universität wien

MASTERARBEIT

Titel der Masterarbeit

„Insertion of Modifications into Proteins by Semisynthesis: Generating and Testing of new *NpuDnaE*-Protein Constructs and Expression and Purification of a recombinantly expressed GPI-anchored Peptide for Expressed Protein Ligation“

verfasst von

Alanca Schmid BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2015

Studienkennzahl lt. Studienblatt:	A 066 862
Studienrichtung lt. Studienblatt:	Masterstudium Chemie
Betreut von:	Univ.-Prof. Dr. Christian Friedrich Wilhelm Becker

Table of contents

Abstract

Posttranslational modifications (PTMs) are crucial in biology since they severely influence protein structure, activity and localization, as well as interactions with other cellular components, degradation and signaling. Therefore, while studying proteins, PTMs need to be taken into account. Since the study on isolated modifications is not possible within native, multiply modified proteins, the introduction of site specific PTMs into large proteins is an essential step towards the elucidation of protein function. The prion protein is known to cause fatal neurological disorders such as Creutzfeldt Jakob disease or Bovine Spongiform Encephalopathy by conversion into a pathogenic isoform during posttranslational modification. Membrane anchoring *via* a GPI anchor seems to play a crucial role in neuroinvasion and persistence of the disease. For further studies concerning GPI-anchored protein functions, an easy and fast synthesis strategy is required. Due to the demanding carbohydrate chemistry, a molecular biological approach would be preferable. Therefore, a construct containing a GPI anchoring machinery was expressed in yeast and purified using antigen-antibody interactions. By incorporating a TEV cleavage site, the homogeneous GPI anchor was subsequently released containing an N-terminal cysteine. After purification the obtained membrane anchor can be attached to a C-terminal thioester containing prion protein as well as other thioester carrying proteins *via* Expressed Protein Ligation (EPL). To create recombinantly expressed proteins carrying a C-terminal thioester, inteins are the tool of choice. Fusing the protein of interest (POI) upstream to a mutated intein, N-terminal cleavage can be induced by addition of thiols releasing the POI with the respective thioester on its C-terminus. Since DnaE inteins from different cyanobacteria are known to provide extremely fast splicing kinetics and very good cleavage efficiencies (Shah *et al*., 2012), an artificially fused, mutated version of DnaE intein from *Nostoc punctiforme* was cloned in frame into challenging proteins such as prion protein and the tau protein. Kinetics of thioester generation were observed via SDS-PAGE and compared to the results of previously used MxeGyrA constructs.

Zusammenfassung

Posttranslational Modifikationen (PTMs) sind wichtige Prozesse in der Biologie, da sie sowohl die Proteinstruktur und -aktivität stark beeinflussen wie auch eine Rolle in Abbau- und Signalwegen, Wechselwirkungen mit anderen Zellbestandteilen und der räumlichen Anordnung von Proteinen spielen. Deshalb ist es wichtig, PTMs bei der Untersuchung von Proteinen zu berücksichtigen. Da native Proteine zumeist mehrfach posttranslational modifiziert vorliegen, ist die Beobachtung und Funktionsbestimmung einzelner PTMs schwierig. Dieses Problem kann durch die ortsspezifische Einführung von PTMs in Proteine gelöst werden. Das Prion Protein ist bekannt dafür, durch posttranslationale Umwandlung in eine pathogene Isoform, tödliche, neurologische Erkrankungen wie z.B. Creutzfeld Jakobs Krankheit oder Bovine Spongiforme Enzephalopatie auszulösen. Dabei scheint die Membranverankerung über einen GPI Anker eine wichtige Rolle bei der Neuroinvasion und Persistenz der Krankheit zu spielen. Um die Funktionen GPI verankerter Proteine weiter zu untersuchen, wird eine einfache und schnelle Synthesestrategie benötigt. Da die rein chemische Synthese auf Grund der komplexen Kohlehydratstrukturen sehr aufwendig ist, wäre ein molekularbiologischer Ansatz zu bevorzugen. Daher wurde ein Konstrukt, welches alle für eine GPI Verankerung benötigten Signalsequenzen besitzt, in Hefezellen exprimiert und unter Ausnutzung von Antigen-Antikörper Wechselwirkungen aufgereinigt. Eine eingebaute TEV Protease Erkennungssequenz ermöglicht die Freisetzung des homogenen GPI Ankers mit einem Cystein am N-terminus. Dieser so erhaltene Membrananker kann anschließend durch Expressed Protein Ligation (EPL) an verschiedene C-terminale Thioester enthaltene Proteine angeknüpft werden. Um rekombinant exprimierte Proteine mit einem C-terminalen Thioester zu generieren, sind Inteine das Hilfsmittel der Wahl. Wird das jeweilige Protein vor ein mutiertes Intein kloniert, kann die N-terminale Spaltung durch die Zugabe eine Thiols ausgelöst werden, was zur Freisetzung des Proteins mit einem C-terminalem Thioester führt. Da DnaE Inteine verschiedener Cyanobakterien bekannt sind für ihre schnelle Spleißkinetik und sehr gute Spaltungseffektivität wurde eine künstlich verbundene, mutierte Version des DnaE Inteins von *Nostoc Punctiforme* in frame in anspruchsvolle Proteine wie das Prion Protein oder das Tau Protein kloniert. Die Kinetik der Thioester Bildung wurde mittels SDS-PAGE beobachtet und die Ergebnisse mit den zuvor verwendeten MxeGyrA Konstrukten verglichen.

Abbreviations

AGE	agarose gelelectrophoresis
APS	ammonium persulfate
BSE	bovine spongiforme encephalopathy
bp	base pairs
CBD	chitin binding domain
CJD	Creutzfeldt Jakob´s disease
CMC	critical micelle concentration
CNS	central nervous system
CPS	conditional protein splicing
CV	column volume
Da	Dalton
DNA	deoxyribonucleic acid
DnaE	catalytic α-subunit of DNA polymerase III
DTT	dithiothreitol
ECL	electrochemiluminescence
E.coli	*Escherichia coli*
EPL	expressed protein ligation
ER	endoplasmic reticulum
FT	flow through
GPI	glycosylphosphatidylinositol
GPI-AP	glycosylphosphatidylinositol-anchored protein
GyrA	gyrase subunit A
HILIC	hydrophilic interaction liquid chromatography
HPLC	high performance liquid chromatography
HRP	Horseradish peroxidase
IPTG	isopropyl β-D-1-thiogalactopyranoside
LB	lysogeny broth
LC-MS	liquid chromatography mass spectrometry
LMW marker	low molecular weight marker
M	molar, mol/L
MALDI	matrix-assisted laser desorption/ionization
MESNa	sodium 2-sulfanylethanesulfonate

MSC	multiple cloning site
Mxe	*Mycobacterium* xenopi
NCL	native chemical ligation
NMR	nuclear magnetic resonance
NP-HPLC	normal phase high performance liquid chromatography
Npu	*Nostoc punctiforme*
OD_{600}	optical density at 600 nm
OG	octyl-β-D-glucopyranoside
PAGE	polyacrylamide gelelectrophoresis
PCR	polymerase chain reaction
PI-PLC	phosphoinositide phospholipase C
POI	protein of interest
PrP	prion protein
PTS	protein trans splicing
PTM	posttranslational modification
PVDF	polyvinylidene fluoride
rSAP	recombinant shrimp alkaline phosphatase
rt	room temperature
SC	synthetic complete
SDS	sodiumdodecylsulfate
SPPS	solid phase peptide synthesis
TAE	tris-acetate-EDTA
TBS	tris-buffered saline
TCA	trichloroacetic acid
TCEP	tris(2-carboxyethyl)phosphine
TEMED	tetramethylethylenediamine
TEV	tobacco etch virus
Tris	tris(hydroxymethyl)aminomethane
TSE	transmissible spongiform encephalopathy
Ub	ubiquitin
UDP-GlcNAc	uridine diphosphate- N-acetylglucosamine

1 Introduction

1.1 Prion proteins and the role of GPI anchors

Until the discovery of prion diseases, all infectious diseases were either ascribed to microbial origin like bacteria and fungus or were assumed to be caused by viruses. It took a long time for the scientific community to accept the theory first reported by Prusiner in 1982 (Prusiner, 1982), indicating a changing protein structure as the cause of disease. The term prion disease or transmissible spongiform encephalopathy (TSE) refers to a family of neurodegenerative disorders which are accompanied by a structural change of the prion protein (Basler *et al.*, 1986) and can be found in humans (e.g. Creutzfeldt Jakob disease), as well as in non-human mammals like for example sheep (Scrapie) or cattle (Bovine spongiform encephalopathy, BSE). The phenotypic heterogeneity displayed by different prion diseases is making distinction often difficult. This is aggravated by the fact that TSE can be either inherited, sporadic or acquired by infection, showing different pathological characteristics (Gambetti *et al.*, 2003). In 1986, two different isoforms of the prion protein could be identified and separated: the cellular, soluble PrP^{c} which is released from the membrane after treating it with phospholipase C (PI-PLC) and the disease related, insoluble and protease resistant form PrP^{Sc} (Meyer *et al.*, 1986). Both variants are encoded by the same DNA sequence, suggesting a posttranslational conversion process (Basler *et al.*, 1986). Following the principle of exclusion, all theories about the infectious particle being either a virus-like structure or linked to nucleic acid were successively dismissed and the "protein-only-hypothesis" became the most accepted explanation (Prusiner, 1998). Some arguments supporting the "protein-only-hypothesis" are the familial form of TSE, which is caused by mutations in the PrP gene, the fact that PrP deficient mice could not be infected and that the infectious particles are only destroyed by treatment attacking proteins not viruses or other nucleic acids. Additionally, the cell-free conversion of PrP^{c} into PrP^{Sc}, catalyzed by individual PrP^{Sc} molecules, leads to the conclusion of PrP^{Sc} being the infectious agent (Soto and Castilla, 2004). The cellular prion protein consists of a well-structured, globular C-terminus formed by three α-helices and two small antiparallel β-sheets. At the very end of the C-terminus, a GPI-anchor is attached to the protein. Asn181 and Asn197 can be posttranslationally modified by glycosylation (Figure 1; Aguzzi and Heikenwalder, 2006). The disordered, flexible N-terminal part contains a so called octarepeat region, which seems to be important for copper binding at least *in vitro*

(Hornshaw *et al.*, 1995). A glycine rich region inside the hydrophobic core, playing a critical role in the PrP conversion process is also part of the N-terminal half of the protein (Figure 1B; Harrison *et al.*, 2010). The exact structure of PrP^{Sc} however, remains mainly unclear (Kraus *et al*., 2013).

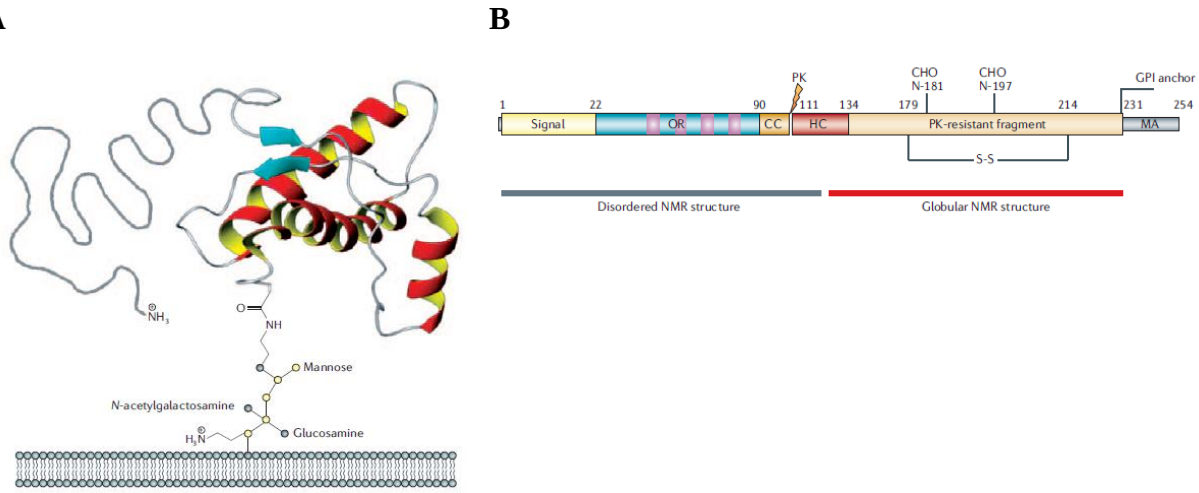


Figure 1: A) Tertiary structure of the cellular prion protein as deduced from NMR spectroscopy. B) An outline of the primary structure of the cellular prion protein including post-translational modifications. A secretory signal peptide resides at the extreme N terminus. The numbers describe the position of the respective amino acids. CC = charged cluster. HC = 'hydrophobic core'. S-S indicates the single disulfide bridge. OR = octarepeat region. The proteinase K (PK) resistant core of PrP^{Sc} is depicted in gold and the approximate cutting site of PK within PrP^{Sc} is indicated by the lightening symbol. MA = membrane anchor region (adapted from Aguzzi and Heikenwalder, 2006).

Despite these insights, there were some unanswered questions left, making the protein-only hypothesis vulnerable. The heterogeneity as well as the fact that it was not possible to generate infectious PrP^{Sc} *in vitro,* that can propagate prion disease *in vivo*, could not be explained by a single infectious protein (Soto and Castilla, 2004). Most of the remaining doubts were cleared up by the results of Wang *et al*. in 2010 demonstrating the *de novo* synthesis of infectious particle from recombinant PrP^{c} (Wang *et al.*, 2010). The different emerged strains could now be explained by the interaction with different cellular co-factors and therefore different conformations of the pathogenic template (Abid *et al.*, 2010). Although not necessary for propagation, the presence of cofactors correlates with *in vivo* infectivity (Wang *et al.*, 2010; Deleault *et al.*, 2007; Kim *et al.*, 2010; Legname *et al.*, 2004; Colby *et al.*, 2010; reviewed by Ma, 2012). Polyanions and lipids play an important role during prion protein conversion, the underlying mechanism, however, still remains unclear (Ma, 2012).

After many years of heated discussions about the identity of the infectious particle and finally the general acceptance of the protein-only hypothesis, another topic became the focus of new discourses: the mechanism of conversion into the pathogenic form of PrP. Two different controversial theories came up: the "template-assisted model" (Figure 2A) and the "seeding/nucleation model" (Figure 2B) (Abid and Soto, 2006). Both are based on the fact that α-helix-rich PrP^{c} turns into PrP^{Sc} by converting helices into β-sheets, as proven by the findings of Pan *et al.* (Pan *et al.*, 1993). Due to the "template-assisted model", interactions between a PrP^{c} intermediate and a stable PrP^{Sc} monomer led to the conversion of normal PrP into the disease related form (Prusiner, 1991). The "seeding/nucleation" hypothesis on the other site excludes a stable PrP^{Sc} monomer which could not be isolated so far.

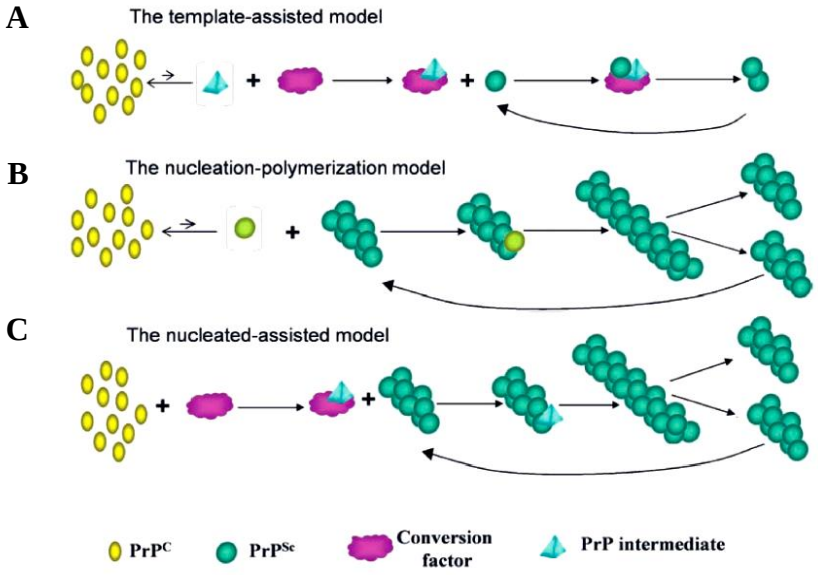


Figure 2: Comparison of three different mechanisms, explaining the conversion of α-helix-rich PrP^{c} into β-sheet-rich PrP^{Sc} (Abid and Soto, 2006).

Supporters of this theory assume a stable equilibrium between PrP^{c} and PrP^{Sc}, where PrP^{Sc} is constantly removed by an initial and rate limiting dimer formation and further aggregation, displacing the equilibrium towards the formation of pathogenic PrP^{Sc} aggregates (Jarrett and Lansbury Jr., 1993). In 2006, Abid and Soto presented a third model, the so called "nucleated-assisted model" (Figure 2C). This theory includes the interaction of PrP^{c} with a conversion factor and avoids the presents of PrP^{Sc} monomers. The PrP^{c} conversion factor intermediate undergoes structural changes, leading to the formation of oligomers and the conversion of α-helices into β-sheets (Abid and Soto, 2006). All three theories point out the fact that PrP^{Sc} is able to convert normal PrP^{c} into a likeness of itself in a self-propagating way by forming amyloid-like fibrils (Prusiner *et al.*, 1983). An assumed subsequent breakdown maybe led to

the formation of smaller and more toxic aggregates which can be used as nuclei for secondary nucleation (Orgel, 1996). Depending on the infectious strain, PrP^{Sc} molecules have a slightly different conformation or even do not match in the primary sequence due to point mutations in distinct species (Safar *et al.,* 1998; Surewicz *et al.,* 2006). The newly formed PrP^{Sc} adopt the same conformation, often leading to variations in incubation time or distinct histological structures of the aggregates (Safar *et al.,* 1998). The concept of conversional protein diseases led to the assumption that maybe other neurodegenerative disorders related to protein misfolding such as Alzheimer´s or Parkinson´s disease also follow the same mechanism of propagation and transmissibility. However, there is only evidence for natural transmission between cells but not between tissues or individuals for misfolded proteins other than PrP so far (Kraus *et al.*, 2013).

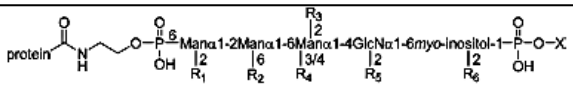
One of the most prominent cases of prion transmissibility even between species was the so called kuru disease, a form of TSE among a cannibalistic tribe in Papua New Guinea, known since the mid-twentieth century. The second important widespread TSE infection led to around 180.000 infected cattle in England during the 1980s and 1990s, followed by more than 200 cases of vCJD due to the consumption of BSE-prion contaminated meat (Weissmann *et al*, 2002; Alpers, 2008; Aguzzi *et al*., 2008).

But how is it possible for PrP^{Sc} to reach the central nervous system (CNS), propagate in different tissues, maintain infectivity and pass over to other individuals? Newly formed PrP^{Sc} has to evade all normal protein folding quality control mechanisms like chaperones, proteolysis or autophagy, in order to avoid degradation (Kraus *et al*, 2013). After ingestion, PrP^{Sc} is transported from the gastrointestinal tract to the CNS (for detailed theories concerning the mechanism view Kraus *et al.,* 2013 or Cobb and Surewicz, 2009).

Transformation already starts in the endoplasmic reticulum (ER) and is completed at the cell surface, where the prion protein is finally embedded into the membrane *via* a GPI anchor (Harris, 2003). Cell-to-cell transmission probably occurs due to intercellular membrane exchange *via* release and uptake of PrP^{Sc} containing exosomes (Fevrier *et al.*, 2004). Nascent PrP seems to be occasionally internalized and thus coming into contact with lysosomal proteases, leading to an N-terminal truncation of PrP^{Sc} (around residue 110, see Figure 1B), and to a complete degradation of PrP^{c} (Caughey *et al.,* 1991; Harris, 2003). On the one hand, this kind of internalization is part of a normal recycling pathway. On the other hand, since endocytosis is triggered by extracellular copper bound to the protein, this process is indicating a physiological role of PrP^{c} in copper metabolism (Pauly and Harris, 1998). However,

comparing binding affinities and physiological copper concentrations, this theory is questionable. Additionally, it could be demonstrated that membrane anchoring is necessary for neuroinvasion and spreading and therefore manifestation of disease (Klingeborn *et al.*, 2011), as well as persistence of the infection (McNally *et al.*, 2009). Mice expressing PrP without a GPI anchor were not able to develop clinical characteristics of the disease, making GPI anchors an interesting research subject in prion disease investigations (Chesebro *et al.*, 2005). Elucidating the role GPI anchors are playing in TSE´s seem to be a promising step towards a fully understanding of prion transmission and infection and will maybe lead to new therapeutic approaches.

The glycosylphosphatidylinositol (GPI) anchor is a ubiquitous posttranslational modification in all forms of eukaryotic organisms. It was discovered indirectly when a new phosphatidylinositol phospholipase (PI-PLC) was described in 1977 (Low and Finean, 1977). The treatment with PI-PLC resulted in a release of different proteins from the cell surface into the extracellular matrix, suggesting all these proteins were inserted into the membrane *via* a phosphatidylinositol structure (Low, 1989).



protein	R_1	R_2	R_3	R_4	R_5	R_6	X
rat brain Thy-1	±Manα1−2	OH	PEtN	±GalNAcβ1−4	OH	OH	alkylacyl-glycerol
human erythrocyte AChE	OH	±PEtN	PEtN	OH	OH	palmitate	alkylacyl-glycerol
hamster brain scrapie prion protein	±Manα1−2	OH	PEtN	(±NANA)- (±Gal)-GalNAcβ1−4	OH	OH	nd
human urine CD59	±Manα1−2	OH	PEtN	±GalNAcβ1−4	OH	palmitate	nd
mouse skeletal muscle NCAM	±Manα1−2	nd	PEtN	±GalNAcβ1−4	OH	OH	nd
bovine liver 5′-nucleotidase	±Manα1−2	±PEtN	PEtN	±HexNAc	OH	OH	nd
human placental APase	OH	±PEtN	PEtN	OH	OH	OH	alkylacyl-glycerol
human CD52	±Manα1−2	±PEtN	PEtN	OH	OH	palmitate	diacyl-glycerol
pig kidney membrane dipeptidase	OH	±PEtN	PEtN	(±Galβ1−3)GalNAcβ1−4 or (±NANA) - GalNAcβ1−4	OH	OH	diacyl-glycerol
human kidney membrane dipeptidase	±Manα1−2	nd	PEtN	(±Galβ1−3)GalNAcβ1−4	OH	OH	nd
T. brucei VSG	OH	OH	OH	±Galα1−2(Galα1−2Galα1−6) Galα1−3	OH	OH	dimyristyl-glycerol
T. cruzi 1G7	±Manα1−2	OH	OH	OH	OH	OH	alkylacyl-glycerol
T. cruzi NETNES	±Manα1−2	OH	OH	OH	PEtN	OH	alkylacyl-glycerol
L. major gp63	OH	OH	OH	OH	OH	OH	alkylacyl-glycerol
S. cerevisiae gp125	±Manα1−2Manα1−2 or Manα1−3Manα1−2	OH	OH	OH	OH	OH	diacyl-glycerol
A. fumigatus PhoAp	±Manα1−3Manα1−2	OH	OH	OH	OH	OH	ceramide
P. communis arabinogalactan proteins	OH	OH	OH	±GalNAcβ1−4	OH	OH	ceramide
D. Discoideum PsA	±Manα1−2	nd	nd	OH	OH	OH	ceramide

[a] Various side chain modifications of carbohydrates, phosphoethanolamine, and/or palmitate (R_1−R_6) are indicated. In some proteins, certain side chains may only be present in a proportion of GPI anchors (indicated by ±). OH indicates that no side chain is known to be present; nd indicates that the side chain or lipid moiety has not been determined. X is the lipid moiety, Man is mannose, Gal is galactose, GalNAc is *N*-acetylgalactosamine, NANA is sialic acid, HexNAc is *N*-acetylhexosamine, and PEtN is phosphoethanolamine.

Table 1: Several structures of known GPI anchors (Paulick and Bertozzi, 2008)

In the following years, the composition and linkage of many variants of GPI anchors from several species were studied and, despite all diversities, a core structure could be identified: The GPI anchor is attached to the C-terminus of the protein *via* phosphoethanolamine, which is linked to a mannose$_3$-glucosamine-phosphatidylinositol chain. This structure can be modified by additional glycans (mannose, galactose, sialic acid, complex

N-acetylgalactosamine containing structures etc.) and phosphoethanolamines, by the acetylation of inositol or different fatty acids (diacetylglycerol or ceramides, saturated or unsaturated and different in length) used for insertion into the lipid bilayer (Nosjean *et al.,* 1997; Paulick and Bertozzi, 2008). There are not only many differences in the modifications of GPI anchors between species but also between different cell types within the same organism, making the generation of homogenous GPI anchors difficult (see Table 1). Primarily, the GPI anchor serves to attach proteins to the exterior surface of cell membranes (Low and Saltiel, 1988). However, the abundant use of this highly complex structure despite the availability of much simpler lipid anchors (palmitoylation, prenylation and other lipidations) indicates further functions (Paulick and Bertozzi, 2008). One advantage of GPI-anchoring at the cell surface is the higher stability in comparison to other lipids, which are only used for intracellular attachment where higher flexibility is beneficial (Low, 1999). Another interesting characteristic of the GPI anchor is the little space it occupies in the membrane as distinct from other protein transmembrane domains. Therefore GPI-anchored proteins (GPI-APs) can be concentrated in small domains on the cell surface (Puig *et al.,* 2014). GPI anchors are also proposed to influence the conformation of the proteins they are linked to. Bütikofer *et al.* could show, that the affinity of some antibodies towards their antigens varies whether they are GPI anchored or not. Because many pathogenic proteins are GPI-linked, this is an important fact which has to be considered while studying GPI-APs connected with various diseases (Bütikofer *et al.,* 2001).

The biosynthesis of the GPI core structure takes place at the cytoplasmic surface of the ER. It starts with the transfer of GlcNAc from UDP-GlcNAc to phosphatidylinositol, followed by a deacetylation of GlcNAc in yeast and mammalian cells (Taylor and Hooper, 2011). In these organisms also *O*-palmitoylation of the inositol is required, although it is often removed later and seems to have quality control functions and is playing a role in protein transport (Fujita *et al.,* 2006). Hereafter, the whole structure flips across the membrane to the luminal side of the ER. Subsequently, three mannoses and one phosphoethanolamine are added, provided by dolicholpyrophosphate-mannose and 3-phosphatidylethanolamine, respectively. After the particular specific modifications are done, attachment to the protein follows. GPI-AP's are synthesized with two signaling sequences. The first N-terminal one is needed to translocate the protein into the ER, the second sequence is a C-terminal hydrophobic peptide which inserts the protein into the ER membrane and also serves as the signal for GPI attachment (Yu *et al*., 2013). The GPI anchor is transferred to the protein after cutting off the C-terminal

signal sequence, catalyzed by a transamidase enzyme complex (Udenfriend and Kodukula, 1995). Afterwards, the so called "fatty acid remodeling" takes place (Masterson *et al.*, 1990). These processes include the replacement of diacylglycerol chains by ceramides in yeast as well as inositol deacylation and the exchange of unsaturated fatty acids with saturated stearic acid at the sn-2 position in mammals (Maeda *et al.*, 2007). The last step of GPI anchor synthesis is then the transportation of the GPI-AP to the outer cell membrane *via* the trans-Golgi network.

To study its role in prion diseases as well as in other pathogenic mechanisms, access to sufficient amounts of homogenous GPI anchors similar to the native structure in mammals is desirable. Since the combination of carbohydrates, lipids, inositol and phosphate groups requires a highly challenging strategy for chemical synthesis (Guo and Bishop, 2004), other methods are desirable. Several approaches were already carried out successfully by using Expressed Protein Ligation (EPL) to attach synthetic GPI anchors or mimics to the C terminus of the protein (Olschewski *et al.*, 2007; Becker *et al.*, 2008). Details of EPL and the required inteins to generate large proteins containing site specific modifications, for example GPI-anchored PrP, will be discussed in the next section.

1.2 Semisynthesis and the use of inteins

Achieving site specific modifications in large proteins was a big challenge for a long time. Although the introduction of Solid Phase Peptide Synthesis (SPPS) in 1963 by Merrifield made site specific modifications and incorporation of unnatural amino acids possible, the length of peptides was generally limited to around 50 amino acids (Merrifield, 1963; Kent, 1988).

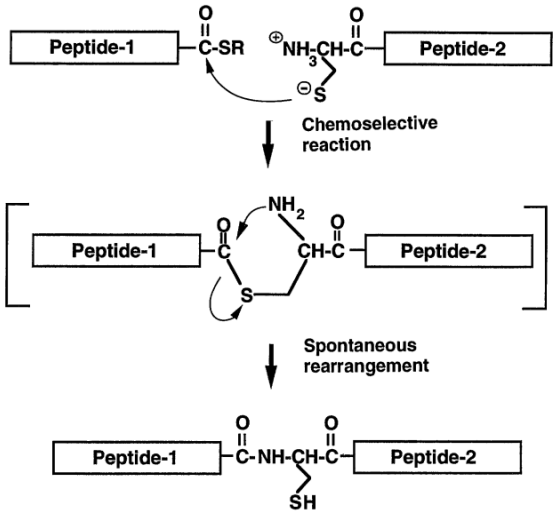


Figure 3: The principal of Native Chemical Ligation. Peptide 1 undergoes a nucleophilic attack by the sidechain of the cysteine residue at the amino terminus of peptide 2. The subsequently formed thioester intermediate then rapidly rearranges to yield a product with a native peptide bond at the ligation site. Both reacting peptide segments are in completely unprotected form, and the target peptide is obtained directly in a native form without the need of further manipulations (Dawson *et al.*, 1994).

Access to larger peptides and small proteins was more easily achieved after introduction of Native Chemical Ligation developed by Dawson *et al*. (Dawson *et al*., 1994). Here, two or more fragments can be fused together through the reaction of a C-terminal thioester moiety of peptide 1 with an N-terminal cysteine of peptide 2 (Figure 3). Since the first reaction step is reversible and only the right intermediate undergoes the last irreversible step, inter- and intramolecular side reactions can be excluded. Therefore, the use of fully unprotected peptides in neutral, aqueous environment became possible (Dawson and Kent, 2000). Furthermore, an amide bond is formed, meaning the obtained peptide is containing a native backbone structure. The addition of specific thiols is improving reaction rate and yield by forming an even more reactive intermediate (Dawson *et al*., 1997). All these facts are making NCL a very interesting method in peptide science. Proteins up to 120 amino acids can be synthesized in satisfying purity and yield (David *et al.,* 2004). Although the synthesis of larger proteins *via* NCL was reported (Becker *et al.,* 2003), the technique has limitations concerning the production of large proteins. Since it is not possible to ligate an endless number of fragments without increasing the probability of purification losses, new methods had to be developed. In 1998, when Muir *et al.* published their paper on Expressed Protein Ligation (EPL), this problem was partially solved (Muir *et al.,* 1998). The semisynthetic approach represents a combination between chemical and biological methods to incorporate modifications into large proteins (Becker, 2011): chemically synthesized, modified peptides were ligated with large, recombinant expressed proteins *via* NCL. The discovery of inteins and the possibility of taking advantage of their natural splicing mechanism to obtain recombinant expressed proteins containing a C-terminal thioester opened up the new field of protein semisynthesis (Shih *et al*., 1988; Xu and Perler, 1996).

The term “Intein” is made up of the words *int*ervening prot*ein* and describes an amino acid sequence which cleaves itself out of the newly synthesized protein by forming a peptide bond between the flanking regions (Kane *et al.,* 1990; Paulus, 2000). This posttranslational process is completely independent of any cofactors or external energy sources, only the intein primary sequence induces the intramolecular bond rearrangement (Paulus, 2000; Xu *et al*., 1993; Kawasaki *et al*., 1997). Although inteins are commonly embedded near to the functional site of regulatory or housekeeping proteins, there is no evidence for biological importance, wherefore they are also sometimes called „selfish genetic elements“ (Gogarten *et al.,* 2002).

The Mechanism of splicing includes four principles steps (Figure 4):

1) Nucleophilic attack followed by an N->S/O acyl shift
2) Transesterification of linear intermediate resulting in the formation of a branched thioester
3) Cyclization of asparagine
4) Rearrangement: S/O->N acyl shift and succinimide hydrolysis

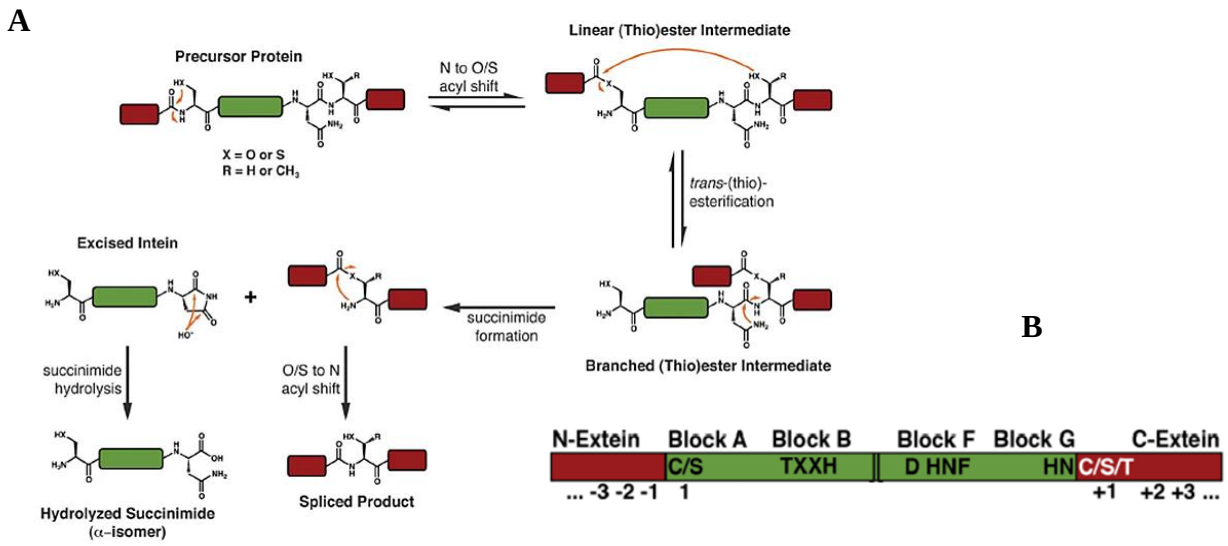


Figure 4: A) Mechanism of protein splicing (Shah and Muir, 2014). B) General scheme of intein structure. Conserved positions are labelled with the corresponding amino acid (Shah and Muir, 2014). Blocks are defined according to Pietrokovski 1994.

Nucleophilic attack followed by an N->S/O acyl shift

Cysteine or serine at position -1 attacks the peptide bond right next to it in direction to the N-terminal extein residue which is already destabilized by interaction with histidine B10. This position is the most conserved one among all inteins, pointing out its importance in the initial splicing step (Figure 5; Mills *et al.*, 2014). Du *et al*. elucidated the mechanism in 2009 by measuring the pKa of histidine B10 before and after the reaction using NMR-based titration experiments. Acid-base catalysis with the involvement of a water molecule could be identified as the underlying reaction (Du *et al.*, 2009).

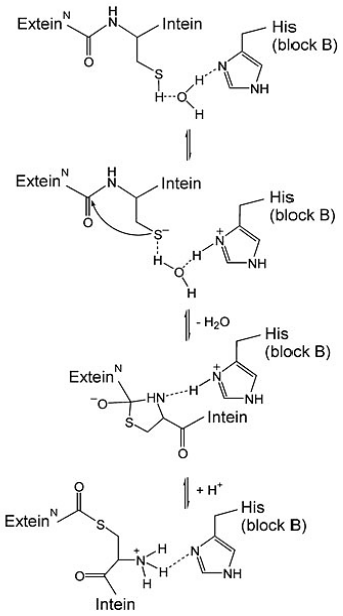


Figure 5: Proposed mechanism for N->S/O acyl shift and role of conserved histidine B10 (Volkmann and Mootz, 2013). Studying the pKa of HisB10 before and after the reaction, it became obvious that and acid-base catalysis takes place by the help of a water molecule (Du *et al.*, 2009).

Transesterification of linear intermediate resulting in the formation of a branched thioester

The transesterification process in protein splicing is a reversible, pH-sensitive reaction and can be initiated by different nucleophiles (Paulus, 2000). Xu *et al.* could elucidate the structure of the resulting branched intermediate and the underlying mechanism of its formation. The first indication of a branched structure was given by the discovery of two free N-termini (Xu *et al.*, 1993). A detailed structure was deduced from the identification of an alkali labile bond between the N-extein and the peptide bond connected intein-C-extein construct (Xu *et al.*, 1994; Figure 4). It is assumed, that the reaction equilibrium is shifted towards the branched intermediate by controlling the protonation state of the nucleophile at position +1, leading to an unusual low pKa and therefore increasing the nucleophilic character of the particular amino acid (Mills *et al.*, 2014; Paulus, 2000; Figure 5).

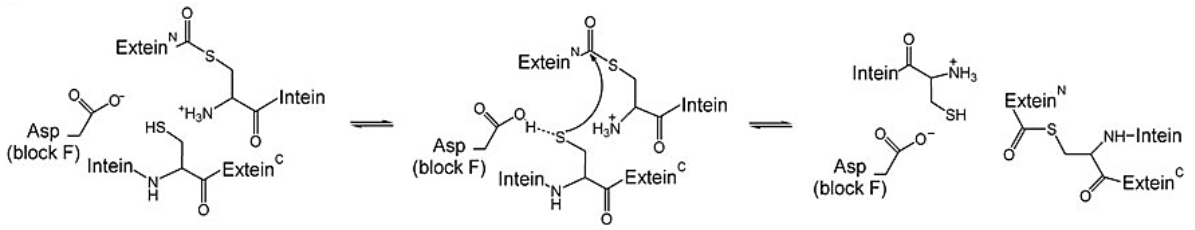


Figure 5: Proposed mechanism for transesterification and involvement of conserved intein residues (Volkmann and Mootz, 2013; modified).

Cyclization of asparagine

The irreversible cleavage is carried out by a nucleophilic attack of asparagine amide (amino acid 7 in block G, see Figure 4B) on the peptide carbonyl group next to it on the extein side. Thereby, a succinimide is formed on the intein C-terminus and the C-extein is released connected to the N-extein *via* a thioester bond (Xu *et al*., 1994; Shao *et al*., 1995; Volkmann and Mootz, 2013). Again, detailed modeling revealed the importance of conserved amino acids within the intein (Mujika *et al*., 2009; Figure 6A). The Muir group proposed a slightly different activation of the asparagine amide (Liu *et al.*, 2014; Figure 6B). However, both agreed on the theory that the branched intermediate serves as a control step since its particular structure seems to directly affect the ability to perform the irreversible asparagine cyclization step. Therefore, the absence of a branched intermediate prevents premature cleavage of linear thioesters (Mujika *et al*., 2009; Frutos *et al*., 2010; Liu *et al*., 2014). It could be also shown that the resolution of the branched intermediate represents the rate limiting step of protein splicing reactions (Frutos *et al*., 2010).

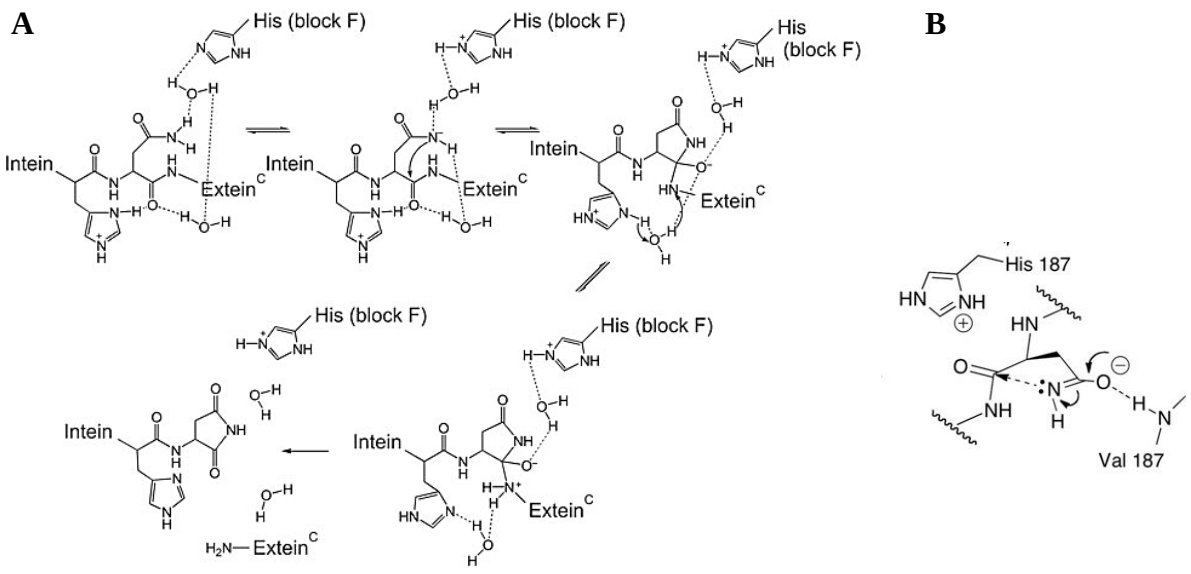


Figure 6: A) Assumed mechanism of asparagine cyclization during protein splicing proposed by Volkmann and Mootz based on the results of Mujika *et al*. (Volkmann and Mootz, 2013; Mujika *et al*., 2009). The carbonyl group as well as the nitrogen of asparagine is activated by the interaction with water molecules and histidine side chains. B) activated asparagine amide proposed by Liu *et al.* (Liu *et al*., 2014)

Rearrangement: S/O->N acyl shift and succinimide hydrolysis

The last steps of intein-mediated protein splicing are the hydrolysis of the C-terminal succinimide on the intein resulting in an Asn or iso-Asn side-chain and the peptide bond formation between the two exteins by an S/O->N acyl shift of the oxo- or thioester. Both

processes are independent of intein catalysis and occur spontaneously due to low activation energies (Shao and Paulus, 1997; Frutos *et al.*, 2010; Xu *et al*., 1994).

Two main side reactions of this canonical mechanism can occur: 1) premature C-terminal asparagine cyclization without previous formation of the branched intermediate, resulting in a separated C-extein, an N-extein-intein construct and the hydrolysis of the asparagine. 2) An external nucleophile attacks the linear or branched intermediate leading to a separated N-extein and an intein-C-extein construct (Shah and Muir, 2014; Figure 7).

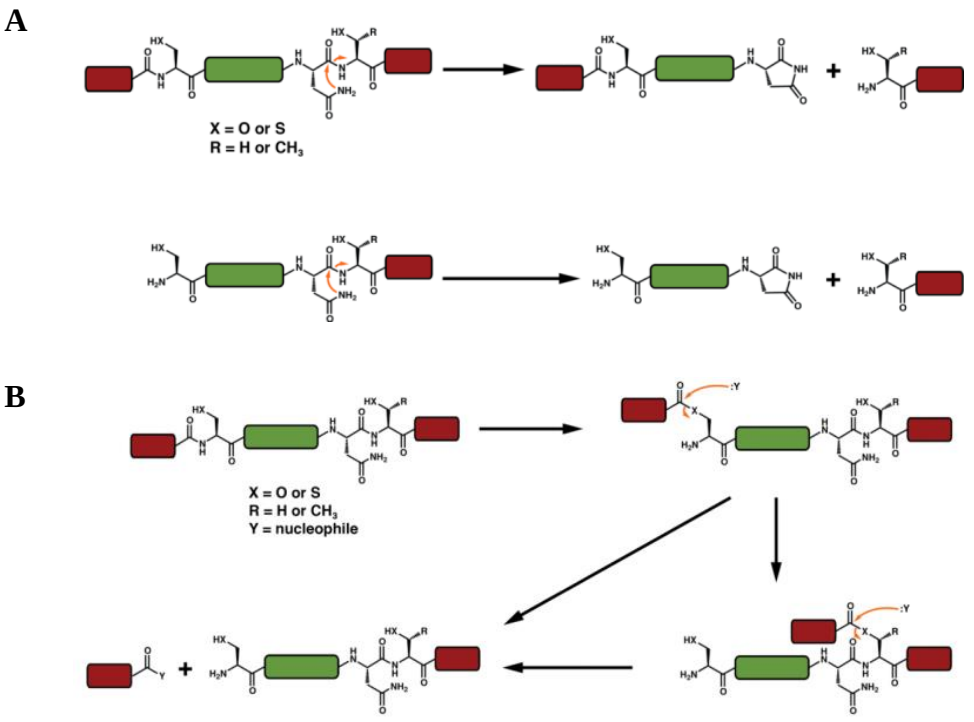


Figure 7: Side reactios of protein splicing A) C-terminal cleavage. B) N-terminal cleavage (Shah and Muir, 2014).

The normal splicing mechanism as well as both side reactions are used in different intein applications. The discovery of inteins influenced many different fields of protein chemistry, maybe the most prominent application today is the tagless purification of proteins. Mutation of block A nucleophile (C/S->A) prevents formation of the intermediate, leading to C-terminal cleavage, which can be induced by either increasing pH or temperature. Mutation of block G asparagine (D->A) instead traps the intermediate and enables N-terminal cleavage after addition of an external nucleophile or again raising the pH (Shah and Muir, 2014; Southworth *et al*., 1999). Fusing the protein of interest upstream or downstream of the intein, respectively, immobilize the construct during the following washing steps *via* an attached affinity tag and induce intein splicing afterwards lead to a purified protein without any

remaining tags. This principle was first made commercial available by NEB (New England Biolabs), bringing the IMPACT^{TM} system to the marked in 1997 (Chong *et al*., 1997). Guan *et al.*, among others, improved the protein purification system by using split inteins with very fast splicing kinetics and varying the affinity tag linkage (Guan *et al.*, 2013).

Another often used technique involving inteins is the already described protein semisynthesis. When fused upstream to the intein, the protein of interest containing a C-terminal thioester is obtained by resolving the branched intermediate through addition of thiols (Chong *et al.*, 1996). The use of an N-terminal cysteine containing peptide as thiol agent leads to a direct ligation to the protein and is called Expressed Protein Ligation (Muir *et al.*, 1998). EPL can be used to modify proteins by incorporating unnatural amino acids, posttranslational modifications or labeling proteins by attaching probes (Shah and Muir, 2014; Vila-Perelló and Muir, 2010). One other important application of EPL is the fusion of building blocks with different isotopic labeling schemes to obtain a full length protein whose NMR-spectra is extremely simplified in comparison to conventionally isotopic-labeled protein variants (Xu *et al*., 1999).

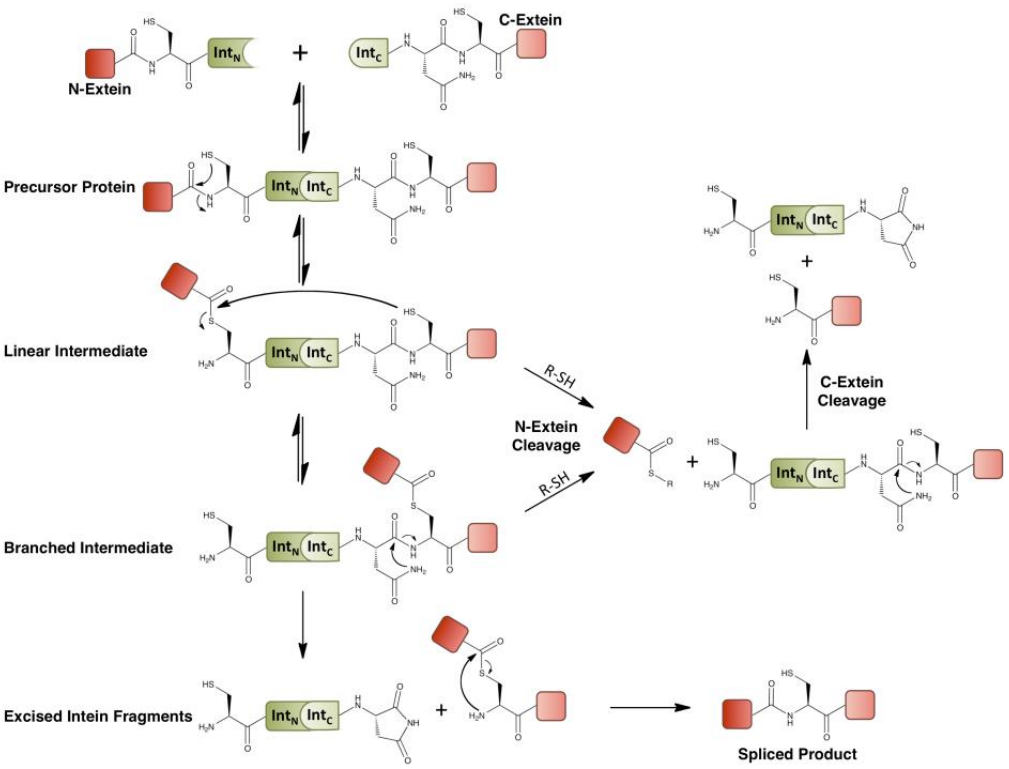


Figure 8: Mechanism of Protein-*Trans*-Splicing and associated side reactions (Shah *et al*., 2012). After Association of the intein fragments, PTS is following the same mechanism as normal protein splicing in *cis*.

A further breakthrough in the history of intein application was the discovery of naturally occurring split inteins and the so called Protein-*Trans*-Splicing (PTS; Figure 8). In contrast to their engineered versions, naturally split inteins do not require any refolding procedures, are mostly soluble and spontaneously associate to form functional inteins. *Trans*-splicing follows the same mechanism as the *cis*-splicing inteins (Aranko *et al.*, 2014; Figure 8). The highly specific ligation of two proteins fused to either the N-terminal fragment (Int^N) or the C-terminal part (Int^C), respectively, opened up the *in vivo* application of EPL in form of PTS. Specific *in vivo* labelling of proteins fused to one split intein fragment after injection of the counterpart fragment fused to any kind of peptide probe became possible (Wood and Camarero, 2014). *In vitro* and *in vivo* backbone cyclization of proteins is an important tool especially in drug discovery to create libraries of cyclic proteins and peptides for screening purposes (Shah and Muir, 2014). Fusing the protein of interest in between the two split intein fragments so that it is representing both exteins at the same time, the two termini became linked together *via* a native peptide bond after protein splicing in *trans*, leaving behind a cyclized protein (Camarero and Muir, 1999).

Streamlined EPL described by Vila-Perelló *et al*. represents a combination of protein purification and semisynthesis in on step: the protein of interest fused to the Int^N of DnaE intein from *Nostoc punctiforme* is bound to a solid support *via* specific interaction to the immobilized, mutated Int^C. After purification, elution is performed by addition of an N-terminal cysteine containing peptide which induces protein splicing by intercepting the reactive thioester intermediate (Vila-Perelló *et al*., 2012; Figure 9).

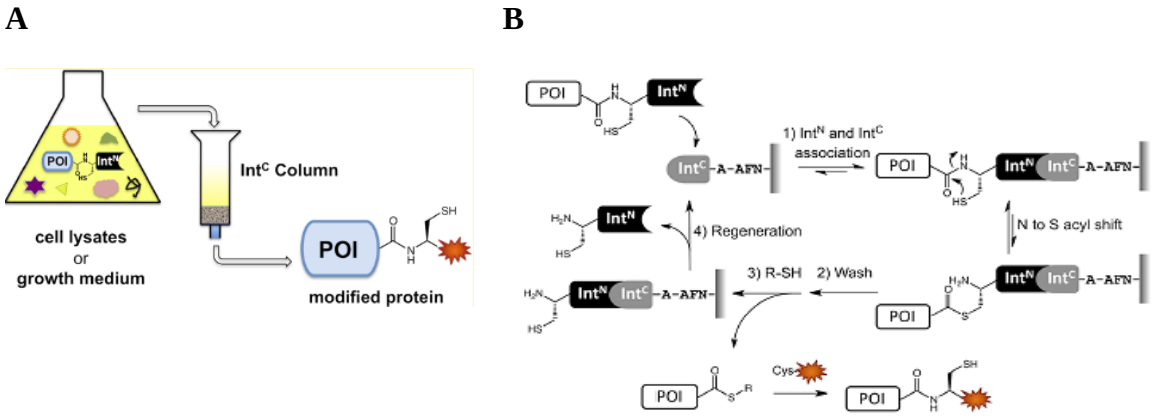


Figure 9: A) Scheme of One-pot streamlined EPL. B) Schematic procedure for the isolation of an α-thioester derivative of a protein of interest (POI) using engineered split intein fragments (Int^N and Int^C) (Vile-Perelló *et al.*, 2012)

Another very important application of split inteins is the specific and very fast activation of protein function, the so called Conditional Protein Splicing (CPS) (Mootz and Muir, 2002; Mootz *et al.*, 2003; Figure 10). Artificially split intein fragments with low affinity towards each other are fused to two domains which can undergo heterodimerization, if a special trigger is present. After dimerization, the intein fragments are in close proximity making the splicing process possible (Figure 10A).

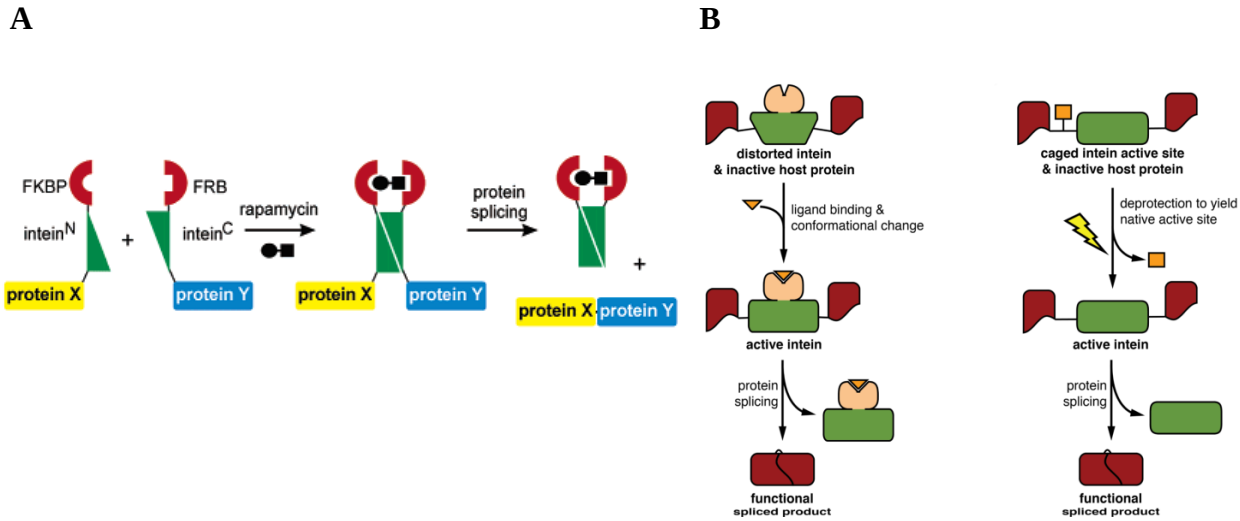


Figure 10: A) The principle of Conditional Protein Splicing (CPS) using small molecules to activate protein functions (Mootz and Muir, 2002). B) Further variations of CPS (Shah and Muir, 2014).

In this case, the FKBP/FRB system was used and rapamycin served as small molecule trigger. Currently there are more variations of this principle developed (Figure 10B). Caging the split or contiguous intein next to the splicing junction with a light- or protease sensitive group enables the induction of splicing at defined time point (Cook *et al.*, 1995; Vila-Perelló *et al.*, 2008; Binschik *et al.*, 2011). A second new principle is the fusion of a small molecule binding domain within a specially designed intein leading to its distortion and therefore inactivation. After addition of the corresponding small molecule, the right conformation is restored and splicing is induced (Skretas and Wood, 2005; Buskirk *et al.*, 2004). CPS is a very useful tool in protein chemistry and could also open new drug delivery and activation strategies in the future (Shah and Muir, 2014).

Split intein are not only interesting because of their unique fields of applications but also because of their extraordinary fast splicing kinetics and high C-extein sequence tolerance. Especially the DnaE intein from the cyanobacterium *Nostoc punctiforme* offers an extremely fast splicing rate constant with a half-live of one minute (Zettler *et al*, 2009) and an efficiency of over 98 % (Iwai *et al.*, 2006) in *trans* splicing applications. To address the question

whether the DnaE intein of *Nostoc punctiforme* is an exception or rather the norm, Shah *et al.* examined splicing kinetics and extein residue tolerance of DnaE inteins from 18 different cyanobacteria (Shah *et al.,* 2012). They could demonstrate that almost all tested inteins exhibit much faster splicing kinetic when previously used inteins such as MxeGyrA and that the C-extein sequence tolerance is high but varies between species. Non-catalytic positions were identified based on which high active inteins and low active inteins can be distinguished, shedding light on how splicing efficiency is controlled. Although not calculating kinetics in detail, they could also show the advantages of artificially fused NpuDnaE fragments in *cis* splicing reactions, generating protein-α-thioesters quicker and with increased efficiency compared to MxeGyrA inteins (Shah *et al.,* 2012).

1.3 Aims of the work

Posttranslational modifications (PTMs) are crucial in biology since they severely influence protein structure, activity and localization, interactions with other cellular components, degradation and signaling. Therefore, while studying proteins, PTMs need to be taken into account. Since the study on isolated modifications is not possible within native, multiply modified proteins, the introduction of site specific PTMs into large proteins is an essential step towards the elucidation of protein function. In the present work, two different projects were addressed, both aiming to create or improve tools for the introduction of PTMs into proteins.

The prion protein is known to be causing fatal neurological disorders such as Creutzfeldt Jakob disease or Bovine Spongiform Encephalopathy by conversion into a pathogenic isoform during posttranslational modification. Membrane anchoring *via* a GPI anchor seems to play a crucial role in neuroinvasion and persistence of the disease (Klingeborn *et al.*, 2011; McNally *et al.,* 2009). Hence, the expression, isolation and purification of a native, homogenous GPI anchor containing an N-terminal cysteine is one of the objectives of this study. The obtained GPI anchor can then be attached to a C-terminal thioester containing prion protein as well as other thioester carrying proteins *via* Expressed Protein Ligation (EPL).

To create recombinantly expressed proteins carrying a C-terminal thioester, inteins are the tool of choice. Fusing the protein of interest (POI) upstream to a mutated intein, N-terminal cleavage can be induced by addition of thiols releasing the POI with the respective thioester on its C-terminus. Since DnaE inteins from different cyanobacteria are known to provide extremely fast splicing kinetics and very good cleavage efficiencies (Shah *et al*., 2012), an

artificially fused, mutated version of DnaE intein from *Nostoc punctiforme* was cloned in frame into challenging proteins such as prion protein and the tau protein. Kinetics of thioester generation were observed via SDS-PAGE and compared to the results of previously used MxeGyrA constructs and product identity was confirmed by LC-MS measurements.

2 Materials and Methods

2.1 Materials

2.1.1 Instruments

Description	Name	Company
Incubator	Infors HT Multitron	Infors HT
Spectrometer	NanoDrop 2000c	Thermo Fisher
Centrifuge	Avanti J-26 XP	Beckman Coulter
Centrifuge	Allegra X-30R	Beckman Coulter
Centrifuge	Centrifuge 5418	Eppendorf
Centrifuge	PerfectSpin Mini	Peqlab
Autoclave	Varioklav	Thermo Fisher
Autoclave	Laborautoklav LA-MCS-203	Sanoclav
Lyophilizer	Christ Alpha 2-4 LD plus	Christ
Cell Disruptor	Cell Disruptor	Constant Systems Ltd
Homogenizer	T18 basic Ultra-Turrax	IKA
Thermomixer	Thermomixer Comfort	Eppendorf
Thermocycler	MJ Mini Thermal Cycler	BioRad
Imaging	ChemiDoc MP	BioRad
HPLC	UltiMate 3000LC System	Thermo Fischer
LC-MS	Waters AutoPurification HPLC/MS System	Waters

2.1.2 Enzymes/Antibodies

Enzyme	Manufacturer
AgeI-HF	NEB
SapI	NEB
NcoI	NEB
AflII	NEB
DpnI	NEB
rSAP	NEB
T4 DNA-ligase	NEB
Q5 DNA-polymerase	NEB
Lyticase	Sigma Aldrich
TEV protease	Kindly provided by Dr. Manuel Brehs (produced according to Tropea *et al.*, 2009)

Antibodies	Species	Manufacturer
Anti-His	Mouse	Sigma Aldrich
Anti-PrP	Mouse	Millipore
Anti-FLAG	Mouse	Sigma Aldrich
Anti-mouse-IgG HRP coupled	Rabbit	Sigma Aldrich
Anti-FLAG M2 affinity resin	Mouse	Sigma Aldrich

2.2 Expression and isolation of a GPI anchor

2.2.1 Recombinant Expression

For the expression of Kre5p-PrP-His-TEV-Flag-Gas1p-GPI, three flasks containing 5 ml SC-medium and 0.1 mg/ml adenine hemisulfate were inoculated with yeast cells transfected

with the required vector. Adenine was added to prevent the enrichment of a red pigment, which slows down growth rate (Smirnov *et al.*, 1967; De la Cruz *et al.*, 1998). Growth took place at 37°C for 7 hours. 500 µl of each culture were transferred into 5 ml new SC-medium + adenine hemisulfate (0.1 mg/ml) and cells continued growing at 37°C overnight. Another over day culture was prepared as described above. Afterwards, each culture was transferred to 100 ml new SC-medium + adenine hemisulphate (0.1 mg/ml) to maintain another overnight culture. 3 x 1 liter main culture was inoculated with these cells to an OD_{600} of 0.6. When OD_{600} reached 1, expression was induced by addition of copper (II) sulfate (0.4 M, 1 ml/l). After three hours at 30°C, cells were harvested by centrifugation (5,000 g, 25 Minutes, 4°C) and stored at -80°C.

SC-medium	6.7	g/l	Bacto-Yeast Nitrogen base without Amino Acids
	2	g/l	Drop out mix (For details see 6)
	20	g/l	Glucose
Adenine hemisulfate	5	g/ml	Adenine hemisulphate in autoclaved water, sterile-filtered

2.2.2 Cell lysis and disruption

After defrosting the cell pellet, lyticase was added to digest cell walls and the mixture was stirred 1 hour at rt in TBS buffer + 0.6 M sorbitol (100 ml buffer/liter main culture). Cell disruption was performed at 2.7 bar using a Constant Systems Ltd cell disruptor, followed by a centrifugation at 75,000 g for 2 hours. To solubilize the membrane protein, the pellet was suspended in buffer A (20 ml buffer/liter main culture) overnight.

TBS	50 150	mM mM	Tris NaCl in water, pH 7.8
Buffer A	6	M	Guanidine hydrochloride in TBS, pH 7.8

2.2.3 Immobilized Nickel Ion Affinity Chromatography

The solubilized membrane proteins were centrifuged (1.5 hours, 75,000 g, 4°C) and the supernatant was purified *via* Immobilized Nickel Ion Affinity Chromatography in one batch. Therefore, the column was calibrated with 2 column volumes (CV = 25 ml) H_2O and 2 CV buffer A. Lysate was incubated with resin for 30 minutes, washed with 4 CV buffer A and protein was eluted after 15 minutes incubation with 2 CV buffer B, respectively. Resin was washed with 2 CV buffer A, 2 CV H_2O, 2 CV ethanol (20 %) and stored in 20% ethanol.

The eluate (200 ml) was concentrated by ultrafiltration using an Amicon 30 kDa MWCO tube to obtain 11 ml concentrated product. The desired protein was identified *via* SDS-PAGE.

Buffer B	300	mM	Imidazole in Buffer A

2.2.4 UV concentration measurements

The concentration of constructs before and after the TEV cleavage was measured using absorbance at 280 nm and calculated based on the Lambert-Beer equation:

$$E = \varepsilon\, c\, d$$

ε_{280} (Kre5p-PrP-His-TEV-FLAG-Gas1p) ≈ 74,830 M^{-1} cm^{-1}

ε_{280} (FLAG-Gas1p) ≈ 1,490 M^{-1} cm^{-1}

Extinction coefficients were obtained using ProtParam from ExPASy. The program is calculating extinction coefficients according to Gill and Hippel (Gill and Hippel, 1989) whose work is built on previous findings from Edelhoch (Edelhoch, 1967) and additional modifications by Pace *et al.* (Pace *et al.*, 1995).

2.2.5 Dialysis and TEV protease cleavage

A stepwise dialysis from 8 M to 2 M urea in TBS buffer before TEV protease cleavage was performed to prevent protein precipitation (8 M -> 6 M -> 4 M -> 2 M). The Protein was mixed 1:10 (w/w) with TEV protease, 5 mM DDT and twice the critical micelle concentration (CMC = 40 mM) of Octyl-β-D-glucopyranoside (OG) to enhance protein solubility. Cleavage took place overnight in 2 M Urea in TBS buffer + 5 mM DTT (Gogolin *et al*., 2010).

2.2.6 SDS-PAGE

To remove the guanidinium which would disturb SDS-PAGE, a TCA precipitation step was performed before loading samples onto the gel. Therefore, 15 µl TCA (40%)/ 50 µl sample solution were incubated on ice for 15 min. Precipitate was centrifuges 15 minutes at 14,000 rpm, supernatant discarded and protein pellet was washed with 200 µl ethanol (20%). After another centrifugation step, pellet was diluted in 10 µl water and 10 µl 2x loading buffer. Samples which did not contain any guanidinium were mixed 1:1 with 2 loading buffer. SDS-PAGE separation was performed using the method developed by Laemmli (Laemmli, 1970). 10 µl of each sample solution were subsequently loaded onto the gel, as well as 7.5 µl of LMW protein marker. Gels were run in Laemmli buffer at 250 V/ 80-100 mA for 30 minutes (gel composition can be found in the supplementary information 6).

Visualization was achieved by either Coomassie-blue staining and a following destaining step or a silver staining procedure.

Coomassie Staining

The polyacrylamide gel was incubated with Coomassie staining solution for about half an hour. To remove background, two washing steps with destaining solution were performed for approximately 10-30 minutes.

Silver Staining

The polyacrylamide gel was incubated with 50 ml incubation solution (10 minutes), washed with water (3 x 5 minutes) followed by 50 ml of nitrate-solution (10 minutes) and another washing step (3 x 20 seconds). Afterwards a developing solution was used (incubation time depended on protein concentration) and the procedure was stopped with 50 ml of stop solution.

Gel composition	Separation gel	1.8	ml	H_2O
		3.75	ml	Acrylamide (30%)
		1.95	ml	Separating buffer (1.5 M Tris, 0.4 % (w/v) SDS, pH 8.8)
		78.5	µl	SDS (10%)
		78.5	µl	APS (10%)
		2.35	µl	TEMED
	Stacking gel	1.8	ml	H_20
		0.45	ml	Acrylamide (30%)

		0.35	ml	Stacking buffer (0.5 M Tris, 0.4 % (w/v) SDS, pH 6.8)
		26.5	µl	SDS (10%)
		26.5	µl	APS (10%)
		2.65	µl	TEMED
Laemmli buffer (10x)		250	mM	Tris
		2	M	Glycine
		1	% (w/v)	SDS
Loading buffer (2x)	(with thiols)	500	mM	Tris
		6	% (w/v)	SDS
		35	% (v/v)	Glycerole
		3.55	% (v/v)	β-mercaptoethanol
		0.05	% (w/v)	Bromophenolblue
				pH 6.8
	(without thiols)	500	mM	Tris
		6	% (w/v)	SDS
		35	% (v/v)	Glycerol
		0.05	% (w/v)	Bromophenolblue
				pH 6.8
Coomassie staining solution		0.1	% (w/v)	Coomassie R250
		10	% (v/v)	Acetic acid
		45	% (v/v)	Methanol
				in water
Coomassie destaining solution		10	% (v/v)	Acetic acid
		40	% (v/v)	Methanol
				in water
Incubation solution (50 ml)		15	ml	Ethanol
		3.4	g	Sodium acetate
		0.1	g	Sodiumthiosulfate
				in water
Nitrate solution (50 ml)		0.05	g	Silver nitrate
		30	µl	Formylaldehyde (36 %)
				in water
Developing solution (50 ml)		1	g	Sodium carbonate
		15	µl	Formylaldehyde (36 %)
				in water

Stop solution (50 ml)	0.04 10	M %	EDTA Glycerol in water

2.2.7 Affinity Chromatography

To separate the Flag-Gas1p-GPI construct from undesired components (Kre5p-PrP-His and TEV protease), an affinity chromatography with agarose-coupled ANTI-FLAG® M2 Affinity Gel was performed. Different conditions were tried to identify optimal conditions for specific binding and elution and therefore obtain the best separation (Table 2).

Since the affinity gel does not tolerate high amounts of chaotropic salts, reducing agents and detergents, the protein construct was dialyzed to a final concentration of 0.5 M urea and thereby most of the detergent and reducing agent was removed as well.

	test 1	**test 2**	**test 3**	**test 4**	**test 5**
Format	batch, gravity flow	batch, gravity flow	batch, gravity flow	batch, centrifugation (4000 g, 1 min, 4°C)	batch, centrifugation (4000 g, 1 min, 4°C)
Sample incubation	1 hour, rt	1 hour, rt	1 hour, rt	overnight, 4°C	1 hour, rt + 20 mM OG + 150 mM NaCl
Elution	6x1 ml 0.1 M glycine, pH 3.5	6x200 µl + 4x1 ml 0.1 M glycine, pH 3.5	4x1 ml + 3x2 ml FLAG-peptide (100 µg/ml) in TBS	2x2 ml 0.1 M glycine, pH 5 2x2 ml 0.1 M glycine, pH 4.5 2x2 ml 0.1 M glycine , pH 4 2x2 ml 0.1 M glycine, pH 3.5	4x1 ml 0.1 M glycine, pH 3.5
Re-generation	washing with TBS until neutral pH	washing with TBS until neutral pH	3x2 ml glycine, pH 3.5, washing with TBS until neutral pH	washing with TBS until neutral pH	washing with TBS until neutral pH

Table 2: Different conditions used during anti-FLAG-affinity chromatography

The affinity gel was equilibrated by washing with several column volumes (1 CV = 1 ml) TBS followed by 3 CV 0.1 M glycine pH 3.5 and another 5 CV TBS until the flow through reached neutral pH again. The sample (0.6-0.9 mg) was incubated with the affinity gel for a certain time (see Table 2) and unbound substance was subsequently removed with 5 CV TBS. After elution of the protein (using different conditions, see Table 2), the gel was recycled and stored in 50 % glycerol in TBS containing 0.02 % sodium azide. All steps were performed at room temperature, if not specified otherwise.

2.2.8 Western Blots

The proteins of interest were transferred from the polyacrylamide gel to a PVDF membrane for 2 hours at 110 mA/ 10-50 V, using semi-dry method in Towbin buffer (Towbin *et al.,* 1979). The membrane was incubated in blocking buffer at 4°C overnight. Thereafter, the specific antibodies were added (anti-His 10 µl, anti-Flag 7.5 µl, anti-PrP 3 µl in 30 ml blocking solution). After 1 hour (anti-Flag and anti-PrP) or 2 hours (anti-His), the membrane was washed with TBS buffer (6x5 minutes, 30 ml, respectively) and incubated with the secondary antibody (1 µl anti-mouse HRP conjugate in 40 ml TBS) for 1 hour. Excess antibody was removed by another washing step with TBS (6x10 minutes, 30 ml respectively). Detection was performed *via* chemi-luminescence after the addition of ECL2-substrate.

Towbin Buffer	250 1.92	mM M	Tris Glycine, pH 8.6
Blocking Buffer	2.5 50	g µl	Skim milk powder Tween in 50 ml TBS buffer, pH 7.8

2.2.9 Dot Blots

The FLAG-Gas1p-GPI construct was too small (≈ 4,807 Da) to be analyzed by common SDS PAGE and subsequent immunodetection *via* Western Blot. Hence, the Dot Blot approach was used to analyze the fractions obtained after affinity chromatography. PVDF membrane was activated with methanol and 5 µl samples were blotted onto the wet membrane. After drying for 30 min, the free parts of the membrane were blocked for unspecific binding by incubating the membrane for one hour (rt) or overnight (4°C) in blocking buffer. The next steps were performed as described in 2.2.8, using specific anti-His-, anti-PrP- and anti-Flag-antibodies as

well as secondary anti-mouse-HRP-coupled antibodies to visualize the Kre5p-PrP-His and Flag-Gas1p-GPI, respectively.

2.3 Molecular Cloning and testing of different ultrafast intein-fusion constructs

2.3.1 Media and bacterial cultivation

2.3.1.1 LB-Agar Plates

LB-medium containing additional 15 mg/L Agar was autoclaved and cooled down to around 55°C. The appropriate amount of antibiotics was added and the mixture distributed into petri dishes. The plates were allowed to cool down and stored at 4°C.

2.3.1.2 Medium

LB-medium supplemented with 100 µg/ml Ampicillin was used.

LB-Medium	10	g/l	Tryptone
	5	g/l	Yeast extract
	10	g/l	NaCl
Ampicillin (1000x)	100	mg/ml	Ampicillin

2.3.1.3 Glycerol stocks

To preserve cells for long term storage, glycerol stocks were prepared. 500 µl of overnight culture were mixed with 500 µl glycerol (40 % v/v) and stored at -80°C.

2.3.2 Transformation of competent cells

Competent cells were slowly thawed on ice. 50-100 ng of dissolved plasmid (or 10 µl ligation mix) were added and the mixture was incubated on ice for 15 minutes. The transformation was performed by heating the cells to 42°C for 90 seconds, then stopping the heat shock immediately afterwards by putting the cells back on ice. After adding 1 ml of LB medium, cells were shaken at 37°C for 1 hour. 200 µl of the solution were put onto a selective LB-agar-plate. The volume of the remaining mixture was reduced by centrifugation and the dissolved pellet was put onto another LB agar plate. Both plates were incubated at 37°C overnight.

2.3.3 Plasmid Extraction

The plasmid extraction was carried out according to the protocol of GeneJET Plasmid Miniprep Kit by Thermo Fisher Scientific using 1 ml of overnight culture.

2.3.4 Polymerase Chain Reaction (PCR)

PCR was performed in two different modes. First, a temperature gradient was used during the annealing step of the primers to identify the optimal annealing temperature (Table 3). Subsequently, a larger amount of template was amplified. To create the appropriate restriction sites, primers depicted in Table 3 were used (synthesized by VBC Biotech Vienna).

Annealing Temperatures:

Template	Primer forward (5' -> 3')	Primer reverse (5' -> 3')	Optimal annealing temperature
NpuDnaE	*GCTA**GCTCTTCC**TGC*T TAAGCTATGAAACGG AAATATTGACAGTAG AATATGG (SapI)	*CTTCC*G**ACCGGT**GAG GCCAGTTTAATG (AgeI)	69.4°C
PrP90-231	TATAA**CCATGG**GCCA AGGAGGG (NcoI)	GATGCA**G*GAAGAGC***G TCTCCCGTCGTAATAG GCCCTGGG (SapI)	63.9°C
Tau1-290	*ATTAA**CC***A**TGG**CTGAG CC*A*CGCCAGGAG (NcoI)	*GTAAAG**CTTAAG***CACT TGGACTGGACGTTGCT AAGATCC (AflII)	69.5°C

Table 3: Annealing temperatures PCR. The nucleotides written in italic letters do not anneal to the template (overhang), the bold characters are showing the recognition site of the respective restriction enzymes.

Program setting:

98°C	4 min
98°C	1 min
Gradient/Temperature	1 min
72°C	30 sec
30 Cycles	-
72°C	5 min
4°C	forever
END	-

Table 4: Programm setting PCR

PCR reaction mixture (200 µl reaction):

		5xQ5 Reaction Buffer	40	µl
10	mM	dNTPs	4	µl
10	µM	Forward Primer	10	µl
10	µM	Reverse Primer	10	µl
≈ 500	ng	Template		
		Q5 High-Fidelity DNA Polymerase	2	µl
		Autoclaved Water	to 200	µl

Table 5: PCR reaction mixture

2.3.5 Agarose Gelelectrophoresis (AGE)

To visualize the results of all PCR reactions, agarose gelelectrophoresis was performed. Considering the small size of the produced inserts, a 1.5 % agarose gel was used. Therefore, 1.5 g agarose was boiled in 100 ml TAE until complete dissolution and filled in gel chambers where it was cooled down to room temperature. Samples were mixed 5:1 with 6 x DNA-Loading Dye containing 2 % ethydiumbromide and were loaded on the gel. Along with the samples, a marker mixture of the same volume was separated on the gel. The separation was achieved in TAE buffer at 120 V/ 80-100 mA for one hour and gels were analyzed by the use of UV light.

Marker mixture	2:1:2		DNA Ladder (100bp/1kb:6x Loading dye:H_2O)
TAE	40	mM	Tris
	20	mM	Acetic acid
	1	mM	EDTA, pH 8.4

2.3.6 Restriction Enzyme Digest

2.3.6.1 Endonuclease II Digest

To produce matching sticky ends, both the vector and the respective inserts were digested with endonuclease type II enzymes. The reaction was performed following the protocol from New England Biolabs for "Optimizing Restriction Endonuclease Reactions" (neb.com) at 37°C for one hour. Since all enzymes used are 100 % active in CutSmart buffer, double digest could be done without any problems (NEB: "Double Digest Protocol with Standard Restriction Enzymes", neb.com). To stop the reaction by inactivating the enzyme, the mixture was heated up to 65°C for 20 minutes.

Restriction Enzyme	10	units (generally, 1 µl is used)
10 x CutSmart buffer	1	µg
DNA	5	µl
H_2O	to 50	µl

2.3.6.2 Digest of methylated DNA by DpnI

To destroy template DNA, a DpnI digest was done after PCR. Due to *E.coli* possessing deoxyribonucleic acid methylase (Dam), generally methylated template DNA is recognized by DpnI, as opposed to the unmethylated PCR product (Geier and Modrich, 1979). Similar conditions as described in 2.3.6.1 were used. However, reaction time was extended to three hours and the heat inactivation was done at 80°C for 20 minutes.

2.3.6.3 5´phosphate removal by rSAP

After restriction digest of the vector, 5´phosphate groups had to be removed to prevent self-ligation. Therefore, one unit of recombinant shrimp-alkaline-phosphatase (rSAP) for each pmol DNA ends (= twice the DNA sample concentration) was added to the restriction mix and incubated for another hour at 37°C. Heat inactivation was performed at 65°C for 10 minutes.

2.3.7 DNA Purification

After PCR and DpnI digest, respectively, DNA samples were purified using the QIAquick® PCR Purification Kit (50) from Qiagen.

2.3.8 Ligation

Ligation was performed at 16°C overnight using the following mixture:

T4 DNA Ligase	1	µl
Vector DNA	0.2	pmol
Insert DNA	0.06	pmol
10x T4 Ligase Buffer	2	µl
H_2O	to 20	µl

10 µl of each ligation mix were used to transformed *E.coli* XL1 cells (see 2.2), which were subsequently grown on agar plates over night at 37°C.

2.3.9 Restriction analysis

To prove the incorporation of the insert into the vector, a restriction analysis was performed. The obtained plasmids were digested with the same restriction enzymes used during the cloning procedure and analyzed afterwards by AGE. In case of successful cloning, the vector as well as the insert should be seen on the gel after staining with ethydiumbromide and excitation by UV light.

2.3.10 Sequencing

Sequencing was done by VBC Biotech Vienna.

2.4 Expression, purification and intein cleavage of NpuDnaE-H_6 fusion constructs

2.4.1 Expression

200 ml of 2YT media were inoculated with *E.coli* Bl21 (DE3) cells containing the vectors pTXB3-PrP90-231-Mxe-H_6-CBD, pTXB3-PrP90-231-NpuDnaE-H_6, pTWIN1-Tau1-290-Mxe-CBD or pTXB3-Tau1-290-NpuDnaE-H_6, respectively. After adding the appropriate amount of Ampicillin (100 µg/ml), cells were grown at 37°C for approximately 16 hours. One overnight culture was used to inoculate one main culture (Table 6). In case of Ub-NpuDnaE, main culture was directly inoculated with appropriate cells. After three hours, protein expression was induced by the addition of IPTG (Isopropyl β-D-1-thiogalactopyranoside).

	Tau	**PrP**	**Ub**
volume pre-culture	2x200 ml	200	-
volume main culture	2x1.5 l	2 l	1 l
amount of IPTG	1 mM	1 mM	0.5 mM
Hours of induction	6	4	3

Table 6: Expression conditions Tau-constructs, PrP-constructs, Ub-construct

Cell growth was followed by measuring the OD_{600}. Furthermore, samples before and after the induction of expression were collected to monitor protein overexpression by SDS-PAGE. To allow a quantitative comparison, each sample must contain a similar number of cells.

Therefore the following formula was used to calculate the required volume of cell culture: $(1/OD_{600})$*200 µl (Table 7).

	OD_{600} main culture before induction
PrP90-231-Mxe-H_6-CBD	0.689
PrP90-231-NpuDnaE-H_6	0.923
Tau1-290-Mxe-CBD	0.833, 1.126
Tau1-290-NpuDnaE-H_6	0.997, 1.178
	OD_{600} after induction (shortly before cell harvesting)
PrP90-231-Mxe-H_6-CBD	1.36
PrP90-231-NpuDnaE-H_6	1.37
Tau1-290-Mxe-CBD	1.667, 1.772
Tau1-290-NpuDnaE-H_6	1.254, 1.599

Table 7: OD_{600} values before and after induction of expression Tau-constructs, PrP-constructs

2YT Medium	16	g	Tryptone
	10	g	Yeast extract
	5	g	NaCl
			H_2O to one liter

2.4.2 Cell harvesting and disruption

After defined time points (Table 5), the cells were harvested by centrifugation (6,000 rpm, 4°C, 30 minutes) and washed with TBS/PBS (4,700 rmp, 4°C, 45 minutes) before the actual cell disruption. Harvested cells were initially homogenized using an IKA T18 basic Ultra-Turrax followed by squeezing in a Constant Systems Ltd fluidizer. Heterologous expression of mammalian proteins in *E.coli* cytoplasm often leads to the formation of insoluble inclusion bodies (Waugh, 2005). No final conclusion could be made so far why these insoluble aggregates are formed in more than half of all recombinantly overexpressed proteins in *E.coli* (Esposito and Chatterjee, 2006; Waugh, 2005). Nevertheless, Esposito and Chatterjee assumed, aggregation occurs due to different translation- and folding-rates in host organism and expressing *E.coli* cells (Esposito and Chatterjee, 2006). Insufficient amounts of chaperones and high concentrations of nascent protein lead to hydrophobic interactions and hereafter to protein aggregation (Wang, 2009). Previous tests showed expression in inclusion bodies for PrP-construct but not for Tau-construct (Olschewski *et al.*, 2007; Görner, 2011).

According to the manufacturer´s recommendation, a pressure of 1.81 kbar was used for proteins expressed in inclusion body instead of the 1.35 kbar used for cells expressing soluble protein.

Prion-protein-constructs

After disruption, a second centrifugation step (20,000 rpm, 4°C, 30 minutes) was performed, the insoluble fraction was again washed with TBS (20,000 rpm, 4°C, 30 minutes) and solubilized for 48 hours in 8 M guanidinium hydrochloride in 50 mM Tris.

Tau-constructs/ Ub-construct

After cell disruption, the mixture was centrifuged (20,000 rpm, 4°C, 20 minutes), the pellet discarded and the supernatant used for further purification.

2.4.3 Purification of Ub-NpuDnaE-H_6, Tau1-290-NpuDnaE-H_6, PrP90-231-Mxe-H_6-CBD and PrP90-231-NpuDnaE-H_6 *via* Nickel-Affinity Chromatography

For PrP-constructs, the same procedure as described in 2.2.3 was used. Ub-NpuDnaE-H_6 and Tau1-290-NpuDnaE-H_6 were washed and eluted in sodium phosphate buffer instead.

Wash buffer	20	mM	Sodium phosphate (dibasic/monobasic)
	500	mM	NaCl
	5	mM	Imidazole pH 7,4
Elution buffer	20	mM	Sodium phosphate (dibasic/monobasic)
	500	mM	NaCl
	500	mM	Imidazole pH 7,4

2.4.4 Intein Cleavage Ub-NpuDnaE-H_6, Tau1-290-NpuDnaE-H_6, PrP90-231-Mxe-H_6-CBD and PrP90-231-NpuDnaE-H_6

Ub-NpuDnaE-H_6

100 µl of the eluate obtained after affinity chromatography purification was mixed with 50 µl thiolysis buffer, 10 mM TCEP and 100 mM MESNa. During the following incubation at 30°C and 450 rpm, samples of 10 µl were collected at various time points, the reaction was

quenched by addition of 30 µl 2x loading buffer without thiols and samples were stored at -20°C until completion of reaction.

Thiolysis buffer	100	mM	Sodium phosphate (dibasic/monobasic)
	150	mM	NaCl
	1	mM	EDTA
	1	mM	TCEP, pH = 7.2

Tau1-290-NpuDnaE-H_6

1 ml affinity chromatography elution fraction was incubated at 30°C with 300 mM MESNa and 1 mM TCEP. 20 µl samples were taken at various time points, quenched with 20 µl 2x loading dye without thiols and stored at -20°C until completion of reaction.

PrP90-231-Mxe-H_6-CBD, PrP90-231-NpuDnaE-H_6

Eluate fractions were concentrated to a final volume of 2.5 ml using Amicon filters 30 kDa or 10kDa MWCO (PrP-Mxe construct or PrP-Npu construct, respectively). Prior to intein cleavage, guanidinium hydrochloride buffer was changed to urea containing buffer using PD10 desalting columns (GE Healthcare Life Science). Columns were washed with 25 ml H_2O and 25 ml 8 M urea in 50 mM Tris, pH 8. 2.5 ml sample were passed through size exclusion resin before elution was performed by addition of 3.5 ml 8 M urea in 50 mM tris, pH 8. Afterwards, samples were dissolved 1:1 in 50 mM Tris to obtain a final urea concentration of 4 mM (Carvajal-Vallejos *et al.*, 2012). 1 ml sample was incubated with 500 mM MESNa at 30°C and 450 rpm. 20 µl samples were taken at various time points, quenched with 20 µl 2x loading dye (Mxe-construct = with thiols, Npu-construct = without thiols) and stored at -20°C until completion of reaction.

2.4.5 Purification of Tau1-290-Mxe-CBD using chitin beads

This construct was purified by taking advantage of the Chitin Binding Domain (CBD). This domain is localized at the C-terminus and binds covalently to chitin beads. After thiol-induced intein cleavage, the protein is released containing a C-terminal thioester, whereas the intein remains on the beads and can be stripped off later by the addition of NaOH solution (300 mM). The resin was equilibrated by washing once with water and three additional times with washing buffer. After incubating the column with cell lysate for 2 hours at 4°C, the resin

was centrifuged and the removed supernatant was stored as flow through. After additional four washing-steps, the cleavage was induced by addition of 200 mM 2-mercaptoethane sulfonate sodium salt (MesNa) in washing buffer. After 48 hours at rt, the supernatant was removed and the resin was washed with three CV column volumes of washing buffer. For regeneration, the resin was incubated three times in NaOH (300 mM) for 30 minutes.

Washing buffer	20	mM	Sodium phosphate (dibasic/monobasic)
	500	mM	NaCl pH 7.7

2.4.6 Final HPLC purification of c-terminal thioester containing proteins

Both, the Tau-thioester and the PrP-thioester, were purified via high performance liquid chromatography using a Varian HPLC system and a preparative C4 column (Table 7).

Tau-thioester	**PrP-thioester**
20-60 % buffer B in buffer A	30-60 % buffer B in buffer A
Flow rate: 1 ml/min	Flow rate: 1 ml/min
1 hour	1 hour
10 minutes desalting step	20 min desalting step

Table 8: HPLC purification conditions thioesters

Fractions were collected every 30 seconds and analyzed via direct injections into an ESI mass spectrometer. Product containing fractions were combined, lyophilized using a Christ Alpha 2-4 LD plus Lyophylizer and stored at -80°C.

Buffer A	0.1 % TFA in water
Buffer B	0.08 % TFA in Acetonitrile

2.4.7 ESI-LC/MS measurements before and after intein cleavage

To identify and compare contents especially before and after intein cleavage, ESI-LC/MS measurements were performed. Therefore, Kromasil C4-analytical HPLC column (Sigma

Aldrich) and gradient of 5-65 % buffer B in buffer A was used to separate sample components prior to mass spectrometric analysis.

2.4.8 Kinetics

To obtain information about the cleavage rates of Mxe-Gyr and NpuDnaE inteins, kinetic calculations were performed by comparing the band intensities of SDS-gels at various time points of splicing reaction. To exclude loading errors, the intensities of splicing product were always expressed as percentage of the corresponding uncut-construct in the same lane. The obtained values and time points were plotted using the curve fitting program Graphpad Prism and the first order rate constant was calculated by the following equation:

Reactant depletion:	$\mathbf{Y = Y_0 - Z\,(e^{-kt}) + Z}$	Y= fraction intensity
Product formation:	$\mathbf{Y = Y_0 + (Z - Y_0)(1 - e^{-kt})}$	Z= Y at time point ∞
		t= time in minutes
Half-life was calculated using the best fit value of k:	$\mathbf{t_{1/2} = ln2/k}$	k= first order rate constant

Equation 1: kinetic calculations

3 Results

3.1 Expression and Isolation of a GPI Anchor

3.1.1 Expression and first purification steps

To obtain a native, logically synthesized GPI anchor, *Saccharomyces cerevisiae* cells were transformed with p425Cup1 vector encoding for PrP22-230 with the N-terminal ER-importing sequence of Kre5p (AA28-39), followed by a C-terminal His-tag, a TEV-protease cleavage site, a FLAG-Tag and the natively GPI-anchored yeast peptide Gas1p511-528 (Figure 11, Schumacher *et al.*, 2010; Amino acid sequence can be found in the supplementary information). Cells were cultured in SC-medium and protein expression was induced by addition of copper (II) sulfate. Since GPI anchors are common eukaryotic posttranslational modifications and are difficult to obtain in prokaryotic *E.coli* systems, the yeast expression system was used (Sahdev *et al.*, 2008).

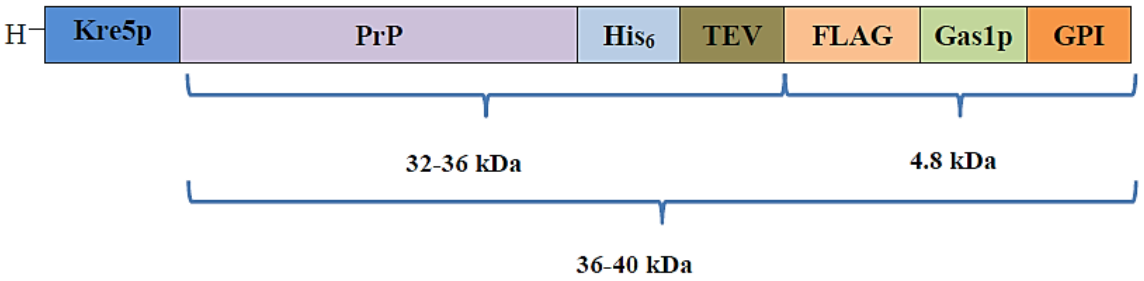


Figure 11: Scheme of expressed protein construct

The signaling sequence of Kre5p, a protein important in the β-1,6-glycan synthesis (Meaden *et al.*, 1990), leads to the translocation of nascent protein into the ER where GPI-synthesis takes place. PrP is serving as a placeholder and is fused to a His-tag, which can be mainly used for purification purposes and in addition for antibody detection. After TEV protease cleavage at the incorporated recognition site, the GPI-anchored Gas1p fragment together with the N-terminally fused FLAG-tag are released containing an N-terminal cysteine which can be used for NCL. The FLAG-tag provides another purification and detection possibility, using antibodies against the highly hydrophilic and therefore surface exposed sequence (Einhauer and Jungbauer, 2001).

After expression, cell harvesting and solubilization of the inclusion bodies, the desired, now soluble protein was purified *via* Immobilized Nickel Ion Affinity Chromatography (see 2.2.3).

The protein containing fractions were precipitated by TCA and the subsequently ran SDS gel (Figure 12) shows that the affinity chromatography was only partially successful since there are still many different protein bands present in the elution fraction and the concentrated eluate.

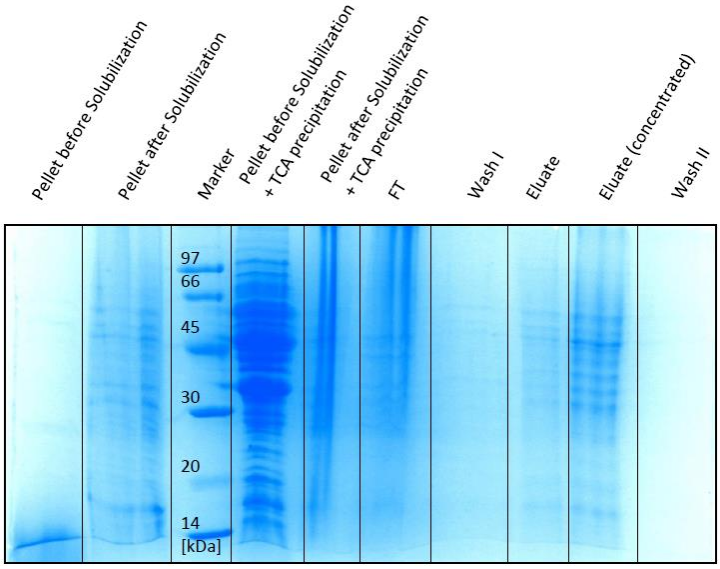


Figure 12: SDS-PAGE affinity chromatography fractions of GPI-anchored construct, Coomassie staining

The remaining lanes are lacking sufficiently resolved protein bands and therefore do not allow conclusive interpretation. To improve visualization of protein bands, silver staining technique as described in 2.2.6 was used (Figure 13). However, the overall protein concentration was too high and due to a lack of time, no additional gels with lower protein concentrations were performed.

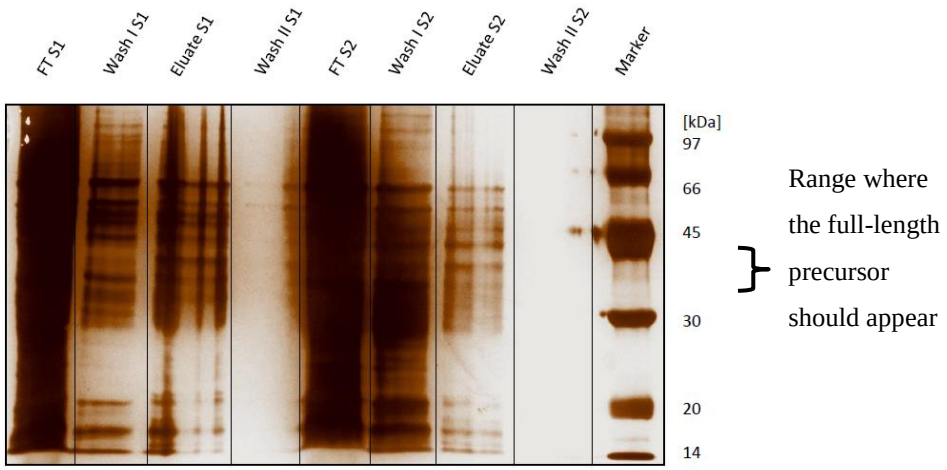


Figure 13: SDS-PAGE affinity chromatography fractions of GPI-anchored construct, silver staining. S1 = sample 1, S2 = sample 2.

To further confirm the presents of the desired construct, Western blots were performed to visualize His-tag, FLAG-tag and PrP after cleavage. The GPI-anchored Gas1p peptide was cut from the reporter prion protein by treating the construct with TEV Protease. To allow protease cleavage the urea concentration was reduced to 4 M (from originally 8 M) by stepwise dialysis (for details see 2.2.5). Before the actual TEV cleavage, the non-ionic detergent OG (*n*-octyl-β-D-glucoside) was added at twice the CMC to keep the GPI-anchored peptide solubilized inside of the micelles. The mixture was dialyzed against 2 M urea and analyzed by Western blots. Since PrP is known to possess a heterogeneous glycosylation pattern and can be either mono- or diglycosylated, the following ranges in calculated molecular weight are caused by different PrP variants (Schumacher *et al.*, 2010). The mass of the full-length construct before the cleavage add up to 36-40 kDa, the His-tag as well as PrP containing fragment should appear at around 32-36 kDa and the desired product has a weight of 4.8 kDa (Figure 11).

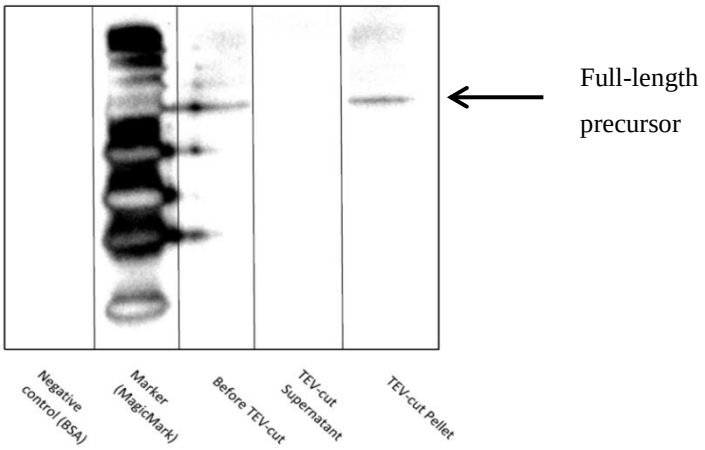


Figure 14: Anti-FLAG Western blot before and after the TEV protease cleavage

The anti-FLAG antibody treated Western blot (Figure 4) provides a first indication of successful cleavage, since no FLAG sequence could be found in the supernatant after TEV-cut. Since the FLAG-containing cleavage product has a molecular weight of only about 4.8 kDa, it was not expected to be visible on the blot. The assumption of successful cleavage was further confirmed by another Western blot, using anti-His antibodies (Figure 15). Here, an obvious mass difference of the His-tag containing constructs before and after the TEV protease treatment can be seen. The mass difference corresponds approximately to the expected mass of 4.8 kDa. Since there was no uncut product left in the supernatant after cleavage, the reaction was completed. Hence, in every immunoreaction based on this fraction,

no full-length PrP-His_6-FLAG-Gas1p-GPI is present and therefore cannot lead to ambiguous signals.

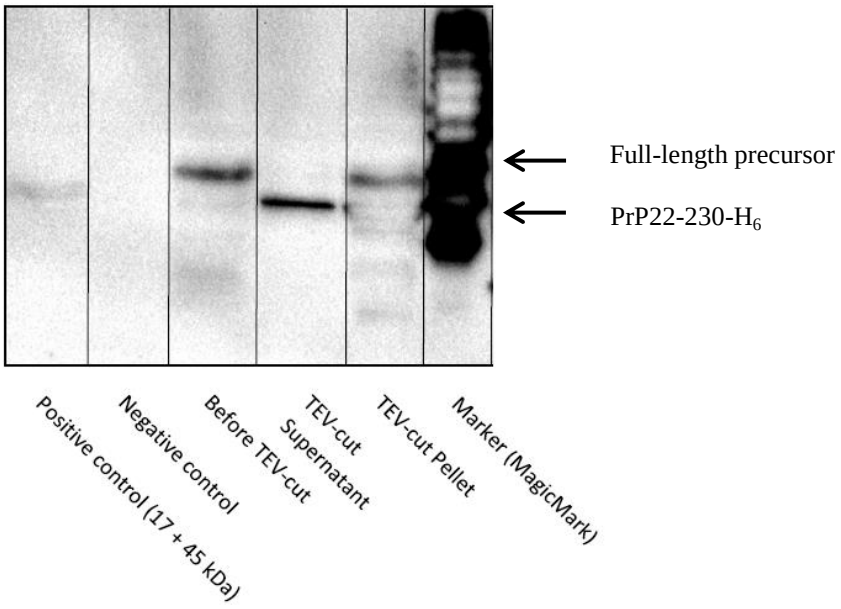


Figure 15: Anti-His-tag Western blot before and after TEV protease cleavage

3.1.2 Affinity chromatography using a FLAG-antibody functionalized agarose matrix

To improve product purity for further reactions, various methods have been tried already. Schumacher *et al.* for example used covalent capture beads (Schumacher *et al.*, 2010). This method is based on the fact, that the GPI-anchored Gas1p peptide is released containing an N-terminal cysteine after TEV protease cleavage. It becomes immobilized by forming a thiazolidine between the cysteine and an aldehyde group on the beads. That way, impurities can be washed out and the desired peptide can be eluted using *O*-methylhydroxylamine (Villain *et al.*, 2001). Since the use of covalent capture beads for purification of the GPI-anchored FLAG-Gas1p was not satisfying and their production is laborious and elaborate (personal communication with Can Araman), finding alternatives for covalent capture beads became one of the subjects of the present work. Methods like purification *via* preparative HPLC seemed to be problematic due to the highly hydrophobic nature of the GPI anchor. Therefore, also detection and verification of product identity is difficult to achieve. Neither MALDI measurements nor direct infusion into an ESI-MS provided any conclusive results. Since the uncut construct contained a FLAG-tag, an easy and highly specific principle of purification should be the interaction between the FLAG-tag and anti-FLAG antibodies.

FLAG antibody factionalized agarose resin was used to isolate the desired FLAG-Gas1p-GPI peptide, while other impurities were washed out and the peptide was eluted either by lowering the pH or by competition with a FLAG peptide. The different fractions were analyzed *via* dot blots using anti-His, anti-FLAG and anti-PrP antibodies. Positive signals in anti-FLAG dot blots indicate the presence of FLAG-Gas1p-GPI (hereinafter referred to as "product"), positive signals in anti-His and anti-PrP dot blots correspond to the presence of PrP-His_6 (hereinafter referred to as "redundant fragment"). As described in 2.2.7 (Table 2) different conditions were tried in order to find the optimal purification terms. Two formats were used, either including washing steps by centrifugation (test 4, 5; Table 2) or by gravity flow (test 1, 2, 3; Table 2). Variations in incubation time and temperature should help to find optimal conditions for antigen binding. The formation of long-range interactions like hydrogen bonds is favored at low temperatures, but at the same time the association rate is decreasing (Reverberi and Reverberi, 2007). Therefore, when lowering the temperature, the incubation time was increased (test 4; Table 2). As already stated above, two different elution methods were tried. Elution at low pH is an easy method to weaken antigen-antibody interaction independently of respective epitope-paratope pairs (test 1, 2, 5; Table 2). To improve the separation during the elution, a gradient was used to slowly decrease pH (test 4; Table 2). One disadvantage of this method is the pH-sensitivity of antibodies, meaning the elution of antigen can lead to irreversible destruction of the immobilized antibody when using slightly too harsh conditions. For competitive elution either a specific high affinity binder has to be used, or the competitor antigen has to be applied in large excess. While this method is highly specific, convenience and costs to create new competitors or to buy purified antigens for every experiment are unsatisfying. In addition the column had to be regenerated by removing the FLAG peptide after every elution, meaning the disadvantage of harsh pH treatment could not be avoided. Because of the mentioned reasons and the unsatisfactory results shown in Figure 18 (test 3; Table 2), this technique was only used in 1 out of 5 test conditions. To improve protein solubility and stabilization, detergent was added to the sample (test 5; Table 2). The selected, non-ionic alkyl sugar detergent OG forms micelles at concentrations higher then 20-25 mM (Chattopadhyay and London, 1984) by exposing the hydrophilic sugar moiety towards the aqueous environment, while hydrophobic alkyl chains are projected into the center. This kind of detergent is able to extract membrane proteins in a mild, non-denaturating way, also allowing to obtaining proteins in their biologically active form (Seddon *et al.*, 2004). Removal of detergent can be achieved by dialysis.

The dot blots of fractions obtained under test 1 conditions are presented in Figure 16. Both, the anti-FLAG and the anti-His antibody treated dot blot showed positive results in the flow through and the first washing fraction and no protein/GPI-anchored peptide in the elution fractions.

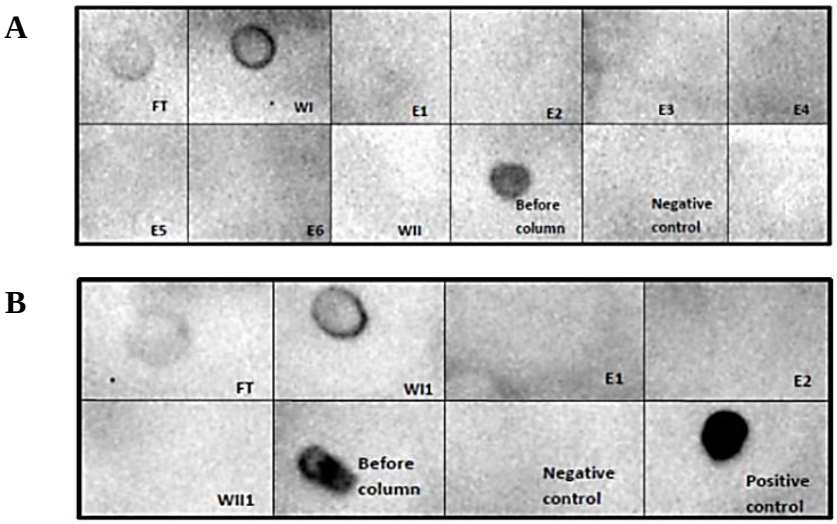


Figure 16: Test 1 dot blots: A) anti-FLAG, B) anti-His. FT = flow through, W = wash, E = elution

Consequently, there was either no GPI-modified peptide at all and the positive signals are generated from uncut precursor, the peptide did not bind to the resin, or the epitope got destroyed by harsh pH conditions during the elution process. Since previous results (Figure 15) render the first explanation quite unlikely, the possibilities of insufficient binding to the resin or degradation are more probable in this context. Destruction of antibodies could be excluded since a quality control with a FLAG-peptide showed good separation results (data not shown).

Under test 2 conditions the flow through fractions showed product signals, indicating a column overload (Figure 17A). This time, however, some product could be identified in the elution fractions (E7-9). Unfortunately, the same fractions still contained the redundant fragment (Figure 17B/C), meaning the separation did not work as intended. Extended wash steps would probably also not help to improve purity of elution fractions, since there is no redundant fragment until fraction E7. These observations indicate unspecific binding of the redundant fragment, which is disrupted under the same conditions as the elution of the product takes place.

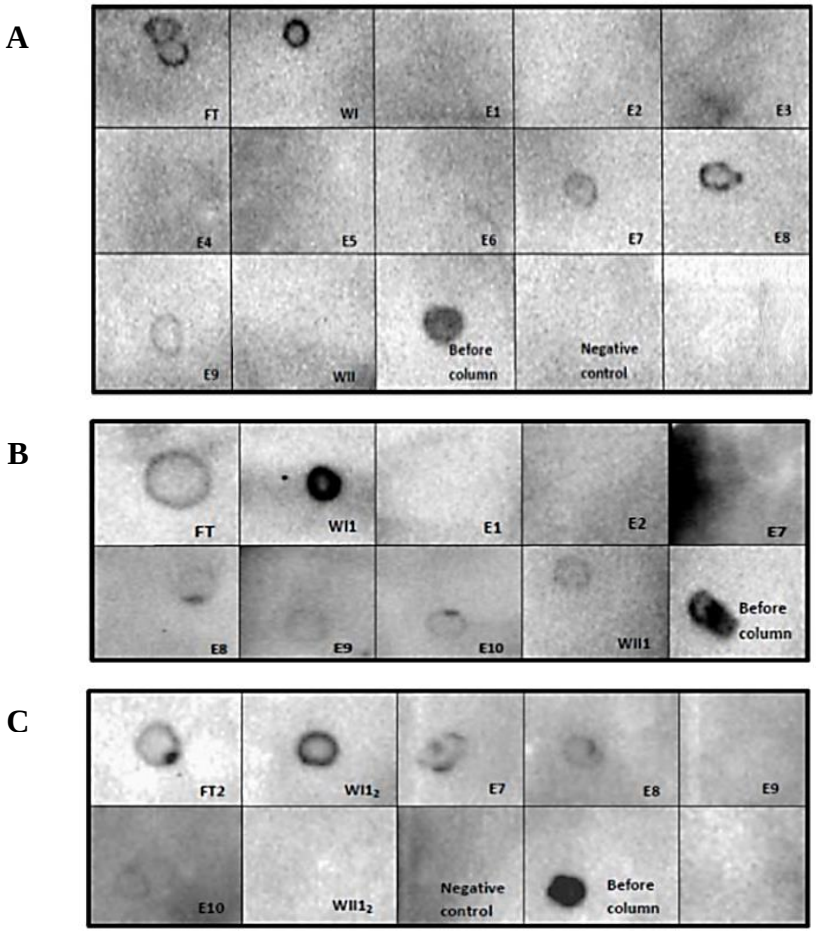


Figure 17: Test 2 dot blots: A) anti FLAG, B) anti-His, C) anti-PrP. FT = flow through, W = wash, E = elution

Fractions after competitive elution using a short FLAG-peptide (DYKDDDDK, Sigma Aldrich) could not be analyzed by anti-FLAG dot blots because of false positive signals caused by the peptide. Nevertheless, anti-His and anti-PrP dot blots revealed redundant fragments in almost every fraction (Figure 18A/B).

Incubation overnight and elution using a pH gradient led to dot blots with quite low signal intensities (Figure 19). However, both product and redundant fragment could be found in the flow through, E4b and WII2, meaning the separation did not work. Considering the long incubation time and the lack of reducing agents, there is a high probability of disulfide bond formation between the two cleaved fragments, making separation inaccessible under the applied conditions impossible.

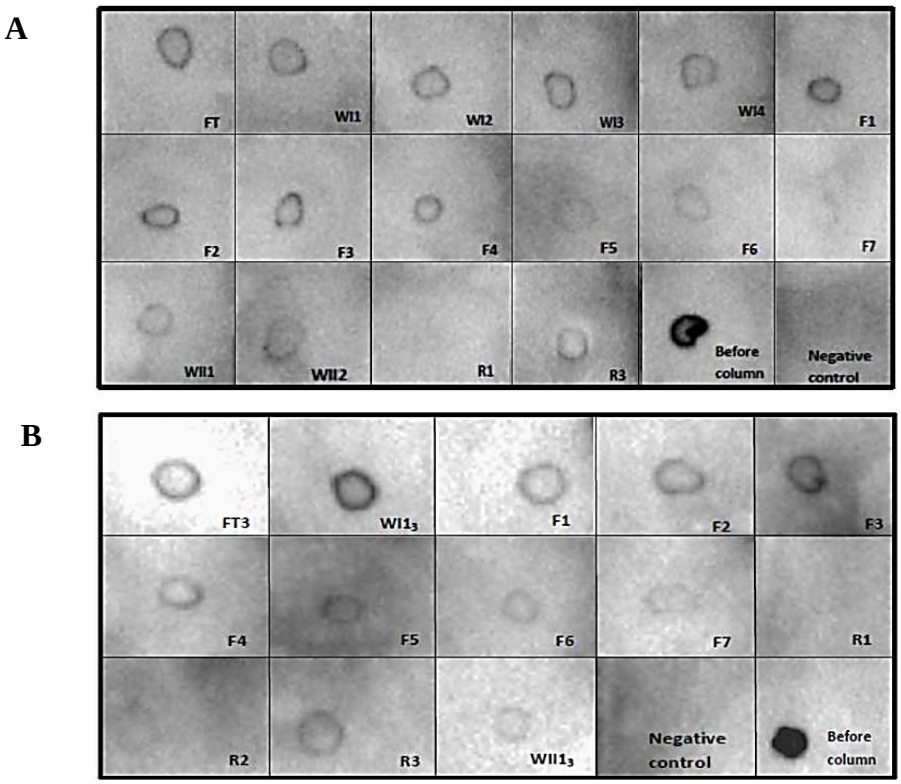


Figure 18: Test 3 dot blots: A) anti-His, B) anti-PrP. FT = flow through, W = wash, F = elution with FLAG-peptide, R = regeneration.

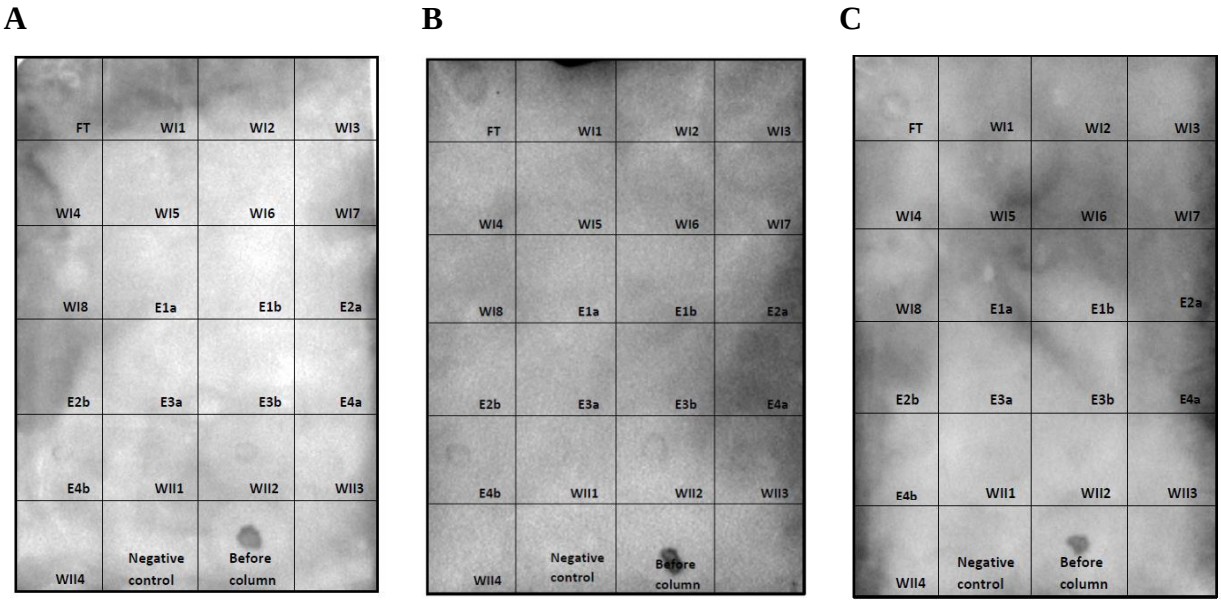


Figure 19: Test 4 dot blots: A) anti-FLAG, B) anti-His, C) anti-PrP

Also the addition of OG did not improve the separation of FLAG-tag containing Gas1p-GPI and PrP-His_6 (test 5, Figure 20), since fractions E2-E4 still contained both cleavage products.

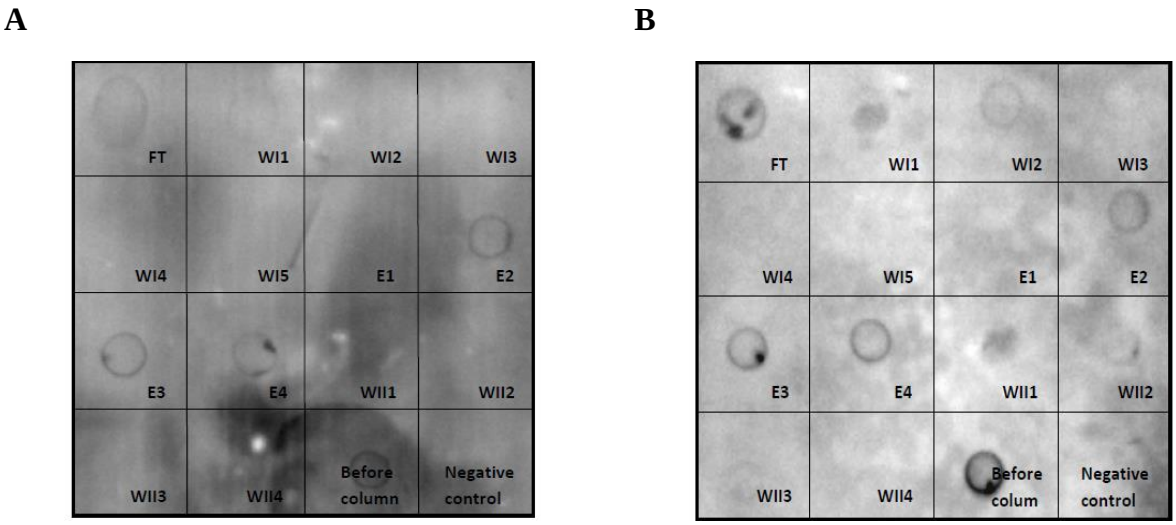


Figure 20: Test 5 dot blots: A) anti-FLAG, B) anti-His

In summary, despite the fact that the antigen-antibody interaction is generally one of the strongest and most specific protein interactions, the antibody-coupled resin is no viable choice in this particular case. The desired product could be easily eluted but only low purity was obtained due to unavoidable co-elution of redundant fragments.

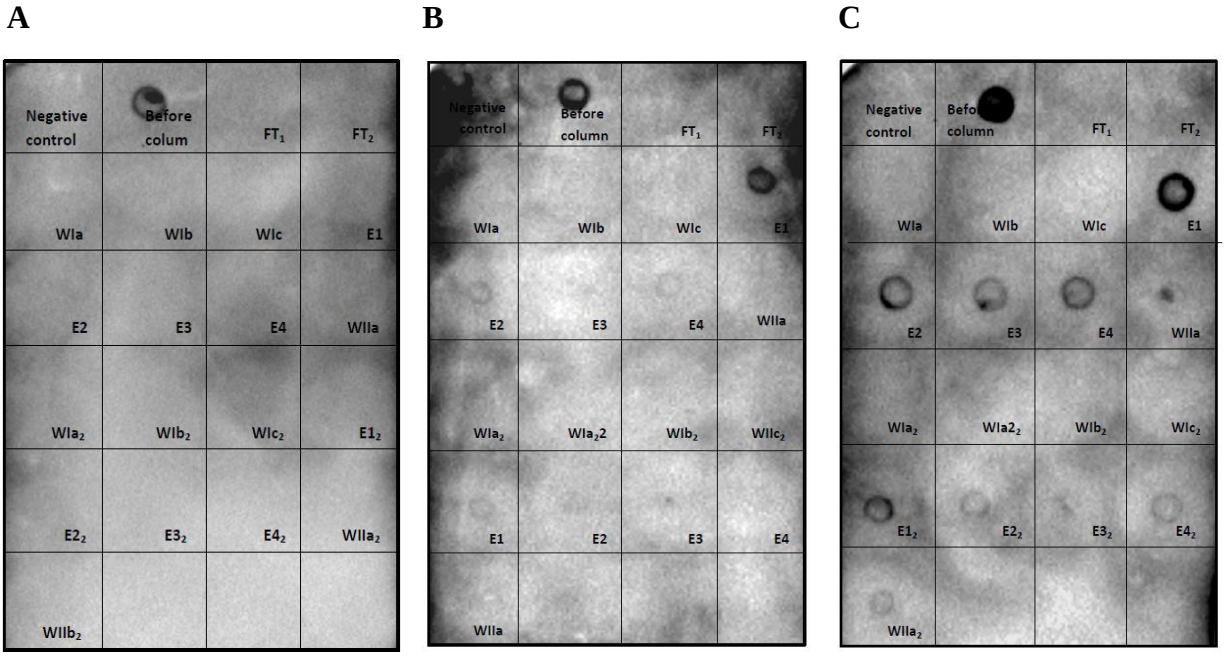


Figure 21: Dot blots after Ni-Affinity Chromatography: A) anti-FLAG, B) anti-His, C) anti PrP

Since the desired product did not contain a His-tag anymore, the possibility of separation using a second Ni-Affinity Chromatography came up. In this case, the product should be found in the flow through, while the His-tag containing redundant fragment should bind to the column. 1 % OG was added to the TEV protease cleavage product and chromatography was

performed according to 2.2.3. Dot blots showed no product at all (Figure 21A), while the redundant fragment could be found in the elution fraction as expected (Figure 21B/C). Again there are two possible explanations for the inability of detecting the product. One is the degradation of FLAG-epitope needed for antibody recognition, the other possibility could be an unspecific, hydrophobic binding of GPI-anchored peptide to the affinity resin. After several washing steps with OG containing urea buffer, 70 % isopropanol and ethanol, there was still no product in the wash fractions (data not shown). Hence epitope degradation is the most likely explanation.

In conclusion, neither the antibody-antigen interaction nor the Ni-Affinity Chromatography was able to efficiently separate the GPI product from undesired His-tagged PrP. Thus, no significant improvement of purification after the cleavage could be achieved, compared to the previously used covalent capture beads. Moreover, antibody-coupled resin is more expensive although not as laborious as producing covalent capture beads, since the resin is commercially available.

3.2 Cloning, expression and comparison of different intein-protein constructs

Due to their expected benefits described in 1.2, the fast DnaE split intein from *Nostoc punctiforme* should be able to enhance the thioesther generation rates when fused to protein of interest compared to previously used intein construct such as MxeGyrA constructs. This has been already demonstrated for Ub-NpuDnaE by Shah *et al.* (Shah *et al.*, 2012). However, their advantage for semisynthesis of more challenging proteins like the naturally unfolded tau protein or prion protein variants still needs to be proven. To address this question, the strategy showed in Figure 22 was used to create new vectors, containing either the DNA-sequence encoding for Tau1-290 or PrP90-231 fused to NpuDnaE inteins and a downstream His-tag for purification reasons.

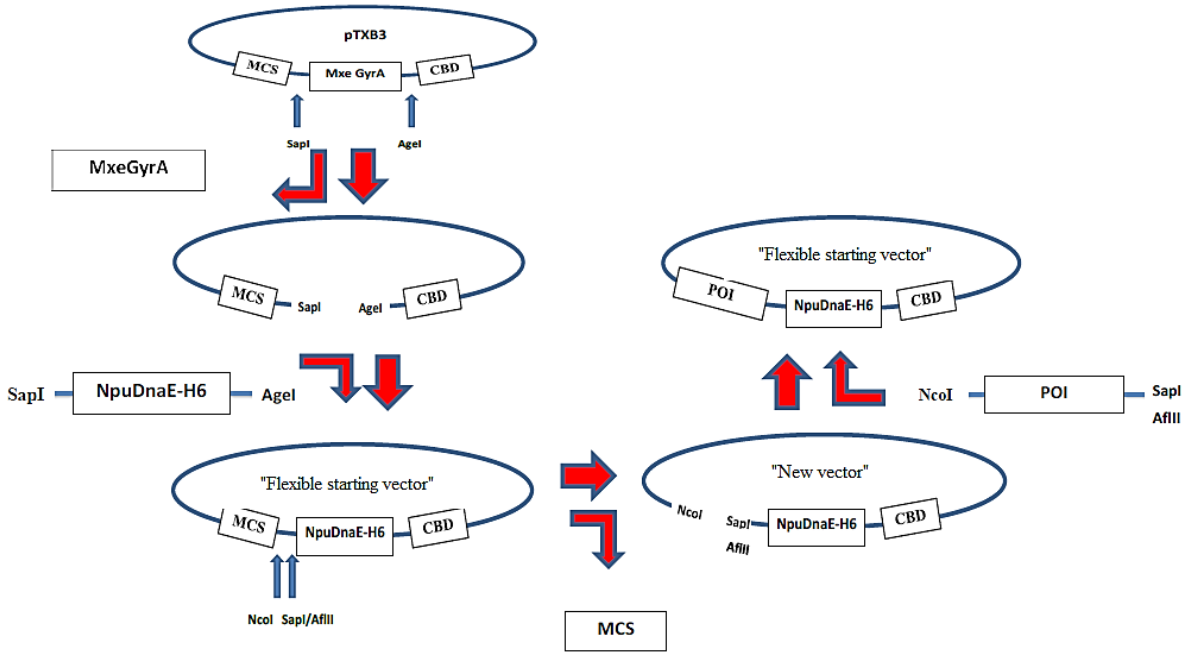


Figure 22: Overall cloning strategy pTXB3-PrP90-231-NpuDnaE-H_6, pTXB3-Tau1-290-NpuDnaE-H_6. MSC = multiple cloning site, POI = protein of interest, CBD = chitin binding domain. SapI, AgeI, NcoI and AflII = endonuclease II enzymes.

3.2.1 Cloning of pTXB3-NpuDnaE-H_6

Since the vector pTXB1-Ub-NpuDnaE-H_6 kindly provided by the Muir lab at Princeton University did not contain an appropriate restriction site to easily change the POI (Figure 23A), the intein had to be transferred into a new vector prior to use a flexible cloning strategy. Therefore, pTXB3 vector was chosen and primers were designed to create new restriction sites on the upstream site of intein (Figure 23B). After amplification of the NpuDnaE insert *via* PCR, restriction digest of the insert as well as the pTXB3 vector with the enzymes AgeI-HF and SapI was performed. The newly generated matching sticky ends were subsequently ligated using T4 ligase. The obtained vector was used to transform *E.coli* XL1 cells, which were subsequently grown on agar plates. Random colonies were picked and a test restriction was carried out.

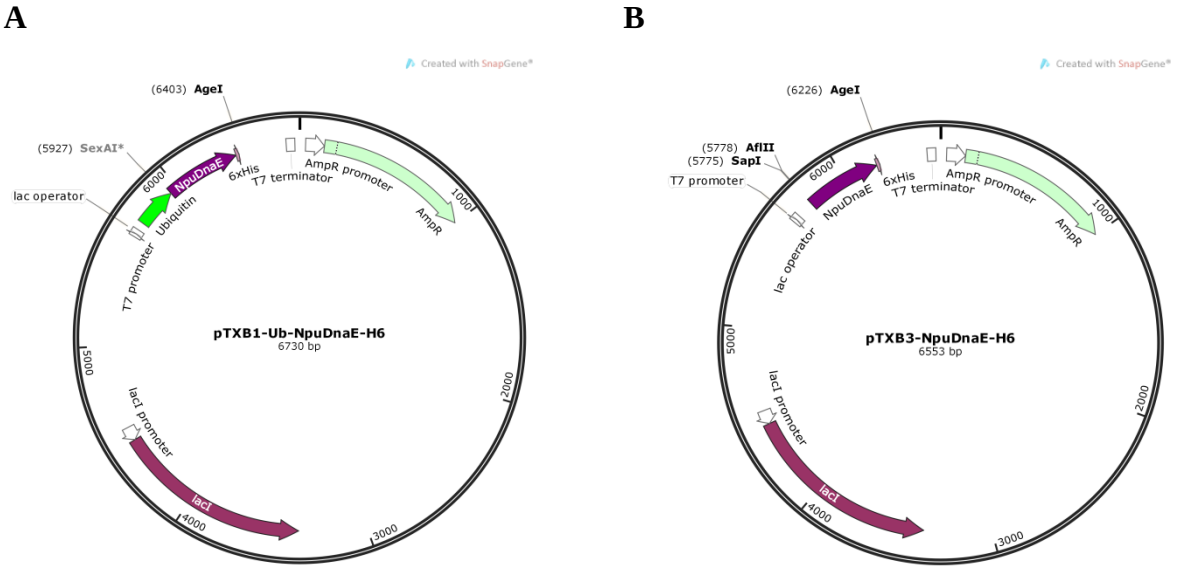


Figure 23: A) Vector map of pTXB1-Ub-NpuDnaE-H_6 (kind donation of Muir lab, Princeton University). B) Vector map of pTXB3-NpuDnaE-H_6 "flexible starting vector"

Agarose gel electrophoresis showed successful cloning and transformation of colony IIa, IIb an IIc, since the obtained small fragment appeared at approximately the size of expected NpuDnaE insert (≈ 450 bp) (Figure 24). The ultimate confirmation of sequence integrity was supplied by DNA-sequencing for clones IIb and IIc.

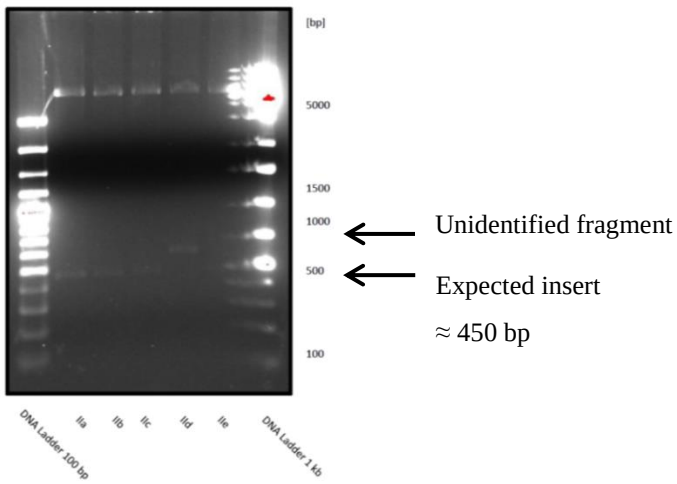


Figure 24: Agarose gel test restriction pTXB3-NpuDnaE-H6 IIa, IIb, IIc, IId, IIe, = random colonies from plate No II.

3.2.2 Cloning of pTXB3-PrP90-231-NpuDnaE-H_6

To create the vector depicted in Figure 25A, PrP90-231 was amplified using a pTXB3-PrP90-231-MxeGyrA-H_6-CBD vector as template and primers containing restriction sites for NcoI

and SapI. After restriction digest of the flexible starting vector and the PrP-based insert, ligation was performed and *E.coli* XL1 cells were transformed.

A **B**

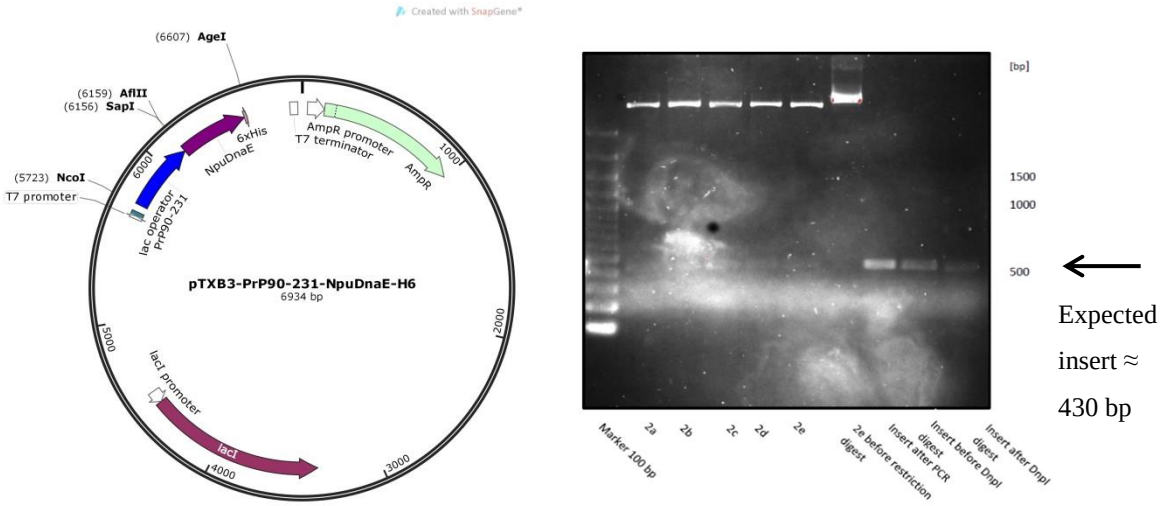


Figure 25: A) Vector map of pTXB3-PrP90-231-NpuDnaE-H_6. B) Agarose gel test restriction pTXB3-NpuDnaE-PrP90-231-H_6. 2a, 2b, 2c, 2d, 2e, = random colonies from plate No II. PrP90-231 ≈ 430 bp.

Again, random colonies were picked, plasmids were extracted and results presented in Figure 25B were obtained from subsequent test restriction and agarose gel electrophoresis. Vector samples from all five colonies were sequenced and 2a, 2b and 2e contained the desired DNA-sequence.

3.2.3 Cloning of pTXB3-Tau1-290-NpuDnaE-H_6

The same procedure as described in 3.2.2 was used to create the vector shown in Figure 26A. Vector pTWIN1-Tau1-290-MxeGyrA-CBD served as a PCR template. Due to difficulties with compatibility of sticky ends generated from different SapI sites, AflII was used instead. Therefore, primers were designed containing AgeI and AflII restriction sites, respectively. Test restriction of extracted plasmids showed positive results for all tested clones (Figure 26B). Sequencing verified the results for Ia, Ic, Ie, IId and IIe.

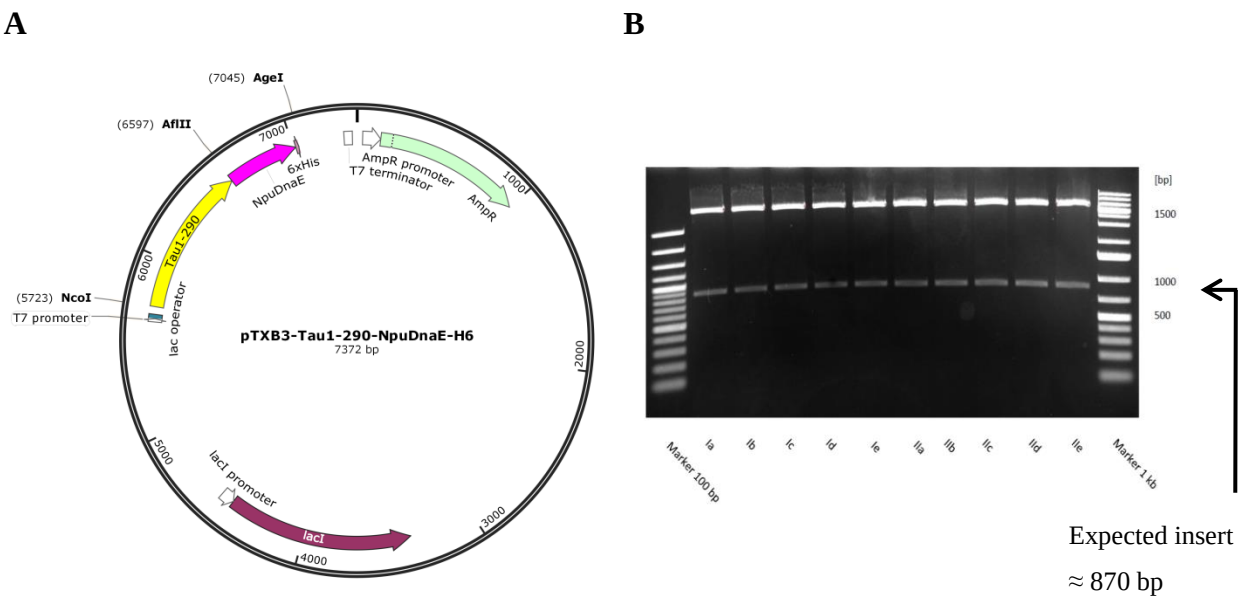


Figure 26: A) Vector map of pTXB3-Tau1-290-NpuDnaE-H_6. B) Agarose gel test restriction pTXB3-Tau1-290-NpuDnaE-H_6. Ia, Ib, Ic, Id, Ie, IIa, IIb, IIc, IId, IIe = random colonies from plate No I/II. Tau1-290 ≈ 870 bp.

3.2.4 Expression, purification and intein cleavage of Ubiquitin-NpuDnaE-H6

To confirm the reproducibility of NpuDnaE intein splicing reported by Shah *et al.*, *E.coli* Bl21(DE3) cells were transformed using the donated pTXB1-Ub-NpuDnaE-H_6 vector and the protein construct was expressed and purified following the published protocol (Shah *et al.*, 2012).

Based on the increasing intensities observed at various time points after induction of expression (Figure 27), the desired fusion protein was significantly overexpressed. However, at this stage the first challenge of very fast splicing inteins becomes obvious: the splicing process already started during expression and especially while being stored at room temperature before purification. The band at around 28-30 kDa most likely belongs to the full-length product (calculated mass: 25.5 kDa), while the band at 22 kDa corresponds to the cleaved NpuDnaE-H_6 intein (calculated mass: 16.9 kDa). Prematurely cleaved ubiquitin appears at around 17 kDa (calculated mass: 8.6 kDa) (Figure 27). All three components show a higher apparent than calculated mass in the SDS-gel, what is not in agreement with the results presented by Shah *et al.* 2012 (Figure 30A).

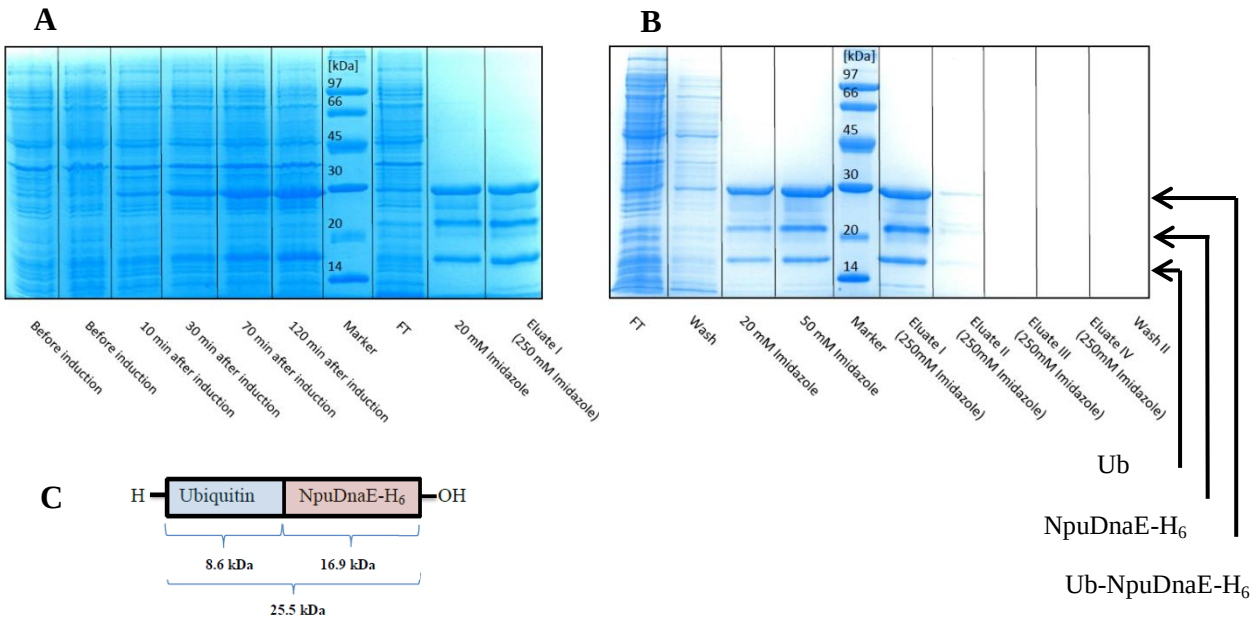


Figure 27: A) SDS gel of samples collected before and during expression of Ub-NpuDnaE-H_6. B) SDS gels of Ni-affinity chromatography fractions of Ub-NpuDnaE-H_6. C) Scheme of expressed contruct.

For more precise identification of all proteins, LC/MS measurements were carried out. Under non-reducing conditions, the following spectra were obtained (Figure 18).

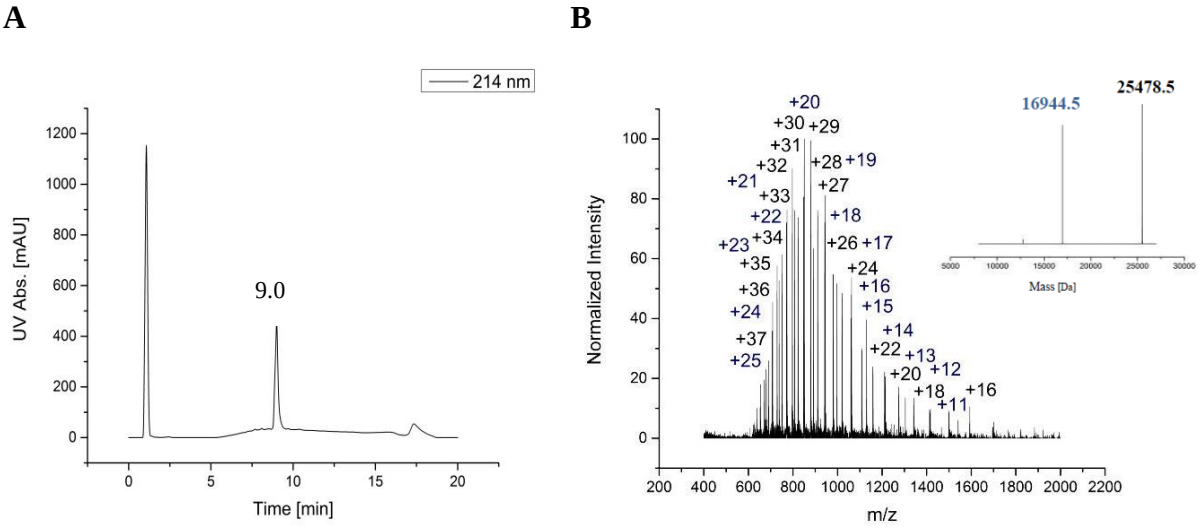


Figure 28: A) HPLC-Chromatogram Ub-NpuDnaE-H_6 before the intein cleavage. B) LC-MS Ub-NpuDnaE-H_6 peak at 9 min: 1) = NpuDnaE-H_6. Calculated masses: 16945.1 Da., observed masses: 16944.5 Da. 2) = Ub-NpuDnaE-H_6. Calculated mass: 25492 Da., observed mass: 25478.5 Da.

The HPLC-chromatogram peak at 9 minutes was analyzed and deconvolution showed masses of 16944.5 Da and 25478.5 Da, corresponding to the calculated masses of NpuDnaE-H_6 (16945.15 Da) and Ub-NpuDnaE-H_6 (25491.98 Da. Difference of 13.5 Da maybe due to modifications during the ionization process). Since both, the cleaved intein and the uncut

precursor could be found in the same peak, a separation of still intact construct was not possible under the applied conditions. The cleaved ubiquitin could not be found in the chromatography fractions.

To determine reaction rates intein cleavage was performed according to protocol described in 2.4.4. At various time points, samples of 10 µl were collected, quenched by addition of TCEP and 2x gel loading buffer (without thiols) and stored at -20°C until completion of the reaction. The subsequently performed SDS-PAGE delivered the following results (Figure 29):

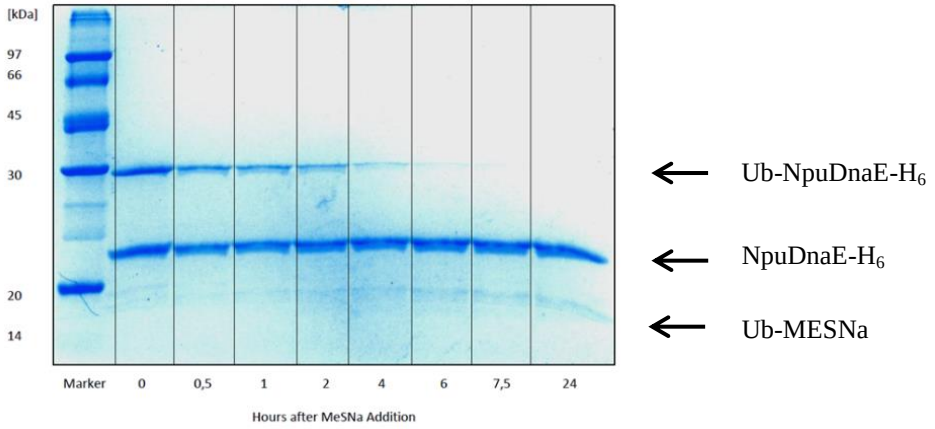
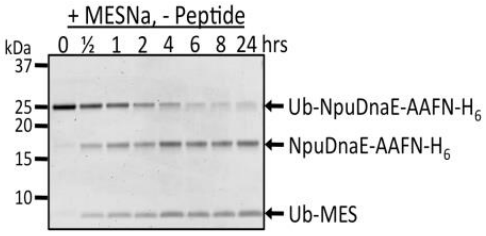


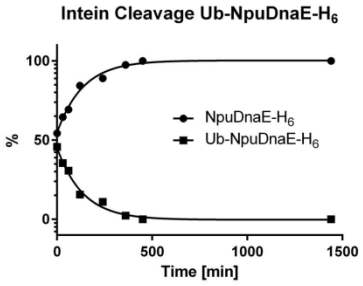
Figure 29: SDS gel of intein cleavage Ub-NpuDnaE-H_6

The band intensity of NpuDnaE-H_6 is increasing while the intensity of the uncut construct is decreasing, consistent with theory of intein cleavage. Between 4-6 hours, almost the whole precursor is cleaved.

A



B



C

k [min^{-1}]	0.0073
half-life [min]	94
uncut product after 24 hours [%]	0

Figure 30: A) SDS gel intein splicing Ub-NpuDnaE-H_6, Shah *et al.*, 2012. B) Percentage of spliced construct versus time Ub-NpuDnaE-H_6, using normalized band intensities obtained from Figure 29. C) Calculation of rate constant and half-life according to equations 1.

Compared to results obtained by Shah *et al.* (Shah *et al.*, 2012; Figure 30A), cleavage time was even shorter and quantitative turnover was achieved after 24 hours (Figure 29B/C). Rate constant and half-life were calculated according to equations 1. The obtained values

confirmed fast splicing times (Figure 30C). Nevertheless, judging by the initial cut-to-uncut ratio, around 50 % of product was lost due to irreversible premature cleavage which occurred before the reaction was started (Figure 29). That means, although the turnover was complete, only half of the theoretical thioester containing ubiquitin yield could be obtained.

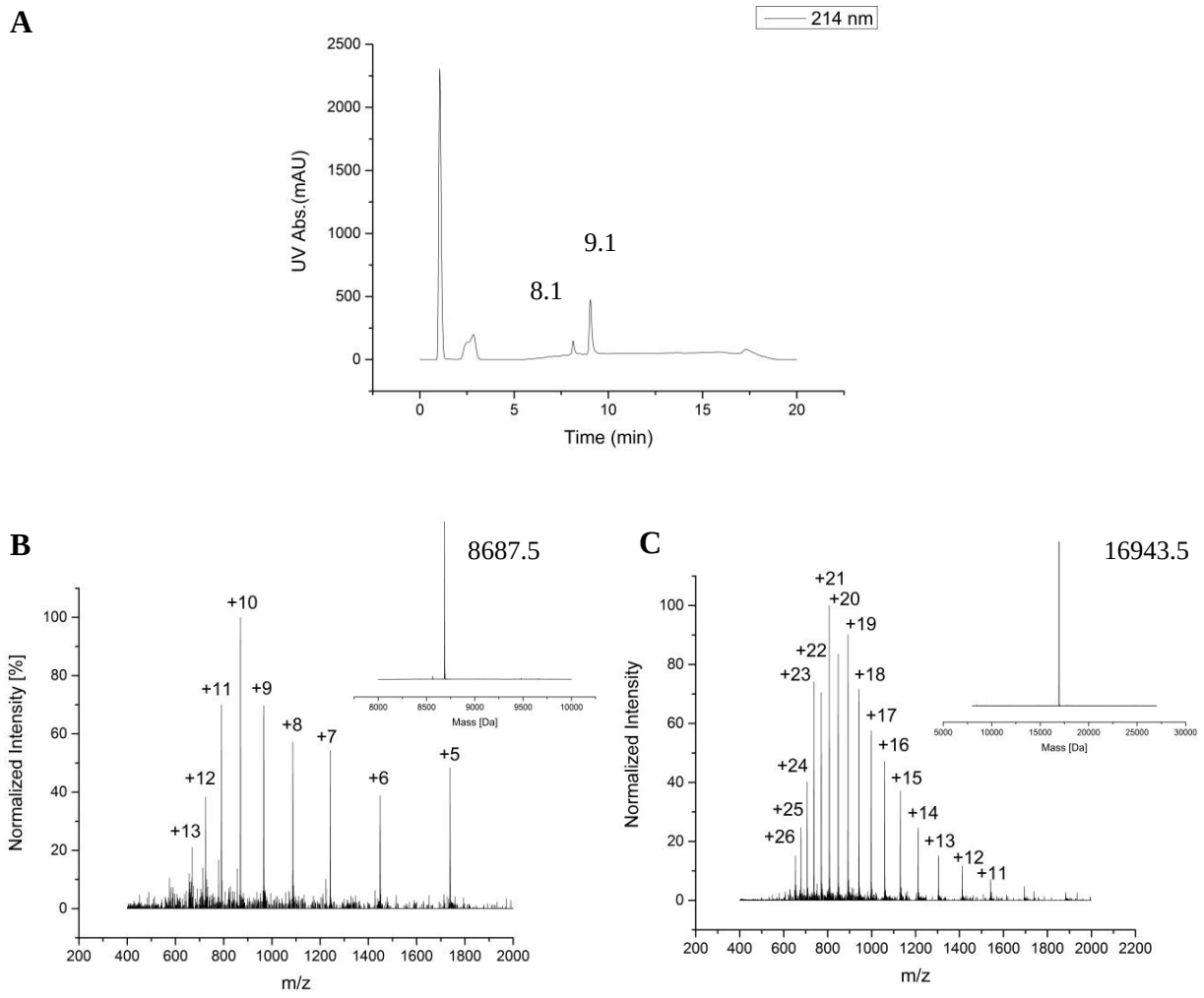


Figure 31: LC-MS Ub-NpuDnaE-H6 after intein cleavage. A) HPLC chromatogramm. B) Spectrum of peak at 9.1 min and deconvolution = Ub-MESNa. Calculated mass: 8683.6 Da., observed mass: 8687.5 Da. C) Spectrum ofpeak at 8.1 min and deconvolution= NpuDnaE-H_6. Calculated masses: 16945.1 Da., observed masses: 16943.5 Da.

To sum it up, thiolysis of Ub-NpuDnaE-H_6 described by Shah *et al.*, 2012 was well reproducible, showing even faster splicing and complete substrate turn over. Unfortunately, premature hydrolysis decreased yield by around 50 %. Since Shah *et al.* only calculated rate constants for NpuDnaE *trans* splicing, a direct quantitative comparison was not possible.

Finally, it can be stated that the advantage of ultrafast splicing times can also turn into disadvantage, considering the diminishing yield, due to premature thioester cleavage. Expression time and temperature as well as the presence of nucleophiles can influence the premature cleavage rate and should be optimizes to achieve the best possible thioester yield.

To observe splicing behavior of ultrafast NpuDnaE inteins fused to more challenging proteins and compare it to previously used MxeGyrA variants, four different expressions were carried out. BL21 (DE3) cells transfected with pTXB3-PrP90-231-Mxe-H_6-CBD, pTXB3-PrP90-231-NpuDnaE-H6, pTWIN1-Tau1-290-Mxe-CBD and pTXB3-Tau1-290-NpuDnaE-H6, respectively, were grown and protein expression was induced by addition of IPTG.

3.2.5 Expression, purification and intein cleavage of PrP-90-231-Mxe-CBD-H_6 and PrP90-231-NpuDnaE-H_6

Both variants were expressed simultaneously, using identical conditions to ensure differences in cleavage assays will only be based on differences in intein performance. After cell growth, expression, cell harvesting, cell disruption and solubilization, Ni-Affinity Chromatography was performed to isolate the desired constructs.

After TCA-precipitation, fractions were loaded onto SDS gels, electrophoresis was performed and the following pictures were obtained (Figure 32). In case of PrP90-231-Mxe-H_6-CBD, the column was clearly overloaded, since the flow through showed an intense band at a similar apparent weight as the elution fractions. The following three wash fractions as well as the elution fractions and the subsequent second wash fractions all contained the some product. During the elution another band appeared as well.

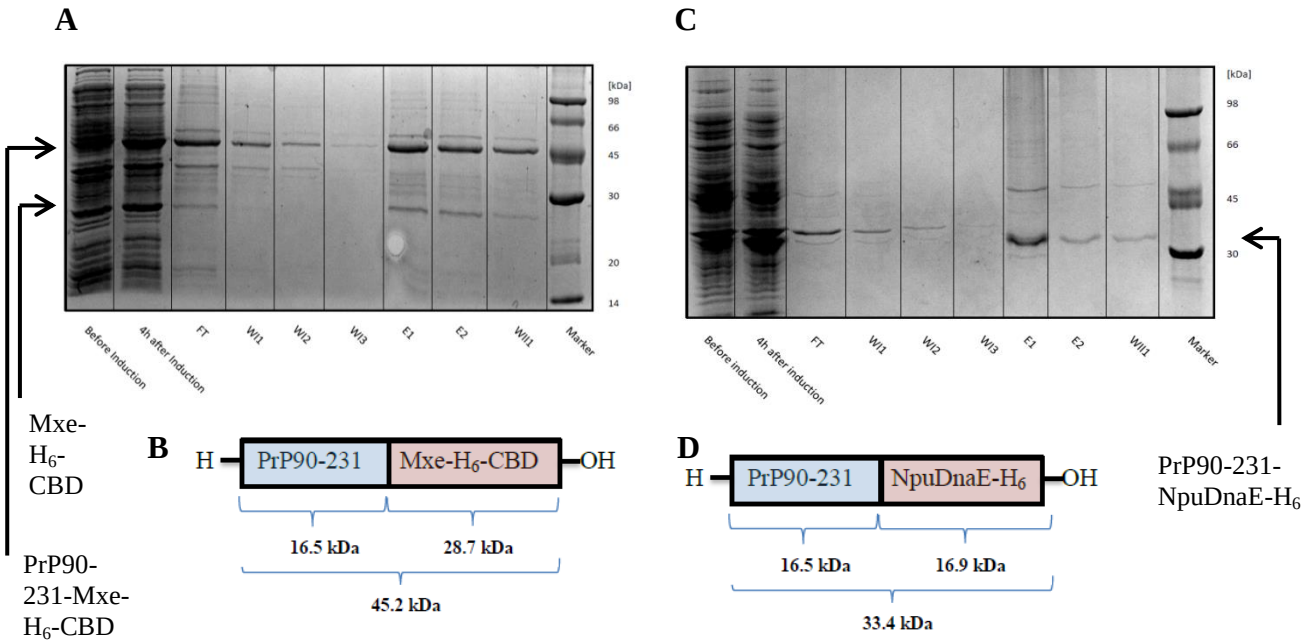


Figure 32: A) SDS gel of Ni-affinity chromatography fractions of PrP90-231-Mxe-H_6-CBD. B) Scheme of expressed construct purified in A. C) SDS gel of Ni-affinity chromatography fractions of PrP90-231-NpuDnaE-H_6. D) Scheme of expressed constructs purified in C.

Judged by size, this could be the already cleaved intein still containing a His-tag, although Figure 34A did not show premature cleavage. SDS gel of PrP90-231-NpuDnaE-H_6 did not show any overload. In this particular case, running time was definitely chosen too long and all fragments smaller than approximately 20 kDa leaked out of the gel, making it impossible to determine the amount of premature hydrolysis. In both gels, a protein band of about 50 kDa could be already found during expression and it was not possible to remove it during affinity chromatography. Based to the size, it is unlikely any kind of dimer containing a His-tag and moreover, it seemed to be the same protein in both gels. Therefore unspecific binding of an undefined *E.coli* protein is the most likely explanation.

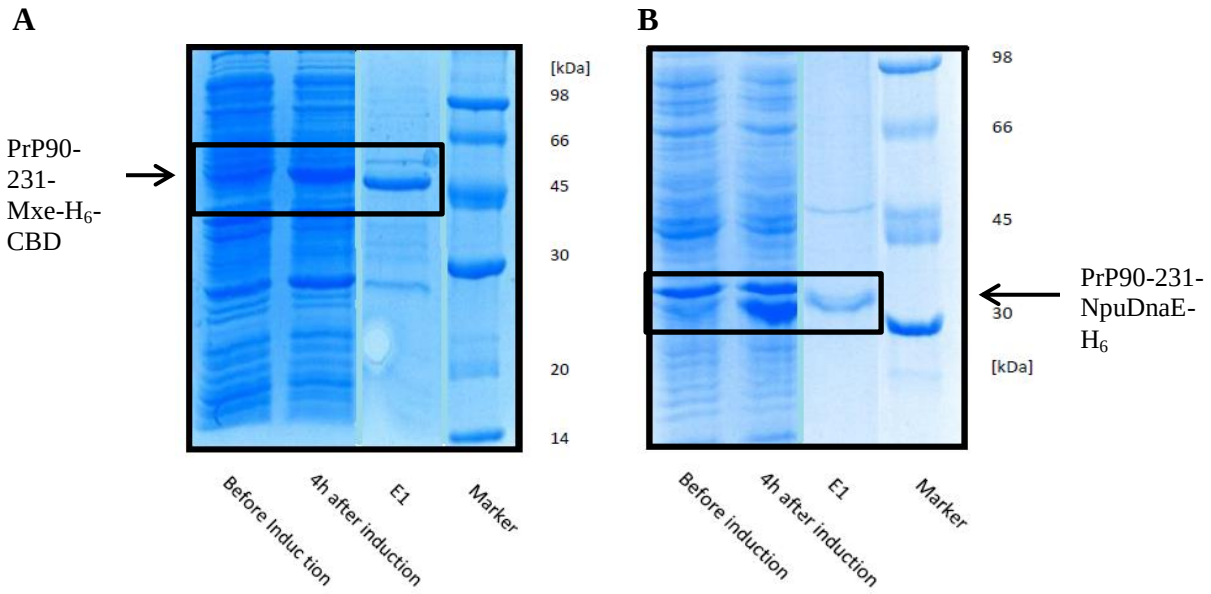


Figure 33: Comparision of overexpression. A) PrP90-231-Mxe-H_6-CBD. B) PrP90-231-NpuDnaE-H_6.

Comparing product band intensities of both variants before and after induction of expression, another advantage of PrP90-231-NpuDnaE-H_6 was demonstrated. A much higher overexpression could be seen in case of PrP90-231-NpuDnaE-H_6 compared to PrP90-231-Mxe-H_6-CBD (Figure 33).

Intein cleavage was performed as described (2.4.4). SDS gel of PrP90-231-Mxe-H_6-CBD showed continuous depletion of the fusion protein over time and increasing intensity of the cleavage fragment band (Figure 34A). The sample at T_0 only showed low amounts of cut fragment, so it can be assumed that the bands at around 17 kDa belong to the newly formed PrP-thioester rather than the hydrolyzed protein. After 24 hours, the splicing activity seemed to stagnate and only a little amount of construct is cut during the next 24 hours.

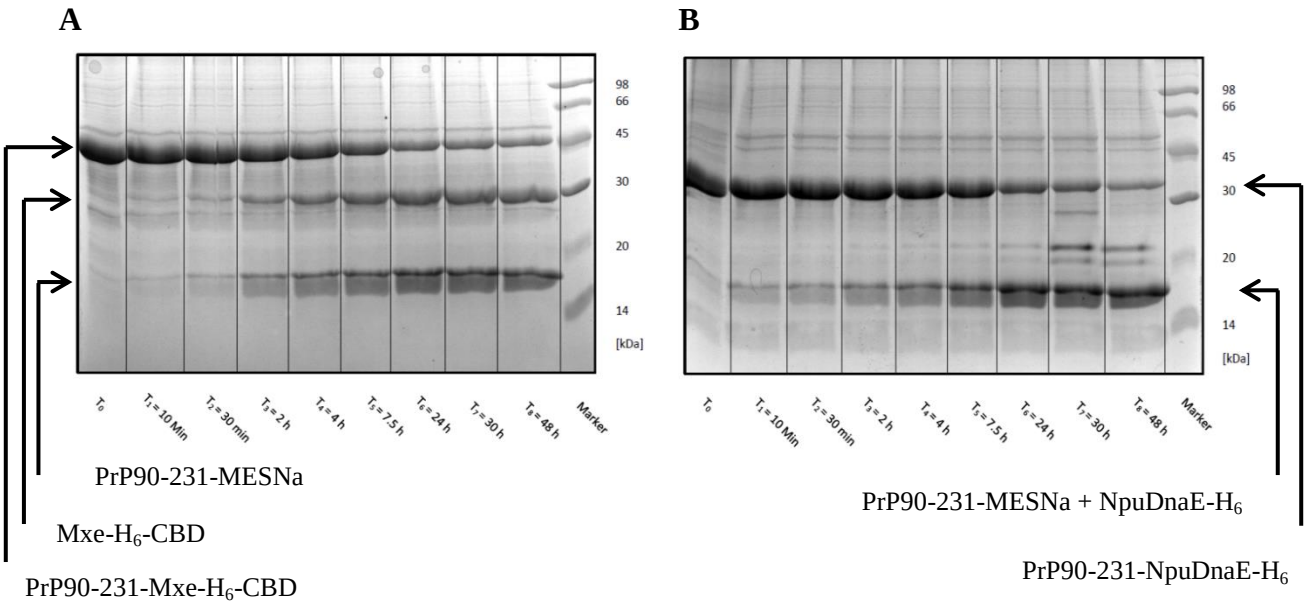
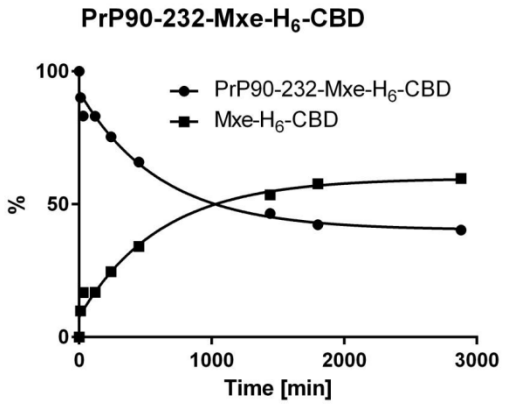


Figure 34: SDS gels of intein cleavage. A) PrP90-231-Mxe-H_6-CBD. B) PrP90-231-NpuDnaE-H_6.

Despite incomplete turnover, the reaction was stopped after 48 hours. Intensity analysis calculated 40 % of remaining uncleaved precursor at the end of reaction (Figure 35B). PrP-thioester created undefined bands ranging from the calculated MW to lower masses, but since repeated SDS-PAGE showed an equal picture, kinetic calculations were performed by using distinct bands on the top of each smear.

Since PrP90-231 and NpuDnaE-H_6 are similar in size (16.5 kDa and 16.9 kDa), the two corresponding bands cannot be distinguished on gel (Figure 34B). Therefore only one cut fragment band with increasing intensity and one whole construct band with decreasing intensity could be observed during splicing reaction. After 24 hours, other additional bands appeared. They seemed to belong to degradation fragments. This conclusion has been drawn because of the size which cannot be explained with dimer formation or the like. Additionally these bands are increasing over time during the reaction at 30°C, which makes degradation a likely explanation. Possibly introduced proteases could have helped to fragment the protein. Once again, there was still uncleaved precursor left after the end of reaction after 48 hours. By calculating percentage intensities, uncut precursor added up to 22 % (Figure 35D). Although still not completed, this reaction left only half of the amount of uncut product compared to intein cleavage using Mxe-variants (Identity of all components was further confirmed by LS-MS, see 6). Figure 35 is again revealing a larger amount of premature cleavage in case of very fast splicing NpuDnaE intein.

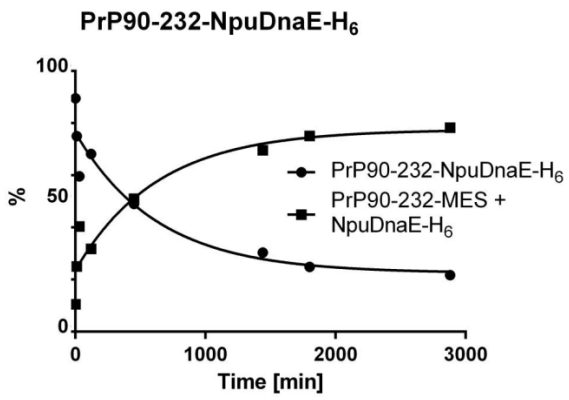
A



B

k [min^{-1}]	$7.6x10^{-3}$
half-life [min]	425
uncut product	
after 24 hours [%]	46
after 48 hours [%]	40

C



D

k [min^{-1}]	$1.6x10^{-3}$
Half Life [min]	435
uncut product	
after 24 hours [%]	30
after 48 hours [%]	22

Figure 35: Relative band intensity versus time using normalized band intensities obtained from Figure 34.Calculation of rate constant and half-life according to equations 1. A) + B) PrP90-231-Mxe-H_6-CBD, C) + D) PrP90-231-NpuDnaE-H_6.

Rate constants are almost identical, meaning there is no advantage concerning cleavage time in case of prion protein fused NpuDnaE-H_6 intein compared to previously used MxeGyrA variants. Nevertheless, the new construct showed higher overexpression and a much better, even if not completed turnover (Figure 33).

3.2.6 Expression, purification and intein cleavage of Tau1-290-Mxe-CBD and Tau1-290-NpuDnaE-H_6

Also in case of Tau fusion proteins, both variants were expressed under identical conditions. After cell harvesting, further purification steps were split up, due to the lack of His-tag in previously used construct. It has to kept in mind that also the cleavage assays had different basic requirements (cleavage in solution versus cleavage on solid support), making cleavage rates of Tau variants not quantitatively comparable in the proper sense. Tau1-290-NpuDnaE-H_6 was purified *via* affinity chromatography and results of subsequently performed SDS-PAGE can be seen in Figure 36C. Tau1-290-Mxe-CBD was purified by taking advantage of the chitin binding domain fused to the protein-intein-construct. After binding to chitin functionalized resin all impurities were washed out. Elution and intein cleavage were performed in one step by incubation with 200 mM MESNa. Thereby, the

protein was released containing a C-terminal thioester, whereas the intein still remained on the beads and could be stripped off later by the addition of NaOH solution. At every step, samples for SDS-PAGE were taken and results are presented in Figure 36A. SDS gel of Tau1-290-NpuDnaE-H_6 Ni-Affinity Chromatography indicated a lot of already cleaved Tau protein in fractions FT-WI3 (Figure 36C). In elution I, II and the following wash fraction, this protein could only be found in traces, due to the lacking His-tag. However, cleaved intein fragment could not be separated from uncut precursor. Also other impurities were still in the elution fractions when they were combined, concentrated and used for intein cleavage assay. In SDS gels, all Tau protein containing bands appeared approximately 10 kDa higher than their calculated masses.

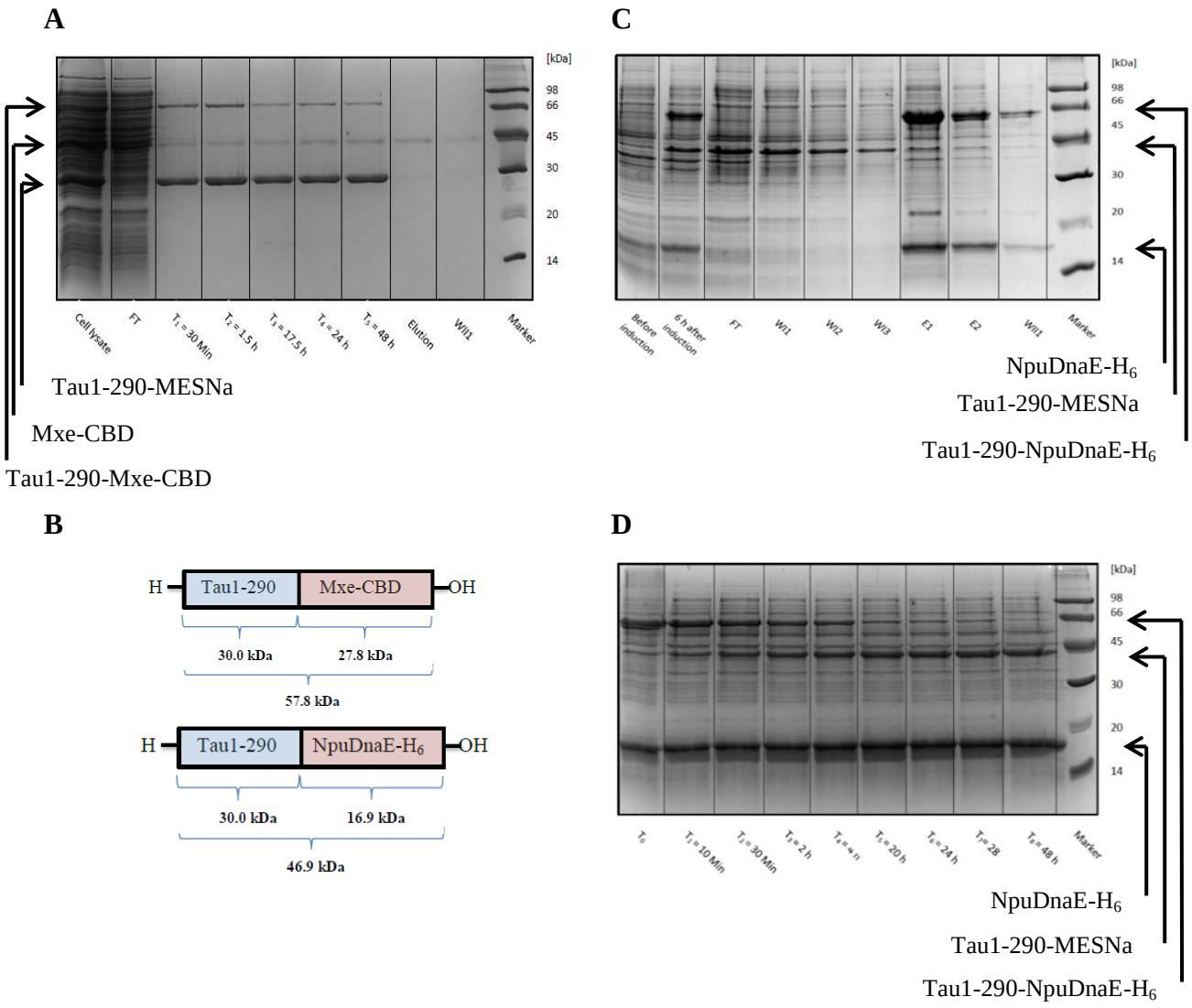


Figure 36: A) SDS gel of chitin beads/intein cleavage fractions Tau1-290-Mxe-CBD . B) Scheme of expressed constructs. C) SDS gel of Ni-affinity chromatography fractions Tau1-290-NpuDnaE-H_6 . D) SDS gel intein cleavage Tau1-290-NpuDnaE-H_6 .

This phenomenon was already observed and explained in 1988 by Lee *et al.* by tracing it back to the unstructured nature and the correlated large surface-to-volume ratio (Lee *et al.*, 1988). LC-MS measurements confirmed identity (see 6). Intein cleavage of Tau1-290-NpuDnaE-H_6 worked quite well, whole construct band intensity is decreasing while thioester formation could be observed by the increasing band intensity in the SDS-gel (Figure 36D). Only little of prematurely cleaved Tau protein was present at t_0, meaning it can be assumed that bands at 44 kDa are corresponding to Tau-MESNa thioester. However, subsequent LC-MS measurements showed three different Tau species: Tau-OH (-Methionine), Tau-MESNa (-Methionine) and Tau-MESNa (+Methionine), although the distinction between Tau-OH (+Methionine) and Tau-MESNa (-Methionine) is difficult since the mass difference is only about 7 Da. (see 6).

One problem of the previously used vector was the unsatisfying yield due to very little overexpression of the desired construct (Görner, 2011). Comparing the band intensities of "new" and "old" variants before and 6 hours after expression in Figure 37, again new NpuDnaE-H_6 containing compound fused into a pTXB3 vector instead of pTWIN1 showed a much larger amount of overexpression.

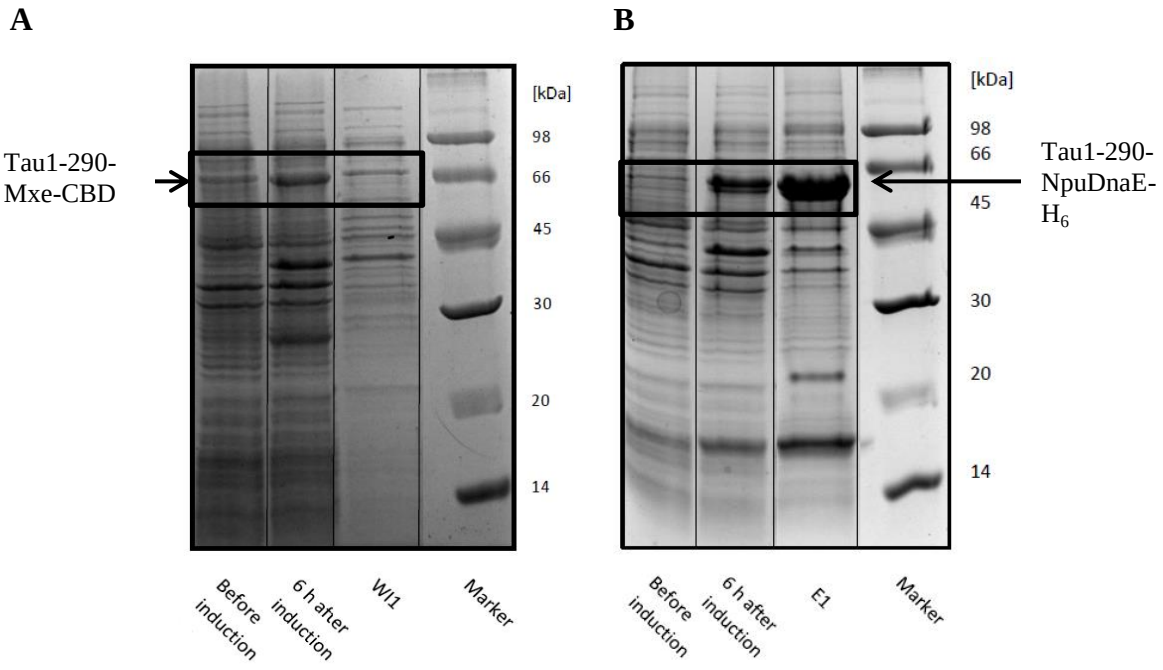


Figure 37: Comparision of overexpression. A) Tau1-290-Mxe-CBD. B) Tau1-290-NpuDnaE-H_6.

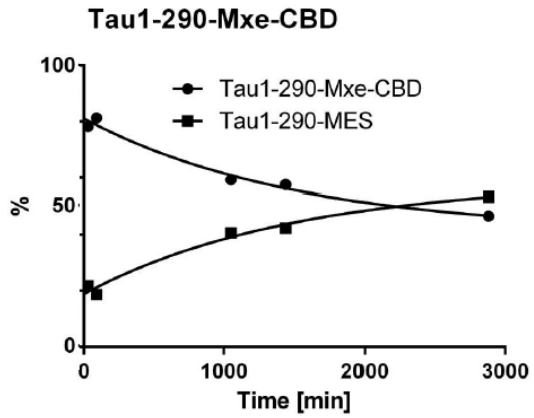
Cleavage of Tau1-290-Mxe-CBD was induced directly after washing while still bound to chitin resin. At various time points, 50µl of resin were collected, boiled in 50µl of 2x gel

loading buffer (without thiols) for 5 minutes and centrifuged afterwards. Supernatant was used for SDS-PAGE (Figure 36A).

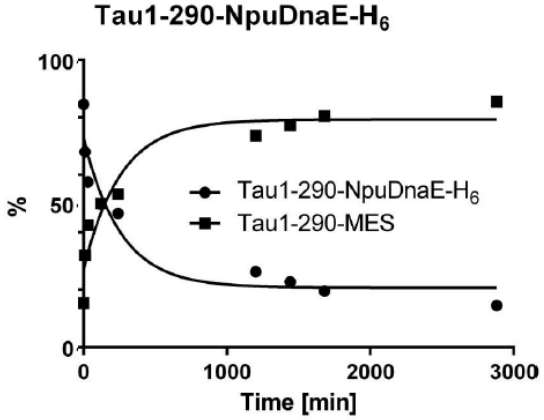
The addition of loading buffer and subsequent boiling lead to the release of precursor as well as cleaved intein from chitin functionalized agarose beads. The ratio of cleaved to uncleaved constructs indicates cleavage efficiency. Since precursor and intein are still bound to resin, there should be only thioester present in elution fraction. Slight contaminations with cleaved intein could be seen in elution fractions and subsequent wash fractions (Figure 36A). Apart from that, Tau-thioester could be obtained in better purity than from the Tau1-290-NpuDnaE-H_6 construct.

When comparing rate constants (Figure 38B/D), a difference of one order of magnitude could be found between MxeGyrA and NpuDnaE fusion constructs. Also turnover is much more advanced after 24 as well as 48 hours in case of the DnaE intein from *Nostoc punctiforme*. Tau-thioester and PrP-thioester were purified separately using a Varian HPLC system and conditions described in 2.4.6. Protein containing fractions were combined, lyophilized and yields were determined by weighing dry protein powder. Since this study was not primarily designed to compare yields, no quantitative numbers are available.

A



C



B

k [min^{-1}]	$6.7x10^{-4}$
half-life [min]	1041
uncut product	
after 24 hours [%]	58
after 48 hours [%]	47

D

k [min^{-1}]	$3.9x10^{-3}$
Half Life [min]	179
uncut product	
after 24 hours [%]	23
after 48 hours [%]	14

Figure 38: Relative band intensity versus time using normalized band intensities obtained from Figure 36.Calculation of rate constant and half-life according to equations 1. A) + B) Tau1-290-Mxe-CBD C) + D) Tau1-290-NpuDnaE-H_6.

4 Discussion and Conclusions

4.1 Isolation of homogenous GPI anchors from yeast

The aim of the experiments described in 2.2 was to obtain GPI anchors *via* anti-FLAG antibody coupled resin as an alternative to the already established but not convenient purification method for the GPI-anchored fusion construct using covalent capture beads (Schumacher *et al.,* 2010). In theory, the new method should be easy and does not require laborious preparation steps, since the antibody-antigen interaction is based on high affinity and specificity. Unfortunately, purification using different reaction conditions was not successful in any case. No separation of desired FLAG-Gas1p-GPI from PrP-H_6 was achieved. Also a repeated Ni-Affinity Chromatography to remove His-tagged PrP did not produce pure FLAG-Gas1p-GPI.

4.1.1 Absence of the FLAG-tag in elution fractions under test 1 conditions

After TEV protease cleavage, the FLAG sequence appended by an N-terminal cysteine residue represented the N-terminus of the desired construct. Since the N-terminal exposed FLAG-tag should have an especially high binding affinity towards the here used antibody (Einhauer and Jungbauer, 2001), inefficient binding to the resin is questionable. According to the ProtParam tool from ExPasy, FLAG-Gas1p511-528 (CG*DYKDDDDK*SSSSSSSA SSSSSSKKN) is categorized as unstable due to an *in vitro* instability index of 85 (proteins with an instability index > 40 are defined as unstable) using the method established by Guruprasad *et al*. (Guruprasad *et al.,* 1990) for calculations. On the other hand, studies showed that FLAG-tag is neither removed by proteases in *E.coli* nor degraded in *Saccharomyces cerevisiae* (Plückthun, 1992; Einhauer und Jungbauer, 2001). In conclusion, the final reason for the absence of positive FLAG signals in the elution fractions under test 1 conditions could not be determined without further examinations.

4.1.2 Presence of redundant fragment in elution fractions under test 2, test 4 and test 5 conditions

Two possibilities could explain this observation: one interpretation would be that PrP bound somewhere unspecifically to the antibody or agarose resin and was released when coming into contact with the elution buffer. Since reducing conditions could not be used due to

incompatibility with antibodies, another cause of separation failure could be the formation of disulfide bonds. N-terminal exposed cysteine of FLAG-Gas1p-GPI could interact with Cys78 or Cys213 of PrP22-230, making a separation impossible. This explanation is most likely under test 4 since incubation was carried out overnight (Table 2).

4.1.3 Presence of redundant fragment in all fractions under test 3 conditions

As already mentioned before, this phenomenon could be caused by insufficient washing steps. However, approximately the same amount of protein was used and also washing steps were not significantly decreased compared to the other test conditions.

4.1.4 Absence of FLAG in flow through and washing fractions after Ni-Affinity Chromatography

Again, degradation or unspecific binding would be in line with the observations. Since there was still no positive FLAG signal in any of the subsequently obtained fractions after treatment with 70% isopropanol or OG containing urea buffer (data not shown), destruction or distortion of the antibody recognizing epitope is the most reasonable explanation.

In summary, the separation of Cys-FLAG-Gas1p511-528-GPI from PrP22-230 with anti-FLAG antibody modified agarose resin did not work potentially due to supposed unspecific binding, epitope degradation, disulfide bond formation or combinations thereof. Another problem could be the insolubility, since only low amounts of OG could be used in combination with the antibodies. Instead, in case of the secondly performed nickel affinity chromatography, OG was used in the CMC but did not seem to improve the protein solubility either. Both, the separation and the analysis have to be improved during further studies. To prevent unspecific binding to the resin, other possibilities such as trying the same procedure with a simpler and less hydrophobic fusion protein instead of PrP should be taken into account when planning new strategies. The already available Kre5p-GFP-H_6-FLAG-Gas1p construct from Schumacher *et al.*, 2010 could be an alternative. Also precipitation techniques using different solvents could be tried, although already performed acetone and ethanol precipitation did not have any separation effect (data not shown). Since the desired product contains a highly charged and hydrophilic peptide and sugar moiety attached to hydrophobic fatty acids, chromatography methods are difficult to apply. Nevertheless, Hydrophilic Interaction Liquid Chromatography (HILIC) or simple Normal Phase High Performance Liquid Chromatography (NP-HPLC) could be tried to overcome drawbacks of

previously used Reverse Phase HPLC, where the hydrophobic character of GPI anchors led to strong interactions with the stationary phase and in consequence caused elution problems. Also detection via immunoreactions cannot provide all information required for precise identification. SDS-PAGE using a higher content of acrylamide could help to assign a defined mass to the FLAG containing construct. Even more information could be gained by proper MALDI measurements, although sample desalting appeared to be a problem and therefore no results were obtained (data not shown).

4.2 Expression, purification and comparison of reactions rates of MxeGyrA and NpuDnaE intein constructs

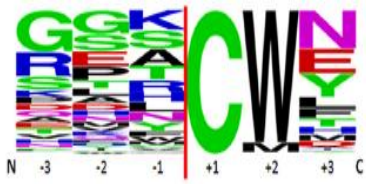
All results presented in 0 suggest the use of ultrafast NpuDnaE for generation of C-terminal thioester containing Tau1-290 and PrP90-231 has some advantages over previously used MxeGyrA variants. However, the advantages are not as dramatic as expected after seeing results for ubiquitin fusion constructs (3.2.4). Altogether, the increased overexpression (Figure 33, Figure 37), faster reaction rates and higher turnover (Figure 35, Figure 38) are advantageous. Nevertheless, there are still some concerns when comparing the excellent ubiquitin results with whose from Tau and PrP, specifically with respect to rate constants and turnover.

It is long known, that reaction kinetics and overall yield are highly influenced by flanking amino acids at position -1, -2, -3, +2 and +3 (Shah and Muir, 2014; Amitai *et al.*, 2009; Shah *et al.*, 2013). Cheriyan *et al.* could demonstrate the connection between splicing efficiency and every possible amino acid combination by observing cell survival in Kan^R-assays. Therefore, NpuDnaE mutants were fused into *C.diphtheriae* aminoglycoside phosphotransferase, an enzyme which provides kanamycin resistance if correctly expressed after successful splicing (Cheriyan *et al.*, 2013). In Figure 39 selectivity ratios for every particular amino acid combination can be seen. Surprisingly, the endogenous sequence is only the best choice for position +3. At position +2, tryptophan is clearly preferred. This fact could be ascribed to a compromise between extein representing protein sequence integrity and splicing kinetics (Cheriyan *et al.*, 2013). At the N-terminal extein a broader range of sequence variation is tolerated. In this regard, a sequence comparison might explain the observed differences in reaction kinetics and yield of tested NpuDnaE fusion constructs.

A

Amino Acid	N-terminal Extein			C-terminal Extein	
	(-3)	(-2)	(-1)	(+2)	(+3)
D	1.39	1.52	0.00	0.00	1.26
E	1.26	*3.51	0.07	0.00	5.50
N	0.93	1.26	2.19	0.00	*8.61
Q	1.39	0.86	0.33	0.00	0.27
H	1.06	1.26	2.25	0.00	1.59
K	1.59	0.93	4.44	0.07	0.00
R	1.81	0.20	0.86	0.00	0.00
S	0.99	1.24	1.37	0.00	0.07
C	0.46	0.40	0.80	0.07	0.07
T	0.63	0.89	1.59	0.00	0.96
P	0.80	1.52	0.00	0.00	0.00
G	4.67	2.68	0.17	0.00	0.03
A	*0.63	0.83	1.92	0.00	0.23
V	0.56	0.50	0.23	0.03	0.10
I	0.07	0.53	0.00	0.00	0.40
L	0.11	0.57	0.75	0.00	0.82
M	0.13	1.13	0.86	2.19	1.46
F	0.13	0.73	1.26	*0.07	2.25
Y	0.20	0.60	*2.05	0.13	5.23
W	0.07	0.40	0.99	29.42	0.07
STOP	0.00	0.00	0.00	0.00	0.00

B



A) The selectivity ratio is defined as the observed frequency divided by the theoretical frequency. Selectivity ratio values near or equal to 1 indicate no selection. Higher selectivity ratios indicate stronger selective pressure for that amino acid, whereas values under 1 indicate negative selective pressure for that amino acid or high selective pressure for another amino acid. The native flanking NpuDnaE extein residues are labeled with asterisks. The+1 position was fixed to the native Cys residue and was thus not included in this table. The intein is present between the -1 and +1 position.

B) Logo diagrams of flanking extein sequences for 483 randomly chosen positive hits selected for splicing of the Npu DnaE intein. The *vertical red line* indicates the position of the intein. The+1 position was fixed as Cys.

Figure 39: A) Selectivity ratio for the plasmid population that survived kanamycin selection (Cheriyan *et al.*, 2013). B) Sequence logo of NpuDnaE flanking extein residues (Cheriyan *et al.*, 2013).

The plasmid obtained from the Muir group contained the endogenous sequence NCFN at the C-extein positions with mutations of asparagine and cysteine to alanine, leading to N-terminal splicing and therefore making the generation of thioesters possible (Shah *et al.*, 2012; Sequence can be found in the supplementary information 6). As can be seen in Table 9, this might not be the best arrangement.

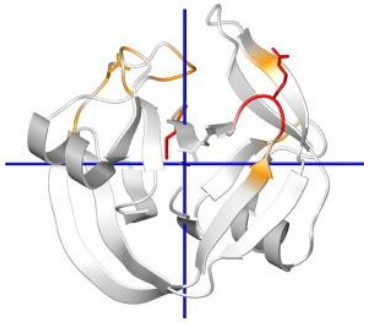
As already mentioned, also residues at position -1, -2 and -3 seem to influence splicing kinetics. In protein-intein fusion constructs, these positions are represented by the C-terminal amino acids of fusion protein. Meaning every protein with a different amino acid sequence at the C-terminus can alter splicing behavior.

	-3	**-2**	**-1**
DnaE	A	E	Y
Ubiquitin	R	G	G
Tau1-290	Q	S	K
PrP90-231	R	S	S
“best combination” (according to Figure 39A)	G	E	K

Table 9: Comparison of N-extein residues in native context, in Tau1-290, PrP90-231 or ubiquitin fusion constructs and “best combination” according to Figure 39A

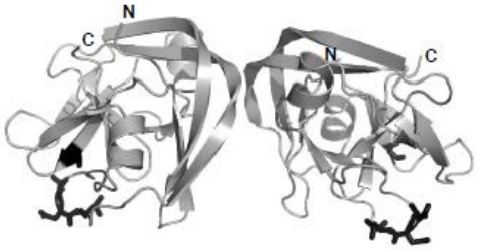
Table 9 shows the final three C-terminal amino acids of native α-subunit of DNA polymerase III N-extein (DnaE), Tau1-290, PrP90-231, ubiquitin and "best combination" according to Figure 39A. At position -3, small uncharged amino acids like alanine and glycine can be found in native context as well as in the proposed "best combination". Nevertheless, also a big and charged amino acid like asparagine can obviously be tolerated since ubiquitin showed acceptable splicing kinetics. PrP90-231 and Tau1-290 also feature bulky amino acids at this location (R, Q). Hence, this position might not be critical in this case. Position -2 is identically represented by glutamic acid in DnaE and "best combination". Again ubiquitin showed an amino acid with highly different characteristics. Instead of a negatively charged amino acid, a glycine can be found at this position. PrP90-231 as well as Tau1-290 showed the polar but uncharged amino acid serine. Lysine should be the most efficient amino acid at position -1 according to the "best combination" and can be also found in Tau1-290. Ubiquitin shows again glycine while PrP90-231 possesses a serine and tyrosine is present in native DnaE N-intein sequence. To sum it up, no obvious pattern that allows improved splicing based on the last three N-terminal amino acids of ubiquitin, PrP90-231, Tau1-290, native DnaE N-extein and "best combination" suggested by Cheriyan *et al.* was identified. Consequently, N-extein sequence variability is indeed not the decisive factor for ultrafast cis splicing and thioester generation of NpuDnaE intein PrP90-231 and Tau1-290 fusion constructs.

A



Ribbon drawing of the lowest energy conformer of the single chain variant of NpuDnaE intein. Calculated axes of the structure are displayed by the three lines, indicating the center of the structure. The horizontal line is the longest axis. The shortest axis of the structure is perpendicular to the plane of the page. The residues of which amide groups were not detectable in the NMR spectra were colored in red. The residues indicating conformational exchanges are colored in orange (Oeemig *et al.*, 2009)

B



The crystal structure of monomeric NpuDnaE intein. Two molecules in the asymmetric unit are shown. Four residues deleted in Δvariant are shown in stick model and in dark grey. N and C denote the N- and C-termini, respectively. (Aranko *et al.*, 2013)

Figure 40: A) Solution structure of NpuDnaE obtained by NMR measurements (Oeemig *et al.*, 2009). B) Crystal structure of NpuDnaE obtained by X-ray analysis (Aranko *et al.*, 2013).

Interaction of the fusion protein with the intein folding could also be a problem for splicing efficiency. Looking at the structure of NpuDnaE intein (Figure 40), it could be possible that the fused protein is somehow interfering with the intein by projecting into the horseshoe-shaped groove and is thus preventing proper splicing.

However, all these explanations are hypotheses, which have to be proven experimentally in further studies either by mutation studies or co-crystal structures. On the basis of the obtained results in this study, no further conclusions about the splicing kinetics of newly produced NpuDnaE fusion constructs could be drawn.

5 Literature

- Abid, K., Morales, R., Soto, C., Cellular Factors implicated in Prion Replication, *FEBS Lett*., **2010**, 584 (11): 2409–2414.
- Abid, K., Soto, C., The intriguing prion disorders, *Cell Mol Life Sci*., **2006**, 63(19-20): 2342–2351.
- Aguzzi, A., Heikenwalder, M., Pathogenesis of prion diseases: current status and future outlook, *Nature*, **2006**, 4: 765-775.
- Aguzzi, A., Sigurdson, C., Heikenwaelder, M., Molecular Mechanisms of Prion Pathogenesis, *Annu. Rev. Pathol. Mech. Dis*., **2008**, 3: 11–40.
- Alpers, M.P., The epidemiology of kuru: monitoring the epidemic from its peak to its end, *Phil. Trans. R. Soc. B*, **2008**, 363: 3707–3713.
- Amitai, G., Callahan, B.P., Stanger, M.J., Belfort, G., Belfort, M., Modulation of intein activity by its neighboring extein substrates, *PNAS*, **2009**, 106 (27): 11005-11010.
- Aranko, A.S., Oeemig, J.S., Kajander, T., Iwaï, H., Supplementary information of: Intermolecular domain swapping induces intein-mediated protein alternative splicing, *Nature Chem Biol*, **2013**, 9: 616-622.
- Aranko, A.S., Wlodawer, A., Iwai, H., Nature's recipe for splitting inteins, *Protein Eng Del Sel*, **2014**, 27 (8): 263-271.
- Basler, K., Oesch, B., Scott, M., Westaway, D., Wälchli, M., Groth, D.F., McKinley, M.P., Prusiner, S.B., Weissmann, C., Scrapie and cellular PrP isoforms are encoded by the same chromosomal gene, *Cell*, **1986**, 46(3): 417-28.
- Becker C. F., Hunter C. L., Seidel R., Kent, S. B., Goody R. S., Engelhard M., Total chemical synthesis of a functional interacting protein pair: the protooncogene H-Ras and the Rasbinding domain of its effector c-Raf1, *Proc. Natl Acad. Sci. USA*, **2003**, 100: 5075–5080.
- Becker, C.F.W., Liu, X., Olschewski, D., Castelli, R., Seidel, R., Seeberger, P.H., Semisynthesis of a Glycosylphosphatidylinositol-Anchored Prion Protein, *Angew. Chem. Int. Ed*., **2008**, 47: 8215 –8219
- Becker C.F.W., Proteinsemisynthese — ein Werkzeug für komplex modifizierte Proteine, *BIOspektrum*, **2011**, 17(3): 291-293.

- Binschik, J., Zettler, J., Mootz, H.D., Photocontrol of Protein Activity Mediated by the Cleavage Reaction of a Split Intein, *Angew. Chem. Int. Ed*., **2011**, 50: 3249 –3252.
- Buskirk, A.R., Ong, Y.-C., Gartner, Z.J., Liu, D.R., Directed evolution of ligand dependence: Small-molecule-activated protein splicing, *PNAS*, **2004**, 101 (29): 10505-10510.
- Bütikofer, P., Malherbe, T., Boschung, M., Roditi, I., GPI-anchored proteins: now you see'em, now you don't, *FASEB J*., **2001**, 15: 545–548.
- Camarero, J.A., Muir, T.W., Biosynthesis of a Head-to-Tail Cyclized Protein with Improved Biological Activity, *J. Am. Chem. Soc.*, **1999**, *121*, 5597-5598.
- Carvajal-Vallejos, P., Pallissé, R., Mootz, H.D., Schmidt, S.R., Unprecedented Rates and Efficiencies Revealed for New Natural Split Inteins from Metagenomic Sources, *J Biol Chem*, **2012**, 287 (34): 28686-28696.
- Caughey, B., Raymond, G.J., Ernst, D., Race, R.E., N-Terminal Truncation of the Scrapie-Associated Form of PrP by Lysosomal Protease(s): Implications Regarding the Site of Conversion of PrP to the Protease-Resistant State, *J Virol*, **1991**, 65 (12): 6597-6603.
- Chattopadhyay, A., London, E., Fluorimetric Determination of Critical Micelle Concentration Avoiding Interference from Detergent Charge, *Anal Biochem*, **1984**, 139: 408-412.
- Cheriyan, M., Pedamallu, C.S., Tori, K., Perler, F., Faster Protein Splicing with the *Nostoc punctiforme* DnaE Intein Using Non-native Extein Residues, *J Biol Chem*, **2013**, 288 (9): 6202-6211.
- Chesebro, B., Trifilo, M., Race, R., Meade-White, K., Teng, C., LaCasse, R., Raymond, L., Favara, C., Baron, G., Priola, S., Caughey, B., Masliah, E., Oldstone, M., Anchorless Prion Protein Results in Infectious Amyloid Disease Without Clinical Scrapie, *Science*, **2005**, 308 (3): 1435-1439.
- Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F.B., Xu, M.-Q., Protein Splicing Involving the *Saccharomyces cerevisiae* VMA Intein: the steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an *in vitro* splicing system, *J Biol Chem*, **1996**, 271 (36): 22159-22168.

- Chong, S., Mersha, F.B., Comb, D.G, Scott, M.E., Landry, D., Vence, L.M., Perler, F.B., Benner, J., Kucera, R.B., Hirvonen, C.A., Pelletier, J.J., Paulus, H., Xu, M.-Q., Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element, *Gene*, **1997**, 192: 271–281.
- Cobb, N.J., Surewicz, W.K., Prion Diseases and Their Biochemical Mechanisms, *Biochem.*, **2009**, 48 (12): 2574–2585.
- Cook, S.N., Jack, W.E., Xiong, X., Danley, L.E., Ellman, J.A., Schultz, P.G., Noren, C.J., Photochemically lnitiated Protein Splicing, *Angew. Chem. In. Ed. Engl*., **1995**, 34 (15): 1629-1630.
- Colby, D.W., Wain, R., Baskakov, I.V., Legname, G., Palmer, C.G., Nguyen, H.-O.B., Lemus, A., Cohen, F.E., DeArmond, S.J., Prusiner, S.B., Protease-Sensitive Synthetic Prions, *PLoS Pathog*, **2010**, 6 (1): e1000736.
- David R., Richter M.P.O., Beck-Sickinger A.G., Expressed protein ligation, *Eur. J. Biochem.*, **2004**, 271: 663–677.
- Dawson P.E., Muir T.W., Clark-Lewis I., Kent S.B., Synthesis of proteins by native chemical ligation, *Science*, **1994**, 266 (5186): 776-779.
- Dawson P.E. , Churchill M.J. , Ghadiri M.R., Kent S.B.H., Modulation of Reactivity in Native Chemical Ligation through the Use of Thiol Additives, *J. Am. Chem. Soc.*, **1997**, 119 (19): 4325–4329.
- Dawson, P.E., Kent, S.B.H., Synthesis of native proteins by chemical ligation, *Ann. Rv. Biochem.*, **2000**, 69: 923-960.
- De la Cruz, J., Daugeron, MC., Linder, P., Smart genetic screens, *Yeast Gene Analysis (Brown, A. J. P., Tuite M. F., eds.), Academic Press, Boca Raton, FL*, **1998**, 269–295.
- Deleault, N.R., Harris, B.T., Rees, J.R., Supattapone, S., Formation of native prions from minimal components in vitro, *PNAS*, **2007**, 104 (23): 9741–9746.
- Du, Z., Shemella, P.T., Liu, Y., McCallum, S.A., Pereira, B., Nayak, S.K., Belfort, G., Belfort, M., Wang, C., Highly conserved histidine plays a dual catalytic role in protein splicing: a pKa shift mechanism, *J Am Chem Soc*, **2009**, 131:11581–11589.
- Edelhoch, H., Spectroscopic determination of tryptophan and tyrosine in proteins, *Bio Chem*, **1967**, 6 (7): 1948-1954.

- Einhauer, A., Jungbauer, A., The FLAGe peptide, a versatile fusion tag for the purification of recombinant proteins, *J. Biochem. Biophys. Methods*, **2001**, 49: 455-465.
- Esposito, D., Chatterjee, D.K., Enhancement of soluble protein expression through the use of fusion tags, *Curr Opin Biotech*, **2006**, 17: 353-358.
- Fevrier, B., Vilette, D., Archer, F., Loew, D., Faigle, W., Vidal, M., Laude, H., Raposo, G., Cells release prions in association with exosomes, *PNAS*, **2004**, 101 (26): 9683-9688.
- Fujita, M., Yoko-o, T., Jigami, Y., Inositol deacylation by BST1P is required for the quality control of glycosylphosphatidylinositol-anchored proteins, *Mol. Biol. Cell*., **2006**, 17: 834-850.
- Frutos, S., Goger, M., Giovani, B., Cowburn, D., Muir, T.W., Branched Intermediate Formation Stimulates Peptide Bond Cleavage in Protein Splicing, *Nat Chem Biol*, **2010**, 6 (7): 527-533.
- Gambetti, P., Kong, Q., Zou, W., Parchi, P., Chen, S.C., Sporadic and familial CJD: classification and characterization, *Brit Med Bull*, **2003**, 66: 213–239.
- Geier, G.E., Modrich, P., Recognition Sequence of the dam Methylase of Escherichia coli K12 and Mode of Cleavage of Dpn I Endonuclease, *J Biol Chem*, **1979**, 254 (4): 1408-1413.
- Gill, S.C., Hippel, P.H., Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data, *Anal Biochem*, **1989**, 182:319-326.
- Gogarten, J.P., Senejani, A.G., Zhaxybayeva, O., Olendzenski, L., Hilario, E., Inteins**:** Structure, Function, and Evolution, *Annu. Rev. Microbiol*., **2002**, 56: 263–87.
- Gogolin, L., Seidel, R., Engelhard, M., Goody, R.S., Becker, C.F.W., Semisynthesis of human thymidine monophosphate kinase, *Biopolymers Pept Sci*, **2010**, 94 (4): 433-440.
- Görner, C., Masterarbeit: Semisynthesis of Tau Protein Variants, Technische Universität München, **2011.**
- Guan, D., Ramirez, M., Chen, Z., Split Intein Mediated Ultra-Rapid Purification of
- Tagless Protein (SIRP), *Biotechnol and Bioeng*, **2013**, 110 (9): 2471-2481.
- Guo, Z., Bishop, L., Chemical Synthesis of GPIs and GPI-Anchored Glycopeptides, *Eur. J. Org. Chem.*, **2004**, 17: 3585-3596.

- Guruprasad, K., Reddy, B.V, Pandit, M.W, Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence, *Protein Eng*, **1990**, 4 (2): 155–61.
- Harris, D.A., Trafficking, turnover and membrane topology of PrP, *Brit Med Bull*, **2003**, 66: 71–85.
- Harrison, C.F., Lawson, V.A., Coleman, B.M., Kim, Y.-S., Masters, C.L., Cappai, R., Barnham, K.J., Hill, A.F., Conservation of a Glycine-rich Region in the Prion Protein Is Required for Uptake of Prion Infectivity, *J Biol Chem*, **2010**, 285 (26): 20213–20223.
- Hornshaw, M.P., McDermott, J.R., Candy, J.M., Lakey, J.H., Copper binding to the N-terminal tandem repeat region of mammalian and avian prion protein: structural studies using synthetic peptides, *Biochem Bioph Res Co*, **1995**, 214 (3): 993-999.
- Iwai, H., Züger, S., Jin, J., Tam, P.-H., Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme, *FEBS*, **2006**, 580: 1853-1858.
- Jarrett, J.T., Lansbury, P.T. Jr., Seeding 'one-dimensional crystallization' of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie?, *Cell*, **1993**, 73: 1055–1058.
- Kane, P.M., Yamashiro, C.T., Wolczyk, D.F., Neff, N., Goebl, M., Stevens, T.H., Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H(+)-adenosine triphosphatase, *Science*, **1990**, 250 (4981): 651-657.
- Kawasaki, M., Satow, Y., Ohya, Y., Anraku, Y., Protein splicing in the yeast Vma1 protozyme: evidence for an intramolecular reaction, FEBS, **1997**, 412 (3): 518-520.
- Kent, S.B.H., Chemical synthesis of peptides and proteins, *Ann. Rev. Biochem*., **1988**, 57: 957-989.
- Kim, J.-I., Cali, I., Surewicz, K., Kong, Q., Raymond, G.J., Atarashi, R., Race, B., Qing, L., Gambetti, P., Caughey, B., Surewicz, W.K., Mammalian Prions Generated from Bacterially Expressed Prion Protein in the Absence of Any Mammalian Cofactors, *J Biol Chem*, **2010**, 285 (19): 14083–14087.
- Klingeborn, M., Race, B., Meade-White, K.D., Rosenke, R., Striebel, J.F., Chesebro, B., Crucial Role for Prion Protein Membrane Anchoring in the Neuroinvasion and Neural Spread of Prion Infection, *J Virol*, **2011**, 85 (4): 1484-1494.
- Kraus, A., Groveman, B.R., Caughey, B., Prions and the Potential Transmissibility of Protein Misfolding Diseases, *Annu. Rev. Microbiol*., **2013**, 67: 543–64.

- Laemmli, UK, Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4, *Nature*, **1970**, 227 (5259) 680-685.
- Lee G., Cowan N., Kirschner M., The primary structure and heterogeneity of tau protein from mouse brain, *Science*, **1988**, 239: 285-288.
- Legname, G., Baskakov, I.V., Nguyen, H.O., Riesner, D., Cohen, F.E., DeArmond, S.J., Prusiner, S.B., Synthetic mammalian prions, *Science*, **2004**, 305: 673-676.
- Liu, Z., Frutos, S., Bick, M.J., Vila-Perelló, M., Debelouchina, G.T., Darst, S.A., Muir, T.W., Structure of the branched intermediate in protein splicing, *PNAS*, **2014**, 111 (23): 8422-8427.
- Low, M.G., Finean, J.B., Release of Alkaline Phosphatase from Membranes by a Phosphatidylinositol-Specific Phospholipase C, *Biochem. J.*, **1977**, 167: 281-284.
- Low, M.G., Saltiel, A.R., Structural and functional Roles of Glycosyl-phosphatidylinositol in Membranes, *Science*, **1988**, 239: 268–275.
- Low, M.G., The Glycosyl-phosphatidylinositol Anchor of Membrane Proteins, *Biochim. Biophys. Acta*, **1989**, 988: 427–454.
- Low, M., GPI-anchored Biomolecules: an Overview, in *GPI-anchored Membrane Proteins and Carbohydrates* (Hoessli, DC., and Ilangumaran, S., Eds.) Landes Company, Austin (TX), **1999**, pp 1-14.
- Ma, J., The Role of Cofactors in Prion Propagation and Infectivity, *PLoS Pathog*, **2012**, 8 (8): e1002589.
- Maeda, Y., Tashima, Y., Houjou, T., Fujita, M., Yoko-o, T., Jigami, Y., Taguchi, R., Kinoshita, T., Fatty Acid Remodeling of GPI-anchored Proteins Is Required for Their Raft Association, *Mol Biol Cell*, **2007**, 18: 1497-1506.
- Masterson, W.J., Raper, J., Doering, T.L., Hart, G.W., Englund, P.T., Fatty acid remodeling: a novel reaction sequence in the biosynthesis of trypanosome glycosyl phosphatidylinositol membrane anchors, *Cell*, **1990**, 62: 73-80.
- McNally, K.L., Ward, A.E., Priola, S.A., Cells Expressing Anchorless Prion Protein Are Resistant to Scrapie Infection, *J Virol*, **2009**, 83 (9): 4469-4475.
- Meaden, P., Hill, K., Wagner, J., Slipetz, D., Sommer, S.S., Busseyl, H., The Yeast KRE5 Gene Encodes a Probable Endoplasmic Reticulum Protein Required for (1 >6)-13-D-Glucan Synthesis and Normal Cell Growth, *Mol Cell Biol*, **1990**, 10 (6): 3013-3019.

- Meyer, R.K., McKinley, M.P., Bowman, K.A., Braunfeld, M.B., Barry, R.A., Prusiner, S.B., Separation and properties of cellular and scrapie prion proteins, *Proc. Natl. Acad. Sci. USA*, **1986**, 83: 2310-2314.
- Merrifield R.B., Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide, *J. Am. Chem. Soc.*, **1963**, 85(14): 2149–2154.
- Mills, K.V., Johnson, M.A., Perler, F.B., Protein Splicing: How Inteins Escape from Precursor Proteins, *J Biochem*, **2014**, 289 (21): 14498-14505.
- Mootz, H.D., Muir, T.W., Protein Splicing Triggered by a Small Molecule, *J. Am. Chem. Soc*., 2**002**, 124: 9044-9045.
- Mootz, H.D., Blum, E.S., Tyszkiewicz, A.B., Muir, T.W., Conditional Protein Splicing: A New Tool to Control Protein Structure and Function in Vitro and in Vivo, *J. Am. Chem. Soc*., **2003**, 125: 10561-10569.
- Muir T.W., Sondhi D., Cole P.A., Expressed protein ligation: a general method for protein engineering, *Proc. Natl Acad. Sci. USA*, **1998**, 95: 6705–6710.
- Mujika, J.I., Lopez, X., Mulholland, A.J., Modeling Protein Splicing: Reaction Pathway for C-Terminal Splice and Intein Scission, *J Phys Chem B*, **2009**, 113: 5607-5616.
- Nosjean, O., Briolay, A., Roux, B., Mammalian GPI proteins: sorting, membrane residence and functions, *Biochim. Biophys. Acta*, **1997**, 1331, (2): 153-186.
- Oeemig, J.S., Aranko, A.S., Djupsjöbacka, J., Heinämäki, K., Iwaï, H., Solution structure of DnaE intein from Nostoc punctiforme: Structural basis for the design of a new split intein suitable for site-specific chemical modification, *FEBS Letters*, **2009**, 583: 1451-1456.
- Olschewski, D., Seidel, R., Miesbauer, M., Rambold, A.S., Oesterhelt, D., Winklhofer, K.F., Tatzelt, J., Engelhard, M., Becker, C.F.W., Semisynthetic Murine Prion Protein Equipped with a GPI Anchor Mimic Incorporates into Cellular Membranes, *Chem Biol*, **2007**, 14: 994-1006.
- Orgel, L., Prion replication and secondary nucleation, *Chem Biol*, **1996**, 3: 413-414.
- Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., Gray, T., How to measure and predict the molar absorption coefficient of a protein, *Protein Science*, **1995**, 4: 2411-2423.
- Pan, K.-M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R.J., Cohenu, F.E., Prusiner, S.B., Conversion of α-helices into β-

sheets features in the formation of the scrapie prion proteins, *Proc. Natl. Acad. Sci. USA*, **1993**, 90: 10962-10966.

- Paulick, MG., Bertozzi, CR., The Glycosylphosphatidylinositol Anchor: A Complex Membrane-Anchoring Structure for Proteins, *Biochemistry*, **2008**, 47: 6991-7000.
- Paulus, H., Protein splicing and related forms of protein autoprocessing, *Annu. Rev. Biochem*., **2000**, 69: 447-496.
- Pauly, P.C., Harris, D.A., Copper Stimulates Endocytosis of the Prion Protein, *J Biol Chem*, **1998**, 273 (50): 33107-33110.
- Pietrokovski, S., Conserved sequence features of inteins (protein introns) and their use in identifying new inteins and related proteins, *Protein Sci*, **1994**, 3: 2340-2350.
- Plückthun, A., Mono- and bivalent antibody fragments produced in *Escherichia coli*: engineering, folding and antigen binding, *Immunol Rev*, **1992**, 130:151–88.
- Prusiner, S.B., Novel proteinaceous infectious particles cause scrapie, *Science*, **1982**, 216(4542): 136-144.
- Prusiner, S.B., McKinley, M.P., Bowman, K.A., Bolton, D.C., Bendheim, P.E., Groth, D.F., Glenner, G.G., Scrapie prions aggregate to form amyloid-like birefringent rods, *Cell*, **1983**, 35 (2): 349-358.
- Prusiner, S.B., Molecular biology of prion diseases, *Science*, **1991**, 252 (5012):1515-22.
- Prusiner, S.B., Prions, *Proc. Natl. Acad. Sci. USA*, **1998**, 95: 13363-13383.
- Puig, B., Altmeppen, H., Glatzel, M., The GPI-anchoring of PrP: Implications in sorting and pathogenesis, *Prion*, **2014**, 8 (1): 8-11.
- Reverberi, R., Reverberi, L., Factors affecting the antigen-antibody reaction, *Blood Transfus*, **2007**, 5(4): 227-240.
- Safar, J., Wille, H., Itri, V., Groth, D., Serban, H.,Torchia, M., Cohen, F.E., Prusiner, S.B., Eight prion strains have PrPSc molecules with different conformations, *Nat Med*, **1998**, 4 (10): 1157-1165.
- Sahdev, S., Khattar, S.K., Singh Saini, K., Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies, *Mol Cell Biochem*, **2008**, 307: 249-264.

- Schumacher, M.C., Resenberger, U., Seidel, R.P., Becker, C.F.W., Winklhofer, K.F., Oesterhelt, D., Tatzelt, J., Engelhard, M., Synthesis of a GPI Anchor Module Suitable for Protein Post-Translational Modification, *Biopolymers Pept Sci*, **2010**, 94, (4): 457–464.
- Seddon, A.M., Curnow, P., Booth, P.J., Membrane proteins, lipids and detergents: not just a soap opera, *Biochim Biophys Acta*, **2004**, 1666: 105-117.
- Shah, N.H., Dann, G.P., Vila-Perelló, M., Liu, Z., Muir, T.W., Ultrafast Protein Splicing is Common among Cyanobacterial Split Inteins: Implications for Protein Engineering, *J. Am. Chem. Soc.*, **2012**, 134: 11338−11341.
- Shah, N.H., Eryilmaz, E., Cowburn, D., Muir, T.W., Extein Residues Play an Intimate Role in the Rate-Limiting Step of Protein Trans-Splicing, *J. Am. Chem. Soc.*, **2013**, 135: 5839-5847.
- Shah, N.H., Muir, T.W., Inteins: Nature's Gift to Protein Chemists, *Chem Sci*., **2014**, 5 (1): 446–461.
- Shao, Y., Xu, M.-Q., Paulus, H., Protein Splicing: Characterization of the Aminosuccinimide Residue at the Carboxyl Terminus of the Excised Intervening Sequence, *BioChem*, **1995**, 34: 10844-10850.
- Shao, Y., Paulus, H., Protein splicing: estimation of the rate of 0-N and S-N acyl rearrangements, the last step of the splicing process, *J Peptide Res*, **1997**, 50: 193-198.
- Shih C.K., Wagner R., Feinstein S., Kanik-Ennulat C., Neff N., A dominant trifluoperazine resistance gene from Saccharomyces cerevisiae has homology with F0F1 ATP synthase and confers calcium-sensitive growth, *Mol Cell Biol*., **1988**, 8(8): 3094–3103.
- Skretas, G., Wood, D.W., Regulation of protein activity with small-molecule-controlled inteins, *Protein Science*, **2005**, 14: 523-532.
- Smirnov, MN., Smirnov, VN., Budowsky, EI., Inge-Vechtomov, SG., Serebrjakov, NG., Red pigment of adenine-deficient yeast Saccharomyces cerevisiae, *Biochem Bioph Res Co*, **1967**, 27 (3): 299-304.
- Soto, C., Castilla, J., The controversial protein-only hypothesis of prion propagation, *Nature Medicine*, **2004**, 10: 63-67.

- Southworth, M.W., Amaya, K., Evans, T.C., Xu, M.-Q., Perler, F.B. Purification of Proteins Fused to Either the Amino or Carboxy Terminus of the *Mycobacterium xenopi* Gyrase A Intein, *BioTechniques*, **1999**, 27:110-120.
- Surewicz, W.K., Jones, E.M., Apetri, A.C., The Emerging Principles of Mammalian Prion Propagation and Transmissibility Barriers: Insight from Studies in Vitro, *Acc. Chem. Res*., **2006**, 39: 654-662.
- Taylor D.R., M. Hooper N.M., GPI-Anchored Proteins in Health and Disease, in: C.J. Vidal, *Posttranslational Modifications in Health and Disease*, Protein Reviews 13, **2011**.
- Towbin H., Staehelin T., Gordon J., Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications, *PNAS*, **1979**, 76 (9): 4350-4354.
- Tropea, J.E., Cherry, S., Waugh, D.S., Expression and Purification of Soluble His_6-tagged TEV Protease, *Methods Mol Biol*, **2009**, 498:297-307
- Udenfriend, S., Kodukula, K., How glycosylphosphatidylinositol anchored membrane proteins are made, *Annu. Rev. Biochem*., **1995**, 64: 563-591.
- Vila-Perelló, M., Hori, Y., Rib, M., Muir, T.W., Activation of Protein Splicing by Protease- or Light-Triggered O to N Acyl Migration, *Angew. Chem. Int. Ed.*, **2008**, 47: 7764 –7767.
- Vila-Perelló, M., Muir, T.W., Biological Applications of Protein Splicing, *Cell*, **2010**, 143 (2): 191-200.
- Vila-Perelló, M., Liu, Z., Shah, N.H., Willis, J.A., Idoyaga, J., Muir, T.W., Streamlined Expressed Protein Ligation Using Split Inteins, *J Am Chem Soc*, **2012**, 135: 286-292.
- Villain, M., Vizzavona, J., Rose, K., Covalent capture: a new tool for the purification of synthetic and recombinant polypeptides, *Chem Biol*, **2001,** 8: 673-679.
- Volkmann, G., Mootz, H.D., Recent progress in intein research: from mechanism to directed evolution and applications, *Cell Mol Life Sci*, **2013**, 70: 1185-1206.
- Wang, F., Wang, X., Yuan, C.-G., 2, Ma, J., Generating a Prion with Bacterially Expressed Recombinant Prion Protein, *Science*, **2010**, 327 (5969): 1132-1135.
- Wang, L., Towards revealing the structure of bacterial inclusion bodies, *Prion*, **2009**, 3(3): 139-145.

- Waugh, D.S., Reprint of: Making the most of affinity tags, *TRENDS Biotechnol*, **2005**, 23 (6): 316-320.
- Weissmann, C., Enari, M., Klöhn, P.-C., Rossi, D., Flechsiga, E., Transmission of prions, *J Infect Dis*, **2002**, 186 (2): S157-165.
- Wood, D.W., Camarero, J.A., Intein Applications: From Protein Purification and Labeling to Metabolic Control Methods, *J Biol Chem*, **2014**, 289 (21): 14512–14519.
- Xu, M.-Q., Southworth, M.W., Mersha, F.B., Hornstra, L.J., Perler, F.B., In vitro protein splicing of purified precursor and the identification of a branched intermediate, *Cell*, **1993**, 75 (7): 1371-1377.
- Xu, M.-Q., Comb, D.G., Paulus, H., Noren, C.J., Shao, Y., Perler, F.B., Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation, *EMBO*, **1994**, 13 (23): 5517-5522.
- Xu, M.-Q., Perler, F.B., The mechanism of protein splicing and its modulation by mutation, *EMBO J.*, **1996**, 15 (19): 5146-5153.
- Xu, R., Ayers, B., Cowburn, D., Muir, T.W., Chemical ligation of folded recombinant proteins: Segmental isotopic labeling of domains for NMR studies, *Proc. Natl. Acad. Sci. USA*, **1999**, 96: 388-393.
- Yu, S, Guo, Z., Johnson, C., Gu, G., Wu, Q., Recent progress in synthetic and biological studies of GPI anchors and GPI-anchored proteins, *Curr. Opin. Chem. Biol.*, **2013**, 17: 1006–1013.
- Zettler, J., Schütz, V., Mootz, H.D., The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction, *FEBS*, **2009**, 583: 909-914.

6 Supplementary information

Tau1-290-NpuDnaE-H_6 cleavage

Calculated masses:

Tau1-290-NpuDnaE-H6		**46913.1 Da**
	-Met	**46781.9 Da**
Tau1-290-OH		**29986.0 Da**
	-Met	**29854.8 Da**
Tau1-290-MES		**30110.0 Da**
	-Met	**29978.8 Da**
NpuDnaE-H6		**16945.1 Da**

Tau1-290-NpuDnaE-H_6 before intein cleavage

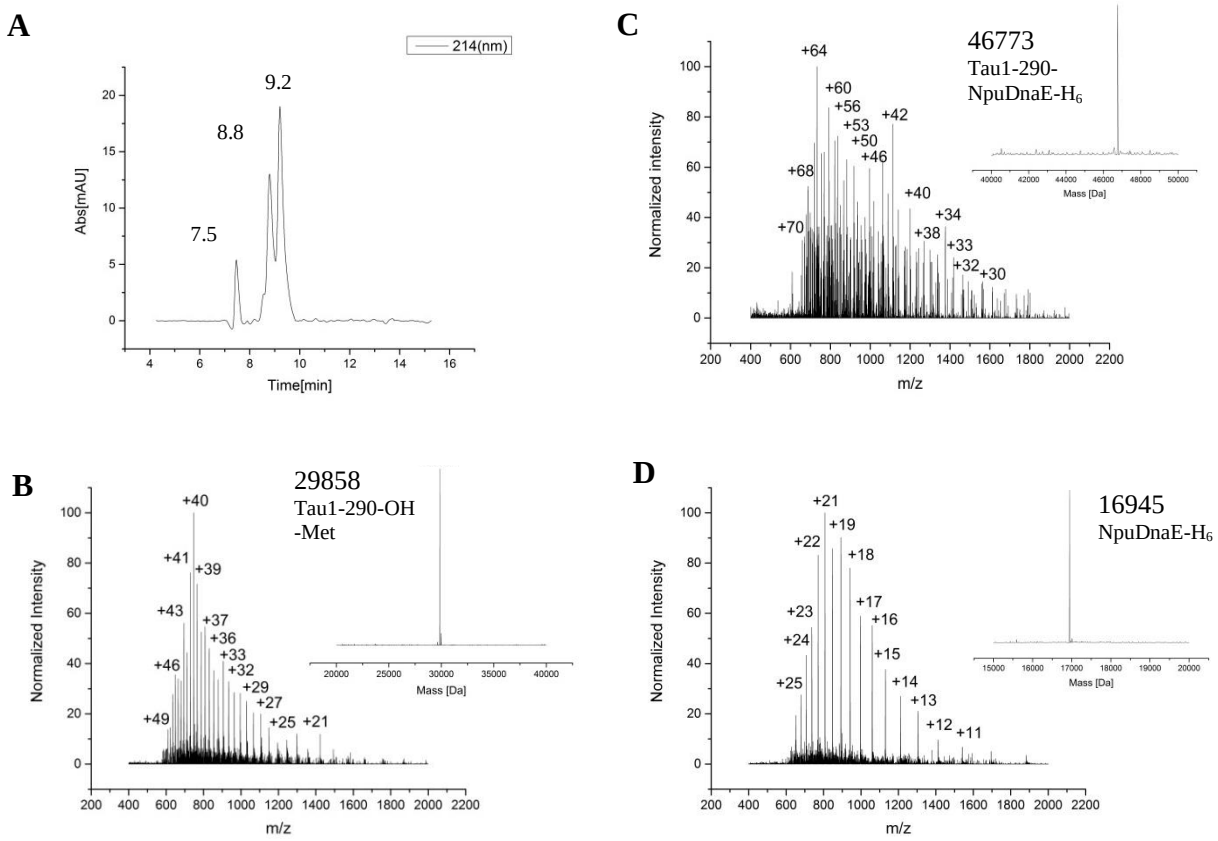


Figure 41: LC-MS Tau1-290-NpuDnaE-H_6 before intein cleavage. A) HPLC chromatogramm. B) Spectrum of peak at 7.5 min = Tau1-290-OH –Met. Calculated mass: 29854.8 Da, observed mass: 29858 Da. C) Spectrum of peak at 9.2 min = Tau1-290-NpuDnaE-H_6 –Met. Calculated mass: 46781.9 Da, observed mass: 46773 Da. D) Spectrum of peak at 9.2 min: peak at 9.5 min = NpuDnaE-H_6. Calculated mass: 16945.1 Da, observed mass: 16945 Da.

Tau-1-290-NpuDnaE-H6 after intein cleavage

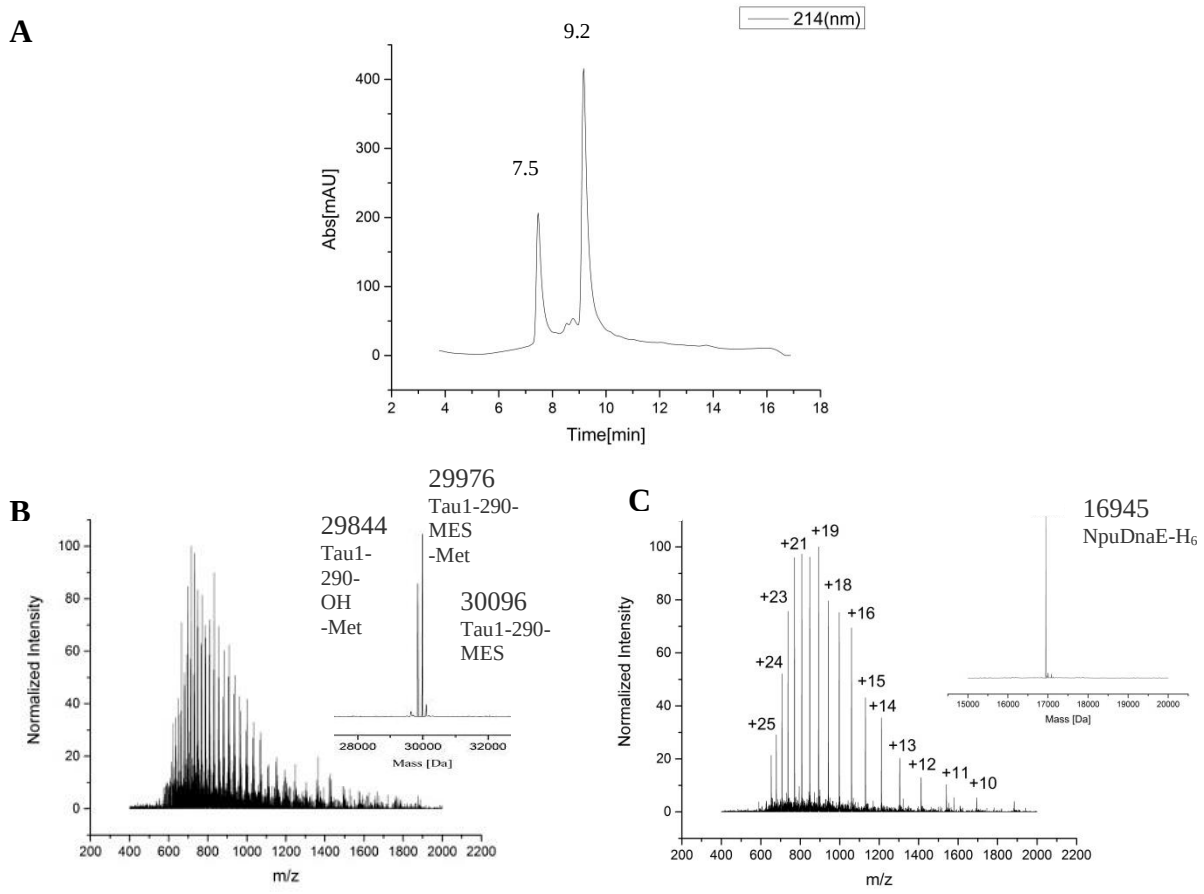


Figure 42: LC-MS Tau1-290-NpuDnaE-H_6 after intein cleavage. A) HPLC chromatogram. B) Spectrum of peak at 7.5 min = 1) Tau1-290-OH -Met, 2) Tau1-290-MESNa –Met, 3) Tau1-290-MESNa. Calculated masses: 1) 29854.8 Da, 2) 30110 Da, 3) 29978.8 Da. Observed masses: 1) 29844 Da, 2) 29976 Da, 3) 30096 Da. C) Spectrum of peak at 9.2 min = NpuDnaE-H_6. Calculated mass: 16945.1 Da, observed mass: 16945 Da.

Tau1-290-Mxe-CBD cleavage

Calculated masses:

Tau1-290-Mxe-CBD		**57826.3 Da**
	-Met	**57695.1 Da**
Tau1-290-OH		**29986.0 Da**
	-Met	**29854.8 Da**
Tau1-290-MES		**30110.0 Da**
	-Met	**29978.8 Da**
Mxe-CBD		**27858.4 Da**

Tau1-290-Mxe-CBD after intein cleavage

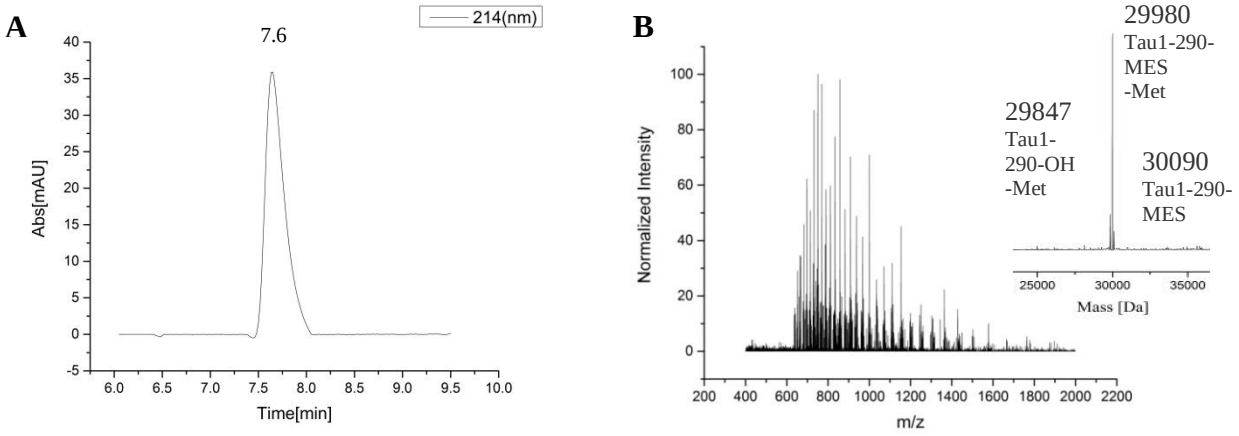


Figure 43: LC-MS Tau1-290-Mxe-CBD after intein cleavage. A) HPLC chromatogramm. B) Spectrum peak at 7.6 min = 1) Tau1-290-OH -Met, 2) Tau1-290-MESNa –Met, 3) Tau1-290-MESNa. Calculated masses: 1) 29854.8 Da, 2) 30110 Da, 3) 29978.8 Da. Observed masses: 1) 29847 Da, 2) 29980 Da, 3) 30090 Da.

PrP90-231-NpuDnaE-H_6 cleavage

Calculated masses:

PrP-90-231-NpuDnaE-H6		**33394.5 Da**
	-Met	**33263.3 Da**
PrP90-231-OH		**16467.4 Da**
	-Met	**16336.2 Da**
PrP90-231-MES		**16591.4 Da**
	-Met	**16460.2 Da**
NpuDnaE-H6		**16945.1 Da**

PrP90-231-NpuDnaE-H_6 before intein cleavage

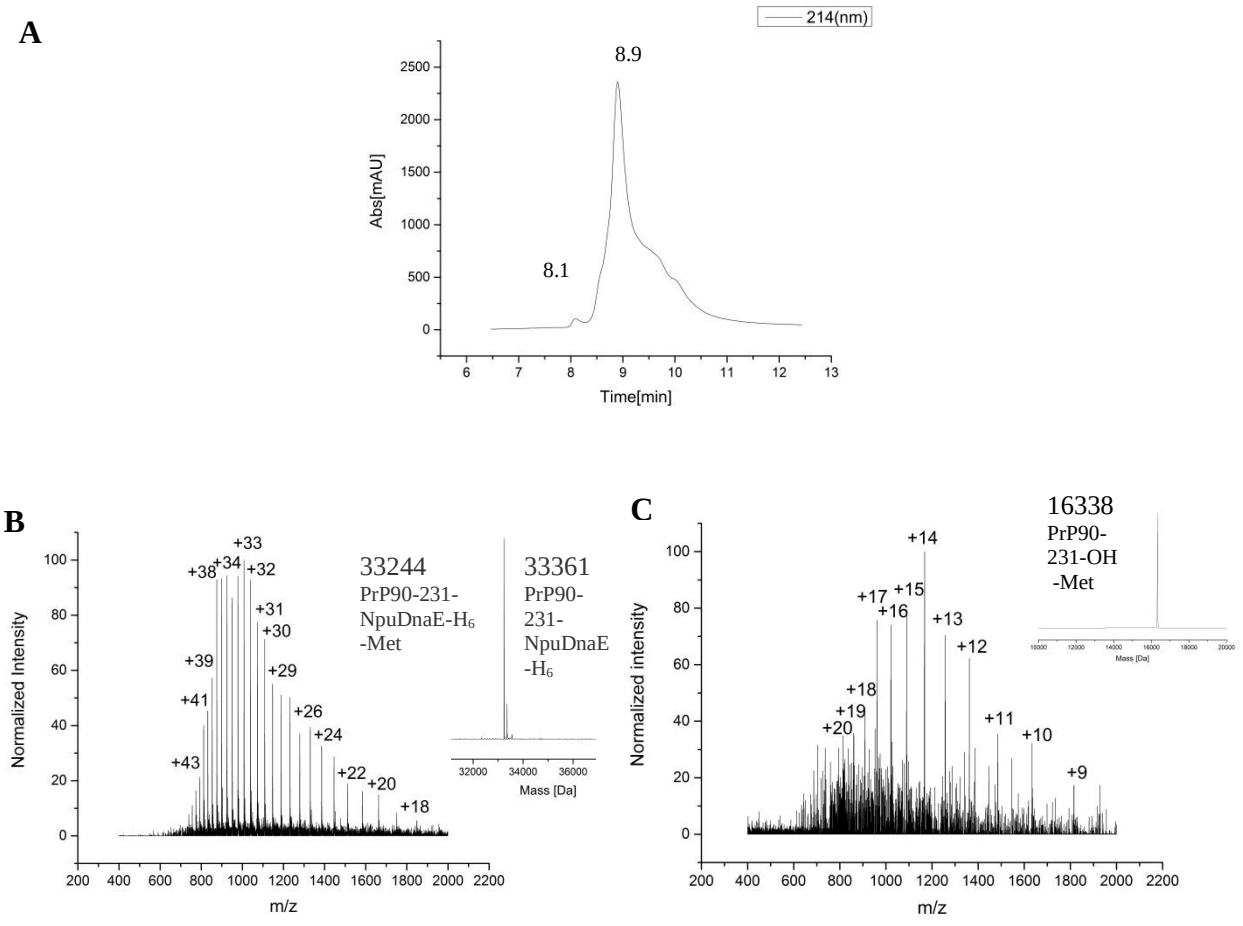


Figure 44: LC-MS PrP90-231-NpuDnaE-H_6 before intein cleavage. A) HPLC chromatogramm. B) Spectrum of peak at 8.9 min = 1) PrP90-231-NpuDnaE-H_6 -Met, 2) PrP90-231-NpuDnaE-H_6. Calculated masses: 1) 33263.3 Da, 2) 33394.5 Da. Observed masses: 1) 33244 Da, 2) 33361 Da. C) Spectrum of peak at 8.1 min = PrP90-231-OH -Met. Calculated mass:16336.2 Da, observed mass: 16338 Da.

PrP90-231-NpuDnaE-H_6 after intein cleavage

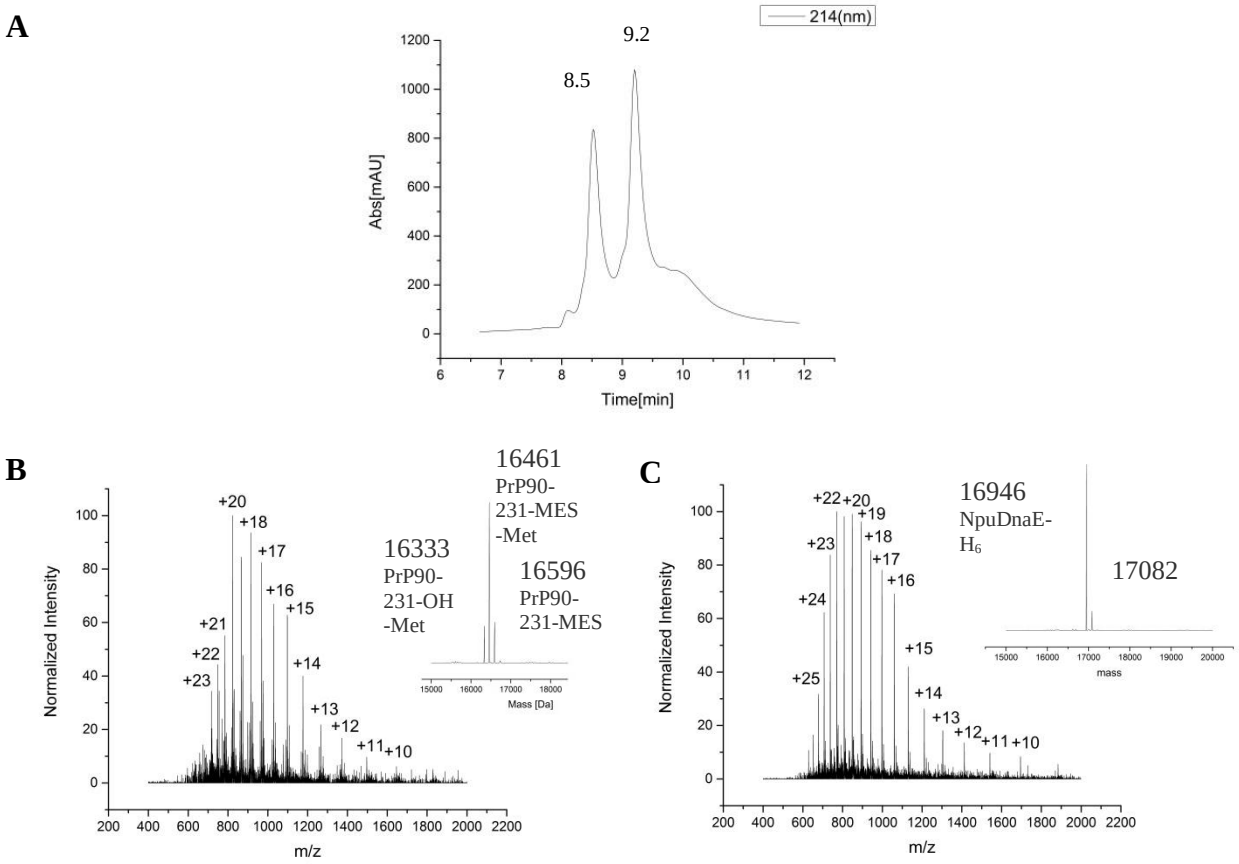


Figure 45: LC-MS PrP90-231-NpuDnaE-H_6 after intein cleavage. A) HPLC chromatogram. B) Spectrum of peak at 8.5 min = 1) PrP90-231-OH -Met, 2) PrP90-231-MESNa –Met, 3) PrP90-231-MESNa. Calculated masses: 1) 16336.2 Da, 2) 16460.2 Da, 3) 16591.4 Da. Observed masses: 1) 16333 Da, 2) 16461 Da, 3) 16596 Da. C) Spectrum of peak at 9.2 min = 1) NpuDnaE-H_6, 2) Probably an NpuDnaE-H_6-MESNa derivative. Calculated mass: 1)16945.1 Da, 2) 17069.1 Da. Observed mass: 1) 16946 Da, 2) 17082 Da.

PrP90-231-Mxe-H_6-CBD cleavage

Calculated masses:

PrP-90-231-Mxe-H6-CBD		**45525.0 Da**
	-Met	**45400.9 Da**
PrP90-231-OH		**16467.4 Da**
	-Met	**16336.2 Da**
PrP90-231-MES		**16591.4 Da**
	-Met	**16460.2 Da**
Mxe-H6-CBD		**29075.6 Da**

PrP90-231-Mxe-H_6-CBD before intein cleavage

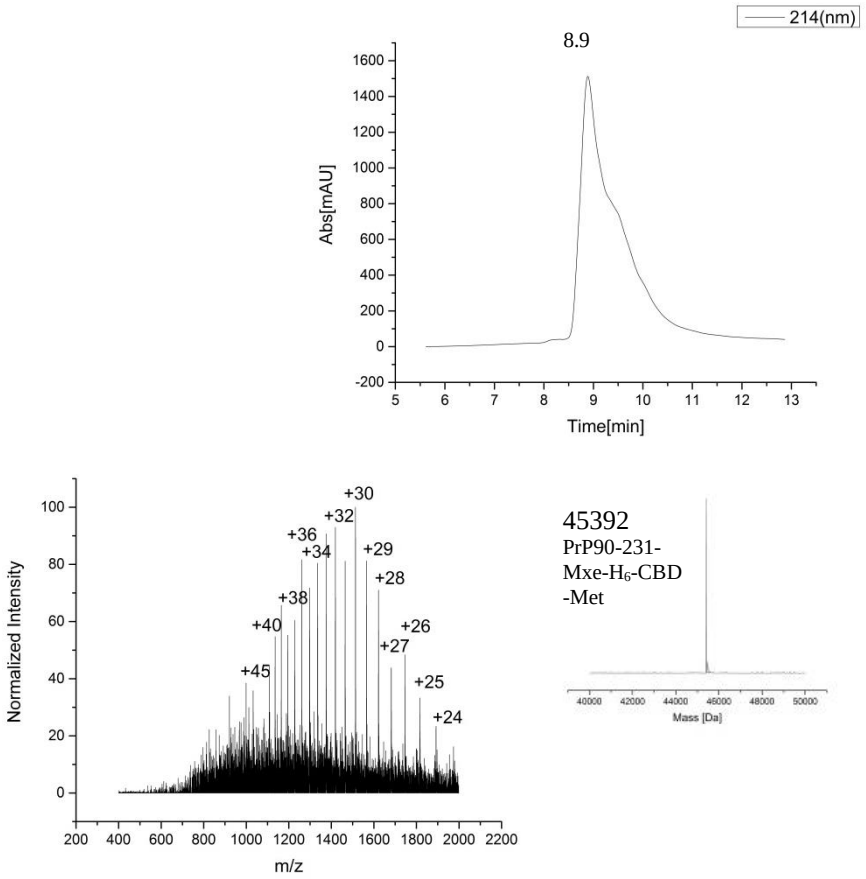


Figure 46: LC-MS PrP90-231-Mxe-H_6-CBD before intein cleavage. A) HPLC chromatogram B) Spectrum of peak at 8.9 min = PrP90-231-Mxe-H6-CBD - Met. Calculated mass: 45401 Da, observed mass: 45392 Da.

PrP90-231-Mxe-H_6-CBD after intein cleavage

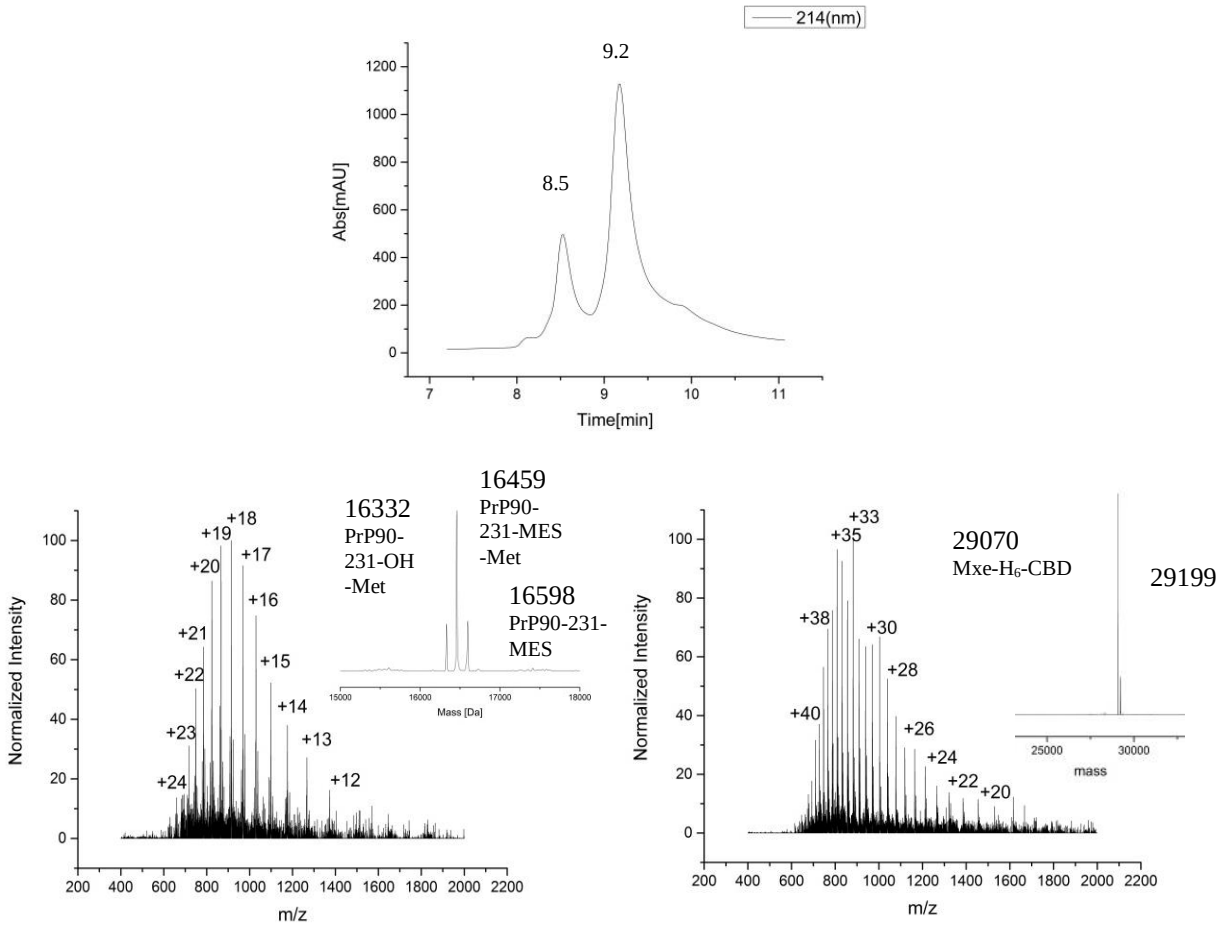


Figure 47: LC-MS PrP90-231-Mxe-H_6-CBD after intein cleavage. A) HPLC chromatogram. B) Spectrum of peak at 8.5 min = 1) PrP90-231-OH -Met, 2) PrP90-231-MESNa –Met, 3) PrP90-231-MESNa. Calculated masses: 1) 16336.2 Da, 2) 16460.2 Da, 3) 16591.4 Da. Observed masses: 1) 16332 Da, 2) 16459 Da, 3) 16598 Da. C) Spectrum of peak at 9.2 min = 1) Mxe-H_6-CBD. 2) Probably Mxe-H6-CBD-Mesna derivative. Calculated mass: 1) 29075.5 Da. 2) 29199.5. Observed mass: 1) 29070 Da, 2) 29199 Da.

WTYSLRYGIPEFKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGTWGQPHGGG
WGQPHGGSWGQPHGGSWGQPHGGGWGQGGGTHNQWNKPSKPKTNMKHMAGA
AAAGAVVGGLGGYMLGSAMSRPMIHFGNDWEDRYYRENMYRYPNQVYYRPVDQ
YSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVVEQMCVTQYQKESQ
AYYDGRES**LQ**HHHHHH**KASGGSG**ENLYFQCGDYKDDDDKSSSSSSSASSSSSSKKN

● = Kre5p28-39, ● = PrP22-230, ● = His-tag, ● = TEV recognition site, ● = FLAG-tag,
● = Gas1p511-528, ● = linker sequences

Figure 48: Amino acid sequence of expressed GPI-anchored construct

MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYN
IQKESTLHLVLRLRGG**CLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPV**
AQWHDRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLP
NIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASAAFNHHHHHH

Figure 49: Amino acid sequence of Ub-NpuDnaE-H_6 (Shah *et al.*, 2012)

Drop-out mix	
Substance	**Amount [g]**
Adenin	0.5
Alanin	2
Argenin	2
Asparagin	2
Aspartat	2
Cystein	2
Glutamin	2
Glutamat	2
Glycin	2
Histidin	2
Inositol	2
Isoleucin	2
Lysin	2
Methionin	2
p-Aminobenzoesäure	0.2
Phenylalanin	2
Prolin	2
Serin	2
Threonin	2
Tryptophan	2
Tyrosin	2
Uracil	2
Valin	2

Figure 50: Composition Drop-out mix

	H_2O [ml]	Acrylamide [ml]	Buffer [ml]	10% SDS [µl]	10% APS [µl]	TEMED [µl]
	Separating Gel 15% Laemmi					
2 gels	3,6	7,5	3,9	157	157	4,7
4 gels	4,8	10	5,2	209	209	9,4
9 gels	19,2	40	20,8	837	837	38
	Separating Gel 18% Laemmi					
2 gels	2,1	9	3,9	157	157	4,7
4 gels	2,8	12	5,5	209	209	9,4
9 gels	11,2	48	20,8	837	837	38
	Stacking Gel					
2 gels	3,6	0,9	0,7	53	53	5,3
4 gels	6,3	1,6	1,3	93	93	9,4
9 gels	29	7,2	5,6	426	426	42

Figure 51: SDS gel composition

Acknowledgement

At first, I would like to thank Prof. Christian Becker for the opportunity to work in his group. I really appreciate the great supervision and the continuous support while working in the lab as well as afterwards when writing the thesis.

I would also like to thank the whole working group: thanks to you all for the friendly, instructive, supporting and productive working atmosphere! Special thanks are going to Can Araman for his constructive support and Manuel Brehs for taking the trouble to proofread my thesis.

I am grateful to have friends like Maria Bassermann, Solveig Harm, Rocio Burchard and Rahel Köpf who helped me by reviewing my writing and provided moral support since we were all in the same situation.

I am thankful for my mother who supported and motivated me during all the highs and lows you experience while writing a thesis and letting me blow up her phone bill.

Curriculum vitae

Name: Alanca Schmid

Date of Birth: 30.06.1988

Place of Birth: Kaufbeuren

Nationality: German

10.2015	Masterthesis: „Insertion of Modifications into Proteins by Semisynthesis: Generating and Testing of new *NpuDnaE*-Protein Constructs and Isolation and Purification of a recombinantly expressed GPI-anchored Peptide for Expressed Protein Ligation“ Supervisor: Professor Christian Friedrich Wilhelm Becker, Institut für Biologische Chemie, Universität Wien
03.2013-10.2015	Master studies, Master Chemie, Universität Wien
02.2013	Bachelorthesis: “Semisynthese Post-translational Modifizierter Tau40 Varianten“ Supervisor: Dr. Manuel Brehs, Institut für Biologische Chemie, Universität Wien
10.2009-02.2013	Bachelor studies, Bachelor Chemie, Universität Wien
06.2008	Abitur, Waldorfschule Landsberg/Lech
2004-2008	Waldorfschule Landsberg/Lech
1994-2004	Waldorfschule Kempten/Allgäu