Identifizierung, Charakterisierung und Optimierung von neuen kovalenten Split-Isopeptid Systemen

Der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg
zur
Erlangung des Doktorgrades Dr. rer. nat.

vorgelegt von
Marlene Pröschel
aus Nürnberg

Als Dissertation genehmigt von der
Naturwissenschaftlichen Fakultät der
Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung: 19.03.2018

Vorsitzende/r des Promotionsorgans: Prof. Dr. Georg Kreimer

Gutachter/in: Prof. Dr. Uwe Sonnewald

Prof. Dr. Jutta Eichler

Teile dieser Arbeit sind in folgenden Publikationen enthalten:

Pröschel, M., Detsch, R., Boccaccini, A. R., & Sonnewald, U. (2015). Engineering of Metabolic Pathways by Artificial Enzyme Channels. *Frontiers in Bioengineering and Biotechnology*, *3*(October), 168. https://doi.org/10.3389/fbioe.2015.00168

Pröschel, M., Kraner, M. E., Horn, A. H. C., Schäfer, L., Sonnewald, U., & Sticht, H. (2017). Probing the potential of CnaB-type domains for the design of tag/catcher systems. *PLoS ONE*, *12*(6), 1–26. https://doi.org/10.1371/journal.pone.0179740

Inhaltsverzeichnis

Inhaltsverzeichnis I

Zusammenfassung 1

Summary 3

1 Einleitung 5

1.1 Proteine als modulare Bausteine des Lebens und der Synthetischen Biologie 5

1.1.1 Verschiedene Ebenen der Proteinstruktur – modularer Aufbau 5

1.1.2 Strukturmotive und Domänen 8

1.1.2.1 Proteindomänen zur Vermittlung von spezifischen Proteininteraktionen 8

1.1.2.2 Immunglobulindomänen 9

1.2 Isopeptidbindungen und deren natürliches Vorkommen 12

1.2.1 Intermolekulare Isopeptidbindungen 13

1.2.2 Intramolekulare Isopeptidbindungen 15

1.2.2.1 Entdeckung und biologische Funktion 15

1.2.2.2 Allgemeiner Reaktionsmechanismus 17

1.3 Split-Domänen Systeme basierend auf Isopeptidbindungen – Split-Isopeptid Systeme 19

1.3.1 Erste Generation von Split-Isopeptid Systemen: Isopeptag-N/ Pilin-N und Isopeptag/ Pilin-C 20

1.3.2 Zweites Split-Isopeptid System: SpyCatcher-SpyTag System (Spy-System) 21

1.3.3 SnoopCatcher-SnoopTag System 24

1.3.4 Anwendungen kovalenter Catcher-Tag Systeme 26

1.4 Zielsetzung der Arbeit 27

2 Material und Methoden 29

2.1 Chemikalien, Enzyme und Verbrauchsmaterialien 29

2.2 Antikörper 29

2.3 Vektoren und rekombinante Plasmide 30

2.4 Oligonukleotide, synthetische DNA-Fragmente und Sequenzierungen 36

2.5 Biologisches Material 37

2.5.1 Bakterien 37

2.5.1.1 Stämme 37

2.5.1.2 Kultivierung 37

2.5.1.3 Transformation von *Agrobacterium tumefaciens* 38

2.5.2 Pflanzenmaterial – *Nicotiana benthamiana* 38
2.5.2.1 Anzucht 38
2.5.2.2 Transiente Agrobakterium-vermittelte Transformation von *N. benthamiana* Blättern 38
2.6 Molekularbiologische Methoden 39
2.6.1 DNA-Standardmethoden 39
2.6.2 Klonierung von pQE-9 Expressionskonstrukten 40
2.6.3 Klonierung von pRB-35S Expressionskonstrukten 42
2.6.4 Protein-Methoden 43
2.6.4.1 Probenvorbereitung für die gelelektrophoretische Auftrennung von Proteinen 43
2.6.4.1.1 *E. coli* Proben 43
2.6.4.1.2 *N. benthamiana* Blatt-Zellaufschluss mit Homogenisator 44
2.6.4.2 Bis-Tris SDS-Polyacrylamid-Gelektrophorese (SDS-PAGE) 44
2.6.4.3 Coomassie-Färbung 45
2.6.4.4 Densitometrie zur Quantifizierung von Proteinen im Gel 46
2.6.4.5 Elektrotransfer von Proteinen auf Nitrocellulosemembran und Immundetektion (Western Blot) 46
2.6.4.6 Heterologe Expression rekombinanter Proteine in *E. coli* M15 [pREP4] Zellen 48
2.6.4.7 *E. coli* Zelllyse und Aufreinigung 6x His-markierter Proteine über Ni-NTA Affinitätschromatographie unter nativen Bedingungen 49
2.6.4.8 Dialyse, Konzentrationsbestimmung und Aufkonzentrierung von Proteinen 50
2.6.4.9 Kovalente Interaktionsanalysen von Catcher- und Tag-Proteinen ..... (*In vitro* Rekonstitutionsexperimente) 51
2.7 Mikroskopische Methoden 52
2.7.1 Konfokale Laserscanning Mikroskopie (KLSM) 52
2.7.2 Fluoreszenzmikroskopie 53
2.8 Chloroplasten-Isolierung über SpyCatcher-beladene Magnetpartikel (Chloroplasten-Trap) 53
2.8.1 Beladung von Maleimid-*beads* mit Cystein-SpyCatcher-Protein 53
2.8.2 Isolierung von Chloroplasten aus Gesamtblattfiltrat 55
2.9 Software und Webserver 56
3 Ergebnisse 57
3.1 Identifizierung neuer Split-Isopeptid Systeme 57
3.1.1 Bioinformatische Analysen (*in silico* Analysen) 57
3.1.1.1 Strukturbasierte Datenbanksuche 57
3.1.1.2 Selektion der Kandidatendomänen 57

3.1.1.2.1 3phs-System 59
3.1.1.2.2 4oq1-System 59
3.1.1.2.3 3kptN- und 3kptC-System 60
3.1.1.3 Domänenspaltung und Validierung 60
3.1.2 Catcher- und Tag-DNA-Konstrukte 62
3.1.3 Heterologe Expression der Catcher- und Tag-Konstrukte in *E. coli* M15 [pREP4] Zellen 63
3.1.4 *In vitro* Analyse der Split-Domänen Rekonstitution basierend auf Isopeptidbindungen zwischen Catcher und Tag (*in vitro* Rekonstitutionsexperimente) 65
3.2 Charakterisierung der neu identifizierten Split-Isopeptid Systeme 68
3.2.1 Verifizierung der reaktiven Triade 69
3.2.2 Effizienz der kovalenten Reaktion zwischen Catcher und Tag 72
3.2.3 Spezifität der Catcher-Tag Reaktionen 74
3.2.4 Optimale Reaktionsbedingungen der kovalenten Catcher-Tag Reaktion 76
3.2.4.1 Analyse der Wirkung verschiedener chemischer Zusätze auf die Rekonstitutionseffizienz 76
3.2.4.2 Optimales Catcher/Tag Verhältnis der Reaktion 79
3.2.4.3 Temperaturabhängigkeit der Catcher-Tag Reaktion 82
3.2.4.4 pH-Abhängigkeit der Catcher-Tag Reaktion 84
3.3 Optimierung der neuen Split-Isopeptid Systeme durch rationale, gerichtete Mutagenese 89
3.3.1 Optimierung der initialen Catcher-Tag Spaltungsstelle (verlängerte und verkürzte Catcher- und Tag-Proteinvarianten) 90
3.3.1.1 N- und C- terminal modifizierte 4oq1Tag-Varianten 91
3.3.1.2 4oq1Catcher-Varianten 99
3.3.1.3 3kptCTag-Varianten 100
3.3.2 Stabilisierung der Catcher-Proteine: gezielte Aminosäureaustausche durch ortsspezifische Mutagenese 102
3.3.2.1 Entfernung von hydrophoben Aminosäuren an der Proteinoberfläche 103
3.3.2.2 Strukturbasierte Mutationen 106
3.3.2.2.1 Zugänglichkeit der Tag-Bindestelle im Catcher-Protein 106
3.3.2.2.2 Prolin-Substitutionen 107
3.3.2.2.3 Glycin-Substitutionen 110
3.3.2.3 B-Faktor-basierte Substitutionen (Austausch besonders flexibler Aminosäuren) 111
3.3.2.4 PROSS – Protein Repair One Stop Shop 113

3.3.2.5 ,,SpyTag“-Mutationen: N→D Substitution in den Tag-Peptiden 116
3.4 Kovalente Split-Isopeptid Systeme *in planta* 118
3.4.1 Spy-System *in planta* 120
3.4.2 Snoop-System *in planta* 122
3.4.3 3kptC-System *in planta* 123
3.4.4 4oq1-System *in planta* 123
3.5 Anwendungen der kovalenten Catcher-Tag Systeme 125
3.5.1 Protein-Immobilisierung an Magnetpartikeln durch Catcher-Tag Funktionalisierung 126
3.5.2 Generierung von stabilen Multiproteinkomplexen 129
3.5.3 Isolierung von Chloroplasten über das SpyCatcher-SpyTag System 130
4 Diskussion 133
4.1 Proteine: Direkte Sequenz-Struktur-Funktionsbeziehung – Segen und Fluch für das ,,Protein Engineering“ 133
4.1.1 Proteine besitzen nur eine geringe Konformationsstabilität 133
4.1.2 Proteine sind leicht zu manipulieren 134
4.1.3 Heterologe Proteinexpression als erste kritische und aufschlussreiche Hürde des ,,Protein Engineerings“ 135
4.2 Verschiedene Strategien des ,,Protein Engineerings“ 137
4.2.1 Rationaler Ansatz – gezielte Mutagenese 137
4.2.2 Evolutionärer Ansatz – zufällige Mutagenese & *in vitro* Evolution 137
4.2.3 Vergleich der verschiedenen Ansätze des ,,Protein Engineerings“ im Hinblick auf die Verbesserung der Catcher-Tag Rekonstitutionseffizienz 138
4.3 Rationale Strategien zur Optimierung von kovalenten Catcher-Tag Systemen . 139
4.3.1 Ermittlung optimaler Reaktionsbedingungen bereits als erster Optimierungsschritt der neu identifizierten Catcher-Tag Systeme 140
4.3.2 Überprüfung der initialen Produkte der CnaB-Domänenspaltung durch N- und C-terminale Proteinvarianten 142
4.3.3 Stabilisierung der Catcher-Proteine 143
4.3.3.1 Richtiges Maß an Flexibilität und Starrheit in Proteinstrukturen 144
4.3.3.1.1 Prolin- und Glycin-Substitutionen 144
4.3.3.1.2 B-Faktor basierte Substitutionen 146
4.3.3.2 Einfluss des hydrophoben Effekts auf die Proteinstabilität 146
4.3.3.3 Vergleichendes ,,Proteindesign“ - PROSS 148
4.3.4 Unterschiedliche strukturelle Voraussetzungen 149

4.3.5 Unterschiedliche chemische Voraussetzungen verschiedener Catcher-Tag Systeme ... 153
4.3.6 Unterschiedliche intrinsische Eigenschaften der CnaB-Domänen ... 155
4.3.7 Fazit und Ausblick der Mutationsanalysen ... 157
4.4 Mögliche neue Anwendungen der kovalenten Catcher-Tag Systeme ... 159
4.4.1 Entwicklung von synthetischen, stabilen Metabolonen ... 160
4.4.2 Anwendungen bei BioLEDs ... 161
4.4.3 Pflanzenbiotechnologische Anwendungen ... 161
4.4.3.1 Organellen-Trap ... 161
4.4.3.2 Spy-Trap als kovalente Alternative zur GFP-Trap ... 162
4.4.3.3 *Metabolic engineering* in Pflanzen ... 163
4.4.4 Weiterer Optimierungsbedarf der Catcher-Tag Systeme ... 164
Literaturverzeichnis ... 167
Abkürzungsverzeichnis ... 179
Anhang ... 182

Zusammenfassung

Die überraschende Neuentdeckung intramolekularer Isopeptidbindungen innerhalb Immunglobulin-ähnlicher Domänen von adhäsiven, extrazellulären Proteinen gram-positiver Bakterien hat die Biotechnologie und Synthetische Biologie revolutioniert. Das aus der Spaltung der CnaB2-Domäne des Fibronektinbindeproteins von *Streptococcus pyogenes* resultierende ,,Superkleber"-System, bestehend aus dem Protein SpyCatcher und seinem spezifischen Peptidpartner SpyTag, ist ein häufig und vielfältig verwendetes molekulares Werkzeug für Biokonjugationen. Die über dieses SpyCatcher-SpyTag System vermittelte post-translationale Kontrolle über Proteine (Verknüpfung und Immobilisierung von Proteinen, Funktionalisierung von Materialien) profitiert dabei von dem kovalenten, irreversiblen, stabilen, spezifischen und robusten Charakter der genetisch kodierten Catcher-Tag Reaktionspartner, die unter Ausbildung einer autokatalytischen Isopeptidbindung sehr effizient miteinander reagieren. Um die Anwendungsmöglichkeiten dieser Technologie vor allem durch Programmierbarkeit und Kontrollierbarkeit zu erweitern und zu verbessern, wurden in dieser Arbeit alternative, neue Catcher-Tag Systeme, basierend auf Isopeptidbindungen, identifiziert, charakterisiert und optimiert. Über strukturbasierte Datenbanksuche innerhalb von Zelloberflächenproteinen gram-positiver Bakterien, die aus Ig-ähnlichen Domänen des CnaB-Typs aufgebaut sind, wurden vier potentielle Split-Domänen Kandidaten für genauere Analysen ausgewählt. Die CnaB-Domänen wurden mittels *in silico* Analysen in zwei Teile, Catcher-Protein und Tag-Peptid, gespalten und anschließend *in vitro* bezüglich ihrer Fähigkeit zur Split-Domänen Rekonstitution, unter Ausbildung einer intermolekularen Isopeptidbindung, analysiert. Zwei der vier CnaB-Kandidatendomänen, 4oq1 und 3kptC, konnten in dieser Arbeit mittels Rekonstitutionsexperimenten erfolgreich als neue Split-Isopeptid Systeme identifiziert werden. Das 4oq1Catcher-4oq1Tag System basiert auf der D2-Domäne des basalen Pilins RrgC von *Streptococcus pneumoniae*, während sich das 3kptCCatcher-3kptCTag System von der CNA3-Domäne des Hauptpilins BcpA von *Bacillus cereus* ableitet. An die Identifizierung schloss sich die genauere Charakterisierung der neuen Catcher-Tag Systeme an. Nach der Verifizierung der Isopeptid-Triade, der Analyse der Rekonstitutionseffizienz und Orthogonalitätstests der neuen Catcher-Tag Systeme, wurden die optimalen Reaktionsbedingungen (optimales Catcher-Tag Verhältnis, Testen bestimmter chemischer Zusätze, Temperatur- und pH-Optimum) ermittelt. Deutlich wurde, dass die neu identifizierten Catcher-Tag Systeme relativ langsam und ineffizient waren, vor allem im direkten Vergleich mit dem hocheffizienten SpyCatcher-SpyTag System. Unter optimalen Bedingungen konnte die Reaktivität der Systeme bereits deutlich verbessert werden. Um die

Effizienz der kovalenten Reaktion zwischen entsprechenden Catcher- und Tag-Reaktionspartnern weiter zu optimieren, wurden verschiedene Strategien des ,,Protein Engineerings" angewendet und die Proteinkomponenten genetisch manipuliert. Zuerst wurden die initialen Produkte der CnaB-Domänenspaltungen überprüft, indem N- und C-terminale Proteinvarianten im Hinblick auf eine verbesserte Rekonstitutionseffizienz getestet wurden. Sowohl essentielle als auch für die Reaktion verzichtbare Aminosäuren konnten dadurch identifiziert werden. Vielfältige rationale Mutagenese-Strategien, vor allem zur Stabilisierung der größeren Catcher-Proteine durch gezielte Aminosäuresubstitutionen, folgten. Einige, wenige Mutationen konnten die kovalente Reaktion zwischen Catcher und Tag verbessern, wohingegen die meisten Mutationen keine Verbesserung oder sogar eine Verschlechterung der Rekonstitutionseffizienz mit sich brachten, was die Komplexität des ,,Protein Engineerings" verdeutlicht. Nach Bestätigung der Funktionalität des SpyCatcher-SpyTag Systems und des 4oq1Catcher-4oq1Tag Systems *in planta*, wurde eine erste pflanzenbiotechnologische Anwendung (Chloroplasten-Isolierung) dieser kovalenten Catcher-Tag Systeme getestet. Mögliche andere und neue Anwendungen der Catcher-Tag Systeme wurden, basierend auf ersten Vorexperimenten (Herstellung eines weißen Trimers für BioLED-Anwendungen, SpyCatcher- oder SpyTag-funktionalisierte Magnetpartikel), diskutiert. Die in dieser Arbeit erforschten Catcher-Tag Systeme erweitern somit das Repertoire an nützlichen, vielseitig einsetzbaren molekularen Werkzeugen zur irreversiblen, post-translationalen Kontrolle über Proteine und eröffnen neue Anwendungsmöglichkeiten, auch in der Pflanzenbiotechnologie.

Summary

The discovery of unprecedented intramolecular isopeptide bonds within immunoglobulin-like domains of adhesive extracellular proteins from gram-positive bacteria has revolutionized Biotechnology and Synthetic Biology. The protein superglue system, comprising of the protein SpyCatcher and its specific peptide partner SpyTag, is based on the splitting of the CnaB2 domain of the fibronectin-binding protein from *Streptococcus pyogenes* and is a convenient and widely used molecular tool for any kind of bioconjugation. The SpyCatcher-SpyTag mediated post-translational control over proteins (protein-coupling, protein-immobilization, functionalization of materials) benefits from the covalent, irreversible, stable, specific and robust character of the genetically encoded reaction partners, Catcher and Tag, that efficiently interact to form an autocatalytic isopeptide bond. To extend and improve the functional potential of this technology, especially by making it more controllable and programmable, the focus of this work was to identify, characterize and optimize new Catcher-Tag systems based on isopeptide bonds. By structure-based data base search within cell surface proteins from gram-positive bacteria consisting of Ig-like domains of the CnaB-folding type four potential split-domain candidates were selected and subjected to further detailed investigation. Based on *in silico* analyses, the CnaB domain splitting, which results in one Catcher protein and one Tag peptide, was performed. Subsequent *in vitro* split-domain reconstitution experiments to test for the underlying isopeptide bond formation were carried out. Two out of the four CnaB domain candidates, 4oq1 and 3kptC, turned out to be functional split-isopeptide systems. The 4oq1Catcher-4oq1Tag system is based on the D2 domain of the basal pilin RrgC from *Streptococcus pneumoniae*, while the 3kptCCatcher-3kptCTag system is derived from the CNA3 domain of the major pilin BcpA from *Bacillus cereus*. After the identification, the new Catcher-Tag systems were further characterized. Verification of the isopeptide triad, analysis of the efficiency and specificity of the Catcher-Tag systems were followed by identifying the optimal reaction conditions (optimal Catcher-Tag ratio, chemical additives, optimal temperature and pH). Compared to the established, very efficient SpyCatcher-SpyTag system the newly identified Catcher-Tag systems are relatively slow and inefficient with respect to isopeptide bond formation. Under optimal conditions the reactivity of the systems was already improved significantly. To further increase the reaction efficiency between respective Catcher and Tag reaction partners different protein engineering strategies to genetically manipulate the protein components were applied. To analyse the initial products of CnaB domain splitting first, N- and C-terminally modified protein variants were tested for improved reconstitution efficiency. Essential as well as dispensable amino acids for the reaction were identified. Various rational

mutagenesis design strategies, especially to improve protein stability of the bigger Catcher proteins, were applied next. Just a few mutations were identified that could improve the covalent reaction between Catcher and Tag, whereas most of the mutations did not affect the reaction at all or made it even worse. This clearly shows the complexity of protein engineering. The SpyCatcher-SpyTag as well as the 4oq1Catcher-4oq1Tag system were proven to be functional in plants and therefore a first plant biotechnological application (isolation of chloroplasts) of these covalent Catcher-Tag systems was tested. Based on some pre-experiments (generation of a white trimer for BioLED applications, SpyCatcher- or SpyTag-functionalized magnetic beads) new future applications of these Catcher-Tag systems were discussed. The Catcher-Tag systems discovered in this thesis enlarge the toolbox of convenient molecular devices for irreversible, post-translational control over proteins and open up new opportunities, also in plant biotechnology.

1 Einleitung

So wie sich die Natur auch immer wieder modularer Grundbausteine bedient und diese im Laufe der Evolution je nach spezifischer Funktion und Anforderung anpasst und perfektioniert, nutzen Biologen im Bereich der Synthetischen Biologie ebenfalls solche modularen Systeme als Ausgangspunkte ihrer Arbeit. Durch Manipulation und Änderungen sollen so Systeme mit neuartigen und verbesserten Eigenschaften hervorgebracht werden, die vor allem im Bereich der Biotechnologie Anwendung finden sollen. Proteine stellen solche natürlichen modularen Bausteine dar.

1.1 Proteine als modulare Bausteine des Lebens und der Synthetischen Biologie

1.1.1 Verschiedene Ebenen der Proteinstruktur – modularer Aufbau

Proteine sind die wichtigsten Bestandteile der meisten biologischen Prozesse, die sie aufgrund ihrer komplexen dreidimensionalen Struktur mit höchster Effizienz, Präzision und Spezifität ausführen. Die dreidimensionale Struktur der Proteine, die mehrere Organisationsstufen beinhaltet (siehe Abb.1.2), bestimmt somit maßgeblich die Proteinfunktion (Baker und Young, 2015, Alberts, Stryer).

Proteine (=Polypeptide) sind lineare Polymere und setzen sich in ihrer Primärstruktur aus monomeren Untereinheiten, den L-α-Aminosäuren zusammen, die über Peptidbindungen miteinander verknüpft sind (Stryer, Alberts) (siehe Abb.1.1).

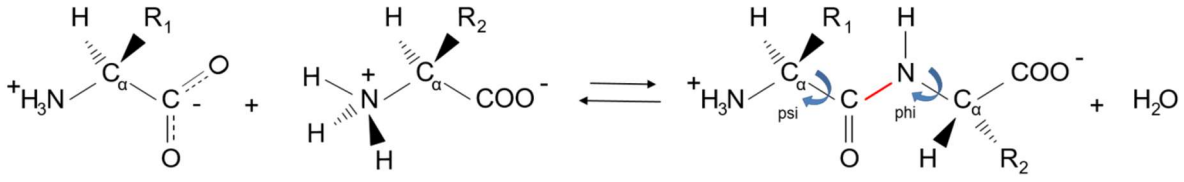


Abb.1.1: Bildung einer Peptidbindung zwischen zwei Aminosäuren (Dipeptid). Gezeigt sind zwei Aminosäuren (zentrales C_α-Atom mit vier Substituenten: Aminogruppe, Carboxylgruppe, Wasserstoffatom und Seitenkette/ Rest: R) die unter Wasserabspaltung zu einem Dipeptid reagieren. Die Peptidbindung (rot), die die beiden Aminosäuren miteinander verknüpft, entsteht durch die Reaktion der α-Aminogruppe einer Aminosäure mit der α-Carboxylgruppe einer anderen Aminosäure. Durch Addition weiterer Aminosäuren über Peptidbindungen entsteht eine polare Polypeptidkette, deren konstantes Rückgrat aus sich wiederholenden Einheiten besteht. Den variablen Anteil eines Polypeptids bilden die einzelnen Seitenketten. Die Torsionswinkel phi und psi beschreiben die Rotationsmöglichkeiten um die strukturbestimmenden Einfachbindungen einer Aminosäure.

An der Proteinhauptkette sind beliebige der 20 verschiedenen Aminosäure-Seitenketten (R) angehängt (siehe Anhang Tabelle A3), die sich aufgrund ihrer unterschiedlichen funktionellen Gruppen (z.B. Alkohole, Thiole, Thioether, Carbonsäuren, Carboxamide, alkalische Gruppen) und physikalischen und chemischen Eigenschaften auszeichnen. So unterscheiden sich die 20 verschiedenen Aminosäuren-Seitenketten in Größe, Gestalt/ Geometrie, Ladung, Polarität bzw. Fähigkeit zur Ausbildung von Wasserstoffbrücken, Hydrophobizität, sowie chemischer Reaktivität und bestimmen dadurch maßgeblich die Funktionen, die die Proteine vermitteln.

Eine besondere Eigenschaft von Proteinen besteht darin, dass sie sich spontan zu einer dreidimensionalen Struktur falten können, die allein durch die individuelle Aminosäuresequenz (**Primärstruktur**) des Proteins festgelegt wird (siehe Abb.1.2). Polypeptidketten falten sich in ihrer **Sekundärstruktur** zu regelmäßigen und wohldefinierten Strukturelementen wie α-Helix und β-Faltblatt, Kehren (*turns*) und Schleifen (*loops*). Diese lokalen Sekundärstrukturelemente basieren vor allem auf der Ausbildung stabilisierender Wasserstoffbrückenbindungen zwischen nahe beieinander liegenden Peptidgruppen der Proteinhauptketten und stellen die stabilste Konformation, bzw. die Konformation geringster potenzieller Energie, in diesem Bereich der Polypeptidkette dar (Ponnuswamy und Gromiha, 1994, Pace et al., 1996). Verschiedene Aminosäuren besitzen vor allem aufgrund ihrer physikochemischen Eigenschaften (wie z.B. der strukturbestimmenden Torsionswinkel phi und psi, siehe Abb.1.1) unterschiedliche Präferenzen für die Ausbildung von α-Helices, β-Faltblatt-Strukturen, β-Kehren und Schleifen und kommen deshalb unterschiedlich häufig in diesen Sekundärstrukturelementen vor (Otaki et al., 2010). Diese lokalen Sekundärstrukturbereiche und deren räumliche Lage relativ zueinander leisten einen wichtigen Beitrag zur endgültigen Proteinstruktur bzw. der Gesamtanordnung der Polypeptidkette eines Proteins, die als **Tertiärstruktur** bezeichnet wird. Die gesamte Raumstruktur einer Polypeptidkette ergibt sich v.a. aus den Wechselwirkungen zwischen Aminosäureresten, die innerhalb der linearen Sequenz weit voneinander entfernt sind. Aber nicht nur die zahlreichen, hauptsächlich nicht-kovalenten Wechselwirkungen (Van-der-Waals Wechselwirkungen, hydrophobe Wechselwirkungen, elektrostatische Wechselwirkungen und Wasserstoffbrücken) zwischen verschiedenen Aminosäuren bestimmen die Gesamtanordnung von Polypeptidketten, sondern auch die Wechselwirkung und Interaktion mit dem jeweiligen Milieu. So falten sich wasserlösliche Proteine in wässriger Umgebung aufgrund des hydrophoben Effekts spontan zu kompakten Strukturen mit einem unpolaren Kern. Hydrophobe Seitenketten werden hauptsächlich im Proteininneren verborgen und vor allem über Van-der-Waals Wechselwirkungen dicht gepackt und zusammengehalten, während polare, geladene Reste eher an der Proteinoberfläche liegen und

Wasserstoffbrückenbindungen mit dem wässrigem Milieu ausbilden (Pace et al., 1996, Yan et al., 2008). Somit falten sich Proteine natürlicherweise immer in die thermodynamisch stabilste Konformation mit der geringsten Energie (Kumar und Nussinov, 2001, Pace et al., 1996, Alberts). Diese dreidimensionale Struktur ist essentiell für die Funktion eines Proteins. Eine weitere Organisationsebene des strukturellen Aufbaus von Proteinen, die **Quartärstruktur**, ergibt sich durch die hauptsächlich nicht-kovalente Zusammenlagerung von mehreren unterschiedlichen Polypeptidketten (Untereinheiten) zu funktionellen Komplexen. Die Quartärstruktur beschreibt somit die räumliche Anordnung der Untereinheiten und die Art ihrer Wechselwirkungen miteinander (Lehrbücher: Stryer, Alberts, Biotechnologie).

Die modulare Natur von Proteinen kann gut am Beispiel von Pilinen, den Proteinbausteinen des Pilus von gram-positiven Bakterien verdeutlicht werden (siehe Abb.1.2) (Kang und Baker, 2012, Kang et al., 2014).

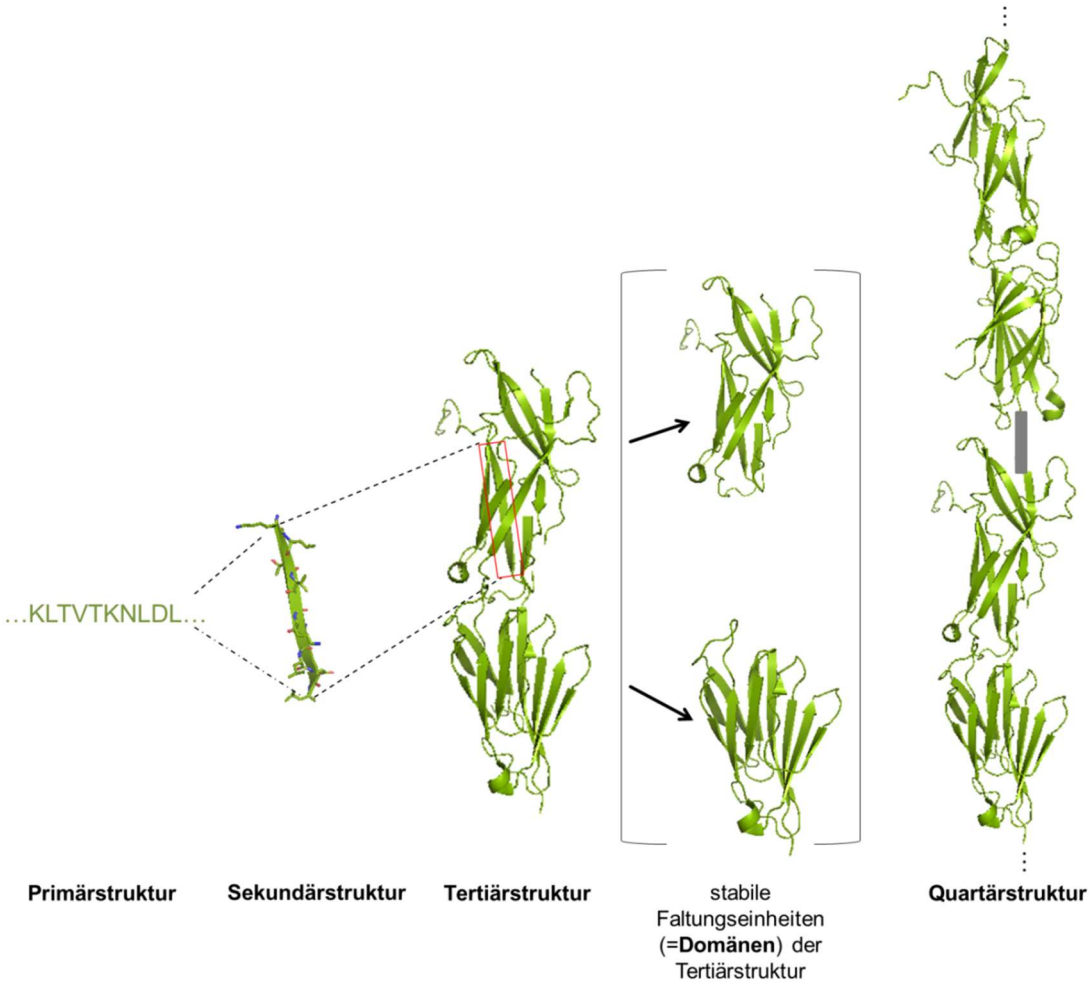


Abb.1.2: Ebenen der Proteinstruktur am Beispiel des Hauptpilins Spy0128 aus *Streptococcus pyogenes*. Proteine weisen vier Hauptebenen des strukturellen Aufbaus auf. Die Primärstruktur (Aminosäuresequenz), Sekundärstruktur (β-Strang), Tertiärstruktur (definierte Gesamtstruktur der Polypeptidkette) und die Quartärstruktur (Zusammenlagerung mehrerer Polypeptidketten). Große Proteine bestehen zudem häufig aus kompakten, strukturellen Grundeinheiten, die individuell und

unabhängig falten und spezifische Funktionen vermitteln (=Proteindomänen). Das Hauptpilin Spy0128 besteht in seiner Tertiärstruktur aus zwei reinen β-Faltblatt-Domänen (N- und C-terminale Domäne), die über einen Serinrest (Serin-172, Kang et al., 2007) verbunden sind. Die β-Faltblätter sind aus einzelnen β-Strängen, die über Wasserstoffbrücken miteinander interagieren, aufgebaut. Die Seitenketten benachbarter Aminosäuren in einem β-Strang weisen in entgegengesetzte Richtungen, also abwechselnd Richtung Proteinkern (hydrophob) oder Milieu (hydrophil). In der Quartärstruktur lagern sich mehrere monomere Spy0128 Piline zusammen (hier sind zwei Spy0128 Piline gezeigt), die kovalent über spezifische Sortase-vermittelte intermolekulare Isopeptidbindungen miteinander verknüpft sind. Die repetitiven Spy0128 Hauptpiline bilden das Rückgrat des *S. pyogenes* Pilus. Durch ebenfalls kovalente Assoziation von Basis- und Kopf-Pilinen wird der Pilus vervollständigt.

1.1.2 Strukturmotive und Domänen

Proteine sind zwar aufgrund ihrer individuellen Aminosäureabfolge enorm divers, jedoch wiederholen sich bestimmte Kombinationen von Sekundärstrukturen in zahlreichen Proteinen und zeigen häufig auch ähnliche Funktionen. Diese spezifischen Kombinationen bezeichnet man als **Strukturmotive** (Stryer). Allerdings ähneln sich nicht nur bestimmte Abfolgen von Sekundärstrukturelementen in vielen Proteinen, sondern auch deren räumliche Anordnung, da sich Proteine aus einem Grundbausatz unabhängig faltender Module aufbauen. Einige, vor allem größere Proteine, falten sich in zwei oder mehr kompakte, unabhängige Bereiche, die oft durch ein flexibles Polypeptidsegment verbunden sind. Solche autonomen, kompakten globulären Faltungseinheiten werden **Domänen** genannt und stellen die modularen Bausteine der Protein-Welt dar (Dawson und Kent, 2000, Stryer, Alberts). Oft sind diese Faltungseinheiten auch mit spezifischen Funktionen assoziiert (Dawson und Kent, 2000), deren Zusammenwirken die Gesamtfunktion eines Proteins ausmacht. Proteine können trotz unterschiedlicher Gesamttertiärstruktur gleiche Domänen besitzen (Stryer). So sind verschiedene und ganz diverse Proteine (auch von verschiedenen Spezies), die gleiche oder ähnliche Funktionen erfüllen oder ähnliche Prozesse in der Zelle vermitteln, oft aus denselben, zum Teil weit verbreiteten Strukturdomänen aufgebaut, die sich jedoch im Laufe der Evolution durch Anpassungen an ganz spezifische Anforderungen diversifiziert haben (Bork et al., 1994).

1.1.2.1 Proteindomänen zur Vermittlung von spezifischen Proteininteraktionen

In der Zelle, dem natürlichen Wirkort der meisten Proteine, sind einzelne Proteine nie isoliert, sondern immer umgeben von zahlreichen anderen Molekülen, mit denen sie wechselwirken (Long et al., 2005, Good et al., 2011, Conrado et al., 2008, Alberts). Diese physikalischen Wechselwirkungen bestimmen die biologischen Eigenschaften eines Proteins, denn die Aktivität vieler Proteine hängt allgemein von ihrer Fähigkeit ab, andere Moleküle zu erkennen und spezifisch zu binden (Smith und Kortemme, 2010, Alberts). Die Komplexität vieler Funktionen, die Proteine übernehmen und ausführen (Zakeri, 2015), erfordert daher eine genaue und regulierte Kooperation interagierender Proteine, um die Spezifität und Effizienz

molekularer Prozesse zu garantieren. Diese Kooperation und auch Koordination der richtigen Interaktionspartner wird oft über spezifische Protein-Protein oder Protein-Peptid Interaktionen gewährleistet und spielt vor allem in Signaltransduktionskaskaden der Metazoen eine essentielle Rolle (Good et al., 2011, Horn und Sticht, 2015, Cohen et al., 2008, Yan et al., 2008, Gfeller, 2012). Die meisten natürlich vorkommenden Protein-Protein Interaktionen werden durch hochspezifische, nicht-kovalente Affinitätsinteraktionen zwischen einer Proteindomäne (Adapterdomäne) und einem Peptidliganden (Peptidmotiv) vermittelt und sind deshalb schwacher und transienter Natur (Zakeri, 2015, Zakeri et al., 2012, Veggiani et al., 2014). Prominente Beispiele an natürlichen, weitverbreiteten eukaryotischen Protein-Protein Interaktionsdomänen sind die SH3 (Src-Homologie-3)-Domäne und ihr spezifischer SH3-Ligand bzw. die SH2 (Src-Homologie-2)-Domäne und ihr SH2-Ligand, die PDZ (PSD95-Discs large_ZO1)-Domäne und ihr PDZ-Ligand sowie die GBD (GTPase-Bindedomäne)-Domäne mit ihrem GBD-Ligand (Horn und Sticht, 2015, Dueber et al., 2009). Spezifische Proteininteraktionen werden aber nicht nur von Eukaryoten zur Koordination von Signaltransduktionsprozessen verwendet. Prokaryoten, wie *Clostridium thermocellum*, verwenden spezifische, hochaffine Protein-Module (Cohesin und Dockerin) zur Ausbildung von extrazellulären Multienzymkomplexen zur effizienten Degradation von Cellulose (Cellulosom: Bayer et al., 1994, Bayer et al., 2004). Viele dieser modularen Protein-Peptid Interaktionssysteme finden auch in der Synthetischen Biologie, vor allem bei der post-translationalen Ligation von Proteinen, zahlreiche Anwendung (Designer-Cellulosome und synthetische Protein-Scaffolds: Bayer et al., 2004, Stern et al., 2016, Conrado et al., 2008, Moraïs et al., 2010, Moraïs et al., 2012, Zhang, 2011, You und Zhang, 2013; Dueber et al., 2009, Moon et al., 2010, Lee et al., 2012, Chen et al., 2014). Während der transiente, nicht-kovalente Charakter dieser Protein-Protein Interaktionen für die dynamische Regulation natürlicher, zellulärer Prozesse essentiell ist, haben diese spezifischen aber schwachen, reversiblen Interaktionen große Nachteile für viele biotechnologische Anwendungen, bei denen oft harsche, nicht-physiologische Konditionen und Krafteinwirkungen herrschen (Veggiani et al., 2014, Pröschel et al., 2015, Pröschel et al., 2017).

1.1.2.2 Immunglobulindomänen

Prominente und für diese Arbeit relevante Beispiele für modulare Multidomänen-Proteine liefert die Welt der Zelloberflächenproteine von gram-positiven Bakterien. Pilus-Proteine, die Piline, oder auch andere extrazelluläre Adhäsine (MSCRAMMs: *microbial surface components recognizing adhesive matrix molecules*) von gram-positiven Bakterien sind sehr häufig aus mehreren, individuellen Domänen aufgebaut, die eine ganz bestimmte Faltung und Funktion besitzen (Vengadesan und Narayana, 2011). Domänen, die sehr häufig von gram-positiven Bakterien als Bausteine zum Aufbau ihrer extrazellulären, adhäsiven Proteine

genutzt werden, sind Abwandlungen der konstanten Domäne des humanen Immunglobulin G (IgG-C), sogenannte Ig-ähnliche Domänen (siehe Abb.1.3). Die Immunglobulinfaltung ist eine der wichtigsten und häufigsten Strukturdomänen in der Biologie, die essentielle Funktionen in der Immunabwehr von Wirbeltieren (modulare Grundbausteine der Antikörper/ Immunglobuline, Zelloberflächenrezeptoren), bei der Zelladhäsion und vielen anderen Prozessen besitzen (Fraser et al., 2006, Kang und Baker, 2011, Mirny und Shakhnovich, 1999, Bodelón et al., 2013). Ig-ähnliche Domänen, die trotz zum Teil extrem unterschiedlicher Sequenzen dieselbe, konservierte Grundstruktur wie Immunglobulindomänen aufweisen, wurden durch Strukturanalysen in zahlreichen, funktionell diversen Proteinen innerhalb aller Lebensformen (Prokaryoten und Eukaryoten), sowie bei Viren und Phagen gefunden (Bork et al., 1994, Kang und Baker, 2011, Fraser et al., 2007, Fraser et al., 2006). Trotz ihrer Diversität und Modifikationen scheint allen Ig-ähnlichen Domänen ihre Rolle bei der Vermittlung von Protein-Protein Interaktionen durch spezifische Erkennung und Bindung von Molekülen gemeinsam zu sein (Bork et al., 1994, Shaik et al., 2014, Mirny und Shakhnovich, 1999). Zwei Faltungsvarianten dieser IgG-C Grundbaustein-Domäne kommen besonders häufig in extrazellulären Adhäsinen (Piline und MSCRAMMs) von gram-positiven Bakterien vor (Vengadesan und Narayana, 2011, Krishnan, 2015) (siehe Abb.1.3). Je nach topologischer Anordnung der einzelnen β-Stränge werden zwei Faltungstypen, CnaA (DEv-IgG) und CnaB (reverse/ inverse IgG Faltung oder IgG-rev), unterschieden, die näher beschrieben werden sollen, da vor allem die CnaB-Domänen Ausgangspunkt und Grundlage sowohl der etablierten als auch der in dieser Arbeit entwickelten, kovalenten Split-Isopeptid Systeme sind. Cna steht für *collagen adhesin*, das Kollagenbindeprotein aus *Staphylococcus aureus*, in dessen Domänen A (Kollagen-Binderegion) und B, diese beiden Faltungstypen zuerst beschrieben wurden (Deivanayagam et al., 2000). Beide Faltungstypen leiten sich, wie schon erwähnt, von der konstanten IgG Domäne (IgG-C) ab, die aus einem antiparallelen β-Faltblatt aus 4 β-Strängen (ABED) auf einer Seite und einem antiparallelen β-Faltblatt aus drei Strängen (CFG) auf der anderen Seite der β-*barrel* Struktur besteht (β-Sandwich Struktur, Bork et al., 1994, Deivanayagam et al., 2000) (siehe Abb.1.3). Die zwei β-Faltblätter umgeben einen hydrophoben Kern und sind durch hydrophobe Wechselwirkungen sowie eine einzelne Disulfidbrücke (zwischen Strang B und F), die die Struktur stabilisiert, verbunden (Bork et al., 1994, Halaby et al., 1999, Stryer). Ähnlich wie die IgG-C Domäne besitzt die CnaA-Typ Domäne (DEv-IgG) zwei β-Faltblattstrukturen (siehe Abb.1.3). Die β-Stränge ABED bilden das erste β-Faltblatt aus, die Stränge CFG das zweite. Jedoch besitzt der DEv-IgG Faltungstyp oft zwei weitere β-Stränge zwischen den Strängen E und D (siehe Abb.1.3). Der CnaB-Faltungstyp der IgG-C Domäne besteht aus sieben antiparallelen β-Strängen, die sich zu zwei β-Faltblättern (β-Sandwich Struktur) arrangieren und zwar genau umgekehrt wie die β-Stränge in der IgG-C Domäne

(inverse IgG Faltung) (Krishnan et al., 2007, Deivanayagam et al., 2000, Vengadesan und Narayana, 2011) (siehe Abb.1.3). Entweder gruppieren sich die sieben β-Stränge zu drei- (DAG) und vier-strängigen (FEBC) antiparallelen β-Faltblattstrukturen oder zu vier (DAGF)- und drei (EBC) –strängigen antiparallelen β-Faltblattstrukturen (Vengadesan und Narayana, 2011, Deivanayagam et al., 2000). Aufgrund der speziellen Struktur und Topologie dieser Domänen werden bestimmte konservierte Sequenzmotive zwischen verschiedenen Proteinen oder strategisch positionierte, konservierte Aminosäuren des N- und des C-Terminus innerhalb einer Proteindomäne in räumliche Nähe zueinander gebracht, was spezifische inter- und intramolekulare Reaktionen und Interaktionen ermöglicht. Innerhalb der CnaA-Faltungsdomänen kann eine intramolekulare Isopeptidbindung zwischen den antiparallel angeordneten β-Strängen A (enthält Lysin) und F (enthält Asparagin) erfolgen (siehe Abb.1.3 schwarzer Balken). Ein konservierter Asparaginsäurerest (D-Typ) auf Strang C katalysiert die Isopeptid-Verknüpfung der beiden β-Faltblattstrukturen. Eine mögliche Isopeptidbindung innerhalb der CnaB-Domäne kann zwischen zwei benachbarten, parallel angeordneten β-Strängen desselben β-Faltblatts gebildet werden (Vengadesan und Narayana, 2011). Die intramolekulare Isopeptidbindung findet dabei immer zwischen dem ersten (Strang A mit konserviertem Lysinrest) und dem letzten β-Strang (G-Strang mit konserviertem Asparagin) innerhalb desselben β-Faltblatts statt und wird durch einen konservierten Glutaminsäurerest (E-Typ) auf dem E-Strang im zweiten β-Faltblatt katalysiert (Vengadesan und Narayana, 2011) (siehe Abb.1.3). Durch diese intramolekulare, kovalente Verknüpfung werden der N- und der C-Terminus dieser Proteindomäne stabil miteinander quervernetzt (Kang und Baker, 2009). So nutzen gram-positive Bakterien spezielle Varianten dieser einfachen humanen IgG-C Domäne, um inter- und intramolekulare Isopeptidbindungen zu beherbergen, die den polymeren Adhäsinen besondere Stabilität und Resistenz gegenüber proteolytischem Abbau verleihen und die kovalente kettenartige Polymerisation von einzelnen Pilin-Untereinheiten erlaubt (Vengadesan und Narayana, 2011).

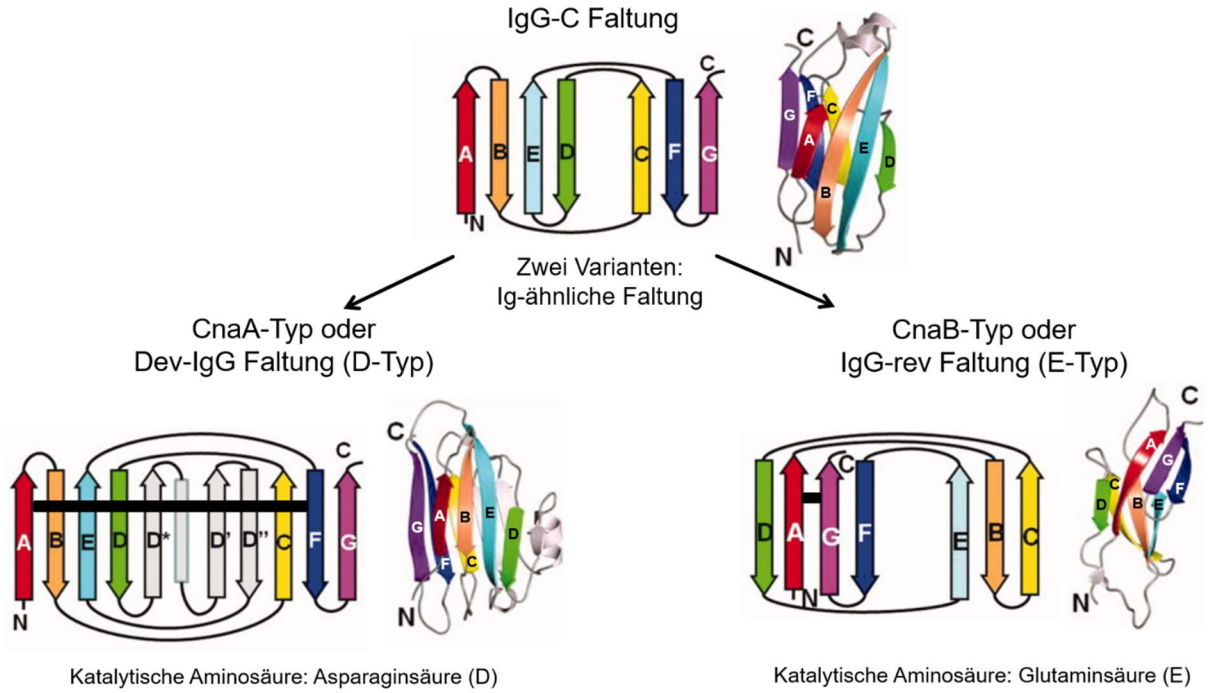


Abb.1.3: Schematische Darstellung der konstanten Domäne des Immunglobulin G (IgG-C) und der zwei IgG-ähnlichen Domänen des CnaA- und CnaB-Typs (modifiziert nach Vengadesan und Narayana, 2011). Gezeigt ist die Standard-Immunglobulinfaltung der konstanten Domäne des Immunglobulin G (IgG-C: zwei β-Faltblätter aus antiparallelen β-Strängen, die einen hydrophoben Proteinkern umgeben und durch eine Disulfidbrücke stabilisiert werden). Zwei topologische Varianten dieser Immunglobulinfaltung (Ig-ähnliche Domänen), der CnaA-Typ und der CnaB-Typ, kommen häufig in adhäsiven Zelloberflächenproteinen (Piline, MSCRAMMs) von gram-positiven Bakterien vor. Mögliche intramolekulare Isopeptidbindungen zwischen verschiedenen β-Strängen innerhalb dieser beiden Domänen sind mit schwarzen Balken gekennzeichnet. In der CnaB-Domäne werden der erste und der letzte β-Strang kovalent miteinander verknüpft, während in der CnaA-Domäne der erste und der vorletzte β-Strang quervernetzt werden.

1.2 Isopeptidbindungen und deren natürliches Vorkommen

Isopeptidbindungen sind Amidbindungen, die dieselbe Struktur wie Peptidbindungen (siehe Abb.1.1) besitzen. Im Gegensatz zur Peptidbindung, bei der die kovalente Bindung immer zwischen den reaktiven α-Gruppen (Carboxyl- und Aminogruppe der Proteinhauptkette) der Aminosäuren stattfindet, ist bei der Isopeptidbindung jedoch mindestens eine der reaktiven Gruppen (Carboxyl-/ Carboxamid- und/oder Aminogruppe) nicht Teil der Proteinhauptkette. So stellen Isopeptidbindungen kovalente Verknüpfungen zwischen zwei Seitenketten oder zwischen Seitenkette und Hauptkette dar (Veggiani et al., 2014, Kang und Baker, 2011). Isopeptidbindungen treten in der Natur sowohl inter-, d.h. zwischen verschiedenen Molekülen, als auch intramolekular, d.h. innerhalb eines Moleküls, auf.

1.2.1 Intermolekulare Isopeptidbindungen

Es gibt zahlreiche, in der Natur vorkommende biologische Prozesse, bei denen intermolekulare Isopeptidbindungen eine wesentliche Rolle spielen (Kang et al., 2007, Kang und Baker, 2009, Kang und Baker, 2011). Durch die kovalente Reaktion einer ε-Aminogruppe eines Lysinrestes eines Proteins mit einer α-Carboxylgruppe der Hauptkette eines anderen Proteins werden essentielle biologische Prozesse wie zum Beispiel die post-translationalen Modifikationen zur spezifischen Proteinmarkierung wie Ubiquitinylierung und Sumoylierung vermittelt (Kang und Baker, 2011, Kang et al., 2007). Die Energie zur Bildung einer Isopeptidbindung (endergone Reaktion) stammt dabei aus der Hydrolyse von ATP. Bei der Transglutaminierung wird eine kovalente Verknüpfung von Proteinen durch die Reaktion einer ε-Aminogruppe eines Lysinrestes mit einer Carboxamidgruppe einer Glutamin-Seitenkette geschlossen (Kang und Baker, 2011). Eine essentielle Rolle spielen intermolekulare Isopeptidbindungen auch bei der Verankerung von adhäsiven Oberflächenproteinen (Pili oder andere Adhäsine/ MSCRAMMs) an der Zellwand von gram-positiven Bakterien, sowie bei der End-zu-End Assemblierung einzelner Pilus-Untereinheiten (Piline) zu einem vollständigen, gänzlich kovalenten Pilus-Polymer (Kang und Baker, 2011, Hendrickx et al., 2011, Kang und Baker, 2009, Krishnan, 2015) (siehe Abb.1.4). Beide Prozesse, die essentiell zur Virulenz und Pathogenität gram-positiver Bakterien beitragen, werden von spezifischen Cystein-Transpeptidasen (TLXT**C** Motiv im aktiven, katalytischen Zentrum), den Sortasen vermittelt, die konserviert und ausschließlich bei gram-positiven Bakterien zu finden sind (Hendrickx et al., 2011, Kang et al., 2007, Vengadesan und Narayana, 2011). Sortasen erkennen und spalten Amidbindungen innerhalb spezifischer Sequenzmotive (*sorting signal*: LPX**TG**-ähnliche Motive) von sekretierten Zielproteinen und knüpfen neue Isopeptidbindungen, entweder zwischen verschiedenen Pilus-Proteinen (Pilin-Motiv YP**K**N) oder zwischen Pilinen oder anderen MSCRAMMs und der bakteriellen Zellwand (Peptidoglycanschicht).

Alle diese Mechanismen zur Bildung von intermolekularen Isopeptidbindungen haben gemeinsam, dass sie Enzym-vermittelt sind (Enzyme mit reaktivem Cysteinrest im aktiven Zentrum) und über ein transientes Thioester-Intermediat (Thioacyl-Enzym Intermediat) ablaufen, welches über einen nukleophilen Angriff einer ε-Aminogruppe eines Lysinrestes aufgelöst wird, sodass es schließlich zur Ausbildung einer Isopeptidbindung kommt (Kang und Baker, 2011).

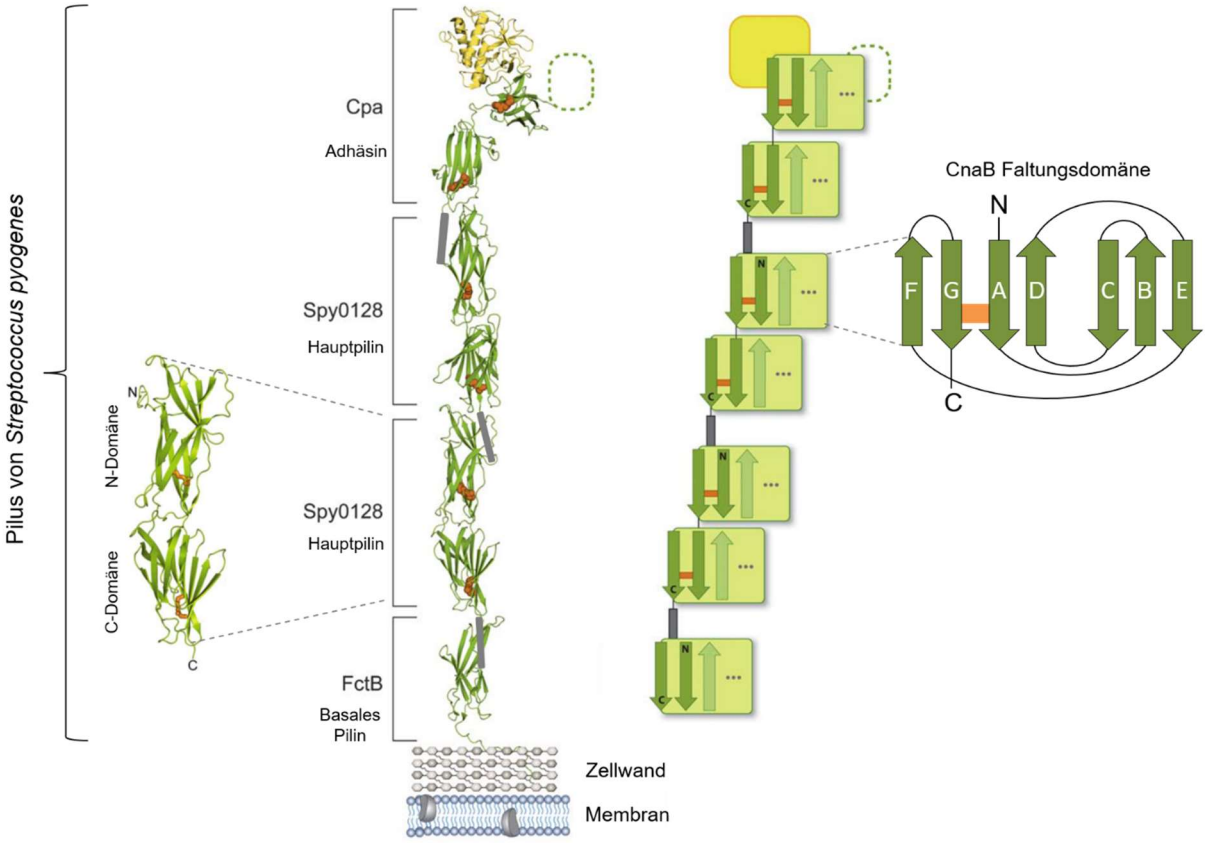


Abb.1.4: Schematische Darstellung des Pilus von *Streptococcus pyogenes* (modifiziert nach Kang und Baker, 2011). Das Modell zeigt das Pilus-Polymer aufgebaut aus einzelnen Pilinen (Pilus-Komponenten: basales Pilin FctB, das den Pilus an der bakteriellen Zellwand verankert, repetitive Einheiten des Grundgerüst-Pilins Spy0128, das selbst aus 2 Domänen besteht, Cpa-Adhäsin an der Spitze des Pilus), die über Sortase-vermittelte intermolekulare Isopeptidbindungen (graue Balken) kovalent miteinander verknüpft sind (*head-to-tail manner*, Kang et al., 2007). Alle diese Piline bestehen aus Domänen, die Ig-ähnliche Faltungen des CnaB-Typs annehmen (siehe Vergrößerung). Innerhalb dieser CnaB-Domänen können sich intramolekulare Isopeptidbindungen (oranger Balken) zwischen dem ersten und dem letzten β-Faltblattstrang innerhalb eines β-Faltblatts ausbilden, was die Proteinstabilität drastisch erhöht.

Eine autokatalytisch gebildete, das heißt nicht Enzym-vermittelte, intermolekulare Isopeptidbindung wurde dagegen in viralen Kapsidproteinen des Bakteriophagen HK97 gefunden, wo sie essentiell zur Stabilisierung der Kapsid-Struktur beiträgt (Kang und Baker, 2011). Diese Isopeptidbindung findet spontan statt, wenn sich die Kapsid-Untereinheiten rearrangieren, sodass die drei an der Ausbildung der Isopeptidbindung beteiligten Aminosäuren (Triade: Asparagin, Lysin, Glutaminsäure), in räumliche Nähe und in die richtige, bindungskompatible Position gebracht werden (Kang et al., 2007, Kang und Baker, 2011).

1.2.2 Intramolekulare Isopeptidbindungen

1.2.2.1 Entdeckung und biologische Funktion

Bis vor ein paar Jahren herrschte die Meinung, dass Disulfidbrücken die einzigen kovalenten Bindungen zwischen Aminosäure-Seitenketten darstellen (Baker und Young, 2015, Kang und Baker, 2011, Kang und Baker, 2009, Kwon et al., 2014). Doch der Aufbau des Pilus von *Streptococcus pyogenes* (siehe Abb.1.4) macht deutlich, dass neben den Sortase-vermittelten intermolekularen Isopeptidbindungen, die die einzelnen Piline kovalent miteinander und mit der Zellwand verknüpfen, auch intramolekulare Isopeptidbindungen in der Natur auftreten.

Im Jahre 2007 wurde erstmals eine neue kovalente, intramolekulare Verknüpfung zwischen Aminosäure-Seitenketten innerhalb von Oberflächenproteinen gram-positiver Bakterien entdeckt (Kang et al., 2007). Kang et al. identifizierten die erste intramolekulare Isopeptidbindung in der Kristallstruktur des Hauptpilins Spy0128 des Humanpathogens *Streptococcus pyogenes* (Kang et al., 2007). Das Hauptpilin Spy0128 besteht aus zwei reinen β-Faltblatt-Domänen, die Ig-ähnliche Faltungsdomänen des CnaB-Typs darstellen (siehe Abb.1.5 und Kapitel 1.1.2.2). Innerhalb dieser CnaB-Domänen kommt es jeweils zur Ausbildung einer Isopeptidbindung zwischen der reaktiven ε-Aminogruppe der Seitenkette von Lysin (K) und der reaktiven δ-Carboxamidgruppe der Seitenkette von Asparagin (N) (K36-N168 in der N-terminalen Domäne und K179-N303 in der C-terminalen Domäne, siehe Abb.1.5). Benachbarte Glutaminsäurereste (E117 bzw. E258) katalysieren diese Reaktion und hydrophobe Aminosäuren umgeben die drei reaktiven Aminosäuren (Kang et al., 2007) (siehe Abb.1.5).

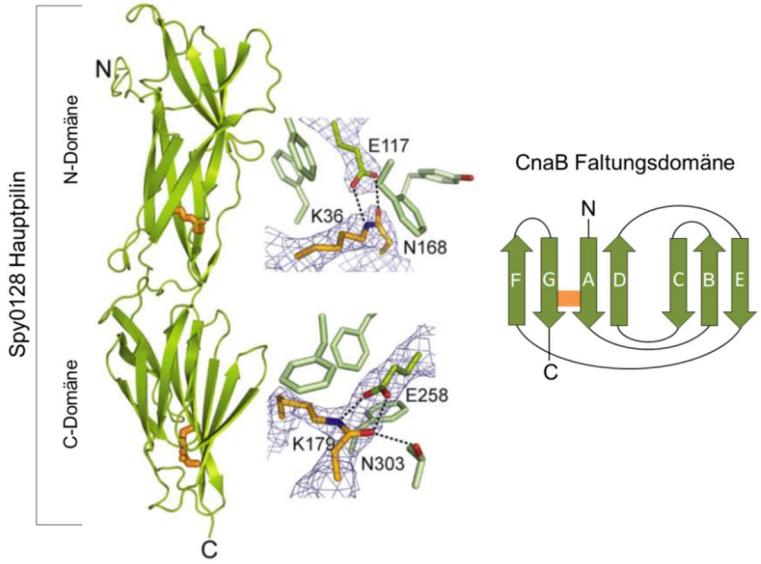


Abb.1.5: Schematische Darstellung des Hauptpilins Spy0128 aus *Streptococcus pyogenes* (modifiziert nach Kang und Baker, 2011). Das Spy0128 Protein ist aus zwei Ig-ähnliche Domänen (N-

und C-terminale Domäne) des CnaB-Faltungstyps (inverse IgG-C Faltung) aufgebaut. Innerhalb des hydrophoben Proteinkerns dieser CnaB-Domänen kommt es zur Ausbildung je einer spontanen, intramolekularen Isopeptidbindung zwischen Seitenketten von Lysin (K) und Asparagin (N). Katalysiert wird die kovalente Reaktion durch eine benachbarte, gegenüberliegende Glutaminsäure (E), die stabilisierende Wasserstoffbrücken (gestrichelte Linien) mit den Isopeptid-Gruppen ausbildet. Hydrophobe, aromatische Aminosäuren schirmen die Isopeptidbindung zusätzlich ab und schaffen eine hydrophobe Mikroumgebung.

Diese intramolekularen Isopeptidbindungen wurden kristallographisch und massenspektrometrisch verifiziert (Kang et al., 2007, Kang und Baker, 2011). *In silico* Analysen bakterieller Genome (pdb-Datenbanksuche mit reaktiver Lysin-Asparagin-Glutaminsäure Triade des Spy0128 Pilins als strukturelle Matrize) und weitere kristallographische Strukturaufklärungen bakterieller Oberflächenproteine ergaben, dass intramolekulare Isopeptidbindungen bei Pilus-Proteinen (Piline) und auch anderen extrazellulären, adhäsiven Zelloberflächenproteinen (MSCRAMMs) von gram-positiven Bakterien sehr weit verbreitet sind (Kang et al., 2007, Kang und Baker, 2011). Alle bis heute identifizierten intramolekularen Isopeptidbindungen wurden ausschließlich in Immunglobulin-ähnlichen Proteindomänen des CnaA- oder des CnaB-Faltungstyps gefunden (Kang und Baker, 2011, Hagan et al., 2010, siehe Kapitel 1.1.2.2). Da diese Ig-ähnlichen Faltungsdomänen des CnaA- und CnaB-Typs in tausenden von Zelloberflächenproteinen gram-positiver Bakterien vorkommen, wird vermutet, dass dort auch ähnlich viele intramolekulare Isopeptidbindungen existieren (Kang und Baker, 2011, Schwarz-Linek und Banfield, 2014). Interessanterweise ist das Muster der Bildung der intramolekularen Isopeptidbindungen innerhalb von Pilinen und vielen MSCRAMMs strukturell und räumlich konserviert, obwohl sie nur eine geringe Sequenzhomologie (<18%) und eine große Diversität bezüglich der Größe und Domänenorganisation aufweisen (Vengadesan und Narayana, 2011, Kang und Baker, 2011, Kang und Baker, 2012). Diese intramolekularen Quervernetzungen besitzen eine einzigartige, auch evolutionär konservierte, biologische Funktion. Sie vermitteln den Proteinen, innerhalb der sie gebildet werden, durch die Stabilisierung der Tertiärstruktur eine enorme mechanische, thermische und pH-Stabilität, sowie oft auch eine erhöhte Resistenz gegenüber proteolytischem Abbau (Kang und Baker, 2011, Veggiani et al., 2014, Zakeri et al., 2012, Zakeri, 2015, Kang und Baker, 2009). Analog den intramolekularen Disulfidbrücken in Ig-ähnlichen Domänen bei gram-negativen Bakterien, nutzen gram-positive Bakterien diese intramolekularen Isopeptidbindungen somit, um ihre extrazellulären, harschen Konditionen ausgesetzten Virulenzfaktoren (Pili und Adhäsine/ MSCRAMMs) zu stabilisieren und vor proteolytischem Abbau zu schützen (Kang et al., 2007, Kang und Baker, 2011, Schwarz-Linek und Banfield, 2014, Veggiani et al., 2014).

1.2.2.2 Allgemeiner Reaktionsmechanismus

Die Kombination aus kristallographischen-, massenspektrometrischen-, NMR-Analysen und theoretischen Berechnungen (Kang und Baker, 2011, Veggiani et al., 2014, Reddington und Howarth, 2015, Hagan et al., 2010, Hu et al., 2011) konnte zur Aufklärung des Mechanismus der Bildung der intramolekularen Isopeptidbindungen innerhalb dieser Ig-ähnlichen Faltungsdomänen beitragen (siehe Schema in Abb.1.6). Der aus den verschiedenen Analysen abgeleitete Reaktionsmechanismus scheint, wie auch die strukturellen Grundvoraussetzungen, die zur Ausbildung einer intramolekularen Isopeptidbindung führen, konserviert zu sein (Kang und Baker, 2011, Kang et al., 2007).

Bei der intramolekularen Isopeptidbindung handelt es sich um eine autokatalytische, spontane, *single-turnover* Reaktion, bei der weder Enzyme noch Cofaktoren nötig sind. Die Reaktion findet zwischen reaktiven Gruppen der Seitenketten von Lysin und Asparaginsäure (Amino- und Carboxylgruppe: Abspaltung von Wasser) oder häufiger zwischen reaktiven Gruppen von Lysin- und Asparagin-Seitenketten (Amino- und Carboxamidgruppe: Abspaltung von Ammoniak) statt (Kang et al., 2007, Kang und Baker, 2011, Hu et al., 2011, Schwarz-Linek und Banfield, 2014) (siehe Abb.1.6). Die kovalente Reaktion wird stets durch eine benachbarte Glutaminsäure (E-Typ) oder Asparaginsäure (D-Typ) katalysiert, die Protonen aufnehmen und abgeben kann (Protonenrelais-System) und dadurch den Übergangszustand über die Ausbildung von Wasserstoffbrückenbindungen stabilisiert und die Reaktion ermöglicht (Veggiani et al., 2014, Hu et al., 2011). Die drei essentiellen, direkt an der Isopeptidbindung beteiligten Aminosäuren (Lysin, Asparagin/ Asparaginsäure, Glutaminsäure/ Asparaginsäure) nennt man auch reaktive Triade oder Isopeptid-Triade (Veggiani et al., 2014). Initial kommt es zu einem direkten nukleophilen Angriff der unprotonierten ε-Aminogruppe eines Lysinrestes an das γ-Carbonyl-Kohlenstoffatom einer Asparagin- oder Asparaginsäure-Seitenkette. Dadurch, dass die katalytische Aminosäure protoniert vorliegt, wird die γ-Carbonylgruppe der Asparagin- oder Asparaginsäure-Seitenkette polarisiert, sodass das γ-C-Atom eine partielle positive Ladung erhält und leichter nukleophil durch das freie Elektronenpaar der deprotonierten ε-Aminogruppe des Lysinrestes angegriffen werden kann. Zuerst bildet sich ein Intermediat, das durch Protonen-*shuttle* aufgelöst wird und schließlich zur Ausbildung einer kovalenten Isopeptidbindung führt, indem die katalytische Aminosäure Protonen von der ε-Aminogruppe des Lysinrestes auf die Abgangsgruppe (-NH_2 oder –OH) überträgt (Kang und Baker, 2011, Hu et al., 2011). Essentiell für die Ausbildung dieser intramolekularen Isopeptidbindungen ist außerdem eine hydrophobe Proteinumgebung, die die Aktivierungsenergie zur Bildung einer Amidbindung senkt und damit das Gleichgewicht der Reaktion auf die Seite der Isopeptidbildung verschiebt (Zakeri et al., 2012). Aus diesem Grund liegt die reaktive Triade innerhalb des

hydrophoben Proteinkerns der Ig-ähnlichen Domänen verborgen und ist oft zusätzlich von konservierten hydrophoben, vor allem aromatischen Aminosäure-Seitenketten umgeben (Kang et al., 2007, Kang und Baker, 2011, Hu et al., 2011) (siehe Abb.1.5). Diese hydrophobe Mikroumgebung beeinflusst den pK_S-Wert ionisierbarer Gruppen, sodass der neutrale, reaktivere Zustand der Aminosäure-Seitenketten (*inverse protonated state*, Hu et al., 2011) favorisiert wird (Kang et al., 2007, Veggiani et al., 2014, Kang und Baker, 2011, Pace et al., 2009). So wird der pK_S-Wert von Glutaminsäure erhöht (Protonierung der Carboxylgruppe) und der pK_S-Wert der ε-Aminogruppe des Lysinrestes erniedrigt (Deprotonierung der Aminogruppe), was den nukleophilen Charakter der Aminogruppe verstärkt (Kang et al., 2007, Kang und Baker, 2011, Veggiani et al., 2014, Hu et al., 2011). Spontane, autokatalytische intramolekulare Isopeptidbindungen werden also nur gebildet, wenn die beteiligten Aminosäuren (konservierte reaktive Triade) durch Proteinfaltung in räumliche Nähe (*proximity-induced*) und in geeignete Position kommen und in einer hydrophoben Umgebung sind (Kang und Baker, 2011). Demnach ist die Anwesenheit und räumliche Nähe der drei reaktiven Aminosäuren in der richtigen Position/ Orientierung und Umgebung notwendig und ausreichend für die Ausbildung einer autokatalytischen intramolekularen Isopeptidbindung (Schwarz-Linek und Banfield, 2014).

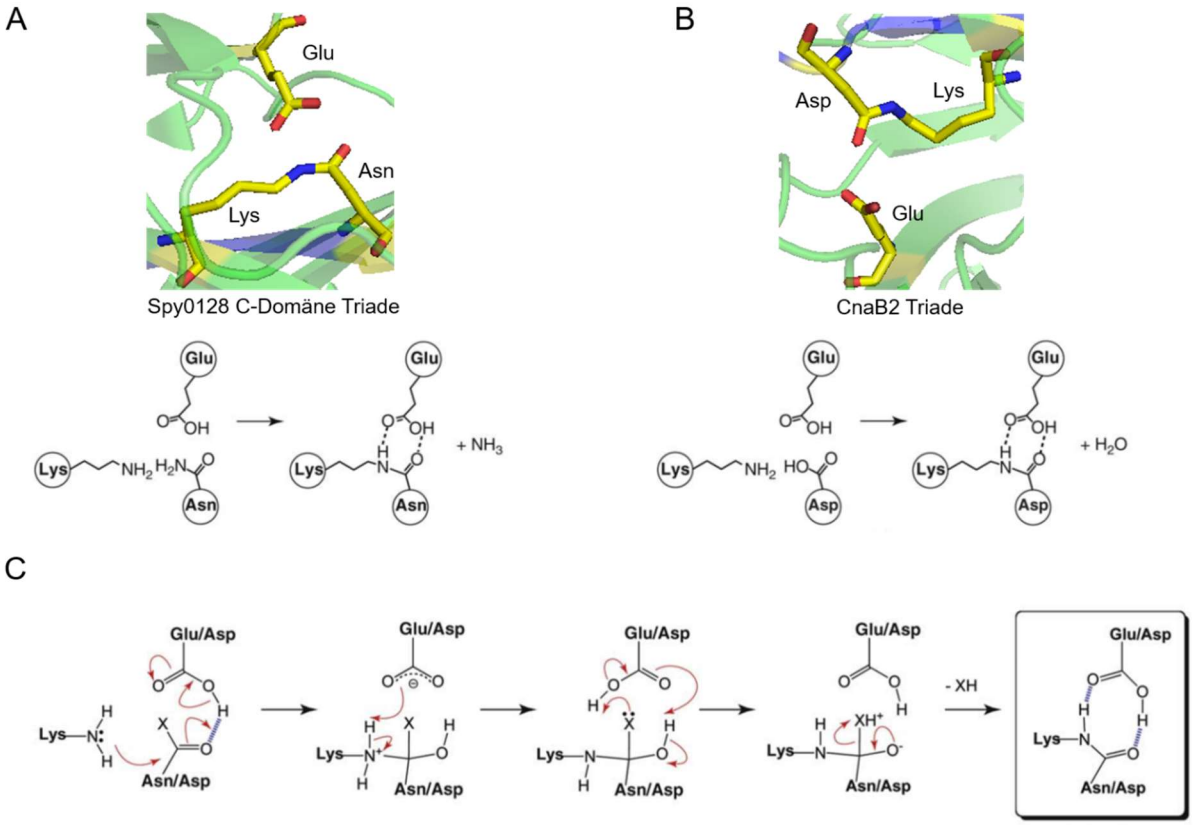


Abb.1.6: Angenommener Reaktionsmechanismus der Bildung von intramolekularen Isopeptidbindungen (modifiziert nach Kang und Baker, 2011). Der Reaktionsmechanismus zur Ausbildung von spontanen, autokatalytischen Isopeptidbindungen wurde mittels kristallographischer- (Kang et al., 2007), massenspektrometrischer- (Kang und Baker, 2011), NMR (*nuclear magnetic*

resonance) -Analysen (Hagan et al., 2010) und QM/MM-Simulationen (*quantum mechanics/molecular mechanics*) (Hu et al., 2011, Hagan et al., 2010) analysiert. Schematische Darstellung der reaktiven Triade und der Ausbildung von intramolekularen Isopeptidbindungen zwischen einer Lysin- und Asparagin-Seitenkette unter Ammoniakabspaltung (A) oder zwischen einer Lysin- und Asparaginsäure-Seitenkette unter Wasserabspaltung (B). Abgeleiteter Reaktionsmechanismus der Ausbildung von intramolekularen Isopeptidbindungen (C). Es kommt zu einer kovalenten Reaktion zwischen der Aminogruppe der Lysin-Seitenkette und der Carboxamid- oder Carboxlgruppe der Asparagin- bzw. Asparaginsäure-Seitenkette, wobei die komplette Reaktion durch eine benachbarte Glutaminsäure/ Asparaginsäure über Protonen-*shuttle* katalysiert wird. X steht für NH_2 bei der Lysin-Asparagin Bindung, was in der Abspaltung von Ammoniak resultiert bzw. für OH bei der Reaktion zwischen Lysin und Asparaginsäure, wobei Wasser abgespalten wird. Sowohl die Übergangszustände als auch das finale Isopeptid-Produkt werden über Wasserstoffbrücken, die die Carboxylgruppe der katalytischen Aminosäure ausbildet, stabilisiert.

1.3 Split-Domänen Systeme basierend auf Isopeptidbindungen – Split-Isopeptid Systeme

Aufgrund der kovalenten Natur sind diese neu entdeckten intramolekularen Isopeptidbindungen innerhalb von Proteindomänen des CnaA- oder CnaB-Faltungstyps sehr attraktive Module und Werkzeuge für die Vermittlung von stabilen, irreversiblen Proteininteraktionen bei biotechnologischen Anwendungen und stellen eine echte Alternative zu den vom Redox-Potential abhängigen Disulfidbrücken und den nicht-kovalenten Proteindomänen-Peptid Interaktionen dar (Zakeri, 2015, Zakeri and Howarth, 2010, Kang und Baker, 2011, Kwon et al., 2017). Wie bereits erwähnt, sind Proteindomänen unabhängige Faltungseinheiten, die auch in Isolation ihre Faltung und damit Funktion erhalten und somit als eigenständige, stabile Module verwendet werden können (Rossi et al., 2012). Zahlreiche Split-Protein Reportersysteme zum Nachweis von Protein-Protein Interaktionen, wie das Split-GFP und das Split-Luciferase System, existieren bereits und finden häufig Anwendung (Shekhawat und Gosh, 2011, Ozawa et al., 2006). Bei diesen Split-Protein Systemen wurden die Reporterproteine in zwei Domänen geteilt, die über nicht-kovalente Interaktionen bei räumlicher Nähe rekonstituieren können und ihre Funktionalität erhalten (Shekhawat und Gosh, 2011, Ozawa et al., 2006, Veggiani et al., 2014, Zakeri und Howarth, 2010). Doch selbst Proteindomänen stellen nicht zwangsläufig die kleinsten funktionellen Einheiten eines Proteins dar, die in der Biotechnologie genutzt werden können (Veggiani et al., 2014). In Anlehnung an die Split-Protein Systeme sollte getestet werden, ob die kompakten Proteindomänen, innerhalb der sich die intramolekularen Isopeptidbindungen ausbilden, ebenfalls in zwei Teile gespalten werden können ohne die Fähigkeit der Bildung einer Isopeptidbindung zu verlieren. Im Jahr 2010 machten sich Howarth und Zakeri erstmals die Entdeckung der intramolekularen Isopeptidbindungen innerhalb von Ig-ähnlichen Domänen zu Nutze und entwickelten das erste kovalente Split-Domänen System, das nach rationaler Spaltung der Proteindomänen in zwei Teile (Protein- und Peptidfragment)

weiterhin die Fähigkeit besitzt, Isopeptidbindungen, nun intermolekularer Natur, zwischen den gespaltenen Teilen zu bilden (kovalente Split-Proteindomänen Rekonstitution, Veggiani et al., 2014).

1.3.1 Erste Generation von Split-Isopeptid Systemen: Isopeptag-N/ Pilin-N und Isopeptag/ Pilin-C

Die erste Generation an Protein-Peptid Split-Systemen, basierend auf Isopeptidbindungen, leitet sich von dem Hauptpilin Spy0128 des gram-positiven Bakteriums *Streptococcus pyogenes* ab (siehe Abb.1.5 und Abb.1.7). Das Hauptpilin Spy0128 besteht aus zwei Ig-ähnlichen Domänen des CnaB-Typs (N-terminale und C-terminale Domäne), innerhalb der es zur Ausbildung je einer intramolekularen Isopeptidbindung kommt (Kang et al., 2007). Die Isopeptid-Triade besteht in der C-terminalen Domäne aus den Aminosäuren Lysin-179-Asparagin-303-Glutaminsäure-258 und in der N-terminalen Domäne aus den Aminosäuren Lysin-36-Asparagin-168-Glutaminsäure-117 (genauer Reaktionsmechanismus siehe Kapitel 1.2.2.2). Beide Domänen wurden jeweils rational in ein Peptidfragment (Isopeptag bzw. Isopeptag-N), bestehend aus einem β-Strang, der entweder den reaktiven Asparaginrest oder den reaktiven Lysinrest beinhaltet, und in ein großes Proteinfragment (Pilin-C bzw. Pilin-N), das entweder das reaktive Lysin bzw. das reaktive Asparagin enthält, gespalten (siehe Abb.1.7). Sowohl *in vitro* Rekonstitutionsexperimente als auch massenspektrometrische Analysen ergaben für beide Systeme (Isopeptag und Pilin-C bzw. Isopeptag-N und Pilin-N) die Bildung eines kovalenten Komplexes als Ergebnis der Rekonstitution der reaktiven Triade und der Ausbildung einer Isopeptidbindung zwischen den einzelnen Domänenfragmenten (Zakeri und Howarth, 2010, Kang und Baker, 2011). Die Protein-Peptid Interaktionen waren dabei spezifisch, sodass Pilin-C spezifisch mit Isopeptag, aber nicht mit Isopeptag-N reagierte (Zakeri und Howarth, 2010). So wurde 24h nach Reaktionsstart bei dem effizienteren der beiden Systeme, dem Isopeptag-MBP/Pilin-C Split-System (Zakeri, 2015), eine ca. 50%-ige Ausbeute an rekonstituiertem kovalenten Komplex erreicht (Zakeri und Howarth, 2010). Überraschenderweise war diese Reaktion größtenteils unabhängig von Temperatur (4°C - 37°C), pH-Wert (pH 5 - 8) und Pufferzusammensetzungen (auch Zugabe von Detergenzien), da die Ausbeute und Effizienz/ Geschwindigkeit der kovalenten Produktbildung zwischen Isopeptag und Pilin-C bei allen getesteten Bedingungen vergleichbar waren (Zakeri und Howarth, 2010). Um zu zeigen, dass die Ausbildung einer spezifischen Isopeptidbindung zwischen Isopeptag und Pilin-C auch in komplexeren Umgebungen als im Reaktionsgefäß ablaufen kann, wurden *in vivo* Experimente durchgeführt (Rekonstitution von koexprimiertem Isopeptag-MBP und Pilin-C im

Cytosol von *E. coli* und spezifische Amidbindung an der Oberfläche von Säugerzellen, Zakeri und Howarth, 2010).

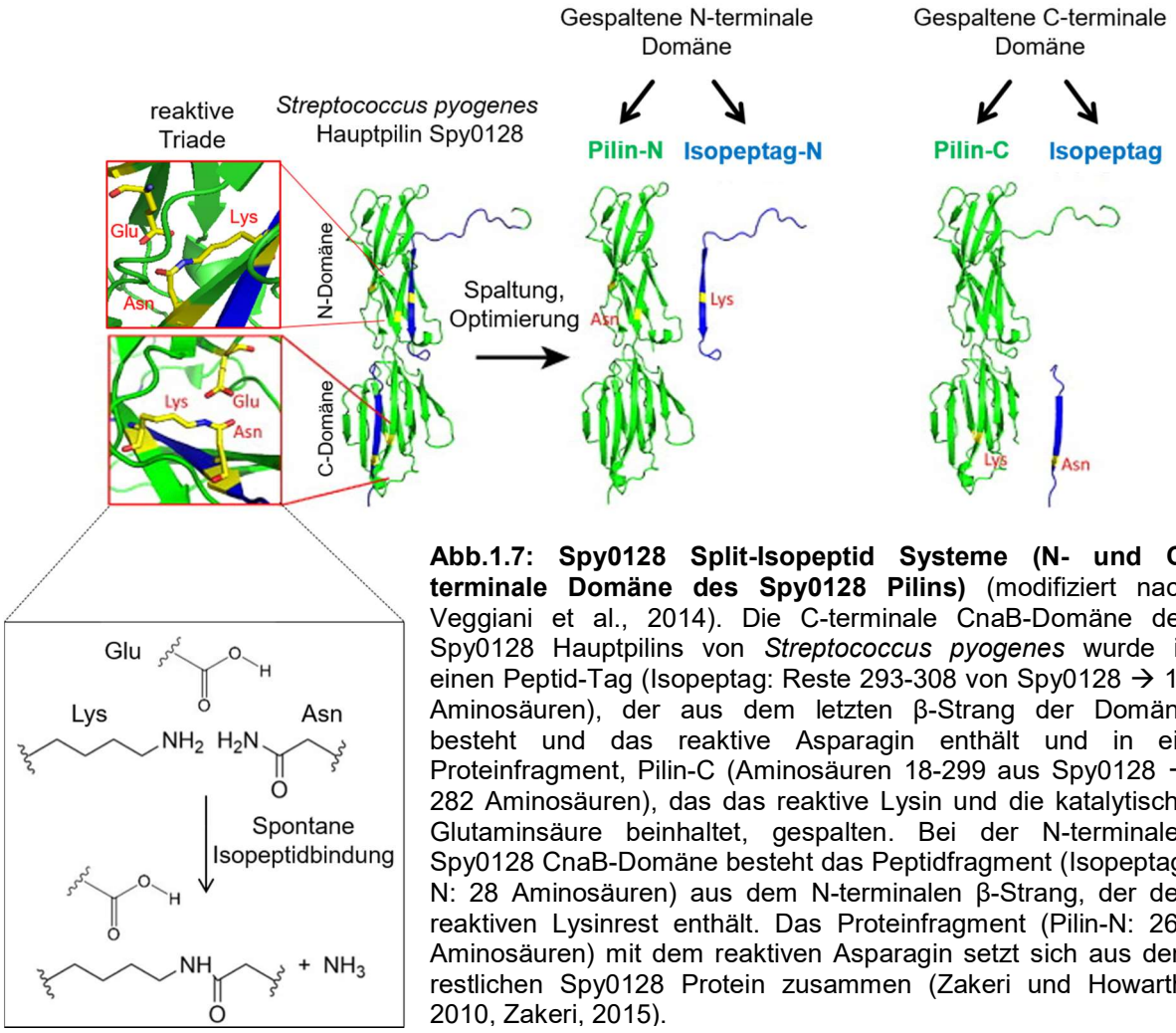


Abb.1.7: Spy0128 Split-Isopeptid Systeme (N- und C-terminale Domäne des Spy0128 Pilins) (modifiziert nach Veggiani et al., 2014). Die C-terminale CnaB-Domäne des Spy0128 Hauptpilins von *Streptococcus pyogenes* wurde in einen Peptid-Tag (Isopeptag: Reste 293-308 von Spy0128 → 16 Aminosäuren), der aus dem letzten β-Strang der Domäne besteht und das reaktive Asparagin enthält und in ein Proteinfragment, Pilin-C (Aminosäuren 18-299 aus Spy0128 → 282 Aminosäuren), das das reaktive Lysin und die katalytische Glutaminsäure beinhaltet, gespalten. Bei der N-terminalen Spy0128 CnaB-Domäne besteht das Peptidfragment (Isopeptag-N: 28 Aminosäuren) aus dem N-terminalen β-Strang, der den reaktiven Lysinrest enthält. Das Proteinfragment (Pilin-N: 263 Aminosäuren) mit dem reaktiven Asparagin setzt sich aus dem restlichen Spy0128 Protein zusammen (Zakeri und Howarth, 2010, Zakeri, 2015).

1.3.2 Zweites Split-Isopeptid System: SpyCatcher-SpyTag System (Spy-System)

Nach der Verifizierung der generellen Funktionalität der Split-Proteindomänen Systeme, basierend auf Isopeptidbindungen (Howarth und Zakeri, 2010), wurde 2012 ein neues, aus kleineren Komponenten bestehendes und schneller reagierendes Split-Isopeptid System, das SpyCatcher-SpyTag System, entwickelt (Zakeri et al., 2012, Veggiani et al., 2014, Zakeri, 2015). Das SpyCatcher-SpyTag System basiert auf der CnaB2-Domäne (*second immunoglobulin-like collagen adhesin domain*) des extrazellulären Oberflächenproteins FbaB (*fibronectin binding protein*) aus dem gram-positiven Bakterium *Streptococcus pyogenes* (Spy) (Zakeri et al., 2012) (siehe Abb.1.8). Innerhalb dieser CnaB2-Domäne kommt es während der Proteinfaltung autokatalytisch zu einer spontanen intramolekularen Isopeptidbindung, die dem extrazellulären Protein FbaB eine enorme thermische,

mechanische und pH-Stabilität verleiht (Kang und Baker, 2011, Schwarz-Linek und Banfield, 2014, Zakeri et al., 2012, Veggiani et al., 2014). Die Isopeptidbindung findet zwischen den Seitenketten des reaktiven Lysinrestes-31 (Lys31) und des reaktiven Asparaginsäurerestes-117 (Asp117) statt und wird durch einen benachbarten, gegenüberliegenden Glutaminsäurerest (Glu77) katalysiert (siehe Abb.1.8). Durch die rationale Spaltung der CnaB2-Domäne und einigen weiteren rationalen Optimierungen im Zuge des ,,Protein Engineerings“ zur Verbesserung der Rekonstitutionseffizienz (Zakeri et al., 2012 Veggiani et al., 2014) wurden zwei genetisch kodierte Reaktionspartner, SpyCatcher-Protein und SpyTag-Peptid, generiert (siehe Abb.1.8). Die SpyCatcher-SpyTag Rekonstitution zu einem kovalenten Komplex, basierend auf einer Isopeptidbindung, findet sowohl *in vitro*, *in vivo* intrazellulär als auch extrazellulär statt (z.B. im Cytosol von *E. coli* Zellen, Säugerzellen) und kann unter den verschiedensten Bedingungen (pH-Wert, Temperatur, Puffer-Bedingungen, Zugabe von Detergenzien) ablaufen (Zakeri et al., 2012, Veggiani et al., 2014). Auch reduzierende Substanzen und der Redox-Status spielen bei der Reaktion keine Rolle, da weder das Catcher-Protein noch das Tag-Peptid Cysteine enthalten. Die kovalente SpyCatcher-SpyTag Interaktion, die auch massenspektrometrisch (Zakeri et al., 2012) und kristallographisch (Li et al., 2014) verifiziert wurde, stellt die stärkste gemessene Proteininteraktion überhaupt dar, die Kräften bis zu 1900 Piconewton standhält (Zakeri et al., 2012, Li et al., 2014, Veggiani et al., 2014). Die Split-Domänen Rekonstitution erfolgt zudem sehr effizient und schnell (ca. 50%-ige Rekonstitution der Fragmente zu einem kovalenten SpyCatcher-SpyTag Komplex innerhalb der ersten Reaktionsminute, Zakeri et al., 2012, Zakeri, 2015), ist stabil (irreversibel), robust (insensitiv), spezifisch und die beiden Reaktionspartner müssen einfach nur, ohne weitere Zusätze, gemischt werden. Dadurch ergeben sich zahlreiche, v.a. biotechnologische Anwendungsmöglichkeiten dieses kovalenten SpyCatcher-SpyTag Systems, die näher in Kapitel 1.3.4 beschrieben werden. So können z.B. Zielproteine durch einfache Genfusion an SpyTag bzw. SpyCatcher post-translational und irreversibel miteinander verknüpft werden (biologisches Protein-Protein Konjugationssystem). Vorteilhaft dabei ist, dass der kleine SpyTag (13 Aminosäuren, Zakeri et al., 2012, Zakeri, 2015, Li et al., 2014) sowohl C-, N-terminal als auch intern an Zielproteine fusioniert werden kann, ohne die Fähigkeit zu verlieren, kovalente Komplexe mit SpyCatcher zu bilden (Zakeri et al., 2012).

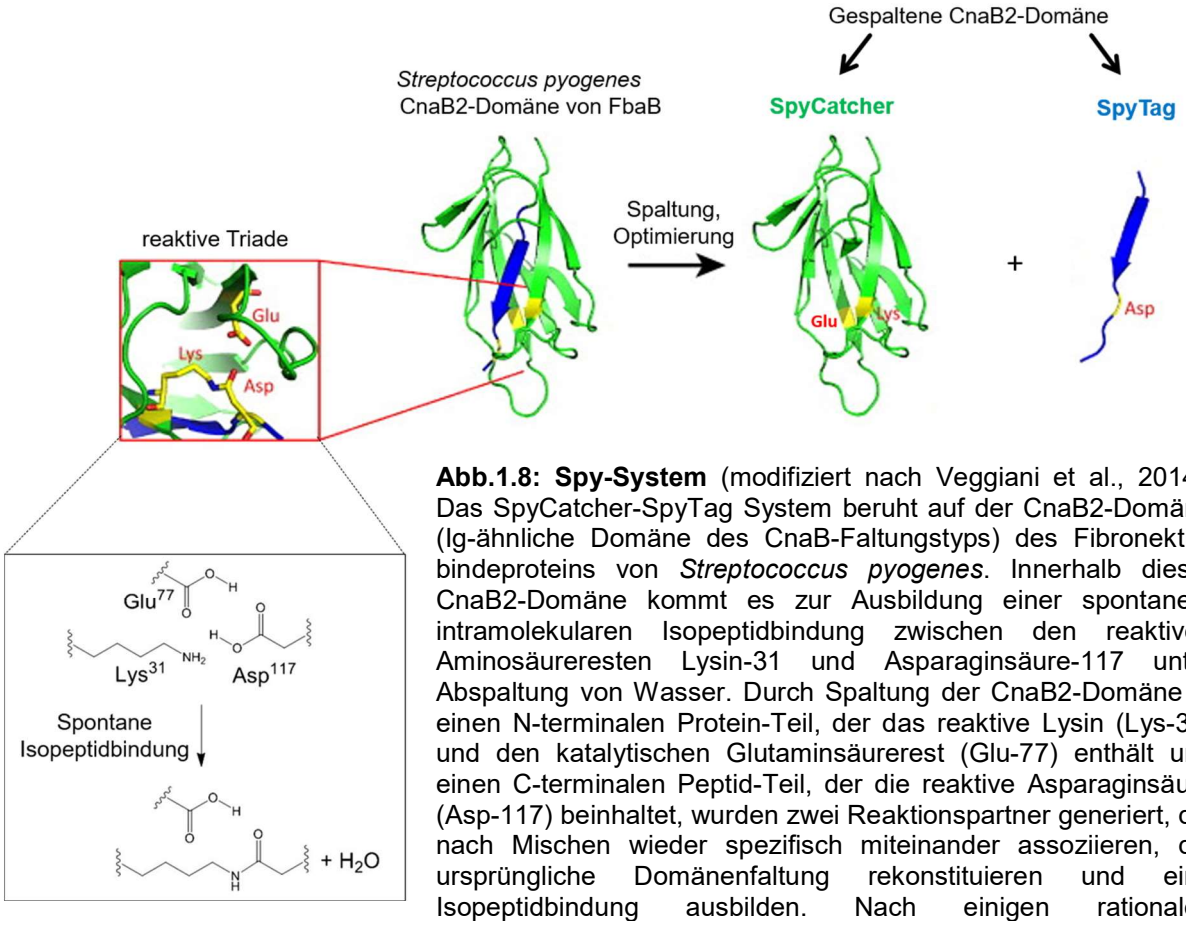


Abb.1.8: Spy-System (modifiziert nach Veggiani et al., 2014). Das SpyCatcher-SpyTag System beruht auf der CnaB2-Domäne (Ig-ähnliche Domäne des CnaB-Faltungstyps) des Fibronektin-bindeproteins von *Streptococcus pyogenes*. Innerhalb dieser CnaB2-Domäne kommt es zur Ausbildung einer spontanen, intramolekularen Isopeptidbindung zwischen den reaktiven Aminosäureresten Lysin-31 und Asparaginsäure-117 unter Abspaltung von Wasser. Durch Spaltung der CnaB2-Domäne in einen N-terminalen Protein-Teil, der das reaktive Lysin (Lys-31) und den katalytischen Glutaminsäurerest (Glu-77) enthält und einen C-terminalen Peptid-Teil, der die reaktive Asparaginsäure (Asp-117) beinhaltet, wurden zwei Reaktionspartner generiert, die nach Mischen wieder spezifisch miteinander assoziieren, die ursprüngliche Domänenfaltung rekonstituieren und eine Isopeptidbindung ausbilden. Nach einigen rationalen Optimierungen zur Verbesserung der Effizienz der kovalenten Reaktion der Reaktionspartner, wurden die Split-Domänenfragmente SpyCatcher (Proteinfragment: initial 116 Aminosäuren (Zakeri et al., 2012), später 84 Aminosäuren (Li et al., 2014)) und SpyTag (Peptidfragment: 13 Aminosäuren) genannt (Zakeri et al., 2012).

Eine Erweiterung des zweiteiligen Spy-Systems, bestehend aus den Proteinpartnern SpyCatcher und SpyTag, ergibt sich durch die Spaltung der CnaB2-Domäne des FbaB Proteins von *Streptococcus pyogenes* in drei Teile (Veggiani et al., 2014, Fierer et al., 2014) (siehe Abb.1.9). Dabei wurde ein weiterer β-Faltblattstrang des früheren SpyCatcher-Proteins, der den reaktiven Lysinrest beinhaltet, abgespalten und KTag genannt. Folglich besteht das System nun aus zwei kleinen Peptid-Tags (SpyTag: 13 Aminosäuren und KTag: 10 Aminosäuren), deren kovalente Interaktion durch den übrig gebliebenen Teil des SpyCatchers, genannt SpyLigase (107 Aminosäuren), katalysiert wird (Fierer et al., 2014, Veggiani et al., 2014) (siehe Abb.1.9). Somit verteilen sich die drei reaktiven Aminosäuren (Asparaginsäure, Lysin und Glutaminsäure) nun auf drei verschiedene Proteinfragmente (siehe Abb.1.9), die bei gemeinsamer Inkubation die reaktive Triade rekonstituieren und eine kovalente Isopeptidbindung zwischen SpyTag und KTag ausbilden (Veggiani et al., 2014). Vorteilhaft ist dieses System im Vergleich zum zweiteiligen SpyCatcher-SpyTag System, da die Peptidfragmente (Tags) sehr klein sind und bei Genfusion an Zielproteine nur minimale

oder keine Effekte auf die native Faltung des Zielproteins haben. Aufgrund der nur minimalen Modifikationen von Zielproteinen stellt dieses dreiteilige kovalente Peptid-Peptid Interaktionssystem den perfekten Ausgangspunkt für die Verknüpfung von Proteinbausteinen dar (Veggiani et al., 2014). Der Nachteil dieses dreiteiligen Systems ist aber, dass es, verglichen mit dem zweiteiligen SpyCatcher-SpyTag System, weniger effizient (langsamer und geringere Ausbeute) ist, spezielle Pufferbedingungen braucht und die Funktionalität in Zellen noch nicht gezeigt wurde (Veggiani et al., 2014).

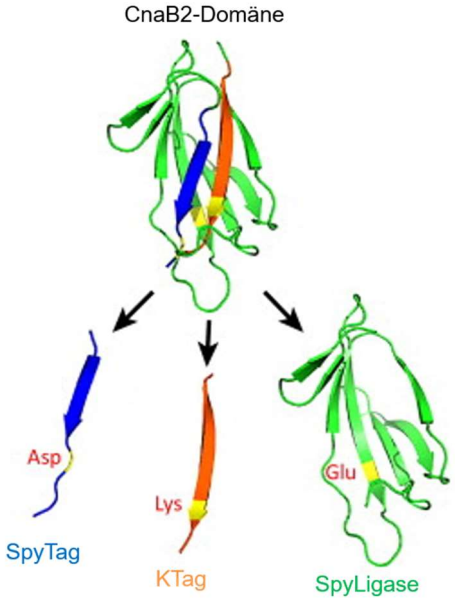


Abb.1.9: 3-teiliges Spy-System (SpyLigase) (modifiziert nach Veggiani et al., 2014). Die CnaB2-Domäne wurde rational in drei Teile gespalten, sodass jedes Fragment (SpyTag, KTag und SpyLigase) eine der drei an der Isopeptidbindung beteiligten Aminosäuren enthält. Während der SpyTag mit dem reaktiven Asparaginrest (13 Aminosäuren) im Vergleich zum zweiteiligen Spy-System unverändert blieb, wurde ein weiterer β-Strang (KTag: 10 Aminosäuren), der das reaktive Lysin enthält, aus der CnaB2-Domäne herausgeschnitten. Das Proteinfragment mit der katalytischen Glutaminsäure, SpyLigase, kann an beide Peptid-Tags andocken, die reaktive Triade rekonstituieren und die Ausbildung einer Isopeptidbindung zwischen den beiden Peptiden vermitteln (Veggiani et al., 2014).

1.3.3 SnoopCatcher-SnoopTag System

Das Snoop-System stellt das neueste Split-Isopeptid System dar und leitet sich von der D4 Ig-ähnlichen Domäne (CnaB-Typ) des RrgA Pilus-Adhäsins des gram-positiven Bakteriums *Streptococcus pneumoniae* ab (Veggiani et al., 2016) (siehe Abb.1.10). Innerhalb dieser D4-Domäne kommt es zur Ausbildung einer intramolekularen, spontanen Isopeptidbindung zwischen den Seitenketten der Aminosäuren Lysin-742 und Asparagin-854 (siehe Abb.1.10). Die kovalente, autokatalytische Reaktion zwischen Lysin und Asparagin wird durch die benachbarte Glutaminsäure-803 katalysiert. Analog zu den Vorgänger Split-Isopeptid Systemen wurde die Domäne, innerhalb der die intramolekluare Isopeptidbindung ausgebildet wird, rational in zwei Teile gespalten (siehe Abb.1.10) und die Reaktion über *rational mutagenesis design* vor allem durch Stabilisierung des SnoopCatcher-Proteins optimiert (Veggiani et al., 2016). Die beiden Reaktionspartner, SnoopTag (N-terminaler β-Strang der D4-Domäne) mit dem reaktiven Lysinrest und SnoopCatcher mit der katalytischen Glutaminsäure und dem reaktiven Asparaginrest, rekonstituierten und bildeten einen kovalenten Komplex, nachdem sie gemischt wurden (Verifizierung auch mittels

Massenspektrometrie, Veggiani et al., 2016). Auch die kovalente Reaktion zwischen SnoopCatcher und SnoopTag toleriert verschiedenste Reaktionsbedingungen (Temperatur, pH-Werte, Pufferzusammensetzungen, Detergenzien) und kann über ein breites Reaktionsspektrum hinweg effizient ablaufen. Das SnoopCatcher-SnoopTag System ist zudem ähnlich effizient wie das SpyCatcher-SpyTag System und die beiden Catcher-Tag Systeme können parallel, z.B. zur stabilen, programmierbaren Assemblierung von multiplen Proteinen (Entwicklung von ,,Polyproteams“, Veggiani et al., 2016), verwendet werden, da sie orthogonal sind und keine Kreuzreaktionen zeigen (Veggiani et al., 2016). Die Orthogonalität des Snoop- und des Spy-Systems wurde vor allem durch die gezielte Verteilung der reaktiven Aminosäuren der Isopeptid-Triade beim rationalen Spalten der Domänen gewährleistet. Anders als bei der CnaB2-Domänenspaltung zur Generierung des Spy-Systems wurde beim Snoop-System nicht der C-terminale, sondern der N-terminale β-Strang der zugrundeliegenden CnaB-Domäne abgespalten (Zakeri et al., 2012, Veggiani et al., 2016). So trägt SnoopTag wie auch SpyCatcher das reaktive Lysin, während SpyTag und SnoopCatcher den reaktiven Asparaginsäurerest bzw. Asparaginrest enthalten. Folglich können nur zusammengehörige Catcher und Tags miteinander reagieren.

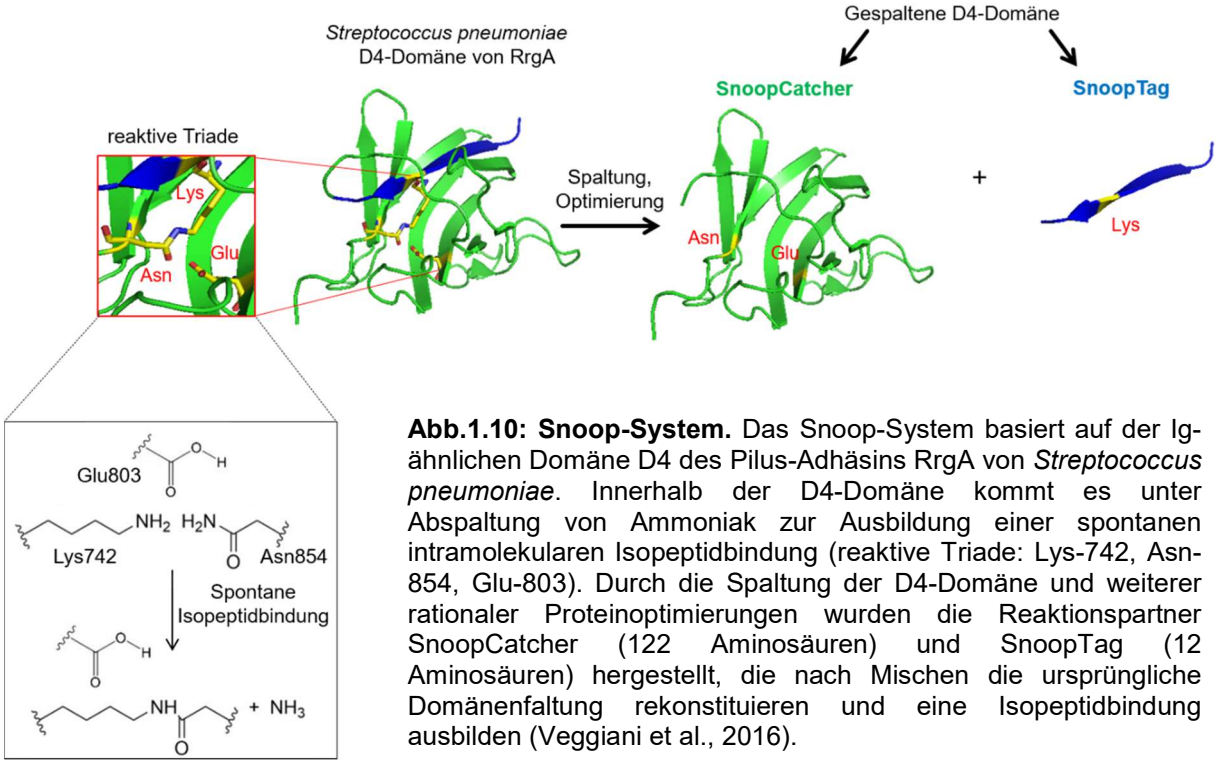


Abb.1.10: Snoop-System. Das Snoop-System basiert auf der Ig-ähnlichen Domäne D4 des Pilus-Adhäsins RrgA von *Streptococcus pneumoniae*. Innerhalb der D4-Domäne kommt es unter Abspaltung von Ammoniak zur Ausbildung einer spontanen intramolekularen Isopeptidbindung (reaktive Triade: Lys-742, Asn-854, Glu-803). Durch die Spaltung der D4-Domäne und weiterer rationaler Proteinoptimierungen wurden die Reaktionspartner SnoopCatcher (122 Aminosäuren) und SnoopTag (12 Aminosäuren) hergestellt, die nach Mischen die ursprüngliche Domänenfaltung rekonstituieren und eine Isopeptidbindung ausbilden (Veggiani et al., 2016).

1.3.4 Anwendungen kovalenter Catcher-Tag Systeme

Die für biotechnologische Anwendungen idealen Protein-Protein oder Protein-Peptid Interaktionen sollen spezifisch, kovalent (stabil auch über einen längeren Zeitraum), robust (insensitiv gegenüber verschiedenen Reaktionsbedingungen), generalisierbar (einsetzbar *in vitro* und *in vivo*) und praktikabel (keine chemischen Modifikationen und Cofaktoren sind nötig) sein (Veggiani et al., 2014). Da die Catcher-Tag Systeme, auch im Vergleich mit allen anderen post-translationalen Proteinligationstechniken, alle diese Kriterien erfüllen, finden die genetisch kodierten Catcher-Tag Systeme, und vor allem das am besten funktionierende und charakterisierte SpyCatcher-SpyTag System, zahlreiche und vielfältige Anwendung in der Synthetischen Biologie und in der Biotechnologie (Reddington und Howarth, 2015, Veggiani et al., 2014, Zakeri, 2015, Zakeri et al., 2012, Gilbert et al., 2017, Pröschel et al., 2017, Banerjee und Howarth, 2017). Beispielsweise wurde, basierend auf dem SpyCatcher-SpyTag System, im Jahr 2017 die Firma SpyBiotech gegründet (http://www.spybiotech.com/web/). Auch die Firma kerafast (http://www.kerafast.com) vertreibt SpyTag- und SpyCatcher-Proteine als irreversible Werkzeuge, sogenannte ,,Protein Coupling Reagents“, für jede Art von post-translationalen Biokonjugationen (Reddington und Howarth, 2015). Anwendungen des SpyCatcher-SpyTag Systems umfassen spezifische Strukturmarkierungen in der Fluoreszenzmikroskopie (Zakeri et al., 2012), Stabilisierung von Enzymen durch Ringschluss (Schoene et al., 2014, Schoene et al., 2016, Wang et al., 2016, Gilbert et al., 2017, Liu et al., 2017), bioaktive Hydrogele und funktionalisierte Materialien (Sun et al., 2014, Gao et al., 2016, Botyanszki et al., 2015, Chen et al., 2014, Siegmund et al., 2016), sowie vielfältige Protein-Protein Ligationstechniken zur Generierung von bestimmten Proteinarchitekturen (Zhang et al., 2013, Liu et al., 2017, Fairhead et al., 2014) und profitablen Multiproteinkomplexen (Veggiani et al., 2016, Banerjee und Howarth, 2017). Durch die Verknüpfung von Proteinen können neben effizienten Multienzymkomplexen (Myhrvold et al., 2016, Giessen und Silver, 2016, Gilbert et al., 2017, Alves et al., 2015) auch effektive Impfstoffe durch die Präsentation multipler Antigene (Liu et al., 2015, Brune et al., 2016, Brune et al., 2017, Thrane et al., 2016) und multivalente Affibodies zur Erhöhung der Avidität generiert werden (Fierer et al., 2014). Das neu entdeckte, ebenso effiziente und vor allem orthogonale Snoop-System erweitert zudem das Anwendungsgebiet und leistet einen großen Beitrag zur Kontrollierbarkeit und Programmierbarkeit aller dieser Anwendungen (Veggiani et al., 2016, Brune et al., 2017). Aufgrund dieser attraktiven und vielfältigen Anwendungsmöglichkeiten ist das Forschungsgebiet zur Entwicklung neuer kovalenter Split-Systeme zur post-translationalen Kontrolle von Proteinen sehr aktiv und schnell wachsend. Vorangetrieben wird diese Bewegung auch dadurch, dass ständig neue, intramolekulare kovalente Bindungen (Isopeptidbindungen, Ester- und Thioesterbindungen) in adhäsiven

Oberflächenproteinen von gram-positiven Bakterien entdeckt werden (Tan et al., 2016, Bonnet et al., 2017, Kwon et al., 2014, Young et al., 2017, Walden et al., 2015, Schwarz-Linek und Banfield, 2014, Pointon et al., 2010, Howarth, 2017). Das neueste Split-Isopeptid System stellt das SdyCatcher-SdyTag System dar, das auf einer CnaB-Faltungsdomäne des Fibronektinbindeproteins von *Streptococcus dysgalactiae* basiert (Tan et al., 2016). Ein kovalentes Split-Proteindomänen System, das auf einer Esterbindung zwischen Glutamin und Threonin beruht, wurde auch kürzlich generiert (Split-Ester System: Young et al., 2017). Das Split-Ester System leitet sich von einer Ig-ähnlichen Domäne des Adhäsins Cpe0147 von *Clostridium perfringens* ab, innerhalb der es zu einer autokatalytischen, intramolekularen Esterbindung zwischen Seitenketten von Threonin und Glutamin kommt (Kwon et al., 2014).

1.4 Zielsetzung der Arbeit

Seit der Entdeckung von kovalenten, intramolekularen Isopeptidbindungen innerhalb Ig-ähnlicher Domänen (CnaA- und CnaB-Typ) in adhäsiven Zelloberflächenproteinen gram-positiver Bakterien versuchen Protein-Ingenieure sich diese Entdeckung zu Nutze zu machen. Im Jahre 2012 wurde, basierend auf der Spaltung der CnaB2-Domäne des Fibronektinbindeproteins von *Streptococcus pyogenes*, das bis heute effizienteste, kovalente Split-Isopeptid System, bestehend aus SpyCatcher und SpyTag, das zahlreiche Anwendungen als Biokonjugationssystem findet, entwickelt. Da vermutlich tausende weitere Ig-ähnliche Domänen mit wahrscheinlich ähnlich vielen intramolekularen Isopeptidbindungen in der Natur, in adhäsiven Zelloberflächenproteinen von gram-positiven Bakterien, vorkommen, scheint die Quelle an alternativen kovalenten Protein-Peptid Interaktionssystemen zur post-translationalen Manipulation von Zielproteinen unerschöpflich. Aber tolerieren alle diese CnaB-Faltungsdomänen eine Spaltung, sodass die Fähigkeit zur Ausbildung einer Isopeptidbindung erhalten bleibt? Und können neu identifizierte Catcher-Tag Systeme durch verschiedene ,,Protein Engineering“-Strategien so optimiert werden, dass sie SpyCatcher-SpyTag Niveau erreichen und somit echte Alternativen darstellen?

Das Ziel dieser Arbeit war (I) die Identifizierung neuer Split-Isopeptid Systeme innerhalb von CnaB-Domänen von extrazellulären Adhäsinen gram-positiver Bakterien mittels bioinformatischer (*in silico*) und experimenteller (*in vitro*) Analysen. Neu identifizierte Catcher-Tag Systeme sollten zudem charakterisiert (II) und mittels rationaler Herangehensweise und gezielter Mutagenese deren Split-Domänen Rekonstitutionseffizienz optimiert werden (III). Zudem sollte die Funktionalität sowohl der etablierten als auch der neu entdeckten Catcher-Tag Systeme *in planta* überprüft (IV) und neue Anwendungsbeispiele dieser kovalenten Catcher-Tag Systeme getestet werden (V).

Notwendig war die Entdeckung von neuen, alternativen Catcher-Tag Systemen vor allem, da zu Beginn dieser Arbeit ausschließlich das SpyCatcher-SpyTag System als effizientes, kovalentes Protein-Peptid Split-System bekannt war (Zakeri et al., 2012). Um aber eine bessere post-translationale Kontrolle über Proteine zu haben, z.B. um mehrere verschiedene Proteine in programmierbarer Weise und am besten gleichzeitig stabil und irreversibel miteinander verknüpfen oder an bestimmten Trägern (ko)immobilisieren zu können, sind alternative kovalente Split-Isopeptid Systeme unabdingbar. Aus diesem Grund sollte diese Arbeit zur Erweiterung des modularen Baukastensystems an kovalenten Catcher-Tag Systemen, basierend auf Isopeptidbindungen, beitragen.

2 Material und Methoden

2.1 Chemikalien, Enzyme und Verbrauchsmaterialien

Alle verwendeten Chemikalien, Enzyme und Verbrauchsmaterialien wurden, sofern nicht anders erwähnt, von den Firmen Roth GmbH & Co. KG (Karlsruhe), Thermo Fisher Scientific (Waltham, Massachusetts), Sigma-Aldrich Chemie GmbH (Steinheim), Applichem GmbH (Darmstadt), Roche Diagnostics GmbH (Mannheim), Invitrogen (Karlsruhe), New England Biolabs GmbH (Frankfurt am Main) und VWR International GmbH (Darmstadt) bezogen.

Für die Isolierung von Plasmid-DNA aus Bakterien (Miniprep Kit, QIAprep Spin Miniprep Kit), die Aufreinigung von PCR-Produkten oder präparativen Verdaus (PCR Purification Kit) und für die Extraktion von DNA-Fragmenten aus Agarosegelen (Gel Extraction Kit) wurden Kits der Firma QIAGEN GmbH (Hilden) verwendet.

2.2 Antikörper

Für die immunchemische Detektion unterschiedlich markierter Proteine im Western Blot wurden folgende Antikörper eingesetzt.

Antikörper	Verdünnung	Referenz/Hersteller
Anti-His-POD (monoklonaler Anti-Polyhistidin Klon HIS-1 Peroxidasekonjugat)	1:2000 - 1:5000	Sigma-Aldrich Chemie GmbH (Steinheim)
Anti-HA-Peroxidase High Affinity (3F10) Ratte monoklonaler Antikörper (Klon 3F10)	1:500	Roche Diagnostics GmbH (Mannheim) REF 12013819001
Anti-GFP (Maus IgG monoklonaler Antikörper; Klone 7.1 und 13.1)	1:1000	Roche Diagnostics GmbH (Mannheim) REF 11814460001
Rabbit anti-Mouse IgG (H+L) Sekundärantikörper – HRP-	1:10000	Thermo Fisher Scientific (Waltham, Massachusetts)

2.3 Vektoren und rekombinante Plasmide

Folgende Vektoren wurden in dieser Arbeit verwendet:

Vektor	Resistenz	Verwendung	Referenz/Hersteller
pCRBlunt	Kan^R	Klonierungsvektor	Invitrogen (Karlsruhe)
pQE-9	Amp^R	Bakterieller Expressionsvektor (N-terminale 6x His Fusion)	Qiagen (Hilden)
pRB-35S	$Spec^R$, $Strep^R$	Binärer Vektor, konstitutive Expression in Pflanzen	F. Börnke

Anhand der Ausgangsvektoren wurde mittels sequenzspezifischer Primer (siehe Tabelle A2 im Anhang) und geeigneter DNA-Matrizen (siehe Tabelle A1 im Anhang), eine Vielzahl von Plasmiden kloniert, welche in der folgenden Tabelle aufgelistet sind (Tabelle 1.1).

Tabelle 1.1: Rekombinante Plasmide.

Name	Klonierung [Primer (MP / Ox), Matrize (#), Größe (kDa)]
Spy-System	
pCRBlunt/GS-SpyCatcher (#15)	Synthetisches Gen
pQE-9/mCherry-GS-SpyCatcher-HA (#35)	MP44+MP66, mCherry MP65+MP60, #15 → overlap PCR: MP44+MP60, 39 kDa
pQE-9/eGFP-GS-SpyCatcher-HA (#36)	MP33+MP64, eGFP MP63+MP55, #15 → overlap PCR: MP33+MP55, 40 kDa
pQE-9/mTagBFP-GS-SpyCatcher-HA (#38)	MP52+MP68, mTagBFP MP67+MP55, #15

	→ overlap PCR: MP52+MP55, 39 kDa
pQE-9/SpyTag-GS-MBP (#44)	Synthetisches Gen, 44 kDa
pQE-9/SpyTag-GS-eGFP-GS-SpyTag (#68)	MP76+MP77, #36, 32 kDa
pQE-9/SpyTag-GS-eGFP-GS-SnoopTag (#81)	MP76+MP89, #68, 32 kDa
pQE-9/SpyCatcher only (#87)	MP101+MP82, #15, 11 kDa
pQE-9/Cystein-TEV-SpyCatcher (#147)	MP200+MP82, #87, 12 kDa
pRB-35S/OEP7-GS-SpyTag-HA (#167)	MP42+MP187, OEP7 (At3g52420), 10 kDa
pRB-35S/eGFP-GS-SpyCatcher-HA (#171)	MP194+MP32, #36, 38 kDa
Snoop-System	
pQE-9/mCherry-GS-SnoopCatcher (#82)	MP87+MP86, #35 MP85+MP88, #111 → overlap PCR: MP87+MP88, 41 kDa
pQE-9/SnoopTag-GSGESG-MBP (#108)	MP121+MP107, #98, 44 kDa
pQE-9/SnoopCatcher only (#111)	MP122+MP88, SnoopCatcher (Prof. M. Howarth), 14 kDa
pRB-35S/OEP7-GS-SnoopTag-HA (#168)	MP42+MP188, OEP7 (At3g52420), 10 kDa
pRB-35S/eGFP-GS-SnoopCatcher-HA (#173)	MP194+MP197, #36 MP198+MP196, #111 → overlap PCR: MP194+MP196, 41 kDa
3phs-System	
pQE-9/3phsCatcher-HA (#52)	Synthetisches Gen, 13 kDa
pQE-9/3phsTag-GS-MBP (#53)	Synthetisches Gen, 44 kDa
pQE-9/mCherry-GS-3phsCatcher (#96)	MP87+MP97, #35 MP96+MP98, #52 → overlap PCR: MP87+MP98, 39 kDa
3kptN-System	
pQE-9/3kptNCatcher-HA (#56)	Synthetisches Gen, 13 kDa

pQE-9/3kptNTag-GS-MBP (#57)	Synthetisches Gen, 44 kDa
pQE-9/mCherry-GS-3kptNCatcher (#94)	MP87+MP91, #35 MP90+MP92, #56 → overlap PCR: MP87+MP92, 39 kDa
4oq1-System	
pQE-9/4oq1Catcher-HA (#58)	Synthetisches Gen, 13 kDa
pQE-9/4oq1Tag-GS-MBP (#59)	Synthetisches Gen, 44 kDa
pQE-9/mCherry-GS-4oq1Catcher (#95)	MP87+MP94, #35 MP93+MP95, #58 → overlap PCR: MP87+MP95, 39 kDa
pQE-9/4oq1Tag-GSGESG-MBP (#98)	MP106+MP107, #59, 44 kDa
pQE-9/4oq1Catcher K155→A (#99)	MP116+MP95, #58, 12 kDa
pQE-9/4oq1Tag N252→A (#100)	MP115+MP107, #98,44 kDa
pQE-9/4oq1Catcher C-terminal VDHQ Verlängerung (#102)	MP110+MP111, #58, 13 kDa
pQE-9/4oq1Tag C-terminal LPRGN Verkürzung (#103)	MP109+MP107, #98, 43 kDa
pQE-9/4oq1Tag N-terminal H243-L245 Verlängerung (#104)	MP108+MP107, #98, 44 kDa
pQE-9/4oq1Tag C-terminal R257-N259 Verkürzung (#105)	MP113+MP107, #98, 44 kDa
pQE-9/4oq1Tag N-terminal L245 Verlängerung (#106)	MP112+MP107, #98, 44 kDa
pQE-9/4oq1Catcher C-terminal Q239-L240 Verkürzung (#107)	MP110+MP114, #58, 12 kDa
pQE-9/4oq1Catcher E222→Q (#110)	MP110+MP140, #107,12 kDa
pQE-9/4oq1Catcher only (#112)	MP110+MP95, #58, 12 kDa
pQE-9/4oq1Catcher I154→T (#126)	Ox31+MP161, #102 MP160+Ox32, #102 → Gibson assembly, 12 kDa

pQE-9/4oq1Catcher I185→E (#127)	Ox29+MP161, #102 MP160+Ox30, #102 → Gibson assembly, 13 kDa
pQE-9/4oq1Catcher F209→Y (#128)	Ox37+MP161, #102 MP160+Ox38, #102 → Gibson assembly, 13 kDa
pQE-9/4oq1Catcher L215→K (#129)	Ox25+MP161, #102 MP160+Ox26, #102 → Gibson assembly, 13 kDa
pQE-9/4oq1Catcher L226→K (#130)	Ox35+MP161, #102 MP160+Ox36, #102 → Gibson assembly, 13 kDa
pQE-9/4oq1Catcher V196→N (#131)	Ox21+MP161, #102 MP160+Ox22, #102 → Gibson assembly, 13 kDa
pQE-9/4oq1Catcher Y229→S (#132)	Ox39+MP161, #102 MP160+Ox40, #102 → Gibson assembly, 12 kDa
pQE-9/4oq1Catcher V177→N (#134)	Ox33+MP161, #102 MP160+Ox34, #102 → Gibson assembly, 13 kDa
pQE-9/4oq1Tag N252→D (#135)	MP156+MP107, #98, 44 kDa
pQE-9/4oq1Tag +2nd β-Strang (#139)	MP165+MP107, #104, 45 kDa
pQE-9/4oq1Catcher-PROSS (#148)	Synthetisches Gen, 13 kDa
pQE-9/4oq1Catcher D176→G (#150)	MP173+MP161, #102 MP160+MP174, #102 → Gibson assembly,12 kDa
pQE-9/4oq1Catcher M147→L (#151)	MP175+MP111, #102,12 kDa
pQE-9/4oq1Catcher N212→G (#152)	MP176+MP161, #102

	MP160+MP177, #102 → Gibson assembly,12 kDa
pQE-9/4oq1Catcher V223→T (#153)	MP178+MP161, #102 MP160+MP179, #102 → Gibson assembly,13 kDa
pQE-9/4oq1Catcher L234→N (#154)	MP180+MP161, #102 MP160+MP181, #102 → Gibson assembly,13 kDa
pQE-9/4oq1Catcher N161→P (#155)	MP182+MP161, #102 MP160+MP183, #102 → Gibson assembly, 12 kDa
pQE-9/4oq1Catcher V231→P (#156)	MP184+MP161, #102 MP160+MP185, #102 → Gibson assembly,13 kDa
pQE-9/4oq1Catcher L234→P (#157)	MP110+MP186, #102, 12 kDa
pRB-35S/OEP7-GS-4oq1Tag-HA (#170)	MP42+MP190, OEP7 (At3g52420), 10 kDa
pRB-35S/mCherry-GS-4oq1Catcher-HA (#175)	MP195+MP199, #95, 39 kDa
3kptC-System	
pQE-9/3kptCCatcher-HA (#54)	Synthetisches Gen, 13 kDa
pQE-9/3kptCTag-GS-MBP (#55)	Synthetisches Gen, 44 kDa
pQE-9/eGFP-GS-3kptCCatcher-HA (#65)	MP78+MP79, #36 MP80+MP81, #54 → overlap PCR: MP78+MP81, 40 kDa
pQE-9/mCherry-GS-3kptCCatcher (#115)	MP87+MP145, #35 MP144+MP146, #54 → overlap PCR: MP87+MP146, 39 kDa
pQE-9/3kptCTag-GSGESG-MBP (#136)	MP157+MP107, #98, 44 kDa
pQE-9/3kptCTag NQ Verlängerung (#137)	MP158+MP107, #136, 44 kDa

pQE-9/3kptCTag GWI (#138)	MP159+MP107, #136, 44 kDa
pQE-9/3kptCCatcher only (#142)	MP169+MP146, #54, 11 kDa
pQE-9/3kptCTag N512→A (#143)	MP168+MP107, #136, 44 kDa
pQE-9/3kptCCatcher E472→Q (#144)	MP171+MP161, #142 MP160+MP172, #142 → Gibson assembly,11 kDa
pQE-9/3kptCCatcher K417→A (#145)	MP201+MP146, #142, 11 kDa
pQE-9/3kptCTag N512→D (#146)	MP202+MP107, #136, 44 kDa
pQE-9/3kptCCatcher-PROSS (#149)	Synthetisches Gen, 11 kDa
pRB-35S/OEP7-GS-3kptCTag-HA (#169)	MP42+MP189, OEP7 (At3g52420), 10 kDa
pRB-35S/eGFP-GS-3kptCCatcher-HA (#174)	MP194+MP32, #65, 39 kDa
pQE-9/3kptCCatcher I418→K (#176)	MP216+MP146, #142, 12 kDa
pQE-9/3kptCCatcher N422→P (#177)	MP207+MP146, #142, 11 kDa
pQE-9/3kptCCatcher K425→P (#178)	MP208+MP146, #142, 11 kDa
pQE-9/3kptCCatcher I440→K (#179)	MP214+MP161, #142 MP160+MP215, #142 → Gibson assembly, 12 kDa
pQE-9/3kptCCatcher E454→P (#180)	MP209+MP161, #142 MP160+MP210, #142 → Gibson assembly, 11 kDa
pQE-9/3kptCCatcher F471→Y (#181)	MP212+MP161, #142 MP160+MP213, #142 → Gibson assembly, 12 kDa
pQE-9/3kptCCatcher E481→S (#182)	MP203+MP161, #142 MP160+MP204, #142 → Gibson assembly, 11 kDa
pQE-9/3kptCCatcher D482→N (#183)	MP205+MP161, #142

	MP160+MP206, #142 → Gibson assembly, 11 kDa
pQE-9/3kptCCatcher Y488→P (#184)	MP169+MP211, #142, 11 kDa

HA ist die Abkürzung für den Hämagglutinin Affinitäts-tag. GS steht allgemein für die Aminosäureabfolge GSGSGSG. Alle Tag-Konstrukte sind über eine linker-Sequenz (entweder GSGSGSG oder GSGESG) an ihrem C-Terminus an das Maltosebindeprotein von *E. coli* (malE Gen ohne Signalpeptid) fusioniert. Sequenzen der synthetischen Gene und Primersequenzen sind im Anhang in Tabelle A1 und A2 hinterlegt.

2.4 Oligonukleotide, synthetische DNA-Fragmente und Sequenzierungen

Alle verwendeten Oligonukleotide (siehe Anhang Tabelle A2) wurden von der Firma Metabion International AG (Martinsried) bezogen. Synthetische DNA-Fragmente (DNA strings) wurden von der Firma life technologies/Thermo Fisher Scientific (Waltham, Massachusetts) hergestellt und sind ebenfalls im Anhang aufgelistet (siehe Tabelle A1). Sanger Sequenzierungen wurden von der GATC Biotech GmbH (Konstanz) durchgeführt.

Synthetische DNA-Fragmente der neuen, alternativen Catcher-Tag Systeme wurden Codon optimiert für die Expression in *E. coli* und mit BamHI und SalI Schnittstellen bestellt, damit nach präparativem Restriktionsverdau direkt und gerichtet in den bakteriellen Expressionsvektor pQE-9, der ebenfalls mit BamHI und SalI geschnitten wurde, ligiert werden konnte. Diese rekombinanten Plasmide dienten vor allem als DNA-Matrizen für die Konstruktion weiterer Genkonstrukte über PCR-Methoden mit spezifischen Primern (siehe auch Klonierung von pQE-9 Expressionskonstrukten 2.6.2).

Alle Sequenzen, die für die alternativen Catcher- und Tag-Konstrukte kodieren, wurden in Abstimmung mit Prof. Dr. Heinrich Sticht und Dr. Anselm Horn generiert. Die Sequenzen und DNA-Matrizen der Snoop-Konstrukte (SnoopTag und SnoopCatcher) stammen von Prof. Mark Howarth (Department of Biochemistry, University of Oxford, Sequenzen siehe Tabelle A1 im Anhang).

2.5 Biologisches Material

2.5.1 Bakterien

2.5.1.1 Stämme

Folgende Bakterienstämme wurden in dieser Arbeit verwendet.

Stamm	Genotyp/Resistenz	Verwendung	Referenz/Hersteller
E. coli DH5α	F– Φ80*lacZ*ΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	Klonierung	Invitrogen
E. coli XL1Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZ ΔM15 Tn10 (Tet^R)]	Klonierung von pQE-9 Expressionskonstrukten	Bullock et al., 1987
E. coli M15 [pREP4]	Basiert auf *E. coli* K12 (Nal^S, Str^S, Rif^S, $Thi^–$, $Lac^–$, Ara^+, Gal^+, $Mtl^–$, $F^–$, $RecA^+$, Uvr^+, Lon^+) Trägt Repressorplasmid pREP4 das Kanamycin Resistenz vermittelt (Kan^R)	High-level Expression rekombinanter Proteine, streng kontrolliert über das Repressorplasmid pREP4 und das IPTG induzierbare lac Operon	Qiagen (Hilden), The QIAexpressionist Fifth Edition (Juni 2003)
A.tumefaciens C58C1	Rif^R mit Helferplasmid pGV2260 (Amp^R)	Pflanzentransformation	Van Larebeke et al., 1974, Deblaere et al., 1985

2.5.1.2 Kultivierung

Die Kultivierung von *Escherichia coli* (*E. coli)* Stämmen erfolgte bei 37°C in LB-Medium (10 g/l Bacto-Trypton, 10 g/l NaCl, 5 g/l Hefeextrakt), die Kultivierung von *Agrobacterium tumefaciens* (*A. tumefaciens)* bei 28°C in YEB-Medium (5 g/l Rinderextrakt, 1 g/l Hefeextrakt, 5 g/l Bacto-Trypton, 5 g/l Saccharose, 2 mM $MgSO_4$). Zur Selektion wurden die Antibiotika in folgenden Endkonzentrationen eingesetzt:

Antibiotika	Endkonzentration
Ampicillin (Amp)	100 µg/ml (MQ)
Kanamycin (Kan)	50 µg/ml (MQ)
Rifampicin (Rif)	50 µg/ml (DMSO)
Spectinomycin (Spec)	100 µg/ml (MQ)
Streptomycin (Strep)	40 µg/ml (MQ)

Die Antibiotika-Endkonzentrationen der Hauptkulturen für die rekombinante Proteinexpression in *E. coli* M15 [pREP4] Zellen (siehe Kapitel 2.6.4.6) betrugen 200 µg/ml Amp und 25 µg/ml Kan (Qiagen, The QIAexpressionist).

2.5.1.3 Transformation von *Agrobacterium tumefaciens*

Für die Transformation von Agrobakterien wurde zu ca. 100 µl von auf Eis aufgetauten kompetenten Zellen ca. 1 – 2 µg Plasmid-DNA gegeben (Höfgen und Willmitzer, 1988). Einem kurzen Kälteschock durch Einfrieren in flüssigem Stickstoff folgte das Auftauen der Zellen für fünf Minuten bei 37°C. Im nächsten Schritt wurde 1 ml LB-Medium zugegeben und der Ansatz bei 28°C für drei bis vier Stunden auf dem Schüttler bei 200 rpm zur Expression der Resistenzgene inkubiert. Die Zellernte erfolgte bei 13000 rpm für eine Minute. Der Überstand wurde verworfen, das Zellpellet im Rückfluss resuspendiert und auf antibiotikahaltigen YEB-Platten ausplattiert. Die Inkubation erfolgte bei 28°C für zwei bis drei Tage.

2.5.2 Pflanzenmaterial – *Nicotiana benthamiana*

2.5.2.1 Anzucht

Nicotiana benthamiana Pflanzen wurden auf Erde im Gewächshaus bei 22°C, 16 h Licht (30 klux Lichtintensität) und einer konstanten Luftfeuchtigkeit von ca. 50% mit täglicher Bewässerung kultiviert.

2.5.2.2 Transiente Agrobakterium-vermittelte Transformation von *N. benthamiana* Blättern

Zur transienten Expression von Proteinen in *N. benthamiana* wurde die Agrobakterien-vermittelte Transformation (ATMT) eingesetzt. Hierfür wurden Agrobakterien-Übernachtkulturen, die mit den zu untersuchenden Genkonstrukten/ Plasmiden transformiert wurden, mit den entsprechenden Antibiotika bei 28°C unter Schütteln über Nacht inkubiert.

Die Bakteriensuspension wurde anschließend zur Zellernte bei 4500 rpm und 4°C für 15 Minuten abzentrifugiert und nach photometrischer Ermittlung der Bakterienkonzentration (Photometer UVIKON XL, Goebel Instrumentelle Analytik, Ludwigshafen) die oD_{600} mit MilliQ Wasser auf 1,0 eingestellt. Zur Analyse mehrerer Genkonstrukte gleichzeitig, zum Beispiel für Kolokalisierungsexperimente oder Interaktionsstudien, wurden die einzelnen Agrobakteriensuspensionen (oD_{600} = 1) im Verhältnis 1:1 gemischt.

Die Infiltration der *N. benthamiana* Pflanzen erfolgte mit einer nadelfreien 1 ml-Spritze auf der Blattunterseite.

Die zu analysierenden Konstrukte wurden stets mit einem zweiten rekombinanten Agrobakterien-Stamm koinfiltriert, der das Gen für das Protein p19 aus dem Tomatenzwergbuschvirus enthält. Dieses virale Protein unterdrückt das *gene silencing* in der Pflanze und erlaubt eine erhöhte transiente Expression der in die Pflanze eingebrachten Genkonstrukte (Voinnet et al., 2003). Für die Infiltration wurden der p19-exprimierende Agrobakterien-Stamm, der auch auf eine oD_{600} von 1 eingestellt wurde, und der Stamm, bzw. die Stämme mit dem zu untersuchenden Konstrukt in einem Verhältnis von 1:1 eingesetzt. Bei Infiltrationen für Interaktions- bzw. Lokalisierungsstudien am KLSM wurde die p19 Konzentration auf eine oD_{600} von 10 eingestellt und später beim Mischen der einzelnen Konstrukte 1/10 Volumen des Gesamtansatzes an p19-Bakterienkultur eingesetzt.

2.6 Molekularbiologische Methoden

2.6.1 DNA-Standardmethoden

Grundlegende Techniken der Nukleinsäuremanipulation, vor allem im Rahmen klassischer Klonierungen, wie die Amplifikation von DNA durch die Polymerase-Kettenreaktion (PCR), Restriktionsverdau, Ligation, Aufreinigung von DNA-Fragmenten, Agarose-Gelelektrophorese von Nukleinsäuren, Transformation von *E. coli* Zellen, sowie Präparation von Plasmiden wurden nach Sambrook und Russell (2001) durchgeführt. Für die Aufreinigung von PCR-Produkten wurde das Qiaquick® PCR Purification Kit (Qiagen) verwendet. Die Extraktion von DNA-Fragmenten aus Agarosegelen erfolgte durch das Qiaquick® Gel Extraction Kit (Qiagen). Für besonders saubere Plasmidisolierungen aus *E. coli* Zellen wurde das QIAprep® Spin Miniprep Kit (Qiagen) verwendet. Dabei wurde nach Protokoll des Herstellers vorgegangen.

2.6.2 Klonierung von pQE-9 Expressionskonstrukten

Zur Klonierung von N-terminal 6x His-markierten Genkonstrukten für die rekombinante Proteinexpression in *E. coli* M15 [pREP4] Zellen und die anschließende Affinitätsaufreinigung dieser Proteine über Ni-NTA Agarose wurde das induzierbare pQE-9 Vektorsystem (Qiagen, The QIAexpressionist) verwendet (siehe Abb.2.1). Der 6x His Affinitäts-tag ist in dem pQE-9 Vektor kodiert, so dass über die multiple Klonierungsstelle (MCS) effizient und leicht N-terminal 6x His-markierte Genfusionen generiert werden können.

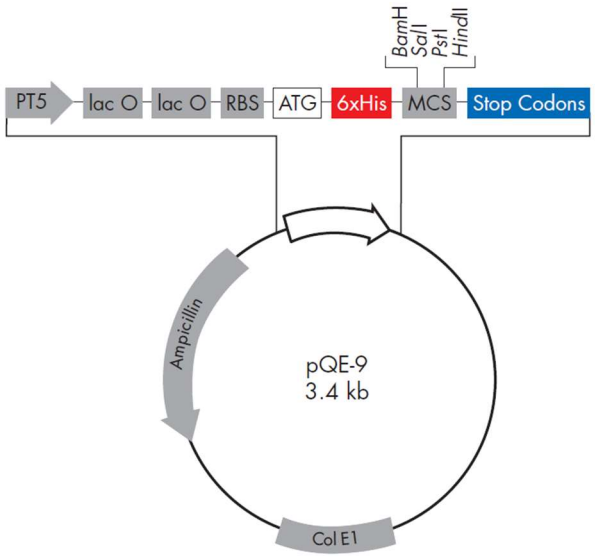


Abb.2.1: Schematische Darstellung des pQE-9 Expressionsvektors (The QIAexpressionist): Der pQE-9 Vektor (*low-copy* Plasmid) besitzt den starken T5-Promotor (PT5) aus dem Bakteriophagen T5. Die hohe Transkriptionsrate des T5-Promotors kann nur in Anwesenheit von hohen Mengen an lac Repressor-Protein effizient reguliert bzw. verhindert werden. Das spezielle "zweifach Operator''-System (2x lac O) in dem pQE Expressionsvektor in Kombination mit dem pREP4 Repressorplasmid (in *E. coli* M15 [pREP4] Zellen vorhanden), das hohe Mengen an lac Repressor-Protein generiert, sichert eine strikte transkriptionelle Kontrolle der klonierten Genkonstrukte. Außerdem besitzt der Vektor eine synthetische ribosomale Bindestelle (RBS) für eine hohe Translationsrate. Das Start Codon (ATG) mit anschließendem 6x His Affinitäts-tag sorgt für eine N-terminale Fusion des His-tags an das in die multiple Klonierungsstelle (MCS) inserierte Genkonstrukt, sodass N-terminale 6xHis-Fusionskonstrukte entstehen. Das β-Lactamase Gen vermittelt eine Ampicillin-Resistenz. Zwei starke Terminatoren sorgen für ein korrektes Transkriptionsende. ColE1 ist der *origin of replication*.

Die Klonierung der pQE-9 Expressionskonstrukte (siehe Tabelle 1.1) erfolgte vorwiegend über klassische Klonierung (siehe Abb.2.2). Das DNA-Insert, das für das jeweilige zu exprimierende Zielprotein kodiert, wurde stets in den Leserahmen des 6x His-tags des pQE-9 Vektors über kompatible Enden kloniert.

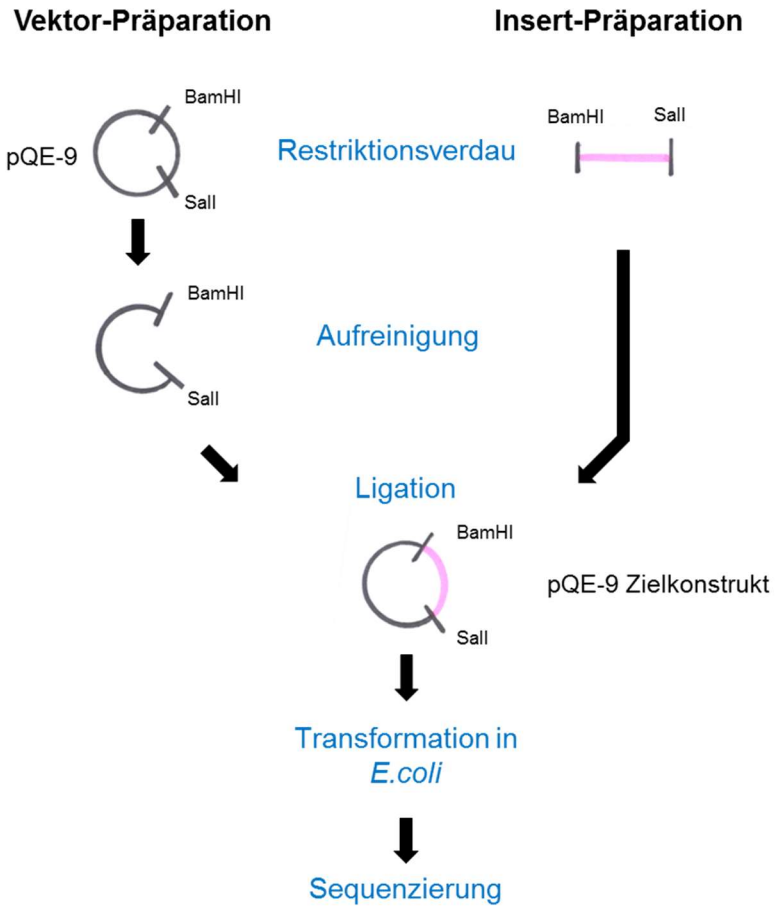


Abb.2.2: Schema der Klonierung von pQE-9 Expressionskonstrukten. Der pQE-9 Vektor und das Insert, welches über PCR (Phusion DNA-Polymerase) hergestellt und amplifiziert wurde, wurden mit denselben Restriktionsenzymen (meist BamHI und SalI) verdaut, um kompatible Enden zu generieren. Aufgereinigte, verdaute DNA-Fragmente wurden ligiert, in *E. coli* XL1Blue Zellen transformiert, Plasmide isoliert und die Sequenzen der Zielkonstrukte über Sanger Sequenzierung überprüft (GATC).

Für die Generierung und Amplifikation der gewünschten DNA-Inserts (Zielgene) mittels PCR-Methoden wurde stets die High-Fidelity Phusion DNA-Polymerase (Thermo Fisher Scientific, Waltham, Massachusetts) verwendet. Über sequenzspezifische Primer (siehe Anhang Tabelle A2) und geeignete DNA-Matrizen (siehe Tabelle 1.1, Sequenzen im Anhang Tabelle A1) wurden die jeweiligen Sequenzen generiert. Vorwärts- und Rückwärtsprimer wurden verwendet, um die entsprechenden DNA-Matrizen in PCR-Reaktionen N- und C-terminal zu verlängern oder zu verkürzen bzw. um N- und C-terminale Aminosäureaustausche durch gezielte Nukleotidveränderungen vorzunehmen (*site-directed mutagenesis*). Nukleotide, die mitten in der Sequenz der DNA-Matrize lagen, wurden über das Gibson Assembly® Cloning Kit (NEB), *origin of replication* Primer (MP160 und MP161) und komplementäre Mutagenese-Primer, die die zu mutierende Sequenz enthielten, mutiert. Bei allen durchgeführten DNA-Manipulationen wurde stets die *E. coli Codon usage* beachtet. Bei der Verwendung von zwei verschiedenen DNA-Matrizen, v.a. zur Konstruktion von Fusionskonstrukten (wie z.B. N-

terminale mCherry-GS-Catcher Fusionen), wurde die Methode der *overlap-extension*-PCR (Sambrook und Russell, 2001, Bryksin und Matsumura, 2010) mit überlappenden, komplementären Primern verwendet. Bei der klassischen, gezielten Klonierung über Restriktionsenzyme und kompatible Enden wurden die PCR-Produkte aus dem Agarosegel ausgeschnitten (Gel Extraction Kit, Qiagen) und präparativ mit Restriktionsenzymen (meist BamHI/SalI, Thermo Fisher Scientific) verdaut und über Säulen aufgereinigt (PCR Purification Kit, Qiagen). Anschließend erfolgte die gerichtete Ligation der verdauten Inserts (T4 DNA Ligase, Thermo Fisher Scientific) in den *E. coli* Expressions-Zielvektor pQE-9, der zuvor ebenfalls mit denselben Restriktionsenzymen geschnitten und über das PCR Purification Kit (Qiagen) aufgereinigt wurde. Die Transformation chemisch kompetenter *E. coli* XL1Blue Zellen, Plasmidisolierung und Sanger-Sequenzierung (GATC) zur Verifizierung der Sequenzen schlossen sich an. Bei der Gibson Klonierungsstrategie wurden ebenfalls die PCR-Produkte aus dem Agarosegel ausgeschnitten (Gel Extraction Kit, Qiagen) und für die Gibson Assembly Reaktion unter Verwendung des 2x Gibson Assembly Master Mixes (NEB) eingesetzt (Durchführung gemäß Herstellerangaben, NEB). Es folgten, analog der klassischen Klonierung, die Transformation in *E. coli* XL1Blue Zellen, Plasmidisolierung und Sequenzüberprüfung (Sanger Sequenzierung, GATC).

2.6.3 Klonierung von pRB-35S Expressionskonstrukten

Konstrukte, die in *Nicotiana benthamiana* konstitutiv über den CaMV35S-Promotor überexprimiert werden sollten (siehe Tabelle 1.1), wurden klassisch über den bakteriellen Zwischenvektor pCRBlunt in den pflanzlichen Expressionsvektor pRB-35S kloniert (siehe Abb.2.3). Die zugrundeliegenden DNA-Sequenzen (inklusive N-terminaler Kozak-Sequenz AACA vor dem Start-Codon) der zu analysierenden Proteine wurden über PCR-Methoden mit spezifischen Primern (siehe Anhang Tabelle A2) mittels der Phusion-DNA Polymerase (Thermo Fisher Scientific) hergestellt. Die PCR-Produkte wurden aus dem Gel extrahiert und über glatte Enden in den StuI vorverdauten pCRBlunt Vektor mittels T4 DNA-Ligase (Thermo Fisher Scientific) ligiert. Anschließend wurden chemisch kompetente *E. coli* DH5α Zellen transformiert, Plasmidpräparationen gemacht und die korrekten DNA-Sequenzen mittels Sanger Sequenzierung (GATC) verifiziert. Die Umklonierung in den pRB-35S Zielvektor erfolgte über präparative Restriktionsspaltungen sowohl des Zielvektors als auch der Inserts, Aufreinigungen der DNA-Fragmente und gerichtete T4 Ligation (siehe Abb.2.3). Plasmidpräparationen mit anschließenden Restriktionsverdaus bestätigten die korrekten pRB-35S Expressionskonstrukte.

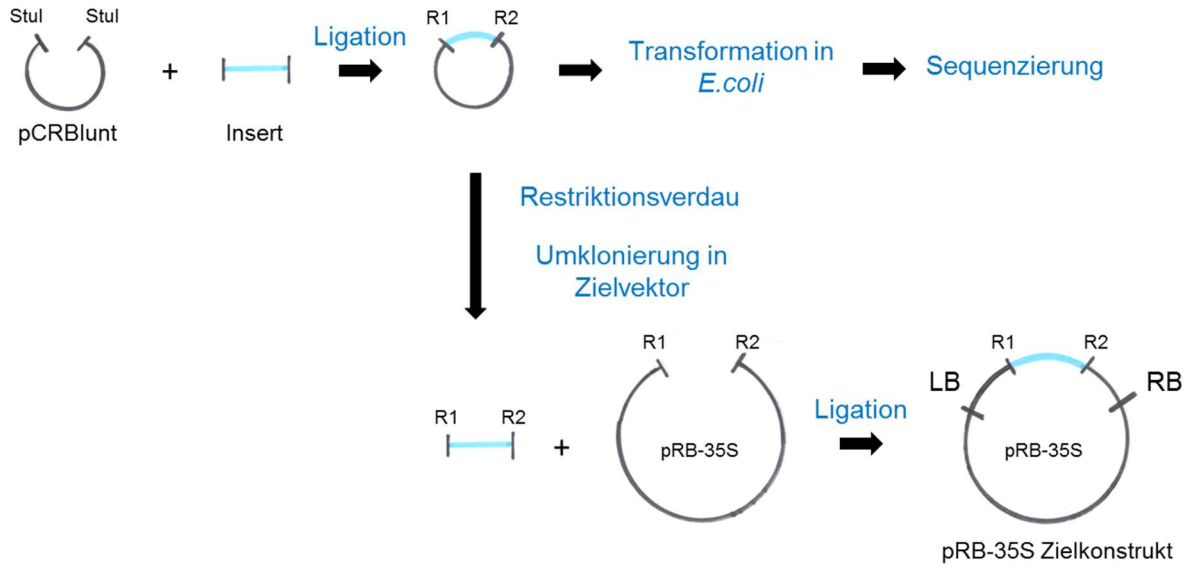


Abb.2.3: Schema der Klonierung von pRB-35S Expressionskonstrukten. Der Zwischenvektor pCRBlunt wurde mit dem Restriktionsenzym StuI (Thermo Fisher Scientific) zur Generierung von glatten Enden verdaut und aufgereinigt. Das Zielgen (Insert) wurde mittels Phusion DNA-Polymerase (produziert glatte Enden im PCR-Produkt) in PCR-Reaktionen generiert und anschließend aufgereinigt. Nach der Ligation folgten die Transformation in *E. coli* DH5α Zellen, Plasmidisolierungen und Verifizierung der Insertsequenz (GATC). Korrekte Inserts wurden über spezifische Restriktionsschnittstellen (R1 und R2) aus dem Zwischenvektor geschnitten und über Ligation in den analog verdauten (R1 und R2) Pflanzenexpressionsvektor pRB-35S ligiert (LB: *left border*, RB: *right border*). Plasmidisolierungen und Kontrollverdaus identifizierten positive Zielkonstrukte.

2.6.4 Protein-Methoden

2.6.4.1 Probenvorbereitung für die gelelektrophoretische Auftrennung von Proteinen

2.6.4.1.1 *E. coli* Proben

Alle *E. coli* Proben, sowohl Gesamtzell-Rohextrakte bei Expressionskinetiken (eingestellt auf dieselbe oD_{600}), verschiedene Fraktionen der Proteinaufreinigung (Pellet, Überstand, Durchfluss, Waschschritte, Eluate) als auch aufgereinigte Proteine bei Interaktionskinetiken/ Rekonstitutionsexperimenten, wurden mit denaturierendem, SDS-haltigem 4x Lämmli-Probenpuffer versetzt und 10 min bei 95°C gekocht bevor sie auf 12%-ige Bis Tris Gele zur SDS-PAGE aufgetragen wurden. Gleiche Volumina der Proben wurden stets auf die Gele aufgetragen, um verschiedene Proben direkt miteinander vergleichen zu können.

4x Lämmli-Probenpuffer	
200 mM	Tris HCl pH 6,8
18%	β-Mercaptoethanol (v/v)
40 %	Glycerin (v/v)
0,01 %	Bromphenolblau (w/v)
8 %	SDS (w/v)

2.6.4.1.2 *N. benthamiana* Blatt-Zellaufschluss mit Homogenisator

Für die Analyse von Gesamtproteinextrakten aus Pflanzen, vor allem für Protein-Expressionsnachweise, wurden in flüssigem Stickstoff eingefrorene *N. benthamiana* Blattscheiben (5er Korkbohrer, 0,9 cm Durchmesser) mit Hilfe des Rotor-Stator-Homogenisators (Heidolph, Schwabach) aufgeschlossen und mit 70 µl 4-fach konzentriertem, denaturierendem Lämmli-Probenpuffer versetzt. Fertig homogenisierte Proben wurden auf Eis gehalten bis alle Proben aufgearbeitet wurden. Anschließend wurden die Proben für 10 Minuten bei 95°C denaturiert und für eine Minute bei 13000 rpm zentrifugiert. Entweder wurde der Überstand der Proben, direkt auf ein 12%-iges Bis-Tris Polyacrylamidgel aufgetragen oder bei -20°C bis zur Weiterverwendung gelagert.

2.6.4.2 Bis-Tris SDS-Polyacrylamid-Gelektrophorese (SDS-PAGE)

Zur gelelektrophoretischen Auftrennung von Proteinen anhand ihres Molekulargewichts wurden 12%-ige Bis-Tris-Polyacrylamidgele (PAA) verwendet.

Trenn- und Sammelgel des Bis-Tris-PAA-Gels setzten sich wie folgt zusammen:

12% Trenngel		Sammelgel	
1,43 ml	3,5x Bis Tris Puffer (1,25 M) pH 6,5-6,8	0,71 ml	3,5x Bis Tris Puffer (1,25 M) pH 6,5-6,8
2 ml	30% AA / 0,8% Bis-AA	0,35 ml	30% AA / 0,8% Bis-AA
1,57 ml	MQ-Wasser	1,44 ml	MQ-Wasser
20 µl	10% APS (v/v)	25 µl	10% APS (v/v)
2 µl	TEMED	4 µl	TEMED

1x MOPS-Tris-Puffer diente als Laufpuffer der Gelelektrophorese.

5x MOPS-Tris-Laufpuffer	
250 mM	MOPS
250 mM	Tris
5 mM	EDTA
0,5%	SDS (w/v)

Als Standard wurde stets der “Page Ruler Prestained Protein Ladder” (Thermo Fisher Scientific) verwendet.

2.6.4.3 Coomassie-Färbung

Die Coomassie-Färbung erlaubt es, Proteine nach der gelelektrophoretischen Auftrennung durch Bindung des Farbstoffes an basische Aminosäure-Seitenketten (im sauren Milieu) im Gel zu visualisieren. Verwendet wurde eine Mischung der Farbstoffe Coomassie Brilliant Blue R-250 und G-250. Zum Anfärben der Proteine wurden die Bis-Tris Gele nach der Gelelektrophorese mit Coomassie-Lösung bedeckt und für mindestens 10 min schüttelnd inkubiert. Durch anschließendes Entfärben der Gele mit Entfärberlösung wird der unspezifisch gefärbte Hintergrund entfärbt und die spezifischen Proteinbanden werden deutlich sichtbar. Die Intensitäten der Bandenfärbung lassen dabei erste Rückschlüsse auf die Quantität der Proteine zu, wobei eine genaue Quantifizierung nur über densitometrische Analysen (Kapitel 2.6.4.4) möglich ist.

Coomassie-Lösung	
0,05%	Coomassie Brilliant Blue R-250 / G-250 (w/v)
40%	Methanol
10%	Essigsäure
ad. 50%	H_2O

Entfärberlösung	
40%	Methanol
10%	Essigsäure
50%	H_2O

2.6.4.4 Densitometrie zur Quantifizierung von Proteinen im Gel

Zur genaueren Quantifizierung der Proteine im Proteingel wurden densitometrische Analysen, das heißt quantitative Messungen der Farbdichte (Bandendicke und Intensität), mittels des Bildbearbeitungs- und Bildverarbeitungsprogramms ImageJ durchgeführt. Die in Replikaten durchgeführten Experimente wurden einzeln ausgewertet und die Mittelwerte und Standardabweichungen in Microsoft Excel berechnet. Die Effizienz der kovalenten Reaktion zwischen Catcher und Tag wurde auf zwei Arten bestimmt. Entweder wurde die Zunahme der Catcher-Tag Produktbande im Vergleich zur 0h-Kontrolle, oder die Abnahme der Input-Banden der Substrate, Catcher und Tag (Input-Banden als 100%-Referenz), quantifiziert.

2.6.4.5 Elektrotransfer von Proteinen auf Nitrocellulosemembran und Immundetektion (Western Blot)

Bei einem Western Blot (Immunblot) werden die zuvor durch SDS-PAGE der Größe nach aufgetrennten Proteine unmittelbar vom Gel auf eine proteinbindende Nitrocellulosemembran (feste Phase) transferiert und dort einer Nachweisreaktion mittels spezifischer Antikörper unterzogen. Der Transfer der Proteine vom Gel auf die Membran (GE Healthcare Life Sciences) im Halbtrockenverfahren (Membran und Whatman-Filterpapiere wurden in 1x Transferpuffer getränkt) erfolgte durch Anlegen eines elektrischen Feldes (Elektrotransfer: 80 bis 90 Minuten bei 0,8 mA pro cm² Gel).

1x Transferpuffer	
39 mM	Glycin
48 mM	Tris
20%	Methanol (v/v)
pH mit HCl auf 8,2 bis 8,3 eingestellt	

Vor Beginn der Nachweisreaktion der Proteine mittels Antikörpern wurden freie, unspezifische Proteinbindestellen auf der Membran durch mindestens einstündiges Schwenken in Blockierlösung (5% Magermilch in 1x TBS-T) abgesättigt. Nach dem Blockieren wurde die Nitrocellulosemembran mit Antikörper-Lösung (Antikörper verdünnt in Verdünnungspuffer: 1% Magermilch in 1x TBS-T) für ein bis zwei Stunden bei Raumtemperatur unter Schütteln inkubiert. Die spezifischen Verdünnungen der einzelnen Antikörper sind unter Gliederungspunkt 2.2 aufgeführt.

	1x TBS-Puffer	**1x TBS-T-Puffer**	**Blockier-lösung**	**Verdünnungs-puffer**
Tris HCl, pH 7,5	20 mM	20 mM	20 mM	20 mM
NaCl	500 mM	500 mM	500 mM	500 mM
Tween 20 (v/v)	--	0,1%	0,1%	0,1%
Magermilchpulver (w/v)	--	--	5%	1%

Anschließend wurde die Membran dreimal in 1x TBS-T-Puffer für jeweils 15 Minuten gewaschen, um nicht oder nur schwach gebundene Antikörper zu entfernen. Handelte es sich bei dem Antikörper bereits um ein Peroxidase-Konjugat (Anti-HA-POD, Anti-His-POD) konnte die Detektion der Zielproteine direkt durchgeführt werden (siehe unten). Die Inkubation mit Sekundärantikörper-Peroxidase-Konjugat (rabbit anti-mouse-POD-Konjugat in Verdünnungspuffer) war für den Anti-GFP Antikörper notwendig und erfolgte für weitere ein bis zwei Stunden bei Raumtemperatur unter Schütteln. Nach erneutem dreimaligen Waschen mit 1x TBS-T-Puffer und MilliQ-Wasser, erfolgte die Visualisierung des markierten Zielproteins mittels ECL-Reaktion (*enhanced chemiluminescence*). Dabei wurde die Membran in einem 1:1 Verhältnis der ECL-Lösungen 1 und 2 für eine Minute unter ständigem Schütteln inkubiert. Bei der ECL-Reaktion oxidiert die Peroxidase, in Anwesenheit von Wasserstoffperoxid als Oxidationsmittel, das Substrat Luminol in eine lumineszierende Persäure, und die emittierte Lumineszenz kann durch Auflegen eines Röntgenfilms sichtbar gemacht werden.

ECL-Lösungen			
Lösung 1		**Lösung 2**	
5 ml	100 mM Tris HCl pH 8,5	5 ml	100 mM Tris HCl pH 8,5
72 µl	CuLu-Mix	3 µl	H_2O_2

CuLu-Mix	
250 mM	Luminol
90 mM	Cumarinsäure

2.6.4.6 Heterologe Expression rekombinanter Proteine in *E. coli* M15 [pREP4] Zellen

Nach Klonierung der pQE-9 Expressionskonstrukte (siehe Kapitel 2.6.2) und deren Transformation in chemisch kompetente M15 [pREP4] *E. coli* Zellen, wurden die gewünschten Proteine als N-terminale 6x His-Fusionen in dem heterologen Wirt *E. coli* M15 [pREP4] nach IPTG-Zugabe (Isopropyl-x-D-thiogalactopyranosid, Roth) überexprimiert. *E. coli* Vorkulturen (Einzelkolonie in LB-Medium versetzt mit 200 µg/ml Amp und 25 µg/ml Kan Endkonzentration), die bei 28°C unter Schütteln (180 rpm) über Nacht inkubiert wurden, wurden eingesetzt, um die Expressions-Hauptkulturen (LB-Medium mit 200 µg/ml Amp und 25 µg/ml Kan Endkonzentration) anzuimpfen. Anfänglich wurde die Hauptkultur auf eine oD_{600} von 0,2 angeimpft. Im Zuge von Optimierungen wurde dieses Protokoll umgestellt und die Hauptkulturen mit 1/100 der Vorkulturen (10 ml Vorkultur auf 1 l Hauptkultur) auf eine oD_{600} von ca. 0,02 angeimpft. Die Zellen wurden bei 28°C schüttelnd (200 rpm) inkubiert bis sie eine oD_{600} von ca. 0,5 erreichten. Anschließend erfolgte die Induktion der Proteinexpression durch Zugabe von 1 mM bzw. 0,5 mM IPTG. Eine IPTG-Endkonzentration von 0,5 mM wurde im Rahmen der Optimierung der Expressionsbedingungen schließlich beibehalten. Die Bakterienzellen wurden weitere vier Stunden bei 28°C und 200 rpm inkubiert, bevor sie durch Zentrifugation bei 5000 g für 20 min und 4°C geerntet wurden. Zellpellets wurden bei -20°C bis zur Weiterverwendung (Proteinaufreinigung) gelagert. Für Expressionskinetiken (Überprüfung der Expression rekombinanter Proteine über die Zeit) wurden stündlich Proben aus der Hauptkultur genommen, die oD_{600} am Photometer

bestimmt und die pelletierten Zellen in entsprechenden Mengen 4x Lämmli-Puffer aufgenommen, sodass alle Proben auf dieselbe oD_{600} eingestellt wurden, bevor sie 10 min bei 95°C gekocht und auf Proteingele (12% Bis-Tris Gele) aufgetragen wurden.

2.6.4.7 *E. coli* Zelllyse und Aufreinigung 6x His-markierter Proteine über Ni-NTA Affinitätschromatographie unter nativen Bedingungen

Zellpellets der heterologen Proteinexpression (Pellet aus 1 l Hauptkultur) wurden auf Eis aufgetaut und in jeweils 8 ml Lysepuffer (The QIAexpressionist, Qiagen: 2 – 5 ml Lysepuffer/g Nassgewicht) mit Pefabloc (1 mM Endkonzentration, Roth) und Protease Inhibitor Cocktail (cOmplete ULTRA Tablets EDTA-free, Roche) aufgenommen und resuspendiert. Die Zelllyse der Bakterienzellen erfolgte durch Ultraschallbehandlung auf Eis (6x 10 sec sonifizieren mit jeweils 10 sec Pause zwischen den Impulsen; SONOPULS Ultraschall-Homogenisator, Bandelin electronic GmbH & Co. KG, Berlin) im Kühlraum. Zur Trennung/ Fraktionierung von löslichen (Überstand) und unlöslichen Proteinen und Zellbestandteilen (Pellet) wurde das Zelllysat anschließend 30 min bei 10000 g und 4°C zentrifugiert (Sorvall RC-6 Plus, Thermo Fisher Scientific, Thermo Electron LED GmbH, Osterode). Der Lysat-Überstand, der die löslichen Proteine (v.a. auch die gewünschten 6x His-markierten Proteine) enthält, wurde für die Aufreinigung weiterverwendet. Ungefähr 4 ml Lysat-Überstand wurden mit 1 ml Ni-NTA Agarose (Ni-NTA = Nickel-Nitrilotriessigsäure, Qiagen), die zuvor mit Lysepuffer äquilibriert wurde, gemischt (The QIAexpressionist) und für ca. 2 h im Kühlraum im Überkopfroller in Polypropylensäulen mit Polyethylen-Filterscheiben (5 ml, Qiagen) inkubiert, damit die 6x His-markierten Proteine an die Ni-NTA Matrix binden (siehe Abb.2.4). Alles, was nicht an die Ni-NTA Agarose Matrix gebunden hatte, wurde mit dem Durchfluss entfernt. Schwach und unspezifisch gebundene Proteine wurden durch dreimaliges Waschen mit Waschpuffer von der Ni-NTA Matrix gewaschen. Die gewünschten 6x His-markierten Proteine wurden schließlich mit Elutionspuffer von der Ni-NTA Matrix eluiert. Zur Überprüfung der Proteinaufreingung wurden Proben aller relevanten Schritte (Lysat-Überstand, Lysat-Pellet, Durchfluss, Waschschritte, Eluate) genommen, mit 4x Lämmli-Probenpuffer versetzt, 10 min bei 95°C gekocht und SDS-PAGE Analysen mit anschließender Coomassie-Färbung unterzogen.

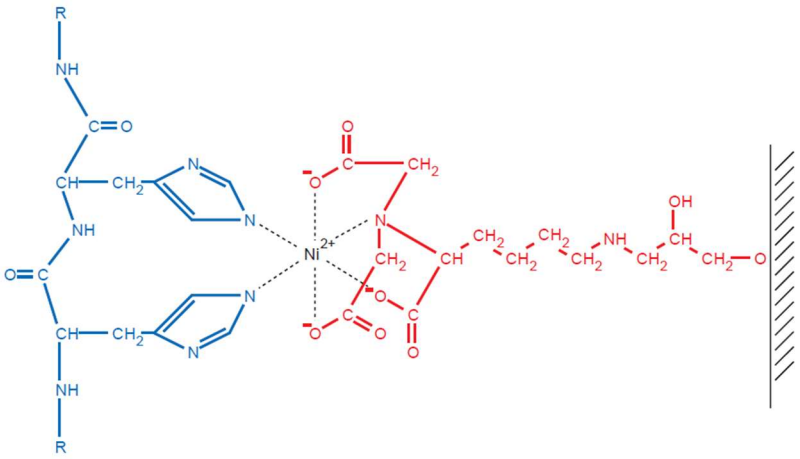


Abb.2.4: Schematische Darstellung der Interaktion der Ni-NTA Matrix mit benachbarten Histidinresten der 6x His-markierten rekombinanten Proteine (The QIAexpressionist). Durch molekulare Interaktionen der Histidinreste der rekombinanten Proteine (blau) mit der Säulenmatrix aus Ni-NTA (Nickel-Nitrilotriessigsäure) Agarose (Qiagen, Hilden) (rot mit zentralem Nickel-Ion in schwarz), werden die Zielproteine in der festen Phase immobilisiert, können gewaschen und mit Hilfe eines stark imidazolhaltigen Puffers spezifisch durch Kompetition um die Bindestellen/ Koordinationsstellen am zentralen Nickel-Ion, eluiert werden. Die Immobilisierung erfolgt über die Koordination der freien Elektronenpaare der Imidazol-Strukturen der Histidinreste an das Nickel-Zentralatom, welches über vier Bindestellen koordinativ an die Nitrilotriessigsäure gebunden ist. Dabei stellt der Chelator NTA alleine die freien, bindenden Elektronenpaare (NTA besitzt 4 freie Elektronenpaare) bereit, die dann mit dem Nickel-Ion an vier Bindestellen einen stabilen Chelat-Komplex ausbilden.

	Lysepuffer	**Waschpuffer**	**Elutionspuffer**
NaH_2PO_4 pH 8,0	50 mM	50 mM	50 mM
NaCl	300 mM	300 mM	300 mM
Imidazol	10 mM	20 mM	250 mM
Frisch zugeben			
1 mM Pefabloc	✓	---	---
Protease Inhibitor Cocktail	✓	---	---

2.6.4.8 Dialyse, Konzentrationsbestimmung und Aufkonzentrierung von Proteinen

Nach der Aufreinigung wurden die rekombinanten Proteine, die noch in Elutionspuffer gelöst waren, mittels Dialyseschläuchen (Servapor dialysis tubing, MWCO 12000-14000 bzw. MWCO 3500, Serva Electrophoresis GmbH, Heidelberg) in 1xPBS pH 7,4 dialysiert. Dafür

wurden die Proteine in die hydrierten Dialyseschläuche mit einem bestimmten Größenausschluss (MWCO: *molecular weight cut off*), entsprechend ihres Molekulargewichts, eingefüllt und die Dialyse in 5 l 1xPBS pH 7,4 im Kühlraum über Nacht unter konstantem Durchmischen inkubiert. Am Folgetag wurde der Dialysepuffer (5 l 1xPBS pH 7,4) gewechselt und einige Stunden weiter inkubiert. Für die Bestimmung der Proteinkonzentration wurde die Absorption bei 280 nm am NanoDrop Spektrophotometer ND-1000 (Peqlab) bestimmt (Programm A280). Unter Berücksichtigung der proteinspezifischen Extinktionskoeffizienten (ExPASy ProtParam tool, Gasteiger et al., 2005) wurde die Molarität der Proteine bestimmt. Falls nötig, wurden die Proteine mittels Amicon Säulen (Amicon® Ultra Centrifugal Filter Devices 3K, 10K, 30K; Merck Millipore Ltd., Tullagreen, Carrigtwohill, Co. Cork, Ireland) einer bestimmten Ausschlussgröße durch Zentrifugieren aufkonzentriert. Proteinaliquots wurden bis zur Weiterverwendung bei -80°C gelagert.

10xPBS	
1,36 M	NaCl
26,82 mM	KCl
80,90 mM	Na_2HPO_4
17,53 mM	KH_2PO_4
pH auf 7,4 bzw. pH 7,5 eingestellt mit NaOH	

2.6.4.9 Kovalente Interaktionsanalysen von Catcher- und Tag-Proteinen (*In vitro* Rekonstitutionsexperimente)

Um die Bildung von kovalenten Komplexen basierend auf Isopeptidbindungen zwischen Catcher und Tag zu untersuchen, wurden immer rekombinant in *E. coli* M15 [pREP4] Zellen exprimierte (2.6.4.6), über Ni-NTA Affinitätschromatographie aufgereinigte (2.6.4.7) und in 1xPBS pH 7,4 dialysierte (2.6.4.8) Proteine einer bestimmten Molarität (M) verwendet. Die zu untersuchenden Proteinpartner, Catcher und Tag-MBP, wurden in einem bestimmten Verhältnis gemischt, bei 25°C und in 1xPBS pH 7,5 gemeinsam inkubiert und die Bildung von kovalenten Catcher-Tag Komplexen über die Zeit mittels gezielter Probennahme analysiert. Die Proben der Interaktionsanalysen wurden zum Abstoppen der Reaktion in denaturierendem 4x Lämmli-Probenpuffer aufgenommen und für 10 min bei 95°C gekocht, um den kovalenten Charakter der Interaktion der Proteine verifizieren zu können, bevor sie

auf ein 12%-iges Bis-Tris Gel zur SDS-PAGE mit anschließender Coomassie-Brilliant Blue Färbung aufgetragen wurden. Für die Untersuchung der biochemischen Eigenschaften und zur Bestimmung der idealen Reaktionsbedingungen der kovalenten Komplexbildung zwischen Catcher und Tag wurden die Proteininteraktionsstudien unter verschiedenen Bedingungen (pH-Wert, Temperatur, verschiedene Zusätze, verschiedene Verhältnisse der Reaktionspartner) durchgeführt. Vergleichende Analysen, bei denen bis auf den zu untersuchenden Parameter die Reaktionsbedingungen konstant gehalten wurden, wurden durchgeführt. Die zu quantifizierenden Proben wurden in Replikaten angefertigt, um diese anschließend densitometrisch auswerten zu können (siehe 2.6.4.4). Versuche zur Ermittlung des pH-Optimums wurden in 1xSPG-Puffer (Veggiani et al., 2016) verschiedener pH-Werte durchgeführt, während alle anderen Rekonstitutionsexperimente mit 1xPBS pH 7,4 - 7,5 als Reaktionspuffer durchgeführt wurden. Als Maß für die Effizienz der kovalenten Catcher-Tag Reaktion wurden sowohl die Zunahme der Catcher-Tag Produktbande im Vergleich zur 0h-Kontrolle (keine Interaktion) als auch die Abnahme der Catcher- und Tag-Eduktbanden (entweder Input-Banden als Referenz (100%) oder 0h-Kontrollen) herangezogen.

10xSPG	
125 mM	Bernsteinsäure/ Succinylsäure
437,5 mM	NaH_2PO_4
437,5 mM	Glycin
Gewünschter pH (4 - 9) wurde mit NaOH eingestellt	

2.7 Mikroskopische Methoden

2.7.1 Konfokale Laserscanning Mikroskopie (KLSM)

Für konfokale Analysen wurde ein Leica TCS SP5 II (AOBS) inverses Laserscanning Mikroskop (KLSM; Leica Mikrosysteme Vertrieb GmbH, Wetzlar) verwendet. Die Anregung des grün-fluoreszierenden Proteins (eGFP = enhanced GFP) erfolgte unter Verwendung eines Argon-Lasers (Blau-Anregung) bei 488 nm, während das rot-fluoreszierende Protein mCherry über einen DPSS-Laser (Hellgrün-Anregung) bei 561 nm angeregt wurde. eGFP wurde in einem Detektionsbereich von 496 nm bis 560 nm aufgenommen, mCherry von 580 nm bis 650 nm und die Chlorophyll-Autofluoreszenz von 680 nm bis 780 nm. Die konfokalen Datensätze wurden mit dem Programm Omero web client bearbeitet.

Magnetische Partikel, die mit drei verschiedenen Fluoreszenzproteinen gleichzeitig beladen waren (mTagBFP, eGFP und mCherry), wurden mittels des konfokalen Laserscanning Mikroskops im *sequential scan*-Modus mit folgenden Einstellungen untersucht: Das blau-fluoreszierende mTagBFP wurde mittels UV-Laser 405 (405 Diode) angeregt und bei 415 nm-490 nm detektiert, eGFP und mCherry wurden wie oben beschrieben angeregt und detektiert.

2.7.2 Fluoreszenzmikroskopie

Einfache fluoreszenzmikroskopische Analysen, wie die Detektion von eGFP im GFP-Kanal (Blau-Anregung) und die Aufnahme der Chlorophyll-Autofluoreszenz der Chloroplasten (mCherry-Kanal: Grün-Anregung oder *long path* Kanal: Blau-Anregung), sowie Durchlichtaufnahmen wurden am Leica DMR-Mikroskop (Leica Mikrosysteme Vertrieb GmbH, Bensheim) durchgeführt.

2.8 Chloroplasten-Isolierung über SpyCatcher-beladene Magnetpartikel (Chloroplasten-Trap)

2.8.1 Beladung von Maleimid-*beads* mit Cystein-SpyCatcher-Protein

Die Beladung der magnetischen Maleimid-*beads* verschiedener Größen mit Cystein-SpyCatcher-Protein erfolgte über eine stabile, irreversible Thioether-Bindung der freien Thiolgruppe des artifiziellen N-terminalen Cysteinrestes des SpyCatcher-Proteins mit der Maleimid-Gruppe auf der Oberfläche der Magnetpartikel (*Maleimide-surface activated beads*) (siehe Abb.2.5). Das Cystein-SpyCatcher-Konstrukt (#147 siehe Tabelle 1.1) wurde extra kloniert (siehe 2.6.2), da das native SpyCatcher-Protein keine Cysteine enthält.

Maleimid aktivierte beads	Größe	Firma
PureCube Maleimide-Activated MagBeads	25 µm	Cube Biotech GmbH, Monheim, Deutschland
Mono Mag Maleimide Activated Magnetic Beads	0,5 µm	Ocean NanoTech, San Diego, CA 92126 US
Mono Mag Maleimide Activated Magnetic Beads	4,5 µm	Ocean NanoTech, San Diego, CA 92126 US

Die kommerziell erworbenen Maleimid-Magnetpartikel wurden in 1xPBS pH 7,5 äquilibriert und rekombinant in *E. coli* hergestelltes, aufgereinigtes (+ 10 mM EDTA im Elutionspuffer, 20 mM β-Mercaptoethanol im Waschpuffer) und in 1xPBS pH 7,4 + 0,5 mM EDTA dialysiertes Cystein-SpyCatcher Protein (#147 in Tabelle 1.1) im Überschuss zu den Maleimid-Partikeln gegeben (Cube Biotech Angaben als Orientierung https://cube-biotech.com/files/protocols/Coupling%20Procedure%20%20Maleimide%20MagBeads.pdf: 2 - 6 mg/ml Proteinlösung zu 1 ml *bead slurry*; anstatt *coupling buffer* I wurde 1xPBS pH 7,4 verwendet) und über Nacht bei 4°C oder für 2 bis 3 Stunden bei Raumtemperatur unter Schütteln inkubiert. Nicht gebundenes Cystein-SpyCatcher Protein wurde abgenommen, die Proteinkonzentration am NanoDrop bei 280 nm bestimmt und die *beads* dreimal mit je 1 ml 1xPBS pH 7.5 gewaschen. Beim Arbeiten mit magnetischen Partikeln, vor allem zum Separieren der *beads* von der restlichen Flüssigkeit (Überstand), wurde immer ein Magnetständer (DynaMagTM-2, Life Technologies AS, Oslo) verwendet. Um unspezifische Proteinbindestellen auf der *bead*-Oberfläche abzusättigen, wurden bei Bedarf die Magnetpartikel nach Cystein-SpyCatcher-Beladung mit 5% Magermilchpulver in 1xPBS pH 7,5 über Nacht blockiert und am Folgetag dreimal mit je 1 ml 1xPBS pH 7,4 gewaschen. Zur Überprüfung der SpyCatcher-Beladung wurden die *beads* mit SpyTag-eGFP-Protein, das rekombinant in *E. coli* hergestellt, aufgereinigt und in 1xPBS pH 7,4 dialysiert wurde, für mindestens 30 min bei 4°C im Überkopfschüttler inkubiert. Nach der Inkubation wurde der Überstand abgenommen und die Partikel dreimal mit je 1 ml 1xPBS pH 7,4 gewaschen. Die Magnetpartikel wurden anschließend in 1xPBS aufgenommen und ein Aliquot unter dem Fluoreszenzmikroskop (Leica) im GFP-Kanal (Anregung mit blauem Licht) betrachtet.

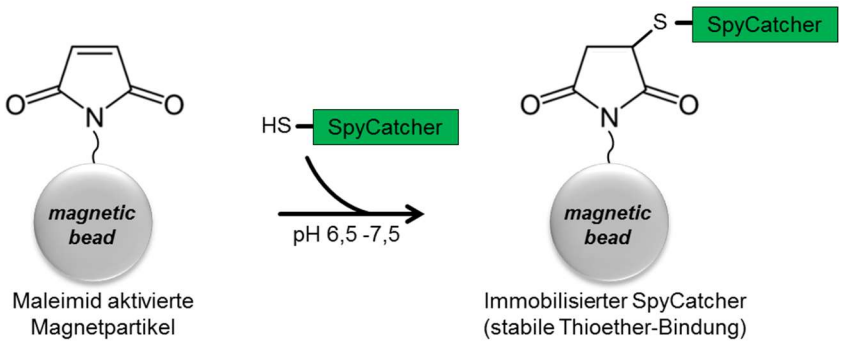


Abb.2.5: Schema der Beladung von Maleimid-aktivierten Magnetpartikeln mit Cystein-SpyCatcher-Protein. Unter pH-Bedingungen zwischen 6,5 und 7,5 bildet das Maleimid auf der *bead*-Oberfläche mit freien Sulfhydryl-Gruppen (Thiolen) stabile, kovalente, irreversible Thioether aus.

2.8.2 Isolierung von Chloroplasten aus Gesamtblattfiltrat

Frisch geerntete *Nicotiana benthamiana* Blattproben (4 Blattscheiben Korkbohrer #18, ca. 0,5 g Frischgewicht), die 48h zuvor über Agrobakterien-vermittelte transiente Transformation mit dem Chloroplasten-*targeting* Zielkonstrukt OEP7-SpyTag (#167 in Tabelle 1.1) infiltriert wurden, wurden frisch, das heißt ohne Einfrieren, in 2,5 ml eiskaltem Isolierpuffer gemörsert. Das homogenisierte Blattlysat wurde anschließend durch Miracloth (EMD Millipore Corporation, Billerica, MA 01821 USA) filtriert und das Filtrat auf mit Cystein-SpyCatcher beladene magnetische Partikel (4,5 µm *Maleimide-activated beads*) ungefähr im Verhältnis 4 zu 1 (*beads* aus 50 µl *bead*-Lösung + 200 µl Blattfiltrat) gegeben und bei 4°C unter leichtem Schütteln für mindestens 30 min inkubiert. Nach der Inkubation der SpyCatcher-*beads* mit dem Blattlysat wurden die *beads* mittels Magnetständer vom Überstand getrennt, der Überstand abgenommen und die *beads* anschließend dreimal mit je 1 ml Isolierpuffer gewaschen. Es folgte die visuelle Analyse der Chloroplasten-Isolierung über SpyCatcher-*beads* mittels Fluoreszenzmikroskopie (Leica) und konfokaler Mikroskopie (KLSM).

Chloroplasten-Isolierpuffer	
10 mM	Hepes pH 7,0 (KOH)
0,33 M	Sorbitol
0,4 mM	KCl
40 µM	EDTA
1%	PVP
0,1%	BSA
Frisch zugeben	
1 mM	Pefabloc Protease Inhibitor (Roth)
Complete ULTRA Tablets EDTA-free (Roche)	
2 mM	DTT

2.9 Software und Webserver

Für Molekülvisualisierungen dreidimensionaler Proteinstrukturen wurde das Programm Open-Source PyMOL™ 1.8.x verwendet (PyMOL (TM) Molecular Graphics System, Version 1.8.6.1. Copyright (c) Schrodinger, LLC., Created by Warren L. DeLano, Ph.D., http://www.pymol.org). Unter Zuhilfenahme des Rosetta Backrub Servers (Lauck et al., 2010, Smith und Kortemme, 2008, 2010, 2011) wurden Aminosäuresubstitutionen geplant (Sequence Tolerance). Der PROSS-Webserver (Protein Repair One Stop Shop: Goldenzweig et al., 2016) wurde verwendet, um stabilisierte Proteinvarianten durch gezielte Aminosäureaustausche vorherzusagen. Die Torsionswinkel phi und psi eines Proteins wurden mithilfe des Internetprogramms MolProbity ermittelt (Chen et al., 2010, Davis et al., 2007, Lovell et al., 2003). Für alle diese Webserver, die beim rationalen Mutagenese-Design helfen sollten, wurden die pdb-Einträge (Kristallstrukturen) der Proteindatenbank (PDB: Berman et al., 2000) als Grundlage und Inputs verwendet.

3 Ergebnisse

3.1 Identifizierung neuer Split-Isopeptid Systeme

Um neue Split-Isopeptid Systeme als Alternative zu dem Spy- (Zakeri et al., 2012) und dem Snoop-System (Veggiani et al., 2016) zu finden, wurde eine Kombination von computerbasierten und experimentellen Methoden verwendet. Die *in silico* Analysen wurden von Prof. Dr. Heinrich Sticht und Dr. Anselm Horn durchgeführt. Die experimentelle Überprüfung der Vorhersagen war Aufgabe der vorliegenden Arbeit.

3.1.1 Bioinformatische Analysen (*in silico* Analysen)

3.1.1.1 Strukturbasierte Datenbanksuche

Zur Identifizierung von alternativen Split-Isopeptid Systemen wurde zunächst eine strukturbasierte Datenbanksuche innerhalb der Proteindatenbank (DALI-Analyse, PDB90 mit dem pdb-Eintrag 2x5p als Suchanfrage) durchgeführt (Pröschel et al., 2017). Der Fokus lag auf der Struktur, da Isopeptidbindungen aufgrund fehlender Sequenzmotive durch reine Sequenzanalysen fast unmöglich zu identifizieren sind (Kang und Baker, 2011, Kang et al., 2007). Bei der DALI-Analyse (Holm et al., 2008, Holm und Rosenström, 2010) wurde nach Strukturen mit ähnlicher dreidimensionaler Faltung (Topologie) wie die der CnaB2-Domäne des Spy-Systems gesucht. Weitere Suchkriterien waren eine möglichst geringe Sequenzidentität sowohl der Kandidaten untereinander als auch der Kandidaten mit dem Spy-System (<30%) und das Vorhandensein einer, in der Proteindatenbank hinterlegten Kristallstruktur mit verifizierter Isopeptid-Triade (Pröschel et al., 2017). Die Datenbanksuche konzentrierte sich dabei auf Oberflächenproteine (v.a. Pilus-Proteine) gram-positiver Bakterien, die Faltungsdomänen des CnaB-Typs (siehe Kapitel 1.1.2.2) besitzen, da das bis dahin potenteste kovalente Split-Isopeptid System, das Spy-System, und auch alle weiteren bekannten Catcher-Tag Systeme, auf solchen CnaB-Domänen basieren.

3.1.1.2 Selektion der Kandidatendomänen

An die Datenbanksuche schloss sich eine genauere, lokale Untersuchung der Proteinstrukturen an, die hauptsächlich von Dr. Anselm Horn durchgeführt wurde. Bei der Strukturanalyse der Proteindomänen wurden u.a. die Domänenstabilität (geringer Kontakt der Kandidatendomäne zu Nachbardomänen) und potentielle Auswirkungen der Spaltung der Domänen in einen Catcher- und einen Tag-Teil (kurze Verbindungssequenz zwischen Tag- und Catcher-Region, Stabilität der Spaltungsprodukte) berücksichtigt. Schließlich

ergaben die Datenbanksuche und die anschließende lokale Proteinstrukturanalyse der Kandidatendomänen vier potentielle Split-Isopeptid Systeme, die näher analysiert wurden.

Die vier Kandidatensysteme 3phs, 4oq1, 3kptN und 3kptC, benannt nach den pdb-Einträgen, leiten sich von vier verschiedenen CnaB-Domänen (N2, D2, CNA2, CNA3) aus drei verschiedenen Pilus-Proteinen (GBS52, RrgC, BcpA) von gram-positiven Bakterien ab (siehe Abb.3.1).

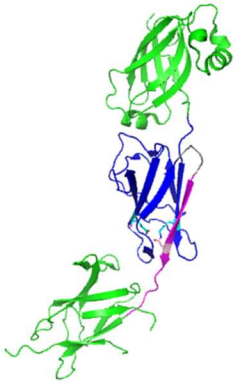
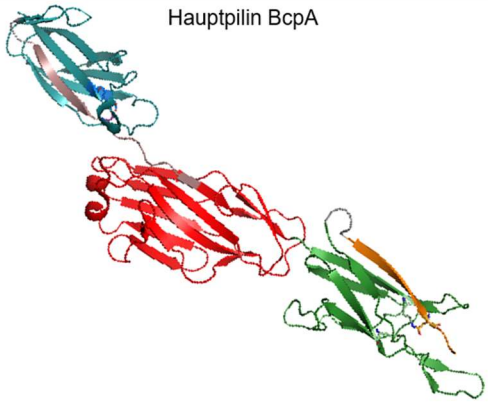
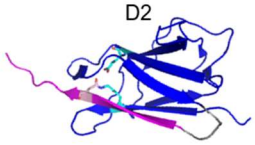
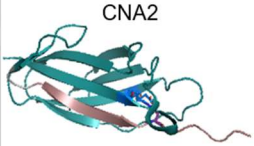
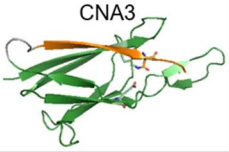
System	**3phs**	**4oq1**	**3kptN**	**3kptC**
Organismus	*Streptococcus agalactiae*	*Streptococcus pneumoniae*	*Bacillus cereus*	*Bacillus cereus*
Protein	Pilin GBS52	Basales Pilin RrgC 	Hauptpilin BcpA 	
Domäne	N2	D2 	CNA2 	CNA3 
PDB code	3PHS	4OQ1	3KPT	3KPT
Aminosäuren Catcher / Tag	152-245 / 248-263	146-240 / 246-259	164-255 / 258-272	409-500 / 505-515+PTK
Isopeptid-Triade	Glu226 Lys160 Asn256	Glu222 Lys155 Asn252	Glu223 Lys174 Asn265	Glu472 Lys417 Asn512

Abb.3.1: Zusammenstellung der vier Kandidatensysteme zur Generierung neuer Split-Isopeptid Systeme (modifiziert nach Pröschel et al., 2017). Die Molekülvisualisierungen wurden mittels PyMOL Open Source basierend auf den pdb-Einträgen (Kristallstrukturen) angefertigt. Gezeigt sind die Proteine (Piline) der gram-positiven Bakterien, deren Bestandteil die einzelnen Kandidatendomänen (N2 des 3phs-Systems, D2 des 4oq1-Systems, CNA2 des 3kptN-Systems und CNA3 des 3kptC-Systems), innerhalb der die intramolekularen Isopeptidbindungen stattfinden (Isopeptid-Triade), bilden.

3.1.1.2.1 3phs-System

Das 3phs-System leitet sich von der N2-Domäne des Pilus-Proteins GBS52 (GBS: group B *streptococcus*) aus dem gram-positiven Bakterium *Streptococcus agalactiae* ab (siehe Abb.3.1). Innerhalb dieser N2-Domäne kommt es zur Ausbildung einer intramolekularen Isopeptidbindung (Kang et al., 2007, Kang und Baker, 2011). Die Isopeptid-Triade setzt sich aus den Aminosäuren Lysin-160, Asparagin-256 und Glutaminsäure-226 zusammen. Das Pilin GBS52 (*minor pilin*) ist neben dem Pilin GBS104 (*minor pilin*) und dem Hauptpilin GBS80 Teil der Pilus-Struktur (Krishnan et al., 2007; Rosini et al., 2006). Ob es sich bei dem Pilin GBS52 um ein Adhäsin an der Spitze des Pilus handelt (Krishnan et al., 2007), dessen Funktion die spezifische Erkennung und Adhäsion der Bakterien an Wirtszellen ist, oder ob GBS52 ein basales Pilin ist (Linke et al., 2010, Kang und Baker, 2011, Kang und Baker, 2012), welches an der Verankerung des Pilus in der bakteriellen Zellwand beteiligt ist, ist nicht ganz geklärt. Das GBS52 Pilin selbst besteht aus zwei IgG-ähnlichen Faltungsdomänen des CnaB-Typs, N1 und N2, die über eine kurze Linker-Sequenz miteinander verbunden sind (siehe Abb.3.1) (Krishnan et al., 2007).

3.1.1.2.2 4oq1-System

Das 4oq1-System basiert auf der D2-Domäne des basalen Pilus-Proteins RrgC von *Streptococcus pneumoniae* (siehe Abb.3.1). Der Pneumokokken-Pilus setzt sich aus drei Strukturproteinen, RrgA, RrgB und RrgC, zusammen, die essentiell für die funktionale Pilus-Ausbildung und Oberflächen-Adhäsion sind. RrgA ist das Adhäsin, das die Spitze des Pilus bildet (*minor tip pilin*), während RrgB das Grundgerüst des Pilus ausmacht (*major backbone pilin*). RrgC ist das basale Pilin (*minor base pilin*), das den Pilus an der bakteriellen Zellwand über die Aktivität der Sortase A verankert (Shaik et al., 2014; LeMieux et al., 2006, Nelson et al., 2007, El Mortaji et al., 2010, 2012). Während die Strukturen von RrgA und RrgB schon länger bekannt sind, wurde die Kristallstruktur von RrgC erst 2014 von Shaik et al. aufgeklärt. Das RrgC Protein besteht aus drei unabhängigen Domänen, D1, D2 und D3 (Shaik et al., 2014). Diese drei Domänen falten sich zwar unabhängig voneinander, leiten sich aber von einer Kerndomäne (β-*barrel* Struktur) ab, die sich durch kleine Veränderungen der Sekundärstrukturelemente zwischen den Domänen unterscheidet. Die Aufklärung der Kristallstruktur hat ergeben, dass die Domänen D2 und D3 des RrgC Proteins in anti-parallele β-Faltblätter falten (IgG-reverse fold) und jeweils durch eine intramolekulare Isopeptidbindung, die den ersten mit dem letzten β-Strang des β-Faltblattes verknüpft, stabilisiert werden (CnaB-Typ, siehe 1.1.2.2). Die intramolekulare Isopeptidbindung in der D2-Domäne bildet sich zwischen den Seitenketten der Reste Lysin-155 und Asparagin-252, wobei Glutaminsäure-222 die Reaktion katalysiert (Shaik et al., 2014).

3.1.1.2.3 3kptN- und 3kptC-System

Das 3kptC-System hat seinen Ursprung in der CNA3-Domäne des Hauptpilins BcpA aus dem gram-positiven Bakterium *Bacillus cereus*, während das 3kptN-System auf der CNA2-Domäne von BcpA basiert (siehe Abb.3.1). Das BcpA Protein bildet das Grundgerüst des Pilus und ist zudem verantwortlich für die Verankerung des Pilus mit der bakteriellen Zellwand, während die Adhäsion an das Wirtsgewebe über das Pilin/ Adhäsin BcpB an der Spitze des Pilus von *B. cereus* stattfindet. Die Aufklärung der Kristallstruktur ergab, dass das rekombinant hergestellte BcpA Protein aus den drei Domänen – CNA2, XNA und CNA3 - besteht, die alle β-Faltblatt-Strukturen annehmen. Bei den Domänen CNA2 und CNA3 handelt es sich um reverse Ig-Faltungen (CnaB-Typ), die jeweils eine intramolekulare Isopeptidbindung aufweisen (Budzik et al., 2009). Die Isopeptidbindung in der CNA2-Domäne bildet sich zwischen Lysin-174 und Asparagin-265 und wird durch Glutaminsäure-223 katalysiert. Die Isopeptid-Triade in der CNA3-Domäne besteht aus Lysin-417, Asparagin-512 und Glutaminsäure-472 (Budzik et al., 2009).

3.1.1.3 Domänenspaltung und Validierung

Nach der Kandidatenselektion wurden geeignete Spaltungsstellen der ausgewählten CnaB-Domänen anhand von Strukturanalysen und unter Berücksichtigung der reaktiven Triade vorgeschlagen (siehe Abb.3.1). Die initiale, rationale Spaltung der CnaB-Domänen in ein Catcher-Protein (kompakte, globuläre Kerndomäne) und ein Tag-Peptid (ein β-Strang) wurde ebenfalls von Prof. Dr. Heinrich Sticht und Dr. Anselm Horn übernommen. Im Anschluss daran wurden von Lena Schäfer im Rahmen ihrer Masterarbeit Moleküldynamik-Simulationen durchgeführt, um die Catcher-Tag Komplexe genauer zu untersuchen und sie bezüglich ihrer Stabilität mit den zugrundeliegenden Domänen zu vergleichen. Als wichtige Unterscheidungsparameter wurden dabei die Van-der-Waals-Wechselwirkungen und die Wasserstoffbrückenbindungen zwischen Catcher- und Tag-Komponenten verwendet. Eine genaue Analyse der Simulationen empfahl die weitere experimentelle Untersuchung aller vier Kandidatensysteme 3phs, 4oq1, 3kptN und 3kptC (Pröschel et al., 2017) (siehe Abb.3.1).

Die Vorgehensweise der computergestützten Kandidatensuche, Selektion, Spaltung und Validierung bis hin zur tatsächlichen experimentellen *in vitro* Analyse der neuen, potentiellen Split-Isopeptid bzw. Catcher-Tag Systeme ist schematisch in Abbildung 3.2 dargestellt.

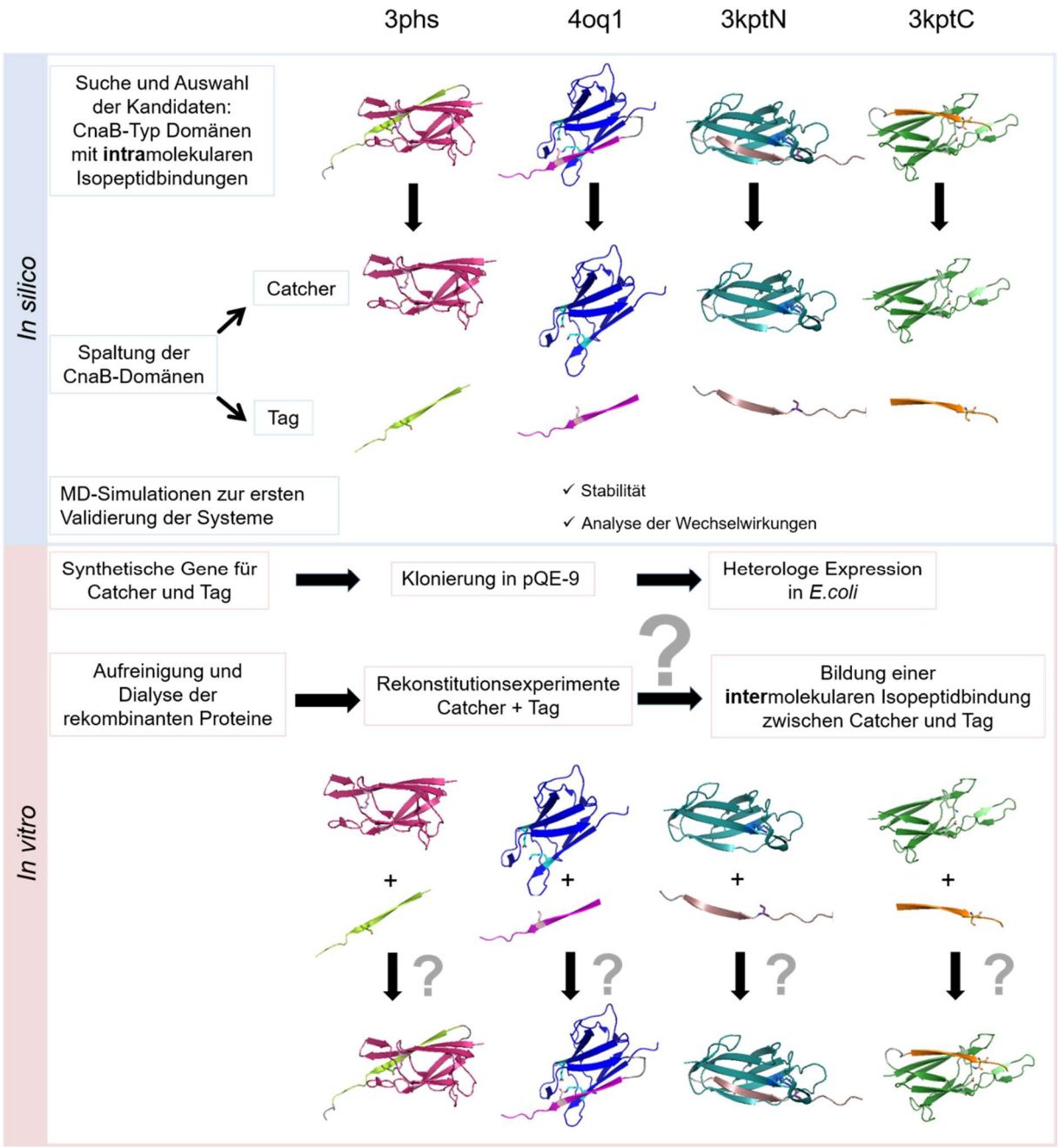


Abb.3.2: Zusammenfassendes Schema der Vorgehensweise zur Identifizierung neuer, kovalenter Split-Isopeptid Systeme. Mit der CnaB2-Domäne des Spy-Systems als Suchgrundlage wurden mittels bioinformatischer, strukturbasierter Analysen die vier potentesten CnaB-Domänen aus drei verschiedenen Pilus-Proteinen als mögliche Split-Isopeptid Kandidaten gewählt. Die vier ausgewählten CnaB-Domänen, innerhalb der sich je eine intramolekulare Isopeptidbindung zwischen reaktiven Seitenketten von Lysin und Asparagin ausbildet, wurden basierend auf bioinformatischen Analysen in zwei Teile, Catcher-Protein und Tag-Peptid, gespalten. Die potentiellen Catcher-Tag Systeme (3phs, 4oq1, 3kptN und 3kptC) wurden anschließend mittels Moleküldynamik (MD)-Simulationen genauer untersucht, um die Stabilität und die molekularen Interaktionen zwischen Catcher und Tag zu ermitteln. Nach erfolgreicher *in silico* Analyse wurden die Gene, die für Catcher bzw. Tag kodieren synthetisch hergestellt, in den *E. coli* Expressionsvektor pQE-9 kloniert, die rekombinanten Proteine heterolog in *E. coli* exprimiert und über Ni-NTA Affinitätschromatographie aufgereinigt und dialysiert. In *in vitro* Split-Domänen Rekonstitutionsexperimenten sollte nun geklärt werden, ob die Catcher-Proteine mit ihren spezifischen Tag-Peptiden einen kovalenten Komplex bilden, wenn sie gemischt werden, das heißt die Fähigkeit zur Ausbildung einer Isopeptidbindung – diesmal einer intermolekularen - trotz Spaltung beibehalten haben oder nicht.

3.1.2 Catcher- und Tag-DNA-Konstrukte

Nach erfolgreicher *in silico* Analyse wurden die DNA-Sequenzen, die für die zu untersuchenden Tag-Peptide bzw. Catcher-Proteine kodieren, ermittelt. Dafür wurden die Proteinsequenzen der Catcher- und Tag-Komponenten der vier Split-Isopeptid Kandidatensysteme (3phs, 3kptC, 3kptN, 4oq1: siehe Abb.3.1 und 3.2) aus der strukturbasierten Datenbanksuche mithilfe des Programms tblastn (translated BLAST, NCBI: National Center for Biotechnology Information, U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda MD, 20894 USA) analysiert. Die besten Treffer der BLAST-Suche (Standardparameter) wurden dann in einem ,,Alignment“ (GeneWise, European Bioinformatics Institute, EMBL-EBI) zwischen der ursprünglichen Input-Proteinsequenz aus der Datenbanksuche und der bei der BLAST-Suche gefundenen Nukleotidsequenz nochmals verifiziert und synthetische Genkonstrukte unter Beachtung der *E. coli Codon usage* unter Rücksprache mit Prof. Dr. Heinrich Sticht bestellt (siehe Kapitel 2.4). Die initialen Catcher-Sequenzen wurden mit einem Hämagglutinin-Tag (HA) versehen und die Tag-Sequenzen wurden stets als C-terminale Maltosebindeprotein (MBP)-Fusion (malE Gen aus *E. coli* ohne Signalpeptid) generiert (siehe Abb.3.3, Sequenzen der DNA-Konstrukte #52 - #59 siehe Anhang Tabelle A1). Für die heterologe Proteinexpression, Ni-NTA Affinitätsaufreinigung und anschließende *in vitro* Rekonstitutionsexperimente wurden die Catcher- und Tag-DNA-Konstrukte über BamHI und SalI Schnittstellen gezielt in den bakteriellen Expressionsvektor pQE-9 kloniert (siehe Kapitel 2.6.2). Diese Plasmide stellten die Grundlage für alle weiteren DNA-Konstrukte der neu identifizierten Catcher-Tag Systeme dar. Über sequenzspezifische Mutagenese-Primer (siehe Anhang Tabelle A2) wurden diese DNA-Konstrukte über PCR-Methoden gezielt genetisch manipuliert, sodass Nukleotide substituiert, deletiert und inseriert wurden (siehe Kapitel 2.6.2 und Tabelle 1.1).

Abb.3.3: Schematische Darstellung der initialen Catcher- und Tag-DNA-Konstrukte. Die zugrundeliegenden DNA-Sequenzen (#52 bis #59 siehe Tabelle 1.1, Sequenzen im Anhang Tabelle A1), die für die (A) Catcher-Proteine bzw. (B) Tag-Peptide der vier Split-Isopeptid Kandidatensysteme aus der *in silico* Analyse kodieren, wurden über *translated* BLAST-Analysen mit Aminosäuresequenzen als Input ermittelt, Codon optimiert auf die Expression in *E. coli* Zellen und als synthetische DNA-Fragmente bestellt. Die Klonierung in den bakteriellen Expressionsvektor pQE-9 erfolgte über kompatible Enden, erzeugt durch BamHI/SalI Verdaus sowohl der DNA-Inserts als auch des Zielvektors (siehe Kapitel 2.6.2). Der GS-linker besteht aus der Aminosäureabfolge GSGSGSG.

3.1.3 Heterologe Expression der Catcher- und Tag-Konstrukte in *E. coli* M15 [pREP4] Zellen

Zuerst wurden Expressionsanalysen der Catcher- und Tag-MBP Konstrukte (#52 bis #59 siehe Tabelle 1.1 und Tabelle A1) durchgeführt, um zu sehen, ob die Proteine in dem heterologen Wirt *E. coli* M15 [pREP4] auch hergestellt werden (siehe Kapitel 2.6.4.6). Alle Tag-MBP Konstrukte wurden erfolgreich in *E. coli* M15 [pREP4] Zellen bei 28°C unter Schütteln nach Induktion mit 1 mM IPTG-Endkonzentration überexprimiert (siehe Abb.3.4), wohingegen lediglich der 4oq1Catcher und nach einigen Optimierungen der Expressionsbedingungen (Hauptkultur auf oD_{600} ~0,02 animpfen und Induktion mit 0,5 mM IPTG als Endkonzentration) auch der 3kptCCatcher exprimiert wurden (siehe Abb.3.4 B und C). Weder der 3phsCatcher noch der 3kptNCatcher konnten rekombinant in *E. coli* M15 [pREP4] Zellen hergestellt werden (siehe Abb.3.4 A und D).

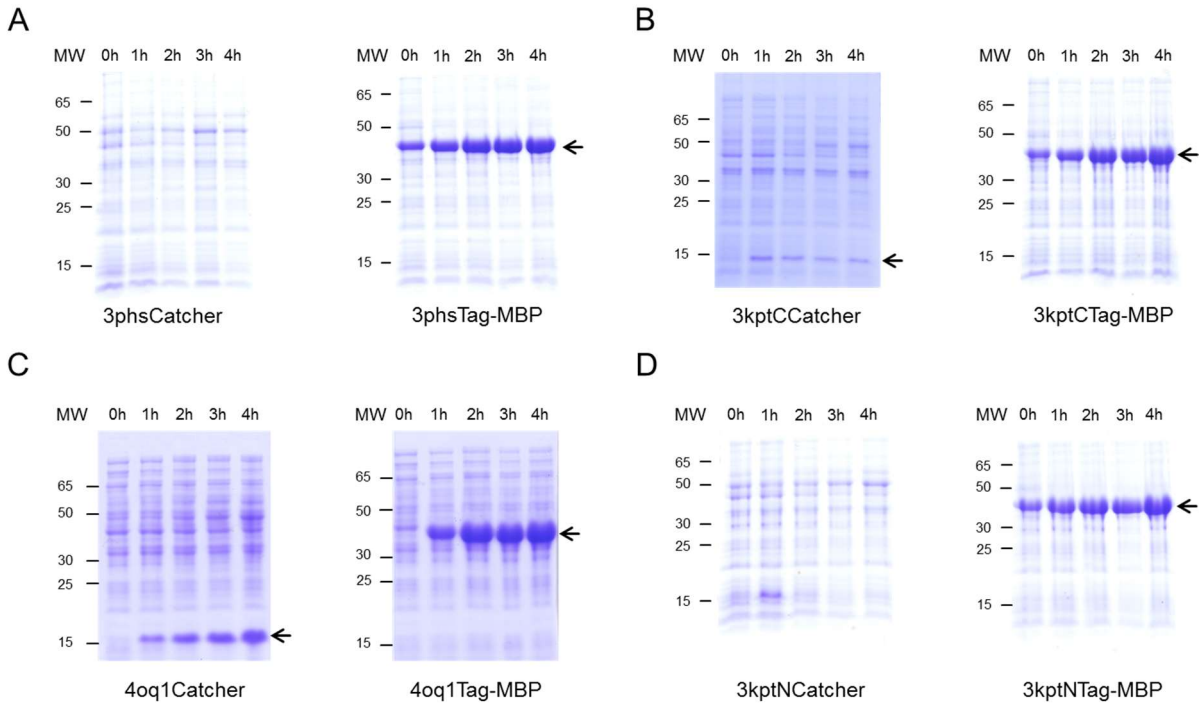


Abb.3.4: Expressionskinetiken der initialen Catcher- bzw. Tag-MBP Proteine (#52 - #59) in M15 [pREP4] *E. coli* Zellen nach IPTG-Induktion. Die heterologe Expression der rekombinanten Proteine wurde bei 28°C für 4 Stunden unter ständigem Schütteln (180 - 200 rpm) durchgeführt. Die Proteinexpression wurde ab einer oD_{600} von ca. 0,5 mit Zugabe von 1mM bzw. 0,5mM IPTG-Endkonzentration induziert. Die Expression der rekombinanten Proteine wurde mittels SDS-PAGE und anschließender Coomassie-Färbung über die Zeit analysiert. Die Produktbanden der rekombinanten Proteine sind mit einem Pfeil markiert. (A) Heterologe Expression von 3phsCatcher und 3phsTag-MBP, (B) heterologe Expression von 3kptCCatcher und 3kptCTag-MBP, (C) heterologe Expression von 4oq1Catcher und 4oq1Tag-MBP, (D) heterologe Expression von 3kptNCatcher und 3kptNTag-MBP. Alle Catcher-Proteine sind ca. 13 kDa und alle Tag-MBP-Proteine ca. 44 kDa groß. MW steht für *molecular weight* (Angaben in kDa).

Aufgrund der schlechten bzw. fehlenden Expression der initialen Catcher-Konstrukte wurden mCherry-GS-Catcher Genfusionen über overlap-PCR generiert (Klonierung siehe Kapitel 2.6.2 und Tabelle 1.1: #94, #95, #96 und #115, Primersequenzen siehe Anhang Tabelle A2) und deren Expression in M15 [pREP4] *E. coli* Zellen überprüft (siehe Abb.3.5). Das rote Fluoreszenzprotein mCherry wurde als Fusionspartner gewählt, da es in vorherigen Expressionsanalysen immer zu einer großen Menge an löslichem Protein geführt hat und die erfolgreiche Expression nach Induktion leicht durch Rotfärbung der Kultur mit der Zeit visuell nachweisbar ist. Die mCherry-Catcher Fusionsproteine wurden erfolgreich in M15 [pREP4] *E. coli* Zellen bei 28°C unter ständigem Schütteln und IPTG-Induktion (1 mM) exprimiert und standen nun für Interaktionsanalysen mit den entsprechenden Tag-MBP Konstrukten bereit (siehe Abb.3.5). Hier soll noch erwähnt werden, dass alle erfolgreich exprimierten, rekombinanten Proteine auch in löslicher Form über Ni-NTA Affinitätschromatographie (siehe Kapitel 2.6.4.7) aufgereinigt werden konnten (siehe Input-Spuren aller Interaktionsanalysen).

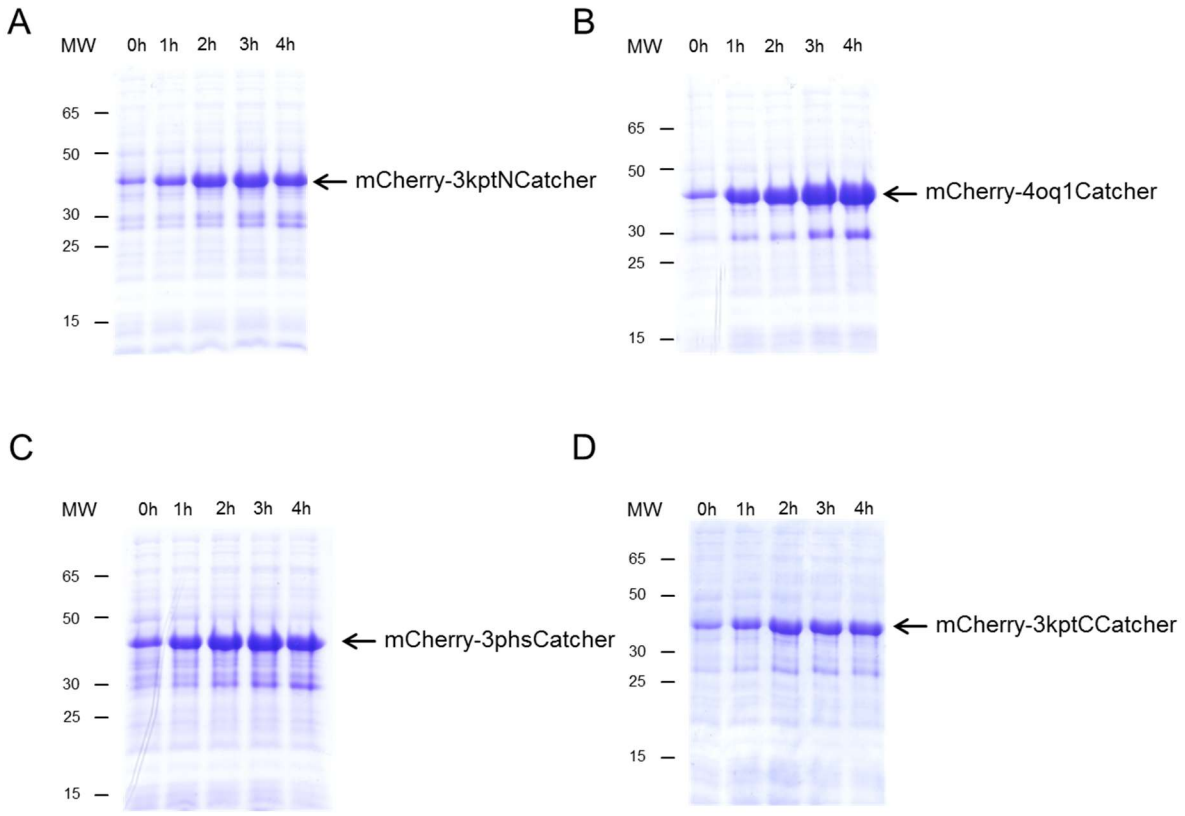


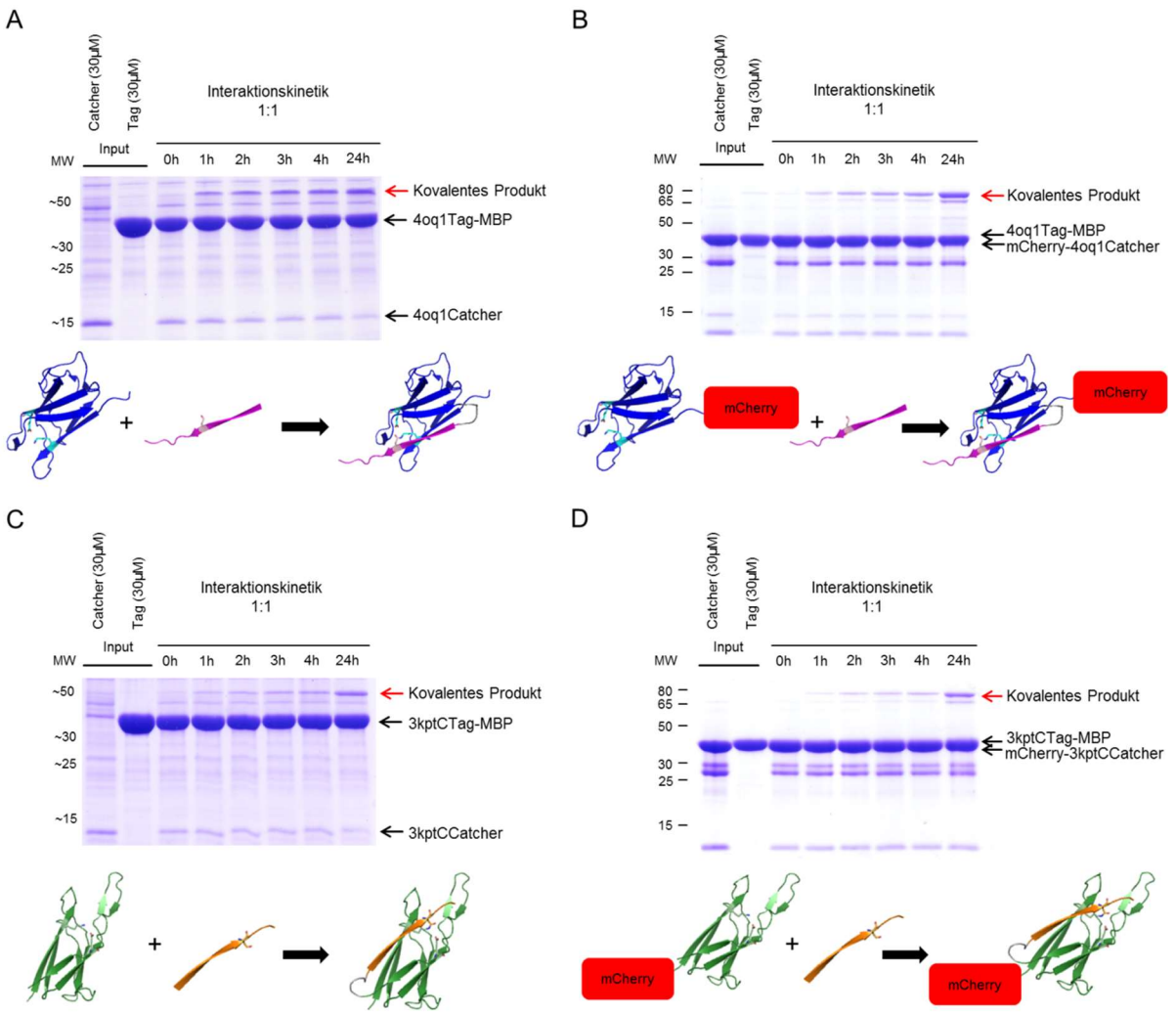
Abb.3.5: Expressionskinetiken der mCherry-Catcher Fusionsproteine (#94, #95, #96 und #115) in M15 [pREP4] *E. coli* Zellen nach IPTG-Induktion. Die heterologe Expression der rekombinanten Proteine wurde bei 28°C für 4 Stunden unter ständigem Schütteln (180 - 200 rpm) durchgeführt. Die Proteinexpression wurde ab einer oD_{600} von ca. 0,5 mit Zugabe von 1mM IPTG-Endkonzentration induziert. Die Expression der rekombinanten Proteine wurde mittels SDS-PAGE und anschließender Coomassie-Färbung über die Zeit analysiert. Die Produktbanden der rekombinanten Proteine sind mit einem Pfeil gekennzeichnet. (A) Heterologe Expression von mCherry-3kptNCatcher, (B) heterologe Expression von mCherry-4oq1Catcher, (C) heterologe Expression von mCherry-3phsCatcher, (D) heterologe Expression von mCherry-3kptCCatcher. Alle mCherry-Catcher Proteine sind ca. 39 kDa groß. MW steht für *molecular weight* (Angaben in kDa).

3.1.4 *In vitro* Analyse der Split-Domänen Rekonstitution basierend auf Isopeptidbindungen zwischen Catcher und Tag (*in vitro* Rekonstitutionsexperimente)

In *in vitro* Split-Domänen Rekonstitutionsexperimenten sollte nun geklärt werden, ob die Catcher-Proteine mit ihren spezifischen Tag-Peptiden einen kovalenten Komplex bilden, wenn sie gemischt werden, das heißt die Fähigkeit zur Ausbildung einer Isopeptidbindung trotz Spaltung beibehalten haben oder nicht (siehe Abb.3.2). Über die Beantwortung der Frage, welche der vier vorgeschlagenen Systeme tatsächlich nach der Spaltung in Catcher und Tag noch die Fähigkeit besitzen eine Isopeptidbindung auszubilden, sollte neben der tatsächlichen Identifzierung von neuen Split-Isopeptid Systemen auch das Potential der *in silico* Analyse zur Vorhersage von solchen Split-Domänen Systemen ermittelt werden (siehe Pröschel et al., 2017).

Für die kovalenten Interaktionsanalysen wurden Catcher- und Tag-Proteine verwendet, die rekombinant in *E. coli* M15 [pREP4] Zellen hergestellt (2.6.4.6), über Ni-NTA Affinitätschromatographie aufgereinigt (2.6.4.7) und in 1xPBS pH 7,4 dialysiert wurden (2.6.4.8). Nach Bestimmung der Proteinkonzentration (NanoDrop A280) wurden Catcher- und Tag-Proteine unter Berücksichtigung der proteinspezifischen Extinktionskoeffizienten auf dieselbe Molarität (30 µM als Input-Molarität) eingestellt und in äquimolarem Verhältnis (1:1 Verhältnis mit 15 µM Catcher- bzw. Tag-Endkonzentration) gemischt und bei 25°C inkubiert (siehe Kapitel 2.6.4.9). Die Rekonstitutionsexperimente ergaben, dass sowohl das 4oq1Catcher-4oq1Tag System als auch das 3kptCCatcher-3kptCTag System *in vitro* funktionell sind und sich ein kovalentes Produkt, das resistent gegenüber Kochen in SDS-haltigem 4x Lämmli-Puffer ist, bildet (siehe neu enstandene, höhermolekulare Banden in Abb.3.6, jeweils gekennzeichnet durch rote Pfeile). Die neu entstandene Produktbande nahm bei beiden Systemen über den gesamten Analysezeitraum (24h) stetig zu, wobei die korrelierende Abnahme der Catcher- bzw. Tag Input-Banden nur in den Rekonstitutionsexperimenten ohne mCherry-Fusion sichtbar war, da die mCherry-Catcher Fusionsproteine die gleiche Größe wie die Tag-MBP Konstrukte besitzen (siehe Abb.3.6 B, D, E, F). Die kovalente Interaktion zwischen 4oq1Catcher und 4oq1Tag bzw. zwischen 3kptCCatcher und 3kptCTag konnte jedoch eindeutig in beiden Situationen, das heißt sowohl mit dem Catcher alleine als auch mit N-terminaler mCherry-Fusion nachgewiesen werden (siehe Abb.3.6). In den Rekonstitutionsexperimenten mit den mCherry-Catcher Fusionsproteinen sind mehrere weitere Banden im Gel zu sehen, die vor allem Abbauprodukte des mCherry-Proteins darstellen (Pröschel et al., 2017). In massenspektrometrischen Analysen, durchgeführt von Max Kraner, wurde nochmals

verifiziert, dass es sich bei der neu entstandenen Bande im Proteingel tatsächlich um das Produkt aus Catcher und entsprechendem Tag handelt (Pröschel et al., 2017). Weder das 3phsCatcher-3phsTag System noch das 3kptNCatcher-3kptNTag System ergaben ähnliche Ergebnisse. Nach dem Mischen und Inkubieren von mCherry-3phsCatcher mit 3phsTag-MBP bzw. mCherry-3kptNCatcher mit 3kptNTag-MBP konnte keine neu entstandene, höhermolekulare Catcher-Tag Produktbande, die stabil gegenüber Kochen in 4x Lämmli-Puffer ist, in den Coomassie-Gelen detektiert werden (siehe Abb.3.7). Das bedeutet, dass weder das 3phs-System noch das 3kptN-System unter diesen Bedingungen und mit den initial gewählten Spaltungsstellen als Split-Domänen Systeme funktionieren. Folglich konnten aus zwei der vier CnaB-Kandidatendomänen, die über *in silico* Analysen entdeckt, ausgewählt, analysiert und gespalten wurden, funktionelle *in vitro* Split-Isopeptid Systeme entwickelt werden. Im weiteren Arbeitsverlauf wurden diese beiden neuen Catcher-Tag Systeme, 4oq1-System und 3kptC-System, weiter untersucht.



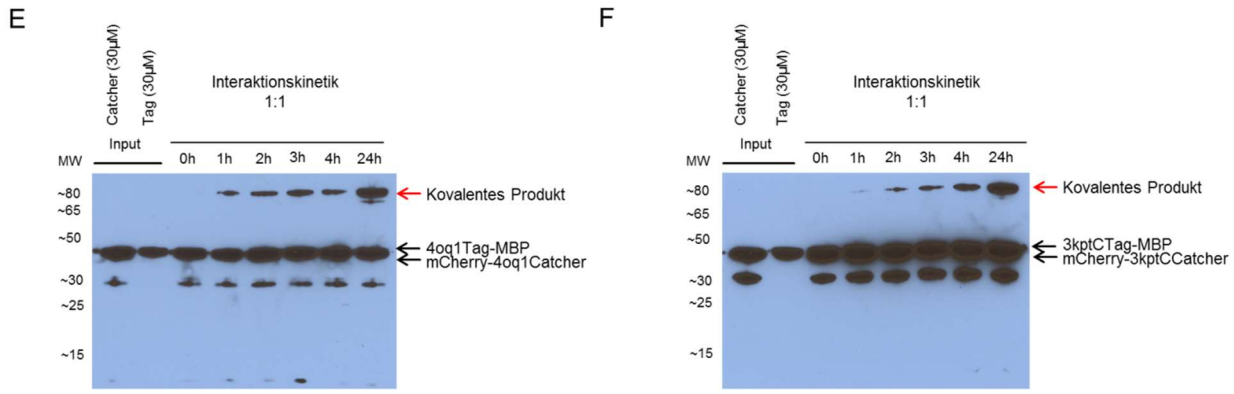


Abb.3.6: *In vitro* Rekonstitutionskinetiken der Catcher- und Tag-Proteine des 4oq1- und des 3kptC-Systems. Aufgereinigte und in 1xPBS pH 7,4 dialysierte Proteine wurden in äquimolarem Verhältnis (15 µM Endkonzentration von Catcher und Tag) in 1xPBS pH 7,5 Reaktionspuffer gemischt und unter Schütteln (500 rpm) bei 25°C inkubiert. Die Kinetik der Bildung von kovalenten Interaktionsprodukten zwischen Catcher- und Tag-Proteinen wurde mittels SDS-PAGE und Coomassie-Färbung bzw- Anti-His Western Blot analysiert. Die Proben wurden stündlich aus dem Interaktionsgemisch entnommen, 4x Lämmli-Probenpuffer zugegeben und 10 min bei 95°C gekocht, bevor sie auf Proteingele zur Elektrophorese aufgetragen wurden. Gezeigt sind kovalente Interaktionsanalysen zwischen (A) 4oq1Catcher und 4oq1Tag-MBP, (B und E) mCherry-4oq1Catcher und 4oq1Tag-MBP, (C) 3kptCCatcher und 3kptCTag-MBP und (D und F) mCherry-3kptCCatcher und 3kptCTag-MBP.

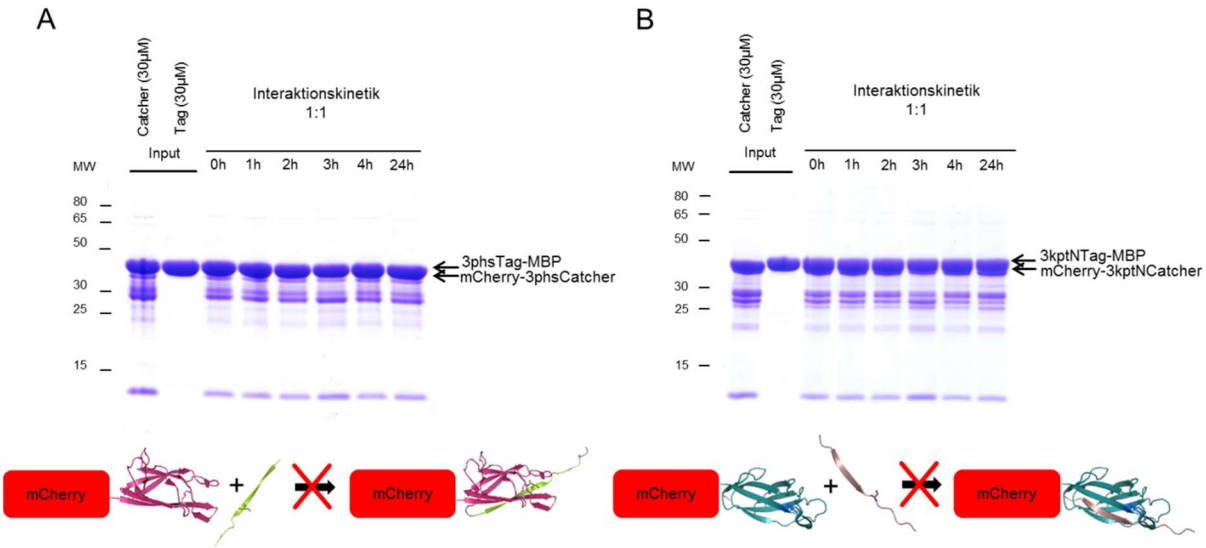


Abb.3.7: *In vitro* Rekonstitutionskinetik zwischen mCherry-3phsCatcher und 3phsTag-MBP (A) bzw. mCherry-3kptNCatcher und 3kptNTag-MBP (B). Aufgereinigte und in 1xPBS pH 7,4 dialysierte Proteine wurden in äquimolarem Verhältnis (15 µM Endkonzentration von Catcher und Tag) in 1xPBS pH 7,5 Reaktionspuffer gemischt und unter Schütteln (500 rpm) bei 25°C inkubiert. Die Kinetik der Bildung von kovalenten Interaktionsprodukten zwischen Catcher- und Tag-Proteinen wurde mittels SDS-PAGE und Coomassie-Färbung analysiert. Die Proben wurden stündlich aus dem Interaktionsgemisch entnommen, 4x Lämmli-Probenpuffer zugegeben und 10 min bei 95°C gekocht, bevor sie auf Proteingele zur Elektrophorese aufgetragen wurden.

3.2 Charakterisierung der neu identifizierten Split-Isopeptid Systeme

Nach der Identifizierung der neuen Catcher-Tag Systeme, 4oq1 und 3kptC, sollten die beiden Systeme zunächst sowohl genetisch als auch biochemisch näher charakterisiert werden (siehe Abb.3.8). Zuerst sollte die reaktive Triade in *in vitro* Analysen bestätigt werden. Anschließend wurde die Rekonstitutionseffizienz der neuen Systeme im Vergleich zu den etablierten Split-Isopeptid Systemen, Spy und Snoop (siehe 1.3.2, 1.3.3), untersucht. Kreuzreaktionen wurden getestet, um die Orthogonalität/ Spezifität der Systeme zu analysieren. Biochemische Analysen des 4oq1- und des 3kptC-Systems schlossen sich an, um optimale Reaktionsbedingungen für die kovalenten Catcher-Tag Reaktionen zu ermitteln (siehe Abb.3.8).

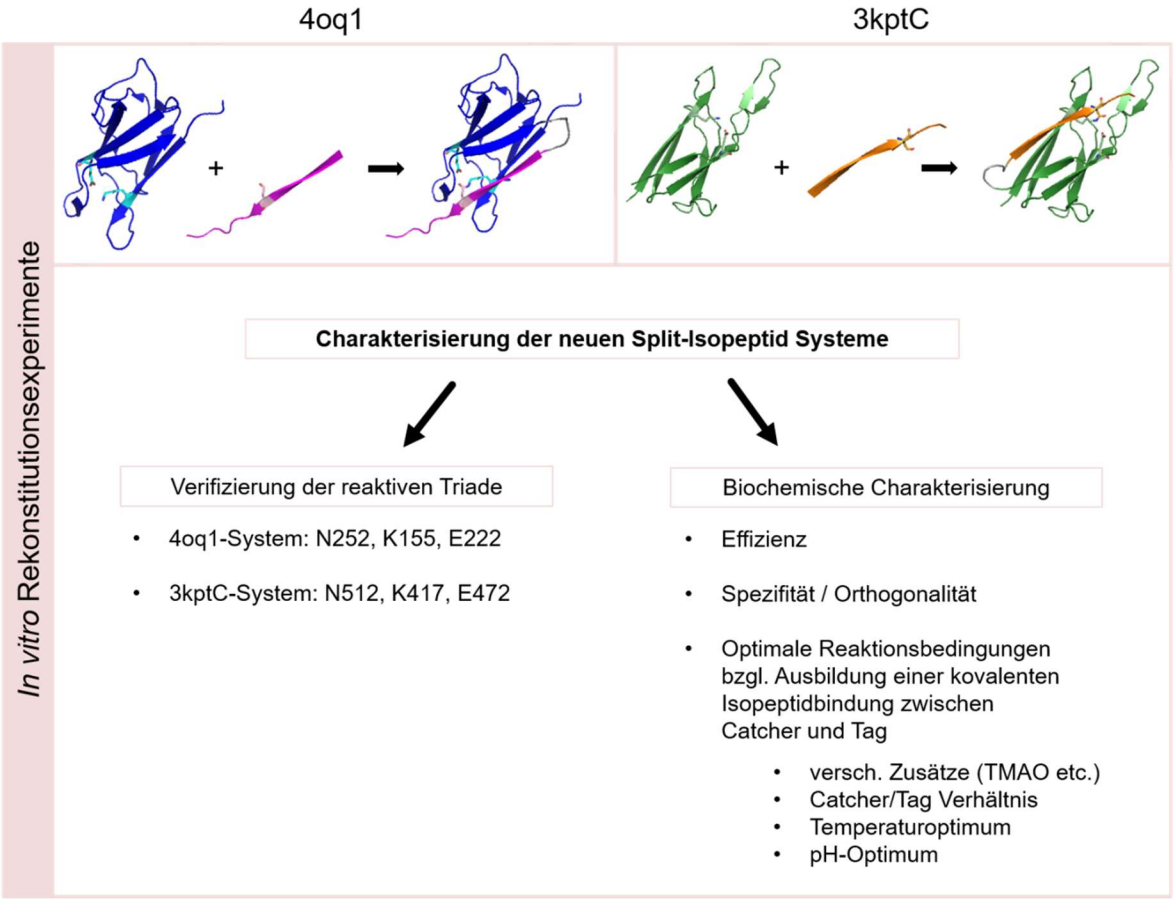


Abb.3.8: Zusammenfassendes Schema der Vorgehensweise zur Charakterisierung der neu identifizierten, kovalenten Split-Isopeptid Systeme, 4oq1 und 3kptC.

3.2.1 Verifizierung der reaktiven Triade

Im Zuge der genetischen Charakterisierung des 4oq1- und des 3kptC-Systems sollten zunächst die drei reaktiven Aminosäuren (bekannt aus den Kristallstrukturen), die direkt an der Bildung der Isopeptidbindung beteiligt sind, in *in vitro* Rekonstitutionsexperimenten verifiziert werden (siehe Kapitel 2.6.4.9). Dazu wurden entsprechende nicht-reaktive Catcher- und Tag-Mutanten über ortsspezifische Mutagenese mit Primern, die die Nukleotidveränderungen einführen, hergestellt (Klonierung siehe Kapitel 2.6.2 und Tabelle 1.1: #99, #100, #110, #143, #144, #145, Primersequenzen siehe Anhang Tabelle A2). Die Lysin (K)- und die Asparaginreste (N), zwischen denen die Isopeptidbindung stattfindet, wurden in Alanine (A) mutiert und die katalytischen Glutaminsäuren (E) zu Glutamin (Q). Die nicht-reaktiven Catcher- und Tag-Mutanten wurden einzeln in Rekonstitutionsanalysen auf ihre Fähigkeit zur Bildung von kovalenten Catcher-Tag Komplexen getestet (siehe Abb.3.9 und 3.10).

Während die Reaktion zwischen dem Wildtyp-4oq1Tag und dem Wildtyp-4oq1Catcher erwartungsgemäß in der Bildung eines kovalenten Produktes resultierte (siehe Abb.3.9 A), führte keine der drei Mutanten (4oq1Tag N252→A, 4oq1Catcher K155→A, 4oq1Catcher E222→Q) bei der Reaktion mit dem jeweiligen Wildtyp-Proteinpartner zu einer neuen Produktbande, die im Proteingel nach Kochen in 4x Lämmli-Puffer sichtbar war (siehe Abb.3.9 B - D). Analog verhielt es sich mit dem 3kptCCatcher-3kptCTag System. Die Wildtyp-3kptC-Proteinpartner bildeten bei gemeinsamer Inkubation ein kovalentes Produkt mit steigender Intensität über die Zeit (siehe Abb.3.10 A). Sobald eine der drei reaktiven Aminosäuren der Isopeptid-Triade (3kptCTag N512→A, 3kptCCatcher K417→A, 3kptCCatcher E472→Q) mutiert war, fand diese kovalente Reaktion zwischen Catcher und Tag nicht mehr statt (siehe Abb.3.10 B - D).

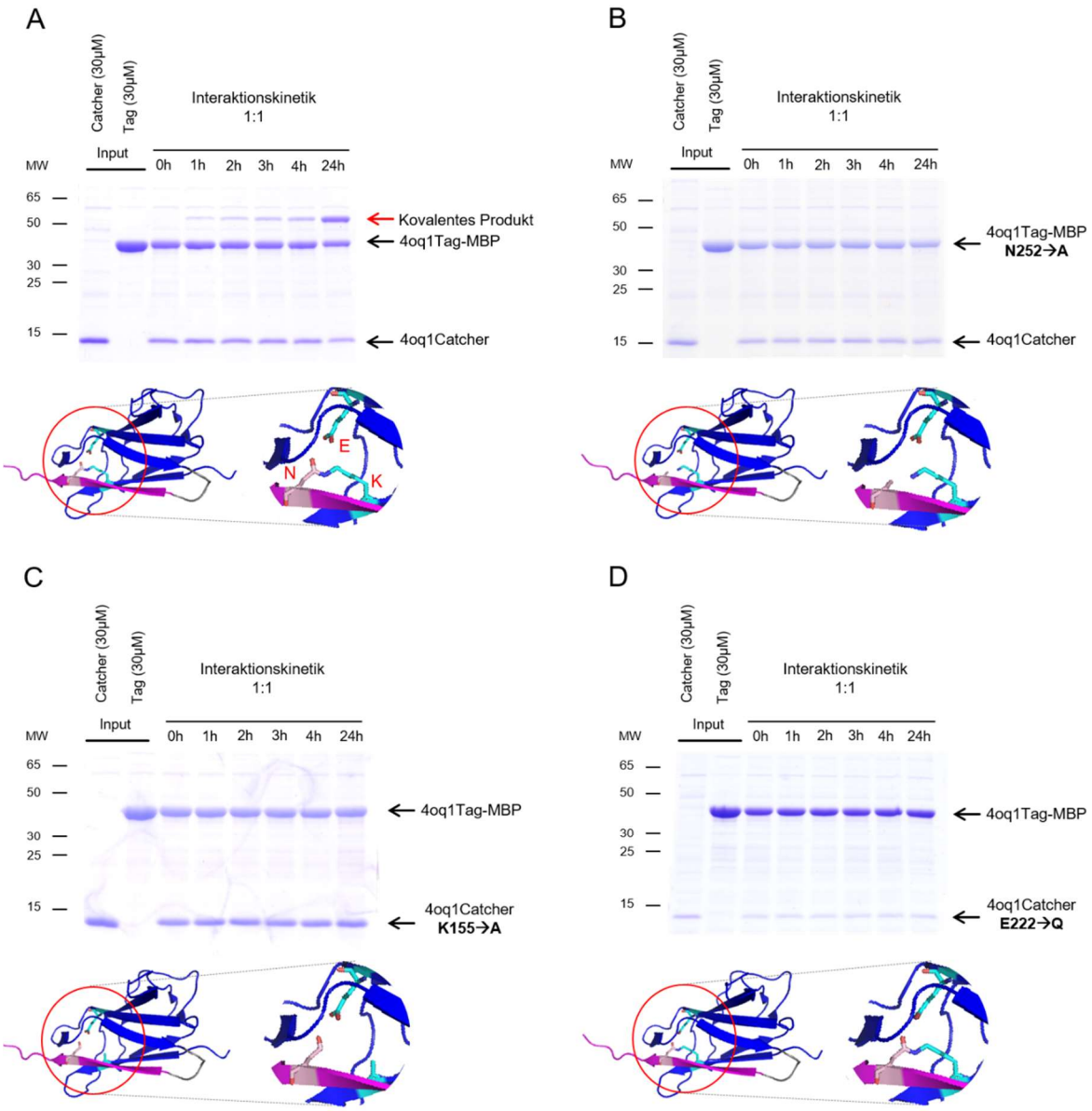


Abb.3.9: *In vitro* Rekonstitutionskinetik der Isopeptid-Triade Mutanten des 4oq1Catcher-4oq1Tag Systems. Aufgereinigte und in 1xPBS pH 7,4 dialysierte Proteine wurden in äquimolarem Verhältnis (15 µM Endkonzentration von Catcher und Tag) in 1xPBS pH 7,5 Reaktionspuffer gemischt und unter Schütteln (500 rpm) bei 25°C inkubiert. Die Kinetik der Bildung von kovalenten Interaktionsprodukten zwischen Catcher- und Tag-Proteinen wurde mittels SDS-PAGE und Coomassie-Färbung analysiert. Die Proben wurden stündlich aus dem Interaktionsgemisch entnommen, 4x Lämmli-Probenpuffer zugegeben und 10 min bei 95°C gekocht, bevor sie auf Proteingele zur Elektrophorese aufgetragen wurden. Gezeigt sind kovalente Interaktionsanalysen zwischen (A) 4oq1Catcher (#112) und 4oq1Tag-MBP (#98) (Wildtyp), (B) 4oq1Catcher und 4oq1Tag N252→A-MBP (#100), (C) 4oq1Catcher K155→A (#99) und 4oq1Tag-MBP und (D) 4oq1Catcher E222→Q (#110) und 4oq1Tag-MBP.

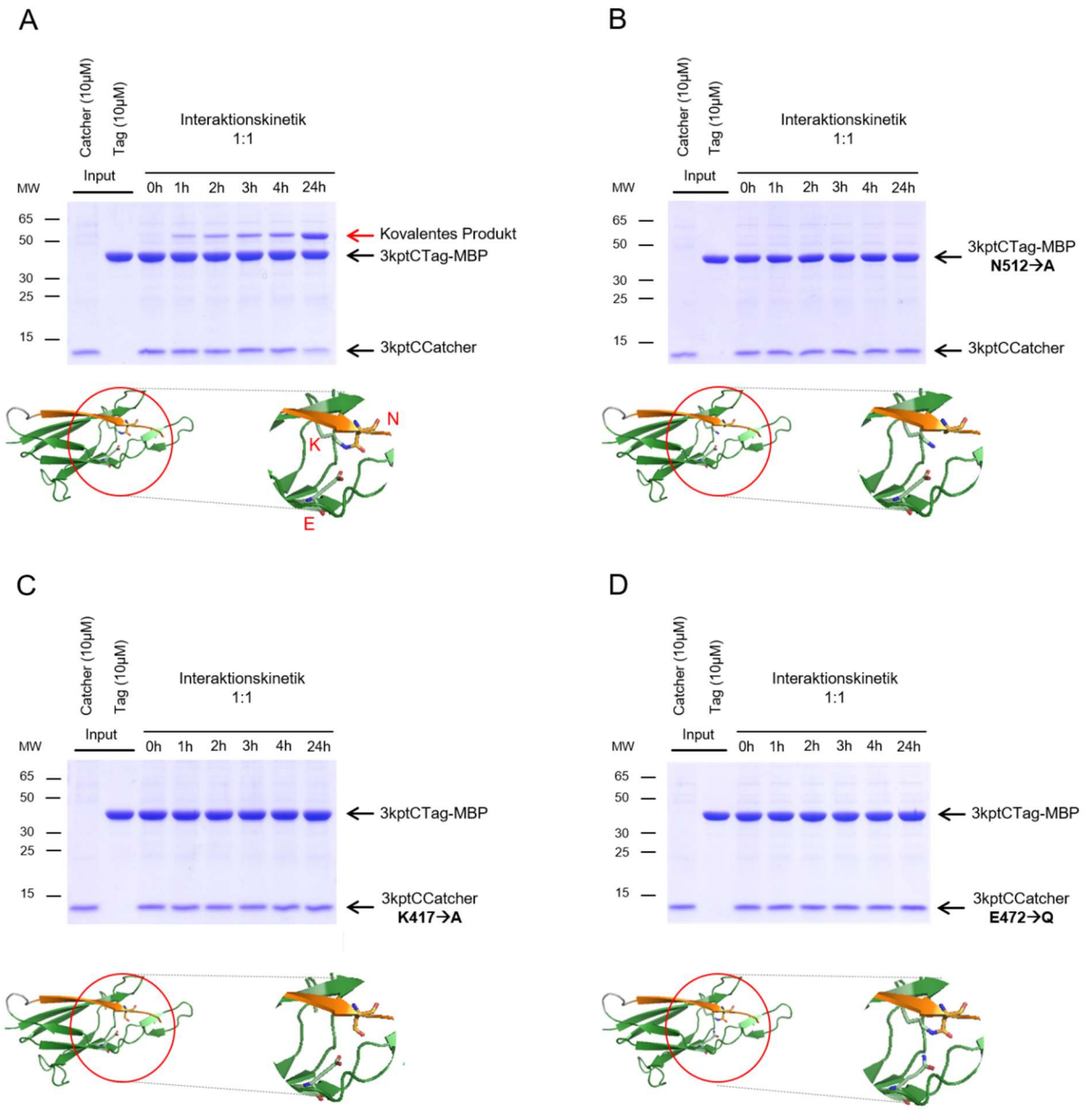


Abb.3.10: *In vitro* Rekonstitutionskinetik der Isopeptid-Triade Mutanten des 3kptCCatcher-3kptCTag Systems. Aufgereinigte und in 1xPBS pH 7,4 dialysierte Proteine wurden in äquimolarem Verhältnis (10 µM Endkonzentration von Catcher und Tag) in 1xPBS pH 7,5 Reaktionspuffer gemischt und unter Schütteln (500 rpm) bei 25°C inkubiert. Die Kinetik der Bildung von kovalenten Interaktionsprodukten zwischen Catcher- und Tag-Proteinen wurde mittels SDS-PAGE und Coomassie-Färbung analysiert. Die Proben wurden stündlich aus dem Interaktionsgemisch entnommen, 4x Lämmli-Probenpuffer zugegeben und 10 min bei 95°C gekocht, bevor sie auf Proteingele zur Elektrophorese aufgetragen wurden. Gezeigt sind kovalente Interaktionsanalysen zwischen (A) 3kptCCatcher (#142) und 3kptCTag-MBP (#136) (Wildtyp), (B) 3kptCCatcher und 3kptCTag N512→A-MBP (#143), (C) 3kptCCatcher K417→A (#145) und 3kptCTag-MBP und (D) 3kptCCatcher E472→Q (#144) und 3kptCTag-MBP.

3.2.2 Effizienz der kovalenten Reaktion zwischen Catcher und Tag

Zur ersten kinetischen Beurteilung der neuen Split-Isopeptid Systeme, 4oq1 und 3kptC, wurde die Effizienz der Bildung von kovalenten Catcher-Tag Komplexen im Vergleich zu den bereits etablierten und optimierten Split-Isopeptid Systemen, Spy und Snoop, in analog durchgeführten Interaktionsanalysen (gleiche Reaktionsbedingungen: Molaritäten, Puffer, pH, Temperatur) ermittelt (siehe 2.6.4.9). Laut Literatur reagieren sowohl das Spy- als auch das Snoop-System innerhalb von Minuten sehr effizient mit hoher Ausbeute miteinander (Zakeri et al., 2012, Veggiani et al., 2016, siehe Einleitung). Auch in den eigens durchgeführten Rekonstitutionsexperimenten wiesen die Interaktionen zwischen SpyCatcher und SpyTag bzw. zwischen SnoopCatcher und SnoopTag schon nach einer Stunde gemeinsamer Inkubation sehr starke Produktbanden auf, die im weiteren Reaktionsverlauf nur noch marginal zunahmen (siehe Abb.3.11 A und B). Die Produktbanden der neuen Catcher-Tag Systeme, 4oq1 und 3kptC, waren dagegen in der 1h-Probe sehr schwach ausgeprägt und wuchsen langsam und stetig über die Zeit (siehe Abb.3.11 C und D). Das Niveau der Produktausbeute des Spy- und Snoop-Systems wurde aber nie erreicht. Auf das gleiche Ergebnis kommt man, wenn man die Abnahme von Catcher- und Tag-Komponenten im Reaktionsverlauf (0h - 24h) betrachtet. Bei der SpyCatcher-SpyTag- und bei der SnoopCatcher-SnoopTag-Reaktion war nach einstündiger, gemeinsamer Inkubation von Catcher und Tag schon eine deutliche Abnahme der jeweiligen Edukt-Banden im Vergleich zu der jeweiligen 0h-Probe zu sehen (siehe Abb.3.11 A und B). Densitometrische Abschätzungen (es wurden keine Replikate angefertigt) der Spy- und der Snoop-Reaktion mittels ImageJ (siehe Kapitel 2.6.4.4) ergaben, dass eine Stunde nach Mischen der Catcher- und Tag-Komponenten schon ca. 50% der Catcher-Proteine abreagiert haben. Der SpyTag war sowohl 1h als auch 24h nach Inkubationsbeginn zu ca. 60% im Vergleich zur 0h-Probe aufgebraucht, während ungefähr die Hälfte der SpyCatcher-Proteine über den kompletten Analysezeitraum (1h - 24h) abreagiert war, verglichen mit der 0h-Probe. Eine Stunde nach Mischen von SnoopCatcher und SnoopTag haben 55% des SnoopCatchers und 26% des SnoopTags abreagiert. Die Abreaktion nahm über den Reaktionsverlauf etwas zu und erreichte 24h nach Inkubationsbeginn 88% an abreagiertem SnoopCatcher und 43% an abreagiertem SnoopTag, wobei die 0h-Probe immer als Referenz diente. Eine genauere Quantifizierung wurde erschwert durch die Tatsache, dass bereits in der 0h-Probe, also unmittelbar nach Mischen der jeweiligen Catcher- und Tag-Proteinkomponenten, eine leichte Produktbande zu sehen war, was für die enorm schnelle und effiziente Reaktion zwischen SpyTag und SpyCatcher, bzw. zwischen SnoopTag und SnoopCatcher spricht. Die Abnahme der Reaktanten für die Reaktionen des 4oq1- und des 3kptC-Systems war nur in der 24h-Probe, verglichen mit der 0h-Probe, wirklich sichtbar und spricht für eine ineffiziente

Reaktion, vor allem im Vergleich mit den höchst effizienten Catcher-Tag Systemen, Spy und Snoop (siehe Abb.3.11). Densitometrische Abschätzungen im Vergleich mit den jeweiligen 0h-Proben ergaben, dass bei dem 4oq1-System 62% des 4oq1Catchers und 7% des 4oq1Tags 24h nach Mischen der Komponenten abreagiert haben. Bei dem 3kptC-System waren 66% des 3kptCCatchers und nicht detektierbare Mengen an 3kptCTag reaktiv. Eine Stunde nach Mischen der Komponenten haben, verglichen mit der 0h-Probe, lediglich etwa 12% des 4oq1Catchers und nicht detektierbare Mengen an 4oq1Tag abreagiert. Bei dem 3kptC-System haben ca. 30% des 3kptCCatchers und auch nicht detektierbare Mengen an 3kptCTag abreagiert, ebenfalls im Vergleich zur 0h-Probe.

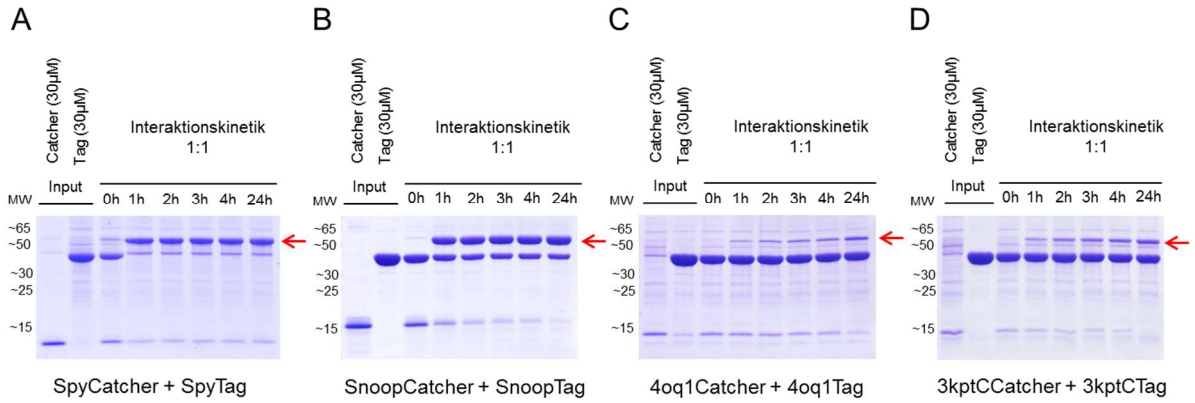


Abb.3.11: *In vitro* Rekonstitutionskinetiken zur Ermittlung der Reaktionseffizienz der Catcher-Tag Systeme. Aufgereinigte und in 1xPBS pH 7,4 dialysierte Proteine wurden in äquimolarem Verhältnis (15 µM Endkonzentration von Catcher und Tag) in 1xPBS pH 7,5 Reaktionspuffer gemischt und unter Schütteln (500 rpm) bei 25°C inkubiert. Die Kinetik der Bildung von kovalenten Interaktionsprodukten zwischen Catcher- und Tag-Proteinen wurde mittels SDS-PAGE und Coomassie-Färbung analysiert. Die Proben wurden stündlich aus dem Interaktionsgemisch entnommen, 4x Lämmli-Probenpuffer zugegeben und 10 min bei 95°C gekocht, bevor sie auf Proteingele zur Elektrophorese aufgetragen wurden. Gezeigt sind kovalente Interaktionsanalysen zwischen (A) SpyCatcher und SpyTag-MBP, (B) SnoopCatcher und SnoopTag-MBP, (C) 4oq1Catcher und 4oq1Tag-MBP und (D) 3kptCCatcher und 3kptCTag-MBP. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

Bei Betrachtung einer kleineren, engmaschigeren Zeitskala im Minutenbereich (1 min, 3 min, 5 min, 10 min, 30 min) ist ebenfalls zu sehen, dass die kovalente Reaktion zwischen 4oq1Catcher und 4oq1Tag bzw. zwischen 3kptCCatcher und 3kptCTag sehr langsam abläuft und sich eher im Stundenbereich abspielt (siehe Abb.3.12). Die kovalente Komplexbildung aus 4oq1Catcher und 4oq1Tag bzw. 3kptCCatcher und 3kptCTag konnte nach ca. 5 Minuten gemeinsamer Catcher-Tag Inkubation erstmals in Form einer sehr schwachen Produktbande, die über den gesamten Analysezeitraum langsam aber stetig zunahm, visuell im Proteingel detektiert werden (siehe Abb.3.12). Im Gegensatz dazu bilden SpyCatcher und SpyTag bzw. SnoopCatcher und SnoopTag unmittelbar nach Mischen der Reaktionspartner einen kovalenten Catcher-Tag Komplex, was an den Produktbanden in den 0h-Proben zu

sehen ist (siehe Abb.3.11). Das bedeutet, dass die kovalente Reaktion zwischen SpyCatcher und SpyTag bzw. zwischen SnoopCatcher und SnoopTag sofort (innerhalb von Sekunden) beginnt, sobald sich beide Komponenten treffen und die Zeit vom Mischen der Reaktionspartner bis zum Abstoppen der Reaktion durch Zugabe von 4x Lämmli-Puffer und Probennahme nicht ausreicht, um eine saubere 0h-Probe (keine kovalente Catcher-Tag Komplexbildung) zu bekommen.

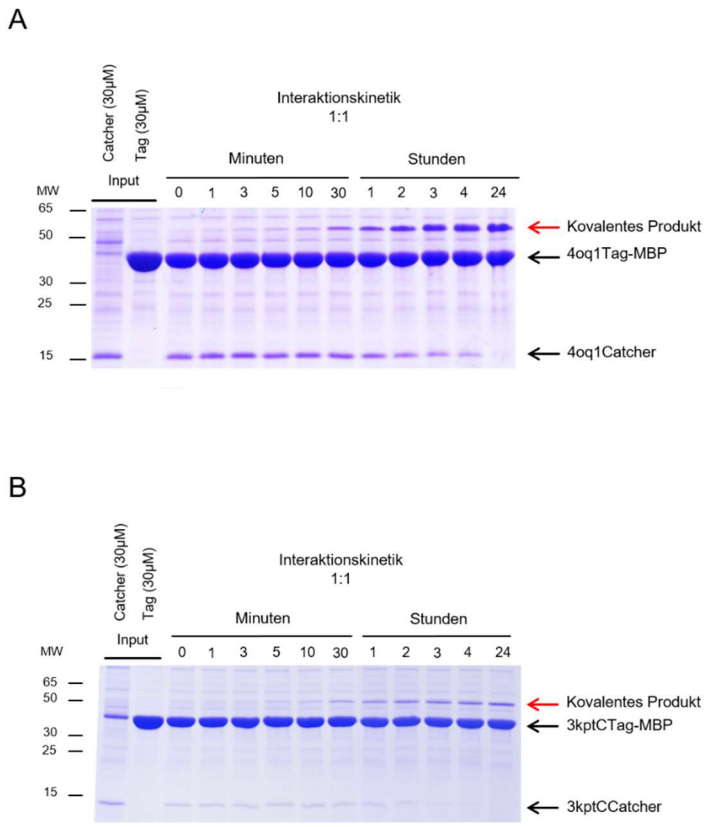


Abb.3.12: Genauere Analysen der Rekonstitutionseffizienz der neuen Catcher-Tag Systeme. Aufgereinigte und in 1xPBS pH 7,4 dialysierte Proteine wurden in äquimolarem Verhältnis (15 µM Endkonzentration von Catcher und Tag) in 1xPBS pH 7,5 Reaktionspuffer gemischt und unter Schütteln (500 rpm) bei 25°C inkubiert. Die Minuten- und Stundenkinetik der Bildung von kovalenten Interaktionsprodukten zwischen Catcher- und Tag-Proteinen wurden mittels SDS-PAGE und Coomassie-Färbung analysiert. Die Proben wurden aus dem Interaktionsgemisch entnommen, 4x Lämmli-Probenpuffer zugegeben und 10 min bei 95°C gekocht, bevor sie auf Proteingele zur Elektrophorese aufgetragen wurden. Gezeigt sind kovalente Interaktions-analysen zwischen (A) 4oq1Catcher und 4oq1Tag-MBP, (B) 3kptCCatcher und 3kptCTag-MBP.

3.2.3 Spezifität der Catcher-Tag Reaktionen

Neben der Effizienz als wichtigem Parameter für die Beurteilung der kovalenten Reaktion zwischen Catcher und Tag spielte die Analyse der Spezifität/ Orthogonalität eine essentielle Rolle, vor allem im Hinblick auf die Kontrollierbarkeit und Programmierbarkeit bei gleichzeitigem Einsatz mehrerer dieser Catcher-Tag Systeme (Spy, Snoop, 4oq1, 3kptC). Aus diesem Grund wurden alle möglichen Kreuzreaktionen in kovalenten Interaktionsanalysen unter konstanten Bedingungen (gleiche Proteinstammlösungen und Reaktionsbedingungen) untersucht. Bei den Kreuzreaktionsexperimenten fiel auf, dass der SpyCatcher sowohl mit dem 4oq1Tag als auch mit dem 3kptCTag kovalente Komplexe bildet, da die Coomassie-Gele neu entstandene Produktbanden bei diesen Interaktionen zeigten (siehe Abb.3.13, rote Pfeile). Anhand der Bandendicke und -intensität scheint es so, als wäre die Kreuzreaktion zwischen SpyCatcher und 4oq1Tag ausgeprägter als die

Reaktion zwischen SpyCatcher und 3kptCTag. Eine weitere Kreuzreaktion konnte zwischen dem 3kptCCatcher und dem 4oq1Tag festgestellt werden, was an der über die Zeit wachsenden neuen Bande, im Vergleich zu der 0h-Probe, sichtbar war (siehe Abb.3.13 C). Alle anderen möglichen Catcher-Tag Kombinationen führten zu keinem kovalenten Produkt, das resistent gegenüber Kochen in 4x Lämmli-Puffer war, da neue Banden auf der erwarteten Catcher-Tag-Produktgröße im Gel ausblieben (siehe Abb.3.13). Die Positivkontrollen, das heißt die jeweiligen Catcher-Proteine mit ihren entsprechenden Tag-Peptiden verhielten sich wie erwartet und bildeten kovalente Komplexe, die sich im Gel als neue hochmolekulare Banden, im Vergleich zur 0h-Probe, zeigten (siehe Abb.3.13 erste Spalte in A - D).

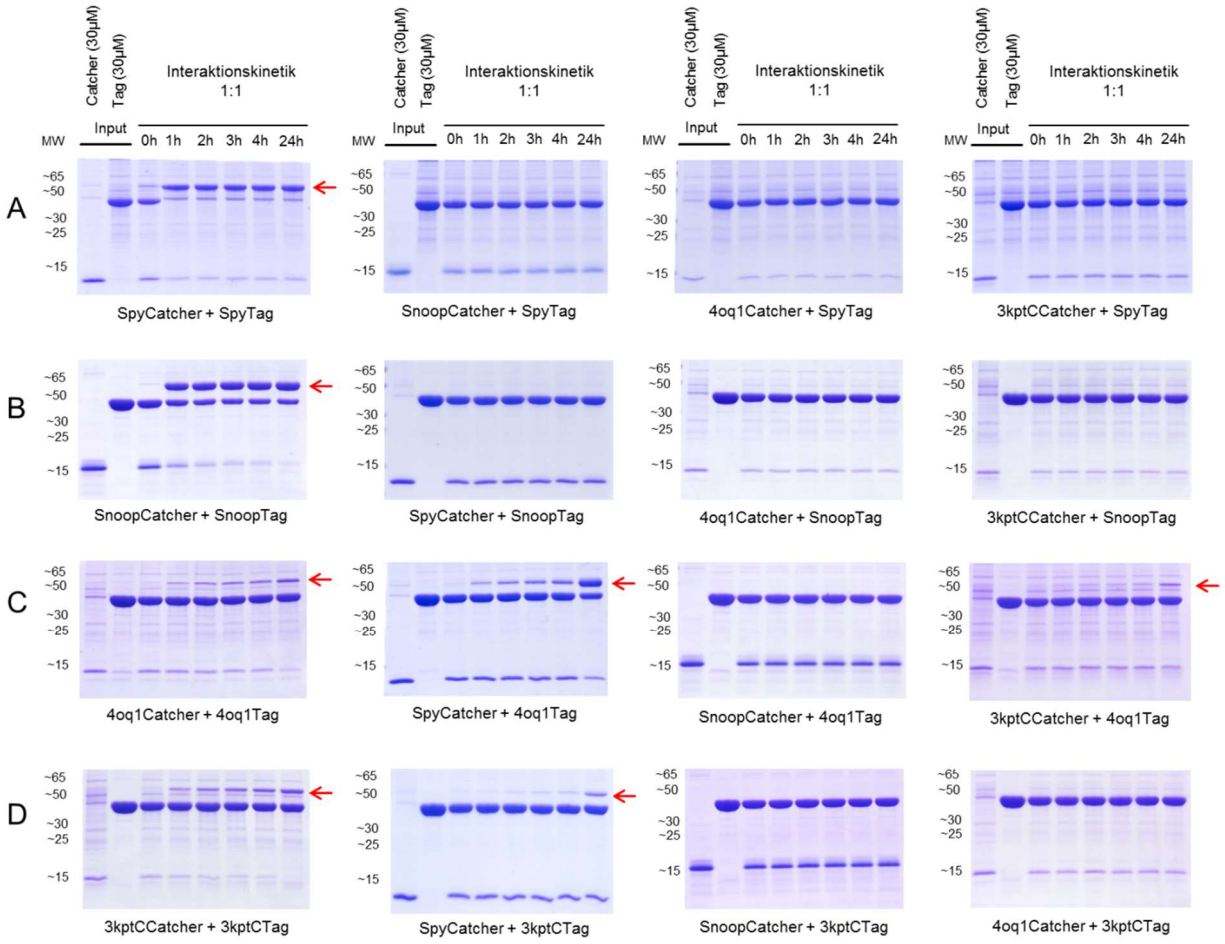


Abb.3.13: Analyse der Kreuzreaktivitäten der verschiedenen Catcher-Tag Systeme untereinander. Aufgereinigte und in 1xPBS pH 7,4 dialysierte Proteine wurden in äquimolarem Verhältnis (15 µM Endkonzentration von Catcher und Tag) in 1xPBS pH 7,5 Reaktionspuffer gemischt und unter Schütteln (500 rpm) bei 25°C inkubiert. Die Kinetik der Bildung von kovalenten Interaktionsprodukten zwischen Catcher- und Tag-Proteinen wurde mittels SDS-PAGE und Coomassie-Färbung analysiert. Die Proben wurden stündlich aus dem Interaktionsgemisch entnommen, 4x Lämmli-Probenpuffer zugegeben und 10 min bei 95°C gekocht, bevor sie auf Proteingele zur Elektrophorese aufgetragen wurden. Gezeigt sind kovalente Interaktionsanalysen zwischen (A) SpyTag-MBP und Spy-, Snoop-, 4oq1-, und 3kptCCatcher, (B) SnoopTag-MBP und Snoop-, Spy-, 4oq1-, und 3kptCCatcher, (C) 4oq1Tag-MBP und 4oq1-, Spy-, Snoop-, und 3kptCCatcher, (D) 3kptCTag-MBP und 3kptC-, Spy-, Snoop-, und 4oq1Catcher. Kovalente Catcher-

Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Der Auftrag der Spuren ist bei allen Gelen derselbe und entspricht der Beschriftung in Zeile A.

3.2.4 Optimale Reaktionsbedingungen der kovalenten Catcher-Tag Reaktion

Zur weiteren biochemischen Analyse des 4oq1Catcher-4oq1Tag Systems und des 3kptCCatcher-3kptCTag Systems sollten die optimalen Reaktionsbedingungen (Wirkung chemischer Zusätze, optimales Catcher/Tag Verhältnis, Temperaturoptimum, pH-Optimum), unter denen sich kovalente Catcher-Tag Komplexe am effizientesten ausbilden können, analysiert werden (siehe Abb.3.8). Dafür wurden vergleichende Rekonstitutionsexperimente zwischen aufgereinigten und dialysierten Catcher- (#102 bzw. #142 siehe Tabelle 1.1) und Tag-Komponenten (#104 bzw. #136 siehe Tabelle 1.1) in Replikaten durchgeführt und die Effizienz der kovalenten Reaktionen densitometrisch quantifiziert (siehe Kapitel 2.6.4.4). Als Maß der Effizienz der kovalenten Reaktion zwischen Catcher und Tag wurden sowohl die Zunahme der Produktbande (0h-Probe als Referenz) als auch die Abnahme der Catcher- oder Tag-Input-Banden herangezogen (Input-Banden als 100%-Referenz).

3.2.4.1 Analyse der Wirkung verschiedener chemischer Zusätze auf die Rekonstitutionseffizienz

Basierend auf der Tatsache, dass die Zugabe einer 1,5 M Lösung des chemischen Chaperons Trimethylamin N-oxid (TMAO) die Reaktionsrate sowohl des Spy-Systems als auch des Snoop-Systems verbesserte (Fierer et al., 2014, Veggiani et al., 2016, Yancey, 2004), wurde getestet, ob die Reaktivität der neu identifizierten Systeme, 4oq1 und 3kptC, ebenfalls durch Zugabe von 1,5 M TMAO (Sigma-Aldrich) verbessert wird. Zusätzlich wurde die Wirkung von 40% Glycerin als kosmotrope Chemikalie (Russo, 2008) auf die Ausbildung der kovalenten Catcher-Tag Komplexe getestet. Als Referenz dienten analog durchgeführte Interaktionen (gleiche Reaktionsbedingungen) nur ohne den Zusatz von Chemikalien.

Bei der äquimolaren Interaktion zwischen 4oq1Catcher und 4oq1Tag konnte eine leichte Verbesserung der Rekonstitutionseffizienz nach Zugabe von 1,5 M TMAO oder 40% Glycerin detektiert werden (siehe Abb.3.14 A - C). Deutlicher wurde der positive Effekt von 1,5 M TMAO aber, wenn fünfmal so viel 4oq1Catcher wie 4oq1Tag in die Reaktion eingesetzt wurde (Ermittlung des optimalen Catcher/Tag Verhältnisses siehe Kapitel 3.2.4.2) (siehe Abb.3.14 D). Densitometrische Analysen ergaben, dass 2,5 Stunden nach Mischen von 4oq1Catcher und 4oq1Tag ca. 70% des 4oq1Tags und ca. 50% des 4oq1Catchers abreagiert haben (Input-Banden als Referenz), wenn dem Reaktionsgemisch 1,5 M TMAO

zugesetzt wurde. Ohne Zugabe von TMAO waren zum gleichen Zeitpunkt lediglich ca. 40% des 4oq1Tags und 25% des 4oq1Catchers verbraucht (siehe Abb.3.14 D und E).

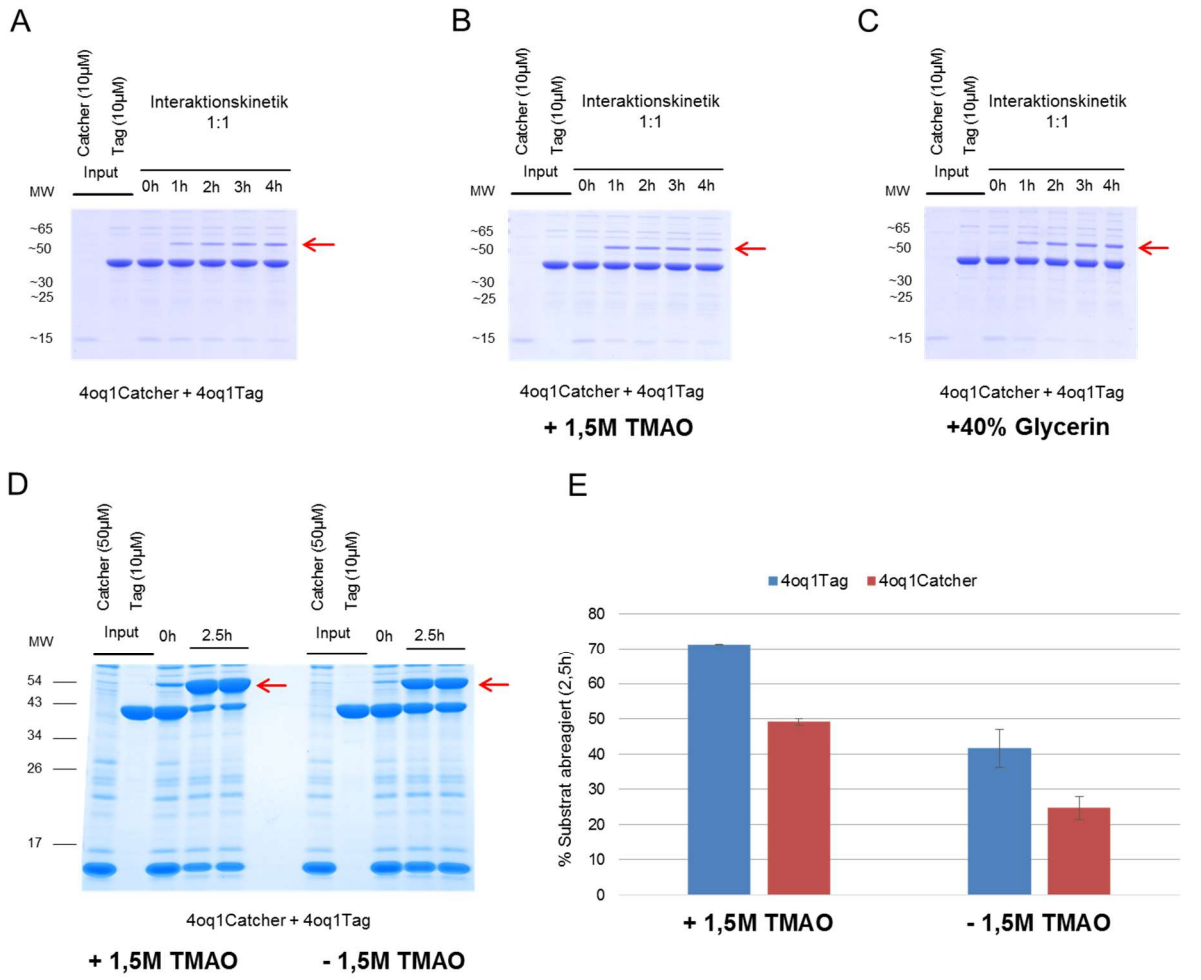


Abb.3.14: Einfluss von 1,5M TMAO und 40% Glycerin auf die Rekonstitutionseffizienz von 4oq1Catcher und 4oq1Tag. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in unterschiedlichen Catcher/Tag Verhältnissen (A - C: 1:1, D: 5:1) in 1xPBS pH 7,5 Reaktionspuffer mit und ohne 1,5M TMAO bzw. 40% Glycerin gemischt und bei 25°C inkubiert (A - D). Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. (E) Das Diagramm spiegelt die densitometrische Auswertung des Proteingels in D wider und zeigt die Mittelwerte und Standardabweichungen der in Duplikaten durchgeführten Reaktionen 2,5h nach Reaktionsstart mit und ohne 1,5M TMAO. Als Maß der Rekonstitutionseffizienz wurde die Abreaktion von 4oq1Catcher- und 4oq1Tag im Vergleich zu den Input-Spuren (Input-Banden als 100%) quantifiziert. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

Ähnlich verhielt es sich mit der kovalenten Reaktion zwischen 3kptCCatcher und 3kptCTag. Durch Zugabe von 1,5 M TMAO (Abb.3.15 B und E) oder 40% Glycerin (Abb.3.15 C und F) konnte die Reaktionsrate eindeutig verbessert werden, was an der dickeren Produktbande und der stärkeren Abnahme der Input-Banden im Vergleich zu der Reaktion ohne Zugabe der Chemikalien (Abb.3.15 A und D) zu sehen war. Die densitometrische Quantifizierung der in Triplikaten durchgeführten Interaktionen zwei Stunden nach Inkubationsstart ergab, dass bei Zugabe von 1,5 M TMAO 25% des 3kptCTags und 76% des 3kptCCatchers abreagiert

haben. Ähnlich verhielt es sich bei der Zugabe von 40% Glycerin mit 23% abreagiertem 3kptCTag und 74% abreagiertem 3kptCCatcher. Ohne jegliche Zugabe von Chemikalien waren nach zweistündiger Reaktionszeit nicht detektierbare Mengen an 3kptCTag und 28% an 3kptCCatcher verbraucht (siehe Abb.3.15 G). Da 1,5 M TMAO und 40% Glycerin ähnlich positive Effekte auf die kovalente Reaktion zwischen 3kptCCatcher und 3kptCTag hatten, wurde mit 40% Glycerin weitergearbeitet, da TMAO auch einen Einfluss auf den pH-Wert des Reaktionsgemisches hat.

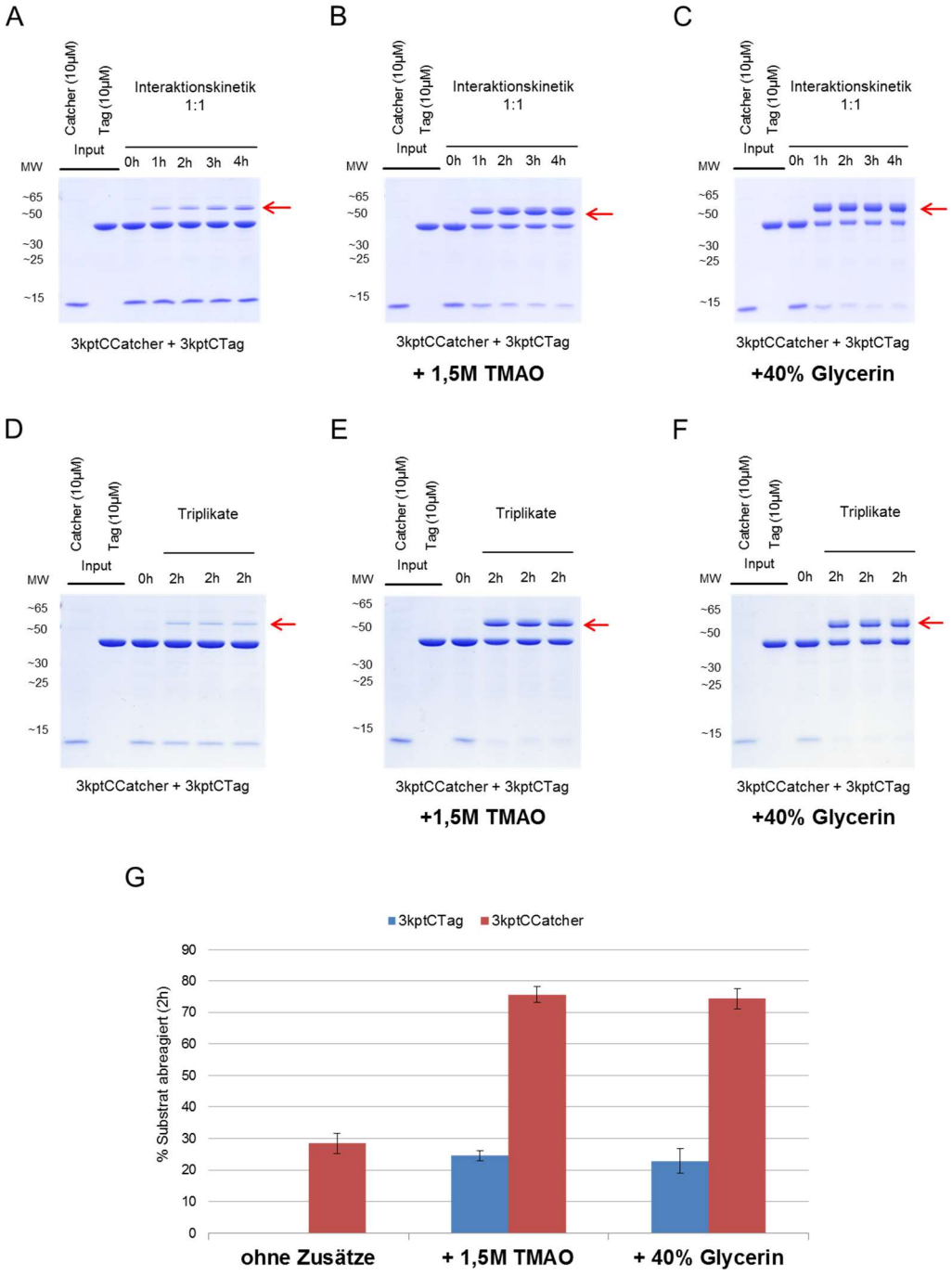


Abb.3.15: Einfluss von 1,5M TMAO und 40% Glycerin auf die kovalente Reaktion zwischen 3kptCCatcher und 3kptCTag. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in äquimolarem Verhältnis in 1xPBS pH 7,5 Reaktionspuffer gemischt und bei 25°C inkubiert (A – F). Die

Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. (G) Das Diagramm spiegelt die densitometrische Auswertung der Proteingele D - F wider und zeigt die Mittelwerte und Standardabweichungen der in Triplikaten durchgeführten Reaktionen 2h nach Reaktionsstart mit und ohne 1,5M TMAO bzw. 40% Glycerin. Als Maß der Rekonstitutionseffizienz wurde die Abreaktion von 3kptCCatcher und 3kptCTag im Vergleich zu den Input-Spuren (Input-Banden als 100%) quantifiziert. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

3.2.4.2 Optimales Catcher/Tag Verhältnis der Reaktion

Wahrscheinlich ist die Effizienz bzw. Weiterreaktion der kovalenten Interaktion zwischen Catcher und Tag durch den vorzeitigen Verbrauch oder partiellen Nichtgebrauch (partielle Fehlfaltung) eines der beiden Reaktanten limitiert. Bei den neu identifizierten Systemen 4oq1 und 3kptC waren vor allem die Catcher-Komponenten problematisch. Schon die Expression der Catcher-Proteine in *E. coli* M15 [pREP4] Zellen war um einiges schlechter als die Expression der jeweiligen Tag-MBP Konstrukte und die Reinheit der Catcher-Eluate nach der Affinitätsaufreinigung war ebenfalls schlechter als die der Tag-Eluate (siehe Kapitel 3.1.3 bzw. Input-Spuren der Interaktionsanalysen). Um das für die Bildung von kovalenten Catcher-Tag Komplexen optimale Verhältnis von Catcher und Tag herauszufinden wurden Rekonstitutionsexperimente mit steigenden Catcher-Konzentrationen durchgeführt, während die Tag-Konzentration konstant gehalten wurde.

Mit steigender Konzentration des 4oq1Catchers konnte die kovalente Reaktion zwischen 4oq1Catcher und 4oq1Tag deutlich verbessert werden (siehe Abb.3.16 A und B), was die Konzentrationsabhängigkeit der Reaktion zeigt. Die im Vergleich zur 0h-Probe neu entstandene Produktbande wurde mit steigender Catcher-Konzentration immer dicker und die korrelierende Abnahme der Edukt-Banden immer deutlicher im Vergleich zu der äquimolaren Interaktion zwischen 4oq1Catcher und 4oq1Tag. Während der 4oq1Tag in der 1:1 Interaktion kaum abreagieren konnte (nur 13% nach 2,5-stündiger Inkubation), da der 4oq1Catcher schon nahezu aufgebraucht war (71% des 4oq1Catchers waren nach 2,5 Stunden aufgebraucht), konnten durch 5-fachen 4oq1Catcher-Überschuss schließlich 68% des 4oq1Tags reagieren und kovalente Komplexe mit dem 4oq1Catcher bilden, von dem 50% abreagiert haben (siehe Abb.3.16 A und B). Beobachtet man den Trend der Banden-abnahme von 4oq1Catcher und 4oq1Tag unter verschiedenen Verhältnissen sieht man, dass eine weitere Erhöhung der 4oq1Catcher-Endkonzentration bei konstanten 4oq1Tag-Mengen zu keiner weiteren Verbesserung der kovalenten Komplexbildung führen würde. Der Prozentsatz an reaktivem 4oq1Catcher-Protein nimmt bei Reaktionen, die das 3:1 Verhältnis überschreiten wieder ab und größere Mengen an 4oq1Catcher-Protein liegen wieder unverbraucht vor (siehe Abb. 3.16 A und B), da nun wahrscheinlich der 4oq1Tag der limitierende Faktor einer Weiterreaktion ist.

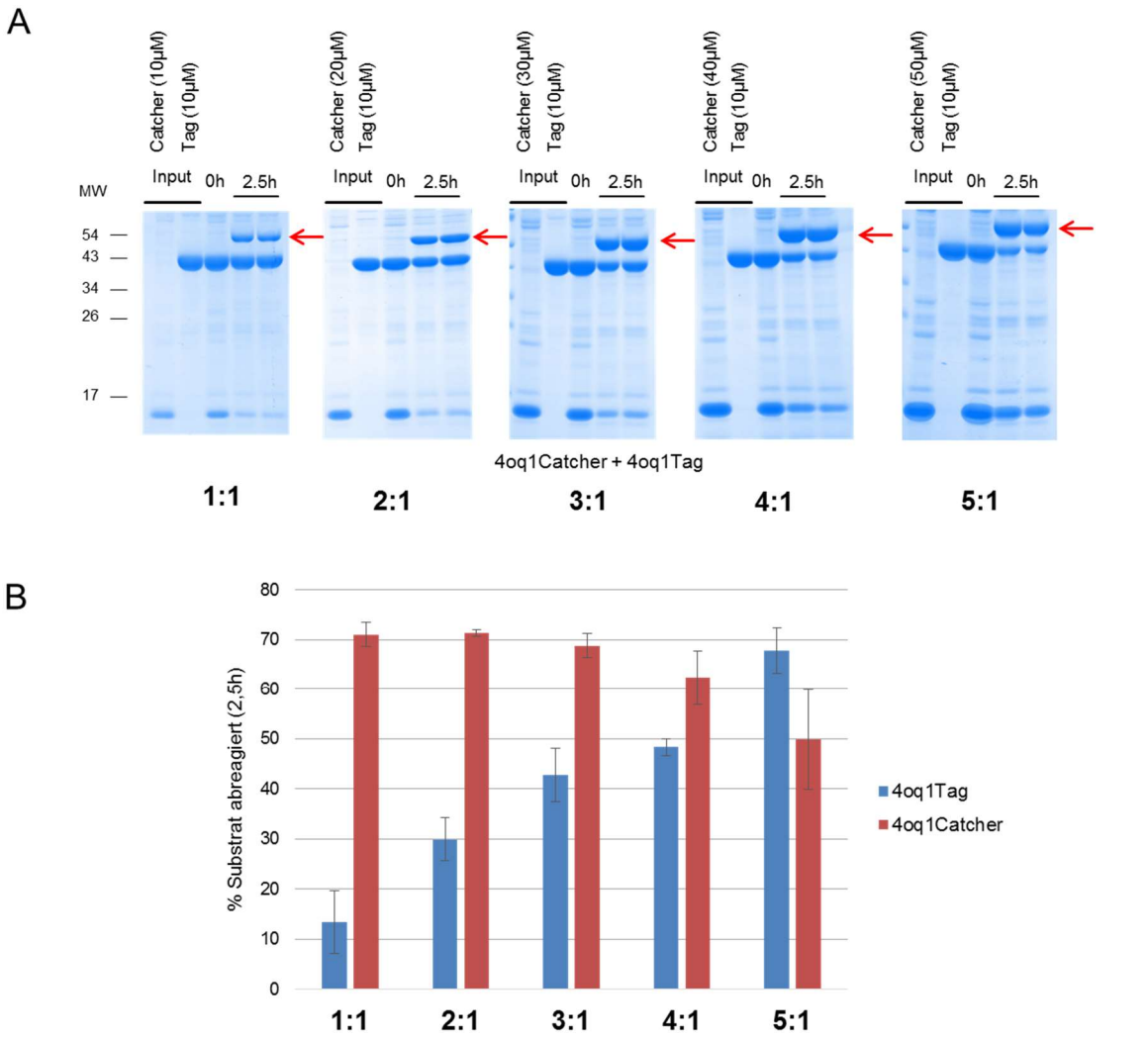


Abb.3.16: Ermittlung des optimalen Catcher/Tag Verhältnisses der kovalenten Reaktion zwischen 4oq1Catcher und 4oq1Tag. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in verschiedenen Catcher/Tag Verhältnissen (1:1 bis 5:1) und unter Zugabe von 1,5M TMAO zu dem 1xPBS pH 7,5 Reaktionspuffer gemischt und bei 25°C inkubiert (A). Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. (B) Das Diagramm spiegelt die densitometrische Auswertung der Proteingele unter A wider und zeigt die Mittelwerte und Standardabweichungen der in Duplikaten durchgeführten Reaktionen 2,5h nach Reaktionsstart. Als Maß der Rekonstitutionseffizienz wurde die Abreaktion von 4oq1Catcher und 4oq1Tag im Vergleich zu den Input-Spuren (Input-Banden als 100%) quantifiziert. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

Ähnlich verhielt es sich mit dem 3kptC-System, dessen Reaktionsrate ebenfalls konzentrationsabhängig ist. Bei einer äquimolaren Interaktion zwischen 3kptCCatcher und 3kptCTag haben nach zweistündiger gemeinsamer Inkubation 34% des 3kptCTags und 75% des 3kptCCatchers zu kovalenten Komplexen (neu entstandene Produktbande im Vergleich zu 0h-Probe) abreagiert (siehe Abb.3.17). Steigert man die Konzentration des 3kptCCatchers im Reaktionsgemisch, so konnte eine besser ausgeglichene Abreaktion von 3kptCCatcher und 3kptCTag erzielt werden, bei der beide Reaktionspartner in ähnlichem Maße verbraucht wurden. Das optimale Ergebnis bezüglich der Zunahme an kovalentem

Produkt und korrelierender Abnahme der Eduktbanden/ Input-Banden lieferte ein 2:1 Verhältnis von 3kptCCatcher zu 3kptCTag (siehe Abb.3.17). So haben zwei Stunden nach Reaktionsstart 60% des 3kptCTags und 67% des 3kptCCatchers zu einem kovalenten Komplex reagiert. Im Gegensatz dazu blieben bei einem 1:1 Verhältnis von 3kptCCatcher und 3kptCTag, 66% des Tags unreaktiv, wobei 75% des Catchers abreagiert haben. Umgekehrt verhielt es sich bei Catcher-Überschüssen, die das 2:1 Catcher zu Tag Verhältnis überstiegen. Während der Prozentsatz an reaktiven 3kptCTag-Peptiden mit steigender 3kptCCatcher-Konzentration stets zunahm, wurde die Menge an unreaktiven 3kptCCatcher-Proteinen immer größer (siehe Abb.3.17; 74% unreaktiver 3kptCCatcher bei den Verhältnissen 4:1 und 5:1).

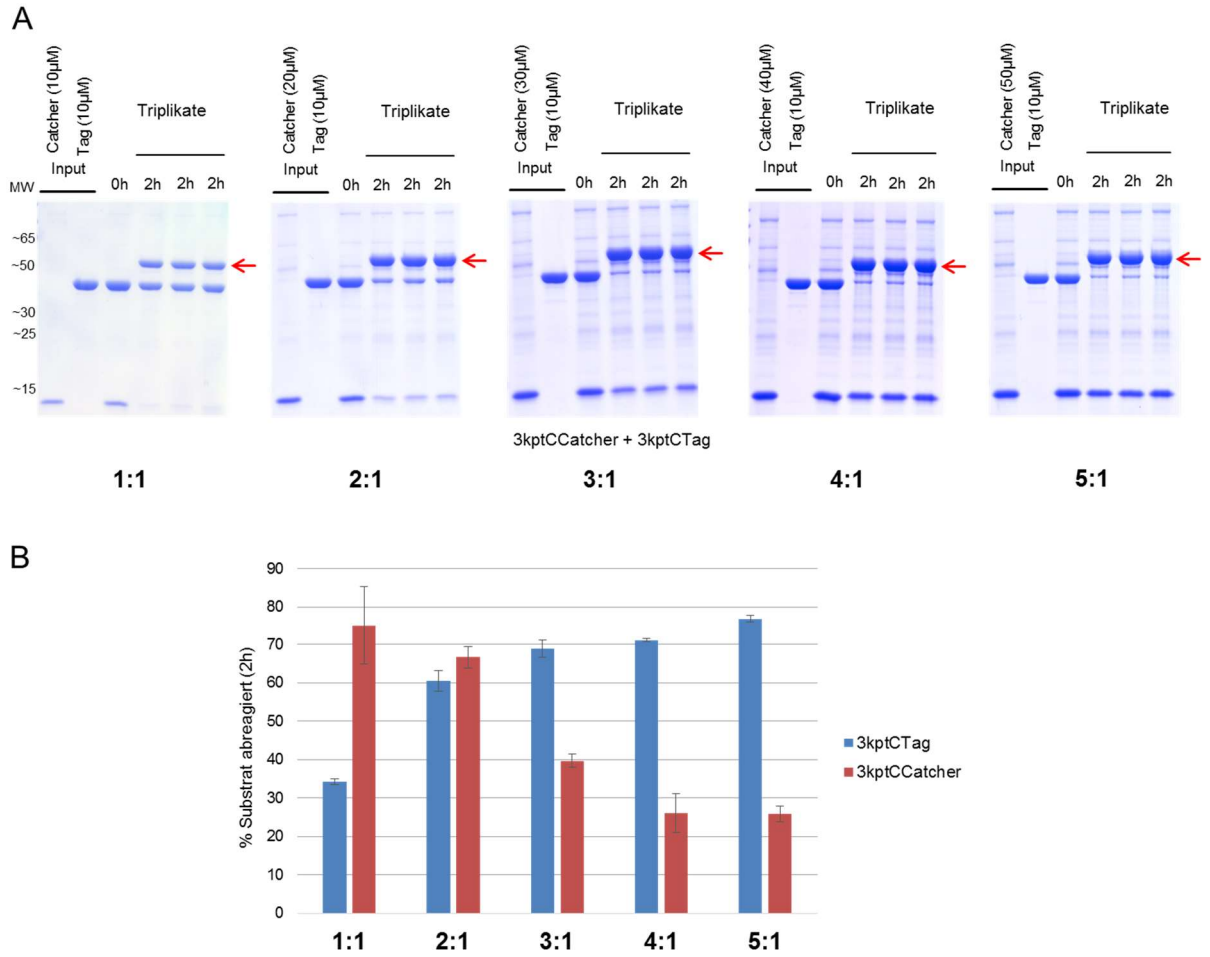


Abb.3.17: Ermittlung des optimalen 3kptCCatcher/3kptCTag Verhältnisses. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in verschiedenen Catcher/Tag Verhältnissen (1:1 bis 5:1) und unter Zugabe von 40% Glycerin zu dem 1xPBS pH 7,5 Reaktionspuffer gemischt und bei 25°C inkubiert. (A) Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. (B) Das Diagramm spiegelt die densitometrische Auswertung der Proteingele unter A wider und zeigt die Mittelwerte und Standardabweichungen der in Triplikaten durchgeführten Reaktionen 2h nach Reaktionsstart. Als Maß der Rekonstitutionseffizienz wurde die Abreaktion von 3kptCCatcher und 3kptCTag im Vergleich zu den Input-Spuren (Input-Banden als 100%) quantifiziert. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

3.2.4.3 Temperaturabhängigkeit der Catcher-Tag Reaktion

Biochemische Reaktionen und vor allem deren Ablaufgeschwindigkeit sind allgemein temperaturabhängig (RGT-Regel). Zudem hat die Temperatur einen großen Einfluss auf die Konformation und Stabilität von Proteinen (z.B. Hitzedenaturierung von Proteinen als Folge der Proteinentfaltung) sowie auf den pK_S-Wert von Aminosäuren bzw. ionisierbaren Gruppen (Stryer). Folglich sollte die Temperaturabhängigkeit der kovalenten Reaktion zwischen Catcher und Tag untersucht werden.

In vergleichenden Rekonstitutionsexperimenten, bei denen nur die Temperatur als einziger Parameter verändert wurde, konnte gezeigt werden, dass die kovalente Reaktion zwischen 4oq1Catcher und 4oq1Tag unter allen getesteten Temperaturen funktioniert, da im Vergleich zu den Input-Proben eine neue, höhermolekulare Produktbande zwei Stunden nach Reaktionsstart zu sehen war (siehe Abb.3.18 A). Dass eine Produktbande bereits in der 0h-Probe zu sehen war, könnte daran liegen, dass das Abstoppen der Reaktion mit SDS-Probenpuffer durch unzureichendes Mischen nicht komplett erfolgreich war. Deutlich zu sehen war außerdem, dass die kovalente Reaktion zwischen 4oq1Catcher und 4oq1Tag mit steigender Temperatur effizienter wird (siehe Abb.3.18). Vergleicht man die Produktbanden 2,5 Stunden nach Reaktionsstart bei 4°C und 12°C mit denen von 25°C, 37°C und 42°C so ist eine deutliche Effizienzsteigerung zu sehen (siehe Abb.3.18 A), die sich auch in der Quantifizierung der Abnahme der Catcher- und Tag-Komponenten widerspiegelt (siehe Abb.3.18 B). Eine Erhöhung der Temperatur von 4°C auf 12°C und dann auf 25°C führte zu einer gesteigerten Reaktivität des 4oq1Catchers und des 4oq1Tags um jeweils mindestens 10% (siehe Abb.3.18 B). Die Temperaturerhöhung von 25°C auf 37°C bzw. 42°C führte zu keiner weiteren Steigerung der Effizienz der kovalenten Reaktion zwischen 4oq1Catcher und 4oq1Tag, sondern blieb auf ähnlichem Niveau, was für ein breites Temperaturoptimum dieser Reaktion spricht.

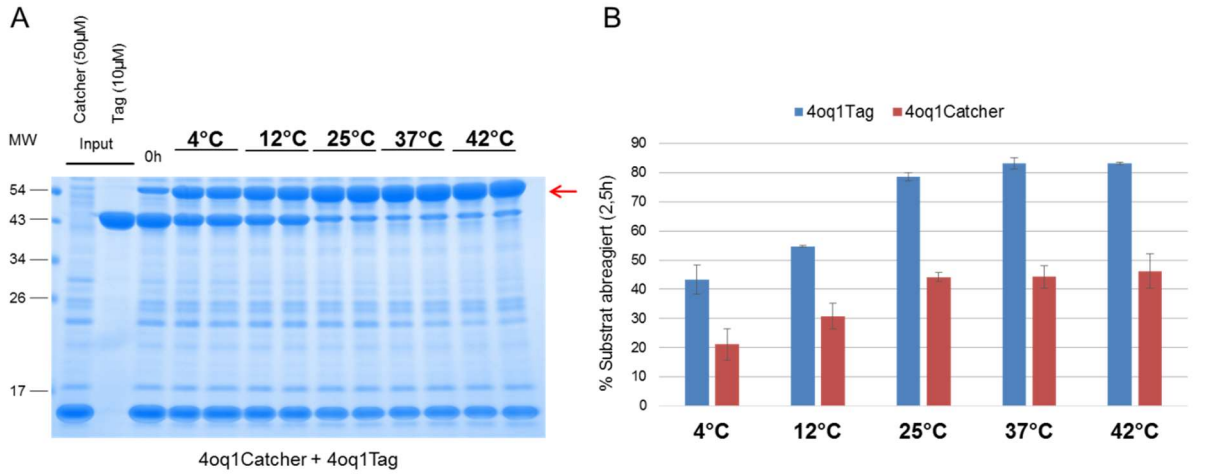


Abb.3.18: Temperaturabhängigkeit der kovalenten Reaktion zwischen 4oq1Catcher und 4oq1Tag. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher/Tag Verhältnis in 1xPBS pH 7,5 Reaktionspuffer unter Zugabe von 1,5M TMAO gemischt und bei verschiedenen Temperaturen (4°C bis 42°C) inkubiert (A). Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. (B) Das Diagramm spiegelt die densitometrische Auswertung des Proteingels unter A wider und zeigt die Mittelwerte und Standardabweichungen der in Duplikaten durchgeführten Reaktionen 2,5h nach Reaktionsstart. Als Maß der Rekonstitutionseffizienz wurde die Abreaktion von 4oq1Catcher und 4oq1Tag im Vergleich zu den Input-Spuren (Input-Banden als 100% Referenz) quantifiziert. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

Die kovalente Reaktion zwischen 3kptCCatcher und 3kptCTag konnte über den kompletten, getesteten Temperaturbereich detektiert werden (siehe neue Produktbande bei allen Temperaturbedingungen in Abb.3.19 A) und hatte ihr Temperaturoptimum bei 12°C bzw. 25°C. Nach zweistündiger Inkubation bei 12°C haben 62% des 3kptCTags und 72% des 3kptCCatchers abreagiert. Die analoge Inkubation bei 25°C erbrachte 57% an reaktivem Tag und 71% an reaktivem Catcher. Während die 4oq1-Reaktion ein breites Temperaturoptimum von 25°C - 42°C aufwies, verschlechterte sich die Effizienz der 3kptC-Reaktion deutlich bezüglich der Bildung von kovalenten Catcher-Tag Komplexen bei Temperaturen über 25°C. Selbst die kovalente Reaktion zwischen 3kptCCatcher und 3kptCTag bei 4°C (nach zweistündiger Inkubation haben sowohl 50% des Catchers als auch 50% des Tags abreagiert) war effizienter als die analogen Reaktionen bei 37°C oder noch dramatischer bei 42°C, wo zwei Stunden nach Reaktionsstart nur 8% des Tags und 14% des Catchers reaktiv waren (siehe Abb.3.19).

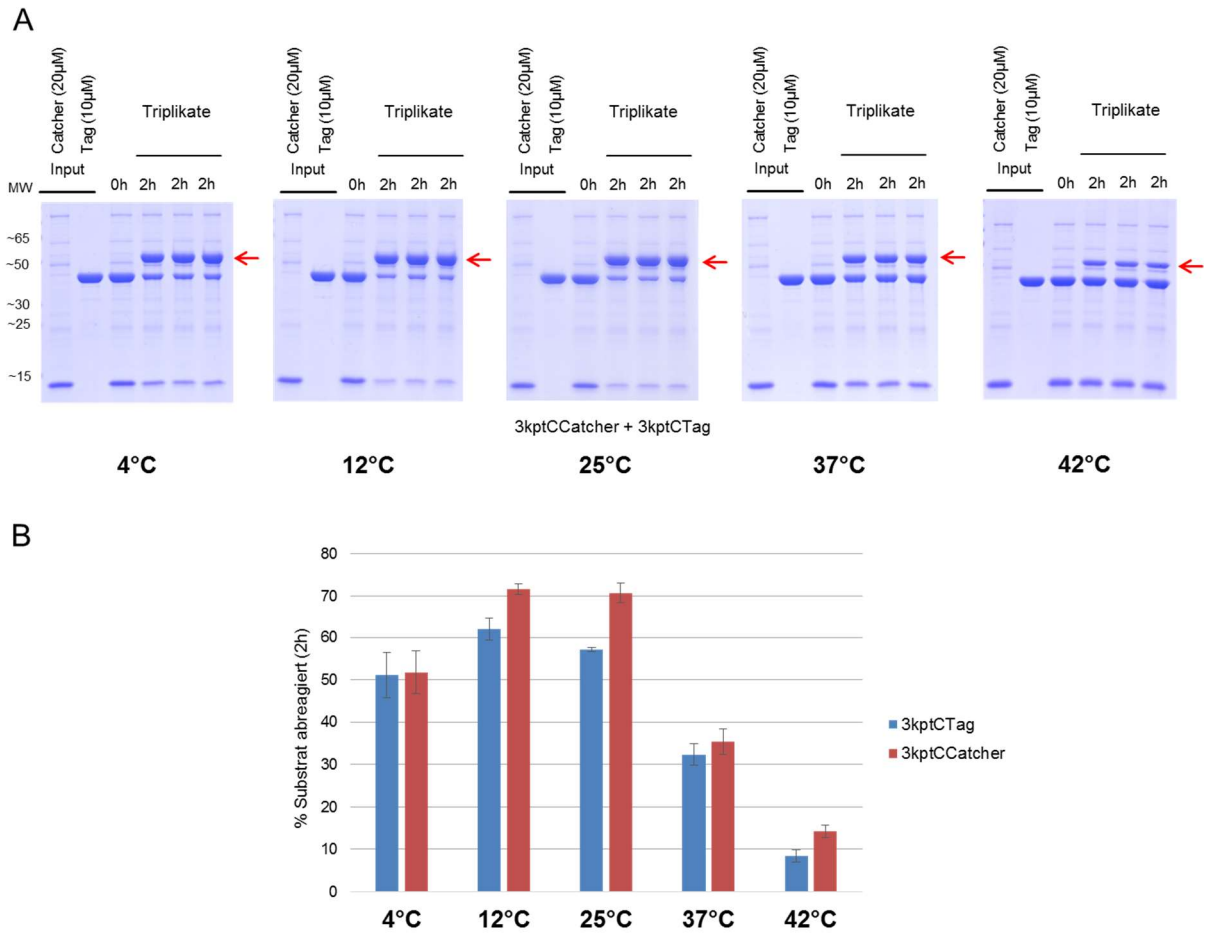


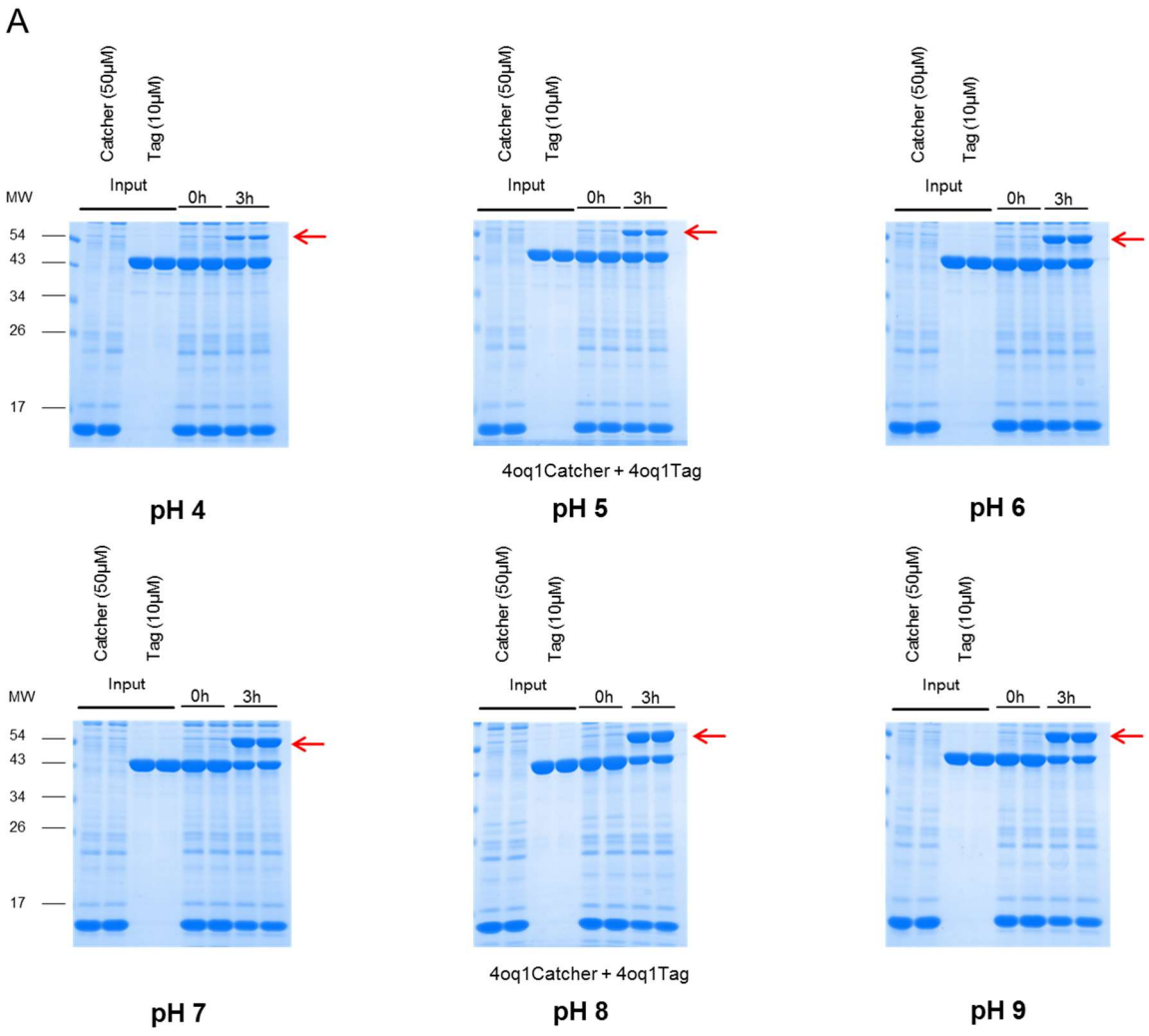
Abb.3.19: Temperaturabhängigkeit der kovalenten Reaktion zwischen 3kptCCatcher und 3kptCTag. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 2:1 Catcher/Tag Verhältnis in 1xPBS pH 7,5 Reaktionspuffer unter Zugabe von 40% Glycerin gemischt und bei verschiedenen Temperaturen (4°C bis 42°C) inkubiert. (A) Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. (B) Das Diagramm spiegelt die densitometrische Auswertung der Proteingele unter A wider und zeigt die Mittelwerte und Standardabweichungen der in Triplikaten durchgeführten Reaktionen 2h nach Reaktionsstart. Als Maß der Rekonstitutionseffizienz wurde die Abreaktion von 3kptCCatcher und 3kptCTag im Vergleich zu den Input-Spuren (Input-Banden als 100% Referenz) quantifiziert. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

3.2.4.4 pH-Abhängigkeit der Catcher-Tag Reaktion

Als nächstes sollte die pH-Wert Abhängigkeit der kovalenten Catcher-Tag Reaktionen für das 4oq1- und das 3kptC-System untersucht werden. Jedes Protein besitzt ein pH-Optimum, bei dem es maximal stabil und aktiv ist (Talley und Alexov, 2011, Alexov, 2004). Die Struktur, Stabilität, Löslichkeit, Aktivität und Funktion eines Proteins ist von dessen Nettoladung und dem Ionisierungsgrad der einzelnen reaktiven Gruppen abhängig, was wiederum vom pH-Wert der Proteinumgebung und den pK_S-Werten der Aminosäuren beeinflusst wird (Pace et al., 2009, Alexov, 2004, Talley und Alexov, 2011). Speziell für die Isopeptid-Triade und die Ausbildung einer Isopeptidbindung ist der Ionisierungsgrad der reaktiven Gruppen

entscheidend (siehe Einleitung Kapitel 1.2.2.2). Deshalb sollte getestet werden, inwieweit der pH-Wert des Reaktionspuffers die Rekonstitutionseffizienz der neuen Split-Isopeptid Systeme beeinflusst.

Vergleichende Rekonstitutionsexperimente mit 4oq1Catcher und 4oq1Tag, bei denen ausschließlich der pH-Wert des Reaktionspuffers (1xSPG Puffer) geändert wurde, ergaben, dass die kovalente Reaktion zwischen 4oq1Catcher und 4oq1Tag bei allen getesteten pH-Werten ablaufen kann, aber mit steigenden pH-Werten deutlich effizienter wird (siehe Zunahme der Produktbanden oder Abnahme der Input-Banden in Abb.3.20 A). Am effizientesten ist die Reaktion bei pH-Werten zwischen 7 und 9, da drei Stunden nach Reaktionsstart ca. 50% des 4oq1Tags und 25% des 4oq1Catchers zu einem kovalenten Komplex abreagiert haben (siehe Abb.3.20 A und B). In sauren pH-Bereichen (pH 4 oder pH 5) haben nach dreistündiger Inkubation dagegen etwa nur 15% des 4oq1Tags und 5% des 4oq1Catchers abreagiert (siehe Abb.3.20 A und B).



B

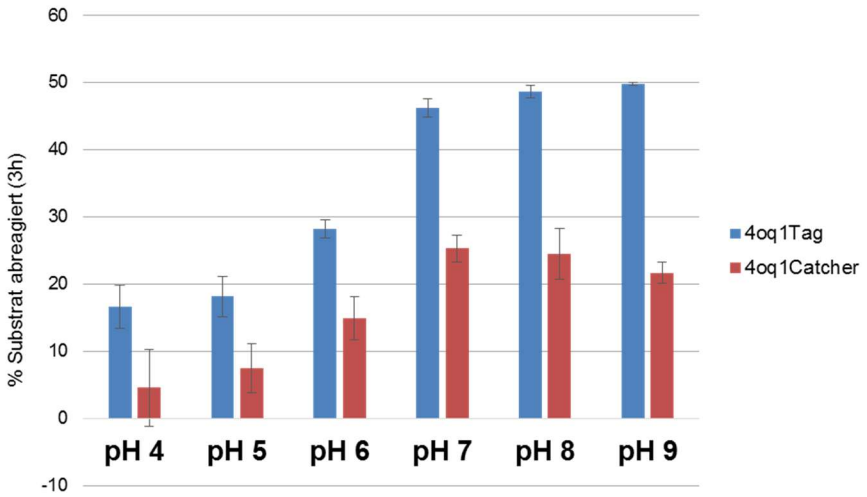


Abb.3.20: Analyse der pH-Abhängigkeit der kovalenten Reaktion zwischen 4oq1Catcher und 4oq1Tag. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher/Tag Verhältnis in 1xSPG Reaktionspuffer mit verschiedenen pH-Werten (pH 4 bis 9) gemischt und bei 25°C inkubiert (A). Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. (B) Das Diagramm spiegelt die densitometrische Auswertung der Proteingele unter A wider und zeigt die Mittelwerte und Standardabweichungen der in Duplikaten durchgeführten Reaktionen 3h nach Reaktionsstart. Als Maß der Rekonstitutionseffizienz wurde die Abreaktion von 4oq1Catcher und 4oq1Tag im Vergleich zu den Input-Spuren (Input-Banden als 100%) quantifiziert. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

Das 3kptC-System funktionierte ebenfalls bei allen getesteten pH-Werten (pH 4 - 9), wobei die kovalente Reaktion, wie auch beim 4oq1-System, bei neutralen bis basischen pH-Werten (pH 7 - 9) bevorzugt ablief (siehe Abb.3.21 A). Bei pH 4 in Abbildung 3.21 A ist zwar eine neue Produktbande im Vergleich zur 0h-Probe im Proteingel zu sehen, jedoch konnten bei der densitometrischen Auswertung keine detektierbaren Mengen an abreagiertem 3kptCTag oder 3kptCCatcher, verglichen mit den Input-Banden, quantifiziert werden, was für eine sehr ineffiziente Reaktion im sauren Bereich spricht (siehe Abb.3.21 B). Die kovalente Reaktion zwischen 3kptCCatcher und 3kptCTag verbesserte sich mit steigenden pH-Werten, wobei das pH-Optimum bei neutralem pH 7 erreicht wurde. 77% des 3kptCTags und 48% des 3kptCCatchers haben bei pH 7 nach zweistündiger Inkubation kovalent miteinander reagiert (siehe Abb.3.21). Erwähnenswert ist die Tatsache, dass die kovalente Catcher-Tag Rekonstitution des 3kptC-Systems in einem breiten pH-Bereich von pH 6 bis 9 ähnlich effizient ablief.

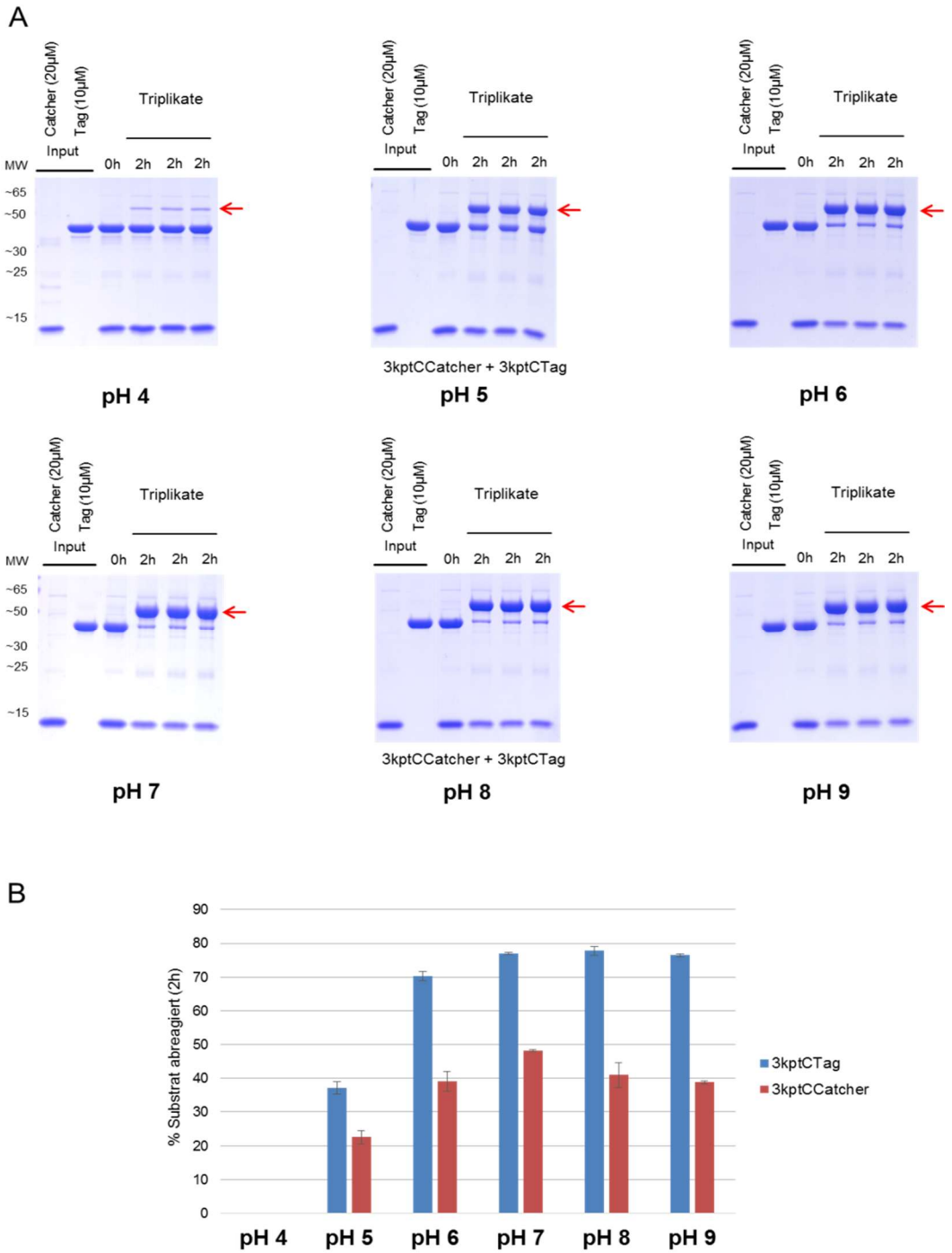


Abb.3.21: Analyse der pH-Abhängigkeit der kovalenten Reaktion zwischen 3kptCCatcher und 3kptCTag. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 2:1 Catcher/Tag Verhältnis in 1xSPG Reaktionspuffer mit verschiedenen pH-Werten (pH 4 bis 9) und unter Zugabe von 40% Glycerin gemischt und bei 25°C inkubiert (A). Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. (B) Das Diagramm spiegelt die densitometrische Auswertung der Proteingele unter A wider und zeigt die Mittelwerte und Standardabweichungen der in Triplikaten durchgeführten Reaktionen 2h nach Reaktionsstart. Als Maß der Rekonstitutionseffizienz wurde die Abreaktion von 3kptCCatcher und 3kptCTag im Vergleich zu den Input-Spuren (Input-Banden als 100% Referenz) quantifiziert. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

Das Spy-System hat ein pH-Optimum von pH 5 und die Rekonstitutionseffizienz nimmt mit steigenden pH-Werten (pH 8) ab (Zakeri et al., 2012). Anders verhält es sich mit den neu identifizierten Split-Isopeptid Systemen 4oq1 und 3kptC, deren Reaktionsrate mit steigenden, basischen pH-Werten im Hinblick auf die Bildung von kovalenten Catcher-Tag Komplexen verbessert wird (siehe oben). Durch Ausnutzung der unterschiedlich effizienten Reaktionen bei unterschiedlichen pH-Werten sollte untersucht werden, ob die unter Kapitel 3.2.3 beobachtete Kreuzreaktion von 4oq1Tag bzw. 3kptCTag mit dem SpyCatcher-Protein eventuell minimiert und die Spezifität der Systeme dadurch verbessert werden kann. Diese Theorie sollte in vergleichenden *in vitro* Rekonstitutionsexperimenten sowohl bei pH 7,5 als auch bei pH 9 analysiert werden. Rein von der qualitativen, visuellen Analyse der Coomassie-Gelbilder konnte keine eindeutige Minimierung der Kreuzreaktion zwischen SpyCatcher und 4oq1Tag bzw. zwischen SpyCatcher und 3kptCTag detektiert werden (siehe Abb.3.22). Bei jeder getesteten Interaktion (siehe Abb.3.22 A - D) konnten sowohl bei pH 7,5 als auch bei pH 9 neue, ähnlich stark ausgeprägte Produktbanden im Vergleich zur 0h-Probe detektiert werden, die auch im weiteren Reaktionsverlauf zunahmen. Grund dafür dürfte die enorm hohe Reaktivität des SpyCatcher-Proteins unter nahezu allen Reaktionsbedingungen sein (Zakeri et al., 2012), sodass selbst vermeintlich ungünstigere Konditionen die Kreuzreaktion nicht minimieren können. Um eindeutige und aussagekräftigere Ergebnisse zu erzielen, müssten Triplikate der Interaktionen mit anschließender densitometrischer Auswertung durchgeführt werden.

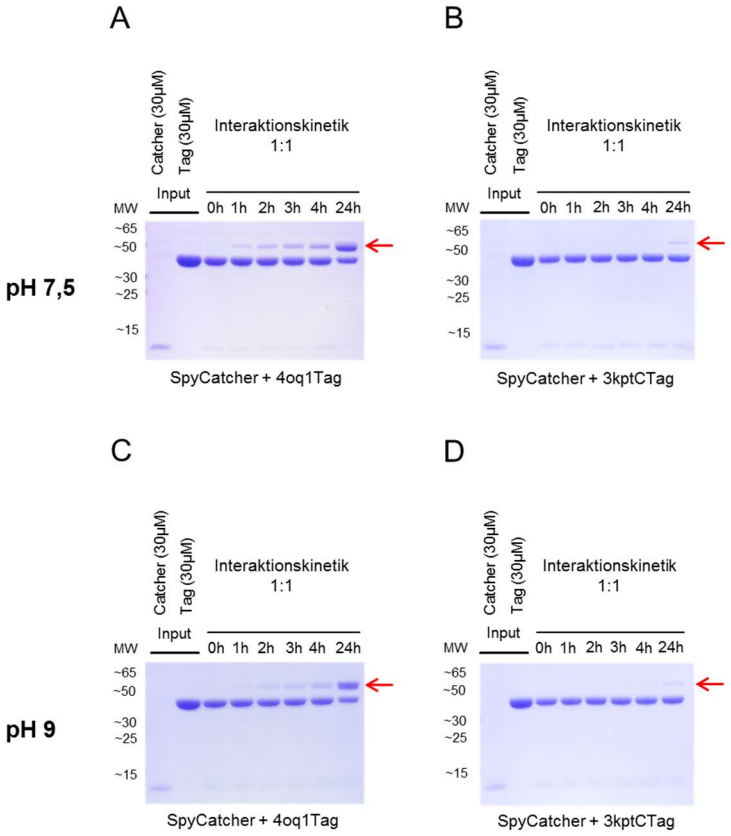
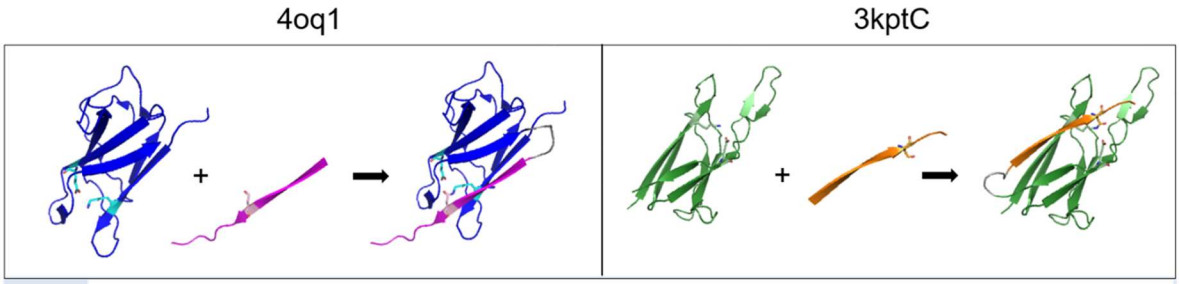


Abb.3.22: Minimierung der 4oq1Tag- bzw. 3kptCTag-Kreuzreaktion mit dem SpyCatcher-Protein durch unterschiedliche pH-Bedingungen. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem äquimolaren Verhältnis (15 µM Catcher- und Tag-Endkonzentration) gemischt und bei 25°C inkubiert, wobei der 1xSPG Reaktionspuffer einmal pH 7,5 (A und B) und einmal pH 9 (C und D) hatte. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

3.3 Optimierung der neuen Split-Isopeptid Systeme durch rationale, gerichtete Mutagenese

Aufgrund der schlechten Rekonstitutionseffizienz (siehe Kapitel 3.2.2) der neuen kovalenten Split-Isopeptid Systeme, 4oq1 und 3kptC, vor allem im Vergleich mit den etablierten Systemen, Spy und Snoop, wurden vielfältige rationale Optimierungsstrategien im Zuge des ,,Protein Engineerings“ unternommen. Zuerst sollten durch Verlängerungen und Verkürzungen der Catcher- bzw. Tag-Konstrukte die initialen CnaB-Domänenspaltungen analysiert werden (siehe Abb.3.23). Im Fokus weiterer Optimierungen stand vor allem die Stabilisierung der größeren Catcher-Proteine durch spezifische, gerichtete Mutationen einzelner Aminosäuren *(site-directed mutagenesis)* (siehe Abb.3.23). Die rationale Vorgehensweise zur Optimierung der Proteine basierte dabei auf dem vorhandenen Wissen über die Sequenz, die 3D-Struktur (pdb-Kristallstrukturen der Domänen) und die Eigenschaften der Aminosäuren (siehe u.a. Tabelle A3 im Anhang), und wurde durch bereits durchgeführte Strategien zur Verbesserung der etablierten Catcher-Tag Systeme (Zakeri und Howarth, 2010, Zakeri et al., 2012, Veggiani et al., 2016) und geeignete bioinformatische Programme gestützt.



computergestütztes „rational mutagenesis design"

Optimierung der neuen kovalenten Split-Isopeptid Systeme durch N- und C-terminale Proteinvarianten (*splitting-site mutants*)

- N- und C-terminale Verlängerungen oder Verkürzungen der Produkte der initialen Catcher-Tag Spaltung der jeweiligen CnaB-Domänen (D2-Domäne für 4oq1 und CNA3-Domäne für 3kptC)

Optimierung der neuen kovalenten Split-Isopeptid Systeme durch Einzel-Aminosäureaustausche v.a. zur Stabilisierung der Catcher-Proteine

- Austausch von exponierten, hydrophoben Aminosäuren an der Proteinoberfläche durch hydrophilere Aminosäuren
- Strukturbasierte Aminosäure-Mutationen (z.B. Prolin- und Glycin-Substitutionen)
- Mutation besonders flexibler Aminosäuren (B-Faktor)
- PROSS (*Protein Repair One Stop Shop*); Mutation mehrerer Aminosäuren gleichzeitig

Catcher-Mutanten

- ,,SpyTag“-Mutation → Asparaginsäure (D) statt reaktivem Asparagin (N) in 4oq1Tag bzw. 3kptCTag

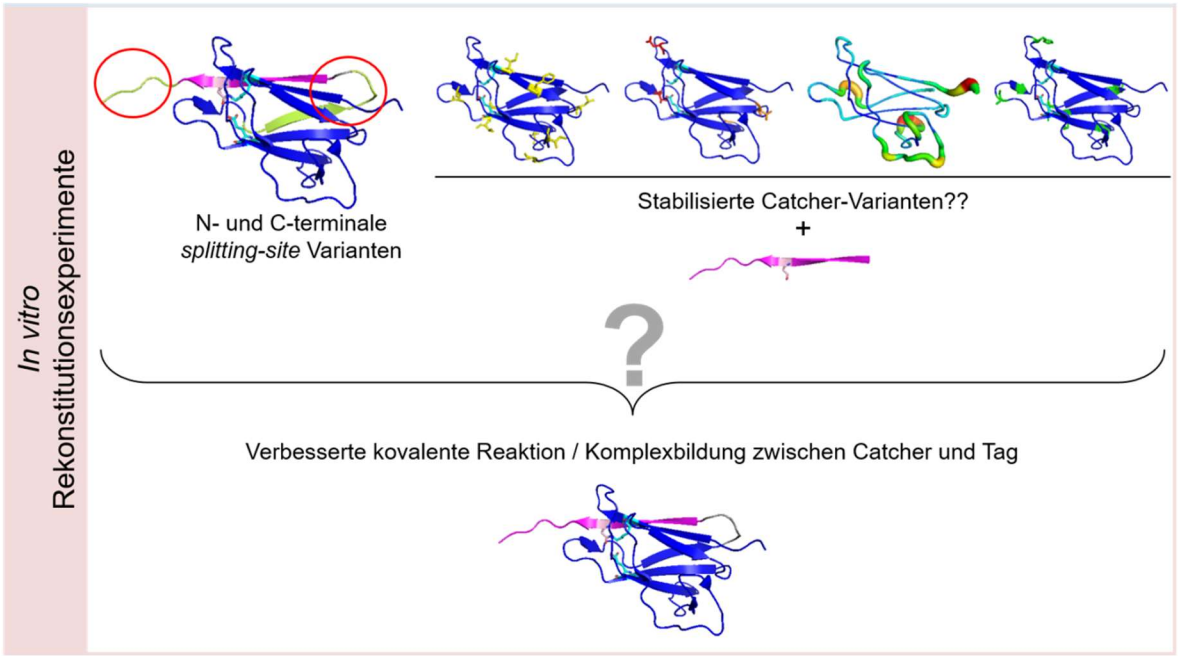


Abb.3.23: Zusammenfassendes, vereinfachtes Schema der Vorgehensweise zur Optimierung der neuen, kovalenten Split-Isopeptid Systeme 4oq1 und 3kptC. Die verschiedenen Strategien sind beispielhaft für das 4oq1-System gezeigt, wurden aber analog für das 3kptC-System angewendet. Zur Verbesserung der Rekonstitutionseffizienz wurden neben N- und C-terminalen *splitting-site* Varianten (zu mutierende N- und C-terminale Proteinbereiche sind durch rote Kreise gekennzeichnet) der Catcher- und Tag-Proteine gezielte Aminosäuresubstitutionen vor allem zur Stabilisierung der Catcher-Proteine getestet. Zu mutierende Aminosäuren des 4oq1Catcher-Proteins sind farbig hervorgehoben (gelb: exponierte hydrophobe Aminosäuren an der Proteinoberfläche, rot: Aminosäuren mit einem bestimmten phi-Winkel in *loops*, die zu Prolinen mutiert wurden, Farbcode der B-Faktor Werte, wobei rot für besonders hohe B-Werte und daher besonders flexible Bereiche der Proteinstruktur steht, grün: PROSS-vorhergesagte stabilisierende Mutationen). Abgesehen von den PROSS-Catcher-Mutanten wurden ausschließlich Einzel-Aminosäureaustausche durchgeführt, um die Auswirkung jeder einzelnen Aminosäuremutation auf die kovalente Rekonstitution von Catcher und Tag zu testen. Die veränderten reaktiven Tag-Mutanten (N→D) sind in diesem Schema nicht gezeigt, wurden aber ebenfalls in kovalenten Rekonstitutionsanalysen mit ihren entsprechenden Catcher-Reaktionspartnern getestet.

3.3.1 Optimierung der initialen Catcher-Tag Spaltungsstelle (verlängerte und verkürzte Catcher- und Tag-Proteinvarianten)

In vitro Rekonstitutionsexperimente ergaben, dass die kovalente Reaktion zwischen den Reaktionspartnern 4oq1Catcher und 4oq1Tag bzw. 3kptCCatcher und 3kptCTag nicht sehr effizient ist (Kapitel 3.2.2). Dies könnte zum einen an ungünstigen Reaktionsbedingungen liegen, weshalb die Interaktionen auch biochemisch charakterisiert wurden (siehe 3.2.4) oder an den initialen Catcher- und Tag-Produkten der CnaB-Domänenspaltung (siehe Abb.3.1). Die initiale Spaltung der CnaB-Domänen (D2-Domäne für 4oq1 und CNA3-Domäne für 3kptC) in einen Catcher- und einen Tag-Teil wurde basierend auf bioinformatischen Analysen, u.a. unter Beachtung der stabilisierenden, molekularen Wechselwirkungen wie Van-der-Waals Kräfte und Wasserstoffbrückenbindungen, vorgenommen (Pröschel et al.,

2017). Jedoch ist die genaue und optimale Definierung der Grenze zwischen Catcher und Tag *in silico* kritisch zu bestimmen und kann nur experimentell durch schrittweise Modifikationen der initialen Aminosäuresequenzen genauer ermittelt werden. Um nun zu testen, ob die Effizienz der Reaktion zwischen Catcher und Tag durch unterschiedlich N- und C- terminal verkürzte oder verlängerte Proteinvarianten verbessert werden kann, wurden Catcher- und Tag-Varianten, ebenfalls basierend auf bioinformatischen Analysen, über ortsspezifische Mutagenese hergestellt (Klonierungen siehe Kapitel 2.6.2 und Tabelle 1.1, Primersequenzen sind im Anhang in Tabelle A2 hinterlegt) und vergleichende *in vitro* Rekonstitutionsexperimente durchgeführt (siehe Kapitel 2.6.4.9). Verglichen wurde die Effizienz der Reaktionen der modifizierten Systeme immer mit dem jeweiligen Catcher-Tag Wildtyp-System, bestehend aus den Produkten der initialen Domänenspaltung (siehe Abb.3.1).

3.3.1.1 N- und C- terminal modifizierte 4oq1Tag-Varianten

Basierend auf bioinformatischen Analysen (Prof. Dr. Heinrich Sticht, Dr. Anselm Horn), vor allem unter Beachtung der molekularen Wechselwirkungen zwischen 4oq1Catcher und 4oq1Tag, wurden zwei N-terminal verlängerte 4oq1Tag-Varianten hergestellt (4oq1Tag-L (#106 in Tabelle 1.1) und 4oq1Tag-HQL (#104 in Tabelle 1.1), siehe Abb.3.24 A II & III), da die N-terminalen Reste des Tags, laut *in silico* Analysen, Wasserstoffbrücken ausbilden. Diese zwei N-terminal verlängerten 4oq1Tag-Varianten sollten die Interaktion zwischen Catcher und Tag durch die Ausbildung zusätzlicher Wechselwirkungen verbessern (Pröschel et al., 2017). Neben den verlängerten Varianten wurden auch zwei C-terminal verkürzte 4oq1Tag-Varianten hergestellt (siehe Abb.3.24 A IV & V), um zu sehen welche Aminosäuren essentiell für die Ausbildung eines kovalenten Catcher-Tag Komplexes sind und auf welche Aminosäuren in Zukunft bei der Entwicklung neuer, vielleicht auch noch kürzerer, Tag-Peptide verzichtet werden kann. Die Klonierungen der DNA-Konstrukte mit sequenzspezifischen Primern (siehe Tabelle A2 im Anhang) erfolgte wie unter Kapitel 2.6.2 und in Tabelle 1.1 beschrieben. Vergleichende *in vitro* Rekonstitutionsexperimente ergaben, dass beide N-terminal verlängerten 4oq1Tags (4oq1Tag-L und 4oq1Tag-HQL) zu einer leichten Verbesserung der Bildung eines kovalenten Catcher-Tag Komplexes (stärkere Produktbande, vor allem zu den Zeitpunkten 1h - 4h nach Reaktionsstart) im Vergleich zu dem Wildtyp-System führten (siehe Abb.3.24 B I - III). Um die Produktbande deutlicher sehen und auch quantifizieren zu können, wurde das mCherry-4oq1Catcher Fusionsprotein verwendet und analoge Rekonstitutionsexperimente mit den jeweiligen 4oq1Tag-Varianten durchgeführt (siehe Abb.3.25 B I - III). Die Interaktionen wurden in Triplikaten durchgeführt und die Produktbanden anschließend mittels ImageJ densitometrisch analysiert und quantifiziert (Kapitel 2.6.4.4). Die Ergebnisse der Densitometrie (siehe Abb.3.26) bestätigten

die zuvor gemachten Beobachtungen, nämlich dass die N-terminal verlängerten 4oq1Tag-Varianten die Effizienz der kovalenten Reaktion zwischen Catcher und Tag im Vergleich zum Wildtyp-System verbesserten. So war die Zunahme der kovalenten Catcher-Tag Produkte im Vergleich zu der 0h-Probe innerhalb des Zeitraums 1h bis 4h nach Inkubationsstart stets höher als die des Wildtyp-Systems. Die Level der Produktzunahme der verschiedenen Systeme glichen sich jedoch 24h nach Inkubationsstart an. Folglich lief die kovalente Reaktion zwischen 4oq1Catcher und den N-terminal verlängerten 4oq1Tag-Varianten effizienter, das heißt schneller ab, da im Vergleich zum Wildtyp-System mehr kovalentes Produkt zu frühen Reaktionszeitpunkten (1h - 4h) detektiert wurde. Die Erweiterung des 4oq1Tags um den Rest L245 scheint dabei entscheidend zur Verbesserung der Reaktivität des 4oq1-Systems beizutragen, da sich die um die Reste L245 oder H243-L245 verlängerten 4oq1Tag-Varianten bezüglich ihrer Rekonstitutionseffizienz kaum unterschieden (siehe Abb.3.24 – 3.26). Die C-terminal verkürzten 4oqTag-Varianten dagegen, führten zu keinem Zeitpunkt zu einer Verbesserung der kovalenten Produktbildung im Vergleich zum Wildtyp-System (siehe Abb.3.24 I, IV und V). Die 4oq1Tag-Verkürzung um die Reste L255 - N259 (#103 in Tabelle 1.1) verschlechterte die Reaktion sogar erheblich, weshalb die Rekonstitutionsexperimente zur Quantifizierung mit dem mCherry-4oq1Catcher Konstrukt nur mit der RGN verkürzten 4oq1Tag-Version (#105 in Tabelle 1.1) wiederholt wurden (siehe Abb.3.25 B IV). Auch die densitometrische Analyse der dreifach unabhängig durchgeführten Interaktionen zwischen mCherry-4oq1Catcher und 4oq1Tag-ΔRGN ergab keine verbesserte kovalente Produktausbeute im Vergleich mit dem 4oq1-Wildtyp-System (siehe Abb.3.26). Folglich scheinen die Reste L255 und P256 besonders wichtig für die Ausbildung von kovalenten Catcher-Tag Komplexen zu sein, da die Deletion der Reste RGN keinen Effekt auf die Rekonstitutionseffizienz verglichen mit dem Wildtyp-System hatte, wohingegen die 4oq1Tag-ΔLPRGN Mutante fast zu einem kompletten Verlust der kovalenten Reaktion zwischen Catcher und Tag führte (Abb.3.24 IV und V).

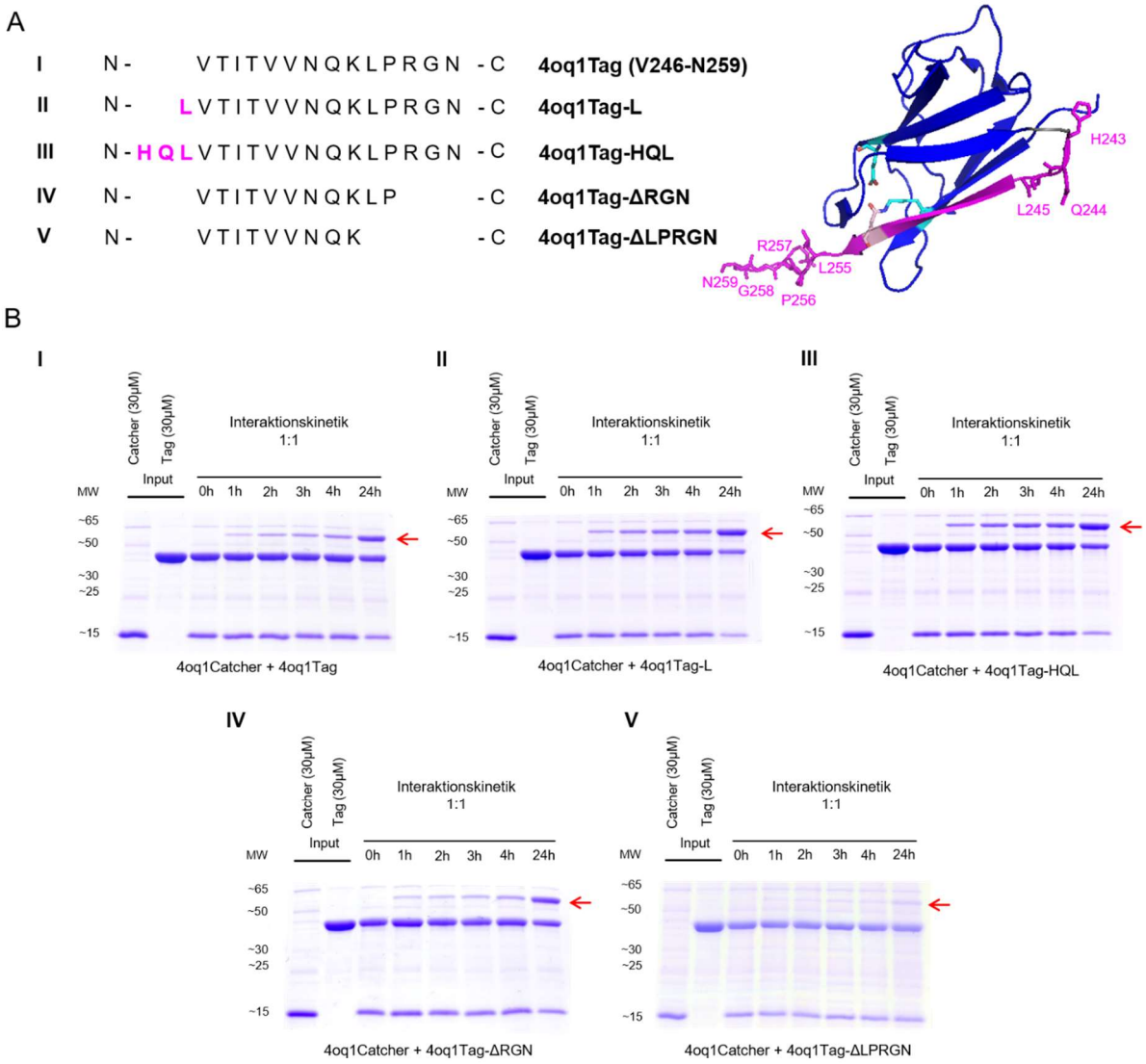


Abb.3.24: Analyse der Rekonstitutionseffizienz von N- und C-terminalen 4oq1Tag-Varianten in der Reaktion mit 4oq1Catcher. (A) Sequenzvergleiche der verschiedenen 4oq1Tag-Varianten (I: 4oq1Tag-Wildtyp (Reste V246 - N259), II: N-terminale L245 Verlängerung des 4oq1Tags, III: N-terminale Verlängerung des 4oq1Tags um die Reste H243 bis L245, IV: C-terminale Verkürzung des 4oq1Tags um die Reste R257 bis N259, V: C-terminale Verkürzung des 4oq1Tags um die Reste L255 bis N259). Die Molekülvisualisierung wurde mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag). (B) Rekonstitutionsexperimente zwischen 4oq1Catcher und verschiedenen 4oq1Tag-Varianten (I bis V). Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem äquimolaren Verhältnis (15 µM Catcher- und Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

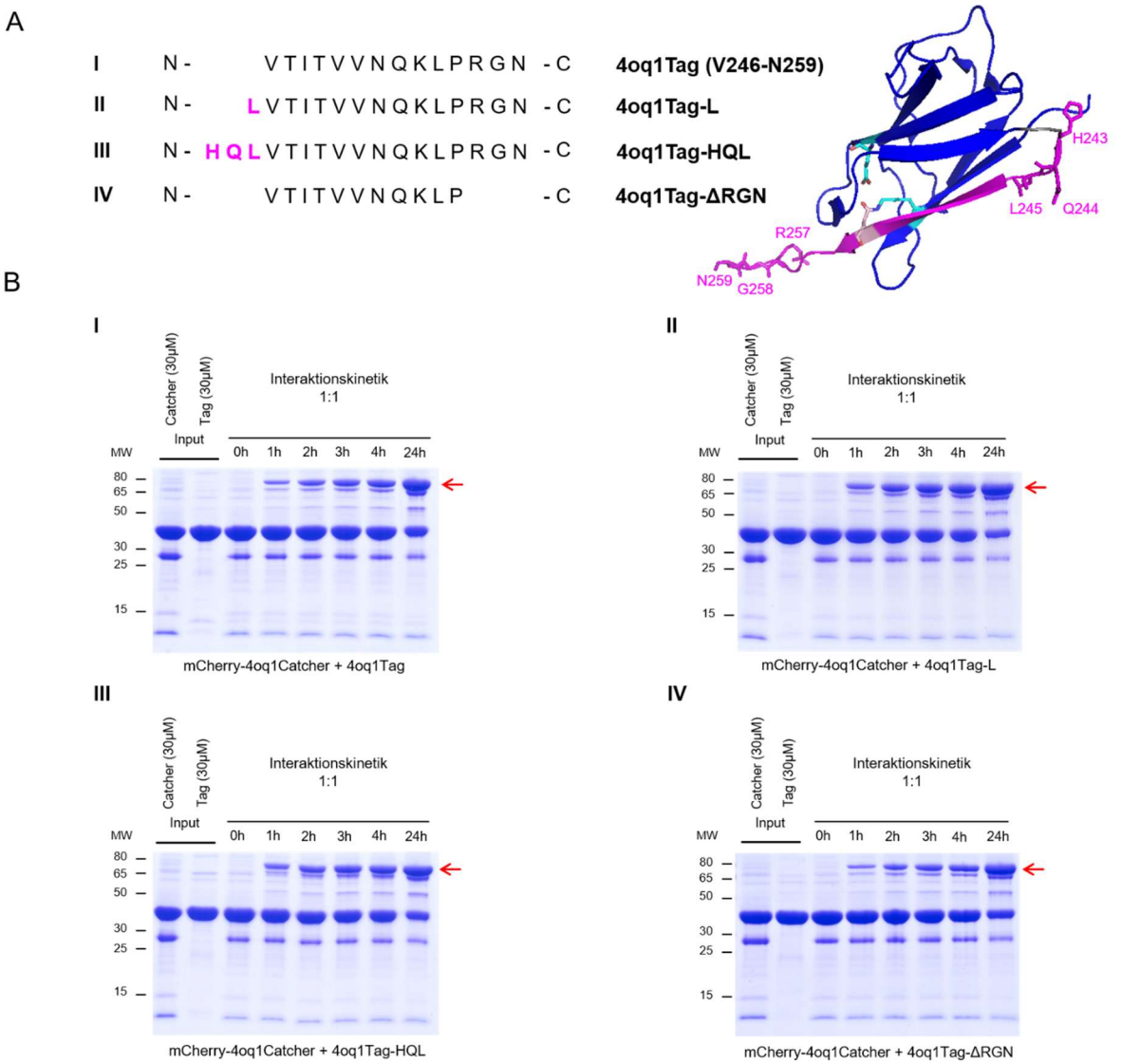


Abb.3.25: Analyse der Rekonstitutionseffizienz von N- und C-terminalen 4oq1Tag-Varianten in der Reaktion mit mCherry-4oq1Catcher. (A) Sequenzvergleiche der verschiedenen 4oq1Tag-Varianten (I: 4oq1Tag-Wildtyp (Reste V246 - N259), II: N-terminale L245 Verlängerung des 4oq1Tags, III: N-terminale Verlängerung des 4oq1Tags um die Reste H243 bis L245, IV: C-terminale Verkürzung des 4oq1Tags um die Reste R257 bis N259). Die Molekülvisualisierung wurde mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag). (B) Rekonstitutionsexperimente zwischen mCherry-4oq1Catcher und verschiedenen 4oq1Tag-Varianten (I bis IV). Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem äquimolaren Verhältnis (15 µM Catcher- und Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

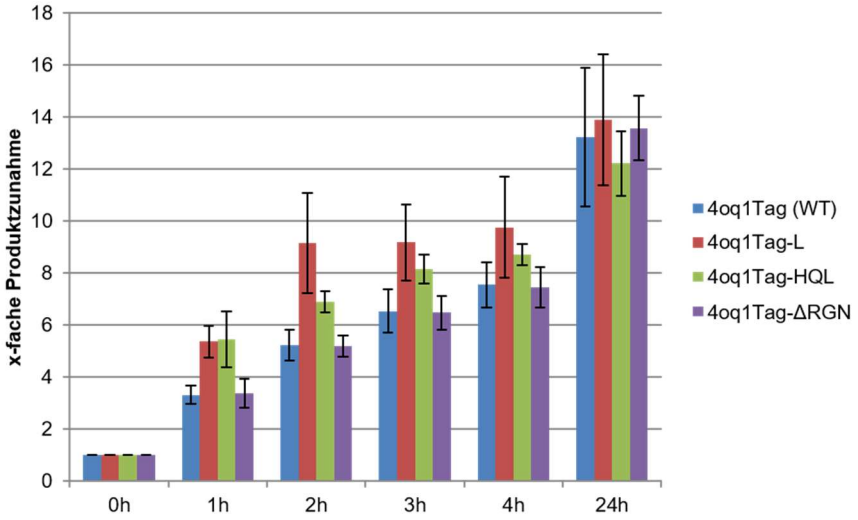


Abb.3.26: Quantifizierung der Zunahme an kovalentem Produkt, resultierend aus der Reaktion zwischen mCherry-4oq1Catcher und verschiedenen 4oq1Tag-Varianten mittels Densitometrie. Zu quantifizierende Reaktionen wurden in Triplikaten durchgeführt (Bedingungen siehe Bildunterschrift in Abb.3.25) und die Produktbanden der Proteingele anschließend densitometrisch mittels ImageJ analysiert (0h-Probe als Referenzpunkt =1). Gezeigt sind die Mittelwerte und Standardabweichungen aus drei unabhängig durchgeführten Rekonstitutionsexperimenten.

Nachdem gezeigt werden konnte, dass die N-terminal verlängerten 4oq1Tag-Varianten (4oq1Tag-L und 4oq1Tag-HQL) einen positiven Einfluss auf die Effizienz der Bildung eines kovalenten Catcher-Tag Komplexes haben, sollte untersucht werden, ob die verlängerten 4oq1Tag-Varianten auch einen positiven Einfluss auf die SpyCatcher-Kreuzreaktion haben, das heißt in geringerem Maße mit dem SpyCatcher-Protein unspezifische, kovalente Komplexe bilden. Dazu wurden vergleichende Rekonstitutionsanalysen mit allen notwendigen Kontrollen durchgeführt (Abb.3.27 A - D: Positivkontrollen). Abbildung 3.27 zeigt, dass die N-terminal verlängerten 4oq1Tag-Varianten zu keiner deutlichen Verbesserung der Spezifität der kovalenten Interaktion führten. Auch die verlängerten 4oq1Tag-Peptide kreuzreagierten mit dem SpyCatcher-Protein und zwar in ähnlichem Maße wie der Wildtyp-4oq1Tag (ähnliche Intensitäten der Produktbanden siehe Abb.3.27). Eine Stunde nach dem Mischen von SpyCatcher und den verlängerten 4oq1Tag-Varianten konnte eine leichte Produktbande im Vergleich zur 0h-Kontrolle detektiert werden, die über den gesamten Reaktionsverlauf zunahm (siehe Abb.3.27 E - G). Zwar ist die Kreuzreaktion (siehe Abb.3.27 E - G) im Vergleich zu den systeminternen Catcher-Tag Reaktionen schwächer (siehe Abb.3.27 A - D), kann jedoch nicht ungeachtet bleiben. Während der SpyCatcher mit seinem SpyTag schon nach einstündiger gemeinsamer Inkubation eine dicke Produktbande aufwies (siehe Abbildung 3.27 A und siehe auch Kapitel 3.2.2), produzierte die Reaktion aus SpyCatcher und 4oq1Tag (Wildtyp und Varianten) nur eine sehr schwache Bande bei gleicher Inkubationszeit und unter gleichen Bedingungen.

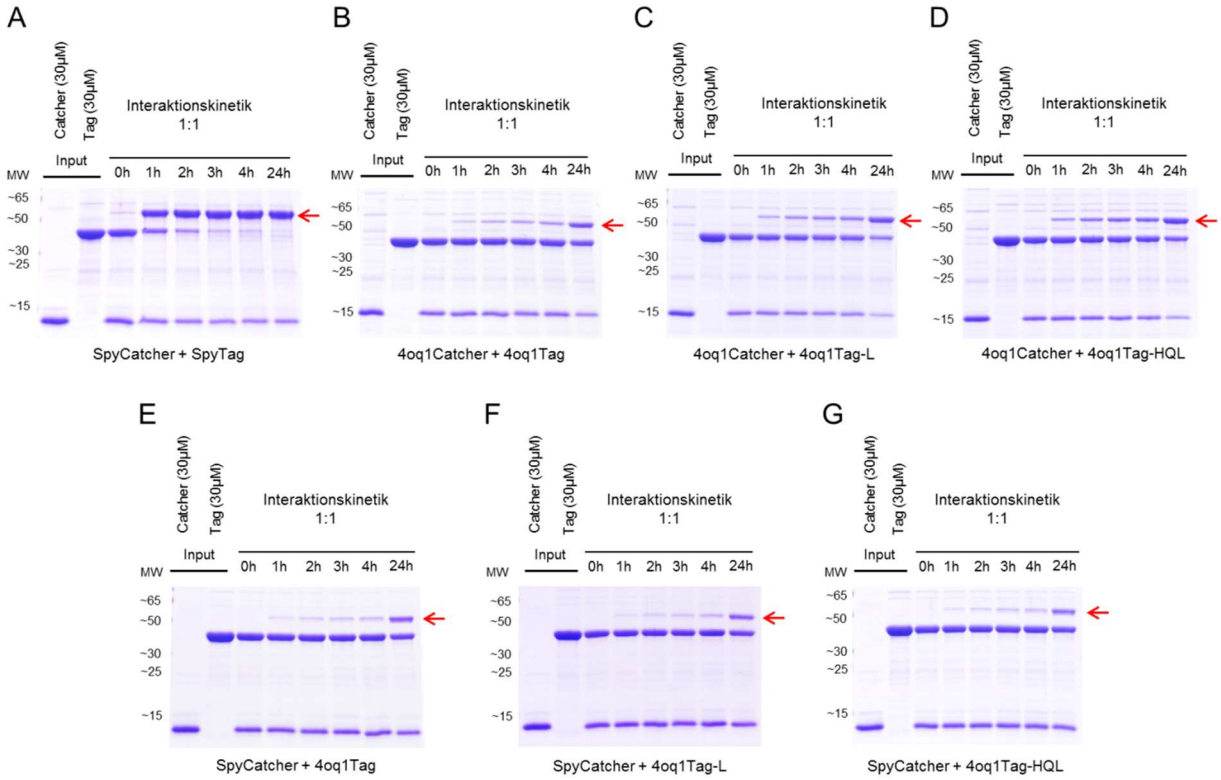


Abb.3.27: Analyse der Kreuzreaktivität der N-terminal verlängerten 4oq1Tag-Varianten in der Reaktion mit SpyCatcher. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem äquimolaren Verhältnis (15 µM Catcher- und Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

Um weitere stabilisierende Wechselwirkungen zwischen Catcher und Tag zuzulassen, wurde eine weitere 4oq1Tag-Variante, bei der ein kompletter, zweiter β-Faltblattstrang (L234 - L240: β-Strang) in die Tag-Sequenz integriert wurde, hergestellt (Klonierung siehe Kapitel 2.6.2 und Tabelle 1.1: #139) und in *in vitro* Rekonstitutionsexperimenten getestet (siehe Kapitel 2.6.4.9). Die Integration des kompletten β-Faltblattstranges führte zu einer N-terminalen Verlängerung des Wildtyp-4oq1Tags (V246 - N259) um die nativen Aminosäuren L234 - L245 der D2-Domäne des RrgC Proteins und resultierte in einem 26 Aminosäure langen 4oq1Tag (L234 - N259, siehe Abb.3.28 A). Verglichen wurde der Einfluss des zweiten Faltblattstranges im 4oq1Tag auf die Effizienz der Bildung eines kovalenten Catcher-Tag Komplexes mit den ,,besten", bisher ermittelten, 4oq1Tag- (4oq1Tag-HQL) und 4oq1Catcher- (4oq1Catcher-VDHQ, siehe Kapitel 3.3.1.2) Varianten. Die Effizienz der kovalenten Reaktion zwischen 4oq1Catcher und 4oq1Tag konnte durch die Integration des zweiten β-Faltblattstranges in die 4oq1Tag-Sequenz gesteigert werden (siehe Abb.3.28 B). Bei analogen Reaktionsbedingungen und stets im Vergleich zu dem Wildtyp-System (hier: 4oq1Tag-HQL und 4oq1Catcher-VDHQ) war sowohl die Catcher-Tag Komplex repräsentierende Produktbande im Gel als auch die Abnahme der Catcher- und Tag-Banden

stärker ausgeprägt, wenn der 4oq1Tag um den zweiten β-Strang verlängert wurde (siehe Abb.3.28 B I und II). Die Effizienzsteigerung der kovalenten Reaktion zwischen 4oq1Catcher und 4oq1Tag durch die Integration des zweiten β-Faltblattstranges konnte sowohl ohne als auch mit Zugabe von 1,5 M TMAO in den Reaktionspuffer (siehe Kapitel 3.2.4.1) eindeutig detektiert werden (siehe Abb.3.28 B I und II).

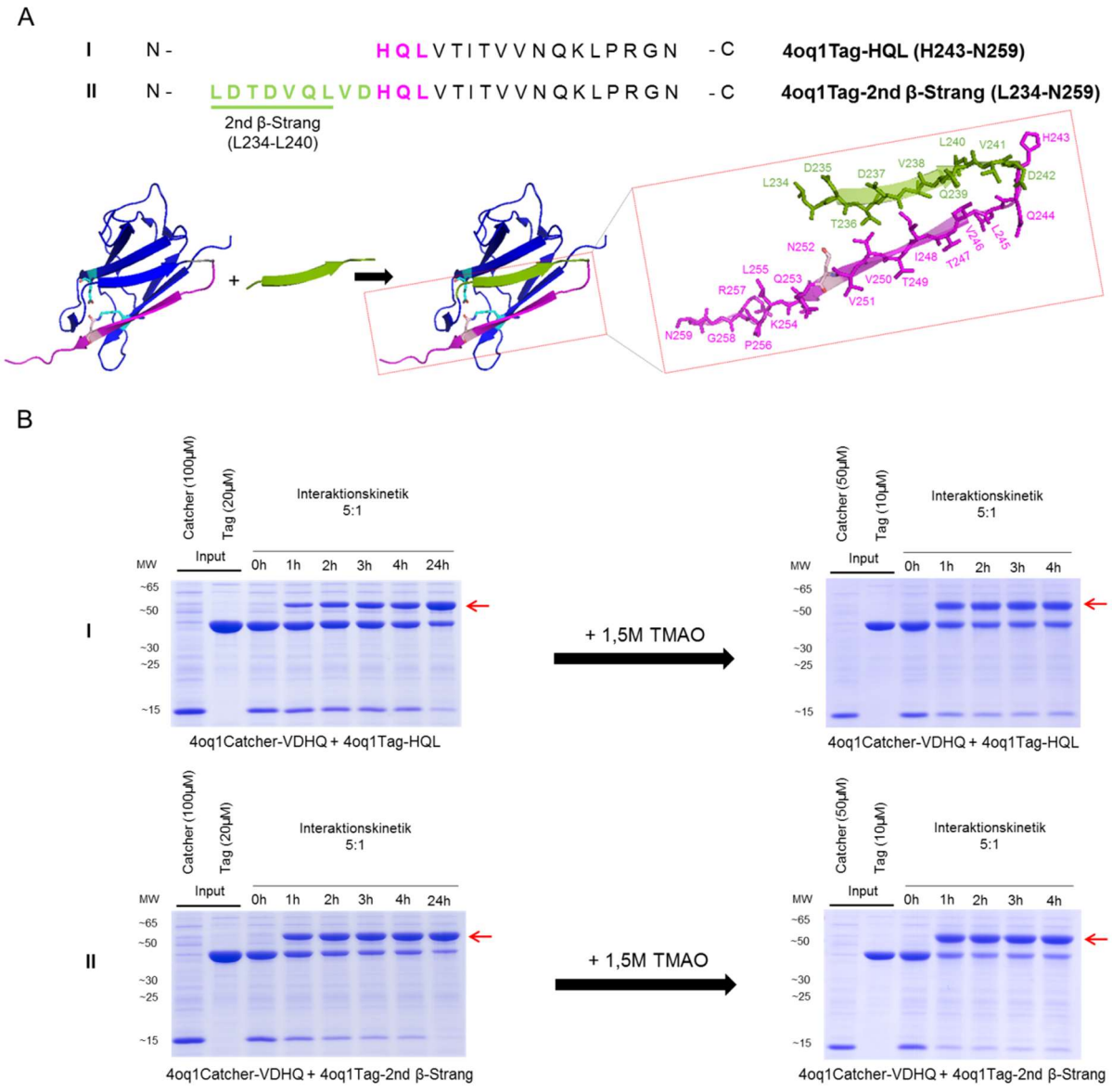


Abb.3.28: Untersuchung der Wirkung der Verlängerung des 4oq1Tags um einen weiteren β-Faltblattstrang auf die Effizienz der Ausbildung eines kovalenten 4oq1Catcher-4oq1Tag Komplexes. (A) Sequenzvergleiche der 4oq1Tag-Varianten (I: 4oq1Tag-HQL: H243-N259, II: 4oq1Tag-2nd β-Strang: L234-N259). Die Molekülvisualisierungen wurden mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag, grün: zusätzlicher β-Faltblattstrang). (B) Rekonstitutionsexperimente zwischen 4oq1Catcher-VDHQ und verschiedenen 4oq1Tag-Varianten (I und II) unter verschiedenen Bedingungen. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher zu Tag Verhältnis (50 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit und ohne Zugabe von 1,5M TMAO inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

Rational betrachtet sollte die drastische Verlängerung des 4oq1Tags um einen zweiten β-Faltblattstrang die Wahrscheinlichkeit einer spezifischen Reaktion zwischen 4oq1Tag und 4oq1Catcher, aufgrund zusätzlicher, spezifischer nicht-kovalenter Wechselwirkungen erhöhen. Aus diesem Grund sollte getestet werden, ob die Integration des β-Faltblattstranges in den 4oq1Tag die Spezifität verbessert und somit die vorher beobachtete Kreuzreaktion mit dem SpyCatcher- und dem 3kptCCatcher-Protein minimiert (siehe Kapitel 3.2.3). Die Positivkontrollen (siehe Abb.3.29 A - D) erbrachten die erwarteten kovalenten Catcher-Tag-Komplexe (rote Pfeile) und bestätigten die Funktionalität der einzelnen Systeme, auch wenn die Catcher-Proteine (v.a. 4oq1Catcher) nicht oder nur schwach im Proteingel zu sehen waren. Vergleicht man die Kreuzreaktionen des Wildtyp-4oq1Tags mit SpyCatcher bzw. 3kptCCatcher (Abb.3.29 E und G) mit den Reaktionen zwischen 4oq1Tag-2nd β-Strang und SpyCatcher bzw. 3kptCCatcher (Abb.3.29 F und H), konnten keine visuell detektierbaren Unterschiede weder bezüglich der Zunahme und Intensität der Produktbande noch bezüglich der Abnahme der Catcher- bzw. Tag-Substratbanden verzeichnet werden.

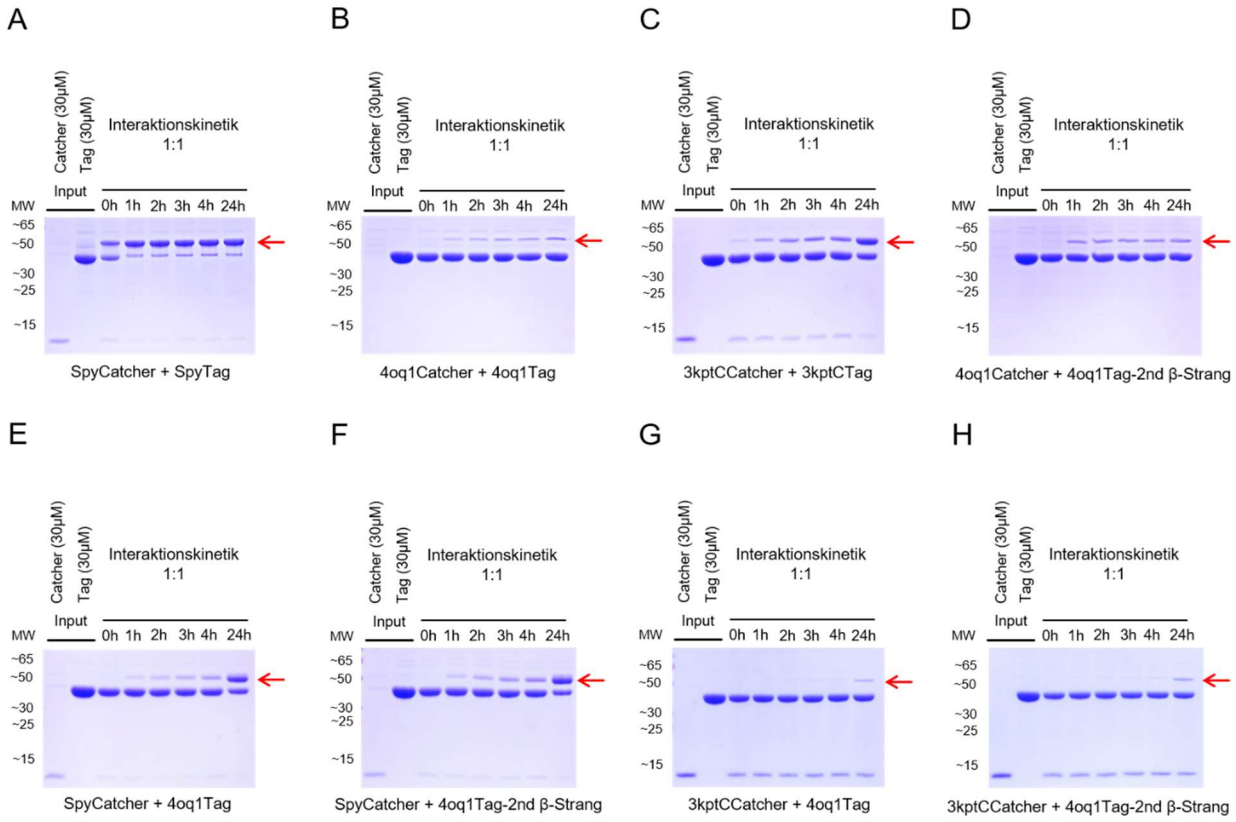


Abb.3.29: Analyse der Wirkung der Verlängerung des 4oq1Tags um einen weiteren β-Faltblattstrang auf die Spezifität der Ausbildung eines kovalenten 4oq1Catcher-4oq1Tag Komplexes. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem äquimolaren Verhältnis (15 µM Catcher- und Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

3.3.1.2 4oq1Catcher-Varianten

Mit dem Ziel, die kovalente Interaktion zwischen Catcher und Tag zu verbessern, wurde neben dem 4oq1Tag auch das initiale 4oq1Catcher-Spaltungsprodukt (T146 - L240, siehe Abb.3.1) in seiner Länge, beruhend auf *in silico* Analysen, verändert (siehe Abb.3.30 A, Klonierungen siehe 2.6.2 und Tabelle 1.1). In vergleichenden *in vitro* Rekonstitutionsexperimenten (siehe Kapitel 2.6.4.9) konnte gezeigt werden, dass der C-terminal verlängerte 4oq1Catcher-VDHQ (#102 in Tabelle 1.1) auf Wildtyp-Niveau oder evtl. etwas besser bezüglich der Ausbildung eines kovalenten Komplexes mit dem 4oq1Tag-Peptid als Reaktionspartner ist (siehe Abb.3.30 B I und II). Während in dem Zeitraum 1h bis 4h nach Reaktionsstart kaum Unterschiede in der Produktzunahme zu sehen waren, war der 4oq1Catcher-VDHQ 24 Stunden nach gemeinsamer Catcher-Tag Inkubation nahezu vollständig verbraucht, wohingegen der Wildtyp-4oq1Catcher noch eine deutlichere Bande im Proteingel aufwies (siehe Abb.3.30 B I und II). Der C-terminal verkürzte 4oq1Catcher-ΔQL (#107 in Tabelle 1.1) dagegen, verschlechterte die kovalente Reaktion zwischen Catcher und Tag enorm im Vergleich zum Wildtyp-System (siehe Abb.3.30 I und III), obwohl die Reste Q239 und L240 in *in silico* Analysen nur Fluktuationen und keine Sekundärstrukturen aufwiesen und deshalb auch entfernt wurden (Professur für Bioinformatik). Lediglich nach 24-stündiger, gemeinsamer Catcher-Tag Inkubation war eine sehr schwache Produktbande zwischen 4oq1Catcher-ΔQL und 4oq1Tag im Gel zu sehen, während die Abnahme der Substratbanden im Vergleich mit der 0h-Probe nicht detektierbar war (siehe Abb.3.30 B III).

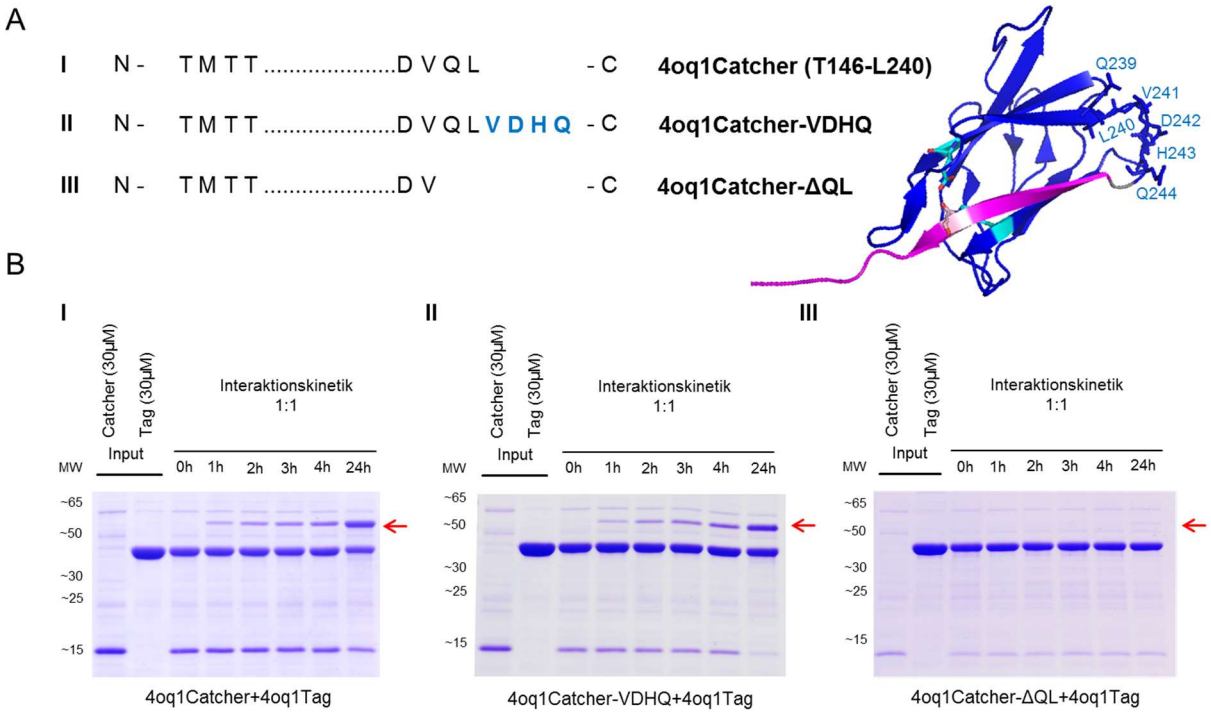


Abb.3.30: Analyse der Rekonstitutionseffizienz von 4oq1Catcher-Varianten in der Reaktion mit 4oq1Tag. (A) Sequenzvergleiche der verschiedenen 4oq1Catcher-Varianten (I: 4oq1Catcher-Wildtyp

(Reste T146 - L240), II: C-terminale Verlängerung des 4oq1Catchers um die Reste V241 - Q244, III: C-terminale Verkürzung des 4oq1Catchers um die Reste Q239 und L240). Die Molekülvisualisierung wurde mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag). (B) Rekonstitutionsexperimente zwischen verschiedenen 4oq1Catcher-Varianten und 4oq1Tag (I bis III). Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem äquimolaren Verhältnis (15 µM Catcher- und Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

3.3.1.3 3kptCTag-Varianten

Die initiale Spaltung der CNA3-Domäne des 3kptC-Systems in ein Catcher-Protein und ein Tag-Peptid (siehe Abb.3.1) sollte ebenfalls mithilfe von bioinformatischen Analysen (Beachtung der molekularen Interaktionen) optimiert werden. Zwei verschiedene 3kptCTag-Varianten wurden über ortsspezifische Mutagenese mit spezifischen Primern hergestellt (siehe Kapitel 2.6.2 und Tabelle 1.1, Primersequenzen siehe Tabelle A2 im Anhang) und die Wirkung der Mutationen auf die Effizienz der Bildung von kovalenten Catcher-Tag Komplexen in vergleichenden Rekonstitutionsexperimenten analysiert (siehe Kapitel 2.6.4.9). Bei der einen 3kptCTag-Variante wurden N-terminal die Reste N503 und Q504 hinzugefügt (siehe Abb.3.31 A II; #137 in Tabelle 1.1). Außerdem wurden die artifiziell, analog dem SpyTag hinzugefügten, unphysiologischen C-terminalen Aminosäuren P516, T517 und K518 durch die Reste G, W und I ersetzt, die laut *in silico* Analysen in einigen homologen Proteinen als physiologische Verlängerung des Tags vorkommen (siehe Abb.3.31 A III; #138 in Tabelle 1.1).

Weder der N-terminal verlängerte 3kptCTag-NQ noch die GWI-Substitutionsvariante des 3kptCTags führten zu einer verbesserten Produktausbeute/ Rekonstitutionseffizienz im Vergleich zu der Wildtyp-Reaktion aus 3kptCCatcher bzw. mCherry-3kptCCatcher und 3kptCTag (siehe Abb.3.31 B und C). Die kovalente Komplexbildung aus Catcher und Tag war in keinem Fall verbessert, was auch die anschließende densitometrische Quantifizierung der Gelbilder bestätigte (siehe Abb.3.32). Quantifiziert wurden die Produktbanden der in Triplikaten durchgeführten Interaktionen zwischen mCherry-3kptCCatcher und den jeweiligen 3kptCTag-Varianten. Während die C-terminale GWI-Substitution zu jedem Zeitpunkt Wildtyp-Niveau erreichte, bezüglich der Produktzunahme (immer 0h als Referenz), war die Komplexbildung aus Catcher und Tag eher verschlechtert, wenn die Reaktion mit 3kptCTag-NQ durchgeführt wurde (siehe Abb.3.31 und 3.32).

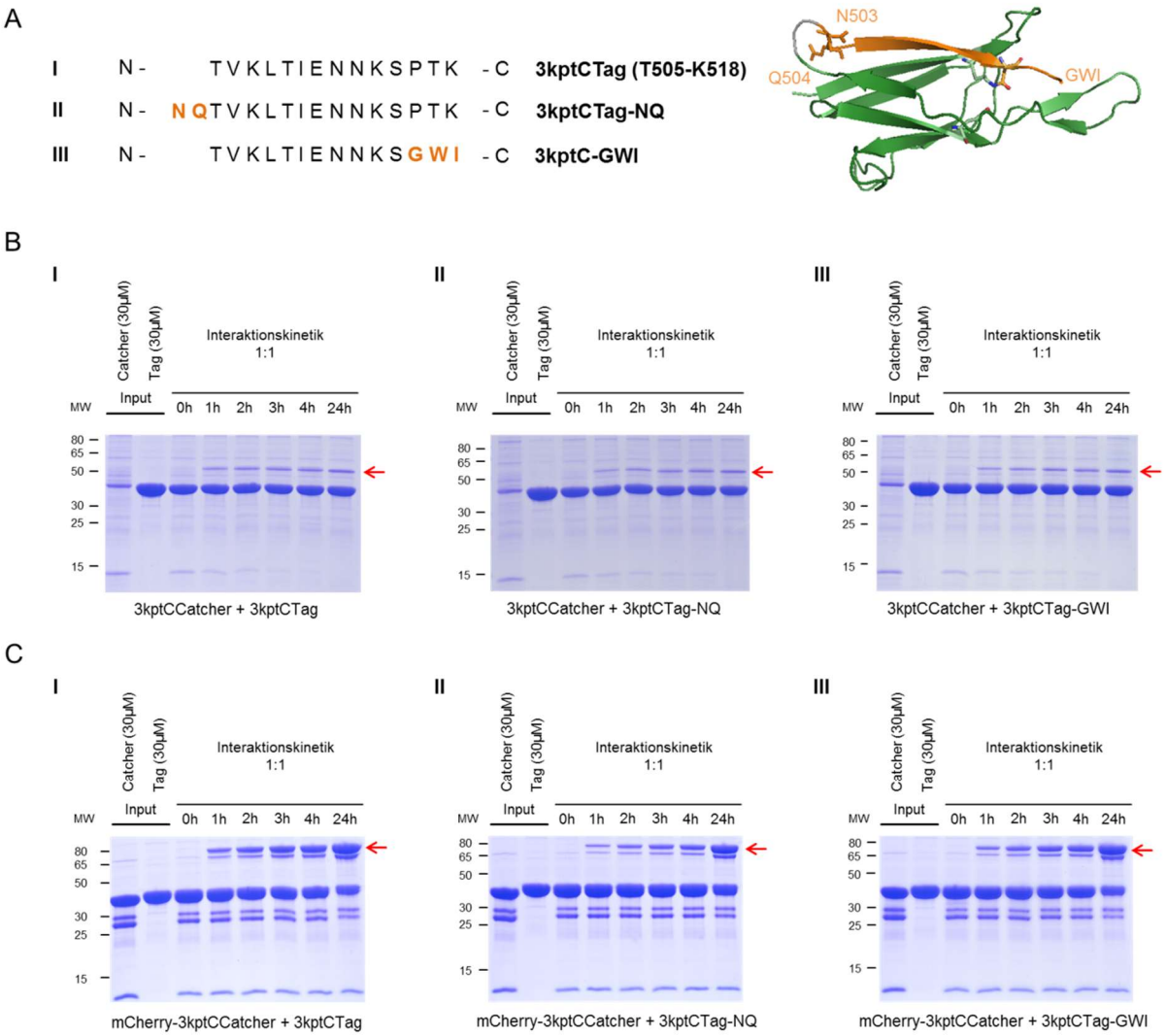


Abb.3.31: Analyse der Rekonstitutionseffizienz von N- und C-terminalen 3kptCTag-Varianten in der Reaktion mit 3kptCCatcher bzw. mCherry-3kptCCatcher. (A) Sequenzvergleiche der verschiedenen 3kptCTag-Varianten (I: 3kptCTag-Wildtyp (Reste T505 - K518), II: N-terminale Verlängerung des 3kptCTags um die Reste N503 und Q504, III: C-terminale Aminosäuresubstitution der Reste PTK durch GWI). Die Molekülvisualisierung wurde mittels PyMOL angefertigt (grün: 3kptCCatcher, orange: 3kptCTag). Rekonstitutionsexperimente zwischen 3kptCCatcher (B) bzw. mCherry-3kptCCatcher (C) und verschiedenen 3kptCTag-Varianten (I bis III). Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem äquimolaren Verhältnis (15 µM Catcher- und Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet.

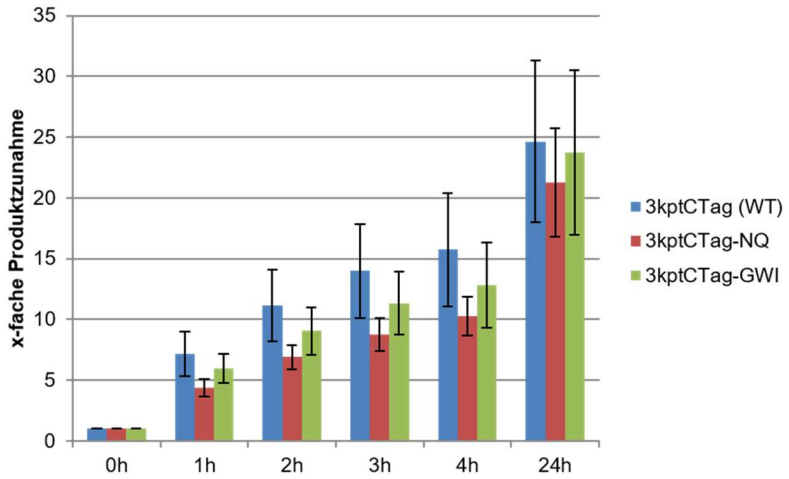


Abb.3.32: Quantifizierung der Zunahme an kovalentem Produkt, resultierend aus der Reaktion zwischen mCherry-3kptCCatcher und verschiedenen 3kptCTag-Varianten mittels Densitometrie. Zu quantifizierende Reaktionen wurden in Triplikaten durchgeführt (Bedingungen siehe Bildunterschrift in Abb.3.31) und die Produktbanden der Proteingele anschließend densitometrisch mittels ImageJ analysiert (0h-Probe als Referenzpunkt =1). Gezeigt sind die Mittelwerte und Standardabweichungen aus drei unabhängig durchgeführten Rekonstitutionsassays.

3.3.2 Stabilisierung der Catcher-Proteine: gezielte Aminosäure-austausche durch ortsspezifische Mutagenese

Bei allen Rekonstitutionsexperimenten und vor allem am Anfang dieser Arbeit, war zu sehen, dass die Catcher-Proteine, 4oq1Catcher und 3kptCCatcher, deutlich schwächer als ihre Tag-MBP Pendants im heterologen Wirt *E. coli* exprimiert wurden (siehe Abb.3.4), was auf Probleme der intrinsischen Proteinstabilität der Catcher-Proteine hindeuten könnte. Aus diesem Grund wurde der Fokus der Optimierung der neuen Split-Isopeptid Systeme vor allem auf die Stabilisierung der größeren Catcher-Proteine gelegt. Die Idee war, die Catcher-Proteine durch gezielte Aminosäuremutationen zu stabilisieren und somit auch die kovalente Interaktion mit den entsprechenden Tag-Peptiden zu verbessern. So wurden verschiedene rationale Strategien in Kombination mit nützlichen, geeigneten Webservern (siehe Material und Methoden Kapitel 2.9) zur Stabilisierung der Catcher-Proteine verwendet (siehe Abb.3.23). Die Klonierungen der zugrundeliegenden DNA-Konstrukte erfolgten über PCR-Methoden mit spezifischen Mutagenese-Primern stets unter Beachtung der *E. coli Codon usage* (siehe 2.6.2 und Tabelle 1.1, Primersequenzen siehe Anhang Tabelle A2). Vergleichende Rekonstitutionsexperimente wurden parallel mit denselben Proteinstammlösungen und denselben Reaktionsbedingungen durchgeführt und die Effizienz der Catcher-Tag Reaktionen visuell und qualitativ mit den entsprechenden Wildtyp-Referenzen (#102 + #104 für das 4oq1-System und #142 + #136 für das 3kptC-System, siehe Tabelle 1.1) verglichen.

3.3.2.1 Entfernung von hydrophoben Aminosäuren an der Proteinoberfläche

In wässriger Lösung falten sich Proteine in ihrer dreidimensionalen Struktur natürlicherweise aufgrund des hydrophoben Effekts in die energetisch günstigste, thermodynamisch stabilste Konformation (siehe Einleitung). So nehmen die meisten Proteine in wässriger Umgebung kompakte globuläre Strukturen an, in der die hydrophoben Aminosäuren üblicherweise im Proteininneren (hydrophober Proteinkern) verborgen sind, während hydrophile Aminosäuren eher dem Lösungsmittel zugewandt sind (Pace et al., 1996). Basierend auf dieser physiologischen Proteinarchitektur wurden exponierte, hydrophobe Aminosäuren an der Oberfläche der Catcher-Proteine durch hydrophilere Aminosäuren unter Beachtung der Aminosäure-Präferenzen für bestimmte Sekundärstrukturelemente (siehe Einleitung) ersetzt, um die Catcher-Proteine zu stabilisieren.

In *in vitro* Rekonstitutionsexperimenten (siehe Kapitel 2.6.4.9) wurden die Effekte der Einzel-Aminosäureaustausche der Catcher-Mutanten auf die Ausbildung von kovalenten Catcher-Tag Komplexen jeweils mit der entsprechenden Catcher-Tag Wildtypreaktion verglichen. Die Substitutionen hydrophober Aminosäuren an der Proteinoberfläche, die ins Lösungsmittel reichen, durch Aminosäuren mit hydrophilen Seitenketten führten sowohl beim 4oq1-System als auch beim 3kptC-System in den meisten Fällen zu keiner verbesserten Rekonstitutionseffizienz im Vergleich zu dem jeweiligen Wildtyp-System (siehe Abb.3.33 und 3.34). Für das 4oq1-System zeigte lediglich die 4oq1Catcher-Mutante, in der Phenylalanin-209 zu Tyrosin mutiert wurde eine verbesserte Reaktivität zur Ausbildung eines kovalenten Catcher-Tag Komplexes (siehe Abb.3.33 I). So war sowohl die Produktbande zu jedem Zeitpunkt der Interaktionskinetik als auch die graduelle Abnahme der Catcher- und Tag-Banden ausgeprägter als im Wildtyp-System. Die 4oq1Catcher-Aminosäureaustausche Valin-196 zu Asparagin und Leucin-215 zu Lysin zeichneten sich durch eine vergleichsweise starke Abnahme der Catcher-Banden von der 0h-Probe/ Input-Probe zur 1h-Probe aus (siehe Abb.3.33 H und J), was für eine gute Reaktivität spricht. Die restlichen 4oq1Catcher-Mutanten waren auf Wildtyp-Niveau bezüglich der Rekonstitutionseffizienz oder schlechter. Die Coomassie-Proteingelbilder zeigten auch, dass die rekombinante Expression mancher 4oq1Catcher-Mutanten, wie z.B. der Isoleucin-185 zu Glutaminsäure-Mutante, sehr schlecht war (siehe Abb.3.33 G), was eine Interpretation der Effizienz der kovalenten Interaktion zwischen Catcher und Tag erschwerte.

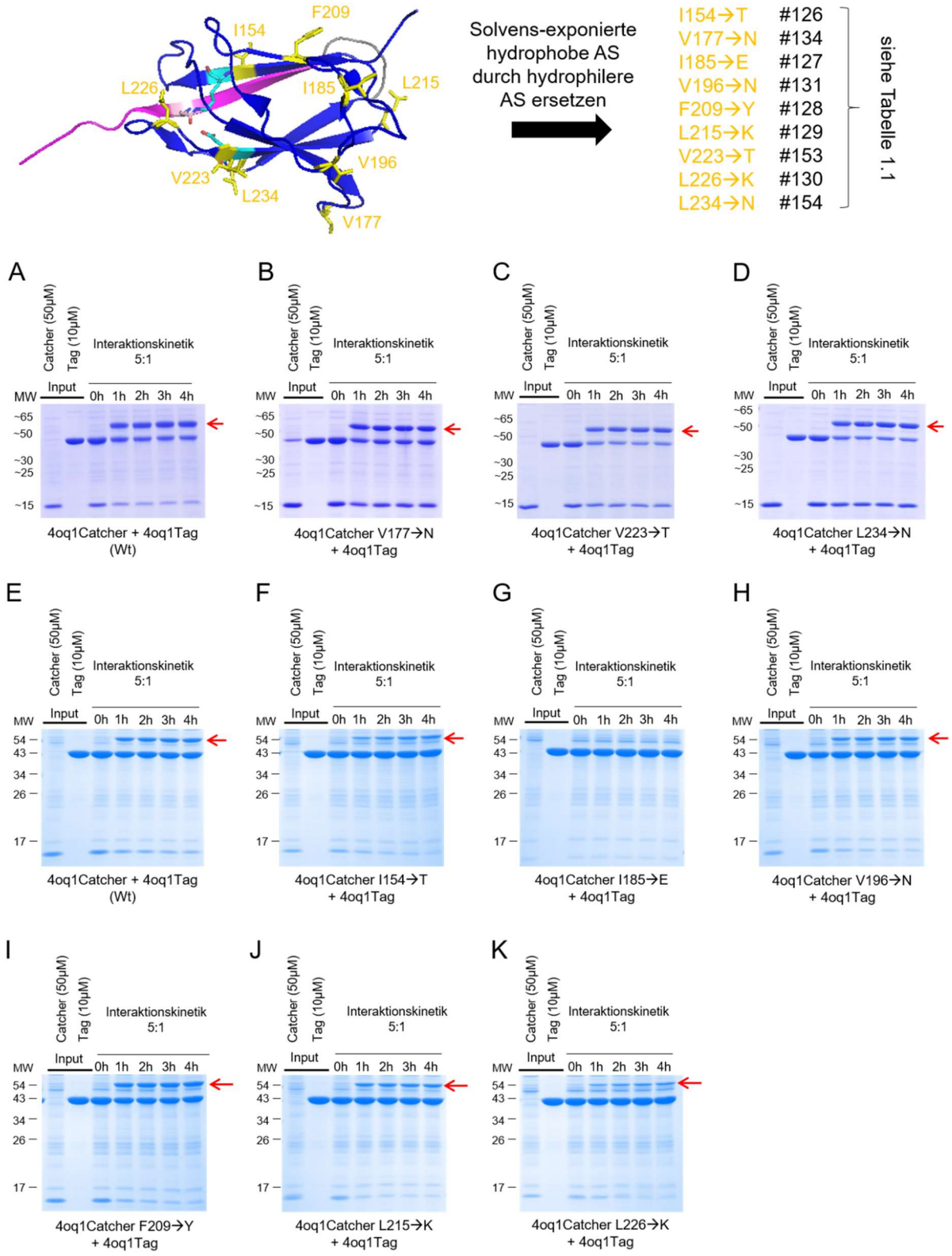


Abb.3.33: Analyse der Mutationseffekte hydrophober Aminosäuren an der Proteinoberfläche des 4oq1Catchers. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher zu Tag Verhältnis (50 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 1,5M TMAO inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierung wurde mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag, zu mutierende Reste sind in gelb hervorgehoben).

Ähnlich verhielt es sich mit dem 3kptC-System. Nur die Mutation von Phenylalanin an der Stelle 471 zu Tyrosin im 3kptCCatcher führte bei der Reaktion mit 3kptCTag zu einer im Vergleich zum Wildtyp-System verbesserten Produktausbeute, was vor allem an der effizienteren Abnahme der 3kptCCatcher F471→Y korrespondierenden Bande im Gel festgemacht werden konnte (siehe Abb.3.34 B). Die beiden anderen 3kptCCatcher-Mutanten erbrachten keine Effizienzsteigerung in der Ausbildung eines kovalenten Produktes aus Catcher und Tag (siehe Abb.3.34 C und D). Im Gegenteil, wurde Isoleucin-418 zu Lysin mutiert so verschlechterte sich die kovalente Reaktion zwischen Catcher und Tag sogar im direkten Vergleich mit dem 3kptC-Wildtyp-System (siehe Abb.3.34 A und D).

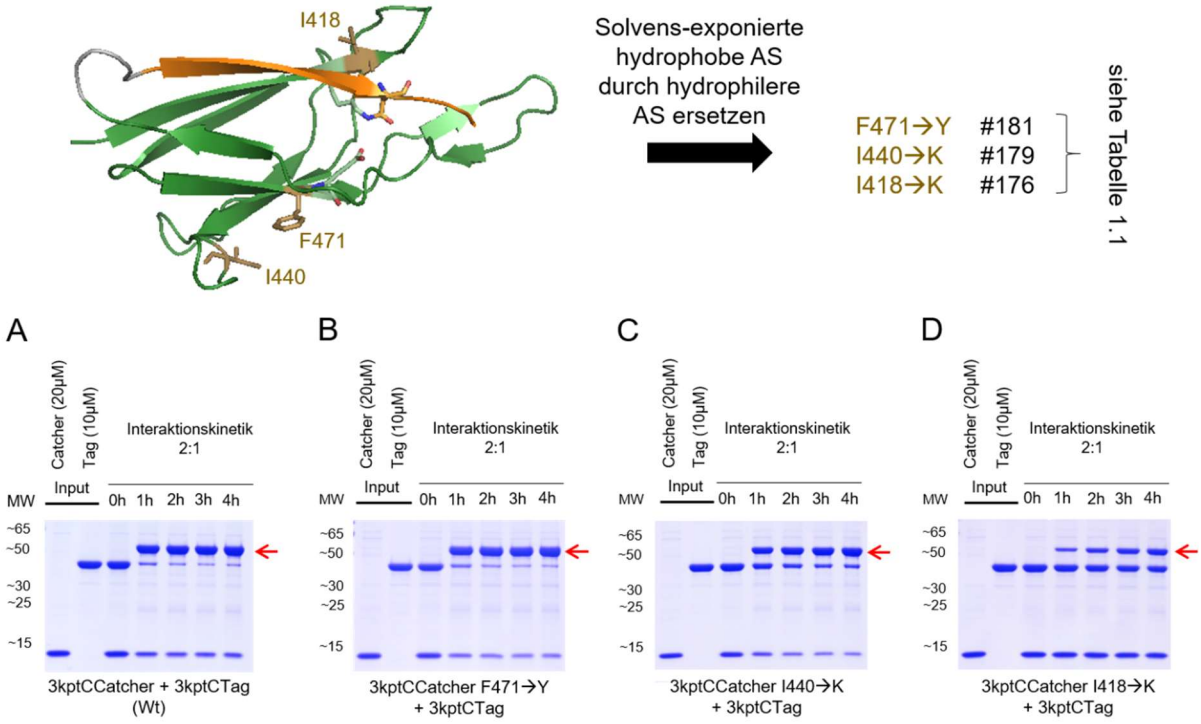


Abb.3.34: Analyse der Mutationseffekte hydrophober Aminosäuren an der Proteinoberfläche des 3kptCCatchers. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 2:1 Catcher zu Tag Verhältnis (20 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 40% Glycerin inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierung wurde mittels PyMOL angefertigt (grün: 3kptCCatcher, orange: 3kptCTag, zu mutierende Reste sind in braun hervorgehoben).

3.3.2.2 Strukturbasierte Mutationen

3.3.2.2.1 Zugänglichkeit der Tag-Bindestelle im Catcher-Protein

Bei rationaler Betrachtung der 3D-Proteinstruktur des 4oq1Tag-4oq1Catcher Komplexes mithilfe des Visualisierungsprogramms PyMOL fiel auf, dass im 4oq1Catcher-Protein ein Tyrosinrest (Y229) direkt an der Eintrittsstelle des 4oq1Tags und in unmittelbarer Nähe zur Isopeptidbindung lokalisiert ist (siehe Abb.3.35). Aufgrund der Hypothese, dass die große, aromatische Ringstruktur des Tyrosins den 4oq1Tag-Eintritt und damit die Ausbildung eines kovalenten Catcher-Tag Komplexes eventuell stört, wurde dieser Tyrosinrest an Position 229 in eine kleinere Aminosäure mit einer Hydroxylgruppe, nämlich Serin, mutiert (siehe Abb.3.35). Vergleichende Rekonstitutionsexperimente mit dem 4oq1-Wildtyp-System wurden durchgeführt, um zu testen ob die Mutation einen positiven Effekt auf die kovalente Interaktion zwischen Catcher und Tag hat. Das Ergebnis war deutlich und zeigte, dass die Y229→S 4oq1Catcher-Mutante die kovalente Reaktion mit dem 4oq1Tag-Reaktionspartner deutlich verschlechterte, verglichen mit dem 4oq1-Wildtyp-System (siehe Abb.3.35 A und B). Sowohl die Produktbande, die den kovalenten Komplex aus Catcher und Tag repräsentiert, als auch die korrelierende Abnahme der Eduktbanden waren deutlich schwächer ausgeprägt. Die Ergebnisse widerlegen nicht nur die anfänglich aufgestellte Hypothese, dass dieser Tyrosinrest die Catcher-Tag Interaktion stören könnte, sondern deuten auf eine essentielle Rolle des Tyrosins-229 für die Interaktion zwischen Catcher und Tag hin, denn die alleinige Mutation dieser Aminosäure führte beinahe zu einem kompletten Verlust der kovalenten Reaktion zwischen 4oq1Catcher und 4oq1Tag (siehe Abb.3.35).

Analog verhält es sich mit der Tyrosin-Mutation an Stelle 488 zu Prolin im 3kptCCatcher (siehe Abb.3.37 E). Auch im 3kptCCatcher liegt die aromatische Aminosäure Tyrosin-488 in unmittelbarer Nachbarschaft der Isopeptidbindung zwischen Catcher und Tag (siehe Abb.3.37). Die Substitution dieses Tyrosinrestes durch Prolin führte zu einem kompletten Funktionsverlust der kovalenten Reaktion zwischen 3kptCCatcher und 3kptCTag (siehe Abb.3.37 E) und zeigt somit die essentielle Rolle dieser Aminosäure an dieser Stelle für die Ausbildung der Isopeptidbindung.

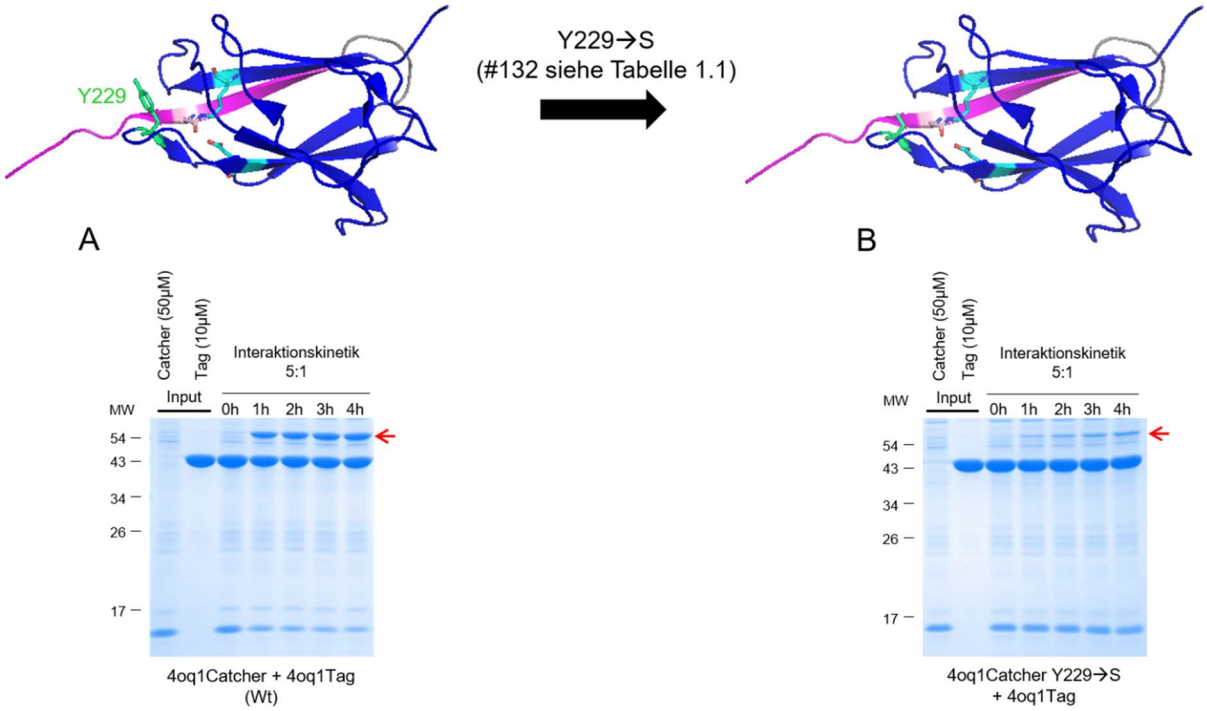


Abb.3.35: Analyse der Tyrosin-229 Mutation des 4oq1Catchers. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher zu Tag Verhältnis (50 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 1,5M TMAO inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierungen wurden mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag, der zu mutierende Tyrosinrest-229 ist in türkis hervorgehoben).

3.3.2.2.2 Prolin-Substitutionen

Proline sind, wie auch Glycine, strukturell besonders wichtige Aminosäuren, die aufgrund ihrer physikochemischen Eigenschaften (v.a. Größe, Torsionswinkel) die Konformation eines Proteins in hohem Maße beeinflussen (Morgan und Rubenstein, 2013, Choi et al., 2006). Dabei verhalten sich Prolin und Glycin quasi gegenteilig. Während Glycin die kleinste (R: H) und flexibelste Aminosäure darstellt, die in alle Strukturen problemlos passt und daher keine Sekundärstrukturelemente besonders bevorzugt, gilt Prolin als Sekundärstruktur (α-Helix und β-Faltblatt)-Zerstörer (Choi et al., 2006). Prolin besitzt eine unpolare, rigide Pyrrolidinring-Seitenkette, die sowohl mit dem C_α-Atom als auch mit dem Stickstoffatom verbunden ist. Daher fehlt Prolin eine freie NH-Gruppe, die Wasserstoffbrücken ausbilden kann (Choi et al., 2006, Morgan und Rubenstein, 2013). Die Ringstruktur schränkt die Rotationsfähigkeit um die C_α-N Bindung und damit die Konformation erheblich ein und fixiert den phi-Winkel bei ca. -60°, was nicht zu α-Helices und β-Faltblatt-Strukturen passt (Stryer, Choi et al., 2006). Zudem liegen mit Prolin geknüpfte Peptidbindungen häufig in der seltenen cis-Konfiguration vor, was ebenfalls die Proteinkonformation aufgrund sterischer Hinderungen weitreichend

beeinflusst (Stryer). Die strukturelle Starrheit, bzw. die eingeschränkte Flexibilität die Proline aufgrund ihrer physikochemischen Eigenschaften innerhalb der Polypeptidkette vermitteln, reduziert jedoch destabilisierende entropische Kräfte des Proteinrückgrats im ungefalteten Zustand und erhöht somit die Proteinstabilität (Choi et al., 2006). Aus diesem Grund werden Prolin-Mutationen vor allem in *loop*-Regionen zur Stabilisierung von Proteinstrukturen eingesetzt (Choi et al., 2006).

Zur Selektion geeigneter Aminosäuren als Angriffspunkte für Prolin-Substitutionen in Schleifen-Regionen der Catcher Proteine, 4oq1Catcher und 3kptCCatcher, wurden zunächst die Torsionswinkel (phi und psi) der Catcher-Proteine mithilfe des Programms MolProbity (siehe 2.9) ermittelt. Ausgewählte Aminosäuren (Nähe zur Isopeptidbindung als ein Kriterium) mit phi-Winkeln zwischen -50° und -70° in flexiblen *loop*-Regionen der Proteine wurden durch die Aminosäure Prolin mittels ortsspezifischer Mutagenese auf Nukleinsäureebene ersetzt. Vergleichende *in vitro* Interaktionsexperimente sollten nun klären, ob durch diese Erhöhung der strukturellen Starrheit die Proteinstruktur der Catcher-Proteine stabilisiert wird und somit einen positiven Einfluss auf die Rekonstitutionseffizienz zwischen Catcher und Tag hat.

Für das 4oq1-System kam es bei keiner der beiden Prolin-4oq1Catcher-Substitutionen (N161→P und V231→P) zu einer verbesserten Effizienz der kovalenten Reaktion zwischen 4oq1Catcher und 4oq1Tag im Vergleich zum Wildtyp-System (siehe Abb.3.36 A - C). Die kovalente Interaktion zwischen 4oq1Catcher und 4oq1Tag wurde dagegen eher verschlechtert und führte im Falle der Mutation von Asparagin-161 zu Prolin zu einem vollständigen Verlust der kovalenten Catcher-Tag Reaktion (siehe Abb.3.36 B).

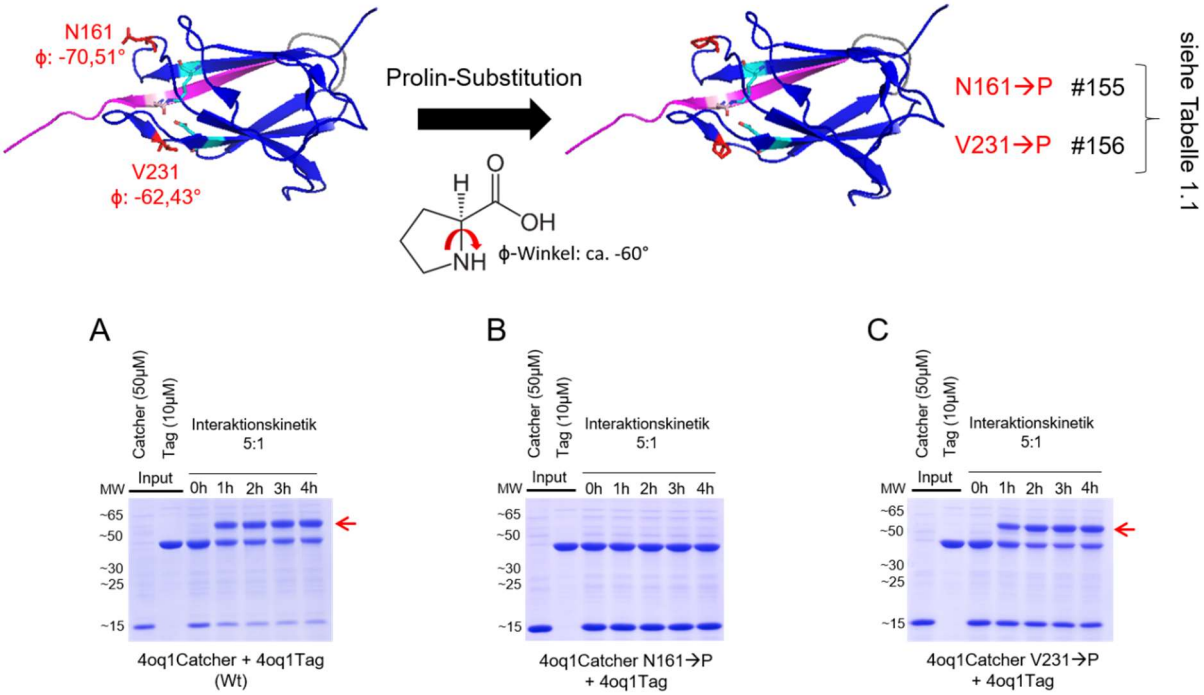


Abb.3.36: Analyse der Mutationseffekte der Prolin-Substitutionen des 4oq1Catchers. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher zu Tag Verhältnis (50 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 1,5M TMAO inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierungen wurden mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag, zu mutierende Reste sind in rot hervorgehoben).

Drei der vier Prolin-Substitutionen des 3kptCCatcher-Proteins (N422→P, K425→P, E454→P, Y488→P) führten, im Vergleich zu dem 3kptC-Wildtyp-System, zu einer teilweise dramatischen Verschlechterung der kovalenten Reaktion zwischen Catcher und Tag (siehe Abb.3.37). Die bereits angesprochene Mutation von Tyrosin-488 zu Prolin führte zu einem kompletten Verlust der kovalenten Catcher-Tag Reaktion, während die Mutation von Asparagin-422 zu Prolin in einer starken Reduktion der Interaktion resultierte (siehe Abb.3.37). Lediglich die Substitution von Lysin an Position 425 durch Prolin schien einen positiven Effekt auf die Effizienz der Ausbildung von kovalenten 3kptCCatcher-3kptCTag Komplexen zu haben (siehe Abb.3.37 C), was v.a. anhand der raschen und starken Abnahme der Catcher- bzw. v.a. der Tag-Banden im Vergleich zu den Input-Banden und der 0h-Kontrolle deutlich wurde.

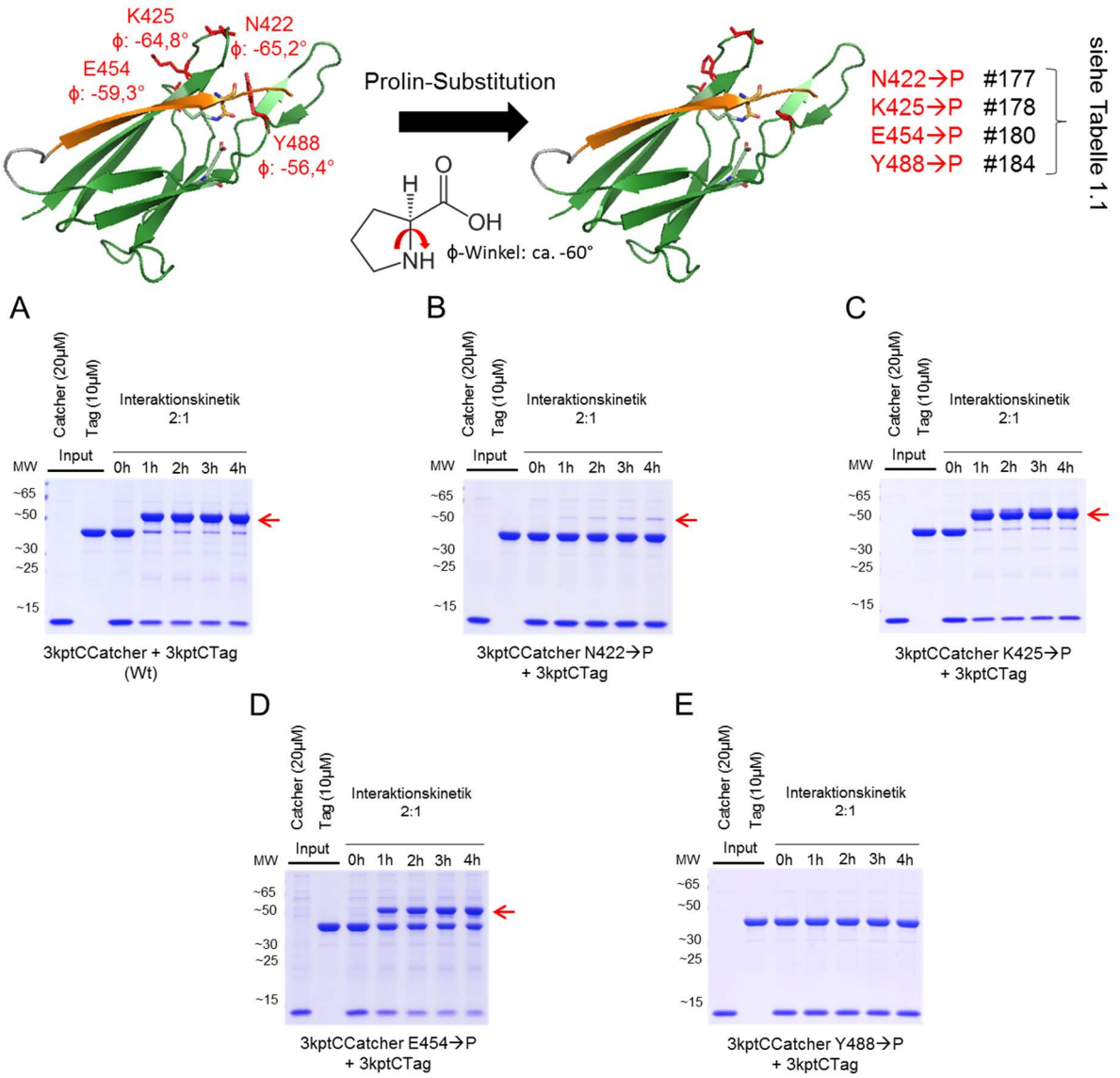


Abb.3.37: Analyse der Mutationseffekte der Prolin-Substitutionen des 3kptCCatchers. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 2:1 Catcher zu Tag Verhältnis (20 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 40% Glycerin inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierungen wurden mittels PyMOL angefertigt (grün: 3kptCCatcher, orange: 3kptCTag, zu mutierende Reste sind in rot hervorgehoben).

3.3.2.2.3 Glycin-Substitutionen

Ein anderer, eher entgegengesetzter, Ansatz als die Erhöhung der Rigidität der Proteinstruktur zur Stabilisierung der Catcher-Proteine durch Einführen von Prolinen, wurde durch Glycin-Substitutionen in *loop*-Regionen zur Erhöhung der Flexibilität der Proteinstruktur erzielt. Exemplarisch wurden zwei ausgewählte Aminosäuren im 4oq1Catcher-Protein, Asparagin-212 und Asparaginsäure-176, in Glycine mutiert (siehe Abb.3.38). Analog durchgeführte Rekonstitutionsexperimente zeigten jedoch keine

verbesserte kovalente Catcher-Tag Komplexbildung der 4oq1Catcher-Glycin-Mutanten im Vergleich zum 4oq1-Wildtyp-System (siehe Abb.3.38 A – C).

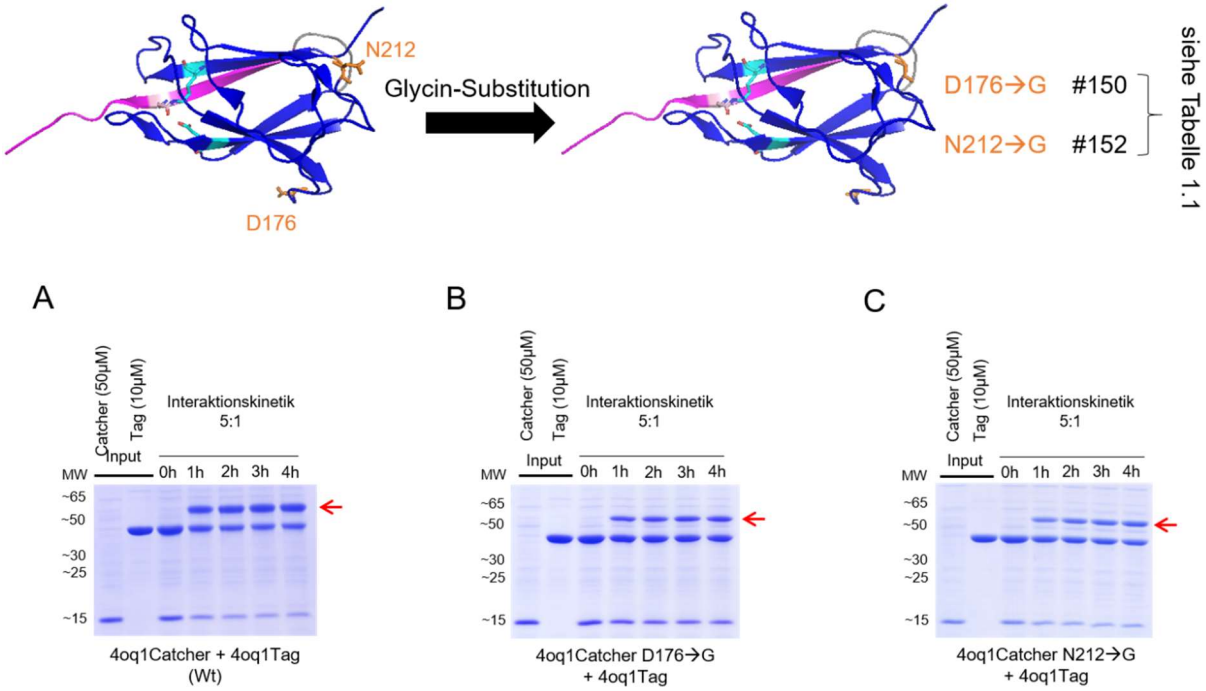


Abb.3.38: Analyse der Mutationseffekte der Glycin-Substitutionen des 4oq1Catchers. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher zu Tag Verhältnis (50 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 1,5M TMAO inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierungen wurden mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag, zu mutierende Reste sind in orange hervorgehoben).

3.3.2.3 B-Faktor-basierte Substitutionen (Austausch besonders flexibler Aminosäuren)

Der B-Faktor, oder auch Temperatur-Faktor (Debye-Waller Faktor), ist ein Parameter aus der Kristallographie und beschreibt die positionelle Auslenkung/ Mobilität von einzelnen Atomen (Abweichung der Elektronendichte) in der Kristallstruktur und kann somit als Maß für die Flexibilität von Proteinstrukturen dienen (Trueblood et al., 1996, Lumry, 2002). Je größer der B-Faktor, desto größer ist die räumliche Ausdehnung der Elektronenwolke, desto flexibler, mobiler ist ein Atom bzw. die Proteinkonformation in diesem Bereich. Die B-Faktor Werte jeder Aminosäure der Catcher-Proteine wurden visuell bei PyMOL ermittelt (b-factor putty). Besonders flexible Aminosäuren in *loop*-Regionen der Catcher-Proteine (rote Regionen, hohe B-Faktor Werte siehe Abb.3.39 und 3.40) wurden mutiert und durch energetisch günstige, stabilisierende Aminosäuren, ermittelt über Rosetta Backrub (siehe 2.9), ersetzt. Vergleichende *in vitro* Rekonstitutionsexperimente, mit den jeweiligen Wildtyp-Systemen als

Referenzen, sollten den Einfluss der Einzel-Aminosäureaustausche auf die Effizienz der kovalenten Reaktionen zwischen Catcher und Tag aufklären.

Die kovalente Interaktion zwischen 4oq1Catcher und 4oq1Tag konnte durch keine der beiden B-Faktor basierten Catcher-Mutanten (M147→L, D176→G siehe auch bei Glycin-Substitutionen) im Vergleich zum 4oq1-Wildtyp-System verbessert werden (siehe Abb.3.39 A - C).

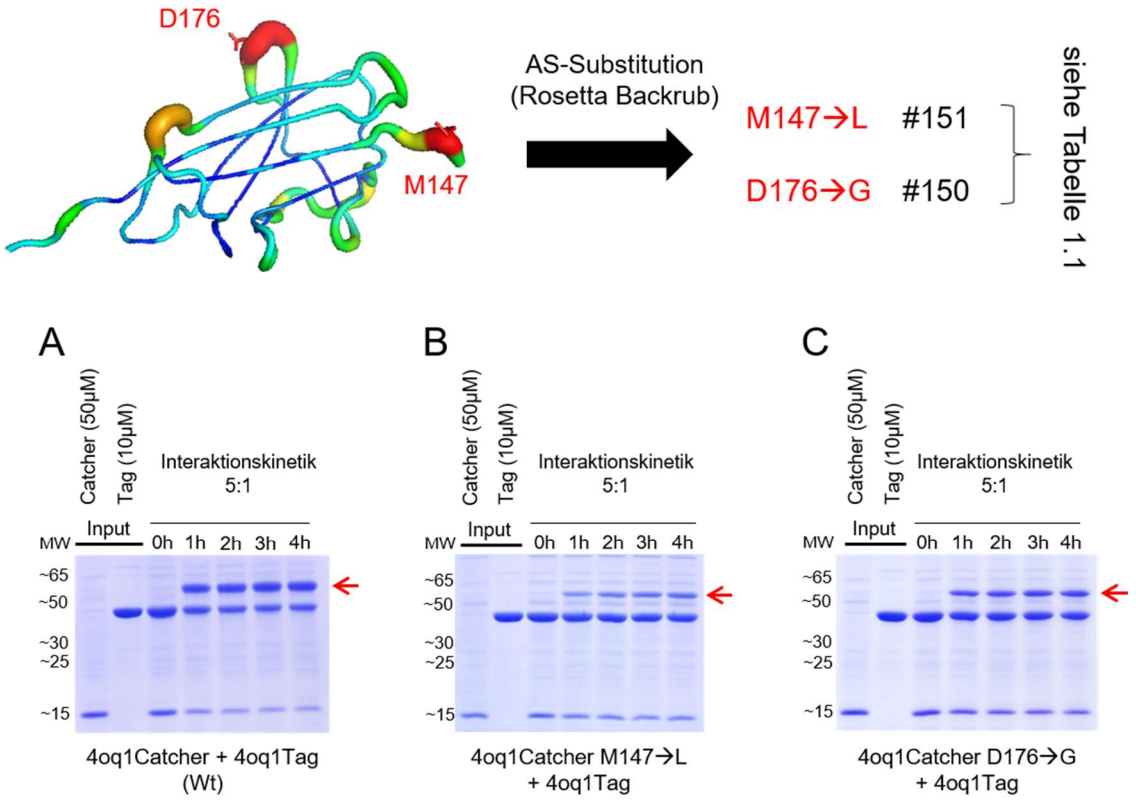


Abb.3.39: Analyse der Aminosäureaustausche im 4oq1Catcher basierend auf B-Faktor Werten. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher zu Tag Verhältnis (50 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 1,5M TMAO inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierung wurde mittels PyMOL angefertigt (zu mutierende Reste sind in rot hervorgehoben).

Potentiell stabilisierende Mutationen besonders flexibler Aminosäuren, basierend auf hohen B-Faktor Werten (siehe Abb.3.40), des 3kptCCatcher-Proteins (E481→S und D482→N) wurden in kovalenten Bindungstests im Vergleich zum 3kptC-Wildtyp-System analysiert (siehe Abb.3.40 A - C). Die Reaktion beider 3kptCCatcher-Mutanten mit 3kptCTag-Peptid erreichte Wildtyp-Niveau bezüglich der Bildung von kovalenten Catcher-Tag Komplexen. Die Mutation von Glutaminsäure an Position 481 zu Serin führte sogar zu einer leichten

Steigerung der Rekonstitutionseffizienz (siehe Abb.3.40 B). So schien der 3kptCTag schneller mit dem 3kptCCatcher E481→S zu reagieren als mit dem Wildtyp-3kptCCatcher, was an der starken Abnahme der 3kptCCatcher E481→S korrespondierenden Bande vor allem innerhalb der ersten Stunde festgemacht werden kann (siehe Abb.3.40 A und B).

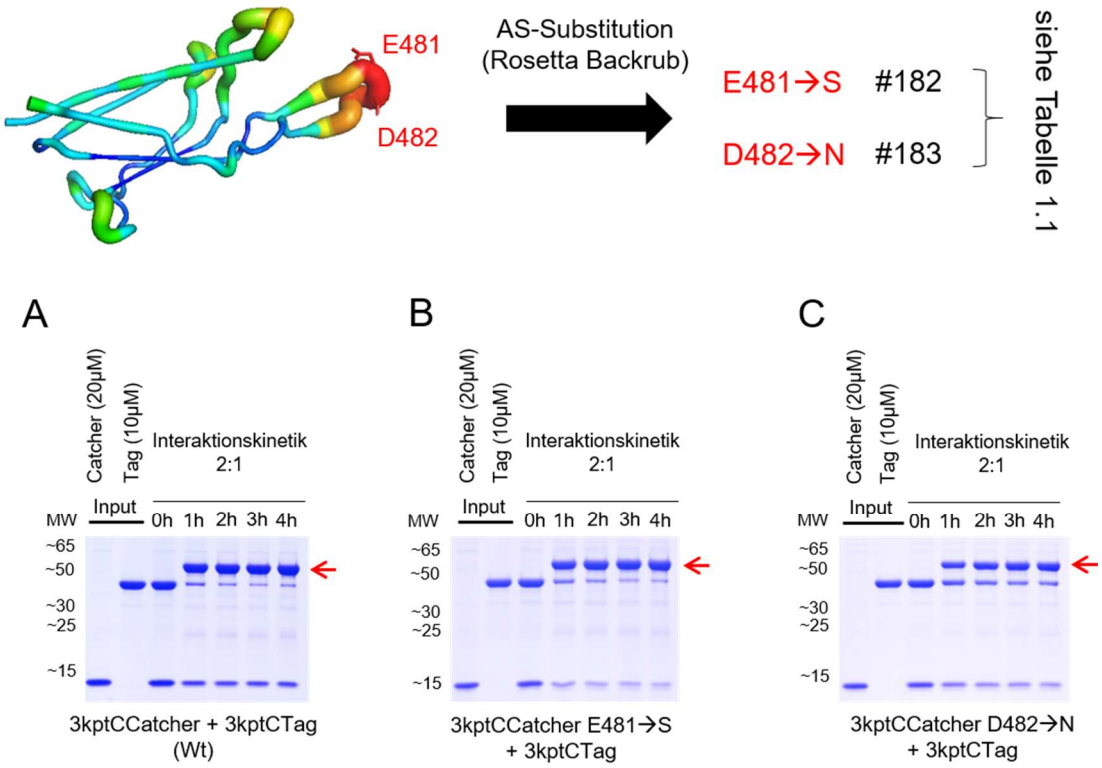


Abb.3.40: Analyse der Aminosäureaustausche im 3kptCCatcher basierend auf B-Faktor Werten. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 2:1 Catcher zu Tag Verhältnis (20 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 40% Glycerin inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierung wurde mittels PyMOL angefertigt (zu mutierende Reste sind in rot hervorgehoben).

3.3.2.4 PROSS – Protein Repair One Stop Shop

Eine weitere ,,Protein Engineering"-Strategie zur Verbesserung der Proteinstabilität der Catcher-Proteine, 4oq1Catcher und 3kptCCatcher, basierte auf dem Internetserver PROSS (Protein Repair One Stop Shop; Weizmann Institut Fleishman Labor; Stable by Design, Goldenzweig et al., 2016). Bei dem PROSS-Webserver handelt es sich um ein vollautomatisiertes struktur- und sequenzbasiertes ,,Proteindesign"-Programm, das auf die Verbesserung von Proteinen im Hinblick auf deren Stabilität und heterologe (Über)Expression in Bakterien bzw. Proteinlöslichkeit abzielt (Goldenzweig et al., 2016). Es

verspricht die Herstellung von stabilisierten Proteinvarianten von nahezu jedem Zielprotein, ohne deren Funktion zu ändern (z.B. Enzyme wie die humane Acetylcholinesterase und die bakterielle Phosphotriesterase, Goldenzweig et al., 2016). Als Input werden die Aminosäuresequenz und die dreidimensionale Struktur (pdb-Eintrag, Kristallstruktur) des Proteins eingelesen und das Programm berechnet, basierend auf ,,Multiple Sequence Alignments“ (MSA) mit homologen Proteinen (*homologous sequence analysis*) und bestimmten Algorithmen (*computational mutation scanning*) energetisch günstigere Proteinvarianten (*favorable native-state energy*). Als Ergebnis werden verschiedene Aminosäuresubstitutionen vorgeschlagen, die das Protein schließlich stabilisieren sollen (Goldenzweig et al., 2016). Als Input wurden die Catcher-Tag Komplexe (4oq1: Reste T146 - N259, 3kptC: Reste T409 - S515) verwendet, also nicht nur die Catcher-Proteine alleine, da es letztlich um die Interaktion zwischen Catcher und Tag geht. Um die reaktive Triade, die wahrscheinlich nicht die energetisch günstigsten Aminosäuren in diesem Proteinbereich darstellen, zu schützen, wurden diese Reste fixiert und aus den Berechnungen ausgeschlossen. Vorgeschlagene Catcher-Mutanten (siehe Abb.3.41 und 3.42) wurden als synthetische Gene bestellt und in den pQE-9 Expressionsvektor kloniert (Klonierungen siehe Kapitel 2.6.2, Sequenzen der synthetischen Gene #148 und #149 in Tabelle A1 im Anhang) und in Rekonstitutionsexperimenten (siehe Kapitel 2.6.4.9) mit den jeweiligen Tag-Peptiden getestet.

Vergleichende Interaktionsanalysen wurden sowohl mit den Wildtyp-Catchern, 4oq1Catcher und 3kptCCatcher, als auch mit den PROSS-basierten Catcher-Proteinen, 4oq1Catcher-PROSS (7 Mutationen) und 3kptCCatcher-PROSS (5 Mutationen), durchgeführt und die Effizienz der Bildung von kovalenten Catcher-Tag Produkten untersucht. Während der 4oq1Catcher-PROSS die Rekonstitutionseffizienz im Vergleich zum 4oq1-Wildtyp-System nicht verbessern konnte (siehe Abb.3.41), schien die 3kptCCatcher-PROSS Mutante effizienter mit 3kptCTag zu kovalenten Komplexen zu reagieren als der 3kptCCatcher-Wildtyp (siehe Abb.3.42). So waren die Catcher- und Tag-Eduktbanden bereits eine Stunde nach Reaktionsstart in Abb.3.42 B, im Vergleich zur 0h-Kontrolle und den Input-Banden, schon dramatisch reduziert, was für eine effiziente Reaktion zwischen 3kptCCatcher-PROSS und 3kptCTag spricht. Auffallend war zudem die Doppelbande, die sowohl bei den 3kptCCatcher-PROSS Banden als auch bei den Produktbanden (1h - 4h) auftauchte. Scheinbar zerfällt der 3kptCCatcher-PROSS in einen zweiten, etwas kleineren reaktiven Teil, der ebenfalls kovalente Komplexe mit 3kptCTag ausbilden kann (siehe Abb.3.42 B).

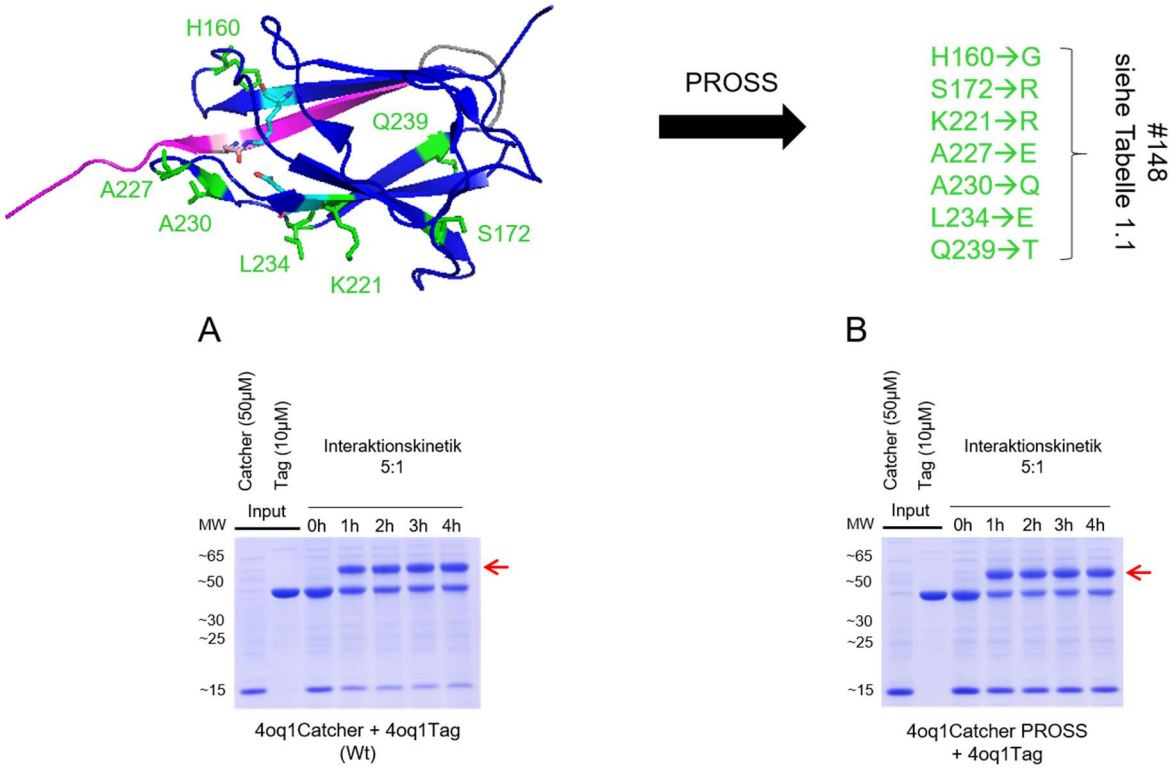


Abb.3.41: Analyse der PROSS-basierten 4oq1Catcher-Mutante. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher zu Tag Verhältnis (50 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 1,5M TMAO inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierung wurde mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag, zu mutierende Reste sind in grün hervorgehoben, sieben Aminosäuren wurden gleichzeitig mutiert).

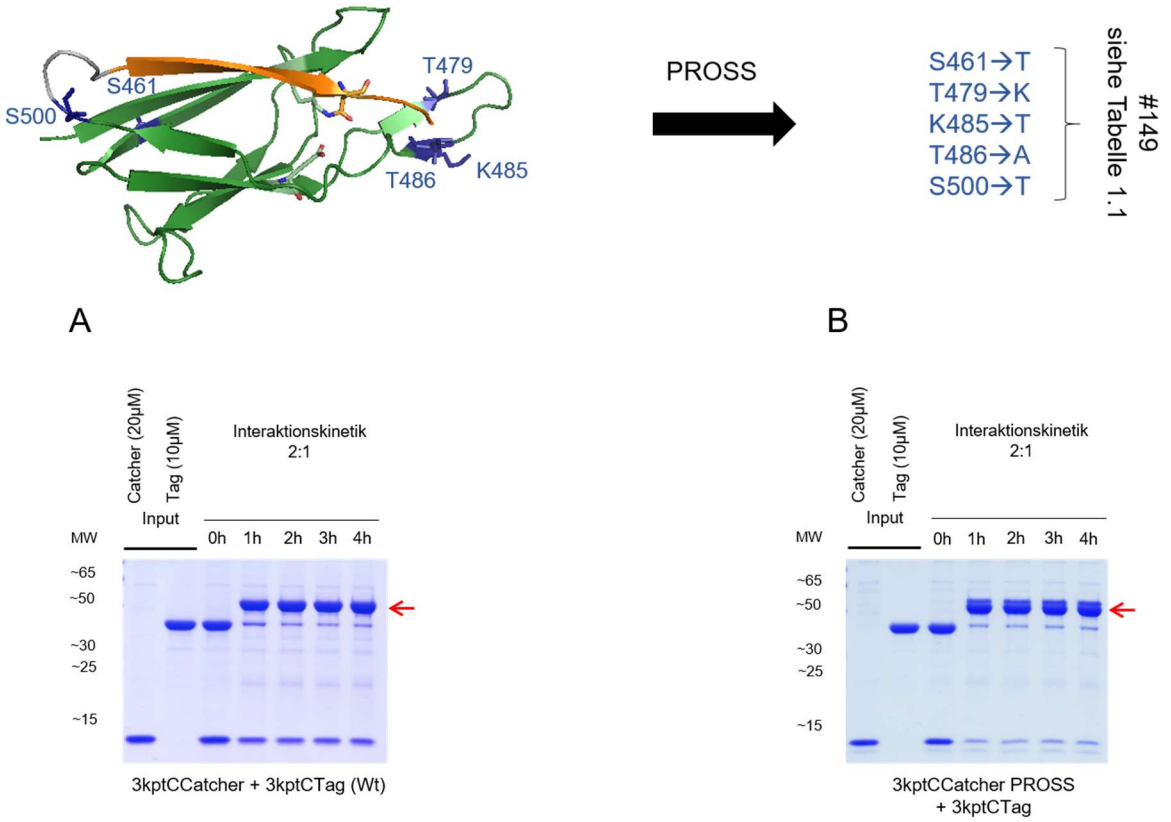


Abb.3.42: Analyse der PROSS-basierten 3kptCCatcher-Mutante. Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 2:1 Catcher zu Tag Verhältnis (20 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 40% Glycerin inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierung wurde mittels PyMOL angefertigt (grün: 3kptCCatcher, orange: 3kptCTag, zu mutierende Reste sind in blau hervorgehoben, fünf Aminosäuren wurden gleichzeitig mutiert).

3.3.2.5 „SpyTag“-Mutationen: N→D Substitution in den Tag-Peptiden

Die neu identifizierten, kovalenten Split-Isopeptid Systeme 4oq1 und 3kptC leiten sich, wie auch das Spy-System und alle weiteren bis heute entwickelten Catcher-Tag Systeme, von Ig-ähnlichen Faltungsdomänen des CnaB-Typs ab (siehe Einleitung). Jedoch sind die reaktiven Aminosäuren, die direkt an der Isopeptidbindung beteiligt sind, andere. Während die intramolekulare Isopeptidbindung der CnaB2-Domäne des Fibronektinbindeproteins, auf der das Spy-System basiert, zwischen Seitenketten von Lysin (K) und Asparaginsäure (D) unter Abspaltung von Wasser gebildet wird (Zakeri et al., 2012), stammen die Systeme 4oq1 und 3kptC von CnaB-Domänen, innerhalb der die Isopeptidbindung zwischen Seitenketten von Lysin (K) und Asparagin (N) unter Abspaltung von Ammoniak entsteht (Shaik et al., 2014, Budzik et al., 2009). Glutaminsäure katalysiert in jedem Fall die Ausbildung der Isopeptidbindung, was charakteristisch für CnaB-Domänen ist (siehe Einleitung 1.1.2.2). Aufgrund der enormen Effizienzunterschiede der neuen Systeme, 4oq1 und 3kptC, im

Vergleich zu dem SpyCatcher-SpyTag System bezüglich der Split-Domänen Rekonstitution, sollte überprüft werden, ob die Mutation des reaktiven Asparagins zu Asparaginsäure (siehe Abb.3.43 und Abb.3.44) die Reaktion der neu identifizierten Systeme, 4oq1 und 3kptC, verbessern kann.

Weder die Asparagin zu Asparaginsäure-Mutante des 4oq1Tags (4oq1Tag N252→D) noch die Asparagin zu Asparaginsäure-Mutante des 3kptCTags (3kptCTag N512→D) zeigten in Rekonstitutionsexperimenten, im Vergleich zu den jeweiligen Wildtyp-Systemen, verbesserte Produktausbeuten (siehe Abb.3.43 und Abb.3.44). Die N→D Tag-Mutanten verhielten sich dagegen wie die nicht-reaktiven N→A Mutanten (siehe Kapitel 3.2.1) und führten zu einem kompletten Verlust der Fähigkeit zur Bildung von kovalenten Catcher-Tag Komplexen (siehe Abb.3.43 und Abb.3.44).

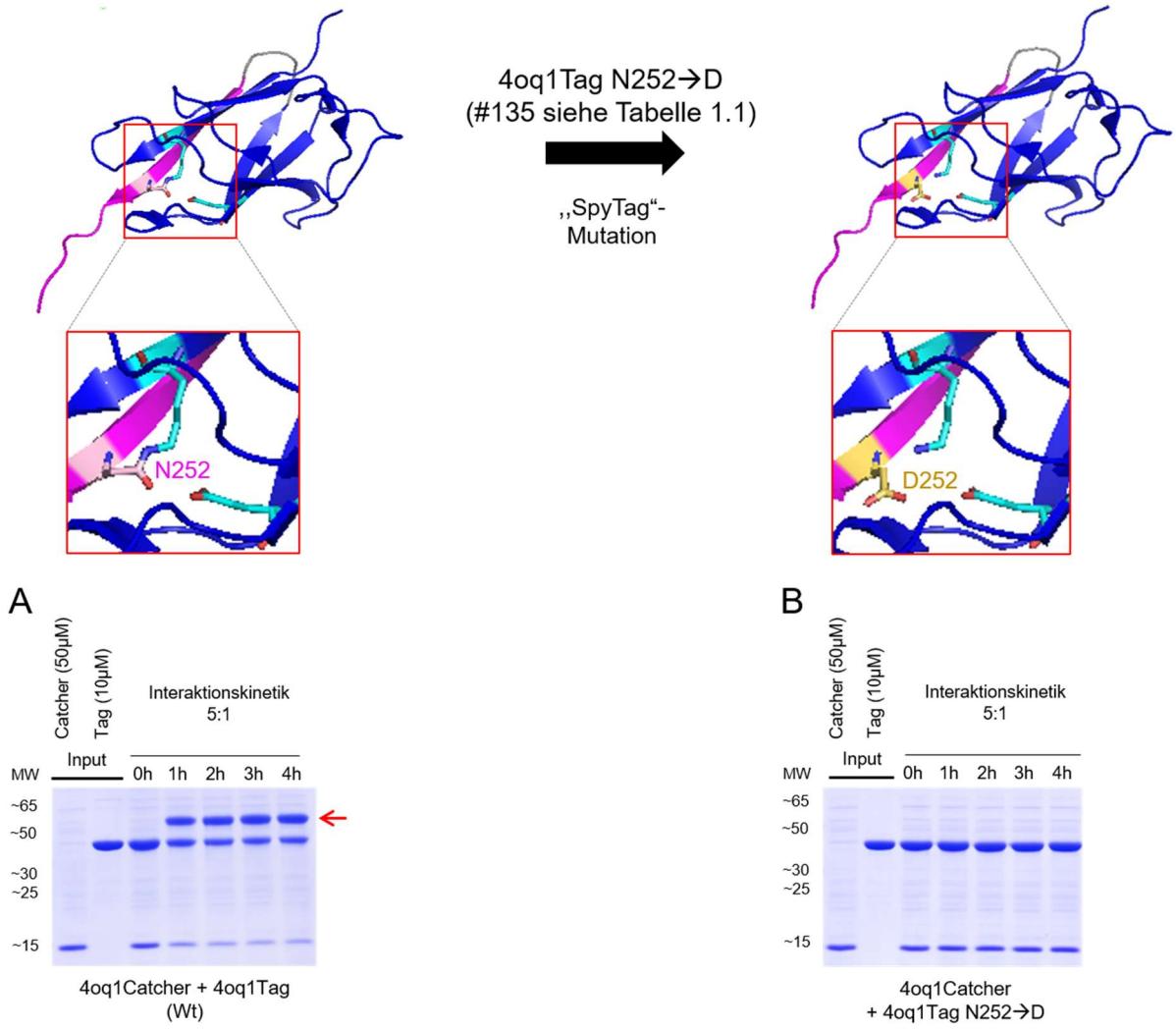


Abb.3.43: Analyse der „SpyTag"-Mutation des 4oq1Tags (N252→D). Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem 5:1 Catcher zu Tag Verhältnis (50 µM Catcher- und 10 µM Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer mit Zugabe von 1,5M TMAO inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierungen wurden mittels PyMOL angefertigt (blau: 4oq1Catcher, rosa: 4oq1Tag).

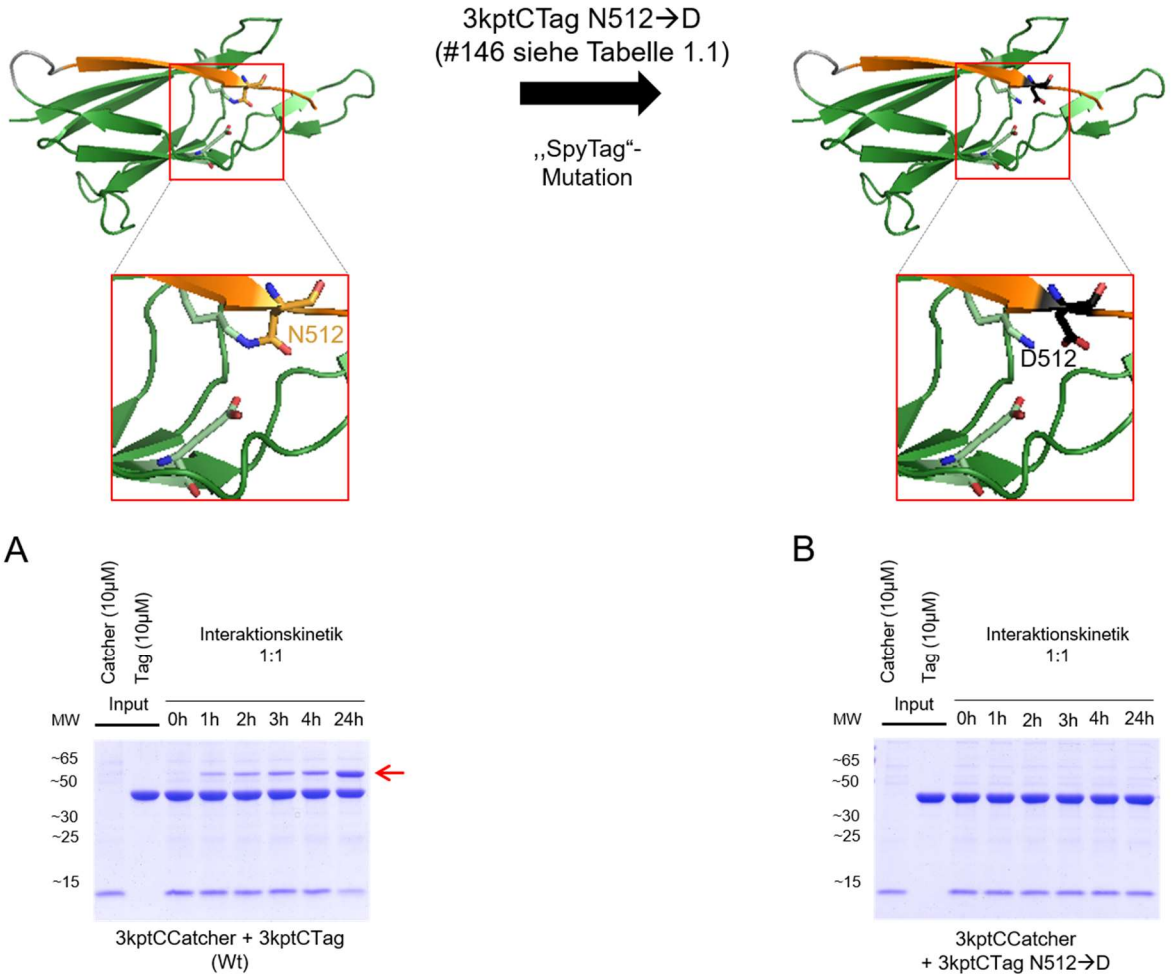


Abb.3.44: Analyse der ,,SpyTag"-Mutation des 3kptCTags (N512→D). Aufgereinigte und in 1xPBS pH 7,5 dialysierte Proteine wurden in einem äquimolaren Verhältnis (10 µM Catcher- und Tag-Endkonzentration) gemischt und bei 25°C in 1xPBS pH 7,5 als Reaktionspuffer inkubiert. Die Proben für SDS-PAGE Analysen mit anschließender Coomassie-Färbung wurden vor dem Gelauftrag mit SDS-Probenpuffer versetzt und 10 min bei 95°C gekocht. Kovalente Catcher-Tag Produkte sind mit einem roten Pfeil gekennzeichnet. Die Molekülvisualisierungen wurden mittels PyMOL angefertigt (grün: 3kptCCatcher, orange: 3kptCTag).

3.4 Kovalente Split-Isopeptid Systeme *in planta*

Nach den *in vitro* Experimenten zur Identifizierung, Charakterisierung und ersten Optimierung der neuen, kovalenten Split-Isopeptid Systeme, sollte die Funktionalität dieser Systeme auch *in planta* getestet werden. Die stabile Verknüpfung oder Immobilisierung von Zielproteinen über die kovalente und spezifische Interaktion zwischen Catcher-Protein und Tag-Peptid könnte viele Einsatzmöglichkeiten in der Pflanzenbiotechnologie finden und neue Anwendungen ermöglichen oder bereits bestehende Methoden durch den kovalenten Charakter dieser Catcher-Tag Systeme revolutionieren (siehe Diskussion). Neben den neu identifizierten Catcher-Tag Systemen, 4oq1 und 3kptC, wurden auch das Spy-System und das Snoop-System *in planta* analysiert, da zu diesem Zeitpunkt nicht bekannt war, ob diese

etablierten Systeme in der Pflanze auch die Fähigkeit zur Ausbildung von kovalenten Catcher-Tag Komplexen, basierend auf Isopeptidbindungen, besitzen. Die Fähigkeit zur Interaktion zwischen Tag-Peptid und jeweiligem Catcher-Protein *in planta* wurde indirekt über die spezifische Direktion von Catcher-Reporterkonstrukten an OEP7 (*outer envelope membrane protein* 7, Sequenz abgeleitet von At3g52420)-SpyTag markierte Chloroplasten am konfokalen Mikroskop (KLSM) analysiert (siehe Schema Abb.3.45).

Für die *in planta* Analysen wurden Genfusionskonstrukte (siehe Abb.3.45) über PCR-Methoden mit spezifischen Primern generiert und in den pflanzlichen Expressionsvektor pRB-35S ligiert (Klonierungen der Pflanzenexpressionskonstrukte siehe 2.6.3 und Tabelle 1.1, Primersequenzen siehe Tabelle A2 im Anhang), in Agrobakterien transformiert (2.5.1.3) und in *Nicotiana benthamiana* über *Agrobacterium tumefaciens*-vermittelte Transformation transient koexprimiert (2.5.2.2). Bei erfolgreicher Interaktion von Catcher und Tag *in planta*, wird das Reporter-Catcher Fusionsprotein an die Chloroplasten, an deren äußerer Membran OEP7-Tag Proteine gebunden sind, dirigiert (siehe Abb.3.45). Im Mikroskop sollte, im Falle einer Interaktion zwischen Catcher und Tag, eine Kolokalisierung der Reporterfluoreszenz (eGFP oder mCherry) mit der Autofluoreszenz der Chloroplasten (Chlorophyll) zu sehen sein. Die kovalente Natur dieser Catcher-Tag Interaktion wurde anschließend in Western Blot-Analysen, bei denen die homogenisierten Blattproben vor der SDS-PAGE in denaturierendem 4x Lämmli-Probenpuffer gekocht wurden (siehe Gliederungspunkt 2.6.4.1.2), analysiert. Bei positiver kovalenter Interaktion zwischen Catcher- und Tag-Fusionsproteinen ist neben den Substratbanden eine neue, höhermolekulare Produktbande im Western Blot zu erwarten.

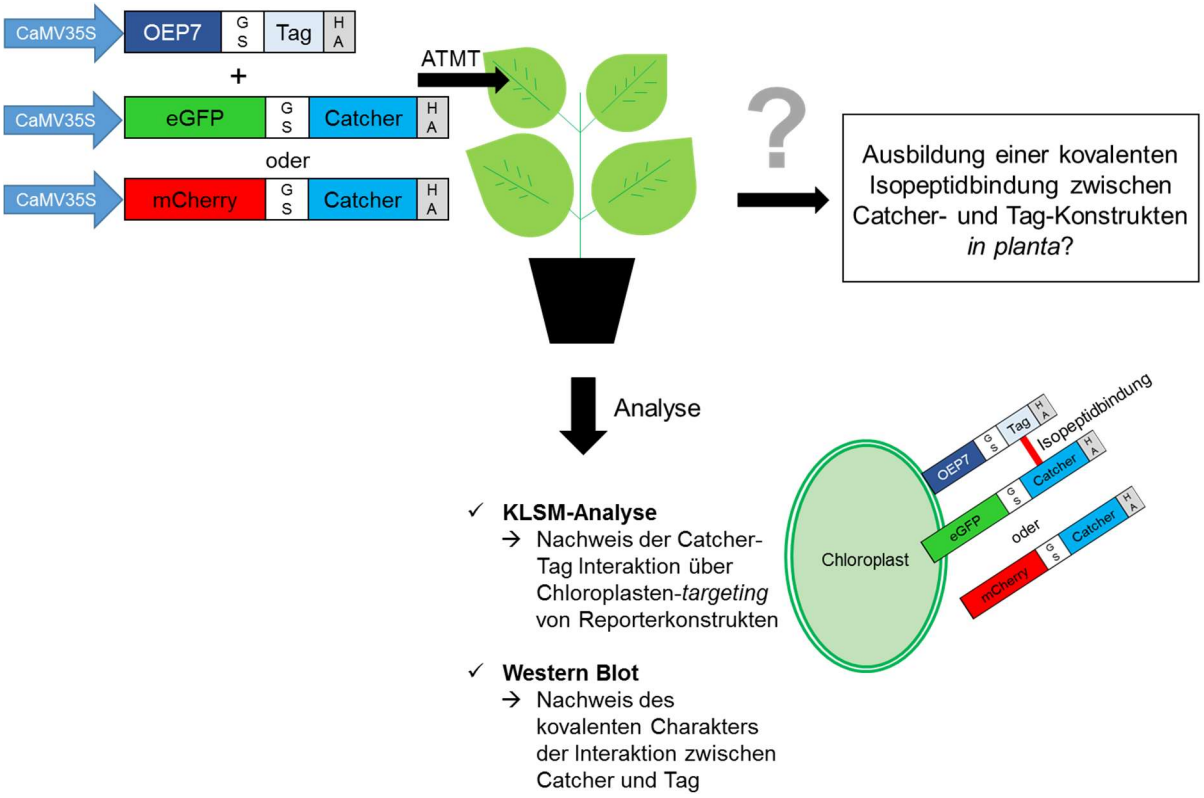


Abb.3.45: Schematische Darstellung des Vorgehens zur Analyse der Funktionalität kovalenter Catcher-Tag Systeme *in planta* (*Nicotiana benthamiana*). Catcher und Tag stehen hier für die kovalenten Split-Isopeptid Systeme Spy, Snoop, 4oq1 und 3kptC, wobei stets nur zusammengehörige Catcher-Tag Paare *in planta* getestet wurden. Die pRB-35S-Zielkonstrukte (OEP7-GS-Tag-HA und eGFP-GS-Catcher-HA oder mCherry-GS-Catcher-HA) wurden koinfiltriert und mittels *Agrobacterium tumefaciens*-vermittelter transienter Transformation (ATMT) in *Nicotiana benthamiana* Blätter koexprimiert (CaMV35S Promotor). Die Interaktion zwischen Catcher- und Tag-Fusionsproteinen wurde 72h nach Infiltration am konfokalen Mikroskop (KLSM-Analyse) untersucht. Die kovalente Natur dieser Catcher-Tag Interaktion wurde in Western Blot-Analysen (Kinetik-Experimente) nach Kochen der Proteinproben in denaturierendem 4x Lämmli-Puffer und SDS-PAGE verifiziert. GS steht für die Aminosäureabfolge GSGSGSG.

3.4.1 Spy-System *in planta*

Das SpyCatcher-SpyTag System funktioniert *in vitro* und *in vivo* intra- und extrazellulär (Zakeri et al., 2012, Veggiani et al., 2014, Gilbert et al., 2017). Bisher unbekannt ist jedoch, ob das etablierte und vielfältig eingesetzte Spy-System auch *in planta* funktioniert und spezifische, kovalente Catcher-Tag Komplexe ausgebildet werden.

In den Abbildungen 3.46 A und B konnte gezeigt werden, dass sowohl das OEP7-GS-SpyTag-HA Fusionsprotein (#167 in Tabelle 1.1) als auch das eGFP-GS-SpyCatcher-HA Fusionsprotein (#171 in Tabelle 1.1) erfolgreich transient in *Nicotiana benthamiana* exprimiert wurden. 24 Stunden nach Infiltration der *N. benthamiana* Blätter konnten bereits spezifische Signale der HA-markierten Zielproteine auf den erwarteten Größen im Western Blot detektiert werden (siehe Abb.3.46 A und B). Die kovalente Interaktion zwischen dem

SpyCatcher- und dem SpyTag-Fusionsprotein nach Koinfiltration und Koexpression konnte sowohl in Anti-HA Western Blots als auch in Anti-GFP Western Blots in Form von einer neu entstandenen Produktbande bei ca. 50 kDa eindeutig gezeigt werden (siehe Abb.3.46 C und D; rote Pfeile markieren die Produktbande).

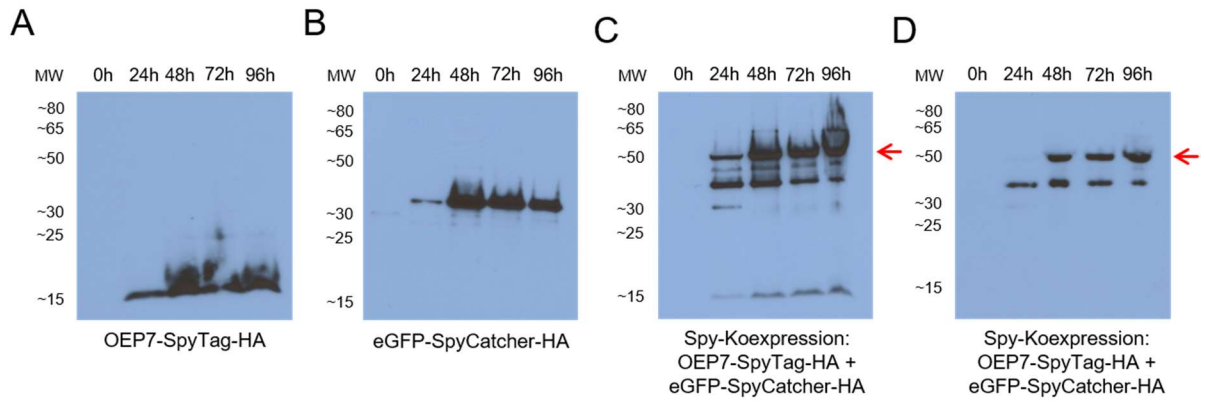


Abb.3.46: Western Blot-Analysen zum Nachweis der Expression und der kovalenten Interaktion der SpyCatcher- bzw. SpyTag-Konstrukte *in planta*. Die Spy-Konstrukte (OEP7-GS-SpyTag-HA #167 und eGFP-GS-SpyCatcher-HA #171) wurden über transiente Agrobakterium-vermittelte Transformation in *Nicotiana benthamiana* Blätter sowohl einzeln (A und B) als auch gemeinsam infiltriert (C und D). Für Expressions- bzw. Interaktionskinetiken wurden Blattproben der infiltrierten Stellen genommen und Gesamtproteinextrakte in denaturierendem SDS-Probenpuffer (4x Lämmli) hergestellt, die vor der SDS-PAGE mit anschließender Western Blot-Analyse 10 min bei 95°C gekocht wurden. Gezeigt sind: (A) Anti-HA-POD Western Blot (1:500) des OEP7-SpyTag-HA Konstruktes (10 kDa), (B) Anti-GFP Western Blot (Anti-GFP Antikörper 1:1000, rabbit anti-mouse-POD 1:10000) des eGFP-SpyCatcher-HA Konstruktes (38 kDa), (C) Anti-HA-POD Western Blot (1:500) der Spy-Koexpression, (D) Anti-GFP Western Blot (Anti-GFP Antikörper 1:1000, rabbit anti-mouse-POD 1:10000) der Spy-Koexpression. Kovalente Catcher-Tag Produkte (ca. 50 kDa) sind mit einem roten Pfeil gekennzeichnet.

Die Fähigkeit von SpyCatcher und SpyTag *in planta* miteinander zu interagieren, wurde zudem durch KLSM-Analysen, die 72h nach Infiltration durchgeführt wurden, bestätigt (siehe Abb.3.47). Nach Koinfiltration des SpyTag-Peptids, fusioniert an den OEP7-Chloroplastenanker, und des eGFP-SpyCatcher Fusionsproteins konnte die GFP-Fluoreszenz eindeutig und überwiegend an den Chloroplasten (Chlorophyll-Autofluoreszenz: blau) detektiert werden (siehe Abb.3.47 A Kolokalisierung des GFP-Signals mit der Autofluoreszenz der Chloroplasten). Ohne gleichzeitige Expression des Chloroplastenanker-SpyTag Konstruktes lokalisierte das eGFP-SpyCatcher Protein dagegen eindeutig im Cytoplasma und im Zellkern (siehe Abb.3.47 B). Über diese Kolokalisierung der GFP-Fluoreszenz und der Autofluoreszenz der Chloroplasten nach Koexpression von OEP7-SpyTag und eGFP-SpyCatcher konnte gezeigt werden, dass das eGFP-SpyCatcher Protein über die Interaktion mit OEP7-SpyTag an die Chloroplasten dirigiert wird (siehe Abb.3.47 A), da das eGFP-SpyCatcher Protein, wenn es alleine transient in *N. benthamiana* exprimiert wird, überwiegend im Cytoplasma und im Nukleus lokalisiert ist (siehe Abb.3.47 B).

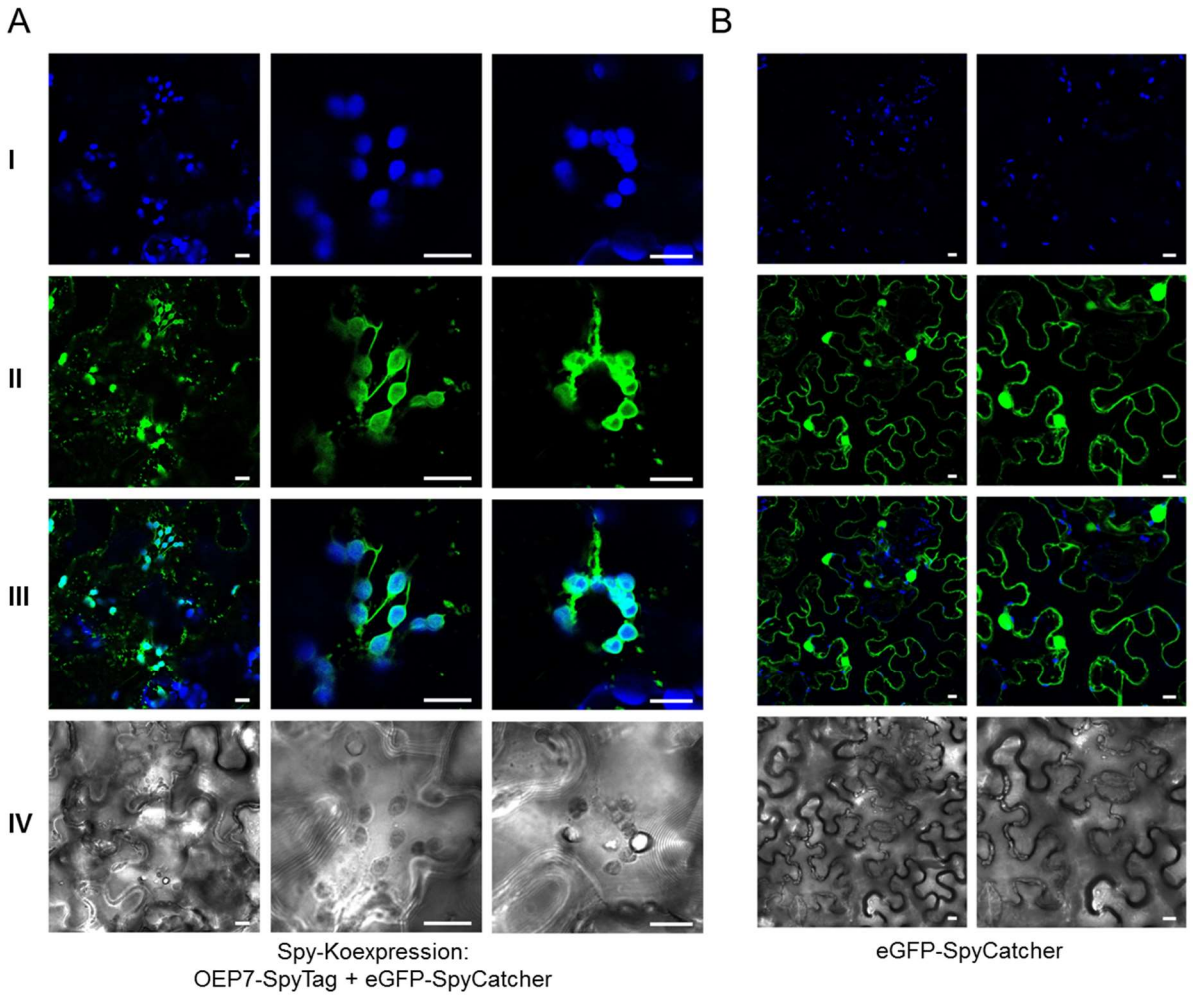


Abb.3.47: Repräsentative KLSM-Bilder 72h nach Infiltration von *Nicotiana benthamiana* Blättern mit den Konstrukten OEP7-SpyTag und/oder eGFP-SpyCatcher. Die Konstrukte OEP7-GS-SpyTag-HA (#167) und eGFP-GS-SpyCatcher-HA (#171) wurden mittels Agrobakterien-vermittelter Transformation transient in *Nicotiana benthamiana* Blättern exprimiert und infiltrierte Blattstellen 72h nach der Infiltration unter dem KLSM betrachtet. (A) Koexpression von OEP7-SpyTag und eGFP-SpyCatcher, (B) alleinige Expression des Reporterkonstruktes eGFP-SpyCatcher. I: Chlorophyll-Autofluoreszenz (blau), II: eGFP-Fluoreszenz (grün), III: Überlagerung, IV: Durchlicht. Der Größenstandard beträgt 10µm.

3.4.2 Snoop-System *in planta*

Nach Koinfiltration von *N. benthamiana* Blättern mittels *Agrobacterium tumefaciens*-vermittelter Transformation konnte weder die Proteinexpression von OEP7-GS-SnoopTag-HA (#168 in Tabelle 1.1, 10 kDa) noch von eGFP-GS-SnoopCatcher-HA (#173 in Tabelle 1.1, 41 kDa) in Anti-HA Western Blot-Analysen nachgewiesen werden (siehe Abb.3.48). Dass der Anti-HA Western Blot aber technisch funktioniert hat, zeigte die Positivkontrolle (PK) (siehe Abb.3.48). Aussagen über eine mögliche kovalente Interaktion der SnoopCatcher- und SnoopTag-Fusionskonstrukte *in planta* konnten deshalb nicht gemacht werden.

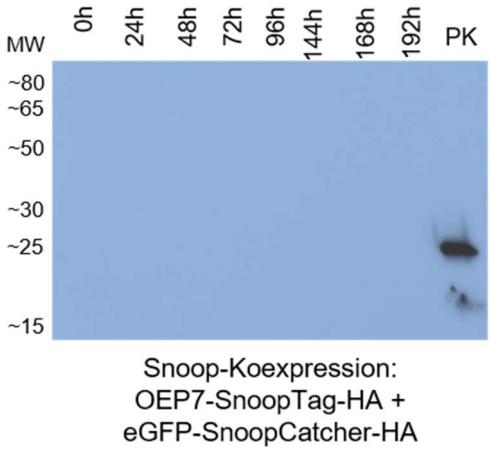


Abb.3.48: Anti-HA (1:500) Western Blot-Analyse zum Nachweis der Expression und der kovalenten Interaktion zwischen SnoopCatcher- und SnoopTag-Konstrukten *in planta*. Die Snoop-Konstrukte (OEP7-GS-SnoopTag-HA #168 und eGFP-GS-SnoopCatcher-HA #173) wurden über transiente Agrobakterium-vermittelte Transformation in *Nicotiana benthamiana* Blätter koinfiltriert. Für Expressions- bzw. Interaktionskinetiken wurden Blattproben infiltrierter Stellen genommen und Gesamtproteinextrakte in denaturierendem SDS-Probenpuffer (4x Lämmli) hergestellt, die vor der SDS-PAGE mit anschließender Western Blot-Analyse 10 min bei 95°C gekocht wurden.

3.4.3 3kptC-System *in planta*

Die Fähigkeit des 3kptC-Systems zur Ausbildung von kovalenten Catcher-Tag Komplexen *in planta* konnte ebenfalls nicht beurteilt werden, da nur das OEP7-GS-3kptCTag-HA Fusionsprotein (#169 in Tabelle 1.1), aber nicht das eGFP-GS-3kptCCatcher-HA Fusionsprotein (#174 in Tabelle 1.1) in *N. benthamiana* exprimiert wurde (siehe Abb.3.49). Für das OEP7-3kptCTag-HA Konstrukt (10 kDa) konnte 48h nach der Infiltration in *N. benthamiana* ein spezifisches Anti-HA Signal im Western Blot detektiert werden, während das eGFP-3kptCCatcher-HA (39 kDa) korrespondierende Signal über den gesamten Analysezeitraum ausblieb (siehe Abb.3.49).

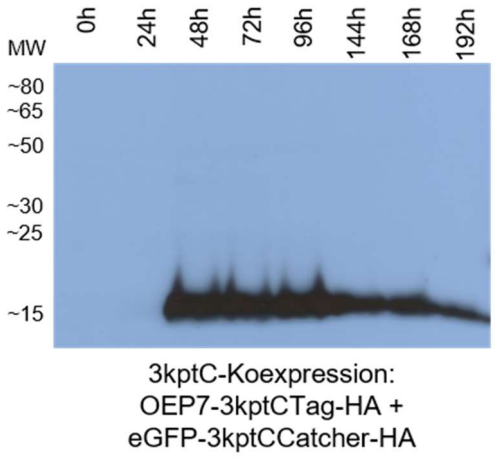


Abb.3.49: Anti-HA (1:500) Western Blot-Analyse zum Nachweis der Expression und der kovalenten Interaktion zwischen 3kptCCatcher- und 3kptCTag-Konstrukten *in planta*. Die 3kptC-Konstrukte (OEP7-GS 3kptCTag-HA #169 und eGFP-GS-3kptCCatcher-HA #174) wurden über transiente Agrobakterium-vermittelte Transformation in *Nicotiana benthamiana* Blätter koinfiltriert. Für Expressions- bzw. Interaktionskinetiken wurden Blattproben infiltrierter Stellen genommen und Gesamtproteinextrakte in denaturierendem SDS-Probenpuffer (4x Lämmli) hergestellt, die vor der SDS-PAGE mit anschließender Western Blot-Analyse 10 min bei 95°C gekocht wurden.

3.4.4 4oq1-System *in planta*

Bei der Analyse der transienten Expression der 4oq1-Fusionsproteine (OEP7-GS-4oq1Tag-HA (#170 in Tabelle 1.1) und mCherry-GS-4oq1Catcher-HA (#175 in Tabelle 1.1)) in *Nicotiana benthamiana* mittels Anti-HA Western Blot-Analysen konnte gezeigt werden, dass

beide Konstrukte erfolgreich exprimiert wurden (siehe Abb.3.50 A und B). Das OEP7-4oq1Tag Konstrukt (10 kDa) wurde 2 Tage nach der Infiltration über ATMT (*Agrobacterium tumefaciens*-vermittelte Transformation) transient in *N. benthamiana* Blättern exprimiert, während die Expression des mCherry-4oq1Catcher Konstruktes (39 kDa) schon 24h nach der Infiltration detektiert werden konnte (siehe Abb.3.50 A und B). Beide Konstrukte wurden über einen längeren Zeitraum (Analyse bis 192h nach Infiltration) transient *in planta* exprimiert, wobei sich mCherry-typische Abbauprodukte zeigten (siehe Abb.3.50 B und C). Die Analyse der Koexpression beider Konstrukte ergab, dass 72h nach der Koinfiltration eine neue, höhermolekulare Bande im Anti-HA Western Blot auftauchte (siehe Abb.3.50 C, roter Pfeil), die dem kovalenten Produkt aus OEP7-4oq1Tag und mCherry-4oq1Catcher entspricht. Somit konnte gezeigt werden, dass die Fusionsproteine OEP7-4oq1Tag und mCherry-4oq1Catcher kovalent *in planta* miteinander reagieren und einen stabilen Catcher-Tag Komplex bilden.

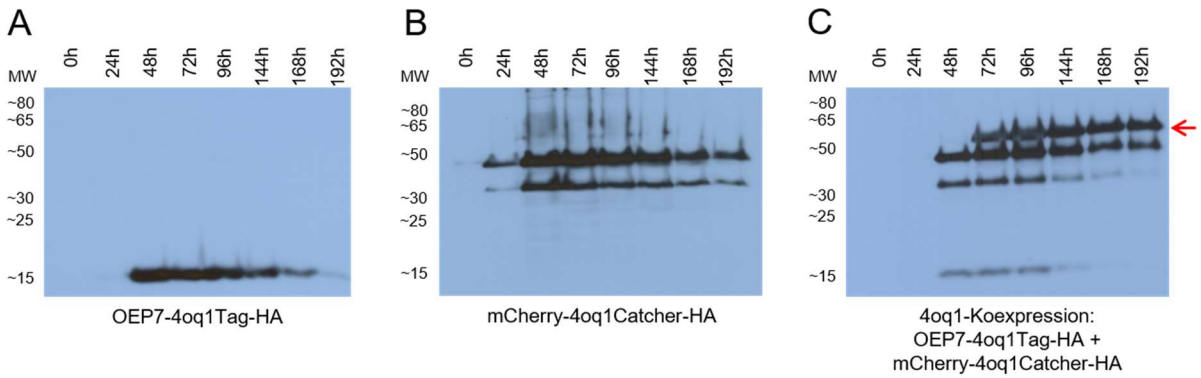


Abb.3.50: Anti-HA Western Blot-Analysen zum Nachweis der Expression und der kovalenten Interaktion der 4oq1Catcher- bzw. 4oq1Tag-Konstrukte *in planta*. Die 4oq1-Konstrukte (OEP7-GS-4oq1Tag-HA #170 und mCherry-GS-4oq1Catcher-HA #175) wurden über transiente Agrobakterium-vermittelte Transformation in *Nicotiana benthamiana* Blättern sowohl einzeln (A und B) als auch gemeinsam infiltriert (C). Für Expressions- bzw. Interaktionskinetiken wurden Blattproben infiltrierter Stellen genommen und Gesamtproteinextrakte in denaturierendem SDS-Probenpuffer (4x Lämmli) hergestellt, die vor der SDS-PAGE mit anschließender Western Blot-Analyse 10 min bei 95°C gekocht wurden. Gezeigt sind Anti-HA Western Blot-Analysen (1:500) (A) des OEP7-4oq1Tag-HA Konstruktes (10 kDa), (B) des mCherry-4oq1Catcher-HA Konstruktes (39 kDa) und (C) der 4oq1-Koexpression. Kovalente Catcher-Tag Produkte (ca. 50 kDa) sind mit einem roten Pfeil gekennzeichnet.

Die Interaktion zwischen den 4oq1-Fusionsproteinen OEP7-4oq1Tag und mCherry-4oq1Catcher in *N. benthamiana* Blättern 72h nach transienter Agrobakterium-vermittelter Transformation konnte ebenfalls in mikroskopischen Kolokalisierungsanalysen am KLSM verifiziert werden (siehe Abb.3.51 A). Wurden die Proteine OEP7-4oq1Tag und mCherry-4oq1Catcher gemeinsam exprimiert, wurde 72h nach Infiltration die rote mCherry-Fluoreszenz vermehrt und überwiegend an den Chloroplasten detektiert (siehe Abb.3.51 A), wohingegen das mCherry-4oq1Catcher Fusionsprotein im Cytoplasma und im Zellkern

lokalisierte, wenn es alleine, das heißt ohne Chloroplastenanker-4oq1Tag, in *N. benthamiana* transient exprimiert wurde (siehe Abb.3.51 B).

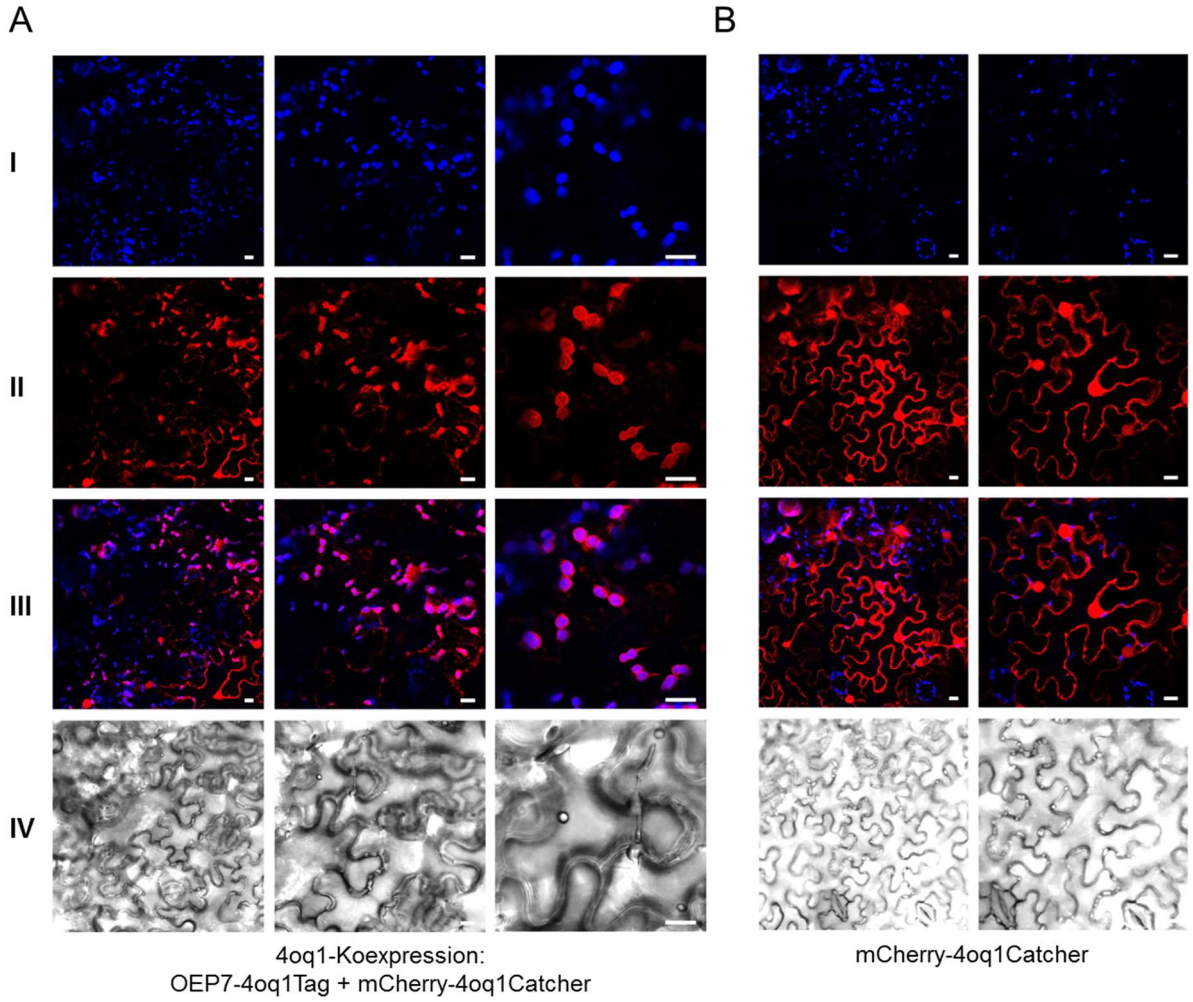


Abb.3.51: Repräsentative KLSM-Bilder 72h nach Infiltration von *Nicotiana benthamiana* Blättern mit den Konstrukten OEP7-4oq1Tag und/oder mCherry-4oq1Catcher. Die Konstrukte OEP7-GS-4oq1Tag-HA (#170) und mCherry-GS-4oq1Catcher-HA (#175) wurden mittels Agrobakterien-vermittelter Transformation transient in *Nicotiana benthamiana* Blättern exprimiert und infiltrierte Blattstellen 72h nach der Infiltration unter dem KLSM betrachtet. (A) Koexpression von OEP7-4oq1Tag und mCherry-4oq1Catcher, (B) alleinige Expression des Reporterkonstruktes mCherry-4oq1Catcher. I: Chlorophyll-Autofluoreszenz (blau), II: mCherry-Fluoreszenz (rot), III: Überlagerung, IV: Durchlicht. Der Größenstandard beträgt 10µm.

3.5 Anwendungen der kovalenten Catcher-Tag Systeme

Aufgrund der Stabilität, Effizienz, Spezifität und Robustheit der Split-Isopeptid Systeme, vor allem des Spy- und Snoop-Systems, ergeben sich zahlreiche, vor allem biotechnologische, Anwendungen. Theoretisch können diese kovalenten Split-Isopeptid Systeme für alle Anwendungen und Techniken interessant sein, die von stabilen, irreversiblen Protein-Protein Interaktionen profitieren.

3.5.1 Protein-Immobilisierung an Magnetpartikeln durch Catcher-Tag Funktionalisierung

Über die kovalente und spezifische Interaktion zwischen Catcher und Tag (vollständige Orthogonalität zumindest für das Spy- und das Snoop-System, Veggiani et al., 2016) können Catcher/Tag-funktionalisierte Proteine (SpyTag- oder SpyCatcher-Fusionsproteine) an bestimmten Catcher/Tag-funktionalisierten Trägern oder Matrizes immobilisiert werden. Werden multiple Proteine auf demselben Träger koimmobilisiert, können stabile Multiproteinkomplexe generiert werden. Bei der Verwendung spezifischer Systeme kann sowohl die Reihenfolge als auch die Struktur dieser Proteinkomplexe zeitlich und räumlich genau kontrolliert und programmiert werden. Um das Prinzip der irreversiblen Kopplung von Proteinen über kovalente Split-Isopeptid Systeme zu überprüfen, wurde das bislang am besten charakterisierte und optimierte SpyTag-SpyCatcher System als Modellsystem verwendet. Für ein *proof of principle*, ob Proteine an bestimmte Träger über das Spy-System gekoppelt werden können, wurden Fluoreszenzproteine als Reporter verwendet, deren Lokalisierung einfach über mikroskopische Studien analysiert werden konnte. Für einen ersten Test wurden Streptavidin-funktionalisierte magnetische Partikel (GE Healthcare Life Sciences: 37 – 100 µm) verwendet, die über die sehr starke Affinitätsinteraktion mit Biotin-modifizierten SpyTag-Proteinen beladen wurden (Kooperation mit der Professur für Pharmazeutische Chemie unter der Leitung von Prof. Dr. Jutta Eichler, Dr. Andrea Lischke: Herstellung der synthetischen Biotin-SpyTag Peptide und Beladung der Streptavidin-*beads*). Der zu immobilisierende Gegenpart waren rekombinant hergestellte Reporter-SpyCatcher Fusionsproteine (mCherry-GS-SpyCatcher (#35 in Tabelle 1.1), eGFP-GS-SpyCatcher (#36 in Tabelle 1.1) und mTagBFP-GS-SpyCatcher (#38 in Tabelle 1.1); Genkonstrukte siehe Abb.3.52 A), die nach der Funktionalisierung der Partikel mit dem SpyTag-Protein kovalente und spezifische Catcher-Tag Komplexe ausbilden sollten (siehe Schema in Abb.3.52 B). Die KLSM-Analyse, nach der Beladung der magnetischen Partikel, bestätigte die Immobilisierung der SpyCatcher-Reporter Fusionsproteine an den magnetischen Partikeln (siehe Abb.3.52). Auch die Kolokalisierung bzw. Koimmobilisierung der Reporterproteine, eGFP-GS-SpyCatcher, mTagBFP-GS-SpyCatcher und mCherry-GS-SpyCatcher an demselben Partikel konnte über spezifische Anregungen der verschiedenen Fluoreszenzproteine nachgewiesen werden (siehe Abb.3.52 C - G).

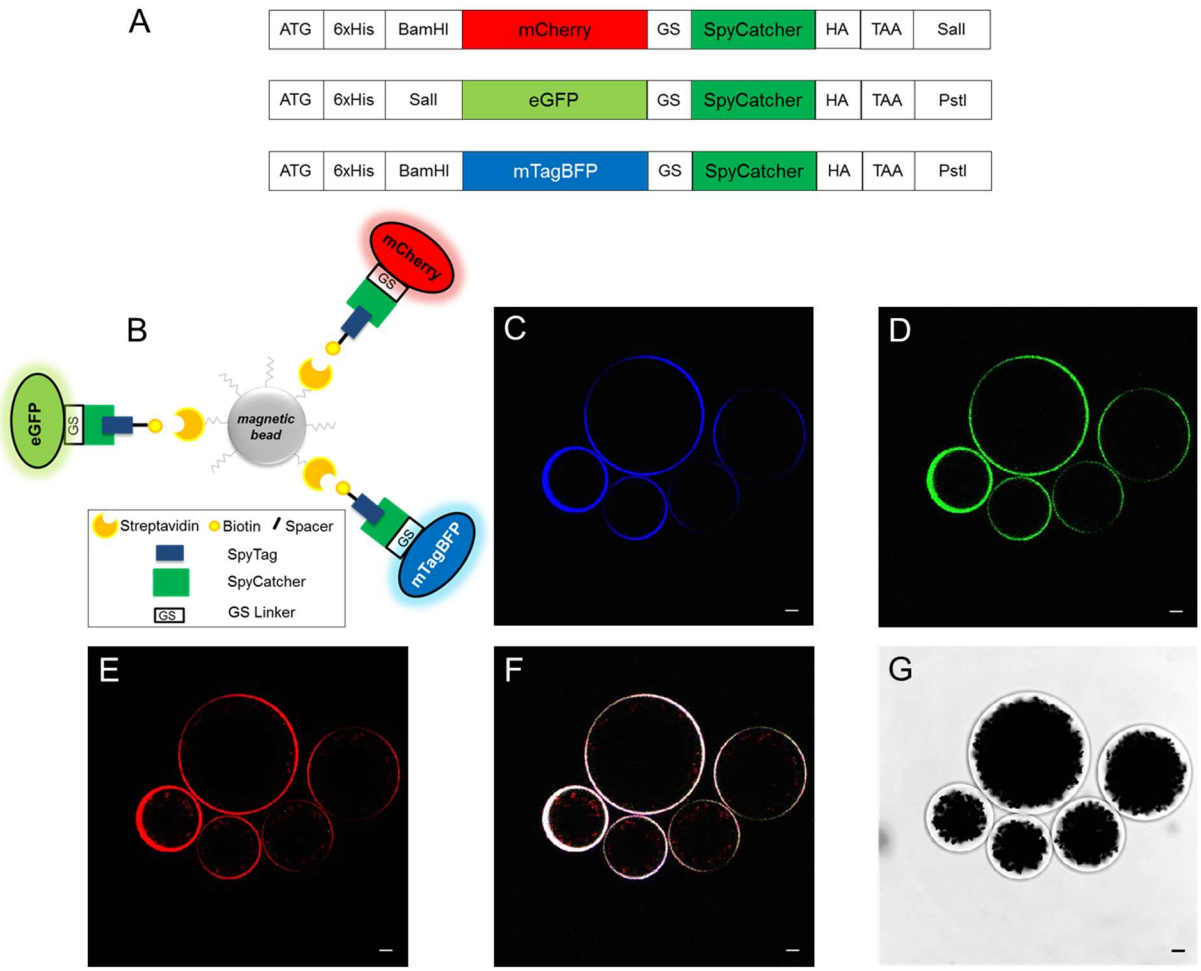


Abb.3.52: Analyse nicht-kovalenter mit SpyTag-funktionalisierter Streptavidin-Magnetpartikel. Zur Beladung wurden die SpyTag-funktionalisierten GE-Magnetpartikel (37 - 100 µm) mit einer äquimolaren Mischung (1:1:1) aus rekombinant in *E. coli* hergestellten und aufgereinigten Reporter-SpyCatcher Fusionsproteinen (mTagBFP-GS-SpyCatcher (#38), eGFP-GS-SpyCatcher (#36) und mCherry-GS-SpyCatcher (#35)) inkubiert. Schematische Darstellung der Genkonstrukte, die für die Reporter-SpyCatcher Fusionsproteine kodieren (A). Gezeigt sind repräsentative KLSM-Bilder (C - G) trifunktionalisierter GE-Streptavidin-Magnetpartikel (Schema in B). Angeregt wurden die verschiedenen Fluorophore mit verschiedenen Wellenlängen im *sequential scan*-Modus (UV-Laser 405 (405 Diode) 415nm-490nm → mTagBFP (C); 488 Laser (Argon-Laser) 496nm-556nm → eGFP (D); 561 Laser (DPSS561) 571nm-631nm → mCherry (E)). In F sind die Fluoreszenzsignale aller drei Kanäle überlagert und G ist das Durchlichtbild. Der Größenstandard beträgt 10 µm. GS steht für die Aminosäureabfolge GSGSGSG.

Um ein vollständig kovalentes System zu etablieren, wurden die Streptavidin-Partikel durch magnetische Partikel, deren Oberfläche mit Maleimid-Gruppen aktiviert ist, ersetzt (Cube Biotech, Ocean NanoTech). Über die kovalente, irreversible Thioether-Reaktion des Maleimids mit freien Thiolgruppen können Proteine über zugängliche, exponierte Cysteinreste gezielt, direkt und stabil an die Magnetpartikel gekoppelt werden (siehe Abb.2.5). Aus diesem Grund wurde über gezielte Mutagenese unter Verwendung spezifischer Primer ein SpyCatcher-Genkonstrukt erstellt, das für einen N-terminalen Cysteinrest am SpyCatcher-Protein kodiert (siehe Abb.3.53 A I, #147 in Tabelle 1.1, Primersequenzen im Anhang Tabelle A2). Nach heterologer Expression in M15 [pREP4] *E.*

coli Zellen und Ni-NTA Affinitätsaufreinigung des Cystein-SpyCatcher Proteins (siehe Abb.3.53 A II) wurden die Maleimid-*beads* über Thioether-Kopplung mit Cystein-SpyCatcher Protein beladen/ funktionalisiert (siehe Kapitel 2.8.1). Über die kovalente, spezifische SpyCatcher-SpyTag Interaktion sollten dann SpyTag-Fusionsproteine an die Partikel immobilisiert werden (siehe Schema in Abb.3.53 B). Die Funktionalität dieser kovalenten Kopplung wurde über aufgereinigte SpyTag-eGFP Fusionsproteine am Fluoreszenzmikroskop getestet. Es konnte gezeigt werden, dass alle verwendeten Partikel, unabhängig von ihrer Größe, erfolgreich mit Cystein-SpyCatcher Protein beladen wurden, da die Inkubation mit SpyTag-eGFP Fluoreszenzproteinen zu einem spezifischen GFP-Signal an den *beads* führte (siehe Abb.3.53 C I und II).

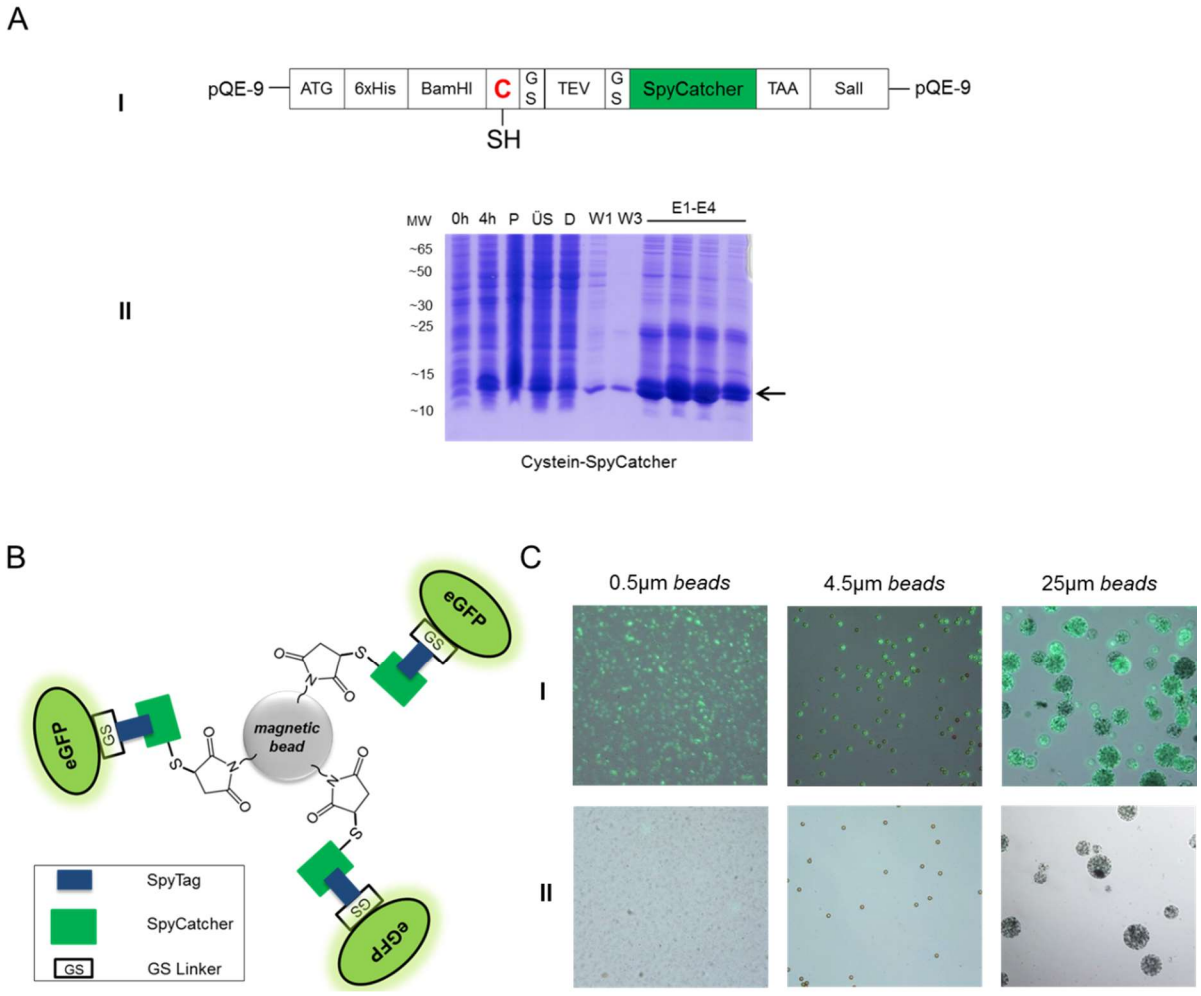


Abb.3.53: Analyse kovalenter mit SpyCatcher-funktionalisierter Maleimid-Magnetpartikel. A: (I) Schema des Cystein-GS-TEV-GS-SpyCatcher DNA-Konstruktes (#147), (II) Coomassie-Gel der Ni-NTA Affinitätsaufreinigung des in *E. coli* M15 [pREP4] Zellen rekombinant hergestellten Cystein-SpyCatcher Proteins (Proben der Expressionskinetik eingestellt auf dieselbe oD_{600}: 0h (vor Induktion) und 4h nach Induktion mit 0,5mM IPTG, Proben der Aufreinigung: P:Pellet nach Zelllyse, ÜS: Überstand nach Zelllyse, D: Durchfluss, W1, W3: Waschschritte, E1-E4: Eluate). B: Schema der Beladung der über Thioether-Bindung gekoppelten SpyCatcher-*beads* mit dem Reporterprotein SpyTag-GSGSGSG-eGFP, C: Analyse am Fluoreszenzmikroskop (I: Durchlicht + Fluoreszenz (GFP-Kanal) der SpyCatcher-*beads* inkubiert mit SpyTag-eGFP, II: Durchlicht + Fluoreszenz (GFP-Kanal) der Negativkontrolle: unbeladene Maleimid-*beads* (ohne SpyCatcher) inkubiert mit SpyTag-eGFP.

3.5.2 Generierung von stabilen Multiproteinkomplexen

Neben der kovalenten Immobilisierung von Proteinen über Catcher-Tag Systeme durch reziproke Funktionalisierung sowohl des Trägers als auch der zu immobilisierenden Proteine mit Catcher-Protein bzw. Tag-Peptid, können Proteine auch post-translational, stabil miteinander verknüpft werden. Dadurch, dass die Catcher- und Tag-Module genetisch kodiert sind, können Catcher- bzw. Tag-Genfusionen einfach kloniert (2.6.2), rekombinant in *E. coli* Zellen hergestellt (2.6.4.6) und aufgereinigt werden (2.6.4.7). Durch einfaches Mischen der kompatiblen Reaktionspartner sollen kovalente Proteinkomplexe entstehen. Um auch dieses Prinzip zu testen, wurden Catcher- bzw. Tag-modifizierte Reporterfusionsproteine verwendet (siehe Abb.3.54 A). Zur besseren Kontrollierbarkeit und Programmierbarkeit der Zusammensetzung (Reihenfolge und Anzahl der Proteine in einem Komplex) der Multiproteinkomplexe wurde zusätzlich zu dem SpyCatcher-SpyTag System das orthogonale SnoopCatcher-SnoopTag System verwendet (siehe Abb.3.54 B). Sowohl Coomassie-, Western Blot-Analysen (siehe Abb.3.54 C und D) als auch die Visualisierung der Proteingele unter UV-Licht (siehe Abb.3.54 E) bestätigten die Generierung eines funktionellen, kovalent verknüpften Trimers aus blau, grün und rot fluoreszierenden Proteinen (Schema siehe Abb.3.54 B).

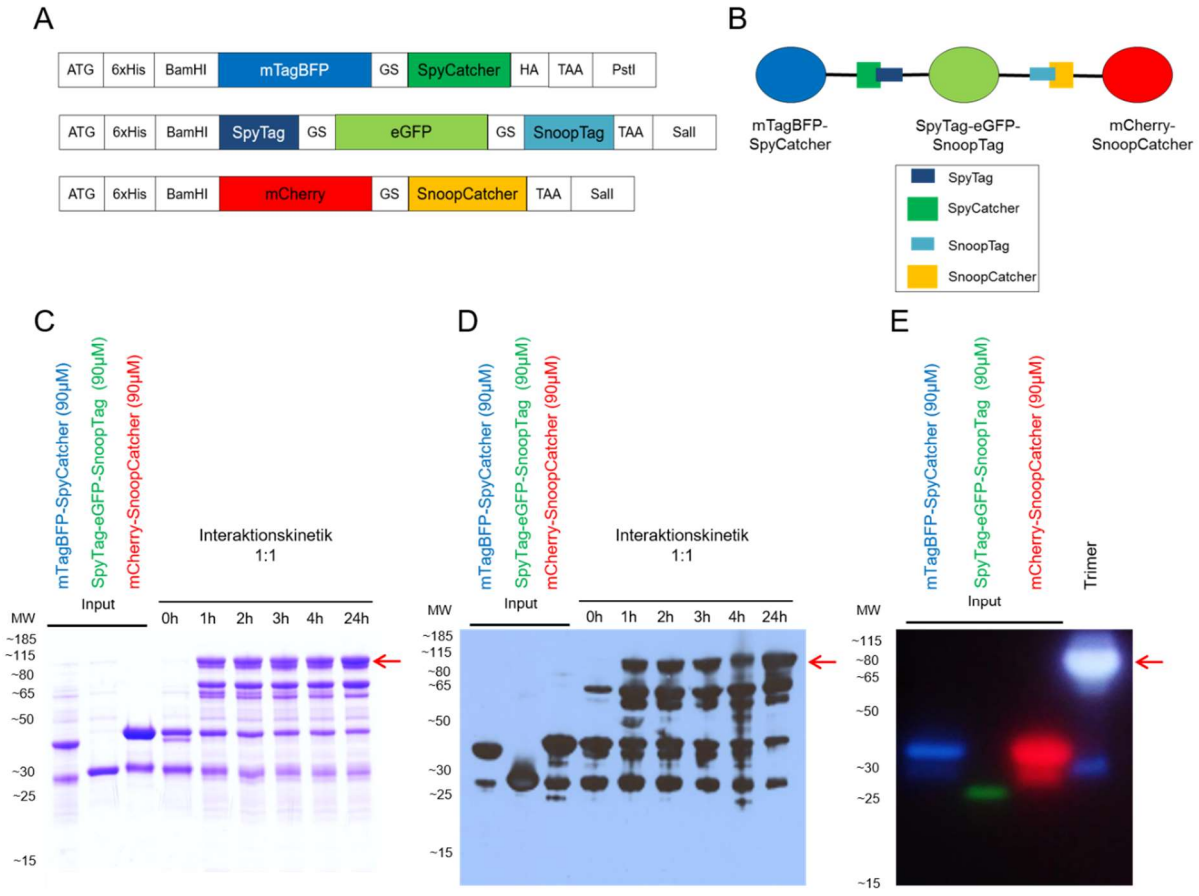


Abb.3.54: Herstellung eines Proteinkomplexes (weißes Trimer) vermittelt über die orthogonalen Split-Isopeptid Systeme SpyCatcher-SpyTag und SnoopCatcher-SnoopTag. Gezeigt ist die schematische Darstellung der zugrundeliegenden DNA-Einzelkomponenten des Trimers (#38, #81 und #82 in Tabelle 1.1); GS-linker besteht aus GSGSGSG (A), die schematische Assemblierung der Proteine zu einem Trimer über das Spy- und das Snoop-System (B), ein Coomassie-Proteingel (C) und Anti-His Western Blot (1:5000) (D) der äquimolaren Interaktionskinetiken (30 µm Endkonzentration) der drei heterolog exprimierten und aufgereinigten Proteinkomponenten (siehe Input). Die Proben wurden vor dem Gelauftrag 10 min bei 95°C in 4x Lämmli-Puffer zur Verifizierung des kovalenten Charakters der Bindungen gekocht. In E ist ein natives Proteingel der drei einzelnen Proteinkomponenten (siehe A bzw. B) und des weißen Trimers, das nach der kovalenten Interaktion nochmals über Gelfiltration (Äkta pure Superdex Highload, Dr. Jörg Hofmann) aufgereinigt und aufkonzentriert wurde (20K MWCO Pierce Concentrator), bei UV-Anregung gezeigt. Die Proben wurden für das native Proteingel zwar mit 4x Lämmli-Puffer versetzt, aber nicht gekocht (E). Das Trimer ist durch rote Pfeile hervorgehoben.

3.5.3 Isolierung von Chloroplasten über das SpyCatcher-SpyTag System

Aufgrund der Spezifität, Effizienz und Stabilität des SpyCatcher-SpyTag Systems ergeben sich noch weitere Anwendungsmöglichkeiten, weit über die Arbeit mit isolierten, aufgereinigten Proteinen hinaus. Nachdem auch gezeigt werden konnte, dass das SpyCatcher-SpyTag System *in planta* funktioniert und SpyCatcher- bzw. SpyTag-Fusionsproteine kovalente Komplexe ausbilden (siehe Kapitel 3.4.1), sollte getestet werden, ob das Spy-System für die spezifische Isolierung von Chloroplasten aus homogenisiertem Pflanzenfiltrat geeignet ist. Mit Cystein-SpyCatcher (#147 in Tabelle 1.1) beladene Magnetpartikel der Größe 4,5 µm (Abb.3.53, Beladung siehe Kapitel 2.8.1) wurden verwendet um Chloroplasten, die zuvor durch OEP7-SpyTag Fusionsprotein (#167 in Tabelle 1.1) markiert wurden (transiente Agrobakterium-vermittelte Transformation des Plasmids #167 in *N. benthamiana*), über die irreversible SpyCatcher-SpyTag Interaktion aus einem Proteingemisch/ Pflanzenfiltrat und unter Anwendung von Kraft (magnetische Kräfte) zu isolieren (Schema siehe Abb.3.55).

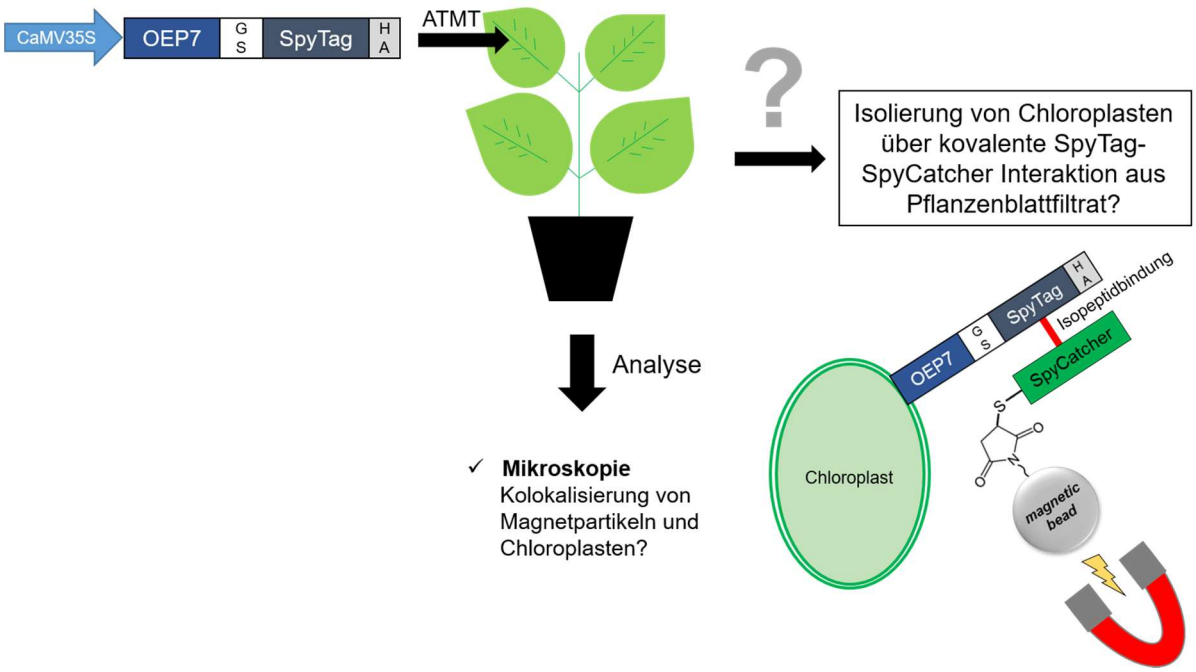


Abb.3.55: Schematische Darstellung der Chloroplasten-Isolierung über das SpyCatcher-SpyTag System (Chloroplasten-Trap). Das Konstrukt zur spezifischen SpyTag-Markierung der Chloroplasten (OEP7-GS-Tag-HA #167) wurde mittels *Agrobacterium tumefaciens*-vermittelter Transformation (ATMT) transient in *Nicotiana benthamiana* Blättern exprimiert. 48h nach Infiltration wurden Chloroplasten aus Gesamtblattfiltrat über SpyCatcher-funktionalisierte Magnetpartikel extrahiert.

Erste Versuche zur Isolierung von Chloroplasten (siehe Kapitel 2.8.2) zeigten, dass unter Verwendung von SpyCatcher-Magnetpartikeln und SpyTag-markierten Chloroplasten im Vergleich mit den durchgeführten Negativkontrollen, wo je eine der beiden Spy-Komponenten (entweder SpyCatcher auf den Magnetpartikeln oder OEP7-GS-SpyTag-HA an den Chloroplasten) fehlte, mehr Chloroplasten (rote Chlorophyll-Autofluoreszenz) aus dem Pflanzenblattfiltrat extrahiert wurden (siehe Abb.3.56 A und B). Waren beide Spy-Reaktionspartner vorhanden, war unter analogen Versuchsbedingungen und analoger Durchführung stets die Ausbeute an Chloroplasten an den *beads* größer als wenn eine der Komponenten fehlte. Dass über einige Magnetpartikel der Negativkontrollen trozdem Chloroplasten aus dem Gesamtblattfiltrat isoliert wurden, könnte durch unspezifische Bindungen entweder der Chloroplasten oder der SpyCatcher-Proteine auf den Magnetpartikeln erklärt werden.

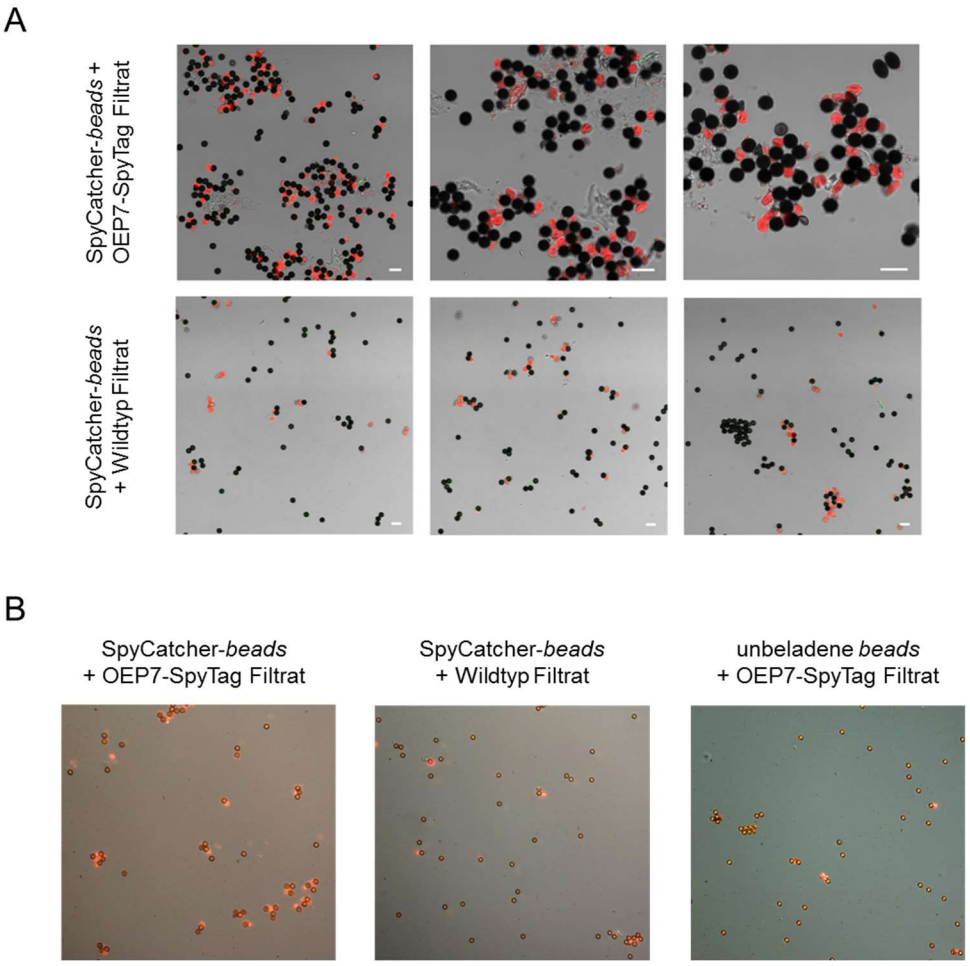


Abb.3.56: Repräsentative Mikroskopiebilder der Chloroplasten-Isolierung über das SpyCatcher-SpyTag System. Die Chloroplasten-Isolierung mithilfe von SpyCatcher-Magnetpartikeln (4,5 µm) wurde 48h nach Infiltration der *Nicotiana benthamiana* Blätter, die über Agrobakterium-vermittelte Transformation das Protein OEP7-GS-SpyTag-HA (#167) exprimieren, durchgeführt. SpyCatcher-*beads* wurden mit Blattfiltrat inkubiert, gewaschen und anschließend sowohl unter dem KLSM (A; Größenstandard beträgt 10 µm) als auch unter dem normalen Fluoreszenzmikroskop (B) betrachtet.

4 Diskussion

4.1 Proteine: Direkte Sequenz-Struktur-Funktionsbeziehung – Segen und Fluch für das „Protein Engineering“

4.1.1 Proteine besitzen nur eine geringe Konformationsstabilität

Proteine sind komplexe und fragile Gebilde, die selbst im gefalteten Zustand nur eine geringe Stabilität aufweisen (Ponnuswamy und Gromiha, 1994, Kang und Baker, 2011, Gromiha, 2007, Kwon et al., 2017, Kang und Baker, 2009). Die Konformationsstabilität von Proteinen im gefalteten und im ungefalteten Zustand unterscheidet sich kaum voneinander, sodass die Differenz der freien Energie zwischen dem gefalteten und dem ungefalteten Zustand unter physiologischen Bedingungen sehr gering ist (ΔG = 5 - 20 kcal mol^{-1}) (Ponnuswamy und Gromiha, 1994, Pace et al., 1996). Der Grad zwischen nativ gefaltetem, aktivem Protein und ungefaltetem, inaktivem Protein ist somit nur sehr schmal. Der Hauptunterschied dabei ist, dass die Interaktionen im gefalteten Zustand hauptsächlich intramolekularer Natur sind, während intermolekulare Interaktionen, auch mit Molekülen des Lösungsmittels, im ungefalteten Zustand überwiegen (Ponnuswamy und Gromiha, 1994). Folglich hängt die Stabilität gefalteter Proteinstrukturen maßgeblich von den zahlreichen, nicht-kovalenten Interaktionen (hydrophobe, elektrostatische und Van-der-Waals Wechselwirkungen und Wasserstoffbrücken) innerhalb einer Polypeptidkette ab, die den hohen destabilisierenden entropischen Kräften des ungefalteten Zustands entgegenwirken, diese überwinden und schließlich die Proteinfaltung begünstigen (Pace et al., 1996). Der hydrophobe Effekt spielt dabei eine Hauptrolle und stellt die treibende, initiale Kraft zur Faltung der Proteine dar, während die anderen Kräfte, wie Wasserstoffbrückenbindungen, Van-der-Waals Kräfte und elektrostatische Wechselwirkungen vor allem die Struktur und Form definieren, stabilisieren und einer erneuten Entfaltung des Proteins entgegenwirken (Ponnuswamy und Gromiha, 1994). Das Zusammenspiel und die feine Balance zwischen stabilisierenden und destabilisierenden Kräften, die bestimmen ob sich ein Protein faltet oder nicht bzw. gefaltet bleibt oder wieder entfaltet, macht deutlich wie empfindlich und komplex die Prozesse sind, die zu einem nativ gefalteten, aktiven Protein führen und wie leicht dieser Zustand wieder umgekehrt werden kann.

4.1.2 Proteine sind leicht zu manipulieren

Durch die genetische Kodierung von Proteinen und die direkte Beziehung zwischen Aminosäuresequenz, Proteinstruktur und Proteinfunktion (siehe Einleitung) birgt das Forschungsgebiet des ,,Protein Engineerings“ ein unerschöpfliches Potential bei der Generierung von Proteinen oder Proteinmodulen mit bestimmten, veränderten Eigenschaften und Funktionen. So können alle natürlichen Proteinbausteine bzw. deren zugrundeliegende Nukleinsäuresequenz als Ausgangspunkt benutzt und durch bestimmte Strategien der DNA-Manipulation gezielt, einfach und schnell verändert werden (Pace et al., 2014). Jedoch hat diese direkte und komplexe Beziehung auch einige Nachteile für das ,,Protein Engineering“, vor allem wenn man sich des fragilen Charakters von Proteinen bewusst ist (siehe Kapitel 4.1.1). Artifizielle Veränderungen der Aminosäuresequenz können sich natürlich nicht nur positiv auf die Eigenschaften eines Zielproteins auswirken. Durch die Mutation von nativen Aminosäuren werden immer Wechselwirkungen beeinflusst und verändert, die je nach Position und Beteiligung am Interaktionsnetzwerk einen mehr oder weniger großen Beitrag zur Stabilität, Struktur und Funktion von Proteinen leisten (Pace et al., 1996). In dieser Arbeit konnte durch einige wenige Aminosäureaustausche die Rekonstitutionseffizienz zwischen Catcher und Tag verbessert werden. Einige Mutationen dagegen beeinflussten die kovalente Reaktion zwischen Catcher und Tag kaum und die meisten mutierten Varianten hatten negative Auswirkungen auf die Reaktivität im Vergleich zu den Wildtyp-Systemen. Somit wird deutlich, dass der Austausch einzelner Aminosäuren schon massive, negative Auswirkungen auf die native Struktur und Funktion von Proteinen haben kann. Das größte Problem ist jedoch, dass der Effekt von Mutationen, selbst wenn nur eine einzelne Aminosäure betroffen ist, oft nicht genau vorhergesagt werden kann, was zu unerwarteten Fehlfaltungen und partiellem oder komplettem Funktionsverlust von Proteinen führen kann. Besonders erschwert wird die Vorhersage von Mutationseffekten auf die Faltung und Stabilität von Proteinen dadurch, dass Mutationen aufgrund der zahlreichen Wechselwirkungen, durch die sich die Aminosäuren gegenseitig beeinflussen, oft Strukturveränderungen in weiter entfernten Proteinbereichen verursachen und nicht nur am tatsächlichen Mutationsort (Pace et al., 1996). Aus diesem Grund führt das ,,Protein Engineering“, auch trotz vorhandener Informationen über die Sequenz und Struktur von Proteinen, oft nur durch zahlreiche experimentelle Analysen nach dem Prinzip von Versuch und Irrtum zum Erfolg. Genauso wie einzelne destabilisierende Aminosäuremutationen bereits zu einem völligen Funktionsverlust des Zielproteins führen können, trägt die Veränderung von individuellen Aminosäuren in anderen Fällen häufig kaum zu einer Stabilisierung des Proteins oder der gezielten Veränderung von bestimmten Proteineigenschaften bei (Goldenzweig et al., 2016). Das dürfte vor allem daran liegen, dass Proteinstrukturen aufgrund der Konservierung

bestimmter, funktionsgebender Faltungsbereiche doch einige Sequenzmutationen tolerieren (Borthakur et al., 2011, Tiana et al., 1998, Mirny und Shakhnovich, 1999). Aus diesem Grund müssen vor allem große Proteine durch die Kombination mehrerer Mutationen stabilisiert/ verändert werden, was wiederum schwer vorhersagbar und einzuschätzen ist (Goldenzweig et al., 2016). So muss die Kombination einzelner stabilisierender Mutationen nicht zwangsläufig zu einem noch stabileren Protein führen und erfordert daher individuelle Analysen.

4.1.3 Heterologe Proteinexpression als erste kritische und aufschlussreiche Hürde des ,,Protein Engineerings“

Ein weiteres Problem, das oft mit einer geringen Proteinstabilität des nativen Zustands und der direkten Sequenz-Struktur-Funktionsbeziehung zusammenhängt, ist, dass rekombinante Proteine oft Probleme bei der heterologen Expression/ Überexpression und der Löslichkeit aufweisen. Viele Proteine sind falsch gefaltet und/oder aggregieren, wenn sie in einem fremden Wirt überexprimiert werden (Goldenzweig et al., 2016, Baneyx und Mujacic, 2004). Somit stellt die Herstellung von stabilen, nativ gefalteten Proteinen die erste kritische Hürde dar, die bei der Arbeit mit Proteinen überwunden werden muss, denn ohne die korrekte Proteinfaltung büßen die wertvollen Proteine ihre Funktion ein und sind komplett unbrauchbar (siehe 4.1.2). Stabilisierende Mutationen von Zielproteinen im Bereich des ,,Protein Engineerings“ zur Verbesserung der heterologen Produktion von rekombinanten Proteinen, vor allem für biotechnologische und industrielle Anwendungen, sind deshalb von sehr großem Interesse. In dieser Arbeit konnten zwei der vier ausgewählten Kandidatensysteme, 3phs und 3kptN, nicht direkt auf ihre Fähigkeit zur kovalenten Split-Domänen Rekonstitution getestet werden, da die heterologe Expression der Catcher-Komponenten, 3phsCatcher und 3kptNCatcher, in *E. coli* Zellen nicht funktionierte (siehe 3.1.3). Die Expression der mCherry-Catcher Fusionsproteine war dann zwar erfolgreich, jedoch konnte in Interaktionsanalysen mit den jeweiligen Tag-Peptiden keine Ausbildung eines kovalenten Komplexes beobachtet werden. Da es einen Zusammenhang zwischen der Proteinstabilität und der heterologen Expression oder Überexpression gibt, ist es sehr wahrscheinlich, dass die Proteine 3phsCatcher und 3kptNCatcher besonders instabil bzw. falsch gefaltet sind (Pröschel et al., 2017). Es ist sogar fraglich, ob die Catcher-Proteine innerhalb des Fusionsproteins korrekt, nativ gefaltet und aktiv sind. Jedenfalls kam es unter gewählten Bedingungen und mit den initialen Produkten der CnaB-Domänenspaltungen nicht zu einer kovalenten Reaktion zwischen mCherry-3phsCatcher und 3phsTag bzw. zwischen mCherry-3kptNCatcher und 3kptNTag. Das heißt aber nicht, dass eine Split-Domänen Rekonstitution nicht möglich ist, wenn man die CnaB-Domänenspaltung anders vornimmt,

sodass die Produkte eventuell stabiler gefaltet sind und sich auch rekombinant herstellen lassen. Möglich ist aber auch, dass die Domänen, auf denen das 3phs- und das 3kptN-System beruhen, schon intrinsisch instabil sind. Die heterologe Expression der Ursprungsdomänen könnte hier nähere Einblicke liefern. In der Literatur sind einige Domänen von Pilinen bekannt, die sehr flexibel (hohe B-Faktor Werte), instabil oder besonders sensitiv gegenüber proteolytischem Abbau sind (Kang et al., 2014). Meistens handelt es sich dabei um die N-terminalen Domänen von Hauptpilinen, wie beispielsweise der CNA1-Domäne des BcpA Hauptpilins von *B. cereus* oder der D1-Domäne des RrgB Hauptpilins von *S. pneumoniae* (Krishnan, 2015, Budzik et al., 2009, Kang et al., 2014, Paterson und Baker, 2011). Folglich sind nicht alle Ig-ähnlichen Domänen gleichermaßen geeignet für eine Spaltung in einen Catcher- und einen Tag-Teil. Das Hauptkriterium ist neben der Stabilität der Ursprungsdomäne, die auch herausgelöst aus der Pilin-Struktur bestehen bleiben muss, natürlich die Stabilität der Catcher- und Tag-Produkte der Domänenspaltung. Durch die rekombinante Proteinherstellung können somit erste Informationen über die Domänenstabilität und damit über das Potential dieser Domänen, als Basis für die Entwicklung von Catcher-Tag Split-Isopeptid Systemen zu fungieren, gewonnen werden. Beide Systeme, 4oq1 und 3kptC, deren Catcher- und Tag-Reaktionspartner erfolgreich in *E. coli* exprimiert wurden, zeigten anschließend auch die Fähigkeit zur kovalenten Domänen-Rekonstitution. Das spricht dafür, dass die Catcher-Proteine, 4oq1Catcher und 3kptCCatcher, auch herausgelöst aus ihren Ursprungsdomänen größtenteils stabil und nativ gefaltet vorliegen (hoher Anteil an β-Faltblattstrukturen), sodass die strukturellen Grundvoraussetzungen für die Rekonstitution der reaktiven Triade und der Ausbildung einer Isopeptidbindung gegeben sind. CD (*circular dichroism)*-Analysen, durchgeführt am Lehrstuhl für Biotechnik (Dr. Benedikt Schmid) zur Ermittlung der Sekundärstrukturanteile, bestätigten diese Hypothese (Pröschel et al., 2017). Jedoch bestimmt nicht die Faltung oder Stabilität der Gesamtstruktur der Proteinfragmente alleine, ob sich eine Isopeptidbindung zwischen Catcher und Tag ausbildet. Ein wichtiges Kriterium ist zudem, ob die Reaktionspartner in einer reaktionskompatiblen Konformation (z.B. richtige Orientierung und räumliche Nähe der reaktiven Reste, siehe Einleitung) vorliegen, die schließlich die Ausbildung einer Isopeptidbindung erst ermöglicht (siehe Kapitel 4.3.4 der Diskussion).

4.2 Verschiedene Strategien des ,,Protein Engineerings“

,,Protein Engineering“ ist ein Teilgebiet der Biotechnologie/ Synthetischen Biologie und beschreibt die Veränderung und Neukonstruktion von Proteinen, um bestimmte neue anwendungsrelevante Eigenschaften zu generieren oder bereits existierende Proteineigenschaften je nach Anwendungsziel zu optimieren (Biotechnologie Buch, Walsh, 2007, Kazlauskas und Bornscheuer, 2009). Bei den Veränderungen kann es sich sowohl um Mutationen einzelner oder mehrerer unabhängiger Aminosäuren (Substitutionen, Deletionen, Insertionen) handeln oder kann ganze Proteindomänen oder nur post-translationale Modifikationen betreffen (Walsh, 2007). Prinzipiell gibt es zwei Herangehensweisen beim ,,Protein Engineering“. Das rationale ,,Design“ von Proteinen und die gerichtete Evolution. Bei den neuen Catcher-Tag Systemen, dem 4oq1- und dem 3kptC-System, sollte es in dieser Arbeit konkret um die Optimierung der Split-Domänen Rekonstitutionseffizienz gehen. Die Ausbildung von spontanen Isopeptidbindungen erscheint zwar einfach (Schwarz-Linek und Banfield, 2014: ,,all that is required is three appropriately positioned amino acids in a hydrophobic environment“), aber trotzdem müssen bestimmte Voraussetzungen, die vor allem die Proteinstruktur betreffen (,,conserved structural context“, Kang und Baker, 2011), gegeben sein. Folglich ist das ,,Protein Engineering“ dieser Split-Isopeptid Systeme, bei denen die Proteinstruktur über alles entscheidet, alles andere als trivial (siehe Kapitel 4.1).

4.2.1 Rationaler Ansatz – gezielte Mutagenese

Beim rationalen ,,Protein Engineering“ werden üblicherweise bereits existierende Proteine mit bekannter Sequenz und aufgeklärter Kristallstruktur unter Zuhilfenahme von geeigneten Computerprogrammen modelliert und anschließend gezielt über ortsspezifische Mutagenese mittels PCR-Methoden und unter Verwendung von spezifischen Mutagenese-Primern verändert. Bestimmte Aminosäuren werden unter Berücksichtigung ihrer spezifischen Wechselwirkungen (z.B. Catcher-Tag Interaktionen) und ihrem Beitrag zur zu optimierenden Proteinfunktion, hier der Ausbildung einer Isopeptidbindung, gezielt mutiert.

4.2.2 Evolutionärer Ansatz – zufällige Mutagenese & *in vitro* Evolution

Bestehende Proteine oder Proteinmodule müssen sich im Laufe der Evolution ständig an spezifische, neue Funktionen durch veränderte Proteineigenschaften anpassen. Dieser evolutionäre Vorgang ist natürlicherweise sehr erfolgreich und kann als Grundlage für die Konstruktion neuer Proteine beim ,,Protein Engineering“ dienen. Die evolutionäre Herangehensweise des ,,Protein Engineerings“ beginnt bei der zufälligen Mutagenese (z.B. *error-prone* PCR Methode) zur Erzeugung von möglichst großen und diversen Bibliotheken

an Proteinvarianten, die im Anschluss anhand bestimmter Eigenschaften in mehreren Runden mit steigender Stringenz selektioniert werden (,,Biopanning“, *in vitro* Evolution). Zur Verbesserung von Protein-Protein Interaktionen werden sehr häufig Phage Display-Techniken angewendet, wo die Zielproteine als Hüllproteinfusionen auf der Oberfläche von Bakteriophagen präsentiert und gegen immobilisierte Interaktionspartner getestet werden (Zakeri et al., 2012, Bazan et al., 2012, Hoffmann et al., 2010, Keeble und Banerjee et al., 2017).

4.2.3 Vergleich der verschiedenen Ansätze des ,,Protein Engineerings“ im Hinblick auf die Verbesserung der Catcher-Tag Rekonstitutions-effizienz

Beide Ansätze des ,,Protein Engineerings” haben ihre Vor- und Nachteile, die hier kurz umrissen werden sollen. Bei der rationalen Vorgehensweise sind sowohl ein fundiertes Wissen über Proteine, deren Aufbau, Faltung und Struktur als auch geeignete *in silico* Werkzeuge nötig. Native Aminosäuren werden zwar gezielt aufgrund von bestimmten Strategien mutiert, jedoch kann die Auswirkung auf die gesamte Proteinstruktur niemals exakt vorhergesagt werden, sodass sogar theoretisch stabilisierende, energetisch günstigere Mutationen, z.B. berechnet über bestimmte Algorithmen, in der Praxis das Protein destabilisieren und funktionell ruinieren (siehe 4.1.2). Die rationale Herangehensweise erscheint aber schneller und einfacher als der evolutionäre Ansatz des ,,Protein Engineerings“, wo zuerst durch aufwändige Klonierungsarbeiten ganze, möglichst große, Bibliotheken an zufälligen Protein- oder Peptidvarianten generiert werden müssen. Allerdings erfolgt die anschließende Durchmusterung der einzelnen Kandidaten über *in vitro* Evolution und braucht nicht zwangsläufig eine Kristallstruktur oder ähnliche Informationen über die Zielproteine. Jedoch können bei *in vitro* Evolutionsstrategien, wie Phage Display, Proteinvarianten nur aufgrund von Bindeeigenschaften selektioniert werden. Über evolutionären Selektionsdruck ausgewählte Varianten erscheinen zudem vielversprechender im Hinblick auf die Verbesserung bestimmter Proteineigenschaften, als durch rationale Überlegungen und eventuelle Computersimulationen und -analysen generierte Proteinvarianten, die ihre *in vitro* Testphase noch vor sich haben. Aus diesem Grund wäre z.B. ein Phage Display-Experiment zur Verbesserung der Bindeeigenschaften zwischen Catcher-Protein und Tag-Peptid sehr sinnvoll (Keeble und Banerjee et al., 2017). Da die Reaktionsrate selbst bei dem effizientesten Catcher-Tag System, dem Spy-System, weit unter dem Diffusionslimit liegt (Zakeri et al., 2012, Veggiani et al., 2014), könnte insbesondere die initiale Phase der Reaktion, die Assoziation und Komplexbildung von Catcher und Tag (*on-rate*), von einer *in vitro* Evolutionsstrategie zur Selektion der stärksten

Protein-Peptid Interaktionen profitieren (Zakeri et al., 2012). Durch Phage Display-Techniken, die kovalente Bindungen selektionieren, konnte die Ausbildung der spontanen Isopeptidbindung zwischen SpyCatcher und SpyTag bereits beschleunigt werden (Keeble und Banerjee et al., 2017). Kristallographische Strukturaufklärungen des SpyCatcher-SpyTag Komplexes haben gezeigt, dass sich zuerst ein hochaffiner, nicht-kovalenter Komplex zwischen SpyCatcher und SpyTag bildet (Komplexbildung), der sich vor allem über Wasserstoffbrücken stabilisiert, wenn sich die Reaktionspartner finden (Findungsphase) (Li et al., 2014). Anschließend reagieren SpyCatcher und SpyTag sehr schnell, aufgrund einer gut zugänglichen SpyTag-Bindestelle und einer reaktionskompatiblen Konformation des SpyCatcher-Proteins kovalent miteinander, indem sich eine Isopeptidbindung ausbildet (Reaktion siehe 1.2.2.2) (Li et al., 2014). Da es wahrscheinlich ist, dass alle Catcher-Tag Rekonstitutionen, die über die Bildung von Isopeptidbindungen laufen, diesem dreistufigen Prozess (Findungsphase, Komplexbildung, Reaktion) folgen, stehen die Chancen gut, dass auch die neu identifizierten Split-Isopeptid Systeme über Methoden wie Phage Display, wo verbesserte Bindeeigenschaften selektioniert werden, sowohl in ihrer nicht-kovalenten Assoziationsphase als auch in der tatsächlichen, kovalenten Reaktionsphase deutlich verbessert werden können. Zudem könnte über den Einsatz von Kompetitoren die Stringenz der Selektion erhöht werden, sodass man nicht nur effizientere, sondern auch spezifischere Protein-Peptid Reaktionspartner identifiziert. Ideal zur Optimierung der Rekonstitutionseffizienz der neu identifizierten Catcher-Tag Systeme wäre deshalb eine Kombination aus rationaler Mutagenese und Phage Display.

4.3 Rationale Strategien zur Optimierung von kovalenten Catcher-Tag Systemen

Zur Optimierung der Rekonstitutionseffizienz der neuen Catcher-Tag Systeme, 4oq1 und 3kptC, wurden in dieser Arbeit unterschiedliche, rationale Strategien verfolgt. Diese Strategien bzw. deren Auswirkungen auf die kovalente Reaktion zwischen den jeweiligen Catcher- und Tag-Reaktionspartnern sollen nun diskutiert und mit den angewendeten Strategien, die zu den effizienten, etablierten Systemen geführt haben, verglichen werden (Zakeri und Howarth, 2010, Zakeri et al., 2012, Veggiani et al., 2016, Tan et al., 2016). Dabei sollen auch mögliche Ursachen der unterschiedlichen Reaktivitäten der Systeme besprochen werden.

4.3.1 Ermittlung optimaler Reaktionsbedingungen bereits als erster Optimierungsschritt der neu identifizierten Catcher-Tag Systeme

Wie auch die etablierten Catcher-Tag Systeme (Isopeptag/Pilin-C, das Spy- und das Snoop-System: Zakeri und Howarth, 2010, Zakeri et al., 2012, Veggiani et al., 2016) waren die neu identifizierten Systeme, 4oq1 und 3kptC, weitgehend unabhängig von Reaktionsbedingungen wie der Temperatur und dem pH-Wert. Sowohl das 4oq1- als auch das 3kptC-System funktionierten bei allen getesteten Bedingungen, sodass es nach Mischen von Catcher- und Tag-Komponenten zur Ausbildung von kovalenten Komplexen kam. Jedoch wurden physiologische Bedingungen deutlich bevorzugt. Dass die neuen Systeme, 4oq1Catcher-4oq1Tag und 3kptCCatcher-3kptCTag, vielleicht etwas sensitiver gegenüber bestimmten Reaktionsbedingungen reagieren als die etablierten Catcher-Tag Systeme, dürfte vor allem daran liegen, dass es sich bei den etablierten Systemen, vor allem bei Spy- und SnoopCatcher, um bereits genetisch optimierte, intrinsisch stabile Proteine handelt, die optimale strukturelle Voraussetzungen für die Rekonstitution der reaktiven Triade und der Ausbildung einer Isopeptidbindung besitzen (Li et al., 2014). Die neu entdeckten Systeme und vor allem die großen Catcher-Proteine, 4oq1Catcher und 3kptCCatcher, liegen unoptimiert und unstabilisiert vor und sind deshalb eventuell noch anfälliger oder abhängiger von äußeren Faktoren, die neben den reinen physikochemischen Eigenschaften auch die Struktur, Stabilität und Aktivität von Proteinen beeinflussen können (Alexov, 2004, Talley und Alexov, 2011, Cao et al., 2017). Während beim Spy- und beim Snoop-System erhöhte Rekonstitutionseffizienzen vor allem durch genetische Manipulationen erreicht wurden (Zakeri et al., 2012, Veggiani et al., 2016), erbrachte die biochemische Charakterisierung (siehe 3.2) der neuen Catcher-Tag Systeme, 4oq1 und 3kptC, dagegen bereits eine deutliche Optimierung der Reaktivität. Das macht deutlich, dass für die Effizienz einer chemischen Reaktion, hier die Ausbildung einer Isopeptidbindung, immer auch die äußeren Faktoren, wie pH-Wert und Temperatur eine Rolle spielen, vor allem wenn die Proteine an sich nicht besonders stabil und optimiert sind. Besonders die Erhöhung des Catcher-Anteils im Reaktionsansatz konnte die Rekonstitutionseffizienz der neu identifizierten Systeme deutlich verbessern. Das dürfte vor allem daran liegen, dass die größeren Catcher-Proteinfragmente, 4oq1Catcher und 3kptCCatcher, wahrscheinlich nicht komplett stabil sind und eventuell partiell ungefaltet oder in einer weniger reaktionskompatiblen Konformation vorliegen. Indizien, die für eine partiell ungefaltete Struktur bzw. eine gewisse Instabilität der Catcher-Proteine sprechen, lieferten die Ergebnisse der heterologen Überexpression (siehe 3.1.3). So lieferte die heterologe Expression in *E. coli* Zellen immer relativ geringe Mengen an rekombinanten Catcher-Proteinen, vor allem im Vergleich mit den Tag-MBP Konstrukten. Folglich verlief die anschließende Aufreinigung der Catcher-Proteine über Ni-NTA auch nicht

optimal, da immer Kontaminationen mit unspezifischen Proteinen detektiert wurden, was zudem die Bestimmung der Proteinkonzentration und somit die Ermittlung der Molarität erschwerte. Durch die simple Erhöhung der Anzahl an Catchermolekülen im Reaktionsansatz und dadurch, dass die Rekonstitutionsrate direkt proportional zur Konzentration der Edukte ist (Zakeri und Howarth, 2010, Zakeri et al., 2012), wird die Wahrscheinlichkeit einer erfolgreichen Reaktion mit entsprechenden Tag-Peptiden erhöht und somit die Reaktionseffizienz deutlich verbessert. Auch die Zugabe von proteinstabilisierenden, kosmotropen Additiven wie Glycerin (Russo, 2008) und dem chemischen Chaperon TMAO (Yancey, 2004, Fierer et al., 2014, Veggiani et al., 2016) konnte die kovalente Reaktion zwischen Catcher und Tag für die beiden neuen Split-Isopeptid Systeme deutlich verbessern. Die positive Wirkung von TMAO auf die Rekonstitutionseffizienz wurde ebenfalls für das Spy- und das Snoop-System beschrieben (Fierer et al., 2014, Veggiani et al., 2016). Es wird deutlich, dass die Proteinstruktur der Catcher- bzw. Tag-Proteine und die Orientierung und Position der Aminosäuren der reaktiven Triade bei der Ausbildung von Isopeptidbindungen eine weitaus größere Rolle spielen als die Reaktionsbedingungen (Schwarz-Linek und Banfield, 2014, siehe Einleitung). So führen selbst optimale Reaktionsbedingungen nicht zu der Ausbildung einer Isopeptidbindung, wenn die passende Proteinkonformation rund um die reaktive Triade nicht gegeben ist. Die Effizienz der Ausbildung von Isopeptidbindungen kann jedoch, wenn die essentiellen strukturellen Grundvoraussetzungen (siehe Einleitung 1.2.2.2) gegeben sind, durch optimale Reaktionsbedingungen positiv beeinflusst werden. Da die Catcher-Tag Reaktionen konzentrationsabhängig sind (siehe auch Zakeri und Howarth, 2010, Zakeri et al., 2012), kann die Effizienz durch Erhöhung der Konzentration der Edukte erhöht werden. Das für die Reaktion optimale Verhältnis zwischen Catcher und Tag hängt dabei auch von der Faltung und Stabilität, widergespiegelt durch die Ausbeute an rekombinant hergestelltem Protein, der individuellen Komponenten ab. Reaktionsadditive, die die Proteinstrukturen stabilisieren, hatten stets einen positiven Effekt auf die Reaktivität. Ob die Zugabe von Glycerin bzw. TMAO die Reaktionsrate tatsächlich aufgrund einer stabilisierten Proteinstruktur verbesserte oder ob die Catcher-Tag Reaktion auch aufgrund der veränderten Viskosität des Reaktionspuffers und damit einer verlangsamten Diffusionsrate effektiver wurde, ist nicht eindeutig geklärt.

4.3.2 Überprüfung der initialen Produkte der CnaB-Domänenspaltung durch N- und C-terminale Proteinvarianten

Aufgrund der Tatsache, dass die langsame Split-Domänen Rekonstitution des 4oq1- und des 3kptC-Systems, seinen Ursprung in der initialen Spaltung der zugrundeliegenden CnaB-Domänen haben könnte, wurden zuerst verschiedene Spaltungsvarianten getestet, indem die Catcher- und Tag-Sequenzen N- und C-terminal verkürzt oder um die nativen Sequenzen der jeweiligen Ausgangsdomäne verlängert wurden. Ausschließlich N-terminale Verlängerungen des 4oq1Tags (4oq1Tag-L und 4oq1Tag-HQL und vor allem 4oq1Tag+2nd β-Strang) konnten die kovalente Split-Domänen Rekonstitution verbessern, was wahrscheinlich durch die Ausweitung des Netzwerkes an möglichen, stabilisierenden Interaktionen (z.B. Wasserstoffbrücken) zwischen 4oq1Catcher und 4oq1Tag hervorgerufen wird. Vor allem die Integration eines zweiten β-Stranges in die 4oq1Tag-Sequenz ließ die kovalente Reaktion zwischen Catcher und Tag effektiver werden. Weder 4oq1Catcher- noch 3kptCTag-Varianten (Verlängerungen, Verkürzungen oder Substitutionen) konnten zu einer wirklichen Steigerung der Reaktivität führen. Im Gegenteil, die C-terminale Verkürzung des 4oq1Catchers um die Reste Glutamin-239 und Leucin-240 verschlechterte die kovalente Reaktion zwischen 4oq1Catcher und 4oq1Tag deutlich, genauso wie die C-terminale Deletion des 4oq1Tags um die Reste Leucin-255 bis Asparagin-259 (siehe Kapitel 3.3.1). Diese Mutanten lassen darauf schließen, dass die deletierten Aminosäuren bzw. einzelne dieser deletierten Aminosäuren wichtig für die Funktion zur Ausbildung eines kovalenten Catcher-Tag Komplexes sind. So stellen diese Aminosäuren wahrscheinlich wichtige Interaktionspunkte zwischen Catcher und Tag dar, die vor allem die initiale, nicht-kovalente Komplexbildung betreffen. Allgemein lässt sich sagen, dass sowohl das 4oq1- als auch das 3kptC-System ziemlich robust gegenüber kleineren Veränderungen der Sequenzlänge der Catcher-Tag Reaktionspartner waren (Pröschel et al., 2017). Ein deutlicher Effekt war nur bei der drastischen Erweiterung des 4oq1Tags um einen zweiten β-Strang zu sehen, was dafür spricht, dass mehrere stabilisierende Wechselwirkungen nicht primär funktionsbestimmender Aminosäuren zwischen Catcher und Tag geschlossen werden müssen, um zu einer detektierbaren Steigerung der Rekonstitutionseffizienz zu führen (Quantität statt Qualität). Handelt es sich aber um funktionsdefinierende Aminosäuren, kann die Deletion oder Insertion einzelner essentieller Aminosäuren dagegen sofort zu einer Verschlechterung, bis hin zu einem kompletten Funktionsverlust, oder einer Verbesserung der Reaktion führen (Qualität statt Quantität). Sobald beispielsweise eine der Aminosäuren der Isopeptid-Triade mutiert ist, kann keine Isopeptidbindung mehr ausgebildet werden. Die Verlängerung der Catcher- bzw. Tag-Fragmente wird außerdem generell besser toleriert als deren Verkürzung. Besonders vorsichtig muss bei der Verkürzung der sehr kurzen Tag-Peptide vorgegangen

werden, da es aufgrund der geringen Größe wahrscheinlich ist, dass die meisten Aminosäuren auch wichtige Kontaktpunkte zum Catcher-Protein herstellen. Tag-Varianten haben im Gegensatz zu Catcher-Varianten ebenfalls beim Isopeptag/Pilin-C System und beim SpyCatcher-SpyTag System zu dramatischen Veränderungen der Reaktionseffizienz geführt (Zakeri et al., 2012, Zakeri und Howarth, 2010). So führte die C-terminale Verlängerung des SpyTag-Peptids um fünf Aminosäuren (YKPTK) der nativen FbaB Sequenz zu einer deutlichen Verbesserung der Rekonstitutionseffizienz (Zakeri et al., 2012), während sich die Reaktivität durch Verkürzung des Isopeptags um fünf Aminosäuren deutlich verschlechterte (Zakeri und Howarth, 2010). Somit scheint die Integration nicht essentieller Reste, die auch doppelt, d.h. also sowohl im Catcher- als auch im Tag-Fragment vorkommen können, im Hinblick auf die Reaktivität zwischen Catcher und Tag weniger kritisch. Bezüglich der Verwendung der Catcher- bzw. Tag-Fragmente als Fusionsproteine zur Modifikation von Zielproteinen sind jedoch minimale Größen von Vorteil, weshalb es schon wichtig wäre, nicht essentielle Reste zukünftig zu entfernen. Kristallographische Strukturaufklärungen der Catcher-Tag Komplexe könnten hier sowohl Aufschluss über tatsächliche Aminosäure-Interaktionspunkte zwischen Catcher und Tag liefern als auch unbeteiligte Aminosäuren, die verzichtbar für eine Catcher-Tag Reaktion sind, identifizieren. So konnten auch Li et al. (2014) zeigen, dass spezifische Aminosäurekontakte/ Wechselwirkungen zwischen SpyTag und SpyCatcher wichtig für eine schnelle Split-Domänen Rekonstitution sind (Li et al., 2014, Zakeri et al., 2012, Veggiani et al., 2014). So führte beispielsweise die Substitution der Aminosäure Isoleucin-34 durch Glutaminsäure im SpyCatcher zu einer deutlichen Verbesserung der Rekonstitutionseffizienz, da die Interaktion zwischen SpyCatcher und SpyTag durch die Ausbildung einer weiteren Wasserstoffbrücke gestärkt wurde (Zakeri et al., 2012, Li et al., 2014). Somit scheinen, wie auch allgemein in Protein-Protein Interaktionsoberflächen, sogenannte *hot-spot* Aminosäuren zu existieren, die den Hauptbeitrag zur Komplexbildung leisten (Yan et al., 2008). Durch die gezielte Stärkung dieser *hot-spot* Interaktionen könnte die Split-Domänen Rekonstitution, und vor allem die initiale Assoziationsphase von Catcher und Tag, optimiert werden.

4.3.3 Stabilisierung der Catcher-Proteine

Die mäßige Rekonstitutionseffizienz der neuen Catcher-Tag Systeme könnte natürlich auch an einer geringen Stabilität, insbesondere der größeren Catcher-Proteine liegen (siehe Kapitel 3.3.2), weshalb verschiedene potentiell stabilisierende Aminosäuresubstitutionen, basierend auf unterschiedlichen rationalen Ansätzen, getestet wurden. Die Erhöhung der intrinsischen Stabilität der Catcher-Proteine hat zumindest für das Spy- und das Snoop-System große Erfolge erzielt (Zakeri et al., 2012, Veggiani et al., 2016). Während beim

SpyCatcher vor allem die Entfernung exponierter hydrophober Aminosäuren an der Proteinoberfläche die Reaktionsrate verbesserte (Zakeri et al., 2012), erbrachten Aminosäuresubstitutionen zur Stabilisierung von Sekundärstrukturelementen im SnoopCatcher den gewünschten Effekt (Veggiani et al., 2016).

4.3.3.1 Richtiges Maß an Flexibilität und Starrheit in Proteinstrukturen

Der richtige Grad an Flexibilität und Starrheit innerhalb von Polypeptidketten ist wohl enorm kritisch beim ,,Protein Engineering“. Auf der einen Seite ist strukturelle Flexibilität enorm wichtig für die Faltung und Funktion von Proteinen (Potapov et al., 2004). So wird vermutet, dass sich flexible Strukturen durch positive Selektion aufgrund funktioneller Vorteile entwickelt haben (Borthakur et al., 2011). Auf der anderen Seite erlaubt eine gewisse Einschränkung der Rotationsfreiheit innerhalb von Polypeptidketten durch Herabsenken der hohen entropischen Kräfte des ungefalteten Zustands quasi erst die native Faltung von Proteinen und vermittelt strukturelle Stabilität (Pace et al., 1996, Choi et al., 2006). Besonders essentiell erscheint diese Balance, wenn es um bestimmte Interaktionen geht, die durch Positionseffekte, strukturelle Neuordnung oder Faltungen, hervorgerufen werden, wie der Prozess der Ausbildung von Isopeptidbindungen (siehe Einleitung). Flexibilität in Proteinstrukturen erlaubt eine gewisse intramolekulare Umstrukturierung, sodass sich Reste rearrangieren können und auch gezielt exponiert und dadurch in räumliche Nähe gebracht werden können, um spezifische Interaktionen und Bindungen zu ermöglichen und zu erleichtern (Krishnan, 2015, Kang et al., 2014, Borthakur et al., 2011: *coupled folding and binding*, Cao et al., 2017). Doch flexiblere Proteinstrukturen zeichnen sich oftmals auch durch eine ungenügende Stabilität und erhöhte Anfälligkeit gegenüber proteolytischem Abbau aus (Krishnan, 2015, Kang et al., 2014). Unter diesen Gesichtspunkten wurden verschiedene, potentiell stabilisierende Catcher-Varianten hergestellt.

4.3.3.1.1 Prolin- und Glycin-Substitutionen

Der rigide Pyrrolidinring von Prolinen reduziert die Rotationsfreiheit der Proteinhauptkette um die konformationsbestimmende Einfachbindung zwischen dem C_α und dem N-Atom (phi-Torsionswinkel) des ungefalteten Zustands und erhöht somit durch die Reduktion destabilisierender, entropischer Kräfte die Proteinstabilität (Choi et al., 2006). So wurde beschrieben, dass Prolin-Substitutionen verschiedener Aminosäuren in bestimmten Proteinbereichen die Proteinstruktur stabilisierten (Choi et al., 2006, Matthews et al., 1987, Watanabe et al., 1994), weshalb die möglicherweise instabilen Catcher-Proteine ebenfalls dieser Strategie unterzogen wurden. Die meisten in dieser Arbeit durchgeführten Prolin-Substitutionen sowohl des 4oq1- als auch des 3kptC-Systems führten zu keiner Verbesserung der kovalenten Reaktion zwischen zusammengehörigen Catcher- und Tag-

Komponenten (siehe 3.3.2.2.2). Im Gegenteil, sie verschlechterten die Reaktion. Zwei der insgesamt sechs getesteten Mutanten führten sogar zu einem kompletten Verlust der Bildung eines kovalenten Catcher-Tag Komplexes. Lediglich eine Prolin-Mutante des 3kptCCatchers (K425→P) zeigte eine bessere Rekonstitutionseffizienz im Vergleich zum Wildtyp. Die Starrheit, die Proline aufgrund ihrer Struktur innerhalb von Proteinkonformationen vermitteln (Choi et al., 2006, Morgan und Rubenstein, 2013), könnte in manchen Fällen dazu führen, dass reaktionsinkompatible Konformationen fixiert werden. Die reaktive Triade zur Ausbildung einer autokatalytischen Isopeptidbindung formt sich natürlicherweise während der Proteinfaltung, wenn die involvierten Aminosäuren im hydrophoben Proteinkern in geeigneter Position in räumliche Nähe gebracht werden (Kang und Baker, 2011, Kang et al., 2014, Schwarz-Linek und Banfield, 2014). Dieser Prozess der Isopeptid-Triaden Rekonstitution könnte durch die rigide Konformation, die weniger Rotationsfreiheit bei der Proteinfaltung zulässt, gestört sein, weshalb sich die Prolin-Mutagenese Strategie auch negativ auf die Reaktivität der Catcher-Tag Systeme auswirken könnte. Somit sind Prolin-Substitutionen zur Stabilisierung oder Fixierung von Catcher-Strukturen wahrscheinlich erst dann wirklich sinnvoll, wenn die reaktionskompatible Form vorherrscht. Hinzu kommt, dass insbesondere Proline auch aufgrund sterischer Hinderungen weitreichende Auswirkungen auf die gesamte Proteinstruktur haben und das komplette Aminosäure-Interaktionsnetzwerk sowie Sekundärstrukturelemente beeinflussen bzw. zerstören können (Choi et al., 2006, Morgan und Rubenstein, 2013). So kann der Einsatz von Prolinen sowohl stabilisierend als auch hochgradig destabilisierend für Proteine sein und der Effekt jeder Mutation muss individuell beurteilt werden und ist nicht generalisierbar (Choi et al., 2006). Glycin-Substitutionen haben gegenteilige Effekte, sodass Strukturbereiche durch Glycine flexibler werden und sich bestimmten Faltungen besser anpassen können. Im Fall der Verbesserung der Rekonstitutionseffizienz sollten Glycin-Substitutionen intramolekulare Umstrukturierungen und die Proteinfaltung, durch die die Rekonstitution der Isopeptid-Triade und die Ausbildung der Isopeptidbindung eingeleitet werden, erleichtern. Beide Glycin-Catcher-Mutanten führten zu einer kovalenten Produktbildung bei der Reaktion mit 4oq1Tag, jedoch konnte die Reaktivität im Vergleich zum 4oq1Catcher-4oq1Tag Wildtyp-System nicht verbessert werden. Das zeigt, dass die Erhöhung der Flexibilität von Proteinstrukturen durch Glycin-Substitutionen weniger dramatische Auswirkungen auf die Proteinfunktion hatte als die Erhöhung der Rigidität durch Prolin-Substitutionen, jedenfalls solange die Proteinstruktur noch stabil genug ist. Ein Indiz für die Proteinstabilität wäre wiederum die heterologe Überexpression (siehe Kapitel 4.1.3).

4.3.3.1.2 B-Faktor basierte Substitutionen

Die B-Faktor basierte Strategie wurde auch verfolgt, um die Proteinstabilität durch Ersetzen besonders flexibler Aminosäuren (hohe B-Faktor Werte) zu verbessern und dadurch die Rekonstitutionseffizienz zu erhöhen. Jedoch wurden flexible Aminosäuren nicht durch Proline substituiert, sondern durch andere, energetisch günstigere Aminosäuren (Rosetta Backrub). Eine der vier durchgeführten Mutationen erbrachte eine Steigerung der Reaktionseffizienz, während keine der Aminosäuresubstitutionen zu einem kompletten Verlust der Fähigkeit zur Ausbildung von kovalenten Catcher-Tag Komplexen führte. Dies spricht für einen weniger dramatischen Eingriff in die Proteinstruktur im Vergleich zu den Prolin-Substitutionen, die wohl eine sehr radikale Strategie des ,,Protein Engineerings“ darstellen (s.o.).

4.3.3.2 Einfluss des hydrophoben Effekts auf die Proteinstabilität

Durch die Entfernung hydrophober Aminosäuren, wie Phenylalanin, Isoleucin, Leucin, Methionin, Valin und Tryptophan, an der Proteinoberfläche, die energetisch eher ungünstig sind und somit destabilisierend auf die Gesamtstruktur wirken können (siehe Einleitung), sollte die Proteinstabilität der Catcher-Proteine erhöht werden. Diese Strategie war zum Beispiel bei der Optimierung des SpyCatcher-Proteins enorm erfolgreich (Zakeri et al., 2012, M69→Y). In dieser Arbeit führten lediglich zwei der insgesamt zwölf getesteten Mutationen zu einer leicht verbesserten kovalenten Reaktion zwischen Catcher und Tag. Beide Male handelte es sich um die Entfernung von großen, weit ins wässrige Milieu ragenden Phenylalaninen (F209→Y bei 4oq1Catcher und F471→Y bei 3kptCCatcher), die zu hydrophileren, polaren Tyrosinen mutiert wurden. Dadurch, dass Phenylalanin durch Tyrosin ersetzt wird, bleibt die aromatische Ringstruktur erhalten und es wird eine freie Hydroxylgruppe eingeführt, die eine gewisse Polarität erzeugt (Tyrosin als amphiphatische Aminosäure). Potentielle, stabilisierende Wasserstoffbrücken mit Wassermolekülen der Hydrathülle sind nun denkbar. Das Problem dieser Strategie ist jedoch, dass selbst energetisch ungünstige Aminosäuren, wie hydrophobe Aminosäuren an der Proteinoberfläche oder hydrophile, und vor allem geladene Aminosäuren im Proteininneren, bestimmte Funktionen erfüllen können, die essentiell für die Stabilität, Struktur und Funktion von Proteinen sind. Dies trifft aber vor allem auf verborgene hydrophile Reste zu. So kostet das Verbergen von hydrophilen Aminosäuren im hydrophoben Proteinkern zwar Energie, die aber durch andere, stabilisierende Wechselwirkungen dieser Reste, wie Wasserstoffbrücken oder elektrostatische Interaktionen, kompensiert und positiviert werden kann (Pace et al., 1996). Folglich leisten diese, auf den ersten Blick energetisch ungünstigen Reste, auch ihren stabilisierenden Beitrag zur Gesamtproteinkonformation. Im Grunde handelt es sich bei den

Aminosäuren der reaktiven Triade auch um energetisch ungünstige Aminosäuren (geladene bzw. polare Aminosäuren) im sonst hydrophoben Proteinkern der Ig-ähnlichen Domänen. Der energetische Vorteil bzw. der Vorteil für die Proteinstabilität, der durch die Ausbildung dieser intramolekularen Isopeptidbindung entsteht, ist jedoch enorm und überwiegt. Aus diesem Grund sind intramolekulare Isopeptidbindungen innerhalb von Ig-ähnlichen Domänen evolutionär konserviert und als Strategie zur Proteinstabilisierung vor allem bei gram-positiven Bakterien, denen die zelluläre Maschinerie zur Ausbildung von stabilisierenden Disulfidbrücken fehlt, sehr weit verbreitet (Zakeri et al., 2012, Kang und Baker, 2011, Veggiani et al., 2014, siehe Einleitung). Die Substitution hydrophober Aminosäuren an der Proteinoberfläche dagegen, scheint keine großen Auswirkungen auf die Proteinkonformation und Proteinstabilität zu haben, jedenfalls solange man keine besonders strukturbestimmenden Aminosäuren, wie z.B. Prolin, einsetzt oder besonders raumgreifende Aminosäuren, wie z.B. Phenylalanin und Tryptophan, betroffen sind. So führte keine der Aminosäuremutationen zu einem Funktionsverlust der Bildung von kovalenten Catcher-Tag Komplexen (siehe Kapitel 3.3.2.1). Trotzdem sollte beim Austausch von hydrophoben Aminosäuren berücksichtigt werden, dass dies auch einen Effekt auf die Flexibilität und Stabilität von Proteinen haben kann (Borthakur et al., 2011, Cao et al., 2017). So zeichnen sich intrinsisch flexible, unstrukturierte Proteine, wie z.B. die konstante Domäne des IgE Moleküls, unter anderem durch eine relativ niedrige Hydrophobizität aus (Borthakur et al., 2011). Hydrophobe und vor allem aromatische Aminosäuren an Proteinoberflächen haben aber oft eine andere biologische Funktion, die ihre energetisch gesehen eher ungünstige Lage erklärt und rechtfertigt. So sind diese Aminosäuren in Protein-Protein Interaktionsoberflächen (z.B. Antigen-Bindestellen in Antikörpern, Mian et al., 1991) überrepräsentiert, wo sie an der Interaktion und Bindung von anderen Molekülen/ Proteinen beteiligt sind (Mian et al., 1991, Yan et al., 2008). Dieser wichtige biologische Beitrag zur Vermittlung von spezifischen und affinen Protein-Protein Interaktionen (Yan et al., 2008) spielt jedoch beim ,,Protein Engineering“ zur Entwicklung von möglichst effizienten Split-Isopeptid Systemen keine Rolle, da die Catcher-Tag Systeme als individuelle Module, herausgelöst aus dem Pilus-Polymer, arbeiten sollen. Jedoch ist bekannt, dass Domänen-Domänen Interaktionen sowie die Pilus-Assemblierung über intermolekulare Isopeptidbindungen durchaus einen Einfluss auf die Ausbildung von intramolekularen Isopeptidbindungen haben können und sich diese Prozesse gegenseitig positiv beeinflussen (Hu et al., 2011, Budzik et al., 2009, Paterson und Baker, 2011, Kang et al., 2014, Krishnan, 2015, Shaik et al., 2014, El Mortaji et al., 2010).

4.3.3.3 Vergleichendes ,,Proteindesign“ - PROSS

Evolutionäre Prozesse führen natürlicherweise zur Optimierung von funktionsbestimmenden Proteineigenschaften wie der Stabilität oder der Faltungsgeschwindigkeit (Borthakur et al., 2011, Debès et al., 2013). Aus diesem Grund sind spezifische Aminosäuren, die essentielle Kontaktpunkte darstellen, innerhalb von Faltungsmotiven, wie beispielsweise der Immunglobulinfaltung, konserviert (Borthakur et al., 2011, Mirny und Shakhnovich, 1999). Somit stellt vergleichendes ,,Proteindesign“ auch eine wichtige Strategie innerhalb des rationalen ,,Protein Engineerings“ dar. Auf Sequenz- und Strukturvergleichen (3D-Strukturen) mit homologen Proteinen (*multiple sequence alignments*) basiert der PROSS-Server, der mittels vollautomatisierter Algorithmen stabile Proteinvarianten durch gezielte Aminosäuresubstitutionen eines Zielproteins vorhersagt (Goldenzweig et al., 2016). Beide PROSS-basierte Catcher-Mutanten waren in der Lage, kovalente Komplexe mit ihren jeweiligen Tag-Peptiden zu bilden. Während der 4oq1Catcher-PROSS von der Rekonstitutionseffizienz her ungefähr Wildtyp-Niveau erreichte, verbesserte der 3kptCCatcher-PROSS die Reaktion. Das spricht dafür, dass dieser PROSS-Server ein geeignetes Werkzeug und Hilfsmittel beim rationalen ,,Protein Engineering“ darstellt und auch in Zukunft verwendet werden kann. Dreidimensionale Strukturen sind, evolutionär betrachtet, höher konserviert als Sequenzen, was man z.B. gut an den Ig-ähnlichen Domänen sehen kann, die alle die konservierten, strukturellen Grundeigenschaften der Immunglobulinfaltung aufweisen (β-Sandwich Struktur mit hydrophobem Kern, siehe Einleitung 1.1.2.2), obwohl sie sehr variable Sequenzen besitzen können (Bork et al., 1994). Vergleichendes ,,Proteindesign“ profitiert daher von der Kombination aus Sequenz- und Strukturinformation eines Proteins bzw. mehrerer, homologer Proteine, die miteinander verglichen werden. So wurde z.B. auch das SdyCatcher-SdyTag System durch Sequenz- und Strukturvergleiche mit dem homologen Spy-System optimiert (Tan et al., 2016). Durch Sequenzvergleiche der zu stabilisierenden Zielproteine mit möglichst vielen homologen Proteinen, die durch divergente Evolution aus einem gemeinsamen Vorläufer entstanden sind, können sowohl aufschlussreiche Sequenzunterschiede (Angriffspunkte für Mutationen) als auch essentielle, konservierte Aminosäuren (eher Ausschluss von der Mutagenese) identifiziert werden (Borthakur et al., 2011). Die Strukturinformation liefert dabei weitere wichtige Informationen über das Vorkommen bestimmter, eventuell stabilisierender (energetisch günstiger) Aminosäuren in bestimmten Faltungen bzw. Strukturbereichen und zwar über die normale Aminosäure-Präferenz für bestimmte Sekundärstrukturelemente hinaus (Borthakur et al., 2011). Eine weitere Herangehensweise zur Stabilisierung von Proteinen könnte der Sequenz- und Strukturvergleich mit thermophilen gram-positiven Bakterien liefern (z.B. thermophile Stämme von Bacillus und Clostridien, optimales

Wachstum bei 45-75°C), deren Proteine besonders stabil und resistent sein müssen, um bei Hitze nicht zu entfalten/ denaturieren. Ideal für die Optimierung von Catcher-Tag Systemen wäre die Existenz von Ig-ähnlichen Domänen mit intramolekularen Isopeptidbindungen in Thermophilen, was einen direkten Vergleich ermöglichen würde. Einerseits ist diese Existenz wahrscheinlich, da Ig-ähnliche Domänen innerhalb von adhäsiven Zelloberflächenproteinen von gram-positiven Bakterien extrem weit verbreitet sind (Kang et al., 2007, Kang und Baker, 2011). Andererseits ist der Selektionsdruck zur Proteinstabilisierung bei diesen thermophilen Bakterien wahrscheinlich ein anderer, da deren Proteine sowieso intrinsisch stabiler sein müssen, um die immer vorherrschenden, extremen Temperaturen tolerieren zu können. Überleben und Wachstum bei den thermophilen Bakterien steht also gegen Virulenz und Pathogenität bei den in dieser Arbeit untersuchten gram-positiven Bakterien. Folglich ist die Strategie der Ausbildung von intramolekularen Isopeptidbindungen als lokale, stabilisierende Proteinquervernetzungen zur Erhöhung der Proteinstabilität eventuell nicht so sehr verbreitet oder evolutionär konserviert innerhalb thermophiler gram-positiver Bakterien, die zahlreiche andere sequenz- und strukturbezogene Strategien zur Erhöhung der Hitzeresistenz ihrer Proteine entwickelt haben. So besitzen Proteine thermophiler Bakterien eine besondere Art der Proteinfaltung, die sich durch eine allgemein dicht gepackte, wenig flexible Struktur auszeichnet, die unter anderem durch lokale Interaktionen, vor allem elektrostatischer Natur (Salzbrücken), hervorgerufen wird. Deletionen flexibler *loop*-Regionen oder häufiges Vorkommen von Prolinen in *loop*-Regionen zur Senkung der Entropie, sowie die Überrepräsentation verzweigter, hydrophober Aminosäuren, sind weitere Merkmale thermoresistenter Proteine (Kumar und Nussinov, 2001, Jaenicke, 1996, Razvi und Scholtz, 2006). Auch wenn thermophile gram-positive Bakterien eventuell nicht über solche Ig-ähnlichen Domänen mit intramolekularen Isopeptidbindungen verfügen sollten, könnte man trotzdem einige dieser natürlich evolvierten Strategien zur Proteinstabilisierung beim ,,Protein Engineering“ mitberücksichtigen und imitieren.

4.3.4 Unterschiedliche strukturelle Voraussetzungen

Der Grad an geometrischer und chemischer Komplementarität von Proteinoberflächen entscheidet über die Effizienz (Assoziationsgeschwindigkeit, Stärke und Dauer) und Spezifität einer Interaktion (Yan et al., 2008). Die beobachteten Effizienzunterschiede der kovalenten Catcher-Tag Reaktion verglichen mit den etablierten Systemen, Spy und Snoop, könnten deshalb auch auf strukturelle Unterschiede, wie unzugängliche Tag-Bindestellen in den Catcher-Proteinen, 4oq1Catcher und 3kptCCatcher, zurückzuführen sein. So ist z.B. das Isopeptag/Pilin-C System auch deutlich ineffizienter als das SpyCatcher-SpyTag System (Zakeri, 2015). Es wird angenommen, dass die Assoziation des Isopeptags mit dem Pilin-C

in einer reaktionskompatiblen Konformation wahrscheinlich der limitierende Reaktionsschritt ist (Zakeri und Howarth, 2010). Die Tag-Bindestelle des SpyCatcher-Proteins hingegen, scheint besonders gut exponiert zu sein, sodass es erst zu einer schnellen nicht-kovalenten Komplexbildung von SpyCatcher und SpyTag kommt, die fast unmittelbar zur Ausbildung einer Isopeptidbindung führt (Li et al., 2014). Der Komplex wird vor allem über Wasserstoffbrückenbindungen zwischen SpyCatcher und SpyTag stabilisiert (Li et al., 2014), was ebenfalls einen wichtigen Beitrag zu der schnellen, stabilen Assoziation der Proteinpartner leistet. Aus diesem Grund wurden die Wasserstoffbrückennetzwerke zwischen Catcher und Tag der neu identifizierten Systeme, 4oq1 und 3kptC, untersucht (Professur für Bioinformatik, Pröschel et al., 2017). Im Vergleich zu ihren Ausgangsdomänen zeigten die gespaltenen Domänen jedoch kaum Unterschiede in der Fähigkeit zur Ausbildung von Wasserstoffbrücken (Pröschel et al., 2017). Vielleicht verhält es sich mit dem 4oq1Catcher-4oq1Tag und dem 3kptCCatcher-3kptCTag System ähnlich wie mit dem Isopeptag/Pilin-C System (Zakeri und Howarth, 2010) und die Rekonstitutionseffizienz/ Reaktivität ist deshalb schlechter, da die Catcher- und Tag-Komponenten in einer weniger reaktionskompatiblen Konformation vorliegen. Eventuell müssen die neu identifizierten Catcher- und Tag-Komponenten erst erhöhte Energiebarrieren überwinden und sich über eine Art *induced fit*-Mechanismus (*fold upon binding* oder *induced folding*, Cao et al., 2017) in die richtige Faltung und Position bringen, während der SpyCatcher wahrscheinlich schon in einer vorgefalteten, reaktiven Struktur vorliegt, bevor der SpyTag bindet (Cao et al., 2017, Li et al., 2014). Die Kreuzreaktivität des SpyCatcher-Proteins sowohl mit 4oq1Tag als auch mit 3kptCTag unterstreicht die optimale, reaktionskompatible Konformation, die der SpyCatcher von sich aus annimmt, da eine Reaktion zur Ausbildung von Isopeptidbindungen sogar mit weniger passenden Tag-Peptiden erfolgt. Ob die initiale nicht-kovalente Assoziation von Catcher und Tag oder die tatsächliche Isopeptidbildung, oder beides, die kovalente Reaktion der neuen Split-Isopeptid Systeme limitieren, gilt es künftig zu klären. Analysen der dreidimensionalen Proteinstrukturen zeigten sowohl im 4oq1Catcher als auch im 3kptCCatcher jeweils einen Tyrosinrest, der direkt an der Eintrittsstelle des jeweiligen Tag-Peptids und in unmittelbarer Umgebung zur Isopeptid-Triade lokalisiert ist (siehe 4.1 und 4.2 A). Mittels ortsspezifischer Mutagenese wurde der Tyrosinrest-229 im 4oq1Catcher durch die kleinere, hydrophile Aminosäure Serin ersetzt, um zu testen, ob die relativ große Aminosäure mit ihrer aromatischen Ringstruktur eventuell das 4oq1Tag-Peptid bei der initialen Komplexbildung mit dem 4oq1Catcher-Protein stört. Anstatt die Rekonstitutionseffizienz zu verbessern, führte diese Tyrosin-Substitution zu einer deutlichen Verschlechterung der kovalenten Reaktion zwischen 4oq1Catcher und 4oq1Tag. Die Substitution des Tyrosinrestes-488 durch Prolin im 3kptCCatcher an ähnlicher Stelle führte sogar zu einem kompletten Funktionsverlust der Bildung von kovalenten 3kptCCatcher-

3kptCTag Komplexen. Diese Ergebnisse deuten ganz stark auf eine essentielle Rolle dieser Tyrosinreste für die Ausbildung einer Isopeptidbindung hin. Dass an ähnlicher Position sowohl beim 4oq1Catcher als auch beim 3kptCCatcher ein Tyrosin lokalisiert ist, spricht eventuell für eine funktionsbedingte Konservierung dieses Restes. Struktur- und Sequenzvergleiche der etablierten Catcher-Tag Systeme (Spy-, Snoop-, Isopeptag/Pilin Systeme) im Hinblick auf die Existenz von Tyrosinresten nahe der reaktiven Triade ergaben, dass jedes der Catcher-Tag Systeme über solch einen Tyrosinrest verfügt (siehe Abb.4.3), was die Hypothese einer essentiellen Funktion und Konservierung dieses Restes bekräftigt. Möglicherweise besitzen diese Tyrosinreste, direkt an der Tag-Eintrittsstelle und in unmittelbarer Nähe zur Isopeptidbindung, eine protektive Funktion, indem sie wie eine Art Deckel oder Schutzschild fungieren (siehe Abb.4.1 und 4.2), das den hydrophoben Proteinkern nach außen hin abschließt und den Eintritt von Wassermolekülen und eventuelle kompetitive Reaktionen verhindert. Die Isopeptid-Triade ist von außen betrachtet, also wenn man die Proteinoberfläche, die von dem wässrigen Milieu umgeben ist, anschaut, komplett im Inneren der Ig-ähnlichen Proteindomäne verborgen und nicht zu sehen (siehe Abb.4.1 und 4.2). Die Tyrosinreste spielen in jedem Fall eine wichtige Rolle, denn sie schließen eine Lücke in dieser Struktur, die sich genau am Ort der Isopeptidbindung (reaktive Triade) befindet (siehe Abb.4.1 und 4.2 B - D). So werden optimale Bedingungen, wie die hydrophobe Proteinumgebung (siehe Einleitung), für die Ausbildung von spontanen Isopeptidbindungen garantiert. Es ist wahrscheinlich, dass primär die große aromatische Ringstruktur des Tyrosins und nicht die freie Hydroxylgruppe, die die Fähigkeit zur Ausbildung von Wasserstoffbrücken besitzt, funktionsentscheidend ist. Die kleinere Aminosäure Serin, die auch über eine Hydroxylgruppe verfügt, konnte das Fehlen der Tyrosin-Seitenkette nicht funktionsgemäß kompensieren, sodass die Bildung von kovalenten Catcher-Tag Komplexen deutlich reduziert war (Tyrosin-229→Serin). Eine Substitution der Tyrosinreste durch Phenylalanin, das ebenfalls einen aromatischen Ring, allerdings ohne Hydroxylgruppe besitzt, könnte die Rolle der polaren Hydroxylgruppe bzw. der aromatischen, hydrophoben Ringstruktur bei der Ausbildung einer Isopeptidbindung zwischen Catcher und Tag noch genauer aufklären. Hinzu kommt, dass Tyrosin eine besondere Funktion bei der Vermittlung von Protein-Protein Interaktionen hat und aufgrund seines amphiphatischen Charakters oft in Interaktionsoberflächen wie Antigen-Bindestellen vorkommt (Fellouse et al., 2006, Yan et al., 2008, Mian et al., 1991). So scheinen neben der reaktiven Triade und den hydrophoben Aminosäuren, die diese Triade umgeben (Kang et al., 2007, Kang und Baker, 2011), noch weitere Aminosäuren innerhalb dieser Ig-ähnlichen Domänen (CnaB-Typ), die zu einer möglichst effizienten Ausbildung von intramolekularen Isopeptidbindungen beitragen, konserviert zu sein.

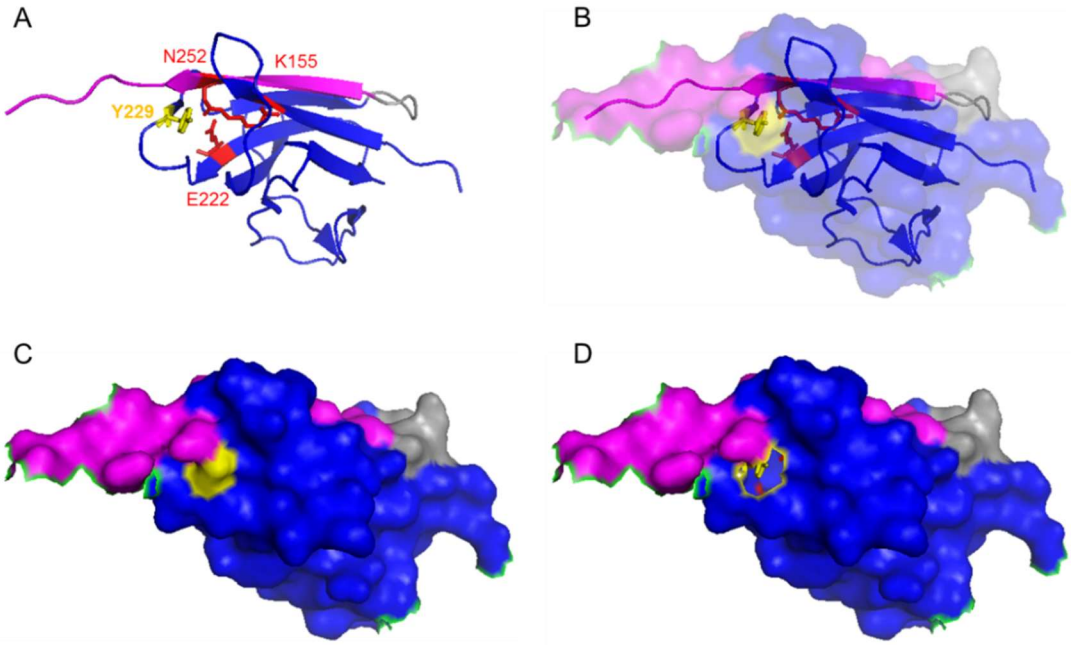


Abb.4.1: Analyse der Rolle des Tyrosins-229 bei der Ausbildung einer Isopeptidbindung zwischen 4oq1Catcher und 4oq1Tag. Der Tyrosinrest 229 des 4oq1Catchers ist in gelb dargestellt. Die Aminosäuren der Isopeptid-Triade sind in rot hervorgehoben (A). Abbildungen B - D zeigen die Proteinoberfläche. Die Molekülvisualisierungen wurden mittels PyMOL Open Source erstellt (blau: 4oq1Catcher, rosa: 4oq1Tag).

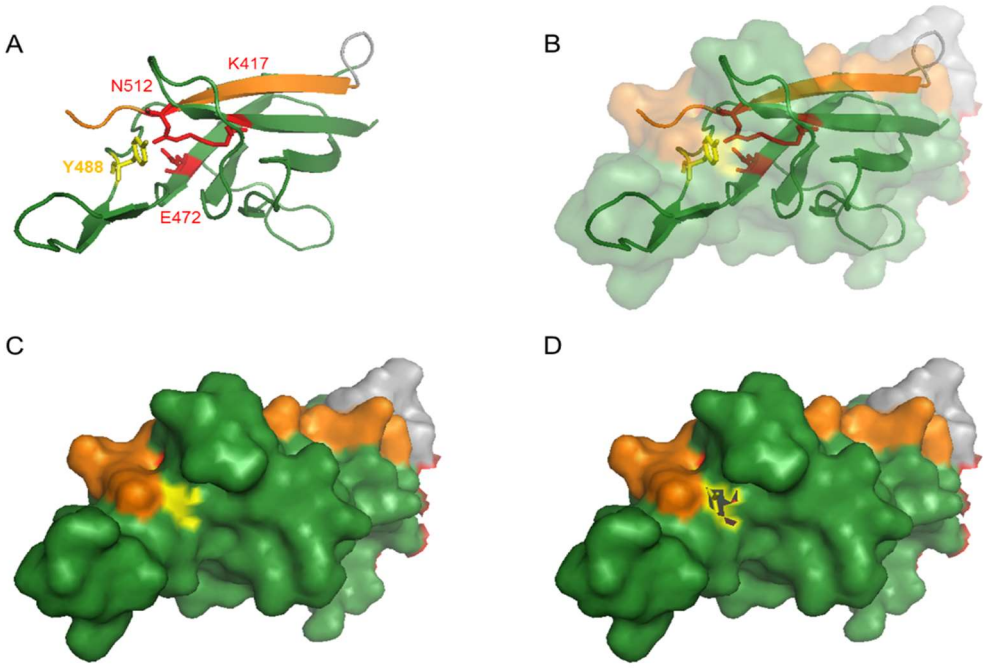


Abb.4.2: Analyse der Rolle des Tyrosins-488 bei der Ausbildung einer Isopeptidbindung zwischen 3kptCCatcher und 3kptCTag. Der Tyrosinrest 488 des 3kptCCatchers ist in gelb dargestellt. Die Aminosäuren der Isopeptid-Triade sind in rot hervorgehoben (A). Abbildungen B - D zeigen die Proteinoberfläche. Die Molekülvisualisierungen wurden mittels PyMOL Open Source erstellt (grün: 3kptCCatcher, orange: 3kptCTag).

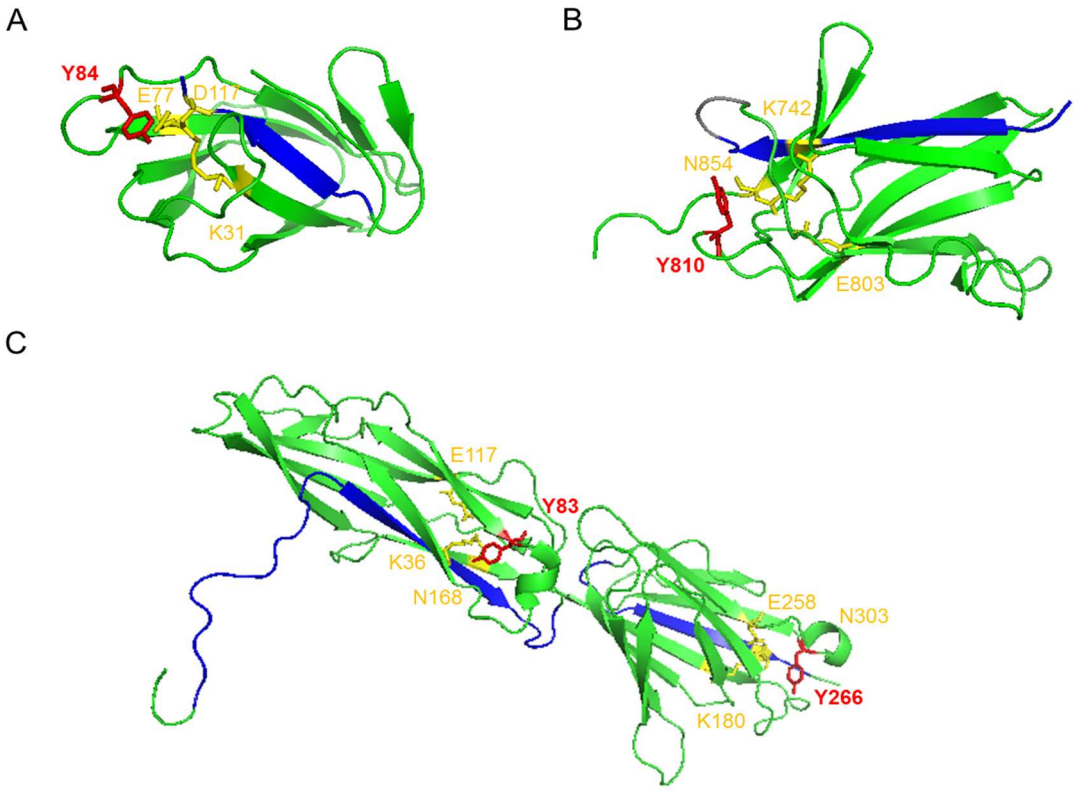


Abb.4.3: Existenz von Tyrosinresten nahe der Isopeptid-Triade in den etablierten Catcher-Tag Systemen. Tyrosinreste sind in rot dargestellt, während die jeweiligen Isopeptid-Triaden in gelb hervorgehoben sind. (A) Spy-System (Y84), (B) Snoop-System (Y810), (C) Spy0128 (N- und C-terminale Domäne: Y83 und Y266) Die Molekülvisualisierungen wurden mittels PyMOL Open Source erstellt (hellgrün: Catcher, blau: Tag).

4.3.5 Unterschiedliche chemische Voraussetzungen verschiedener Catcher-Tag Systeme

Möglich ist aber auch, dass sich neben den strukturellen Unterschieden (Faltung, Stabilität, reaktionskompatible Konformation, Zugänglichkeit der Tag-Bindestelle) die allgemeine Chemie der verschiedenen Catcher-Tag Systeme unterscheidet und deshalb die Reaktionen unterschiedlich schnell und effizient ablaufen. Indizien für eine unterschiedliche chemische Grundlage der Systeme für die Ausbildung von Isopeptidbindungen liefern beispielsweise die Ergebnisse bezüglich der pH-Optima (siehe 3.2.4.4), die sowohl die Proteinstabilität als auch die Aktivität beeinflussen (Alexov, 2004, Talley und Alexov, 2011, Cao et al., 2017). Die unterschiedlichen pH-Optima zwischen dem Spy-System (pH 5, Zakeri et al., 2012) und den neuen Systemen, 4oq1 und 3kptC (pH 7 - 9), können durch unterschiedliche Aminosäurezusammensetzungen und der Organisation leicht ionisierbarer Gruppen innerhalb der dreidimensionalen Proteinstruktur hervorgerufen werden (Alexov, 2004). Sieben der 20 Aminosäuren (Tyrosin: pK_S= 10,9, Cystein: pK_S= 8,3, Arginin: pK_S= 12,5, Lysin: pK_S= 10,8, Histidin: pK_S= 6,0, Asparaginsäure: pK_S= 4,1 und Glutaminsäure: pK_S= 4,1) verfügen über leicht ionisierbare Seitenketten, die abhängig von ihrem pK_S-Wert, entweder

geladen oder ungeladen vorliegen (Pace et al., 2009, Stryer). Diese Eigenschaft befähigt diese Aminosäuren, abhängig vom pH-Wert der Umgebung und ihrem individuellen pK_S-Wert, entweder als Säure (Protonendonor) oder als Base (Protonenakzeptor) in der Interaktion mit anderen Aminosäuren zu fungieren. Durch diesen Protonen-*shuttle* (Protonenrelais-System) können Reaktionen ermöglicht oder Ionenbindungen eingegangen werden (Stryer). Besonders das Vorkommen von Histidinen, die nahe dem neutralen pH-Wert ihren Ionisierungsgrad ändern können, könnte den Unterschied in den pH-Optima der verschiedenen Systeme ausmachen. Sequenzanalysen ergaben jedoch, dass das 3kptC-System über keine Histidine verfügt, während das 4oq1- und das Spy-System jeweils zwei Histidinreste besitzen (siehe Abb.4.4 A und B). Lediglich die Nähe dieser Histidinreste relativ zur Isopeptid-Triade unterscheidet sich etwas (siehe Abb.4.4). Weitere Untersuchungen sind hier nötig, um die Rolle dieser Histidine für die Ausbildung von Isopeptidbindungen aufzuklären. Zudem könnten Analysen anderer leicht ionisierbarer Aminosäuren rund um die Isopeptid-Triade weitere Einblicke liefern.

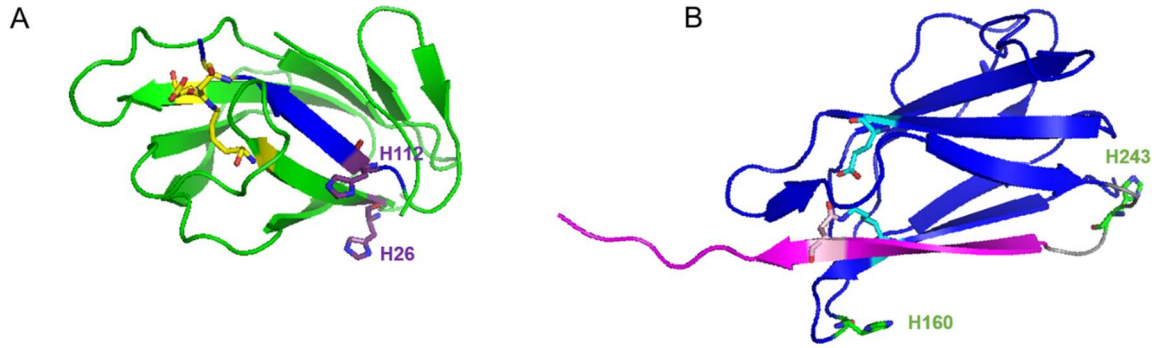


Abb.4.4: Vorkommen von Histidinen in dem Spy-System und dem 4oq1-System. (A) Gezeigt ist die Molekülstruktur des SpyCatcher (grün)-SpyTag (blau) Systems mit der reaktiven Triade (gelb) und den Histidinresten (lila: H26, H112). (B) Gezeigt ist die Molekülstruktur des 4oq1Catcher (blau)-4oq1Tag (rosa) Systems mit der reaktiven Triade (hellblau und rosa) und den Histidinresten (grün: H160, H243).

Unterschiedliche isoelektrische Punkte (pI) der Catcher- und Tag-Proteine wurden ebenfalls in Betracht gezogen, ein Grund für die unterschiedliche Reaktivität der Systeme zu sein, wobei das pH-Optimum eines Proteins nicht mit dem isoelektrischen Punkt korreliert (Talley und Alexov, 2011, Pace et al., 2009). Zudem wurde versucht die reaktiven Asparaginreste (N) des 4oq1- und des 3kptC-Systems zu Asparaginsäure (D) wie im effizienten Spy-System zu mutieren, um dadurch eine eventuelle Effizienzsteigerung der Reaktion zu erzielen. Die Asparagin zu Asparaginsäure-Substitutionen führten jedoch bei beiden Systemen zu einem kompletten Verlust der Bildung von kovalenten Catcher-Tag Komplexen (siehe 3.3.2.5). Diese Beobachtung spricht dafür, dass die N→D-Mutationen in den neu identifizierten Systemen komplexere Auswirkungen auf die chemische Reaktion zur Ausbildung von Isopeptidbindungen haben, als nur die Tatsache, dass entweder Ammoniak oder Wasser

freigesetzt wird. Jedoch konnten sowohl 4oq1Tag (reaktives Asparagin) als auch 3kptCTag (reaktives Asparagin) mit SpyCatcher kreuzreagieren und kovalente Catcher-Tag Komplexe bilden, wenn auch sehr viel schlechter als SpyTag (reaktive Asparaginsäure) mit SpyCatcher (siehe Kapitel 3.2.3). Die analoge Substitution von Asparaginsäure durch Asparagin im Spy-System führte ebenfalls zur Ausbildung einer Isopeptidbindung, auch wenn dies viel langsamer erfolgte als die Wildtyp-Reaktion zwischen Asparaginsäure und Lysin (Hagan et al., 2010, Kang und Baker, 2011). Dies lässt vermuten, dass theoretisch jedes Catcher-Tag System in der Lage ist, Isopeptidbindungen sowohl zwischen Lysin und Asparagin als auch zwischen Lysin und Asparaginsäure auszubilden, nur eben mit unterschiedlicher Effizienz. Neben den strukturellen Voraussetzungen und der Proteinstabilität, vor allem der größeren Catcher-Proteine, könnten feine Unterschiede in der Proteinumgebung der Isopeptidbindung bestimmen, ob eher Asparaginsäure (Wasserabspaltung) oder Asparagin (Freisetzung von Ammoniak) begünstigt wird (Kang und Baker, 2011, Hagan et al., 2010). Hinzu kommt, dass Asparagin im Gegensatz zur geladenen Asparaginsäure nicht nur als Wasserstoffakzeptor, sondern auch als Wasserstoffdonor bei der Ausbildung von Wasserstoffbrückenbindungen fungieren kann (Mian et al., 1991). Dieser Unterschied könnte über den Stabilisierungsgrad, sowohl der Übergangszustände als auch der resultierenden Isopeptid-Triade, die Effizienz der Reaktionen zur Ausbildung von Isopeptidbindungen beeinflussen (siehe Kapitel 1.2.2.2, Kang et al., 2007, Kang und Baker, 2011, Veggiani et al., 2014).

4.3.6 Unterschiedliche intrinsische Eigenschaften der CnaB-Domänen

In der Literatur sind Proteine bzw. Proteindomänen beschrieben, bei denen die Ausbildung einer intramolekularen Isopeptidbindung natürlicherweise langsam oder überhaupt nur während der Pilus-Polymerisation, also nicht in isolierten rekombinanten Proteinen, erfolgt (Krishnan, 2015). Häufig handelt es sich dabei um die N-terminalen Domänen (D1-Domänen) von Hauptpilinen gram-positiver Bakterien (z.B. D1-Domäne von SpaD aus *Corynebacterium diphteriae*, von RrgB aus *Streptococcus pneumoniae* und von BcpA aus *Bacillus cereus*) (Kang et al., 2014, Krishnan, 2015, Budzik et al., 2009, Paterson und Baker, 2011, Pröschel et al., 2017). Die N-terminalen Domänen der Grundgerüstpiline zeichnen sich generell durch eine erhöhte intrinsische Flexibilität, Instabilität und Sensitivität gegenüber proteolytischem Abbau aus. Zudem besitzen diese Domänen fast immer das essentielle Lysin, das von der Sortase zur Pilus-Assemblierung durch die Knüpfung intermolekularer Isopeptidbindungen erkannt wird (siehe Einleitung 1.2.1), als Teil des konservierten Pilin-Motivs (YP**K**N) am C-Terminus der Domäne (Kang et al., 2014, Krishnan, 2015). Basierend auf Strukturanalysen wird vermutet, dass die ineffiziente Isopeptidbildung durch die relative Lage der reaktiven Aminosäuren der Isopeptid-Triade zueinander und zu dem Ort der

Ausbildung von intermolekularen Isopeptidbindungen (Pilin-Motiv) erklärt werden kann (Kang et al., 2014, Krishnan, 2015, Paterson und Baker, 2011). In der D1-Domäne des rekombinant hergestellten SpaD Pilins, bei dem die intramolekulare Isopeptidbindung langsam erfolgt, liegt das essentielle Lysin (Lys-179) des Pilin-Motivs genau neben dem reaktiven Asparaginrest (Asn-180), der Teil der Isopeptid-Triade ist (Kang et al., 2014). Folglich ist es sehr wahrscheinlich, dass die Ausbildung von intramolekularen Isopeptidbindungen durch die Ausbildung von intermolekularen Isopeptidbindungen und den daraus resultierenden Konformationsänderungen stark beeinflusst wird (Kang et al., 2014). Zudem liegt das Pilin-Motiv der D1-Domäne direkt an der Domänengrenze in unmittelbarer Nähe zur D1-D2 Interaktionsoberfläche, was darauf hindeutet, dass Domänen-Domänen Interaktionen auch einen Einfluss auf die Ausbildung von intramolekularen Isopeptidbindungen haben können (Kang et al., 2014, Hu et al., 2011). Im Fall der N-terminalen D1-Domäne des RrgB Pilins scheinen die reaktiven Aminosäuren der Isopeptid-Triade zu weit enfernt voneinander zu liegen, sodass es zu keiner intramolekularen Isopeptidbindung in den rekombinant exprimierten Proteinen kommt (Paterson und Baker, 2011, Kang et al., 2014). Dass es aber im vollständig assemblierten Pilus zur Ausbildung der intramolekularen Isopeptidbindung kommt, kann ebenfalls durch die unmittelbare Nähe des Asparaginrestes der reaktiven Triade zum Pilin-Motiv mit dem essentiellen Lysinrest erklärt werden. Durch die intermolekulare Verknüpfung der einzelnen Piline kann es zu Konformationsänderungen kommen, die die reaktiven Reste schließlich in räumliche Nähe zueinander bringen (Paterson und Baker, 2011, Kang et al., 2014). Somit müssen bei manchen Domänen zuerst Energiebarrieren überwunden werden, bevor die reaktiven Aminosäuren der Isopeptid-Triade in räumliche Nähe gebracht werden können. Diese Energiebarrieren scheinen in den isolierten Domänen höher zu sein, als wenn die Domänen Teil des Pilins (Domänen-Domänen Interaktionen) bzw. des gesamten assemblierten Pilus-Polymers (intermolekulare Verknüpfungen der Piline) sind (Kang et al., 2014, El Mortaji et al., 2012; Paterson und Baker, 2011). Wieder anderen N-terminalen Ig-ähnlichen Domänen von Hauptpilinen fehlen die reaktiven Aminosäuren zur Ausbildung einer intramolekularen Isopeptidbindung komplett (z.B. N-terminale CnaB-Domäne des SpaD-Homologs SpaA, Kang et al., 2014). Das zeigt deutlich, dass sich die Pilin-Domänen, die die Basis der Catcher-Tag Split-Isopeptid Systeme bilden, durchaus unterscheiden können. Neben den Grundkriterien, wie der Existenz der reaktiven Triade, unterscheiden sich oftmals auch die intrinsischen Eigenschaften dieser Domänen, wie die Geschwindigkeit zur Ausbildung von intramolekularen Isopeptidbindungen, z.B. aufgrund von konformationsbedingten Energiebarrieren, wie im Fall der N-terminalen D1-Domänen der Hauptpiline SpaD und RrgB (Kang et al., 2014, Krishnan, 2015). Ob es sich bei der langsamen Isopeptidbildung des 4oq1- und des 3kptC-Systems wirklich um eine intrinsische Eigenschaft der neu identifizierten Systeme handelt, könnten

Untersuchungen der Effizienz der Ausbildung der Isopeptidbindung innerhalb der CnaB-Ursprungsdomänen (D2 aus RrgC für das 4oq1-System bzw. CNA3 aus BcpA für das 3kptC-System) aufklären (Kang et al., 2014). Im Fall des 4oq1-Systems spricht zumindest die verbesserte Reaktivität durch Erhöhung der Temperatur (siehe 3.2.4.3) für eine Energiebarriere, die *in vitro* durch erhöhte Temperaturen überwunden werden kann (Kang et al., 2014). Die 3kptCatcher-3kptCTag Reaktion dagegen, wird mit Temperaturen über 25°C wieder ineffizienter (siehe 3.2.4.3), was dann eher gegen diese Theorie beim 3kptC-System spricht. Alle diese Punkte sprechen für eine Einzigartigkeit der Systeme und deren Catcher-Tag Proteinkomponenten, die trotz großer Ähnlichkeiten, individuelle Optimierungsstrategien erfordern. Außerdem wird deutlich, dass trotz des allgemein großen Potentials der CnaB-Domänen neue Split-Isopeptid Systeme hervorzubringen, lange nicht alle diese Domänen (z.B. flexible, N-terminale Domänen von Hauptpilinen) für die Spaltung in einen Catcher- und Tag-Teil geeignet sind (fehlende Stabilität, fehlende Fähigkeit zur Ausbildung von Isopeptidbindungen), und dass eine genaue Analyse der Ursprungsdomänen, bzw. der zugrundeliegenden Piline, sehr hilfreich ist. Die *in silico* Analyse kann dabei enorm helfen und eine Vorauswahl treffen, indem völlig ungeeignete Domänen direkt ausgeschlossen werden. Das Potential der bioinformatischen Analysen wird deutlich, da bereits ohne weitere Optimierungen aus zwei der vier Kandidatendomänen tatsächlich funktionelle Split-Isopeptid Systeme generiert werden konnten.

4.3.7 Fazit und Ausblick der Mutationsanalysen

Die folgende Abbildung 4.5 soll das Potential der verschiedenen Proteinvarianten, basierend auf den angewendeten, rationalen Mutagenese-Strategien (siehe 3.3, Abb.3.23), zur Verbesserung der Split-Domänen Rekonstitutionseffizienz der neu identifizierten Systeme, 4oq1 und 3kptC, zusammenfassen. Die PROSS-Server basierten Catcher-Varianten (siehe 3.3.2.4), die multiple Aminosäuresubstitutionen gleichzeitig betreffen, sind in Abbildung 4.5 nicht enthalten. Deutlich wird, dass es kein allgemeingültiges Erfolgsrezept für die Optimierung von Proteinen bzw. bestimmten Proteineigenschaften gibt, sodass ein und dieselbe Strategie das eine Proteinsystem verbessert, während das andere Proteinsystem dadurch nicht beeinflusst oder verschlechtert wird (siehe 4.1.2). Folglich muss jedes Protein individuell analysiert und in Zukunft weiter optimiert werden. Sobald die Analyse der optimalen Tag-Peptid-Länge, die aufgrund der Catcher-Tag Wechselwirkungen in der Interaktionsoberfläche eine wichtige Rolle bei der Split-Domänen Rekonstitution spielt (Zakeri et al., 2012, Zakeri und Howarth, 2010), abgeschlossen ist, bieten die größeren Catcher-Proteine, deren intrinsische Stabilität und Struktur maßgeblich über die Ausbildung von autokatalytischen Isopeptidbindungen entscheiden, allgemein mehr Angriffspunkte für

gezielte Mutationen als die kleinen Tag-Peptide. So nehmen die Tag-Peptide aufgrund ihrer geringen Größe keine Tertiärstruktur oder bestimmte dreidimensionale Faltung, außer eventuell die ihres Fusionsproteins, an und liegen wahrscheinlich vor der Assoziation mit den großen Catcher-Proteinen, wie die meisten Peptide, weitgehend unstrukturiert vor (Veggiani et al., 2014).

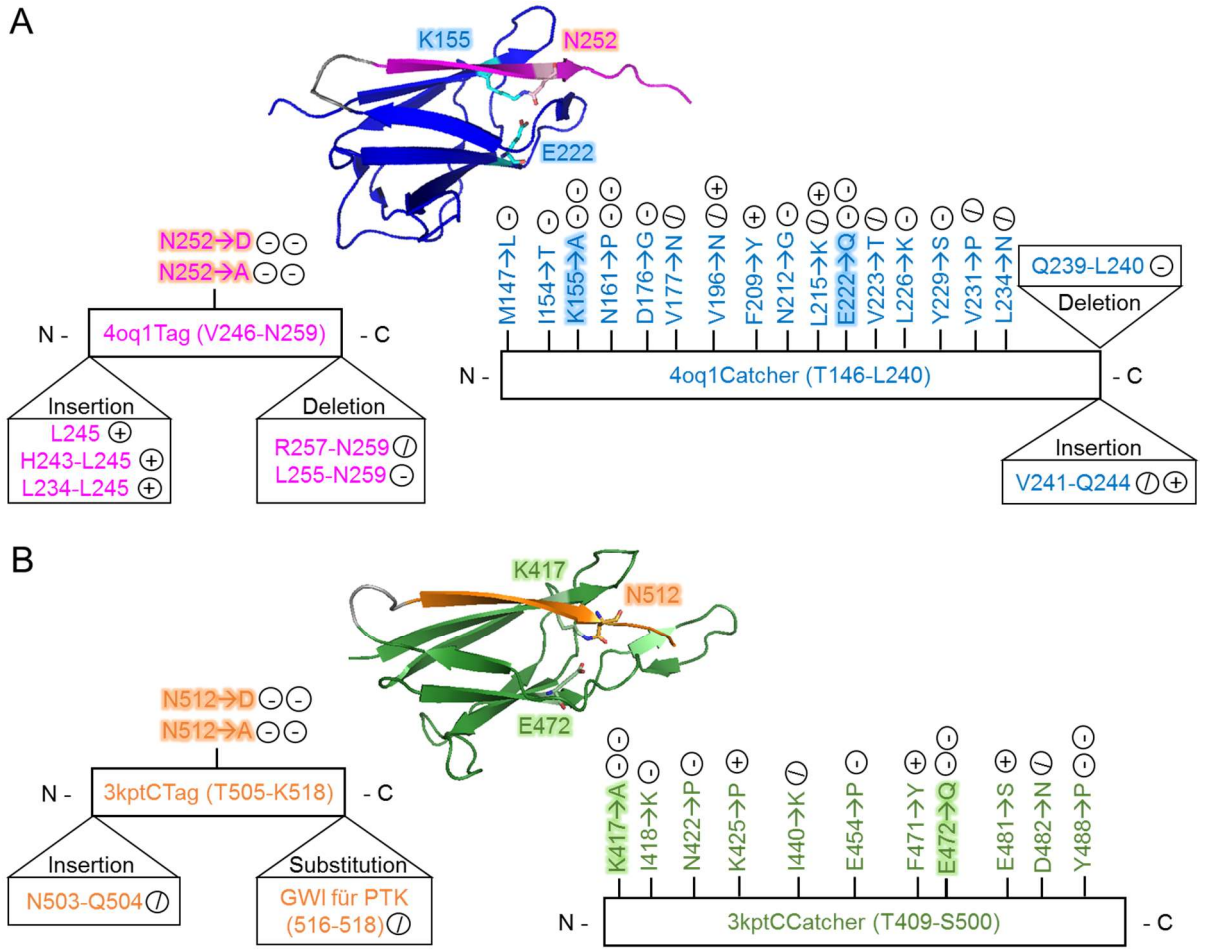


Abb.4.5: Zusammenfassung der getesteten Proteinvarianten basierend auf unterschiedlichen, rationalen Mutagenese-Strategien zur Verbesserung der Rekonstitutionseffizienz. Die einzelnen Proteinvarianten wurden in Rekonstitutionsexperimenten getestet und mit den entsprechenden Wildtyp-Reaktionen bezüglich ihrer Reaktivität zur Ausbildung von kovalenten Catcher-Tag Komplexen rein qualitativ verglichen (keine Quantifizierung). Die Einzel-Aminosäureaustausche bzw. die N- und C-terminalen Insertionen, Deletionen und Substitutionen mehrerer Aminosäuren, die die Reaktivität des 4oq1 (A)- bzw. des 3kptC (B)-Systems verbesserten, sind mit einem (+) gekennzeichnet. Die Proteinvarianten, die die Reaktivität im Vergleich mit dem Wildtyp-System verschlechterten, sind mit einem (–) hervorgehoben. (-) (-) bedeutet, dass die Fähigkeit zur Ausbildung eines kovalenten Catcher-Tag Komplexes komplett verloren gegangen ist. Alle anderen Mutationen hatten im Vergleich zum Wildtyp-System keinen oder kaum einen Effekt auf die Rekonstitutionseffizienz (/) und zeigten manchmal nur eine leichte Tendenz (/) (+).

In Zukunft gilt es zu testen, ob sich die positiven Effekte einzelner Aminosäuremutationen auf die Rekonstitutionseffizienz addieren, synergetisch wirken oder sich vielleicht sogar wieder aufheben und die Reaktivität zwischen Catcher- und Tag-Proteinpartnern wieder verschlechtern, wenn die besten Mutationen kombiniert werden (siehe Kapitel 4.1.2). Um

eindeutige Aussagen treffen zu können, müssten die Interaktionsanalysen in Replikaten mit anschließenden densitometrischen Auswertungen durchgeführt werden.

4.4 Mögliche neue Anwendungen der kovalenten Catcher-Tag Systeme

Durch die Verknüpfung von mehreren verschiedenen Proteinen als modulare Komponenten können Funktionen kombiniert, verstärkt oder für bestimmte Anwendungen zweckentfremdet werden, um Systeme mit neuen, oft multifunktionalen Eigenschaften zu erzeugen (Reddington und Howarth, 2015, Gilbert et al., 2017, Veggiani et al., 2016, Zakeri, 2015: *mix-and-match*). Zudem profitieren viele biologische Prozesse, natürlicherweise und artifiziell erzeugt, von der räumlichen Nähe, der an dem jeweiligen Prozess beteiligten Proteine (Good et al., 2011, Polka et al., 2016, Veggiani et al., 2016, Dueber et al., 2009, Bayer, 2017, Sweetlove und Fernie, 2013, Lee et al., 2012, Chen et al., 2014, Pröschel et al., 2015, Banerjee und Howarth, 2017). Aufgrund der vielen Vorteile der Catcher-Tag Systeme, die vor allem für die Reaktion zwischen SpyCatcher und SpyTag beschrieben wurden (siehe Kapitel 1.3.2), werden diese Systeme vielfältig als modulare Werkzeuge zur irreversiblen Verknüpfung von Proteinen eingesetzt und anderen Techniken zur Proteinligation vorgezogen (Veggiani et al., 2014: ,,Proteinsuperkleber“). Alternative post-translationale kovalente Proteinligationsstrategien sind oft abhängiger von experimentellen Bedingungen, rufen unerwünschte und unspezifische Nebenreaktionen hervor, sind nur *in vitro* anwendbar oder benötigen Enzyme, Cofaktoren oder chemische Modifikationen (Zakeri, 2015, Zakeri et al., 2012, Li et al., 2014, Veggiani et al., 2014). Das Potential und die Ideen der Anwendungsbereiche dieser kovalenten Catcher-Tag Systeme sind aber noch lange nicht ausgeschöpft und werden durch den stetigen Informationszugewinn und Forschungsfortschritt vorangetrieben. Insbesondere durch die Entdeckung neuer und im Idealfall auch orthogonaler kovalenter Catcher-Tag Systeme werden zahlreiche biotechnologische Anwendungen, die von dem modularen Baukastensystem profitieren, möglich. Die Spezifität verschiedener Systeme ermöglicht die programmierbare, kontrollierbare und irreversible Verknüpfung, aber auch Immobilisierung von Proteinen (Veggiani et al., 2016). Ob es sich bei den Zielproteinen um Enzyme, Fluoreszenzproteine oder jede beliebige andere Art von Proteinen handelt, hängt dabei alleine von der Anwendung und dem Verwendungsziel ab.

4.4.1 Entwicklung von synthetischen, stabilen Metabolonen

Handelt es sich bei den Zielproteinen um Enzyme, können beispielsweise stabile Multienzymkomplexe (Metabolone) über post-translationale, gezielte Catcher-Tag-Verknüpfungen generiert werden. Inspiriert von effizienten metabolischen Systemen in der Natur, wie dem Cellulosom oder der Tryptophansynthase, soll durch die räumliche Annäherung sequentieller Enzyme einer Reaktionskaskade *substrate channeling* ermöglicht werden, was die Effizienz der Reaktion und letztlich die Produktausbeute deutlich steigert (Conrado et al., 2008, Dueber et al., 2009, Zhang, 2011, Pröschel et al., 2015). Die Generierung von Multienzymkomplexen kann dabei über mehrere Strategien erreicht werden (siehe Abb.4.6) (Lee et al., 2012, Chen et al., 2014, Pröschel et al., 2015). Durch Funktionalisierung der Zielenzyme mit Tag oder Catcher können stabile Enzymketten gebildet werden (siehe Abb.4.6 C). Metabolone können auch über Koimmobilisierung der Zielenzyme auf demselben Träger (z.B. Magnetpartikel, Kang et al., 2014, Jia et al., 2013, Jia et al., 2014), derselben Matrix (z.B. Alginat, Polyethylenoxide, Hydrogele) oder auf Scaffolds (z.B. aufgebaut aus Catcher-Proteinen) generiert werden (siehe Abb.4.6 A und B). Alle diese Strategien sind einfach zu realisieren, da die Catcher-Tag Systeme genetisch kodiert sind und sowohl zur Funktionalisierung von Materialien (biologischer oder chemischer/ synthetischer Natur) als auch von Zielproteinen über Genfusionen verwendet werden können. Im Vergleich zu bereits vorhandenen synthetischen Metabolonen (siehe Dueber et al., 2009, Moon et al., 2010) sind die über die Catcher-Tag Interaktion vermittelten Metabolone rein kovalenter Natur und sind deshalb stabil und auch unter harschen Bedingungen, z.B. in Bioreaktoren oder *microfluidic* Systemen, wo Krafteinwirkungen herrschen, verwendbar (Pröschel et al., 2015, Pröschel et al., 2017).

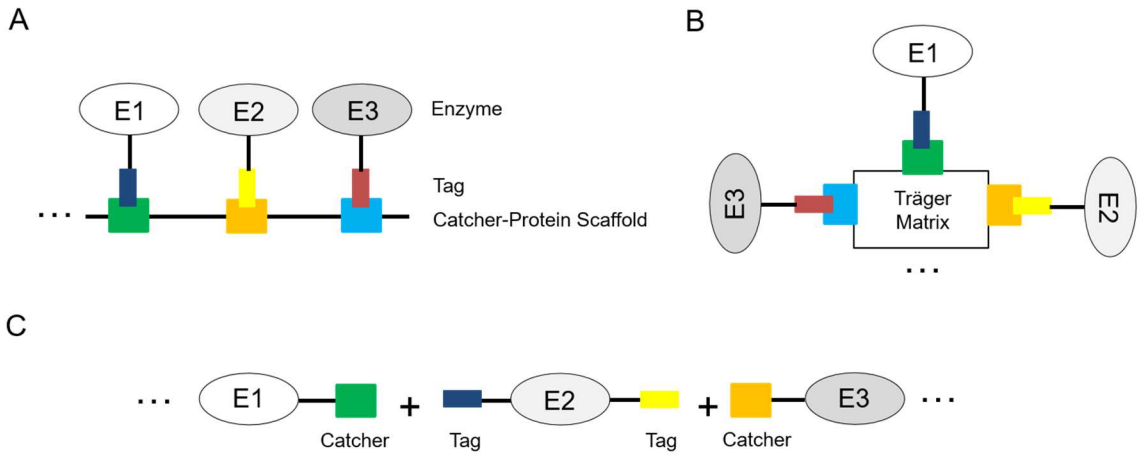


Abb.4.6: Verschiedene Strategien zur Generierung von post-translationalen Metabolonen über kovalente Catcher-Tag Systeme. Multiple Enzyme (E1 - E3) sind über spezifische, kovalente Catcher-Tag Interaktionen (bunte Kästchen) zu stabilen Multienzymkomplexen assembliert. (A) Kolokalisierung der Tag-Enzym Fusionsproteine an einem Catcher-Scaffold. (B) Koimmobilisierung von Tag-Enzym Fusionsproteinen an Catcher-funktionalisierten Materialien. (C) Catcher-Tag vermittelte Enzymkette.

4.4.2 Anwendungen bei BioLEDs

Werden rekombinant hergestellte und aufgereinigte Fluoreszenzproteine post-translational über diese Catcher-Tag Systeme kovalent miteinander verknüpft, können stabile Oligomere mit bestimmten Farbeigenschaften für die Lichtindustrie hergestellt werden. Durch die Verknüpfung von blau-, grün- und rot-fluoreszierenden Proteinen (mTagBFP, eGFP und mCherry) konnte ein stabiles, weiß-fluoreszierendes Trimer hergestellt werden (siehe Abb.3.54 im Ergebnisteil), das bei der Entwicklung von weißen BioHLEDs (*Bioinspired Hybrid Light-Emitting Diodes*) Anwendung finden könnte. Proteinbasierte, weiße LEDs, die aus drei Polyethylenoxid (PEO)-Schichten mit integrierten Proteinen bestehen, wurden bereits generiert (3-*layer cascade coating*: Weber und Niklaus et al., 2015). Das weiße Trimer stellt eine einfache und elegante Alternative zu diesen drei übereinander gelagerten PEO-Proteinschichten, die in der Summe weißes Licht erzeugen, dar. Durch den Einsatz von alternativen, orthogonalen Catcher-Tag Systemen, wie dem Spy- und dem Snoop-System, können exakt programmierbare Trimere generiert werden, was die Variation und Feinjustierung der einzelnen Farbanteile erlaubt. Durch die räumliche Nähe der Fluoreszenzproteine innerhalb des Trimers ist zudem ein verbesserter Energietransfer (verbesserte Emissionseffizienz/ *quantum yield*) denkbar.

4.4.3 Pflanzenbiotechnologische Anwendungen

4.4.3.1 Organellen-Trap

Die Entdeckung der generellen Funktionalität des Spy-Systems und des 4oq1-Systems in Pflanzen (siehe 3.4) eröffnet zahlreiche, neue Anwendungen in der Pflanzenbiotechnologie. Neben der einfachen und schnellen Isolierung von mit Tag-Peptiden markierten Chloroplasten über SpyCatcher-funktionalisierte Magnetpartikel (siehe 3.5.3), könnten auf diese Weise theoretisch alle Zellorganellen oder Zielproteine spezifisch aus Pflanzenfiltrat/ Gesamtzellextrakt extrahiert oder in der Zelle spezifisch angesteuert und markiert werden (Zakeri et al., 2012). Organellen-spezifische Signalpeptide oder *targeting*-Sequenzen zur Markierung der gesuchten Organellen oder Strukturen und die Generierung von Tag-modifizierten Genfusionen sind dafür nötig. Inwieweit diese Catcher-Tag Systeme im Vergleich zu den bereits etablierten Methoden zur Aufreinigung von Zellorganellen Anwendung finden, muss noch genauer analysiert werden. Sowohl der Zeit- und Materialaufwand, die Quantität als auch die Qualität der Isolierung müssen dabei berücksichtigt und mit den gängigen, bereits existierenden Methoden verglichen werden. Chloroplasten beispielsweise, lassen sich relativ einfach und schnell durch grobe subzelluläre Fraktionierung (Zentrifugation) aus Gesamtzellextrakt isolieren, sodass die

Isolierung über Catcher-Tag Systeme, wo die Pflanzen erst mittels Agrobakterien transformiert werden müssen, weder einen zeitlichen Vorteil noch einen Vorteil bezüglich der Chloroplastenausbeute bringt. Außerdem wird deutlich, dass die Transformationseffizienz der Tag-Signalpeptid Fusionen eine direkte Auswirkung auf die Ausbeute an isolierten Strukturen hat, jedenfalls bei transienten Transformationen. Lediglich ein Qualitätsvorteil bezüglich der Reinheit durch das spezifische Herausfischen von markierten Organellen über Catcher-Tag Interaktionen ist denkbar. Besonders interessant ist die Methode daher für die Isolierung von Organellen, die mit üblichen Methoden nur schwer und aufwändig zu isolieren sind oder für Anwendungen, die von der gleichzeitigen Isolierung von verschiedenen Organellen über verschiedene, orthogonale Catcher-Tag Systeme profitieren. Auf kovalenten Catcher-Tag Interaktionen basierende Abwandlungen der INTACT (*isolation of nuclei tagged in specific cell types*)-Methode sind theoretisch ebenfalls denkbar. Bei der INTACT-Methode werden Zellkerne eines bestimmten Zelltyps in transgenen *Arabidopsis thaliana* Pflanzen durch ein biotinyliertes Kernhüllenprotein spezifisch markiert und anschließend über Streptavidin-gekoppelte magnetische Partikel über Affinitätsinteraktion isoliert (Deal und Henikoff, 2011). Die Expression von Chloroplastenanker-Tag Fusionsproteinen (z.B. OEP7-SpyTag) unter der Kontrolle eines zelltypspezifischen Promotors könnte so die spezifische Markierung von Chloroplasten nur eines bestimmten Zelltyps, wie beispielsweise Schließzellen oder Geleitzellen, und die anschließende Aufreinigung über Catcher-gekoppelte magnetische Partikel (z.B. SpyCatcher-*beads*, Abb.3.53) aus Gesamtzellextrakt ermöglichen.

4.4.3.2 Spy-Trap als kovalente Alternative zur GFP-Trap

Eine weitere mögliche pflanzenbiotechnologische Anwendung dieser Catcher-Tag Systeme wäre die Entwicklung einer Alternative zur GFP- oder RFP-Trap zur Identifizierung und Verifizierung von Protein-Protein Interaktionen (siehe Abb.4.7). Somit könnte die Spy-Trap eine kovalente Alternative zur GFP-Trap darstellen, die nicht auf einer Antigen-Antikörper basierenden Affinitätsinteraktion beruht, sondern auf der spezifischen und kovalenten Interaktion zwischen SpyCatcher auf den Magnetpartikeln (siehe Abb.3.53) und SpyTag-Fusionsproteinen (Köderproteinen), von denen man mögliche Interaktionspartner (Beuteproteine) identifizieren und verifizieren möchte (siehe Abb.4.7). Für eine Spy-Trap müssten anstatt GFP-Fusionen, SpyTag-Zielproteinfusionen und anstatt kommerziellen Anti-GFP Magnetpartikeln, SpyCatcher-Magnetpartikel hergestellt werden (siehe Kapitel 2.8.1). Ein großer Vorteil der Spy-Trap gegenüber der GFP-Trap liegt in der Größe des Proteins, das an das Zielprotein über Genfusion angehängt werden muss. Während der SpyTag lediglich aus 13 Aminosäuren besteht, besteht eGFP (*enhanced* GFP) aus 239 Aminosäuren (ca. 27 kDa) (http://www.uniprot.org/uniprot/C5MKY7). Vor allem große Genfusionen bringen

aber die Gefahr mit sich, die native Faltung und Struktur ihrer Fusionspartner zu stören bzw. zu verändern (Yang et al., 2016, Zakeri, 2015, Veggiani et al., 2014, Levary et al., 2011, Chen et al., 2014), sodass eventuell native Interaktionspartner des Zielproteins nicht gefunden bzw. falsch positive Interaktionspartner identifiziert werden. Die große β-*barrel* Struktur des GFP-Proteins verhindert zudem, eventuell über sterische Hinderung, die Integration der Zielproteine in Interaktionskomplexen, sodass native Interaktionen verhindert werden. In diesem Zusammenhang könnte es aber auch umgekehrt bei der Spy-Trap Probleme geben, sodass der kleine SpyTag als Köderproteinfusion in Interaktionskomplexe gezogen wird und anschließend aufgrund von Unzugänglichkeit nicht mehr von SpyCatcher-Magnetpartikeln erkannt wird.

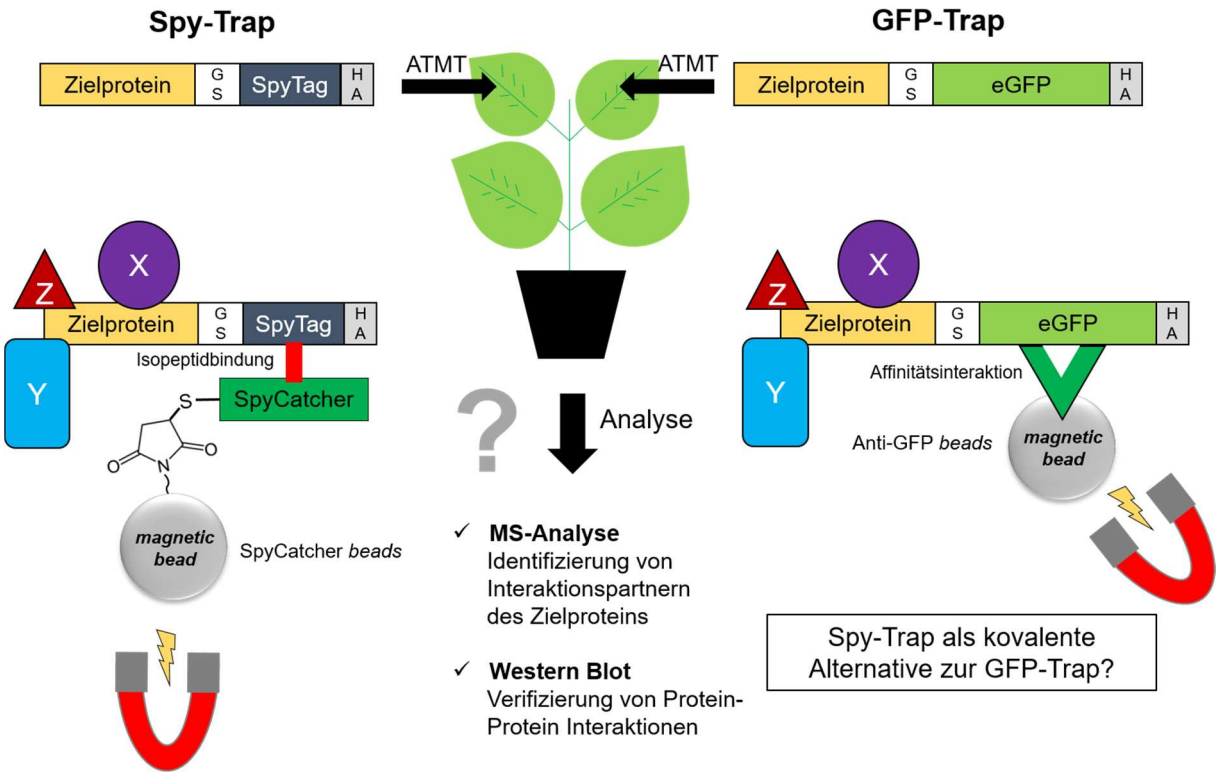


Abb.4.7: Vergleichendes Schema zwischen Spy-Trap und GFP-Trap. Beide Methoden erfordern die Agrobakterium-vermittelte Transformation (ATMT) der Pflanze mit den Zielproteinfusionen (Köderproteine), entweder Zielprotein-SpyTag oder Zielprotein-eGFP. Nach Expression der Köderproteine *in planta* werden entweder SpyCatcher-*beads* oder Anti-GFP-*beads* zur Isolierung der Zielproteine samt Interaktionspartnern (X, Y, Z: Beuteproteine) aus dem Gesamtproteinextrakt verwendet.

4.4.3.3 *Metabolic engineering* in Pflanzen

Durch die spezifische Markierung von Organellen oder Strukturen innerhalb der Pflanzenzelle über Catcher- oder Tag-markierte Signalpeptide werden auch *metabolic engineering* Anwendungen zur gezielten Manipulation endogener Stoffwechselwege, z.B. zur verbesserten Ausbeute eines bestimmten Produktes, möglich (Capell und Christou, 2004,

Sonnewald, 2003, Weber und Bräutigam, 2013, Jørgensen et al., 2005). Über die kovalente und spezifische Catcher-Tag Interaktion können Enzyme, die über Genfusion entweder mit Catcher-Protein oder Tag-Peptid modifiziert wurden, ebenfalls an bestimmte Orte in der Zelle dirigiert und dauerhaft immobilisiert werden (Polka et al., 2016). Die Fixierung von Enzymen an ihrem natürlichen Wirkort oder die Kolokalisierung von sequentiellen Enzymen einer Reaktionskaskade an einem strategisch günstigen Ort in der Zelle (z.B. Kopplung glykolytischer Enzyme an die Mitochondrienmembran, Giegé et al., 2003, Sweetlove und Fernie, 2013) könnten aufgrund verkürzter Diffusionswege und *substrate channeling* zu metabolischen Vorteilen führen. Sowohl endogene Stoffwechselprozesse könnten effizienter ablaufen oder gezielt durch die artifizielle Veränderung des Metabolitflusses zur Produktion bestimmter Produkte manipuliert werden als auch heterologe Enzyme und Stoffwechselwege könnten an einem gezielten subzellulären Ort eingeführt werden (siehe Abb.4.8).

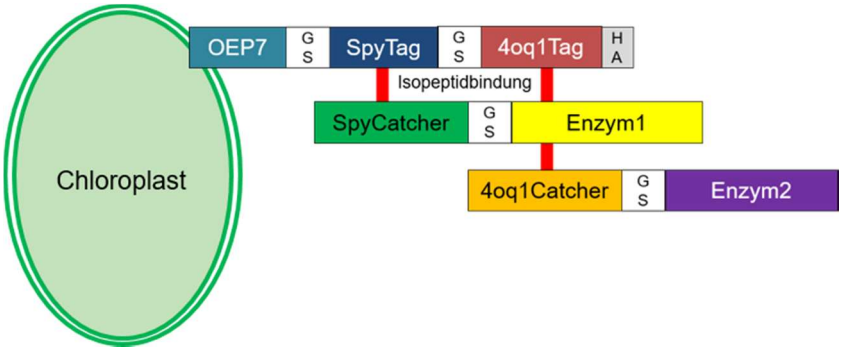


Abb.4.8: Artifizielles Metabolon an den Chloroplasten durch Kolokalisierung von Enzymen über kovalente Catcher-Tag Interaktionen. Die in Pflanzen funktionellen Split-Isopeptid Systeme, SpyCatcher-SpyTag und 4oq1Catcher-4oq1Tag, können verwendet werden, um die Enzyme 1 und 2 über den Chloroplastenanker OEP7 an Chloroplasten zu kolokalisieren. Die Isopeptidbindungen, die sich zwischen den Catcher- und Tag-Fusionsproteinen ausbilden, sind in roten Balken dargestellt.

4.4.4 Weiterer Optimierungsbedarf der Catcher-Tag Systeme

Die Split-Isopeptid Systeme, rund um das Vorzeigesystem SpyCatcher-SpyTag, bieten viele Vorteile gegenüber allen anderen Techniken zur kovalenten, post-translationalen Verknüpfung von Proteinen (Zakeri et al., 2012, Veggiani et al., 2014, Gilbert et al., 2017, Zakeri, 2015, Zakeri und Howarth, 2010). Das SpyCatcher-SpyTag System besteht aus genetisch kodierten Teilen, ist stabil, irreversibel, robust (weitgehend unabhängig von allen Reaktionsbedingungen, auch Redox-insensitiv, da keine Cysteine enthalten sind), spezifisch, sehr effizient und kann sowohl *in vitro* als auch *in vivo* (intra- und extrazellulär) angewendet werden (siehe auch 1.3.2). Zudem können SpyCatcher und SpyTag überall innerhalb der Polypeptidkette positioniert werden und brauchen keinerlei chemische Modifikationen oder Cofaktoren um autokatalytisch miteinander zu reagieren (Zakeri et al., 2012, Veggiani et al., 2014, Zakeri, 2015). Somit ist die SpyCatcher-SpyTag Methode ziemlich simpel, einfach

handzuhaben und erfordert nur die rekombinante Expression der Protein-Reaktionspartner (Veggiani et al., 2016). Doch wie jedes System hat auch das SpyCatcher-SpyTag System seine Einschränkungen. Dadurch, dass SpyCatcher und SpyTag genetisch kodiert sind, müssen sie an Zielproteine als Genfusionen kloniert werden. Wie bei allen Genfusionen kann das, vor allem bei dem größeren SpyCatcher-Protein, zu Problemen der nativen Faltung der Zielproteine und damit zu Fehlfunktionen führen (Veggiani et al., 2014, Zakeri, 2015). Außerdem hinterlässt die Verknüpfung von Proteinen über diese Catcher-Tag Systeme Spuren, was bedeutet, dass die resultierenden Proteinkonjugate immer auch die Catcher- und Tag-Sequenzen als Artefakte der Assemblierung tragen (Veggiani et al., 2014, Zakeri et al., 2012). Obwohl die SpyCatcher-SpyTag-Reaktion in *in vitro* Systemen mit Proteinkonzentrationen im µM-Bereich höchst effizient innerhalb von Minuten abläuft, liegt die Reaktionsrate weit unter dem Diffusionslimit, was vor allem die initiale Komplexbildung (*on-rate*) zwischen den Reaktionspartnern erschwert (Veggiani et al., 2014, Zakeri et al., 2012). Dies schränkt insbesondere die Effizienz der SpyCatcher-SpyTag-Reaktion *in vivo* ein, wenn Konzentrationen im nM-Bereich oder geringer vorherrschen. Aus diesem Grund werden nicht-kovalente Affinitätssysteme wie Biotin und Streptavidin, deren Assoziationsrate nahe dem Diffusionslimit liegt, bei geringen Konzentrationen noch häufig dem SpyCatcher-SpyTag System vorgezogen (Veggiani et al., 2014). Ein weiterer eventueller Nachteil des Spy-Systems ist, dass die Reaktion spontan und irreversibel abläuft und dadurch nicht rückgängig ist und vor allem auch nicht wirklich kontrolliert werden kann, außer die kovalente Reaktion zwischen SpyCatcher und SpyTag wird aktiv, z.B. durch sterische Blockierung der Bindestelle, verhindert (Zakeri, 2015, Matsunaga et al., 2013). Im Zusammenhang mit der Herstellung von Catcher-/ Tag-basierten Impfstoffen (Brune et al., 2016, Brune et al., 2017) wurden zudem eventuelle immunogene Eigenschaften dieser Systeme diskutiert, die bakteriellen Ursprung haben (Veggiani et al., 2016). Die CnaB-Domänen, von denen sich die Catcher-Tag Systeme ableiten, stellen jedoch Immunglobulin-ähnliche Domänen dar, die Varianten der konstanten Domäne des humanen Immunglobulins sind (siehe Einleitung). Folglich könnte das humane Immunsystem deshalb weniger stark reagieren, was übrigens auch den Bakterien bei der Adhäsion des Wirtsgewebes zugute kommt (Kang und Baker, 2012, Baker und Young, 2015, Vengadesan und Narayana, 2011). Einschränkungen in der Anwendung, vor allem der neu entdeckten Catcher-Tag Systeme, ergeben sich durch die fehlende vollständige Orthogonalität der Systeme untereinander. Auch wenn jeder Catcher seinen spezifischen Tag bevorzugt, kann eine Kreuzreaktion nicht ausgeschlossen werden, weshalb die gleichzeitige Verwendung dieser Systeme noch limitiert ist. Gezielte, sowohl rationale als auch evolutionäre (z.B. Phage Display) ,,Protein Engineering“-Strategien zur Verbesserung der Spezifität der Catcher-Tag Systeme sind hier nötig. Die Frage ist jedoch, inwieweit unterscheiden sich die Systeme und kann eine vollständige Spezifität überhaupt

erreicht werden in Anbetracht der Tatsache, dass sich alle Catcher-Tag Systeme strukturell sehr ähneln und diese Grundstruktur auch konserviert ist, damit sich die Isopeptidbindung ausbilden kann (siehe Einleitung). Verschiedene 4oq1Tag-Varianten, sogar die Integration eines weiteren β-Faltblattstranges in die 4oq1Tag-Sequenz, konnten die Kreuzreaktionen nicht minimieren. Folglich muss entweder die Proteinstruktur deutlicher verändert werden, was natürlich auch zum kompletten Verlust der Fähigkeit zur Ausbildung von Isopeptidbindungen führen kann, oder eventuell spezifischere Interaktionen zwischen zusammengehörigen Catcher- und Tag-Proteinen künstlich eingeführt werden. Theoretisch wären spezifische elektrostatische Wechselwirkungen (Salzbrücken) zwischen dem Catcher-Protein und dem Tag-Peptid eines Systems denkbar (Yan et al., 2008). Während sich zusammengehörige Reaktionspartner über gegenteilig geladene Aminosäuren anziehen, sollen sich Catcher- und Tag-Proteine von unterschiedlichen Systemen durch gleichartig geladene Aminosäuren innerhalb der Interaktionsoberfläche sogar abstoßen.

Aufgrund der vielen Vorteile, die das Spy-System nach wie vor gegenüber allen Proteinligationssystemen bietet, ist das Potential riesig und rechtfertigt jede Bemühungen. Das langfristige Ziel ist es, alle Catcher-Tag Systeme durch weitere Optimierungen auf ein ähnliches Niveau zu bringen, um echte Alternativen im modularen Baukastensystem zur post-translationalen Kontrolle von Proteinen zu haben. Diese Arbeit soll einen wichtigen Beitrag dazu leisten.

Literaturverzeichnis

Alberts, Lehrbuch der Molekularen Zellbiologie 3. Auflage, Alberts, Bray, Hopkin, Johnson, Lewis, Raff, Roberts, Walter WILEY-VCH Verlag GmbH & Co. KGaA, ISBN-13: 978-3-527-31160-6, ISBN-10: 3-527-31160-2

Alexov, E. (2004). Numerical calculations of the pH of maximal protein stability: The effect of the sequence composition and three-dimensional structure. *European Journal of Biochemistry*, *271*(1), 173–185. http://doi.org/10.1046/j.1432-1033.2003.03917.x

Alves, N. J., Turner, K. B., Daniele, M. A., Oh, E., Medintz, I. L., & Walper, S. A. (2015). Bacterial Nanobioreactors-Directing Enzyme Packaging into Bacterial Outer Membrane Vesicles. *ACS Applied Materials and Interfaces*, *7*(44), 24963–24972. http://doi.org/10.1021/acsami.5b08811

Baker, E. N., und Young, P. G. (2015). Host-microbe Interactions: Convergent weaponry in a biological arms race. eLife 2015;4:e08710. DOI: 10.7554/eLife.08710

Banerjee, A., & Howarth, M. (2017). Nanoteamwork: covalent protein assembly beyond duets towards protein ensembles and orchestras. *Current Opinion in Biotechnology*, *10cc*(Figure 1), 16–23. http://doi.org/10.1016/j.copbio.2017.10.006

Baneyx, F., & Mujacic, M. (2004). Recombinant protein folding and misfolding in Escherichia coli. *Nature Biotechnology*, *22*(11), 1399–1407. http://doi.org/10.1038/nbt1029

Bayer, E. A. (2017). Cellulosomes and designer cellulosomes: why toy with Nature? *Environmental Microbiology Reports*, *9*(1), 14–15. https://doi.org/10.1111/1758-2229.12489

Bayer, E. a, Belaich, J.-P., Shoham, Y., & Lamed, R. (2004). The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annual Review of Microbiology*, *58*, 521–54. http://doi.org/10.1146/annurev.micro.57.030502.091022

Bayer, E. a, Morag, E., & Lamed, R. (1994). The cellulosome - a treasure- trove for biotechnology. *Trends in Biotechnology*, *12*(9), 259–265.

Bazan, J., Całkosiński, I., & Gamian, A. (2012). Phage display — A powerful technique for immunotherapy 1. Introduction and potential of therapeutic application. *Human Vaccine and Immunotherapics*, *8*(December), 1817–1828. http://doi.org/10.4161/hv.21703

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The protein data bank. *Nucleic Acids Research*, *28*(1), 235–242. http://doi.org/10.1093/nar/28.1.235

Biotechnologie, William J. Thieman, Michael A. Palladino, Pearson Studium 2007, ISBN 978-3-8273-7236-9

Bodelón, G., Palomino, C., & Fernández, L. Á. (2013). Immunoglobulin domains in Escherichia coli and other enterobacteria: From pathogenesis to applications in antibody technologies. *FEMS Microbiology Reviews*, *37*(2), 204–250. https://doi.org/10.1111/j.1574-6976.2012.00347.x

Bonnet, J., Cartannaz, J., Tourcier, G., Contreras-Martel, C., Kleman, J. P., Morlot, C., Vernet, T., & Di Guilmi, A. M. (2017). Autocatalytic association of proteins by covalent bond formation: a Bio Molecular Welding toolbox derived from a bacterial adhesin. *Scientific Reports*, *7*(January), 43564. http://doi.org/10.1038/srep43564

Bork, P., Holm, L., Sander, C. (1994). The Immunoglobulin Fold: Structural Classification,

Sequence Patterns and Common Core. *J. Mol. Biol.* 242(4), 309-320. https://doi.org/10.1006/jmbi.1994.1582

Borthakur, S., Andrejeva, G., & McDonnell, J. M. (2011). Basis of the intrinsic flexibility of the Cε3 domain of IgE. *Biochemistry*, *50*(21), 4608–4614. https://doi.org/10.1021/bi200019y

Botyanszki, Z., Tay, P. K. R., Nguyen, P. Q., Nussbaumer, M. G., & Joshi, N. S. (2015). Engineered catalytic biofilms: Site-specific enzyme immobilization onto E. coli curli nanofibers. *Biotechnology and Bioengineering*, *112*(10), 2016–2024. http://doi.org/10.1002/bit.25638

Brune, K. D., Buldun, C. M., Li, Y., Taylor, I. J., Brod, F., Biswas, S., & Howarth, M. (2017). Dual Plug-and-Display Synthetic Assembly Using Orthogonal Reactive Proteins for Twin Antigen Immunization. *Bioconjugate Chemistry*, *28*(5), 1544–1551. http://doi.org/10.1021/acs.bioconjchem.7b00174

Brune, K. D., Leneghan, D. B., Brian, I. J., Ishizuka, A. S., Bachmann, M. F., Draper, S. J., Biswas, S., & Howarth, M. (2016). Plug-and-Display: decoration of Virus-Like Particles via isopeptide bonds for modular immunization. *Scientific Reports*, *6*(October), 19234. http://doi.org/10.1038/srep19234

Bryksin, A. V., & Matsumura, I. (2010). Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. NIH Public Access. *BioTechniques*, *48*(6), 463–465. http://doi.org/10.2144/000113418.

Budzik, J. M., Poor, C. B., Faull, K. F., Whitelegge, J. P., He, C., & Schneewind, O. (2009). Intramolecular amide bonds stabilize pili on the surface of bacilli. *Proceedings of the National Academy of Sciences*, *106*(47), 19992–19997. http://doi.org/10.1073/pnas.0910887106

Bullock, W. O., Fernandez, J. M., and Short, J. M., XL1-Blue: a high-efficiency plasmid transforming recA Escherichia coli strain with β-galactosidase selection. Biotechniques, 5, 376–378 (1987).

Cao, Y., Liu, D., & Zhang, W.-B. (2017). Supercharging SpyCatcher toward an intrinsically disordered protein with stimuli-responsive chemical reactivity. *Chem. Commun.*, *53*(63), 8830–8833. http://doi.org/10.1039/C7CC04507G

Capell, T., & Christou, P. (2004). Progress in plant metabolic engineering. *Current Opinion in Biotechnology*, *15*(2), 148–154. http://doi.org/10.1016/j.copbio.2004.01.009

Chen, A. Y., Deng, Z., Billings, A. N., Seker, U. O. S., Lu, M. Y., Citorik, R. J., Zakeri, B., & Lu, T. K. (2014). Synthesis and patterning of tunable multiscale materials with engineered cells. *Nature Materials*, *13*(5), 515–23. http://doi.org/10.1038/nmat3912

Chen, R., Chen, Q., Kim, H., Siu, K. H., Sun, Q., Tsai, S. L., & Chen, W. (2014). Biomolecular scaffolds for enhanced signaling and catalytic efficiency. *Current Opinion in Biotechnology*, *28*, 59–68. http://doi.org/10.1016/j.copbio.2013.11.007

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., & Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, *66*(1), 12–21. http://doi.org/10.1107/S0907444909042073

Choi, E. J., & Mayo, S. L. (2006). Generation and analysis of proline mutants in protein G. *Protein Engineering, Design and Selection*, *19*(6), 285–289. http://doi.org/10.1093/protein/gzl007

Cohen, M., Reichmann, D., Neuvirth, H., & Schreiber, G. (2008). Similar chemistry, but

different bond preferences in inter versus intra-protein interactions. *Proteins: Structure, Function and Genetics*, *72*(2), 741–753. http://doi.org/10.1002/prot.21960

Conrado, R. J., Varner, J. D., & DeLisa, M. P. (2008). Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Current Opinion in Biotechnology*, *19*(5), 492–499. http://doi.org/10.1016/j.copbio.2008.07.006

Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall III, W. B., Snoeyink, J., Richardson, J. S., & Richardson, D. C. (2007). MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research*, *35*(SUPPL.2), 375–383. http://doi.org/10.1093/nar/gkm216

Dawson, P. E., & Kent, S. B. H. (2000). Synthesis of Native Proteins By Chemical Ligation. *Annual Review of Biochemistry*, *69*, 923–960. http://doi.org/10.1146/annurev.biochem.69.1.923

Deal, R. B., & Henikoff, S. (2011). The INTACT method for cell type–specific gene expression and chromatin profiling in Arabidopsis thaliana. *Nature Protocols*, *6*(1), 56–68. http://doi.org/10.1038/nprot.2010.175

Debès, C., Wang, M., Caetano-Anollés, G., Gräter, F. (2013). Evolutionary Optimization of Protein Folding. PLoS Comput Biol 9(1): e1002861. doi:10.1371/ journal.pcbi.1002861

Deblaere, R., Bytebier, B., De Greve, H., Deboeck, F., Schell, J., Van Montagu, M., Leemans, J. (1985). Efficient octopine Ti plasmid-derived vectors for Agrobacterium-mediated gene transfer to plants. *Nucleic Acids Res.* 13(13): 4777-88

Deivanayagam, C. C. S., Rich, R. L., Carson, M., Owens, R. T., Danthuluri, S., Bice, T., Höök, M., & Narayana, S. V. L. (2000). Novel fold and assembly of the repetitive B region of the Staphylococcus aureus collagen-binding surface protein. *Structure*, *8*(1), 67–78. http://doi.org/10.1016/S0969-2126(00)00081-2

Deivanayagam, C. C. S., Wann, E. R., Chen, W., Carson, M., Rajashankar, K. R., Höök, M., & Narayana, S. V. L. (2002). A novel variant of the immunoglobulin fold in surface adhesins of Staphylococcus aureus: Crystal structure of the fibrinogen-binding MSCRAMM, clumping factor A. *EMBO Journal*, *21*(24), 6660–6672. http://doi.org/10.1093/emboj/cdf619

Dombkowski, A. A. (2003). Disulfide by Design™: A computational method for the rational design of disulfide bonds in proteins. *Bioinformatics*, *19*(14), 1852–1853. http://doi.org/10.1093/bioinformatics/btg231

Dombkowski, A. A., Sultana, K. Z., & Craig, D. B. (2014). Protein disulfide engineering. *FEBS Letters*, *588*(2), 206–212. http://doi.org/10.1016/j.febslet.2013.11.024

Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V, Prather, K. L. J., & Keasling, J. D. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnology*, *27*(8), 5–8. http://doi.org/10.1038/nbt.1557

El Mortaji, L., Fenel, D., Vernet, T., & Di Guilmi, A. M. (2012). Association of RrgA and RrgC into the Streptococcus pneumoniae pilus by sortases C-2 and C-3. *Biochemistry*, *51*(1), 342–352. http://doi.org/10.1021/bi201591n

El Mortaji, L., Terrasse, R., Dessen, A., Vernet, T., & Di Guilmi, A. M. (2010). Stability and assembly of pilus subunits of Streptococcus pneumoniae. *Journal of Biological Chemistry*, *285*(16), 12405–12415. http://doi.org/10.1074/jbc.M109.082776

Fairhead, M., Veggiani, G., Lever, M., Yan, J., Mesner, D., Robinson, C. V., Dushek, O., van der Merwe, P. A., & Howarth, M. (2014). SpyAvidin hubs enable precise and ultrastable

orthogonal nanoassembly. *Journal of the American Chemical Society*, *136*(35), 12355–12363. http://doi.org/10.1021/ja505584f

Fellouse, F. A., Barthelemy, P. A., Kelley, R. F., & Sidhu, S. S. (2006). Tyrosine plays a dominant functional role in the paratope of a synthetic antibody derived from a four amino acid code. *Journal of Molecular Biology*, *357*(1), 100–114. http://doi.org/10.1016/j.jmb.2005.11.092

Fierer, J. O., Veggiani, G., & Howarth, M. (2014). SpyLigase peptide-peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, E1176–81. http://doi.org/10.1073/pnas.1315776111

Fraser, J. S., Maxwell, K. L., & Davidson, A. R. (2007). Immunoglobulin-like domains on bacteriophage: weapons of modest damage? *Current Opinion in Microbiology*, *10*(4), 382–387. http://doi.org/10.1016/j.mib.2007.05.018

Fraser, J. S., Yu, Z., Maxwell, K. L., & Davidson, A. R. (2006). Ig-Like Domains on Bacteriophages: A Tale of Promiscuity and Deceit. *Journal of Molecular Biology*, *359*(2), 496–507. http://doi.org/10.1016/j.jmb.2006.03.043

Gao, X., Fang, J., Xue, B., Fu, L., & Li, H. (2016). Engineering Protein Hydrogels Using SpyCatcher-SpyTag Chemistry. *Biomacromolecules*, *17*(9), 2812–2819. http://doi.org/10.1021/acs.biomac.6b00566

Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D. and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*, 571–608.

Gfeller, D. (2012). Uncovering new aspects of protein interactions through analysis of specificity landscapes in peptide recognition domains. *FEBS Letters*, *586*(17), 2764–2772. http://doi.org/10.1016/j.febslet.2012.03.054

Giegé, P., Heazlewood, J. L., Millar, A. H., Fernie, A. R., Leaver, C. J., Sweetlove, L. J. (2003). Enzymes of Glycolysis Are Functionally Associated with the Mitochondrion in Arabidopsis Cells. *The Plant Cell* 15, 2140-2151

Giessen, T. W., & Silver, P. A. (2016). A Catalytic Nanoreactor Based on in Vivo Encapsulation of Multiple Enzymes in an Engineered Protein Nanocompartment. *ChemBioChem*, *17*(20), 1931–1935. http://doi.org/10.1002/cbic.201600431

Gilbert, C., Howarth, M., Harwood, C. R., & Ellis, T. (2017). Extracellular Self-Assembly of Functional and Tunable Protein Conjugates from Bacillus subtilis. *ACS Synthetic Biology*, *6*(6), 957–967. http://doi.org/10.1021/acssynbio.6b00292

Goldenzweig, A., Goldsmith, M., Hill, S. E., Gertman, O., Laurino, P., Ashani, Y., Dym, O., Unger, T., Albeck, S., Prilusky, J., Lieberman, R. L., Aharoni, A., Silman, I., Sussman, J. L., Tawfik, D. S., & Fleishman, S. J. (2016). Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. *Molecular Cell*, *63*(2), 337–346. http://doi.org/10.1016/j.molcel.2016.06.012

Good, M. C., Zalatan, J. G., & Lim, W. A. (2011). Scaffold Proteins: Hubs for Controlling the Flow of Cellular Information. *Science*, *332*(6030), 680–686. http://doi.org/10.1126/science.1198701

Gromiha, M.M. (2007). Prediction of protein stability upon point mutations. Biochemical Society Transactions Dec 01, 2007, 35 (6) 1569-1573; DOI: 10.1042/BST0351569

Hagan, R. M., Björnsson, R., McMahon, S. A., Schomburg, B., Braithwaite, V., Bühl, M., Naismith, J. H., & Schwarz-Linek, U. (2010). NMR spectroscopic and theoretical

analysis of a spontaneously formed lys-asp isopeptide bond. *Angewandte Chemie - International Edition*, *49*(45), 8421–8425. http://doi.org/10.1002/anie.201004340

Halaby, D. M., Poupon, A., & Mornon, J.-P. (1999). The immunoglobulin fold family: sequence analysis and 3D structure comparisons. *Protein Engineering, Design and Selection*, *12*(7), 563–571. http://doi.org/10.1093/protein/12.7.563

Hendrickx, A. P. A., Budzik, J. M., Oh, S., & Schneewind, O. (2011). Architects at the bacterial surface — sortases and the assembly of pili with isopeptide bonds. *Nature Publishing Group*, *9*(3), 166–176. http://doi.org/10.1038/nrmicro2520

Hoffmann, S., Funke, S. A., Wiesehan, K., Moedder, S., Glück, J. M., Feuerstein, S., Gerdts, M., Mötter, J., & Willbold, D. (2010). Competitively selected protein ligands pay their increase in specificity by a decrease in affinity. *Molecular bioSystems*, *6*(1), 126–33. http://doi.org/10.1039/b910945e

Holm, L., Kääriäinen, S., Rosenström, P., & Schenkel, A. (2008). Searching protein structure databases with DaliLite v.3. *Bioinformatics*, *24*(23), 2780–2781. http://doi.org/10.1093/bioinformatics/btn507

Holm, L., & Rosenström, P. (2010). Dali server: Conservation mapping in 3D. *Nucleic Acids Research*, *38*(SUPPL. 2), 545–549. http://doi.org/10.1093/nar/gkq366

Horn, A. H. C., & Sticht, H. (2015). Synthetic Protein Scaffolds Based on Peptide Motifs and Cognate Adaptor Domains for Improving Metabolic Productivity. *Frontiers in Bioengineering and Biotechnology*, *3*(November), 191. http://doi.org/10.3389/fbioe.2015.00191

Howarth, M. (2017). Smart superglue in streptococci? the proof is in the pulling. *Journal of Biological Chemistry*, *292*(21), 8998–8999. http://doi.org/10.1074/jbc.H117.777466

Höfgen, R. und Willmitzer, L. (1988). Storage of competent cells for agrobacterium transformation. Nucleic Acids Res. 1988;16(20):9877. doi: 10.1093/nar/16.20.9877.

Hu, X., Hu, H., Melvin, J. A., Clancy, K. W., McCafferty, D. G., und Yang, W. (2011). Autocatalytic Intramolecular Isopeptide Bond Formation in Gram-Positive Bacterial Pili: A QM/MM Simulation. *J Am Chem Soc*. 133(3): 478–485. doi:10.1021/ja107513t.

Jaenicke, R. (1996). Stability and folding of ultrastable proteins: eye lens crystallins and enzymes from thermophiles. *The FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *10*(1), 84–92.

Jia, F., Narasimhan, B., & Mallapragada, S. (2014). Materials-based strategies for multi-enzyme immobilization and co-localization: A review. *Biotechnology and Bioengineering*, *111*(2), 209–222. http://doi.org/10.1002/bit.25136

Jia, F., Narasimhan, B., & Mallapragada, S. K. (2013). Biomimetic Multienzyme Complexes Based on Nanoscale Platforms. *AIChE Journal*, *59*(2), 355–360. http://doi.org/10.1002/aic.13992

Jørgensen, K., Rasmussen, A. V., Morant, M., Nielsen, A. H., Bjarnholt, N., Zagrobelny, M., Bak, S., & Møller, B. L. (2005). Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current Opinion in Plant Biology*, *8*(3 SPEC. ISS.), 280–291. http://doi.org/10.1016/j.pbi.2005.03.014

Kang, H. J., & Baker, E. N. (2009). Intramolecular Isopeptide Bonds Give Thermodynamic and Proteolytic Stability to the Major Pilin Protein of Streptococcus pyogenes, *284*(31), 20729–20737. http://doi.org/10.1074/jbc.M109.014514

Kang, H. J., & Baker, E. N. (2011). Intramolecular isopeptide bonds: Protein crosslinks built

for stress? *Trends in Biochemical Sciences*, *36*(4), 229–237. http://doi.org/10.1016/j.tibs.2010.09.007

Kang, H. J., & Baker, E. N. (2012). Structure and assembly of Gram-positive bacterial pili: Unique covalent polymers. *Current Opinion in Structural Biology*, *22*(2), 200–207. http://doi.org/10.1016/j.sbi.2012.01.009

Kang, H. J., Coulibaly, F., Clow, F., Proft, T., & Baker, E. N. (2007). Stabilizing Isopeptide Bonds Revealed in Gram-Positive Bacterial Pilus Structure. *Science*, *318*, 1625–1628. http://doi.org/10.1126/science.1145806

Kang, H. J., Paterson, N. G., Kim, C. U., Middleditch, M., Chang, C., Ton-That, H., & Baker, E. N. (2014). A slow-forming isopeptide bond in the structure of the major pilin SpaD from Corynebacterium diphtheriae has implications for pilus assembly. *Acta Crystallographica Section D: Biological Crystallography*, *70*(5), 1190–1201. http://doi.org/10.1107/S1399004714001400

Kang, W., Liu, J., Wang, J., Nie, Y., Guo, Z., & Xia, J. (2014). Cascade Biocatalysis by Multienzyme − Nanoparticle Assemblies.

Kazlauskas, R. J., & Bornscheuer, U. T. (2009). Finding better protein engineering strategies. *Nature Chemical Biology*, *5*(8), 526–9. http://doi.org/10.1038/nchembio0809-526

Keeble, A. H., Banerjee, A., Reddington, S. C., Ferla, M. P., Howarth, M., & Khairil Anuar, I. N. A. (2017). Evolving accelerated amidation by SpyTag/SpyCatcher to analyze membrane dynamics. *Angewandte Chemie International Edition*. http://doi.org/10.1002/anie.201707623

Krishnan, V. (2015). Pilins in gram-positive bacteria: A structural perspective. *IUBMB Life*, *67*(7), 533–543. http://doi.org/10.1002/iub.1400

Krishnan, V., Dwivedi, P., Kim, B. J., Samal, A., Macon, K., Ma, X., Mishra, A., Doran, K. S., Ton-That, H., & Narayana, S. V. L. (2013). Structure of Streptococcus agalactiae tip pilin GBS104: A model for GBS pili assembly and host interactions. *Acta Crystallographica Section D: Biological Crystallography*, *69*(6), 1073–1089. http://doi.org/10.1107/S0907444913004642

Krishnan, V., Gaspar, A. H., Ye, N., Mandlik, A., Ton-That, H., Narayana, S. V. (2007). An IgG-like domain in the minor pilin GBS52 of Streptococcus agalactiae mediates lung epithelial cell adhesion. Structure. 2007; 15 (8):893±903. Epub 2007/08/19. https://doi.org/10.1016/j.str.2007.06.015 PMID: 17697995.

Kumar, S., & Nussinov, R. (2001). How do thermophilic proteins deal with heat? *Cellular and Molecular Life Sciences*, *58*(9), 1216–1233. https://doi.org/10.1007/PL00000935

Kwon, H., Squire, C. J., Young, P. G., & Baker, E. N. (2014). Autocatalytically generated Thr-Gln ester bond cross-links stabilize the repetitive Ig-domain shaft of a bacterial cell surface adhesin. *Proceedings of the National Academy of Sciences*, *111*(4), 1367–1372. http://doi.org/10.1073/pnas.1316855111

Kwon, H., Young, P. G., Squire, C. J., & Baker, E. N. (2017). Engineering a Lys-Asn isopeptide bond into an immunoglobulin-like protein domain enhances its stability. *Scientific Reports*, *7*(January), 42753. http://doi.org/10.1038/srep42753

Lauck, F., Smith, C. A., Friedland, G. F., Humphris, E. L., & Kortemme, T. (2010). RosettaBackrub-a web server for flexible backbone protein structure modeling and design. *Nucleic Acids Research*, *38*(SUPPL. 2), 569–575. http://doi.org/10.1093/nar/gkq369

Lee, H., DeLoache, W. C., & Dueber, J. E. (2012). Spatial organization of enzymes for

metabolic engineering. *Metabolic Engineering*, *14*(3), 242–251. https://doi.org/10.1016/j.ymben.2011.09.003

LeMieux, J., Hava, D. L., Basset, A., & Camilli, A. (2006). RrgA and RrgB are components of a multisubunit pilus encoded by the Streptococcus pneumoniae rlrA pathogenicity islet. *Infection and Immunity*, *74*(4), 2453–2456. https://doi.org/10.1128/IAI.74.4.2453-2456.2006

Levary, D. A., Parthasarathy, R., Boder, E. T., & Ackerman, M. E. (2011). Protein-Protein Fusion Catalyzed by Sortase A. *PLoS ONE*, *6*(4), 3–8. http://doi.org/10.1371/journal.pone.0018342

Li, L., Fierer, J. O., Rapoport, T. A., & Howarth, M. (2014). Structural analysis and optimization of the covalent association between SpyCatcher and a peptide tag. *Journal of Molecular Biology*, *426*(2), 309–317. http://doi.org/10.1016/j.jmb.2013.10.021

Linke, C., Young, P. G., Kang, H. J., Bunker, R. D., Middleditch, M. J., Caradoc-Davies, T. T., Proft, T., & Baker, E. N. (2010). Crystal structure of the minor pilin FctB reveals determinants of group A streptococcal pilus anchoring. *Journal of Biological Chemistry*, *285*(26), 20381–20389. http://doi.org/10.1074/jbc.M109.089680

Liu, D., Wu, W.-H., Liu, Y.-J., Wu, X.-L., Cao, Y., Song, B., Li, X., & Zhang, W.-B. (2017). Topology Engineering of Proteins in Vivo Using Genetically Encoded, Mechanically Interlocking SpyX Modules for Enhanced Stability. *ACS Central Science*, *3*(5), 473–481. http://doi.org/10.1021/acscentsci.7b00104

Liu, T., Wang, Y., Luo, X., Li, J., Reed, S. A., Xiao, H., Young, T. S., & Schultz, P. G. (2016). Enhancing protein stability with extended disulfide bonds. *Proceedings of the National Academy of Sciences of the United States of America*, 1605363113–. http://doi.org/10.1073/pnas.1605363113

Liu, Z., Zhou, H., Wang, W., Tan, W., Fu, Y.-X., & Zhu, M. (2015). A novel method for synthetic vaccine construction based on protein assembly. *Scientific Reports*, *4*(1), 7266. http://doi.org/10.1038/srep07266

Long, M. S., Jones, C. D., Helfrich, M. R., Mangeney-Slavin, L. K., & Keating, C. D. (2005). Dynamic microcompartmentation in synthetic cells. *Proc Natl Acad Sci U S A*, *102*(17), 5920–5925. http://doi.org/10.1073/pnas.0409333102

Lovell, S. C., Davis, I. W., Adrendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S., & Richardson, D. C. (2003). Structure validation by C alpha geometry: phi,psi and C beta deviation. *Proteins-Structure Function and Genetics*, *50*(August 2002), 437–450. http://doi.org/10.1002/prot.10286

Lumry, R. (2002). Protein substructures and folded stability. *Biophysical Chemistry*, *101-102*(March), 81–92. http://doi.org/10.1016/S0301-4622(02)00156-4

Matsunaga, R., Yanaka, S., Nagatoishi, S., & Tsumoto, K. (2013). Hyperthin nanochains composed of self-polymerizing protein shackles. *Nature Communications*, *4*(May), 1–10. http://doi.org/10.1038/ncomms3211

Matthews, B. W., Nicholson, H., & Becktel, W. J. (1987). Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(19), 6663–7. http://doi.org/10.1073/pnas.84.19.6663

Mian, I. S., Bradwell, A. R., & Olson, A. J. (1991). Structure, function and properties of antibody binding sites. *Journal of Molecular Biology*, *217*(1), 133–151. http://doi.org/10.1016/0022-2836(91)90617-F

Mirny, L. A., & Shakhnovich, E. I. (1999). Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function. *Journal of Molecular Biology*, *291*(1), 177–196. http://doi.org/10.1006/jmbi.1999.2911

Moon, T. S., Dueber, J. E., Shiue, E., & Prather, K. L. J. (2010). Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered E. coli. *Metabolic Engineering*, *12*(3), 298–305. http://doi.org/10.1016/j.ymben.2010.01.003

Moraïs, S., Barak, Y., Caspi, J., & Hadar, Y. (2010). Cellulase-xylanase synergy in designer cellulosomes for enhanced degradation of a complex cellulosic substrate. *MBio*, *1*(5), 3–10. http://doi.org/10.1128/mBio.00285-10.Editor

Moraïs, S., Morag, E., Barak, Y., Goldman, D., Hadar, Y., Lamed, R., Shoham, Y., Wilson, D. B., & Bayer, E. A. (2012). Deconstruction of lignocellulose into soluble sugars by native and designer cellulosomes. *mBio*, *3*(6), 1–11. http://doi.org/10.1128/mBio.00508-12

Morgan, A. A., & Rubenstein, E. (2013). Proline: The Distribution, Frequency, Positioning, and Common Functional Roles of Proline and Polyproline Sequences in the Human Proteome. *PLoS ONE*, *8*(1), 1–9. http://doi.org/10.1371/journal.pone.0053785

Myhrvold, C., Polka, J. K., & Silver, P. A. (2016). Synthetic Lipid-Containing Scaffolds Enhance Production by Colocalizing Enzymes. *ACS Synthetic Biology*, *5*(12), 1396–1403. http://doi.org/10.1021/acssynbio.6b00141

Nelson, A. L., Ries, J., Bagnoli, F., Dahlberg, S., Fälker, S., Rounioja, S., Tschöp, J., Morfeldt, E., Ferlenghi, I., Hilleringmann, M., Holden, D. W., Rappuoli, R., Normark, S., Barocchi, M. A., & Henriques-Normark, B. (2007). RrgA is a pilus-associated adhesin in Streptococcus pneumoniae. *Molecular Microbiology*, *66*(2), 329–340. https://doi.org/10.1111/j.1365-2958.2007.05908.x

Niklaus, L., Tansaz, S., Dakhil, H., Weber, K. T., Pröschel, M., Lang, M., Kostrzewa, M., Coto, P. B., Detsch, R., Sonnewald, U., Wierschem, A., Boccaccini, A. R., & Costa, R. D. (2017). Micropatterned Down-Converting Coating for White Bio-Hybrid Light-Emitting Diodes. *Advanced Functional Materials*, *27*(1), 1–10. http://doi.org/10.1002/adfm.201601792

Otaki, J.M., Tsutsumi, M., Gotoh, T., & Yamamoto, H. (2010). Secondary structure characterization based on amino acid composition and availability in proteins. *Journal of Chemical Information and Modeling*, *50*(4), 690–700. DOI: 10.1021/ci900452z

Ozawa, T. (2006). Designing split reporter proteins for analytical tools. *Analytica Chimica Acta*, *556*(1), 58–68. http://doi.org/10.1016/j.aca.2005.06.026

Pace, C. N., Grimsley, G. R., & Scholtz, J. M. (2009). Protein ionizable groups: pK values and their contribution to protein stability and solubility. *Journal of Biological Chemistry*, *284*(20), 13285–13289. http://doi.org/10.1074/jbc.R800080200

Pace, C. N., & Shirley, B. A. (1996). Forces contributing proteins of proteins. *Faseb*, *10*(1), 75–83.

Pace, C. N., Scholtz, J. M., & Grimsley, G. R. (2014). Forces stabilizing proteins. *FEBS Letters*, *588*(14), 2177–2184. http://doi.org/10.1016/j.febslet.2014.05.006

Paterson, N. G., & Baker, E. N. (2011). Structure of the full-length major pilin from Streptococcus pneumoniae: Implications for isopeptide bond formation in gram-positive bacterial pili. *PLoS ONE*, *6*(7). http://doi.org/10.1371/journal.pone.0022095

Pointon, J. A., Smith, W. D., Saalbach, G., Crow, A., Kehoe, M. A., & Banfield, M. J. (2010). A highly unusual thioester bond in a pilus adhesin is required for efficient host cell interaction. *Journal of Biological Chemistry*, *285*(44), 33858–33866.

http://doi.org/10.1074/jbc.M110.149385

Polka, J. K., Hays, S. G., & Silver, P. A. (2016). Building spatial synthetic biology with compartments, scaffolds, and communities. *Cold Spring Harbor Perspectives in Biology*, *8*(8). https://doi.org/10.1101/cshperspect.a024018

Ponnuswamy, P. K., und Gromiha, M. M. (1994). On the Conformational Stability of Folded Proteins. *J. theor. Biol.* 166(1), 63-74. https://doi.org/10.1006/jtbi.1994.1005

Potapov, V., Sobolev, V., Edelman, M., Kister, A., & Gelfand, I. (2004). Protein-protein recognition: Juxtaposition of domain and interface cores in immunoglobulins and other sandwich-like proteins. *Journal of Molecular Biology*, *342*(2), 665–679. http://doi.org/10.1016/j.jmb.2004.06.072

Pröschel, M., Detsch, R., Boccaccini, A. R., & Sonnewald, U. (2015). Engineering of Metabolic Pathways by Artificial Enzyme Channels. *Frontiers in Bioengineering and Biotechnology*, *3*(October), 168. http://doi.org/10.3389/fbioe.2015.00168

Pröschel, M., Kraner, M. E., Horn, A. H. C., Schäfer, L., Sonnewald, U., & Sticht, H. (2017). Probing the potential of CnaB-type domains for the design of tag/catcher systems. *PLoS ONE*, *12*(6), 1–26. http://doi.org/10.1371/journal.pone.0179740

The QIA*expressionist*™ 06/2003 A handbook for high-level expression and purification of 6xHis tagged proteins Fifth Edition (Juni 2003), Qiagen (Hilden)

Razvi, A., & Scholtz, J. M. (2006). Lessons in stability from thermophilic proteins. *Protein Science*, *15*(7), 1569–1578. http://doi.org/10.1110/ps.062130306

Reddington, S. C., & Howarth, M. (2015). Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher. *Current Opinion in Chemical Biology*, *29*, 94–99. http://doi.org/10.1016/j.cbpa.2015.10.002

Rosini, R., Rinaudo, C. D., Soriani, M., Lauer, P., Mora, M., Maione, D., Taddei, A., Santi, I., Ghezzo, C., Brettoni, C., Buccato, S., Margarit, I., Grandi, G., & Telford, J. L. (2006). Identification of novel genomic islands coding for antigenic pilus-like structures in Streptococcus agalactiae. *Molecular Microbiology*, *61*(1), 126–141. https://doi.org/10.1111/j.1365-2958.2006.05225.x

Rossi, E. A., Goldenberg, D. M., & Chang, C. H. (2012). Complex and defined biostructures with the dock-and-lock method. *Trends in Pharmacological Sciences*, *33*(9), 474–481. http://doi.org/10.1016/j.tips.2012.06.001

Russo, D. (2008). The impact of kosmotropes and chaotropes on bulk and hydration shell water dynamics in a model peptide solution. *Chemical Physics*, *345*(2-3), 200–211. http://doi.org/10.1016/j.chemphys.2007.08.001

Sambrook, J., Russell, D. (2001). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press 3: ISBN 978-0879695774

Schoene, C., Bennett, S. P., & Howarth, M. (2016). SpyRing interrogation: analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries. *Scientific Reports*, *6*(1), 21151. http://doi.org/10.1038/srep21151

Schoene, C., Fierer, J. O., Bennett, S. P., & Howarth, M. (2014). SpyTag/SpyCatcher cyclization confers resilience to boiling on a mesophilic enzyme. *Angewandte Chemie - International Edition*, *53*(24), 6101–6104. http://doi.org/10.1002/anie.201402519

Schwarz-Linek, U., & Banfield, M. J. (2014). Yet more intramolecular cross-links in Gram-positive surface proteins, *111*(4), 1073–1074. http://doi.org/10.1073/pnas.1322482111

Shaik, M. M., Maccagni, A., Tourcier, G., Di Guilmi, A. M., & Dessen, A. (2014). Structural basis of pilus anchoring by the ancillary pilin RrgC of Streptococcus pneumoniae. *Journal of Biological Chemistry*, *289*(24), 16988–16997. http://doi.org/10.1074/jbc.M114.555854

Shekhawat, S. S., und Ghosh, I. (2011). Split-Protein Systems: Beyond Binary Protein-Protein Interactions. *Curr Opin Chem Biol.*; 15(6): 789–797. doi:10.1016/j.cbpa.2011.10.014.

Siegmund, V., Piater, B., Zakeri, B., Eichhorn, T., Fischer, F., Deutsch, C., Becker, S., Toleikis, L., Hock, B., Betz, U. A. K., & Kolmar, H. (2016). Spontaneous Isopeptide Bond Formation as a Powerful Tool for Engineering Site-Specific Antibody-Drug Conjugates. *Scientific Reports*, *6*(1), 39291. http://doi.org/10.1038/srep39291

Smith, C. A., & Kortemme, T. (2008). Backrub-Like Backbone Simulation Recapitulates Natural Protein Conformational Variability and Improves Mutant Side-Chain Prediction. *Journal of Molecular Biology*, *380*(4), 742–756. http://doi.org/10.1016/j.jmb.2008.05.023

Smith, C. A., & Kortemme, T. (2010). Structure-based prediction of the peptide sequence space recognized by natural and synthetic PDZ domains. *Journal of Molecular Biology*, *402*(2), 460–474. http://doi.org/10.1016/j.jmb.2010.07.032

Smith, C. A., & Kortemme, T. (2011). Predicting the Tolerated Sequences for Proteins and Protein Interfaces Using RosettaBackrub Flexible Backbone Design. *PLoS ONE*, *6*(7). http://doi.org/10.1371/journal.pone.0020451

Sonnewald, U. (2003). Plant Biotechnology : From basic science to industrial applications, *725*, 723–725.

Stern, J., Kahn, A., Vazana, Y., Shamshoum, M., Moraïs, S., Lamed, R., & Bayer, E. A. (2015). Significance of Relative Position of Cellulases in Designer Cellulosomes for Optimized Cellulolysis. *PloS One*, *10*(5), e0127326. http://doi.org/10.1371/journal.pone.0127326

Stern, J., Moraïs, S., Lamed, R., & Bayer, E. A. (2016). Adaptor scaffoldins: An original strategy for extended designer cellulosomes, inspired from nature. *mBio*, *7*(2), 1–10. http://doi.org/10.1128/mBio.00083-16

Stryer, Biochemie 7. Auflage, Springer Spektrum, ISBN 978-3-8274-2988-9

Sun, F., Zhang, W.-B., Mahdavi, A., Arnold, F. H., & Tirrell, D. A. (2014). Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(31), 11269–74. http://doi.org/10.1073/pnas.1401291111

Sweetlove, L. J., & Fernie, A. R. (2013). The spatial organization of metabolism within the plant cell. *Annual Review of Plant Biology*, *64*, 723–46. https://doi.org/10.1146/annurev-arplant-050312-120233

Talley, K., & Alexov, E. (2011). On the pH-optimum of activity and stability of proteins. *Proteins*, *78*(12), 2699–2706. http://doi.org/10.1002/prot.22786.On

Tan, L. L., Hoon, S. S., & Wong, F. T. (2016). Kinetic controlled Tag-Catcher interactions for directed covalent protein assembly. *PLoS ONE*, *11*(10), 1–15. http://doi.org/10.1371/journal.pone.0165074

Thrane, S., Janitzek, C. M., Matondo, S., Resende, M., Gustavsson, T., de Jongh, W. A., Clemmensen, S., Roeffen, W., van de Vegte-Bolmer, M., van Gemert, G. J., Sauerwein, R., Schiller, J. T., Nielsen, M. A., Theander T. G., Salanti, A., & Sander, A. F. (2016). Bacterial superglue enables easy development of efficient virus-like particle based

vaccines. *Journal of Nanobiotechnology*, *14*(1), 30. http://doi.org/10.1186/s12951-016-0181-1

Tiana, G., Broglia, R. A., Roman, H. E., Vigezzi, E., und Shakhnovich, E. (1998). Folding and misfolding of designed proteinlike chains with mutations. The Journal of Chemical Physics 108, 757. https://doi.org/10.1063/1.475435

Ton-That, H., & Schneewind, O. (2004). Assembly of pili in Gram-positive bacteria. *Trends in Microbiology*, *12*(5), 228–234. http://doi.org/10.1016/j.tim.2004.03.004

Trueblood, K. N., Bürgi, H.-B., Burzlaff, H., Dunitz, J. D., Gramaccioli, C. M., Schulz, H. H., Shmueli, U., & Abrahams, C (1996). Atomic Dispacement Parameter Nomenclature. Report of a Subcommittee on Atomic Displacement Parameter Nomenclature. *Acta Crystallographica Section A*, *52*(5), 770–781. http://doi.org/10.1107/S0108767396005697

Van Larebeke, N., Engler, G., Holsters, M., Van den Elsacker, S., Zaenen, I., Schilperoort, R. A., Schell, J. (1974). Large plasmid in *Agrobacterium tumefaciens* essential for crown gallinducing ability. *Nature* 252(5479): 169-70

Veggiani, G., Nakamura, T., Brenner, M. D., Gayet, R. V., Yan, J., Robinson, C. V., & Howarth, M. (2016). Programmable polyproteams built using twin peptide superglues. *Proceedings of the National Academy of Sciences*, *113*(5), 1202–1207. http://doi.org/10.1073/pnas.1519214113

Veggiani, G., Zakeri, B., & Howarth, M. (2014). Superglue from bacteria: Unbreakable bridges for protein nanotechnology. *Trends in Biotechnology*, *32*(10), 506–512. http://doi.org/10.1016/j.tibtech.2014.08.001

Vengadesan, K., & Narayana, S. V. L. (2011). Structural biology of Gram-positive bacterial adhesins. *Protein Science*, *20*(5), 759–772. http://doi.org/10.1002/pro.613

Voinnet, O., Rivas, S., Mestre, P., Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant Journal* 33, 949-956

Walden, M., Edwards, J. M., Dziewulska, A. M., Bergmann, R., Saalbach, G., Kan, S.-Y., Miller, O. K., Weckener, M., Jackson, R. J., Shirran, S. L., Botting, C. H., Florence, G. J., Rohde, M., Banfield, M. J., Schwarz-Linek, U. (2015). An internal thioester in a pathogen surface protein mediates covalent host binding. *eLife*, *4*(JUNE), 1–24. http://doi.org/10.7554/eLife.06638

Walsh, G. (2007). Protein engineering: Case studies of commercialized engineered products. *Biochemistry and Molecular Biology Education*, *35*(1), 2–8. http://doi.org/10.1002/bmb.21

Wang, B., Xiao, S., Edwards, S. A., & Gra, F. (2013). Isopeptide Bonds Mechanically Stabilize Spy0128 in Bacterial Pili, *104*(May), 2051–2057. http://doi.org/10.1016/j.bpj.2013.04.002

Wang, J., Wang, Y., Wang, X., Zhang, D., Wu, S., & Zhang, G. (2016). Enhanced thermal stability of lichenase from Bacillus subtilis 168 by SpyTag/SpyCatcher-mediated spontaneous cyclization. *Biotechnology for Biofuels*, *9*(1), 79. http://doi.org/10.1186/s13068-016-0490-5

Watanabe, K., Masuda, T., Ohashi, H., Mihara, H., & Suzuki, Y. (1994). Multiple Proline Substitutions Cumulatively Thermostabilize Bacillus Cereus ATCC7064 Oligo-1,6-Glucosidase: Irrefragable Proof Supporting the Proline Rule. *European Journal of Biochemistry*, *226*(2), 277–283. http://doi.org/10.1111/j.1432-1033.1994.tb20051.x

Weber, A. P. M., & Bräutigam, A. (2013). The role of membrane transport in metabolic engineering of plant primary metabolism. *Current Opinion in Biotechnology*, *24*(2), 256–262. http://doi.org/10.1016/j.copbio.2012.09.010

Weber, M. D., Niklaus, L., Pröschel, M., Coto, P. B., Sonnewald, U., & Costa, R. D. (2015). Bioinspired Hybrid White Light-Emitting Diodes. *Advanced Materials*, *27*(37), 5493–5498. http://doi.org/10.1002/adma.201502349

Yan, C., Wu, F., Jernigan, R. L., Dobbs, D., & Honavar, V. (2008). Characterization of protein-protein interfaces. *Protein Journal*, *27*(1), 59–70. http://doi.org/10.1007/s10930-007-9108-x

Yancey, P. H. (2004). Compatible and counteracting solutes: protecting cells from the Dead Sea to the deep sea. *Science Progress*, *87*(Pt 1), 1–24. http://doi.org/10.3184/003685004783238599

Yang, H., Liu, L., & Xu, F. (2016). The promises and challenges of fusion constructs in protein biochemistry and enzymology. *Applied Microbiology and Biotechnology*, *100*(19), 8273–8281. http://doi.org/10.1007/s00253-016-7795-y

You, C., & Zhang, Y. H. P. (2013). Self-assembly of synthetic metabolons through synthetic protein scaffolds: One-step purification, co-immobilization, and substrate channeling. *ACS Synthetic Biology*, *2*(2), 102–110. http://doi.org/10.1021/sb300068g

Young, P. G., Yosaatmadja, Y., Harris, P. W. R., Leung, I. K. H., Baker, E. N., & Squire, C. J. (2017). Harnessing ester bond chemistry for protein ligation. *Chem. Commun.*, *53*(9), 1502–1505. http://doi.org/10.1039/C6CC09899A

Zakeri, B. (2015). Synthetic Biology: A New Tool for the Trade. *ChemBioChem*, *16*(16), 2277–2282. http://doi.org/10.1002/cbic.201500372

Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., & Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(12), E690–7. http://doi.org/10.1073/pnas.1115485109

Zakeri, B., & Howarth, M. (2010). Spontaneous Intermolecular Amide Bond Formation between Side Chains for Irreversible Peptide Targeting, 4526–4527.

Zhang, W. Bin, Sun, F., Tirrell, D. A., & Arnold, F. H. (2013). Controlling macromolecular topology with genetically encoded SpyTag-SpyCatcher chemistry. *Journal of the American Chemical Society*, *135*(37), 13988–13997. http://doi.org/10.1021/ja4076452

Zhang, Y. H. P. (2011). Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnology Advances*, *29*(6), 715–725. http://doi.org/10.1016/j.biotechadv.2011.05.020

Zou, Q., Bennion, B. J., Daggett, V., & Murphy, K. P. (2002). The molecular mechanism of stabilization of proteins by TMAO and its ability to counteract the effects of urea. *Journal of the American Chemical Society*, *124*(7), 1192–1202. http://doi.org/10.1021/ja004206b

Abkürzungsverzeichnis

% Prozent

°C Grad Celsius

µ mikro

AS Aminosäure (siehe Tabelle A3)

APS Ammoniumpersulfat

A. tumefaciens *Agrobacterium tumefaciens*

ATMT *Agrobacterium tumefaciens mediated transformation*

ATP Adenosintriphosphat

B. cereus *Bacillus cereus*

BioHLEDs *Bioinspired Hybrid Light-Emitting Diodes*

BLAST Basic Local Alignment Search Tool

bzw. beziehungsweise

ca. circa

CaMV *Cauliflower Mosaic Virus*

CD *circular dichroism*

cm Zentimeter

Cna Collagen adhesin

d.h. das heißt

DNA Desoxyribonukleinsäure

ECL *enhanced chemiluminescence*

E. coli *Escherichia coli*

et al. et alii

evtl. eventuell

FbaB *fibronectin binding protein*, Fibronektinbindeprotein

g Erdbeschleunigung g

g Gramm

GFP/ **eGFP** grün-fluoreszierendes Protein/ *enhanced* grün-fluoreszierendes Protein

h Stunden

HA Hämagglutinin

Ig Immunglobulin

INTACT *isolation of nuclei tagged in specific cell types*

IPTG Isopropyl-x-D-thiogalactopyranosid

K kDa

kDa Kilodalton

KLSM Konfokale(s) Laserscanning Mikroskopie/Mikroskop

l Liter

m milli

m Meter

M Molarität/ molar

mA Milliampere

MBP Maltosebindeprotein

MD Moleküldynamik

min Minute(n)

MSCRAMMs *microbial surface components recognizing adhesive matrix molecules*

MW *molecular weight*

MWCO *molecular weight cut off*

n nano

N. benthamiana *Nicotiana benthamiana*

Ni-NTA Nickel-Nitrilotriessigsäure

NMR *nuclear magnetic resonance*

OEP7 *outer envelope membrane protein 7*

oD600 optische Dichte bei 600 nm

PAGE Polyacrylamid-Gelektrophorese

PCR Polymerase-Kettenreaktion

PDB Proteindatenbank

PEO Polyethylenoxid

PROSS Protein Repair One Stop Shop

QM/MM *quantum mechanics/molecular mechanics*

rpm *revolutions per minute*

RFP rot-fluoreszierendes Protein

sec Sekunde(n)

SDS Natriumdodecylsulfat

S. pyogenes *Streptococcus pyogenes*

S. pneumoniae *Streptococcus pneumoniae*

TMAO Trimethylamin N-oxid

u.a. unter anderem

UV ultraviolett

v.a. vor allem

z.B. zum Beispiel

Anhang

Tabelle A1: DNA-Sequenzen:

Synthetische DNA-Fragmente	Sequenz 5´→3´
GSGSGSG-SpyCatcher #15	GGATCTGGTTCAGGCTCCGGAGAAGAAGATAGTGCTACCCATATT AAATTCTCAAAACGTGATGAGGACGGCAAAGAGTTAGCTGGTGCA ACTATGGAGTTGCGTGATTCATCTGGTAAAACTATTAGTACATGGA TTTCAGATGGACAAGTGAAAGATTTCTACCTGTATCCAGGAAAATA TACATTTGTCGAAACCGCAGCACCAGACGGTTATGAGGTAGCAAC TGCTATTACCTTTACAGTTAATGAGCAAGGTCAGGTTACTGTAAAT GGCAAAGCAACTAAAGGT
SpyTag-GSGSGSG-MBP #44	GGATCCGCGCATATTGTGATGGTGGATGCGTATAAACCGACCAAA GGATCTGGTTCAGGCTCCGGAAAAATCGAAGAAGGTAAACTGGTA ATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTC GGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAG CATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACT GGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGT GGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAA GCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGT TACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTAT CGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTG GGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAA GAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCC GCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGG CAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAA AGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATG AATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAG GCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAAC ATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACC TTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCA GGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAAGAGTTC CTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATA AAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAG AGTTGGCGAAAGATCCACGTATTGCCGCCACCATGGAAAACGCCC AGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCT GGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGT CAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAAGTCGAC
3phsCatcher-HA #52	CCTGGGATCCTCTACCGGCGAATTAGATTTACTGAAAGTGGGCGT GGATGGCGATACCAAAAAACCGCTGGCGGGCGTGGTGTTTGAACT GTATGAAAAAAACGGCCGTACCCCGATTCGTGTGAAAAACGGCGT GCATTCTCAGGATATTGATGCGGCGAAACATTTAGAAACCGATTCT TCTGGCCATATTCGTATTTCTGGCCTGATTCATGGCGATTATGTGT TAAAAGAAATTGAAACCCAGTCTGGCTATCAGATTGGCCAGGCGG AAACCGCGGTGACCATTGAATACCCTTATGATGTGCCAGACTACG CATAAGTCGACCCTG
3phsTag-	CCTGGGATCCAAAACCGTGACCGTGACCATTGAAAACAAAAAAGT GCCGACCCCGAAAGGATCTGGTTCAGGCTCCGGAAAAATCGAAGA AGGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGT

GSGSGSG- MBP #53	CTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAA GTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAG GTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACAC GACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATC ACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGG GATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCT GTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACC CGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGA AAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGT ACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCA AGTATGAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATA ACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTA AAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGC TGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTG GGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAAC GGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTGG CGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCT GGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCT GGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAA GTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCAC CATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCA GATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGC CGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGC AGACTTAAGTCGACCCTG
4oq1Catcher- HA #58	CCTGGGATCCACCATGACCACCAAAGTGAAACTGATTAAAGTGGA TCAGGATCATAACCGTTTAGAAGGCGTGGGCTTTAAATTAGTGTCT GTGGCGCGTGATGTGTCTGAAAAAGAAGTGCCGTTAATTGGCGAA TATCGTTATAGCTCTTCTGGCCAGGTGGGCCGTACCCTGTATACC GATAAAAACGGCGAAATTTTTGTGACCAACCTGCCGCTGGGCAAC TATCGTTTTAAAGAAGTGGAACCGCTGGCGGGCTATGCGGTGACC ACCCTGGATACCGATGTGCAGCTGTACCCTTATGATGTGCCAGAC TACGCATAAGTCGACCCTG
4oq1Tag- GSGSGSG- MBP #59	CCTGGGATCCGTGACCATTACCGTGGTGAACCAGAAATTACCGCG TGGCAACGGATCTGGTTCAGGCTCCGGAAAAATCGAAGAAGGTAA ACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCT GAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACC GTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCG GCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGC TTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCG GACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAA GCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAA AAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGA AAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATG AAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTG GCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACA AACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATG GTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACT GCCGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCT GAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAA AGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGC GGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTA CGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACCATGGA

	AAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTC CGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAG CGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTA AGTCGACCCTG
3kptCCatcher- HA #54	CCTGGGATCCACCACCGGCATTATTGAATTAACCAAAATTGATAGC GCGAACAAAAACAAATTAAAAGGCGCGGAATTTGTGCTGAAAGATA ACAACGGCAAAATTGTGGTGGTGGCGGGCAAAGAAGTGACCGGC GTGTCTGATGAAAACGGCGTGATTAAATGGTCTAACATTCCGTATG GCGATTATCAGATTTTTGAAACCAAAGCGCCGACCTATACCAAAGA AGATGGCACCAAAACCTCTTATCAGTTATTAAAAGATCCGATTGAT GTGAAAATTAGCTACCCTTATGATGTGCCAGACTACGCATAAGTCG ACCCTG
3kptCTag- GSGSGSG- MBP #55	CCTGGGATCCACCGTGAAATTAACCATTGAAAACAACAAAAGCCC GACCAAAGGATCTGGTTCAGGCTCCGGAAAAATCGAAGAAGGTAA ACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCT GAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACC GTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCG GCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGC TTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCG GACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAA GCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAA AAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGA AAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATG AAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTG GCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACA AACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATG GTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACT GCCGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCT GAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAA AGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGC GGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTA CGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACCATGGA AAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTC CGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAG CGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTA AGTCGACCCTG
3kptNCatcher- HA #56	CCTGGGATCCGAAATTAAACGTGGCGCGGTGGATTTAATTAAAAC CGGCGTGAACGAAAAAGCGTTAGCGGGCGCGGTGTTTTCTTTATT TAAAAAAGATGGCACCGAAGTGAAAAAAGAATTAGCGACCGATGC GAACGGCCATATTCGTGTGCAGGGCTTAGAATATGGCGAATATTAT TTTCAGGAAACCAAAGCGCCGAAAGGCTATGTGATTGATCCGACC AAACGTGAATTTTTTGTGAAAAACTCTGGCACCATTAACGAAGATG GCACCATTACCTCTTACCCTTATGATGTGCCAGACTACGCATAAGT CGACCCTG
3kptNTag- GSGSGSG- MBP #57	CCTGGGATCCGTGGTGAAAATTGAAGTGAAAAACAACGAAGAACC GACCATTGATGGATCTGGTTCAGGCTCCGGAAAAATCGAAGAAGG TAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTC GCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCA CCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTG CGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACC GCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCC CGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATG

	CCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTG AAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCC AAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGC GAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTT CACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTA TGAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATAACGC TGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAA CAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGC CTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGC ATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTA CTGCCGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTG CTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCA AAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAG CGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTT ACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACCATGG AAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGT CCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCA GCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTT AAGTCGACCCTG
4oq1Catcher PROSS #148	CCTGGGATCCACCATGACCACCAAAGTGAAACTGATTAAAGTGGA TCAGGATGGCAACCGTTTAGAAGGCGTGGGCTTTAAATTAGTGCG CGTGGCGCGTGATGTGTCTGAAAAAGAAGTGCCGTTAATTGGCGA ATATCGTTATAGCTCTTCTGGCCAGGTGGGCCGTACCCTGTATACC GATAAAAACGGCGAAATTTTTGTGACCAACCTGCCGCTGGGCAAC TATCGTTTTCGCGAAGTGGAACCGCTGGAAGGCTATCAGGTGACC ACCGAAGATACCGATGTGACCCTGGTGGATCATCAGTAAGTCGAC CCTG
3kptCCatcher PROSS #149	CCTGGGATCCACCACCGGCATTATTGAATTAACCAAAATTGATAGC GCGAACAAAAACAAATTAAAAGGCGCGGAATTTGTGCTGAAAGATA ACAACGGCAAAATTGTGGTGGTGGCGGGCAAAGAAGTGACCGGC GTGTCTGATGAAAACGGCGTGATTAAATGGACCAACATTCCGTATG GCGATTATCAGATTTTTGAAACCAAAGCGCCGACCTATAAAAAAGA AGATGGCACCACCGCGTCTTATCAGTTATTAAAAGATCCGATTGAT GTGAAAATTACCTAAGTCGACCCTG
Snoop-System DNA-Sequenzen (Prof. Mark Howarth, Veggiani et al., 2016)	
SnoopTag	AAACTGGGCGATATTGAATTTATTAAAGTGAACAAA
SnoopCatcher	AAGCCGCTGCGTGGTGCCGTGTTTAGCCTGCAGAAACAGCATCCC GACTATCCCGATATCTATGGCGCGATTGATCAGAATGGGACCTATC AAAATGTGCGTACCGGCGAAGATGGTAAACTGACCTTTAAGAATCT GAGCGATGGCAAATATCGCCTGTTTGAAAATAGCGAACCCGCTGG CTATAAACCGGTGCAGAATAAGCCGATTGTGGCGTTTCAGATTGTG AATGGCGAAGTGCGTGATGTGACCAGCATTGTGCCGCAGGATATT CCGGCTACATATGAATTTACCAACGGTAAACATTATATCACCAATG AACCGATACCGCCGAAA
Fluoreszenzproteine	
eGFP	ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCAT CCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCG TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACC CTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCC ACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGC

	TACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATG CCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGAC GGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACAC CCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCC ACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGG TGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGC TCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCC GTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTG AGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGA GTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT ACAAGTAA
mCherry	ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGA GTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCA CGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAG GGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCT GCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTC CAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAA GCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTT CGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGC AGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACT TCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGG GAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAA GGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACT ACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTG CAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACC TCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCC GAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTA A
mTagBFP	ATGAGCGAGGAGCTGATCAAAGAGAACATGCACATGAAGCTGTAC ATGGAAGGCACCGTTGACAACCACCACTTTAAGTGCACGTCTGAG GGTGAGGGTAAGCCGTACGAAGGCACCCAAACCATGCGTATCAAA GTTGTGGAGGGCGGTCCACTGCCGTTCGCTTTTGACATTCTGGCG ACCAGCTTCCTGTACGGTTCCAAAACGTTCATTAACCATACTCAGG GCATTCCGGATTTCTTCAAACAGAGCTTTCCGGAAGGTTTCACCTG GGAGCGTGTCACCACGTATGAAGATGGTGGTGTGTTGACCGCCAC CCAAGATACCTCCCTGCAAGATGGCTGTCTGATCTATAACGTGAAA ATTCGTGGCGTCAACTTTACGAGCAATGGTCCGGTGATGCAGAAG AAAACCCTGGGTTGGGAGGCGTTTACGGAAACCCTGTATCCGGCC GATGGTGGCCTGGAGGGCCGTAACGACATGGCACTGAAGCTGGT TGGTGGCAGCCATTTGATCGCAAATATCAAGACGACGTACCGCAG CAAGAAACCGGCGAAAAATCTGAAGATGCCGGGTGTTTACTATGT CGACTACCGTCTGGAACGCATTAAAGAAGCGAATAATGAGACTTAC GTGGAGCAGCACGAGGTTGCAGTCGCGCGCTATTGCGACTTGCC TAGCAAGCTGGGTCATAAACTGAATTAA

Tabelle A2: Primersequenzen

Name	Sequenz 5´→3´
MP32	CAGGGTCGACTGCGTAGTCTGGCACATCATAAGG
MP33	GTCGACGTGAGCAAGGGCGAGGAGCTGTTC
MP42	GGATCCAACAATGGGAAAAACTTCTGGAGCTAAGC
MP44	GGATCCGTGAGCAAGGGCGAGGAGGATAAC
MP52	GGATCCAGCGAGGAGCTGATCAAAGAG
MP55	CTGCAGTTATGCGTAGTCTGGCACATCATAAGGGTAACCTTTAGTTGCTTTG CCATTTAC
MP60	GTCGACTTATGCGTAGTCTGGCACATCATAAGGGTAACCTTTAGTTGCTTTG CCATTTAC
MP63	GTCGACGTGAGCAAGGGCGAGGAGCTGTTC
MP64	CTTCTCCGGAGCCTGAACCAGATCCCTTGTACAGCTCGTCCATGCCGAGAG
MP65	CCGGCGGCATGGACGAGCTGTACAAGGGATCTGGTTCAGGCTCCGGAG AAG
MP66	CTTCTCCGGAGCCTGAACCAGATCCCTTGTACAGCTCGTCCATGCCGC CGG
MP67	CTAGCAAGCTGGGTCATAAACTGAATGGATCTGGTTCAGGCTCCGGAGAAG
MP68	CTTCTCCGGAGCCTGAACCAGATCCATTCAGTTTATGACCCAGCTTGCTAG
MP76	CCTGGGATCCGCGCATATTGTGATGGTGGATGCGTATAAACCGACCAAAG GATCTGGTTCAGGCTCCGGAGTGAGCAAGGGCGAGGAGCTGTTC
MP77	CAGGGTCGACTTATTTGGTCGGTTTATACGCATCCACCATCACAATATGCG CTCCGGAGCCTGAACCAGATCCCTTG
MP78	CCTGGGATCCGTGAGCAAGGGCGAGGAGCTGTTC
MP79	GGTTAATTCAATAATGCCGGTGGTTCCGGAGCCTGAACCAGATCCCTTG
MP80	CAAGGGATCTGGTTCAGGCTCCGGAACCACCGGCATTATTGAATTAACC
MP81	CAGGGTCGACTTATGCGTAGTCTGGCACATCATAAGG
MP82	CAGGGTCGACTTAACCTTTAGTTGCTTTGCCATTTACAG
MP85	GGATCTGGTTCAGGCTCCGGAAAGCCGCTGCGTGGTGCCGTG
MP86	CACGGCACCACGCAGCGGCTTTCCGGAGCCTGAACCAGATCC

MP87	CCTGGGATCCGTGAGCAAGGGCGAGGAGGATAAC
MP88	CAGGGTCGACTTATTTCGGCGGTATCGGTTCATTGGTG
MP89	CAGGGTCGACTTATTTGTTCACTTTAATAAATTCAATATCGCCCAGTTTTCC GGAGCCTGAACCAGATCCCTTGTACAGCTCGTCCATGCC
MP90	GGATCTGGTTCAGGCTCCGGAGAAATTAAACGTGGCGCGGTG
MP91	CACCGCGCCACGTTTAATTTCTCCGGAGCCTGAACCAGATCC
MP92	CAGGGTCGACTTAAGAGGTAATGGTGCCATCTTC
MP93	GGATCTGGTTCAGGCTCCGGAACCATGACCACCAAAGTGAAAC
MP94	GTTTCACTTTGGTGGTCATGGTTCCGGAGCCTGAACCAGATCC
MP95	CAGGGTCGACTTACAGCTGCACATCGGTATCCAG
MP96	GGATCTGGTTCAGGCTCCGGATCTACCGGCGAATTAGATTTAC
MP97	GTAAATCTAATTCGCCGGTAGATCCGGAGCCTGAACCAGATCC
MP98	CAGGGTCGACTTATTCAATGGTCACCGCGGTTTC
MP101	CCTGGGATCCGAAGAAGATAGTGCTACCCAT
MP106	CCTGGGATCCGTGACCATTACCGTGGTGAACCAGAAATTACCGCGTGGCA ACGGTAGTGGTGAAAGTGGTAAAATCGAAGAAGGTAAACTGG
MP107	CAGGGTCGACTTACGAGCTCGAATTAGTCTGCGCGTCTTTCAGGGCTTC ATC
MP108	CCTGGGATCCCATCAGCTGGTGACCATTACCGTGGTGAACCAG
MP109	CCTGGGATCCGTGACCATTACCGTGGTGAACCAGAAAGGTAGTGGTGAAA GTGGTAAAATC
MP110	CCTGGGATCCACCATGACCACC
MP111	CAGGGTCGACTTACTGATGATCCACCAGCTGCACATCGGTATCCAGGGTG GTC
MP112	CCTGGGATCCCTGGTGACCATTACCGTGGTGAACC
MP113	CCTGGGATCCGTGACCATTACCGTGGTGAACCAGAAATTACCGGGTAGTG GTGAAAGTGGTAAAATC
MP114	CAGGGTCGACTTACACATCGGTATCCAGGGTGGTC
MP115	CCTGGGATCCGTGACCATTACCGTGGTGGCGCAGAAATTACCGCGTGGC AAC

MP116	CCTGGGATCCACCATGACCACCAAAGTGAAACTGATTGCGGTGGATCAGG ATCATAACCG
MP121	CCTGGGATCCAAACTGGGCGATATTGAATTTATTAAAGTGAACAAAGGTAGT GGTGAAAGTGGTAAAATC
MP122	CCTGGGATCCAAGCCGCTGCGTGGTGCCGTGTTTAG
MP140	CAGGGTCGACTTACAGCTGCACATCGGTATCCAGGGTGGTCACCGCATAG CCCGCCAGCGGTTCCACCTGTTTAAAACG
MP144	GGATCTGGTTCAGGCTCCGGAACCACCGGCATTATTGAATTAAC
MP145	GTTAATTCAATAATGCCGGTGGTTCCGGAGCCTGAACCAGATCC
MP146	CAGGGTCGACTTAGCTAATTTTCACATCAATCG
MP156	CCTGGGATCCGTGACCATTACCGTGGTGGATCAGAAATTACCGCGTGGC AAC
MP157	CCTGGGATCCACCGTGAAATTAACCATTGAAAACAACAAAAGCCCGACCAA AGGTAGTGGTGAAAGTGGTAAAATCG
MP158	CCTGGGATCCAACCAGACCGTGAAATTAACCATTGAAAAC
MP159	CCTGGGATCCACCGTGAAATTAACCATTGAAAACAACAAAAGCGGCTGGAT TGGTAGTGGTGAAAGTGGTAAAATC
MP160	CTCAAGTCAGAGGTGGCGAAAC
MP161	GTTTCGCCACCTCTGACTTGAG
MP165	CCTGGGATCCCTGGATACCGATGTGCAGCTGGTGGATCATCAGCTGGTGA CCATTACCGTG
MP168	CCTGGGATCCACCGTGAAATTAACCATTGAAGCGAACAAAAGCCCGACCAA AGGTAG
MP169	CCTGGGATCCACCACCGGCATTATTGAATTAAC
MP171	GTATGGCGATTATCAGATTTTTCAGACCAAAGCGCCGACCTATAC
MP172	GTATAGGTCGGCGCTTTGGTCTGAAAAATCTGATAATCGCCATAC
MP173	CTTTAAATTAGTGTCTGTGGCGCGTGGCGTGTCTGAAAAAGAAGTGCCG
MP174	CGGCACTTCTTTTTCAGACACGCCACGCGCCACAGACACTAATTTAAAG
MP175	CCTGGGATCCACCCTGACCACCAAAGTGAAACTGATTAAAG
MP176	CGGCGAAATTTTTGTGACCGGCCTGCCGCTGGGCAACTATCG
MP177	CGATAGTTGCCCAGCGGCAGGCCGGTCACAAAAATTTCGCCG

MP178	GGCAACTATCGTTTTAAAGAAACCGAACCGCTGGCGGGCTATGCG
MP179	CGCATAGCCCGCCAGCGGTTCGGTTTCTTTAAAACGATAGTTGCC
MP180	GCGGGCTATGCGGTGACCACCAACGATACCGATGTGCAGCTGGTG
MP181	CACCAGCTGCACATCGGTATCGTTGGTGGTCACCGCATAGCCCGC
MP182	GATTAAAGTGGATCAGGATCATCCGCGTTTAGAAGGCGTGGGC
MP183	GCCCACGCCTTCTAAACGCGGATGATCCTGATCCACTTTAATC
MP184	GAACCGCTGGCGGGCTATGCGCCGACCACCCTGGATACCGATGTG
MP185	CACATCGGTATCCAGGGTGGTCGGCGCATAGCCCGCCAGCGGTTC
MP186	CAGGGTCGACTTACTGATGATCCACCAGCTGCACATCGGTATCCGGGGTG GTCACCGCATAGCCCGC
MP187	GTCGACTTATGCGTAGTCTGGCACATCATAAGGGTATTTGGTCGGTTTATAC GCATCCACCATCACAATATGCGCTCCGGAGCCTGAACCAGATCC
MP188	GTCGACTTATGCGTAGTCTGGCACATCATAAGGGTATTTGTTCACTTTAATA AATTCAATATCGCCCAGTTTTCCGGAGCCTGAACCAGATCC
MP189	GTCGACTTATGCGTAGTCTGGCACATCATAAGGGTATTTGGTCGGGCTTTT GTTGTTTTCAATGGTTAATTTCACGGTTCCGGAGCCTGAACCAGATCC
MP190	GTCGACTTATGCGTAGTCTGGCACATCATAAGGGTAGTTGCCACGCGGTAA TTTCTGGTTCACCACGGTAATGGTCACTCCGGAGCCTGAACCAGATCC
MP194	GGATCCAACAATGGTGAGCAAGGGCGAGGAGCTG
MP195	GGATCCAACAATGGTGAGCAAGGGCGAGGAGGATAAC
MP196	GTCGACTTATGCGTAGTCTGGCACATCATAAGGGTATTTCGGCGGTATCGG TTCATTGGTG
MP197	CACGGCACCACGCAGCGGCTTTCCGGAGCCTGAACCAGATCC
MP198	GGATCTGGTTCAGGCTCCGGAAAGCCGCTGCGTGGTGCCGTG
MP199	GTCGACTTATGCGTAGTCTGGCACATCATAAGGGTACAGCTGCACATCGGT ATCCAG
MP200	CCTGGGATCCTGCGGCAGCGAAAACCTGTATTTTCAGGGTTCTGAAGAAGA TAGTGCTACCCATATTAAATTC
MP201	CCTGGGATCCACCACCGGCATTATTGAATTAACCGCGATTGATAGCGCGAA CAAAAAC
MP202	CCTGGGATCCACCGTGAAATTAACCATTGAAGATAACAAAAGCCCGACCAA AGG

MP203	CAAAGCGCCGACCTATACCAAAAGCGATGGCACCAAAACCTCTTATC
MP204	GATAAGAGGTTTTGGTGCCATCGCTTTTGGTATAGGTCGGCGCTTTG
MP205	CGCCGACCTATACCAAAGAAAACGGCACCAAAACCTCTTATCAG
MP206	CTGATAAGAGGTTTTGGTGCCGTTTTCTTTGGTATAGGTCGGCG
MP207	CCTGGGATCCACCACCGGCATTATTGAATTAACCAAAATTGATAGCGCGCC GAAAAACAAATTAAAAGGCG
MP208	CCTGGGATCCACCACCGGCATTATTGAATTAACCAAAATTGATAGCGCGAA CAAAAACCCGTTAAAAGGCGCGGAATTTG
MP209	GAAGTGACCGGCGTGTCTGATCCGAACGGCGTGATTAAATGGTC
MP210	GACCATTTAATCACGCCGTTCGGATCAGACACGCCGGTCACTTC
MP211	CAGGGTCGACTTAGCTAATTTTCACATCAATCGGATCTTTTAATAACTGCGG AGAGGTTTTGGTGCCATC
MP212	CCGTATGGCGATTATCAGATTTATGAAACCAAAGCGCCGACC
MP213	GGTCGGCGCTTTGGTTTCATAAATCTGATAATCGCCATACGG
MP214	CTGAAAGATAACAACGGCAAAAAAGTGGTGGTGGCGGGCAAAG
MP215	CTTTGCCCGCCACCACCACTTTTTTGCCGTTGTTATCTTTCAG
MP216	CCTGGGATCCACCACCGGCATTATTGAATTAACCAAAAAAGATAGCGCGAA CAAAAAC
Ox21	GTTATAGCTCTTCTGGCCAGAACGGCCGTACCCTGTATAC
Ox22	GTATACAGGGTACGGCCGTTCTGGCCAGAAGAGCTATAAC
Ox25	GTGACCAACCTGCCGAAAGGCAACTATCGTTTTAAAG
Ox26	CTTTAAAACGATAGTTGCCTTTCGGCAGGTTGGTCAC
Ox29	GTCTGAAAAAGAAGTGCCGTTAGAAGGCGAATATCGTTATAGC
Ox30	GCTATAACGATATTCGCCTTCTAACGGCACTTCTTTTTCAGAC
Ox31	CCACCAAAGTGAAACTGACCAAAGTGGATCAGGATC
Ox32	GATCCTGATCCACTTTGGTCAGTTTCACTTTGGTGG
Ox33	GTGTCTGTGGCGCGTGATAACTCTGAAAAAGAAGTGCCG
Ox34	CGGCACTTCTTTTTCAGAGTTATCACGCGCCACAGACAC
Ox35	CGTTTTAAAGAAGTGGAACCGAAAGCGGGCTATGCGGTG

Ox36	CACCGCATAGCCCGCTTTCGGTTCCACTTCTTTAAAACG
Ox37	GATAAAAACGGCGAAATTTATGTGACCAACCTGCCGC
Ox38	GCGGCAGGTTGGTCACATAAATTTCGCCGTTTTTATC
Ox39	GAACCGCTGGCGGGCAGCGCGGTGACCACCCTG
Ox40	CAGGGTGGTCACCGCGCTGCCCGCCAGCGGTTC

Tabelle A3: Die 20 Aminosäuren und ihre wichtigsten Eigenschaften (Stryer).

Aminosäure (AS)	**Wichtige Eigenschaften der Seitenkette**
Alanin (A, Ala)	unpolar (hydrophob), klein, nicht reaktiv, Präferenz für α-Helices
Glycin (G, Gly)	unpolar, kleinste und einfachste Aminosäure, achiral (R: H), strukturell wichtige Aminosäure: flexible Konformation, keine bestimmte Sekundärstrukturelement-Präferenz (passt in alle Strukturen), Präferenz für Kehren
Valin (V, Val)	unpolar, aliphatisch, nicht reaktiv, Verzweigung am β-C-Atom kann aufgrund sterischer Behinderung α–Helices destabilisieren, Präferenz für β-Stränge
Leucin (L, Leu)	unpolar, aliphatisch, nicht reaktiv, Präferenz für α-Helices
Isoleucin (I, Ile)	unpolar, aliphatisch, nicht reaktiv, zusätzliches Chiralitätszentrum, Verzweigung am β-C-Atom kann aufgrund sterischer Behinderung α–Helices destabilisieren, Präferenz für β-Stränge
Prolin (P, Pro)	unpolar, strukturell wichtige Aminosäure, die die Konformation eines Proteins in hohem Maße beeinflusst (Gegenteil von Glycin): Seitenkette ist sowohl mit dem Cα-Atom als auch mit dem Stickstoffatom verbunden (freie NH-Gruppe fehlt). Diese Ringstruktur schränkt Konformation ein → fixierter, begrenzter phi-Winkel von 60° → Sekundärstrukturelement-Zerstörer (Prolin unterbricht α-Helices und β-Faltblätter), Peptidbindung liegt häufig in cis-Konfiguration vor, Präferenz für Kehren

Phenylalanin (F, Phe)	unpolar, aromatisch
Methionin (M, Met)	unpolar, aliphatisch, enthält Thioethergruppe (-S-)
Tryptophan (W, Trp)	unpolar, aromatisch
Cystein (C, Cys)	unpolar, klein, zwei Cysteinreste können unter oxidativen Bedingungen kovalente Disulfidbrücken ausbilden → essentielle Funktion bei Proteinstabilisierung, enthält reaktionsfreudige Thiolgruppe (-SH), pK_S-Wert: 8,3
Asparagin (N, Asn)	ungeladen polar (hydrophil), stört tendenziell die α-Helix (Seitenketten-Gruppe konkurriert mit CO- und NH-Gruppen der Hauptkette um Wasserstoffbrücken), Präferenz für Kehren
Glutamin (Q, Gln)	ungeladen polar
Serin (S, Ser)	ungeladen polar, klein, hydroxylierte Version von Alanin (reaktiver als hydrophobe Gegenstücke), stört tendenziell die α-Helix (Seitenketten-Gruppe konkurriert mit CO- und NH-Gruppen der Hauptkette um Wasserstoffbrücken)
Threonin (T, Thr)	ungeladen polar, verzweigt, hydroxylierte Version von Valin, zusätzliches Assymetriezentrum, Verzweigung am β-C-Atom kann aufgrund sterischer Behinderung α–Helices destabilisieren, Präferenz für β-Faltblätter
Tyrosin (Y, Tyr)	ungeladen polar, aromatisch, hydroxyliertes Derivat von Phenylalanin, pK_S-Wert: 10,9
Asparaginsäure (D, Asp) oder Aspartat	bei physiologischem pH-Wert gewöhnlich deprotoniert also negativ geladen (saure Seitenkette), pK_S-Wert: 4,1, stört tendenziell die α-Helix (Seitenketten-Gruppe konkurriert mit CO- und NH-Gruppen der Hauptkette um Wasserstoffbrücken)
Glutaminsäure (E, Glu) oder Glutamat	bei physiologischem pH-Wert gewöhnlich deprotoniert also negativ geladen (saure Seitenkette), pK_S-Wert: 4,1, Präferenz für α-Helices

Arginin (R, Arg)	bei neutralem pH-Wert positiv geladen (sehr hydrophil), lang, pK_S-Wert: 12,5
Lysin (K, Lys)	bei neutralem pH-Wert positiv geladen, lang, primäre Aminogruppe am Ende des Restes, pK_S-Wert: 10,8
Histidin (His, H)	enthält Imidazolgruppe (aromatische Ringstruktur), die einen pK_S-Wert von 6 besitzt und je nach lokaler Umgebung entweder positiv geladen oder ungeladen vorliegt. Die Imidazolgruppe kann also in der Nähe des physiologischen pH-Wertes seinen Ionisierungsgrad ändern und sowohl Protonen binden als auch freisetzen (Protonen-*shuttle*, Protonenrelais)