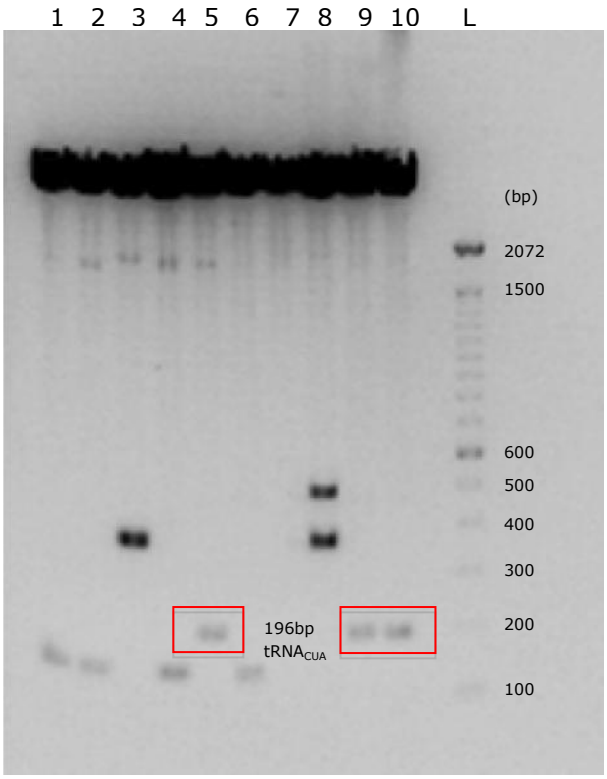

Fig. 10. Restriction digestion of pCR2.1/EctRNA with EcoRI. 1-10= clones 1-10; L= 100bp ladder (Invitrogen). Fragments of the expected length was present in clone 5, 9 and 10. (2% agarose)

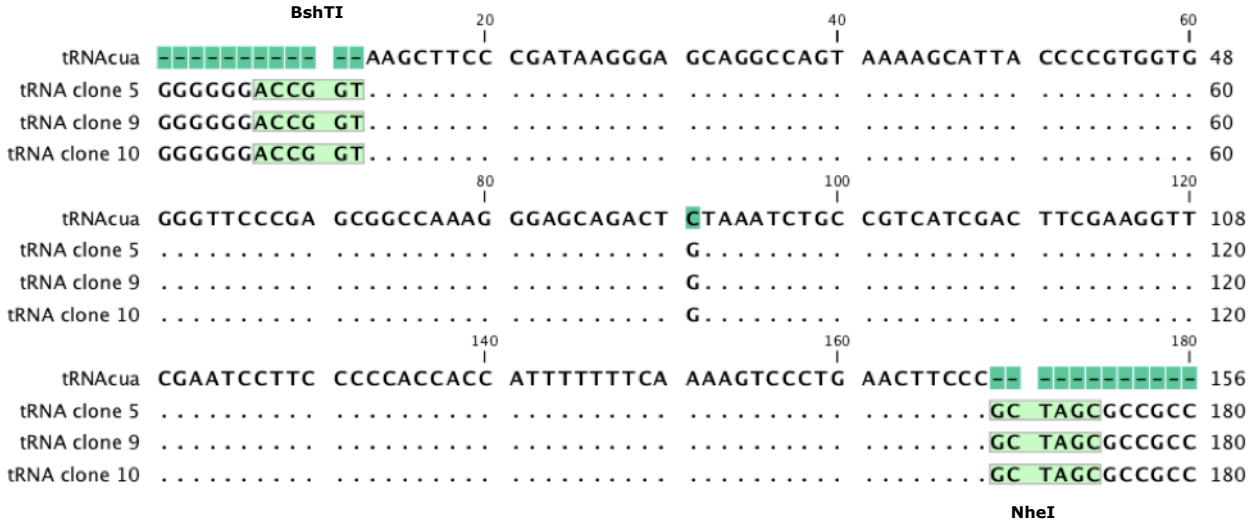

Fig. 11. Alignment of the $tRNA_{CUA}$ sequence with the sequences of clone 5, 9 and 10. The sequences differ by one nucleotide, indicated in blue. The restriction sites (NheI and BshTI) are indicated in green

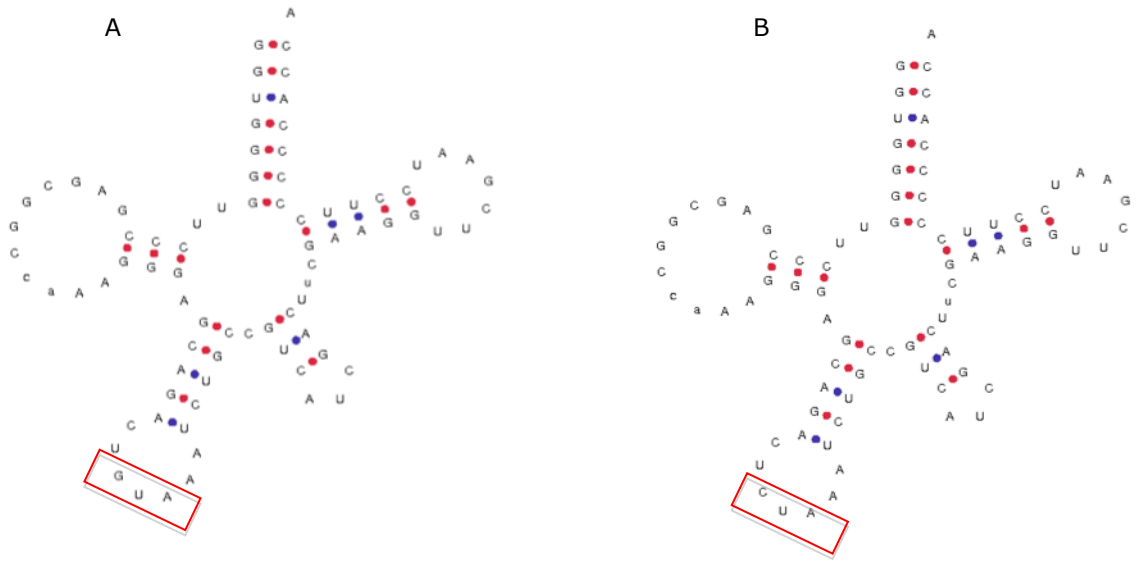

Fig. 12. Secondary structure of the tRNA amplified from *E. coli* BL21 (DE3) pLys (A) and the *E. coli* amber suppressor $tRNA_{CUA}$ (B). The two tRNA molecules have a different anticodon. The amplified tRNA is not the amber suppressor $tRNA_{CUA}$, but the conventional tyrosyl $tRNA_{GUA}$.

Modeling the secondary structure of both the *E. coli* $tRNA_{CUA}$ and the amplified tRNA revealed that the differing nucleotide was located in the anticodon. All the clones that were amplified were the conventional *E. coli* tyrosyl tRNA with anticodon GUA, and not the amber suppressor tRNA with anticodon CUA. This is not so surprising, because the tRNA sequences differ only by one nucleotide and the tyrosyl tRNA is much more abundant than the amber suppressor tRNA. Chances are that if more colonies were picked from the plate, the wanted tRNA would be found. However, probably a large number of colonies would need to be sequenced before the $tRNA_{CUA}$ was found. Therefore, the anticodon of the amplifed tyrosyl tRNA was modified by site-directed mutagenesis. After the site-directed mutagenesis, sequencing verified that the anticodon had changed from GUA to CUA, which recognizes the amber stop codon (results not shown).

2.3 Cloning in pEcTyrRS vector

In order to create the pEcTyrRS/$tRNA_{CUA}$ vector, the *E. coli* $tRNA_{CUA}$ was cloned in the plasmid pEcTyrRS. The *E. coli* $tRNA_{CUA}$ was first excised from the pCR2.1 vector by restriction digestion with restriction enzymes BshTI (AgeI) and NheI. The plasmid pEcTyrRS was cut with the same enzymes. As recommended by the DoubleDigest™ tool from Fermentas, 1X Tango™ buffer was used with a 4-fold excess of BshTI (AgeI). After digestion of the pEcTyrRS vector, 2 fragments were expected: a first fragment of approximately 6kb, which represents the open vector, and a second fragment of approximately 0.5kb. Digestion of pCR2.1/EctRNA_{CUA} should result in a fragment of 166bp, the $tRNA_{CUA}$ insert, and the pCR2.1 vector with an approximate length of 3.9kbp. After the restriction digestion, the reaction mixtures were analysed by agarose gel electrophoresis. The open pEcTyrRS vector, indicated with a red frame (Fig. 13, lane1), was purified from the agarose gel. $tRNA_{CUA}$ on the other hand, also indicated with a red frame (Fig. 13, lane2), was to faint under the UV-light without the help of imaging software and thus not suitable to excise from the gel. This can be explained by the small size of $tRNA_{CUA}$. The lower the amount of DNA, the lower the amount of GelRed™ that can intercalate and thus the lower the fluorescence.

An additional explanation could be the bad compatibility of the two restriction enzymes, with as a consequence an incomplete digestion of the DNA. As a result, only a fraction of the $tRNA_{CUA}$ is excised from the pCR2.1 vector, whereby the visibility of the fragment is even more reduced.

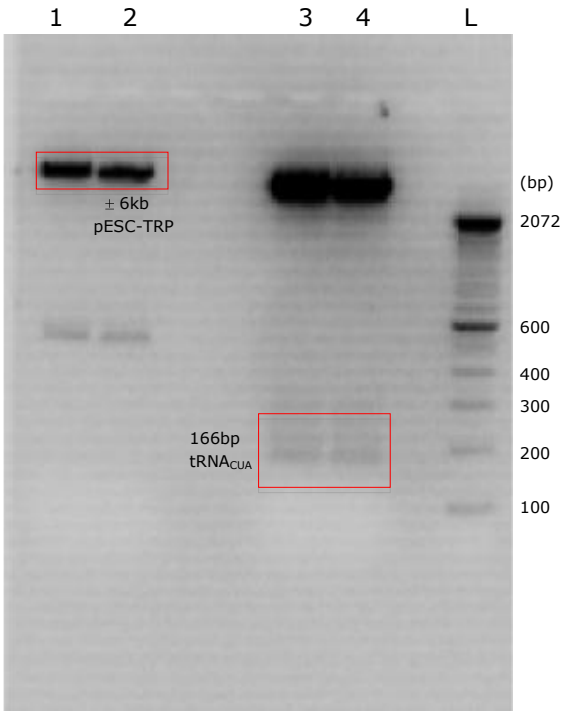

Fig. 13. Restriction digestion of the pEcTyrRS vector and the pCR2.1/ECtRNA$_{CUA}$ vector with restriction enzymes BshTI (AgeI) and NheI. 1+2= the highest bands, indicated with a red frame, represent the opened pEcTyrRS vector (6kb). The lowest bands represent a fragment of 500bp, which will be replaced by the tRNA$_{CUA}$. 3+4= the highest bands represent the pCR2.1 vector. The lowest bands, indicated by a red frame, represent the tRNA$_{CUA}$ insert. L= 100bp ladder (Invitrogen). (2%agarose)

In a second attempt, the restriction reaction was carried out in two steps, to solve the compatibility problem, and on a larger amount of DNA. In a first step, the $tRNA_{CUA}$ was digested with the first restriction enzyme, NheI. The partially digested DNA was separated on an agarose gel. As expected, a fragment of approximately 4,1kbp was seen after the first restriction digestion step, representing the open vector with $tRNA_{CUA}$ (Fig. 14A). After purification, the DNA was digested with the second restriction enzyme BshTI (AgeI) resulting in the excision of $tRNA_{CUA}$. Next to a band of approximately 3,9kb (the open pCR2.1 vector) a second fragment of 166bp was expected, representing the excised $tRNA_{CUA}$. However, that band was not visible on the gel, as shown in Fig. 14B. Again, this is probably to blame to the small size of the $tRNA_{CUA}$. After the first digestion step, the purified DNA was eluted with 50µl MQ, resulting in a large reaction volume for the second restriction reaction. As a consequence, the second restriction reaction had to be divided over two slots of the gel, implying dilution of the DNA. Additionally, with each purification step, a fraction of the DNA is lost. So, although the restriction digestion was started with a larger amount of DNA, the $tRNA_{CUA}$ was still not visible on the agarose gel.

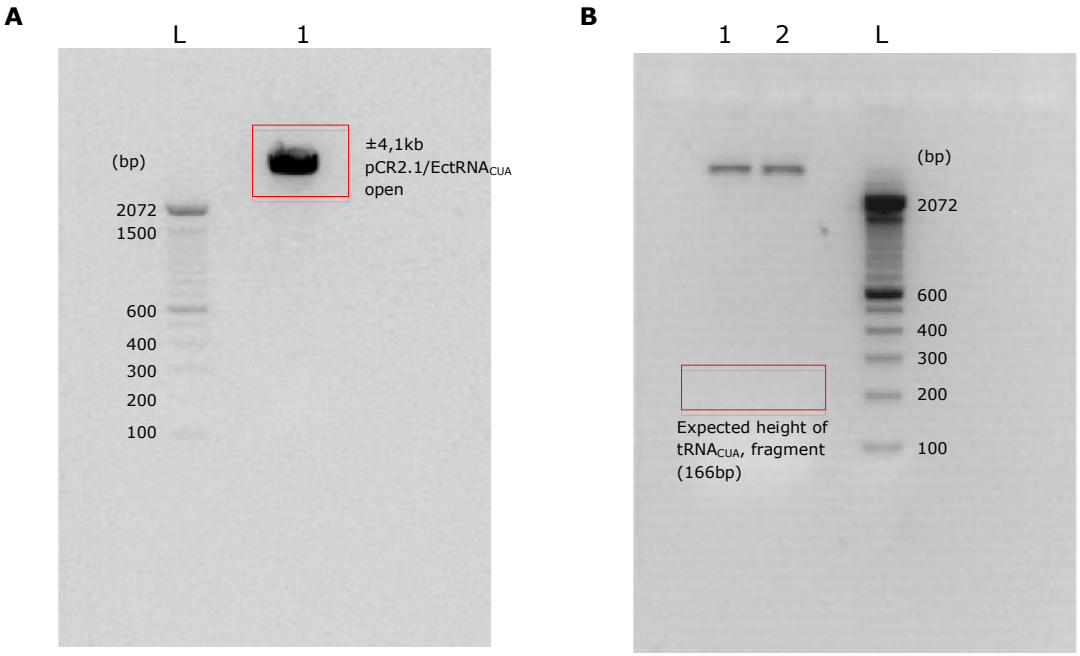

Fig. 14. Restriction digestion of pCR2.1/EctRNA$_{CUA}$ with restriction enzymes BshTI (AgeI)and NheI. A) 1= First restriction step with NheI. Open pCR2.1/EctRNA$_{CUA}$ vector (indicated with the red frame); L= 100bp ladder (Invitrogen) (1% agarose). B) 1+2= Second restriction step with BshTI. The red frame indicates the height where the excised tRNA$_{CUA}$ fragment was expected. The highest band represents the empty pCR2.1 vector (1,5% agarose).

In order to avoid diluting the sample too much and get round the compatibility problem, FastDigest® restriction enzymes were used in a third effort to isolate the tRNA$_{CUA}$. FastDigest® enzymes are an advanced line of restriction enzymes that are 100% active in the universal FastDigest® buffers. Using these enzymes made it possible to perform the restriction reaction in one step, reducing the reaction volume and avoiding loss of material. The result of the restriction digestion is shown in Fig. 15A. As expected, both a fragment of approximately 166bp (the excised tRNA$_{CUA}$) and a fragment of approximately 3,9kb (the open pCR2.1 vector) were visible on the agarose gel, implying that the restriction digestion was successful. A small volume of the purified tRNA$_{CUA}$ was loaded on an agarose gel, together with the purified open plasmid pESC-TRP (Fig. 15B) in order to estimate the proportion insert/vector for the ligation reaction. Taken the size of both fragments into account, the concentration of the DNA of both the vector and the insert were in the same range, but both low. For the ligation, three ratios of vector over insert were tested; 1:1, 1:2 and 1:3. However, after transformation, no noticeable difference was present between the three different ratios and the negative control plate (vector without insert). A possible explanation is that the concentration of the insert was to low, or that a fraction of the pESC-TRP vector was not completely digested and thus able to self-ligate. Self-ligated vector is the most efficiently transformed and as result, a large amount of background colonies were present.

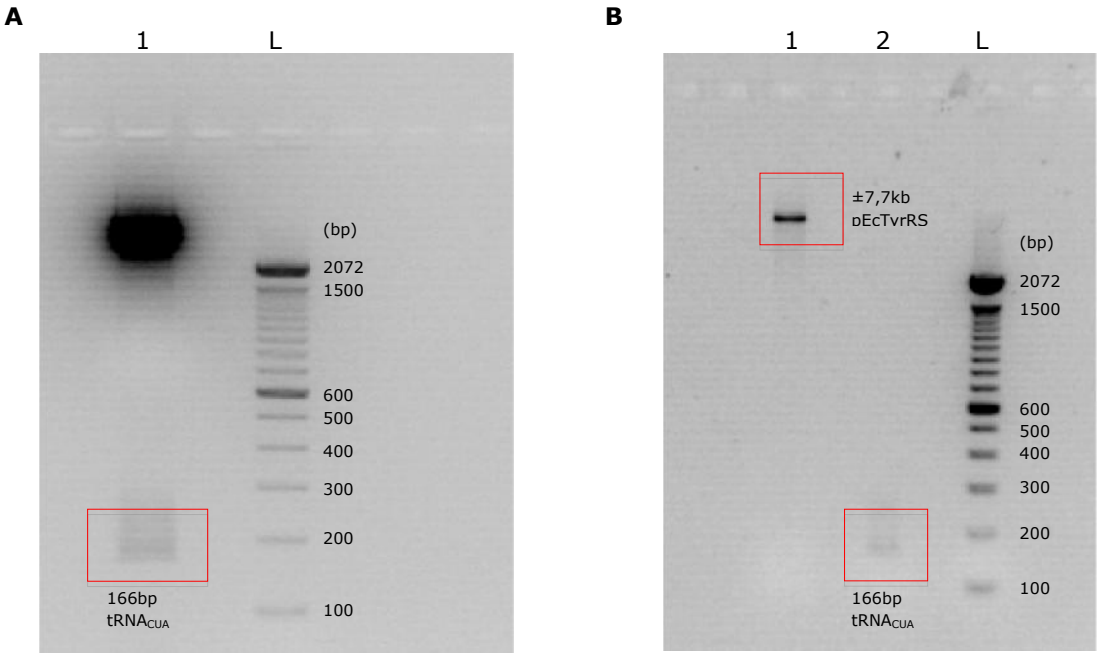

Fig. 15. Restriction digestion of pCR2.1/EctRNA$_{CUA}$ vector with FastDigest restriction enzymes NheI and AgeI. A) 1=The highest band represents the pCR2.1 vector (3,9kb). The lowest band, indicated with a red frame, indicates the tRNA$_{CUA}$ insert. L= 100bp ladder (Invitrogen) (2% agarose). B) The open pEcTyrRS vector and the tRNA$_{CUA}$ insert after purification from the agarose gel. 1= open pEcTyrRS (7,7kb); 2= tRNA$_{CUA}$ insert (166bp). L= 100bp ladder (Invitrogen) (1,5% agarose).

Increasing both the duration of the restriction digestion (in order to obtain complete digestion) and the concentration of the DNA, solved this problem. In order to digest a substantial amount of DNA without increasing the reaction volume, the concentration of the DNA of both pESC-TRP and pCR2.1/EctRNA$_{CUA}$ was increased a ten-fold by ethanol precipitation. This last restriction digestion of both tRNA$_{CUA}$ and pEcTyrRS resulted in clearly visible fragments on the agarose gel, which could be easily purified. In order to obtain large amounts of tRNA, both restriction reactions (of both lane 1 and 2) of tRNA$_{CUA}$ were purified from the gel (Fig 16). After purification of both fragments from the agarose gel, the tRNA$_{CUA}$ insert was cloned in the pEcTyrRS vector, resulting in the pEcTyrRS/tRNA$_{CUA}$. For the ligation reaction, three vector:insert ratios were tested (1:1, 1:3 and 1:6); The ratio 1:6 gave the best results. The pEcTyrRS/tRNA$_{CUA}$ vector was transformed into Top10 F' competent *E. coli* cells. Plasmid DNA was isolated from 10 colonies, grown overnight, and analysed by restriction digestion (data not shown). From the clones containing an insert of the expected length, 4 clones were sent for sequencing. All 4 clones contained the *E. coli* tRNA$_{CUA}$ insert without mutations (data not shown).

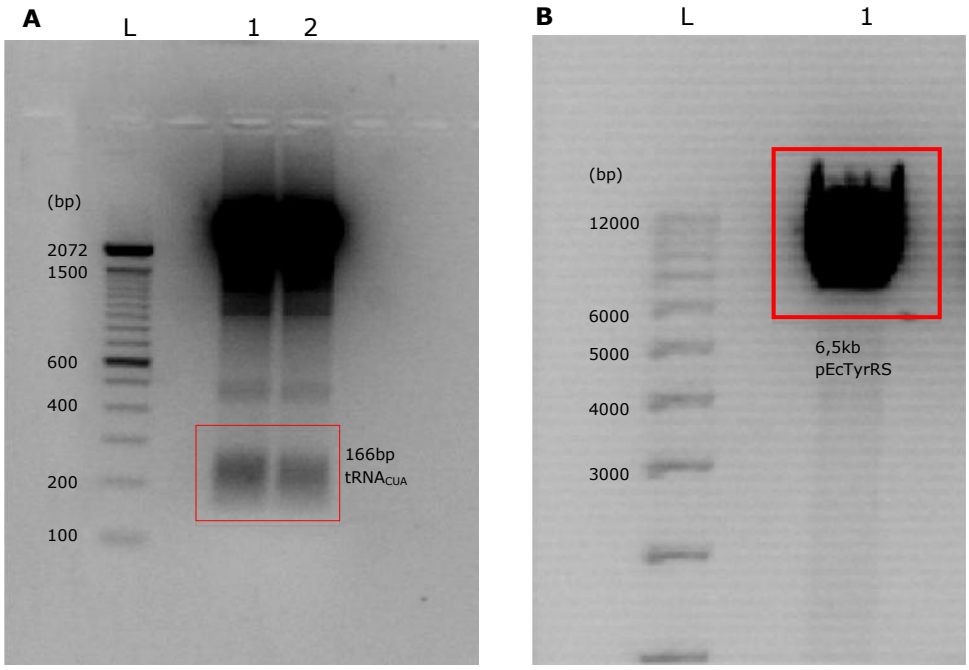

Fig. 16. Restriction digestion of pCR2.1/tRNA$_{CUA}$ and pEcTyrRS, with FastDigest restriction enzymes NheI and AgeI. A) Restriction digestion of pCR2.1/EctRNA$_{CUA}$. L=100bp ladder (Invitrogen); 1+2= the highest bands represent the pCR2.1 vector (3,9kb). The lowest bands, indicated with a red frame, represent the tRNA$_{CUA}$ insert (200bp). B) Restriction digestion of pEcTyrRS. L=1kb plus ladder (Invitrogen). 1= The band, indicated with a red frame, represents the open pEcTyrRS.

3 Amplification and cloning of the *S. cerevisiae* ADH1 promoter

3.1 Optimization of the PCR, restriction analysis and sequencing

In order to amplify the yeast full length ADH1 promoter, a PCR was performed on the plasmid pMA424. After the amplification reaction, the reaction mixtures were analysed by agarose gel electrophoresis. AHD1 has a length of 1443bp. With the extra nucleotides added by the primers and the extra adenosine added by Taq polymerase included, a fragment of 1468 bp was expected. PCR was performed on two minipreps. Both minipreps showed successful amplification; a fragment of the expected height was seen for both PCR reaction mixtures (Fig 17, lane 1 and 2). In lane 3, a band is present for the negative control (mastermix without plasmid DNA). This is probably the result of leakage from one of the other lanes. ADH1, amplified from miniprep 1, was cloned in the pCR2.1 vector and transformed in Top10 F' competent *E. coli* cells.

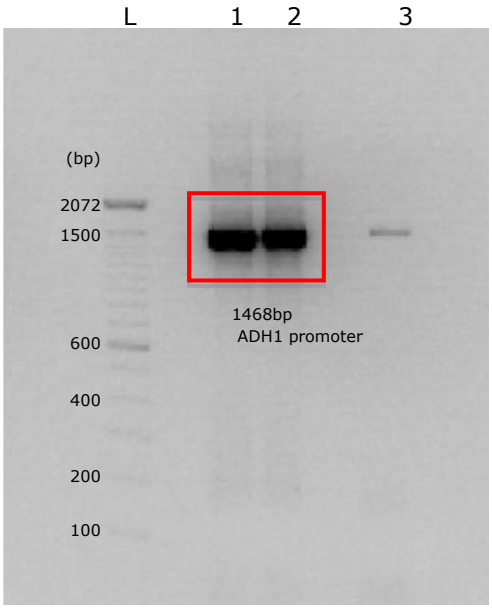

Fig. 17. Amplification of the *S. cerevisiae* ADH1 promoter from the vector pMA424. L=100bp ladder (Invitrogen); 1= ADH1 promoter amplified from miniprep 1; 2= ADH1 promoter amplified from miniprep 2; 3= negative control I (mastermix without DNA); 4= negative control II= mastermix without primers and Taq polymerase. The band in lane 3 is probably a result from leakage from one of the other lanes.

pCR2.1/ADH1 plasmid DNA was isolated from 4 overnight grown colonies and the insert was analysed by restriction digestion with restriction enzyme EcoRI. In Fig. 18, the results of the restriction analyses are visible in lanes 5 to 8. Fragments of the expected length (1483 bp for the ADH1 promoter insert and 3,9kb for the pCR2.1 vector) were present in all four lanes. In order to verify these results, the clones were sequenced and the resulting sequences were aligned with the sequence of the full length ADH1 promoter of the pDBLeu vector (Invitrogen) (Fig. 19). From the sequencing results could be concluded that none of the clones was suitable to clone in the pEcTyrRS/$tRNA_{CUA}$ vector. The BshTI (AgeI) restriction site was absent in clone 3 and all three clones contained several mutations. Not only compared to the reference sequence, but also compared to each other. This can possibly be explained by the polymerase used for the amplification, Taq polymerase, which has no proofreading function. Another factor that could cause these mutations is the amount of DNA used for the amplification. When to much DNA is added to the reaction mixtures in proportion to the other products, it is possible that those products are almost exhausted before the end of the amplification reaction, creating mismatches in the amplified DNA.

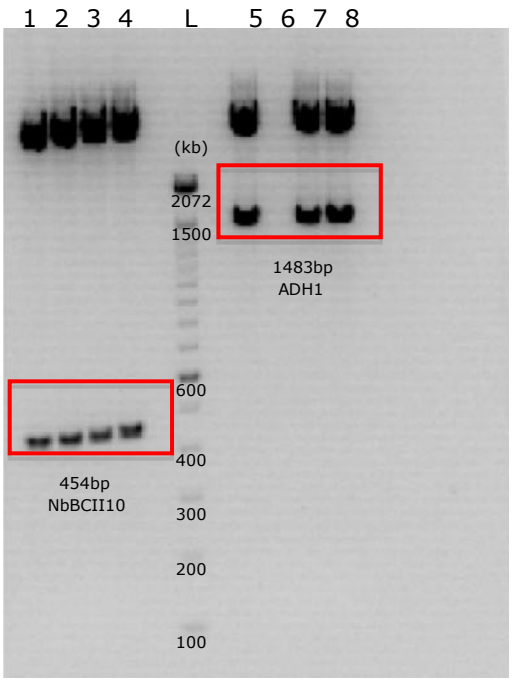

Fig. 18. Restriction digestion of pCR2.1/NbBCII10 and pCR2.1/ADH1 with EcoRI. 1-4= clones 1-4 from pCR2.1/NbBCII10; L= 100bp ladder (Invitrogen). 5-8= clones 1-4 from pCR2.1/ADH1. A fragment of the expected length was present in clones 1, 2, 3 and 4 from pCR2.1/NbBCII10, and in clones 5, 7 and 8 of pCR2.1/ADH1. (1,5% agarose)

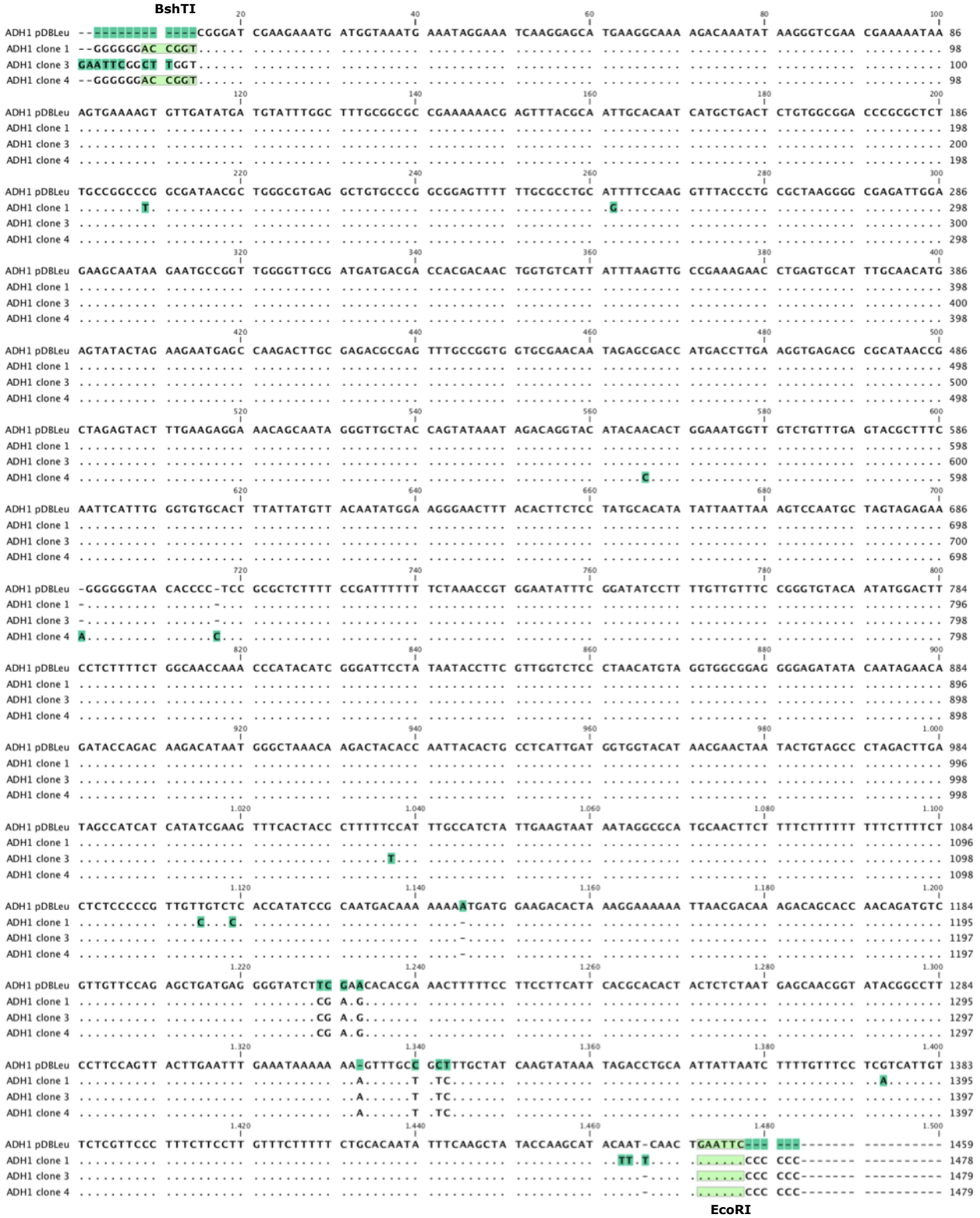

Fig. 19. Alignment of the pDBLeu ADH1 promoter with the sequences of clones 1, 3 and 4. Clone 3 does not contain the BshTI restriction site and there are several mutations in the three clones. Not only compared to the reference ADH1 promoter, but also compared to each other. The mutations are indicated in blue, the restriction sites (BshTI and EcoRI) with a green box.

In order to solve the problem of the mutations, a new gradient PCR was carried out with different dilutions of the pMA424 miniprep and instead of Taq polymerase, Pfu polymerase was used. Pfu

polymerase is a high fidelity polymerase with proofreading function. In order to add the adenosines to the amplified genes, necessary for TA cloning, Taq polymerase was added to the reaction mixtures at the end of the amplification reaction and incubated for an additional 10 minutes at 72°C. After the PCR, the reaction mixtures were analysed by agarose gel electrophoresis. Based on the agarose gel, no noticeable difference could be detected between the different concentrations of plasmid DNA (Fig. 20). The reaction mixture with the lowest DNA concentration was used for the TA cloning. The resulting pCR2.1/ADH1 construct was transformed into Top10 F' competent *E. coli* cells and the next day, plasmid DNA was isolated from 4 overnight grown colonies. The presence of the correct insert was checked by restriction digestion and sequencing. The results of the restriction analysis are shown in Fig. 21. Again, two fragments were expected, a first fragment of approximately 3,9kb (the pCR2.1 vector) and a second fragment of 1483bp (the ADH1 promoter). Fragments of the expected length were present in lanes 3, 4 and 5. Lanes 2 also contains two fragments of the expected length, but not clearly visible. This is possibly the result from leakage from one of the adjacent lanes. Clone 3, 4 and 5 were sequenced. Clone 4 contained two mutations compared to clones 3 and 5. Clone 3 and 5 showed complete sequence similarity. There were still several differences compared to the ADH1-promoter from the pDBLeu vector, but that is probably due to natural variations, which do not interfere with normal promoter function (Fig. 22).

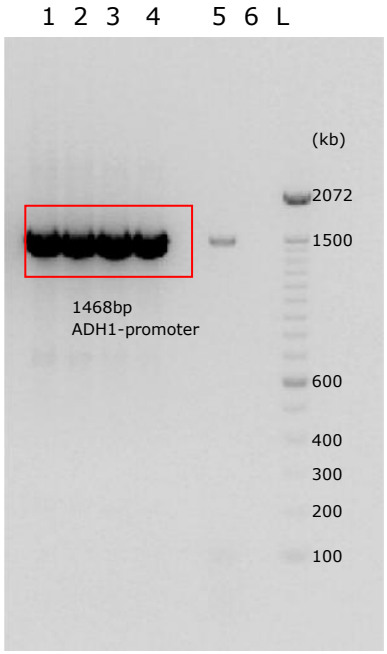

Fig. 20. Amplification of the *S. cerevisiae* ADH1 promoter from the vector pMA424. In order to optimize the reaction, a concentration gradient of the pMA424 plasmid DNA was applied. 1= miniprep diluted 1/100; 2= miniprep diluted 1/500; 3= miniprep diluted 1/1000; 4= miniprep diluted 1/2000; 5= negative control I (mastermix without DNA); 6= negative control II (mastermix without primers and polymerase); L=100bp ladder (Invitrogen); The band in lane 5 is probably a result from leakage from one of the other lanes or from contamination from one of the other samples.

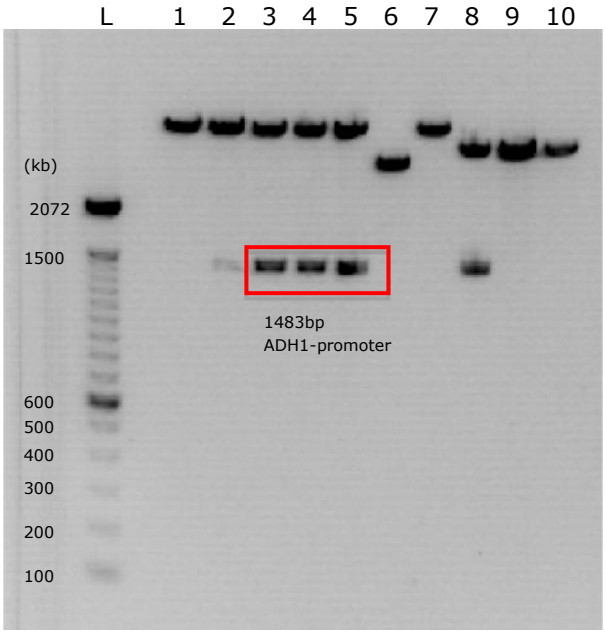

Fig. 21. Restriction digestion of pCR2.1/ADH1 with EcoRI. L= 100bp ladder (Invitrogen); 1-10= clones 1-10 from pCR2.1/ADH1. Fragments of the expected length were clearly visible in clones 3, 4 and 5. (1,5% agarose)

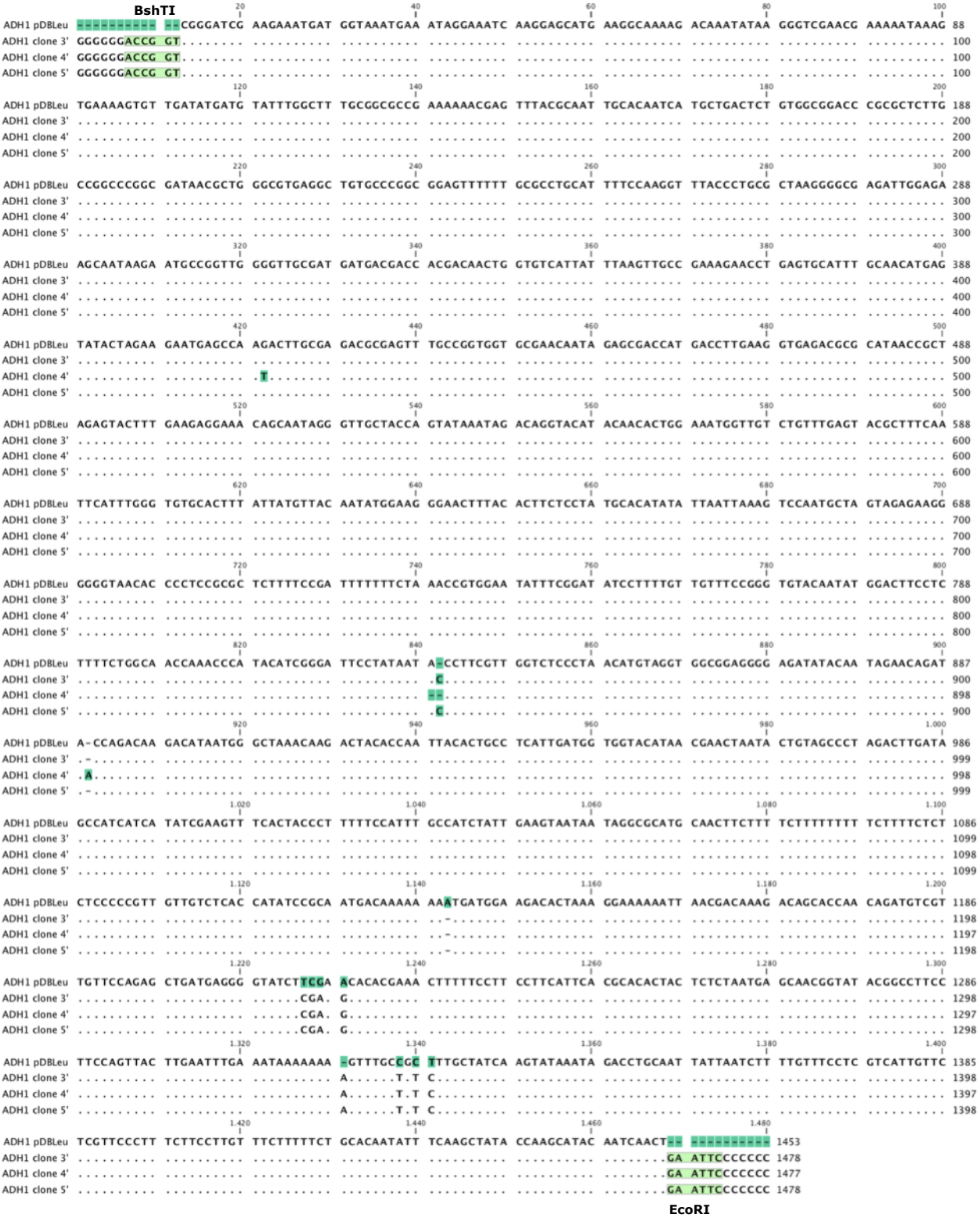

Fig. 22. Alignment of the pDBLeu ADH1 promoter with the sequences of clones 3, 4 and 5. Clone 4 has 2 mutations compared to clone 3 and 5. Clone 3 and 5 show complete sequence similarity. There are still several differences compared to the ADH1 promoter of the pDBLeu vector. However, this is probably not a problem. Differences between the different sequences are indicated in blue, the restriction sites (BshTI and EcoRI) in green.22

3.2 Cloning in pEcTyrRS/tRNA$_{CUA}$

In order to complete the pEcTyrRS/tRNA$_{CUA}$ construct, clone 3 of the amplified ADH1 promoter was cloned in the pEcTyrRS/tRNA$_{CUA}$ vector. The ADH1 promoter was first excised from the pCR2.1 vector by restriction digestion, with the conventional restriction enzymes EcoRI and BshTI (AgeI). The pEcTyrRS/tRNA$_{CUA}$ vector was opened using the same enzymes. Restriction digestion of pEcTyrRS/tRNA$_{CUA}$ should generate two fragments: a first fragment of approximately 6,2 kb (the open vector) and a second fragment of approximately 300bp. Digestion of pCR2.1/ADH1 should result in a fragment of 1453bp (the ADH1 promoter insert) and a fragment of 3,9kb (the pCR2.1 vector). After the restriction reaction, the fragments were analysed by agarose gel electrophoresis. The restriction digestion was successful, as fragments of the expected length appeared on the agarose gel (Fig 23). The ADH1 promoter and the open pEcTyrRS vector were excised from the agarose gel and after purification of the DNA, the ADH1 promoter was cloned into the pEcTyrRS/tRNA$_{CUA}$ vector, resulting in the complete pEcTyrRS/tRNA$_{CUA}$ vector construct. The pEcTyrRS/tRNA$_{CUA}$ vector with the ADH1 promoter was transformed into Top10 F' competent *E. coli* cells. Plasmid DNA was isolated from 5 colonies, grown overnight, and analysed by restriction digestion. All the clones contained an insert from the expected length (data not shown). Of these clones, 3 were sent for sequencing. The sequencing results showed that all clones contained the ADH1 promoter without mutations (data not shown).

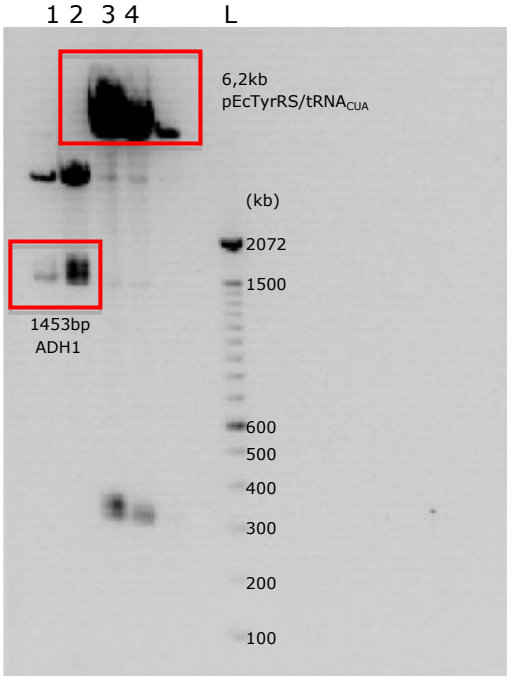

Fig. 23. Restriction digestion of the pCR2.1/ADH1 vector and the pEcTyrRS/tRNA$_{CUA}$ vector with restriction enzymes BshTI (AgeI) and EcoRI. 1+2= restriction reaction of pCR2.1/ADH1. The highest bands (approximately 3,9kb) represent the empty pCR2.1 vector. The bands indicated with a red frame represent the excised ADH1 promoter (1449b). 3+4= restriction reaction of pEcTyrRS/tRNA$_{CUA}$. The highest bands indicated with a red frame represent the opened pEcTyrRS/tRNA$_{CUA}$ vector. The lowest bands represent the fragment excised from the pEcTyrRS/tRNA$_{Cua}$ vector, which will be replaced by the ADH1 promoter. L= 100bp ladder (Invitrogen). (1%agarose)

II. Construction of the pTEF-MF/BCII10 vector

1 Amplification and cloning of NbBCII10

1.1 PCR

For the construction of the pTEF-MF/BCII10 vector, the NbBCII10 was amplified from the pHEN6 vector by PCR. The PCR reaction mixture was analysed by agarose gel electrophoresis. The BCII10 nanobody has a length of 402bp. With the 6-His tag, the extra nucleotides added by the primers and the extra adenosine added by Taq polymerase taken into account, a fragment of 439bp was expected. In Fig. 24, the results of the amplification are shown. A clear band is visible at the expected height. The BCII10 Nb was cloned in the pCR2.1 vector. Subsequently, pCR2.1/ADH1 was transformed in Top10F' competent *E. coli* cells.

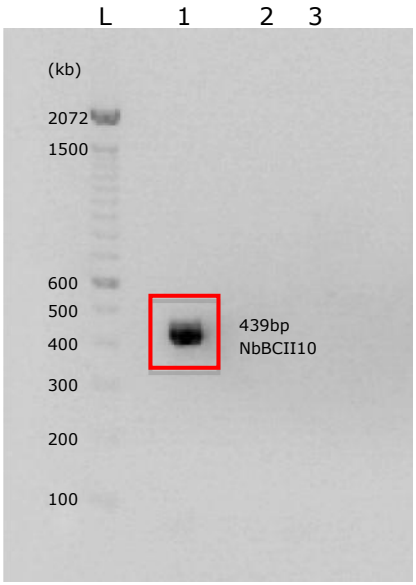

Fig. 24. Amplification of NbBCII10 from the pHEN6 vector. L=100bp ladder (Invitrogen); 1= amplified NbBCII10 (438bp, indicated with a red frame); 2= negative control I (reaction mixture without DNA); 3= negative control II (reaction mixture without primers and Taq polymerase).

1.2 Restriction analysis and sequencing

pCR2.1/BCII10 plasmid DNA was isolated from four overnight grown colonies of the transformed Top10 F' *E. Coli* cells and the insert was analysed by restriction digestion with restriction enzyme EcoRI. In Fig. 18, the results of the restriction analyses are visible in lanes 1 to 4. Fragments of the expected length (454bp for the NbBCII10 insert and 3,9kb for the pCR2.1 vector) are present in all four lanes. In order to verify these results, the clones were sequenced and the resulting sequences were aligned with the sequence of BCII10 (Fig 25). All sequences showed one mutation compared to the BCII10 sequences, but that was a silent mutation and thus not problematic.

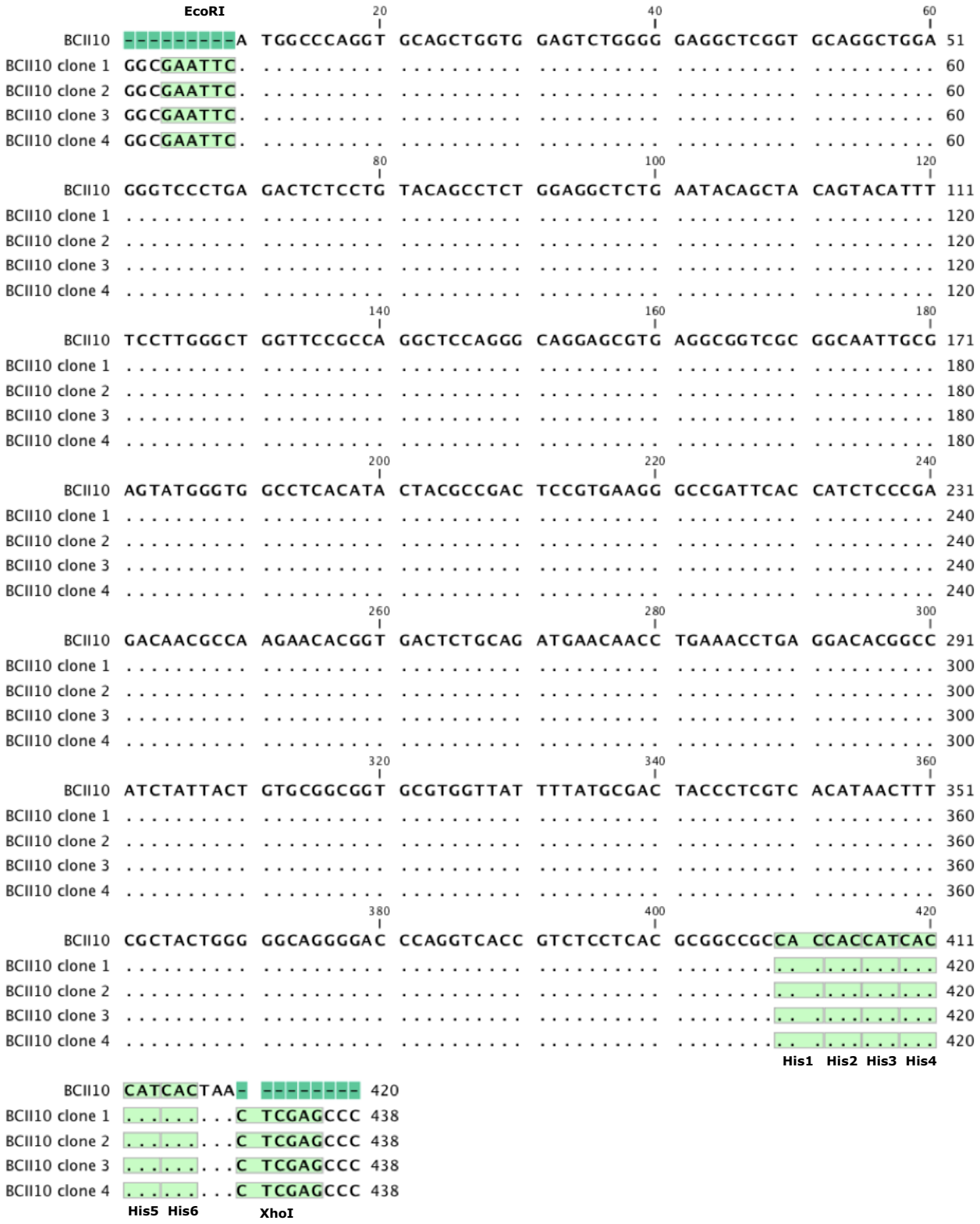

Fig. 25. Alignment of the BCII10 nanobody with the different BCII10 clones, amplified by PCR. All clones contain one silent mutation compared to the BCII10 sequence. The mutation is indicated in blue. The restriction sites and the 6His-Tag are indicated in green.

1.3 Cloning in pTEF-MF

With the purpose of cloning the BCII10 nanobody in the pTEF-MF yeast expression vector, the pCR2.1/BCII10 vector and pTEF-MF vecter were digested with the restriction enzymes XhoI and EcoRI. Restriction digestion of pTEF-MF should result in a fragment of 6575bp, and a second fragment of 33bp (not visible on an agarose gel). Digestion of pCR2.1/BCII10 should result in a fragment of 3.9kb (the empty pCR2.1 vector) and a fragment of 430bp (the BCII10 insert). The fragments obtained after restriction digestion were analysed by agarose gel electrophoresis. The restriction digestion was a success, as fragments of the expected length are visible on the agarose gel (Fig 26). The BCII10 nanobody and the open pTEF-MF vector were excised from the agarose gel and after purification of the DNA, the BCII10 nanobody was cloned into the pTEF-MF vector, resulting in the pTEF-MF/BCII10 vector construct. The vector construct was transformed into Top10 F' competent *E. coli* cells. Plasmid DNA was isolated from four colonies, grown overnight, and analysed by restriction digestion and sequencing. All clones contained the BCII10 nanobody without mutations and in frame, which is important for the expression of the nanobody under the control of the TEF1 promoter (data not shown).

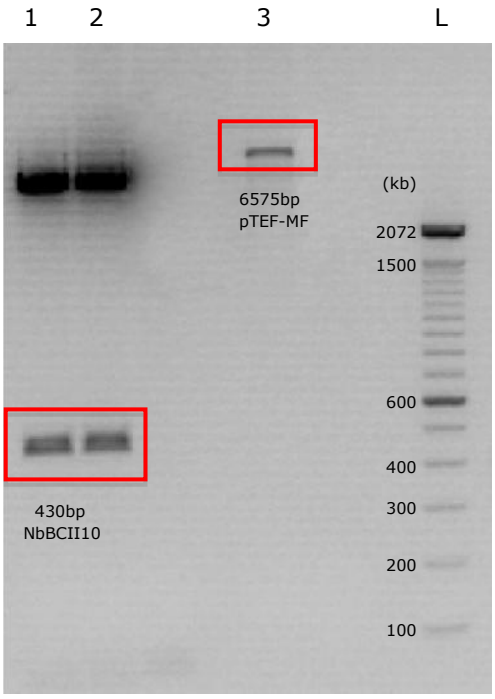

Fig. 26. Restriction digestion of the pCR2.1/BCII10 vector and the PTEF-MF vector with restriction enzymes EcoRI and XhoI. 1+2= restriction reaction of pCR2.1/BCII10. The highest bands (approximately 3,9kb) represent the empty pCR2.1 vector. The bands indicated with a red frame represent the excised BCII10 nanobody (430bp). 3= restriction reaction of pTEF-MF. The band, indicated with a red frame, represents the opened pTEF-MF vector. L= 100bp ladder (Invitrogen). (1,5%agarose)

CONCLUSION

This senior internship was part of a larger project, with as ultimate goal the site-specific immobilization of nanobodies on a solid substrate using 'click' chemistry. In order to accomplish this, a 'click' functionalized amino acid will be introduced into the protein during the translation process. For this, the 'amber' stop codon will be reassigned with the help of a mutant *E. coli* TyrRS/$tRNA_{CUA}$ pair, which recognizes the 'click' functionalized amino acid instead of tyrosine. This EcTyrRS/$tRNA_{CUA}$ pair will be used as an orthogonal pair in the yeast *S. cerevisiae* to introduce the 'click' functionality into the nanobody. The purpose of this senior practical training was to carry out some preliminary steps of this project and to develop the necessary vectors.

A first objective of the practical training was to amplify the genetically encoded *E. coli* TyrRS/$tRNA_{CUA}$ pair and the *S. cerevisiae* ADH1 promoter, and clone them into the pESC-TRP vector. A second objective of the practical training was to amplify the BCII10 nanobody and clone it into the pTEF-MF yeast expression vector. The PCR reactions for the amplification of the genes were optimized by applying both temperature and $MgCl_2$ gradients. After amplification, the four genes were cloned in the pCR2.1 high-copy vector and the resulting pCR2.1 constructs were transformed in Top10 F' competent *E. coli* cells. Plasmid DNA was isolated and the presence of the insert was verified by restriction analysis and sequencing. The sequence of the amplified *S. cerevisiae* ADH1 promoter showed several point mutations. A second PCR of the ADH1 promoter was carried out to solve this problem. This time, a DNA concentration gradient was applied and instead of Taq polymerase, a high fidelity polymerase was used. The amplified genes were then cloned into their destination vector. Because $tRNA_{CUA}$ is a small fragment, it was difficult to clone the $tRNA_{CUA}$ in the pEcTyrRS vector. To solve this issue, the DNA was ethanol precipitated before cloning, in order to increase the DNA concentration. The resulting pEcTyrRS/$tRNA_{CUA}$ and pTEF-MF/BCII10 constructs were analysed by restriction digestion and sequencing. Both revealed that the construction of both the pEcTyrRS/$tRNA_{CUA}$ vector and the pTEF-MF/BCII10 vector were successful.

In a next step, the pEcTyrRS/$tRNA_{CUA}$ vector will be used for the construction of a library of mutant TyrRSs. This will be accomplished by a random site-directed mutagenesis reaction, targeting 5 amino acids involved in amino acid binding of the tyrosine RNA synthetase. From this library, a TyrRS that recognizes only the artificial amino acid will be selected. The pTEF-MF/BCII10 vector will also be the target for site-directed mutagenesis, in order to introduce an 'amber' mutation at a well-defined and strategic position into the BCII10 nanobody. The EcTyrRS/$tRNA_{CUA}$ pair, selected from the library, will be used as an orthogonal pair in the yeast *S. cerevisiae* to introduce the 'click' functionalized amino acid in response to the amber codon in the BCII10 nanobody, resulting in a 'click' functionalized protein. In a later phase, this 'click' functionalized protein could possibly be coupled to a solid substrate, modified with the complementary 'click' functionality.

The hypothesis is that site-specific immobilization of proteins on a biosensor for example, will largely increase the sensitivity of the sensor. For example in the healthcare sector, this would be a considerable improvement. It would be possible to detect diseases in an earlier stage, which would

allow to start more rapidly with the treatment and consequently improve the prognosis and lower health care costs. Highly sensitive biosensors would also be helpful in the food industry, for the detection of contaminants and pathogens in food products, or in the environmental sector for monitoring environmental pollution via water and soil samples.

REFERENCES

[1] Saerens D, Huang L, Bonroy K, et al. Antibody fragments as probe in biosensor development. Review article. Sensors. 2008; 8: 4669-4686.

[2] Campàs M, Prieto-Simon, Marty JL. A review of the use of genetically engineered enzymes in electorchemical biosensors. Seminars in Cell & Developmental Biology. 2009; 20: 3-9.[3] Chambers JP, Arulanadam BP, Matta LL et al. Biosensor recognition elements. Current issues in molecular biology. 2008; 10: 1-12.

[3] Chambers JP, Arulanadam BP, Matta LL et al. Biosensor recognition elements. Current issues in molecular biology. 2008; 10: 1-12.

[4] Ghindilis AL, Atanasov P, Wilkins M et al. Immunosensors: electrochemical sensing and other engineering approaches. Biosensors & Bioelectronics. 1998; 13 (1): 113-131.

[5] Nairn R, Helbert M. Antigens and Antibody Structure. In: Nairn R, Helbert M. Immunology for medical students. Second edition. Philadelphia: Mosby Elsevier; 2007. p. 19-25

[6] Saerens D, Ghassabeh GH, Muyldermans S. Single-domain antibodies as building blocks for novel therapeutics. Current opinion in pharmacology. 2008; 8: 600-608.

[7] Hamers-Casterman C, Atarhuoch T, Muyldermans S et al. Naturally occurring antibodies devoid of light chains. Letters to Nature. 1993; 363: 446-448.

[8] Muyldermans S, Baral TN, Retamozzo CV et al. Camelid immunoglobulins and nanobody technology. Veterinary Immunology and Immunopathology. 2009; 128: 178-183.

[9] Muyldermans S, Baral TN, Retamozzo CV et al. Camelid immunoglobulins and nanobody technology. Veterinary Immunology and Immunopathology. 2009; 128: 178-183.

[10] Thomassen YE, Meijer W, Sierkstra L, et al. Large-scale production of VHH antibody fragments by

Saccharomyces cerevisiae. Enzyme and Microbial Technology. 2002; 30: 273-278.

[11] Loris R, Marinez-Rodriguez S, Conrath K, et al. Recognition of the active site of the β-lactamase BCII by a camelid V_HH, cAbBCII-10. Annual reports. 2006: 167-168.

[12] Frere JM. Beta-lactamases and bacterial resistance to antibiotics. Molecular Microbiology. 1995; 16: 385–395

[14] Rusmini R, Zhong Z, Feijen J. Protein immobilization strategies for protein biochips. Biomacromolecules. 2007; 8: 1775-1789.

[15] Vijayendran RA, Leckband DE. A quantitative assessment of heterogeneity for surface-immobilized proteins. Analytical Chemistry. 2001; 73: 471-480.

[16] MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. Science. 2000; 289: 1760-1763.

[17] Arenkov P, Kukhtin A, Gemmell A. Protein microchips: use for immunoassay and enzymatic reactions. Analytical Biochemistry. 2000; 278: 123-131.

[18] Peluso P, Wilson DS, Do D et al. Optimizing antibody immobilization strategies for the construction of protein microarrays. Analytical Biochemistry. 2003; 312: 113-124.

[19] Camarero JA. Recent developments in site-specific immobilization of proteins onto solid supports Review article. PeptideScience. 2007; 90 (3): 450-458.

[20] Hartmuth C.K, Finn M.G, Sharpless K.B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Review article. Angewandte Chemie. 2001; 40: 2004-2021.

[21] Hartmuth C, Kolb M, Finn B et al. Click chemistry - diverse chemical functions from a few good reactions. Angewandte Chemie International Edition. 2001; 40: 2004-2021

[22] Hein CD, Liu X, Wang D. Click Chemistry, A Powerful Tool for Pharmaceutical Sciences. Pharmaceutical Research. 2008, 25 (10): 2216-2230.

[23] Xu MQ, Perler FB. The mechanism of protein splicing and its modulation by mutation. The EMBO Journal. 1996; 15 (19): 5146-5153.

[24] Muir TW, Sondhi D, Cole PA. Expressed protein ligation: A general method for protein engineering. Proceedings of the National Academy of Science USA. 1998; 95: 6705-6710.

[25] Muir TW. Semisynthesis of proteins by expressed protein ligation. Annual Review of Biochemistry. 2003; 72: 249-289.

[26] Zhang Z, Smith BAC, Wang L, et al. A new strategy for site-specific modification of proteins in vivo. Biochemistry. 2003; 42: 6735-6746.

[27] Xie J, Schultz PG. Adding amino acids to the genetic repertoire. Current Opinion in Chemical Bioloty. 2005; 9: 548-554.

[28] Chen S, Schultz P, Brock A. An improved system for the generation and analysis of mutant proteins containing unnatural amino acids in *Saccharomyces cerevisiae.* Journal of Molecular Biology. 2007; 371: 112-122.

[29] Kohrer C, Xie L, Kellerer S, et al. Import of amber and ochre suppressor tRNAs into mammalian cells: A general approach to site-specific insertion of amino acid analogues into proteins. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98(25): 14310-14315.

[30] Sakamoto K, Hayashi A, Sakamota A, et al. Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. Nucleic Acids Research. 2002; 30(21):4692-4699.

[31] Chin JW, Cropp TA, Chu S, et al. Progress toward an expanded eukaryotic genetic code. Chemistry and Biology. 2003; 10: 511-519.

[32] Edwards H, Schimmel P. A bacterial amber suppressor in Saccharomyces cerevisiae is selectively recognized by a bacterial aminoacyl-tRNA synthetase. Molecular and cellular biology. 1990; 10: 1633-1641.

[33] Doctor BP, Loebel JE, Kellogg DA. Studies on the species specificity of yeast and E. Coli tyrosine tRNAs. Cold Spring Harbor Symposia on Quantitative Biology. 1966; 31: 543-548.

[34] Deiters A, Cropp AT, Mukherji M, et al. Adding amino acids with novel reactivity to the genetic code of Saccharomyces cerevisiae. Journal of the American Chemical Society. 2003; 125(39): 11782-11783.

[35] Ferrer E. Spotlight on targeting aminoacyl-tRNA synthetases for the treatment of fungal infections. Drug News & Perspectives. 2006; 19(6): 347-348.

SUPPLEMENTARY INFORMATION

Supplement I. Lysogeny Broth medium/agar plates

Bacto-tryptone (BD)	10g
Bacto-yeast extract (BD)	5g
NaCl (Sigma)	10g
NaOH (1M, pH 7)	1ml
MilliQ	up to 1L

If the LB-medium is used to pour agar plates, 15g of agar (BD) is added before autoclaving.
*BD = Becton, Dickinson and Company

Supplement II. Composition of PCR reaction mixtures and PCR programmes

Table S1. Reaction mixtures for the PCR of *E. coli* tRNA, TyrRS, the *S. cerevisiae* ADH1 promoter and NbBCII10. For tRNA, EcTyrRS and the ADH1 promoter, four different $MgCl_2$ concentrations were tested. NbBCII10 was amplified with 2mM $MgCl_2$. For *E. coli* tRNA and TyrRS, each $MgCl_2$ reaction mixture was made for 8 different temperatures. The ADH1 promoter and the BCII10 nanobody were amplified from plasmid DNA. For this, 1µl of DNA was added to the reaction mixture instead of 5µl. The amount of MQ was adjusted to a final volume of 50 µl.

	2mM $MgCl_2$ (µl)	4mM $MgCl_2$ (µl)	6mM $MgCl_2$ (µl)	8mM $MgCl_2$ (µl)
PCR buffer 10x	5	5	5	5
dNTP's (2,5mM each TaKaRa)	4	4	4	4
Forward primer (10µM)	1	1	1	1
Reverse primer (10µM)	1	1	1	1
$MgCl_2$ (50µM, Fermentas)	2	4	6	8
Polymerase (Fermentas)	0,25	0,25	0,25	0,25
DNA	5	5	5	5
MQ	31,75	29,75	27,75	25,75

Table S2. Program of the gradient PCR of *E. coli* tRNA

Temperature	Time	# Cycles
95°C	5'	1
95°C 45-70°C 72°C	30" 30" 30"	30
72°C	5'	1

Table S3. Program of the gradient PCR of *E. coli* TyrRS

Temperature	Time	# Cycles
95°C	5'	1
95°C 35-60°C 72°C	30" 30" 1'30"	30
72°C	5'	1

Table S4. Program of the PCR of *S. cerevisiae* ADH1 promoter. For the amplification reaction with Pfu, the elongation time was prolonged to 4 minutes.

Temperature	Time	# Cycles
95°C	5'	1
95°C 60°C 72°C	30" 30" 1'30"	30
72°C	5'	1

Table S5. Program of the PCR of the BCII10 nanobody

Temperature	Time	# Cycles
95°C	5'	1
95°C 50°C 72°C	30" 30" 1'30"	30
72°C	5'	1

Supplement III. Sequences

TyrRS

TCATAACGA**GAATTC**ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGGGGCTGGTAGCCCAGGT
GACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCGCTCTATTGCGGCTTCGATCCTA
CCGCTGACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGCCACA
AGCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAA
GCTGAACACCGAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATTT
CGACTGTGGAGAAAACTCTGCTATCGCGGCGAACAACTATGACTGGTTCGGCAATATGAATGTGCTGACCTT
CCTGCGCGATATTGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAA
CCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTGTTGCAGGGTTATGACTTCGCCTGTCTG
AACAAACAGTACGGTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGAC
CTGACCCGTCGTCTGCATCAGAATCAGGTGTTTGGCCTGACCGTTCCGCTGATCACTAAAGCAGATGGCACC
AAATTTGGTAAAACTGAAGGCGGCGCAGTCTGGTTGGATCCGAAGAAAACCAGCCCGTACAAATTCTACCAG
TTCTGGATCAACACTGCGGATGCCGACGTTTACCGCTTCCTGAAGTTCTTCACCTTTATGAGCATTGAAGAGA
TCAACGCCCTGGAAGAAGAAGATAAAAACAGCGGTAAAGCACCGCGCGCCCAGTATGTACTGGCGGAGCAG
GTGACTCGTCTGGTTCACGGTGAAGAAGGTTTACAGGCGGCAAAACGTATTACCGAATGCCTGTTCAGCGGT
TCTTTGAGTGCGCTGAGTGAAGCGGACTTCGAACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGA
AAAGGGCGCAGACCTGATGCAGGCACTGGTCGATTCTGAACTGCAACCTTCCCGTGGTCAGGCACGTAAAA
CTATCGCCTCCAATGCCATCACCATTAACGGTGAAAAACAGTCCGATCCTGAATACTTCTTTAAAGAAGAAGA
TCGTCTGTTTGGTCGTTTTACCTTACTGCGTCGCGGTAAAAAGAATTACTGTCTGATTTGCTGGAAATAA**GCG
GCCGC**ACGTAGTAA

- TyrRS= **1275bp**
- Extra nucleotides from primers= 32bp, indicated in blue. The introduced restriction sites are indicated in bold. 1275bp + 32bp= **1307bp**
- Taq-polymerase= during the amplification, 1 extra adenosine is added by Taq polymerase. 1307bp + 1= **1308bp** (Fig. S1)
- pCR2.1 vector= 15bp between the PCR product and the flanking EcoRI restriction sites. 1308bp + 15bp= **1323bp**
- Cloning in pEcTyrRS= 1275bp + 11bp (from the introduced restriction sites)= **1286bp**

Fig. S1. Cloning site of the pCR2.1 vector (TA cloning kit, Invitrogen).

$tRNA_{CUA}$

GGGGGG**ACCGGT**AAGCTTCCCGATAAGGGAGCAGGCCAGTAAAAGCATTACCCCGTGGTGGGGTTCCCGA
GCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCAT
TTTTTTCAAAAGTCCCTGAACTTCCC**GCTAGC**GCCGCC

- $tRNA_{CUA}$= **156bp** (including pre-tRNA)
- Extra nucleotides from primers= 24bp, indicated in blue. The introduced restriction sites are indicated in bold. 156bp + 24bp= **180bp**
- Taq-polymerase= during the amplification, 1 extra adenosine is added by Taq polymerase. 180bp + 1= **181bp** (Fig. S1)
- pCR2.1 vector= 15bp between the PCR product and the flanking EcoRI restriction sites. 181bp + 15bp= **196bp**
- Cloning in pEcTyrRS= 156bp + 10bp (from the introduced restriction sites)= **166bp**

ADH1-promoter

GGGGGG**ACCGGT**CGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAAGGCAAAAG
ACAAATATAAGGGTCGAACGAAAAATAAAGTGAAAAGTGTTGATATGATGTATTTGGCTTTGCGGCGCCGAA
AAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGGACCCGCGCTCTTGCCGGCCCGGCGATA
ACGCTGGGCGTGAGGCTGTGCCCGGCGGAGTTTTTTGCGCCTGCATTTTCCAAGGTTTACCCTGCGCTAAGG
GGCGAGATTGGAGAAGCAATAAGAATGCCGGTTGGGGTTGCGATGATGACGACCACGACAACTGGTGTCAT
TATTTAAGTTGCCGAAAGAACCTGAGTGCATTTGCAACATGAGTATACTAGAAGAATGAGCCAAGACTTGCG
AGACGCGAGTTTGCCGGTGGTGCGAACAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATAACCGCT
AGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTACATACAACACTGGAA
ATGGTTGTCTGTTTGAGTACGCTTTCAATTCATTTGGGTGTGCACTTTATTATGTTACAATATGGAAGGGAACT
TTACACTTCTCCTATGCACATATATTAATTAAAGTCCAATGCTAGTAGAGAAGGGGGGTAACACCCCTCCGCG
CTCTTTTCCGATTTTTTTCTAAACCGTGGAATATTTCGGATATCCTTTTGTTGTTTCCGGGTGTACAATATGGA
CTTCCTCTTTTCTGGCAACCAAACCCATACATCGGGATTCCTATAATACCTTCGTTGGTCTCCCTAACATGTAG
GTGGCGGAGGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCTAAACAAGACTACACCAAT
TACACTGCCTCATTGATGGTGGTACATAACGAACTAATACTGTAGCCCTAGACTTGATAGCCATCATCATATC
GAAGTTTCACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCAACTTCTTTTCTTTTTTT
TTCTTTTCTCTCTCCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAAAATGATGGAAGACACTAAAGG
AAAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGAGGGGTATCTTCGAACAC
ACGAAACTTTTTCCTTCCTTCATTCACGCACACTACTCTCTAATGAGCAACGGTATACGGCCTTCCTTCCAGTT
ACTTGAATTTGAAATAAAAAAAGTTTGCCGCTTTGCTATCAAGTATAAATAGACCTGCAATTATTAATCTTTTG
TTTCCTCGTCATTGTTCTCGTTCCCTTTCTTCCTTGTTTCTTTTTCTGCACAATATTTCAAGCTATACCAAGCAT
ACAATCAACT**GAATTC**CCCCCC

- ADH1-promoter= **1443bp**
- Extra nucleotides from primers= 24bp, indicated in blue. The introduced restriction sites are indicated in bold. 1443bp + 24bp= **1467bp**
- Taq-polymerase= during the amplification, 1 extra adenosine is added by Taq polymerase. 1467bp + 1= **1468bp** (Fig. S1)
- pCR2.1 vector= 15bp between the PCR product and the flanking EcoRI restriction sites. 1468bp + 15bp= **1483bp**
- Cloning in pEcTyrRS= 1443bp + 10bp (from the introduced restriction sites)= **1453bp**

NbBCII10

GGC**GAATTC**CCATGGCCCAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTCGGTGCAGGCTGGAGGGTCCC
TGAGACTCTCCTGTACAGCCTCTGGAGGCTCTGAATACAGCTACAGTACATTTTCCTTGGGCTGGTTCCGCCA
GGCTCCAGGGCAGGAGCGTGAGGCGGTCGCGGCAATTGCGAGTATGGGTGGCCTCACATACTACGCCGAC
TCCGTGAAGGGCCGATTCACCATCTCCCGAGACAACGCCAAGAACACGGTGACTCTGCAGATGAACAACCTG
AAACCTGAGGACACGGCCATCTATTACTGTGCGGCGGTGCGTGGTTATTTTATGCGACTACCCTCGTCACAT
AACTTTCGCTACTGGGGGCAGGGGACCCAGGTCACCGTCTCCTCACGCGGCCGCCACCACCATCACCATCA
CTAA**CTCGAG**CCC

- NbBCII10= **420bp** (including the 6-His tag in green)
- Extra nucleotides from primers= 18bp, indicated in blue. The introduced restriction sites are indicated in bold. 420bp + 18bp= **438bp**
- Taq-polymerase= during the amplification, 1 extra adenosine is added by Taq polymerase. 438bp + 1= **439bp** (Fig. S1)
- pCR2.1 vector= 15bp between the PCR product and the flanking EcoRI restriction sites. 439bp + 15bp= **454bp**
- Cloning in pEcTyrRS= 420bp + 10bp (from the introduced restriction sites)= **430bp**

Supplement IV. Restriction digestion reactions

Tabel S6. Resctriction digestion of *E. coli* TyrRS and pESC-TRP. *E. coli* TyrRS and the plasmid pESC-TRP were digested by the restriction enzymes EcoRI and NotI.

	TyrRS (µl)	pESC-TRP (µl)
MQ	21	23
10X Buffer O	3	3
DNA	3 (±1µg)	1µl (±0,5µg)
Restriction enzyme EcoRI	1	1
Restriction enzyme NotI	2	2

Table S7. Restriction digestion of pCR2.1/tRNA$_{CUA}$ and pEcTyrRS after ethanol precipitation. The DNA was digested in two steps.) First digestion step with restriction enzyme NheI. B) Second digestion step with restriction enzyme BshTI.

A	pCR2.1/tRNA$_{CUA}$ (µl)	pEcTyrRS (µl)
MQ	14	14
10x Buffer Tango™	2	2
DNA	3 (10,2 µg)	3 (6 µg)
Restriction enzyme NheI	1	1

B	pCR2.1/tRNA$_{CUA}$ (µl)	pEcTyrRS (µl)
10X Buffer Tango™	5,5	5,5
DNA (digested with NheI)	50	50
Restriction enzyme BshTI	1	1

Supplement V. Cloning reactions

Table S8. Cloning reaction of TyrRS in pESC-tRP. Three different vector:insert ratios were tested. A negative control (vector without insert) was included. The reaction mixtures were incubated overnight at 22°C.

	3:1(µl)	1:1 (µl)	1:3 (µl)	Neg. cntr. (µl)
MQ	13	15	13	16
10x Buffer	2	2	2	2
T4 ligase	1	1	1	1
TyrRS	1	1	3	/
pESC-TRP	3	1	1	1
	20	**20**	**20**	**20**

Table S9. Cloning of $tRNA_{CUA}$ in pEcTyrRS. Three different vector:insert ratios were rested. A negative control (vector without insert) was included. The reaction mixtures were incubated overnight at 22°C.

	1:1(µl)	1:3 (µl)	1:6 (µl)	Neg. cntr. (µl)
MQ	13	15	13	16
10x Buffer	2	2	2	2
T4 ligase	1	1	1	1
$tRNA_{CUA}$	4,5	13,5	27	/
pEcTyrRS	4,5	4,5	4,5	4,5
	20	**30**	**40**	**20**

Table S10. Cloning of the ADH1 promoter in pEcTyrRS/$tRNA_{CUA}$. Three different vector:insert ratios were rested. A negative control (vector without insert) was included. The reaction mixtures were incubated overnight at 22°C.

	1:1(µl)	1:3 (µl)	1:6 (µl)	Neg. cntr. (µl)
MQ	8	8	3,5	12,5
10x Buffer	2	3	4	2
T4 ligase	1	1	1	1
ADH1 promoter	4,5	13,5	27	/
pEcTyrRS/$tRNA_{CUA}$	4,5	4,5	4,5	4,5
	20	**30**	**40**	**20**

Table S11. Cloning of the BCII10 nanobody in the pTEF-MF yeast expression vector. Three different vector:insert ratios were rested. A negative control (vector without insert) was included. The reaction mixtures were incubated overnight at 22°C.

	1:1(µl)	1:3 (µl)	1:6 (µl)	Neg. cntr. (µl)
MQ	8	15	13	16
10x Buffer	2	2	2	2
T4 ligase	1	1	1	1
BCII10	4,5	13,5	27	/
pTEF-MF	4,5	4,5	4,5	4,5
	20	**30**	**40**	**20**

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
In vivo site-specific modification of proteins with artificial click functionalized amino acids

Richting: **master in de biomedische wetenschappen-klinische moleculaire wetenschappen**
Jaar: **2011**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Baré, Birgit

Datum: **14/06/2011**