AUS DEM LEHRSTUHL
FÜR MEDIZINISCHE MIKROBIOLOGIE UND HYGIENE
PROF. DR. RALF WAGNER
DER FAKULTÄT FÜR MEDIZIN
DER UNIVERSITÄT REGENSBURG

Quantification of the suppression efficiency at different positions of Nisin in *E. coli*

Inaugural - Dissertation
zur Erlangung des Doktorgrades
der Medizin

der
Fakultät für Medizin
der Universität Regensburg

vorgelegt von
Miriam Thewes

2020

AUS DEM LEHRSTUHL
FÜR MEDIZINISCHE MIKROBIOLOGIE UND HYGIENE
PROF. DR. RALF WAGNER
DER FAKULTÄT FÜR MEDIZIN
DER UNIVERSITÄT REGENSBURG

Quantification of the suppression efficiency at different positions of Nisin in *E. coli*

Inaugural - Dissertation
zur Erlangung des Doktorgrades
der Medizin

der
Fakultät für Medizin
der Universität Regensburg

vorgelegt von
Miriam Thewes

2020

Dekan:	Prof. Dr. Dirk Hellwig
1. Berichterstatter:	Prof. Dr. Ralf Wagner
2. Berichterstatter:	Prof. Dr. Bernd Salzberger
Tag der mündlichen Prüfung:	10. Juni 2021

Table of Contents

Abstract

With the continuous emergence of resistant bacterial strains, the urgent need for novel antibiotics is unquestionable. Lantibiotics are a group of ribosomally produced and posttranslationally modified peptides, which display antibiotic activity. The lantibiotic nisin, a 34 amino acid peptide, shows antibiotic activity mainly against gram-positive bacteria. By introducing non-canonical amino acids (ncAA) to nisin, chemical properties can be altered, potentially facilitating therapeutic usage or increasing its antibiotic range.

The aim of this work was the identification of suitable positions of nisin for incorporation of the non-canonical amino acid Boc-Lysin (Boc-K) using amber suppression in *E. coli* and quantification of the suppression efficiency.

First, vector constructs containing nisin, fused in-frame with the reporter protein GPF, as well as the assay set up were optimized. Expression of nisin variants was quantified using a fluorescence assay. A screening of all variants indicated high GFP expression levels compared to wild-type nisin constructs. The maximum was measured for position 13 (175 %) and 16 (189 %), depending on the suppressor pair used. However, quality control experiments revealed high fluorescence intensity levels even when the ncAA was not present in the media. Further analysis of the expressed proteins showed that the methionines within nisin allowed internal translation starts after interruption of translation due to the amber stop codon. For this reason, only results from the rear part of nisin, with amber suppression taking place after the last internal methionine were analyzed for the initial objective. These results from stop codon suppression from amino acid position 21 of nisin and further towards the C-terminus showed satisfactory suppression efficiencies compared to nisin wild-type expression, ranging from 15 % to 68 %. Two amino acid positions in nisin were particularly promising regarding protein yields, namely position 21 (66 %) and 31 (68 %), depending on the suppressor pair used.

In summary, this work showed that incorporation of non-canonical amino acids in nisin using amber suppression is feasible and provides satisfactory suppression efficiencies. However, internal translation starts impeded analysis of positions 1 to 20. For avoiding the latter, further work, e.g. substitution of the methionines within nisin with other amino acids is necessary.

Zusammenfassung

Mit kontinuierlicher Zunahme resistenter Bakterienstämme ist die dringende Notwendigkeit neuer Antibiotika unstrittig. Lantibiotika sind eine Gruppe ribosomal produzierter und posttranslational modifizierter Peptide mit antibiotischer Wirkung. Das Lantibiotikum Nisin, ein Peptid aus 34 Aminosäuren, ist überwiegend gegen grampositive Bakterien antibiotisch wirksam. Durch den Einbau nicht-kanonischer Aminosäuren (ncAA) in Nisin können chemische Eigenschaften verändert werden, welche potentiell die therapeutische Anwendung erleichtern oder das antibiotische Wirkspektrum erweitern können.

Das Ziel dieser Arbeit war die Identifikation passender Positionen in Nisin für den Einbau der nicht-kanonischen Aminosäure Boc-Lysin (Boc-K) mittels Amber Suppression in *E. coli* und Quantifizierung der Suppressionseffizienz.

Hierfür wurden zunächst die Vektorkonstrukte, die Nisin und das daran *in-frame* gekoppelte Reporter Protein GFP enthalten, sowie der Versuchsaufbau optimiert. Die Expression der Nisin Varianten wurde durch eine Fluoreszenzmessung quantifiziert. Ein *Screening* aller Varianten zeigte hohe GFP Expressionslevel im Vergleich zu Nisin Wildtyp Konstrukten. Maximalwerte wurden für die Positionen 13 (175 %) und 16 (189 %), abhängig vom verwendeten Suppressorpaar, gemessen. Experimente zur Qualitätskontrolle zeigten jedoch, dass hohe Fluoreszenz-intensitäts-Level gemessen wurden, auch wenn die ncAA nicht im Medium vorhanden war. Weiterführende Analysen der exprimierten Proteine zeigten, dass Methionine in Nisin interne Translationsstarts ermöglichten, nachdem die Translation durch das Amber Stopp Codon abgebrochen war. Aus diesem Grund konnten nur Ergebnisse aus dem hinteren Anteil von Nisin für die ursprüngliche Fragestellung verwertet werden, da die Amber Suppression hier nach dem letzten internen Methionin stattfand. Die Ergebnisse der Stopp Codon Suppression ab Nisin-Position 21 zeigten zufriedenstellende Suppressionseffizienzwerte zwischen 15% und 68% im Vergleich zur Nisin Wildtyp Expression. Zwei Aminosäurepositionen von Nisin zeigten besonders vielversprechende Proteinerträge: Position 21 (66 %) und 31 (68 %), abhängig vom verwendeten Suppressorpaar.

Zusammenfassend konnte in dieser Arbeit gezeigt werden, dass der Einbau von ncAA in Nisin mittels Amber Suppression funktioniert und zufriedenstellende Suppressionseffizienz-werte erzielt wurden. Interne Translationsstarts verhinderten jedoch die Analyse der Positionen 1 bis 20. Um dieses Problem zu umgehen sind weitere Arbeiten nötig, beispiels-weise um die Methionine innerhalb von Nisin durch andere Aminosäuren zu ersetzen.

1. Introduction

1.1 Antibacterial resistance

When Sir Alexander Fleming discovered the antibiotic effect of Penicillin in 1928, a big step towards an effective therapy of formerly potentially deadly diseases caused by bacterial infections was taken (1). Modern medicine as it is known today, which, amongst others includes major surgery, organ transplantation and improvement in childhood survival, would not be possible without the existence of an adequate antimicrobial treatment (2). However, resistance mechanisms in bacteria, viruses, fungi and parasites are endangering the effectiveness of these drugs. For this reason, the World Health Organization (WHO) has listed antimicrobial resistance as one of the "ten threats to global health" in 2019 (3), which require special awareness. The WHO demands an international action plan to combat antimicrobial resistance, as well as a global surveillance plan (4,5).

Although antibacterial resistance has only caught increasing attention in the last few decades, knowledge about the existence of antibacterial resistance is not new. Even before Penicillin, the first antibiotic drug in history was available on the market in 1943, bacterial strains which were immune to the antibiotic agent had already been described (6). Figure 1 provides a brief overview on introduction of antibiotics and identification of antibiotic resistance. When Sir Alexander Fleming was awarded with the Nobel Prize for his achievements in 1945, he warned the audience in his lecture that the wrong use of the drug could cause the appearance of resistant strains, which would not be treatable anymore (7).

The development of antibiotic resistances in bacteria is a natural process, which occurs due to selection pressure. The human application of antibiotic drugs has accelerated the spread of resistant bacterial strains. Misuse has caused an increased selection pressure. Erroneous application in human medicine includes antibiotic treatment of infections which are not caused by sensitive bacteria or even viral infections, as well as not indicated long-term application. Likewise, underuse of antibiotics, such as stopping a treatment early, or the intake of an underdose has contributed to the development of resistant strains. Also, the application of antibiotic drugs in veterinary medicine, food animals and agriculture, particularly when used for growth promotion and disease prevention, facilitates the spreading of resistances (8,9). As a result of decades of use and misuse of antibiotics, an increasing part of multi-drug resistant pathogens can be found, especially in the clinically most relevant bacterial strains (10).

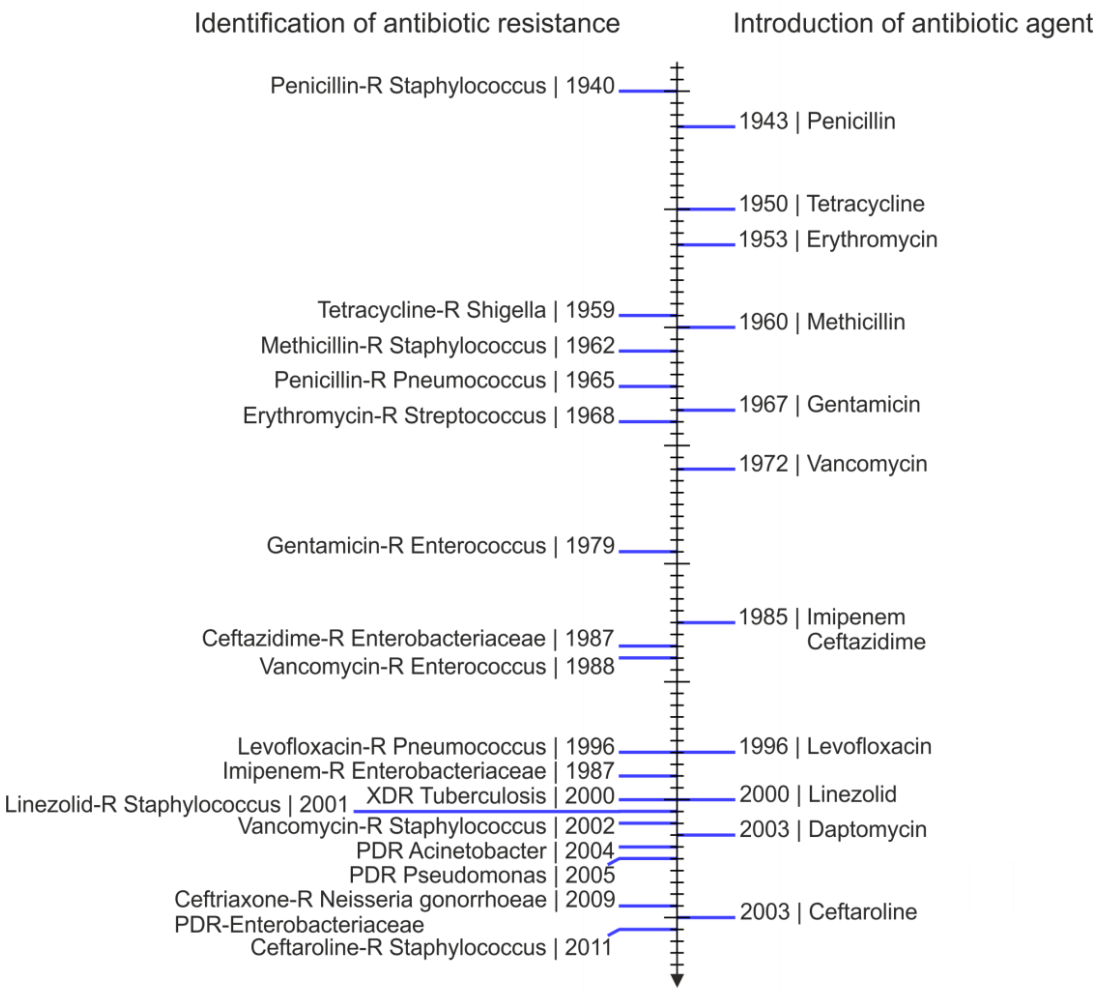

Figure 1: Timeline of introduction of antibiotic agents and identification of resistant bacteria
Antibiotic resistances are displayed on the left side of the timeline. The suffix -R refers to a bacterial strain, which is resistant to the preceding antibiotic drug. XDR is short for extensively drug resistant bacteria and PDR for pandrug resistant bacteria. Figure adapted from U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2013 (11).

Drug- and multidrug resistance is not only a threat to global health, but also causes immense costs. Infections with resistant bacterial strains have been shown to increase hospitalization costs and to lengthen hospital stays. When health service costs and loss of productivity are accounted, it was estimated that in the year 2004, antibacterial resistance cost the United States 55 billion US dollar, which is more than one third of the estimated cost for Diabetes (12,13).

In order to prevent a return of conditions similar to those in the pre-antibiotic era, the need for novel antibiotics, as well as for strategies to prevent the spread and further development of resistant bacterial strains is unquestionable.

1.2 Lantibiotics

Whilst searching for new antibiotic agents, lantibiotics have attracted attention as a new potential class of therapeutics. The term lantibiotic is short for "lanthionine-containing antibiotic" and refers to a group of peptides, which are ribosomally produced by gram-positive bacteria that undergo posttranslational modifications (14). Most lantibiotics display antibiotic activity against gram-positive pathogens, such as staphylococci, streptococci, enterococci and clostridia, while some are also effective against certain gram-negative bacteria, such as *Haemophilus* or *Neisseria* strains (15,16). The eponymous posttranslational modifications (PTM) of this group of peptides creates lanthionine and/or methyl-lanthionine containing ring structures. In this process, serines (Ser) and threonines (Thr) of the ribosomally produced precursor peptide are dehydrated by a specific dehydratase (LanB), forming the rare amino acids 2,3-didehydroalanine (Dha) and (Z)-2,3-didehydrobutyrine (Dhb), respectively. In a second step, cyclization catalyzed by a specific cyclase (LanC) allows the characteristic ring formation by using a cysteine from the propeptide and linking it to one of the above named newly formed amino acids, resulting in lanthionine or methyllanthionine rings, respectively. All lantibiotics are synthesized with a leader sequence, which is removed proteolytically either before, during or after the peptide is exported from the cell (17,18). When the leader is cleaved off, all steps of posttranslational modifications are completed. Figure 2 shows the posttranslational modifications lantibiotics undergo exemplarily for the best-studied lantibiotic, nisin A.

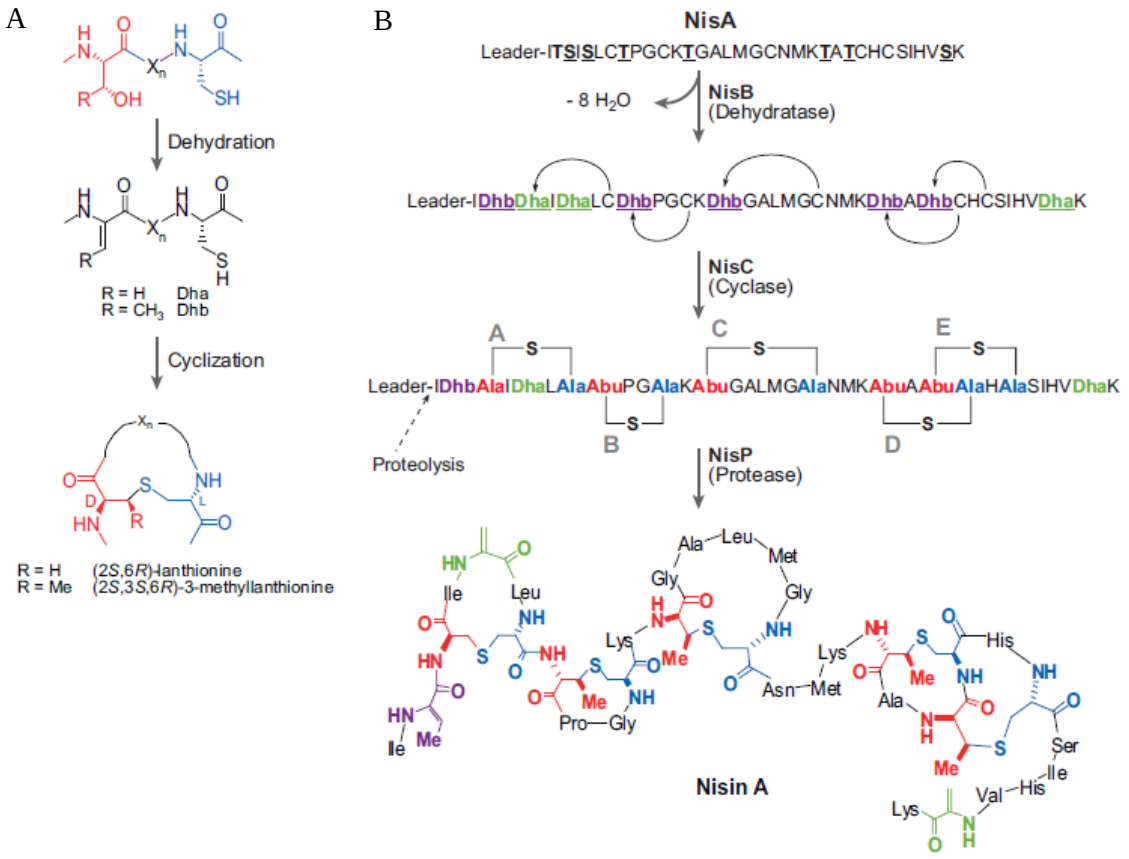

Figure 2: Posttranslational modifications common to all lantibiotics
(A) Serines and Threonines of the propeptides are subjected to dehydration, resulting in 2,3-didehydroalanine (Dha) and (Z)-2,3-didehydrobutyrine (Dhb), respectively. Following, cyclization leads to formation of the eponymous amino acids lanthionine and methyllanthionine and the characteristic ring formation. (B) Posttranslational modifications shown exemplarily at the lantibiotic Nisin A. The precursor peptide first undergoes dehydration, catalyzed by NisB, a dehydratase. In a second step NisC, a specific Cyclase, causes the formation of lanthionine and methyllanthionine rings. Finally, the leader of the propeptide is cleaved off by the protease NisP, resulting in the fully active nisin A. In the lanthionine and methyllanthionine rings, parts that derived from Ser/Thr are shown in red and parts that derived from Cys are blue. Dha is marked green and Dhb blue. Figure from Willey et al. (17).

Once exported to the extracellular space, lantibiotics act mainly against gram-positive bacteria, since their target site is lipid II, which is part of the pepdidoglycan cell wall of gram-positive bacteria (19). In gram negative bacteria, the lipid II containing part of the cell wall is enclosed by an outer membrane. For this reason, lantibiotics are generally less effective against gram negative pathogens, unless special circumstances are given, for example treatment with an agent facilitating access to lipid II prior to exposure to the lantibiotic (20). Two main modes of action for lantibiotics are known to date. Inhibition of cell wall synthesis as well as pore formation in the cell wall can be induced after binding to lipid II (21,22). Whether only one of these modes of actions or a combination of the two is used depends on the specific lantibiotic.

In order not to be harmed by their own product, producer strains of lantibiotics are protected by lantibiotic immunity. Two mechanisms are known to ensure self-protection, namely specific immunity peptides and a specialized ATP binding cassette (ABC) transporter system. In some cases both mechanisms are used simultaneously (23–25). The provided protection is highly specific and ensures immunity against the produced lantibiotic, yet usually not against other lantibiotics produced by different bacterial strains. Cross-immunity against other lantibiotics is rare (26), which is one of the aspects that likely contributes to low levels of resistance against lantibiotics.

Known resistance mechanisms against lantibiotics include general resistance mechanisms such as alteration of the cell wall or production of biofilms, which are common responses to the presence of antimicrobials (27). However, in some cases very specific resistance mechanisms against certain lantibiotics have been identified. For example, a nisin resistance protein has been shown to be produced by some non-nisin-producer strains of *Lactococcus lactis*. The nisin resistance protein cleaves the last six amino acids of nisin, leading to a truncated and significantly less effective bactericidal lantibiotic (28).

Generally speaking, the tendency to generate resistances against lantibiotics currently appears to be low (29), which allows the group of lantibiotics to be seen as a promising candidate for future alternatives to established antibiotic drugs.

1.3 Nisin

Due to its early discovery, nisin is one of the best-studied lantibiotics to date. It was first described in 1928 in milk cultures (30) and has been in commercial use since the 1950s in England as a food-preservative in dairy products (31). Nisin is considered a safe food additive and was approved by the Joint Food and Agriculture Organization/World Health Organization in 1969. Almost 20 years later, it was also approved by the Food and Drug Administration of the United States of America and today is used in more than 50 countries in the food industry (32). During the long history of nisin's use as food preservative it was observed that the antibiotic action of nisin was not limited to typical foodborne bacteria, but also effective against a much broader spectrum of pathogens (33).

Nisin A is a 34 amino acid protein produced by some strains of *Lactococcus lactis* and undergoes the typical posttranslational modifications as described above (see chapter 1.2, Figure 2). Several naturally occurring nisin variants, such as nisin Z (34) or nisin Q (35), have been isolated from different producing strains (36).

Bioengineering of this lantibiotic by modifying the nisin encoding gene has led to even more variants. Attempts were made to influence specific properties of the protein, such as solubility or antibiotic activity and spectrum (37). For example, certain bioengineered variants displayed higher antibiotic activity against gram-negative pathogens such as *Salmonella* or *Pseudomonas* (38) or enhanced antibiotic activity against clinically relevant resistant pathogens, e.g. methicillin-resistant *Stapylococcus aureus* (MRSA) and vancomycin-resistant *enterococci* (VRE) (39).

1.4 Expanding the genetic code by introduction of non-canonical amino acids

The possibilities of changing a peptide's properties by altering the encoding gene are restricted due to the limited number of amino acids that can naturally be encoded by the genetic code, i.e. the twenty proteinogenic amino acids. Yet, there are numerous non-canonical amino acids (ncAA) which bear possibly interesting chemical functions, such as fluorescence, photoreactivity or redox-activity (40). By incorporation of non-canonical amino acids into lantibiotics their biochemical properties could be changed in a desired way. However, due to the complex structure of lantibiotics, chemical synthesis of these proteins suffers from low yields (41–43). For this reason, two main methods for *in vivo* incorporation of ncAA were developed. Their main difference lies in whether the ncAA shall generally be incorporated in a protein or rather the host cell or if a site-specific incorporation is desired.

The first method, namely selective pressure incorporation (SPI) aims at a residue-specific incorporation of non-canonical amino acids in an auxotrophic strain. Using this technology, one canonical amino acid is completely depleted from the growth medium and replaced by a ncAA. Since the endogenous aminoacyl tRNA synthetase and tRNA, as well as the translation apparatus are tolerant to a certain extent, the non-canonical amino acid will be incorporated instead of the canonical amino acid. Advantages of this method are that no gene modification is necessary and high levels of replacement of the natural amino acid can be achieved. However, depending on the purpose of amino acid exchange, the latter might at the same time be a disadvantage, since the replacement of the amino acids takes place at all sites encoding the canonical amino acid and cannot be limited to certain exchanges. Furthermore, a similarity in the size and structure of the amino acids to be interchanged must be given, since otherwise the ncAA is not recognized by the endogenous aminoacyl tRNA synthetase (aaRS) and tRNA (44–47).

The second method, which was used in this work, aims at a site-specific incorporation of non-canonical amino acids and is called stop codon suppression (SCS). As the name suggests, the function of a stop codon is reprogrammed to a sense codon. Since it is least frequently used to terminate translation in *E. coli*, usually the amber stop codon TAG serves as suppressed codon (48). For this technique, an exogenous, specifically evolved tRNA and the associated aminoacyl tRNA synthetase, as well as a suitable non-canonical amino acid must be available in the producing cell. When an amber stop codon is recognized on the mRNA in the translation process, regularly, translation would be terminated at this point by the binding of a release factor (RF1), which leads to the dissociation of the polypeptide from the ribosomal complex. However, when a suitable tRNA is available that recognizes the amber stop base triplet UAG on the mRNA, the tRNA can bind to the mRNA in the ribosomal complex (Figure 3). The charged ncAA is incorporated in the nascent polypeptide and translation proceeds until one of the other two stop codons are recognized (ochre stop codon: TAA or opal stop codon: TGA).

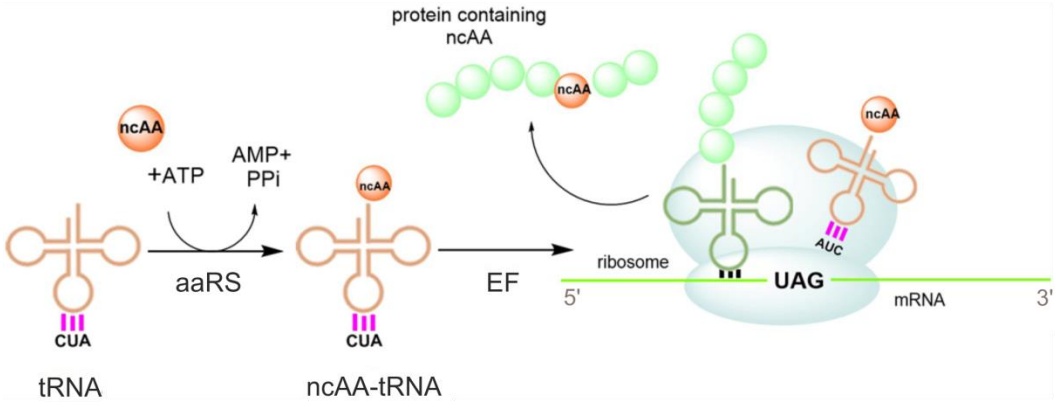

Figure 3: Stop codon suppression and incorporation of ncAA during translation
An exogenous tRNA with the anticodon AUC (3' → 5') is charged with the non-canonical amino acid by the orthogonal aminoacyl tRNA synthetase (aaRS). When the stop codon UAG is recognized on the mRNA translation is either terminated, or as shown here in case of successful stop codon suppression, the exogenous aminoacyl tRNA, which is brought to the ribosome by the elongation factor (EF), binds the stop codon in the ribosome and the ncAA is introduced to the nascent protein. Figure adapted from Crnković et al. (49).

It is required that the exogenous aaRS specifically only interacts with the non-canonical amino acid and only aminoacylates the exogenous tRNA in order to prevent canonical amino acids to be incorporated at this site. Vice versa, the ncAA is not supposed to be incorporated by endogenous aaRS/tRNA pairs. This property of non-promiscuity is referred to as orthogonality to the endogenous system (50–52).

For introduction of more than one non-canonical amino acid, quadruplet codon reassignment was developed from stop codon suppression. This method requires the usage of more than one orthogonal suppressor pair (53).

For this work, the SCS method was chosen, since it allows site specific incorporation of the desired amino acid. Replacement of the codon which encodes the amino acid that is supposed to be exchanged with the ncAA against the amber stop codon TAG is necessary. The method is limited in its efficiency due to competition with naturally occurring processes, such as translation termination in the case of stop codon reassignment. Furthermore, it has been shown that suppression efficiency highly depends on the surrounding sequence context (54). In order to find suitable positions of nisin which allow the incorporation of a non-canonical amino acid, an amber suppression screening at each amino acid site of nisin was performed. However, the identification of promising sites for the incorporation of amino acids does not at the same time guarantee the production of a fully functional protein, which is able to undergo the PTM machinery and displays its antibiotic properties. This is further to be investigated.

1.5 Objective

This work is part of the SYNPEPTIDE project, which was funded by the *Seventh Framework Programme* of the European Union and brought together five academic groups and three industrial partners from four different European countries. Generally speaking, the project aimed at designing novel peptide antibiotics and to introduce new features and properties to the lantibiotics, as well as at the establishment of a method for screening the novel peptides with regard to their functionality (55). Diversification of lantibiotics was targeted by various approaches, such as modular shuffling of existing lantibiotics (56), addition of posttranslational modifications or incorporation of ncAA to existing lantibiotics (47). This work was part of the latter subdivision.

In previous work in the group of Prof. Wagner, an in-frame scanned nisin A TAG library was generated and successful test runs for relative quantification of nisin wild-type expression were performed. For quantification of protein expression, a nisin/GFP fusion protein was generated and the fluorescence intensity was used as measure for protein quantification (Bachelor's thesis Maximilian Fischer, 57).

Aim of this work was the identification of positions in the lantibiotic nisin A, which allow the incorporation of the non-canonical amino acid Boc-K via amber stop codon suppression in *E. coli* and quantification of the suppression efficiency using a fluorescence assay (Figure 4).

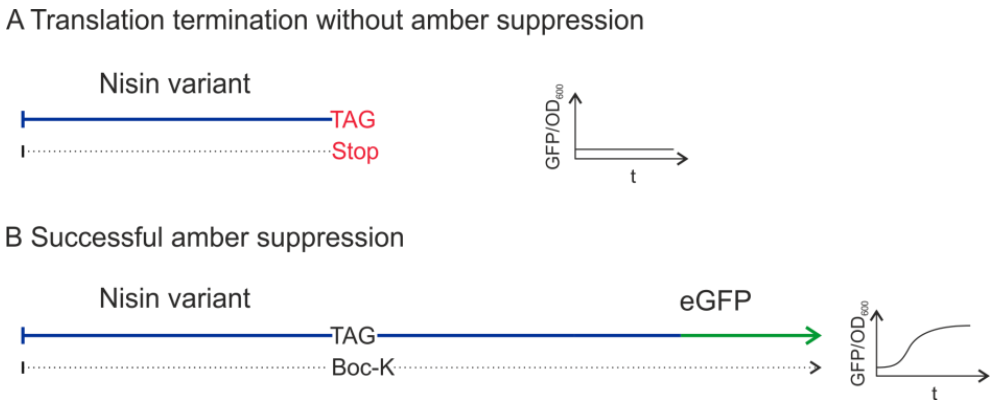

Figure 4: Fluorescence intensity as measure for amber suppression
(A) In absence of the non-canonical amino acid Boc-K, the translation terminates at the stop codon TAG of the nisin variant. A C-terminally truncated protein results, which does not contain GFP and therefore cannot be detected in the fluorescence assay. (B) In case of successful amber suppression, Boc-K is incorporated instead of translation termination at the stop codon TAG. The whole protein including eGFP can be translated and the fluorescence intensity can be used as measure for protein expression.

This work focused on optimization of the fluorescence assay and assessing its reliability as well as improvement of a plasmid used for incorporation of transgenes in a first step. Second, a screening of the suppression efficiency of every nisin A mutant with the incorporated non-canonical amino acid was performed and production of nisin as well as its variants was indirectly measured via analysis of the fluorescence intensity of GFP.

In order to test the reliability of the used system, quality controls were performed and possibly critical component parts were checked independently from the whole assay. Since the results did not confirm expectations, the suppression assay was critically scrutinized concerning possible sources of error.

2. Material

2.1 Antibodies

Antibody	Usage	Supplier	Cat. No.
Anti GFP	Western Blot	Santa Cruz (Dallas, USA)	sc-8334
Anti-His_6-Peroxidase	Western Blot	Roche (Basel, Switzerland)	11 965 085 001
Goat-Anti-Rabbit	Western Blot	Thermo Fischer Scientific (Waltham, USA)	31460

Table 1: Antibodies

2.2 Bacterial strains and media

Cell Strains	Genotype
E. coli BL21(DE3)	*huA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*
E. coli NEB Turbo	F' *proA+B+ lacIq ΔlacZM15 / fhuA2 Δ(lac-proAB) glnV galK16 galE15 R(zgb-210::Tn10)*TetS *endA1 thi-1 Δ(hsdS-mcrB)5*
E. coli T7 Express I^q	MiniF *lacI*q(CamR) */ fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10--*TetS*)2 [dcm] R(zgb-210::Tn10--*TetS*) endA1 Δ(mcrC-mrr)114::IS10*

Table 2: Bacterial strains

Bacterial media	Composition
LB medium	1% Bacto tryptone; 0.5% Bacto yeast extract; 1% NaCl; pH 7.5
LB_{Amp} medium	LB medium; 100 µg/ml ampicillin
$LB_{Amp+Chlor}$ medium	LB medium; 100 µg/ml ampicillin; 30 µg/ml chloramphenicol
LB_{Chlor} medium	LB medium; 30 µg/ml chloramphenicol

Table 3: Bacterial media

2.3 Buffers and solutions

Buffer/Solution	Composition
Coomassie staining solution	0.2 % (w/v) Coomassie Blue R-250; 50 % (v/v) methanol; 10 % (v/v) acetic acid
Laemmli buffer	312.5 mM Tris; 5 % (w/v) SDS; 25 % (v/v) ß-Mercaptoethanol; 25 % (v/v) Glycerol; 2.5 mM EDTA; pH 6.8
PBS	140 mM NaCl; 2.7 mM KCl; 10 mM Na_2HPO_4;1.8 mM KH_2PO_4
TBE buffer	90 mM Tris; 0.55 % (w/v) H_3BO_3; 2 mM EDTA
WB transfer buffer	25 mM Tris; 1.13 % (w/v) Glycine; 10 % MeOH
TBS	150 mM NaCl; 50 mM Tris; pH 7.4
TTBS	150 mM NaCl; 50 mM Tris; 0,05 % Tween-20; pH 7.4
ECL 1	2.5 mM Luminol; 0,4 mM Coumaric acid; 100 mM Tris HCL; pH 8.5
ECL 2	100 mM Tris HCL pH 8.5; 0,00183 % (v/v) H_2O_2

Table 4: Buffers and solutions

2.4 Chemicals and consumables

The non-canonical amino acids were purchased from Sigma-Aldrich (St. Louis, USA) (Boc-K, Cat. No. 359661) and Activate Scientific (Prien, Germany) (pBF, Cat. No. AS43582), respectively. Part of the used Boc-K was kindly provided by the group of Prof. Budisa, TU Berlin. 5 ml HisTrap columns used for chromatographic protein purification were obtained from GE Healthcare (Chalfont St Giles, Great Britain) (Cat. No. 17-5248).

All other chemicals and consumables were purchased from Merck (Darmstadt, Germany), NEB (Ipswich, USA), Roth (Karlsruhe, Germany), Sarstedt (Nuembrecht,Germany), Sigma-Aldrich (Steinheim, Germany) or Thermo Scientific (Waltham, USA), if not specified otherwise in the respective sections.

2.5 Commercial Kits

Name	Supplier	Cat. No.
GeneJet Plasmid Miniprep Kit	Thermo Fischer Scientific (Waltham, USA)	K0503
QIAquick Gel Extraction Kit	Qiagen (Hilden, Germany)	28706
Quick Ligation™ Kit	NEB (Ipswich, USA)	M2200S

Table 5: Commercial Kits

2.6 Enzymes

Enzyme	Supplier	Cat. No.
Phusion HF DNA polymerase	NEB (Ipswich, USA)	M0530
CIP	NEB (Ipswich, USA)	M0290
Nde I restriction enzyme	NEB (Ipswich, USA)	R0111
Not I HF restriction enzyme	NEB (Ipswich, USA)	R3189

Table 6: Enzymes

2.7 Instruments

Name	Manufacturer
ÄKTAbasic P-900	Amersham/GE Healthcare (Chalfont St Giles, Great Britain)
BlueFlash semi-dry blotting device	SERVA Electrophoresis (Heidelberg, Germany)
Branson Sonifier	Branson (Danbury, USA)
ChemiLux Pro ECL imager	Intas (Goettingen, Germany)
GeneAmp PCR System 2400	Perkin Elmer (Waltham, USA)
Infinite 200 PRO microplate reader	Tecan (Maennedorf, Switzerland)
NanoDrop ND-1000	Peqlab (Erlangen, Germany)

Table 7: Instruments

2.8 Plasmids and Oligonucleotides

Plasmids containing the orthogonal suppressor pairs pTB77 and pTB290, that were used for incorporation of the ncAA Boc-K were obtained from Dr. Tobias Baumann, group of Prof.

Budisa, TU Berlin. Further information on these plasmids is provided in Table 8 below. Plasmid pEVOL-pBpF was kindly provided by Peter G. Schultz, Scripps Research Institute.

Plasmid	Specification	
pTB77	pJZ MmpylT Strep-MmpylS(Y384F)	aaRS not codon-optimized N-terminal Strep-tag
pTB290	pJZ MbpylTS(Y349F)	aaRS codon-optimized no Strep-tag

Table 8: Orthogonal suppressor pair plasmids pTB77 and pTB290

Figure 5 displays a vector map of the pET21a derivative, which encodes for the Nis A/GFP fusion protein.

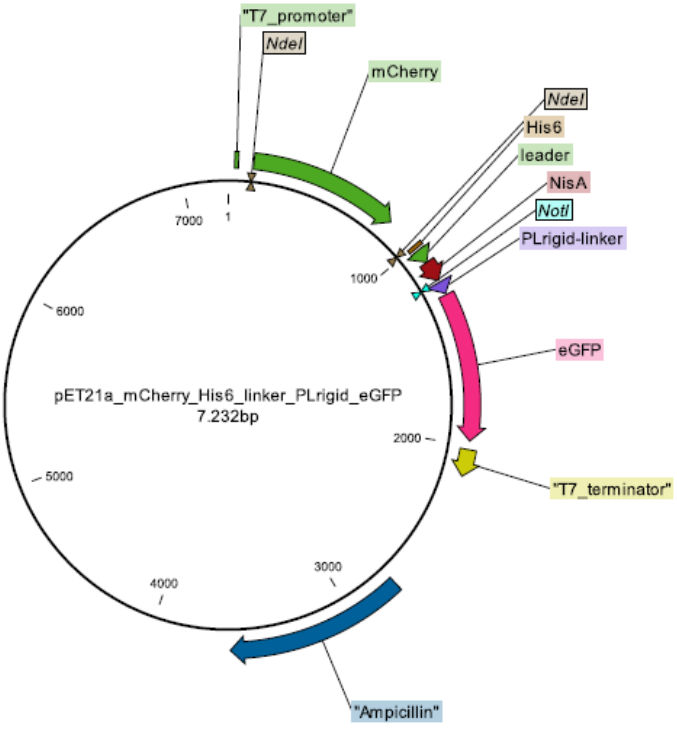

Figure 5: Plasmid pET21a derivative
This vector map contains the ampicillin resistance gene as well as the restriction sites that were relevant for this work. It also includes the mCherry sequence which was later cut out.

A list of oligonucleotides used as forward and reverse primers for amplification of nisin wild-type and Nis A variants is given below (Table 9).

Primer	Use	Nucleic acid sequence
#050_DP	Forward primer for all Nis A variants	TAATACGACTCACTATAGGG
nisA-lib-rev-wt	Reverse primer for Nis A variants Nis A 1 to Nis A 27	GGCAGCGCGGCCGCCTTTGCTTACGTGAATACTACA
TAG_34_rev_wo_TAA	Reverse primer for Nis A 28	GGCAGCGCGGCCGCCTTTGCTTACGTGAATACT**CTA**
TAG_33_rev_wo_TAA	Reverse primer for Nis A 29	GGCAGCGCGGCCGCCTTTGCTTACGTGAAT**CTA**ACA
TAG_32_rev_wo_TAA	Reverse primer for Nis A 30	GGCAGCGCGGCCGCCTTTGCTTACGTG**CTA**ACTACA
TAG_31_rev_wo_TAA	Reverse primer for Nis A 31	GGCAGCGCGGCCGCCTTTGCTTAC**CTA**AATACTACA
TAG_30_rev_wo_TAA	Reverse primer for Nis A 32	GGCAGCGCGGCCGCCTTTGCT**CTA**GTGAATACTACA
TAG_29_rev_wo_TAA	Reverse primer for Nis A 33	GGCAGCGCGGCCGCCTTT**CTA**TACGTGAATACTACA
TAG_28_rev_wo_TAA	Reverse primer for Nis A 33	GGCAGCGCGGCCGCC**CTA**GCTTACGTGAATACTACA

Table 9: Primers

2.9 Software

Software	Version
Magellan	Magellan Standard
CLC Main Workbench	7.6.1
GraphPad PRISM	8.3.1
UNICORN	5.31

Table 10: Software

3. Methods

3.1 Microbial techniques

3.1.1 Growth and storage of bacteria

Bacteria were cultivated overnight in 5 ml Lysogeny broth (LB) medium at 37 °C and shaken at 220 rpm, if not specified otherwise. If necessary, antibiotics were added to the medium (ampicillin 100 µg/ml, chloramphenicol 30 µg/ml or a combination of both).

Bacterial cultures which were used repeatedly for inoculation of new cultures were stored as glycerol stocks (50% bacterial culture in LB medium, 50% Glycerol) at -80 °C.

3.1.2 Transformation of chemically competent *E. coli*

Chemically competent cells were prepared as previously described (58) and stored at -80 °C. For transformation, 200 µl of cell suspension were thawed on ice. After addition of 50 ng plasmid DNA or 10-20 µl ligation mix, cells were stirred gently and incubated on ice for 5 minutes. Cells were heat shocked for 45 seconds at 42 °C and incubated on ice for another 5 minutes. Subsequently, 800 µl LB medium was added and cells were shaken at 300 rpm and 37 °C for 1 h. Cells were plated on LB-agar plates containing ampicillin and/or chloramphenicol and kept at 37 °C overnight.

3.2 Molecular biology

3.2.1 Plasmid purification

DNA was purified using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) following manufacturer's instructions and eluted with 30 µl ddH_2O. DNA concentration and purity were determined photometrically using a NanoDrop 1000 spectrophotometer (Peqlab).

3.2.2 Polymerase chain reaction

DNA fragments were amplified by polymerase chain reaction (PCR). Phusion High-fidelity DNA polymerase (NEB) was used. Annealing temperatures for primers were calculated using the NEB Tm calculator (www.tmcalculator.neb.com). For details on PCR sample preparation and PCR program see Table 11 and Table 12 below.

Component	20 µl reaction
Template DNA	Variable
Forward Primer	1 µl
Reverse Primer	1 µl
dNTP mix (10 mM each)	1 µl
5 x Phusion HF buffer	4 µl
ddH_2O	To 19 µl
Phusion HF Polymerase	1 µl

Table 11: PCR sample preparation

Step	Temperature	Time	No. of cycles
Initial Denaturation	94 °C	2 min	
Denaturation	94 °C	30 s	
Annealing	55° C	30 s	25
Elongation	72 °C	30 s / kb	
Final Extension	72 °C	5 min	
Hold	4 °C		

Table 12: Standard PCR program

3.2.3 Agarose gel electrophoresis

For separation of DNA fragments according to their size, agarose gel electrophoresis was performed. Depending on the size of expected DNA fragments, TBE buffer containing 1% or 2% agarose (w/v) was boiled. Next, 50 ng/ml ethidium bromide was added. When hardened, the gel was overlaid with TBE buffer. DNA samples mixed with 6 x purple gel loading dye (NEB) were loaded onto the gel, as well as a DNA ladder to estimate a given DNA fragment's size (100 bp DNA ladder, NEB or O'GeneRuler 1 kb DNA Ladder, Thermo Scientific). Electrophoresis was performed at 90 V for 30 to 45 min. DNA bands were visualized by UV light (302 nm for documentation purposes or 365 nm when the gel fragment was to be excised, respectively).

For extraction from the gel, desired DNA bands were excised and purified using the QIAquick Gel Extraction Kit (Qiagen). DNA was eluted with 30 µl ddH_2O and stored at -20 °C.

3.2.4 Restriction digestion

Restriction digestions for analytical and preparative purposes were all carried out using NEB restriction enzymes according to the manufacturer's instructions. For details on the enzymes used, see section 2.6.

3.2.5 Ligation

For ligation of linearized plasmids, 1 µl Quick Ligase (2000 units) per 50 ng vector was used. For ligation of inserts into vectors, 50 ng vector DNA was mixed with a threefold molar excess of insert DNA prior to addition of 1 µl Quick Ligase. Preparations were incubated for 30 min at room temperature. All ligations were performed using the Quick Ligation Kit (NEB), for details on the preparations see Table 13.

Component	**Ligation of linearized plasmids (20 µl)**	**Ligation of vector and insert (20 µl)**
Quick Ligase buffer (2 x)	10 µl	10 µl
Vector DNA	50 ng	50 ng
Insert DNA	-	Depending on insert size
ddH_2O	to 19 µl	to 19 µl
Quick Ligase	1 µl	1 µl

Table 13: Preparations for ligations

3.2.6 Cloning

For cloning, molecular biology methods were performed as described in the previous sections. Vector and insert were each digested with suitable restriction enzymes. Subsequently, the vector was treated with calf intestinal alkaline phosphatase (CIP, NEB) to prevent religation. For this purpose, 2 µl CIP (20 units) were added to 50 µl digestion mix and incubated for 30 min at 37 °C. Following, vector and insert were separated from cropped DNA fragments by performing agarose gel electrophoresis. After extraction from the gel and ligation of vector and insert, chemically competent *E. coli* (Turbo competent *E. coli*, NEB) were transformed with the ligated DNA. Following, the plasmid was purified from the expression cells. The final cloning products were subjected to Sanger sequencing for confirmation that the generated DNA corresponded to the desired sequence (performed by GeneArt or SeqLab).

3.3 Protein biochemistry

3.3.1 Relative protein quantification by fluorescence measurement

Nisin expression was relatively quantified through fluorescence measurement of the linked reporter proteins eGFP (enhanced green fluorescent protein) or RFP (red fluorescent protein) using an Infinite 200 Pro microplate reader (Tecan) and the associated Magellan Software.

The excitation wavelengths used were 480 nm (eGFP) and 580 nm (RFP) (excitation bandwidth: 9 nm) and the emission wavelengths were 510 nm and 610 nm (emission bandwidth 20 nm) for eGFP and RFP fluorescence signals, respectively.

Cultures which were to be measured were grown overnight, then inoculated to an OD_{600} of 0.1 in 6 ml LB medium containing suitable antibiotics and incubated at 37 °C for 1h 30 min. During this time, bacterial cultures reached an OD_{600} of 0.6-0.7, which was always controlled by re-measurement of the OD_{600}. Cultures with deviating densities were discarded.

The fluorescence assay was carried out in a 24-well format, with each sample being measured in technical triplicates.

Each well was inoculated with bacterial culture, LB medium containing antibiotics, arabinose, IPTG and the ncAA when indicated, leading to final concentrations as specified in Table 14. The total volume was 1.5 ml per well.

Component		**Final concentration**
Bacterial culture		OD_{600}=0.25
Arabinose		0.02% w/w
IPTG		1 mM
Ampicillin		100 µg/ml
Chloramphenicol		30 µg/ml
ncAA	**Boc-K**	1 mM
	pBF	1 mM

Table 14: Sample preparation for plate reader measurements

Subsequently, the plate reader measurement was started, providing stable conditions for the cultivation of bacteria at a temperature of 37 °C and shaking at 218 rpm. Fluorescence intensity and OD_{600} were monitored for 8.4 h with measurement points every 12.9 min.

3.3.2 Protein purification

Overnight cultures, which had been grown at 37 °C in 5 ml LB medium with a suitable antibiotic, were inoculated to an OD_{600} of 0.1. After two hours of incubation, when an OD_{600} of 0.6 to 0.7 was reached, the samples were induced with IPTG and arabinose to a final concentration of 1 mM IPTG and 0.02% w/w arabinose. When necessary, Boc-K was simultaneously added to the samples to a final concentration of 1 mM. Cells were incubated for 5 hours at 37 °C and shaken at 220 rpm. Samples were centrifuged at 4000 rpm for 15 minutes at 4 °C and supernatant was discarded. The pellets were resuspended in 20 ml PBS and centrifuged again at 4000 rpm for 15 minutes. The resulting pellets were stored at -20° C, if not used immediately. For further processing, pellets were resuspended in 10 ml binding buffer. Cell disruption was achieved by high energy ultrasound application for 15 minutes using a Branson Sonifier. After centrifugation at 14 000 rpm at 4 °C for 20 minutes, the supernatant containing the proteins was sterile filtrated (pore size 0.22 µm). The following steps were performed using the Äkta high-performance liquid chromatography system (GE Healthcare) and a 5 ml HisTrap column (GE Healthcare). After the column had been loaded with the sample at a flow rate of 2 ml/min, it was washed with eight column volumes (40 ml) running buffer to remove non-bound proteins. Following, the protein was eluted from the column using elution buffer containing 500 mM imidazole at a flow rate of 3 ml/min. Elution fractions of 2 ml were collected. Fractions of interest (fraction 6 and 14) were subjected to SDS-PAGE and Western blot or Coomassie staining.

3.3.3 SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed to separate proteins according to their size. For sample preparation, 50 µl of each sample were mixed with 12.5 µl of 5 x Laemmli buffer and incubated at 95 °C for 5 min. Depending on the further use of the gels, 10 µl or 5 µl per sample for Coomassie Staining or Western Blots, respectively, were loaded onto a 17% SDS-gel, as well as 6 µl PageRuler Plus Prestained Protein Ladder (Thermo Scientific) for later size determination. Electrophoresis was started with an initial voltage of 40 V, until samples had passed the stacking gel. The residual electrophoresis was performed at 70-100 V for 1 h.

3.3.4 Coomassie staining

After electrophoresis, proteins were visualized by incubating SDS-gels with Coomassie Staining Solution for 5 minutes. Subsequently, the gel was washed with ddH_2O until protein bands were sufficiently visible.

3.3.5 Western blot

Semi-dry electroblotting was performed for the transfer of proteins from SDS-gels onto 0.45 µm pore size nitrocellulose membranes. The latter were incubated in WB transfer buffer for 15 min. The transfer stack was set up as following: three Whatman papers the size of the nitrocellulose membrane were soaked in WB transfer buffer and placed on the anode of the blotting device, followed by the nitrocellulose membrane. The SDS-gel was placed on top of the membrane and overlaid by another three soaked Whatman papers on the cathode side. Blotting was performed at 1.5 mA/cm² membrane for 1 h.

Subsequently, non-specific binding of antibodies was prevented by incubation of the membrane in 50 ml TBS containing 5% non-fat dry milk overnight at 4 °C. After three washing steps (5 min each) with TTBS (TBS containing 0.2% Tween), the membrane was overlaid with antibody-solution.

In case of the HRP-conjugated anti-6xHis-antibody (Roche), which was applied in a 1:1000 dilution in 1% non-fat dry milk TBS solution, 2 h incubation at room temperature was sufficient. The membrane was washed three times for 5 min in TTBS.

For GFP detection, the membrane was overlaid with anti-GFP-antibody (Santa Cruz), which was diluted 1:4000 in 1% non-fat dry milk TBS. After an incubation time of 1 h at room temperature and three washing steps of 5 min with TTBS, the secondary goat-anti-rabbit-antibody (Thermo Fischer Scientific) was applied in a 1:2000 solution in 1% non-fat dry milk TBS. The membrane was incubated for 1 h at room temperature and washed three times as described above.

Finally, the membrane was exposed to ECL-substrate solution, consisting of a 1:1 mixture of ECL-1 and ECL-2 reagent for 5 minutes and signals were visualized using a ChemiLux Pro Imager (Intas).

4. Results

4.1 Optimization of the fluorescence assay

One of the aims of this work was the further development and improvement of the fluorescence assay for relative protein quantification, based on previous work in the group of Prof. Wagner (57), which had already borne successful test runs for relative quantification of nisin wild-type expression.

Nisin variants should be quantified using a two plasmid system. The first plasmid (Figure 6), a derivative of a pET21a vector, includes the nisin sequence, which is fused C-terminally with enhanced green fluorescent protein (eGFP, also referred to as GFP in the following) through a linker. Furthermore, it contains an mCherry sequence, which encodes red fluorescent protein (RFP), a 6x-His-Tag, which is situated at the N-terminus of nisin, as well as an ampicillin resistance gene.

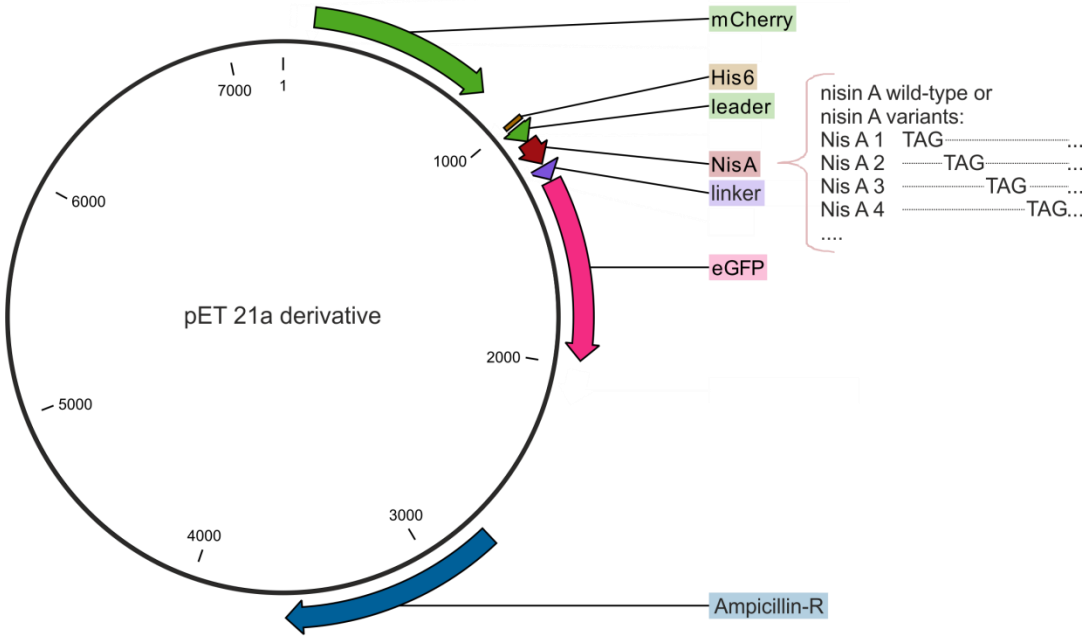

Figure 6: Vector map of the pET21a derivative containing nisin A
The vector map provides an overview of the components which were relevant for this work and partly subjected to modifications.

This vector was subjected to modifications which are described in detail in the following sections. A set of 34 nisin variants was then cloned into the final vector for the amber suppression screening. Each nisin variant differs in one codon from the wild-type, replacing

the naturally occurring base triplet with the amber stop codon TAG. The nisin mutants used were also a product of above-mentioned work in the group of Prof. Wagner (57).

For measurements of the nisin amber suppression screening, all cells were transformed with a second plasmid, encoding an orthogonal suppressor pair and a chloramphenicol resistance gene. The suppressor pair, which is necessary for suppression of the amber stop codon TAG, consists of a tRNA and an aminoacyl tRNA synthetase. Two different sets of suppressor pairs were used in this work, pTB77 and pTB290 (both were kindly provided by Dr. Tobias Baumann, group of Prof. Budisa, Technical University Berlin). While pTB77 (=pJZ MmpylT Strep-MmpylS(Y384F)) contains an N-terminal Step-tag and encodes a non-codon optimized aminoacyl tRNA synthetase, the second pair, pTB290 (=pJZ MbpylTS(Y349F) is not Strep-tagged and its aminoacyl-tRNA-synthetase is codon-optimized for *E. coli* (see section 2.8).

Before starting the nisin amber suppression screening, preliminary experiments aiming for optimization of the vector and assay set-up were performed.

4.1.1 Choice of expression cells

Two different *E. coli* expression cell strains were compared regarding growth kinetics and GFP expression in order to determine which one is better suited for high protein yields. *E. coli* BL21(DE3) and *E. coli* T7 Express I^q cells are both engineered for high-level protein expression using the T7 expression system, with the latter strain being an enhanced derivative of BL21 with reduced expression leakiness of T7-promoter-controlled genes, when they are not induced.

Both cell strains were transformed with the pET21a vector containing the nisin wildtype construct and four different linkers between the protein of interest and the GFP gene (see section 4.1.2 below), respectively, and subjected to the fluorescence assay for relative protein quantification. Each construct was measured in biological quadruplicates, as well as each sample in technical triplicates. *E. coli* BL21(DE3) cells displayed a slightly faster and higher overall growth, but both expression cell strains grew steadily and consistently from an OD_{600} of 0.25 when inoculated to an OD_{600} of approximately 0.9 to 1.1 at the measurement end-point after 8.4 h (Figure 7 A). However, a difference in protein expression kinetics of the cell strains was observed. *E. coli* T7 Express I^q cells started GFP expression with a slight temporal delay of approximately 20 minutes compared to *E. coli* BL21(DE3) cells. Apart from this, the log phase in the first three hours of the experiment was comparable in terms of GFP fluorescence intensity. Following, *E. coli* BL21(DE3) constructs entered a stationary phase of

GFP fluorescence intensity values, while they gently kept rising in *E. coli* T7 Express I^q cells, leading to higher overall protein expression (Figure 7 B).

By tendency, these effects occurred independently from the linker sequence used (data not shown).

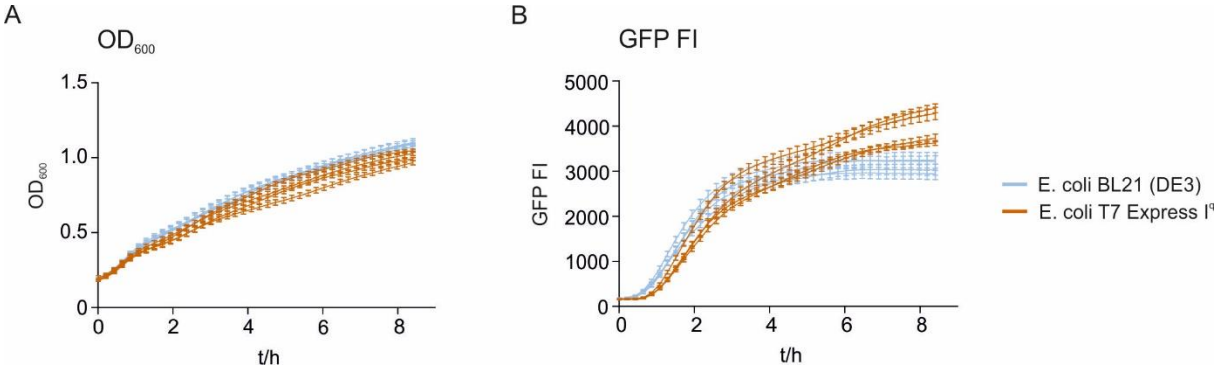

Figure 7: Comparison of two different expression cell strains for the fluorescence assay
E. coli BL21(DE3) and *E. coli* T7 Express I^q cells were transformed with four variants of nisin wild-type plasmid, differing in a linker sequence between nisin and the reporter protein eGFP. Four clones of each variant in the two cell strains were picked and subjected to the fluorescence assay. Cell growth kinetics were monitored by measurement of the OD_{600} (A) and GFP fluorescence intensity signals served as marker for protein expression (B). *E. coli* BL21(DE3) measurements are displayed in light blue, whereas *E. coli* T7 Express I^q cells measurements are shown in orange. In both graphs, error bars indicate the SEM of the biological replicates for each sample.

Since the fluorescence intensity and therefore GFP expression was higher in *E. coli* T7 Express I^q cells despite slightly reduced overall growth compared to *E. coli* BL21(DE3) cells and GFP fluorescence intensity values continued rising after entering the stationary phase of cell growth, *E. coli* T7 Express I^q cells were used for all following experiments, if not specified otherwise.

4.1.2 Insertion of different linkers between nisin A and GFP

In order to prevent unwanted interactions between nisin A and the reporter protein eGFP, a linker was inserted between the two proteins, which should disturb expression as little as possible. For the choice of a suitable linker, four different ones had previously been cloned into the nisin wild-type construct (57), differing in length, physiochemical properties and the secondary structure they form. Nucleotide and amino acid sequence of the different linkers are shown in Figure 8.

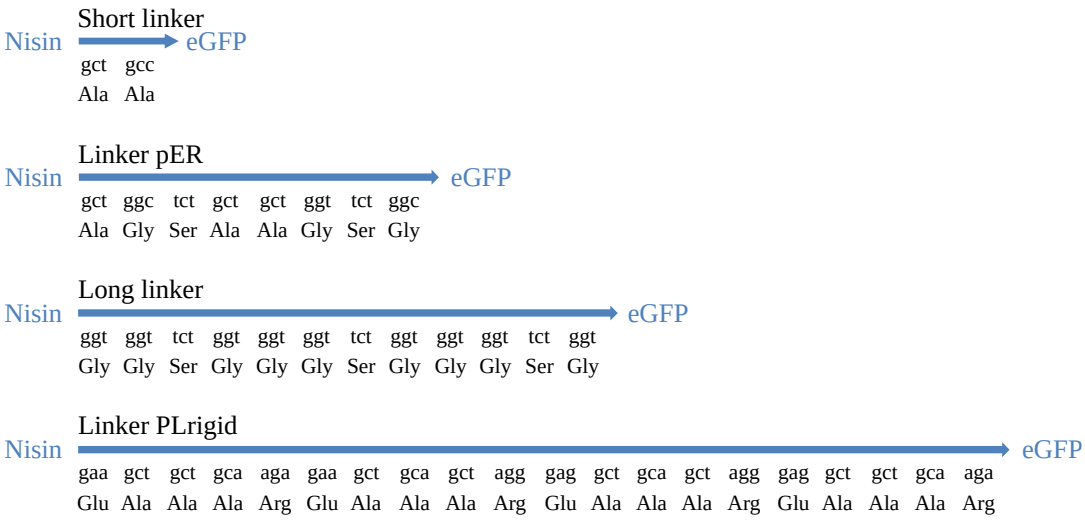

Figure 8: Nucleotide and amino acid sequence of the tested linkers
Displayed linkers were cloned into nisin wild-type constructs in between nisin and the eGFP sequence. The arrows point from 5' to 3' of the nucleotide sequence or N- to C-terminus of the amino acid sequence, respectively. Figure adapted from Bachelor's thesis Maximilian Fischer (57).

The linker consisting of two alanines (short linker) is particularly short and flexible and might therefore have only minimal influence on protein translation. Linker pER is present in the pER13a vector the constructs had originally been developed in. The long linker is characterized by a repetitive sequence of glycines and serines, which are both small and hydrophilic amino acids and thus expected to have little influence on the linked proteins. While this linker does not build any secondary structure, the forth linker which was used, linker PLrigid, is characterized by the rigid α-helix it forms (59). Separating nisin and GFP by a long linker that might minimize possible sterical hindrances between the two proteins, which could otherwise affect protein translation and folding.

In order to find out which linker is best suited for the planned screening, the fluorescence assay was performed in technical triplicates and in biological quadruplicates for each linker-construct. None of the different linker constructs disturbed growth of the *E. coli* T7 Express I^q expression cells remarkably (Figure 9 A). Expression kinetics were also rather similar, meaning GFP expression started after approximately 40 min, followed by a steep log phase for about 2 h 20 min and an ongoing milder increase of the GFP fluorescence intensity values in the stationary phase. However, overall GFP expression did vary depending on which linker was used. Constructs containing the linker PLrigid displayed the steepest increase of fluorescence intensities in the log phase, while all other linker constructs did not differ from each other significantly. In the stationary phase, the least additional GFP expression was observed when constructs contained the short or the long linker, whereas when linker pER

was used, GFP expression values gained on to the level of linker PLrigid constructs after 6 hours (Figure 9 B).

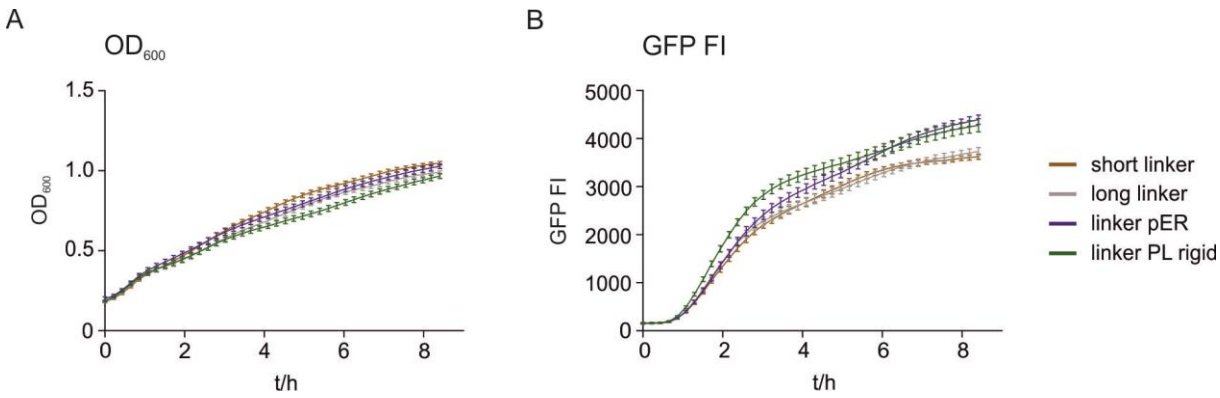

Figure 9: Comparison of four different linkers between nisin and eGFP
E. coli T7 Express I^q cells were transformed with four nisin wild-type variants differing in the linkers between nisin and the reporter protein eGFP. Four clones of each linker construct were picked and used to perform the fluorescence assay for examination of cell growth monitored by the OD_{600} (A) and GFP fluorescence intensity signals indicating protein expression (B) with each sample being measured in technical triplicates. In both graphs, error bars indicate the SEM of the biological quadruplicates for each sample.

For all further experiments, constructs containing linker PLrigid were used, since they provided high overall protein expression and allowed for a good distinction between log and stationary phase.

4.1.3 Normalization of protein expression

The assay used in this work measures fluorescence intensities of the reporter protein eGFP to identify positions within nisin which will tolerate the incorporation of a non-canonical amino acid. Since protein expression strongly depends on bacterial growth, fluorescence values were normalized to the respective OD_{600} values for each measurement point in order to correct the obtained signal for the number of bacterial cells. Furthermore, it should also be tested in this work if a second reporter protein encoded on the same plasmid as nisin and eGFP is suitable as an additional normalization parameter. The nisin wild-type constructs, which had been generated in previous work, contain an mCherry sequence, which encodes an RFP (red fluorescent protein) as a second reporter protein besides eGFP. Transcription of both genes is controlled by an IPTG-inducible T7 promoter. While eGFP is directly linked to the C-terminus of nisin, the RFP gene is separate from nisin (see vector map, Figure 6).

To test if RFP provides a suitable normalization signal, *E. coli* T7 Express I^q bacteria were transformed with nisin wild-type plasmids containing eGFP and RFP or eGFP only. 12 bacterial clones of each type were used for inoculation of liquid cultures and subjected to fluorescence measurement in technical triplicates over a time of 8.4 hours after induction of nisin expression (Figure 10). One culture transformed with the plasmid in which the mCherry sequence was removed was discarded due to insufficient growth (data not shown).

The presence of RFP did not appear to have a major influence on bacterial growth since the measured OD_{600} values were similar irrespective of additional RFP expression (Figure 10 A). However, GFP fluorescence intensity was lower at any time when the mCherry gene was transcribed from the same plasmid (Figure 10 B). It was also observed that the expression of RFP was delayed compared to eGFP. This might lead to skewed normalization values as is indicated in Figure 10 C. Finally, during exponential growth (t= 2 hours), it was observed that for a given clone, the overall deviation of normalized eGFP values from the mean value of all clones was higher when normalizing the signal to RFP than to the OD_{600} (Figure 10 D). Thus, it was concluded that the presence of RFP did not provide any additional benefit in terms of normalization.

Since the overall eGFP fluorescence was lower when RFP was co-expressed, the latter was removed from the plasmid and the screening was performed with eGFP normalization to the OD_{600} only. It was also observed in this experiment that variances in protein expression were significantly lower between different bacterial clones when normalizing eGFP to OD_{600} (Figure 10 C and D). This made it acceptable to pick a single clone for each nisin variant from which glycerol stocks were then prepared. These glycerol stocks were used to inoculate liquid cultures during amber suppression screening in which each variant was screened three times on different days and in technical triplicates. To account for possible variances between different assay runs, wild-type nisin, to which obtained variant values were normalized, was always measured as well.

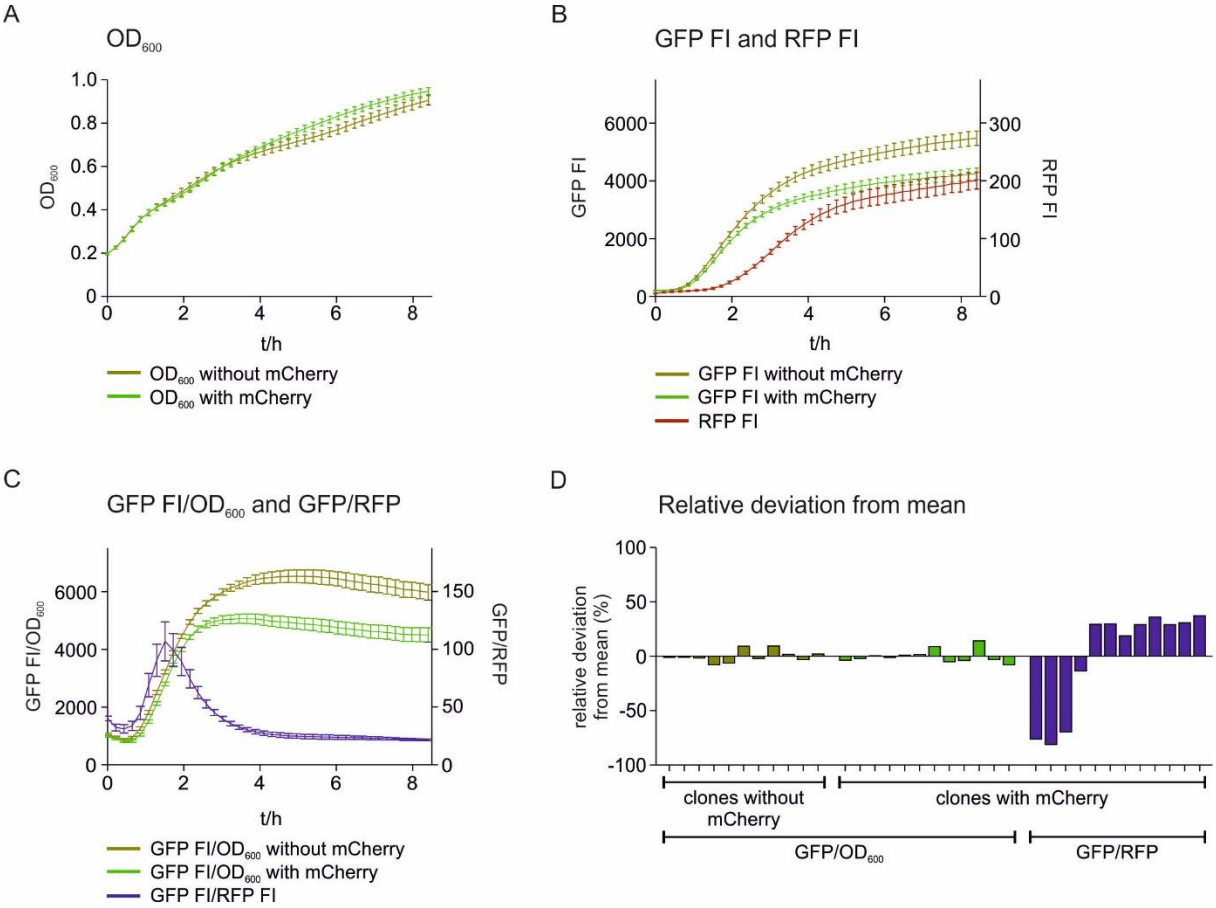

Figure 10: Evaluation of RFP as second reporter protein
E. coli T7 Express I^q cells were either transformed with nisin wild-type plasmids encoding RFP (mCherry) and GFP as reporter proteins or GFP only. 11 clones without RFP and 12 clones containing RFP were subjected to the fluorescence assay in technical triplicates. (A) The OD_{600} was used to monitor cell growth. (B) Fluorescence intensities of clones containing the GFP sequence only (dark green), as well as the GFP and RFP sequences (light green and red) indicate protein expression. GFP fluorescence intensity signals are plotted on the left y-axis, while the RFP fluorescence intensity signal is plotted on the right y-axis. (C) GFP FI signal normalized to OD_{600} for constructs containing the mCherry sequence are shown in light green, whereas the GFP FI values normalized to the OD_{600} for constructs without the RFP gene are displayed in dark green. The alternative GFP FI to RFP FI normalized graph is shown in purple. GFP FI/OD_{600} normalization values are plotted on the left y-axis, while GFP FI/RFP FI values are plotted on the right y-axis. (A-C) Error bars indicate the SEM of the biological replicates. (D) Normalized fluorescence intensity values in the log phase (t = 2 hours) of each clone were compared to the mean of the respective group (dark green: GFP FI/OD_{600} for constructs without mCherry, light green: GFP FI/OD_{600} for constructs which contain mCherry, purple: GFP FI/RFP FI for constructs that contain mCherry). Relative deviation from the average is displayed in percent for the individual clones.

4.1.4 Timing of suppressor pair induction and ncAA addition

Arabinose and IPTG, as well as a suitable ncAA were added manually to the bacterial cultures right before the start of a plate reader measurement for the fluorescence assay. Arabinose functioned as inducer for the suppressor pair since both of its genes are controlled by a

tryptophan promoter, while IPTG was used to induce nisin and therefore eGFP expression, which is controlled by a T7 promoter.

To assess if the timing of suppressor pair induction or ncAA addition has an influence on overall protein expression or expression kinetics, several chronologically different schemes were tested. For this purpose, the fluorescence assay was performed with nisin wild-type as well as with two variants: Nis A 5 and Nis A 29 (i.e. the fifth and the twenty-ninth codon were substituted by the stop codon TAG). Since two different suppressor pairs were available and to be tested in the nisin amber suppression screening, all experiments on induction times were carried out with both of them, each in combination with both tested variants and the wild-type.

Following temporal schemes were tested: Boc-K and arabinose were each added individually to sample preparations half an hour (t= -0.5 h), one hour (t= -1 h) or one and a half hours (t= -1.5 h) before the start of the plate reader run, respectively (Figure 11, schemes 3 to 8). Furthermore, the ncAA and arabinose were added simultaneously to the liquid cultures, one and a half hours (t= -1.5 h) or right before (t= 0 h) starting the fluorescence assay (Figure 11, schemes 1 and 2). The time period of 1.5 hours which was analyzed corresponds to the time in which cells reach the log phase again after being re-inoculated after overnight growth.

Two main points should be analyzed with this experiment. First, it should be assessed if induction of the suppressor pair genes prior to induction of nisin benefits the production of GFP. For the incorporation of a ncAA into nisin, the suppressor pair is a necessary factor, since it provides the specific tRNA and tRNA-synthetase. With the addition of arabinose prior to the addition of IPTG, both suppressor pair components are already present when nisin translation is initiated. Second, the ncAA (in this case Boc-K) was added manually to the preparations and needed to be taken up by the cells before further processing, e.g. integration into nisin variants, was possible. By the addition of Boc-K prior to the induction of nisin, the possibility should be excluded that delayed up-take of the ncAA decelerated the production of nisin.

Both factors were expected to potentially affect nisin variant synthesis. However, neither the induction of the suppressor pair, nor the addition of Boc-K should directly influence the expression of nisin wild-type. Still, the production of the suppressor pair and the presence of Boc-K could have general effects on the expressing cell, such as influences on cell growth or on overall protein production. For this reason, wild-type nisin was also included in the tests.

Bacterial growth and reporter protein expression were analyzed for 4.5 hours, at which time cells had reached the stationary phase. Figure 11 shows the normalized results of the

fluorescence assay (GFP/OD_{600}). When induction and addition of Boc-K took place simultaneously at t=0 h, a higher GFP fluorescence value per OD_{600} was reached than for all other tested schemes. The second highest level of protein expression was detectable when arabinose and the ncAA were added simultaneously at t= -1.5 h, even though in this case, the increase of fluorescence signal per cell mostly applied to the nisin wild-type. All other schemes tested produced results which resembled each other concerning GFP expression per OD_{600}, with an overall lover protein expression.

Since maximization of nisin production efficiency presumably increases the sensitivity of the amber suppression screening, the scheme resulting in the highest fluorescence values was once more chosen. Thus, all further experiments were carried out with the simultaneous induction of both inducible systems, as well as the addition of the ncAA right before the plate reader measurement start (scheme 1: t = 0 h).

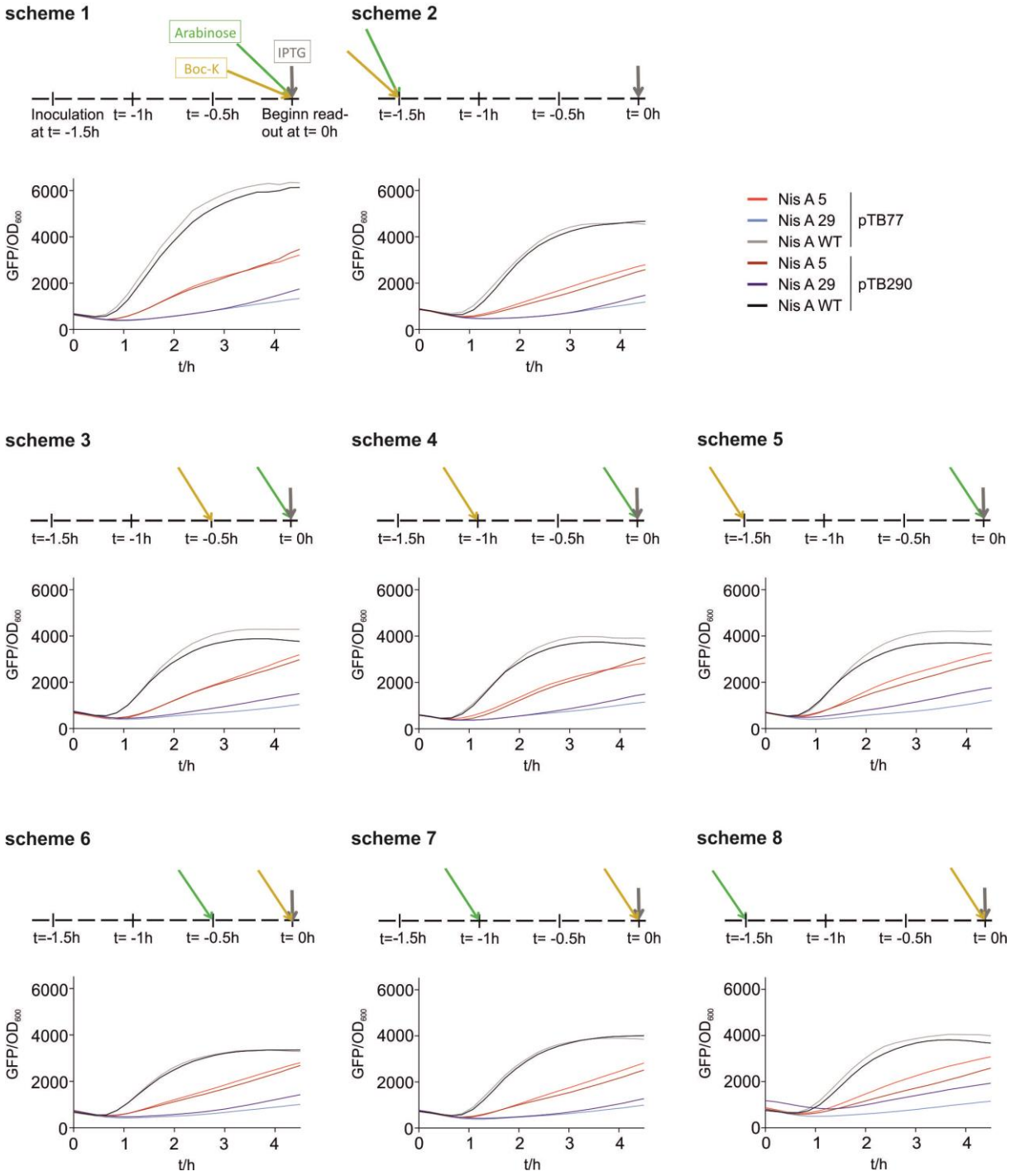

Figure 11: Timing of suppressor pair induction and ncAA addition
The timeline above the graphs indicates at which time the respective agents were added to the culture before the start of the plate reader measurements. Yellow arrows indicate the addition of the ncAA Boc-K and green arrows the addition of arabinose and induction of suppressir pair expression. For better orientation, induction of nisin expression is shown with a grey arrow even though the addition time of IPTG was not altered. The fluorescence assay was performed with the nisin wild-type construct and the variants Nis A 5 and Nis A 29, each one tested with the suppressor pairs pTB77 and pTB290, respectively. Normalized protein expression is displayed in black/grey (Nis A wild-type), red (Nis A 5) and blue (Nis A 29), with the lighter color for each sample representing measurements with pTB77 and the darker one pTB290. Scheme 8 displayed a growth irregularity of Nis A 29/pTB290,

which is the reason for the skewed dark blue line. GFP FI was normalized to the OD_{600}. Each sample was measured in technical triplicates, with the plotted line corresponding to the mean values.

4.2 Nisin amber suppression screening

After all preliminary experiments concerning vector design and assay settings were completed, the screening of the nisin amber library was performed under these optimized conditions.

For better comparability of the different samples, all conditions were kept constant, except for the nisin variants the cells were transformed with. This included that cells expressing wild-type nisin were also transformed with the respective suppressor pair. Since only 8 samples could be measured at a time, the obtained values were normalized to those of wild-type nisin, which was always measured on the same plate during an assay run. For this reason, fluorescence intensity values that were first normalized to the OD_{600} are presented as percentage of the wild-type, which was set to 100 %. The normalization to the respective wild-type should ensure that differences in between nisin variants could be attributed to the variants themselves, and lower the influence of possible variability in between different plate reader runs. Samples were tested in biological triplicates in different plate reader runs, with each replicate being tested in technical triplicates during one plate reader measurement.

For a better visualization of differences in protein expression, three representative time-points were chosen to display the results: one hour (Figure 12 A) , two hours (Figure 12 B) and six hours (Figure 12 C).

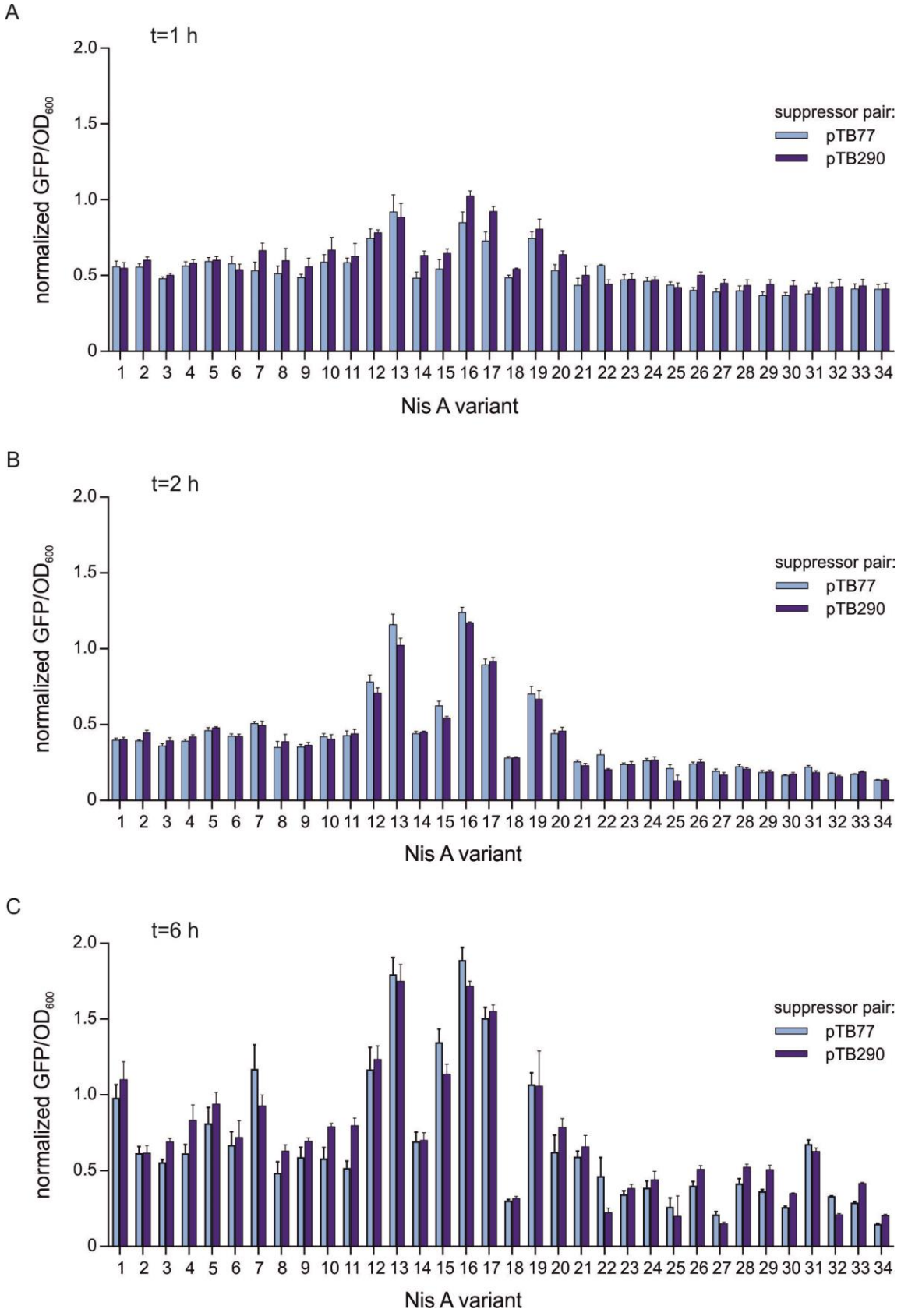

Figure 12: Amber suppression screening for all Nis A variants
The fluorescence assay was carried out in biological triplicates with each variant of Nis A. For better comparability, each plate reader run contained a nisin wild-type sample, which was also transformed

with the suppressor pair to be tested in the same run. GFP FI/OD_{600} values of the variants were normalized to nisin wild-type GFP FI/OD_{600}, which was set to 1. Values greater than 1 indicate higher levels of protein expression than the wild-type. Diagrams display the respective normalized Nis A variant GFP FI/OD_{600} values after 1 hour (A), 2 hours (B) or 6 hours (C). Error bars indicate the SEM of biological replicates. Parts of the experiments were performed by Malin Zaddach.

One hour values were chosen, since they displayed differences in the early phase of the log-phase of protein expression (Figure 12 A). At this time point already, several observations were made. First, all variants produced eGFP at a significant level of at least 36 % of nisin wild-type expression. Second, while on average variants expressed 53 % (pTB77) or 58 % (pTB290) of nisin wild-type expression, several mutants stood out with higher values (positions 12, 13, 16, 17 and 19). Notable nisin variants are displayed in Table 15 below. Third, a split-up in a first part of nisin from positions 1 to 17 and a second part of nisin (Nis A 18-34) was evident. On average, mutants in the front part expressed slightly more than 50 % of Nis A wild-type GFP expression, while the back part variants stayed a little below this value, with the exception of the above named outstanding variants. Finally, the obtained values for a given position were comparable for both suppressor pairs, indicating that differences in fluorescence intensity were mainly due to positional effects and less dependent on the suppressor pair used.

The two hour time point for data analysis was chosen, since it represented the log phase of bacterial growth and GFP expression (Figure 12 B). At this point in the experiment, similar effects to those after one hour were observed. The same variants as named above displayed remarkably high GFP fluorescence intensity values (see Table 15) and the difference between the front part and the back part of nisin was even more distinct than one hour after protein induction. Again, none of the suppressor pairs was superior to the other one regarding the expression of the reporter protein. Furthermore, the results showed that after two hours, the average percentage of nisin variant GFP expression in relation to wild-type nisin GFP expression was lower than after one hour. This allowed the assumption that the increase of nisin wild-type eGFP expression was steeper than the increase of nisin variant GFP expression in this phase.

After 6 hours of measurement time, cells had reached their maximum GFP values. Some variants displayed a slight increase of GFP fluorescence signals even after this period, but in general, the stationary phase had been reached at this time point. Observations which were made previously after one and two hours, also applied to the normalized GFP expression after 6 hours (Figure 12 C). However, differences among the particular variants were more distinct. Furthermore, the difference between the normalized fluorescence values of the front and back

part of nisin was bigger than in the first two time points. While the average ratio of normalized GFP expression in the first part of the protein was 96 % of nisin wild-type values, it was significantly lower in the back part from position 18 on (43 % of Nis A wild-type expression). Outstanding variants at this point were Nis A 1 and Nis A 7, which expressed GFP just as well as the wild-type, Nis A 12, Nis A 15 and Nis A 19, which expressed approximately 20 % more GFP than the wild-type and variants Nis A 13, Nis A 16 and Nis A 17, which produced at least 50 % more of the reporter protein compared to the wild-type. For an overview of normalized GFP expression relative to wild-type expression of highly expressing variants Nis A 12, Nis A 13, Nis A 16, Nis A 17 and Nis A 19 after 1, 2 and 6 hours see Table 15.

Nisin variant	suppressor pair	1h	2h	6h
Nis A 12	pTB77	75%	78%	117%
	pTB290	78%	71%	124%
Nis A 13	pTB77	92%	116%	180%
	pTB290	89%	102%	175%
Nis A 16	pTB77	85%	124%	189%
	pTB290	103%	117%	172%
Nis A 17	pTB77	73%	90%	151%
	pTB290	92%	92%	155%
Nis A 19	pTB77	75%	70%	107%
	pTB290	81%	67%	106%

60% GFP expression relative to wild-type 200%

Table 15: Highly expressing nisin variants normalized to wild-type expression
Nisin wild-type GFP expression normalized to the OD_{600} was set to 100 %. Data was obtained from the amber suppression screening. The displayed variants were chosen due to high expression levels throughout the monitored time, even though these are not the only ones reaching values comparable to wild-type level (see Figure 12).

4.3 Quality control of the suppression system

In the amber suppression system, a specifically designed aminoacyl-tRNA-synthetase is used to link a non-canonical amino acid to an artificial tRNA molecule. While this tRNA/amino acid complex is present, the non-canonical amino acid can be specifically integrated during

translation. In its absence, translation stops at the amber codon TAG, since it naturally functions as a stop codon.

The nisin constructs used for the amber suppression screening contain the reporter protein eGFP at the C-terminus of nisin. When protein expression is terminated at any position of nisin, eGFP should consequently not be translated either, which would lead to non-detectable fluorescence signals in the assay.

To test if this is the case here, eight different nisin variants were chosen and subjected to fluorescence measurement without the addition of Boc-K to the bacterial cultures. Criteria for the choice of the eight samples were that variants from the whole length of nisin should be represented, and variants that displayed particularly high or low eGFP expression values in the amber suppression screening should be included. As a control, the same eight variants were tested in the same plate reader run but with Boc-K being present during the fluorescence assay. All samples were tested in technical triplicates. The experiment was carried out with both suppressor pairs that were used in the amber suppression screening.

Surprisingly, the expectations of no or only slight background fluorescence in the absence of Boc-K were not met at any time in the experiment, with neither of the suppressor pairs and none of the different variants (Figure 13).

For comparability, the same time points as in the nisin amber suppression screening are displayed. After one hour, differences in fluorescence intensity per OD_{600} values in the presence or absence of Boc-K were marginal (Figure 13 A). The same effect was visible two hours after induction (Figure 13 B). It should be mentioned that the trend of higher protein expression in the front part of nisin compared to the back part, which was already evident in the nisin amber suppression screening (see section 4.2, Figure 12), was observed here once more.

After six hours of protein expression, normalized GFP values were higher when Boc-K was added to the preparations for almost every tested variant irrespective of the used suppressor pair, even though the differences did not reach the same extent for all samples (Figure 13 C).

Finally, it was noteworthy that protein expression values did not increase in between the two hour and the six hour measurement for some variants when Boc-K was not added. This was the case for the variant Nis A 22 when suppressor pair pTB77 was used and for variants Nis A 29 and Nis A 32 independently of the used suppressor pair. For all other nisin mutants, fluorescence per OD_{600} values were higher after six hours than after two hours, irrespective of whether the non-canonical amino acid was added or not.

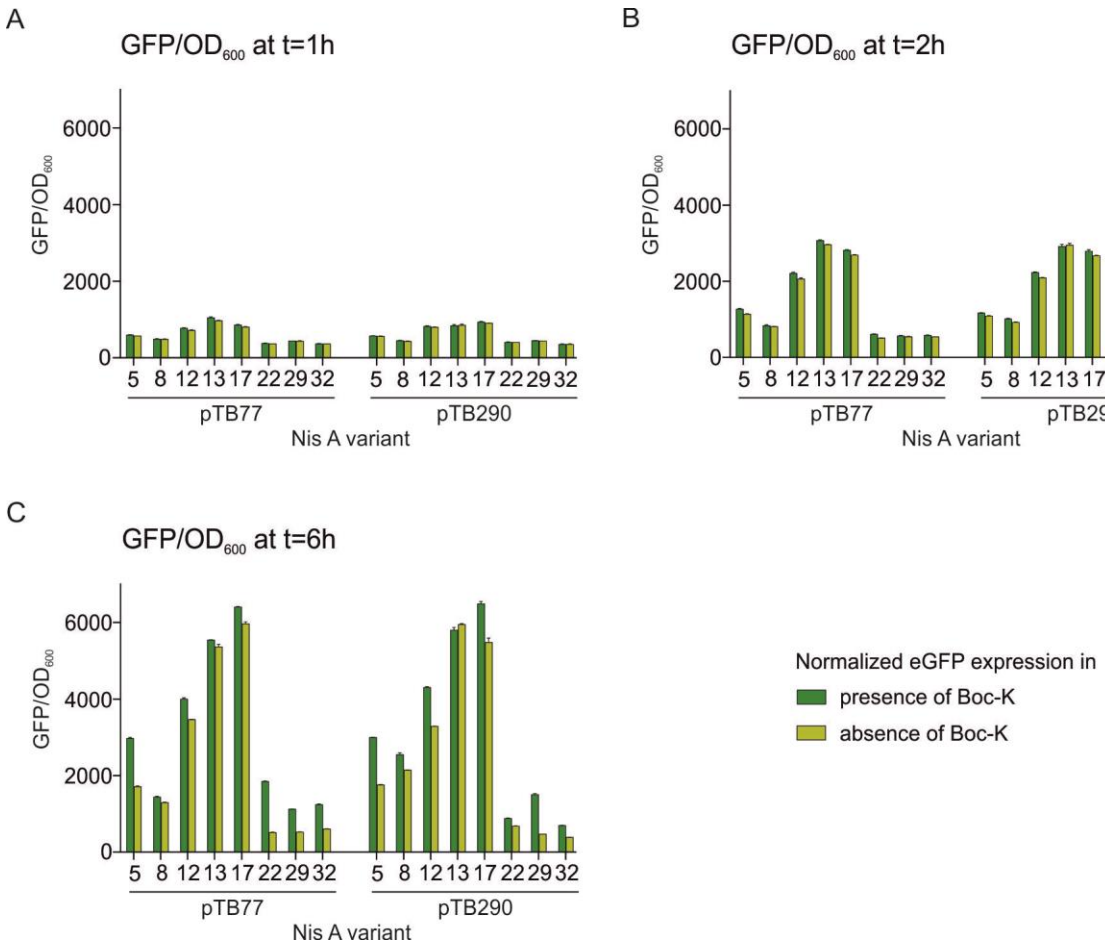

Figure 13: Protein expression in presence or absence of Boc-K
The fluorescence assay was performed with eight different Nis A variants for testing protein expression in absence of the ncAA Boc-K (light green). As a reference, fluorescence intensity when Boc-K is available was also determined in the same plate reader runs (dark green). GFP FI was normalized to the OD_{600}. For better comparability, the same time points were chosen for data analysis as in the amber suppression screening analysis (1 hour (A), 2 hours (B) and 6 hours (C)). Error bars indicate the SEM of the technical triplicates.

The variably high GFP/OD_{600} values for the different variants indicated that the observed fluorescence signals are not background signals from the cells or the constructs themselves. In order to estimate the actual background fluorescence, GFP values were determined when the nisin/GFP plasmid was not present (Figure 14). *E. coli* T7 Express I^q cells, which only carried the suppressor pair pTB77 plasmid, but not the nisin/GFP plasmid were subjected to the fluorescence assay. Furthermore, it was assessed if the amber codon TAG really functioned as a stop codon during translation in the absence of the synthetic tRNA/amino acid complex. For this purpose, the variants Nis A 5 and Nis A 29 were tested without prior transformation of the expression cells with the suppressor pair. As a positive control, a nisin wild-type sample

including all other components required for amber suppression was tested in the same plate reader run.

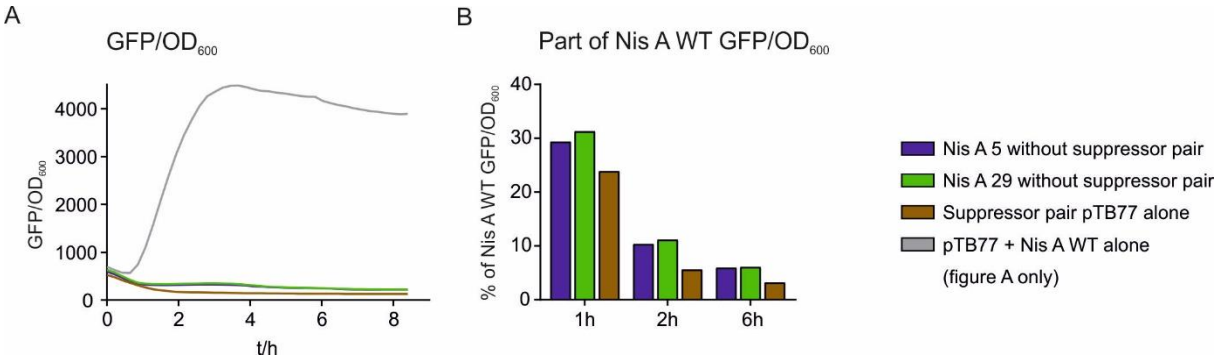

Figure 14: Background signals in the fluorescence assay
The variants Nis A 5 (blue) and Nis A 29 (green) were subjected to the fluorescence assay without the prior transformation of a suppressor pair to the expression cells. The brown graph/bar indicates the fluorescence signals which were obtained when no nisin/GFP plasmid, but only the suppressor pair plasmid was present, representing background signals. Expression cells including the nisin wild-type plasmid and pTB77 served as a positive control. (A) GFP FI normalized to the OD_{600} of the different samples. (B) In this figure, the obtained normalized signal is shown as a part of the positive control for each other sample after one, two and six hours of protein expression. The legend on the right side of the figure applies to both graphs.

Data for both these experiments showed that only low fluorescence signals derived from the cells themselves or other components in the assay set-up (Figure 14 A). When the nisin plasmid was not present, only marginal amounts of fluorescence were detected. Also, the data clearly showed that the suppressor pair was necessary for suppression and only a very small amount of GFP was formed due to skipping of the amber stop codon. This was represented by the difference between the graph of the suppressor pair alone and the variants that were tested without the suppressor pair (Figure 14 A).

Further data analysis (Figure 14 B) showed which part of the fluorescence signal must be ascribed to background signals at the relevant time-points. Since the values for the tested variants Nis A 5 and Nis A 29 did not differ significantly from each other, it could be assumed that background fluorescence for other variants is similarly low, too.

Further steps were taken to examine where the fluorescence signals in absence of Boc-K derived from. For this purpose, the following components of the nisin amber suppression screen were assessed separately from each other regarding functionality: the bacterial expression strain, the plasmid which encodes the nisin/GFP fusion protein and the helper plasmid which carries the gene for the orthogonal tRNA and the tRNA-synthetase.

4.3.1 Exchange of the expression strain to *E. coli* BL21(DE3) cells

One of the first experiments in this work during optimization of the fluorescence assay set-up was the choice of expression cells from two different cell strains (see section 4.1.1). *E. coli* T7 Express I^q cells were chosen for the following experiments, since they displayed higher GFP expression values for the nisin wild-type construct. Yet, *E. coli* BL21(DE3) cells also provided GFP expression levels which would have allowed their use as expression cells in the fluorescence assay. In order to evaluate if the choice of expression cells had an influence on GFP expression in presence or absence of Boc-K, *E. coli* BL21(DE3) cells were transformed with the same eight nisin variants that were tested in the quality control experiment from section 4.3 (Figure 13) and then subjected to the fluorescence assay.

The most distinct difference between GFP values of constructs in *E. coli* T7 Express I^q cells and those in *E. coli* BL21(DE3) cells was, that when expressed in the former, relevantly higher normalized fluorescence intensity values resulted (Figure 15, compared to Figure 13). Surprisingly, differences in GFP values between the two cell strains were higher in this experiment compared to the nisin wild-type measurements which were carried out in the preliminary experiments for expression cell determination (section 4.1.1). However, this result was considered to be valid, since the 16 measurements of the different positions were performed on four different days with comparably low GFP expression in all cases.

In general, overall GFP expression levels were lower when *E. coli* BL21(DE3) cells were used, irrespective of whether the ncAA was added or not. After one and two hours (Figure 15 A and B), fluorescence intensity signals were comparable for all tested positions with variant Nis A 32 in combination with the suppressor pair pTB290 being the only exception. After 2 hours, the variant Nis A 32 produced about 50% less GFP than all other variants. Finally, after six hours (Figure 15 C), more distinct differences in the amount of protein expression could be observed in between the different variants.

Surprisingly, the pattern of well expressing variants and nisin mutants that express less GFP did not coincide with the pattern shown when *E. coli* T7 Express I^q cells were used. While the variants Nis A 13 and Nis A 17 produced remarkably high amounts of GFP in the latter expression cells, GFP values were below average in *E. coli* BL21(DE3) cells.

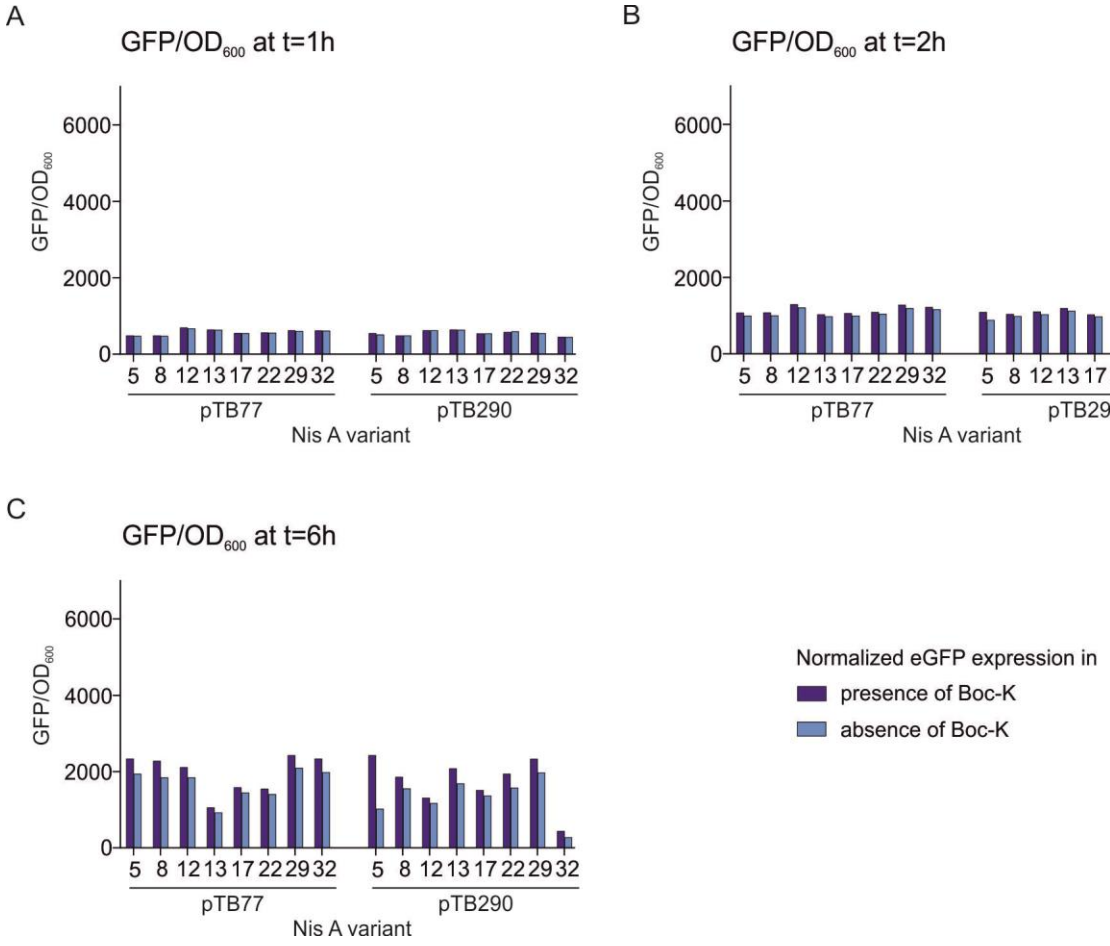

Figure 15: Protein expression in presence or absence of Boc-K in *E. coli* BL21(DE3) cells
The fluorescence assay was performed with the same eight Nis A variants which were previously used to determine protein expression in presence or absence of the ncAA (Figure 13). For this experiment, the expression cells *E. coli* BL21(DE3) were transformed with one of the eight exemplarily used Nis A variants, as well as the suppressor pair plasmid pTB77 or pTB290, respectively. Light blue bars show fluorescence intensity values when Boc-K was not added. As a reference, fluorescence intensity when Boc-K is available was also determined in the same plate reader runs (dark blue). GFP FI was normalized to the OD_{600}. For better comparability, the same time points as in the previous experiments were chosen for data analysis (1 hour (A), 2 hours (B) and 6 hours (C)).

However, when examining the influence of the presence or absence of the non-canonical amino acid, a similar trend was observed as for the expression in *E. coli* T7 Express I^q cells. Normalized GFP values were lower when Boc-K was not added than when it was present in the experiments, but once more, the differences were only marginal.

As a summary of this experiment, it can be stated that the observed phenomenon of high fluorescence values, irrespective of the presence or absence of Boc-K, cannot be explained by the choice of the expression cell strain.

4.3.2 Substitution of the orthogonal suppressor pair and the associated ncAA

A possible explanation for the high fluorescence values in the absence of Boc-K could be that the artificial synthetases that were used did not specifically link only the ncAA to the respective tRNA, but also other naturally occurring amino acids that were present in the culture medium. This hypothesis should be examined by the usage of a third suppressor pair (pEVOL-pBpF). pEVOL-pBpF (pEVOL) is an orthogonal tRNA/tRNA-synthetase pair which was developed by Peter G. Schultz (60) for the incorporation of the non-canonical amino acid p-benzoylphenylalanine (pBF) into proteins using the amber suppression system. In the manufacturing process of a suppressor pair, the synthetase undergoes several steps of positive and negative selection to ensure its specificity.

E. coli BL21(DE3) bacteria were transformed with the suppressor pair pEVOL and the above tested eight nisin variants. The samples were subjected to the fluorescence assay under the same conditions as described in section 4.3. Each variant was measured twice, one time with pBF being added to the preparations before the plate reader measurement and one time in absence of the ncAA. All samples were measured in technical triplicates. The results of this experiment (Figure 16) corresponded very well to the observations made in the first experiment on protein expression in presence or absence of the ncAA used (section 4.3). The overall normalized eGFP fluorescence values in this experiment did not reach the ones in the first experiment. However, despite the usage of *E. coli* BL21(DE3) cells, overall GFP values were rather high which stood in contrast to the previous experiment (section 4.3.1, Figure 15), but corresponded well to the initial tests, in which *E. coli* T7 Express I^q cells were chosen for the screening (section 4.1.1, Figure 7). In these experiments, it was already shown that overall fluorescence values of *E. coli* BL21(DE3) cells were lower compared to *E. coli* T7 Express I^q cells.

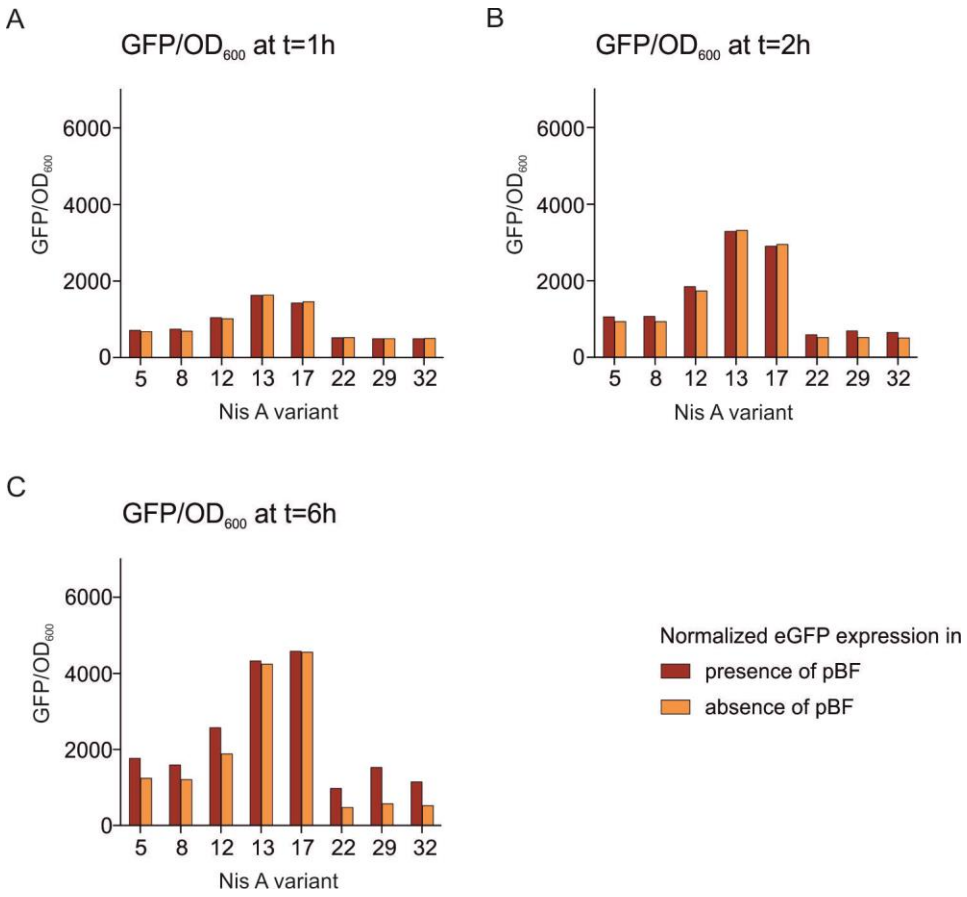

Figure 16: Protein expression using suppressor pair pEVOL in presence or absence of the ncAA
E. coli BL21(DE3) cells were transformed with one out of eight different Nis A variants and the suppressor pair pEVOL, respectively. The constructs were subjected to the fluorescence assay, once with the corresponding amino acid pBF available (red bars) and once in absence of the ncAA (orange bars). GFP FI was normalized to the OD_{600}. For better comparability, the same time points were chosen for data analysis as in the amber suppression screening analysis (1 hour (A), 2 hours (B) and 6 hours (C)).

When comparing the fluorescence signals in presence or absence of the ncAA (Figure 16), pEVOL provided very similar data as the two suppressor pairs pTB77 and pTB290, which were used in the amber suppression screening. For comparability, the same time points for data analysis were chosen as for previous experiments. After one hour (Figure 16 A), differences between samples with or without the addition of pBF were marginal. Furthermore, due to the earlier beginning of the log phase in *E. coli* BL21(DE3) cells (see section 4.1.1, Figure 7), it already became apparent at this point, that the same positions that displayed higher fluorescence values when *E. coli* T7 Express I^q cells and the suppressor pairs pTB77 and pTB290 were used, do also lead to good GFP expression with the suppressor pair pEVOL. Two hours after induction (Figure 16 B), the same tendencies that could be observed after one hour were visible. The highest increase of normalized GFP values could be observed

for the nisin variants Nis A 12, Nis A 13 and Nis A 17, just like when the other suppressor pairs were used. At this point, six out of the eight tested variants displayed higher fluorescence intensity values when pBF was added. Once more however, the differences were lower than expected. After six hours (Figure 16 C), the discrepancies between samples which contained the ncAA and those which did not were more distinct in the same six variants that displayed differences after two hours already. Namely these variants were Nis A 5, Nis A 8, Nis A 12, Nis A 22, Nis A 29 and Nis A 32. Interestingly, for the two variants which provided the highest fluorescence values (Nis A 13 and Nis A 17), no significant difference could be observed in the fluorescence values as a result of ncAA addition. Finally, it was remarkable that for the variants Nis A 22, Nis A 29 and Nis A 32, the fluorescence values in absence of pBF did not or only slightly increase in between the 2 hour and the 6 hour measurement point, while there was a gain of fluorescence intensities when the ncAA was added.

Overall, very similar observations were made in this experiment with a different suppressor pair and its corresponding non-canonical amino acid, when compared to the original experiment on protein expression in presence or absence of Boc-K in *E. coli* T7 Express I^q cells with the suppressor pairs pTB77 and pTB290 (see section 4.3, Figure 13). For this reason, it may be assumed that the usage of the latter suppressor pairs is not decisive for the high eGFP values when Boc-K is not added to the preparations.

4.3.3 Purification and further analysis of expressed proteins

Since to this point it remained unclear, what caused the high eGFP values in absence of the respective ncAA in the fluorescence assay, some protein variants were purified and subjected to further investigations. All constructs that were used throughout this work contained an N-terminal 6xHis-tag, which makes purification via Ni-Ion chromatography possible. Six variants spanning the entire protein (Nis A 2, Nis A 5, Nis A 13, Nis A 17, Nis A 22 and Nis A 29) were chosen and fractions from the purification were analyzed by Western blot and Coomassie staining of SDS gels. As controls, the nisin wild-type and *E. coli* T7 Express I^q cells without any plasmids added were included. For all other samples, including the nisin wild-type, pTB77 was used as suppressor pair. Since it had become evident in previous experiments that the respective suppressor pair was not responsible for the high fluorescence signal in the absence of the ncAA (section 4.3.2), this experiment was carried out with only one suppressor pair. All constructs were expressed in *E. coli* T7 Express I^q cells. Each variant was subjected to the purification and the above mentioned tests in presence as well as in absence of Boc-K.

Samples were prepared as described in section 3.3.2 and analyzed using the ÄKTA protein purification system. Elution profiles of nisin wild-type, untransformed *E. coli* T7 Express I^q cells and the exemplary picked samples Nis A 5 and Nis A 29 with and without the addition of Boc-K to the preparations are shown below (Figure 17).

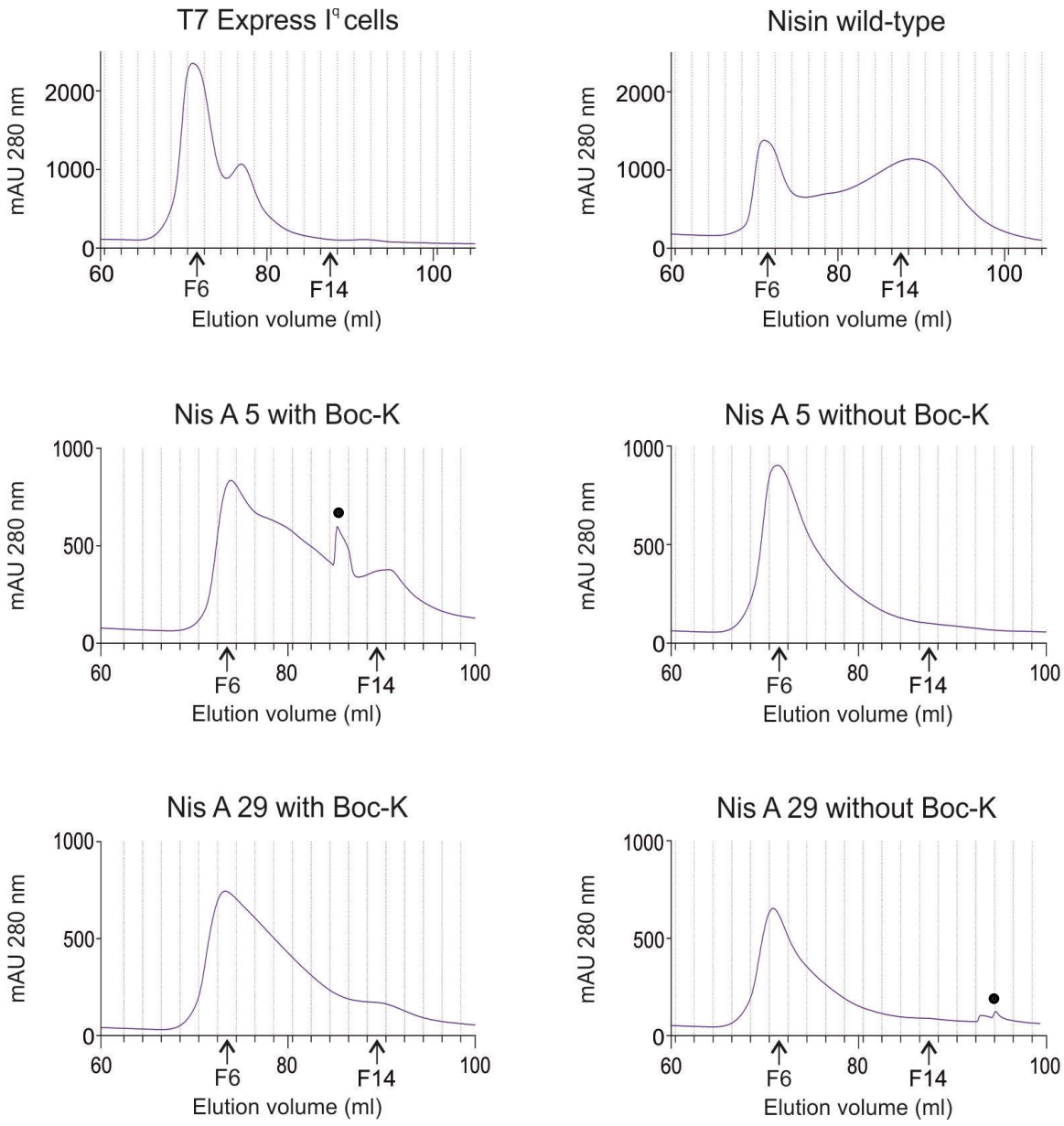

Figure 17: Elution profiles of Nis A constructs in presence or absence of Boc-K
E. coli T7 Express I^q cells were transformed with the suppressor pair pTB77 as well as the indicated Nisin variants or left entirely untransformed. Total cell lysates were subjected to Ni-Ion chromatography and eluted using an imidazole gradient. X-axis of the graphs indicate the elution volume (ml), y-axis display absorbance units as an indirect measure of relative protein amount. Blue lines indicate the absorption of the respectively measured variant at 280 nm. Fractions of further interest (fraction 6 and 14) are marked with an arrow on the x-axis. Fractionator errors are labeled with a black dot.

E. coli T7 Express I^{q} cells, which did not contain any plasmid, displayed one big peak spreading from fractions 5 to 7 with an additional shoulder peak in fraction 9. Since no 6xHis-tag is encoded in the *E. coli* T7 Express I^{q} cells, presumably this was unspecifically bound protein. Nisin wild-type showed two peaks with their maxima in fraction 6 and 14 to 16. Though less distinct, a similar profile could be seen in the elution profile of variant Nis A 5 when Boc-K was added. Furthermore, an artifact was visible in this graph (fraction 12), which occurred due to a fractionator error. When no ncAA was added to variant Nis A 5, the second peak could not be observed. A similar pattern was visible in the other tested variant, Nis A 29. When Boc-K was added, a second peak was shown, which did not become apparent when the non-canonical amino acid was not available in the sample. This indicated, that the second peak contained the full-length nisin variant. In fractions 16 and 17, an irregularity was visible, which again most likely originated from a fractionator error. Furthermore, it was notable that in the presence of Boc-K both peaks, but especially the second one, were smaller in variant Nis A 29 than in Nis A 5. This corresponded to the results from the fluorescence assay measurements, where the latter variant displayed higher eGFP values than Nis A 29.

The other tested variants (Nis A 2, Nis A 13, Nis A 17 and Nis A 22, each one with and without Boc-K) also displayed one peak with its maximum in fraction 6. Samples which contained the ncAA additionally showed a second, smaller peak in fraction 14 (data not shown).

For further examination of the expressed protein, the crude extract (cell lysate), as well as the two fractions of interest from the Äkta protein purification run, namely 6 and 14, were loaded onto SDS-gels. The fractions were chosen since the observed peaks had their maxima in these areas. Samples were subjected to SDS-PAGE runs followed by Coomassie staining or Western blot using an anti-6xHis antibody or an anti-GFP antibody (*E. coli* T7 Express I^{q} cells, Nis A wild-type, Nis A 5 and Nis A 29: see Figure 18). While the anti-6xHis antibody binds to the N-terminal Hexahistidine-Tag, the anti-GFP antibody recognizes the reporter protein eGFP, which is located C-terminally of nisin. Only completely translated proteins are recognized by both antibodies. The molecular weight of nisin and the linked eGFP protein is approximately 37 kDa. eGFP alone has a MW of 26.6 kDa while nisin is 3.5 kDa in size. In addition, the 6xHis-tag as well as the linker also contribute to the molecular weight of the protein.

The analysis of *E. coli* T7 Express I^{q} cells (Figure 18 A) showed, that the results from the Coomassie stained SDS-gel correspond to the Äkta elution profile. Fraction 6 contained many times more protein than fraction 14. Since the expression cells did not contain a

hexahistidine-Tag in the untransformed status, no signal was expected, which was confirmed in this Western blot. In contrast, signals were detected when the anti-GFP antibody was used, which was probably due to unspecific background binding. The strongest band in fraction 6 was about 30 kDa big, while in fraction 14 the molecular weight was approximately 40 kDa. Interestingly, these bands were not very distinct in the crude extract. This indicated that the elution fractions were enriched for proteins that the GFP antibody unspecifically binds to. When using the anti-6xHis antibody, the gels which were loaded with Nis A wild-type samples (Figure 18 B) showed a well defined band at approximately 37 kDa in fraction 14 and the crude extract. This corresponded to the calculated size of nisin and led to the assumption that the wanted protein can be found in this fraction in the other samples, too. The second Western Blot, performed with the anti-GFP antibody confirmed the presence of the protein at 37 kDa in fraction 14, but furthermore displayed a weaker band at approximately 30 kDa in fraction 6. Both bands were also visible in the respective other fraction, but by far less distinct, as well as clearly visible in the crude extract. More bands, which could not be assigned to any expected protein were visible between 60 and 80 kDa in all lanes. These might have been dimers of the respective proteins at 30 or 37 kDa or once more, unspecific binding.

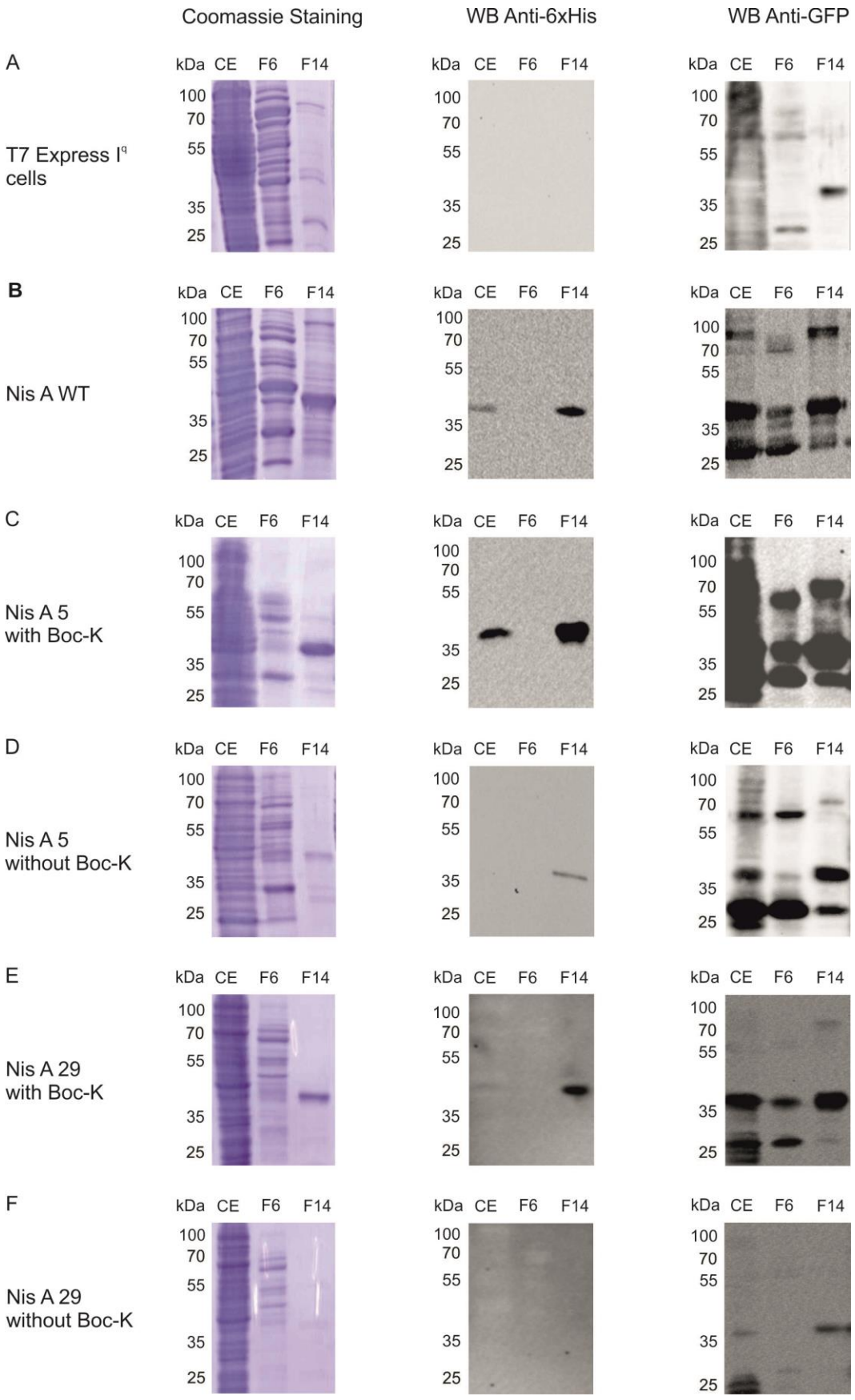

Figure 18: Western blot analysis of the expressed proteins

E. coli T7 Express I^q cells were transformed with the suppressor pair pTB77 as well as the indicated nisin variants or left entirely untransformed. Total cell lysates were subjected to Ni-Ion chromatography and eluted using an imidazole gradient. The crude cell extract (CE) as well as elution fractions 6 and 14 (F6 and F14) were separated according to their size via SDS page. Proteins were either stained with Coomassie or specifically detected by the indicated antibodies (anti-6xHis or anti-GFP) after Western blotting.

In the presence of Boc-K, Western blots for variants Nis A 5 and Nis A 29 displayed comparable bands as for the nisin wild-type sample (Figure 18 C and E). However, differences were notable in the signal intensities. When Boc-K was omitted and samples of Nis A 5 and Nis A 29 were subjected to anti-6xHis Western Blots, the sole signal which could be detected was a weak band at 37 kDa in fraction 14 for Nis A 5 only. When the anti-GFP antibody was used, several proteins could be detected, mainly of 30 and 37 kDa size or twice this size in fractions 6 and 14, respectively (Figure 18 D and F).

The results of the other variants which were analyzed were either similar to the ones from Nis A 5 or Nis A 29. Nis A 2, Nis A 13 and Nis A 17 displayed similar results as the samples from Nis A 5 (data not shown). When the anti-6xHis antibody was used, a band was always visible in the presence of the non-canonical amino acid in fraction 14 at 37 kDa. When Boc-K was absent, the respective band was not present. When the Western blot was performed with the anti-GFP antibody, two bands were visible in each fraction, at 30 and 37 kDa, respectively. The 30 kDa band was more distinct than the one at 37 kDa in fraction 6 and vice versa in fraction 14.

Nis A 22 results corresponded rather to the ones from Nis A 29 samples (data not shown). When Boc-K was added to the preparation, these variants resembled all other variants. However, when no ncAA was present and the anti-6xHis antibody was used for the Western blot, no signal was detectable in either of the examined fractions. When the analysis was performed with the anti-GFP antibody, slight bands were visible at 37 kDa in all lanes and at 30 kDa in the crude extract.

Overall, the results of the analysis of the expressed proteins indicated, that for the samples Nis A 22 and Nis A 29, the presence of Boc-K was crucial for the complete translation of the nisin/GFP fusion protein, whereas for the other tested nisin variants this did not appear to be the case.

5. Discussion

5.1 Optimized fluorescence assay provides high protein yields

All optimizations of the fluorescence assay (section 4.1) aimed at maximizing protein yields of nisin wild-type in order to make differences of Nis A variant expression in the amber suppression screening more easy to distinguish.

For this reason, *E. coli* T7 Express I^q cells were chosen as expression cells. When compared to *E. coli* BL21(DE3) cells, the former displayed higher overall protein yields, and a continuous rise in eGFP expression, even in the stationary phase (section 4.1.1).

The second aspect optimized was the linkage between nisin and the reporter protein GFP. Linker PLrigid, which forms an α-helix (59), appeared to be most suitable. Since it is the longest of all tested linkers and forms a stable secondary structure, it might minimize sterical hindrances of the two proteins most efficiently (section 4.1.2).

The initially used constructs which originated from previous work in the group of Prof. Wagner by Maximilian Fischer (57) contained an mCherry sequence, encoded by a separate gene, as second reporter protein besides GFP. However, RFP fluorescence developed with a temporal delay, possibly due to slower protein maturation kinetics. After translation, fluorescent proteins undergo posttranslational modifications resulting in the formation of a chromophore. This process, which does not require modifying enzymes, is also referred to as maturation. Several previous works independently demonstrated that the maturation time of mCherry, the red fluorescent protein that was tested in this work, is markedly higher than that of eGFP at different temperatures including 37°C (61–64). This might explain the delayed detection of RFP signals in comparison to GFP, which was observed here. Since the co-expression of RFP lowered the expression of GFP and also, RFP did not turn out to be further beneficial as a normalization parameter, the mCherry sequence was removed from the vector for further experiments (section 4.1.3).

The final aspect of optimization in the fluorescence assay set-up was the timing of suppressor pair induction as well as timing of Boc-K addition. Three different strategies were tested. A simultaneous pre-induction of the suppressor pair and addition of the ncAA was tested to examine if the expression of the suppressor pair prior to nisin transcription, combined with the possibility of earlier linkage of Boc-K to the synthetic tRNA provided a faster start of nisin and GFP expression. Furthermore, the suppressor pair was solely pre-induced allowing for it to be already present when nisin translation starts. To test whether the uptake of the

ncAA would possibly be rate-limiting in protein expression, Boc-K was added to the preparations in advance. However, highest fluorescence values were reached when inductions and addition of Boc-K all took place simultaneously right at the start of the plate reader measurement (section 4.1.4). Several mechanisms might contribute to the finding that the premature presence of these components did not provide any benefit in terms of GFP-synthesis. Most non-canonical amino acids are metabolized within the cell, which might result in the formation of toxic intermediates. Indeed, toxicity has been observed for the non-canonical amino acid p-benzoylphenylalanine (pBF) (60), which was also used in this work (section 4.3.2). Furthermore, the sole presence of the exogenous tRNA and aaRS can have toxic effects on expression cells, since interactions with endogenous translation components may occur (60). This effect is supposedly more distinct when the corresponding non-canonical amino acid is not yet available for proper translation. In addition, the translation of endogenous mRNAs is presumably disturbed by the presence of the suppressor pair if the stop codon TAG is suppressed. Even though the amber codon is the least frequently used of the three stop codons and terminates only 7-9% of all genes in *E. coli*, most of whom are non-essential (60,65), this might have a detrimental influence on cellular fitness and lead, as a result in this work, to reduced levels of nisin/GFP expression. The data summarized in Figure 11 supports the idea that simultaneous induction of both inducible systems as well as addition of the ncAA is least harmful and therefore results in highest GFP levels.

5.2 GFP formation in the absence of Boc-K

The performance of the fluorescence assay with different nisin variants in the absence of Boc-K revealed that GFP was not only formed when correct amber suppression and incorporation of the ncAA took place. Observed fluorescence signals by far exceeded values which could be explained by auto-fluorescence of expression cells or other components in the assay apart from the reporter protein itself (section 4.3).

Further experiments showed that high GFP signals could also be detected when nisin variants were expressed in a different cell strain, namely *E. coli* BL21(DE3), in the absence of Boc-K (section 4.3.1). Additionally, the tested suppressor pairs pTB77 or pTB290 were exchanged by the orthogonal tRNA/aaRS pair pEVOL (60) and the corresponding ncAA pBF. Results of the fluorescence assay with and without the ncAA using this exchanged suppressor pair corresponded to data which was collected using pTB77 and pTB290 (section 4.3.2).

However, while in the fluorescence assay samples which were prepared with Boc-K being present and ones without the ncAA in the preparations displayed rather similar GFP signals, differences were clearly visible in the respective elution profiles of nisin variants which had been subjected to Ni-ion affinity chromatography (section 4.3.3). Two elution peaks were obtained for nisin wild-type and Nis A variants which were translated in the presence of Boc-K, while only one peak could be observed when the ncAA was omitted. The respective peaks always appeared in the same elution fractions. Both fractions of interest were subjected to Coomassie staining and Western blot analysis using an anti-6xHis-antibody and an anti-GFP-antibody for each variant, including samples which did not display a peak in the second fraction. Analysis of the above-named experiments suggests that several different effects might contribute to the high GFP signals in absence of the non-canonical amino acid, which are to be discussed in the following.

5.2.1 Translational Readthrough

Partly, expression of full length nisin/GFP proteins might be due to translational readthrough of the TAG stop codon without amber suppression. Technically, TAG suppression by an unnatural tRNA/aaRS pair and incorporation of a ncAA, can be considered as translational readthrough as well. However, in this work, translational readthrough is used as a term describing the effect of protein translation not being terminated when the ribosome encounters a stop codon on the mRNA which results in the incorporation of a natural amino acid. This mechanism could cause eGFP signals irrespective of amber suppression and consequently also in the absence of the non-canonical amino acid Boc-K.

Readthrough mechanisms are known to occur regularly during the replication of some viruses (66,67), but they have also been observed in *Drosophila* (68), mammalian cells (69,70), yeast (71,72) and *E. coli* (73). For *E. coli*, it is proposed that readthrough events are growth phase dependent and more likely to happen during active growth compared to the stationary phase. A given explanation suggests that in this phase, a limited amount of release factor is available. As a result, the termination of translation is impeded, thereby increasing the likelihood of readthrough events occurring (74). These hypotheses were made for the stop codon TGA, but might also apply to the amber stop codon TAG.

Quality control experiments from this work (section 4.3, Figure 14) indicate that a certain amount of GFP is formed with only the nisin/GFP plasmid being present in expression cells, but not the suppressor pair. For differentiation of readthrough signals and background fluorescence, measurements without the nisin/GFP plasmid being present were performed.

The subtraction of the background fluorescence levels from fluorescence intensity values obtained from experiments with the nisin/GFP plasmid being present, but not the suppressor plasmid, allows to estimate the part of translational readthrough signals. The results from this experiment do support the idea, that skipping of the amber stop codon does occur under the given circumstances. However, when compared to regular Nis A wild-type expression, the part of readthrough expression is small (Figure 14 B). Only 4,7 % or 5,5 % of the signal for the variants Nis A 5 or Nis A 29, respectively, can be attributed to readthrough fluorescence after two hours and 2,4 % or 2,9 % after six hours.

The settings may be compared to the ones in other quality control experiments, when expression cells are additionally transformed with the suppressor pair, but no suitable ncAA is added, so that the requirements for amber suppression are not yet fully met. Still, subjecting the incomplete preparations to the fluorescence assay results in fluorescence signals (see section 4.3, Figure 13). These observations suggest that a part of the nisin/GFP proteins is translated by skipping the stop codon TAG. However, when directly compared to the amount of fluorescence signals produced by translational readthrough, fluorescence intensity (FI) signals from expression cells containing the suppressor pair do by far outreach signals from cells in which it was omitted (Figure 19).

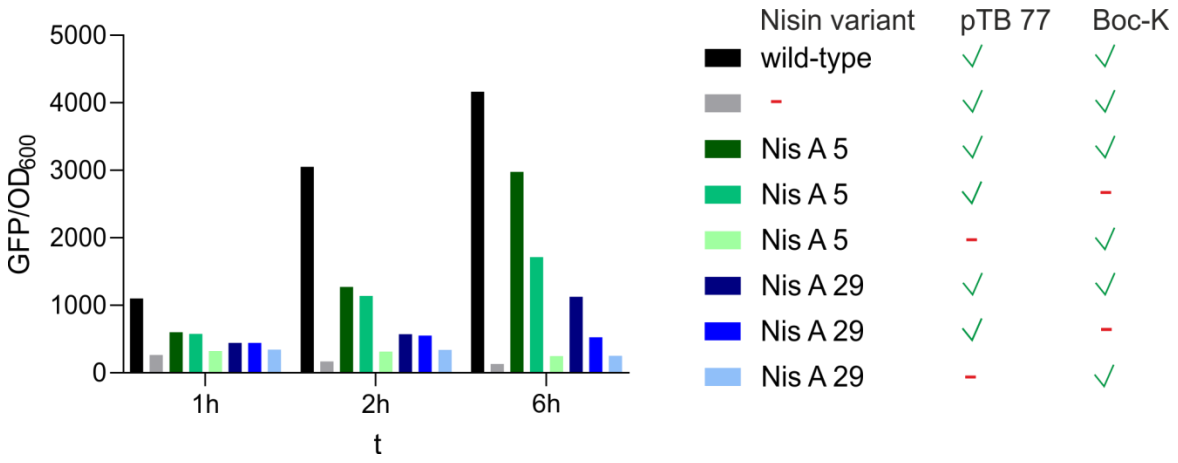

Figure 19: FI signals in absence of the suppressor pair and in absence of Boc-K
E. coli T7 Express I^q cells were transformed with either a nisin variant and the suppressor pair pTB77 and Boc-K was added, or one of the three components was omitted. The graph combines results from different quality control experiments. Nisin wild-type fluorescence serves as positive control (black), while fluorescence detected without the nisin/GFP plasmid represents background signals (grey). Nisin variants 5 and 29 were used as exemplary samples to demonstrate the part of the fluorescence intensity signals deriving from experiments containing all necessary components or only parts of them. The legend indicates which components were absent or present.

Considering the small impact of the readthrough mechanism on nisin/GFP protein translation in the absence of a suitable suppressor pair for amber suppression, this appears to be a negligible source of error for the amber suppression screening.

During amber suppression, TAG readthrough events are probably even diminished compared to their occurrence in the quality control experiments, since in this case, three competitors are available: first: RF-1, second: the exogenous tRNA, which has been evolved to specifically recognize the amber codon and third: random tRNAs, occasionally allowing TAG readthrough. It appears to be more likely in this scenario that a suitable tRNA binds to the mRNA than that random tRNA binding occurs. Furthermore, in the quality control experiments it has already been shown that regular translation termination appears to happen by far more efficient than readthrough events. Also, there is no reason to assume that the frequency of readthrough events would increase in the presence of a suitable suppressor pair and its corresponding ncAA. These arguments indicate that TAG stop codon readthrough does probably falsify the results of the amber suppression screening only to a very low extend.

Another point supporting this conclusion is the following: if readthrough events occurred, the molecular weight of the generated protein would presumably be very similar to a Nisin/GFP protein containing Boc-K, since the two products would only differ in one amino acid. However, nisin purification from cell lysate suggests that protein species were generated that deviate considerably from the expected molecular weight (section 4.3.3). It is therefore unlikely that readthrough events explain the formation of GFP in the absence of Boc-K to a major extent. For proof of these presumptions, analysis of expressed proteins using mass spectrometry could be useful.

To prevent amber codon readthrough, enhanced expression of RF-1, which is the respective release factor terminating translation at TAG codons, could be attempted (75). However, this would at the same time hinder amber suppression, given that the exogenous suppressor pair competes with the release factor as well (see section 5.3.1).

5.2.2 Synthetase unspecificity and insufficient orthogonality

The simultaneous detection of anti-6xHis and anti-GFP signals in Western blots of the nisin variant Nis A 5 which had been cultivated in the absence of Boc-K (section 4.3.3, Figure 18 D) suggests, that full length nisin/GFP protein was translated even though not all necessary components for correct amber suppression were present.

Several requirements must be fulfilled by the aaRS pair and the ncAA for successful amber suppression (48,50,51). For instance, the ncAA must provide good bioavailability. In previous work it was shown that most non-canonical amino acids are taken up by *E. coli*, unless they are highly charged (65,76). Furthermore, the ncAAs must not be recognized by any endogenous aminoacyl-tRNA-synthetase, which might otherwise lead to incorporation into proteins at other positions besides the amber stop codon. However, a mischarge of endogenous tRNAs with the ncAA by the endogenous aaRS is unlikely, since various mechanisms control the correct matching of tRNA with its corresponding specific canonical amino acid (77). Additionally, the newly introduced tRNA/aaRS pair must be orthogonal to all endogenous tRNAs and aminoacyl-tRNA-synthetases, meaning that no cross-reaction between the two systems may occur. This includes that the introduced aaRS must only load the non-canonical amino acid to the orthogonal tRNA. In the manufacturing process, improved orthogonality and synthetase specificity can be achieved by different methods, but most commonly by two sets of positive and negative selection rounds (50,51,78), which are explained in the following.

The first selection set shall ensure that exogenous tRNAs are not recognized and loaded with their respective amino acids by endogenous aaRS, but rather only by their cognate aminoacyl-tRNA-synthetase. For this purpose, *E. coli* expression cells are transformed with a library of mutated exogenous tRNAs. In the negative selection round, expression cells encode for a barnase gene containing TAG mutations. When no non-canonical amino acid is added, only those tRNAs which are recognized by *E. coli* synthetases allow amber suppression and therefore translation of the toxic barnase protein, which leads to cell death. Only cells containing tRNAs which do not interact with any endogenous aaRS survive (Figure 20 A).

Resulting exogenous tRNAs from the negative selection round are then used for positive selection, for which cells express a matching orthogonal aminoacyl-tRNA-synthetase, as well as a ß-lactamase gene, which contains a TAG mutation. Only cells carrying tRNAs which are recognized by their cognate aaRS can express ß-lactamase and survive, since ampicillin is present in the media (Figure 20 B).

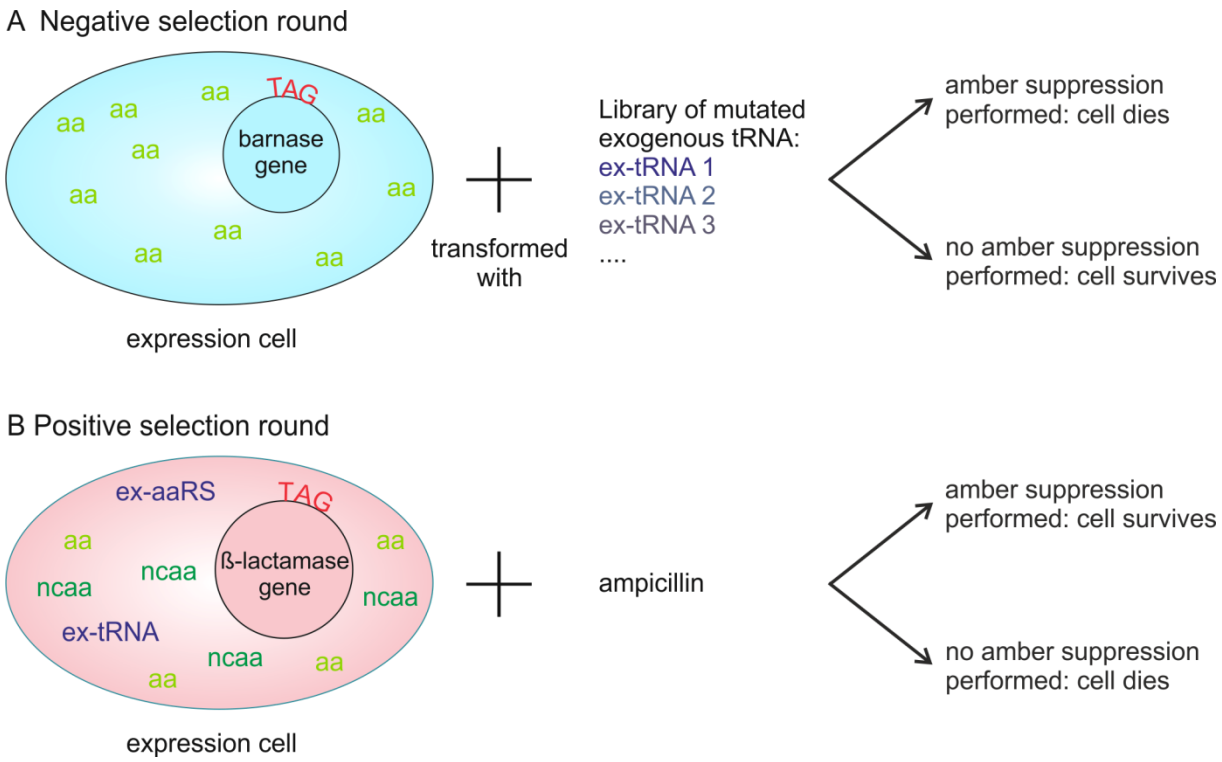

Figure 20: First selection set for suppressor pairs for improved orthogonality
(A) Expression cells are transformed with a toxic barnase gene containing a TAG mutation. A library of mutated exogenous tRNAs (ex-tRNAs) is screened by transformation of one tRNA, respectively to the expression cell. Exogenous tRNAs which are loaded by endogenous aaRS with amino acids die, since the toxic barnase gene is expressed, while when the exogenous tRNA remains unloaded the cell survives. These tRNAs are then subjected to the positive selection round. (B) Expression cells are transformed with the chosen ex-tRNAs from the first round and the orthogonal aaRS, as well as with a ß-lactamase gene containing a TAG mutation. ß-lactamase expression by amber suppression is necessary for the cell to survive, since ampicillin is present in the media.

The second selection set aims at enhancing the specificity of the aminoacyl-tRNA-synthetase, so it will only aminoacylate the correct non-canonical amino acid. For the following steps, tRNA mutants, which survived the first set of selection rounds are used. In the positive selection rounds, amber suppression enables transformation of the chloramphenicol acetyl transferase (CAT), which allows bacteria to grow in the presence of the antibiotic chloramphenicol, which is added to the media bacterial cultures are grown in. For this, TAG mutations are introduced in the CAT gene. The non-canonical amino acid or other naturally occurring amino acids are present in the media and may be aminoacylated by one of the tRNA-synthetase mutants which are tested, and be incorporated in the nascent chloramphenicol acetyl transferase protein. Independently from which amino acid is incorporated at the amber suppression site, the function of CAT is not affected (Figure 21 A). In negative selection rounds, a barnase gene containing TAG mutations is present in the cells,

while the non-canonical amino acid the respective suppressor pair is being evolved for, is not added to the media. Bacterial clones containing aminoacyl-tRNA-synthetase mutants, which will charge tRNA with any available amino acid and by this allow suppression of the amber stop codon and therefore expression of barnase, will die due to the high toxicity of their protein (Figure 21 B). Only bacterial clones with an aminoacyl-tRNA-synthetase mutant which reliably and uniquely aminoacylates the ncAA it was evolved for can survive this process.

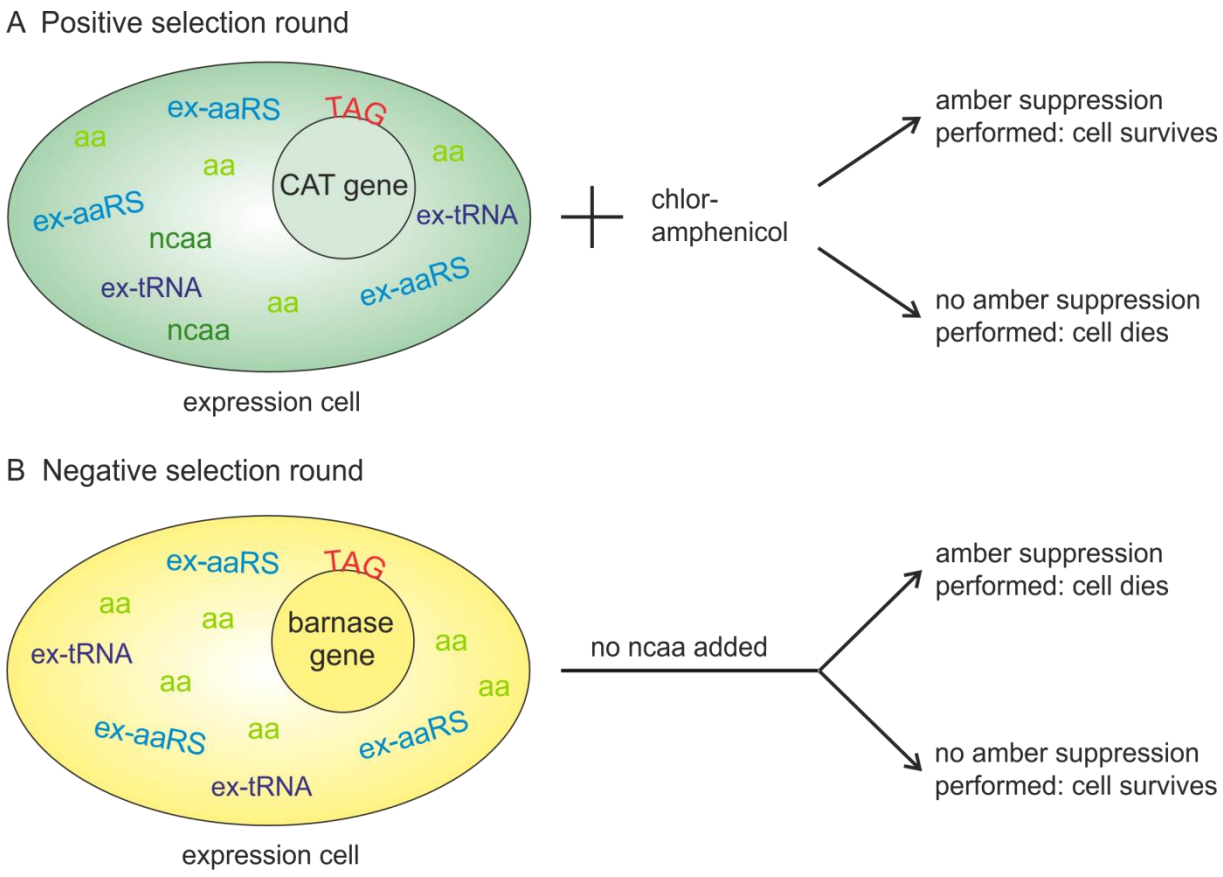

Figure 21: Second selection set for suppressor pairs for enhanced specificity
(A) In positive selection rounds, expression cells are transformed with a CAT gene containing a TAG mutation, the exogenous tRNA from the first set and the orthogonal aaRS. Expression of the CAT allows bacteria to grow in the presence of the antibiotic chloramphenicol and is only possible by performance of amber suppression, irrespectively whether a canonical amino acid or a non-canonical amino acid is incorporated. Orthogonal suppressor pairs which do not function together do not lead to expression of the CAT and therefore cell death. (B) Expression cells are transformed with a barnase gene, which contains a TAG mutation. No non-canonical amino acid is added to the media. When unspecific loading of the exogenous tRNA takes place, barnase is expressed and leads to cell death due to its toxicity. Only bacterial clones with an aminoacyl-tRNA-synthetase mutant which uniquely aminoacylates the ncAA it was evolved for survive this process.

Typically, two to three rounds of the entire selection procedure are performed and should ensure for a highly orthogonal suppressor pair, which only recognizes the one non-canonical amino acid it was evolved for. This selection scheme is the most common one and was used for the suppressor pairs in this work. However, other selection schemes and additional approaches to improve the suppressor pairs do exist (79,80).

As mentioned above, results from the fluorescence assay of variant Nis A 5 in the absence of Boc-K (section 4.3), as well as Western blot analysis of the expressed protein (section 4.3.3) may lead to the idea, that synthetase specificity in this experiment was not sufficient.
It is therefore conceivable, that the suppressor pair which was used in this experiment (pTB77) is either not sufficiently orthogonal, meaning that *E. coli* aminoacyl-tRNA-synthetases might have charged the exogenous tRNA with their cognate amino acids, or that the exogenous aaRS is promiscuous in the aminoacylation of amino acids. To this point, it remains unclear, which one of these mechanisms leads to the unspecific charge of the tRNA. The observed unspecificity might also occur with the use of suppressor pair pTB290, although this is yet to be tested. Since for the generation of further suppressor pairs it might be helpful to know which selection step is rather error-prone or if both equally contribute to the unspecificity, this could easily be determined by reassessment of the negative selection rounds of both selection steps as described. For the first step, the tRNA needs to be available separately from the aaRS. In both steps, in case of no interaction with endogenous aaRS or naturally occurring amino acids, cell death is expected. In case of survival of bacteria in either of the steps, other mutations of the respectively tested component (tRNA or aaRS) could be subjected to further selection rounds. If errors repeatedly occur in the same selection step with different suppressor pairs that display unspecificity, alteration of the respective set might be necessary.
Following simple modifications of the above described selection process could be helpful tools for altering the orthogonality of tRNA and aaRS and enhancing the specificity of the aminoacyl-tRNA-synthetase towards its non-canonical amino acid. Negative selection rounds are typically carried out with two or three TAG mutations in the barnase gene (78,81). The reduction to only one TAG codon in the gene might increase the error rate of wrongly sorted out tRNAs or tRNA/aaRS pairs due to other mechanisms than amber suppression leading to barnase expression. However, it would also improve sensitivity of the selection, since every single amber suppression event would directly lead to expression of the toxic protein. This change could be applied to both sets of selection rounds.

Additionally, positive selection rounds of both sets could be repeated more often with rising concentrations of the antibiotic used, as it is suggested by some authors (78). When ß-lactamase or the chloramphenicol acetyl transferase are not expressed at all, or only at low rates due to inconsistent amber suppression, cell death occurs. For this reason tRNAs and aaRS/tRNA pairs working with a high fidelity will tolerate higher concentrations of the respective antibiotic. Furthermore, current methods do typically require the amber suppression of just one amber stop codon in the genes of the respectively used antibiotics. By introducing a second or even third TAG mutation at permissive sites, only those constructs which provide reliable amber suppression would survive the selection round.

For further improvement of the aminoacyl-tRNA-synthetase specificity, a different method shall briefly be mentioned: the fluorescence-based cell sorting for choice of efficiently incorporating synthetases (82,83). It could either replace the second selection set, or additionally be performed with the most promising aaRS candidates from the second selection set. In both cases, cloning of tRNA and aaRS into a different vector would be required, since this method is based on a single plasmid assay. In addition to the suppressor pair, the plasmid carries a modified eGFP gene which contains a TAG mutation at a permissive site. Amber suppression allows expression of the fluorescent protein. Bacteria which were transformed with this vector are then grown in six selection rounds with and without the addition of the suitable non-canonical amino acid. When the ncAA is present, amber suppression and therefore eGFP expression is desired. For this reason cells are sorted by FACS (fluorescence activated cell sorting) according to their eGFP expression. Highly fluorescent bacteria are then subjected to the following selection round in which the ncAA is not added to the media. Under these conditions, no amber suppression should be possible and only cells which do not express eGFP are then used for the next positive selection round. However, with this method even after six cycles, still 12.5 % of the sorted cells display eGFP expression in the absence of the non-canonical amino acid (82).

A combination of both described methods, even though time-consuming and complex, might result in better synthetase specificity.

Various efforts could be made to enhance the suppressor pair orthogonality and synthetase specificity. However, in this work, unspecificity was only observed for one out of six samples tested and was not mainly responsible for the high GFP fluorescence intensity values in the absence of Boc-K. For this reason, other possible causes for the translation of green fluorescent protein after the stop codon should be investigated.

5.2.3 Internal translation starts

Translational readthrough (section 5.2.1) as well as synthetase unspecificity (section 5.2.2) do likely only explain a small part of the fluorescence signals which can be detected when cultures for the fluorescence assay are prepared without the addition of the non-canonical amino acid. Further investigation of the expressed protein (section 4.3.3) brings up another source of error where measured signals might derive from. Western blot analysis comparing exemplary samples from the fluorescence assay performed in presence or absence of Boc-K showed that the C-terminal GFP is produced in both cases. In contrast, the hexahistidine-Tag, which is situated at the N-terminus of the protein, is only part of the translated fusion protein, when the ncAA is present in the media (Figure 18). These findings indicate that N-terminally truncated proteins are produced when no Boc-K is added.

Consequently, translation of these proteins must start at some point downstream of the 6xHis-Tag. For protein translation initiation, a start codon AUG, which encodes for the amino acid methionine is required. With regard to the amino acid sequence of nisin following the 6xHis-Tag, three methionines are present that might serve as additional start codons (Figure 22). The first eligible start codon initiates the leader of nisin, a part of the protein which is translated, but later on removed by posttranslational modifications when the protein is naturally expressed by *L. lactis* (37). The remaining two AUG codons can be found within nisin itself, at amino acid positions 17 and 21 of the protein.

Figure 22: Methionines that might serve as internal start codons
Schematic view of the parts of which the fully translated nisin A fusion protein consists of. The nisin amino acid sequence is fully displayed, whereas in the leader only the methionine, as it is the amino acid of interest, is shown. Methionines that might serve as internal translation starts are highlighted in green.

Other codons such as GUG, UUG or CUG, which occasionally initiate translation as well do exist, but since they are by far not used as frequently as AUG (84), they will not be considered in this work.

In general, for translation initiation in bacteria, a Shine-Dalgarno (SD) sequence (85) upstream of the start codon AUG is required. The recognition of this nucleotide arrangement

on the mRNA leads to its binding to the anti-Shine-Dalgarno sequence, which is found on the ribosomal RNA. Following, three initiation factors (IF-1, IF-2 and IF-3), the 30S and the 50S subunits of the ribosome and the start tRNA, which is charged with methionine, bind to the mRNA (86–89).

Every methionine within the fusion protein is led by a so-called Shine-Dalgarno-like sequence in an acceptable distance to the AUG codon (90). The term SD-like sequences refers to nucleotide motifs within the coding region of a gene, which are complementary to the anti-SD sequence. These internal SD sites can cause translational pausing (91), which among other effects may facilitate co-translational folding (92), but have also been shown to allow minor amounts of internal starts (93). Disrupted translation due to an incorporated TAG stop codon within nisin could lead to increased rates of internal starts and therefore explain the N-terminally truncated proteins causing fluorescence signals in the absence of Boc-K. However, this theory presumably only applies to the AUG codons within nisin itself (at positions 17 and 21), since the start codon located at the beginning of the leader is not preceded by a stop codon and would therefore only allow negligible amounts of internal starts. This idea is supported by Western blot results from section 4.3.3 (Figure 18), which show that in case of nisin wild-type translation, which does not contain any inserted stop codons, no truncated proteins are produced.

Apart from translational starts triggered by a Shine Dalgarno (SD) sequence, more recent findings have shown that translation in prokaryotes can also be initiated SD-independently (94,95). For instance, this option is mandatory for leaderless genes, which lack an 5' untranslated region providing the SD sequence (96). Another SD-independent mechanism includes the ribosomal protein S1, which was identified to initiate protein translation autonomously (97).

Assuming that one of these mechanisms might allow an internal translation start, independently of whether the start is based on a Shine-Dalgarno mechanism or not, fluorescence can be detected without amber suppression taking place. In the following, possible scenarios of protein truncation are discussed. A schematic visualization is provided in Figure 23.

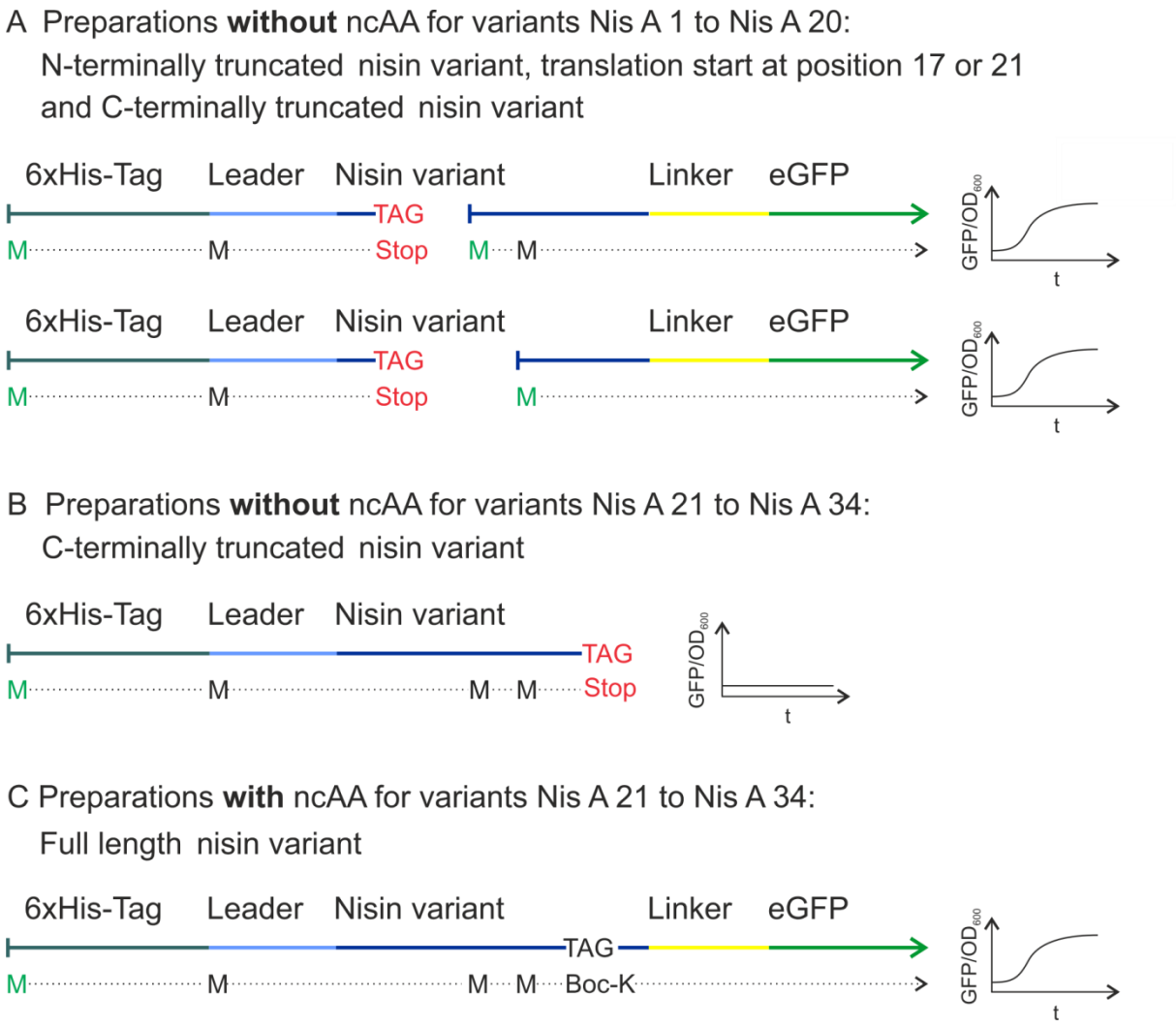

Figure 23: Theory of internal translation starts
(A) When the ncAA is not added to the preparations for variants Nis A 1 to Nis A 20, a start codon is present at positions 17 or 21 of nisin, which can lead to translation of an N-terminally truncated protein, which expresses GFP. Furthermore, due to the translation stop, a C-terminally truncated protein is expressed. (B) When the ncAA is not present in preparations for the samples Nis A 21 to Nis A 34, no start codon is available downstream of the stop codon, no internal start is possible and no GFP is expressed. A C-terminally truncated protein results. (C) For variants Nis A 21 to Nis A 34 which were tested in the presence of Boc-K, amber suppression must be performed in order for GFP to be expressed.

Considering positions 17 and 21 as internal start sites, two groups can be differentiated. The first one includes all nisin variants which contain the TAG mutation at position 1 to 20 (Nis A 1 to Nis A 20). In this group, at least one start codon is present downstream of the newly incorporated stop codon TAG. An internal start at position 17 or 21 of nisin is possible, resulting in an N-terminally truncated protein which expresses GFP and therefore allows fluorescence detection (Figure 23 A).

In the second group, all nisin variants that underwent the TAG mutation at positions 21 to 34 (Nis A 21 to Nis A 34) can be found. In this case, no methionine is available downstream of the stop codon and therefore no internal start is possible. When no ncAA is added to the preparations, translation stops, resulting in a C-terminally truncated protein (Figure 23 B). However, when Boc-K is present in the media, the only option how translation can be continued is amber suppression, which was originally aimed at (Figure 23 C). This indicates that fluorescence intensity signals, which were obtained for variants Nis A 21 to Nis A 34 in the amber suppression screening do actually derive from suppression of the stop codon TAG and incorporation of the ncAA Boc-K (except for negligible amounts due to translational readthrough and synthetase unspecificity, see sections 5.2.1 and 5.2.2).
This thesis is supported by results from the experiments of certain Nis A variants with and without the respective non-canonical amino acid. In T7 Express I^q cells, which were tested with the suppressor pairs pTB77, pTB290 and the suppressor pair pEVOL, a significant difference between the variants with and without the respective suppressor pair may be noted (see Figure 13 and Figure 16). For an unknown reason, this observation does not apply, when the same experiment was carried out in BL21(DE3) expression cells with the suppressor pairs pTB77 and pTB290 (see Figure 15).

In order to eliminate the possibility of internal translation starts, the questionable methionines could be substituted by other amino acids. Since methionine is not charged and non-polar, a rather similar amino acid such as alanine might be suitable. Excision of the respective methionine and incorporation of the new amino acid could be achieved by overlap extension PCR (98) or quick change mutagenesis (99).

5.3 Amber suppression in the performed screening

In a synopsis of the previous discussion, amber suppression and translation of the whole fusion protein may only be assumed for nisin variants Nis A 21 to Nis A 34. The protein yielded for all other variants may include truncated proteins to an unknown extend (see Figure 23). The part of fully translated nisin fusion proteins could be assessed by mass spectrometry of the resulting proteins (100). Since for above named variants, it may be assumed that amber suppression was carried out sufficiently, conclusions can be made about amber suppression in the rear part of the protein. For this reason, the following section only refers to nisin variants Nis A 21 to Nis A 34, if not specified otherwise.

5.3.1 Suppression efficiency

Overall yields for the expressed fusion proteins in the amber suppression screening were satisfactory, ranging from 15 % (Nis A 34 for pTB 77 and Nis A 27 for pTB 290) to 68 % for Nis A 31 using plasmid pTB 77 and 66 % for Nis A 21 coupled with plasmid pTB 290 (see Figure 24) after 6 hours compared to nisin wild-type expression, which was set to 100 %. Literature research on protein yields for comparable experiments for amber suppression suggests values between 35 % and 50 % (60,101).

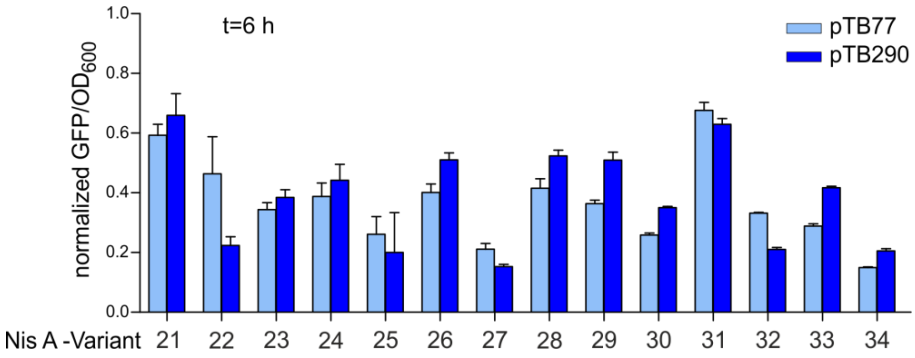

Figure 24: Amber suppression in the variants Nis A 21 to Nis A 34
This figure is a cutout from Figure 12 and only displays results for variants Nis A 21 to Nis A 34, since in this part of the protein, measured GFP FI signals do mainly derive from amber suppression. The fluorescence assay was carried out in biological triplicates with each variant of Nis A. For better comparability, each plate reader run contained a nisin wild-type sample, which was also transformed with the suppressor pair to be tested in the same run. GFP FI/OD_{600} values of the variants were normalized to nisin wild-type GFP FI/OD_{600}, which was set to 1. Diagrams display the respective normalized Nis A variant GFP FI/OD_{600} values after 6 hours. Error bars indicate the SEM of biological replicates.

Since one of the original aims of this project is to find a way of identifying possible therapeutics, the possibility of large-scale production of the resulting protein is crucial. Various efforts, such as the following ones might be taken for increasing yields.

Amber suppression and translation termination are the two mechanisms competing against each other when a TAG codon is encountered. When the stop codon is recognized by RF-1, the latter causes the ribosome to dissociate from the mRNA (102). Due to this natural effect, proteins which contain the TAG codon as a sense codon for amber suppression are often translated incompletely resulting in C-terminally truncated proteins. In order to eliminate this problem, *Wang et. al.* created an *E. coli* RF-1 knockout cell strain, which enables amber suppression efficiency even surpassing that of wild-type expression (103). Similar approaches use several mutations to inactivate release factor 1, also showing significantly higher yields for amber suppressed proteins (104,105). Since the remaining stop codons (TAA and TGA)

both use the release factor 2 (106), translation termination is not affected for genes terminated by TAA or TGA, as long as RF-2 expression is constitutively over-expressed. Nisin expression in an RF-1 knockout strain might even allow incorporation of the ncAA at more than one site of the protein.

On the other side, the usage of an RF-1 knockout strain might at the same time promote unwanted TAG readthrough, which has been shown to contribute to unspecific background fluorescence signals in this work (see section 5.2.1).

Other authors suggest an *E. coli* based cell-free method for enhancing protein yields. Depending on the respective work, the ncAA to be incorporated and the used tRNAs and aaRS, protein yields range from 50 % to 120 % compared to the respective wild-type expression level (107,108).

A different option to increase the incorporation of ncAAs during amber suppression is lowering the affinity of the ribosome to RF-1 by usage of evolved ribosomes (109).

5.3.2 Promising nisin A variants

Regarding promising sites within nisin for amino acid incorporation, two positions are especially remarkable, namely Nis A 21 and Nis A 31. High expression levels are provided here, independently of which suppressor pair is used (see Figure 24). While at position 21, a methionine is substituted by Boc-K, at position 31, the ncAA takes the place of a histidine. Both amino acids are not parts of ring formation in nisin (see Figure 2 B), which might facilitate expression of the modified protein.

In further steps, a screening of the whole protein should be aimed at, in order to identify promising Nis A variants over the whole span of the protein. For this, methionines serving as internal translation starts should be substituted by a different amino acid. After identification of positions which allow incorporation of the non-canonical amino acid and provide high yields, functionality of the engineered proteins must be tested, i.e. using a minimum inhibitory concentration (MIC) assay in order to test antibiotic effectiveness.

When promising positions for incorporation of ncAAs are identified and functionality is ensured, the next aim should be to incorporate different non-canonical amino acids, in order to tap the full potential, that ncAA bring with them for alteration of chemical properties of the engineered proteins.

In conclusion, data from this work has shown that incorporation of non-canonical amino acids in nisin using amber suppression is achievable. However, further optimizations will be

required to efficiently inhibit usage of internal translation starts and the production of truncated protein variants.

Appendix

List of figures

List of tables

Abbreviations

aaRS	Aminoacyl-tRNA-synthetase
Amp	Ampicillin
Boc-K	Boc-Lysine
Chor	Chloramphenicol
Dha	2,3-didehydroalanine
Dhb	(z)-2,3-didehydrobutyrine
FI	Fluorescence intensity
IF-1 / IF-2 / IF-3	Initiation factor 1/initiation factor 2/initiation factor 3
LB	Lysogeny broth
MIC	Minimum inhibitory concentration
MRSA	Methicillin-resistant staphylococcus aureus
ncAA	Non-canonical amino acid
Nis A	Nisin A
OD	Optical density
pBF	p-benzoylphenylalanine
PDR	Pandrug resistant
RF-1 / RF-2 / RF-3	Release factor 1/release factor 2/release factor 3
SCS	Stop codon suppression
SD	Shine-Dalgarno
SEM	Standard error of the mean
Ser	Serine
SPI	Selective pressure incorporation
Thr	Threonine
VRE	Vancomycin-resistant enterococci
WB	Western Blot
XDR	Extensively drug resistant

References

1. Fleming A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. Bull World Health Organ. 2001;79(8):780–90.
2. Laxminarayan R, Duse A, Wattal C, Zaidi AKM, Wertheim HFL, Sumpradit N, u. a. Antibiotic resistance-the need for global solutions. Lancet Infect Dis. Dezember 2013;13(12):1057–98.
3. Ten threats to global health in 2019 [Internet]. [zitiert 27. Januar 2019]. Verfügbar unter: https://www.who.int/emergencies/ten-threats-to-global-health-in-2019
4. WHO | Global action plan on antimicrobial resistance [Internet]. WHO. [zitiert 17. Januar 2019]. Verfügbar unter: http://www.who.int/antimicrobial-resistance/publications/global-action-plan/en/
5. GLASS | Global antimicrobial resistance surveillance system (GLASS) report [Internet]. WHO. [zitiert 14. Januar 2019]. Verfügbar unter: http://www.who.int/glass/resources/publications/early-implementation-report/en/
6. Abraham EP, Chain E. An enzyme from bacteria able to destroy penicillin. 1940. Rev Infect Dis. August 1988;10(4):677–8.
7. Fleming, Alexander. Penicillin [Internet]. Nobel Lecture gehalten auf der; 1945 Dez 11 [zitiert 17. Januar 2019]. Verfügbar unter: https://www.nobelprize.org/prizes/medicine/1945/fleming/lecture/
8. Davies J, Davies D. Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev. September 2010;74(3):417–33.
9. Holmes AH, Moore LSP, Sundsfjord A, Steinbakk M, Regmi S, Karkey A, u. a. Understanding the mechanisms and drivers of antimicrobial resistance. The Lancet. 9. Januar 2016;387(10014):176–87.
10. Levy SB, Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. Nature Medicine. 30. November 2004;10(12s):S122–9.
11. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Herausgeber. ANTIBIOTIC RESISTANCE THREATS in the United States, 2013 [Internet]. 2013 [zitiert 12. Dezember 2019]. Verfügbar unter: https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf
12. Watkins RR, Bonomo RA. Overview: Global and Local Impact of Antibiotic Resistance. Infectious Disease Clinics of North America. 1. Juni 2016;30(2):313–22.
13. Smith R, Coast J. The true cost of antimicrobial resistance. BMJ. 11. März 2013;346:f1493.
14. Sahl HG, Bierbaum G. Lantibiotics: biosynthesis and biological activities of uniquely modified peptides from gram-positive bacteria. Annu Rev Microbiol. 1998;52:41–79.
15. Castiglione F, Lazzarini A, Carrano L, Corti E, Ciciliato I, Gastaldo L, u. a. Determining the structure and mode of action of microbisporicin, a potent lantibiotic active against multiresistant pathogens. Chem Biol. Januar 2008;15(1):22–31.
16. Dischinger J, Basi Chipalu S, Bierbaum G. Lantibiotics: promising candidates for future applications in health care. Int J Med Microbiol. Januar 2014;304(1):51–62.
17. Willey JM, van der Donk WA. Lantibiotics: peptides of diverse structure and function. Annu Rev Microbiol. 2007;61:477–501.
18. Chatterjee C, Paul M, Xie L, van der Donk WA. Biosynthesis and mode of action of lantibiotics. Chem Rev. Februar 2005;105(2):633–84.
19. Brötz H, Sahl HG. New insights into the mechanism of action of lantibiotics--diverse biological effects by binding to the same molecular target. J Antimicrob Chemother. Juli 2000;46(1):1–6.
20. Stevens KA, Sheldon BW, Klapes NA, Klaenhammer TR. Nisin treatment for inactivation of Salmonella species and other gram-negative bacteria. Appl Environ Microbiol. Dezember 1991;57(12):3613–5.
21. Wiedemann I, Breukink E, van Kraaij C, Kuipers OP, Bierbaum G, de Kruijff B, u. a. Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. J Biol Chem. 19. Januar 2001;276(3):1772–9.
22. Hasper HE, Kramer NE, Smith JL, Hillman JD, Zachariah C, Kuipers OP, u. a. An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid II. Science. 15. September 2006;313(5793):1636–7.

23. Otto M, Peschel A, Götz F. Producer self-protection against the lantibiotic epidermin by the ABC transporter EpiFEG of Staphylococcus epidermidis Tü3298. FEMS Microbiol Lett. 15. September 1998;166(2):203–11.
24. Kuipers OP, Beerthuyzen MM, Siezen RJ, De Vos WM. Characterization of the nisin gene cluster nisABTCIPR of Lactococcus lactis. Requirement of expression of the nisA and nisI genes for development of immunity. Eur J Biochem. 15. August 1993;216(1):281–91.
25. Draper LA, Ross RP, Hill C, Cotter PD. Lantibiotic immunity. Curr Protein Pept Sci. Februar 2008;9(1):39–49.
26. Aso Y, Okuda K, Nagao J, Kanemasa Y, Thi Bich Phuong N, Koga H, u. a. A novel type of immunity protein, NukH, for the lantibiotic nukacin ISK-1 produced by Staphylococcus warneri ISK-1. Biosci Biotechnol Biochem. Juli 2005;69(7):1403–10.
27. Draper LA, Cotter PD, Hill C, Ross RP. Lantibiotic resistance. Microbiol Mol Biol Rev. Juni 2015;79(2):171–91.
28. Sun Z, Zhong J, Liang X, Liu J, Chen X, Huan L. Novel mechanism for nisin resistance via proteolytic degradation of nisin by the nisin resistance protein NSR. Antimicrob Agents Chemother. Mai 2009;53(5):1964–73.
29. van Heel AJ, Montalban-Lopez M, Kuipers OP. Evaluating the feasibility of lantibiotics as an alternative therapy against bacterial infections in humans. Expert Opin Drug Metab Toxicol. Juni 2011;7(6):675–80.
30. Rogers LA, Whittier EO. LIMITING FACTORS IN THE LACTIC FERMENTATION. J Bacteriol. Oktober 1928;16(4):211–29.
31. Delves-Broughton J, Blackburn P, Evans RJ, Hugenholtz J. Applications of the bacteriocin, nisin. Antonie Van Leeuwenhoek. Februar 1996;69(2):193–202.
32. Shin JM, Gwak JW, Kamarajan P, Fenno JC, Rickard AH, Kapila YL. Biomedical Applications of Nisin. J Appl Microbiol. Juni 2016;120(6):1449–65.
33. Le Blay G, Lacroix C, Zihler A, Fliss I. In vitro inhibition activity of nisin A, nisin Z, pediocin PA-1 and antibiotics against common intestinal bacteria. Lett Appl Microbiol. September 2007;45(3):252–7.
34. Mulders JW, Boerrigter IJ, Rollema HS, Siezen RJ, de Vos WM. Identification and characterization of the lantibiotic nisin Z, a natural nisin variant. Eur J Biochem. 1. November 1991;201(3):581–4.
35. Zendo T, Fukao M, Ueda K, Higuchi T, Nakayama J, Sonomoto K. Identification of the lantibiotic nisin Q, a new natural nisin variant produced by Lactococcus lactis 61-14 isolated from a river in Japan. Biosci Biotechnol Biochem. Juli 2003;67(7):1616–9.
36. Wirawan RE, Klesse NA, Jack RW, Tagg JR. Molecular and genetic characterization of a novel nisin variant produced by Streptococcus uberis. Appl Environ Microbiol. Februar 2006;72(2):1148–56.
37. Field D, Cotter PD, Ross RP, Hill C. Bioengineering of the model lantibiotic nisin. Bioengineered. 2015;6(4):187–92.
38. Yuan J, Zhang Z-Z, Chen X-Z, Yang W, Huan L-D. Site-directed mutagenesis of the hinge region of nisinZ and properties of nisinZ mutants. Appl Microbiol Biotechnol. Juni 2004;64(6):806–15.
39. Field D, Connor PMO, Cotter PD, Hill C, Ross RP. The generation of nisin variants with enhanced activity against specific gram-positive pathogens. Mol Microbiol. Juli 2008;69(1):218–30.
40. Budisa N. Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire. Angew Chem Int Ed Engl. 3. Dezember 2004;43(47):6426–63.
41. Ross AC, Liu H, Pattabiraman VR, Vederas JC. Synthesis of the lantibiotic lactocin S using peptide cyclizations on solid phase. J Am Chem Soc. 20. Januar 2010;132(2):462–3.
42. Fukase K, Kitazawa M, Sano A, Shimbo K, Horimoto S, Fujita H, u. a. Synthetic Study on Peptide Antibiotic Nisin. V. Total Synthesis of Nisin. BCSJ. 1. August 1992;65(8):2227–40.
43. Ongey EL, Neubauer P. Lanthipeptides: chemical synthesis versus in vivo biosynthesis as tools for pharmaceutical production. Microb Cell Fact. 7. Juni 2016;15:97.
44. Johnson JA, Lu YY, Van Deventer JA, Tirrell DA. Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. Curr Opin Chem Biol. Dezember 2010;14(6):774–80.
45. Zhou L, Shao J, Li Q, van Heel AJ, de Vries MP, Broos J, u. a. Incorporation of tryptophan analogues into the lantibiotic nisin. Amino Acids. 2016;48:1309–18.

46. Nickling JH, Baumann T, Schmitt F-J, Bartholomae M, Kuipers OP, Friedrich T, u. a. Antimicrobial Peptides Produced by Selective Pressure Incorporation of Non-canonical Amino Acids. J Vis Exp. 04 2018;(135).
47. Baumann T, Nickling JH, Bartholomae M, Buivydas A, Kuipers OP, Budisa N. Prospects of In vivo Incorporation of Non-canonical Amino Acids for the Chemical Diversification of Antimicrobial Peptides. Front Microbiol. 2017;8:124.
48. Wals K, Ovaa H. Unnatural amino acid incorporation in E. coli: current and future applications in the design of therapeutic proteins. Front Chem [Internet]. 1. April 2014 [zitiert 25. November 2016];2. Verfügbar unter: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3982533/
49. Crnković A, Suzuki T, Söll D, Reynolds NM. Pyrrolysyl-tRNA synthetase, an aminoacyl-tRNA synthetase for genetic code expansion. Croat Chem Acta. Juni 2016;89(2):163–74.
50. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of Escherichia coli. Science. 20. April 2001;292(5516):498–500.
51. Liu CC, Schultz PG. Adding New Chemistries to the Genetic Code. Annual Review of Biochemistry. 2010;79(1):413–44.
52. Ryu Y, Schultz PG. Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. Nat Methods. April 2006;3(4):263–5.
53. Dumas A, Lercher L, Spicer CD, Davis BG. Designing logical codon reassignment - Expanding the chemistry in biology. Chem Sci. 1. Januar 2015;6(1):50–69.
54. Pott M, Schmidt MJ, Summerer D. Evolved sequence contexts for highly efficient amber suppression with noncanonical amino acids. ACS Chem Biol. 19. Dezember 2014;9(12):2815–22.
55. Synthetic Biology for the production of functional peptides | SYNPEPTIDE Project | FP7 | CORDIS | European Commission [Internet]. [zitiert 11. Oktober 2019]. Verfügbar unter: https://cordis.europa.eu/project/rcn/110110/factsheet/de
56. Schmitt S, Montalbán-López M, Peterhoff D, Deng J, Wagner R, Held M, u. a. Analysis of modular bioengineered antimicrobial lanthipeptides at nanoliter scale. Nat Chem Biol. 2019;15(5):437–43.
57. Fischer M. Optimierungen zur Expression von Nisin als Grundlage für den späteren Einbau nicht-kanonischer Aminosäuren in Escherichia coli. Universität Regensburg; 2014.
58. Inoue H, Nojima H, Okayama H. High efficiency transformation of Escherichia coli with plasmids. Gene. 30. November 1990;96(1):23–8.
59. Yan W, Imanishi M, Futaki S, Sugiura Y. Alpha-helical linker of an artificial 6-zinc finger peptide contributes to selective DNA binding to a discontinuous recognition sequence. Biochemistry. 24. Juli 2007;46(29):8517–24.
60. Young TS, Ahmad I, Yin JA, Schultz PG. An enhanced system for unnatural amino acid mutagenesis in E. coli. J Mol Biol. 15. Januar 2010;395(2):361–74.
61. Macdonald PJ, Chen Y, Mueller JD. Chromophore maturation and fluorescence fluctuation spectroscopy of fluorescent proteins in a cell-free expression system. Anal Biochem. 1. Februar 2012;421(1):291–8.
62. Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol. Dezember 2004;22(12):1567–72.
63. Merzlyak EM, Goedhart J, Shcherbo D, Bulina ME, Shcheglov AS, Fradkov AF, u. a. Bright monomeric red fluorescent protein with an extended fluorescence lifetime. Nat Methods. Juli 2007;4(7):555–7.
64. Iizuka R, Yamagishi-Shirasaki M, Funatsu T. Kinetic study of de novo chromophore maturation of fluorescent proteins. Anal Biochem. 15. Juli 2011;414(2):173–8.
65. Xie J, Schultz PG. An expanding genetic code. Methods. Juli 2005;36(3):227–38.
66. Skuzeski JM, Nichols LM, Gesteland RF, Atkins JF. The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. J Mol Biol. 20. März 1991;218(2):365–73.
67. Li G, Rice CM. The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream of the termination codon. J Virol. August 1993;67(8):5062–7.
68. Klagges BR, Heimbeck G, Godenschwege TA, Hofbauer A, Pflugfelder GO, Reifegerste R, u. a. Invertebrate synapsins: a single gene codes for several isoforms in Drosophila. J Neurosci. 15. Mai 1996;16(10):3154–65.

69. Loughran G, Chou M-Y, Ivanov IP, Jungreis I, Kellis M, Kiran AM, u. a. Evidence of efficient stop codon readthrough in four mammalian genes. Nucleic Acids Res. 18. August 2014;42(14):8928–38.
70. Eswarappa SM, Potdar AA, Koch WJ, Fan Y, Vasu K, Lindner D, u. a. Programmed Translational Readthrough Generates Anti-Angiogenic VEGF-Ax. Cell. 19. Juni 2014;157(7):1605–18.
71. Namy O, Duchateau-Nguyen G, Hatin I, Hermann-Le Denmat S, Termier M, Rousset J-P. Identification of stop codon readthrough genes in Saccharomyces cerevisiae. Nucleic Acids Res. 1. Mai 2003;31(9):2289–96.
72. Namy O, Duchateau-Nguyen G, Rousset J-P. Translational readthrough of the PDE2 stop codon modulates cAMP levels in Saccharomyces cerevisiae. Mol Microbiol. Februar 2002;43(3):641–52.
73. Jalajakumari MB, Thomas CJ, Halter R, Manning PA. Genes for biosynthesis and assembly of CS3 pili of CFA/II enterotoxigenic Escherichia coli: novel regulation of pilus production by bypassing an amber codon. Mol Microbiol. Dezember 1989;3(12):1685–95.
74. Wenthzel A-MK, Stancek M, Isaksson LA. Growth phase dependent stop codon readthrough and shift of translation reading frame in Escherichia coli. FEBS Letters. 16. Januar 1998;421(3):237–42.
75. Drugeon G, Jean-Jean O, Frolova L, Le Goff X, Philippe M, Kisselev L, u. a. Eukaryotic release factor 1 (eRF1) abolishes readthrough and competes with suppressor tRNAs at all three termination codons in messenger RNA. Nucleic Acids Res. 15. Juni 1997;25(12):2254–8.
76. Wang L, Xie J, Schultz PG. Expanding the genetic code. Annu Rev Biophys Biomol Struct. 2006;35:225–49.
77. Reynolds NM, Lazazzera BA, Ibba M. Cellular mechanisms that control mistranslation. Nat Rev Microbiol. Dezember 2010;8(12):849–56.
78. Wang L, Schultz PG. A general approach for the generation of orthogonal tRNAs. Chemistry & Biology. September 2001;8(9):883–90.
79. Guo J, Melançon CE, Lee HS, Groff D, Schultz PG. Evolution of Amber Suppressor tRNAs for Efficient Bacterial Production of Unnatural Amino Acid-Containing Proteins. Angew Chem Int Ed Engl. 2009;48(48):9148–51.
80. Wang N, Ju T, Niu W, Guo J. Fine-tuning interaction between aminoacyl-tRNA synthetase and tRNA for efficient synthesis of proteins containing unnatural amino acids. ACS Synth Biol. 20. März 2015;4(3):207–12.
81. Young TS, Schultz PG. Beyond the Canonical 20 Amino Acids: Expanding the Genetic Lexicon. J Biol Chem. 9. April 2010;285(15):11039–44.
82. Kuhn SM, Rubini M, Fuhrmann M, Theobald I, Skerra A. Engineering of an orthogonal aminoacyl-tRNA synthetase for efficient incorporation of the non-natural amino acid O-methyl-L-tyrosine using fluorescence-based bacterial cell sorting. J Mol Biol. 19. November 2010;404(1):70–87.
83. Santoro SW, Wang L, Herberich B, King DS, Schultz PG. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. Nat Biotechnol. Oktober 2002;20(10):1044–8.
84. Blattner FR. The Complete Genome Sequence of Escherichia coli K-12. Science. 5. September 1997;277(5331):1453–62.
85. Shine J, Dalgarno L. The 3’-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA. April 1974;71(4):1342–6.
86. Kozak M. Initiation of translation in prokaryotes and eukaryotes. Gene. 8. Juli 1999;234(2):187–208.
87. Cooper GM. Translation of mRNA. 2000 [zitiert 3. Oktober 2017]; Verfügbar unter: https://www.ncbi.nlm.nih.gov/books/NBK9849/
88. Laursen BS, Sørensen HP, Mortensen KK, Sperling-Petersen HU. Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev. März 2005;69(1):101–23.
89. Gualerzi CO, Pon CL. Initiation of mRNA translation in bacteria: structural and dynamic aspects. Cell Mol Life Sci. 2015;72:4341–67.
90. Chen H, Bjerknes M, Kumar R, Jay E. Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of Escherichia coli mRNAs. Nucleic Acids Res. 25. November 1994;22(23):4953–7.
91. Wen J-D, Lancaster L, Hodges C, Zeri A-C, Yoshimura SH, Noller HF, u. a. Following translation by single ribosomes one codon at a time. Nature. 3. April 2008;452(7187):598–603.

92. Fluman N, Navon S, Bibi E, Pilpel Y. mRNA-programmed translation pauses in the targeting of E. coli membrane proteins. eLife [Internet]. 3. Verfügbar unter: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4359368/
93. Li G-W, Oh E, Weissman JS. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature. 28. März 2012;484(7395):538–41.
94. Malys N, McCarthy JEG. Translation initiation: variations in the mechanism can be anticipated. Cell Mol Life Sci. März 2011;68(6):991–1003.
95. Chang B, Halgamuge S, Tang S-L. Analysis of SD sequences in completed microbial genomes: non-SD-led genes are as common as SD-led genes. Gene. 24. Mai 2006;373:90–9.
96. Moll I, Grill S, Gualerzi CO, Bläsi U. Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. Mol Microbiol. Januar 2002;43(1):239–46.
97. Boni IV, Isaeva DM, Musychenko ML, Tzareva NV. Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. Nucleic Acids Res. 11. Januar 1991;19(1):155–62.
98. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene. 15. April 1989;77(1):51–9.
99. Liu H, Naismith JH. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. BMC Biotechnol. 4. Dezember 2008;8:91.
100. Zhang G, Annan RS, Carr SA, Neubert TA. Overview of peptide and protein analysis by mass spectrometry. Curr Protoc Protein Sci. November 2010;Chapter 16:Unit16.1.
101. Smolskaya S, Andreev YA. Site-Specific Incorporation of Unnatural Amino Acids into Escherichia coli Recombinant Protein: Methodology Development and Recent Achievement. Biomolecules [Internet]. 28. Juni 2019;9(7). Verfügbar unter: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6681230/
102. Poole E, Tate W. Release factors and their role as decoding proteins: specificity and fidelity for termination of protein synthesis. Biochim Biophys Acta. 7. September 2000;1493(1–2):1–11.
103. Johnson DBF, Xu J, Shen Z, Takimoto JK, Schultz MD, Schmitz RJ, u. a. RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. Nat Chem Biol. 18. September 2011;7(11):779–86.
104. Mukai T, Hayashi A, Iraha F, Sato A, Ohtake K, Yokoyama S, u. a. Codon reassignment in the Escherichia coli genetic code. Nucleic Acids Res. Dezember 2010;38(22):8188–95.
105. Mukai T, Hoshi H, Ohtake K, Takahashi M, Yamaguchi A, Hayashi A, u. a. Highly reproductive Escherichia coli cells with no specific assignment to the UAG codon. Sci Rep [Internet]. 18. Mai 2015;5. Verfügbar unter: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4434889/
106. Scolnick E, Tompkins R, Caskey T, Nirenberg M. Release factors differing in specificity for terminator codons. Proc Natl Acad Sci USA. Oktober 1968;61(2):768–74.
107. Smolskaya S, Zhang ZJ, Alfonta L. Enhanced yield of recombinant proteins with site-specifically incorporated unnatural amino acids using a cell-free expression system. PLoS ONE. 2013;8(7):e68363.
108. Goerke AR, Swartz JR. High-level cell-free synthesis yields of proteins containing site-specific non-natural amino acids. Biotechnol Bioeng. 1. Februar 2009;102(2):400–16.
109. Wang K, Neumann H, Peak-Chew SY, Chin JW. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. Nat Biotechnol. Juli 2007;25(7):770–7.

Eidesstattliche Erklärung

Ich, Thewes, Miriam, geboren am 07.11.1991 in München, erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.
Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Insbesondere habe ich nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen.
Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Puchheim, den 01.11.2020

Miriam Thewes

Danksagung

Ich möchte mich sehr herzlich bei meinem Doktorvater, Prof. Dr. Ralf Wagner für die Vergabe des spannenden Themas und die hervorragende Betreuung während meiner Doktorarbeit bedanken. Danke für deinen Input, Motivation und Geduld über die gesamte Zeit.
Großer Dank gilt auch Prof. Dr. Bernd Salzberger für die Übernahme des Zweitgutachtens.
Vielen Dank an Dr. David Peterhoff für die Anleitung und Begleitung meiner Arbeit im Labor, die vielen konstruktiven Diskussionen und Hilfestellungen. Danke, dass du immer ein offenes Ohr für mich hattest.
An dieser Stelle sei auch Malin Zaddach für die Hilfe bei der Durchführung der Experimente gedankt.
Bei Prof. Dr. Nediljko Budisa und Dr. Tobias Baumann bedanke ich mich für das Überlassen der Suppressorpaare pTB77 und pTB290 und den wissenschaftlichen Austausch.

Danke an die "Wagners" für die herzliche Aufnahme in die AG, die vielen kleinen und großen Hilfen im Labor, aber auch für die schöne und unvergessliche Zeit zusammen, im Labor, beim Kaffeetrinken, Volleyballspielen, Grillen oder gemeinsamen Abenden in Regensburg.

Grazie di cuore an meine Eltern Franca und Roland und meine Schwester Luisa. Danke, dass ihr mich auf meinem Weg immer unterstützt habt, insbesondere auch während des gesamten Studiums und meiner Doktorarbeit.
Danke an Richie, für deine moralische Unterstützung während der Examensphase und Doktorarbeit. Ich schätze mich glücklich, dich an meiner Seite zu haben.