Carina Andrea Sommer, BSc.

NcTrp as building blocks: Preparative *in vitro* conversion and *in vivo* biosynthesis with parallel incorporation into a model protein

MASTER‘S THESIS

to achieve the university degree of

Diplom-Ingenieurin

Master’s degree programme: Biotechnology

submitted to

Graz University of Technology

Supervisor
UNIV.-PROF. Dipl.-Ing. Dr. techn. Helmut Schwab

Institute of Molecular Biotechnology
Graz University of Technology

Dipl.-Ing. Dr. techn. Birgit Wiltschi
Austrian Centre of Industrial Biotechnology

Graz, August 2016

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material, which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis dissertation.

Date

Signature

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank Univ.-Prof. Dipl.-Ing. Dr.techn. Helmut Schwab for being the main examiner of this work, and without whose permission it would not have been possible to conduct my research within the Junior Group Synthetic Biology. Furthermore, I would like to thank the Austrian Centre of Industrial Biotechnology and FFG for funding this project.

I would also like to thank my supervisor, Dipl.-Ing. Dr. techn. Birgit Wiltschi, for giving me the opportunity to work in the field of incorporating non-canonical amino acids, and of course for her continuous support and guidance throughout the overall procedure of my master thesis.

I would also like to extend my gratitude to Dipl.-Ing. Dr. techn. Corinna Odar, for working with me on this project, for sharing her expertise, and for her support in the lab. I also want to thank every member of the Junior Group Synthetic Biology for their support in discussing lab issues. Special thanks go Patrik Fladischer, MSc, who was always with me and inspired me with our late-night discussions about new and future achievements in the field of Synthetic Biology. Furthermore, I want to thank Manuela Killinger, BSc, who is an expert in cloning, and who was always very ambitious in helping me with cloning issues that appeared to be insolvable in the beginning.

I would also like to thank my dearest friends Chrisi, Mohamed and Paul, who are always with me. Special thanks go to Lena for her steady support and her patience. She makes me laugh even in the hard times. Last, but certainly not least, I would like to thank family. I want to thank my sister for being awesome and my grandparents for teaching me to have the right set of priorities. I want to thank my parents for making it possible for me to study, and for their support during the years.

ABSTRACT

The incorporation of non-canonical amino acids (ncAA) in proteins is one of the main topics of Synthetic Biology. This strategy enables the introduction of novel chemical and structural properties into proteins. Special attention is drawn to the incorporation of Trp analogs (ncTrp) into proteins, due to the extraordinary status of Trp in the pool of natural amino acids. Trp is encoded by a single triplet codon (UGG); it is a relatively rare amino acid (Ross et al., 1997; Lepthien et al., 2008), and is involved in molecular interactions such as π-π stacking, hydrogen bonding, and cation-π interactions (Budisa et al., 2004). The spectroscopic properties of Trp, which are based on the indole side chain, mainly contribute to the spectroscopic properties of proteins (Lakowicz, 1983; Lepthien et al., 2008). However, a significant downside to ncTrps is that they are expensive, and commercially not available. To overcome this problem, we first established an *in vitro* procedure for the preparative production of ncTrp (Winn et al., 2008). This was achieved by the enzymatic condensation of indole analogs with Ser to ncTrp, which was catalyzed by the tryptophan synthase from *Salmonella enterica* subsp. *enterica* serovar Typhimurium (Miles, 2001). In place of all available Trp analogs, we chose 5OH-indole and 5F-indole for the preparative production of the corresponding ncTrp.

Based on the previously published supplementation based incorporation (SPI) method for the residue-specific incorporation of ncAA in proteins (Budisa et al., 1998), we established an *in vivo* method for the parallel enzymatic synthesis of the ncTrp, and the residue-specific incorporation into a recombinantly expressed protein in *E. coli*. With this method, we achieved the full labeling of the target protein, the enhanced cyan fluorescent protein (ECFP) with ncTrp.

KURZFASSUNG

Der Einbau von nicht-kanonischen (nc) Aminosäuren in Proteine ist eine wichtiges Teilgebiet der Synthetischen Biologie und ermöglicht die chemische und strukturelle Veränderung von Proteinen. Aufgrund der Einzigartigkeit des Tryptophans (Trp) unter den kanonischen Aminosäuren, kommt dem Austausch von Trp zu nicht-kanonischen Trp Analoga (ncTrp) ganz besondere Aufmerksamkeit zuteil. Trp ist eine seltene Aminosäure (Ross et al., 1997; Lepthien et al., 2008) und sie wird nur von einem einzigen Triplett-Codon codiert. Sie ist außerdem an wichtigen molekularen Interaktionen beteiligt (Budisa et al., 2004). Des Weiteren trägt Trp durch seine spektroskopischen Eigenschaften, die auf den Indolring zurückzuführen sind, hauptsächlich zu den spektroskopischen Eigenschaften von Proteinen bei (Lakowicz, 1983; Lepthien et al., 2008). Der einzige Nachteil beim Einbau von ncTrp in Proteine besteht darin, dass diese sehr teuer oder kommerziell nicht verfügbar sind.

Um dieses Problem zu lösen, haben wir eine *in vitro* Synthese (Winn et al., 2008) für die präparative Herstellung von ncTrp etabliert. Diese Synthese besteht in der Kondensation von Indolanalogen und Serin zu ncTrp, die von der Tryptophan Synthase aus *Salmonella enterica* subsp. *enterica* serovar Typhimurium katalysiert wird (Miles, 2001). Stellvertretend für alle verfügbaren nicht-kanonischen Indole, haben wir 5OH-Indol und 5F-Indol für die Kondensation zu den entsprechenden ncTrp verwendet.

Basierend auf der bereits publizierten *supplementation based incorporation* (SPI) Methode für den aminosäure-spezifischen Einbau von nicht-kanonischen Aminosäuren in Proteine (Budisa et al., 1998), haben wir ein eine *in vivo* Methode entwickelt, um die ncTrp in der *E. coli* Zelle zu synthetisieren und diese direkt in ein rekombinant exprimiertes Modellprotein einzubauen. Als Modellprotein diente in dieser Studie das „enhanced cyan fluorescent protein“ (ECFP), bei dem es uns gelang alle im Protein vorkommenden Trp durch ncTrp zu ersetzen.

1 INTRODUCTION

In the last couple of years, the incorporation of non-canonical amino acids (ncAAs) into proteins became a very popular field of research. This is based on the possibility of expanding the chemical properties by substituting a canonical amino acid (cAA) for the ncAA. The method for the incorporation of ncAAs in *E. coli* can be divided into two principles; site-specific incorporation and residue-specific incorporation (Walsh et al., 2013; Liu & Schultz, 2010; Betts & Russel, 2003).

Site-specific incorporation is a method for the incorporation of ncAAs at a defined position within a target protein. This can be achieved with the strategy of stop codon suppression, or the usage of a quadruplet codon. These strategies require a heterologous orthogonal pair consisting of a suppressor tRNA and an aminoacyl-tRNA synthetase, which is specific for an ncAA. The aminoacyl-tRNA synthetase often has to be engineered in order to recognize an ncAA of choice (Zheng & Kwon, 2012; Hoesl & Budisa, 2012; Liu & Schultz, 2010).

Residue-specific incorporation enables a global replacement of one cAA with its non-canonical analog. The residue-specific incorporation of an ncAA within a target protein can be achieved by *in vitro* or *in vivo* protein synthesis. Cell free protein synthesis separates the cell growth from protein expression, and it is advantageous if certain reagents, the ncAA, or the expressed protein are harmful for the host (Singh-Blom et al., 2014; Hong et al., 2014; Hoesl & Budisa, 2012). Initial experiments of *in vivo* incorporation of ncAAs began in the 1950s and 1960s (Brawerman & Yčas, 1957; Schlesinger, 1968). Experiments in this field were not efficient, until the introduction of auxotrophic expression hosts (Minks et al., 1999). In such strains, an essential gene for the expression of the cAA is knocked out in the genome. This enables to set the titer of the cAA by the composition of the medium (Datsenko & Wanner, 2000). The concentration of the cAA within the cells should be at a minimum, because the ncAA and the cAA compete to interact with the corresponding host aminoacyl-tRNA synthetase. Due to the substrate tolerance of the endogenous aminoacyl-tRNA synthetase for sterochemically similar amino acids, an additional heterologously expressed orthogonal pair of an aminoacyl-tRNA synthetase and a suppressor tRNA is not necessary (Zheng & Kwon, 2012; Budisa, 2004; Kiick et al., 2000). To bypass the bottleneck of the stereochemical similarity, the substrate tolerance of the aminoacyl-tRNA synthetase can be broadened by site directed mutagenesis (Johnson et al., 2010).

The established strategies for *in vivo* incorporation of ncAAs are grounded on the supplementation-based incorporation method (SPI) and can be implemented with a limited concentration of the cAA in the medium or with medium shift. The medium shift approach is based on an initial cell growth in the presence of all 20 cAAs. For protein expression, the cells are harvested

and re-suspended in fresh synthetic medium, in which one cAA is exchanged with its non-canonical counterpart. This strategy ensures that the ncAA does not have to compete with the cAA for interacting with the aminoacyl-tRNA synthetase. Mostly, cell growth is performed in a complex medium to reach a high cell density before the medium is exchanged (Van Hest et al., 2000; Kiick et al., 2000). Alternatively, full synthetic medium can be used.

In contrast to the medium shift method, cell growth and target protein expression can be performed in synthetic medium if a limited concentration is used of the cAA that is replaced. The cells grow until the cAA is consumed entirely. At this point, the cell growth stalls because of the exhaustion of the limited cAA in the medium. Now, the ncAA is added to the medium, and the protein expression is induced (Budisa et al., 1998). This strategy ensures that the cAA is no longer present in the medium, when the ncAA is introduced. Since the exchange of a cAA to an ncAA facilitates an alteration of the protein properties, several cAAs were previously exchanged: Leu, His, Ile, Val, Trp, Phe, Met, Cys, Lys, Pro, Tyr, Trp (Zheng & Kwon, 2012).

To name but a few prominent examples, this method was used to improve the stability of the green fluorescent protein from *Aequoria victoria*, and of ubiqutin, by a global replacement of Pro to 4-flouroproline (Steiner et al., 2008; Crespo & Rubini, 2011). Enhanced activity, enhanced substrate tolerance, changes in the optimal temperature, as well as changes in the optimal pH, could all be achieved by the exchange of Met, Pro, Phe and Tyr, in a lipase from *Thermoanaerobacter thermohydrosulficus* (Hoesl et al., 2011). However, the globally replaced cAA by its non-canonical counterpart can also be disadvantageous for a protein. The incorporation of an ncAA can also lead to denaturation, or loss of activity due to partial unfolding (Zheng & Kwon, 2012). For this reason, the residue-specific incorporation method of Met analogs was previously used for protein conformation studies (Wolschner et al., 2009).

1.1 TRP AS A TARGET FOR THE EXCHANGE TO TRP ANALOGS

Trp is an attractive target for the exchange to a non-canonical Trp (ncTrp), because of its extraordinary status within the pool of natural amino acids. It is encoded by a single triplet codon (UGG), and it is a relatively rare amino acid, with an abundance of approximately 1% in proteins (Ross et al., 1997; Lepthien et al., 2008). Trp is also involved in molecular interactions such as π-π stacking, hydrogen bonding, and cation-π interactions, which play a crucial role in the folding, activity and stability of the proteins (Budisa et al., 2004). An advantage in the use of Trp is its spectroscopic properties, which are based on the indole side chain of the amino acid. Therefore, Trp mainly contributes to the spectroscopic properties of the proteins (Lakowicz, 1983; Lepthien et al., 2008). However, only a couple of Trp analogs are commercially available and they are very expensive.

To overcome this problem, the analogs can be enzymatically synthesized using tryptophan synthase (TrpS). The enzyme is an $\alpha_2\beta_2$ tetramer, and forms an intermolecular tunnel. The α subunit catalyzes a reversible retroaldol cleavage of indole-3-glycerol phosphate to indole and D-glycerolandehyde-3-phosphate (Figure 1, panael A). The β subunit catalyzes the condensation of indole with Ser to Trp (Figure 1, panael B). For condensation, the β subunit requires pyridoxal 5'-phosphate (PLP) as a co-factor (Miles, 2001; Dunn et al., 2008).

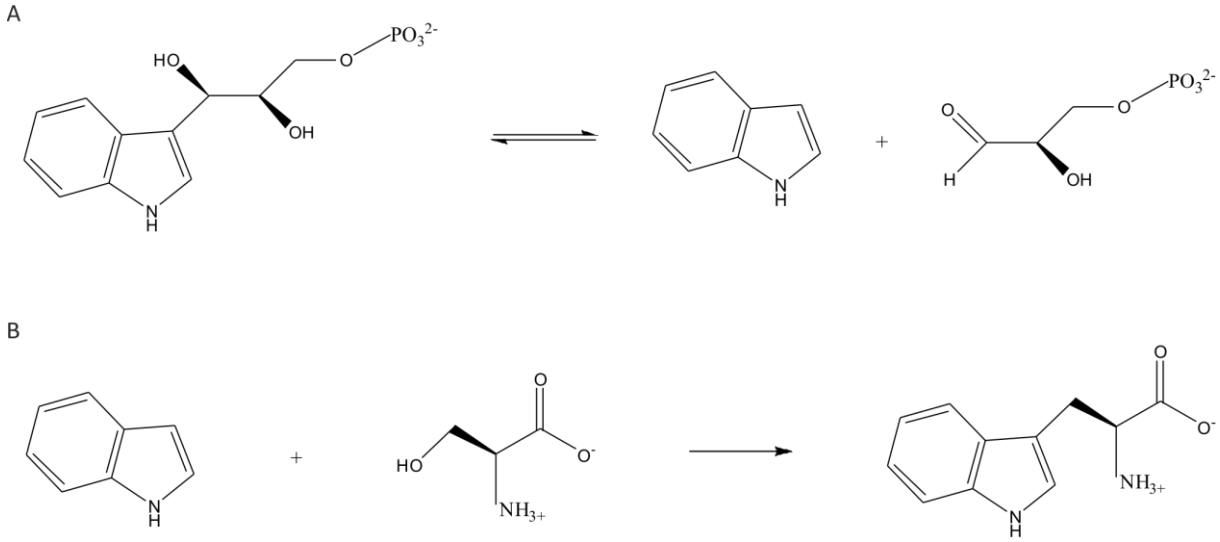


Figure 1: Catalytic activity of the TrpS.
Panel A shows the reversible retroaldol cleavage of indole-3-glycerol phosphate to indole and D-glycerolandehyde-3-phosphate, which is catalyzed by the α subunit. Panel B shows the condensation of indole with Ser to Trp, which is catalyzed by the β subunit (Miles, 2001).

Initial experiments for the enzymatic synthesis of Trp analogs from indole analogs with the tryptophan synthase from *E. coli* (*Ec*TrpS) began in the 1970s (Wilcox, 1974). Until today, various Trp analogs have been synthesized with *Ec*TrpS from the corresponding indole analogs (Phillips, 2004). However, other TrpSs have also been used for the enzymatic synthesis of Trp analogs, for instance, the tryptophan synthase from *Salmonella enterica*. With this enzyme, it is even possible to synthesize L-halotryptophans in a cell-free way (Smith et al., 2014), as well as in a biofilm (Perni et al., 2013). Most commonly, the TrpS from *Salmonella enterica* subsp. *enterica* serovar Typhimurium (*St*TrpS) has been used (Miles, 2001). The *St*TrpS is well characterized, and the synthesis of aza-, thio-, fluoro-, chloro and methyltryptophan was previously reported with this enzyme (Sloan & Phillips, 1992; Phillips et al., 1995; Goss et al., 2006). The synthesized Trp analogs can be directly applied for the residue-specific incorporation, by conducting the cell-free incorporation method or the SPI method. This strategy was chosen for the enzymatic synthesis of β-selenolo[3,2-*b*]pyrrolyl-alanine from seleno[3,2-*b*]pyrrole (Welch & Phillips, 1999). The incorporation of selenium into the aromatic system of the Trp should provide a stable and planar Trp analog. The incorporation of a selenium-containing Trp analog into proteins can solve the phase problem in crystallography (Bae et al., 2001).

Due to the low occurrence of Trp in proteins, it seems to be a promising alternative to SeMet and offers a site-specific resonance center for X-ray crystallography by the MAD method, stated by Bae et al. (2001).

1.2 ENHANCED CYAN FLUORESCENT PROTEIN (ECFP) AS MODEL PROTEIN

ECFP is one of several variants of the green fluorescent protein (GFP) from *Aequoria victoria* (Chalfie et al., 1994) which were produced by classical protein engineering methods (Heim et al., 1994; Heim & Tsien, 1996; Tsien, 1998). Their fluorescence depends on the formation of a chromophore of the residues at the positions 57 - 65. At position 66, an aromatic amino acid is crucial (Tsien, 1998). A Trp at position 66 provokes cyan fluorescing properties, but the exchange of Trp at position 66 for an amino acid with an electron-donating amino group leads to a red shift in fluorescence. This Stoke shift is caused by a lone pair of electrons of the nitrogen atom in the $4NH_2$-Trp, which are conjugated to the delocalized π-system of Trp. Thus, the red shift in fluorescence is a result of intramolecular migration of the charge. This spectral change is visible as a golden fluorescence and is responsible for the name, gold fluorescent protein (ECFP[$4NH_2$-Trp]) (Budisa et al., 2002; Bae et al., 2003).

1.3 THESIS OBJECTIVES

In this thesis, we focused on ncTrp due to the important role of Trp in proteins. As aforementioned, the high costs of the ncTrp are the main drawback of working with these compounds. Therefore, the first objective was to address the preparative enzymatic production of ncTrp, in order to enable convenient use of ncTrp for the pre-established incorporation strategies. This enzymatic conversion was catalyzed by *St*TrpS in an *in vitro* assay, which was described by Winn et al. (2008). To demonstrate the effective *in vitro* condensation of non-canonical indole with Ser to ncTrp, we chose 5F-indole and 5OH-indole as model substrates. The 5F-indole and 5OH-indole were condensed with Ser to 5F-Trp and 5OH-Trp. The 5F-indole was chosen since the successful conversion of fluorinated Trp had already been described by Winn et al. (2008), and the further incorporation of fluorinated Trps is of common interest (Budisa & Pal, 2004). The 5OH-indole was chosen due to its unique absorption and fluorescence emission properties (i.e. a 20 nm red-shift in fluorescence compared to Trp). The incorporation of 5OH-Trp into proteins enables a selective excitation in the range of 315 to 320 nm (Budisa & Pal, 2004; Ross et al., 1997). Based on this, it proved to be a great advantage that 5OH-Trp and 5F-Trp could be enzymatically synthesized in a preparative way for a convenient use.

The second, and also main objective, was to establish an efficient incorporation protocol in *E. coli* for the *in vivo* biosynthesis and the parallel incorporation of ncTrp analogs into the heterologously

expressed enhanced cyan fluorescent protein (ECFP), based on the pre-established SPI method (Budisa et al., 1998). To establish such a protocol, $4NH_2$-indole was chosen as a model substrate to convert it to $4NH_2$-Trp, and to incorporate it simultaneously into ECFP. The conversion of $4NH_2$-indole to $4NH_2$-Trp and the incorporation of $4NH_2$-Trp into ECFP was chosen, because it leads to simple and visible proof of a successful incorporation. To show that this method is also applicable for other indole analogs, we conducted the *in vivo* synthesis and incorporation protocol with 5OH-indole and 5F-indole. These two analogs were chosen because the incorporation of 5OH-Trp into fluorescent proteins from *Aequoria victoria* has not been accomplished before (Budisa et al., 2004), and the incorporation of 5F-Trp into proteins is considered attractive, due to the improved properties of fluoro-variants (Minks et al., 1999; Bonskill & Wong, 1988; Budisa et al., 2004; Parsons et al., 1998).

2 MATERIAL AND METHODS

2.1 CLONING AND PREPARATION OF THE EXPRESSION STRAIN

The functional parts of the plasmid p15A-AraC-H6-*St*TrpS (Supplementary Figure 1, panel A) for the *St*TrpS, originated from three different plasmids. The promoter P_{araBAD}, together with *araC*, were prepared by digesting the pQEara-gdb (Marzluf, 2011) with EcoRI and NotI (Thermo Fisher Scientific, Waltham, MA). The p15A origin of replication, the *kanR* resistance marker and the rrnB terminator originated from the pLEU^{fbr} (Anderhuber et al., 2016), and were released by cutting with the same restriction enzymes. In both cases, 3.5 µg plasmid DNA were used for the digestion, and the DNA fragments were gel-purified using the Promega Wizard Gel Purification kit (Madison,WI), according to the manufacturer's protocol. The 300 ng fragment comprising the p15A origin of replication, the *kanR* resistance marker, the rrnB terminator, and the resulting 60 ng of the fragment with P_{araBAD} and *araC* were ligated with T4 DNA ligase (Thermo Fisher Scientific), in a total volume of 20 µl. These fragments formed the p15A-AraC backbone plasmid. 5.5 µg of the p15A-AraC backbone plasmid were digested for 1 h at 37 °C with EcoRI for linearization, and dephosphorylated by FastAP Thermosensisitive Alkaline Phosphatase (Thermo Fisher Scientific). The genes, which encode the *St*TrpS, the *trpA* and the hexahistidine-tagged *trpB*, originated from plasmid pSTWS2H (Hofmann, 2012). The *trpA* and the N-terminal hexahistidine-tagged *trpB* were amplified by PCR (Primer 1 and Primer 2; IDT Inc., Coralville, IA; Supplementary Table 1) with the TaKaRa Ex Taq® DNA polymerase (Takara Bio, Saint-Germain-en-Laye, France). The TaKaRa Ex Taq® DNA polymerase has a proofreading function. The resulting DNA fragments of the PCR amplification were gel-purified, and inserted into the linearized p15A-araC backbone plasmid by Gibson assembly (Gibson et al., 2009). Therefore, the Gibson assembly mix was used with 5.5 µg of the backbone plasmid and 300 ng of the DNA sequence, coding for *St*TrpS. The Gibson assembly consisted of: 25 U/mL Phusion® High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA); 4000 U/mL Taq DNA ligase (New England Biolabs); 4 U/mL T5 exonuclease (BioZym, Hessisch Oldendorf, Germany); 5% PEG-8000 (Sigma-Aldrich, St. Louis, MO); 1 mM NAD; 0.2 mM dNTPs; 10 mM $MgCl_2$; and 10 mM DTT in 100 mM Tris/Cl (all Thermo Fisher Scientific).

The expression construct with the ECFP coding sequence was prepared by digestion of plasmid pQE80L (Qiagen, Venlo, Netherlands) with XhoI and XbaI (Thermo Fisher Scientific), and the insertion of ECFP by Gibson assembly (kindly provided by P. Fladischer, ACIB GmbH, Austria). The ECFP coding sequence originated from pET30a-ECFP (kindly provided by K. Steiner, ACIB GmbH, Austria) and was amplified by PCR (Primer 3 and Primer 4; IDT Inc., Coralville, IA; Supplementary Table 1) with the

TaKaRa Ex Taq® DNA polymerase (Takara Bio). The plasmid map of the resulting pQE80L-H6-ECFP is shown in Supplementary Figure 1, panel B. The DNA sequences of the pQE80L-H6-ECFP and the p15A-araC-H6-*St*TrpS can be found in Supplementary Sequence 1 and Supplementary Sequence 2, respectively.

100 ng of the p15A-AraC-H6-*St*TrpS were transformed into the Trp auxotrophic BL21(DE3) gold Δ*trpC*::FRT strain (kindly provided by N. Anderhuber., ACIB GmbH, Austria) by electroporation (Seidman et al., 2001) with a voltage of 1.5 V and a resistance of 100 Ω. The resulting BL21(DE3) gold ΔtrpC::FRT{p15A-araC-H6-StTrpS} strain was further electroporated (1.5 V and 100 Ω) with pQE80L-H6-ECFP, which yielded the expression strain BL21(DE3) gold Δ*trpC*::FRT {p15A-AraC-H6-*St*TrpS; pQE80L-H6-ECFP}.

2.2 *St*TrpS EXPRESSION AND *IN VITRO* CONVERSION

2.2.1 EXPRESSION AND IMMOBILIZATION OF THE *St*TrpS

The Trp auxotrophic *E. coli* expression strain, BL21(DE3) Δ*trpC*::FRT {p15A-AraC-H6-*St*TrpS; pQE80L-H6-ECFP}, was cultivated in 500 ml Lennox LB medium (Roth, Karlsruhe, Germany) at 37 °C and 120 rpm. The *St*TrpS was induced at D_{600} of 0.8 with 2 mg/L arabinose (Sigma-Aldrich). The expression was performed overnight. The cells were harvested and resuspended in 20 mM Tris/Cl, 150 mM NaCl, pH 7.8. 1 mg/mL Lysozyme and 0.003 mg/mL DNAse I were added (all Roth) for a better cell lysis, which was performed by sonication (output control 8, duty cycle 70 - 80 %, 6 min). The lysed cells were centrifuged at 20000 g, 4 °C for 40 min to separate soluble and insoluble cell components. The soluble *St*TrpS was immobilized overnight at 4°C on 2 ml nickel-nitrilotriacetic acid (Ni-NTA) resin (1 ml bed volume) (Qiagen, Hilden, Germany). The Ni-NTA resin with the immobilized *St*TrpS was centrifuged at 66 g (4°C) for 5 min, to separate the resin from the cell lysate. The Ni-NTA resin was washed with 50 mM $NaH_2PO_4*H_20$, 300 mM NaCl, 20 mM imidazole (all Roth), until the Bradford protein assay (BioRad, Vienna, Austria) did not detect any proteins in the flow through. A final washing step with of 0.1 M K-PO_4 buffer (Roth) pH 7.8 was performed to remove the imidazole.

2.2.2 ENZYMATIC CONVERSION OF 5OH-INDOLE TO 5OH-TRP AND 5F-INDOLE TO 5F-TRP

The *in vitro* enzymatic conversion was based on the procedure published by Winn et al. (2008). The conversion was performed in 50 ml reaction tubes with 0.1 M K-PO_4 buffer (Roth) pH 7.8 which additionally contained 1 mM Ser (Roth), 0.06 mM pyridoxal 5'phosphate monohydrate (PLP)

(Sigma-Aldrich), 1 mM 5F-indole or 5OH-indole (Molekula, GmBH, München Germany), and the *St*TrpS, immobilized on 2 ml Ni-NTA resin (1 ml bed volume). The incubation lasted for 6 days at 37 °C. The conversion product and the immobilized *St*TrpS were separated by centrifugation at 411 g 4 °C for 15 min.

2.2.3 PURIFICATION OF 5OH-TRP AND 5F-TRP

45 ml of the conversion product were lyophilized and dissolved in 3 ml 0.1 M K-PO_4 buffer (Roth) pH 7.8 and in 3 ml solution consisting of propanol, 28% ammonium hydroxide (both Roth) and H_2O, mixed in a volumetric ratio of 8 : 1 : 1, respectively. To remove the remaining Ser, the indole analog, and residual PLP, the concentrate was applied onto a silica gel 60 (Roth) column (10 cm). Propanol, H_2O and 28% ammonium hydroxide, in the same volumetric ratio as described above, were used as mobile phase. Fractions of 1 ml were taken, and 3 µl of each fraction were spotted onto a silica gel 60 sheet (Merck Millipore, Billerica, MA). The silica gel sheets were stained with either ninhydrin (Roth) or with the van Urk-Salkowski reagent. The van Urk-Salkowski reagent was a 3-to-1-part mixture of the Salkowaski and the van Urk reagent. The van Urk reagent consisted of 1 g p-dimethylaminibenzaldehyde (Sigma-Aldrich), which was dissolved in 50 ml conc. HCL (Roth), and diluted with 50 ml of H_2O. For the Salkowsky reagent, 2.03 g of $FeCl_3$ * H_2O (Roth) were dissolved in 500 ml H_2O, and 300 ml conc. H_2SO_4 (Roth) were added (Ehmann, 1977).

2.2.4 HPLC DETERMINATION OF THE ENZYMATIC CONVERSION OF 5OH-INDOLE TO 5OH-TRP AND 5F-INDOLE TO 5F-TRP

The HPLC analysis (Agilent 1100) of the *in vitro* conversion was based on the method of Szkop & Bielawski (2013), which we adapted for our purpose. We used a ZORBAX Eclipse XDB-C8 (4,6 x 150 mm, 5-Micron) column (Agilent Technologies, Santa Clara, CA). Two eluents were used for installing a gradient elution. Eluent A consisted of 2.5 : 97.5 % (v/v) acetic acid (Roth) and H_2O, with a pH of 3.8 (calibrated with KOH). Eluent B was composed of 80 : 20 % (v/v) of acetonitrile (Roth) and H_2O. First, the Trp analog was eluted for 25 min with eluent A and B mixed at 80 : 20 % (v/v). Then, the ratio was changed to 50 : 50 % for 6 min and finally to 0 : 100 % for 2 min. The column was regenerated in A/B 80 : 20 % for 7 min. A total run took 40 min, with a flow rate of 1 mL/min. 10 µL of each sample were injected. The Trp analogs were detected with a fluorescence detector, by using the excitation wavelengths of 280 nm for 5F-Trp, and 350 nm for 5OH-Trp. The emissions were measured at 350 nm and 380 nm for 5F-Trp and 5OH-Trp, respectively.

2.3 IN VIVO BIOSYNTHESIS OF TRP ANALOGS AND ECFP VARIANT EXPRESSION

2.3.1 CULTIVATION CONDITIONS FOR THE INCORPORATION PROCEDURES

The BL21(DE3) Δ*trpC*::FRT{p15A-AraC-H6-*St*TrpS; pQE80L-H6-ECFP} gold strain was cultured in a synthetic minimal medium, which contained 20 mM glucose as a C-source, trace elements, salts and all 20 amino acids. The concentration of Trp was limited to 18 µM, whereas the other 19 amino acids were supplemented at 50 mg/L. The detailed composition of the medium can be found in Supplementary Table 3. The cells grew at 37 °C and 120 rpm. The optimal arabinose concentration and time for the induction of the *St*TrpS were determined as indicated in the results section.

For the incorporation procedures, the cells were grown as indicated. The two incorporation procedures differed only in the time for the addition of 1 mM $4NH_2$-indole (ABCR, Karlsruhe, Germany), details are described in the results section. The *St*TrpS was induced with 2 mg/L arabinose (Sigma-Aldrich) at D_{600} of 1. ECFP was induced with 0.1 mM IPTG (Biosynth, Staad, Switzerland) and the temperature was decreased to 28 °C. The D_{600} was recorded in three technical replicates. Sample preparation and expression analysis by sodium dodecyl sulfate polyacryl gel electrophoresis (SDS-PAGE).

The SDS-PAGE was performed with 12% polyacrylamide gels as described by Laemmli (1970). The electrophoresis was performed for 15 min with 100 V, and for 40 min with 180 V. Afterwards, the SDS gels were stained with Coomassie Brilliant Blue G250 (Roth). The staining solution consisted of 2.5 g Coomassie Briliant Blue G250, 7.5% (v/v) acetic acid, 50% (v/v) ethanol and 42.5% (v/v) H_2O. The destaining solution was a mixture of 7.5% (v/v) acetic acid, 20% ethanol and 72.5% H_2O. The gels were distained until distinct protein bands were visible. For the estimation of the protein sizes, a prestained protein ladder (Thermo Fisher Scientific) was used.

To monitor protein expression, samples were collected before *St*TrpS induction, before addition of $4NH_2$-indole, after 3 h of expression, and after overnight expression. Culture volumes corresponding to 1 D_{600} were harvested by centrifugation for 1 min at 13000 rpm at room temperature and the culture supernatant was discarded. The cell pellets were resuspended in 100 µL CelLyticB 2x Cell lysis reagent (2x conc.) (Sigma-Aldich) for chemical cell lysis. The chemical lysis was performed in thermo mixer for 10 min at 25 °C and 900 rpm. Afterwards, the samples were centrifuged for 2 min at 17 g at room temperature and 100 µL of the supernatant (soluble protein faction) were mixed with 25 µl 5x SDS loading dye (composition see Supplementary Table 2). The pellet (insoluble protein faction) was re-solubilized with 100 µL 6 M urea (Roth), and afterwards, 25 µl 5x SDS loading dye was added. The soluble and insoluble protein fractions were incubated for 5 min at 90 °C and 600 rpm, and 9.4 µl (0.075 D_{600}) were applied onto the SDS gel.

For the visualization of the soluble proteome including the expressed ECFP[5F-Trp] and ECFP[5OH-Trp], the cells were harvested after the overnight cultivation, by centrifugation at 4000 g at 4 °C for 20 min. The cell pellets were re-suspended in 20 mM Tris/Cl, 150 mM NaCl at pH 7.8. For a better cell lysis, 1 mg/mL lysozyme (Roth) and 0.003 mg/mL DNAse (Sigma-Aldrich) were added. The cells were lysed by sonication (output control 8, duty cycle 70 - 80 %, 6 min), and afterwards the soluble and insoluble cell components were separated by centrifugation at 20000 g, 4 °C for 40 min. The insoluble protein fraction was re-suspended in 6 M urea. The protein concentration was determined with a Bradford protein assay. The samples for the SDS-PAGE had a total volume of 200 µl. They contained 50 µl NuPAGE (Thermo Fisher Scientific) sample buffer and 1 µl beta-mercaptoethanol (Roth). The samples were heated for 20 min at 70 °C, before 3 µg total protein were applied onto the gel.

2.3.2 WESTERN BLOT

For the immunodetection (Haid & Suissa, 1983)of the hexahistidine-tagged proteins, a SDS-PAGE was performed as described above. The proteins were electrotransfered from within the SDS-Gel onto a nitrocellulose membrane (GE Healthcare, Little Chalfont, UK) (160 mA, 25 V 120 min). The membrane was blocked overnight at 4 °C with 3% bovine serum albumin (BSA) (Roth) in Tris-buffered saline with Tween 20 (TTBS) (20 mM Tris pH 7.5, 150 mM NaCl, 0.1 Tween 20). The immunodetection was performed with a primary anti-hexahistidine epitope tag antibody from mouse (1:3000 dilution in TTBS, 1 h incubation at 4 °C) and a goat anti-mouse IgG + IgM (H+L) secondary antibody conjugated to horseradish peroxidase (1:10000 dilution in TTBS, 1 h incubation at room temperature) (all from Life Technologies, Carlsbad, CA). The SuperSignal® West Dura Extended Duration Substrate (Life Technologies) was used for the chemiluminescent detection as indicated by the manufacturer.

2.3.3 PROTEIN CONCENTRATION DETERMINATION

The protein concentration was determined by Bradford protein assay (BioRad), according to the manufacturer's protocol at 595 nm. The calibration curve was performed with 0.0625, 0.125, 0.25, 0.5 and 1,2 mg /ml albumin G (Roth) in 50 mM Tris/Cl pH 7.5. The determinations of the protein concentration of the samples as well as the calibration curve were conducted in 96-well microtiter plates with a plate reader.

2.3.4 VARIANT PROTEIN PURIFICATION

The *E. coli* host cells were harvested by centrifugation at 4000 g, 4 °C for 20 min and resuspended in 20 mM Tris/Cl, 150 mM NaCl, pH 7.8. Then, 1 mg/mL lysozyme and 0.003 mg/mL DNAse were added

(all Roth) to improve the cell lysis. The cell lysis was performed by sonication (output control 8, duty cycle 70 - 80 %, 6 min) and the lysed cells were centrifuged at 20000 g, at 4 °C for 40 min to separate soluble and insoluble cell components.

The ECFP[4NH_2-Trp] variant was purified by two-phase partitioning based on the method described by Samarkina et al. (2009). 7.5 ml 5 M NaCl, 58.25 ml saturated $(NH_4)_2SO_2$ (pH 7,8) and 30 ml EtOH (all Roth) were added to 25 ml cell lysate and mixed for 30 sec. The mixture was allotted to 50 mL reaction tubes for centrifugation for 7 min at 3000 g at room temperature. After centrifugation, two phases were visible. The supernatant contained the fluorescent protein, which was transferred to a fresh tube. One volume nButatnol (Roth) was added to four volumes of supernatant, mixed for 30 sec and centrifuged for 7 min at 3000g at room temperature. After centrifugation, the fluorescent protein was located in the lower, aqueous phase and was recovered using a syringe. Afterwards, the samples were dialyzed three times against 2000 volumes of 50 mM Tris/Cl at pH 7.5 and 4 °C using SnakeSkin™ Dialysis Tubing (Thermo Fisher Scientific).

The ECFP[5F-Trp] and ECFP[5OH-Trp] were purified via their N-terminal hexahistidine tag by Ni^{2+} affinity chromatography on 500 µl bed volume of Ni-NTA resin in batch mode, according to the manufacturer's protocol. The cell lysate was prepared as described above and 1 M urea was added to prevent the aggregation of the *St*TrpS. The cell lysates and the Ni-NTA resin were incubated for 1 h at 4 °C with slight shaking. Afterwards, the Ni-NTA resin with the bound hexahistidine-tagged proteins and the unbound proteins were separated by centrifugation at room temperature with 4000 rpm for 20 min. The Ni-NTA resin was resuspended in 1.5 ml wash buffer, containing 3.75 M urea, 50 mM NaH_2PO_4*H_2O, 300 mM NaCl and 20 mM imidazole (all Roth). Again, the Ni-NTA was separated by centrifugation at room temperature with 13000 rpm for 2 min. The washing and centrifugation steps were repeated twice. The bound proteins were eluted by 3 successive incubations of the Ni-NTA resin in 250 ml elution buffer, which contained 50 mM NaH_2PO_4*H_2O, 300 mM NaCl and 250 mM imidazole. Afterwards, the Ni-NTA resin was centrifuged as described above for 5 min. In order to remove all Ni-NTA resin before the dialysis, an additional final centrifugation step was performed. The samples were dialyzed against 2000 volumes of 50 mM Tris/Cl at pH 7.5 and 4 °C using SnakeSkin™ Dialysis Tubing (Thermo Fisher Scientific) in the micro dialysis protocol for 2 ml reaction tubes by the manufacturer.

After dialysis, the purified ECFP variant and the ECFP parent protein were stored in a 50 mM Tris/Cl, pH 7.5 buffer.

2.3.5 FLUORESCENCE SPECTROSCOPY

The fluorescence of the isolated proteins was analyzed in a 96-well microtiter plate by exciting the ECFP and ECFP[4NH_2-Trp] at their specific wavelengths of 434 nm and 466 nm, respectively (Bae et al., 2003). The fluorescence spectrum was recorded from 460 nm to 700 nm excitation wavelength. The fluorescence spectrometry analysis was performed on a Synergy MX spectrometer (BioTek, Winooski,VT) at 20 °C.

2.3.6 MASS SPECTROMETRY OF THE ECFP[NCTRP]

For mass analysis, the ECFP[ncTrp] variants were sent to the ACIB Core Facility Functional Proteomics (Assoc. Prof. PD DI Dr.tech. Ruth Birner-Grünberger, Medical University of Graz, Austria). The samples were ionized by electrospray ionization, and their intact mass was analyzed by the Time of Flight (TOF) method. The theoretical mass of the ECFP[ncTrp] was calculated by using MassXpert (Rusconi & Belghazi, 2002).

3 RESULTS AND DISCUSSION

3.1 *IN VITRO* SYNTHESIS OF 5F-TRP AND 5OH-TRP

The incorporation of ncAA into proteins is a promising method of expanding the chemical properties of the target proteins. The previously published incorporation methods included the supplementation of the ncAA to the growth medium (Zheng & Kwon, 2012). However, ncAAs, particularly ncTrp are considerably expensive, and many analogs are commercially unavailable. In order to decrease the costs of the incorporation of ncTrp into proteins, we established the preparative *in vitro* synthesis of Trp analogs by the enzymatic condensation of the corresponding indole analogs with Ser (Winn et al., 2008).

In place of all available Trp analogs, we chose the enzymatic conversion of 5F-indole to 5F-Trp, and 5OH-indole to 5OH-Trp, with the *St*TrpS enzyme. The *St*TrpS was immobilized on a Ni-NTA resin *via* its N-terminal hexahistidine tag. The conversion was performed with 1 mM indole analog, equimolar amounts of Ser and 0.06 mM PLP, as the co-factor. The reaction was completed in 6 days. Detailed information can be found in the methods section.

To remove the residual indole analogs, PLP and Ser, the *in vitro* conversion product was purified by silica gel chromatography. Eluent fractions were collected in regular intervals. To identify those fractions containing the Trp analog, a small volume of each fraction was spotted on a silica gel sheet and stained with ninhydrin and the van Urk-Salkowski reagent (Figure 2). The van Urk-Salkowski reagent specifically stains indoles (Ehmann, 1977), and so it indicated the indole analog as well as the Trp analog on the silica gel sheet (Figure 2, panel A and C). The ninhydrin reagent stains ammonia, primary and secondary amines (Yemm et al., 1955) and so, Ser as well as the Trp analog could be visualized on the silica gel sheet (Figure 2, panel B and D). Thus, fractions stained with both reagents contained the Trp analog. The colorimetric analysis of the elution fractions showed that, in case of the *in vitro* synthesis of 5F-indole to 5F-Trp, the fractions 8 to 24 were stained with the van Urk-Salkowski and the ninhydrin reagent (Figure 2, panel A and B). The fractions 6 to 21 of the 5OH-indole to 5OH-Trp conversion also stained with both reagents. Hence, we concluded that the Trp analogs were present in these fractions (Figure 2, panel C and D).

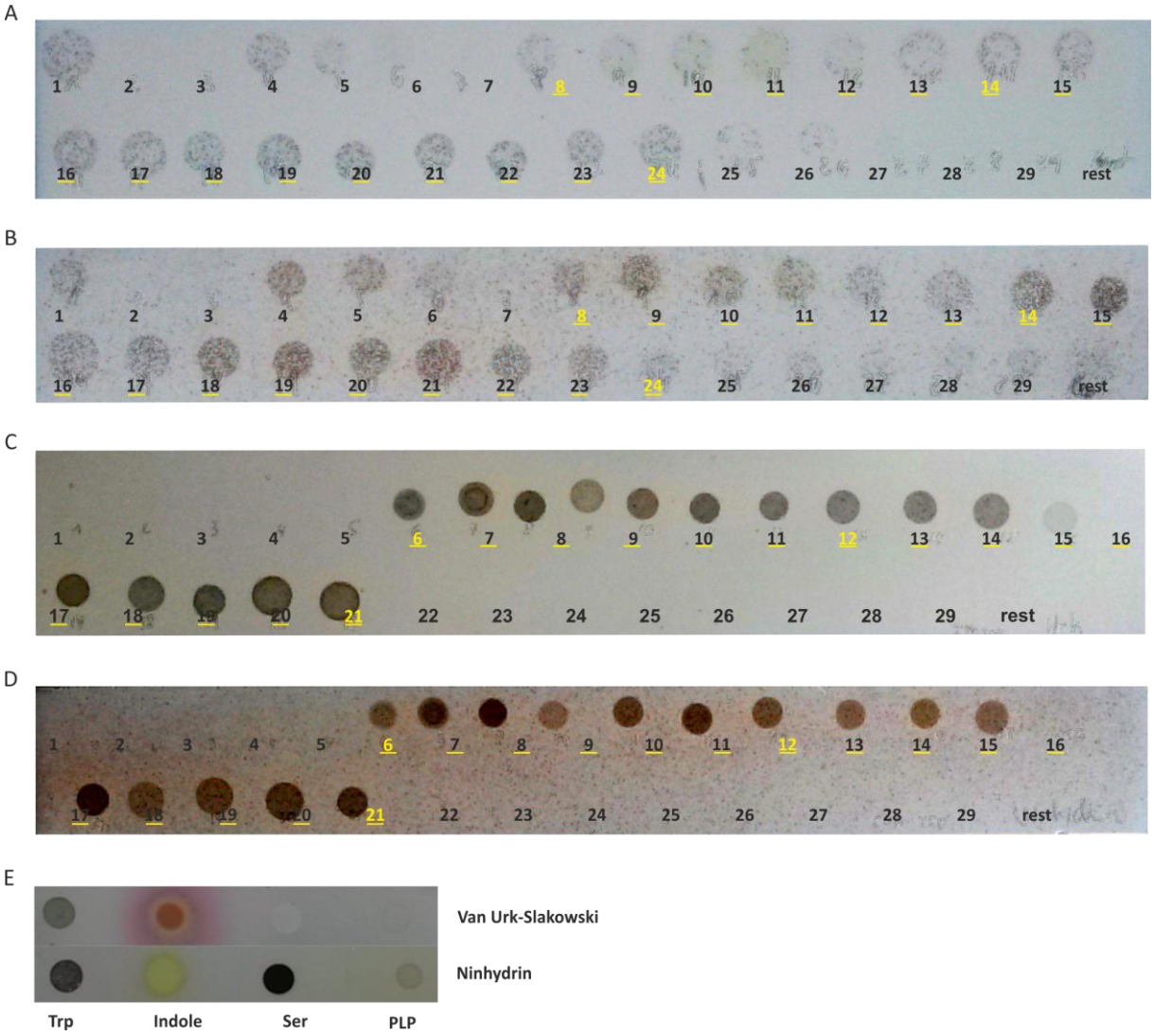


Figure 2: Stained samples of the enzymatic *in vitro* conversion of the indole analogs to ncTrp analogs.
Three µl of each fraction of the conversion of 5F-indole to 5F-Trp (A and B) and 5OH-indole to 5OH-Trp (C and D), were spotted onto a silica gel sheet. The van Urk-Salkowski reagent stains the indole ring (A and C) and the ninhydrin reagent reacts with the primary amines of amino acids (B and D). The highlighted fractions that were used for further HPLC analysis and the fractions, which were selected for preparative purification, are underlined. Trp, Indole, Ser and PLP were spotted for a pre-test on a silical gel sheet and stained with the van Urk-Salkowski or ninhydrin reagent (E).

After we had identified the fractions containing 5OH-Trp and 5F-Trp, we performed an HPLC analysis to verify the purity of the single fractions (Figure 3). The fractions 8, 14, 24 of the 5F-Trp *in vitro* synthesis, and the fractions 6, 12, 21 containing 5OH-Trp were chosen for the HPLC analysis. Since the samples had already been purified from the other components of the *in vitro* biotransformation by silica gel chromatography, we expected only a single peak of the respective Trp analog in the HPLC chromatograms. Thus, the major peaks in Figure 3 could be assigned unambiguously to the expected Trp analogs. The HPLC chromatograms of fractions 6, 12, 21 showed only a single peak at 1.8 min, indicating 5OH-Trp (Figure 3, panels B, D and F). The HPLC chromatograms of the fractions containing 5F-Trp showed a prominent peak at 3.4 min, and a small peak at 5 min (Figure 3, panels A, C and E). Although we did not have commercial reference compounds, we assumed that the peaks at 3.4 and 1.8 min were 5F- and 5OH-Trp, respectively. 5F-Indole eluted much later, at 27 min, than 5F-Trp; the

5-OH indole and 5OH-Trp peaks were also well separated (6.6 min *vs* 1.8 min). PLP eluted at 1.3 min, which is close to 5OH-Trp at 1.8 min, however, the peak intensity is much higher than it would have been expected for PLP. Ser is not visible by the detection method used. The minor peak at 5 min in Figure 3, panels A, C and E, remained unidentified. 5F- and 5OH-Trp were prepared very efficiently and in high purity by the enzymatic conversion.

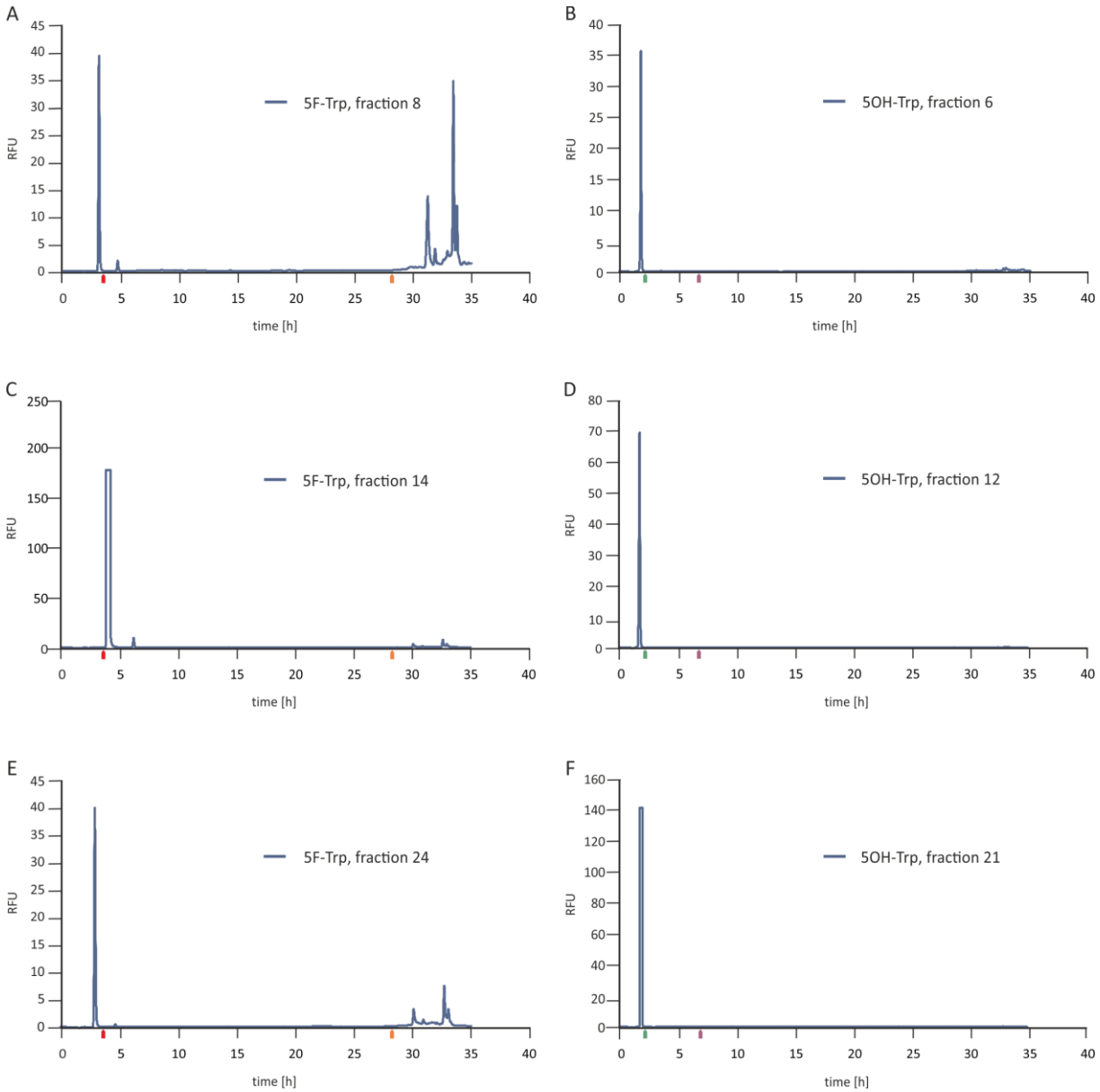


Figure 3: HPLC chromatograms of selected fractions of the enzymatic *in vitro* conversions.
The fractions 8, 14 and 24 of the enzymatic conversion of 5F-indole to 5F-Trp (A, C, E), and the fractions 6, 12 and 21 of the enzymatic conversion of 5OH-indole to 5OH-Trp (B, D and F), are shown. Retention times of all related compounds in the in vitro assay: Trp: 2.5 min; indole: 20 min; 5F-indole: 27 min (orange arrow); 5F-Trp: 3.4 min (red arrow); 5OH-indole: 6.6 min (green arrow); 5OH-Trp : 1.8 (purple arrow); PLP 1.3 min; Ser: not visible.

The HPLC results were confirmed by NMR (Ao. Univ.-Prof. Dipl.-Ing. Dr. techn. Hans-Jörg Weber, Institute for organic chemistry, Graz University of Technology) (Supplementary Figure 3, Supplementary Figure 4). The fractions 8 to 24 (5F-Trp) (Figure 3, panels A, C and E) and the fractions 12 to 21 (5OH-Trp) (Figure 3, panels B, D and F) were pooled, and the solvent was evaporated. With

the *in vitro* conversion of 5OH-indole to 5OH-Trp and 5F-indole to 5F-Trp, we reached 79 % and 53 % of the theoretical yield, respectively. In absolute numbers, 1 mol of each indole analog was used to yield 0.526 mol (173 mg) of 5F-Trp and 0.785 mol (117 mg) of 5OH-Trp.

The successful enzymatic *in vitro* synthesis of the 5OH-indole and the 5F-indole to 5OH-Trp and 5F-Trp, catalyzed by *St*TrpS, is in line with the report of Winn et al. 2008, who published that the *St*TrpS is an efficient Trp synthase for the conversion of indole analogs to Trp analogs. Hence, the *St*TrpS seemed to be promising for an enzymatic *in vivo* conversion of indole analogs to Trp analogs, and for a parallel incorporation of the ncTrp into proteins.

3.2 *IN VIVO* SYNTHESIS OF TRYPTOPHAN ANALOGS AND INCORPORATION INTO ECFP

Based on the already published SPI method for the incorporation of ncAA in proteins (Budisa et al., 1995), we aimed to establish an *in vivo* method for the parallel enzymatic synthesis of the ncTrp, and the incorporation into a target protein. In this chapter, I will describe two assayed incorporation procedures and their evaluation.

3.2.1 INDOLE ANALOG SUPPLEMENTATION SHORTLY BEFORE TARGET PROTEIN INDUCTION

We established an *in situ* biotransformation and incorporation procedure for the labeling of ECFP with ncTrp. The incorporation procedure is based on a co-expression of *St*TrpS, and on the depletion of the Trp auxotrophic host cells for Trp before target gene expression (Figure 4, panel A). The expression of *St*TrpS was decoupled from the expression of the target protein. We used a Trp auxotrophic strain with a deletion of *trpC.* This ensured that the *E. coli* cells could not synthesize Trp on their own, and that the intracellular Trp depended on the Trp concentration in the medium. The use of auxotrophic stains is crucial for a quantitative residue-specific incorporation (Budisa, 2004). Our protocol distinguishes itself from the SPI method (Budisa et al., 1998) by using two plasmids for separately induced expressions of two recombinant proteins. The two cultivation phases (Figure 4, panel A; growth phase, production phase) were performed without shifting the medium. One plasmid was the p15A-AraC-H6-*St*TrpS that carried a hexahistidine-tagged *St*TrpS, which was under the control of the arabinose inducible P_{araBAD} promoter. The plasmid was, due to its p15A origin of replication, a medium copy plasmid. The second expression construct was a high copy plasmid with a ColE1 origin of replication, termed pQE80L-H6-ECFP. The plasmid carried a hexahistidine-tagged ECFP, which was under the control of an IPTG inducible T5/*lacO* promoter. The cells were cultivated under limited Trp conditions (Supplementary Table 1). Trp was titrated in this minimal medium so that cell growth stalled at a D_{600} of 1.8 – 2 due to Trp depletion. In the first phase of the cultivation,

the cells grew because they were supplied with all cAAs and during this phase the expression of *St*TrpS was induced (Figure 4, panel A). Since the Trp in the medium was growth limiting, we aimed to ensure sufficient *St*TrpS in the cells before they ran out of Trp. The second phase started after the depletion of Trp, when the indole analog was added. The *St*TrpS, produced during the growth phase, converted the indole analog to the Trp analog. The expression of the target gene was induced 30 min after the addition of the indole analog, to give the *St*TrpS time to catalyze the conversion.

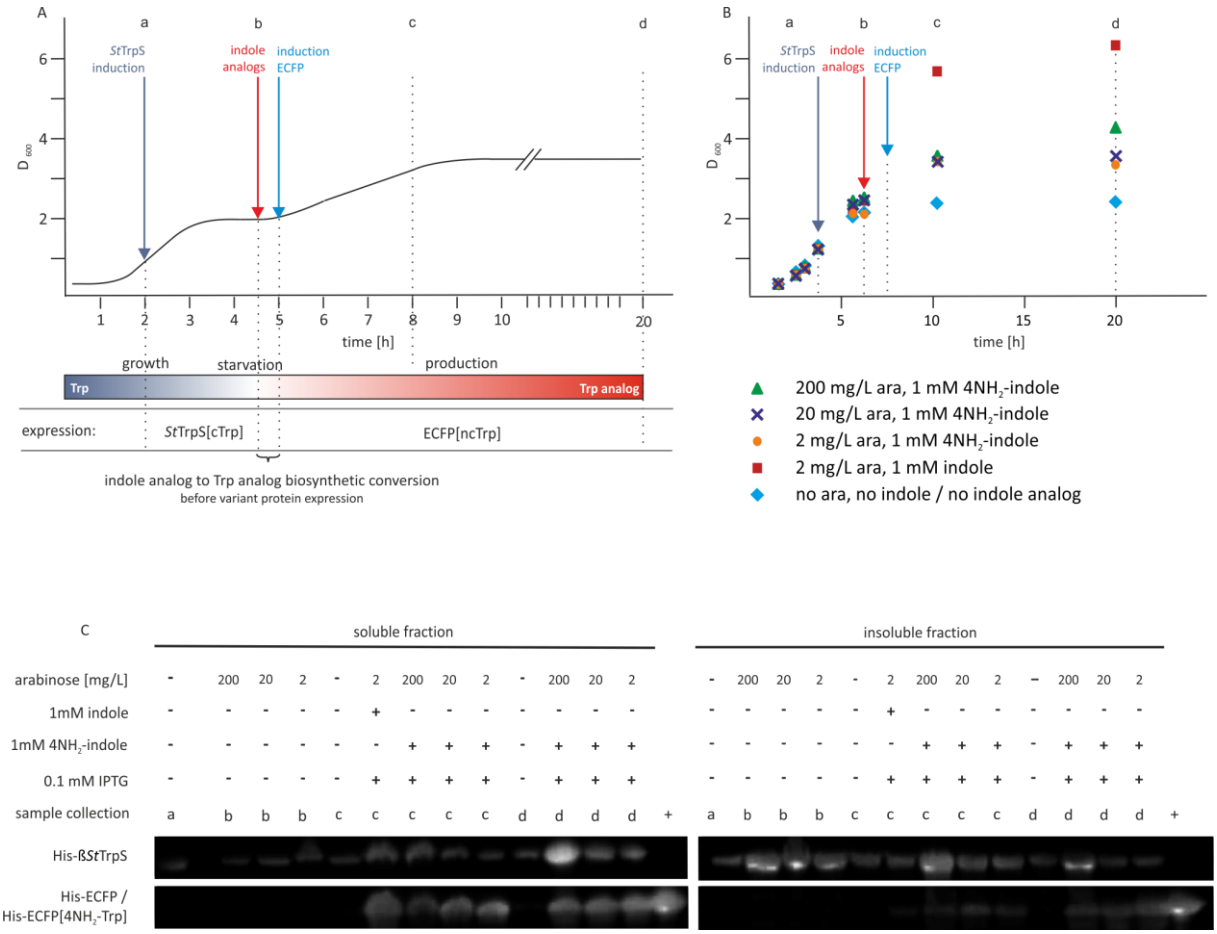


Figure 4: Schematic illustration of the predicted growth patten of the procedure with the indole analog supplementation 30 min before target protein induction (A), growth pattern of indole and $4NH_2$-indole fed cultures (B), and a Western blot of the hexahistidine-tagged (His-) proteins: β*St*TrpS, ECFP and ECFP[$4NH_2$-Trp] (C).
The image in (A) shows the schematic growth behavior and variant protein production (ECFP) in relation to the Trp consumption and enzymatic ncTrp synthesis. Arrows indicate the induction of the expression of the *St*TrpS by the addition of arabinose, the addition of $4NH_2$-indole, and the induction of the variant protein (ECFP) with IPTG. The samples for the SDS-PAGE and the Western blot (C) were collected at the instances marked with letters (a, before TrpS induction; b, before addition of indole or indole analog; c, after 3h of variant protein expression; d, overnight expression).
The outlined procedure in A was applied and the SDS-PAGE samples were collected as indicated. D_{600} was determined in triplicates and the calculated relative mean values are visualized (B).
We analyzed the collected samples by immunodetection, as shown in (B). The His-β*St*TrpS, His-ECFP and His-ECFP[$4NH_2$-Trp] are visible as fluorescing bands. Purified His-ECFP was used as a positive control and is indicated with + (C).

For an optimal utilization of the C-source, the optimal D_{600} and arabinose concentration had to be examined for the induction of *St*TrpS. Sufficient amounts of *St*TrpS were to be present in the cells for the catalysis of the conversion of the indole analog to the Trp analog. However, the expression of the *St*TrpS should not use up all cellular nutrients for its production. We examined the induction of

*St*TrpS at D_{600} of 0.5, 1, and 1.5 with 200 mg/L, 20 mg/L and 2 mg/L arabinose (Supplementary Figure 5). When we induced the *St*TrpS at D_{600} of 1 or 1.5, the depletion of Trp occurred as expected, regardless of the arabinose concentration (Supplementary Figure 5, panel B and C). The induction of the *St*TrpS at D_{600} of 0.5 with 2 mg/L and 20 mg/L also showed the expected Trp depletion, resulting in growth arrest at D_{600} of ~1.8-2. However, the induction of *St*TrpS at D_{600} of 0.5 with 200 mg/L arabinose showed no growth arrest. This was in contrast to our expectations since we anticipated the cells to be Trp limited. The host strain carried a block at the end of the Trp biosynthesis pathway (Δ*trpC*). It is unclear, how the comparably high amount of 200 mg/L arabinose for induction of *St*TrpS expression obviously enabled the cells to bypass their Trp auxotrophy (Supplementary Figure 5, panel A). Because we repetitively observed this growth phenomenon, we decided to induce the expression of *St*TrpS at D_{600} of 1. After the growth arrest, we added Trp to the cultures to confirm that the depletion of Trp was the reason for the growth arrest and that no other nutrient was limiting.

Following the identification of the optimal D_{600} for the induction of *St*TrpS with arabinose, we wanted to investigate the influence of the arabinose concentration on the expression level of the *St*TrpS and the ECFP. Therefore, the *St*TrpS was induced at D_{600} of 1 with 200 mg/L, 20 mg/L or 2 mg/L arabinose. After the growth arrest, when indole or $4NH_2$-indole were added, the cells apparently started to grow again (Figure 4, panel B). In the indole fed cultures, the final D_{600} reached a value of 6; with $4NH_2$-indole, a final D_{600} of 3.2 - 4. The indole fed culture (Figure 4, panel B, red squares) received indole (1 mM, corresponds to 117 mg/L) which was biosynthesized to Trp to overcome the depletion of Trp, owing to the limited Trp concentration in the minimal medium. In contrast, the cultures fed with $4NH_2$-indole were not able to reach the same final D_{600} as the culture that was supplemented with indole. Nevertheless, the $4NH_2$-indole fed cultures also restart growing when $4NH_2$-indole was supplemented after the depletion Trp. This indicates that the $4NH_2$-indole was converted to $4NH_2$-Trp and could be utilized in ribosomal translation but not that efficient than Trp. A control culture that neither received arabinose for the induction of *St*TrpS nor was fed indole or an indole analog remained depleted for Trp throughout the entire production phase (Figure 4, panel B, blue diamonds). Particularly, the culture that received a comparably high arabinose dose (200 mg/L) for the induction of *St*TrpS, grew to a higher cell density than the cultures with a lower arabinose concentration (20 mg/L and 2 mg/L). In this case, arabinose could have been utilized after exhausting the available C-source (Desai & Rao, 2010).

Furthermore, the observation of the growth restart of the $4NH_2$-indole fed cultures is of concern. In contrast to the work of Hoesl et al. (2015) and Bacher & Ellington (2001), it showed the utilization of ncTrp for proliferation without evolving the genome of the host cells with a selective pressure for a long time. Hoesl et al. (2015) and Bacher & Ellington (2001) described how to evolve the genomes of the host cells for being capable to grow on L-β-(thieno[3,2-*b*]pyrrolyl)alanine and 4-fluorotryptophan,

respectively, but not for $4NH_2$-Trp. Supposedly, the utilization of ncTrp for the cellular metabolism without evolving the host strain depends on the ncTrp. Another hypothesis was that the co-expressed *St*TrpS played a crucial role. The *St*TrpS is able to synthesize indole analogs efficiently to Trp analogs (Winn et al., 2008), as we already showed with the *in vitro* conversion of 5F-indole and 5OH-indole to 5F-Trp and 5OH-Trp. Based on the assumption that the movement of indole analogs across the membrane of *E. coli* behave similar than indole (Piñero-Fernandez et al., 2011), we speculate that $4NH_2$-Trp was abundant in the cellls.

To scrutinize the expression levels of the hexahistidine-tagged *St*TrpS and ECFP variant by immunodetection, samples at defined growth stages were collected. The samples were collected before the induction of *St*TrpS with arabinose (a), before the addition of indole/$4NH_2$-indole and the induction of ECFP with IPTG (b), after three hours of expression (c), and after expression overnight (d) (Figure 4, panel B, samples at defined growth stage marked with a, b, c and d). A Western blot of the soluble and insoluble proteome was performed (Figure 4, panel C). It showed a basal expression of *St*TrpS, under the control of the P_{araBAD} promoter, without induction (Figure 4, panel c, sample a). The expression of ECFP was regulated with an IPTG inducible T5/*lacO* promoter, which appeared to be tighter than P_{araBAD} (Figure 4, panel c, sample a and b). The Western blot demonstrated that the $4NH_2$-Trp could be utilized for the ECFP expression (Figure 4, panel c, sample c and d with (+) for $4NH_2$-indole), because it didn't show any ECFP expression (Figure 4, panel c, sample d with (-) for $4NH_2$-indole) in the samples which originated from cultures without $4NH_2$-indole supplementation. This and the growth pattern of the cell cultures (Figure 4, panel B) assisted the evidence that Trp was depleted before the cultures were fed with $4NH_2$-indole. The variations in the arabinose concentration only affected the *St*TrpS expression, but not the ECFP expression. This means that higher amounts of *St*TrpS in the cells did not cause a higher expression level of the ECFP variants. Based on this finding, we decided to induce *St*TrpS with 2 mg/L arabinose at D_{600} of 1.

3.2.2 INDOLE ANALOG SUPPLEMENTATION DURING CELL GROWTH

In the review of Richmond (1962), he concluded that ncAA can cause growth inhibition, but do not have preferential inhibitory effects on protein synthesis. Minks et al. (1999) also observed the reported growth inhibitory effect when they incorporated 4F-Trp, 5F-Trp and 6F-Trp into annexin V by using a Trp auxotrophic *E. coli* strain but without the co-expression of *St*TrpS. Based on these findings, we devised a procedure, where the indole analog was supplemented during the cell growth. With this approach, we wanted to assess the addition of $4NH_2$-indole before the depletion of Trp, so the *St*TrpS had more time to convert the $4NH_2$-indole to the $4NH_2$-Trp. The concept of a separate induction of the *St*TrpS and ECFP remained the same. The *St*TrpS was induced during the exponential growth phase, at D_{600} of 1 with 2 mg/L arabinose. In both procedures, the ECFP expression was

induced with IPTG simultaneously. In contrast to the incorporation procedure described above, the $4NH_2$-indole was added immediately after the induction of *St*TrpS.

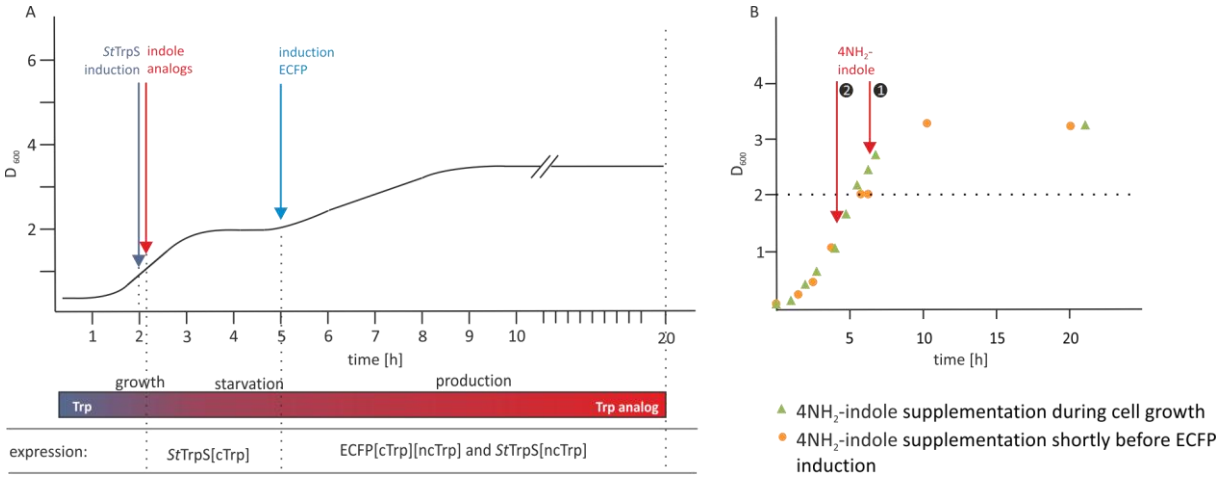


Figure 5: Schematic illustration of the predicted growth pattern of the procedure with the indole analog supplementation during cell growth (A) and growth patterns of the $4NH_2$-indole-fed cultures resulting from both described procedures (B).
The figure in (A) shows the schematic growth pattern and the variant protein expression (ECFP) in context with the Trp consumption and enzymatic ncTrp analog synthesis. The induction of the expression of the *St*TrpS by the addition of arabinose, the addition of $4NH_2$-indole and the induction of the variant protein with IPTG are indicated by arrows.
The D_{600} values of the "indole analog supplementation during cell growth" (green) and the "indole analog supplementation shortly before target protein induction" procedure (orange) are given in (B). The addition of the $4NH_2$-indole is highlighted with a red arrow, which is labeled ❶ for the procedure with the indole analog supplementation shortly before target protein induction and ❷ for the procedure with the indole analog supplementation during cell growth.

If $4NH_2$-indole or rather $4NH_2$-Trp affected exponential growth we expected to observe a preliminary growth arrest or a linearization of the proliferation rate. To test our hypothesis, we compared the different growth patterns of the cells resulting from the application of both procedures (Figure 5, panel B). Both cultures reached the same final D_{600}. In case of performing the procedure in which we added the $4NH_2$-indole shortly before the induction of the ECFP expression, the cells behaved as already shown in Figure 4 panel B. They stopped growing when they had exhausted Trp and restarted growing after the addition of $4NH_2$-indole (Figure 5, panel B, orange dots). In case of performing the procedure in which we added the $4NH_2$-indole during cell growth, we were not able to detect the anticipated deviations of the exponential cell growth. These observations allowed speculating, whether the tryptophanyl-tRNA synthetase from *E. coli* consumes the Trp before the $4NH_2$-Trp, or not. The common notion is, that the cAAs are the preferred substrates of the aminoacyl-tRNA synthetases. An enzymatic characterization of the tryptophanyl-tRNA synthetase from *E. coli* with different substrates would be necessary. Based solely on the cell densities, we cannot answer this question. Changes in the morphology of the cells, which can be caused by heterologous expression, can influence the D_{600} values. Neither this was not examined under the microscope nor the cell dry weights of the cells, with or without heterologous expression of *St*TrpS and ECFP, were compared.

3.2.3 Evaluation of the $4NH_2$-Trp incorporation efficiencies of the two different incorporation procedures

To assess the incorporation of $4NH_2$-Trp, the ECFP variant proteins produced with the two procedures were purified by the method of Samarkina et al. (2009). The procedure, in which we added the $4NH_2$-indole shortly before the induction of the ECFP expression, produced ECFP[$4NH_2$-Trp]-❶. The ECPF variant, which was produced by acompolishing the procedure with the addition of the $4NH_2$-indole during the cell growth, was named ECFP[$4NH_2$-Trp]-❷. The incorporation of $4NH_2$-Trp was analyzed by fluorescence spectrometry (Figure 6, panel A) and mass spectroscopy (Table 1).

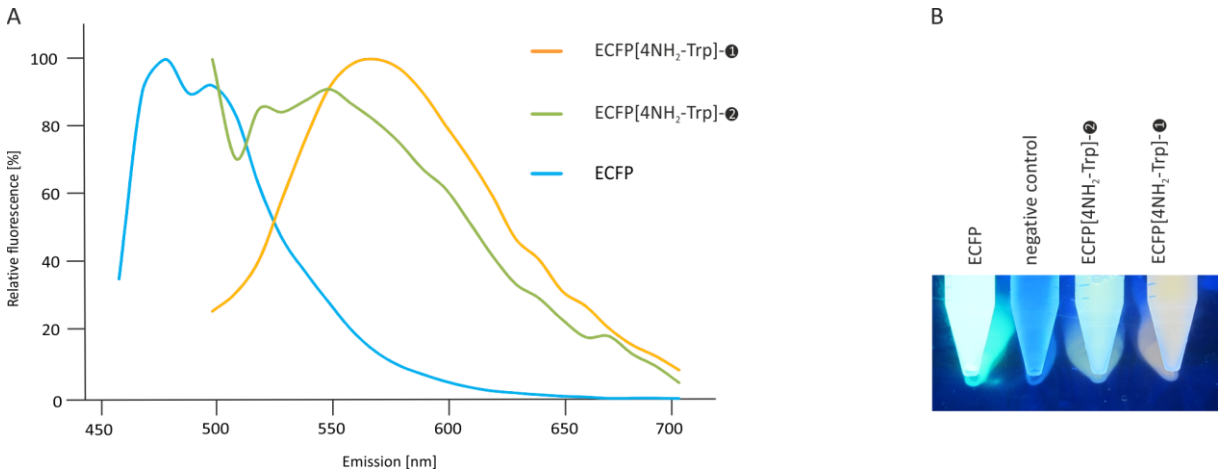


Figure 6: Fluorescence emission spectra (A) and fluorescing samples (B) of the purified ECFP and ECFP[$4NH_2$-Trp]
The graphs in (A) show the fluorescence emission spectrum of ECFP and the different fluorescence emission spectra of the ECFP[$4NH_2$-Trp]-❶ and ECFP[$4NH_2$-Trp]-❷. The samples were excited at their expected specific wavelengths (ECFP: λ_{ex} 434 nm, ECFP[$4NH_2$-Trp]: λ_{ex} 466 nm). The fluorescence signals were set to 100%.
The purified ECFP samples, which are produced by accomplishing the two procedures, and an empty vector control sample under UV light, are shown in (B).

The amino acid sequence of ECFP is shown in Supplementary Sequence 3. ECFP contains two Trp residues. One of them is essential for forming the chromophore. The incorporation of $4NH_2$-Trp into ECFP at position 66, which is part of the chromophore, can be assessed easily based on a red shift in fluorescence. The ECFP[$4NH_2$-Trp] appears golden and shows an emission maximum at 576 nm. The parent protein, the ECFP, appears cyan fluorescing and shows two emission maxima at 475 nm and 506 nm (Bae et al., 2003).

We excited 2.3 µg of the ECFP parent protein at 434 nm and detected the emission wavelength after every 10 nm from 460 nm to 700 nm. The resulting fluorescence spectrum showed the expected characteristics (Figure 6, panel A, cyan line) with an emission maximum of 480 nm and 500 nm which corresponded nicely to the published value of 475 nm and 506 nm. We excited 1.5 µg of ECFP[$4NH_2$-Trp]-❶ and of ECFP[$4NH_2$-Trp]-❷ at 466 nm, which is the specific excitation wavelength

of ECFP[4NH_2-Trp]. The fluorescence emission spectra were measured as described for the ECFP parent protein. The ECFP[4NH_2-Trp]-❶ showed the fluorescence red shift characteristic for ECFP[4NH_2-Trp], which is already known from literature, (Bae et al., 2003) and an emission maximum at 570 nm (Figure 6, panel A, gold line). In contrast, the fluorescence spectrum of the ECFP[4NH_2-Trp]-❷ with its two emission maxima at 500 nm and 550 nm, let us presume a mixture of ECFP and ECFP[4NH_2-Trp] (Figure 6, panel A, green line). It showed characteristics of the ECFP and the ECFP[4NH_2-Trp].

We were even able to see the different fluorescence of the purified ECFP, ECFP[4NH_2-Trp]-❶ and ECFP[4NH_2-Trp]-❷, under a UV light lamp (Figure 6, panel B).

The mass spectrometry analysis confirmed the results of the fluorescence spectrometry (Table 1). We calculated the masses of the ECFP variants, containing none, one or two 4NH_2-Trp, as well as with or without a formed chromophore, respectively. The calculated masses for each option and the found masses are given in Supplementary Table 4. We could assign all found masses to species with the chromophore. The relative abundance of all found species are given in Table 1. The ECFP[4NH_2-Trp]-❶ was fully labeled. The ECFP[4NH_2-Trp]-❷ was composed of 28.7% of full labeled ECFP[4NH_2-Trp], 12.3% ECFP and 58.9% of single labeled ECFP[4NH_2-Trp]. Thus, the full labeled ECFP[4NH_2-Trp]-❷ had a golden fluorescence signal and the unlabeled ECFP in this sample appeared cyan fluorescing. The single labeled ECFP[4NH_2-Trp]-❷ could appear golden or cyan fluorescing because only the incorporation of 4NH_2-Trp in the chromophore, at position 66, causes the golden fluorescence. The second Trp of the ECFP does not contribute to a change of the fluorescence. The mass spectrometry cannot distinguish between the positions of the 4NH_2-Trp, so it could not be determined, how much of the single labeled ECFP[4NH_2-Trp]-❷ appeared cyan or gold fluorescing.

Table 1: Comparison of the incorporation protocols in terms of 4NH_2-Trp incorporation into ECFP.
The relative abundance of the incorporated 4NH_2-Trp into ECFP[4NH_2-Trp]-❶ and ECFP[4NH_2-Trp]-❷ was shown in the table. The abundance of the species were calculated based on the generated data by mass spectrometric analysis. The calculated and found masses are listed in Supplementary Table 4. 'n.d' refers to 'not detected'.

species number of incorporated 4NH_2-Trp	**abundance of species [%]**	
	ECFP[4NH_2-Trp]-❶	ECFP[4NH_2-Trp]-❷
0	n.d.	12.3
1	n.d.	58.9
2	100	28.7

This was an evidence to confirm that E. coli does not consume Trp before 4NH_2-Trp for the incorporation into proteins. Only the procedure with a growth arrest, which was caused by the depletion of the limited Trp, resulted in fully labelded ECFP[4NH_2-Trp]. The procedure in which we added the 4NH_2-indole during the cell growth to the same medium, showed exponential growth and a mixture of ECFP, single labelded ECFP[4NH_2-Trp] and fully labeled ECFP[4NH_2-Trp].

3.3 INCORPORATION OF 5OH-TRP AND 5F-TRP INTO ECFP

After our success in full labeling of ECFP with $4NH_2$-Trp, we aimed to examine the suitability to label ECFP with other ncTrp by this procedure. Further, we questioned the necessity of the co-expressed *St*TrpS in this procedure for yielding a fully labeled ECFP.

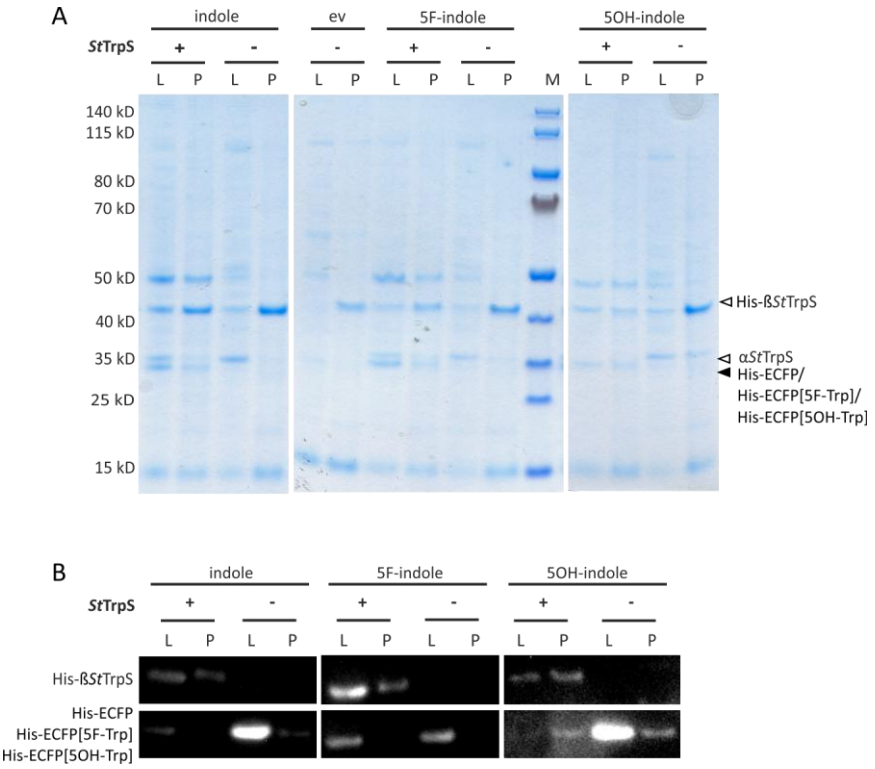


Figure 7: Expression and solubility of the ECFP, ECFP[5F-Trp] and ECFP[5OH-Trp] (A) and the corresponding immunodetection of the hexahistidine-fusion tag (B)
The soluble (L) and insoluble (P) fractions of the whole-cell expression are shown on a 12% Coomassie-stained SDS gel. The calculated molecular weights for His-β*St*TrpS, α*St*TrpS and the His-ECFP/ECFP[5F-Trp]/ECFP[5OH-Trp] are 45 kDa, 29 kDa and 28 kDa, respectively. The supplemented indole analog is labeled with 5F-indole and 5OH-indole. M, molecular weight marker; ev, empty vector control (A).
The soluble (L) and insoluble (P) His-β*St*TrpS and His-ECFP variants were analyzed by Western blot (B).
The presence or absence of the *St*TrpS is indicated with either (+) or (-) and the indole analogs with 5F-indole and 5OH-indole.

We chose 5OH-indole and 5F-indole for a parallel *in vivo* biosynthesis to the corresponding Trp analogs and the incorporation into ECFP. The expression of the ECFP variant was performed with (*St*TrpS(+) strain) and without the co-expression of *St*TrpS (*St*TrpS(-) strain). The expression was performed as described in 3.2.1. The cultures without a co-expression of *St*TrpS and with a supplementation of 5F-indole, grew to a final D_{600} of 3.5. The same strain and with a supplementation of 5OH-indole reached a final D_{600} of 2.5. The *St*TrpS(+) strain, which was supplemented with 5F-indole and 5OH-indole, grew to a final D_{600} of 4.5 and 3.4 respectively. From these observations we concluded that the *St*TrpS in the cells works as efficiently as *in vitro* and the *St*TrpS(+) strain are better provided with intracellular 5F-Trp and 5OH-Trp than the *St*TrpS(-) strain.

An SDS-Gel of the soluble and insoluble proteome of the cultures fed with the 5OH-indole or with the 5F-indole are shown in Figure 7, panel A. The SDS-Gel of the cultures fed with 5F-indole showed an overexpression band for the ECFP[5F-Trp] in the *St*TrpS(+) and the *St*TrpS(-) strain. In contrast, the cells fed with 5OH-indole showed an overexpression band of ECFP[5OH-Trp] only in the *St*TrpS(-) strain. The ECFP[5OH-Trp] band is only very slightly visible in the *St*TrpS(+) strain. These results were confirmed by a Western blot (Figure 7, panel B). The Western blot showed an expression of both variant proteins, albeit it was more pronounced in the *St*TrpS(-) strain.

However, we performed a mass spectrometry analysis to assess if full labeling with the Trp analogs can be achieved without the co-expression of *St*Trp. Therefore, the ECFP[5F-Trp] and ECFP[5OH-Trp], which both carried a hexahistidine fusion-tag, were purified by Ni^{2+} affinity chromatography.

The calculated and found masses of the ECFP[5F-Trp] and ECFP[5OH-Trp] are indicated in Supplementary Table 5, and all possible variants are: none, one or two 5F-Trp and 5OH-Trp, and each of them with and without formed chromophore. The mass spectrometry analysis revealed that all ECFP[5F-Trp] and all ECFP[5OH-Trp], expressed with the *St*TrpS(+) strain, were fully labeled. (Table 2) Without a co-expression of *St*TrpS, only 32% of the ECFP[5F-Trp] were fully labeled and the residual 68% showed single labeling. Also the ECFP[5OH-Trp], which was expressed without the co-expression of *St*TrpS, was mostly unlabeled and only 18% showed an incorporation of one 5OH-Trp residue.

Table 2: Comparison of the (+)*St*TrpS and (-)*St*TrpS strains in terms of 5F-Trp and 5OH-Trp incorporation into ECFP
The ECFP[5OH-Trp] and ECFP[5F-Trp] were expressed with the (+)*St*TrpS and (-)*St*TrpS strains. The relative abundances of incorporated 5F-Trp and 5OH-Trp into ECFP are shown in the table. The data was calculated from the generated data by mass spectrometric analysis. The calculated and found masses are listed in Supplementary Table 5. 'n.d' refers to 'not detected'.

		abundance of species [%]	
variant protein	**species** number of incorporated ncTrp	**in (+)*St*TrpS strain**	**in (-)*St*TrpS strain**
ECFP[5OH-Trp]	0	n.d.	82
	1	n.d.	18
	2	100	n.d.
ECFP[5F-Trp]	0	n.d.	n.d.
	1	n.d.	68
	2	100	32

Thus, the procedure with a clear depletion of Trp, an over-expression of *St*TrpS and with the addition of the ncTrp shortly before the induction of the ECFP, was the first successful incorporation procedure for the incorporation of 5OH-Trp into a fluorescent protein from *Aequoria victoria*.

From our observations, we hypothesize that the overexpression of *St*TrpS prevents the onset of the stringent response even in the absence of Trp. Although the cells were depleted of Trp, which was indicated by the growth arrest, unlabeled and single labeled ECFP[5OH-Trp], as well as single labeled ECFP[5F-Trp] could be found in the mass spectrometry analysis (Supplementary Table 5). We assumed that the *St*TrpS(-) strain could not provide sufficient amounts of 5F-Trp or 5OH-Trp to serve

a recombinant expression under the control of an IPTG inducible T5/*lacO* promoter. Thus, besides the starvation for Trp, the cells also starved for the ncTrp. As a consequence of starvation, the intracellular concentration of uncharged aminoacyl-tRNAs increased, which enhanced the intracellular protein breakdown (Goldberg, 1971: Schlessinger & Ben-Hamida, 1966). Thus, the liberated Trp was incorporated into ECFP as well.

4 CONCLUSION

As one part of my Master Thesis, we established an *in vitro* procedure for the preparative production of ncTrp. The Trp analogs are usually rather expensive or commercially unavailable. To overcome this drawback within our work with ncTrp, we used the *St*TrpS for the enzymatic condensation of simple indole precursors with Ser to ncTrp (Miles, 2001). The *St*TrpS has a broad substrate spectrum (Phillips, 2004) and in place of all available Trp analogs, we chose 5OH-indole and 5F-indole for the enzymatic synthesis. The procedure of the enzymatic synthesis was simple and the purified and dehydrated ncTrp can be stored easily. The "homemade" ncTrp can be used in exact amounts for further incorporation experiments, which was only possible with purchased ncTrp so far.

The other part of my Master Thesis was to establish an enzymatic *in vivo* synthesis of ncTrp and parallel residue-specific incorporation of the synthesized ncTrp into ECFP. This procedure is based on the co-expression of *St*TrpS and so, a separate synthesis of ncTrp is redundant. With this procedure, we achieved the expression of fully labeled model protein, ECFP, with 5OH-Trp, 5F-Trp and $4NH_2$-Trp. urther, we observed that the expression of ECFP variants without the co-expression of *St*TrpS yielded labeled ECFP[ncTrp] but not all ECFP[ncTrp]were fully labeled with the ncTrp.

Based on these findings, it could be interesting to conduct this successful procedure with other model proteins, and to expand the spectrum of used ncTrp. In particular, a model protein which harbors more than two Trp residues would validate the proficiency of our procedure.

This method can serve as an example for other ways of an enzymatic *in vivo* synthesis of ncAA and the parallel residue-specific incorporation into model proteins. An adapted two-step incorporation protocol for other ncAA might be especially attractive for the residue-specific incorporation of expensive ncAAs.

5 REFERENCES

Anderhuber, N. et al., 2016. High-level biosynthesis of norleucine in E. coli for the economic labeling of proteins. *Journal of Biotechnology*.

Bacher, J.M. & Ellington, A.D., 2001. Selection and Characterization of Escherichia coli Variants Capable of Growth on an Otherwise Toxic Tryptophan Analogue. *Journal of Bacteriology*, 183(18), pp.5414–5425.

Bae, J.H. et al., 2003. Expansion of the Genetic Code Enables Design of a Novel "Gold" Class of Green Fluorescent Proteins. *Journal of Molecular Biology*, 328(5), pp.1071–1081.

Bae, J.H. et al., 2001. Incorporation of β-selenolo[3,2-b]pyrrolyl-alanine into proteins for phase determination in protein X-ray crystallography. *Journal of Molecular Biology*, 309(4), pp.925–936.

Betts, M.J. & Russel, R.B., 2003. Amino Acid Properties and Consequences of Substitutions. *Bioinformatics for Geneticists.*, 4, pp.289–304.

Bonskill, P.M. & Wong, T.-F.J., 1988. Suppression of fluorescence of tryptophan residues in proteins by replacement with 4-fluorotryptophan. *Biochem. J*, pp.305 – 308.

Brawerman, G. & Yčas, M., 1957. Incorporation of the amino acid analog tryptazan into the protein of Escherichia coli. *Archives of Biochemistry and Biophysics*, 68(1), pp.112–117.

Budisa, N. et al., 2002. Global Replacement of Tryptophan with Aminotryptophans Generates Non-Invasive Protein-Based Optical pH Sensors. *Angew. Chem. Int. Ed*, (21), pp.4066–4069.

Budisa, N. et al., 1995. High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in Escherichia coli. *European Journal of Biochemistry*, 230(2), pp.788–796.

Budisa, N. et al., 2004. Probing the role of tryptophans in Aequorea victoria green fluorescent proteins with an expanded genetic code. *Biological chemistry*, 385(2), pp.191–202.

Budisa, N., 2004. Prolegomena to Future Experimental Efforts on Genetic Code Engineering by Expanding Its Amino Acid Repertoire Angewandte. *Angew. Chem. Int. Ed*, 43, pp.6426–6463.

Budisa, N. et al., 1998. Residue-specific bioincorporation of non-natural, biologically active amino acids into proteins as possible drug carriers: structure and stability of the per-thiaproline mutant of annexin V. *Proceedings of the National Academy of Sciences of the United States of America*, 95(2), pp.455–9.

Budisa, N. & Pal, P.P., 2004. Designing novel spectral classes of proteins with a tryptophan-expanded genetic code. *Biol Chem*, 385(10), pp.893–904.

Chalfie, M. et al., 1994. Green fluorescent protein as a marker for gene expression. *Science (New York, N.Y.)*, 263(5148), pp.802–5.

Crespo, M.D. & Rubini, M., 2011. Rational design of protein stability: effect of (2S,4R)-4-fluoroproline on the stability and folding pathway of ubiquitin. *PloS one*, 6(5), p.e19425.

Datsenko, K. & Wanner, B., 2000. One-step inactivation of chromosomal genes in Escherichia

coli K-12 using PCR products. *Proceedings of the National …*, 2000.

Desai, T.A. & Rao, C. V, 2010. Regulation of arabinose and xylose metabolism in Escherichia coli. *Applied and environmental microbiology*, 76(5), pp.1524–32.

Dunn, M.F. et al., 2008. Tryptophan synthase: the workings of a channeling nanomachine. *Trends in Biochemical Sciences*, 33(May), pp.254–264.

Ehmann, A., 1977. The van URK-Salkowski reagent — a sensitive and specific chromogenic reagent for silica gel thin-layer chromatographic detection and identification of indole derivatives. *Journal of Chromatography A*, 132(2), pp.267–276.

Gibson, D.G. et al., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5), pp.343–345.

Goldberg, A.L., 1971. A Role of Aminoacyl-tRNA in the Regulation of Protein Breakdown in Escherichia coli. *Proceedings of the National Academy of Sciences*, 68(2), pp.362–366.

Goss, R.J.M. et al., 2006. A convenient enzymatic synthesis of l-halotryptophans. *Chemical Communications*, 36(47), p.4924.

Haid, A. & Suissa, M., 1983. Immunochemical identification of membrane proteins after sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Methods in enzymology*, 96, pp.192–205.

Heim, R., Prasher, D.C. & Tsien, R.Y., 1994. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, 91(26), pp.12501–4.

Heim, R. & Tsien, R.Y., 1996. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology*, 6(2), pp.178–182.

Van Hest, J.C.M., Kiick, K.L. & Tirrell, D. a., 2000. Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *Journal of the American Chemical Society*, 122(7), pp.1282–1288.

Hoesl, M.G. et al., 2015. Chemical Evolution of a Bacterial Proteome. *Angewandte Chemie International Edition*, 54(34), pp.10030–10034.

Hoesl, M.G. et al., 2011. Lipase Congeners Designed by Genetic Code Engineering. *ChemCatChem*, 3(1), pp.213–221.

Hoesl, M.G. & Budisa, N., 2012. Recent advances in genetic code engineering in Escherichia coli. *Current Opinion in Biotechnology*, 23(5), pp.751–757.

Hofmann, M., 2012. *Constitutive expression of the tryptophan synthase of Salmonella typhimurium in E. coli*,

Hong, S.H., Kwon, Y.-C. & Jewett, M.C., 2014. Non-standard amino acid incorporation into proteins using Escherichia coli cell-free protein synthesis. *Frontiers in Chemistry*, 2(June), pp.1–7.

Johnson, J.A. et al., 2010. Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Current opinion in chemical biology*, 14(6), pp.774–80.

Kiick, K.L., van Hest, J.C.M. & Tirrell, D.A., 2000. Expanding the scope of protein biosynthesis

by altering the methionyl-tRNA synthetase activity of a bacterial expression host. *Angewandte Chemie International Edition in English*, 39(12), pp.2148–2152.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259), pp.680–685.

Lakowicz, J.R., 1983. *Principles of Fluorescence Spectroscopy*, Boston, MA: Springer US.

Lepthien, S. et al., 2008. Azatryptophans endow proteins with intrinsic blue fluorescence. *PNAS*, 105(42).

Liu, C.C. & Schultz, P.G., 2010. Adding New Chemistries to the Genetic Code. *Annual Review of Biochemistry*, 79(1), pp.413–444.

Marzluf, T., 2011. *Gamma hydroxybutyrate dehydrogenase (GHBdH) congeners*. Albert-Ludwigs-University of Freiburg.

Miles, E.W., 2001. Tryptophan synthase: a multienzyme complex with an intramolecular tunnel. *Chemical record (New York, N.Y.)*, 1(2), pp.140–51.

Minks, C. et al., 1999. Atomic mutations at the single tryptophan residue of human recombinant annexin V: Effects on structure, stability, and activity. *Biochemistry*, 38(33), pp.10649–10659.

Parsons, J.F. et al., 1998. Enzymes harboring unnatural amino acids: mechanistic and structural analysis of the enhanced catalytic activity of a glutathione transferase containing 5-fluorotryptophan. *Biochemistry*, 37(18), pp.6286–94.

Perni, S. et al., 2013. Optimisation of engineered Escherichia coli biofilms for enzymatic biosynthesis of l-halotryptophans. *AMB Express*, 3(1), p.66.

Phillips, R.S. et al., 1995. Enzymatic synthesis of Thia-L-tryptophans. *Bioorganic & Medicinal Chemistry Letters*, 5(11), pp.1133–1134.

Phillips, R.S., 2004. Synthetic applications of tryptophan synthase. *Tetrahedron: Asymmetry*, 15(18), pp.2787–2792.

Piñero-Fernandez, S. et al., 2011. Indole transport across Escherichia coli membranes. *Journal of Bacteriology*, 193(8), pp.1793–1798.

Richmond, M.H., 1962. The effect of amino acid analogues on growth and protein synthesis in microorganisms. *Bacteriological reviews*, 26, pp.398–420.

Ross, A.J.B., Szabo, A.G. & Hogue, C.W.V., 1997. *Flourescence Spectroscopy*, Elsevier.

Rusconi, F. & Belghazi, M., 2002. Desktop prediction/analysis of mass spectrometric data in proteomic projects by using massXpert. *Bioinformatics*, 18(4), pp.644–645.

Samarkina, O.N. et al., 2009. Universal and rapid method for purification of GFP-like proteins by the ethanol extraction. *Protein expression and purification*, 65(1), pp.108–13.

Schlesinger, S., 1968. The effect of amino acid analogues on alkaline phosphatase. Formation in Escherichia coli K-12. II. Replacement of tryptophan by azatryptophan and by tryptazan. *The Journal of biological chemistry*, 243(14), pp.3877–83.

Schlessinger, D. & Ben-Hamida, F., 1966. Turnover of protein in Escherichia coli starving for nitrogen. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis*, 119(1), pp.171–182.

Seidman, C.E. et al., 2001. Introduction of Plasmid DNA into Cells. In *Current Protocols in Molecular Biology*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 1.8.1–1.8.10. Available at: http://doi.wiley.com/10.1002/0471142727.mb0108s37 [Accessed July 21, 2016].

Singh-Blom, A., Hughes, R. a & Ellington, A.D., 2014. An amino acid depleted cell-free protein synthesis system for the incorporation of non-canonical amino acid analogs into proteins. *Journal of biotechnology*, 178, pp.12–22.

Sloan, M.J. & Phillips, R.S., 1992. ENZYMATIC SYNTHESIS OF AZA-L-TRYPTOPHANS. , (9), pp.1053–1056.

Smith, D.R.M. et al., 2014. The First One-Pot Synthesis of L- 7-Iodotryptophan from 7-Iodoindole and Serine, and an Improved Synthesis of Other L-7-Halotryptophans. *Organic Letters*, 16, pp.2622 – 2625.

Steiner, T. et al., 2008. Synthetic biology of proteins: tuning GFPs folding and stability with fluoroproline. *PloS one*, 3(2), p.e1680.

Szkop, M. & Bielawski, W., 2013. A simple method for simultaneous RP-HPLC determination of indolic compounds related to bacterial biosynthesis of indole-3-acetic acid. *Antonie van Leeuwenhoek*, 103(3), pp.683–91.

Tsien, R.Y., 1998. THE GREEN FLUORESCENT PROTEIN. *Annual Review of Biochemistry*, 67(1), pp.509–544.

Walsh, C.T., O'Brien, R. V. & Khosla, C., 2013. Nonproteinogenic Amino Acid Building Blocks for Nonribosomal Peptide and Hybrid Polyketide Scaffolds. *Angewandte Chemie International Edition*, 52(28), pp.7098–7124.

Welch, M. & Phillips, R.S., 1999. Enzymatic syntheses of 6-(4H-Selenolo[3,2-b]pyrrolyl)-L-alanine, 4-(6H-selenolo[2,3-b]pyrrolyl)-L-alanine, and 6-(4H-furo[3,2-b]pyrrolyl)-L-alanine. *Bioorganic & Medicinal Chemistry Letters*, 9(5), pp.637–640.

Wilcox, M., 1974. The Enzymatic Synthesis of L-Tryptophan. *Reactions*, 440, pp.436–440.

Winn, M. et al., 2008. A convenient one-step synthesis of L -aminotryptophans and improved synthesis of 5-fluorotryptophan. *Bioorganic & medicinal chemistry letters*, 18(16), pp.4508–10.

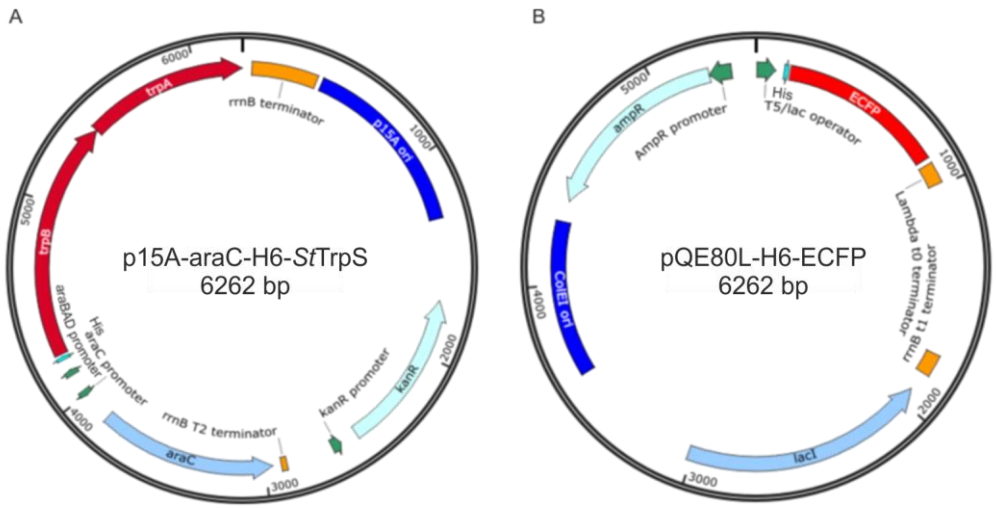
Wolschner, C. et al., 2009. Design of anti- and pro-aggregation variants to assess the effects of methionine oxidation in human prion protein. *Proceedings of the National Academy of Sciences of the United States of America*, 106(19), pp.7756–61.

Yemm, E.W., Cocking, E.C. & Ricketts, R.E., 1955. The determination of amino-acids with ninhydrin. *The Analyst*, 80(948), p.209.

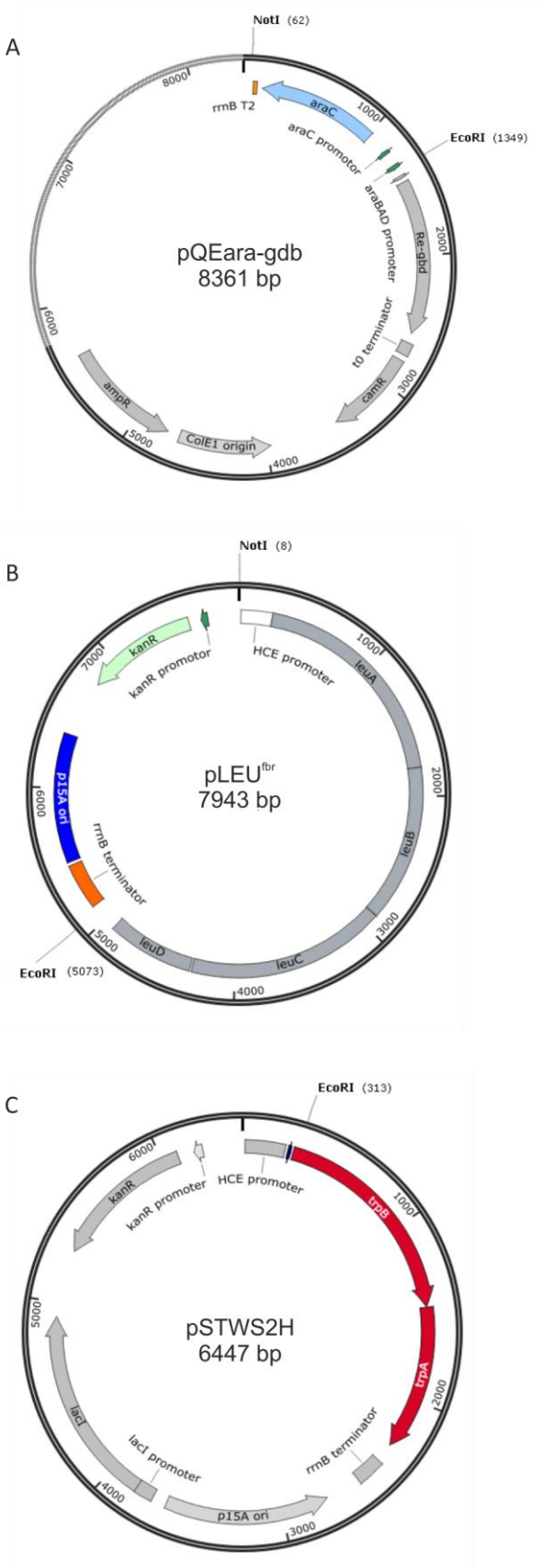
Zheng, S. & Kwon, I., 2012. Manipulation of enzyme properties by noncanonical amino acid incorporation. *Biotechnology journal*, 7(1), pp.47–60.

6 SUPPLEMENTARY INFORMATION

6.1 CLONING OF THE EXPRESSIONS PLASMIDS



Supplementary Figure 1: Plasmid maps for the recombinant expression of *St*TrpS ECFP.
Panel A displays the features of the p15A-AraC-H6-*St*TrpS expression construct. p15A, origin of replication; *kanR*, kanamycin resistance; *araC*, regulator of P_{araBAD} promoter; *araC,* repressor, His, hexahistidine-tag; P_{araBAD}, promoter; *trpB,* coding sequence for the β subunit of *St*TrpS); *trpA,* coding sequence for α subunit of *St*TrpS; rrnB, terminator. Panel B displays the features the pQE80L-H6-ECFP expression construct. ColE1, origin of replication; *ampR,* ampicillin resistance marker; *lacI,* repressor; His, hexahistidine-tag; T5/*lacO,* promoter and lac operator; ECFP/GfFP, the coding sequence of target gene; lambda t0, terminator; rrnB t1, terminator.



Supplementary Figure 2: Plasmid maps for the assembly of the expression construct. p15A-AraC-H6-*St*TrpS.
Panel A displays in its plasmid map the features of the pQEara-gdb. *araC,*repressor; P_{araBAD}, promoter. Panel B shows pLEU^{fbr}. p15A, origin of replication; *kanR*, kanamycin resistance marker; rrnB, terminator; Panel C displays the features the pSTWS2H. *trpB,* coding sequence for the β subunit of *St*TrpS; *trpA,* coding sequence for α subunit of *St*TrpS.
The mentioned features were used for the assembly of the p15A-araC-H6-*St*TrpS.

Supplementary Sequence 1: Nucleotide sequence pQE80L-H6-ECFP.

T5/*lacI* promoter	7-87
6x His	127-144
ECFP	151-870
lambda t0	895-989
rrnBTI	1751-1848
lacI	1936-3018
ColEI	3596-4278
ampR	4373-5233

```
CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCA
ATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGA
GGATCGCATCACCATCACCATCACGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTCACC
GGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTG
TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACC
ACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTGGGGCGTGCAG
TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCC
GAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGC
GCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGAC
TTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACATCAGCCACAAC
GTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCAC
AACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGC
GACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAA
GACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATC
ACTCTCGGCATGGACGAGCTGTACAAGTGATAAAAGCTTAATTAGCTGAGCTTGGACTCC
TGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTT
GCCGCCGGGCGTTTTTTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAGCT
AAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGG
CATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACC
GTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTAT
CCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGGCA
ATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCAT
GAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTT
CTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAA
GGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTT
GATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATAT
TATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGT
GATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAG
GGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCCTGGGGTAATGACT
CTCTAGCTTGAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTT
TATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCCTCTAGATTACG
TGCAGTCGATGATAAGCTGTCAAACATGAGAATTGTGCCTAATGAGTGAGCTAACTTACA
TTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCAT
TAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTT
TCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAG
TTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGT
TAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATC
CGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATC
GTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTTGTTG
AAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCG
AGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCC
CGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGT
ACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAA
TAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGG
ATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACA
GGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGC
GCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGC
AACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTA
ATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGC
CTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTA
TAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGC
CATACCGCGAAAGGTTTTGCACCATTCGATGGTGTCGGAATTTCGGGCAGCGTTGGGTCC
TGGCCACGGGTGCGCATGATCTAGAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAAC
CTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGC
AGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACC
CAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTG
TACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACC
GCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGC
GGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATA
ACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCG
CGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCT
CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAA
GCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTC
TCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGT
AGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGG
```

CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT
TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGC
TGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCG
CTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTC
AAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT
AAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAA
AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT
GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG
CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAG
CCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTA
ATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTG
CCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCG
GTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCT
CCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTA
TGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTG
GTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCC
CGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTG
GAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGA
TGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTG
GGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAAT
GTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTC
TCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCA
CATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCT
ATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAC

Supplementary Sequence 2: Nucleotide sequence p15A-AraC-H6-*St*TrpS.

rrnB	50-375
p15A origin	403-1314
kanR	1723-2655
rrnB T2	2894-2921
araC	2964-4021
P_{araBAD}	4118-4145
6x His	4203-4220
trpB	4233-5426
trpA	5426-6232

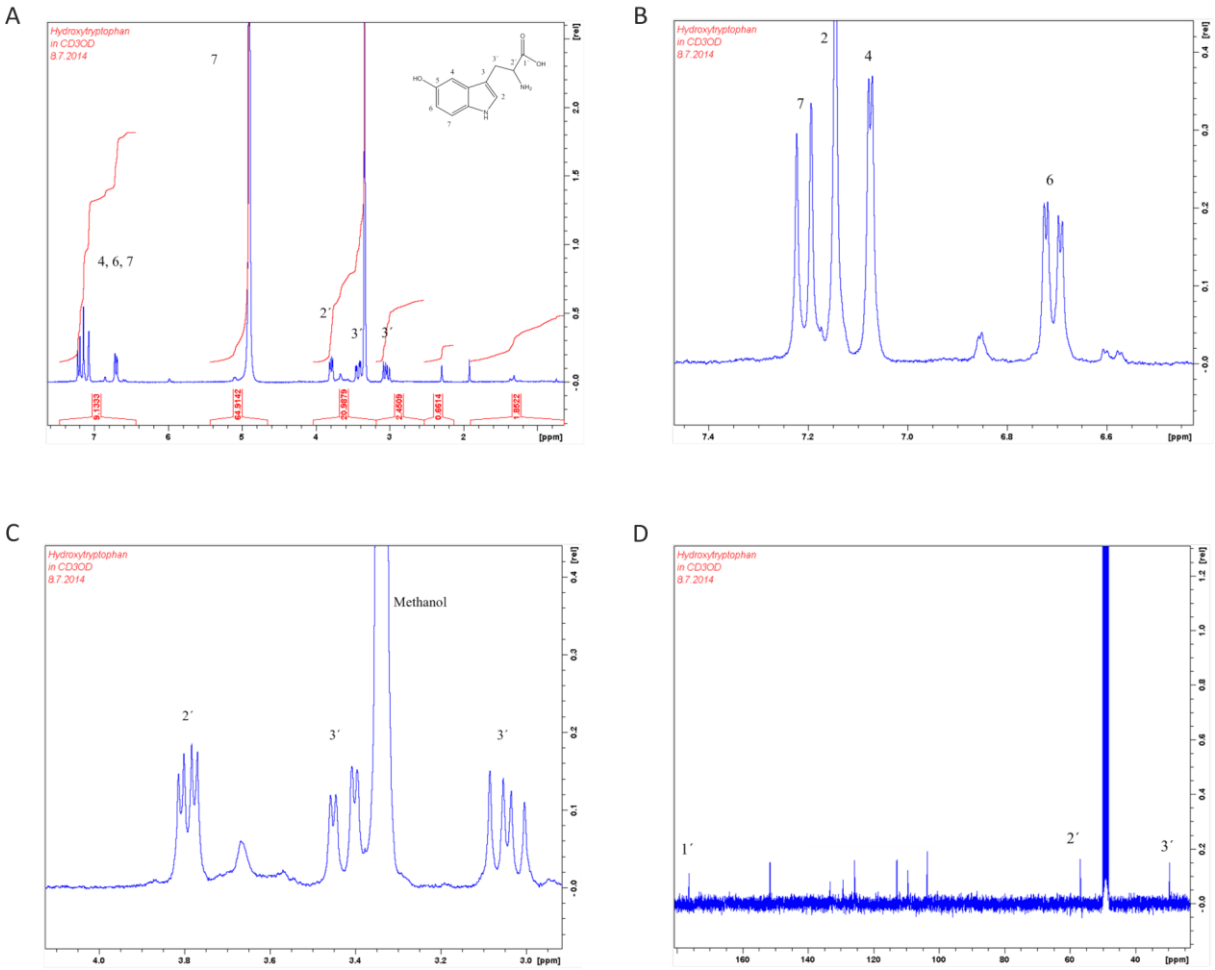
TTAATTAAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGC
AGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCC
GATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACG
AAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCT
CCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGG
GTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCT
GACGGATGGCCTTTTTGGATAAGCTGTCAAACATGAGAATTAACAACTTATATCGTATGG
GGCTGACTTCAGGTGCTACATTTGAAGAGATAAATTGCACTGAAATCTAGAAATATTTTA
TCTGATTAATAAGATGATCTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTG
AAAACGAAAAAACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTT
TGAACCGAGGTAACTGGCTTGGAGGAGCGCAGTCACCAAAACTTGTCCTTTCAGTTTAGC
CTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAGTGGCTGCTGCC
AGTGGTGCTTTTGCATGTCTTTCCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGAACTGCCTAC
CCGGAACTGAGTGTCAGGCGTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACC
GGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCGGGTTTCGCCACCACTGATTTGAGCGTCAGATTTCGTGATGC
TTGTCAGGGGGGCGGAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGT
TAAGTATCTTCCTGGCATCTTCCAGGAAATCTCCGCCCCGTTCGTAAGCCATTTCCGCTC
GCCGCAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGT
ATCACATATTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCACATGAAGCACTTC
ACTGACACCCTCATCAGTGCCAACATAGTAAGCCAGTATACACTCCGCTAGCGCTGATGT
CCGGCGGTGCTTTTGCCGTTACGCACCACCCCGTCAGTAGCTGAACAGGAGGGACAGGGT
CGACCAAAGCGGCCATCGTGCCTCCCCACTCCTGCAGTTCGGGGGCATGGATGCGCGGAT
AGCCGCTGCTGGTTTCCTGGATGCCGACGGATTTGCACTGCCGGTAGAACTCCGCGAGGT
CGTCCAGCCTCAGGCAGCAGCTGAACCAACTCGCGAGGGGATCGAGCCCGGGGTGGGCGA
AGAACTCCAGCATGAGATCCCCGCGCTGGAGGATCATCCAGCCGGCGTCCCGGAAAACGA
TTCCGAAGCCCAACCTTTCATAGAAGGCGGCGGTGGAATCGAAATCTCGTGATGGCAGGT
TGGGCGTCGCTTGGTCGGTCATTTCGAACCCCAGAGTCCCGCTCAGAAGAACTCGTCAAG
AAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAA
GCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTC
CTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATT
TTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCACGACGAGATCCTCGCCGTC
GGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTC
GTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCG
ATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCAT
TGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTG

CCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCAC
AGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGCAG
TTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCTGA
CAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTCATAGCCGAA
TAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATCCATCTTGTTCAAGCATGCG
AAACGACCGTCATCCTGTCTCTTGATCAGATCTTGATCCCCTGCGCCATCAGATCCTTGG
CGGCAAGAAAGCCATCCAGTTTACTTTGCAGGGCTTCCCAACCTTACCAGAGGGCGCCCC
AGCTGGCAATTCCGGTTCGCTTGCTGTCCATAAAACCGCCCAGTCTAGCTATCGCCATGT
AAGCCCACTGCAAGCTACCTGCTTTCTCTTTGCGCTTGCGTTTTCCCTTGTCCAGATAGC
CCAGTAGCTGACATTCATCCGGGGTCAGCACCGTTTCTGCGGACTGGCTTTCTACGTGTT
CCGCTTCCTTTAGCAGCCCTTGCGCCCTGAGTGCTTGCGGCAGCGTGAAGCTTATCGATG
CGGCCGCCTCGAGAAAAGGCCATCCGTCAGGATGGCCTTCTTCCGCGCACATTTCCCCGA
AAAGTGCCACCTGCATCGATTTATTATGACAACTTGACGGCTACATCATTCACTTTTTCT
TCACAACCGGCACGGAACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAATACCCGCGAG
AAATAGAGTTGATCGTCAAAACCAACATTGCGACCGACGGTGGCGATAGGCATCCGGGTG
GTGCTCAAAAGCAGCTTCGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAAGACGCTA
ATCCCTAACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGACAAGCAAACATGCTGT
GCGACGCTGGCGATATCAAAATTGCTGTCTGCCAGGTGATCGCTGATGTACTGACAAGCC
TCGCGTACCCGATTATCCATCGGTGGATGGAGCGACTCGTTAATCGCTTCCATGCGCCGC
AGTAACAATTGCTCAAGCAGATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGC
CCGGCGTTAATGATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCGCTTCATCCGGG
CGAAAGAACCCCGTATTGGCAAATATTGACGGCCAGTTAAGCCATTCATGCCAGTAGGCG
CGCGGACGAAAGTAAACCCACTGGTGATACCATTCGCGAGCCTCCGGATGACGACCGTAG
TGATGAATCTCTCCTGGCGGGAACAGCAAAATATCACCCGGTCGGCAAACAAATTCTCGT
CCCTGATTTTTCACCACCCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATT
CCCAGCGGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTTAAACCC
GCCACCAGATGGGCATTAAACGAGTATCCCGGCAGCAGGGGATCATTTTGCGCTTCAGCC
ATACTTTTCATACTCCCGCCATTCAGAGAAGAAACCAATTGTCCATATTGCATCAGACAT
TGCCGTCACTGCGTCTTTTACTGGCTCTTCTCGCTAACCAAACCGGTAACCCCGCTTATT
AAAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGTC
TATAATCACGGCAGAAAAGTCCACATTGATTATTTGCACGGCGTCACACTTTGCTATGCC
ATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTGACGCTTTTTATCGCAACTCTCT
ACTGTTTCTCCATACCCGTTTTTTTGGTACCGGAAAAAGGAGATCTGCATATGAGAGGAT
CGCATCACCACCACCACCATAGCAGCGGCCATATGACAACACTTCTCAACCCCTACTTTG
GTGAATTCGGCGGCATGTATGTGCCGCAGATCCTGATGCCTGCGCTGAACCAGCTTGAAG
AGGCCTTCGTCAGCGCGCAAAAAGATCCTGAATTTCAGGCGCAATTCGCCGATCTGCTAA
AAAACTACGCGGGACGCCCCACCGCGCTGACGAAATGCCAGAACATTACCGCCGGTACGC
GTACCACGTTGTATTTAAAGCGCGAAGATCTACTGCACGGCGGCGCGCACAAAACCAATC
AGGTACTGGGTCAGGCGCTGCTGGCCAAACGGATGGGTAAAAGCGAGATTATCGCTGAAA
CCGGCGCCGGTCAGCACGGCGTCGCCTCTGCGCTCGCCAGCGCCCTGCTGGGTCTGAAAT
GCCGTATCTATATGGGCGCCAAAGACGTTGAGCGCCAGTCGCCGAACGTCTTCCGTATGC
GTCTGATGGGCGCTGAGGTCATCCCGGTTCATAGCGGCTCCGCTACGCTAAAAGATGCCT
GTAACGAGGCGCTGCGCGACTGGTCCGGTAGTTACGAAACCGCGCACTATATGCTCGGCA
CGGCGGCAGGACCGCATCCCTATCCCACCATCGTTCGCGAGTTCCAGCGCATGATTGGCG
AAGAGACGAAAGCGCAAATCCTCGACAAAGAGGGCCGTCTGCCAGATGCCGTTATCGCTT
GCGTCGGTGGCGGCTCAAACGCTATCGGGATGTTTGCCGATTTTATTAATGATACCAGCG
TCGGGCTAATAGGCGTTGAACCTGGTGGTCATGGTATTGAAACCGGCGAGCATGGCGCGC
CGCTTAAACATGGTCGCGTTGGCATCTATTTCGGGATGAAAGCGCCGATGATGCAAACAG
CAGACGGGCAAATTGAAGAGTCCTATTCCATTTCCGCCGGGCTCGATTTCCCGTCCGTTG
GGCCGCAGCATGCGTACCTGAACAGCATCGGACGCGCGGATTATGTCTCCATTACCGATG
ATGAGGCGCTGGAAGCCTTCAAAACGTTGTGCCGCCATGAGGGAATTATCCCGGCGCTGG
AGTCCTCCCACGCGTTGGCGCACGCTCTGAAAATGATGCGCGAGCAGCCGGAAAAAGAGC
AACTGCTGGTGGTCAATCTCTCTGGCCGCGGAGATAAAGACATCTTTACCGTACACGATA
TCCTGAAAGCGCGAGGGGAAATCTGATGGAACGCTACGAAAATTTATTTGCCCAACTCAA
CGATCGCCGGGAAGGCGCTTTTGTCCCCTTCGTGACCCTGGGCGACCCTGGCATTGAACA
GTCACTGAAAATTATTGACACACTGATTGACGCCGGCGCCGACGCTCTAGAACTGGGGGT
TCCCTTCTCCGATCCGCTGGCCGATGGCCCTACCATCCAGAATGCGAACTTACGCGCCTT
CGCCGCTGGCGTCACGCCGGCTCAGTGTTTTGAAATGCTGGCGCTGATTCGTGAAAAACA
CCCGACCATTCCGATTGGCCTGCTAATGTACGCGAATCTGGTGTTCAATAACGGCATAGA
TGCGTTCTATGCCCGTTGTGAACAGGTTGGCGTAGATTCCGTGCTGGTCGCAGATGTCCC
GGTTGAAGAATCGGCCCCCTTCCGCCAGGCAGCGTTACGGCATAATATCGCGCCGATCTT
CATCTGCCCGCCAAATGCGGATGACGATCTTCTGCGCCAGGTCGCATCTTACGGCCGCGG
TTACACCTACCTGCTTTCGCGTTCGGGTGTCACCGGCGCGGAAAACCGTGGCGCATTGCC
GTTGCATCATCTCATTGAGAAGCTTAAAGAGTACCATGCCGCGCCTGCGTTACAGGGCTT
CGGTATCTCCTCGCCGGAACAGGTGTCTGCGGCCGTGCGTGCCGGGGCGGCTGGCGCTAT
CTCCGGCTCAGCCATTGTCAAGATTATCGAGAAAAACCTCGCGTCTCCCAAACAGATGTT
GGCGGAGCTCAGGTCCTTTGTCTCAGCCATGAAAGCCGCCAGCCGCGCATAA

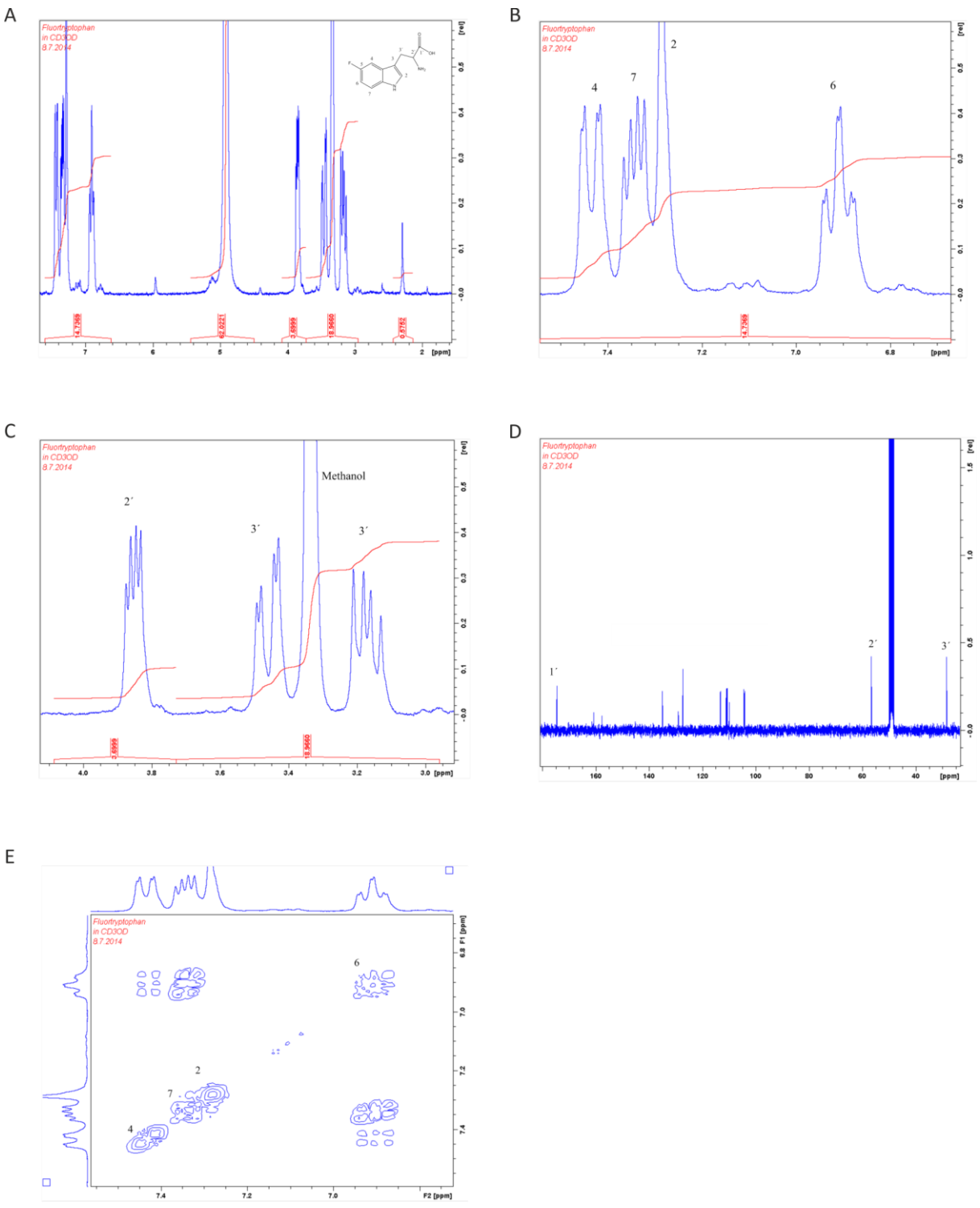
Supplementary Table 1: Used primers.

Primer name	Sequence
Primer 1	CGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTTGGTACCGGAAAAAGGAGATCTGCATATGAG
Primer 2	TGTTTTATCAGACCGCTTCTGCGTTCTGATTTAATCTTAATTAATTATGCGCGGCTGGCG
Primer 3	CGCATCACCATCACCATCACGGATCCATGGTGAGCAAGGGCGAG
Primer 4	CAAGCTCAGCTAATTAAGCTTTTATCACTTGTACAGCTCGTCCATGC

6.2 NMR SPECTRA OF 5OH-TRP AND 5F-TRP



Supplementary Figure 3: ^{1}H NMR spectra of 5OH-Trp.



Supplementary Figure 4: ^{19}F NMR spectra of 5F-Trp.

Supplementary Sequence 3: Amino acid sequence of ECFP.

Single letter code of the N-terminally hexahistidine-tagged ECFP (251 amino acids, 2 Trp).The Trp residues are highlighted in blue with the corresponding amino acid residue numbers in subscript.

MRGSHHHHHHGSM$_1$VSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVP**W$_{57}$**PTLVTTLT**W$_{66}$**GVQCFSRYPDHM
KQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYISHNVYITADKQKNGIKANFKIRHNIE
DGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

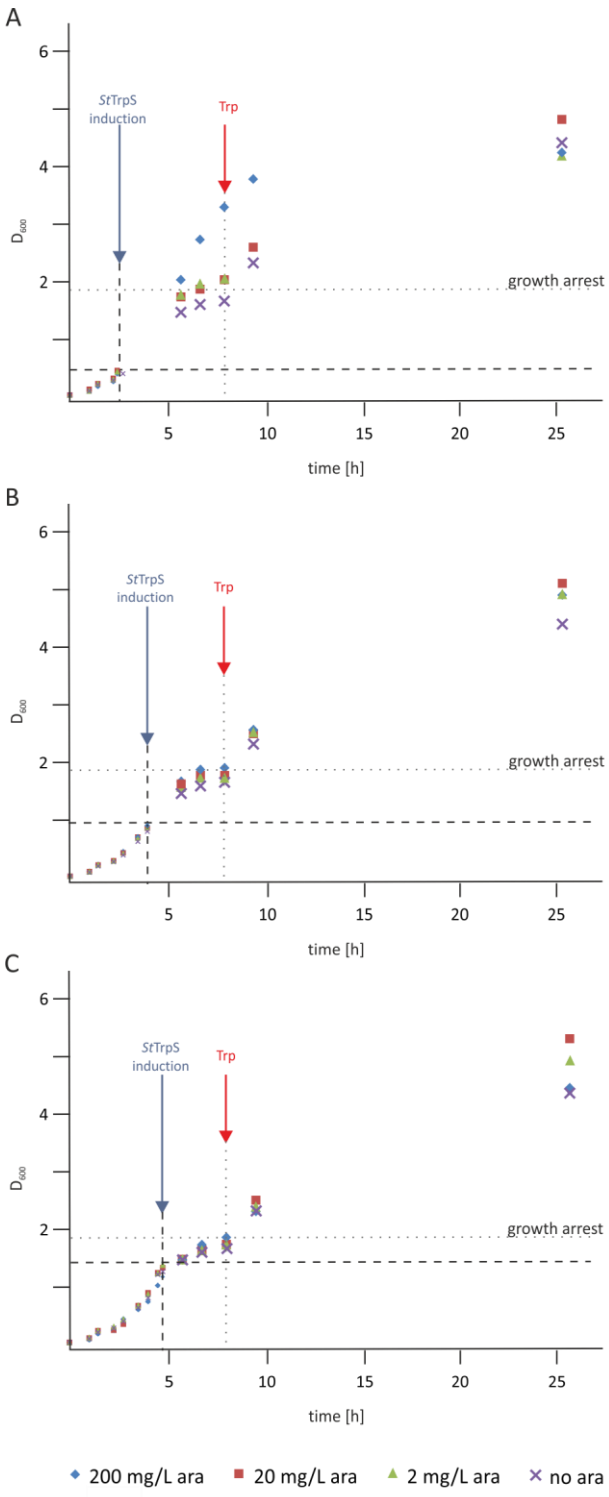
6.3 EXPRESSION AND PURIFICATION OF ECFP[NCTRP]

Supplementary Table 2: Composition of 5x SDS loading dye.
All chemicals are commercially available at Roth.

Component	Concentration
Sodium dodecyl sulfate	5 % (w/v)
Glycerol	1.6 % (v/v)
Bromphenol blue	0.4 % (w/v)
β-mercaptoethanol	0.16 % (v/v)
Tris/Cl pH 6.8	6.4 mM

Supplementary Table 3: Minimal medium composition
The chemicals are commercially available at Roth, unless indicated otherwise.

Description	Component	Concentration
Trace elements	$FeSO_4 \cdot 7\ H_2O$	9 µM
	$MnSO_4 \cdot H_2O$	3.5 µM
	$AlCl_3 \cdot 6\ H_2O$	2.5 µM
	$CoCl_2 \cdot 6$ H2O	2 µM
	$ZnSO_4 \cdot 7\ H_2O$	0.4 µM
	$Na_2MoO_4 \cdot 2\ H_2O$	0.5 µM
	$CuCl_2 \cdot 2\ H_2O$	0.4 µM
	H_3BO_3	0.5 µM
Salts	Na_2HPO_4	47.8 mM
	KH_2PO_4	22.0 mM
	NaCl	8.6 mM
	NH_4Cl	18.6 mM
Other ingredients	Glucose	20.0 mM
	$MgSO_4 \cdot 7\ H_2O$	1.0 mM
	$CaCl_2 \cdot 2\ H_2O$	7 µM $CaCl_2$
	19 amino acid supplementation (without Trp)	50 mg/L
	Trp	18 µM
	D(+)-Biotine	4.0 µM
	Thiamine hydrochloride	3.3 µM
	Ampicillin (Sigma-Aldrich)	100 mg/L
	Kanamycin	50 mg/L



Supplementary Figure 5: Effect of the different arabinose concentrations on growth of the Trp auxotrophic *E. coli* strain during the co-expression of the tryptophan synthase from *Samonella typhimurium* (*St*TrpS).
The expression of *St*TrpS was induced with different arabinose concentrations: 200 mg/L, 20 mg/L and 2 mg/L at D_{600} of 0.5 (A), 1 (B) and 1.5 (C). The induction at D_{600} of 0.5 with 200 mg/L arabinose showed a different growth behavior compared to the other cultures, which were induced with less arabinose (A). We observed homogeneous growth patterns of the expression strains, when the defined arabinose concentrations were added at D_{600} of 1 (B) and 1.5 (C). The addition of 25 µM Trp after the growth arrest (dotted horizontal line) is marked with an arrow. D_{600} values were determined in triplicates and the calculated relative mean deviations are shown.

6.4 MASS ANALYSIS OF ECFP[NCTRP]

Supplementary Table 4: Calculated and found masses for ECFP variants produced with the one-step incorporation protocol and the two-step incorporation protocol.
An undefined mass peak was detected at 28,281.19 (labeled with *) in the sample produced with the one-step incorporation protocol. This was not the case with the two-step incorporation protocol.

ECFP[ncTrp]	chromophore formation	$mass_{calculated}$ / Da	number of $ncTrp_{incorporated}$	$mass_{found}$ with protocol two-step incorporation protocol / Da	$mass_{found}$ with protocol one-step incorporation protocol / Da
ECFP[4NH_2-Trp]	no	28,286.12	none	n.d.	n.d.
		28,301.13	1	n.d.	n.d.
		28,316.14	2	n.d.	n.d.
	yes	28,266.09	none	n.d.	28,265.20
		28,281.10	1	n.d.	28,280.29*
		28,296.11	2	28,296.26	28,296.28

Supplementary Table 5: Calculated and found masses for ECFP[5OH-Trp] and ECFP[5F-Trp] expressed with (+)*St*TrpS and (-)*St*TrpS strains.

ECFP[ncTrp] variant	chromophore formation	$mass_{calculated}$ / Da	number of $ncTrp_{incorporated}$	$mass_{found}$ in (+)*St*TrpS / Da	$mass_{found}$ in (-)*St*TrpS / Da
ECFP[5OH-Trp]	no	28,286.12	none	n.d.	n.d.
		28,302.11	1	n.d.	n.d.
		28,318.11	2	n.d.	n.d.
	yes	28,266.09	none	n.d.	28,266.47
		28,282.08	1	n.d.	28,282.44
		28,298.08	2	28,297.48	n.d.
ECFP[5F-Trp]	no	28,286.12	none	n.d.	n.d.
		28,304.11	1	n.d.	28304.36
		28,322.100	2	n.d.	n.d.
	yes	28,266.09	none	n.d.	n.d.
		28,284.08	1	n.d.	28,284.40
		28,302.07	2	28302.46	28,301.65

6.5 LISTED DEVICES, INSTRUMENTS CHEMICALS AND REAGENTS

Supplementary Table 6: Used instruments and devices.

Instrument /Device	Supplier
Analytical scale	Sartorius, Göttingen, Germany
Avanti J-20 XP centrifuge	Beckmann Coulter Inc.; Brea, CA
Centrifuge tubes	Thermo Fischer Scientific Inc.; Waltham, MA
Centrifuges	Centrifuge 5415R: Eppendorf; Hamburg, Germany Centrifuge 5424: Eppendorf
ddH20 device (arium®basic)	Sartorius
Eppendorf tubes	Sarstedt, Nümbrecht, Germany
Flasks	2000 mL (Schott Duran); Bartelt; Graz, Austria 250 mL (Schott Duran); Bartelt
HPLC	Agilent 1100, Agilent Technologies, Santa Clara, CA
Incubator HT MultitronII	InforsAG, Bottmingen, Switzerland
Lab scale	Binder GmbH, Tuttlingen, Germany
Laminar air flow hood AirClean	Woerden, Netherlands
Petri dishes	Greiner Bio-one International AG, Kremsmünster, Austria
Photometer	Beckman Clouter Inc BioPhotometer; Eppendorf
Pipette tips	Greiner Bio-one International AG
Pipettes	1000 µL, 200 µL, 20 µL; Denville; South Plainfield, NJ 10 µL (Biohit); Sartorius
Plate reader(SPECTRAmax Plus384)	Molecular Devices, Sunnyvale, CA
Scanner	Tevion USB Scanner, Mühlheim, Germany
Sonifier	Branson, Danbury, CT
Sterile filters	Sartorius
UV cuvettes	Greiner Bio-one International AG
Vortex	IKA®-Werke GembH & Co. K; Staufen, Germany
ZORBAX Eclipse XDB-C8	Agilent Technologies, Santa Clara, CA

Supplementary Table 7: Used chemicals, regents and enzymes.

Chemical /Reagent /Enzyme	Cat. Nr./ order code	Supplier
α-D-glucose monohydrate	6780.2	Roth
β-mercaptoethanol	4227.1	Roth
β-Nicotinamid-adenin-dinucleotid	AE11.2	Carl Roth
1,4-dithiotreit (DTT)	6908.1	Roth
4-Aminoindole	5192-23-4	ABCR, Karlsruhe, Germany
5-Fluoroindole	399-52-0	Molekula
5-Hydroxyindole	1953-54-4	Molekula
Acetic Acid	6755.2	Roth
Acetonitrile	75-05-8	Roth
Acrylamide / Bis	161-0156	BioRad
Agarose	840004	BioZyme; Hessisch-Oldendorf, Germany
Amersham Protran 0.45 NC	10600002	GE Helathcare, Little Chalfont, UK
Ammonium chloride	6923-52-0	Roth
Ammonium Persulfate (APS)	13375.01	Serva; Heidelbrerg, Gemany
Ammonium Sulfate	7783-20-2	Roth; Karlsruhe, Gemany
Bovine Serum Albumin	9048-46-8	Sigma-Aldrich
Bradford protein assay	5000201	BioRad, Vienna, Austria
Calciumchloride dihydrate	CN93.1	Roth
Cell lytic B reagent	MFCD02097909	Sigma-Aldrich
Coomassie Blue R-250	6104-59-2	Roth
D(+)-Biotin	58-85-5	Roth
Di-sodium hydrogen phosphate	T875.2	Roth
DNAse I	DN25-100mg	Sigma-Aldrich
dNTPs	K039.1	Roth
D-α-Aminobenzylpenicillin sodium salt	69-52-3	Roth
Ethanol	20821.330	VWR international, Radnor, PA
FastAP Thermosensitive Alkaline Phosphatase (1 U/µL)	EF0654	Thermo Fisher Scientific
FD EcoRI	FD0274	Thermo Fisher Scientific
FD NotI	FD0596	Thermo Fisher Scientific
FD XhoI	FD0695	Thermo Fisher Scientific
Fermentas Gene Jet TM Plasmid	K0502	Thermo Fisher Scientific
Glycerol	3908.3	Roth
Hydrochloric acid, fuming	4625.1	Roth
Imidazole	1047161000	Merck, Billerica, MA
Indole	120-72-9	Sigma-Aldrich
Iron(III)chloride	7705-08-0	Roth
Isopropyl β-D-1thiogalactopyranoside (IPTG)	CN03.3	Roth
Kanamycin sulfate	T832.2	Roth
L-(+)-Arabinose	5328-37-0	Sigma-Aldrich
LB-Agar Lennox	X65.3	Roth
LB-medium Lennox	X964.2	Roth
L-Tryptophan	13831082-100G	Molekula
Lysozyme	8259-2	Roth
Magnesium chloride	8.14733.0500	Merck
Magnesium sulfate heptahydrate	A537.4	Roth
N, N, N', N'-tetramethylethylenediamine	161-0800	BioRad
nButanol	71-36-3	Roth
Ninhydrin	485-47-2	Roth
Ni-NTA agarose beads	30250	Qiagen, Hilden, Germany
NuPAGE sample buffer	NP0008	Thermo Fisher Scientific
PageRuler prestained protein ladder	SM0671	Thermo Fisher Scientific
P-dimethylaminibenzaldehyde	100-10-7	Sigma-Aldrich
PEG-8000	P5413-500G	Sigma-Aldrich
Phusion® High-Fidelity DNA Polymerase 2U/µl	M0530S	New England Biolabs
Potassium dihydrogen phosophate	P018.2	Roth
Propanol	71-23-8	Roth
Pyridoxal 5'phosphate monohydrate	41468-25-1	Sigma-Aldrich
Silica gel 60	7631-86-9	Roth
SOB Medium	AE27.1	Roth
Sodium chloride	9265.1	Roth
Sodium dihydrigen phosphate monohydrate	T879.2	Roth
Sodium dodecyl sulfate (SDS)	2326.1	Roth
Sulfuric acid	7664-93-9	Roth
SuperSignal® West Dura Extended Duration Substrate	34075	Life Technologies
T4 Ligase	EL0014	Thermo Fisher Scientific

T5 Exonuclease	162340	BoiZym
TaKaRa Ex Tag Polymerase (5U/ml)	RR001A	Takara Bio, Saint-Germain-en-Laye
Thiamine hydrochloride	67-03-8	Sigma-Aldrich
TLC Silica gel 60 F_{254}	1055490001	Merck Millipore, Billerica, MA
Tris(hydroxymethyl)-aminomethane	1185-53-1	Roth
Triton-X100	3051.3	Roth
Tween20	9005-64-5	Sigma-Aldrich
Urea	57-13-6	Roth
Wizard® SV Gel and PCR Clean-Up System	A9285	Promega; Madison, WI