The advantages and disadvantages of bioorthogonal proteins
Groenewold, G.J.M.

Citation
Groenewold, G. J. M. (2021, February 17). *The advantages and disadvantages of bioorthogonal proteins*. Retrieved from https://hdl.handle.net/1887/3142384

Cover Page

Universiteit Leiden

Leiden University
Repository

The handle https://hdl.handle.net/1887/3142384 holds various files of this Leiden University dissertation.

Author: Groenewold, G.J.M.
Title: The advantages and disadvantages of bioorthogonal proteins
Issue Date: 2021-02-17

The advantages and disadvantages of bioorthogonal proteins

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus Prof. dr. ir. H. Bijl,

volgens besluit van het College voor Promoties

te verdedigen op woensdag 17 februari 2021

klokke 15:00 uur

door

Gerdientje Johanna Mirjam Groenewold

Geboren te Harderwijk in 1990

Promotiecommissie

Promotor: Dr. S.I. van Kasteren

Prof. dr. H.S. Overkleeft

Overige leden: Prof. dr. J.M.F.G. Aerts (voorzitter)

Prof. dr. M. van der Stelt (secretaris)

Dr. A.L. Boyle

Prof. dr. G.P. van Wezel

Prof. dr. G. van den Bogaart   *Rijksuniversiteit Groningen*

Prof. dr. D.J. Scheffers   *Rijksuniversiteit Groningen*

Table of contents

Chapter 1

Introduction

Abstract

It is possible nowadays to recombinantly express most proteins. However, these proteins will never be fully identical to those found in nature, due to differences in post-translational modifications. In order to mimic these modifications, various synthetic and semi-synthetic approaches have been developed, such as the *de novo* synthesis of proteins, their enzymatic modification or residue-specific chemical modification. An overview of these approaches will be given in this Chapter.

Introduction

Proteins play an important role in nearly all biological processes, such as catalysis [1], transport [2, 3], cell-cell interactions [4, 5] and our immune system [6, 7]. In early days, the study of a protein of interest required its isolation from its natural source. With the exception of highly abundant proteins, isolation of native proteins requires large – sometimes kilogram – quantities of tissue (plants, animals) and even then, often only small quantities of the desired protein are obtained [8]. Today it is possible – for most proteins – to either prepare a synthetic version [9] or to express and purify the protein of interest recombinantly [10]. Both approaches allow mutation of a protein of interest, which in turn allows the study of disease-associated mutations, but also the introduction of non-canonical amino acids including ones bearing chemical functionalities not encountered in nature [11]. This Chapter discusses current technologies for recombinant protein expression. Following a basic introduction into protein expression in general (section **1.1-1.3**), the main focus will be on the expression of proteins containing a non-canonical, non-natural functionality.

1.1 Basic principles of protein expression

Proteins are expressed from its corresponding coding sequence, a process which involves two major steps: transcription and translation [12]. During transcription the information encoded in deoxyribonucleic acid (DNA) is transcribed into pre-messenger ribonucleic acid (pre-mRNA) which is next processed by RNA polymerase II to form mature mRNA [13]. This mRNA serves as a template for protein synthesis with stretches of three specific bases encoding the twenty natural (also called canonical) amino acids as well as three signals to halt protein synthesis (stop codons) [14]. This process, called translation, occurs in the ribosomal compartment [15]. The ribosome consists of a small and a large subunit, of which the small subunit exhibits three different binding sites oriented from 5' to 3': the exit site (E), the peptidyl site (P) and the aminoacyl site (A) (Figure 1). When the mRNA is located inside the ribosome, the A-site binds the aminoacyl-transfer RNA (aa-tRNA, a complex which will be discussed in more detail later in this Chapter) bearing the complementary codon on the mRNA. The P-site holds the complex with the growing polypeptide chain and lastly, the E-site holds the tRNA without its amino acid, before it is released [16]. This process is terminated when a stop
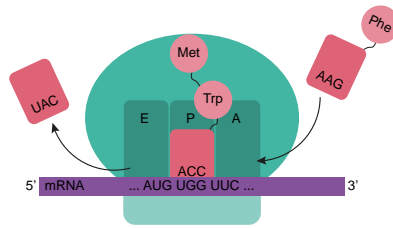
Figure 1. Stages of translation. The aa-tRNA binds to the A-site, hereafter the polypeptide chain will grow in the P-site. The empty aa-tRNA will leave via the E-site.
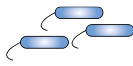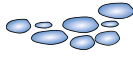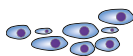
codon enters the A-site. Since there are no native aa-tRNA that recognize stop codons a termination factor will bind, releasing the polypeptide chain [17, 18]. Hereafter, the protein will either spontaneously or with the help of chaperones fold in its native state [19].

1.2 Protein expression hosts

Protein translation can be exploited to recombinantly produce proteins in a variety of expression systems, either prokaryotic or eukaryotic. The advantage of the former is the ease of use and scalability. The advantage of the latter is that proteins that have undergone mammalian post-translational modifications, in particular glycosylation, can be obtained. The first step in the production of recombinant proteins entails cloning of the underlying recombinant DNA, that codes for the protein of interest, into an expression host that supports gene expression [20]. Protein expression is most commonly conducted in bacteria, yeasts, insect cells or mammalian cells [21], but proteins can also be expressed in algae [22] or in cell free cultures [23]. Each of these expression systems have their advantages and disadvantages. The four most used expression systems, with their advantages and disadvantages, are discussed in some more detail below (Table 1).

Bacterial expression systems are commonly used because high expression levels can often be reached at low cost, thus allowing the production of large quantities of protein [24]. However, post-translational modifications (PTMs) such as may occur on mammalian proteins are not part of bacterial protein synthesis machineries (although some strains have been engineered to perform *N*-glycosylation [25]). Other features that complicate bacterial expression are the presence of endotoxins, such as lipopolysaccharides (LPS), as impurities can affect the immunological properties of a

Table 1: The advantages and disadvantages of different protein expression systems.

Expression systems	Advantages	Challenges	Most common applications
Bacterial	• Low cost • Rapid expression • Scalable • Most widely used expression system	• Inclusion bodies • No eukaryotic PTMs • Difficult to express higher MW proteins • Protein specific optimization	• Functional assays • Structural analysis • Bacterial proteins • Antigen proteins • Cytokines • Enzymes
Yeast	• Low cost • Rapid expression • High yield • Scalable • Diverse PTMs	• Improper glycosylation • High mannose modification • Growth optimization • Fermentation required for high yields	• Functional assays • Structural analysis • Antibody production • Lower MW proteins • Cytokines • Enzymes
Baculovirus - insect	• High capacity genes • Soluble proteins • Suitable for toxic proteins • PTMs similar to mammalian expression • Static or suspension cultures	• More demanding culture conditions • Lack of partial glycosylation • Time consuming	• Functional assays • Structural analysis • Cytoplasmic proteins • Toxic proteins • Transmembrane proteins (TP) • Protein complexes
Mammalian	• Soluble proteins • Lower endotoxin • Better bioactivity • PTMs • Transient and stable expression	• More demanding culture conditions • High-yield only in suspension conditions	• Functional assays • Structural analysis • Recombinant antibodies • Antibody production • Virus production • Expression of difficult proteins

protein sample. Their removal prior to use requires additional purification steps, whereas the production in an endotoxin free environment is also an option [26]. The reducing environment of the cytoplasm in for instance *Escherichia coli* (*E. coli*) [27] can also limit the yields of expression of proteins that require disulfide bridges for stable folding, and can result in the expression in inclusion bodies, vesicles containing partially or incorrectly folded proteins [28 - 30]. One of the possibilities to avoid the protein of interest ending up in inclusion bodies is by expressing the protein in the periplasm [31], where the oxidizing environment and the presence of proteins like disulfide oxidoreductase (DsbA) allow disulfide bonds to be formed more easily [27, 32].

Yeast expression systems support both intra- and extracellular expression of prokaryotic and eukaryotic proteins [33, 34]. The systems, in contrast to most contemporary bacterial ones, offer the possibility for expression levels in a minimal expression time in simple media [35]. Yeast expression allows for the introduction of post-translational modifications, including both *N*- and *O*-linked oligosaccharides, however often hyperglycosylated proteins are obtained [36].

Protein expression in insect cells makes use of baculovirus infection to introduce the gene of interest [37]. The system can be used for large scale production and allows for PTMs [38]. Post-translational modifications are chemical modifications of a protein,
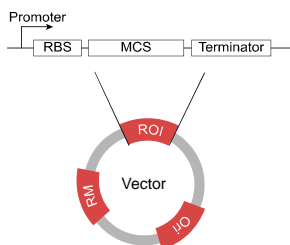
Figure 2. Minimal components of an expression vector. An expression vector, which at its region of interest (ROI), contains at least a promoter, a ribosomal binding site, a multiple cloning site and a corresponding terminator. Moreover, a vector contains an origin of replication (Ori) and a resistance marker (RM).

which play an important role in the human proteome, which will be described in more detail in section 1.4 [39]. The proteins will be, depending on the chosen system, either secreted intra- or extracellularly [40, 41]. On the downside, protein expression using this system is time-consuming at relatively high costs [37].

Despite the extensive developments in non-mammalian eukaryotic expression, most mammalian proteins, especially those that are endowed with post-translational modifications and/or those that rely on mammalian folding systems, are still best expressed in a mammalian expression system [42]. There are various ways to express a protein in mammalian cells: via plasmid transfection [43], via retrovirus or lentivirus transduction [44] or via modified mRNA transfection using both transient expression and stable cell cultures [45]. However, the use of mammalian cells comes with more challenging culture conditions, such as slow cell growth and high production costs by the use of expensive media and culture conditions [46]. High expression levels can only be obtained in suspension cultures, cultures in which cells are homogeneously suspended and grown in agitating liquid medium [47].

1.3 Recombinant protein expression

Once the expression host is selected, cloning the recombinant DNA encoding the protein of interest requires the selection and design of an appropriate expression vector, which should be compatible with the selected expression host.

For bacterial expression, two expression vector systems are commonly used, the pET vector system of Novagen/EMD Millipore [48, 49] and the pQE vector system of Qiagen [50]. The pET systems contain at least a T7 RNA polymerase promoter, a

ribosomal binding site (RBS) – the Shine-Dalgarno sequence –, a multiple cloning site (MCS) and a T7 terminator (Figure 2) [51]. When the gene is cloned into the MCS, the resulting plasmid is transformed into a lambda DE3 lysogen strain, containing the T7 RNA polymerase under a *lac* repressor, and subsequently protein expression is induced using isopropyl-β-D-1-thiogalactopyranoside (IPTG) [52]. The pQE vectors contain the same minimal components, with the exception that the gene will be cloned behind a T5 promoter [53]. This promoter will be recognized by the *E. coli* RNA polymerase and therefore does not require the T7 RNA polymerase [54]. Besides these two often-used vectors, other vectors are used as well which are inducible by other molecules [55-57].

The basic components, a promoter, a ribosomal binding site – called the Kozak sequence in mammalian cells – and a multiple cloning site, are the same for each of the other described expression systems in section 1.2. Every expression host requires a different promoter. For instance, in mammalian protein expression a cytomegalovirus (CMV) promoter can be used [42]; in insect cells a polyhedron promoter [58]; and in yeast a galactose inducible GAL1 promoter [59].

1.4 Post-translational modifications of proteins

During and after protein expression, the nascent protein can be subjected to multiple modifications. These, so called post-translational modifications (PTMs) extend the repertoire of chemical functionalities beyond those available from the 20 canonical amino acids [60]. The changes in protein properties imbued by PTMs – which can be both permanent and reversible – modulate and regulate cellular processes [61]. The change in properties the introduction of *N*- and *O*-linked glycans has, plays important roles in protein folding and stability [62], mucosal barrier function [63], cell-cell communication [64], and immune activation [65], to name but a few examples. Phosphorylation of serine/threonine/tyrosine (and other) residues, GlcNAcylation, and ubiquitination are reversible PTMs that play key roles in transducing signals within the cell [66]. They affect processes ranging from cell division and growth to apoptosis [67] by regulating intracellular signalling, transcription, protein routing, and degradation (Figure 3) [68, 69].
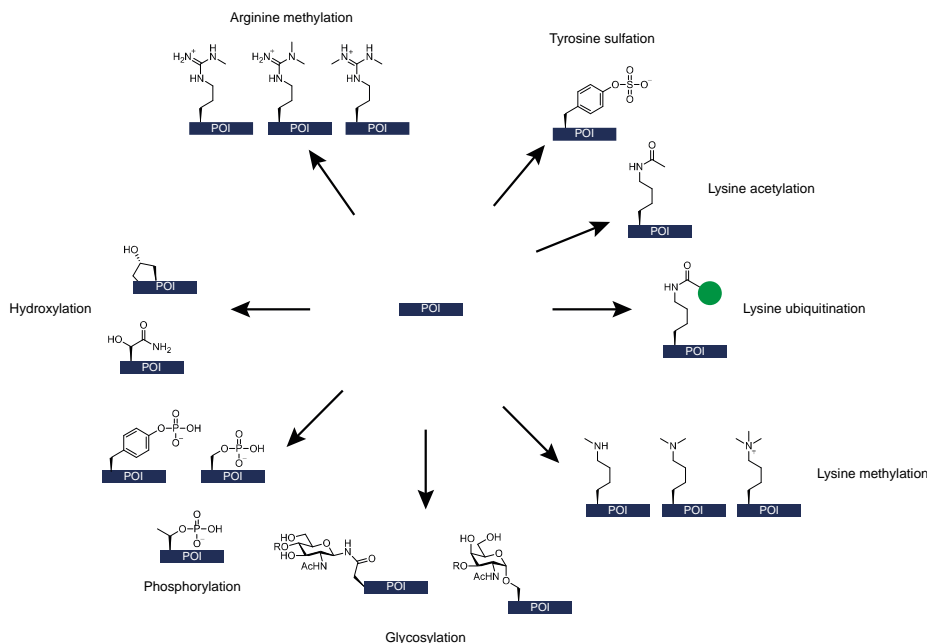
Figure 3. A few examples of post-translational modifications of proteins.

PTMs play a major role in a variety of human diseases [70], such as cancers [71], auto-immune diseases [72], heart diseases [73] and neurodegenerative diseases [74]. However, and unlike the amino acid backbone that is templated by the DNA of the original genetic code, PTMs are not encoded and therefore often heterogeneous, which makes the study of their precise function more difficult. This holds particularly true when multiple PTMs compete for the same site in the protein, such as for example serine and threonine residues that can be either phosphorylated or GlcNAcylated [75], – resulting in a heterogeneous protein mixture – or when, as is often the case, PTMs are reversible [76].

In recent years, extensive effort has gone into methods that allow the production of homogeneous PTM-modified proteins [77]. This has yielded three strategies for their production: chemical total synthesis of the PTM-modified protein [78]; site-selective enzymatic modification [79]; or the tag-and-modify approach, in which a small unique chemical functionality is introduced at a specific site in the protein and subsequently reacted with a PTM-reagent (or mimic) [80]. The latter approach has the advantages of its highly versatile and modular nature that gives full choice over the site of modification

[81]. One obvious caveat that the linkage of the PTM to backbone is non-native in most cases.

1.4.1 Synthesis of homogeneous N-linked glycosylated proteins through enzymatic reactions

The first two strategies to incorporate PTMs in proteins involve synthesizing the PTM-modified protein *de novo,* or via enzymatic reactions [82]. Both these methods have been largely used to produce proteins that carry glycans at specific sites. The synthesis of glycoproteins [83] can be performed combining solid phase peptide synthesis and native chemical ligation reactions [78], and has yielded some impressive examples such as synthetic single glycoforms of erythropoietin (EPO) [84] and the primary gp120 V1V2 glycopeptide of human immunodeficiency virus (HIV) [85].

The second approach to synthesize single glycoforms of proteins has been to engineer the existing mixtures of glycans on expressed proteins. In this approach, *N*-linked glycans (there is no universal protocol for site-specific *O*-linked glycan modification yet [86]) are hydrolyzed down to a single GlcNAc-asparagine. This minimal motif is then elaborated using engineered endoglycosidases that favor the reverse reaction, so called glycosynthases [87] (Scheme 1A). This method has been successfully used to synthesize single glycoforms of RNAse [88], but also of antibodies [89].

1.4.2. The "tag-and-modify" approach on native residues

The "tag-and-modify" method is the third method to produce glycoproteins and other PTMs, and compared to the above two methods is less cumbersome and more versatile. It is dependent on first introducing a reactive amino acid into the protein backbone at the prospective site of PTM, and the subsequent selective chemical modification of this tag [80]. This tag can be either a reactive natural amino acid (usually a cysteine [90]), or a noncanonical amino acid [81, 91].

1.4.3 Site-specific chemical ligations to canonical amino acids

Bioconjugation has by and large relied on the reaction of electrophiles with the nucleophilic amino acids cysteine (Cys) and lysine (Lys) for (site-specific) modifications [92]. The scarcity of free cysteines in proteins has made this amino acid particularly

attractive for tag-and-modify approaches based on natural residues as this increases the chance of a predefined protein. However, other amino acids have also been studied, but – like lysines – their abundance has limited site-specific modification [93-99].

Cysteines are often used for modification as they contain a strong soft nucleophile under physiological conditions, that offers unique chemical reactivity. In combination with its low abundance in proteins [100] it is one of the more suitable amino acids that has been used for site-selective modification of a protein [90, 101]. One of the most commonly used cysteine modification techniques comprises alkylation with iodoacetamide-based reagents [102], and this technique is commonly used for capping free thiols for proteomics experiments [103]. This strategy has for instance been used by the Flitsch-group to create synthetic glycoproteins by reacting engineered thiols with glycosyl-iodoacetamides [104]. However, this reaction is not fully selective for thiols, with cross-reactivity to lysine residues observed at basic pH required to drive the reaction to completion. Maleimides react with a cysteine thiol in a Michael addition [105, 106], and has been employed for instance to produce Pneumococcal surface adhesin A (PsaA) mutants linked to a tetrasaccharide in a study on glycoconjugate vaccine candidates [107]. Moreover, a maleimide-thiol ligation reaction was used to site-specifically glycosylate the HIV-1 gp41 peptide, which is a potent anti-HIV agent [108].

Combining cysteine modification with the "tag-and-modify" approach, Davis and coworkers developed a toolbox to modify native and non-canonical amino acids, which allowed them to methylate, glycosylate, and phosphorylate proteins [81]. With this approach a library of PTM-modified histones – histones are known to undergo acetylation, phosphorylation, and ADP-ribosylation – were produced [109]. Except for histone H3, none of the human histones encompasses a natural cysteine, which makes this an obvious amino acid to introduce in recombinant histone proteins, and then modify. Cysteine-mutant histones were for instance reacted with a dibromide reagent to form a dehydroalanine (Dha), which served as precursor for the construction of close analogues (one sulfur substituting for one carbon) of six PTMs [110] (Scheme 1B).

Lysines can be efficiently modified by acylation using cyanates and isocyanates [111, 112], sulfonylation using sulfonyl fluorides [113] and, via reductive amination, with aldehydes [114]. Although aldehydes can be used for selective modification, reduction

A

B

C
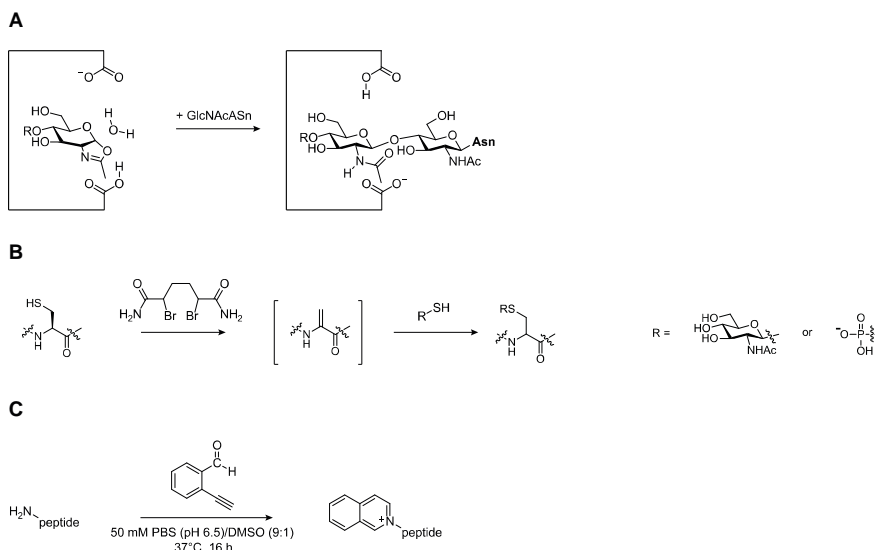
Scheme 1. Schematic overview of different chemical amino acid modifications. A) Glycosylation of an asparagine via the oxazoline intermediate [116]. B) The modification of a cysteine via the Dha intermediate [110]. C) The modification of the N-terminal amine with the use of 2-EBA [121].

of the transiently formed imines can be detrimental to the structural, and therefore, functional integrity of the underlying protein [115]. Lysine is a highly abundant amino acid, and as such of much more limited use for site-selective modification than cysteines are.

The exception to this rule is perhaps the N-terminal amine residue, which under natural circumstances can be, for instance, acylated, methylated and ubiquitylated at the N-terminus [117-120]. The difference in pKa of the N-terminus compared to Lys-amines (pKa of about 8 for α- and 10 for ε-amino acids) allows for its selective modification at lower pH values [92, 120]. Under mildly acidic conditions, the use of 2-ethynylbenzaldehydes (2-EBA) has allowed selective modification of the N-termini [121], but modifications of this terminus with PTMs has not been reported (Scheme 1C). Moreover, the N-terminal amine residue can be selectively modified using 1.75 eq of a diazotransfer reagent – imidazole-1-sulfonyl azide – at pH 8.5 [122, 123].

1.5 Bioorthogonal chemistry for site-specific PTM introduction

All the above-described procedures are based on protein expression with the 20 proteinogenic amino acids. The next sections will describe methods to introduce additional chemical functionality during protein expression using non-canonical amino acids with functional groups the reactivity of which is orthogonal to those found in canonical amino acids. Three methods have been reported to achieve this: cell free translation [124, 125], the use of heterologous aaRS/tRNA pairs (for instance, amber codon suppression), and the use of non-canonical amino acids that are isosteric to canonical ones [126].

1.5.1 Incorporation of non-canonical amino acids via aaRS/tRNA pairs

Non-canonical amino acids can be incorporated in proteins using heterologous aaRS/tRNA pairs that bind certain stop codons [127]. This allows, amongst others, for site-specific introduction of PTMs [128], the introduction of amino acids containing fluorophores [129], photocrosslinking moieties [130, 131] as well as ligation handles [132].

The technique is centered around tRNAs – isolated from extremophilic bacteria – that recognise the codons normally assigned to bind the release factors RF1 or RF2 [133]. The three stop codons used for this method are known as amber (UAG), ochre (UAA) and opal/umber (UGA) [134]. For the incorporation of non-canonical amino acids via a heterologous aaRS/tRNA pair, the amber stop codon is most often used [135] (Figure 4), because of its limited use inside bacterial cells (~9%) [136, 137] and human cells (~23%) [138]. Several conditions have to be met to allow expression of a protein containing a bioorthogonal functionality with this system: the tRNA-synthase should not recognise any canonical amino acid, but should recognise the non-canonical one used and load it on the correct tRNA. This tRNA then, once charged with the desired amino acid should recognise the stop codon exclusively [127, 139]. It should be noted that the thus created UAC-tRNA bearing the non-canonical amino acid will be in competition with the release factor(s) that normally bind the stop codon, with a potentially reduced protein yield accompanied by the expression of truncated proteins as a result [140].

Early experiments on the termination codons were performed using an in vitro translation system [141]. Using this system, the *E. coli* chloramphenicol acetyltransferase

(cat) gene was modified [142] and mutating the anticodon sequence of the initiator tRNA from CAU to CUA could be used to initiate protein expression [143]. This research eventually led to the first active aaRS/tRNA pair for the incorporation of noncanonical amino acids (ncAAs) in *E. coli*: the tyrosyl pair of *Methanocaldococcus jannaschii*, that allowed incorporation of *O*-methyl-L-tyrosine in dihydrofolate reductase [144, 145]. Since then other aaRS/tRNA pairs have been developed, including the pyrrolysyl pair from *Methanosarcina barkeri* [146] and *Methanosarcina mazei* [147], the lysyl [148] and glutamyl [149] pair of *Pyrococcus horikoshii*, and the tryptophanyl pair of *Saccharomyces cerevisiae* [150]. The aminoacylation of these pairs was first shown using Northern blotting [151] or in stop codon read-through assays [152]. Quite recently, a tRNA extension method (tREX) was developed, utilizing a rapid screen and discovery of orthogonal aminoacyl-tRNA synthetase-tRNA pairs [153, 154]. This method uses an online database containing tRNA sequences, which can identify candidate orthogonal tRNAs [155].

Today, genetic code expansion technologies have reached maturity, at least for the design of non-canonical proteins in bacterial expression systems. A wide variety of proteins incorporating a non-canonical amino acid has seen the light in recent years and for a variety of purposes. Amongst these proteins it was green fluorescent protein (GFP), which was expressed with the incorporation of for instance p-azidophenylalanine (pAzpa) and p-propargyloxyphenylalanine (pPpa) in CHO and human 293T cells, using the amber codon suppressor - *Bst*RNA$^{Tyr}_{CUA-}$ from *B. stearothermophilus* [138].

Now that incorporation of one single non-canonical amino acid can be accomplished through genetic code expansion with considerable confidence, focus today is moving towards the incorporation of several non-canonical amino acids into a single protein [139]. Huang *et al.* were the first to introduce three acetylated lysine residues in model protein GFP. To do so, they introduced three amber stop codons at predetermined sites. To get useful quantities of the target protein, they had to suppress release factor 1, which was achieved by preparing a knockout of the N-terminal domain of ribosomal protein L11 [155]. Following these studies Xiao *et al.* showed that incorporation of two different non-canonical amino acids, namely O-methyltyrosine and ε-*tert*-Boc-lysine, could be achieved by using two aaRS/tRNA pairs in conjunction

Figure 4. Schematic overview of the incorporation of unnatural amino acids into proteins using amber codon suppression. Left: The introduction of a natural amino acid. Right: The incorporation of an unnatural amino acid, using a modified tRNA synthetase.

with two different stop codons. These authors also showed that an antibody could be expressed with a p-acetylphenylalanine within the heavy chain and azidolysine within the light chain, using the amber (TAG) and ochre (TAA) stop codons respectively [156].

There are three natural stop codons that can be used to incorporate unnatural amino acids. This led the Chatterjee-group to expand genetic code expansion even further towards the incorporation of 5-hydroxytryptophan, p-azidophenylalanine and cyclopropene-lysine in GFP. With this, the limit of incorporation via manipulation of the three native stop codon usage was reached [157].

Genetic code expansion allows for the efficient incorporation of a variety of non-canonical amino acids in bacterial, yeast and mammalian cells [158, 159]. Protein yields proved a limiting factor in the early days [158, 160], but today proteins encompassing a single non-canonical amino acid can be obtained in good quantities: grams per liter

have been achieved in *E. coli* expression systems and 100s of milligrams in CHO cells [160].

1.5.2 Incorporation of non-canonical amino acids via metabolic labeling

The main alternative method to genetic code expansion for incorporating non-canonical amino acids into proteins is via metabolic labeling [127]. This method requires the activation of a non-canonical amino acid by a naturally occurring aminoacyl-tRNA synthetase that is in effect 'fooled' by the isosteric/isoelectronic nature of the new amino acid [161]. This method was first established with the incorporation of selenomethionine substituting for methionine to produce proteins suitable for structure determination by crystallography [162, 163]. Tirrell and coworkers then chose to use this methionine replacement method to attempt the incorporation of additional chemical functionalities into proteins to alter protein function. Even before the onset of bioorthogonal ligation chemistry they had reported the incorporation of amino acids, such as homoallylglycine [164], homopropargylglycine [165] and azidohomoalanine [166] into recombinant proteins for the use in conjugation reactions. The methodology is intrinsically more restricted in the diversity of amino acids that can be introduced in comparison with genetic code expansion: the non-canonical amino acid has to be structurally highly similar to the canonical one it is to substitute [167]. This canonical amino acid also should be absent, as otherwise it would compete with the unnatural analogue for incorporation [168], which dictates that auxotrophic strains are needed: ones that cannot produce the natural amino acid *de novo* [169]. Overcoming these limitations, Tirrell and coworkers revealed that proteins could be expressed with specific incorporation of non-canonical amino acids with a translation fidelity of >99% as shown by mass spectrometry [158].

Following the original report by Bertozzi and coworkers on the development of the Staudinger ligation as a bioorthogonal reaction [170], several groups employed Staudinger ligation chemistries for the modification of azide-containing proteins [171, 172]. Azide-containing proteins can also be modified using copper(I)-catalyzed azide alkyne [3+2] cycloaddition 'click' reactions [173, 174] or copper-free, strain-promoted varieties with for instance bicyclononyne (BCN) [175, 176]. Complementary chemistries
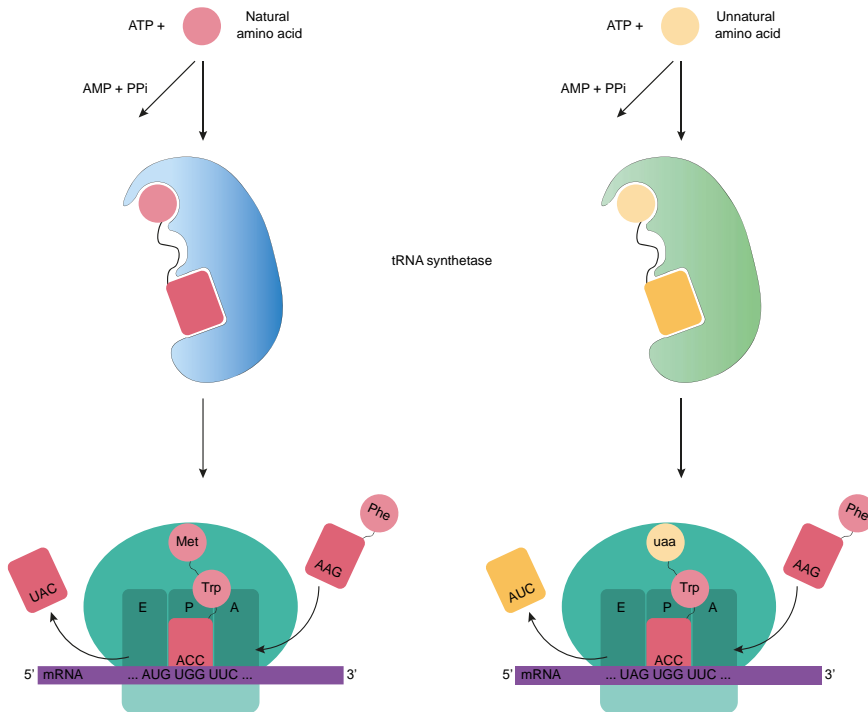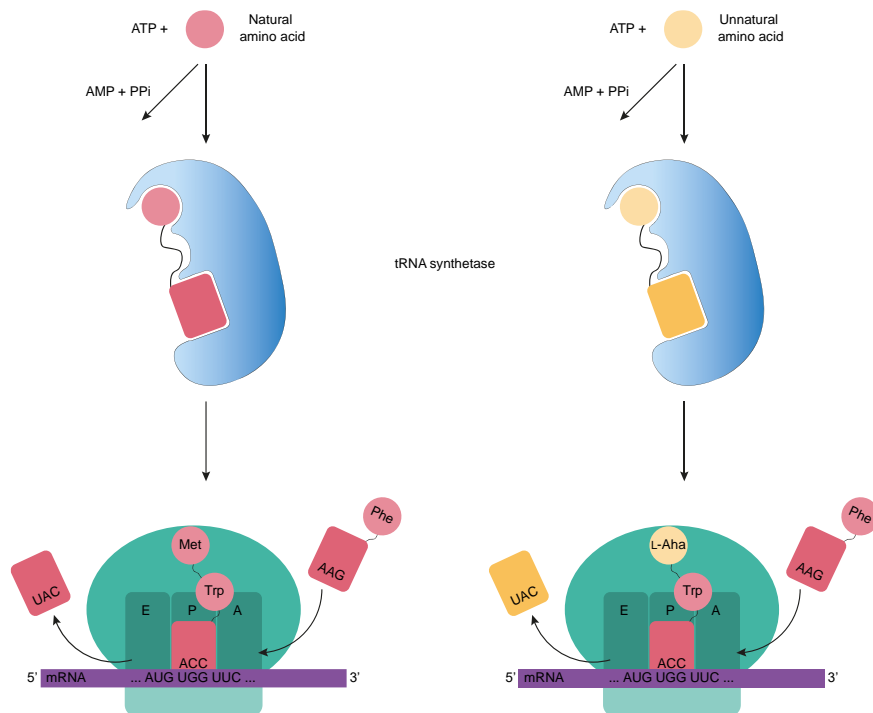
Figure 5. Schematic overview of the incorporation of unnatural amino acids in proteins using the native tRNA synthetase. Left: The introduction of a natural amino acid. Right: The incorporation of an unnatural amino acid, using the canonical tRNA synthetase.

have been reported as well and that make use of non-canonical amino acids introduced following the Tirrell protocol, for instance ones containing ketone moieties [177], which can react with functionalized hydrazides to form a hydrazine [178].

The above-described examples made use of methionine-auxotrophic *E. coli* expression strains. Since these reports, the approach has been expanded to other amino acid analogues, such as those of L-isoleucine [179], L-leucine [180], L-phenylalanine [181], L-proline [182] and L-tryptophan [183], using the corresponding auxotrophic expression strains.

1.5.3 Utilization of bioorthogonal amino acids incorporated in proteins.

Azidohomoalanine (L-Aha) and homopropargylglycine (L-Hpg) are extensively used as close L-methionine (L-Met) analogues. Their azide or alkyne functional groups can be used for bioorthogonal chemistry (Figure 5) [127, 184]. As the incorporation of

these isosteric amino acids in engineered proteins results in minimal differences in the protein construct compared to wild type counterparts, it offers a plethora of ligation possibilities.

Bioorthogonal labelling of proteins has been used by amongst others the groups of Bertozzi [92], Davis [185] and van Kasteren [186] in the visualization of specific proteins in biological systems. For instance, van Kasteren and Davies could site-selectively mutate all natural methionines in LacS into isoleucine residues and at the same time reintroduce a methionine at a preselected position. When the protein was expressed under Tirrell conditions (methionine-auxotroph strain, no methionine in the medium that did include azidohomoalanine) azide-modified LacS mutants were obtained that could be subsequently modified to carry glycan structures [186].

A variety of biological processes have been evaluated using the described metabolic labelling (section **1.5.2**). Among these is the labelling of immunogenic proteins for the visualization of antigen degradation inside antigen presenting cells [187]. Furthermore, metabolic labelling can be used for proteome labelling of bacteria, which can be subsequently used for detection purposes using correlated light-electron microscopy [188, 189] and for protein enrichment in secretome studies [190].

1.6 Aim of this Thesis

In this Thesis, results on the optimization of expression, and evaluation of bioorthogonal proteins in immunological systems is presented.

Chapter 2 of this Thesis describes the production and purification of ovalbumin (Ova), a common antigenic protein used for studies of both major histocompatibility complex (MHC) class I and II. In this Chapter it is attempted to optimize the production and purification of Ova containing the non-canonical amino acids L-Aha and L-Hpg instead of the natural methionine residue. This optimization was attempted using pET16b containing the following constructs: 6His-TEV-Ova, 10His-TEV-Ova and native Ova. It is shown that the first could be expressed in LB medium and with the incorporation of both unnatural amino acids. However, the yields proved to be low, most likely because of the formation of inclusion bodies during protein expression. For this reason, the second construct – 10His-TEV-Ova – was designed. Expressing this ovalbumin variant resulted in the formation of inclusion bodies even when the construct

was further extended with solubilization tags. For this reason, inclusion body purification was explored and although the yields for the methionine containing protein were substantial, it proved not applicable to Ova with the incorporation of L-Hpg. With this knowledge a final construct was designed, in which the purification tag was removed. This led to the expression of ovalbumin with the incorporation of L-Aha, but not L-Hpg and only in the in-house produced BL21::MetA auxotroph strain. **Chapter 2** ends with the comparison between the multiple expressed proteins.

As protein production as aimed for in **Chapter 2** proved troublesome, it was decided to switch attention to tetanus toxin C fragment (TTCF), results on which are described in **Chapter 3**. It was explored whether or not the MHC class II antigen presentation is altered when the antigenic epitope contains a ligation handle. For this, *TTCF* was modified at the antigenic region and the protein was produced with the incorporation of the unnatural amino acid L-Aha. After confirming the ligation possibilities, degradation of these modified proteins was visualized using bone marrow derived dendritic cells (BMDCs). Next, the antigen processing and presentation of the proteins is analyzed using the 2F2 T cell hybridoma. Presentation of each of the resulting 20 proteins proved possible, however, the IL-2 expression levels remained low. For this reason, the presentation was exemplified using ovalbumin MHC class II modified peptides and the use of A20s and DO11.10. Lastly, these azido-containing peptides were used for an initial immunoprecipitation assay.

 In **Chapter 4**, horseradish peroxidase (HRP) was modified using a diazotransfer reaction on the lysine residues. This chemically modified protein was first assessed for its ability for chemical ligation reactions and afterwards used as a visualization tool. Using CuAAC ligation reactions, it was exemplified that alkylated proteins and bacteria can be ligated to the azidylated HRP. Moreover, it was attempted to ligate the protein to alkylated probes.

Finally, **Chapter 5** summarizes this Thesis and outlines some of the future aims using bioorthogonal proteins. It highlights the potential of this method, for instance by describing the potential protein-protein ligation, in which it is envisioned to homogenously ligate two proteins together by labelling one protein with an alkyne ligation handle and one with an azide handle. Furthermore, the site-specific ligation of

a glycan onto a protein is evaluated and initial antigen presentation studies are described in this Chapter. Moreover, alternative prospects of using protein ligation chemistry is described, of which the use of labelled fluorescent proteins is exemplified.

Chapter 1

References

1. Chen, Z. and Zeng, A-P. (2016). Protein engineering approaches to chemical biotechnology. *Curr. Opin. Biotechnol. 42,* p.198-205

2. Reboul, E. and Borel, P. (2011). Proteins involved in uptake, intracellular tansport and basolateral secretion of fat-soluble vitamins carotenoids by mammalian enterocytes. *Prog. Lipid Res. 50,* p.388-402

3. Jaehme, M. and Slotboom, D.J. (2015). Diversity of membrane transport proteins for vitamins in bacteria and archaea. *Biochim. Biophys. Acta. 1850,* p.565-576

4. Movva, N.R., Nakamura, K. and Inouye M. (1980). Gene structure of the OmpA protein, a major surface protein of *Escherichia coli* required for cell-cell interaction. *J. Mol. Biol. 143,* p.317-328

5. Gloushankova, N.A., Rubtsova, S.N. and Zhitnyak, I.Y. (2017). Cadherin-mediated cell-cell interactions in normal and cancer cells. *Tissue Barriers. 5,* e1356900

6. Woo, S-R., Corrales, L. and Gajewski, T.F. (2015). Innate immune recognition of cancer. *Annu. Rev. Immunol. 33,* p.445-474

7. Lubbers, R., van Essen, M.F., van Kooten, C. and Trouw, L.A. (2017). Production of complement components by cells of the immune system. *Clin. Exp. Immunol. 188,* p.183-194

8. Rosano, G.L. and Ceccarelli, E.A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol. 5,* 172

9. Hartrampf, N., Saebi, A., Poskus, M., Gates, Z.P., Callahan, A.J., Cowfer, A.E., Hanna, S., Antilla, S., Schissel, C.K., Quartararo, A.J., Ye, X., Mijalis, A.J., Simon, M.D., Loas, A., Liu, S., Jessen, C., Nielsen, T.E. and Pentelute, B.L. (2020). Synthesis of proteins by automated flow cytometry. *Science. 368,* p.980-987

10. Young, C.L., Britton, Z.T. and Robinson, A.S. (2012). Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnol. J. 7,* p.620-634

11. Notarangelo, L.D., Kim, M-S., Walter, J.E. and Lee, Y.N. (2016). Human RAG mutations: Biochemistry and clinical implications. *Nat. Rev. Immunol. 16,* p.234-246

12. Crick, F. (1970). Central dogma of molecular biology. *Nature. 227,* p.561-563

13. Moore, M.J. and Proudfoot, N.J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell. 136,* p.688-700

14. Koonin, E.V. and Novozhilov, A.S. (2017). Origin and evolution of the universal genetic code. *Annu. Rev. Genet. 51,* p.45-62

15. Johnson, D.B.F and Wang, L. (2010). Imprints of the genetic code in the ribosome. *Proc. Natl. Acad. Sci. U.S.A. 107,* p.8298-8303

16. Rodina, M.V. and Wintermeyer, W. (2011). The ribosome as a molecular machine: The mechanism of tRNA-mRNA movement in translocation. *Biochem. Soc. Trans. 39,* p.658-662

17. Nakamura, Y. and Ito, K. (1998). How protein reads the stop codon and terminates translation. *Genes Cells. 3,* p.265-278

18. Dabrowski, M., Bukowy-Bieryllo, Z. and Zietkiewics, E. (2015). Translational readthrough potential of natural termination codons in eucaryotes – The impact of RNA sequence. *RNA Biol. 12,* p.950-958

19. Walter, S. and Buchner, J. (2002). Molecular chaperones – cellular machines for protein folding. *Angew. Chem. Int. Ed. Engl. 41,* p.1098-1113

20. Wang, R.F. and Kushner, S.R. (1991). Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli. Gene. 100,* p.195-199

21. Schmidt, F.R. (2004). Recombinant expression systems in the pharmaceutical industry. *Appl. Microbiol. Biotechnol. 65,* p.362-372

22. Mayfield, S.P., Franklin, S.E. and Lerner, R.A. (2003). Expression and assembly of a fully active antibody in algae. *Proc. Natl. Acad. Sci. U.S.A. 100,* 438-442

23. Sawasaki, T., Ogasawara, T., Morishita, R. and Endo, Y. (2002). A cell-free protein synthesis system for high-throughput proteomics. *Proc. Natl. Acad. Sci. U.S.A. 99,* p.14652-14657

24. Swartz, J.R. (2001). Advances in *Escherichia coli* production of therapeutic proteins. *Curr. Opin. Biotechnol. 12,* p.195-201

25. Valderrama-Rincon, J.D., Fisher, A.C., Merritt, J.H., Fan, Y-Y., Reading, C.A., Chhiba, K., Heiss, C., Azadi, P., Aebi, M. and DeLisa, M.P. (2012). An engineered eukaryotic protein glycosylation pathway in *Escherichia coli. Nat. Chem. Biol. 8,* p.434-436

26. Mamat, U., Wilke, K., Bramhill, D., Schromm, A.B., Lindner, B., Kohl, T.A., Corchero, J.L., Villaverde, A., Schaffer, L., Head, S.R., Souvignier, C., Meredith, T.C. and Woodard, R.W. (2015). Detoxifying *Escherichia coli* for endotoxin-free production of recombinant proteins. *Microb. Cell Fact. 14,* 57

27. Fabianek, R.A., Hennecke, H. and Thöny-Meyer, L. (1998). The active-site cysteines of the periplasmic thioredoxin-like protein CcmG of *Escherichia coli* are important but not essential for cytochrome *c* maturation in vivo. *J. Bacteriol. 180,* p.1947-1950

28. Fischer, B., Sumner, I. and Goodenough, P. (1993). Isolation, renaturation, and formation of disulfide bonds of eukaryotic proteins expressed in *Escherichia coli* as inclusion bodies. *Biotechnol. Bioeng. 41,* p.3-13

29. Carrío, M.M. and Villaverde, A. (2002). Construction and deconstruction of bacterial inclusion bodies. *J. Biotechnol. 96,* p.3-12

30. Singh, A., Upadhyay, V., Upadhyay, A.K., Singh, S.M. and Panda, A.K. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microb. Cell. Fact. 14,* 41

31. Jong, W.S.P., Vikström, D., Houben, D., van den Berg van Saparoea, H.B., de Gier, J-W. and Luirink, J. (2017). Application of an *E. coli* signal sequence as a versatile inclusion body tag. *Microb. Cell Fact. 16,* 50

32. Missiakas, D. and Raina, S. (1997). Protein folding in the bacterial periplasm. *J. Bacteriol. 179,* p.2465-2471

33. Schwientek, T. and Ernst, J.F. (1994). Efficient intra- and extracellular production of human beta-1,4-galactosyltransferase in *Saccharomyces cerevisiae* is mediated by yeast secretion leaders. *Gene. 145,* p.299-303

34. Cregg, J.M., Cereghino, J.L., Shi, J. and Higgins, D.R. (2000). Recombinant protein expression in *Pichia pastoris. Mol. Biotechnol. 16,* p.23-52

35. Bonander, N. and Bill, R.M. (2012). Optimising yeast as a host for recombinant protein production (review). *Methods Mol Biol. 866,* p.1-9

36. Conde, R., Cueva, R., Pablo, G., Polaina, J. And Larriba, G. (2004). A search for hyperglycosylation signals in yeast glycoproteins. *J. Biol. Chem. 279,* p.43789-43798

37. Sampaio de Oliveira, K.B., Leite, M.L., Rodrigues, G.R., Morales Duque, H., Andrade da Costa, R., Cunha, V.A., de Loiola Costa, L., da Cunha, N.B., Franco, O.L. and Dias, S.C. (2020). Strategies for recombinant production of antimicrobial peptides with pharmacological potential. *Expert Rev. Clin. Pharmacol. 13,* p.367-390

38. Berger, I., Fitzgerald, D.J. and Richmond, T.J. (2004). Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol. 22,* p.1583-1587

39. Khoury, G.A., Baliban, R.C. and Floudas, C.A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep. 1,* 90

40. Cereghino, J.L. and Cregg, J.M. (2000). Heterologous protein expression in the methylotrophic yeasts *Pichia pastoris. FEMS Microbiology reviews. 24,* p.45-66

41. Vieira Gomes, A.M., Souza Carmo, T., Silva Carvalho, L., Mendonça Bahia, F. and Parachin, N.S. (2018). Comparison of yeasts as hosts for recombinant protein production. *Microorganisms. 6,* 38

42. Khan, K.H. (2013). Gene expression in mammalian cells and its application. *Adv. Pharm. Bull. 3,* p.257-263

43. Chen, C. and Okayama, H. (1987). High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell Biol. 7,* p.2745-2752

44. Cooray, S., Howe, S.J. and Thrasher, A.J. (2012). Retrovirus and lentivirus vector design and methods of cell conditioning. *Methods Enzymol. 507,* p.29-57

45. Kim, T.K. and Eberwine, J.H. (2010). Mammalian cell transfection: The present and the future. *Anal. Bioanal. Chem. 397,* p.3173-3178

46. Portolano, N., Watson, P.J., Fairall, L., Millard, C.J., Milano, C.P., Song, Y., Cowley, S.M. and Schwabe, J.W.R. (2014). Recombinant protein expression for structural biology in HEK 293F suspension cells: A novel and accessible approach. *J. Vis. Exp. 92,* 51897

47. Moore, G.E. and Ulrich, K. (1965). Suspension cultures of mammalian cells. A review. *J. Surg. Res. 5,* p.270-282

48. Studier, F.W. (1991). Use of the bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol. 219,* p.37-44

49. Shilling, P.J., Mirzadeh, K., Cumming, A.J., Widesheim, M., Köck, Z. and Daley, D.O. (2020). Improved designs for pET expression plasmids increase protein production yield in *Escherichia coli. Commun. Biol. 3,* 214

50. Yin, J., Li, G., Ren, X. And Herrler, G. (2007). Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. *J. Biotechnol. 127,* p.335-347

51. Mierendorf, R.C., Morris, B.B., Hammer, B. and Novy, R.E. (1998). Expression and purification of recombinant proteins using the pET system. *Methods Mol. Med. 13,* p.257-292

52. Wurm, D.J., Veiter, L., Ulonska, S., Eggenreich, B., Herwig, C. and Spadiut, O. (2016). The *E. coli* pET expression system revisited – mechanistic correlation between glucose and lactose uptake. *Appl. Microbiol. Biotechnol. 100,* p.8721-8729

53. Bertani, I., Devescovi, G. and Venturi, V. (1999). Controlled specific expression and purification of 6x His-tagged proteins in *Pseudomonas. FEMS. Microbiol. Lett. 179,* p.101-106

54. Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.T. and Dobberstein, B. (1987). A T5 promoter-based transcription-translation system for the analysis of proteins *in vitro* and *in vivo. Methods Enzymol. 155,* p.416-433

55. Khlebnikov, A., Risa, Ø., Skaug, T., Carrier, T.A. and Keasling, J.D. (2000). Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture. *J. Bacteriol. 182,* p.7029-7034

56. Bateman, B.T., Donegan, N.P., Jarry, T.M., Palma, M. and Cheung, A.L. (2001). Evaluation of a tetracycline-inducible promoter in *Staphylococcus aureus* in vitro and in vivo and its application in demonstrating the role of *sigB* in microcolony formation. *Infect. Immun. 69,* p.7851-7857

57. Wegerer, A., Sun, T. and Altenbuchner, J. (2008). Optimization of an *E. coli* L-rhamnose-inducible expression vector: Test of various genetic module combinations. *BMC Biotechnol. 8,* 2

58. Ooi, B.G., Rankin, C. and Miller, L.K. (1989). Downstream sequences augment transcription from the essential initiation site of a baculovirus polyhedrin gene. *J. Mol. Biol. 210,* p.721-736

59. Flick, J.S. and Johnston, M. (1990). Two systems of glucose repression of the GAL1 promoter in *Saccharomyces cerevisiae. Mol. Cell Biol. 10,* p.4757-4769

60. Uversky, V.N. (2013). Posttranslational modification. Brenner's Encyclopedia of Genetics. Academic Press. 2nd edition. p.425-430

61. Walsh, C.T., Garneau-Tsodikova, S. and Gatto, G.J. Jr. (2005). Protein posttranslational modifications: The chemistry of proteome diversifications. *Angew. Chem. Int. Ed. Engl. 44,* p.7342-7372

62. Shental-Bechor, D. and Levy, Y. (2008). Effect on glycosylation on protein folding: A close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U.S.A. 105,* p.8256-8261

63. Corfield, A.P. (2014). Mucins: a biologically relevant glycan barrier in mucosal protection. *Biochim. Biophys. Acta. 1850,* p.236-252

64. Iwona, R. and Katarzyna, S. (2020). Glycosylation of proteins of human skin fibroblasts is changed by rosmarinic acid. *Naunyn Schmiedebergs Arch. Pharmacol. 393,* p.419-427

65. Marth, J.D. and Grewal, P.K. (2008). Mammalian glycosylation in immunity. *Nat. Rev. Immunol. 8,* p.874-887

66. Wang, Y-C., Peterson, S.E. and Loring, J.F. (2014). Protein post-translational modifications and regulation of pluripotency in human stem cells. *Cell Res. 24,* p.143-160

67. Pucci, B., Kasten, M. and Giordano, A. (2000). Cell cycle and apoptosis. *Neoplasia. 2,* p.291-299

68. Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev. 12,* p.599-606

69. Lecker, S.H., Goldberg, A.L. and Mitch, W.E. (2006). Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J. Am. Soc. Nephrol. 17,* p.1807-1819

70. Audgnotto, M. and Peraro, M.D. (2017). Protein post-translational modifications: *In silico* prediction tools and molecular modeling. *Comput. Struct. Biotechnol. J. 15,* p.307-319

71. Varier, R.A. and Timmers, H.T.M. (2011). Histone lysine methylation and demethylation pathways in cancer. *Biochim. Biophys. Acta. 1815,* p.75-89

72. Pugliese, A. (2017). Autoreactive T cells in type 1 diabetes. *J. Clin. Invest. 127,* p.2881-2891

73. Chung, H.S., Wang, S-B., Venkatraman, V., Murray, C.I. and van Eyk, J.E. (2013). Cysteine oxidative posttranslational modifications: Emerging regulation in the cardiovascular system. *Circ. Res. 112,* p.382-392

74. Longo, V.D. and Kennedy, B.K. (2006). Sirtuins in aging and age-related disease. *Cell. 126,* p.257-268

75. Hart, G.W., Slawson, C., Ramirez-Correa, G. and Lagerlof, O. (2011). Cross talk between O-GlcNAcylation and phosphorylation: Roles in signaling, transcription, and chronic disease

76. Craveur, P., Narwani, T.J., Rebehmed, J. and de Brevern, A.G. (2019). Investigation of the impact of PTMs on the protein backbone conformation. *Amino Acids. 51,* p.1065-1079

77. Barber, K.W. and Rinehart, J. (2018). The ABCs of PTMs. *Nat. Chem. Biol. 14,* p.188-192

78. Fernández-Tejada, A., Brailsford, J., Zhang, Q., Shieh, J-H., Moore, M.A.S. and Danishefsky, S.J. (2016). Total synthesis of glycosylated proteins. *Top. Curr. Chem. 362,* p.1-26

79. Zhang, Y., Park, K-Y., Suazo, K.F. and Distefano, M.D. (2018). Recent progress in enzymatic protein labelling techniques and their applications. *Chem. Soc. Rev. 47,* p.9106-9136

80. Chalker, J.M., Bernardes, G.J.L. and Davis, B.G. (2011). A "tag-and-modify" approach to site-selective protein modification. *Acc. Chem. Res. 44,* p.730-741

81. Wright, T.H., Bower, B.J., Chalker, J.M., Bernardes, G.J.L., Wiewiora, R., Ng, W-L., Raj, R., Faulkner, S., Vallée, M.R.J., Phanumartwiwath, A., Coleman, O.D., Thézénas, M-L., Khan, M., Galan, S.R.G., Lercher, L., Schombs, M.W., Gerstberger, S., Palm-Espling, M.E., Baldwin, A.J., Kessler, B.M., Claridge, T.D.W., Mohammed, S. and Davis, B.G. (2016). Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity. *Science. 354,* aag1465

82. Grotenberg, G. & Ploegh, H. (2007). Dressed-up proteins. *Nature. 446,* p.994-995

83. Davis, B.G. (2002). Synthesis of glycoproteins. *Chem. Rev. 102,* p.579-602

84. Wang, P., Dong, S., Shieh, J-H., Peguero, E., Hendrickson, R., Moore, M.A.S. and Danishefsky, S.J. (2013). Erythropoietin derived by chemical synthesis. *Science. 342,* p.1357-1360

85. Fernández-Tejada, A., Haynes, B.F. and Danishefsky, S.J. (2015). Designing synthetic vaccines for HIV. *Expert Rev. Vaccines. 14,* p.815-831

86. Gorelik, A. and van Aalten, D.M.F. (2020). Tools for functional dissection of site-specific O-GlcNAcylation. *RSC Chem. Biol. 1,* p.98-102

87. Chao, Q., Ding, Y., Chen, Z-H., Xiang, M-H., Wang, N. and Gao, X-D. (2020). Recent progress in chemo-enzymatic methods for the synthesis of N-glycans. *Front. Chem. 8,* 513

88. Huang, W., Li, C., Li, B., Umekawa, M., Yamamoto, K., Zhang, X. and Wang, L-X. (2010). Glycosynthases enable a highly efficient chemoenzymatic synthesis of N-glycoproteins carrying intact natural N-glycans. *J. Am. Chem. Soc. 131,* p.2214-2223

89. Shivatare, S.S., Huang, L-Y., Zeng, Y-F., Liao, J-Y., You, T-H., Wang, S-Y., Cheng, T., Chiu, C-W., Chao, P., Chen, L-T., Tsai, T-I., Huang, C-C., Wu, C-Y., Lin, N-H. and Wong, C-H. (2018). Development of glycosynthases with broad glycan specificity for the efficient glycol-remodeling of antibodies. *Chem. Commun. (Camb). 54,* p.6161-6164

90. Chalker, J.M., Bernardes, G.J.L., Lin, Y.A. and Davis, B.G. (2009). Chemical modification of proteins at cysteine: Opportunities in chemistry and biology. *Chem. Asian J. 4,* p.630-640

91. Spicer, C.D. and Davis, B.G. (2014). Selective chemical protein modification. *Nat. Commun. 5,* 4740

92. Sletten, E.M. and Bertozzi, C.R. (2009). Bioorthogonal chemistry: Fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed. Engl. 48,* p.6974-6998

93. Ward, C.C., Kleinman, J.I. and Nomura, D.K. (2017). NHS-esters as versatile reactivity-based probes for mapping proteome-wide ligandable hotspots. *ACS Chem. Biol. 12,* p.1478-1483

94. Bloom, S., Liu, C., Kölmel, D.K., Qiao, J.X., Zhang, Y., Poss, M.A., Ewing, W.R. and MacMillan, D.W.C. (2018). Decarboxylative alkylation: An approach to site-selective bioconjugation of native proteins via oxidation potentials. *Nat. Chem. 10,* p.205-211

95. deGruyter, J.N., Malins, L.R. and Baran, P.S. (2017). Residue-specific peptide modification: A chemist's guide. *Biochemistry. 56,* p.3863-3873

96. Ban, H., Nagano, M., Gavrilyuk, J., Hakamata, W., Inokuma, T. and Barbas, C.F. (2013). Facile and stabile linkages through tyrosine: Bioconjugation strategies with the tyrosine-click reaction. *Bioconjugate Chem. 24,* p.520-532

97. Cong, Y., Pawlisz, E., Bryant, P., Balan, S., Laurine, E., Tommasi, R., Singh, R., Dubey, S., Peciak, K., Bird, M., Sivasankar, A., Swierkosz, J., Muroni, M., Heidelberg, S., Farys, M., Khayrzad, F., Edwards, J., Badescu, G., Hodgson, I., Heise, C., Somavarapu, S., Liddell, J., Powel, K., Zloh, M., Choi, J-W., Godwin, A. and Brocchini, S. (2012). Site-specific PEGylation at histidine tags. *Bioconjugate Chem. 23,* p.248-263

98. Antos, J.M. and Francis, M.B. (2004). Selective tryptophan modification with rhodium carbenoids in aqueous solution. *J. Am. Chem. Soc. 126,* p.10256-10257

99. Taylor, M.T., Nelson, J.E., Suero, M.G. and Gaunt, M.J. (2018). A protein functionalization platform based on selective reactions at methionine residues. *Nature. 562,* p.563-568

100. Rudolph, T.K. and Freeman, B.A. (2009). Transduction of redox signaling by electrophile-protein reactions. *Sci. Signal. 2,* re7

101. Gunnoo, S.B. and Madder, A. (2016). Chemical protein modification through cysteine. *Chembiochem. 17,* p.529-553

102. Creighton, T.E. (1984). Disulfide bond formation in proteins. *Methods Enzymol. 107,* p.305-329

103. Sechi, S. and Chait, B.T. (1998). Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem. 70,* p.5150-5158

104. Macmillan, D., Bill, R.M., Sage, K.A., Fern, D. and Flitsch, S.L. (2001). Selective in vitro glycosylation of recombinant proteins: semi-synthesis of novel homogeneous glycoforms of human erythropoietin. *Chem. Biol. 8,* p.133-145

105. Nair, D.P., Podgórski, M., Chatani, S., Gong, T., Xi, W., Fenoli, C.R. and Bowman, C.N. (2014). The thiol-michael addition click reaction: A powerful and widely used tool in materials chemistry. *Chem. Mater. 26,* p.724-744

106. Ravesco, J.M.J.M., Faustino, H., Trindade, A. and Gois, P.M.P. (2019). Bioconjugation with maleimides: A useful tool for chemical biology. *Chemistry. 25,* p.43-59

107. Pillot, A., Defontaine, A., Fateh, A., Lambert, A., Prasanna, M., Fanuel, M., Pipelier, M, Csaba, N., Violo, T., Camberlain, E. and Grandjean, C. (2019). Site-specific

conjugation for fully controlled glycoconjugate vaccine preparation. *Front Chem. 7,* 726

108. Ni, J., Singh, S. and Wang, L-X. (2003). Synthesis of maleimide-activated carbohydrates as chemoselective tags for site-specific glycosylation of peptides and proteins. *Bioconjug. Chem. 14,* p.232-238

109. Nadal, S., Raj, R., Mohammed, S. and Davis, B.G. (2018). Synthetic post-translational modification of histones. *Curr. Opin. Chem. Biol. 45,* p.35-47

110. Chalker, J.M., Lercher, L., Rose, N.R., Schofield, C.J. and Davis, B.G. (2012). Conversion of cysteine into dehydroalanine enables access to synthetic histones bearing diverse post-translational modifications. *Angew. Chem. Int. Ed. Engl. 51,* p.1835-1839

111. Boutureira, O. and Bernardes, G.J.L. (2015). Advances in chemical protein modification. *Chem. Res. 115,* p.2174-2195

112. Stark, G.R., Stein, W.H. and Moore, S. (1960). Reactions of the cyanate present in aqueous urea with amino acids and proteins. *J. Biol. Chem. 235,* p.3177-3181

113. Narayanan, A. and Jones, L.H. (2015). Sulfonyl fluorides as privileged warheads in chemical biology. *Chem. Sci. 6,* p.2650-2659

114. Spears, R.J. and Fascione, M.A. (2016). Site-selective incorporation and ligation of protein aldehydes. *Org. Biomol. Chem. 14,* p.7622-7638

115. Apel, C., Kasper, M-A., Stieger, C.E., Hackenberger, C.P.R. and Christmann, M. (2019). Protein modification of lysine with 2-(2-styrylcyclopropyl)ethanal. *Org. Lett. 21,* p.10043-10047

116. Fairbanks, A.J. (2017). The ENGases: Versatile biocatalysts for the production of homogeneous *N*-linked glycopeptides and glycoproteins. *Chem. Soc. Rev. 46,* p.5128-5146

117. Meinnel, T., Serero, A. and Giglione, C. (2006). Impact of the N-terminal amino acid on targeted protein degradation. *Biol. Chem. 387,* p.839-851

118. Varland, S., Osberg, O. and Arnesen, T. (2015). N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects. *Proteomics. 15,* p.2385-2401

119. Rosen, C.B. and Francis, M.B. (2017). Targeting the N terminus for site-selective protein modification. *Nat. Chem. Biol. 13,* p.697-705

120. Koniev, O. and Wagner, A. (2015). Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem. Soc. Rev. 44,* p.5495-5551

121. Deng, J-R., Lai, N.C-H., Kung, K. K-Y., Yang, B., Chung, S-F., Leung, A.S-L., Choi, M-C., Leung, Y-C. and Wong, M-K. (2020). N-terminal selective modification of peptides and proteins using 2-ethynylbenzaldehydes. *Comm. Chem. 3,* 67

122. Lohse, J., Swier, L.Y.M., Oudshoorn, R.C., Médard, G., Kuster, B., Slotboom, D-J., Witte, M.D. (2017). Targeted diazotransfer reagents enable selective modification of proteins with azides. *Bioconjugate Chem. 28,* p.913-917

123. van Dongen, S.F., Teeuwen, R.L., Nallani, M., van Berkel, S.S., Cornelissen, J.J., Notte, R.J. and van Hest, J.C. (2009). Single-step azide introduction in proteins via an aqueous diazo transfer. *Bioconjug. Chem. 20,* p.20-23

124. Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K. and Ueda, T. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol. 19,* p.751-755

125. Smolskaya, S., Logashina, Y.A. and Andreev, Y.A. (2020). *Escherichia coli* extract-based cell-free expression system as an alternative for difficult-to-obtain protein biosynthesis. *Int. J. Mol. Sci. 21,* 928

126. Melnikov, S.V. and Söll, D. (2019). Aminoacyl-tRNA synthetases and tRNAs for an expanded genetic code: What makes them orthogonal? *Int. J. Mol. Sci. 20,* 1929

127. Lang, K. and Chin, J.W. (2014). Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev. 114,* p.4764-4806

128. Mukai, T., Kobayashi, T., Hino, N., Ynagisawa, T., Sakamoto, K. and Yokoyama, S. (2008). Adding L-Lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem. Biophys. Res. Commun. 371,* p.818-822

129. Katritzky, A.R. and Narindoshvili, T. (2009). Fluorescent amino acids: Advances in protein-extrinsic fluorophores. *Org. Biomol. Chem. 7,* p.627-634

130. Chin, J.W., Martin, A.B., King, D.S., Wang, L. and Schultz, P.G. (2002). Addition of photocrosslinking amino acid to the genetic code of *Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 99,* p.11020-11024

131. Guo, J., Melançon, C.E. III, Lee, H.S., Groff, D. and Schultz, P.G. (2009). Evolution of amber suppressor tRNAs for efficient bacterial production of unnatural amino acid-containing proteins. *Angew. Chem. Int. Ed. Engl. 48,* p.9148-9151

132. Kim, C.H., Axup, J.Y. and Schultz, P.G. (2015). Protein conjugation with genetically encoded unnatural amino acids. *Curr. Opin. Chem. Biol. 17,* p.412-419

133. Nakamura, Y. and Ito, K. (1998). How protein read the stop codon and terminates translation. *Genes Cells. 3,* p.265-278

134. Clark, D.P., Pazdernik, N.J. and McGehee, M.R. (2019). Chapter 26 – Mutations and Repair. *Molecular Biology* 3rd ed. p.832-879

135. Lin, X., Yu, A.C.S. and Chan, T.F. (2017). Efforts and challenges in engineering the genetic code. *Life. 7,* 12

136. Nakamura, Y., Gojobori, T. and Ikemura, T. (2000). Codon usage tabulated from international DNA sequence databases: Status for the year 2000. *Nucleic Acids. Res. 28,* p.292

137. Korkmaz, G., Holm, M., Wiens, T. and Sanyal, S. (2014). Comprehensive analysis of stop codon usage in bacteria and its correlation with release factor abundance. *J. Biol. Chem. 289,* p.30334-30342

138. Liu, W., Brock, A., Chen, S., Chen, S. and Schultz, P.G. (2007). Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat. Methods. 4,* p.239-244

139. Xiao, H. and Schultz, P.G. (2016). At the interface of chemical and biological synthesis: An expanded genetic code. *Cold Spring Harb. Perspect Biol. 8,* a023945

140. Wang, K., Neumann, H., Peak-Chaw, S.Y. and Chin, J.W. (2007). Evolved orthogonal ribosomes enhance the effieciency of synthetic genetic code expansion. *Nat. Biotechnol. 25,* p.770-777

141. Anthony-Cahill, S.J., Griffith, M.C., Noren, C.J., Suich, D.J. and Schultz, P.G. (1989). Site-specific mutagenesis with unnatural amino acids. *Trends Biochem. Sci. 14,* p.400-403

142. Capone, J.P., Sedivy, J.M., Sharp, P.A. and RajBahndary, U.L. (1986). Introduction of UAG, UAA, and UGA nonsense mutations at a specific site in the *Escherichia coli* chloramphenicol acetyltransferase gene: Use in measurement of amber, ochre, and opal suppression in mammalian cells. *Mol. Cell. Biol. 6,* p.3059-3067

143. Varshney, U. and RajBhandary, U.L. (1990). Initiation of protein synthesis from a termination codon. *Proc. Natl. Acad. Sci. U.S.A. 87,* p.1586-1590

144. Wang, L., Magliery, T.J., Liu, D.R. and Schultz, P.G. (2000). A new functional suppressor tRNA/aminoacyl-tRNA synthetase pair for the in vivo incorporation of unnatural amino acids into proteins. *J. Am. Chem. Soc. 122,* p.5010-2011

145. Wang, L., Brock, A., Herberich, B. and Schultz, P.G. (2001). Expanding the genetic code of *Escherichia coli. Science. 292,* p.498-500

146. Ambrogelly, A., Gundllapalli, S., Herring, S., Polycarpo, C., Frauer, C. and Söll, D. (2007). Pyrrolysine is not hardwired for cotranslational insertion at UAG codons. *Proc. Natl. Acad. Sci. U.S.A. 104,* p.3141-3146

147. Namy, O., Zhou, Y., Gundllapalli, S., Polycarpo, C.R., Denise, A., Rousset, J-P., Söll, D. and Ambrogelly, A. (2007). Adding pyrrolysine to the *Escherichia coli* genetic code. *FEBS Lett. 581,* p.5282-5288

148. Anderson, J.C., Wu, N., Santoro, S.W., Lakshman, V., King, D.S. and Schultz, P.G. (2004). An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U.S.A. 101,* p.7566-7571

149. Santoro, S.W., Anderson, J.C., Lakshman, V. and Schultz, P.G. (2003). An archaebacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in *Escherichia coli. Nucleic Acids Res. 31,* p.6700-6709

150. Chatterjee, A., Xiao, H., Yang, P-Y., Soundararajan, G and Schultz, P.G. (2013). A tryptophanyl-tRNA synthetase/tRNA pair for unnatural amino acid mutagenesis in *E. coli. Angew. Chem. Int. Ed. Engl. 52,* p.5106-5109

151. Varshney, U., Lee, C.P. and RajBhandary, U.L. (1991). Direct analysis of aminoacylation levels of tRNAs *in vivo. J. Biol. Chem. 266,* p.24712-24718

152. Wenthzel, A.M., Stancek, M. and Isaksson, L.A. (1998). Growth phase dependent stop codon readthrough and shift of translation reading frame in *Escherichia coli. FEBS Lett. 421,* p.237-242

153. Cervettini, D., Tang, S., Fried, S.D., Willis, J.C.W., Funke, L.F.H., Colwell, L.J. and Chin, J.W. (2020). Rapid discovery and evolution of orthogonal aminoacyl-tRNA synthetase-tRNA pairs. *Nat. Biotechnol. Online ahead of print.*

154. Koehler, C., Girona, G.E., Reinkemeier, C.D. and Lemke, E.A. (2020). Inducible genetic code expansion in eukaryotes. *Chembiochem. Online ahead of print.*

155. Huang, Y., Russell, W.R., Wan, W., Pai, P-J., Hussell, D.H. and Liu, W. (2010). A convenient method for genetic incorporation of multiple noncanonical amino acids into one protein in *Escherichia coli. Mol. Biosyst. 6,* p.683-686

156. Xiao, H., Chatterjee, A., Choi, S-H., Bajjuri, K.M., Sinha, S.C. and Schultz, P.G. (2013). Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. *Angew. Chem. Int. Ed. Engl. 52,* p.14080-14083

157. Italia, J.S., Addy, P.S., Erickson, S.B., Peeler, J.C., Weerapana, E. and Chatterjee, A. (2019). Mutually orthogonal nonsense-suppression systems and conjugation chemistries for precise protein labeling at up to three distinct sites. *J. Am. Chem. Soc. 141,* p.6204-6212

158. Xie, J. and Schultz, P.G. (2005). An expanding genetic code. *Methods. 36,* p.227-238

159. Liu, C.C. and Schultz, P.G. (2010). Adding new chemistries to the genetic code. *Annu. Rev. Biochem. 79,* p.413-444

160. Chin, J.W. (2017). Expanding and reprogramming the genetic code. *Nature. 550,* p.53-60

161. Jafarnejad, S.M., Kim, S-H. and Sonenberg, N. (2018). Aminoacylation of proteins: New targets for the old ARSenal. *Cell Metab. 27,* p.1-3

162. Cohen, G.N. and Cowie, D.B. (1957). Total replacement of methionine by selenomethionine in proteins of *Escherichia coli. C. R. Hebd Seances Acad. Sci. 244,* p.680-683

163. Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science. 254,* p.51-58

164. van Hest, J.C.M., Kiick, K.L. and Tirrell, D.A. (2000). Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *J. Am. Chem. Soc. 122,* p.1282-1288

165. Kiick, K.L., Saxon, E., Tirrell, D.A. and Bertozzi, C.R. (2002). Incorporation of azides into recombinant proteins for chemoselective modification by Staudinger ligation. *Proc. Natl. Acad. Sci. U.S.A. 99,* p.19-24

166. Kiick, K.L., van Hest, J.C.M. and Tirrell, D.A. (2000). Expanding the scope of protein biosynthesis by altering the methionyl-tRNA synthetase activity of a bacterial expression host. *Angew. Chem. Int. Ed. Engl. 39,* p.2148-2152

167. Budisa, N., Karnbrock, W., Steinbacher, S., Humm, A., Prade, L., Neuefeind, T., Moroder, L. and Huber, R. (1997). Bioincorporation of telluromethionine into

proteins: A promising new approach for X-ray structure analysis of proteins. *J. Mol. Biol. 270,* p.616-623

168. Tang, Y. and Tirrell, D.A. (2002). Attenuation of the editing activity of the *Escherichia coli* leucyl-tRNA synthetase allows incorporation of novel amino acids into proteins in vivo. *Biochemistry. 41,* p.10635-10645

169. Voloshchuk, N. and Montclare, J.K. (2010). Incorporation of unnatural amino acids for synthetic biology. *Mol. Biosyst. 6,* p.65-80

170. Saxon, E., Armstrong, J.I. and Bertozzi, C.R. (2000). A "traceless" Staudinger ligation for the chemoselective synthesis of amide bonds. *Org. Lett. 2,* p.22141-2143

171. Köhn, M. and Breinbauer, R. (2004). The staudinger ligation – a gift to chemical biology. *Angew. Chem. Int. Ed. Engl. 43,* p.3106-3116

172. Meyers, E.L. and Raines, R.T. (2011). A phosphine-mediated conversion of azides to diazo-compounds. *Angew. Chem. Int. Ed. Engl. 48,* p.2359-2363

173. Rostovtsev, V.V., Green, L.G., Fokin, V.V. and Sharpless, K.B. (2002). A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Ed. Eng. 41,* p.2596-2599

174. Wang, Q., Chan, T.R., Hilgraf, R., Fokin, V.V., Sharpless, K.B. and Finn, M.G. (2003). Bioconjugation by copper(I)-catalyzed azide-alkyne [3+2] cycloaddition. *J. Am. Chem. Soc. 125,* p.3192-3193

175. Agard, N.J., Prescher, J.A. and Bertozzi, C.R. (2004). A strain-promoted [3+2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc. 126,* p.15046-15047

176. Lallana, E., Riguera, R. and Fernandez-Megia, E. (2011). Reliable and efficient procedures for the conjugation of biomolecules through Huisgen azide-alkyne cycloadditions. *Angew. Chem. Int. Ed. Engl. 50,* p.8794-8804

177. Datta, D., Wang, P., Carrico, I.S., Mayo, S.L. and Tirrell, D.A. (2002). A designed phenylalanyl-tRNA synthetase variant allows efficient in vivo incorporation of aryl ketone functionality into proteins. *J. Am. Chem. Soc. 124,* p.5652-5653

178. Ngo, J.T. and Tirrell, D.A. (2011). Non-canonical amino acids in the interrogation of cellular protein synthesis. *Acc. Chem. Res. 44,* p.677-685

179. Wang, P., Tang, Y. and Tirrell, D.A. (2003). Incorporation of trifluoroisoleucine into proteins in vivo. *J. Am. Chem. Soc. 125,* p.6900-6906.

180. Tang, Y., Ghirlanda, G., Petka, W.A., Nakajima, T., DeGrado, W.F. and Tirrell, D.A. (2001). *Angew. Chem. Int. Ed. Engl. 40,* p.1494-1496

181. Sharma, N., Furter, R., Kast, P. and Tirrell, D.A. (2000). Efficient introduction of aryl bromide functionality into proteins in vivo. *FEBS lett. 467,* p.37-40

182. Renner, C., Alefelder, S., Bae, J.H., Budisa, N, Huber, R. and Moroder, L. (2001). Fluoroprolines as tools for protein design and engineering. *Angew. Chem. Int. Ed. Engl. 40,* p.923-925

183. Budisa, N., Alefelder, S., Bae, J.H., Golbik, R., Minks, C., Huber, R. and Moroder, L. (2001). Proteins with ß-(thienopyrrolyl)alanines as alternative chromophores and pharmaceutically active amino acids. *Protein Sci. 10,* p.1281-1292

184. Fang, K.Y., Lieblich, S.A. and Tirrell, D.A. (2018). Incorporation of non-canonical amino acids into proteins by global reassignment of sense codons. *Methods Mol. Biol. 1798,* p.173-186

185. Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A. and Schuman, E.M. (2006). Selective identification of newly synthesized proteins in mammalian cells using biorthogonal noncanonical amino acid tagging (BONCAT). *PNAS, 103,* p.9482-9487

186. van Kasteren, S.I., Kramer, H.B., Jensen, H.H., Campbell, S.J., Kirkpatrick, J., Oldham, N.J., Anthony, D.C. and Davis, B.G. (2007). Expanding the diversity of chemical protein modification allows post-translational mimicry. *Nature. 446,* p.1105-1109

187. Araman, M.C., Pieper-Pournara, L., van Leeuwen, T., Kampstra, A.S.B., Bakkum, T., Marqvorsen, M.H.S., Nascimento, C.R., Groenewold, G.J.M., Wulp, W., Camps, M.G.M., Overkleeft, H.S., Ossendorp, F.A., Toes, R.E.M. and van Kasteren, S.I. (2019). Bioorthogonal antigens allow the unbiased study of antigen processing and presentation. *BioRxiv.* p.439323

188. van Elsland, D.M., Bos, E., de Boer, W., Overkleeft, H.S., Koster, A.J. and van Kasteren, S.I. (2016). Detection of bioorthogonal groups by correlative light and electron microscopy allows imaging of degraded bacteria in phagocytes. *Chem. Sci. 7,* p.752-758

189. van Elsland, D.M., Pujals, S., Bakkum, T., Bos, E., Oikonomeas-Koppasis, N., Berlin, I., Neefjes, J., Meijer, A.H., Koster, A.J., Albertazzi, L. and van Kasteren, S.I. (2018). Ultrastructural imaging of Salmonella-host interactions using super-resolution correlative light-electron microscopy of bioorthogonal pathogens. *Chembiochem. 19,* p.1766-1770

190. Eichelbaum, K., Winter, M., Berriel Diaz, M., Herzig, S. and Krijgsveld, J. (2012). Selective enrichment of newly synthesized proteins for quantitative secretome analysis. *Nat. Biotechnol. 30,* p.984-990

Chapter 2

Expression and purification of bioorthogonal ovalbumin

Abstract

   Ovalbumin is an oft-used model protein in immunology. In this Chapter the recombinant expression is explored, and the recombinant expression using bioorthogonal methionine substitution. It was shown that the protein solubility was substantially reduced when the protein was N-terminally ligated to a linker sequence or a purification tag. The protein was therefore expressed without a purification tag.

Introduction

Chicken ovalbumin (Ova), a member of the serine protease inhibitors (serpin) protein family, is the most abundant (~58% by weight) protein in hen egg white [1]. Due to its high abundance and the fact that Ova is antigenic to mice, it has become a stalwart in immunological research [1]. When used immunologically, it is of paramount importance to consider the structural features of ovalbumin, particularly its post-translational modifications, as these can have profound effects on the interactions with immune cells [2].

Nascent Ova may undergo a number of post-translational modifications that combined make that the protein exists in hen egg white in a variety of isoforms. For instance, the Ova polypeptide contains six cysteine (Cys) residues of which two (Cys73 and Cys120) form a disulfide bridge [3-5]. Mature Ova features one *N*-linked glycan at asparagine (Asn) 292 [6, 7], with a second potential glycosylation site at Asn311 [7-9]. In addition, the N-terminus may exist as free amine as well as in N-acylated forms and finally three serine residues within the Ova sequence may isomerize from L to D configuration over time [10].

These and other properties can vary strongly between different Ova-isolates. This is particularly troublesome in the context of the immunological use of the antigen: changes in glycosylation, for example, can change the immunological properties of the protein (by altering protein/lectin interactions) [11]. Changes in stability – such as those that occur by Ser-isomerization – can also alter the rate of proteolysis and antigenicity of antigens [12].

These problems are further compounded when traceable variants of Ova are used. Fluorophore conjugation to the protein has been used extensively to track its intracellular fate [13-16]. However, this conjugation alters the surface charge and lipophilicity by replacing the positive charges of lysine (Lys)-residues with large – often hydrophobic – groups. This results in a full loss of secondary structure of the protein [16]. The features of these reagents resulted in a significantly different T cell activation [16].

To prevent both these problems, it was hypothesized that a recombinant, bioorthogonal, form of the protein would be a powerful reagent. It would result in a reduction of the batch-to-batch variation that complicates the use of this model antigen.

Expression in a prokaryotic expression system, such as *Escherichia coli* (*E. coli*) has, for instance, been reported to yield homogeneous Ova lacking any post-translational modification [4, 5, 17-20], whilst keeping other structural features intact. It would also allow the production of the protein with non-canonical amino acids, using for example the bioorthogonal non-canonical amino acid tagging strategy (BONCAT) [21]. In this approach the detectable groups are not appended from Lys/Cys sidechains, but incorporated as amino acids isosteric to methionine [22]. It was further hypothesized that this Ova-variant would be very suitable as a detectable reagent that would be compatible with the wealth of other available immune reagents for Ova, including T cell reagents targeting specific MHC-I and MHC-II restricted regions of the protein [23-26].

Bioorthogonal Ova variants are attractive commodities, as they allow for the introduction of desirable functionalities pre-, ante- or post-experimentation, thus enabling flexibility in immunological studies, especially given the increase in specificity, selectivity and sensitivity in bioorthogonal reactions that are continuously becoming available in present times. The availability of flexible and effective expression systems is key for the use of bioorthogonal Ova in immunological studies. Recent studies [16] have addressed this issue and set the stage for the experimental work presented in this Chapter.

Thus, this Chapter presents optimized protocols for the introduction of bioorthogonal amino acids in well-defined Ova constructs. Of the various methodologies that have seen the light in the past decades for introducing such non-canonical commodities, the strategy developed initially by Tirrell and co-workers was selected. The reason for this was the relative ease of execution and the relatively large quantities of modified protein that can be obtained.

The strategy is rooted in supplementing culture media of an amino acid auxotrophic bacterial strain with an amino acid that structurally closely resembles that of the amino acid the bacteria cannot produce itself. Methionine-auxotrophic *E. coli* strains are most often used for this purpose and this determines also the nature of the substituting amino acid, which needs to resemble methionine and at the same time needs to harbor a bioorthogonal group. Ova encompasses 17 methionine residues, that in theory may all be substituted for a bioorthogonal lookalike, and the research

described here therefore also aimed to establish which methionine, and what number of methionine, could be supplanted depending on the conditions used.

Results

The introduction of non-canonical amino acids in proteins through modified expression systems can be achieved through expanding the genetic code, or by substitution in growth media of a canonical amino acid for a close structural analogue. The first strategy, originally developed by Schultz *et al.*, allows for introduction of a single bioorthogonal amino acid at a precise position within the polypeptide sequence [27]. The methodology is however laborious, fraught with failure and oftentimes yields only small quantities of the desired protein. The second strategy, originating from work by Tirrell and co- workers, is experimentally more robust and allows for the generation of large quantities of the desired protein, but has the caveat that a canonical amino acid is potentially replaced by a non-canonical one throughout the polypeptide [22, 28]. Work described in this chapter followed the Tirrell procedure, in particular the substitution of methionine for either azidohomoalanine or homopropargylglycine, both bioorthogonal amino acids and both close structural methionine analogues [22, 28].

The general workflow followed is depicted in Figure 1 and comprises selection of a suitable expression host (here: a methionine-auxotrophic *E. coli* strain), cloning of the desired plasmid into this, expression of the protein of interest using culture media in which the canonical amino acid is substituted for the desired non-canonical one (here: methionine for either azidohomoalanine or homopropargylglycine). This is followed by protein isolation, purification and finally analysis of the thus obtained protein product to establish structural integrity (have the bioorthogonal amino acids been incorporated in the predetermined positions, and if so, which of the 17 possible positions, and to what extent) and functionality (does the bioorthogonal protein act as the wild type one does). The following sections detail results on the design, cloning, expression, purification, and analysis of a number of bioorthogonal Ova constructs. The Chapter will conclude with general observations on the obtained results, will weigh advantages and disadvantages for the various constructs and their preparation and will end with recommendations
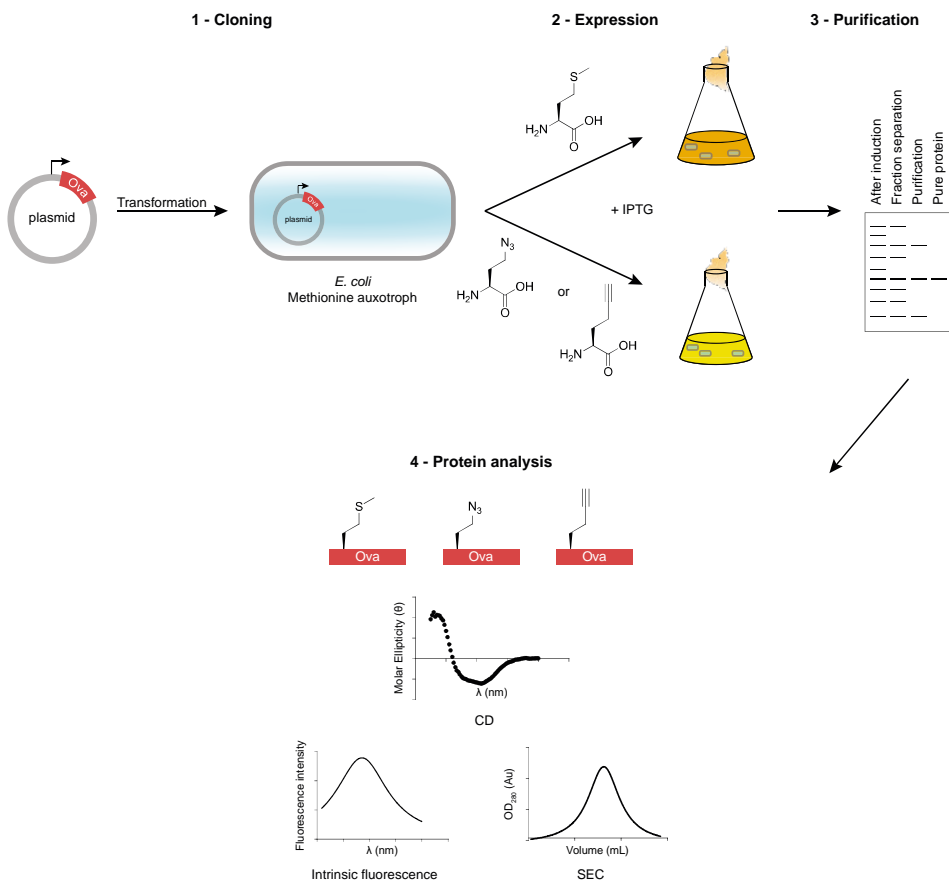
Figure 1. General strategy for the production of bioorthogonal Ova as described in this chapter. Individual steps subject of here-described studies are the design of Ova constructs and their incorporation into appropriate plasmids, which are then used to transform methionine-auxotrophic bacteria (**1**); induced expression in the presence of either methionine or analogous, bioorthogonal amino acids (**2**); protein separation and detection of (modified) Ova (**3**); and functional analysis of thus obtained (bioorthogonal) Ova (**4**).

which procedure to follow in function of the type of bioorthogonal Ova construct desired.

2.1 Expression of 6His-TEV-Ova

As a first starting point for producing bioorthogonal Ova, the previously reported expression construct bearing a six-histidine (His) tag and a Tobacco Etch Virus (TEV) cleavage site at its N-terminus (6His-TEV-Ova), was taken [29]. The expression plasmid pMSCG7 was kindly provided by N. Del Cid. The 6-His-tag allows for immobilized metal

affinity chromatography (IMAC) purification, whereas the TEV cleavage site facilitates proteolytic removal [30] of the His-tag from the target protein following nickel-nitrilotriacetic acid (Ni-NTA) purification to yield a minimally modified protein.

With the aim to substitute methionine residues for bioorthogonal isosteres, the auxotrophic *E. coli* strain B834(DE3) [31] was transformed with the pMSCG7 plasmid bearing ovalbumin. Expressions were first attempted in (methionine-containing) lysogeny broth (LB) medium and expression was induced by the addition of isopropyl-β-D-1-thiogalactopyranoside (IPTG) [32]. After harvesting, the cells were subjected to a freeze-thaw cycle followed by suspension in lysis buffer and subsequent lysis via French press. The soluble fraction was separated by ultracentrifugation and protein fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) showing robust overexpression of a protein at the correct size (45 kDa) (Figure 2A), further confirmed by Western blot as well as its presence in both the soluble and insoluble fractions (Figure 2B). However, some pre-induction background expression was observed in this system as revealed by Western blot analysis (Figure 2A). As this can lead to lower incorporation of the unnatural amino acids, it was attempted to suppress this using the addition of 1% glucose (w/v) to the pre-expression culture medium (Figure 2A right lanes) [32].

Next, Ova expression in methionine replacement medium was tested. This SelenoMet-medium is of slightly different nutrient composition, which may affect expression levels. Expression in SelenoMet-medium augmented with methionine (L-Met; 40 mg/L) showed robust expression of recombinant Ova (Figure 2C).

Following these successful pilot experiments, expression with bioorthogonal amino acids was attempted next with this construct. After growing *E. coli* B834-6His-TEV-Ova to the exponential phase (OD_{600} of ~0.8) in glucose-augmented LB medium, the cells were harvested by centrifugation and a medium switch was performed to SelenoMet medium supplemented with azidohomoalanine (L-Aha; 72 mg/L) or homopropargylglycine (L-Hpg; 40 mg/L). Expression was then induced with IPTG as before and continued overnight. Ensuing sample preparation and SDS-PAGE analysis as before showed overexpression of a protein band at 45 kDa, that was present in both fractions.
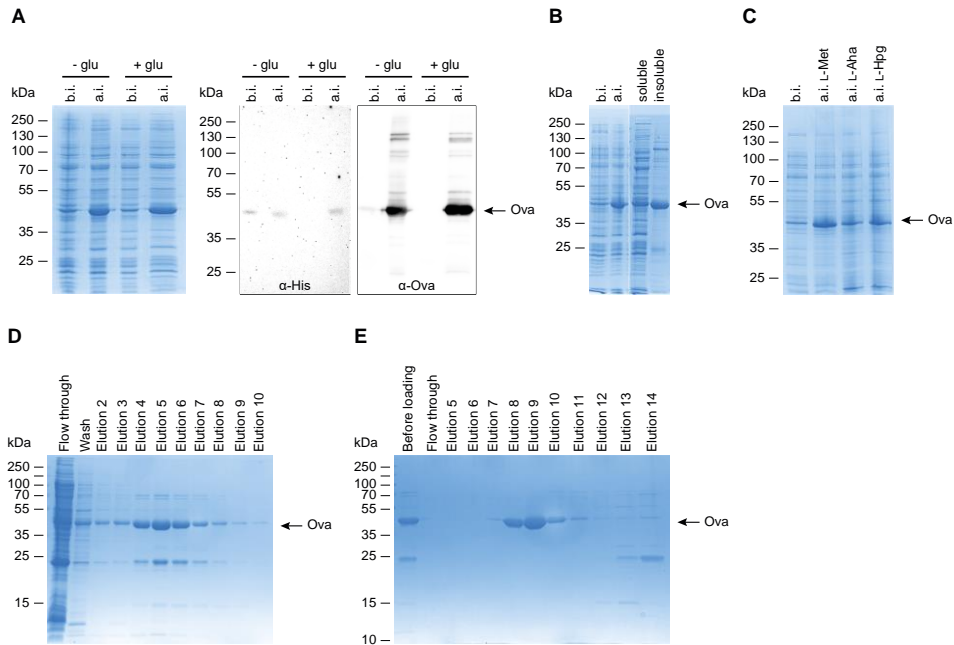
Figure 2. Expression of 6His-TEV-Ova. A) Expression of Ova with and without the suppression of glucose before induction. Protein expression was suppressed in the presence of 1% glucose, as revealed by Western blot analysis using α-His (middle) and α-Ova (right) antibodies. B) Expressed Ova separated into the soluble and insoluble protein fractions, with most of the protein ending up in the insoluble fraction. C) Expression of Ova in media supplemented with L-Met, L-Aha or L-Hpg. D) Ni-NTA purification of Ova-Aha, of which fractions 4 - 7 were further purified over E) a Q column, resulting in >95% pure protein. A-E) Proteins were resolved in a 10% or 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie Brilliant Blue G-250.

The recombinant 6His-Ova-Aha was then purified by IMAC using a Ni-NTA column (Figure 2D), followed by anionic ion exchange (IEX) chromatography using a Q column (Figure 2E). This resulted in >95% pure ovalbumin (as determined by SDS-PAGE analysis) in a (low) yield of 0.4 - 0.8 mg per liter expression culture.

2.2 Expression of 10His-TEV-Ova

The low recovery yield of the bioorthogonal Ova from expression experiment **2.1** could be because Ova proteins were produced in inclusion bodies; that the 6His-tag proved too short to effect complete binding to the Ni-NTA column; that the B834(DE3) strain proved insufficiently efficient in expressing the target protein; or a combination
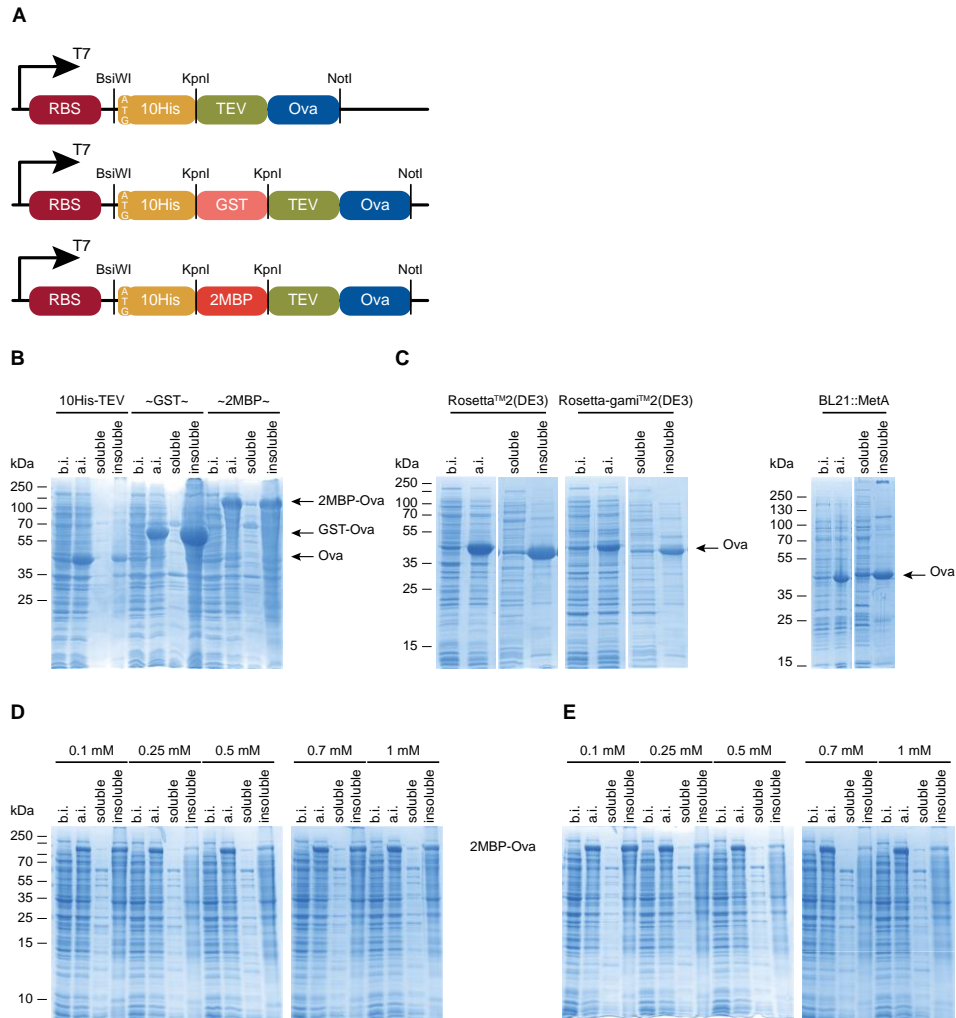
Figure 3. Comparison of soluble expression levels of 10His-x-TEV-Ova. A) Schematic overview of the used gene constructs. B) Protein expression of 10His-TEV-Ova, 10His-GST-TEV-Ova and 10His-2MBP-TEV-Ova in B834, as separated in soluble and insoluble protein fractions. C) 10His-TEV-Ova expressed in Rosetta™ 2(DE3), Rosetta-gami™ 2(DE3) and BL21::MetA as separated in soluble and insoluble fractions. D) Expression of 10His-2MBP-TEV-Ova at 18˚C or E) 30˚C with different concentrations of IPTG, as separated in soluble and insoluble protein fractions. B-E) Proteins were resolved in a 10% or a 12.5% SDS-PAGE and stained with Coomassie Brilliant Blue G-250 or R-250.

of these. Experiments discussed in this section were conducted with the aim to address these issues.

With the aim to potentially increase protein binding to the Ni-NTA column [33], a plasmid allowing the expression of Ova with a larger His-tag was constructed:

pET28a_10His-TEV-Ova. The multiple cloning site of plasmid pET28a was extended with a linker sequence containing 10, rather than 6, histidine residues, again linked to *Ova* via a TEV cleavage site. The full-length *Ova*-gene was ligated N-terminally to the 10His-TEV linker sequence into the multiple cloning site (Figure 3A). The plasmid was then transformed into the same auxotrophic expression strain B834(DE3) as used under **2.1**, but also into the non-auxotrophs Rosetta™ 2(DE3) [34], Rosetta-gami™ 2(DE3) and the in-house produced Met-auxotroph BL21::MetA [35]. After expression in LB medium, harvesting of the cells and their lysis, the soluble and insoluble fractions were separated by centrifugation [20]. Both fractions were analyzed by SDS-PAGE, which revealed overexpression of a protein at the correct size (45 kDa). Most of the protein was found in the insoluble fraction (Figure 3B and 3C). It was first attempted to increase the solubility of the Ova-constructs by creating genetic fusions to the solubility-enhancing proteins GST and MBP [36-38]. For this reason, plasmids pET28a_10His-GST-TEV-Ova and pET28a_10His-2MBP-TEV-Ova were designed. The plasmids were obtained by extending the 10His-TEV linker with either a *GST* gene [39] or a *MalE* gene [40]. Again, Ova was N-terminally ligated (Figure 3A) and the resulting plasmids were transformed to B834(DE3). The protein constructs were expressed and solubilized as described under **2.1** and the resulting fractions were analyzed by SDS-PAGE showing expression of 10His-GST-TEV- Ova (at 71 kDa) and 10His-2MBP-TEV-Ova (at 129 kDa). Again, fusion proteins proved to localize predominantly to the insoluble fraction during expression (Figure 3B).

After the failure of these solubility enhancing proteins, it was next attempted to enhance solubility by lowering the rate of expression. This was done in two ways: lowering the IPTG concentration [41-43], and expression at lower temperatures [44-46]. Expression of the 10His-2MBP-TEV-Ova by induction with either 0.1, 0.25, 0.5, 0.7 or 1 mM IPTG, at either 18°C or 30°C was attempted. Again, none of these measures improved the fraction of soluble protein (Figure 3D and 3E). Based on these results it was decided to investigate the possibility to obtain Ova constructs from the insoluble protein fraction.
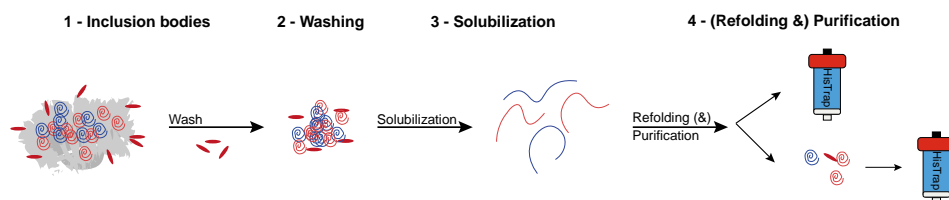
Figure 4. Workflow of inclusion body purification. The inclusion bodies were washed, solubilized and either loaded directly onto a Ni-NTA column, or refolded before loading onto a Ni-NTA column.

2.3 Isolation and purification of 10His-TEV-Ova from inclusion bodies

As all efforts to maximize expression of the protein in the soluble fraction proved fruitless (see section 2.2), it was chosen to explore the option of obtaining Ova from the insoluble fraction. For this, protocols based on either a solubilization-refolding-purification strategy [47, 48], or a solubilization-purification-refolding strategy were investigated (Figure 4).

In both approaches the insoluble fraction after washing and harvesting is first denatured in a chaotropic buffer to solubilize the protein content (Figure 4 – step 1-3). Chaotropic buffers (buffers containing high molarity urea or guanidine) have been shown to disrupt secondary structures of proteins leading to solutions of unfolded polypeptides. The thus obtained solution of denatured proteins was either directly loaded onto the Ni-NTA column (Figure 4 – step 4 upper) or subjected to refolding using a flash-dilution protocol prior to loading onto the column (Figure 4 – step 4 lower) [20].

The latter approach was explored first. After expression and lysis, the soluble protein fractions were removed by centrifugation and the remaining insoluble fractions were washed twice with water. Next, the inclusion bodies were solubilized in 50 mM Tris-HCl pH 8.5, 8 M urea and 5 mM 1,4-dithiothreitol (DTT), with the use of sonication. The resulting soluble fraction was cleared by centrifugation and then flash diluted (10x) in 50 mM Tris-HCl pH 8.5 [20]. Aggregates were removed by centrifugation and the solute (assumed to contain the refolded Ova) was then loaded onto the Ni-NTA column. The column was then washed, and the protein eluted using increasing imidazole concentrations (Figure 5A). For the solubilization-purification-refolding route (Figure 4 – step 4 upper route), the inclusion bodies were harvested as before (see "lower route"),

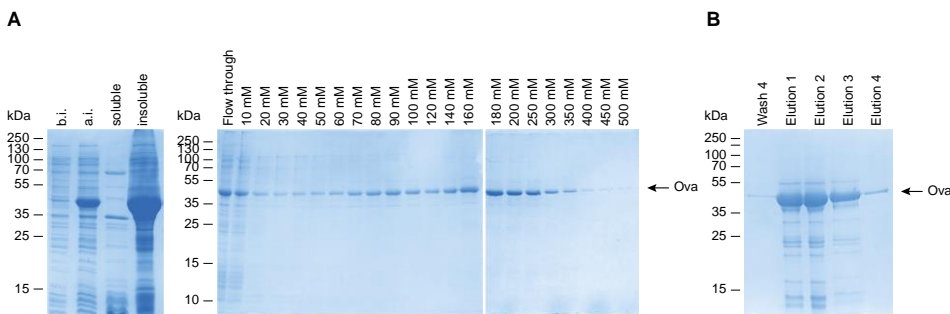**A**   **B**

Figure 5. Ni-NTA protein purification profile of native and denatured 10His-TEV-Ova. A) Expressed 10His-TEV-Ova as separated in soluble and insoluble protein fractions were subsequently subjected to refolding conditions and loaded onto a Ni-NTA column. The protein eluted in all elution fractions. B) Purification of the same construct under denaturing conditions. A-B) Proteins were resolved in a 12.5% SDS-PAGE and stained with Coomassie Brilliant Blue R-250.

with the exception that lysis was performed by French press in these particular experiments.

The insoluble fraction was then washed once with lysis buffer containing Triton X-100, with the aim to remove membrane and cell wall associated proteins [49], and twice with lysis buffer. Finally, the washed inclusion bodies were solubilized in a chaotropic buffer (6 M guanidine-HCl in 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole), cleared from residual debris by centrifugation and then loaded directly onto the column in this unfolded state. The column was then washed with 2 x 5 column volumes (CV) of a wash buffer containing 6 M guanidine-HCl and twice with 5 CV of a wash buffer containing 6 M urea. The urea wash step was necessary because of the disruptive properties of guanidine in subsequent analyses. The protein was eluted with a single CV of wash buffer containing 500 mM imidazole.

SDS-PAGE of the eluted fractions obtained under native conditions showed Ova in the flow through, wash, and each of the elution fractions (Figure 5A). The same was observed for the protein isolated under denaturing conditions (Figure 5B). It was therefore concluded that Ni^{2+}-affinity chromatography was unsuitable for purifying these constructs. An alternative purification procedure was therefore investigated next (section 2.4) with the aim to purify 10His-TEV-Ova under native conditions using ion exchange chromatography only.
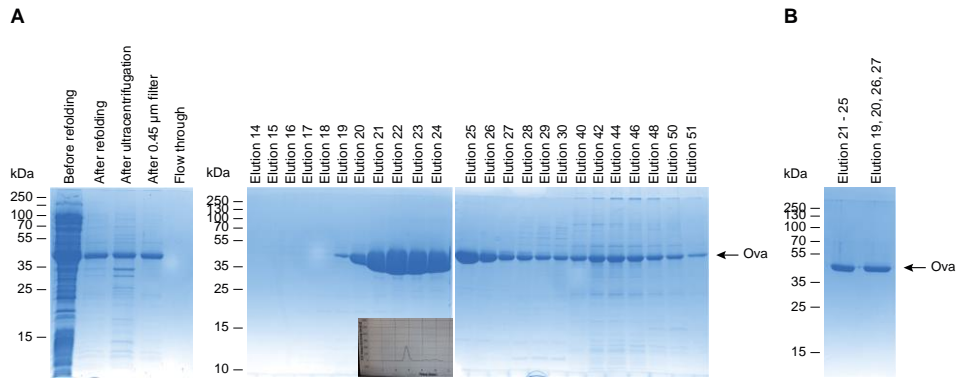
Figure 6. 10His-TEV-Ova purification using a Q column. A) Purification and elution profile of 10His-TEV-Ova, elution fractions were collected in 0.2 CV fractions. Insert shows an almost symmetrical elution pattern. B) Combined elution fractions 21 – 25; 1.5 µg and elution fractions 19, 20, 26 and 27; 1.5 µg. A-B) Proteins were resolved in a 12.5% SDS-PAGE and stained with Coomassie Brilliant Blue R-250.

2.4 Isolation and purification of 10His-TEV-Ova by ion exchange chromatography

To purify Ova proteins by ion exchange, the quaternary ammonium anion exchange column as described in section **2.1** was used. However, rather than first attempting a Ni^{2+}-affinity purification, the proteins were loaded directly onto this column. For this, the inclusion bodies were denatured and flash diluted as described in section **2.3**, and then left standing for 4.5 days at 4°C. After removal of the aggregates, the remaining solution was loaded onto a Q-column which was then washed with 10 CV of buffer (50 mM Tris-HCl pH 8.5) and eluted by increasing the salt concentration from 0 - 500 mM NaCl. SDS-PAGE analysis revealed complete binding to the column (no putative Ova observed in the flow through, see Figure 6A lane "Flow through") and emergence of a protein at the correct molecular weight (MW) following elution at a salt concentration of 190 - 270 mM. (Figure 6A and insert). The protein obtained was >95% pure as determined by SDS-PAGE (Figure 6B) and the yield after purification was 15.1 mg/500 mL expression culture with a single purification step.

Having an optimized isolation/refolding/purification protocol in hand, this protocol was next used to investigate whether bioorthogonal Ova could be obtained. To this end, Ova-Hpg was expressed as described in section **2.2**, and the resulting insoluble fraction was subjected to the above purification protocol. Unfortunately, the protein
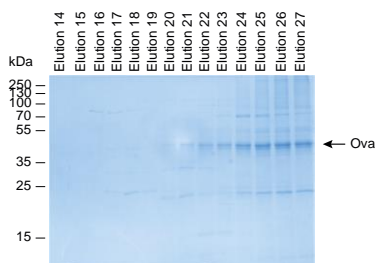
Figure 7. 10His-TEV-Ova-Hpg purification using a Q column. A) Elution fractions validated the presence of Ova only in the higher salt elution fractions. Proteins were resolved in a 12.5% SDS-PAGE and stained with Coomassie Brilliant Blue R-250.

only eluted at higher salt concentrations (Figure 7). This indicated that the protein was either unfolded or aggregated, suggesting that the presence of the unnatural amino acids represents a complicating factor in refolding.

2.5 Expression of native Ova

One of the reasons for the trouble with refolding of 10His-TEV-Ova as cloned and expressed in section **2.2** could be the N-terminal 10His-TEV-appendage. As is evident from the experiments described under section **2.4**, purification protocols do not necessarily include nickel columns which would make a His-tag obsolete [50, 51] and therefore an expression plasmid for Ova lacking this sequence was constructed.

With the aim to investigate whether Ova constructs lacking a His-tag would better segregate into the soluble fraction or alternatively be easier to retain from the insoluble fraction in appropriately folded form, native, full-length *Ova* was cloned into the multiple cloning site of plasmid pET16b (Figure 8A) and the resulting plasmid was transformed into the methionine auxotrophs B834(DE3), and BL21::MetA. Next, the protein was expressed in LB medium and solubilized as described under section **2.1** for 6His-TEV-Ova. SDS-PAGE analysis of both soluble and insoluble fractions showed that expression was successful in both expression strains and that the Ova protein ended up predominantly in the soluble fraction in each case (Figure 8B).

Purification of soluble expressed Ova might require the introduction of an additional purification step prior to Q column loading. This because, loading the cleared lysate directly onto the column might result in the co-elution of multiple proteins together with the target protein.
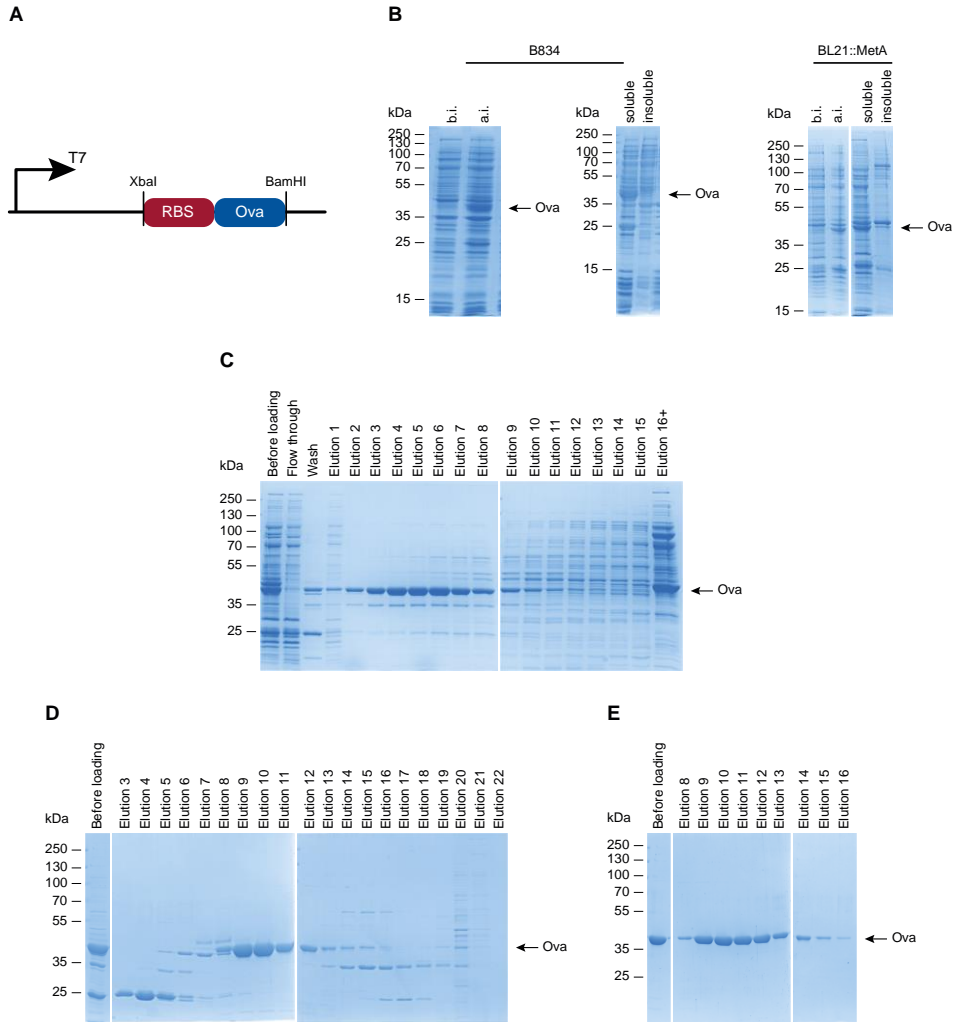
Figure 8. Expression and purification of native Ova. A) Schematic overview of the used gene construct. B) Expression of Ova in B834(DE3) and BL21::MetA and its separation in the soluble and insoluble protein fractions. C) Protein elution profile of Ova over a DEAE column in a 15 CV gradient 20 - 150 mM NaCl in 50 mM MOPS pH 7.5. D) Protein elution profile over a Q HP column in a 20 CV gradient 40 - 150 mM NaCl in 50 mM NaHCO$_3$ pH 8.0. E) Protein elution profile over a Q HP column in a 30 CV gradient 40 - 150 mM NaCl in 50 mM Tris-HCl pH 8.5. A-E) Proteins were resolved in a 10% or a 12.5% SDS-PAGE and stained with Coomassie Brilliant Blue G-250.

With the aim to remove some of the co-expressed proteins before loading the soluble fraction to the column, it was decided to assess the use of osmotic stress and ammonium sulfate precipitation [52, 53]. As lysis by osmotic stress can result in the clearance of
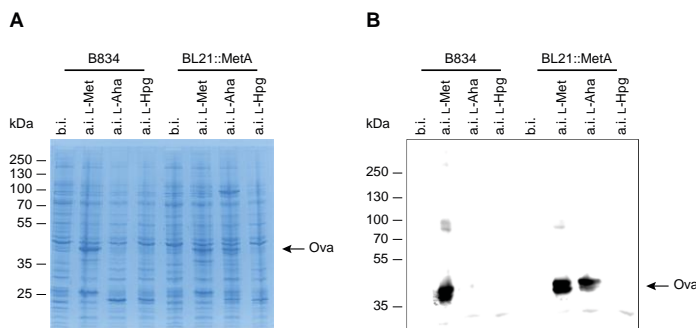
Figure 9. Expression of native Ova with the incorporation of L-Met, L-Aha and L-Hpg. A) Expression of Ova in B834(DE3) and BL21::MetA. B) Protein expression was confirmed by Western blotting against Ova. A) Proteins were resolved in a 10% SDS-PAGE and stained with Coomassie Brilliant Blue G-250.

protein without disturbing the cell wall with e.g. lysozyme or the use of French press. Additionally, ammonium sulfate precipitation after cell lysis can result in the removal of co-expressed proteins with a solubility difference as opposed to Ova in the presence of salt [53] (data not shown).

As the use of these two methods in combination with a Q column proved unsuccessful (data not shown), it was decided to perform a first purification over diethylaminoethyl (DEAE) Sepharose instead. The filtered soluble fraction was loaded onto a DEAE column and the protein was partially purified by eluting in a 10 CV gradient of 20 - 300 mM NaCl in MOPS buffer (Figure 8C). The Ova-containing fractions were next buffer exchanged to 50 mM NaHCO$_3$ pH 8.0 and loaded onto a Q column. The protein was eluted with a 20 CV gradient 40 - 150 mM NaCl in 50 mM NaHCO$_3$ pH 8.0 (Figure 8D). The proteins in elution fractions 9 - 12 were next combined, buffer exchanged to 50 mM Tris-HCl pH 8.0 and loaded again onto a Q column. Elution with a 30 CV gradient 40 - 150 mM NaCl followed by SDS-PAGE analysis of the fractions eluted from this final column showed Ova in >95% purity, in a yield of 5.2 mg/250 mL expression culture (Figure 8E).

2.6 Expression of bioorthogonal variants of native Ova

To obtain the Hpg- and Aha-containing variants of untagged Ova, which was obtained in pure form under section **2.5**, the expression with the incorporation of these
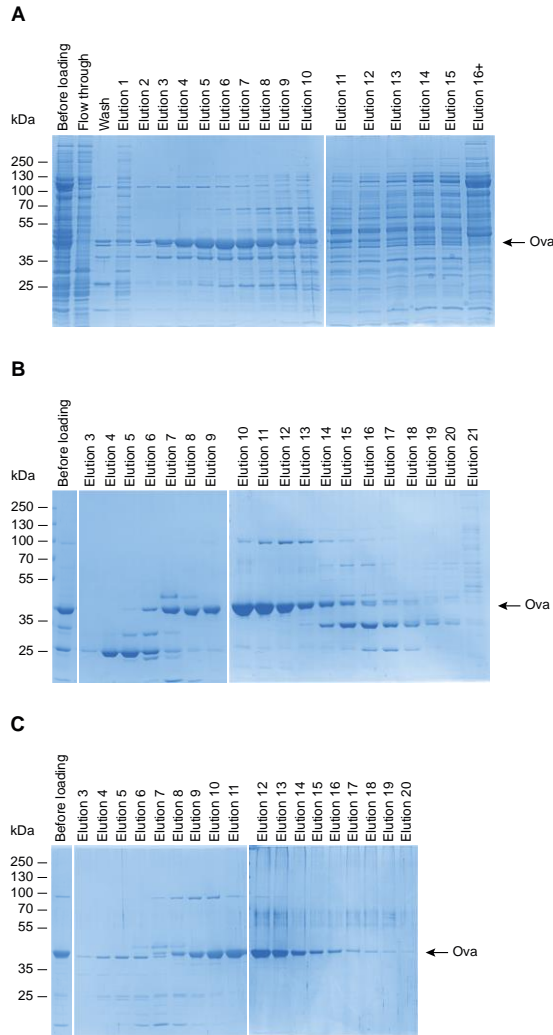
Figure 10. Purification of native Ova-Aha. A) Protein elution profile of Ova over a DEAE column in a 15 CV gradient 20 - 150 mM NaCl in 50 mM MOPS pH 7.5. B) Protein elution profile over a Q HP column in a 20 CV gradient 40 - 150 mM NaCl in 50 mM NaHCO₃ pH 8.0. C) Protein elution profile over a Q HP column in a 30 CV gradient 40 - 150 mM NaCl in 50 mM Tris-HCl pH 8.5. A-C) Proteins were resolved in a 10% SDS-PAGE and stained with Coomassie Brilliant Blue G-250.

unnatural amino acids in expression strains B834 and BL21::MetA in otherwise unaltered conditions was assessed next (see section **2.1**). The resulting before and after induction samples were analyzed by SDS-PAGE and Western blot. These showed that Ova-Hpg was not present in these after induction samples as derived from either of the

two expression strains (Figure 9). The same holds true for Ova-Aha in case of B834 usage. However, Ova-Aha was produced in the experiments done with BL21::MetA (Figure 9).

Next, native Ova-Aha was purified using the established (see section 2.5) protocol. SDS-PAGE analysis of the various fractions obtained in this manner showed (Figure 10A-C) a major band corresponding to the size of Ova (and Ova-Aha) in >95% purity, in a yield of 4.3 mg/250 mL expression culture (Figure 10C).

2.7 Analysis of 6His-TEV-Ova, native Ova and the L-Aha counterparts

Having successfully expressed native Ova and native Ova-Aha, it was decided to compare these structurally to 6His-TEV-Ova, 6His-TEV-Ova-Aha and Ova extracted from chicken egg (ggOva) in order to analyze whether the modified proteins are similar and as stable as its wild type counterpart.

First, the purity of the different proteins was visualized via SDS-PAGE and Native PAGE. This showed a purity for all of the proteins of >95% (Figure 11A and 11B). However, in some of the samples a protein band at ~35 kDa was visible. Further α-Ova Western blot analysis confirmed this protein to be Ova related (Figure 11C).

In addition, for some of the protein samples a minor protein band was present. Which was especially noticeable for native Ova and native Ova-Aha at ~35 kDa, a protein band mentioned earlier. It is known that a protein in a different folding-state can run differently over SDS-PAGE, it was aimed to analyze if this was the case for these proteins. For this, native Ova-Aha was combined with Laemmli sample buffer containing different concentrations of DTT. Subsequent SDS-PAGE analysis showed that with increasing concentrations of reducing agent, an increase in the 43 kDa – ovalbumin – band was seen (Figure 11D). However, this was not the case for the complete fraction and LC-MS/MS analysis confirmed the presence of a protein of ~35 kDa in the native Ova sample.

With no major in gel differences between 6His-TEV-Ova, native Ova, and their Aha counterparts compared to ggOva, the reactivity of the proteins towards Alexa Fluor 647 (AF647) alkyne under copper-catalyzed Huisgen cycloaddition conditions was assessed, in order to test the incorporation of Aha (Figure 11E). For this each of the proteins was

combined with a copper(I)-catalyzed azide alkyne cycloaddition (CuAAC) ligation mixture containing the fluorophore. The resulting in-gel fluorescence showed a higher intensity for 6His-Ova-TEV-Aha as compared to native Ova-Aha. This could be a first indication that not all L-Met residues were replaced by L-Aha in the latter protein.

Next, the multimerization of the protein was analyzed via analytical Size Exclusion Chromatography Multi-Angle Light Scattering (SEC-MALS). The plots resulting from these measurements showed the presence of two or three different peaks, representing the monomeric and higher oligomeric states of Ova (Figure 11F).

The analysis was continued with the identification of the proteins secondary and tertiary structure. The secondary structure was measured via far-UV CD measurements (Figure 11G). The proteins were diluted to 0.2 mg/mL in 50 mM $NaHCO_3$ pH 8.0, before recording the spectra. ggOva showed a 43% α-helical and 22% β-sheet content in this buffer. The α-helical content of the protein was only different for native Ova and the β-sheet content only for 6His-TEV-Ova. Surprisingly, the azidylated versions showed to be less different from ggOva than the methionine containing proteins.
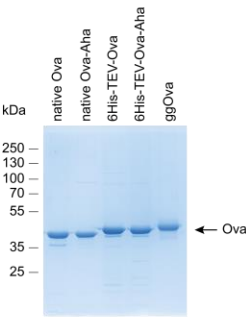
The tertiary structure was characterized via fluorescence spectroscopy [54]. For this, the proteins were diluted to 0.5 mg/mL in 50 mM $NaHCO_3$ pH 8.0 and the emission spectra of the 3 tryptophan residues were recorded after an excitation at 280 nm. An emission peak ~340 nm was measured for ggOva (Figure 11H), as reported in literature [20, 55, 56]. This peak was lacking for 6His-TEV-Ova-Aha, which indicated a folding difference not visible in other experiments [54, 57].

Another method to assess correct protein folding and stability is via thermal stability measurements. The discussed proteins showed, except for ggOva, melting temperatures (T_m) corresponding to those found in literature (Figure 11I) [19]. Both 6His-TEV-Ova and native Ova showed a T_m of ~78-79˚C and the azidylated proteins showed a T_m of ~72-73˚C. Interestingly, Ova-Aha also showed a T_m of 78.1˚C, which might be a second indication of incomplete L-Met to L-Aha replacement.
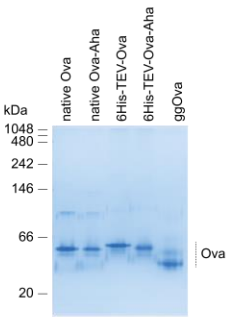
It was shown earlier that ovalbumin is resistant to trypsin when it is in its native protein conformation [20]. To test if the Ova samples are indeed in the native
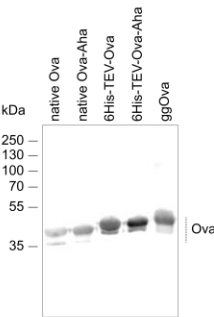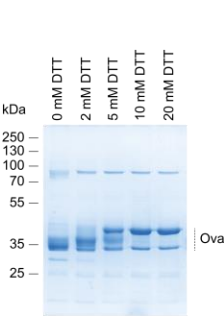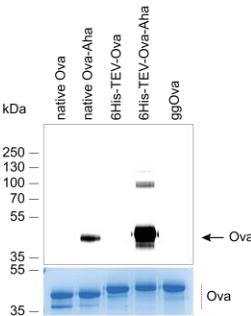
A

	native Ova	native Ova-Aha	6His-TEV-Ova	6His-TEV-Ova-Aha	ggOva

kDa
250
130
100
70
55
← Ova
35
25

B

	native Ova	native Ova-Aha	6His-TEV-Ova	6His-TEV-Ova-Aha	ggOva

kDa
1048
480
242
146
66   Ova
20

C

	native Ova	native Ova-Aha	6His-TEV-Ova	6His-TEV-Ova-Aha	ggOva

kDa
250
130
100
70
55
Ova
35

D

	0 mM DTT	2 mM DTT	5 mM DTT	10 mM DTT	20 mM DTT

kDa
250
130
100
70
55
Ova
35
25

E

	native Ova	native Ova-Aha	6His-TEV-Ova	6His-TEV-Ova-Aha	ggOva

kDa
250
130
100
70
55
← Ova
35

55
Ova
35

F

Multimeric protein   Monomeric protein

Relative intensity vs Time (min)

- ▼ native Ova
- ♦ native Ova-Aha
- ■ 6His-TEV-Ova
- ▲ 6His-TEV-Ova-Aha
- ● ggOva

G

θ (deg × cm² × dmol⁻¹) vs λ (nm)

- ▼ native Ova
- ♦ native Ova-Aha
- ■ 6His-TEV-Ova
- ▲ 6His-TEV-Ova-Aha
- ● ggOva

H

Relative intensity vs λ (nm)

- native Ova
- native Ova-Aha
- 6His-TEV-Ova
- 6His-TEV-Ova-Aha
- ggOva

I

Ration 350/330 nm vs Temperature (°C)

- native Ova
- native Ova-Aha
- 6His-TEV-Ova
- 6His-TEV-Ova-Aha
- ggOva

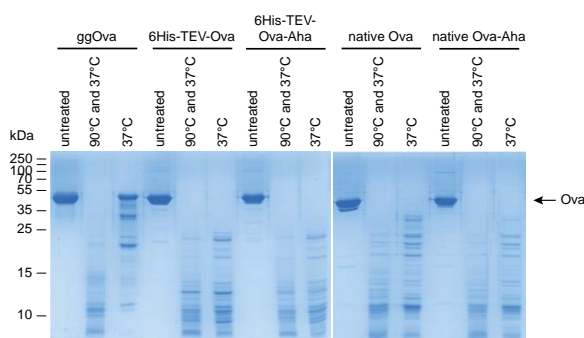| Sample	T_m (°C)	
native Ova		78.7
native Ova-Aha	71.8	78.1
6His-TEV-Ova		77.8
6His-TEV-Ova-Aha	72.8	
ggOva		83.2

J

Figure 11. Complete analysis of recombinant ovalbumin as compared to ggOva. A) Protein purity of ovalbumin as assessed by SDS-PAGE. B) Protein purity as evaluated by Native PAGE. C) Western blot analysis against Ova showed the same protein purity as SDS-PAGE. D) SDS-PAGE analysis of Ova-Aha, showing the existence of different sub forms with increasing DTT concentrations. E) Protein ligation to AF647 alkyne, which resulted in fluorescence of the azidylated proteins (upper panel). Protein loading was confirmed by Coomassie staining (lower panel). F) SEC-MALS traces of the different protein samples, measured by relative intensity. G) Far-UV CD analysis of the different protein samples measured in 50 mM NaHCO$_3$ pH 8.0. H) Fluorescence emission spectra of Ova, measured by relative intensity. I) Melting temperatures of the Ova samples, measured by the 350/330 nm ratio and sorted by equal Tm. J) Trypsin resistance assay for the protein samples. With untreated samples as negative control and 90˚C treated samples as positive control. A, B, D, E and J) Proteins were resolved in a 10% or a 15% SDS-PAGE and stained with Coomassie Brilliant Blue G-250.

conformation, the protein was analyzed via a trypsin resistance assay. In this study, the protein was either heated at 90˚C before treatment with 0.25 µg of trypsin for 10 min at 37˚C or only treated with the protease at 37˚C. As a negative control, the protein was combined with buffer. Subsequently, the trypsin cleavage was analyzed via SDS-PAGE. The protein remained intact only in the negative control sample (Figure 11J). Whereas pretreatment of the sample at 90˚C or treatment at 37˚C alone, showed a complete degradation profile for each of the recombinant proteins. ggOva remained partly intact when treated with trypsin at 37˚C.

Discussion and conclusion

In this Chapter the expression and purification of different Ova constructs was described. In sector 2.1 it was shown that 6His-TEV-Ova could be expressed even with the incorporation of the unnatural amino acids L-Aha and L-Hpg. The purification of this protein was exemplified by 6His-TEV-Ova-Aha, but was the same for 6His-TEV-Ova and 6His-TEV-Ova-Hpg.

In sector **2.2-2.4** the expression and purification of 10His-TEV-Ova was described. And though protein expression was successful when the protein was either grown in LB medium or with the incorporation of L-Hpg, purification of this protein construct was more troublesome. For each purification, the same protocol was followed. Nonetheless, refolding and Q column purification showed varying amounts of protein aggregates in different purifications. Moreover, it was so far impossible to refold and purify 10His-TEV-Ova-Hpg. Expression and purification of 10His-TEV-Ova with the incorporation of L-Aha was not assessed, but due to the mentioned issues it is not suggested to try this approach.

The expression and purification of native Ova was successful in the preferred methionine auxotroph *E. coli* strain B834(DE3). However, incorporation of the unnatural amino acids was troublesome and therefore it was necessary to use the in-house produced BL21::MetA strain to express Ova-Aha. One of the explanations for the expression differences between B834(DE3) and BL21::MetA could be a discrepancy in expression rates, which was observed earlier by van Elsland *et al.* [58], which could contribute to premature protein degradation. The irreproducible protein expression when L-Hpg was incorporated in native ova, which was sometimes observed after 1 or 2 hours of expression, supported the idea that higher expression levels of the protein triggered protein degradation. A second explanation for this could be that BL21::MetA was not a complete methionine auxotroph, which means unnatural amino acids compete with methionine to be incorporated [59]. This consequently lowers the amount of unnatural amino acids in the protein and might therefore result in a heterogeneous protein product bearing these two amino acid residues. This idea was supported by genomic DNA extraction and subsequent sequencing of the MetA region, which showed to be fully intact. However, as it is not known how the strain was designed, this data remains inconclusive.

The only conclusive way to confirm the incorporation of L-Aha residues in native Ova (and 6His-TEV-Ova), is via LC-MS/MS analysis. Comparison of the LC-MS/MS spectra of both native Ova and native Ova-Aha showed no differences. This was most
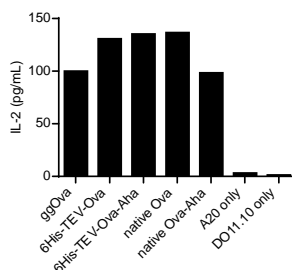
Figure 12: Antigen presentation of different Ova proteins by A20 to DO11.10 T cell hybridomas. DO11.10 activation after antigen presentation of ggOva, 6His-TEV-Ova, 6His-TEV-Ova-Aha, native Ova or native Ova-Aha via A20. IL-2 read out was measured by an ELISA assay.

likely due to insufficient ionization capacities of azidylated ovalbumin. In future experiments, the protein should be digested with trypsin and subsequent peptide analysis should be performed.

All these proteins were expressed and purified with the aim to subsequently use them in T cell assays. For this reason, 0.25 mg/mL (final concentration) of the five different proteins as described in section 2.7 were given to A20 B cells and after 4 hours of initial pulse, the antigen was presented for 20 hours to DO11.10 T cell hybridomas and their activation was subsequently measured via an IL-2 ELISA assay. The preliminary data, of this single replicate, hinted at equal processing of the antigens (Figure 12), but a proper full scale experiment will be conducted in future.

In conclusion, this Chapter showed the successful expression and purification of 6His-TEV-Ova and 6His-TEV-Ova-Aha, moreover it showed the successful optimization of these for native Ova and native Ova-Aha. Which methionine residues were replaced remains inconclusive. Nevertheless, initial results suggest the successful use of the proteins in T cell assays.

Acknowledgements

Materials and Methods

General

All chemicals and reagents were purchased at Sigma-Aldrich, Alfa Aesar, Acros, Merck or VWR, unless stated otherwise. SDS-PAGE, Native PAGE and Western blot materials were purchased at Bio-Rad. Cloning reagents were ordered at Thermo Fisher Scientific. DNA primers were ordered at Sigma-Aldrich or Integrated DNA Technologies. Protease inhibitors were obtained from Roche or Amresco. Cell culture disposables were from Greiner or Sarstedt.

Solutions

PBS contained 5 mM KH_2PO_4, 15 mM Na_2HPO_4, 150 mM NaCl, pH 7.4 and PBST was PBS supplemented with 0.05% Tween-20. TBS contained 50 mM Tris-HCl, 150 mM NaCl and TBST was TBS supplemented with 0.05% Tween-20. Laemmli sample buffer 4* contained 60 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.01% (v/v) bromophenol blue.

Strains and plasmids

E. coli strains XL10, B834(DE3), Rosetta-gami™ 2(DE3), Rosetta™ 2(DE3) and BL21::MetA were used as cloning and expression strains. Plasmid pMSCG7 was used for the expression of 6His-Ova, pET28a was modified accordingly and used for the expression of 10His-TEV-Ova, 10His-2MBP-TEV-Ova and 10His-GST-TEV-Ova, and pET16b was used for the expression of native Ova. All plasmids contain the IPTG-inducible T7 promoter.

Cloning

Plasmid design

pMSCG7 containing full-length chicken Ova was a kind gift from the M. Raghavan lab. For bacterial expression constructs, pET28a was modified with a 10His-TEV linker sequence, which was ligated into the vector using the BsiWI and NheI restriction sites. This vector was extended with a *GST* or a 2 *MalE* gene, using the KpnI restriction site, resulting in pET28a_10His-GST-TEV and pET28a_10His-2MBP-TEV respectively. To obtain 10His-TEV-OVA, 10His-2MBP-TEV-Ova or 10His-GST-TEV-Ova the DNA

fragment encoding Ova was amplified by PCR from the pMSCG7 plasmid containing full-length Ova, and ligated into pET28a_10His-TEV vector using the NcoI and NotI restriction sites. To obtain native Ova, Ova was ligated into the pET16b vector using the XbaI and BamHI restriction sites. All sequences were verified by Sanger sequencing (Macrogen).

Genomic DNA extraction

An overnight culture of BL21::MetA in LB medium was sedimented (500 µL) and the pellet was flash frozen in liquid N_2. The pellet was thawed at rt and resuspended in 500 µL NT1 buffer, further purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel). Genomic PCR reactions were performed on 2 µL isolated genomic DNA using Phusion High-Fidelity DNA Polymerase in Phusion GC buffer in a final volume of 50 µL. These products were gel purified and sequences were verified by Sanger sequencing.

Table 1: List of primer sequences

#	primer name	sequence 5' -> 3'
p1	pET28a_His-TEV linker_fwd	CATGTGGTACCATGCATCACCATCATCACCATCACCACCATCACTGTACAGAGAACCTGTACTTCCAAGGCCCATGGG
p2	pET28a_His-TEV linker_rev	CTAGCCCATGGGCCTTGGAAGTACAGGTTCTCTGTACAGTGATGGTGGTGATGGTGATGGTGATGCATGGTACCACATG
p3	10HisTEV_Ova_fwd	GGCGGCGGTCTCCCATGCGGCTCCATCGGCGGCAGC
p4	10HisTEV_Ova_rev	GGCGGCGCGGCCGCTTAAGGGGAAACACATCTGCC
p5	10HisTEV_MBP_fwd	CAGGGTACCAAAATCGAAGA
p6	10HisTEV_MBP_rev	CGTGGTACCGTTGTTGTTGTTATT
p7	10HisTEV_GST_fwd	CAGGGTACCAGCCCTATACTAGG
p8	10HisTEV_GST_rev	CGTGGTACCTTTTGGAGGATGG
p9	pET16b_Ova_fwd	GGCGGCGGATCCTTAAGGGGAAACACATCTGC
p10	pET16b_Ova_rev	GGCGGTTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCTCCATCGGCGCAGC
p11	MetA_genomic DNA_fwd	CATGCAGGCTCGACATTGGC
p12	MetA_genomic DNA_rev	CACCTGCTGAGGTACGTTTC

Protein expression

Expression of 6His-TEV-Ova, 10His-TEV-Ova, 10His-GST-TEV-Ova, 10His-2MBP-TEV-Ova and native Ova

An overnight culture of B834(DE3) containing pMSCG7_Ova, pET28a_10His-TEV-Ova, pET28a_10His-GST-TEV-Ova or pET28a_10His-2MBP-TEV-Ova or pET16b_Ova or BL21::MetA containing pET16b_Ova was diluted 1:100 in LB medium containing 50 µg/mL ampicillin (pMSCG7 and pET16b) or 25 µg/mL kanamycin (pET28a) in the presence or absence of 1% glucose. Cells were grown at 37˚C, 180 rpm to an OD_{600} of ~0.7-1.2 and sedimented (3220 rcf, 20 min, 4˚C) before being resuspended in LB medium containing either 50 µg/mL ampicillin or 25 µg/mL kanamycin. The culture was induced

with either 1 mM or indicated concentrations of IPTG (final concentration) and expression took place for 4 h at 18˚C or 30˚C. After expression, cells were sedimented and washed once with PBS. Pellets were stored at -80˚C until protein purification. The before and after induction samples were suspended in 1* Laemmli buffer and resolved in either a 10% or a 12.5% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue R-250 or G-250) was used for protein analysis.

Expression of 6His-TEV-Ova, 10His-TEV-Ova and native Ova with the incorporation of L-Met, L-Aha or L-Hpg

An overnight culture of B834(DE3; 6His-TEV-Ova and 10His-TEV-Ova) or BL21::MetA (native Ova) was diluted 1:100 - 1:143 in LB medium containing 50 µg/mL ampicillin (pMSCG7) and 1% glucose. Cells were grown at 37˚C, 180 rpm to an OD_{600} of ~0.8-0.9 and sedimented (3428 rcf, 15 min, 4˚C) before being washed twice with SelenoMet medium (Molecular Dimensions). Cells were resuspended in SelenoMet medium containing 50 µg/mL ampicillin (pMSCG7) and depleted for 30 min at 37˚C, 180 rpm before 30-50 min depletion at 30˚C, 130 rpm. After this either L-Methionine (L-Met; 40 mg/L; Ajinomoto) L-azidohomoalanine (L-Aha TFA salt; 72 mg/L) or L-Homopropargylglycine (D/L-Hpg; 80 mg/L) was added to the culture and 15 min later the culture was induced with IPTG (1 mM final concentration), expression took place overnight. After expression, cells were sedimented and washed once with PBS. Pellets were stored at -80˚C until protein purification. The before and after induction samples were suspended in 1* Laemmli buffer and resolved in either a 10% or a 12.5% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue R-250 or G-250) was used for protein analysis.

Protein solubilization
6His-TEV-Ova expressed in B834; 10His-TEV-Ova expressed in Rosetta™ 2(DE3), Rosetta-gami™ 2(DE3) and BL21::MetA; native Ova expressed in BL21::MetA

The bacterial pellets (25 mL or 50 mL culture) were resuspended in 10 mL lysis buffer containing 50 mM Tris-HCl pH 8.5, 100 mM NaCl, 1 mM PMSF, 5 mM EDTA. The bacterial cells were disrupted by sonication on ice, 30% amplitude 10 sec on, 10 sec off

for 6 cycles or by French Press, before separating the soluble from the insoluble protein (15000 rcf, 30 min, 4˚C). The resulting supernatant (soluble protein) and pellet (insoluble protein) were separated and the pellet was resuspended in 10 mL lysis buffer. Or the pellet was washed 2 times with MQ water (1 mL) and pelleted by centrifugation (15000 rcf, 20 min, 4˚C). The pellet was resuspended in 5 mL resuspension buffer containing 50 mM Tris-HCl pH 8.5, 100 mM NaCl, 1 mM PMSF, 5 mM EDTA and incubated for 2 hours on the rollerbench. Pellet and supernatant were separated by centrifugation (15000 rcf, 20 min, 20˚C). The pellet was resuspended in 250 µL or 1 mL solubilization buffer containing 50 mM Tris-HCl pH 8.5, 8 M urea, 5 mM DTT (insoluble fraction). Protein fractions were combined with 4* Laemmli buffer and all resolved in a 10% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue R-250 or G-250) was used for protein analysis.

10His-TEV-Ova, 10His-GST-TEV-Ova and 10His-2MBP-TEV-Ova expressed in B834

The bacterial pellets (10 mL culture) were resuspended in 200 µL resuspension buffer containing 50 mM Tris-HCl pH 8.5, 100 mM NaCl, 1 mM PMSF, 5 mM EDTA. The bacterial cells were disrupted by sonication on ice, 30% amplitude 9 sec on, 6 sec off for 3 min, before separating the soluble from the insoluble protein (30130 rcf, 20 min, 4˚C). The pellet was washed 2 times with MQ water (1 mL) and pelleted by centrifugation (30130 rcf, 20 min, 4˚C). The inclusion bodies were solubilized in 40 µL MQ water and 250 µL solubilization buffer containing 50 mM Tris-HCl pH 8.5, 8 M urea, 5 mM DTT and disrupted by sonication as described before. The solubilized fraction was separated by centrifugation (17500 rpm, 30 min, 25˚C). The protein fractions were combined with 4* Laemmli buffer and all resolved in a 12.5% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue R-250 or G-250) was used for protein analysis.

Native Ova expressed in B834

The bacterial pellet (30 mL culture) was resuspended in 5 mL lysis buffer containing 50 mM Tris-HCl pH 8.5, 100 mM NaCl, 1 mM PMSF, 5 mM EDTA and the bacterial cells were disrupted by sonication on ice, 30% amplitude 10 sec on, 10 sec off for 6 cycles, before separating the soluble from the insoluble protein (15000 rcf, 20 min, 4˚C). The

insoluble protein fraction was suspended in 5 mL lysis buffer for analysis. Subsequently, the protein fractions were combined with 4* Laemmli buffer and all resolved in a 10% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue G-250) was used for protein analysis.

Large scale solubilization of 10His-TEV-Ova and 10His-TEV-Ova-Hpg

As described previously with the following adjustments: The bacterial pellet (500 mL) was resuspended in 10 mL lysis buffer. The bacterial cells were disrupted by sonication on ice 30% amplitude 9 sec on, 1 sec off for 15 min (L-Hpg), before separating the soluble from the insoluble protein fraction (15000 rcf, 20 min, 4˚C). The pellet was washed 2 times with MQ water (1 mL) and the washed inclusion bodies were solubilized in 12.5 mL solubilization buffer for 1 - 2 hours rolling at RT.

Guanidine-HCl solubilization of 10His-TEV-Ova

The bacterial pellet (250 mL culture) was resuspended in 10 mL lysis buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 1 mg/mL lysozyme, 5 mM β-ME, 250 U benzonase. The bacterial cells were disrupted by French press at 1.9 kbar or by sonication on ice, 30% amplitude 9 sec on, 6 sec off for 3 minutes, before separating the soluble from the insoluble protein (15000 rcf, 30 min, 4˚C). The pellet was washed 1x with lysis buffer containing 1% Triton X-100 and 2x with lysis buffer (1 mL each) and pelleted by centrifugation (30130 rcf, 20 min, 4˚C). The inclusion bodies were solubilized in 2.5 mL solubilization buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 6 M guanidine-HCl, 10 mM imidazole overnight at 4˚C.

Refolding

The solubilized protein as described under "large scale solubilization of 10His-TEV-Ova and 10His-TEV-Ova-Hpg" was flash refolded by 10x dilution in refolding buffer (50 mM Tris-HCl pH 8.5) and either used directly or left for 4.5-5.5 days at 4˚C. Aggregates were removed by centrifugation (17500 rpm, 30 min, 4˚C or 25000 rpm, 30 min, 4˚C; Ti70 rotor, Beckman Coulter) and the supernatant was filtered over 0.2 or 0.45 µm (Filtropur S, Sarstedt).

Chapter 2

Purification

6His-TEV-Ova; 6His-TEV-Ova-Aha and 6His-TEV-Ova-Hpg

The bacterial pellets (2 L culture) were suspended in 20 mL lysis buffer containing 50 mM NaH_2PO_4 pH 8.0, 500 mM NaCl, protease inhibitor, 250 U benzonase. The bacterial cells were disrupted by sonication on ice, 25% amplitude 4 sec on, 9 sec off for 30 cycles, before separating the soluble from the insoluble protein (15000 rcf, 30 min, 4˚C). The soluble fraction was loaded on a prewashed (50 mM NaH_2PO_4 pH 8.0, 500 mM NaCl, 10 mM imidazole) 1 mL Ni-NTA agarose column, connected to an ÄKTA start FPLC system equipped with UV, pH and conductance detectors at a flow rate of 1 CV/min. Next, the column was washed with 50 CV wash buffer (50 mM NaH_2PO_4 pH 8.0, 500 mM NaCl, 20 mM imidazole), before being eluted with a 20 CV gradient of wash buffer to elution buffer (50 mM NaH_2PO_4 pH 8.0, 500 mM NaCl, 500 mM imidazole). Ova containing fractions were dialyzed extensively against 10 mM $NaHCO_3$ and loaded onto a 1 mL Q HP column. The protein was eluted with a 15 CV gradient 0 – 500 mM NaCl in 10 mM $NaHCO_3$. Protein concentrations were determined using the mass extinction coefficient (ε is 36900/ (M cm) and MW 43063 g/mol, Eppendorf BioPhotometer). The protein fractions were combined with 4* Laemmli buffer and all resolved in a 12.5% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue R-250) was used for protein analysis.

10His-TEV-Ova - Ni-NTA agarose

The refolded protein as described under "refolding" was loaded onto a 250 µL Ni-NTA agarose column and the protein was eluted in 2 CV fractions with increasing concentrations of imidazole (0 – 500 mM) in a buffer containing 50 mM Tris-HCl, 500 mM NaCl. The protein fractions were combined with 4* Laemmli buffer and all resolved in a 12.5% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue R-250) was used for protein analysis.

The solubilized inclusion bodies as described under "large scale solubilization of 10His-TEV-Ova and 10His-TEV-Ova-Hpg" were loaded onto a 1 mL Ni-NTA agarose column and washed 2x with 5 CV wash buffer A containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 6 M guanidine-HCl, 10 mM imidazole and 2x with 5 CV wash buffer B

containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 6 M urea, 20 mM imidazole. The protein was eluted with 4x 1 CV elution buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 6 M urea, 500 mM imidazole. The urea containing protein fractions were combined with 4* Laemmli buffer and resolved in a 12.5% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue R-250) was used for protein analysis.

10His-TEV-Ova and 10His-TEV-OVA-Hpg - Q Sepharose HP IEX

The filtered (Filtropur S 0.2 or 0.45 μm; Sarstedt) refolded protein fraction was loaded on a prewashed (5 CV MQ water, 5 CV high salt buffer (50 mM Tris-HCl pH 8.5, 500 mM NaCl) and 10 CV low salt buffer (50 mM Tris-HCl)) 5 mL Q Sepharose column (HiTrap™ Q HP; GE Healthcare), connected to an ÄKTA start FPLC system equipped with UV, pH and conductance detectors at a flow rate of 1 CV/min. The column was washed with 10 CV of 50 mM Tris-HCl and bound proteins were eluted with a 10 CV gradient of the low salt buffer to the high salt buffer and 1 mL fractions were collected. The protein fractions were combined with 4* Laemmli buffer and all resolved in a 12.5% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue R-250) was used for protein analysis. Fractions were combined according to purity and extensively dialysed (6 - 8 kDa MWCO, 3.3 mL/cm, FisherBrand or 12 - 14 kDa MWCO, 2 mL/cm, Spectra/Por) against 50 mM Tris-HCl, pH 8.5. Protein concentrations were measured using the mass extinction coefficient (ε = 32890; MW = 45493 Da) and protein fractions were combined by concentration.

Native Ova and native Ova-Aha - DEAE Sepharose FF IEX

The bacterial pellet (250 mL) was resuspended in 50 mL lysis buffer containing 50 mM MOPS pH 7.5, 10% glycerol, 10 mM DTT, 500 U benzonase, 1 mM PMSF, 1 mM EDTA and the protocol was continued as described above. The bacterial cells were disrupted by French Press, before separating the soluble from the insoluble protein (15000 rcf, 30 min, 4°C). The filtered (Filtropur S 0.2 μm; Sarstedt) protein fraction was loaded on a prewashed (5 CV MQ water, 5 CV high salt 50 mM MOPS buffer and 10 CV low salt 50 mM MOPS buffer) 5 mL DEAE Sepharose column (HiTrap™ DEAE FF; GE Healthcare), connected to an ÄKTA start FPLC system equipped with UV, pH and

conductance detectors at a flow rate of 1 CV/min. The column was washed with 5 - 10 CV of 20 mM NaCl in the corresponding buffer and bound proteins were eluted with a 10 CV salt gradient of 20 - 300 mM NaCl in 50 mM MOPS buffer containing 10 mM DTT. The eluates were collected in 1 CV fractions and combined with 4* Laemmli buffer. All were resolved in a 10% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue G-250) was used for protein analysis. Fractions were combined according to purity and concentrated via an Amicon Ultra spin filter with a 10 kDa cutoff (Merck). Next, the protein was diluted in 50 mM NaHCO$_3$ pH 8.0, 10 mM DTT.

Native Ova and native Ova-Aha - Q Sepharose HP IEX

The buffer exchanged eluate from the described DEAE Sepharose FF IEX column was loaded onto a prewashed (5 CV MQ water, 5 CV high salt 50 mM NaHCO$_3$ buffer and 10 CV low salt 50 mM NaHCO$_3$ buffer) 5 mL Q Sepharose column (HiTrap™ Q HP; GE Healthcare), connected to an ÄKTA start FPLC system equipped with UV, pH and conductance detectors at a flow rate of 1 CV/min. The column was washed with 5-10 CV of 20 mM NaCl in the corresponding buffer and bound proteins were eluted with a 20 CV gradient 40 - 150 mM NaCl in 50 mM NaHCO$_3$ pH 8.0 containing 10 mM DTT. The eluates were collected in 1 CV fractions and subsequently treated as described for "Native Ova and native Ova-Aha - DEAE Sepharose FF IEX". With the exception that the protein fraction was diluted in 50 mM Tris-HCl pH 8.0, 10 mM DTT.

Subsequently, the protein was loaded as described for the purification in 50 mM NaHCO$_3$ buffer. With the exception that the protein was eluted with a 30 CV gradient 40 - 150 mM NaCl in a 50 mM Tris-HCl pH 8.0 containing 10 mM DTT. The protein was buffer exchanged to 50 mM NaHCO$_3$ buffer, pH 8.0 for subsequent analysis.

Protein concentrations were measured using the mass extinction coefficient (ε = 31775; MW = 42911 Da) and protein fractions were combined by concentration.

Native Ova - Short overview of other tested conditions

The purification conditions as described in the following table resulted in a maximal purity of 80% as analyzed by SDS-PAGE. Some of the purifications were started by osmotic lysis in a low salt buffer, subsequently followed by the addition of lysozyme

to break down the bacterial cell wall. Next, an ammonium sulfate precipitation was performed at a concentration of 40%. Other purifications were performed by loading the lysed cells immediately onto the described protein column.

Table 2: Summary of other tested conditions for native Ova purification

Buffer used for column purification	(NH₄)₂(SO₄)	Column
50 mM Tris-HCl pH 7.5	x	Q
50 mM Tris-HCl pH 8.0	x	Q
50 mM NaHCO₃ pH 7.1	x	Q
50 mM NaHCO₃ pH 8.0	x	Q
50 mM NaHCO₃ pH 8.0, 10% glycerol, 2 mM DTT, 50 U benzonase, 1 mM PMSF, 1 mM EDTA		DEAE
50 mM HEPES pH 8.0, 10% glycerol, 2 mM DTT, 50 U benzonase, 1 mM PMSF, 1 mM EDTA		DEAE

Protein analysis

Western blot analysis

Samples were resolved in a 10 or a 12.5% SDS-PAGE (proteins at indicated concentrations) along with PageRuler™ Plus Protein Marker (Thermo Scientific) and transferred onto a PVDF membrane by Trans-Blot Turbo™ Transfer system directly after scanning. Membranes were washed with TBS and TBST and blocked with 5% nonfat dry milk (Elk; Campina) in TBST at rt for 1.5 h or at 4°C ON. Membranes containing protein were then incubated with primary antibody in 5% milk in TBST (1 h at rt) Membranes were washed 3x with TBST and incubated with matching secondary antibody in 5% milk in TBST (1 h at rt). Subsequently washed three times with TBST and once with TBS. Membranes were developed with luminol (10 mL of 1.4 mM luminol in 100 mM Tris, pH 8.8 + 100 μL of 6.7 mM p-coumaric acid in DMSO + 3 μL of 30% (v/v) H_2O_2) [60] and chemiluminescence was detected on the ChemiDoc™ MP System in the chemiluminescence channel and the protein marker was visualized with Cy3 and Cy5 settings.

Primary antibodies: polyclonal rabbit anti-Ova (1:2500, 1:5000 or 1:10000, LifeSpan BioSciences, LS-C59287), polyclonal rabbit anti-6x His epitope tag (1:1000, Rockland, 600-401-382)

Secondary antibody: mouse anti-rabbit IgG-HRP (1:5000, Santa Cruz, sc-2357).

Change of protein folding

Ova-Aha was diluted to 3.33 mg/mL and combined with 4* Laemmli sample buffer without β-mercaptoethanol, containing either 0, 2, 5, 10 or 20 mM DTT. The protein

fractions were resolved uncooked in a 10% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Subsequent Coomassie staining (Coomassie Brilliant Blue G-250) was used for protein analysis.

Fluorescent analysis of 6His-Ova-Aha and native Ova-Aha

6His-Ova-Aha and native Ova-Aha were characterized by ligating Alexa Fluor 647 (AF647; Invitrogen) alkyne to the azide click handle. This was done via CuAAC reaction, the protein was combined 2:1 (v/v) with click mix (containing 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand, 30 mM aminoguanidine-HCl and 14 µM AF647-alkyne in 88 mM HEPES pH 7.2, final concentration in click mix) and the reaction took place for 1 hour at rt in the dark. The reaction was quenched by the addition of 4* Laemmli buffer and subsequently samples were resolved in a 10% SDS-PAGE gel along with PageRuler™ Plus Protein Marker (Thermo Scientific), before scanning Cy3 and Cy5 multichannel settings (605/50 and 695/55 filters, respectively; ChemiDoc™ MP System, Bio-Rad). Coomassie staining (Coomassie Brilliant Blue G-250) was used for the confirmation of protein loading.

SEC-MALS

Purified ovalbumin variants were assessed for their molecular weight and multimerisation states using multi-angle light scattering (MALS) using the µDAWN detector and OptiLab RI detector (both Wyatt Instruments) in line with a Superdex200 10/300 gel filtration column (GE Healthcare). Molecular weights for the appropriate peaks were determined with the refractive index by using ASTRA software (Wyatt Instruments).

Circular Dichroism (CD) analysis

The secondary structure of ovalbumin was assessed using CD spectroscopy. Far UV-CD spectra were recorded using a Jasco J815 CD spectrometer equipped with a Jasco PTC 123 Peltier temperature controller (Easton, MD) between 190-260 nm. A minimum of five spectra with an acquisition time of 70 seconds (s) for each scan in a 1 mm quartz cuvette at 1 nm resolution were acquired at rt and averaged. Protein concentration used

for Ova was 0.2 mg/mL diluted in 50 mM $NaHCO_3$ pH 8.0. Analysis of the secondary structure was performed by BeStSel test version [61].

Intrinsic tryptophan fluorescence

Intrinsic tryptophan fluorescence emission spectra were recorded using a LS-55 Luminescence Spectrophotometer (Perkin Elmer). Protein samples were diluted to 50 µg/mL in 100 µL Tris-HCl, pH 8.5 or to 0.5 mg/mL in 75 µL $NaHCO_3$ pH 8.0. The protein was excited at 280 nm and spectra were recorded in the wavelength range of 290 - 400 nm with a 5 nm slit.

Thermostability measurements

Sample quality was assessed using Tycho (NanoTemper Technologies) in which intrinsic fluorescence was detected during a thermal ramp. The ratio between 350 nm and 330 nm was plotted and the inflection point (Ti) was determined as an indication of thermal stability.

Tryptic digest

Protein samples were diluted to 12.5 µg in 25 µL 50 mM $NaHCO_3$ pH 8.0. Each of the samples was supplemented with 1 mM $CaCl_2$. One of the series was heated for 15 minutes at 90°C to obtain denatured Ova. Each of the samples was subsequently treated with 0.25 µg (0.5 µL) trypsin or with the same amount of buffer and the samples were incubated for 10 min at 37°C, before adding a final concentration of 0.5 mM PMSF to quench trypsin activity. Samples were combined with 4* Laemmli sample buffer and resolved in a 15% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue G-250) was used for protein analysis.

Cell culture

General

A20 B cell was a kind gift of C. Watts and was tested on regular basis for mycoplasma contamination. Cultures were discarded after 2 months of use. The cells were cultured at 37°C under 5% CO_2 in RPMI 1640 (containing 25 mM HEPES) supplemented with

stable glutamine (2 mM), heat inactivated fetal calf serum (10% v/v; Biowest), β-mercaptoethanol (50 μM), penicillin (200 IU/mL; Duchefa) and streptomycin (200 μg/mL; Duchefa). Cells were passaged every 2 - 3 days.

DO11.10 was a kind gift of F. Ossendorp and was tested on regular basis for mycoplasma contamination. Cultures were discarded after 2 months of use. The cells were cultured at 37˚C under 5% CO_2 in IMDM supplemented with stable glutamine (2 mM), heat inactivated fetal calf serum (10% v/v; Biowest), β-mercaptoethanol (50 μM), penicillin (200 IU/mL; Duchefa) and streptomycin (200 μg/mL; Duchefa). Cells were passaged every 2 - 3 days.

Antigen presentation assay

A20s (50000 cells/well) were seeded in a 96-well tissue-culture treated microtiter plate. Adherence was allowed at 37˚C under 5% CO_2 for at least 1 h prior to the addition of the proteins at 0.25 mg/mL final concentration. The cells were incubated with the antigens for 4 hours, followed by the addition of DO11.10 T cell hybridoma (50000 cells/well) were added to the pulsed A20s and co-cultured for 20 h for antigen recognition and IL-2 production by the T cells at 37˚C under 5% CO_2. After overnight incubation, cells were sedimented by centrifugation (360 rcf, 5 min, rt) and supernatant was transferred to a new 96-wells plate. Stimulation of the T cell hybridoma was measured by IL-2 readout using an ELISA assay according to manufacturer's protocol (Invitrogen).

Graphical analysis

All analysis was determined using GraphPad Prism® 6 or 8 or Microsoft Excel 2016.

References

1. Dhanapala, P. de Silva, C., Doran, T. and Suphioglu, C. (2015). Cracking the egg: An insight into egg hypersensitivity. *Mol. Immunol. 66,* p.375-383

2. Burgdorf, S., Lukacs-Kornek, V. and Kurts, C. (2006). The mannose receptor mediates uptake of soluble but not of cell-associated antigen for cross-presentation. *J. Immunol. 176,* p.6770-6776

3. Huntington, J.A. and Stein, P.E. (2001). Structure and properties of ovalbumin. *J. Chromatogr. B. Biomed. Sci. Appl. 756,* p.189-198

4. Takahashi, N., Koseki, T., Doi, E. and Hirose, M. (1991). Role of an intrachain disulphide bond in the conformation and stability of ovalbumin. *J. Biochem. 109,* p.846-851

5. Takahashi, N. and Hirose, M. (1992). Reversible denaturation of disulfde-reduced ovalbumin and its reoxidation generating the native cystine cross-link. *J. Biol. Chem. 267,* p. 11565-11572

6. Harvey, D.J., Wing, D.R., Küster, B. and Wilson, I.B.H. (2000). Composition of N-linked carbohydrates from ovalbumin and co-purified glycoproteins. *J. Am. Soc. Mass Spectr. 11,* p.564-571

7. Suzuki, T., Kitajima, K., Emore, Y., Inoue, Y. and Inoue, S. (1997). Site-specific de-N-glycosylation of diglycosylated ovalbumin in hen oviduct by endogenous peptide: N-glycosidase as a quality control system for newly synthesized proteins. *PNAS. 94,* p.6244-6249

8. Kato, Y., Iwase, H. and Hotta, K. (1986). Characterization of a highly glycosylated biosynthetic intermediate of ovalbumin. *Arch. Biochem. Biophys. 244,* p.408-412

9. Glabe, C.G., Hanover, J.A. and Lennarz, W.J. (1980). Glycosylation of ovalbumin nascent chains. *J. Biol. Chem. 255,* p.9236-9242

10. Miyamoto, T., Takahashi, N., Sekine, M., Ogawa, T., Hidaka, M., Homma, H. and Masaki, H. (2015). Transition of serine residues to the D-form during the conversion of ovalbumin into heat stable S-ovalbumin. *J. Pharmaceut. Biomed. 116,* p.145-149

11. Wolfert, M.A. and Boons, G-J. (2013). Adaptive immune activation: Glycosylation does matter. *Nat. Chem. Biol. 9,* p.776-784

12. Delamarre, L., Pack, M., Chang, H., Mellman, I. and Trombetta, E.S. (2005). Differential lysosomal proteolysis in antigen-presenting cells determines antigen fate. *Science. 307,* p.1630-1634

13. Van Montfoort, N., Camps, M.G., Khan, S., Filippov, D.V., Weterings, J.J., Griffith, J.M., Geuze, H.J., van Hall, T., Verbeek, J.S., Melief, C.J. and Ossendorp, F. (2009). Antigen storage compartments in mature dendritic cells facilitate prolonged cytotoxic T lymphocyte cross-priming capacity. *PNAS. 106,* p.6730-6735

14. Fear, V.S., Burchell, J.T., Lai, S.P., Wikstrom, M.E., Blank, F., von Garnier, C., Turner, D.J., Sly, P.D., Holt, P.G., Strickland, D.S. and Stumbles, P.A. (2011). Restricted aeroallergen access to airway mucosal dendritic cells in vivo limits allergen-specific

CD4$^+$ T cell proliferation during the induction of inhalation tolerance. *J. Immunol. 187,* p.4561-4570

15. Wahid, R., Cannon, M.J. and Chow, M. (2005). Dendritic cells and macrophages are productively infected by poliovirus. *J. Virol. 79,* p.401-409

16. Araman, M.C., Pieper-Pournara, L., van Leeuwen, T., Kampstra, A.S.B., Bakkum, T., Marqvorsen, M.H.S., Nascimento, C.R., Groenewold, G.J.M., Wulp, W., Camps, M.G.M., Overkleeft, H.S., Ossendorp, F.A., Toes, R.E.M. and van Kasteren, S.I. (2019). Bioorthogonal antigens allow the unbiased study of antigen processing and presentation. *BioRxiv.* p.439323

17. Ito, K. and Matsudomi, N. (2005). Structural characteristics of hen egg ovalbumin expressed in yeast Pichia pastoris. *Biosci. Biotechnol. Biochem. 69,* p.755-761

18. Takahashi, N., Orita, T. and Hirose, M. (1995). Production of chicken ovalbumin in *Escherichia coli. Gene. 161,* p.211-216

19. Takahashi, N., Onda, M., Hyashi, K., Yamasaki, M., Mita, T. and Hirose, M. (2005). Thermostability of refolded ovalbumin and S-ovalbumin. *Biosci. Biotechnol. Biochem. 69,* p.922-931

20. Upadhyay, V., Singh, A. and Panda, A.K. (2016). Purification of recombinant ovalbumin from inclusion bodies of *Escherichia coli. Protein Expr. Purif. 117,* p. 52-58

21. Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A. and Schuman, E.M. (2006). Selective identification of newly synthesized proteins in mammalian cells using biorthogonal noncanonical amino acid tagging (BONCAT). *PNAS, 103,* p.9482-9487

22. van Hest, J.C.M., Kiick, K.L. and Tirrell, D.A. (2000). Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *J. Am. Chem. Soc. 122,* p.1282-1288

23. Karttunen, J., Sanderson, S. and Shastri, N. (1992). Detection of rare antigen-presenting cells by the lacZ T-cell activation assay suggests an expression cloning strategy for T-cell antigens. *PNAS. 89,* p.6020-6024

24. Barnden, M.J., Allison, J., Heath, W.R. and Carbone, F.R. (1998). Defective TCR expression in transgenic mice constructed using cDNA-based α- and ß-chain genes under the control of heterologous regulatory elements. *Immunol. Cell Biol. 76,* p.34-40

25. Hogquist, K.A., Jameson, S.C., Health, W.R., Howard, J.L., Bevan, M.J. and Carbone, F.R. (1994). T cell receptor antagonist peptides induce positive selection. *Cell. 76,* p.17-27

26. White, J., Haskins, K.M., Marrack, P. and Kappler, J. (1983). Use of I region-restricted, antigen-specific T cell hybridomas to produe idiotypically specific anti-receptor antibodies. *J. Immunol. 130,* p.1033-1037

27. Young, T.S. and Schultz, P.G. (2010). Beyond the canonical 20 amino acids: Expanding the genetic lexicon. *J. Biol. Chem. 285,* p.11039-11044

28. Kiick, K.L., Saxon, E., Tirrell, D.A. and Bertozzi, C.R. (2002). Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *PNAS. 99,* p.19-24

29. Del Cid, N., Shen, L., Bellelse, J. and Raghavan, M. (2012). Assessment of roles for calreticulin in the cross-presentation of soluble and bead-associated antigens. *PLoS One. 7,* e.41727

30. Kapust, R.B. and Waugh, D.S. (2000). Controlled intracellular processing of fusion proteins by TEV protease. *Protein Expr. Purif. 19,* p.312-318

31. Wood, W.B. (1966). Host specificity of DNA produced by *Escherichia coli*: Bacterial mutations affecting the restriction and modification of DNA. *J. Mol. Biol. 16,* p.118-133

32. Novy, R. and Morris, B. Use of glucose to control basal expression in the pET system. *inNovations. 13,* p.8-10

33. Mohanty, A.K. and Wiener, M.C. (2004). Membrane protein expression and purification: Effects of polyhistidine tag length and position. *Protein Expr. Purif. 33,* p.311-325

34. (2003). Rosetta™ 2(DE3) Competent cells for enhanced coverage of codon bias in *E. coli. inNovations. 18,* 28

35. Sohi, M., Alexandrovich, A., Moolenaar, G., Visse, R., Goosen, N., Vernede, X., Fontecilla-Camps, J.C., Champness, J. and Sanderson, M.R. (2000). Crystal structure of *Escherichia coli* UvrB C-terminal domain, and a model for UvrB-uvrC interaction. *FEBS Lett. 465,* p.161-164

36. Costa, S., Almeida, A., Castro, A. and Domingues, L. (2014). Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: The novel Fh8 system. *Front. Microbiol. 5,* 63

37. Esposito, D. and Chatterjee, D.K. (2006). Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol. 17,* p.353-358

38. Kapust, R.B. and Waugh, D.S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci. 8,* p.1668-1674

39. Nygren, P.A. Ståhl, S. and Uhlén, M. (1994). Engineering proteins to facilitate bioprocessing. *Trends Biotechnol. 12,* p.184-188

40. Pryor, K.D. and Leiting, B. (1997). High-level expression of soluble protein in *Escherichia coli* using a His6-tag and maltose-binding-protein double affinity fusion system. *Protein Expr. Purif. 10,* p.309-319

41. Castellanos-Mendoza, A., Castro-Acosta, R.M., Olvera, A., Zavala, G., Mendoza-Vera, M., García-Hernándeze, E., Alagón, A., Trujillo-Roldán, M.A. and Valdez-Cruz, N.A. (2014). Influence of pH control in the formation of inclusion bodies during production of recombinant sphingomyelinase-D in *Escherichia coli. Microb. Cell Fact. 13,* 137

42. Neubauer, P. and Hofmann, K. (1994). Efficient use of lactose for the lac promoter-controlled overexpression of the main antigenic protein of the foot and mouth disease virus in *Escherichia coli* under fed-batch fermentation conditions. *FEMS Microbiol. Rev. 14,* p.99-102

43. Wurm, D.J., Veiter, L., Ulonska, S., Eggenreich, B., Herwig, C. and Spadiut, O. (2016). The *E. coli* pET expression system revisited-mechanistic correlation between glucose and lactose uptake. *Appl. Microbiol. Biotechnol. 100,* p.8721-8729

44. San-Miguel, T., Pérez-Bermúdez, P. and Gavidia, I. (2013). Production of soluble eukaryotic recombinant proteins in *E. coli* is favoured in early log-phase cultures induced at low temperature. *Springerplus, 2,* 89

45. de Groot, N.S. and Ventura, S. (2006). Effect of temperature on protein quality in bacterial inclusion bodies. *FEBS Lett. 580,* p.6471-6476

46. Singh, A., Upadhyay, V., Upadhyay, A.K., Singh, S.M. and Panda, A.K. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilisation process. *Microb. Cell Fact. 14,* 41

47. Jungbauer, A. and Kaar, W. (2007). Current status of technical protein refolding. *J. Biotechnol. 128,* p.587-596

48. Singh, S.M. and Panda, A.K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng. 99,* p.303-310

49. Palmer, I. and Wingfield, P.T. (2004). Preparation and extraction of insoluble (inclusion-body) proteins from *Escherichia coli. Curr. Protoc. Protein Sci.* Chapter: Unit-6.3

50. Booth, W.T., Schlachter, C.R., Pote, S., Ussin, N., Mank, N.J., Klapper, V., Offermann, L.R., Tang, C., Hurlburt, B.K. and Chruszcz, M. (2018). Impact of an N-terminal polyhistidine tag on protein thermal stability. *ACS Omega. 3,* p.760-768

51. Fink, A.L. (1998). Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Fold. Des. 3,* p.9-23

52. Islam, M.T. (2017). Radiation interactions with biological systems. *Int. J. Radiat. Biol. 93,* p.487-493

53. Duong-Ly, K.C. and Gabelli, S.B. (2014). Salting out of proteins using ammonium sulfate precipitation. *Methods Enzymol. 541,* p.85-94

54. Ghisaidoobe, A.B.T. and Chung, S.J. (2014). Intrinsic tryptophan fluorescence in the detection and analysis of proteins: A focus on förster resonance energy transfer techniques. *Int. J. Mol. Sci. 15,* p.22518-22538

55. Onda, M., Tatsumi, E., Takahashi, N. and Hirose, M. (1997). Refolding process of ovalbumin from urea-denatured state. Evidence for the involvement of nonproductive side chain interactions in an early intermediate. *J. Biol. Chem. 272,* p.3973-3979

56. Ishimaru, T., Ito, K., Tanaka, M. and Matsudomi, N. (2010). Thermostabilization of ovalbumin by alkaline treatment: Examination of the possible roles of D-serine residues. *Protein Sci. 19,* p.1205-1212

57. Duy, C. and Fitter, J. (2006). How aggregation and conformational scrambling of unfolded states govern fluorescence emission spectra. *Biophys. J. 90,* p.3704-3711

58. van Elsland, D.M., Pujals, S., Bakkum, T., Bos, E., Oikonomeas-Koppasis, N., Berlin, I., Neefjes, J., Meijer, A.H., Koster, A.J., Albertazzi, L. and van Kasteren S.I. (2018). Ultrastructural imaging of salmonella-host interactions using super-resolution correlative light-electron microscopy of bioorthogonal proteins. *Chembiochem. 19,* p.1766-1770

59. Link, A.J. and Tirrell, D.A. (2005). Reassignment of sense codons in vivo. *Methods. 36,* p.291-298

60. Mruk, D.D. and Cheng, C.Y. (2011). Enhanced chemiluminescence (ECL) for routine immunoblotting: An inexpensive alternative to commercially available kits. *Spermatogenesis. 1,* p.121-122

61. Micsonai, A., Wien, F., Kernya, L., Lee, Y.H., Goto, Y. Réfrégiers, M. and Kardos, J. (2015). Accurate secondary structure prediction and folEd recognition for circular dichroism spectroscopy. *PNAS. 112,* e.3095-3103

Chapter 3

T cell activation by bioorthogonally
labelled tetanus toxin C fragment

Abstract

In this Chapter the bioorthogonal expression and initial immunological evaluation of the antigen tetanus toxin C fragment was assessed. A mutant library of the protein was made that contained a bioorthogonal amino acid at every position in the known MHC-II restricted epitopes. The effect of these mutations on antigen processing and T cell activation was analyzed.

Introduction

Antigen presentation plays a vital role in the T cell immune response. To induce this response, proteins are cleaved into peptide fragments, which are presented on either the major histocompatibility complex (MHC) class I or MHC class II [1, 2]. Peptides presented via MHC class II are presented to CD4+ T helper cells [3]. These particular T cells are involved in immunity and inflammation and, depending on their cytokine production, can be classified into different subtypes [4, 5], such as Th1, Th2, Treg and Th17. Their activation is therefore pivotal for many branches of the adaptive immune response [6-11].

To obtain an MHC class II antigen induced T cell responses, peptides arising from exogenous proteins need to be internalized by professional antigen presenting cells (APCs) – expressing the MHC class II molecules – via for instance endocytosis [12]. After this, proteins are enclosed in endosomes and processed into peptide fragments, which will encounter the MHC class II molecules. These are concurrently transported to the MHC class II loading compartment (MIIC) (Figure 1A - step 1 and 2), where the peptides can be loaded onto the open-ended groove of MHC class II molecules. MHC class II molecules bind a core epitope of nine amino acid residues – with strong anchor points at p1, p4, p6 and p9 –, but can hold peptides of ~13-25 residues [13-15]. Next, the complex is transported to the plasma membrane (Figure 1A - step 3), where the epitope is recognized by a T cell receptor with a possible antigenic reaction of the T cell as the result [2, 16].

As described above, antigens are taken up by a professional APC, processed and subsequently loaded onto a newly synthesized MHC class II molecule [17]. This, however, is the route that the majority of the proteins will follow [18]. In some cases, the antigenic protein can bind directly – or after partially unfolding – to these molecules [18, 19]. The mainly intact proteins will not bind to the newly synthesized MHC class II molecules, but will bind instead to the ones recycled from the cell surface and/or from instable formed complexes [18, 20, 21].

Some of the particulars of MHC class II antigen processing remain unknown. For example, it is known that peptide cleavage plays an important role in T helper cell activation, but the order of cleavage/binding is still under debate: does the protein bind
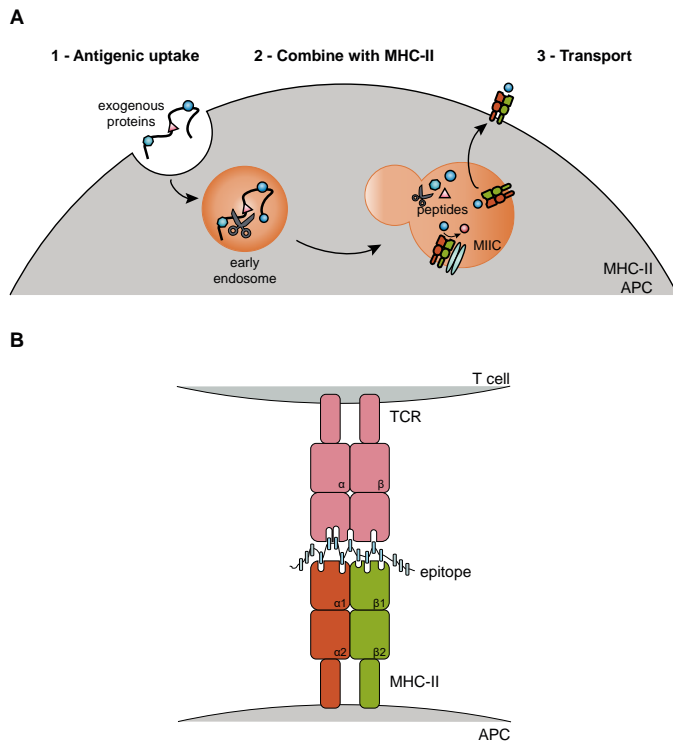
Figure 1. Schematic overview of MHC class II antigen presentation. A) Schematic overview of the uptake of an antigen, which is cleaved by peptides before being loaded in the MIIC complex on MHC class II molecules. The combined MHC class II peptide complex is transported to the plasma membrane. B) At the plasma membrane a T cell can recognize the epitope and will become activated.

the MHC class II molecule first and is trimmed afterwards? Or is the protein processed first and the liberated peptides then loaded (as depicted in Figure 1A) [22]. A recent study of Roche *et al.* shows that the order of processing and loading may be dependent on the activation-status of the APC [23]. A second parameter that remains enigmatic is that the *in vitro* binding affinity of both the peptide for the MHC class II, and that between the T-cell receptor and the peptide-MHC-II complex does not correlate to immunogenicity [24]. This suggests that other factors – either following T cell recognition or that take place during peptide loading are essential determinants for immunogenicity. For instance, the open-ended nature of the peptide binding groove of MHC class II makes it possible that overhanging amino acids in the flanking region of the core epitope are involved, and that instead of a single peptide epitope an array of peptides differing in the length/nature of the overhang are presented [25].

To study MHC class II antigen processing and presentation, model proteins are often used. One such model protein is tetanus toxin C fragment (TTCF; residues 865-1315 of tetanus toxin) [26, 27], which has been earlier used to gain knowledge about the adaptive immune response [25, 27]. Presentation of this protein only requires limited processing by asparagine-specific endopeptidase (AEP) [19, 28, 29]. Antigen presentation can be traced using fluorescent antigen variants to correlate a visual image to the overall immune outcome. However, it has been shown that this strongly reduces antigen presentation for ovalbumin [30].

To prevent this problem, it was postulated that a small ligation handle – such as an azide or an alkyne moiety – inside the TTCF epitope would result in presentation of an antigen closer in structure to the wild type one, yet allowing detection of the antigen using click chemistry (see **Chapter 1**) – a postulate that was investigated in research described in this Chapter. For this, a strategy developed by Tirrell and co-workers was selected, where a model protein is expressed with non-canonical amino acids, allowing for i.e. the bioorthogonal non-canonical amino acid tagging strategy (BONCAT) [31]. In this method, an amino acid auxotrophic bacterial strain is used for protein expression. Furthermore, the canonical methionine is removed from culture media and supplemented with the isosteric azidohomoalanine [32] instead using the natural methionine tRNA synthetase for incorporation [33].

In this Chapter, it is tested whether or not it is possible to create a protein in which an amino acid in the epitope of interest is replaced with an L-azidohomoalanine (L-Aha) residue to allow the tracking of the epitope all the way from uptake to presentation on MHC-II. Furthermore, it is investigated if these mutated proteins will still be processed in an APC [30] and are able to induce T cell activation. This Chapter will conclude with general observations on the obtained results and will end with recommendations for future experiments.

Results

The aim of this Chapter is to follow antigen processing from uptake to MHC class II presentation and subsequent T cell activation using a family of bioorthogonal variants
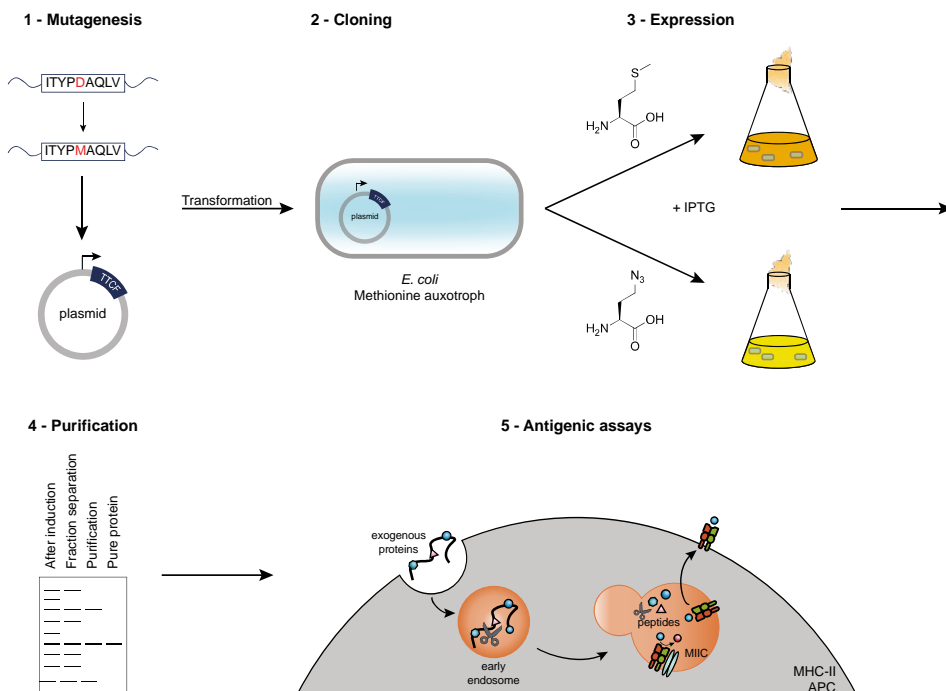
Figure 2. Schematic overview of the workflow. Experimental setup describing the mutagenesis of *TTCF* of which the resulting plasmid was transformed into the *E. coli* auxotroph B834(DE3). The protein was expressed in either LB medium or in SelenoMet medium to incorporate L-Met or L-Aha respectively. The protein was purified, before using the resulting TTCF (mutant) proteins in antigen presentation assays.

of TTCF (Figure 2), where each position in the predicted core epitope was mutated to methionine (L-Met) to allow expression with L-azidohomoalanine at these sites.

2.1 Epitope mutation and protein expression

Multiple T cell hybridomas, specific for different epitopes of the model protein TTCF, have been reported by the group of C. Watts [27]. In this Chapter the 2F2 hybridoma was used [27], which recognises a peptide spanning the residues 900 - 916 of tetanus toxin (SGFNSSVITYPDAQLVP; the predicted core epitope is underlined) (PDB: 1A8D; Figure 3A).

Core epitope mutants were therefore made from the previously reported pEH101 plasmid containing *TTCF* C-terminally ligated to a ten-histidine (His)-tag, kindly provided by C. Watts [34]. This resulted in the generation of nine mutant genes, each of
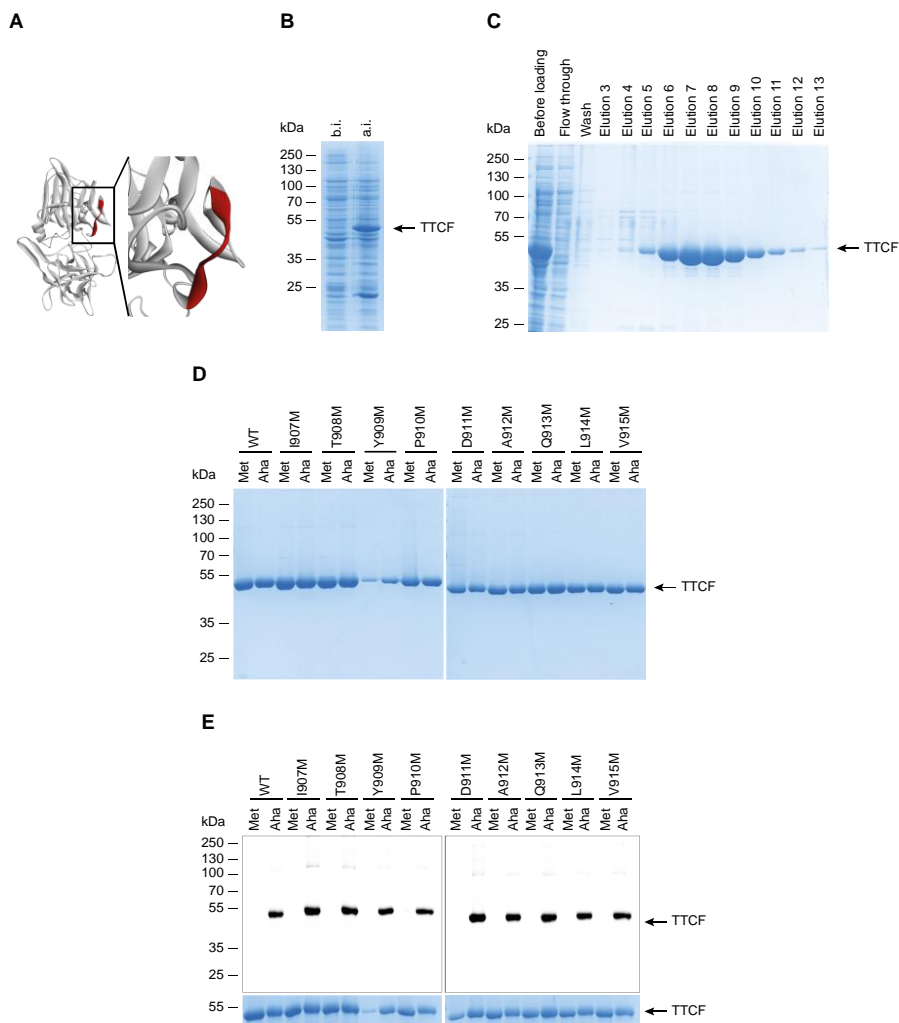
Figure 3. The use of the 2F2 epitope and its mutants. A) Crystal structure of TTCF (PDB: 1A8D) with in red the position of the epitope in the protein; visualized in Discovery Studio. B) Representative picture of the before and after induction samples of TTCF Aha. C) Representative picture of the purification of TTCF Aha. D) The purified proteins, all with a purity of >98% as analyzed by SDS-PAGE. E) The resulting azide-containing proteins were ligated to AF647 alkyne (upper panel) with as control the coomassie staining of all the proteins. B-E) Proteins were resolved in a 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie Brilliant Blue G-250.

which was cloned into the multiple cloning site of vector pET16b and transformed into *E. coli* B834(DE3) [35]. Subsequent expression in LB medium gave overexpression of a protein as revealed by the emergence of a band at 53 kDa after SDS-PAGE. This protein

Table 1: List of expressed proteins

Mutation	Met incorporation	Aha incorporation
-	ITAPDAQLV	ITAPDAQLV
I907M	MTAPDAQLV	AhaTAPDAQLV
T908M	IMAPDAQLV	IAhaAPDAQLV
Y909M	ITMPDAQLV	ITAhaPDAQLV
P910M	ITAMDAQLV	ITAAhaDAQLV
D911M	ITAPMQLV	ITAPAhaQLV
A912M	ITAPDMQLV	ITAPDAhaQLV
Q913M	ITAPDAMLV	ITAPDAAhaLV
L914M	ITAPDAQMV	ITAPDAQAhaV
V915M	ITAPDAQLM	ITAPDAQLAha

proved present in the after induction sample (Figure 3B), and could be purified using the 10His-tag with Ni^{2+} affinity chromatography [34] (Figure 3C).

Expression with the bioorthogonal amino acid L-Aha was next attempted. For this the cells were grown in SelenoMet-medium that was depleted of L-Met and augmented with L-Aha, thus conditions that are expected to yield proteins in which all (or most) Met-residues are expected to be replaced by this azide-containing counterpart [36]. The cells were incubated overnight after IPTG induction and all mutants showed – as displayed by SDS-PAGE analysis – the appearance of a robust band for overexpression at the predicted 53 kDa. Purification as before yielded the protein with >98% purity (Figure 3D) as determined by SDS-PAGE. The thus obtained twenty TTCF and mutant TTCF proteins, containing either eight (wild type) or nine internal (mutants) L-Met or L-Aha residues are summarized in Table 1.

2.2 TTCF-Aha ligations

It was next determined whether these proteins were reactive under a Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) also called a Cu(I)-catalyzed Huisgen cycloaddition conditions (ccHc) towards Alexa Fluor 647 (AF647) alkyne. Separation by SDS-PAGE of the thus reacted proteins and subsequent in-gel fluorescence showed that all the azidylated proteins had reacted with AF647 alkyne, whereas their methionine containing counterparts had not (Figure 3E).

2.3 TTCF degradation

Next, processing of the modified TTCF proteins by antigen presenting cells (APCs) was assessed [30]. For this bone marrow derived dendritic cells (BMDCs) were pulsed with the antigens and chased for 0, 1, 4 or 24 hours. Afterwards, the cells were
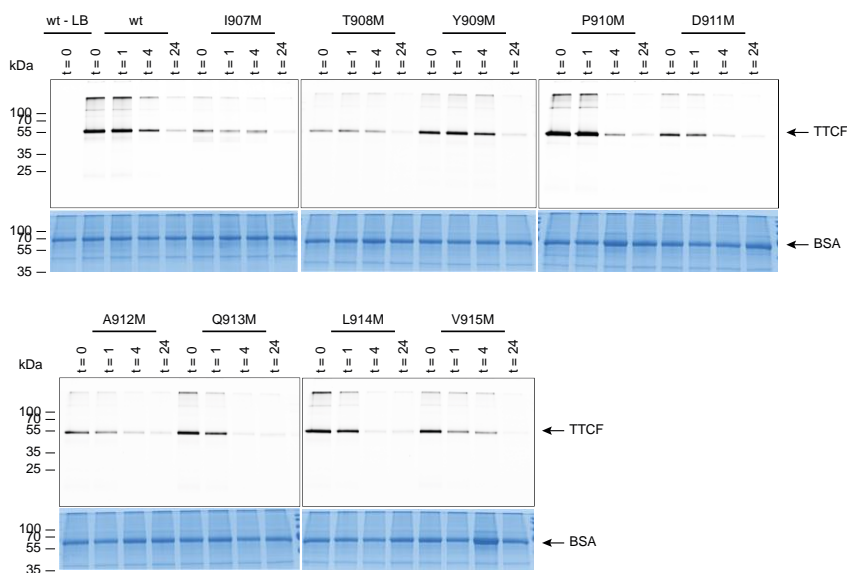
Figure 4. Degradation assay of TTCF by BMDCs. The different azide-containing proteins were used to pulse BMDCs for 1 hour, after which a chase of 0, 1, 4 or 24 hours took place. Cells were lysed together with the wild type protein containing L-Met as negative control, each of the protein samples showed degradation over time as visualized after ligation to AF647 alkyne. Proteins were resolved in a 10% SDS-PAGE and stained with Coomassie Brilliant Blue G-250.

washed, lysed and subjected to the above mentioned CuAAC prior to SDS-PAGE separation and in-gel fluorescence analysis (Figure 4). From these experiments it appeared that after 4 hours the level of intact azidylated protein was reduced by >25% and >50% after a 24 hours chase (Figure 4). For some of the mutants, a lower uptake was observed. Surprisingly, all the mutants degraded over time with no observation of intermediates, a process also observed by Araman *et al.* [37]. This could either be because the proteins are rapidly degraded after initial cleavage or because the azide handle is reduced over time. The latter was disproved in a previous study by Bakkum *et al.* [38].

Protein degradation within the first hour of pulse is thought to be related to destructive antigen processing [36]. This is associated with rapid release of the chain-derived class II-associated invariant chain peptide (CLIP), whereby the protein can be loaded into the MHC class II molecule on the cell surface or in the early endosome [39, 40]. This process could lead to different epitopes, a process which could lead to
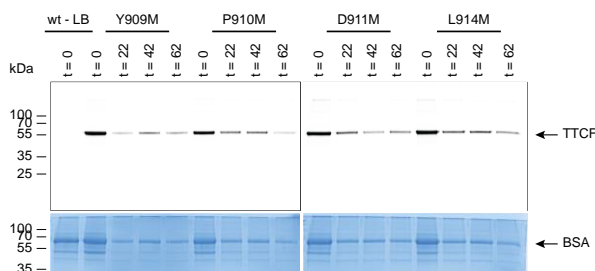
Figure 5. Pulse assay of TTCF to BMDCs. The randomly chosen azide-containing proteins were used to pulse BMDCs for 1 hour. Cells were lysed together with the wild type protein containing L-Met as negative control, none of the samples showed degradation over time as visualized after ligation to AF647 alkyne. Proteins were resolved in a 10% SDS-PAGE and stained with Coomassie Brilliant Blue G-250.

suboptimal T cell presentation. To analyze the rate of protein degradation during the first hour of antigen processing, the cells were pulsed with selected mutant proteins for 22, 42 and 62 minutes. The cells were washed, lysed and subjected to a ligation mixture containing AF647 alkyne as described for the antigen degradation experiment. The normalized in-gel fluorescence showed no decrease in fluorescent signal at these shorter pulse times (Figure 5).

2.4 T cell activation by Aha-epitopes

It was next determined whether the in-epitope mutations affected activation of the 2F2-hybridoma. BMDCs were therefore pulsed for 4 hours with each of the TTCF proteins (0.1 mg/mL final concentration), washed twice with PBS and incubated with 2F2 overnight. T cell activation – as analyzed via an IL-2 ELISA – showed that mutating each of the amino acid residues individually to methionine in this ITYPDAQLV epitope (as depicted in Table 1) still resulted in T cell activation (Figure 6A). In addition, when this assay was performed using the proteins containing L-Aha residues, IL-2 production was observed (Figure 6B). To more precisely study this phenomenon, presentation efficiency of a minimal synthetic epitope modified with a bioorthogonal group was analyzed. No antigen presentation was observed for this 2F2 epitope.

In order to assess the effect of modification in the peptides with bioorthogonal groups on only the presentation, a library of peptides based on the well-studied OT-II epitope (ISQAVHAAHAEINEAGR (Figure 7A)) of ovalbumin was synthesized. Different
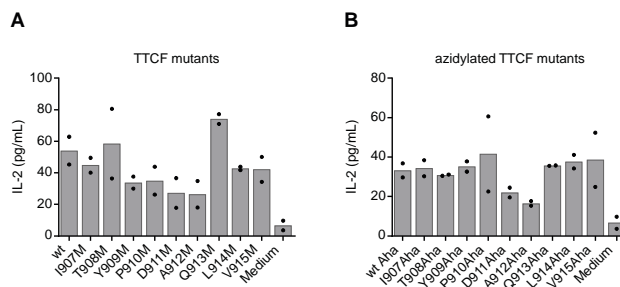
Figure 6. Antigen presentation of TTCF epitope mutants by BMDCs to 2F2 T cell hybridomas. A) BMDCs were pulsed for 4 hours with either wild type TTCF or a mutant in which a point mutation to methionine was made in the epitope. Followed by 20 hours presentation to the T cells (N = 2, n = 3). B) As for A), but containing an azidohomoalanine residue at each of the methionine positions (N = 2, n = 3).

residues of the predicted core epitope – AAHAEINEA – were mutated to L-Aha [41]. The immunogenic properties of the wild type and resulting mutant peptides were evaluated with an antigen presentation assay, using A20 and the Ova MHC-II specific T cell line DO11.10 [42].

In order to obtain antigen presentation, the I-Ad-positive B cell line A20 was pulsed for 4 hours with each of the long ova peptides prior to the addition of the DO11.10 cells for a 20 hour chase period. T cell activation was subsequently analyzed via an IL-2 ELISA assay. This showed that peptides containing mutations at positions 1 and 8 still activated the T cells, whereas antigen presentation assays with mutations at positions 2, 3, 5 and 7 did not (Figure 7C). In line with literature data, mutating position 5 – facing towards the TCR – did not show any response. Both positions capable of inducing T cell responses after mutation are positioned towards the T cell receptor (TCR) [41]. With mutant position 8 giving a higher initial response, it was decided to continue with this peptide for future ligation reactions.

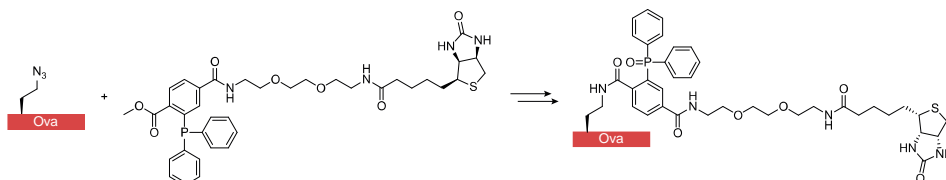To determine to what extent chemical modifications are allowed within the epitope, it was decided to react the azidylated p8 peptide (ISQAVHAAHAEINAhaAGR) with a biotin-phosphine (ISQAVHAAHAEINbiotinAGR) and the resulting peptide was used for antigen presentation studies (Figure 7B). This showed a decrease of approximately 50% activation as compared to the azidylated mutant peptide p8 (Figure 7D).
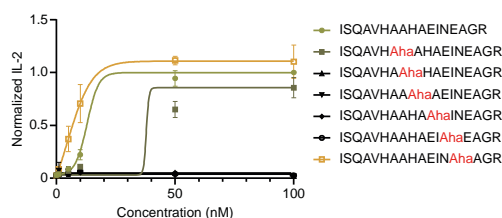
A

amino acid #	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339
amino acid	I	S	Q	A	V	H	A	A	H	A	E	I	N	E	A	G	R
position in MHC II							p1	p2	p3	p4	p5	p6	p7	p8	p9		

B

C

- ISQAVHAAHAEINEAGR
- ISQAVH**Aha**AHAEINEAGR
- ISQAVHA**Aha**HAEINEAGR
- ISQAVHAA**Aha**EINEAGR
- ISQAVHAAHA**Aha**INEAGR
- ISQAVHAAHAEI**Aha**EAGR
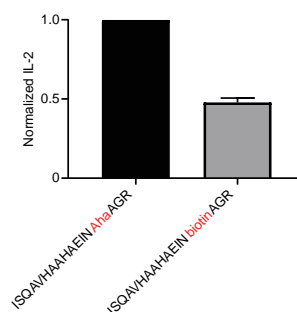- ISQAVHAAHAEIN**Aha**AGR

D

Figure 7. Antigen presentation of Ova peptides by A20 to DO11.10. A) Amino acid sequence of the ovalbumin MHC class II ISQAV long peptide. B) Schematic reaction of azidylated ovalbumin with biotin-phosphine. C) A20s were pulsed for 4 hours with either a wild type or an azidylated mutant peptide, which was followed by 20 hours presentation to the T cells (N = 3, n = 2/3). D) As for C), but using a mutant peptide and its biotinylated counterpart (N = 2, n = 2/3). All T cell activation studies were analyzed by ELISA IL-2 readout

Next, it was attempted to use ISQAVHAAHAEINAhaAGR in an immunoprecipitation assay [43, 44]. For this, A20s were pulsed for 24 hours with either the unmodified peptide, mutant p8 peptide or the p8 mutant peptide, pre-modified with biotin (using a Staudinger ligation reaction). After this chase period, the cell lysate containing the mutant p8 peptide was subjected to a Staudinger ligation using biotin-phosphine (2 hours at 37˚C). All cell lysates were then immunoprecipitated using streptavidin-coated beads. Subsequent Western blot analysis of the eluate showed no detectable MHC-II for the unmodified peptide and the pre-modified peptide, but did show signal when staining for the β-chain of MHC class II in the elution fractions (Figure 8). This confirms the suitability of the p8 modified antigens as a ligation reagent.
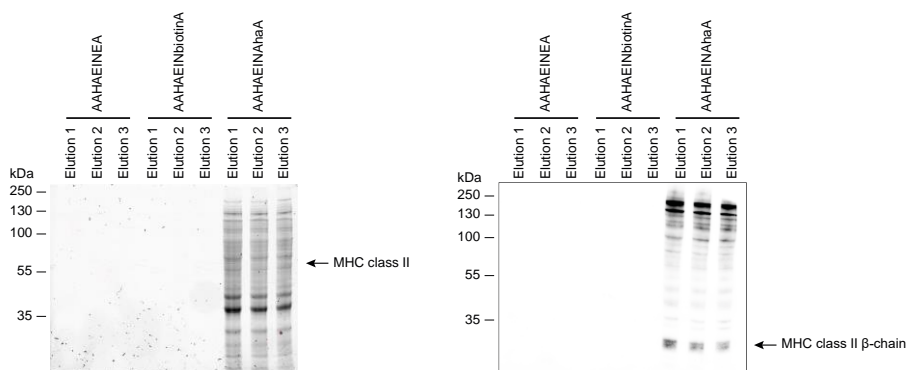
Figure 8. MHC class II immunoprecipitation. Left: 10% TGX Stain Free gel containing the elution fractions of the three used peptides: the wild type peptide (AAHAEINEA), the premodified peptide (AAHAEINbiotinA) and the azidylated peptide (AAHAENAhaA) subjected to a Staudinger ligation with biotin-phosphine in cell lysate. Right: The corresponding Western blot, using an MHC Class II (β-chain) antibody.

Discussion and conclusion

In this Chapter it was shown that it was possible to produce antigens with the incorporation of L-Aha, which could be followed in antigen degradation experiments. Furthermore, subsequent antigen presentation studies showed that these azidylated proteins still induced T cell activation. Due to issues with the IL-2 levels of the TTCF antigens and difficulties with expressing azidylated Ova proteins (in more details discussed in **Chapter 2**) it was chosen to perform experiments using both the TTCF proteins and Ova peptides.

One area in which these reagents could be of future interest would be in providing reagents that could block T cell activation (Figure 9A-C). As the azides can be ligated with bulky groups under physiological conditions, the ligation of a bulky group could be used as a stop-signal for T cell activation (as shown for the Ova p8 peptide). In a preliminary experiment, it was opted to use the incorporated azide handle in the TTCF epitope and ligate a bulky group to this position after antigen processing in order to prevent the T cell receptor to recognize this. However, an initial experiment in which the azides were ligated with either biotin-phosphine, bicyclononyne (BCN)-biotin, dibenzocyclooctyne (DBCO)-AF647, or reduced to an amine with tris(2-carboxyethyl)phosphine (TCEP) did not induce a detectable blockade of T cell recognition (Figure 9D). The reasons for this
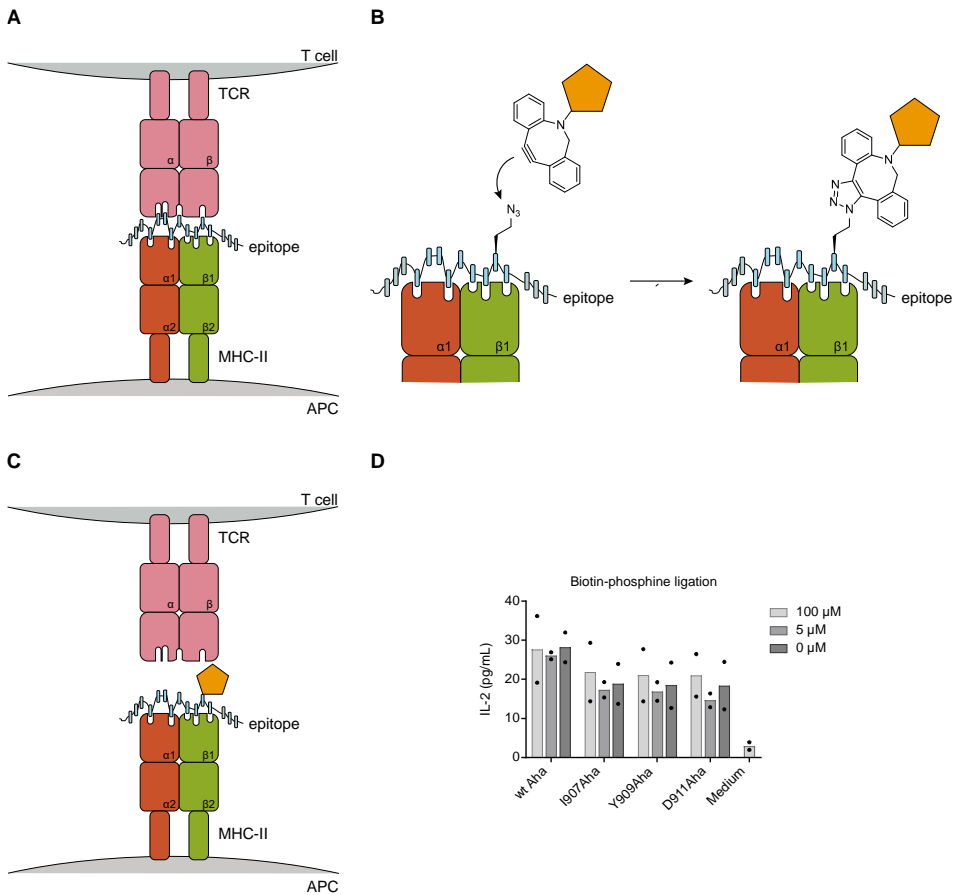
Figure 9. Antigen presentation of TTCF by BMDCs to 2F2 T cell hybridomas. A) At the plasma membrane a T cell can recognize the epitope and will become activated. This activation might be blocked by ligation to a B) bulky group on the presentation site. Which results in a block of C) recognition by the T cell receptor. D) BMDCs pulsed with different proteins, which were subjected to biotin-phosphine and subsequent overnight antigen presentation (N = 2, n = 3). T cell activation studies were analyzed by ELISA IL-2 readout.

could be either a broad tolerance at the tested positions, or inefficient ligation reactions. However, due to time constraints this could not be further investigated.

Acknowledgements

Materials and methods

General

All chemicals and reagents were purchased at Sigma-Aldrich, Alfa Aesar, Acros, Merck or VWR, unless stated otherwise. Reagents were used without further purification. SDS-PAGE, TGX Stain Free and Western blot materials were purchased at Bio-Rad. Cloning reagents were ordered at Thermo Fisher Scientific. DNA primers were ordered at Sigma-Aldrich or Integrated DNA Technologies. Protease inhibitors were obtained from Roche. Peptides were in-house synthesized containing a C-terminal amide and characterized by NMR and/or LC-MS. Cell culture disposables were from Sarstedt.

Solutions

PBS contained 5 mM KH_2PO_4, 15 mM Na_2HPO_4, 150 mM NaCl, pH 7.4 and PBST was PBS supplemented with 0.05% Tween-20. TBS contained 50 mM Tris-HCl, 150 mM NaCl and TBST was TBS supplemented with 0.05% Tween-20. Laemmli sample buffer 4* contained 60 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.01% (v/v) bromophenol blue.

Compound synthesis

Imidazole-1-sulfonyl azide hydrogen sulfate

As described by Potter *et al.* [45]. A 250 ml round-bottomed flask with a stirring bar was brought under nitrogen atmosphere and cooled to 0˚C. Sodium azide (5.2 g, 80 mmol, 1.0 eq.) was added before the addition of dry EtOAc (80 mL). After adding SO_2Cl_2 (6.5 ml, 80 mmol, 1.0 eq.) over 10 min the reaction was slowly warmed to rt and left stirring under an inert atmosphere overnight. The reaction mixture was then recooled to 0˚C, before the addition of imidazole (10.9 g, 160 mmol, 2.0 eq.) and the resulting suspension was stirred for 3.5 h at 0˚C. The reaction mixture was basified with saturated aqueous $NaHCO_3$ (150 mL). The organic layer was separated from the aqueous layer, washed with H_2O and dried with $MgSO_4$. The suspension was filtered and cooled to 0˚C. Concentrated H_2SO_4 (4.4 ml, 80 mmol, 1.0 eq.) was added over 5 min. The mixture was warmed to rt and stirred until a precipitate had formed, which was filtered and washed with EtOAc. The remaining white crystals were dried under vacuum to afford the title compound pure (12.8 g, 47 mmol, 59%).

Chapter 3

Boc-Dab-COOH

As described by More and Vince [46]. At 4°C the solution of *N*-Boc-L-glutamine (27.9 g, 113 mmol, 1.0 eq.) in THF (0.27 L) and water (68 mL) was combined with (Diacetoxyiodo)benzene (39 g, 136 mmol, 1.2 eq.). After 6.5 h of stirring, the mixture was evaporated to dryness, dissolved in H_2O (0.13 L) and washed four times with EtOAC (100 mL). The aqueous layer was evaporated to dryness and 3x co-evaporated with toluene and dried under vacuum to afford a pure compound (14.6 g, 67 mmol, 59%).

(S)-2-amino-4-azidobutanoic acid (L-azidohomoalanine)

Adjusted from Zhang *et al.* [47]. Boc-Dab-COOH (4.97 g, 23 mmol, 1.0 eq.) was suspended with K_2CO_3 (8.03 g, 58 mmol, 2.5 eq.) and $CuSO_4.5H_2O$ (59 mg, 0.23 mmol, 0.01 eq.) in MeOH (125 mL). To a stirred suspension, imidazole-1-sulfonyl azide hydrogen sulfate (6.23 g, 23 mmol, 1.0 eq.) was added and the mixture was stirred overnight at rt. Upon completion on TLC, the mixture was concentrated and diluted in EtOAC (100 mL). The organic layer was washed with 1% HCl (500 mL) twice and H_2O (500 mL) once. The crude was purified by column chromatography and fractions with over 90% purity as confirmed by NMR data were treated with 20% TFA in DCM for 4 h at rt. Upon completion on TLC, DCM was evaporated and the resulting oil was 2x co-evaporated with H_2O. The product was dissolved in MQ water and lyophilized 3 times to obtain the product (3.7 g, 14 mmol, 61%).

Tris-hydroxypropyltriazolylmethylamine (THPTA)

A two-step synthesis as described by Hong *et al.* [48].

Strains and plasmids

E. coli strains XL10 and B834(DE3) were used as cloning and expression strains. Plasmid pET16b was used for the expression of 10His-TTCF and its mutant variants and contained the IPTG-inducible T7 promoter.

Cloning

pEH101 containing full-length *TTCF* with a linker sequence was a gift from the Watts lab [18]. To obtain pET16b_10His-TTCF, the DNA fragment encoding TTCF was

Table 3: List of primer sequences

#	primer name	sequence 5' -> 3'	#	primer name	sequence 5' -> 3'
p1	TTCF_fwd	CGCAAAATTCCCTCTAGAATAATTTTTG	p11	TTCF_D911M_fwd	CATATCCAATGGCTCAATTGG
p2	TTCF_rev	GTTAGCAGCCGGATCCTTAGTCGTTGG	p12	TTCF_D911M_rev	CCAATTGAGCCATTGGATATG
p3	TTCF_I907M_fwd	CCTCTGTTATGACATATCC	p13	TTCF_A912M_fwd	CATATCCAGATATGCAATTGGTGCC
p4	TTCF_I907M_rev	GGATATGTCATAACAGAGG	p14	TTCF_A912M_rev	GGCACCAATTGCATATCTGGATATG
p5	TTCF_T908M_fwd	CTGTTATCATGTATCCAGATG	p15	TTCF_Q913M_fwd	CCAGATGCTATGTTGGTGCCGG
p6	TTCF_T908M_rev	CATCTGGATACATGATAACAG	p16	TTCF_Q913M_rev	CCGGCACCAACATAGCATCTGG
p7	TTCF_Y909M_fwd	GTTATCACAATGCCAGATGCTC	p17	TTCF_L914M_fwd	GATGCTCAAATGGTGCCGGGC
p8	TTCF_Y909M_rev	GAGCATCTGGCATTGTGATAAC	p18	TTCF_L914M_rev	GCCCGGCACCATTTGAGCATC
p9	TTCF_P910M_fwd	CACATATATGGATGCTCAATTGG	p19	TTCF_V915M_fwd	GTTGATGCCCGGCATCAATTGAG
p10	TTCF_P910M_rev	CCAATTGAGCATCCATATATGTG	p20	TTCF_V915M_rev	CTCAATTGATGCCGGGCATCAAC

amplified by PCR from the pEH101 plasmid. This fragment was ligated into the pET16b vector using the XbaI and BamHI restriction sites. All sequences were verified by Sanger sequencing (Macrogen).

Mutations of TTCF were created using site-directed mutagenesis. The resulting DNA fragment was ligated into the pET16b vector as described for the wild type gene.

Bacterial growth and protein expression

Expression of TTCF

An overnight culture of B834 containing pET16b_TTCF (earlier cloned from the pEH101 plasmid, kindly provided by prof. C. Watts) [34], or one of the mutants, was diluted 1:100 in LB medium containing 50 µg/mL ampicillin and 1% glucose. Cells were grown at 37˚C, 180 rpm to an OD_{600} of ~0.8 and sedimented (3428 rcf, 15 min, 4˚C) before being resuspended in LB medium containing 50 µg/mL ampicillin. The culture was induced with IPTG (1 mM final concentration) and expression took place ON at 30˚C, 130 rpm. After overnight expression, cells were pelleted and washed once with PBS. Pellets were stored at -80˚C until protein purification.

Expression of azidylated TTCF

An overnight culture of B834 containing pET16b_TTCF (earlier cloned from the pEH101 plasmid, kindly provided by prof. C. Watts) [34], or one of the mutants, was diluted 1:100 in LB medium containing 50 µg/mL ampicillin and 1% glucose. Cells were grown at 37˚C, 180 rpm to an OD_{600} of ~0.8 and sedimented (3428 rcf, 15 min, 4˚C) before being washed twice with SelenoMet medium (Molecular Dimensions). Cells were resuspended in SelenoMet medium containing 50 µg/mL ampicillin and depleted for 30 min at 37˚C, 180 rpm before 30 min depletion at 30˚C, 130 rpm. After this azidohomoalanine (L-Aha TFA salt; 72 mg/L) was added to the culture and 15 min later

the culture was induced with isopropyl-β-D-thiogalactopyranoside (IPTG; 1 mM final concentration) and expression took place ON. After overnight expression, cells were pelleted and washed once with PBS. Pellets were stored at -80˚C until protein purification.

Protein purification and analysis
Purification of TTCF and azidylated TTCF

Adjusted from Antoniou *et al.* (2000) [34]. A 1L culture pellet was resuspended in 15 mL lysis buffer containing 100 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 1 mg/ml lysozyme, 250 U benzonase. Lysis took place for 20 min at rt, before separating the soluble from the insoluble protein (10000 rcf, 15 min, 4˚C). The supernatant was filtered over a 0.2 µm filter (Filtropur S, Sarstedt) and loaded on a 5 mL His-column (Ni-NTA Superflow Cartridge; Qiagen). After loading, the column was washed with 5 CV of IMAC25 buffer (100 mM Tris-HCl pH 8.0, 500 mM NaCl, 25 mM imidazole), before eluting with a 15 CV gradient 25 - 500 mM imidazole. The protein fractions were monitored by resolving samples in a 10% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue G-250) revealed the purity of the protein. The fractions containing the most pure protein were combined and extensively dialysed (6 - 8 kDa MWCO, 3.3 mL/cm, FisherBrand or 12 - 14 kDa MWCO, 2 mL/cm, Spectra/Por) against PBS. The protein was concentrated over an amicon spin filter 10 kDa MWCO, before being aliquoted, flash frozen and stored at -80˚C until further use.

Cell culture
General

A20s and the T cell hybridoma 2F2 were a kind gift of C. Watts and were tested on regular basis for mycoplasma contamination. Cultures were discarded after 2 months of use. A20s were cultured at 37˚C under 5% CO_2 in RPMI 1640 (containing 25 mM HEPES) supplemented with stable glutamine (2 mM), heat inactivated fetal calf serum (10% v/v; Biowest), β-mercaptoethanol (50 µM), penicillin (200 IU/mL; Duchefa) and streptomycin (200 µg/mL; Duchefa). Cells were passaged every 2 - 3 days. The 2F2 T cell hybridoma was cultured at 37˚C under 5% CO_2 in RPMI 1640 (containing 25 mM HEPES)

supplemented with stable glutamine (2 mM), heat inactivated fetal calf serum (10% v/v; Biowest), pyruvate (1 mM), NEAA (1x), β-mercaptoethanol (50 μM), penicillin (200 IU/mL; Duchefa) and streptomycin (200 μg/mL; Duchefa). Cells were passaged every 2 - 3 days.

The MHC class II Ova (323-339) specific T cell DO11.10 was a kind gift of F. Ossendorp. The cells were cultured at 37˚C under 5% CO_2 in IMDM supplemented with stable glutamine (2 mM), heat inactivated fetal calf serum (10% v/v; Biowest), β-mercaptoethanol (50 μM), penicillin (200 IU/mL; Duchefa) and streptomycin (200 μg/mL; Duchefa). Cells were passaged every 2 - 3 days.

BMDCs were isolated from the bone marrow from the tibia and femurs of a C57BL/6 mice. Cells were cultured at 37˚C under 5% CO_2 in IMDM medium supplemented with stable glutamine (2 mM), heat inactivated fetal calf serum (10% v/v; Biowest), β-mercaptoethanol (20 μM), GM-CSF (20 ng/mL, Immunotools), penicillin (200 IU/mL; Duchefa) and streptomycin (200 μg/mL; Duchefa). Cells were passaged on day 4 and day 7. Generally cell experiments were performed on day 7 or 8.

Degradation assay

BMDCs (50000 cells/well) were seeded in a 96-well tissue-culture treated microtiter plate. Adherence was allowed at 37˚C under 5% CO_2 for at least 1 h prior to the addition of 100 μg of protein. After 1 h cells were sedimented and protein was chased for 0, 1, 4 or 24 h at 37˚C under 5% CO_2. Subsequently, cells were sedimented by centrifugation (360 rcf, 5 min, rt) and stored at -20˚C before fluorescent analysis.

Pulse assay

As described for degradation assay with the following adjustments. Protein was pulsed for 0, 22, 42 and 62 min. Cells were sedimented by centrifugation (360 rcf, 5 min, 4˚C) and the pellets were stored at -20˚C before fluorescent analysis.

Antigen presentation assay – TTCF proteins

BMDCs (50000 cells/well) were seeded in a 96-well tissue-culture treated microtiter plate. Adherence was allowed at 37˚C under 5% CO_2 for at least 1 h prior to the addition of peptides or proteins at indicated concentrations. The cells were incubated with the

antigens for 4 hours, followed by 2 times washing with PBS. The corresponding T cell hybridoma (50000 cells/well) was added to the pulsed BMDCs and cells were co-cultured for 17 h or 20 h for antigen recognition and IL-2 production by the T cells at 37˚C under 5% CO_2. After overnight incubation, cells were sedimented by centrifugation (360 rcf, 5 min, rt) and supernatant was transferred to a new 96-wells plate. Stimulation of the T cell hybridoma was measured by IL-2 readout using an ELISA assay according to manufacturer's protocol (Invitrogen).

Antigen presentation assay – Ova peptides

A20s (50000 cells/well) were seeded in a 96-well tissue-culture treated microtiter plate. Adherence was allowed at 37˚C under 5% CO_2 for at least 1 h prior to the addition of the peptides at indicated concentrations. The cells were incubated with the antigens for 4 hours, followed by the addition of DO11.10 (50000 cells/well). Subsequently, the cells were co-cultured for 20 h for antigen recognition and IL-2 production by the T cells at 37˚C under 5% CO_2. After overnight incubation, cells were sedimented by centrifugation (360 rcf, 5 min, rt) and supernatant was transferred to a new 96-wells plate. Stimulation of the T cell was measured by IL-2 readout using an ELISA assay according to manufacturer's protocol (Invitrogen).

Antigen presentation blocking assay – TTCF proteins

BMDCs (50000 cells/well) were seeded in a 96-well tissue-culture treated microtiter plate. Adherence was allowed at 37˚C under 5% CO_2 for at least 1 h prior to the addition of 0.3 mg/mL of protein and 1 ng/μL lipopolysaccharide (LPS). After 6 h of pulse, cells were sedimented and washed once with PBS. Protein was chased in BMDCs overnight in medium containing LPS. After 18 h the cells were sedimented (360 rcf, 5 min, rt) and washed once with PBS. Subsequently, cells were fixed on ice with 50 μL ice cold 0.2% glutaraldehyde. After 60 sec the reaction was quenched by the addition of 200 μL of ice cold BMDC medium. Cells were sedimented by centrifugation (300 rcf, 5 min, rt) and washed once with 100 μL PBS. Subsequently biotin-phosphine was added at indicated concentrations for 2 hours at 37˚C under 5% CO_2 in BMDC medium. The cells were washed once with PBS before the 2F2 T cell hybridoma (50000 cells/well) were added to the pulsed BMDCs and co-cultured for 20 h 2F2 T cell hybridoma at 37˚C under 5% CO_2.

After overnight incubation, cells were sedimented by centrifugation (360 rcf, 5 min, rt) and supernatant was transferred to a new 96-wells plate. Stimulation of the T cell hybridoma was measured by IL-2 readout using an ELISA assay according to manufacturer's protocol (Invitrogen).

Fluorescent analysis

Fluorescent analysis of azidylated TTCF

TTCF-Aha and its mutants were characterized by ligating Alexa Fluor 647 (AF647; Invitrogen) alkyne to the azide click handle. This was done via CuAAC reaction, 2.5 µg protein was combined 2:1 (v/v) with click mix (containing 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand, 30 mM aminoguanidine-HCl and 14 µM AF647-alkyne in 88 mM HEPES pH 7.2, final concentration in click mix) reaction for 1 hour at rt in the dark. The reaction was quenched by the addition of 4* Laemmli buffer. Samples were resolved in a 10% SDS-PAGE gel along with PageRuler™ Plus Protein Marker (Thermo Scientific), before scanning Cy3 and Cy5 multichannel settings (605/50 and 695/55 filters, respectively; ChemiDoc™ MP System, Bio-Rad). Coomassie staining (Coomassie Brilliant Blue G-250) was used for correct protein loading.

Degradation and pulse assay

Cells were lysed in 50 µL lysis buffer (100 mM HEPES pH 7.2, 50 mM NaCl, 1x EDTA free protease inhibitor (Roche), 0.25% CHAPS, 25 U benzonase) for 30 min on ice. Protein concentration was measured by Qubit according to manufacturer's protocol (Invitrogen). The cell lysate was diluted to 1.2-1.3 mg/mL in 20 µL and combined 2:1 (v/v) with click mix (containing 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand, 30 mM aminoguanidine-HCl and 14 µM AF647-alkyne in 88 mM HEPES pH 7.2, final concentration in click mix) to react the remaining azidylated protein for 1 hour at rt in the dark. The reaction was quenched by the addition of 4* Laemmli buffer. Samples (30 µL, 18-20 µg of total lysate) were resolved in a 15% SDS-PAGE gel along with PageRuler™ Plus Protein Marker (Thermo Scientific), before scanning Cy3 and Cy5 multichannel settings (605/50 and 695/55 filters, respectively; ChemiDoc™ MP System, Bio-Rad). Coomassie staining (Coomassie Brilliant Blue G-250) was used for correct protein loading.

Chapter 3

Immunoprecipitation

$5.0 \cdot 10^7$ A20 cells were seeded ($1 \cdot 10^6$/mL medium) in a culture flask and left for 1 hour before a 24 h incubation with either ISQAVHAAHAEINEAGR, ISQAVHAAHAEINAhaAGR or ISQAVHAAHAEINbiotinAGR (100 nM). Cells were then washed twice with PBS and subjected to a freeze/thaw cycle. Cells were lysed in lysis buffer containing 0.25% deoxycholate, 0.5% IGEPAL CA-630, 1 mM EDTA, 1 mM PMSF, Roche EDTA free protease inhibitor in PBS for 1 hour, rotating at 4°C and the solubilized proteins were separated by centrifugation (15000 rcf, 15 min, 4°C). The cell lysate containing ISQAVHAAHAEINAhaAGR was next reacted with biotin-phosphine (2 hours at 37°C) and the resulting peptide was purified over a Zeba spin desalting column. The cleared lysates were diluted to 30 mg/mL in lysis buffer and to 5 mg/mL in PBS, before being subjected to immunoprecipitation using Dynabeads™ MyOne™ Streptavidin C1 magnetic beads. After binding for 1 hour at 4°C and a subsequent binding at rt for 30 min, the beads were washed with buffer A (150 mM NaCl in 20 mM Tris-HCl, pH 7.4), buffer B (400 mM NaCl in 20 mM Tris-HCl, pH 7.4), buffer A and buffer C (20 mM Tris-HCl, pH 8.0). Subsequently the protein was eluted by suspending the beads in 3x 50 μL 1* Laemmli sample buffer (5 min at 65°C).

Western blot analysis

Samples were resolved in a 10% TGX Stain Free gel along with PageRuler™ Plus Protein Marker (Thermo Scientific) and transferred onto a PVDF membrane by Trans-Blot Turbo™ Transfer system directly after scanning. Membranes were washed with TBS and TBST and blocked with 5% BSA in TBST at rt for 1 h. Subsequently, the membranes were incubated with primary antibody in 5% BSA in TBST (1 h at rt) Membranes were washed 3x with TBST and incubated with matching secondary antibody in 5% milk in TBST (1 h at rt). Subsequently washed three times with TBST and once with TBS. Membranes were developed with luminol (10 mL of 1.4 mM luminol in 100 mM Tris, pH 8.8 + 100 μL of 6.7 mM p-coumaric acid in DMSO + 3 μL of 30% (v/v) H_2O_2) [49] and chemiluminescence was detected on the ChemiDoc™ MP System in the chemiluminescence channel and the protein marker was visualized with Cy3 and Cy5 settings.

Primary antibodies: monoclonal InVivoMAb anti-mouse MHC Class II (β-chain), clone KL277 (1:1000, BioXCell, BE0140)

Secondary antibody: Rabbit Anti-Armenian hamster IgG H&L (1:5000, Abcam, ab5745).

Graphical analysis

All analysis was determined using GraphPad Prism® 6 and 8 or Microsoft Excel 2016.

References

1. Blum, J.S., Wearsch, P.A. and Cresswell, P. (2013). Pathways of antigen processing. *Annu. Rev. Immunol. 31,* p.443-473

2. Neefjes, J., Jongsma, M.L.M., Paul, P. and Bakke, O. (2011). Towards a systems understanding of MHC Class I and MHC Class II antigen presentation. *Nat. Rev. Immunol. 11,* p.823-836

3. Luckheeram, R.V., Zhou, R., Verma, A.D. and Xia, B. (2012). CD4+ T cells: Differentiation and functions. *Clin. Dev. Immunol.* 925135

4. Mosmann, T.R., Cherwinski, H., Bond, M.W., Giedlin, M.A. and Coffman, R.L. (1986). Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. *J. Immunol. 136,* p.2348-2357

5. Zhu, J. and Paul, W.E. (2008). CD4 T cells: Fates, functions, and faults. *Blood. 112,* p.1557-1569

6. Mucida, D. and Cheroutre, H. (2010). The many face-lifts of CD4 T helper cells. *Adv. Immunol. 107,* p.139-152

7. den Haan, J.M.M., Arens, R. and van Zelm, M.C. (2014). The activation of the adaptive immune system: Cross-talk between antigen-presenting cells, T cells and B cells. *Immunol. Lett. 162,* p.103-112

8. Gagliani, N. and Huber, S. (2017). Basic aspects of T helper cell differentiation. *Methods Mol. Biol. 1514,* p.19-30

9. Nelson, B.H. (2004). IL-2, regulatory T cells, and tolerance. *J. Immunol. 172,* p.3983-3988

10. Sakaguchi, S., Sakaguchi, N., Asano, M., Itoh, M. and Toda, M. (1995). Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J. Immunol. 155,* p.1151-1164

11. Dooms, H., Wolslegel, K., Lin, P. and Abbas, A.K. (2007). Interleukin-2 enhances CD4+ T cell memory by promoting the generation of IL-7Rα-expressing cells. *J. Exp. Med. 204,* p.547-557

12. Roche, P.A. and Furuta, K. (2015). The ins and outs of MHC class II0mediated antigen processing and presentation. *Nat. Rev. Immunol. 15,* 203-216

13. Chicz, R.M., Urban, R.G., Lane, W.S., Gorga, J.C., Stern, L.J. Vignali, D.A.A. and Strominger, J.L. (1992). Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature. 358,* p.764-768

14. Holland, C.J., Cole, D.K. and Godkin, A. (2013). T cell responses with the flanking residues of MHC class II-bound peptides: The core is not enough. *Front. Immunol. 4,* 172

15. Zhang, L., Udako, K., Mamitsuka, H. and Zhu, S. (2012). Toward more accurate pan-specific MHC-peptide binding prediction: A review of current methods and tools. *Brief. Bioinform. 13,* p.350-364

16. ten Broeke, T., Wubbolts, R. and Stoorvogel, W. (2013). MHC class II antigen presentation by dendritic cells regulated through endosomal sorting. *Cold Spring Harb. Perspect. Biol. 5,* p.a016873

17. Davidson, H.W., Reid, P.A., Lanzavecchia, A. and Watts, C. (1991). Processed antigen binds to newly synthesized MHC class II molecules in antigen-specific B lymphocytes. *Cell. 67,* p.105-116

18. Reid, P.A. and Watts, C. (1992). Constitutive endocytosis and recycling of major histocompatibility complex class II glycoproteins in human B-lymphoblastoid cells. *Immunology. 77,* p.539-542

19. Moss, C.X., Tree, T.I. and Watts, C. (2007). Reconstruction of a pathway of antigen processing and class II MHC peptide capture. *EMBO J. 26,* p.2137-2147

20. Sinnathamby, G. and Eisenlohr, L.C. (2003). Presentation by recycling MHC class II molecules of an influenza hemagglutinin-derived epitope that is revealed in the early endosome by acidification. *J. Immunol. 170,* p.3504-3513

21. Watts, C., Reid, P.A., Simitsek, P.D. and West, M.A. (1991). Antigen processing and class II MHC-restricted antigen presentation in B-lymphocytes. *Biochem. Soc. Trans. 19,* p.263-265

22. Pathak, S.S. and Blum, J.S. (2000). Endocytic recycling is required for the presentation of an exogenous peptide via MHC class II molecules. *Traffic. 1,* p.561-569

23. Roche, P.A., Ishido, S., Eisenlohr, L.C. and Cho, K-J. (2020). Activation of dendritic cells alters the mechanism of MHC class II antigen presentation to CD4 T cells. *J. Immunol. 204,* p.1621-1629

24. Barra, C., Alvarez, B., Paul, S., Sette, A., Peters, B., Andreatta, M., Buus, S. and Nielsen, M. (2018). Footprints of antigen processing boost MHC class II natural ligand predictions. *Genome med. 10,* 84

25. Watts, C. (2012). The endosome-lysosome pathway and information generation in the immune system. *Biochim. Biophys. Acta. 1824,* p.14-21

26. Fairweather, N.F. and Lyness, V.A. (1986). The complete nucleotide sequence of tetanus toxin. *Nucleic Acids Res. 14,* p.7809-7812

27. Matthews, S.P., Werber, I., Deussing, J., Peters, C., Reinheckel, T. and Watts, C. (2010). Distinct protease requirements for antigen presentation *in vitro* and *in vivo. J. Immunol. 184,* p.2423-2431

28. Manoury, B., Hewitt, E.W., Morrice, N., Dando, P.M., Barrett, A.J. and Watts, C. (1998). An asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation. *Nature. 396,* p.695-699.

29. Watts, C. (2001). Antigen processing in the endocytic compartment. *Curr. Opin. Immunol. 13,* p.26-31

30. Araman, M.C., Pieper-Pournara, L., van Leeuwen, T., Kampstra, A.S.B., Bakkum, T., Marqvorsen, M.H.S., Nascimento, C.R., Groenewold, G.J.M., Wulp, W., Camps, M.G.M., Overkleeft, H.S., Ossendorp, F.A., Toes, R.E.M. and van Kasteren, S.I. (2019). Bioorthogonal antigens allow the unbiased study of antigen processing and presentation. *BioRxiv.* p.439323

31. Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A. and Schuman, E.M. (2006). Selective identification of newly synthesized proteins in mammalian cells using biorthogonal noncanonical amino acid tagging (BONCAT). *PNAS, 103,* p.9482-9487

32. Kiick, K.L., Saxon, E., Tirrell, D.A. and Bertozzi, C.R. (2002). Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *PNAS. 99,* p.19-24

33. van Hest, J.C.M., Kiick, K.L. and Tirrell, D.A. (2000). Efficient incorporation of unsaturated methionine analogues into proteins *in vivo. J. Am. Chem. Soc. 122,* p.1282-1288

34. Antoniou, A.N., Blackwood, S-L., Mazzeo, D. and Watts, C. (2000). Control of antigen presentation by a single protease cleavage site. *Immunity. 12,* p.391-398

35. Wood, W.B. (1966). Host specificity of DNA produced by *Escherichia coli*: Bacterial mutations affecting the restriction and modification of DNA. *J. Mol. Biol. 16,* p.118-133

36. van Kasteren, S.I., Kramer, H.B., Jensen, H.H., Campbell, S.J., Kirkpatrick, J., Oldham, N.J., Anthony, D.C. and Davis, B.G. (2007). Expanding the diversity of chemical protein modification allows post-translational mimicry. *Nature. 446,* p.1105-1109

37. Araman, M.C., Pieper-Pournara, L., van Leeuwen, T., Kampstra, A.S.B., Bakkum, T., Marqvorsen, M.H.S., Nascimento, C.R., Groenewold, G.J.M., Wulp, W., Camps, M.G.M., Overkleeft, H.S., Ossendorp, F.A., Toes, R.E.M. and van Kasteren, S.I. (2019). Bioorthogonal antigens allow the unbiased study of antigen processing and presentation. *BioRxiv.* p.439323

38. Bakkum, T., van Leeuwen, T., Sarris, A.J.C., van Elsland, D.M., Poulcharidis, D., Overkleeft, H.S. and van Kasteren, S.I. (2018). Quantification of bioorthogonal stability in immune phagocytes using flow cytometry reveals rapid degradation of strained alkynes. *ACS Chem. Biol. 13,* p.1173-1179

39. Robinson, J.H. and Delvig, A.A. (2002). Diversity in MHC class II antigen presentation. *Immunology. 105,* p.252-262

40. Ito, Y., Ashenberg, O., Pyrdol, J., Luoma, A.M., Rozenblatt-Rosen, O., Hofree, M., Christian, E., Ferrari de Andrada, L., en Tay, R., Teyton, L., Regev, A., Dougan, S.K. and Wucherpfennig, K.W. (2018). Rapid CLIP dissociation from MHC II promotes an unusual antigen presentation pathway in autoimmunity. *J. Exp. Med. 215,* p.2617-2635

41. Roy, B.M., Zhukov, D.V. and Maynard, J.A. (2012). Flanking residues are central to DO11.10 T cell hybridoma stimulation by ovalbumin 323-339. *PLoS One. 7,* e47585

42. Robertson, J.M., Jensen, P.E. and Evavold, B.D. (2000). DO11.10 and OT-II T cells recognize a C-terminal ovalbumin 323-339 epitope. *J. Immunol. 164,* p.4706-4712

43. Sofron, A., Ritz, D., Neri, D. and Fugmann, T. (2016). High-resolution analysis of the murine MHC class II immunopeptidome. *Eur. J. Immunol. 46,* p.319-328

44. Bassani-Sternberg, M., Barnea, E., Beer, I., Avivi, I., Katz, T. and Admon, A. (2010). Soluble plasma HLA peptidome as a potential source for cancer biomarkers. *Proc. Natl. Acad. Sci. U.S.A. 107,* p.18769-18776

45. Potter, G.T., Jayson, G.C., Miller, G.J. and Gardiner, J.M. (2016). An updated synthesis of the diazo-transfer reagent imidazole-1-sulfonyl azide hydrogen sulfate. *J. Org. Chem. 81,* p.3443-3446

46. More, S.S. and Vince, R. (2009). Inhibition of Glyoxalase I: The first low-nanomolar tight-binding inhibitors. *J. Med. Chem. 52,* p.4650-4656

47. Zhang, M.M., Tsou, L.K., Charron, G., Raghavan, A.S. and Hang, H.C. (2010). Tandem fluorescence imaging of dynamic S-acylation and protein turnover. *PNAS. 107,* p.8627-8632

48. Hong, V., Presolski, S.I., Ma, C. and Finn, M.G. (2009). Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew. Chem. Int. Ed. Engl. 48,* p.9879-9883

49. Mruk, D.D. and Cheng, C.Y. (2011). Enhanced chemiluminescence (ECL) for routine immunoblotting: An inexpensive alternative to commercially available kits. *Spermatogenesis. 1,* p.121-122

Chapter 4

Exploring azido-HRP as a tool in immunocytochemistry

Abstract

Previously it was shown that horseradish peroxidase (HRP) could be modified with azide residues without losing its enzymatic activity. It was postulated that this modified protein could potentially be used as a new tool in immunocytochemistry. The production of this HRP-N$_3$ and its use as an enzymatic detection reagent for bioorthogonal groups was assessed, which proved to be useful for the detection of recombinant bioorthogonal proteins by dot blot and SDS-PAGE/Western blot, but not for the detection of bioorthogonal labels in an immunocytochemical setting.

Introduction

Immunohistochemistry and immunocytochemistry staining are well-established tools to visualize the presence of specific antigens in biological samples [1-3]. In these approaches, antibodies raised against antigens of interest conjugated to reporter proteins are used to detect their presence in tissue samples. Despite its long history, this process is still of pivotal importance in the clinic for the identification of primary tumors [4-6], pathogen infections [7-9], auto-immune disease [10-12] and other pathologies.

Paul Ehrlich is commonly regarded as the founding father of the fields of histochemistry and histopathology. His pioneering work on the use of aniline dyes on animal sections [13] allowed for the first time to directly visualize the presence of pathogenic bacteria in tissue samples [14]. In 1890 Emil von Behring reported on the harvesting and use of antibodies [15-17]. Marrack and co-workers subsequently and inspired by the work of Ehrlich and Behring came up with the idea to combine a dye with an antibody and in this manner color typhoid and cholera in tissue samples [18]. A decade later, Coons *et al.* performed what may be considered as the first true immunohistochemistry experiments by using antibody fluorophore conjugates to visualize pneumococci in histopathological samples [19]. Chemical conjugation of an antibody to reporter enzymes, technologies developed for the first time in the sixties of last century [20, 21], gave another major impetus to the field. Signal amplification, created through enzymatic processing of a reporter substrate, now allowed the detection of rare antigens. The use of substrate/enzyme pairs that yielded osmiophilic products, also allowed the combination of histochemistry with electron microscopy (EM)-based imaging techniques [1, 22].

Immunostaining thus pertains the formation of antigen-antibody complexes, and the subsequent detection of these complexes through visualizing the antibody that is selected and/or engineered for these purposes. The antigen-recognizing antibody can be directly conjugated to the reporter entity, for instance and as is depicted in Figure 1, horseradish peroxidase (HRP). Alternatively, the antigen-specific antibody may be designed such that it is recognized by a generic, secondary antibody that in turn is conjugated to HRP. In both these examples, antigens are detected by adding an HRP
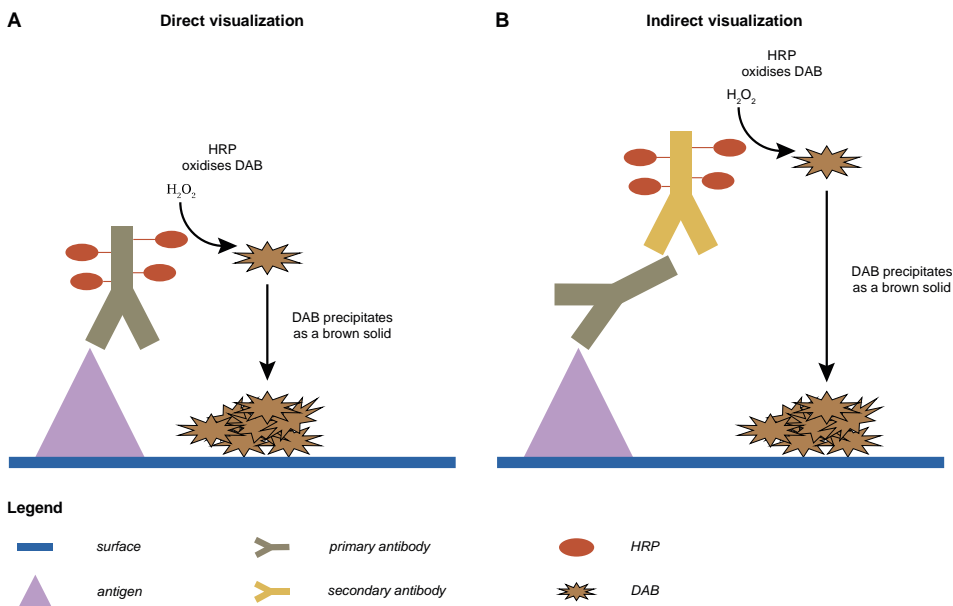
Figure 1. Different immunostaining approaches. A) Direct visualization of the antigen: The primary antibody, which recognizes a specific antigen, is conjugated to HRP. Hereafter HRP can precipitate for instance 3,3'-diaminobenzidine (DAB). B) Indirect visualization of the antigen: The primary antibody recognizes a specific antigen. The secondary will recognize this primary antibody and is conjugated to the visualization moiety, again used for DAB precipitation.

substrate that upon HRP-mediated oxidation yields a luminescent signal, or as depicted here, a colorimetric deposit.

Two of the used reporter enzymes are β-galactosidase (β-Gal) [23] and HRP [24]. When active, both can enzymatically convert so-called chromogens, compounds that produce a colored product. HRP, for example, can oxidize different substrates, such as 3,3'-diaminobenzidine (DAB) (Scheme 1A) [25, 26], 3-amino-9-ethylcarbazole (AEC) [27, 28] and 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) (Scheme 1B) [29, 30], moreover it can oxidize chemiluminescent reagents such as luminol (Scheme 1C) [31]. The reactions as depicted in the forward direction in Scheme 1, are chemically as follows: 1. DAB reacts with H_2O_2 in the presence of the catalyst HRP. In the first step, a quinone iminium cation radical is formed after the donation of an electron to HRP, hereafter in the second step DAB polymerizes into a complex brown precipitate after the donation of another electron [32]. 2. The simplified reaction of 2 molecules of ABTS with H_2O_2 under the catalyzation of HRP resulting 2 ABTS radicals [33, 34]. 3. Luminol reacts

A

B

C
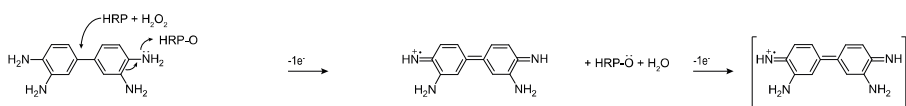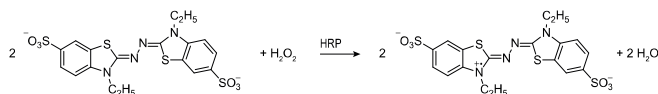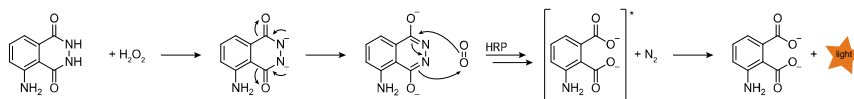
Scheme 1. Schematic overview of conversions by HRP. All of the here presented molecules react with H_2O_2 using HRP as a catalyst. A) DAB conversion into its insoluble brown precipitate. B) The formation of green ATBS radicals from ABTS. C) Oxidation of luminol resulting in luminescence.

with hydroxyl groups to form a dianion intermediate. After tautomerization and the reaction with hydrogen peroxide, an unstable peroxide is formed. The resulting compound, 3-aminophthalate, will be in its excited state and with the return to its ground state emit light [35, 36].

One caveat in immunohistochemistry is that certain antigens defy antibody recognition. Immunization of an antibody-producing host with small molecular antigens for instance and because of a limited number of chemical functionalities for antibody-antigen interaction, often yield low-affinity antibodies at most. Moreover, some antigens are simply not immunogenic, meaning no antibodies will be produced. Another caveat is the possible lack of specificity of an antibody towards the antigen it has been developed for, in which case cross-reactivity may yield false positives. Complementary histological techniques not reliant on antibodies or random chemical dye labelling could be an addition to further advance histological approaches.

One technique that has the potential to play a role in immunohistochemistry is bioorthogonal chemistry. With the introduction of this technique, it became possible to perform a chemical reaction in a biological setting without or with minimal interference of the natural processes in experimental settings [38]. In order to use bioorthogonal chemistry in an histochemistry approach, e.g. an azide group can be introduced into a
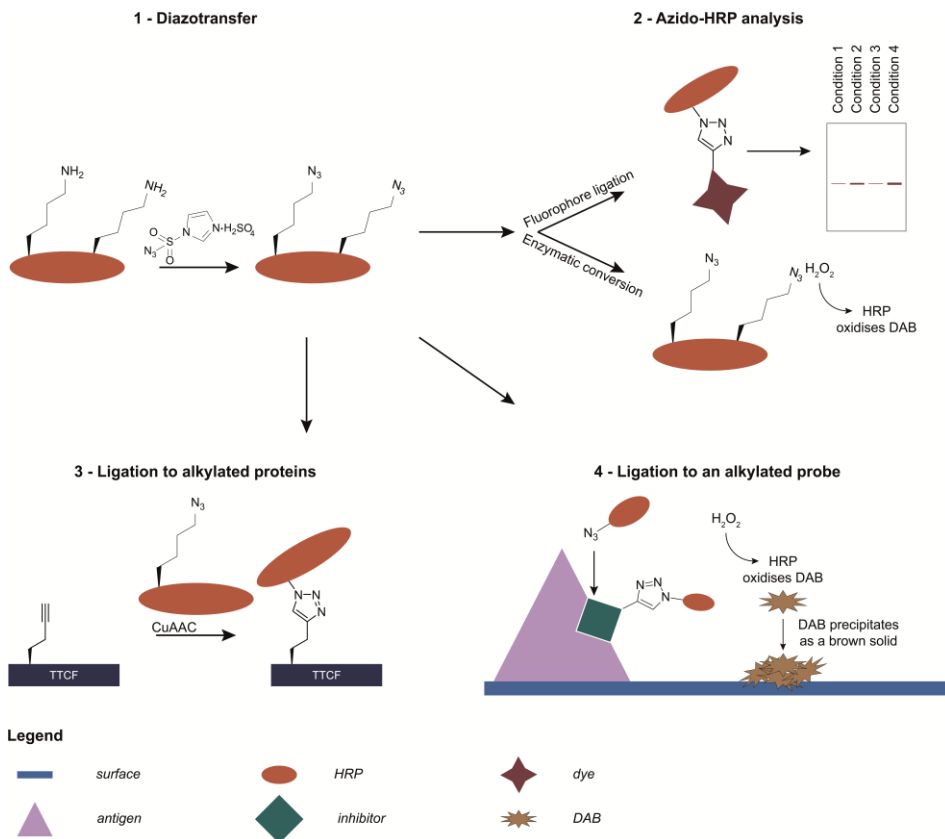
Figure 2. Schematic overview of the use of HRP-N₃. Different diazotransfer conditions to HRP were explored (**1**) and analyzed via e.g. fluorescent labeling or DAB precipitation studies (**2**). Next, the possibility to ligate this moiety to alkynylated proteins and bacteria were analyzed (**3**) and subsequently followed by testing its use in probe targeting studies (**4**).

biomolecule and a Staudinger ligation [39] or a Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) [40] can be used to ligate a reporter molecule to the azide. In terms of histology, click chemistry, a fast ligation reaction between i.e. an alkyne and an azide [40], has been used to successfully image the bacterial peptidoglycan layer in sputum samples [41], but its use in histology has been relatively unexplored. The work in this Chapter describes research efforts to develop a click-histology workflow, which would expand the histology toolbox.

In 2007, Davis and co-workers described the production of azide/alkyne-containing β-galactosidase [42], which could serve as an enzyme in the click-histology workflow, as this histology-compatible enzyme remained active even after chemical

modification using ligation reactions. In 2009, Van Dongen *et al.* described the use of the diazotransfer reagent developed by Stick and co-workers [43] to azidylate the six lysine residues and the N-terminal amine of HRP without loss of peroxidase activity [44, 45]. It was thus postulated that these reagents could serve as histological tools for visualizing the presence of click handles. In this Chapter it was chosen to explore azido-HRP as a reagent, due to the tetrameric nature and more complex purification of β-Gal [42].

Results

The first steps in the exploration of azido-HRP as a histological click reagent was the optimization of the diazotransfer conditions to maximize the number of azide groups per protein without affecting its catalytic ability (Figure 2 - step 1). This was done in two steps. The efficiency of the diazotransfer reaction was assessed by qualifying their reactivity using a Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) also known as Cu(I)-catalyzed Huisgen cycloaddition (ccHc) with a fluorescent (Alexa Fluor 647 (AF647)-conjugated) alkyne. The enzymatic activity of the HRP proteins that had undergone diazotransfer under various conditions was assessed by their ability to convert ABTS and DAB (Figure 2 - step 2). The azido-HRP constructs featuring optimal enzymatic activity is evaluated as click reagent in the second part of the Chapter (Figure 2 - step 3 and 4).

4.1 Optimization of the diazotransfer reaction

In first instance, the diazotransfer conditions as described by van Dongen *et al.* were evaluated [44]. In their original protocol, van Dongen *et al.* reacted 1 eq. of HRP with 13 eq. of imidazole-1-sulfonyl azide hydrogen sulfate (Stick's reagent) and 8.8 eq. of copper in 4.3 mM K_2CO_3 buffer. To see whether this reaction could be further optimized, the pH, Stick-reagent and copper concentrations were varied (Table 1). After purification over an Amicon Ultra spin column the efficiency of the reaction was determined by conjugation of an AF647 alkyne using CuAAC conditions.

In terms of the reactivity of HRP, van Dongen *et al.* modified on average four amines into azido residues [44]. In this study, the reaction could not be further

Table 1. Overview of the different diazotransfer conditions. The most optimal conditions as derived from the fluorescent ligation reaction and the initial ABTS analysis. The azidylated conditions used for further investigation (condition E, I, J and van Dongen HRP) are noted with an asterisk.

Reaction	HRP (eq.)	K₂CO₃ (conc. or eq.)	Cu(II)SO₄.5H₂O (eq.)	Stick (H₂SO₄; eq.)	Fluorescent signal (%)	Activity (eq.)
A	1.0	100 mM, pH 8	1.1	20	38 ± 35	81
B	1.0	100 mM, pH 8	1.1	10	46 ± 4	103
C	1.0	100 mM, pH 8	1.1	5	28 ± 1	107
D	1.0	100 mM, pH 8	1.1	1	8 ± 1	109
E*	1.0	100 mM, pH 9	1.1	20	88 ± 15	154
F	1.0	100 mM, pH 9	1.1	10	79 ± 21	106
G	1.0	100 mM, pH 9	1.1	5	54 ± 6	99
H	1.0	100 mM, pH 9	1.1	1	13 ± 1	93
I*	1.0	100 mM, pH 10	1.1	20	129 ± 11	142
J*	1.0	100 mM, pH 10	1.1	10	98 ± 10	82
K	1.0	100 mM, pH 10	1.1	5	64 ± 10	103
L	1.0	100 mM, pH 10	1.1	1	12 ± 1	117
van Dongen HRP*	1.0	127.3	8.8	12.6	100 ± 0	100
HRP	1.0	-	-	-	0 ± 0	100

optimized: lowering the pH to 8 reduced the diazotransfer yield (conditions A-D; Table 1), as did lowering of the amount of Stick's reagent. The optimal ligation range was found to be pH 9-10 with 10-20 equivalents Stick's reagent (entries E, I, J; Table 1 and Figure 3A and B). The residual catalytic activity did not vary significantly between the different reaction conditions when assessed with an ABTS assay [44]. All of the azidylated proteins remained catalytically active with an observed activity ranging from 81 - 154% of that of the unmodified wild type counterpart (Table 1 and Figure 3C). Which means that the
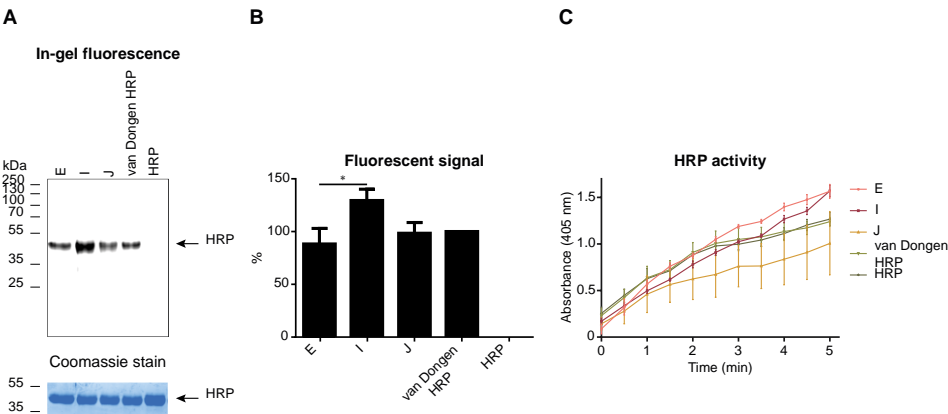

Figure 3. Analysis of the different azidylated HRP proteins compared to native HRP. A) Representative in-gel fluorescence after AF647 alkyne ligation to the differently derived azidylated HRP proteins (conditions E, I, J and van Dongen HRP). B) Intensity in % of the fluorescent signal compared to azido-HRP. C) Colorimetric conversion of ABTS by HRP and its azidylated versions with in Table 1 the corresponding activity as compared to unmodified HRP. Data is expressed as mean ± SD (N = 3, n = 1). Statistical analysis was performed using Tukey's multiple comparisons test. * P ≤ 0.05.
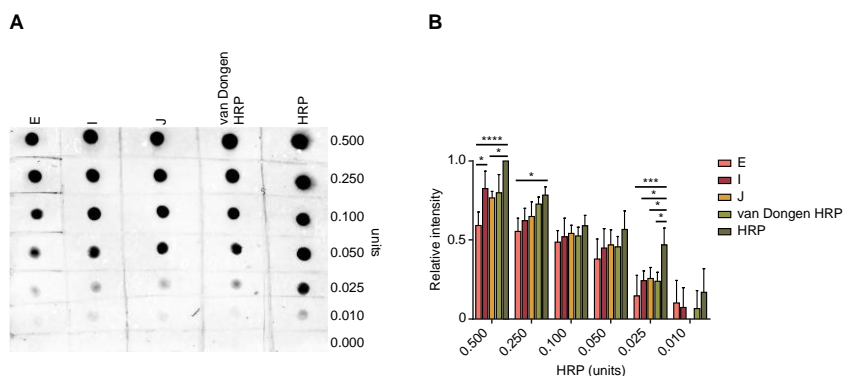
Figure 4. Dot blot analysis of various amounts of modified HRP and native HRP. The dots were dotted onto a PVDF membrane and visualized with commercially available DAB and imaged after 3-4 minutes of enzymatic reaction. A) Representative dot blot of the differently derived HRP proteins. This shows in B) the relative intensity of the dots, normalized to HRP 0.500 units. The activity of the azidylated proteins was in some cases significantly reduced, with condition I seen as most comparable to native HRP. Data represent means ± SD (N = 3, n = 1). Statistical analysis was performed using Tukey's multiple comparisons test. **** $P \leq 0.0001$; *** $P \leq 0.001$; * $P \leq 0.05$.

conditions as described by van Dongen *et al.* were near optimal for our purposes [44].

Azido-HRP produced by using conditions E, I and J, and the van Dongen conditions were next compared with wild type HRP, using a concentration range of the modified enzymes in a dot blot analysis. After blocking and washing, the membrane containing HRP-N$_3$ was incubated with DAB substrate and the amount of brown precipitate formed was assessed. Despite an absence in activity differences in the ABTS assay, here significant variance between the differentially produced HRP was observed (Figure 4A and B). The limit of detection of van Dongen, condition E, condition I and wild type was >0.010 units, whereas the limit of detection for condition J was >0.025 units. For all subsequent experiments, condition I (pH 10, 20 eq. Stick's Reagent) was used to produce HRP-N$_3$.

4.2 Ligation reactions

It was next tested whether HRP-N$_3$ could be used for the visualization of alkynylated proteins on membrane. For this, either alkynyl-TTCF (TTCF-Hpg; produced as described in more detail in **Chapter 2** and **3** [46, 47]), containing eight homopropargylglycine (Hpg) residues per protein (including the start codon), or wild
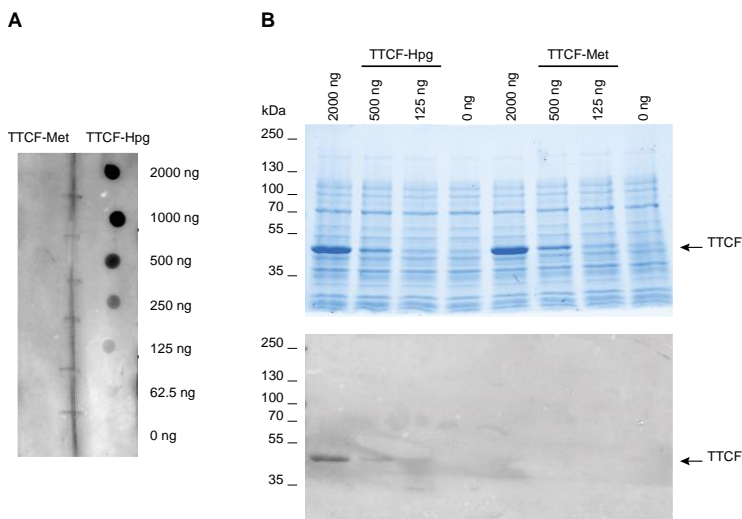
Figure 5. HRP-N₃ ligation to folded and unfolded TTCF-Hpg. A) The TTCF-Hpg and TTCF-Met proteins were dotted in a non-denatured state on a PVDF membrane, and after CuAAC reaction with HRP-N₃ as reporter molecule, HRP was visualized with DAB precipitation. This resulted in visualization of only TTCF-Hpg as low as 62.5 ng of protein. B) TTCF-Hpg and TTCF-Met were combined with B834 cell lysate in different concentrations and run over a 10% TGX Stain-Free gel. After visualization of the proteins in a Stain-Free gel (upper panel), the gel was transferred to a PVDF membrane and a CuAAC reaction was performed on blot with HRP-N₃ as ligating molecule. HRP was visualized with DAB precipitation (lower panel). The resulting precipitate was only visible at positions where TTCF-Hpg was present (2000, 500 and 125 ng of protein).

type TTCF was serially diluted and dotted onto a polyvinylidene difluoride (PVDF) membrane followed by a blocking step with 5% nonfat dry milk. The membrane was then subjected to a CuAAC ligation mixture, containing 13 μM azido-HRP in presence of 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand, 30 mM aminoguanidine-HCl in 88 mM HEPES pH 7.2. The copper ions of the ligation reaction were removed to prevent interference with the DAB oxidation, by washing with 100 mM EDTA [20, 48]. Next, the blots were subjected to DAB-staining, which resulted in visible DAB precipitation for the serially-diluted TTCF-Hpg (limit of detection 62.5 ng/1.2 pmol protein; Figure 5A). No signal was observed at any concentration in the negative control (use of native TTCF). This experiment showed the possibility to ligate HRP-N₃ to a non-denatured protein and to visualize this via DAB precipitation.

Next, the specificity of the CuAAC ligation reaction was assessed by serially diluting the TTCF-Hpg and TTCF-Met in *Escherichia coli* (*E. coli*) lysate. The resulting mixtures
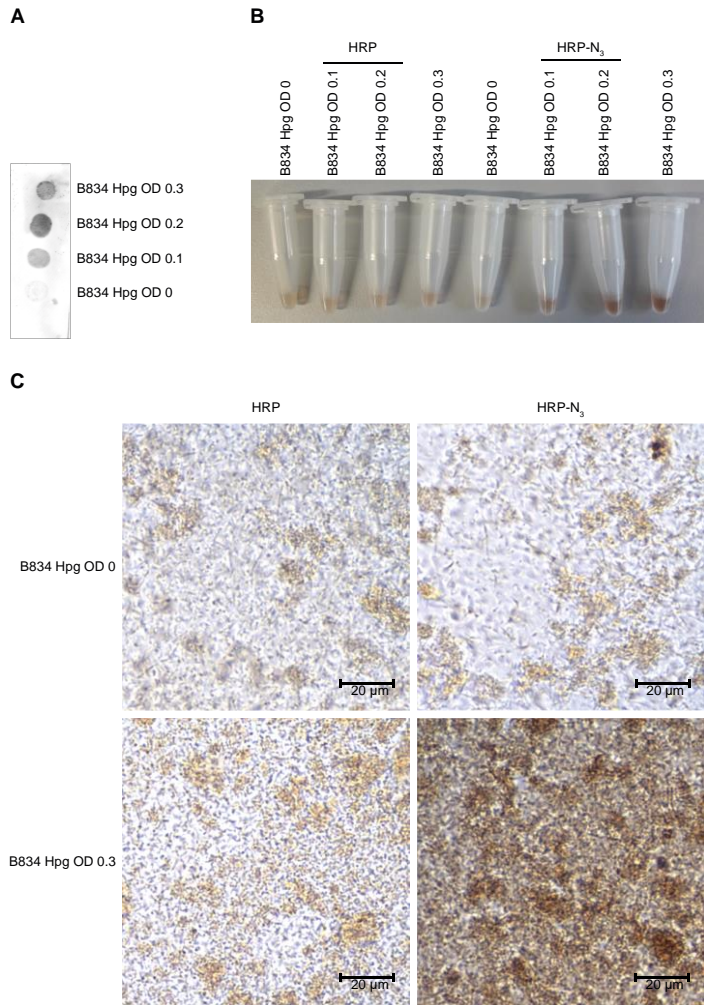
Figure 6. HRP-N₃ ligation to alkynylated bacteria, enzymatically reacted with DAB. A) Representative blot of a serial dilution of alkynylated bacteria, ligated to HRP-N₃. Which showed an increase in DAB precipitation with an increased concentration of alkynylated bacteria. The dot blot was representative for the performed triplicate. B) Visualization of the DAB precipitation performed in solution phase after ligation to HRP and HRP-N₃ and C) the microscopic analysis of the ligation reaction. The DAB signal of the alkynylated bacteria reacted with HRP-N₃ was higher than background.

were separated in a TGX Stain-Free gel and transferred to a PVDF-membrane. The resulting blot was reacted with 40 eq. of HRP-N₃, followed by washing with buffer containing imidazole [49], citric acid [50] and EDTA [51], again to remove residual copper ions. Reaction with DAB yielded a detectable signal at the expected protein size of 53 kDa, with a detection limit of 125 ng (2.3 pmol, corresponding to 18.7 pmol of

alkyne moieties; Figure 5B, lower panel). Furthermore, the blot and the accompanying gel confirmed the selectivity of the ligation reaction (Figure 5B).

4.3 Bacterial cell labelling

After establishing that HRP-N$_3$ can be used to visualize alkynylated proteins, it was next explored whether this approach could also be used to detect alkynes in a model immunocytochemical setting. This setting was simulated on dot blot and in a solution phase experiment, using *E. coli* bacteria of which one population was labeled with L-Hpg and the other one was unlabeled. For this, the *E. coli* methionine auxotroph B834 was used and either grown to exponential phase in LB medium or in the presence of L-Hpg for metabolic labelling [46]. After growth, the mixtures of alkynylated and unmodified bacteria were washed, fixed with 4% PFA, permeabilized with either 0.1% IGEPAL or 0.15% Triton X-100, washed and subsequently incubated with 1% H$_2$O$_2$ to block endogenous peroxidases [52, 53]. After an additional washing step, the mixtures were dotted onto PVDF membrane, blocked with 5% milk and subjected to a CuAAC ligation mixture. Subsequently, the membrane was washed as before and stained with DAB. Imaging the blots showed the DAB precipitation to be dependent on the percentage Hpg-labeled *E. coli* present in each sample. (Figure 6A). Similar results were obtained when the reaction was performed in liquid phase (Figure 6B and C). These experiments showed that the approach is compatible with a histological workflow.

4.4 In-cell labelling of an enzyme activity

One of the main areas where bioorthogonal reactions are truly additive to antibody stainings is in the in-cell detection of enzyme activities [54, 55]. By introducing ligation handles into covalent enzyme inhibitors, active populations of a particular enzyme can be visualized (Figure 2 - step 4); something that is not possible using antibody techniques. In this manner, cathepsins [56], post-proline proteases [57], cysteine proteases [58, 59] and serine proteases [60, 61] have been visualized, to name but a few. The approach is also moving to the clinic for e.g. the identification of tumor boundaries [62]. As conventional tumor histology is done using immunohistochemical approaches,
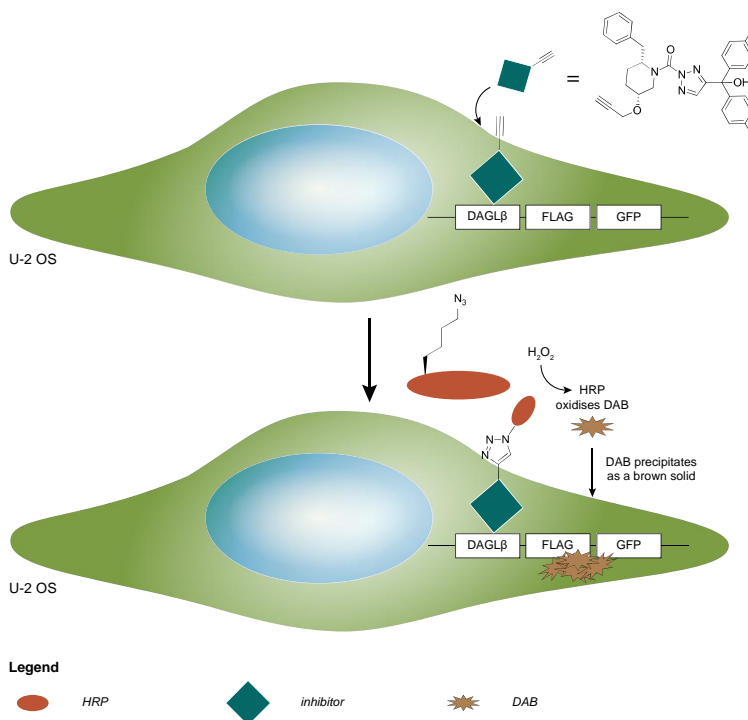
Figure 7. Schematic overview of HRP-N₃ ligation to probe DH376 which can inhibit i.e. DAGLβ. This overview shows the inhibition of DAGLβ by DH376 (left) and the ligation of HRP-N₃ with a CuAAC reaction. When reacting with DAB, precipitation should occur only at the site of the protein inhibited by the alkyne containing probe.

it was attempted to combine activity-based protein profiling (ABPP) and CuAAC ligation with HRP-N₃ staining.

Two stable U-2 OS cell lines expressing the serine hydrolases DAGLβ (unpublished) and ABHD6 [63] combined with a FLAG-tag, for immunoblotting, and a GFP reporter gene, for fluorescent colocalization, were used to demonstrate this concept. The two serine hydrolases were labeled with an alkyne-containing inhibitor of the enzymes, DH376 (Figure 7) [64].

At first, it was tested what the limit of detection was for detecting an ABP-labeled enzyme with HRP-N₃ by Western blot using luminol as substrate for the DAGLβ-FLAG-GFP cell construct. The protein was reacted with 100 nM of the alkyne-containing ABP DH376 *in situ* and the cell lysate resolved in a TGX Stain-Free gel, prior to being transferred to a PVDF membrane. Western blot analysis was performed using antibodies
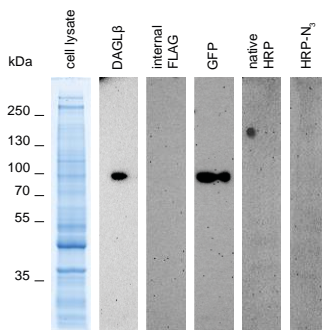
Figure 8. Western blot analysis of the construct and HRP-N₃ ligation to probe DH376. Cell lysate and Western blot analysis of a DAGLβ-FLAG-GFP cell line treated with DH376. Western blot was performed against DAGLβ, FLAG and GFP. The PVDF membrane containing the cell lysate was furthermore reacted with HRP and HRP-N₃ in the presence of a CuAAC ligation mixture. HRP-N₃ could not be visualized using a chemiluminescent readout.

against DAGLβ, FLAG and GFP. Alternatively, the blots were exposed to a CuAAC ligation mixture containing either HRP or HRP-N₃. These experiments show that the protein could readily be visualized by Western blot (using a chemiluminescent readout) against DAGLβ (~100 kDa protein (DAGLβ is a 74 kDa and GFP a 27 kDa protein) and GFP (Figure 8, panel "DAGLβ" and "GFP"), but not the FLAG-tag (Figure 8, panel "internal FLAG"). Moreover, the ligation reaction with HRP(-N₃), failed to yield a detectable signal (Figure 8, panel "native HRP" and "HRP-N₃").

The detection was also attempted in whole cells using microscopy and DAB-precipitation. For this, cells containing either the DAGLβ or ABHD6 construct were treated with 100 nM DH376 and fixed with 4% paraformaldehyde (PFA). These cells were permeabilized with 0.15% Triton-X100 and a CuAAC ligation reaction was performed in the presence of HRP-N₃ or HRP (as negative control). After washing the cells with a 100 mM EDTA wash buffer and PBS, the cells were subjected to DAB precipitation and subsequently washed with PBS. Microscopic analysis of the cells showed that no DAB precipitation could be observed that correlated with enzyme presence for DAGLβ and only marginally for ABHD6 (Figure 9). The presence of DAB precipitate led to a decrease in GFP signal [65, 66].

It was therefore hypothesized that either the alkyne handle of DH376 was buried in the active site of the enzyme, or the fixation and/or permeabilization of the cells was
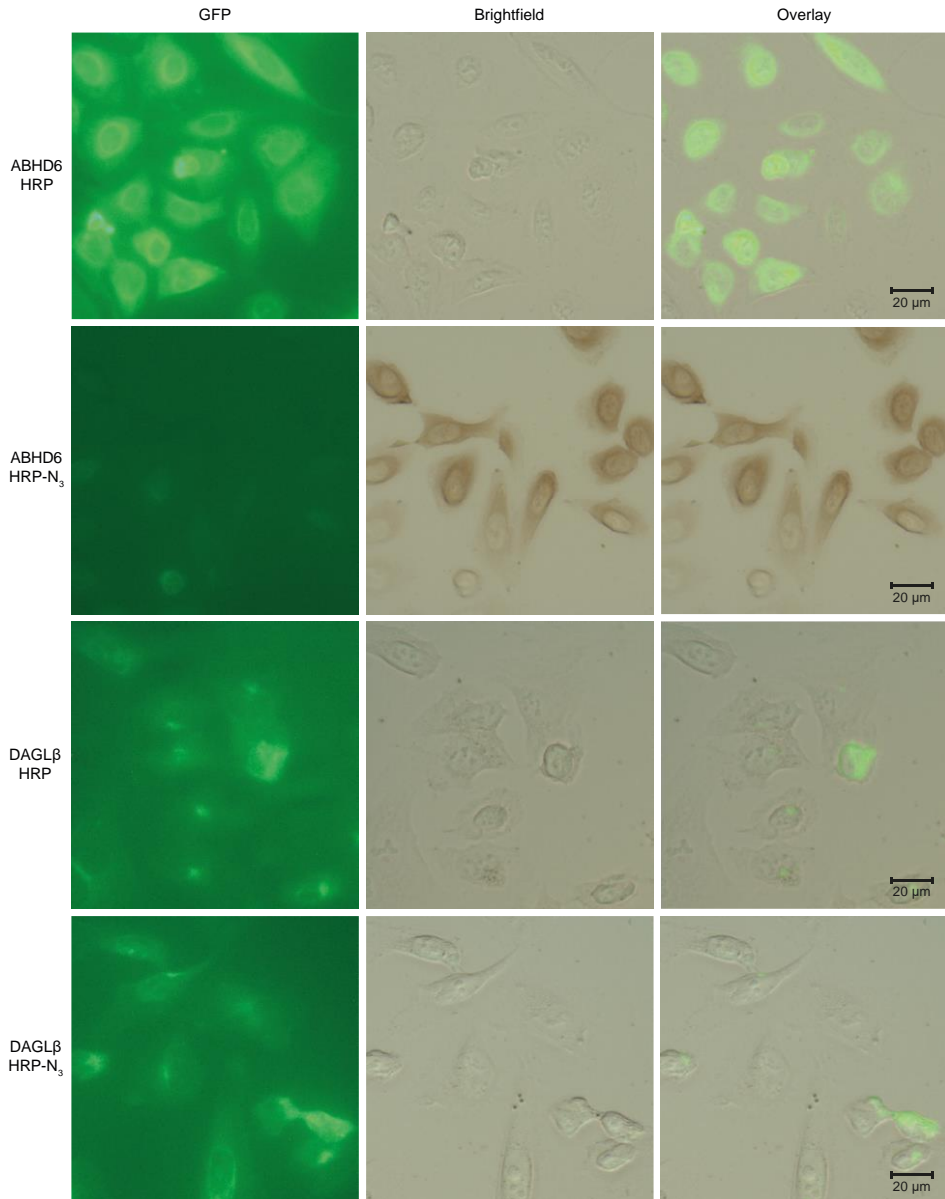
Figure 9. Microscopic analysis of DAB precipitation after probe treatment of DAGLβ-FLAG-GFP containing cells. Both the GFP signal, the brightfield signal and an overlay of these two signals were displayed. DAB precipitation is visible on cells treated with HRP-N₃, but not in cells reacted with HRP. None of the treated cells showed a GFP signal.
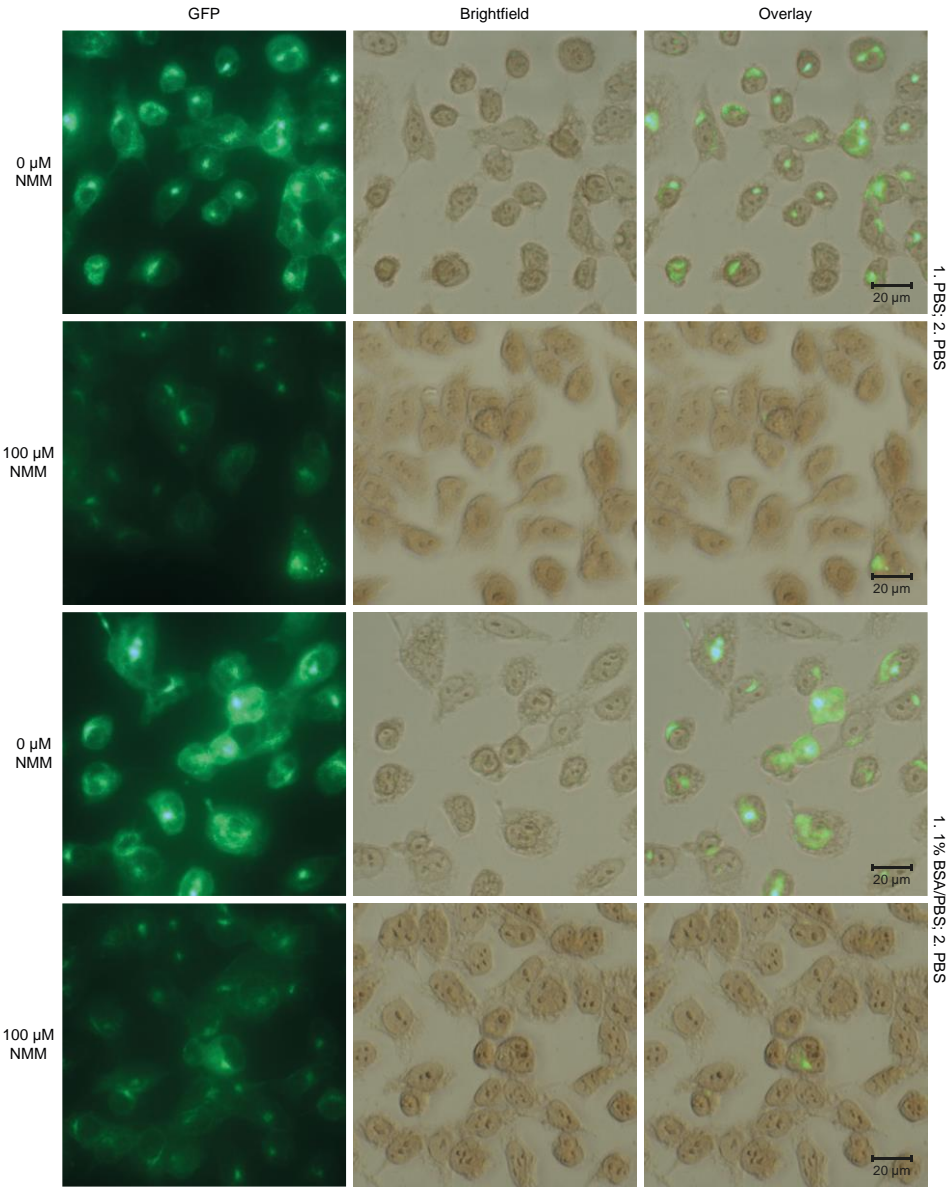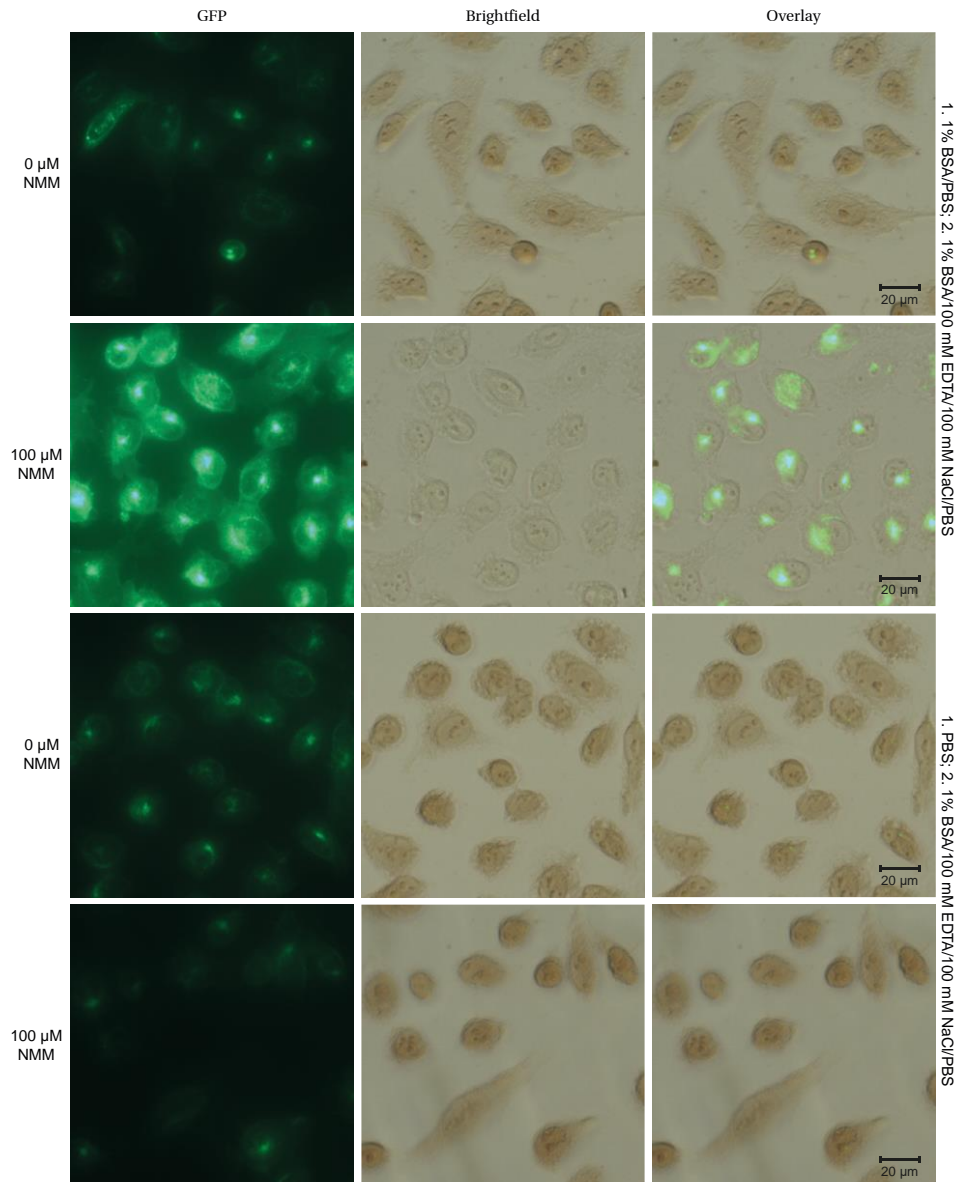
Figure 10. The use of different blocking agents. Different wash steps were combined with and without blocking the free amines with NMM. Wash step 1, with either PBS or 1% BSA/PBS, was performed before the addition of HRP and wash step 2, with either PBS or 1% BSA/100 mM EDTA/100 mM NaCl/PBS was performed after the addition of HRP.

unsuccessful. The fixation and permeabilization steps were varied in order to obtain increased DAB precipitation. Cells were treated with the following different

fixative/permeabilizer agents: acetone, methanol, methanol:acetone 1:1 and methanol:ethanol 1:1. Moreover, PFA fixation was combined with permeabilization by digitonin and saponin. These conditions resulted in the formation of brown precipitate for both the HRP-N$_3$ ligation and the negative control (HRP).

The formation of brown precipitate in the negative control was hypothesized to stem from aspecific binding and therewith an issue with incomplete blocking. Different blocking agents were therefore also assessed. First, immediately after methanol fixation the cells were washed with a mix of glycine/PBS (blocks free aldehydes) [67] and after blocking potential internal peroxidases with H$_2$O$_2$, free amines were blocked with N-methylmorpholine (NMM) buffered in PBS [68]. Furthermore, a combination of wash steps with BSA/PBS and BSA/EDTA/NaCl/PBS was introduced, the first before and the latter after the addition of HRP. The only blocking method that prevented cells to be stained brown was blocking with both 1% BSA/PBS and 1% BSA/100 mM EDTA/100 mM NaCl/PBS (Figure 10).

Some of the alternative wash steps showed samples free of background. This combination, 1% BSA/PBS and 1% BSA/100 mM EDTA/100 mM NaCl/PBS, was therefore used in a follow up experiment, using either HRP or HRP-N$_3$ in the ligation mix. Unfortunately, the DAB precipitation was higher in the negative control samples as compared to the samples (Figure 11), leading to the conclusion that this method is not sensitive enough for detecting the ABP.

Discussion and conclusion

This Chapter describes a new method for the visualization of alkynylated proteins and bacteria in immunohistological settings. In the first part of this Chapter the optimization of HRP-azidylation was attempted. The best condition – after analysis in a fluorescent and DAB precipitation assay – was found to be condition I (pH 10, 20 eq. Stick's Reagent). This means at higher pH and higher equivalents of Stick's Reagent, the diazotransfer to HRP resulted in the best azidylation with respect to its ligation and residual DAB activity.
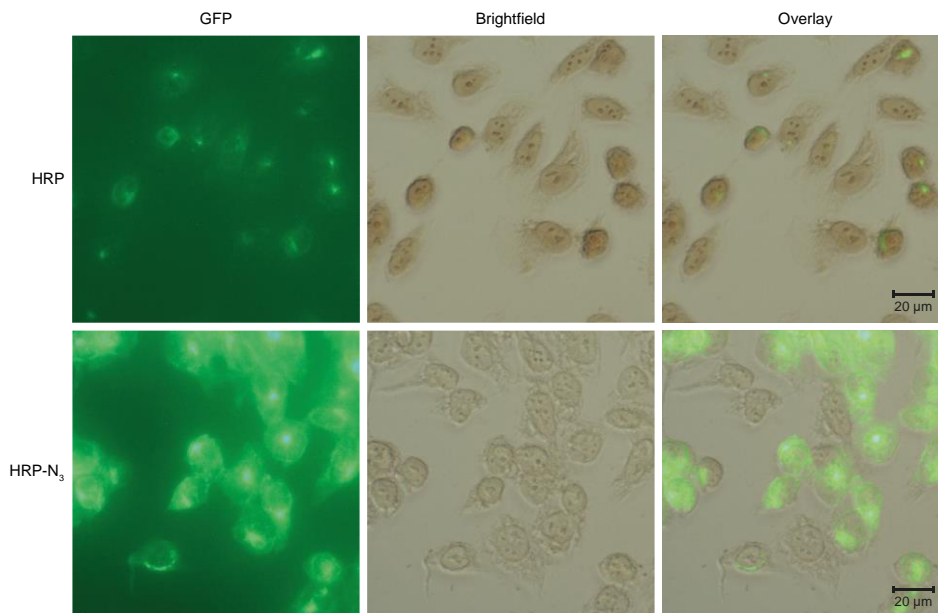
Figure 11. The use of NMM with blocking agents 1% BSA/PBS and 1% BSA/100 mM EDTA/100 mM NaCl/PBS. Upper panels: Background signal using HRP in the ligation mixture. Lower panels: The ligation reaction performed using HRP-N$_3$ in the ligation mixture.

This reagent was next used to detect the presence of alkyne groups in a variety of samples. For instance, the approach could be used to selectively detect an alkynylated protein in cell lysate, and the presence of labelled bacteria in a mixed population. It could not, however, be used to identify an alkyne-modified activity based probe in mammalian cell samples.

When working on the ligation of the bacterial cell lysates, it was shown that each of the experiments showed background staining with DAB. In each of them, 1% H$_2$O$_2$ was used to block endogenous peroxidases. However, literature data is available that this might not be the optimal condition for peroxidase blockade [49]. In future, it is suggested to repeat the ABTS reactivity experiment as described for alkynylated bacteria, using a concentration range of H$_2$O$_2$ with the aim to increase the signal-to-noise in the microscopic analysis which was described in the in cell labeling section of this Chapter.

The experiments in mammalian cells were less successful. The reaction of azido-HRP with an alkyne containing active site probe did not yield a detectable signal, either in a blot-experiment or in cell-based imaging assays. For now, it can be concluded that

azido-HRP is a useful tool for the detection of abundant loci of alkynes in samples, but not for approaches where few and/or diffuse labels are present.

Acknowledgements

Materials and methods

General

All chemicals and reagents were purchased at Sigma-Aldrich, Alfa Aesar, Acros, Merck or VWR, unless stated otherwise. Reagents were used without further purification. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), Stain-Free gel and Western blot materials were purchased at Bio-Rad. DH376 was previously synthesized in-house and characterized by NMR and LC-MS [64]. Cell culture disposables were from Sarstedt.

Solutions

PBS contained 5 mM KH_2PO_4, 15 mM Na_2HPO_4, 150 mM NaCl, pH 7.4 and PBST was PBS supplemented with 0.05% Tween-20. TBS contained 50 mM Tris-HCl, 150 mM NaCl and TBST was TBS supplemented with 0.05% Tween-20. Laemmli sample buffer 4* contained 60 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.01% (v/v) bromophenol blue.

Compound synthesis

Imidazole-1-sulfonyl azide hydrogen sulfate

As described by Potter *et al.* [45]. A 250 ml round-bottomed flask with a stirring bar was brought under nitrogen atmosphere and cooled to 0˚C. Sodium azide (5.2 g, 80 mmol, 1.0 eq.) was added before the addition of dry EtOAc (80 mL). After adding SO_2Cl_2 (6.5 ml, 80 mmol, 1.0 eq.) over 10 min the reaction was slowly warmed to rt and left stirring under an inert atmosphere overnight. The reaction mixture was then recooled to 0˚C, before the addition of imidazole (10.9 g, 160 mmol, 2.0 eq.) and the resulting suspension was stirred for 3.5 h at 0°C. The reaction mixture was basified with saturated aqueous $NaHCO_3$ (150 mL). The organic layer was separated from the aqueous layer, washed with H_2O and dried with $MgSO_4$. The suspension was filtered and cooled to 0˚C. Concentrated H_2SO_4 (4.4 ml, 80 mmol, 1.0 eq.) was added over 5 min. The mixture was warmed to rt and stirred until a precipitate had formed, which was filtered and washed with EtOAc. The remaining white crystals were dried under vacuum to afford the title compound pure (12.8 g, 47 mmol, 59%).

Chapter 4

Synthesis of tris-hydroxypropyltriazolylmethylamine (THPTA)

A two-step synthesis as described before by Hong *et al.* [70].

Synthesis of (S)-2-Aminohex-5-ynoic acid (L-Hpg)

Synthesis as described before by Li *et al.* [71], followed by a procedure based on Biagini *et al.* [71], Dong *et al.* [72] Chenault *et al.* [73] to obtain the chirally pure variant.

HRP modification and analysis

Diazo-transfer to HRP

Adjusted from van Dongen *et al.* (2009) [44]. To a solution of HRP (263-325 units/mg; P8375 10kU) in MQ (2.5 mg/mL) an aqueous solution of K_2CO_3 (either 2 mg/mL or 100 mM, pH 8-10) was added. Subsequently followed by the addition of $CuSO_4.5H_2O$ (1 mg/mL) and Stick's reagent (2 mg/mL) and the reaction was left shaking at 20˚C 500 rpm overnight. The mixture was transferred to an Amicon Ultra spin filter with either a 3 or 10 kDa cutoff and centrifuged to dryness. The protein was redissolved in MQ water and centrifuged for another 5-8 washings. The end product was redissolved in MQ water and the presence of click handles was visualized via fluorescence analysis. Bradford analysis, with BSA as standard, was used to measure protein concentrations.

Fluorescent analysis of HRP-N$_3$

HRP-N$_3$ was characterized by ligating Alexa Fluor 647 (AF647) alkyne to the azide click handle. This was done via CuAAC reaction, protein was combined 2:1 (v/v) with click mix (containing 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand, 30 mM aminoguanidine-HCl and 14 µM AF647-alkyne in 88 mM HEPES pH 7.2, final concentration in click mix) reaction for 1 hour at rt in the dark. The reaction was quenched by the addition of 4* Laemmli buffer. Samples were resolved in a 10% SDS-PAGE gel (180 V, 75 min) along with PageRuler^{TM} Plus Protein Marker (Thermo Scientific), before scanning Cy3 and Cy5 multichannel settings (605/50 and 695/55 filters, respectively; ChemiDoc™ MP System, Bio-Rad). Coomassie staining (Coomassie Brilliant Blue R-250) was used for correcting protein loading and quantification was done using Image Lab (Bio-Rad).

Activity assay

Adjusted from Van Dongen *et al.* [44]. All solutions were kept on ice until measurement. A solution of ABTS/H_2O_2 (8 mM ABTS in 20 mM sodium phosphate buffer, pH 7.4 containing 0.2% H_2O_2) was prepared fresh before each measurement. HRP or HRP-N_3 was diluted to 2.5 units/mL in 20 mM sodium phosphate buffer, pH 7.4 and fractionate 100 µL per well of a 96 well plate. As a control 100 µL of phosphate buffer was used. To each well, 5 µL of ABTS/H_2O_2 solution was added immediately before measuring the absorption at 405 nm on a TECAN Infinite M1000 Pro plate reader.

Bacterial growth and protein expression

B834 growth

B834 was inoculated in triplicate with 1:100 diluted ON culture grown in LB medium. The cells were grown until an OD_{600} of ~0.8 and sedimented in aliquots to a final OD_{600} of 0.2 and stored until use at -20˚C.

Expression and purification of recombinant TTCF

Adjusted from Antoniou *et al.* (2000) [75]. An overnight culture of B834 containing pET16b_TTCF (earlier cloned from the pEH101 plasmid, kindly provided by prof. C. Watts) was diluted 1:100 in LB medium containing 50 µg/mL ampicillin and 1% glucose. Cells were grown at 37˚C, 180 rpm to an OD_{600} of ~0.8 and sedimented (3428 rcf, 15 min, 4˚C) before being washed twice with SelenoMet medium (Molecular Dimensions). Cells were resuspended in SelenoMet medium containing 50 µg/mL ampicillin and depleted for 30 min at 37˚C, 180 rpm before 30 min depletion at 30˚C, 130 rpm. After this either L-Methionine (L-Met; 40 mg/L; Ajinomoto) or L-Homopropargylglycine (L-Hpg; 40 mg/L) was added to the culture and 15 min later the culture was induced with isopropyl-β-D-thiogalactopyranoside (IPTG; 1 mM final concentration) and expression took place ON. After overnight expression, cells were pelleted and washed once with PBS. Pellets were stored at -80˚C until protein purification. A 1 L culture pellet was resuspended in 15 mL lysis buffer containing 100 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 1 mg/mL lysozyme, 250 U benzonase. Lysis took place for 20 min at rt, before separating the soluble from the insoluble protein (10000 rcf, 15 min, 4˚C). The supernatant was filtered over a 0.2 µm filter and loaded on a 5 mL His-column (Ni-NTA

Superflow Cartridge; Qiagen). After loading, the column was washed with 5 CV of IMAC25 buffer (100 mM Tris-HCl pH 8.0, 500 mM NaCl, 25 mM imidazole), before eluting with a 15 CV gradient 25 - 500 mM imidazole. The protein fractions were monitored by SDS-PAGE as described before and Coomassie staining revealed the purity of the protein. The fractions containing the most pure protein were combined and extensively dialysed (6 - 8 kDa MWCO, 3.3 mL/cm, FisherBrand or 12 - 14 kDa MWCO, 2 mL/cm, Spectra/Por) against PBS. The protein was concentrated over an Amicon spin filter 10 kDa MWCO, before being aliquoted, flash frozen and stored at -80˚C until further use.

Dot blot and Western blot analysis
HRP and HRP-N_3

A serial dilution of HRP and HRP-N_3 was made and 2 μL was pipetted onto a PVDF membrane to form a dot. The protein was spotted on three separate membranes, dried for 1 h and blocked in 5% nonfat dry milk (Elk; Campina) in TBST at room temperature for 1.5 h or at 4˚C ON. Subsequently washed 3x with TBST and 1x with TBS before adding DAB substrate (BD Biosciences). After 3-4 minutes the membranes were imaged on a ChemiDoc™ MP System, Bio-Rad and quantified by ImageJ.

Recombinant protein

A serial dilution of TTCF-Hpg was made and 2 μL was pipetted onto PVDF membrane to form a dot. The protein was spotted on three separate membranes, dried for 1 h and blocked in 5% nonfat dry milk (Elk; Campina) in TBST for 1.5 h at room temperature. HRP-N_3 was ligated to the proteins as described for fluorescence analysis, using 12.8 μM (349 eq. N_3 - alkyne) HRP-N_3 instead of a fluorophore. The blot was washed 3x with 100 mM EDTA in TBST, 3x with TBST and 1x with TBS. Subsequently DAB was added and after 3-5 minutes the membranes were imaged on a ChemiDoc™ MP System, Bio-Rad and quantified by ImageJ.

Western blot analysis

Samples were resolved in a 10% Stain-Free gel (proteins at indicated concentrations or 20 μg cell lysate) along with PageRuler™ Plus Protein Marker (Thermo

Scientific) and transferred onto a PVDF membrane by Trans-Blot Turbo™ Transfer system directly after scanning. Membranes were washed with TBS and TBST and blocked with 5% nonfat dry milk (Elk; Campina) in TBST at rt for 1.5 h or at 4°C ON.

Membranes containing cell lysates were then incubated with primary antibody in 5% nonfat dry milk (Elk; Campina) in TBST (GFP; FLAG; 1 h at rt) or washed 3x times with TBST and incubated with DAGLβ in 5% BSA in TBST (4°C, ON). Membranes were washed 3x with TBST and incubated with matching secondary antibody in 5% nonfat dry milk (Elk; Campina) (1 h at rt). Subsequently washed three times with TBST and once with TBS. Membranes were developed with luminol (10 mL of 1.4 mM luminol in 100 mM Tris, pH 8.8 + 100 μL of 6.7 mM p-coumaric acid in DMSO + 3 μL of 30% (v/v) H_2O_2) [76] and chemiluminescence was detected on the ChemiDoc™ MP System in the chemiluminescence channel and protein marker was visualized with Cy3 and Cy5 settings (605/50 and 695/55 filters, respectively).

Membranes used for HRP-N_3 ligation were washed 3x times with TBST and a ligation mixture containing HRP or HRP-N_3 (42 eq. or 0.86 nmol) was added (1 h at rt). Subsequently washed 3x with 10% (w/v) citric acid, 100 mM EDTA, pH 8.0, 500 mM imidazole, 3x with TBST and 1x with TBS. Before detecting with DAB substrate or chemiluminescence as described before.

Primary antibodies: monoclonal mouse anti-FLAG M2 (1:5000, Sigma Aldrich, F3156), monoclonal anti-GFP (1:2500, Thermo-Fisher, GF28R), Monoclonal anti-DAG Lipase β (1:1000, Cell Signaling, D4P7C)

Secondary antibodies: mouse IgGκ BP-HRP (1:5000, Santa Cruz, sc-516102), mouse anti-rabbit IgG-HRP (1:5000, Santa Cruz, sc-2357).

Analysis of alkynylated bacteria
B834 growth and storage

B834 was inoculated in triplicate with 1:100 diluted ON culture grown in LB medium. The cells were grown until an OD$_{600}$ of ~0.7-0.8 and sedimented before being resuspend in SelenoMet medium containing either L-Met or L-Hpg. Cells were grown for another 2 hours, before an OD of 0.5 was sedimented. Cells were resuspended in 1 mL of 4% PFA and fixed for 20 min at rt. Fixed cells were washed once with PBS and stored at an OD$_{600}$ of 0.4 in PBS at 4°C.

DAB visualization

Bacterial cells were combined to reach the wanted OD of alkynylated bacteria and solubilized with either 0.1% IGEPAL or 0.15% Triton X-100 for 20-25 min at rt, before being washed once with PBS. Afterwards cells were incubated with 1% H_2O_2 to block endogenous peroxidases. Washed once more with PBS and resuspended in a CuAAC ligation mixture (containing 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand, 30 mM aminoguanidine-HCl and 1 nmol HRP(-N_3) in 88 mM HEPES pH 7.2, final concentration in click mix) for 1 h at rt in the dark. After ligation, cells were washed 3x with 100 mM EDTA and 1x with PBS. Pellet was resuspended in DAB substrate and visualized by visible imaging and microscopic analysis (EVOS FL Auto 2).

Or as described for *recombinant protein*, with minor adjustments.

Cell culture
General cell culture

Cell lines were purchased at ATCC and tested on regular basis for mycoplasma contamination. Cultures were discarded after 2-3 months of use. U-2 OS cell lines expressing either ABHD6 or DAGLβ were kindly provided by A.C.M. van Esbroeck and cultured at 37˚C under 7% CO_2 in DMEM containing phenol red, stable glutamine, newborn bovine serum (10% v/v; Seradigm), G418 (400 µg/mL; supplier) penicillin (200 IU/mL; Duchefa) and streptomycin (200 µg/mL; Duchefa). Medium was refreshed every 2-3 days and cells were passaged two times a week at 80-90% confluence.

ABHD6/DAGLβ inhibition by DH376

U-2 OS cells were seeded at appropriate density one day before treatment (1:250 - 1:400 for 8-well chamber slides (Ibitreat; Ibidi) and 1:10 - 1:15 in culture dishes for the preparation of cell lysates). Culture medium was removed and treatment medium, containing 100 nM DH376, was added.

Cells were trypsinized after incubation for 1 h at 37˚C and 7% CO_2, washed 1x with PBS and spun down (3000 rcf, 5 min, rt). Cell pellets were flash frozen and stored at -80˚C until further use.

For microscopic analysis, medium was aspirated and cells were washed 2x with PBS following the experimental procedure as described below.

Microscopic analysis of treated U-2 OS cells

Cells used for different permeabilization methods were either fixed with 4% PFA in PBS (15 min, rt), washed 1x with PBS and permeabilized with 0.15% Triton X-100, 100 µM Digitonin or 0.5% Saponin in PBS or in one step fixed and permeabilized with solvents acetone, methanol, 1:1 (v/v) methanol:acetone or 1:1 (v/v) methanol:ethanol for 10 min at -20˚C and washed 2x with PBS. Subsequently washed 2 or 3x with PBS(T) and 1% H_2O_2 in PBS (10 min, rt). A 2x wash with PBS(T) was performed before adding 10 µM (1.4 - 2 nmol) HRP($-N_3$) 10 min prior to the click mix (containing 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand and 30 mM aminoguanidine-HCl in 88 mM HEPES pH 7.2, final concentration in the mix) or in the click mix itself. After 1 hour, cells were washed 3x with 100 mM EDTA in PBS and 3x with PBS(T), before the addition of DAB substrate (15 min, rt) and cells were subsequently washed 2x with PBS and stored at 4˚C until microscopic analysis (EVOS FL Auto 2).

Cells used for wash optimization methods were fixed and permeabilized with ice-cold methanol for 10 min at rt. Before washing 3x with 100 mM glycine in PBS and 1x with 1% H_2O_2 in PBS (10 min, rt). Cells were washed 2x with PBS and blocked with 0 or 100 µM NMM in PBS (10 min, rt) and washed 1x with PBS. Followed by incubation with 1% BSA in PBS or PBS (20 min, rt) and washed 1x with PBS. HRP($-N_3$) (1 nmol; 10 µM) was added and after 10 min, when necessary, a click mix (containing 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand and 30 mM aminoguanidine-HCl in 88 mM HEPES pH 7.2, final concentration in the mix) was added and left for 1 h at rt. And added after ligation, or after 10 min HRP incubation, followed by 3x washing with 1% BSA/100 mM EDTA/100 mM NaCl in PBS or PBS and 3x PBS. DAB substrate was added to the chambers for 5 min, rt and cells were subsequently washed 2x with PBS and stored in PBS at 4˚C until microscopic analysis.

Preparation of cell lysates

Cell pellets were thawed on ice and resuspended in lysis buffer (250 mM sucrose, 20 mM HEPES pH 7.2, 1 mM $MgCl_2$, 2 mM DTT, 50 U/mL benzonase). Protein concentrations were determined using Quick Start™ Bradford Protein Assay (Bio-Rad) and diluted to convenient concentration in lysis buffer.

Statistical analysis

Replicates shown represent biological replicates and data represent means ± SD. All statistical analysis was determined using GraphPad Prism® 7/8 or Microsoft Excel 2016. **** $P \leq 0.0001$; *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; NS if $P > 0.05$. All statistical analyses were conducted using GraphPad Prism® 7 or Microsoft Excel 2016.

References

1. Nakane, P.K. and Pierce, G.B. (1966). Enzyme-labeled antibodies: Preparation and application for the localization of antigens. *J. Histochem. Cytochem. 14,* p.929-931

2. Saper, C.B. and Sawchenko, P.E. (2003). Magic peptides, magic antibodies: Guidelines for appropriate controls for immunohistochemistry. *J. Comp. Neurol. 465,* p.161-163

3. Burry, R.W. (2011). Controls for immunocytochemistry: An update. *J. Histochem. Cytochem. 59,* p.6-12

4. Bochtler, T., Löffler, H. and Krämer, A. (2018). Diagnosis and management of metastatic neoplasms with unknown primary. *Semin. Diagn. Pathol. 35,* p.199-206

5. Geramizadeh, B., Marzban, M. and Churg, A. (2016). Role of immunohistochemistry in the diagnosis of solitary fibrous tumor, a review. *Iran J. Pathol. 11,* p.195-203

6. Krajewska, M., Krajwski, S., Epstein, J.I., Shabaik, A., Sauvageot, J., Song, K., Kitada, S. and Reed, J.C. (1996). Immunohistochemical analysis of bcl-2, bax, bcl-X, and mcl-1 expression in prostate cancers. *Am. J. Pathol. 148,* p.1567-1576

7. Guarner, J. and Brandt, M.E. (2011). Histopathologic diagnosis of fungal infections in the 21st century. *Clin. Microbiol. Rev. 24,* p.247-280

8. Haines, D.M. and West, K.H. (2005). Immunohistochemistry: Forging the links between immunology and pathology. *Vet. Immunol. Immunopathol. 108,* p.151-156

9. Eyzaguirre, E. and Hague, A.K. (2008). Application of immunohistochemistry to infections. *Arch. Pathol. Lab Med. 132,* p.424-431

10. Tiniakos, D.G., Brain, J.G. and Bury, Y.A. (2015). Role of histopathology in autoimmune hepatitis. *Dig. Dis. 33,* p.53-64

11. Fujimura, Y. (1996). Pathogenesis of aphtoid ulcers in Crohn's disease: Correlative findings by magnifying colonoscopy, electron microscopy, and immunohistochemistry. *Gut. 38,* p.724-732

12. Larsson, Å., Bredberg, A., Henriksson, G., Manthorpe, R. and Sallmyr, A. (2005). Immunohistochemistry of the B-cell component in lower lip salivary glands of Sjögren's syndrome and healthy subjects. *Scand. J. Immunol. 61,* p.98-107

13. Ehrlich, P. (1877). Contributions to the knowledge of the aniline dyes and their use in the microscopic technique. *Arch. Microsc. Anat. 13,* p.263-278

14. Valent, P., Groner, B., Schumacher, U., Superti-Furga, G., Busslinger, M., Kralovics, R., Zielinski, C., Penninger, J.M., Kerjaschki, D., Stingl, G., Smolen, J.S., Valenta, R., Lassmann, H., Kovar, H., Jäger, U., Kornek, G., Müller, M. and Sörgel, F. (2016). Paul Ehrlich (1854-1915) and his contributions to the foundation and birth of translational medicine. *J. Innate Immun. 8,* p.111-120

15. Behring, E. and Kitasato, S. (1890). Über das zustandekommen der diphtheria-immunität und der tetanus-immunität bei thieren. *Dtsch. Med. Wochenschrift. 49,* p.1113-1114

16. Kaufmann, S.H. (2017). Remembering Emil von Behring: From tetanus treatment to antibody cooperation with phagocytes. *mBio. 8,* e00117-17

17. Childs, G.V. (2014). History of Immunohistochemistry. In: *Pathobiology of human disease.* p.3775-3796

18. Marrack, J. (1934). Nature of antibodies. *Nature. 133,* p.292-293

19. Coons, A.H., Creech, H.J. and Jones, R.N. (1941). Immunological properties of an antibody containing a fluorescent group. *Proc. Soc. Exp. Biol. Med. 47,* p.200-202

20. Vandesande, F. (1979). A critical review of immunocytochemical methods for light microscopy. *J. Neurosci. Methods. 1,* p.3-23

21. Sternberger, L.A. (1969). Some new developmens in immunocytochemistry. *Mikroskopie, 25,* p.346-361

22. Nakane, P.K. and Pierce, G.B. (1967). Enzyme-labeled antibodies for the light and electron microscopic localization of tissue antigens. *J. Cell. Biol. 33,* p.307-318

23. Cotson, S., Holt, S.J. (1958). Studies in enzyme cytochemistry. IV. Kinetics of aerial oxidation of indoxyl and some of its halogen derivatives. *Proc. R. Soc. Lond. B. Biol. Sci. 148,* p.506-519

24. Graham, R.C. and Karnovsky, M.J. (1966). The early stages of absorption of injected horseradish peroxidase in the proximal tubules of mouse kidney: Ultrastructural cytochemistry by a new technique. *J. Histochem. Cytochem. 14,* 291-302

25. Greenwalt, T.J., Swierk, E.M. and Steane, E.A. (1975). Use of horseradish peroxidase-labelled antiglobulin for the colorimetric quantitation of erythrocyte antibodies. *J. Immunol. Methods. 8,* p.351-361

26. Morrell, J.I., Greenberger, L.M. and Pfaff, D.W. (1981). Comparison of horseradish peroxidase visualization methods: Quantitative results and further technical specifics. *J. Histochem. Cytochem. 29,* p.903-916

27. Trojanowski, J.Q., Obrocka, M.A. and Lee, V.M. (1983). A comparison of eight different chromogen protocols for the demonstration of immunoreactive neurofilaments or glial filaments in rat cerebellum using the peroxidase-antiperoxidase method and monoclonal antibodies. *J. Histochem. Cytochem. 31,* p.1217-1223

28. Nakata, T. and Suzuki, N. (2012). Chromogen-based immunohistochemical method for elucidation of the coexpression of two antigens using antibodies from the same species. *J. Histochem. Cytochem. 60,* p.611-619

29. Porstmann, B., Porstmann, T., Nugel, E. and Evers, U. (1985). Which of the commonly used marker enzymes gives the best results in colorimetric and fluorimetric enzyme immunoassays: Horseradish peroxidase, alkaline posphatase or beta-galactosidase? *J. Immunol. Methods. 79,* p.27-37

30. Roberts, I.M., Jones, S.L., Premier, R.R. and Cox, J.C. (1991). A comparison of the sensitivity and specificity of enzyme immunoassays and time-resolved fluoroimmunoassay. *J. Immunol. Methods. 143,* 49-56

31. Kurien, B.T. and Scofield, R.H. (2006). Western blotting. *Methods. 38,* p.283-293

32. Nurjayadi, M., Apriyani, D., Hasan, U., Santoso, I., Kurniadewi, F., Kartika, I.R., Agustini, K., Puspasari, F., Natalia, D. and Mangunwardoyo, W. (2016). Immunogenicity and specificity of anti-recombinant protein Fim-C-*Salmonella typhimurium* antibody as a model to develop Typhoid vaccine. *Procedia Chem. 18,* p.237-245

33. Gräfe, D., Gaitzsch, J., Appelhans, D. and Voit, B. (2014). Cross-linked polymersomes as nanoreactors for controlled and stabilized single and cascade enzymatic reactions. *Nanoscale. 6,* p.10752-10761

34. Chen, J., Zhang, Y., Cheng, M., Guo, Y., Šponer, J., Monchaud, D., Mergny, J-L., Ju, H. and Zhou, J. (2018). How proximal nucleobases regulate the catalytic activity of G-quadruplex/hemin DNAzymes. *ACS Catal. 8,* p.11352-11361

35. Kim, H-S. and Pyun, J-C. (2009). Hyper sensitive strip test with chemiluminescence signal band. *Procedia Chem. 1,* p.1043-1046

36. Ramkiran, A., Mohan, V.G., Mishra, P. and Padmavathy, N. (2018). Enhanced chemiluminescence of luminol by metal peroxides nanoparticles. *Chem. Adv. Mater. 3,* p.16-22

37. Hein, C.D., Liu, X-M, Wang, D. (2008). Click chemistry, a powerful tool for pharmaceutical sciences. *Pharm. Res. 25,* p.2216-2230

38. Sletten, E.M. and Bertozzi, C.R. (2011). From mechanism to mouse: A tale of two biorthogonal reactions. *Acc. Chem. Res. 44,* p.666-676

39. Saxon, E., Armstrong, J.I. and Bertozzi, C.R. (2000). A "traceless" Staudinger ligation for the chemoselective synthesis of amide bonds. *Org. Lett. 2,* p.2141-2143

40. Agard, N.J., Prescher, J.A. and Bertozzi, C.R. (2004). A strain-promoted [3+2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc. 127,* p.15046-15047

41. Kamariza, M., Shieh, P., Faland, C.S., Peters, J.S., Chu, B., Rodriguez-Rivera, F.P., Babu Sait, M.R., Treuren, W.V., Martinson, N., Kalscheuer, R., Kana, B.D. and Bertozzi, C.R. (2018). Rapid detection of *Mycobacterium tuberculosis* in sputum with a solvatochromic trehalose probe. *Sci. Transl. Med. 430,* e.aam6310

42. van Kasteren, S.I., Kramer, H.B., Jensen, H.H., Campbell, S.J., Kirkpatrick, J., Oldham, N.J., Anthony, D.C and Davis, B.G. (2007). Expanding the diversity of chemical protein modification allows post-translational mimicry. *Nature. 446,* p.1105-1109

43. Goddard-Borger, E.D. and Stick, R.V. (2011). An efficient, inexpensive, and shelf-stable diazotransfer reagent: Iamidazole-1-sulfonyl azide hydrochloride. *Org. Lett. 9,* p.3797-3800

44. van Dongen, S.F., Teeuwen, R.L., Nallani, M., van Berkel, S.S., Cornelissen, J.J., Notte, R.J. and van Hest, J.C. (2009). Single-step azide introduction in proteins via an aqueous diazo transfer. *Bioconjug. Chem. 20,* p.20-23

45. Veitch, N.C. (2004). Horseradish peroxidase: A modern view of a classic enzyme. *Phytochemistry. 65,* p.249-259

46. van Hest, J.C.M., Kiick, K.L. and Tirrell, D.A. (2000). Efficient incorporation of unsaturated methionine analogues into proteins *in vivo*. *J. Am. Chem. Soc. 122,* p.1282-1288

47. Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A. and Schuman, E.M. (2006). Selective identification of newly synthesized proteins in mammalian cells using biorthogonal noncanonical amino acid tagging (BONCAT). *Proc. Natl. Acad. Sci. U.S.A. 103,* p.9482-9487

48. Saka, S.K., Wang, Y., Kishi, J.Y., Zhu, A., Zeng, Y., Xie, W., Kirli, K., Yapp, C., Cicconet, M., Beliveau, B.J., Lapan, S.W., Yin, S., Lin, M., Boyden, E.S., Kaeser, P.S., Pihan, G., Church, G.M. and Yin, P. (2019). Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nat. Biotechnol. 37,* p.1080-1090

49. Bauman, J.E., and Wang, J.C. (1964). Imidazole complexes of nickel (II), copper (II), zinc (II), and silver (I). *Inorg. Chem. 3,* p.368-373

50. Hamada, Y.Z., Cox, R., Hamada, H. (2015). Cu^{2+}-citrate dimer complexes in aqueous solutions. *J. Basic Appl. Sci. 11,* p.583-589

51. Gyliene, O., Vengris, T., Nivinskiene, O. and Binkiene, R. (2010). Decontamination of solutions containing Cu(II) and ligands tartrate, glycine and quadrol using metallic iron. *J. Hazard Mater. 175,* p.452-459

52. Boenisch, T. (2003). Handbook of immunochemical staining methods. 3rd ed. Dako Cytomation Corp., Carpinteria, CA, USA.

53. Radulescu, R.T. and Boenisch, T. (2007). Blocking endogenous peroxidases: A cautionary note for immunohistochemistry. *J. Cell. Mol. Med. 11,* p.1419

54. Adibekian, A., Martin, B.R., Chang, J.W., Hsu, K-L., Tsuboi, K., Bachovchin, D.A., Speers, A.E., Brown, S.J., Spicer, T., Fernandez-Vega, V., Ferguson, J., Hodder, P.S., Rosen, H. and Cravatt, B.F. (2012). Confirming target engagement for reversible inhibitors *in vivo* by kinetically tuned activity-based probes. *J. Am. Chem. Soc. 134,* p.10345-10348

55. van Esbroeck, A.C.M., Janssen, A.P.A., Cognetta, A.B. 3rd, Ogosawara, D., Shpak, G., van der Kroeg, M., Kantae, V., Baggelaar, M.P., de Vrij, F.M.S., Deng, H., Allarà, M., Lin, Z., van der Wel, T., Soethoudt, M., Mock, E.D., den Dulk, H., Baak, I.L., Florea, B.I., Hendriks, G., De Petrocellis, L., Overkleeft, H.S., Hankemeier, T., de Zeeuw, C.I., Di Marzo, V., Maccarone, M., Cravatt, B.F., Kushner, S.A. and van der Stelt, M. (2017). Activity-based protein profiling reveals off-target proteins of the FAAH inhibitor BIA 10-2474. *Science. 356,* p.1084-1087

56. Dana, D., Garcia, J., Bhuiyan, A.I., Rathop, P., Joo, L., Novoa, D.A., Paroly, S., Fath, K.R., Chang, E.J. and Pathak, S.K. (2019). Cell penetrable, clickable and tagless activity-based probe of human cathepsin L. *Bioorg. Chem. 85,* p.505-514

57. Sabidó, E., Tarragó, T., Niessen, S., Cravatt, B.F. and Giralt, E. (2009). Activity-based probes for monitoring post-proline protease activity. *Chembiochem. 10,* p.2361-2366

58. Kato, D., Boatright, K.M. Berger, A.B., Nazif, T., Blum, G., Ryan, C., Chehade, K.A., Salvesen, G.S. and Bogyo. M. (2005). Activity-based probes that target diverse cysteine protease families. *Nat. Chem. Biol. 1,* p.33-38

59. Greenbaum, D., Medzihradszky, K.F., Burlingame, A. and Bogyo, M. (2000). Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem. Biol. 7,* p.569–581

60. Bachovchin, D.A. and Cravatt, B.F. (2012). The pharmacological landscape and therapeutic potential of serine hydrolases. *Nat. Rev. Drug. Discov. 11,* p.52-68

61. Liu, Y., Patricelli, M.P. and Cravatt. B.F. (1999). Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci. U.S.A. 96,* p.14694–14699.

62. Widen, J.C., Tholen, M., Yim, J.J., Antaris, A., Casey, K.M., Rogalla, S., Klaassen, A., Sorger, J. and Bogyo, M. (2019). Multivariate AND-gate substrate probes as enhanced contrast agents for fluorescence-guided surgery. *BioRxiv,* 695403

63. van Rooden, E.J., Kohsiek, M., Kreekel, R., van Esbroeck, A.C.M., van den Nieuwendijk, A.M.C.H., Janssen, A.P.A., van den Berg, R.J.B.H.N., Overkleeft, H.S. and van der Stelt, M. (2018). Design and synthesis of quenched activity-based probes for diacylglycerol lipase and α,ß-hydrolase domain containing protein. *Chem. Asian. J. 13,* p.3491-3500

64. Deng, H. and van der Stelt, M. (2018). Chemical tools to modulate 2-arachidonoylglycerol biosynthesis. *Biotechnol. Appl. Biochem. 65,* p.9-15

65. Meisslitzer-Ruppitsch, C., Rörhl, C., Neumüller, J., Pavelka, M. and Ellinger, A. (2009). Photooxidation technology for correlated light and electron microscopy. *J. Microsc. 235,* p.322-335

66. Ellisman, M.H., Deerinck, T.J., Shu, X. and Sosinsky, G.E. (2012). Picking faces out of a crowd: Genetic labels for identification of proteins in correlated light and electron microscopy imaging. *Methods Cell. Biol. 111,* p.139-155

67. Rosas-Arellano, A., Villalobos-González, J.B., Palma-Tirado, L., Beltrán, F.A., Cárabez-Trejo, A., Missirlis, F. and Castro, M.A. (2016). A simple solution for antibody signal enhancement in immunofluorescence and triple immunogold assays. *Histochem. Cell. Biol. 146,* p.421-430

68. Jones, L.H., Beal, D., Selby, M.D., Everson, O., Burslem, G.M., Dodd, P., Millbank, J., Tran, T.D., Wakenhut, F., Graham, E.J. and Targett-Adams, P. (2011). In-cell click labelling of small molecules to determine subcellular localisation. *J. Chem. Biol. 4,* p.49-53

69. Potter, G.T., Jayson, G.C., Miller, G.J. and Gardiner, J.M. (2016). An updated synthesis of the diazo-transfer reagent imidazole-1-sylfonyl azide hydrogen sulfate. *J. Org. Chem. 81.* p.3443-3446

70. Hong, V., Presolski, S.I., Ma, C. and Finn, M.G. (2009). Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew. Chem. Int. Ed. Engl. 48,* p.9879-9883

71. Li, N., Lim, R.K., Edwardraja, S. and Lin, Q. (2011). Copper-free Sonogashira cross-coupling for functionalization of alkyne-encoded proteins in aqueous medium and in bacterial cells. *J. Am. Chem. Soc. 133,* p.15316-15319

72. Biagini, S.C.G., Gibsoné, T., Keen, S.E. and Stephen, P. (1998). Cross-metathesis of unsaturated α-amino acid derivatives. *J. Chem. Soc. Perkin Transactions 1, 16,* p.2485-2500

73. Dong, S., Merkel, L., Moroder, L. and Budisa, N. (2008). Convenient syntheses of homopropargylglycine. *J. Pept. Sci. 14,* p.1148-1150

74. Chenault, H.K., Dahmer, J. and Whitesides, G.M. (1989). Kinetic resolution of unnatural and rarely occurring amino acids: Enantioselective hydrolysis of N-acyl amino acids catalyzed by acylase I. *J. Am. Chem. Soc. 111,* p.6354-6364

75. Antoniou, A.N., Blackwood, S.L., Mazzeo, D. and Watts, C. (2000). Control of antigen presentation by single protease cleavage site. *Immunity, 12,* p.391-398

76. Mruk, D.D. and Cheng, C.Y. (2011). Enhanced chemiluminescence (ECL) for routine immunoblotting: An inexpensive alternative to commercially available kits. *Spermatogenesis, 1,* p.121-122

Chapter 5
Summary and future prospects

Abstract

    This Chapter describes various future applications of bioorthogonal protein expression, such as its use to synthesize bioorthogonal fluorescent proteins, and glycosylated antigen libraries.

The aim of this thesis was to develop bioorthogonal antigenic proteins and to use these as tools to study antigen processing, presentation and T cell activation.

In **Chapter 1** an introduction to the field of protein modification was given. This chapter focused on the synthesis of proteins with bioorthogonal ligation handles; in particular those that can be incorporated via modified protein expression systems.

Chapter 2 challenged the production and purification of model protein ovalbumin with the incorporation of the unnatural amino acids azidohomoalanine (L-Aha) and homopropargylglycine (L-Hpg). The metabolic labeling of the protein with L-Aha was indeed possible in BL21::MetA. Successful purification of the protein enabled the study of differences between the azidylated and the wild type version of this protein and gives one the possibility to use this modified antigenic protein for antigen presentation studies.

In **Chapter 3** the bioorthogonal incorporation of L-Aha and L-Hpg into the immunological model protein tetanus toxin C fragment (TTCF) was presented. A series of mutants of the protein was made, their ability to activate TTCF-specific T cell clones was assessed, as well as the effect of the bioorthogonal modifications on T cell activation. The latter were shown to be tolerated by the T cell receptor. This scope of this tolerance was broadened to the MHC-II restricted epitope of ovalbumin, which was assessed using a series of peptide antigens.

In **Chapter 4** horseradish peroxidase (HRP) was explored as new tool for immunohistology. HRP was subjected to diazotransfer conditions and the resulting HRP-N_3 was used in dot blot and immunoblot experiments. These were successfully performed both on purified alkylated protein and alkylated bacteria. Furthermore, a ligation reaction between the peroxidase and an alkylated activity-based probe (ABP) was presented. However, this should be further optimized, as the detection limit was reached when this method was applied in cellular ABPP studies.

With this broadening of the expertise to produce large amounts of recombinant bioorthogonal proteins, some new approaches can be envisaged. The first steps towards some of these are described below, but a thorough assessment should be performed in
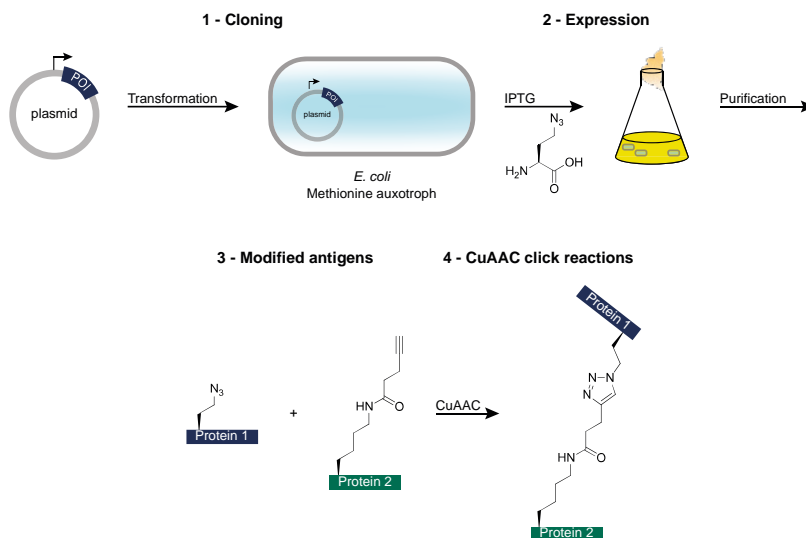
Figure 1. General strategy for the ligation of two proteins. Individual steps subject of here-described studies are the transformation of pEH101 to methionine-auxotrophic *E. coli* B834 (**1**); induced expression in the presence of the methionine analogue azidohomoalanine and subsequent purification (**2**); Combining this with the alkyne modified β-casein (**3**); and the subsequent ligation reaction, resulting in a protein-protein ligation (**4**).

future experiments.

5.1 Bioorthogonal chemistry for protein-protein ligation

The first use of this approach is to ligate two proteins together. Normal ligations rely either on the introduction of a sortag – an LPXTG tag which can be used for ligation reactions using the Sortase A enzyme – [1], or on the random ligation between cysteine (Cys) and lysine (Lys) residues. The former method is relatively low yielding and requires genetic modification of the protein; as well as a free N- and C-terminus. The most common method of Lys/Cys ligation suffers from poor control reactions, polymerization and homomultimerization. It was thus envisaged that bioorthogonal ligation reactions could be used to prevent homodimerization, whilst still getting high enough yields. This was done using the approach outlined in Figure 1: one protein is modified with an alkyne handle and the other with an azide handle and the two are ligated together using the copper(I)-catalyzed azide alkyne cycloaddition (CuAAC) [2]. This approach would give
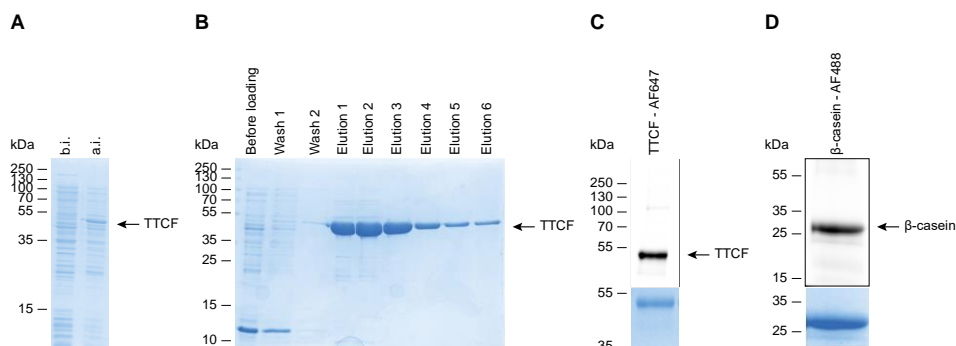
Figure 2. Expression of TTCF-Aha and fluorescent analysis of modified TTCF and β-casein. A) The before and after induction (b.i. and a.i. respectively) samples showing the expression of TTCF-Aha. B) Ni-NTA purification of TTCF-Aha, of which fractions 1-6 resulted in >95% pure protein and were subsequently used for C) fluorescent analysis by ligation to AF647 alkyne. D) Fluorescent analysis of β-casein-alkyne by ligation to AF488 azide. A-D) Proteins were resolved in a 10% or 12.5% SDS-PAGE and stained with Coomassie Brilliant Blue R-250.

far better control over the protein:protein ratio and preclude the aforementioned homodimerization.

With this aim in mind, two different model proteins, TTCF and β-casein, were modified with either an azide or alkyne handle. The former was made using the recombinant expression and purification protocols outlined in **Chapter 3** (Figure 2A and B). The second protein, β-casein, was chemically modified with an alkyne handle using pentynoic acid *N*-hydroxysuccinimide (HOSu) ester as a quick method for achieving this reagent. For this, 100 µM β-casein in 0.1 M NaHCO$_3$ pH 8.6 was combined with 1 eq. of pentynoic acid HOSu ester and this coupling took place overnight at 4°C. Subsequently, the reaction was desalted and buffer exchanged against 100 mM HEPES pH 8.4.

The accessibility of the ligation handles was first assessed using fluorophores with the opposite ligation handle, Alexa Fluor 647 (AF647) alkyne (for ligation to TTCF) and AF488 azide (for ligation to β-casein) under CuAAC-conditions. The resulting in-gel fluorescence showed that both proteins were ligated to the fluorophore (Figure 2C and D).

It was next assessed if it was possible to ligate the two proteins using the CuAAC reaction. For this, β-casein-azide and TTCF-alkyne were ligated in a 1:0.5 or in a 1:1 molar ratio in the presence of ligation mixture containing either 0, 1 or 5 eq. of Cu-
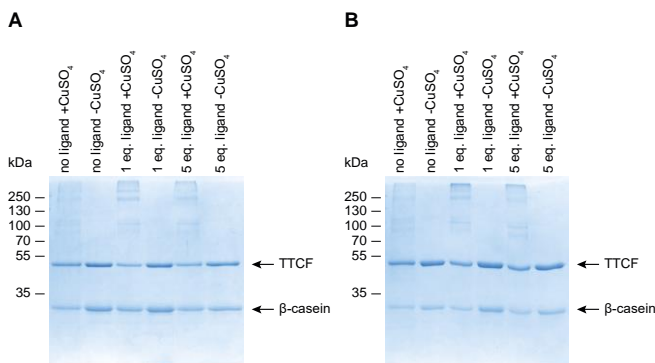
Figure 3. Ligation reaction between β-casein-alkyne and TTCF-azide. This reaction was performed in a A) 1:0.5 or B) 1:1 molar ratio in the presence of 0, 1 or 5 eq. of THPTA ligand. The 1:1 molar ratio showed a higher likelihood of aggregation of the proteins. A-B) Proteins were resolved in a 10% SDS-PAGE and stained with Coomassie Brilliant Blue R-250.

stabilizing THPTA ligand, in presence/absence of $CuSO_4$ for 15 min at rt. SDS-PAGE analysis of the reaction showed that the ligation efficiency was independent of the protein ratio and of the ligand concentration, but – as expected – only proceeded in presence of $CuSO_4$ (Figure 3A and B). It was, however, observed that an increase in the concentration of ligand present in the ligation mixture led to an increase in higher molecular weight bands (Figure 3), likely due to the presence of multiple ligation handles in both proteins. However, this approach does offer the opportunity to produce protein-protein heterodimers without the presence of homodimer contaminants.

5.2 The importance of protein glycosylation in antigen presentation

A second area in which the high yielding production of bioorthogonal proteins can be of interest is in producing homogeneous glycoforms of these proteins [3]. As for the protein dimerization reactions described above, the common approaches to produce such proteins – such as enzymatic glycan remodeling [4, 5] or lysine modification [6], are either limited in the nature and sites of modification and/or result in heterogeneous mixtures in terms of modified sites. This can have extensive immunological consequences, probably best exemplified by the disagreements regarding the role of the mannose receptor (MR) in enhancing antigen cross-presentation. Burgdorf *et al.* published the observation that glycosylated Ova-antigen can be internalized via the MR
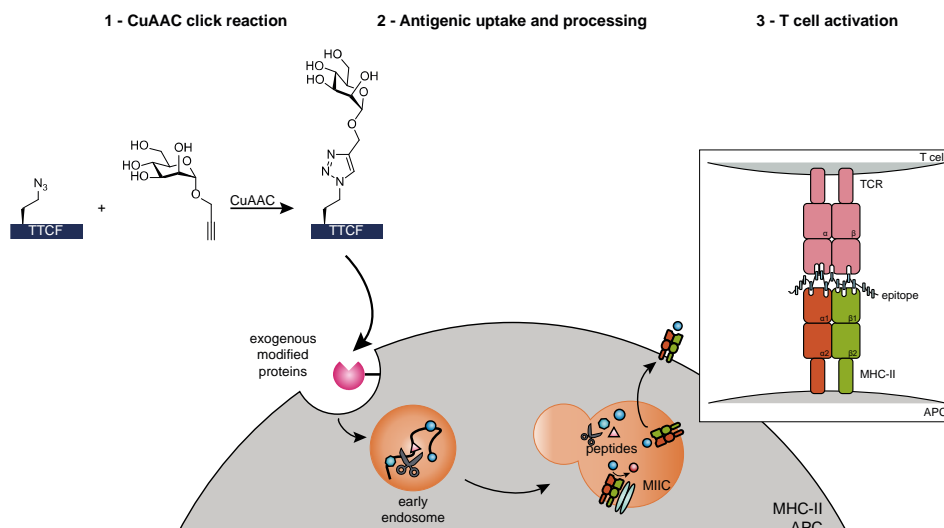
Figure 4. General strategy for MHC class II antigen presentation using glycosylated proteins. Individual steps subject of here-described studies are the azidylation of TTCF and the glycosylation of this protein by for instance alkylated mannose (**1**). This modified protein was next subject to antigenic uptake and processing, before T cell activation by means of IL-2 production was measured (**2** and **3**).

leading to routing of this antigen to a 'cross-presentation enabled'-endosome, leading to enhanced CD8 T cell activation [7]. However, other studies of the same system [8] showed quite the opposite effect for the same system. The difference can only be rationalized by the fact that different glycoform mixtures and modifications were used in these systems and that these have differential interactions with the immune lectins found on APCs.

The MR receptor has several binding domains which are able to recognize several carbohydrate moieties, besides mannose monosaccharides. Examples of these are fucose, glucose, *N*-acetyl glucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), mannose disaccharides, and more complex trisaccharides such as Lewis[x] and tri-GlcNAc [9-12]. For the mannose disaccharides the α1-6, and α1-3 linkages are preferred over α1-2 and α1-4 linkages [9, 10].

The effect of the trisaccharides was evaluated by the group of Van Kooijk, using ovalbumin as a model antigen [13]. In this study, they modified ovalbumin with either Lewis[A] or tri-GlcNAc and illustrated that both glycosylated proteins showed enhanced

binding to the MR compared to the wild type protein. Furthermore, glycosylation of the protein showed an increased MHC class I-mediated cross presentation.

With the aim to analyze how glycosylation influences MHC presentation and T cell activation – as depicted in Figure 4 – the bioorthogonal glycosylation of antigens was attempted – much along the lines of the previously described strategy [14]. For this, it was decided to use an azidylated TTCF (produced as earlier described in Chapter 2 and section 5.1), which was ligated to different alkylated glycans using the CuAAC [2, 3].

The azidylated protein was next modified with three different alkylated sugar moieties which are known to bind the mannose receptor, namely propargylated GlcNAc [9], propargylated mannose [9] and propargylated 4-SO_4-GalNAc [15]. For this, a literature protocol was followed, with minor adjustments [3]. In short: the azidylated protein was combined with the corresponding glycan and buffered with sodium phosphate pH 8.3. The reaction – under inert conditions – was started by the addition of a Cu(I) solution. After brief vortexing, the reaction was left for 5 min at rt. After two more additions of the Cu(I) solution, the protein was immediately desalted and buffer exchanged to PBS.

With the aim to analyze if every azide moiety of the protein was modified with the respective glycan, it was decided to expose the unmodified protein (Met), the azidylated protein (Aha) and the three resulting glycosylated proteins to a CuAAC ligation mixture containing AF647 alkyne. SDS-PAGE-fluorimetry showed no fluorescent signal for the glycosylated proteins and the negative control (Met), whereas the positive control (Aha) did show the AF647 signal (Figure 5A). It was therefore assumed that every L-Aha moiety in the protein was modified with the sugar residue. However, in the future this must be confirmed, for instance using mass spectrometry analysis.

Next, the immunogenicity of these modified proteins was analyzed. For this, the glycosylated proteins were given to bone marrow derived dendritic cells (BMDCs), which acted as APCs in this experiment. After 4 hours of pulse, the cells were washed twice with PBS and the TTCF-specific 2F2 T cell hybridoma was added. After overnight chase, T cell activation was measured via an IL-2 readout. This showed no significant difference between the glycosylated proteins (Figure 5B).
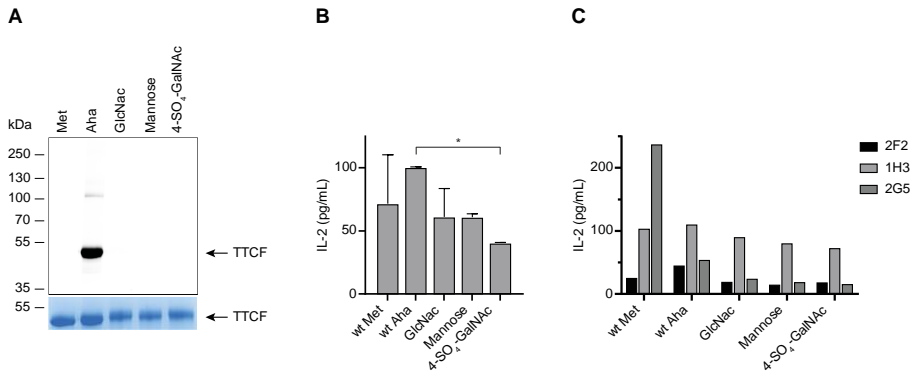
Figure 5. T cell activation by glycosylated proteins. A) Representative in-gel fluorescence after AF647 alkyne ligation to the differently derived azidylated or glycosylated proteins. Proteins were resolved in a 7.5% SDS-PAGE and stained with Coomassie Brilliant Blue G-250. B) BMDCs were pulsed for 4 hours with different concentrations of TTCF wild type and glycosylated versions, before being presented to T cell hybridoma 2F2. The produced IL-2 concentration was not significantly different for any of the glycosylated proteins. C) The same experiment as described for B) presented to T cell hybridomas 2F2, 1H3 and 2G5. 2G5 did not produce IL-2 when using the modified protein. Data was expressed as mean (± SD) (N = 1 or N = 2, n = 3). Statistical analysis was performed using Tukey's multiple comparisons test. * P ≤ 0.05.

In addition, it was analyzed whether the T cell activation of glycosylated proteins was epitope specific. Therefore, an initial study was performed in which besides the activation of the 2F2, also the activation of the 1H3 and the 2G5 T cell hybridomas was measured. This showed that the IL-2 levels of both the 2F2 and 1H3 T cells remained mostly unaltered upon activation when using the glycosylated proteins (Figure 5C). However, in line with the expectations (see **Chapter 2**), the 2G5 epitope showed no T cell activation when any of the modified proteins was used.

The experiments described in this section showed that glycosylation of TTCF does not alter the MHC class II antigen response. This was earlier reported for Ova-sulfo-Le[A] and Ova-tri-GlcNAc [13], however for these glycoproteins it was shown that MHC class I antigen presentation was enhanced. Therefore, it is recommended to evaluate the glycosylation of recombinant ovalbumin in an antigen presentation assay, using both OT-I and OT-II as responding T cells.

5.3 Bioorthogonal fluorescent bacteria

Fluorescent proteins have been transformative in the study of living systems [16, 17]. GFP is the most commonly used fluorescent protein, exciting mostly at 395 nm, and
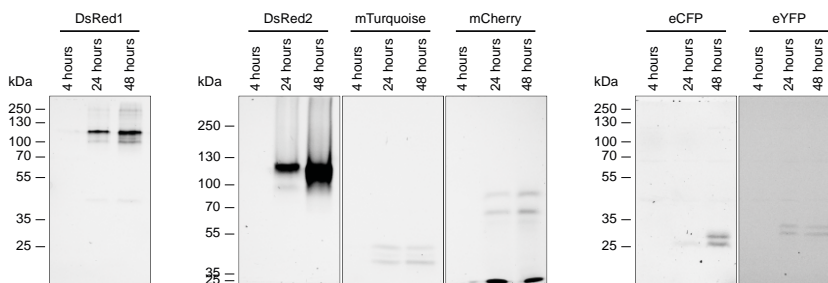
Figure 6. In-gel fluorescence of the constitutively expressed fluorophores. In-gel fluorescence of DsRed1, DsRed2, mTurquoise, mCherry, eCFP and eYFP after 4, 24 and 48 hours of culture growth. Proteins were resolved in a 10% SDS-PAGE and stained with Coomassie Brilliant Blue G-250.

emitting at 509 nm [16]. In the study of antigen processing, they can also serve as excellent markers: the loss of fluorescence can be used to determine whether a protein or bacterium is degraded [18]. After this degradation, however, the protein fragments are invisible. It has previously been shown that bioorthogonal amino acids incorporated into recombinant proteins served as excellent markers to study precisely this later degradation and that combining these two types of reagents would yield an informative tool to study degradation: The presence of both fluorescent signals in a cell could indicate the presence of intact fluorescent protein in a particular organelle, whereas the presence of only the bioorthogonal label would indicate protein degradation [19, 20].

Within the framework of studying the degradation of bacteria by APCs, a variety of fluorescent protein-producing bacteria – also capable of BONCAT labeling – were produced. For this, expression vectors for DsRed1, DsRed2, mTurquoise, mCherry, eCFP and eYFP, were produced and expression of the fluorescent proteins was first assessed in *Escherichia coli* (*E. coli*) B834(DE3). For this, the genes coding for each of the aforementioned fluorophores: DsRed1, DsRed2_S4T, mTurquoise, mCherry, eCFP and eYFP were cloned behind the hns promoter as earlier described by Van Elsland *et al.* [18]. Next, the resulting plasmids were transformed into *E. coli* B834 and protein expression was monitored. For this, *E. coli* B834 bearing the plasmids were grown in LB medium and after 4, 24 and 48 hours of growth samples were taken. SDS-PAGE analysis showed that all of the mentioned fluorophores were constitutively expressed (Figure 6). DsRed1 and DsRed2, mCherry and eYFP needed at least 24 hours to become fluorescent, whereas
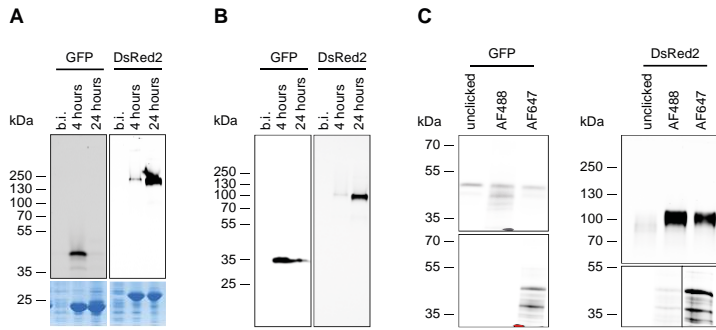
Figure 7. In-gel fluorescence of the recombinant fluorophorescent proteins. A) In-gel fluorescence of GFP and DsRed2 expressed in LB medium and Coommassie staining (lower panel). B) In-gel fluorescence of GFP and DsRed2 expressed in the presence of L-Aha. C) Left: Fluorescence analysis of GFP: AF488 analysis in the upper panel and AF647 analysis in the lower panel. Right: Fluorescence analysis of DsRed2: AF550 analysis in the upper panel, AF488 in the left lower panel and AF647 in the right lower panel. A-C) Proteins were resolved in a 10% SDS-PAGE and stained with Coomassie Brilliant Blue G-250.

mTurquoise and eCFP already showed the emergence of fluorescence after 4 hours. Of note is the visualization of both DsRed1 and DsRed2 in its tetrameric form [21, 22] and mCherry in both its monomeric and weak dimeric form [23]. With the fluorescence of the constitutively expressed proteins confirmed, the constructs can now be used in for instance lysosomal degradation experiments.

Besides the constitutively expressed fluorescent proteins, it was also decided to prepare two inducible constructs, containing either DsRed2_S4T or GFP_A206K. Both were expressed in *E. coli* with and without the incorporation of L-Aha using the standard protocol (Chapter 2 and 3 and sections 1 and 2). Fluorescence analysis of the expression samples showed the presence of GFP and DsRed2 when grown in LB medium (Figure 7A) and when expressed in the presence of L-Aha (Figure 7B). Coomassie staining confirmed expression and revealed that DsRed2 is only fluorescent in its tetrameric folding as described in literature [21].

Next, both cell lysates containing wild type and azidylated GFP and DsRed2 were subjected to a CuAAC ligation mixture containing either AF488 alkyne or AF647 alkyne. The resulting fluorescent image showed as expected an overlap between AF488 and GFP (Figure 7C; left). Surprisingly, no overlap was seen between AF647 and GFP. In the presence of the reducing conditions of the CuAAC ligation mixture, GFP diffused in

multiple protein bands. For DsRed2, the tetrameric protein did not ligate to the fluorophores, whereas the monomeric version did ligate to both (Figure 7C; right). In future, the ligation reactions should be confirmed using either LC-MS/MS and/or Native PAGE analysis.

Acknowledgements

Materials and methods
General

All chemicals and reagents were purchased at Sigma-Aldrich, Alfa Aesar, Acros, Merck or VWR, unless stated otherwise. Reagents were used without further purification. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) materials were purchased at Bio-Rad.

Solutions

PBS contained 5 mM KH_2PO_4, 15 mM Na_2HPO_4, 150 mM NaCl, pH 7.4. Laemmli sample buffer 4* contained 60 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.01% (v/v) bromophenol blue.

Compound synthesis
Chemical modification of β-casein

Dissolve pentynoic acid HOSu ester to a concentration of 100 mM in DMSO (19.5 mg/mL). Dissolve β-casein to a concentration of 100 µM (2.4 mg/mL) in 0.1 M $NaHCO_3$ pH 8.6 and add 1 eq. of pentynoic acid HOSu ester. After overnight coupling at 4°C, the reactions were desalted over a Zeba Spin Desalting Column (Thermo Scientific), 40 kDa MWCO according to manufacturer's protocol and simultaneously buffer exchanged against 100 mM HEPES pH 8.4. The modified protein was stored at -20°C until further use.

Strains and plasmids

E. coli strains XL10 and B834(DE3) were used as cloning and expression strains respectively. Plasmid pEH101 containing full-length *TTCF* with a linker sequence was a gift from the Watts lab [24] and was used for the expression of 10His-TTCF and contained the IPTG-inducible T7 promoter.

Cloning
TTCF

pEH101 containing full-length *TTCF* with a linker sequence was a gift from the Watts lab [24]. To obtain pET16b_10His-TTCF, the DNA fragment encoding TTCF was amplified by PCR from the pEH101 plasmid. This fragment was ligated into the pET16b

vector using the XbaI and BamHI restriction sites. All sequences were verified by Sanger sequencing (Macrogen).

Table 1: List of primer sequences

#	primer name	sequence 5' -> 3'
p1	TTCF_fwd	CGCAAAATTCCCTCTAGAATAATTTTTG
p2	TTCF_rev	GTTAGCAGCCGGATCCTTAGTCGTTGG
p3	T7_GFP_fwd	GGCGGCCGTCTCCCATGAGTAAAGGAGAAGAAC
p4	T7_GFP_rev	GGCGGCGGATCCTTATTTGTATAGTTCATCC
p5	T7_DsRed2_S4T_fwd	GGCGGCCGTCTCCCATGGCCTCCACCGAGAACG
p6	T7_DsRed2_rev	GGCGGCCGTCTCGGATCCTTTATCTAGATCCGGTGG
p7	hns_DsRed1_fwd	GGCGGTATGCATATGGTGCGCTCCTCC
p8	hns_DsRed1_rev	GGCGGTACGCGTCTACAGGAACAGGTG
p9	hns_DsRed2_S4T_fwd	GGCGGTATGCATATGGCCTCCaCCGAG
p10	hns_DsRed2_rev	GGCGGTACGCGTTTATCTAGATCCGG
p11	hns_mCherry_fwd	GGCGGTATGCATATGGTGAGCAAGGGC
p12	hns_mCherry_rev	GGCGGTACGCGTCTAGACTCGAGATCTG
	hns_mTurqouise_fwd	
p13	hns_eCFP_fwd	GGCGGTATGCATATGGTGAGCAAGGGC
	hns_eYFP_fwd	
p14	hns_mTurqouise_rev	
	hns_eCFP_rev	GGCGGTACGCGTTTACTTGTACAGCTC
	hns_eYFP_rev	

Inducible fluorophores

To obtain pET16b_GFP and pET16b_DsRed2_S4T, the DNA fragments encoding the fluorophores were amplified by PCR. Using this PCR reaction, DsRed2 was mutated to DsRed2_S4T, to enhance fluorescent signal [25]. The resulting fragments were ligated into the pET16b vector using the NcoI and BamHI restriction sites. All sequences were verified by Sanger sequencing (Macrogen).

Constitutively expressed fluorophores

To obtain the constitutively expressed constructs containing DsRed1, DsRed2_S4T, mTurquoise, mCherry, eCFP and eYFP, the DNA fragments encoding the fluorophores were amplified by PCR. Using this PCR reaction, DsRed2 was mutated to DsRed2_S4T, to enhance fluorescent signal [25]. The resulting fragments were ligated into the hns_GFP vector [18] using the NsiI and MluI restriction sites. All sequences were verified by Sanger sequencing (Macrogen).

Bacterial growth and protein expression

Expression of TTCF, GFP and DsRed2_S4T

An overnight culture of B834 containing the corresponding plasmid was diluted 1:100 in LB medium containing 50 µg/mL ampicillin (and 1% glucose in case of TTCF).

Cells were grown at 37°C, 180 rpm to an OD_{600} of ~0.8 and sedimented (3428 rcf, 15 min, 4°C) before being resuspended in LB medium containing 50 μg/mL ampicillin. The culture was induced with IPTG (1 mM final concentration) and expression took place for 4 hours or ON at 30°C, 130 rpm. After overnight expression, cells were pelleted and washed once with PBS. Pellets were stored at -80°C until protein purification. The before and after induction samples were suspended in 1* Laemmli buffer and the protein fractions were combined with 4* Laemmli buffer and all resolved in a 10% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific), before scanning Cy3 and Cy5 multichannel settings (605/50 and 695/55 filters, respectively; ChemiDoc™ MP System, Bio-Rad) in case of the fluorescent proteins. Coomassie staining (Coomassie Brilliant Blue G-250) was used for protein analysis.

Expression of azidylated TTCF, GFP and DsRed2_S4T

An overnight culture of B834 containing the corresponding plasmid was diluted 1:100 in LB medium containing 50 μg/mL ampicillin (and 1% glucose in case of TTCF). Cells were grown at 37°C, 180 rpm to an OD_{600} of ~0.8 and sedimented (3428 rcf, 15 min, 4°C) before being washed twice with SelenoMet medium (Molecular Dimensions). Cells were resuspended in SelenoMet medium containing 50 μg/mL ampicillin and depleted for 30 min at 37°C, 180 rpm before 30 min depletion at 30°C, 130 rpm. After this azidohomoalanine (L-Aha TFA salt; 72 mg/L) was added to the culture and 15 min later the culture was induced with isopropyl-β-D-thiogalactopyranoside (IPTG; 1 mM final concentration) and expression took place ON. After overnight expression, cells were pelleted and washed once with PBS. Pellets were stored at -80°C until protein purification. The before and after induction samples were suspended in 1* Laemmli buffer and the protein fractions were combined with 4* Laemmli buffer and all resolved in a 10% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific), before scanning Cy3 and Cy5 multichannel settings (605/50 and 695/55 filters, respectively; ChemiDoc™ MP System, Bio-Rad) in case of the fluorescent proteins. Coomassie staining (Coomassie Brilliant Blue G-250) was used for protein analysis.

Chapter 5

Constitutive expression of DsRed1, DsRed2_S4T, mTurquoise, mCherry, eCFP and eYFP

An overnight culture of B834(DE3) containing the corresponding plasmid was diluted 1:100 in LB medium containing 50 μg/mL ampicillin. Cells were grown at 37˚C, 180 rpm and samples were taken after 4, 24 and 48 hours of incubation.

The samples were suspended in 1* Laemmli buffer and the protein fractions were combined with 4* Laemmli buffer and all resolved in a 10% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific), before scanning Cy2, Cy3 and Cy5 multichannel settings (532/528, 605/50 and 695/55 filters, respectively; ChemiDoc™ MP System, Bio-Rad). Coomassie staining (Coomassie Brilliant Blue G-250) was used for protein analysis.

Protein purification
Purification of azidylated TTCF

Adjusted from Antoniou *et al.* (2000) [24]. A 1 L culture pellet was resuspended in 15 mL lysis buffer containing 100 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 1 mg/mL lysozyme, 250 U benzonase. Lysis took place for 20 min at rt, before separating the soluble from the insoluble proteins (10000 rcf, 15 min, 4˚C). The supernatant was filtered over a 0.2 μm filter (Filtropur S, Sarstedt) and loaded on a 2 mL His-column (Ni-NTA Agarose; Qiagen). After loading at 4˚C for 2 hours, the column was washed twice with 2.5 CV of IMAC25 buffer (100 mM Tris-HCl pH 8.0, 500 mM NaCl, 25 mM imidazole), before eluting with 7 CV at 200 mM imidazole. The protein fractions were monitored by A_{280} values and resolving samples on a 10% SDS-PAGE along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue G-250) revealed the purity of the protein. The fractions containing the most pure protein with an A_{280} value over 0.1 were combined and extensively dialysed (6 - 8 kDa MWCO, 3.3 mL/cm, FisherBrand or 12 - 14 kDa MWCO, 2 mL/cm, Spectra/Por) against PBS. The protein was concentrated over an Amicon spin filter 10 kDa MWCO, before being aliquoted, flash frozen and stored at -80˚C until further use.

Alternatively, the lysed cells were loaded on a 5 mL His-column (Ni-NTA Superflow Cartridge; Qiagen). After loading, the column was washed with 5 CV of IMAC25 buffer (100 mM Tris-HCl pH 8.0, 500 mM NaCl, 25 mM imidazole), before eluting with a 15 CV gradient 25 - 500 mM imidazole. Subsequently followed by the protocol described above.

Ligation reactions

Protein-protein

β-casein-alkyne and TTCF-Aha were ligated via CuAAC reaction. 0.5 or 1 eq. of TTCF-Aha was diluted in 100 mM HEPES pH 8.4 and combined 2:1 (v/v) with click mix (containing 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand, 30 mM aminoguanidine-HCl and 1 eq. of β-casein-alkyne in 88 mM HEPES pH 7.2, final concentration in click mix). After 15 min ligation at rt in the dark, the reaction was quenched by the addition of 4* Laemmli buffer. Samples were resolved in a 10% SDS-PAGE gel along with PageRuler™ Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue R-250) was used to examine protein ligation.

Protein-glycan

Adjusted from van Kasteren *et al.* (2007) [3]. The reaction should be executed under inert conditions and therefore nitrogen was bubbled through each solution before use.

Dissolve ligand (39.1 mg/mL in MQ:acetonitrile 1:1, 90 mM) and CuBr (10 mg/mL in acetonitrile, 70 mM). Combine prior to addition in a 5:6 ratio respectively. Dilute the protein to 2 mg/mL in PBS (500 μL). Add a dissolved glycan solution (250 μL, 31.2 mM, 50 eq. in MQ water) and 100 μL of sodium phosphate, pH 8.3. To this a Cu(I)/ligation solution (25 μL) was added and after 10 sec of vortexing, left for 5 min at rt. This was repeated another two times using a freshly prepared Cu(I)/ligation solution. After the third time, the glycosylated protein was immediately buffer exchanged over a Zeba Spin Desalting Column (Thermo Scientific), 7 kDa MWCO according to manufacturer's protocol and simultaneously buffer exchanged against PBS. Protein concentrations were determined using Quick Start™ Bradford Protein Assay (Bio-Rad) and the modified proteins were stored at 4˚C or flash-frozen and stored at -80˚C until further use.

Protein-fluorophore

Proteins were ligated via CuAAC reaction to AF647-alkyne. The protein was diluted to a concentration of 0.5 μg/μL in 100 mM HEPES pH 7.2 or MQ water and combined 2:1 (v/v) with click mix (containing 3 mM copper sulfate, 30 mM sodium ascorbate, 3 mM THPTA ligand, 30 mM aminoguanidine-HCl and 14 μM AF647-alkyne in 88 mM HEPES pH 7.2, final concentration in click mix). After 1 h ligation at rt in the dark, the reaction

was quenched by the addition of 4* Laemmli buffer. Samples were resolved in a 7.5% or a 10% SDS-PAGE gel along with PageRuler^{TM} Plus Protein Marker (Thermo Scientific). Coomassie staining (Coomassie Brilliant Blue G-250) was used as protein loading control.

Cell culture

General

T cell hybridomas 2F2, 1H3 and 2G5 were a kind gift of C. Watts and were tested on regular basis for mycoplasma contamination. Cultures were discarded after 2 months of use. The cells were cultured at 37˚C under 5% CO_2 in RPMI 1640 (containing 25 mM HEPES) supplemented with stable glutamine (2 mM), heat inactivated fetal calf serum (10% v/v; Biowest), pyruvate (1 mM), NEAA (1x), β-mercaptoethanol (50 µM), penicillin (200 IU/mL; Duchefa) and streptomycin (200 µg/mL; Duchefa). Cells were passaged every 2 - 3 days.

BMDCs were isolated from the bone marrow from the tibia and femurs of a C57BL/6 mice. Cells were cultured at 37˚C under 5% CO_2 in RPMI medium supplemented with stable glutamine (2 mM), heat inactivated fetal calf serum (10% v/v; Biowest), β-mercaptoethanol (20 µM), GM-CSF (20 ng/mL, Immunotools), penicillin (200 IU/mL; Duchefa) and streptomycin (200 µg/mL; Duchefa). Cells were passaged on day 4 and day 7. Generally cell experiments were performed on day 7 or 8.

Antigen presentation assay

BMDCs (50000 cells/well) were seeded in a 96-well tissue-culture treated microtiter plate. Adherence was allowed at 37˚C under 5% CO_2 for at least 1 h prior to the addition of peptides or proteins at indicated concentrations. The cells were incubated with 0.1 mg/mL (final conc.) of the antigens for 4 hours, followed by 2 times washing with PBS. The corresponding T cell hybridoma (50000 cells/well) was added to the pulsed BMDCs and co-cultured for 20 h for antigen recognition and IL-2 production by the T cells at 37˚C under 5% CO_2. After overnight incubation, cells were sedimented by centrifugation (360 rcf, 5 min, rt) and supernatant was transferred to a new 96-wells plate. Stimulation of the T cell hybridoma was measured by IL-2 readout using an ELISA assay according to manufacturer's protocol (Invitrogen).

Statistical analysis

Replicates shown represent biological replicates and data represent means ± SD. All statistical analysis was determined using GraphPad Prism® 8 or Microsoft Excel 2016. **** $P \leq 0.0001$; *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; NS if $P > 0.05$.

Chapter 5

References

1. Popp, M.W., Antos, J.M., Grotenberg, G.M., Spooner, E. and Ploegh, H.L. (2007). Sortagging: A versatile method for protein labelling. *Nat. Chem. Biol. 3,* p.707-708

2. Agard, N.J., Prescher, J.A. and Bertozzi, C.R. (2004). A strain-promoted [3+2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc. 126,* p.15046-15047

3. van Kasteren, S.I., Kramer, H.B., Gamblin, D.P. and Davis, B.G. (2007). Site-selective glycosylation of proteins: Creating synthetic glycoproteins. *Nat. Protoc. 2,* p. 3185-3194.

4. Moremem, K.W., Tiemeyer, M. and Nairn, A.V. (2012). Vertebrate protein glycosylation: Diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol. 13,* p.448-462

5. Ploeg, H. and Neefjes, J.J. (1990). Protein glycosylation. *Curr. Opin. Cell Biol. 2,* p.1125-1130

6. Oliver, C.M., Melton, L.D. and Stanley, R.A. (2006). Creating proteins with novel functionality via the Maillard reaction: A review. *Crit. Rev. Food. Sci. Nutr. 46,* p.337-350

7. Burgdorf, S., Kautz, A., Böhnert, V., Knolle, P.A. and Kurts, C. (2007). Distinct pathways of antigen uptake and intracellular routing in CD4 and CD8 T cell activation. *Science. 316,* p.612-616

8. Segura, E., Gupta, N., Albiston, A.L., Wicks, I.P., Chai, S.Y. and Villandangos J.A. (2010). Reply to Burgdorf *et al.*: The mannose receptor is not involved in antigen cross-presentation by steady-state dendritic cells. *Proc. Natl. Acad. Sci. U.S.A. 107,* E50-51

9. van Die, I. and Cummings, R.D. (2017). The mannose receptor in regulation of helminth-mediated host immunity. *Front. Immunol. 8,* 1677

10. Kéry, V., Křepinský, J.J.F., Warren, C.D., Capek, P. and Stahl, P.D. (1992). Ligand recognition by purified human mannose receptor. *Arch. Biochem. Biophys. 298,* p.49-55

11. Leteux, C., Chai, W., Loveless, R.W., Yuen, C-T., Uhlin-Hansen, L., Combarnous, Y., Jankovic, M., Maric, S.C., Misulovin, Z., Nussenzweig, M.C. and Feizi, T. (2000). The cysteine-rich domain of the macrophage mannose receptor is a multispecific lectin that recognizes chondroitin sulfates A and B and sulphated oligosaccharides of blood group Lewis[a] and Lewis[x] types in addition to the *N*-glycans of lutropin. *J. Exp. Med. 191,* p.1117-1126

12. Liu, Y., Chirino, A.J., Misulovin, Z., Leteux, C., Feizi, T., Nussenzweig, M.C. and Bjorkman, P.J. (2000). Crystal structure of the cysteine-rich domain of mannose receptor complexed with a sulphated carbohydrate ligand. *J. Exp. Med. 191,* p.1105-1115

13. Singh, S.K., Streng-Ouwehand, I., Litjens, M., Kalay, H., Burgdorf, S. Saeland, E., Kurts, C., Unger, W.W. and van Kooijk, Y. (2011). Design of neo-glycoconjugates that

target the mannose receptor and enhance TLR-independent cross-presentation and Th1 polarization. *Eur. J. Immunol. 41,* p.916-925

14. Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A. and Schuman, E.M. (2006). Selective identification of newly synthesized proteins in mammalian cells using biorthogonal noncanonical amino acid tagging (BONCAT). *PNAS. 103,* p.9482-9487

15. Liu, Y., Misulovin, Z. and Bjorkman, P.J. (2001). The molecular mechanism of sulphated carbohydrate recognition by the cysteine-rich domain of mannose receptor. *J. Mol. Biol. 305,* p.481-490

16. Snapp, E. (2005). Design and use of fluorescent fusion proteins in cell biology. *Curr. Protoc. Cell Biol.* Chapter: Unit-21.4

17. Dunn, K.W., Kamocka, M.M. and McDonals, J.H. (2011). A practical guide to evaluating colocalization in biological microscopy. *Am. J. Physiol. Cell Physiol. 300,* c.723-742

18. van Elsland, D.M., Bos, E., de Boer, W., Overkleeft, H.S., Koster, A.J. and van Kasteren, S.I. (2016). Detection of bioorthogonal groups by correlative light and electron microscopy allows imaging of degraded bacteria in phagocytes. *Chem. Sci. 7,* p.752-758

19. van Elsland, D.M., Pujals, S., Bakkum, T., Bos, E., Oikonomeas-Koppasis, N., Berlin, I., Neefjes, J., Meijer, A.H., Koster, A.J., Albertazzi, L. and van Kasteren, S.I. (2018). Ultrastructural imaging of salmonella-host interactions using super-resolution correlative light-electron microscopy of bioorthogonal proteins. *Chembiochem. 19,* p.1766-1770

20. Araman, M.C., Pieper-Pournara, L., van Leeuwen, T., Kampstra, A.S.B., Bakkum, T., Marqvorsen, M.H.S., Nascimento, C.R., Groenewold, G.J.M., Wulp, W., Camps, M.G.M., Overkleeft, H.S., Ossendorp, F.A., Toes, R.E.M. and van Kasteren, S.I. (2019). Bioorthogonal antigens allow the unbiased study of antigen processing and presentation. *BioRxiv.* p.439323

21. Bevis, B.J. and Glick, B.S. (2002). Rapidly maturing variants of the discosoma red fluorescent protein. *Nat. Biotechnol. 20,* 83-87

22. Wall, M.A., Socolich, M. and Ranganathan, R. (2000). The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nat. Struct. Biol. 7,* p.1133-1138

23. Suzuki, T., Arai, S., Takeuchi, M., Sakurai, C., Ebana, H., Highashi, T., Hashimoto, H., Hatsuzawa, K. and Wada, I. (2012). Development of cysteine-free fluorescent proteins for the oxidative environment. *PLoS One. 7,* e.37551

24. Antoniou, A.N., Blackwood, S-L., Mazzeo, D. and Watts, C. (2000). Control of antigen presentation by a single protease cleavage site. *Immunity. 12,* p.391-398

25. Sörenson, M., Lippuner, C., Kaiser, T., Mißlitz, A., Aebischer, T. and Bumann, D. (2003). Rapidly maturing red fluorescent protein variants with stronly enhanced brightness in bacteria. *FEBS lett. 552,* p.110-114

Nederlandse samenvatting

Het onderzoek dat in dit proefschrift wordt beschreven richt zich op het gebruik van bioorthogonaal gelabelde eiwitten in immunologische analyses. In deze thesis wordt vooral gefocust op MHC-II antigenen.

In **hoofdstuk 1** wordt een algemene omschrijving van eiwitexpressie gegeven. Hierna wordt besproken hoe belangrijk posttranslationele modificaties zijn en hoe eiwitten posttranslationeel gemodificeerd kunnen worden door gebruik te maken van bioorthogonale eiwitten. Vervolgens worden verschillende methodes beschreven waarmee eiwitten bioorthogonaal gemodificeerd kunnen worden. Allereerst wordt er aandacht besteed aan het modificeren van de N- en C-terminus en de zijketens van de aminozuren waaruit eiwitten zijn opgebouwd. Dan volgt een introductie in verschillende eiwitexpressie systemen en de stammen waarin recombinante eiwitten tot expressie kunnen worden gebracht. Achtereenvolgens worden twee methoden besproken waarbij onnatuurlijke aminozuren, die een orthogonale groep bevatten, ingebouwd kunnen worden. Als eerste wordt het gebruik van onder andere amber codon suppressie besproken om vervolgens metabole labeling te benoemen. Tot slot wordt in het eerste hoofdstuk een aantal toepassingen van bioorthogonale eiwitten besproken.

In **hoofdstuk 2** wordt de expressie en zuivering beschreven van ovalbumine (Ova) met de inbouw van L-Aha en L-Hpg. Allereerst wordt het N. Del Cid construct gebruikt om 6His-TEV-Ova tot expressie te brengen. Het bleek mogelijk om dit eiwit met de inbouw van deze onnatuurlijke aminozuren te produceren en te zuiveren, maar dit gaf een lage opbrengst. Om deze opbrengst te verhogen wordt het gen aan een 10His-tag gekloneerd. De expressie van 10His-TEV-Ova resulteert in de productie van zogenaamde inclusion bodies. Met het toevoegen van solubiliserende tags kon dit niet veranderd worden en het bleek ook niet mogelijk om het eiwit te zuiveren uit de inclusion bodies wanneer L-Hpg in het eiwit was ingebouwd. Een laatste eiwit expressie wordt geprobeerd waarbij het eiwit zonder tags tot expressie wordt gebracht. Dit resulteerde in de expressie van het eiwit in LB medium en met de inbouw van L-Aha, maar dit laatste alleen in BL21::MetA en niet in de veelgebruikte B834(DE3).

Na de in **hoofdstuk 2** besproken moeilijkheden met de expressie en zuivering van bioorthogonaal Ova wordt in **hoofdstuk 3** tetanus toxin C fragment (TTCF) gebruikt om antigeen presentatie te bestuderen. Hiervoor worden de voorspelde negen aminozuren, – de onderstreepte aminozuren in de volgende sequentie: SGFNSSV<u>ITYPDAQLV</u>P – die binden tussen het MHC klasse II epitoop en de T-cel receptor van de 2F2 T-cel hybridoma, gemuteerd naar methionine. Dit resulteert in 9 mutant genen, die vervolgens gekloneerd worden in het expressieplasmide pET16b. Na transformatie van deze plasmiden in B834(DE3) is het mogelijk om deze tot expressie te brengen in LB medium of SelenoMet medium waaraan L-Aha is toegevoegd. Dit resulteert in 20 (mutant) eiwitten die, na zuivering, gebruikt kunnen worden voor diverse antigeen experimenten. Als eerste wordt bekeken of de bioorthogonaal gelabelde aminozuren afgebroken worden in BMDCs, waarna, na bevestiging van de afbraak, wordt gekeken of het eiwit in staat is om de corresponderende T-cel hybridoma te activeren. Dit bleek het geval voor zowel de azide-gelabelde eiwitten als voor de methionine bevattende eiwitten, alleen voor beiden bleek de IL-2 productie te laag om verschil te kunnen ontdekken tussen eventueel geblokkeerde antigeen presentatie en wild-type niveaus. Om deze reden werden verschillende MHC-II peptide varianten gesynthetiseerd gebaseerd op residuen 323-339 (ISQAV<u>HAAHAEINE</u>AGR) van het modeleiwit Ova. Hierbij bevatte elk peptide een andere aminozuur vervanging naar L-Aha. Een antigeen presentatie assay via A20s naar DO11.10 liet zien dat verschillende mutant peptiden nog in staat waren om de T-cellen te activeren. Het veelbelovendste peptide, ISQAVHAAHAEINAhaAGR, werd vervolgens gebruikt voor een immunoprecipitatie tegen MHC-II, waarvan de eerste resultaten veelbelovend lijken.

In **hoofdstuk 4** wordt een nieuwe methode geïntroduceerd voor het visualiseren van antigenen in immunocytochemie en immunohistochemie. Voor deze methode wordt gebruik gemaakt van het bekende horseradish peroxidase (HRP). Als eerste wordt gekeken of de diazotransfer zoals beschreven bij van Dongen *et al.* verbeterd kan worden. Het bleek dat de modificatie van HRP tot HRP-N_3 via de methode van van Dongen bijna optimaal is. Vervolgens wordt gekeken of HRP-N_3 geligeerd kan worden aan gealkyleerd TTCF en specificiteit van de reactie wordt gecontroleerd door dit eiwit

in cellysaat te verdunnen. Na bevestiging van deze specificiteit wordt de ligatiereactie ook uitgevoerd met verschillende concentraties gealkyleerde bacteriën. Dit liet een duidelijk verschil zien tussen niet gemodificeerde en volledig gealkynyleerde bacteriën. Als laatste wordt in dit hoofdstuk HRP-N$_3$ gebruikt om gealkyleerde probes te visualiseren na inhibitie van het bijbehorende eiwit. Dit resulteerde niet in een verschil tussen signaal en achtergrondsignaal.

List of publications

Van Dalen, F.J., Bakkum, T., van Leeuwen, T., Groenewold, G.J.M., Deu, E., Koster, A., van Kasteren, S.I. and Verdoes, M. (2020). Application of a highly selective cathepsin S two-step activity-based probe in multicolour bio-orthogonal correlative light electron microscopy. *Front. Chem.* Manuscript submitted.

van Leeuwen, T., Araman, M.C., Pieper-Pournara, L., Kampstra, A.S.B., Bakkum, T., Marqvorsen, M.H.S., Nascimento, C.R., Groenewold, G.J.M., Wulp, W., Camps, M.G.M., Janssen, G.M., van Veelen, P.A., van Westen, G.J.P., Janssen, A.P.A., Florea, B.I., Overkleeft, H.S., Ossendorp, F.A., Toes, R.E.M. and van Kasteren, S.I. (2020). Bioorthogonal protein labelling enables the study of antigen processing of citrullinated and carbamylated auto-antigens. *Chem. Sci.* Manuscript submitted.

Bakkum, T., Heemskerk, M.T., Bos, E., Groenewold, G.J.M., Oikonomeas-Koppasis, N., Walburg, K.V., van Veen, S., van der Lienden, M.J.C., van Leeuwen, T., Haks, M.C., Ottenhoff, T.H.M., Koster, A.J. and van Kasteren, S.I. (2020). Bioorthogonal correlative light-electron microscopy of *Mycobacterium tuberculosis* in macrophages reveals the effect of anti-tuberculosis drugs on subcellular bacterial distribution. *ACS Cent. Sci. 6,* 1997-2007

de Geus, M.A.R., Groenewold, G.J.M., Maurits, E., Araman, M.C. and van Kasteren, S.I. Synthetic methodology towards allylic *trans*-cyclooctene-ethers enables modification of carbohydrates: bioorthogonal manipulation of the *lac* repressor. (2020). *Chem. Sci. 11,* 10175-10179

Poolman, J.M., Maity, C., Boekhoven, J., van der Mee, L., le Sage, V.A.A., Groenewold, G.J.M., van Kasteren, S.I., Versluis, F., van Esch, J.H. and Eelkema, R. (2016). A toolbox for controlling the properties and functionalisation of hydrozone-based supramolecular hydrogels. *J. Mater. Chem. B. 4,* p.852-858

Curriculum vitae – Nederlands

Mirjam Groenewold werd geboren op 23 februari 1990 te Harderwijk. In 2009 behaalde zij het gymnasium diploma (met het profiel Natuur & Techniek) aan het Kellebeek College te Breda. Vervolgens begon ze in hetzelfde jaar aan de bachelor opleiding Life Science and Technology met als hoofdrichting Moleculaire Biologie aan de Rijksuniversiteit Groningen. Deze rondde zij in 2012 af met een bachelor project in de onderzoeksgroep Protein Crystallography, waar ze onder leiding van dr. A. Terwisscha van Scheltinga en H. Schäfer werkte aan "An activity assay for chitin synthase".

In 2012 vervolgde zij haar studie aan dezelfde universiteit met de research master Molecular Biology. Hierbinnen voerde ze vervolgens 2 verschillende masterstages uit. De eerste in de onderzoeksgroep Moleculaire Microbiologie onder leiding van prof. dr. A.J.M. Driessen. Dit project, getiteld "Engineering non-ribosomal peptide synthetases in *Penicillium chrysogenum*", werd uitgevoerd onder de dagelijkse begeleiding van H. Ali. Voor haar tweede stage ging zij naar het Max Planck Instituut voor Terrestrische Microbiologie, alwaar zij in de groep van prof. dr. S.V. Albers werkte aan een project getiteld "*In vivo* and *in vitro* analysis of FlaH of *Sulfolobus acidocaldarius*". Dit onderzoek werd begeleid door dr. T. Neiner en dr. M. van Wolferen. Zij vulde haar vrij besteedbare ruimte met een deelname in het iGEM Groningen team van 2013.

Na het afronden van haar master in 2014, begon ze in maart 2015 als PhD student "doctors voor de klas" in de vakgroep Bio-organische Synthese aan de Universiteit Leiden onder leiding van prof. dr. H.S. Overkleeft en dr. S.I. van Kasteren, resulterend in het onderzoek zoals beschreven staat in dit proefschrift. Een eerstegraads bevoegdheid Scheikunde werd behaald in 2018. Delen van het onderzoek zijn gepresenteerd als poster op NWO-CHAINS te Veldhoven (2017 & 2018) en het Reedijk Symposium (2019). Verder heeft zij een poster pitch gegeven op Bioorthogonal & Bioresponsive te Edinburgh (2019) en een mondelinge presentatie tijdens NWO-CHAINS te Veldhoven (2019).

Thans is zij als post-doctoraal onderzoeker werkzaam in de onderzoeksgroep Medische Microbiologie onder leiding van prof. dr. E.J. Snijder en dr. M.J. van Hemert in het Leids Universitair Medisch Centrum.

Curriculum vitae – English

Mirjam Groenewold was born in Harderwijk (NL) on the 23th of February 1990. In 2009, she obtained her gymnasium diploma (majoring in Science and Technology) at the Kellebeek College in Breda. In the same year, she started the bachelor degree 'Life Science and Technology' at the University of Groningen. Here her major was Molecular Biology. She successfully completed a bachelor project in the Protein Crystallography research group under the supervision of dr. A. Terwisscha van Scheltinga and H. Schäfer on "An activity assay for chitin synthase".

In 2012 she continued her studies with a research master in Molecular Biology at the same university. She performed two internships as part of this degree. The first was conducted at the Molecular Microbiology group under supervision of prof. dr. A.J.M. Driessen and titled "Engineering non-ribosomal peptide synthetases in *Penicillium chrysogenum*". For this project she was supervised by dr. H. Ali. For her second internship she moved abroad to the Max Planck Institute for Terrestrial Microbiology in Marburg, where she joined the group of prof. dr. S.V. Albers. Under the daily supervision of dr. T. Neiner and dr. M. van Wolferen, she worked on a project entitled "*In vivo* and *in vitro* analysis of FlaH of *Sulfolobus acidocaldarius*". Her elective space was filled with a participation in the iGEM Groningen team of 2013.

After the successful completion of her master's degree in 2014, she started as a PhD student in March 2015. She was the first crop of students that were part of the "doctors voor de klas"-program, in which the candidates would both obtain a PhD and a first degree teaching qualification. For the former, she joined the Bio-organic Synthesis group at the University of Leiden where under the supervision of prof. dr. H.S. Overkleeft and dr. S.I. van Kasteren, she produced the work described in this thesis. She also obtained her teaching qualification at master-level in Chemistry in 2018.

Parts of the research described in this Thesis were presented as a poster at NWO-CHAINS at Veldhoven (2017 & 2018) and the Reedijk Symposium (2019), as a flash

presentation at Bioorthogonal & Bioresponsive at Edinburgh (2019) and an oral presentation at NWO-CHAINS at Veldhoven (2019).

Currently she is working as a postdoctoral researcher in the department of Medical Microbiology under supervision of prof. dr. E.J. Snijder and dr. M.J. van Hemert at the Leids Universitair Medisch Centrum.

Dankwoord

Zoals iedereen weet is het onmogelijk om te promoveren zonder de steun van heel veel mensen. Deze zou ik dan ook graag willen bedanken voor al hun lieve en bemoedigende woorden tijdens mijn promotieonderzoek.

Misschien was jij wel degene die vertrouwen in mij had als PhD student en die mij daarom deze positie heeft aangeboden. En misschien was jij het wel die altijd oog had voor het welbevinden van een ander.

Misschien waren jullie wel degenen die mij tijdens het promotietraject hebben ondersteund en mij uiteindelijk nog een snelcursus wetenschappelijk schrijven hebben gegeven.

Misschien was jij het wel bij wie ik altijd terecht kon om te praten over moeilijkheden en die die gesprekken altijd afsloot met "kop op en doorgaan".†

Misschien heb ik met jullie ontzettend kunnen lachen op het lab of tijdens de lunchpauzes. Of misschien konden wij hele serieuze kantoorgesprekken afwisselen met momenten van heel veel lol.

Misschien kon ik bij jou altijd terecht voor een goed gesprek en herinnerde jij mij eraan dat het ook belangrijk is om andere dingen te doen dan alleen labwerk.

Misschien heb ik met jullie samengewerkt om bepaalde projecten te kunnen afronden. Enerzijds om chemische verbindingen in biologische experimenten te kunnen testen, anderzijds omdat de biologische experimenten om chemische verbindingen vroegen.

Misschien hebben jullie wel een deel van mijn promotieonderzoek uitgevoerd en heeft elk van jullie mij een inkijkje gegeven in het leerproces van jezelf. Misschien vind je zelfs delen van je werk terug in mijn boekje.

Misschien kon ik bij jullie altijd terecht om te sparren over mijn ideeën over eiwitzuiveringen. En misschien stimuleerde jij mij wel om zo vroeg mogelijk te beginnen met schrijven, omdat je er dan achter komt welke experimenten je nog mist voor een sluitend verhaal.

Misschien zorgden jullie er voor dat wij alle experimenten konden uitvoeren, doordat jullie altijd op tijd de buffers voor ons leverden.

Misschien hadden jullie wel het vertrouwen in mij dat ik een goede docent zou kunnen zijn en hebben jullie mij hier altijd in ondersteund. Nog altijd ben ik welkom om een praatje te komen maken op school en dit waardeer ik enorm.

Misschien waren jullie niet fysiek aanwezig op de universiteit, maar onze virtuele gesprekken hebben mij zeker geholpen om de moeilijkere momenten in het traject in het juiste perspectief te zien. Maybe you are not close by, but I am sure we will keep in touch!

Misschien was jij het met wie ik heb samengewerkt om een SPOC te maken en zoals jij zelf zei: "we vormden een super team". Nog altijd waardeer ik de vriendschap die daaruit is voortgekomen!

Misschien kon ik met jou altijd sparren over wetenschappelijk onderzoek. Meer dan eens zijn hier nieuwe ideeën en inzichten uit voortgekomen en dit bewijst maar weer dat je altijd van elkaars talenten gebruik moet maken.

En tot slot, maar niet het minst belangrijk. Misschien hebben jullie mij altijd vrij gelaten om mijn eigen pad te kiezen en hebben jullie mij geleerd dat het belangrijk is om die dingen te doen in het leven die je leuk vindt.