



TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt

Lehrstuhl für Chemie der Biopolymere

BLaTM: A novel method to measure transmembrane domain interactions

Christoph Schanzenbach

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. D. Frishman

Prüfer der Dissertation:

1. Univ.-Prof. Dr. D. Langosch
2. Univ.-Prof. Dr. W. Liebl

Die Dissertation wurde am 21.11.2016 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 29.01.2017 angenommen.

Für Miriam

Zusammenfassung

Membranproteine sind der Hauptangriffspunkt von Medikamenten. 20% bis 30% aller offenen Leserahmen auf dem Genom kodieren für diese Proteine, welche über eine oder mehrere hydrophobe Transmembransegmente (TMS_e) in zelluläre Membranen eingebettet sind. Membranproteine können über diese Segmente sehr spezifisch oligomerisieren. Diese Interaktionen sind häufig essentiell für die Bildung von tertiären Proteinkomplexen und damit die Funktion dieser Proteine. Obwohl diese Proteine von großer biologischer Relevanz sind, sind sie im Gegensatz zu löslichen Proteinen schlecht erforscht. Um die Funktion von TMS_e bei den Wechselwirkungen zwischen (unterschiedlichen) Untereinheiten eines Komplexes zu verstehen, benötigt man einen sensitiven Assay, welcher sowohl homo- als auch heterotypische TMS-TMS Interaktionen messen kann. Die Entwicklung eines derartigen, neuen Assays war das Ziel dieser Forschungsarbeit. Der neue Assay sollte bewährte Assays wie ToxR, GALLEX oder BACTH übertreffen und dabei dennoch weiterhin auf *E. coli* basieren. Um die komplexen Mechanismen, die anderen Assays zu Grunde liegen, wie etwa die Gentranskriptionsaktivierung zu umgehen, beruht der neue Assay BLaTM auf der direkten Aktivitätsbestimmung eines rekonstituierten „Split-Proteins“. Als Split-Protein wurde hierbei das Antibiotikumresistenzprotein β -Lactamase TEM-1 ausgewählt, welches β -Lactame wie Ampicillin enzymatisch hydrolysieren kann. Die Fragmente des Split-Proteins werden auf DNA Ebene an zwei zu untersuchende TMS_e fusioniert. Der Anteil an rekonstituierten Enzym ist hierbei proportional zur Affinität der beiden TMS_e. Die Vorteile dieser Methode sind die gut quantifizierbare enzymatische Aktivität und der gleichzeitige Nachweis der Membraninsertion. Der BLaTM Assay wurde mit Hilfe der TMS_e von Glycophorin A (GpA), von Sulfhydryl Oxidase 2 (QSOX2) und des hochaffinen LS46 TMS optimiert und verifiziert. Die TMS-Affinität korreliert mit dem zuverlässigen und genau bestimmbaren LD₅₀ Wert, also der Ampicillinkonzentration welche für 50% aller *E. coli* letal ist. Es konnte hier gezeigt werden, dass mit diesem Assay mit hoher Reproduzierbarkeit sowohl homo- als auch heterotypische TMS-TMS Interaktionen vermessen werden können. Außerdem konnten die bereits bekannten Effekte bestimmter Mutationen, welche zuvor mit anderen Methoden wie ToxR, GALLEX oder NMR nachgewiesen worden waren, erfolgreich reproduziert werden. Zusätzlich kann der Assay an verschiedene TMS-

Affinitäten angepasst werden, wodurch sowohl die Identifizierung von hochaffinen TMSen, als auch der Nachweis von schwächeren Dimeren möglich ist. Zusammenfassend ist BLaTM eine neue Methode für die Erforschung von intramembranen Protein-Protein Wechselwirkungen durch die Bestimmung von homo- und heterotypischen TMS-TMS Affinitäten.

Abstract

The main drug targets in humans are membrane proteins. 20% to 30% of all open reading frames are coding for these proteins which are embedded into cellular membranes by one or several hydrophobic transmembrane domains (TMDs). It is known that these membrane protein regions can form oligomers via distinct amino acid motifs which are very sensitive to mutations. These interactions may be essential for quaternary interactions during the assembly of proteins complexes and thus support protein function. Despite their outstanding biological relevance, membrane proteins are only poorly investigated compared to their soluble counterparts. To uncover the role of TMDs in protein assembly a new assay is needed which enables exact measurements of heterotypic TMD-TMD interactions. This method should surpass previously established assays, such as ToxR, GALLEX or BACTH, but remain *E. coli*-based. To overcome the complexities of gene transcription activation used in those assays, the new so called BLaTM assay is based on the direct activity-measurement of a reconstituted split-protein. The antibiotic resistance protein β -lactamase TEM-1, which hydrolyses β -lactams such as ampicillin, was chosen because of the ease of its quantification and because it simultaneously works as membrane insertion control. By split protein fusion to TMDs of interest at the DNA-level, the enzyme activity is directly proportional to the TMD-TMD affinity. The BLaTM assay was optimized and verified using the well-investigated TMDs Glycophorin A (GpA), Sulphydryl oxidase 2 (QSOX2) and the artificial high dimerizing LS46 TMD. TMD affinities were defined by the ampicillin concentration lethal to 50% of expressing *E. coli* cells - the LD₅₀. It has been successfully demonstrated here that the BLaTM assay can be used with high reproducibility for the quantification of homo- and especially heterotypic TMD-TMD interactions. Disrupting effects of distinct mutations in the GpA and QSOX2 TMDs that were known from ToxR, GALLEX, BACTH or NMR measurements, could be reproduced with BLaTM. Additionally, the assay can be adjusted to different TMD affinities, allowing the identification of strong dimers as well as low affinity. Summarizing, the BLaTM is a novel method for the investigation of intramembraneous homo- and heterotypic protein-protein interactions which allows the determination of affinities of low and high affinity TMD dimers.

Acknowledgements

I want to use this possibility to thank all the people who supported me – scientifically and personally - during the last four years and enabled me to accomplish my dissertation. They all together helped me to enjoy the lab work and to overcome less successful periods.

Prof Dr. Dieter Langosch: for giving me the opportunity to write my dissertation at his chair. He enabled me more scientific freedom than I have ever expected and supported me in my project all the time.

Wolfgang Liebl and Prof. Dr. Dmitrij Frishman: for agreeing to examine my thesis.

Dr. Mark Teese, Ayşe Julius and Fabian Schmidt: for the many fruitful discussions, new ideas and trust in my work. They encouraged me to confide the results in an overwhelming manner. Especially Ayşe for her detailed review of my thesis and the many helpful comments and Mark for the support in programming and sharing his scientific expertise.

Barbara Rauscher and Doreen Tetzlaff: for their daily support in the lab, sequencing of a lot of DNA samples, preparing an uncountable number of LB-plates and many nice conservations.

Dr. Markus Gütlich: for sharing his in-depth knowledge in almost everything and answering all questions about technical and scientific problems.

All my present and past colleagues (Katja, Philipp, Philipp, Christoph, Alex, Martina, Elke, Walter, Ellen, Yao, Martin, Christian, Jan, Eliane, Ute, Sevnur, Oli, Steven): for providing the good working atmosphere, answering my questions and all the cake.

My students: for their input into my projects and the knowledge I gained from their questions.

My parents and my two sisters: for continuous supporting me in everything I do and being the best family I can imagine.

And in particular I want to thank my wife Corina who supported me far in excess of what would have been normal and trusted my work and my skills without any doubt. Thank you for your love and being part of my life – together we will succeed in everything!

Content

Zusammenfassung	I
Abstract	III
Acknowledgements	IV
Content	VI
1 Introduction.....	1
1.1 Cell membranes.....	1
1.2 Membrane proteins	1
1.2.1 Classification.....	2
1.2.2 Recognition and translocation.....	2
1.2.3 Insertion and folding	4
1.3 TMD-TMD Interaction	6
1.3.1 Interaction forces.....	6
1.3.2 In vivo TMD-TMD interaction assays.....	8
1.3.3 Orientation dependence.....	17
1.4 Protein complementation assays	19
1.5 β -Lactam antibiotics.....	20
1.5.1 Mechanism of antibiotic function	20
1.5.2 β -Lactamase	21
1.5.3 Split β -lactamase complementation assay	23
2 Motivation.....	25
3 Material and Methods	27
3.1 Preparation of chemical competent cells	27
3.2 Transformation.....	27
3.3 Polymerase chain reaction (PCR)	28

3.4	Molecular cloning.....	28
3.4.1	Restriction-based cloning.....	29
3.4.2	Cassette cloning.....	29
3.4.3	Q5 site directed mutagenesis.....	31
3.4.4	Quikchange mutagenesis.....	32
3.4.5	Transfer-PCR	32
3.5	Plasmid propagation	33
3.6	DNA purification.....	33
3.7	DNA quantification	34
3.8	Agarose gel electrophoresis.....	34
3.9	gDNA extraction from <i>Saccharomyces cerevisiae</i>	35
3.10	DNA Sequencing.....	35
3.11	GFP expression test	37
3.12	β -Lactamase activity test.....	38
3.13	SDS polyacrylamide gel electrophoresis.....	39
3.14	Western blot	40
3.15	BLaTM Assay	42
3.15.1	Choose of vectors	42
3.15.2	Vector construction	44
3.15.3	Assay protocol.....	50
3.15.4	Data analysis	52
4	Results	54
4.1	Initial experiments	54
4.1.1	Characterizing the activity of full length β -lactamase	54
4.1.2	Complementation of soluble split β -lactamase	55
4.1.3	Evaluation of ampicillin resistance data.....	57
4.2	BLaTM 1.1	58

4.2.1	Quantification method	59
4.2.2	Protein expression level	61
4.2.3	Influence of the linker between the β -lactamase fragment and TMD	63
4.2.4	Orientation dependence.....	65
4.2.5	Disruption index.....	66
4.2.6	Heterotypic interacting TMDs	67
4.3	BLaTM 1.2.....	71
4.3.1	Adjustment to different TMD-TMD affinities.....	71
4.3.2	Comparison to BLaTM 1.1	73
5	Discussion	75
5.1	Initial experiments.....	75
5.1.1	Optimization of protein constructs.....	75
5.1.2	Quantification of reconstituted β -lactamase	76
5.1.3	Data evaluation	77
5.2	Linker optimization.....	78
5.3	Heterotypic interaction.....	79
5.4	Orientation dependence.....	82
5.5	Characterization of β -lactamase fusion proteins	83
5.6	Tuning BLaTM for different TMD-TMD affinities	84
5.7	Comparison of BLaTM 1.1 and BLaTM 1.2	85
5.8	Comparison to established assays	86
5.9	Conclusion and outlook	87
6	Bibliography	90
7	List of Figures	106
8	List of Abbreviations	108
9	Appendix.....	113
9.1	TMD cassettes.....	113

9.2	PCR primers	115
9.3	Sequencing primers	116
9.4	Bacteria strains	116
9.5	Protein sequences	116
9.6	Plasmid sequences	117
9.6.1	N-BLa 1.1	117
9.6.2	C-BLa 1.1	119
9.6.3	N-BLa 1.2	121
9.6.4	C-BLa 1.2	123
	Publications	125
	Lebenslauf	126

1 Introduction

1.1 Cell membranes

Biological membranes are semipermeable bilayers, which separate different compartments in a cell and define its shape. They consist of amphipathic lipids and sometimes also contain sterols, like cholesterol in vertebrates [1]. Depending on the species and compartment, the lipid composition differs highly [1, 2]. The hydrophilic headgroups facing the aqueous surrounding are forming a double layer with a highly hydrophobic core. This core, filled with the acyl chains of the lipids, has a diameter of about 30 Å [3]. In contrast to small hydrophobic molecules, charged and large molecules cannot cross the hydrophobic core, so that it acts as an essential selective barrier. Together with the lipid headgroups on both sides, the membrane double layer has a diameter of about 60 Å. The length and the saturation state of the acyl chains is temperature dependent which ensures a liquid crystalline phase of the membrane [4]. Moreover, it is assumed that the different lengths of the acyl chains also fulfill other functions such as grouping proteins and lipids by hydrophobic mismatch [5, 6]. In addition to the differences in lipid composition, the lipids are not distributed equally between the inner and outer layer of the membrane. This asymmetry is maintained by specific floppases and flippases and is essential for cell survival. The collapse of the asymmetric distribution in mitochondria, for instance, is an apoptotic signal, and thus, lethal for cells [7]. However, biological membranes are more than simple barriers. They hold a place for incorporated membrane proteins that may for example be part of highly enzymatic activities such as the respiratory chain or act as an interface for intercellular communication.

1.2 Membrane proteins

Membrane proteins are not only coded by 20% to 30% of all open reading frames [8, 9], but also represent more than 60% of all drug targets [10], pointing out the importance of this protein class. Nevertheless, soluble proteins are far better investigated than membrane proteins. This disproportion is reflected by the number of entries in the RCSB protein database (PDB) with less than 3% membrane proteins [11]. The concentration of this kind of protein in membranes is highly compartment specific and depends on its enzymatic

activity. Therefore, the membrane protein concentration is no more than about 20% in myelin and 23% in the human red blood cell plasma membrane, but 50% in other plasma membranes and up to 76% in the mitochondrial membrane [12, 13].

1.2.1 Classification

Membrane proteins are classified according to their α -helical and β -barrel secondary structure and their topology. Until today, β -barrel membrane proteins are only found in the outer membranes of Gram-negative bacteria, mitochondria and chloroplasts [14]. The α -helical integral membrane proteins are subdivided by their topology into partial (monotopic) or complete (bitopic or polytopic) membrane penetrating proteins [15]. Furthermore, bitopic proteins (single-pass) are distinguished by their periplasmic terminus (type I: N_{out} or type II: C_{out}), whereas polytopic proteins (multi-pass) are generally termed type III and β -barrel proteins type IV [16]. Examples for bitopic proteins are receptors (e.g. EGFR), cell adhesion proteins (e.g. integrin) or protein anchors. Examples for polytopic proteins are transporters (e.g. ABC-transporter), channels (e.g. voltage-gated ion channel), receptors (e.g. G protein-coupled receptor) or proteases (e.g. rhomboids). The topology of the proteins is characterized by the “positive-inside rule” which implies that positively charged amino acids are more prone to face the cytoplasm [17, 18].

The membrane spanning regions are called transmembrane domains (TMDs) and generally consist of more than 20, mostly hydrophobic, amino acids [3]. As these stretches are longer and more hydrophobic in the membrane than in soluble protein regions, they can be easily predicted as TMDs by simple hydropathy plots as shown by Kyte and Doolittle in 1982 [19-21]. By the introduction of a “Biological hydrophobicity scale”, which describes the enthalpy ΔG_{app}^{aa} (apparent free energy of membrane insertion of every single amino acid) of membrane insertion for each amino acid at each TMD position, the prediction was improved more than 20 years later (s. 1.2.3) [22, 23]. Attempts to improve the discrimination efficacy further, are still ongoing. In this context, the terminal hydrophobic helices rule for multi-pass proteins, saying that most C-terminal TMDs are more hydrophobic than internal helices [24, 25], is utilized, for example.

1.2.2 Recognition and translocation

To ensure the specific recognition of membrane or periplasmic proteins, there are complex membrane protein insertion and transportation machineries in every organism. In

prokaryotic cells the main protein for membrane protein insertion, embedded in the cytosolic membrane, is the SecYEG translocon which is homologous to the eukaryotic Sec61-type translocon. It acts in concert with several cofactors and is responsible for the actual insertion in most cases. The translocation of proteins through the membrane is mainly realized by two different pathways, which are named after their unique signal peptide recognition proteins SecA or SRP (signal recognition particle), respectively [26]. The membrane proteins are recognized via a pathway specific N-terminal signal peptide [27]. It consists of a positively charged N region, followed by a hydrophobic H region and a polar C region, which also contains the signal peptidase cleavage site [28]. The regions, but not the amino acid sequences for recognition, are conserved over eukaryotes and prokaryotes so that signal peptides are in general interchangeable [29]. Though a wide range of sequences can be recognized as signal peptides [30] it was shown that they can contain additional information regarding for example the translocation pathway [27], translocation efficacy [31, 32] or signal peptide cleavage rate [33, 34]. Furthermore, they may contain post-cleavage functions such as self-antigens on immune cells [35] or cofactors [36].

One of the two possible pathways for translocation is the co-translational SRP-dependent pathway where the ribosomal mRNA to protein translation occurs in parallel to the translocon-mediated protein translocation and membrane insertion. In this pathway, the very hydrophobic signal peptide is recognized by the SRP which binds the ribosome [37-39]. This pathway is mainly used for integral inner membrane proteins as this process avoids solvent exposure of hydrophobic, membrane embedded protein regions. The signal sequence is usually the first TMD of a multi-pass protein and is also called signal anchor, as it is not cleaved by a signal peptidase [26]. However, there are also well investigated exceptions like the human calcitonin receptor which have a cleavable signal peptide [40]. Hence, it is problematic to use the SRP-dependent pathway for the expression of recombinant membrane proteins, if they must not have a signal anchor. The SRP, together with the ribosome and the nascent polypeptide chain, bind the SRP-receptor FtsY on the cytoplasmic membrane. The hydrolysis of two GTPs (one by FtsY, and one by SecYEG) leads to the release of the nascent chain from the SRP [41]. In the following, this newly synthesized polypeptide is translocated into and through the membrane by the holo-translocon (SecYEG, YidC, SecDF/YajC) or YidC or SecYEG only.

The SecA-dependent pathway on the other hand is preliminary used for secreted and outer membrane proteins. Furthermore, it has been shown that this pathway is efficient for the insertion of inner membrane proteins in artificial systems [42, 43]. Usually, it is a post-translational process that is supported by the ATPase SecA [38], which recognizes the less hydrophobic signal peptide [37, 44]. The two chaperones SecB and TF keep the completely translated nascent polypeptide in a translocation competent unfolded state [45-47]. After signal peptide recognition, SecA binds to the SecYEG complex and drives translocation through the SecYEG channel by ATP hydrolysis and the proton motive force [48]. Subsequently, the signal peptide is cleaved off by a signal peptidase.

Besides the two main pathways the Tat pathway (twin-arginine translocation) is an additional, important post-translational secretion system in bacteria. In contrast to the other two pathways the Tat pathway can transport folded proteins across the inner membrane to the periplasm [49]. These folded proteins are also recognized by a signal peptide. This peptide, however, is less hydrophobic than in the SecA or SRP-dependent pathway and contains a twin-arginine motif in the N region [50, 51]. This system for the export of folded proteins is necessary for proteins which have to bind for example complex redox cofactors in the cytoplasm [52].

1.2.3 Insertion and folding

The mechanism of TMD insertion from the translocon into the membrane is under heavy debate. The central protein of the bacterial translocon, stabilized by SecE, is SecY, which is composed of ten TMDs that form a channel through which nascent polypeptide chains can pass [53, 54]. The SecY protein is U-formed with a lateral gate between the TMD 2b and TMD 7 and a plug inside the channel, which prevents ion leakage through an inactive translocon. The bound signal peptide introduces conformational rearrangements, which widen the channel and remove the plug, so that the nascent polypeptide can enter the translocon [55].

In the relatively simple “in-out” model (Figure 1 A) a passing hydrophobic polypeptide segment is recognized as TMD by a hydrophobic ring [53, 56, 57], so that the lateral gate opens and the TMD enters the phospholipid bilayer [58]. In the “sliding” model (Figure 1 B) the hydrophobic polypeptide segment does not enter the central channel, but slides along the lateral gate - always exposed to the acyl chains of the lipids. The position in the

lateral gate represents the hydrophobicity of the polypeptide. If the region of the polypeptide is hydrophobic enough it enters the lipid bilayer. It is suggested that there are no distinct open and close states. TMD insertion rather is a continuous process, in which all energetic available conformations are explored by the potential TMD [59]. This scanning of conformations could explain why even mutations that are not located in the lateral gate region influence the insertion efficacy or topology.

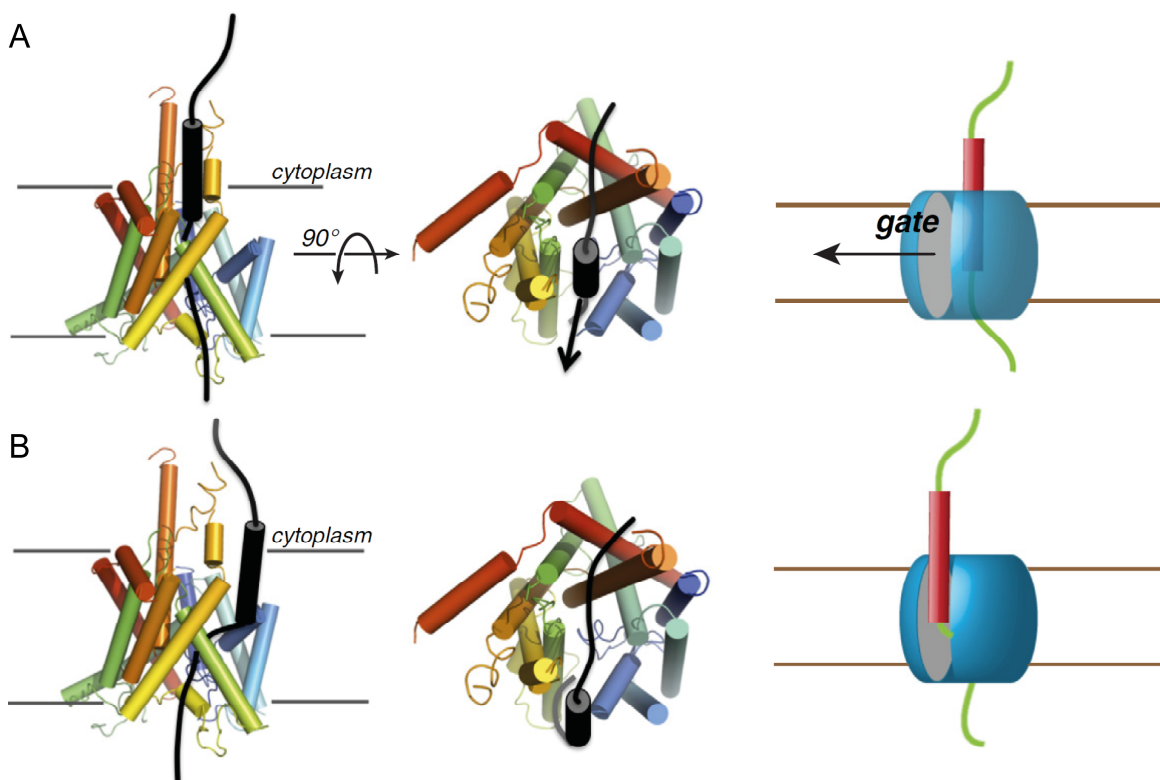


Figure 1: Two models for membrane insertion of TMDs. Left side: colored cylinders: TMDs of a translocon (10 TMDs of SecY and 2 TMDs of SecE). Black: translocated polypeptide. Black cylinder: TMD of a translocated polypeptide. Right side: schematic illustration of the two models. Blue: translocon. Red: TMD of a translocated polypeptide. Green: hydrophilic, not inserted region of a translocated polypeptide. **A)** “in-out” model: the TMD moves through the central channel and exits the translocon sideways through the lateral gate. **B)** “sliding” model: the TMD does not enter the central channel but slides along the lateral gate until it enters the lipid bilayer in an equilibrium. Adapted from Cymer et al. 2015 [59].

In this context the question arises how and when a sequence is recognized as a TMD to be inserted into the lipid bilayer or transported through the translocon into the periplasm. For this purpose, Hessa et al. [22] defined a biological hydrophobicity scale using ΔG_{app} (apparent free energy of membrane insertion) by measuring the impact of each amino acid at each TMD position ΔG_{app}^{aa} to the insertion efficacy [23]. To calculate ΔG_{app} for a TMD all ΔG_{app}^{aa} values are summed up. Hence, apolar amino acids can compensate for unfavored amino acids and thus, facilitate membrane insertion. It was, however, shown, that this

value only gives a hint, as ionic amino acids can be overestimated and the sequence context plays a role in membrane insertion efficacy, too [23, 60].

1.3 TMD-TMD Interaction

Interactions between TMDs play an important role in many cellular processes. First, they are essential for the correct folding of multi-pass membrane proteins. The interactions of the membrane embedded α -helices ensure the correct tertiary structure which is crucial for the function. Furthermore, they are responsible for quaternary structures, such as the formation of translocons or the respiratory chain. Additionally, intermolecular interactions can be responsible for receptor activation, which is for example the case for the oncogene HER2 in breast cancer [61, 62]. The importance of TMD-TMD interaction was shown for instance by the determination of the complete membrane protein interactome of *Arabidopsis thaliana*, where 12102 interactions between membrane proteins and soluble or membrane proteins could be found with a split-ubiquitin yeast two-hybrid screen [63]. Babu et al. determined membrane protein complexes in *Saccharomyces cerevisiae* were determined by copurification followed by mass spectroscopy [64]. Here, 1726 membrane protein - protein interactions and 501 putative heteromeric complexes were found.

1.3.1 Interaction forces

The geometry of interhelical interactions is defined by the crossing angle Ω and the direction of the two α -helices (parallel or anti-parallel). In parallel formation the conformation is called left-handed if $\Omega > 0^\circ$, or right-handed if $\Omega < 0^\circ$ [65]. In anti-parallel conformations the crossing angle of left-handed pairs is about -155° , for right-handed about 145° . Left-handed conformations are more abundant and have a heptad motif [66], the right-handed have a tetrad motif [65] which terms the periodic appearance of small amino acids (G, A, S). The heptad motif is derived from the “knobs into holes” packing, which is found in soluble coiled-coil structures and is also characteristic for “leucine-zippers” [67, 68]. These idealized motifs define the interaction interface of α -helices and hence, the position of all other interacting amino acids.

Inter- and intramolecular interactions of soluble proteins are well investigated and the impacts of the different forces on the protein stability are well defined. Even though the environment in a lipid bilayer is completely different, the same fundamental forces including van der Waals packing, hydrogen bonding, salt-bridge and aromatic π

interactions can be found. The main difference is the absence of water for what reason (partially) charged atoms cannot be complexed and stabilized by hydrogen bonds. In soluble proteins the buried amino acids are more apolar than the solvent exposed ones, which causes the hydrophobic collapse during protein folding [12, 69]. In contrast to this, in TMDs it is the other way around since the non-polar lipid bilayer is the solvent and the interfacial and buried amino acids are more polar [70-72].

The most common and best investigated interaction motif is the GxxxG, or more general (small)xxx(small) ($x = \text{any amino acid, small} = \text{G, A, S}$) [73, 74]. The small amino acids can maximize van der Waals interactions and enable intermolecular hydrogen bonds between the backbones ($\text{C}\alpha\text{-H} \cdots \text{O}=\text{C}$) or the backbone and the side chains of some amino acids, such as cysteine, threonine or phenylalanine ($-\text{SH}/-\text{OH} \cdots \text{O}=\text{C}$) [75]. For example the TMD of Glycophorin A [76, 77], the homo- and heterotypic interacting TMD of Integrin αIIb [78-81] or the homotypic interacting TMD 5 of the multi-pass protein TXA₂ [82] are proven to interact via GxxxG motifs. ~12% of all single-pass TMDs contain at least one GxxxG motif [74], which means an overrepresentation of only 32% compared to a random distribution of glycines [83]. Hence, it is not surprising that there are also GxxxG independent interactions such as it is the case for DAP12, which instead interacts homotypically via an aspartic acid and a threonine, though it also consists of a GxxxG motif [84, 85]. Even though, the specificity of GxxxG dependent interacting TMDs can be explained by the sequence context and additional interacting amino acids. Serine for example does not only work as a small amino acid facilitating close contact between two TMDs, but can also form hydrogen bonds to other amino acids, as well as other polar or ionizable amino acids can do. They are the basis for distinct motifs, such as serine zippers or polar clamps [70]. Though the serines can form hydrogen bonds, the interaction providing motif is probably more related to the small side chain at every seventh position as described for the leucine-zipper motif [86, 87]. The existence and function of salt-bridges in TMDs is widely discussed as the membrane integration of ionized amino acids is energetically unfavored due to high desolvation energies. In the proposed two-step model [88], which separates TMD membrane integration and TMD-TMD interaction into different processes, salt-bridges should be very stable, once the ionized amino acids are integrated. This can lead to irreversible interactions, such as the T cell receptor-CD3 complex [89, 90]. Furthermore, it is hard to distinguish hydrogen-bonds and salt-bridges in membranes, even in high resolution structures [91]. Therefore, other methods are necessary

to analyze distinct bonds [92], which dramatically limits the dataset. Nevertheless, it was shown in a library screening that homotypic interacting TMDs containing ionizable amino acids can have a very high dimerizing propensity [93]. Additionally, to their stabilizing function, salt-bridges are necessary for various functions of membrane proteins, such as transporter activity of the lactose carrier LacY [94], activation of rhodopsin [95] or gating of the viral voltage-gated potassium channel K_{cv} [96]. The aromatic side chains of phenylalanine, tyrosine and tryptophan can interact with themselves ($\pi - \pi$) and with basic residues such as lysine, arginine or histidine (cation - π) due to their quadrupole moment which was already shown in the TOXCAT assay [97]. Furthermore, homotypic interacting TMDs with $\pi - \pi$ interacting residues were enriched in several library screenings using degenerated codons [93, 98, 99]. In summary, the correct folding of membrane proteins is governed by various weak interactions [100] including a complementary shape of the interacting TMDs.

TMD-TMD interaction interfaces can overlap with TMD-lipid interfaces [5]. This is for example the case for C99, the cleavage product of the amyloid precursor protein (APP) [101]. On the one hand it was shown that the C99 forms a homodimer mediated by a GxxxG motif [102] and on the other hand the same motif is involved in cholesterol binding [103, 104]. As the equilibrium dissociation constant for C99 homodimerization is ~ 0.5 mol%, which is about 1000 times higher than in natural membranes [101, 105], the importance of homodimerization *in vivo* is under discussion [5]. Especially, as the equilibrium constant for cholesterol binding is ~ 3 mol%, which is at the lower end of physiological cholesterol concentrations in mammalian cells [106]. Thus, most APP or C99 proteins are binding cholesterol. As dimeric complexes are very unlikely, it is not surprising that cholesterol containing lipid rafts and cholesterol concentration are correlated with Alzheimer's Disease [107]. Hence, multifunctional interfaces can cause false results in TMD-TMD interaction measurements at non-physiological concentrations or without competitive binding partners, although for this special case in Alzheimer's Disease it is not proven *in vivo* yet.

1.3.2 In vivo TMD-TMD interaction assays

All genetic TMD-TMD interaction assays are based on transcription activation of a reporter protein like β -galactosidase (ToxR [77], GALLEX [108], BACTH [42, 109]), chloramphenicol acetyltransferase (TOXCAT [110], POSSYCCAT [111]) or GFP (AraTM

[43]). The DNA-binding domain is fused to the cytoplasmic end of the TMD and thus, directly or indirectly influences the transcription. It was demonstrated for the ToxR [112], GALLEX [113] and BACTH [114] assays that they work with full length membrane proteins as well.

Homotypic ToxR-related, GALLEX and AraTM assays

The principle of the homotypic interaction assays ToxR [77], TOXCAT [110], POSSYCCAT [111], GALLEX [108] and AraTM [43] is very similar (Figure 2). The main difference is the sensor domain (ToxR, LexA or AraC) which influences the gene transcription (*lacZ*: β -galactosidase, *gfp*: GFP, *cat*: chloramphenicol acetyltransferase). In all cases, the fusion domain has to form a dimer to bind the promoter/operator region (*ctx*, *op⁺/op⁺* or *p_{BAD}*) and alter the reporter gene transcription [115]. In the ToxR and the AraTM system the transcription is activated, in the GALLEX system it is repressed. Hence, the mRNA level is assumed to be proportional to the amount of expressed reporter protein. The β -galactosidase activity can be quantified by ortho-Nitrophenyl- β -galactoside (ONPG) turnover, which is measured colorimetrically [116]. The activity is measured in Miller units [117] and has to be normalized either to the positive control Glycophorin A wild type (GpA wt) which dimerizes well in the ToxR/TOXCAT/POSSYCCAT assay or to the negative control GpA G83I, which dimerizes poorly in the GALLEX assay, respectively. The ToxR and the TOXCAT/POSSYCCAT systems differ in the reporter protein, which is β -galactosidase in the ToxR system, or chloramphenicol acetyltransferase (CAT) in the other systems, respectively [77, 110, 111]. CAT can be quantified by measuring the resistance against the antibiotic chloramphenicol. In the TOXCAT assay the promoter/reporter gene *ctx::cat* is plasmid coded, whereas in the POSSYCCAT it is chromosomally coded. The sensitivity for dimerized ToxR proteins is higher in the TOXCAT assay, as there are more available binding sites in the cell and no interference with the chromosomal DNA [115].

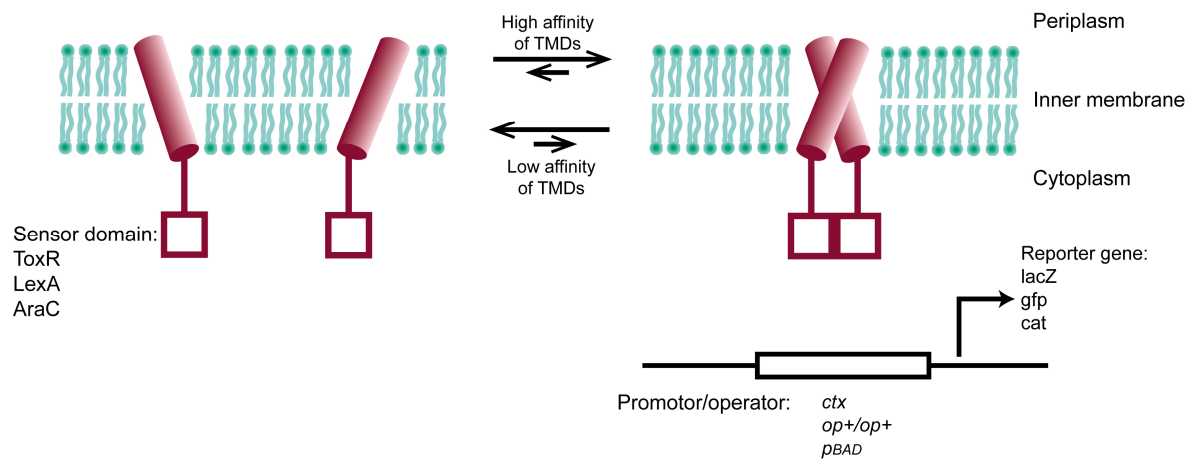


Figure 2: Schematic illustration of the homotypic ToxR/TOXCAT/POSSYCCAT [77, 110, 111], GALLEX [108] and AraTM [43] assay. Green: Lipid bilayer of the *E. coli* inner membrane. Cylinder: TMD. Square: sensor domain (ToxR/TOXCAT/POSSYCCAT: ToxR, GALLEX: LexA, AraTM: AraC). Periplasmic domains are not shown. The amount of dimeric sensor domains depends on the affinity of the TMDs. Monomeric sensor domains (left side) cannot bind to the promoter/operator region (ToxR/TOXCAT/POSSYCCAT: *ctx*, GALLEX: *op⁺/op⁺*, AraTM: *p_{BAD}*). Dimeric sensor domains (right side) bind to the promoter/operator region and alter reporter gene transcription (*lacZ*, *gfp* or *cat*).

The principle of these dimerization assays has proven its worth in the last 20 years since the ToxR assay was published in 1996 [77]. The ToxR assay was the template for the newer TOXCAT [110], GALLEX [108], AraTM [43] or CadC based [118] assays whereas mainly ToxR, TOXCAT and GALLEX are widely used. They are suitable for the determination of the impact of single amino acids to the affinity by e.g. alanine or leucine scanning. The TOXCAT [73, 97, 110, 119], as well as the POSSYCCAT [93, 98, 111, 120] system, were successfully used for library screening. The randomized amino acid positions have to be chosen carefully as it was shown that the GALLEX and the ToxR-like assays are TMD orientation and TMD length dependent [108, 121]. Hence, usually four different orientations of each TMD have to be tested to get reliable results. All proteins contain a periplasmic MalE (Maltose-binding periplasmic protein, UniProtKB [122] accession number P0AEX9) which has a molecular weight of about 41 kDa. On the one hand, this large domain could sterically inhibit, or influence the interaction of two proteins and lead to false negative result. On the other hand, it ensures proper topology and facilitates insertion into the membrane [123, 124]. The ToxR and GALLEX proteins especially profit from this as they are type II membrane proteins and thus, do not contain a signal peptide. The insertion efficacy and the topology can be proven in a maltose complementation assay [125, 126]. The ToxR and GALLEX assays are using the GpA TMD for the necessary normalization though it is a type I membrane protein in human erythrocytes [127]. The influence of this opposite topology in the *E. coli* membrane to for

example the dimerization propensity is not investigated, yet. The AraTM protein is a type I membrane protein and although it was emphasized that this topology is an advantage over the previous established assays, no results are published for GpA [43]. One problem of all these systems is that they are transcription-based and thus, many “dark” steps lie in between the translation, the actual interaction and the output signal. First, the cytoplasmically translated proteins have to be inserted into the membrane. This step might be TMD-dependent but can be checked by maltose complementation. Nevertheless, non-integrated proteins can persist in the cytoplasm as inclusion bodies or soluble complexes. As DNA binding also occurs in this compartment, false-positive signals are possible but have not been observed, so far. Second, the dimerization of the TMDs results in the formation of an active dimeric transcription factor (ToxR, LexA, AraC) which binds its specific promoter region. After recognition, the reporter gene first has to be transcribed and translated into the reporter protein, whereby the efficacy of both steps is cell and environment dependent. For β -galactosidase detection the cells have to be lysed first and the ONPG turnover is measured in the following. The Miller units are normalized to a GpA positive control, which then gives the final result.

Heterotypic GALLEX, DN-ToxR and DN-AraTM

Originally developed for the measurement of homotypic TMD-TMD interactions, all three assays were modified to measure heterotypic interactions, too. In the GALLEX assay the binding domain and the DNA binding site were modified (Figure 3 A) [108] in a way that only a heterodimer, but no homodimers of LexA wt and LexA 408 can bind to the new promoter/operator region (op408/op+) on the genome and thus repress the transcription. The measurable output is the reduction of the β -galactosidase expression caused by the heterodimer.

In the dominant-negative ToxR system (DN-ToxR) [128] and the dominant-negative AraTM system (DN-AraTM) [129] (Figure 3 B) only the DNA binding domains were modified. Here one TMD is fused to the wild type domain and the other to an inactive mutant. Only the wild type homodimer can activate the transcription by binding the operator region (*ctx* or *p_{BAD}*). When the wild type and a nonfunctional mutant proteins (*: ToxR S87H, AraC R210A) are coexpressed some wild type proteins form inactive heterodimers. So these proteins are not available for homodimerization anymore and there is less active wild type homodimer. The signal decreases if the homotypic interaction is

suppressed by the heterotypic interaction. As an output the difference of the signal between the wt/mut cells and the wild type only expressing cells is calculated.

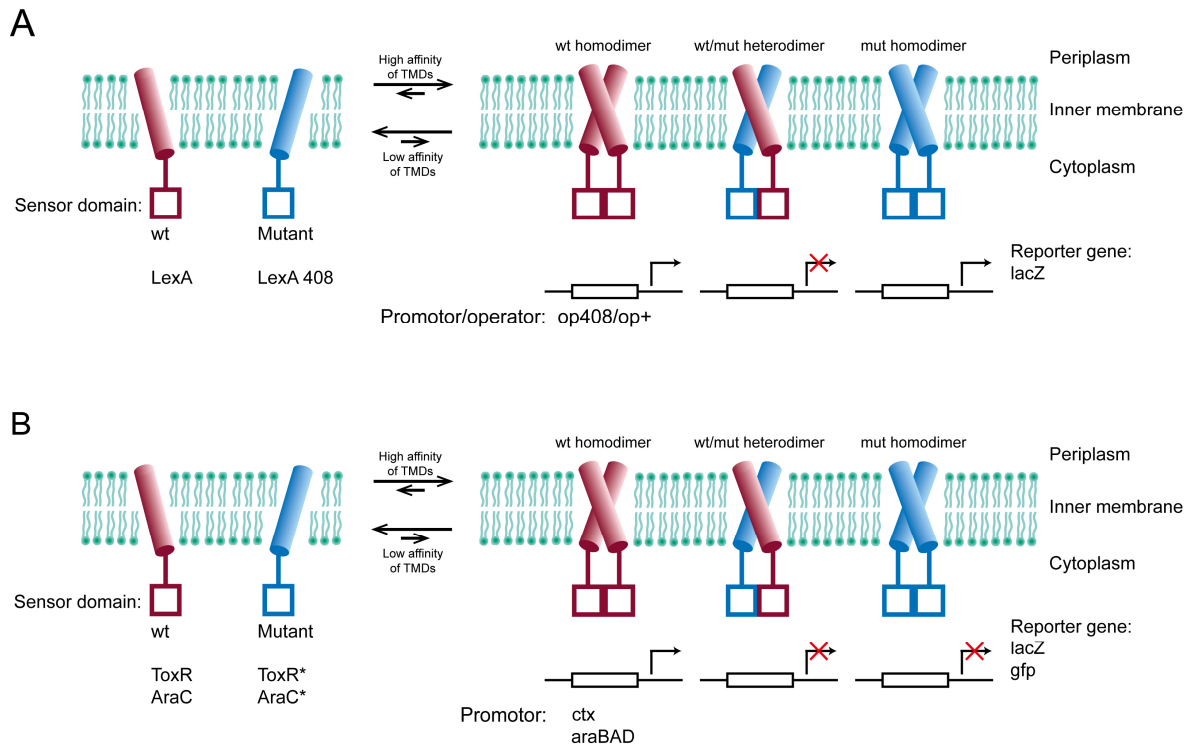


Figure 3: Schematic illustration of the heterotypic variants of the GALLEX [108], dominant-negative ToxR [128] and dominant negative AraTM [129] assays. Green: Lipid bilayer of the *E. coli* inner membrane. Cylinder: TMD. Red square: sensor domain wild type (wt) (GALLEX: LexA, DN-ToxR: ToxR, DN-AraTM: AraC). Blue square: sensor domain mutant (GALLEX: LexA 408, DN-ToxR: ToxR*, DN-AraTM: AraC*). Periplasmic domains are not shown. The amount of dimeric sensor domains depends on the affinity of the TMDs. **A**) GALLEX: the monomeric TMDs (left side), the wild type (wt) homodimer (LexA/ LexA) and the mutant (mut) homodimer (LexA 408/ LexA 408) cannot bind to the promotor/operator *op408/op+* (right side). Only the wt/mut heterodimer (LexA/LexA 408) binds to the promotor/operator and represses the transcription of the reporter gene *lacZ*. **B**) Dominant-negative ToxR system and DN-AraTM: the monomeric TMDs (left side), wt/mut heterodimer (ToxR/ToxR* or AraC/AraC*) and mutant homodimer (ToxR*/ToxR* or AraC*/AraC*) (right side) cannot bind to the promotor (*ctx* or *p_{BAD}*). Only the wt homodimer (ToxR/ToxR or AraC/AraC) binds to the promotor and activates reporter gene transcription (*lacZ* or *gfp*).

In general the variants of the ToxR, GALLEX and AraTM assay used to detect heterotypic interactions have similar advantages and disadvantages as their original versions. It is noteworthy that all assays can only look at one dimer (wild type homodimer or heterodimer) but its amount always depends on the affinity of the other two possible dimer formations, too. The most direct assay is the GALLEX assay as the result is absolute and not influenced by the homotypic interactions. Furthermore, the ToxR DN-system depends on one high- and one low-copy plasmid so that the expression of the two chimeric proteins differs which can influence the result. The DN-AraTM proteins are coded on a low- and a medium-copy plasmid whose copy numbers still differ by a factor of more than five [130, 131]. Even though the GALLEX proteins are coded on two low-copy plasmids and the

copy numbers are more similar, an influence of copy number cannot be excluded, too [124]. Neither the dominant negative assays nor the heterotypic GALLEX could be demonstrated to be suitable for heterotypic screening purposes, although the GALLEX assay should be if using for example MacConkey's selective agar [115]. This agar is selective for Gram-negative bacteria and sensitive to lactose-fermenting bacteria but not sufficiently quantitative to distinguish middle and high affinity TMDs.

BACTH

Another transcription based system is the BACTH assay (Bacterial adenylate cyclase two-hybrid) [109]. Originally developed for the measurement of the interaction of soluble proteins it was recently adapted to interactions of transmembrane domains (Figure 4) [42]. In contrast to the previously described assays ToxR, GALLEX and AraTM (Figure 2) the BACTH assay does not depend on a dimerization of transcription domain but on protein complementation (s. 1.4). The *Bordetella pertussis* adenylate cyclase is split on DNA level into an N-terminal fragment (T25) and a C-terminal fragment (T18) which are both fused C-terminal to the TMDs of interest. Each fragment alone is inactive and cannot catalyze cAMP formation from ATP. If the TMDs dimerize, the two fragments T25 and T18 come into proximity and can fold to an active enzyme. The reconstituted adenylate cyclase can now convert ATP to cAMP. The catabolite activator protein (CAP) binds cAMP and thereby activates the *cAMP/CAP dependent* promoter [132] which controls a *lacZ* gene. The expression of β -galactosidase can be quantified colorimetrically by ONPG turnover [117]. As *E. coli* cells express endogenous adenylate cyclase, only adenylate cyclase deficient strains (*cya*⁻) can be used for this assay.

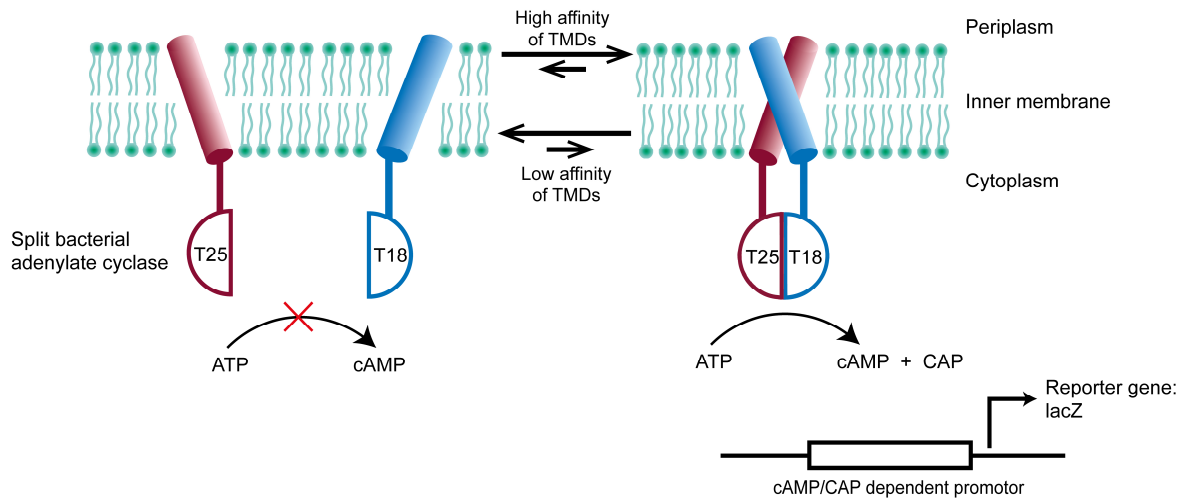


Figure 4: Schematic illustration of the BACTH assay (bacterial adenylate cyclase two-hybrid) [42, 109]. Green: Lipid bilayer of the *E. coli* inner membrane. Cylinder: TMD. Red semicircle: N-terminal T25 fragment of a bacterial adenylate cyclase (BAC). Blue semicircle: C-terminal T18 fragment of a BAC. The amount of reconstituted BAC depends on the affinity of the TMDs. Monomeric TMDs (left side): The BAC cannot reconstitute thus ATP cannot be circulated to cAMP. Heterodimeric TMDs (right side): The two BAC fragments reconstitute to an active enzyme which converts ATP to cAMP. cAMP and CAP form a complex which binds to the *cAMP/CAP* dependent promotor and activates reporter gene transcription (*lacZ*).

The BACTH assay is optimal for the measurement of heterotypic interactions as only the heterodimer can generate a signal and is in this respect similar to the heterotypic GALLEX. Hence, homotypic dimers are disregarded in this assay, too. It was successfully shown that the BACTH assay is suitable for the screening of interactions of soluble proteins [133, 134] but not yet for membrane proteins. The BACTH assay is not only applicable for parallel dimers but could theoretically be adaptable for antiparallel TMD-TMD interactions. However up to now, this has not been tested systematically. One advantage over the previously described direct transcription based assays is that the DNA does not have to be in spatial proximity to the membrane. The cAMP diffuses to the CAP and the complex then binds to the promotor. On the other hand this means that additionally to the described problems with transcription based methods, another step is introduced. So depending on the growth and stress conditions there may not be enough ATP for the reaction or an altered CAP concentration in the cell. Additionally it is very problematic that the two proteins are coded by one low- and one high copy plasmid which may result in different protein expression levels. As a consequence the results have to be interpreted with great care. Furthermore, the assay is not performed at the physiological temperature of 37 °C but at 30 °C which causes an altered lipid and acyl chain length composition [4] which could indirectly influence the dimerization propensity [135].

MYTH/MaMTH

Other split protein based systems are the MYTH assay (membrane yeast two-hybrid) [136, 137] and the adapted MaMTH assay (mammalian-membrane two-hybrid) [138]. In principle they both operate on a split ubiquitin. The MYTH assay was developed as a yeast two-hybrid system (YTH) for membrane proteins because the original system is only suitable for soluble, cytoplasmic proteins [139]. In the first YTH system the dimerized proteins had to enter the yeast nucleus to activate the reporter gene transcription. Both parts of the split transcription factor GAL4 had to bind simultaneously to the DNA. The fragments were fused to the investigated proteins that consequently had to enter the nucleus, too. Hence, in the next generation (MYTH) the transcription factor is decoupled from the dimerizing proteins (Figure 5). A full-length transcription factor (PLV: protein A-LexA-VP16) is fused C-terminal to the C-terminal ubiquitin fragment (Cub: amino acids 35 – 76) of a split ubiquitin. This fusion protein is again fused to the first transmembrane protein. The N-terminal ubiquitin fragment (NubG: amino acids 1-34, I13G) is obtained by fusion to the second transmembrane protein. As the transcription factor (PLV) is bound to a membrane protein it cannot enter the nucleus (left side). Due to dimerization of the two transmembrane domains the ubiquitin can reconstitute. Therefore, it can be recognized by endogenous ubiquitin specific proteases (UBPs) which cut off the PLV fragment. This can enter the yeast nucleus and finally activate the transcription of a reporter gene which can be detected. These reporter genes may be *lacZ*, coding for β -galactosidase, or *HIS3*, coding for imidazoleglycerol-phosphate dehydratase which is essential for histidine synthesis [140]. β -Galactosidase activity can be quantified by ONPG turnover [117] or by imidazoleglycerol-phosphate dehydratase activity on histidin deficient agar plates [141].

As there is no nucleus in *E. coli* cells the chromosome is always accessible for the transcription factor. Hence, this assay is not applicable in *E. coli* cells as the transcription would be activated independently from the dimerization and cleavage events. Besides UBPs would have to be expressed additionally to enable the cleavage as they are not endogenous in prokaryotic cells.

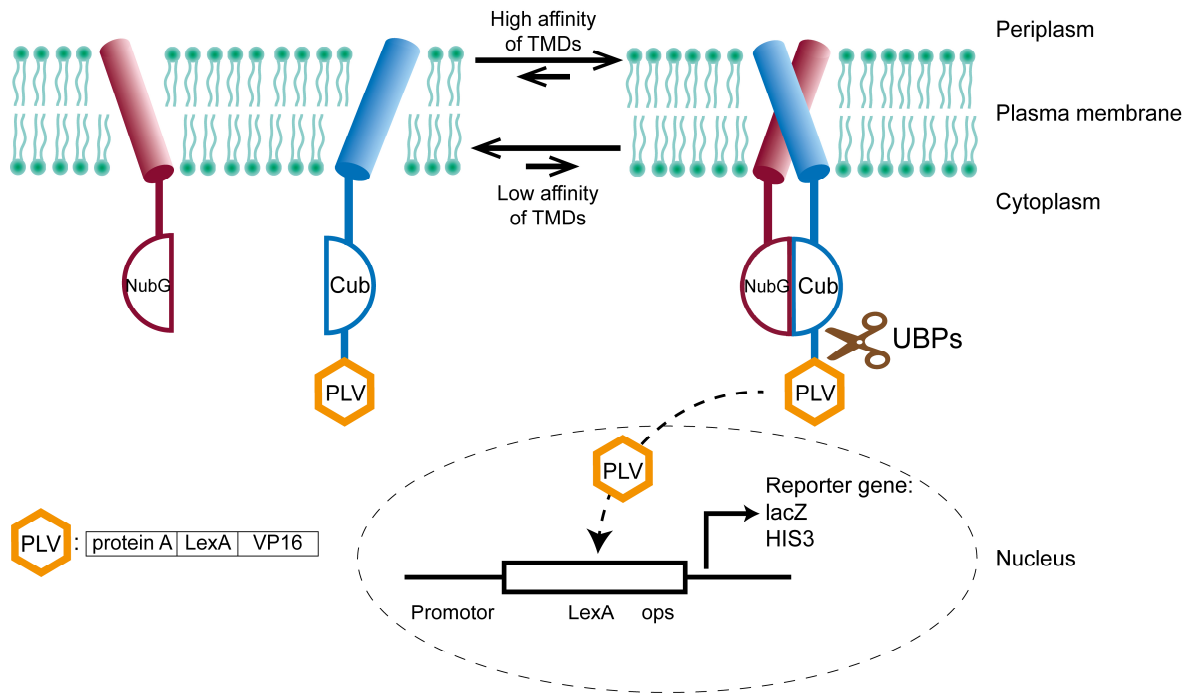


Figure 5: Schematic illustration of the MYTH assay (Membrane Yeast Two-Hybrid) [136, 137]. Periplasmatic domains are not shown. The N-terminal ubiquitin fragment (NubG) is fused C-terminally to one transmembrane protein (red), the C-terminal fragment (Cub) to another (blue). A PLV domain (orange; protein A-LexA-VP16) is again fused C-terminally to the Cub domain. After dimerizing of the transmembrane proteins the ubiquitin reconstitutes. Ubiquitin specific proteases (UBPs) recognize the ubiquitin and cut off the PLV domain which can enter the yeast nucleus and activate gene transcription. *lacZ*: β -galactosidase, *HIS3*: imidazoleglycerol-phosphate dehydratase.

Both the MYTH and MaMTH assays are only suitable for yeast and mammalian cells which is more challenging than *in vivo* experiments in *E. coli*, but enables dimerization experiments in native membranes and natural environments. Endogenous proteins can be tagged with the split proteins (iMYTH) [142] in the yeast chromosome. So the expression is still driven by the native promoter and ensures unaltered protein expression. The yielded protein concentration is different for each protein, which makes the data analysis difficult and complex. It is hard to distinguish whether two signals are unequal due to varying expression rates or due to different affinities. In the tMYTH variant the proteins are expressed ectopically from a plasmid [137, 143]. For yeast proteins this means that they are overexpressed and could thus influence any aspect in the cell's physiology. The transcription factor is expressed functionally and only due to membrane protein fusion not accessible for DNA binding in the nucleus. If there are Cub proteins left in the cytoplasm, which were not recognized by the translocon machinery, they could independently activate the gene transcription. This could be in particular problematic for overexpressed or non-yeast proteins. Overall the MYTH and MaMTH assays are suited for heterotypic membrane protein interaction measurements, especially because they can measure protein-

protein interaction in their natural environment. Up to now this system was not used systematically for TMD-TMD interactions.

1.3.3 Orientation dependence

A general problem in the ToxR [77, 121] and GALLEX [108] assays is the orientation dependence, which describes the influence of how the TMD interface is geometrically related to the sensor domain on the signal. Up to now there are no reports about this phenomenon in the AraTM or BACTH assay. It is caused by the architecture of α -helical TMDs (Figure 6 A), which is characterized by intramolecular hydrogen bonding of the peptide backbone between the carbonyl (i) and the amino hydrogen four amino acids further (i+4). This results in negative phi (-60°) and psi (-50°) angles which results in a very stable helical structure of 3.6 amino acids per turn and a rise of 1.5 Å per amino acid [12, 144, 145]. The backbone hydrogen bonding shields the polar carbonyl and amino groups from the hydrophobic surrounding area in the lipid bilayer and thus stabilizes the polypeptide. The complete unfolding of a 20 amino acid TMD backbone in a membrane would cost about 80 kcal mol⁻¹ which explains why unfolded polypeptides cannot persist in the bilayer [146]. Nevertheless, about one half of all TMDs are bent or contain disorders such as kinks caused by helix-interrupting prolines [147]. These flexible regions are thought to be important for function or positioning of important amino acids.

Assuming that there is an optimal TMD interface to sensor domain orientation for DNA-binding, which thus generates the highest possible signal, and that a deviation causes false-negative results, the challenge is to find their interface (Figure 6 B). If there is only one distinct, stiff TMD/TMD interface and a rigid linker between the TMD and the sensor domain, the α -helix architecture should allow pre-defined adjustments of the relative orientation. Addition of each N-terminal amino acid to the TMD helix turns sensor domain located upstream about 100° anticlockwise relative to the TMD-TMD interface [12]. Hence, for a complete coverage of all putative TMD-TMD interface to sensor domain combinations at least three amino acids have to be inserted (Figure 6 B, orientations 0, 1, 2 and 3). To avoid extension of the TMD, the same number of amino acids is deleted at its C-terminus. Thus, the crossing angle of the two TMDs and a potential hydrophobic mismatch are likely to remain constant. Comparing maximal and minimal signals of all orientations yields the “orientation dependence”.

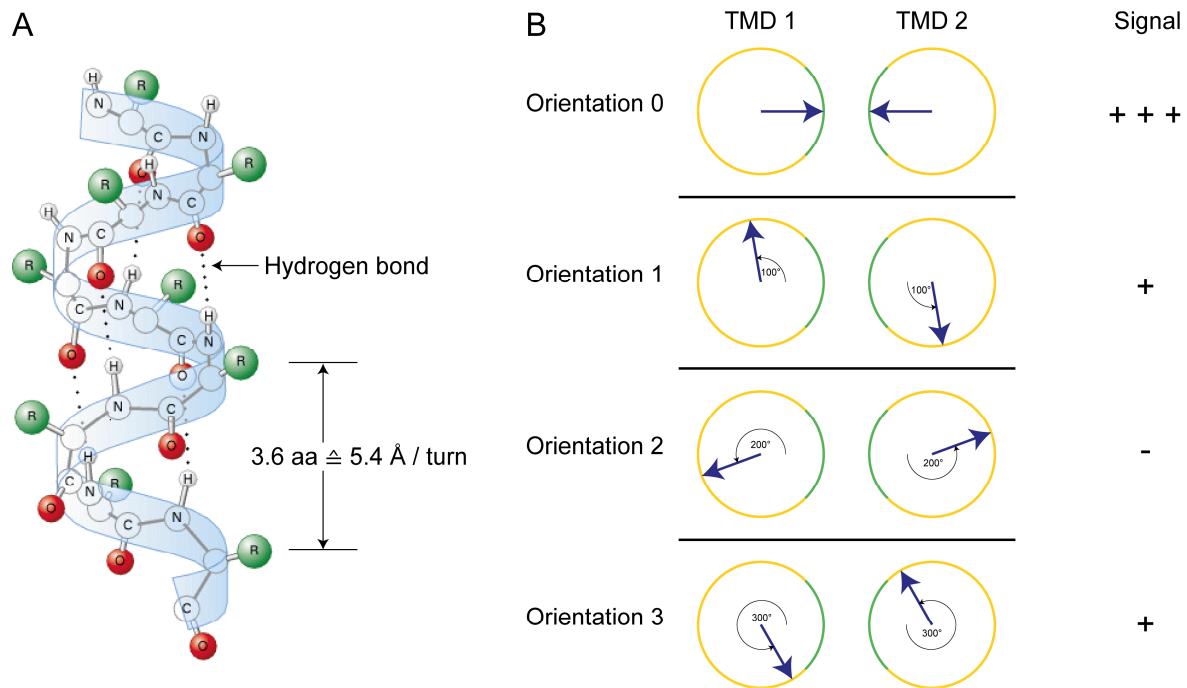


Figure 6: Relationship of the molecular structure of an α -helix and orientation dependence. **A**) Schematic illustration of the molecular structure of an α -helix (top to down: N- to C-terminus). Blue band: peptide backbone. Green balls (R): any amino acid side chains. Atoms of peptide bond (O, C, N, H) are labeled. Dotted line: intramolecular hydrogen bond between the carbonyl and i+4 amino hydrogen of the backbone. Each turn consists of 3.6 amino acids (aa) and the rise is 5.4 Å. Adapted and modified from Cruz 2011 [148]. **B**) Schematic illustration of the influence of added amino acids in an α -helical conformation between the TMD and a sensor domain to their relative orientation and the signal (right). Circle: top view of the TMD. Blue arrow: Orientation of the sensor domain (ToxR or LexA). Green: interfacial side of the TMD. Yellow: non-interfacial side of the TMD. Orientation 0: start construct. Orientation 1: one additional amino acid at the N-terminus of the TMD. Orientation 2: two additional amino acids at the N-terminus of the TMD. Orientation 3: three additional amino acids at the N-terminus of the TMD. Each inserted amino acid turns the sensor domain by 100° (orientations 1, 2 and 3) relative to the interfacial side of the TMD.

If the connection between the TMD and the sensor domain is very rigid and if there is one explicit interaction motif, only one or two of the four orientations should generate a signal. The more flexible the linker is the less orientation dependent the signal should be as the sensor domains would adapt to any situation. In general neighboring orientations (0,1; 1,2; 2,3; 3,0) should show more similar signals than opposite ones (0,2; 1,3).

Indeed the affinity of the TMDs could also be altered due to added or deleted amino acids and accordingly change the signal. For example the large cyto- and periplasmic domains could collide in one particular TMD interface to sensor domain orientation so that not all sides of the TMD are available for dimerization. Furthermore, the penetration depth of the interface could influence the signal if the head groups of the lipids or the hydrophobicity, which increases to the center of the membrane, have an impact on the interaction. If so, the first or the last orientation (0 or 3) should show the highest signal because the respective amino acids are in the correct neighborhood. Furthermore, some of the added or deleted

amino acids could be part of the interface. This could result in a clear signal shift between two orientations. In reality, any mixture of all these phenomena could lead to a false-negative measurement. This is not only the case for the ToxR and GALLEX assay but for all experiments which focus on single parts of proteins as for example the TMDs.

A very long and flexible linker between the TMD and the sensor domain should abolish orientation-dependence because it can structurally decouple two domains [149]. On the other hand a too long linker increases the degrees of freedom of the domains thus hindering transcription.

1.4 Protein complementation assays

The first split-protein-complementation assay (PCA) was reported in 1994 by Johnsson and Varshavsky [136] and utilized ubiquitin which is only processed by ubiquitin-specific proteases in its quasi-native, reconstituted conformation. Since then many different split proteins for the measurement of soluble protein-protein interactions were designed. All these proteins have to fulfill various requirements: (i) The amount of active protein has to be measurable. It may either be quantified directly (e.g. GFP [150]), by the products of its enzymatic activity (e.g. luciferase [151] or β -lactamase [152, 153]), by gene transcription activation due to educts (e.g. ubiquitin [136], adenylate cyclase [109] or TEV protease [154]) or by another secondary effect such as survival (e.g. DHFR [155]). (ii) There has to be a site where splitting the protein generates two inactive fragments. These fragments should (iii) not dimerize themselves but (iv) forced dimerization via a fused dimerization domain would reconstitute the protein's activity. (v) Smaller proteins are of advantage as the expression is more efficient and the possibility of steric hindrance is smaller. (vi) Optimally the reconstitution is independent of N- or C-terminal fusions of the investigated proteins.

A commonly used PCA is the Bimolecular Fluorescence Complementation (BiFC) assay which was developed for high-throughput applications and is basically a PCA utilizing a fluorescent protein [156-158]. The fluorescence can be quantitatively detected directly or used for library analysis by fluorescence activated cell sorting (FACS). Furthermore, it is also suitable for intracellular localization of interactions [156]. However, the method is critical because there is self-affinity of the fragments and the reconstitution is irreversible [159, 160]. Other common PCAs are based on the complementation of ubiquitin and

bacterial adenylate cyclase which were both also adapted for the measurement of membrane protein interactions (s. 1.3.2).

PCAs were used for the measurement of distinct protein-protein interactions in vivo, for library screenings and for in vitro applications. Up to now only the bacterial adenylate cyclase (BACTH [42]) was adapted for the systematical measurement of TMD-TMD interactions (s. 1.3.2). The ubiquitin based MYTH/MaMTH [137, 138] and BiFC [161-168] assays were only used for full length membrane proteins so far.

1.5 β -Lactam antibiotics

The first and most famous β -lactam antibiotic is penicillin which was discovered in 1928 by Alexander Fleming [169] and isolated in 1940 by Ernst Chain [170]. Since then many derivatives were developed to overcome resistant bacteria or to extend their activity spectrum to for example Gram-negative bacteria (e.g. ampicillin) [171]. The members of this antibiotic class differ in their core ring structure which consists of a β -lactam ring and mostly a fused, carboxylated five or six-membered ring which can be unsaturated and which can contain a single oxygen or sulfur atom [172]. Furthermore, they vary in the side groups to avoid binding and degradation by β -lactamases (s. 1.5.2).

1.5.1 Mechanism of antibiotic function

β -Lactam antibiotics inhibit the cell wall synthesis by irreversible binding of penicillin-binding proteins (PBP) due to structural similarity to their natural substrate D-Ala-D-Ala [173-175]. PBPs are a protein group and UniProtKB [122] lists 11 different PBP in the *E. coli* lab-strain K12. The high molecular mass (HMM) proteins (PBPs 1-3) are DD-transpeptidases and essential for cell elongation, shape determination and septation [176]. The low molecular mass (LMM) proteins (PBPs 4-7 and AmpH) are DD-carboxypeptidases whose functions are less clear but which are possibly involved in murein (peptidoglycan) rearrangement. DD-transpeptidases cross-link glycan chains via for example pentapeptides (Figure 7) and build up the murein of the bacterial cell wall [177]. Without these linkages the cell wall becomes unstable and cannot withstand the internal turgor pressure [178]. The common feature among all PBPs is an active-site serine which can be acetylated by β -lactams or by the substrate mimic diacetyl-L-Lys-D-Ala-D-lactate [179-182]. The complete blocking of the active site by this covalent binding explains the high potency of β -lactams as antibiotics.

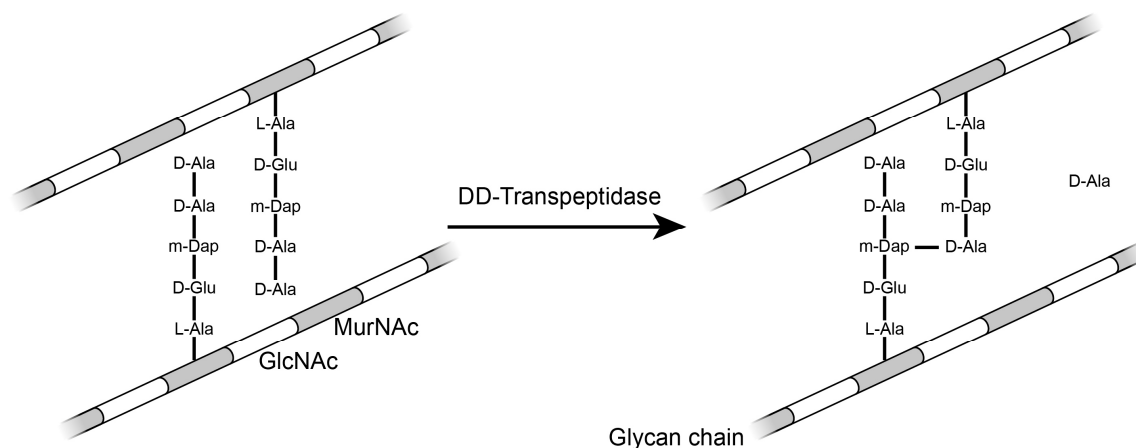


Figure 7: Structure of murein and cross-linking by DD-transpeptidase. Glycan chains consist of alternating GlcNAc (N-acetylglucosamine) and MurNAc (N-acetylmuramic acid) groups. Some MurNAc contain pentapeptides (L-Ala, D-Glu, m-Dap (meso-diaminopimelic acid), D-Ala, D-Ala). The DD-transpeptidase cross-links the amino group of one m-Dap with the first D-Ala of another peptide by cleavage of the peptide bond to the terminal D-Ala.

1.5.2 β-Lactamase

β-Lactamases (EC 3.5.2.6) are a group of periplasmic enzymes which can cleave β-lactam bonds by hydrolysis. They play an important role in modern medicine as they mediate antibiotic resistance by the degradation of β-lactam antibiotics such as penicillin or ampicillin. β-Lactamases are classified into the Ambler classes A through D, based on amino acid homology [183-188] or into the less common Bush-Jacoby-Medeiros groups 1 through 4, based on substrate and inhibitor profile [189]. The Ambler classes A, C and D are structurally homologous serine hydrolases whereas class B are Zn²⁺ dependent metallohydrolases [190]. The two main representatives of class A are TEM [191] and SHV [192] which share 68% sequence homology and differ in the size of the active site. The proposed reaction mechanism is shown in Figure 8 A [190]. In brief, the substrate β-lactam antibiotic is bound and positioned by the amino group of Lys₂₃₄, the amide hydrogens of Ser₇₀ and of the amino acid at position 237. Ser₇₀, supported by Glu₁₆₆, forms a high-energy, tetrahedral acylation complex with the β-lactam via a nucleophilic attack to the carbonyl of the amide bond. Due to protonation of the amide nitrogen and regeneration of the Glu₁₆₆ the amide bond is cleaved. An activated catalytic water molecule hydrolyses the ester bond between the β-lactam and the Ser₇₀ and thus regenerates the complete active pocket of the β-lactamase so that the deactivated β-lactam antibiotic can be released.

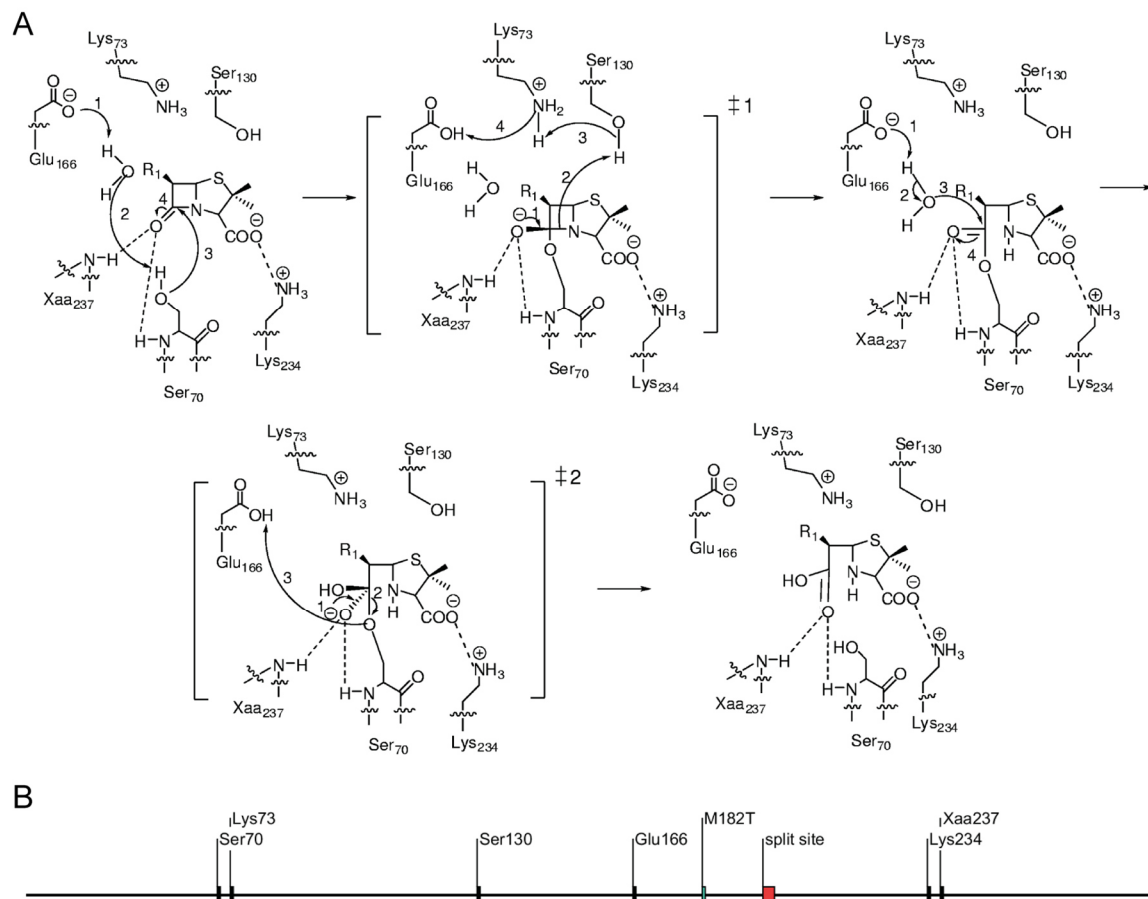


Figure 8: Catalytic center and reaction mechanism of class A serine β -lactamases. **A**) Reaction mechanism of class A serine β -lactamases. After the hydroxyl group of Ser₇₀ is activated by the protonation of the carboxy group of Glu₁₆₆ it can perform a nucleophile attack on the carbonyl group of the β -lactam bond. The tetrahedral acylation intermediate ($\ddagger 1$) is unstable and after protonation of the β -lactam nitrogen the β -lactam bond is cleaved and the carboxy group of Glu₁₆₆ regenerated. This group activates a catalytic H₂O molecule which can perform a nucleophile attack on the carboxy carbon of the broken β -lactam bond. The second tetrahedral deacylation intermediate ($\ddagger 2$) is also unstable and by hydrolysis of the bond between Ser₇₀ and the substrate the complete active center is regenerated. The deactivated β -lactam can be released. Figure adapted and modified from Drawz and Bonomo, 2010 [190]. **B**) Scheme of class A serine β -lactamases amino acid sequence. Catalytic active amino acids (black) [190], stabilizing mutation M182T (blue) [193] and split site for β -lactamase protein complementation assay (red) [153] are emphasized.

Due to high clinical relevance of β -lactamases the protein is well investigated. Originally the active enzyme was quantified colorimetrically by hydrolysis of the β -lactam ring of the chromophore nitrocefin which thus changes its absorption maximum from 390 nm to 492 nm and was used for resistance detection in various bacterial strains [194, 195]. Later the enzyme was utilized for different issues in microbiology and protein chemistry. As nitrocefin and β -lactam antibiotics are not membrane permeable [196] they are perfectly suited for the quantification of periplasmatic proteins [197-199], as membrane integration control and for topology determination [200, 201]. On the other hand nitrocefin has to traverse the outer membrane via porins which can cause detection problems if the permeability is reduced [202, 203]. Another application of this enzyme was the use for

investigation of the haemolysin transport pathway as it also folds in the culture medium [204].

β -Lactamase is also very suitable for measurements in eukaryotic cells as those contain neither orthologs nor paralogs, which reduces the background activity. As nitrocefin cannot access the cytoplasm, another substrate was needed. Therefore, the cell permeable fluorescent compound CCF2/AM was developed which contains an intrinsic Förster resonance energy transfer (FRET) pair that is coupled by a cephalosporin β -lactam [205]. After hydrolysis of the β -lactam bond by a β -lactamase, the unstable intermediate spontaneously rearranges and the fluorophores separate. Hence, there is no FRET anymore and the donor fluorophore emits the light directly which can be quantified. As the turnover is proportional to the amount of protein, this method was used for protein expression studies [205, 206]. However, for activation the β -lactamase substrate has to be deacetylated after cellular uptake which occurs in eukaryotic cells but not in *E. coli* where a specific esterase is lacking. Missing esterase activity can be overcome by coexpression of the *Fusarium solani pisi* esterase cutinase [207]. This adapted method is advantageous for single cell measurements such as FACS, as the fluorophore is trapped in the periplasmic space after deacetylation. Hence, this extended assay was used for the discovery of new amyloid-beta aggregation inhibitors in *E. coli* [208].

1.5.3 Split β -lactamase complementation assay

The next step of utilizing the TEM-1 β -lactamase in molecular biology was the development of a split variant for protein-protein interaction measurements. The groups of Wehrman and Galarneau developed this protein complementation assay independently in 2002 [152, 153]. In both assays the β -lactamase is split at amino acid Gly₁₉₆/Leu₁₉₈ and generates two inactive fragments. Galarneau added the mutation M182T which stabilizes the enzyme and improves folding but does not influence its activity [193, 209]. Figure 8 B shows a scheme of a class A β -lactamase amino acid sequence with the positions of the enzymatic active amino acids as shown in Figure 8 A, the described M182T mutation and the splitting site. The two fragments show no enzymatic activity as the N-terminal fragment lacks the amino acids Lys₂₃₄ and Ala₂₃₇ (in case of TEM-1) so that the substrate β -lactam cannot be bound and positioned correctly. On the other hand the C-terminal fragment lacks the enzymatic active amino acids Ser₇₀, Lys₇₃, Ser₁₃₀ and Glu₁₆₆, which excludes any enzymatic activity [190]. It was shown that the proteins of interest can be

fused either to the C- or N-terminus of the C-terminal β -lactamase fragment and preferable to the C-terminus of the N-terminal fragment [153].

Hence, this protein complementation assay is suitable for many applications. For example this approach was already used for screening of G-protein-coupled-receptor inhibitors [210] and combined with FACS for toll-like receptor inhibitors [211]. The ability to facilitate ampicillin resistance was used in library screenings for the generation of high affinity TNF- α binding affibodies [212]. Furthermore, it can also be used for in vitro applications such as content mixing assays for liposome fusions to investigate the function of SNARE proteins [213].

2 Motivation

As described in the previous chapter, protein-protein interactions in biological membranes, mediated by TMDs, play an important role in signal transduction, protein folding and assembly. Nevertheless, many interfacial amino acids have not been investigated yet due to the limitation of existing methods. Especially, there is a need for new methods to quantify interactions between pairs of candidate TMDs and to screen for strongly interacting helices.

Hence, the aim of this thesis was the development of a new assay for the measurement of heterotypic TMD-TMD interactions, which could furthermore be adapted for library screenings in the future. All available assays depend on the activation of gene transcription. The new approach should circumvent the problems of these assays by reducing the number of steps between interaction and output signal and enable the measurement of homo- and heterotypic TMD-TMD interactions. To reach this goal, a more direct method, a split protein complementation assay (PCA), was designed. This comprised the use of the well investigated β -lactamase TEM-1, which can be split for PCAs and be colorimetrically quantified [153]. The complete assay was designed as a genetic assay where the TMDs of interest interact in the inner cell membrane of *E. coli* similar to ToxR or GALLEX [77, 108]. This circumvents (i) work intensive protein purification, (ii) increases the throughput, (iii) allows the measurement of mediated ampicillin resistance and (iv) forms the basis for screening set-ups.

Several features, which are necessary for a robust assay, were defined. First, a protein containing a full-length β -lactamase had to be designed which is a template for the following constructs. Second, it had to be proven whether the protein complementation works in the periplasmic space of *E. coli*, which had never been shown before. Therefore, an established soluble dimerization domain was used as a test case to demonstrate the feasibility. Third, a reliable quantification method for reconstituted β -lactamase had to be chosen. Forth, the dimerization has to be reproducibly detected and quantified. Fifth, the complete fusion proteins, containing the split fragments of the β -lactamase and a TMD of choice, have to be detectable as well. For this purpose, the proteins and the assay protocol had to be tested and optimized on the basis of several well investigated TMDs to

achieve a robust, reproducible, reliable and simple new method for the exact measurement of the strength of homo- and heterotypic TMD-TMD interactions.

3 Material and Methods

All buffers were prepared with deionized water (dH₂O) if nothing else is indicated. The compositions of all buffers are described in the corresponding sections. All chemicals were purchased from Applichem (Darmstadt, Germany), Carl Roth (Karlsruhe, Germany) or Sigma-Aldrich (St. Louis, MO, US) and all consumables were purchased from Sarstedt (Nürnbrecht, Germany) if nothing else is indicated.

3.1 Preparation of chemical competent cells

Escherichia coli (*E. coli*) cells were made chemically competent using the protocol developed by Chung et al. [214]. 100 mL preheated LB medium (tryptone and yeast extract purchased from Carl Roth (Karlsruhe, Germany) with an appropriate antibiotic (100 µg/mL streptomycin for JM83, 12.5 µg/mL tetracyclin for *E. coli* XL1-Blue, none for *E. coli* BL21) were inoculated with 1 mL fresh overnight culture of the required *E. coli* strain and incubated in a 250 mL shaking flask at 37 °C and 140 rpm until a OD₆₀₀ of 0.3 was reached. Then the culture was chilled on ice for 10 min and afterwards pelleted for 10 min at 1000 xg and 4 °C (Hermle Z513K centrifuge, Wehingen, Germany). The supernatant was discharged and the pellet resuspended gently in 10 mL ice cold TSS buffer. The cells were aliquoted in 100 µL per tube, immediately frozen in liquid nitrogen and stored at -80 °C.

LB medium (pH 7.0)

1% (w/v)	Tryptone
0.5% (w/v)	Yeast extract
171 mM	NaCl
Autoclaved	

TSS buffer

	LB medium
5% (v/v)	DMSO
10% (w/v)	PEG-3350
50 mM	MgCl ₂
Freshly prepared, sterile filtered	

3.2 Transformation

The chemical competent cells (s. 3.1) were transformed with plasmid DNA by heat shock. A cell aliquot was thawed on ice for 10 min and up to 10 µL total volume DNA was added. After 30 min incubation on ice, the cells were incubated in a water bath at 42 °C for 1 min and then again incubated on ice for 2 min. Afterwards 900 µL LB medium containing

20 mM glucose was added and incubated for 1 h in a turning wheel at 37 °C to allow development of the plasmid coded antibiotic resistance. Subsequently 100 µL cells were plated on LB-agar plates (LB medium with 1.5% agar) containing an appropriate antibiotic (34 µg/mL chloramphenicol (Cm) for N-BLa plasmids, 35 µg/mL kanamycin sulfate (Kan) for C-BLa plasmids).

3.3 Polymerase chain reaction (PCR)

Many molecular methods (s. 0) are strongly dependent on PCR. If not indicated differently, the Phusion High-fidelity DNA Polymerase (cat. no. M0530) from NEB (Ipswich, MA, US) was used. If the PCR product had to be purified by agarose gel purification a 50 µL reaction was prepared, in other respects 25 µL. The general protocol for a 25 µL reaction was as follows.

5 µL	5x Phusion buffer	30 sec	98 °C		Initial denaturation
5 ng	Template DNA	10 sec	98 °C		Denaturation
1.25 µL	Forward primer (10 mM)	30 sec	50-65 °C	30x	Annealing
1.25 µL	Reverse primer (10 mM)	1 min/kb	72 °C		Elongation
0.5 µL	dNTPs (10 mM each)	10 min	72 °C		Final elongation
0 µL - 2.5 µL	DMSO	Hold at	4 °C		
0.25 µL	Phusion polymerase				
Add to 25 µL	dH ₂ O				

The reaction could be optimized by varying the amount of DMSO (standard 3%) and the annealing temperature, depending on primer melting temperature (T_M), which was calculated with the web tool “OligoCalc” (<http://biotools.nubic.northwestern.edu/OligoCalc.html>) [215] using the “Salt Adjusted” T_M (50 mM Na⁺).

3.4 Molecular cloning

In the section molecular cloning all procedures modifying DNA to create new DNA plasmids with new features are described. Depending on the region on the plasmid and the number of modified base pairs different techniques were used. All restriction enzymes, T4 DNA ligase and T4 DNA polynucleotide kinase (PNK) were purchased from Thermo Fisher Scientific (Waltham, MA, US), the Phusion High-fidelity DNA Polymerase from

North England Biolabs (Ipswich, MA, US) and the PfuUltra II Fusion HS DNA Polymerase from Agilent (Santa Clara, CA, US).

All reactions were performed in a thermocycler (Mastercycler or Mastercycler personal from Eppendorf (Hamburg, Germany)). 10 μ L of the product were transformed into chemical competent *E. coli* XL1-Blue plated on a LB-agar plate containing appropriate antibiotic and incubated for 14 h to 20 h at 37 °C until single colonies were visible on the plates. After that the plates could be stored at 4 °C up to several weeks.

3.4.1 Restriction-based cloning

Transfer of DNA regions longer than approximately 200 bp from one plasmid to another was done by restriction based cloning. A linear vector DNA (plasmid backbone) and a linear insert DNA (DNA region from a donor plasmid) with compatible, usually sticky ends were created and mixed for reaction.

In case of compatible restriction sites on the donor and acceptor plasmid, both plasmids (1 μ g) were digested for 1 h with 10 U suitable restriction endonucleases (Thermo Fisher Scientific) in corresponding buffers. If no compatible restriction site was available, the insert was amplified from the donor plasmid by PCR (s. 3.3) using primers that introduce a new restriction site by 5' extension. To ensure high enzyme efficiency all buffer salts from the PCR reaction were removed by column purification (s. 3.6) before endonuclease digestion. The digested samples were isolated by agarose gel purification (s. 3.6) and merged by T4 DNA ligase. 50 ng vector DNA and the threefold molar mass of insert DNA were used.

2 μ L	10x T4 DNA Ligase buffer	20 min	22 °C	
50 ng	Vector DNA	20 min	16 °C	Ligation
$m = \frac{bp(insert)}{bp(vector)} \times 3 \times 50 \text{ ng}$	Insert DNA	20 min	12 °C	
1 U	T4 DNA Ligase	10 min	70 °C	Deactivation
Add to 20 μ L	dH ₂ O	Hold at	4 °C	

3.4.2 Cassette cloning

New TMDs, which were meant to be measured in the BLaTM assay, had to be integrated into the N-BLa and C-BLa cloning vectors, which contain an “Integrin α 5 GP” TMD with an *Apa*I restriction site. To perform cassette cloning the vectors were digested using the

two endonuclease restriction enzymes *NheI* and *BamHI* which cut precisely upstream and downstream of the “Integrin $\alpha 5$ GP” TMD to remove it and thus create two sticky ends (Figure 9, red bases). The new inserts consist of two oligos, whose centers code for the new TMD sequence (Figure 9, black bases, e.g. GpA₁₉_1_wt) and are reverse complement to each other so that they can hybridize. By adding a corresponding set of bases (Figure 9, blue bases) to the ends of the two primers, the same sticky ends as on the digested vector were created on the cassette fragment.

```

amino acid  n r a s I I F G V M A G V I G T I L L I S Y A i h k
sense       5'...TCGAGdCTAGCATTATTTTGGCGTGATGGCGGGCGTGATTGGCACCATTCTGCTGATTAGCTATGCGGdGATCCACA...3'
antisense   3'...AGCTCGATCdgSTAATAAAAACCGCACTACCGCCCGCACTAACCGTGGTAAGACGACTAATCGATACGCCCTAGGTGT...5'
                NheI                                     BamHI

```

Figure 9: Oligo design for cassette cloning. Each cassette consists of one sense and one reverse complementary antisense primer. Amino acids: capital letters: TMD, lowercase letters: juxtamembrane region. Primer: grey bases: vector backbone, red bases: sticky ends of the vector, blue bases: sticky ends of the insert., black bases: coding sequence for the TMD (e. g. GpA₁₉_1_wt). Restriction sites are indicated.

The cassette was created by hybridizing two complementary oligos due to slowly decreasing temperature. Hence, monomeric oligo nucleotides can specifically dimerize to a double-strand DNA cassette.

10 μ L	10x Tango buffer	10 min	90 °C	
1 μ L	Sense oligo (100 mM)	45 sec	-1 °C/cycle	70x
1 μ L	Antisense oligo (100 mM)	Hold at	4 °C	
88 μ L	dH ₂ O			

The cloning vector had to be digested in two steps because of incompatibility of the two restriction enzymes *NheI* and *BamHI*. First 3 μ g DNA were digested with 10 U *NheI* in 50 μ L 1x Tango buffer for 1.5 h at 37 °C. Then 10 U *BamHI* and 5.5 μ L 10x Tango buffer were added to double the buffer concentration. After 1.5 h incubation at 37 °C the linearized vector was isolated by agarose gel purification (s. 3.6). Afterwards the cassette was phosphorylated by T4 DNA polynucleotide kinase (PNK) and ligated into the linearized vector:

2 μ L	10x T4 DNA Ligase buffer	30 min	37 °C	Phosphorylation
1 μ L	ATP (10 mM)	20 min	22 °C	
35 ng	Linearized vector	20 min	16 °C	Ligation
1 μ L	Hybridized cassette	20 min	12 °C	
1 U	T4 DNA Ligase	10 min	70 °C	Deactivation
5 U	PNK	Hold at	4 °C	
Add to 20 μ L	dH ₂ O			

To avoid false positive clones, in the last step remaining cloning vector was removed by digestion using 5 U *ApaI* for at least 30 min at 30 °C.

3.4.3 Q5 site directed mutagenesis

For deleting regions of any length, the insertion of short sequences up to 60 bp or substitution of single amino acids Q5[®] site directed mutagenesis, developed by NEB (Ipswich, MA, US), was used. In this method the complete plasmid was amplified by PCR (s. 3.3) and then circularized. The binding sites of the primers define the region on the plasmid for the plasmid. For deletions the corresponding regions are omitted between the 5' binding sites of the primers. For insertions the corresponding regions are introduced by 5' extensions of the primers. For substitutions the corresponding bases are replaced by 5' base exchanges of the primers.

Following, the success of the PCR was checked by agarose gel electrophoresis (0.7% (w/v) agarose). Then the PCR product was circularized:

2 μ L	10x T4 DNA Ligase buffer	30 min	37 °C	Phosphorylation
1 μ L	ATP (10 mM)	20 min	22 °C	
10 μ L	PCR product	20 min	16 °C	Ligation
1 U	T4 DNA Ligase	20 min	12 °C	
5 U	PNK	10 min	70 °C	Deactivation
Add to 20 μ L	dH ₂ O	Hold at	4 °C	

In the last step before transformation the template DNA was removed by digestion for at least 2 h at 37 °C with 10 U *DpnI* which is specific for methylated DNA and cuts on average every 256 bp.

3.4.4 Quikchange mutagenesis

The Quikchange mutagenesis [216], marketed by Stratagene (La Jolla, CA, US), can be used to mutate sequences between single bases and 60 bases. It is more suitable for shorter insertions or substitutions. In this method two complete reverse complementary oligos were needed, whose centers code for the desired mutation and have downstream and upstream binding sites with annealing temperatures between 55 °C and 60 °C. The PCR will create linear sense and antisense vector DNA fragments, whose 5' region contains the desired mutation. The two fragments can hybridize and form the complete vector with the mutation and two nicked regions. The mutation was inserted by PCR using following conditions:

5 µL	10x PfuUltraII buffer	2 min	95 °C	
20 ng	Template DNA	20 sec	95 °C	
125 ng	Sense oligo	20 sec	58 °C	16x - 18x
125 ng	Antisense oligo	5 min	72 °C	
1.5 µL	DMSO	10 min	72 °C	
1 µL	dNTP (10 mM)	Hold at	4 °C	
1 µL	PfuUltraII DNA Polymerase			
Add to 50 µL	dH ₂ O			

In the last step before transformation the template DNA was removed by digestion for at least 2 h at 37 °C with 10 U *DpnI* which is specific for methylated DNA and cuts on average every 256 bp. After transformation in *E. coli* the cells will repair these regions and the desired mutation is inserted.

3.4.5 Transfer-PCR

Transfer-PCR (TPCR) is an alternative method to restriction based cloning to transfer large DNA regions from one vector to another without using endonucleases [217]. It combines PCR (s. 3.3) with the Quikchange mutagenesis (s. 3.4.4). The oligos consist of two parts: the 3' parts are binding with an annealing temperature of about 55 °C to the donor plasmid and amplify the desired new sequence. The 5' ends are binding with an annealing temperature of about 65 °C to the acceptor plasmid and enable the new DNA fragment to insert site specific into the acceptor plasmid. A two-step protocol was used in this thesis as it has a higher efficacy than a one-step protocol. Thus first a megaprimer was produced by PCR (s. 3.3) containing the new sequence and extensions on both sites for targeting the

correct position on the acceptor plasmid. Then the megaprimer was purified via agarose gel electrophoresis (s. 3.6) and transferred to the acceptor plasmid following the Quikchange protocol (s. 3.4.4).

10 μ L	5x Phusion buffer	30 sec	98 °C	
10 ng	Vector DNA	10 sec	98 °C	
50 ng	Megaprimer DNA	30 sec	68 °C	18x
1 μ L	dNTPs (10 mM)	6 min	72 °C	
1.5 μ L	DMSO	10 min	72 °C	
1 μ L	Phusion polymerase	Hold at	4 °C	
Add to 50 μ L	dH ₂ O			

In the last step before transformation the template DNA was removed by digestion for at least 2 h at 37 °C with 10 U *DpnI* which is specific for methylated DNA and cuts on average every 256 bp.

3.5 Plasmid propagation

The success of every cloning had to be checked. Therefore, single colonies were picked from LB agar plates with a sterile 200 μ L tip and transferred into a glass tube with 8 mL LB medium containing an appropriate antibiotic. Alternatively overnight cultures were inoculated with fresh transformations directly if homogenous plasmid solution was transformed.

For the preparation of plasmid DNA the NucleoSpin® Plasmid Purification Kit from Macherey-Nagel was used accordingly to the manufacture's protocol for high-copy plasmids (p_{BAD} variants) or for low-copy plasmids (N-BLa and C-BLa plasmids), respectively. All optional washing and heating steps were conducted.

3.6 DNA purification

Depending on the cloning method, DNA fragments had to be separately isolated or purified from buffer salts to ensure high enzyme activity.

If a fragment had to be isolated from a DNA mixture, the sample was separated by agarose gel electrophoresis and stained with ethidium bromide (s. 3.8). The band containing the desired DNA fragment was cut under UV light ($\lambda = 365$ nm) and transferred into a 1.5 mL tube. For the isolation of the DNA from the agarose the NucleoSpin® Gel and PCR clean-

up kit from Macherey-Nagel was used accordingly to the manufacture's protocol. All optional washing and heating steps were conducted.

If no separation but only a desalting was necessary, the sample was loaded directly onto the column according to the manufacture's protocol of the NucleoSpin® Gel and PCR clean-up kit from Macherey-Nagel. All optional washing and heating steps were conducted.

3.7 DNA quantification

The concentration of prepared DNA was determined by measuring the extinction at 260 nm and 280 nm in a quartz cuvette using an Ultrospec 3100*pro* photometer (Amersham Biosciences, Amersham, UK). The DNA was diluted 1:40 in dH₂O and the extinction was measured against a dH₂O blank. The quotient of 260 nm and 280 nm should be around 1.8 for pure DNA. Higher values are indicating RNA contaminations, lower values remaining protein. Values up to 2.0 were accepted for further experiments.

3.8 Agarose gel electrophoresis

Agarose gel electrophoresis was used for DNA purification, quality control of PCR reactions and the analysis of control digestions of cloning products. The used buffer system was Tris-acetate/EDTA (TAE). The agarose concentration defines the optimal separation range. If the desired fragment was shorter than 1000 bp 1.5% agarose was used, if it was longer than 5000 bp 0.7% agarose and otherwise 1% agarose.

TAE buffer (pH 8.0)

40 mM	Tris
20 mM	Acetic acid
1 mM	EDTA

Stored at room temperature

50 mL TAE buffer with the desired amount of agarose low EEO (Appllichem, Darmstadt, Germany) was heated in the microwave at maximum power for 100 sec and afterwards the evaporated volume was filled-up with dH₂O. The complete solution was filled in a combined casting and electrophoresis system and 0.03 µg/mL ethidium bromide was added to visualize the DNA after the beforehand separation. A comb was inserted to form pockets for the samples and after cooling down to room temperature the gel could be used. The

chamber was filled with TAE buffer and the comb and the gel-casting gates were removed carefully. The DNA samples were mixed with DNA Gel Loading Dye (6X) (Thermo Fisher Scientific) and loaded into the pockets. Additionally 6 μ L GeneRuler 1 kb DNA Ladder (Thermo Fisher Scientific) or 6 μ L GeneRuler 100 bp DNA Ladder (Thermo Fisher Scientific) was loaded, depending on the investigated samples. The fragments were separated at 70 V for 50 min and following visualized under UV light ($\lambda = 312$ nm).

3.9 gDNA extraction from *Saccharomyces cerevisiae*

A few mg baker yeast (Weininger Hefe, Rewe) was resuspended in 25 μ L 20 mM NaOH and incubated at 95 °C for 2 min. To remove cell debris and intact cells the suspension was centrifuged for 5 min at 13000 rpm (Heraeus Biofuge fresco, Thermo Fisher Scientific, Waltham, MA, US). 1 μ L supernatant was used as template DNA for the PCR reaction (s. 3.3).

3.10 DNA Sequencing

DNA plasmids were either sequenced by GATC Biotech (Konstanz, Germany) according to their protocol or in-house using the DNA gel sequencer LONG READIR 4200 (LI-COR Biosciences, Lincoln, NE, US). In the following in-house protocol for each template four sequencing reactions (one different ddNTP each) with a fluorescent dye labeled primer (IRD-700 or IRD-800) were conducted.

Sample preparation

For each sample one mastermix was pipetted. The Tth inorganic pyrophosphatase was purchased from Genecraft (Köln, Germany), the Taq polymerase was expressed and purified in-house. If the plasmid DNA concentration was below 120 ng/ μ L, 8.2 μ L DNA solution was used instead of 1 μ g.

3.5x Sequencing buffer (pH 9.0)		Storage buffer (for enzyme dilution)	
175 mM	KCl	10 mM	K ₃ PO ₄ , pH 7.0
35 mM	Tris	100 mM	NaCl
0.35%	Triton X-100	0.5 mM	EDTA
12.3 mM	MgCl ₂	1 mM	DTT
Stored at 4 °C		0.01%	Tween 20
Pyrophosphatase dilution (0.5 U/μL)		50% (v/v)	Glycerol
Storage buffer		Stored at 4 °C	
10% (v/v)	Tth inorganic Pyrophosphatase (5 U/μL)	Stop/loading buffer	
Stored at -20 °C		95% (v/v)	Formamide
Master mix		10 mM	EDTA, pH 9.0
7.8 μL	3.5x Sequencing buffer	0.1% (w/v)	Basic fuchsin
1 μg	Template DNA	0.01% (w/v)	Bromophenol blue
1.1 μL	Labeled primer (2 μM)	Stored at -20 °C	
1.1 μL	Pyrophosphatase dilution		
0.3 μL	Taq polymerase		
Add to 18.5 μL	dH ₂ O		

For the sequencing reaction 2 μL of each nucleotide mix was pipetted into one PCR tube. Last 4 μL master mix were added to each tube, gently mixed and the sequencing reaction was started. Before loading on the sequencing gel 3 μL stop/loading buffer were pipetted to each sample and incubated for 4 min at 95 °C. If the samples were not sequenced immediately they were stored at -20 °C.

Nucleotide mixes, store at -20 °C

A		C		G		T	
15 μM	dNTP	15 μM	dNTP	15 μM	dNTP	15 μM	dNTP
0.21 μM	ddATP	0.21 μM	ddCTP	0.38 μM	ddGTP	0.38 μM	ddTTP

Sequencing protocol

2 min	95 °C	
20 s	95 °C	
20 s	57 °C	30x
1 min	70 °C	
1 min	70 °C	
Hold at	4 °C	

Sequencing gel preparation

All polyacrylamide sequencing gel components were combined except APS and TEMED. Glass plates were washed with 10% (w/v) SDS and isopropanol. Plates, spacers and clamps were assembled. APS and TEMED were added to the sequencing gel mix and the solution was filled between the glass plates with a syringe immediately. Then the comb was inserted with its smooth side, the buffer place holder was fixed on top and the screws were tightened. After polymerization the buffer place holder and the comb were removed, the gel assembled to the sequencing machine and the comb inserted. The two buffer tanks were filled with 1x TBE long run buffer.

Sequencing gel		1x TBE long run buffer (pH 8.3-8.7)	
5 mL	40% Acrylamide-Bisacrylamide 32:1	134 mM	Tris
10.5 g	Urea	45 mM	Boric acid
2.5 mL	10x TBE long run buffer	2.5 mM	EDTA
3.75 mL	Formamide	Stored at room temperature	
Add to 25 mL	dH ₂ O		
25 µl	TEMED		
175 µl	10% (w/v) APS		

Electrophoresis

0.5 µL of each reaction was loaded onto the gel. During electrophoresis (1200 V, 37 mA, 40 W, 50 °C) the fluorescence was recorded. The sequencing pattern was analyzed with Base ImageIR Image Analysis 4.0 software (LI-COR Biosciences, Lincoln, NE, US) and evaluated with CLC Main Workbench 6.9.1 (CLC Bio, Aarhus, Denmark).

3.11 GFP expression test

As all β -lactamase hybrid protein constructs contain a sfGFP at their C-terminus, the cloning was tested for frameshifts by verifying GFP expression. Emission at 520 nm after

excitation at 485 nm indicated expression of GFP and thus no frameshift mutation in the cassette cloning procedure (s. 3.4.2). For the test of BLaTM 1.1 plasmids two 200 μ L aliquots of each overnight culture were prepared – one with 1.33 mM arabinose and another without inducer. Both samples were incubated for 2 h at 37 °C in the turning wheel. 10 μ L of each culture were diluted with 90 μ L PBS in a black 96-well plate (Nunc, Roskilde, Denmark) and the fluorescence intensity was measured ($\lambda_{\text{Ex}} = 485$ nm, $\lambda_{\text{Em}} = 520$ nm, PolarStar, BMG Labtech, Ortenberg, Germany). If the induced sample showed an at least three times higher fluorescence than the control, a correct expression and thus the absence of a frameshift was assumed. Because of lower expression levels of BLaTM 1.2 proteins, the overnight cultures had to be diluted 1:10 in LB-medium to ensure higher expression. Again, one sample of each culture was induced with 1.33 mM arabinose and another one used as control. After 4 h expression at 37 °C 200 μ L cells were centrifuged for 2 min at 11000 rpm (Biofuge fresco, Heraeus) and resuspended in the same volume low-fluorescent PBS (flow cytometry grade, Thermo Fisher Scientific). Then 100 μ L cell suspension was pipetted into a black 96-well plate (Nunc, Roskilde, Denmark) and the fluorescence intensity was measured ($\lambda_{\text{Ex}} = 485$ nm, $\lambda_{\text{Em}} = 520$ nm, PolarStar, BMG Labtech, Ortenberg, Germany). If the induced sample showed an at least twofold higher fluorescence than the control, a correct open reading frame was assumed.

For quantitative GFP expression measurements of the BLaTM 1.2 constructs the expression was monitored directly in 12-well plates (Greiner Bio-one, Kremsmünster, Austria). Corresponding to the BLaTM 1.2 protocol (s. 3.15.3) 2 mL of medium were inoculated 1:10 with an overnight culture in the wells and induced with appropriate amounts of arabinose and IPTG (133 μ M arabinose and 0.1 – 0.7 mM IPTG). After 4 h expression the fluorescence intensity ($\lambda_{\text{Ex}} = 485$ nm, $\lambda_{\text{Em}} = 520$ nm) and the absorption at 544 nm was measured in the plates in a microplate reader (FluoStar, BMG Labtech, Ortenberg, Germany). For analysis the background fluorescence (mean of three non-induced samples) was subtracted from all values. The corrected values were normalized to the cell density by dividing the values by A_{544} .

3.12 β -Lactamase activity test

The activity of β -lactamase was tested by the enzymatic cleavage of the β -lactam ring of the chromogenic cyclosporine derivate nitrocefin. After cleavage its absorption maximum switches from 390 nm to 492 nm [153, 194, 199]. The nitrocefin (Merck Millipore,

Darmstadt) was dissolved 10 mM in DMSO and stored at -20 °C. An overnight culture of JM83 cells containing the desired plasmids was diluted 1:10 in fresh LB medium, induced with 0.2% arabinose and incubated for another 4 h at 37 °C. 200 µL cells were centrifuged (2 min, 7000 rpm) and resuspended in 200 µL sodium phosphate buffer (100 mM PO₄³⁻, pH 7.0). 20 µL of the resuspended *E. coli* culture, 100 µL sodium phosphate buffer/EDTA (100 mM Na₃PO₄, pH 7.0, 2 mM EDTA), 78 µL water and 2 µL nitrocefin (final concentration 100 µM) were mixed. Then the β-lactamase activity was determined over 20 min by measuring the extinction increase at 492 nm in a 96-well plate in microplate reader (VersaMmax, Molecular Devices, Biberach an der Riss).

3.13 SDS polyacrylamide gel electrophoresis

Proteins were separated depending on their molecular weight by the discontinuous sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) [218]. The combined casting and running system for gels (10 cm x 10 cm x 0.8 mm) “PerfectBlue Dual Gel Twin S” from PEQLAB (Erlangen, Germany) was used.

Gel preparation

The glass plates were cleaned with 70% ethanol, separated with one spacer on each side and fixed with the clamps, notched plate facing to the middle, to the chamber of the electrophoresis device. The position was defined by the gel casting base. The base was turned around so that the gaskets were facing up and the chamber was fixed to seal the glass plates. All resolving gel components were combined except ammonium persulfate (APS) and tetramethylethylenediamine (TEMED). To start the polymerization APS and TEMED were added and the solution was pipetted between the plates immediately to height of about 6 cm. Then the gel was overlaid with isopropanol to protect it from aerial oxygen which inhibits polymerization. 30 min later the gel was completely polymerized and the isopropanol was removed completely. All stacking gel components were combined except APS and TEMED. To start the polymerization APS and TEMED were added and the solution was pipetted between the plates immediately to the bottom of the notch. To create pockets for the samples a comb was inserted. After about 30 min the stacking gel was polymerized too and the gel was ready for use. It could be stored for several days at 4 °C.

Stacking gel buffer (pH 6.8)

250 mM	Tris
0.2% (w/v)	SDS
Stored at room temperature	

Resolving gel buffer (pH 8.5)

1 M	Tris
0.26% (w/v)	SDS
Stored at room temperature	

Stacking gel

1.6 mL	Stacking gel buffer
0.4 mL	30% Acrylamid- bisacrylamid (37.5:1)
1.2 mL	dH ₂ O
25 µL	APS (10% (w/v))
2.5 µL	TEMED

12.5% Resolving gel

2.8 mL	Resolving gel buffer
3.0 mL	30% Acrylamid- bisacrylamid (37.5:1)
1.5 mL	dH ₂ O
50 µL	APS (10% (w/v))
2.5 µL	TEMED

Sample preparation and electrophoresis

The *E. coli* culture samples had to be concentrated and denatured. For concentration 0.5 OD₆₀₀ cell sample was centrifuged for 1 min at 6000 rpm (Biofuge fresco, Heraeus), the supernatant discarded and the pellet resuspended in 20 µL reducing 1x Laemmli loading buffer. Then it was denatured for 5 min at 95 °C.

The gel casting base was removed and if only one gel was used the other site of the chamber was sealed with a blocking plate. The chamber was placed in the buffer tank and the chamber and the tank were filled with Laemmli buffer. The comb was removed carefully and all pockets were washed with buffer. The samples and 4 µL PageRuler™ Prestained Protein Ladder (Thermo Fischer Scientific) were loaded. The proteins were separated for 80 min at 200 V.

Laemmli buffer

20 mM	Tris
192 mM	Glycine
0.1% (w/v)	SDS
Stored at room temperature	

5x Laemmli loading buffer (pH 6.8)

50 mM	Tris
5% (w/v)	SDS
20% (v/v)	Glycerol
0.02% (w/v)	Bromphenol blue
10% (v/v)	β-Mercaptoethanol
Stored at 4 °C	

3.14 Western blot

For the specific detection of proteins, separated by SDS-PAGE (s. 3.13), semi-dry Western blotting was used. The colorimetric detection was carried out after a two-step immuno staining of the proteins. Alkaline phosphatase, fused to a secondary antibody, hydrolyzes

5-bromo-4-chloro-3-indolyl phosphate (BCIP). The product 5-bromo-4-chloro-3-indoxyl can be oxidized by nitro blue tetrazolium (NBT) to the blue dye 5,5'-dibromo-4,4'-dichloro-indigo, which is insoluble. The twice reduced NBT forms an insoluble, blue diformazan dye. Both products precipitate on the Western blot nitrocellulose membrane where protein was stained by the antibodies.

Protein transfer

For an efficient transfer of proteins from SDS-PAGE gel to a Western blot nitrocellulose membrane, it was essential that all layers were soaked with blotting buffer and put air bubble-free one on the other. Four filter papers (Munktell & Filtrak, Bärenstein, Germany) were placed on the middle of the anode of the blotting device (Modell SD 1, cti, Idstein, Germany). The stacking gel of the SDS-PAGE gel was removed and the resolving gel with the separated proteins was placed on the filter papers. Then a nitrocellulose blotting membrane (Berrytec, Grünwald, Germany), activated for at least 1 min in blotting buffer, was put on the gel, followed by another four layers of soaked filter papers. Following the cathode was placed on the transfer stack and weighted. The proteins were transferred to the nitrocellulose membrane for 1.5 h at 70 mA (1 mA/cm²).

Protein detection

After that, all proteins on the blotting membrane were stained for 1 min with PonceauS solution to check for transfer efficiency. Then the membrane was washed with water to remove all unbound PonceauS. Afterwards the stained membrane was documented and the bound PonceauS was stripped by washing with alkalized water. Then the membrane was blocked with 10 mL 3% (w/v) skimmed milk powder in TBS for 1 h at room temperature or overnight at 4 °C. Next the membrane was washed for 5 min with TBS and incubated for 1 h at room temperature with 10 mL 3% (w/v) skimmed milk powder in TBS containing 0.01% (v/v) Monoclonal ANTI-FLAG M2 antibody (Sigma-Aldrich, St. Louis, MO, US) as a primary antibody. The membrane was washed three times for 5 min with TBS-T and following incubated for 1 h at room temperature with 10 mL 3% (w/v) skimmed milk powder in TBS containing 0.01% (v/v) Anti-Mouse IgG AP Conjugate (Promega, Madison, WI, US) as a secondary antibody. Before detection the unbound antibodies were removed by another three 5 min washing steps with TBS-T. The secondary antibodies, and so the FLAG tagged proteins, were detected with the NBT/BCIP

solution until bands were clearly visible. To stop the reaction the membrane was rinsed several times with water. For documentation the membrane was dried at room temperature and scanned.

Blotting Buffer

	Laemmli buffer
20% (v/v)	Methanol
Stored at room temperature	

PonceauS solution

3% (w/v)	Trichloroacetic acid
0.3% (w/v)	PonceauS
Stored at room temperature	

TBS (pH 7.4)

20 mM	Tris
150 mM	NaCl
Stored at room temperature	

TBS-T (pH 7.4)

TBS	
0.5% (v/v)	Tween 20
Stored at room temperature	

BCIP solution

	Dimethylformamid
5% (w/v)	BCIP
Stored at -20 °C	

NBT solution

70% (v/v)	Dimethylformamid
5% (w/v)	NBT
Stored at -20 °C	

AP buffer (pH 9.5)

100 mM	Tris
100 mM	NaCl
5 mM	MgCl ₂
Stored at room temperature	

Staining solution

20 mL	AP buffer
120 µL	NBT solution
60 µL	BCIP solution
Freshly prepared	

3.15 BLaTM Assay

3.15.1 Choose of vectors

As there are two different proteins involved in the assay, two plasmids with several important requirements were needed. (i) First the vectors have to be compatible to each other. That is to say they have to use two different replication systems. Otherwise one of the plasmids would get lost during the growth of *E. coli* and after the complete degradation of the antibiotics. (ii) For selection the plasmids have to contain two different antibiotic resistance genes. (iii) Furthermore, the plasmids should have a similar copy number to achieve an equal protein expression. (iv) Low copy plasmids are preferred to accomplish a low protein expression, which is essential for membrane protein interaction assays. (v) The protein expression has to be inducible. That means it can be induced at a defined time point and the strength can be controlled by the amount of inductor. (vi) The vectors have to be

compatible with the cassette cloning technique (s. 3.4.2) which forbids *Bam*HI and *Nhe*I restriction sites on the plasmid backbone.

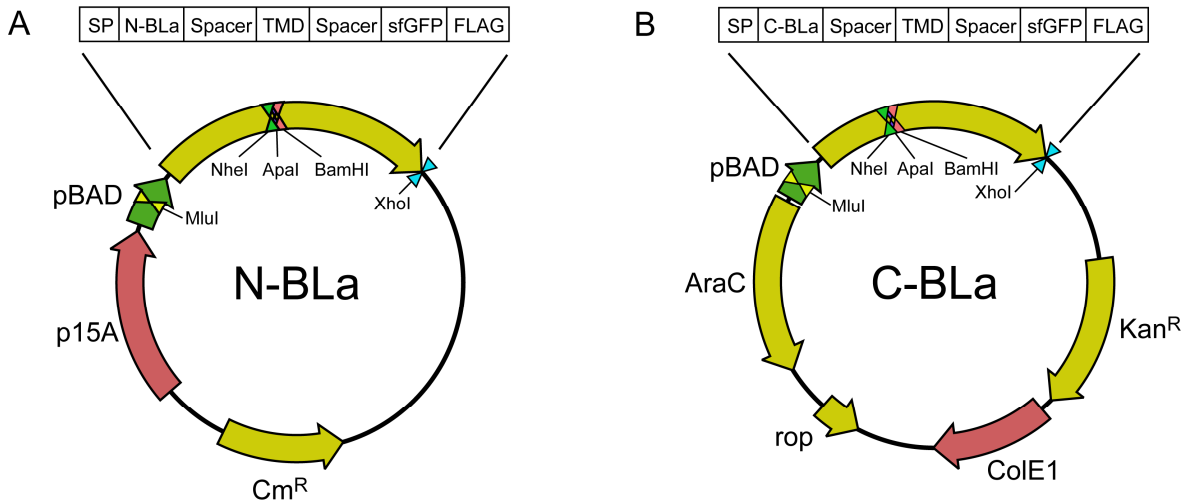


Figure 10: Illustration of the final vectors containing the BLaTM proteins. Yellow: open reading frames, red: origin of replication, green: operon region. Important restriction sites are indicated. BLaTM proteins (s. 3.15.2) contain a signal peptide (SP), the N- or C-terminal fragment of the split β -lactamase (N-BLa, C-BLa), a spacer between the fragment and the transmembrane domain (TMD), the TMD of interest, a spacer between the TMD and the sfGFP, a sfGFP and a FLAG-epitope for detection. **A)** N-BLa: pToxRVII [219] based vector containing a p15A origin of replication, a chloramphenicol acetyltransferase gene (Cm^R , chloramphenicol resistance) and a p_{BAD} operator. **B)** C-BLa: pBAD322K [220] based vector containing a pMB1 origin of replication, a *rop* gene (copy number repression), an aminoglycoside O-phosphotransferase gene (Kan^R , kanamycin resistance), an *AraC* gene (regulator protein of the p_{BAD} operon) and a p_{BAD} operator.

The first chosen vector was the pToxRVII vector (Figure 10 A), created by Eric Lindner [219]. The backbone is the pACAY184 plasmid containing the multiple cloning site and the T7 terminator region of the pET22b plasmid. It has a low copy p15A origin of replication [221], a chloramphenicol resistance gene (Cm^R , chloramphenicol acetyltransferase), a p_{BAD} operator [222] and is compatible with the cassette cloning technique.

The second vector was the pBAD322K vector (Figure 10 B), created and provided by John Cronan [220]. It is a pBAD24 vector with an inserted *rop* gene [223, 224] to reduce the copy number and has a pMB1 origin of replication [225] which is compatible to the p15A origin of replication [130]. Furthermore, it contains the p_{BAD} operator, the corresponding *AraC* gene and a kanamycin resistance gene (Kan^R , aminoglycoside O-phosphotransferase). As it contained two additional *Nhe*I restriction sites in a non-coding region, these had to be removed first by Q5 mutagenesis (primers: pBAD322_fwd, pBAD322_rev; 55 °C, 3% DMSO), resulting in the pBAD322K_Δ plasmid. The modified version is compatible to the cassette cloning technique.

3.15.2 Vector construction

This section describes the construction of the plasmids N-BLa 0.1, C-BLa 0.1, N-BLa 1.1, C-BLa 1.1, N-BLa 1.2, C-BLa 1.2 and their variants. Figure 11 gives an overview of the different cloning steps which are explained in detail in the following chapters. For the BLa 1.1 plasmids the flexible linker length is indicated in subscript (5, 9, 13 or 25) whereas BLa 1.2 plasmids always contain a flexible 13 amino acids linker. The TMD of all cloning vectors is “Integrin- α 5-GP” (UniProtKB [122] accession number: P06756; amino acids 996 – 1008, L1002G, A1003P), which contains an *Apa*I restriction site for cloning control.

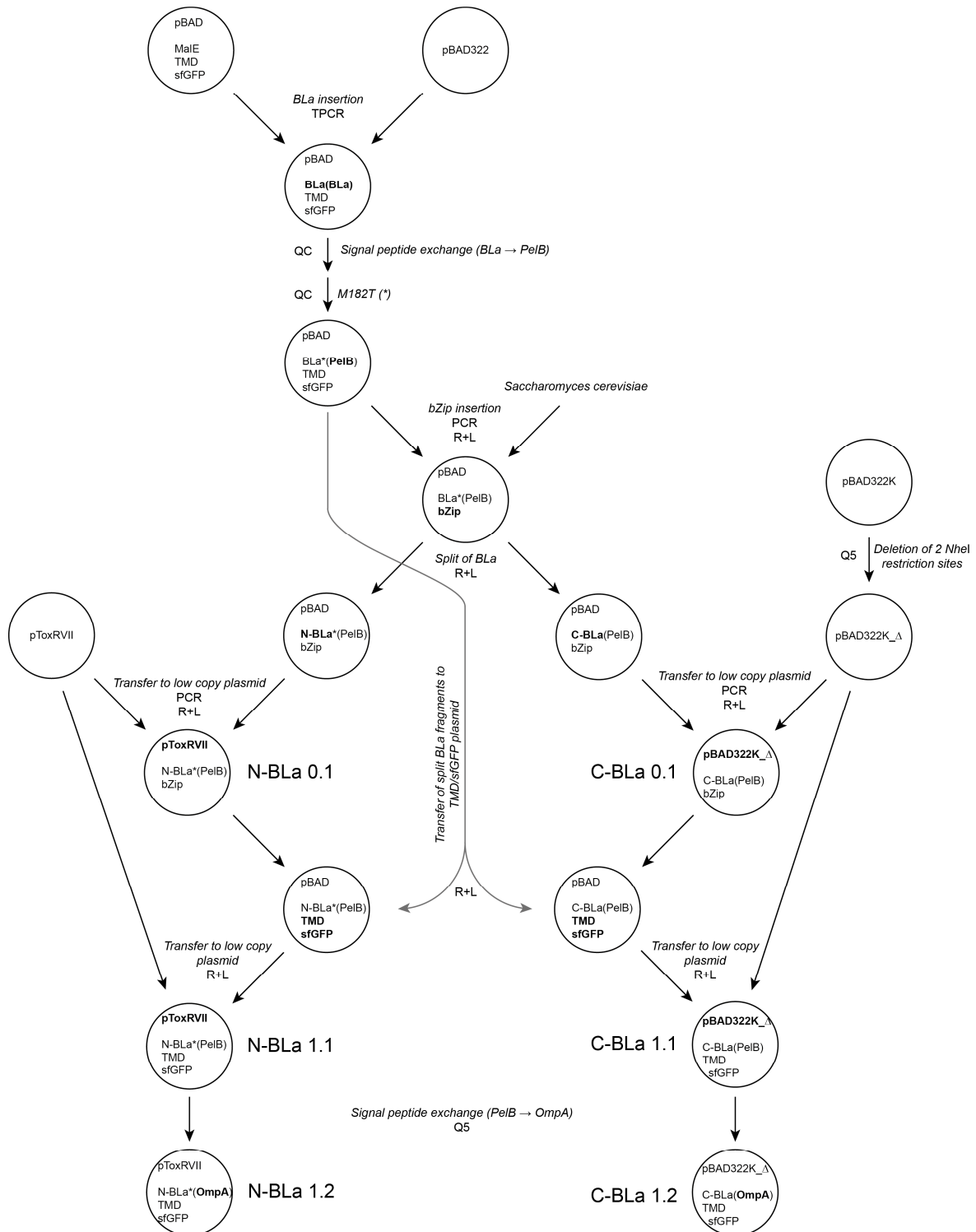


Figure 11: Overview of the vector construction. Vector backbone and segments of fusion proteins are indicated. Modified parts in bold. MalE: maltose binding protein, BLa: full-length TEM-1 β -lactamase (residues 23 – 286), N-BLa: N-terminal fragment of TEM-1 β -lactamase (residues 23 – 194), C-BLa: C-terminal fragment of TEM-1 β -lactamase (residues 196 – 286), *: β -lactamase M182T mutation, TMD: Integrin- α 5-GP, sfGFP: superfolder GFP, bZip: basic leucine-zipper of GCN4 (residues 235 – 281), signal peptide in brackets (BLa: TEM-1 β -lactamase (residues 1 – 22); PelB: pectate lyase B (residues 1 – 20); OmpA: outer membrane protein A (residues 1 – 23)). Methods: TPCR: transfer-PCR, QC: Quikchange mutagenesis, Q5: Q5-mutagenesis, PCR: insert amplification by PCR, R+L: restriction and ligation based cloning.

Construction of BLaTM full length protein

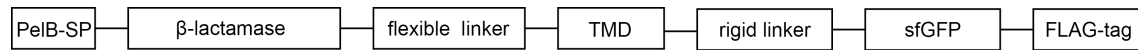


Figure 12: Schema of BLaTM hybrid protein

The primary constructs are based on the $p_{BAD}::malE$ -TMD-sfGFP expression vector designed by Christian Ried [226-228]. The most important modifications were performed in this high copy vector. The MalE and the first rigid linker were replaced by the full length β -lactamase and a five amino acids flexible linker by TPCR (template: pBAD322, primers: MP_AmpR_fwd, MP_AmpR_rev). The M182T mutation was introduced by Quikchange mutagenesis to improve the stability of the enzyme (BLa_M182T_s, BLa_M182T_as; 12 cycles) [153, 193]. The original β -lactamase signal peptide was replaced by a PelB signal peptide (UniProtKB [122] accession number Q00205) by another Quikchange mutagenesis (BLa_PelB_s, BLa_PelB_as; 18 cycles) to achieve a good membrane insertion. This plasmid was termed $p_{BAD}::bla(PelB)M182T$ -5x-TMD-sfGFP.

Construction of BLa-bZIP full length protein

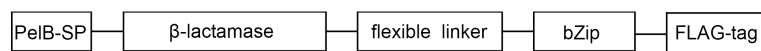


Figure 13: Schema of BLa-bZIP hybrid protein

For the proof-of-principle a construct containing the homodimerizing basic leucine-zipper domain (bZIP, residues 235–281) of the leucine-zipper GCN4 (UniProtKB [122] accession number P03069) [229] was created. The leucine-zipper was amplified from the gDNA of *Saccharomyces cerevisiae* by PCR (primers: GCN4_fwd_15x_speI, GCN4_rev_ft_PstI; 0% DMSO; protocol see below) and replaced the “5x-TMD-sfGFP” fragment of $p_{BAD}::bla(PelB)M182T$ -5x-TMD-sfGFP via the restriction sites *SpeI* and *PstI*. This plasmid was termed $p_{BAD}::bla(PelB)M182T$ -15x- bZIP.

3 min	95 °C	
30 sec	95 °C	
20 sec	40 °C	5x
+ 0.5 °C/sec		
1 min	72 °C	
10 sec	95 °C	
30 sec	55 °C	30x
1 min	72 °C	
5 min	72 °C	
Hold at	4 °C	

Construction of BLa-bZIP split proteins (N-BLa 0.1 and C-BLa 0.1)

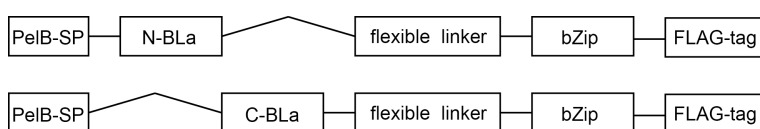


Figure 14: Scheme of BLa-bZIP split proteins (N-BLa 0.1 and C-BLa 0.1)

For the final constructs the β -lactamase was split to an N-terminal fragment (residues A23-G194) attached to the PelB signal peptide from *Aspergillus niger* (Pectin lyase B, UniProtKB [122] accession number Q00205, residues 1 - 20) and a C-terminal fragment (residues L196-W286) attached to the same signal peptide [153]. First the split BLa-bZIP vectors (N-BLa 0.1 and C-BLa 0.1) were constructed by deleting the needless parts of $p_{BAD}::bla(PelB)M182T-15x-bZIP$ by Q5 mutagenesis resulting in the vector $p_{BAD}::N-bla(PelB)M182T-15x-bZIP$ and the vector $p_{BAD}::C-bla(PelB)-15x-bZIP$ (N-*bla* primers: BLa_194G_rev, BLa_speI-15x_fwd, 60 °C, 3% DMSO; C-*bla*: primers: BLa_PelB+H_rev, BLa_196L_fwd, 56 °C, 3% DMSO). The N-BLa containing insert had to be amplified by PCR first to add one restriction site (pToxRV_fwd, pToxRV_rev_XhoI, 55 °C, 3% DMSO). Subsequently it was inserted into the pToxRVII vector by restriction based cloning using the restriction sites *Mlu*I and *Xho*I. This plasmid was termed N-BLa_0.1. Next the C-BLa containing insert was transferred from $p_{BAD}::C-bla(PelB)-15x-bZIP$ vector to the final vectors pBAD322K by restriction based cloning using the restriction sites *Mlu*I and *Pst*I. This plasmid was termed C-BLa_0.1.

Construction of BLaTM 1.1 split proteins

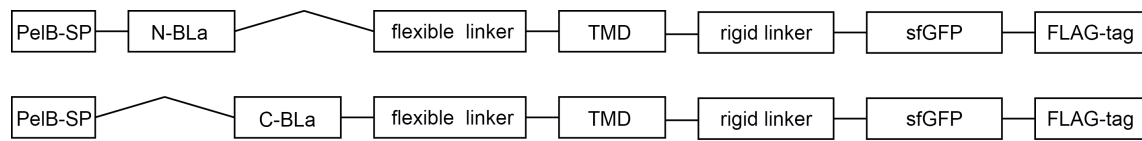


Figure 15: Scheme of BLaTM 1.1 split proteins.

First the split regions were transferred from the $p_{BAD}::N\text{-}bla(\text{PelB})M182T\text{-}15\text{x-bZIP}$ and $p_{BAD}::C\text{-}bla(\text{PelB})\text{-}15\text{x-bZIP}$ vectors to the $p_{BAD}::bla(\text{PelB})M182T\text{-}5\text{x-TMD-sfGFP}$ vector by restriction based cloning using the restriction sites *MluI* and *SpeI*. This led to the vectors $p_{BAD}::N\text{-}bla(\text{PelB})M182T\text{-}5\text{x-TMD-sfGFP}$ and $p_{BAD}::C\text{-}bla(\text{PelB})\text{-}5\text{x-TMD-sfGFP}$. Following they were transferred to the target vectors $p_{ToxRVII}$ and $p_{BAD322K_}\Delta$ using the same procedure as for the creation of the N-BLa 0.1 and C-BLa 0.1 plasmids. These plasmids were termed N-BLa 1.1₅ and C-BLa 1.1₅ and were used in the BLaTM 1.1 assay. The index indicates the length of the linker between the β -lactamase fragments and the TMD. (For DNA and protein sequences of N-BLA 1.1₁₃ and C-BLA 1.1₁₃ see appendix).

Construction of different linker lengths

For assay optimization the flexible linker between the β -lactamase fragments and the TMD (Figure 15) was extended by Q5 mutagenesis (Table 1). Here the linker was elongated by one, two or four GGGS segments resulting in 9, 13 or 25 amino acids linkers, respectively (Figure 25 A). For the BLa 1.1₉ constructs both Q5 mutagenesis worked (template: BLa 1.1₅, 58 °C, 3% DMSO), but for the BLa 1.1₁₃ and BLa 1.1₂₅ construct, which was a side product of a C-BLa 1.1₁₃ PCR reaction, only the C-BLa PCR reactions worked (template: C-BLa_1.1₅, 55 °C, 3% DMSO). So the extended linkers had to be transferred from the C-BLa 1.1 vectors to the N-BLa 1.1 vectors by restriction based cloning using the restriction sites *SpeI* and *XhoI*.

Table 1: Q5 primers for linker extensions. The C-BLa 1.1₂₅ was a side product of a C-BLa 1.1₁₃ PCR reaction.

linker length	N-BLa 1.1	C-BLa 1.1
9 aa	BLa_Spacer_fwd N-BLa_Spacer_GGGS_rev	BLa_Spacer_fwd C-BLa_Spacer_GGGS_rev
13 aa	transfer from C-BLa_1.1 ₁₃	BLa_Spacer_fwd C-BLa_2xGGGS_rev
25 aa	transfer from C-BLa_1.1 ₂₅	BLa_Spacer_fwd C-BLa_2xGGGS_rev

Construction of BLaTM 1.2 split proteins

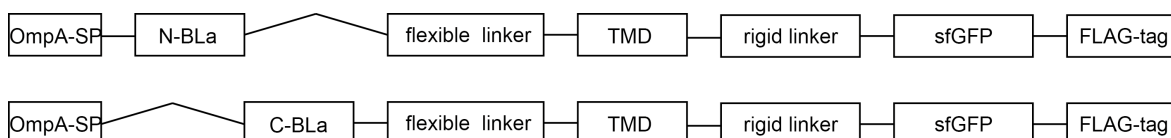


Figure 16: Scheme of BLaTM 1.2 split proteins.

To obtain less inclusion bodies and a better membrane insertion of the proteins in BLaTM 1.2, the signal peptide PelB of the BLa 1.1 proteins was changed to the one of OmpA (UniProtKB [122] accession number P0A910). The exchange was conducted by Q5 mutagenesis with the plasmids N-BLa 1.1₁₃ and C-BLa 1.1₁₃ as templates. The used primers for the N-BLa vector were BLa_delta_SP_fwd and OmpA_rev (3% DMSO, 55 °C), for the C-BLa vector BLa_196L_fwd and OmpA_rev (3% DMSO, 55 °C). These plasmids were termed N-BLa 1.2 and C-BLa 1.2 and were used in the BLaTM 1.2 assay. (For DNA and protein sequences see appendix).

3.15.3 Assay protocol

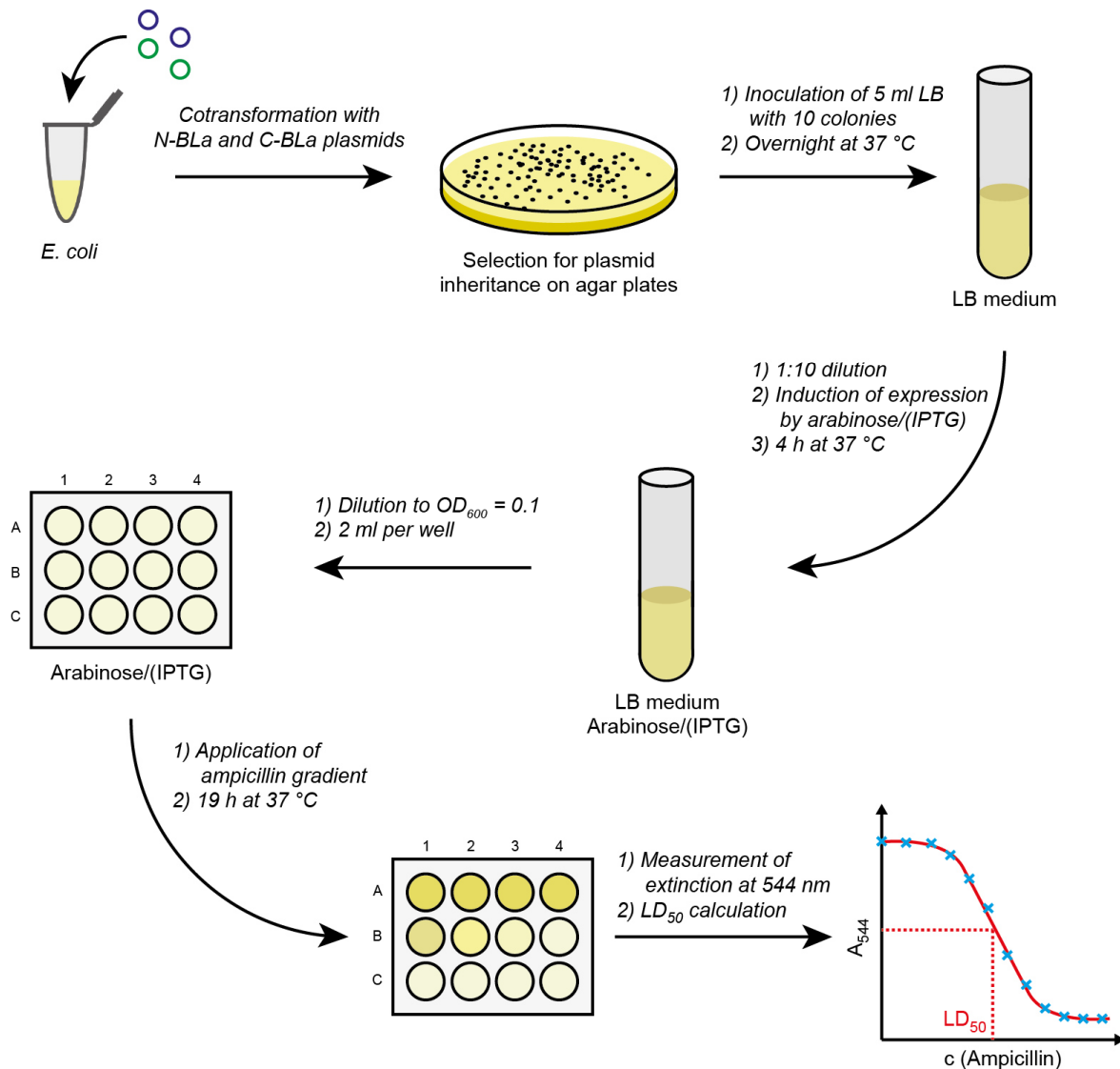


Figure 17: Protocol for the BLaTM Assay. Chemical competent *E. coli* cells were cotransformed with N-BLa and C-BLa plasmids and plated on agar plates. 10 colonies were pooled in 5 mL LB medium and incubated overnight. Induction with 1.33 mM arabinose for BLa 0.1 and BLa 1.1 plasmids; 133 μ M arabinose and 0.1 to 0.7 mM IPTG for BLa 1.2 plasmids.

N-BLa and C-BLa plasmids containing the TMD pair to be measured, were cotransformed into *E. coli* BL21 (BLaTM 1.1) or *E. coli* JM83 (BLaTM 1.1 or BLaTM 1.2), respectively and plated on LB-agar plates containing Cm and Kan for selection for plasmid inheritance. *E. coli* BL21 [230, 231] cells are a protein expression strain, *E. coli* JM83 cells [232] cannot metabolize arabinose and show a higher transformation competence. After incubation for 14 h to 18 h at 37 °C the plates were sealed with Parafilm (Bemis, Oshkosh, WI, US), stored at 4 °C and could be used for inoculation for up to one week. It should be noted that lower plate quality, such as age or moisture, tended to lower the LD50 values.

Hence, the fitness of the cells on the transformation plates is already very crucial for the final measurements. However, if this fact was taken into account, it has been successfully avoided

For the BLaTM assay 10 colony forming units from one agar plate were combined in 5 mL LB-medium (Cm, Kan) and incubated in a turning wheel for 14 h to 18 h at 37 °C. The overnight culture was diluted 1:10 in 5 mL LB-medium (Cm, Kan) containing inducer (BLaTM 0.1 and BLaTM 1.1: 1.33 mM arabinose, BLaTM 1.2: 133 μ M arabinose and 0.1 mM – 0.7 mM IPTG) and incubated again. After 4 h expression the OD₆₀₀ was measured and 25 mL LB-medium (Cm, Kan) of an OD₆₀₀ 0.1 containing the same concentration inducer. 2 mL of the dilution were pipetted in each cavity of a 12-well plate (Greiner Bio-one, Kremsmünster, Austria). Different volumes (Figure 18) of a freshly prepared ampicillin stock solution (concentration depending on the TMD pair: 5 mg/mL – 40 mg/mL) were added resulting in final ampicillin concentrations between 0 μ g/mL – 75 μ g/mL and 0 μ g/mL – 600 μ g/mL. The ampicillin concentration used was adjusted to the LD₅₀ values of the measured TMD to secure high resolution. Applying small concentration steps of ampicillin concentration ensures accurate LD₅₀ determination.

The plates were incubated in a moisturized sealed container for 19 h at 37 °C and 200 rpm on a shaker (shaking amplitude 10 mm, Orbital Shaker 3005, GFL (Burgwedel, Germany)). Then the extinction at $\lambda = 544$ nm was measured directly with a microplate reader (FluoStar, BMG Labtech). For the calculation of the LD₅₀ value, the absorbance data were fitted with the Hill equation (s. 3.15.4). At least two independent cultures for each TMD pair were subjected to the ampicillin gradient in duplicate experiments.

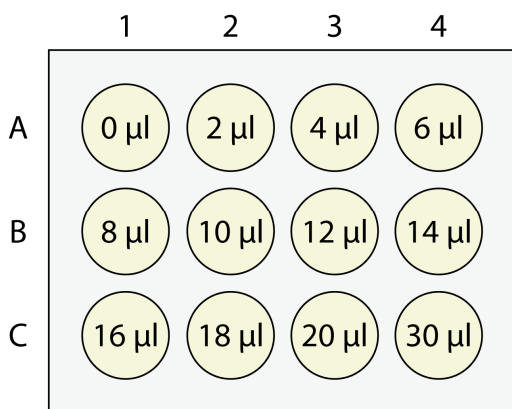


Figure 18: Pipetting scheme for BLaTM assay in a 12-well plate. Volumes given correspond to added ampicillin stock solution per 2 mL of culture ($c(\text{Amp}) = 5 \text{ mg/mL} - 40 \text{ mg/mL}$). The final concentration depends on affinity of TMD pair.

3.15.4 Data analysis

LD₅₀ values

The dimerization of a TMD pair was calculated from the cell densities of *E. coli* cultures at different ampicillin concentrations after 19 h growth in a 12-well plate. The relative affinity is described by the LD₅₀ value, which is the ampicillin concentration where half of the cells are dead. This is represented by the half of the maximum cell density. The extinction data at $\lambda = 544$ nm, which are the measured cell densities at different ampicillin concentrations, were fitted with the Hill equation to calculate the LD₅₀ values [233, 234] (equation 1), y : normalized absorption. c : maximum of sigmoidal curve of the fit. k : LD₅₀. x : normalized ampicillin concentration to highest applied concentration. g : Hill coefficient (maximum slope of the sigmoidal curve). The Hill equation fits the growth data to a sigmoidal curve whereas the inflection point is the LD₅₀. For the calculation the data were normalized to the maximum absorption value and to the maximum applied ampicillin concentration. The minimum of the sigmoidal curve was assumed to be 0.

$$y = \frac{c}{1 + \left(\frac{x}{k}\right)^{-g}} \quad (1)$$

The Hill equation was integrated in Python script by Mark Teese and Alexander Götz to simplify the analysis and increase the data throughput. A quality criterion for the fit is the difference between the values of the inflection point and the curve center, which should be 0. Differences of 0.05 were accepted for the y-values and 0.001 for x-values. Failed fits were revised manually and may not be used for further calculations.

Disruption index

The disruption index describes the impact of amino acids mutations of a TMD on the relative affinity defined by the LD₅₀. For comparability of the impact of distinct mutations in different TMDs, the disruption index is normalized to the LD₅₀ of the wild type of the investigated TMD (equation 2, [wt]: LD₅₀ of the wild type TMD, [M]: LD₅₀ of the mutated TMD).

$$disruption\ index = \frac{[wt] - [M]}{[wt]} \quad (2)$$

If the mutant has a lower affinity than the wild type, the value is between 0 and 1 at which 0 means no effect and 1 complete disruption of the interaction. Negative values indicate a positive impact of a mutation to the ability to dimerize.

4 Results

4.1 Initial experiments

4.1.1 Characterizing the activity of full length β -lactamase

Initial experiments were conducted to optimize split β -lactamase fusion proteins. Therefore, the expression, the localization and the activity of different full-length β -lactamase fusion proteins were monitored and compared. To ensure proper membrane insertion, the native β -lactamase signal peptide was exchanged by the signal peptide of Pectin lyase B (PelB, UniProtKB [122] accession number Q00205) which increased the protein expression to an amount detectable by Western blot (Figure 24). Furthermore, the enhancing β -lactamase stability mutation M182T [193] and the cytoplasmic marker proteins green fluorescent protein (EGFP) and super folder green fluorescent protein (sfGFP) were investigated [227, 235]. Therefore, β -lactamase activity was quantified by the turnover of the chromogenic cephalosporin nitrocefin [194] (Figure 19 A). Due to hydrolysis, the absorption maximum switches from 390 nm to 492 nm. The increase of the absorption at 492 nm is proportional to the β -lactamase activity. The EGFP fluorescence indicates an N_{out} orientation of the fusion protein as GFP is only active under reducing conditions in the cytoplasm but not in the periplasmic oxidizing environment [236, 237]. sfGFP is active in the periplasm as well, but only if it is exported by the post-translational Tat-pathway [238]. This was avoided by a signal peptide specific for the SecYEG-dependent pathway. To ensure diffusion controlled access of nitrocefin to periplasmic β -lactamase, EDTA was added to the buffer, to weaken the *E. coli* cell wall by complexation of divalent cations which are necessary for cell wall integrity [239]. Combination of both parameters allow topology determination which is aimed to be N_{out}/C_{in} . This is the case at high GFP fluorescence and high β -lactamase activity.

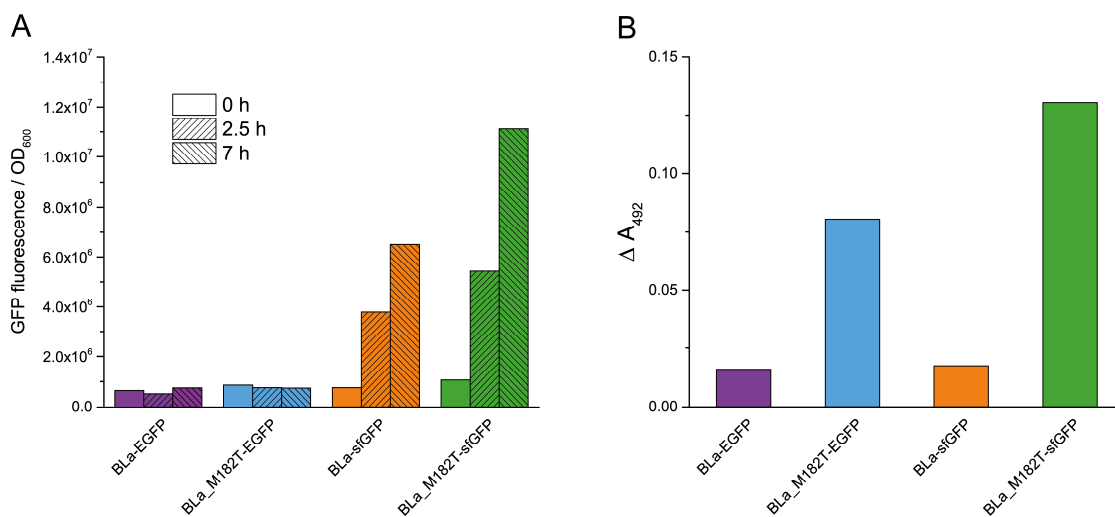


Figure 19: β -Lactamase hybrid protein activity and GFP fluorescence. EGFP: enhanced GFP. sfGFP: superfolder GFP. M182T: stabilizing β -lactamase mutation. All constructs contain a PelB signal peptide, full length β -lactamase (BLA), a 5 amino acids linker and the Integrin- α V-GP TMD. **A)** GFP fluorescence ($\lambda_{\text{Ex}} = 485 \text{ nm}$, $\lambda_{\text{Em}} = 520 \text{ nm}$) normalized to OD_{600} . Before induction (0 h), 2.5 h and 7 h after induction with $166 \mu\text{M}$ arabinose. **B)** β -lactamase activity measured by nitrocefin turnover. Buffer is supplemented with 1 mM EDTA for cell wall disruption. The nitrocefin cleavage product is quantified at $\lambda = 492 \text{ nm}$. Expression for 5 h and induction with $166 \mu\text{M}$ arabinose.

Figure 19 shows that GFP fluorescence (A) of full-length β -lactamase constructs does not correlate with the nitrocefin turnover (B). Western blot analysis had shown that all proteins were equally and correctly expressed (data not shown, because proteins contain different antibody epitopes). EGFP containing proteins (purple and blue) do not show any fluorescence increase after protein expression induction, whereas the superfolder GFP (orange and green) seems to be correctly folded and is fluorescent. In contrast to β -lactamase wild type, only the proteins containing the β -lactamase enhancing M182T mutation [193, 209] can degrade the β -lactam nitrocefin (blue and green). The combination BLA-M182T-sfGFP (green) was used as a base for all subsequent fusion proteins, because (i) Western blot analysis demonstrated full length protein expression, (ii) nitrocefin turnover indicates periplasmic, active β -lactamase and (iii) the high GFP fluorescence level proves the existence of correctly folded, cytoplasmic GFP. It was previously shown that the M182T mutation support the split-protein-complementation [153].

4.1.2 Complementation of soluble split β -lactamase

After the successful demonstration of membrane anchored β -lactamase showing enzymatic activity, it had to be proven that split β -lactamase complementation also works in the periplasm of *E. coli* [153]. Therefore, two compatible plasmids with a similar low copy numbers and the same p_{BAD} promoter sequence were used as vectors for all

complementation experiments (Figure 10). As published in 2002, the homodimerizing leucine-zipper bZIP from the GCN4 transcription factor conveys nitrocefin hydrolysis by the reconstitution of a split β -lactamase [153, 229]. Hence, the homodimerizing leucine-zipper domain bZIP of the transcription factor GCN4 from *S. cerevisiae* [67, 68, 240] was fused to the C-terminus of both split β -lactamase fragments and periplasmic secretion was enforced by an N-terminal PelB signal peptide. The activity of reconstituted β -lactamase was quantified by nitrocefin turnover and additionally by ampicillin resistance. Furthermore, it should be shown that each fragment alone does not generate any nitrocefin hydrolysis or ampicillin resistance higher than the background.

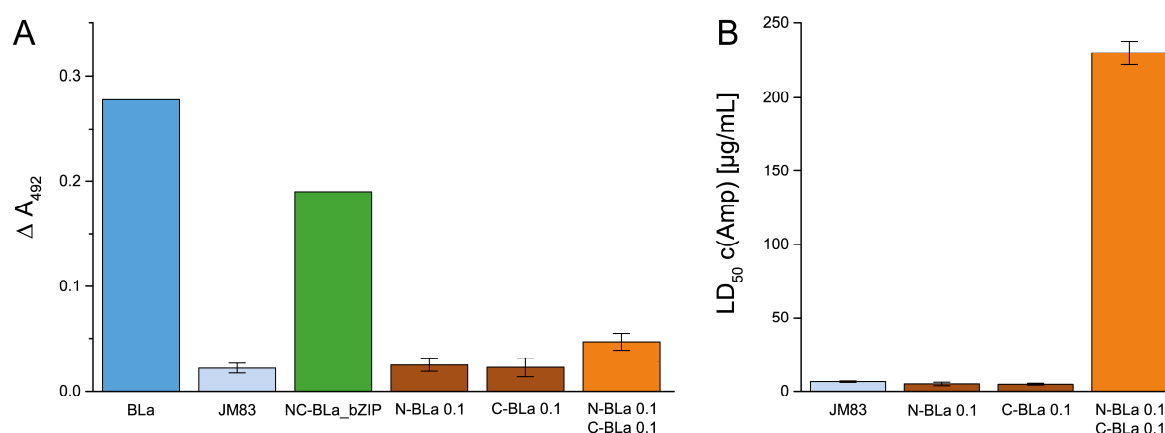


Figure 20: Initial experiments with split β -lactamase fragments fused to soluble leucine-zipper bZIP. Plasmids transformed in *E. coli* JM83. BLa (blue): native TEM-1 β -lactamase coded on the pBAD322 plasmid (constitutively expressed). JM83 (light blue): *E. coli* JM83 only. NC-BLa_bZIP (green): full length (PelB)BLa_M182T_bZIP. N-BLa 0.1 (brown): N-terminal fragment (PelB)N-BLa_M182T_bZIP. C-BLa 0.1 (brown): C-terminal fragment (PelB)C-BLa_bZIP. Orange: cotransformed N-BLa 0.1 and C-BLa 0.1 plasmids. **A**) β -lactamase activity measured by nitrocefin turnover after 20 min. Induction with 1.33 mM arabinose. Means \pm SEM, $n = 1-4$. **B**) β -lactamase activity measured by mediated ampicillin resistance (LD₅₀). Induction with 1.33 mM arabinose. LD₅₀ values of full length β -lactamase containing constructs were unquantifiable. Means \pm SEM, $n = 3$.

Figure 20 A shows the β -lactamase activity of different constructs or cotransformed constructs, respectively. The low signal of the non-transformed *E. coli* JM83 (light blue) indicates only minor background activity due to unspecific hydrolysis of nitrocefin. The activity of the native soluble β -lactamase (blue: BLa) is only one third above that of the soluble β -lactamase bZIP fusion protein (green: NC-BLa_bZIP). This indicates that the protein was successfully transported into the periplasm. Additionally, the C-terminal bZIP domain does not disturb enzymatic activity. The nitrocefin turnover mediated by the individual β -lactamase fragments (brown: N-BLa 0.1 or C-BLa 0.1) is at background level (light blue: JM83 cells only). The cotransformation of these two constructs (orange), which reconstitute the split β -lactamase to an active enzyme after bZIP-mediated dimerization,

leads to a two-fold increased activity. However, this trend is not significant. In a further step, the activity of the negative control, the fragmented β -lactamase alone and coexpressed fragments were quantified by ampicillin resistance (Figure 20 B). For the quantification of β -lactamase activity by ampicillin resistance (LD₅₀-values) the density of bacterial cells grown at increasing ampicillin concentrations was monitored after 19 h. It was measured as absorption at 544 nm directly in a 12-well plate. For each LD₅₀ calculation the growth at 12 different ampicillin concentrations was determined in parallel (Figure 21). *E. coli* JM83 transformed with only one of the two β -lactamase fragments (brown) showed no enhanced resistance to ampicillin relative to the *E. coli* JM83 without any plasmid (light blue). By contrast, the *E. coli* cells which were cotransformed with both plasmids (orange) and contain both parts of the soluble β -lactamase are resistant to ampicillin concentrations by more than 200 $\mu\text{g}/\text{mL}$ which exceeds the negative controls for more than 40 times. The LD₅₀ values of the full length β -lactamase constructs could not be quantified using the same set-up because the ampicillin resistance was too high ($> 1000 \mu\text{g}/\text{mL}$).

4.1.3 Evaluation of ampicillin resistance data

The data for the calculation of the LD₅₀-value was directly obtained in 12-well plates as cell density, measured as absorption at 544 nm. The absorption data are plotted against the ampicillin concentrations and the LD₅₀-value is the ampicillin concentration where the absorption is half maximal.

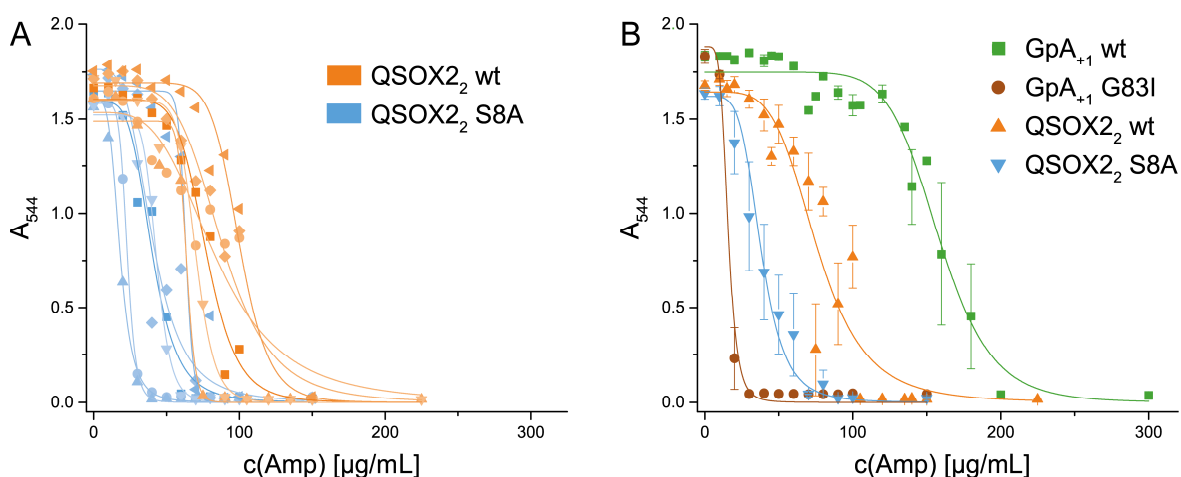


Figure 21: Illustration of fitting absorption raw data of the BLaTM 1.1 assay with Hill equation. **A)** Absorption data of 6 independent measurements each of QSOX2₂ wt and QSOX2₂ S8A. **B)** Averaged absorption data of four different TMDs. Single data points of QSOX2 wt and QSOX2 S8A are shown in A). Means \pm SEM, $n = 4 - 6$.

Figure 21 A shows the absorption data of six independent measurements each of the model TMDs QSOX2₂ wt (orange) and QSOX2₂ S8A (blue) in the BLaTM 1.1 assay. The sigmoidal behavior is the basis for the LD₅₀ calculation. Each data set was separately fitted with the Hill equation (Equation 1, p. 52) for LD₅₀ calculation. It illustrates that in all, except of one mutant outlier, measurements of the QSOX2₂ wild type TMD (orange) enables growth up to higher ampicillin concentrations than its non-dimerizing mutant S8A (blue). The absorption data of all measurements for four different exemplary TMDs were combined (Figure 21 B). It clearly shows TMD-dependent survival of *E. coli* at different ampicillin concentrations. The high error bars at certain data points are always at medium absorption levels because small differences of the resistance cause large absorbance differences.

4.2 BLaTM 1.1

In the following the protocol was optimized with respect to (i) *E. coli* strain, (ii) cotransformation, (iii) inoculation strategy, (iv) incubation time, (v) strength of induction, (vi) density of the cells, (vii) type of shaker and (viii) shaker speed. The final protocol is described in chapter 3.15.3.

Although the system is perfectly suited for the measurement of heterotypically interactions, the optimization experiments were conducted with homotypic interacting TMDs. These model TMDs are very well investigated and their interfaces are known. Furthermore, heterotypic measurements are more complex by inactive homotypic dimers of the two TMDs which may reduce the amounts of protein available for heterotypic dimers (Figure 33). To prove the generality of the findings, two strongly dimerizing TMDs with known interfaces were used for assay implementation. The TMD of the human Glycophorin A (GpA; UniProtKB [122] accession number P02724) was chosen as a test construct, as it is intensively investigated and a very well understood right-handed homodimerizing TMD [77, 241-245]. Furthermore, it is an accepted model TMD for the establishment of TMD-TMD interaction assays such as ToxR [77], TOXCAT [110], GALLEX [108] or BACTH [42] and is used as a reference in these assays for normalization. In addition, the TMD of the human sulfhydryl oxidase 2 (QSOX2; UniProtKB [122] accession number Q6ZRP7) was selected [246]. The S8A mutation of the QSOX2 TMD decreases the signal to GpA G83A level, which proved that serine 8 is part of the interaction interface.

In order to investigate the orientation-dependence (s. 1.3.3) of TMD-TMD interactions observed in other assays [77, 108, 121] the flexible linker between the β -lactamase fragments and the TMD was optimized to achieve a high signal-to noise signal and low orientation-dependence at the same time.

4.2.1 Quantification method

The primary experiments using soluble dimerization domains (s. 4.1.2) already indicated the low sensitivity of quantification of β -lactamase by nitrocefin turnover in the periplasmic space (Figure 20). For this reason, the β -lactamase fragments, fused to TMDs, were tested by nitrocefin turnover and ampicillin resistance.

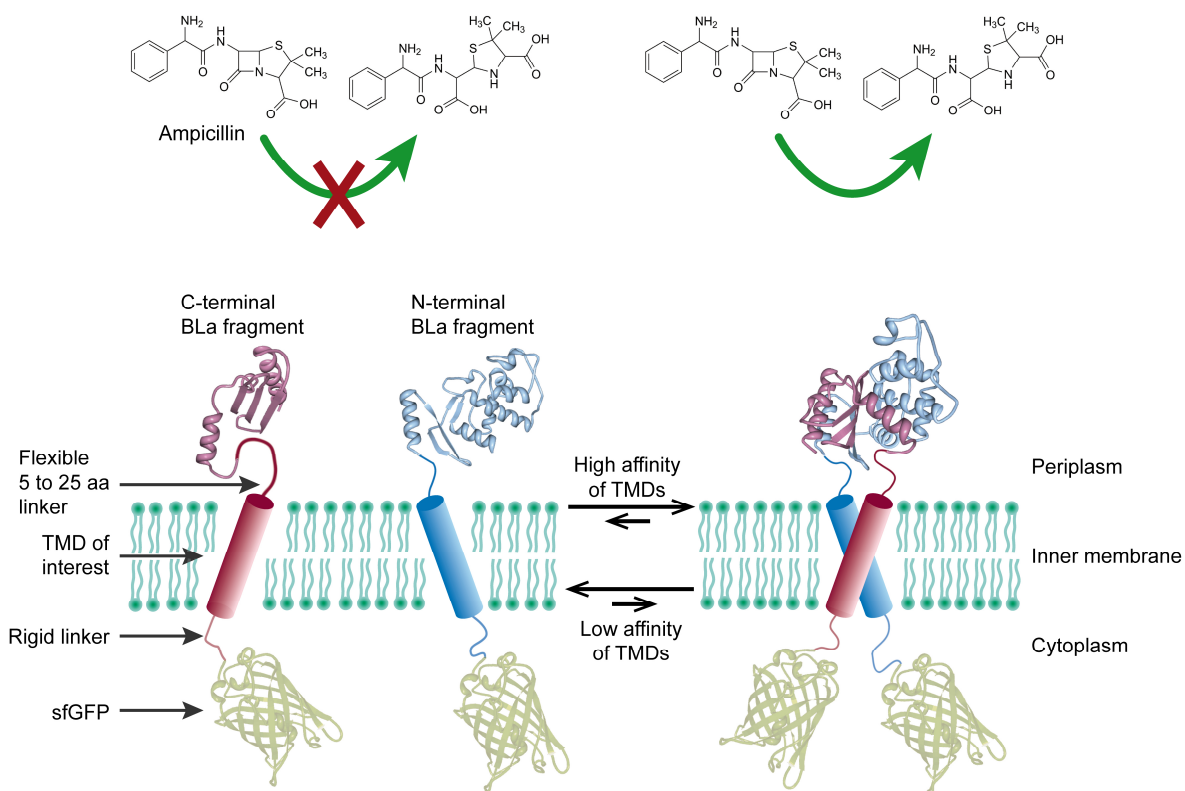


Figure 22: Schematic illustration of the BLaTM assay (not to scale). Green: inner *E. coli* membrane. Red: C-terminal split β -lactamase fragment (C-BLa) fused to TMD (cylinder). Blue: N-terminal split β -lactamase fragment (N-BLa) fused to TMD (cylinder). Light green: cytoplasmic superfolder GFP (sfGFP). All labels count for both fusion proteins. Left side: Low affinity between the TMDs. Fusion proteins are monomeric and ampicillin or any other substrates cannot be degraded. Right side: High affinity between two TMDs. Fusion proteins form a heterodimer and the split β -lactamase fragments reconstitute to an active enzyme. Ampicillin can be degraded which leads to antibiotic resistance. PDB IDs: TEM-1 β -lactamase: 1axb, sfGFP: 4lqu.

Figure 22 illustrates the working principle of the BLaTM system. The fusion proteins N-BLa (blue) and C-BLa (red) consist of five parts (s. appendix 9.5). They differ in their large periplasmic domain, which is either the N-terminal fragment, containing the

stabilizing M182T mutation, or the C-terminal fragment of the TEM-1 β -lactamase, respectively. Both fragments are fused to a TMD of interest via a flexible linker, which enables free rotation and creates space to the membrane. The stabilizing and detectable cytoplasmic sfGFP is fused via rigid linker to each TMD.

If there is no or low affinity between the two TMDs of interest (left), both fusion proteins will be monomeric and the β -lactamase fragments cannot reconstitute. Thus, the protein is not active and ampicillin cannot be degraded. The cells die. The higher the affinity, the more proteins will form heterodimers (right). Hence, the N-terminal β -lactamase can reconstitute and form an enzymatic active enzyme which can degrade its substrate ampicillin. The more proteins form dimers, the more active enzymes are in the periplasm, the higher is the lethal ampicillin concentration.

The experiments were conducted with three different TMDs in the BLaTM 1.1₅ system (5 amino acids linker between β -lactamase fragment and TMD constructs) and it is likely that the results are similar with the other versions of the BLaTM fusion protein.

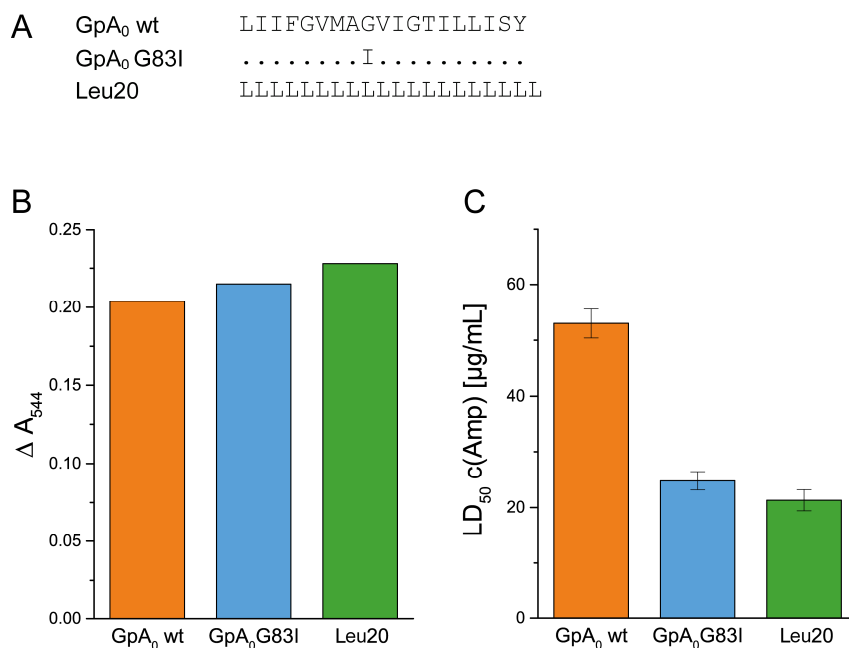


Figure 23: Initial experiments with membrane-bound split β -lactamase constructs using BLaTM 1.1₅ (5 amino acids linker between BLa fragment and TMD). **A**) TMD sequences **B**) Enzymatic activity of reconstituted β -lactamase after 5 h induction with 166 μ M arabinose, quantified by nitrocefin turnover after 20 min. The nitrocefin cleavage product is quantified at $\lambda = 492$ nm. $n = 1$. **C**) β -lactamase reconstitution efficiency measured by mediated ampicillin resistance (LD₅₀). Induction with 1.33 mM arabinose. Means \pm SEM, $n = 5$.

It was expected that the GpA wt TMD containing constructs mediate a higher enzyme activity than the Leu20 TMD and the negative control GpA G83I. G83I reduces the high

affinity of the GpA wt TMD by ~70 – 90%, depending on the assays used [42, 77, 108]. However, it was demonstrated (Figure 23 A) that the nitrocefin turnover is independent from the affinity of the measured TMD. Hence, it is not possible to use this method for the measurement and discrimination of different TMD-TMD affinities because TMDs of different known affinities cannot be distinguished. Consequently, a more sensitive method for the reliable detection of reconstituted β -lactamase was needed. The same constructs were thus tested for their ability to confer ampicillin resistance of expression *E. coli* (Figure 23 B). Indeed, different LD₅₀ values discriminated the positive control GpA from the negative control GpA G83I and from Leu20 (Figure 23 C). In conclusion the established nitrocefin based method, which is suitable to detect soluble protein-protein interactions using split β -lactamase, is not suitable for the measurement of TMD-TMD interactions [247-249]. The more sensitive ampicillin resistance is the method of choice for comparing TMD-TMD affinities.

4.2.2 Protein expression level

All N-BLa 1.1 and C-BLa 1.1 proteins contain a FLAG-epitope at their C-terminus [250]. Hence, full length proteins and C-terminal cleavage products can be specifically detected via Western blot. Figure 24 B shows the Western blots of *E. coli* cells expressing all eight cotransformed GpA TMD constructs and all eight QSOX2 TMD constructs in the BLaTM 1.1₁₃ version. Additionally, constructs lacking a signal peptide (Δ SP) were analyzed for insertion control. Proteins recognized by the membrane protein insertion machinery were processed by a signal peptidase which cleaves off the signal peptide (s. 1.2.2). Following, correctly inserted proteins should have the same molecular weight as the Δ SP control proteins.

molecular weight as the Δ SP_C-BLa protein. This suggests correct membrane insertion and processing of this protein. On the other hand, the signal peptide PelB of the N-BLa 1.1 protein seems not to be cleaved off, as suggested by its larger mass, which suggests insufficient protein recognition, membrane insertion and/or processing by a signal peptidase. The corresponding LD₅₀ values of all 17 constructs (Figure 24 A), show that the ampicillin resistance and the protein expression are completely independent.

4.2.3 Influence of the linker between the β -lactamase fragment and TMD

First measurements of GpA TMDs were conducted using a 5 amino acids linker between the β -lactamase fragments and the TMDs. As significant differences in the LD₅₀ values of different orientations were obtained in these experiments (Figure 25 B), it was assumed that the linker was too short to ensure free orientation of the β -lactamase fragments such that the positions of both fragments relative to each other was coupled to the relative position of the fused TMD. Therefore, the aim was the creation of a linker which is flexible enough to achieve the independent movement of the fragments of the fusion proteins and thus, enables reconstitution, on the one hand, but does not increase the background signal, on the other. Consequently, the linker was extended by one, two or four GGGS segments (Figure 25 A) to find the optimal length. To test for orientation dependence, the strongly dimerizing GpA and QSOX2 wild type TMDs and their mutants G83I (GpA) and S8A (QSOX2), respectively, were used. In order to cover all interaction interfaces all four orientations for each wild type and mutant TMD were investigated (Figure 25 B and C).

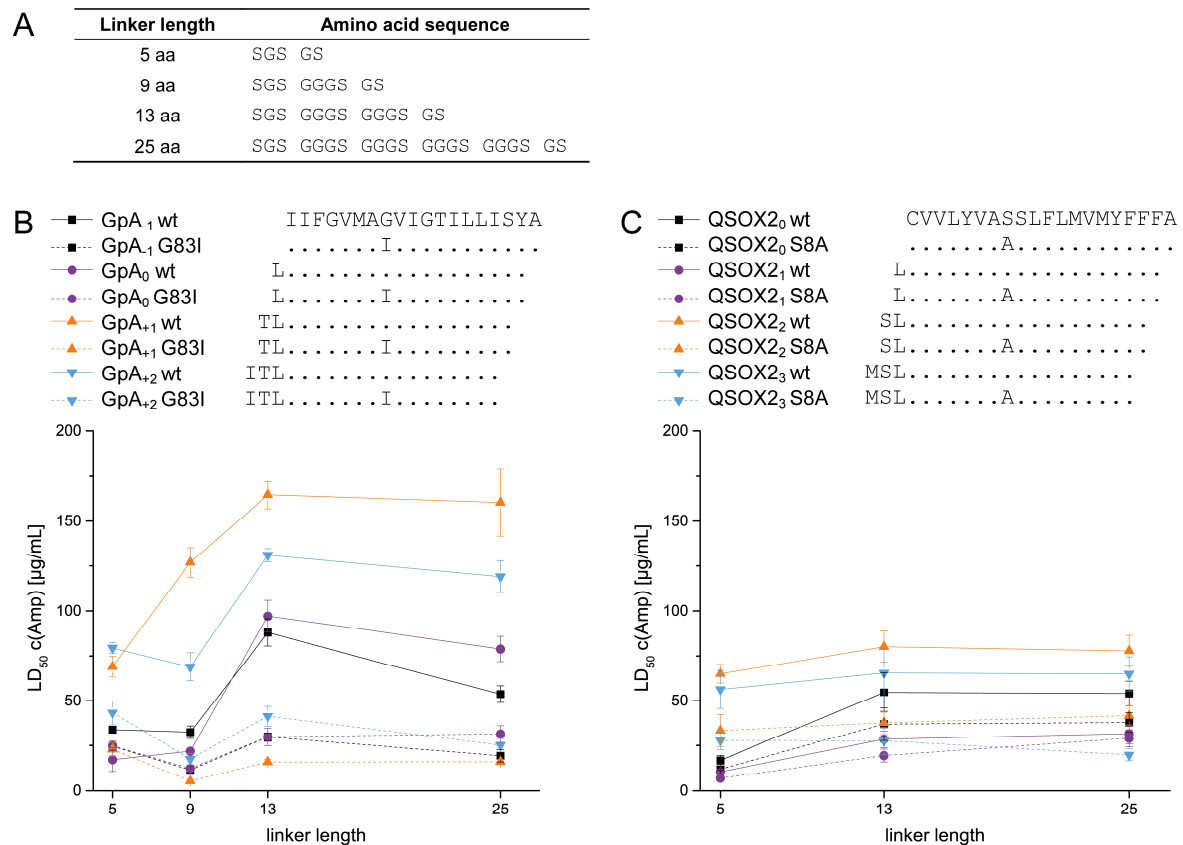


Figure 25: Effect of the linker length between the β -lactamase fragments and the TMD to the LD_{50} in the BLaTM 1.1 assay. **A)** Amino acid sequences of the tested linkers between β -lactamase fragments and TMD. **B)** top: Amino acid sequences of all used Glycophorin A (GpA) model TMDs. Bottom: LD_{50} of all GpA TMDs measured with a 5, 9, 13 and 25 amino acids linker between split β -lactamase fragments and TMD. Means \pm SEM, $n = 3 - 4$. **C)** Top: Amino acid sequences of all used Sulphydryl oxidase 2 (QSOX2) model TMDs. Bottom: LD_{50} of all QSOX2 TMDs measured with a 5, 13 and 25 amino acids linker between split β -lactamase fragments and TMD. Means \pm SEM, $n = 3 - 4$.

The homotypic affinity of the different GpA TMDs with different linkers (Figure 25 B) was characterized by a number of phenomena. First, the LD_{50} values of the negative control G83I (dashed lines) are more or less independent from the linker length for all four orientations. The smallest values were obtained with the 9 amino acids linker. The LD_{50} values of the wild-type TMDs increase up to 3-fold from the 5 amino acids linker to the 13 amino acids linker. In some cases (GpA₋₁ wt and GpA₀ wt) the resistance levels decrease with a further extension of the linker to 25 amino acids. The wild-type signals of the four orientations of the GpA wt TMDs are the closest to each other with the 13 amino acids linker.

The measurements of the different QSOX2 TMDs in the BLaTM 1.1 system with different flexible linkers (Figure 25 C) show also that some signals increase with a longer linker, nevertheless to a lesser extent than with the GpA TMDs. There is no effect on the two stronger dimerizing TMDs QSOX2₂ (orange) and QSOX2₃ (blue) – neither on the wild-

types nor on the mutants S8A. However, there is an effect on the signal intensity of the wild-type and mutant TMDs QSOX₀ (black) between the 5 and 13 amino acids linker, which results in more homogenous signal intensities of the four wild-type orientations and their S8A mutants of the QSOX2 TMD, respectively.

4.2.4 Orientation dependence

Figure 25 B and C lead to the assumption that the linker extension could not completely abolish orientation dependence. Orientation dependence describes the maximal difference of the LD₅₀ of the four orientations of one TMD. It is the quotient of the lowest LD₅₀ and the highest LD₅₀. Hence, the value is always between 0 and 1, whereas 1 means no difference and no orientation dependence. The higher is the orientation dependence. For in-depth analysis, the orientation dependence was calculated for the TMDs GpA, QSOX2 and LS46 (Figure 26 and Figure 29) and averaged for each linker.

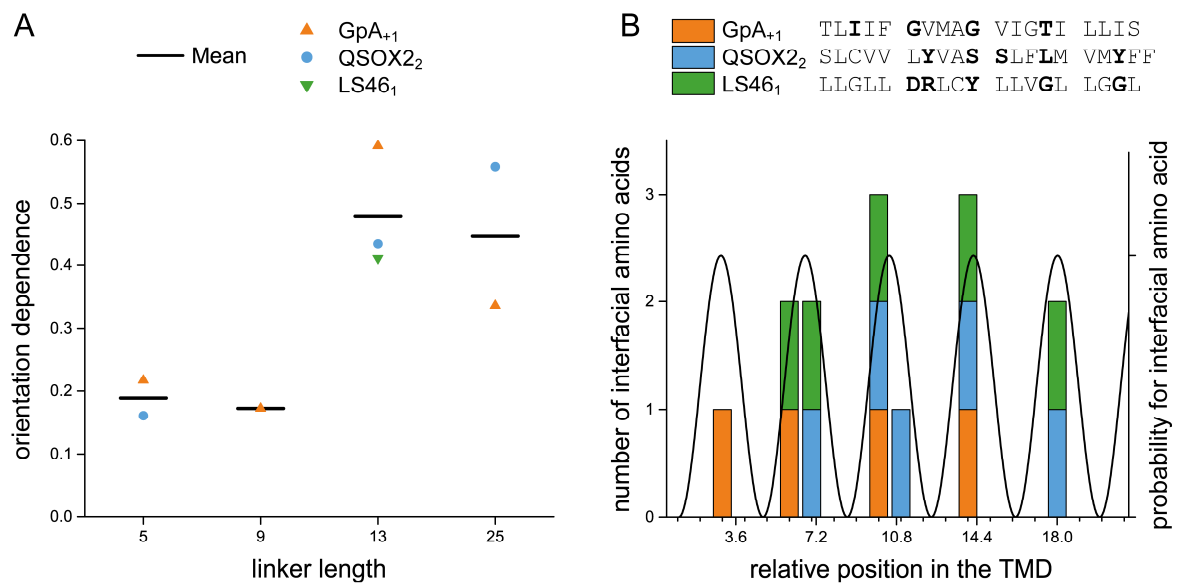


Figure 26: Orientation dependence of different TMDs and site directness of the interfacial amino acids. **A**) Absolute and averaged orientation dependence of GpA, QSOX2 and LS46 TMDs for each linker length (Figure 25 and Figure 29) in the BLaTM 1.1 system. **B**) Relative position of interfacial amino acids of the model TMDs GpA, QSOX2 and LS46 to the periplasmic split β -lactamase fragments. Relative probability of each amino acid position is indicated. The most important interfacial amino acids [77, 226, 252] (in bold) of the best orientations (Figure 25, Figure 29) are blotted against their position in the TMD. Each helix turn consists of 3.6 amino acids as indicated on the x-axis. The periodicity of probability is illustrated by a fitted sinus curve.

The summary of all GpA, QSOX2 and LS46 data shows (Figure 26 A) that there is a very high orientation dependence for short linkers (5 and 9) and a considerably lower one for the two longer ones. The LS46 TMD was only tested with a 13 amino acids linker. The orientation dependence might result from deleted or added amino acids located at the

termini, yet important for interaction. The orientation dependence could also systemically result from steric hindrance or necessary positioning of the β -lactamase fragments relative to the interaction interface. If the second case is correct, the interfacial amino acids of all TMDs should always be positioned at about the same relative positions in the TMD. To investigate these two hypotheses, the most strongly dimerizing orientations of the three tested TMDs GpA, QSOX2 and LS46 were chosen for analysis (Figure 26 B). The interfacial amino acids were defined according to published ToxR experiments [77, 246, 252] and blotted against their positions in the TMD. The hits are summed up to emphasize the important positions. The distribution of the interfacial amino acids follows the α -helical pattern which says that each turn exists of 3.6 amino acids [12]. Hence, the interfacial amino acids of all three TMDs are on the same side, which indicates a systematic impact of the system to the orientation dependence. Derived from this periodicity the relative probability, that a distinct amino acid is a part of the interaction interface, can be estimated for each position in the TMD (Figure 26 B, black curve).

4.2.5 Disruption index

One important aspect in assay development is the optimization in respect to the resolution, which is defined here by the difference in LD_{50} values between positive and negative controls. This relationship can be described by the disruption index. It normalizes the impact of a mutation to the wild type signal resulting in a relative number. If the mutation has a negative impact on dimerization, the value is between 0 and 1, where 0 means no effect and 1 signifies the complete disruption of the interaction. Complete disruption would mean no growth at all. Negative values indicate a positive impact of the mutation on dimerization. As there is always a background signal due to natural ampicillin resistance of *E. coli*, a value of 1 cannot be reached. The disruption index was calculated for each GpA and QSOX2 TMD (Figure 25) and plotted against the linker length.

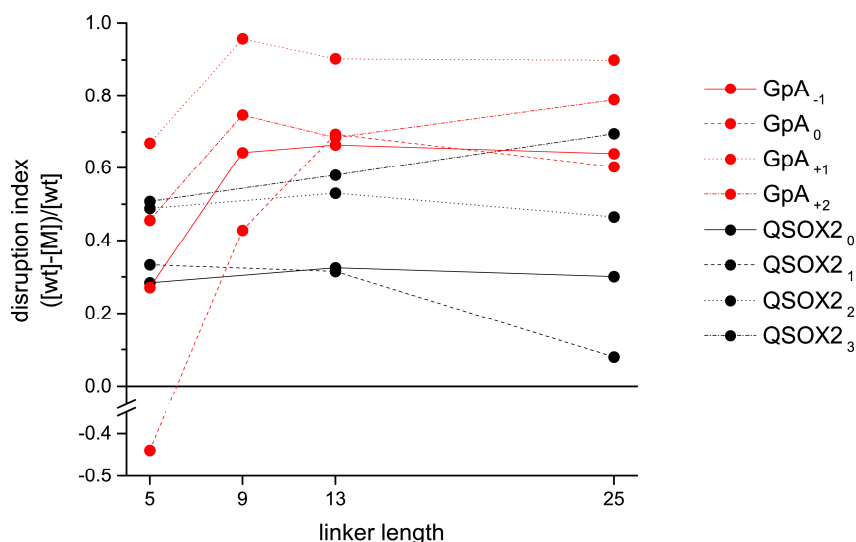


Figure 27: Disruption indices ($([wt]-[M])/[wt]$; $[wt] = LD_{50}$ of wild type, $[M] = LD_{50}$ of GpA G83I or QSOX2 S8A) of all orientations of GpA and QSOX2 BLaTM 1.1 constructs plotted against the linker length.

The analysis of the disruption indices (Figure 27) shows that they vary depend on the linker length. For the GpA TMDs (red), the disrupting impact of the mutation increases between the linkers 5 to 13, especially between linker 5 and 9. The disruption indices with the 13 amino acids and 25 amino acids linker are quite similar, while the 13 amino acids linker shows the better results, as the values are closer to each other. The shortest linker has the poorest resolution. The analysis of the QSOX2 TMD data (black) shows more similar disruption indices, especially between the 5 and 13 amino acids linkers. Here, the variations between the different orientations increase with the 25 amino acids linker, relative to the other two. This is due to two variations of the mutant signals (Figure 25 C). The wild type signals are constant for all four orientations. In sum, a linker of 13 amino acids leads to the highest resolution, overall consistent with the least orientation dependence.

4.2.6 Heterotypic interacting TMDs

After it was successfully demonstrated that the assay works for homotypic TMD-TMD interactions, it should be shown that the assay is also suitable for the measurement of heterotypic TMD-TMD interaction. As a model, the LS46 TMD was chosen. This high-affinity, homotypically interacting TMD was found by Herrmann et al. in a library screening using the ToxR/POSSYCCAT system [252, 253]. The interaction motif includes the two ionizable amino acids aspartic acid (D) and arginine (R) (Figure 28). These amino acids can be (partially) charged, where D is negative and R is positive. This results in two

D-R interaction sites. The mutation of the D on one TMD and the R on the other one to e.g. leucine (L) results in one, heterotypic D-R interaction site. These mutant TMDs should have a lower propensity for homotypic interactions than the original double charged TMDs.

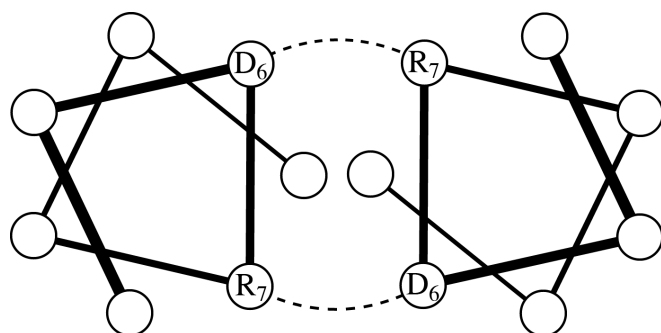


Figure 28: Main interaction interface of TMD LS46 of the ionizable amino acids aspartic acid (D) and arginin (R). Solid lines: peptide backbone, dashed lines: salt-bridge like bonds. Helix wheel drew with DrawCoil 1.0 and modified (<http://www.grigoryanlab.org/drawcoil/>) [254].

As first experiments indicated orientation dependence (Figure 25), all four orientations of the LS46 model TMD had to be tested (Figure 29 A). The best orientation was used for the generation of the artificial heterotypic interaction interface. The orientation 3 (LS46₃) is one amino acid longer than the others. Otherwise the C-terminal GxxxG motif would be destroyed which is assumed to be essential for a strong interaction [252].

```

LS460      LGLLDRLCYLLVGLLGLL
LS461      L.....
LS462      LL.....
LS463      LLL.....

```

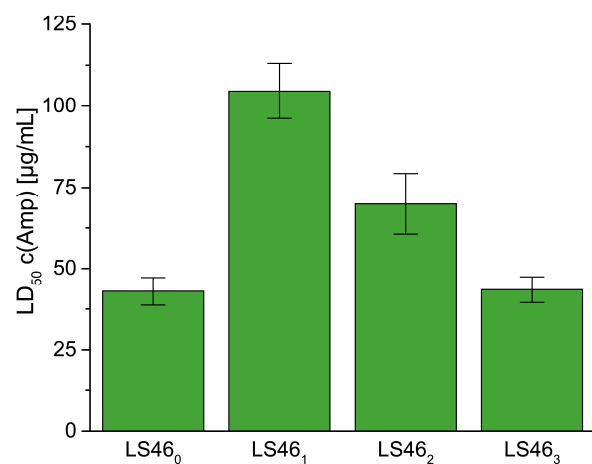


Figure 29: Homotypic interaction of the artificial model TMD LS46 in the BLaTM 1.1₁₃ system. Top: Amino acid sequences of the used TMDs. Ionic amino acids and GxxxG motifs emphasized in bold. Bottom: LD₅₀ of LS46 TMDs. Means ± SEM, n = 4 - 5.

The orientation LS46₁ generates an LD₅₀ which is significantly higher ($p \leq 0.05$) than all the others (Figure 29). It is about $\frac{2}{3}$ of the signal of the optimal GpA wt orientation. This is much lower than in the ToxR assay where the signal is about twice that of GpA wt [252]. Nevertheless, LS46₁ was taken as a base for the generation of heterotypically interacting TMDs. Therefore, a leucine backbone was used with the two ionic amino acids and the GxxxG at the positions corresponding to the LS46₁ TMD (L19_G₁₄G₁₈_D₆R₇). This reductionism should exclude a possible influence of other amino acids in the TMD and ensure good membrane integration.

In theory, the signal should be independent from the TMD/vector combination. That means that the LD₅₀ values of the TMD pair N-BLa::TMD#1/C-BLa::TMD#2 (e.g. L19_G₁₄G₁₈_D₆ and L19_G₁₄G₁₈_R₇) and the LD₅₀ value of the pair N-BLa::TMD#2/C-BLa::TMD#1 would be equal. Thus, it would be sufficient to measure all orientations of N-BLa::TMD#1 against all of C-BLa::TMD#2, but not necessarily in addition all N-BLa::TMD#2 against all C-BLa::TMD#1.

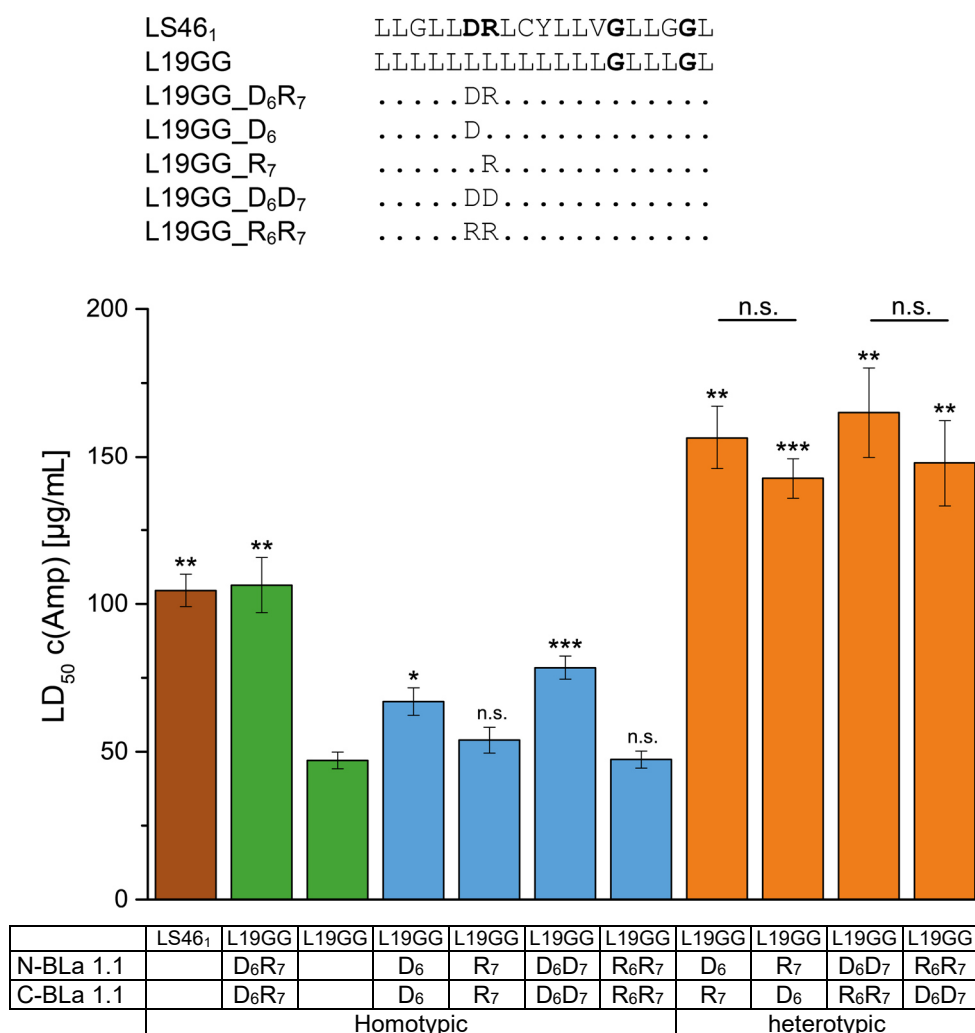


Figure 30: Homo- and heterotypic interactions of TMDs containing ionizable amino acids based on TMD L19_G₁₄G₁₈ in the BLaTM 1.1₁₃ system. Top: Amino acid sequences of the used model TMDs. Proposed interacting amino acids emphasized in bold. Bottom: Brown: LS46₁, Green: Positive (D₆R₇) and negative control (L19_G₁₄G₁₈). Blue: homotypic TMD pairs with one or two ionizable amino acids each. Orange: heterotypic TMD pairs with one or two ionizable amino acids each. Significance levels to negative control. n.s. = not significant, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$. Means \pm SEM, $n = 4$.

Figure 30 shows the results of switching N-BLa for C-BLa plasmids (D₆/R₇ vs R₇/D₆) and demonstrates the impact of the ionizable amino acids arginine (R) and aspartic acid (D) on the homo- (blue) and heterotypic (orange) dimerization propensity of TMDs. Moreover, the significance levels compared to the negative control L19_G₁₄G₁₈ (green; L19GG) are annotated. The highly dimerizing LS46₁ TMD (Figure 29) and the poly-leucine TMD containing the LS46 interaction motifs D₆R₇ and G₁₄xxxG₁₈ (green; positive control) are showing the same LD₅₀s of about 100 µg/mL and are about the double of the negative control. Arginines do not support homotypic dimerization. In contrast to this, one aspartic acid at position 6 significantly increases homo dimerization by about 40%, two aspartic acids at positions 6 and 7 even by more than 60%, which is still considerably less than the positive control. The ionizable amino acids of heterotypic pairs (orange bars) are in

optimal positions to each other (Figure 28) because they are derived from the orientation-optimized TMD LS46 (Figure 29). The LD₅₀ values of all four measured pairs are between 3 and 3.5 times higher than L19GG and are significant ($p < 0.01$). Above that, they are between 34% and 58% higher than LS46. In sum, heterotypic interactions can be measured and that the signals do not differ significantly between the two TMD/vector combinations (orange).

4.3 BLaTM 1.2

Expression analysis of the BLaTM 1.1 proteins by Western blot had shown that the signal peptide of the N-BLa proteins was not cleaved off (Figure 24). This could mean that the signal peptidase did not recognize its cleavage site. To test this issue, the signal peptidase recognition site was mutated to represent the native sequence (H21A) or the one of the C-BLa (A22L, A22L/P23L), respectively. All three modifications did not reduce the molecular weight suggesting that this strategy was not successful (data not shown). Next, the complete signal peptide was replaced by the commonly used signal peptide of Outer Membrane Protein A (OmpA, 1-22, UniProtKB [122] accession number P0A910) of *E. coli*, resulting in system BLaTM 1.2. As a consequence, the protein expression decreased dramatically to below the detection limit of the monoclonal ANTI-FLAG® M2 antibody in a Western blot. The ampicillin resistance strongly increased, however, as growth of neither the GpA₊₁ wt nor the GpA₊₁ G83I constructs was inhibited at ampicillin concentrations of 600 µg/mL. Thus, it was decided to use 133 µM arabinose and additionally inhibit the promoter activation with its inhibitor IPTG [255]. As a result, a tightly adjustable system which can be optimized for high and low affinity TMDs (s. 4.3.1) has been developed.

4.3.1 Adjustment to different TMD-TMD affinities

In order to tune expression using BLaTM 1.2 system, advantage was taken of the fact that the p_{BAD} promotor can be inhibited by IPTG [255]. As it was shown that inhibition with IPTG and influences the ampicillin resistance as well, several IPTG concentrations were tested to achieve the best resolution, i.e., the biggest differences between the positive (wt) and negative controls (G83I). For this purpose, the weakest (GpA₀) and the high interacting (GpA₊₁) orientation were chosen (Figure 31 B, Figure 32). The expression was induced with 133 µM arabinose and 0.1 to 0.7 mM IPTG. IPTG is a competitive inhibitor

of the p_{BAD} promoter [222, 255], meaning that an increase of the IPTG concentration decreases the protein expression and thus the LD_{50} . To verify that the resistance resolution is not influenced by different TMDs, the GFP fluorescence was measured.

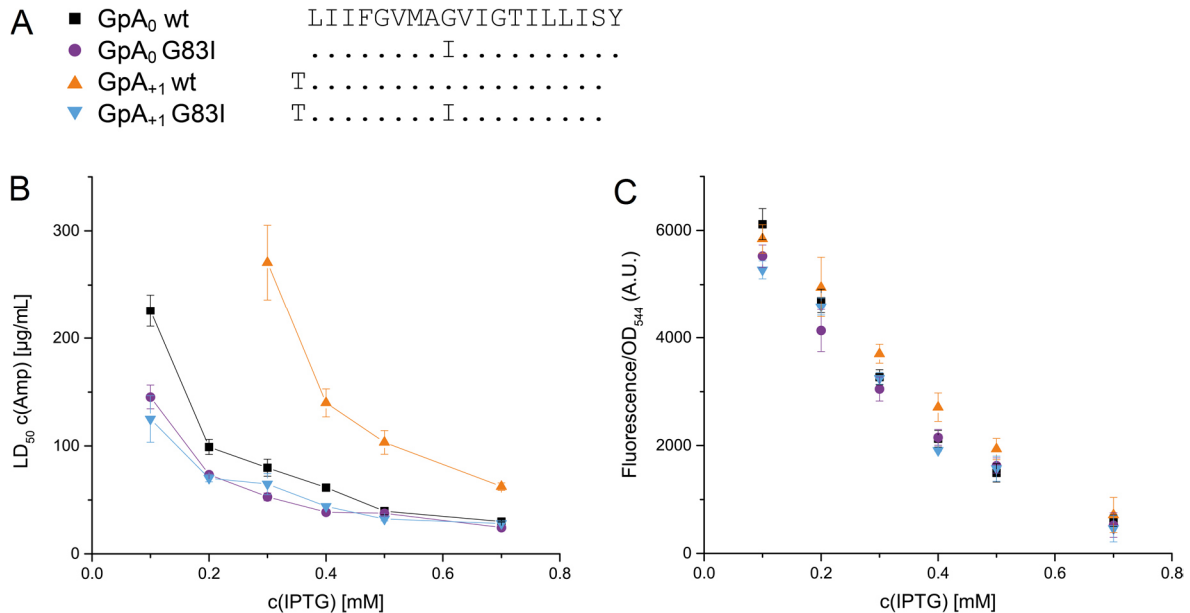


Figure 31: Relationship of protein expression and LD_{50} in the BLaTM 1.2 system. Activation of p_{BAD} promoter by arabinose is repressed by the competitive inhibitor IPTG. **A)** Amino acid sequences of the used TMDs. **B)** Induction dependent LD_{50} s of GpA constructs. Induction with 133 μ M arabinose and variable IPTG concentrations. Means \pm SEM, $n = 1 - 9$. **C)** Induction dependent GFP fluorescence. Fluorescence ($\lambda_{Ex} = 485$ nm, $\lambda_{Em} = 520$ nm) measured after 4 h expression. Corrected to non-induced sample and normalized to A_{544} . Means \pm SEM, $n = 3$.

Figure 31 B shows that increasing IPTG concentrations reduce the LD_{50} for all measured GpA TMDs. At high IPTG concentrations the GpA₀ wt (black) and GpA₀ G83I (purple) constructs cannot be distinguished. The resolution between the wild type and the mutant TMDs increases with decreasing IPTG concentration. Although, the background resistance increases at lower IPTG concentrations, the wt signal is about 50% higher than the signal of the mutant at 0.1 mM IPTG. The LD_{50} of the GpA₊₁ wt (orange) is clearly higher than the LD_{50} of the GpA₊₁ G83I (blue) at all measured induction levels. The resolution between the two TMDs increases with lower IPTG concentrations from a factor of 2.2 at 0.7 mM IPTG to a factor of 4.2 at 0.3 mM IPTG. At higher expression levels the resistance raise to > 600 μ g/mL. The LD_{50} values of the two mutant TMDs are very similar at all IPTG concentrations. Fluorescence ($\lambda_{Ex} = 485$ nm, $\lambda_{Em} = 520$ nm) of the fused GFP (Figure 31 C) demonstrates a linear correlation between the concentration of IPTG and protein expression. At 0.7 mM IPTG, a 5-fold excess of IPTG relative to arabinose, the expression is almost completely repressed. These results (Figure 32) indicate that the concentration of

IPTG in the assay may be adjusted to the affinity of the TMD under investigation. While 0.5 mM IPTG allows the investigation of high-affinity TMD pairs, 0.3 mM IPTG is recommended for lower affinities.

4.3.2 Comparison to BLaTM 1.1

To compare the results of the BLaTM 1.1 assay with the BLaTM 1.2 assay, the same model TMD GpA was used. Again, all four orientations, each as wild type and G83I mutant, were measured. The basis for the 1.2 vectors was the BLaTM_1.1₁₃ constructs. Other linkers between the split β -lactamase fragments and the TMD were not tested again, because the β -lactamase domains should not be influenced by the signal peptide, as it ought to be cleaved off by a signal peptidase after insertion.

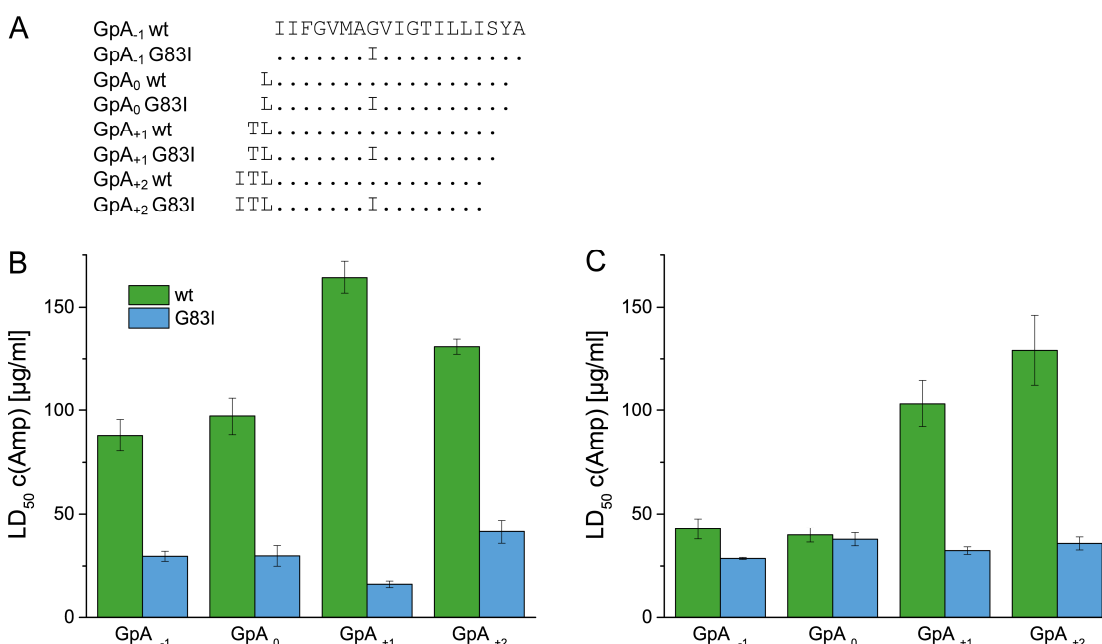


Figure 32: Homotypic interaction of all orientations of GpA wt (green) and GpA G83I TMDs (blue) in the BLaTM systems. **A)** Amino acid sequences of the used TMDs. Orientation as index. **B)** LD₅₀ values in the BLaTM 1.1 assay. Induction with 1.33 mM arabinose. Means \pm SEM, n = 5 - 6. **C)** LD₅₀ values in the BLaTM 1.2 assay. Induction with 133 μ M arabinose and 0.5 mM IPTG. Means \pm SEM, n = 5.

As already shown, the GpA wt (green) and its mutant G83I (blue) can be clearly distinguished in all four orientations ($p < 0.001$) in the BLaTM 1.1 system (Figure 32 B). The biggest difference of the LD₅₀ values with a factor of 10 was measured with the GpA₊₁ TMD.

The LD₅₀ measurements in the BLaTM 1.2 system of four different orientations of the GpA TMD wild type (Figure 32 B, green) and G83I mutant (blue) show, that the assay also

works with the OmpA signal peptide, as well. As GpA is a highly dimerizing TMD, an induction level of 133 μ M arabinose and 0.5 mM IPTG was chosen. The LD₅₀s of the two TMDs GpA₋₁ wt and GpA₀ wt are very low at around mutant level. Thus, the positive (wt) and negative (G83I) control can hardly be distinguished in these two cases although the difference between the GpA₋₁ wt TMD and its mutant is significant ($p < 0.05$). In contrast to this, the LD₅₀s of GpA₊₁ wt and GpA₊₂ wt are 3 to 4 times higher than their corresponding mutants ($p < 0.001$).

5 Discussion

5.1 Initial experiments

5.1.1 Optimization of protein constructs

The different parts of the BLaTM fusion proteins were optimized in respect to enzyme activity and membrane topology. The modified parts are (i) the cytoplasmic domain (EGFP or sfGFP) [227, 235] (ii) a stabilizing M182T mutation in the periplasmic β -lactamase [193] and (iii) the signal peptide.

EGFP is only active in the reducing environment of the cytoplasm but not in the oxidizing periplasm [236, 237]. sfGFP is more stable and is also fluorescent in the periplasm if it was exported via the post-translational Tat-pathway [238]. However, as the used signal peptides lack the highly conserved twin-arginine motif SRRXFLK, the proteins cannot be recognized and transported via the Tat-pathway [256]. Thus, fluorescent protein has to be cytoplasmic and functions as proof for a C_{in} -topology. As only the sfGFP containing constructs are fluorescent but not the EGFP fusion proteins (Figure 19 A), there seems to be either inappropriate folding or periplasmic localization of the EGFP proteins. Both could be influenced by the different jxtamembrane regions of the two fusion proteins. Following these findings sfGFP was chosen for further assay development.

The measurement of nitrocefin turnover provides information about the folding, the activity and thus the compartment of the β -lactamase. A high turnover reflects functional, periplasmic β -lactamase [201] and thus requires signal peptide recognition and translocation of the fusion protein to the periplasm. Only the fusion proteins containing the β -lactamase stabilizing M182T mutation can degrade nitrocefin (Figure 19 B). The M182T mutation seems to be essential for the folding and the enzymatic activity of the periplasmic, membrane-bound β -lactamase [153]. The fact that EGFP and sfGFP fusion proteins can degrade nitrocefin implies correct insertion of both proteins.

Thus, the sfGFP/M182T construct was used as a template for all following experiments, as it has the desired type I topology and an enzymatic active periplasmic β -lactamase.

5.1.2 Quantification of reconstituted β -lactamase

Up to now protein complementation of split β -lactamase was only used in the cytoplasm of *E. coli* and not in its periplasm [153]. The homo-dimerizing basic leucine zipper (bZIP) was used as a soluble domain for complementation [67, 68, 240]. Only heterodimers of the N- and C-terminal fragment of the reporter protein β -lactamase, fused to the bZIP, can confer enzymatic activity as homodimers lack essential functional domains (s. 1.5.3).

Nitrocefin hydrolysis

The β -lactam nitrocefin can be used for quantification of β -lactamase activity as it changes its absorption maximum due to hydrolysis of the β -lactam ring [194]. Indeed, it is proposed that nitrocefin can solely be used in cell lysates and not in intact cells as it is not membrane permeable [257]. Harsh lysis, such as french press or sonication, of the cells is no option for the measurement of membrane protein interactions as the complete membrane and embedded membrane proteins would be fragmented and complexed by detergents. Consequently, it was attempted to quantify active β -lactamase in the periplasm which should be accessible for nitrocefin [201].

It was proven that neither the N- nor C-fragment alone can hydrolyze nitrocefin (Figure 20 A, brown). The hydrolysis rate of the cotransformed *E. coli* (orange) is only about twice of the negative control (light blue). The measurement of the full-length β -lactamase with the same features as the fragmented constructs (green) showed a high signal. Thus, the low signal of the reconstituted fragments is not due to insufficient export to the periplasm but due to low regained enzymatic activity. One reason is the unfavored construction of the fusion-proteins as both dimerization domains are attached to the C-terminus of each β -lactamase fragment. This decreases the reconstitution efficiency [153]. Next, the bZIP domains were exchanged to a TMD-sfGFP domain to measure the reconstitution of membrane bound β -lactamase conferred by TMD-TMD interaction. Although the used model TMDs GpA₀ wt, GpA₀ G83I and Leu20 have different reconstitution propensities [110, 258], they are indistinguishable in the nitrocefin assay (Figure 23). Thus, this method cannot be used for the split β -lactamase based quantification of TMD-TMD interactions in the experimental set-up.

Ampicillin resistance

Another method for quantification of the reconstitution efficiency of split β -lactamase is the detection of ampicillin resistance due to ampicillin hydrolysis. This new approach was first tested with soluble bZIP constructs (Figure 20 B).

It was successfully demonstrated that neither the N-terminal, nor the C-terminal β -lactamase fragment alone can increase the ampicillin resistance relative to the background *E. coli* JM83 resistance. This means, neither monomers, nor homodimers of split β -lactamase generate a survival advantage. In contrast to this, coexpressing *E. coli* survive ampicillin concentrations of up to more than 200 $\mu\text{g/mL}$. This is, however, much less than expected, as it was reported that *E. coli* expressing the full length protein are resistant to ampicillin concentrations of up to 1600 $\mu\text{g/mL}$ on agar plates [198]. Here, it must be taken into account that in the used set-up (bZIP domain fused C-terminal to both split fragments) the proteins have to reconstitute in an unfavored conformation [153]. Additionally, it was shown that the resistance levels are very environment specific, for example between agar plates and liquid culture (data not shown). Nevertheless, the results were promising as the measured LD_{50} of the cotransformed cells is about 40 times higher than of cells with either only one of the fragments or without any plasmid. Thus, the soluble bZIP dimerization domain was exchanged to a TMD dimerizing domain (Figure 23 B). In this approach the GpA₀ wt and its negative mutant G83I could be clearly distinguished by a factor of more than two. The LD_{50} of the Leu20 TMD is at GpA G83I level, which means very low affinity. It was expected that the Leu20 TMD forms more dimers than the GpA G83I TMD [258].

In these primary experiments it was successfully verified, that (i) the N- and C-terminal fragment of the β -lactamase alone are inactive, (ii) that the fusion of a dimerization domain to the C-terminus of both β -lactamase fragments enables reconstitution of active β -lactamase and (iii) that the dimerization affinities can be measured by an increase of ampicillin resistance. (iv) The small errors proofed reliability of the assay.

5.1.3 Data evaluation

The fitness of *E. coli* depends on the amount of ampicillin in the periplasm. This antibiotic shows only bacteriostatic effects at low concentrations but is bacteriolytic at higher concentrations [259]. If there is enough active β -lactamase to degrade the β -lactam

antibiotic faster than it enters the periplasm and inhibits the PBP (s. 1.5.1), there will be non-inhibited growth. β -Lactamase expressing, inhibited *E. coli* will decrease the ampicillin concentration in the surrounding medium until a specific, non-inhibiting concentration is reached, allowing the cells to grow and divide. It is assumed that this mechanism is the case for the used set-up. As a consequence, the calculated LD₅₀ value will increase with longer incubation times. Hence, the LD₅₀ values, which are used for data evaluation in this work, always represent the LD₅₀ after 19 h incubation. Longer or shorter incubation times will probably shift the complete curve and lead to higher or smaller calculated LD₅₀ values, respectively. Consequently, the amount of active, reconstituted split β -lactamase is proportional to the time until the ampicillin is degraded to non-inhibiting concentration. An alternative output would be the minimal bacteriolytic concentration whose determination would require longer incubation times. The prolonged incubation time would however increase the probability of random natural resistance events in *E. coli* which causes false-positive results and is thus unsuitable.

The LD₅₀ data were obtained by the determination of the cell density at different ampicillin concentrations. For each measurement, 12 ampicillin concentrations were chosen, where the concentration steps should be as small as possible and the median concentration should be about the mean LD₅₀. This ensures exact LD₅₀ values based on fitting the data with the Hill equation (equation 1, chapter 3.15.4). The algorithm used for fitting needs several high cell densities and several low cell densities. The maximum A₅₄₄ values are up to 1.8 (Figure 21) which is far higher than the linear measurement range for light scattering [260]. Hence, the maximum value is underestimated. As the LD₅₀ value is defined as the ampicillin concentration at half of the maximum absorption, the value is not correctly defined. In this special case, the effect is negligible, because the most important output is the drop of absorption, which is usually very sharp at one distinct ampicillin concentration (Figure 21).

5.2 Linker optimization

First results with a flexible 5 amino acids linker (Figure 12 and Figure 25 A) between β -lactamase fragments and the TMD indicated that there is high orientation dependence (Figure 26 A) as well as low LD₅₀ signals (Figure 25 B). The orientation dependence might be caused by a too short linker as it could sterically restrict the two β -lactamase fragments to reconstitute if the distance or flexibility between the fixed dimerization site and the

fragments is too short. Furthermore, the nearby membrane could influence the reconstitution efficacy by its properties (e.g. charges of the lipid head groups) by incorporation of parts of the β -lactamase fragments. Hence, in the next step the linker was step-wisely extended by introduction of GGGS units. These units are characterized by their flexibility due to glycines and their solubility, which is caused by the serine (Figure 25 A) [261]. The test of four different linkers (5, 9, 13 and 25 amino acids) with the two TMDs GpA and QSOX2 (Figure 25 B,C) showed only minor differences in the LD₅₀ values of the 5 and 9 or 13 and 25 amino acids linker, respectively. Thus, a linker longer than 13 amino acids does not improve the assay further. This is in line with linker analysis using a coupled soluble FRET pair [149], which shows that there is an increase of linker flexibility up to a distinct linker length and a further elongation has only minor effects. As a tendency to lower signals is visible, it is even possible that the distance between the two β -lactamase fragments becomes too large for an efficient reconstitution of the enzyme. Unspecific interactions do not seem to increase due to longer linker, as most signals of the negative control TMDs did not increase. Another feature tested using linker variations was the disruption index, representing the impact of a mutation to the dimerization signal (Figure 27). The higher the index, the better is the discrimination of different affinities. The analysis shows that the disruption indices increase for the GpA TMD with longer linkers but for the QSOX2 TMDs, they stay almost constant. The low disruption indices of the QSOX2₀ and QSOX2₁ TMD are due to low LD₅₀s of the positive controls and not due to higher negative controls. It was proven that the background signal due to unspecific reconstitution is more or less constant for all negative control TMDs and all linker lengths.

In summary, the 13 amino acids linker is preferred for further experiments due to its superior properties in (i) orientation dependence (Figure 26 A), (ii) the absolute LD₅₀ values (Figure 25 B and C) and (iii) the average disruption indices (Figure 27). I note that the extension of the flexible linker could not abrogate the orientation dependence completely.

5.3 Heterotypic interaction

In measurements of TMD-TMD interactions, only half of the homotypic TMD dimers produce a signal as only the split β -lactamase heterodimers [$N \cdot C$] generate a signal but not the homodimers [$N \cdot N$] and [$C \cdot C$] (Figure 33). Thus, the signal generated by a hetero interaction is influenced by potential homo-interactions of the individual TMDs. If the

homotypic affinity of one TMD is very high, the equilibrium is shifted towards this homodimer and consequently the measured signal is lower. For complete interpretation of heterotypic affinity data it is always necessary to determine the homotypic affinities, too. This is the only possibility to ensure that for example a measured negative impact of one amino acid mutation on heterodimerization is really due to decreased heterotypic affinity and not due to increased homotypic affinity of the mutated TMD. This can provide important additional data.

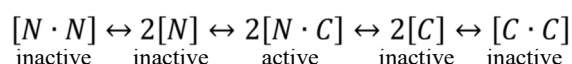


Figure 33: Possible dimers, their activity and the dependence on each other. $[N]$ N-BLa, $[C]$ C-BLa, $[N \cdot N]$ N-BLa homodimer, $[C \cdot C]$ C-BLa homodimer, $[N \cdot C]$ N-BLa/C-BLa heterodimer.

To demonstrate that the assay is also suitable for the measurement of heterotypic interactions, a well investigated model TMD was used [252]. The used LS46 TMD was evolved from a homotypic library screening and originally contains the two ionizable amino acids aspartic acid and arginine at position six and seven. Furthermore, there is an essential GxxxG motif at the C-terminus of the TMD. The D_6 of one TMD interacts with the R_7 of another (Figure 28), which generates a heterotypic interaction motif with low homotypic affinities. Assuming that there is a D_6 on one TMD and an R_7 on the other, the counterpart in homodimers would be an inert L. To demonstrate the effect of D or R on homodimerization, the interface was extended to D_6D_7 or R_6R_7 so that there is always a complementary, possibly interacting amino acid [262]. It was expected that the heterodimer D_6D_7/R_6R_7 should be stronger than the D_6/R_7 pair, as there are two D/R interaction sites.

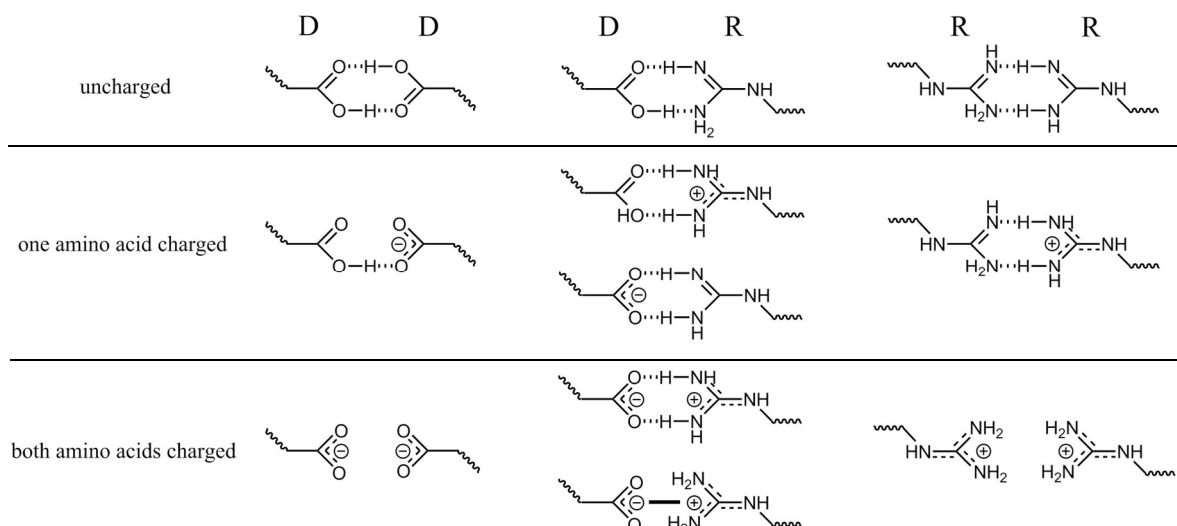


Figure 34: Intermolecular interactions between the carboxyl group of aspartic acid (D) and the guanidinium group of arginine (R) in different charge states. Hashed lines: Hydrogen bonds, bold line: Salt bridge. Uncharged: Medium strong hydrogen bonds, one amino acid charged: Strong hydrogen bonds due to the negative charge of the oxygen or stronger polarization of the amino hydrogen; both amino acids charged: repulsion between amino acids of the same kind due to same charge; very strong hydrogen bonds or salt bridge for heterotypic interacting D and R.

The following results were obtained: First, the LS46₁ TMD and the L19GG_D6R₇ TMD show the same LD₅₀ which is contrary to the published data obtained with ToxR, which is ascribed to the different methods [252]. Nevertheless, the contribution of the ionic amino acid pair D₆R₇ to the interaction becomes apparent.

Second, the D₆/D₆ pair interacts significantly better than the control, which suggests that the side chains form hydrogen bonds [263] (Figure 34). The pK_a of the β-carboxyl group is 3.90 in water [264] but it is unlikely that it is deprotonated in the apolar membrane as well. Nevertheless, it was shown in the ζζ_{TM} homodimer, which is essential for its assembly with the T-cell-receptor, that the two aspartic acids of the dimerization interface share one proton, resulting in -1 net charge [265, 266]. The BLaTM assay cannot determine the charge state in the used model TMD but the net charge is probably 0 or -1 as interaction was measured (Figure 34). The D₆D₇/D₆D₇ pair dimerizes even better, as two carboxyl groups of each TMD can form hydrogen bonds.

Third, arginine containing TMDs do not mediate homotypic interactions. The pK_a, and thus the charge state, of the guanidinium group is very solvent dependent. The pK_a is 12.48 in water [267] but decreases to less than 2 in the center of cholesterol containing membranes. It depends on the position in the membrane, the saturation of the acyl chains and the lipid head groups [268]. Nevertheless, our results indicate that the side chain is protonated as

there is no interaction measured (Figure 30 and Figure 34). Probably, the lipids in the cytoplasmic membrane of *E. coli* stabilize the protonated state of the guanidinium group [4, 268-271]. Additionally, the arginine is not in the center of the TMD which could result in snorkeling effects, which stabilize the positive charge as well [272].

Forth, the heterotypic interactions of the TMD combinations D₆/R₇ and D₆D₇/R₆R₇ were more than 3 times stronger than the negative control (L19GG) and up to 1.5 times higher than the positive control (L19GG_D₆R₇). This indicates strong heterotypic interactions, and, due to low amount of homodimers, a high content of active heterodimers (Figure 33). The charge state of the two amino acids is unclear. However it is unlikely that both amino acids are completely uncharged because in close proximity a proton will migrate from the aspartic acid to the arginine resulting in a negatively charged carboxyl group and a positively charged guanidinium group. The two unfavoured charges can stabilize each other in the apolar environment and the pK_a values are more similar to ones in aqueous solvent. These charged groups can strongly interact via two very strong hydrogen bonds or one salt-bridge. Additionally, the combination “uncharged aspartic acid/charged arginine” is possible as the hydrogens of the positively charged guanidinium group are strongly polarized which results in strong hydrogen bonds. However, the single charge in the apolar environment is unfavored and less likely. Two oppositely charged amino acids do not further increase the LD₅₀. Probably, the additional artificial R₆/D₇ interface does not have the same impact on the dimerization as the D₆/R₇, or the heterotypic TMD-TMD affinity with one interacting pair is so strong, that the signal is already saturated.

5.4 Orientation dependence

In this work three different TMDs – GpA, QSOX2 and LS46 – were intensively investigated. In all three cases, there was orientation dependence between 0.4 and 0.6 with the 13 amino acids linker (Figure 26 A). For analysis, the interfacial amino acids in their position of the best dimerizing orientation were plotted against their position in the TMD (Figure 26 B) [77, 246, 252]. The analysis of all tested TMDs shows, that there is a periodicity of the interacting amino acids of 3 to 4 which corresponds to the 3.6 periodicity of α -helices (s. 1.3.3) [12]. Therefore, there appears to be a preferred interfacial side of the TMDs. This could be used for the prediction of probable interfacial amino acids from the signal representing homotypic interaction in the four orientations. This prediction would make the identification of interacting amino acids more efficient by reducing the number

of residues to be tested. However, more TMDs have to be analyzed for the establishment of this prediction method. Extending this approach to heterotypically interacting TMDs would require the systematic testing of four against four different orientations, resulting in 16 combinations.

5.5 Characterization of β -lactamase fusion proteins

The Western blot analysis of the BLaTM 1.1 protein expression showed a very similar protein expression for both N-BLa and C-BLa proteins of all GpA and QSOX2 constructs, (Figure 24). The N-BLa protein, including the signal peptide, has a molecular weight of 54.5 kDa, the C-BLa protein of 45.5 kDa. The migration of the proteins is not as expected, which is typical for membrane proteins [251]. SDS-PAGE is sensitive to hydrophobic amino acid stretches, such as TMDs or signal peptides. These regions bind more SDS, which alters the migration of the protein in the gel, also called “gel shifting”. In addition, the unspecifically cleaved off C-terminal GFP fragment, containing the FLAG epitope and probably the rigid linker (~ 30 kDa) was detected in a very similar amount. Assuming an N_{out} TM topology, it was expected that the 2 kDa N-terminal signal peptide would be cleaved off by a signal peptidase (s. 1.2.2). To confirm signal peptide removal, control proteins with cut signal peptide were constructed whose molecular weight, determined by SDS-PAGE, should be equal to the molecular mass of correctly inserted proteins. This expectation was met by the C-BLa proteins but not by the N-BLa proteins, which are clearly heavier than the control proteins without signal peptide. The C-BLa protein as either not recognized by the signal peptidase (s. 1.2.2), or the molecular mass shift after signal peptide cleavage is only visible for the N-BLa proteins.

To ascertain that this problem does not simply arise from using the unusual signal peptide of pectin lyase B (PelB, UniProtKB) from *Aspergillus niger*, which is a fungus, the resistance measurements, as well as the quantification of periplasmic full-length β -lactamase by nitrocefin turnover, showed, that at least some of the proteins were exported to the periplasm or incorporated into the inner *E. coli* membrane, respectively. On this account, the signal peptide was modified by optimizing the signal peptide cleavage site, which did neither alter the Western blot, nor the ampicillin resistance values.

The complete signal peptide was exchanged to the commonly used signal peptide of pectate lyase 2 (Pel2) from *Erwinia carotovora* [273], the original β -lactamase signal

peptide (TEM-1) [274] or the signal peptide of Outer membrane protein A (OmpA) from *E. coli* [275].

According to Western blot analysis, all these signal peptides dramatically decreased the protein expression to beyond the detection limit. The Pel2 and TEM-1 signal peptide containing constructs could not generate any ampicillin resistance at all, whereas the constructs containing the OmpA signal peptide convey extremely high resistance above detection limit using the established protocol. Accordingly, the amount of protein necessary for gaining ampicillin resistance is very small and it is very likely that most of both proteins containing the PelB signal peptide is not inserted into the membrane but aggregated as inclusion bodies. However, as ampicillin is not membrane permeable these aggregates cannot generate false positive signals due to unspecific β -lactamase reconstitution. It is not clear why the expression of PelB proteins is generally higher, or why and how a signal peptide can influence protein expression.

5.6 Tuning BLaTM for different TMD-TMD affinities

For the OmpA signal peptide containing proteins, hereinafter called BLa 1.2 proteins, the induction strength had to be reduced to avoid unspecific reconstitution of the split β -lactamase insertion. The arabinose concentration was decreased by a factor of 10 to 133 μ M, which resulted in still very high, yet sequence-specific ampicillin resistance values (GpA₊₁ wt vs GpA₊₁ G83I) and is at the border to full activation of the araBAD promotor [276]. The promotor p_{BAD} used here is assumed to have an all-or-none behavior, which means that protein expression is either on or off in a given cell. Additionally, there is a maintenance behavior, so that cells which started expression stay activated due to auto-activating themselves, by regulator protein upregulation [222, 277]. In the BLaTM 1.1 system this could be ignored, as the protein expression was induced by a high arabinose concentration, which constantly activated the promoters in all cells [277]. Expression in the BLaTM 1.2 assay was reduced by repressing the induction with isopropyl β -D-1-thiogalactopyranoside (IPTG). IPTG is a competitive inhibitor of the regulator protein AraC, because it inhibits the transcription in a dose dependent manner [222, 255]. The inhibition can be circumvented by high arabinose concentrations or by distinct mutations in the AraC protein, which indicates a direct impact of IPTG on the regulator protein [255]. The impact of IPTG concentration on β -lactamase expression was proven by quantification of protein expression by GFP fluorescence and of split β -lactamase reconstitution

efficiency via ampicillin resistance measurements (Figure 31). Protein expression for different GpA TMD orientations is indirectly proportional to the IPTG concentration. Orientation-dependence and the effect of the G83I mutation are preserved. Driving expression of GpA₀ wt and GpA₀ G83I TMD by decreasing IPTG concentrations allowed to distinguish the LD₅₀ values associated with them. This means that a relatively low affinity TMD-TMD interaction can be measured at higher protein concentration.

The numerical K_d values cannot be determined, because this would require determining the concentration of both β -lactamase proteins in the bacterial membrane. Further, this would require reaching saturating LD₅₀ values which has not been tested here.

This experiment showed that the induction can be adjusted to different affinities of the investigated TMDs. High affinity TMD-TMD interactions can be measured at low expression levels to get a signal as stable as possible. Less well interacting TMDs can be measured at a less inhibiting IPTG concentration, to get higher protein concentrations. High LD₅₀ values, as seen for GpA₊₁ wt at 0.3 mM IPTG, go along with higher errors on the one hand. This is because the signal is more sensitive to small IPTG and arabinose concentration differences or protein expression, respectively.

Summarizing, new TMD pairs should be investigated first by intermediate induction, such as 133 μ M arabinose and 0.3 mM IPTG. For further detailed analysis, such as alanine scanning, the IPTG concentration should be adjusted to reach high signal separation with low errors. It should also be noted that although the cleaved off GFP cannot be distinguished from the full length proteins by fluorescence measurements, a link between induction and amount of periplasmic split β -lactamase can be assumed, because the ratio of the two proteins should be stable.

5.7 Comparison of BLaTM 1.1 and BLaTM 1.2

The high induction of gene transcription in the BLaTM 1.1 system starts full protein expression in all *E. coli* cells in the culture [276, 277]. This causes inclusion bodies which are an indicator of insufficient protein folding or too high protein concentrations [278]. Due to suboptimal protein recognition of the membrane proteins the proteins can accumulate in the cytoplasm and form such aggregates. Reduction of the expression strength led to non-differentiable low LD₅₀ values. The insufficient membrane integration only allows the measurement of high and middle affine TMD-TMD interactions, such as

for example GpA wt, QSOX2 wt or LS46 in their optimal orientation (Figure 26 and Figure 29). Low affinity TMD-TMD pairs, such as for example GpA G83I and Leu20 (Figure 23 B) cannot be distinguished, probably because the protein concentrations in the inner membrane are below their K_d . Additionally, high affinity TMDs and their mutants can hardly be distinguished if they are measured in a low-signal orientation (Figure 26). The switch to the BLaTM 1.2 system allows, additionally to the measurement of high- and middle-affine TMDs, the quantification of low-signal interactions. It has to be noted that probably the use of the additional competitive inhibitor IPTG of the inducer arabinose in the BLaTM 1.2 system, results in higher errors compared to the BLaTM 1.1 system. Summarizing, both systems are able to efficiently measure TMD-TMD interactions, but the BLaTM 1.2 system is more flexible and can be adjusted to TMD pairs of different affinities.

5.8 Comparison to established assays

One major difficulty in the development of BLaTM was the orientation-dependence. However, it was also reported for the ToxR and GALLEX assay [77, 108, 121] and is not a unique characteristic of the BLaTM system. There are no published results about orientation dependence, neither positive nor negative, in the AraTM, BACTH, MYTH or MaMTH systems. Beside this, BLaTM improved existing methods for the measurement of TMD-TMD interactions in different aspects. First, instead of colorimetric or fluorogenic detection methods [152, 153], BLaTM uses acquired antibiotic resistance of *E. coli* as a quantitative output. The LD_{50} directly depends on the concentration and activity of reconstituted split β -lactamase and accordingly on the affinity of both interacting TMDs fused to the β -lactamase fragments. Thus, there is a minimal number of steps from dimerization to an activation of the reporter enzyme β -lactamase. By contrast, other methods, such as GALLEX or BACTH, are based on a cascade of events [42, 108, 109]. Second, antibiotic resistance leads to survival of bacterial cells which makes the BLaTM assay perfectly suitable for screenings, where high-affinity TMD pairs can be isolated from combinatorial libraries. Third, the β -lactamase substrate ampicillin is not membrane permeable. Thus, only the reconstitution of the split β -lactamase fragments in the periplasm, after targeting the proteins into the inner *E. coli* membrane by a cleavable N-terminal signal peptide, produces ampicillin resistance. In contrast to the existing TMD-TMD interaction assays, proteins that are not membrane-integrated proteins, aggregated or

localized in the cytoplasm cannot cause false-positive results (s. 1.3.2). Forth, superfolder GFP (sfGFP) [227] is fused to the C-terminus of the TMD at its cytoplasmic end to enhance the stability, improve membrane insertion and allow direct protein quantification. It was shown that sfGFP is fluorescent in the oxidizing periplasm only if it was exported by the post-translational Tat-pathway, or by the co-translational SRP-dependent SecYEG pathway and not by the post-translational SecA-dependent pathway (s. 1.2.2) [37, 238, 279-281]. This indicates, that the sfGFP folding kinetic is too fast for the post-translational SecYEG dependent translocation, as it cannot be recognized by SecA and is thus, not translocated through the translocon [238]. For this reason, β -lactamase fusion proteins with their C-terminal sfGFP cannot be completely transferred into the periplasm, but are forced to incorporate into the inner *E. coli* membrane as shown for multi-pass proteins [24]. This forced membrane integration avoids false positive signals if a TMD is too hydrophilic for membrane insertion and interacts in the periplasm (s. 1.2.3). In case of complete backward movement of the fusion protein into the cytoplasm, no ampicillin resistance would exist. In that case this would be an indicator for a too hydrophilic and thus not suitable TMD. Fifth, in contrast to other systems, there is no need for distinct features of the used *E. coli* strain as for example in the BACTH assay [109]. This flexibility allows the investigation of other features such as the influence of an altered lipid composition on the affinity of a TMD pair [101, 282, 283].

5.9 Conclusion and outlook

The BLaTM assay is a novel method to measure the dimerization of homo- and heterotypically interacting TMDs. It is based on the TMD-dependent reconstitution of split TEM-1 β -lactamase which mediates ampicillin resistance in its active form. Accordingly, the resistance level, represented by the LD₅₀, is directly proportional to the dimerization propensity. Thus, this method allows the very direct measurement of TMD-TMD interactions. BLaTM is an important tool to improve the understanding of intra-membranous protein-protein interactions.

On the basis of well investigated TMD pairs, it was successfully demonstrated, that the BLaTM assay can reproducibly measure TMD-TMD interactions. Nevertheless, this assay can be improved and modified to extend the areas of application or the throughput. At the moment one complete 12-well plate is needed for each data point to gain exact values. It should be possible to increase the number of samples per plate by upscaling to 24- or 48-

well plates. The shift to the 96-well format is more complicated, as the system and the protocol has to be converted to deep-well plates where, for example, the oxygen supply of *E. coli* cultures is strongly reduced compared to plates with larger wells.

One application is the creation of TMD libraries and the subsequent screening for high-interacting TMDs on ampicillin containing agar plates. As the two proteins are coded separately, the assay is especially precious for the screening of heterotypic interacting TMDs. Up to now, there is no other method capable of achieving this. Further, the BLaTM method can be used to identify specific peptide inhibitors against a TMD dimerization interface. Consequently, this method could be used to find inhibitors against distinct, pathogenic mutants of a TMD, due to the high specificity of TMD-TMD interactions. One example is the Alzheimer precursor protein (APP), which can be proteolysed by the γ -secretase as a monomer but not as a dimer [284]. Following, the formation of stable APP/peptide complexes would reduce the formation of amyloid beta and thus of amyloid plaques.

Furthermore, the designed fusion proteins can be directly purified and used for in vitro experiments, for example in order to determine the influence of different lipids to the dimerization. Here, the interaction can be quantified by nitrocefin turnover, or by anisotropy measurements using the C-terminal sfGFP. Up to now, the assay is limited to parallel TMD pairs, but can be easily adjusted to antiparallel TMD pairs, by the creation of a new, type II fusion protein. It would be the first assay to measure this kind of interactions in vivo. This could be used to find totally new interacting TMD pairs and give insights to the function of protein complexes or substrate recognition in a cellular membrane. Intramolecular, antiparallel interactions are very common in multi-pass proteins, but there are also intermolecular, antiparallel interactions, such as in the small multidrug resistance protein EmrE [285, 286]. As very little is known about this interaction mode, it would be worth to extend the assay to this application.

The cytoplasmic domain sfGFP is not necessary for the generation of the output signal, as this is generated by the periplasmic domains. Thus, it could be exchanged to another functional domain. For example it could be used to quantify the homodimerization propensity of the TMDs in the same experiment. One possibility is the combination with the AraTM system which utilizes AraC as a sensor protein. This would need another expression system such as T7 and a third plasmid containing a pBAD promotor and a

reporter protein such as β -galactosidase. Another possibility would be the use of the sfGFP for homo-FRET measurements to quantify homodimers, however this would need extensive improvements regarding the protein expression and an almost complete establishment of a novel method. A fusion protein, which allows the control of membrane integration would be desirable, but is hard to implement because the sensor protein has to be on the cytoplasmic side of the membrane.

Furthermore, the IPTG-dependent adjustable protein expression is a tool to control protein expression in each cell and reduces the problem of the on/off behavior of the p_{BAD} promoter, as it inhibits the transcription start itself. In the current method, this allows the adjustment of the assay to different K_{ds} of the investigated TMD pairs. This method could be transferred to other protein concentration dependent systems as well to reduce unspecific signal due to protein overexpression. Moreover, it is useful for the expression of toxic proteins to reduce the selection for low or non-expressing *E. coli* in protein expression cultures.

Summing up, the current assay can be used for the reliable quantification of homo- and heterotypic TMD-TMD interactions with the possibility of fine tuning to different TMD affinities and can be adapted to a heterotypic TMD screening which is not possible with conventional methods.

6 Bibliography

- [1] Hannich JT, Umebayashi K, Riezman H. Distribution and functions of sterols and sphingolipids. *Cold Spring Harb Perspect Biol.* 2011;3.
- [2] Opekarova M, Tanner W. Specific lipid requirements of membrane proteins--a putative bottleneck in heterologous expression. *Biochimica et biophysica acta.* 2003;1610:11-22.
- [3] White SH, Wimley WC. Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct.* 1999;28:319-65.
- [4] Morein S, Andersson AS, Rilfors L, Lindblom G. Wild-type *Escherichia coli* cells regulate the membrane lipid composition in a "window" between gel and non-lamellar structures. *Journal of Biological Chemistry.* 1996;271:6801-9.
- [5] Stangl M, Schneider D. Functional competition within a membrane: Lipid recognition vs. transmembrane helix oligomerization. *Biochimica et biophysica acta.* 2015;1848:1886-96.
- [6] Lomize MA, Lomize AL, Pogozheva ID, Mosberg HI. OPM: orientations of proteins in membranes database. *Bioinformatics.* 2006;22:623-5.
- [7] Wang X, Yang C, Chai J, Shi Y, Xue D. Mechanisms of AIF-mediated apoptotic DNA degradation in *Caenorhabditis elegans*. *Science.* 2002;298:1587-92.
- [8] Wallin E, von Heijne G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein science : a publication of the Protein Society.* 1998;7:1029-38.
- [9] Fagerberg L, Jonasson K, von Heijne G, Uhlen M, Berglund L. Prediction of the human membrane proteome. *Proteomics.* 2010;10:1141-9.
- [10] Arinaminpathy Y, Khurana E, Engelman DM, Gerstein MB. Computational analysis of membrane proteins: the largest class of drug targets. *Drug Discov Today.* 2009;14:1130-5.
- [11] Berman HM. The Protein Data Bank. *Nucleic Acids Research.* 2000;28:235-42.
- [12] Berg JM, Tymoczko JL, Stryer L. *Biochemie.* 6 ed. Heidelberg: Spektrum Akademischer Verlag; 2007.
- [13] Dupuy AD, Engelman DM. Protein area occupancy at the center of the red blood cell membrane. *Proceedings of the National Academy of Sciences of the United States of America.* 2008;105:2848-52.
- [14] Tamm LK, Hong H, Liang B. Folding and assembly of beta-barrel membrane proteins. *Biochimica et biophysica acta.* 2004;1666:250-63.
- [15] Blobel G. Intracellular protein topogenesis. *Proceedings of the National Academy of Sciences of the United States of America.* 1980;77:1496-500.
- [16] Singer SJ. The structure and insertion of integral proteins in membranes. *Annu Rev Cell Biol.* 1990;6:247-96.
- [17] Heijne G. The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *Embo J.* 1986;5:3021-7.
- [18] von Heijne G, Gavel Y. Topogenic signals in integral membrane proteins. *European journal of biochemistry / FEBS.* 1988;174:671-8.

- [19] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982;157:105-32.
- [20] Eisenberg D, Schwarz E, Komaromy M, Wall R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol.* 1984;179:125-42.
- [21] Zhao G, London E. An amino acid "transmembrane tendency" scale that approaches the theoretical limit to accuracy for prediction of transmembrane helices: relationship to biological hydrophobicity. *Protein science : a publication of the Protein Society.* 2006;15:1987-2001.
- [22] Hessa T, Kim H, Bihlmaier K, Lundin C, Boekel J, Andersson H, et al. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature.* 2005;433:377-81.
- [23] Hessa T, Meindl-Beinker NM, Bernsel A, Kim H, Sato Y, Lerch-Bader M, et al. Molecular code for transmembrane-helix recognition by the SecE1 translocon. *Nature.* 2007;450:1026-30.
- [24] Virkki MT, Peters C, Nilsson D, Sorensen T, Cristobal S, Wallner B, et al. The positive inside rule is stronger when followed by a transmembrane helix. *J Mol Biol.* 2014;426:2982-91.
- [25] Peters C, Tsirigos KD, Shu N, Elofsson A. Improved topology prediction using the terminal hydrophobic helices rule. *Bioinformatics.* 2015.
- [26] Collinson I, Corey RA, Allen WJ. Channel crossing: how are proteins shipped across the bacterial plasma membrane? *Philos Trans R Soc Lond B Biol Sci.* 2015;370.
- [27] Thie H, Schirrmann T, Paschke M, Dubel S, Hust M. SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in *E. coli*. *N Biotechnol.* 2008;25:49-54.
- [28] von Heijne G. Signal sequences. The limits of variation. *J Mol Biol.* 1985;184:99-105.
- [29] von Heijne G. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 1986;14:4683-90.
- [30] Kaiser CA, Preuss D, Grisafi P, Botstein D. Many random sequences functionally replace the secretion signal sequence of yeast invertase. *Science.* 1987;235:312-7.
- [31] Shaffer KL, Sharma A, Snapp EL, Hegde RS. Regulation of protein compartmentalization expands the diversity of protein function. *Dev Cell.* 2005;9:545-54.
- [32] Rane NS, Yonkovich JL, Hegde RS. Protection from cytosolic prion protein toxicity by modulation of protein translocation. *Embo J.* 2004;23:4550-9.
- [33] Li Y, Bergeron JJ, Luo L, Ou WJ, Thomas DY, Kang CY. Effects of inefficient cleavage of the signal sequence of HIV-1 gp 120 on its association with calnexin, folding, and intracellular transport. *Proceedings of the National Academy of Sciences of the United States of America.* 1996;93:9606-11.
- [34] Bae T, Schneewind O. The YSIRK-G/S motif of staphylococcal protein A and its role in efficiency of signal peptide processing. *J Bacteriol.* 2003;185:2910-9.
- [35] Ulbrecht M, Martinozzi S, Grzeschik M, Hengel H, Ellwart JW, Pla M, et al. Cutting edge: the human cytomegalovirus UL40 gene product contains a ligand for HLA-E and prevents NK cell-mediated lysis. *Journal of immunology.* 2000;164:5019-22.
- [36] Eichler R, Lenz O, Strecker T, Eickmann M, Klenk HD, Garten W. Identification of Lassa virus glycoprotein signal peptide as a trans-acting maturation factor. *EMBO Rep.* 2003;4:1084-8.
- [37] Lee HC, Bernstein HD. The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proceedings of the National Academy of Sciences of the United States of America.* 2001;98:3471-6.
- [38] Karamyshev AL, Johnson AE. Selective SecA association with signal sequences in ribosome-bound nascent chains: a potential role for SecA in ribosome targeting to the bacterial membrane. *The Journal of biological chemistry.* 2005;280:37930-40.

- [39] Nilsson I, Lara P, Hessa T, Johnson AE, von Heijne G, Karamyshev AL. The code for directing proteins for translocation across ER membrane: SRP cotranslationally recognizes specific features of a signal sequence. *J Mol Biol.* 2015;427:1191-201.
- [40] Zhang Z, Henzel WJ. Signal peptide prediction based on analysis of experimentally verified cleavage sites. *Protein science : a publication of the Protein Society.* 2004;13:2819-24.
- [41] Saraogi I, Akopian D, Shan SO. A tale of two GTPases in cotranslational protein targeting. *Protein science : a publication of the Protein Society.* 2011;20:1790-5.
- [42] Sawma P, Roth L, Blanchard C, Bagnard D, Cremel G, Bouveret E, et al. Evidence for new homotypic and heterotypic interactions between transmembrane helices of proteins involved in receptor tyrosine kinase and neuropilin signaling. *J Mol Biol.* 2014;426:4099-111.
- [43] Su PC, Berger BW. Identifying key juxtamembrane interactions in cell membranes using AraC-based transcriptional reporter assay (AraTM). *The Journal of biological chemistry.* 2012;287:31515-26.
- [44] Gelis I, Bonvin AM, Keramisanou D, Koukaki M, Gouridis G, Karamanou S, et al. Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell.* 2007;131:756-69.
- [45] Randall LL, Topping TB, Smith VF, Diamond DL, Hardy SJ. SecB: a chaperone from *Escherichia coli*. *Methods in enzymology.* 1998;290:444-59.
- [46] Watanabe M, Blobel G. SecB functions as a cytosolic signal recognition factor for protein export in *E. coli*. *Cell.* 1989;58:695-705.
- [47] Hoffmann A, Becker AH, Zachmann-Brand B, Deuerling E, Bukau B, Kramer G. Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. *Molecular cell.* 2012;48:63-74.
- [48] Lill R, Cunningham K, Brundage LA, Ito K, Oliver D, Wickner W. SecA protein hydrolyzes ATP and is an essential component of the protein translocation ATPase of *Escherichia coli*. *Embo J.* 1989;8:961-6.
- [49] Patel R, Smith SM, Robinson C. Protein transport by the bacterial Tat pathway. *Biochimica et biophysica acta.* 2014;1843:1620-8.
- [50] Sanders C, Wethkamp N, Lill H. Transport of cytochrome c derivatives by the bacterial Tat protein translocation system. *Mol Microbiol.* 2001;41:241-6.
- [51] DeLisa MP, Samuelson P, Palmer T, Georgiou G. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *The Journal of biological chemistry.* 2002;277:29825-31.
- [52] Berks BC. A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol.* 1996;22:393-404.
- [53] Van den Berg B, Clemons WM, Jr., Collinson I, Modis Y, Hartmann E, Harrison SC, et al. X-ray structure of a protein-conducting channel. *Nature.* 2004;427:36-44.
- [54] Kihara A, Akiyama Y, Ito K. Ftsh Is Required for Proteolytic Elimination of Uncomplexed Forms of SecY, an Essential Protein Translocase Subunit. *Proceedings of the National Academy of Sciences of the United States of America.* 1995;92:4532-6.
- [55] Hizlan D, Robson A, Whitehouse S, Gold VA, Vonck J, Mills D, et al. Structure of the SecY complex unlocked by a preprotein mimic. *Cell reports.* 2012;1:21-8.
- [56] Junne T, Kocik L, Spiess M. The hydrophobic core of the Sec61 translocon defines the hydrophobicity threshold for membrane integration. *Mol Biol Cell.* 2010;21:1662-70.

- [57] Pitonzo D, Yang Z, Matsumura Y, Johnson AE, Skach WR. Sequence-specific retention and regulated integration of a nascent membrane protein by the endoplasmic reticulum Sec61 translocon. *Mol Biol Cell*. 2009;20:685-98.
- [58] Egea PF, Stroud RM. Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;107:17182-7.
- [59] Cymer F, von Heijne G, White SH. Mechanisms of integral membrane protein insertion and folding. *J Mol Biol*. 2015;427:999-1022.
- [60] Stone TA, Schiller N, von Heijne G, Deber CM. Hydrophobic blocks facilitate lipid compatibility and translocon recognition of transmembrane protein sequences. *Biochemistry*. 2015;54:1465-73.
- [61] Papewalis J, Nikitin A, Rajewsky MF. G to A polymorphism at amino acid codon 655 of the human erbB-2/HER2 gene. *Nucleic Acids Res*. 1991;19:5452.
- [62] Xie D, Shu XO, Deng Z, Wen WQ, Creek KE, Dai Q, et al. Population-based, case-control study of HER2 genetic polymorphism and breast cancer risk. *J Natl Cancer Inst*. 2000;92:412-7.
- [63] Jones AM, Xuan Y, Xu M, Wang RS, Ho CH, Lalonde S, et al. Border control--a membrane-linked interactome of Arabidopsis. *Science*. 2014;344:711-6.
- [64] Babu M, Vlasblom J, Pu S, Guo X, Graham C, Bean BD, et al. Interaction landscape of membrane-protein complexes in *Saccharomyces cerevisiae*. *Nature*. 2012;489:585-9.
- [65] Walters RF, DeGrado WF. Helix-packing motifs in membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 2006;103:13658-63.
- [66] Langosch D, Heringa J. Interaction of transmembrane helices by a knobs-into-holes packing characteristic of soluble coiled coils. *Proteins*. 1998;31:150-9.
- [67] Landschulz WH, Johnson PF, McKnight SL. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science*. 1988;240:1759-64.
- [68] O'Shea EK, Rutkowski R, Kim PS. Evidence that the leucine zipper is a coiled coil. *Science*. 1989;243:538-42.
- [69] Udgaonkar JB. Polypeptide chain collapse and protein folding. *Arch Biochem Biophys*. 2013;531:24-33.
- [70] Adamian L, Liang J. Interhelical hydrogen bonds and spatial motifs in membrane proteins: polar clamps and serine zippers. *Proteins*. 2002;47:209-18.
- [71] Adamian L, Nanda V, DeGrado WF, Liang J. Empirical lipid propensities of amino acid residues in multispan alpha helical membrane proteins. *Proteins*. 2005;59:496-509.
- [72] Ried CL, Kube S, Kirrbach J, Langosch D. Homotypic interaction and amino acid distribution of unilaterally conserved transmembrane helices. *J Mol Biol*. 2012;420:251-7.
- [73] Russ WP, Engelman DM. The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol*. 2000;296:911-9.
- [74] Teese MG, Langosch D. Role of GxxxG Motifs in Transmembrane Domain Interactions. *Biochemistry*. 2015;54:5125-35.
- [75] Hong H. Toward understanding driving forces in membrane protein folding. *Arch Biochem Biophys*. 2014;564:297-313.
- [76] Lemmon MA, Flanagan JM, Treutlein HR, Zhang J, Engelman DM. Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry*. 1992;31:12719-25.
- [77] Langosch D, Brosig B, Kolmar H, Fritz HJ. Dimerisation of the glycoporphin A transmembrane segment in membranes probed with the ToxR transcription activator. *J Mol Biol*. 1996;263:525-30.

- [78] Li R, Gorelik R, Nanda V, Law PB, Lear JD, DeGrado WF, et al. Dimerization of the transmembrane domain of Integrin alphaIIb subunit in cell membranes. *The Journal of biological chemistry*. 2004;279:26666-73.
- [79] Lau TL, Kim C, Ginsberg MH, Ulmer TS. The structure of the integrin alphaIIb beta3 transmembrane complex explains integrin transmembrane signalling. *Embo J*. 2009;28:1351-61.
- [80] Li R, Mitra N, Gratkowski H, Vilaire G, Litvinov R, Nagasami C, et al. Activation of integrin alphaIIb beta3 by modulation of transmembrane helix associations. *Science*. 2003;300:795-8.
- [81] Yin H, Litvinov RI, Vilaire G, Zhu H, Li W, Caputo GA, et al. Activation of platelet alphaIIb beta3 by an exogenous peptide corresponding to the transmembrane domain of alphaIIb. *The Journal of biological chemistry*. 2006;281:36732-41.
- [82] Frey AJ, Ibrahim S, Gleim S, Hwa J, Smyth EM. Biased suppression of TP homodimerization and signaling through disruption of a TM GxxxGxxxL helical interaction motif. *J Lipid Res*. 2013;54:1678-90.
- [83] Senes A, Gerstein M, Engelman DM. Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J Mol Biol*. 2000;296:921-36.
- [84] Call ME, Wucherpennig KW, Chou JJ. The structural basis for intramembrane assembly of an activating immunoreceptor complex. *Nat Immunol*. 2010;11:1023-9.
- [85] Wei P, Zheng BK, Guo PR, Kawakami T, Luo SZ. The association of polar residues in the DAP12 homodimer: TOXCAT and molecular dynamics simulation studies. *Biophys J*. 2013;104:1435-44.
- [86] North B, Cristian L, Fu Stowell X, Lear JD, Saven JG, DeGrado WF. Characterization of a membrane protein folding motif, the Ser zipper, using designed peptides. *J Mol Biol*. 2006;359:930-9.
- [87] Ruan W, Becker V, Klingmuller U, Langosch D. The interface between self-assembling erythropoietin receptor transmembrane segments corresponds to a membrane-spanning leucine zipper. *The Journal of biological chemistry*. 2004;279:3273-9.
- [88] Popot JL, Engelman DM. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry*. 1990;29:4031-7.
- [89] Engelman DM. Electrostatic fasteners hold the T cell receptor-CD3 complex together. *Molecular cell*. 2003;11:5-6.
- [90] Call ME, Pyrdol J, Wiedmann M, Wucherpennig KW. The organizing principle in the formation of the T cell receptor-CD3 complex. *Cell*. 2002;111:967-79.
- [91] Schobert B, Cupp-Vickery J, Hornak V, Smith S, Lanyi J. Crystallographic structure of the K intermediate of bacteriorhodopsin: conservation of free energy after photoisomerization of the retinal. *J Mol Biol*. 2002;321:715-26.
- [92] Marti T, Rosselet SJ, Otto H, Heyn MP, Khorana HG. The retinylidene Schiff base counterion in bacteriorhodopsin. *The Journal of biological chemistry*. 1991;266:18674-83.
- [93] Herrmann JR, Panitz JC, Unterreitmeier S, Fuchs A, Frishman D, Langosch D. Complex patterns of histidine, hydroxylated amino acids and the GxxxG motif mediate high-affinity transmembrane domain interactions. *J Mol Biol*. 2009;385:912-23.
- [94] Lee JI, Hwang PP, Hansen C, Wilson TH. Possible salt bridges between transmembrane alpha-helices of the lactose carrier of *Escherichia coli*. *The Journal of biological chemistry*. 1992;267:20758-64.
- [95] Kim JM, Altenbach C, Kono M, Oprian DD, Hubbell WL, Khorana HG. Structural origins of constitutive activation in rhodopsin: Role of the K296/E113 salt bridge. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101:12508-13.

- [96] Hertel B, Tayefeh S, Kloss T, Hewing J, Gebhardt M, Baumeister D, et al. Salt bridges in the miniature viral channel Kcv are important for function. *Eur Biophys J.* 2010;39:1057-68.
- [97] Johnson RM, Hecht K, Deber CM. Aromatic and cation-pi interactions enhance helix-helix association in a membrane environment. *Biochemistry.* 2007;46:9208-14.
- [98] Ridder A, Skupjen P, Unterreitmeier S, Langosch D. Tryptophan supports interaction of transmembrane helices. *J Mol Biol.* 2005;354:894-902.
- [99] Unterreitmeier S, Fuchs A, Schaffler T, Heym RG, Frishman D, Langosch D. Phenylalanine promotes interaction of transmembrane domains via GxxxG motifs. *J Mol Biol.* 2007;374:705-18.
- [100] Baker RP, Urban S. Architectural and thermodynamic principles underlying intramembrane protease function. *Nature chemical biology.* 2012;8:759-68.
- [101] Song Y, Hustedt EJ, Brandon S, Sanders CR. Competition between homodimerization and cholesterol binding to the C99 domain of the amyloid precursor protein. *Biochemistry.* 2013;52:5051-64.
- [102] Munter LM, Voigt P, Harmeier A, Kaden D, Gottschalk KE, Weise C, et al. GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of Abeta42. *Embo J.* 2007;26:1702-12.
- [103] Beel AJ, Mobley CK, Kim HJ, Tian F, Hadziselimovic A, Jap B, et al. Structural studies of the transmembrane C-terminal domain of the amyloid precursor protein (APP): does APP function as a cholesterol sensor? *Biochemistry.* 2008;47:9428-46.
- [104] Nierzwicki L, Czub J. Specific Binding of Cholesterol to the Amyloid Precursor Protein: Structure of the Complex and Driving Forces Characterized in Molecular Detail. *J Phys Chem Lett.* 2015;6:784-90.
- [105] Munter LM, Botev A, Richter L, Hildebrand PW, Althoff V, Weise C, et al. Aberrant amyloid precursor protein (APP) processing in hereditary forms of Alzheimer disease caused by APP familial Alzheimer disease mutations can be rescued by mutations in the APP GxxxG motif. *The Journal of biological chemistry.* 2010;285:21636-43.
- [106] van Meer G, Voelker DR, Feigenson GW. Membrane lipids: where they are and how they behave. *Nature reviews Molecular cell biology.* 2008;9:112-24.
- [107] Vetrivel KS, Thinakaran G. Membrane rafts in Alzheimer's disease beta-amyloid production. *Biochimica et biophysica acta.* 2010;1801:860-7.
- [108] Schneider D, Engelman DM. GALLEX, a measurement of heterologous association of transmembrane helices in a biological membrane. *The Journal of biological chemistry.* 2003;278:3105-11.
- [109] Karimova G, Pidoux J, Ullmann A, Ladant D. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proceedings of the National Academy of Sciences of the United States of America.* 1998;95:5752-6.
- [110] Russ WP, Engelman DM. TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proceedings of the National Academy of Sciences of the United States of America.* 1999;96:863-8.
- [111] Gurezka R, Langosch D. In vitro selection of membrane-spanning leucine zipper protein-protein interaction motifs using POSSYCCAT. *The Journal of biological chemistry.* 2001;276:45580-7.
- [112] Joce C, Wiener AA, Yin H. Multi-Tox: application of the ToxR-transcriptional reporter assay to the study of multi-pass protein transmembrane domain oligomerization. *Biochimica et biophysica acta.* 2011;1808:2948-53.

- [113] Klein N, Kummerer N, Hobernik D, Schneider D. The AQP2 mutation V71M causes nephrogenic diabetes insipidus in humans but does not impair the function of a bacterial homolog. *FEBS Open Bio*. 2015;5:640-6.
- [114] Karimova G, Dautin N, Ladant D. Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. *J Bacteriol*. 2005;187:2233-43.
- [115] Schneider D, Finger C, Prodohl A, Volkmer T. From interactions of single transmembrane helices to folding of alpha-helical membrane proteins: analyzing transmembrane helix-helix interactions in bacteria. *Current protein & peptide science*. 2007;8:45-61.
- [116] Kolmar H, Frisch C, Kleemann G, Gotze K, Stevens FJ, Fritz HJ. Dimerization of Bence Jones proteins: linking the rate of transcription from an *Escherichia coli* promoter to the association constant of REIV. *Biol Chem Hoppe Seyler*. 1994;375:61-70.
- [117] Miller JH. *Experiments in molecular genetics*: Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory , 1972. - XVI, 466 S. : Ill., graph. Darst.; 1972.
- [118] Lindner E, White SH. Topology, dimerization, and stability of the single-span membrane protein CadC. *J Mol Biol*. 2014;426:2942-57.
- [119] Dawson JP, Melnyk RA, Deber CM, Engelman DM. Sequence context strongly modulates association of polar residues in transmembrane helices. *J Mol Biol*. 2003;331:255-62.
- [120] Lindner E, Unterreitmeier S, Ridder AN, Langosch D. An extended ToxR POSSYCCAT system for positive and negative selection of self-interacting transmembrane domains. *Journal of microbiological methods*. 2007;69:298-305.
- [121] Kirrbach J, Krugliak M, Ried CL, Pagel P, Arkin IT, Langosch D. Self-interaction of transmembrane helices representing pre-clusters from the human single-span membrane proteins. *Bioinformatics*. 2013;29:1623-30.
- [122] Consortium TU. UniProt: a hub for protein information. *Nucleic Acids Res*. 2015;43:D204-12.
- [123] Kolmar H, Hennecke F, Gotze K, Janzer B, Vogt B, Mayer F, et al. Membrane insertion of the bacterial signal transduction protein ToxR and requirements of transcription activation studied by modular replacement of different protein substructures. *Embo J*. 1995;14:3895-904.
- [124] Tome L, Steindorf D, Schneider D. Genetic systems for monitoring interactions of transmembrane domains in bacterial membranes. *Methods in molecular biology*. 2013;1063:57-91.
- [125] Bedouelle H, Duplay P. Production in *Escherichia coli* and one-step purification of bifunctional hybrid proteins which bind maltose. Export of the Klenow polymerase into the periplasmic space. *European journal of biochemistry / FEBS*. 1988;171:541-9.
- [126] Brosig B, Langosch D. The dimerization motif of the glycophorin A transmembrane segment in membranes: importance of glycine residues. *Protein science : a publication of the Protein Society*. 1998;7:1052-6.
- [127] Furthmayr H, Marchesi VT. Glycophorins: isolation, orientation, and localization of specific domains. *Methods in enzymology*. 1983;96:268-80.
- [128] Lindner E, Langosch D. A ToxR-based dominant-negative system to investigate heterotypic transmembrane domain interactions. *Proteins*. 2006;65:803-7.
- [129] Su PC, Berger BW. A novel assay for assessing juxtamembrane and transmembrane domain interactions important for receptor heterodimerization. *J Mol Biol*. 2013;425:4652-8.
- [130] Sambrook J, Fritsch EF, Maniatis T. *Molecular cloning: A laboratory manual*. 2nd ed ed. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press; 1989.

- [131] Tolia NH, Joshua-Tor L. Strategies for protein coexpression in *Escherichia coli*. *Nat Methods*. 2006;3:55-64.
- [132] Riggs AD, Reiness G, Zubay G. Purification and DNA-binding properties of the catabolite gene activator protein. *Proceedings of the National Academy of Sciences of the United States of America*. 1971;68:1222-5.
- [133] Ostermeier M, Nixon AE, Shim JH, Benkovic SJ. Combinatorial protein engineering by incremental truncation. *Proceedings of the National Academy of Sciences of the United States of America*. 1999;96:3562-7.
- [134] Dautin N, Karimova G, Ullmann A, Ladant D. Sensitive genetic screen for protease activity based on a cyclic AMP signaling cascade in *Escherichia coli*. *J Bacteriol*. 2000;182:7060-6.
- [135] Anbazhagan V, Schneider D. The membrane environment modulates self-association of the human GpA TM domain--implications for membrane protein folding and transmembrane signaling. *Biochimica et biophysica acta*. 2010;1798:1899-907.
- [136] Johnsson N, Varshavsky A. Split ubiquitin as a sensor of protein interactions in vivo. *Proceedings of the National Academy of Sciences of the United States of America*. 1994;91:10340-4.
- [137] Stagljar I, Korostensky C, Johnsson N, te Heesen S. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proceedings of the National Academy of Sciences of the United States of America*. 1998;95:5187-92.
- [138] Petschnigg J, Groisman B, Kotlyar M, Taipale M, Zheng Y, Kurat CF, et al. The mammalian-membrane two-hybrid assay (MaMTH) for probing membrane-protein interactions in human cells. *Nat Methods*. 2014;11:585-92.
- [139] Fields S, Song O. A novel genetic system to detect protein-protein interactions. *Nature*. 1989;340:245-6.
- [140] Struhl K, Davis RW. Promoter Mutants of the Yeast His3 Gene. *Journal of Molecular Biology*. 1981;152:553-68.
- [141] Fetchko M, Stagljar I. Application of the split-ubiquitin membrane yeast two-hybrid system to investigate membrane protein interactions. *Methods*. 2004;32:349-62.
- [142] Paumi CM, Chuk M, Snider J, Stagljar I, Michaelis S. ABC transporters in *Saccharomyces cerevisiae* and their interactors: new technology advances the biology of the ABCC (MRP) subfamily. *Microbiol Mol Biol Rev*. 2009;73:577-93.
- [143] Petschnigg J, Moe OW, Stagljar I. Using yeast as a model to study membrane proteins. *Curr Opin Nephrol Hypertens*. 2011;20:425-32.
- [144] Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. *J Mol Biol*. 1963;7:95-9.
- [145] Hovmoller S, Zhou T, Ohlson T. Conformations of amino acids in proteins. *Acta crystallographica Section D, Biological crystallography*. 2002;58:768-76.
- [146] White SH, Ladokhin AS, Jayasinghe S, Hristova K. How membranes shape protein structure. *The Journal of biological chemistry*. 2001;276:32395-8.
- [147] Meruelo AD, Samish I, Bowie JU. TMKink: a method to predict transmembrane helix kinks. *Protein science : a publication of the Protein Society*. 2011;20:1256-64.
- [148] Cruz WS. Peptide Bonds and Protein Structure. In: Laboratory BC, editor.: Washington University; 2011.
- [149] Evers TH, van Dongen EM, Faesen AC, Meijer EW, Merx M. Quantitative understanding of the energy transfer between fluorescent proteins connected via flexible peptide linkers. *Biochemistry*. 2006;45:13183-92.

- [150] Cabantous S, Terwilliger TC, Waldo GS. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nature biotechnology*. 2005;23:102-7.
- [151] Ozawa T, Kaihara A, Sato M, Tachihara K, Umezawa Y. Split luciferase as an optical probe for detecting protein-protein interactions in mammalian cells based on protein splicing. *Analytical chemistry*. 2001;73:2516-21.
- [152] Wehrman T, Kleaveland B, Her JH, Balint RF, Blau HM. Protein-protein interactions monitored in mammalian cells via complementation of beta -lactamase enzyme fragments. *Proceedings of the National Academy of Sciences of the United States of America*. 2002;99:3469-74.
- [153] Galarneau A, Primeau M, Trudeau LE, Michnick SW. Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nature biotechnology*. 2002;20:619-22.
- [154] Wehr MC, Laage R, Bolz U, Fischer TM, Grunewald S, Scheek S, et al. Monitoring regulated protein-protein interactions using split TEV. *Nat Methods*. 2006;3:985-93.
- [155] Pelletier JN, Arndt KM, Pluckthun A, Michnick SW. An in vivo library-versus-library selection of optimized protein-protein interactions. *Nature biotechnology*. 1999;17:683-90.
- [156] Hu CD, Chinenov Y, Kerppola TK. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Molecular cell*. 2002;9:789-98.
- [157] Kerppola TK. Visualization of molecular interactions using bimolecular fluorescence complementation analysis: characteristics of protein fragment complementation. *Chem Soc Rev*. 2009;38:2876-86.
- [158] Miller KE, Kim Y, Huh WK, Park HO. Bimolecular Fluorescence Complementation (BiFC) Analysis: Advances and Recent Applications for Genome-Wide Interaction Studies. *J Mol Biol*. 2015;427:2039-55.
- [159] Kodama Y, Hu CD. An improved bimolecular fluorescence complementation assay with a high signal-to-noise ratio. *BioTechniques*. 2010;49:793-805.
- [160] Pusch S, Dissmeyer N, Schnittger A. Bimolecular-fluorescence complementation assay to monitor kinase-substrate interactions in vivo. *Methods in molecular biology*. 2011;779:245-57.
- [161] Fdez E, Martinez-Salvador M, Beard M, Woodman P, Hilfiker S. Transmembrane-domain determinants for SNARE-mediated membrane fusion. *J Cell Sci*. 2010;123:2473-80.
- [162] Dong CH, Jang M, Scharein B, Malach A, Rivarola M, Liesch J, et al. Molecular association of the Arabidopsis ETR1 ethylene receptor and a regulator of ethylene signaling, RTE1. *The Journal of biological chemistry*. 2010;285:40706-13.
- [163] Ali GS, Reddy A. PAMP-triggered immunity: Early events in the activation of FLAGELLIN SENSITIVE2. *Plant Signal Behav*. 2008;3:423-6.
- [164] Chang J, Clay JM, Chang C. Association of cytochrome b5 with ETR1 ethylene receptor signaling through RTE1 in Arabidopsis. *The Plant journal : for cell and molecular biology*. 2014;77:558-67.
- [165] Sjöhamn J, Hedfalk K. Unraveling aquaporin interaction partners. *Biochimica et biophysica acta*. 2014;1840:1614-23.
- [166] Talaty P, Emery A, Everly DN, Jr. Characterization of the latent membrane protein 1 signaling complex of Epstein-Barr virus in the membrane of mammalian cells with bimolecular fluorescence complementation. *Virology*. 2011;414:414-23.
- [167] Talaty P, Emery A, Holthusen K, Everly DN, Jr. Identification of transmembrane protein 134 as a novel LMP1-binding protein by using bimolecular fluorescence complementation and an enhanced retroviral mutagen. *Journal of virology*. 2012;86:11345-55.

- [168] Vijayapalani P, Maeshima M, Nagasaki-Takekuchi N, Miller WA. Interaction of the trans-frame potyvirus protein P3N-PIPO with host protein PCaP1 facilitates potyvirus movement. *PLoS pathogens*. 2012;8:e1002639.
- [169] Fleming A. On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to their Use in the Isolation of *B. influenzae*. *British journal of experimental pathology*. 1929;10:226-36.
- [170] Chain E, Florey HW, Gardner AD, Heatley NG, Jennings MA, Orr-Ewing J, et al. Penicillin as a Chemotherapeutic Agent. *The Lancet*. 1940;236:226-8.
- [171] Hamilton-Miller JM. Chemical manipulations of the penicillin nucleus: a review. *Chemotherapy*. 1967;12:73-88.
- [172] Dalhoff A, Janjic N, Echols R. Redefining penems. *Biochem Pharmacol*. 2006;71:1085-95.
- [173] Wise EM, Jr., Park JT. Penicillin: its basic site of action as an inhibitor of a peptide cross-linking reaction in cell wall mucopeptide synthesis. *Proceedings of the National Academy of Sciences of the United States of America*. 1965;54:75-81.
- [174] Strominger JL, Tipper DJ. Bacterial cell wall synthesis and structure in relation to the mechanism of action of penicillins and other antibacterial agents. *Am J Med*. 1965;39:708-21.
- [175] Chen Y, Zhang W, Shi Q, Hesk D, Lee M, Mobashery S, et al. Crystal structures of penicillin-binding protein 6 from *Escherichia coli*. *Journal of the American Chemical Society*. 2009;131:14345-54.
- [176] Frere JM, Page MG. Penicillin-binding proteins: evergreen drug targets. *Curr Opin Pharmacol*. 2014;18:112-9.
- [177] Vollmer W, Bertsche U. Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. *Biochimica et biophysica acta*. 2008;1778:1714-34.
- [178] Weidel W, Pelzer H. Bagshaped Macromolecules--a New Outlook on Bacterial Cell Walls. *Adv Enzymol Relat Areas Mol Biol*. 1964;26:193-232.
- [179] Frere JM, Joris B. Penicillin-sensitive enzymes in peptidoglycan biosynthesis. *Crit Rev Microbiol*. 1985;11:299-396.
- [180] Frère JM, Nguyen-Distèche M, Coyette J, Joris B. Mode of action: interaction with the penicillin binding proteins. In: Page M, editor. *The Chemistry of β -Lactams*: Springer Netherlands; 1992. p. 148-97.
- [181] Frere JM, Duez C, Ghuysen JM, Vandekerckhove J. Occurrence of a serine residue in the penicillin-binding site of the exocellular DD-carboxy-peptidase-transpeptidase from *Streptomyces R61*. *FEBS letters*. 1976;70:257-60.
- [182] Yocum RR, Amanuma H, O'Brien TA, Waxman DJ, Strominger JL. Penicillin is an active-site inhibitor for four genera of bacteria. *Journal of Bacteriology*. 1982;149:1150-3.
- [183] Ambler RP. The structure of beta-lactamases. *Philos Trans R Soc Lond B Biol Sci*. 1980;289:321-31.
- [184] Ambler RP, Coulson AF, Frere JM, Ghuysen JM, Joris B, Forsman M, et al. A standard numbering scheme for the class A beta-lactamases. *The Biochemical journal*. 1991;276 (Pt 1):269-70.
- [185] Galleni M, Lamotte-Brasseur J, Rossolini GM, Spencer J, Dideberg O, Frere JM, et al. Standard numbering scheme for class B beta-lactamases. *Antimicrobial agents and chemotherapy*. 2001;45:660-3.
- [186] Jaurin B, Grundstrom T. *ampC* cephalosporinase of *Escherichia coli* K-12 has a different evolutionary origin from that of beta-lactamases of the penicillinase type. *Proceedings of the National Academy of Sciences of the United States of America*. 1981;78:4897-901.

- [187] Ouellette M, Bissonnette L, Roy PH. Precise insertion of antibiotic resistance determinants into Tn21-like transposons: nucleotide sequence of the OXA-1 beta-lactamase gene. *Proceedings of the National Academy of Sciences of the United States of America*. 1987;84:7378-82.
- [188] Hall BG, Barlow M. Revised Ambler classification of {beta}-lactamases. *The Journal of antimicrobial chemotherapy*. 2005;55:1050-1.
- [189] Bush K, Jacoby GA, Medeiros AA. A functional classification scheme for beta-lactamases and its correlation with molecular structure. *Antimicrobial agents and chemotherapy*. 1995;39:1211-33.
- [190] Drawz SM, Bonomo RA. Three decades of beta-lactamase inhibitors. *Clin Microbiol Rev*. 2010;23:160-201.
- [191] Datta N, Kontomichalou P. Penicillinase synthesis controlled by infectious R factors in Enterobacteriaceae. *Nature*. 1965;208:239-41.
- [192] Matthew M, Hedges RW, Smith JT. Types of beta-lactamase determined by plasmids in gram-negative bacteria. *J Bacteriol*. 1979;138:657-62.
- [193] Sideraki V, Huang W, Palzkill T, Gilbert HF. A secondary drug resistance mutation of TEM-1 beta-lactamase that suppresses misfolding and aggregation. *Proceedings of the National Academy of Sciences of the United States of America*. 2001;98:283-8.
- [194] Bourgault AM, Rosenblatt JE. Characterization of Anaerobic Gram-Negative Bacilli by Using Rapid Slide Tests for Beta-Lactamase Production. *Journal of Clinical Microbiology*. 1979;9:654-6.
- [195] Uri JV. Detection of beta-lactamase activity with nitrocefin of multiple strains of various microbial genera. *Acta Microbiol Hung*. 1985;32:133-45.
- [196] Zimmermann W, Rosselet A. Function of the outer membrane of *Escherichia coli* as a permeability barrier to beta-lactam antibiotics. *Antimicrobial agents and chemotherapy*. 1977;12:368-72.
- [197] Fisher AC, Kim W, DeLisa MP. Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. *Protein science : a publication of the Protein Society*. 2006;15:449-58.
- [198] Mansell TJ, Linderman SW, Fisher AC, DeLisa MP. A rapid protein folding assay for the bacterial periplasm. *Protein science : a publication of the Protein Society*. 2010;19:1079-90.
- [199] Lam VM, Beerepoot P, Angers S, Salahpour A. A novel assay for measurement of membrane-protein surface expression using a beta-lactamase. *Traffic*. 2013;14:778-84.
- [200] Broome-Smith JK, Spratt BG. A vector for the construction of translational fusions to TEM beta-lactamase and the analysis of protein export signals and membrane protein topology. *Gene*. 1986;49:341-9.
- [201] Broome-Smith JK, Tadayyon M, Zhang Y. Beta-lactamase as a probe of membrane protein assembly and protein export. *Mol Microbiol*. 1990;4:1637-44.
- [202] Angus BL, Carey AM, Caron DA, Kropinski AM, Hancock RE. Outer membrane permeability in *Pseudomonas aeruginosa*: comparison of a wild-type with an antibiotic-supersusceptible mutant. *Antimicrobial agents and chemotherapy*. 1982;21:299-309.
- [203] Fisher JF, Meroueh SO, Mobashery S. Bacterial resistance to beta-lactam antibiotics: compelling opportunism, compelling opportunity. *Chem Rev*. 2005;105:395-424.
- [204] Chervaux C, Sauvonnet N, Le Clainche A, Kenny B, Hung AL, Broome-Smith JK, et al. Secretion of active beta-lactamase to the medium mediated by the *Escherichia coli* haemolysin transport pathway. *Mol Gen Genet*. 1995;249:237-45.

- [205] Zlokarnik G, Negulescu PA, Knapp TE, Mere L, Burres N, Feng L, et al. Quantitation of transcription and clonal selection of single living cells with beta-lactamase as reporter. *Science*. 1998;279:84-8.
- [206] Zlokarnik G. Fusions to β -lactamase as a reporter for gene expression in live mammalian cells. *Methods in enzymology* 2000. p. 221-44.
- [207] Nord O, Gustrin A, Nygren PA. Fluorescent detection of beta-lactamase activity in living *Escherichia coli* cells via esterase supplementation. *FEMS microbiology letters*. 2005;242:73-9.
- [208] Lee LL, Ha H, Chang YT, DeLisa MP. Discovery of amyloid-beta aggregation inhibitors using an engineered assay for intracellular protein folding and solubility. *Protein science : a publication of the Protein Society*. 2009;18:277-86.
- [209] Wang X, Minasov G, Shoichet BK. Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J Mol Biol*. 2002;320:85-95.
- [210] Zhao CK, Yin Q, Li SY. A high-throughput screening system for G-protein-coupled receptors using beta-lactamase enzyme complementation technology. *Acta Pharmacol Sin*. 2010;31:1618-24.
- [211] Lee HK, Brown SJ, Rosen H, Tobias PS. Application of beta-lactamase enzyme complementation to the high-throughput screening of toll-like receptor signaling inhibitors. *Molecular pharmacology*. 2007;72:868-75.
- [212] Lofdahl PA, Nord O, Janzon L, Nygren PA. Selection of TNF-alpha binding affibody molecules using a beta-lactamase protein fragment complementation assay. *N Biotechnol*. 2009;26:251-9.
- [213] Jun Y, Wickner W. Assays of vacuole fusion resolve the stages of docking, lipid mixing, and content mixing. *Proceedings of the National Academy of Sciences of the United States of America*. 2007;104:13010-5.
- [214] Chung CT, Niemela SL, Miller RH. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Sciences of the United States of America*. 1989;86:2172-5.
- [215] Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res*. 2007;35:W43-6.
- [216] Wang W, Malcolm BA. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. *BioTechniques*. 1999;26:680-2.
- [217] Erijman A, Dantes A, Bernheim R, Shifman JM, Peleg Y. Transfer-PCR (TPCR): a highway for DNA cloning and protein engineering. *J Struct Biol*. 2011;175:171-7.
- [218] Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 1970;227:680-5.
- [219] Lindner E. Identifikation heterotypischer TMD-TMD Interaktionen 2007.
- [220] Cronan JE. A family of arabinose-inducible *Escherichia coli* expression vectors having pBR322 copy control. *Plasmid*. 2006;55:152-7.
- [221] Cozzarelli NR, Kelly RB, Kornberg A. A minute circular DNA from *Escherichia coli* 15. *Proceedings of the National Academy of Sciences of the United States of America*. 1968;60:992-9.
- [222] Schleif R. AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMS Microbiol Rev*. 2010;34:779-96.
- [223] Cesareni G, Muesing MA, Polisky B. Control of ColE1 DNA replication: the rop gene product negatively affects transcription from the replication primer promoter. *Proceedings of the National Academy of Sciences of the United States of America*. 1982;79:6313-7.

- [224] Cesareni G, Cornelissen M, Lacatena RM, Castagnoli L. Control of pMB1 replication: inhibition of primer formation by Rop requires RNA1. *Embo J*. 1984;3:1365-9.
- [225] Sakakibara Y, Tomizawa J. Replication of colicin E1 plasmid DNA in cell extracts. II. Selective synthesis of early replicative intermediates. *Proceedings of the National Academy of Sciences of the United States of America*. 1974;71:1403-7.
- [226] Ried CL. Unilaterale Konserviertheit und homotypische Interaktion von Transmembranhelices im humanen bitopischen Membranproteom: Technische Universität München; 2013.
- [227] Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineering and characterization of a superfolder green fluorescent protein. *Nature biotechnology*. 2006;24:79-88.
- [228] Arai R, Ueda H, Kitayama A, Kamiya N, Nagamune T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein engineering*. 2001;14:529-32.
- [229] Pelletier JN, Campbell-Valois FX, Michnick SW. Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proceedings of the National Academy of Sciences of the United States of America*. 1998;95:12141-6.
- [230] Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*. 1986;189:113-30.
- [231] Wood WB. Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA. *J Mol Biol*. 1966;16:118-33.
- [232] Messing J, Gronenborn B, Muller-Hill B, Hans Hopschneider P. Filamentous coliphage M13 as a cloning vehicle: insertion of a HindII fragment of the lac regulatory region in M13 replicative form in vitro. *Proceedings of the National Academy of Sciences of the United States of America*. 1977;74:3642-6.
- [233] Hill AV. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *The Journal of physiology*. 1910;40:iv-vii.
- [234] Goutelle S, Maurin M, Rougier F, Barbaut X, Bourguignon L, Ducher M, et al. The Hill equation: a review of its capabilities in pharmacological modelling. *Fundam Clin Pharmacol*. 2008;22:633-48.
- [235] Cormack BP, Valdivia RH, Falkow S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*. 1996;173:33-8.
- [236] Feilmeier BJ, Iseminger G, Schroeder D, Webber H, Phillips GJ. Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. *Journal of Bacteriology*. 2000;182:4068-76.
- [237] Toddo S, Soderstrom B, Palombo I, von Heijne G, Norholm MH, Daley DO. Application of split-green fluorescent protein for topology mapping membrane proteins in *Escherichia coli*. *Protein science : a publication of the Protein Society*. 2012;21:1571-6.
- [238] Fisher AC, DeLisa MP. Laboratory evolution of fast-folding green fluorescent protein using secretory pathway quality control. *PloS one*. 2008;3:e2351.
- [239] Hancock RE. Alterations in outer membrane permeability. *Annu Rev Microbiol*. 1984;38:237-64.
- [240] Talanian RV, McKnight CJ, Kim PS. Sequence-specific DNA binding by a short peptide dimer. *Science*. 1990;249:769-71.
- [241] Schulte TH, Marchesi VT. Self-association of human erythrocyte glycophorin A. Appearance of low mobility bands on sodium dodecyl sulfate gels. *Biochimica et biophysica acta*. 1978;508:425-30.

- [242] Treutlein HR, Lemmon MA, Engelman DM, Brunger AT. The glycoporphin A transmembrane domain dimer: sequence-specific propensity for a right-handed supercoil of helices. *Biochemistry*. 1992;31:12726-32.
- [243] Smith SO, Song D, Shekar S, Groesbeck M, Ziliox M, Aimoto S. Structure of the transmembrane dimer interface of glycoporphin A in membrane bilayers. *Biochemistry*. 2001;40:6553-8.
- [244] Doura AK, Fleming KG. Complex interactions at the helix-helix interface stabilize the glycoporphin A transmembrane dimer. *J Mol Biol*. 2004;343:1487-97.
- [245] Lemmon MA, Flanagan JM, Hunt JF, Adair BD, Bormann BJ, Dempsey CE, et al. Glycoporphin A dimerization is driven by specific interactions between transmembrane alpha-helices. *The Journal of biological chemistry*. 1992;267:7683-9.
- [246] Ried C, Scharnagl C, Langosch D. Entrapment of water at the transmembrane helix-helix interface of quiescin sulphydryl oxidase 2 (QSOX2). *Biochemistry*. 2016.
- [247] Secco P, D'Agostini E, Marzari R, Licciulli M, Di Niro R, D'Angelo S, et al. Antibody library selection by the β -lactamase protein fragment complementation assay. *Protein engineering, design & selection : PEDS*. 2009;22:149-58.
- [248] Porter JR, Stains CI, Segal DJ, Ghosh I. Split beta-lactamase sensor for the sequence-specific detection of DNA methylation. *Analytical chemistry*. 2007;79:6702-8.
- [249] Ou W, Marino MP, Lu C, Reiser J. Rapid titration of retroviral vectors using a beta-lactamase protein fragment complementation assay. *Gene Ther*. 2013;20:43-50.
- [250] Park SH, Cheong C, Idoyaga J, Kim JY, Choi JH, Do Y, et al. Generation and application of new rat monoclonal antibodies against synthetic FLAG and OLLAS tags for improved immunodetection. *J Immunol Methods*. 2008;331:27-38.
- [251] Rath A, Glibowicka M, Nadeau VG, Chen G, Deber CM. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106:1760-5.
- [252] Herrmann JR, Fuchs A, Panitz JC, Eckert T, Unterreitmeier S, Frishman D, et al. Ionic interactions promote transmembrane helix-helix association depending on sequence context. *J Mol Biol*. 2010;396:452-61.
- [253] Herrmann JR. Rolle polarer und ionisierbarer Aminosäuren bei der Interaktion membranständiger α -Helices 2010.
- [254] Grigoryan G, Keating AE. Structural specificity in coiled-coil interactions. *Curr Opin Struct Biol*. 2008;18:477-83.
- [255] Lee SK, Chou HH, Pflieger BF, Newman JD, Yoshikuni Y, Keasling JD. Directed evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters. *Applied and environmental microbiology*. 2007;73:5711-5.
- [256] Palmer T, Berks BC. The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol*. 2012;10:483-96.
- [257] Remy I, Ghaddar G, Michnick SW. Using the beta-lactamase protein-fragment complementation assay to probe dynamic protein-protein interactions. *Nature protocols*. 2007;2:2302-6.
- [258] Ruan W, Lindner E, Langosch D. The interface of a membrane-spanning leucine zipper mapped by asparagine-scanning mutagenesis. *Protein science : a publication of the Protein Society*. 2004;13:555-9.
- [259] Masuda G, Tomioka S, Uchida H, Hasegawa M. Bacteriostatic and bactericidal activities of selected beta-lactam antibiotics studied on agar plates. *Antimicrobial agents and chemotherapy*. 1977;11:376-82.

- [260] Myers JA, Curtis BS, Curtis WR. Improving accuracy of cell and chromophore concentration measurements using optical density. *BMC Biophys.* 2013;6:4.
- [261] Argos P. An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *J Mol Biol.* 1990;211:943-58.
- [262] Meindl-Beinker NM, Lundin C, Nilsson I, White SH, von Heijne G. Asn- and Asp-mediated interactions between transmembrane helices during translocon-mediated membrane protein assembly. *EMBO Rep.* 2006;7:1111-6.
- [263] Steiner T. The hydrogen bond in the solid state. *Angewandte Chemie.* 2002;41:49-76.
- [264] Wagenknecht H, Knapp H. L-Asparaginsäure. RÖMPP ONLINE. Stuttgart: Georg Thieme Verlag KG; 2006.
- [265] Call ME, Schnell JR, Xu C, Lutz RA, Chou JJ, Wucherpfennig KW. The structure of the zeta-zeta transmembrane dimer reveals features essential for its assembly with the T cell receptor. *Cell.* 2006;127:355-68.
- [266] Sharma S, Lensink MF, Juffer AH. The structure of the CD3zeta-zeta transmembrane dimer in lipid bilayers. *Biochimica et biophysica acta.* 2014;1838:739-46.
- [267] Wagenknecht H, Knapp H. L-Arginin. RÖMPP ONLINE. Stuttgart: Georg Thieme Verlag KG; 2006.
- [268] Yoo J, Cui Q. Chemical versus mechanical perturbations on the protonation state of arginine in complex lipid membranes: insights from microscopic pKa calculations. *Biophys J.* 2010;99:1529-38.
- [269] Li L, Vorobyov I, MacKerell AD, Jr., Allen TW. Is arginine charged in a membrane? *Biophys J.* 2008;94:L11-3.
- [270] Choe S, Hecht KA, Grabe M. A continuum method for determining membrane protein insertion energies and the problem of charged residues. *J Gen Physiol.* 2008;131:563-73.
- [271] Yoo J, Cui Q. Does arginine remain protonated in the lipid membrane? Insights from microscopic pKa calculations. *Biophys J.* 2008;94:L61-3.
- [272] Schow EV, Freitas JA, Cheng P, Bernsel A, von Heijne G, White SH, et al. Arginine in membranes: the connection between molecular dynamics simulations and translocon-mediated insertion experiments. *J Membr Biol.* 2011;239:35-48.
- [273] Lei SP, Lin HC, Wang SS, Callaway J, Wilcox G. Characterization of the *Erwinia carotovora* pelB gene and its product pectate lyase. *J Bacteriol.* 1987;169:4379-83.
- [274] Beha D, Deitermann S, Muller M, Koch HG. Export of beta-lactamase is independent of the signal recognition particle. *The Journal of biological chemistry.* 2003;278:22161-7.
- [275] Movva NR, Nakamura K, Inouye M. Amino acid sequence of the signal peptide of ompA protein, a major outer membrane protein of *Escherichia coli*. *The Journal of biological chemistry.* 1980;255:27-9.
- [276] Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol.* 1995;177:4121-30.
- [277] Siegele DA, Hu JC. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proceedings of the National Academy of Sciences of the United States of America.* 1997;94:8168-72.
- [278] Hoffmann F, van den Heuvel J, Zidek N, Rinas U. Minimizing inclusion body formation during recombinant protein production in *Escherichia coli* at bench and pilot plant scale. *Enzyme Microb Technol.* 2004;34:235-41.
- [279] Dammeyer T, Tinnefeld P. Engineered fluorescent proteins illuminate the bacterial periplasm. *Computational and structural biotechnology journal.* 2012;3:e201210013.

- [280] Aronson DE, Costantini LM, Snapp EL. Superfolder GFP is fluorescent in oxidizing environments when targeted via the Sec translocon. *Traffic*. 2011;12:543-8.
- [281] Cava F, de Pedro MA, Blas-Galindo E, Waldo GS, Westblade LF, Berenguer J. Expression and use of superfolder green fluorescent protein at high temperatures in vivo: a tool to study extreme thermophile biology. *Environmental microbiology*. 2008;10:605-13.
- [282] Sanders CR, Mittendorf KF. Tolerance to changes in membrane lipid composition as a selected trait of membrane proteins. *Biochemistry*. 2011;50:7858-67.
- [283] Kuznetsov AS, Volynsky PE, Efremov RG. Role of the Lipid Environment in the Dimerization of Transmembrane Domains of Glycophorin A. *Acta Naturae*. 2015;7:122-7.
- [284] Winkler E, Julius A, Steiner H, Langosch D. Homodimerization Protects the Amyloid Precursor Protein C99 Fragment from Cleavage by gamma-Secretase. *Biochemistry*. 2015;54:6149-52.
- [285] Lloris-Garcera P, Bianchi F, Slusky JS, Seppala S, Daley DO, von Heijne G. Antiparallel dimers of the small multidrug resistance protein EmrE are more stable than parallel dimers. *The Journal of biological chemistry*. 2012;287:26052-9.
- [286] Banigan JR, Gayen A, Cho MK, Traaseth NJ. A structured loop modulates coupling between the substrate-binding and dimerization domains in the multidrug resistance transporter EmrE. *The Journal of biological chemistry*. 2015;290:805-14.
- [287] Bullock WO, Fernandez JM, Short JM. XL1-Blue: a high efficiency plasmid transforming recA Escherichia coli strain with beta-galactosidase selection. *BioTechniques* 5:376-379. 1987.

7 List of Figures

Figure 1: Two different discussed models for membrane insertion of TMDs.....	5
Figure 2: Schematic illustration of the homotypic ToxR/TOXCAT/POSSYCCAT, GALLEX and AraTM assay.....	10
Figure 3: Schematic illustration of the heterotypic variants of the GALLEX , dominant-negative ToxR and dominant negative AraTM assays.	12
Figure 4: Schematic illustration of the BACTH assay (bacterial adenylate cyclase two-hybrid)	14
Figure 5: Schematic illustration of the MYTH assay (membrane yeast two-hybrid).....	16
Figure 6: Relationship of the molecular structure of an α -helix and orientation dependence	18
Figure 7: Structure of murein and cross-linking by DD-transpeptidase.....	21
Figure 8: Catalytic center and reaction mechanism of class A serine β -lactamases.....	22
Figure 9: Oligo design for cassette cloning	30
Figure 10: Illustration of the final vectors containing the BLaTM proteins.....	43
Figure 11: Overview of the vector construction	45
Figure 12: Scheme of BLaTM hybrid protein	46
Figure 13: Scheme of BLa-bZIP hybrid protein	46
Figure 14: Scheme of BLa-bZIP split proteins (N-BLa 0.1 and C-BLa 0.1)	47
Figure 15: Scheme of BLaTM 1.1 proteins.	48
Figure 16: Scheme of BLaTM 1.2 proteins	49
Figure 17: Protocol for the BLaTM Assay	50
Figure 18: Pipetting scheme for BLaTM assay in a 12-well plate	51
Figure 19: β -lactamase activity and protein expression.....	55
Figure 20: Primary experiments with split β -lactamase fragments fused to soluble leucine-zipper bZIP	56
Figure 21: Absorption raw data of the BLaTM 1.1 ₁₃ assay fitted with Hill equation	57
Figure 22: Schematic illustration of the BLaTM assay	59
Figure 23: Primary experiments with membrane bound split β -lactamase constructs.. ..	60
Figure 24: Western blots of BLa 1.1 ₁₃ Δ SP and BLa 1.1 ₁₃ constructs	62
Figure 25: Effect of the linker length between the β -lactamase fragments and the TMD to the LD ₅₀ in the BLaTM 1.1 assay.....	64
Figure 26: Orientation dependence of different TMDs and site directness of the interfacial amino acids.....	65

Figure 27: Disruption indices of all orientations of GpA and QSOX2.....	67
Figure 28: Main interaction interface of TMD LS46.....	68
Figure 29: Homotypic interaction of the artificial model TMD LS46 in the BLaTM 1.1 ₁₃ system	68
Figure 30: Homo- and heterotypic interactions of TMDs containing ionizable amino acids based on TMD L19_G ₁₄ G ₁₈ in the BLaTM 1.1 ₁₃ system.....	70
Figure 31: Relationship of protein expression and LD ₅₀ in the BLaTM 1.2 system.	72
Figure 32: Homotypic interactions of all orientations of GpA wt and GpA G83I TMDs in the BLaTM 1.2 system.	73
Figure 33: Possible dimers, their activity and the dependence on each other.....	80
Figure 34: Intermoleculare interactions between the carboxyl group of aspartic acid (D) and the guanidinium group of arginine (R) in different charge states..	81
Figure 35: Amino acid sequences (single letter code) of the proteins used in BLaTM 1.1 (PelB) and BLaTM 1.2 (OmpA) assay	116

8 List of Abbreviations

ΔG_{app}	Change of Gibbs free energy for each amino acid due to membrane insertion
ΔG_{app}^{aa}	Change of Gibbs free energy for each amino acid at each TMD position due to membrane insertion
$^{\circ}\text{C}$	Degree Celsius
μg	10^{-6} gram
μL	10^{-6} liter
μM	10^{-6} mol/liter
A	Alanine
Å	Ångström (10^{-10} meter)
A_{544}	Absorption at 544 nm
Aa	Amino acid
Amp	Ampicillin
AP	Alkaline phosphatase
APP	Alzheimer precursor protein
APS	Ammonium persulfate
Ara	Arabinose
ATP	Adenosine triphosphate
BCIP	5-bromo-4-chloro-3-indolyl phosphate
BLa	β -Lactamase
<i>bla</i>	Gene coding for β -lactamase
bp	Base pair
bZIP	Basic leucine zipper
C	Cysteine
cAMP	Cyclic adenosine monophosphate
CAP	Catabolite activator protein
CAT	Chloramphenicol acetyltransferase
<i>cat</i>	Gene coding for chloramphenicol acetyltransferase
C-BLa	Fusion protein containing C-terminal fragment of β -lactamase

Cm	Chloramphenicol
<i>Cm^R</i>	Gene coding for chloramphenicol acetyltransferase
<i>ColE1</i>	ColE1/pMB1 origin of replication
<i>ctx</i>	Promoter of the ToxR transcription factor
Cub	C-terminal ubiquitin fragment
D	Aspartic acid
DAP12	TYRO protein tyrosine kinase-binding protein
ddNTP	Didesoxynucleotide
dH ₂ O	Distilled water
DHFR	Dihydrofolate reductase
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Desoxynucleotide
DTT	Dithiothreitol
<i>E. coli</i>	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
F	Phenylalanine
FACS	fluorescence activated cell sorting
FRET	Förster resonance energy transfer
G	Glycine
<i>gfp</i>	Gene coding for green fluorescent protein
GFP	green fluorescent protein
Glu	Glutamic acid
GpA	Glycophorin A, UniprotKB P02724
H	Histidine
h	Hour
HER2	Human epidermal growth factor receptor 2
<i>HIS3</i>	Gene coding for imidazoleglycerol-phosphate dehydratase
I	Isoleucine
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K	Lysine
Kan	Kanamycin sulfate

<i>Kan^R</i>	Gene coding for aminoglycoside-3'-phosphotransferase
kb	10 ³ base pairs
KCl	Potassium chloride
K _d	Dissociation constant
kDa	10 ³ Dalton
L	Leucine
<i>lacZ</i>	Gene coding for β-galactosidase
LB medium	Lysogeny broth medium
LD ₅₀	Median lethal dose, ampicillin concentration at half maximum A ₅₄₄ after 19 h incubation
Leu	Leucine
Lys	Lysine
M	Methionine
mA	10 ⁻³ ampere
MaMTH	Mammalian-membrane two-hybrid
MgCl ₂	Magnesium chloride
min	Minute
mL	10 ⁻³ liter
mM	10 ⁻³ mol/liter
MYTH	Membrane yeast two-hybrid
N	Asparagine
n	Number of measurements
Na ⁺	Sodium cation
Na ₃ PO ₄	Sodium phosphate
NaCl	Sodium chloride
N-BLa	Fusion protein containing N-terminal fragment of β-lactamase
NBT	Nitro blue tetrazolium
NC-BLa	Full-length β-lactamase
ng	10 ⁻⁹ gram
nm	10 ⁻⁹ meter
NubG	N-terminal ubiquitin fragment
OD ₆₀₀	Optical density at 600 nm
OmpA	Signal peptide of Outer Membrane Protein A (UniProtKB P0A910)

ONPG	Ortho-nitrophenyl- β -galactoside
op ⁺ /op ⁺	Operator/promotor in homotypic GALLEX assay
op408/op ⁺	Operator/promotor in heterotypic GALLEX assay
P	Proline
<i>p15a</i>	p15a origin of replication
p _{BAD}	Promoter for AraBAD proteins
PBP	Penicillin-binding protein
PBS	Phosphate buffered saline
PCA	Protein complementation assay
PCR	Poly-chain-reaction
PDB ID	Protein data base identifier
PEG-3350	Polyethylene glycol, ~3350 g/mol
PelB	Signal peptide of Pectin lyase B, UniProtKB Q00205
PLV	protein A-LexA-VP16
PNK	T4 DNA polynucleotide kinase
QSOX2	Sulfhydryl Oxidase 2 , UniprotKB Q6ZRP7
R	Arginine
RNA	Ribonucleic acid
<i>rop</i>	Gene coding for rop protein
rpm	Revolutions per minute
S	Serine
SDS	Sodium dodecyl sulfate
SDS-PAGE	Polyacrylamide gel electrophoresis
SEM	Standard error of mean
Ser	Serine
sfGFP	Super folder green fluorescent protein
SP	Signal peptide
SRP	Signal recognition particle
T	Threonine
TEMED	Tetramethylethylenediamine
T _M	Primer melting temperature
TMD	Transmembrane domain
TMS	Transmembransegment

TPCR	Transfer PCR
Tris	Tris(hydroxymethyl)aminomethane
U	Unit
UBP	Ubiquitin specific proteases
V	Valine
V	Volt
v/v	Volume per volume
W	Tryptophan
w/v	Weight per volume
wt	Wildtype
X/x	Any amino acid
xg	x gravity
Y	Tyrosine
YTH	Yeast two-hybrid
λ	Wave length
λ_{Em}	Emission wave length
λ_{Ex}	Excitation wave length

Name	Sequence
QSOX2_3_wt_s	CTAGCATGTCCTGTGCGTGGTGTGTATGTGGCAAGCTCTCTGTTTCTGATGGTGTATGTACTTCGG
QSOX2_3_wt_as	GATCCCGAAGTACATACCATCAGAAACAGAGAGCTTGCCACATACAGCACACCACGCACAGACATG
LS46_0_s	CTAGCCTGGCCCTGCTGGATCGCCTGTGTCTATCTGCTGGTGGCCTGCTGGGGCCCTGCTGGG
LS46_0_as	GATCCCAGCAGGCCGCCAGCAGGCCACCAGCAGATAGCACAGGCGATCCAGCAGGCCCCAGG
LS46_1_s	CTAGCCTGCTGGCCCTGCTGGATCGCCTGTGTCTATCTGCTGGTGGCCTGCTGGGGCCCTGCTGGG
LS46_1_as	GATCCCAGGCCGCCAGCAGGCCACCAGCAGATAGCACAGGCGATCCAGCAGGCCCCAGCAGG
LS46_2_s	CTAGCCTGCTGGCCCTGCTGGATCGCCTGTGTCTATCTGCTGGTGGCCTGCTGGGGCCCTGCTGGG
LS46_2_as	GATCCCAGGCCGCCAGCAGGCCACCAGCAGATAGCACAGGCGATCCAGCAGGCCCCAGCAGCAGG
LS46_3_s	CTAGCCTGCTGGCCCTGCTGGATCGCCTGTGTCTATCTGCTGGTGGCCTGCTGGGGCCCTGCTGGG
LS46_3_as	GATCCCAGGCCGCCAGCAGGCCACCAGCAGATAGCACAGGCGATCCAGCAGGCCCCAGCAGCAGCAGG
L19_G14G18_s	CTAGCCTGCTTACTATTGCTCTTACTGTTACTGCTGTTACTGGTCTCTTCTGTAGGCCCTCGG
L19_G14G18_as	GATCCCAGGCCCTAGCAAGAGACCCAGTAACAGCAGTAAACAGTAAAGCAATAGTAAAGAGCAGG
L19_G14G18_D6R7_s	CTAGCCTGCTTACTATTGGATCGCCTGCTTACTTACTTACTGGTCTCTTGTAGGCCCTCGG
L19_G14G18_D6R7_as	GATCCCAGGCCCTAGCAAGAGACCCAGTAATAGTAAAGAGACCCAGTAAAGAGCAGGCCGATCCAA
L19_G14G18_D6_s	CTAGCCTGCTTACTATTGGATTTACTGTTACTGCTGTTACTGGTCTCTTGTAGGCCCTCGG
L19_G14G18_D6_as	GATCCCAGGCCCTAGCAAGAGACCCAGTAACAGCAGTAAACAGTAAATCCAAATAGTAAAGAGCAGG
L19_G14G18_R7_s	CTAGCCTGCTTACTATTGCTCCGCCCTGTTACTGCTGTTACTGGTCTCTTGTAGGCCCTCGG
L19_G14G18_R7_as	GATCCCAGGCCCTAGCAAGAGACCCAGTAACAGCAGTAAACAGCAGTAAAGAGCAGGCCGCAATAGTAA
L19_G14G18_D6D7_s	CTAGCCTGCTTACTATTGGATGATCTGTTACTGTTACTGTTACTGGTCTCTTGTAGGCCCTCGG
L19_G14G18_D6D7_as	GATCCCAGGCCCTAGCAAGAGACCCAGTAACAGCAGTAAACAGATCAATCCAAATAGTAAAGAGCAGG
L19_G14G18_R6R7_s	CTAGCCTGCTTACTATTGCGCCGCCCTGTTACTGCTGTTACTGGGTCCTTGTAGGCCCTCGG
L19_G14G18_R6R7_as	GATCCCAGGCCCTAGCAAGAGACCCAGTAACAGCAGTAAACAGGCCGCCGCAATAGTAAAGAGCAGG

9.2 PCR primers

Name	Sequence
Q6ZRP7 S8A +	CCATCAGAAAACAGAGAGGCTGCCACATACAGCACC
Q6ZRP7 S8A -	GCTGTATGTGGCAAGCCCTGTTCTGTGATGGTG
MP AmpR fwd	GTTTTTTTGGCTCTAGATAACGAGGGCAAAAAATGAGTATTCAACATTTCCGTGTC
MP AmpR rev	CCAGAAATGATCACCGTAGCTCGATTAGCAGAACCGCTGCCACTAGTCCAATGCTTAATCAGTGAGGCAC
BLa M182T s	CGACGAGCGTGACACCCACGACGCGCTGTAGCAATGGCAACAAC
BLa M182T as	GTTGTTGCCATTGCTACAGGCGTGGTGTACGCTCGTGC
BLa PelB s	GATAACGAGGGCAAAAAATGCATATATAAACTGCTGTTTTGGGCCCGGGCCGCGAGCCTGGCGAGCGCGGTGAGCGCTACCCAGAAAACGCT
BLa PelB as	AGCGTTCTGGGTGAGCGCTACCGCGCTCGCCAGGCTCGCGGCGGGCCGCAACAGCTTATAATGCATTTTTTGGCCCTCGTTATC
GCN4 fwd 15x_speI	ATCTGCAGTTACTTGTGCTCGTCTGCTTGTAGTCGGCTTCGCCAACTAATTTCTTTAAATCTG
GCN4 rev ft PstI	GACTAGTGGCGGTGGCGGCAGTGGCGGTGGTGGCAGCGGGCGGGTGGCAGCAACACTGAAGCCCGCAGGCGGTTTC
BLa 194G rev	GCCAGTTAATAGTTTGGCGCAAC
BLa speI-15x_fwd	ACTAGTGGCGGTGGCGGC
BLa PelB+H_rev	GTGAGCGCTCACCCGCGC
BLa 196L_fwd	CTACTTACTCTAGCTTCCCGG
pToxRV_fwd	GCAATCTGTAACAAGAAGCGGG
pToxRV_rev XhoI	AACTCGAGGAGGCTTATGGACTGACTTG
BLa Spacer_fwd	GGTTCTGCTAATCGAGCTAGC
N-BLa Spacer_GGGS_rev	GCTACCGCCACCGCTGCCACTAGTCCAGTTA
C-BLa Spacer_GGGS_rev	GCTACCGCCACCGCTGCCACTAGTCCAAATGC
C-BLa 2xGGGS_rev	GCTGCCCGCCCGCTGCCCGCCGCGCTGCCACTAGTCCAAATGC
BLa delta SP_fwd	CCAGAAAACGCTGGTGAAGTA
OmpA_rev	CGCCGCTGGCCACGGTCGCAAAAGCCCGCCAGGCCACCAGCAATCGCGGTCTTTTCAATTTTTGGCCCTCGTTATCTAG

9.3 Sequencing primers

Name	Sequence	Plasmids	Platform
GATC-pBAD-FP	ATGCCATAGCATT TTTTATCC	N-BLa and C-BLa	GATC SupremeRun
BFP_Y66H_rev	GGGTCAGGGTGGTAACCAGG	N-BLa and C-BLa	GATC LightRun
N-BLa_fwd_IRD700	CACGACGCCTGTAGCAATG	N-BLa	In-house
C-BLa_fwd_IRD800	CGAAATAGACAGATCGCTG	C-BLa	In-house

9.4 Bacteria strains

<i>E. coli</i> strain	Genotype	References
BL21	F ⁻ <i>e14</i> (<i>McrA</i>) <i>hsdR</i> (<i>rK</i> ⁻ <i>mK</i> ⁻) <i>glnV44 thr-1 leuB6 thi-1 lacY1 fhuA21 mcrB hflA150::Tn10</i> (Tet ^R)	[230, 231]
JM83	F ⁻ <i>ara</i> Δ (<i>lac-proAB</i>) <i>rpsL</i> (Str ^R)[ϕ 80 <i>dlac</i> Δ (<i>lacZ</i>) <i>M15</i>] <i>thi</i>	[232]
XL1-blue	F ['] :: <i>Tn10 proA</i> ⁺ <i>B</i> ⁺ <i>lacI</i> ^q Δ (<i>lacZ</i>) <i>M15</i> / <i>recA1 endA1 gyrA96</i> (Nal ^R) <i>thi hsdR17</i> (<i>rK</i> ⁻ <i>mK</i> ⁺) <i>glnV44 relA1 lac</i>	[287]

9.5 Protein sequences

A BLaTM 1.1: ^{PeIB} **MHYKLLFAAAAAASLASAVSAH...**
 BLaTM 1.2: ^{OmpA} **MKKTAIAIAVALAGFATVAQAA...**

B

```

# 23   BLa_N-term
... APETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMS
BLa_N-term
TFKVLCCGAVLSRVDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMT
BLa_N-term
VRELCSAAITMSDNTAANLLLTITGGPKELTAFLNMGDHYVTRLDRWE
BLa_N-term      M182T      # 194 Spacer
PELNEAIPNDERDTTTPVAMATTLRKLTTGSGSGGGSGGGSGSANRA
Int a5 GP      Spacer
SVITLAVGPGLLLLGIHKLAVNVQLPAEAAAKEAAAKEAAAKEAAAKE
Spacer  sfGFP
AAAKAASKGEELFTGVVPIVLVDGDVNGHKFSVRGEGEGDATNGKLT
sfGFP
LKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEG
sfGFP
YVQERTISFKDDGYKTRAEVKFEGDGLVNRIELKGI DFKEDGNILGH
sfGFP
KLEYNFNHSHVYITADKQKNGIKANFKIRHNVEDGQVQLADHYQQNTP
sfGFP
IGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMLLEFVTAAGITHGMDE
sfGFP FLAG-tag
LYKGDYKDHDG*

```

C

```

# 196   BLa_C-term
... LLTLASRQQLIDWMEADKVAGPLLRSLPAGWFIADKSGAGERG
BLa_C-term
SRGIIAALGPDGKPSRIVVIYTTGSOATMDERNRQIAEIGASLIKHWT
Spacer      Int a5 GP      Spacer
SGSGGGSGGGSGSANRASVILAVGPGLLLLGIHKLAVNVQLPAEAAA
Spacer      sfGFP
KEAAAKEAAAKEAAAKEAAAKAASKGEELFTGVVPIVLVDGDVNGHK
sfGFP
FSVRGEGEGDATNGKLT LKFICTTGKLPVPWPTLVTTLTYGVCFSRY
efGFP
PDHMKQHDFFKSAMPEGYVQERTISFKDDGYKTRAEVKFEGDGLVNR
sfGFP
IELKGI DFKEDGNILGHKLEYNFNHSHVYITADKQKNGIKANFKIRHN
sfGFP
VEDGQVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDH
sfGFP      FLAG-tag
VLEFVTAAGITHGMDELYKGDYKDHDG*

```

Figure 35: Amino acid sequences (single letter code) of the BLaTM proteins. **A**) Signal peptides used in BLaTM 1.1 (PeIB) and BLaTM 1.2 (OmpA). **B**) N-BLa protein with a 13 residue spacer and the integrin $\alpha 5$ GP TMD. Numbers at the start and end of the sequence denote residues of TEM-1 β -lactamase (UniProtKB annotation number P62593). **C**) C-BLa protein with a 13 residue spacer and the integrin $\alpha 5$ GP TMD. Numbers at the start of the sequences denote residue of TEM-1 β -lactamase (UniProtKB annotation number P62593).

9.6 Plasmid sequences

9.6.1 N-BLa 1.1

(AGTCAGCCCC ATACGATATA AGTTGTAATT CTCATGTTTG ACAGCTTATC AATCGATAAG **arabinose_O2_operator_region** AAACCAATTG TCCATATTGC ATCAGACATT GCCGCTACTG
 20 40 60 80 100
 CGTCTTTTAC TGGCTCTTCT CGCTAACCA ACCGGTAACC CCGCTTATTA AAAGCATTCT GTAACAAAGC GGGACCAAAG CCATGACAAA **MluI** AACCGCTAAC
 120 140 160 180 200
 AAAAGTGCTC ATAATCACGG CAGAAAAGTC CACATTGATT ATTTGCACGG CGTCACACTT **arabinose_O1_operator_region** TGCTATGCCA TAGCATTTTT ATCCATAAGA TTAGCGGATG
 240 280 300
 CTACCTGACG CTTTTATCG CAACTCTCTA CTGTTTCTCC ATACCCGTTT TTTTGGGCTC TAGATAACGA GGGCAAAAAA **pelB signal peptide** TGCAATTATA ACTGCTGTTT
 340 360 400
pelB signal peptide 420 440 **N-BLa** 460 480 500
 GCGGCGCGGG CCGCGAGCCT GGGGAGCGCG GTGAGCGCTC **N-BLa** ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT
 520 540 560 580 600
N-BLa 620 640 660 680 700
 ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAAGCTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC
N-BLa 720 740 760 780 800
 GGTATTATCC CGTGTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGTGACTC CACCAGTCAC AGAAAAGCAT
N-BLa 820 840 860 880 900
 CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA
N-BLa 920 940 960 980 1,000
 AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA
N-BLa 920 940 960 980 1,000
 CACCACGACG CCGTGTAGCAA TGGCAACAAC GTTGGCAAA CTATTAACCT GCACTAGTGG CAGCGGCGGC GGCAGCGGCG GCGGACGCGG TTCGTGCTAAT
NheI **SpeI** **JM**
JM **alphaVmut-TMD** **Apal** **BamHI** **rigid linker**
 CGAGCTAGCG TGATCATCTT GCGGGTGGGG CCGGCCCTGC TGCTCCTGGG GATCCACAAA CTGGCAGTTA ATGTTCCAGCT GCCTGCCGAA GCAGCAGCGA
rigid linker **sfGFP**
 AAGAAGCCGC TGCCAAAGAA CCGGACGCGA AAGAGGCTGC CGCGAAAGAG GCAGCAGCTA AAGCCGCAAG CAAAGGTGAA GAACTGTTT CCGGTGTTGT
sfGFP
 TCCGATTCTG GTTGAATGG ATGGTATGT TAATGGCCAC AAATCTCAG TTCGTGGTGA AGGCGAAGGT GATGCAACCA ATGGTAACT GACCTGAAA
sfGFP
 TTTATCTGTA CCACGGCAA ACTGCCGTT CCGTGGCCGA CCCTGGTTAC CACCCTGACC TATGGTGTTC AGTGTTCAG CCGTTATCCG GATCATATGA
sfGFP
 AACAGCACGA TTTCTTCAA TCTGCAATGC CGGAAGGTTA TGTTCAGAA CGTACCATCT CCTTTAAGA TGATGGCACC TATAAGACCC GTGCCGAAGT
sfGFP
 TAAATTTGAA GGTGATACCC TGGTGAATCG CATTGAACTG AAAGGCATCG ATTTCAAAGA AGATGGTAAAT ATCCTGGGCC ATAACTGGA ATATAATTTT
sfGFP
 AACAGCCACA ACGTGTATAT CACCGCAGAT AAACAGAAGA ATGGCATCAA AGCCAACCTT AAGATCCGCC ATAATGTTGA AGATGGCAGC GTTCAGCTGG
sfGFP
 CAGATCATTG TCAGCAGAA ACACCGATTG GTGATGGTCC GGTCTGCTG CCGGATAATC ATTATCTGAG CACCAGAGC GTTCTGAGCA AAGATCCGAA
sfGFP **FLAG-tag**
 TGAGAAACGT GATCACATGG TGCTGCTGGA ATTTGTTACC GCAGCAGGTA TTACCATGG TATGGATGAA CTGTATAAG GCGATTATAA AGATCACGAC
FLAG-tag **XhoI**
 GGCTAACTCG AGCACCACCA CCACCACCAC TGAGATCCGG CTGCTAACAA AGCCCGAAAG GAAGCTGAGT TGCTGCTGC CACCCTGAG CAATAACTAG
T7 terminator
 CATAACCCTT TGGGGCCTCT AAACGGGTCT TGAGGGGTTT TTTGACCGAT GCCCTTGAGA GCCTTCAACC CAGTCAGCTC CTTCCGGTGG GCGCGGGGCA
 2,020 2,040 2,060 2,080 2,100
 TGACTATCGT CGCCGCACCT ATGACTGTCT TCTTTATCAT GCAACTCGTA GGACAGGTGC CCGCAGCGCT CTGGTCAATT TTCGGCAGG ACCGCTTTTC
 2,120 2,140 2,160 2,180 2,200
 CTGGAGCGCG ACGATGATCG GCCTGTCTGT TGCCTGATTC GGAATCTTGC ACGCCCTCGC TCAAGCCTTC GTCAGTGGTC CCGCCACCAA ACGTTTCGGC
 2,220 2,240 2,260 2,280 2,300
 GAGAAGCAGG CCATTATCGC CGGCATGGCG GCCGACGCGC TGGGCTACGT CTTGCTGGCC TTCGCGACGC GAGGCTGGAT GGCCTTCCCC ATTATGATTC
 2,320 2,340 2,360 2,380 2,400
 TTCTCGCTTC CCGCGGCATC GGGATGCCCC CGTTGCAGGC CATGCTGTCC AGGCAGGTAG ATGACGACCA TCAGGGACAG CTTCAAGGAT CGCTCGCGGC
 2,420 2,440 2,460 2,480 2,500
 TCTTACCAGC CTAACCTCGA TCACTGGACC GCTGATCGTC ACGCGATTT ATGCCGCTC GCGGAGCACA TGAACGGGT TGCCATGGAT TGTAGGCGCC
 2,520 2,540 2,560 2,580 2,600
 GCCCTATACC TTGCTGCTC CCGCGGTTG CGTGCAGGTC CATGGAGCCG GGCCACCTCG ACCTGAATGG AAGCCGCGGC CACCTCGCTA ACGGATTCAC
 2,620 2,640 2,660 2,680 2,700
 CACTCCAAGA ATGGAGGCCA ATCAATCTCT GCGGAGAAGT GTGAATGGCC AAACCAACCC TTGGCAGAAC ATATCCATCG CGTCCGGCAT CTCCAGCACG
 2,720 2,740 2,760 2,780 2,800

```

      2.820      2.840      2.860      2.880      2.900
      |         |         |         |         |
CGCACGCGGC GCATCTCGGG CAGCGTTGGG TCCTGGCCAC GGGTGCGCAT GATCGTGCTC CTGTCTGTGA GGACCCGGCT AGGCTGGCGG GGTTCGCCTTA
      2.920      2.940      2.960      2.980      3.000
      |         |         |         |         |
CTGGTTAGCA GAATGAATCA CCGATACGGC AGCGAACGTG AAGGACTGCG TGCTGCAAAA CGTCTGCGAC CTGAGCAACA ACATGAATGG TCTTCGGTTT
      3.020      3.040      3.060      3.080      3.100
      |         |         |         |         |
CCGTGTTTCG TAAAGTCTGG AAACGCGGAA GTCCCTACG TGCTGCTGAA GTTGCCCGCA ACAGAGAGTG GAACCAACCG GTGATACCAC GATACTATGA
      3.120      3.140      3.160      3.180      3.200
      |         |         |         |         |
CTGAGAGTCA ACGCCATGAG CGGCCTCATT TCTTATTCG AGTTACAACA GTCCGACCG CTGTCCGGTA GCTCCTCCG GTGGGCGCGG GGCATGACTA
      3.220      3.240      3.260      3.280      3.300
      |         |         |         |         |
TCGTGCGCGC ACTTATGACT GTCTTCTTTA TCATGCAACT CGTAGGACAG GTGCCGGCAG CGCCCAACAG TCCCCCGGCC ACGGGGCTG CCACCATACC
      3.320      3.340      3.360      3.380      3.400
      |         |         |         |         |
CACGCCGAAA CAAGCGCCTC GCACCATTAT GTTCCGGATC TGCATCGCAG GATGCTGCTG GCTACCCTGT GGAACACCTA CATCTGTATT AACGAAGCGC
      3.420      3.440      3.460      3.480      3.500
      |         |         |         |         |
TAACCGTTTT TATCAGGCTC TGGGAGGCAG AATAAATGAT CATATCGTCA ATTATTACCT CCACGGGGAG AGCCTGAGCA AACTGGCCTC AGGCATTTGA
      3.520      3.540      3.560      3.580      3.600
      |         |         |         |         |
GAAGCACACG GTCACACTGC TTCGGTAGT CAATAAACCG GTAACACAGC AATAGACATA AGCGGCTATT TAACGACCCT GCCTGAACC GACGACCGGG
      3.620      3.640      3.660      3.680      3.700
      |         |         |         |         |
TCGAATTTGC TTTCGAATTT CTGCCATTCA TCCGCTTATT ATCATTATT CAGGCGTAGC ACCAGGCGTT TAAGGGCACC AATAACTGCC TTAATAAAT
      3.720      3.740      3.760      3.780
      |         |         |         |
TACGCCCGCG CGTGCCACTC ATCGCAGTAC TGTGTAAAT CATAAAGCAT TCTGCCGACA TGAAGCCAT CACAGACGGC ATGATGAACC TGAATCGCCA
      3.820      3.840      3.860      3.880
      |         |         |         |
CGGGCATCAG CACCTTGTGC CCTTGCATAT AATATTTGCC CATGGTGAAG ACGGGGGCGA AGAAGTTGTC CATATTTGCC ACGTTTAAAT CAAAACCTGT
      3.920      3.940      3.960      3.980
      |         |         |         |
GAAACTCACC CAGGGATTGG CTGAGACGAA AAACATATTC TCAATAAAC CTTTAGGAA ATAGGCCAGG TTTTCACCGT AACACGCCAC ATCTTGGCAA
      4.020      4.040      4.060      4.080
      |         |         |         |
TATATGTGTA GAAACTGCGG GAAATCGTCG TGGTATTAC TCCAGAGCGA TGAAAACGTT TCAGTTTGTCT CATGGAAAAC GGTGTAACAA GGGTGAACAC
      4.120      4.140      4.160      4.180
      |         |         |         |
TATCCCATAT CACCAGCTCA CCGTCTTTCA TTGCCATACG GAATCCCGGA TGAGCATTCA TCAGGCGGGC AAGAATGTGA ATAAAGGCCG GATAAAACTT
      4.220      4.240      4.260      4.280
      |         |         |         |
GTGCTTATTT TTTCTTACGG TCTTTAAAAA GGCCGTAATA TCCAGCTGAA CGGTCTGTTT ATAGGTACAT TGAGCAACTG ACTGAAATGC CTCAAAATGT
      4.320      4.340      4.360      4.380
      |         |         |         |
TCTTTACGAT GCCATTGGGA TATATCAACG GTGGTATATC CAGTGATTTT TTTCTCCATT TTAGCTTCCT TAGCTCCTGA AAATCTCGAT AACTCAAAAA
      4.420      4.440      4.460      4.480
      |         |         |         |
ATACGCCCGG TAGTGATCTT ATTTCAATTAT GGTGAAAGTT GGAACCTCTT ACGTGCCGAT CAACGTCTCA TTTTCGCCAA AAGTTGGCCC AGGGCTTCCC
      4.520      4.540      4.560      4.580
      |         |         |         |
GGTATCAACA GGGACACCAG GATTTATTTA TTCTGCGAAG TGATCTTCCG TCACAGGTAT TTATTCGGCG CAAAGTGCCT CGGGTGATGC TGCCAACCTA
      4.620      4.640      4.660      4.680
      |         |         |         |
CTGATTTAGT GTATGATGGT GTTTTGAGG TGCTCCAGTG GCTTCTGTTT CTATCAGCTG TCCCTCCTGT TCAGCTACTG ACGGGGTGGT CGGTAACGGC
      4.720      4.740      4.760      4.780
      |         |         |         |
AAAAGCACCG CCGGACATCA GCGCTGGCGG AGTGTATACT GGCTTACTAT GTTGGCACTG ATGAGGGTGT CAGTGAAGTG CTTCATGTGG CAGGAGAAAA
      4.820      4.840      4.860      4.880
      |         |         |         |
AAGGCTGCAC CGGTGCGTCA GCAGAAATAG TGATACAGGA TATATCCGC TCCCTCGCTC ACTGACTCGC TAGCTCGGT CGTTGCACTG CGGCGAGCGG
      4.920      4.940      4.960      4.980
      |         |         |         |
AAATGGCTTA CGAACGGGGC GGAGATTTCC TGAAGATGC CAGGAAGATA CTAAACAGGG AAGTGAGAGG GCCCGGGCAA AGCCGTTTTT CCATAGGCTC
      5.020      5.040      5.060      5.080      5.100
      |         |         |         |         |
CGCCCCCTG ACAAGCATCA CGAAATCTGA CGCTCAAATC AGTGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC GTTTCGCCCT GCGGCTCCC
      5.120      5.140      5.160      5.180      5.200
      |         |         |         |         |
TCGTGCGCTC TCCTGTTCTT GCCTTTCGGT TTACCGGTGT CATTCCGCTG TTATGGCCGC GTTTGTCTCA TTCCACGCC T GACACTCAGT TCCGGGTAGG
      5.220      5.240      5.260      5.280      5.300
      |         |         |         |         |
CAGTTCGCTC CAAGCTGGAC TGTATGCACG AACCCCCCGT TCAGTCCGAC CGCTGCGCCT TATCCGGTAA CTATGCTTT GAGTCCAACC CGGAAAGACA
      5.320      5.340      5.360      5.380      5.400
      |         |         |         |         |
TGCAAAAGCA CCACTGGCAG CAGCCACTGG TAATTGATTT AGAGGAGTTA GTCTTGAAGT CATGCGCCGG TTAAGGCTAA ACTGAAAGGA CAAGTTTTGG
      5.420      5.440      5.460      5.480      5.500
      |         |         |         |         |
TGACTGCGCT CCTCCAAGCC AGTTACCTCG GTTCAAAGAG TTGGTAGCTC AGAGAACCTT CGAAAAACCG CCCTGCAAGG CGGTTTTTTC GTTTTCAGAG
      5.520      5.540      5.560      5.580      5.600
      |         |         |         |         |
CAAGAGATTA CGCGCAGACC AAAACGATCT CAAGAAGATC ATCTTATTAA TCAGATAAAA TATTTCTAGA TTTCAAGTGA ATTTATCTCT TCAAATGTAG
      5.620      5.640      5.660      5.680
      |         |         |         |
CACCTGA

```

alphaVmut-TMD = TMD Int_a5GP

JM = Juxtamembrane region

CAT = Chloramphenicol acetyltransferase

9.6.2 C-BLa 1.1

20 40 60 80 AraC
 <<TGACAACCTTG ACGGCTACAT CATTCACTTT TTCTTCACAA CCGGCACGGA ACTCGCTCGG GCTGGCCCCG GTGCATTTTT TAAATACCCG CGAGAAATAG
 120 140 160 180 AraC
 AGTTGATCGT CAAAACCAAC ATTGCGACCG ACGGTGGCGA TAGGCATCCG GGTGGTGCTC AAAAGCAGCT TCGCCTGGCT GATACGTTGG TCCTCGCGCC
 220 240 260 280 AraC
 AGCTTAAGAC GCTAATCCCT AACTGCTGGC GGAAAAGATG TGACAGACGC GACGGCGACA AGCAAACATG CTGTGCGACG CTGGCGATAT CAAAATTGCT
 320 340 360 380 AraC
 GTCTGCCAGG TGATCGCTGA TGTACTGACA AGCCTCGCGT ACCCGATTAT CCATCGGTGG ATGGAGCGAC TCGTTAATCG CTTCCATGCG CCGCAGTAAC
 420 440 460 480 AraC
 AATTGCTCAA GCAGATTTAT CGCCAGCAGC TCCGAATAGC GCCCTTCCCC TTGCCCGCGG TTAATGATTT GCCCAAACAG GTCGCTGAAA TGGCGCTGGT
 520 540 560 580 AraC
 GCGCTTCATC CCGGCGAAG AACCCTGTAT TGGCAAATAT TGACGGCCAG TTAAGCCATT CATGCCAGTA GGCGCGCGGA CGAAAGTAAA CCCACTGGTG
 620 640 660 680 AraC
 ATACCATTTC CGAGCCTCCG GATGACGACC GTAGTGATGA ATCTCTCCTG GCGGGAACAG CAAAATATCA CCCGGTCCGG AAACAAATTC TCGTCCCTGA
 720 740 760 780 AraC
 TTTTTCACCA CCCCTGACC GCGAATGGT AGATTGAGAA TATAACCTTT CATTCCACGC GGTCGGTCTGA TAAAAAATC GAGATAACCG TTGCCTCAA
 820 840 860 880 AraC
 TCGGCGTTAA ACCCGCCACC AGATGGGCAT TAAACGAGTA TCCCGGCAGC AGGGGATCAT TTTGCGCTTC AGCCATACTT TTCATACTCC CGCCATTGAG
 arabinose_O2_operator_region
 AraC
 940 960 980 1,000
 AGAAGAAACC AATTGTCCAT ATTGCATCAG ACATTGCCGT CACTGCGTCT TTTACTGGCT CTTCTCGCTA ACCAAACCGG TAACCCCGCT TATTAAGAGC
 1,020 1,040 1,060 1,080
 ATTTCTGTAAC AAAGCGGGAC CAAAGCCATG ACAAAAACGC GTAACAAAAG TGCTATAAT CACGGCAGAA AAGTCCACAT TGATTATTTG CACGGCGTCA
 CAP binding site
 1,120 1,140 1,160 1,180 1,200
 CACTTTGCTA TGCCATAGCA TTTTATCCA TAAGATTAGC GGATGCTACC TGACGCTTTT TATCGCAACT CTCTACTGTT TCTCCATACC CGTTTTTTTTG
 XbaI
 1,220 1,240 1,260 1,280 1,300
 GGCTCTAGAT AACGAGGGCA AAAAATGCAT TATAAATGCG TGTTTGCGGC CGCGGGCCGG AGCCTGGCGA GCGCGGTGAG CGCTCACTTA CTACTCTAG
 C-BLa
 1,320 1,340 1,360 1,380 1,400
 CTTCCCGGCA ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACTTCTGC GCTCGGCCCT TCCGGTCCGG TGGTTTATTG CTGATAAATC
 C-BLa
 1,420 1,440 1,460 1,480 1,500
 TGGAGCCGGT GAGCGTGGT CTGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGACGGG GAGTCAGGCA
 C-BLa
 1,520 1,540 1,560 1,580 1,600
 ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT TAAGCATTGG ACTAGTGGCA GCGGGCGCGG CAGCGGGCGG GGCAGCGGTT
 SpeI
 flexible linker
 1,640 1,660 1,680 1,700
 CTGCTAATCG AGCTAGCGTG ATCATTCTGG CGGTGGGGCC CGGCCCTGCT CTCCTGGGGA TCCACAAACT GGCAGTTAAT GTTCAGCTGC CTGCCGAAGC
 flexJM linker
 NheI
 alphaVmut-TMD
 1,720 1,740 1,760 1,780 1,800
 AGCAGCGAAA GAAGCCCGCT CCAAAGAAGC GGCAGCGAAA GAGGCTGCGC CGAAAGAGCC AGCAGCTAAA GCCGCAAGCA AAGGTGAAGA ACTGTTTACC
 rigid linker
 sfGFP⁰
 1,820 1,840 1,860 1,880 1,900
 GGTGTTGTTT CGATTCTGGT TGAAGTGGAT GGTGATGTTA ATGGCCACAA ATTCTCAGTT CGTGGTGAAG GCGAAGGTGA TGCAACCAAT GGTAAACTGA
 sfGFP
 1,920 1,940 1,960 1,980 2,000
 CCCTGAAATT TATCTGTACC ACCGGCAAAC TGCCGGTTCG GTGGCCGACC CTGTTACCA CCCTGACCTA TGGTGTTCAG TGTTTCAGCC GTTATCCGGA
 sfGFP
 2,020 2,040 2,060 2,080 2,100
 TCATATGAAA CAGCAGGATT TCTTCAAATC TGCAATGCCG GAAGGTTATG TTCAAGAAGC TACCATCTCC TTTAAAGATG ATGGCACCTA TAAGACCCGT
 sfGFP
 2,120 2,140 2,160 2,180 2,200
 GCCGAAGTTA AATTGGAAGG TGATACCCTG GTGAATCGCA TTGAAGTAA AGGCATCGAT TTCAAAGAAG ATGGTAATAT CCTGGGCCAT AAACCTGGAAT
 sfGFP
 2,220 2,240 2,260 2,280 2,300
 ATAATTTCAA CAGCCACAAC GTGTATATCA CCGCAGATAA ACAGAAGAAT GGCATCAAAG CCAACTTTAA GATCCGCCAT AATGTTGAAG ATGGCAGCGT
 sfGFP
 2,320 2,340 2,360 2,380 2,400
 TCAGCTGGCA GATCATTATC AGCAGAATAC ACCGATTGGT GATGGTCCGG TTCTGCTGCC GGATAATCAT TATCTGAGCA CCCAGAGCGT TCTGAGCAAA
 sfGFP
 2,420 2,440 2,460 2,480 2,500
 GATCCGAATG AGAAACGTGA TCACATGGTG CTGCTGGAAT TTGTTACCGC AGCAGGTATT ACCCATGGTA TGGATGAECT GTATAAAGGC GATTATAAGG
 FLAG-Tag
 XhoI
 PstI
 HindIII
 2,520 2,540 2,560 2,580 2,600
 ATCACGACGG CTAACCTCGAG GTCGACCTGC AGGCATGCAA GCTTGGCTGT TTTGGCGGAT GAGAGAAGAT TTTACGCTG ATACAGATTA AATCAGAACG
 2,620 2,640 2,660 2,680 2,700
 CAGAAGCGGT CTGATAAAAC AGAATTGGC TGGCGGCAGT AGCGCGGTGG TCCCACCTGA CCCCATGCCG AACTCAGAAG TGAACCGCCG TAGCGCCGAT
 rmB_transcriptional_termination...
 2,720 2,740 2,760 2,780 2,800
 GGTAGTGTGG GGTCTCCCA TGCGAGAGTA GGAACCTGCC AGGCATCAA TAAAACGAAA GGCTCAGTCG AAAGACTGGG CCTTTCGTTT TATCTGTTGT

2,820 | 2,840 | 2,860 | 2,880 | 2,900
 TTGTCCGTGA ACGCTCTCCT GAGTAGGACA AATCCGCCGG GAGCGGATTT GAACGTTGCG AAGCAACGGC CCGGAGGGTG GCGGGCAGGA CGCCCCCAT
 2,920 | 2,940 | 2,960 | 2,980 | 3,000
 AACTGCCAG GCATCAAAT AAGCAGAAGG **rrnB_T2_transcrip_term** CCATCCTGAC GGATGGCCTT TTTCGTTTC TACAACTCT TTTGTTTAT TTTCTAAATA CATTCAAATA
 3,020 | 3,040 | 3,060 | 3,080 | 3,100
 TGTATCCGCT CATGGAGCTC AAAGCCACGT TGTGTCTCAA AATCCTGTAT GTTACATTGC ACAAGATAAA AATATATCAT CATGAACAAT AAAACTGTCT
 3,120 | 3,140 | 3,160 | 3,180 | 3,200
 GCTTACATAA ACAGTAATAC AAGGGGTGTT **KanR** ATGAGCCATA TTCAACGGGA AACGCTTGC TCCAGGCCGC GATTAATTC CAACATGGAT GCTGATTTAT
KanR 3,220 | 3,240 | 3,260 | 3,280 | 3,300
 ATGGGTATAA ATGGGCTCGC GATAATGTCG GGCAATCAGG TGCACAATC TATCGTTTGT ATGGGAAGCC CGATGCGCCA GAGTTGTTTC TGAAACATGG
KanR 3,320 | 3,340 | 3,360 | 3,380 | 3,400
 CAAAGGTAGC GTTGCCAATG ATGTTACAGA TGAGATGGTC AGACTAAACT GGCTGACGGA ATTTATGCCT CTTCCGACCA TCAAGCATT TATCCGGACT
KanR 3,420 | 3,440 | 3,460 | 3,480 | 3,500
 CCTGATGAAG CATGGTACT CACCCTGCG ATCCCCGGA AAACAGCATT CCAGGTATTA GAAGAATATC CTGATTCAGG TGAAATATT GTTGATGCGC
KanR 3,520 | 3,540 | 3,560 | 3,580 | 3,600
 TGGCAGTGT CCTGCGCCGG TTGCATTGCA TTCCTGTTG TAATGTCTCT TTTAAGCAGG ATCGCGTATT TCGTCTCGCT CAGGCACAAT CACGAATGAA
KanR 3,620 | 3,640 | 3,660 | 3,680 | 3,700
 TAACGGTTTG GTTGATGCGA GTGATTTTGA TGACGAGCGT AATGGCTGGC CTGTTGAACA AGTCTGAAA GAAATGCACA AACTTTTGCC ATTCTCACC
KanR 3,720 | 3,740 | 3,760 | 3,780 | 3,800
 GATTCAGTCG TCACCTATGG TGATTTCTCA CTTGATAACC TTATTTTGA CGAGGGGAAA TTAATAGTT GTATTGATGT TGGACGAGTC GGAATCGCAG
KanR 3,820 | 3,840 | 3,860 | 3,880 | 3,900
 ACCGATACCA GGATCTTGCC ATCCTATGGA ACTGCCTCGG TGAGTTTTCT CTTCTATTAC AGAAACGGCT TTTTCAAAA TATGGTATTG ATAATCTGA
KanR 3,920 | 3,940 | 3,960 | 3,980 | 4,000
 TATGAATAAA TTGCAGTTTC ATTTGATGCT CGATGAGTTT TCTAAGAGC TCTCCATGAC CAAAATCCCT TAACGTGAGT TTTCTTCCA CTGAGCGTCA
 4,020 | 4,040 | 4,060 | 4,080 | 4,100
 GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGCAA CAAAAAAC ACCGCTACCA GCGGTGGTTT
 4,120 | 4,140 | 4,160 | 4,180 | 4,200
 GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGT AACTGGCTTC AGCAGAGCGC AGATACCAA TACTGTCTT CTAGTGTAGC CGTAGTTAGG
 4,220 | 4,240 | 4,260 | 4,280 | 4,300
 CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGC ATAAGTCGT TCTTACCGGG
 4,320 | 4,340 | 4,360 | 4,380 | 4,400
 TTGACTCAA GACGATAGT ACCGATAAG GCGCAGCGT CGGGCTGAAC GGGGGTTTCG TGCACACAGC CCAGCTTGA GCGAACGACC TACCCGAAC
 4,420 | 4,440 | 4,460 | 4,480 | 4,500
 TGAGATACCT ACAGCGTAG CTATGAGAAA GCGCCACGCT TCCGAAGGG AGAAAGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGAA CAGGAGAGCG
 4,520 | 4,540 | 4,560 | 4,580 | 4,600
 CACGAGGAG CTTCCAGGGG GAAACGCTG GTATCTTTAT AGTCTGTGCG GGTTCGCGA CCTCTGACTT GAGCGTCGAT TTTTGTGAT CTCGTCAGGG
 4,620 | 4,640 | 4,660 | 4,680 | 4,700
 GGGCGGAGCC TATGGAAAA CGCCAGCAAC GCGGCTTTT TACGGTCTCT GCGCTTTTGC TGGCTTTTG CTCACATGTT CTTTCTCGG TTATCCCTG
 4,720 | 4,740 | 4,760 | 4,780 | 4,800
 ATTCTGTGGA TAACCGTATT ACCGCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCGGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA
 4,820 | 4,840 | 4,860 | 4,880 | 4,900
 GCGCCTGATG CGGTATTTTCT TCTTACGCA TCTGTGCGGT ATTTACACCC GCATATGGTG CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAAGC
 4,920 | 4,940 | 4,960 | 4,980 | 5,000
 CAGTATACAC TCCGCTATCG CTACGTGACT GGGTCATGGC TGCGCCCGA CACCCGCCAA CACCCGCTGA CGCGCCCTGA CGGGCTTGTG TGCTCCCGC
 5,020 | 5,040 | 5,060 | 5,080 | 5,100
 ATCCGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC ATGTGTCAGA **GGTTTTTACC** **GTCATCACC** **AAACGCGGA** **GGCAGCTGG** **GTAAGCTCA** **ROP_ORF**
 5,120 | 5,140 | 5,160 | 5,180 | 5,200
 TCAGCGTGGT CGTGAAGCGA TTCACAGATG TCTGCCTGTT CATCCGCGTC CAGCTCGTTG AGTTTCTCCA GAAGCGTTAA TGTCTGGCTT CTGATAAAGC **ROP_ORF**
 5,220 | 5,240 | 5,260 | 5,280 | 5,300
 GGGCCATGTT AAGGGCGGT TTTTCTGTT TGCTCACTGA TGCCCTCGTG TAAGGGGAT TTCTGTTTCT GGGGTAATG ATACCGATGA AACGAGAGAG
 5,320 | 5,340 | 5,360 | 5,380 | 5,400
 GATGCTCACG ATACGGGTTA CTGATGATGA ACATGCCCGG TTACTGGAAC GTTGTGAGGG TAAACAACCTG GCGGTATGGA TCGCGCGGGA CCAGAGAAAA
 5,420 | 5,440 | 5,460 | 5,480 | 5,500
 ATCACTCAGG GTCATGCGCA GCGCGATGCA TAATGTGCTT GTCAAATGGA CGAAGCAGGG ATTCTGCAA CCCTATGCTA CTCGCTCAAG CCGTCAATTG
 TCTGATTCGT TACCAATTA)

alphaVmut-TMD = TMD Int_a5GP

JM = Juxtamembrane region

KanR = Aminoglycoside O-phosphotransferase

9.6.3 N-BLa 1.2

```

20          40          60          80          100
|AGTCAGCCCC ATACGATATA AGTTGTAATT CTCATGTTTG ACAGCTTATC AATCGATAAG AAACCAATTG TCCATATTGC ATCAGACATT GCCGCTACTG
|
120          140          160          180          200
|CGTCTTTTAC TGGCTCTTCT CGTAAACCAA ACCGTAACC CGGCTTATTA AAAGCATTCT GTAACAAAGC GGGACCAAAG CCATGACAAA AACCGCTAAC
|
220          240          260          280          300
|AAAAGTGTCT ATAATCACGG CAGAAAAGTC CACATTGATT ATTTGCACGG CGTCACACTT TGCTATGCCA TAGCATTTTT ATCCATAAGA TTAGCGGATG
|
320          340          360          380          400
|CTACCTGACG CTTTTTATCG CAACTCTCTA CTGTTTCTCC ATACCCGTTT TTTTGGGCTC TAGATAACGA GGGCAAAAAA TGAAAAAGAC CGCGATTGCG
|
420          440          460          480          500
|ATTGCGGTGG CGCTGGCGGG CTTTGCAGCC GTGGCGCAGG CGGCGCCAGA AACGCTGGTG AAAGTAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG
|
520          540          560          580          600
|GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCITGA GAGTTTTCGC CCCGAAGAAG GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG
|
620          640          660          680          700
|CGCGGTATTA TCCCGTGTG ACGCCGGGCA AGAGCAACTC GGTCCGCCGA TACACTATTC TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG
|
720          740          760          780          800
|CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCGAGT CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC
|
820          840          860          880          900
|CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACTCGC CTTGATCGTT GGGAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG
|
920          940          960          980          1000
|TGACACCACG ACGCCTGTAG CAATGGCAAC AACGTTGCGC AAACATTAA CTGGCACTAG TGGCAGCGGC GCGCGCAGCG CGCGCGCGAG CGGTTCTGCT
|
1020          1040          1060          1080          1100
|AATCGAGCTA GCGTGATCAT TCTGGCGGTG GGGCCCGGCC TGCTGCTCCT GGGGATCCAC AAAGTGGCAG TTAATGTCCA GCTGCCGTGC GAAGCAGCAG
|
1120          1140          1160          1180          1200
|CGAAAGAAGC CGTGCCAAA GAAGCGGCAG CGAAAGAGGC TGCCCGGAAA GAGGCAGCAG CTAAAGCCGC AAGCAAAGGT GAAGAACTGT TTACCGGTGT
|
1220          1240          1260          1280          1300
|TGTTCCGATT CTGGTTGAA C TGGATGGTGA TGTTAATGGC CACAAATTC CAGTTCGTGG TGAAGGCGAA GGTGATGCAA CCAATGGTAA ACTGACCCCTG
|
1320          1340          1360          1380          1400
|AAATTTATCT GTACCACCGG CAAACTGCCG GTTCCGTGGC CGACCCCTGT TACCACCTGT ACCTATGGTG TTCAGTGTTC CAGCCGTTAT CCGGATCATA
|
1420          1440          1460          1480          1500
|TGAAACAGCA CGATTTCTTC AAATCTGCAA TGCCGGAAGG TTATGTTCAA GAACGTACCA TCTCCTTTAA AGATGATGGC ACCTATAAGA CCCGTGCCGA
|
1520          1540          1560          1580          1600
|AGTTAAATTT GAAGGTGATA CCCTGGTGAA TCGCATTGAA CTGAAAGGCA TCGATTTCAA AGAAGATGGT AATATCCTGG GCCATAAACT GGAATATAAT
|
1620          1640          1660          1680          1700
|TTCAACAGCC ACAACGTGTA TATCACCAGCA GATAAACAGA AGAATGGCAT CAAAGCCAAC TTTAAGATCC GCCATAAATG TGAAGATGGC AGCGTTCCAGC
|
1720          1740          1760          1780          1800
|TGGCAGATCA TTATCAGCAG AATACACCGA TTGGTGATGG TCCGGTTCTG CTGCCGGATA ATCATTATCT GAGCACCCAG AGCGTTCTGA GCAAAGATCC
|
1820          1840          1860          1880          1900
|GAATGAGAAA CGTGATCACA TGGTGTCTGCT GGAATTTGTT ACCGCGACGAG GTATTACCCA TGGTATGGAT GAAGCTGATA AAGCGGATTA TAAAGATCAC
|
1920          1940          1960          1980          2000
|GACGGCTAAC TCGAGCACCA CCACCACCAC CACTGAGATC CGGCTGCTAA CAAAGCCCGA AAGGAAGCTG AGTTGGCTGC TGCCACCCTG GAGCAATAAC
|
2020          2040          2060          2080          2100
|TAGCATAACC CCTTGGGGCC TCTAAACGGG TCTTGAGGGG TTTTTTGACC GATGCCCTTG AGAGCCTTCA ACCCAGTCAG CTCCTTCCGG TGGCGCGGGG
|
2120          2140          2160          2180          2200
|GCATGACTAT CGTCGCCGA CTTATGACTG TCTTCTTTAT CATGCAACTC GTAGGACAGG TGCCGGCAGC GCTCTGGGTC ATTTTCGGCG AGGACCGGTT
|
2220          2240          2260          2280          2300
|TCGCTGGAGC GCGACGATGA TCGCCTGTCT GCTTGGGTA TTCGGAATCT TGCACGCCCT CGCTCAAGCC TTCGTCACGT GTCCCGCCAC CAAACGTTTC
|
2320          2340          2360          2380          2400
|GGCGAGAAGC AGGCCATTAT CGCCGGCATG GCGGCCGACG CGCTGGGCTA CGTCTTGCTG GCGTTCCGGA CCGGAGGCTG GATGGCCTTC CCCATTATGA
|
2420          2440          2460          2480          2500
|TTCTTCTCGC TTCCGGCGGC ATCGGGATGC CCGCGTTGCA GGCCATGCTG TCCAGGCAGG TAGATGACGA CCATCAGGGA CAGCTTCAAG GATCGCTCGC
|
2520          2540          2560          2580          2600
|GGCTCTTACC AGCCTAACTT CGATCACTGG ACCGCTGATC GTCACGGCGA TTTATGCCGC CTCGGCGAGC ACATGGAACG GTTTGGCATG GATTGTAGGC
|
2620          2640          2660          2680          2700
|GCCGCCCTAT ACCTTGTCTG CCTCCCCGCG TTGCGTCCGC GTGCATGGAG CCGGCCACCC TCGACCTGAA TGAAGCCCGG CGGCACCTCG CTAACGGATT
|
2720          2740          2760          2780          2800
|CACCCTCCA AGAATTGGAG CCAATCAATT CTTGCGGAGA ACTGTGAATG CGAAACCAA CCCTTGGCAG AACATATCCA TCGCGTCCGC CATCTCCAGC

```



```

      2,820      2,840      2,860      2,880      2,900
      |         |         |         |         |
AGCCGCACGC GCGCATCTC GGCAGCGTT GGGTCTGGC CACGGGTGCG CATGATCGT CTCCTGTCTG TGAGGACCCG GCTAGGCTGG CGGGGTTGGC

      2,920      2,940      2,960      2,980      3,000
      |         |         |         |         |
TTACTGGTTA GCAGAATGAA TCACCGATAC GCGAGCGAAC GTGAAGCGAC TGCTGCTGCA AAACGTCTGC GACCTGAGCA ACAACATGAA TGGTCTTCGG

      3,020      3,040      3,060      3,080      3,100
      |         |         |         |         |
TTTCCGTGTT TCGTAAAGTC TGGAAACGCG GAAGTCCCCT ACGTGCTGCT GAAGTTGCCG GCAACAGAGA GTGGAACCAA CCGGTGATAC CACGATACTA

      3,120      3,140      3,160      3,180      3,200
      |         |         |         |         |
TGACTGAGAG TCAACGCCAT GAGCGGCCTC ATTTCTTATT CTGAGTTACA ACAGTCCGCA CCGCTGTCCG GTAGCTCCTT CCGGTGGGCG CGGGGCATGA

      3,220      3,240      3,260      3,280      3,300
      |         |         |         |         |
CTATCGTTCG CGCACTTATG ACTGTCTTCT TTATCATGCA ACTCGTAGGA CAGGTGCCGG CAGCGCCCAA CAGTCCCCTG GCCACGGGGC CTGCCACCAT

      3,320      3,340      3,360      3,380      3,400
      |         |         |         |         |
ACCCACGCGG AAACAAGCGC CCTGCACCAT TATGTTCCGG ATCTGCATCG CAGGATGCTG CTGGCTACCC TGTGGAACAC CTACATCTGT ATTAACGAAG

      3,420      3,440      3,460      3,480      3,500
      |         |         |         |         |
CGCTAACCGT TTTTATCAGG CTCTGGGAGG CAGAATAAAT GATCATATCG TCAATTATTA CCTCCACGGG GAGAGCCTGA GCAAACCTGGC CTCAGGCATT

      3,520      3,540      3,560      3,580      3,600
      |         |         |         |         |
TGAGAAGCAC ACGGTCACAC TGCTTCCGGT AGTCAATAAA CCGTAAAC AGCAATAGAC ATAAGCGGCT ATTTAACGAC CTGCGCTGA ACCGACGACC

      3,620      3,640      3,660      3,680      3,700
      |         |         |         |         |
GGGTGCAATT TGCTTTCGAA TTTCTGCCAT TCATCCGCTT ATTATCACTT ATTCAGGCGT AGCACCAGGC GTTTAAGGGC ACCAATAACT GCCTTAAAAA

      3,720      3,740      3,760      3,780
      |         |         |         |
AATTAGCGCC CGCCCTGCCA CTCATCGCAG TACTGTTGTA ATTCATTAAG CATTCTGCCG ACATGGAAGC CATCACAGAC GGCATGATGA ACCTGAATCG CAT

      3,820      3,840      3,860      3,880
      |         |         |         |
CCAGCGGCAT CAGCACCTTG TCGCCTTGCG TATAATATT GCCCATGGTG AAAACGGGGG CGAAGAAGTT GTCCATATTG GCCACGTTTA AATCAAAACT CAT

      3,920      3,940      3,960      3,980
      |         |         |         |
GGTGAAACTC ACCCAGGGAT TGGCTGAGAC GAAAAACATA TTCTCAATAA ACCCTTAGG GAAATAGGCC AGGTTTTTAC CGTAACACGC CACATCTTGC CAT

      4,020      4,040      4,060      4,080
      |         |         |         |
GAATATATGT GTAGAAACTG CCGGAAATCG TCGTGGTATT CACTCCAGAG CGATGAAAC GTTTCAGTTT GCTCATGGAA AACGGGTGAA CAAGGGTGAA CAT

      4,120      4,140      4,160      4,180
      |         |         |         |
CACTATCCCA TATCACCAGC TCACCGTCTT TCATTGCCAT ACGGAATTCG GGATGAGCAT TCATCAGGCG GGCAAGAATG TGAATAAAGG CCGGATAAAA CAT

      4,220      4,240      4,260      4,280
      |         |         |         |
CTTGTGCTTA TTTTCTTTA CCGTCTTTAA AAAGGCCGTA ATATCCAGCT GAACGGTCTG GTTATAGGTA CATTGAGCAA CTGACTGAAA TGCCTCAAAA CAT

      4,320      4,340      4,360      4,380      4,400
      |         |         |         |         |
TGTTCTTTAC GATGCCATTG GGATATATCA ACGGTGGTAT ATCCAGTGAT TTTTCTTCC ATTTTAGCTT CCTTAGCTCC TGAANAATCTC GATAACTCAA CAT

      4,420      4,440      4,460      4,480      4,500
      |         |         |         |         |
AAAATACGCC CGGTAGTGAT CTTATTTTCA TATGGTAAA GTTGAACCT CTTACGTGCC GATCAACGTC TCATTTTTCG CAAAAGTTGG CCCAGGGCTT CAT

      4,520      4,540      4,560      4,580      4,600
      |         |         |         |         |
CCCAGGATCA ACAGGGACAC CAGGATTTAT TTATTTCTGC AAGTGATCTT CCGTCACAGG TATTTATTCG GCGCAAAGTG CGTCGGGTGA TGCTGCCAAC CAT

      4,620      4,640      4,660      4,680      4,700
      |         |         |         |         |
TTACTGATTT AGTGATGAT GGTGTTTTTG AGGTGCTCCA GTGGCTTCTG TTTCTATCAG CTGTCCCTCC TGTTACGCTA CTGACGGGGT GGTGCGTAAC CAT

      4,720      4,740      4,760      4,780
      |         |         |         |
GGCAAAAGCA CCGCCGGACA TCAGCGCTGG CGGAGTGAT ACTGGCTTAC TATGTTGGCA CTGATGAGGG TGTCAGTGAA GTGCTTCATG TGGCAGGAGA p15a_ORI

      4,820      4,840      4,860      4,880
      |         |         |         |
AAAAAGGCTG CACCGGTGGC TCAGCAGAAT ATGTGATACA GGATATATTC CGCTTCTCG CTCACTGACT CGCTACGCTC GGTCGTTTCA CTGCGGCGAG p15a_ORI

      4,920      4,940      4,960      4,980      5,000
      |         |         |         |         |
CGGAAATGCG TTACGAACGG GCGGAGATT TCCTGGAAGA TGCCAGGAAG AACTTAACA GGAAGTGAG AGGGCCGCG CAAAGCCGTT TTTCCATAGG p15a_ORI

      5,020      5,040      5,060      5,080      5,100
      |         |         |         |         |
CTCCGCCCCC CTGACAAGCA TCACGAAATC TGACGCTCAA ATCAGTGGTG GCGAAACCGG ACAGGACTAT AAAGATACCA GCGGTTTCCC CTGGCGGCT CAT

      5,120      5,140      5,160      5,180      5,200
      |         |         |         |         |
CCCTCGTGCG CTCTCCTGTT CCTGCCTTTC GGTTACC GGTCATTCCG CTGTTATGGC CGCGTTTGTG TCATTCCACG CCTGACACTC AGTTCGGGT CAT

      5,220      5,240      5,260      5,280      5,300
      |         |         |         |         |
AGGCAGTTCG CTCCAAGCTG GACTGTATGC ACGAACCCCG CGTTCAGTCC GACCGCTGCG CCTTATCCGG TAACATATCGT CTTGAGTCCA ACCCGAAAG CAT

      5,320      5,340      5,360      5,380      5,400
      |         |         |         |         |
ACATGCAAAA GCACCACTGG CAGCAGCCAC TGTAATTGA TTTAGAGGAG TTAGTCTTGA AGTCATGCGC CGGTTAAGGC TAAACTGAAA GGACAAGTTT CAT

      5,420      5,440      5,460      5,480      5,500
      |         |         |         |         |
TGGTGACTGC GCTCTCCAA GCCAGTTACC TCGGTTCAA GAGTTGGTAG CTCAGAGAAC CTTGAAAAA CCGCCCTGCA AGGCGGTTTT TTCGTTTTCA CAT

      5,520      5,540      5,560      5,580      5,600
      |         |         |         |         |
GAGCAAGAGA TTACGCGCAG ACCAAAACGA TCTCAAGAAG ATCATCTTAT TAATCAGATA AAATATTTCT AGATTTCAGT GCAATTTATC TCTTCAAATG CAT

TAGCACCTGA)

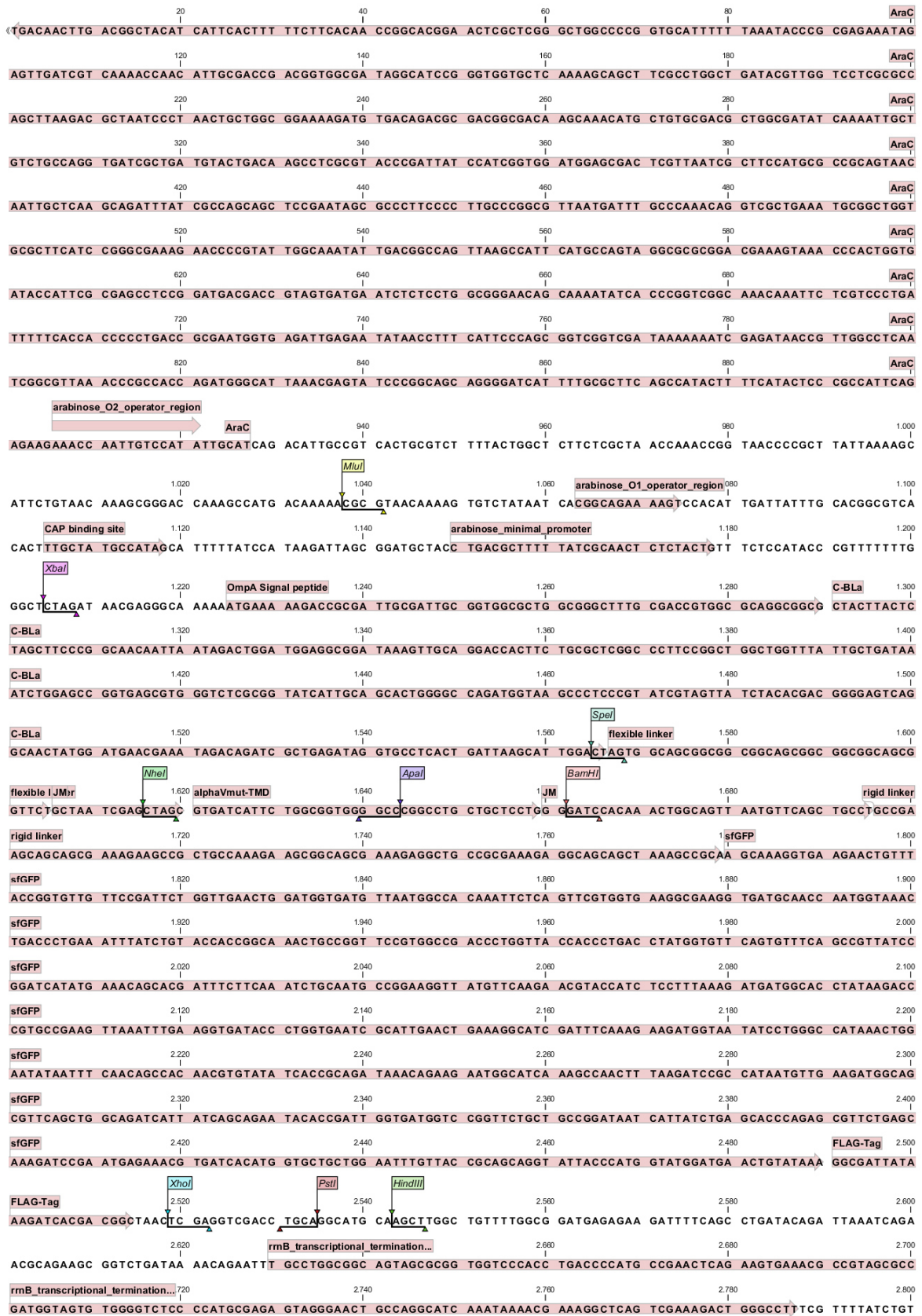
```

alphaVmut-TMD = TMD Int_a5GP

JM = Juxtamembrane region

CAT = Chloramphenicol acetyltransferase

9.6.4 C-BLa 1.2



```

                2,820                2,840                2,860                2,880                2,900
TGTTTGTCCG TGAACGCTCT CTTGAGTAGG ACAAATCCGC CGGGAGCGGA TTTGAACGTT GCGAAGCAAC GGCCCGGAGG GTGGCGGGCA GGACGCCCGC
                2,920                rrnB_T2_transcrip_term0                2,960                2,980                3,000
CATAAACTGC CAGGCATCAA ATTAAGCAGA AGGCCATCCT GACGGATGGC CTTT TGCCTT TCTACAAAC TCTTTTGTTT ATTTTCTAA ATACATTCAA
                3,020                3,040                3,060                3,080                3,100
ATATGTATCC GCTCATGGAG CTCAAAGCCA CGTTGTGTCT CAAAATCTCT GATGTTACAT TGCACAAGAT AAAAATATAT CATCATGAAC AATAAAACTG
                3,120                KanR                3,140                3,160                3,180                3,200
TCTGCTTACA TAAACAGTAA TACAAGGGGT GTTATGAGCC ATATTCAACG GGAAACGCTCT TGCTCCAGGC CGCGATTAAA TTCCAACATG GATGCTGATT
KanR                3,220                3,240                3,260                3,280                3,300
TATATGGGTA TAAATGGGCT CGCGATAATG TCGGGCAATC AGGTGCGACA ATCTATCGTT TGTATGGGAA GCCCGATGCG CCAGAGTTGT TTCTGAAACA
KanR                3,320                3,340                3,360                3,380                3,400
TGCCAAAGGT AGCGTTGCCA ATGATGTTAC AGATGAGATG GTCAGACTAA ACTGGCTGAC GGAATTTATG CCTCTTCCGA CCATCAAGCA TTTTATCCGG
KanR                3,420                3,440                3,460                3,480                3,500
ACTCCTGATG AAGCATGGTT ACTCACCCT GCGATCCCG CCAAACAGC ATTCCAGTA TTAGAAGAAT ATCCTGATTC AGGTGAAAAT ATTGTTGATG
KanR                3,520                3,540                3,560                3,580                3,600
CGCTGGCAGT GTTCTCGCCG CGGTTGCATT CGATTCCCTGT TTGTAATTTG CCTTTTAAACA GCGATCGCGT ATTTTCGTC GCTCAGGCGC AATCACGAAT
KanR                3,620                3,640                3,660                3,680                3,700
GAATAACGGT TTGGTTGATG CGAGTGATTT TGATGACGAG CGTAATGGCT GGCCTGTTGA ACAAGTCTGG AAAGAAATGC ACAAACTTTT GCCATTCTCA
KanR                3,720                3,740                3,760                3,780                3,800
CCGGATTGAG TCGTCACTCA TGGTGATTTT TCACCTGATA ACCTTATTTT TGACGAGGGG AAATTAATAG GTTGTATTGA TGTGGACGA GTCGGAATCG
KanR                3,820                3,840                3,860                3,880                3,900
CAGACCAGTA CCAGGATCTT GCCATCCTAT GGAACCTGCT CGGTGAGTTT TCTCCTCAT TACAGAAACG GCTTTTTCAA AAATATGGTA TTGATAATCC
KanR                3,920                3,940                3,960                3,980                4,000
TGATATGAAT AAATGTCAGT TTCATTGAT GCTCGATGAG TTTTCTAAG AGCTCTCCAT GACCAAAATC CCTTAACGTG AGTTTTCTGT CCACTGAGCG
                4,020                4,040                4,060                4,080                4,100
TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG
                4,120                4,140                4,160                4,180                4,200
TTTGTGTTCC GGATCAAGAG CTACCAACTC TTTTCCGAA GGTAAGTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT
                4,220                4,240                4,260                4,280                4,300
AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC
                4,320                4,340                4,360                4,380                4,400
GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTGCGGCTG AACGGGGGGT TCGTGCACAC AGCCAGCTT GGAGCGAAGC ACCTACCCG
                4,420                4,440                4,460                4,480                4,500
AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA
                4,520                4,540                4,560                4,580                4,600
GCGCACGAGG GAGCTCCAG GGGGAAACGC CTGGTATCTT TATAGTCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTGTG ATGCTCGTCA
                4,620                4,640                4,660                4,680                4,700
GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGGCC TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTTCTT GCGTTATCC
                4,720                4,740                4,760                4,780                4,800
CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGACGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA
                4,820                4,840                4,860                4,880                4,900
AGAGCGCCTG ATGCGGTATT TTCTCCTTAC GCATCTGTGC GGTATTTTAC ACCGCATATG GTGCACTCTC AGTACAATCT GCTCTGATGC CGCATAGTTA
                4,920                4,940                4,960                4,980                5,000
AGCCAGTATA CACTCCGCTA TCGTACGTG ACTGGGTCAT GGCTGCGCCC CGACACCCGC CAACACCCGC TGACGCGCCC TGACGGGCTT GTCTGCTCCC
                5,020                5,040                5,060                5,080                ROP_ORF
GGCATCCGCT TACAGACAAG CTGTGACCGT CTCCGGGAGC TGCATGTGTC AGAGGTTTTT ACCGTCATCA CCGAAACCGC CGAGGCAGCT GCGGTAAGGC
                5,120                5,140                5,160                5,180                ROP_ORF
TCATCAGCGT GGTCGTGAAG CGATTACAG ATGTCTGCC TTTATCCGC GTCCAGCTCG TTGAGTTTCT CCAGAAGCGT TAATGTCTGG CTTCTGATAA
                5,220                ROP_ORF0                5,260                5,280                5,300
AGCGGGCCAT GTTAAGGGCG GTTTTTTCTT GTTTGGTCAC TGATGCCTCC GTGTAAGGGG GATTTCTGTT CATGGGGGTA ATGATACCGA TGAACAGAGA
                5,320                5,340                5,360                5,380                5,400
GAGGATGCTC ACGATACGGG TTAAGTATGA TGAACATGCC CGGTACTGAG AACGTTGTGA GGGTAACAA CTGGCGGTAT GGATGCGGCG GGACCAGAGA
                5,420                5,440                5,460                5,480                5,500
AAAATCACTC AGGGTCAATG CCAGCGCGAT GCATAATGTG CCTGTCAAAT GGACGAAGCA GGGATTCTGC AAACCTATG CTAAGCTCGTC AAGCCGTCAG
                5,520
TTGTCTGATT CGTTACCAAT TA3

```

alphaVmut-TMD = TMD Int_a5GP

JM = Jxtamembrane region

KanR = Aminoglycoside O-phosphotransferase

Publications

Schanzenbach C, Schmidt FC, Breckner P, Teese MG, Langosch D. Identifying ionic interactions within a membrane using BLaTM, a genetic tool to measure homo- and heterotypic transmembrane helix-helix interactions. *Sci. Rep.* **7**, 43476; (2017).

Schanzenbach C, Ilver D, Blomgren J, Jonasson C, Johansson C, Krozer A, Ye L, Rustas BO. Preparation and characterisation of a sensing system for wireless pH measurements in vivo, in a rumen of a cow. *Sensors and Actuators, B: Chemical.* 2017 Apr 1;242:637-644.

Lebenslauf

Persönliche Daten

Name Christoph Schanzenbach
Geburtsdatum/-ort 09.03.1987 in München
Wohnsitz Mellingerstrasse 37, 5400 Baden, Schweiz

Schulische Ausbildung

1993-1997 Volksschule Zorneding

1997-2006 Humboldt-Gymnasium Vaterstetten
Abschluss: Allgemeine Hochschulreife

Akademische Ausbildung

2006-2009 Studium Molekulare Biotechnologie an der TU München
Abschluss: Bachelor of Science

Bachelorarbeit am Lehrstuhl für Biologische Chemie
„Untersuchungen zur Bindung von Imatinib an das α_1 -Acid-Glykoprotein“

2009-2012 Studium Molekulare Biotechnologie an der TU München
Abschluss: Master of Science

Masterarbeit am Lehrstuhl für Chemie der Biopolymere
„Einfluss der Phosphatidylinositol-Phosphat Verteilung auf die Liposomen Fusionsrate in einem in vitro Assay“

2012-2017 Promotion am Lehrstuhl für Chemie der Biopolymere von Prof. Dr. Dieter Langosch, TU München

“BLaTM: A novel method to measure transmembrane domain interactions”

Sonstige Tätigkeiten

2011/04-2011/09 Praktikum IMEGO AB, Göteborg, Schweden
Bereich Sensorentwicklung

2007/05-2011/03 Laborhilfskraft bei Sequiserve GmbH, Vaterstetten