Co- and posttranslational engineering of the therapeutic glycoprotein erythropoietin with unnatural amino acids

Dissertation

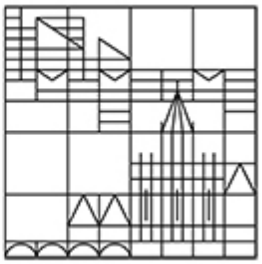
submitted for the degree of Doctor of Natural Sciences

Presented by

Katharina Streichert

at the

Universität Konstanz



Faculty of Sciences

Department of Chemistry

Konstanz, 2016

Date of the oral examination: 13th of July 2016

First supervisor: Dr. Marina Rubini

Second supervisor: Prof. Dr. Jörg Hartig

Abstract

Erythropoietin (EPO) is the most important drug to treat different types of anaemia arising from chronic kidney disease, cancer or AIDS. Four of the top ten biotech drugs are a form of recombinant EPO and the market for EPO is steadily increasing. EPO contains three *N*-linked and one *O*-linked glycosylations at the positions N24, N38, N83 and S126, which account for 40% of the molecular weight. Glycosylation is very important for the protein stability against thermal and proteolytic degradation. This posttranslational modification enhances the circulatory half-life and therefore the *in vivo* activity of the protein.

Until now, therapeutic glycoproteins are produced in eukaryotic cells resulting in a heterogeneous glycosylation pattern. This structural heterogeneity of *N*-linked oligosaccharides encumbers the correlation of glycan structure with glycoprotein function.

In addition, therapeutic proteins are often PEGylated, as this modification increases the solubility, protects from proteolytic degradation and enhances the circulatory half-life. However, PEGylation has the big disadvantage of being unspecific or even random.

Herein, a semi-synthetic approach is presented for the production of EPO with well-defined and uniform glycan structures at specific positions. As bacterial cells do not posttranslationally glycosylate proteins, *E. coli* cells were used to incorporate non-natural amino acids into EPO by amber stop codon suppression methodology. These non-natural amino acids bear a specific bio-orthogonal chemical function. For example, the pyrrolysine derivative Plk has an alkyne group. After expression and purification, EPO could be coupled at the natural glycosylation sites to defined, synthesized oligosaccharides or purchased PEG chains by copper-catalysed 1,3-dipolar Huisgen cycloaddition between alkynes and azides.

Each individual step from synthesizing the unnatural amino acids, amber stop codon suppression in *E. coli*, purification, refolding to click chemistry had to be optimised in order to produce homogeneously glycosylated or PEGylated EPO in high yields and in a convenient, low-cost manner.

The generated EPO variants were characterised biophysically by mass spectrometry and circular dichroism (CD). All of them displayed the correct molecular mass and the secondary structure of EPO is not disturbed by incorporated unnatural amino acids or coupled decorations. Moreover, CD spectra measurements showed a large protective effect of one short glycan (under 2 kDa) or of a 5kDa-PEG on the secondary structure of EPO upon repeated freezing-thawing cycles.

The biological activity was investigated by cell differentiation and proliferation assays with different cell lines. It could be shown that all EPO variants have a positive effect on cell differentiation of haematopoietic stem cells from mouse bone marrow. Results that are more detailed were found in cell proliferation assays. In contrast to small PEG-chains, even one coupled glycan with a molecular weight of 1.7 kDa is sufficient to decrease the EC_{50}-value, and consequently, to increase the biological activity of EPO. Two terminal sialic acids further augment this effect.

To sum it up, a new procedure for engineering glycoproteins, such as EPO, is presented, which combines non-canonical amino acids and click chemistry. With this new methodology in hand, first

steps were undertaken towards elucidating the impact of each glycosylation position and pattern on the function of EPO, as an example for therapeutic glycoproteins.

Zusammenfassung

Erythropoetin (EPO) ist das wichtigste Arzneimittel, um verschiedenartige Anämien zu behandeln, die zum Beispiel mit chronischem Nierenversagen, Krebs oder AIDS einhergehen. Von den zehn meist verkauften biotechnologischen Medikamenten, sind vier rekombinantes EPO. Der Markt für EPO expandiert Zusehens. EPO enthält drei *N*- und eine *O*-Glykosylierung an den Positionen N24, N38, N83 und S126, welche zusammen 40% der Molekülmasse ausmachen. Die Glykosylierung von Proteinen spielt eine wichtige Rolle für die thermische und proteolytische Proteinstabilität. Diese posttranslationale Modifikation erhöht die Halbwertszeit im Blutkreislauf und damit auch die *in vivo* Aktivität des Proteins.

Bis jetzt werden therapeutische Glykoproteine in eukaryotischen Zellen produziert, was zu heterogenen Glykosylierungsmustern führt. Diese strukturelle Heterogenität der *N*-Glykosylierungen verhindert die Korrelation der Zuckerstruktur mit der Glykoproteinfunktion.

Zusätzlich zu Glykosylierungen werden therapeutische Proteine häufig PEGyliert, weil dies die Löslichkeit erhöht, ebenfalls vor proteolytischen Abbau schützt, sowie auch die Halbwertszeit im Blutkreislauf erhöht. Der größte Nachteil dieser Methode ist jedoch, dass sie unspezifisch oder sogar an wahllosen Positionen erfolgt.

Diese Arbeit präsentiert eine halbsynthetische Vorgehensweise, wie EPO mit definierten und gleichförmigen Zuckerketten an definierten Positionen hergestellt werden kann. Da Bakterien Proteine nicht posttranslational modifizieren, wurde *E. coli* verwendet, um nicht-kanonische Aminosäuren mittels Amber-Suppression in EPO einzubauen. Diese nicht-kanonische Aminosäuren besitzen spezielle, bio-orthogonale chemische Funktionen. Zum Beispiel trägt das Pyrrolysinderivat Plk eine Alkylgruppe. Nach erfolgter Expression und Reinigung konnte EPO an den natürlichen Glykosylierungsstellen mit definiert synthetisierten Zuckerketten oder mit käuflich erworbenen PEG-Ketten gekoppelt werden. Hierfür wurde die Kupfer-katalysierte 1,3-dipolare Huisgen Cycloaddition zwischen Alkinen und Aziden verwendet.

Jeder einzelne Schritt, von der Synthese nicht-natürlicher Aminosäuren, Amber-Suppression in *E. coli*, Reinigung, Rückfaltung bis zur Click-Chemie, musste optimiert werden, um homogen glykosyliertes oder PEGyliertes EPO in ausreichenden Mengen, kostengünstig und praktisch anwendbar zu produzieren.

Die hergestellten EPO-Varianten wurden biophysikalisch durch Massenspektrometrie und Circulardichroismus (CD) charakterisiert. Alle Varianten zeigten die korrekte Molekülmasse auf. Außerdem wird die Sekundärstruktur von EPO durch die eingeführten, nicht-kanonischen Aminosäuren oder durch die gekoppelten Modifizierungen nicht beeinflusst. Des Weiteren konnte durch CD-Messungen festgestellt werden, dass bereits eine kleine Zuckerkette mit einem

Molekulargewicht von unter 2 kDa oder eine 5 kDa-PEG-Kette einen deutlich schützenden Effekt auf die Sekundärstruktur von EPO bei wiederholten Zyklen des Einfrierens und Auftauens haben.

Die biologische Aktivität wurde durch Zelldifferenzierungs- und Proliferationsassays mit unterschiedlichen Zelllinien untersucht. Es konnte gezeigt werden, dass alle EPO-Varianten einen positiven Einfluss auf die Differenzierung hämatopoetischer Stammzellen aus Knochenmark der Maus nehmen. Detailliertere Ergebnisse wurden bei den Zellproliferationsassays erreicht. Im Gegensatz zu kleinen PEG-Ketten, reicht nur eine gekoppelte Zuckerkette mit einem Molekulargewicht von 1,7 kDa aus, um den EC_{50}-Wert zu verkleinern und somit, die biologische Aktivität von EPO zu erhöhen. Zwei hinzugefügte Sialinsäuren am Ende der Zuckerkette verstärken diesen Effekt.

Zusammenfassend wird eine neue Vorgehensweise vorgestellt, welche nicht-natürliche Aminosäuren mit Click-Chemie kombiniert. Mit dieser Methode kann man Glykoproteine, wie EPO, gezielt nach seinen Wünschen entwickeln. Erste Schritte wurden unternommen, um den spezifischen Einfluss der Glykosylierungsstelle und –struktur auf die Funktion von EPO zu entziffern, als Beispiel eines therapeutischen Glykoproteins.

Table of Content

1. Introduction

1.1. The genetic code

1.1.1. The central dogma of molecular biology

Soon after Crick and Watson postulated the three-dimensional structure of the DNA, Crick wondered about the functional relation of DNA and proteins. His thoughts ended up to the central dogma of molecular biology (figure 1). It says that the flow of information is allowed only from nucleic acids to nucleic acids and from nucleic acids to proteins (Crick 1970).

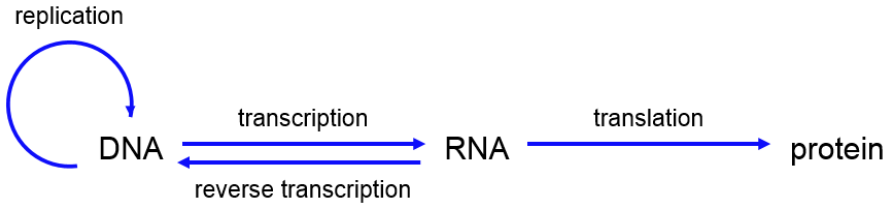


Figure 1: The central dogma of molecular biology postulated by (Crick 1970).

1.1.2. The genetic code

After the elucidation of the triplet genetic code (figure 2), it became clear how the information for amino acid sequences was contained in nucleic acid sequences. The fact that 20 canonical amino acids are assigned by 61 coding triplet combinations and three termination signals is nowadays considered as basic molecular biological knowledge. The first experiment in order to decipher the genetic code was performed by Matthaei and Nirenberg in 1961 (Matthaei and Nirenberg 1961). They used an artificial mRNA, poly-uracil, and a bacterial extract with ribosomes and a mixture of all aminoacyl-tRNAs. Their finding was that the only possible amino acid assigned to UUU is phenylalanine (Budisa 2004).

The genetic code is degenerated because two to four triplets are assigned to one amino acid. Leucine, arginine and serine are even represented by six different triplets. Only tryptophan and methionine have one specific triplet. Triplets that code for the same amino acid, so-called synonymous triplets, are often similar to each other. For example, all four glycine triplets start with GG. More precisely, two triplets XYC and XYU are always assigned to the same amino acid (Knippers 2006).

The genetic code used by all known forms of life is nearly universal with minor variations. Examples for minor variations are mitochondrial DNA and DNA from simple organisms as *mycoplasma* and *paramecium*. The appearance of unusual codon usage is interpreted as a sign for an early separation of the respective genomes from the main branch of evolution (Knippers 2006).

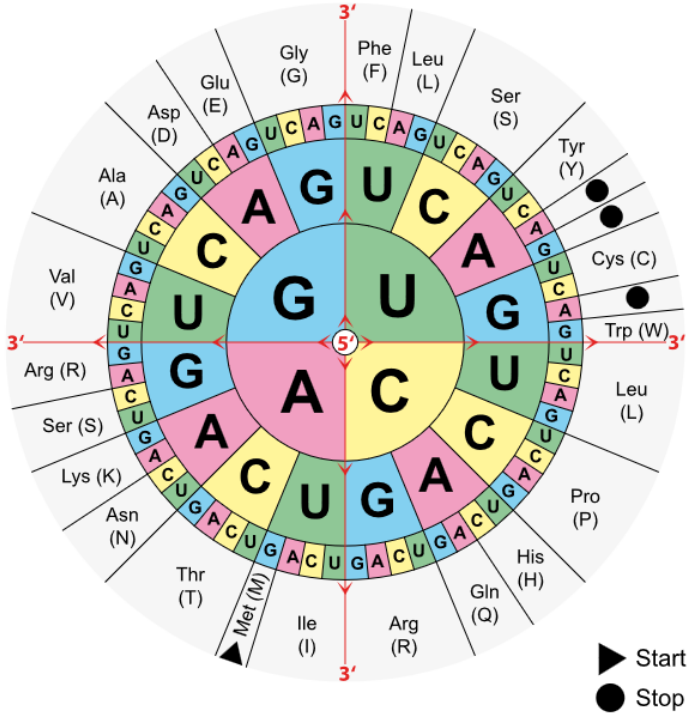


Figure 2: The genetic code – 61 triplets encode twenty amino acids; three codons (UAG, UAA and UGA) represent termination signals

1.1.3. Further natural amino acids: selenocysteine and pyrrolysine

In a few enzymes of bacteria and eukaryotes, the unusual amino acid selenocysteine, often called the 21st amino acid, is present (figure 3A). Selen has similar properties as sulphur, but is much more reactive. This higher reactivity is mandatory for the enzyme functionality. Examples are the formate-dehydrogenase in *E. coli* and the glutathione peroxidase in mammals. Selenocysteine is incorporated in response to the codon UGA, which was defined as stop codon in the previous chapter (1.1.2). This is consistent to the universality of the genetic code because of the special structure of the selenocysteyl-tRNA. First, serine is loaded onto this tRNA and in serial reactions, the OH-group is exchanged by a selenol group. A special translation factor is needed at the ribosome to incorporate the amino acid in response to the UGA codon on the mRNA. This does not occur, if UGA is used as stop codon at the end of the gene. Consequently, the selenocysteyl-tRNA in the ternary complex with its own translation factor and GTP has to recognise the UGA codon, but also the nucleotide sequence and the secondary structure in its vicinity (Knippers 2006).

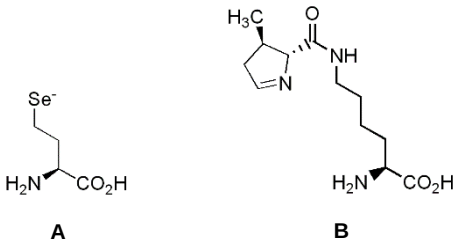


Figure 3: Structures of the 21st amino acid selenocysteine (A) and the 22nd amino acid pyrrolysine (B)

In 2002, the 22nd amino acid pyrrolysine (figure 3B), a derivative of lysine, was found to be encoded by the UAG amber codon, present in the gene of monomethylamine methyltransferase of *Methanosarcina barkeri*, an anaerobe that has been isolated from mud samples in lakes and bogs (Hao, Gong et al. 2002, Srinivasan, James et al. 2002). This is a special case because UAG is normally used as a stop codon. Pyrrolysine behaves like a typical canonical amino acid. It is directly charged onto its cognate $tRNA^{Pyl}$ with the CUA anticodon by its own pyrrolysyl-tRNA synthetase (PylRS) (Blight, Larue et al. 2004, Polycarpo, Ambrogelly et al. 2004). Its presence is not very widespread. Pyrrolysine was found in limited numbers of organisms including some other members of the *Methanosarcinaceae* family and several bacteria (Herring, Ambrogelly et al. 2007, Fekner and Chan 2011).

1.2. The ribosome-mediated protein synthesis

Protein synthesis is performed by ribosomes in all organisms. These translate the genetic information, which was previously transcribed from DNA to mRNA, into the amino acid sequence of a protein. The translation can be divided into two crucial recognition events: the codon-anticodon interaction between tRNA and mRNA on the ribosome and the amino acid aminoacylation by specific aminoacyl-tRNA synthetases (Budisa 2004).

All ribosomes consist out of two unequal subunits, in prokaryotes, they are designated as 50S and 30S subunits. Each subunit is built out of one to three RNA molecules and various proteins. 60% is RNA and 40% consists out of protein. At the 30S subunit, the mRNA encounters the tRNA and the 50S subunit mediates the amino acid linkages. Translation can be divided into three steps: initiation, elongation – translocation and termination (Knippers 2006).

1.2.1. The tRNA and its aminoacylation

tRNAs consist out of 74 to 94 ribonucleotides (figure 4) and the properties that have all tRNAs in common are the following (Knippers 2006):

- The acceptor stem with seven to nine base pairs is made by the base pairing of the 5'-terminal nucleotide with the 3'-terminal nucleotide.
- The 3'-end has a CCA-tail. The amino acid is linked to the 3'-hydrosyl group of the terminal adenine to form the aminoacyl-tRNA.
- The anticodon is located in the middle of a seven-nucleotide loop in the anticodon arm.
- The D arm often contains dihydrouridine.
- The TΨC arm always contains the sequence thymidine, pseudouridine (modified uridine), cytosine (TΨC).
- Between anticodon loop and TΨC loop, a variable loop is located, which has a variable length for different tRNAs.

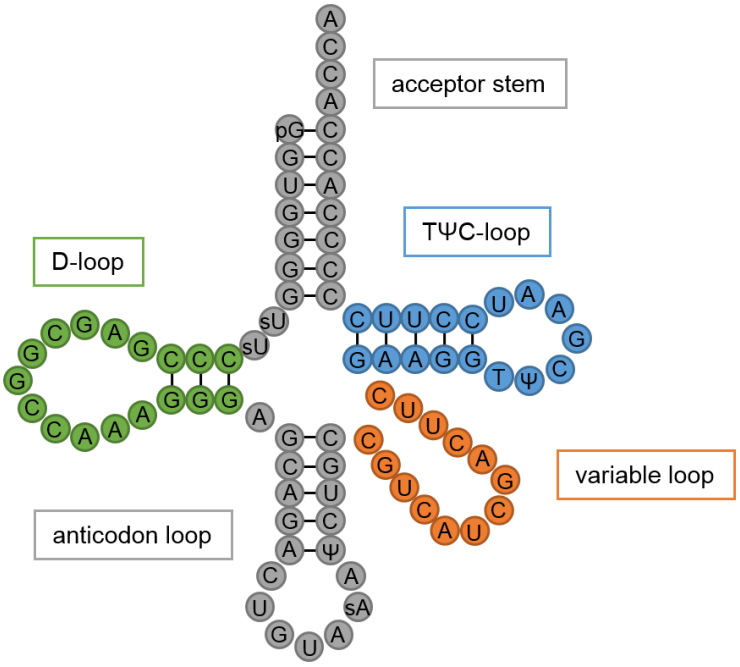


Figure 4: Secondary structure of $tRNA^{Tyr}$ from *E. coli* as an example; figure is inspired by (Knippers 2006)

Aminoacyl tRNA synthetases are enzymes that load the respective amino acid onto the tRNA. There are at least twenty different aminoacyl tRNA synthetases in each cell, one for each amino acid. They aminoacylate tRNAs in two steps:

- ATP is cleaved and aminoacyl-AMP is formed.
- Aminoacyl-AMP bound to the synthetase reacts with tRNA, an aminoacyl-ester bond is formed between the carboxyl group of the amino acid and either the 2'- or the 3'-OH of the last tRNA nucleotide at the 3'-end, meanwhile AMP is released.

There are two classes of aminoacyl tRNA synthetases: Class I contains the ubiquitous Rossmann fold at the active centre, which is composed of up to seven mostly parallel β-strands. The first two strands are connected by an α-helix. This protein structural motif is found in all proteins that bind nucleotides. Class I synthetases aminoacylate at the 2'-OH of the terminal adenosine nucleotide on tRNA. Class II synthetases have a distinct active centre and aminoacylate at the 3'-OH of the terminal adenosine on tRNA. Synthetases also have a proofreading function to ensure the high fidelity of tRNA charging. The aminoacyl-tRNA bond can be hydrolysed or aminoacyl-AMP can be cleaved through a weak esterase activity (Knippers 2006).

1.2.2. Initiation

The methionine codon AUG is the start codon for nearly all open reading frames in bacteria and eukaryotes. However, a 5'-non-coding region of about 4-14 nucleotides before the start codon, also called Shine-Dalgarno sequence, is important for the initiation. This region builds up base pairs with complementary sequences at the 3'-end of the 16S rRNA of the ribosome. The length of the Shine-

Dalgarno sequence and its distance to the AUG start codon determine the stability of the initiator complex (Shine and Dalgarno 1974).

The initiator tRNA, which is loaded with formyl methionine, is responsible for the exact start of protein synthesis because the 5'-nucleotide from the acceptor stem does not form base pairs, the anticodon loop contains three GC base pairs and the adenine at position 37 is not modified. These characteristics may provide a certain flexibility in codon-anticodon binding and an own type of ribosome binding for the formyl-methionine-tRNA. The formyl residue occupies the amino group of the first amino acid and hence, defines the synthesis direction. During protein synthesis, polypeptide deformylases remove the formyl residue at the N-terminus and methionine aminopeptidases cut off the start methionine from nascent polypeptide chains. Moreover, three initiation factors (IF1-3) are needed. These factors encounter the mRNA at the 30S subunit. IF1 activates IF2 and IF3 and keeps the ribosome subunits separately. IF3 suppresses non-initiator tRNAs from the ribosome and increases the binding specificity of formyl-methionine-tRNA. IF2 is activated by binding of GTP. Then, it binds and translocates formyl-methionine-tRNA to the P-site under GTPase activity. Lastly, IF1 and IF3 leave the ribosome. The 70S subunit can bind and IF2-GDP is released (Knippers 2006).

1.2.3. Elongation and translocation

The ribosome has two different binding pockets for loaded tRNAs, the aminoacyl site (A-site), which recognise the incoming tRNA with the complementary codon matching the mRNA codon, and the peptidyl site (P-site), where the peptide bond is formed (figure 5). First, the initiator tRNA is located at the P-site and the next triplet on the mRNA at the A-site. The ternary complex arrives at the A-site. It consists out of the respective aminoacyl-tRNA bound to the elongation factor EF-Tu, which is activated through binding to GTP. GTP is converted to GDP, EF-Tu/GDP leaves the ribosome, *N*-formyl-Met is removed from the tRNA and a new peptide bound is formed between its carboxyl group and the amino group of the next amino acid. Temporarily, a dipeptidyl-tRNA is now located at the A-site. Then, translocation takes place, for which EF-G is needed (also activated through binding to GTP). The ribosome shifts one triplet length on the mRNA, whereby GTP is converted to GDP. Lastly, the peptidyl-tRNA is at the P-site, the empty tRNA leaves the ribosome through the E-site and a new cycle can start at the empty A-site (Knippers 2006).

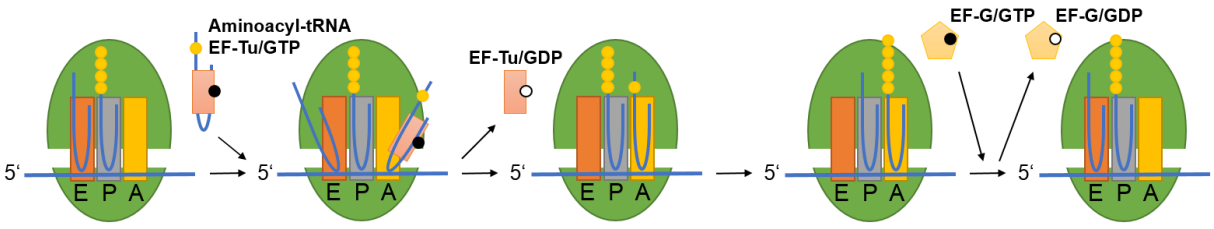


Figure 5: Schematic overview of the elongation and translocation steps during translation; figure is inspired by (Knippers 2006)

1.2.4. Termination

One of the three stop codons determines the termination of translation. Because there are not any tRNAs with complemental anticodons to stop codons, the ribosome stops at a stop codon. Then, release factors RF1 and RF2 place themselves at the A-site and RF3 removes RF1 or RF2 from the ribosome. EF-G together with the ribosome-recycling factor mediates the separation of the two ribosomal subunits and the release of mRNA and tRNA (Knippers 2006).

1.3. Protein engineering with non-natural amino acids

1.3.1. Selective pressure incorporation

Selective pressure incorporation (SPI) is based on the use of auxotrophic strains. This means that the cells are not able to biosynthesise one or more canonical amino acid themselves. Its roots lie in the classical experiment of (Cowie and Cohen 1957). They reported the incorporation of seleno-methionine into the whole proteome, using an *E. coli* methionine-auxotroph mutant strain. The bacterial growth rate was dependent on the external methionine supply. Therefore, it was possible to replace methionine by selenomethionine. In such cultures, the cells grew more slowly but exponential. Selenomethionine was found to completely and uniformly substitute methionine in all cellular proteins and thus, an “unnatural microorganism” was obtained.

With some exceptions as selenomethionine, all non-canonical amino acids that are not metabolic intermediates are toxic. However, it was observed that toxic analogues might serve as substrates in protein synthesis. If such toxic analogues are added together with their canonical counterparts in the growth media, usually lower incorporation levels in all cellular proteins are obtained. For substitutions in single target proteins, this is a major problem to overcome in order to achieve full substitution. The use of auxotrophic strains provided a solution to circumvent toxic metabolic effects (figure 6). However, the auxotrophic approach for complete substitution of target proteins could be fully generalised to a single target protein only after the introduction of recombinant DNA techniques. The basic requirements for a successful SPI-experiment include:

- Selection of a proper cell and expression system
- Control of fermentation conditions (for example the environment)
- Selective pressure for the replacement of the amino acid (for example the reassignment of a sense-codon in a single protein)

The amino acid analogues need to be sterically almost identical to the canonical ones and are called isosteres or surrogates. They have to fulfil three conditions:

- The uptake of the non-canonical amino acid
- Its attachment onto the tRNA
- Its incorporation into the nascent polypeptide chain

In such approaches, the amino acid, which the cells cannot produce themselves, is supplied in restricted amounts for cellular growth. As the stationary phase is reached, the culture is transferred into another minimal medium depleted from the parental amino acid and with a high concentration

of the unnatural amino acid analogue. From this point on, the host cells serve only as a “factory” to produce the recombinant protein. In that way, the cells are forced to incorporate the unnatural amino acid instead of the missing canonical amino acid due to the lack of an absolute substrate specificity of the aminoacyl-tRNA synthetase. Thus, the toxicity can be circumvented in this straightforward way. An alternative would be to block biosynthetic pathways of the host cells by proper inhibitors (Budisa 2004).

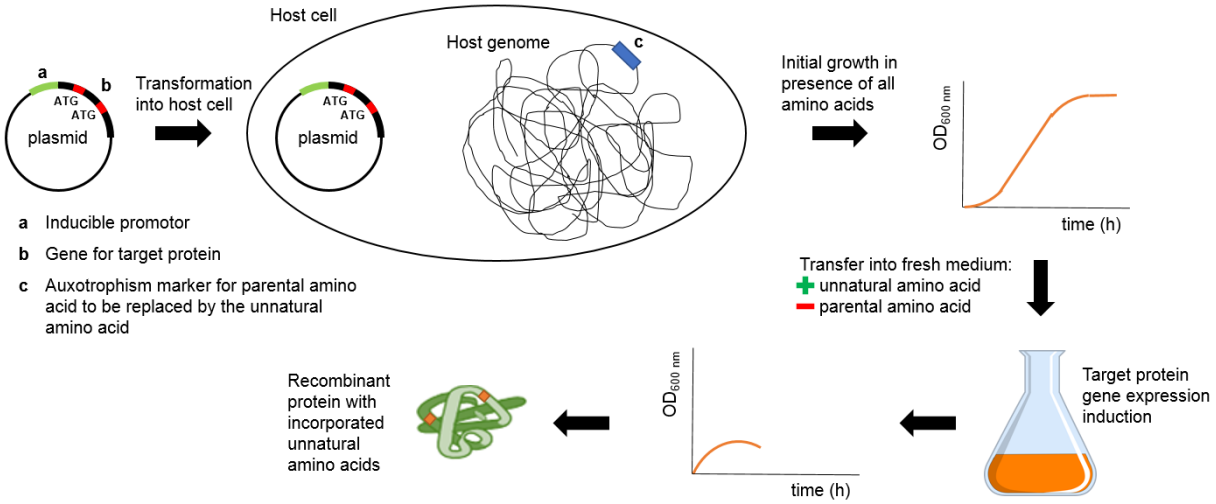


Figure 6: Schematic overview of the selective pressure incorporation method; strong host auxotrophism and control of the fermentation conditions are crucial for effective unnatural amino acid incorporation. Figure is inspired by (Budisa and Biava 2014).

Azidohomoalanine (AHA) and homopropargylglycine (HPG) are two examples for methionine analogues, which can be introduced into proteins via SPI (figure 7). Bertozzi et al. have successfully demonstrated the incorporation of these two unnatural amino acids into the protein murine dihydrofolate reductase using methionine auxotrophic *E. coli* (Kiick, Saxon et al. 2002).

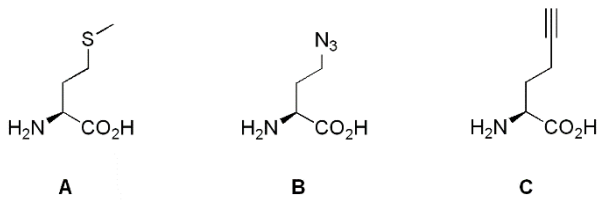


Figure 7: Structure of methionine (A) and its analogues azidohomoalanine (B) and homopropargylglycine (C)

1.3.2. (Amber) stop codon suppression

Another possibility to introduce unnatural amino acids into proteins is by stop codon suppression. Therefore, a tRNA/aminoacyl-tRNA synthetase pair from another organism is needed, which recognises one of the three stop codons. This pair needs to be orthogonal to the host organism, which means that there are not any cross-reactions: The unnatural amino acid is not recognised by endogenous aminoacyl-tRNA synthetases, nor the orthogonal synthetase recognises one of the

canonical amino acids. Moreover, the stop codon is distinctly assigned to the unnatural amino acid. One of the remaining two stop codons must then serve as stop signal.

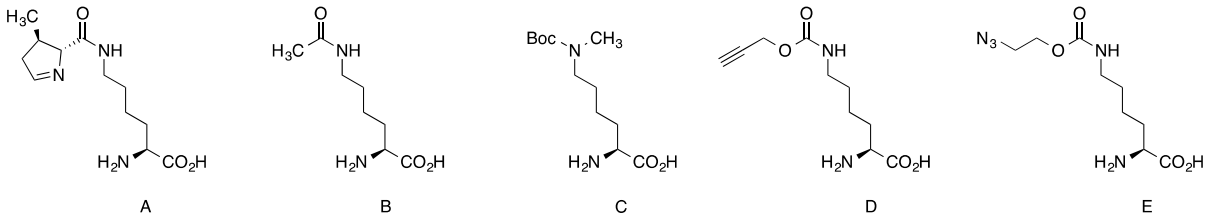


Figure 8: Structure of pyrrolysine (A) and its derivatives (B-E); Plk (D), Pln (E)

In this study, the $tRNA_{CUA}$/pylRS pair from *Methanosarcina barkeri* was used, which recognises the amber stop codon UAG on the mRNA and assigns pyrrolysine to it. By introducing its genes into *E. coli*, it is possible to incorporate pyrrolysine into any recombinant protein opposite an amber stop codon (Blight, Larue et al. 2004). This is also possible for the structural similar derivatives of pyrrolysine (figure 8). A schematic overview of the strategy to incorporate the pyrrolysine derivative Plk (figure 8D) into a target protein, is depicted in figure 9.

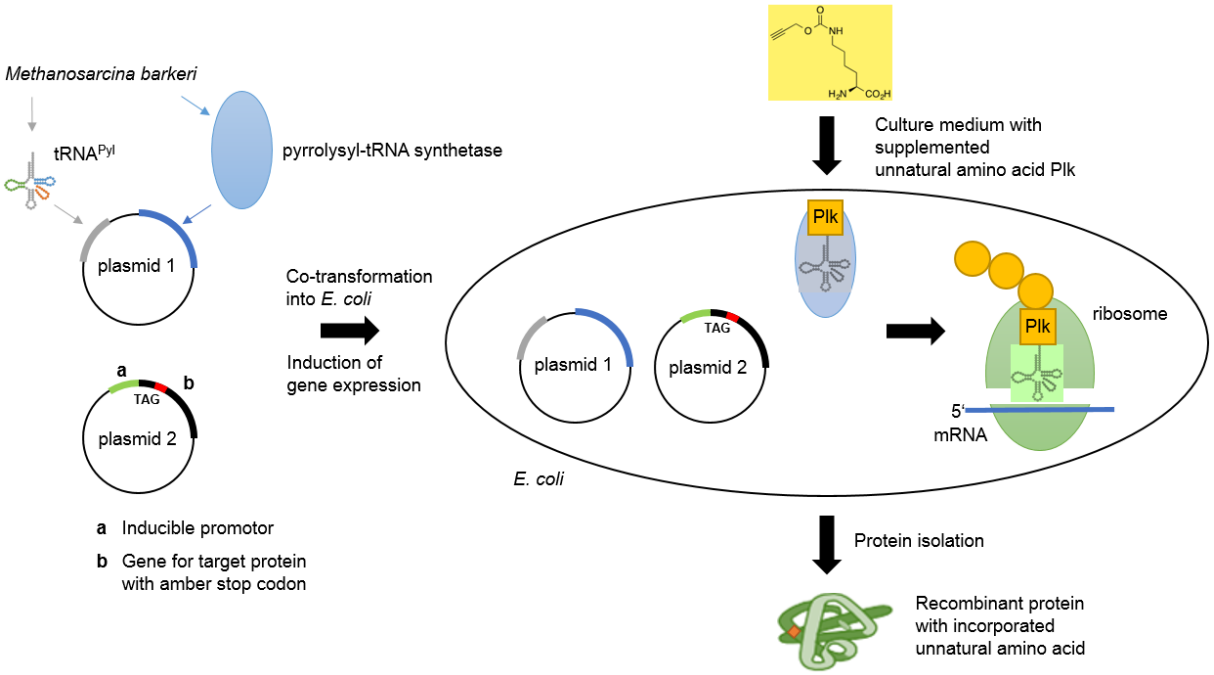


Figure 9: Schematic overview of Plk incorporation into a target protein as an example for amber suppression

An advantage of amber suppression over SPI is that the incorporation of non-canonical amino acids is exclusively opposite to the amber stop codon. Hence, canonical amino acids in the proteome are not replaced. The drawbacks of this method compared to SPI are the possible truncation of the target protein, due to a stop of the translational machinery at the amber stop codon and a tremendously decreased yield, if more than one amber stop codon is used within one target protein.

1.4. Click chemistry

Sharpless et al. defined the term “click chemistry” in 2001. It is the superordinate concept of all reactions that quickly and reliably join two functional groups together with respect to the following criteria. A desirable click reaction would be modular with broad application possibilities. The stereospecific reaction would result in high yields with only non-hazardous by-products. Moreover, the reaction conditions should be simple; preferably, the reaction would be possible in water without any solvent. A large thermodynamic driving force that favours a reaction with a single reaction product would be desirable (Kolb, Finn et al. 2001).

The most popular reaction that fits this concept best is the azide-alkyne Huisgen cycloaddition. Huisgen was the first to understand the scope of the 1,3-dipolar cycloaddition (Huisgen 1984). However, the enormous drawback in this time was the fact that an elevated temperature (over 100 °C) was needed, and still, it takes hours to days to react. Under these conditions, it was not possible to use the reaction for biomolecules. This problem has been solved when Cu(I) was used as a catalyst. The copper-catalysed cycloaddition was reported in 2002 simultaneously by two independent groups (Rostovtsev, Green et al. 2002, Tornoe, Christensen et al. 2002). It transforms azides and terminal alkynes exclusively into the corresponding 1,4-disubstituted 1,2,3-triazoles, in contrast to the uncatalysed reaction, which provides mixtures of 1,4- and 1,5-triazole regioisomers (Hein and Fokin 2010). A proposed mechanism is depicted in figure 10.

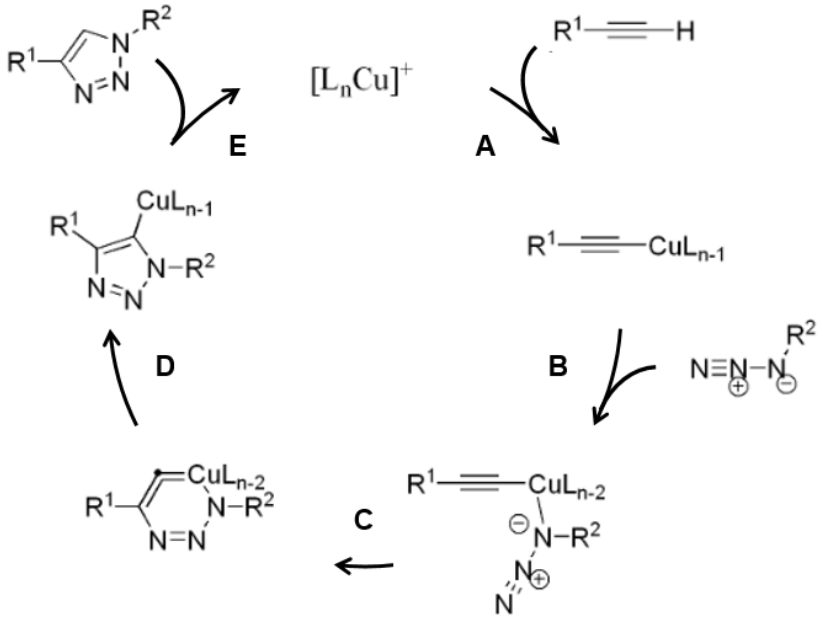


Figure 10: Proposed mechanism of the Cu(I)-catalysed Huisgen 1,3-dipolar cycloaddition of an azide and an alkyne forming a 1,2,3-triazole (Himo, Lovell et al. 2005)

It starts with the formation of the copper(I) acetylide (**A**) and then, the azide replaces one of the ligands and binds to the copper atom by the nitrogen, which is proximal to the carbon (**B**). Subsequently, the distal nitrogen of the azide attacks the C2 carbon of the acetylide, forming the unusual six-membered copper(III) metallacycle (**C**). From there, the barrier for ring contraction,

which forms the triazolyl-copper derivative, is very low (**D**). Proteolysis releases the triazole product completing the catalytic cycle (**E**) (Himo, Lovell et al. 2005).

There are several possibilities to introduce the Cu(I) species. First, a direct source of copper(I) could be used, for example, the complex tetrakis(acetonitrile)copper(I) hexafluorophosphate, which is used in this work and depicted in figure 11. Second, Cu(I) is generated within the reaction. $CuSO_4$ is commonly used in presence of a reducing agent like sodium ascorbate. An alternative would be the oxidation of Cu(0) metal (Meldal and Tornoe 2008).

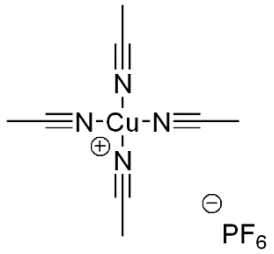


Figure 11: Cu(I)-complex for copper-catalysed azide-alkyne cycloaddition

Polytriazoles are suitable ligands for the reaction, especially derivatives of propargylamines, for example TBTA (figure 12A). It is proposed that the tertiary amine and the 1,2,3-triazole functionalities likely work in concert to make TBTA an efficient ligand (Chan, Hilgraf et al. 2004). A further improvement of the click reaction was the use of THPTA (figure 12B), which has a similar structure to TBTA, but has the advantage to be water-soluble (Hong, Presolski et al. 2009). Therefore, it is the preferred ligand for reactions with proteins under physiological conditions, as it is the case in this study.

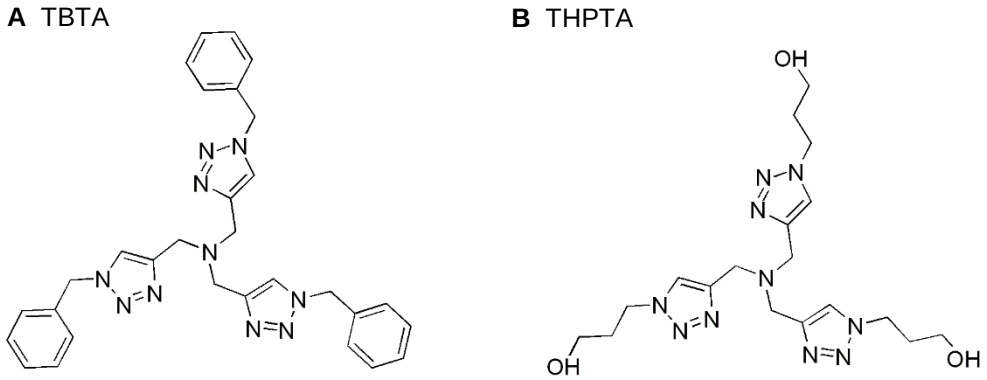


Figure 12: Two possible ligands, TBTA (A) and THPTA (B) for copper-catalysed azide-alkyne cycloaddition

1.5. Glycosylated proteins

Glycosylation is one of the most prominent posttranslational modification. More than half of all proteins are glycosylated (Apweiler, Hermjakob et al. 1999), among them almost all secretory proteins. Glycosylation has numerous functional consequences on proteins solubility, folding,

assembly to complexes and specific biological interactions, for example in cell-cell recognition, immune response and development (Arnold, Wormald et al. 2007). As such, it has also become very important for the biotechnical production of drugs containing glycoproteins. Therefore, the regulation of protein function by these modifications has to be understood. However, this is difficult because of the high heterogeneity of natural glycoproteins in their oligosaccharide structure (Thobhani, Yuen et al. 2009), which is due to their biosynthesis.

1.5.1. Classification of glycoproteins

There are three types of *N*-linked oligosaccharides: high-mannose, complex and hybrid (figure 13). The complex type shows the highest diversity with multiple branching possibilities and optional terminal capping with sialic acids. Higher eukaryotes often have the high-mannose type of *N*-glycan with weakly processed structures bearing terminal mannose residues derived from the nascent 14-mer (see also chapter 1.5.2). Hybrid *N*-glycans possess properties of both other types: in the 1,3-branch, complex-type termini are present and the 1,6-branch consists out of a high-mannose part (Kajihara, Yamamoto et al. 2010, Unverzagt and Kajihara 2013).

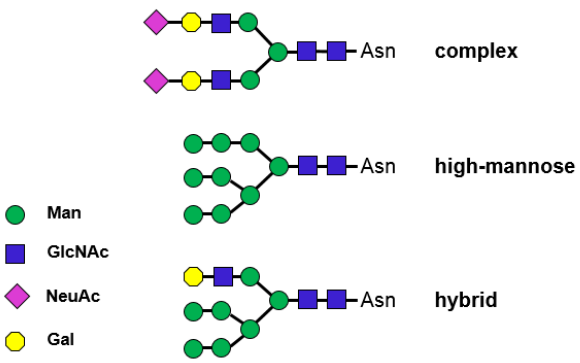


Figure 13: Overview of different types of *N*-glycans found on *N*-glycoproteins

1.5.2. Biosynthesis of glycosylated proteins

Glycosylation can be classified into *O*- and *N*-linked. *O*-linked glycans are attached to the hydroxyl group of serine or threonine residues and are normally rather short in mammalian cytosolic glycoproteins (Wells, Vosseller et al. 2001). The first residue is an *N*-acetyl-α-D-galactosamine. In the Golgi apparatus additional oligosaccharide units are transferred, which leads to eight basic core structures of moderate complexity. *N*-linked glycans are more complex and so, their biosynthesis (figure 14). Within the endoplasmic reticulum (ER), a 14-mer oligosaccharide is built up onto a dolichol phosphate embedded in the lipid bilayer of the ER and is transferred in its entirety to the amide of an asparagine residue of the nascent peptide chain. Subsequently, it is enzymatically elongated and a folding process takes place, either spontaneously or with the help of chaperones, which can discriminate between properly folded and misfolded proteins (Helenius and Aebi 2004, Lizak, Gerber et al. 2011). The correctly folded glycoproteins are transferred into the Golgi apparatus, where the initial high-mannose type glycans are converted into the complex- or hybrid-type. Finally, the completed *N*-linked glycoprotein is translocated to the cell surface or is secreted. During

remodelling, a vast variety of the final oligosaccharides is generated. The resulting mixture of glycoforms makes it difficult to elucidate to what extend one individual glycoform is involved in protein functions such as trafficking, secretion and bioactivity (Kajihara, Yamamoto et al. 2010, Unverzagt and Kajihara 2013).

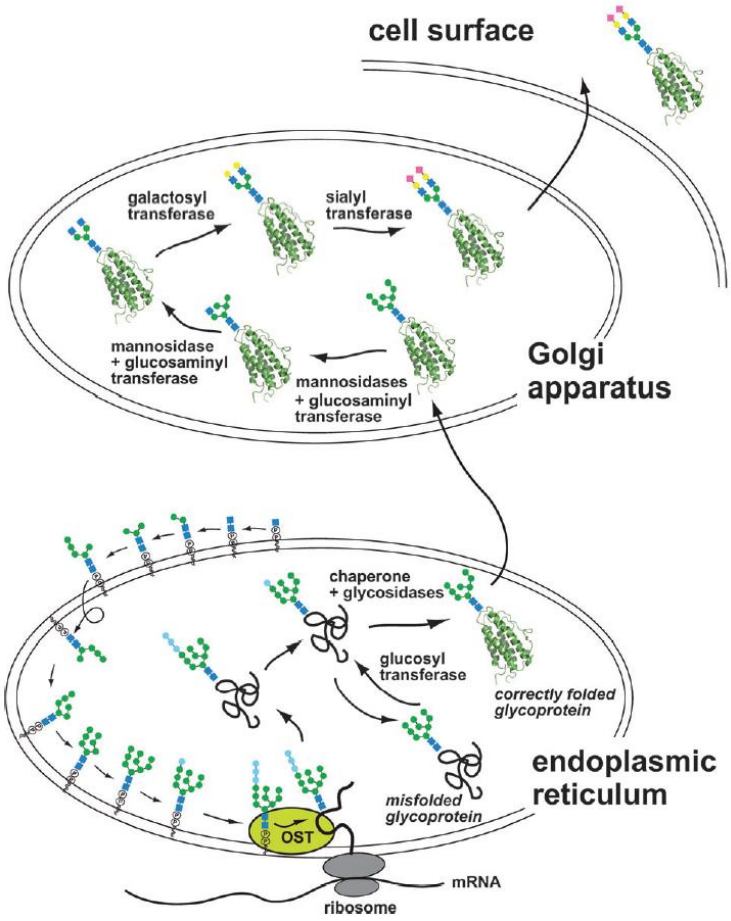


Figure 14: Biosynthesis of *N*-glycoproteins, figure is adopted from (Unverzagt and Kajihara 2013)

1.5.3. Production of glycosylated proteins

There are several possibilities for the production of glycoproteins. Many approaches for synthesis of glycoproteins and their analogues exist. However, only in a few cases homogeneous glycoproteins were obtained because chemical synthesis of glycoproteins is still in its fledging stages. Often only low yields are obtained in laborious multi-step synthesis (Hackenberger and Schwarzer 2008, Kajihara, Yamamoto et al. 2010, Payne and Wong 2010, Yuan, Chen et al. 2010). Therefore, pharmaceutic companies use in general expression systems for the commercial production of glycosylated proteins, either in eukaryotic systems or in bacterial hosts. Examples for eukaryotic systems are Chinese hamster ovary (CHO) cells and yeast species such as *Saccharomyces cerevisiae* and *Pichia pastoris*. The advantages of such systems are that tetra-antennary glycans are linked to the proteins by the host cells themselves. However, there are also several disadvantages: The production is very costly and time intensive and the risk of obtaining hypermannosylated proteins is very high. Those proteins cannot be commercialised. In general, glycosylation patterns of proteins

produced by eukaryotic systems are heterologous and are distinct to glycoforms in humans. In contrast, the production in bacterial systems, such as in *E. coli* has the benefit that high protein yields at low cost levels are reached. However, post-translational modifications like glycosylation do not occur. Non-glycosylated proteins are less soluble and prone to precipitation and aggregation.

1.6. PEGylated proteins

PEGylation defines the linkage of one or more polyethylene glycol (PEG) chains to proteins or peptides and has become a simple alternative to glycosylation with similar effects. In the 1970s, Davis first described PEGylation and concluded that the hydrophilic polymer link could increase the half-life of conjugated proteins *in vivo* and reduce immunogenicity (Davis 2002). Since then, the procedure of PEGylation was developed and now, a vast range of chemical and enzymatic methods for conjugation is available. PEG is the most successful covalent linked polymer to pharmaceutically active molecules with the following benefits (Nucci, Shorr et al. 1991, Pasut, Guiotto et al. 2004):

- Stabilisation of labile drugs from chemical degradation
- Protection from proteolytic degradation
- Reduction of immunogenicity, antigenicity and toxicity
- Enhancement of water solubility
- Increase of the circulatory half-life
- Reduction of renal clearance, mainly due to the increased molecular weight

Proteins are usually PEGylated by a chemical reaction between the suitably activated PEGylation reagents and various chemical groups on the amino side chains of the protein, such as carboxyl, hydroxyl, amino and thiol groups. Most frequently, random PEGylation is applied on the ε-amino groups of the relative abundant amino acid lysine usually located on the protein surface. The result will be a complex mixture of conjugates with various numbers and sites of PEGylation. Furthermore, most of the employed PEGylation reagents are not strongly specific for the reaction with amino groups of the lysine residues, but react also with other protein nucleophiles, for example N-terminal amino groups and the side chains of serine, threonine, tyrosine and cysteine residues (Jevsevar, Kunstelj et al. 2010).

Examples for a site-specific approach are the N-terminal and the cysteine-specific PEGylation. N-terminal PEGylation is achieved by a reductive alkylation step with a PEG-aldehyde reagent and a reducing agent. Thiol-specific reagents for cysteine-specific PEGylation are maleimide, pyridyl disulphide and vinyl sulfone. By this method, the PEG chain is coupled to natural or genetically introduced unpaired cysteines. This could be a drawback because in native proteins, cysteine residues are required in disulphide bridges or are responsible for the interaction with other proteins (Jevsevar, Kunstelj et al. 2010).

The establishment of PEGylated proteins as pharmaceuticals can be divided into two generations (Pasut, Guiotto et al. 2004):

- PEG chains with low molecular weight (under 12 kDa) are considered as the first PEG generation. They contain a relevant percentage of PEG diol impurities. Moreover, the

employed chemistry was not optimised, as side reactions and weak or reversible linkages have been described. Examples for first generation drugs are Adagen®, a PEG-adenosine deaminase for the treatment of severe combined immunodeficiency disease or Oncaspar®, a PEG-asparaginase for the treatment of leukaemia both from Enzon Pharmaceuticals, Inc (Levy, Hershfield et al. 1988, Graham 2003).

- The second generation of conjugates were an advancement over the first, as the impurities and polydispersity were reduced. Furthermore, selectivity of protein modification and availability of activated PEGs were improved and spacers between PEG and protein were investigated. PEG-Interferon-α2b marketed as PEG-Intron® from Schering-Plough and a branched 40-kDa-PEG-Interferon-α2a marketed as Pegasys® from Roche Pharmaceuticals are two examples for successful second generation PEGylated drugs (Bailon, Palleroni et al. 2001, Rajender Reddy, Modi et al. 2002, Wang, Youngster et al. 2002).

1.7. Erythropoietin

Erythropoietin is a globular glycoprotein hormone (Davis, Arakawa et al. 1987) and the primary haematopoietic growth factor cytokine for the maturation of erythrocyctes from precursors in the bone marrow (Stephenson, Axelrad et al. 1971, Goldwasser 1984). Therefore, it is used as a pharmaceutical to treat anaemia associated with chronic kidney disease (CKD), chemotherapy and AIDS treatment (Cazzola, Mercuriali et al. 1997, Sowade, Sowade et al. 1998).

1.7.1. Structural properties

1.7.1.1. Amino acid sequence of EPO

EPO has 166 amino acids (figure 15) and without glycosylation a theoretical molecular weight of about 18400 kDa and a theoretical pI of 8.75.

```
        10         20         30         40         50         60
APPRLICDSR VLERYLLEAK EAENITTGCA EHCSLNENIT VPDTKVNFYA WKRMEVGQQA

        70         80         90        100        110        120
VEVWQGLALL SEAVLRGQAL LVNSSQPWEP LQLHVDKAVS GLRSLTTLLR ALGAQKEAIS

       130        140        150        160
PPDAASAAPL RTITADTFRK LFRVYSNFLR GKLKLYTGEA CRTGDR
```

Figure 15: Amino acid sequence of EPO

1.7.1.2. Secondary structure of EPO

The crystal structure of EPO was resolved by (Syed, Reid et al. 1998) (figure 16). They used a non-glycosylated EPO variant with the following mutations: N24K, N38K, N83K, P121N and P122S. The two proline residues were mutated based on the possible *cis-trans* conformation heterogeneity, which was observed in ^{15}N-NMR relaxation data (Cheetham, Smith et al. 1998). They observed extremely low-order parameters for the residues E117-A128 in the loop with conformational heterogeneity in the backbone in the vicinity of the proline residues.

EPO is a member of the cytokine hormone family, which shares a four helical bundle "up-up-down-down" motif. This requires two long loops between the helix A and B and between helix C and D. A disulphide bridge holds together the antiparallel helices A and D at positions C7 and C161. The second disulphide bridge between C29 and C33 links the end of the A helix with part of the AB loop (Cheetham, Smith et al. 1998, Syed, Reid et al. 1998). Additionally, EPO has two short helices and two small antiparallel ß-strands. The later ones are typical for the short-chain class, to which belong the macrophage colony-stimulating factor, stem cell factor, interleukin-4 and -5 (Rozwarski, Gronenborn et al. 1994).

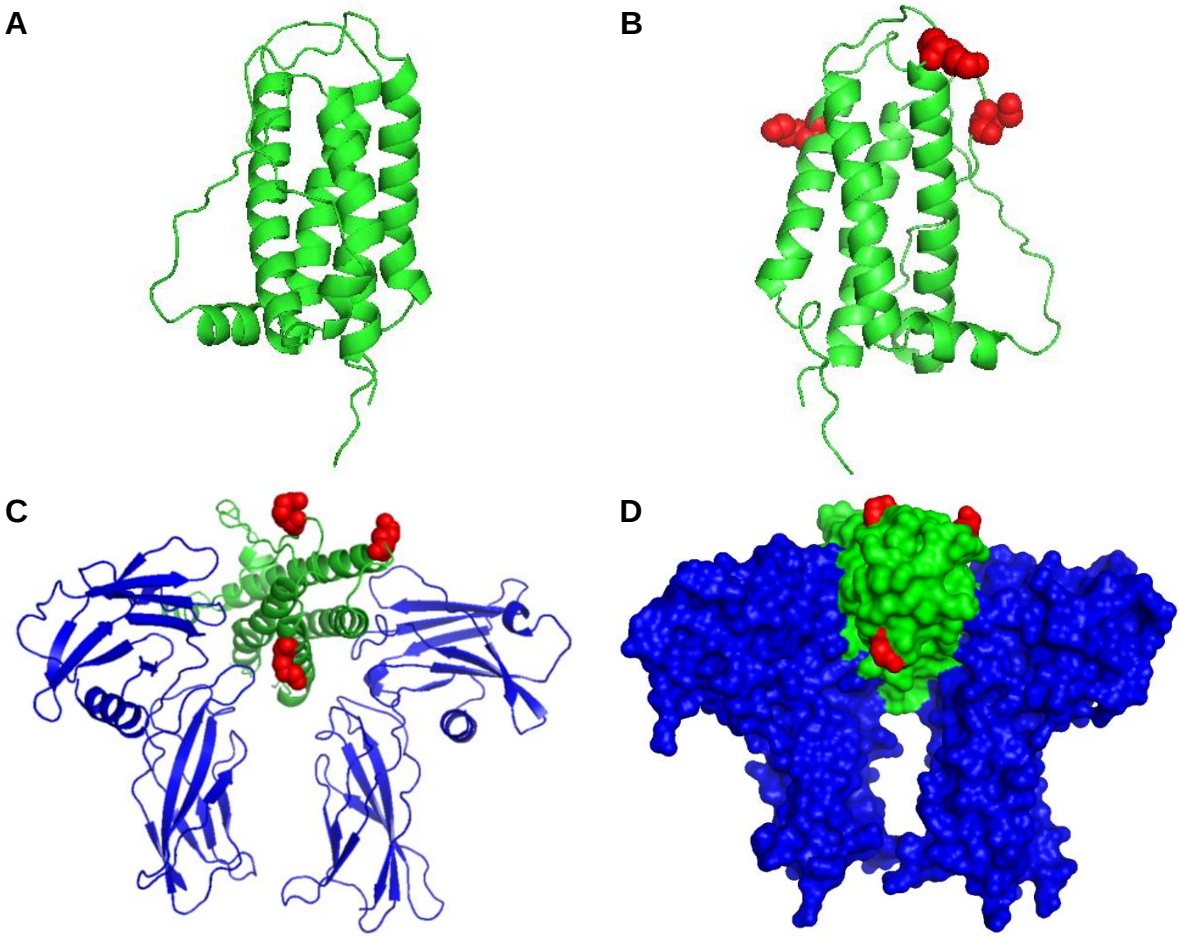


Figure 16: Structure of EPO (A), EPO (green) with three mutated N-to-K residues at the natural glycosylation sites (red spheres) (B), EPO bound to its receptor (blue) (C-D), (Syed, Reid et al. 1998)

1.7.1.3. Glycosylation pattern of EPO

EPO has one *O*-glycosylation (S126) and three *N*-glycosylation (N24, N38 and N83) sites. Glycosylation accounts for 40% of the molecular weight of EPO, which is approximately 30 to 34 kDa (Takeuchi, Takasaki et al. 1988). The *N*-linked carbohydrate chains have two to four branches often with terminal sialic acids. These carbohydrate chains are not required for receptor binding *in vitro*, but are important for the *in vivo* activity of EPO (Delorme, Lorenzini et al. 1992).

As a high degree of heterogeneity in the sialic acid distribution and in the branching is observed within each *N*-glycosylation site, as well as between each of the sites, there are different EPO

glycoforms with a maximum of 14 negatively charged sialic acids. Most of the glycan chains on EPO belong to the complex type with fucosylated tri- or tetraantennae. Bi-antennary sugar chains are less abundant (Sasaki, Bothner et al. 1987, Sasaki, Ochi et al. 1988). Sialic acid residues increase the solubility of the protein and thus, are more effective in stimulating erythropoiesis *in vivo* (Egrie and Browne 2001).

1.7.2. Biological function

1.7.2.1. Erythropoiesis

Erythrocyte production is dynamic and tightly regulated. A total number of 2-3 x 10^{13} erythrocytes is maintained by healthy adults, which are approximately 5 million erythrocytes per microlitre blood. The life span of an erythrocyte is about 120 days. Therefore, 1% of the circulating erythrocytes is replaced each day (Torbett and Friedman 2009).

Erythropoiesis is the term for the regulated process of proliferation and differentiation of haematopoietic progenitor cells into mature red blood cells (figure 17). The process can be divided into three major steps (Baron and Fraser 2005, McGrath and Palis 2008):

- Generation of erythroid committed blast cells from multipotent haematopoietic progenitors
- Division and differentiation of erythroid progenitor cells
- Terminal cellular morphologic changes (including enucleation) to produce reticulocytes and eventually mature red cells

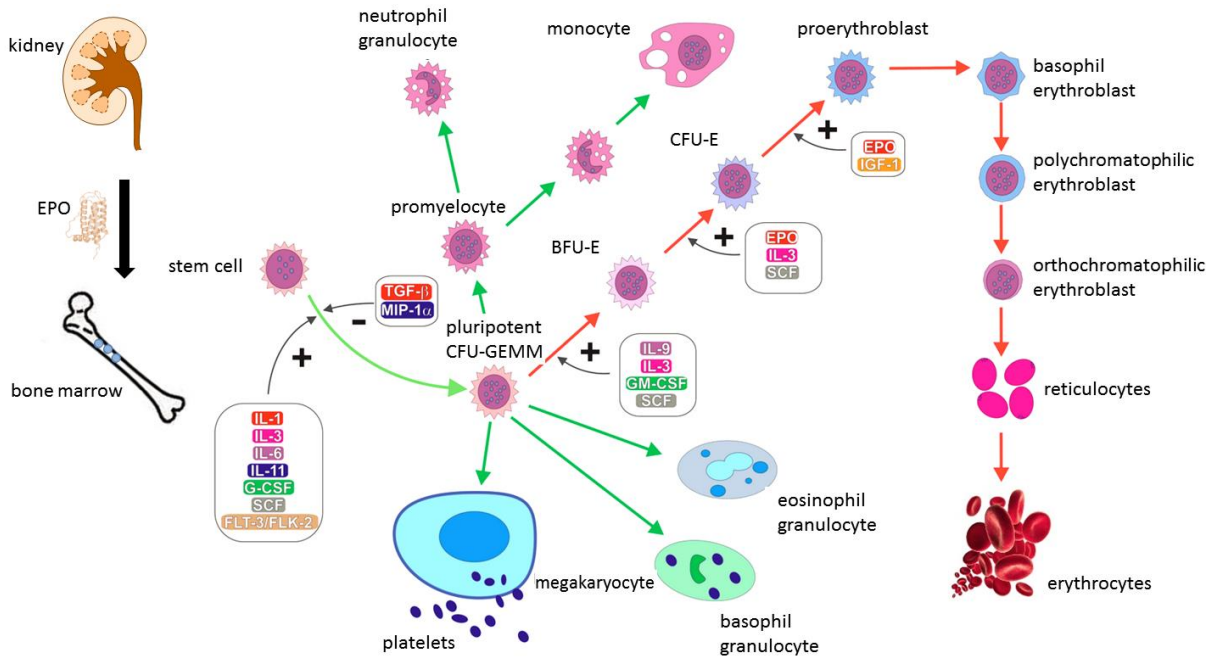


Figure 17: Schematic overview of the role of EPO in erythropoiesis

Erythroblasts are produced in adults from committed progenitors in the bone marrow. During this process, the multipotency is lost and there is an increased lineage restriction (Baron and Fraser 2005,

McGrath and Palis 2008). The stages in erythrocyte formation are defined by their ability to form colonies in semisolid medium supplemented with specific cytokines (CFU, colony-forming units). At early stages, the cells respond to a broad variety of cytokines, but as differentiation progresses, the cytokine responses become more specific to erythroid progenitors (Migliaccio and Migliaccio 1988, Koury and Bondurant 1990, McGrath and Palis 2008).

A common myeloid progenitor gives rise to bipotential progenitors restricted to either the erythroid/megakaryocyte or the granulocyte/macrophage pathways. *In vitro,* this stage of development is represented by the colony-forming unit – granulocyte, erythrocyte, macrophage, megakaryocyte (CFU-GEMM) precursor (Debili, Coulombel et al. 1996, Akashi, Traver et al. 2000). Only the erythroid/megakaryocyte-restricted progenitors express the erythropoietin receptor (EPOR) and are responsive to EPO. The most immature erythroid-restricted progenitor is the burst-forming unit – erythroid (BFU-E) (Stephenson, Axelrad et al. 1971, Heath, Axelrad et al. 1976). They are highly proliferative blast-like cells, express EPOR only moderately and give rise to CFU-E progenitors, which are highly EPO-responsive. On these cells, EPOR is expressed in high amounts. CFU-E progenitors begin to express haemoglobin and generate smaller colonies. It was found that EPO stimulate division and prevent apoptosis (Heath, Axelrad et al. 1976).

Several stages of morphologically identifiable nucleated precursors arise from CFU-E to reticulocytes (Stephenson, Axelrad et al. 1971). Important cellular processes take places during this development (Torbett and Friedman 2009):

- Accumulation of haemoglobin
- Decrease in cell size
- Nuclear condensation
- Final enucleation

Erythroblasts mature in the erythroblastic island, which is a specialised microenvironmental niche in the bone marrow (Manwani and Bieker 2008). These islands consist of a central macrophage that extends cytoplasmic protrusions to a ring of surrounding erythroblasts (Gifford, Derganc et al. 2006). The macrophage serves as a source for nutrients, survival and proliferative signals to the erythroblasts. Finally, the reticulocytes mature into erythrocytes with the help of the central macrophage (Manwani and Bieker 2008).

1.7.2.2. Hypoxic regulation of EPO

The regulation of EPO production during normoxia and hypoxia is depicted in figure 18. Hypoxia is primarily sensed in the kidneys and will lead to an increase in EPO production. The renal produced EPO stimulates the maturation of the erythroid progenitors in the bone marrow. The increased number of red blood cells carrying oxygen, results in a corrected oxygen state of the tissue. Therefore, the concentration of EPO in blood serum is inverse proportional to the haematocrit as an indicator for hypoxia (Torbett and Friedman 2009).

In adult mammals, peritubular interstitial fibroblasts in the kidney are the major EPO production site (Lacombe, Da Silva et al. 1988, Bachmann, Le Hir et al. 1993). At the molecular level, EPO expression is coupled to prolyl hydroxylase domain (PHD) proteins, which are oxygen sensors, and to the

transcription factor hypoxia inducible factor (HIF). HIF is a heterodimeric transcription factor consists of a labile α- and a constitutively expressed β-subunit (Wang and Semenza 1993, Wang and Semenza 1993). Under normoxic conditions, PHD proteins constitutively hydroxylate two specific proline residues in HIF-α, which then, can be bound by the von Hippel-Lindau protein. Subsequently, this leads to ubiquitination and proteosomal degradation of HIF-α (Semenza 2001). However, under hypoxia conditions, HIF-α is stabilised by HIF-β and subsequently, the transcription of the EPO gene and other hypoxia response genes is activated. The α-subunit is rate-limiting in the transcription complex and its destruction is controlled by the amount of cellular oxygen (Torbett and Friedman 2009).

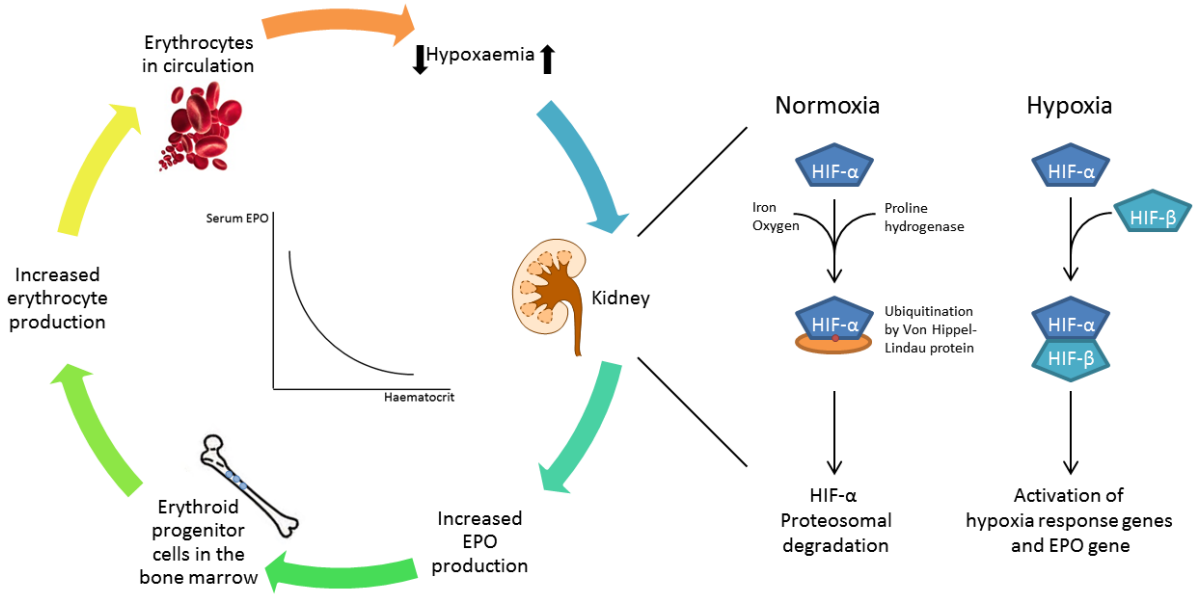


Figure 18: Schematic overview of the regulation of EPO production during normoxia and hypoxia; figure is inspired by (Torbett and Friedman 2009)

1.7.2.3. EPO receptor signalling processes

Figure 19 depicts an overview of EPOR signalling processes. EPOR is a member of the cytokine-receptor superfamily, which is characterised by an extracellular-binding region, a transmembrane region and an intracellular domain (Youssoufian, Longmore et al. 1993). Upon binding of EPO to EPOR, a tighter connection of the two homodimers of EPOR is induced, due to a conformational change (Cheetham, Smith et al. 1998) and two Janus kinase 2 (JAK2) tyrosine kinase molecules are activated (Witthuhn, Quelle et al. 1993, Remy, Wilson et al. 1999). This leads to phosphorylation of several tyrosine residues in the intracellular region of EPOR, which is a docking site for signalling proteins with phospho-tyrosine binding motifs of several pathways including STAT5, phosphatidyl-inositol 3-kinase (PI3K/Akt) and Ras/MAPK (Richmond, Chohan et al. 2005, Watowich 2011). Finally, these pathways lead to transcription of genes for survival, proliferation and differentiation of the cell (Jelkmann 2004). The signal transduction is terminated by the haematopoietic cell phosphatase, which catalyses the dephosphorylation of JAK2 (Klingmuller, Lorenz et al. 1995). The EPO/EPOR complex is internalised after dephosphorylation of the receptor. The duration of EPO signalling is

controlled by the proteasome, which inhibits the renewal of receptor molecules on the cell surface (Verdier, Walrafen et al. 2000).

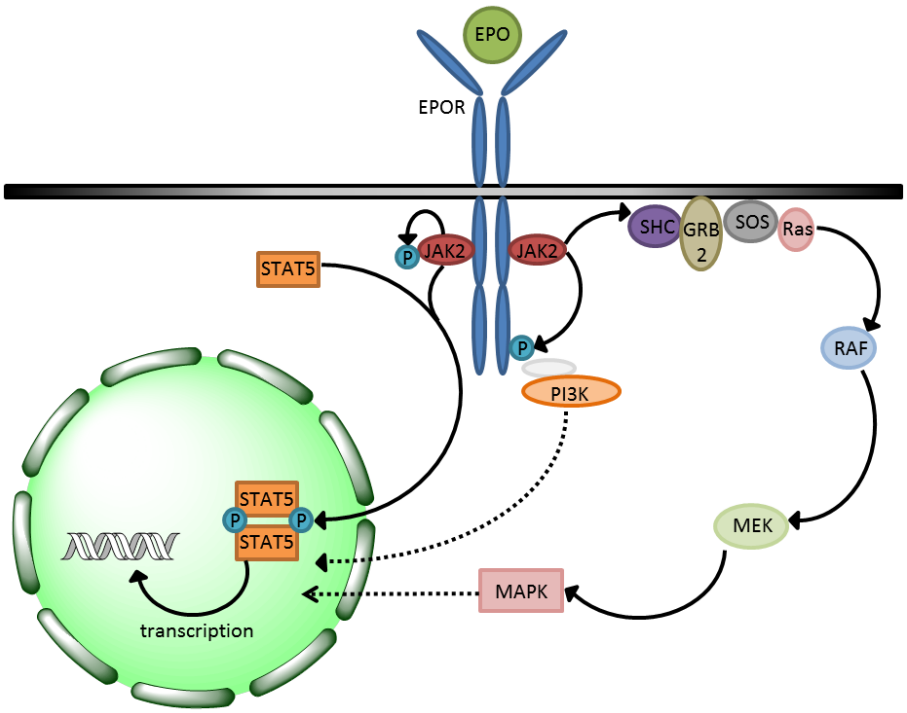


Figure 19: Schematic overview of intracellular signalling processes upon EPO receptor binding; figure is inspired by (Jelkmann 2004)

1.7.3. EPO as a pharmaceutical

1.7.3.1. Historical aspects

Jourdanet was the first, who discovered the relationship between altitude and blood viscosity in 1863. Viault contributed several years later by the finding that the altitude has an increasing effect on the number of red blood cells in circulation. In 1906, Carnot proposed a model for regulation of "haemopoietine" by injecting serum from anaemic rabbits into normal rabbits, which caused an elevated number of red blood cells. He concluded that the serum contained "haemopoietine", which was responsible for the increase in cell number (Foote 2009). 30 years later, Erslev extended the original study by investigating the number of nucleated red blood cells in the bone marrow, the number of peripheral reticulocytes and the haematocrit of the normal rabbits after injection of large amounts of plasma from anaemic rabbits. This showed that the red blood cell production is mediated by a humoral factor (Erslev 1953). In 1957, it was found that EPO is produced in the kidney by (Jacobson, Goldwasser et al. 1957).

Until 1977, only insufficient amounts of EPO were isolated and purified for characterisation and a potentially development of EPO as a therapeutic protein. Then, a group around Goldwasser was able to isolate and purify EPO in milligram amounts from 1500 L of aplastic anaemia patients' urine (Miyake, Kung et al. 1977). Even though, the amounts of EPO were still small and the purification

procedure took seven steps, it was enough material to partially characterise the protein (Foote 2009).

Lin successfully cloned the EPO gene, after many failed trials. In a novel approach, multiple sets of fully degenerate oligonucleotide probes were used to screen a human genomic library. In this case, two small pools of oligonucleotides corresponded to short fragmented samples of EPO amino acid sequences. 128 different probes were needed in each pool for every possible codon that encoded these putative amino acid sequences. These probes were labelled radioactively to identify any matches of a single probe with the human genome (Lin, Suggs et al. 1985). In that way, they found the entire coding region of the human gene for EPO, which was the basis for the development of the expression system in CHO cells. Another approach for cloning the human EPO gene using degenerate oligonucleotides was also successful (Jacobs, Shoemaker et al. 1985).

1.7.3.2. Classification of erythropoiesis-stimulating agents (ESAs)

Recombinant human EPOs that have the same amino acid sequence including disulphide bridges and glycosylation sites as endogenous EPO are called “epoetins”, according to the International Nonproprietary Name (INN), experts of the World Health Organization (WHO 2007). However, all recombinant forms have a distinct glycosylation pattern, compared to native EPO, indicated by Greek letters. The glycan composition depends on the cell line used for expression and the protein purification procedures (Jelkmann 2007). CHO cells are used as expression hosts for the production of Epoetin α and β. Epoetin β is characterised by a higher sialylation percentage of the tetra-antennary glycans. Epoetin ω is expressed in baby hamster kidney (BHK) cells and has a distinct glycan pattern and antennary structure than epoetins α and β.

1.7.3.3. Clinical use of ESAs in anaemia

Some of the benefits of epoetin α in therapy are listed below (Foote 2009):

- Increased exercise tolerance
- Improved central nervous system function
- Reduced heart enlargement
- Reduced extreme fatigue
- Increased ability to perform daily functions of life
- Reduced risk of alloimmunisation in transplant recipients
- Improved coagulation

1.7.3.4. Commercial production of recombinant EPO

The market for EPO has increased steadily for all indications: chronic kidney disease, oncology, dialysis and pre-dialysis. In a data review from 2006, four of the top ten biotech drugs were a form of recombinant EPO with darbepoetin α (Aranesp® from Amgen) ranking third globally, which is an engineered epoetin with two additional *N*-glycosylation sites (Lawrence 2006) (figure 20). Other currently marketed forms of recombinant EPO include Epogen (Epoetin α from Amgen) with three

N-glycans and PEGylated EPOs, such as Continuous Erythropoietin Receptor Activator (CERA, Mircera from Hoffmann-La Roche). This recombinant EPO is PEGylated either at the N-terminal amino group or at the ε-amino group of lysines (predominantly K52 or K45) (Macdougall 2005).

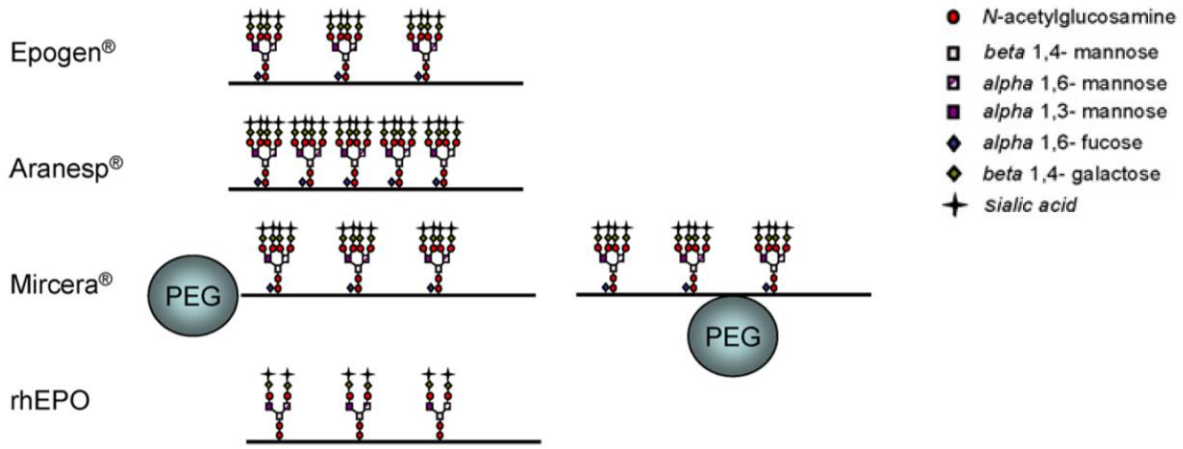


Figure 20: Overview of three currently marketed ESAs in comparison to recombinant human EPO produced in CHO cells; figure is adopted from (Nett, Gomathinayagam et al. 2012)

1.8. Objective of this work

Aim of this study was to develop and optimise a new semi-synthetic approach to produce homogeneous glycosylated EPO in order to study the effects of carbohydrate composition and position on the properties and function of EPO (figure 21). Natural human EPO, as well as current recombinant EPO species are determined in their amino acid sequence and position of glycosylation. However, the heterogeneity of the glycosylation structure hinders the correlation between glycan composition and glycoprotein function.

The strategy to obtain homogenous glycosylated EPO was to use *E. coli* as bacterial host, which does not glycosylate proteins as a posttranslational modification and to incorporate non-canonical amino acids with a specific bio-orthogonal chemical function at defined sites that can be selectively linked *in vitro* with defined, synthetic oligosaccharides by click chemistry. In detail, the applied strategy combines the benefits of the following methods:

- The expression in a bacterial system (*E. coli*) produces high protein yields.
- The incorporation of non-natural amino acids is site-specific and bio-orthogonal adding new properties to the protein (for example alkyne or azide functions).
- The chemical synthesis of carbohydrates leads to homogeneous and defined glycan chains.
- The convenient copper-catalysed 1,3-dipolar Huisgen cycloaddition specifically couples glycan chains to EPO at the introduced azide- or alkyne groups.

Furthermore, the developed method could be used to not only glycosylate proteins, but also to PEGylate them site-specifically without any side reaction, which is a major advantage over traditional PEGylation methods.

The produced glycosylated and PEGylated EPO variants were subsequently characterised biophysically and –chemically and their biological activity was assessed by *in vitro* proliferation and differentiation assays. These studies showed position- and composition-specific effects of glycosylation and PEGylation on EPO.

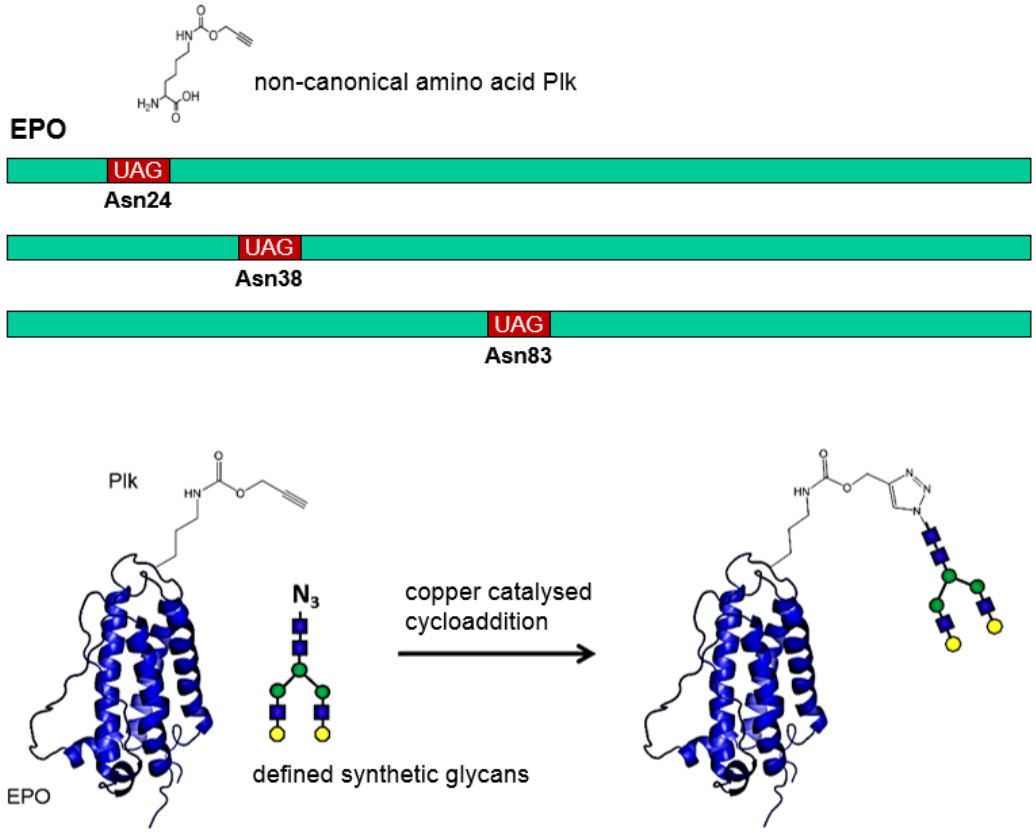


Figure 21: Semi-synthetic approach for the synthesis of glycosylated EPO with defined and uniform oligosaccharides.

2. Material

2.1. Disposables

Disposable	Supplier
96 well plates, flat bottom, transparent	Greiner bio-one
96 well plates, flat bottom, white polystyrol	Greiner bio-one
Electroporation gene pulser cuvettes, 0.5 mm	Biorad
Falcon tubes, 15 and 50 mL	Roth
Immobilon-P transfer membrane	Millipore
Injection needles	Braun
Low profile thermo-stripes (12 x 0.2 mL)	ABgene
Parafilm	Parafilm
PD-10 desalting columns	GE Healthcare
Petri dishes	Peske
Reaction tubes, 1.5 and 2.0 mL	Peske

Reaction tubes, 200 µL	ABgene
Scalpels	Bayha
Snake skin pleated dialysis tubing	Thermo Scientific
Sterile filtration units	Nalgene
Syringe	Peske
Syringe sterile filters	Milipore
Tips for laboratory pipettes	Peske
Tips for multichannel pipettes	Peske
UV-cuvettes	Roth
Vivaspin columns	Sartorius
Whatman paper, 3 mm	Merck Eurolab

Table 1: Disposables

2.2. Software

Name	Company
Clone Manager 5	Scientific and Educational Software
Microsoft Office Excel 2010	Microsoft
Microsoft Office PowerPoint 2010	Microsoft
Microsoft Office Word 2010	Microsoft
PyMOL	Schrödinger
Quantity One®	Biorad
Origin 2015G	OriginLab Corporation
EndNote X7.1	Thomson Reuters
ChemDraw Ultra 14.0	Cambridge Soft

Table 2: Software

2.3. Equipment

Device	Identification	Producer
96-well plate reader	Infinite 200 Pro	Tecan
Agarose gel racks		Fisher Scientific
Autoclave		Tecnoclav 50
Balances	PJ3000	Mettler
	PG403S	Mettler
CD cuvettes (Quartz SUPRASIL®)	110-QS, light path: 1 mm	Hellma Analytics
CD spectropolarimeter	J-815 with accessory MPTC-490S	Jasco
Centrifuges	5810R	Eppendorf
	Multifuge 4KR	Heraeus
Chromatography columns	XK and C columns	GE Healthcare
Electrophoresis device		Biorad
Electroporator	Gene Pulser Xcell	Biorad
Gel documentation device	Chemidoc XRS	Biorad
Gel drier		Biorad
Heating blocks		Fisher Scientific
Incubation shaker for tubes and flasks	Innova4430	New Brunswick Scientific
Magnetic stirrer	MR 3000 D	Heidolph
microplate scintillation and	TopCount® NXT™	PerkinElmer

luminescence counter		
Multichannel pipettes	Transferpette	Brand
Overhead shaker		Heidolph
PCR thermocycler		Miometra
pH-meter	Seven Easy	Mettler Toledo
Pipettes		Eppendorf
Pipettor	Pipetboy	Eppendorf
Power supply unit	Power Pac 3000	Biorad
Refrigerated centrifuge	Biofuge Primo R	Heraeus
SDS-PAGE racks		Biorad
SEC column	Superdex 75 10/300 GL	GE Healthcare
Spectrophotometer	Nanodrop	Peqlab
Speedvac	Concentrator 5301	Eppendorf
Sterile bench		HERA safe
Table top centrifuge	5417C Eppendorf Mini Spin	Eppendorf
Table top shakers	KS 260 basic	IKA
Thermocycler	T gradient	Biometra
Thermomixer	Thermomixer comfort	Eppendorf
Ultrapure water installation		Sartorius
Ultrasonic homogenizers	Sonifier 250	Branson
UV/VIS-photometer	BioPhotometer	Eppendorf
Vortexer	7-2020	Neolab
Water baths		Memmert
Wet blot cell	Mini Trans-Blot® cell	Biorad

Table 3: Equipment

2.4. Chemicals

2.4.1. Chemicals for molecular biology

Reagent	**Producer**
1,4-Dithiothreitol (DTT)	Roth
2-Mercaptoethanol	Roth
2-propanol	Riedel-de-Haen
3-(*N*-morpholino)propanesulfonic acid (MOPS)	Sigma-Aldrich
Acetic acid	Norma Pur
Agar	Roth
Agarose	Invitrogen
Ammonium persulphate (APS)	Fluka
Bis-Acrylamide	Roth
Boc-Lys-OH	Sigma-Aldrich
Boric acid	Fluka
Bromphenol blue	Fluka
Carbenicillin disodium salt	Roth
Chloramphenicol	Roth
CM Sepharose fast flow	GE Healthcare
Coomassie brilliant blue R 250	Thermo Scientific
Disodium hydrogenphosphate x 2 H_20	Merck
Ethanol	Roth
Ethidiumbromide	Roth

Ethylenediaminetetraacetic acid (EDTA)	Roth
Glucose	Riedel-de-Haen
Glycerol	Merck
Glycine	Roth
Guanidine hydrochloride	Sigma-Aldrich
Imidazole	Merck
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Roth
Kanamycin sulphate	Roth
L-Arginine monohydrochloride	Roth
LB broth	Roth
L-Cystine	Roth
Magnesium chloride	Acros Organics
Magnesium sulphate	ICN Biomedicals
Ni-NTA agarose	Qiagen and Biozyme
N-lauroylsarcosine	Sigma-Aldrich
Penicillin-Streptomycin (10,000 u Pen, 10 mg Strep per mL in 0.9% NaCl)	Sigma-Aldrich
Phenol red	Sigma-Aldrich
Phenylmethylsulfonyl fluoride (PMSF)	Roth
Sephadex G50	Amersham
Sodium chloride	Roth
Sodium dodecyl sulphate (SDS)	Roth
Sodium hydroxide	Merck
Sodium phosphate dibasic anhydrous	Sigma-Aldrich
Sodium phosphate monobasic monohydrate	Sigma-Aldrich
sodium thiosulfate	Sigma-Aldrich
SP Sepharose fast flow	GE Healthcare
TEMED (N,N,N',N'-Tetramethylethylenediamine)	Roth
Tetrakis(acetonitrile)copper(I) hexafluorophosphate	AG Winter, University Konstanz
Tritium thymidine	PerkinElmer
Triton X-100 (Polyethylene glycol p-(1,1,3,3-tetramethylbutyl)-phenyl ether)	Roth
Trizma® base (Tris)	Roth
Trypsin-EDTA solution (1 x, sterile)	Sigma-Aldrich
Tryptone	Roth
Tween® 20	Riedel-de-Haen
Urea	Roth
Yeast extract	Roth

Table 4: Chemicals for molecular biology

2.4.2. Components for click chemistry

The azido-carbohydrate were obtained from collaboration partner Prof. Unverzagt from University of Bayreuth. The small PEGs with azide group were purchased from Jena Bioscience and the PEG-20kDa from Iris Biotech GmbH. An mPEG-alkyne of 5 kDa was ordered from Creative PEGWorks (figure 22).

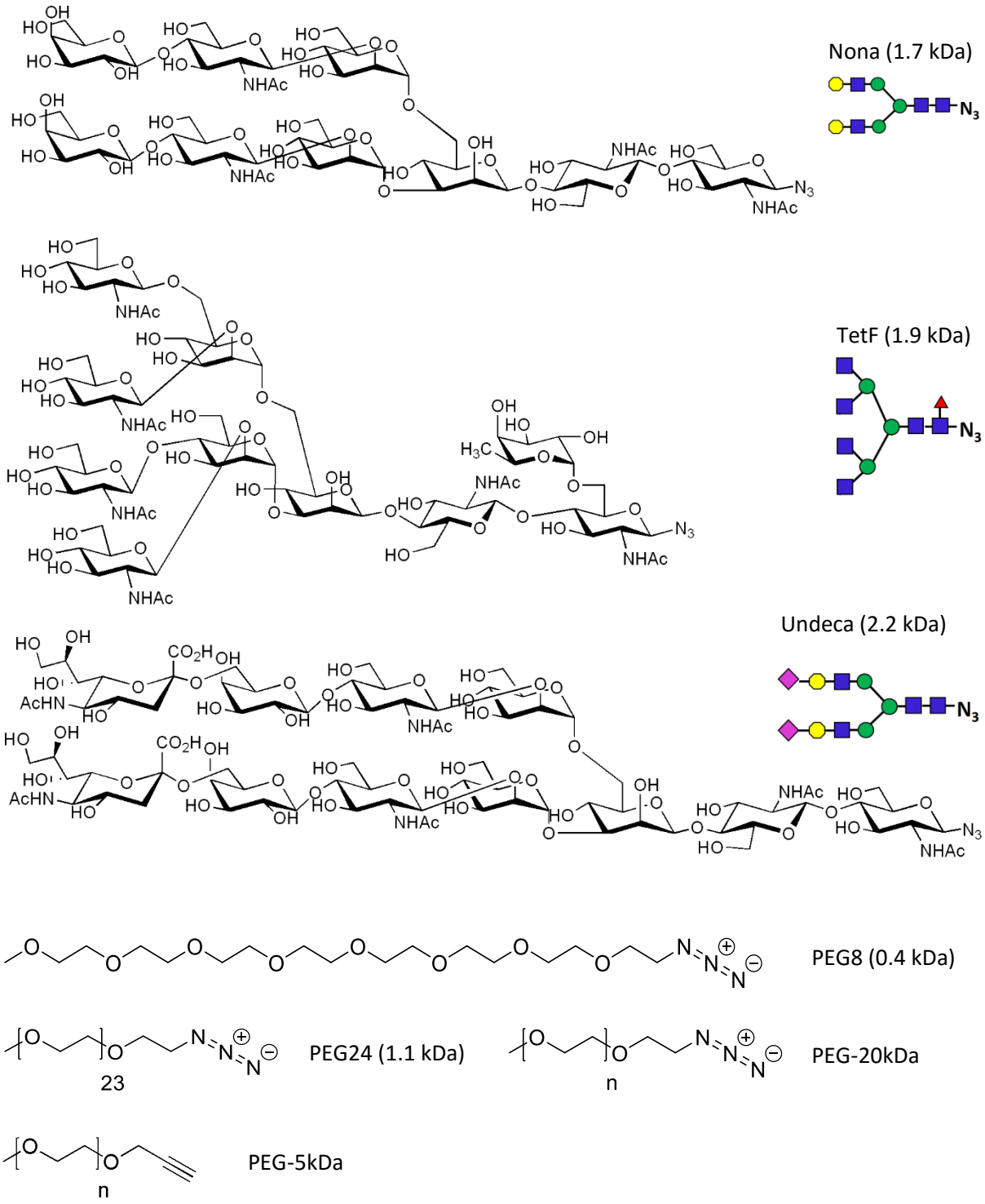


Figure 22: Structures of glycan- and PEG chains that were used as click-ligands in click reactions

A GlcNAc-ligand for the Cu(I)-complex was also provided from the group of Prof. Unverzagt, University of Bayreuth (figure 23).

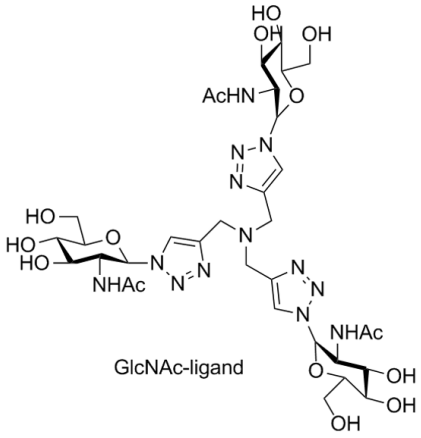


Figure 23: GlcNAc-ligand for click reactions

2.5. Buffers and solutions

TAE-buffer	
Tris-acetic acid, pH 8.3	40 mM
EDTA	1 mM

Table 5: TAE-buffer

Glycine SDS electrophoresis buffer		Sarcosyl running buffer	
Tris-HCl, pH 8.9	25 mM	MOPS, pH 7.7	50 mM
Glycine	200 mM	Tris	50 mM
SDS	1% (w/v)	*N*-lauroylsarcosine	0.1%
		EDTA	1 mM

Table 6: Glycine SDS electrophoresis buffer and sarcosyl running buffer

6 x sample buffer for SDS-PAGE		4 x sample buffer for Sarcosyl-PAGE (10 mL)	
Tris-HCl, pH 6.8	22.5% (v/v)	Tris-HCl, pH 8.5	988 mM
Glycerol	50% (v/v)	*N*-lauroylsarcosine	8%
SDS	5% (w/v)	Glycerol	40%
Bromphenol blue	0.5% (w/v)	EDTA	2.04 mM
2-mercaptoethanol	12.5% (v/v)	Phenol red	3 mg

Table 7: Sample buffers for SDS- and Sarcosyl-PAGE

Separating gel (15%)	volume in mL for four gels	Stacking gel (4%)	volume in mL for four gels
H_2O	3.75	H_2O	4.65
1.5 M Tris-HCl, pH 8.8	3.75	1 M Tris-HCl, pH 6.8	1.875
10% SDS	0.15	10% SDS	0.03
30% Bis-acrylamide	7.5	30% Bis-acrylamide	0.975
10% APS	0.112	10% APS	0.053
TEMED	0.015	TEMED	0.008

Table 8: Composition of separating and stacking gel for SDS-PAGE

Western blot transfer buffer	
Glycine SDS electrophoresis buffer	20%
Methanol	20%
H_2O	60%

Table 9: Western blot transfer buffer

AP buffer	
Tris-HCl, pH 9.5	100 mM
NaCl	100 mM
$MgCl_2$	100 mM

Table 10: AP buffer

TBS-Tween®	
Tris-HCl, pH 7.5	20 mM
NaCl	500 mM
Tween® 20	0.05%

Table 11: TBS-Tween®

Western blot detection solution	
AP buffer	10 mL
NBT Solution (one tablet containing 10 mg dissolved in 1 mL H_2O)	330 µL
BCIP Solution (one tablet containing 50 mg in 1 mL DMF)	33 µL

Table 12: Western blot detection solution

Cell lysis buffer	
Tris-HCl, pH 8.0	20 mM
EDTA	1 mM
PMSF	1 mM
Lysozyme	50 µg/mL

Table 13: Cell lysis buffer

Denaturation buffer	
Tris-HCl, pH 8.0	20 mM
Guanidine hydrochloride	6 M
2-mercaptoethanol	10 mM
EDTA	0.5 mM

Table 14: Denaturation buffer

Ni-NTA Purification Buffer	
Tris-HCl, pH 8.0	20 mM
Guanidine hydrochloride	3 M
2-mercaptoethanol	10 mM
EDTA	0.5 mM

Table 15: Ni-NTA Purification buffer

Gel Filtration Buffer, sterile filtrated	
Tris-HCl, pH 7.5	20 mM
NaCl	300 mM

Table 16: Gel filtration buffer

Refolding buffer	
Tris-HCl, pH 8.5	20 mM
L-arginine hydrochloride	1 M
L-glutathione, reduced	1 mM
L-glutathione, oxidised	0.3 mM

Table 17: Refolding buffer

2.6. Cell culture media

2.6.1. Bacterial media

LB medium	
Tryptone	1% (w/v)
Yeast extract	0.5% (w/v)
NaCl	1% (w/v)

Table 18: LB medium

LB agar	
Tryptone	1% (w/v)
Yeast extract	0.5% (w/v)
NaCl	1% (w/v)
Agar	2% (w/v)

Table 19: LB agar

SOC medium	
Tryptone	2% (w/v)
Yeast extract	0.5% (w/v)
NaCl	0.05% (w/v)
$MgCl_2$	10 mM
$MgSO_4$	10 mM
Glucose 20% (w/v)	2% (v/v)

Table 20: SOC medium

Amino acid solution	250 mL
Alanine	0.3 g
Arginine	0.3 g
Asparagine	0.3 g
Aspartic acid	0.3 g
Cysteine	0.3 g
Glutamic acid	0.3 g
Glutamine	0.3 g
Glycine	0.3 g
Histidine	0.3 g
Isoleucine	0.3 g
Leucine	0.3 g
Lysine	0.3 g
Phenylalanine	0.3 g
Proline	0.3 g
Serine	0.3 g
Threonine	0.3 g
Tryptophan	0.3 g
Valine	0.3 g

Table 21: Amino acid solution without methionine

Minimal medium	1 L
5 x M9 salts	200 mL
1M $MgSO_4$	2 mL
20% glucose	20 mL
100 mM $CaCl_2$	1 mL
Amino acid solution	50 mL
Biotin	10 mg
10 mg/mL thiamine	1 mL
50 mM tyrosine	10 mL
Antibiotics	Appropriate

Table 22: Minimal medium

5 x M9 salts	500 mL
Na_2HPO_4 x $2H_2O$	22 g
KH_2PO_4	7.5 g
NaCl	1.25 g
NH_4Cl	2.5 g

Table 23: 5 x M9 salts

2.6.2. Medium for eukaryotic cell culture

Freezing medium	
Medium	50%
Foetal bovine serum	40%
DMSO	10%

Table 24: Freezing medium

Erythrocyte lysis buffer, pH 7.3	
NH_4Cl	150 mM
$KHCO_3$	10 mM
EDTA	0.1 mM

Table 25: Erythrocyte lysis buffer

Benzidine staining solution (Gallicchio and Murphy 1979)	
Acetic acid	500 mM
Benzidine	0.2%
30% H_2O_2	0.4% (directly prior to staining)

Table 26: Benzidine staining solution

Methylcellulose (1500 cps) in Iscove's Modified Dulbecco's Medium (IMDM)	1.4%
Foetal bovine serum	15%
Bovine serum albumin	2%
L-glutamine	2 nM
2-mercaptoethanol	50 µM
Recombinant human insulin	10 µg/mL
Human transferrin	200 µg/mL
Recombinant mouse SCF	50 ng/mL
Recombinant mouse IL-3	10 ng/mL
Recombinant mouse IL-6	10 ng/mL

Table 27: Methylcellulose in IMDM medium, cat. number: HSC008, (R&Dsystems® 2015)

2.7. Enzymes

Enzymes	Supplier
Calf intestine alkaline phosphatase (CIAP)	Fermentas
Enterokinase, bovine, recombinant	Sigma-Aldrich
Lysozyme	Sigma
Phusion DNA-polymerase	New England Biolabs
Restriction enzymes	New England Biolabs / Fermentas
Shrimp alkaline phosphatase (SAP)	Fermentas
T4 DNA ligase	New England Biolabs
T4 polynucleotide kinase (PNK)	New England Biolabs

Table 28: Enzymes

2.8. Standards and kits

Standards and kits	Supplier
Bovine serum albumin standard	Thermo Scientific
Gel loading dye, blue (6 x)	New England Biolabs
Gene ruler DNA ladder mix	Fermentas
HisDetector™ Western blot kit, AP colorimetric	KPL
Micro BCA ™ protein assay kit	Thermo Scientific
MinElute reaction clean up kit	Qiagen
PageRuler prestained protein ladder	Fermentas
PageRuler unstained protein ladder	Fermentas
Plasmid midi kit	Qiagen
QIAprep spin miniprep kit	Qiagen
QIAquick gel extraction kit	Qiagen

Table 29: Standards and kits

2.9. Oligonucleotides and plasmids

2.9.1. Primers

All oligonucleotides were purchased from Metabion. The sequences are depicted from 5'- to 3'-end.

2.9.1.1. PCR-Primers

Name	Sequence
EPO-HPG-his-for	CGTGGACATATGGATGATGATGATAAAGCACCGCCTCG
EPO-HPG-his-rev	CGAGCTGGATCCTTATTAATGATGGTGGTGATGATGACG
His-EPO-HPG-for	AGTGGACATATGCATCATCATCACCATCACGATGAT
His-EPO-HPG-rev	GACGATGGATCCTTATTAACGATCACCGGTACGACATGC
NdeI-MKAhisEPO-for	ATACATATGAAAGCACATCATCATCACCATCAC
NcoI-MAKhisEPO-for	GAGATATACCCATGGCAAAACATCATCATCACCATCAC

Table 30: PCR-Primers

2.9.1.2. Primers for site-directed mutagenesis

Name	Sequence
Lys24TAG-for	AAGAAGCCGAATAGATTACCACCGGTTGTGCA
Lys24TAG-rev	T TGCTTCCAGCAGATAACG
Lys83TAG-for	TGGTTTAGAGCAGCCAGCCGTGGGAA
Lys83TAG-rev	GCAGTGCCTGACCACGCAGAACT
Lys38TAG-for	CTGAATGAATAGATTACAGTGCCGGATACC
Lys38TAG-rev	GCTACAATGTTCTGCACAA
Lys24Met-for	GAAGCCGAAATGATCACCACCGGTTGTGCA
Lys24Met-rev	TTTTGCTTCCAGCAGATAACGTTC
Lys83Met-for	CTGGTTATGAGCAGCCAGCCGTGGGAACC
Lys83Met-rev	CAGGGCCTGACCACGCAGAACTGCTTC
Lys38Met-for for template 24Met	CTGAATGAAATGATCACCGTTCCGGATACCAAAGTGAAC
Lys38Met-for for template 83Met	CTGAATGAGATGATCACCGTTCCGGATACCAAAGTGAAC
Lys38Met-rev	GCTACAATGTTCTGCACAACCGGTGG

Table 31: Primers for site-directed mutagenesis

2.9.1.3. Sequencing Primers

Name	Sequence
pET-RP	CTAGTTATTGCTCAGCGG
SpHI-FP	AAGATCGGGCTCGCCACTT
SpHI-RP	CCTGCATTAGGAAGCAGC
DuetDOWN1	GATTATGCGGCCGTGTACAA
DuetUP2	TTGTACACGGCCGCATAATC
T7-Terminator	GCTAGTTATTGCTCAGCGG
ACYCDuetUP1	GGATCTCGACGCTCTCCCT

Table 32: Sequencing Primers

2.9.2. Plasmids

Plasmid	Resistance	Characteristics	Supplier	Reference
pET-11a	AmpR	T7-promotor, lacI	Novagen	
pRSFDuet-1	KanR	Two MSCs, T7-promotor, lacI	Novagen	
pETDuet-1	AmpR	Two MSCs, T7-promotor, lacI	Novagen	
pTARA	CamR	Insert: T7 RNA polymerase	Addgene	(Wycuff and Matthews 2000)

Table 33: Plasmids

2.10. *E.coli* strains

Cells	Genotype	Supplier
BL21 (DE3)	*E. coli* B F– *dcm ompT hsdS*(r_{B-} m_{B-}) *gal* λ(DE3)	Stratagene
XL10-Gold (Ultracompetent)	Tet^rΔ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F' *proAB* $lacI^q$ ZΔ*M15* Tn*10* (Tet^r) Amy Cam^r]	Stratagene
B834 (DE3)	F^- *ompT* $hsdS_B$(r_B^- m_B^-) *gal dcm met* (DE3)	Novagen

Table 34: *E. coli* strains

2.11. Human cell lines

Cells	Cell type	Supplier
TF-1	Erythroleukemia	DSMZ
5637	Urinary bladder carcinoma	DSMZ

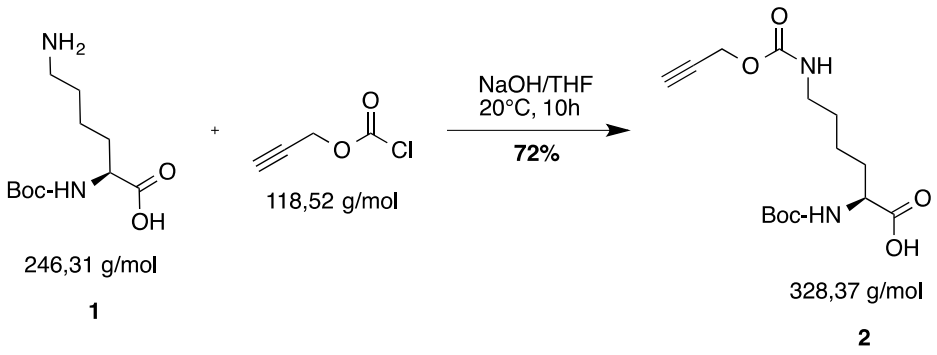
Table 35: Human cell lines

3. Methods

3.1. Synthesis of non-natural amino acids and other reagents for click chemistry

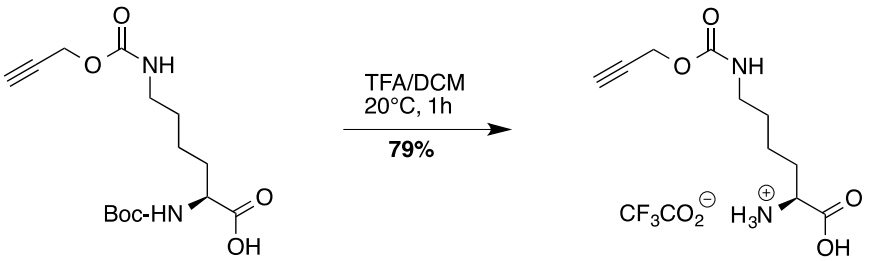
3.1.1. Synthesis of Plk, a pyrrolysine derivative with an alkyne functionality

N^2-(tert-butoxycarbonyl)-N^6-((prop-2-yn-1-yloxy)carbonyl)-L-lysine (**2**)



10 g of Boc-Lys-OH (**1**) (40.6 mmol) were used in one approach and were dissolved in 100 mL of 0.5 M NaOH and 50% THF (v/v). After one hour of stirring the solution was cooled down to 0 °C in an ice bath and 3.2 mL of 4 °C cold propargyl chloroformate (32,4 mmol) were added dropwise over ten minutes and warmed overnight to RT. The next day, the solution was cooled down to 0 °C, washed with ice cold Et_2O and acidified by ice cold 1 M HCl. Subsequently, the product was extracted with ice cold EtOAc. The organic layers were dried over $MgSO_4$ and the solvents were evaporated to clean give Boc-Plk (**2**) (9.6 g, 29.2 mmol) as a white foam in 79% yield (literature 83% (Nguyen, Lusic et al. 2009)).

N^6-((prop-2-yn-1-yloxy)carbonyl)-L-lysine TFA salt (Plk) (**3**)

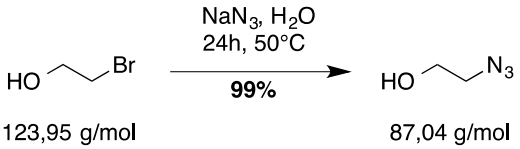


2 **3**

9.6 g of the propargyl carbamate (**2**) (29.24 mmol) was dissolved in 96 mL dry DCM and 96 mL of TFA was added dropwise. It was allowed to stir for one hour. The solvents were evaporated and the product was precipitated through addition of Et_2O and incubation at 4 °C overnight. Lastly, it was filtered and dried, affording clean Plk TFA salt (**3**) (7.86 g, 22.96 mmol) as a white solid in 79% yield (literature 96% (Nguyen, Lusic et al. 2009)). ^{1}H NMR (400 MHz, D_2O): δ (ppm) = 4.65-4.76 (m, 2H), 4.00 (t, *J* = 6.3 Hz, 1H), 3.19 (t, *J* = 6.7 Hz, 2H), 2.93 (t, *J* = 2.4 Hz, 1H), 1.88-2.06 (m, 2H), 1.38-1.64 (m, 4H).

3.1.2. Synthesis of Pln, a pyrrolysine derivative with an azide functionality

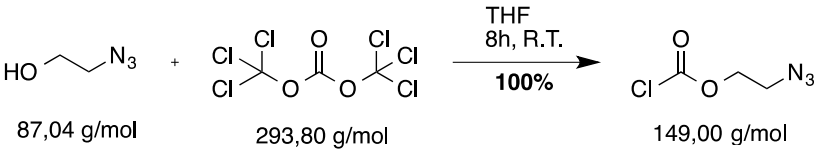
2-Azidoethanol (**4**)



4

4 g of 2-bromoethanol (32 mmol) were dissolved in 45 mL of H_2O. 4.16 g of NaN_3 (64 mmol) were added and the solution was stirred at 50 °C for 24 h. After cooling down to RT, it was extracted with Et_2O (3 x 50 mL). The combined organic layers were allowed to stir for 7 h in the presence of $MgSO_4$. Lastly, it was filtered and the solvents were evaporated, affording clean 2-azidoethanol (**4**) (2.76 g, 31.73 mmol) as a clear liquid in 99% yield (Li, Yuan et al. 2015).

2-Azidoethyl chloroformate (**5**)

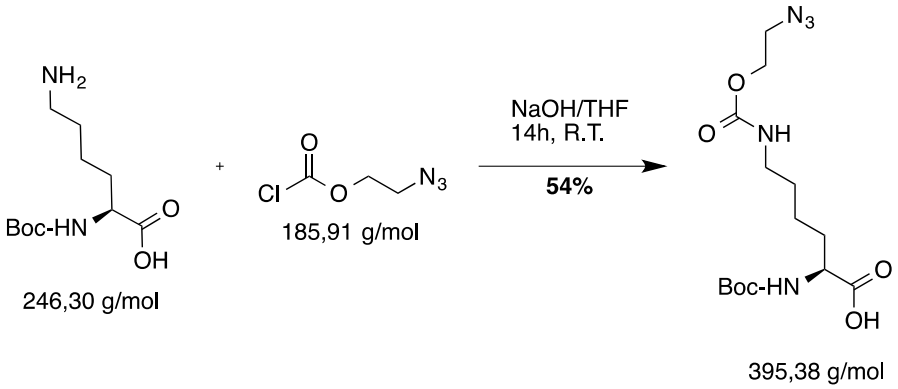


4 **5**

1.47 g of triphosgene (16.82 mmol) were dissolved in 30 mL of dry THF and cooled down to 0 °C under N_2 atmosphere. 1.47 g of 2-azidoethanol (**4**) (16.82 mmol) was added dropwise for 10 min and

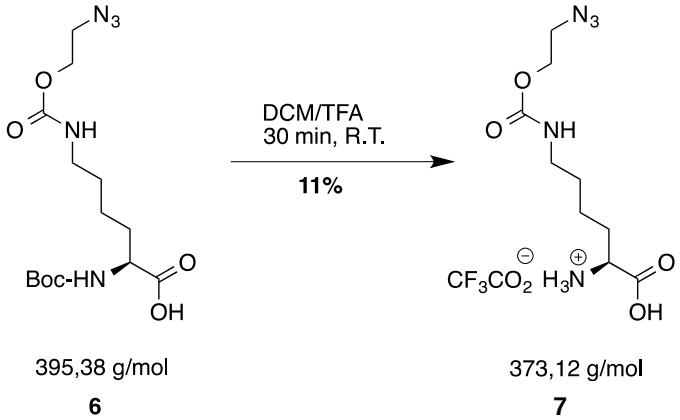
the solution was allowed to stir at RT for 8 h. The solvent was evaporated and the white solid was dried for 1 h. The product (**5**) was used directly for the next step (Nguyen, Lusic et al. 2009).

N^6-((2-azidoethoxy)carbonyl)-N^2-(tert-butoxycarbonyl)-L-lysine (**6**)



3.14 g of 2-azidoethyl chloroformate (**5**) (16.88 mmol) were dissolved in 4.4 mL of THF and 5 g of Boc-Lys-OH (**1**) (20.3 mmol) were dissolved in a solution containing 58.8 mL of 1 M NaOH and 17.4 mL of THF (4:1). At 0 °C, the 2-azidoethyl chloroformate solution was added dropwise to the Boc-Lys-OH solution and was stirred for 12 h slowly warming to RT. Then, it was cooled again to 0 °C, acidified to pH 2-3 by ice-cold 1 M HCl solution. The aqueous layer was extracted with 100 mL of EtOAc and the organic layer was subsequently washed with brine (5 x 100 mL), dried over $MgSO_4$ and filtered. The solvent was evaporated affording product (**6**) (3.6 g, 9.11 mmol) as a yellowish oil in 54% yield (literature: 80% (Nguyen, Lusic et al. 2009)).

N^6-((2-azidoethoxy)carbonyl)-L-lysine TFA salt (Pln) (**7**)



3.6 g of compound **6** (9.11 mmol) was dissolved in 72 mL of DCM/TFA (1:1) and was stirred at RT for 30 min. The solvents were evaporated under vacuum. The residual oil was re-dissolved in 5 mL of MeOH, precipitated into Et_2O and filtered. The precipitation was repeated several times and the collected white solid was dried under vacuum, affording product **7** (704 mg, 1.89 mmol) in 11% yield

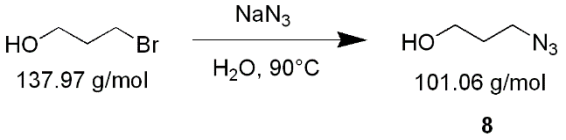
(literature 93% (Nguyen, Lusic et al. 2009)). ^{1}H NMR (400 MHz, D_2O): δ (ppm) = 4.30 (t, J = 4.8 Hz, 2H), 3.77-4.04 (m, 1H), 3.58 (t, J = 4.6 Hz, 1H), 3.11-3.27 (m, 2H), 1.86-2.09 (m, 2H), 1.40-1.65 (m, 4H).

3.1.3. Supply of other non-natural amino acids

L-Homopropargylglycine (HPG) was purchased from Santa Cruz Biotechnology.

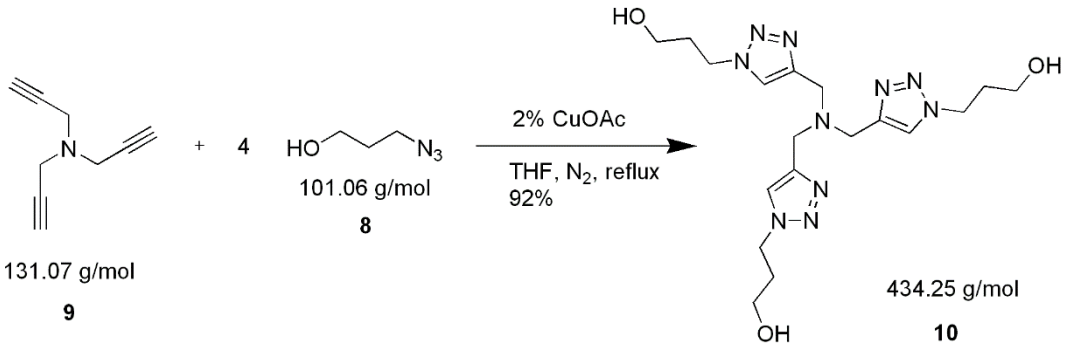
3.1.4. Synthesis of THPTA, a water-soluble ligand for click reactions

3-Azidopropan-1-ol (**8**)



20 g of 3-bromopropan-1-ol (144 mmol) and 18.7 g of sodium azide (288 mmol) were dissolved in 150 mL water and the resulting solution was stirred at 90 °C overnight. The mixture was extracted with dichloromethane (3 x 150 mL). The combined organic layers were dried over $MgSO_4$ and concentrated by rotary evaporation under reduced pressure using a bath kept at RT, to obtain 3-azidopropan-1-ol (**8**) as a pale yellow oil (11.9 g) in 82% yield (Hong, Presolski et al. 2009).

Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) (**10**)



Tripropargylamine (**9**) (27 mmol) and 3-azidopropan-1-ol (**8**) (108 mmol) was dissolved in 90 mL of THF and stirred under N_2 atmosphere. Cuprous acetate (2 mol%) was added and the solution was refluxed overnight under inert atmosphere. The mixture was concentrated, dissolved in 50 mL of water, and stirred with 2-3 g of Cuprisorb resin to remove the copper ions. The solution was filtered, the solid washed, and the combined solutions concentrated under vacuum to provide a yellow oil which solidified under high vacuum. The resulting yellow solid was dispersed in acetonitrile, sonificated to further break up the solid, filtered, washed with acetonitrile, and dried under vacuum to obtain 10.8 g of THPTA (**10**) as an off-white solid in 92% yield (Hong, Presolski et al. 2009).

3.2. Molecular biological methods

Methods like restriction digest, PCR and all types of gel-electrophoresis were performed according to the protocols listed in (Sambrook and Russell 2001) under consideration of the producer's instructions for the respective equipment and substances.

3.2.1. Plasmid preparation and DNA concentration measurement

The QIAprep Spin Miniprep kit was used to isolate plasmids out of *E. coli* cells. The culture size was 5 mL. For larger amounts (50 mL), the plasmid midi kit was used according to the manufacturer's instructions. The exact quantity of DNA could be determined by measuring the extinction at 260 nm. The Lambert-Beer's law allows the direct quantification via UV spectrometry, as the UV extinction is directly linear to the DNA concentration. DNA absorbs UV radiation from 250 to 270 nm with its maximum at 260 nm. The following formula was used for quantification:

$$\text{dsDNA concentration in } \mu g/\mu L = OD_{260\,nm} \text{ x dilution factor x } 0.05$$

The quotient OD_{260} / OD_{280} was checked in order to estimate the DNA purity. A value from 1.8 to 2.0 indicates high purity, whereas protein or phenol contaminations cause a lower and RNA contaminations a higher value (Sambrook and Russell 2001).

3.2.2. Sequencing

5 µL of the appropriate primer (5 µM) was mixed with 5 µL plasmid DNA (80 – 100 ng/mL) and was sent to GATC Biotech AG for sequencing analysis.

3.2.3. Agarose gel electrophoresis

For separation of DNA molecules, a gel-electrophoresis was performed on a 0.8% agarose gel in TAE buffer, which is also the running buffer (see table 5 in chapter 2.5). To each DNA sample, the appropriate volume of 6 x gel loading dye was added. 3 µL of the appropriate DNA ladder were filled in a separate pocket allowing the determination of DNA sample sizes. The voltage for the electrophoresis was set to 110 V. The agarose gels were stained with ethidium bromide (0.5 µg/mL) and documented under UV light with chemidoc XRS.

3.2.4. Restriction digest

A restriction digest of double stranded DNA is useful to obtain cohesive ends for cloning a DNA substrate into a vector. Position and direction of the insert within the vector are controlled by choosing two restriction endonucleases that are not isoschizomeric. A double or single digest was performed for two hours at 37 °C. The amount of units, buffer conditions and inactivation could be learnt from the manufacturer's protocol.

3.2.5. Ligation

Plasmids for transformation into *E. coli* cells were generated by ligation of an insert with a vector. Ligation was performed overnight at 16 °C with 1 µL of T4 DNA ligase (400 u/µL). 100 ng of vector and three times the amount in mol of the insert were incubated with the ligase in the appropriate buffer. In general, the reaction volume was 20 µL. A negative control without insert was always treated exactly under the same conditions in order to check, if re-ligation of the vector occurred. A small aliquot of the reaction was loaded onto a 0.8% agarose gel as a control. After ligation, the reaction samples were heated at 65 °C for 10 min to inactivate the enzyme and the plasmid concentration was measured.

3.2.6. Site-directed mutagenesis

In order to mutate one single triplet of the EPO gene, primers were designed (listed in table 31 in chapter 2.9.1.2), as it is described in the protocol of (Hsieh and Vaisvila 2013). In table 36, the components for gradient PCR are listed.

Component	Stock concentration	Volume for 100 µL
Template	100 ng/µL	1 µL
dNTPs	10 mM	2 µL
Forward primer	10 µM	5 µL
Reverse primer	10 µM	5 µL
HF buffer	5 x	20 µL
HF-Phusion DNA Polymerase	2 u/µL	1 µL

Table 36: Components for site-directed-mutagenesis PCR

The 100 µL master mix was divided into four 25 µL reaction samples with four different annealing temperatures during gradient PCR (table 37). A small aliquot of each PCR approach was loaded onto an agarose gel to determine the best annealing conditions. The respective template DNA was digested by DpnI for 1h at 37 °C. (DpnI was subsequently heat inactivated at 80 °C for 20 min.)

Step	Temperature	Time	Cycles
Denaturation	98 °C	30 s	
Denaturation	98 °C	10 s	x 30
Annealing	65 °C – 85 °C	20 s	
Extension	72 °C	3.5 min	
Final extension	72 °C	10 min	

Table 37: Gradient PCR program for site-directed mutagenesis

The PCR product was phosphorylated by T4 polynucleotide kinase (PNK) for 1h at 37 °C (table 38). PNK was heat inactivated for 20 min at 65 °C, prior to addition of 1 µL T4 DNA ligase for overnight ligation at 16 °C, as it is described in 3.2.5.

Components	Stock concentration	Volume for 20 µL
PCR product	> 5 ng/µL	7 µL
T4 DNA ligase buffer	10 x	2 µL
T4 polynucleotide kinase	10 u/µL	1 µL
ddH_2O		10 µL

Table 38: Approach for PNK phosphorylation

3.2.7. Construction of the expression plasmids

3.2.7.1. Construct for expression of unmodified EPO

The codon-optimised EPO gene with a C-terminal polyhistidine-tag (six consecutive histidine residues), flanked by restriction sites NdeI and BamHI, was inserted into the cloning site of a pET11a vector under the control of a T7 promoter. (For information on the EPO sequences, see chapter 6.1.1.) pET11a has a carbenicillin resistance. The plasmid was electro-transformed into *E. coli* BL21 (DE3).

3.2.7.2. Constructs for amber stop codon suppression system

The codon-optimised EPO genes with a C-terminal polyhistidine-tag, the pylRS gene and the $tRNA^{Pyl}$ sequence were ordered from Geneart® (Thermo Fisher Scientific). (For information on the EPO sequences, see chapter 6.1.1.) The constructs were flanked by the restriction sites NdeI and BamHI in order to clone them into the multiple cloning site (MCS) under control of a T7-promotor of the pET-11a vector (figure 24). The sequence for the $tRNA^{Pyl}$ from *Methanosarcina barkeri* under the control of a consecutive lpp-promoter was also inserted into pET-11a using the restriction site SphI. On pRSFduet, the corresponding pylRS was inserted into the first MCS under the control of a T7-promoter between the restriction sites NcoI and SacI. The TAG stop codons were introduced into the EPO gene by site-directed mutagenesis with the primers listed in table 31 in chapter 2.9.1.2.

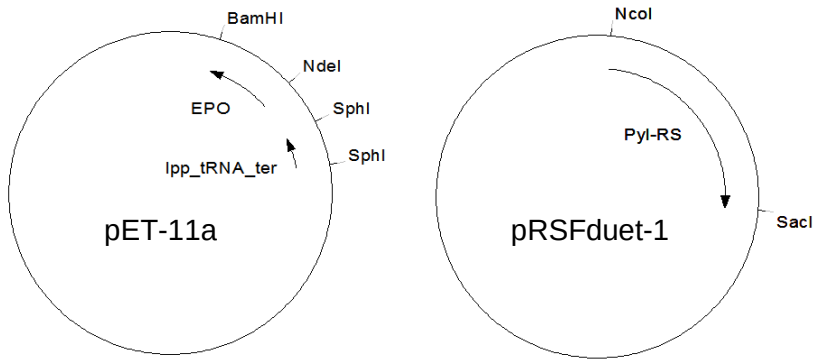


Figure 24: Construction of the expression plasmids

3.2.7.3. Constructs for selective pressure incorporation system

The codon-optimised EPO genes with an N-terminal polyhistidine-tag were ordered from Geneart® (Thermo Fisher Scientific). (For information on the EPO sequences, see chapter 6.1.1.) The three ATG codons were introduced at positions 24, 38 and 83 by site-directed mutagenesis with the primers listed in table 31 in chapter 2.9.1.2. To transfer the gene, a PCR with primers listed in table 30 in chapter 2.9.1.1 was performed, whereby the restriction sites NdeI or NcoI and BamHI were added in order to clone them into the MCS under control of a T7-promotor of pRSFDuet-1, pETDuet-1 or pET-11a. The plasmid pTARA with the insert T7 RNA polymerase was a gift from Kathleen Matthews (see table 33 in 2.9.2).

3.3. Microbiological methods

3.3.1. Preparation of chemical competent *E. coli*

$CaCl_2$-competent *E. coli* XL10 were prepared as follows. 200 mL of LB medium was inoculated 1:200 with an overnight culture and grown shaking in an incubator at 37 °C to an $OD_{600\ nm}$ of 0.6 (see table 18 in chapter 2.6.1). The cells were pelleted by centrifugation (4000 rpm, 4 °C, 20 min) and re-suspended in 160 mL of an ice cold, sterile 100-mM $MgCl_2$ solution. After another centrifugation step, the cells were re-suspended in 80 mL of an ice cold, sterile 50 mM $CaCl_2$ solution and centrifuged again. Lastly, they were re-suspended in 8 mL of solution containing 50 mM $CaCl_2$ and 15% glycerol. 100 µL aliquots of the cell suspension were stored at -80 °C.

3.3.2. Chemical transformation

A 100 µL aliquot of $CaCl_2$-competent *E. coli* XL10 was thawed on ice and put into a 1 mL reaction tube together with 5 µL plasmid DNA on ice for 30 min. For 5 min, the mixture was heated to 37 °C and 1 mL of pre-warmed SOC medium was added (see table 20 in chapter 2.6.1). The incubation at 37 °C was prolonged and the cell suspension was spread over LB agar plates with the appropriate antibiotics (see table 19 in chapter 2.6.1).

3.3.3. Preparation of electro-competent *E. coli*

Electro-competent *E. coli* BL21 (DE3) cells were used for transformation. They were prepared after Ragsdale (Dower, Miller et al. 1988). 200 mL of LB medium was inoculated 1:200 with an overnight culture and grown shaking in an incubator at 37 °C to an OD_{600nm} of 0.6 (see table 18 in chapter 2.6.1). The culture was cooled down on ice for 30 min and centrifuged at 4000 rpm for 10 min at 4 °C. The pellet was re-suspended in 50 mL of ice cold sterile water and put on ice for 15 min. This washing step was repeated three times with reduced volumes of water (20 mL and 10 mL). Lastly, the pellet was re-suspended in 2 mL of ice cold 10% glycerol and the cell suspension was distributed in 100 µL aliquots that were frozen immediately at -80 °C.

3.3.4. Electro-transformation

2 µL of each plasmid were mixed with a 100 µL aliquot of electro-competent *E. coli* BL21 (DE3), which had been thawed on ice. The mixture was transferred into an ice cold 1-mm cuvette for electroporation. The Biorad gene pulser X cell produced a pulse of 1.8 kV. Immediately, the cells were transferred into 1 mL of 37 °C warm SOC medium and incubated at 37 °C for 1 h (see table 20 in chapter 2.6.1). Subsequently, they were spread on LB agar plates containing the appropriate antibiotics and incubated at 37 °C overnight (see table 19 in chapter 2.6.1).

3.4. Protein biochemical methods

3.4.1. Expression methods and incorporation of non-natural amino acids

3.4.1.1. Expression of human recombinant wt EPO

An overnight culture was used to inoculate several litres of LB containing 100 µg/mL carbenicillin (see table 18 in chapter 2.6.1). The culture was grown at 37 °C until an OD_{600} of one was reached. Protein expression was induced by adding IPTG to a final concentration of 1 mM. The incubation was prolonged overnight. The next day, the cells were harvested at 4000 rpm for 20 min at 4 °C. The supernatant was discarded and the pellets could be frozen at -20 °C until further use.

3.4.1.2. Amber stop codon suppression

The EPO gene with a C-terminal polyhistidine-tag and an amber stop codon (TAG) at specific positions (for information on the EPO sequences, see chapter 6.1.1) was inserted into the cloning site of a pET11a vector under the control of a T7 promoter. Moreover, the sequence for the $tRNA^{Pyl}$ from *Methanosarcina barkeri* under the control of a consecutive lpp-promoter was also cloned into pET11a. On a second plasmid pRSFduet with a kanamycin resistance, the corresponding PylRS was inserted into the first cloning site under the control of the T7 promoter.

The two plasmids were electro-co-transformed into *E. coli* BL21 (DE3). An overnight culture was used to inoculate several litres of LB containing carbenicillin (100 µg/mL) and kanamycin (34 µg/mL) (see table 18 in chapter 2.6.1). The culture was grown at 37 °C until an OD_{600} of one was reached. Then, defined concentrations of Plk or Pln were added to the culture and the incubation was prolonged for 15 min, in order to allow complete dissolving of the unnatural amino acid and its uptake into the cells. Protein expression was then induced adding IPTG to a final concentration of 1 mM. The incubation was prolonged overnight. The next day, the cells were harvested at 4000 rpm for 20 min at 4 °C. The supernatant was discarded and the pellets could be frozen at -20 °C until further use.

3.4.1.3. Selective pressure incorporation

The EPO gene with a C-terminal or N-terminal polyhistidine-tag and the ATG-codon at defined positions was inserted into the cloning site of a pET-11a vector, pRSFDuet-1 or pETDuet-1 under the control of a T7 promoter (for information on the EPO sequences, see chapter 6.1.1). The naturally occurring methionine at position 54 was mutated to leucine and a cleavable tag was included after the start codon and the N-terminal polyhistidine-tag, as the unnatural amino acid would also be incorporated at these undesired positions. The plasmid was transformed into methionine auxotroph B834 cells.

An overnight culture was grown in minimal medium supplemented by 0.1 mM methionine and was used to inoculate a defined volume of minimal medium supplemented by 30 µM methionine (see table 22 in chapter 2.6.1). The culture was incubated at 37 °C until the cells reached the stationary phase because of depletion of methionine. This was determined by measuring the OD each hour. Then, the culture was pelleted by centrifugation, re-suspended in fresh minimal medium and divided

into three separate cultures. 1 mM methionine was added to the positive control and no methionine was added to the negative control. The non-natural amino acid HPG (1mM) was added to the third culture. The incubation was prolonged overnight and the next day, the cells were centrifuged. The supernatant was discarded and the pellets could be frozen at -20 °C until further use.

3.4.2. Solubilisation of inclusion bodies

EPO is expressed insolubly within inclusion bodies, which need to be disintegrated (Wang, Liu et al. 2010). The pellets were re-suspended in cell lysis buffer and incubated on ice for 30 min (see table 13 in chapter 2.5). Subsequently, they were treated with ultrasound four times for 1 min, 5 cycles, 20%, with a pause of 2 min between each sonication step. To get rid of other cell material, it was ultra-centrifuged at 16000 x g for 20 min at 4 °C and the supernatant was discarded. The pellets were washed once with 20 mM Tris-HCl, pH 8.0. After a second ultra-centrifugation step, they were re-suspended in denaturation buffer (see table 14 in chapter 2.5), treated with a Potter-Elvehjem homogenizer and stirred overnight at RT to obtain soluble EPO, which is now in the supernatant after a last ultra-centrifugation step at 16000 x g for 40 min.

3.4.3. Protein purification methods

3.4.3.1. Affinity-tag purification

The supernatant was incubated with prepared Ni-NTA agarose beads for 2 hours or overnight at 4 °C in an overhead-shaker. As a preparation, the slurry Ni-NTA agarose beads hab been centrifuged in 15 mL or 50 mL vials at 1000 x g for 5 min. The ethanol was removed and the beads were washed three times with Ni-NTA purification buffer (see table 15 in chapter 2.5). After incubation, the loaded Ni-NTA beads were transferred into a 15 mL or 50 mL column with a 20 µm frit and were washed with Ni-NTA purification buffer and subsequently with 20 mM of imidazole added. The proteins labelled by a polyhistidine tag were eluted with a stepwise imidazole gradient. 10 mL of 50 mM, 100 mM, 200 mM, 300 mM and 500 mM of imidazole in Ni-NTA purification buffer were used and 2 mL fractions were collected from the elution, which were then analysed by spectrophotometer and SDS-PAGE.

3.4.3.2. Size-exclusion chromatography (SEC)

All EPO samples for SEC were and centrifuged at 10,000 rpm for 5 min at 4 °C. The column Superdex 75 10/300 GL was prepared on a FPLC system by equilibrating it with gel filtration buffer (see table 16 in chapter 2.5). The flow rate of the run was 0.3 mL/min and the fraction volume was either 0.5 mL or 0.3 mL. The peak fractions were analysed by spectrophotometer and SDS-PAGE.

3.4.4. Refolding of EPO

Affinity-tag purified EPO in denaturing Ni-NTA purification buffer was concentrated to a protein concentration of 5 – 10 mg/mL. 1 – 2 mL of protein solution were diluted 1:200 to an end concentration of 25 – 50 µg/mL in 4 °C cold refolding buffer (see table 17 in chapter 2.5) by first

pipetting the EPO solution onto the bottom of a flask and then, pouring over the refolding buffer in order to avoid guanidine micelles. It was incubated at 4 °C overnight to enable refolding. Subsequently, it was concentrated to the smallest volume possible by the use of Vivaspin® centrifugal concentrators.

3.4.5. Protein concentration determination

3.4.5.1. Photometrical determination by nanodrop spectrophotometer

The protein concentration in mg/mL was determined by measurement of the absorbance at 280 nm. The nanodrop spectrophotometer was used. The concentration was calculated as followed, where ε is the specific extinction coefficient:

$$c\ [mg/mL] = \frac{Abs_{280}}{\epsilon}$$

EPO has an extinction coefficient of $\epsilon = 1.24$ L/(mol*cm). This value was used regardless of glycosylation to obtain the concentration of the protein part only.

3.4.5.2. Micro BCA protein assay

If protein concentration determination was not possible by absorbance measurement, because of concentrations lower than 0.1 mg/mL, a micro BCA protein assay in 96-well plates was performed with usage of the kit, following the instructions. As a standard, the included BSA was used.

3.4.6. Biochemical protein characterisation

3.4.6.1. SDS-PAGE

The samples which should be analysed by glycine SDS-PAGE were mixed with 6 x sample buffer and denatured at 95 °C for 5 min (see table 7 in chapter 2.5). If they contained whole cells, they were treated with ultrasound for 30 s, 5 cycles, 20%, prior to heat denaturation. For native PAGE analysis, 2-mercaptoethanol was omitted in the 6 x sample buffer and the protein samples were not denatured. Then, the samples were loaded onto a gel that consists of a stacking (above) and a resolving part (below). For composition details, see table 8 in chapter 2.5. The gel-electrophoresis was conducted in glycine SDS electrophoresis buffer at a constant voltage of 200 V (see table 6 in chapter 2.5).

3.4.6.2. Sarcosyl-PAGE

For analysis of proteins coupled to PEG-20kDa, a Sarcosyl-PAGE after Reichel was done (Reichel 2012). Therefore, 15 µL of the sample was mixed with 5 µL of sarcosyl sample buffer and 1 µL reducing agent (2 M DTT) and boiled at 95 °C for 5 min. The gels were put between two different buffers. The catholyte solution consists out of 200 mL sarcosyl running buffer and 500 µL antioxidant, which is prepared immediately before use: 380 mg sodium thiosulfate in 1 mL water. The anolyte

solution is pure sarcosyl running buffer. For composition details of sarcosyl running and sample buffer, see tables 6 and 7 in chapter 2.5.

3.4.6.3. Coomassie staining

The gels were stained by Coomassie Brilliant Blue R-250 solution for 1h and subsequently, were destained by a solution of 50% methanol and 10% acetic acid for 3 h. The gels were scanned for documentation.

3.4.6.4. Western blot

For specific detection of polyhistidine tagged proteins using Ni-NTA conjugates, a Western blot was performed after SDS electrophoresis. Therefore, the PVDF Immobilon-P transfer membrane was activated in methanol for 5 min before incubating it together with the gel in transfer buffer (see table 9 in chapter 2.5) for 10 min. Then, a wet blot was performed. The "sandwich" was built in that way, that the gel is on the negative pool in close contact to the membrane on the positive pole. Air bubbles were avoided. The blot was conducted at 300 mA for 35 min. Subsequently, the HisDetector™ Western blot kit, AP colorimetric was used and the detection was performed following the instructions of KPL. The HisDetector™ Nickel-AP was diluted 1 to 2000. The membrane was washed with TBS-Tween and AP buffer (see tables 10 and 11 in chapter 2.5). Finally, the membrane was incubated with a detection solution containing NBT and BCIP in AP buffer for 5 to 10 min in the dark (see table 12 in chapter 2.5). After drying, the membrane was photographed for documentation.

3.4.6.5. Circular dichroism spectrometry

Circular dichroism is an adequate technique for the examination of the secondary structure of EPO in solution. The CD spectrum was measured from 250 nm to 195 nm at a scanning speed of 200 nm/min with a data interval of 0.2 nm.

3.4.6.6. Melting curves

Melting curves of proteins were assessed by the same spectropolarimeter as for CD spectra (see chapter 3.4.6.5). The secondary structure peaks of EPO at 208 nm and 222 nm were measured while heating the sample up from 20 °C to 85 °C. The heating rate was 0.5 °C per min. For not glycosylated EPO, only the apparent melting point can be determined because the process is irreversible and the protein precipitates above 44 °C (Narhi, Arakawa et al. 1991).

3.5. Click chemistry

Various carbohydrates with an azide group (see figure 22 in chapter 2.4.2) were coupled to EPO with alkyne functionality by copper-(I)-catalysed azide-alkyne Huisgen cycloaddition (Kolb, Finn et al. 2001, Tornoe, Christensen et al. 2002). An overview is depicted in figure 25.

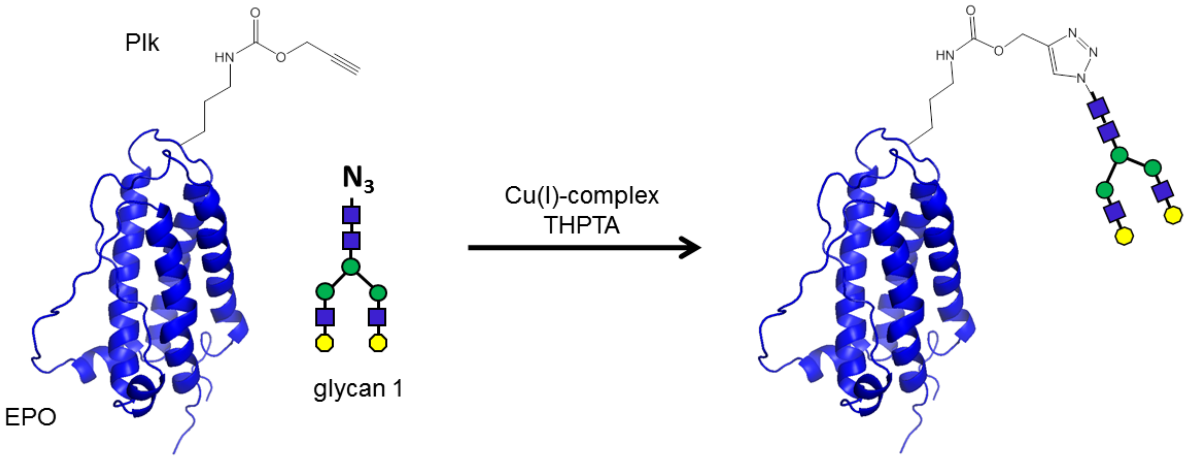


Figure 25: Schematic overview of the copper(I)-catalysed azido-alkyne Huisgen cycloaddition (click reaction)

In table 39, an example for a calculation of all components for click reaction is shown. All components were dissolved in water with the exception of Cu(I)-complex, which is dissolved in acetonitrile. It is important that the reaction takes place under N_2-atmosphere and that O_2 is excluded due to potentially oxidation of Cu(I) to Cu(II), which is detrimental for several components of the reaction. The Cu(I)-THPTA-premix is degassed three times before it is added to the protein click-ligand mixture, which is degassed only once because EPO is prone to precipitation under harsh conditions. Various protein : click-ligand ratios were tested. A ratio of 1:10, as depicted in table 39 was considered the best ratio for high EPO concentrations to keep the balance between a good turnover and not wasting the click-ligand. The reaction was processed for one hour at room temperature. Immediately, the samples were centrifuged at 10000 rpm for 5 min prior to purification via SEC (see chapter 3.4.3.2).

	stock in mg/mL	**stock in mM**	**conc in mM**	**vol in µL**
EPO-Plk	1.13	0.058	0.047	40
glycan-N_3		6	0.47	3.88
Cu(I)-complex		80	4	2.5
THPTA		400	8	1
Tris-buffer, pH 7.5		20		2.62
				50
Cu(I)-THPTA-premix	1 x	6 x		
Cu(I)-complex	2.5 µL	15 µL		
THPTA	1 µL	6 µL		
	3.5 µL	21 µL		

Table 39: Example of a click reaction approach

3.6. Cell culture and biological assays

3.6.1. Cell culture

Cell culture methods were performed according to the protocols listed in the laboratory handbook of fundamental techniques in cell culture (ECACC 2010) under consideration of the producer's instructions for the respective equipment and substances.

3.6.1.1. Resuscitation of frozen cell lines

An ampoule of cells frozen in liquid nitrogen was collected and quickly transferred to a 37 °C water bath for one or two minutes until only a small ice crystal remained. Subsequently, it was wiped with a tissue soaked in 70% alcohol prior to opening in the microbiological safety cabinet. The cells were transferred into 30 mL of pre-warmed cell culture medium and centrifuged for 5 min at 400 x g at RT. The pellet was re-suspended in the appropriate medium and transferred into a cell culture flask.

3.6.1.2. Suspension cell culture

The cell culture medium was exchanged after 48 to 72 h or when a yellow colour of the medium indicates an acidic pH due to phenol red. The total amount of viable cells was determined and the cell suspension was centrifuged for 5 min at 400 x g at RT. The appropriate volume of fresh medium was calculated, in which the cells were re-suspended for further incubation.

3.6.1.3. Adherent cell culture

After 48 or 72 h of cultivation, the medium was collected or discarded by carefully pipetting from one corner of the cell culture flask. The cells were washed with 10 mL PBS prior to incubation with 3 mL trypsin/EDTA for 5 to 10 min at 37 °C. The side of the flask might be tapped to release any remaining attached cells. An inverted microscope was used to examine, if all cells were detached and floated. Then, the trypsin reaction was stopped by adding 7 mL of medium. 10 µL of the cell suspension was removed to determine the amount of cells. The appropriate volume of the cell suspension was centrifuged for 5 min at 400 x g at RT and re-suspended in fresh medium and transferred into a cell culture flask for further incubation.

3.6.1.4. Cell quantification

Adherent cells were brought into suspension using trypsin/EDTA as previously described. Suspension cells were re-suspended in order to obtain a homogeneous cell suspension. 10 µL were removed under sterile conditions and 10 µL of Trypan blue (dilution factor: 2) was added and mixed by gentle pipetting. The chamber of a clean haemocytometer was filled with the cell suspension and viewed under an inverted phase contrast microscope using x 10 magnification. The number of viable cells (bright cells) was counted. The concentration of viable cells was calculated using the equation below:

$$\text{Viable cells per mL} = \frac{\text{number of live cells counted}}{\text{number of large corner squares } (1\ \text{mm}^2) \text{ counted}} \text{ x dilution x } 10000$$

3.6.1.5. *Cryopreservation of cell lines*

Adherent cells were brought to into suspension using trypsin/EDTA as described previously. Suspension cells could be used directly. The cells were counted, centrifuged at 400 x g for 5 min at RT and re-suspended at a concentration of 2 – 4 x 10^6 cells per mL in freezing medium (see table 24 in chapter 2.6.2). 1 mL aliquots were pipetted into cryoprotective ampoules. These were placed inside a pre-cooled (4 °C) passive freezer (Nalgene Mr Frosty box, Sigma), which was filled with isopropyl alcohol and put at -80 °C. After one to three days, the frozen ampoules were transferred into a liquid nitrogen storage vessel and the locations were recorded.

3.6.2. Cell differentiation assay

The bone marrow was isolated following the instructions from Mosser (Zhang, Goncalves et al. 2008). The ends of femurs and tibiae from mice were cut off with a pair of scissors. PBS was flushed through the bones to obtain the bone marrow. Therefore, a 1 mL syringe with a 27 gauge needle was used. The cell suspension was led through a cell strainer into a 50 mL vial and centrifuged at 500 x g for 7 min. The erythrocytes were lysed by incubating with erythrocyte lysis buffer for 5 min at RT and occasionally agitated (see table 25 in chapter 2.6.2) (Kruisbeek 2001). Subsequently, the bone marrow cells were washed twice with IMDM medium (Lonza) and counted.

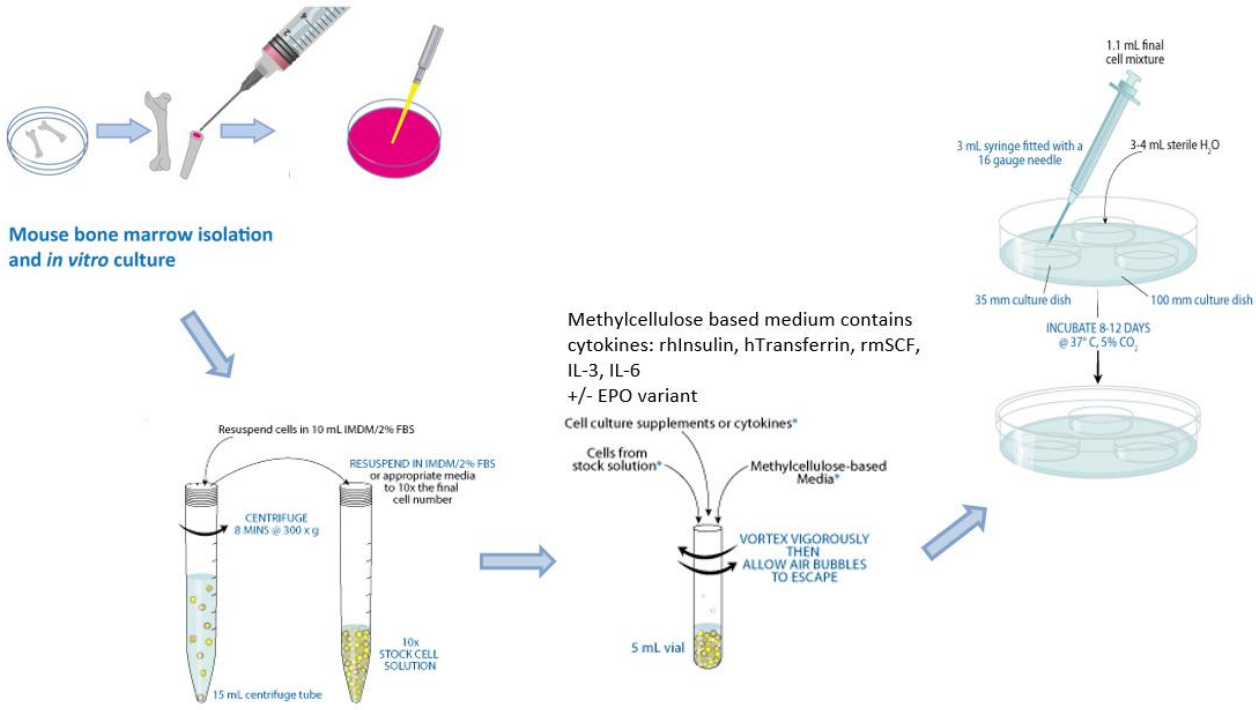


Figure 26: Schematic overview of the procedure for the cell differentiation assay with primary mouse bone marrow cells, modified from (R&Dsystems® 2015).

It was preceded following the instructions of Miller (Wognum, Yuan et al. 2013) and of R&Dsystems® (R&Dsystems® 2015). 400 µL of the cell suspension with a defined cell number (for example 1 x 10^5 per mL) was added to 4 mL methylcellulose complete medium in a 15 mL falcon (see table 27 in

chapter 2.6.2). Lastly, a solution with a defined EPO concentration was added. The vial was vigorously vortexed to thoroughly mix the cells with the media and after 20 min incubation at RT to allow air bubbles to escape, three times 1.2 mL (for triplicates) of the 4.4 mL final cell mixture is put into 35 mm culture dishes using a 5 mL syringe fitted with a 16 gauge needle. Two sample dishes and an uncovered dish containing 5 mL of sterile water were placed into one 100 mm culture dish and cover. The sterile water dish serves to maintain the humidity necessary for colony development. The cells were incubated for 8 days at 37 °C and 5% CO_2. The procedure is depicted as a flow chart in figure 26.

The different types of colonies were investigated under the microscope. For better visualization of haemoglobin-containing colonies, the plates were stained with benzidine staining solution for 30 min, which stains haemoglobin dark blue (see table 26 in chapter 2.6.2) (Gallicchio and Murphy 1979). The colonies were counted and pictures of the plates were taken.

3.6.3. Cell proliferation assay

The human erythroleukemia cell line TF-1 is described to be proliferately responsive to several haematopoietic growth factors including EPO (Kitamura, Tange et al. 1989).

The cultivation medium was RPMI 1640, supplemented with 16% heat inactivated foetal bovine serum (FBS), a mixture of penicillin and streptavidin, and 10% conditioned medium from cell line 5637. The latter contains several cytokines, which are obligatory for the survival of TF-1 cells.

The cell line 5637 derived from primary bladder carcinoma and was described to produce several growth factors, e.g. SCF, IL-1 and GM-CSF, but not EPO (Quentmeier, Zaborski et al. 1997). It was cultured in RPMI 1640 medium supplemented with 10% heat inactivated FBS. These cells are adherent, therefore, the supernatant could be easily removed after two to three days of culture. Subsequently, it was sterile filtrated and frozen at -20 °C until use.

3.6.3.1. alamarBlue® assay

A defined volume of TF-1 cell suspension was washed twice with cultivation medium without conditioned medium 5637 and the cell number was adjusted to 1 x 10^5 per mL. The cell suspension was distributed into the wells of a 96-well plate. The final cell number was 1 x 10^4 cells per well. After incubating the cells at 37 °C, 5% CO_2 for four hours without any cytokines, the dilutions of several EPO variants in triplicates were added to each well. Final EPO concentrations ranging from 0.001 to 1000 ng/mL were used. The alamarBlue® reagent was added avoiding light after 48 hours of incubation, and after further 48 hours, the fluorescence was read out (excitation wavelength: 560 nm, emission wavelength: 590 nm). The procedure is depicted as a flow chart in figure 27.

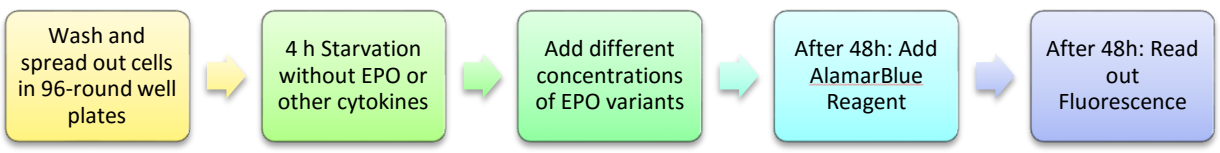


Figure 27: Overview of the procedure for the TF-1 proliferation assay

3.6.3.2. ^{3}H-thymidine assay

For convenient reasons, the ^{3}H-thymidine assay was performed on the same cells in 96-well plates, which were available from the alamarBlue® assay because this reagent is not toxic to the cells and the cells were still proliferating upon EPO stimulation. The cells were pulsed with 50 µCi per well, incubated at 37 °C, 5% CO_2 for 24 h and lysed by water. The DNA was transferred onto a fibreglass membrane, overlaid with scintillation agent and the radiation was measured by the TopCount® NXT™ microplate scintillation and luminescence counter.

4. Results and Discussion

4.1. EPO expression

4.1.1. Expression of wt EPO

The sequence of wt EPO (figures 61 and 62A-B in chapter 6.1.1) was cloned into the pET11a vector, which was transformed subsequently into *E. coli* strains BL21 (DE3) and B834 (DE3). The cell cultures were grown until an OD_{600} of one was reached, gene expression was induced by adding IPTG and occurred overnight. In figure 28, the expression profile is shown. The protein band with a molecular weight of about 19 kDa clearly shows overexpression of the protein.

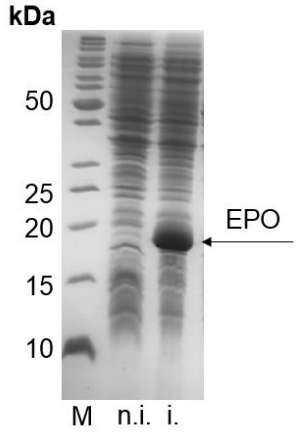


Figure 28: 15% SDS-PAGE showing the expression profile of EPO; M: marker; n. i.: not induced; i.: induced.

4.1.2. Incorporation of non-canonical amino acids by amber stop codon suppression

4.1.2.1. Expression of EPO incorporating Plk

The pET11a vector with the EPO gene (figures 61 and 62A-B in chapter 6.1.1) was used as a template for site directed mutagenesis. On positions K24, K38 and K83, the amino acids were individually replaced by an UAG stop codon. The successful mutagenesis was confirmed by sequencing the EPO gene. The vectors pET11a and pRSFduet pylS were subsequently co-transformed into *E. coli* BL21 (DE3). For the expression of EPO38Plk, the elongation factor EF-Tu from *E. coli* was needed in order

to obtain sufficient yields. Therefore, the gene *tufB* was cloned into the second MCS of pRSFduet pylS. The cell cultures were grown until an OD_{600} of one was reached, 2 mM of the unnatural amino acid Plk was added and gene expression was induced by adding IPTG and occurred overnight.

In figure 29, the expression profiles are shown for each individual incorporation site. The overexpression of EPO24Plk and EPO83Plk (figures 29A-B) is very high, deduced from the broad and intensive bands at 19 kDa, in contrast to the expression of EPO38Plk (figure 29C). Here, a smaller band is visible at 19 kDa. Moreover, in each expression profile the overexpressed PylRS can be recognised with a molecular weight of 48 kDa. Due to the amber suppression methodology, there is always truncated EPO, but it can be observed only for position 83 because the polypeptide chain of the EPO residues 1-82 has a molecular weight of about 9 kDa and the respective band appears at the bottom of the SDS gel (figure 29B). For position 38, incorporation of Plk resulted in such a low expression that no EPO band was visible under the same conditions as for positions 24 and 83 (data not shown). Higher yields were obtained in presence of elongation factor Tu, which has a molecular weight of 43 kDa.

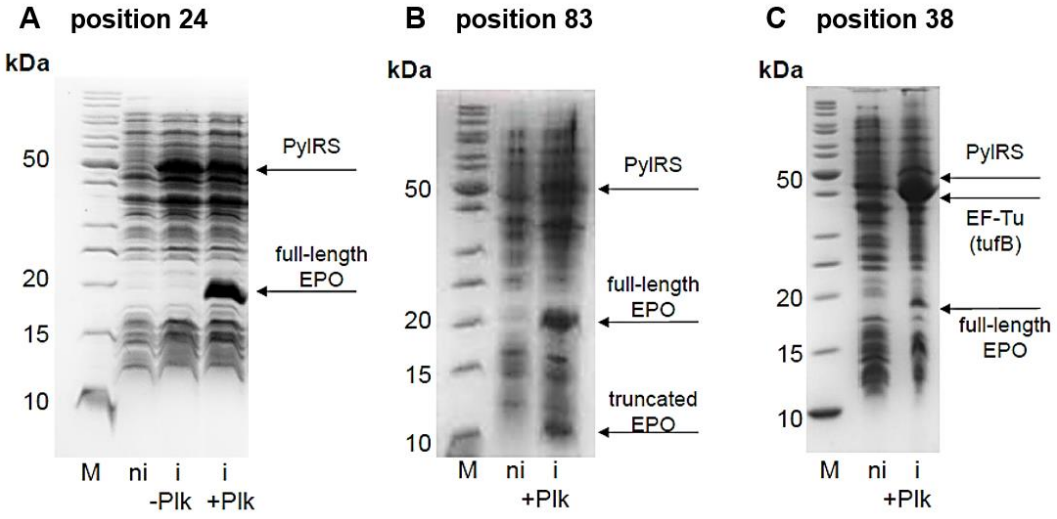


Figure 29: 15% SDS-PAGE showing the expression profile of EPO with incorporated Plk at the natural glycosylation sites

4.1.2.2. Expression of EPO incorporating Pln

For incorporation of Pln into EPO the same expression plasmids as for Plk were used (see 4.1.2.1). Pln was added to the culture in concentrations ranging from 0.5 to 1.6 mM (depending on purity). In figure 30, the expression profiles are depicted. Analogous to Plk incorporation, the expression showed highest yields at positions 24 and 83 (figure 30A-B). For position 38, EF-Tu is necessary to see EPO38Pln expression (figure 30C). The gene *tufB* is highly overexpressed overlaying the band of PylRS, which cannot be seen. However, it has to be there, because successful incorporation is only possible in presence of PylRS.

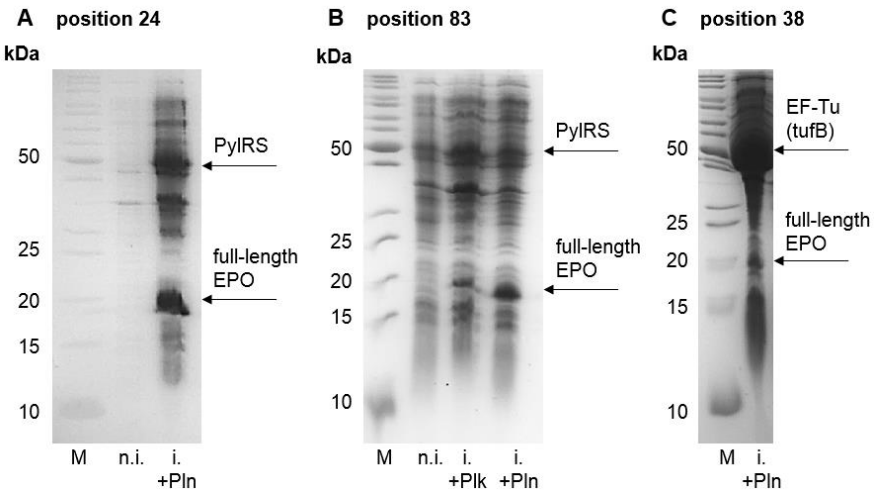


Figure 30: 15% SDS-PAGE showing the expression profile of EPO with incorporated Pln at the natural glycosylation sites

4.1.3. Incorporation of HPG by selective pressure incorporation

The sequence of HPG-EPO (figure 62C in chapter 6.1.1) was cloned into the pRSFduet vector, which was subsequently co-transformed together with pTARA into *E. coli* B834 (DE3). This plasmid has a T7 RNA polymerase under the control of AraPBAD promotor and therefore arabinose control. It is thought that overexpression of T7 RNA polymerase increases the yield of the target protein.

HPG will be incorporated by SPI in response to the ATG codon, which is present at the three natural glycosylation sites simultaneously. In figure 31, the expression profile is depicted. SPI experiments were always performed with several controls: a not induced control, an induced control without methionine or any analogue and an induced control in presence of methionine. The yield of expressed EPO with incorporated HPG is decreased compared to wtEPO. However, it was enough for purification and peptide mass fingerprint analysis.

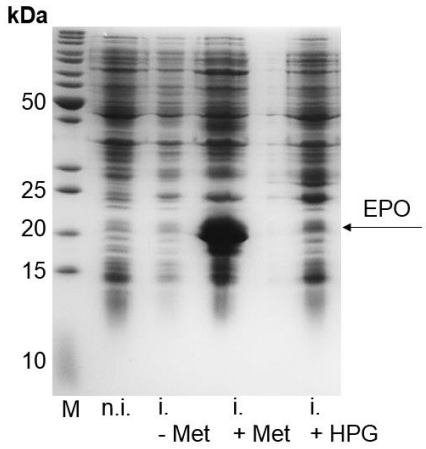


Figure 31: 15% SDS-PAGE showing the expression profile of EPO with incorporated HPG at positions 24, 38 and 83

4.1.4. Western Blot

Western blots were performed to prove that the observed 19-kDa band is indeed expressed recombinant EPO, particularly, if only a very faint band could be found on the gel. The detection is based on Ni-NTA conjugates, which bind to the N- or C-terminal polyhistidine-tag of EPO. The successful expression of recombinant EPO (see figures 28-31 in chapters 4.1.1 to 4.1.3) could be confirmed by Western blot analysis (data not shown). Moreover, the Western blot proved the complete production of EPO, as the polyhistidine-tag is at the C-terminus with the exception of HPG-EPO. There, it is at the N-terminus.

4.1.5. Peptide mass fingerprint

In order to prove the successful incorporation of unnatural amino acids into EPO, a peptide mass fingerprint was performed by the proteomic facility of the University of Konstanz. It could be shown that Plk was successfully incorporated at positions 24 and 38 individually and HPG at all three positions (24, 38 and 83) simultaneously. The sequence coverage ranged from 75 to 82%, whereby the unnatural amino acid positions were covered.

4.1.6. Conclusions of unnatural amino acid incorporation into EPO

The non-canonical amino acids Plk and Pln could be incorporated individually at each naturally occurring glycosylation site in high yields by amber stop codon suppression. However, the incorporation of unnatural amino acids Plk and Pln in response to the amber stop codon at position 38 is less efficient than at positions 24 and 83. This could be due to the position of incorporation. Because of unknown reasons, this position is not as suitable, as the other two positions in EPO to incorporate unnatural amino acid in response to the amber stop codon.

In figure 30B, a direct comparison of the expression profiles at position 83 is shown. In this case, incorporation of Pln results in higher protein yields compared to incorporation of Plk. It has to be said that overall expression of EPO incorporating unnatural amino acids such as Plk and Pln was not very reproducible. Indeed, the yields varied in each individual experiment. Nevertheless, a trend was observable: the protein yields were higher for incorporation of Pln than of Plk, although Chin et al. reported that the expression yields for Plk were higher than for Pln (see also 4.4) (Nguyen, Lusic et al. 2009).

Selective pressure incorporation was used to incorporate an unnatural amino acid at three positions simultaneously. The expression yields remain high, irrespective to the amount of incorporation positions. In contrast, using amber suppression, the yields decrease dramatically when two or three unnatural amino acids are incorporated. It was tried to incorporate Plk at two and three positions simultaneously, but the yields were too low for purification (data not shown).

The produced HPG-EPO (amino acid sequence, see figure 62C in chapter 6.1.1) had an N-terminal enterokinase tag. In this way, it was thought to cleave off the start HPG and the polyhistidine tag. However, tag cleavage by enterokinase was not successful (data not shown). A possible explanation could be that the cleavage site is buried within the protein and not accessible for the enterokinase.

4.2. Affinity-tag purification

EPO is expressed in inclusion bodies, in which the major percentage is denatured and only a minor part is folded or misfolded. To extract EPO from the other cell compartments, the inclusion bodies are solubilised under denaturing conditions and the solubilised EPO with a terminal polyhistidine-tag is collected by a first crude Ni-NTA affinity purification by a step-wise elution with increasing imidazole concentrations. A sample of all collected fractions was loaded onto a SDS-PAGE to determine the yield and purity of denatured EPO (figure 32). In the depicted gel example, EPO is eluted by an imidazole concentration of 100 mM. It is still not pure, as mainly upper bands appear in the same fractions as EPO. Because of these impurities, it was not possible to determine the exact yield of EPO expression at this state. Only an approximate value could be obtained by measuring the absorbance at 280 nm, as the extinction coefficients of the contaminants were unknown. However, at least another purification step is followed, in which EPO will completely be purified. Therefore, it was possible to continue with these partially pure EPO fractions.

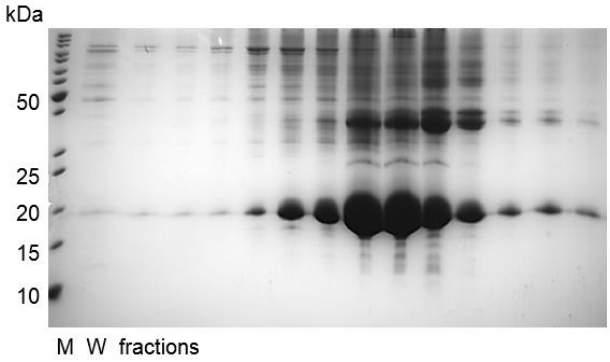


Figure 32: 15% SDS-PAGE of fractions from affinity-tag purification of EPO under denaturing conditions with Ni-NTA beads; W: washing fraction

4.3. Refolding

After Ni-NTA purification, EPO is still denatured in a buffer containing 3 M of guanidine hydrochloride. The step of refolding, which means to bring back EPO to its native state with the two structural disulphide bonds, had to be optimised.

In literature, several methods for refolding EPO are delineated:

- The groups of Bradburne, Kochendoerfer and Kajihara used an old dialysis protocol for refolding their synthesised and glycosylated or polymer-attached EPO derivatives from Saxena and Wetlaufer: (Saxena and Wetlaufer 1970) (Kochendoerfer, Chen et al. 2003, Chen, Cressman et al. 2005, Hirano, Macmillan et al. 2009). Folding was proceeded by a three-step dilution of guanidine hydrochloride starting at a concentration of 3 M in the presence of a glutathione or a cysteine-cystine redox system.
- Su et al. refolded *E. coli* derived non-glycosylated EPO via dilution and tested the following three additives: 0.5 M arginine, 1 M guanidine hydrochloride and 2 M urea in comparison to no additives (Wang, Liu et al. 2010). They reported that all three additives were effective in

suppressing the formation of insoluble aggregates during refolding with a mass recovery over 90%, but only in the presence of 0.5 M arginine, the yield of correctly refolded EPO was as high as 73% after SEC and RP-HPLC. Their final optimised refolding conditions adopted were 20 mM Tris-HCl (pH 8.5) containing 0.5 M arginine at 4 °C with a protein concentration of 100 µg/mL.

- Hamann added a glutathione redox system to the refolding dilution protocol (Hamann 2010). The best concentrations to obtain highest yields were 1 mM reduced (GSH) and 0.3 mM oxidised glutathione (GSSG).
- Kent et al. used the dialysis protocol for a not glycosylated EPO derivative (Liu, Pentelute et al. 2012). First, they dissolved EPO in a 6 M guanidine hydrochloride solution. The next dialysis buffer contained 3 M of guanidine hydrochloride, 4 mM of cysteine and 0.5 mM of cystine (pH 8.5). The last dialysis buffer contained 10 mM of Tris-HCl (pH 7.0).
- Danishefsky et al. orientated themselves on a protocol of Strickland et al., where EPO is solubilised with *N*-lauroylsarcosine and is allowed to oxidise with the addition of $CuSO_4$ (Narhi, Arakawa et al. 1991, Wang, Dong et al. 2012).

First, it was tried to refold EPO through dilution. The basic buffer contained 20 mM of Tris-HCl (pH 8.5), 0.5 M of arginine. The effect of both redox systems (GSH/GSSG and cysteine/cystine) on the yield of soluble EPO was investigated. The tested glutathione redox system consisted out of 1 mM GSH and 0.3 mM GSSG and the Cys-redox system contained of 1 mM cysteine and 50 µM cystine. As protein concentrations were measured by UV-VIS and the present high concentrations of arginine falsify the measurement positively, the absolute protein yield could not be determined. Nevertheless, it was possible to compare the resulting EPO yields to each other because of equal arginine concentrations. The yield of the best refolding system, the one with glutathione, was set as arbitrary 100% and the other systems' relative yields to this system are shown in table 40. A sodium phosphate buffer (20 mM, pH 8.5) could replace the Tris-buffer system without any changes in the yield.

Redox system	yield
none	30%
cysteine/cystine	80%
GSH/GSSG	100%

Table 40: Resulting yields in refolding by dilution experiments with several redox systems

In parallel, several refolding conditions by dialysis with a two-step dilution of guanidine hydrochloride were assessed. Three different versions of the first dialysis buffer were tested (see conditions in table 41). The second dialysis buffer was sodium phosphate buffer (20 mM, pH 7.5) in all cases.

	first dialysis buffer I	first dialysis buffer II	first dialysis buffer III
buffer	20 mM sodium phosphate, pH 8.0	20 mM sodium phosphate, pH 8.0	20 mM sodium phosphate, pH 8.0
additives	0.5 M arginine	1 M guanidine hydrochloride	1 M guanidine hydrochloride
redox system	GSH/GSSG	GSH/GSSG	cysteine/cystine

Table 41: Several conditions tested as first dialysis buffer in refolding by dialysis

The relative yields compared to the yield of refolding by dilution in presence of the GSH/GSSG system (that was set to 100%) are depicted in table 42.

Additives	yield
0.5 M arginine, GSH/GSSG	22%
1 M guanidine hydrochloride, GSH/GSSG	15%
1 M guanidine hydrochloride, cysteine/cystine	13%

Table 42: Resulting yields in refolding by dialysis experiments with several additives in the first dialysis buffer

Moreover, further additives for refolding by dilution were tested: Triton X-100, *N*-lauroylsarcosine and PEG 3550. All supplements help to recover a higher yield of refolded EPO, but they stick to the protein and it was impossible to remove them again, which is necessary for further applications. Therefore, they were not considered as reliable additives.

To sum it up, the glutathione redox system appeared to be more preferable than the cysteine/cystine redox system and refolding by dilution has been shown to be more beneficial than refolding by dialysis. The second finding is very reasonable, because EPO concentrations are much higher in the dialysis procedure compared to the dilution procedure. The compactness of EPO in dialysis can easily lead to aggregation and precipitation of the protein.

The resulting optimal refolding conditions are described in the method section (see 3.4.4) with the optimised refolding buffer listed in the material section (see table 17 in chapter 2.5).

4.4. Purification of refolded EPO

After refolding, EPO is present in refolding buffer containing high concentrations of arginine. Moreover, there are misfolded EPO molecules, which need to be removed from the correctly folded part. Therefore, another purification step is essential. Various purification methods and columns were assessed for optimal recovery of correctly folded EPO.

First, ion exchange chromatography (IEC) was tested. As EPO has a pI 9 (isoelectric point), a cation exchanger would be first choice. SP (sulfopropyl) sepharose is a strong and CM (carboxymethyl) sepharose is a weak cation exchanger. However, it is necessary to change the buffer system before IEC because the high concentration of arginine is interfering and the pH value has to be exactly adjusted to pH 7.5. This was done by dialysis and by disposable PD-10 desalting columns. In both cases, 20 mM Tris- or sodium phosphate buffer was used. As mentioned above, EPO is prone to

precipitation within dialysis. PD-10 columns have not shown better results, as most of EPO has been stuck inside the column. Yields in IEC itself were very low, about 1% of the protein input before refolding was recovered after IEC. Consequently, IEC was not considered for EPO purification.

Last, SEC with various columns was tested. Superdex was by far the best gel filtration medium compared to Sephadex or Sephacryl. For SEC, it is not necessary to exchange buffer before purification, which is an important advantage over IEC because buffer exchange was always related to high loss of protein. After refolding, the protein solution was concentrated with Vivaspin® centrifugal concentrators, which was a time-consuming process, as several hundred millilitres had to be concentrated to one or few millilitres. The smaller the load input, the better and clearer the separation after SEC would be. In figure 33A, a chromatogram is depicted as an example. EPO was eluted in a clear peak after 12 – 16 mL elution volume. The smaller peak before the EPO peak consisted out of impure EPO. When the peak after the EPO peak was loaded onto a SDS-PAGE, no band could be seen. It is proposed, that the peak originate from small fragments of EPO. Another notice that could be recognised from the chromatogram is the relative UV absorbance signal of the different wavelengths. For better comparison, the UV spectrum of EPO is depicted in figure 33B. The peak at 200 nm originates from the peptide bond and the peak at 280 nm originates from aromatic ring systems of mainly tryptophan and tyrosine, and to a smaller extent of phenylalanine. The absorbance at 280 nm, together with EPO's extinction coefficient of $\epsilon = 1.24$ L/(mol*cm), is used to determine the protein concentration (see 3.4.5.1).

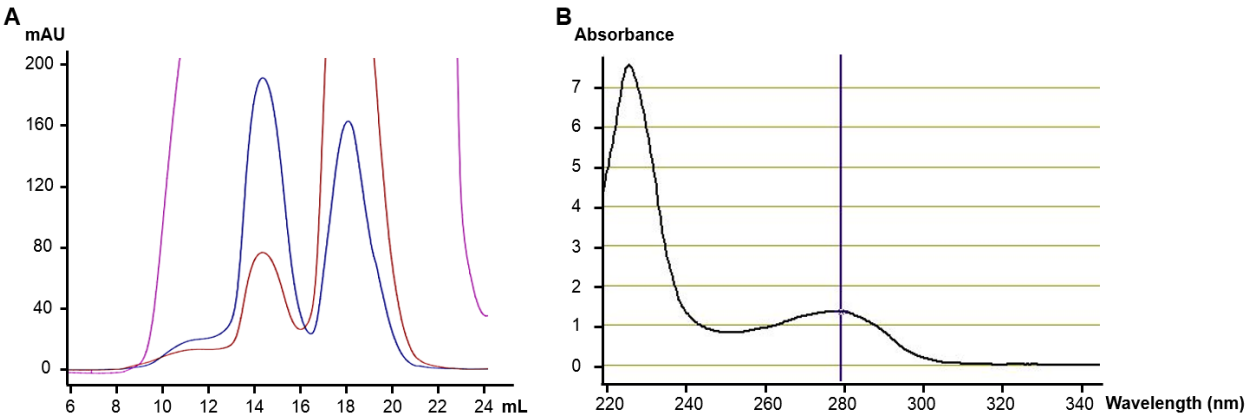


Figure 33: (A) Example of a chromatogram of EPO purification via SEC, UV absorbance at 280 nm (blue), 254 nm (red) and 214 nm (magenta); (B) Example of an UV spectrum of EPO

Fractions of 0.25 to 0.5 mL were collected and samples of these were loaded onto a SDS-PAGE (figure 34). The first small peak corresponds indeed to EPO with some impurities, the large peak is pure EPO and the third peak on the chromatogram could not be seen on SDS-PAGE. Incorporation of unnatural amino acids, e.g. Plk or Pln does not alter the elution behaviour.

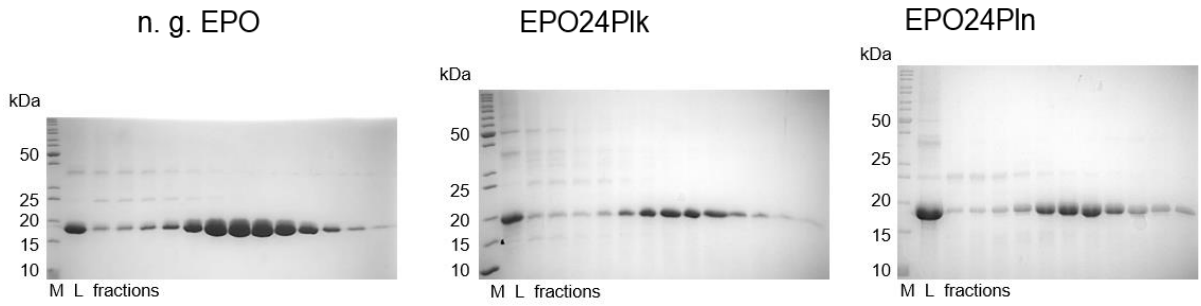


Figure 34: SDS-PAGE of the load and fractions containing EPO after SEC; M: marker, L: load

The clean EPO fractions were pooled and concentrated. The protein concentration was determined by measurement of the absorbance at 280 nm for the correct input in further experiments. In this way, it was possible to determine the yield of purified and refolded EPO per litre of expression at this stage (table 43). For unmodified EPO, the yield was 11.8 mg per litre of bacterial culture. This value was set as a reference for the EPO variants with an incorporated amino acid. The yield for incorporating Pln into EPO and the subsequent isolation procedure is 9% for positions 24 and 83, whereas it is decreased to 3% for position 38.

For incorporation of Plk, the yields further decreases to 3% and 2% for positions 24 and 83. Only 0.5% protein yield was found incorporating Plk at position 38. These findings confirm the observation that yields are better for incorporating Pln, as it is described in chapter 4.1.6. Moreover, even the elongation factor Tu 2 (tufB) is added to support the translation, this location is disfavoured for incorporation of Plk or Pln at position 38. The yields are 15% to 40%, when compared to the respective two other positions of incorporation.

	expression volume in L	**yield in mg**	**yield in mg/L**	**yield in % to EPO**
EPO	0.20	2.35	11.77	100
EPO24Plk	3.00	1.15	0.38	3.25
EPO38Plk	6.00	0.35	0.06	0.50
EPO83Plk	3.00	0.71	0.24	2.01
EPO24Pln	1.00	1.10	1.10	9.35
EPO38Pln	0.50	0.20	0.40	3.40
EPO83Pln	0.50	0.50	1.00	8.50

Table 43: Yields of purified and refolded EPO per litre of bacterial culture

A native PAGE could quickly show a first evidence, if EPO is correctly folded (figure 35). Native EPO appears at a lower molecular weight than reduced EPO because the two disulphide bridges are intact, thus the protein is more compact, in comparison to denatured EPO, which is boiled and treated with 2-mercaptoethanol.

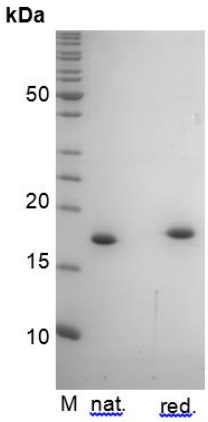


Figure 35: Native PAGE showing refolded (nat.) and reduced (red.) EPO

4.5. Click chemistry

4.5.1. EPO-Plk coupled to azido-glycans

The click reaction between azido-glycans and EPO containing Plk with an alkyne group has been optimised. In figure 36, the results of the first optimisation experiments with low EPO concentrations are shown on 15% SDS-PAGE. As some Cu(II)-species would always be present, because it is not possible to exclude aerial oxygen to 100%, as less as possible Cu(I)-complex was used. Figure 36A shows that 4 mM Cu(I)-complex increased the yield of click reaction to about 50% from about 30% with 1 mM Cu(I)-complex. In figure 36B, it was realised that the optimal ratio of Cu(I)-complex to the ligand THPTA was 1:2 and not 1:5, as reported in literature (Hong, Presolski et al. 2009). Lastly, the EPO-alkyne : azido-glycan ratio was investigated in figure 36C. It was found that the more azido-glycan is added, the more EPO is linked to an azido-glycan by click reaction. For economically reasons, an EPO : glycan ratio of 1:50 was used in further experiments, which yields in more than 90% product formation.

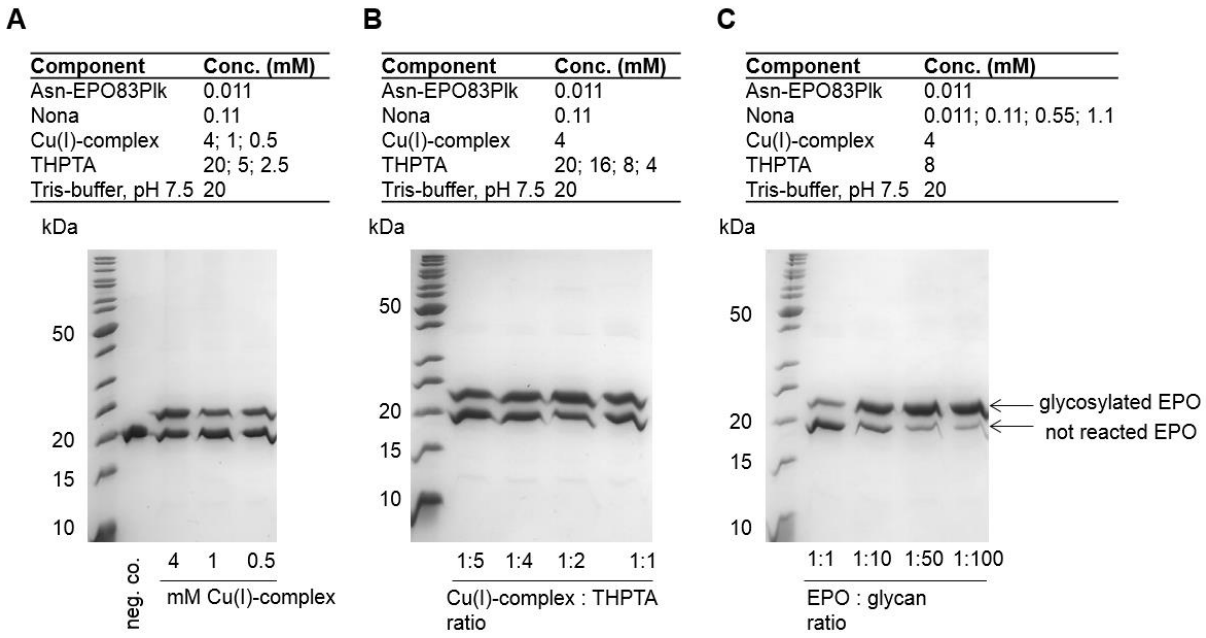


Figure 36: SDS-PAGE of the optimisation of the click-reaction (4 h reaction time): (A) Optimisation of the concentration of Cu(I)-complex, neg. co.: EPO without incorporated Plk; (B) Optimisation of Cu(I)-complex : THPTA ratio; (C) Optimisation of EPO-alkyne : azido-glycan ratio

In a next step, the click reaction was repeated with high EPO concentrations (over 30 µM) to have enough glycosylated EPO for purification and further experiments (figure 37). It was found that the EPO:glycan ratio could be decreased again to 1:10, as the click product formation is apparently the same for 1:10, 1:20 and 1:30. The resulting optimal conditions for click reactions of EPO-Plk with azido-glycans are summarised in table 44.

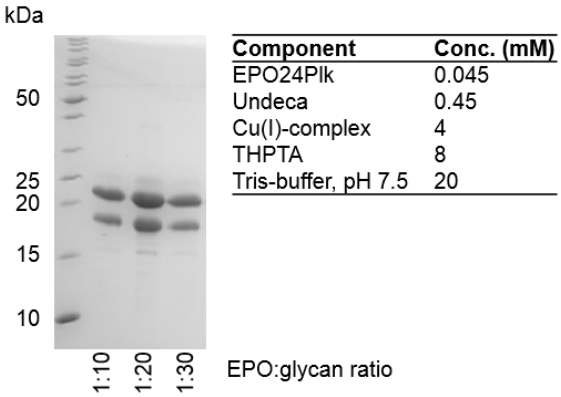


Figure 37: SDS-PAGE of the optimisation of the click reaction with high EPO concentrations

	low EPO concentration	high EPO concentration
EPO-Plk	8 µM	40 µM
azido-glycan	400 µM	400 µM
Cu(I)-complex	4 mM	4 mM
THPTA	8 mM	8 mM

Table 44: Optimal click-reaction conditions

Further conditions were tested to optimise the reaction yields (figure 38). However, all of them had either a negative or no effect on the yield of the click reaction.

- First, it was tested, if the reaction would be more efficient in sodium phosphate buffer to follow some advices of Finn et al. (Hong, Presolski et al. 2009). Though, figure 38A shows clearly that best click reaction results are obtained using Tris-buffer.
- BSA was added to the click reaction because it was thought, it might protect the protein from precipitation during click reaction (figure 38B). To investigate this, the samples were centrifuged (13000 rpm, 4 °C, 5 min) after click reaction. Then, the supernatant was transferred into a new reaction tube and the pellet was dissolved in Tris-buffer. BSA has only a minor effect on the precipitation of EPO during click reaction. Another finding of this gel is that unreacted EPO precipitate to a higher extent than glycosylated EPO, which is a first hint that one oligosaccharide of 2.2 kDa with two sialic acid residues is able to protect EPO from aggregation.
- Marx et al. could improve click product formation with yields up to 40% by adding at least 0.25 mM of SDS to the reaction (Schneider, Schneider et al. 2013). However, in click reactions of EPO to azido-glycans, where the yield is already over 90%, adding SDS to the reaction has no further effect (figure 38C).
- Instead of using Cu(I) salts, the catalyst can be prepared *in situ*, for example by reduction of $CuSO_4$ (Rostovtsev, Green et al. 2002). Sodium ascorbate and TCEP were shown to be a suitable reductant, when proteins are present (Wang, Chan et al. 2003). These traditional conditions for protein click reactions were assessed in figure 38D. In all cases, THPTA was used as a ligand. EPO does not precipitate under these conditions, but the click product yield decreases dramatically using $CuSO_4$ and sodium ascorbate. Even no click product at all was obtained using TCEP as a reductant, even though no precipitation was observed.
- Next, sodium ascorbate was added to the reaction with Cu(I)-complex and THPTA. The idea was that sodium ascorbate will reduce present Cu(II)-species originating from oxidised Cu(I) by leaking areal oxygen (figure 38E). However, the expected increase in yield was not observed. Instead, almost all the protein precipitated and product yields were reduced enormously to about 15%.
- Click reaction was conducted in Tris-buffer with various pH values ranging from 6.5 to 8.3. No effect of the pH value on the click reaction could be shown (figure 38F).
- Lastly, THPTA was replaced by a synthesised GlcNAc-ligand from the group of Prof. Unverzagt, University of Bayreuth (figures 23 and 38G). Unfortunately, no click product was formed adding this ligand.

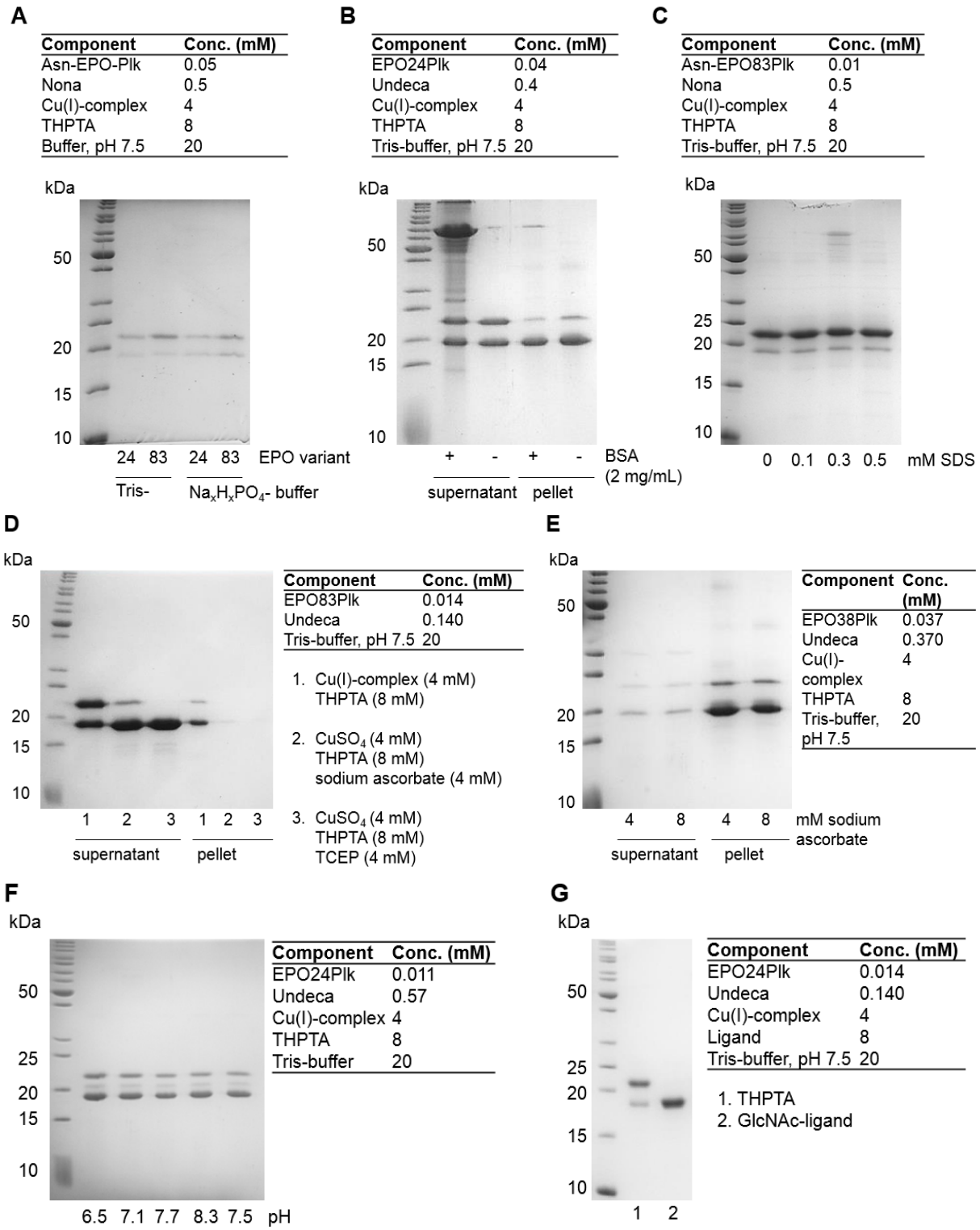
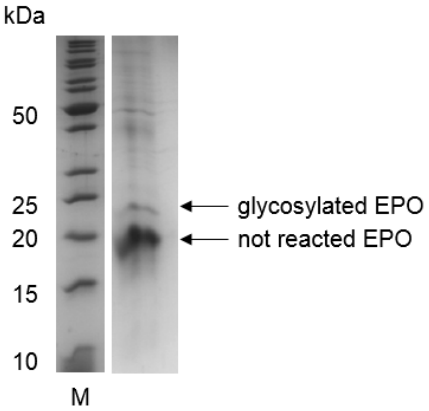


Figure 38: Further studies of click reactions on 15% SDS-PAGE: (A) Comparison of click reactions in Tris- and $Na_xH_xPO_4$-buffer, (B-C) Investigating the effect of BSA (B) and various concentration of SDS (C) in click reactions; (D) Comparison of different Cu(I)-sources and reductants; (E) Testing the effect of sodium ascorbate in click reactions; (F) Comparison of click reactions in buffers with different pH-values; (G) Comparison of THPTA as a ligand to a GlcNAc-ligand from the group of Prof. Unverzagt, University of Bayreuth

Another idea was to couple EPO to oligosaccharides under denaturing conditions (figure 39). In this way, the procedure from expressing EPO to the final glycosylated EPO would be one purification step shorter (see figure 40 for an overview). Therefore, the click reaction was performed using EPO purified by Ni-NTA in purification buffer containing 3 M GdmCl. However, the click reaction is very inefficient under these conditions. Consequently, this strategy was not pursued any longer after this experiment.



Component	Conc. (mM)
Asn-EPO24Plk	0.015
Nona	0.750
Cu(I)-complex	4
THPTA	8
GdmCl	3000

Figure 39: SDS-PAGE of the click reaction of EPO24Plk to glycan Nona in 3 M GdmCl after 3 h

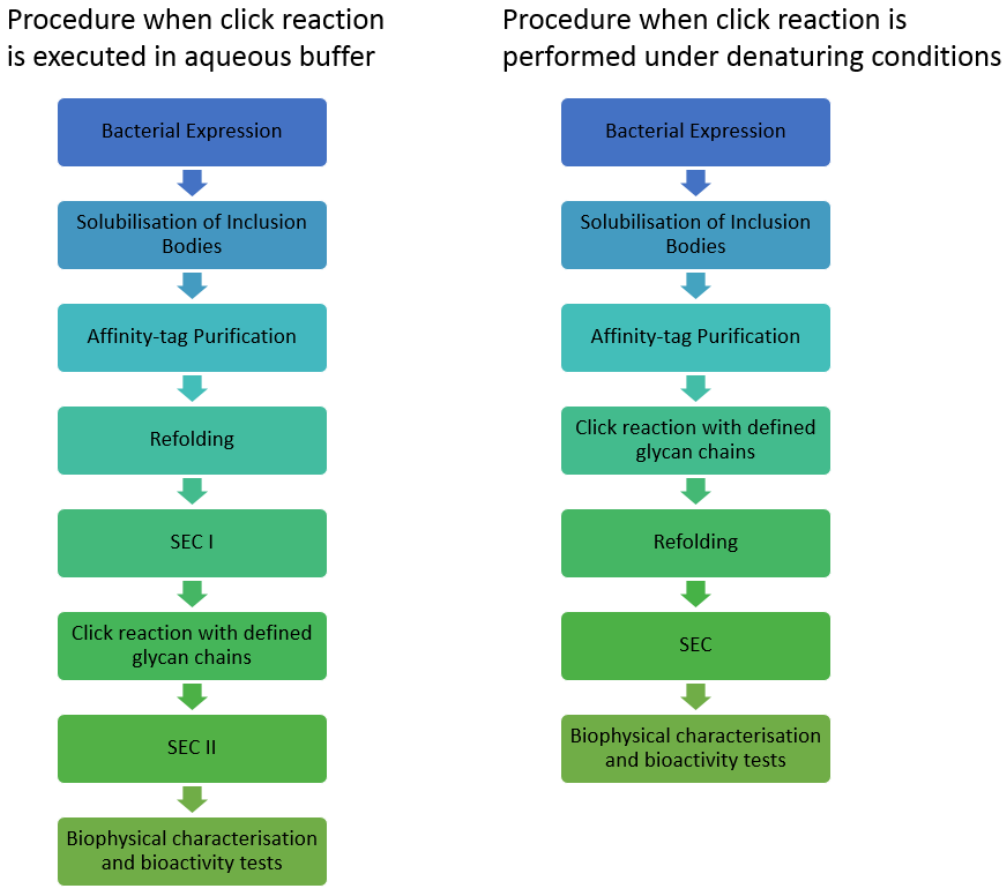
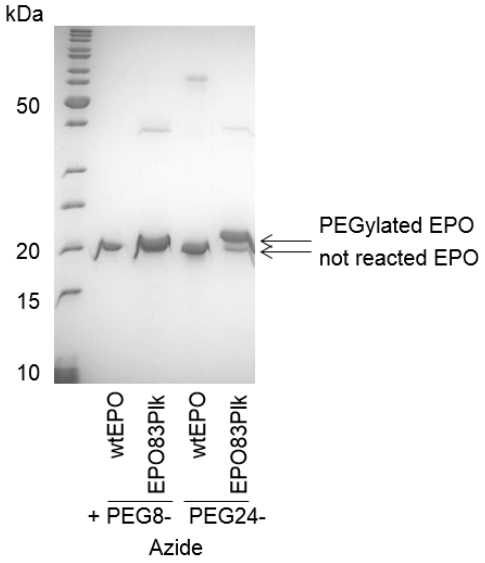


Figure 40: Overview of two possible procedures from bacterial expression of EPO to homogeneously glycosylated EPO

4.5.2. EPO-Plk coupled to azido-PEGs

4.5.2.1. Click reaction to short PEG chains

Instead of oligosaccharides, EPO could be also successfully coupled to commercially available PEG-azides (figure 41). The PEG8 is too small to see a distinct band for the EPO83-PEG8. Nevertheless, as the band appears to be thicker than the EPO band in the negative control, it is concluded that the click reaction was as efficient as for PEG24, where a turnover of over 80% can be observed.

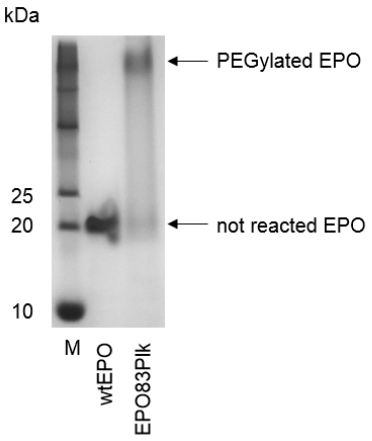


Component	Conc. (mM)
Asn-EPO(83Plk)	0.011
PEG8/24-Azide	1.1
Cu(I)-complex	4
THPTA	8
Tris-buffer, pH 7.5	20

Figure 41: Click reaction of EPO83Plk to short PEG chains

4.5.2.2. Click reaction to PEG-20kDa

As it was not possible to see the band of EPO coupled to the PEG-20KDa on SDS-PAGE (data not shown), a Sarcosyl-PAGE was performed (see chapter 3.4.6.2 for details). On this Sarcosyl-PAGE in figure 42, it is shown that almost all of the EPO is coupled to the 20-kDa-PEG, as the major band appears at the top of the gel. The PEGylated EPO should have a molecular weight of about 40 kDa. PEGylated proteins can always be found at higher molecular weights than estimated because PEG interacts with SDS (Reichel 2012). This effect was also observed previously (Wang, Liu et al. 2010).

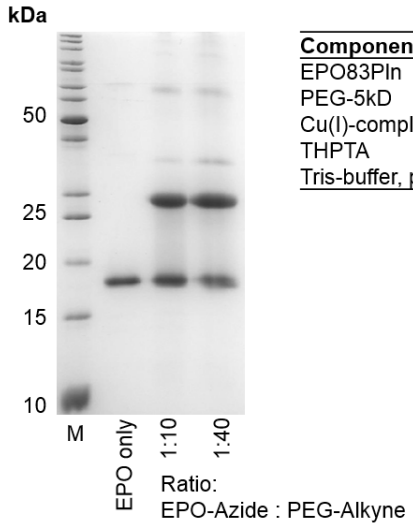


Component	Conc. (mM)
Asn-EPO83Plk	0.011
PEG-20kD	0.112
Cu(I)-complex	4
THPTA	8
Tris-buffer, pH 7.5	20

Figure 42: Sarcosyl-PAGE of click reaction with EPO83Plk to PEG-20kDa

4.5.3. EPO-Pln coupled to alkyne-PEG

Pln has an azide group, therefore a click ligand with an alkyne group was necessary. A 5-kDa-PEG-alkyne was purchased from Creative PEGWorks. The click reaction conditions were set as before, however, now with an excess of the alkyne. Click product formation could be observed with a turnover rate of about 60-70% (figure 43).



Component	Conc. (mM)
EPO83Pln	0.03
PEG-5kD	0.30; 1.22
Cu(I)-complex	4
THPTA	8
Tris-buffer, pH 7.5	20

Figure 43: Click reaction of EPO83Pln to alkyne-PEG-5kDa

4.6. Purification of glycosylated and PEGylated EPO

After click reaction to PEG or glycan chains, the coupled EPO molecules were separated from the not reacted EPO and other components of the click reaction by SEC purification. In figure 44, the chromatograms and the respective 15% SDS-PAGEs of some examples for SEC purification after click reaction are depicted. Normally, two peaks could be observed: one peak for the clicked EPO product and another peak for the uncoupled EPO. For smaller decorations, the two peaks were not separated completely and a few fractions containing both species had to be removed.

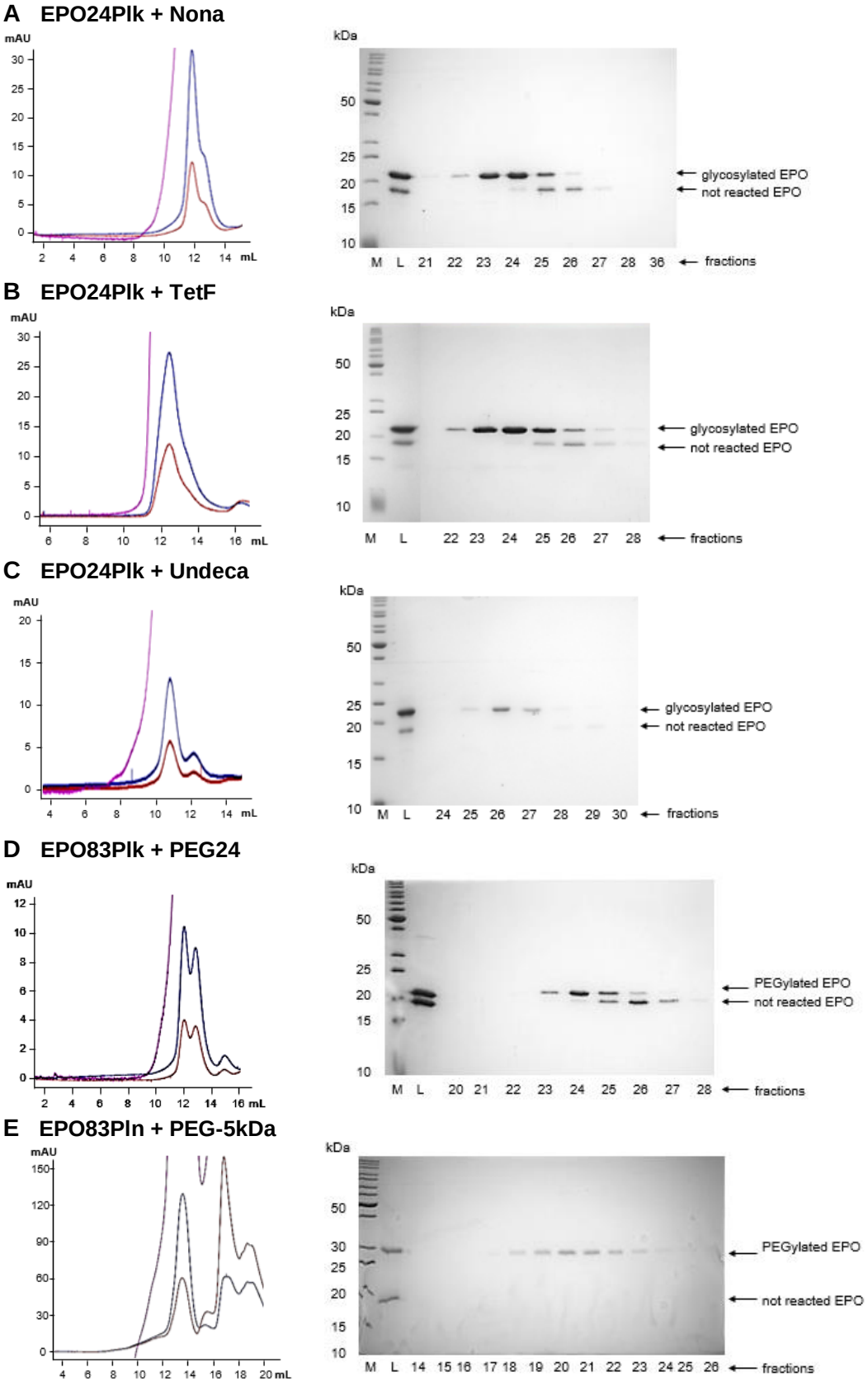


Figure 44: Chromatograms and 15% SDS-PAGEs of SEC purifications of glycosylated and PEGylated EPO after click reaction, M: Marker, L: Load

For bigger click ligands, for example PEG-5kDa and Undeca, the two peaks eluted totally separated from each other. Here, all fractions with clicked EPO could be collected, pooled and concentrated. The concentration was determined measuring the absorbance at 280 nm and the yield of click reaction and the subsequent purification was calculated. The average protein yield was about 10.5% of the amount of EPO set in the click reaction. The coupled EPO variants were used for biophysically characterisation and biological activity assays.

4.7. Mass analysis

The EPO variants expressed in *E. coli* with or without the incorporated unnatural amino acid Plk were refolded and coupled to various glycans. In order to prove that all steps provide the whole EPO protein with all amino acids and glycan residues, the EPO samples were sent to the Functional Genomics Center of ETH Zurich for ESI-MS analysis. As it is shown in figure 45, the correct mass for EPO, EPO with incorporated Plk and EPO coupled to various glycans could be confirmed. No deleterious effects of copper on the glycan backbone could be detected.

	AsnEPO	AsnEPO24Plk	AsnEPO24-Nona	EPO24-TetF	AsnEPO24-Undeca
calculated	19362.3	19410.1	21075.7	21240.8	21657.9
found	19362.5	19410.5	21076.5	21240.0	21659.9

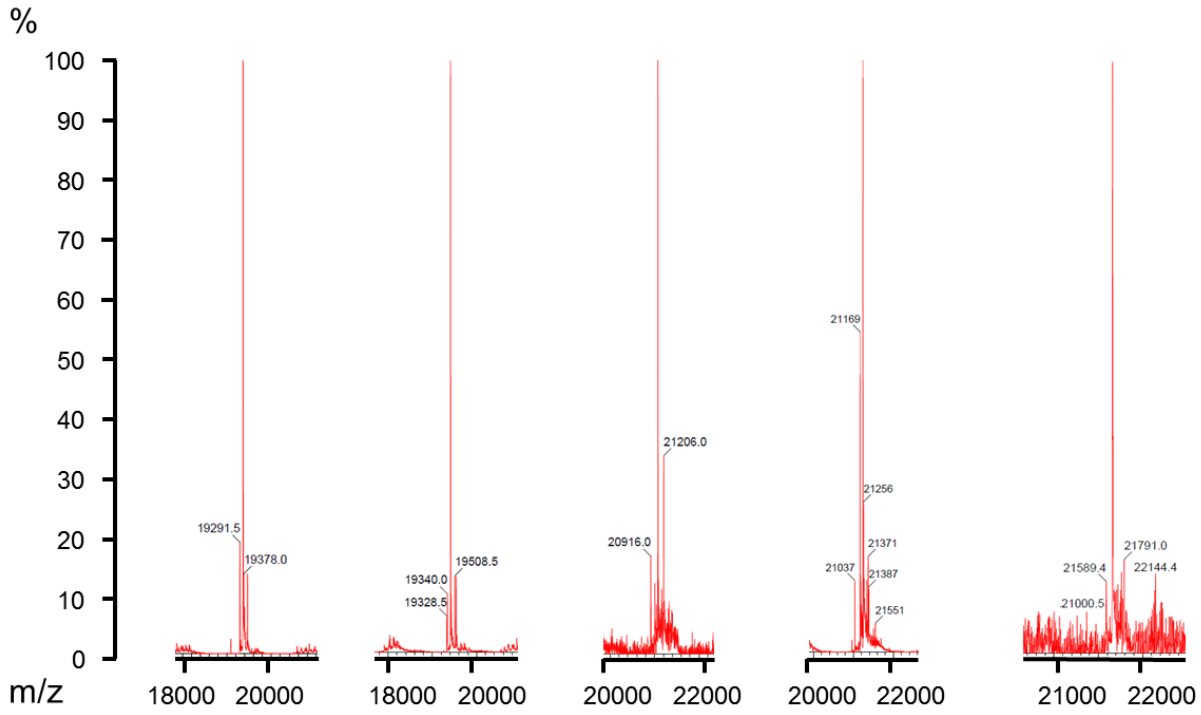


Figure 45: Calculated and found mass of several EPO variants and the respective deconvulted spectra of the ESI-MS analysis

4.8. Biophysical characterisation

4.8.1. CD spectrometry

After refolding and purification of EPO, it was important to assess the secondary structure by CD spectrometry. In this way, it is possible to detect α-helices and β-strand structures, which are a hint for correctly folded EPO. The characteristic minimum at 208 nm and the saddle point at 222 nm for proteins consisting primarily of α-helices, as it is the case for EPO, could be observed for all EPO variants (figure 46) (Kelly and Price 2000, Kelly, Jess et al. 2005). For better comparison, all glycosylated and PEGylated samples were overlaid by the respective EPO variant with the incorporated unnatural amino acid and adjusted to the same concentration. It could be shown that the click reaction had no effect on the secondary structure of folded EPO.

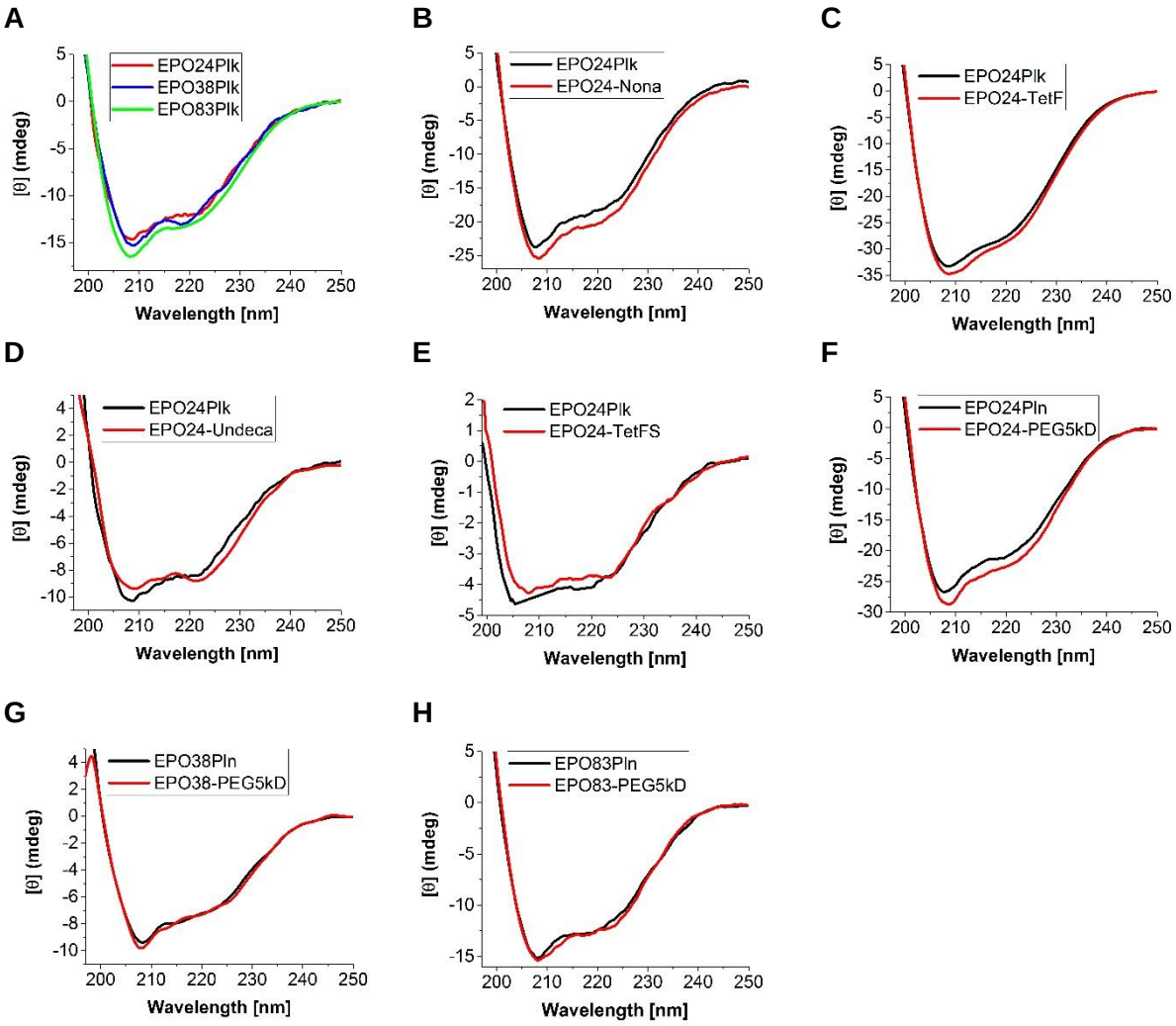


Figure 46: CD spectra of several glycosylated and PEGylated EPO variants in comparison to the respective EPO-Plk variant; (A) conc.: 5 µM (B) conc.: 6 µM (C) conc.: 8 µM (D) conc.: 4 µM (E) conc.: 3 µM (F) conc.: 8 µM (G) conc.: 3 µM (H) conc.: 5 µM; Concentrations were determined by micro BCA protein assay and by UV measurement in parallel.

4.8.1.1. EPO stability upon repeated freezing-thawing cycles

Circular dichroism was a perfect method to assess the effect of freezing-thawing cycles and of prolonged incubation at 37 °C. The relative concentration could be estimated as an attempt based on the assumption that the secondary structure is not altered. The concentration of uncoupled EPO (e.g. with Plk) was adjusted to the concentration of EPO with decorations by UV measurement, in order to have similar starting CD signal intensities (figures 47-48). Subsequently, the EPO samples were frozen in liquid nitrogen and thawn again to RT three times. The precipitated EPO was removed by centrifugation (5 min at 13000 rpm) and the supernatant was transferred to a new reaction tube.

Both carbohydrates, Nona and TetF protect EPO efficiently from aggregation due to repeated freezing-thawing cycles, which can be observed in figure 47. In table 45, the CD spectrum minima at 208 nm and 222 nm are listed for glycosylated and not glycosylated EPO variants before and after three freezing-thawing cycles. Moreover, the percentage of soluble EPO after treatment was calculated. Glycosylated EPO (with Nona and TetF) stays to a percentage of about 67-69% in solution after freezing-thawing cycles, whereas not glycosylated EPO only to a percentage of about 44-47%. EPO with the incorporated unnatural amino acid is even worse. Here, only 23-24% stay in solution after treatment.

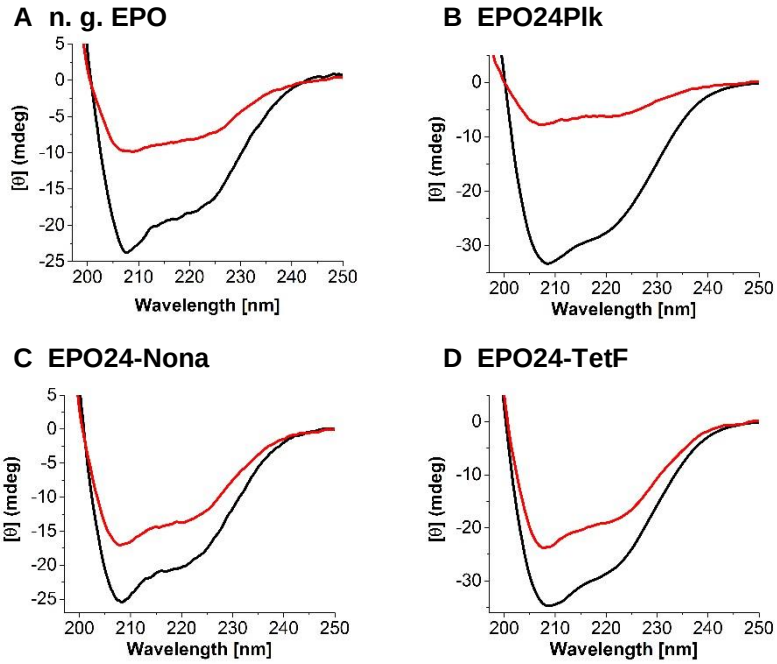


Figure 47: CD spectra of several (glycosylated) EPO variants before (black line) and after (red line) three freezing-thawing cycles

The same effect could be noticed for EPO coupled to a 5-kDa-PEG chain (figure 48 and table 46). Here, PEGylation at position 24 protects EPO best against precipitation upon repeated freezing-thawing cycles. 91-93% of EPO remain soluble. The effect of a 5-kDa-PEG is also significant at the other two positions 38 and 83. PEGylated EPO could be found in solution after freezing-thawing cycles to 66-77% for position 38 and to 76-77% for position 83.

	208 nm			222 nm		
	before	after	%	before	after	%
n. g. EPO	-19.9	-9.3	47	-15.2	-6.6	44
EPO24-Plk	-33.2	-7.8	23	-26.1	-6.1	24
EPO24-Nona	-25.4	-17.1	67	-19.4	-1.3	68
EPO24-TetF	-34.5	-23.8	69	-27.4	-18.4	67

Table 45: CD signal intensity in mdeg at 208 nm and 222 nm and the percentage of recovered soluble glycosylated EPO after three freezing-thawing cycles

In contrast, EPO with the unnatural amino acid Pln incorporated is prone to precipitation upon three freezing-thawing cycles. The unnatural amino acid at position 24 seems to be the most acceptable position, as the loss of the soluble EPO fraction is only about 41-42%. For position 38, about 45-51% of EPO-Pln precipitates upon freezing-thawing cycles. Position 83 seems to be the most unfavourable because only 22-24% of the protein could be recovered after treatment.

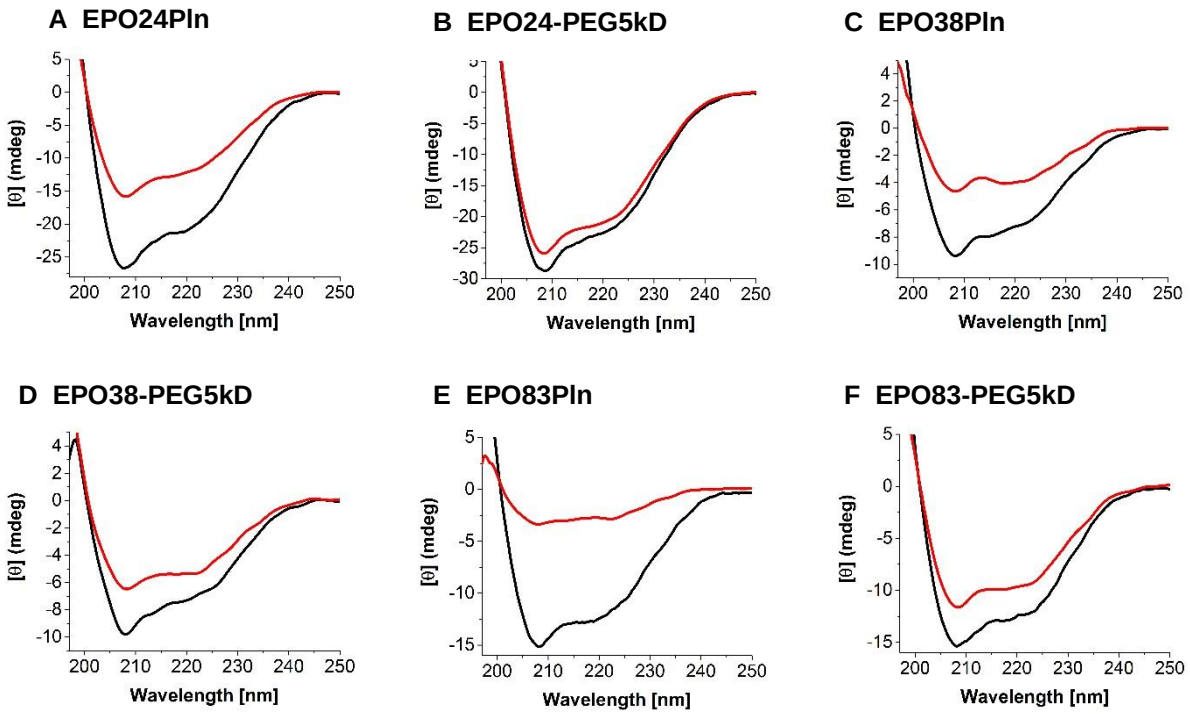


Figure 48: CD spectra of PEGylated-(5kD) and Pln-EPO variants before (black line) and after (red line) three freezing-thawing cycles

To sum it up, only one coupled glycan chain of a mean molecular weight of 1.8 kDa is already able to protect EPO from precipitation upon three repeated freezing-thawing cycles, as well as a 5-kDa-PEG chain.

	208 nm			222 nm		
	before	after	%	before	after	%
EPO24Pln	-26.7	-15.8	59	-20.0	-11.7	58
EPO24-PEG5kD	-28.5	-25.8	91	-21.8	-20.2	93
EPO38Pln	-9.4	-4.6	49	-7.0	-3.9	55
EPO38-PEG5kD	-9.8	-6.5	66	-6.9	-5.3	77
EPO83Pln	-15.1	-3.4	22	-11.8	-2.9	24
EPO83-PEG5kD	-15.4	-11.6	76	-12.3	-9.5	77

Table 46: CD signal intensity in mdeg at 208 nm and 222 nm and the percentage of recovered soluble PEGylated EPO after three freezing-thawing cycles.

4.8.1.2. Stability of PEGylated EPO against unspecific aggregation

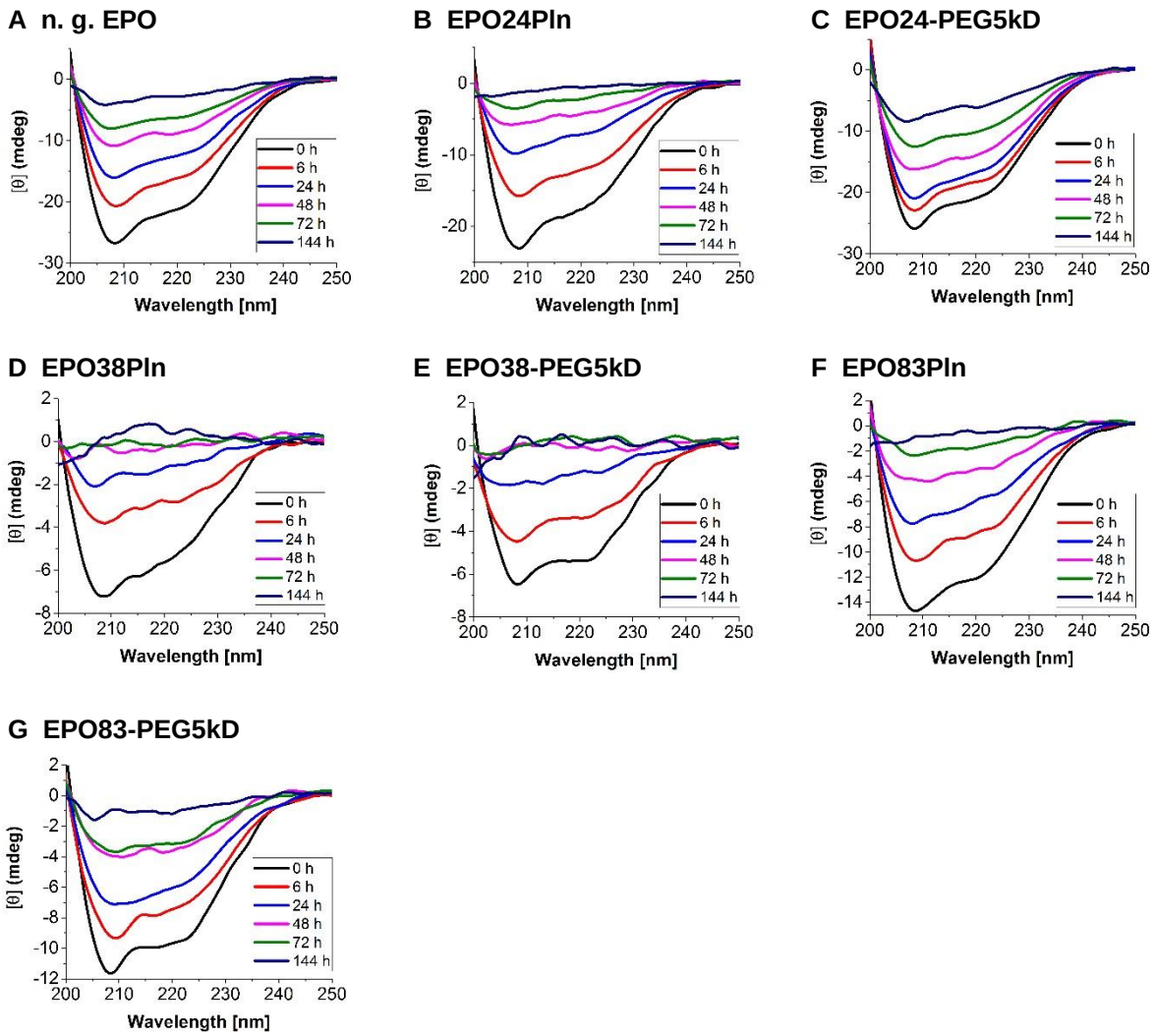


Figure 49: CD spectra after various incubation times at 37 °C of PEGylated-(5kD) and Pln-EPO variants

The stability of PEGylated EPO against unspecific aggregation was also assessed by circular dichroism. Therefore, the EPO-Pln concentrations were adjusted to the concentrations of the respective EPO-PEG (5kDa) variants. UV measurements verified the concentration estimation by CD. Then, all samples were incubated at 37 °C. At various time points, the samples were centrifuged and the supernatant was transferred into a new reaction tube to eliminate precipitated protein. The amount of soluble EPO was assessed by CD measurement. In figure 49, CD spectra of various time points during the incubation at 37 °C of each EPO variant are overlaid.

To present the results in more clarity, the following calculations were done. The CD signals at 208 nm and 222 nm before incubation were set to 100% and the decrease in signal intensity, which is directly proportional to the concentration of soluble protein, is listed in percentage in table 47. The percentage of soluble EPO during incubation at 37 °C is represented in a chart in figure 50.

The effect of PEGylation on EPO is best observed at position 24 (figure 50A). The percentage of soluble EPO24-PEG5kD is at least 11% higher than for n. g. EPO at each time point. After 144 h of incubation at 37 °C, 32% is still soluble, in comparison to 15% of n. g. EPO and only 6% of EPO24Pln at that time point. EPO38-PEG5kD is after 6 h of incubation still at 100%. However, at later time points, the recovery of soluble EPO is comparable to EPO38Pln. EPO83-PEG5kD is slightly more stable against thermal unfolding with a maximal difference to EPO83Pln of 15% after 72 h. The percentage respective to the starting point was 31% for EPO83-PEG5kD and only 16% for EPO83Pln at this time point.

208 nm	0 h	%	6 h	%	24 h	%	48 h	%	72 h	%	144 h	%
n. g. EPO	-26.7	100	-20.6	77	-16.1	60	-10.9	41	-8.0	30	-4.0	15
EPO24Pln	-23.0	100	-15.7	68	-9.8	43	-5.7	25	-3.5	15	-1.3	6
EPO24-PEG5kD	-25.8	100	-22.9	88	-20.9	81	-16.2	63	-12.5	48	-8.2	32
EPO38Pln	-7.2	100	-3.8	52	-2.0	28	-0.1	2	-0.2	3	0.1	-2
EPO38-PEG5kD	-6.5	100	-6.5	100	-1.8	28	0.0	0	-0.1	1	0.3	-4
EPO83Pln	-14.6	100	-10.6	73	-7.8	53	-4.2	29	-2.4	16	-1.0	7
EPO83-PEG5kD	-11.6	100	-9.1	78	-7.0	60	-3.9	33	-3.6	31	-1.0	9

222 nm	0 h	%	6 h	%	24 h	%	48 h	%	72 h	%	144 h	%
n. g. EPO	-20.5	100	-15.6	76	-12.0	59	-8.4	41	-6.0	29	-2.7	13
EPO24Pln	-16.8	100	-11.7	69	-6.9	41	-4.0	24	-1.9	11	-0.5	3
EPO24-PEG5kD	-20.2	100	-17.9	89	-16.2	80	-13.2	66	-9.6	48	-5.8	29
EPO38Pln	-5.3	100	-2.8	53	-1.1	21	-0.2	4	-0.2	4	0.4	-7
EPO38-PEG5kD	-5.3	100	-3.3	62	-1.2	23	-0.2	3	0.2	-4	0.1	-3
EPO83Pln	-11.6	100	-8.2	71	-5.6	48	-3.4	29	-1.5	13	-0.5	5
EPO83-PEG5kD	-9.5	100	-7.2	76	-5.8	61	-3.4	36	-3.1	33	-0.9	10

Table 47: CD signal intensity in mdeg at 208 nm and 222 nm and percentage of recovered soluble PEGylated EPO after various incubation time points at 37 °C

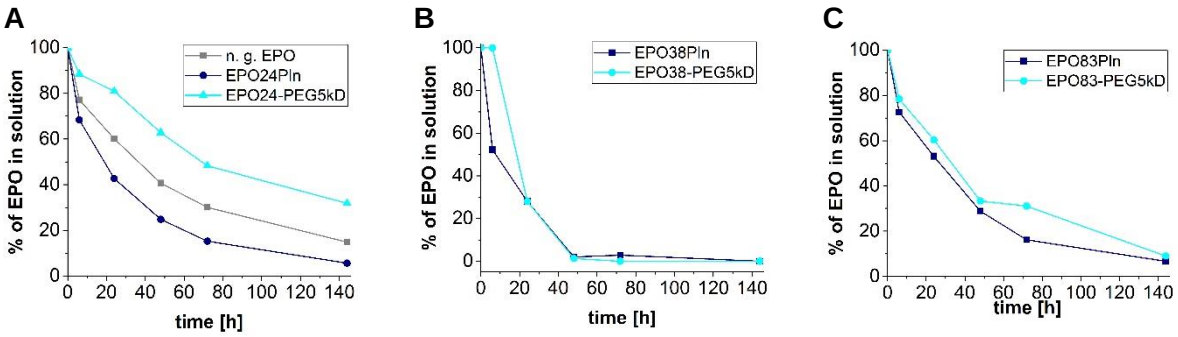
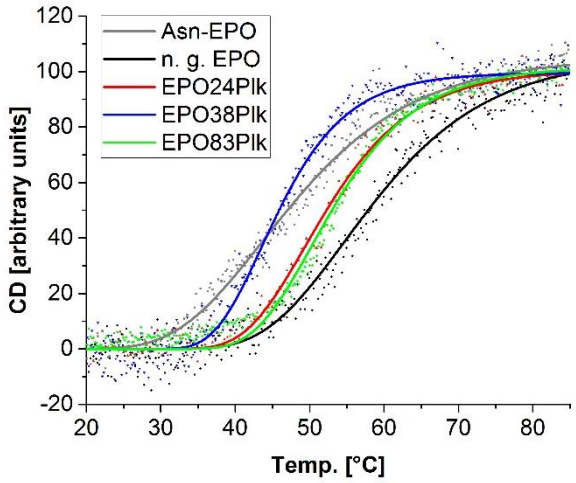


Figure 50: Percentage of soluble PEGylated EPO in comparison to soluble EPO-Pln during incubation at 37°C based on CD signals at 208 nm; PEGylation of EPO at position 24 (A), 38 (B) and 83 (C)

4.8.2. Melting curves

The apparent melting point for not glycosylated EPO is T_M = 56 °C and 12 °C higher than the not glycosylated AsnEPO (T_M = 44 °C) (see figure 51 and table 48). Furthermore, for AsnEPO the secondary structure starts to break up early at very low temperatures of about 30 to 35 °C. As the melting point decreases upon incorporation of the unnatural amino acid Plk of about 2 to 11 °C it was really necessary to use the EPO-sequence, which is shown in figure 62A. This sequence has the following amino acid mutations: N24K, N38K, N83K, P121N and P122S. For comparison, the AsnEPO amino acid sequence is depicted in figure 62B. The three N to K mutations at the positions that are naturally glycosylated, has been shown to decrease the aggregation of *E. coli*-derived EPO (Narhi, Arakawa et al. 2001). This finding could be confirmed by assessing the melting curves (figure 51). The additional two mutations P121N and P122S were inserted in an attempt to reduce the conformational heterogeneity in this loop based on a possible *cis-trans* isomerisation of the two prolines reported by (Cheetham, Smith et al. 1998). It was thought that this could be disadvantageous for the folding procedure. Additionally, the crystal structure of EPO with these five mutations was resolved by (Syed, Reid et al. 1998).



EPO variant	T_M [°C]
AsnEPO	44
n. g. EPO	56
EPO24Plk	53
EPO38Plk	45
EPO83Plk	54

Table 48: Apparent melting points of EPO-Plk variants

Figure 51: Melting curves of EPO-Plk variants in comparison to AsnEPO

4.9. Biological activity assays

4.9.1. Cell differentiation assay

In differentiation assays, haematopoietic stem cells from mouse bone marrow differentiate and proliferate upon stimulation with a cytokine cocktail including EPO. There are several possibilities for colonies that can be observed after eight days: BFU-E, CFU-M, CFU-M, CFU-GM, CFU-GEMM. In figure 52, obvious BFU-E are depicted, as they are clearly red from the present haemoglobin. In contrast, in figure 53, some of the other observed colony types are represented. As it was not always easy to identify the colony type, the plates were incubated with a benzidine staining solution, which stained haemoglobin-containing colonies dark blue (figure 54). To get an overview of benzidine positive colonies, the whole plates were photographed (figure 55).

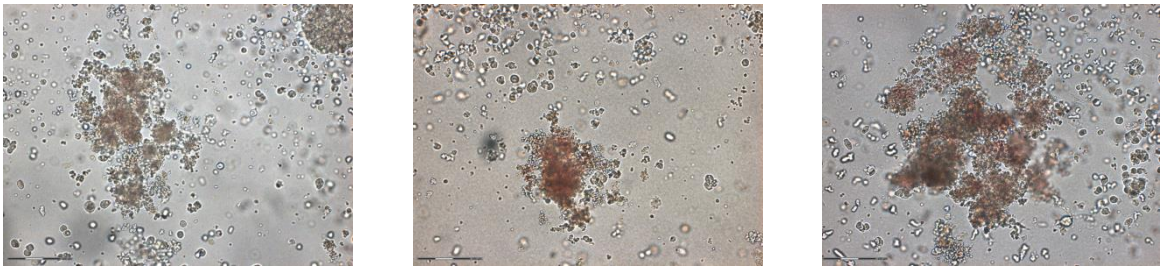


Figure 52: Examples of BFU-E

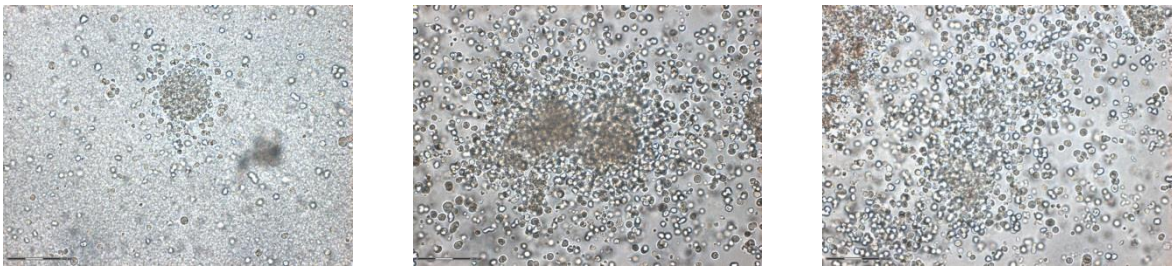


Figure 53: Examples for other cell colonies, which do not contain erythrocytes or their precursors

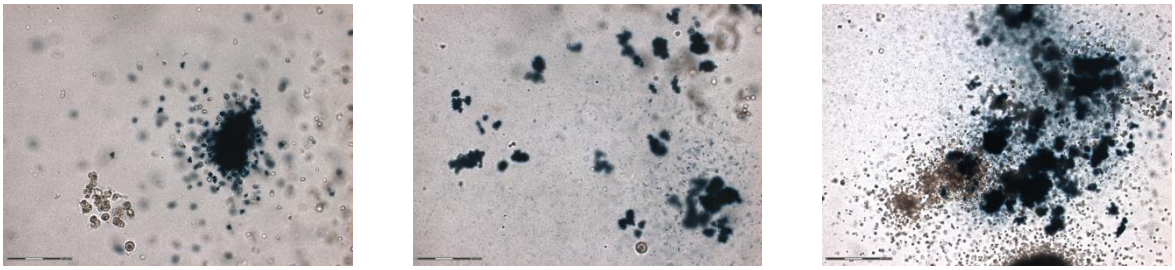


Figure 54: Examples for benzidine stained colonies containing haemoglobin, which is stained dark blue

It was tried to establish a dose-response curve with not glycosylated EPO and purchased CHO-derived EPO as a reference. In order to be able to test several EPO concentrations as triplicates, smaller plates were assessed. However, then the cell number decreases under a representative value. A representative dose-response curve could not be established because this system only responded totally or not. Therefore, the number of colonies can give only little hints about which EPO variant is more active than the other.

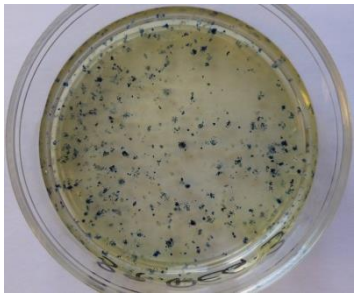


Figure 55: Example for a 35 mm plate containing methylcellulose medium and 8-day old colonies derived from mouse bone marrow stained with benzidine

However, it is clear that all tested EPO variants are biological active and produce a significant higher amount of benzidine-positive colonies than the negative control without any EPO (figure 56 and table 49). Moreover, all the produced EPO variants also show a higher degree of erythroid activity than the bought CHO-derived EPO, which was less active. EPO24-Undeca generates 274 colonies, clearly the highest amount of colonies in this experiment. This could be a hint that sialylated glycan chains increase the *in vitro* activity of EPO.

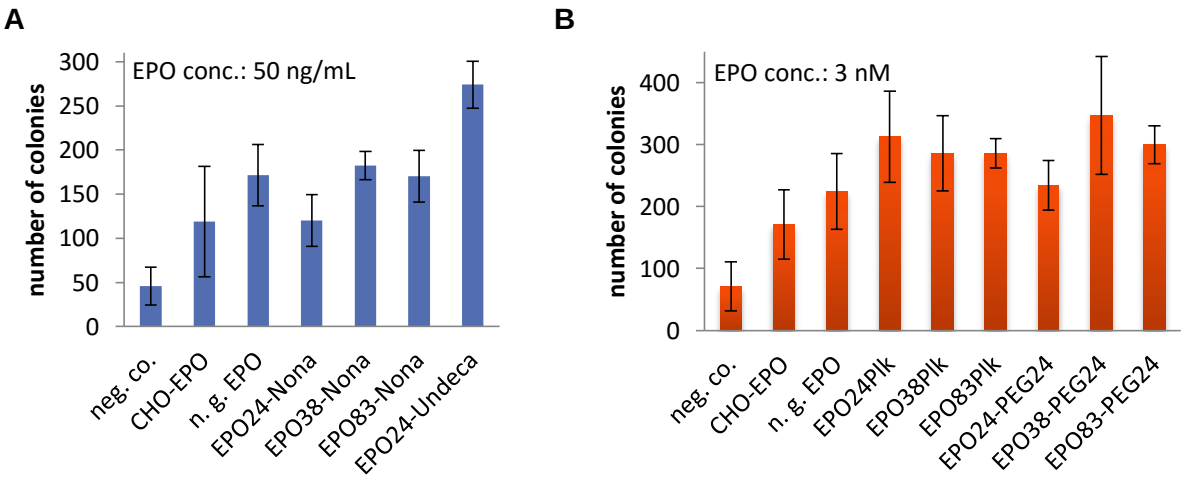


Figure 56: Amount of haemoglobin-containing colonies, n = 3 for each experiment, glycosylated EPO variants (A), PEGylated EPO variants (B)

Glycosylation experiment		PEGylation experiment	
EPO variant	**number of colonies**	**EPO variant**	**number of colonies**
negative control	46	negative control	71
CHO-EPO	119	CHO-EPO	171
n. g. EPO	172	n. g. EPO	224
EPO24-Nona	120	EPO24Plk	312
EPO38-Nona	182	EPO38Plk	286
EPO83-Nona	170	EPO83Plk	286
EPO24-Undeca	274	EPO24-PEG24	234
		EPO38-PEG24	347
		EPO83-PEG24	300

Table 49: Amount of haemoglobin-containing colonies, n = 3 for each experiment

Adding PEGylated EPO variants to the experiment indeed resulted in colony numbers until 347. However, the individual experiments cannot be compared to each other because the viability of primary cells varies from mouse to mouse. Coupling PEG24 to EPO does not significantly increase colony numbers compared to not glycosylated EPO or the EPO-Plk variants.

4.9.2. Cell proliferation assays

Although the differentiation assay represents the situation *in vivo* more exactly than a proliferation assay, such assays were performed with two different cell types to obtain a dose-response curve with a respective EC_{50}-value for each EPO variant. In this way, it was possible to compare all EPO variants to each other and to EC_{50}-values from literature.

4.9.2.1. TF-1 cell assays

Ten different EPO concentrations ranging from 0.01 to 500 ng/mL were used to stimulate 1 x 10^4 EPO dependent TF-1 cells in each well. Fully glycosylated EPO, derived from CHO cells, was employed as positive control in each experiment. The measured value of unstimulated TF-1 cells was set to 1 and the relative proliferation rate was plotted against the EPO concentration. The data were fitted to a non-cooperative binding reaction with a single binding site (Hill coefficient = 1). For each EPO variant, two values were considered to assess the biological activity: the EC_{50} and the maximal proliferation rate.

The results for glycosylated EPO and EPO-Plk in comparison with CHO-EPO are summarised in figure 57 and table 50. In all cases, CHO-EPO had the smallest EC_{50}-value (0.06 nM) and the highest maximal proliferation rate (3.77), in contrast to not glycosylated EPO, expressed in *E. coli*, which has an EC_{50}-value of 0.9 nM and a maximal proliferation rate of 3.0. The EPO variants with incorporated Plk have slightly higher EC_{50}-values, ranging from 1.2 to 1.8 nM and comparable maximal proliferation rates. The EPO-Plk are less active because the unnatural amino acid is replacing a natural one with different chemical properties. By comparing the effect of the incorporation site on the EC_{50}-value, the position 24 appeared to be the most tolerant.

For the glycosylated EPO variants, a trend is observable: One Nona glycan reduces the EC_{50}-value at least by half and one Undeca glycan by third to fifth at each of the three glycosylation positions, whereby the maximal proliferation rates constantly stayed around 3.0. In addition, both glycans, Nona and Undeca, at position 24 show the lowest EC_{50}-values of 0.48 and 0.23 nM, respectively. For Undeca, even the maximal proliferation rate is elevated to 3.2 compared to 2.97 for not glycosylated EPO.

To sum it up, only one glycan with a molecular size of 1.7 kDa is able to ameliorate the cell proliferation activity of EPO on TF-1 cells. This positive effect in biological activity can be increased by adding two terminal sialic acids, which is the only difference between Nona and Undeca.

If comparing glycosylation at each position, glycosylation at position 24 showed the highest decrease of the EC_{50}-value and glycosylation on positions 38 and 83 are comparable good.

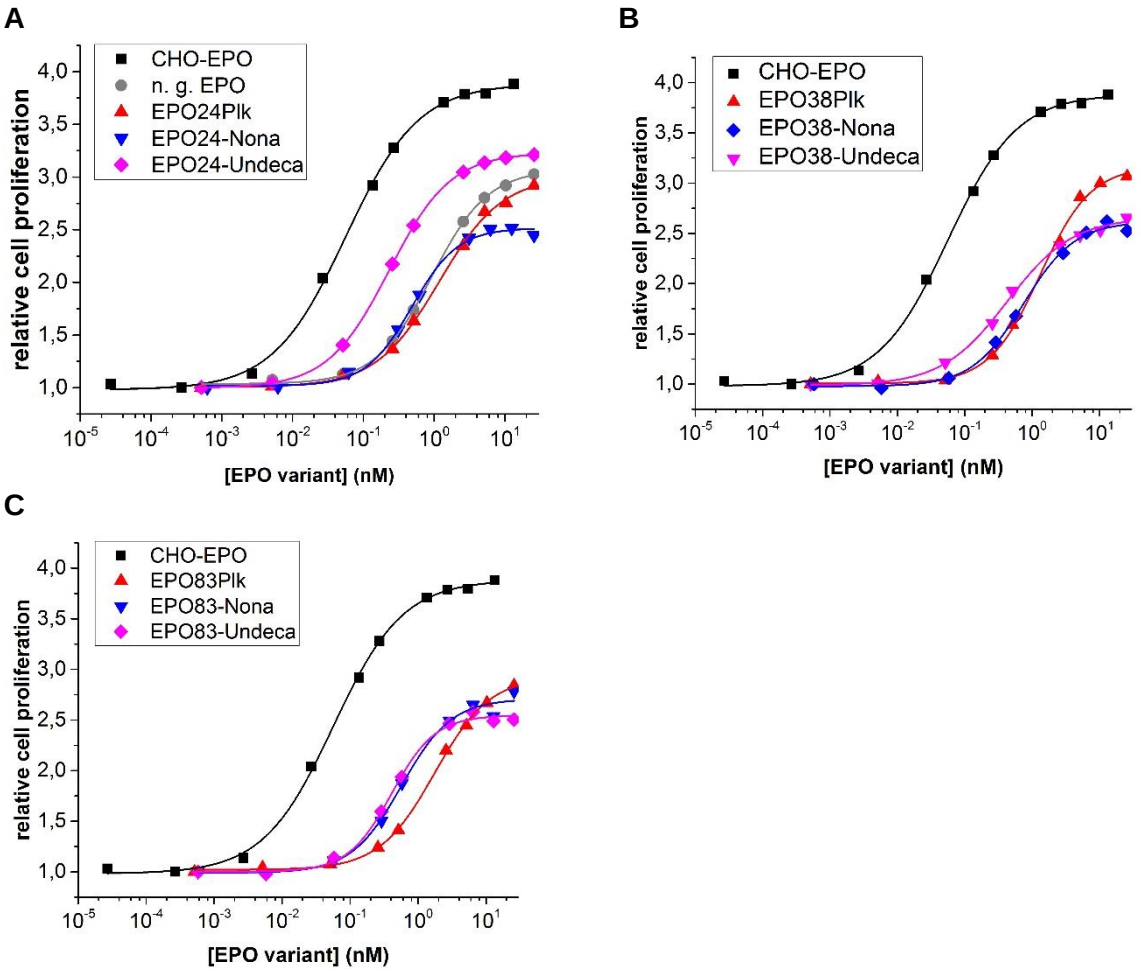


Figure 57: Dose-response curves of glycosylated EPO variants obtained from alamarBlue assays with TF-1 cells. Glycosylation at position 24 (A), 38 (B) and 83 (C). n = 3

EPO Variant	EC_{50} (nM)	max. cell proliferation	% max. cell proliferation
CHO-EPO	**0.056** ± 0.003	3.77 ± 0.03	100
n. g. EPO	**0.923** ± 0.042	2.97 ± 0.02	79
EPO24Plk	**1.189** ± 0.100	2.85 ± 0.04	75
EPO24-Nona	**0.483** ± 0.037	2.52 ± 0.03	67
EPO24-Undeca	**0.233** ± 0.017	3.22 ± 0.03	85
EPO38Plk	**1.321** ± 0.117	3.10 ± 0.05	82
EPO38-Nona	**0.742** ± 0.081	2.62 ± 0.05	69
EPO38-Undeca	**0.434** ± 0.038	2.67 ± 0.03	71
EPO83Plk	**1.753** ± 0.106	2.93 ± 0.04	78
EPO83-Nona	**0.570** ± 0.085	2.71 ± 0.06	72
EPO83-Undeca	**0.410** ± 0.032	2.55 ± 0.03	67

Table 50: EC_{50}-values and maximal cell proliferation rate of TF-1 stimulated with various glycosylated EPO variants, n = 3

To prove the reliably of the alamarBlue assay, which measures the respiratory activity of the TF-1 cells, the relative cell proliferation was measured by simultaneously performed ^{3}H-thymidine assays from the same cell culture plates. ^{3}H-thymidine is taken up and incorporated by the cells and can be quantified subsequently. The same trend is visible as in the alamarBlue assay (data not shown). CHO-EPO was revealed to have the best EC_{50}-value and maximal proliferation rate. The incorporation of Plk had no or a very little effect on both values. Glycosylation generally reduced the EC_{50}-values, no matter which glycan was introduced at which position. EPO24-Undeca was the most active produced EPO variant irrespective of the measurement method for cell proliferation.

Results from both assays are consistent, however the alamarBlue assay is more reliable than the ^{3}H-thymidine assay because of higher sensitivity. Additionally, radioactive tritium was necessary for the ^{3}H-thymidine assay, therefore experiments were continued by assessing the relative proliferation rate by non-toxic alamarBlue.

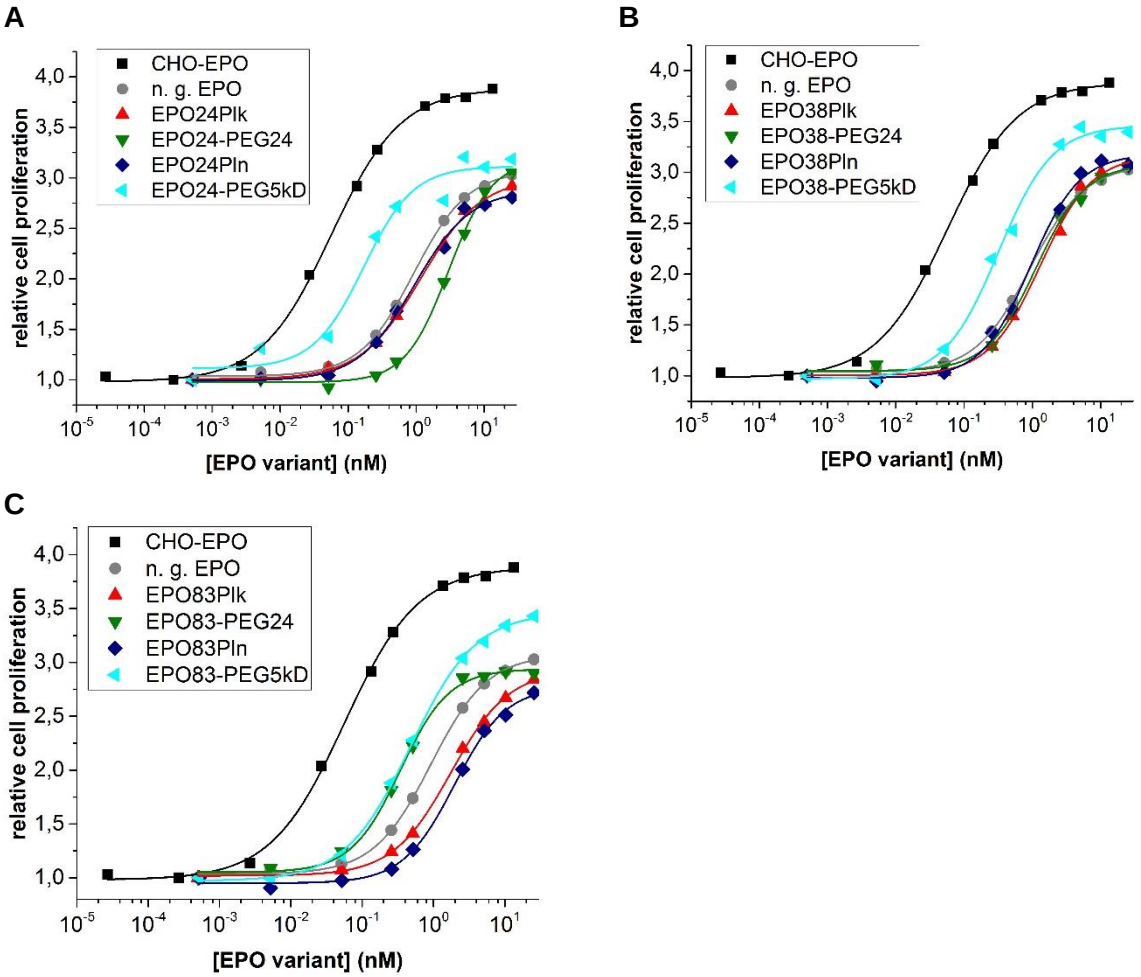


Figure 58: Dose-response curves of PEGylated EPO variants obtained from alamarBlue assays with TF-1 cells. PEGylation at position 24 (A), 38 (B) and 83 (C). n = 3

EC_{50}-values and the maximal proliferation rates were also determined for PEGylated EPO variants (figure 58 and table 51). Again, CHO-derived fully-glycosylated EPO was used as positive control and not glycosylated EPO as negative control. Two different PEG-chains were coupled to EPO: The PEG24

to EPO-Plk and the 5-kDa-PEG chain to EPO-Pln. Similar to Plk, Pln has also not a significant effect on EC_{50}-values and maximal proliferation rates. All three natural glycosylation positions were tested. The PEG24 with a size of 1.1 kDa has no or only little effect, as the respective curves of PEGylated and not PEGylated EPO are overlying in figure 58B and are relative near to each other in figure 58A and C. In contrast, the PEG5kD has a significant effect on the EC_{50}-values and the maximal proliferation rates at each of the three positions: The PEG chain is responsible for a 2- to 6-fold decrease of EC_{50}-values depended on the PEGylation position. The resulting EC_{50}-values ranging from 0.17 nM (for EPO24-PEG5kD) to 0.48 nM (for EPO83-PEG5kD) are compared to not glycosylated EPO (0.92 nM). PEGylation with the 5kD-PEG chain elevated the maximal proliferation rates from 3.07 (EPO without decorations) to 3.11 (EPO24-PEG5kD), 3.46 (EPO38-PEG5kD and EPO83-PEG5kD).

EPO Variant	EC_{50} (nM)	max. cell proliferation	% max. cell proliferation
CHO-EPO	**0.056** ± 0.003	3.89 ± 0.03	100
n. g. EPO	**0.923** ± 0.047	3.07 ± 0.03	79
EPO24Plk	**1.177** ± 0.098	3.00 ± 0.05	77
EPO24-PEG24	**2.995** ± 0.176	3.18 ± 0.06	82
EPO24Pln	**0.943** ± 0.137	2.88 ± 0.07	74
EPO24-PEG5kD	**0.169** ± 0.049	3.11 ± 0.10	80
EPO38Plk	**1.316** ± 0.128	3.17 ± 0.06	82
EPO38-PEG24	**1.144** ± 0.120	3.08 ± 0.06	79
EPO38Pln	**0.944** ± 0.111	3.18 ± 0.07	82
EPO38-PEG5kD	**0.319** ± 0.036	3.46 ± 0.07	89
EPO83Plk	**1.753** ± 0.107	2.93 ± 0.04	75
EPO83-PEG24	**0.342** ± 0.022	2.94 ± 0.03	76
EPO83Pln	**1.969** ± 0.177	2.77 ± 0.06	71
EPO83-PEG5kD	**0.477** ± 0.022	3.46 ± 0.03	89

Table 51: EC_{50}-values and maximal cell proliferation rate of TF-1 stimulated with various PEGylated EPO variants

To facilitate comparisons between glycosylated and PEGylated EPO, the dose response curves for EPO-Undeca and EPO-PEG5kD variants from figures 57 and 58 are overlaid in figure 59 for each modification position. In table 52, the EC_{50}-values and the maximal cell proliferation rates of these two modifications are listed next to each other.

In contrast to the other two positions, the dose-response curves of EPO24-PEG5kD and EPO24Undeca are very close to each other (see figure 59A). Their EC_{50}-values and maximal proliferation rates are very similar (around 0.2 nM and 80-85% of the maximal proliferation rate of CHO-EPO). Although the EC_{50}-values of these two EPO variants for positions 38 and 83 are corresponding to each other as well, the maximal proliferation rates are distinct (see figures 59B

and C). The EC_{50}-values are ranging from 0.3 nM to 0.5 nM for these two decorations and positions, whereas the maximal proliferation rate is 89% for EPO38-PEG5kD and EPO83-PEG5kD of the maximal proliferation rate of CHO-EPO. For EPO38-Undeca and EPO83-Undeca, the maximal proliferation rate is reduced to 71% and 67%.

To sum it up, decorations at position 24 showed the highest effects on the dose-response curves. They are overlaying for EPO24-Undeca and EPO24-PEG5kD. It can be concluded that an oligosaccharide with two terminal sialic acids and a molecular weight of about 2.2 kDa has the same effect as a 5-kDa-PEG at the natural glycosylation position 24. For the other two modification positions, the EC_{50}-values are slightly higher but comparable to each other (Undeca and PEG5kDa). However, the maximal proliferation rate stays above 80% for PEG5kDa. In contrast, it is reduced to around 70% for Undeca.

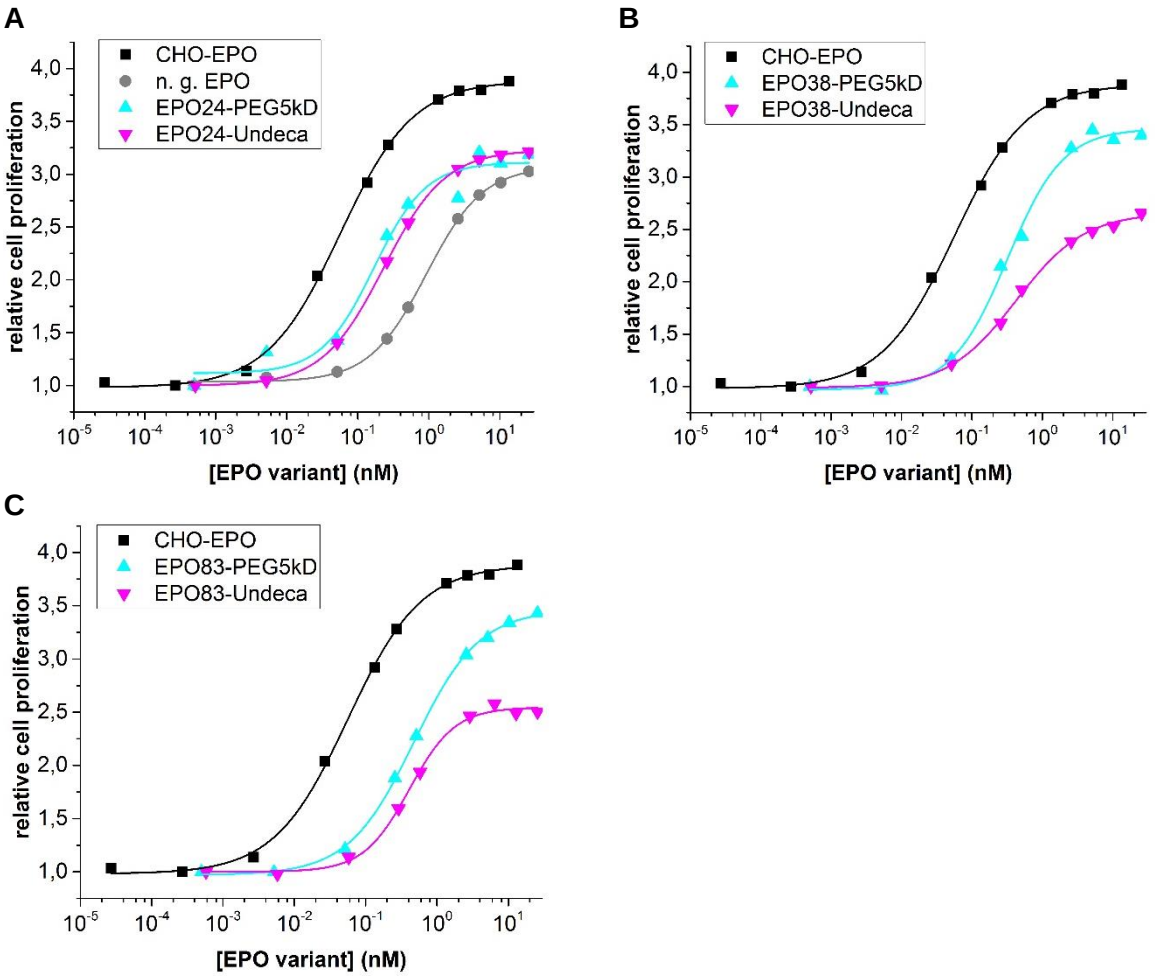


Figure 59: Comparison of the dose-response curves of EPO-PEG5kD and EPO-Undeca obtained from alamarBlue assays with TF-1 cells. Decoration at position 24 (A), 38 (B) and 83 (C). n = 3

EPO Variant	EC_{50} (nM)	max. cell proliferation	% max. cell proliferation
CHO-EPO	**0.056** ± 0.003	3.89 ± 0.03	100
n. g. EPO	**0.923** ± 0.047	3.07 ± 0.03	79
EPO24-PEG5kD	**0.169** ± 0.049	3.11 ± 0.10	80
EPO24-Undeca	**0.233** ± 0.017	3.22 ± 0.03	85
EPO38-PEG5kD	**0.319** ± 0.036	3.46 ± 0.07	89
EPO38Undeca	**0.434** ± 0.038	2.67 ± 0.03	71
EPO83-PEG5kD	**0.477** ± 0.022	3.46 ± 0.03	89
EPO83-Undeca	**0.410** ± 0.032	2.55 ± 0.03	67

Table 52: Comparison of the EC_{50}-values and maximal cell proliferation rate of TF-1 stimulated with EPO-PEG5kD and EPO-Undeca variants

4.9.2.2. UT-7 cell assays

Additionally, most of the produced EPO variants were sent to Prof. Jelkmann, University of Lübeck for testing cell proliferation by MTT assay with UT-7/EPO cells. In the MTT assay, the metabolic activity of the cells is measured, based on the reduction of MTT to formazan. Thus, this assay is comparable to the alamarBlue assay. UT-7 is a human leukemic cell line, which grows when IL-3, GM-CSF or EPO is supplemented (Komatsu, Nakauchi et al. 1991). Later, the same group established a subline UT-7/EPO that is solely EPO-dependent (Komatsu, Yamamoto et al. 1993). The cells were incubated in 96-well plates without any serum nor EPO for 16 h. Then, the EPO variants were supplemented in various concentrations and the cells were further incubated for one week. Finally, MTT was added for absorbance measurement of formazan at 570 nm.

The respective EC_{50}-values and maximal proliferation rates are represented in table 53. As a positive control, Epoetin-β was used and its maximal proliferation rate was set to 100%. The dose-response curve of Epoetin-β is very close to those of Komatsu et al. (Komatsu, Yamamoto et al. 1993).

An outstanding maximal proliferation rate of 134% was found for EPO24-Undeca. However, its EC_{50}-value was 0.136 nM. This is five times more than the EC_{50}-value for Epoetin-β (0.026 nM). EPO24Plk was as potent as the negative control n. g. EPO (without unnatural amino acid). EPO24-Nona has an EC_{50}-value of 2.69 nM due to unknown reasons.

All variants with Plk and/or glycosylation at position 38 showed a smaller potency compared to the other two positions.

The maximal proliferation rates for EPO83Plk, EPO83-Nona and EPO83-Undeca are slightly higher than for Epoetin-β. However, the EC_{50}-values ranged from 0.51 to 1.0 nM. Hence, they are averagely three times higher than for Epoetin-β.

EPO Variant	EC_{50} (nM)	max. cell proliferation	% max. cell proliferation
Epoetin-beta	0.026 ± 0.003	9.75 ± 0.60	100
n. g. EPO	0.373 ± 0.049	9.69 ± 0.35	99
EPO24Plk	0.435 ± 0.069	9.86 ± 0.54	101
EPO24-Nona	2.686 ± 0.042	7.36 ± 0.09	76
EPO24-Undeca	0.136 ± 0.010	13.01 ± 0.13	134
n. g. EPO	0.272 ± 0.032	4.65 ± 0.09	48
EPO38Plk	0.273 ± 0.046	3.33 ± 0.10	34
EPO38-Nona	0.293 ± 0.038	3.19 ± 0.07	33
EPO38-Undeca	Determination was not possible.		
n. g. EPO	0.272 ± 0.032	4.65 ± 0.09	48
EPO83Plk	0.836 ± 0.036	9.76 ± 0.13	100
EPO83-Nona	1.014 ± 0.008	10.29 ± 0.03	106
EPO83-Undeca	0.511 ± 5423	10.10 ± 0.17	104

Table 53: EC_{50}-values and maximal cell proliferation rate of UT-7 cells stimulated with various glycosylated EPO variants

4.9.3. Conclusions of biological activity assays

Two different cell assays, a differentiation and a proliferation assay, were performed on the produced EPO variants in parallel in order to elucidate distinct properties of them. The cell differentiation assay is a closer model to the real conditions found *in vivo*. Thus, the experiments take place in a more complex model with many variables that could not be standardised easily. By the establishment of the cell differentiation assay for the produced EPO variants, it was possible to demonstrate that they stimulate the cell differentiation from BFU-E to haemoglobin containing reticulocytes. Interestingly, the produced EPO variants were all more active than CHO-EPO, whereas in cell proliferation assays, CHO-EPO showed the maximal proliferation rate. Hence, glycosylation of EPO has the opposite effect in these two assays. A possible explanation to this could be that binding of EPO to its receptor once at the beginning of the experiment is sufficient to trigger differentiation until reticulocyte development, whereas there are constantly new cells arising in the cell proliferation assays, which need to be stimulated by EPO continuously. In cell proliferation assays, EPO needs to be functionally active over a longer period compared to cell differentiation assays. It is thought that the positive effect of glycosylation on the stability of EPO against thermal unfolding and proteolysis is higher than the negative effect of glycosylation on binding to the receptor.

It was found that only one glycan with a molecular size of 1.7 kDa is able to ameliorate the cell proliferation activity of EPO on TF-1 cells. This positive effect in biological activity can be increased by adding two terminal sialic acids, which is the only difference between Nona and Undeca. Outstanding effects on cell proliferation and differentiation could be shown in all assays for EPO24-Undeca. In the UT-7/MTT assay, this EPO variant even stimulate the cell proliferation more than the positive control

Epoetin-β and in cell differentiation assays, its addition produces the highest amount of benzidine-positive colonies.

The results represented in the previous chapters are compared to EC_{50}-values found in literature:

- Kent et al. tested the *in vitro* biological activity of their synthetic not glycosylated EPO in a cell proliferation assay with TF-1 cells and quantified the uptake of ^{3}H-thymidine (Liu, Pentelute et al. 2012). They found an EC_{50}-value of 0.02 µg/mL or 1.09 nM, which is very close to the result of 0.923 nM for n. g. EPO in the alamarBlue assay.
- Sjodin et al. tested CHO-derived fully glycosylated EPO and calculated an EC_{50}-value of about 18 pM, which is about half or third the values found here (42 pM in ^{3}H-thymidine uptake and 56 pM in alamarBlue assay) (Hammerling, Kroon et al. 1996).
- Kajihara et al. synthesised an EPO analogue with two complex-type sialyloligosaccharides at positions 24 and 30 (Hirano, Macmillan et al. 2009). They used the TF-1 cell proliferation assay and found an EC_{50}-value of 4 ng/mL, which correspond to 0.2 nM. This value correlates to the identified EC_{50}-values of 0.2 nM for EPO24-Undeca and 0.4 nM for EPO38-Undeca and EPO83-Undeca. Moreover, they also assessed CHO-derived fully glycosylated EPO as a control. They detected an EC_{50}-value of 4 ng/mL, which correspond to 0.1 nM. This is double the EC_{50}-value of 0.056 nM, which was determined here for CHO-derived fully glycosylated EPO.

5. Summary and Outlook

5.1. Summary

A new semi-synthetic approach was presented to produce homogeneously glycosylated EPO with the help of unnatural amino acids and click chemistry. As bacterial cells do not posttranslationally glycosylate proteins, *E. coli* cells were used to incorporate non-natural amino acids into EPO by amber stop codon suppression methodology. These non-natural amino acids bear a specific bio-orthogonal chemical function, for example the pyrrolysine derivative Plk bearing an alkyne group. After expression and purification, EPO could be coupled at the natural glycosylation sites to defined synthesized oligosaccharides or purchased PEG chains by copper-catalysed 1,3-dipolar Huisgen cycloaddition between alkynes and azides.

Each individual step from synthesizing the unnatural amino acids, amber stop codon suppression in *E. coli*, purification, refolding to click chemistry had to be optimised in order to produce homogeneously glycosylated or PEGylated EPO in high yields and in a convenient, low-cost manner. For reasons of clarity, an overview of all the steps of the procedure is depicted in figure 60.

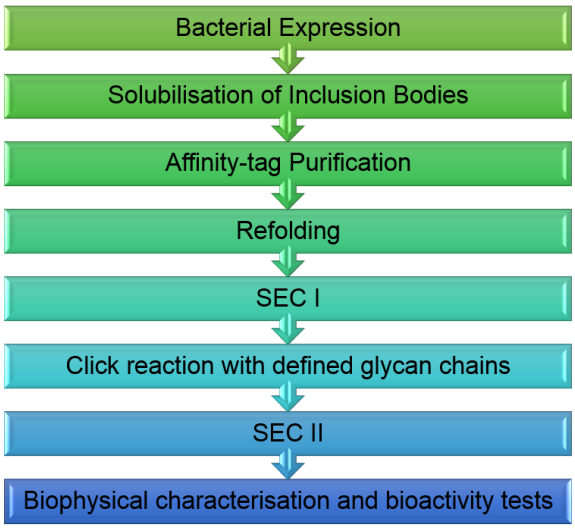


Figure 60: Overview of the procedure from bacterial expression of EPO to the characterisation of homogeneously glycosylated EPO

The generated EPO variants were characterised biophysically by mass spectrometry and circular dichroism (CD). All of them displayed the correct molecular mass and the secondary structure of EPO is not disturbed by incorporated unnatural amino acids or coupled decorations. Moreover, CD spectra measurements showed a large protective effect of one short glycan (under 2 kDa) or of a 5kDa-PEG on the secondary structure of EPO upon repeated freezing-thawing cycles.

The biological activity was investigated by cell differentiation and proliferation assays with different cell lines. It could be shown that all EPO variants have a positive effect on cell differentiation of haematopoietic stem cells from mouse bone marrow.

Results that are more detailed were found in cell proliferation assays. In contrast to small PEG-chains, only one coupled glycan with a molecular weight of 1.7 kDa is able to decrease the EC_{50}-value, and consequently, to increase the biological activity of EPO. Two terminal sialic acids even further augment this effect.

In detail, the effect of the glycan chain Undeca to the dose-response curve of EPO is comparable to the effect of a 5-kDa-PEG (see figure 59 and table 52). These two decorations could have a similar umbrella effect on the corresponding EPO surface. Whereas PEG24 is too small to show a similar effect on the biological activity and the glycan chain Nona without terminal sialic acids showed an intermediate effect (see figures 57-58 and tables 50-51).

To sum it up, a new procedure for engineering glycoproteins, such as EPO, is presented, which combines non-canonical amino acids and click chemistry. With this new methodology in hand, first steps were undertaken towards elucidating the impact of each glycosylation position and pattern on the function of EPO, as an example for therapeutic glycoproteins.

5.2. Outlook

For further knowledge, *in vivo* studies in mice are necessary to learn about the circulatory half-life of the EPO variants and their effect on the haematocrit and the percentage of reticulocytes in blood.

Furthermore, two or three identical oligosaccharides could be coupled to EPO. It would be convenient to use the SPI methodology to incorporate HPG at three or even more positions because the expression yields are not decreased for more than one unnatural amino acid incorporation, as it is the case for amber stop codon suppression. However, another cleavable tag should be use to remove the start HPG and the polyhistidine tag, as the enterokinase cleavage was not successful. For example, an intein could be used for this purpose. Subsequently, the produced EPO with HPG could be coupled by click chemistry, as it is described for EPO-Plk in this thesis.

In addition, the experiment using click chemistry to couple EPO to a 20-kDa PEG chain could be repeated in concentrations high enough for purification and subsequent characterisation. All further generated EPO variants could be compared with the described EPO variants from this study.

A methodology combination of the amber stop codon and the selective pressure incorporation would allow incorporating two or more different unnatural amino acids with specific chemical functions at defined positions in order to couple EPO with two different decorations. For example, *para*-azidophenylalanine could be incorporated in response to a TAG stop codon and simultaneously, HPG in response to an ATG codon by SPI. As Staudinger phosphite reaction is possible under denaturising conditions, phosphite activated PEG chains could be coupled to *para*-azidophenylalanine in high concentrations of GdmCl. After refolding, azido-glycans could be reacted to the alkyne group of the HPG residues by click chemistry. In this way, various combination possibilities to design EPO with position specific decorations are available to elucidate synergistic or additives effects of oligosaccharides and/or PEG chains on EPO.

6. Appendix

6.1. Sequences

6.1.1. EPO sequence

```
        10          20          30          40          50          60
ATGGCACCGC CTCGTCTGAT TTGTGATAGC CGTGTTCTGG AACGTTATCT GCTGGAAGCA

        70          80          90         100         110         120
AAAGAAGCCG AAAAAATTAC CACCGGTTGT GCAGAACATT GTAGCCTGAA TGAAAAAATT

       130         140         150         160         170         180
ACAGTGCCGG ATACCAAAGT GAATTTTTAT GCCTGGAAAC GTATGGAAGT TGGTCAGCAG

       190         200         210         220         230         240
GCAGTTGAAG TTTGGCAGGG TCTGGCACTG CTGAGCGAAG CAGTTCTGCG TGGTCAGGCA

       250         260         270         280         290         300
CTGCTGGTTA AAAGCAGCCA GCCGTGGGAA CCGCTGCAGC TGCATGTTGA TAAAGCAGTT

       310         320         330         340         350         360
AGCGGTCTGC GTAGCCTGAC CACCCTGCTG CGTGCACTGC GTGCCCAGAA AGAAGCAATT

       370         380         390         400         410         420
TCTAATAGCG ATGCAGCATC TGCAGCACCG CTGCGTACCA TTACCGCAGA TACCTTTCGT

       430         440         450         460         470         480
AAACTGTTTC GCGTGTATAG CAATTTTCTG CGTGGCAAAC TGAAACTGTA TACCGGTGAA

       490         500         510         520
GCATGTCGTA CCGGTGATCG TCATCACCAT CATCATCATT AA
```

Figure 61: Codon optimised EPO sequence; the yellow highlighted triplets were exchanged by a TAG stop codon individually to obtain three EPO sequences with one TAG stop codon each. For incorporation of HPG, they were replaced by the ATG-codon and the polyhistidine tag was moved to the N-terminus.

A Amino acid sequence of EPO

```
         10          20          30          40          50          60
APPRLICDSR VLERYLLEAK EAEKITTGCA EHCSLNEKIT VPDTKVNFYA WKRMEVGQQA

         70          80          90         100         110         120
VEVWQGLALL SEAVLRGQAL LVKSSQPWEP LQLHVDKAVS GLRSLTTLLR ALRAQKEAIS

        130         140         150         160         170
NSDAASAAPL RTITADTFRK LFRVYSNFLR GKLKLYTGEA CRTGDRHHHH HH
```

B Amino acid sequence of AsnEPO

```
         10          20          30          40          50          60
APPRLICDSR VLERYLLEAK EAENITTGCA EHCSLNENIT VPDTKVNFYA WKRMEVGQQA

         70          80          90         100         110         120
VEVWQGLALL SEAVLRGQAL LVNSSQPWEP LQLHVDKAVS GLRSLTTLLR ALGAQKEAIS

        130         140         150         160         170
PPDAASAAPL RTITADTFRK LFRVYSNFLR GKLKLYTGEA CRTGDRHHHH HH
```

C Amino acid sequence of HPG-EPO

```
         10          20          30          40          50          60
MAKHHHHHHD DDDKAPPRLI CDSRVLERYL LEAKEAEMIT TGCAEHCSLN EMITVPDTKV

         70          80          90         100         110         120
NFYAWKRLEV GQQAVEVWQG LALLSEAVLR GQALLVMSSQ PWEPLQLHVD KAVSGLRSLT

        130         140         150         160         170         180
TLLRALRAQK EAISNSDAAS AAPLRTITAD TFRKLFRVYS NFLRGKLKLY TGEACRTGDR
```

Figure 62: Amino acid sequences of EPO (A), AsnEPO (B); at the positions of the yellow highlighted amino acids an unnatural amino acid was incorporated individually in response to a TAG stop codon to obtain three EPO proteins with an unnatural amino acid at one of the following positions: 24, 38 and 83. For the amino acid sequence of HPG-EPO (C) the three highlighted methionine residues (M) were replaced by HPG simultaneously.

6.2. Index of abbreviations

General abbreviations	
AP	alkaline phosphatase
A-site	aminoacyl-site
CD	circular dichroism
CERA	continuous erythropoietin receptor activator
CM	carboxymethyl group
conc.	concentration
DNA	deoxyribonucleic acid
DSMZ	„Deutsche Sammlung von Mikroorganismen und Zellkulturen“
E. coli	Escherichia coli
EF	elongation factor
ER	endoplasmic reticulum
E-site	exit-site
FBS	foetal bovine serum
GDP	guanosine diphosphate

GTP	guanosine triphosphate
i.	induced
IEC	ion exchange chromatography
IF	initiation factor
IF	initiation factor
IMDM	Iscove's Modified Dulbecco's Medium
INN	International Nonproprietary Name
L	load
LB	lysogeny broth
M	marker
MCS	multiple cloning site
mRNA	messenger ribonucleic acid
n. i.	not induced
NMR	nuclear magnetic resonance
OD	optical density
OST	oligosaccharyl-transferase
PAGE	polyacrylamide gel electrophoresis
pI	isoelectric point
PNK	T4 polynucleotide kinase
P-site	peptidyl-site
PylRS	pyrrolysyl-tRNA synthetase
R	resistance
RF	release factor
RNA	ribonucleic acid
RP-HPLC	reversed phase high performance liquid chromatography
rRNA	ribosomal ribonucleic acid
RT	room temperature
S	sedimentation coefficient
SEC	size exclusion chromatography
SOC	super optimal broth with glucose
SP	sulfopropyl group
SPI	selective pressure incorporation
TBS	Tris-buffered saline
tRNA	transfer ribonucleic acid
Uni	university
W	washing fraction
WHO	World Health Organization
wt	wild type

Table 54: General abbreviations

Chemicals	
^{3}H	tritium
AHA	L-azidohomoalanine
Amp	ampicillin
APS	ammonium persulphate
BCIP	5-bromo-4-chloro-3-indolyl phosphate
Boc-Lys-OH	N_{α}-(*tert*-butoxycarbonyl)-L-lysine
Boc-Plk	(S)-2-(*tert*-butoxycarbonylamino)-6-((prop-2-ynyloxy)carbonylamino)hexanoic acid
$CaCl_2$	calcium chloride
Cam	chloramphenicol

Carb	carbencillin
Cu(I)-complex	tetrakis(acetonitrile)copper(I) hexafluorophosphate or tetrafluoroborate
$CuSO_4$	copper(II) sulphate
D_2O	deuterium oxide (heavy water)
DCM	dichloromethane
DMF	dimethylformamide
EDTA	ethylenediaminetetraacetic acid
Et_2O	diethyl ether
EtOAc	ethyl acetate
GdmCl	guanidinium hydrochloride
GlcNAc	*N*-acetylglucosamine
H_2O_2	hydrogen peroxide
HCl	hydrochloric acid
HPG	L-homopropargylglycine
Kan	kanamycin
$KHCO_3$	potassium bicarbonate
MeOH	methanol
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
MOPS	3-(*N*-Morpholino)propanesulfonic acid
NaCl	sodium chloride
NaN_3	sodium azide
NaOH	sodium hydroxide
NBT	nitro blue tetrazolium chloride
NH_4Cl	ammonium chloride
NTA	nitrilotriacetic acid
PEG	polyethylene glycol
Plk	N^6-((prop-2-yn-1-yloxy)carbonyl)-L-lysine
Pln	N^6-((2-azidoethoxy)carbonyl)-L-lysine
PMSF	phenylmethylsulfonyl fluoride
SDS	sodium dodecyl sulphate
TBTA	tris(benzyltriazolmethyl)amine
TCEP	tris(2-carboxyethyl)phosphine
TEMED	tetramethylethylenediamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
THPTA	tris(3-hydroxypropyltriazolylmethyl)amine
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol

Table 55: Abbreviation of chemicals

Abbreviations concerning cells or cytokines	
BFU-E	Burst forming unit - erythroid
BHK	Baby hamster kidney
CFU-E	Colony forming unit - erythroid
CFU-G	Colony forming unit – granulocyte
CFU-GEMM	Colony forming unit – granulocyte/erythrocyte/monocyte/megakaryocyte
CFU-GM	Colony forming unit – granulocyte/macrophage
CFU-M	Colony forming unit – macrophage
CHO	Chinese hamster ovary
EPO	Erythropoietin

EPOR	Erythropoietin receptor
ESA	Erythropoiesis-stimulating agent
GM-CSF	Granulocyte-macrophage colony-stimulating factor
HIF	Hypoxia inducible factor
IGF	Insulin-like growth factor
IL	Interleukin
JAK	Janus kinase
MAPK	Mitogen-activating protein kinase
MIP	Macrophage inflammatory protein
PHD	Prolyl hydroxylase domain
PI3K	Phosphatidyl-inositol 3-kinase
SCF	Stem cell factor
SCF	Stem cell factor
STAT5	Signal transducer and activator of transcription 5
TGF	Transforming growth factor

Table 56: Abbreviations concerning cells or cytokines

6.3. Index of figures

6.4. Index of tables

7. Acknowledgement

This dissertation was worked out in the section of organic and cellular chemistry at the department of chemistry from University of Konstanz from 2012 to 2015.

First, I would like to express my deep gratitude to my “Doktormutter” Dr. Marina Rubini. She enabled my PhD position in her group and gave me continuous support and advices concerning not only science, but also life. I have learned so many things from you. Thank you!

I would like to thank Prof. Dr. Andreas Marx for integrating me in his research group and the financial and personal support.

Thanks to Prof. Dr. Jörg Hartig responsible for the second opinion.

I would like to thank my team colleague Eugenia Hoffmann for our scientific discussions about our very special “model protein” erythropoietin. Working on this protein was not always easy. It was motivating to have a co-fighter against protein precipitation.

Thanks to Vlasta Radusevic and Anke Gerull for excellent technical help. Especially, I would like to express my gratitude to Vlasta, who always had open ears for me. Moreover, thank you for all the coffee and cookies!

I would like to thank Prof. Dr. Carlo Unverzagt and Prof. Dr. Wolfgang Jelkmann for their collaboration.

Furthermore, I want to thank the whole Brunner group, especially Prof. Dr. Thomas Brunner and Carina Seitz for their collaboration.

Thanks to all members of the Marx group for the inspiring working atmosphere.

Lastly, I would like to express my heartfelt gratitude to my family and friends, especially my parents and Daniel Buchlaub, for their continuous encouragement and support.

8. Bibliography

- Akashi, K., D. Traver, T. Miyamoto and I. L. Weissman (2000). "A clonogenic common myeloid progenitor that gives rise to all myeloid lineages." Nature **404**(6774): 193-197.
- Apweiler, R., H. Hermjakob and N. Sharon (1999). "On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database." Biochim Biophys Acta **1473**(1): 4-8.
- Arnold, J. N., M. R. Wormald, R. B. Sim, P. M. Rudd and R. A. Dwek (2007). "The impact of glycosylation on the biological function and structure of human immunoglobulins." Annu Rev Immunol **25**: 21-50.
- Bachmann, S., M. Le Hir and K. U. Eckardt (1993). "Co-localization of erythropoietin mRNA and ecto-5'-nucleotidase immunoreactivity in peritubular cells of rat renal cortex indicates that fibroblasts produce erythropoietin." J Histochem Cytochem **41**(3): 335-341.
- Bailon, P., A. Palleroni, C. A. Schaffer, C. L. Spence, W. J. Fung, J. E. Porter, G. K. Ehrlich, W. Pan, Z. X. Xu, M. W. Modi, A. Farid, W. Berthold and M. Graves (2001). "Rational design of a potent, long-lasting form of interferon: a 40 kDa branched polyethylene glycol-conjugated interferon alpha-2a for the treatment of hepatitis C." Bioconjug Chem **12**(2): 195-202.
- Baron, M. H. and S. T. Fraser (2005). "The specification of early hematopoiesis in the mammal." Curr Opin Hematol **12**(3): 217-221.
- Blight, S. K., R. C. Larue, A. Mahapatra, D. G. Longstaff, E. Chang, G. Zhao, P. T. Kang, K. B. Green-Church, M. K. Chan and J. A. Krzycki (2004). "Direct charging of tRNA(CUA) with pyrrolysine in vitro and in vivo." Nature **431**(7006): 333-335.
- Budisa, N. (2004). "Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire." Angew Chem Int Ed Engl **43**(47): 6426-6463.
- Budisa, N. and H. Biava (2014). "Evolution of fluorinated enzymes: An emerging trend for biocatalyst stabilization." Eng. Life Sci. **14**: 340-351.
- Cazzola, M., F. Mercuriali and C. Brugnara (1997). "Use of recombinant human erythropoietin outside the setting of uremia." Blood **89**(12): 4248-4267.
- Chan, T. R., R. Hilgraf, K. B. Sharpless and V. V. Fokin (2004). "Polytriazoles as copper(I)-stabilizing ligands in catalysis." Org Lett **6**(17): 2853-2855.
- Cheetham, J. C., D. M. Smith, K. H. Aoki, J. L. Stevenson, T. J. Hoeffel, R. S. Syed, J. Egrie and T. S. Harvey (1998). "NMR structure of human erythropoietin and a comparison with its receptor bound conformation." Nat Struct Biol **5**(10): 861-866.
- Chen, S. Y., S. Cressman, F. Mao, H. Shao, D. W. Low, H. S. Beilan, E. N. Cagle, M. Carnevali, V. Gueriguian, P. J. Keogh, H. Porter, S. M. Stratton, M. C. Wiedeke, L. Savatski, J. W. Adamson, C. E. Bozzini, A. Kung, S. B. Kent, J. A. Bradburne and G. G. Kochendoerfer (2005). "Synthetic erythropoietic proteins: tuning biological performance by site-specific polymer attachment." Chem Biol **12**(3): 371-383.
- Cowie, D. B. and G. N. Cohen (1957). "Biosynthesis by Escherichia coli of active altered proteins containing selenium instead of sulfur." Biochim Biophys Acta **26**(2): 252-261.
- Crick, F. (1970). "Central dogma of molecular biology." Nature **227**(5258): 561-563.
- Davis, F. F. (2002). "The origin of pegnology." Adv Drug Deliv Rev **54**(4): 457-458.
- Davis, J. M., T. Arakawa, T. W. Strickland and D. A. Yphantis (1987). "Characterization of recombinant human erythropoietin produced in Chinese hamster ovary cells." Biochemistry **26**(9): 2633-2638.
- Debili, N., L. Coulombel, L. Croisille, A. Katz, J. Guichard, J. Breton-Gorius and W. Vainchenker (1996). "Characterization of a bipotent erythro-megakaryocytic progenitor in human bone marrow." Blood **88**(4): 1284-1296.
- Delorme, E., T. Lorenzini, J. Giffin, F. Martin, F. Jacobsen, T. Boone and S. Elliott (1992). "Role of glycosylation on the secretion and biological activity of erythropoietin." Biochemistry **31**(41): 9871-9876.

- Dower, W. J., J. F. Miller and C. W. Ragsdale (1988). "High efficiency transformation of E. coli by high voltage electroporation." Nucleic Acids Res **16**(13): 6127-6145.
- ECACC (2010). Fundamental Techniques in Cell Culture - Laboratory Handbook.
- Egrie, J. C. and J. K. Browne (2001). "Development and characterization of novel erythropoiesis stimulating protein (NESP)." Nephrol Dial Transplant **16 Suppl 3**: 3-13.
- Erslev, A. (1953). "Humoral regulation of red cell production." Blood **8**(4): 349-357.
- Fekner, T. and M. K. Chan (2011). "The pyrrolysine translational machinery as a genetic-code expansion tool." Curr Opin Chem Biol **15**(3): 387-391.
- Foote, M. A. (2009). Studies of erythropoiesis and the discovery and cloning of recombinant human erythropoietin. Erythropoietins, Erythropoietic Factors and Erythropoiesis. S. G. Elliott, M. A. Foote and G. Molineux. Switzerland, Birkhäuser Verlag.
- Gallicchio, V. S. and M. J. Murphy, Jr. (1979). "In vitro erythropoiesis. II. Cytochemical enumeration of erythroid stem cells (CFU-e and BFU-e) from normal mouse and human hematopoietic tissues." Exp Hematol **7**(5): 219-224.
- Gifford, S. C., J. Derganc, S. S. Shevkoplyas, T. Yoshida and M. W. Bitensky (2006). "A detailed study of time-dependent changes in human red blood cells: from reticulocyte maturation to erythrocyte senescence." Br J Haematol **135**(3): 395-404.
- Goldwasser, E. (1984). "Erythropoietin and its mode of action." Blood Cells **10**(2-3): 147-162.
- Graham, M. L. (2003). "Pegaspargase: a review of clinical studies." Adv Drug Deliv Rev **55**(10): 1293-1302.
- Hackenberger, C. P. and D. Schwarzer (2008). "Chemoselective ligation and modification strategies for peptides and proteins." Angew Chem Int Ed Engl **47**(52): 10030-10074.
- Hamann, M. (2010). Engineering recombinant human erythropoietin (rhEPO) in E. coli with non-natural amino acids. Master Thesis, University of Konstanz.
- Hammerling, U., R. Kroon, T. Wilhelmsen and L. Sjodin (1996). "In vitro bioassay for human erythropoietin based on proliferative stimulation of an erythroid cell line and analysis of carbohydrate-dependent microheterogeneity." J Pharm Biomed Anal **14**(11): 1455-1469.
- Hao, B., W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki and M. K. Chan (2002). "A new UAG-encoded residue in the structure of a methanogen methyltransferase." Science **296**(5572): 1462-1466.
- Heath, D. S., A. A. Axelrad, D. L. McLeod and M. M. Shreeve (1976). "Separation of the erythropoietin-responsive progenitors BFU-E and CFU-E in mouse bone marrow by unit gravity sedimentation." Blood **47**(5): 777-792.
- Hein, J. E. and V. V. Fokin (2010). "Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides." Chem Soc Rev **39**(4): 1302-1315.
- Helenius, A. and M. Aebi (2004). "Roles of N-linked glycans in the endoplasmic reticulum." Annu Rev Biochem **73**: 1019-1049.
- Herring, S., A. Ambrogelly, C. R. Polycarpo and D. Soll (2007). "Recognition of pyrrolysine tRNA by the Desulfitobacterium hafniense pyrrolysyl-tRNA synthetase." Nucleic Acids Res **35**(4): 1270-1278.
- Himo, F., T. Lovell, R. Hilgraf, V. V. Rostovtsev, L. Noodleman, K. B. Sharpless and V. V. Fokin (2005). "Copper(I)-catalyzed synthesis of azoles. DFT study predicts unprecedented reactivity and intermediates." J Am Chem Soc **127**(1): 210-216.
- Hirano, K., D. Macmillan, K. Tezuka, T. Tsuji and Y. Kajihara (2009). "Design and synthesis of a homogeneous erythropoietin analogue with two human complex-type sialyloligosaccharides: combined use of chemical and bacterial protein expression methods." Angew Chem Int Ed Engl **48**(50): 9557-9560.
- Hong, V., S. I. Presolski, C. Ma and M. G. Finn (2009). "Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation." Angew Chem Int Ed Engl **48**(52): 9879-9883.

- Hsieh, P. C. and R. Vaisvila (2013). "Protein engineering: single or multiple site-directed mutagenesis." Methods Mol Biol **978**: 173-186.
- Huisgen, R. (1984). 1,3-Dipolar Cycloaddition Chemistry. New York, Wiley.
- Jacobs, K., C. Shoemaker, R. Rudersdorf, S. D. Neill, R. J. Kaufman, A. Mufson, J. Seehra, S. S. Jones, R. Hewick, E. F. Fritsch and et al. (1985). "Isolation and characterization of genomic and cDNA clones of human erythropoietin." Nature **313**(6005): 806-810.
- Jacobson, L. O., E. Goldwasser, W. Fried and L. Plzak (1957). "Role of the kidney in erythropoiesis." Nature **179**(4560): 633-634.
- Jelkmann, W. (2004). "Molecular biology of erythropoietin." Intern Med **43**(8): 649-659.
- Jelkmann, W. (2007). "Recombinant EPO production--points the nephrologist should know." Nephrol Dial Transplant **22**(10): 2749-2753.
- Jevsevar, S., M. Kunstelj and V. G. Porekar (2010). "PEGylation of therapeutic proteins." Biotechnol J **5**(1): 113-128.
- Kajihara, Y., N. Yamamoto, R. Okamoto, K. Hirano and T. Murase (2010). "Chemical synthesis of homogeneous glycopeptides and glycoproteins." Chem Rec **10**(2): 80-100.
- Kelly, S. M., T. J. Jess and N. C. Price (2005). "How to study proteins by circular dichroism." Biochim Biophys Acta **1751**(2): 119-139.
- Kelly, S. M. and N. C. Price (2000). "The use of circular dichroism in the investigation of protein structure and function." Curr Protein Pept Sci **1**(4): 349-384.
- Kiick, K. L., E. Saxon, D. A. Tirrell and C. R. Bertozzi (2002). "Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation." Proc Natl Acad Sci U S A **99**(1): 19-24.
- Kitamura, T., T. Tange, T. Terasawa, S. Chiba, T. Kuwaki, K. Miyagawa, Y. F. Piao, K. Miyazono, A. Urabe and F. Takaku (1989). "Establishment and characterization of a unique human cell line that proliferates dependently on GM-CSF, IL-3, or erythropoietin." J Cell Physiol **140**(2): 323-334.
- Klingmuller, U., U. Lorenz, L. C. Cantley, B. G. Neel and H. F. Lodish (1995). "Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals." Cell **80**(5): 729-738.
- Knippers, R. (2006). Molekulare Genetik. Stuttgart, Georg Thieme Verlag.
- Kochendoerfer, G. G., S. Y. Chen, F. Mao, S. Cressman, S. Traviglia, H. Shao, C. L. Hunter, D. W. Low, E. N. Cagle, M. Carnevali, V. Gueriguian, P. J. Keogh, H. Porter, S. M. Stratton, M. C. Wiedeke, J. Wilken, J. Tang, J. J. Levy, L. P. Miranda, M. M. Crnogorac, S. Kalbag, P. Botti, J. Schindler-Horvat, L. Savatski, J. W. Adamson, A. Kung, S. B. Kent and J. A. Bradburne (2003). "Design and chemical synthesis of a homogeneous polymer-modified erythropoiesis protein." Science **299**(5608): 884-887.
- Kolb, H. C., M. G. Finn and K. B. Sharpless (2001). "Click Chemistry: Diverse Chemical Function from a Few Good Reactions." Angew Chem Int Ed Engl **40**(11): 2004-2021.
- Komatsu, N., H. Nakauchi, A. Miwa, T. Ishihara, M. Eguchi, M. Moroi, M. Okada, Y. Sato, H. Wada, Y. Yawata and et al. (1991). "Establishment and characterization of a human leukemic cell line with megakaryocytic features: dependency on granulocyte-macrophage colony-stimulating factor, interleukin 3, or erythropoietin for growth and survival." Cancer Res **51**(1): 341-348.
- Komatsu, N., M. Yamamoto, H. Fujita, A. Miwa, K. Hatake, T. Endo, H. Okano, T. Katsube, Y. Fukumaki, S. Sassa and et al. (1993). "Establishment and characterization of an erythropoietin-dependent subline, UT-7/Epo, derived from human leukemia cell line, UT-7." Blood **82**(2): 456-464.
- Koury, M. J. and M. C. Bondurant (1990). "Erythropoietin retards DNA breakdown and prevents programmed death in erythroid progenitor cells." Science **248**(4953): 378-381.

- Kruisbeek, A. M. (2001). "Isolation of mouse mononuclear cells." Curr Protoc Immunol **Chapter 3**: Unit 3 1.
- Lacombe, C., J. L. Da Silva, P. Bruneval, J. G. Fournier, F. Wendling, N. Casadevall, J. P. Camilleri, J. Bariety, B. Varet and P. Tambourin (1988). "Peritubular cells are the site of erythropoietin synthesis in the murine hypoxic kidney." J Clin Invest **81**(2): 620-623.
- Lawrence, S. (2006). "Biotech blockbusters consolidate markets." Nat Biotechnol **24**(12): 1466.
- Levy, Y., M. S. Hershfield, C. Fernandez-Mejia, S. H. Polmar, D. Scudiery, M. Berger and R. U. Sorensen (1988). "Adenosine deaminase deficiency with late onset of recurrent infections: response to treatment with polyethylene glycol-modified adenosine deaminase." J Pediatr **113**(2): 312-317.
- Li, Z., D. Yuan, X. Fan, B. H. Tan and C. He (2015). "Poly(ethylene glycol) conjugated poly(lactide)-based polyelectrolytes: synthesis and formation of stable self-assemblies induced by stereocomplexation." Langmuir **31**(8): 2321-2333.
- Lin, F. K., S. Suggs, C. H. Lin, J. K. Browne, R. Smalling, J. C. Egrie, K. K. Chen, G. M. Fox, F. Martin, Z. Stabinsky and et al. (1985). "Cloning and expression of the human erythropoietin gene." Proc Natl Acad Sci U S A **82**(22): 7580-7584.
- Liu, S., B. L. Pentelute and S. B. Kent (2012). "Convergent chemical synthesis of [lysine(24,38,83)] human erythropoietin." Angew Chem Int Ed Engl **51**(4): 993-999.
- Lizak, C., S. Gerber, S. Numao, M. Aebi and K. P. Locher (2011). "X-ray structure of a bacterial oligosaccharyltransferase." Nature **474**(7351): 350-355.
- Macdougall, I. C. (2005). "CERA (Continuous Erythropoietin Receptor Activator): a new erythropoiesis-stimulating agent for the treatment of anemia." Curr Hematol Rep **4**(6): 436-440.
- Manwani, D. and J. J. Bieker (2008). "The erythroblastic island." Curr Top Dev Biol **82**: 23-53.
- Matthaei, J. H. and M. W. Nirenberg (1961). "Characteristics and stabilization of DNAase-sensitive protein synthesis in E. coli extracts." Proc Natl Acad Sci U S A **47**: 1580-1588.
- McGrath, K. and J. Palis (2008). "Ontogeny of erythropoiesis in the mammalian embryo." Curr Top Dev Biol **82**: 1-22.
- Meldal, M. and C. W. Tornoe (2008). "Cu-catalyzed azide-alkyne cycloaddition." Chem Rev **108**(8): 2952-3015.
- Migliaccio, A. R. and G. Migliaccio (1988). "Human embryonic hemopoiesis: control mechanisms underlying progenitor differentiation in vitro." Dev Biol **125**(1): 127-134.
- Miyake, T., C. K. Kung and E. Goldwasser (1977). "Purification of human erythropoietin." J Biol Chem **252**(15): 5558-5564.
- Narhi, L. O., T. Arakawa, K. Aoki, J. Wen, S. Elliott, T. Boone and J. Cheetham (2001). "Asn to Lys mutations at three sites which are N-glycosylated in the mammalian protein decrease the aggregation of Escherichia coli-derived erythropoietin." Protein Eng **14**(2): 135-140.
- Narhi, L. O., T. Arakawa, K. H. Aoki, R. Elmore, M. F. Rohde, T. Boone and T. W. Strickland (1991). "The effect of carbohydrate on the structure and stability of erythropoietin." J Biol Chem **266**(34): 23022-23026.
- Nett, J. H., S. Gomathinayagam, S. R. Hamilton, B. Gong, R. C. Davidson, M. Du, D. Hopkins, T. Mitchell, M. R. Mallem, A. Nylen, S. S. Shaikh, N. Sharkey, G. C. Barnard, V. Copeland, L. Liu, R. Evers, Y. Li, P. M. Gray, R. B. Lingham, D. Visco, G. Forrest, J. DeMartino, T. Linden, T. I. Potgieter, S. Wildt, T. A. Stadheim, M. d'Anjou, H. Li and N. Sethuraman (2012). "Optimization of erythropoietin production with controlled glycosylation-PEGylated erythropoietin produced in glycoengineered Pichia pastoris." J Biotechnol **157**(1): 198-206.
- Nguyen, D. P., H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters and J. W. Chin (2009). "Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a

pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry." J Am Chem Soc **131**(25): 8720-8721.
- Nucci, M. L., R. Shorr and A. Abuchowski (1991). "The therapeutic value of poly(ethylene glycol)-modified proteins." Adv. Drug Deliv. Rev. **6**: 133-151.
- Pasut, G., A. Guiotto and F. M. Veronese (2004). "Protein, peptide and non-peptide drug PEGylation for therapeutic application." Expert Opinion on Therapeutic Patents **14**(6): 859-894.
- Payne, R. J. and C. H. Wong (2010). "Advances in chemical ligation strategies for the synthesis of glycopeptides and glycoproteins." Chem Commun (Camb) **46**(1): 21-43.
- Polycarpo, C., A. Ambrogelly, A. Berube, S. M. Winbush, J. A. McCloskey, P. F. Crain, J. L. Wood and D. Soll (2004). "An aminoacyl-tRNA synthetase that specifically activates pyrrolysine." Proc Natl Acad Sci U S A **101**(34): 12450-12454.
- Quentmeier, H., M. Zaborski and H. G. Drexler (1997). "The human bladder carcinoma cell line 5637 constitutively secretes functional cytokines." Leuk Res **21**(4): 343-350.
- R&Dsystems®. (2015). "The Mouse/Rat Colony Forming Cell (CFC) Assay using Methylcellulose-based Media." from http://www.rndsystems.com/literature_MouseMethylcellulose.aspx.
- Rajender Reddy, K., M. W. Modi and S. Pedder (2002). "Use of peginterferon alfa-2a (40 KD) (Pegasys) for the treatment of hepatitis C." Adv Drug Deliv Rev **54**(4): 571-586.
- Reichel, C. (2012). "SARCOSYL-PAGE: a new electrophoretic method for the separation and immunological detection of PEGylated proteins." Methods Mol Biol **869**: 65-79.
- Remy, I., I. A. Wilson and S. W. Michnick (1999). "Erythropoietin receptor activation by a ligand-induced conformation change." Science **283**(5404): 990-993.
- Richmond, T. D., M. Chohan and D. L. Barber (2005). "Turning cells red: signal transduction mediated by erythropoietin." Trends Cell Biol **15**(3): 146-155.
- Rostovtsev, V. V., L. G. Green, V. V. Fokin and K. B. Sharpless (2002). "A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes." Angew Chem Int Ed Engl **41**(14): 2596-2599.
- Rozwarski, D. A., A. M. Gronenborn, G. M. Clore, J. F. Bazan, A. Bohm, A. Wlodawer, M. Hatada and P. A. Karplus (1994). "Structural comparisons among the short-chain helical cytokines." Structure **2**(3): 159-173.
- Sambrook, J. and D. W. Russell (2001). Molecular Cloning - A laboratory manual, Cold Spring Harbor Laboratory Press.
- Sasaki, H., B. Bothner, A. Dell and M. Fukuda (1987). "Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA." J Biol Chem **262**(25): 12059-12076.
- Sasaki, H., N. Ochi, A. Dell and M. Fukuda (1988). "Site-specific glycosylation of human recombinant erythropoietin: analysis of glycopeptides or peptides at each glycosylation site by fast atom bombardment mass spectrometry." Biochemistry **27**(23): 8618-8626.
- Saxena, V. P. and D. B. Wetlaufer (1970). "Formation of three-dimensional structure in proteins. I. Rapid nonenzymic reactivation of reduced lysozyme." Biochemistry **9**(25): 5015-5023.
- Schneider, D., T. Schneider, D. Rosner, M. Scheffner and A. Marx (2013). "Improving bioorthogonal protein ubiquitylation by click reaction." Bioorg Med Chem **21**(12): 3430-3435.
- Semenza, G. L. (2001). "HIF-1, O(2), and the 3 PHDs: how animal cells signal hypoxia to the nucleus." Cell **107**(1): 1-3.
- Shine, J. and L. Dalgarno (1974). "The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites." Proc Natl Acad Sci U S A **71**(4): 1342-1346.

- Sowade, B., O. Sowade, J. Mocks, W. Franke and H. Warnke (1998). "The safety of treatment with recombinant human erythropoietin in clinical use: a review of controlled studies." Int J Mol Med **1**(2): 303-314.
- Srinivasan, G., C. M. James and J. A. Krzycki (2002). "Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA." Science **296**(5572): 1459-1462.
- Stephenson, J. R., A. A. Axelrad, D. L. McLeod and M. M. Shreeve (1971). "Induction of colonies of hemoglobin-synthesizing cells by erythropoietin in vitro." Proc Natl Acad Sci U S A **68**(7): 1542-1546.
- Syed, R. S., S. W. Reid, C. Li, J. C. Cheetham, K. H. Aoki, B. Liu, H. Zhan, T. D. Osslund, A. J. Chirino, J. Zhang, J. Finer-Moore, S. Elliott, K. Sitney, B. A. Katz, D. J. Matthews, J. J. Wendoloski, J. Egrie and R. M. Stroud (1998). "Efficiency of signalling through cytokine receptors depends critically on receptor orientation." Nature **395**(6701): 511-516.
- Takeuchi, M., S. Takasaki, H. Miyazaki, T. Kato, S. Hoshi, N. Kochibe and A. Kobata (1988). "Comparative study of the asparagine-linked sugar chains of human erythropoietins purified from urine and the culture medium of recombinant Chinese hamster ovary cells." J Biol Chem **263**(8): 3657-3663.
- Thobhani, S., C. T. Yuen, M. J. Bailey and C. Jones (2009). "Identification and quantification of N-linked oligosaccharides released from glycoproteins: an inter-laboratory study." Glycobiology **19**(3): 201-211.
- Torbett, B. E. and J. S. Friedman (2009). Erythropoiesis: an overview. Erythropoitins, Erythropoietic Factors and Erythropoiesis. M. A. F. S. G. Elliott, G. Molineux, Birkhäuser Verlag/Switzerland.
- Tornoe, C. W., C. Christensen and M. Meldal (2002). "Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(i)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides." J Org Chem **67**(9): 3057-3064.
- Unverzagt, C. and Y. Kajihara (2013). "Chemical assembly of N-glycoproteins: a refined toolbox to address a ubiquitous posttranslational modification." Chem Soc Rev **42**(10): 4408-4420.
- Verdier, F., P. Walrafen, N. Hubert, S. Chretien, S. Gisselbrecht, C. Lacombe and P. Mayeux (2000). "Proteasomes regulate the duration of erythropoietin receptor activation by controlling down-regulation of cell surface receptors." J Biol Chem **275**(24): 18375-18381.
- Wang, G. L. and G. L. Semenza (1993). "Characterization of hypoxia-inducible factor 1 and regulation of DNA binding activity by hypoxia." J Biol Chem **268**(29): 21513-21518.
- Wang, G. L. and G. L. Semenza (1993). "Desferrioxamine induces erythropoietin gene expression and hypoxia-inducible factor 1 DNA-binding activity: implications for models of hypoxia signal transduction." Blood **82**(12): 3610-3615.
- Wang, P., S. Dong, J. A. Brailsford, K. Iyer, S. D. Townsend, Q. Zhang, R. C. Hendrickson, J. Shieh, M. A. Moore and S. J. Danishefsky (2012). "At last: erythropoietin as a single glycoform." Angew Chem Int Ed Engl **51**(46): 11576-11584.
- Wang, Q., T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless and M. G. Finn (2003). "Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition." J Am Chem Soc **125**(11): 3192-3193.
- Wang, Y. J., Y. D. Liu, J. Chen, S. J. Hao, T. Hu, G. H. Ma and Z. G. Su (2010). "Efficient preparation and PEGylation of recombinant human non-glycosylated erythropoietin expressed as inclusion body in E. coli." Int J Pharm **386**(1-2): 156-164.
- Wang, Y. S., S. Youngster, M. Grace, J. Bausch, R. Bordens and D. F. Wyss (2002). "Structural and biological characterization of pegylated recombinant interferon alpha-2b and its therapeutic implications." Adv Drug Deliv Rev **54**(4): 547-570.
- Watowich, S. S. (2011). "The erythropoietin receptor: molecular structure and hematopoietic signaling pathways." J Investig Med **59**(7): 1067-1072.

- Wells, L., K. Vosseller and G. W. Hart (2001). "Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc." Science **291**(5512): 2376-2378.
- WHO (2007). International nonproprietary names (INN) for biological and biotechnological substances. I. W. d. 05.179.
- Witthuhn, B. A., F. W. Quelle, O. Silvennoinen, T. Yi, B. Tang, O. Miura and J. N. Ihle (1993). "JAK2 associates with the erythropoietin receptor and is tyrosine phosphorylated and activated following stimulation with erythropoietin." Cell **74**(2): 227-236.
- Wognum, B., N. Yuan, B. Lai and C. L. Miller (2013). "Colony forming cell assays for human hematopoietic progenitor cells." Methods Mol Biol **946**: 267-283.
- Wycuff, D. R. and K. S. Matthews (2000). "Generation of an AraC-araBAD promoter-regulated T7 expression system." Anal Biochem **277**(1): 67-73.
- Youssoufian, H., G. Longmore, D. Neumann, A. Yoshimura and H. F. Lodish (1993). "Structure, function, and activation of the erythropoietin receptor." Blood **81**(9): 2223-2236.
- Yuan, Y., J. Chen, Q. Wan, R. M. Wilson and S. J. Danishefsky (2010). "Toward fully synthetic, homogeneous glycoproteins: advances in chemical ligation." Biopolymers **94**(4): 373-384.
- Zhang, X., R. Goncalves and D. M. Mosser (2008). "The isolation and characterization of murine macrophages." Curr Protoc Immunol **Chapter 14**: Unit 14 11.