Extending the scope of covalent peptide-protein conjugation: from purification to expanded nanoassembly



Irsyad Noor Abadi Khairil Anuar

Department of Biochemistry
St John's College
University of Oxford

Michaelmas Term 2019

This thesis is submitted to the Department of Biochemistry, University of Oxford, in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Abstract

The surprising discovery of intramolecular isopeptide bonds within Gram-positive bacteria pili and surface adhesins gave rise to the development of a new type of protein conjugation method, namely the Tag/Catcher system. Among the different kinds of Tag/Catcher platforms, SpyTag/SpyCatcher is the leading system with more than 200 articles utilising the technology due to the rapid and buffer-independent nature of the reaction. In this thesis, I have developed a purification method called Spy&Go to complement the SpyTag/SpyCatcher platform. Spy&Go enables the purification of proteins using SpyTag as the affinity peptide tag. Spy&Go thereby eliminates the dependence on extraneous tags for purification of proteins intended for SpyCatcher-based protein conjugation reactions. Spy&Go was developed by engineering SpyCatcher2.1, a fast reacting variant of SpyCatcher, to become SpyDock by removing reactivity whilst maintaining tight binding to SpyTag. The purity of proteins purified through Spy&Go was comparable to the purity of proteins purified by Ni-NTA using a polyhistidine tag. Subsequently, a different route for elution of SpyTag-proteins was explored by conferring pH-sensitivity to a SpyDock variant through the introduction of histidine residues at the SpyTag003-binding interface. With the combined effects of three histidine mutations, SpySwitch was developed, which can bind SpyTag003 at pH 8.0 and release at pH 5.0. SpyTag/SpyCatcher has been used to modularly multimerise proteins for vaccination, nanomedicine and making enzymatic nanoreactors. To extend the scope of SpyCatcher-based protein multimerisation platforms, I developed a Spy oligomerisation toolbox composed of SpyCatcher002 fused to coiled coils, to multimerise SpyTag-proteins into dimer, trimer, tetramer, pentamer, hexamer and heptamer assemblies. The SpyCatcher002-oligomers enable a rapid and scalable route to multimerise different proteins into defined assemblies. An agonistic nanobody against Death Receptor 5 was multimerised through SpyCatcher002-oligomers and showed a stoichiometry-dependent effect on breast cancer cell line apoptosis upon receptor activation.

Authorship Declaration

I declare that this thesis is entirely my own work, except for cases described below. This thesis has not been submitted for any other degree at the University of Oxford or any other institution.

Experiments performed by others:

Chapter 3:

Isothermal titration calorimetry was performed by Dr Anthony Keeble.

Mass spectrometry was performed by Dr Anthony Tumber.

Chapter 5:

Size exclusion chromatography–multi-angle light scattering was performed by Dr David Staunton.

Cell killing assay was performed by Dr Anusuya Banerjee.

Publications arising from this thesis are detailed below:

Khairil Anuar, I.N.A., Banerjee, A., Keeble, A.H., Carella, A., Nikov, G.I., and Howarth, M. (2019). Spy&Go purification of SpyTag-proteins using pseudo-SpyCatcher to access an oligomerization toolbox. *Nature Communications*, *10*(1), 1734.

Acknowledgements

I would like to thank my supervisor, Professor Mark Howarth, for accepting me into his lab and having the patience and persistence to guide and teach me on how to be a good scientist from day one. His creativity and entrepreneurial endeavours still amaze me, and I wish to emulate those characteristics in the future. I would also like to thank my funding bodies, Yayasan Khazanah and Oxford Centre for Islamic Studies, and my college, St John's College, for their financial support that enabled me to pursue my DPhil in the University of Oxford.

To the past and current members of the Sherratt, Whitby and Uphoff labs – thank you very much for making the lab very accepting and lively for me on a daily basis. I could not have asked to share the lab area with any better teams.

To the Howarth lab members, past and present, I thank you with all my heart for creating an atmosphere that encouraged scientific discussion and interesting informal conversations. Over time, you have become my extended family and I have learnt so much from our times together. The times I have spent with all of you in and out of the lab will be held dearly in my memories. There are too many people to thank individually, which may take too much of the space here, but do know that I appreciate every single one of you immensely and hope that our paths will cross again.

To Tofu and Tomo, your presence at home made a long day more manageable. Thank you for always making me smile with your playfulness. I miss you two very much.

To my family, particularly my parents, thank you for believing in me and always encouraging me to strive for the best. You have been the best role models any son could have asked for and I still look to you for inspiration.

To my best friend, my wife and my soulmate, Mariana – without you, there is no doubt I will not be where I am today. You have been with me through thick and thin, always welcoming and always patient. Thank you for always being by my side no matter the circumstance. Here's to the rest of our lives together.

Abbreviations

AIM	Autoinduction media
BLAST	Basic Local Alignment Search Tool
ChBD	Chitin-binding domain
CIDR	Cysteine-rich interdomain region
COMP	Cartilage oligomeric matrix protein
CyRPA	Cysteine-rich Protective Antigen
DBC	Dynamic binding capacity
DISC	Death-inducing signalling complex
DLS	Dynamic light scattering
DMEM	Dulbecco's Modified Eagle Medium
EDTA	Ethylenediaminetetraacetic acid
EG	Ethylene glycol
EpCAM	Epithelial Cell Adhesion Molecule
FBS	Foetal bovine serum
FPLC	Fast protein liquid chromatography
FRET	Fluorescence Resonance Energy Transfer
GFP	Green fluorescent protein
GST	Glutathione S-transferase
GuHCl	Guanidine hydrochloride
HPLC	High Performance Liquid Chromatography
IDT	Integrated DNA Technologies
IMAC	Immobilised metal ion affinity chromatography
ITC	Isothermal titration calorimetry

KIH	Knobs-into-holes
MBP	Maltose-binding protein
MD	Molecular dynamics
MWCO	Molecular weight cutoff
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
POI	Protein of interest
ppm	parts per million
SBC	Static binding capacity
SBD	Starch-binding domain
SEC-MALS	Size-exclusion chromatography-multi-angle light scattering
SDS	Sodium dodecyl sulfate
SPR	Surface plasmon resonance
TCEP	(tris(2-carboxyethyl)phosphine)
TEV	Tobacco Etch Virus
TRAIL	TNF-related apoptosis-inducing ligand
VASP	Vasodilator-stimulated phosphoprotein
VLP	Virus-like particle

Table of Contents

Abstract **1**

Authorship Declaration **2**

Acknowledgements **3**

Abbreviations **4**

Table of Contents **6**

List of Figures and Tables **10**

1 Introduction **12**

1.1 Tools for protein conjugation *12*

1.1.1 Development of SpyTag/SpyCatcher and faster reacting iterations 14

1.1.2 Expansion to the Tag/Catcher toolkit 17

1.1.2.1 Isopeptag/Pilin-C 18

1.1.2.2 SnoopTag/SnoopCatcher 18

1.1.2.3 SdyTag/SdyCatcher 21

1.1.2.4 Cpe peptide/E-Ligase 22

1.1.2.5 A_W/B_{VA} (SpyTag I3W/SpyCatcher F77V F94A) 24

1.1.2.6 Screening of CnaB domain proteins for more Tag/Catcher pairs 24

1.1.2.7 Smaller scar conjugation 25

1.1.3 Limitations to SpyTag/SpyCatcher 27

1.2 Purification of recombinant proteins *28*

1.2.1 Current affinity tags 29

1.2.1.1 Polyhistidine tag 30

1.2.1.2 Epitope tags 33

1.2.1.3 Streptavidin-binding tags 35

1.2.1.4 Tags to non-proteinaceous substrates 36

1.2.2 Tag cleavage 38

1.3 Defined oligomeric display of proteins *39*

1.3.1 Strategies to oligomerise proteins 41

1.3.1.1 Tandem genetic fusion 41

1.3.1.2 Genetic fusion to multimerising domains 42

1.3.1.3 Modular assembly using multimerising domains 43

1.3.2 SpyCatcher-based nanoassembly using multimerising domains 44

1.4 Thesis objectives *46*

2 Methods 48

2.1 Cloning 48

2.1.1 Polymerase Chain Reaction (PCR) 48

2.1.2 Agarose gel electrophoresis and gel extraction 48

2.1.3 Gibson isothermal assembly 49

2.1.4 *E. coli* transformation 49

2.1.5 Plasmid preparation 49

2.1.6 Plasmid sequence verification 50

2.2 Protein expression and purification 50

2.2.1 Media and growth conditions 50

2.2.2 Protein expression 50

2.2.2.1 Bacterial expression 50

2.2.2.2 Mammalian expression 51

2.2.3 Ni-NTA purification of soluble His-tag-proteins 52

2.2.4 Ion exchange purification of SpyDock 53

2.2.5 Denaturing purification of insoluble His-tag-proteins 53

2.2.6 Gel filtration purification of SpyCatcher002-oligomers 54

2.2.7 Spy&Go purification of soluble SpyTag-proteins 55

2.2.7.1 Elution condition screening 55

2.2.7.2 Washing condition 56

2.2.7.3 Purification from different binding conditions 56

2.2.7.4 Purification of soluble proteins 57

2.2.7.5 Purification by competitive peptide elution 58

2.2.7.6 Purification by tag cleavage 59

2.2.8 Quantification of protein concentration 59

2.2.9 Quantification of protein purity and densitometry 60

2.3 Protein characterisation 60

2.3.1 SDS-PAGE 60

2.3.2 Mass spectrometry 61

2.3.3 Dynamic light scattering (DLS) 61

2.3.4 Size exclusion chromatography/multi-angle light scattering (SEC-MALS) 62

2.3.5 Isothermal titration calorimetry (ITC) 62

2.3.6 Fluorescence intensity after Spy&Go purification 63

2.4 Coupling to solid support 63

2.5 Regeneration of Spy&Go resin 64

2.6 In vitro protein conjugation and excess protein recapture 64

2.7 DR5-dependent cell killing assay 65

2.8 Rational selection of histidine mutants 65

2.9 pH elution screening of histidine mutants 66
2.9.1 pH tolerance of SpyTag003-sfGFP 66
2.9.2 Stepwise pH elution of SpyTag003-sfGFP 67

3 Development of SpyDock for Spy&Go purification of SpyTag-proteins 70

3.1 Engineering SpyCatcher for reversible binding with SpyTag 71
3.1.1 Covalent bond-abolishing mutants 72
3.1.2 Specific solid-phase attachment of SpyDock 75
3.1.3 Binding affinity characterisation 78

3.2 Optimising washing and elution conditions 79
3.2.1 Screening elution conditions 79
3.2.2 Optimising washing conditions 81
3.2.3 Purification of SpyTag-proteins 82

3.3 Spy&Go resin characterisation 87
3.3.1 Binding buffer compatibility 88
3.3.2 Long-term storage 89
3.3.3 Regeneration and reusability 90

3.4 Discussion 92

4 Mild Spy&Go elution and design of a pH-switchable SpyTag binder 96

4.1 Mild Spy&Go elution conditions 97
4.1.1 Competitive peptide elution 97
4.1.2 On-resin tag cleavage 98

4.2 Engineering of a pH-switchable SpySwitch 100
4.2.1 Histidine scanning on SpyTag:SpyCatcher interface 101
4.2.2 Investigating tolerance of sfGFP to low pH 104
4.2.3 Comparative stepwise pH elution from histidine mutants 106
4.2.4 pH-dependent elution of SpyTag003-fusions from lysate 110

4.3 Discussion 112

5 Expansion of Spy oligomerisation toolbox to probe multivalent cell signalling 119

5.1 SpyCatcher oligomers development 120

5.2 Characterisation of SpyCatcher002-oligomers assembly 122

5.3 Oligomeric assembly of a nanobody 126
5.3.1 Reaction with αDR5-SpyTag and excess reactant recapture 128
5.3.2 Characterisation of conjugated SpyCatcher002-oligomer complexes 130

5.4 Proof of concept for stoichiometry-based assay 135

5.5 Discussion 137

6 Summary and general discussion 141

Bibliography 152

List of Figures and Tables

Figure 1.1 Design of SpyTag/SpyCatcher from *S. pyogenes* CnaB2 domain 14

Figure 1.2 Design of Tag/Catcher platforms from the transamidation route of isopeptide bond formation 20

Figure 1.3 Similarity and reaction rate comparison between SpyTag/SpyCatcher and SdyTag/SdyCatcher 22

Figure 1.4 Reaction and design of Cpe peptide and E-Ligase 23

Figure 1.5 Different strategies for smaller scar conjugation 26

Figure 1.6 Strategies to multimerise proteins 46

Figure 0.1 Flow-chart of stepwise pH elution of SpyTag003-sfGFP……...68

Figure 3.1 Abolishing isopeptide bond formation 75

Figure 3.2 Covalent immobilisation of SpyDock to solid-phase support 77

Figure 3.3 Binding affinity of SpyDock to SpyTag and SpyTag002 78

Figure 3.4 Elution buffer screening for Spy&Go 80

Figure 3.5 Imidazole wash improves Spy&Go purity 82

Figure 3.6 Spy&Go purification of different SpyTag generations 84

Figure 3.7 Spy&Go purification of different bacterial proteins 85

Figure 3.8 Spy&Go purification of different mammalian proteins 87

Figure 3.9 Spy&Go binding buffer compatibility 89

Figure 3.10 Spy&Go resin is stable for long-term storage 90

Figure 3.11 Spy&Go resin can be regenerated and reused 92

Figure 4.1 Competitive peptide elution of different SpyTag-MBP generations from Spy&Go 98

Figure 4.2 On-resin tag cleavage of SpyTag-TEVsite-MBP using SuperTEV protease 100

Figure 4.3 SpyTag:SpyCatcher binding interface and residue selection for histidine substitution 103

Figure 4.4 SpyTag003-sfGFP low pH-tolerance and fluorescence 106

Figure 4.5 Rational determination of pH-sensitive SpySwitch using stepwise elution of SpyTag003-sfGFP 110

Figure 4.6 Structural analysis of selected histidine mutants ..... **Error! Bookmark not defined.**

Figure 4.7 Purification of SpyTag003-sfGFP from cell lysate 112

Figure 5.1 Structure of coiled coils and SpyCatcher002 used to construct SpyCatcher002-oligomers toolbox 122
Figure 5.2 Gel filtration and assembly size determination of SpyCatche002-oligomers 123
Figure 5.3 SpyCatcher002-oligomers assembly sizes determined by SEC-MALS 125
Figure 5.4 Schematic of multivalent DR5 activation using valency-defined SpyCatcher002-oligomers 128
Figure 5.5 Removal of excess αDR5-SpyTag by recapturing using Spy&Go resin 129
Figure 5.6 Intact protein mass spectrometry analysis of SpyCatcher002-oligomer subunit sizes 132
Figure 5.7 SpyCatcher002-oligomer:αDR5-SpyTag assembly sizes after conjugation determined by SEC-MALS 133
Figure 5.8 Hydrodynamic radius distribution of SpyCatcher002-oligomers determined by DLS 135
Figure 5.9 Effect of multivalent αDR5 on cancer cell apoptosis 136

Table 4.1 Residues of SpyCatcher003 S49C E77A for histidine substitution 104
Table 5.1 Size determination of SpyCatcher002-oligomers from gel filtration 124
Table 5.2 Size determination of SpyCatcher002-oligomers from SEC-MALS 126
Table 5.3 Size determination of conjugated SpyCatcher002-oligomer:αDR5-SpyTag from SEC-MALS 133

1 Introduction

1.1 Tools for protein conjugation

Protein-protein conjugation is the formation of a covalent linkage between two distinct protein units into a single entity, bringing together the function of each moiety into the super-assembly to create a protein complex with non-natural functions for application in different fields. Although many strong non-covalent interactions are available such as biotin-avidin with K_d of 10^{-15} M (Wilchek and Bayer, 1988), antigen-antibody with K_d of typically 10^{-7}–10^{-11} M (Braden et al., 1995), and colicin-immunity proteins with K_d of 10^{-14}–10^{-17} M (Wallis et al., 1995b, 1995a), covalently conjugating proteins provides a stronger force-resistant linkage and eliminates protein component reshuffling that can result in heterogeneity. Traditional genetic fusion of different protein moieties into a single construct may be advantageous to achieve site-specificity, however large genetic fusions may cause complications with protein expression and solubility (Brune and Howarth, 2018; Esposito and Chatterjee, 2006). Additionally, spatial orientation is restricted to only the termini.

Post-translational protein conjugation circumvents this problem by linking the proteins only after complete folding of each component. Additionally, post-translational protein conjugation provides modularity by separating the components before reaction, to help obtain optimal protein expression and relevant post-translational modifications dependent on the expression host (Brune and Howarth, 2018). Protein conjugation can be achieved by chemically coupling proteins through the reaction of specific amino acid residues (Koniev and Wagner, 2015), but problems of heterogeneity limit its usage in many applications. An assortment of tools in the protein conjugation toolkit has been developed over time (Banerjee and Howarth, 2018). Key examples are described below.

Sortase mediates the transpeptidation between an N-terminal oligoglycine and a C-terminal LPXTG motif (X = any amino acid). However, the primary substrate has a high millimolar K_m, thus limiting its use in low concentration reactions (Jacobitz et al., 2017).

Nonetheless, newer engineered sortases are available with lower K_m values, but still requiring a high substrate concentration of 1 mM biotinylated GGGYK with 100 μM tetramutant sortase (Chen et al., 2011b; Jacobitz et al., 2017). Butelase 1 mediates the transpeptidation of a C-terminal BHV (B = D or N) and any N-terminal residue (Nguyen et al., 2014). Butelase 1 regularly reaches 95% substrate conversion compared to ~75% for sortase A (Nguyen et al., 2014). However, butelase-1 requires a 5-fold excess of one substrate over the other for intermolecular conjugation and off-target conjugation to other proteins at the N-terminus is possible (Nguyen et al., 2016). For both systems, only the termini are accessible for conjugation, limiting the option for specific protein orientation that internal loop tagging can offer. Transglutaminases catalyse isopeptide bond formation between Gln and Lys of various proteins and have extensive use in the food industry, biopolymer production and non-specific labelling of proteins (Deweid et al., 2018; Kieliszek and Misiewicz, 2014). However, the lack of substrate specificity is the major limitation of transglutaminases for precise protein conjugation. A site-specific transglutaminase from *Kutzneria albida* was discovered that can conjugate two peptides with minor off-target reactions to other sequences, but the catalytic efficiency is still low (Steffen et al., 2017). Split inteins work differently from the previous three enzymes, whereby the reconstitution of N-intein and C-intein results in the self-excision of the intein, thus ligating the flanking protein sequences called exteins (Shah and Muir, 2014). Although initially slow, the kinetics of split inteins have been improved by protein engineering (Aranko et al., 2014; Stevens et al., 2016). Potential drawbacks are the limited position of inteins at the target protein termini and poor solubility for some inteins (Oeemig et al., 2009; Zheng et al., 2012).

The SpyTag/SpyCatcher conjugation technology was introduced by the Howarth lab from the University of Oxford in 2012 to develop a simple and rapid method of protein conjugation with high substrate conversion (Zakeri et al., 2012). The technology has seen further progress with the expansion of more Tag/Catcher pairs (Liu et al., 2017; Pröschel et al., 2017; Tan et al., 2016; Veggiani et al., 2016; Young et al., 2017) and the development of

faster reacting SpyTag/SpyCatcher pairs (Keeble et al., 2017, 2019). Since its introduction, the SpyTag/SpyCatcher system has seen a boost in usage across the scientific community, with more than 200 articles utilising the technology.

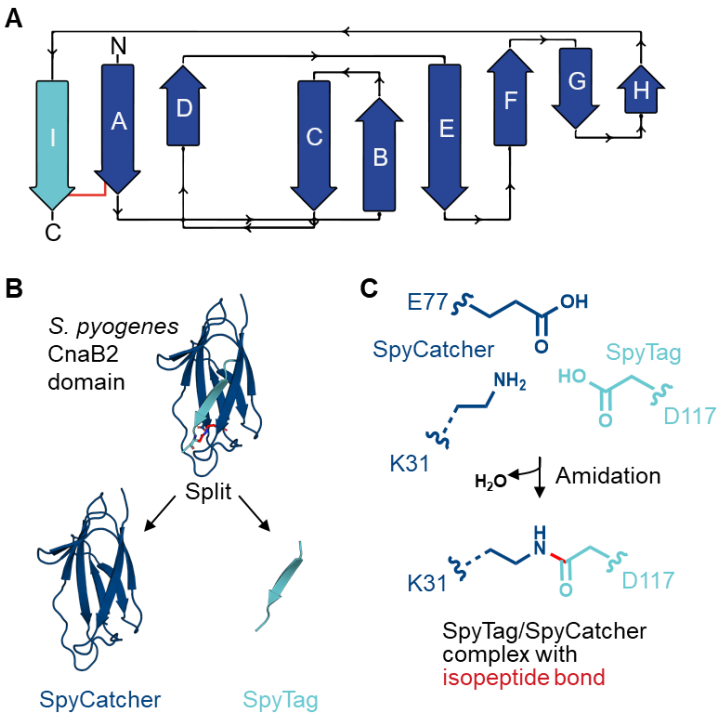


Figure 1.1 Design of SpyTag/SpyCatcher from *S. pyogenes* CnaB2 domain. (A) A topological view of the FbaB adhesin from *S. pyogenes* containing a CnaB2 domain comprised of 9 β-sheets stabilised by an intramolecular isopeptide bond (red) between strand A and strand I. (B) The CnaB2 domain (PDB ID: 2X5P) was split at the C-terminus to produce SpyTag (cyan) and SpyCatcher (dark blue) after rational engineering through tag extension and increasing surface charge in the protein partner. (C) SpyTag contains the reactive Asp (D117) and SpyCatcher contains the reactive Lys (K31) and catalytic Glu (E77). Upon mixing, E77 initiates the formation of an isopeptide bond between the side chains of K31 and D117 to lock the two components irreversibly, producing water as by-product.

1.1.1 Development of SpyTag/SpyCatcher and faster reacting iterations

Isopeptide bonds are amide bonds formed outside of the main polypeptide backbone, particularly between ε-amino group of Lys and the carboxyl groups of Asp and Glu or the carboxamide groups of Asn and Gln residues (Asquith et al., 1974). Intermolecular isopeptide bonds catalysed by enzymes have been documented for decades such as from ubiquitin

ligase (Pickart and Eddins, 2004) and transglutaminase (Griffin et al., 2002). Subsequently, the surprising structural discovery of the first intramolecular isopeptide bond in the major pilin protein Spy0128 of *Streptococcus pyogenes* helped explain the high stability and integrity of Gram-positive bacteria pili during host cell infection and colonization, when combined with the sortase-catalysed intermolecular bridging (Hae et al., 2007). This phenomenon is a proximity-induced reaction that is autocatalysed when the reactive groups are brought together in a hydrophobic environment to favour isopeptide bond formation from a nucleophilic attack of the Lys *ε*-amino group on the carbonyl of the receiving group, releasing ammonia or water depending on the type of isopeptide bond formed (Figure 1.1 and Figure 1.2) (Kang and Baker, 2011). Since then, many other intramolecular isopeptide bonds were discovered and re-discovered in Gram-positive bacteria pili and surface adhesins, revealing the widespread nature of the adaptation (Forsgren et al., 2010; Izoré et al., 2010; Paterson and Baker, 2011; Sridharan and Ponnuraj, 2016; Zähner et al., 2011). An isopeptide bond was also engineered into an immunoglobin-like protein natively lacking any intramolecular side chain reactions, which enhanced its stability (Kwon et al., 2017).

One particular protein, the adhesin fibronectin-binding protein FbaB from *S. pyogenes*, contains an intramolecular isopeptide bond in its CnaB2 domain formed between Lys-Asp (Hagan et al., 2010; Oke et al., 2010). The Howarth lab developed the SpyTag/SpyCatcher technology by splitting this CnaB2 domain and rationally engineering the protein to exploit the spontaneous isopeptide bond formation for recombinant protein conjugation (Figure 1.1) (Zakeri et al., 2012).

The CnaB2 domain is an immunoglobulin-like fold containing 9 β-strands from A to I (Figure 1.1 A). The N-terminal strand A consists of 21 amino acids poorly resolved in the crystal structure (Li et al., 2014). Strand A contains K31 which forms an isopeptide bond with strand I‘s D117, following an amidation reaction catalysed by E77 on strand E. Following the direction of established split proteins (Ozawa, 2006), Zakeri *et al.*, (2012) split the domain to remove the shorter C-terminal strand I to make SpyTag and left the rest to be SpyCatcher,

following rational engineering to increase reactivity (Figure 1.1 B and C). SpyTag fused to maltose-binding protein (MBP; SpyTag-MBP) can react completely with SpyCatcher in under 1 h to reconstitute the whole CnaB2 domain in a range of different buffers, pH values, temperatures, and tag localisation either at N-terminus, C-terminus or internal (Zakeri et al., 2012). The reconstituted complex, SpyTag-MBP:SpyCatcher, regained the high thermostability and mechanostability seen in the native CnaB2 (Hagan et al., 2010).

Although SpyTag/SpyCatcher had advantages in its specificity and ease of use, the technology still had limitations. The second order rate constant to form the SpyTag:SpyCatcher complex was 1.4×10^3 M^{-1} s^{-1}, far below the orientation-restricted diffusion limit of 10^5-10^6 M^{-1} s^{-1} (Qin et al., 2011; Schlosshauer and Baker, 2004). The slow kinetic is particularly restrictive for applications requiring low concentration of reactants, such as *in vivo* cell surface work due to the limited collisional interactions between the two substrates. To account for this issue, the Howarth lab created a faster second-generation pair called SpyTag002/SpyCatcher002 through directed evolution by phage display selection (Keeble et al., 2017). Keeble *et al.*, (2017) found eight mutations that served to stabilise intramolecular interactions within SpyCatcher002, allowing it to adopt a more stable conformation before reaction. SpyCatcher002 also had a negatively-charged surface residue to complement the positively-charged SpyTag002 that is thought to increase the rate of association between the pair. SpyTag002/SpyCatcher002 had a second order reaction rate of 2.0×10^4 M^{-1} s^{-1}, 12-fold faster than the original pair (Keeble et al., 2017).

Building on the foundation discovered from phage display, the newest iteration termed SpyTag003/SpyCatcher003 was developed (Keeble et al., 2019). The insight on electrostatic interactions between SpyTag002 and SpyCatcher002 was exploited to rationally design mutants for the third iteration by improving charge complementarity between the pair. Further negatively-charged residues were incorporated into SpyCatcher003, whilst SpyTag003 contained an extra basic Arg. SpyTag003/SpyCatcher003 had a reaction rate of 5.5×10^5 M^{-1} s^{-1}, well within the rate for diffusion-limited protein-protein interactions (Keeble et al., 2019;

Qin et al., 2011; Schlosshauer and Baker, 2004). Thus, SpyTag003/SpyCatcher003 can react to completion within minutes at nanomolar concentrations. Howarth and colleagues dissected the steps leading up to isopeptide bond formation between SpyTag003 and SpyCatcher003 through a variety of biophysical analyses to understand the effects of the mutations on the non-covalent association rate and covalent reaction rate (Keeble et al., 2019). From stopped-flow FRET and fluorescence gel densitometry analyses, the second-order rate constant for covalent bond formation (5.5×10^5 M^{-1} s^{-1}) approaches the k_{on} for non-covalent association ($6.2 \pm 0.5 \times 10^5$ $M^{-1}s^{-1}$), suggesting that the 003 mutations enabled isopeptide bond reaction by firstly aiding SpyCatcher003 binding to SpyTag003. Nonetheless, the association of SpyTag002 and SpyCatcher002 also had a k_{on} of the same magnitude ($4.2 \pm 0.5 \times 10^5$ M^{-1} s^{-1}; k_{on} for SpyTag and SpyCatcher not measurable), suggesting that the 003 mutations confer other effects (Keeble et al., 2019). Further dissection of the FRET curves and analysis of hydrogen-deuterium exchange for structural dynamics investigation suggested that the 003 mutations also help to accelerate conformational changes within the peptide-protein complex to result in faster isopeptide bond formation, compared to the previous two iterations (Keeble et al., 2019). Thus, in terms of assisting non-covalent association, the effects were already noticeable in the 002 generation, but augmented in the 003 pair. Fast conformational change due to reduced structural dynamics in the 003 pair lead to faster bond formation.

1.1.2 Expansion to the Tag/Catcher toolkit

The promising results of SpyTag/SpyCatcher as an easy genetically-encodable protein conjugation technique invited the development of additional split isopeptide bond-containing proteins. Since the initial discovery of an intramolecular isopeptide bond in *S. pyogenes* Spy0128 and the subsequent influx of other isopeptide bond-containing proteins, the reservoir of possible new Tag/Catcher pairs has gradually increased. Discussed below are examples of other Tag/Catcher pairs or similar technologies that were generated to accompany SpyTag/SpyCatcher.

1.1.2.1 Isopeptag/Pilin-C

Prior to the development of SpyTag/SpyCatcher, the Howarth lab initially trialled a different protein as the basis for the Tag/Catcher system. The major pilin protein Spy0128 from *S. pyogenes* was the first protein to be discovered with spontaneous intramolecular isopeptide bond formation (Hae et al., 2007). The protein contains 2 domains, N-terminal and C-terminal, both of which contain an isopeptide bond between a Lys and an Asn catalysed by a Glu to release ammonia (Hae et al., 2007) (Figure 1.2 A and B). Spy0128 was split at the final β-strand of the C-terminal domain to produce a 16-residue peptide called Isopeptag containing the reactive Asn residue, while the rest formed the 31 kDa protein partner called Pilin-C (Zakeri and Howarth, 2010) (Figure 1.2 B). Isopeptag/Pilin-C was the first instance of a split protein that would reconstitute via a covalent bond, using the isopeptide bond formed between the peptide-protein pair (Ozawa, 2006; Zakeri and Howarth, 2010). The pair was successfully used to cross-link fluorescent proteins for fluorescence resonance energy transfer, immobilise proteins on a solid-phase, and form redox-sensitive nanochains (Abe et al., 2013; Matsunaga et al., 2013). However, the reconstituted protein was big and the reaction between the components was limited, with the product detectable only after 1 h and reaching a mere 60% reconstitution after 72 h (Zakeri and Howarth, 2010).

1.1.2.2 SnoopTag/SnoopCatcher

The Howarth lab dissected *Streptococcus pneumoniae*'s pilus-associated adhesin protein RrgA, which contains isopeptide bonds in the D2 and D4 domains (Izoré et al., 2010). The smaller D4 domain was split by removing the first N-terminal β-strand containing the reactive Lys to make a peptide tag, with the remainder of the protein containing the reactive Asn and catalytic Glu. After rational design of the peptide tag and protein partner, SnoopTag and SnoopCatcher were generated (Veggiani et al., 2016) (Figure 1.2 C). The design of the SnoopTag/SnoopCatcher platform enabled complete orthogonality with SpyTag/SpyCatcher;

SnoopTag, which bears the reactive Lys, cannot react with SpyCatcher that also contains a reactive Lys. Conversely, the same is true with SpyTag (reactive Asp) and SnoopCatcher (reactive Asn) (Veggiani et al., 2016; Zakeri et al., 2012). Furthermore, SnoopTag and SnoopCatcher can react to completion in just 2 h when one component is in 2-fold molar excess. The orthogonality between SpyTag/SpyCatcher and SnoopTag/SnoopCatcher was capitalised on to create higher-order macroassemblies. Examples include chains of multi-component protein units analogous to solid-phase peptide synthesis (Veggiani et al., 2016), a vaccine against two different antigens (Brune et al., 2017), functionalised hydrogels (Wieduwild and Howarth, 2018), protein multilayers (Zhang, Wang, Da, et al 2018, Zhang, Wang, Sun, et al 2018), and a white-light emitting biohybrid diode (Aguino et al., 2018). A different group split the larger D2 domain of RrgA to form a protein-protein “Bio Molecular Welding” platform consisting of Jo and In proteins (Bonnet et al., 2017) (Figure 1.2 C). Jo and In can also reconstitute completely after 2 h, however the large size of each component (10 and 16 kDa), and hence the size of the reconstituted complex, reduced its general applicability. Nonetheless, the isopeptide bond between Jo and In is catalysed by an Asp instead of a Glu; this geometrical constraint may reduce cross-reactivity to other Glu-catalysed Tag/Catcher pairs (Kang and Baker, 2011). Although cross-reactivity has not been tested, Jo and In have potential in orthogonality with SpyTag/SpyCatcher and SnoopTag/SnoopCatcher.

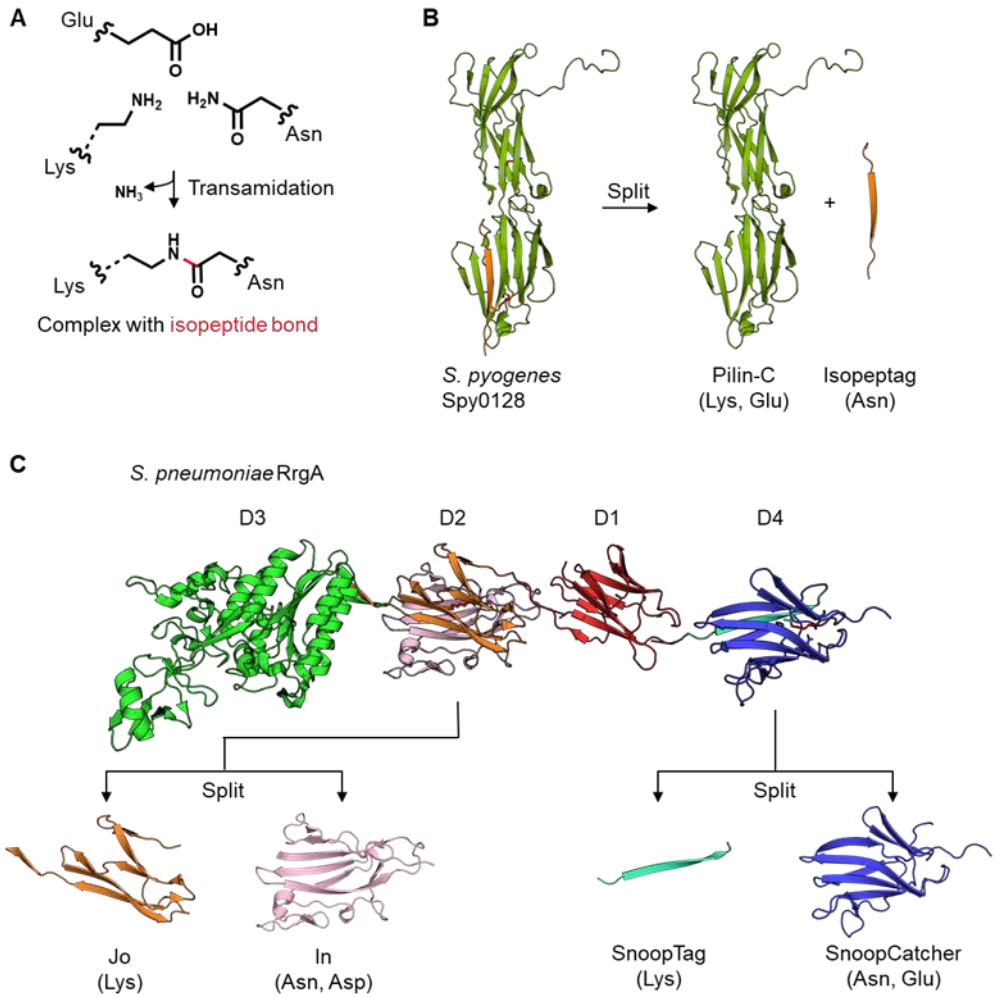
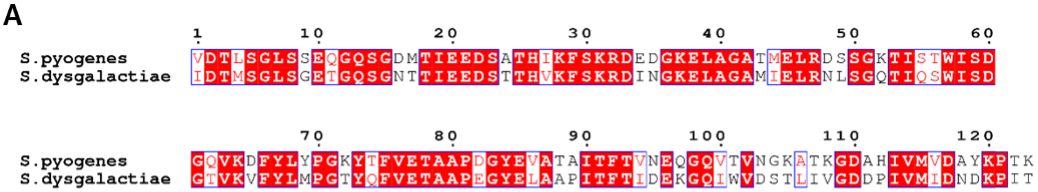


Figure 1.2 Design of Tag/Catcher platforms from the transamidation route of isopeptide bond formation. (A) Chemistry of isopeptide bond formation between a reactive Lys and a reactive Asn catalysed by Glu or Asp with ammonia as by-product. Upon mixing, an irreversible isopeptide bond is formed (red) between the side chains of Lys and Asn. (B) Isopeptag and Pilin-C were developed by splitting the Spy0128 protein from *S. pyogenes*. (C) *S. pneumoniae* RrgA protein has four domains with D2 and D4 containing intramolecular isopeptide bonds. D4 is split to make SnoopTag and SnoopCatcher whereas D2 was split to make Jo and In. For all split domains, the residue(s) in brackets indicate the amino acids in the component partaking in isopeptide bond formation. The first residue is the reactive and second residue (if any) is the catalytic residue.

1.1.2.3 SdyTag/SdyCatcher

In the same year that SnoopTag/SnoopCatcher was introduced, a group in Singapore opted for a different approach to develop an additional Tag/Catcher pair. The group used a fibronectin-binding protein containing a CnaB domain from *Streptococcus dysgalactiae* found from a homology Basic Local Alignment Search Tool (BLAST) search using *S. pyogenes* FbaB CnaB2 domain (parent protein of SpyTag/SpyCatcher) as the query (Tan et al., 2016). The isopeptide formation chemistry is the same as in SpyTag/SpyCatcher, occurring via an amidation reaction between Lys and Asp, instead of a transamidation reaction between Lys and Asn found in SnoopTag/SnoopCatcher and Isopeptag/Pilin-C. The *S. dysgalactiae* CnaB domain protein was split according to Zakeri *et al.*, (2012) at the C-terminal β-strand to create a 13-residue peptide called SdyTag containing the reactive Asp and a partner protein called SdyCatcher containing the reactive Lys and catalytic Glu (Tan et al., 2016). Because of the similar isopeptide chemistry, moderate sequence similarity (Figure 1.3 A), and similar nature of the split, SdyTag/SdyCatcher suffers from cross-reactivity with the SpyTag/SpyCatcher platform, with SdyCatcher having a 2.5-fold higher kinetic preference to react with SpyTag than its cognate SdyTag partner (Tan et al., 2016) (Figure 1.3 B). The cross-reactivity issue hinders the wide adoption of the system. However, the authors suggest that the biased kinetic preferences between each Tags and their Catchers can be exploited for a systematic kinetic control of multiprotein assembly.

A



B

Rate of reaction with different Tags and Catchers (M^{-1} s^{-1})

	SpyCatcher	SdyCatcher
SpyTag	1930	1130
SdyTag	6	450

Figure 1.3 Similarity and reaction rate comparison between SpyTag/SpyCatcher and SdyTag/SdyCatcher. (A) Sequence alignment using ESPript 3.0 revealed a 63% sequence identity between the parent proteins of SpyTag/SpyCatcher and SdyTag/SdyCatcher. (B) Comparison of reaction rates from Tan, Hoon & Wong (2016) between the two Tags and two Catchers showed high cross-reactivity between the different components. SdyCatcher even has a 2.5-fold higher preference for reactivity with SpyTag than SdyTag.

1.1.2.4 Cpe peptide/E-Ligase

Seven years after the surprising discovery of an intramolecular isopeptide bond, the same group led by Edward Baker discovered another unprecedented intramolecular covalent bond chemistry in a protein. Cpe0147 is a cell wall-anchored adhesion protein from *Clostridium perfringens* composed of 11 repeat domains, of which the structure of the first 2 domains were solved (Kwon et al., 2014). Within each domain, an ester bond between Thr and Gln was found between the first and last β-strand, catalysed by a His (Figure 1.4 A). They hypothesised that the intramolecular ester bond serves a similar role to intramolecular isopeptide bonds, which is to stabilise cell surface proteins during host cell invasion. Removal of the ester bond through site-directed mutagenesis drastically reduced thermal stability and increased susceptibility to proteolysis (Kwon et al., 2014). Although ester bonds are naturally reversible, the covalent linkage between Thr-Gln remained intact after maturation reaction and crystallisation (Kwon et al., 2014; Young et al., 2017). The authors hypothesised that an adjacent Asp hydrogen bonds to the catalytic His to block water ingress from hydrolysing the ester bond (Kwon et al., 2014). The same group then developed Cpe peptide and E-Ligase from Cpe0147 to add to the already increasing toolkit of peptide-protein covalent assembly. Cpe peptide was made by splitting the second domain of Cpe0147 at the final β-strand to make a 23-residue peptide containing the reactive Gln and catalytic His, whereas the rest of the domain became the protein partner E-Ligase containing the reactive Thr (Young et al., 2017) (Figure 1.4 B). When each component was fused to a cargo protein and reacted in a 1:1 molar ratio in the optimised buffer consisting of glycerol and calcium, the substrates reacted to >90% completion after 20 h (Young et al., 2017).

By substituting the Thr to Ser in E-Ligase, they made RE-Ligase, which can reverse the reaction upon an increase in pH and removal of glycerol to hydrolyse the ester bond for controlled assembly and disassembly of protein conjugation (Young et al., 2017). Because E-Ligase proceeds via a completely different chemistry than the canonical isopeptide bond-containing proteins, this platform is the most promising to be added to the efficient and orthogonal Tag/Catcher toolkit already made up of SpyTag/SpyCatcher and SnoopTag/SnoopCatcher. A theorised N-terminal split of Cpe0147 to make a Thr-containing peptide and a Gln-containing protein partner may be orthogonal to Cpe peptide and E-Ligase, similar to the orthogonality achieved between SpyTag/SpyCatcher and SnoopTag/SnoopCatcher (Veggiani et al., 2016). Additionally, the authors postulate that since ester bond-containing proteins are highly sequence-diverse, orthogonality can be achieved by selecting proteins from distant phylogenetic branches (Young and Squire, 2020).

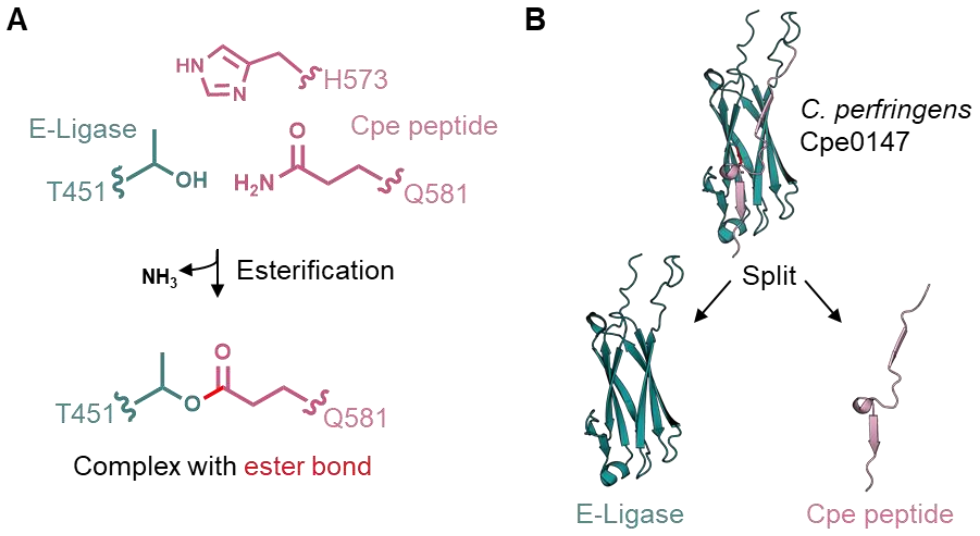


Figure 1.4 Reaction and design of Cpe peptide and E-Ligase. (A) In *C. perfringens* Cpe0147 protein, H573 catalyses a reaction between the side chains of T451 and Q581 to form an ester bond (red), releasing ammonia as by-product. (B) Cpe0147 was split to become Cpe peptide and E-Ligase, where Cpe peptide contains the reactive Q581 and catalytic H573, whereas E-Ligase contains the reactive T451, different from the design of other Tag/Catcher pairs where the Tag contains only one residue partaking in the bond formation.

1.1.2.5 A_W/B_{VA} (SpyTag I3W/SpyCatcher F77V F94A)

Rather than splitting other isopeptide-containing proteins, the Zhang group from Beijing used SpyTag/SpyCatcher as the template to engineer a Tag/Catcher pair with a selective orthogonality to SpyTag/SpyCatcher (Liu et al., 2017). A bulky Trp residue was introduced into SpyTag to make SpyTag I3W (A_W), which impaired its reactivity with SpyCatcher to 39% reconstitution relative to 100% reconstitution between SpyTag and SpyCatcher at 4 °C after 5 h in phosphate buffered saline (PBS). The authors then constructed a site-saturated library of SpyCatcher and selected for a reactive partner to SpyTag I3W through directed evolution. SpyCatcher F77V F94A (B_{VA}) was found to have restored reactivity with A_W, yielding 60% reconstitution. However, B_{VA} still cross-reacts to completion with the native SpyTag, thus the new pair is not fully orthogonal to SpyTag/SpyCatcher (Liu et al., 2017).

1.1.2.6 Screening of CnaB domain proteins for more Tag/Catcher pairs

As with SdyTag/SdyCatcher, further CnaB domain-containing proteins may be amenable to splitting to make more Tag/Catcher pairs. In a single study, 4 different CnaB domains that natively have intramolecular isopeptide bonds with <30% sequence identity to SpyTag/SpyCatcher were examined to expand the Tag/Catcher toolkit: 3phs, 4oq1, 3kptN, and 3kptC (N and C being the N- or C-terminal domain) – named after their Protein Data Bank (PDB) ID. All the proteins undergo the transamidation route of isopeptide bond formation (Pröschel et al., 2017). In the study, they used molecular dynamics (MD) to select for appropriate split points based on the comparative stability of split domain and parent native domain. 3phs and 3kptN were unsuccessful in reconstituting after domain splitting, indicating that not all isopeptide-containing proteins are suitable as split proteins (Budzik et al., 2009; Paterson and Baker, 2011). Although split 4oq1 and 3kptC managed to reconstitute, the reaction was very slow and did not react to completion even after 24 h (Pröschel et al., 2017). The slow kinetics may be due to insufficient initial Tag and Catcher interaction to bring the two

components together which can be resolved through directed evolution and rational engineering (Keeble et al., 2017, 2019; Zakeri et al., 2012). However, it is also possible that the domains are natively slow in undergoing reaction (Kang et al., 2014). Nonetheless, the methodology of incorporating MD to find optimal split points may be beneficial to screen a larger number of protein domains for more Tag/Catcher pairs.

1.1.2.7 Smaller scar conjugation

The primary disadvantage in the Tag/Catcher system compared to other protein conjugation methods previously mentioned is the relatively big protein scar size after reconstitution. Fusing a recombinant protein to smaller sized peptides may be more beneficial to minimise the effect on structural or functional properties of the cargo protein (Kimple et al., 2013). To tackle this problem, the *S. pyogenes* CnaB2 domain was split into three fragments resulting in SpyLigase containing the catalytic Glu to conjugate SpyTag (with reactive Asp) and KTag (with reactive Lys) together (Fierer et al., 2014) (Figure 1.5 A). The study was a good proof of concept for the tripartite split of isopeptide proteins. However, the non-physiological reaction conditions, slow rate with incomplete substrate conversion, and the presence of SpyLigase in the SpyTag-KTag complex limit the general usability of the technology (Fierer et al., 2014).

To address these problems, the Howarth lab then developed the SnoopLigase peptide-peptide ligation platform through a tripartite split of the SnoopTag/SnoopCatcher (Buldun et al., 2018) (Figure 1.5 B). SnoopLigase contains the catalytic Glu to join SnoopTag (with reactive Lys) and DogTag (with reactive Asn). SnoopLigase had better conjugation efficiency than SpyLigase under milder conditions, and the bound SnoopLigase after SnoopTag:DogTag conjugation can be removed through solid-phase attachment of SnoopLigase, leaving only a small peptide-peptide ligation scar (Buldun et al., 2018). This method also showed no immunogenicity against the bacteria-derived SnoopLigase after its removal in mouse model vaccination (Andersson et al., 2019).

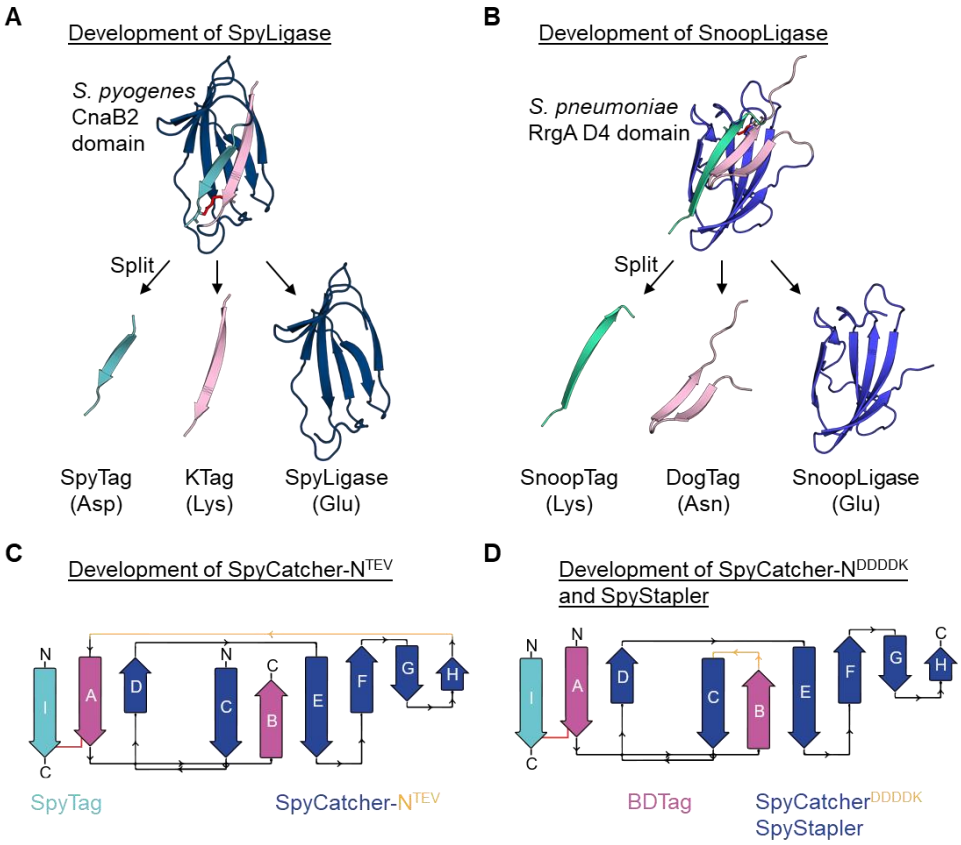


Figure 1.5 Different strategies for smaller scar conjugation. Tripartite split of (A) *S. pyogenes* CnaB2 domain to make SpyLigase and (B) *S. pneumoniae* RrgA D4 domain to make SnoopLigase. Each Ligase mediates isopeptide bond formation between its respective Tags. Residue in brackets indicate the amino acids in the component partaking in isopeptide bond formation. (C) Topological view of the circularly-permutated SpyCatcher-N^{TEV} reacted with SpyTag. Upon covalent conjugation, strands C-H are excised out by TEV protease (cleavage at yellow region), leaving only SpyTag and strands A-B from SpyCatcher-N^{TEV}. (D) Topological view of SpyCatcher^{DDDDK} or SpyStapler reacted with SpyTag. For SpyCatcher^{DDDDK}, no circular permutation was involved, but the enterokinase cleavage site, DDDDK, was introduced in-between strand B and C (yellow). Upon reaction, enterokinase cleaves this position, excising strands C-H. For SpyStapler, SpyCatcher was further split into BDTag (containing Lys) and SpyStapler (blue, without yellow region). SpyStapler mediates the ligation of BDTag with SpyTag, similar to SpyLigase.

Adding to the SpyLigase and SnoopLigase platforms, the Zhang group from Beijing created several SpyCatcher-derived technologies that aim to reduce the scar size of protein assemblies after conjugation. Firstly, SpyCatcher-N^{TEV} was developed through circular permutation: after reaction with SpyTag, the non-isopeptide-containing β-strands can be

cleaved off through an inserted Tobacco Etch Virus (TEV) site using TEV protease (Zhang et al., 2018d) (Figure 1.5 C). They built upon this work further by inserting an enterokinase recognition site, DDDDK, in the 2nd loop between strands B and C (Figure 1.5 D), where upon reaction with SpyTag, the other strands were removed to produce a smaller scar (Zhang *et al.*, 2019). This version had a 10-fold faster second order reaction rate than SpyCatcher-N^{TEV}, but still needed the fusion of SpyCatcher^{DDDDK} variant to the protein of interest, unlike with SpyLigase. This prompted the group to develop SpyStapler (analogous to SpyLigase) that can ligate SpyTag and BDTag (analogous to KTag) together (Wu et al., 2018) (Figure 1.5 D). However, SpyStapler was still bound to the peptide substrates after reaction, similar to SpyLigase and SnoopLigase (Buldun et al., 2018; Fierer et al., 2014; Wu et al., 2018).

1.1.3 Limitations to SpyTag/SpyCatcher

Although many Tag/Catchers have been developed since the inception of the first split protein conjugation system, Isopeptag/Pilin-C (Zakeri and Howarth, 2010), the SpyTag/SpyCatcher platform still dominates over the other platforms (Keeble and Howarth, 2019). The preference for SpyTag/SpyCatcher can be attributed to the rapid reaction, complete substrate conversion and mild reaction conditions.

The first generation saw relatively low reaction rate, but directed evolution and rational engineering improved its rate by approximately 390-fold to achieve diffusion-limited protein-protein interaction rate (Keeble et al., 2019). The reconstituted complex as a conjugation scar is also inherently large at about 14 kDa, and fusing SpyCatcher to a protein of interest may not always be suitable to limit the detrimental effects of the overall protein structure and function (Kimple et al., 2013). Different elegant methods were created to solve this issue, such as SpyLigase, SpyStapler and SpyCatcher-N^{TEV} as previously discussed (Fierer, Veggiani and Howarth, 2014; Wu *et al.*, 2018; Zhang, Wu, *et al.*, 2018; Zhang *et al.*, 2019). Nonetheless, several issues still remain that may hinder the technology from being more widespread.

Currently, many SpyTag-proteins used *in vitro* were purified using a secondary purification tag, most predominantly a polyhistidine tag. Purification tags are powerful to isolate and enrich target proteins from a complex mixture, but they lose utility once the process is completed.

Additionally, SpyTag/SpyCatcher is useful for modular protein assemblies. As such, different protein scaffolds fused to SpyCatcher have been established to accommodate the mixing-and-matching of SpyTag-proteins. For rapid and scalable vaccine generation, the scaffolds include a virus-like particle AP205 (Brune et al., 2016), a coiled coil domain from chicken complement inhibitor C4-binding protein (C4bp) (Brune et al., 2017), and a computationally-designed protein nanoparticle (Bruun et al., 2018). For nanomedicine, a gold nanoparticle scaffold mediated by glutathione S-transferase (GST) gold-sulphur bond was developed (Ma et al., 2018), whereas a self-assembling protein tile enabled coimmobilisation of cascade enzymes (Zhang et al., 2018a). Further scaffolds may be needed to increase the reach of SpyTag/SpyCatcher into other scientific fields.

1.2 Purification of recombinant proteins

The area of protein study and utilisation has been promoted by the advances in protein production and purification. The separation of proteins from a liquid mixture was explored as early as 1910 when α-amylase was captured on a solid-phase support containing insoluble starch, based on the enzyme's natural interaction with the substrate (Starkenstein, 1910). This method would then open the gates for further protein isolations based on natural protein-ligand interactions, that can be capitalised on by immobilising the substrate on a solid support (Hage and Matsuda, 2015). Since then, many other methods of protein isolation and purification were developed that exploited the different characteristics of proteins based on solubility, size, charge and binding affinity (Labrou, 2014; Saraswat et al., 2013). These methods can generally be classed based on the level of purity: either low, medium or high purity. Separation

based on solubility gives lower purity whereas chromatographic methods enable higher resolution (Labrou, 2014).

Among the different chromatographic methods, affinity chromatography enables the highest sensitivity and specificity for high purity purification, often in a single step (Kimple et al., 2013; Saraswat et al., 2013). Affinity chromatography, which has its roots in the seminal discovery by Emil Starkenstein in 1910, utilises the principle of biomolecular recognition whereby biological molecules participate in a specific and reversible interaction based on the high affinity towards their native or designed ligand. Initially, affinity chromatography was restricted to purifying proteins with natural interactions to substrates found in solid form such as starch, edestin, alginic acid, and elastin (Hage and Matsuda, 2015). After more than half a century since 1910, affinity chromatography began to be more applicable to a wider selection of proteins with the combination of beaded agarose as the support matrix and a general protein immobilisation technique using cyanogen bromide (Cuatrecasas et al., 1968). The chief advance was recombinant DNA technology to genetically fuse a protein or peptide sequence having affinity to a matrix column, called affinity tags, to a target protein of interest that has no available binding partner (Kimple et al., 2013; Terpe, 2003). Recombinant DNA has allowed the effective and scalable purification of any protein without prior knowledge of its solubility, size, or charge (Gräslund et al., 2008; Kimple et al., 2013; Vassylyeva et al., 2017).

1.2.1 Current affinity tags

Affinity tags consist of either polypeptides or peptides that have an inherent affinity towards their cognate ligand which is immobilised on a solid-phase support. Traditional affinity tags were predominantly large protein fusions with a known binding partner, such as MBP (di Guana et al., 1988) and GST (Smith and Johnson, 1988). Apart from assisting purification through affinity chromatography, these protein tags also confer increased solubility for better expression of the target protein due to several features such as having an efficient translation initiation when placed at the N-terminus, attracting chaperones during protein folding, having

intrinsic chaperone-like activity, and having net surface charges (Costa et al., 2014; Esposito and Chatterjee, 2006; Waugh, 2005). However, the expression of a large protein affinity tag alongside the protein of interest requires high amino acid and energy reserves for transcription and translation that can exert substantial metabolic burden to the cellular expression system (Costa et al., 2014; Waugh, 2005). Smaller peptide affinity tags have become preferred as the fusion partner to purify proteins through affinity chromatography.

1.2.1.1 Polyhistidine tag

In 1975, the general ability of metal ion interaction with proteins were exploited to fractionate different untagged proteins through immobilised metal ion affinity chromatography (IMAC) (Porath et al., 1975). As histidine was found to have the strongest metal ion interaction, the polyhistidine tag (also His-tag or Hexahistidine tag) was developed by Roche to impart better specificity and sensitivity for protein purification (Hochuli et al., 1988). The polyhistidine tag is composed of tandem repeats of histidines, usually 6-12 residues and they suitably interact with divalent metal cations such as Ni^{2+}, Co^{2+}, Zn^{2+} and Cu^{2+} that are immobilised on affinity matrices through a covalently-bound chelator that traps the metal ions by coordination (Block et al., 2009). Different chelators coordinate the metal ions differently; iminodiacetic acid (IDA) is a tridentate chelator that coordinates Ni^{2+} through 3 coordination bonds, leaving the remaining 3 to bond with electron donor groups from histidine, although it is theorised that only 2 are suitably positioned to participate in the protein-metal coordination (Gaberc-Porekar and Menart, 2005). As IDA is a tridentate chelator, it suffers from metal ion leaching during washes and elution (Block et al., 2009; Bornhorst and Falke, 2000; Gaberc-Porekar and Menart, 2005). Tetradentate chelators such as nitrilotriacetic acid (NTA) for Ni^{2+} and carboxylmethylaspartate (CMA) for Co^{2+} were developed as a response to this issue as a coordination number of 4 to the metal ions would increase their binding affinity to the matrix thereby reducing metal ion leaching (Chaga et al., 1999; Hochuli et al., 1987). However, traces of Ni^{2+} can still be detected in the elution pool from Ni-NTA (Block et al., 2009). A pentadentate

chelator tris(carboxymethyl)ethylene diamine (TED) for Ni^{2+} was then developed but still suffers from metal leakage, albeit minimised, and had compromised protein binding capacity, since there is only 1 coordination bond left to interact with the polyhistidine-tagged protein (Block et al., 2009).

The polyhistidine tag is currently the most popular affinity tag in use (Gräslund et al., 2008; Vassylyeva et al., 2017). In a retrospective analysis of crystallisation of key membrane proteins and nucleic acid-binding proteins, the polyhistidine tag was the first choice for protein purification in 86% and 73% of the cases, suggesting the high preference of the tag across the scientific community (Vassylyeva et al., 2017). Proteins purified by polyhistidine tags can regularly achieve >80% purity in a single step (Kimple et al., 2013), however, co-purification of endogenous proteins are common in IMAC (Bolanos-Garcia and Davies, 2006; Gräslund et al., 2008). The polyhistidine tag is popular because the small size may minimise any perturbation of the protein of interest to be purified. The affinity matrix can also theoretically be used indefinitely as there is no proteinaceous element and it has a high binding capacity of up to about 50 mg of protein per mL of packed resin (Qiagen). These features also make the resins for polyhistidine tags cheaper for purchase. Additionally, the coordination bond between histidine and the metal ion can withstand harsh denaturants such as 6 M guanidine hydrochloride (GuHCl) and 8 M urea (Spriestersbach et al., 2015). Polyhistidine-tagged proteins can be eluted from the affinity resin matrix using a relatively mild elution condition of 150-500 mM imidazole that competes for binding to the immobilised metal ions with histidines, as well as pH 4.5-6.0 or EDTA as alternative elution conditions (Spriestersbach et al., 2015). The mild elution conditions are advantageous over epitope tags (Section 1.2.1.2) to preserve protein structure and function. The tag can also be used for tandem affinity purification (TAP) to study protein complexes, of which the polyhistidine tag is the most popular affinity tag for TAP among >15 different tags (Li, 2010). Because of these reasons and more, the polyhistidine tag remain the most used affinity tag in many laboratories.

However, even the most ubiquitous affinity tag also has down-sides. Metal ion leaching, as discussed previously, has been the highest concern for IMAC methodologies. In nickel-based IMAC, leaching can be detrimental to health as nickel is a toxic heavy metal and is implicated in carcinogenesis and nickel-induced oxidative stress (Bal et al., 2000; Das et al., 2019). Some individuals may also have dermatological nickel sensitivity (Torres et al., 2009). Aside from health issues, polyhistidine tagging usually suffers from non-specific binding of other proteins with high histidine content or with histidine clusters and their subsequent co-purification (Block et al., 2009; Bornhorst and Falke, 2000). However, non-specific binding can be circumvented by washing in the presence of low concentration of imidazole or using higher number of histidines in the tag. The presence of a polyhistidine tag has also been implicated with low protein solubility, particularly as a C-terminal fusion tag (Block et al., 2009; Woestenenk et al., 2004). This issue can be overcome by fusing a secondary solubility tag but will then make the construct bigger and incur higher metabolic burden (Section 1.2.1).

More importantly, polyhistidine tags have been reported to have direct negative effects to the structure and function of certain proteins. A mutant form of a DNA-binding protein, pi200(30.5), was monomeric in solution but existed as a dimer when fused with a polyhistidine tag (Wu and Filutowicz, 1999). Khan *et al.*, (2012) found that a vaccine candidate protein from *Plasmodium falciparum* called $MSP1_{42}$ had more unfolded and aggregated species when a polyhistidine tag was present. Although the polyhistidine-tagged $MSP1_{42}$ induced more immunogenicity, there was less antigen-specific antibodies compared to native protein. The position of polyhistidine tags also matter, as when fused to either the N- or C-terminus of corticotropin-releasing factor receptor type 2a, different disulfide bridging arrangements may occur (Klose et al., 2004). A minimal zinc finger has been shown to undergo a conformational change when tagged with polyhistidine, although the change in structure did not affect its function (Chant et al., 2005). A survey of crystallised proteins in the PDB with and without a polyhistidine sequence also confirmed that, while most structures are unaffected by the extra histidines, some proteins had greater structural disorder (Carson et al., 2007). While changes

in structure may be inconsequential for some applications, changes in the kinetics, function, and stability of proteins may be more severe. Two different enzymes had lower kinetics and catalytic efficiency when a polyhistidine tag is fused (Panek et al., 2013; Sabaty et al., 2013). A polyhistidine tag stably interacted with a split green fluorescent protein (GFP) system, which may explain inconsistencies in similar split systems (Deng and Boxer, 2018). Additionally, the thermal stability of 10 proteins with or without a polyhistidine tag was compared using differential scanning fluorimetry and it was found that the tag conferred increased stability to only one protein, MtDapB from *Mycobacterium tuberculosis*, whereas 4 proteins had no significant differences and the remaining 5 had worse thermal stability (Booth et al., 2018). However, the emergence of these and other cases of affected protein structure and function with a polyhistidine tag may be more reported compared to other affinity tags because there is a higher usage of polyhistidine tag.

1.2.1.2 Epitope tags

Epitope tags are peptide sequences recognised by antibodies or antibody mimetics. Tagging proteins with the epitope tag enables proteins to be detected and purified by the epitope's antibody without generating an antibody against the target protein, thereby simplifying workflow. The first epitope tag, a 9 or 14 amino acid residue epitope of the neuropeptide substance P, was used to detect and monitor the intracellular distribution of mutated hsp70 proteins in monkey COS cells using a fluorescent antibody against neuropeptide substance P (Munro and Pelham, 1984). The method enabled rapid visualisation of protein localisation compared to predecessor techniques, but a chemical modification to convert protein carboxyl groups to amides was needed to assist efficient antibody recognition and binding (Munro and Pelham, 1984). Subsequently, the first fully synthetic epitope tag was developed for protein detection and purification, called FLAG-tag (Hopp et al., 1988). Initially, FLAG-tag was position-sensitive, being able to bind to its cognate antibody M1 only when placed at the N-terminus. Nonetheless, the mild elution condition of calcium removal by

chelating agent and the presence of an enterokinase cleavage site within the FLAG-tag sequence were attractive to encourage high adoption of FLAG-tag within the scientific community. Over time, further FLAG-tag advancements were made such as the generation of different anti-FLAG-tag antibodies, M2 and M5, that are position-insensitive (Einhauer and Jungbauer, 2001), and an arginine-based calcium-independent elution (Futatsumori-Sugai et al., 2009) to make FLAG-tag as one of the most used epitope tags for immunodetection and pull-down assays. Many other epitope tags were made synthetically or by exploiting the natural antigenic properties of certain amino acid sequences (Kimple et al., 2013). Prominent epitope tags include the HA-tag (Field et al., 1988), myc-tag (Evan et al., 1985), and C-tag (De Genst et al., 2010), all working under the same principle.

Affinity chromatography using epitope tags is frequently lower in binding capacity, and hence purification yield, compared to IMAC polyhistidine tag because a protein element (antibody or antibody mimetic) is immobilised onto the solid matrix support instead of metal ions. Steric clashes from the immobilised antibody or antibody mimetic reduce the surface area of contact with the tagged protein. Nonetheless, the purity of epitope tag affinity chromatography often surpasses that of IMAC polyhistidine tag due to the high binding affinity between the epitope and its binder, as is usual for antigen-antibody binding. For example, the recently developed ALFA-tag has a picomolar affinity to its nanobody binder (Götzke et al., 2019). Apart from purification, epitope tags are also commonly used to show protein expression, membrane protein topology, and protein cellular localisation and dynamics (Jarvik and Telmer, 1998). The variety of epitope tags available facilitates orthogonal multi-labelling of tagged proteins.

However, most antibodies and mimetics raised against these epitopes contain intramolecular disulfide bonds, thus rendering the usage of epitope tags impractical under reducing conditions. Additionally, individual epitope tags may be disadvantageous under certain settings. The Tyr in FLAG-tag was shown to be prone to sulfation in some mammalian and insect expression systems, leading to reduced affinity to the cognate anti-FLAG-tag

antibody (Schmidt et al., 2012). 3×HA-tag has also been shown to impair protein structure and function (Hurst et al., 2018; Saiz-Baggetto et al., 2017). The acidic elution condition of some epitope tagged proteins is commonly the leading issue with the method since it may irreversibly damage the target protein.

1.2.1.3 Streptavidin-binding tags

The binding between streptavidin and biotin remains one of the most well-known non-covalent interactions in nature due to its extremely high binding affinity at K_d = 4 x 10^{-14} M (Wilchek and Bayer, 1988). Proteins that are overexpressed and biotinylated *in vivo* can be captured and immobilised onto streptavidin- or avidin-based solid matrix support for purification. Many groups have developed a protocol for the purification of biotinylated substrates using immobilised streptavidin, but the elution conditions were very harsh or not applicable to all substrate types. 2 M 2-mercaptoethanol at room temperature can elute only biotinylated RNA (Jenne and Famulok, 1999), pure water at 70 °C can elute only biotinylated DNA (Holmberg et al., 2005), and 2% sodium dodecyl sulfate (SDS) with 30 mM biotin, 6 M urea and 2 M thiourea was needed for quantitative release of biotinylated proteins (Rybak et al., 2004). Alternatively, streptavidins were modified for lower binding affinity, commonly by disrupting the tetramerising interface (Laitinen et al., 2006; Lim et al., 2013; Wu and Wong, 2005), but utilising an affinity-compromised streptavidin reduced the initial appeal of the very high affinity of native biotin-streptavidin interaction. Switchavidin, with an introduced histidine, allows elution of biotinylated substrates, albeit with the harsh elution of citric acid pH 2.0 and 0.25% SDS (Taskinen et al., 2014). *In vivo* biotinylation of proteins is also required, adding to the number of steps and material for protein purification.

Because of these issues, genetically-encoded peptide tags were explored as an alternative to tagging with biotin. Strep-tag was developed from a peptide library to bind streptavidin and eluted through competition with biotin (Schmidt and Skerra, 1993), but its dependence for C-terminal localisation limited its use in protein purification. The eight amino

acid Strep-tag II was then developed to overcome this limitation and replace Strep-tag, enabling position-independent purification, but the interaction between Strep-tag II and streptavidin was weak at K_d = 13 μM (Schmidt et al., 1996). Through a series of mutations in a flexible streptavidin loop region, the authors generated Strep-Tactin that has a 10-fold lower K_d to Strep-tag II at 1 μM (Voss and Skerra, 1997). Further developments to increase the binding affinity was made, including having two Strep-tag II sequences in the same tag known as Twin-Strep-tag for increased avidity (Schmidt et al., 2013) and a second-generation binder called Strep-Tactin XT with tighter binding to Strep-tag II to lower the K_d to the nM range (IBA Lifesciences).

Other streptavidin-binding peptides were developed to provide a solution to the low binding affinity of Strep-tag II and Strep-Tactin. The streptavidin-binding peptide, SBP-tag, is a 38 amino acid long peptide that binds to streptavidin with a K_d of 2.5 nM, but the larger size of the tag may not be applicable to all proteins (Keefe et al., 2001).

The streptavidin-binding peptide tags have the advantage over other affinity chromatography methods of mild elution conditions, using the presence of biotin or biotin analogues in a wide range of buffers. Nonetheless, some proteins are endogenously biotinylated or bind biotin and thus may be co-purified as contaminants (De Boer et al., 2003; Tytgat et al., 2015). Many cell culture media also contain biotin, which will compete with binding to the streptavidin or Strep-Tactin resin and reduce binding capacity for target protein purification (Schmidt et al., 2013). IBA Lifesciences recognised this problem and developed a solution in the form of BioLock which contains egg white avidin to capture naturally-present biotin when added to cell culture supernatants (Schmidt et al., 2013), but the price of the extra reagent may be uneconomical for large scale use.

1.2.1.4 Tags to non-proteinaceous substrates

Numerous other affinity tags are available that bind to their cognate protein ligands (Kimple et al., 2013; Terpe, 2003). In recent years, affinity tags were developed that exploited

the binding of peptides or proteins to inexpensive non-proteinaceous substrates for purification in a non-chromatographic manner. A bacterial ribosomal L2 protein binds strongly to silica and was made into a 273 amino acid-long Si-tag to assist purification of different proteins by centrifugation (Ikeda et al., 2010). However, the tag size hindered wide adoption. Another group identified a 70 amino acid long region within L2 that can bind diatomaceous earth and silica with good binding affinity to reduce the size of the original Si-tag substantially (Yang et al., 2013). Separately, a 12-residue peptide termed Car9-tag and a 14-residue peptide termed CotB1p from *Bacillus cereus* spore coat protein were developed that can bind silica and be used as affinity tags (Abdelhamid et al., 2014; Coyle and Baneyx, 2014; Soto-Rodríguez et al., 2017). Elution of CotB1p-tagged proteins was by SUMO protease tag cleavage, while Car9-tagged proteins can be eluted by 1 M lysine or arginine at pH 7.5 (Abdelhamid et al., 2014; Soto-Rodríguez et al., 2017). The chitin-binding domain (ChBD) has been used commercially for affinity purification for decades through the NEB IMPACT line (Chong et al., 1997), with the release of proteins catalysed by the self-cleaving intein fused in between the ChBD and the target protein.

Similarly, families of the cellulose binding domain have been exploited for affinity purification since the 1980s and further improvements were made including mild elution with water at 25 °C after incubation with cellulose column under 1 M ammonium sulfate (Ong et al., 1989; Sugimoto et al., 2012). Other protein or peptide tags that bind to other material include the starch-binding domain (SBD) that binds to starch which can be eluted with pH 11 (Lin et al., 2009) or β-cyclodextrin (Guillén et al., 2013), a heptamer peptide tag that binds to the ceramic fluorapatite that can be eluted with 600 mM sodium phosphate (Islam et al., 2014), and a glucan-recognition protein, GRP-tag, that binds to curdlan and elutes with proteases (Horiuchi et al., 2012). Using biopolymers or minerals as the binding surface for these peptide/protein tags presents a low-cost alternative for affinity purification since these materials are inexpensive and can be reused potentially indefinitely without loss of binding capacity. However, some of these tags are large (CeBD, GRP-tag), do not tolerate reducing

agents during binding and washing (ChBD, SBD), and are location-specific (Car9-tag, CotB1p-tag, fluorapatite-binding peptide), thus limiting their use.

1.2.2 Tag cleavage

Often, an affinity tag exerts minimal effect on protein structure and function. However, cases of impaired protein stability, folding and function were reported for many affinity tags as mentioned in section 1.2.1. These tags can be cleaved by proteases during on-resin purification or in-solution after the elution step to remove any disadvantageous characteristics imparted by the affinity tag. Commonly, peptide sequences recognised by endoproteases are cloned in between the affinity tag and the target protein to yield a tagless protein after protease treatment. Many endoprotease cleavage sites have been used to facilitate this process, with the most prominent enzymes being enterokinase, thrombin, factor Xa, TEV protease, and PreScission protease (Waugh, 2011). Endoproteases provide stringent sequence specificity, however cleavage at other sites still occur and the proteolytic reaction can be inefficient in certain buffers and conditions (Shahravan et al., 2008; Waugh, 2011). Furthermore, the workflow requires the addition of endoproteases to the reaction and the removal of the affinity tag and tagged endoprotease to recover the tagless protein in the flow-through; such a workflow and the low efficiency of cleavage reaction can lower the final yield of the target protein and increase the cost for additional proteases. SNAC-tag was developed to remove the need for an enzyme-dependent cleavage reaction. The cleavage reaction is catalysed by Ni^{2+} at room temperature and pH 8.6 buffer for 16 h to obtain higher efficiency and final yield (Dang et al., 2019). With SNAC-tag, no subsequent protease removal is necessary. However, nickel is a toxic heavy metal and care should be taken during handling.

In comparison, secondary self-cleaving protein domains have been incorporated downstream of a primary affinity tag, to allow a combined effect of protein purification and subsequent protein release after induction of self-cleavage. The primary example of this tag combination is the ChBD-intein domains discussed in section 1.2.1.4 (Chong et al., 1997),

where the ChBD binds to chitin and the intein is used to excise the ChBD-intein domains from the target protein after the introduction of reducing agents. Other inteins were explored to enable domain excision including under pH induction (Fong et al., 2010). Apart from inteins, other self-cleaving domains were explored such as the FrpC protein from *Neisseria meningitidis* whereby the self-cleaving was induced by Ca^{2+} (Sadilkova et al., 2008). Sortase A transpeptidase (SrtAc) from *Staphylococcus aureus* self-cleaved in the presence of Ca^{2+} and triglycine (Mao, 2004). However, the FrpC-induced cleavage produces a reactive anhydride at the C-terminal end of the target protein that can undergo side-reactions (Osička et al., 2004), whereas SrtAc can hydrolyse endogenous proteins *in vivo* during overexpression due to it peptidase function (Kruger et al., 2004). The size of the combined affinity tag and self-cleaving proteins are also big and thus may perturb target protein structure prior to cleavage and exert metabolic burden to cells during protein expression.

1.3 Defined oligomeric display of proteins

Protein oligomerisation is an important biological phenomenon that spans a range of different endogenous functions (Ali and Imperiali, 2005; Liu, 2015). Oligomerisation and multimerisation have been exploited in biotechnology to accomplish different objectives. The best known objective is to increase the avidity of proteins to their binding partner when the inherent affinity is low (Vauquelin and Charlton, 2013). Increasing the avidity of low affinity ligands can modulate cellular specificity, and increase potency and circulation time of the ligand in the body (Rudnick and Adams, 2009; Slaga et al., 2018).

A stronger immune response from vaccination is also promoted by antigen multimerisation, due to the interaction between B cells and the highly repetitive, rigid and ordered arrangement of antigens (Bennett et al., 2015; Zinkernagel, 2003). Additionally, superselectivity arising from multivalent interactions between a ligand and a receptor was recently proposed as a property emergent from protein multimerisation. Superselectivity is where a small increase in the number of ligands or receptors per area can modulate a big

effect in cellular behaviour (Dubacheva et al., 2014). The application of superselectivity is beneficial to understand concentration thresholds in cell surface signalling, as illustrated by the non-linear near-sigmoidal increment of bound multivalent hyaluronic acid (HA)-β-cyclodextrin in response to different ferrocene surface area coverages (Dubacheva et al., 2014). Understanding this superselective relationship can help optimise HA-β-cyclodextrin concentration to effectively bind to ferrocene.

The non-linear relationship between surface coverage and effective concentration can also be extended to other interactions. In particular, the trimeric ligand TRAIL (TNF-related apoptosis-inducing ligand) which binds to its native receptor, a class of death receptors (Ashkenazi, 2002), presents an attractive model to study superselectivity in the cellular context. Death receptor family members are often upregulated in cancer cells compared to normal cells and their activation through TRAIL binding results in apoptosis (Ashkenazi, 2008; Ma et al., 2005; MacFarlane, 2003), thus presenting an avenue for targeted cancer killing. The natively trimeric TRAIL binds to and clusters three molecules of Death Receptor 4 (DR4) and Death Receptor 5 (DR5), which contain death domains to recruit the downstream death-inducing signalling complex (DISC) to induce pro-apoptotic signalling cascades. However, TRAIL also binds to decoy death receptors, DcR1 and DcR2, which have incomplete or no death domain and so decrease the pro-apoptotic signals (Ashkenazi, 2002).

Because of the promising cancer-selective premise, recombinant TRAIL has been tested as a therapy in clinical trials (Von Karstedt et al., 2017), however the outcomes have been lacklustre. Although *in vitro* and *in vivo* studies showed the feasibility of TRAIL to induce apoptosis in cancer cells (Ashkenazi, 2008), in-patient induction of apoptosis has been elusive because cancer cells employ a variety of strategies to gain resistance against TRAIL-activating death receptor signalling. Dulanermin, the first recombinant TRAIL for use in clinical trials, did not show superiority over current standards of care in lung, lymphoma and colorectal cancers, thus the development was halted (Lemke et al., 2014; Quintavalle and Condorelli, 2012). One of the mechanisms utilised by cancer cells to resist TRAIL-induced apoptosis is

by upregulating more DcR1 and DcR2 to compete with DR4 or DR5 on binding to dulanermin, thus decreasing pro-apoptotic signalling (Kretz et al., 2019; Lemke et al., 2014). Because of this, antibodies against either DR4 or DR5 were raised to circumvent binding to the decoy receptors (Dubuisson and Micheau, 2017). Select anti-DR4 and anti-DR5 antibodies were tested in clinical trials, but still showed insignificant outcomes compared to standards of care (Kretz et al., 2019; Lemke et al., 2014). A theory on the poor performance of the antibodies include the inability of a bivalent antibody in clustering sufficient DR4 or DR5 molecules to effectively signal apoptosis (Dubuisson and Micheau, 2017; Naval et al., 2019). Thus, antibody mimetics such a scFvs, nanobodies or monobodies against DR4 or DR5 were generated which can be multimerised to target the multimeric death receptors. The valencies ranged from three to eight and have these multimers have shown preclinical promise in inducing apoptosis in different cancer cell lines (Dubuisson and Micheau, 2017; Gieffers et al., 2013; Huet et al., 2014; Papadopoulos et al., 2015; Swers et al., 2013; Wang et al., 2013). Further comparative study of different anti-death receptor valencies can shed more light into possible superselective effects on cancer cell apoptosis to design better drugs to target cancer.

1.3.1 Strategies to oligomerise proteins

1.3.1.1 Tandem genetic fusion

Proteins can be oligomerised through various strategies (Figure 1.6). The most basic strategy is to tandemly repeat the number of recombinant protein units through genetic fusion (Figure 1.6 A). Genetic fusion allows precise control over the number of desired protein units and the separation distance can be manipulated by the use of different linker lengths. For example, increasing the number of an epitope from herpes simplex virus type 1 linearly through sequential genetic fusion revealed higher immunogenicity from one repeat to four (Zheng et al., 1993). However, cloning the same protein sequence tandemly may not be straightforward because a low sequence variability between the protein units, even after varying codon usage, may raise complications in vector design and DNA assembly (Tørresen

et al., 2019). Additionally, tandem genetic fusion can only generate a linear terminus-to-terminus arrangement of the repeat protein units, which may not have optimised spatial presentation to its partner binder (Ding et al., 2017; Lennartz et al., 2018).

1.3.1.2 Genetic fusion to multimerising domains

Alternatively, a single protein unit of interest can be fused to a multimerising domain, which self-associates after expression to form multimers (Figure 1.6 B). Virus-like particles (VLPs) resemble the structure of viruses and are usually composed of highly ordered and repetitive subunits of a viral coat protein. A chimeric VLP can be made that consists of the coat protein fused to a single copy of an antigen of interest which will be displayed multimerically at the surface of a self-assembled VLP (Mohsen et al., 2017). Some commercial VLP-based vaccines are already available and many more are undergoing clinical trials (Mohsen et al., 2017; Roldão et al., 2010).

Alongside VLPs, other protein motifs can also self-assemble to form oligomers. The most well-studied is the coiled coil motif, which is an α-helical peptide that self-assembles with itself (homomeric) or another α-helical peptide (heteromeric) to form supramolecular assemblies (Woolfson, 2005). These motifs are found abundantly in nature to facilitate protein oligomerisation for tasks varying from vesicle tethering to DNA recognition (Truebestein and Leonard, 2016). α-helical coils assemble into coiled coils through knobs-into-holes (KIH) interactions formed by the side chain protrusion from one helix, the “knob”, and the “hole” formed by the diamond-shaped arrangement of four residues on the partnering helix, resulting in an extremely tight-knit helix packing (Crick, 1953; Pauling and Corey, 1953). With the abundant experimental and theoretical studies on this motif, systematic rules are in place that relate the peptide sequence to the coiled coil structure. These rules have allowed researchers to design *de novo* coiled coils or change the oligomeric nature of existing coiled coils, thus generating a database of coiled coils with varying oligomericities (dimer, trimer, tetramer etc),

orientations (parallel or antiparallel) and subunit composition (homomeric or heteromeric) (Woolfson, 2017).

The oligomerising nature of coiled coils have been exploited to multimerise monomeric proteins. A single-chain variable fragment (scFv) was fused to the dimeric GCN4 coiled coil and the mutant tetrameric GCN4 form where the “tetravalent miniantibody” showed higher functional affinity (avidity) than the bivalent form (Pack et al., 1995). A “peptabody” and a “combody” were also developed by utilising the pentamerising domain of the cartilage oligomeric matrix protein (COMP) to increase the avidity of different receptor binders (Terskikh et al., 1997; Wang et al., 2013; Zhu et al., 2010). The engineered heptameric protein, IMX313, from chicken complement inhibitor C4-binding protein was also fused to a malaria antigen, which showed a good immunogenic response (Ogun et al., 2008). Apart from coiled coils, other oligomerising domains have been used to increase protein valency, including the “trimerbody” from the collagen trimerisation domain (Blanco-Toribio et al., 2013), tetravalent scFv multimerised through the p53 tetramerisation domain (Willuda et al., 2001) and a suite of nanopolygons ranging from dimer to decamer using an engineered self-assembling GFP mutant (Kim et al., 2015). Nonetheless, problems may occur with direct fusions to oligomerising domains. A study showed that oligomerising GFP through direct fusion to a suite of dimeric to pentameric coiled coils give unpredictable assembly, which may be rescued by varying linker length; this approach thus requires bespoke optimisation for different proteins fused to different coiled coils (Cristie-David et al., 2017). Additionally, some proteins of interest may require specific expression conditions not optimal for the expression of the oligomerising domain, which can impact the yield of the fusion protein (Brune and Howarth, 2018).

1.3.1.3 Modular assembly using multimerising domains

The problems described above can be alleviated by separating the oligomerising domain (e.g. coiled coils) from the proteins of interest to achieve modularity (Figure 1.6 C). Modularity allows a one-time optimisation process of the oligomerising platform to produce

multimers with correct assembly at high yield, in parallel to the optimised expression of the protein of interest. Biotinylated proteins can be multimerised by binding to tetrameric streptavidin, which has been widely used to make MHC tetramers (Altman et al., 1996). Recently, rhizavidin has been engineered to unlock more multivalency, enabling the display of dimeric (natural rhizavidin structure), tetrameric, hexameric and octameric biotinylated ligands (Yoon et al., 2018). The valency of biotin-binding proteins can be expanded further through the SpyTag/SpyCatcher technology to achieve dodecameric assembly (Fairhead et al., 2014). However, protein biotinylation constitutes an extra step in the process and the structures are not available to cover a continuous range of valencies. Multimerising protein units linearly end-over-end to form protein chains can be achieved using protein ligation methods such as sortase or SpyLigase (Fierer et al., 2014; Li et al., 2017), but control over the number of protein units ligated is hard to achieve. A method to achieve this control is to sequentially add a protein unit end to end as illustrated by the generation of “polyproteams” using the orthogonal SpyTag/SpyCatcher and SnoopTag/SnoopCatcher platforms (Veggiani et al., 2016). However, the process is resource- and time-intensive.

1.3.2 SpyCatcher-based nanoassembly using multimerising domains

The SpyTag/SpyCatcher technology can be used to enable modular assembly of multivalent proteins by fusing the SpyCatcher component to a multimeric scaffold and reacting with a SpyTag-fused protein of interest or vice versa, thus allowing valence-controlled supramolecular assembly. The Howarth lab explored SpyTag/SpyCatcher-mediated modular nanoassembly by fusing SpyCatcher to the 180-mer coat protein of the bacteriophage AP205 and conjugating two SpyTag-fused malarial antigens, Pfs25 and cysteine-rich interdomain region (CIDR) to the nanoparticle, thereby allowing highly repetitive antigen display for effective immunisation (Brune et al., 2016). Thereafter, the group expanded the valency for SpyCatcher-mediated nanoassembly by establishing a heptameric SpyCatcher-IMX313-

SnoopCatcher nanoparticle for dual protein display and a 60-mer SpyCatcher-mi3 engineered nanocage, both for immunisation (Brune et al., 2017; Bruun et al., 2018).

Other groups also created more nanoparticles for the decoration of proteins via the SpyTag/SpyCatcher technology, including using a 24-mer ferritin for vaccination (Wang *et al.*, 2019), or other 60-mers using encapsulin for cell imaging (Bae et al., 2018), E2 core of a pyruvate dehydrogenase complex for mAb quantification (Swartz and Chen, 2018), parvovirus B19 capsid protein for nanomedicine (Cayetano-Cruz et al., 2018), or lumazine synthase for vaccination (Krebs et al., 2019) and nanoreactors (Choi et al., 2018). Another 180-mer from a MS2 phage capsid was used for nanoreactors (Giessen and Silver, 2016). SpyTag-AP205 was independently explored by another team for vaccination (Thrane et al., 2016). The generation of more SpyCatcher-based oligomeric nanoassemblies with varying defined valencies can be beneficial for further exploration into protein space.

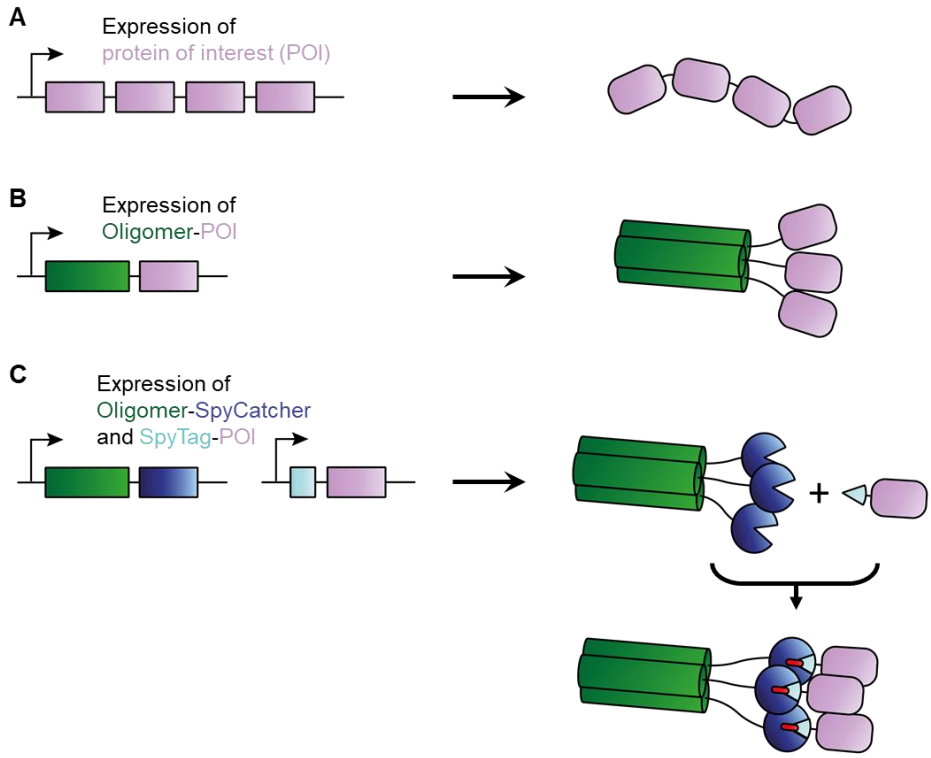


Figure 1.6 Strategies to multimerise proteins. (A) Tandem genetic fusion. A tandem repeat of units of the protein of interest (POI) is cloned into a nucleotide vector linked by linkers. Expression of the tandem genetic fusion results in a linear end-to-end arrangement of POI. (B) Genetic fusion to a multimerising domain. A single copy of the POI is genetically fused to a multimerising domain, whereby upon expression, the multimerising domain will self-assemble and drive multimerisation of the POI. (C) Modular POI assembly using a multimerising domain. Similar to (B), the multimerisation of POI is driven by the self-assembly of the multimerising domain. However, the cloning and expression of the POI and the multimerising domain are separated to enable a modular post-translational assembly. In this example, the POI is tagged with a SpyTag and the multimerising domain is fused to a SpyCatcher, which can covalently bind post-expression.

1.4 Thesis objectives

The first objective of this thesis was to develop a purification method utilising SpyTag to enable a minimal and streamlined approach for downstream SpyCatcher-mediated conjugation reactions, avoiding the need for any extraneous affinity tags. The initial focus was to match the purity of purified SpyTag-proteins compared to the current most widely used approach, by engineering SpyCatcher to non-covalently bind SpyTag and release upon a

change in buffer conditions. The generality of the approach would have to be tested on various proteins, different SpyTag localisation (N-terminal, C-terminal or internal) and from different expression systems. The subsequent focus was to improve the purification condition to enable a milder purification process for potentially sensitive proteins, particularly by exploring pH-induced release of SpyTag003 from a non-reactive SpyCatcher mutant.

The second objective was to generate a valency-defined protein oligomerisation toolbox by fusing a panel of coiled coils with SpyCatcher002 and characterising its assembly with and without SpyTag-proteins. With the Spy oligomerisation toolbox, protein multimerisation can be achieved in a stepwise and modular manner. The focus was to develop optimised SpyCatcher002-oligomers to modularly multimerise proteins, particularly to assess the stoichiometric threshold for activation of surface receptor-mediated intracellular signalling.

2 Methods

2.1 Cloning

2.1.1 Polymerase Chain Reaction (PCR)

For DNA amplification, PCR mixes were prepared in 25 µL final volume that contained 10 pg of DNA plasmid template, 0.5 µM of forward and reverse primers purchased lyophilised from Integrated DNA Technologies (IDT), and 12.5 µL of 2× Q5 high-fidelity DNA polymerase master mix (New England Biolabs, NEB). The standard PCR conditions in a C1000 Thermal Cycler (Bio-Rad) were set as: initial denaturation at 95 °C for 60 s, 30 cycles of denaturation at 95 °C for 30 s, primer annealing at 60 °C for 15 s and elongation at 72 °C for 30 s per kb of amplicon size, and one final elongation step at 72 °C for 5 min.

2.1.2 Agarose gel electrophoresis and gel extraction

Whole PCR mixture samples of 25 µL were mixed with 5 µL of 6× DNA loading dye (1.2 M sucrose, 0.7 mM bromophenol blue, 0.9 mM xylene cyanol FF) and loaded on 0.7-1.5% (w/v) agarose gels, depending on the expected amplicon size, containing 0.01% (v/v) SYBR Safe (Thermo Fisher Scientific). Running conditions were 130 V for 45 min in TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA pH 8.0). Agarose gels were imaged using a ChemiDoc XRS+ Imager (Bio-Rad) with an XcitaBlue Conversion Screen and Filter (Bio-Rad). Gel images were analysed using Image Lab version 6.0.1 (Bio-Rad). To purify PCR products, the bands of interest were excised, and DNA fragments extracted using Wizard SV Gel and PCR Clean-Up Kit (Promega). DNA concentrations were determined using ND-1000 Nanodrop (NanoDrop).

2.1.3 Gibson isothermal assembly

Gibson isothermal assembly of PCR products to make a circular plasmid was performed using the standard Gibson assembly protocol (Gibson et al., 2009). Phusion DNA polymerase was cloned, expressed and purified by Karl Brune in the Howarth lab. Taq DNA ligase and T5 exonuclease were bought from NEB. 2× Gibson assembly master mix was made and stored as 10 µL aliquots at -20 °C. PCR fragments extracted and purified from agarose gels or gBlocks gene fragments purchased from IDT were mixed on ice at equimolar mass ratio to a final volume of 10 µL and added to the 2× Gibson assembly master mix. Samples were incubated for 1-3 h at 50 °C in C1000 Thermal Cycler.

2.1.4 *E. coli* transformation

8 µL of newly constructed plasmids from Gibson assembly were transformed into 100 µL of heat competent *E. coli* NEB Turbo. Heat shock was performed at 42 °C for 45 s in a water bath followed by incubation on ice for 2 min. Alternatively, 2 µL of Gibson assembly mixes were electroporated into 50 µL of electrocompetent XL1-Blue cells. Transformed cells were grown at 37 °C for 1 h before plating on an LB agar plate with the appropriate antibiotic (100 µg/mL ampicillin, 50 µg/mL kanamycin or 34 µg/mL chloramphenicol) overnight at 37 °C.

2.1.5 Plasmid preparation

Colonies from overnight grown LB plates were picked and grown in 5-10 mL LB overnight. The grown cells were spun down at 4,000 *g* for 10 min at 4 °C. Plasmid from the pelleted cells were prepared using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific). For maxiprep, 200 mL of culture was grown, pelleted at 4000 *g* and plasmid was extracted using ZymoPURE II Plasmid Maxiprep Kit (Zymo Research). Plasmid concentration was determined using ND-1000 Nanodrop.

2.1.6 Plasmid sequence verification

For all constructs, the genes of interest were sequenced with Sanger sequencing by Source BioScience (Nottingham, UK) or Eurofins Genomics (Ebersberg, Germany). DNA sequence was validated by sequence alignment using Nucleotide BLAST (NCBI).

2.2 Protein expression and purification

2.2.1 Media and growth conditions

For *E. coli* grow-ups, media were obtained from the Media Glassware Sterilisation Laboratory of the Department of Biochemistry University of Oxford. LB consisted of 10 g tryptone, 5 g yeast extract and 10 g NaCl dissolved in Milli-Q water to a final volume of 1 L. Autoinduction media (AIM) in LB were purchased from Formedium and consisted of 10 g tryptone, 5 g yeast extract, 3.3 g ammonium sulfate, 6.8 g potassium dihydrogen phosphate, 7.1 g disodium hydrogen phosphate, 0.5 g glucose, 2.0 lactose, 0.15 g magnesium sulfate and 0.03 g trace elements dissolved in Milli-Q water to 1 L. All media were sterilised by autoclaving for 20 min at 121 °C. Bacterial cultures were grown in an Infors Multitron shaker (Infors UK Ltd) at 200 rpm.

For mammalian grow-ups, Dulbecco's Modified Eagle's Medium high glucose for HEK293T cells were purchased from Sigma-Aldrich and Expi293 Expression Medium for Expi293HEK cells were purchased from Thermo Fisher Scientific. HEK293T cells were grown in a stationary incubator and Expi293HEK cells were grown in a shaking incubator at 110-125 rpm.

2.2.2 Protein expression

2.2.2.1 Bacterial expression

Proteins were expressed in *E. coli* BL21 (DE3) RIPL (Agilent), except for scPvuII-SpyTag in *E. coli* T7 Express *lysY/I*^{q} (NEB), αDR5-SpyTag in *E. coli* BL21 (DE3) RIPL

containing a gene encoding phosphogluconolactonase to reduce protein gluconoylation (Aon et al., 2008), and all derivatives of SpyCatcher proteins in C41 (DE3) (a kind gift from Anthony Watts, University of Oxford). Single colonies from overnight grown LB plates were picked and grown in 10 mL LB overnight at 200 rpm 37 °C with the appropriate antibiotic concentrations (100 µg/mL ampicillin, 50 µg/mL kanamycin or 34 µg/mL chloramphenicol). The 10 mL culture was added to 1 L of LB containing 0.8% (w/v) glucose and the appropriate antibiotic. Growth of SpyCatcher002-oligomers was not supplemented with glucose. The cells were grown shaking at 200 rpm 37 °C to A_{600} of 0.5-0.6 and the expression was induced with 0.42 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Fluorochem) for 4 h at 30 °C. For all SpyCatcher002-oligomers except SpyCatcher002-oHex, SpyCatcher002-oHept and SpyCatcher002-CC-Hept, the expression was set at 22 °C for 16 h. For SpyCatcher002-oHex, SpyCatcher002-oHept and SpyCatcher002-CC-Hept, the expression was set at 37 °C for 4 h to drive inclusion body formation.

For αDR5-SpyTag and all histidine mutants of SpyCatcher003 S49C E77A, a single colony was picked, inoculated into 1 L of AIM media with the appropriate antibiotic concentration, and incubated at 30 °C shaking at 200 rpm for 24 h. All *E. coli*-expressed proteins were pelleted by centrifugation at 4,000 *g* for 15 min at 4 °C prior to purification.

2.2.2.2 Mammalian expression

EpCAM-SpyTag was expressed in HEK293T cells by culturing in T175 adhesive culture flaks (Corning) with Dulbecco's Modified Eagle Media (DMEM) high glucose with 10% (v/v) foetal bovine serum (Sigma-Aldrich), 2 mM L-glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin (Thermo Fisher Scientific) at 37 °C with 5% CO_2. Prior to DNA transfection, HEK293T cells were seeded into a T875 5-layer flask (Corning) and upon 50% confluency, the media was changed to DMEM, 2 mM glutamine, 50 U/mL penicillin, 25 mM HEPES. During transfection, 30 µg plasmid per 7.5 mL media for each flask layer was added. After 15 min, 2.5 mL media containing 36 µg/mL polyethyleneimine (Sigma-Aldrich) was

added to each layer. 10 mL media containing 4.4 µM valproic acid (Sigma-Aldrich), 100 U/mL penicillin, and 100 µg/mL streptomycin was added to each layer 16–20 h later. Cells were then incubated at 37 °C with 5% CO_2 for another 6 days before harvesting.

CyRPA-SpyTag was expressed in Expi293HEK cells by culturing in Expi293 Expression Media with 50 U/mL penicillin/streptomycin at 37 °C with 7% CO_2 shaking at 110–125 rpm. Transient transfection of the plasmid was done using the ExpiFectamine 293 transfection kit (Thermo Fisher Scientific). Cells were at a density of 2.5×10^6 cells/mL before being transfected with 2.7 µL ExpiFectamine 293 Reagent per 1 µg of plasmid. After 16–18 h, ExpiFectamine transfection enhancers (Thermo Fisher Scientific) were added and the cell supernatant was harvested 4 days post-transfection.

The cell supernatants were harvested by addition of cOmplete™, Mini, EDTA-free protease inhibitor cocktail (Roche), centrifuged at 1000 *g* for 3 min, and filtered through a 0.45 µm syringe filter to remove cell debris. 2.5% of 10× Ni-NTA (500 mM Tris-HCl, 3 M NaCl pH 7.8) or 10× Tris-phosphate buffer (250 mM orthophosphoric acid adjusted to pH 7.0 with Tris base) compared to the volume of cell supernatant was added for pH adjustment before affinity chromatography purification.

2.2.3 Ni-NTA purification of soluble His-tag-proteins

Protein purifications were done at 4 °C throughout, unless otherwise stated. *E. coli*-grown pellets were resuspended in 1× Ni-NTA buffer (50 mM Tris-HCl, 300 mM NaCl pH 7.8); any proteins with disulfide bonds were reduced by addition of 10 mM 2-mercaptoethanol. The resuspended pellets were supplemented with protease inhibitor cocktail, 1 mM phenylmethylsulfonyl fluoride (PMSF), 100 µg/mL lysozyme (Sigma-Aldrich) and 2 U/mL benzonase (Sigma-Aldrich). The lysates were left for rotation at 25 °C for 30 min and further lysed by sonication on ice for 4× 1 min with 1 min rest period at 50% duty cycle using a Sonics Vibra-Cell sonicator (Ultrasonic Processor). Cell lysates were centrifuged at 30,000 *g* for 30 min and incubated with Ni-NTA beads (Qiagen) at 1.5 mL of bead per litre of culture on a

rotary shaker for 1 h. Samples were applied to polyprep gravity columns (Bio-Rad) and the flow-through collected. The resin was washed with 20 packed resin volumes of Ni-NTA wash buffer (10 mM imidazole in Ni-NTA buffer, pH 7.8; proteins with disulfide bonds supplemented with 10 mM 2-mercaptoethanol for the first 10 packed resin volumes; SpyCatcher002-oligomers used 75 mM imidazole in Ni-NTA buffer) and eluted with Ni-NTA elution buffer (200 mM imidazole in Ni-NTA buffer, pH 7.8; SpyCatcher002-oligomers used 350 mM imidazole in Ni-NTA buffer). Elutions were monitored by A_{280} and stopped once A_{280} was <1.0. Proteins were dialysed against 20 mM Tris-HCl pH 8.0 using Spectra/Por 3 3,500 Da molecular weight cutoff (MWCO) membrane and concentrated, if necessary, using Vivaspin centrifugal concentrator 5 kDa cutoff (GE Healthcare).

2.2.4 Ion exchange purification of SpyDock

After Ni-NTA purification, SpyDock was further purified through ion exchange on a HiTrap Q HP anion exchange chromatography column (GE Healthcare) connected to an ÄKTA Pure 25 (GE Healthcare) fast protein liquid chromatography (FPLC) system at 4 °C. SpyDock was eluted with a linear salt gradient of 0.2–0.35 M NaCl in 10 mM Tris-HCl pH 8.0 with 1 mM dithiothreitol at a flow-rate of 2 mL/min. Peak fractions were verified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, dialysed against 20 mM Tris-HCl pH 8.0, and concentrated using Vivaspin centrifugal concentrator 5 kDa cutoff.

2.2.5 Denaturing purification of insoluble His-tag-proteins

SpyCatcher002-oHex, SpyCatcher002-oHept and SpyCatcher002-CC-Hept were found in the insoluble fraction. Cells were resuspended in 100 mM Tris-HCl pH 8.0 supplemented with protease inhibitor cocktail, 1 mM PMSF, 100 µg/mL lysozyme and 2 U/mL benzonase. Cells were lysed by rotation at 25 °C for 30 min and one freeze-thaw cycle from -80 °C to 25 °C, followed by addition of 0.5% (v/v) Triton X-100 and sonication on ice for 4× 1 min with 1 min rest period at 50% duty cycle. The cell lysate was pelleted by centrifugation

at 30,000 *g* for 30 min and washed with 2× PBS + 0.01% (v/v) Triton X-100 and then 2× PBS, with centrifugation at 25,000 *g* for 20 min between steps. The washed inclusion body pellet was resuspended in 8 M urea in 50 mM Tris-HCl pH 8.0 and refolded by diluting 50-fold into refolding buffer of 0.1 M Tris-HCl pH 8.0, 0.4 M L-arginine and 0.1 M PMSF (for SpyCatcher002-oHept, 5 mM reduced glutathione and 0.5 mM oxidized glutathione were present) for 40 h. The refolding mixture was filtered through a 0.45 µm membrane before incubation with Ni-NTA beads on a rotary shaker for 1 h. The lysate-bead mixture was added onto a polyprep gravity column and washed with 20 packed resin volumes of 75 mM imidazole in Ni-NTA buffer (SpyCatcher002-oHept had 10 mM 2-mercaptoethanol for the first 10 packed resin volumes). Elution was with 350 mM imidazole in Ni-NTA buffer. Elutions were monitored by A_{280} and stopped once A_{280} was <1.0. Proteins were dialysed against 20 mM Tris-HCl pH 8.0 using Spectra/Por 3 3,500 Da MWCO membrane and concentrated to suitable working concentrations, if needed, using Vivaspin centrifugal concentrator 5 kDa cutoff (GE Healthcare).

2.2.6 Gel filtration purification of SpyCatcher002-oligomers

After Ni-NTA purification, all SpyCatcher002-oligomers were further purified by gel filtration chromatography on a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) connected to an ÄKTA Pure 25 FPLC system at 4 °C. The mobile phase was 50 mM Tris-HCl, 150 mM NaCl pH 8.0 which ran at a flow-rate of 1 mL/min with A_{280} monitored throughout. Peak fractions were collected for each SpyCatcher002-oligomer, verified by SDS-PAGE gels and concentrated using a Vivaspin centrifugal concentrator 5 kDa cutoff if necessary. Gel Filtration Standard (Bio-Rad, catalog number 151-1901) was run at 1 mL/min at 4 °C on HiLoad 16/600 Superdex 200 pg column by Anne-Marie Andersson from the Howarth laboratory. Gel filtration run was analysed using UNICORN 7 (GE Healthcare) and presented using Microsoft Excel.

2.2.7 Spy&Go purification of soluble SpyTag-proteins

To determine the best cysteine anchoring site and E77X mutation to make SpyDock, 50 µL of resin with covalently-attached SpyCatcher2.1 E77A N-terminal cysteine, S49C and S55C, and SpyCatcher2.1 S49C E77X where X = A, D, G, N, Q, S, T ,V were mixed with 500 µL Tris-phosphate buffer (25 mM orthophosphoric acid adjusted to pH 7.0 with Tris base) containing 0.09 mg Ni-NTA-purified SpyTag-MBP (10× lower amount than available SpyCatcher variant on the resin). The protein-resin mixture was incubated at 4 °C mixing end-over-end for 1 h. The mixture was then centrifuged at 1,000 *g* for 3 min and the flow-through aspirated. The resin was washed 4× with 10 packed resin volumes (500 µL) of Tris-phosphate wash buffer (500 mM imidazole in Tris-phosphate buffer) by shaking at 1,200 rpm for 3 min at 4 °C, centrifuged at 1,000 *g* for 3 min at 4 °C and supernatant aspirated out. The protein was eluted with 4× 1.5 packed resin volumes (75 µL) of Tris-phosphate elution buffer (2.5 M imidazole in Tris-phosphate buffer) by shaking at 1,200 rpm for 5 min at 4 °C, centrifuged at 1,000 *g* for 3 min at 4 °C and eluate aspirated out. Alternatively, the protein-resin mixture mixed and incubated after 1 h was transferred to a micro spin column and the flow-through was captured in an Eppendorf tube by centrifuging the micro spin column in a microcentrifuge at 300 *g* for 30 s. The resin was then washed with Tris-phosphate wash buffer 4× 10 packed resin volumes (500 µL) and centrifuged at 300 *g* for 30 s each time. Elution was initiated by incubating the resin with 4× 1.5 packed resin volumes (75 µL) of Tris-phosphate elution buffer for 5 min at 4 °C and centrifuging at 300 *g* for 30 s each time.

2.2.7.1 Elution condition screening

To determine the elution condition, 0.09 mg of purified SpyTag-MBP in 500 µL Tris-phosphate buffer was mixed with 50 µL of Spy&Go resin for 1 h at 4 °C. The protein-resin mixture was transferred to a micro spin column and the flow-through was captured in an Eppendorf tube by centrifuging the micro spin column in a microcentrifuge at 300 *g* for 30 s. The resin was then washed with Tris-phosphate wash buffer 4× 10 packed resin volumes (500

µL) and centrifuged at 300 *g* for 30 s each time. Elution was initiated by incubating the resin with 4× 1.5 packed resin volumes (75 µL) of either (i) 50 mM glycine 150 mM NaCl at pH 2.0, (ii) 50 mM acetate 150 mM NaCl at pH 5.0, (iii) 50 mM Tris pH 9.0, (iv) 2 M imidazole in 50 mM Tris-HCl 300 mM NaCl at pH 7.8, (v) 2 M $MgCl_2$ in 20 mM Tris-HCl pH 7.0, (vi) 50% ethylene glycol (EG) in 50 mM Tris-borate at pH 10.0, (vii) 2 M NaI in 50 mM Tris at pH 8.0, or (viii) Tris-phosphate elution buffer (2.5 M imidazole in 25 mM orthophosphoric acid buffered to pH 7.0 with Tris base). Resin was incubated with different elution buffers for 5 min at 4 °C and centrifuging at 300 *g* for 30 s each time.

2.2.7.2 Washing condition

To determine the washing condition, 0.09 mg of purified SpyTag-MBP in 500 µL Tris-phosphate buffer was mixed with 50 µL of Spy&Go resin for 1 h at 4 °C. The protein-resin mixture was transferred to a micro spin column and the flow-through was captured in an Eppendorf tube by centrifuging the micro spin column in a microcentrifuge at 300 *g* for 30 s. The resin was then washed 4× with either Tris-phosphate buffer (physiological wash), alternating washes of high and low salt (high salt: 50 mM Tris 1 M NaCl at pH 7.8, low salt: 50 mM Tris without NaCl at pH 7.8), or 500 mM imidazole in Tris-phosphate buffer (Tris-phosphate wash buffer). Elution was initiated by incubating the resin with 4× 1.5 packed resin volumes (75 µL) of Tris-phosphate elution buffer for 5 min at 4 °C and centrifuging at 300 *g* for 30 s each time.

2.2.7.3 Purification from different binding conditions

0.09 mg of purified SpyTag-MBP was added into 300 µL of different binding conditions: TP buffer, PBS buffer (10 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl at pH 7.4), Ni-NTA buffer, 1% and 2% (v/v) Tween-20 in TP buffer, 1% and 2% (v/v) Triton X-100 in TP buffer, urea in 50 mM Tris-HCl pH 8.0 at 1 M, 2 M, 4 M and 6 M, and guanidine hydrochloride at pH 2.0 at 0.5 M and 1 M. SpyTag-MBP in these binding buffers was mixed

with 50 µL of Spy&Go resin and incubated with mixing end-over-end for 45 min. The mixture was centrifuged at 1,000 *g* for 5 min to settle the resin at the bottom of the tube and transferred to a pre-wetted AcroPrep Advance 0.45 µm Supor membrane 96-well plate (Pall Corporation). The plate was centrifuged at 500 *g* for 30 s and the flow-through was collected in a non-treated 96-well plate. The resins were washed with 4× 250 µL Tris-phosphate wash buffer and wash fractions were collected in non-treated 96-well plates by centrifugation at 500 *g* for 30 s. SpyTag-MBP was eluted by adding 4× 50 µL of Tris-phosphate elution buffer, incubating at 4 °C for 5 min and centrifuged at 500 *g* for 30 s to collect the elution fractions.

2.2.7.4 Purification of soluble proteins

For purification of SpyTag-MBP, SpyTag002-MBP and SpyTag003-MBP from lysate, 0.09 mg of each protein was added into ~0.25 g wet cell weight of induced *E. coli* BL21 (DE3) RIPL cleared lysate resuspended in 500 µL Tris-phosphate buffer. The protein-lysate mixture was added to 50 µL of packed Spy&Go resin and incubated at 4 °C, with mixing end-over-end for 1 h. The mixture was transferred to a micro spin column and the flow-through was captured in an Eppendorf tube by centrifuging the micro spin column in a microcentrifuge at 300 *g* for 30 s. The resin was then washed with Tris-phosphate wash buffer 4× 10 packed resin volumes (500 µL) and centrifuged at 300 *g* for 30 s each time. Elution was initiated by incubating the resin with 4× 1.5 packed resin volumes (75 µL) of Tris-phosphate elution buffer for 5 min at 4 °C and centrifuging at 300 *g* for 30 s each time.

For purification of EpCAM-SpyTag, 50 mL of the supernatant from HEK293T culture was mixed with 500 µL of Spy&Go resin. For CyRPA-SpyTag, 3.5 mL of the supernatant from Expi293HEK was mixed with 50 µL of SpyGo resin. Both supernatant-resin mixture was incubated end-over-end for 1 h at 4 °C. Proteins were purified by washing with 4× 10 packed resin volumes of Tris-phosphate buffer (5 mL for EpCAM-SpyTag, 500 µL for CyRPA-SpyTag) in a polyprep column for EpCAM-SpyTag or micro spin column for CyRPA-SpyTag. Elution was with 6× 1 packed resin volume of Tris-phosphate elution buffer for EpCAM-SpyTag and

4× 1 packed resin volume for CyRPA-SpyTag incubated for 5 min at 4 °C and centrifuging at 300 *g* for 30 s each time. For SpyTag-MBP and EpCAM-SpyTag, parallel comparison with Ni-NTA was done. The methods described above were performed with an equivalent volume of packed Ni-NTA resin and buffers, but with the following changes to buffer composition: the cell pellet was resuspended in 50 mM Tris-HCl, 300 mM NaCl pH 7.8, the wash buffer was Ni-NTA wash buffer (30 mM imidazole for SpyTag-MBP, 10 mM imidazole for EpCAM-SpyTag) and elution was carried out with Ni-NTA elution buffer.

For purification of scPvuII-SpyTag and SpyTag-mClover3, the cell pellets from 0.25 L of *E. coli* culture was lysed with 5 mL of BugBuster reagent (Merck) per gram of wet cell weight in the presence of 100 µg/mL lysozyme and 2 U/mL benzonase by incubation at 25 °C for 20 min. The lysate was pelleted by centrifugation at 30,000 *g* for 20 min to remove cell debris, and the cleared lysate was mixed with Spy&Go resin mixing end-over-end at 4 °C for 1 h. The mixtures were transferred to a micro spin column and the flow-through was captured in an Eppendorf tube by centrifuging the micro spin column in a microcentrifuge at 300 *g* for 30 s. The resin was then washed with Tris-phosphate wash buffer 4× 10 packed resin volumes (500 µL) and centrifuged at 300 *g* for 30 s each time. Elution was initiated by incubating the resin with 4× 1.5 packed resin volumes (75 µL) of Tris-phosphate elution buffer for 5 min at 4 °C and centrifuging at 300 *g* for 30 s each time.

2.2.7.5 Purification by competitive peptide elution

Lyophilised SpyTag003 peptide (RGVPHIVMVDAYKRYK) was purchased from Insight Biotechnology Ltd and dissolved in Tris-phosphate buffer to a concentration of 1 mM. 0.09 mg of purified SpyTag-MBP, SpyTag002-MBP or SpyTag003-MBP was added into ~0.25 g wet cell weight of induced *E. coli* BL21 (DE3) RIPL cleared lysate resuspended in 500 µL Tris-phosphate buffer. The protein-lysate mixture was added to 50 µL of packed Spy&Go resin and incubated at 4 °C with mixing end-over-end for 1 h. The mixture was transferred to a micro spin column and the flow-through was captured in an Eppendorf tube by centrifuging the micro

spin column in a microcentrifuge at 300 *g* for 30 s. The resin was then washed with Tris-phosphate wash buffer 4× 10 packed resin volumes (500 µL) and centrifuged at 300 *g* for 30 s each time. Each protein was eluted by incubating with 75 µL of 1 mM SpyTag003 peptide for 2 h at 30 °C and centrifuged at 300 *g* into an Eppendorf tube for 30 s. The elution was repeated another time.

2.2.7.6 Purification by tag cleavage

The cell pellet of SpyTag-TEVsite-MBP from 0.25 L of *E. coli* culture was lysed with 5 mL of BugBuster reagent (Merck) per gram of wet cell weight in the presence of 100 µg/mL lysozyme and 2 U/mL benzonase by incubation at 25 °Cfor 20 min. The lysate was pelleted by centrifugation at 30,000 *g* for 20 min to remove cell debris, and the cleared lysate was mixed with Spy&Go resin with mixing end-over-end at 4 °C for 1 h. The mixture was transferred to a micro spin column and the flow-through was captured in an Eppendorf tube by centrifuging the micro spin column in a microcentrifuge at 300 *g* for 30 s. The resin was then washed with Tris-phosphate wash buffer 4× 10 packed resin volumes (500 µL) and centrifuged at 300 *g* for 30 s each time. 0.5 mg/mL of polyhistidine-tagged SuperTEV fused to MBP (Correnti et al., 2018) was added to the resin and incubated for 2, 4, 20, and 24 h at 4 °C. Cleaved MBP was recovered by centrifuging the micro spin column at 300 *g* for 30 s.

2.2.8 Quantification of protein concentration

Protein concentrations for all proteins except SpyCatcher002-oligomers, αDR5-SpyTag, EpCAM-SpyTag and CyRPA-SpyTag were quantified through absorption at 280 nm (A_{280}) using ND-1000 Nanodrop. The concentrations were calculated using the extinction coefficients predicted by ExPASy ProtParam (Gasteiger et al., 2005). For SpyCatcher002-oligomers, αDR5-SpyTag, EpCAM-SpyTag and CyRPA-SpyTag, the concentrations were determined using bicinchoninic acid (BCA) assay according to the manufacturer's protocol (Thermo Fisher Scientific).

2.2.9 Quantification of protein purity and densitometry

Quantification of protein bands (densitometry) from SDS-PAGE gels to compare the amount of SpyTag-MBP eluted under different non-reactive SpyCatcher2.1 mutants and different wash and elution conditions was done using the Image Lab software version 6.0.1. Background subtraction was set at a disk size of 2 mm. Band intensity was obtained using the parameter adjusted volume using the equation: relative product eluted (%) = 100 × [adjusted volume of SpyTag-MBP eluted in sample lane/adjusted volume of reference SpyTag-MBP in protein control lane].

Protein purity for SpyTag-MBP, SpyTag002-MBP and SpyTag002-MBP purified through either Spy&Go or Ni-NTA was determined using gel densitometry analysis using Image Lab software version 6.0.1. Low sensitivity band detection was used to detect bands in the final eluted lane (T) and compared with the protein control lane (Protein) at a background subtraction of disk size 2 mm. Percentage purity was defined as 100 × [target protein band % in lane (T)/target protein band % in lane (Protein)].

2.3 Protein characterisation

2.3.1 SDS-PAGE

SDS-PAGE was performed using 16% Tris-glycine gels in XCell SureLock Mini-cell gel chambers (Thermo Fisher Scientific). Samples were mixed with 6× SDS loading buffer (0.23 M Tris-HCl pH 6.8, 24% (v/v) glycerol, 120 µM bromophenol blue, 0.23 M SDS) at a 6:1 ratio and heated to 95 °C for 5 min using a C1000 Thermal Cycler. For reduced samples, 100 mM 2-mercaptoethanol was added. Samples in Tris-glycine gels were run at 190 V in Tris-glycine running buffer (25 mM Tris-HCl, 192 mM glycine, 0.1% (w/v) SDS, pH 8.5) for approximately 1 h 15 min. Pre-stained PageRuler (Thermo Fisher Scientific) was used a molecular weight marker. Gels were stained with InstantBlue Coomassie stain (Expedeon), destained with Milli-Q, imaged using ChemiDoc XRS+ imager, and analysed using Image Lab version 6.0.1.

2.3.2 Mass spectrometry

Intact protein mass spectrometry was performed using an Agilent RapidFire 365 platform, coupled to an Agilent 6550 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) mass spectrometer in positive ion mode with a jet stream electrospray ion source (Agilent). Mass spectrometry was performed by Dr Anthony Tumber of the University of Oxford Department of Chemistry. Protein samples at 10 µM in 50 µL volume [2.5 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl) in 50 mM Tris-HCl with 5 mM EDTA at pH 8.5 was added to proteins with disulfide bonds for 1 h] were acidified to 1% (v/v) formic acid on a 384-well polypropylene plate (Greiner). Samples were aspirated under vacuum for 0.4 s on the RapidFire sampling platform and loaded onto a C4 solid-phase extraction cartridge. The samples were washed with 0.1% (v/v) formic acid at 1.5 mL/min flow-rate for 5.5 s and eluted to the mass spectrometer with deionised water containing 85% (v/v) acetonitrile and 0.1% (v/v) formic acid at 1.25 mL/min for 5.5 s. The cartridge was equilibrated with deionised water for 0.5 s. Nitrogen drying gas for the ionisation source was operated at 13 L/min at 225 °C, with the jet stream sheath gas at a flow-rate of 12 L/min at 350 °C and the nozzle voltage at 1500 V. Data analysis to determine the mass of protein samples was done using Mass Hunter Qualitative Analysis software version 7.0, with the protein ionisation data deconvoluted using the maximum entropy algorithm. Predicted protein mass was calculated using ExPASy ProtParam (Gasteiger et al., 2005) based on reduced disulfides and cleavage of the N-terminal formylmethionine. Data were exported and presented using Microsoft Excel.

2.3.3 Dynamic light scattering (DLS)

Samples were prepared at 1 mg/mL in 50 mM Tris-HCl, 150 mM NaCl pH 8.0. Aggregates were removed by centrifugation at 16,900 *g* for 30 min at 4 °C and aspirating the soluble samples away from the pelleted aggregates. Alternatively, samples were centrifuged through a 0.22 µm filter to remove aggregates at 16,900 *g* for 15 min at 4 °C. 20 µL of each sample was loaded into a reusable quartz cuvette and measurements were taken at 20 °C

using an Omnisizer (Viscotek) with 10 scans of 10 s duration each. Data were analysed using OmniSIZE 3.0 and intensity distributions were plotted in Microsoft Excel.

2.3.4 Size exclusion chromatography/multi-angle light scattering (SEC-MALS)

SEC-MALS runs were performed by Dr David Staunton from the University of Oxford Department of Biochemistry Biophysical Suite. Samples were prepared with 100 µL of 2 mg/mL protein before αDR5-SpyTag conjugation and 1 mg/mL after conjugation with αDR5-SpyTag due to dilution after substrate mixing. The running buffer was 50 mM Tris-HCl, 150 mM NaCl pH 8.0. Samples were injected into a Superdex 200 HR 10/30 column (GE Healthcare) at 25 °C at a flow-rate of 0.5 mL/min connected to a Shimadzu High Performance Liquid Chromatography (HPLC) system comprising LC-20AD pump, SIL-20AC autosampler and SPD20A UV/Vis detector. Light scattering was detected by a Wyatt Dawn HELEOS-II 8-angle light scattering detector and Wyatt Optilab rEX refractive index monitor. The resulting light scattering, refractive index and UV traces were processed in ASTRA 6 (Wyatt Technologies). Data were presented using GraphPad Prism v7.0.

2.3.5 Isothermal titration calorimetry (ITC)

This experiment was performed by Dr Anthony Keeble from the Howarth lab. Free SpyDock was trapped as a monomer by reduction with 2.5 mM TCEP-HCl in 50 mM Tris-HCl with 5 mM EDTA at pH 8.5 and reacting with 20 mM iodoacetamide in 50 mM Tris-HCl with 5 mM EDTA at pH 8.5 for 30 min in the dark to produce carbamidomethylated SpyDock. Excess TCEP-HCl and iodoacetamide were removed by dialysis. SpyTag-MBP, SpyTag002-MBP and carbamidomethylated SpyDock were dialysed into 20 mM HEPES, 150 mM NaCl pH 7.0 twice using SpectraPro 3 3,500 Da membrane. ITC was performed using a Microcal PEAQ-ITC calorimeter (Malvern) at 25 °C in 20 mM HEPES, 150 mM NaCl pH 7.0. 20 µM SpyDock was used in the cell and titrated with 20 injections of 210 µM SpyTag-MBP or

SpyTag002-MBP in the syringe. Analyses were carried out using a 1:1 binding model with the MicroCal PEAQ-ITC Analysis software version 1.1.0.1262.

2.3.6 Fluorescence intensity after Spy&Go purification

Following purification through either Spy&Go or Ni-NTA, a concentration gradient of SpyTag-mClover3 in PBS was prepared in triplicates from 15 µM to 0.47 µM at 25 µL in a black, flat-bottom half area 96-well plate (Corning). Fluorescence intensity ($\lambda_{ex} = 482 \pm 16$ nm, $\lambda_{em} = 530 \pm 40$ nm) was detected at 40 flashes per well using a CLARIOstar plate reader (BMG Labtech) at 30 °C. The fluorescence intensity for Spy&Go- and Ni-NTA-purified SpyTag-mClover3 was plotted using Microsoft Excel.

2.4 Coupling to solid support

Purified SpyCatcher2.1 S49C E77X, where X = A, D, G, N, Q, S, T, V, SpyCatcher2.1 N-terminal cysteine E77A, SpyCatcher2.1 S55C E77A, and SpyCatcher003 S49C E77A with different histidine mutants were coupled to SulfoLink Coupling Resin (Thermo Fisher Scientific) according to the manufacturer's protocol. Briefly, for every 1 mL of packed SulfoLink resin, 20 mg of protein was reduced by 1 mM TCEP-HCl for 30 min. The reduced samples were then mixed with SulfoLink resin pre-equilibrated with coupling buffer (50 mM Tris-HCl, 5 mM EDTA, pH 8.5) by end-over-end mixing at 25 °C for 15 min and were left stationary for 30 min. Flow-through was removed by passing through a polyprep column or micro spin column (Bio-Rad) and washing with coupling buffer for 10 packed resin volumes. The resin reactive residues were blocked with 50 mM L-cysteine-HCl (MP Biomedicals) in coupling buffer by mixing end-over-end for 15 min and being left stationary for 30 min at 25 °C. The resin was washed with 10 packed resin volumes of 1 M NaCl and stored in 1× Tris-phosphate buffer with 0.05% (w/v) NaN_3 at 4 °C.

2.5 Regeneration of Spy&Go resin

After purification from cell lysate, Spy&Go resin was equilibrated with 2× 10 packed resin volumes of Tris-phosphate buffer before regeneration. Regeneration was done at 25 °C by incubating with 3× 10 packed resin volumes of 4 M imidazole in Tris-phosphate buffer pH 7.0 for 5 min each time, washing with 1× Tris-phosphate buffer, incubating with 3× 10 packed resin volumes of 6 M guanidine hydrochloride pH 2.0 for 5 min each time, washing with 1× Tris-phosphate buffer, incubating with 3× 10 packed resin volumes of 0.1 M NaOH for 1 min each time, and washing with 2× Tris-phosphate buffer before storage in 20% (v/v) ethanol in Milli-Q water at 4 °C.

2.6 *In vitro* protein conjugation and excess protein recapture

To investigate any covalent bond formation from SpyCatcher2.1 S49C E77X mutants, 5 μM of each mutant was reacted with 2× molar excess of SpyTag-MBP (10 μM) in 30 μL total volume for 24 h at 25 °C in 1× Tris-phosphate buffer. The reaction was quenched with SDS loading buffer and heated to 95 °C for 5 min in a C1000 Thermal Cycler. The samples were then loaded onto SDS-PAGE.

To multimerise αDR5-SpyTag using SpyCatcher002-oligomers, 100 μM monomeric concentration of SpyCatcher002-oligomer was mixed with 3× molar excess of αDR5-SpyTag (300 μM) in 200 μL total volume for 16 h at 25 °C in 50 mM Tris-HCl, 150 mM NaCl, pH 8.0. Excess αDR5-SpyTag was removed by capture of free αDR5-SpyTag using 50 μL Spy&Go resin by incubation with the sample at 25 °C for 2 h with mixing end-over-end. The mixture was loaded onto a micro spin column and the conjugated SpyCatcher002-oligomer:αDR5-SpyTag was recovered by centrifugation at 300 *g* for 1 min at 4 °C into a microcentrifuge tube. This process was repeated once. The protein concentrations were measured by BCA assay.

2.7 DR5-dependent cell killing assay

This experiment was performed by Dr Anusuya Banerjee from the Howarth lab. MDA-MB-231 cells were used for this assay and were obtained from the American Type Culture Collection. The cells were grown at 37 °C with 5% CO_2 in DMEM containing 10% (v/v) foetal bovine serum (FBS), 50 U/mL penicillin and 50 µg/mL streptomycin. Cells were maintained by passaging upon confluency of 70-80% for less than 3 months in total. MDA-MB-231 cells were seeded into a 96-well plate at a concentration of 40,000 cells per well in 100 µL DMEM with 50 U/mL penicillin, 50 µg/mL streptomycin and 1% (v/v) FBS for 16 h at 37 °C with 5% CO_2. To activate DR5 signalling, each conjugated SpyCatcher002-oligomer:αDR5-SpyTag was prepared at 100 nM monomeric concentration and serial dilutions thereof in DMEM with 50 U/mL penicillin, 50 µg/mL streptomycin and 1% (v/v) FBS. SpyTag-MBP was used as a negative control in parallel. Cells were washed twice with sterilised PBS before adding the proteins. Proteins were incubated with the cells for 24 h at 37 °C with 5% CO_2.

Cell viability was measured by addition of 40 µL of 0.15 mg/mL resazurin (Sigma-Aldrich) in PBS at 37 °C with 5% CO_2 for 4 h and measuring the fluorescence (λ_{ex} 544 nm, λ_{em} 590 nm) using a SpectraMax3 plate reader with SoftMax Pro 5.4 software (Molecular Devices). Cell viability was calculated as 100 × (signal of treated cells - signal without cells)/(signal untreated cells–signal without cells). The signal without cells was taken as the resazurin fluorescence in the absence of cells, whereas the signal of untreated cells came from the fluorescence of resazurin with cells that were incubated with DMEM with antibiotics and 1% (v/v) FBS.

2.8 Rational selection of histidine mutants

The crystal structure of SpyTag:SpyCatcher (PDB ID: 4MLI) (Li et al., 2014) was manually inspected for sites to introduce histidines. The residues contributing to the binding interface between SpyTag and SpyCatcher were determined by using the open source InterfaceResidues script by Jason Vertrees (URL for script webpage:

https://pymolwiki.org/index.php/InterfaceResidues) on PyMOL using PDB 4MLI. Following the identification of two binding interface stretches in SpyCatcher, these residues were inspected for potential electrostatic positive-positive charge repulsion between itself (after substitution to charged histidine) and any adjacent positively-charged or partial positive charged residues on SpyTag and SpyCatcher within a ~5 Å radius (Donald et al., 2011) as measured computationally using PyMOL 2.0.

To target the positively-charged N-terminus of SpyTag003, which is absent in the available crystal structure, the N-terminus of SpyTag in PDB ID 4MLI was computationally extended with Arg-Gly-Val from the terminal Ala111 whilst maintaining backbone β-sheet hydrogen bonding to ~3 Å (Berg et al., 2002). SpyCatcher residues within a ~5 Å radius to the new terminal R108 side chain and main chain amine group were selected for histidine substitution. 16 residues on SpyCatcher003 were selected to be mutated to histidine presented in Table 4.1.

2.9 pH elution screening of histidine mutants

2.9.1 pH tolerance of SpyTag003-sfGFP

pH 8.0 buffer was 50 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) and 150 mM NaCl. pH 6.0 buffer was 50 mM MES (2-(N-morpholino)ethanesulfonic acid) and 150 mM NaCl. pH 5.0 and 4.0 buffers were 50 mM acetic acid (made from different combinations of glacial acetic acid and sodium acetate) and 150 mM NaCl. pH 2.0 buffer was 50 mM glycine with 150 mM NaCl. All buffers were made up to 500 mL with cold Milli-Q water and the pH corrected with HCl or NaOH while stirring in an ice bath to maintain low temperature between 4-8 °C during pH measurement. Neutralisation buffer was 1 M Tris-HCl pH 8.0 at 4 °C. To investigate the amount of neutralisation buffer needed to neutralise buffers of pH 2.0, 4.0, 5.0 or 6.0 to approximately 7.0, 10 mL of each buffer was added to a 50 mL tube with a small magnetic stirrer and placed in a beaker with ice. 1 M Tris-HCl pH 8.0

neutralisation buffer was titrated to each buffer sample whilst monitoring the change in pH and the volume of neutralisation buffer added was recorded until the pH was approximately 7.0.

To investigate the tolerance of purified SpyTag003-sfGFP at low pH, 5 µL of 80 µM SpyTag003-sfGFP was added to 50 µL of pH 2.0, 4.0, 5.0 or 6.0 buffers and pre-neutralised pH 2.0, 4.0, 5.0, or 6.0 buffers. The samples were left to incubate at 4 °C for 1 h. Neutralisation buffer was added at a volume of 23 µL for pH 2.0 buffer, 20 µL for pH 4.0 buffer, 10 µL for pH 5.0 buffer and 23 µL for pH 6.0 buffer. For pre-neutralised samples, pre-neutralised pH 2.0, 4.0, 5.0, or 6.0 buffers were added using the same volumes (23 µL for pH 2.0 buffer, 20 µL for pH 4.0 buffer, 10 µL for pH 5.0 buffer and 23 µL for pH 6.0 buffer) to match the dilution factors. Samples were left to incubate for 30 min at 4 °C to allow full refolding of SpyTag003-sfGFP. The final volumes for all samples were stratified by adding pre-neutralised buffers to samples in pH 5.0 (13 µL) and pH 4.0 (3 µL) to match final volumes of pH 2.0 and pH 6.0. For the “no neutralisation” controls, 23 µL of relevant buffers were added to each pH buffer. 20 µL of each sample was transferred to a black, flat-bottom half area 96-well plate. Peak fluorescence intensity at $\lambda_{ex} = 485 \pm 12$ nm, $\lambda_{em} = 520 \pm 20$ nm, excitation spectra from $\lambda_{ex} = 420 \pm 10$ nm to 515 ± 10 nm with $\lambda_{em} = 543 \pm 16$ nm, and emission spectra from $\lambda_{em} = 490 \pm 10$ nm to 560 ± 10 nm with $\lambda_{ex} = 462 \pm 16$ nm were detected using a CLARIOstar plate reader at 30 °C.

2.9.2 Stepwise pH elution of SpyTag003-sfGFP

0.05 mg of purified SpyTag003-sfGFP (17.5 µL at 93 µM) was added to 30 µL of different SulfoLink-coupled SpyCatcher003 S49C E77A histidine mutant resins and the volume was made up to 300 µL with pH 8.0 buffer in an Eppendorf tube. The mixture was left to incubate mixing end-over-end at 4 °C for 45 min in the dark. The mixture was centrifuged at 1,000 *g* for 5 min to settle the resin at the bottom of the tube. The mixture was transferred to a pre-wetted AcroPrep Advance 0.45 µm Supor membrane 96-well plate (Pall Corporation) and centrifuged at 500 *g* for 30 s and the flow-through was collected in a non-treated 96-well

plate. The resins were washed with 4× 250 µL pH 8.0 buffer and wash fractions were collected in non-treated 96-well plate by centrifugation at 500 *g* for 30 s. Stepwise elution by decreasing pH buffers was initiated by incubation for 10 min with 2× 50 µL of buffers from pH 6.0, 5.0, 4.0 to 2.0. Before collecting the eluate in the collection plate, neutralisation buffer was added to the wells (23 µL for pH 6.0, 10 µL for pH 5.0, 20 µL for pH 4.0 and 23 µL for pH 2.0). Final volumes were stratified by adding pre-neutralised buffers to samples in pH 5.0 (13 µL) and pH 4.0 (3 µL) to match final volumes of pH 2.0 and pH 6.0. 20 µL of each sample was transferred to a black, flat-bottom half area 96-well plate. Peak fluorescence intensity at $\lambda_{ex} = 485 \pm 12$ nm, $\lambda_{em} = 520 \pm 20$ nm was detected using a FLUOstar plate reader at 30 °C (Figure 2.1). Readings were analysed and plotted in Microsoft Excel.

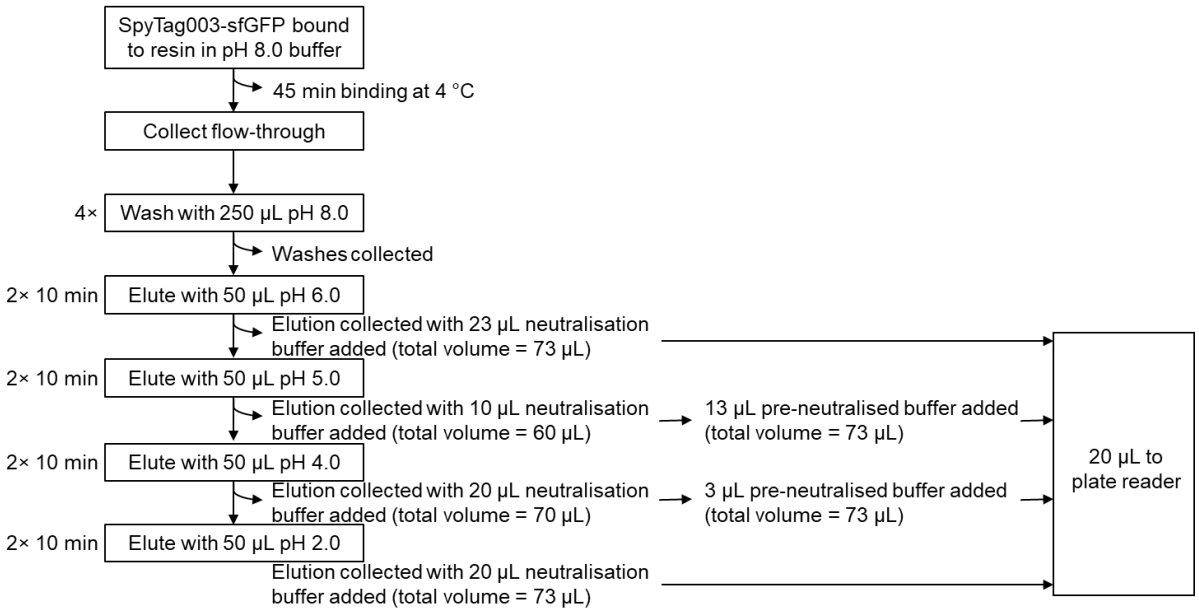


Figure 2.1 Flow-chart of stepwise pH elution of SpyTag003-sfGFP. The process started with SpyTag003-sfGFP bound to different SpyCatcher003 S49C E77A His mutants at pH 8.0, washed and subsequently eluted from pH 6.0 to pH 2.0. Elution fractions at each step was collected and neutralised immediately. Elution fractions from pH 5.0 and pH 4.0 were additionally added with pre-neutralised buffer to make the final volume and dilution factor uniform. All steps were carried out at 4 °C except fluorescence reading using the plate reader at 30 °C.

For the pH elution of SpyTag003-sfGFP from cell lysate, 0.09 mg of SpyTag003-sfGFP was added into ~0.25 g wet cell weight of induced *E. coli* BL21 (DE3) RIPL cleared lysate resuspended in 500 µL 50 mM HEPES pH 8.0 and mixed with 50 µL of SpySwitch resin. The mixture was left to incubate at 4 °C for 45 min. The mixture was transferred into a micro spin

column and flow-through was collected by spinning at 300 *g* for 30 s. The resin was washed 4× with either 50 mM HEPES with 150 mM NaCl at pH 8.0 (physiological wash), alternating washes of high and low salt (high salt: 50 mM HEPES and 1M NaCl pH 8.0, low salt: 50 mM HEPES pH 8.0), or 500 mM imidazole in 50 mM HEPES and 150 mM NaCl pH 8.0. Elution was initiated by incubating 3× 50 µL of pH 5.0 buffer for 10 min per incubation period. The eluate was collected in a tube with 10 µL of neutralisation buffer added. Samples were mixed with SDS loading buffer but not boiled and loaded onto a 16% SDS-PAGE gel. Before Coomassie staining, SpyTag003-sfGFP fluorescence was imaged using ChemiDoc XRS+ Imager (Bio-Rad) with an XcitaBlue Conversion Screen and Filter (Bio-Rad).

3 Development of SpyDock for Spy&Go purification of SpyTag-proteins

The SpyTag/SpyCatcher technology has advantages for covalent conjugation of proteins due to its genetic-encodability and good specificity (Keeble and Howarth, 2019). The simplicity of the platform and ability to react under various conditions (Keeble et al., 2017; Zakeri et al., 2012) enabled the development of over 200 applications in different fields, from immunology to materials (Keeble and Howarth, 2019).

A promising aspect of the Tag/Catcher technology is in the area of rapid and scalable vaccine generation through decoration of virus-like particles, which increase the efficiency of antibody induction to an antigen of interest (Aves et al., 2020; Gomes et al., 2017). Direct genetic fusion of the antigen to the scaffold particle can be done to create a vaccine for immunisation, but often leads to misfolding of the antigen or capsid (Aves et al., 2020). Modular antigen display through SpyTag/SpyCatcher can overcome this challenge (Brune et al., 2016; Thrane et al., 2016). With this method, the individual optimisation usually needed for making a virus-like particle-based vaccine against a specific antigen is alleviated, since the antigen can be optimally produced separately from the immunogenic scaffold (Brune and Howarth, 2018). A SpyTag-fused antigen can be modularly displayed on a multimeric SpyCatcher-VLP to elicit effective immune responses against the antigens. Different SpyCatcher-based nanoparticle scaffolds are already available for various purposes such as vaccination, nanomedicine and enzyme cascades, which illustrates the adaptability of this modular approach (Brune et al., 2017; Bruun et al., 2018; Escolano et al., 2019; Leneghan et al., 2017; Ma et al., 2018; Wang et al., 2019a).

In a SpyTag/SpyCatcher-based vaccine generation process against malaria, the antigen, SpyTag-CIDR, was purified through polyhistidine tag IMAC and required tag cleavage to remove the extraneous affinity tag sequence (Brune et al., 2016). These steps were performed to eliminate the generation of anti-His-tag antibodies and reduce toxic metal ion leaching from

the immobilised metal affinity chromatography column, in accordance with regulations for making biopharmaceuticals (Bornhorst and Falke, 2000; Gaberc-Porekar and Menart, 2005; Jin et al., 2017; Kuo and Chase, 2011; Watson et al., 2011). However, tag cleavage is laborious and entails more production cost, more steps and reduction of protein yield (Waugh, 2005).

A different short epitope tag, C-tag with a tetrapeptide sequence E-P-E-A, was also used, based on capture by a camelid single-chain antibody. The short tag induces minimal antibody response, along with high recovery and purity, making it a suitable tag for antigen purification without tag cleavage (Jin et al., 2017). Nonetheless, the high price of the commercial resin and its persisting non-functional presence after purification are not desirable. It is better to limit the number of functionally-necessary peptide motifs accompanying the antigen protein.

Hence, my aim was to develop a new affinity chromatography system employing SpyTag that can double as a peptide affinity tag. This reduces the complexity of oligonucleotide design and metal leaching, and removes the need for downstream tag cleavage to simplify work-flow for SpyCatcher-based bioconjugation reactions such as for modular vaccine generation. In this chapter, I describe the development of Spy&Go, a process for the purification of SpyTag-fusions from complex cell mixture. I show the engineering of SpyCatcher to generate a high affinity, non-covalent, reversible binder to SpyTags, termed SpyDock, for solid-phase purification of different proteins. I also describe the characterisation of the resin for daily and long-term use.

3.1 Engineering SpyCatcher for reversible binding with SpyTag

SpyTag and SpyCatcher already have an inherent specificity for each other, thus I decided to use SpyCatcher as the base to engineer a reversibly-binding protein partner for SpyTag. Abolished covalent binding between SpyTag and SpyCatcher was already shown by making non-reactive SpyTag (reactive Asp mutated to Ala; SpyTag D117A) or non-reactive SpyCatcher (reactive Lys mutated to Ala; SpyCatcher K31A, or catalytic Glu mutated to Gln;

SpyCatcher E77Q) (Bedbrook et al., 2015; van den Berg van Saparoea et al., 2018; Hagan et al., 2010; Keeble et al., 2017; Schoene et al., 2014; Zakeri et al., 2012). SpyCatcher and SpyTag D117A were also shown through ITC to have a K_d of 0.2 μM; although the covalent linkage has been abolished, this pair can still associate reversibly with high affinity (Zakeri et al., 2012). Nonetheless, rational engineering to find the proper mutations to bind SpyTag and elute effectively was needed to develop a SpyCatcher-based affinity chromatography system.

3.1.1 Covalent bond-abolishing mutants

On SpyCatcher, two amino acid residues are directly involved in the formation of the isopeptide bond with SpyTag, namely the reactive K31 and catalytic E77. Mutating either of this residue has been shown to abolish covalent reaction (Hagan et al., 2010). However, for the purposes of this project, mutating K31 is not pursued as the ε-amino group may form favourable interactions with the reactive D117 from SpyTag even without forming the isopeptide bond. This interaction may contribute to the specificity and stability to bind SpyTag, thus mutation at this site was not pursued to abolish covalent linkage between SpyTag and SpyCatcger. Instead, mutation at the E77 residue was prioritised.

Although SpyCatcher's reactivity has commonly been knocked out using the E77Q mutation (Keeble et al., 2017; Zakeri et al., 2012), this substitution may not be the best mutation to facilitate tight and reversible binding of SpyTag. Eight small amino acids were selected to mutate SpyCatcher's catalytic E77 residue, namely alanine, aspartic acid, glycine, asparagine, glutamine, serine, threonine, and valine. I decided to incorporate these mutations into the latest interim SpyCatcher iteration, SpyCatcher2.1, to promote the best association with SpyTag since SpyCatcher2.1 had the highest rate of isopeptide bond formation compared to SpyCatcher and SpyCatcher002. SpyCatcher2.1 was developed by Dr Anthony Keeble and Mr Alberto Carella from the Howarth lab, building on its predecessor, SpyCatcher002 (Keeble et al., 2017). SpyCatcher2.1 contained three extra mutations obtained by rational design through manual inspection of the SpyTag:SpyCatcher crystal structure (PDB: 4MLI). The

mutations included two substitutions, Q97D and K108E, to complement the increase in positively-charged residues on SpyTag002 and SpyTag003 (Keeble et al., 2019). The third mutation is a loop stabilising substitution, A89P, that was designed to restrict the flexibility of the long A79-A89 loop that contained two important residues, Y83 and E85, that make key interactions with SpyTag at Y119 and K120 (Keeble et al., 2019; Li et al., 2014). Alongside SpyCatcher2.1, a newer SpyTag was also made called SpyTag003 that lengthened SpyTag002 at the N-terminus with the addition of Arg-Gly (Keeble and Howarth, 2019). The three mutations on SpyCatcher2.1 and N-terminal extension on SpyTag003 increased the second order rate constant of reaction to 5.9×10^4 M^{-1} s^{-1}, making the pair three-fold faster than SpyTag002:SpyCatcher002 (2.0×10^4 M^{-1} s^{-1}) (Keeble et al., 2019). At the start of this project, the upgrade from SpyCatcher002 was still being developed, hence SpyCatcher2.1 was the interim SpyCatcher iteration before the finalised development of SpyCatcher003 (two further mutations on top of SpyCatcher2.1) (Keeble et al., 2019).

Among the eight selected SpyCatcher2.1 E77X mutants (where X = A, D, G, N, Q, S, T, or V), all mutations managed to abolish the isopeptide bond formation with SpyTag-MBP after both components were reacted for 24 h at 25 °C, with the exception of E77D (Figure 3.1 A). Since aspartic acid differs from glutamic acid in having one fewer methylene group, SpyCatcher2.1 E77D was expected to still catalyse the formation of the isopeptide bond. However, even after 24 h there was only a fraction of covalent product formed between SpyCatcher2.1 E77D and SpyTag-MBP, reflecting the stringent geometric requirements of the isopeptide-forming residues (Kang and Baker, 2011). Thus, the E77D mutation was no longer pursued for this project.

Thereafter, each SpyCatcher2.1 E77X mutant was covalently coupled onto agarose beads via the introduction of a cysteine to anchor the protein through iodoacetyl-sulfhydryl chemistry. Initially, the cysteine was introduced into SpyCatcher2.1 at the solvent-exposed S49 residue since the residue was previously shown to be amenable to maleimide-sulfhydryl chemistry through the complete labelling of SpyCatcher with a fluorescent probe (Dovala et

al., 2016). After immobilisation to the solid-phase, each SpyCatcher2.1 S49C E77X mutant was subjected to bind and elute the same amount of pure SpyTag-MBP. 50 µL of resin with different non-covalent SpyCatcher variants bound 0.09 mg of pure SpyTag-MBP in 500 µL of Tris-phosphate buffer at pH 7.0 (25 mM orthophosphoric acid buffered with Tris base to pH 7.0). Among the mutants, SpyCatcher2.1 S49C E77A captured and eluted the most SpyTag-MBP (Figure 3.1 B). Thus I took the E77A mutation forward as the covalent bond-abolishing mutation.

The investigation to select the isopeptide bond abolishment mutation was done on SpyCatcher2.1, since the charge-complementary mutations on SpyCatcher2.1 could provide higher affinity interactions with SpyTag to ensure minimal SpyTag-protein loss during binding and washes. To support this hypothesis, SpyCatcher and SpyCatcher002 with the E77A mutation were covalently coupled onto a solid-phase support through the S49C anchoring mutation. Under the same conditions, SpyCatcher S49C E77A achieved only 27% recovery of SpyTag-MBP compared to SpyCatcher2.1 S49C E77A, whereas SpyCatcher002 S49C E77A achieved 42% (Figure 3.1 B). This result demonstrated the requirement for stronger pairing between the immobilised non-covalent SpyCatcher and SpyTag.

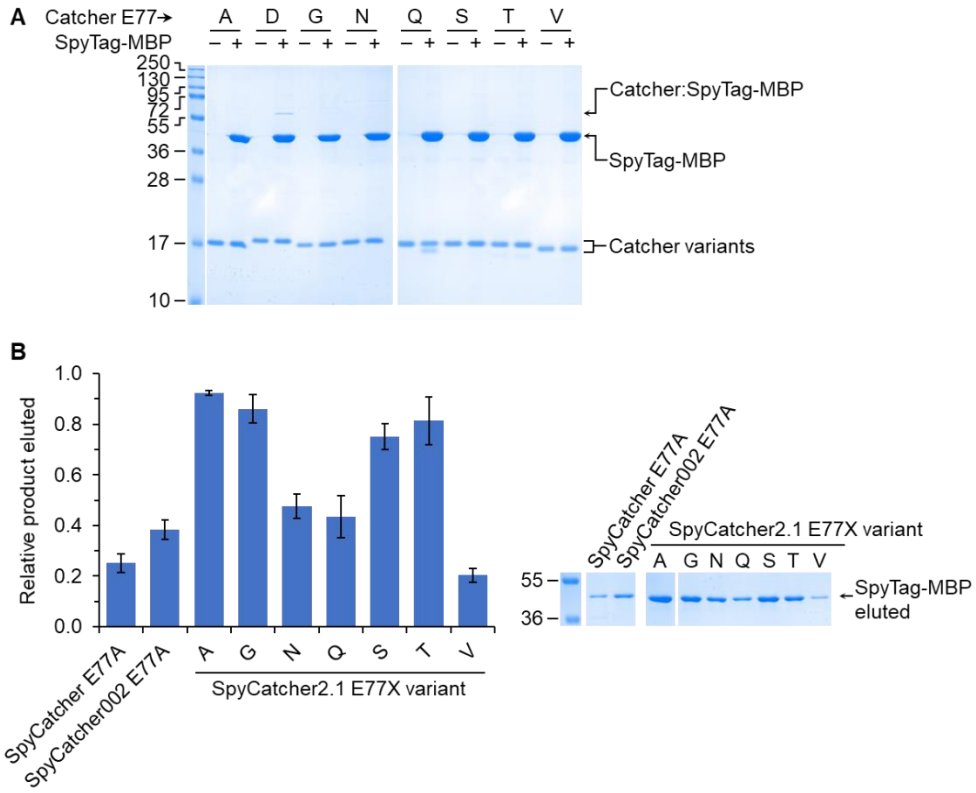


Figure 3.1 Abolishing isopeptide bond formation. (A) All single point mutations at SpyCatcher's catalytic E77 block isopeptide bond formation except E77D. 5 µM of SpyCatcher variants were mixed with 10 µM of SpyTag-MBP for 24 h at 25 °C in Tris-phosphate buffer pH 7.0. (B) Relative quantification (mean ± 1 s.d, *n* = 3) of eluted SpyTag-MBP from different SpyCatcher2.1 E77X mutants and different SpyCatcher generations through gel densitometry analysis. Purified SpyTag-MBP was mixed with immobilised SpyCatcher variants, washed with 500 mM imidazole in Tris-phosphate pH 7.0 and eluted with 2.5 M imidazole in Tris-phosphate pH 7.0. SpyCatcher2.1 E77A was identified as the candidate for SpyDock. Right panel: representative Coomassie-stained SDS-PAGE gel of pooled SpyTag-MBP elutions. Recovery of 1.0 in relative product eluted corresponds to the reference sample of input SpyTag-MBP.

3.1.2 Specific solid-phase attachment of SpyDock

The work in section 3.1.1 was done by anchoring the SpyCatcher variants using iodoacetyl-sulfhydryl chemistry onto agarose beads via the S49C mutation, as this mutation was already previously used (Dovala et al., 2016). However, S49C may not be the best immobilising mutation and so I explored other residues for this purpose.

Proteins can be chemically conjugated by exploiting certain reactive residues, most prominently lysine and cysteine (Koniev and Wagner, 2015). Due to the absence of cysteines and the abundance and importance of lysines in the SpyCatcher sequence, only covalent immobilisation using sulfhydryl chemistry was examined. Single cysteines were introduced through site-directed mutagenesis at surface-exposed positions distal from the residues involved in the formation of isopeptide bonds, namely at the N-terminus, S49, or S55 of SpyCatcher2.1 E77A (Figure 3.2 A). These positions ensure minimal perturbation to the binding of SpyTag and the non-reactive SpyCatcher2.1 E77A. Between the three cysteine anchoring mutations, S49 still prevailed as the best position to immobilise and orient SpyCatcher2.1 E77A on the solid-phase support because the most SpyTag-MBP was eluted from this anchoring mutation (Figure 3.2 B & C).

Thus, SpyCatcher2.1 S49C E77A was selected as the SpyCatcher variant of choice to become SpyDock for the Spy&Go affinity chromatography method.

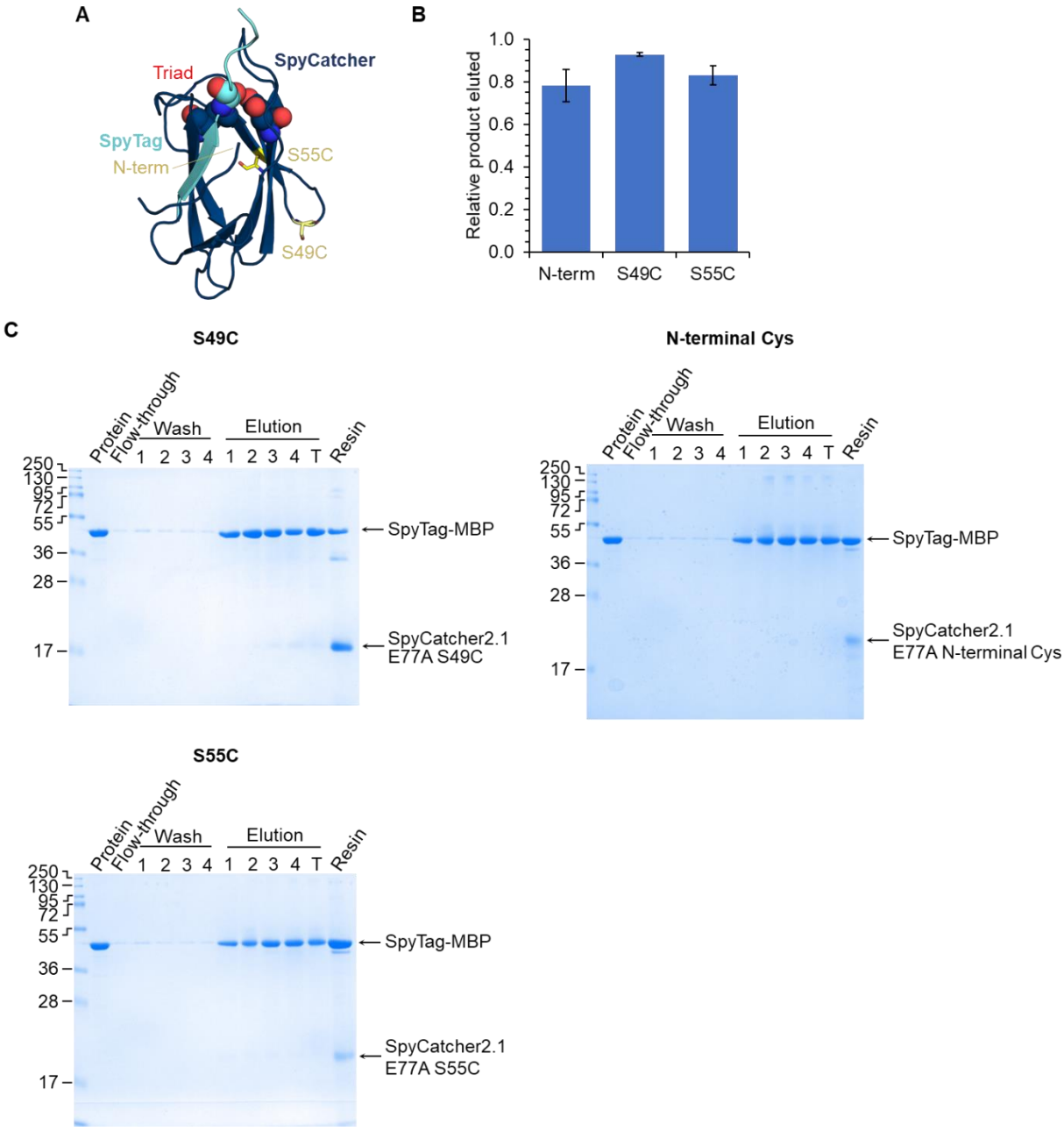


Figure 3.2 Covalent immobilisation of SpyDock to solid-phase support. (A) Positions for the introduction of single cysteines (yellow) for covalent immobilisation to solid-phase support, relative to the isopeptide triad. (B) Relative quantification (mean ± 1 s.d, *n* = 3) of eluted SpyTag-MBP from each SpyCatcher2.1 E77A cysteine anchoring mutant through gel densitometry analysis. Purified SpyTag-MBP was mixed with immobilised SpyCatcher variants, washed with 500 mM imidazole in Tris-phosphate pH 7.0 and eluted with 2.5 M imidazole in Tris-phosphate pH 7.0. SpyCatcher2.1 S49C E77A was identified as the construct for SpyDock. Representative Coomassie-stained SDS-PAGE gels are shown in (C). T: total pooled elutions. Recovery of 1.0 in relative product eluted corresponds to the reference sample of input SpyTag-MBP.

3.1.3 Binding affinity characterisation

After establishing SpyDock, its binding affinities with SpyTag and SpyTag002 were determined. As SpyDock contains a single cysteine which can self-dimerise via a disulfide bond, the protein was first trapped in its monomeric state through an irreversible reaction with iodoacetamide. The dissociation constant, K_d, was then determined by ITC by titrating either SpyTag-MBP or SpyTag002-MBP into the SpyDock sample cell. This experiment was performed by Dr Anthony Keeble from the Howarth lab.

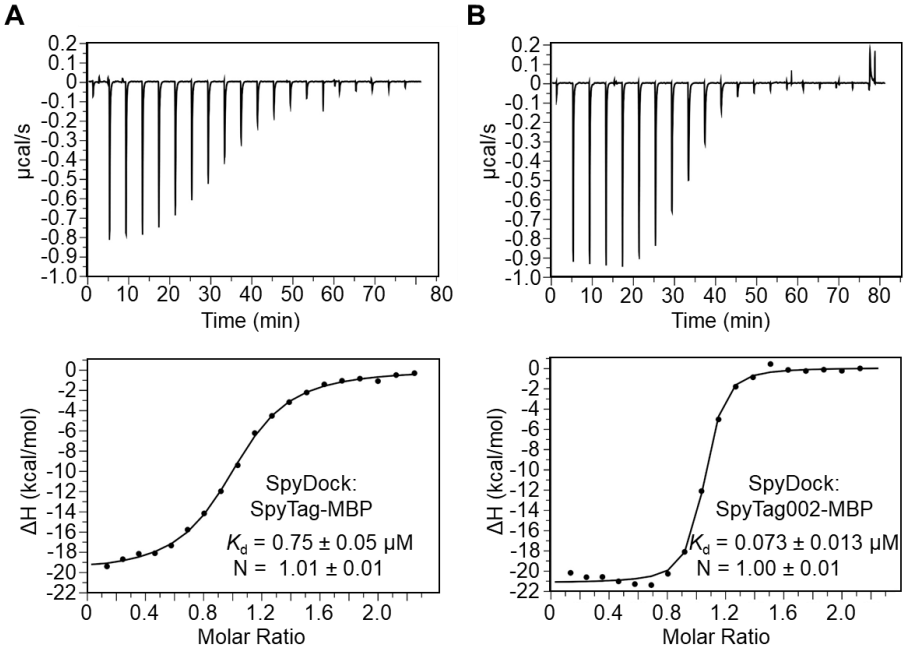


Figure 3.3 Binding affinity of SpyDock to SpyTag and SpyTag002. ITC binding isotherms of SpyDock binding at 25 °C to (A) SpyTag-MBP or (B) SpyTag002-MBP with K_d and stoichiometry (N) shown. Errors bars are the uncertainty in fit to the curve using a 1:1 binding model.

SpyDock and SpyTag-MBP or SpyTag002-MBP bound with 1:1 stoichiometry (Figure 3.3). SpyDock's K_d with SpyTag-MBP (0.75 µM) was approximately 10-fold weaker than with SpyTag002-MBP (0.073 µM), which can be explained by the increased net positively-charged residue interactions found in the second-generation SpyTag002 with the net negatively-charged SpyDock that resulted in a higher rate of reaction (Keeble et al., 2017). The K_d values

determined here are comparable to the values found in current peptide affinity tags mentioned in section 1.2.1.

3.2 Optimising washing and elution conditions

The previous trials to determine the final SpyDock construct were done by binding and eluting purified SpyTag-MBP from a neutral Tris-phosphate buffer (Section 3.1). However, the condition does not emulate real purification where the target protein is found in complex cellular lysate mixtures. In the presence of many cellular molecules, SpyDock may non-specifically bind cellular material other than the target SpyTag-protein. The optimisation to establish a high purity Spy&Go process is described here.

3.2.1 Screening elution conditions

Firstly, I determined the best elution condition to employ in Spy&Go. Protein-protein interactions can be chemically disrupted in a number of ways by breaking the different bonds involved in the protein-protein interface or destabilising protein folding (Firer, 2001). I selected several elution conditions presented in the review and other common elution buffers to determine which buffer can best displace SpyTag from SpyDock. Purified SpyTag-MBP without any cellular material was used for Spy&Go purification to allow for minimal background interference when analysing gel densitometry data (Figure 3.4 A). The panel of buffers were selected from different pH values (glycine pH 2.5, acetate pH 5.0, Tris pH 9.0), high salt concentrations (2 M $MgCl_2$, 2 M imidazole, 2.5 M imidazole), chaotropic activity (2 M sodium iodide) and an organic solvent (50% ethylene glycol). Among the buffers, glycine pH 2.5, 2 M imidazole, and 2.5 M imidazole eluted the most SpyTag-MBP (Figure 3.4), followed by 2 M NaI. Minimal elution of SpyTag-MBP from SpyDock at pH 5 or 9 may be beneficial to ensure proper binding of SpyTag-proteins across this pH range. Between the three best buffers, 2.5 M imidazole in Tris-phosphate at pH 7.0 was chosen as the elution buffer of choice for Spy&Go. Although low pH buffers are used as elution buffers in many affinity chromatography methods

(Kimple et al., 2013), the low pH may be detrimental to the structure and activity of proteins (Talley and Alexov, 2010). 2 M imidazole was insufficient for the elution of SpyTag003-MBP (Figure 3.4 C). 2.5 M imidazole has been shown to be effective in protein elution and permissive in retaining protein activity (Buldun et al., 2018). Nonetheless, for applications that should avoid high imidazole concentrations, a low pH buffer can be used instead.

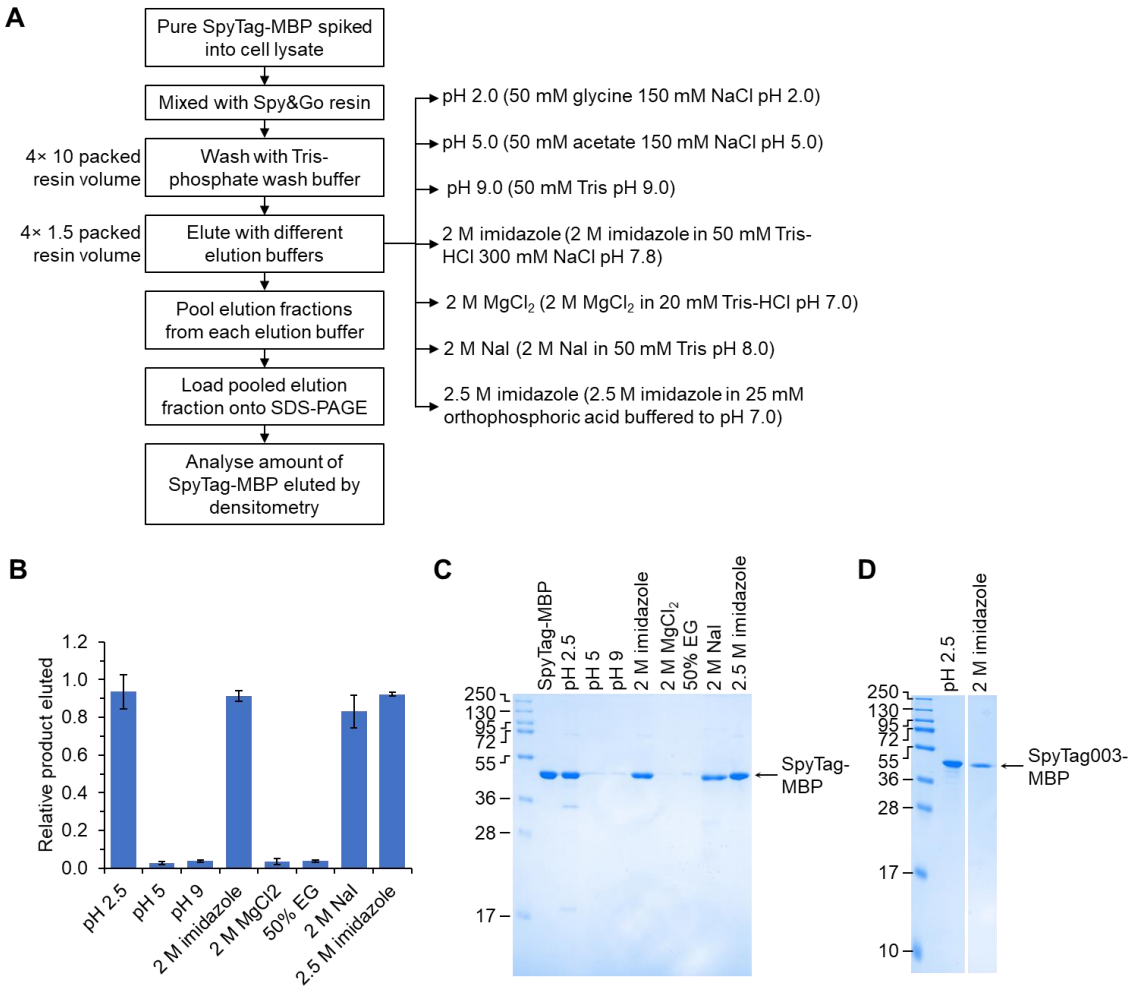


Figure 3.4 Elution buffer screening for Spy&Go. (A) Flow-chart of the steps to determine the Spy&Go elution buffer. (B) Relative quantification (mean ± 1 s.d, *n* = 3) of eluted SpyTag-MBP from SpyDock using different elution buffers through gel densitometry. Purified SpyTag-MBP was mixed with immobilised SpyDock, washed with 500 mM imidazole in Tris-phosphate pH 7.0 and eluted with different elution buffers. pH 2.5, 2 M imidazole and 2.5 M imidazole were identified as potential elution buffer. Representative Coomassie-stained SDS-PAGE of total pooled elution fractions are shown in (C). EG: ethylene glycol. SpyTag-MBP lane as protein size control. Recovery of 1.0 in relative product eluted corresponds to the reference sample of input SpyTag-MBP. (D) Representative Coomassie-stained SDS-PAGE of SpyTag003-MBP eluted from SpyDock using pH 2.5 or 2 M imidazole.

3.2.2 Optimising washing conditions

Following the determination of the elution buffer of choice, the washing conditions were then optimised. For this, purified SpyTag-MBP was spiked into a clarified *E. coli* cell lysate before being incubated with Spy&Go resin to mimic the cellular conditions found when purifying a recombinantly-expressed protein. The resin was then subjected to washes of different buffers, to determine the optimal buffer to remove bound non-specific cellular material. Initially, the physiological Tris-phosphate pH 7.0 binding buffer was tested as this mild buffer condition would be most suitable for protein stability and folding during the washes. Nonetheless, the purity of the elution from purifying spiked SpyTag-MBP into clarified cell lysate was not satisfactory: background protein contamination was still visible in the final pooled elution fractions (Figure 3.5 A).

Subsequently, a regimen of alternating cycles of high and low salt buffers was employed to remove non-specific background protein binding as suggested by Vassylyeva *et al.*, (2017). Here, the resin-incubated cell lysate with spiked SpyTag-MBP was washed with 50 mM Tris + 1 M NaCl at pH 7.8 and then 50 mM Tris without NaCl at pH 7.8 consecutively for two cycles. However, this methodology did not yield good purity elution (Figure 3.5 B). Finally, I emulated the washing conditions from Buldun *et al.*, (2018) by incorporating 500 mM imidazole in the Tris-phosphate pH 7.0 buffer. This resulted in a clean output of eluted SpyTag-MBP from the clarified cell lysate (Figure 3.5 C). Thus, the conditions for the Spy&Go purification regimen was now set: binding in Tris-phosphate at pH 7.0, washing with 500 mM imidazole in Tris-phosphate at pH 7.0, and eluting with 2.5 M imidazole in Tris-phosphate at pH 7.0.

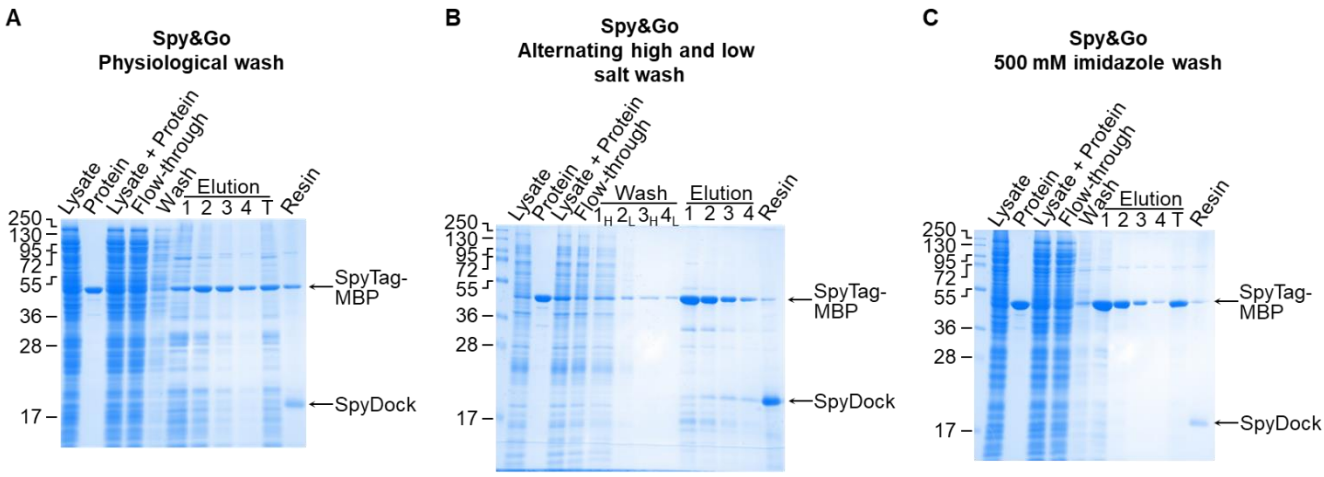


Figure 3.5 Imidazole wash improves Spy&Go purity. SpyTag-MBP was spiked into *E. coli* cell lysate and mixed with immobilised SpyDock, washed with different wash buffers, and eluted with 2.5 M imidazole in Tris-phosphate pH 7.0. Representative Coomassie-stained SDS-PAGE of the purifications are shown. Washes with physiological buffer (A) or alternating high (H: 1 M NaCl) and low (L: no NaCl) salt buffers (B) did not yield good elution purity. (C) Washing with 500 mM imidazole removed non-specific background binding from cell lysate yielding high purity elution. Protein: input SpyTag-MBP protein, T: total pooled elutions.

3.2.3 Purification of SpyTag-proteins

Following the characterisations of SpyDock and optimisation of wash and elution conditions, different proteins were purified to demonstrate the usability of the system across different protein features and cellular expression types. Initially, purified SpyTag-MBP with an N-terminal polyhistidine tag was spiked into clarified *E. coli* cell lysate and purified via Spy&Go or Ni-NTA to compare the two techniques. For Spy&Go, the resin was washed four times with 500 mM imidazole in Tris-phosphate at pH 7.0, and eluted four times with 2.5 M imidazole in Tris-phosphate at pH 7.0, whereas for Ni-NTA, the resin was washed eight times with 30 mM imidazole in 50 mM Tris-HCl and 300 mM NaCl at pH 7.8 and eluted four times with 200 mM imidazole in 50 mM Tris-HCl and 300 mM NaCl at pH 7.8. From Figure 3.6 A & B, the Spy&Go purification system yielded higher purity elutions compared to Ni-NTA, at 98.9 ± 0.5% and 66.4 ± 1.9% respectively. Additionally, spiked SpyTag002-MBP and SpyTag003-MBP were also successfully purified from the clarified *E. coli* lysate with similar purities, albeit with lower yield, since more protein was still found bound on the resin. This finding can be attributed to

the newer SpyTag generations having higher binding affinities to SpyDock than SpyTag-MBP (Figure 3.6). When making the Spy&Go resin, different batches may have different amounts of immobilised SpyDock, thus creating variabilities in protein binding capacity. Indeed, Ni-NTA resin has a higher binding capacity than Spy&Go. This inconsistency was minimised by inputting a low concentration of SpyTag-MBP (0.09 mg, approximately 10-fold lower than the amount of resin-bound SpyDock), thus undersaturating each resin to promote full protein mass transfer from liquid phase to the resin and enabling fairer comparison in eluting the same amount of protein from the resin upon addition of elution buffer.

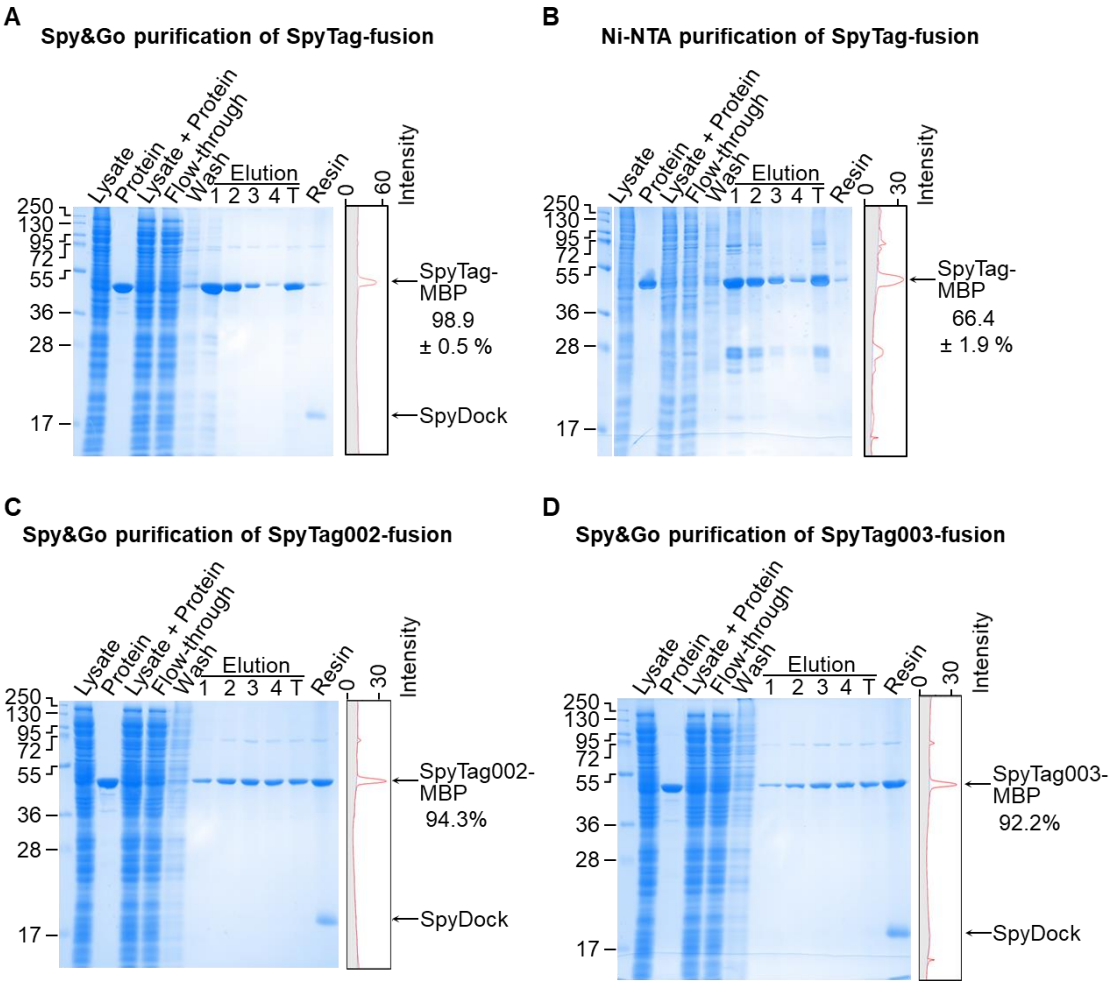


Figure 3.6 Spy&Go purification of different SpyTag generations. SpyTag-MBP was spiked into *E. coli* cell lysate and mixed with immobilised SpyDock, washed with 500 mM imidazole in Tris-phosphate pH 7.0, and eluted with 2.5 M imidazole in Tris-phosphate pH 7.0. Representative Coomassie-stained SDS-PAGE shown here. (A) SpyTag-MBP has higher purity when purified by Spy&Go compared to (B) Ni-NTA (mean ± 1 s.d., *n* = 3). SpyTag002-MBP (C) and SpyTag003-MBP (D) purification from *E. coli* lysate by Spy&Go. Protein: input SpyTag-MBP, SpyTag002-MBP or SpyTag003-MBP protein, T: total pooled elutions. Purity of T determined by gel densitometry.

Other proteins were purified from *E. coli* to illustrate the versatility of SpyTag positions. SpyTag was cloned at the C-terminus of αDR5, an agonistic nanobody against Death Receptor 5, and internally at the flexible loop region of a nuclease-inactivated tandem PvuII restriction enzyme (scPvuII). Purification of these proteins were successful and showed good purity, showing that SpyTag can be positioned at different sites to suit different proteins and applications (Figure 3.7 A-C). Additionally, a fluorescent protein, mClover3, was purified from

E. coli with an N-terminal SpyTag and the fluorescence was compared to mClover3 purified through Ni-NTA at the same concentration ranges. The fluorescence showed little difference between the proteins, indicating that the high imidazole concentration from Spy&Go elution buffer did not affect the structure of mClover3 to substantially diminish its fluorescence (Figure 3.7 D).

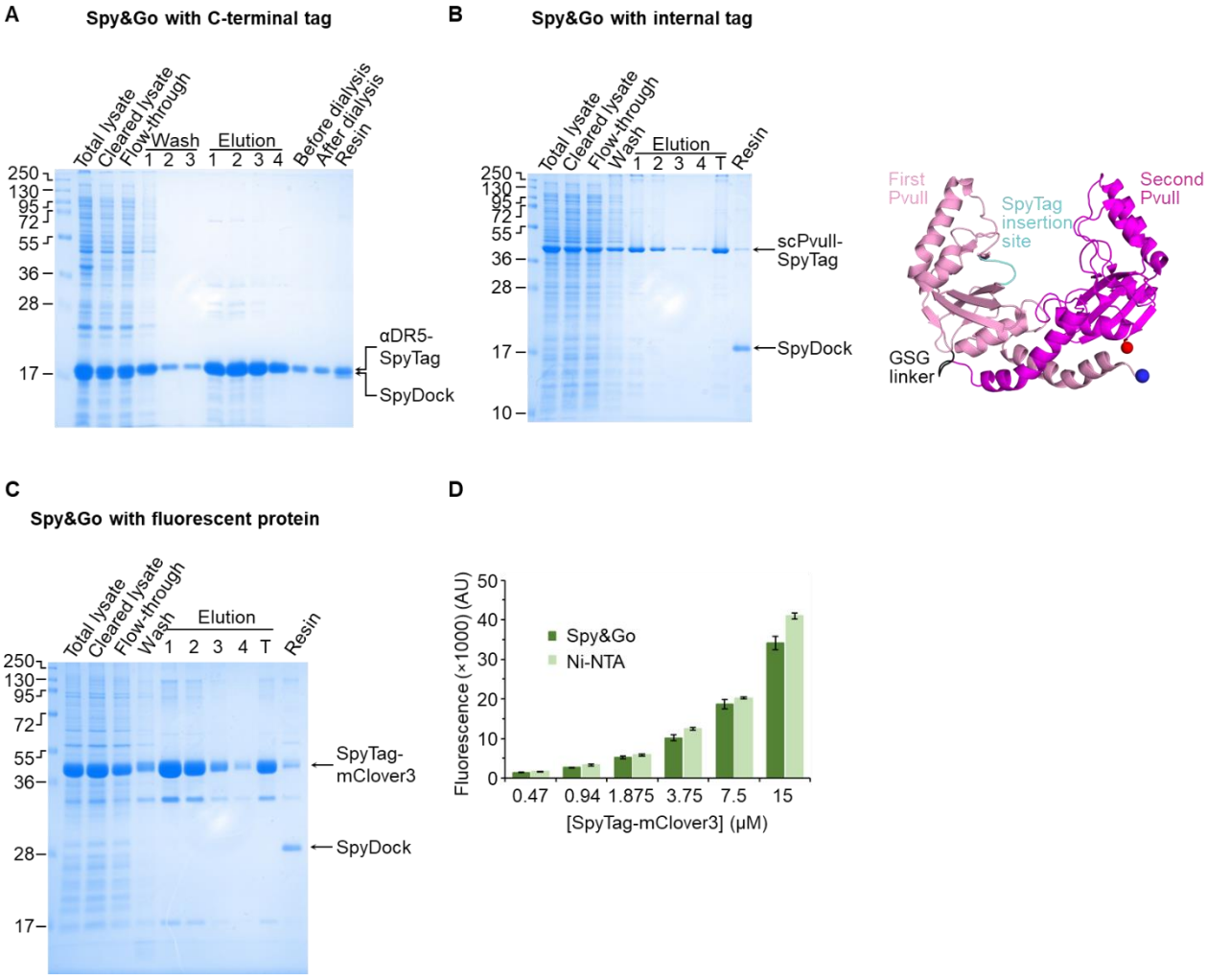


Figure 3.7 Spy&Go purification of different bacterial proteins. Spy&Go purification is feasible with different SpyTag position, either at (A) the C-terminus following αDR5, (B) internal loop of single chain Pvull (scPvull, with internal position illustrated at right panel), or (C) N-terminus of mClover3. Representative Coomassie-stained SDS-PAGE shown here. T: total pooled elutions. (D) The fluorescence of mClover3 purified via Spy&Go was still comparable to Ni-NTA polyhistidine tag-purified mClover3 (mean ± 1 s.d. *n* = 3).

I then explored the feasibility of purifying proteins from mammalian expression systems. Mammalian expression systems enable a stringent protein folding pathway and facilitate post-translational modifications important for complete biological activity, particularly for biopharmaceuticals (Seiradake et al., 2015; Zhu, 2012). The extracellular component of the Epithelial Cell Adhesion Molecule (EpCAM) was tagged with SpyTag and a polyhistidine tag at the C-terminus and expressed in adherent HEK293T cells. The protein was subjected to Spy&Go or Ni-NTA purification and the purity was compared. Different glycoforms of the protein contributed to the heterogeneous size of EpCAM-SpyTag seen on the Coomassie-stained SDS-PAGE. The purity of EpCAM-SpyTag was similar for both purification strategies. However, there was more protein eluted from Ni-NTA, possibly because of the higher protein binding capacity of the Ni-NTA resin. Purification of a malaria antigen, Cysteine-rich Protective Antigen (CyRPA), with a C-terminal SpyTag from suspension Expi293HEK cells was also explored. Again, Spy&Go purification allowed good purity (Figure 3.8).

These purifications show that Spy&Go is a suitable method for the purification of SpyTagged proteins using different expression systems and tag localisations.

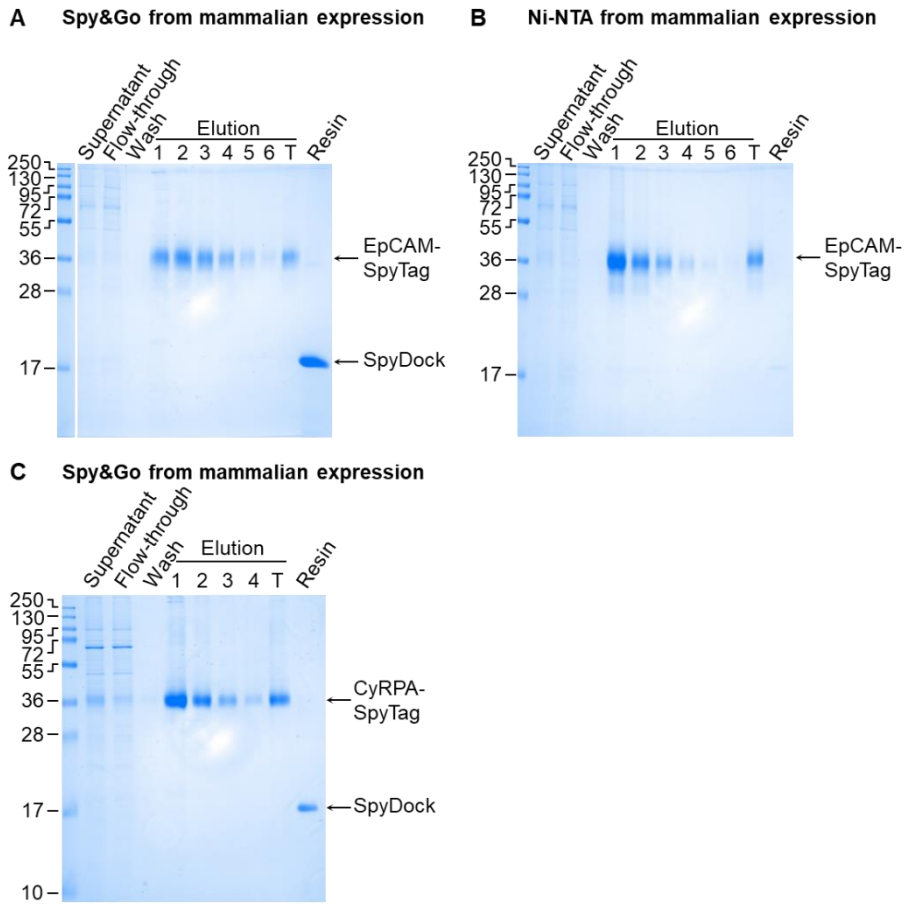


Figure 3.8 Spy&Go purification of different mammalian proteins. Spy&Go purification is effective with different SpyTag-proteins, either in (A) adherent HEK293T cells expressing EpCAM-SpyTag, with Ni-NTA polyhistidine comparison in (B), or in (C) suspension Expi293HEK cells. Representative Coomassie-stained SDS-PAGE shown here. T: total pooled elutions.

3.3 Spy&Go resin characterisation

Carefully characterising the Spy&Go resin is important for adoption of the system among the scientific community. Many researchers require different initial conditions before purification because of varying protein stabilities in different buffers. Additionally, practicality and frugality matter when selecting a resin for protein purification, since the resin needs to tolerate long-term storage and be reusable to be economical. I explored these traits to further characterise the Spy&Go resin.

3.3.1 Binding buffer compatibility

The route for protein stabilisation and purification following recombinant overexpression may differ depending on the physical attributes of the protein. Membrane proteins typically require lipid-solubilising detergents to be extracted from the membrane layer (Anandan and Vrielink, 2016), whilst complex or insoluble proteins may be sequestered in bacterial inclusion bodies that necessitates solubilisation using denaturants prior to purification (Singh and Panda, 2005). Mild non-ionic detergents can also help to decrease non-specific binding (Arkin et al., 2004). To be applicable to a wider variety of proteins, I investigated the buffer compatibility of Spy&Go by spiking purified SpyTag-MBP into different biologically-relevant buffers, binding with Spy&Go resin and attempting to purify the protein following the same wash and elution protocol as above. The buffers include physiological buffers such as PBS and saline Tris-HCl (50 mM Tris-HCl, 300 mM NaCl at pH 7.8), detergents such as Tween 20 (polysorbate 20) and Triton X-100, and denaturants such as urea and GuHCl. The detergents were dissolved in Tris-phosphate buffer at pH 7.0, whereas urea was dissolved in 50 mM Tris-HCl at pH 8.0 and GuHCl was dissolved in distilled water titrated too pH 2.0.

The binding and elution of spiked SpyTag-MBP from these buffers were compared to the Tris-phosphate pH 7.0 buffer described above. The efficiency of binding was compared by loading the final pooled elution fractions onto non-reducing SDS-PAGE; the similar wash and elution should minimise variability between the samples thus the amount of protein eluted should only be affected by binding to the Spy&Go resin. Between the three physiological buffers (Tris-phosphate, PBS and saline Tris-HCl), only saline Tris-HCl impaired the binding of SpyTag-MBP to Spy&Go resin (Figure 3.9). Spy&Go was compatible with up to 2% of Tween 20 and Triton X-100, comparable to other established affinity purification method such as Ni-NTA for polyhistidine tags (Qiagen) and Strep-Tactin for Strep-tag (IBA Lifesciences). For denaturing conditions, the maximum allowable concentration of urea was 2 M, whereas for GuHCl it was 0.5 M without any loss in binding capacity (Figure 3.9). In contrast, other protein-based affinity chromatography methods can go up to 6 M urea (Strep-Tactin XT, IBA

Lifesciences), 8 M urea (CaptureSelect C-tag, Thermo Scientific) or 1 M GuHCl (Strep-Tactin XT, IBA Lifesciences). At low denaturant concentrations of around 2 M, many proteins will not be in unfolded state (Hédoux et al., 2010; Rashid et al., 2005), thus currently Spy&Go is not suitable for denaturing purification.

3.3.2 Long-term storage

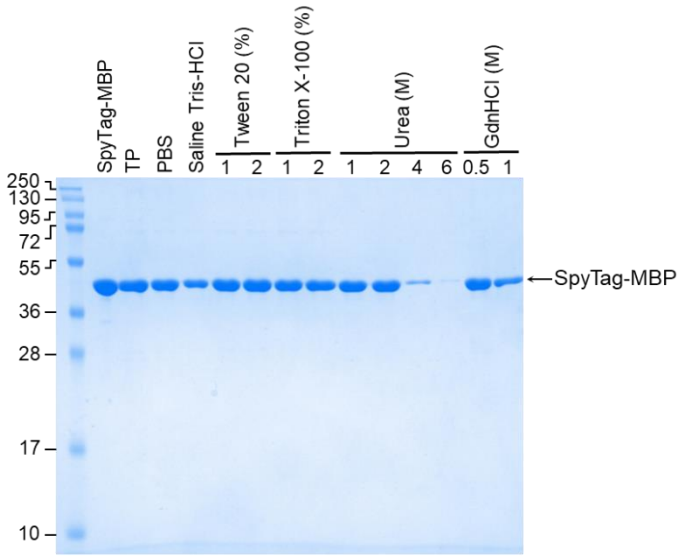


Figure 3.9 Spy&Go binding buffer compatibility. Purified SpyTag-MBP was incubated with Spy&Go resin under the different binding buffers, washed with 500 mM imidazole in Tris-phosphate pH 7.0, and eluted with 2.5 M imidazole in Tris-phosphate pH 7.0. Total pooled elution fractions were loaded onto non-reducing SDS-PAGE and stained with Coomassie.

A key practical feature for affinity chromatography methods is the capability to store the resin long-term in storage buffer without loss in quality. Most resins employ 20% ethanol as the storage buffer to prevent bacterial growth that can proteolyse proteins (Chatterjee *et al.*, 2006; GE Healthcare, 2017). The concentration is also sufficiently low to not precipitate proteins, since higher ethanol concentrations may destabilise protein structure (Schilcher et al., 2013; Yoshikawa et al., 2012). Ethanol is also inexpensive, non-toxic and easily disposed of, unlike preservatives using the more toxic sodium azide. Nonetheless, it must be empirically determined if Spy&Go can tolerate long-term storage in 20% ethanol.

Spy&Go resin was stored in 20% ethanol and its function to purify SpyTag-MBP from lysate was determined and compared with before storage. The resin was still functional after 10 months of storage in 20% ethanol at 4 °C (Figure 3.10).

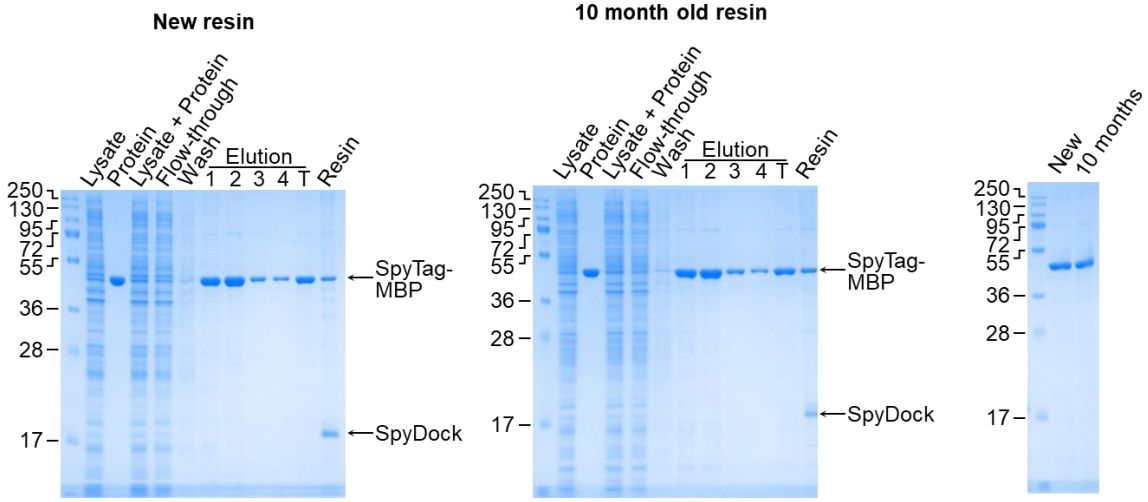


Figure 3.10 Spy&Go resin is stable for long-term storage. Spy&Go resin was stored in 20% ethanol for 10 months and the ability to purify SpyTag-MBP was compared with new Spy&Go resin. SpyTag-MBP was spiked into *E. coli* cell lysate and mixed with immobilised SpyDock, washed with 500 mM imidazole in Tris-phosphate pH 7.0, and eluted with 2.5 M imidazole in Tris-phosphate pH 7.0. Total pooled elution fractions were loaded onto non-reducing SDS-PAGE and stained with Coomassie. No difference in purification quality and capacity was observed. T: total pooled elutions.

3.3.3 Regeneration and reusability

The ability to reuse affinity chromatography resins is another practical feature that minimises material waste and lowers costs compared to single use resins. Post-purification, some tagged proteins may still be bound to the affinity resin. Resins can be reused by removing any residually-bound proteins from the immobilised ligand. For protein-based affinity resins, regeneration is commonly done by destabilising the bonds within and between proteins by subjecting the resin to low pH or denaturing conditions (GE Healthcare, 2017; Johansson *et al.*, 2006; Vassylyeva *et al.*, 2017). Alkaline washes using 0.01-0.5 M NaOH can also be employed to remove protein aggregates, nucleic acids and viral contaminants for column and resin sanitisation (Grönberg et al., 2011).

Although the binding affinity between SpyDock and SpyTag-MBP is only moderate (Figure 3.3), initial trials found it difficult to remove residual bound SpyTag-MBP after elution. Isolated attempts at regeneration of Spy&Go resin using either glycine at pH 2, 4 M imidazole at pH 7.0, 6 M GuHCl at pH 6.5, or 0.1 M NaOH did not fully remove the bound SpyTag-MBP. Subsequently, subjecting the resin with consecutive washes of 4 M imidazole in Tris-phosphate at pH 7.0, 6 M GuHCl at pH 2.0 and 0.1 M NaOH with a physiological wash in between each harsh condition managed to remove the remaining SpyTag-MBP from the used Spy&Go resin to a non-detectable amount on Coomassie-stained SDS-PAGE (Figure 3.11 A).

The regenerated resin was then compared with a new unused resin to purify another protein, αDR5-SpyTag, which has a different molecular weight than SpyTag-MBP (16 kDa versus 45 kDa). The purity and amount of αDR5-SpyTag purified between the two resins were comparable and the regenerated resin did not have apparent carryover of the bigger SpyTag-MBP from a previous purification (Figure 3.11 B). From the regeneration regimen, 4 M imidazole and 6 M GuHCl at pH 2.0 served to remove any remaining bound SpyTag-proteins and other contaminating proteins by interaction destabilisation and unfolding of SpyDock. Although this treatment may not affect the capacity of the resin, prolonged exposure of SpyDock to 0.1 M NaOH will invite its hydrolysis and effectively lower the resin capacity after multiple regeneration, as already seen with some residual SpyDock stripped from the resin in Figure 3.11 A.

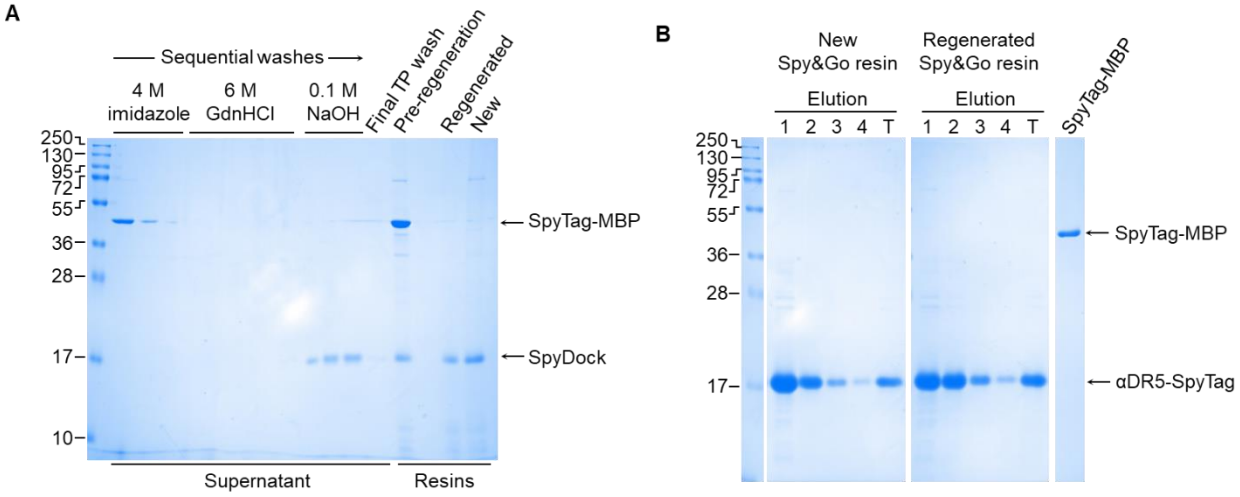


Figure 3.11 Spy&Go resin can be regenerated and reused. (A) Regeneration steps to remove residual SpyTag-MBP from a previous purification. "Pre-regeneration" is the resin after the initial SpyTag-MBP purification. "Regenerated" is the resin following sequential imidazole, guanidinium hydrochloride and NaOH washes, in comparison to a "New" unused resin batch. (B) Spy&Go purification of αDR5-SpyTag using New unused resin or Regenerated resin from (A). Elution fractions 1 – 4 and the total pooled elution (T) fraction are shown.

3.4 Discussion

Affinity chromatography remains the most specific method for recombinant protein purification (Kimple et al., 2013). The ease of endowing a target protein with an affinity tag to enable purification without any prior knowledge of the protein's physicochemical characteristics makes affinity chromatography the method of choice when purifying proteins for many different purposes (Gräslund et al., 2008; Vassylyeva et al., 2017). Nevertheless, the presence of the tag post-purification is an unwanted surplus for certain applications, including for cellular and *in vivo* utilisation of the SpyTag/SpyCatcher technology, particularly in vaccination (Brune et al., 2016; Bruun et al., 2018; Jin et al., 2017). Thus, there was a need to utilise SpyTag itself as the affinity tag for the purification of SpyTag-proteins.

Although a non-covalent binding partner for SpyTag could perhaps be engineered through directed evolution with another binding protein scaffold (Hosse, 2006; Simeon and Chen, 2018), the design or selection may take a significant amount of time to set up. Thus, I decided to use SpyCatcher as the foundation protein to engineer a non-covalent binder of SpyTag, since it is already an innate reaction partner with SpyTag and has been shown to retain a non-covalent interaction when a key isopeptide-forming residue was abolished (Bedbrook et al., 2015; van den Berg van Saparoea et al., 2018; Hagan et al., 2010; Keeble et al., 2017; Schoene et al., 2014; Zakeri et al., 2012). With just as few as 5 mutations on top of SpyCatcher002, including the solid-phase immobilising S49C substitution, SpyDock was developed that can efficiently isolate SpyTag-proteins from a cell lysate mixture with a favourable dissociation constant between 10^{-7}-10^{-8} M, comparable to Strep-Tactin binding to

Strep-tag II (Schmidt et al., 2013) or Ni-NTA binding to the hexahistidine tag (Ackerson et al., 2010). The K_d value of SpyTag003-MBP with SpyDock was not quantified, but is expected to be in the range of 10^{-8}-10^{-9} M based on the lower K_d of SpyTag003 D117A-MBP with SpyCatcher003 (Keeble et al., 2019). it will be helpful to measure this affinity directly in future study. Nonetheless, the K_d values for SpyDock to SpyTag or SpyTag002 are still 1-2 magnitudes higher than epitope tags binding to their antibody/antibody mimetic (Ranawakage et al., 2019; Zhao et al., 2016a). The development of SpyDock was dependent on the mutations found on SpyCatcher2.1, without which the amount of SpyTag-MBP bound and eluted diminished by 54% compared to the latest SpyCatcher002. The three mutations serve to complement the negatively-charged SpyCatcher2.1 to the positively-charged SpyTags, which contributes to the increased interactions between the pair.

Different elution conditions were tested to induce the release of SpyTag-MBP from SpyDock. Although the elution was effective at pH 2.5, as is commonly employed by epitope tag-based affinity chromatography, I decided to proceed with a neutral pH elution containing 2.5 M imidazole since exposure to very low pH can have serious irreversible effects to the structure and hence function of a protein. Although SpyTag-proteins and SpyDock have a moderate affinity interaction, high concentrations of imidazole or NaI were needed to perturb the associated complex. Stop-flow FRET has shown the complexity of SpyTag003 D117A/SpyCatcher003 interaction, showing biphasic association and dissociation kinetics (Keeble et al., 2019), so it would be useful in future work to perform such studies on the peptide interactions of SpyDock. Current methods to elute epitope tags from their binders include using a high salt concentration (e.g. >2 M $MgCl_2$ or >1 M Arg-HCl), whereby the dissociation is induced by salting-in the proteins to reduce protein-protein interaction and favouring protein-salt interaction (Tsumoto et al., 2007). However, the high imidazole concentration is instead theorised to induce a global destabilisation of the protein fold to elute SpyTag-MBP from SpyDock.

The purity of proteins eluted from Spy&Go rivalled the most widely used affinity chromatography tag, the polyhistidine tag. Initial attempts at removing non-specifically bound contaminants from SpyDock using either low salt physiological buffer or high salt buffer were unsuccessful, thus a harsher condition was required by incorporating 500 mM imidazole in the wash buffer. This harsher treatment is especially successful since under-saturating the resin by spiking a small amount of SpyTag-MBP into *E. coli* lysate invites more non-specific binding of contaminants to the immobilised ligand and matrix. However, the purity from Spy&Go purification was still 1.5-times higher than through Ni-NTA. A harsher wash may compromise the overall yield of the purification since some SpyTag-proteins were seen to have come off during the washes. Nonetheless, the ability to purify proteins from different expression systems and different permissible tag locations enable a high flexibility in experimental design.

The binding condition for Spy&Go should be suitable for a majority of soluble proteins, however the binding was compromised in saline Tris-HCl buffer which may be explained by salt sensitivity from charge shielding that can weaken the interaction between SpyDock and SpyTag (Tsumoto et al., 2007). Purifying membrane proteins that require detergents may also be possible because the binding was not affected by the presence of up to 2% Tween 20 or Triton X-100, but such purification will require further validation. Adding 10 mM β-mercaptoethanol during the purification of αDR5-SpyTag also did not impair the process, as expected since there are no cysteines in SpyDock apart from the immobilising mutation. However, purification in denaturing conditions is still a limitation as the binding was severely affected in the commonly used chaotropic denaturants urea and GuHCl. Additionally, the ability to store and regenerate the Spy&Go resin can contribute to a longer resin shelf life, although it is yet to be determined the extent to which the resin can be reused without significantly impairing the capacity and quality of purification.

The development of SpyDock to make the Spy&Go purification platform by having minimal mutations to the sequence of SpyCatcher002 was successfully created in this chapter. However, the high imidazole concentration during elution may not be appropriate for

all applications, such as purifying a salt-sensitive protein or a multi-component complex relying on non-covalent interactions which can be broken by the Spy&Go elution buffer. Additionally, imidazole may compete with metal ion coordination with metal-binding proteins. Nonetheless, the initial results I have obtained here show favourable comparison to the other established purification platforms which have been optimised over many years. Spy&Go also demonstrates the versatility of the SpyTag/SpyCatcher system beyond the original covalent bioconjugation purpose.

4 Mild Spy&Go elution and design of a pH-switchable SpyTag binder

The Spy&Go purification system described in Chapter 3 was successfully established to enable high purity purification of proteins fused to SpyTag. Nonetheless, the current Spy&Go elution condition of 2.5 M imidazole may not be suitable for all proteins.

Using a high imidazole concentration (>1 M) for protein elution has previously been explored: 2 M imidazole to elute a multimeric protein assembly (Brune et al., 2016), 2-3.5 M imidazole to remove ligated peptides from an engineered ligase (Buldun et al., 2018) and up to 5 M imidazole to elute IgG1 from Protein A (Seed, 1994). They report no substantial loss in target protein function when exposed to the high imidazole concentrations, thus demonstrating the mild nature of imidazole compared to low pH for a non-specific elution buffer. This study also showed retained protein activity after Spy&Go purification using 2.5 M imidazole for elution (section 3.2.3 and Figure 3.7).

However, a high salt or solute concentration may affect protein structure and folding to induce a salting out effect, leading to precipitation (Abriata et al., 2016; Arakawa and Timasheff, 1991; Damodaran and Kinsella, 1981). Additionally, due to its structure as the functional group of histidine, imidazole can participate in many histidine-like interactions with other amino acids that may lead to protein destabilisation, especially at such high concentrations (Liao et al., 2013). Although imidazole can be removed by dialysis or buffer exchange, some proteins such as metal-binding proteins may be sufficiently sensitive to be irreversibly destabilised even after a short exposure to imidazole (Huang et al., 2010).

In this chapter, I describe two methods of mild elution for the current Spy&Go system using presently established approaches in the literature. Additionally, I explored the rational engineering of a pH-switchable SpyDock, termed SpySwitch, to bind SpyTag003 at pH 8 and release at a lower pH by incorporating histidines at specific residues.

4.1 Mild Spy&Go elution conditions

Currently, the high imidazole concentration in Spy&Go elution buffer is theorised to result in global destabilisation of some proteins that not only impairs the SpyTag and SpyDock interaction, but also protein folding which can adversely affect the target protein. Other methods are available and routinely utilised to allow a more defined route of protein-protein interaction destabilisation, namely by introducing a competitive binding partner or introducing a peptide cleavage motif for endoprotease processing between the affinity tag and target protein. Here, I explored these two methods to generate milder elution pathways for the Spy&Go purification system.

4.1.1 Competitive peptide elution

Many epitope affinity tags utilise a competitive peptide elution for a milder elution condition instead of the harsher low pH buffer (Kimple et al., 2013). In this method, a peptide with the sequence of the affinity tag is dissolved in physiological buffer and incubated with the target protein and resin mixture to promote displacement of the bound protein. Since SpyTag003 is the latest generation of the SpyTag/SpyCatcher technology with the highest binding affinity to all SpyCatcher generations (Keeble et al., 2019), SpyTag003 was chosen as the elution peptide of choice for Spy&Go, to competitively bind and displace target proteins tagged with either SpyTag, SpyTag002, or SpyTag003. MBP fused with the three SpyTag generations were spiked into clarified *E. coli* lysate, washed with 500 mM imidazole in Tris-phosphate pH 7.0 to remove non-specifically bound cellular material, and eluted by incubating with 1 mM SpyTag003 peptide dissolved in Tris-phosphate at pH 7.0 at 30 °C for 2 h. The elution step was then repeated a second time.

Peptide elution using SpyTag003 as the competitive peptide managed to elute SpyTag-, SpyTag002-, and SpyTag003-MBP from the Spy&Go resin, albeit the elution was still incomplete after 4 h of incubation at an elevated temperature as seen from the remaining target protein bound in the boiled resin fraction (Figure 4.1). Lower peptide concentrations and

a temperature of 4 °C were previously tested, however the elution was poor. As expected, more protein was displaced when MBP fused with the first generation SpyTag compared to the latest generation due to the stronger binding interactions between SpyTag003 and SpyDock. Excess SpyTag003 peptide was present in the eluted fractions, thus buffer exchange, dialysis, or size exclusion chromatography is still required to remove the surplus peptide. With the establishment of the competitive peptide elution, a milder elution condition for Spy&Go is available. However, users will need to consider the higher expense to synthetically produce the SpyTag003 peptide and the elevated temperature at which the elution is needed.

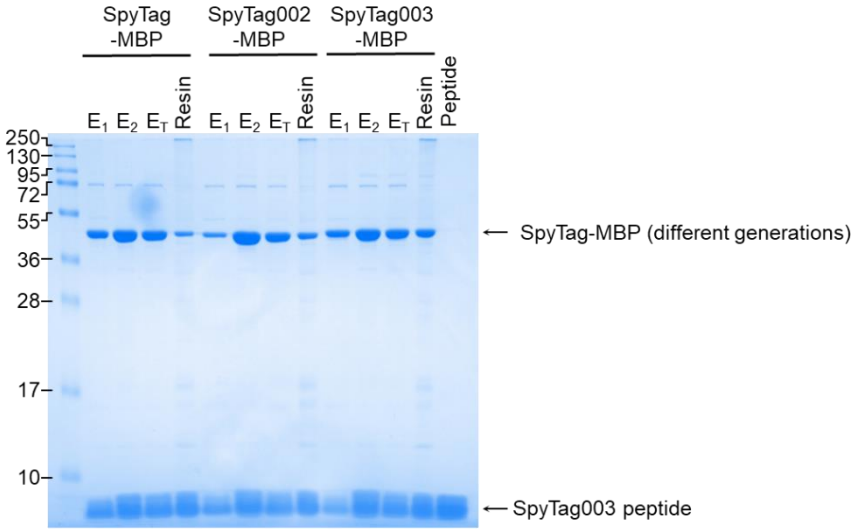


Figure 4.1 Competitive peptide elution of different SpyTag-MBP generations from Spy&Go. 1 mM of SpyTag003 peptide in Tris-phosphate buffer pH 7.0 was incubated with Spy&Go resin bound with SpyTag-, SpyTag002-, or SpyTag003-MBP spiked into *E. coli* lysate and washed. The incubation was done for 2× 2 h at 30 °C and fractions loaded on SDS-PAGE and Coomassie-stained. SpyTag003 peptide manage to elute all generations of SpyTag fused to MBP. E1: first elution, E2: second elution, ET: total pooled elution. Boiled resin fraction also shown.

4.1.2 On-resin tag cleavage

Another common method to elute affinity tagged proteins from resin is through on-resin tag cleavage (Waugh, 2011). This method is accomplished by genetically-encoding a cleavage site recognisable by proteases specific for the site situated between the affinity tag

and the target protein. Cleavage of the site removes the target protein from the affinity tag for elution, leaving the affinity tag still bound on the resin. Although useful for applications requiring only the target protein, this methodology is not advantageous for Spy&Go as the functional SpyTag is excised, leaving the target protein unable to participate in SpyCatcher-based bioconjugation reactions. Nonetheless, SpyTag removal through on-resin cleavage is still useful since the purity of proteins obtained from Spy&Go is favourable.

To demonstrate the feasibility of on-resin cleavage of SpyTag-protein bound to SpyDock, a TEV recognition site consisting of ENLYFQ|G, where | denotes the cleavage site, was inserted downstream of SpyTag and preceding MBP. The protein was expressed in *E. coli*, sonicated to lyse the cells and incubated with Spy&Go resin to allow binding to SpyDock. Non-specific cellular material was then washed with 500 mM imidazole in Tris-phosphate at pH 7.0. 0.5 mg/mL of polyhistidine-tagged SuperTEV fused to MBP (Correnti et al., 2018) was added to the resin to incubate at different time points at 4 °C. Elution of tag-cleaved MBP was already apparent after 2 h, but the reaction was optimal after overnight incubation (Figure 4.2). Steric hindrance from the binding of SpyTag to SpyDock may pose a problem for the accessibility of SuperTEV to the TEV cleavage site, however the construct was already used during phage display selection for the SpyTag002/SpyCatcher002 pair (Keeble et al., 2017) and further accessibility may be provided by elongating the linker preceding the cleavage site.

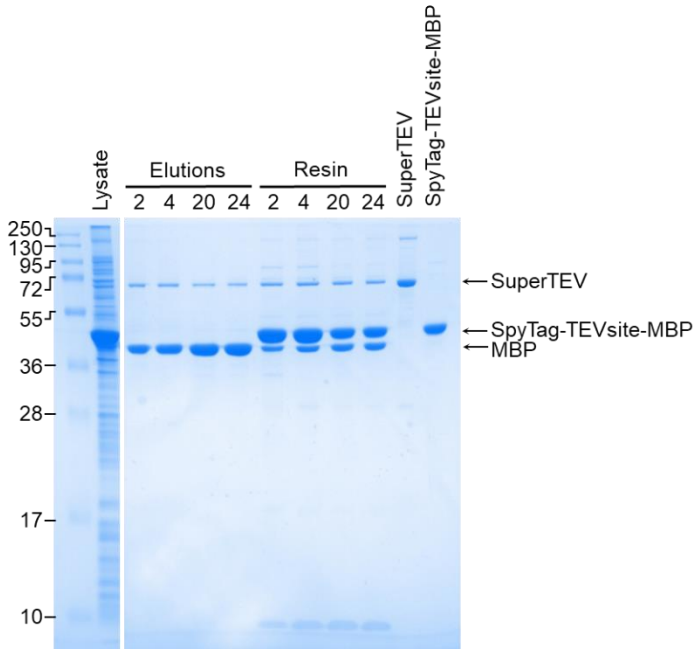


Figure 4.2 On-resin tag cleavage of SpyTag-TEVsite-MBP using SuperTEV protease. Recombinant SpyTag-TEVsite-MBP was incubated with Spy&Go resin and washed. MBP was eluted by incubating with 0.5 mg/mL of SuperTEV protease for 2, 4, 20, or 24 h at 4 °C. The elution and resin fractions were loaded on SDS-PAGE and Coomassie-stained. Controls of SuperTEV and SpyTag-TEVsite-MBP were shown for size comparison. Overnight TEVsite cleavage with SuperTEV for MBP elution is possible.

4.2 Engineering of a pH-switchable SpySwitch

Section 4.1 demonstrates that the elution of SpyTag-proteins from immobilised SpyDock under mild conditions is possible. However, both methods have disadvantages. Competitive peptide elution can be expensive when scaling up due to the high concentration of peptide needed, whereas on-resin tag cleavage removes the functional SpyTag for downstream reactions. A more applicable and cheaper alternative is needed that does not compromise the binding between SpyTag and the non-reactive SpyCatcher variant but is capable of elution under mild condition.

Some proteins in nature undergo a conformational change when exposed to a change in pH of their surrounding environment, which include lysosome-targeted proteins (Kalani et al., 2013; Leloup et al., 2017; Zheng et al., 2014) and spider silk (Dicko et al., 2004). The mechanism of the conformational change depends on the presence and position of ionisable

amino acid residues, particularly histidine, aspartic acid and glutamic acid (Askarieh et al., 2010; Kronqvist et al., 2014; Valéry et al., 2015). Cancer cells have also capitalised on this trait by evolving arginine-to-histidine mutations to confer a gain in pH-sensing function to stabilise the active form of kinases for proliferation (White et al., 2017). Over the years, this phenomenon has been introduced into many engineered proteins to artificially switch binding affinities between the engineered proteins and their ligand from high affinity at high pH to low affinity at low pH. Notable examples include disrupting the binding between an antibody and its native antigen (Murtaugh et al., 2011; Schröter et al., 2015) and disassembling a homotrimeric coiled coil into its respective monomers (Boyken et al., 2019).

The engineered pH-sensitive proteins contain several key features to allow binding affinity switching upon pH change. The amino acid substitutions were found at the protein-protein binding interface, with the key residue substituted into the binding interface being histidine. Histidine was incorporated instead of other amino acids because the pK_a of the imidazole group of histidine's side chain is near neutral at 6.6 (Grimsley et al., 2009; Liao et al., 2013). Additionally, multiple histidines were needed at the binding interface to confer a cooperative effect of decreased binding affinity between the engineered proteins and their native ligand upon a decrease of 1.5-3 pH units (Boyken et al., 2019; Murtaugh et al., 2011).

Engineering a pH-switchable characteristic into a non-reactive SpyCatcher may answer the problem of an efficient and cheap alternative for the elution of SpyTag-proteins by simply changing the pH of the elution buffer. In this section, I describe the development of SpySwitch – a non-reactive SpyCatcher able to bind SpyTag003 at high pH and release at low pH for a more applicable method of purifying recombinant SpyTag003-proteins.

4.2.1 Histidine scanning on SpyTag:SpyCatcher interface

Introducing multiple histidines at the interface between SpyTag and SpyCatcher may compromise the initial binding between the two peptide-protein components. Thus, I selected the latest generation SpyCatcher003 as the starting point for histidine introduction, because

the higher reaction rate contributed by the increased charge complementarity with SpyTag003 may rescue the loss in binding affinity when histidines are mutated in. To develop a pH-switchable non-reactive SpyCatcher003, I rationally introduced single histidines along the binding interface with SpyTag003. I chose SpyTag003 as the binding partner for the pH-switchable non-reactive SpyCatcher003 because it is the latest SpyTag generation which may see an increase in usage due to its higher reaction rate. As the crystal structure for SpyTag003:SpyCatcher003 is not available, I manually inspected the crystal structure of SpyTag:SpyCatcher (PDB ID 4MLI) to select for residues to be mutated (Li et al., 2014).

To focus on the residues involved in the binding interface with SpyTag, I ran the open source InterfaceResidues script on PDB ID 4MLI using PyMOL. Two main stretches of SpyCatcher amino acid residues were identified to be in contact with SpyTag, namely residues 25-35 and 83-92 (Figure 4.3 A). However, SpyCatcher003 residues 104 to 113 are not solved in the 4MLI structure, thus I extended the region of contact to 22-39 and 80-110 to cover more residues (Figure 4.3 B). I manually inspected and selected residues within these contact regions that would create a repulsive destabilising effect against SpyTag003 and local SpyCatcher003 regions upon protonation of histidine, particularly repelling against basic residues (Figure 4.3). Ultimately, 16 residues were chosen to be mutated to histidine from the SpyCatcher003 sequence (Table 4.1). These mutations were singly introduced into SpyCatcher003 S49C E77A to allow consistent solid-phase immobilisation and abolished isopeptide bond formation.

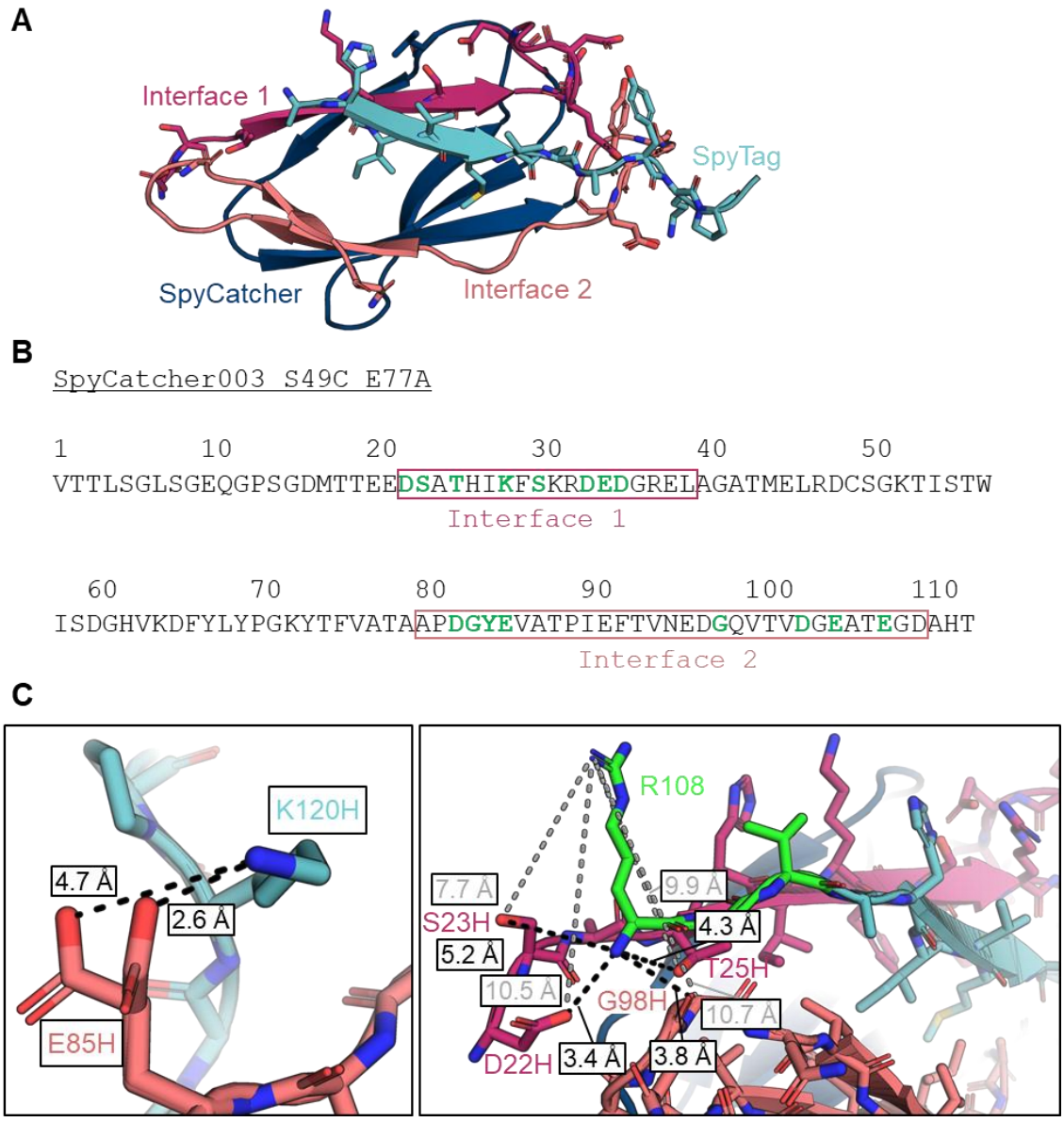


Figure 4.3 SpyTag:SpyCatcher binding interface and residue selection for histidine substitution. (A) Cartoon representation (PDB ID: 4MLI) of SpyTag (cyan) binding to SpyCatcher (blue), with binding interface 1 (magenta) and interface 2 (salmon). SpyCatcher residues for histidine substitution and SpyTag residues shown as sticks. (B) Sequence of SpyCatcher003 S49C E77A with interface 1 and 2 boxed. Residues for histidine substitution in bolded green. (C) Examples of residues selected for histidine substitution. *(Left panel)* E85H mutation may repel against SpyTag003 K120 upon protonation. E85 is available in two conformations, with the A state at 46% and B state at 54% occupancies. A and B states E85 are 4.7 Å and 2.6 Å respectively from K120 ε-amino group. *(Right panel)* SpyTag003 is extendedly N-terminally with residues RGV (green; computer-generated) preceding A111P in the 4MLI structure. SpyCatcher003 D22, S23, T25, and G98 are within 5 Å (black) from R108's N-terminus, thus they were selected for histidine substitutions which, upon protonation, may repel R108. R108 side chain is >5 Å (grey) from D22, S23, T25, and G98, but may still lead to electrostatic repulsion upon conformational change of the complex.

Table 4.1 Residues of SpyCatcher003 S49C E77A for histidine substitution. Residues selected through manual inspection of PDB ID 4MLI for histidine substitution. The residues were selected based on interaction with adjacent basic residues to induce repulsive electrostatic interactions upon histidine protonation at low pH. Navy residues are on SpyCatcher003, cyan residues are on SpyTag003.

Residues on SpyCatcher003 mutated to His	Residues that potentially interact with His mutants (in SpyCatcher003, in SpyTag003)
D22	R108
S23	R108
T25	R108
K28	H26, H112
S30	R32, H112
D33	R37, Y84, Y119
E34	Y119
D35	R37, Y119
D82	K120
G83	K120
Y84	Y119
E85	K120
G98	R108
D103	K123
E105	K123
E108	K123

4.2.2 Investigating tolerance of sfGFP to low pH

A high throughput screening design using fluorescence as a metric was selected to test and compare the 16 different histidine mutations on pH sensitivity of SpyCatcher003 S49C E77A. Superfolder GFP (sfGFP) was chosen as the reporter to fuse to SpyTag003 for determining and comparing its elution from the histidine mutants by measuring its fluorescence intensity after incubation at different pH values. Although the fluorescence of sfGFP is impaired upon exposure to low pH (Roberts et al., 2016; Stepanenko et al., 2014), the fluorescence may be recovered through neutralisation since the protein has a fast refolding rate and high fluorescence recovery upon denaturation by urea, GuHCl and guanidine thiocyanate (Pédelacq et al., 2006; Stepanenko et al., 2012). However, sfGFP fluorescence recovery following neutralisation after denaturation at low pH is yet to be definitively characterised.

To confirm this, I subjected SpyTag003-sfGFP to different pH buffers for 1 h at 4 °C, neutralised to approximately pH 7.5, and compared the excitation and emission spectra with control SpyTag003-sfGFP at the same dilution factor. Control SpyTag003-sfGFP was in 50 mM HEPES, 150 mM NaCl at pH 8.0, whilst the different pH SpyTag003-sfGFP samples were in pH 6.0 buffer (50 mM MES, 150 mM NaCl at pH 6.0), pH 5.0 buffer (50 mM sodium acetate, 150 mM NaCl at pH 5.0), pH 4.0 buffer (50 mM sodium acetate, 150 mM NaCl at pH 4.0), and pH 2.0 buffer (50 mM glycine, 150 mM NaCl at pH 2.0). The pH 2.0-6.0 buffers were neutralised to pH 7.5 by 1 M Tris-HCl at pH 8.0.

Overall, low pH values severely affected the peak fluorescence intensities of SpyTag003-sfGFP after exposure for 1 h (Figure 4.4 A). Upon neutralisation and refolding of sfGFP over an interval of 30 min, fluorescence recovered to approximately the same intensity as SpyTag003-sfGFP already in the pre-neutralised pH buffers (Figure 4.4). Previous attempts to recover sfGFP fluorescence after neutralisation indicated that at least 30 min was needed to fully rescue the fluorescence. Additionally, the peak excitation and emission wavelengths for the renatured SpyTag003-sfGFP and its spectral bandwidth remained the same for all the pH values compared to SpyTag003-sfGFP in pre-neutralised buffers (Figure 4.4). These results show that SpyTag003-sfGFP folding and therefore fluorescence can be rescued after being subjected to low pHs and neutralised even after 1 h of exposure. The volume of neutralisation buffer added to SpyTag003-sfGFP was different for the different pHs, thus the fluorescence intensities across the pH range needed to be corrected for dilution.

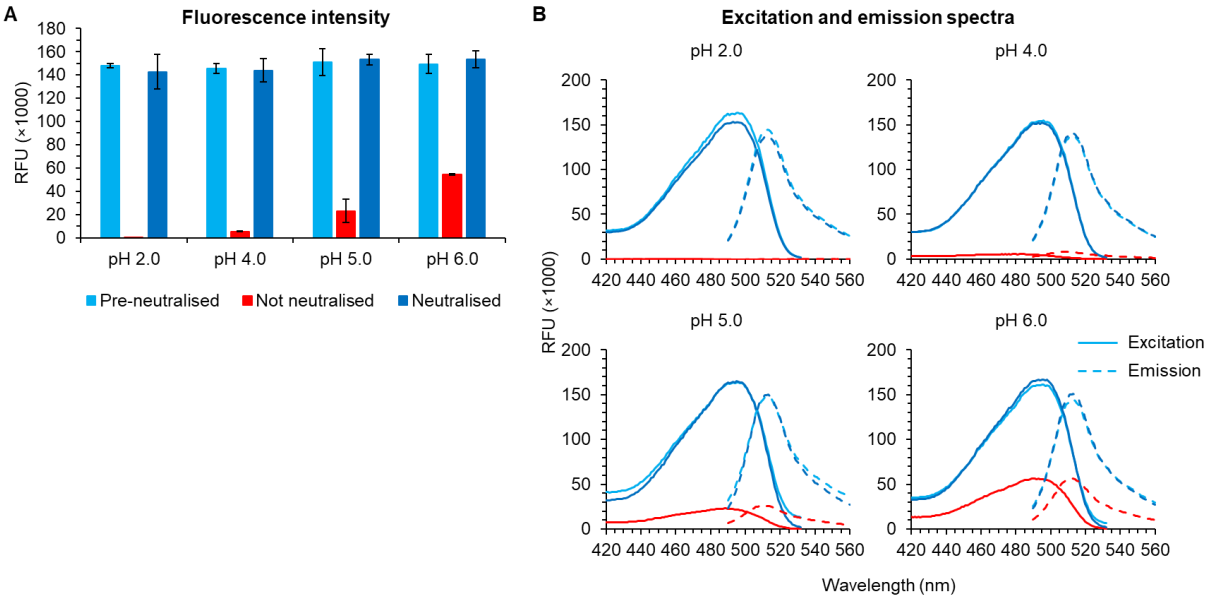


Figure 4.4 SpyTag003-sfGFP low pH tolerance and fluorescence. SpyTag003-sfGFP was incubated in low pH buffers of pH 2.0, 4.0, 5.0 and 6.0 (buffer composition in section 4.2.2) at 4 °C for 1 h and neutralised with 1 M Tris-HCl pH 8.0 to approximately pH 7.0 (Neutralised; navy). (A) The fluorescence intensity at λ_{em} = 520 nm with 20 nm bandpass was measured after being excited at λ_{ex} = 485 nm with 12 nm bandpass. Pre-neutralised (cyan): SpyTag003-sfGFP incubated in pH 2.0-6.0 buffers already neutralised with 1 M Tris-HCl pH 8.0 as folded fluorescent control. Not neutralised (red): SpyTag003-sfGFP incubated in pH 2.0-6.0 buffers without neutralisation after 1 h (mean ± 1 s.d, *n* = 3). (B) Excitation and emission spectra of SpyTag003-sfGFP in pre-neutralised, not neutralised or neutralised state. RFU is relative fluorescence unit in relation to 80 µM purified SpyTag003-sfGFP as reference.

4.2.3 Comparative stepwise pH elution from histidine mutants

To systematically determine pH-sensitive mutants, each SpyCatcher003 S49C E77A with histidine substitutions was immobilised onto a solid-phase matrix similar to section 3.1.2 and incubated with purified SpyTag003-sfGFP at 4 °C for 45 min in HEPES pH 8.0 buffer. After the flow-through was removed, the resin was washed 4 times with HEPES pH 8.0 and subsequently eluted in a stepwise manner from MES pH 6.0 buffer, sodium acetate pH 5.0 buffer, sodium acetate pH 4.0, to glycine pH 2.0. Each pH elution was immediately neutralised with 1 M Tris-HCl pH 8.0 by pre-loading the neutralisation buffer in the collection tubes. Each elution was incubated for 10 min and done twice to fully exchange the pH of each buffer. Eluted SpyTag003-sfGFP was then compared by measuring fluorescence from the different

pH elution fractions. The assay was also compared with SpyDock over the pH range to confirm gain in pH sensitivity of histidine mutants.

15 out of 16 of the histidine mutants retained binding at pH 8.0 when compared to SpyDock, indicating that the mutations are permissible for binding at this pH. The exception was D33H, which had 3-fold more SpyTag003-sfGFP in the flow-through compared to SpyDock. Thus, D33H was removed from consideration due to impaired initial binding. No mutants eluted SpyTag003-sfGFP at pH 6.0 and only minor fluorescence signal was detected at pH 5.0 for 1 mutant, S30H. Among the 15 mutants, five had considerably increased SpyTag003-sfGFP fluorescence in the elution fraction after exposure to pH 4.0 buffer, which were S30H, E34H, G83H, Y84H, and E85H (Figure 4.6 A). In comparison, the remaining 10 mutants had minor SpyTag003-sfGFP elution at pH 4.0 whereas SpyDock still retained binding of SpyTag003-sfGFP at this pH. All constructs had impaired binding with SpyTag003-sfGFP at pH 2.0 with increased observed fluorescence intensities in these elution fractions, confirming the bound state of most SpyTag003-sfGFP from pH 8.0 to pH 4.0. This preliminary result showed promise in the generation of a non-reactive SpyCatcher with reduced binding affinity to SpyTag003 at low pH. Although elution at pH 4.0 is possible with a minimal single mutation, this pH is still too low to be applicable for all proteins. Incorporating multiple histidines into a single construct may shift the transition elution pH higher through cooperativity interaction. Hence, the top five mutants were brought forward and combined to form 10 different double mutants in order to achieve a stronger cooperative effect for increased elution yield upon exposure to low pH.

Each double mutant was subjected to the same assay and compared with the five single mutants plus SpyDock. From the 10 double mutants, six had similar elution profiles as one of the top single mutants, Y84H, indicating that these double mutant combinations did not contribute to the additive effect of their single mutants (Figure 4.6 B). The double mutant S30H + E34H had diminished binding capacity since its flow-through had a high SpyTag003-sfGFP fluorescence intensity relative to the other constructs, thus this double mutant was eliminated.

Three double mutants showed enhanced elution profiles, namely S30H + G83H, S30H + Y84H, and S30H + E85H. These double mutants had more elution of SpyTag003-sfGFP from pH 5.0 buffer compared to the single mutants, albeit most elution still occurred after incubation with pH 4.0 buffer. Nonetheless, the elution yield at the first pH 4 fraction was higher than the single mutants, indicating that multiple histidine cooperativity was effective to give sharper pH 4.0 elution. These three double mutants were then brought forward and combined to form three triple mutants to assess if even more cooperative effect will be produced.

The triple mutants underwent the same assay and the results were compared with the four relevant single mutants (less E34H), the top three double mutants and SpyDock. All the triple mutants had showed some elution of SpyTag003-sfGFP at pH 6.0, however most elution occurred at pH 5.0 (Figure 4.6 C). The triple mutants S30H + G83H + Y84H and S30H + G83H + E85H have similar elution profiles to the double mutant S30H + Y84H, thus they were eliminated. Additionally, S30H + G83H + E85H had impaired initial binding because of the relatively high SpyTag003-sfGFP fluorescence in flow-through, limiting its use. The triple mutant S30H + Y84H + E85H showed good initial binding and a high elution yield on the first pH 5.0 elution fraction, demonstrating that this combination of histidines managed to shift the transition pH towards less acidic conditions compared to its relevant single and double mutants, along with enabling a sharper elution at pH 5.0. These data indicate successful cooperativity from three distinct histidines along the interface with SpyTag003. Therefore, SpyCatcher003 S49C E77A with S30H + Y84H + E85H mutations was selected to become SpySwitch for the affinity purification of SpyTag003-proteins using a mild elution condition of pH 5.0 buffer.

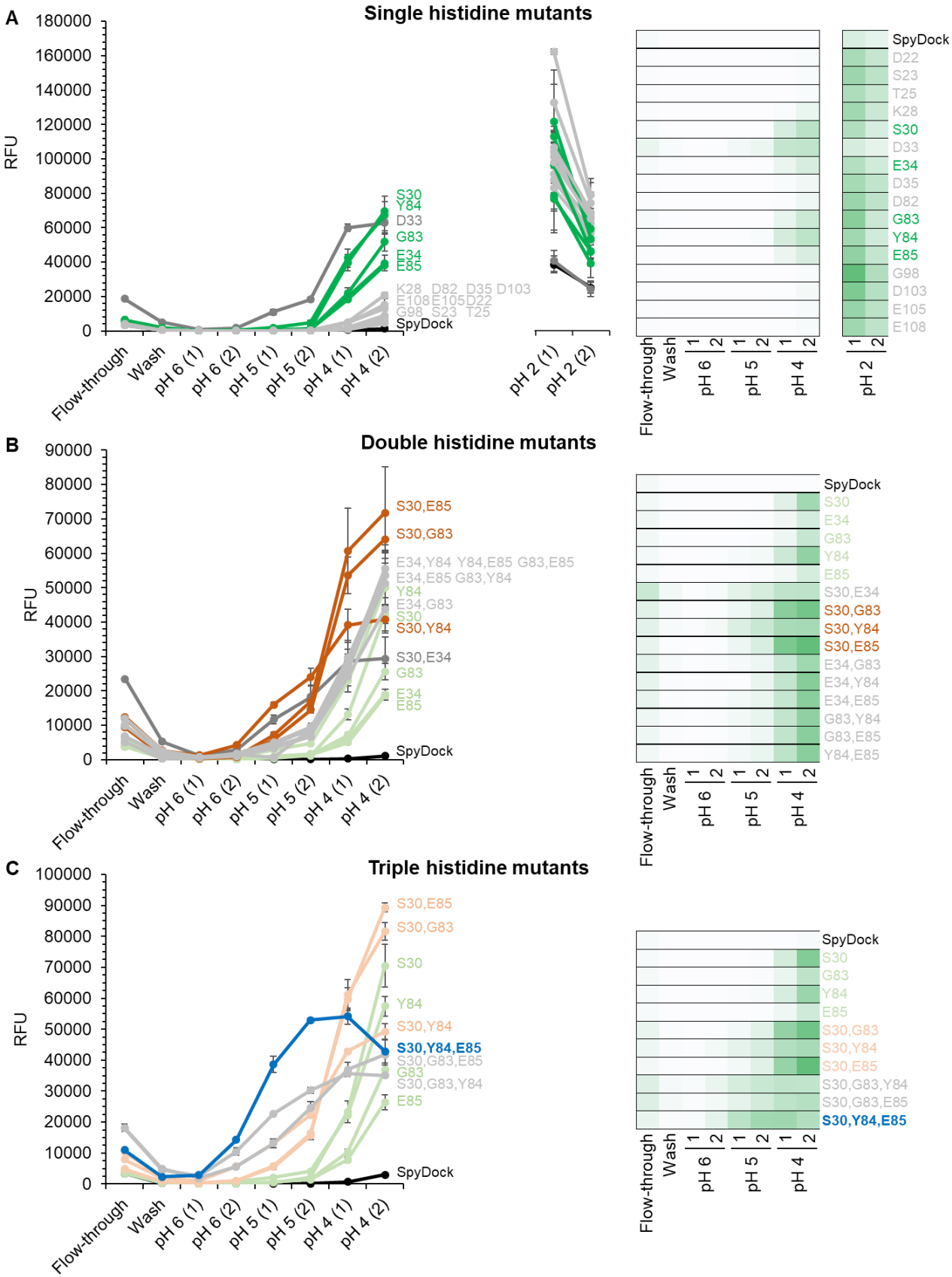

A
Single histidine mutants
RFU
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
S30
Y84
D33
G83
E34
E85
K28 D82 D35 D103
E108 E105 D22
G98 S23 T25
SpyDock
Flow-through
Wash
pH 6 (1)
pH 6 (2)
pH 5 (1)
pH 5 (2)
pH 4 (1)
pH 4 (2)
pH 2 (1)
pH 2 (2)
SpyDock
D22
S23
T25
K28
S30
D33
E34
D35
D82
G83
Y84
E85
G98
D103
E105
E108
pH 6
pH 5
pH 4
pH 2
B
Double histidine mutants
RFU
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
S30,E85
S30,G83
E34,Y84 Y84,E85 G83,E85
E34,E85 G83,Y84
Y84
E34,G83
S30
S30,Y84
S30,E34
G83
E34
E85
SpyDock
S30,E34
S30,G83
S30,Y84
S30,E85
E34,G83
E34,Y84
E34,E85
G83,Y84
G83,E85
Y84,E85
C
Triple histidine mutants
RFU
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
S30,E85
S30,G83
S30
Y84
S30,Y84
S30,Y84,E85
S30,G83,E85
G83
S30,G83,Y84
E85
SpyDock
S30,G83
S30,Y84
S30,E85
S30,G83,Y84
S30,G83,E85
S30,Y84,E85

Figure 4.5 Rational determination of pH-sensitive SpySwitch using stepwise elution of SpyTag003-sfGFP. Purified SpyTag003-sfGFP was bound to variants of immobilised SpyCatcher003 S49C E77A histidine mutant in HEPES pH 8.0 at 4 °C and washed with the same buffer 4 times. Stepwise elution of SpyTag003-sfGFP was triggered by incubating the resin with buffers of decreasing pH from MES pH 6.0, sodium acetate pH 5.0, sodium acetate pH 4.0, to glycine pH 2.0. The fractions were neutralised upon elution and the fluorescence intensity at λ_{em} = 520 nm with 20 nm bandpass was measured (Figure 2.1). The measurements were dilution-corrected and compared among the different histidine mutants and SpyDock. Fluorescence intensities (mean ± 1 s.d, *n* = 3) of (A) single histidine mutants, (B) double histidine mutants, and (C) triple histidine mutants at different purification fractions are shown. Two elution fractions, (1) and (2), for each pH are shown. Fluorescence dot matrices are shown on the right of each corresponding chart. The matrices provide mutant-level detail of the amount of sfGFP fluorescence intensities found in each fraction for fraction-by-fraction comparison (white: low fluorescence, green: high fluorescence). RFU: relative fluorescence units. Eliminated mutants (grey), SpyDock (black), selected single histidine mutants (green), selected double histidine mutants (orange), selected triple histidine mutant (blue).

4.2.4 pH-dependent elution of SpyTag003-fusions from lysate

The experimental design to construct SpySwitch was only analysed by the presence of fluorescence of purified SpyTag003-sfGFP that was eluted in different pH buffers. The workflow was chosen to enable a high throughput method to screen a total of 29 different histidine mutants quickly. However, the usage of purified SpyTag003-sfGFP does not emulate real purification conditions of complex cellular lysate mixtures surrounding the target protein. In section 3.2.2, I tested three washing conditions to determine the buffer of choice to remove non-specifically bound cellular material from SpyDock. Here, I tested the same core conditions in HEPES pH 8.0. They were 1) physiological wash of 50 mM HEPES, 150 mM NaCl at pH 8.0, 2) alternating high and low salt wash of 50 mM HEPES, 1 M NaCl at pH 8.0 (high salt) and 50 mM HEPES no salt at pH 8.0 (low salt) with binding in low salt buffer, and 3) 500 mM imidazole wash of 50 mM HEPES, 500 mM imidazole at pH 8.0. Purified SpyTag003-sfGFP was spiked into clarified *E. coli* lysate resuspended and lysed in HEPES pH 8.0 buffer. The sample was incubated with immobilised SpySwitch for 45 min, washed 4 times with each washing buffer, and eluted 3 times with sodium acetate pH 5.0 buffer into a collection tube pre-loaded with neutralisation buffer.

Between the three wash conditions, all had comparable purity when eluting SpyTag003-sfGFP from immobilised SpySwitch, but the alternating high and low salt washing regimen yielded the most SpyTag003-sfGFP based on SDS-PAGE band intensity of the total pooled elution (Figure 4.7 A). However, both physiological wash and alternating high and low salt wash co-eluted a 17 kDa contaminant. This band may be a cellular contaminant or incompletely-conjugated SpySwitch that was only removed with a harsher 500 mM imidazole wash. If the latter, a more thorough coupling procedure to attach SpySwitch to the solid-phase matrix can be utilised. If the former, more additives in the wash buffer such as a low imidazole concentration or the detergent Tween 20 can be used to remove tightly bound contaminants. The identity of the co-purified band will need to be verified through in-gel digestion mass spectrometry (Shevchenko et al., 2006). In the individual purification fractions from the alternating high and low salt washes (Figure 4.7 B), some SpyTag003-sfGFP was recovered in the flow-through and washes, since these fractions had a fluorescent band corresponding to the size of SpyTag003-sfGFP. This result indicates that the binding between SpySwitch and SpyTag003 may not be optimal to ensure efficient sequestration of SpyTag003-proteins from cell lysate. These two issues demonstrate that SpySwitch requires additional optimisation to enable efficient coupling to a solid-phase matrix and increased binding at pH 8.0. Nonetheless, SpySwitch was successfully shown to facilitate good purity purification of SpyTag003-sfGFP from cell lysate using a mild pH 5.0 elution buffer.

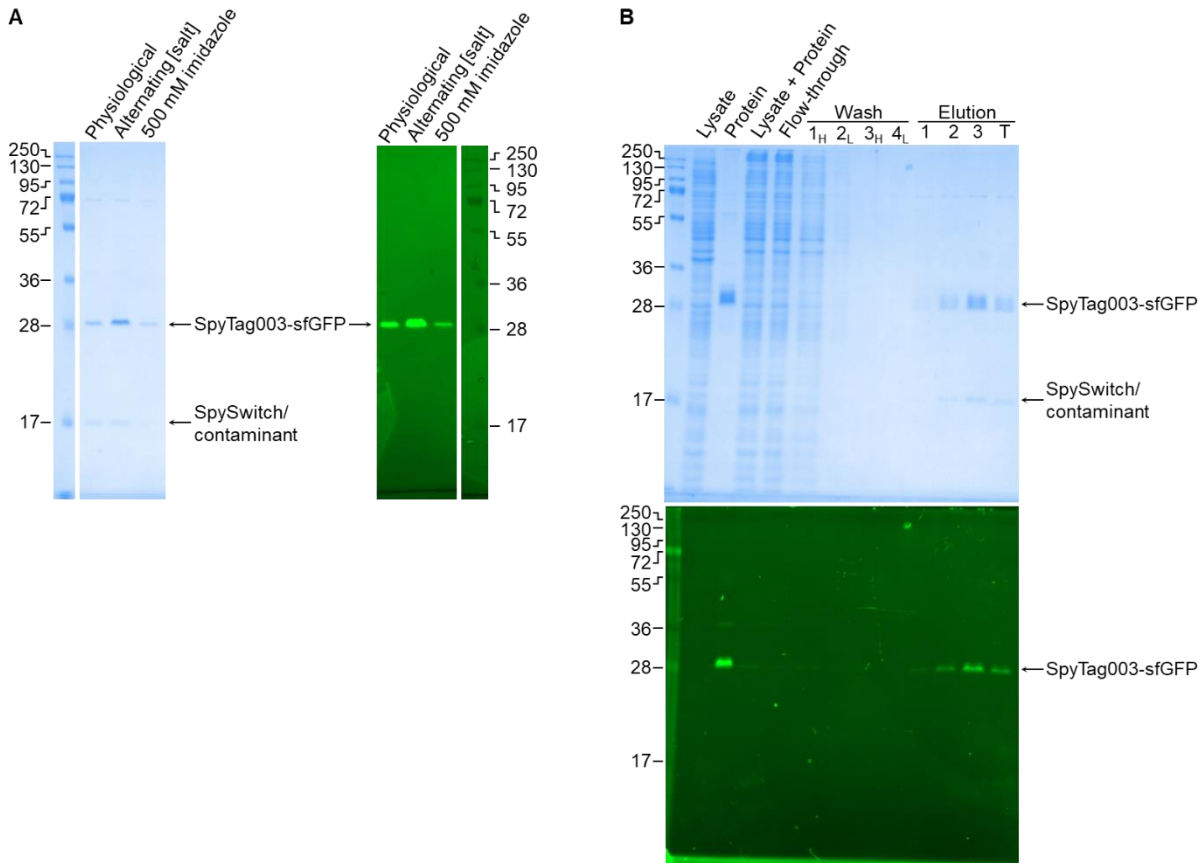


Figure 4.6 Purification of SpyTag003-sfGFP from cell lysate. (A) Optimisation of washing condition for purification of SpyTag003-sfGFP bound on SpySwitch. SpyTag003-sfGFP was spiked into *E. coli* cell lysate mixed with immobilised SpySwitch. Non-specifically bound cellular material was washed with either physiological buffer (50 mM HEPES, 150 mM NaCl pH 8.0), alternating [salt] buffer [50 mM HEPES, 1 M NaCl at pH 8.0 (high salt) and 50 mM HEPES no salt at pH 8.0 (low salt)], or 500 mM imidazole (50 mM HEPES, 500 mM imidazole at pH 8.0). Each purification was eluted with sodium acetate pH 5.0 and the total pooled elution shown in the representative Coomassie-stained SDS-PAGE with accompanying blue light-illuminated SDS-PAG for fluorescence visualisation (*right gel*). (B) Full purification of spiked SpyTag003-sfGFP from *E. coli* lysate washed with alternating [salt] buffer from (A). Washes H: 1 M NaCl, washes L: no NaCl, T: total pooled elution. Accompanying blue light-illuminated SDS-PAGE (*bottom*) for fluorescence visualisation.

4.3 Discussion

Chapter 3 established a purification method to purify SpyTag-proteins. However, the high imidazole elution buffer may not be applicable to all proteins. Here, I explored three different methods to enable a milder elution condition for SpyTag-protein purifications. The first two methods explored alternative elution conditions on the existing Spy&Go platform to

provide a quick solution rather than developing a new purification platform capable of mild elution. Both methods, competitive peptide elution and on-resin tag cleavage, have been widely used for existing affinity purification platforms (Kimple et al., 2013; Waugh, 2011). I optimised the working conditions for these methods to suit Spy&Go.

The latest generation SpyTag003 was selected for competitive peptide elution because the peptide has the highest second order reaction rate constant to all generations of SpyCatcher compared to SpyTag and SpyTag002 (Keeble et al., 2019). SpyTag003 was also designed to have enhanced binding with SpyCatcher003, making it the best candidate peptide to displace SpyTag-, SpyTag002- and SpyTag003-proteins from SpyDock. Indeed, 1 mM SpyTag003 peptide incubated with occupied SpyDock at 30 °C for 2 h managed to elute MBP fused to SpyTag, SpyTag002 and SpyTag003. For established epitope tag affinity chromatography methods such as FLAG-tag, c-Myc tag, HA-tag and V5 tag, the elution peptide concentrations used are between 200 – 500 μM (Tomomori-Sato et al., 2013). Some newly developed epitope tags show elution at concentrations as low as 100 μM (Brown and Takagi, 2018; Takeda et al., 2017), however a high peptide concentration up to 1 mM may still be preferred for better elution yields (Braun et al., 2016; Nogi et al., 2008). The high SpyTag003 peptide concentration required to elute SpyTag-proteins from SpyDock may be cost-limiting since chemical synthesis of peptides is expensive. Producing unmodified peptides recombinantly in bacterial expression system is not a viable option because peptides are prone to aggregation and proteolytic cleavage (Wegmuller and Schmid, 2014). Fusing the peptide to a cleavable folded protein component has been shown to be a practical alternative for recombinant production of peptides (Zhao *et al.*, 2016; Buldun *et al.*, 2018). In this method, the cleavable protein component folds properly in the cell to avoid degradation; once purified, the protein component is cleaved off to release the short peptide. Recombinant peptide production through this method may offer a cheaper alternative for the mild SpyTag003 elution condition. Previously, I conducted initial attempts at mass-production of SpyTag003 from proteolytic cleavage of SUMO-SpyTag003. However, peptide concentrations of 1 mM were

difficult to achieve using this method. Thus, solid-phase synthesised peptide was used instead to investigate the feasibility of competitive peptide elution. Further optimisation efforts in recombinant SpyTag003 production will need to be undertaken.

On-resin tag cleavage can also offer a mild elution alternative for Spy&Go. However, the removal of SpyTag from the protein of interest will defunctionalise the protein from SpyCatcher bioconjugation reactions, thus this method may not be a fully practical alternative. Nonetheless, with the high purity offered by Spy&Go, on-resin tag cleavage can be used as an alternative purification of proteins requiring tag removal. In summary, both methods provide an alternative to the harsher imidazole elution currently used, however further considerations must be taken regarding reagent cost, protein stability in reaction condition, and the need for SpyTag after purification.

Because of the constraints of the mild elution conditions on the current Spy&Go platform, I decided to engineer a new SpyTag binder capable of mild elution. Designing a reversible SpyTag binder with inducible local or global conformational change is a plausible strategy to bind and release SpyTag through a stimulus change. Different methods to incorporate different stimuli-responsive allosteric control into a protein have been extensively developed, for example sensitivity towards temperature (Chakshusmathi et al., 2004), Ca^{2+} (Stratton et al., 2008), maltose (Guntas et al., 2004), light (Lee et al., 2008), and rapamycin (Dagliyan et al., 2013). However, these protein engineering methods usually proceed via the insertion of a stimuli-dependent domain into the protein of interest, which can be unpredictable and dependent on the stability of the target protein (Stein and Alexandrov, 2015). The usage of small molecules as the inducer for conformational switch to deactivate protein function may also increase cost when translated into the area of affinity chromatography. A different stimulus, pH, is more useful since no protein domain insertion is needed to confer stimuli-sensitivity and the cost of buffers is low.

Most methods that engineer pH-switchability into proteins utilise protein display technologies for directed evolution (Könning et al., 2018; Murtaugh et al., 2011; Schröter et

al., 2015) or computational design to deduce the locations to introduce ionisable residues (Ambroggio and Kuhlman, 2006; Boyken et al., 2019; Strauch et al., 2014). Computational design of pH switches can help reduce experimentation time through simulations to provide useful insights, but the method is reliant on the availability of the tertiary and quaternary structure of proteins in complexes, and thus may not be applicable to all proteins. Directed evolution through display technologies can uncover mutations that rational design and computational design may overlook, but the experimental design and setup require optimisations that take time. Thus, as a proof of concept, I decided to rationally introduce histidine mutations into SpyCatcher003 S49C E77A to demonstrate a gain in pH-switchability.

From the 16 residues targeted for histidine substitution, only six showed a gain in conferring noticeable pH sensitivity for immobilised SpyCatcher003 S49C E77A before pH 2.0, which demonstrates the plasticity of SpyCatcher003 to tolerate changes in structure upon histidine substitutions at multiple residues. From the SpyTag:SpyCatcher crystal structure 4MLI, the five selected histidine mutations showed evidence of self-destabilisation and SpyTag003 binding destabilisation upon protonation of histidine at low pH. In S30H$^+$ (whereby H$^+$ denotes protonated histidine), the positive charge of histidine may repel the positive charge of nearby R32 and attract the negatively-charged E38, further weakening the E38-R32 ionic interaction to locally destabilise SpyCatcher003 folding. S30H$^+$ may also repel the protonated H112 of SpyTag003 to reduce binding interactions (**Error! Reference source not found.** A).

The E34H substitution lost its polar interaction from E34 to the amine group of Y119 of SpyTag003. E34 is also postulated to form an anion-quadrupole interaction with the partial positively-charged ring edge of SpyTag003 Y119 (Chakravarty et al., 2018; Jackson et al., 2007). However, this interaction may be lost upon E34H mutation but may be substituted with the partial negative π orbital of E34H to the partial positive Y119 ring edge. Upon protonation, E34H$^+$ would destabilise this interaction to repel Y119 and SpyTag003 as a whole (**Error! Reference source not found.** C).

The imidazole ring of the G83H substitution is theorised to face only the aliphatic carbons of SpyTag003 K120. However, G83H$^+$ may repel the positively-charged K120 of SpyTag003 to interfere with the K120-E85 ionic interaction (**Error! Reference source not found.** B).

In Y84H, the original residue Y84 forms a π-stacking interaction with Y119 of SpyTag003, but this was then replaced with a weaker π-stacking with unprotonated Y84H. Y84H$^+$ would make the π-stacking even weaker to further destabilise binding with SpyTag003 (**Error! Reference source not found.** D).

In the protonated state, E85H$^+$ would incur destabilising and repulsive interactions with the positively-charged SpyTag003 K120 and P121R. This is in contrast to the unprotonated E85H state where the positively-charged K120 can form a cation-π interaction with E85H (Gallivan and Dougherty, 1999; Liao et al., 2013) (**Error! Reference source not found.** B).

One mutation merits discussion regarding the effect of histidine at the binding interface. D33H was the most destabilising mutation, even during binding and washing at pH 8.0. D33 is theorised to form a hydrogen bond with the phenolic hydroxyl group of SpyTag003 Y119, but this is lost upon D33H substitution. The nearby R37 may also form ionic interactions with D33 and D35, but this may be lost upon the bulky and uncharged D33H substitution. These two effects, both SpyTag003 binding destabilisation and local self-destabilisation, may explain the higher intensity of SpyTag003-sfGFP in pH 8.0 fractions. These effects may be worsened upon D33H$^+$ protonation.

Upon the combination of S30H + Y84H + E85H to make the final SpySwitch construct, the three substitutions do not act synergistically to destabilise the same residue. S30H is distant from Y84H and E85H, so its effects are localised to the first binding interface region near the N-terminus of SpyTag003. Y84H and E85H act nearer to the C-terminus of SpyTag003, but the residues are side-by-side, so they face different planes to interact with different SpyTag003 and SpyCatcher003 residues. Nonetheless, the combination of all the SpyTag003 and SpyCatcher003 residue perturbations collectively enable effective

destabilisation of SpyTag003 binding to SpyCatcher003. The destabilising effects within the microenvironment of $E34H^+$ and $G83H^+$ may not be strong enough to elicit a higher pH transition, which may be explained by incomplete protonation of histidines for these residues. However, the hypotheses given for these mutations need confirmation by nuclear magnetic

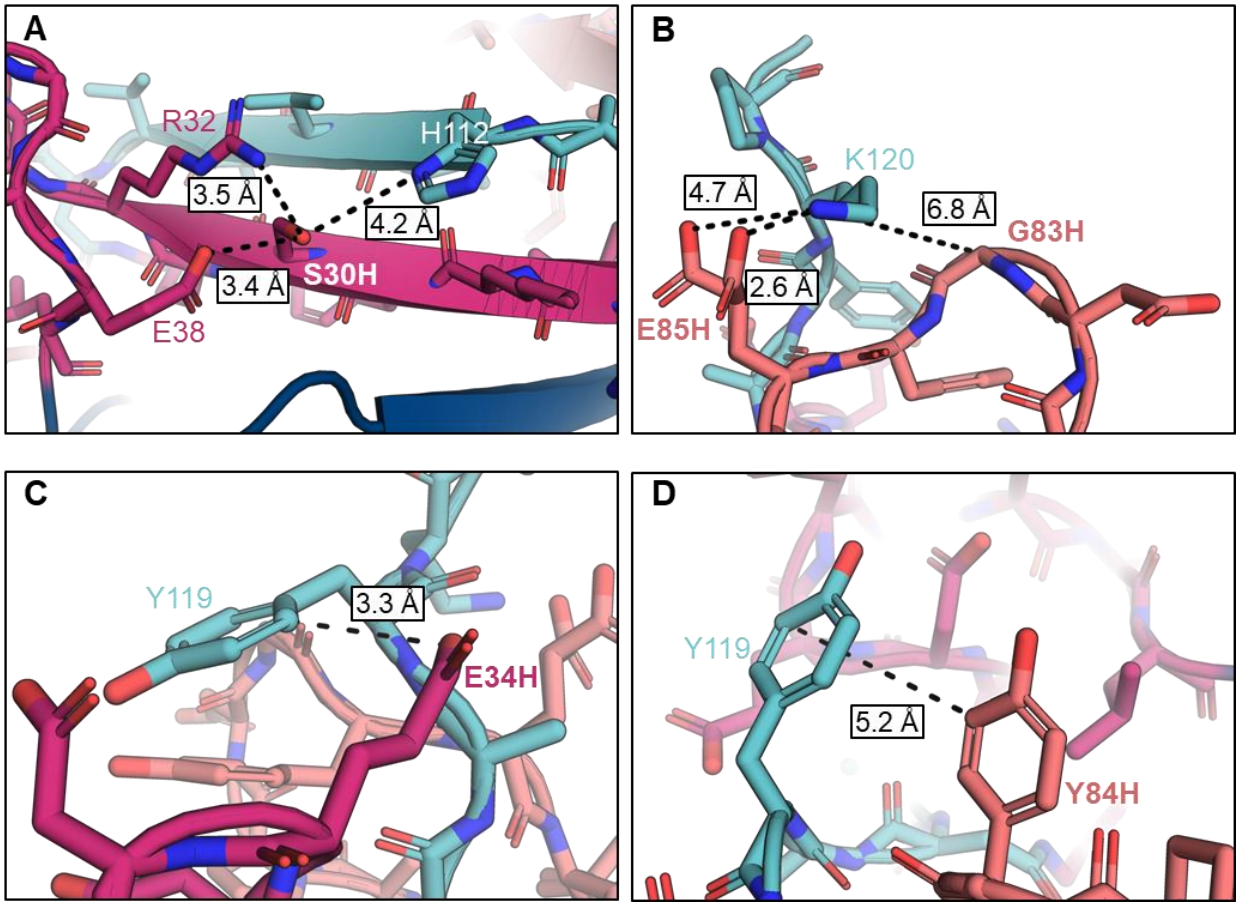


Figure 4.7 Structural analysis of selected histidine mutants. PDB ID 4MLI of SpyTag:SpyCatcher binding shows possible interactions and perturbations with nearby amino acids from the five selected histidine mutants (*bolded*) (A) S30H, (B) G83H, E85H (with two conformers present as detailed in Figure 4.3 C), (C) E34H, and (D) Y84H. SpyTag (cyan), SpyCatcher (blue), interface 1 of SpyCatcher (magenta), interface 2 of SpyCatcher (salmon). Black dotted-lines show polar contacts between residues with distances measured computationally. The explanation for possible interactions is found in-text in section 4.3.

resonance (NMR) spectroscopy and X-ray crystallography studies.

The final SpySwitch construct was tested to purify spiked SpyTag003-sfGFP from cell lysate. The conditions tested, which were binding at pH 8.0, washing with alternating high and low salt at pH 8.0, and eluting at pH 5.0, were successful to yield a SpyTag003-sfGFP free

from most cellular contaminants. Nonetheless, exposing proteins to low pHs such as pH 5.0 may not be ideal to retain protein structure and function. In a study of 310 sampled proteins, only ~10% have optimum pH of stability at pH 5.0 and below, with the mean at pH ~7.0 (Talley and Alexov, 2010). The authors postulated that the optimal pH of protein activity is correlated with the optimal pH of protein stability; thus, as the mean optimal pH of stability is ~pH 7.0, exposure to pH 5.0 such as the elution condition for SpySwitch may not favour retained activity in many proteins. However, the authors also showed that proteins can tolerate pH changes of 1-2 units and still retain protein activity since different subcellular microcompartments with different pHs are available such as organelles along the secretory pathway (Talley and Alexov, 2010; Wu et al., 2000). The plasticity in activity towards changes in pH may alleviate the concerns of the pH 5.0 elution from SpySwitch. During elution from SpySwitch, the protein of interest may be exposed to pH 5.0 for 10 min per elution, which can add up to 40-60 min of exposure. However, immediate neutralisation upon elution may alleviate structural and functional impairments caused by pH 5.0. A database of >5,000 proteomes across various kingdoms of life showed the mean isoelectric point (pI) of proteins is approximately 6.7 (Kozlowski, 2017), but some proteins have a pI close to 5.0. For example, in *Homo sapiens*, the average pI of the proteome (*n* = 92,179) is 6.81, but 14.8% (13,614) of proteins have pI of <5. In *E. coli*, the percentage of proteins with pI <5 is 14.9% (644 out of 4,314) (Kozlowski, 2017). SpySwitch itself has a pI of 4.8 estimated from ProtParam (Gasteiger et al., 2005). For these proteins, elution at pH 5.0 may lead to pI-induced aggregation, which may be reversible upon neutralisation to pH > pI.

SpyTag003:SpyCatcher003 does not have a crystal structure to properly inspect mutations to confer pH-switchability. Thus, I was unable to specifically target mutations in SpyCatcher003 S49C E77A to destabilise the positively-charged N- and C-termini of SpyTag003 (bolded residues are unavailable in the 4MLI structure: **R**GVPHIVMVDAYK**R**Y**K**). However, the results gathered here show promise in the rational design of pH-sensitive

SpySwitch, with only three introduced histidines, for the purification of SpyTag003-proteins under a mild elution condition.

5 Expansion of Spy oligomerisation toolbox to probe multivalent cell signalling

Combining protein domains with different properties into a single entity can create multifunctional protein assemblies for innovative and commercial uses in biotechnology and synthetic biology. The traditional method of genetic fusion to link two proteins together is limited by improper folding, arduous individual optimisation and suboptimal expression systems for post-translational modifications. Modularity can separate protein components for subsequent linkage and is thus preferred over genetic fusion to alleviate these problems. Chemical conjugation is typically used to modularly link the components but is itself limited by heterogeneous coupling and the potential of crosslinking-mediated aggregation (Brune and Howarth, 2018).

To this end, the Howarth lab utilised the Tag/Catcher technology to create platforms to decorate nanoparticles for vaccination. These platforms consist of a central scaffold protein fused to SpyCatcher to post-translationally conjugate SpyTag-antigens. The scaffolds currently explored include a 180-mer virus-like particle AP205 phage coat protein (Brune et al., 2016), a 60-mer engineered nanocage mi3 (Bruun et al., 2018), and a heptameric oligomer IMX313 (Brune et al., 2017). These scaffolds laid the foundation for the establishment of a nanoassembly toolbox for modular genetically-encodable protein conjugation, which has since seen an expansion in the valency and type of scaffold available (Bae et al., 2018; Cayetano-Cruz et al., 2018; Choi et al., 2018; Giessen and Silver, 2016; Krebs et al., 2019; Swartz and Chen, 2018; Wang et al., 2019a). However, alternative assembly scaffolds have not been explored that utilise the Tag/Catcher technology for other defined valencies to enable further investigation into different protein architecture, geometry and stoichiometry.

Multivalency has been known to increase the functional avidity of low binding affinity interactions and may reduce off-target cytotoxicity by enabling better discernment of target cells from healthy cells compared to using high affinity binders (Rudnick and Adams, 2009; Slaga et al., 2018). Recently, an emergent property of multivalency was proposed which was termed superselectivity. Superselectivity refers to the near switch-like effect of multivalent interactions on cell receptors, where a small increase in ligand or receptor density can have a profound effect in modulating cellular behaviour (Curk et al., 2017a; Dubacheva et al., 2014). The current methods to study multivalency are limited by heterogeneous densities of ligands during production of the multivalent probes (Kiessling et al., 2006; Satav et al., 2015). Defined valencies can be achieved through chemical conjugation but requires non-biological material and is limited to conjugating chemical entities, not proteins (Dubacheva et al., 2015; Puffer et al., 2007).

Combining the modularity from the SpyTag/SpyCatcher technology with natural and synthetic oligomeric peptides, I describe here the development of defined SpyCatcher oligomers from dimer to heptamer to enable a faster and simpler plug-and-display approach to study the effect of multivalent binding on cellular signalling. Together with currently established Spy nanoparticles, the SpyCatcher oligomers may have the potential to uncover new protein architecture and function in order to advance biotechnology, nanomedicine and synthetic biology.

5.1 SpyCatcher oligomers development

Coiled coil motifs are peptides consisting of two or more α-helices that self-assemble to wrap around each other through knobs-into-holes interactions to form supramolecular assemblies (Woolfson, 2005). Because of its direct sequence-to-structure relationship, the coiled coil motif has become one of the most extensively studied subject for protein engineering and *de novo* design, yielding hundreds of variations of coiled coils (Woolfson,

2017). Because of their oligomerising nature and variety, I decided to utilise coiled coils as the oligomerising platform to establish the Spy oligomerisation toolbox.

For the purposes of this toolbox, the orientation and subunit composition of coiled coils to be selected were limited to parallel homo-oligomers. Parallel coiled coils have the termini of each oligomer subunit facing the same direction, thus the target protein to be conjugated via SpyTag/SpyCatcher chemistry will face the same plane to exert its gained multivalent effect on cell signalling. Using homo-oligomers can simplify the workflow, to allow the expression and purification of only one self-assembling monomer. The Woolfson lab from the University of Bristol created computationally-designed *de novo* coiled coils from the knowledge of key amino acid residue interactions in the coil's heptad repeat. They established a panel of coiled coils from dimer to heptamer with correct assembly verified through biophysical and structural data (Fletcher et al., 2012; Thomson et al., 2014; Zaccai et al., 2011).

Coiled coils were selected from this panel for the Spy oligomerisation toolbox, namely the CC-Di (PDB ID 4DZM), CC-Tri (PDB ID: 4DZL), CC-Tet (PDB ID 3R4A), and CC-Hex2 (PDB ID: 4PN9) (Fletcher et al., 2012; Thomson et al., 2014). For a pentameric coiled coil fusion, the naturally-occurring oligomerisation domain of the cartilage oligomeric matrix protein (COMP) (PDB ID: 1VDF) (Malashkevich et al., 1996) was used instead of the Woolfson lab's CC-Pent (PDB ID: 4PN8) (Thomson et al., 2014) because expression of CC-Pent was toxic in *E. coli* (personal communication with Dek Woolfson). For the heptameric coiled coil, the oligomerisation domain of the C4-binding protein from chicken, IMX313 (PDB ID: 2YF2) (Ogun et al., 2008), was used instead of the Woolfson lab's CC-Hept (PDB ID: 4PNA) (Thomson et al., 2014) because SpyCatcher was already previously fused to IMX313 for vaccination (Brune et al., 2017). The pentameric COMP and heptameric IMX313 have disulfide bonds between each coil subunit to confer higher stability. The computationally-designed oligomers (CC-Di, CC-Tri, CC-Tet and CC-Hex2) had five heptad repeats in the coil sequence instead of the four heptad repeats reported in the literature, to increase the non-covalent coiled coil stability

(Thomas et al., 2013). These coiled coils were fused to the C-terminus of SpyCatcher002 to enable a faster reaction with SpyTag-proteins. A 9-residue GS linker separated SpyCatcher002 and the coil subunit for flexibility to minimise steric clashes (Figure 5.1).

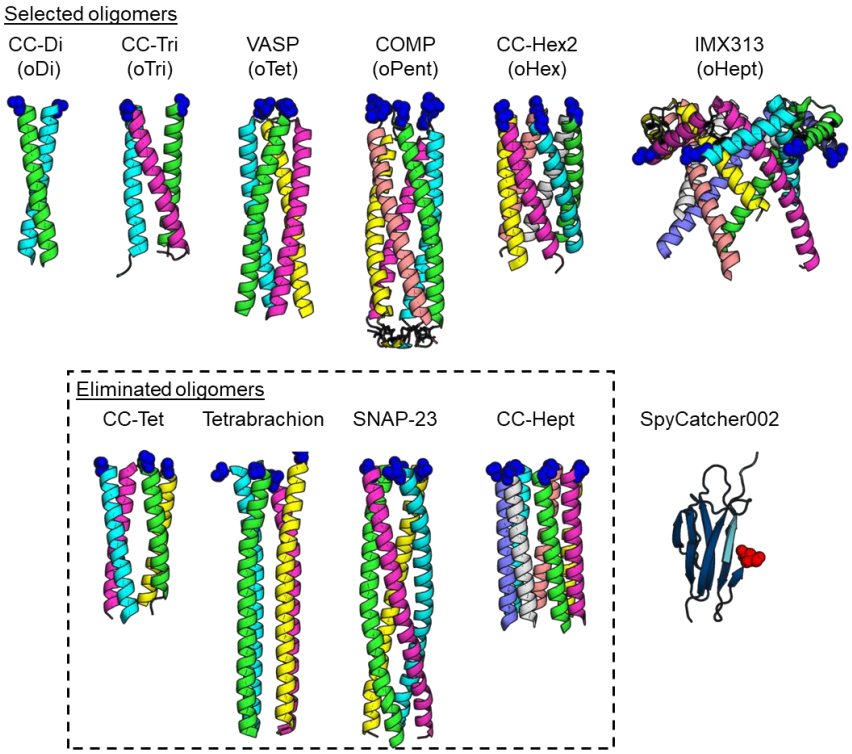


Figure 5.1 Structure of coiled coils and SpyCatcher002 used to construct SpyCatcher002-oligomers toolbox. SpyCatcher002-oligomers were constructed by genetically fusing SpyCatcher002 C-terminus (red ball) to the N-terminus of each subunit of the coiled coils (navy ball). The coiled coils used for the oligomerisation toolbox are presented here, with the top row being the finalised SpyCatcher002-oligomers and the boxed oligomers being the eliminated candidates of tetramers and heptamer. Each coiled coil colour represents one chain. Disulfide bonds are shown as black sticks in COMP (oPent) and IMX313 (oHept). SpyTag (cyan) position is shown in the SpyCatcher002 structure.

5.2 Characterisation of SpyCatcher002-oligomers assembly

Each SpyCatcher002-oligomer construct was expressed in *E. coli* and purified through a polyhistidine tag using Ni-NTA resin. The dimer, trimer, tetramer and pentamer were solubly expressed in the cytosol, whereas the hexamer and heptamer were found in inclusion bodies, necessitating denaturing purification similar to the heptameric SpyCatcher-IMX-SnoopCatcher

previously explored (Brune et al., 2017). To remove soluble aggregates and increase purity, each SpyCatcher002-oligomer construct was subjected to a polishing gel filtration chromatography step (Figure 5.2). From gel filtration, the molecular weight of the assembly, and hence the oligomericity, can be estimated by comparing to a standard curve of known protein molecular weight standards using the equation:

$$K_{av} = \frac{(V_e - V_0)}{(V_c - V_0)}$$

where K_{av} is the partition coefficient of a protein, V_e is the elution volume of a protein, V_0 is the void volume, and V_c is the total geometric volume of the gel filtration column. The main non-aggregate peak of each Ni-NTA-purified SpyCatcher002-oligomer was taken as the sample V_e and K_{av} of each construct was determined. However, the molecular weight of the SpyCatcher002-oligomers from gel filtration was overestimated, showing bigger assembly sizes than expected (Table 5.1). This can be explained by the structure of the coiled coil where the elongated shape gives a higher Stokes radius rather than the idealistic globular structure of the protein standards (Erickson, 2009).

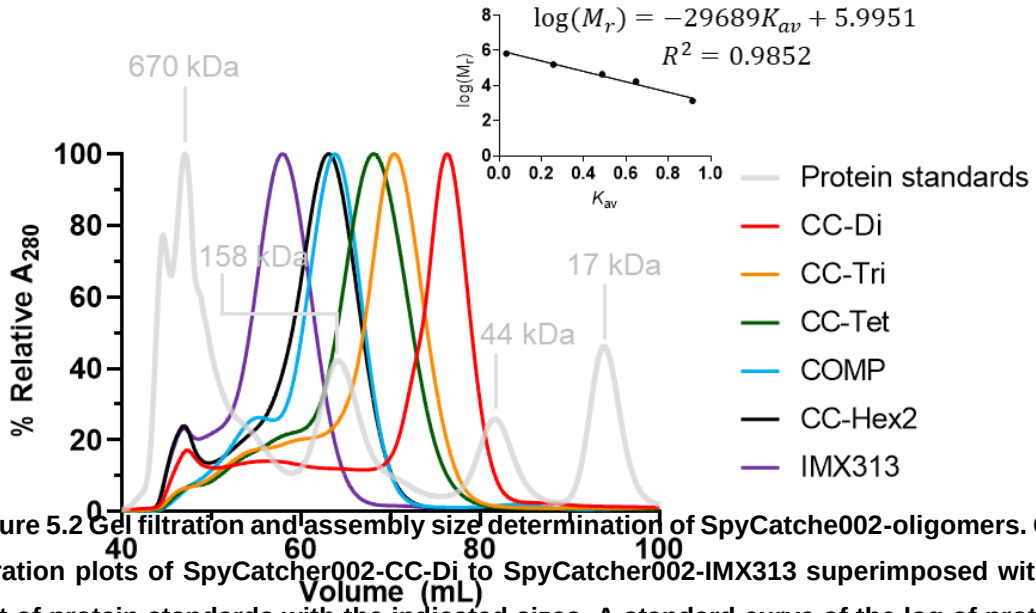


Figure 5.2 Gel filtration and assembly size determination of SpyCatche002-oligomers. Gel filtration plots of SpyCatcher002-CC-Di to SpyCatcher002-IMX313 superimposed with a plot of protein standards with the indicated sizes. A standard curve of the log of protein sizes against protein K_{av} (inset) was plotted based on the known sizes of the protein standards and measured elution volumes, and subsequently fitted with a linear regression model to find the line of best fit. The unknown sizes of SpyCatcher002-oligomers based on the measured elution volumes were then determined using the equations shown. Molecular weights are shown in Table 5.1.

Table 5.1 Size determination of SpyCatcher002-oligomers from gel filtration. The observed sizes were calculated using the equations shown based on the elution volume of each construct.

Protein	V_e (mL)	K_{av}	Log(M_r)	Observed M_r (kDa)	Expected M_r (kDa)	ΔM_r (kDa)
CC-Di	76.3	0.418	4.76	56.9	40.7	+16.2
CC-Tri	70.4	0.341	4.98	96.3	61.5	+34.8
CC-Tet	68.1	0.310	5.07	118.8	81.5	+37.3
COMP	63.8	0.253	5.24	175.0	109.5	+65.5
CC-Hex2	63.1	0.244	5.27	186.2	123.5	+62.7
IMX313	58.0	0.177	5.47	294.9	159.4	+135.5

Because of the overestimated molecular weight of SpyCatcher002-oligomers obtained from gel filtration, light scattering was used in the form of size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). SEC-MALS needs no prior knowledge of protein shape and size to determine the absolute molar mass and hence the oligomeric state of the protein. At 25 °C with a concentration of 2 mg/mL, the sizes of SpyCatcher002-oligomers were determined to be within 10% of the expected assembled molecular weight (Figure 5.3 and Table 5.2). However, this was not the case for SpyCatcher002-CC-Tet because the obtained molecular weight (62.8 g/mol) was 23% below the expected size (81.5 g/mol), nearing a trimeric assembly rather than a tetrameric oligomer. Examining the molar mass plot for SpyCatcher002-CC-Tet, the downward slope implied that two species of oligomers were present in the same peak fraction, indicating incomplete or perturbed assembly of the tetrameric coiled coil (Figure 5.3).

Alternative tetrameric coiled coils were tested to replace CC-Tet. Candidates included the stalk of tetrabrachion protein from *Staphylothermus marinus* (PDB ID: 1FE6) (Stetefeld et al., 2000), the tetramerisation domain of the human vasodilator-stimulated phosphoprotein (VASP) (PDB ID: 1USE) (Kühnel et al., 2004), and the N-terminal coiled coil domain of human SNAP-23 protein (PDB ID: 1NHL) (Freedman et al., 2003). At the same time, the computationally-designed heptameric coiled coil CC-Hept (PDB ID: 4PNA) was also tested.

Upon expression and purification, correct assembly of each coiled coil was examined by SEC-MALS. To note, expression of the tetramer candidate SpyCatcher002-SNAP-23 was low and gave low purification yield, thus this construct was eliminated from consideration. The two remaining tetramer candidates and the computationally-designed heptamer assembled to approximately the correct molecular weight from SEC-MALS (Table 5.2). Nonetheless, SpyCatcher002-CC-Hept was unstable at room temperature, forming precipitated aggregates overnight, which was not observed with the other SpyCatcher002-coiled coils. Between SpyCatcher002-tetrabrachion and SpyCatcher002-VASP, the latter had a smaller deviation of observed molecular weight from the expected molecular weight. Thus, the final constructs for the Spy oligomerisation toolbox consisted of CC-Di, CC-Tri, VASP, COMP, CC-Hex2, and IMX313 coiled coils. They were renamed to SpyCatcher002-oDi to SpyCatcher002-oHept, reflecting their oligomeric nature (“o”), the oligomeric state (e.g. “Di” for dimer, “Tri” for trimer and so on) and fusion to SpyCatcher002 (Figure 5.1).

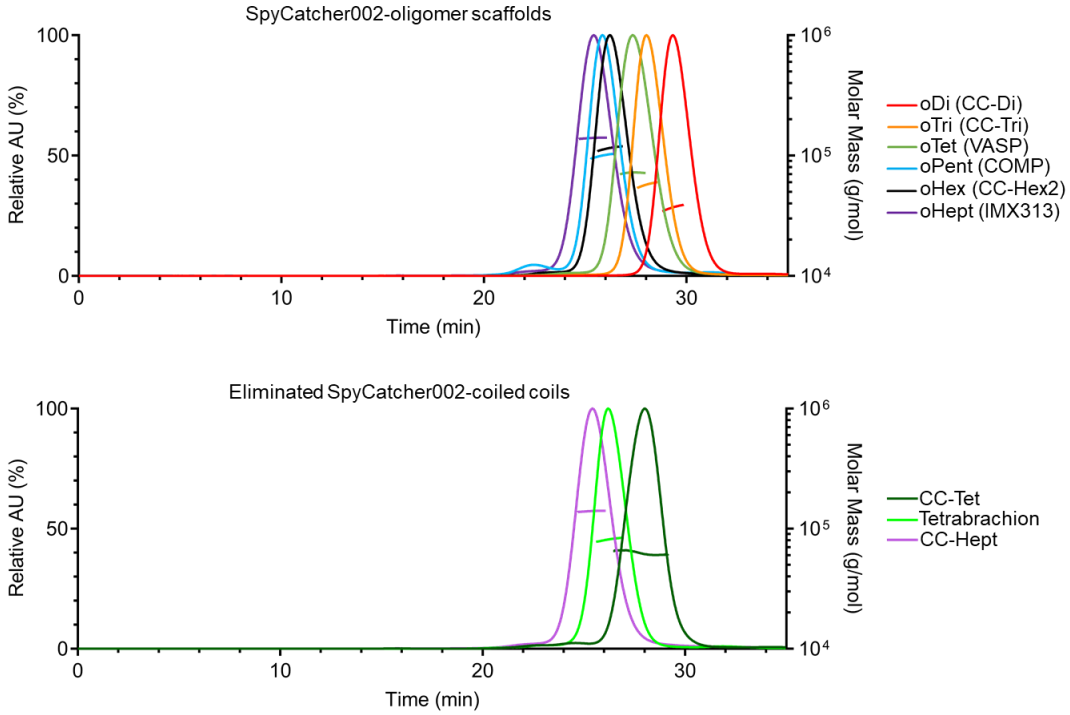


Figure 5.3 SpyCatcher002-oligomers assembly sizes determined by SEC-MALS. SEC-MALS profiles of different SpyCatcher002-oligomers, with the peaks showing the normalised absorbance units (AU) at 280 nm of each SpyCatcher002-oligomer species from SEC. The horizontal lines show the distribution of molar mass of the species within the peak. The observed molecular weights of these species are in Table 5.2. The top SEC-MALS profiles are for the finalised SpyCatcher002-oligomers, and the bottom profiles are for the eliminated candidates. Void volume is at approximately minute 15.

Table 5.2 Size determination of SpyCatcher002-oligomers from SEC-MALS. The observed sizes were determined from multi-angle light scattering (MALS). Coiled coils below the dotted line are the eliminated candidates. Error bars represent the uncertainty in fit to the molar mass curve. % error is 100 × (ΔM_r/ Expected Mr).

Protein	Observed Mr (kDa)	Expected Mr (kDa)	ΔM_r (kDa)	% error
oDi (CC-Di)	37.1 ± 0.4	40.8	-3.7	9.1
oTri (CC-Tri)	57.4 ± 0.3	61.5	-4.1	6.6
oTet (VASP)	79.3 ± 0.2	85.5	-6.2	7.2
oPent (COMP)	99.9 ± 0.2	109.5	-9.6	8.8
oHex (CC-Hex2)	115.5 ± 0.2	123.5	-8.0	6.5
oHept (IMX313)	150.5 ± 0.1	159.4	-8.9	5.6
CC-Tet	62.8 ± 0.5	81.5	-18.7	22.9
Tetrabrachion	81.2 ± 0.2	88.9	-7.7	8.7
CC-Hept	140.0 ± 0.1	143.5	-3.5	2.4

5.3 Oligomeric assembly of a nanobody

To demonstrate the utility of the Spy oligomerisation toolbox, I decided to multimerise an immunologically-relevant protein to study its effect on cellular signalling (Figure 5.4). Death Receptor 5 (DR5) is a member of the tumour-necrosis factor receptor superfamily containing an intracellular cytoplasmic death domain to signal for caspase-8-mediated apoptosis (Ashkenazi, 2002). Other members from this family include Death Receptor 4 (DR4) also containing a death domain, and decoy receptors 1 and 2 (DcR1 and DcR2) which lack a death domain or have a truncated death domain (Ashkenazi, 2002). These receptors are activated upon binding to their native ligand – a trimeric TNF-related apoptosis-inducing ligand (TRAIL). TRAIL clusters three DR5 subunits to activate downstream signalling (Ashkenazi, 2008). Targeting DR5 activation has been recognised as a promising avenue for cancer treatment because TRAIL-activated DR5 showed cytotoxicity for cancer cells while minimally affecting the viability of normal cells (MacFarlane, 2003). Because of this feature, recombinant TRAILs have been tested in many clinical trials (Von Karstedt et al., 2017; Lemke et al., 2014). However, results from the clinical trials were not encouraging, due to the short half-life of recombinant TRAIL and the inherent resistance against TRAIL and TRAIL-induced signalling cascade (Dubuisson and Micheau, 2017; Quintavalle and Condorelli, 2012). Additionally,

TRAIL binding to the decoy receptors DcR1 and DcR2 depletes the reservoir of TRAIL for binding to DR5 (Ashkenazi, 2002; O'Leary et al., 2016).

Agonistic mAbs against DR5 were developed as a response to direct binding away from the decoy receptors for activation of only the relevant receptor DR5 (Dubuisson and Micheau, 2017). However, the bivalent nature of antibodies was insufficient to effectively cluster DR5 into a trimer for activation without Fc crosslinking (Natoni et al., 2007; Watanabe et al., 1998). Antibody mimetics, including single-chain variable fragments (scFv) and nanobodies, against DR5 were then developed to allow multivalency for effective clustering of DR5 upon binding (Dubuisson and Micheau, 2017). One of these constructs is the nanobody 4E6 which was shown to be potent in killing COLO 205, MIA PaCa-2 and BxPC-3 cells when arranged in a linear repeat of 5 subunits compared to fewer subunit repeats (Huet et al., 2014). Using SpyCatcher002-oligomers, the effect of 4E6 αDR5 on DR5 activation for cancer cell apoptosis can be studied with regards to a modular stepwise increase in the number of αDR5 subunits. The circular and planar geometry imparted by the coiled coils may also assist in efficient clustering of DR5 rather than a linear arrangement, as seen from the structure of the subunits once bound with TRAIL (Ashkenazi, 2008).

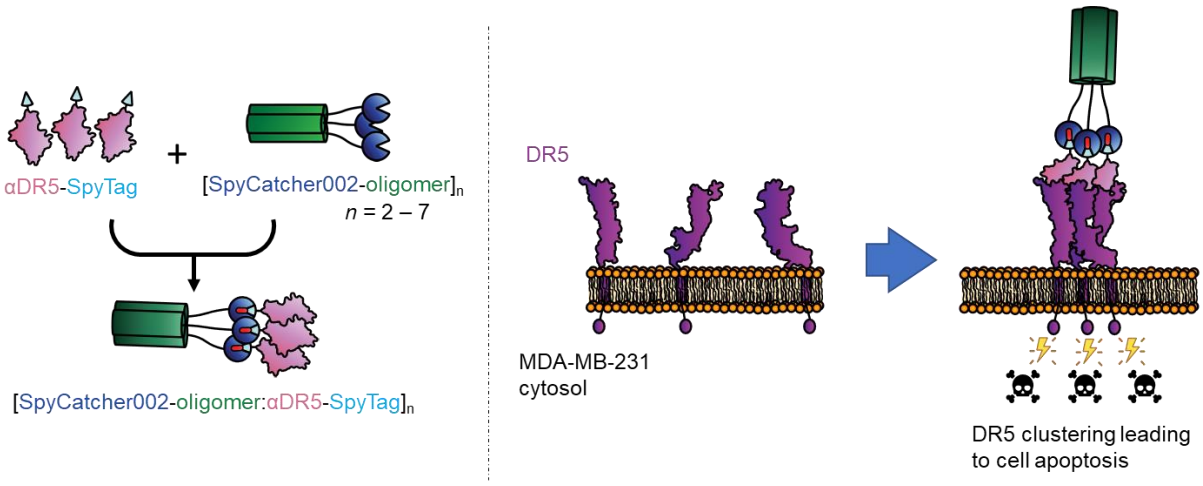


Figure 5.4 Schematic of multivalent DR5 activation using valency-defined SpyCatcher002-oligomers. Purified αDR5-SpyTag and SpyCatcher002-oligomers are reacted to display αDR5 multivalently from dimer to heptamer in a defined assembly. Modular assembly of multivalent αDR5 minimises expression and purification problems. The multivalent αDR5 can then act on cancer cells to cluster DR5 receptors, leading to apoptosis. DR5: Death receptor 5. αDR5: agonistic nanobody against DR5.

5.3.1 Reaction with αDR5-SpyTag and excess reactant recapture

4E6 αDR5 fused to SpyTag (αDR5-SpyTag) was purified by Spy&Go at >95% purity with minimal background contaminants (Figure 3.7). The nanobody was then reacted with SpyCatcher002-oDi to SpyCatcher002-oHept overnight at 25 °C at a 1:3 molar ratio of subunit concentration of SpyCatcher002-oligomers to αDR5-SpyTag to drive complete decoration of SpyCatcher002-oligomers with the nanobody. The reaction was complete for all SpyCatcher002-oligomers as determined by SDS-PAGE, with excess αDR5-SpyTag present in the reaction mixture. Commonly, the removal of excess reactants proceeds by either size exclusion chromatography or dialysis. These methods are time-consuming and not high throughput. Here, Spy&Go can offer an alternative route for the removal of excess SpyTag-protein reactants, demonstrating its utility beyond just primary purification.

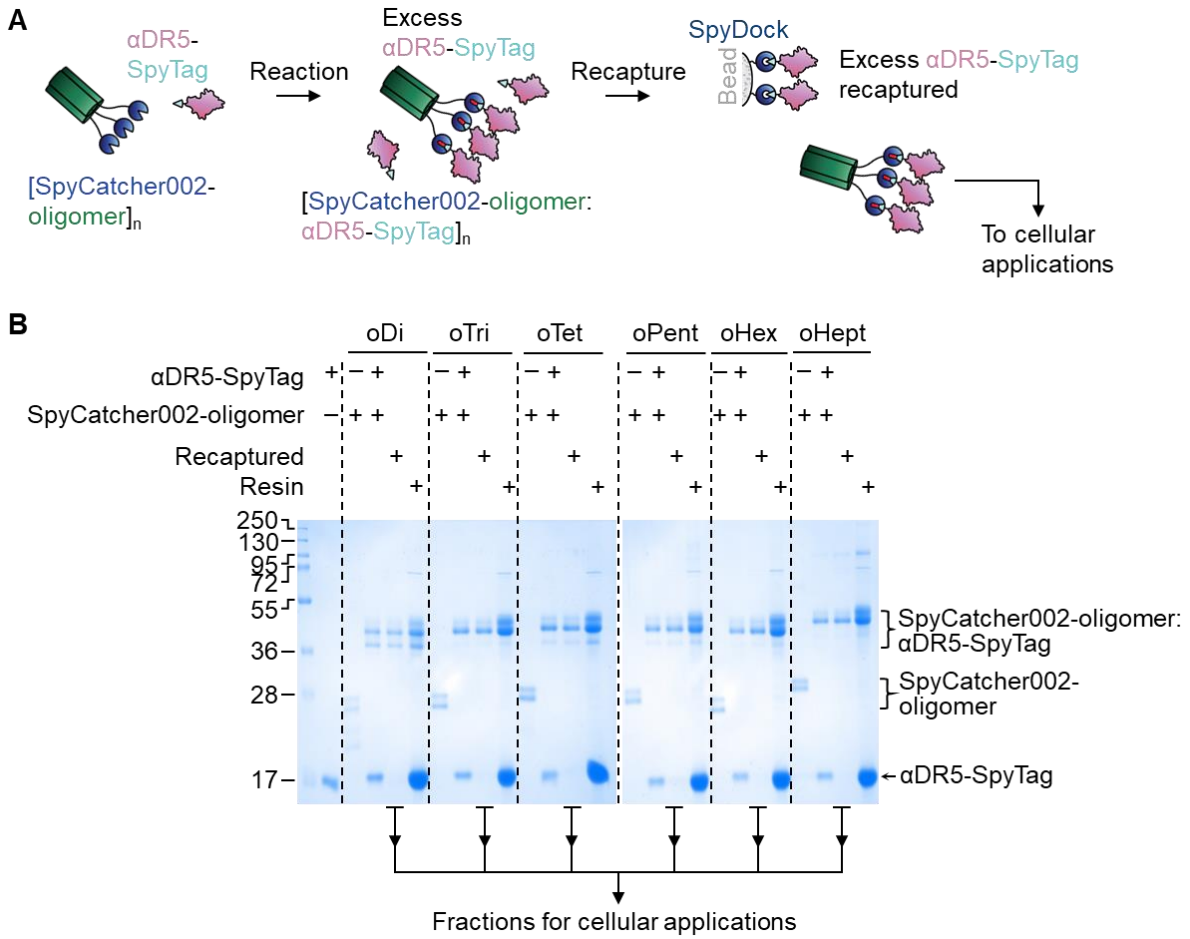


Figure 5.5 Removal of excess αDR5-SpyTag by recapturing using Spy&Go resin. (A) αDR5-SpyTag was present in excess to allow complete conjugation with SpyCatcher002-oligomers. Excess unreacted αDR5-SpyTag can bind to SpyDock present on the Spy&Go resin, leaving the conjugated complex free in the flow-through. (B) Coomassie-stained SDS-PAGE showing the conjugation of αDR5-SpyTag with the SpyCatcher002-oligomer panel. 100 μM of SpyCatcher002-oligomer was reacted with 300 μM of αDR5-SpyTag at room temperature overnight in 50 mM Tris-HCl, 150 mM NaCl at pH 8.0. The Recaptured lanes show the flow-through after incubating the reacted mixture with Spy&Go resin; samples from this fraction proceed to biophysical characterisation and cellular applications. The Resin lanes show boiled resin and the presence of recaptured αDR5-SpyTag and non-specifically-bound SpyCatcher002-oligomer:αDR5-SpyTag complex.

The irreversible reaction between SpyTag and SpyCatcher results in the reconstitution of the parent CnaB2 domain of the *S. pyogenes* FbaB protein, which reduces the availability of the conjugated SpyTag for binding to other SpyCatcher or SpyDock molecules. Spy&Go resin can then be used to remove excess αDR5-SpyTag from the reaction mixture to leave only the reconstituted complex in the flow-through (Figure 5.5 A). After the conjugation

reaction, the mixture was incubated with Spy&Go resin for 2× 2 h at room temperature and the flow-through containing the reconstituted complex was recovered by centrifugation through a mini-spin column. Analysis using SDS-PAGE showed no apparent residual αDR5-SpyTag in the flow-through fraction containing the SpyCatcher002-oligomer:αDR5-SpyTag complex. However, the yields of the reconstituted complex were diminished because some were adsorbed to the Spy&Go resin, possibly via non-specific binding to SpyDock (Figure 5.5 B).

5.3.2 Characterisation of conjugated SpyCatcher002-oligomer complexes

The decoration of SpyCatcher002-oligomers with αDR5-SpyTag increased the overall molecular weight of the reconstituted conjugated complex according to the oligomeric number of the Spy-scaffold. To confirm the increase in the monomeric subunit molecular weight, the SpyCatcher002-oligomers before and after conjugation with αDR5-SpyTag were subjected to intact protein mass spectrometry, performed by Dr Anthony Tumber from the Schofield Group in the University of Oxford. The RapidFire 365 mass spectrometer was calibrated with 10 calibrants of masses ranging from 118.086255 to 2721.894828 g/mol. The calibration error as corrected residuals had a mean of 0.00 ppm with a standard deviation of 0.19 ppm. Corrected residuals ranged from -0.26 to 0.37 ppm (-0.01, 0.06, -0.19, 0.25, -0.05, -0.11, -0.13, 0.37, -0.26, 0.06).

The expected protein masses were calculated from ProtParam (Gasteiger et al., 2005) using the primary sequence of the oligomeric subunit of each SpyCatcher002-oligomer. For disulfide-containing coiled coils (pentameric COMP and heptameric IMX313), the samples were reduced with TCEP before acidification with formic acid. Before SpyTag conjugation, the masses of each SpyCatcher002-oligomer subunit matched the expected masses (Figure 5.6 A). Upon conjugation with αDR5-SpyTag, the subunit molecular weight of each SpyCatcher002-oligomer increased by 16,058 Da from the molecular weight of αDR5-SpyTag with a water molecule removed (16,074-18), from the amidation of Lys and Asp side chains of

SpyCatcher and SpyTag to form the isopeptide bond (Figure 5.6 B). No peaks corresponding to each SpyCatcher002-oligomer subunit were observed, supporting complete decoration of the SpyCatcher002-oligomers with αDR5-SpyTag. For some spectra, a peak corresponding to the size of αDR5-SpyTag was observed, indicating that the removal of αDR5-SpyTag by recapture using Spy&Go resin was not fully effective (Figure 5.6 B). Intact protein mass spectrometry is a qualitative measurement, hence although traces of αDR5-SpyTag were detected in the repurified SpyCatcher002-oligomer:αDR5-SpyTag conjugate samples, αDR5-SpyTag concentration is low as it was undetected by Coomassie staining. Additionally, peaks corresponding to proteolysed SpyCatcher002:SpyTag without the coiled coils and αDR5 were observed in most spectra. SpyCatcher002 of mass 12,486 Da reacted with SpyTag of mass 1,617 Da to form a complex with a mass of 14,085 Da minus a water molecule.

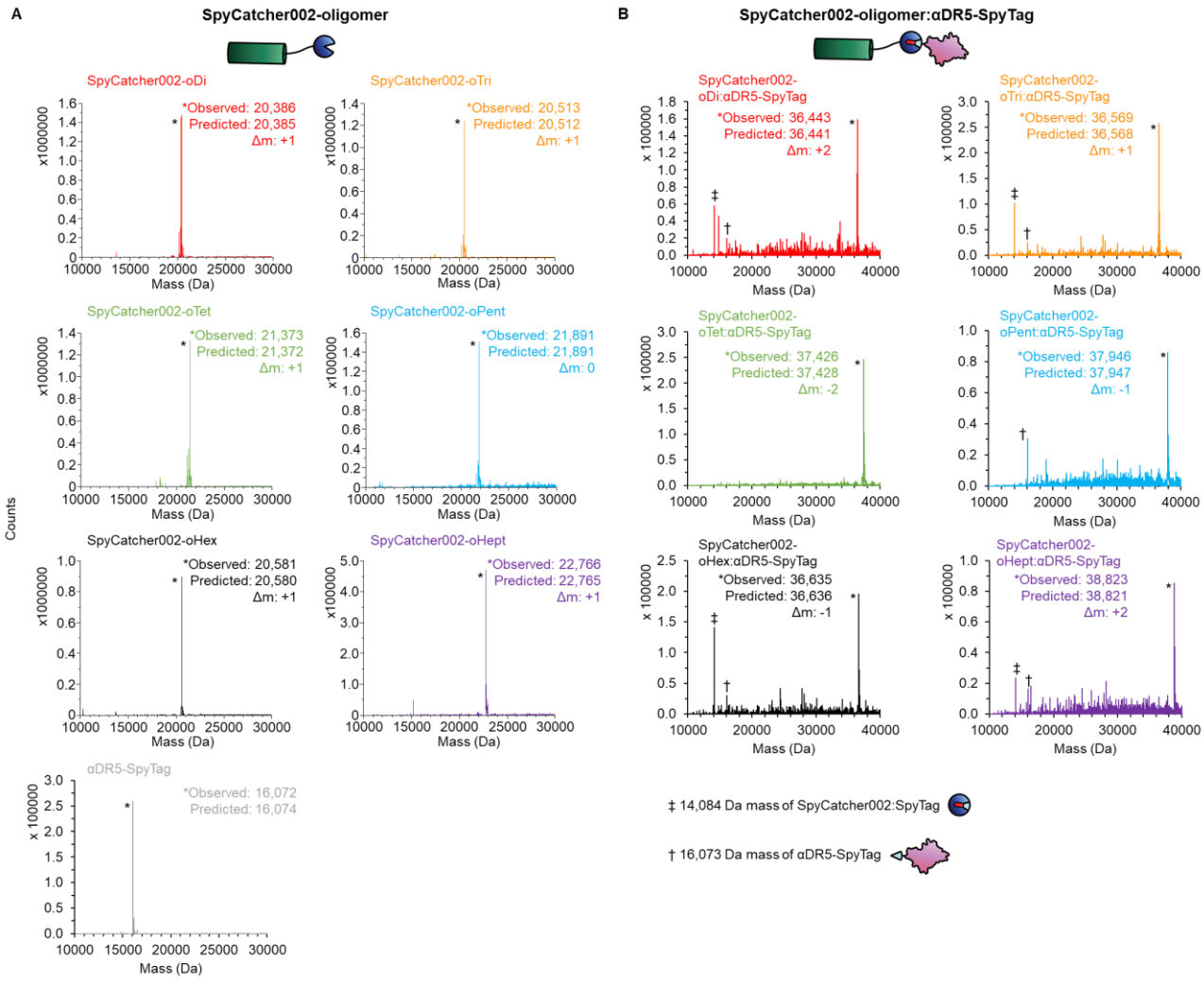


Figure 5.6 Intact protein mass spectrometry analysis of SpyCatcher002-oligomer subunit sizes. (A) SpyCatcher002-oligomer subunit mass before conjugation with αDR5-SpyTag. The highest peak observed was compared with the predicted mass. (B) SpyCatcher002-oligomer:αDR5-SpyTag subunit mass after conjugation with αDR5-SpyTag. The highest peak observed (*) was compared with the predicted mass, with mass error (Δm) shown. Additional relevant peaks marked with ‡ and †. ‡ is the mass of cleaved SpyCatcher002:SpyTag, † is the mass of excess unreacted αDR5-SpyTag.

With the oligomeric subunit masses confirmed through intact protein mass spectrometry, the assembly sizes for each SpyCatcher002-oligomer conjugate were investigated by SEC-MALS, to verify retained oligomer formation after αDR5-SpyTag conjugation. The observed molar masses for each main peak deviated less than 10% of the expected assembly sizes, except for the trimeric SpyCatcher002-oTri:αDR5-SpyTag with an observed size 12.5% smaller than the expected complex size (Figure 5.7 and Table 5.3). Higher-order oligomers, possibly due to aggregation, were present for all conjugated

SpyCatcher002-oligomer complexes, as seen from the heterogeneous peaks at the shoulder of the main peaks but were present in minor populations. SEC-MALS verified that the conjugated SpyCatcher002-oligomer:αDR5-SpyTag complexes formed the assembly close to the expected oligomeric states, albeit with increased formation of higher-order aggregates.

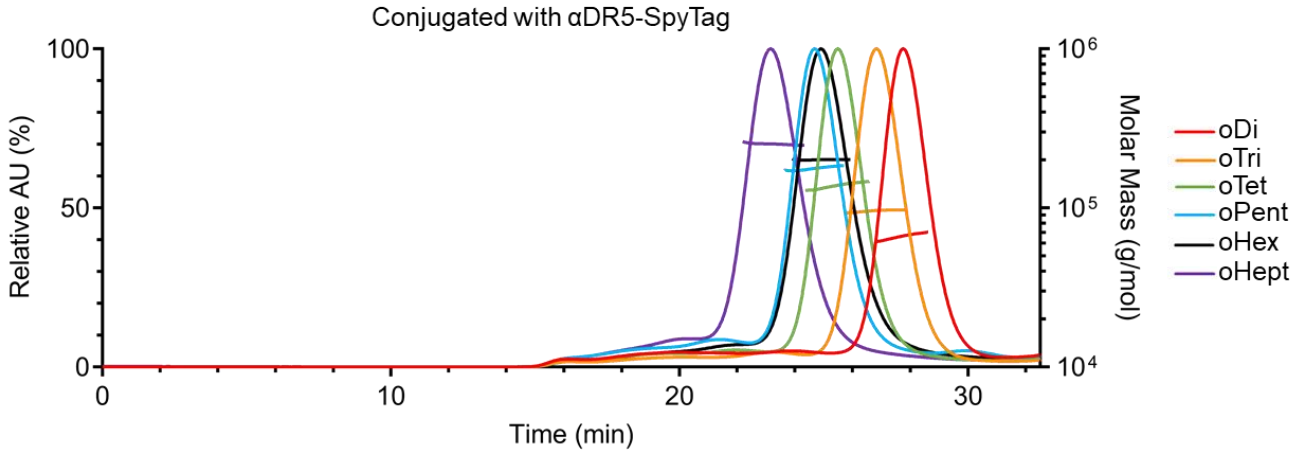


Figure 5.7 SpyCatcher002-oligomer:αDR5-SpyTag assembly sizes after conjugation determined by SEC-MALS. SEC-MALS profiles of different conjugated SpyCatcher002-oligomer:αDR5-SpyTag, with the peaks showing the normalised absorbance units (AU) at 280 nm of each conjugated complex species from SEC. The horizontal lines show the distribution of molar mass of the species within the peak. The observed molecular weights of these species are in Table 5.3. Void volume is at approximately minute 15.

Table 5.3 Size determination of conjugated SpyCatcher002-oligomer:αDR5-SpyTag from SEC-MALS. The observed sizes were determined from multi-angle light scattering (MALS). Error bars represent the uncertainty in fit to the molar mass curve. % error is 100 × (ΔM_r/ Expected Mr).

Protein	Observed Mr (kDa)	Expected Mr (kDa)	ΔM_r (kDa)	% error
oDi:αDR5-SpyTag	66.3 ± 0.5	72.9	-6.6	9.1
oTri:αDR5-SpyTag	96.0 ± 0.2	109.7	-13.7	12.5
oTet:αDR5-SpyTag	138.4 ± 0.3	149.7	-11.3	7.5
oPent:αDR5-SpyTag	177.2 ± 0.3	189.7	-12.5	6.6
oHex:αDR5-SpyTag	201.1 ± 0.2	219.8	-18.7	8.5
oHept:αDR5-SpyTag	252.2 ± 0.1	271.7	-19.5	7.2

Subsequently, the hydrodynamic radius of each SpyCatcher002-oligomer before and after conjugation was investigated using DLS to affirm particle decoration. Before conjugation, the SpyCatcher002-oligomers show homogeneity and increasing particle hydrodynamic size

from dimer to heptamer (Figure 5.8 A). Post-conjugation, the αDR5 conjugates increased in size, indicating covalent linkage formation between SpyCatcher and SpyTag (Figure 5.8 B). Most of the oligomer conjugates maintained good homogeneity post-conjugation but the dimer, trimer and pentamer conjugates formed aggregates at the 100-1000 nm scale, albeit at a smaller proportion. Aggregates were present in 3.1% of αDR5 conjugated to oTet, 0.3% in oHex and 2.5% in oHept. In contrast, αDR5 conjugated to oDi had 22.4% of aggregates, oTri had 6.3% and oPent had 7.4% as measured by area under the curve for the discrete aggregate peaks using OmniSIZE 3.0. These biochemical and biophysical characterisations demonstrate the ease of achieving defined multivalency of a protein ligand and confirm the complete decoration of SpyCatcher002-oligomers with αDR5-SpyTag, the removal of excess αDR5-SpyTag from the reaction mixture and the stability of conjugated SpyCatcher002-oligomers, albeit with increased aggregate formation.

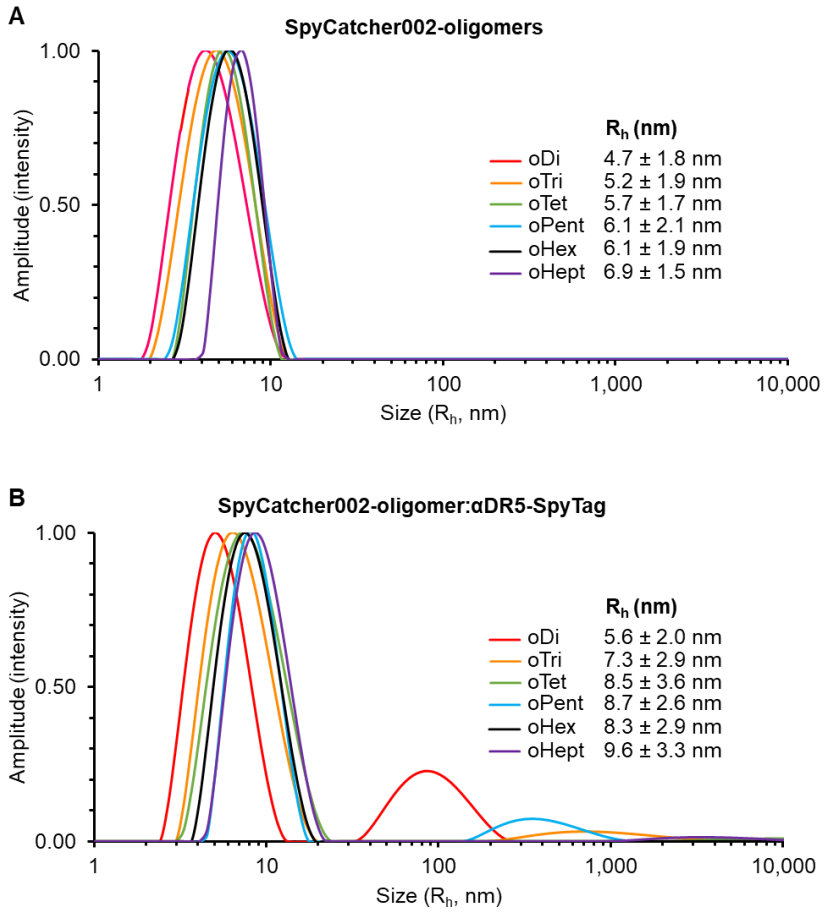


Figure 5.8 Hydrodynamic radius distribution of SpyCatcher002-oligomers determined by DLS. (A) Unconjugated SpyCatcher002-oligomer assembly showing increasing hydrodynamic sizes from oDi to oHept. (B) Conjugated SpyCatcher002-oligomer:αDR5-SpyTag assembly showing increasing hydrodynamic sizes from oDi to oHept. Aggregates at 100-1000 nm range were present in samples of oDi, oTri and oPent. Samples were run at 1 mg/mL at 20 °C. R_h value is mean ± 1 s.d, *n* = 10 scans per run.

5.4 Proof of concept for stoichiometry-based assay

The panel of multivalent αDR5 assembled using the Spy oligomerisation toolbox was purified by Spy&Go and then assessed for stoichiometry-dependent potency against a TRAIL-sensitive human breast cancer cell line, MDA-MB-231 (Zhang et al., 2009). This work was done by Dr Anusuya Banerjee from the Howarth lab.

SpyCatcher002-oligomer:αDR5-SpyTag was added to the cells at different monomeric αDR5 concentrations ranging from 0.1 pM to 100 nM for 24 h at 37 °C with 5% CO_2 to induce DR5-dependent apoptosis. Cell viability, or lack thereof, was tested using resazurin assay to

monitor cellular mitochondrial metabolism for live cells. Monomeric unconjugated αDR5-SpyTag together with dimeric and trimeric conjugates failed to show any cell killing within this range, whereas the tetrameric conjugate only showed modest cell killing above 1 nM of αDR5. In contrast, higher-order oligomers, namely pentameric, hexameric and heptameric αDR5 conjugates, showed a clear stoichiometry-dependent dose-response, decreasing MDA-MB-231 cell viability across the same αDR5 concentration range with potency at the sub-nanomolar concentration (Figure 5.9 A). The half maximal effective concentrations (EC_{50}) for these higher-order αDR5 conjugates were 0.9 nM for the pentameric complex, 0.42 nM for the hexameric complex and 0.31 nM for the heptameric complex, demonstrating that higher valencies of αDR5 induce greater potency of DR5-dependent apoptosis on cancer cells despite the trimeric nature of activated DR5 clustering (Ashkenazi, 2008). As a control, unconjugated SpyCatcher002-oligomers did not show any cell killing from 0.1 pM to 1000 nM. Thus, the cell killing was induced by the multivalent arrangement of αDR5 via SpyCatcher002 fusion.

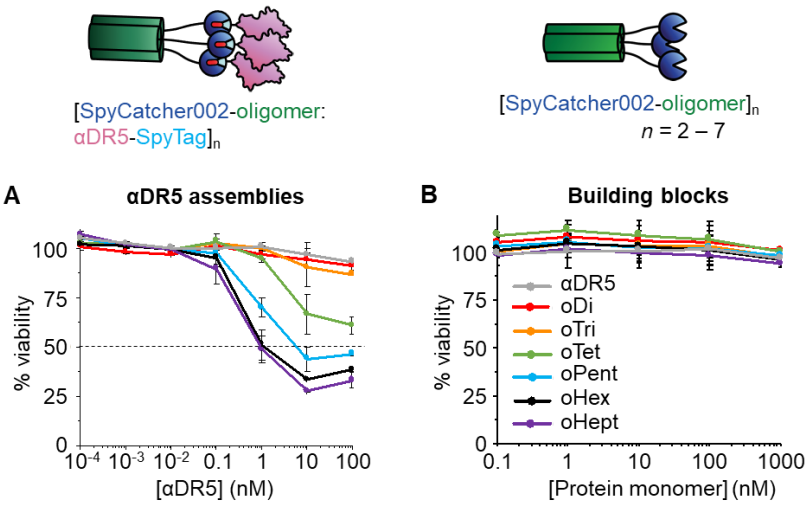


Figure 5.9 Effect of multivalent αDR5 on cancer cell apoptosis. (A) A dose-response curve of MDA-MB-231 cells treated with αDR5 conjugated to SpyCatcher002-oligomers from Figure 5.5 for 24 h before visualising cell viability using resazurin assay. The line at 50% viability is the cutoff for EC_{50} calculation. Error bars represent mean ± 1 s.d, *n* = 2. (B) MDA-MB-231 cells were also treated with the unconjugated SpyCatcher002-oligomers as in (A). Error bars represent mean ± 1 s.d, *n* = 3.

5.5 Discussion

Many cellular processes require receptor oligomerisation for downstream signal transduction upon ligand or stimulus exposure (Hartman and Groves, 2011; Ng et al., 2013). Understanding multivalency is key to exploit protein-protein interactions to control cellular signalling. In particular, the rise of superselectivity whereby a small increase in the number interacting ligands and receptors is able to confer an amplified non-linear signal modulation demonstrates the potential of controlled multivalent ligand assembly (Curk et al., 2017a; Dubacheva et al., 2014). However, current techniques to manipulate nanoparticles for variable ligand valency are either heterogeneous or limited to chemical entities (Dubacheva et al., 2015; Kiessling et al., 2006; Puffer et al., 2007; Satav et al., 2015). Protein-based multivalent ligands can enable easier exploration into protein ligand-receptor interaction.

Without accounting for protein multimerisation on polyhedral protein scaffolds (e.g. viral capsids) due to their big size and multi-planar arrangement, recombinant protein ligands can be multimerised by cloning repeating numbers of the subunit into a monocistronic plasmid coding sequence separated by linkers. However this approach is limited by plasmid size (Silva et al., 2012; Yang and Yang, 2012) and the linear arrangement of the protein construct which may not efficiently mimic the clustering order of receptors. Different protein-based multivalent scaffolds with a multimerising domain have been explored to which proteins of interest can be fused (Blanco-Toribio et al., 2013; Kim and Jung, 2016; Kim et al., 2015; Ogun et al., 2008; Zhu et al., 2010), but they require individual cloning, expression and purification optimisations. Problems can occur such as having heterogeneous populations of 5-7 oligomers instead of a single heptameric species (He et al., 2018), failed assembly of a trimer fusion (Voulgaraki et al., 2005), and degradation of dimer fusions in *E. coli* cytosol (Arndt et al., 2001).

Using protein tags can alleviate such complications to enable a modular assembly of tagged target proteins to a tag-receiving multimerising protein scaffold. Prior examples include a heptameric scaffold using dual Tag/Catcher systems (Brune et al., 2017) and a rhizavidin engineered for monodisperse dimeric, tetrameric, hexameric, and octameric display of

biotinylated ligands (Yoon et al., 2018). However, these methods were limited by the number of the available multivalency. As such, I have developed a nanoassembly toolbox to modularly oligomerise SpyTag-fused proteins in a stepwise manner from dimer to heptamer using characterised coiled coils fused to SpyCatcher002.

The panel of SpyCatcher002-oligomers were made from a mix of naturally occurring and computationally designed coiled coils based on the stability of the fusion protein to maintain oligomeric assembly. Monomeric and oligomeric sizes were verified using intact protein mass spectrometry and SEC-MALS while homogeneity was validated using DLS. SpyCatcher002-oDi, -oTri, -oTet, and -oHex are held together using non-covalent knobs-into-holes interactions at the coiled coil hydrophobic interface, whereas SpyCatcher002-oPent and -oHept have additional disulfide bonds to confer stable association in an oxidising environment. Although hydrophobic interactions can be strong, non-covalently linked SpyCatcher002-oligomers may begin to dissociate at low nanomolar or picomolar concentrations, which can be a limitation in the practical application of these oligomers. Other groups have used non-covalently bonded coiled coils for cellular work at low concentrations (Blanco-Toribio et al., 2013; Cho et al., 2018; Fan et al., 2008). Nonetheless the stability of SpyCatcher002-oligomers at low concentrations needs to be verified by measuring the concentration at which half the complex is assembled (Fletcher et al., 2012).

To demonstrate the utility of the SpyCatcher002-oligomers in understanding multivalent cellular signalling, the clustering of DR5 upon binding with αDR5 for subsequent activation of caspase-8-dependent apoptosis was investigated using dimeric to heptameric αDR5-SpyTag. Interest in this area is high due to the nature of selective toxicity of DR5-targeting to cancer cells. Studies show that higher-order clustering of DR5 above the trimeric DR5 cluster is needed for effective DR5 signalling (Dubuisson and Micheau, 2017; Naval et al., 2019). Current methods to explore the higher-order clustering of DR5 relied on linking αDR5 in a linear arrangement of 4 (Papadopoulos et al., 2015), 5 (Huet et al., 2014), 6 or 8 (Swers et al., 2013) subunits, which is not ideal for high throughput experimentation and not reflective of

radial receptor cluster geometry (Ashkenazi, 2008). αDR5 fused to a COMP coiled coil "combody" was also previously explored for multivalency and circular geometry (Wang et al., 2013), but direct fusion to the pentamerising domain limits modularity for other valencies and target proteins. By conjugating αDR5-SpyTag to SpyCatcher002-oligomers, controlled valency and geometry comparisons can be undertaken with modular convenience to investigate higher-order DR5 or other receptor clustering.

The small size of the αDR5 nanobody can enable its complete conjugation with SpyCatcher002-oligomers because of minimal steric clashes. Antibody mimetics have the advantage of small size while retaining high binding affinity to target molecules, thus the rise in their popularity for diagnostics and therapies (Yu et al., 2017). Larger-sized natural ligands may still be attractive for multivalent study but their conjugation can result in incomplete scaffold decoration (Bruun et al., 2018). Upon conjugation with SpyCatcher002-oligomers, the molar mass and hydrodynamic radius of the assembly increased. Higher-order oligomers of assemblies or aggregates were observed in the characterisation of conjugated SpyCatcher002-oligomer:αDR5-SpyTag from SEC-MALS and DLS, which may be caused by forced oligomerisation of monomeric αDR5 leading to aggregation (Kim and Jung, 2016; Kim et al., 2015). Nonetheless, the proportion of aggregates were minor and further engineering to supercharge the SpyCatcher002-oligomers may alleviate this problem (Lawrence et al., 2007).

Oligomerisation of αDR5 showed that low oligomeric states (n = 2 and 3) were not sufficient to induce apoptosis in MDA-MB-231 cells, which may be counterintuitive considering the trimeric nature of clustered DR5 upon TRAIL binding. A small increase in valency of 2 units to n = 5 managed to activate the signalling pathway effectively to render apoptosis with a sub-nanomolar EC_{50} concentration, suggestive of superselectivity in a protein-protein interaction context (Dubacheva et al., 2014). Further increase in valency to n = 6 and 7 lowered the EC_{50} further, but to a lesser degree than from n = 3 to n = 5. This difference in efficacy can be attributed to crossing a threshold of the number of ligand-receptor interactions necessary

for effective apoptosis. Higher valencies may induce super-clustering of several clustered DR5 receptors, analogous to Fc receptor clustering on immune cells (Naval et al., 2019; Raghavan and Bjorkman, 1996). By identifying the EC_{50} of each oligomeric state, it may be possible to balance cost and efficacy. This proof of concept for modular multivalent protein assembly in DR5 activation shows promise to extend into more diverse cellular signalling processes.

6 Summary and general discussion

The SpyTag/SpyCatcher technology provides a rapid and easy way to link two proteins post-translationally. The usage of the technology has consistently increased since its introduction owing to its general applicability and potential to assist studies in many scientific areas. Many techniques and platforms were created to add on to the SpyTag/SpyCatcher technology, such as increasing its rate of reaction (Keeble et al., 2017, 2019) or expanding the Tag/Catcher toolbox (Liu et al., 2017; Pröschel et al., 2017; Tan et al., 2016; Veggiani et al., 2016; Young et al., 2017; Zakeri and Howarth, 2010). Here, I sought to expand the utility of the technology in the area of protein purification and scaffolds for multivalent nanoassembly.

Protein purification is a key process for the characterisation and understanding of protein structure, function and interaction. Affinity chromatography remains the leading technique for protein purification due to its simplicity and general applicability for a diverse set of proteins without the need for prior knowledge of solubility, size, or charge (Gräslund et al., 2008; Kimple et al., 2013; Vassylyeva et al., 2017). In the preparation for a SpyTag/SpyCatcher bioconjugation reaction, different tags were used for purification of SpyTag-proteins (Keeble and Howarth, 2019), including polyhistidine tag (Brune et al., 2016; Keeble et al., 2017; Stranges et al., 2016), C-tag (Brune et al., 2016; Bruun et al., 2018), FLAG-tag (Pardee et al., 2016), and Strep-tag II (Stein et al., 2017). However, these affinity tags did not provide utility post-purification, which can complicate vector design or require tag cleavage for certain application settings (Brune and Howarth, 2018; Jin et al., 2017). There was a need for a simpler purification approach for SpyTag-proteins.

In this thesis, I have developed a platform to purify SpyTag-proteins by utilising SpyTag itself as the affinity tag, bypassing the need for any extraneous affinity tags to save material cost and yield, whilst simplifying plasmid vector design and size. The purified SpyTag-protein can then be modularly conjugated to assemblies with defined multivalency using the

characterised Spy oligomerisation toolbox consisting of SpyCatcher002-oligomers from a dimeric to heptameric state.

In Chapter 3, I developed Spy&Go purification of SpyTag-proteins. I showed good purity isolation of bacterial- and mammalian-expressed proteins with a high degree of freedom for SpyTag localisation in the protein construct. SpyDock can form a non-covalent association with high binding affinity to different SpyTag generations. The minimal number of mutations needed to develop SpyDock demonstrates the versatility of the system for effective reversible binding. The high binding affinity allows SpyDock to sequester SpyTag-proteins efficiently from cell lysate and overcome stringent wash cycles of 500 mM imidazole to yield proteins with good purity. The dissociation constants determined by ITC between SpyDock and SpyTag- or SpyTag002-MBP were modest, being 1-2 magnitudes above the low nanomolar region of high affinity antibody-antigen binding. Limiting conformational stability through computational protein design (Magliery, 2015) or directed evolution (Lane and Seelig, 2014) can confer higher fidelity and hence higher affinity binding between SpyTag generations and SpyDock (Manivel et al., 2000; Reverberi and Reverberi, 2007). Higher binding affinities can enable more thorough washing without compromising SpyTag-protein sequestration to yield purer products. However, harsher elution conditions such as acidic pH may be needed to destabilise the interaction, which may not be suitable for many proteins and applications.

The elution condition is the main disadvantage of Spy&Go, because many proteins may not be able to tolerate such a high solute concentration which can drive structural unfolding. I have shown the purification of 2 classes of functional proteins: binding proteins and a fluorescent protein. αDR5 and mClover3 retained protein function even after the exposure of 2.5 M imidazole during elution. However, one can argue that these proteins have high fold stability (Bajar et al., 2016; Muyldermans, 2013). Additionally, investigation into the effect of 2.5 M imidazole exposure on the catalytic efficiency of enzymes is yet to be explored. Metalloproteins may also be sensitive to imidazole since imidazole can compete for metal ion binding (Huang et al., 2010; Liao et al., 2013). More purifications of different proteins are

needed to examine purity and functionality after Spy&Go purification. Nonetheless, the alternative elution conditions (pH 2.5 or 1 mM SpyTag003 peptide) available are already widely employed in current affinity chromatography methods (Kimple et al., 2013) and can be used for imidazole-sensitive proteins.

Currently, Spy&Go purification was shown to be compatible to 2% Triton X-100, 2% Tween 20 and 10 mM β-mercaptoethanol, which can facilitate the purification of certain membrane proteins and many sulfhydryl-containing proteins. However, more characterisations into the effect of different chemical compounds on the binding between SpyTag and SpyDock are needed to cover a wider range of proteins, such as the upper limit of different salt concentrations, chelators, ionic and zwitterionic detergents, reductants, alcohols and polyols. The binding between SpyTag and SpyDock was severely compromised in denaturants, limiting the use of Spy&Go in purifying proteins from inclusion bodies. Replacing SpyTag with SpyTag003 may enable better binding tolerance in higher urea concentrations due to the higher binding affinity.

The capability to regenerate Spy&Go resin is economical for repeated usage of the same resin batch. The number of regeneration cycles before severe loss in binding capacity needs to be investigated, especially considering that the 0.1 M NaOH alkaline wash showed some hydrolysis of SpyDock from the solid support. Alkaline stability can be engineered into a protein by substituting Asn to other amino acids without a carboxamide group in the side chain (Gülich et al., 2000, 2003). SpyDock contains two Asn at positions 95 and 103. The latter position is already mutated to N103D in SpyCatcher003, resulting in an increase in reaction rate (Keeble et al., 2019). Further incorporating a non-destabilising substitution at N95 may confer increased alkaline stability to SpyDock for a better regeneration capability.

Another characterisation that would be interesting to measure is the binding capacity of the Spy&Go resin. Binding capacity is economically-relevant to determine the amount of protein that can be purified using a specific volume of chromatography resin. Binding capacity can be classed to static or dynamic binding capacities. Static binding capacity (SBC) reports

the maximum amount of proteins bound to a resin in batch mode at a particular condition, whereas dynamic binding capacity (DBC) reports the amount of proteins bound under operating conditions with a flow-rate until a significant breakthrough of unbound protein occurs. Determination of the more relevant parameter, DBC, requires resin packing into a column and the constant monitoring of protein concentration through absorbance at 280 nm using a HPLC instrument. Once binding capacities are determined, subsequent work on the nature of SpyDock and the solid support can be engineered to increase binding capacities for more efficient purification, such as having repeated number of SpyDock (Müller and Vajda, 2016) or increasing the resin bead particle size and porosity (Pabst et al., 2018).

Due to the unfavourable Spy&Go elution conditions, a different platform for SpyTag-protein purification was explored. Proteins have been engineered to undergo an inducible conformational change when exposed to different stimuli, such as temperature (Chakshusmathi et al., 2004; Megeed et al., 2006), small molecules (Dagliyan et al., 2013; Guntas et al., 2004), metals (Baird et al., 1999; Fanning et al., 2011; Stratton et al., 2008), or light (Lee et al., 2008; Wang et al., 2016a). These conformational changes lower the binding affinity between the protein and its ligand. Engineering sensitivity towards these stimuli can be appealing, but scalability and cost issues may restrict the utility and hence adoptability of the platform. pH-inducible conformational change offers a solution to these issues and has been explored as an approach to diminish protein-protein binding (Boyken et al., 2019; Fanning et al., 2011; Könning et al., 2018; Schröter et al., 2015; Tsukamoto et al., 2014; Watanabe et al., 2013). Thus, I engineered a pH-switchable non-reactive SpyCatcher called SpySwitch in Chapter 4 to allow mild elution of SpyTag-proteins.

Following on previous pH-switchable proteins, I substituted histidines at the interface of SpyTag and SpyCatcher whereby a repulsive interaction can be achieved at pH < pK_a of histidine to destabilise the binding. A clear pH-dependent loss in binding between SpyTag and non-reactive SpyCatcher histidine mutants was observed from pH 8.0 to pH 4.0. Among the 16 single histidine mutants, a majority of 10 mutants showed only minor elution of SpyTag003-

sfGFP at pH 4.0, demonstrating the malleability of SpyCatcher003 S49C E77A to point mutations. Histidine mutants successful in conferring pH sensitivity were then combined to form double histidine mutants and subsequently triple histidine mutants, since having multiple histidines can alter their pK_a at the interface microenvironment cooperatively for increased pH sensitivity (Boyken et al., 2019; Murtaugh et al., 2011). The finalised triple histidine mutant (S30H + Y84H + E85H), now termed SpySwitch, has significantly diminished SpyTag003 binding at pH 5.0, capable of eluting 13-fold more SpyTag003-sfGFP than the best single histidine mutant at this pH (Y84H), and 3-fold more than the best double histidine mutant (S30H + E85H). However, substituting histidines at key interface points with SpyTag was expected to destabilise initial binding at pH 8.0. Accordingly, SpySwitch had 3-fold more unbound SpyTag003-sfGFP in the flow-through than SpyDock, compared to 1-fold and 1.4-fold for the single Y84H and double S30H + E85H mutants. To reduce this compromised binding at pH 8.0, the binding between SpyTag003 and SpySwitch can be increased allosterically or at other pH-insensitive interface residues through phage display (Keeble et al., 2017).

To confirm the cooperativity effect of multiple histidines in SpySwitch over its single and double histidine mutants, it is important to determine the dissociation constant, K_d, of the different constructs over the relevant pH range. Previously, the K_d of any non-covalent interaction between SpyTag and SpyCatcher was determined by ITC (Khairil Anuar et al., 2019; Zakeri et al., 2012), but this method requires dialysing the samples into a running buffer to minimise environmental heat changes. The pIs of SpyCatcher003 and its variants (including SpySwitch) range from 4.5-4.8, thus dialysing the constructs into pH 5.0 for ITC experiments may result in pI-induced aggregation. Using surface plasmon resonance (SPR) to determine K_d is more advantageous because the chip-immobilised SpySwitch is only introduced to low pH during the run, thereby minimising pI-induced aggregation. Furthermore, SPR allows the measurement of protein kinetics such as k_{on} and k_{off}, which can assist in determining the length of incubation time needed for initial binding and elution.

Thereafter, the mechanism of pH sensitivity in SpySwitch can be investigated. The pH sensitivity gained by SpySwitch can be explained in terms of local or global destabilisation in protein structure. Local destabilisation refers to the switch in histidine electrostatic state upon the lowering of pH, which results in nearby electrostatic repulsion from positively-charged residues at the SpyTag and SpySwitch interface. Global destabilisation refers to the inherent instability of the SpySwitch caused by the introduction of histidines that impairs proper folding and structure, especially at lower pH values. Due to its area of application where repeated usage and long-term usability are needed, SpySwitch is required to be globally stable. SpySwitch secondary structure at pH 8.0 and 5.0 can be monitored by circular dichroism spectroscopy to investigate the percentage change in estimated β-strand fold (Greenfield, 2007; Micsonai et al., 2015). For SpySwitch to be stably-folded, its CD spectra should have limited changes to this parameter over the relevant pH range. Its thermal stability can also be determined by differential scanning calorimetry as was documented for the different SpyCatcher generations (Keeble et al., 2017, 2019; Zakeri et al., 2012). To fully understand the effect of pH on residue interplay, SpySwitch can be crystallised at the different pH values with or without SpyTag003 to analyse the structures for interacting residues. Analysing crystal structures may also help in further refining the pH sensitivity through rational design.

Currently, the pH sensitivity of SpySwitch was only demonstrated on the purification of SpyTag003-sfGFP. It will be interesting to purify other proteins that would be problematic for Spy&Go purification, for example metal-binding proteins or multisubunit protein complexes such as a monocistronic RNA polymerase complex (Vassylyeva et al., 2017) because these proteins may be denatured after exposure to 2.5 M imidazole. Additionally, purifying proteins fused to other SpyTag generations will be useful to have backwards compatibility for different SpyTag/SpyCatcher applications. Indeed, the current most used SpyTag generation is the first generation (Keeble and Howarth, 2019). However, SpyTag has the lowest binding affinity and rate of reaction with SpyCatcher variants. Thus, although elution of SpyTag-proteins will be easier at pH 5.0, the initial binding and washes at pH 8.0 may be more compromised due to

the weaker interactions. Furthermore, testing purifications under different buffer conditions to ensure pH-dependence and buffer-independence can be useful to determine the general applicability of SpySwitch pH sensitivity alongside enabling purification under protein-specific buffer composition.

In addition to purification of SpyTag-proteins, establishing a toolbox for easy plug-and-display multivalent assembly can assist in enhancing protein binding avidity and simplifying the study of multivalent ligand-receptor signalling. In Chapter 5, I characterised a Spy oligomerisation toolbox consisting of SpyCatcher002 fused to the oligomerising domains of different coiled coils with valencies from dimer to heptamer. Direct genetic fusion of target proteins (e.g. nanobodies) to the coiled coils require laborious expression and purification optimisations. Indeed, during the selection and characterisation of different SpyCatcher002-oligomers, SpyCatcher002-CC-Tet, SpyCatcher002-SNAP-23 and SpyCatcher002-CC-Hept were eliminated from consideration due to expression and stability issues. Thus, modularly assembling the target protein to the characterised SpyCatcher002-oligomers can circumvent time-consuming optimisation steps.

Pre-conjugation, SpyCatcher002-oligomers assembled to the correct oligomeric state. Post-conjugation with αDR5-SpyTag, the conjugated complexes were still verified to be of correct oligomeric states, but aggregates were found which may be caused by forced oligomerisation of a monomeric protein (Kim and Jung, 2016; Kim et al., 2015). The aggregation may also be caused by the long overnight incubation at room temperature to drive complete conjugation. A time course assay to determine the time required to achieve complete conjugation can minimise the exposure of the complexes to elevated temperatures. Additionally, stabilising agents such as trehalose, glycerol or arginine may reduce protein aggregation (Hamada et al., 2009).

Among the SpyCatcher002-oligomers, oDi, oTri, oTet and oHex are held together by non-covalent knobs-into-holes interactions, which can dissociate into each coil's subunits at very low concentrations and limit their applications. Apart from increasing the heptad repeat

number, increasing the intramolecular electrostatic interactions can also confer higher stability for coiled coils without increasing the peptide length (Drobnak et al., 2017). However, this method still relies on non-covalent interactions. Some alternatives for dimeric, trimeric and tetrameric coiled coils containing disulfide bridges between the coil subunits are found naturally. The dimeric human centriolar CEP135 protein contains a disulfide bridge within the hydrophobic interior at the *d* position of the coiled coil heptad (Kraatz et al., 2016). A trimeric cartilage matrix protein, homologous to the COMP oligomerising domain for oPent, contains two disulfide bridges per coil at the N-terminal end, separate from the main knobs-into-holes interactions (Dames et al., 1998). A disulfide-containing tetrameric coiled coil from the P3 protein of cauliflower mosaic virus contains three distinct coiled coil motifs broken by proline kinks, with one section of the proline kink also containing the disulfide bridges (Hoh et al., 2010). Up until now, there are no discoveries of a naturally-occurring disulfide-linked hexameric coiled coil. Overcoming this, the Woolfson lab computationally-engineered a *de novo* hexameric coiled coil fully linked with disulfide bridges between all subunits (PDB ID: 4H8L). However, this construct also contained a pH-sensitive L24D mutation facing the hydrophobic core (Zaccai et al., 2011). A disulfide-linked hexamer without the L24D mutation has not been made, thus biophysical characterisations are needed to confirm the covalent nature of cysteines in this construct and the stability of the oligomer. Alternatively, the α-helical peptides in SpyCatcher002-oHex can be stapled by a chemical linker connecting introduced cysteines (Skowron et al., 2019), however this method requires further steps and material which can be disadvantageous.

In section 5.4, the Spy oligomerisation toolbox was used to demonstrate the stoichiometry-dependent effect of cancer cell apoptosis through DR5 activation. A switch-like apoptotic response on MDA-MB-231 cells was observed by a small increment of valency from low (n = 2-3) to high (n = 5-7). Different cancer cell lines have different DR5 expression levels (Ma et al., 2005), thus each cell line may require distinct strengths of anti-DR5 binding affinities to attain a potent EC_{50} value. The modularity achievable through SpyCatcher002-oligomers

can enable a faster screening of different anti-DR5 proteins tagged with SpyTag for the different cancer cell lines. Apart from DR5 signalling, other receptors also require multivalent ligand binding for clustering and activation such as the folate receptor (Zhang *et al.*, 2019), CD40 (Trouche et al., 2007) and the T cell receptor (Fairhead et al., 2014).

To complement the current SpyCatcher002-oligomers, further expansion in oligomeric state can be added to the Spy oligomerisation toolbox. Above heptamer, coiled coils are rarely found in higher oligomeric states (Woolfson, 2017), with exceptions being the membrane-spanning octamer (Dong et al., 2006) and dodecamer (Koronakis et al., 2000), and a decameric DNA translocating bacteriophage tail (Sun et al., 2014). Thus, multimeric non-α-helical proteins need to be used to expand the Spy oligomerisation toolbox. Candidates include an octameric endolysin from the streptococcal C1 virus (McGowan et al., 2012; Nelson et al., 2006), a nonameric small terminase from a *Thermus thermophilus* phage G20C (Loredo-Varela et al., 2013), the decameric DNA translocon from bacteriophage ΦX174 (Sun et al., 2014), an undecameric tryptophan RNA-binding attenuation protein from *Geobacillus stearothermophilus* (Chen et al., 1999), a dodecameric tryptophan RNA-binding attenuation protein from *Bacillus halodurans* (Chen et al., 2011a), and another dodecameric proteasome activator protein PafE from *Mycobacterium tuberculosis* (Bai et al., 2016). Above dodecamers, proteins are mostly membrane-spanning, hence have limited practicality. The octamer to dodecamer candidates consist of subunits with multiple α-helices and/or β-sheets folds, except the decameric coiled coil, thus the radial geometry of each SpyCatcher-octamer to SpyCatcher-dodecamer will cover a bigger area than the current SpyCatcher002-oDi to -oHept. Furthermore, the higher-order oligomer candidates have endogenous nucleotide-binding or protein-binding functions that can perturb certain cellular processes upon contact, thus care must be taken when utilising these constructs. Going above monoplanar protein oligomers, more polyhedral protein assemblies can be fused to SpyCatcher or SpyTag than the currently-existing constructs (Keeble and Howarth, 2019) to display and arrange different protein classes in varying architecture for next generation protein nanotechnology.

Aside from exploring higher-order oligomers and polyhedrons, heteromeric coiled coil oligomers can be exploited for the defined display of multiple proteins within a single assembly complex using orthogonal Tag/Catcher chemistries. Heteromeric dimers (Thomas et al., 2013; Thompson et al., 2012; Woolfson, 2017), trimers (Kiyokawa et al., 2000, 2004; Nautiyal and Alber, 2008; Nautiyal et al., 1995), tetramers (Fairman et al., 1996; Root et al., 2009) and hexamers (Spencer and Hochbaum, 2017; Zaccai et al., 2011) are available whereby a unique coil subunit of the heteromer can be fused to either SpyCatcher (amidation isopeptide bond reaction), SnoopCatcher (transamidation isopeptide bond reaction) or E-Ligase (ester bond reaction) for an orthogonal one-pot assembly of Spy-, Snoop-, or Cpe peptide-tagged proteins to make multi-specific protein oligomers. By having multi-specificity to overexpressed oncogenic surface markers using the heteromeric Tag/Catcher complexes, diagnostic discrimination and sensitivity between cancer cells from healthy cells could potentially be enhanced (Curk et al., 2017b; Reubi and Maecke, 2017).

Additionally, one can foresee further utility of SpyDock or SpySwitch in contexts other than protein purification. SpyTag/SpyCatcher technology has been used to increase enzymatic output: SpyTag or SpyCatcher was fused to a multimerising protein scaffold, then enzymes fused to the partner pair are immobilised to the scaffold to increase the local enzyme concentration for increased catalytic activity (Choi et al., 2018; Giessen and Silver, 2016; Lim et al., 2019; Röder et al., 2017; Zhang et al., 2018a). With the reversible binding between SpyTag and SpyDock or SpySwitch, another level of modularity can be achieved by recycling the central scaffold to remove bound enzymes and rebind with other enzymes. This aspect can help with faster optimisation of enzymatic output by rapidly screening different mutant forms of enzymes or investigating different molar ratio combinations of enzymes in the enzyme cascade before establishing the final components with covalent SpyTag/SpyCatcher chemistry.

Due its physiologically-relevant range of pH sensitivity, SpySwitch can be utilised *in vivo*. Serum pH is at 7.4, but certain microenvironments have lower pHs such as the lysosomal

compartment (Mindell, 2012) and the tumour microenvironment (Kato et al., 2013). A tumour microenvironment can be detected by the fall in extracellular pH through FRET pairs fused to SpyTag and SpySwitch, which will dissociate at this low pH. Additionally, the stability of a protein against proteolytic degradation can be pH-controlled by cyclising the protein's termini with SpyTag/SpySwitch, as was done with cyclised β-lactamase, dihydrofolate reductase, phytase, luciferase (Schoene et al., 2016), xylanase (Gilbert et al., 2017), lichenase (Wang *et al.*, 2016), trehalose synthase (Xu et al., 2018), and phenylalanine aldolase (Wang *et al.*, 2019) using SpyTag/SpyCatcher. Naturally-occurring cyclised proteins and peptides have increased resistance to proteolysis because of lower termini flexibility (Trabi and Craik, 2002). A potent protein drug can be cyclised with SpyTag and SpySwitch to stably exert its effect onto target cells, however persistent exposure to the drug may be detrimental. Upon internalisation, the cyclised protein drug will be exposed to lower pH during lysosomal trafficking, thus opening the termini to provide better access for proteolytic degradation and limiting cytotoxicity.

In conclusion, this thesis has presented additional tools to extend the usage of SpyTag/SpyCatcher beyond conventional post-translational covalent bioconjugation into protein purification, defined multivalent nanoassembly and possibly into pH-dependent modulation of stability.

Bibliography

Abdelhamid, M.A.A., Motomura, K., Ikeda, T., Ishida, T., Hirota, R., and Kuroda, A. (2014). Affinity purification of recombinant proteins using a novel silica-binding peptide as a fusion tag. Appl. Microbiol. Biotechnol. *98*, 5677–5684.

Abe, H., Wakabayashi, R., Yonemura, H., Yamada, S., Goto, M., and Kamiya, N. (2013). Split Spy0128 as a potent scaffold for protein cross-linking and immobilization. Bioconjug. Chem. *24*, 242–250.

Abriata, L.A., Spiga, E., and Peraro, M.D. (2016). Molecular Effects of Concentrated Solutes on Protein Hydration, Dynamics, and Electrostatics. Biophys. J. *111*, 743–755.

Ackerson, C.J., Powell, R.D., and Hainfeld, J.F. (2010). Site-specific biomolecule labeling with gold clusters. Methods Enzymol. *481*, 195–230.

Aguino, C.F., Lang, M., Fernández-Luna, V., Pröschel, M., Sonnewald, U., Coto, P.B., and Costa, R.D. (2018). Single-Component Biohybrid Light-Emitting Diodes Using a White-Emitting Fused Protein. ACS Omega *3*, 15829–15836.

Ali, M.H., and Imperiali, B. (2005). Protein oligomerization: How and why. Bioorganic Med. Chem. *13*, 5013–5020.

Altman, J.D., Moss, P.A., Goulder, P.J., Barouch, D.H., McHeyzer-Williams, M.G., Bell, J.I., McMichael, A.J., and Davis, M.M. (1996). Phenotypic analysis of antigen-specific T lymphocytes. Science (80-.). *274*, 94–96.

Ambroggio, X.I., and Kuhlman, B. (2006). Computational design of a single amino acid sequence that can switch between two distinct protein folds. J. Am. Chem. Soc. *128*, 1154–1161.

Anandan, A., and Vrielink, A. (2016). Detergents inmembrane protein purification and

crystallisation. In Advances in Experimental Medicine and Biology, (Springer, Cham), pp. 13–28.

Andersson, A.M.C., Buldun, C.M., Pattinson, D.J., Draper, S.J., and Howarth, M. (2019). SnoopLigase peptide-peptide conjugation enables modular vaccine assembly. Sci. Rep. *9*, 4625.

Aon, J.C., Caimi, R.J., Taylor, A.H., Lu, Q., Oluboyede, F., Dally, J., Kessler, M.D., Kerrigan, J.J., Lewis, T.S., Wysocki, L.A., et al. (2008). Suppressing posttranslational gluconoylation of heterologous proteins by metabolic engineering of Escherichia coli. Appl. Environ. Microbiol. *74*, 950–958.

Arakawa, T., and Timasheff, S.N. (1991). The Interactions of Proteins with Salts, Amino Acids, and Sugars at High Concentration. In Advances in Comparative and Environmental Physiology, (Springer, Berlin, Heidelberg), pp. 226–245.

Aranko, A.S., Oeemig, J.S., Zhou, D., Kajander, T., Wlodawer, A., and Iwaï, H. (2014). Structure-based engineering and comparison of novel split inteins for protein ligation. Mol. Biosyst. *10*, 1023–1034.

Arkin, M.R., Glicksman, M.A., Fu, H., Havel, J.J., and Du, Y. (2004). Inhibition of Protein-Protein Interactions: Non-Cellular Assay Formats (Eli Lilly & Company and the National Center for Advancing Translational Sciences).

Arndt, K.M., Müller, K.M., and Plückthun, A. (2001). Helix-stabilized Fv (hsFv) antibody fragments: Substituting the constant domains of a fab fragment for a heterodimeric coiled-coil domain. J. Mol. Biol. *312*, 221–228.

Ashkenazi, A. (2002). Targeting death and decoy receptors of the tumour-necrosis factor superfamily. Nat. Rev. Cancer *2*, 420–430.

Ashkenazi, A. (2008). Directing cancer cells to self-destruct with pro-apoptotic receptor

agonists. Nat. Rev. Drug Discov. *7*, 1001–1012.

Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C., Rising, A., Johansson, J., and Knight, S.D. (2010). Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. Nature *465*, 236–238.

Asquith, R.S., Otterburn, M.S., and Sinclair, W.J. (1974). Isopeptide Crosslinks—Their Occurrence and Importance in Protein Structure. Angew. Chemie Int. Ed. English *13*, 514–520.

Aves, K.-L., Goksøyr, L., and Sander, A.F. (2020). Advantages and Prospects of Tag/Catcher Mediated Antigen Display on Capsid-Like Particle-Based Vaccines. Viruses *12*, 185.

Bae, Y., Kim, G.J., Kim, H., Park, S.G., Jung, H.S., and Kang, S. (2018). Engineering Tunable Dual Functional Protein Cage Nanoparticles Using Bacterial Superglue. Biomacromolecules *19*, 2896–2904.

Bai, L., Hu, K., Wang, T., Jastrab, J.M., Darwin, K.H., and Li, H. (2016). Structural analysis of the dodecameric proteasome activator PafE in Mycobacterium tuberculosis. Proc. Natl. Acad. Sci. U. S. A. *113*, E1983–E1992.

Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (1999). Circular permutation and receptor insertion within green fluorescent proteins. Proc. Natl. Acad. Sci. U. S. A. *96*, 11241–11246.

Bajar, B.T., Wang, E.S., Lam, A.J., Kim, B.B., Jacobs, C.L., Howe, E.S., Davidson, M.W., Lin, M.Z., and Chu, J. (2016). Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting. Sci. Rep. *6*, 20889.

Bal, W., Kozlowski, H., and Kasprzak, K.S. (2000). Molecular models in nickel carcinogenesis. J. Inorg. Biochem. *79*, 213–218.

Banerjee, A., and Howarth, M. (2018). Nanoteamwork: covalent protein assembly beyond

duets towards protein ensembles and orchestras. Curr. Opin. Biotechnol. *51*, 16–23.

Bedbrook, C.N., Kato, M., Ravindra kumar, S., Lakshmanan, A., Nath, R.D., Sun, F., Sternberg, P.W., Arnold, F.H., and Gradinaru, V. (2015). Genetically Encoded Spy Peptide Fusion System to Detect Plasma Membrane-Localized Proteins In Vivo. Chem. Biol. *22*, 1108–1121.

Bennett, N.R., Zwick, D.B., Courtney, A.H., and Kiessling, L.L. (2015). Multivalent Antigens for Promoting B and T Cell Activation. ACS Chem. Biol. *10*, 1817–1824.

Berg, J., Tymoczko, J., and Stryer, L. (2002). Chemical Bonds in Biochemistry.

van den Berg van Saparoea, H.B., Houben, D., de Jonge, M.I., Jong, W.S.P., and Luirink, J. (2018). Display of recombinant proteins on bacterial outer membrane vesicles by using protein ligation. Appl. Environ. Microbiol. *84*, e02567-17.

Blanco-Toribio, A., Sainz-Pastor, N., Álvarez-Cienfuegos, A., Merino, N., Cuesta, Á.M., Sánchez-Martín, D., Bonet, J., Santos-Valle, P., Sanz, L., Oliva, B., et al. (2013). Generation and characterization of monospecific and bispecific hexavalent trimerbodies. MAbs *5*, 70–79.

Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J., and Schäfer, F. (2009). Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC). A Review. Methods Enzymol. *463*, 439–473.

De Boer, E., Rodriguez, P., Bonte, E., Krijgsveldt, J., Katsantoni, E., Heckt, A., Grosveld, F., and Strouboulis, J. (2003). Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. Proc. Natl. Acad. Sci. U. S. A. *100*, 7480–7485.

Bolanos-Garcia, V.M., and Davies, O.R. (2006). Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. Biochim. Biophys. Acta - Gen. Subj. *1760*, 1304–1313.

Bonnet, J., Cartannaz, J., Tourcier, G., Contreras-Martel, C., Kleman, J.P., Morlot, C., Vernet, T., and Di Guilmi, A.M. (2017). Autocatalytic association of proteins by covalent bond formation: A Bio Molecular Welding toolbox derived from a bacterial adhesin. Sci. Rep. *7*, 43564.

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

Bornhorst, J.A., and Falke, J.J. (2000). Purification of proteins using polyhistidine affinity tags. Methods Enzymol. *326*, 245–254.

Boyken, S.E., Benhaim, M.A., Busch, F., Jia, M., Bick, M.J., Choi, H., Klima, J.C., Chen, Z., Walkey, C., Mileant, A., et al. (2019). De novo design of tunable, pH-driven conformational changes. Science (80-.). *364*, 658–664.

Braden, B.C., Dall'Acqua, W., Eisenstein, E., Fields, B.A., Goldbaum, F.A., Malchiodi, E.L., Mariuzza, R.A., Schwarz, F.P., Ysern, X., and Poljak, R.J. (1995). Protein motion and lock and key complementarity in antigen-antibody reactions. Pharm. Acta Helv. *69*, 225–230.

Braun, M.B., Traenkle, B., Koch, P.A., Emele, F., Weiss, F., Poetz, O., Stehle, T., and Rothbauer, U. (2016). Peptides in headlock - A novel high-affinity and versatile peptide-binding nanobody for proteomics and microscopy. Sci. Rep. *6*, 19211.

Brown, Z.P., and Takagi, J. (2018). The PA tag: A versatile peptide tagging system in the era of integrative structural biology. In Advances in Experimental Medicine and Biology, (Springer, Singapore), pp. 59–76.

Brune, K.D., and Howarth, M. (2018). New routes and opportunities for modular construction of particulate vaccines: Stick, click, and glue. Front. Immunol. *9*, 1432.

Brune, K.D., Leneghan, D.B., Brian, I.J., Ishizuka, A.S., Bachmann, M.F., Draper, S.J.,

Biswas, S., and Howarth, M. (2016). Plug-and-Display: decoration of Virus-Like Particles via isopeptide bonds for modular immunization. Sci. Rep. *6*, 19234.

Brune, K.D., Buldun, C.M., Li, Y., Taylor, I.J., Brod, F., Biswas, S., and Howarth, M. (2017). Dual Plug-and-Display Synthetic Assembly Using Orthogonal Reactive Proteins for Twin Antigen Immunization. Bioconjug. Chem. *28*, 1544–1551.

Bruun, T.U.J., Andersson, A.M.C., Draper, S.J., and Howarth, M. (2018). Engineering a Rugged Nanoscaffold to Enhance Plug-and-Display Vaccination. ACS Nano *12*, 8855–8866.

Budzik, J.M., Poor, C.B., Faull, K.F., Whitelegge, J.P., He, C., and Schneewind, O. (2009). Intramolecular amide bonds stabilize pili on the surface of bacilli. Proc. Natl. Acad. Sci. U. S. A. *106*, 19992–19997.

Buldun, C.M., Jean, J.X., Bedford, M.R., and Howarth, M. (2018). SnoopLigase Catalyzes Peptide-Peptide Locking and Enables Solid-Phase Conjugate Isolation. J. Am. Chem. Soc. *140*, 3008–3018.

Carson, M., Johnson, D.H., McDonald, H., Brouillette, C., and DeLucas, L.J. (2007). His-tag impact on structure. Acta Crystallogr. Sect. D Biol. Crystallogr. *63*, 295–301.

Cayetano-Cruz, M., Coffeen, C.F., Valadez-García, J., Montiel, C., and Bustos-Jaimes, I. (2018). Decoration of virus-like particles with an enzymatic activity of biomedical interest. Virus Res. *255*, 1–9.

Chaga, G., Hopp, J., and Nelson, P. (1999). Immobilized metal ion affinity chromatography on Co2+-carboxymethylaspartate-agarose Superflow, as demonstrated by one-step purification of lactate dehydrogenase from chicken breast muscle. Biotechnol. Appl. Biochem. *29 (Pt 1)*, 19–24.

Chakravarty, S., Ung, A.R., Moore, B., Shore, J., and Alshamrani, M. (2018). A Comprehensive Analysis of Anion-Quadrupole Interactions in Protein Structures.

Biochemistry *57*, 1852–1867.

Chakshusmathi, G., Mondal, K., Lakshmi, G.S., Singh, G., Roy, A., Babu, R.C., Madhusudhanan, S., and Varadarajan, R. (2004). Design of temperature-sensitive mutants solely from amino acid sequence. Proc. Natl. Acad. Sci. U. S. A. *101*, 7925–7930.

Chant, A., Kraemer-Pecore, C.M., Watkin, R., and Kneale, G.G. (2005). Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA causes a conformational change at the DNA-binding site. Protein Expr. Purif. *39*, 152–159.

Chatterjee, I., Somerville, G.A., Heilmann, C., Sahl, H.G., Maurer, H.H., and Herrmann, M. (2006). Very low ethanol concentrations affect the viability and growth recovery in post-stationary-phase Staphylococcus aureus populations. Appl. Environ. Microbiol. *72*, 2627–2636.

Chen, C.S., Smits, C., Dodson, G.G., Shevtsov, M.B., Merlino, N., Gollnick, P., and Antson, A.A. (2011a). How to change the oligomeric state of a circular protein assembly: Switch from 11-subunit to 12-subunit TRAP suggests a general mechanism. PLoS One *6*, e25296.

Chen, I., Dorr, B.M., and Liu, D.R. (2011b). A general strategy for the evolution of bond-forming enzymes using yeast display. Proc. Natl. Acad. Sci. U. S. A. *108*, 11399–11404.

Chen, X.P., Antson, A.A., Yang, M., Li, P., Baumann, C., Dodson, E.J., Dodson, G.G., and Gollnick, P. (1999). Regulatory features of the trp operon and the crystal structure of the trp RNA-binding attenuation protein from Bacillus stearothermophilus. J. Mol. Biol. *289*, 1003–1016.

Cho, J.H., Collins, J.J., and Wong, W.W. (2018). Universal Chimeric Antigen Receptors for Multiplexed and Logical Control of T Cell Responses. Cell *173*, 1426-1438.e11.

Choi, H., Choi, B., Kim, G.J., Kim, H. ul, Kim, H., Jung, H.S., and Kang, S. (2018). Fabrication

of Nanoreaction Clusters with Dual-Functionalized Protein Cage Nanobuilding Blocks. Small *14*, 1801488.

Chong, S., Mersha, F.B., Comb, D.G., Scott, M.E., Landry, D., Vence, L.M., Perler, F.B., Benner, J., Kucera, R.B., Hirvonen, C.A., et al. (1997). Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene *192*, 271–281.

Correnti, C.E., Gewe, M.M., Mehlin, C., Bandaranayake, A.D., Johnsen, W.A., Rupert, P.B., Brusniak, M.Y., Clarke, M., Burke, S.E., De Van Der Schueren, W., et al. (2018). Screening, large-scale production and structure-based classification of cystine-dense peptides. Nat. Struct. Mol. Biol. *25*, 270–278.

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

Coyle, B.L., and Baneyx, F. (2014). A cleavable silica-binding affinity tag for rapid and inexpensive protein purification. Biotechnol. Bioeng. *111*, 2019–2026.

Crick, F.H.C. (1953). The packing of α-helices: simple coiled-coils. Acta Crystallogr. *6*, 689–697.

Cristie-David, A.S., Sciore, A., Badieyan, S., Escheweiler, J.D., Koldewey, P., Bardwell, J.C.A., Ruotolo, B.T., and Marsh, E.N.G. (2017). Evaluation of: De novo -designed coiled coils as off-the-shelf components for protein assembly. Mol. Syst. Des. Eng. *2*, 140–148.

Cuatrecasas, P., Wilchek, M., and Anfinsen, C.B. (1968). Selective enzyme purification by affinity chromatography. Proc. Natl. Acad. Sci. U. S. A. *61*, 636–643.

Curk, T., Dobnikar, J., and Frenkel, D. (2017a). Design Principles for Super Selectivity using Multivalent Interactions. In Multivalency, (Chichester, UK: John Wiley & Sons, Ltd), pp. 75–

101.

Curk, T., Dobnikar, J., and Frenkel, D. (2017b). Optimal multivalent targeting of membranes with many distinct receptors. Proc. Natl. Acad. Sci. U. S. A. *114*, 7210–7215.

Dagliyan, O., Shirvanyants, D., Karginov, A. V., Ding, F., Fee, L., Chandrasekaran, S.N., Freisinger, C.M., Smolen, G.A., Huttenlocher, A., Hahn, K.M., et al. (2013). Rational design of a ligand-controlled protein conformational switch. Proc. Natl. Acad. Sci. U. S. A. *110*, 6800–6804.

Dames, S.A., Kammerer, R.A., Wiltscheck, R., Engel, J., and Alexandrescu, A.T. (1998). NMR structure of a parallel homotrimeric coiled coil. Nat. Struct. Mol. Biol. *5*, 687–691.

Damodaran, S., and Kinsella, J.E. (1981). The effects of neutral salts on the stability of macromolecules. A new approach using a protein-ligand binding system. J. Biol. Chem. *256*, 3394–3398.

Dang, B., Mravic, M., Hu, H., Schmidt, N., Mensa, B., and DeGrado, W.F. (2019). SNAC-tag for sequence-specific chemical protein cleavage. Nat. Methods *16*, 319–322.

Das, K.K., Reddy, R.C., Bagoji, I.B., Das, S., Bagali, S., Mullur, L., Khodnapur, J.P., and Biradar, M.S. (2019). Primary concept of nickel toxicity - An overview. J. Basic Clin. Physiol. Pharmacol. *30*, 141–152.

Deng, A., and Boxer, S.G. (2018). Structural Insight into the Photochemistry of Split Green Fluorescent Proteins: A Unique Role for a His-Tag. J. Am. Chem. Soc. *140*, 375–381.

Deweid, L., Avrutina, O., and Kolmar, H. (2018). Microbial transglutaminase for biotechnological and biomedical engineering. Biol. Chem. *400*, 257–274.

Dicko, C., Vollrath, F., and Kenney, J.M. (2004). Spider silk protein refolding is controlled by changing pH. Biomacromolecules *5*, 704–710.

Ding, P., Zhang, T., Li, Y., Teng, M., Sun, Y., Liu, X., Chai, S., Zhou, E., Jin, Q., and Zhang, G. (2017). Nanoparticle orientationally displayed antigen epitopes improve neutralizing antibody level in a model of porcine circovirus type 2. Int. J. Nanomedicine *12*, 5239–5254.

Donald, J.E., Kulp, D.W., and DeGrado, W.F. (2011). Salt bridges: Geometrically specific, designable interactions. Proteins Struct. Funct. Bioinforma. *79*, 898–915.

Dong, C., Beis, K., Nesper, J., Brunkan-LaMontagne, A.L., Clarke, B.R., Whitfield, C., and Naismith, J.H. (2006). Wza the translocon for E. coli capsular polysaccharides defines a new class of membrane protein. Nature *444*, 226–229.

Dovala, D., Sawyer, W.S., Rath, C.M., and Metzger, L.E. (2016). Rapid analysis of protein expression and solubility with the SpyTag-SpyCatcher system. Protein Expr. Purif. *117*, 44–51.

Drobnak, I., Gradišar, H., Ljubetič, A., Merljak, E., and Jerala, R. (2017). Modulation of Coiled-Coil Dimer Stability through Surface Residues while Preserving Pairing Specificity. J. Am. Chem. Soc. *139*, 8229–8236.

Dubacheva, G. V., Curk, T., Mognetti, B.M., Auzély-Velty, R., Frenkel, D., and Richter, R.P. (2014). Superselective targeting using multivalent polymers. J. Am. Chem. Soc. *136*, 1722–1725.

Dubacheva, G. V., Curk, T., Auzély-Velty, R., Frenkel, D., and Richter, R.P. (2015). Designing multivalent probes for tunable superselective targeting. Proc. Natl. Acad. Sci. U. S. A. *112*, 5579–5584.

Dubuisson, A., and Micheau, O. (2017). Antibodies and Derivatives Targeting DR4 and DR5 for Cancer Therapy. Antibodies *6*, 16.

Einhauer, A., and Jungbauer, A. (2001). The FLAG^{TM} peptide, a versatile fusion tag for the purification of recombinant proteins. J. Biochem. Biophys. Methods *49*, 455–465.

Erickson, H.P. (2009). Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. Biol. Proced. Online *11*, 32–51.

Escolano, A., Gristick, H.B., Abernathy, M.E., Merkenschlager, J., Gautam, R., Oliveira, T.Y., Pai, J., West, A.P., Barnes, C.O., Cohen, A.A., et al. (2019). Immunization expands B cells specific to HIV-1 V3 glycan in mice and macaques. Nature *570*, 468–473.

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

Evan, G.I., Lewis, G.K., Ramsay, G., and Bishop, J.M. (1985). Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. Mol. Cell. Biol. *5*, 3610–3616.

Fairhead, M., Veggiani, G., Lever, M., Yan, J., Mesner, D., Robinson, C. V., Dushek, O., Van Der Merwe, P.A., and Howarth, M. (2014). SpyAvidin hubs enable precise and ultrastable orthogonal nanoassembly. J. Am. Chem. Soc. *136*, 12355–12363.

Fairman, R., Chao, H.G., Lavoie, T.B., Villafranca, J.J., Matsueda, G.R., and Novotny, J. (1996). Design of heterotetrameric coiled coils: Evidence for increased stabilization by Glu--Lys+ ion pair interactions. Biochemistry *35*, 2824–2829.

Fan, C.Y., Huang, C.C., Chiu, W.C., Lai, C.C., Liou, G.G., Li, H.C., and Chou, M.Y. (2008). Production of multivalent protein binders using a self-trimerizing collagen-like peptide scaffold. FASEB J. *22*, 3795–3804.

Fanning, S.W., Murtaugh, M.L., and Horn, J.R. (2011). A combinatorial approach to engineering a dual-specific metal switch antibody. Biochemistry *50*, 5093–5095.

Field, J., Nikawa, J., Broek, D., MacDonald, B., Rodgers, L., Wilson, I.A., Lerner, R.A., and Wigler, M. (1988). Purification of a RAS-responsive adenylyl cyclase complex from Saccharomyces cerevisiae by use of an epitope addition method. Mol. Cell. Biol. *8*, 2159–2165.

Fierer, J.O., Veggiani, G., and Howarth, M. (2014). SpyLigase peptide-peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. Proc. Natl. Acad. Sci. U. S. A. *111*, E1176–E1181.

Firer, M.A. (2001). Efficient elution of functional proteins in affinity chromatography. J. Biochem. Biophys. Methods *49*, 433–442.

Fletcher, J.M., Boyle, A.L., Bruning, M., Bartlett, Sg.J., Vincent, T.L., Zaccai, N.R., Armstrong, C.T., Bromley, E.H.C., Booth, P.J., Brady, R.L., et al. (2012). A basis set of de novo coiled-Coil peptide oligomers for rational protein design and synthetic biology. ACS Synth. Biol. *1*, 240–250.

Fong, B.A., Wu, W.Y., and Wood, D.W. (2010). The potential role of self-cleaving purification tags in commercial-scale processes. Trends Biotechnol. *28*, 272–279.

Forsgren, N., Lamont, R.J., and Persson, K. (2010). Two Intramolecular Isopeptide Bonds are Identified in the Crystal Structure of the Streptococcus gordonii SspB C-terminal Domain. J. Mol. Biol. *397*, 740–751.

Freedman, S.J., Song, H.K., Xu, Y., Sun, Z.Y.J., and Eck, M.J. (2003). Homotetrameric structure of the SNAP-23 N-terminal coiled-coil domain. J. Biol. Chem. *278*, 13462–13467.

Futatsumori-Sugai, M., Abe, R., Watanabe, M., Kudou, M., Yamamoto, T., Ejima, D., Arakawa, T., and Tsumoto, K. (2009). Utilization of Arg-elution method for FLAG-tag based chromatography. Protein Expr. Purif. *67*, 148–155.

Gaberc-Porekar, V., and Menart, V. (2005). Potential for using histidine tags in purification of proteins at large scale. Chem. Eng. Technol. *28*, 1306–1314.

Gallivan, J.P., and Dougherty, D.A. (1999). Cation-π interactions in structural biology. Proc. Natl. Acad. Sci. U. S. A. *96*, 9459–9464.

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. In The Proteomics Protocols Handbook, (Totowa, NJ: Humana Press), pp. 571–607.

De Genst, E.J., Guilliams, T., Wellens, J., Day, E.M., Waudby, C.A., Meehan, S., Dumoulin, M., Hsu, S.T.D., Cremades, N., Verschueren, K.H.G., et al. (2010). Structure and properties of a complex of α-synuclein and a single-domain camelid antibody. J. Mol. Biol. *402*, 326–343.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gieffers, C., Kluge, M., Merz, C., Sykora, J., Thiemann, M., Schaal, R., Fischer, C., Branschädel, M., Abhari, B.A., Hohenberger, P., et al. (2013). APG350 induces superior clustering of trail receptors and shows therapeutic antitumor efficacy independent of cross-Linking via fcγ receptors. Mol. Cancer Ther. *12*, 2735–2747.

Giessen, T.W., and Silver, P.A. (2016). A Catalytic Nanoreactor Based on in Vivo Encapsulation of Multiple Enzymes in an Engineered Protein Nanocompartment. ChemBioChem *17*, 1931–1935.

Gilbert, C., Howarth, M., Harwood, C.R., and Ellis, T. (2017). Extracellular Self-Assembly of Functional and Tunable Protein Conjugates from Bacillus subtilis. ACS Synth. Biol. *6*, 957–967.

Gomes, A., Mohsen, M., and Bachmann, M. (2017). Harnessing Nanoparticles for Immunomodulation and Vaccines. Vaccines *5*, 6.

Götzke, H., Kilisch, M., Martínez-Carranza, M., Sograte-Idrissi, S., Rajavel, A., Schlichthaerle, T., Engels, N., Jungmann, R., Stenmark, P., Opazo, F., et al. (2019). The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. Nat. Commun. *10*, 4403.

Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, B.M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., et al. (2008). Protein production and purification. Nat. Methods *5*, 135–146.

Greenfield, N.J. (2007). Using circular dichroism spectra to estimate protein secondary structure. Nat. Protoc. *1*, 2876–2890.

Griffin, M., Casadio, R., and Bergamini, C.M. (2002). Transglutaminases: Nature's biological glues. Biochem. J. *368*, 377–396.

Grimsley, G.R., Scholtz, J.M., and Pace, C.N. (2009). A summary of the measured pK values of the ionizable groups in folded proteins. Protein Sci. *18*, 247–251.

Grönberg, A., Eriksson, M., Ersoy, M., and Johansson, H.J. (2011). A tool for increasing the lifetime of chromatography resins. MAbs *3*, 195–205.

di Guana, C., Lib, P., Riggsa, P.D., and Inouyeb, H. (1988). Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. Gene *67*, 21–30.

Guillén, D., Moreno-Mendieta, S., Aguilera, P., Sánchez, S., Farres, A., and Rodríguez-Sanoja, R. (2013). The starch-binding domain as a tool for recombinant protein purification. Appl. Microbiol. Biotechnol. *97*, 4141–4148.

Gülich, S., Linhult, M., Nygren, P.Å., Uhlén, M., and Hober, S. (2000). Stability towards alkaline conditions can be engineered into a protein ligand. J. Biotechnol. *80*, 169–178.

Gülich, S., Linhult, M., Ståhl, S., and Hober, S. (2003). Engineering streptococcal protein G for increased alkaline stability. Protein Eng. *15*, 835–842.

Guntas, G., Mitchell, S.F., and Ostermeier, M. (2004). A molecular switch created by in vitro recombination of nonhomologous genes. Chem. Biol. *11*, 1483–1487.

Hae, J.K., Coulibaly, F., Clow, F., Proft, T., and Baker, E.N. (2007). Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. Science (80-.). *318*, 1625–1628.

Hagan, R.M., Björnsson, R., McMahon, S.A., Schomburg, B., Braithwaite, V., Bühl, M., Naismith, J.H., and Schwarz-Linek, U. (2010). NMR spectroscopic and theoretical analysis of a spontaneously formed lys-asp isopeptide bond. Angew. Chemie - Int. Ed. *49*, 8421–8425.

Hage, D.S., and Matsuda, R. (2015). Affinity chromatography: A historical perspective. In Methods in Molecular Biology, (Humana Press, New York, NY), pp. 1–19.

Hamada, H., Arakawa, T., and Shiraki, K. (2009). Effect of Additives on Protein Aggregation. Curr. Pharm. Biotechnol. *10*, 400–407.

Hartman, N.C., and Groves, J.T. (2011). Signaling clusters in the cell membrane. Curr. Opin. Cell Biol. *23*, 370–376.

He, Y.G., Pappworth, I.Y., Rossbach, A., Paulin, J., Mavimba, T., Hayes, C., Kulik, L., Holers, V.M., Knight, A.M., and Marchbank, K.J. (2018). A novel C3d-containing oligomeric vaccine provides insight into the viability of testing human C3d-based vaccines in mice. Immunobiology *223*, 125–134.

Healthcare, G. (2017). Affinity Chromatography Vol.2 Tagged Proteins (GE Healthcare).

Hédoux, A., Krenzlin, S., Paccou, L., Guinet, Y., Flament, M.P., and Siepmann, J. (2010). Influence of urea and guanidine hydrochloride on lysozyme stability and thermal denaturation; A correlation between activity, protein dynamics and conformational changes. Phys. Chem. Chem. Phys. *12*, 13189–13196.

Hochuli, E., Döbeli, H., and Schacher, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. J. Chromatogr. A *411*, 177–184.

Hochuli, E., Bannwarth, W., Dobeli, H., Gentzi, R., and Stuber, D. (1988). Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. Bio/Technology *6*, 1321–1325.

Hoh, F., Uzest, M., Drucker, M., Plisson-Chastang, C., Bron, P., Blanc, S., and Dumas, C. (2010). Structural Insights into the Molecular Mechanisms of Cauliflower Mosaic Virus Transmission by Its Insect Vector. J. Virol. *84*, 4706–4713.

Holmberg, A., Blomstergren, A., Nord, O., Lukacs, M., Lundeberg, J., and Uhlén, M. (2005). The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. Electrophoresis *26*, 501–510.

Hopp, T.P., Prickett, K.S., Price, V.L., Libby, R.T., March, C.J., Cerretti, D.P., Urdal, D.L., and Conlon, P.J. (1988). A short polypeptide marker sequence useful for recombinant protein identification and purification. Bio/Technology *6*, 1204–1210.

Horiuchi, M., Takahasi, K., Kobashigawa, Y., Ochiai, M., and Inagaki, F. (2012). A low-cost affinity purification system using β-1,3-glucan recognition protein and curdlan beads. Protein Eng. Des. Sel. *25*, 405–413.

Hosse, R.J. (2006). A new generation of protein display scaffolds for molecular recognition. Protein Sci. *15*, 14–27.

Huang, A., de Jong, R.N., Folkers, G.E., and Boelens, R. (2010). NMR characterization of foldedness for the production of E3 RING domains. J. Struct. Biol. *172*, 120–127.

Huet, H.A., Growney, J.D., Johnson, J.A., Li, J., Bilic, S., Ostrom, L., Zafari, M., Kowal, C., Yang, G., Royo, A., et al. (2014). Multivalent nanobodies targeting death receptor 5 elicit superior tumor cell killing through efficient caspase induction. MAbs *6*, 1560–1570.

Hurst, C.H., Turnbull, D., Myles, S.M., Leslie, K., Keinath, N.F., and Hemsley, P.A. (2018). Variable effects of C-terminal fusions on FLS2 function: Not all epitope tags are created equal.

Plant Physiol. *177*, 522–531.

Ikeda, T., Ninomiya, K. ichi, Hirota, R., and Kuroda, A. (2010). Single-step affinity purification of recombinant proteins using the silica-binding Si-tag as a fusion partner. Protein Expr. Purif. *71*, 91–95.

Islam, T., Aguilar-Yañez, J.M., Simental-Martínez, J., Ortiz-Alcaraz, C.I., Rito-Palomares, M., and Fernandez-Lahore, M. (2014). A novel strategy for the purification of a recombinant protein using ceramic fluorapatite-binding peptides as affinity tags. J. Chromatogr. A *1339*, 26–33.

Izoré, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., Di Guilmi, A.M., and Dessen, A. (2010). Structural Basis of Host Cell Recognition by the Pilus Adhesin from Streptococcus pneumoniae. Structure *18*, 106–115.

Jackson, M.R., Beahm, R., Duvvuru, S., Narasimhan, C., Wu, J., Wang, H.N., Philip, V.M., Hinde, R.J., and Howell, E.E. (2007). A preference for edgewise interactions between aromatic rings and carboxylate anions: The biological relevance of anion-quadrupole interactions. J. Phys. Chem. B *111*, 8242–8249.

Jacobitz, A.W., Kattke, M.D., Wereszczynski, J., and Clubb, R.T. (2017). Sortase Transpeptidases: Structural Biology and Catalytic Mechanism. Adv. Protein Chem. Struct. Biol. *109*, 223–264.

Jarvik, J.W., and Telmer, C.A. (1998). Epitope Tagging. Annu. Rev. Genet. *32*, 601–618.

Jenne, A., and Famulok, M. (1999). Disruption of the streptavidin interaction with biotinylated nucleic acid probes by 2-mercaptoethanol. Biotechniques *26*, 249–254.

Jin, J., Hjerrild, K.A., Silk, S.E., Brown, R.E., Labbé, G.M., Marshall, J.M., Wright, K.E., Bezemer, S., Clemmensen, S.B., Biswas, S., et al. (2017). Accelerating the clinical development of protein-based vaccines for malaria by efficient purification using a four amino

acid C-terminal ‘C-tag’. Int. J. Parasitol. *47*, 435–446.

Johansson, H., Ljunglof, A., Aberg, P.-M., and GE Healthcare (2006). Regeneration Of A Chromatography Matrix (Worldwide).

Kalani, M.R., Moradi, A., Moradi, M., and Tajkhorshid, E. (2013). Characterizing a histidine switch controlling ph-dependent conformational changes of the influenza virus hemagglutinin. Biophys. J. *105*, 993–1003.

Kang, H.J., and Baker, E.N. (2011). Intramolecular isopeptide bonds: Protein crosslinks built for stress? Trends Biochem. Sci. *36*, 229–237.

Kang, H.J., Paterson, N.G., Kim, C.U., Middleditch, M., Chang, C., Ton-That, H., and 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 Crystallogr. Sect. D Biol. Crystallogr. *70*, 1190–1201.

Von Karstedt, S., Montinaro, A., and Walczak, H. (2017). Exploring the TRAILs less travelled: TRAIL in cancer biology and therapy. Nat. Rev. Cancer *17*, 352–366.

Kato, Y., Ozawa, S., Miyamoto, C., Maehata, Y., Suzuki, A., Maeda, T., and Baba, Y. (2013). Acidic extracellular microenvironment and cancer. Cancer Cell Int. *13*, 89.

Keeble, A.H., and Howarth, M. (2019). Insider information on successful covalent protein coupling with help from SpyBank. Methods Enzymol. *617*, 443–461.

Keeble, A.H., Banerjee, A., Ferla, M.P., Reddington, S.C., Anuar, I.N.A.K., and Howarth, M. (2017). Evolving Accelerated Amidation by SpyTag/SpyCatcher to Analyze Membrane Dynamics. Angew. Chemie - Int. Ed. *56*, 16521–16525.

Keeble, A.H., Turkki, P., Stokes, S., Anuar, I.N.A.K., Rahikainen, R., Hytönen, V.P., and Howarth, M. (2019). Approaching infinite affinity engineering of peptide-protein interaction.

Proc. Natl. Acad. Sci.

Keefe, A.D., Wilson, D.S., Seelig, B., and Szostak, J.W. (2001). One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-tag. Protein Expr. Purif. *23*, 440–446.

Khairil Anuar, I.N.A., Banerjee, A., Keeble, A.H., Carella, A., Nikov, G.I., and Howarth, M. (2019). Spy&Go purification of SpyTag-proteins using pseudo-SpyCatcher to access an oligomerization toolbox. Nat. Commun. *10*, 1734.

Khan, F., Legler, P.M., Mease, R.M., Duncan, E.H., Bergmann-Leitner, E.S., and Angov, E. (2012). Histidine affinity tags affect MSP1 42 structural stability and immunodominance in mice. Biotechnol. J. *7*, 133–147.

Kieliszek, M., and Misiewicz, A. (2014). Microbial transglutaminase and its application in the food industry. A review. Folia Microbiol. (Praha). *59*, 241–250.

Kiessling, L.L., Gestwicki, J.E., and Strong, L.E. (2006). Synthetic multivalent ligands as probes of signal transduction. Angew. Chemie - Int. Ed. *45*, 2348–2368.

Kim, Y.N., and Jung, Y. (2016). Artificial supramolecular protein assemblies as functional high-order protein scaffolds. Org. Biomol. Chem. *14*, 5352–5356.

Kim, Y.E., Kim, Y.N., Kim, J.A., Kim, H.M., and Jung, Y. (2015). Green fluorescent protein nanopolygons as monodisperse supramolecular assemblies of functional proteins with defined valency. Nat. Commun. *6*, 7134.

Kimple, M.E., Brill, A.L., and Pasker, R.L. (2013). Overview of affinity tags for protein purification. Curr. Protoc. Protein Sci. *73*, Unit 9.9.

Kiyokawa, T., Kanaori, K., Tajima, K., and Tanaka, T. (2000). Engineering of the hydrophobic core of an α-helical coiled coil. Biopolym. - Pept. Sci. Sect. *55*, 407–414.

Kiyokawa, T., Kanaori, K., Tajima, K., Kawaguchi, M., Mizuno, T., Oku, J.I., and Tanaka, T. (2004). Selective formation of AAB- and ABC-type heterotrimeric α-helical coiled coils. Chem. - A Eur. J. *10*, 3548–3554.

Klose, J., Wendt, N., Kubald, S., Krause, E., Fechner, K., Beyermann, M., Bienert, M., Rudolph, R., and Rothemund, S. (2004). Hexa-histidin tag position influences disulfide structure but not binding behavior of in vitro folded N-terminal domain of rat corticotropin-releasing factor receptor type 2a. Protein Sci. *13*, 2470–2475.

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

Könning, D., Hinz, S., Grzeschik, J., Schröter, C., Krah, S., Zielonka, S., and Kolmar, H. (2018). Construction of histidine-enriched shark ignar variable domain antibody libraries for the isolation of pH-sensitive vNAR fragments. In Methods in Molecular Biology, (Humana Press, New York, NY), pp. 109–127.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., and Hughes, C. (2000). Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. Nature *405*, 914–919.

Kozlowski, L.P. (2017). Proteome-pI: Proteome isoelectric point database. Nucleic Acids Res. *45*, D1112–D1116.

Kraatz, S., Guichard, P., Obbineni, J.M., Olieric, N., Hatzopoulos, G.N., Hilbert, M., Sen, I., Missimer, J., Gönczy, P., and Steinmetz, M.O. (2016). The Human Centriolar Protein CEP135 Contains a Two-Stranded Coiled-Coil Domain Critical for Microtubule Binding. Structure *24*, 1358–1371.

Krebs, S.J., Kwon, Y.D., Schramm, C.A., Law, W.H., Donofrio, G., Zhou, K.H., Gift, S.,

Dussupt, V., Georgiev, I.S., Schätzle, S., et al. (2019). Longitudinal Analysis Reveals Early Development of Three MPER-Directed Neutralizing Antibody Lineages from an HIV-1-Infected Individual. Immunity *50*, 677-691.e13.

Kretz, A.-L., Trauzold, A., Hillenbrand, A., Knippschild, U., Henne-Bruns, D., von Karstedt, S., and Lemke, J. (2019). TRAILblazing Strategies for Cancer Treatment. Cancers (Basel). *11*, 456.

Kronqvist, N., Otikovs, M., Chmyrov, V., Chen, G., Andersson, M., Nordling, K., Landreh, M., Sarr, M., Jörnvall, H., Wennmalm, S., et al. (2014). Sequential pH-driven dimerization and stabilization of the N-terminal domain enables rapid spider silk formation. Nat. Commun. *5*, 3254.

Kruger, R.G., Otvos, B., Frankel, B.A., Bentley, M., Dostal, P., and McCafferty, D.G. (2004). Analysis of the Substrate Specificity of the Staphylococcus aureus Sortase Transpeptidase SrtA. Biochemistry *43*, 1541–1551.

Kühnel, K., Jarchau, T., Wolf, E., Schlichting, I., Walter, U., Wittinghofer, A., and Strelkov, S. V. (2004). The VASP tetramerization domain is a right-handed coiled coil based on a 15-residue repeat. Proc. Natl. Acad. Sci. U. S. A. *101*, 17027–17032.

Kuo, W.H.K., and Chase, H.A. (2011). Exploiting the interactions between poly-histidine fusion tags and immobilized metal ions. Biotechnol. Lett. *33*, 1075–1084.

Kwon, H., Squire, C.J., Young, P.G., and Baker, E.N. (2014). Autocatalytically generated Thr-Gln ester bond cross-links stabilize the repetitive Ig-domain shaft of a bacterial cell surface adhesin. Proc. Natl. Acad. Sci. U. S. A. *111*, 1367–1372.

Kwon, H., Young, P.G., Squire, C.J., and Baker, E.N. (2017). Engineering a Lys-Asn isopeptide bond into an immunoglobulin-like protein domain enhances its stability. Sci. Rep. *7*, 42753.

Labrou, N.E. (2014). Protein purifi cation: An overview. In Methods in Molecular Biology, (Humana Press, Totowa, NJ), pp. 3–10.

Laitinen, O.H., Hytönen, V.P., Nordlund, H.R., and Kulomaa, M.S. (2006). Genetically engineered avidins and streptavidins. Cell. Mol. Life Sci. *63*, 2992–3017.

Lane, M.D., and Seelig, B. (2014). Advances in the directed evolution of proteins. Curr. Opin. Chem. Biol. *22*, 129–136.

Lawrence, M.S., Phillips, K.J., and Liu, D.R. (2007). Supercharging proteins can impart unusual resilience. J. Am. Chem. Soc. *129*, 10110–10112.

Lee, J., Natarajan, M., Nashine, V.C., Socolich, M., Vo, T., Russ, W.P., Benkovic, S.J., and Ranganathan, R. (2008). Surface sites for engineering allosteric control in proteins. Science (80-.). *322*, 438–442.

Leloup, N., Lössl, P., Meijer, D.H., Brennich, M., Heck, A.J.R., Thies-Weesie, D.M.E., and Janssen, B.J.C. (2017). Low pH-induced conformational change and dimerization of sortilin triggers endocytosed ligand release. Nat. Commun. *8*, 1708.

Lemke, J., Von Karstedt, S., Zinngrebe, J., and Walczak, H. (2014). Getting TRAIL back on track for cancer therapy. Cell Death Differ. *21*, 1350–1364.

Leneghan, D.B., Miura, K., Taylor, I.J., Li, Y., Jin, J., Brune, K.D., Bachmann, M.F., Howarth, M., Long, C.A., and Biswas, S. (2017). Nanoassembly routes stimulate conflicting antibody quantity and quality for transmission-blocking malaria vaccines. Sci. Rep. *7*, 3811.

Lennartz, F., Brod, F., Dabbs, R., Miura, K., Mekhaiel, D., Marini, A., Jore, M.M., Søgaard, M.M., Jørgensen, T., de Jongh, W.A., et al. (2018). Structural basis for recognition of the malaria vaccine candidate Pfs48/45 by a transmission blocking antibody. Nat. Commun. *9*, 3822.

Li, Y. (2010). Commonly used tag combinations for tandem affinity purification. Biotechnol. Appl. Biochem. *55*, 73–83.

Li, K., Zhang, R., Xu, Y., Wu, Z., Li, J., Zhou, X., Jiang, J., Liu, H., and Xiao, R. (2017). Sortase A-mediated crosslinked short-chain dehydrogenases/reductases as novel biocatalysts with improved thermostability and catalytic efficiency. Sci. Rep. *7*, 3081.

Li, L., Fierer, J.O., Rapoport, T.A., and Howarth, M. (2014). Structural analysis and optimization of the covalent association between SpyCatcher and a peptide tag. J. Mol. Biol. *426*, 309–317.

Liao, S.M., Du, Q.S., Meng, J.Z., Pang, Z.W., and Huang, R.B. (2013). The multiple roles of histidine in protein interactions. Chem. Cent. J. *7*, 44.

Lim, K.H., Huang, H., Pralle, A., and Park, S. (2013). Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection. Biotechnol. Bioeng. *110*, 57–67.

Lim, S., Jung, G.A., Glover, D.J., and Clark, D.S. (2019). Enhanced Enzyme Activity through Scaffolding on Customizable Self-Assembling Protein Filaments. Small *15*, 1805558.

Lin, S.C., Lin, I.P., Chou, W.I., Hsieh, C.A., Liu, S.H., Huang, R.Y., Sheu, C.C., and Chang, M.D.T. (2009). CBM21 starch-binding domain: A new purification tag for recombinant protein engineering. Protein Expr. Purif. *65*, 261–266.

Liu, S. (2015). A review on protein oligomerization process. Int. J. Precis. Eng. Manuf. *16*, 2731–2760.

Liu, Y., Liu, D., Yang, W., Wu, X.L., Lai, L., and Zhang, W. Bin (2017). Tuning SpyTag-SpyCatcher mutant pairs toward orthogonal reactivity encryption. Chem. Sci. *8*, 6577–6582.

Loredo-Varela, J., Chechik, M., Levdikov, V.M., Abd-El-Aziz, A., Minakhin, L., Severinov, K.,

Smits, C., and Antson, A.A. (2013). The putative small terminase from the thermophilic dsDNA bacteriophage G20C is a nine-subunit oligomer. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *69*, 876–879.

Ma, W., Saccardo, A., Roccatano, D., Aboagye-Mensah, D., Alkaseem, M., Jewkes, M., Di Nezza, F., Baron, M., Soloviev, M., and Ferrari, E. (2018). Modular assembly of proteins on nanoparticles. Nat. Commun. *9*, 1489.

Ma, Y., Zhang, J., and Zhao, Y. (2005). Correlation between the sensitivity to TRAIL and the expression level of DR5 on the surface of tumor cells. Chinese J. Clin. Oncol. *2*, 599–602.

MacFarlane, M. (2003). TRAIL-induced signalling and apoptosis. Toxicol. Lett. *139*, 89–97.

Magliery, T.J. (2015). Protein stability: Computation, sequence statistics, and new experimental methods. Curr. Opin. Struct. Biol. *33*, 161–168.

Malashkevich, V.N., Kammerer, R.A., Efimov, V.P., Schulthess, T., and Engel, J. (1996). The crystal structure of a five-stranded coiled coil in COMP: A prototype ion channel? Science (80-.). *274*, 761–765.

Manivel, V., Sahoo, N.C., Salunke, D.M., and Rao, K.V.S. (2000). Maturation of an antibody response is governed by modulations in flexibility of the antigen-combining site. Immunity *13*, 611–620.

Mao, H. (2004). A self-cleavable sortase fusion for one-step purification of free recombinant proteins. Protein Expr. Purif. *37*, 253–263.

Matsunaga, R., Yanaka, S., Nagatoishi, S., and Tsumoto, K. (2013). Hyperthin nanochains composed of self-polymerizing protein shackles. Nat. Commun. *4*, 2211.

McGowan, S., Buckle, A.M., Mitchell, M.S., Hoopes, J.T., Gallagher, D.T., Heselpoth, R.D., Shen, Y., Reboul, C.F., Law, R.H.P., Fischetti, V.A., et al. (2012). X-ray crystal structure of the

streptococcal specific phage lysin PlyC. Proc. Natl. Acad. Sci. U. S. A. *109*, 12752–12757.

Megeed, Z., Winters, R.M., and Yamush, M.L. (2006). Modulation of single-chain antibody affinity with temperature-responsive elastin-like polypeptide linkers. Biomacromolecules *7*, 999–1004.

Micsonai, A., Wien, F., Kernya, L., Lee, Y.H., Goto, Y., Réfrégiers, M., and Kardos, J. (2015). Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proc. Natl. Acad. Sci. U. S. A. *112*, E3095–E3103.

Mindell, J.A. (2012). Lysosomal Acidification Mechanisms. Annu. Rev. Physiol. *74*, 69–86.

Mohsen, M.O., Zha, L., Cabral-Miranda, G., and Bachmann, M.F. (2017). Major findings and recent advances in virus–like particle (VLP)-based vaccines. Semin. Immunol. *34*, 123–132.

Müller, E., and Vajda, J. (2016). Routes to improve binding capacities of affinity resins demonstrated for Protein A chromatography. J. Chromatogr. B Anal. Technol. Biomed. Life Sci. *1021*, 159–168.

Munro, S., and Pelham, H.R. (1984). Use of peptide tagging to detect proteins expressed from cloned genes: deletion mapping functional domains of Drosophila hsp 70. EMBO J. *3*, 3087–3093.

Murtaugh, M.L., Fanning, S.W., Sharma, T.M., Terry, A.M., and Horn, J.R. (2011). A combinatorial histidine scanning library approach to engineer highly pH-dependent protein switches. Protein Sci. *20*, 1619–1631.

Muyldermans, S. (2013). Nanobodies: Natural Single-Domain Antibodies. Annu. Rev. Biochem. *82*, 775–797.

Natoni, A., MacFarlane, M., Inoue, S., Walewska, R., Majid, A., Knee, D., Stover, D.R., Dyer, M.J.S., and Cohen, G.M. (2007). TRAIL signals to apoptosis in chronic lymphocytic leukaemia

cells primarily through TRAIL-R1 whereas cross-linked agonistic TRAIL-R2 antibodies facilitate signalling via TRAIL-R2. Br. J. Haematol. *139*, 568–577.

Nautiyal, S., and Alber, T. (2008). Crystal structure of a designed, thermostable, heterotrimeric coiled coil. Protein Sci. *8*, 84–90.

Nautiyal, S., Woolfson, D.N., King, D.S., and Alber, T. (1995). A Designed Heterotrimeric Coiled Coil. Biochemistry *34*, 11645–11651.

Naval, J., Miguel, D. De, Lleyda, A.G., Anel, A., and Lostao, L.M. (2019). Importance of trail molecular anatomy in receptor oligomerization and signaling. Implications for cancer therapy. Cancers (Basel). *11*, 444.

Nelson, D., Schuch, R., Chahales, P., Zhu, S., and Fischetti, V.A. (2006). PlyC: A multimeric bacteriophage lysin. Proc. Natl. Acad. Sci. U. S. A. *103*, 10765–10770.

Ng, S.Y.L., Lee, L.T.O., and Chow, B.K.C. (2013). Receptor oligomerization: From early evidence to current understanding in class B GPCRs. Front. Endocrinol. (Lausanne). *3*, 175.

Nguyen, G.K.T., Wang, S., Qiu, Y., Hemu, X., Lian, Y., and Tam, J.P. (2014). Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. Nat. Chem. Biol. *10*, 732–738.

Nguyen, G.K.T., Qiu, Y., Cao, Y., Hemu, X., Liu, C.F., and Tam, J.P. (2016). Butelase-mediated cyclization and ligation of peptides and proteins. Nat. Protoc. *11*, 1977–1988.

Nogi, T., Sangawa, T., Tabata, S., Nagae, M., Tamura-Kawakami, K., Beppu, A., Hattori, M., Yasui, N., and Takagi, J. (2008). Novel affinity tag system using structurally defined antibody-tag interaction: Application to single-step protein purification. Protein Sci. *17*, 2120–2126.

O'Leary, L., Van Der Sloot, A.M., Reis, C.R., Deegan, S., Ryan, A.E., Dhami, S.P.S., Murillo, L.S., Cool, R.H., De Sampaio, P.C., Thompson, K., et al. (2016). Decoy receptors block TRAIL

sensitivity at a supracellular level: The role of stromal cells in controlling tumour TRAIL sensitivity. Oncogene *35*, 1261–1270.

Oeemig, J.S., Aranko, A.S., Djupsjöbacka, J., Heinämäki, K., and Iwaï, H. (2009). Solution structure of DnaE intein from Nostoc punctiforme: Structural basis for the design of a new split intein suitable for site-specific chemical modification. FEBS Lett. *583*, 1451–1456.

Ogun, S.A., Dumon-Seignovert, L., Marchand, J.B., Holder, A.A., and Hill, F. (2008). The oligomerization domain of C4-binding protein (C4bp) acts as an adjuvant, and the fusion protein comprised of the 19-kilodalton merozoite surface protein 1 fused with the murine C4bp domain protects mice against malaria. Infect. Immun. *76*, 3817–3823.

Oke, M., Carter, L.G., Johnson, K.A., Liu, H., McMahon, S.A., Yan, X., Kerou, M., Weikart, N.D., Kadi, N., Sheikh, M.A., et al. (2010). The scottish structural proteomics facility: Targets, methods and outputs. J. Struct. Funct. Genomics *11*, 167–180.

Ong, E., Greenwood, J.M., Gilkes, N.R., Kilburn, D.G., Miller, R.C., and Warren, R.A.J. (1989). The cellulose-binding domains of cellulases: tools for biotechnology. Trends Biotechnol. *7*, 239–243.

Osička, R., Procházková, K., Šulc, M., Linhartová, I., Havlíček, V., and Šebo, P. (2004). A novel 'clip-and-link' activity of repeat in toxin (RTX) proteins from gram-negative pathogens: Covalent protein cross-linking by an Asp-Lys isopeptide bond upon calcium-dependent processing at an Asp-Pro bond. J. Biol. Chem. *279*, 24944–24956.

Ozawa, T. (2006). Designing split reporter proteins for analytical tools. Anal. Chim. Acta *556*, 58–68.

Pabst, T.M., Thai, J., and Hunter, A.K. (2018). Evaluation of recent Protein A stationary phase innovations for capture of biotherapeutics. J. Chromatogr. A *1554*, 45–60.

Pack, P., MÜller, K., Zahn, R., and Plückthun, A. (1995). Tetravalent Miniantibodies with High

Avidity Assembling inEscherichia colo. J. Mol. Biol. *246*, 28–34.

Panek, A., Pietrow, O., Filipkowski, P., and Synowiecki, J. (2013). Effects of the polyhistidine tag on kinetics and other properties of trehalose synthase from Deinococcus geothermalis. Acta Biochim. Pol. *60*, 163–166.

Papadopoulos, K.P., Isaacs, R., Bilic, S., Kentsch, K., Huet, H.A., Hofmann, M., Rasco, D., Kundamal, N., Tang, Z., Cooksey, J., et al. (2015). Unexpected hepatotoxicity in a phase I study of TAS266, a novel tetravalent agonistic Nanobody® targeting the DR5 receptor. Cancer Chemother. Pharmacol. *75*, 887–895.

Pardee, K., Slomovic, S., Nguyen, P.Q., Lee, J.W., Donghia, N., Burrill, D., Ferrante, T., McSorley, F.R., Furuta, Y., Vernet, A., et al. (2016). Portable, On-Demand Biomolecular Manufacturing. Cell *167*, 248-259.e12.

Paterson, N.G., and 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*, e22095.

Pauling, L., and Corey, R.B. (1953). Compound Helical Configurations of Polypeptide Chains: Structure of Proteins of the α-Keratin Type. Nature *171*, 59–61.

Pédelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. *24*, 79–88.

Pickart, C.M., and Eddins, M.J. (2004). Ubiquitin: Structures, functions, mechanisms. Biochim. Biophys. Acta - Mol. Cell Res. *1695*, 55–72.

Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. Nature *258*, 598–599.

Pröschel, M., Kraner, M.E., Horn, A.H.C., Schäfer, L., Sonnewald, U., and Sticht, H. (2017). Probing the potential of CnaB-type domains for the design of tag/catcher systems. PLoS One *12*, e0179740.

Puffer, E.B., Pontrello, J.K., Hollenbeck, J.J., Kink, J.A., and Kiessling, L.L. (2007). Activating B cells signalling with defined multivalent ligands. ACS Chem. Biol. *2*, 252–262.

Qin, S., Pang, X., and Zhou, H.X. (2011). Automated prediction of protein association rate constants. Structure *19*, 1744–1751.

Quintavalle, C., and Condorelli, G. (2012). Dulanermin in cancer therapy: Still much to do. Transl. Lung Cancer Res. *1*, 158–159.

Raghavan, M., and Bjorkman, P.J. (1996). Fc RECEPTORS AND THEIR INTERACTIONS WITH IMMUNOGLOBULINS. Annu. Rev. Cell Dev. Biol. *12*, 181–220.

Ranawakage, D.C., Takada, T., and Kamachi, Y. (2019). HiBiT-qIP, HiBiT-based quantitative immunoprecipitation, facilitates the determination of antibody affinity under immunoprecipitation conditions. Sci. Rep. *9*, 6895.

Rashid, F., Sharma, S., and Bano, B. (2005). Comparison of guanidine hydrochloride (GdnHCl) and urea denaturation on inactivation and unfolding of human placental cystatin (HPC). Protein J. *24*, 283–292.

Reubi, J.C., and Maecke, H.R. (2017). Approaches to multireceptor targeting: Hybrid radioligands, radioligand cocktails, and sequential radioligand applications. J. Nucl. Med. *58*, 10S-16S.

Reverberi, R., and Reverberi, L. (2007). Factors affecting the antigen-antibody reaction. Blood Transfus. *5*, 227–240.

Roberts, T.M., Rudolf, F., Meyer, A., Pellaux, R., Whitehead, E., Panke, S., and Held, M.

(2016). Identification and Characterisation of a pH-stable GFP. Sci. Rep. *6*, 28166.

Röder, J., Fischer, R., and Commandeur, U. (2017). Engineering Potato Virus X Particles for a Covalent Protein Based Attachment of Enzymes. Small *13*, 1702151.

Roldão, A., Mellado, M.C.M., Castilho, L.R., Carrondo, M.J.T., and Alves, P.M. (2010). Virus-like particles in vaccine development. Expert Rev. Vaccines *9*, 1149–1176.

Root, B.C., Pellegrino, L.D., Crawford, E.D., Kokona, B., and Fairman, R. (2009). Design of a heterotetrameric coiled coil. Protein Sci. *18*, 329–336.

Rudnick, S.I., and Adams, G.P. (2009). Affinity and avidity in antibody-based tumor targeting. Cancer Biother. Radiopharm. *24*, 155–161.

Rybak, J.N., Scheurer, S.B., Neri, D., and Elia, G. (2004). Purification of biotinylated proteins on streptavidin resin: A protocol for quantitative elution. Proteomics *4*, 2296–2299.

Sabaty, M., Grosse, S., Adryanczyk, G., Boiry, S., Biaso, F., Arnoux, P., and Pignol, D. (2013). Detrimental effect of the 6 His C-terminal tag on YedY enzymatic activity and influence of the TAT signal sequence on YedY synthesis. BMC Biochem. *14*, 28.

Sadilkova, L., Osicka, R., Sulc, M., Linhartova, I., Novak, P., and Sebo, P. (2008). Single-step affinity purification of recombinant proteins using a self-excising module from Neisseria meningitidis FrpC . Protein Sci. *17*, 1834–1843.

Saiz-Baggetto, S., Méndez, E., Quilis, I., Igual, J.C., and Bañó, M.C. (2017). Chimeric proteins tagged with specific 3xHA cassettes may present instability and functional problems. PLoS One *12*, e0183067.

Saraswat, M., Musante, L., Ravidá, A., Shortt, B., Byrne, B., and Holthofer, H. (2013). Preparative purification of recombinant proteins: Current status and future trends. Biomed Res. Int. *2013*, 312709.

Satav, T., Huskens, J., and Jonkheijm, P. (2015). Effects of Variations in Ligand Density on Cell Signaling. Small *11*, 5184–5199.

Schilcher, G., Schlagenhauf, A., Schneditz, D., Scharnagl, H., Ribitsch, W., Krause, R., Rosenkranz, A.R., Stojakovic, T., and Horina, J.H. (2013). Ethanol causes protein precipitation - New safety issues for catheter locking techniques. PLoS One *8*, e84869.

Schlosshauer, M., and Baker, D. (2004). Realistic protein-protein association rates from a simple diffusional model neglecting long-range interactions, free energy barriers, and landscape ruggedness. Protein Sci. *13*, 1660–1669.

Schmidt, T.G.M., and Skerra, A. (1993). The random peptide library-assisted engineering of a c-terminal affinity peptide, useful for the detection and purification of a functional ig Fv fragment. Protein Eng. Des. Sel. *6*, 109–122.

Schmidt, P.M., Sparrow, L.G., Attwood, R.M., Xiao, X., Adams, T.E., and McKimm-Breschkin, J.L. (2012). Taking down the FLAG! how insect cell expression challenges an established tag-system. PLoS One *7*, e37779.

Schmidt, T.G.M., Koepke, J., Frank, R., and Skerra, A. (1996). Molecular interaction between the strep-lag affinity peptide and its cognate target, streptavidin. J. Mol. Biol. *255*, 753–766.

Schmidt, T.G.M., Batz, L., Bonet, L., Carl, U., Holzapfel, G., Kiem, K., Matulewicz, K., Niermeier, D., Schuchardt, I., and Stanar, K. (2013). Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. Protein Expr. Purif. *92*, 54–61.

Schoene, C., Fierer, J.O., Bennett, S.P., and Howarth, M. (2014). SpyTag/Spycatcher cyclization confers resilience to boiling on a mesophilic enzyme. Angew. Chemie - Int. Ed. *53*, 6101–6104.

Schoene, C., Bennett, S.P., and Howarth, M. (2016). SpyRings Declassified: A Blueprint for

Using Isopeptide-Mediated Cyclization to Enhance Enzyme Thermal Resilience. Methods Enzymol. *580*, 149–167.

Schröter, C., Günther, R., Rhiel, L., Becker, S., Toleikis, L., Doerner, A., Becker, J., Schönemann, A., Nasu, D., Neuteboom, B., et al. (2015). A generic approach to engineer antibody pH-switches using combinatorial histidine scanning libraries and yeast display. MAbs *7*, 138–151.

Seed, B. (1994). Affinity purification methods involving imidazole elution.

Seiradake, E., Zhao, Y., Lu, W., Aricescu, A.R., and Jones, E.Y. (2015). Production of cell surface and secreted glycoproteins in mammalian cells. In Methods in Molecular Biology (Clifton, N.J.), pp. 115–127.

Shah, N.H., and Muir, T.W. (2014). Inteins: Nature's gift to protein chemists. Chem. Sci. *5*, 446–461.

Shahravan, S.H., Qu, X., Chan, I.S., and Shin, J.A. (2008). Enhancing the specificity of the enterokinase cleavage reaction to promote efficient cleavage of a fusion tag. Protein Expr. Purif. *59*, 314–319.

Shevchenko, A., Tomas, H., Havli, J., Olsen, J. V, and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. *1*, 2856–2860.

Silva, F., Queiroz, J.A., and Domingues, F.C. (2012). Evaluating metabolic stress and plasmid stability in plasmid DNA production by Escherichia coli. Biotechnol. Adv. *30*, 691–708.

Simeon, R., and Chen, Z. (2018). In vitro-engineered non-antibody protein therapeutics. Protein Cell *9*, 3–14.

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

Skowron, K.J., Speltz, T.E., and Moore, T.W. (2019). Recent structural advances in constrained helical peptides. Med. Res. Rev. *39*, 749–770.

Slaga, D., Ellerman, D., Lombana, T.N., Vij, R., Li, J., Hristopoulos, M., Clark, R., Johnston, J., Shelton, A., Mai, E., et al. (2018). Avidity-based binding to HER2 results in selective killing of HER2-overexpressing cells by anti-HER2/CD3. Sci. Transl. Med. *10*, eaat5775.

Smith, D.B., and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene *67*, 31–40.

Soto-Rodríguez, J., Coyle, B.L., Samuelson, A., Aravagiri, K., and Baneyx, F. (2017). Affinity purification of Car9-tagged proteins on silica matrices: Optimization of a rapid and inexpensive protein purification technology. Protein Expr. Purif. *135*, 70–77.

Spencer, R.K., and Hochbaum, A.I. (2017). The Phe-Ile Zipper: A Specific Interaction Motif Drives Antiparallel Coiled-Coil Hexamer Formation. Biochemistry *56*, 5300–5308.

Spriestersbach, A., Kubicek, J., Schäfer, F., Block, H., and Maertens, B. (2015). Purification of His-Tagged Proteins. Methods Enzymol. *559*, 1–15.

Sridharan, U., and Ponnuraj, K. (2016). Isopeptide bond in collagen- and fibrinogen-binding MSCRAMMs. Biophys. Rev. *8*, 75–83.

Starkenstein, E. (1910). Ferment action and the influence upon it of neutral salts. Biochem. Z *24*, 210–218.

Steffen, W., Ko, F.C., Patel, J., Lyamichev, V., Albert, T.J., Benz, J., Rudolph, M.G., Bergmann, F., Streidl, T., Kratzsch, P., et al. (2017). Discovery of a microbial transglutaminase enabling highly site-specific labeling of proteins. J. Biol. Chem. *292*, 15622–15635.

Stein, V., and Alexandrov, K. (2015). Synthetic protein switches: Design principles and applications. Trends Biotechnol. *33*, 101–110.

Stein, V., Nabi, M., and Alexandrov, K. (2017). Ultrasensitive Scaffold-Dependent Protease Sensors with Large Dynamic Range. ACS Synth. Biol. *6*, 1337–1342.

Stepanenko, O. V., Stepanenko, O. V., Kuznetsova, I.M., Shcherbakova, D.M., Verkhusha, V. V., and Turoverov, K.K. (2012). Distinct Effects of Guanidine Thiocyanate on the Structure of Superfolder GFP. PLoS One *7*, e48809.

Stepanenko, O. V., Stepanenko, O. V., Kuznetsova, I.M., Verkhusha, V. V., and Turoverov, K.K. (2014). Sensitivity of superfolder GFP to ionic agents. PLoS One *9*, e110750.

Stetefeld, J., Jenny, M., Schulthess, T., Landwehr, R., Engel, J., and Kammerer, R.A. (2000). Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. Nat. Struct. Biol. *7*, 772–776.

Stevens, A.J., Brown, Z.Z., Shah, N.H., Sekar, G., Cowburn, D., and Muir, T.W. (2016). Design of a Split Intein with Exceptional Protein Splicing Activity. J. Am. Chem. Soc. *138*, 2162–2165.

Stranges, P.B., Palla, M., Kalachikov, S., Nivala, J., Dorwart, M., Trans, A., Kumar, S., Porel, M., Chien, M., Tao, C., et al. (2016). Design and characterization of a nanopore-coupled polymerase for single-molecule DNA sequencing by synthesis on an electrode array. Proc. Natl. Acad. Sci. U. S. A. *113*, E6749–E6756.

Stratton, M.M., Mitrea, D.M., and Loh, S.N. (2008). A Ca2+-sensing molecular switch based on alternate frame protein folding. ACS Chem. Biol. *3*, 723–732.

Strauch, E.M., Fleishman, S.J., and Baker, D. (2014). Computational design of a pH-sensitive IgG binding protein. Proc. Natl. Acad. Sci. U. S. A. *111*, 675–680.

Sugimoto, N., Igarashi, K., and Samejima, M. (2012). Cellulose affinity purification of fusion proteins tagged with fungal family 1 cellulose-binding domain. Protein Expr. Purif. *82*, 290–296.

Sun, L., Young, L.N., Zhang, X., Boudko, S.P., Fokine, A., Zbornik, E., Roznowski, A.P., Molineux, I.J., Rossmann, M.G., and Fane, B.A. (2014). Icosahedral bacteriophage Φx174 forms a tail for DNA transport during infection. Nature *505*, 432–435.

Swartz, A.R., and Chen, W. (2018). Rapid Quantification of Monoclonal Antibody Titer in Cell Culture Harvests by Antibody-Induced Z-ELP-E2 Nanoparticle Cross-Linking. Anal. Chem. *90*, 14447–14452.

Swers, J.S., Grinberg, L., Wang, L., Feng, H., Lekstrom, K., Carrasco, R., Xiao, Z., Inigo, I., Leow, C.C., Wu, H., et al. (2013). Multivalent scaffold proteins as superagonists of TRAIL receptor 2-induced apoptosis. Mol. Cancer Ther. *12*, 1235–1244.

Takeda, H., Zhou, W., Kido, K., Suno, R., Iwasaki, T., Kobayashi, T., and Sawasaki, T. (2017). CP5 system, for simple and highly efficient protein purification with a C-terminal designed mini tag. PLoS One *12*, e0178246.

Talley, K., and Alexov, E. (2010). On the pH-optimum of activity and stability of proteins. Proteins Struct. Funct. Bioinforma. *78*, 2699–2706.

Tan, L.L., Hoon, S.S., and Wong, F.T. (2016). Kinetic controlled Tag-Catcher interactions for directed covalent protein assembly. PLoS One *11*, e0165074.

Taskinen, B., Zauner, D., Lehtonen, S.I., Koskinen, M., Thomson, C., Kähkönen, N., Kukkurainen, S., Määttä, J.A.E., Ihalainen, T.O., Kulomaa, M.S., et al. (2014). Switchavidin: Reversible biotin-avidin-biotin bridges with high affinity and specificity. Bioconjug. Chem. *25*, 2233–2243.

Terpe, K. (2003). Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biotechnol. *60*, 523–533.

Terskikh, A. V., Le Doussal, J.M., Crameri, R., Fisch, I., Mach, J.P., and Kajava, A. V. (1997). ‘Peptabody’: A new type of high avidity binding protein. Proc. Natl. Acad. Sci. U. S. A. *94*,

1663–1668.

Thomas, F., Boyle, A.L., Burton, A.J., and Woolfson, D.N. (2013). A set of de novo designed parallel heterodimeric coiled coils with quantified dissociation constants in the micromolar to sub-nanomolar regime. J. Am. Chem. Soc. *135*, 5161–5166.

Thompson, K.E., Bashor, C.J., Lim, W.A., and Keating, A.E. (2012). Synzip protein interaction toolbox: In vitro and in vivo specifications of heterospecific coiled-coil interaction domains. ACS Synth. Biol. *1*, 118–129.

Thomson, A.R., Wood, C.W., Burton, A.J., Bartlett, G.J., Sessions, R.B., Brady, R.L., and Woolfson, D.N. (2014). Computational design of water-soluble α-helical barrels. Science (80-.). *346*, 485–488.

Thrane, S., Janitzek, C.M., Matondo, S., Resende, M., Gustavsson, T., Jongh, W.A., Clemmensen, S., Roeffen, W., Vegte-Bolmer, M., Gemert, G.J., et al. (2016). Bacterial superglue enables easy development of efficient virus-like particle based vaccines. J. Nanobiotechnology *14*, 30.

Tomomori-Sato, C., Sato, S., Conaway, R.C., and Conaway, J.W. (2013). Immunoaffinity purification of protein complexes from mammalian cells. Methods Mol. Biol. *977*, 273–287.

Torres, F., Das Graças, M., Melo, M., and Tosti, A. (2009). Management of contact dermatitis due to nickel allergy: An update. Clin. Cosmet. Investig. Dermatol. *2*, 39–48.

Tørresen, O.K., Star, B., Mier, P., Andrade-Navarro, M.A., Bateman, A., Jarnot, P., Gruca, A., Grynberg, M., Kajava, A. V, Promponas, V.J., et al. (2019). Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. Nucleic Acids Res. *47*, 10994–11006.

Trabi, M., and Craik, D.J. (2002). Circular proteins - No end in sight. Trends Biochem. Sci. *27*, 132–138.

Trouche, N., Wieckowski, S., Sun, W., Chaloin, O., Hoebeke, J., Fournel, S., and Guichard, G. (2007). Small multivalent architectures mimicking homotrimers of the TNF superfamily member CD40L: Delineating the relationship between structure and effector function. J. Am. Chem. Soc. *129*, 13480–13492.

Truebestein, L., and Leonard, T.A. (2016). Coiled-coils: The long and short of it. BioEssays *38*, 903–916.

Tsukamoto, M., Watanabe, H., Ooishi, A., and Honda, S. (2014). Engineered protein A ligands, derived from a histidine-scanning library, facilitate the affinity purification of IgG under mild acidic conditions. J. Biol. Eng. *8*, 15.

Tsumoto, K., Ejima, D., Senczuk, A.M., Kita, Y., and Arakawa, T. (2007). Effects of salts on protein-surface interactions: Applications for column chromatography. J. Pharm. Sci. *96*, 1677–1690.

Tytgat, H.L.P., Schoofs, G., Driesen, M., Proost, P., Van Damme, E.J.M., Vanderleyden, J., and Lebeer, S. (2015). Endogenous biotin-binding proteins: An overlooked factor causing false positives in streptavidin-based protein detection. Microb. Biotechnol. *8*, 164–168.

Valéry, C., Deville-Foillard, S., Lefebvre, C., Taberner, N., Legrand, P., Meneau, F., Meriadec, C., Delvaux, C., Bizien, T., Kasotakis, E., et al. (2015). Atomic view of the histidine environment stabilizing higher-pH conformations of pH-dependent proteins. Nat. Commun. *6*, 7771.

Vassylyeva, M.N., Klyuyev, S., Vassylyev, A.D., Wesson, H., Zhang, Z., Renfrow, M.B., Wang, H., Higgins, N.P., Chow, L.T., and Vassylyev, D.G. (2017). Efficient, ultra-high-Affinity chromatography in a one-step purification of complex proteins. Proc. Natl. Acad. Sci. U. S. A. *114*, E5138–E5147.

Vauquelin, G., and Charlton, S.J. (2013). Exploring avidity: Understanding the potential gains

in functional affinity and target residence time of bivalent and heterobivalent ligands. Br. J. Pharmacol. *168*, 1771–1785.

Veggiani, G., Nakamura, T., Brenner, M.D., Gayet, R. V., Yan, J., Robinson, C. V., and Howarth, M. (2016). Programmable polyproteams built using twin peptide superglues. Proc. Natl. Acad. Sci. U. S. A. *113*, 1202–1207.

Voss, S., and Skerra, A. (1997). Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. Protein Eng. *10*, 975–982.

Voulgaraki, D., Mitnacht-Kraus, R., Letarte, M., Foster-Cuevas, M., Brown, M.H., and Barclay, A.N. (2005). Multivalent recombinant proteins for probing functions of leucocyte surface proteins such as the CD200 receptor. Immunology *115*, 337–346.

Wallis, R., Leung, K.Y., Pommer, A.J., Videler, H., Moore, G.R., James, R., and Kleanthous, C. (1995a). Protein-Protein Interactions in Colicin E9 DNase-Immunity Protein Complexes. 2. Cognate and Noncognate Interactions That Span the Millimolar to Femtomolar Affinity Range. Biochemistry *34*, 13751–13759.

Wallis, R., Moore, G.R., James, R., and Kleanthous, C. (1995b). Protein-Protein Interactions in Colicin E9 DNase-Immunity Protein Complexes. 1. Diffusion-Controlled Association and Femtomolar Binding for the Cognate Complex. Biochemistry *34*, 13743–13750.

Wang, H., Vilela, M., Winkler, A., Tarnawski, M., Schlichting, I., Yumerefendi, H., Kuhlman, B., Liu, R., Danuser, G., and Hahn, K.M. (2016a). LOVTRAP: An optogenetic system for photoinduced protein dissociation. Nat. Methods *13*, 755–758.

Wang, J., Wang, Y., Wang, X., Zhang, D., Wu, S., and Zhang, G. (2016b). Enhanced thermal stability of lichenase from Bacillus subtilis 168 by SpyTag/SpyCatcher-mediated spontaneous cyclization. Biotechnol. Biofuels *9*, 79.

Wang, W., He, W., Wang, L., Zhang, G., and Gao, B. (2013). Pentamerisation of a scFv directed against TRAIL receptor 2 increases its antitumour efficacy. Immunol. Cell Biol. *91*, 360–367.

Wang, W., Liu, Z., Zhou, X., Guo, Z., Zhang, J., Zhu, P., Yao, S., and Zhu, M. (2019a). Ferritin nanoparticle-based SpyTag/SpyCatcher-enabled click vaccine for tumor immunotherapy. Nanomedicine Nanotechnology, Biol. Med. *16*, 69–78.

Wang, Y., Tian, J., Xiao, Y., Wang, Y., Sun, H., Chang, Y., and Luo, H. (2019b). SpyTag/SpyCatcher cyclization enhances the thermostability and organic solvent tolerance of l-phenylalanine aldolase. Biotechnol. Lett. *41*, 987–994.

Watanabe, H., Matsumaru, H., Ooishi, A., and Honda, S. (2013). Structure-based histidine substitution for optimizing pH-sensitive Staphylococcus protein A. J. Chromatogr. B Anal. Technol. Biomed. Life Sci. *929*, 155–160.

Watanabe, N., Nomura, T., Takai, T., Chiba, T., Honjo, T., and Tsubata, T. (1998). Antigen receptor cross-linking by anti-immunoglobulin antibodies coupled to cell surface membrane induces rapid apoptosis of normal spleen B cells. Scand. J. Immunol. *47*, 541–547.

Watson, D.S., Platt, V.M., Cao, L., Venditto, V.J., and Szoka, F.C. (2011). Antibody response to polyhistidine-tagged peptide and protein antigens attached to liposomes via lipid-linked nitrilotriacetic acid in mice. Clin. Vaccine Immunol. *18*, 289–297.

Waugh, D.S. (2005). Making the most of affinity tags. Trends Biotechnol. *23*, 316–320.

Waugh, D.S. (2011). An overview of enzymatic reagents for the removal of affinity tags. Protein Expr. Purif. *80*, 283–293.

Wegmuller, S., and Schmid, S. (2014). Recombinant Peptide Production in Microbial Cells. Curr. Org. Chem. *18*, 1005–1019.

White, K.A., Ruiz, D.G., Szpiech, Z.A., Strauli, N.B., Hernandez, R.D., Jacobson, M.P., and Barber, D.L. (2017). Cancer-associated arginine-to-histidine mutations confer a gain in pH sensing to mutant proteins. Sci. Signal. *10*, eaam9931.

Wieduwild, R., and Howarth, M. (2018). Assembling and decorating hyaluronan hydrogels with twin protein superglues to mimic cell-cell interactions. Biomaterials *180*, 253–264.

Wilchek, M., and Bayer, E.A. (1988). The avidin-biotin complex in bioanalytical applications. Anal. Biochem. *171*, 1–32.

Willuda, J., Kubetzko, S., Waibel, R., Schubiger, P.A., Zangemeister-Wittke, U., and Plückthun, A. (2001). Tumor Targeting of Mono-, Di-, and Tetravalent Anti-p185HER-2 Miniantibodies Multimerized by Self-associating Peptides. J. Biol. Chem. *276*, 14385–14392.

Woestenenk, E.A., Hammarström, M., Van Den Berg, S., Härd, T., and Berglund, H. (2004). His tag effect on solubility of human proteins produced in Escherichia coli: A comparison between four expression vectors. J. Struct. Funct. Genomics *5*, 217–229.

Woolfson, D.N. (2005). The design of coiled-coil structures and assemblies. In Advances in Protein Chemistry, pp. 79–112.

Woolfson, D.N. (2017). Coiled-coil design: Updated and upgraded. In Sub-Cellular Biochemistry, pp. 35–61.

Wu, J., and Filutowicz, M. (1999). Hexahistidine (His6)-tag dependent protein dimerization: A cautionary tale. Acta Biochim. Pol. *46*, 591–599.

Wu, S.C., and Wong, S.L. (2005). Engineering soluble monomeric streptavidin with reversible biotin binding capability. J. Biol. Chem. *280*, 23225–23231.

Wu, M.M., Llopis, J., Adams, S., McCaffery, J.M., Kulomaa, M.S., Machen, T.E., Moore, H.P.H., and Tsien, R.Y. (2000). Organelle pH studies using targeted avidin and fluorescein-

biotin. Chem. Biol. *7*, 197–209.

Wu, X.L., Liu, Y., Liu, D., Sun, F., and Zhang, W. Bin (2018). An Intrinsically Disordered Peptide-Peptide Stapler for Highly Efficient Protein Ligation Both in Vivo and in Vitro. J. Am. Chem. Soc. *140*, 17474–17483.

Xu, C., Xu, Q., Huang, H., and Jiang, L. (2018). Enhancing the stability of trehalose synthase via spytag/spycatcher cyclization to improve its performance in industrial biocatalysts. Biosci. Biotechnol. Biochem. *82*, 1473–1479.

Yang, J., and Yang, Y. (2012). Plasmid size can affect the ability of Escherichia coli to produce high-quality plasmids. Biotechnol. Lett. *34*, 2017–2022.

Yang, Y., Li, J., and Zhang, Y. (2013). Characterization of the diatomite binding domain in the ribosomal protein L2 from E. coli and functions as an affinity tag. Appl. Microbiol. Biotechnol. *97*, 2541–2549.

Yoon, H.R., Choi, H., Choi, Y.A., Kim, J.A., Jung, J., Kim, H.M., and Jung, Y. (2018). Fabrication of Oligomeric Avidin Scaffolds for Valency-Controlled Surface Display of Functional Ligands. Angew. Chemie - Int. Ed. *57*, 12410–12414.

Yoshikawa, H., Hirano, A., Arakawa, T., and Shiraki, K. (2012). Effects of alcohol on the solubility and structure of native and disulfide-modified bovine serum albumin. Int. J. Biol. Macromol. *50*, 1286–1291.

Young, P.G., and Squire, C.J. (2020). Molecular Superglues: Discovery and Engineering Orthogonalization. In Methods in Molecular Biology, pp. 85–99.

Young, P.G., Yosaatmadja, Y., Harris, P.W.R., Leung, I.K.H., Baker, E.N., and Squire, C.J. (2017). Harnessing ester bond chemistry for protein ligation. Chem. Commun. *53*, 1502–1505.

Yu, X., Yang, Y.-P., Dikici, E., Deo, S.K., and Daunert, S. (2017). Beyond Antibodies as

Binding Partners: The Role of Antibody Mimetics in Bioanalysis. Annu. Rev. Anal. Chem. *10*, 293–320.

Zaccai, N.R., Chi, B., Thomson, A.R., Boyle, A.L., Bartlett, G.J., Bruning, M., Linden, N., Sessions, R.B., Booth, P.J., Brady, R.L., et al. (2011). A de novo peptide hexamer with a mutable channel. Nat. Chem. Biol. *7*, 935–941.

Zähner, D., Gandhi, A.R., Stuchlik, O., Reed, M., Pohl, J., and Stephens, D.S. (2011). Pilus backbone protein PitB of Streptococcus pneumoniae contains stabilizing intramolecular isopeptide bonds. Biochem. Biophys. Res. Commun. *409*, 526–531.

Zakeri, B., and Howarth, M. (2010). Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting. J. Am. Chem. Soc. *132*, 4526–4527.

Zakeri, B., Fierer, J.O., Celik, E., Chittock, E.C., Schwarz-Linek, U., Moy, V.T., and Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. Proc. Natl. Acad. Sci. U. S. A. *109*, E690-7.

Zhang, G., Quin, M.B., and Schmidt-Dannert, C. (2018a). Self-Assembling Protein Scaffold System for Easy in Vitro Coimmobilization of Biocatalytic Cascade Enzymes. ACS Catal. *8*, 5611–5620.

Zhang, X.J., Wang, X.W., Da, X. Di, Shi, Y., Liu, C., Sun, F., Yang, S., and Zhang, W. Bin (2018b). A Versatile and Robust Approach to Stimuli-Responsive Protein Multilayers with Biologically Enabled Unique Functions. Biomacromolecules *19*, 1065–1073.

Zhang, X.J., Wang, X.W., Sun, J.X., Su, C., Yang, S., and Zhang, W. Bin (2018c). Synergistic Enhancement of Enzyme Performance and Resilience via Orthogonal Peptide-Protein Chemistry Enabled Multilayer Construction. Biomacromolecules *19*, 2700–2707.

Zhang, X.J., Wu, X.L., Wang, X.W., Liu, D., Yang, S., and Zhang, W. Bin (2018d). SpyCatcher-N TEV : A Circularly Permuted, Disordered SpyCatcher Variant for Less Trace Ligation.

Bioconjug. Chem. *29*, 1622–1629.

Zhang, X.J., Wu, X.L., Da, X. Di, Wang, X.W., Yang, S., and Zhang, W.-B. (2019a). Engineering SpyCatcher Variants with Proteolytic Sites for Less-Trace Ligation. Chinese J. Chem. *37*, 113–118.

Zhang, Y., Yoshida, T., and Zhang, B. (2009). TRAIL induces endocytosis of its death receptors in MDA-MB-231 breast cancer cells. Cancer Biol. Ther. *8*, 917–922.

Zhang, Y., Cheng, M., Cao, J., Zhang, Y., Yuan, Z., Wu, Q., and Wang, W. (2019b). Multivalent nanoparticles for personalized theranostics based on tumor receptor distribution behavior. Nanoscale *11*, 5005–5013.

Zhao, H., Shen, A., Xiang, Y.K., and Corey, D.P. (2016a). Three recombinant engineered antibodies against recombinant tags with high affinity and specificity. PLoS One *11*, e0150125.

Zhao, Q., Xu, W., Xing, L., and Lin, Z. (2016b). Recombinant production of medium- to large-sized peptides in Escherichia coli using a cleavable self-aggregating tag. Microb. Cell Fact. *15*, 136.

Zheng, A., Yuan, F., Kleinfelter, L.M., and Kielian, M. (2014). A toggle switch controls the low pH-triggered rearrangement and maturation of the dengue virus envelope proteins. Nat. Commun. *5*, 3877.

Zheng, B., Graham, F.L., Johnson, D.C., Hanke, T., McDermott, M.R., and Prevec, L. (1993). Immunogenicity in mice of tandem repeats of an epitope from herpes simplex gD protein when expressed by recombinant adenovirus vectors. Vaccine *11*, 1191–1198.

Zheng, Y., Wu, Q., Wang, C., Xu, M.Q., and Liu, Y. (2012). Mutual synergistic protein folding in split intein. Biosci. Rep. *32*, 433–442.

Zhu, J. (2012). Mammalian cell protein expression for biopharmaceutical production.

Biotechnol. Adv. *30*, 1158–1170.

Zhu, X., Wang, L., Liu, R., Flutter, B., Li, S., Ding, J., Tao, H., Liu, C., Sun, M., and Gao, B. (2010). COMBODY: One-domain antibody multimer with improved avidity. Immunol. Cell Biol. *88*, 667–675.

Zinkernagel, R.M. (2003). On Natural and Artificial Vaccinations. Annu. Rev. Immunol. *21*, 515–546.