Establishing modular ligation through transglutaminase engineering



Dominic Paul Wood

Department of Biochemistry

Wolfson College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2024

Abstract

The ability to precisely direct proteins into ordered complexes is highly valuable for the synthetic biologist. Tools that can direct formation of covalent bonds facilitate particularly robust assemblies, which may enable new properties in areas such as optical control or programming of cell behaviour. Decades of research have produced an extensive range of tools for covalent conjugation of proteins both *in vitro* and *in vivo,* each with specific requirements and characteristics that determine the scope of their application. However, there is still a need for systems that can conjugate endogenous proteins in the extracellular space or precisely assemble endogenous proteins *in vitro*.

Transglutaminases are a family of enzymes characterized by their ability to catalyse formation of a covalent ε-(γ-glutamyl) lysine isopeptide bond between protein substrates. Transglutaminase 2 (TG2) is a well-studied human transglutaminase with potential to provide unique functionality to the protein-protein coupling toolbox. The moderate specificity of TG2 for glutamine donors and low specificity for lysine donors presents the opportunity for covalent decoration of endogenous protein surfaces. This thesis aimed to explore the engineering of TG2 to enhance protein-protein conjugation.

In Chapter 3 I demonstrated that TG2 is a substrate for itself and forms high molecular-weight covalent multimers when activated. A truncated variant of TG2 ($TG2_{465}$) retained transamidase activity while self-aggregation was abolished. I also demonstrated that TG2 interacts robustly with commonly used scaffold proteins and the cell surface. I determined kinetics of transamidase catalysis in different activation states and generated a variant with prolonged activity in oxidising conditions. I introduced glutamine- and lysine-bearing peptide tags to model substrates and demonstrated that these tags impart TG2-reactivity.

In Chapter 4 I investigated SnoopLigase, a protein-protein conjugation system derived from RrgA of *Streptococcus pneumoniae.* SnoopLigase requires very stringent reaction

conditions, which limits application. I demonstrated considerable improvement in reaction efficiency for the SnoopLigase2 system in several common buffers and determined the contribution of each component to the overall increase in performance. I also found that SnoopLigase2 has a strong tolerance to temperature and pH, achieving conjugation in conditions that have not been demonstrated with any competing technologies. These findings open up new possibilities for protein-protein conjugation.

In Chapter 5 I established protocols for efficient assembly of TG2:cargo conjugates using SnoopLigase2. I demonstrated TG2-mediated retention of TG2:cargo constructs to protein surfaces via non-covalent interactions. I generated a TG2:TGFα construct and achieved enhanced duration of TGFα activity due to TG2-mediated localisation at the cell surface.

In this thesis, I have revealed properties of TG2 and SnoopLigase2 that offer interesting functionalities for the synthetic biologist and may be useful for molecular assembly in tissue engineering and bio-photonic applications.

Authorship Declaration

I declare that this thesis is entirely my own work, except for the following:

Figure 4.4B RapidFire Mass Spectrometry experiment was performed by Dr. Anthony Tumber. Data were analysed by Dominic Paul Wood. Figure was created by Dominic Paul Wood.

Appendix Figure 1 SEC-MALS experiment was performed by Dr. David Staunton. Data were analysed by Dr. David Staunton. Figure was created by Dr. David Staunton.

This thesis has not been submitted for any other degree or to any other institution.

The Investigation which is the subject of this Thesis was carried out under the terms of QinetiQ Agreement No: CON11511

Acknowledgements

I would like to thank my primary supervisor, Professor Mark Howarth, for invaluable guidance and excellent feedback through my studies. I would also like to thank my secondary supervisor, Professor Elena Seiradake, for her generosity and kindness. Thank you to the Howarth Lab team for both technical and emotional support throughout my DPhil. In particular, I would like to thank Dr. Anthony Keeble for providing general training and offering technical advice on the SnoopLigase system.

I am grateful to Prof. Sajad Haq and Dr. Nicolette Moreau for giving up their time for valuable discussions on applications of protein conjugation technologies. I am also grateful to the Interdisciplinary Bioscience DTP for the support and training provided, and to both the EPSRC and QinetiQ for funding my studies.

Finally, I would like to thank my wife for her endless support through challenging times.

Abbreviations

BSA	Bovine Serum Albumin
DTT	Dithiothreitol
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
ERK	Extracellular signal-regulated kinase
FXIII	Factor XIII
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GFP	Green fluorescent protein
gpTG	Guinea pig transglutaminase
Grb2	Growth factor receptor-bound protein 2
GST	Glutathione-S-transferase
HBS	HEPES-buffered saline
HEPES	N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid
HRP	Horseradish peroxidase
HSA	Human serum albumin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LACE	Lysine acylation using conjugation enzymes
LB	Luria Bertani
MAPK	Mitogen-activated protein kinase
MBP	Maltose binding protein
mClover-Q	mClover with a C-terminal motif of 2×(GSS)TVQQEL that acts as an acyl donor for TG2-mediated transamidation
mTG	microbial transglutaminase

Ni-NTA	Nickel-nitrilotriacetic acid
PBS	Phosphate-buffered saline
PAGE	Polyacrylamide gel electrophoresis
PAL	Peptidyl asparaginyl ligase
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PI3K	Phosphoinositide 3-kinase
PLC	Phospholipase C
PMSF	Phenylmethylsulfonylfluoride
SDS	Sodium dodecyl sulfate
SEC	Size-exclusion chromatography
sfGFP	Superfolder green fluorescent protein
SPM	Self-processing module
SrtA	Sortase A
STAT1	Signal transducer and activator of transcription 1
SUMO	Small ubiquitin modifier
TBS	Tris-buffered saline
TG	Transglutaminase
TG2	Transglutaminase 2
TG2 FN	TG2 with H30A, R116A and H134A mutations. These mutations decrease affinity for fibronectin
TG2x	TG2 construct truncated to residues 1-465, includes the C230S mutation
$TG2_{465}$	TG2 construct truncated to residues 1-465
$TG2_{465}$ FN	TG2 construct truncated to residues 1-465, includes H30A, R116A, H134A mutations. These mutations decrease affinity for fibronectin

$TG2_{465}$ FN-K	TG2 construct truncated to residues 1-465, with a C-terminal motif of 2×(GSS)GQDPVK that acts as an acyl acceptor for TG2-mediated transamidation
$TG2_{465}$ FN-K2	TG2 construct truncated to residues 1-465, with a C-terminal motif of 5×(GSS)GQDPVK that acts as an acyl donor for TG2-mediated transamidation
$TG2_{465}$ FN-Q	TG2 construct truncated to residues 1-465, with a C-terminal motif of 2×(GSS)TVQQEL that acts as an acyl donor for TG2-mediated transamidation
TGFα	Transforming growth factor alpha
TGFβ	Transforming growth factor beta
TMB	3,3',5,5'-Tetramethylbenzidine
Ub	Ubiquitin
Ub(K0)	Ubiquitin with all lysine residues replaced with arginine
Ub(K0)-Q	Ubiquitin with all lysine residues replaced with arginine, with a C-terminal motif of 2×(GSS)TVQQEL that acts as an acyl donor for TG2-mediated transamidation

Contents

Abstract **2**

Authorship Declaration **4**

Acknowledgements **4**

Abbreviations **5**

Contents **8**

List of figures **12**

Chapter 1 – Introduction **14**

1.1 The Protein-Protein Coupling Toolbox 15

1.1.1. Chemical Methods 15

1.1.2. Peptide/Protein Methods 19

1.1.2.1. Tag-based Methods 20

1.1.2.2. Enzyme-based Methods 22

1.1.3. Gaps in the Protein-Protein Coupling Toolbox 25

1.1.3.1. *In vivo* protein-protein conjugation in the extracellular space 27

1.1.3.2. Fabrication of Precisely-Ordered Endogenous Protein Scaffolds 28

1.2. Transglutaminase 2 for Protein-Protein Coupling 29

1.2.1. Structure of TG2 31

1.2.2. Reaction Mechanism 33

1.2.3. Substrate Scope 34

1.2.4. Regulation of Transamidase Activity 36

1.2.5. Physiological Roles 37

1.2.5.1. Physiological Roles of Intracellular TG2 37

1.2.5.2. Physiological Roles of Extracellular TG2 39

1.3. Thesis Objectives 40

Chapter 2 - Methods **42**

2.1. Molecular Cloning 42

2.1.1. Polymerase Chain Reaction (PCR) 42

2.1.2. Agarose Gel Electrophoresis and DNA Purification 42

2.1.3. Gibson Assembly 42

2.1.4. Plasmid Amplification and Preparation 43
2.1.5. Plasmid Sequence Verification 43
2.1.6. Description of plasmids generated 43
2.2. Protein Expression and Purification 46
2.2.1. Transformation/Transfection 46
2.2.2. Growth and Expression Conditions 47
2.2.3. Purification of Soluble Proteins 47
2.2.4. Purification of TGFα from Inclusion Bodies 48
2.2.5. Refolding of TGFα constructs 49
2.2.6. Protein Dialysis 49
2.3. Protein Quantification 49
2.3.1. Nanodrop 49
2.3.2. BCA assay 50
2.3.3. Gel Densitometry 50
2.4. Protein Analysis 50
2.4.1. SDS-PAGE 50
2.4.2. Western blot 50
2.4.3. Mass Spectrometry 51
2.5. Protein Conjugation Assays 52
2.5.1. Biotin-cadaverine transamidation assay 52
2.5.2. SnoopLigase-mediated conjugation of TG2 to cargo 53
2.5.3. SnoopLigase tag/set comparison assays 53
2.5.4. SnoopLigase reaction condition validation assays 53
2.5.5. SnoopTag/Catcher comparison assays 54
2.5.7. Snooptag/Catcher reaction condition validation assays 55
2.6. SnoopLigase product separation 55
2.7. Plate binding assays 56
2.7.1. ECM protein binding assays 56
2.7.2. TG2-cargo retention assay 57
2.8. Cell surface binding assays 57
2.8.1. Transamidation assay for flow cytometry 57
2.8.2. Cell binding assay for flow cytometry 58
2.8.3. Transamidation assay for Western blot 59
2.9. Redox assays 59
2.9.1. Assay of reduced TG2 in oxidising conditions 59

2.9.2. Assay of oxidised TG2 in reducing conditions....60
2.10. TGFα cell assays....61
2.10.1. Culture and Treatment of Cells....61
2.10.2. Fluorescence microscopy....61
2.10.2. Resazurin assay....63
Chapter 3 – Engineering TG2 for Controllable Covalent Coupling....64
3.1. Introduction....64
3.2. Recombinant Expression of TG2....65
3.3. Transamidase Activity and TG2 Self-Crosslinking....69
3.4. Testing Affinity of $TG2_{465}$ for ECM Proteins and the Cell Surface....73
3.5. Engineering Redox Regulation of $TG2_{465}$....79
3.6. Peptide Tags Facilitate TG2-mediated Protein-Protein Conjugation....83
3.7. Discussion....93
Chapter 4 – Validating Second Generation RrgA-based Technologies....99
4.1. Introduction....99
4.2. Validating the SnoopLigase2 System....105
4.2.1. Comparison of First- and Second-Generation Tags....105
4.2.2. Comparison of First and Second Generation SnoopLigase Sets....106
4.2.3. Validation of SnoopLigase2 in a Range of Reaction Conditions....110
4.2.4. Examining the Mechanism for Enhanced Reactivity of SnoopLigase2....112
4.3. Validating the SnoopTag2/Catcher2 System....114
4.3.1. Comparison of First and Second Generation SnoopTag/Catcher Sets....114
4.3.2. Validation of SnoopTag2/Catcher2 in a Range of Reaction Conditions....115
4.4. Discussion....117
Chapter 5 – Decorating Endogenous Proteins with TG2....120
5.1. Introduction....120
5.2. Modular Assembly of Cargo to TG2x....121
5.2.1. SnoopLigase2-mediated Assembly....121
5.2.3. TG2-mediated Assembly....127
5.3. Retention of $TG2_{465}$-cargo in Protein Scaffolds and at the cell surface....131
5.4. Enhanced Activity of a Growth Factor by TG2-mediated Retention....133
5.5. Discussion....142
Chapter 6 – Summary and General Discussion....146

References **156**

Appendix……..**181**

List of figures

Chapter 1 – Introduction

Figure 1.1 – Chemical methods of protein-protein conjugation……. 18

Figure 1.2 – Protein-based protein-protein conjugation methods……... 19

Figure 1.3 - Similarities and differences within the transglutaminase family…… 30

Figure 1.3 – The structure of TG2……..……..... 32

Figure 1.5 – The transamidation reaction mechanism……..... 33

Figure 1.6 – The role of the catalytic triad in transamidation……. 34

Figure 1.7 - Regulation of TG2 transamidation…….. 37

Chapter 3- Engineering TG2 for Controllable Covalent Coupling

Figure 3.1 – Expression and purification of recombinant TG2 in E. coli with an N-terminal 6 × His tag…….. 66

Figure 3.2 – Premature termination of translation of TG2…… 68

Figure 3.3 – Measurement of TG2 transamidase activity using cadaverine……. 70

Figure 3.4 – C-terminal truncation of TG2…….…….. 71

Figure 3.5 – C-terminal truncation of TG2 abolishes self-crosslinking…… 72

Figure 3.6 - Non-covalent interactions between TG2 and fibronectin…… 75

Figure 3.7 – Non-covalent interactions between TG2 and components of the extracellular space……... 78

Figure 3.8 – Exploring redox regulation of $TG2_{465}$ and variants……..... 82

Figure 3.9 – Using a glutamine peptide tag to direct conjugation of a protein to an endogenous protein scaffold…… 83

Figure 3.10 - Introduction of Qtag to mClover to facilitate TG2-mediated conjugation…….. 85

Figure 3.11 – TG2-mediated conjugation of mClover-Q to the cell surface……..… 87

Figure 3.12 – Transamidation at the cell surface……...……. 89

Figure 3.13 – The anchoring method of protein conjugation……..……..... 90

Figure 3.14 – Introduction of TG2-reactive tags to TG2…… 92

Chapter 4 – Validating Second Generation RrgA-based Technologies

Figure 4.1 – TG2 as a tool for decoration of endogenous proteins……. 100

Figure 4.2 – RrgA-based systems for protein-protein ligation…… 102

Figure 4.3 – Second generation RrgA-based systems……. 104

Figure 4.4 – Mass spectrometry of SnoopLigase2-mediated ligation……. 105

Figure 4.5 – Comparison of reactivity for first and second generation SnoopLigase tags…….. 106

Figure 4.6 – Comparison of reactivity for first and second generation SnoopLigase sets……... 107

Figure 4.7 – Reactivity of original versus modified tags for SnoopLigase2 reaction. 109

Figure 4.8 – SnoopLigase2 reacts in a wide range of conditions……..... 111

Figure 4.9 – SnoopLigase2 is stably anchored to its reaction product…...……..... 113

Figure 4.10 – A comparison between first and second generation SnoopTag/Catcher sets…… 114

Figure 4.11 - SnoopTag2/Catcher2 reacts in a wide range of conditions…… 116

Chapter 5 – Decorating Endogenous Proteins with TG2

Figure 5.1 – TG2-mediated decoration of the cell surface…… 120

Figure 5.2 – Modular coupling of sfGFP to TG2x by SnoopLigase2……... 122

Figure 5.3 – Efficient SnoopLigase2-mediated ligation with the tags at either terminus or an internal site……..……. 124

Figure 5.4 – Preparation of TG2x:sfGFP ligation product……... 126

Figure 5.5 – Transamidation of a model substrate to $TG2_{465}$…… 128

Figure 5.6 – Testing transamidation of isopeptidase-impaired mutant $TG2_{465}$ W278F……. 129

Figure 5.7 – Testing transamidation of LysTag-linked $TG2_{465}$…… 130

Figure 5.8 – Demonstrating robust retention of a TG2:cargo construct in a fibronectin scaffold…….. 131

Figure 5.9 – Demonstrating robust retention of a TG2:cargo construct at the cell surface……. 132

Figure 5.10 - Purification and refolding of SnoopTag2-TGFα……... 134

Figure 5.11 – TG2x can be efficiently connected to TGFα by ligation but not genetic fusion……... 136

Figure 5.12 – TGFα-induced cell proliferation is enhanced by TG2-mediated anchoring to the cell surface……..... 138

Figure 5.13 – TGFα-induced cell morphology changes are enhanced by TG2-mediated anchoring to the cell surface…….…… 140

Figure 5.14 – TGFα-induced changes in EMT marker expression are enhanced by TG2-mediated anchoring to the cell surface…… 141

Chapter 1 – Introduction

The ability to achieve site-specific conjugation of proteins into precise architectures is highly valuable for the synthetic biologist. Protein assembly has been used to functionalise protein scaffolds (Wang et al., 2017a), lock together ligands and receptors at the cell surface (Scheu et al., 2021), construct virus-like particles (Patterson et al., 2017), assemble enzyme complexes (Stern et al., 2016), and build multivalent ligands (Veggiani et al., 2016). The covalent bond is the strongest form of protein-protein interaction and the lack of dissociation of the bonded complex presents the potential for interactions with infinite affinity (Chmura et al., 2001) (Keeble et al., 2019). The advantages of such robust protein-protein interaction have been utilised in nature for a variety of biological processes (Table 1.1).

Table 1.1 Examples of covalent conjugation in nature

Enzyme	Bond	Role	Reference
Ubiquitin	Isopeptide	Labelling proteins for degradation	(Liu et al., 2020)
Sortase	Peptide	Attaching proteins to the cell wall of Gram-positive bacteria	(Marraffini et al., 2006)
FXIII	Isopeptide	Blood coagulation	(Schmitz et al., 2020)
Microbial TG	Isopeptide	Unknown	(Strop, 2014)
PALs	Peptide	Peptide cyclization in plants	(Barber et al., 2013)

Decades of research has produced a range of synthetic tools for covalent assembly of proteins both *in vitro* and *in vivo* (Banerjee and Howarth, 2018), (Taylor et al., 2022). Synthetic biologists can utilise these tools to permanently combine two or more protein molecules to enhance activity and/or extend functionality.

In this chapter I will summarise the available methods of protein-protein covalent assembly, focussing on the mechanism of action and how this mechanism determines the scope of application. I will then contemplate the protein-protein coupling toolbox as a whole and discuss unmet needs that can be addressed by introducing novel technologies to the toolbox.

Finally, I will introduce human transglutaminase 2 (TG2) and consider its potential as a novel protein-protein coupling tool.

1.1 The Protein-Protein Coupling Toolbox

1.1.1. Chemical Methods

Small molecule crosslinkers are available that form covalent bonds with specific functional groups of proteins. Crosslinkers such as genipin and glutaraldehyde react particularly with the primary amines of lysine sidechains and the N-terminus (Sung et al., 1998) (Migneault et al., 2004). Carbodiimides react with the carboxylate groups of aspartate, glutamate, and the C-terminus (Jayachandran et al., 2022). *Group-specific* crosslinkers such as these have been widely used in materials science for altering the mechanical properties of collagen-rich materials (Adamiak and Sionkowska, 2020); applications include fabricating tissue scaffolds (Výborný et al., 2019) and tanning leather (Ding et al., 2020). Primary amines and carboxylate groups are relatively abundant in the proteome (Brooks et al., 2002) (Chu et al., 2018), so crosslinking of proteins rich in these functional groups results in formation of many intermolecular and intramolecular covalent bonds. While formation of many crosslinks is beneficial for improving the rigidity of collagen, and *Group-specific* crosslinkers have been used to produce functional enzyme aggregates (Broun, 1976); this heavy-handed approach is not suitable for directing precise assembly of proteins into orientation-specific multimers. Treatment of proteins with crosslinkers that target primary amines and carboxylates often leads to formation of large covalent aggregates (Clyne et al., 1973) (Karpovsky et al., 1984) and can alter the physicochemical properties of the protein (Wakankar et al., 2010). In one case, orientation-specific conjugation between two single-chain variable fragments was achieved by substituting lysine residues to leave only the desired lysine conjugation sites to create a bispecific assembly (Ueda et al., 2020). However, amino acid substitution is far from an ideal solution because it requires non-trivial protein engineering. Furthermore, replacing lysine

residues may alter the physiochemical properties of the protein; potentially affecting structure, solubility, and activity. Primary amine and carboxylate crosslinkers are best suited to applications where forming many covalent bonds is advantageous and protein activity does not need to be maintained, such as capturing protein conformations (Schumacher et al., 1995) (Sinz, 2017) and protein-protein interactions (Piersimoni et al., 2022).

Thiol conjugation is a *residue-specific* method of covalent coupling which employs functional groups that react with the sulfhydryl group of cysteine, such as maleimide (Smith et al., 2010). The nucleophilic thiol group attacks the carbons adjacent to the maleimide double bond to form a stable thioether bond (Hermanson, 2013). Due to the relative rarity of cysteines (Brooks et al., 2002) (Chu et al., 2018), thiol conjugation is more widely applicable for site-specific conjugation than methods that utilise primary amines or carboxylate groups. Site-specific conjugation is important for production of precise dimer and multimeric architectures where the orientation of the components is important for the function of the assembled product. For example, thiol conjugation has been used to produce an immunotoxin (Betting et al., 2008) where the site of conjugation is important in maintaining the affinity of the antibody. However, thiol conjugation is not without limitations. Disulfide bonds often play important roles in protein structure and function (Bechtel and Weerapana, 2017) and disruption of disulfide bonds by reducing and conjugating cysteines may negatively impact the conjugated protein. Thiol conjugation is thus limited in its application to proteins which lack important disulfide bonds. Another consideration is that the conjugate is susceptible to cleavage by competitive thiols, although this has proven desirable for induced cleavage of antibody-drug conjugates (Saito et al., 2003).

Genetic code expansion allows the introduction of unnatural amino acids with novel functional groups at specific sites in a protein (Ryu and Schultz, 2006). Such groups include photo-reactive and proximity-enabled reactive groups, which form covalent bonds with nearby

residues to capture interacting partners in live cells (Mishra et al., 2020) (Yang et al., 2017). For example, amino acids bearing a diazirine can be photoactivated with ~350 nm light to create reactive carbene intermediates (Dziuba et al., 2020). These intermediates then form covalent bonds via addition reactions with amino acid sidechains or the peptide backbone. The ability to direct conjugation between an engineered protein of interest and any endogenous binding partner is useful for capturing protein-protein interactions in cells. However, genetic manipulation of expression hosts and synthesis of unnatural amino acids are non-trivial techniques and proximity-based conjugation is not suited to directing assembly of precise multi-protein complexes.

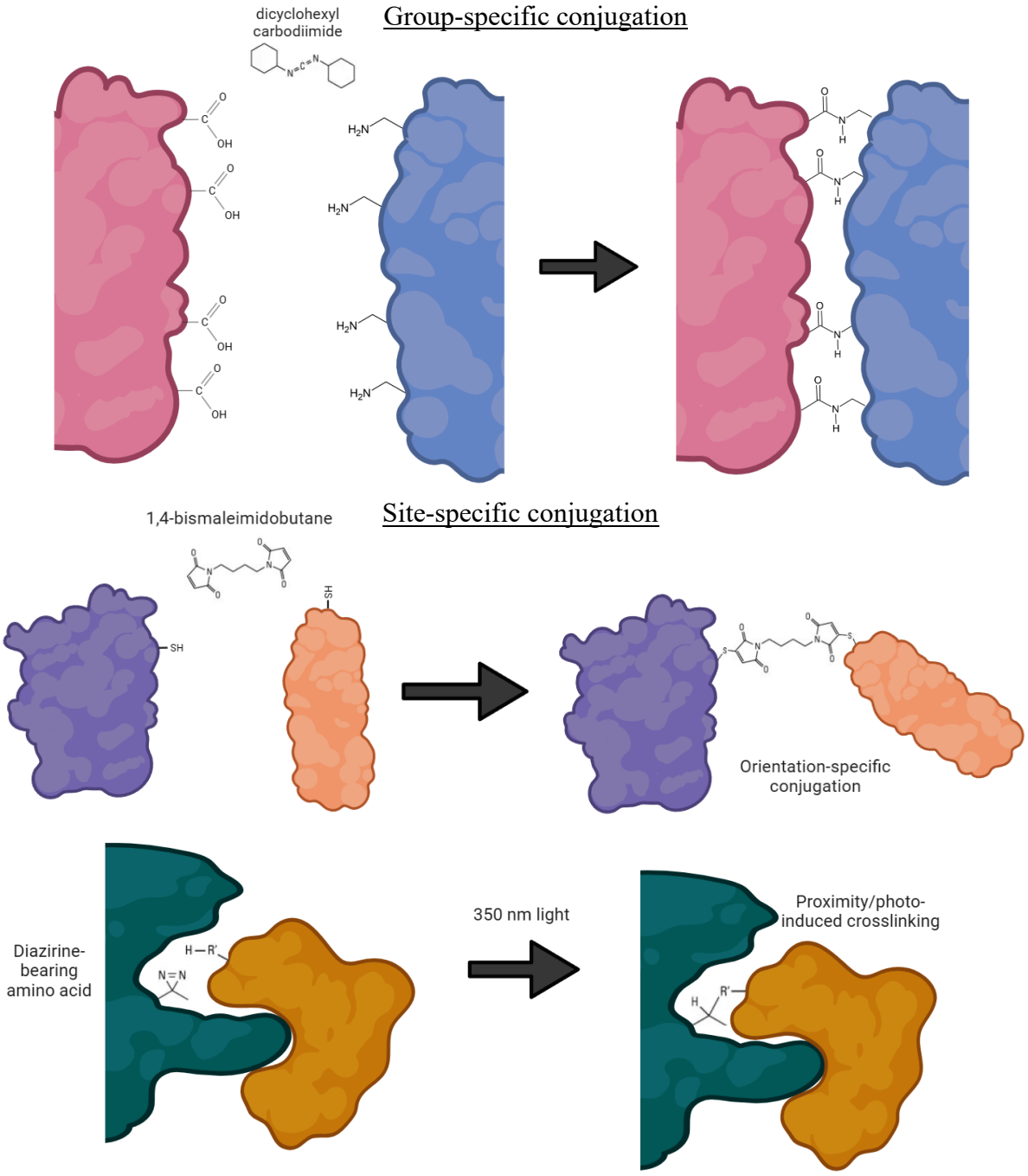


Figure 1.1. Chemical methods of protein-protein conjugation. Group-specific crosslinkers usually result in formation of many intramolecular and intermolecular crosslinks. Site-specific crosslinking methods allow for more precise conjugation. Thiol conjugation allows for orientation-specific assembly, while genetic code expansion enables photo-inducible conjugation of an unnatural amino acid to any nearby reactive residue to capture protein-protein interactions.

1.1.2. Peptide/Protein Methods

Like chemical crosslinkers, some protein-based methods facilitate *group-specific* or *residue-specific* conjugation which similarly limits application. However, some protein-based methods specifically conjugate based on amino acid motifs of several residues. These *motif-specific* methods allow much more precise control of protein-protein assembly.

Protein-based methods can be split into 'tag-based' methods and 'enzyme-based' methods (Figure 1.2), with a few exceptions. Tag-based methods conjugate highly specific amino acid motifs, which allows for precise orientation-specific assembly and conjugation in complex environments. However, substrates must be engineered to introduce the highly specific motifs preventing conjugation of endogenous proteins. Enzyme-based methods typically recognise smaller and simpler motifs, and as such require smaller tags or in some cases no tag at all. Tag-free conjugation allows conjugation of endogenous proteins. However, some enzymatic methods are *group-specific* rather than *residue-specific* which can hinder application both in a cellular environment and in complex *in vitro* protein assembly.

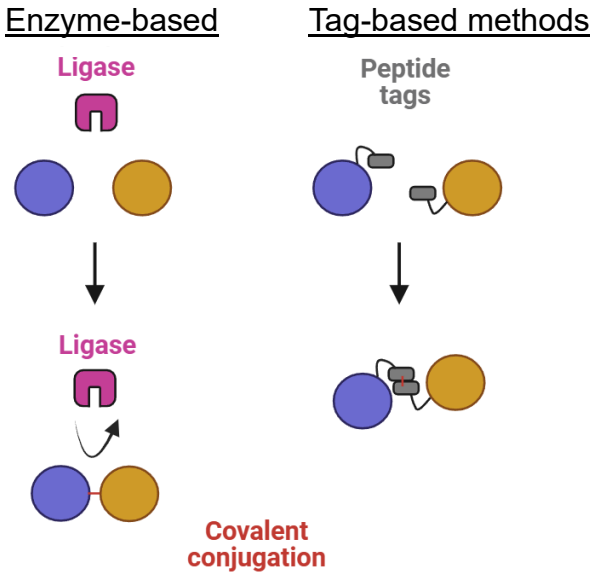


Figure 1.2. Protein-based protein-protein conjugation methods. Enzyme-based methods involve three components. A protein ligase catalyses formation of a covalent bond between preferred motifs in two substrate proteins. Upon formation of the product, the ligase regenerates to catalyse further reactions. In contrast, tag-based methods rely upon covalent bond formation between two peptide tags which are fused to the target proteins. Conjugation is catalysed by the tag itself rather than a separate ligase element.

1.1.2.1. Tag-based Methods

Inteins are protein motifs with the ability to catalyse their own removal from a protein, resulting in formation of a peptide bond between the flanking protein elements (the exteins). In most intein systems, the reaction is driven by two thiol-containing residues and an aspartate. One thiol is at the N-terminus of the intein, the other thiol is at the N-terminus of the C-extein, and the aspartate is at the C-terminus of the intein. Inteins can be split into an N-intein tag and a C-intein tag, each of which are genetically fused to protein exteins of interest (Wang et al., 2022). Upon mixing, the two split intein elements interact to reform the intein and perform ligation of the two extein elements. An attractive feature of split inteins is that the tag sequences are completely excised during conjugation, resulting in scarless ligation. However, solubility issues are often encountered when split intein elements are fused to proteins for conjugation (Pirzer et al., 2018) (Pinto et al., 2020) (Gramespacher et al., 2018). Furthermore, the reaction rate of a split intein is highly dependent on the sequence context surrounding the splice site (Pinto et al., 2020). Thus, a given split intein will work with some exteins and not others, resulting in a lack of modularity. Another consideration is the dependence of most split inteins on thiols, which are not suitable for conjugation in oxidising conditions. However, there are thiol-free split inteins available (Bhagawati et al., 2019). Despite being significantly hampered by the above issues, split inteins have been used for assembly of simple protein conjugates (Han et al., 2017), protein cyclisation (Evans et al., 2000), and cell surface conjugation (Dhar and D. Mootz, 2011).

In contrast to split inteins, the SpyTag/Catcher system has demonstrated modularity and no solubility issues. The SpyTag/Catcher system was created by splitting the FbaB protein of *Streptococcus pyogenes* into a short peptide (SpyTag) and a larger protein element (SpyCatcher) (Zakeri et al., 2012). SpyTag forms a spontaneous isopeptide bond with the 12.5 kDa SpyCatcher protein, between a specific aspartate of the SpyTag and a specific lysine of

the SpyCatcher. The latest iteration of this system has a rate constant of $k = 5.5 \pm 0.6 \times 10^5\ M^{-1}\ s^{-1}$, close to the diffusion limit for protein-protein interactions (Keeble et al., 2019). This rapid system can achieve >90% coupling at 10 nM in 15 min. The SpyTag/Catcher system has been used for *in vitro* assembly (Brune et al., 2016), enzyme cyclization (Schoene et al., 2016), and cell surface conjugation (Keeble et al., 2019).

SnoopTag/Catcher is a similar system to SpyTag/Catcher but derived from the RrgA protein of *Streptococcus pneumoniae* (Veggiani et al., 2016). The two tags form a spontaneous isopeptide bond between a specific lysine of the SnoopTag and a specific asparagine of the SnoopCatcher, which is promoted by an opposing glutamate in the SnoopCatcher. SnoopTag/Catcher is orthogonal to SpyTag/Catcher system and the two have been combined to produce complex multi-protein assemblies *in vitro* (Veggiani et al., 2016).

SnoopLigase is a three-component system which was engineered from the same domain of RrgA as SnoopTag/Catcher (Buldun et al., 2018). In this case the domain was split three ways to generate SnoopTag (containing reactive lysine), DogTag (containing reactive asparagine) and SnoopLigase (containing catalytic glutamate). The SnoopLigase system is a tag-based method, requiring all substrates to be tagged and performing a single reaction with no turnover of the ligase. SnoopLigase has been used for *in vitro* assembly of virus-like particles (Andersson et al., 2019) and bispecific nanobodies (Chen et al., 2022), and for cell surface conjugation (Keeble et al., 2023). Despite finding successful application, the SnoopLigase system has some limitations. SnoopLigase requires very stringent reaction conditions to remain soluble, and micromolar concentrations of components are required along with a substantial excess of one of the two substrates to achieve efficient ligation (Buldun et al., 2018).

NeissLock is an interesting system which is not tag-based nor traditionally enzymatic, but adopts an ‘anchoring’ approach. Derived from the FrpC protein of *Neisseria meningitidis,*

NeissLock employs a self-processing module (SPM) which has calcium-dependent autoproteolytic activity (Scheu et al., 2021). Upon calcium activation, the SPM cleaves itself at an aspartate-proline dipeptide, resulting in generation of an aspartic anhydride. This anhydride can react with a nearby nucleophile, resulting in covalent linkage. SPM can be fused to the C-terminus of an affinity protein and delivered to its ligand, where SPM covalently anchors with only a single amino acid scar. The NeissLock system has very high specificity and has been used for conjugation to endogenous proteins at the cell surface (Scheu et al., 2021).

1.1.2.2 Enzyme-based Methods

Sortase A (SrtA) is a cysteine transpeptidase of *Staphylococcus aureus* (Spirig et al., 2011). SrtA cleaves the C-terminal amino acid motif LPXTG between the threonine and glycine, and subsequently catalyses formation of a peptide bond between the threonine and an N-terminal glycine (Mao et al., 2004). The catalytic activity of SrtA is dependent upon calcium ions binding near the active site. Calcium-dependent activity may be disadvantageous for applications involving calcium-sensitive substrates. SrtA suffers from low conjugation efficiency due to the competing hydrolysis reaction and highly reversible nature of the reaction (Morgan et al., 2022). Despite having a k_{cat} of 1.5 ± 0.2 s^{-1}, the k_{cat}/K_M is only 200 M^{-1} s^{-1} due to low millimolar K_M for the LPXTG tag (Chen et al., 2011). As a result, SrtA reactions require micromolar enzyme concentrations to drive reactions near to completion (Morgan et al., 2022). Engineered variants have been created with considerably improved activity of 2.8×10^4 M^{-1} s^{-1} (Chen et al., 2011). An advantage of SrtA over other enzyme methods is that specificity for the peptide motif is very stringent, decreasing the possibility of off-target conjugation. Sortases have been used *in vitro* for assembly of conjugates (Popp and Ploegh, 2011), scaffold functionalisation (Cambria et al., 2015), and enzyme immobilisation (Fauser et al., 2020). Sortases have also been applied for cell surface conjugation (Kumari et al., 2021).

Peptidyl asparaginyl ligases (PALs) are a subtype of cysteine proteases that perform ligation, which are found in plants (Hemu et al., 2023). PALs cleave the peptide bond C-terminally of an asparagine or aspartate and form a thioester intermediate which is then attacked by an N-terminal nucleophilic acceptor to form a new peptide bond, leaving only a single asparagine/aspartate scar (Nonis et al., 2021). Butelase-1 is an extremely fast PAL with k_{cat} of 17 s^{-1} and k_{cat}/K_M of 5.42 × 10^5 M^{-1} s^{-1} (Nguyen et al., 2014). Despite competing hydrolysis reaction and highly reversible nature of the reaction, Butelase-1 has achieved >95% ligation of a peptide substrate in 20 min at 125 nM enzyme concentration (Nguyen et al., 2014). Butelase-1 recognises an N/DHV motif and has low specificity for the N-terminal substrate (Nguyen et al., 2014). Due to low specificity, Butelase-1 is restricted mostly to simple *in vitro* ligation. Furthermore, application of Butelase-1 is hindered by the requirement for expression as zymogen and processing at low pH to produce the active form (James et al., 2019). OaAEP1 is a PAL which can be expressed in bacteria in a truncated form that does not require activation at low pH (Tang et al., 2020). Despite these restrictions, PALs have been used for enzyme circularization (Nguyen et al., 2015), cell-surface labelling (Harmand et al., 2021), and building protein polymers (Deng et al., 2019).

Subtilisin is a serine protease of *Bacillus amyloliquefaciens*, which has been engineered over several decades for peptide conjugation. Initially, the catalytic serine of Subtilisin was replaced with cysteine to produce thiol-subtilisin which favoured amide synthesis over hydrolysis (Neet and Koshland, 1966). A range of mutations have since been introduced to improve activity, improve stability, and decrease amine substrate specificity (Wu and Hilvert, 1989) (Abrahmsen et al., 1991). Omniligase-1 is a descendent of thiol-subtilisin with broad specificity for the amine donor, allowing ligation to a wide-range of N-termini (Nuijens et al., 2016). Omniligase-1 performs efficient ligation (88%) at nanomolar concentrations in 30 min (Toplak et al., 2021). Reaction efficiency is limited to ~90% due to the reversible nature of the

reaction and the competing hydrolysis reaction. Omniligase-1 has been used for peptide synthesis, peptide cyclization, and simple protein assembly (Nuijens et al., 2016) (Schmidt et al., 2017).

Lysine Acylation using Conjugating Enzymes (LACE) is a system based on ubiquitinylation. Human ubiquitin carrier protein 9 (Ubc9) forms an isopeptide bond between a C-terminal LRLRGG-thioester motif and the lysine in the IKQE amino acid motif (Hofmann et al., 2020). Ubc9 has been used to conjugate Ub-like proteins to recombinant proteins carrying an IKQE peptide tag. The LACE system is limited in its application by the requirement for a C-terminal thioester, which was produced using the GyrA intein in proof-of-concept experiments. The requirement for a C-terminal thioester on protein A and a peptide tag on protein B substrate means both proteins must be engineered and thus LACE cannot be used for conjugation of endogenous proteins. While the reaction was able to reach >90% conversion of substrate, the reaction was relatively slow and required 6 hours to reach this level of conjugation motif (Hofmann et al., 2020). LACE technology is a recent development and has only been used for proof-of-concept assembly of virus-like particles (Levasseur et al., 2022).

Transglutaminases (TGs) are a family of enzymes characterised by their ability to catalyse formation of an isopeptide bond between the γ-carboxamide of a glutamine side chain, and the ε-amino group of a lysine side chain (Lai et al., 2017). TGs prefer certain residues flanking the glutamine and lysine residues, but do not have a strict recognition motif (Esposito and Caputo, 2005) (Gundersen et al., 2014). As a result, TGs have lower specificity than other conjugation methods. Microbial transglutaminase (mTG) of *Streptomyces mobaraensis* has a k_{cat} of 1.39 s^{-1} and k_{cat}/K_M of 1.64×10^3 for a DYALQ motif (Steffen et al., 2017). mTG has been used extensively in the food industry for non-specific crosslinking of protein products (Kieliszek and Misiewicz, 2014). However, mTG is difficult to apply for site-specific conjugation due to low substrate specificity. Human TGs may have higher specificity than

microbial TGs due to the active site being situated in a deep tunnel in the core of the protein, in contrast with the shallow cleft in which the mTG active site resides. Factor XIII (FXIII) is a human transglutaminase that is used as a component of surgical adhesives (Schroth et al., 2006). Both mTG and FXIII require protease cleavage of a propeptide for activation (Zhang et al., 2008) (Gupta et al., 2016).

Oxidases such as peroxidases, tyrosinases and lysyl oxidases have also been explored as protein-protein conjugation tools. Unlike the above methods, oxidases do not directly catalyse the formation of a covalent bond between proteins. Oxidases facilitate protein-protein covalent bond formation by oxidising certain amino acids to more reactive forms, which then spontaneously react. For example, lysyl oxidases oxidise the side chain of lysine to form an aldehyde, reducing O_2 to hydrogen peroxide in the process. The aldehyde can either condense with another aldehyde to form an aldol condensation product, or it can react with an amine to form a Schiff base (Milczek, 2018). Oxidases perform *residue-specific* conjugation rather than *site-specific* conjugation and as such applications of oxidases are limited to non-specific crosslinking in food and textiles industries.

1.1.3 Gaps in the Protein-Protein Coupling Toolbox

Chemical crosslinking methods are unsuitable for some site-specific protein-protein conjugation applications. Protein-based methods have been developed that overcome the limitations of chemical methods, but each protein-based system has unique features which determine the types of applications to which it can be applied (Table 1.2). Low specificity prevents conjugation in complex environments, such as in cellular compartments. The requirement for all substrates to bear reactive tags prevents conjugation of endogenous proteins, such as conjugating proteins to the cell surface. Non-human derived systems present possible immunogenicity issues that restricts *in vivo* applications. Reaction requirements such as calcium or reducing conditions can complicate conjugation in simple *in vitro* contexts.

Table 1.2 Characteristics of protein-based protein-protein coupling tools

Method	Specificity for substrate A	Specificity for substrate B	Substrate engineering requirements	Human-derived	Reaction requirements
Split inteins	High	High	Peptide tags at termini of both substrates	No	Reducing conditions
Tag/Catcher systems	High	High	Peptide tag in substrate A, protein tag in substrate B	No	None
SnoopLigase	High	High	Peptide tags in both substrates	No	None
Butelase-1	Moderate	Low	Peptide tag at C-ter of substrate A	No	Propeptide cleavage
OaAEP1	Moderate	Low	Peptide tag at C-ter of substrate A	No	None
LACE	High	High	C-ter thioester on substrate A, peptide tag in substrate B	Yes	None
Sortase A	High	High	Peptide tag at C-ter of substrate A	No	10 mM Calcium
Omniligase-1	High	Low	Peptide tag at C-ter of substrate A	No	None
mTG	Moderate	Low	Peptide tag in substrate A	No	Propeptide cleavage
FXIII	Moderate	Low	Peptide tag in substrate A	Yes	Propeptide cleavage

1.1.3.1. *In vivo* protein-protein conjugation in the extracellular space

Despite the impressive range of protein-protein conjugation systems, there are applications for which there is not yet an appropriate tool. For *in vivo* conjugation in the extracellular space, only LACE and FXIII are human-derived. LACE is unable to conjugate endogenous proteins, and FXIII is a pro-coagulant which poses a serious risk of thrombosis if allowed to enter systemic circulation (Al-Amer, 2022).

Conjugation in the extracellular space is of interest for robust, long-term localisation of therapeutics to diseased tissues. The efficacy of some drugs has been improved by targeting the drug specifically to diseased cells and tissues (Lewis Phillips et al., 2008). Strategies for targeted delivery of therapeutics usually involve the use of high affinity ligands such as peptides (Berillo et al., 2021), antibodies (Hafeez et al., 2020) or DNA aptamers (Dassie et al., 2009). For example, antibody-drug conjugates have been extensively employed for targeting therapeutics specifically to tumours (Hafeez et al., 2020). However, an approach that covalently bonds a therapeutic molecule to the target site may provide more robust and irreversible localisation compared with non-covalent affinity-based approaches. Furthermore, if the specificity for the target surface is low then much broader decoration may be possible than is achieved by highly specific non-covalent methods.

NeissLock, OaAEP1 and Sortase have been used to covalently anchor therapeutic proteins to the cell surface (Scheu et al., 2021) (Harmand et al., 2021) (Pishesha et al., 2017). However, these methods are limited to *in vitro* cell conjugation because they contain microbially-derived elements with unknown immunogenicity. Furthermore, OaAEP1 and Sortase can only conjugate to exposed N-termini which may be limited at the cell surface and in fibrous protein scaffolds. A method of protein-protein conjugation is required that is human-derived and capable of broad conjugation of the extracellular space.

1.1.3.2. Fabrication of Precisely Ordered Endogenous Protein Scaffolds

Light manipulation is the basis of important biological functions across all kingdoms of life, including photosynthesis, radiative heating, camouflage and light detection (Johnsen, 2012) (Srinivasarao, 1999). Light manipulation in nature is achieved using either pigments or materials with an ordered nanoscale structure, referred to as photonic materials. In contrast to pigmental light manipulation, which is the result of wavelength specific light absorption, structural light manipulation is the result of wavelength specific light reflection (Land, 1972). Photonic materials have a spatially periodic refractive index, with the period near the wavelength of visible light (Pavarini et al., 2005). Examples of photonic materials in nature include the transparent human cornea (Müller et al ., 2004), the iridescent skin of the cephalopods (How et al., 2017), and the striking skin colours of many mammals and birds (Prum, 2003) (Prum et al., 2009). Some photonic materials, including the examples above, are made from proteins. Photonic materials found in nature have evolved to provide exceptional performance, and there is interest in mimicking the features of bio-photonic materials to reproduce these light-manipulating properties in synthetic materials (Zhou et al., 2018) (Zhao et al., 2019).

A key challenge for fabricating bio-photonic materials is achieving precise ordering on a nanometer scale. This has been achieved using lithography to craft photonic materials from silicon (Narasimhan et al., 2018) (Siddique et al., 2017), but these harsh conditions are not amenable to building protein-based bio-photonic structures. A tool for precise covalent conjugation of endogenous proteins may be useful for assembling protein-based bio-photonic materials.

There are multiple protein-protein conjugation methods capable of utilising untagged endogenous proteins as substrates, although none are ideal for fabrication of precisely ordered scaffolds. Sortases and PALs would only conjugate to the N-terminus, which may not be

accessible in fibrous protein scaffolds. NeissLock would require an affinity domain. mTG specificity is too low and would result in extensive cross-linking. FXIII requires activation by thrombin cleavage which is expensive and off-target cleavage may present problems for assembly. A method of protein-protein conjugation is required that can assemble endogenous proteins with simple and controllable activity.

1.2. Transglutaminase 2 for Protein-Protein Coupling

1.2.1. An introduction to Transglutaminases

Transglutaminase activity was first detected in guinea pig liver extract over 60 years ago, and the newly discovered enzyme was dubbed ‘tissue transglutaminase’ (Sarkar, Clarke and Waelsch, 1957). Enzymes with transglutaminase activity are present in all 3 phylogenetic domains (Grenard et al., 2001) (Falcone et al., 1993) (Strop, 2014). All known eukaryotic TGs are evolutionarily related (Lorand and Graham, 2003). Invertebrates have a single TG, while vertebrates have multiple TG homologs as a result of successive gene duplications (Grenard et al., 2001) (Figure 1.3A). Humans have 9 known TG homologs (Lorand and Graham, 2003). Several microbial TGs have been discovered by screening for transglutaminase activity, most of which belong to the *Streptomyces* and *Streptoverticillium* genera (Yu et al., 2008) (Ho et al., 2000) (Cui et al., 2008) (Duran et al., 1998). Microbial TGs do not share sequence homology with eukaryotic TGs (Strop, 2014) (Figure 1.3B), but share a similar active site structure due to convergent evolution (Kashiwagi et al., 2002). A large body of TG literature has accumulated since the first experiments in the 1950s, focussed predominantly on two human TGs (TG2 and FXIII) and one microbial TG from *Streptomyces mobaraensis*. TG2 shares 38% sequence homology with FXIII (Figure 1.3C).

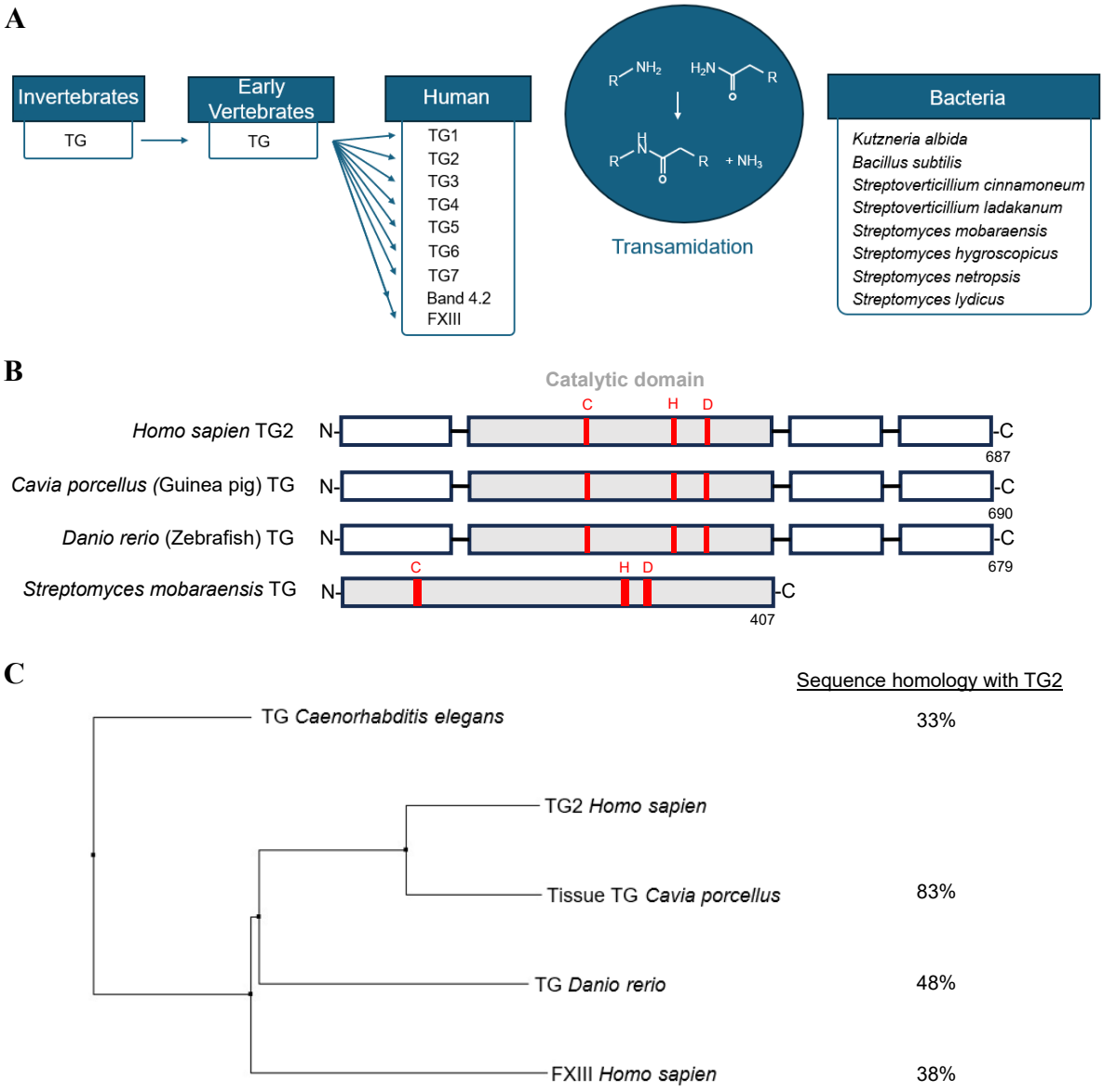


Figure 1.3. Similarities and differences within the transglutaminase family. (A) An illustration of the relatedness of proteins with confirmed transamidase activity. After initial discovery of tissue TG in *Cavia porcellus*, homologs have been discovered in invertebrates, fish, and humans. Some vertebrates, including humans, have several TG homologs due to gene duplication events in early vertebrates. No known microbial TGs share homology with eukaryotic TGs. (B) A comparison of the domain organization of eukaryotic and bacterial TGs, illustrating homology between eukaryotes. Bacterial TGs share the same catalytic triad as eukaryotes but do not share similar domain organisation. (C) Phylogenetic tree calculated from sequence alignment of eukaryotic TGs, calculated using Neighbour Joining method with BLOSUM62.

1.2.1. Structure of TG2

The structure of TG2 has been well studied, with several high-resolution crystal structures available (Jeong et al., 2020) (Liu et al., 2002) (Pinkas et al., 2007) (Figure 1.4A). TG2 forms two distinct conformations which differ in the extent to which the active site is accessible to substrates (Katt et al., 2018) (Figure 1.4B). In the closed form, the two β-barrel domains fold over the catalytic domain and the first β-barrel blocks access to the active site.

TG2 is monomeric at 4 °C to 24 °C (Pinkas et al., 2007), but forms a mixture of dimer, trimer and tetramer at 37 °C (Kim et al., 2017). The only known post-translational modification of TG2 is a transient disulfide bond which is not essential for correct protein folding (Stamnaes et al., 2010). The catalytic core of TG2 sits at the base of a hydrophobic tunnel in the catalytic domain (Keillor et al., 2014) (Figure 1.4C).

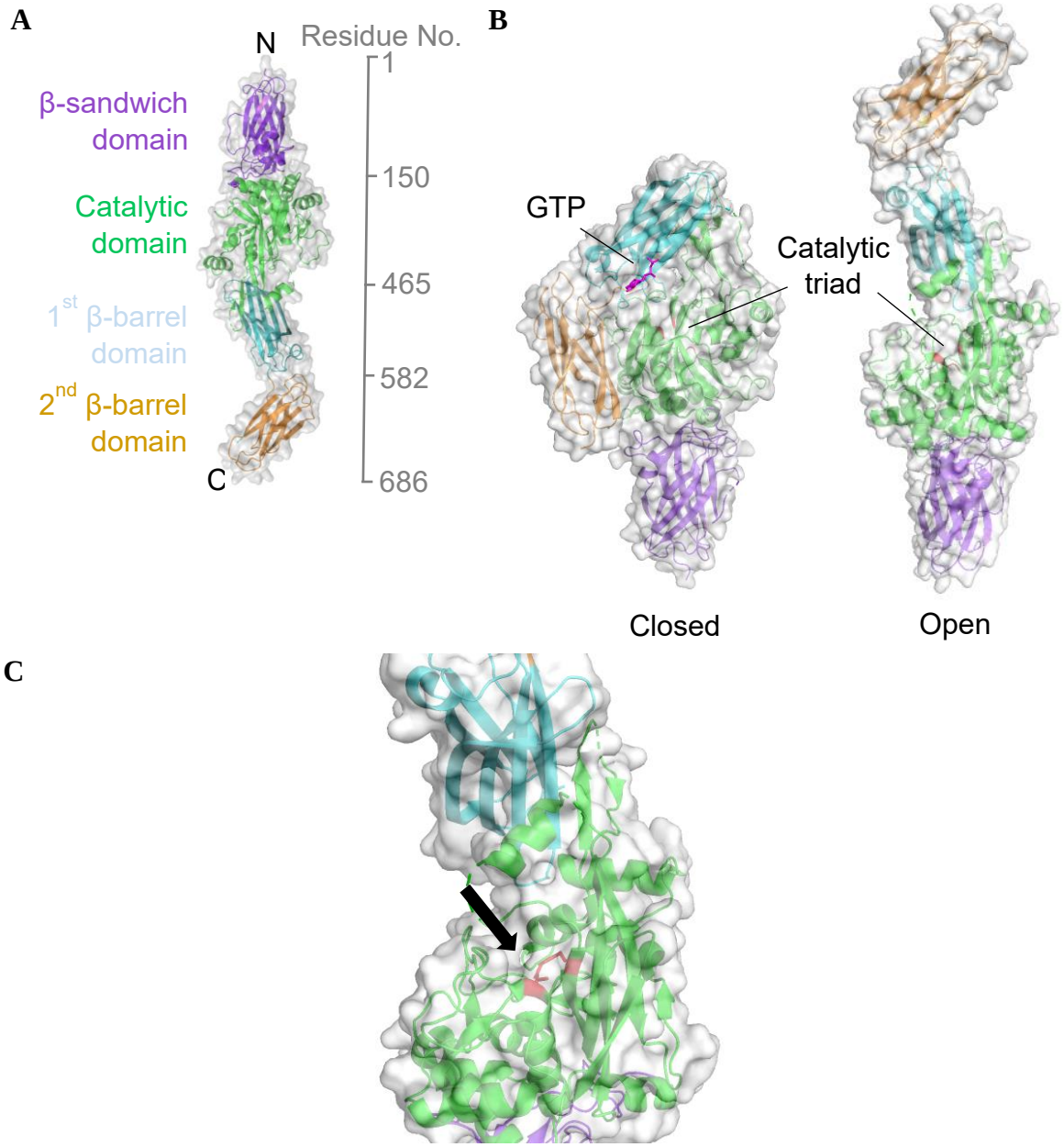


Figure 1.4. The structure of TG2. (A) Crystal structure of TG2 based on PDB 2Q3Z, with each domain coloured. (B) Comparison of the distinct closed and open conformations of TG2, based on PBD 6KZB and PDB 2Q3Z respectively. Bound GTP is shown coloured in magenta, while the catalytic triad is coloured red. (C) The catalytic site of TG2. Black arrow indicates the entrance of the active site tunnel. Catalytic residues coloured red with sidechains displayed as sticks, based on PDB 2Q3Z. Protein surface coloured light grey with cartoon of secondary structure overlaid.

1.2.2. Reaction Mechanism

The transamidation reaction occurs in two steps via a modified ping-pong mechanism whereby glutamine acts as an acyl donor and an amine as an acyl acceptor (Figure. 1.5) (Folk and Cole, 1966). If the amine is provided by a lysine sidechain, then a covalent ε-(γ-glutamyl) lysine isopeptide bond is formed. If the amine is provided by a small molecule, then the glutamine donor protein becomes aminylated. In the absence of free amine, the thioester bond can be hydrolysed, resulting in regeneration of the enzyme and deamidation of the glutamine to glutamate (Keillor et al., 2014).

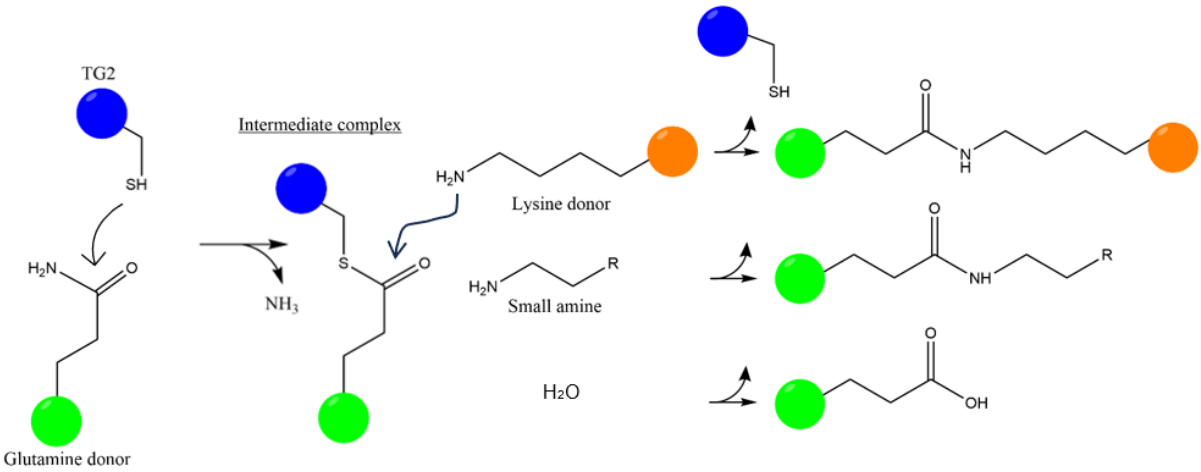


Figure 1.5. The transamidation reaction mechanism. The active site C277 attacks the side chain of a glutamine residue, forming an intermediate enzyme-substrate thioester complex and releasing ammonia. Then an amine or water attacks the thioester bond of the intermediate complex, resulting in regeneration of the free enzyme and release of the transamidation product.

The active site of TG2 forms an acid-base catalytic triad composed of cysteine, histidine and aspartate (Figure 1.6), similar to the active site of cyteine proteases. In contrast to cysteine proteases, the active site of TG2 is buried in a hydrophobic pocket in the core of the catalytic domain (Keillor et al., 2014). Water is excluded from the active site in order to favour transamidation over deamidation of glutamine.

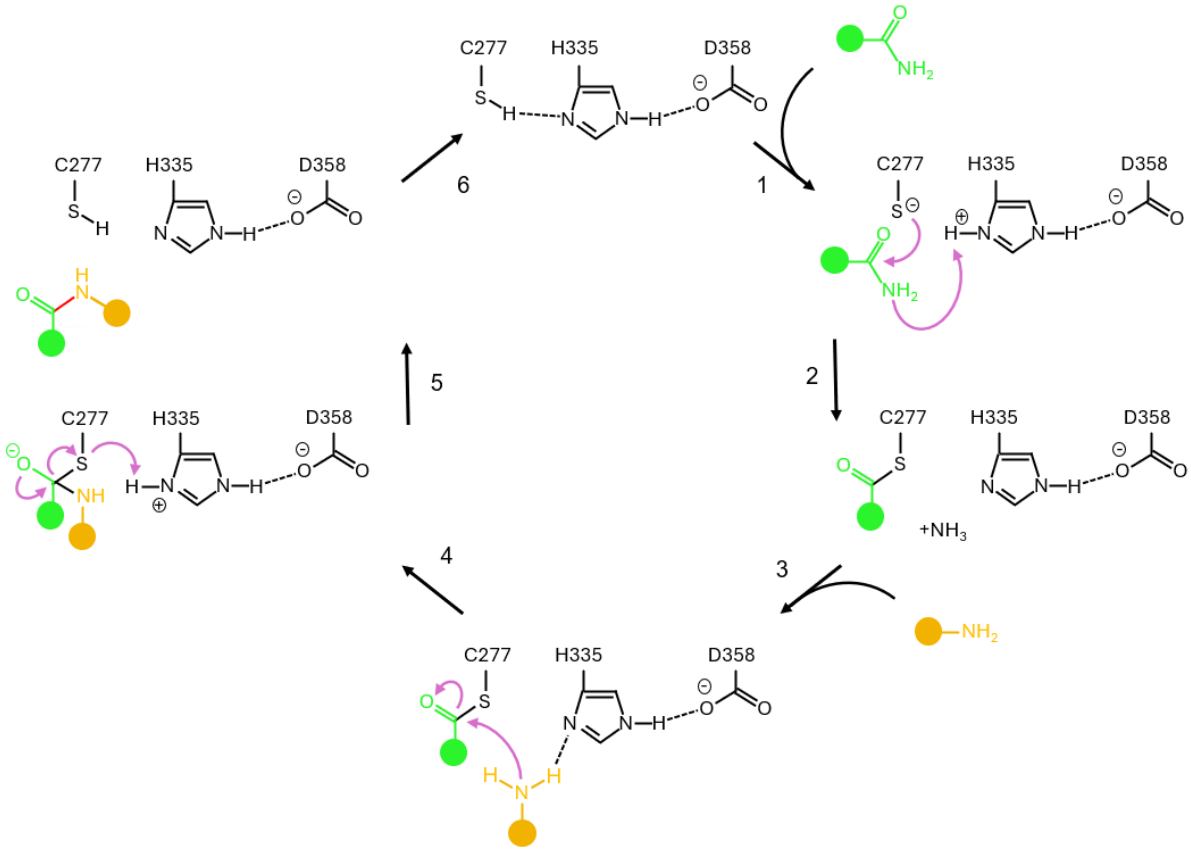


Figure 1.6. The role of the catalytic triad in transamidation. The imidazole group of histidine activates the cysteine by deprotonating the thiol. The protonated histidine is stabilised by the aspartate. The nucleophilic cysteine attacks the carbonyl of the amide substrate to form a thioester intermediate. During this process, the imidazole group is deprotonated resulting in expulsion of ammonia. An amine is then deprotonated by the imidazole group, promoting its attack on the thioester carbonyl. This results in deprotonation of the imidazole group and regeneration of the cysteine thiol.

1.2.3. Substrate Scope

Peptide libraries have been used to examine the preferred sequences for glutamine donors (Sugimura et al., 2006), (Keresztessy et al., 2006), (Ohtsuka et al., 2000). Findings demonstrated that TG2 can utilise glutamine in a broad range of sequence contexts. It has been

suggested that glutamine donors tend to follow the general motif zQxPh where z is polar, x is any residue and h is hydrophobic. However, many natural TG2 substrates do not fit these preferences (Csosz et al., 2009). Furthermore, most proteins contain multiple glutamine residues but few if any are reactive for TG2. For example, Interferon α has 12 glutamine residues and only one is a TG2 substrate (Spolaore et al., 2016). Analysis of TG2 substrates has revealed a preference for glutamine residues in disordered regions (Csosz et al., 2008), suggesting that structure has an important role in determining reactivity. The importance of substrate structure in transglutaminase activity has been demonstrated with type I collagen. This highly structured protein is a substrate for mTG in its folded form. However, after heat-treatment, the denatured substrate can be transamidated by mTG at multiple glutamine and lysine residues (Stachel et al., 2010). The importance of structural context in substrate scope is likely to be due to the narrow tunnel that substrates must occupy to fit in the catalytic site, which requires the residues surrounding the glutamine to assemble linearly (Pinkas et al., 2007).

There are less stringent sequence requirements surrounding the lysine residue, with only a mild preference for leucine at the +1 position (Grootjans, Groenen and de Jong, 1995). Despite this lack of preference, only two of twenty total lysine residues in three natural TG2 substrates act as acyl acceptors for a glutamine tripeptide (Murthy et al., 2009). Both reactive lysine residues are found in flexible regions. Analysis of other TG2 substrates has also demonstrated preference for lysine residues in disordered regions (Lorand et al., 1992). It makes sense that structural context has a key role in determining reactivity because the lysine donor must be sterically compatible with the enzyme-substrate complex to access the thioester bond.

1.2.4. Regulation of Transamidase Activity

The transamidase activity of human TGs are regulated by calcium (Király et al., 2009), but TG2 is unique in that it is also regulated by redox conditions (Stamnaes et al., 2010) (Figure 1.7A). In the closed conformation, GTP stabilises the insertion of the first loop of the first β-barrel domain into the active site (Liu, Cerione and Clardy, 2002), blocking access to the catalytic cysteine and inhibiting transamidase activity. GTP binding is disrupted by calcium, which acts as a cofactor for TG2 by inducing conformational change to the open form and exposing the active site at low mM concentrations (Achyuthan and Greenberg, 1987). TG2 has six distinct calcium binding sites (Király et al., 2009) (Figure 1.7A,B). The exact mechanism by which calcium binding induces conformational change is not known, and small molecule active site inhibitors can also stabilise the open form without calcium (Kerr et al., 2017).

Despite having an exposed active site, the calcium-bound open form of TG2 is not catalytically active in oxidising conditions due to the presence of at least one regulatory disulfide bond in the catalytic domain (Stamnaes et al., 2010). The mechanism by which this disulfide bond(s) inhibits transamidase activity is unknown. Crystal structures have captured C370 forming disulfide bonds with C230 (Jeong et al., 2020) and C371 (Pinkas et al., 2007).

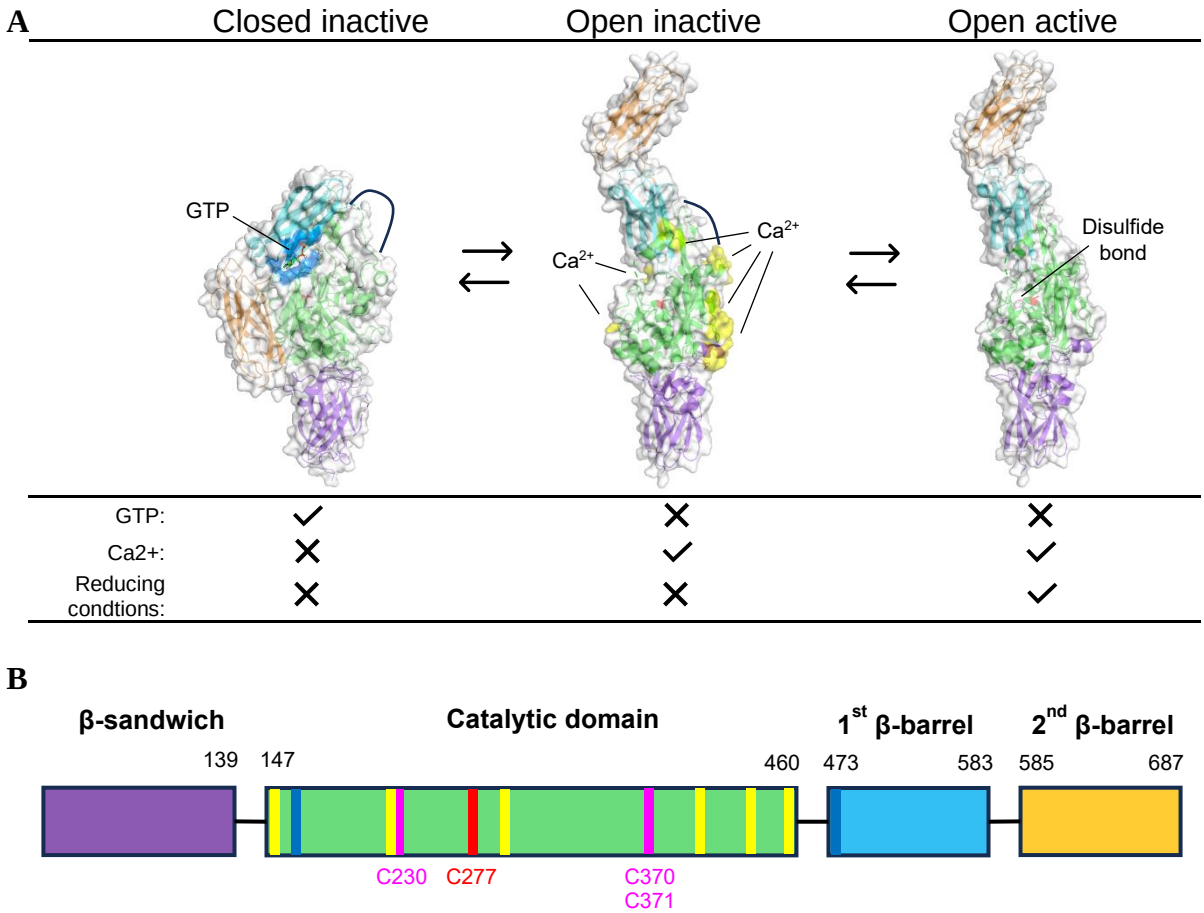


Figure 1.7. Regulation of TG2 transamidation. Calcium and reducing agents combine to activate TG2. Calcium stabilises TG2 in a conformation which exposes the active site, and reducing agents break a regulatory disulfide bond near the active site. (A) Schematics based on PDB 6KZB (closed form) and PDB 2Q3Z (open form). Catalytic cysteine coloured red with side chain displayed, GTP binding site coloured blue, calcium binding sites coloured yellow, cysteines known to form disulfide bonds coloured magenta with side chains displayed. (B) Domain organisation of TG2 with key residues highlighted for regulation of transamidase activity. Calcium binding sites indicated in yellow, GTP binding sites indicated in dark blue.

1.2.5. Physiological Roles

TG2 is present both intracellularly and in the extracellular space (Kuo et al., 2011) (Belkin, 2011). TG2 has two well studied catalytic activities (transamidase and GTPase) and multiple high affinity binding partners (Kanchan et al., 2015). Perhaps unsurprisingly then, TG2 has been implicated in a range of physiological processes.

1.2.5.1. Physiological Roles of Intracellular TG2

TGs were first identified by their transamidase activity (Sarkar, Clarke and Waelsch, 1957), and a clear role for intracellular TG2 transamidation in apoptosis has now been

established (Tatsukawa and Hitomi, 2021). Intracellular TG2 is inactive under normal conditions in all cellular compartments due to the sub-millimolar levels of calcium (Bagur and Hajnóczky, 2017). During cell stress, cytosolic calcium levels are increased (Harr and Distelhorst, 2010) and TG2 expression is upregulated (Jang et al., 2010). Increased TG2 expression leads to alteration of the redox state of mitochondria and release of cytochrome c and apoptosis-inducing factor into the cytosol (Rodolfo et al., 2004). Exactly how intracellular TG2 becomes reduced is not clear, but thioredoxin-1 is a redox cofactor protein with strong specificity for oxidised TG2 (Jin et al., 2011). During apoptosis, thioredoxin-1 may be released into the cytosol and other cellular compartments resulting in activation of intracellular TG2. Once activated, intracellular TG2 contributes to apoptosis through transamidation of key substrates. Cytosolic TG2 crosslinks cytoskeletal proteins such as actin (Nemes et al., 1997). Nuclear TG2 crosslinks transcription factors such as specificity protein 1, preventing expression of growth factors that are essential for cell survival (Tatsukawa et al., 2009).

In addition to transamidation, TG2 has GTPase activity (Mishra and Murphy, 2004). The GTPase activity of TG2 has been shown to mediate α1-adrenergic receptor stimulation of phospholipase C (PLC) (Nakaoka et al., 1994). There is evidence to suggest that TG2 may be acting as a G protein in pancreatic β cells to regulate insulin secretion. PLC is crucial for β cell function and PLC null mice exhibit glucose intolerance (Hwang et al., 2019). TG2 null mice are also glucose intolerant and show impaired insulin secretion (Bernassola et al., 2002). Furthermore, mis-sense mutations in highly conserved residues of TG2 have been found in patients with early-onset type 2 diabetes (Porzio et al., 2007). While evidence for the involvement of TG2 in insulin secretion is compelling, the relevance of GTPase activity in this process is somewhat unclear due to the observation that TG2 may be able to affect signalling by post-translational modification of G proteins. TG2-mediated aminylation of G proteins with histamine and serotonin have been demonstrated *in vitro* by TG2 (Vowinckel et al., 2012) (Dai

et al., 2011). It has been suggested that TG2-mediated aminylation of G proteins is involved in a range of biological processes including insulin secretion (Muma and Mi, 2015).

1.2.5.2. Physiological Roles of Extracellular TG2

While a clear role for intracellular transamidation has emerged, the physiological relevance of extracellular transamidation is clearly defined. Extracellular TG2 is inactive under normal physiological conditions due to the oxidising environment, and is activated upon tissue damage (Siegel et al., 2008). Exactly how TG2 becomes reduced is not fully understood, but it may be via thioredoxin-1. Macrophages exposed to pro-inflammatory stimuli release thioredoxin-1 in sufficient quantities to activate extracellular TG2 (Plugis et al., 2017). Furthermore, intravenous administration of recombinant thioredoxin-1 in mice resulted in increased transamidation in the extracellular space (Plugis et al., 2017). It has been suggested that TG2 cross-linking in the extracellular space improves the mechanical integrity and protease resistance of tissues during ECM remodelling (Aeschlimann et al., 1995). Several extracellular proteins, including components of the ECM, have been identified as substrates of TG2 using *in vitro* assays (Csosz et al., 2009). Also, TG2-mediated crosslinking of ECM protein scaffolds results in increased cell adhesion and migration *in vitro* (Van den Brûle et al., 1998) (Belkin et al., 2005). Another putative role for TG2 transamidation in the extracellular space is in functionalising the ECM with extracellular signalling proteins. TG2 has been shown to conjugate several proteins to the ECM including the latent TGF-β complex (Le et al., 2001), midkine (Mahoney et al., 2000) and protease inhibitor trappin-2 (Guyot et al., 2005).

There are also putative roles for TG2-mediated extracellular aminylation. TG2-mediated aminylation of fibrinogen and glioma cells with serotonin has been demonstrated *in vitro* (Hummerich and Schloss, 2010). It has been suggested that extracellular aminylation of the ECM with serotonin may be involved in ECM remodelling and modulation of cell behaviour (Muma and Mi, 2015).

TG2 null mice are viable with no apparent tissue abnormalities (Nanda et al., 2001), casting doubt on the physiological relevance of extracellular transamidation in tissue development. However, TG2 null mice do display impaired liver regeneration after injury , suggesting a role for TG2 in tissue repair (Nardacci et al., 2003). Hepatocytes from TG2 null mice were found to have impaired α1-adrenergic receptor signalling (Sarang et al., 2005), so the role of TG2 in tissue repair may be via G protein signalling rather than transamidation.

Other observations from TG2 null mice also point towards transamidase independent functions of TG2. The thymus of TG2 null mice appears normal until apoptosis is induced with dexamethasone, in which case the TG2 null thymus is enlarged compared to wildtype (Sarang et al., 2007). Determination of the percentage of accumulated dead cells and viable cells suggests that enlargement is due to impaired phagocytosis as well as decreased apoptosis. Integrin β3 and β5 receptors play an important role in the initiation of phagocytosis (Akakura et al., 2004). Transamidase-independent interactions have been discovered between TG2 and a range of proteins present at the cell surface including β3 integrins (Akimov et al., 2000), which may play a role in phagocytosis. TG2 has also been shown to interact with the extracellular domains of syndecan-4 (Wang et al., 2012), platelet-derived growth factor receptor (Zemskov et al., 2009) and β-catenin (Beazley et al., 2012). In cell assays, increased TG2 expression led to a transamidase-independent increased cell adhesion, migration, survival, and proliferation (Akimov and Belkin, 2001) (Song et al., 2007) (Zemskov et al., 2009). These findings suggest TG2 may have significant transamidase-independent roles in tissue repair.

6.5. Thesis Objectives

TG2 is a well-studied transglutaminase with potential to provide unique functionality as a protein-protein coupling tool. The overall aim of this thesis was to explore the potential of TG2 as a tool for protein-protein conjugation. The physiological role of TG2 in functionalising the ECM with growth factors suggests that TG2 will lend itself well to conjugating proteins to

endogenous protein scaffolds. Furthermore, the ability of TG2 to conjugate a reasonably specific glutamine motif to lysine residues of endogenous proteins with low specificity may allow for broad scaffold decoration. Decoration of protein scaffolds is of interest for cell culture, tissue engineering, and fabricating bio-photonic materials. The goal of Chapter 3 was to establish production of recombinant TG2 and use a range of assays to determine if TG2 is suitable for directing protein-protein conjugation. This included: investigating the propensity for TG2 to crosslink itself, exploring the affinity of TG2 for a range of scaffolds, determining the kinetics of redox regulation, and directing transamidation of model substrates using a glutamine motif.

Modular coupling of cargo proteins to TG2 enhances the range of proteins that can be used to decorate scaffolds, thereby increasing the utility of TG2. The goal of Chapter 4 was to validate the recently-optimised SnoopLigase and SnoopTag/Catcher technologies, with the aim of applying these technologies for modular coupling of cargo proteins to TG2. This involved using model reactions to compare the performance of first and second-generation systems, and to validate both systems in a wide range of reaction conditions.

Localisation of proteins, such as cytokines and growth factors, in scaffolds or at the cell surface can be used to prolong the functional effects of that protein at the target site. The goal of Chapter 5 was to utilise the non-covalent interactions of TG2 to enhance the functionality of other proteins by anchoring them to protein surfaces. TG2 was modified to allow SnoopLigase-mediated conjugation of TG2 to cargo, and TG2:GFP constructs were used to study retention of TG2:cargo constructs to fibronectin and the cell surface. Finally, I used a TG2:TGFα construct to perform proof-of-principle experiments to anchor a therapeutic protein to the cell surface via TG2-mediated interactions.

Chapter 2 – Methods

2.1. Molecular Cloning

2.1.1. Polymerase Chain Reaction (PCR)

Desiccated DNA oligonucleotides were purchased from Integrated DNA Technologies, re-constituted to a final concentration of 100 µM in deionised water and stored at -20 °C. PCRs were prepared at 50 µL final volume, containing 0.2 ng plasmid DNA template, 0.5 µM of each primer, and 25 µL 2× Q5 high-fidelity DNA polymerase master mix (New England Biolabs). The reaction was incubated in a C1000 Thermal Cycler (Bio-Rad). The standard protocol consisted of an initial denaturation step for 30 s at 98 °C, followed by 30 cycles of: 10 s at 98 °C, then 15 s at the desired annealing temperature, then 15 s per kb of product at 72 °C. This was followed by one final step of 2 min at 72°C.

2.1.2. Agarose Gel Electrophoresis and DNA Purification

25 µL of PCR reaction sample was mixed with 5 µL of 6× loading dye (1.2 M sucrose, 0.7 mM bromophenol blue, 0.9 mM xylene cyanol FF) and loaded on a 1% (w/v) agarose gel containing 0.01% (v/v) SybrSafe (Invitrogen). Samples were run for 40 min at 130 V in TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA, pH 8.0). Agarose gels were placed on an XcitaBlue Conversion Screen and Filter (Bio-Rad) and imaged using a Gel Imager Universal Hood II (Bio-Rad). Images were analysed using Image Lab Version 5.2.1 (Bio-Rad). For PCR product purification, PCR product bands of interest were excised and purified using the Wizard SV Gel and PCR Clean-Up Kit (Promega) following the manufacturer's protocol. DNA concentration was determined using an ND-1000 Nanodrop (NanoDrop).

2.1.3. Gibson Assembly

Gibson assembly of PCR fragments was performed as previously described (Gibson et al., 2009) using a home-made 2× Gibson Assembly Master Mix. Agarose gel-purified PCR fragments were added to 10 µL of 2× Gibson Assembly Master Mix at equimolar amounts to

give a final total DNA amount of 0.5 pmol. Deionised water was added, if required, to make the total reaction volume up to 20 µL. Samples were incubated for 1 h at 50 °C in a pre-heated C1000 Thermal Cycler (Bio-Rad).

2.1.4. Plasmid Amplification and Preparation

5 µL of Gibson Assembly reaction was added to 100 µL of heat competent NEB Turbo *E. coli* (New England Biolabs) and incubated on ice for 15 min. Heat shock was performed for 45 s in a 42 °C water bath, followed by 2 min incubation on ice. The entire sample was then added to 350 µL SOC media and incubated at 37 °C for 1 h with 200 rpm shaking. 100 µL of sample was then plated onto an LB agar plate containing the relevant antibiotic (50 µg/mL kanamycin or 100 µg/mL carbenicillin) and incubated overnight at 37 °C. The following day individual colonies were isolated and used to inoculate 10 mL of LB broth containing the relevant antibiotic. Cultures were incubated at 37 °C for 5 h with 200 rpm shaking. Plasmid purification was performed using the QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's protocol.

2.1.5. Plasmid Sequence Verification

All constructs were sequence validated using Sanger sequencing, performed by Source BioScience (Nottingham, UK).

2.1.6. Description of plasmids generated

Plasmid pET28a-His6-3C-TG2 (used for expression of TG2) was generated by Gibson assembly of the pET28a backbone with a codon optimised gBlock (IDT Technologies) comprising residues 1-687 of TG2. Numbering of TG2 is based on crystal structure (PDB 2Q3Z) (Pinkas et al., 2007). The 3C protease site was then introduced by PCR and Gibson assembly. The pET28a-10His-3C-TG2 and pET28a-TG2-10His plasmids used for optimising purification of TG2 were cloned from pET28a-His6-3C-TG2 using PCR and Gibson assembly. pET28a-TG2 C277S-10His was created by site-directed mutagenesis of pET28a-TG2-10His.

$TG2_{465}$ was expressed from the construct pET28a-$TG2_{465}$-10His, which was generated by truncation of pET28a-TG2-10His to contain only residues 1-465 of TG2. $TG2_{465}$ variants FN, C277S, C230S, C370S, C371S and W278F were generated by site-directed mutagenesis of pET28a-$TG2_{465}$-10His. pET-$TG2_{465}$ FN contains the following mutations: K30A, R166A, H134A. pET28a-$TG2_{465}$ FN-Q has the following structure: $TG2_{465}$ FN, GSS linker, Qtag, GSS linker, 10His. pET28a-$TG2_{465}$ FN-Q was generated by insertion of 5x(GSS)TVQQELGSS into pET28a-$TG2_{465}$ FN as primer leader sequence using PCR and Gibson assembly. pET28a-$TG2_{465}$ FN-K has the following structure: $TG2_{465}$ FN, GSS linker, LysTag, GSS linker, 10His. pET28a-$TG2_{465}$ FN-K was generated by insertion of 5x(GSS)GSDPVKGSS into pET28a-$TG2_{465}$ FN as primer leader sequence using PCR and Gibson assembly.

pET28a-mClover3 has the following structure: mClover3, GSG linker, 6His. pET28a-mClover3 was generated from the previously described pET28aSpyTag003-mClover3 (Addgene 133453) (Keeble et al., 2019) by removal of the SpyTag003 using PCR and Gibson assembly. pET28a-mClover-Q has the following structure: mClover3, GSG linker, Qtag, GSG linker, 6His. pET28a-mClover-Q was generated by insertion of TVQQEL in the middle of the GSG linker of pET28a-mClover to give: mClover, GSGSTVQQELSGS, 6His. pET28a-mRuby3 has the following structure: mRuby3, GSS linker, 10His. pET28a-mRuby3 was generated by cloning mRuby3 from pMF070 and inserting it into pET28a-$TG2_{465}$-10His to replace $TG2_{465}$ using PCR and Gibson assembly. pET28a-mRuby3-SnoopTag2 has the following structure: mRuby3, GSS, SnoopTag2, GSS, 6His. pET28a-mRuby3-SnoopTag2 was generated by cloning mRuby3 from pET28a-mRuby3 and inserting into pET28a-AffiTNFα-SnoopTag2 by PCR and Gibson assembly to replace AffiTNFα.

Ub(K0) is a ubiquitin variant with all lysine residues replaced with arginine. pET-Ub(K0) has the following structure: Ub(K0), GSS linker, 6His, GSS linker, Ctag. Ub(K0) was cloned from pET28a-Ub(K0)-ERY-GSY-SPM and inserted into pET28a-mClover by PCR and

Gibson assembly, with Ub(K0) replacing mClover. A Ctag was then introduced in primer leader sequences by PCR and Gibson assembly. pET28a-Ub(K0)-Q has the following structure: Ub(K0), GSS linker, Qtag, GSS linker, 6His, Ctag. pET28a-Ub(K0)-Q was generated by introduction of TVQQEL to pET28a-Ub(K0) as primer leader sequence by PCR and Gibson assembly.

pET28a-SnoopLigase has been previously described (GenBank MG867372) (Buldun et al., 2018). pET28a-SnoopLigase2 has been previously described (GenBank OQ923250, Addgene 201803) (Keeble et al., 2023). pET28a-SUMO-DogTag has been previously described (GenBank MG867376, Addgene 105629) (Buldun et al., 2018). pET28a-SUMO-DogTag2 has been previously described (GenBank OQ923254, Addgene 201807) (Keeble et al., 2023). pET28a-SnoopTagJr-sfGFP (Addgene 201809) and pET28a-SnoopTag2-sfGFP (GenBank OQ923256, Addgene 201810) have been previously described (Keeble et al., 2023). pET28a-Halotag7-SnoopLigase2 (GenBank OQ923249, Addgene 201808) was generated by cloning SnoopLigase2 from pET28a-SnoopLigase2 into the previously described pET28a-HaloTag-SnoopLigase (GenBank MG867371, Addgene 105627) (Buldun et al., 2018) to replace SnoopLigase with SnoopLigase2. pET28a-SnoopCatcher has been previously described (Addgene 72322) (Veggiani et al., 2016). pET28a-SnoopCatcher2 was a gift from Dr. Anthony Keeble and had been generated by introduction of the following mutations to pET28a-SnoopCatcher: K792T, A820S, Q822R,N847D, N780D. Numbering is based on PDB 2WW8 (Izoré et al., 2010). pET28a-AviTag-DogTag2-MBP has been previously described (GenBank MZ365293 Addgene 171773) (Keeble et al., 2022). pET28a-DogTag2-sfGFP loop A was generated by site-directed mutagenesis of the previously described pET28a-DogTag-sfGFP loop A (Addgene 171779) (Keeble et al., 2022) to introduce the I857L mutation to DogTag to create DogTag2. Numbering is based on PDB 2WW8 (Izoré et al., 2010).

pET28a-DogTag2-$TG2_{465}$ and pET28a-$TG2_{465}$-DogTag2 were generated by introduction of DogTag2 and a GSS linker to the N- or C-terminus of pET28a-$TG2_{465}$ as annealed complementary oligonucleotides. Annealing was performed in 30 mM HEPES pH 7.5, 100 mM potassium acetate at 94 °C for 2 min with 2 μM of each oligonucleotide, followed by cooling to 25 °C. pET28a-SnoopTag2-$TG2_{465}$ and pET28a-$TG2_{465}$-SnoopTag2 were generated by introduction of SnoopTag2 and a GSS linker to the N- or C-terminus of pET28a-$TG2_{465}$ as primer leader sequences. TG2x is shorthand for $TG2_{465}$ C230S. pET28a-DogTag2-TG2x (GenBank OQ923258), pET28a-TG2x-DogTag2 (GenBank OQ923259, Addgene 201811), pET28a-SnoopTag2-TG2x (GenBank OQ923260), and pET28a-TG2x-SnoopTag2 (GenBank OQ923261) were generated by site-directed mutagenesis to introduce the C230S mutation into the corresponding construct. pET28a-SnoopTag2-TGFα (GenBank OQ923262, Addgene 201812) has the structure: SnoopTag2, GSS linker, TGFα residues 40-89 (UniProt P01135), GSS linker, 6His, Ctag. pET28a-SnoopTag2-TGFα was generated by cloning of TGFα from previously described pET28a-TGFα-GSY-SPM (Lim, 2022) and insertion into pET-Ub(K0) by PCR and Gibson assembly to replace Ub(K0). pET28a-TG2x-TGFα (GenBank OQ923263) was generated by replacement of SnoopTag2 with human TGFα (codon-optimized for *E. coli*) at the C-terminus of pET28a-TG2x-SnoopTag2.

2.2. Protein Expression and Purification

2.2.1. Transformation/Transfection

For expression in *E. coli*, plasmids were transformed into BL21 (Agilent), BL21 RIPL (Agilent), Rosetta Gami (Merck), T7 Shuffle (New England Biolabs), or T7 Express (New England Biolabs) as described for Plasmid Amplification and Preparation.

For mammalian expression, Expi293F cells (Gibco) were cultured in Expi293F expression media (Gibco) to a density of 5×10^6 cells/mL. Cells were centrifuged at 200 g for 5 min and resuspended in fresh media to give 25 mL of 3×10^6 cells. Cells were transfected with 25 μg

of plasmid DNA using the ExpiFectamine 293 transfection kit (Gibco) following the manufacturers protocol. Cells were then incubated for 5 days at 37 °C with 8 % (v/v) CO_2.

2.2.2. Growth and Expression Conditions

For expression in *E. coli*, single colonies were picked from LB-agar plates and grown overnight at 37 °C in 10 mL LB containing the relevant antibiotic. Plasmids transformed into BL21 RIPL were further supplemented with 35 μg/mL chloramphenicol to maintain the RIPL plasmid. 10 mL of overnight culture was then added to 1 L of prewarmed LB in a 3 L bevelled flask and incubated at 37 ºC with 200 rpm shaking in an Infors Multitron shaker (Infors UK). Cultures were then incubated at 37 °C for ~ 3 h until OD_{600} 0.5. IPTG was added to a final concentration of 420 μM and cultures were either incubated for 5 h at 37 ºC, 5 h at 30 ºC, or 20 h at 18 ºC. At the end of the expression period the cultures were pelleted by centrifugation for 20 min at 4,000 g at 4 °C.

2.2.3. Purification of Soluble Proteins

For GST-tagged proteins, cells were resuspended in 30 mL lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 % (v/v) glycerol, 2 mM dithiothreitol (DTT)) supplemented with 10 mM phenylmethylsulfonylfluoride (PMSF) and cOmplete mini EDTA-free protease inhibitor cocktail (Roche). For His-tagged proteins, cells were resuspended in 30 mL lysis buffer (50 mM Tris-HCl pH 8.0, 400 mM NaCl, 20 mM imidazole) supplemented with 10 mM PMSF and cOmplete mini EDTA-free protease inhibitor cocktail. For C-tagged proteins, cells were resuspended in 30 mL lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl) supplemented with 10 mM PMSF and cOmplete mini EDTA-free protease inhibitor cocktail.

Samples were sonicated on ice four times for 30 s at 50% duty cycle with 1 min breaks between sonication cycles using a Sonics Vibra-Cell sonicator (Ultrasonic Processor), then centrifuged at 17,500 rpm for 25 min at 4 °C with a JA-20 rotor (Beckman Coulter). Cell supernatant was incubated with 500 μL of the relevant affinity resin for 1 h at 4 °C on an end-

over-end tube rotor. GST-tagged proteins were incubated with HiCap glutathione matrix (Qiagen), His-tagged proteins were incubated with 500 µL Nickel-nitrilotriacetic acid (Ni-NTA) Agarose (Qiagen), and C-tagged proteins were incubated with 500 µL CaptureSelect C-tag Affinity Matrix (Thermo Scientific). All subsequent steps were performed at 4 °C. Samples were added to polyprep columns (Bio-Rad) and washed thrice in 20 mL lysis buffer, then protein was eluted six times in 1 mL elution buffer each time. GST tagged proteins were eluted in 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT, 1mM EDTA, 10 % (v/v) glycerol, 10 mM reduced glutathione. His tagged proteins were eluted in 50 mM Tris-HCl pH 8.5, 500 mM NaCl, 10% (v/v) glycerol, 4 mM 2-mercaptoethanol, 200 mM imidazole. C-tagged proteins were eluted in 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% (v/v) glycerol, 4 mM 2-mercaptoethanol, 2 M $MgCl_2$.

2.2.4. Purification of TGFα from Inclusion Bodies

TGFα was expressed in *E. coli* BL21 (Agilent) as described above. Cell pellets were resuspended in 30 mL lysis buffer (50 mM Tris base, 300 mM NaCl, 0.1% (v/v) Triton X-100, pH 8.0) supplemented with 10 mM PMSF and cOmplete mini EDTA-free protease inhibitor cocktail. Samples were sonicated on ice four times for 30 s at 50% duty cycle with 1 min breaks between sonication cycles using a Sonics Vibra-Cell sonicator (Ultrasonic Processor). The cell lysate was then centrifuged at 42,500 rpm for 25 min at 4 °C in a JA-20 rotor (Beckman Coulter). Cell supernatant was discarded and the pellet was resuspended in 4 mL denaturing buffer (100 mM sodium phosphate, 10 mM Tris-HCl, 8 M urea, 2mM 2-mercaptoethanol, pH 8.0) per g of pellet. Pellet was re-homogenised by sonicating on ice three times for 30 s at 50% duty cycle with 1 min breaks between sonication cycles. Denatured cell pellet was incubated for 3 h at 25 °C with gentle shaking. The sample was then centrifuged at 42,500 rpm for 25 min at 4 °C in a JA-20 rotor (Beckman Coulter). Supernatant was collected and 500 µL Ni-NTA Agarose (Qiagen) was added, before incubating for 1 h at 4 °C on an end-over-end tube

rotor. Samples were applied to polyprep columns (Bio-Rad) and washed thrice in 20 mL wash buffer (100 mM sodium phosphate, 10 mM Tris-HCl, 8 M urea, pH 6.3), then protein was eluted six times in 1 mL elution buffer (100 mM sodium phosphate, 10 mM Tris-HCl, 8 M urea, pH 4.5).

2.2.5. Refolding of TGFα constructs

Purified denatured protein samples were refolded by the 'shock dilution' method (Wang et al., 2017b). Sample was added quickly to 4.3 mL of ice-cold refolding buffer (100 mM sodium phosphate, 10 mM Tris-HCl, 1.5 mM reduced glutathione, 0.3 mM oxidised glutathione, 10 mM PMSF, 5 mM EDTA, pH 9.7) per mL of protein sample. Refolding solution was incubated for 24 h at 4 °C with stirring. After refolding, the samples were concentrated in a Vivaspin20 concentrator (Sartorius) and then spun in a microfuge at full speed for 10 min at 4 ºC to collect aggregates and debris. The supernatant was loaded onto a HiLoad 16/600 Superdex 200 pg column pre-equilibrated with HBS pH 7.4 (50 mM HEPES, 150 mM NaCl) at 4 °C. Relevant fractions were collected and concentrated in a Vivaspin20 concentrator (Sartorius).

2.2.6. Protein Dialysis

Proteins which were not purified by SEC were dialyzed into 1,000-fold volume excess of the relevant buffer at 4 °C overnight using 3.5 kDa molecular weight cut-off dialysis tubing (Spectrum Labs). The following day proteins were dialysed twice more in fresh buffer for at least 3 hours each time. After dialysis, samples were centrifuged at 17,000g for 10 min at 4 °C.

2.3. Protein Quantification

2.3.1. Nanodrop

A_{280} was measured using an ND-1000 Nanodrop (NanoDrop); this was used to calculate the protein concentration based on the extinction coefficient of the protein predicted by ExPASy ProtParam (Gasteiger et al., 2003).

2.3.2. BCA assay

Pierce™ BCA Protein Assay Kit (Thermo Scientific) was used following the manufacturer's recommended protocol.

2.3.3. Gel Densitometry

For calculation of the concentration of small proteins which do not contain tryptophan, samples were analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) alongside proteins of similar size and known concentration. Sample quantification of protein bands was performed using Image Lab Software, version 6.1.0 (Bio-Rad).

2.4. Protein Analysis

2.4.1. SDS-PAGE

6× SDS loading buffer (0.23 M Tris–HCl pH 6.8, 24% (v/v) glycerol, 0.23 M sodium dodecyl sulfate (SDS), 120 μM bromophenol blue) was added to protein samples before heating to 99 °C for 3 min and cooling to 25 °C in a thermal cycler (Bio-Rad). Samples were loaded onto Tris-glycine gels with Tris-glycine running buffer (25 mM Tris–HCl, 192 mM glycine, 0.1% (w/v) SDS) in an Xcell SureLock Mini-cell gel chamber (Invitrogen) for 70 min at 190 V at 25 °C. Gels were stained with InstantBlue Coomassie Protein Stain (Abcam) and destained with MilliQ water. Images were acquired using a ChemiDoc XRS+ Imager (Bio-Rad).

2.4.2. Western blot

Samples were dry-transferred from Tris-glycine gels onto a nitrocellulose membrane using the iBlot2 gel transfer device (Invitrogen) at 25 V for 10 min at 25 °C. After transfer, the nitrocellulose membrane was blocked for 16-18 h at 4 °C using blocking buffer (PBS pH 7.4 + 0.05% (v/v) TWEEN 20 + 5% (w/v) skimmed milk powder). The membrane was probed with primary antibody in blocking buffer for 1 h at 25 °C. The membrane was washed 8× with PBS-T (PBS pH 7.4 + 0.05% (v/v) TWEEN 20) before probing with secondary antibody in

blocking buffer for 1 h at 25 °C. For blotting with biotinylated antibodies, skimmed milk was substituted with bovine serum albumin (BSA) in all buffers.

For Western blotting of TG2 and TG2 variants the membrane was blotted with anti-TG2 antibody (CUB 4702, Abcam) diluted 1:10,000 in blocking buffer, followed by blotting with anti-mouse HRP antibody conjugate (A4416, Merck) diluted 1:10,000 in blocking buffer. For Western blotting of His tagged proteins the membrane was blotted with anti-His antibody (HIS.H8, Thermo Fisher) diluted 1:1,000 in blocking buffer, followed by blotting with anti-mouse HRP conjugate (A4416, Merck) diluted 1:10,000 in blocking buffer. For Western blotting of mClover anti-GFP the membrane was blotted with anti-GFP antibody (B34, Biolegend) diluted 1:2,000 in blocking buffer, followed by blotting with anti-mouse HRP conjugate (A4416, Merck) diluted 1:20,000 in blocking buffer. For blotting of proteins aminylated with biotin-cadaverine the membrane was blotted with Pierce high-sensitivity streptavidin-HRP (21130, Thermo Scientific) diluted 1:3,000 in PBS pH 7.4 + 0.05% (v/v) TWEEN 20 + 5% (w/v) BSA. For cell assays, a separate blot was performed with all samples to detect glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as a sample processing control. Membranes were blotted with mouse anti-GAPDH antibody (GA1R, Invitrogen) diluted 1:750 in blocking buffer, followed by blotting with anti-mouse HRP antibody conjugate (A4416, Merck) diluted 1:3,000.

After staining with antibodies, membranes were washed 8× with PBS-T. HRP was detected using SuperSignal West Pico chemiluminescent substrate (Thermo Fisher) as per the manufacturer's protocol. Membranes were imaged using the ChemiDoc XRS+ Imager (Bio-Rad) and analysed using Image Lab version 5.2.1 (Bio-Rad).

2.4.3. Mass Spectrometry

40 μM SUMO-DogTag2 and 200 μM SnoopTag2 solid-phase synthesized peptide (GKLGYIEFYKVEKGY, Insight Biotechnology at 95% purity) were incubated with 60 μM

SnoopLigase2 in TB pH 7.4 (50 mM Tris base adjusted to pH 7.4 with boric acid) and 15% (v/v) glycerol in a total volume of 200 µL for 24 h at 4 °C. The full 200 µL reaction was purified by SEC using a HiLoad 16/600 Superdex 200 pg column, pre-equilibrated with HBS pH 7.4 running buffer. The relevant peak was collected and concentrated in a Vivaspin20 spin concentrator (Sartorius) with 5 kDa MWCO. Analysis was performed by Dr. Anthony Tumber in the Department of Chemistry using a RapidFire 365 platform (Agilent). This setup comprised a jet-stream electrospray ionization source coupled to an Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) (Agilent) detector. Data was analyzed using MassHunter Qualitative Analysis, version (Agilent) with observed molecular weights compared to the expected molecular weights determined by ExPASy ProtParam (Gasteiger et al., 2003) with subtraction of the initiating formyl-methionine. An additional 17.0 Da was also subtracted to account for the loss of ammonia during isopeptide bond formation.

2.5. Protein Conjugation Assays

2.5.1. Biotin-cadaverine transamidation assay

This assay was adapted from (Király et al., 2016) with alterations to better replicate physiological conditions and improve well blocking to reduce background signal. Wells of a 96 well Maxisorp plate were coated with 200 µL of 100 µg/mL N,N-dimethyl casein (Sigma-Aldrich) in HBS (20 mM HEPES, 150 mM NaCl, pH 7.4) overnight at 4°C. Wells were then blocked with 300 µL of Pierce protein-free TBS blocking buffer (Thermo Fisher) for 1 h at 25 °C. Wells were washed thrice with 300 µL of HBS-T (HBS pH 7.4 + 0.5% (v/v) TWEEN 20) before incubation of 200 µL TG2 in reaction buffer (HBS pH 7.4, 1 mM DTT, 1 mM biotin-cadaverine, 1 mM $CaCl_2$) for 30 min at 37 °C. The reaction was quenched by washing wells thrice with 300 µL of HBS-T + 50 mM EDTA before incubation with 150 µL of 0.3 µg/mL Pierce high-sensitivity streptavidin-HRP (21130, Thermo Scientific) diluted in HBS pH 7.4 for 1 h at 25 °C . Wells were washed six times in HBS-T before signal generation by adding 100

µL of 1-Step Ultra TMB-ELISA Substrate Solution (Thermo Scientific). The reaction was stopped by addition of 100 µL of 1 M HCl and absorbance measured at 450 nm on a FLUOstar Omega plate reader (BMG Labtech). Commercially available mTG (Meat Glue, Special Ingredients) was used as a positive control to confirm detection of transamidation.

2.5.2. SnoopLigase-mediated conjugation of TG2 to cargo

TG2x-DogTag2 was incubated with SnoopLigase2 and cargo protein in 2 mL total volume in HBS pH 7.4 at 4 °C for 16 h with gentle rotation. To confirm completion of the reaction, samples were analysed by SDS−PAGE with Coomassie staining. The conjugated product was separated from unreacted cargo by SEC. 2 mL of sample was applied to a previously equilibrated Superdex 200 HiLoad 16/600 column (Cytiva), controlled by an ÄKTA pure (Cytiva) at 4 °C, with HBS pH 7.4 running buffer. The peak corresponding to the conjugate product was concentrated in a Vivaspin 20 MWCO 30 kDa (Sartorius) and analysed by SDS−PAGE with Coomassie staining. % coupling to TG2x was calculated based on the depletion of the band of tagged TG2x:

100 × [1 – (tagged TG2x with SnoopLigase2)/(tagged TG2x without SnoopLigase2)]

2.5.3. SnoopLigase tag/set comparison assays

25 µM equimolar reactions were performed at 21 °C for 3 h in a total volume of 20 µL. Reactions were quenched by addition of 6× SDS loading buffer and 100 mM DTT. Samples were heated at 95 °C for 3 min and then analysed by SDS−PAGE with Coomassie staining. Coupling efficiency (%) was calculated from band intensities as 100× product/(product + unreacted DogTag variant + unreacted SnoopTag variant) for three replicates.

2.5.4. SnoopLigase reaction condition validation assays

Unless stated otherwise, 20 µM of each protein was incubated in 10 mM HEPES pH 7.5 at 25 °C for the desired duration. To terminate the reaction, 6× SDS loading buffer (0.23 M Tris-HCl, pH 6.8, 24% (v/v) glycerol, 120 µM bromophenol blue, 0.23 M SDS) was added.

For temperature-dependence assays, reactions were performed at the indicated range of temperatures for 10 min. For pH-dependence assays, reactions were performed for 9 min in Succinate-Phosphate-Glycine (12.5 mM succinic acid, 50 mM sodium dihydrogen phosphate, 44 mM glycine) to allow buffering over the full range of pH values tested. For the additive-dependence assays, reactions were performed for 8 min in HBS pH 7.0 with: 1% (v/v) Triton X-100, 1% (v/v) TWEEN 20, 1 mM EDTA, or 2 mM DTT. For buffer-dependence assays, reactions were performed for 12 min at pH 7.5 with: 90 mM Tris-borate, 100 mM sodium phosphate buffer (8:2 ratio of Na_2HPO_4 to NaH_2PO_4), 10 mM HEPES, or PBS. Buffers were added to the reactions as 10× stock solutions. For the NaCl-dependence assays, reactions were performed for 10 min with the indicated range of NaCl concentrations added to the buffer.

100 mM DTT was added to the quenched reaction samples, then samples were heated in a thermocycler for 3 min at 95 °C. SDS-PAGE was performed at 190 V in 25 mM Tris-HCl, 192 mM glycine, and 0.1% (w/v) SDS. Gels were stained with InstantBlue Coomassie stain, destained with Milli-Q water, and imaged using a ChemiDoc XRS imager with ImageLab version 6.1.0 software (Bio-Rad). Product formation was quantified using ImageLab version 6.1.0.

2.5.5. SnoopTag/Catcher comparison assays

100 nM equimolar reactions were performed at 21 °C in a total volume of 20 µL HBS pH 7.0 + 2% (w/v) BSA. BSA was added to prevent sticking of proteins to the side of the tube given the low concentrations used. Reactions were quenched by addition of 6× SDS loading buffer and 100 mM DTT. Samples were heated at 95 °C for 3 min and then analysed by SDS−PAGE with Coomassie staining. Coupling efficiency (%) was calculated from band intensities as 100× product/(product + unreacted SnoopTag2 variant) for three replicates.

2.5.7. SnoopTag/Catcher reaction condition validation assays

Unless otherwise stated, 100 nM of each protein was incubated in PBS pH 7.4 + 2% (w/v) BSA for 10 min at 16 °C. To terminate the reaction, 6× SDS loading buffer (0.23 M Tris-HCl, pH 6.8, 24% (v/v) glycerol, 120 μM bromophenol blue, 0.23 M SDS) was added. For temperature-dependence assays, reactions were performed for the indicated duration. For pH-dependence assays, reactions were performed in Succinate-Phosphate-Glycine (12.5 mM succinic acid, 50 mM sodium dihydrogen phosphate, 44 mM glycine) at the indicated pH for 12 min. For additive-dependence assays, reactions were performed in HBS pH 7.0 + 2% (w/v) BSA for 12 min with 1% (v/v) Triton X-100, 1% (v/v) TWEEN 20, 1 mM EDTA, or 2 mM DTT. For buffer-dependence assays reactions were performed for 12 min with the indicated buffer at pH 7.5. For NaCl-dependence assays, reactions were performed with the addition of the indicated range of NaCl concentrations added to the buffer.

100 mM DTT was added to the quenched reaction samples, then samples were heated in a thermocycler for 3 min at 95 °C. SDS-PAGE was performed at 190 V in 25 mM Tris-HCl, 192 mM glycine, and 0.1% (w/v) SDS. Gels were stained with InstantBlue Coomassie stain, destained with Milli-Q water, and imaged using a ChemiDoc XRS imager with ImageLab version 6.1.0 software (Bio-Rad). Product formation was quantified using ImageLab version 6.1.0.

2.6. SnoopLigase product separation

TG2x-DogTag2, SnoopTag2-sfGFP, and HaloTag-SnoopLigase2 were incubated at 10 μM each in HBS pH 7.4 in a total volume of 200 μL for 16 h at 25 °C. To capture HaloTag-SnoopLigase2, 25 μL of washed and equilibrated Magne HaloTag resin (Promega) was added, followed by TWEEN 20 at a final concentration of 0.01% (v/v). Samples were incubated for 1 h at 24 °C on a rotor. The resin was collected using a magnetic rack. After washing the resin twice with 100 μL of HBS-T (HBS pH 7.4 + 0.01% (v/v) TWEEN 20), then thrice with 100 μL

of HBS pH 7.4, the resin was resuspended in one of the three buffers for the elution step. For acid elution, the resin was resuspended in 25 µL of 50 mM glycine−HCl pH 2.5 and incubated in a ThermoMixer (Eppendorf) for 1 min at 37 °C with 800 rpm shaking. The resin was collected with a magnetic rack, and the eluent was removed and neutralised by addition of 2.5 µL 1 M Tris−HCl pH 9.5. Acid elution was then repeated twice more. For imidazole elution, the resin was resuspended in 25 µL of HBS with 3 M imidazole adjusted to pH 7.4. The resin was incubated in a ThermoMixer for 4 h at 37 °C with 800 rpm shaking. The resin was collected, and the eluent was removed. Elution was then repeated twice more. For peptide elution, the resin was resuspended in 25 µL of PBS pH 7.4, with the DogTag:SnoopTagJr conjugated product and incubated in a ThermoMixer for 4 h at 37 °C with 800 rpm shaking. The resin was collected, and the eluent was removed. Elution was then repeated twice more. Beads were then resuspended in 50 µL of PBS pH 7.4 and 6× SDS loading buffer and heated at 98 °C for 3 min to release any proteins bound to the bead-coupled HaloTag-SnoopLigase2.

2.7. Plate binding assays

2.7.1. ECM protein binding assays

Wells of a Nunc MaxiSorp 96-well flat bottom microtiter plate (Thermo Scientific) were coated with 100 µL of 5 µg/mL protein diluted in HBS pH 7.4 overnight at 4 °C. For fibronectin assays, fibronectin 45-kDa proteolytic fragment (F0162, Sigma-Aldrich) was used. For collagen assays, human type I collagen extracted from placenta by pepsin (AB7533, Abcam) was used. For basement membrane assays, MaxGel (E0282, Sigma-Aldrich) was used. MaxGel is a basement membrane extract derived from *in vitro* co-culture of human fibroblasts and human epithelial cells. MaxGel contains collagens, laminin, fibronectin, tenascin, elastin, and several proteoglycans. Wells were blocked with 300 µL of HBS pH 7.4 + 3% (w/v) BSA for 1 h at 25 °C . Wells were washed thrice in 300 µL of HBS-T (HSB pH 7.4 + 0.05% (v/v) TWEEN 20), then 100 µL of TG2 diluted in HBS-T was incubated for 2 h at 37 °C. Wells were

washed thrice in 200 µL of HBS-T. Rabbit anti-TG2 antibody (ab137378, Abcam) was diluted 1:250 in HBS-T + 1% (w/v) BSA and 100 µL of primary antibody solution was incubated in wells for 1 h at 25 °C . Wells were washed thrice in 200 µL HBS-T. Anti-rabbit HRP antibody conjugate (A4416, Merck) was diluted 1:1000 in HBS-T + 1% (w/v) BSA and 100 µL was incubated in wells for 1 h at 25 °C . Wells were washed six times in HBS-T, then signal was generated by adding 100 µL of 1-Step Ultra TMB-ELISA Substrate Solution (Thermo Scientific). The reaction was stopped by addition of 100 µL of 1 M HCl and absorbance was measured at 450 nm on a FLUOstar Omega plate reader (BMG Labtech).

2.7.2. TG2:cargo retention assay

Wells of a Vision Plate^{TM} 98-well black with a clear flat bottom (Thermo Scientific) were coated with 100 µL of 50 µg/mL fibronectin 45-kDa proteolytic fragment (F0162, Sigma-Aldrich) diluted in HBS pH 7.4 overnight at 4 °C. Wells were blocked with 300 µL of PBS pH 7.4 + 2% (w/v) BSA for 1 h at 25 °C . Wells were washed thrice in 300 µL of PBS-T (PBS pH 7.4 + 0.05% (v/v) TWEEN 20), then 100 µL of TG2:cargo conjugate diluted in PBS-T was incubated in the wells for 1 h at 37 °C. Wells were washed for the specified number of times in 200 µL of PBS-T. 100 µL PBS-T was added to the wells and GFP signal was measured with 485 nm excitation and 510 nm emission on a FLUOstar Omega plate reader (BMG Labtech).

2.8. Cell surface binding assays

2.8.1. Transamidation assay for flow cytometry

Expi293F cells (Thermo Scientific) were washed by centrifugation at 125 g for 5 min at 4 °C, followed by resuspension in 50 mL of HBS pH 7.4 + 1% (w/v) BSA. Cells were centrifuged at 125 g for 5 min at 4 °C and resuspended in HBS pH 7.4 + 1% (w/v) BSA to 10^7 cells/mL. 100 µL of cell solution was incubated with 5 µM TG2 for 15 min at 37 °C. 2 mM $CaCl_2$, 1 mM DTT and 5 µM substrate protein (if used) were added and incubated for a further 30 min. Cells were washed thrice in 2 mL of ice cold HBS + 1% (w/v) BSA at 125 g for 5 min

at 4 °C, then resuspended in 100 µL of primary antibody. For detection of TG2 and TG2 variants, cells were incubated on ice for 30 min in anti-TG2 antibody (ab137378, Abcam) diluted 1:100 in HBS + 1% (w/v) BSA. Cells were then washed thrice in 2 mL of HBS + 1% (w/v) BSA at 125 g for 5 min at 4 °C and resuspended in 100 µL of secondary antibody. For detection of TG2 and TG2 variants, cells were incubated on ice for 30 min in anti-rabbit 647 antibody (A-21245, Invitrogen) diluted 200-fold in HBS + 1% (w/v) BSA. Cells were washed thrice in 2 mL of HBS + 1% (w/v) BSA at 125 g for 5 min at 4 °C, then resuspended in 1 mL HBS + 1% (w/v) BSA and 1 µg/mL DAPI, before imaging on a CytoFlex LX flow cytometer. Data were analysed in FlowJo, version 10.8.1 (FlowJo LLC).

2.8.2. Cell binding assay for flow cytometry

A431 cells (ATCC) were cultured to confluency in T75 Nunc™ EasYFlask™ Cell Culture Flasks (Thermo Scientific) with DMEM (Gibco) + 10% (v/v) fetal bovine serum (FBS) (Gibco) + 1% (v/v) penicillin/streptomycin (Gibco). Confluent cells were washed once by addition of 10 mL PBS pH 7.4 to the flask and removal of supernatant with a pipette. Cells were then detached by incubation in 3 mL of prewarmed HBS pH 7.4 + 1% (w/v) BSA + 1 mM EDTA at 37 °C. Detached cells were washed by centrifugation at 125 g for 5 min at 4 °C and resuspended in 10 mL of HBS pH 7.4 + 1% (w/v) BSA. Cells were centrifuged again at 125 g for 5 min at 4 °C and resuspended in HBS pH 7.4 + 1% (w/v) BSA to 10^7 cells/mL. 100 µL of cell solution was incubated with 50 nM – 6 µM TG2 for 30 min at 37 °C. Cells were washed thrice in 2 mL of ice cold HBS + 1% (w/v) BSA at 125 g for 5 min at 4 °C, then resuspended in 100 µL of primary antibody. For detection of TG2 and TG2 variants, cells were incubated on ice for 30 min in anti-TG2 antibody (ab137378, Abcam) diluted 1:100 in HBS + 1% (w/v) BSA. Cells were then washed thrice in 2 mL of HBS + 1% (w/v) BSA at 125 g for 5 min at 4 °C and resuspended in 100 µL of secondary antibody. For detection of TG2 and TG2 variants, cells were incubated on ice for 30 min in anti-rabbit 647 antibody (A-21245, Invitrogen) diluted

200-fold in HBS + 1% (w/v) BSA. Cells were washed thrice in 2 mL of HBS + 1% (w/v) BSA at 125 g for 5 min at 4 °C, then resuspended in 1 mL HBS + 1% (w/v) BSA and 1 µg/mL DAPI, before imaging on a CytoFlex LX flow cytometer. Data were analysed in FlowJo, version 10.8.1 (FlowJo LLC).

2.8.3. Transamidation assay for Western blot

Expi293F cells were washed by centrifugation at 125 g for 5 min at 4 °C and resuspension in 50 mL of HBS pH 7.4. Cells were centrifuged again at 125 g for 5 min at 4 °C and resuspended in HBS pH 7.4 to 10^7 cells/mL. 100 µL of cell solution was incubated with the desired concentration of TG2 for 15 min at 37 °C. 2 mM $CaCl_2$, 1 mM DTT and substrate protein (if used) were added and incubated for a further 30 min. Cells were washed thrice in 2 mL of ice cold HBS pH 7.4 at 125 g for 5 min at 4 °C, then resuspended in 150 µL of lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% (v/v) Triton X-100, 0.1% (w/v) SDS, 0.5% (w/v) sodium deoxycholate pH 7.4) supplemented with cOmplete protease inhibitor mix and 1 mM PMSF. Lysed cells were incubated for 10 min at 4 °C with gentle agitation. Cell debris was removed by centrifugation at 12,000g for 20 min at 4 °C.

2.9. Redox assays

2.9.1. Assay of reduced TG2 in oxidising conditions

Wells of a Nunc MaxiSorp 96-well flat bottom microtiter plate were coated with 100 µL of 5 µg/mL protein diluted in HBS pH 7.4 overnight at 4 °C. Wells were blocked with 300 µL of HBS pH 7.4 + 3% (w/v) BSA for 1 h at 25 °C. Wells were washed thrice in 300 µL of HBS-T (HSB pH 7.4 + 0.05% (v/v) TWEEN 20). TG2 was diluted to 50 nM in HBS-T + 1 mM DTT, then 100 µL of diluted TG2 was incubated in each well for 1.5 h at 4 °C. At the beginning of the assay, the wells were quickly washed thrice in HBS pH 7.4 which was pre-warmed to 37 ˚C. Following washing, the wells were incubated with 100 µL HBS pH 7.4 at 37 °C in a heat block with 300 rpm shaking. At the desired time-point, transamidation was activated by

addition of 50 µL of reaction buffer (HBS pH 7.4, 3 mM biotin-cadaverine, 6 mM $CaCl_2$) followed by 1 min incubation at 37 °C in a heat block with 300 rpm shaking. After 1 min, the reaction was quenched by addition of 150 µL stopping solution (PBS pH 7.4, 100 µM ZDON (Z006, Zedira), 50 mM EDTA). Once all samples were quenched, wells were washed twice in 300 µL PBS-T (PBS pH 7.4 + 0.05% (v/v) TWEEN 20), before incubation with 150 µL of 0.3 µg/mL Pierce high-sensitivity streptavidin-HRP (21130, Thermo Scientific) diluted in PBS-T for 1 h at 25 °C . Wells were washed six times in 300 µL PBS-T, then signal was generated by addition of 100 µL of 1-Step Ultra TMB-ELISA Substrate Solution (Thermo Scientific). The reaction was stopped by addition of 100 µL of 1 M HCl and absorbance measured at 450 nm on a FLUOstar Omega plate reader (BMG Labtech).

2.9.2. Assay of oxidised TG2 in reducing conditions

Wells of a Nunc MaxiSorp 96-well flat bottom microtiter plate (Thermo Fisher) were coated with 100 µL of 5 µg/mL protein diluted in HBS pH 7.4 overnight at 4 °C. Wells were blocked with 300 µL of HBS pH 7.4 + 3% (w/v) BSA for 1 h at 25 °C . Wells were washed thrice in 300 µL of HBS-T (HSB pH 7.4 + 0.05% (v/v) TWEEN 20). TG2 was allowed to oxidise by incubation in HBS pH 7.4 in the absence of reducing agent for 90 min at 4 °C prior to the assay. 100 µL of TG2 diluted to 50 nM in HBS-T was incubated in each well for 1.5 h at 4 °C. At the beginning of the assay, the wells were quickly washed thrice in HBS pH 7.4 which was pre-warmed to 37 °C. Following washing, the wells were incubated with 100 µL HBS pH 7.4 + 1mM DTT at 37 °C in a heat block with 300 rpm shaking. At the desired time-point, transamidation was activated by addition of 50 uL of reaction buffer (HBS pH 7.4, 1 mM DTT, 3 mM biotin-cadaverine, 6 mM $CaCl_2$), followed by 1 min incubation at 37 °C in a heat block with 300 rpm shaking. After 1 min, the reaction was quenched by addition of 150 µL stopping solution (PBS pH 7.4, 100 µM ZDON, 50 mM EDTA). Once all samples were quenched, wells were washed twice in 300 µL PBS-T (PBS pH 7.4 + 0.05% (v/v) TWEEN 20) before incubation

with 150 µL of 0.3 µg/mL Pierce high-sensitivity streptavidin-HRP (21130, Thermo Scientific) diluted in PBS-T for 1 h at 25 °C . Wells were washed six times in 300 µL PBS-T, then signal was generated by addition of 100 µL of 1-Step Ultra TMB-ELISA Substrate Solution (Thermo Scientific). The reaction was stopped by addition of 100 µL of 1 M HCl and absorbance measured at 450 nm on a FLUOstar Omega plate reader (BMG Labtech).

2.10. TGFα cell assays

2.10.1. Culture and Treatment of Cells

DU145 cells (ATCC) were cultured to confluency in T75 Nunc™ EasYFlask™ Cell Culture Flasks (Thermo Scientific) with RPMI 1640 (Gibco) + 10% (v/v) fetal bovine serum (FBS) (Gibco) + 1% (v/v) penicillin/streptomycin (Gibco). Confluent cells were washed twice with RPMI 1640 and serum-starved in RPMI 1640 + 1% (v/v) penicillin/streptomycin for 24 h prior to detachment with Trypsin-EDTA (0.05%) (Gibco). Detached cells were washed once in RPMI 1640 + 10% (v/v) FBS to quench trypsin, then washed in RPMI 1640 + 1% (v/v) penicillin/streptomycin to remove FBS and resuspended in RPMI 1640 + 1% (v/v) penicillin/streptomycin. 9×10^4 DU145 cells were seeded and then incubated for 3 h to allow attachment. Once attached, protein diluted in HBS pH 7.4 was added to a final concentration of 100 nM. Cells were incubated for 30 min, then selected samples were washed five times with RPMI 1640, and all samples were incubated in RPMI 1640 + 1% (v/v) penicillin/streptomycin for the desired duration.

2.10.2. Fluorescence microscopy

9×10^4 DU145 cells were seeded onto 8-well glass coverslips (80807, Ibidi) and treated with protein and washes as detailed above. After treatment, cells were incubated for 3 days at 37 °C with 5% (v/v) CO_2. Cells were then fixed for 20 min at 25 °C in PBS pH 7.4 + 4% (w/v) paraformaldehyde, permeabilised for 5 min at 25 °C in PBS pH 7.4 + 0.1% (v/v) Triton X-100, then blocked for 30 min at 25 °C with PBS + 1% (w/v) BSA + 0.1% (v/v) TWEEN 20. For E-

Cadherin and Vimentin staining, cells were stained with 1.25 μg/mL each of mouse anti-E-Cadherin M168 (ab76055, Abcam) and rabbit anti-Vimentin EPR3776 (ab92547, Abcam) for 1 h followed by 3 washes with PBS-T (PBS pH 7.4 + 1% (v/v) TWEEN 20). Then cells were stained with 2.5 μg/mL of goat anti-mouse Alexa Fluor 647 (A-21236, Invitrogen) and goat anti-rabbit Alexa Fluor 568 (A-11011, Invitrogen) secondary antibodies for 1 h, followed by 3 washes with PBS-T. All antibodies were diluted in PBS pH 7.4 + 1% (w/v) BSA + 0.1% (v/v) TWEEN 20. For phalloidin staining, cells were incubated with 5 U/mL Phalloidin CF647 (BT00041-T, Cambridge Bioscience) diluted in PBS pH 7.4 for 1 h, followed by 3 washes in PBS pH 7.4. All microscopy samples were counter-stained with 1 μg/mL DAPI (Life Technologies) for 5 min prior to imaging. Cells were imaged on an Olympus ScanR wide-field microscope with a scientific Complementary Metal–Oxide–Semiconductor (sCMOS) Hamamatsu Orca Fusion B camera using a 40× LCACHN Air/Dry numerical aperture (NA) 0.55 working distance (WD) 2.2 mm objective. Images were collected using ScanR Acquisition software (version 3.0.0). Excitation was performed with 395/20 excitation and 432/36 emission (DAPI), 575/25 excitation and 595/31 emission (Alexa Fluor 568), and 640/30 nm excitation and 698/70 nm emission (Alexa Fluor 647). Identical exposure times were used within a single channel for all samples. Typical exposure times were 50-200 ms. Images were cropped in ImageJ (version 1.54b) and show representative fields of view. All images in the same figure were prepared, collected and analysed using the same settings. To quantify cell morphology, 3 images were taken per treatment condition of cells stained with Phalloidin CF647 (BT00041-T, Cambridge Bioscience) using 640/30 nm excitation and 698/70 emission. Visualisation of actin filaments was used to manually segment the cells in ImageJ (version 1.54b) to calculate the aspect ratio, defined as (largest diameter along the major axis)/(largest diameter orthogonal to the major axis). Violin plots were generated from 51 cells per treatment condition using GraphPad Prism (version 9.5.0).

2.10.2. Resazurin assay

1.5×10^4 DU145 cells were seeded into wells of a 96-well clear-bottom plate (165305, Thermo Fisher) and treated with protein and washes as detailed above. After treatment the cells were incubated for 3 days at 37 °C with 5% (v/v) CO_2. 10 µL AlamarBlue HS cell viability reagent (A50100, Thermo Fisher) was added to 90 µL cell culture and cells were incubated for 60 min at 37 °C with 5% (v/v) CO_2. Fluorescence was measured using a SpectraMax M3 (Molecular Devices) microplate reader with 560 nm excitation and 590 nm emission. Sample signal was blanked against a control sample of RPMI 1640 media + 10% (v/v) AlamarBlue HS cell viability reagent.

Chapter 3 – Engineering TG2 for Controllable Covalent Coupling

3.1. Introduction

There are several research questions which must be addressed to determine the feasibility of TG2 as a protein-protein coupling tool (Table 3.1). For large scale application and to encourage widespread use, is high yield and purity of recombinant TG2 easily achievable? Secondly, there are previously reported observations of TG2 self-crosslinking both *in vitro* and *in vivo* (Stamnaes et al., 2015). To what extent does self-crosslinking occur in the presence of other substrates? Can TG2 self-crosslinking be eliminated while retaining transamidase activity? Thirdly, TG2 is known to associate with ECM proteins (Belkin, 2011). Can TG2 be easily removed from commonly used protein scaffolds after conjugation has been performed? Fourthly, the kinetics of TG2 redox regulation are unknown. Does TG2 remain active in oxidising conditions for long enough to perform reasonable levels of conjugation? Is TG2 activity fully recovered upon re-reduction, allowing multiple cycles of conjugation? Finally, glutamine-bearing peptide motifs have been used to direct aminylation by TG2, FXIII and mTG (Sugimura et al., 2006) (Oteng-Pabi et al., 2018). Can such motifs be used to direct protein-protein assembly by TG2?

Table 3.1 The key challenges for developing a TG2-based protein-protein coupling

Desirable traits	Challenges to be investigated	Addressed successfully
High yield and purity of TG2	Reports of limited yield and solubility issues	✓
Conjugation is precise	Report of self-crosslinking which may interfere with substrate conjugation	✓
TG2 is biologically inert aside from key function	High-affinity non-covalent interactions with extracellular proteins	✗
TG2 can perform conjugation in oxidising conditions	Inactive in non-reducing conditions, redox regulation poorly characterised	✓
TG2 directs conjugation via a short peptide motif	Glutamine motifs only used previously for aminylation	✓

The aim of this chapter was to explore the properties of TG2 with relation to the desirable traits above to determine the feasibility of TG2 as a protein coupling tool. Firstly, I optimised recombinant expression of TG2 to allow convenient production at reasonable yield and high purity for protein assembly applications. I then removed the two C-terminal β-barrel domains of TG2 to eliminate self-crosslinking, an approach taken from previous research to prevent formation of TG2-gluten multimeric antigens (Stamnaes et al., 2015). I evaluated the retention of TG2 and TG2 variants in ECM scaffolds and at the cell surface. I found robust binding to fibronectin, collagen and the cell surface which was not substantially decreased by disrupting fibronectin and heparan sulfate proteoglycan biding sites. I also explored the kinetics of TG2 redox regulation and developed a TG2 variant with altered redox regulation to increase transamidation in non-reducing conditions. Finally, I introduced TG2-reactive peptide tags to model proteins and to TG2 itself to explore strategies for controlled TG2-mediated conjugation.

3.2. Recombinant Expression of TG2

Recombinant expression of TG2 with high yield and purity is important for convenient use of TG2 as a conjugation tool. I first expressed codon-optimised TG2 in *E. coli* with an N-terminal 6×His tag. Purification by Ni-NTA chromatography resulted in low yield of 1.5 mg/L and low purity of 31% (Figure 3.1A). To improve purity, I extended the length of the His tag to 10×His to allow more stringent 30 mM and 100 mM imidazole washes. Stringent washes improved purity to 58% with slightly improved yield of 2 mg/L, but two bands of impurities were persistent and only eluted with the full-length product (Figure 3.1B).

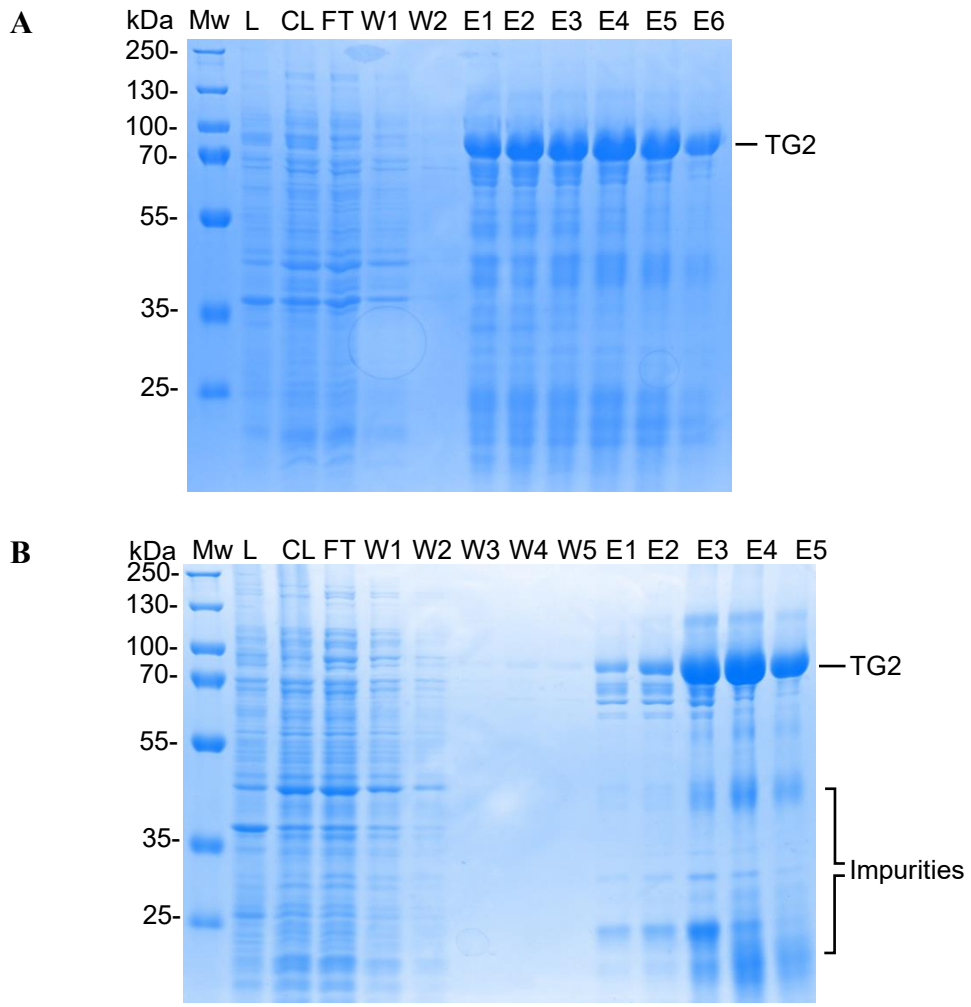


Figure 3.1. Expression and purification of recombinant TG2 in *E. coli* with an N-terminal 6 × His tag. (A) and (B) - analysis by 12% SDS-PAGE with Coomassie staining. All samples were reduced with 100 mM DTT. Mw = molecular weight markers, L = cell lysate, CL = cleared lysate, FT = flow-through, W = wash, E = elution. (A) Ni-NTA purification of 6×His-TG2. Column washes performed with 10 mM imidazole. Elution performed with 200 mM imidazole. One prep performed. (B) Ni-NTA purification of 10×His-TG2 with stringent washing. Column washes W1-3 performed with 10 mM imidazole, washes W4-5 performed with 30 mM imidazole. Elution performed with 200 mM imidazole. One prep performed.

To identify if the observed impurities were non-specific contaminants or truncated His-tagged products, I performed a Western blot of the elution fractions with an anti-His antibody. The presence of multiple bands corresponding in size to the impurities suggests that the contaminants were truncated His-tagged TG2 products (Figure 3.2A). To confirm this, I introduced a 3C protease site to the construct between the 10×His tag and TG2 to give 10×His-3C-TG2. The construct was purified with stringent washes as before, then cleaved from the

column by 3C protease digestion. Elution of two distinct bands of impurities by 3C cleavage construct confirmed that these impurities are TG2 products (Figure 3.2B). Given that the His tag is situated at the N-terminus, these captured TG2 products must be C-terminally truncated. Truncation may have occurred due to premature termination of translation, or TG2 may have been fully translated and then cleaved by proteases during the purification process. To determine the cause of truncation, I switched the 10×His tag to the C-terminus of TG2 to give TG2-10×His. If truncation was due to protease digestion of full length TG2, then both the N-terminus and the C-terminus of TG2 will be present in the cell lysate. In this case, capturing TG2 with a C-terminal tag would also result in capture of truncated products. The C-terminal His tag did not result in capture of truncated products (Figure 3.2C). Absence of truncated TG2 products containing the C-terminus of full length TG2 suggests that the observed truncations were the result of premature termination of transcription. The product was further purified by size-exclusion chromatography (Figure 3.2D). Isolation of the full-length construct from the truncated forms resulted in a yield of 4 mg/L and purity of 91%. TG2 was purified at convenient yield and purity with a C-terminal 10×His tag, this purification method was used for all future TG2 constructs used in this thesis.

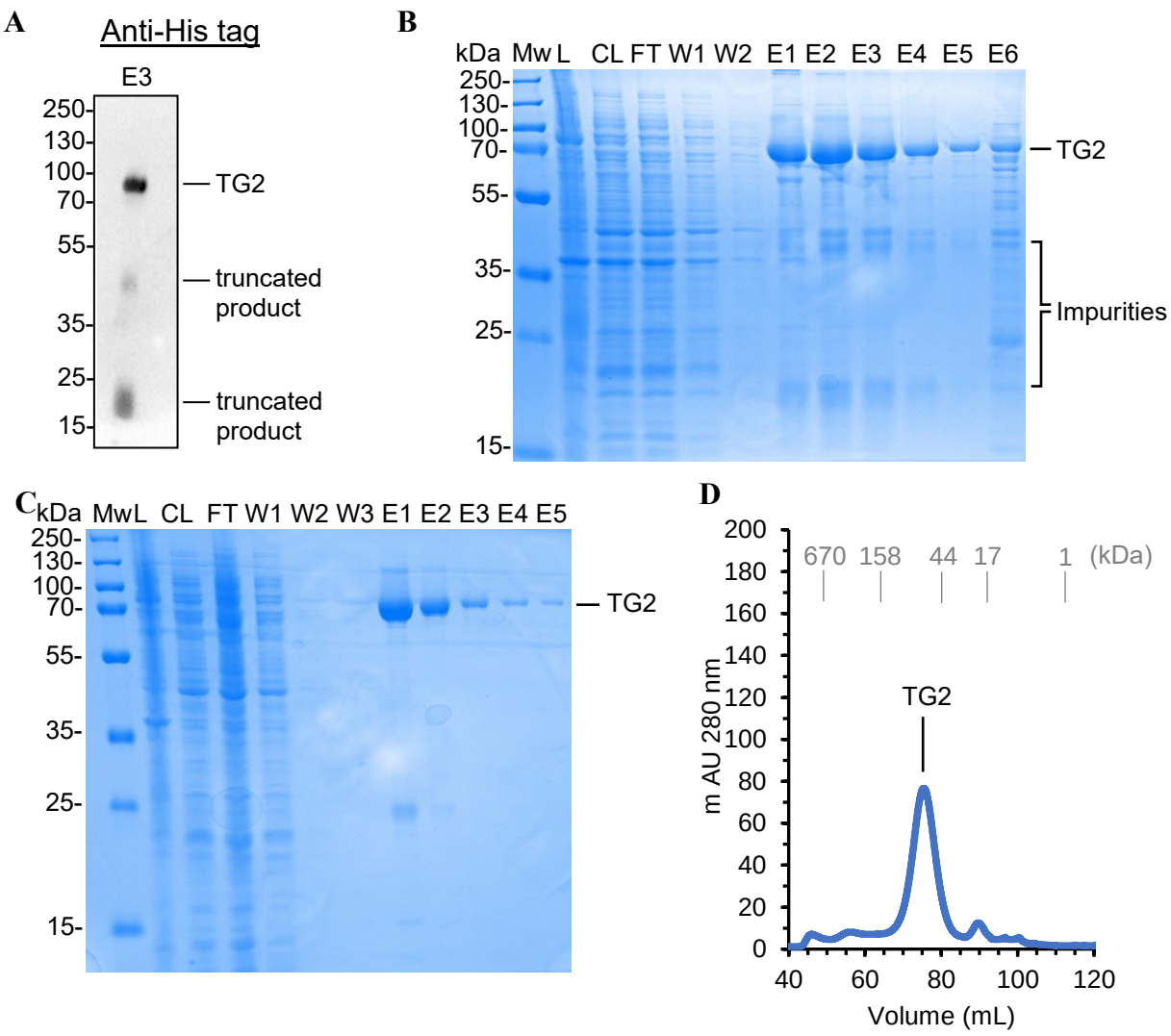


Figure 3.2. Premature termination of translation of TG2. (A) Western blot of elution fraction from (3.1B), His-tagged constructs were detected with mouse anti-His antibody and anti-mouse HRP antibody conjugate. (B) and (C) - analysis by 12% SDS-PAGE with Coomassie staining. All samples were reduced with 100 mM DTT. Mw = molecular weight marker, L = cell lysate, CL = cleared lysate, FT = flow-through, W = wash, E = elution. (B) Ni-NTA purification of 10×His-TG2 with 3C protease cleavage. Column washes performed with 30 mM imidazole. E1-5 are fractions eluted with TBS pH 8.0 after incubation with 150 ug 3C protease. E6 is a fraction eluted with 200 mM imidazole. One prep performed. (C) Ni-NTA purification of TG2-10×His. Column washes performed with 30 mM imidazole. Elution performed with 200 mM imidazole. One representative prep shown of three preparations performed in total. (D) Size-exclusion chromatography of TG2-10×His.

3.3. Transamidase Activity and TG2 Self-Crosslinking

Next, I looked for an assay that measures transglutaminase activity in a biologically relevant way to confirm the ability of my recombinant TG2 to transamidate protein substrates. Multiple assays have been developed for measuring transamidase activity, the most common is the hydroxamate assay. In the hydroxamate assay, TG2 catalyses aminylation of tripeptide acyl donor ZQG with the acyl acceptor hydroxylamine resulting in production of hydroxamate (Folk and Cole, 1966). $FeCl_3$ forms a chromogenic complex with hydroxamate which can be measured colorimetrically. The disadvantage of the hydroxamate assay is that it measures the ability of TG2 to conjugate a tripeptide to a small molecule, a reaction where steric compatibility is not a limiting factor. It has limited relevance for studying the ability of TG2 to conjugate two large protein substrates together. Given that the TG2 active site is buried in a deep hydrophobic tunnel, steric compatibility may play a significant role in transamidation. Thus, I used an alternative assay throughout this thesis in which TG2 catalyses aminylation of a whole protein substrate (casein) with a small molecule amine (cadaverine) (Figure 3.3A) (Slaughter et al., 1992). While this assay does not measure true protein-protein conjugation, the use of a protein as the glutamine donor allows for some steric hindrance to influence reactivity. A disadvantage of this assay is that product formation is not directly measured and thus reaction kinetics cannot be calculated. I optimised this cadaverine assay and used it to confirm that purified TG2 is active (Figure 3.3B), utilising the catalytically inactive variant TG2 C277S as a negative control. These results confirm that the recombinant TG2 sample can catalyse aminylation of a protein substrate with cadaverine.

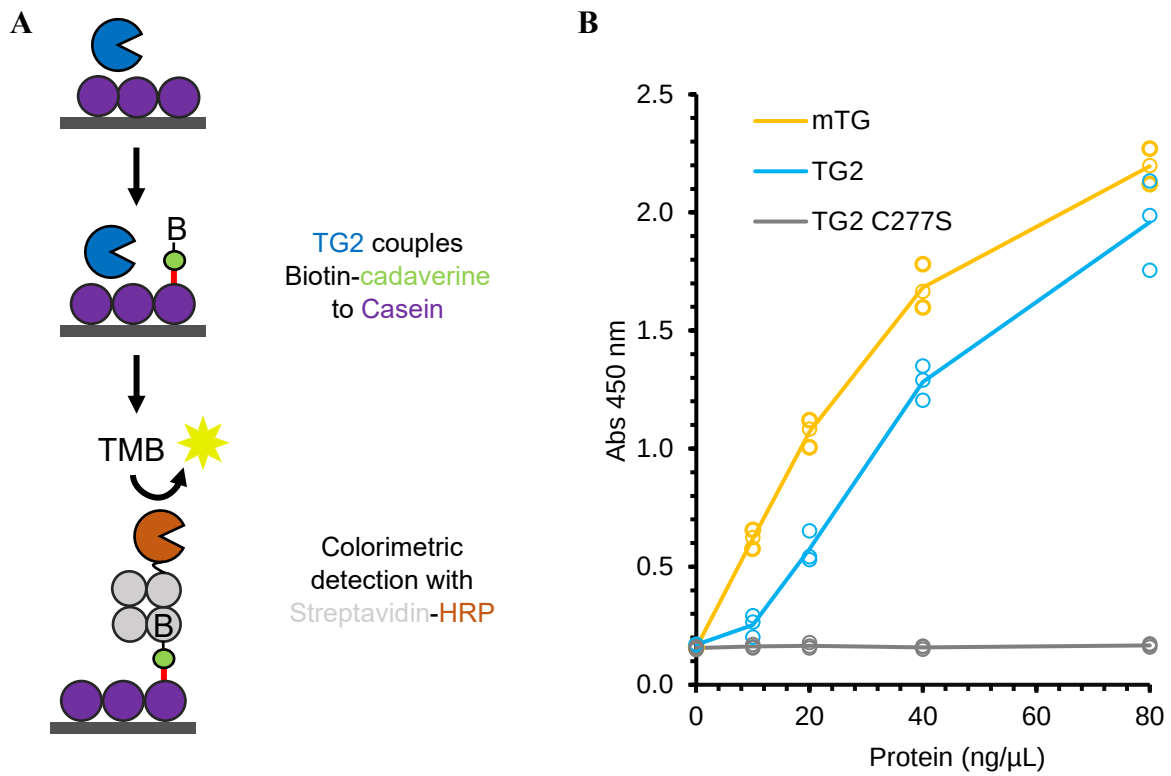


Figure 3.3. Measurement of TG2 transamidase activity using cadaverine. (A) Schematic of the cadaverine assay for transamidase activity of TG2. The reaction is induced by addition of 2 mM $CaCl_2$ and 1 mM DTT, followed by incubation for 30 min at 37 °C. The reaction is quenched by washing away TG2. TMB is tetramethylbenzidine. (B) Cadaverine assay of TG2. Assays were performed with varying concentration of the indicated transglutaminase. Circles show each of the triplicate data points, with the line indicating the mean. One representative preparation shown of three of TG2 and two of TG2 C277S tested in total.

I next analysed some simple TG2 reactions by SDS-PAGE to visualise the reaction products. Samples were treated with hydroxylamine and reducing agent after the reaction and prior to SDS-PAGE to remove any thioester and disulfide bonds, leaving isopeptide linkages intact. The propensity of TG2 to cross-link itself into large covalent aggregates *in vivo* (Stamnaes et al., 2015) may impair directed transamidation of target substrates *in vitro*. To explore this concern, I assayed TG2 with human serum albumin (HSA) to determine if TG2 will preferentially crosslink itself in the presence of a known substrate. After 3 hours at 37 °C the majority of TG2 had formed high molecular weight covalent multimers, indicated by depletion of the monomeric TG2 band and the appearance of a band in the stacking gel. There appears to be minimal depletion of the HSA band, suggesting that TG2 is preferentially crosslinking itself even in the presence of a large excess of HSA (Figure 3.4A).

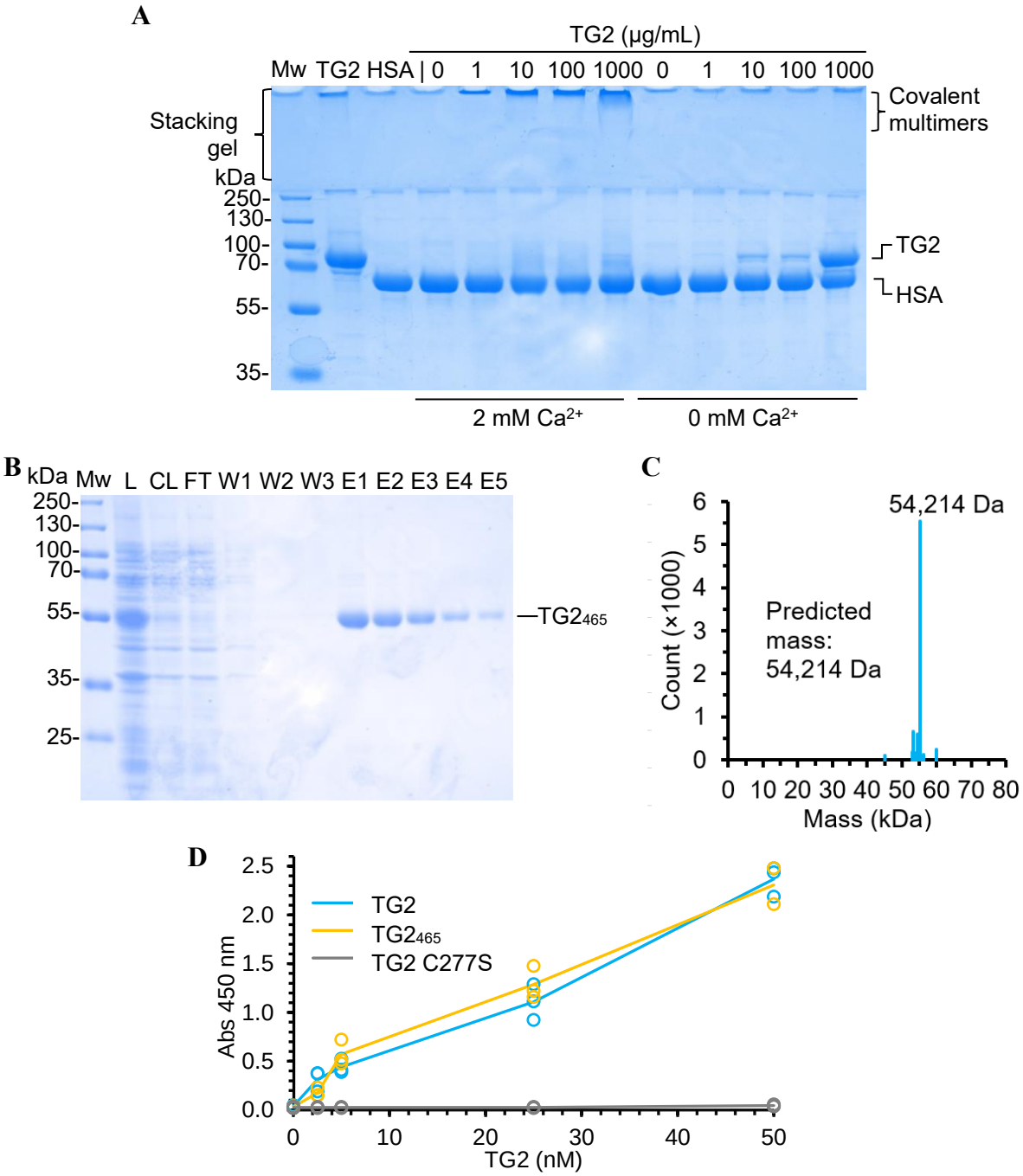


Figure 3.4. C-terminal truncation of TG2. (A) and (B) - analysis by 12% SDS-PAGE with Coomassie staining. All samples were reduced with 100 mM DTT. (A) Transamidation assay of TG2 with known substrate. 15 µM HSA was incubated with the indicated concentration of TG2 for 3 h at 37 °C in TBS pH 7.4, 2 mM $CaCl_2$ and 1 mM DTT. 100 mM hydroxylamine was added to all samples before analysis to deplete the thioester intermediate. One prep performed. (B) Ni-NTA purification of $TG2_{465}$. Column washes performed with 30 mM imidazole. Elution performed with 200 mM imidazole. One representative prep shown of five preparations performed in total. Mw = molecular weight markers, L = cell lysate, CL = cleared lysate, FT = flow-through, W = wash, E = elution.(C) Electrospray ionization mass spectrometry. The 54214 Da product corresponds to $TG2_{465}$- 10×His. (D) Cadaverine assay of truncated TG2. Assays were performed for 30 min at 37 °C with varying concentration of the indicated TG2 variant. Circles show each of the triplicate data points, with the line indicating the mean.

Mass spectrometry of self-crosslinked TG2 has revealed fourteen sites of cross-linking within TG2, eleven of which are in the two C-terminal domains (Stamnaes et al., 2015). A truncated TG2 construct consisting of amino acids 1-465 and without the two C-terminal domains (Figure 1.3A) has previously been shown to abolish self-crosslinking (Stamnaes et al., 2015), but transamidase activity of $TG2_{465}$ was not measured. I generated the truncated TG2 and named it $TG2_{465}$. The $TG2_{465}$ construct was produced at substantially higher yield (8 mg/L) and purity (>98%) than TG2 (Figure 3.4B). The purified product was analysed by mass spectrometry and SEC-MALS to confirm identity (Figure 3.4C) (Appendix figure 1). I tested the activity of $TG2_{465}$ with the cadaverine assay, $TG2_{465}$ showed similar activity to TG2 (Figure 3.4C). However, direct comparisons of activity between TG2 and $TG2_{465}$ cannot be made using the cadaverine assay because the cadaverine assay does not measure total reactivity (including self-reactivity). TG2 and $TG2_{465}$ do not equally self-crosslink, and self-crosslinking of TG2 may compete with aminylation of casein, resulting in lower signal. Thus, it is likely that $TG2_{465}$ has lower overall activity than TG2 but is able to aminylate casein with cadaverine at the same rate as TG2 because it is not performing self-crosslinking.

I then assayed full-length TG2 and $TG2_{465}$ separately in the absence of any other substrate and analysed the reactions with SDS-PAGE (Figure 3.5). After 3 hours incubation at 37 ˚C, a portion of calcium activated TG2 had formed high molecular weight covalent multimers. This finding confirms that full-length TG2 is a substrate for itself in a simple *in vitro* context. In contrast, $TG2_{465}$ did not form any detectable covalent multimers. This result supports previous work that demonstrated $TG2_{465}$ has a considerably lower propensity to crosslink itself, due to removal of the two C-terminal domains (Stamnaes et al., 2015). In summary, TG2 self-crosslinks into covalent multimers in simple *in vitro* assays, even in the presence of a large excess of substrate. Self-crosslinking is diminished in $TG2_{465}$ while considerable activity is retained.

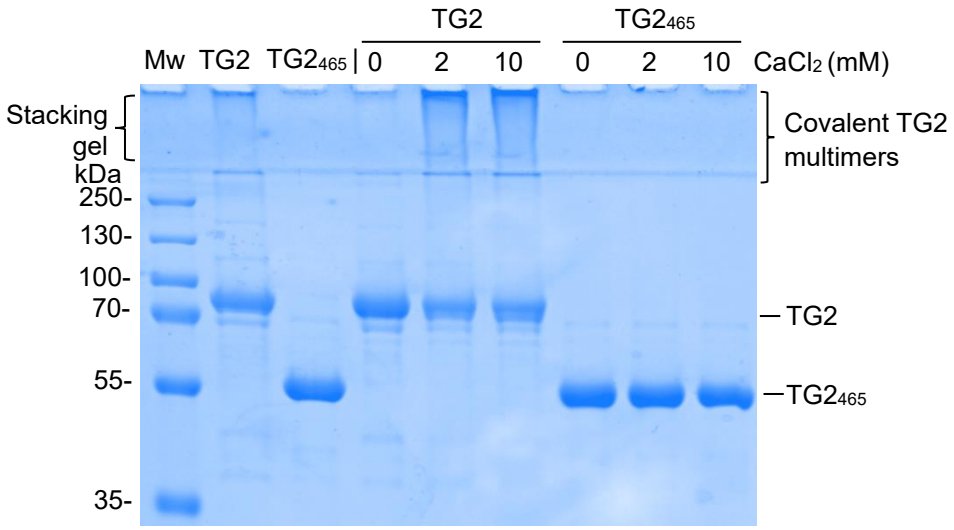


Figure 3.5. C-terminal truncation of TG2 abolishes self-crosslinking. Transamidation assay of TG2 or $TG2_{465}$ alone to determine the extent of self-crosslinking. 15 µM of TG2 was incubated for 30 min at 37 °C in TBS pH 7.4 + 1 mM DTT and the indicated amount of calcium. One representative prep of $TG2_{465}$ shown from five preparations tested in total. Analysis by 12% SDS-PAGE with Coomassie staining. All samples are reduced with 100 mM DTT and thioester intermediates are depleted with 100 mM hydroxylamine before analysis. Mw = molecular weight markers.

3.4. Testing Affinity of $TG2_{465}$ for ECM Proteins and the Cell Surface

Enzyme-mediated protein-protein ligation requires the coordination of at least three structurally complex components. TG2 is known to interact with a range of proteins in the ECM and at the cell surface. Given the concealed nature of the TG2 active site and the importance of steric compatibility between TG2 and its substrates, it is likely that the orientation of TG2 is important for conjugation. Thus, strong non-covalent binding of TG2 to protein surfaces may interfere with the ability of TG2 to covalently conjugate substrates to that surface. Furthermore, in the case of *in vitro* scaffold functionalisation it may be important to be able to wash away TG2 after protein assembly is complete. High non-covalent affinity for scaffold proteins would hinder attempts to remove TG2 from the final product.

I tested the affinity of $TG2_{465}$ for a range of protein surfaces, using a plate assay to gain insight into the propensity of $TG2_{465}$ to interact non-covalently with protein scaffold components. Fibronectin is of interest as a component of tissue scaffolds due to the role it plays

in cell processes such as adhesion, migration, and differentiation (Palomino-Durand et al., 2021). The N-terminal domain of TG2 contains a well characterised fibronectin binding site which confers low nanomolar affinity to fibronectin (Hang et al., 2005) (Cardoso et al., 2015). I incubated $TG2_{465}$ for 2 hours at 37 ˚C in a plate coated with a 45-kda proteolytic fragment of fibronectin derived from human plasma (Figure 3.6A). Wells were washed thoroughly with PBS before detection of residual TG2 with an anti-TG2 antibody. Both TG2 and $TG2_{465}$ were retained in the fibronectin scaffold at low nanomolar concentrations (Figure 3.6B). An anti-TG2 antibody has been identified that is unable to bind TG2 when TG2 is in complex with fibronectin (Iversen et al., 2014). Three residues (K30, R116, and H134) that are important for binding of the antibody are also important for TG2-fibronectin interactions (Cardoso et al., 2015). Mutation of these residues to K30E, R116A, and H134A decreases the affinity of full length TG2 for fibronectin to 2.17 μM (Cardoso et al., 2017). To try to decrease the affinity of $TG2_{465}$ for fibronectin, I introduced these three mutations to the fibronectin-binding domain of $TG2_{465}$ and named the construct $TG2_{465}$ FN. $TG2_{465}$ FN demonstrated impaired fibronectin binding compared to $TG2_{465}$ (Figure 3.6C). However, binding was still detected at mid-to-low nanomolar concentrations for $TG2_{465}$ FN and protein-protein conjugating enzymes are typically used at nanomolar concentrations and above. In conclusion, it is highly likely that $TG2_{465}$ FN would not be easily removable from a fibronectin-based scaffold after directing conjugation.

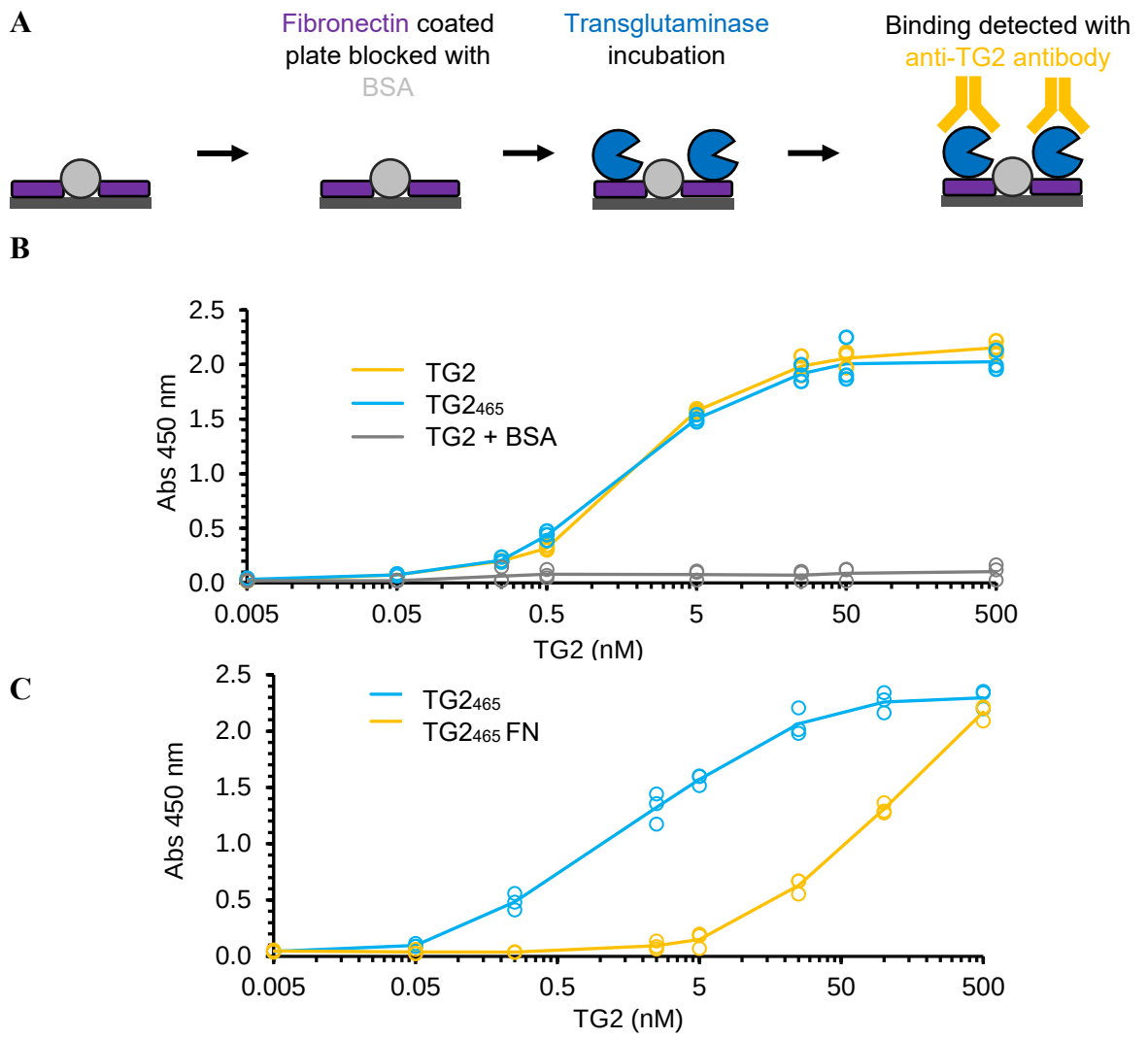


Figure 3.6. Non-covalent interactions between TG2 and fibronectin. (A) Schematic of plate assay used to detect binding. BSA is Bovine Serum Albumin. TG2 is incubated for 2 h at 37 °C. Anti-TG2 antibodies are detected colorimetrically with secondary HRP-conjugated antibody and TMB. (B) Fibronectin binding assay of $TG2_{465}$, performed with varying concentrations of the indicated TG2 variant. Circles show each of the triplicate datapoints with the line indicating the mean. (C) Fibronectin binding assay of $TG2_{465}$ FN performed with varying concentrations of the indicated TG2 variant. Circles show each of the triplicate datapoints with the line indicating the mean.

Collagen is a common component of tissue scaffolds (Smith and Rennie, 2007), and a major component of the ECM (Bosman and Stamenkovic, 2003). TG2 has been shown to colocalise with type VI collagen via the C-terminal domain (Stamnaes et al., 2016) and binding has also been demonstrated (Cardoso et al., 2015) (Stamnaes et al., 2016). However, type VI collagen is not typically used as a component of tissue scaffolds. I used a plate assay to assess the retention of $TG2_{465}$ to a surface of human type I collagen. Retention of detectable levels of

$TG2_{465}$ at a surface of type I collagen was observed at nanomolar concentrations of $TG2_{465}$ (Figure 3.7A). Retention of TG2 and $TG2_{465}$ were equivalent, suggesting that binding does not involve the C-terminal domains (as is the case for type VI collagen). These results suggest that $TG2_{465}$ may not be easily removed from a type I collagen scaffold after directing conjugation.

The basement membrane is a specialized extracellular matrix that surrounds many tissues (Jayadev and Sherwood, 2017). Tissue scaffolds made from basement membrane components are used in cell culture (Chen et al., 2013) and tissue engineering (Nakamura et al., 2017). $TG2_{465}$ may be of interest for functionalising basement membrane scaffolds with proteins such as growth factors and cytokines. MaxGel is a basement membrane extract derived from *in vitro* co-culture of human fibroblasts and human epithelial cells, which has been used for both tissue culture (Lam et al., 2019) and in wound repair (Noorbala et al., 2016). MaxGel contains collagens, laminin, fibronectin, tenascin, elastin, and several proteoglycans. $TG2_{465}$ was incubated on a MaxGel-coated surface for 2 hours at 37 ˚C. Only low levels of $TG2_{465}$ were detected at mid-nanomolar concentrations (Figure 3.7B). Furthermore, retention of $TG2_{465}$ FN was slightly decreased compared to $TG2_{465}$. These results suggest that there would not be significant retention of $TG2_{465}$ FN to a basement membrane scaffold.

For directing conjugation *in vivo*, it is important to try to elucidate how exogenous TG2 would be distributed in the extracellular space. Given that TG2 has known high affinity interactions with ECM components, TG2 may localise in the ECM. Alternatively, TG2 may localise predominantly at the cell surface given the known interactions with integrins and syndecan-4. The localisation of TG2 will determine which endogenous proteins TG2 is exposed to and able to direct conjugation with. Whether TG2 binds to the cell surface or the ECM, or both, may influence application.

A431 cells are a human epidermoid carcinoma cell line. I incubated TG2 with A431 cells for 30 min at 37 ˚C without calcium or reducing agent and subjected the cells to thorough

washes with PBS. Cell-surface bound TG2 was detected with an anti-TG2 antibody using flow cytometry. Both TG2 and $TG2_{465}$ were detected at the cell surface after washing, demonstrating high affinity for the cell surface (Figure 3.7C). $TG2_{465}$ showed slightly weaker binding compared to TG2, suggesting that the C-terminal domains contribute to cell-surface binding. I then tested cell surface binding for $TG2_{465}$ FN and found that it was only slightly decreased (Figure 3.7D). These results demonstrate that exogenous $TG2_{465}$ binds robustly to the cell surface, and mutations to the fibronectin binding domain do not substantially disrupt binding. Based on these results, it is highly likely that at least some $TG2_{465}$ would be non-covalently anchored to the cell surface if applied *in vivo* for conjugation in the extracellular space. This may allow for directed conjugation to the cell surface. It must be noted that the TG2 variants used in these plate assays contained a His tag, and that a His tag control was not performed in this experiment. The possibility of the above result being due to affinity for the His tag cannot be ruled out.

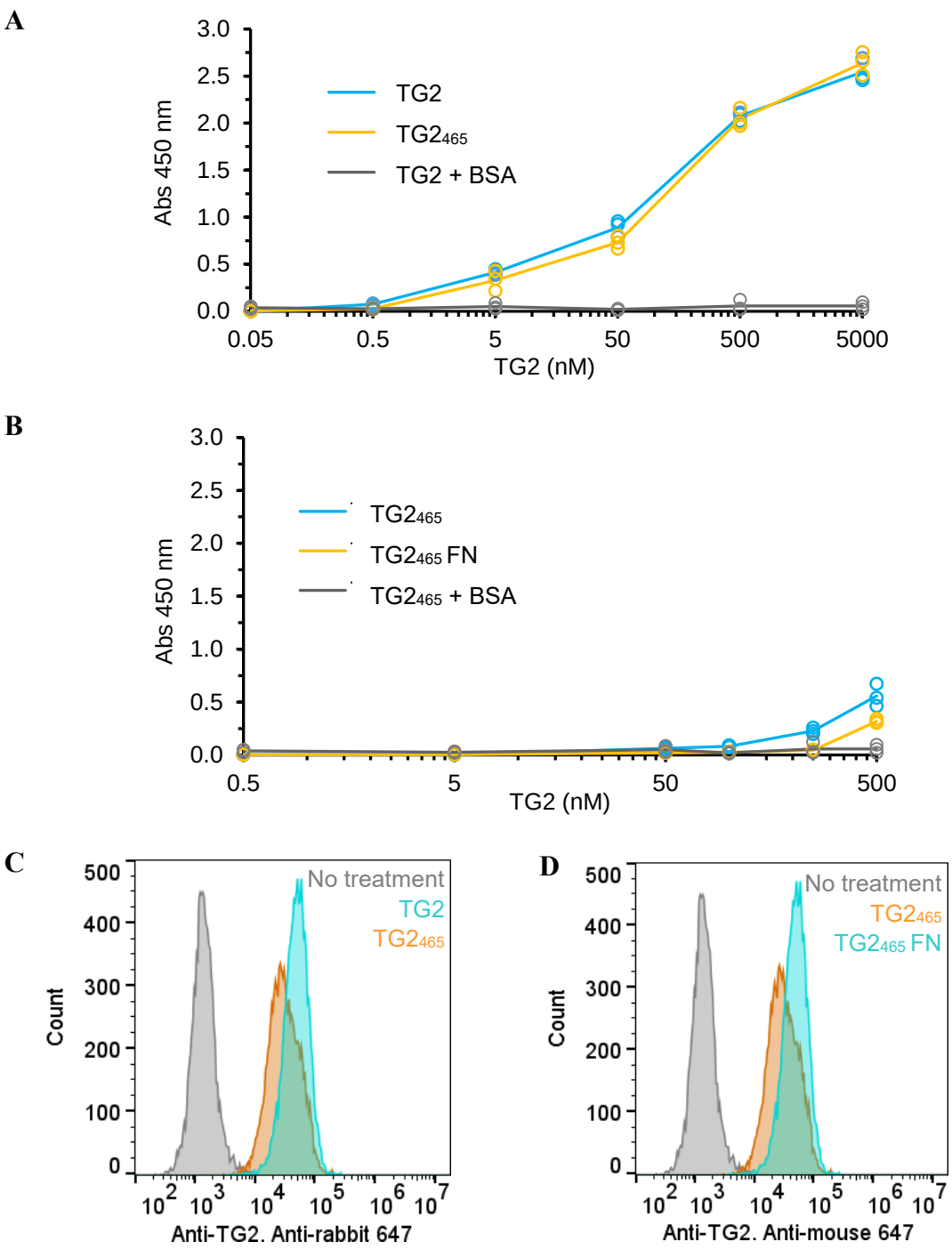


Figure 3.7. Non-covalent interactions between TG2 and components of the extracellular space. (A) Collagen binding assay of TG2. Performed as for the fibronectin binding assay but with human Type I collagen. Assay performed with varying concentrations of the indicated TG2 variant. Circles show each of the triplicate datapoints with the line indicating the mean. (B) MaxGel binding assay of $TG2_{465}$ FN. Assay performed with varying concentrations of the indicated TG2 variant. Circles show each of the triplicate datapoints with the line indicating the mean. (C-D) TG2 cell binding assays. 6 µM of the indicated protein was incubated with 10^7 A431 cells for 30 min at 37 °C in HBS pH 7.4. Cells were washed in HBS pH 7.4 + 1% (w/v) BSA, then stained with rabbit anti-TG2 antibody and anti-rabbit 647 antibody conjugate and analysed by flow cytometry.

3.5. Engineering Redox Regulation of $TG2_{465}$

The redox dependence of TG2 transamidation has implications for both *in vitro* and *in vivo* application. The requirement for addition of a reducing agent is a challenge for *in vivo* conjugation in the oxidising extracellular space, and for *in vitro* conjugation where substrates have important disulfide bonds. The kinetics of TG2 redox regulation are currently unknown. If the regulatory disulfide bond(s) of TG2 form on the order of minutes or hours, then there may be enough residual transamidase activity in a reduced sample of TG2 to perform sufficient conjugation when added to non-reducing conditions before TG2 is inactivated.

Two disulfide bonds have been resolved in crystal structures of TG2: C230-C370 and C370-C371 (Jeong et al., 2020) (Pinkas et al., 2007) (Figure 1.7). Furthermore, mass spectrometry of oxidised TG2 has detected C230-C370 and C370-C371 bonds, but not C230-C371 bonds or any disulfide bonds involving the catalytic cysteine (Stamnaes et al., 2010). Considering these findings, it has been hypothesised that oxidative inactivation of TG2 is due to oxidation of C370 which can form a disulfide bond with either C370 or C371 bond. Therefore, it may be possible to prevent oxidative inactivation of TG2 by mutating these key cysteine residues.

To disrupt formation of the regulatory disulfide bond, I generated a range of mutant $TG2_{465}$ constructs where disulfide bond-forming cysteine residues were replaced with serine. Substitution to serine was preferred to maintain charge and size. The variants were purified to >95% purity (Appendix Figure 2). I then tested transamidase activity of the cysteine mutants with the cadaverine assay. $TG2_{465}$ C370S and $TG2_{465}$ C371S had mean activity of 76% (S.D. = 5.1%) and 69% (S.D. = 1.9%) relative to $TG2_{465}$, while $TG2_{465}$ C230S showed mean activity of 116% (S.D. = 5.6%) relative to $TG2_{465}$ (Figure 3.8A). There are two possible explanations for the observed differences in activity. One explanation is that mutating these residues, which are near the active site, alters the enzyme kinetics of TG2. Another explanation is that mutating

these residues disrupts the kinetics of redox regulation, resulting in the enzyme becoming activated more slowly or more quickly. Kinetic assays are required to establish if the k_{cat} of these variants differ.

I then used a previously developed plate assay to measure activity of $TG2_{465}$ constructs at set timepoints after removal of reducing agent (Figure 3.8B). The activity of all TG2 variants was quickly impaired upon exposure to non-reducing conditions (Figure 3.8C). While the $TG2_{465}$ C370S and $TG2_{465}$ C371S were equivalent to $TG2_{465}$, the C230S mutant demonstrated markedly slower inactivation after removal of reducing agent. $TG2_{465}$ C230S retained 24% activity after 5 min without reducing agent compared to 4% activity for $TG2_{465}$ (Figure 3.8D). The C370S mutation, which abolishes both the C230-C370 and the C370-C371 bond, did not result in prolonged activity in non-reducing conditions. This suggests that other regulatory disulfide bonds can form, whether this occurs only in the absence of the other bonds or in addition to them is unclear. Given the proximity of the catalytic cysteine to other cysteine residues, it is feasible that the catalytic residue itself can form intramolecular disulfide bonds resulting in inactivation.

In summary, $TG2_{465}$ loses almost all activity within 1 minute of removal of reducing agent. This renders $TG2_{465}$ unsuitable for conjugation in non-reducing conditions. $TG2_{465}$ C230S can retain some activity for at least 30 minutes after removal of reducing agent, which may allow a reasonable level of conjugation in non-reducing conditions.

Previous work has shown that the C230-C370 bond forms more quickly in oxidising conditions than the C370-C371 bond, suggesting it is more stable (Stamnaes et al., 2010). One explanation for the observed differences in redox regulation of $TG2_{465}$ C230S and $TG2_{465}$ is that the C230S mutant cannot form the highly stable C230-C370 bond, but can form the less stable C370-C371 bond. The less stable regulatory disulfide bond takes longer to form and is more quickly broken. To test this hypothesis, I performed the reverse version of the previous

assay. $TG2_{465}$ constructs were fully oxidised by prolonged incubation without reducing agent, and then activity measured at set timepoints after addition of reducing agent (Figure 3.8E). Both $TG2_{465}$ and $TG2_{465}$ C230S fully recovered activity over time, but $TG2_{465}$ C230S became fully active within 1 minute while $TG2_{465}$ activity took 15 minutes to fully recover. So $TG2_{465}$ C230S was activated more rapidly and inactivated more slowly than $TG2_{465}$. These results support the hypothesis that $TG2_{465}$ C230S is regulated by a less stable disulfide bond than $TG2_{465}$.

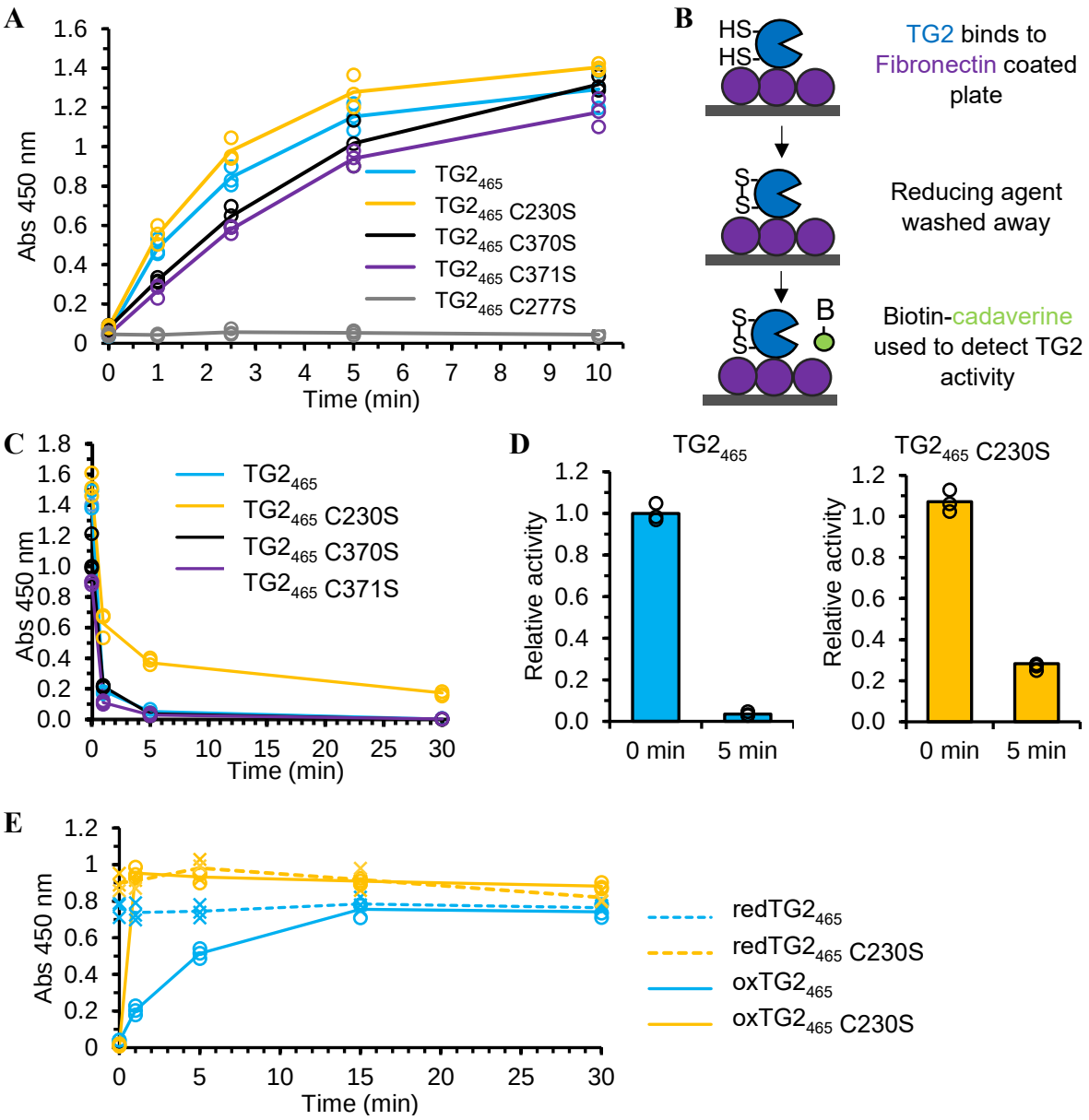


Figure 3.8. Exploring redox regulation of TG2$_{465}$ and variants. (A) Cadaverine assay of TG2 variants. Assays performed with varying concentration of the indicated TG2 variant. Circles show each of the triplicate data points, with the line indicating the mean. (B) Schematic of the plate assay used to test transamidase activity after removal of reducing agent. 50 nM of reduced TG2 is bound to a fibronectin-coated plate. Reducing agent is washed away using HBS. TG2 is then incubated in HBS pH 7.4 at 37 °C for the indicated duration. At the desired end point, TG2 activity is measured by addition of 2 mM $CaCl_2$ and 1 mM Biotin-cadaverine. Incorporation of Biotin-cadaverine into fibronectin is detected colorimetrically with streptavidin-HRP and TMB (C) Oxidation assay of TG2 variants. Proteins were incubated without reducing agent for the indicated duration. Circles show each of the triplicate data points, with the line indicating the mean. (D) Reduction assay in (C) comparing just 0 min and 5 min timepoints. Circles show each of the triplicate data points, with the line indicating the mean. (E) Reduction assay of TG2 variants. Proteins were oxidised by incubation in HBS pH 7.4 for 90 min at 4 °C prior to the assay, then reduced in a plate assay by incubation in HBS pH 7.4 + 1 mM DTT for the indicated duration. Circles show each of the triplicate data points, with the line indicating the mean. Note – datapoints in C-E are normalized against signal for a TG2$_{465}$ C277S control sample.

3.6. Peptide Tags Facilitate TG2-mediated Protein-Protein Conjugation

The moderate specificity of TG2 for glutamine donors and low specificity for lysine donors is well suited for the covalent decoration of protein surfaces with cargo proteins. High specificity for glutamine residues may be leveraged to direct site-specific conjugation of a glutamine-tagged cargo protein, while low specificity for lysine residues may be utilised to allow conjugation of the cargo to endogenous proteins at many sites to maximise the level of scaffold decoration (Figure 3.9).

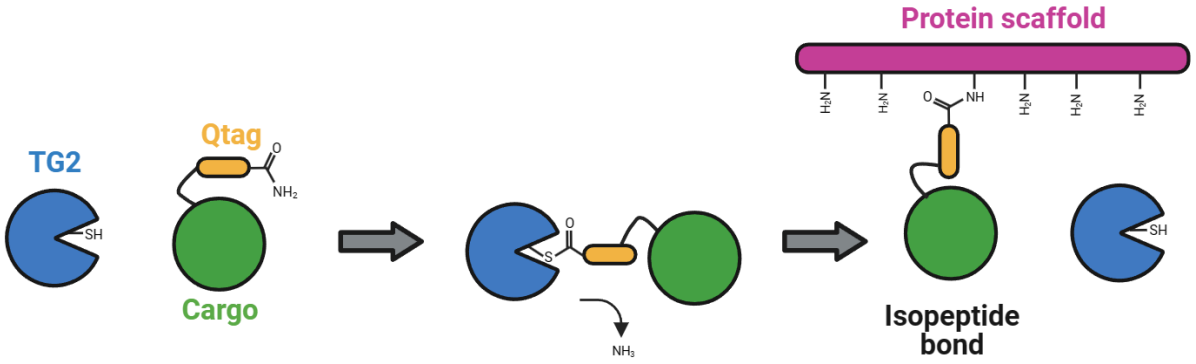


Figure 3.9. Using a glutamine peptide tag to direct conjugation of a protein to an endogenous protein scaffold. The protein of interest is genetically fused to a short TG2-reactive glutamine peptide. Activated TG2 forms a thioester intermediate with the glutamine tag (Qtag). The thioester is subsequently attacked by the lysine sidechain of an endogenous protein scaffold, resulting in covalent conjugation of the cargo protein to the scaffold.

In previous work, short glutamine tags were fused to the N-terminus of Glutathione-S-transferase to impart reactivity for TG2 (Sugimura et al., 2006). These tags were used for facilitating TG-mediated aminylation with small molecular probes. It is not known if these tags are compatible for transamidation to lysine residues in the context of a protein. One of the tags used in that work is a motif from the TG2 substrate human αB crystallin (TVQQEL). I fused the TVQQEL motif (named Qtag) to model proteins to impart TG2 reactivity and performed simple *in vitro* transamidation assays to try to establish TG2-mediated site-specific conjugation. The TVQQEL motif was chosen over alternative tags because of the short length and human origin. Human origin is important for minimising the likelihood of immunogenicity issues for *in vivo* therapeutic applications.

Herein the suffix '-Q' denotes genetic fusion of the Qtag at the C-terminus of a protein. I first assayed TG2 with mClover and mClover-Q. SDS-PAGE analysis revealed novel products in the mClover-Q assay which were not present when TG2 was assayed with mClover (Figure 3.10A). One of the novel bands corresponds in size with an mClover-Q homodimer (58 kDa). The observation of mClover-Q homodimers demonstrates that TG2 can conjugate the Qtag to endogenous lysine residues of mClover. The identity of the product observed at ~80 kDa in the $TG2_{465}$ assay was unclear, as this could have been an mClover-Q homotrimer or a $TG2_{465}$ and mClover-Q heterodimer ($TG2_{465}$:mClover-Q). I repeated the $TG2_{465}$ assay with addition of a competitor amine (biotin-cadaverine) (Figure 3.10B). The addition of biotin-cadaverine inhibited the formation of the 58 kDa product, confirming that it is the product of an ε-(γ-glutamyl) lysine isopeptide linkage. Western blotting of the $TG2_{465}$ assay with an anti-TG2 antibody (Figure 3.10C) and an anti-GFP antibody (Figure 3.10D) confirmed that the ~80kDa band was a $TG2_{465}$:mClover-Q dimer (Figure 3.10C). The presence of $TG2_{465}$:mClover-Q dimers indicates that $TG2_{465}$ also contains at least one reactive lysine that is compatible with the Qtag. This may be problematic when directing conjugation of Q-tagged proteins to scaffolds because transamidation of Qtag and $TG2_{455}$ may be favoured over transamidation of mClover-Q and scaffold.

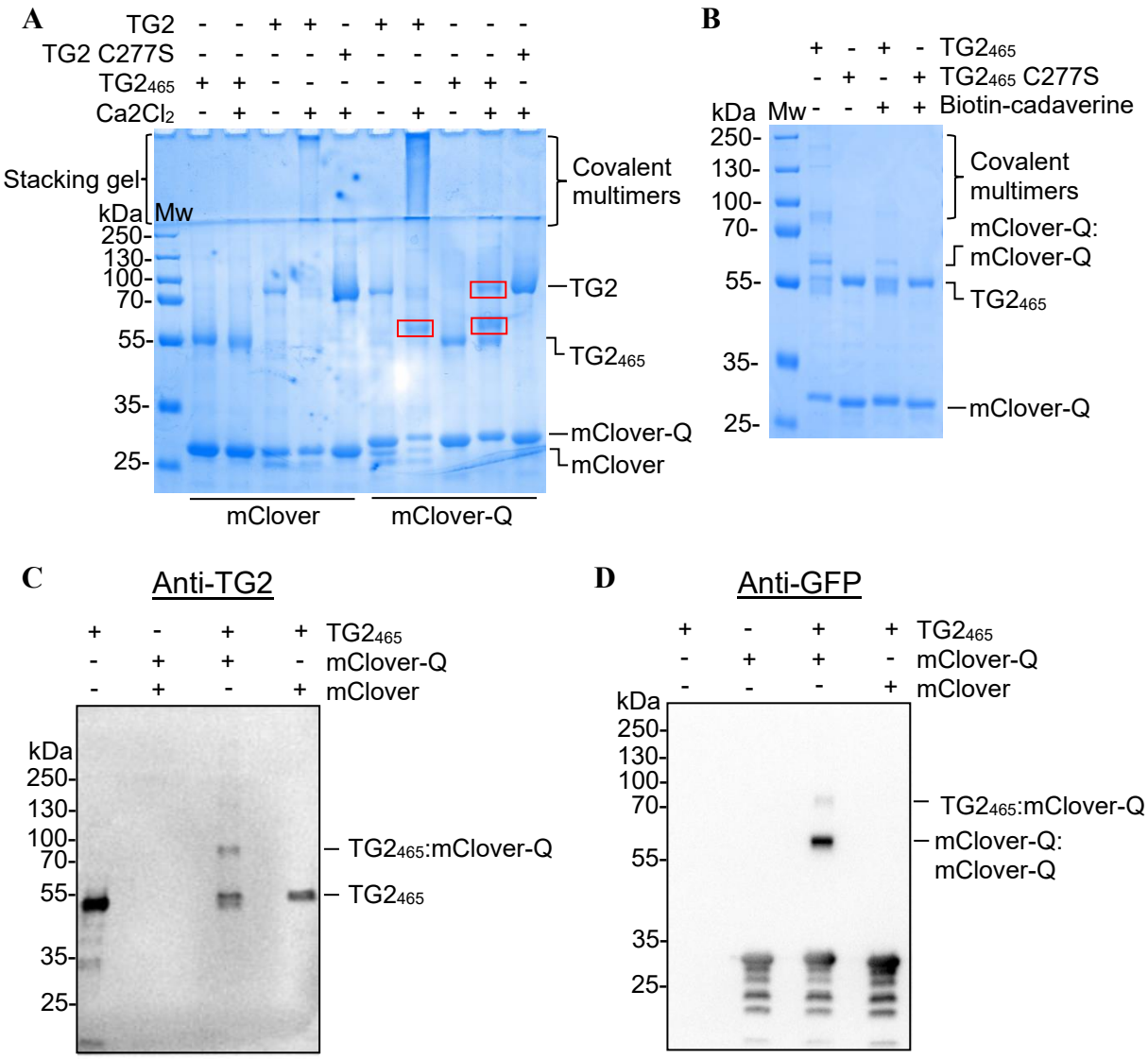


Figure 3.10. Introduction of Qtag to mClover to facilitate TG2-mediated conjugation. (A) and (B) - analysis by 12% SDS-PAGE with Coomassie staining, all samples are reduced with 100 mM DTT before analysis. Thioester intermediates are depleted with 100 mM hydroxylamine before analysis. Mw = molecular weight markers. (A) Transamidation assay of TG2 variants with mClover-Q. 7 µM TG2 and 15 µM mClover-Q were incubated for 3 h at 37 °C in TBS pH 7.4, 2 mM $CaCl_2$ and 1 mM DTT. Red rectangles highlight protein conjugates. (B) Transamidation assay of $TG2_{465}$ and mClover-Q with a competing amine. 7 µM TG2 and 15 µM mClover-Q were incubated for 3 h at 37 °C in TBS pH 7.4, 2 mM $CaCl_2$, 1 mM DTT and 1 mM Biotin-cadaverine. Analysis by SDS-PAGE with Coomassie staining, 100 mM each of DTT and hydroxylamine was added to all samples. (C) Western blot of samples from (A) with mouse anti-TG2 antibody and anti-mouse HRP conjugate. (D) Western blot of samples from (A) with mouse anti-GFP antibody and anti-mouse HRP conjugate.

After confirming that $TG2_{465}$ can conjugate the Qtag in a simple single-substrate assay, I attempted to perform transamidation in a more complex environment i.e. at the cell surface. $TG2_{465}$ and mClover-Q constructs were incubated with Expi293F cells for 30 min at 37 ˚C in the presence of 2 mM calcium. After thorough washing in PBS to remove components of the reaction, cells were analysed by flow cytometry. Incubation of $TG2_{465}$ C230S and mClover-Q together resulted in increased signal in the GFP channel compared to mClover-Q alone, suggesting that TG2 facilitated decoration of the cell surface with mClover-Q (Figure 3.11A). In contrast, incubation of $TG2_{465}$ C277S and mClover-Q together did not result in increased GFP signal at the cell surface. This suggests that TG2-mediated localisation of mClover-Q to the cell surface is dependent upon transamidase activity. This also confirms that binding to the cell surface is not due to the His tag, which is present on the TG2 and mClover variants. Analysis of the same reactions with detection of $TG2_{465}$ using an anti-TG2 antibody revealed that $TG2_{465}$ C230S and $TG2_{465}$ C277S are also present at the cell surface (Figure 3.11B). This demonstrates that, in the presence of the mClover-Q substrate, $TG2_{465}$ binds non-covalently to the cell surface in a manner independent of transamidation. In summary, $TG2_{465}$ facilitated decoration of the cell surface with mClover-Q in a manner that was dependent upon transamidase activity. One possibility is that mClover-Q was covalently conjugated to the cell surface. However, $TG2_{465}$ was also found at the cell-surface. Furthermore, earlier assays demonstrated that $TG2_{465}$ readily conjugates itself to mClover-Q (Figure 3.10). An alternative explanation for the observation of mClover-Q at the cell surface is that $TG2_{465}$ is forming $TG2_{465}$:mClover-Q dimers, and these dimers are non-covalently bound to the cell surface via TG2-mediated interactions.

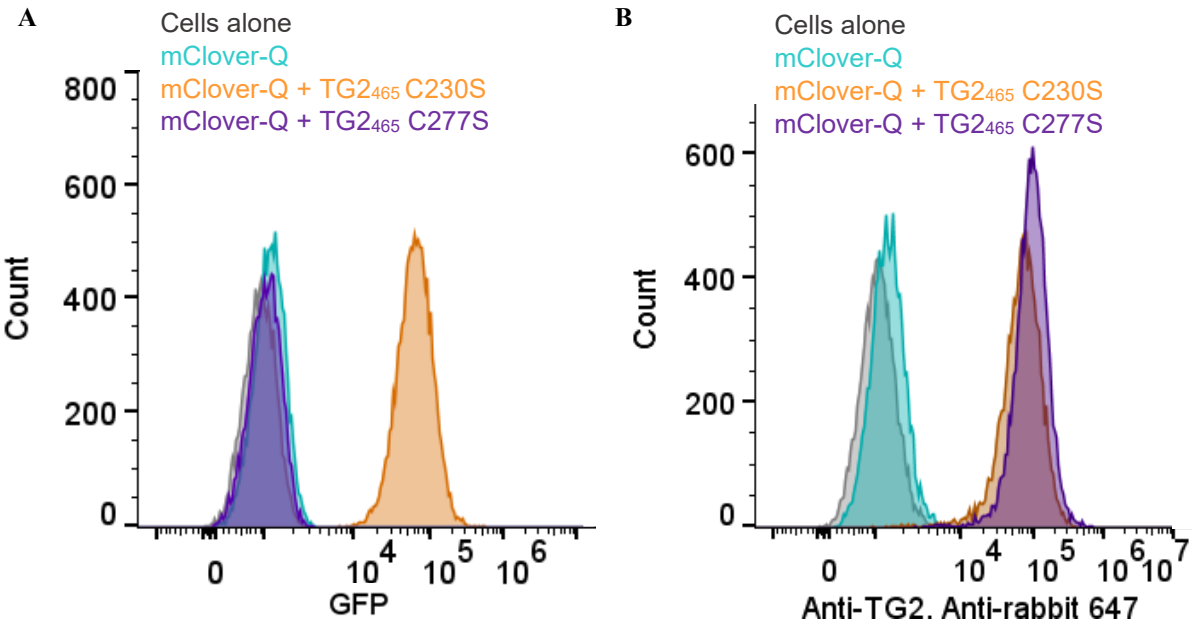


Figure 3.11. TG2-mediated conjugation of mClover-Q to the cell surface. (A-B) 5 µM of the indicated $TG2_{465}$ variant and 10 µM of mClover-Q were incubated with 10^6 Expi293F cells for 30 min at 37 °C in HBS pH 7.4, 2 mM $CaCl_2$ and 1 mM DTT. Cells were then washed in HBS pH 7.4 + 1% (w/v). Cells were stained with rabbit anti-TG2 antibody and anti-rabbit 647 antibody. (A) Cells analysed using 561-585 nm emission channel to detect mClover-Q. (B) Cells analysed using 647 nm emission channel to detect TG2.

To determine if mClover-Q interactions with the cell-surface are covalent or non-covalent I used Western blotting. Firstly, to test if $TG2_{465}$ is active in the cell assay conditions I incubated $TG2_{465}$ C230S with Expi293F cells and biotin-cadaverine for 30 min at 37 ˚C in the presence of 2 mM calcium. Cells were thoroughly washed in PBS to remove reaction components and the cell lysate from these reactions was blotted with a streptavidin-HRP conjugate. Western blotting revealed aminylation of cell-surface proteins with biotin-cadaverine (Figure 3.12A). This result confirms that $TG2_{465}$ C230S is active in these reaction conditions and can utilise cell-surface proteins as substrates.

I then attempted transamidation at the cell-surface by assaying $TG2_{465}$ C230S and mClover-Q with Expi293F cells for 30 min at 37 ˚C in the presence of 2 mM calcium. Western blotting of the cell lysate from these reactions with an anti-GFP antibody did not detect any mClover-Q conjugates (Figure 3.12B). One explanation for this result is that the Qtag is not

compatible with cell-surface lysine residues, an unlikely scenario given the low specificity of TG2 for lysine substrates. A more likely explanation is that $TG2_{465}$ favours formation of mClover-Q homodimers and $TG2_{465}$ C230S:mClover-Q heterodimers over conjugation of mClover-Q to cell surface proteins. Western blotting of the same assays with an anti-TG2 antibody revealed a band for both monomeric $TG2_{465}$ C230S and monomeric $TG2_{465}$ C277S (Figure 3.12C). The presence of the catalytically inactive $TG2_{465}$ C277S in the cell lysate suggests strong non-covalent association with the cell surface at high enough affinity to resist the washes used in the protocol. These findings corroborate the earlier observation of high affinity of $TG2_{465}$ for the cell surface by flow cytometry (Figure 3.11B). Interestingly, when $TG2_{465}$ C230S was assayed with mClover-Q there was a faint band at the correct size for a $TG2_{465}$ C230S:mClover-Q dimer. This suggests that $TG2_{465}$ C230S conjugated itself to mClover-Q, and then this heterodimer non-covalently bound to the cell surface.

Taken together, flow cytometry and Western blotting analysis of $TG2_{465}$ and mClover-Q cell assays demonstrate high-affinity decoration of the cell surface with $TG2_{465}$:mClover-Q dimers and no detectable conjugation of mClover-Q to cell-surface proteins.

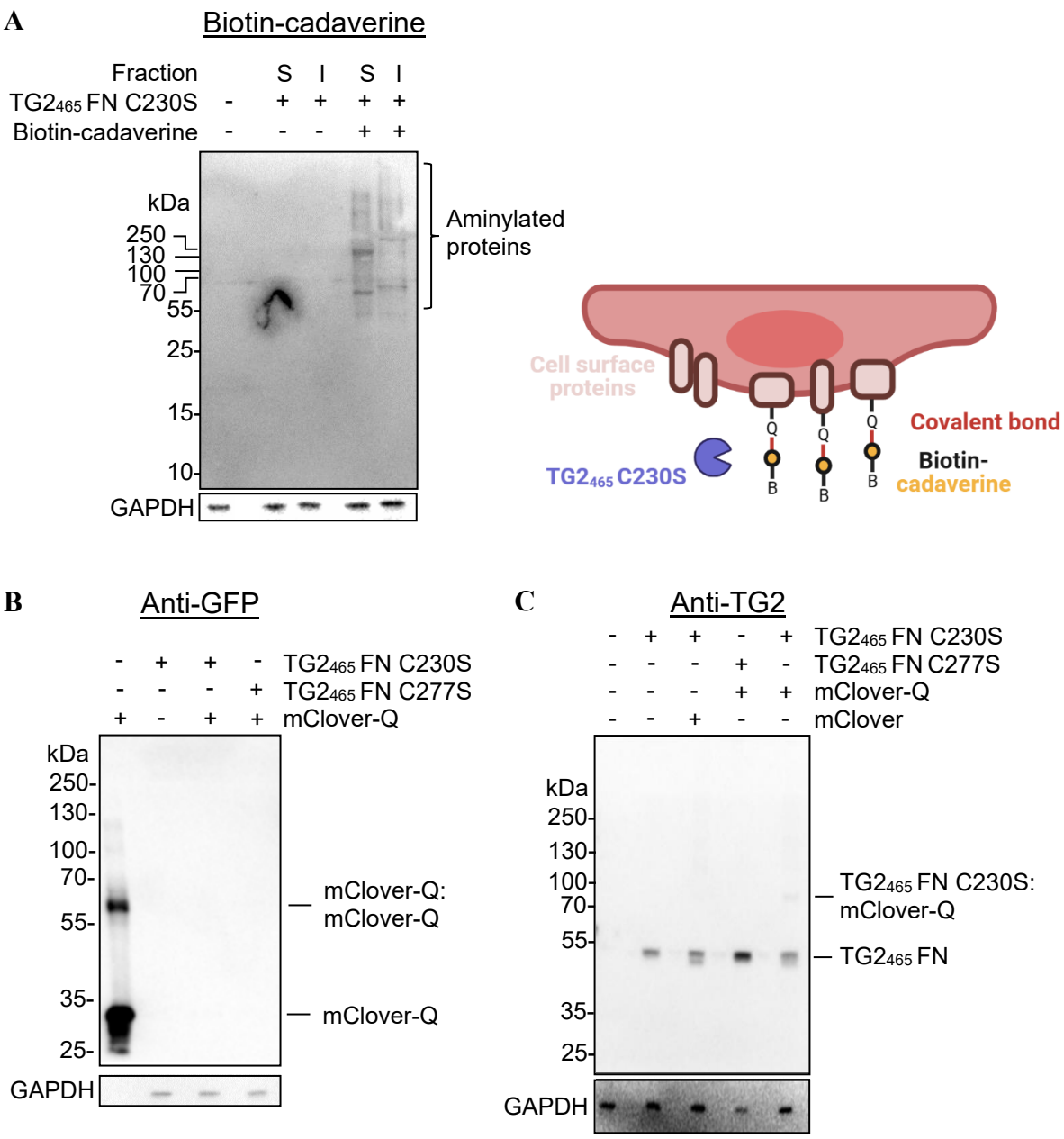


Figure 3.12. Transamidation at the cell surface. (A-C) are Western blots of 12% SDS-PAGE analysis. All samples were reduced with 100 mM DTT before analysis. Thioester intermediates were depleted with 100 mM hydroxylamine before analysis. Separate blots were performed with an anti-GAPDH antibody as a sample processing control. (A) Transamidation assay with biotin-cadaverine at the cell surface. 5 µM of TG2 was incubated with 10^7 Expi293F cells for 30 min at 37 °C in HBS pH 7.4, 10 mM Biotin-cadaverine, 2 mM $CaCl_2$ and 1 mM DTT. S and I indicates whether sample was taken from supernatant or insoluble fraction. Biotin-cadaverine was detected with a streptavidin-HRP. (B) Transamidation assay with mClover-Q at the cell surface. 5 µM of TG2 and 5 µM mClover-Q was incubated with 10^6 Expi293F cells for 30 min at 37 °C in HBS pH 7.4, 2 mM $CaCl_2$ and 1 mM DTT. mClover-Q was detected with a mouse anti-GFP antibody and an anti-mouse HRP conjugate. (C) Western blot of assay in (B) to detect TG2 with mouse anti-TG2 antibody and anti-mouse HRP conjugate.

After failing to produce detectable levels of conjugation of a Q-tagged model protein to the cell surface, I developed an alternative 'anchoring' method in which the protein of interest is fused to TG2 along with the Qtag (Figure 3.13). In this system, TG2 conjugates itself to the scaffold by catalysing transamidation between its own Qtag and lysine residues of the scaffold. Given that TG2 can conjugate the Qtag to its own lysine residues, it is likely that the introduction of the Qtag to TG2 itself will cause some self-multimerization. However, simplifying the reaction from the coordination of three separate molecules to just two molecules may increase the probability of reaction occurring and produce detectable levels of conjugation.

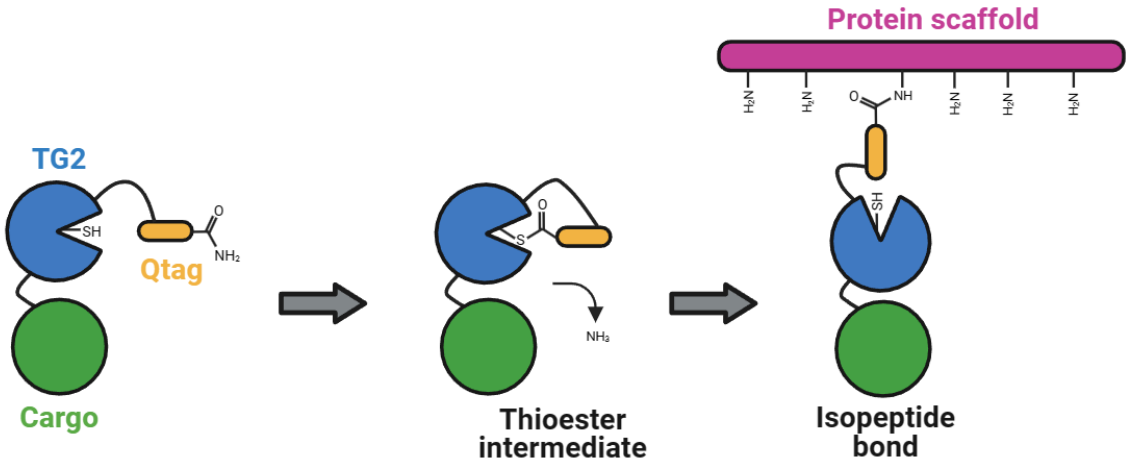


Figure 3.13. The anchoring method of protein conjugation. The protein of interest is fused to TG2. The TG2-reactive Qtag is genetically fused to TG2. Activated TG2 forms a thioester intermediate with its own Qtag. The thioester is subsequently attacked by the lysine sidechain of an endogenous protein scaffold, resulting in covalent conjugation of TG2 and cargo to the scaffold.

I genetically fused Qtag at the C-terminus of $TG2_{465}$ FN. I also used the motif GQDPVK from Trappin as a lysine-bearing tag and named it LysTag. GQDPVK has been used previously as a TG2 substrate in peptide form (Ruoppolo et al., 2003), but not as a tag to introduce reactivity to a protein substrate. Herein, the suffix -K denotes genetic fusion of the LysTag at the C-terminus of a protein. The $TG2_{465}$ variants were purified to >95% purity (Appendix Figure 3). Both $TG2_{465}$ FN-Q and $TG2_{465}$ FN-K display similar levels of transamidase activity to $TG2_{465}$ as determined by cadaverine assay (Figure 3.14A).

To determine if tagging TG2 with Qtag or LysTag results in self-multimerization I incubated each construct on its own for 3 hours at 37 ˚C with 2 mM calcium and 1 mM DTT. As expected, both $TG2_{465}$ FN-Q and $TG2_{465}$ FN-K formed high molecular weight covalent multimers (Figure 3.14B). This result demonstrates that $TG2_{465}$ FN readily conjugates itself to both tags in the absence of other substrates. However, in the presence of other substrates it may still be possible to direct conjugation elsewhere.

To test if tagged-TG2 can covalently conjugate itself to endogenous proteins, I assayed $TG2_{465}$ FN-Q and $TG2_{465}$ FN-K with Expi293F cells for 30 min at 37˚C in the presence of 2 mM calcium. Western blotting of the cell lysates from the reactions with an anti-TG2 antibody detected the monomeric forms of $TG2_{465}$ FN constructs but no covalent multimers (Figure 3.14C). The presence of monomeric $TG2_{465}$ C277S in the blot, which is catalytically inactive, confirms that monomeric $TG2_{465}$ constructs in the blot are the result of non-covalent binding to the cell surface. In conclusion, introduction of Qtag or LysTag to $TG2_{465}$ did not facilitate TG2-mediated covalent anchoring of $TG2_{465}$ to the cell surface.

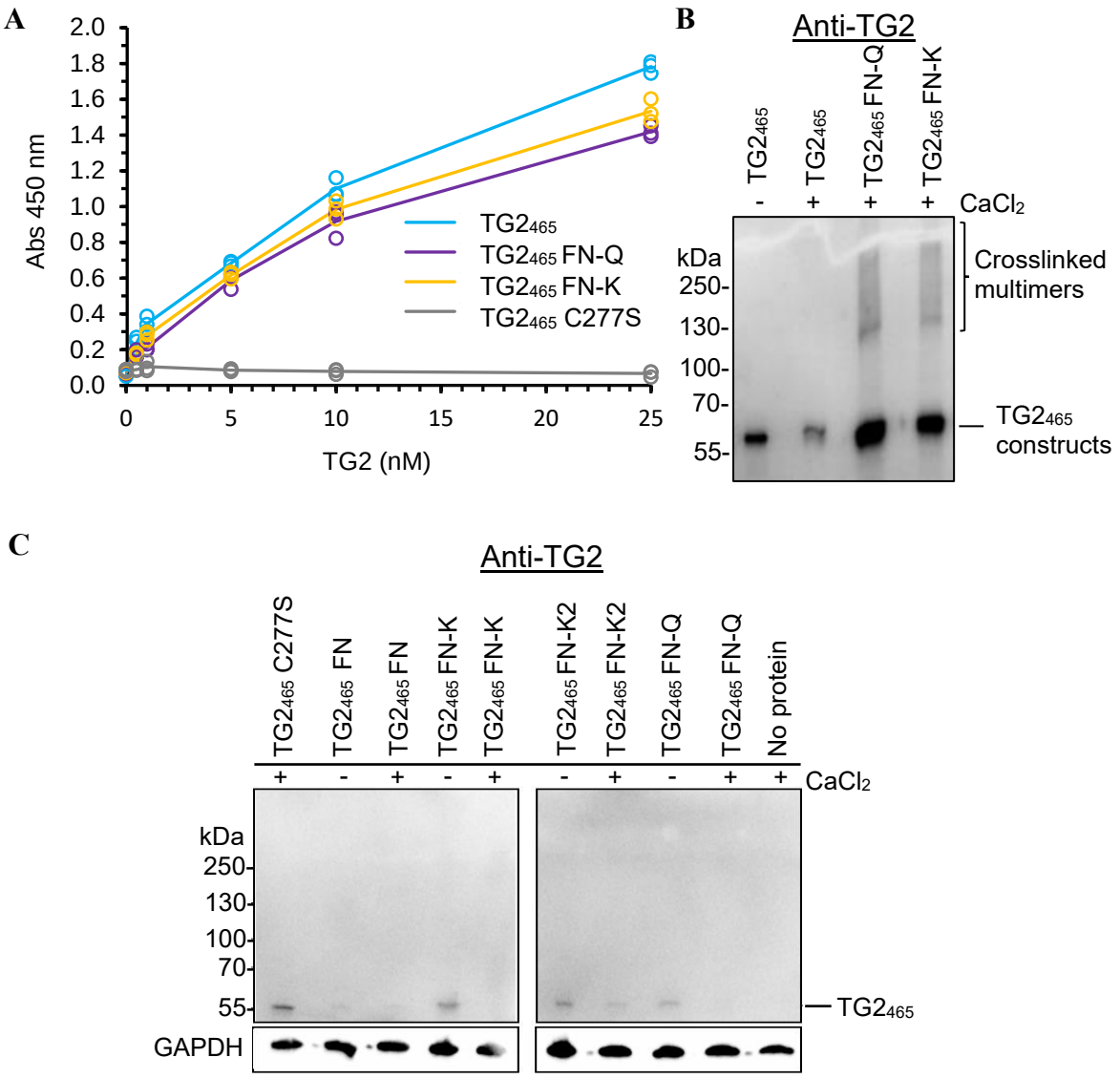


Figure 3.14. Introduction of TG2-reactive tags to TG2. (A) Cadaverine assay of TG2 variants. Assays were performed with varying concentration of the indicated TG2 variant. Circles show each of the triplicate data points, with the line indicating the mean. (B) and (C) are Western blots of 12% SDS-PAGE analysis, blotted with mouse anti-TG2 antibody and anti-mouse HRP conjugate to detect TG2. All samples were reduced with 100 mM DTT before analysis. Thioester intermediates were depleted with 100 mM hydroxylamine before analysis. (B) Transamidation assay of $TG2_{465}$ variants alone to determine the extent of self-crosslinking. 10 µM of the indicated TG2 construct was incubated for 3 h at 37 °C in HBS pH 7.4, 2 mM $CaCl_2$ and 1 mM DTT. Western blot to detect TG2 with mouse anti-TG2 antibody and anti-mouse HRP conjugate. (C) Transamidation assay of tagged $TG2_{465}$ FN at the cell surface. 7 µM of the indicated TG2 construct was incubated with 10^6 Expi293F cells for 30 min at 37 °C with 2 mM $CaCl_2$ and 1 mM DTT. A separate blot was performed with an anti-GAPDH antibody as a sample processing control.

3.7. Discussion

In this chapter I established a protocol for expression and purification of $TG2_{465}$ at high purity and moderate yield (8 mg/L) in *E. coli*, which makes a TG2-based tool convenient to produce. While there are some reports of solubility issues and low yield of recombinant TG2 when expressed in *E. coli* (Roy et al., 2013) (Shi et al., 2002), yields of ~6 mg/L have been achieved previously (Piper et al., 2002). Issues with premature termination of transcription limited yield in this thesis and thus there is scope to further improve expression. Sense codons differing from the stop codon in the wobble position have been shown to cause premature termination of transcription (Freistroffer et al., 2000), initiating termination almost as efficiently as the UAA codon. The TG2 constructs used in this thesis have 10 UAU codons in total, with the codons corresponding to Y168 and Y369 being of particular interest. Premature termination of transcription at Y168 for the TG2 construct used in Figure 3.1 and 3.2A would result in a 20.8 kDa product, which corresponds to lower impurity band. Premature termination of transcription at Y369 for the same construct would result in a 43.8 kDa product, which corresponds to the higher impurity band. These observations suggest it is likely that UAU codons in my TG2 constructs are initiating premature termination of transcription. In hindsight, the avoidance of UAU codons is a key consideration when designing novel open reading frames or performing codon optimization.

I generated the C-terminally truncated variant $TG2_{465}$ and confirmed that it has transamidase activity. In line with previous work, I confirmed that self-crosslinking is diminished in $TG2_{465}$ (Stamnaes et al., 2015). It is worth noting that diminished self-crosslinking means that there are no glutamine resides in $TG2_{465}$ that are compatible with any lysine residues in $TG2_{465}$ for the transamidation reaction. It does not mean that $TG2_{465}$ is devoid of lysine residues or glutamine residues that are compatible with other residues for the transamidation reaction. Later work in this chapter in which Qtag and LysTag were fused to

$TG2_{465}$ demonstrated self-crosslinking, confirming that $TG2_{465}$ does still contain both reactive lysine and glutamine residues. The reason the reactive lysine and glutamine residues of $TG2_{465}$ are not crosslinked can be explained by the strong preference of TG2 for reactive residues in regions of flexibility. Steric hindrance is clearly an important factor in residue compatibility, with the requirement for both the Q and K sites occupying the narrow active site tunnel simultaneously.

Previous work has determined the reaction kinetics of protein conjugating enzymes such as SrtA and Butelase-1 (Chen et al., 2011) (Nguyen et al., 2014). If a truncated variant of TG2 were to be successfully engineered to achieve site-specific conjugation, it would be important to establish the reaction kinetics of the variant to determine if activity is comparable to other established methods. Calculating reaction kinetics was not possible with the cadaverine assay used in this thesis, as this assay does not directly measure product formation.

I demonstrated the propensity of $TG2_{465}$ to remain bound to fibronectin and type I collagen after thorough washing. Fibronectin binding was observed at low nanomolar concentrations of $TG2_{465}$, which is consistent with previous work that measured the affinity of TG2 and $TG2_{465}$ for fibronectin at 0.9 nM and 0.3 nM (Cardoso et al., 2017). While $TG2_{465}$ binds robustly to type I collagen and fibronectin, only low levels of retention were observed for a basement membrane extract. Therefore, it may be the case that $TG2_{465}$ is more suited to conjugation in complex scaffolds, rather than in collagen or fibronectin scaffolds where concentrations of high-affinity binding partners are high.

I also demonstrated high affinity binding of TG2 to the cell surface *in vitro*. Under normal physiological conditions TG2 is found predominantly in the cytosol, with smaller fractions localised at the cell surface and secreted into the extra-cellular space (Eckert et al., 2014). The majority of endogenous extracellular TG2 is known to localise at the cell surface (Gaudry et al., 1999) (Scarpellini et al., 2014). It has been hypothesised that TG2 secretion

occurs via an unusual secretory pathway involving recycling endosomes (Belkin, 2011), whereby TG2 is deposited at the cell surface already in complex with syndecan-4 (Wang et al., 2012). In support of this hypothesis, abolishment of TG2-syndecan-4 interactions results in decreased localisation of TG2 at the cell surface (Wang et al., 2012). It was not known, until now, if free TG2 in solution would bind to the cell surface independently of this unusual secretory pathway. Furthermore, the observation of decreased cell-surface binding for $TG2_{465}$ compared to TG2 suggests that binding is at least partially due to syndecan-4 interactions. TG2 contains a heparan sulfate proteoglycan binding site in the first β-barrel domain, which is absent from the truncated form (Wang et al., 2012). Caution must be taken when extrapolating these results to predict the *in vivo* behaviour of TG2. The cell-binding assays were performed on a single carcinoma cell line. It is not clear how much fibronectin, integrins, syndecan-4 and TG2 these cells express. It is important to expand upon these findings by testing a range of cell types, including more physiologically relevant cells for therapeutic applications such as lymphocytes and primary fibroblasts. In future experiments, further insights into TG2-cell surface interactions could be gained using crosslinking mass spectrometry to identify the proteins that TG2 is interacting with.

I established that $TG2_{465}$ is rapidly inactivated in the absence of reducing conditions, rendering it unsuitable for applications in the extracellular space or *in vitro* applications involving important disulfide bonds. However, $TG2_{465}$ C230S was able to retain some activity for at least 30 minutes, opening the possibility of the aforementioned applications. Given the putative role of transamidation in fibrosis, it may be desirable for *in vivo* applications to have an enzyme which is only briefly active and then switched off before excessive off-target transamidation can occur. Furthermore, transamidase activity can be fully restored after oxidation. Rapid activation/inactivation and full recovery of activity are desirable traits for *in vitro* solid phase protein assembly with repeated cycles of conjugation. The C230S mutant is

an improvement upon $TG2_{465}$ for this application because activity recovers more rapidly, allowing quicker protocols. However, further assays are required to established reaction kinetics. Evaluation of the k_{cat} of the redox mutants is required to confirm if differences in activity are due to altered redox kinetics or altered reaction kinetics.

The observation in this chapter that $TG2_{465}$ activity can be fully recovered after oxidative inactivation contrasts with a previous study which reported that reduction with DTT after oxidation with glutathione results in only a 60% recovery of transamidase activity (Stamnaes et al., 2010). However, the authors reported detection of glutathionylation due to the use of glutathione as an oxidising agent. Glutathionylation of the catalytic cysteine or cysteine residues close to the active site could result in the observed loss of activity after oxidation. The diminished activity of $TG2_{465}$ C230S in non-reducing conditions demonstrated in this chapter is also in contrast with results reported in the above work which showed full-length TG2 C230A was resistant to inactivation even in presence of oxidised glutathione (Stamnaes et al., 2010). In this case, key cysteine residues such as C370 and C371 may have become glutathionylated, preventing formation of the regulatory disulfide bond.

It is of benefit for conjugation in oxidising conditions to engineer $TG2_{465}$ to completely abolish formation of regulatory disulfide bonds and thus abolish oxidative inactivation. I demonstrated that removal of both the C230-C380 and C230-C371 bonds by mutation of C370 does not prevent oxidative inactivation. This finding confirms that, in the absence of the two known regulatory disulfide bonds, other regulatory disulfide bonds can form. Until these alternative disulfide bonds are identified, the easiest approach may be to block all solvent-accessible cysteine residues by glutathionylation or alkylation. Alkylation has previously been used to abolish disulfide bonds in antibodies and in the protease subtilisin from *Bacillus lentus* (Chalker et al., 2011). The challenge with the cysteine blocking approach is to achieve modification of all disulfide-bond forming cysteines while avoiding modification of the

catalytic cysteine. Selective inhibition of alkylation of the catalytic cysteine may be possible by addition of a small peptide active site inhibitor during the blocking reaction.

TG2-mediated conjugation of a model Q-tagged protein to protein scaffolds was unsuccessful, and an alternative anchoring approach whereby the Qtag is fused to TG2 itself was also unsuccessful. One possible explanation is that there is a strong preference for formation of TG2:substrate and TG2:TG2 conjugates. This hypothesis is supported by the observation of TG2:mClover-Q conjugates in the cell assay blot. It may be possible to eliminate the self-reactivity of TG2 by acetylation of reactive lysine residues. It is also possible that covalent conjugation to scaffolds is inhibited by the strong non-covalent interactions observed for the cell surface and ECM proteins. Non-covalent interactions with the scaffold may orient TG2 in a way which prevents active site access to lysine donors. Given that all the known sites of importance for non-covalent interactions are in the N- and C-terminal domains, it may be possible to eliminate such interactions by truncating TG2 to leave only the catalytic domain.

The self-conjugation of TG2-K and TG2-Q into large molecular weight multimers demonstrates that $TG2_{465}$ still contains both reactive lysine and glutamine residues. This highlights the difficulty in predicting reactivity of a given lysine or glutamine in a protein which further complicates application. Almost all proteins of interest will contain at least one glutamine or lysine and it may prove difficult to predict whether they are substrates for TG2 and if so, to which other substrates residues they can be conjugated.

In summary, TG2 is a challenging protein for directing site-specific protein-protein conjugation. Not least due to the relatively low (and difficult to predict) substrate specificity, and the presence of reactive lysine and glutamine residues in TG2 itself. Much more work is required to make TG2 more amenable as a protein-protein conjugation tool, including quantification of the activity of TG2 variants to compare with full length TG2 and other protein conjugation systems.

High affinity of TG2 for protein scaffolds and the cell surface present an interesting opportunity. A high affinity/low specificity binding protein could be used to broadly decorate tissue scaffolds for slow release of therapeutics, or to broadly decorate the cell surface with ligands to modulate cell behaviour. Methods of highly specific targeting of the cell surface such as antibodies and antibody fragments are ideal for targeting a specific receptor in a specific pathway and avoiding off target effects, but it may be of benefit in some cases to broadly decorate the extracellular space with a therapeutic in a low-specificity approach to maximise the local concentration of the cargo. Such a tool could be used to decorate target tissues with growth factors for wound healing or decorate the vasculature of donated organs with immunosuppressing cytokines before implantation.

Chapter 4 – Validating Second Generation RrgA-based Technologies

4.1. Introduction

There is potential to utilise TG2's high affinity interactions with fibronectin and the cell surface for decorating endogenous protein scaffolds, as well as cells, with functional proteins. TG2 could be fused to protein cargo and then applied to a target scaffold (Figure 4.1A). The simplest method of generating TG2:cargo is by genetic fusion. However, recombinant overexpression of TG2 in mammalian systems has not been demonstrated in the literature and my own attempts were unsuccessful. As a result, genetic fusion is limited only to cargo which can be expressed in functional form in *E. coli*. Some proteins which can be expressed in *E. coli* are expressed in insoluble form and the gene products require refolding; such proteins may not refold under the same conditions when fused to TG2. Thus, genetic fusion of cargo proteins to TG2 is limited to a small portion of all the possible proteins that may be of interest.

The rise of Synthetic Biology has seen the introduction of engineering principles to biotechnology to facilitate a transition from 'artisan' to 'industrial' scale. This includes the modularisation of biological components, such as genetic elements in the Registry of Standard Biological Parts (Peccoud et al., 2008) and computational elements in the CellML Model Repository (Cooling et al., 2010). A system for post-translational fusion of cargo to TG2 would vastly increases the range of cargo proteins that can be fused to TG2. Any protein could be expressed as desired and then ligated onto TG2 and delivered to a scaffold (Figure 4.1B). Such modularity would significantly increase the utility of TG2 as a tool.

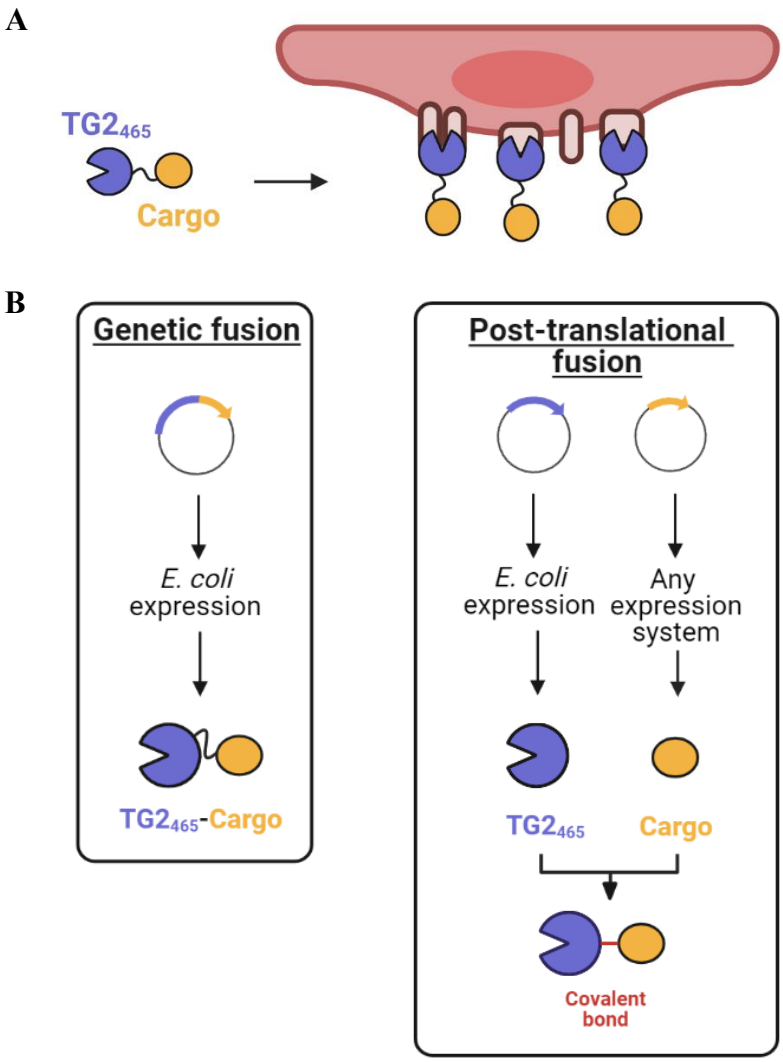


Figure 4.1. TG2 as a tool for decoration of endogenous proteins. (A) A schematic of TG2-mediated decoration of the cell surface with a cargo protein of interest (B) A schematic of methods for generating TG2:cargo constructs. Modular post-translational coupling enhances the range of proteins that can be fused to TG2, thereby increasing the utility of TG2.

RrgA is an adhesin protein from *Streptococcus pneumoniae* which forms a spontaneous intramolecular isopeptide bond between a lysine residue and an asparagine residue in the D4 domain, which is promoted by an opposing glutamate (Izoré et al., 2010) (Figure 4.2A). The D4 domain of RrgA was previously split into two parts such that the reactive lysine was included in a 12 amino acid peptide named SnoopTag, and the reactive asparagine and opposing glutamate were included in a 112 amino acid protein named SnoopCatcher (Figure 4.2B) (Veggiani et al., 2016). SnoopTag and SnoopCatcher spontaneously form an isopeptide bond

when mixed, which was utilised to direct conjugation of two model proteins. The SnoopTag/Catcher system has been used in combination with the orthogonal SpyTag/Catcher system for assembly of polyproteins by solid-phase attachment (Veggiani et al., 2016).

The RrgA D4 domain has also previously been split into three components to produce a ligase-based system named SnoopLigase (Buldun et al., 2018). In this case, the reactive lysine is included in the 12 amino acid SnoopTag peptide, the reactive asparagine is included in a 23 amino acid DogTag peptide, and the catalytic glutamate is included in the 104 amino acid SnoopLigase (Figure 4.2C). The ligase directs the formation of an isopeptide bond between the two tags, which does not occur in the absence of the ligase. SnoopLigase has previously been used for modular assembly of a malaria vaccine candidate (Andersson et al., 2019), bispecific nanobodies (Chen et al., 2022), and enzyme cyclization (Figure 4.2D).

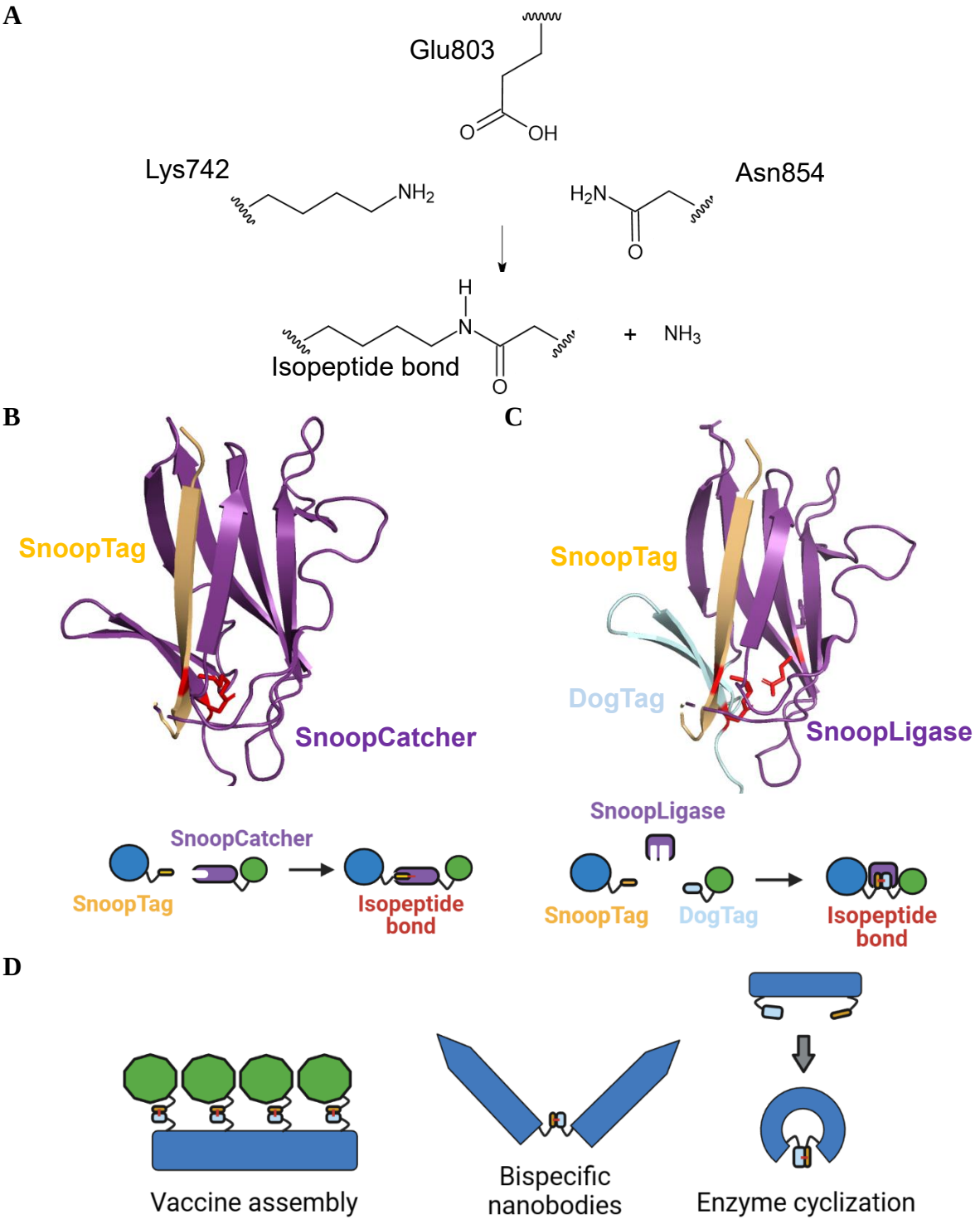


Figure 4.2. RrgA-based systems for protein-protein ligation. (A) Reaction mechanism for isopeptide bond formation in RrgA. Glu803 promotes isopeptide bond formation between Lys742 and Asn854 resulting in loss of ammonia. (B) Two-way split of the RrgA D4 domain into SnoopTag and SnoopCatcher (schematic based on PDB:2WW8). Isopeptide-forming residues are coloured red. SnoopTag and SnoopCatcher are genetically fused to proteins of interest and an isopeptide bond is formed between the two upon mixing. (B) Three-way split of the RrgA D4 domain into SnoopTag, DogTag, and SnoopLigase (schematic based on PDB:2WW8). The catalytic glutamate and isopeptide-forming residues are coloured red. SnoopTag and DogTag are genetically fused to proteins of interest; formation of an isopeptide bond is facilitated by addition of SnoopLigase. (D) Applications of SnoopLigase.

The SnoopLigase system is not without performance limitations. The ligase demonstrates poor solubility in commonly used buffers which limits application. Also, the reaction rate is slow and requires overnight incubation at high concentrations of substrate to drive the reaction near to completion. Improving the efficiency of RrgA-based systems would allow application where substrate concentration is limited, such as when a substrate has low solubility. Also, highly efficient conjugation is more able to overcome limitations such as steric hindrance.

There has been significant effort in our lab to identify mutations in the RrgA D domain which improve performance of RrgA-based systems (Veggiani et al., 2016) (Buldun et al., 2018) (Keeble et al., 2022). Dr. Anthony Keeble recently used a directed evolution approach to identify 3 mutations which may improve the reactivity of SnoopTag, and 1 mutation which may improve the reactivity of DogTag (Figure 4.3A). In addition, previously published work identified mutations in another RrgA-based system which improved solubility and reactivity (Keeble et al., 2022). Four of the mutations identified in this work were introduced to SnoopLigase by Dr. Keeble (Figure 4.3A). The mutations in SnoopTag, DogTag and SnoopLigase have been combined to create the second generation SnoopLigase system named SnoopLigase2. Mutations which were identified in previous work have also been introduced to SnoopCatcher to produce the second generation SnoopCatcher2, which in combination with SnoopTag2 is expected to improve reactivity (Figure 4.3B).

In this chapter I established the reaction rates for first and second generation SnoopLigase and SnoopTag/Catcher systems and explored the contribution of each component of the SnoopLigase system to the observed improvement in performance. I examined the molecular basis for improved performance by comparing ligase-product affinity for first and second generation SnoopLigase systems. Finally, I tested the coupling efficiency of second-

generation systems in a wide range of reaction conditions to determine the scope of potential applications.

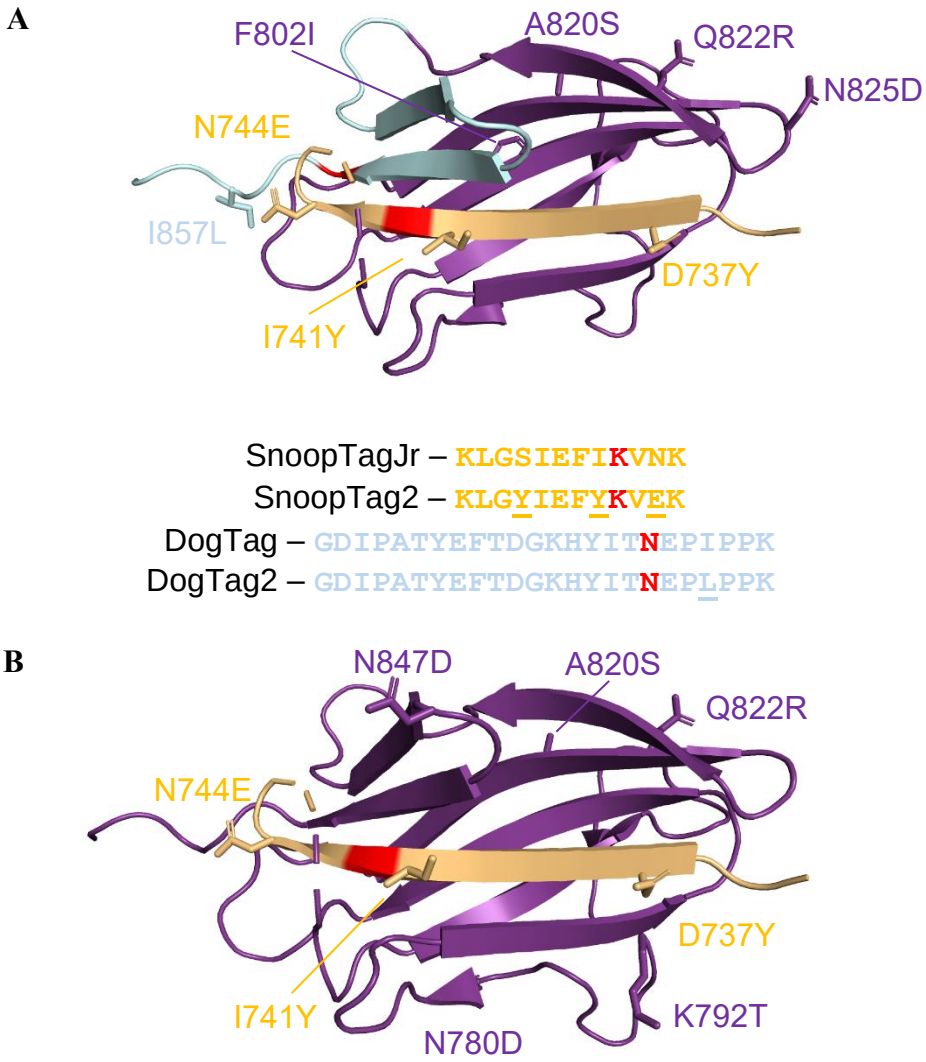


Figure 4.3. Second generation RrgA-based systems. (A) Schematic of SnoopLigase2 based on the structure of the parental domain from PDB 2WW8. The isopeptide-forming residues are highlighted in red. D737Y, I741Y and N744E were identified by directed evolution of SnoopTagJr and combined to create SnoopTag2. I857L was identified by directed evolution of DogTag and introduced to create DogTag2. Amino acid sequences of tags are given with mutations underlined and isopeptide-forming residues coloured red. F802I, A820S, Q822R were identified by directed evolution of DogCatcher and shown to improve reactivity. N825D was added to DogCatcher by rational design to improve solubility. The four DogCatcher mutations were introduced to SnoopLigase to create SnoopLigase2. (B) Schematic of SnoopTag/Catcher based on the structure of the parental domain from PDB 2WW8. The isopeptide-forming residues are highlighted in red. K792T and N780D were previously shown to improve solubility of DogCatcher. N847D was previously added to DogTag to improve electrostatic interactions with a neighbouring Lys. K792T, N780D and N847D were combined with the previously described A820S and Q822R mutations to create SnoopCatcher2.

4.2. Validating the SnoopLigase2 System

4.2.1. Comparison of First- and Second-Generation Tags

Firstly, I validated formation of the expected SnoopTag2/DogTag2 ligation product by SnoopLigase2 using SDS-PAGE and electrospray ionization mass spectrometry (Figure 4.4).

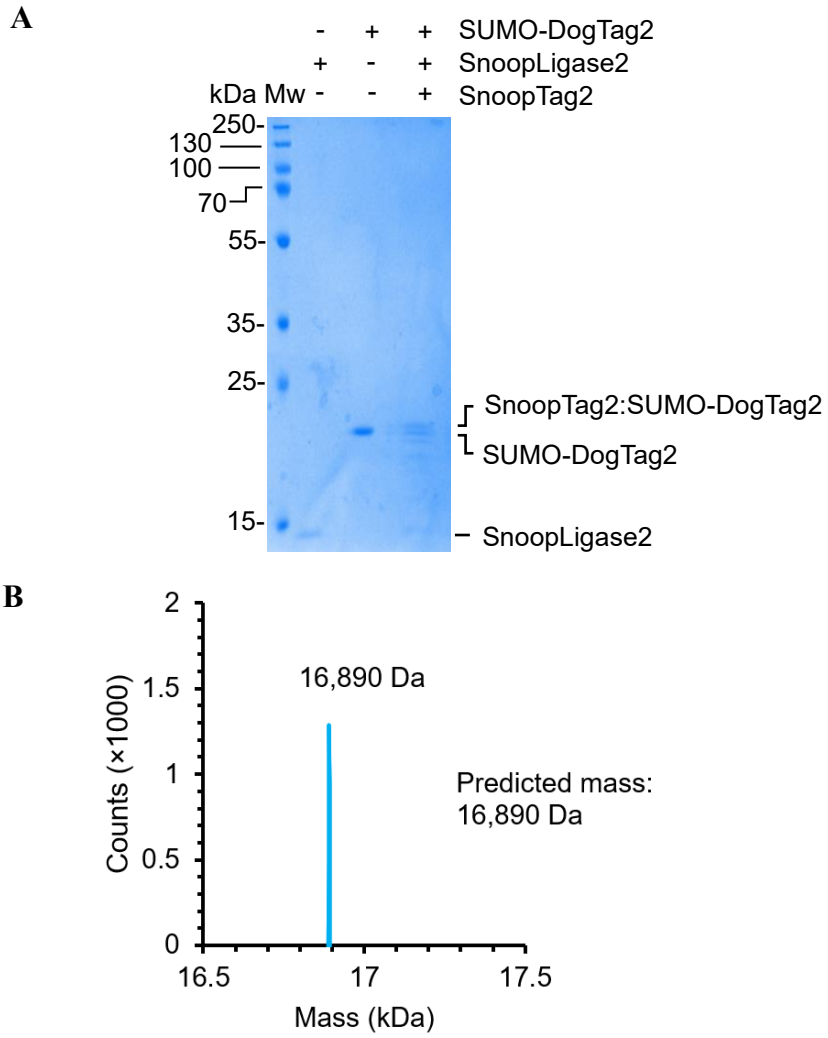


Figure 4.4. Mass spectrometry of SnoopLigase2-mediated ligation. (A) Conjugation of SnoopTag2 to SUMO-DogTag2 by SnoopLigase2. 200 μM SnoopTag2, 40 μM SUMO-DogTag2 and 60 μM SnoopLigase2 were incubated for 24 h at 4 °C. Analysis by 12% SDS-PAGE with Coomassie staining. Mw = molecular weight markers. (B) Electrospray ionization mass spectrometry detects a 16,890 Da product corresponding to SUMO-DogTag2:SnoopTag2 with loss of ammonia.

To test and compare the reactivity of SnoopLigase system components I used model substrates in a reaction where SnoopLigase facilitated ligation between SnoopTagged-sfGFP and DogTagged-small ubiquitin-like modifier (SUMO) (Figure 4.5A). SnoopTag2

demonstrated a major improvement in coupling efficiency compared to SnoopTagJr (Figure 4.5B). Similarly, DogTag2 demonstrated a major improvement in coupling efficiency compared to DogTag (Figure 4.5C).

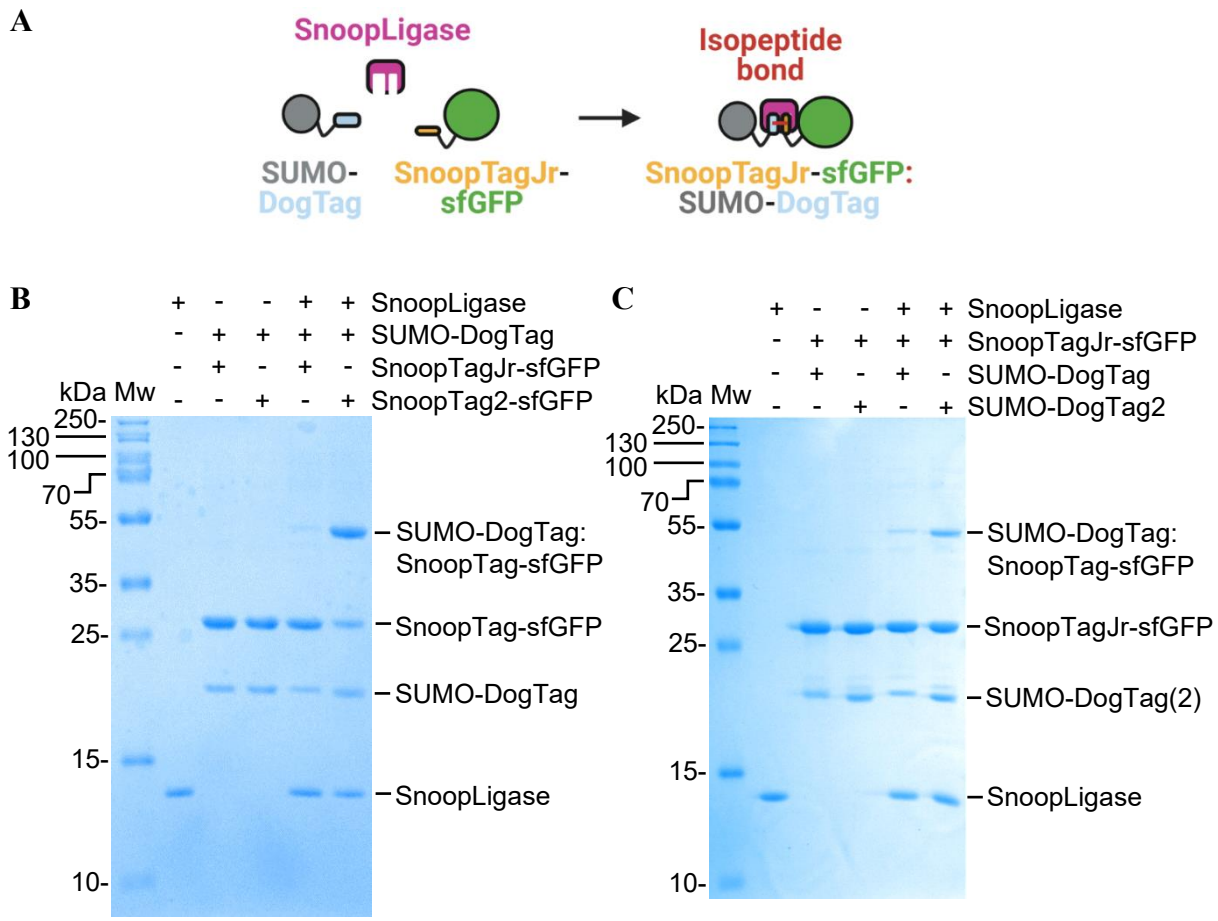


Figure 4.5. Comparison of reactivity for first and second generation SnoopLigase tags. (A) Schematic of the model reaction used to measure and compare reaction efficiency for SnoopLigase components. SnoopTagJr/2-sfGFP and SUMO-DogTag(2) were incubated with SnoopLigase(2). (B) and (C) - analysis of reactions by 12% SDS-PAGE with Coomassie staining. A colon indicates that the proteins are ligated together. All samples are reduced with 100 mM DTT before analysis. Mw = molecular weight markers. (B) Enhanced reactivity of SnoopTag2. SnoopTagJr-sfGFP or SnoopTag2-sfGFP was ligated to SUMO-DogTag using SnoopLigase (25 μM each) for 3 h at 21 °C in Tris-borate pH 7.4. (C) Enhanced activity of DogTag2. SUMO-DogTag or SUMO-DogTag2 was ligated to SnoopTagJr-sfGFP using SnoopLigase (25 μM each) for 3 h at 21 °C in Tris-borate pH 7.4.

4.2.2. Comparison of First and Second Generation SnoopLigase Sets

The original SnoopLigase has low solubility, resulting in high sensitivity to buffer conditions. SnoopLigase was previously found to work optimally in Tris-borate with 15% (v/v) glycerol and minimal NaCl, but was almost completely inactive in commonly used PBS

(Buldun et al., 2018). I first compared the first and second generation SnoopLigase systems in the optimal buffer. The original set demonstrated some reactivity (~40% coupling in 90 min), but the new set demonstrated greatly improved reactivity (~80% coupling in 10 min) (Figure 4.6A). I then compared the first and second generation SnoopLigase systems in PBS. The first generation system hardly showed any reactivity (<5% after 90 min), while the second generation system performed efficient ligation (~80% coupling in 60 min) (Figure 4.6B). Due to the limited reactivity of the original set in PBS, the initial rates of reaction were based on the time to 10% coupling. The second generation set demonstrated an 11-fold improvement in reaction rate in Tris-borate + 15% (v/v) glycerol and a 1,300-fold improvement in PBS (Table 4.1).

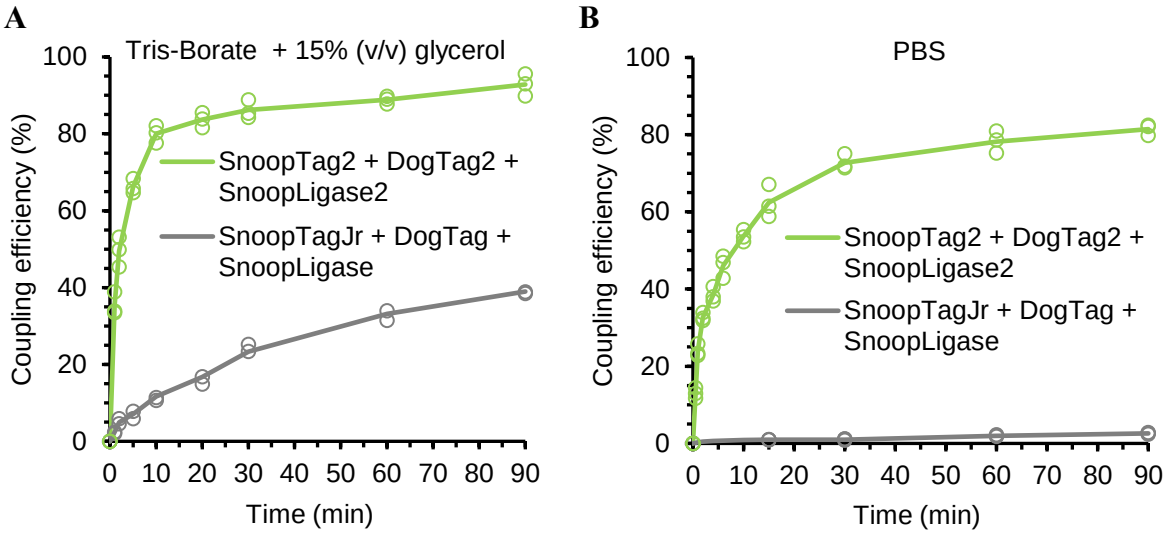


Figure 4.6. Comparison of reactivity for first and second generation SnoopLigase sets. (A) SnoopLigase2 enhances reaction in Tris-borate. 25 μM each of SnoopTagJr/2-sfGFP, SUMO-DogTag(2) and SnoopLigase(2) were incubated for varying times at 21 °C in Tris-borate pH 7.4 with 15% (v/v) glycerol, before SDS−PAGE/Coomassie. Circles show each of the triplicate data points. (B) SnoopLigase2 is functional in the commonly used PBS. As in (A) in PBS pH 7.4. Coupling efficiency is measured as ratio of ligated product to both unreacted substrates. Some data points are overlapping. Experiments in (A) and (B) are a representative example of two experiments performed in total.

Table 4.1 Reaction rates for first and second generation SnoopLigase systems

		Experiment 1 (Figure 4.6)	Experiment 2 (Appendix Figure 4)
Components	**Buffer**	**Time to 10% coupling (s)**	**Time to 10% coupling (s)**
SnoopLigase, SnoopTag, DogTag	PBS	27,120 ± 1,800	27,300 ± 2,004
	Tris-borate + 15% (v/v) glycerol	510 ± 50	480 ± 49
SnoopLigase2, SnoopTag2, DogTag2	PBS	22 ± 2	24 ± 1
	Tris-borate + 15% (v/v) glycerol	15 ± 4	14 ± 3

I then determined the contribution of the second-generation tags to the reactivity of the SnoopLigase2 system by reacting different combinations of first- and second-generation tags with SnoopLigase2. DogTag2 confers a significant increase in coupling efficiency compared to DogTag when reacted with SnoopTagJr ($p < 0.001$, $n = 3$) (Figure 4.7). Similarly, SnoopTag2 confers a significant increase in coupling efficiency compared to SnoopTagJr when reacted with DogTag ($p < 0.001$, $n = 3$) (Figure 4.7). Finally, SnoopTag2 and DogTag2 in combination confer a significant increase in coupling efficiency compared to either SnoopTag2 + DogTag or SnoopTagJr + DogTag2 (Figure 4.7) ($p < 0.05$, $n = 3$).

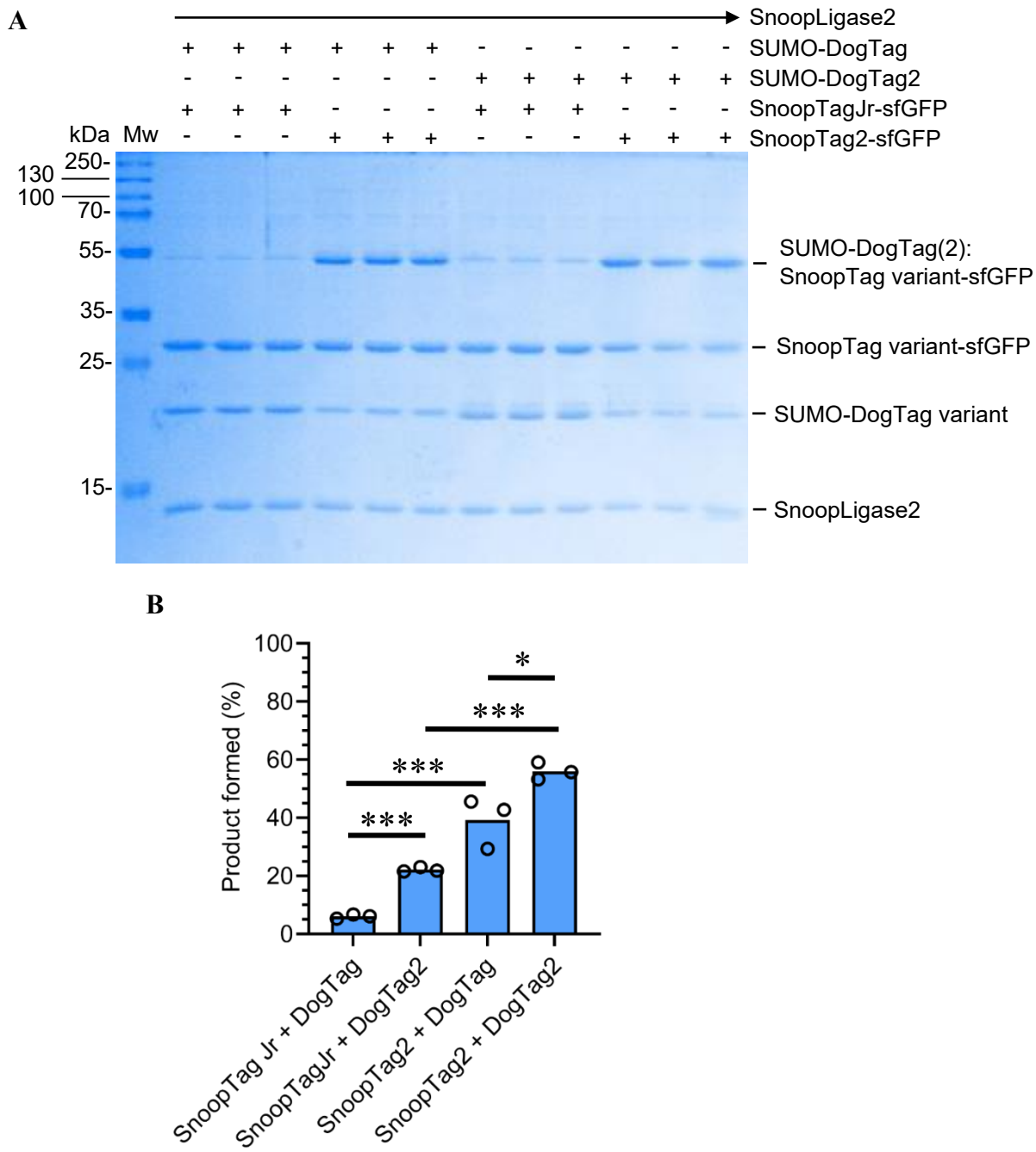


Figure 4.7. Reactivity of original versus modified tags for SnoopLigase2 reaction. (A) Analysis of reactions by 12% SDS-PAGE with Coomassie staining. Reactions were performed in triplicate with each component at 25 μM in Tris-Borate pH 7.4 for 3 h at 21 °C. All samples were reduced with 100 mM DTT before analysis. Mw = molecular weight markers. (**B**) Quantification of reaction in (A). Coupling efficiency is measured as ratio of product to both unreacted substrates. Circles show each of the triplicate data-points for tagged proteins reacting with SnoopLigase2, with the bar indicating the mean. Unpaired two-tailed Student's t test was used to compare selected combinations of tags (* denotes $p < 0.05$, *** denotes $p < 0.001$, $n = 3$).

4.2.3. Validation of SnoopLigase2 in a Range of Reaction Conditions

After establishing that SnoopLigase2 reactivity is not restricted to one specific buffer, I tested SnoopLigase2-mediated ligation of SnoopTag2 and DogTag2 in a wide range of reaction conditions. Reactions were performed at a short time-point where reactivity was ~50% to detect increased or decreased efficiency compared to the control reaction. SnoopLigase2 ligated efficiently at all tested temperatures (4−45 °C) with very little impact on performance (Figure 4.8A). This is in contrast with SnoopLigase which is considerably impaired at temperatures of 37 °C and higher (Buldun et al., 2018).

SnoopLigase2 reaction was also active over a broad pH range, with good reactivity at pH 4.0-10.0 and even some reactivity at pH 4.0 (Figure 4.8B). This is an improvement over SnoopLigase which demonstrates some loss of activity at pH 8.75 and above (Buldun et al., 2018). Commonly-used buffer additives such as the detergent TWEEN 20 had minimal effect on reactivity of SnoopLigase2 (Figure 4.7C). NaCl was well tolerated up to 1 M (Figure 4.8D), which contrasts with the original SnoopLigase system which was considerably impaired at 137 mM NaCl (Buldun et al., 2018). Finally, SnoopLigase2 performed efficient ligation in all four common buffers that were tested (Figure 4.8E), demonstrating much greater buffer tolerability than SnoopLigase.

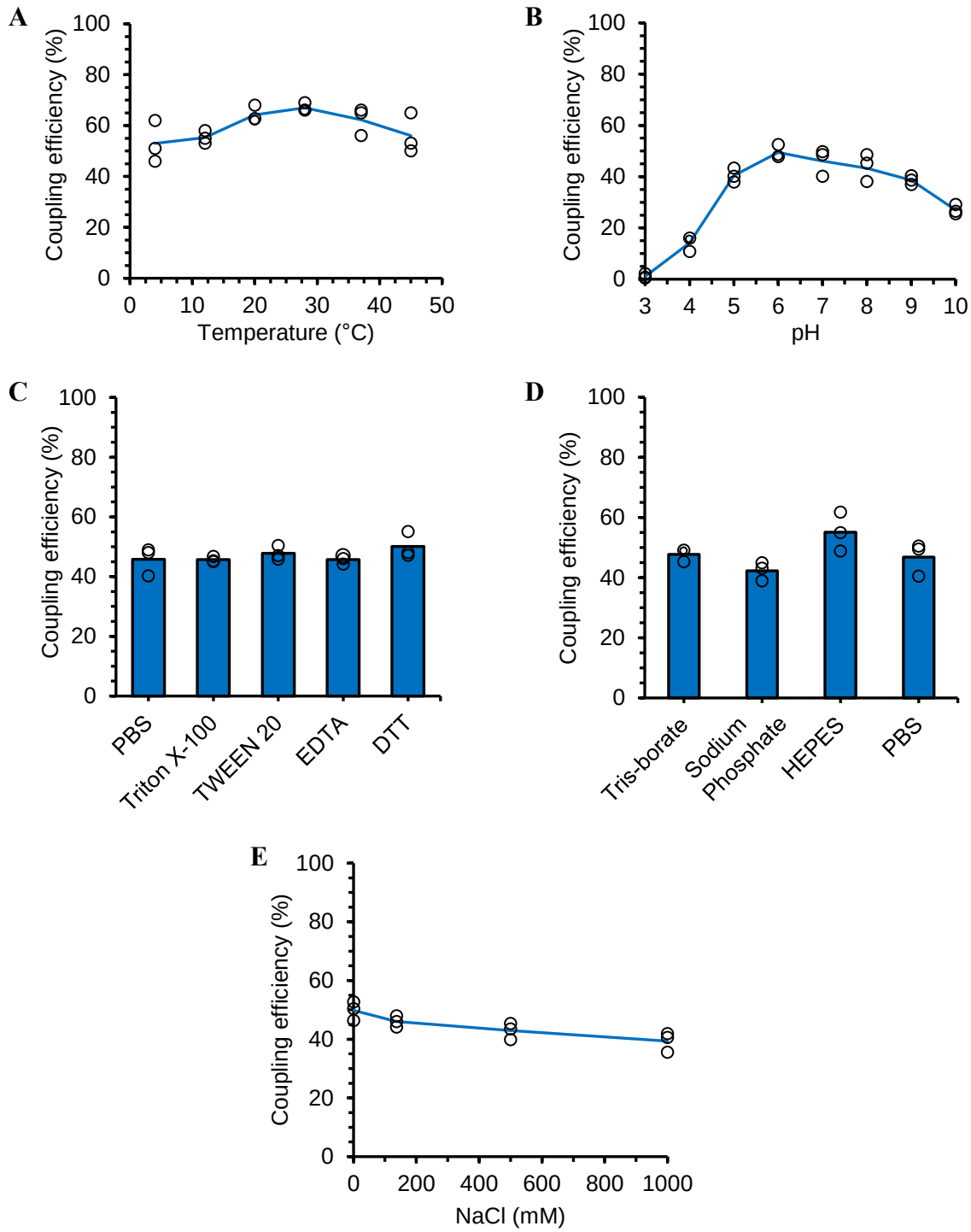


Figure 4.8. SnoopLigase2 reacts in a wide range of conditions. (A) Temperature-dependence of SnoopLigase2 reaction. SnoopLigase2 ligated SnoopTag2-sfGFP and SUMO-DogTag2 (20 μM each) in HEPES pH 7.5 for 10 min at the indicated temperature. (B) pH-dependence. As in (A) in succinate−phosphate−glycine at the indicated pH for 9 min at 25 °C. (C) Additive compatibility. As in (A) in HEPES-buffered saline pH 7.0 for 8 min at 25 °C with 1% (v/v) Triton X-100, 1% (v/v) TWEEN 20, 1 mM EDTA, or 2 mM DTT. (D) Buffer compatibility. As in (A) for 12 min at 25 °C with the indicated buffer. © Salt-dependence. As in (A) in HEPES pH 7.5 for 12 min at 25 °C with the indicated additional NaCl. Coupling efficiency is measured as ratio of ligated product to both unreacted substrates. Circles show each of the triplicate data points, with the line or bar indicating the mean.

4.2.4. Examining the Mechanism for Enhanced Reactivity of SnoopLigase2

SnoopLigase was previously shown to remain non-covalently bound to the SnoopTag-DogTag product after the reaction, revealing that SnoopLigase acts as a single-turnover catalyst (Buldun et al., 2018). A protocol was developed for removal of SnoopLigase from the product by biotinylating SnoopLigase and capturing the SnoopLigase:product complex with Streptavidin-agarose beads (Buldun, Khairil Anuar and Howarth, 2021). The product was dissociated from SnoopLigase by washing with 3 M imidazole, a short pH 2.5 acid wash, or competitive elution with SnoopTag:DogTag peptide. Here, I generated a HaloTag-SnoopLigase2 construct for covalent capture of the SnoopLigase2:product complex and subjected the complex to a range of washes to dissociate the ligated product from SnoopLigase2 (Figure 4.9A). HaloTag-SnoopLigase2 facilitated efficient ligation between SnoopTag2-sfGFP and TG2x-DogTag2 (Figure 4.9B). 3 M imidazole, pH 2.5, or competitive elution was not able to dissociate the product from SnoopLigase2 (Figure 4.9C). SnoopTag2-sfGFP:TG2x-DogTag2 was only dissociated from HaloTag-SnoopLigase2 upon boiling in sodium dodecyl sulfate (SDS).

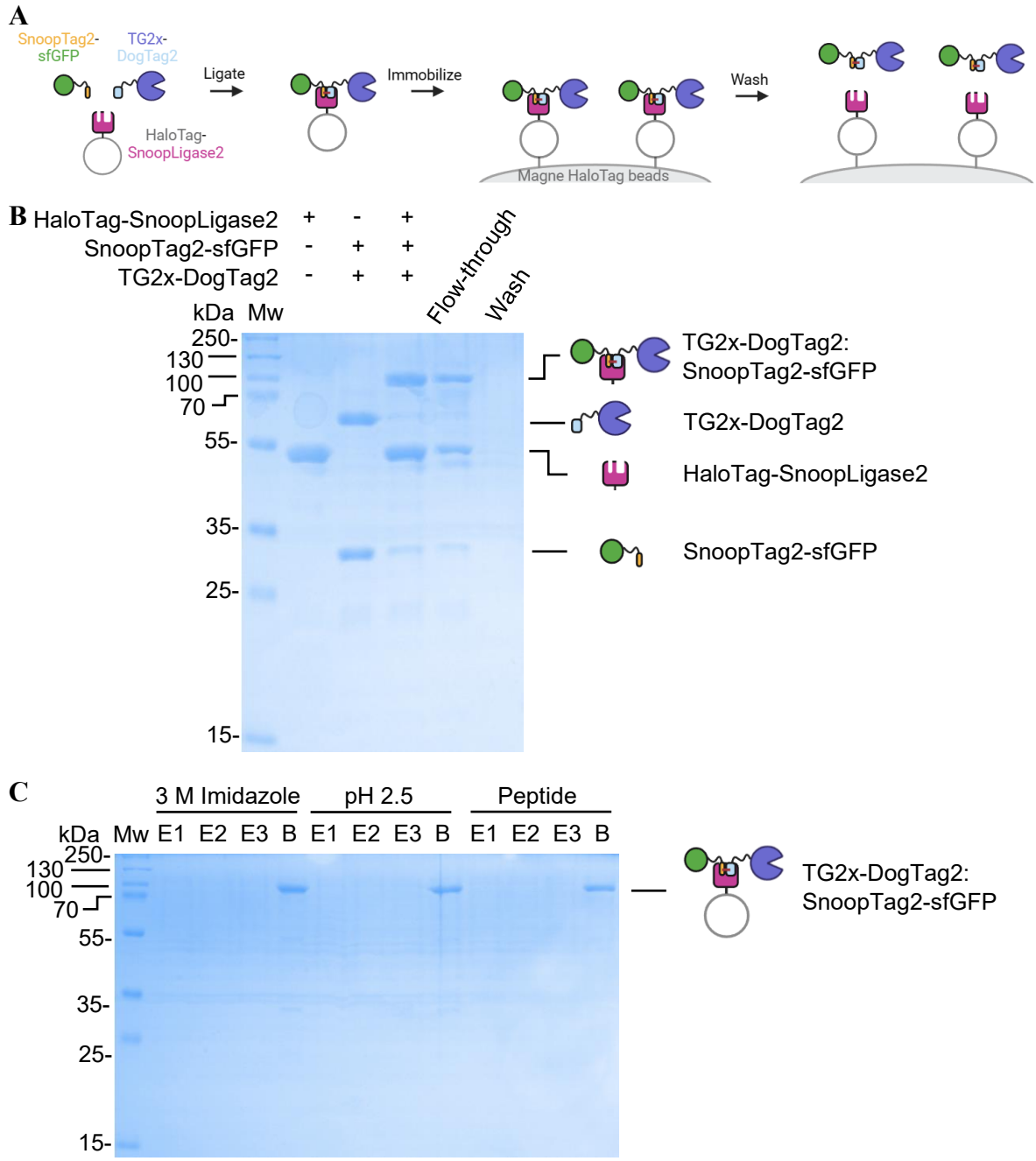


Figure 4.9. SnoopLigase2 is stably anchored to its reaction product. (A) Schematic of solid-phase SnoopLigase2 purification. SnoopTag2- and DogTag2-linked proteins are conjugated by HaloTag-SnoopLigase2. Magne HaloTag beads are used to capture the HaloTag covalently and a range of conditions are used to try to dissociate the ligated product from HaloTag-SnoopLigase2. (B) and (C) - analysis of SnoopLigase elution by 12% SDS-PAGE with Coomassie staining. All samples were reduced with 100 mM DTT before analysis. Mw = molecular weight markers. (B) SnoopTag2-sfGFP and TG2x-DogTag2 were conjugated by HaloTag-SnoopLigase2 (10 µM each) for 16 h at 24 °C. (C) Attempted dissociation of product from SnoopLigase2 using three different elution methods. HaloTag-SnoopLigase2 was captured with Magne HaloTag beads followed by incubation with 3 M imidazole, glycine pH 2.5, or 100 µM SnoopTagJr-DogTag peptide. E1-3 are 50 µL elution fractions. B are the beads resuspended in 50 µL SDS elution buffer, to show the TG2x-DogTag2:SnoopTag2-sfGFP that had been retained on the beads.

4.3. Validating the SnoopTag2/Catcher2 System

4.3.1. Comparison of First and Second Generation SnoopTag/Catcher Sets

To compare the reactivity of first and second generation SnoopTag/Catcher systems, I used a model reaction where a spontaneous isopeptide bond is formed between SnoopTagged-sfGFP and SnoopCatcher(2) (Figure 4.11A). The first generation SnoopTag/Catcher system achieved ~10% coupling after 30 min in HBS pH 7.0 + 2% BSA. The second generation SnoopTag/Catcher system demonstrated a major improvement in reactivity, reaching 95% coupling after 20 min. I then calculated the second-order rate constant (k) for both reactions. The original set had a rate constant of $k = 5.0\times10^{-1}\ M^{-1}\ s^{-1}$, while the second generation set showed a ~82-fold improvement in reaction rate ($k = 4.1\times10^{2}\ M^{-1}\ s^{-1}$) (Figure 4.11B).

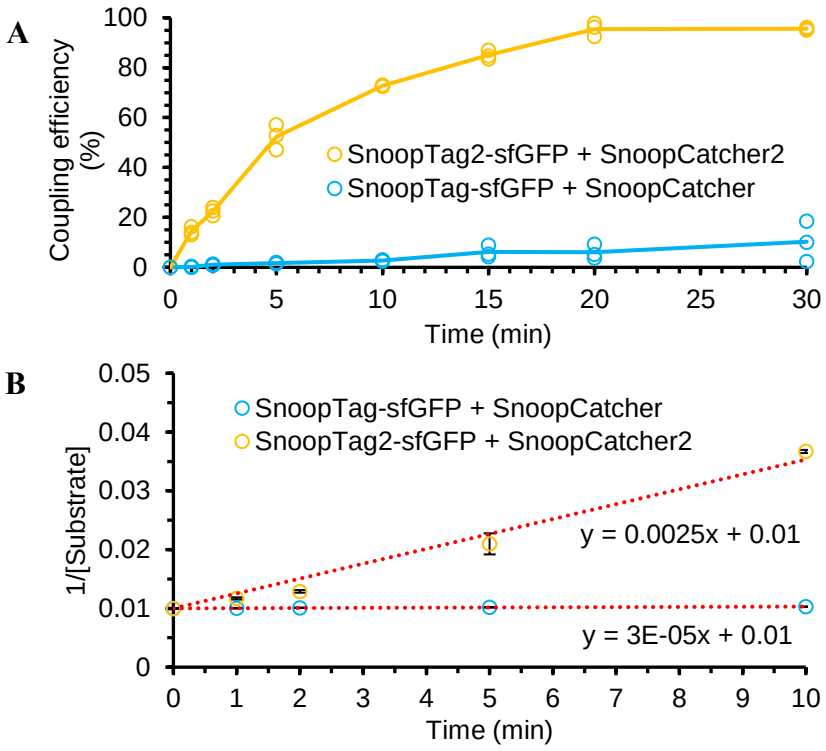


Figure 4.10. A comparison between first and second generation SnoopTag/Catcher sets. (A) Second generation SnoopTag/Catcher system displays considerable higher reactivity than its predecessor. 100 nM SnoopTag(2)-sfGFP was reacted with 100 nM SnoopCatcher(2) in HBS pH 7.0 + 2% (w/v) BSA at 16 ˚C, followed by analysis with SDS−PAGE and fluorescent imaging. Coupling efficiency is measured as ratio of ligated product to both unreacted substrates. Circles show each of the triplicate data points. (B) Rate of reactions in (A) with circles denoting mean of three replicates and error bars denoting ± 1 S.D. Some error bars are too small to be visible.

4.3.2. Validation of SnoopTag2/Catcher2 in a Range of Reaction Conditions

Given the previously demonstrated tolerance of the SnoopLigase2 system to a wide range of reaction conditions (Chapter 4.2.3), I tested the performance of the SnoopTag2/Catcher2 system in various conditions. These reactions were performed at a short time-point where reactivity was ~50%, to detect increased or decreased efficiency compared to the control reaction. SnoopTag2/Catcher2 ligated efficiently at 4−37 °C, but reactivity was impaired at 45 °C (Figure 4.11A). The SnoopTag2/Catcher2 reaction was also active over a broad range of pH values, with optimal reactivity at pH 7.0 and good reactivity at pH 10.0 (Figure 4.11B). As was the case for SnoopLigase2, all tested buffer additives showed no effect on reactivity of Snooptag2/Catcher2 (Figure 4.11C). Four common buffers were tested and efficient reactivity was observed in all cases (Figure 4.11D). Finally, NaCl concentrations up to 1 M were well tolerated (Figure 4.11E).

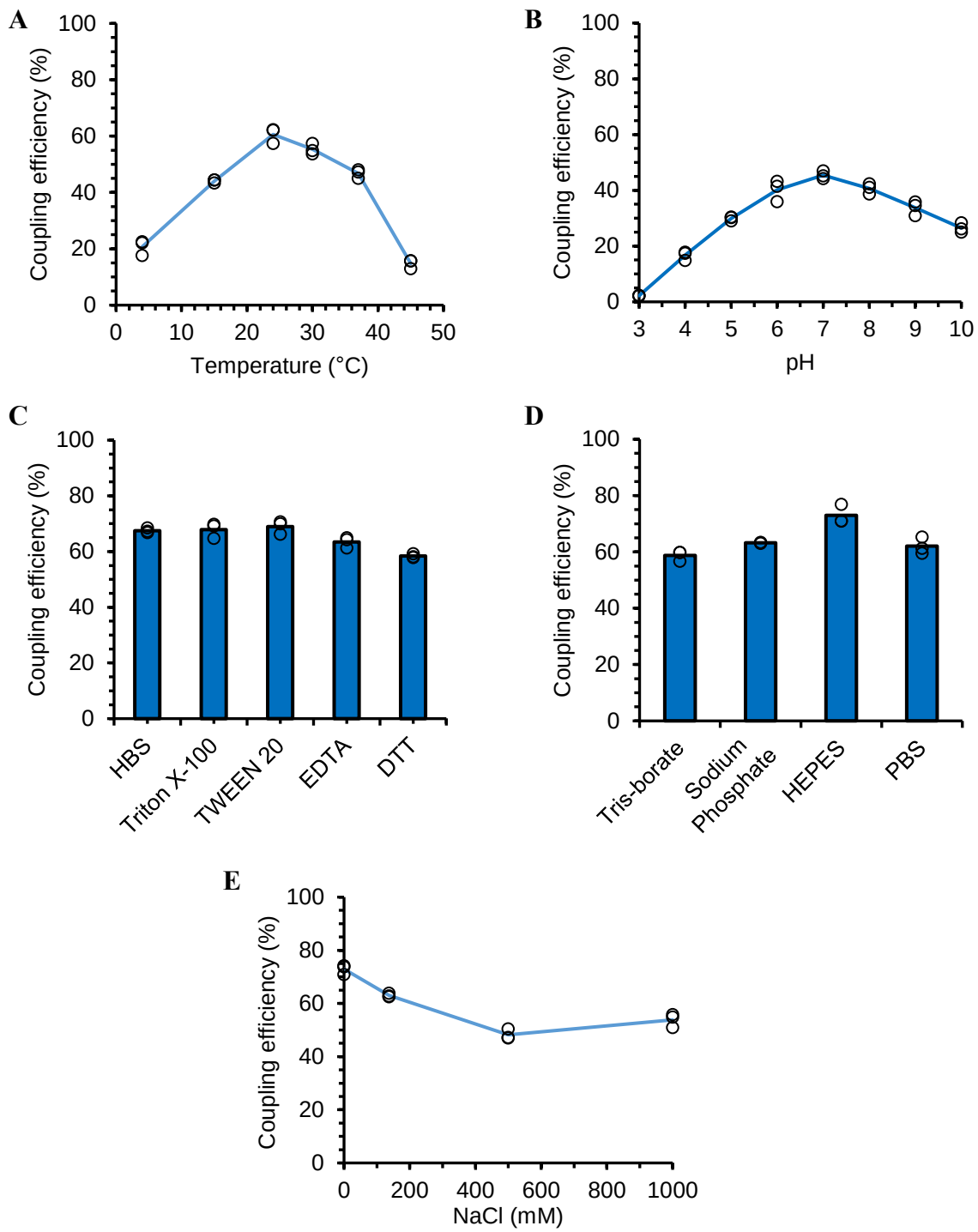


Figure 4.11. SnoopTag2/Catcher2 reacts in a wide range of conditions. (A) Temperature-dependence of the SnoopTag2/Catcher2 reaction. SnoopCatcher2 and SnoopTag2-sfGFP (100 nM each) ligated in PBS pH 7.4 + 2% (w/v) BSA for 10 min at the indicated temperature. (B) pH-dependence. As in (A) with succinate–phosphate–glycine at the indicated pH for 12 min at 16 °C. (C) Additive compatibility. As in (A) with HBS pH 7.0 + 2% (w/v) BSA for 12 min at 16 °C with 1% (v/v) Triton X-100, 1% (v/v) TWEEN 20, 1 mM EDTA, or 2 mM DTT. (D) Buffer compatibility. As in (A) at pH 7.5 for 12 min at 16 °C with the indicated buffer. (E) Salt-dependence. As in (A) with the indicated additional NaCl for 12 min at 16 °C. Coupling efficiency is measured as ratio of ligated product to both unreacted substrates. Circles show each of the triplicate data points, with the line or bar indicating the mean.

4.4. Discussion

For protein-protein coupling tools to be broadly useful to the research community, they must perform rapid conjugation in a wide range of reaction conditions. Mutations introduced into SnoopLigase and SnoopTag/Catcher systems in previous work were shown here to considerably increase reactivity. The 1,300-fold improvement in reaction rate of SnoopLigase2 in PBS and similar performance in other common buffers renders SnoopLigase2 compatible with a much wider range of potential applications. For a single turnover catalyst such as SnoopLigase, the observed increase in reactivity could be due to either an improvement in binding kinetics, or an increase in the speed of the isopeptide bond-forming reaction. The mechanisms by which the mutations improve binding is unknown and difficult to confirm. One clear effect of the mutations introduced to SnoopLigase2 is a drastic improvement in solubility in common buffers.

The stable association of SnoopLigase2 with the SnoopTag2:DogTag2 product must be considered when applying the SnoopLigase2 system and may limit its application in some cases. For example, in therapeutic applications there is a chance of induction of anti-SnoopLigase antibodies. However, when used for VLP assembly the presence of SnoopLigase did not significantly affect the immune response to the target antigen (Andersson et al., 2019). The unusually high affinity of SnoopLigase2 for the ligated product offers the opportunity to create stable heterotrimers of three separately expressed constructs in a single step, which may present alternative applications. For example, in assembly of tri-specific binders for immunotherapeutics or assembly of complex functionalised scaffolds for tissue engineering.

The improved reactivity of the SnoopLigase2 system produces efficient coupling in minutes compared to hours for SnoopLigase. For many *in vitro* protein assembly applications, coupling for several hours overnight is convenient and any increase in reaction rate is only marginally beneficial. However, for time-sensitive experiments such as coupling at the cell

surface this accelerated reactivity may be essential to achieve rapid coupling in seconds to minutes. Also, for nanoparticle assembly reaction rates are typically lower and achieving complete reactivity is challenging. The improved reactivity of SnoopLigase2 may allow application for time-sensitive cell-based coupling and nanoparticle assembly.

The rate of reaction for SnoopLigase2 has been well studied in this chapter, but there are cases where coupling efficiency is more important than the speed of the reaction. For example, 100% coupling efficiency can be important for separating reaction product from substrates to simplify downstream processing. The utility of some other protein-protein coupling systems, such as Sortase, are restricted by poor coupling efficiency due to the reversibility of the reaction (Morgan et al., 2022). In this chapter, coupling efficiencies of up to 95% were demonstrated, but further work is required to optimise reactions to achieve 100% coupling efficiency. In addition, improvements in reactivity were demonstrated in a model reaction with tags only added to one specific terminus. To prove the wide applicability of SnoopLigase2, rapid and efficient coupling should be demonstrated with a range of substrates, and with the tags situated at either terminus.

The improvement in tolerance to reaction conditions for SnoopLigase2 greatly increases the range of potential applications. SnoopLigase2 was shown to have strong temperature and pH tolerance, which is broader than anything demonstrated in the literature to date. Subtiligase, Sortases and inteins have been validated only in mild conditions (pH 7-9, 4-37 ºC) (Antos, Truttmann and Ploegh, 2016) (Weeks and Wells, 2020) (Bhagawati et al., 2020) (Carvajal-Vallejos et al., 2012). The only other systems that have been tested outside the range of pH 7-9 are PALs, Butelase-1 demonstrates optimal activity at pH 5.0-6.5 but is denatured at pH >7.5 (Nguyen et al., 2016). The tolerance of SnoopLigase2 to a wide range of pH values offers the possibility of highly specific covalent coupling within the lysosome, or for immobilising proteins to scaffolds under conditions that inhibit non-specific binding. High

temperature tolerance may allow conjugation to the surface of thermophilic bacteria, which may be of use for assembling enzymes at the cell surface to create cell factories.

In this chapter I have demonstrated that SnoopLigase2 and SnoopTag2/Catcher2 are capable of rapid ligation of model substrates across a wide range of conditions. Both systems are capable of modular coupling of TG2 to cargo proteins. SnoopLigase2 has the advantage over SnoopTag2/Catcher2 in that both partners for coupling only require short peptides to be fused rather than a ~15 kDa Catcher protein. In addition, the stable non-covalent association of SnoopLigase2 with the ligated product offers the opportunity to introduce additional functionality.

Chapter 5 – Decorating Endogenous Proteins with TG2

5.1. Introduction

Second generation systems for SnoopLigase and SnoopTag/Catcher are promising tools for modular coupling of cargo proteins to $TG2_{465}$. Such TG2:cargo constructs could be used to robustly display cargo in protein scaffolds via interactions between TG2 and these surfaces (Figure 5.1). Decoration of endogenous protein scaffolds with proteins may be useful for tissue culture, tissue engineering and bio-inspired photonics (Shi et al., 2020) (Majumdar et al., 2018).

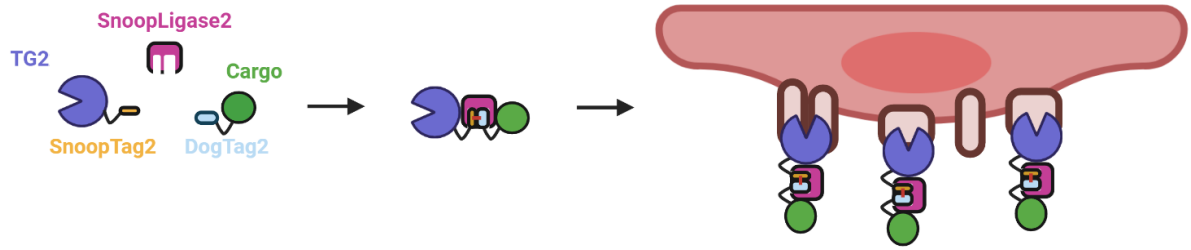


Figure 5.1. TG2-mediated decoration of the cell surface. A schematic illustrating modular coupling of cargo to $TG2_{465}$ to facilitate robust decoration of the cargo at the cell surface through TG2-cell surface interactions.

Transforming Growth Factor α (TGFα) is a ligand for Epidermal Growth Factor Receptor (EGFR). Binding of TGFα to EGFR initiates receptor dimerization, which brings together the cytosolic kinase domains of the EGFRs. Dimerization activates the kinase activity of EGFR, resulting in autophosphorylation. Phosphorylated tyrosine residues of the cytosolic domains act as docking sites for proteins such as Grb2, PI3K and STAT1. Once recruited, these proteins become activated and trigger downstream signalling pathways. Grb2 activates the MAPK/ERK pathway, PI3K activates the Akt pathway, and STAT1 activates the STAT pathway (Krall et al., 2011). TGFα-induced signalling promotes cell proliferation and Epithelial-Mesenchymal Transition (EMT) and is implicated in development and tissue repair (Singh and Coffey, 2014). TGFα has been explored as a therapeutic for tissue repair after ischemic stroke (Leker et al., 2009) (Alipanahzadeh et al., 2014). Previous work has demonstrated that localisation of cytokines to the cell surface can enhance their activity (Curnis et al., 2004)

(Dickerson et al., 2004). Thus, anchoring TGFα in the extracellular space via TG2 may prolong its proliferative effects.

In this chapter, I used SnoopLigase2 to ligate model cargo to TG2. I developed an optimised protocol for production, and isolation of, TG2:cargo products. I demonstrated robust TG2-mediated retention of cargo to a fibronectin scaffold and at the cell surface. I then created a TG2:TGFα conjugate and performed a series of cell assays to demonstrate enhanced activity of TG2:TGFα compared with TGFα alone.

5.2. Modular Assembly of Cargo to TG2x

5.2.1. SnoopLigase2-mediated Assembly

Based on TG2 engineering from Chapter 3, the variant $TG2_{465}$ C230S was used in this chapter and named TG2x for convenience. I genetically fused SnoopTag2 or DogTag2 at the N- or C-terminus of TG2x (Figure 5.2A). All constructs were soluble and properly folded as determined by cadaverine assay of TG2 activity (Figure 5.2B). While all constructs displayed transamidase activity, SnoopTag2 fusions showed lower activity than wild-type or the DogTag2 fusions. A possible explanation is that one or more of the 3 lysine residues in SnoopTag2 are acyl-acceptors for TG2. In this case SnoopTag2 would compete with biotin-cadaverine for conjugation to casein, resulting in a lower signal. Transamidation of lysine residues from SnoopTag2 could be confirmed and specific residues identified by tandem mass spectrometry. The issue of competing transamidation highlights the limitations of the cadaverine assay and the general challenge when comparing activity of different TG2 constructs. Next, I tested SnoopLigase2-mediated ligation of TG2x-DogTag2 to model substrate SnoopTag2-sfGFP. A 3-fold excess of SnoopTag2-sfGFP over TG2x-DogTag2 was enough to achieve 99% depletion of TG2x-DogTag2 (Figure 5.2C,D).

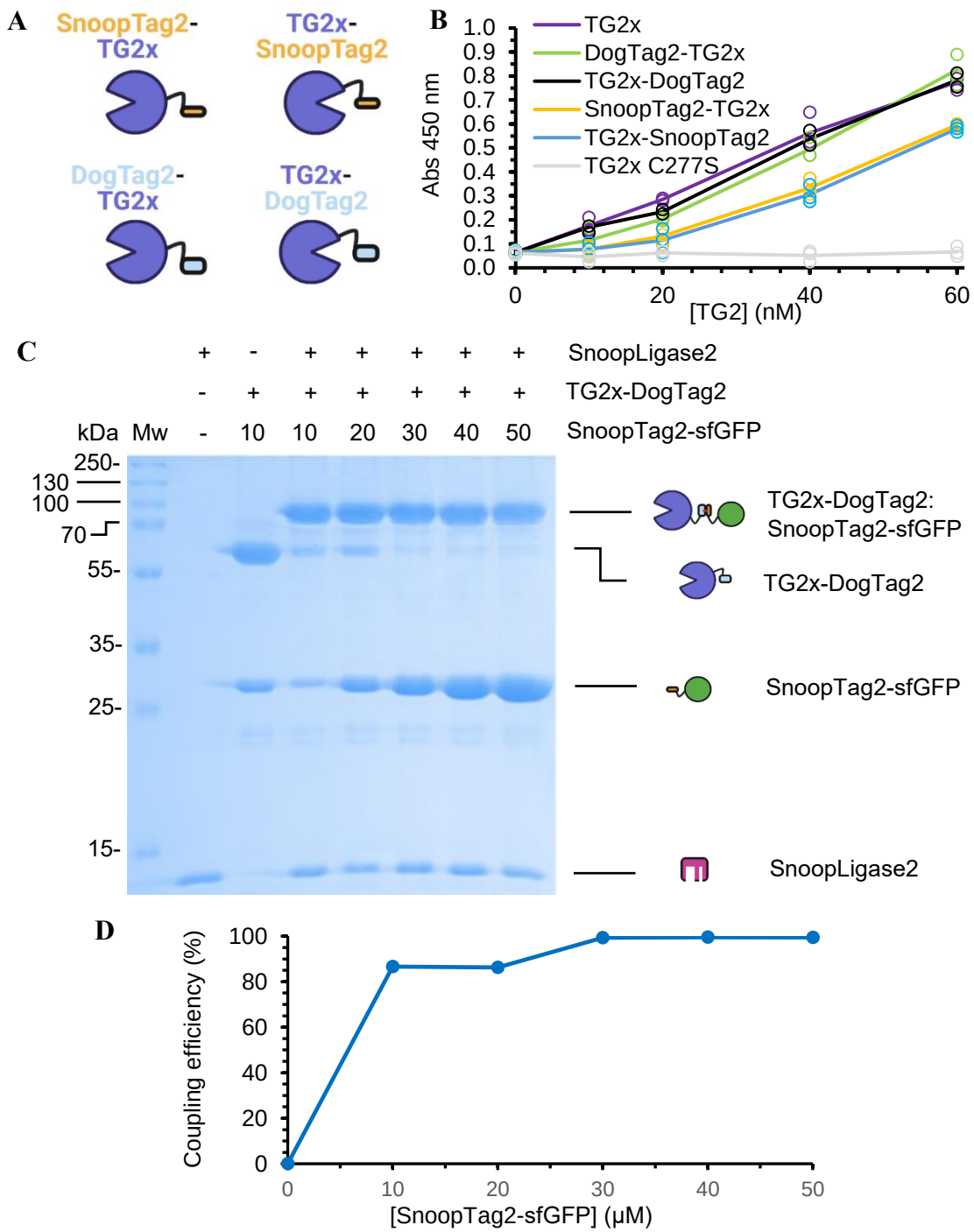


Figure 5.2. Modular coupling of sfGFP to TG2x by SnoopLigase2. (A) Schematic of the tagged TG2x constructs used for testing. DogTag2 or SnoopTag2 are fused to either N- or C- terminus. (B) Cadaverine assay of tagged TG2x constructs. Assays were performed for 30 min at 37 °C with varying concentration of the indicated TG2 variant. Crosses show each of the triplicate data points, with the line indicating the mean. (C) Coupling of TG2x-DogTag2 to SnoopTag2-sfGFP by SnoopLigase2. SnoopTag2-sfGFP was incubated with 10 μM of TG2x-DogTag2 and 20 μM SnoopLigase2 in HEPES-buffered saline pH 7.4 for 16 h at 4 °C. Reactions were analysed by 12% SDS-PAGE with Coomassie staining. All samples were reduced with 100 mM DTT before analysis. Mw = molecular weight markers. Shown is one representative assay of three total assays performed. (D) Quantification of coupling efficiency from (C), measured as depletion of TG2x-DogTag2 substrate.

Highly efficient ligation of TG2x to cargo was demonstrated with SnoopTag2 at either terminus (Figure 5.3A). Ligation of AviTag-DogTag2-MBP to SnoopTag2-TG2x and TG2x-SnoopTag2 depleted 97−98% of the TG2x substrate. I performed the equivalent experiment with DogTag2 at either termini (Figure 5.3B). Ligation of SnoopTag2-sfGFP to DogTag2-TG2x and TG2x-DogTag2 depleted >99% of TG2x substrate. I also demonstrated ligation of TG2x to sfGFP where DogTag2 was situated at an internal loop of sfGFP (Figure 5.3C). In this case, depletion of TG2x substrate was possible but a large excess of DogTag2-sfGFP loop A was required.

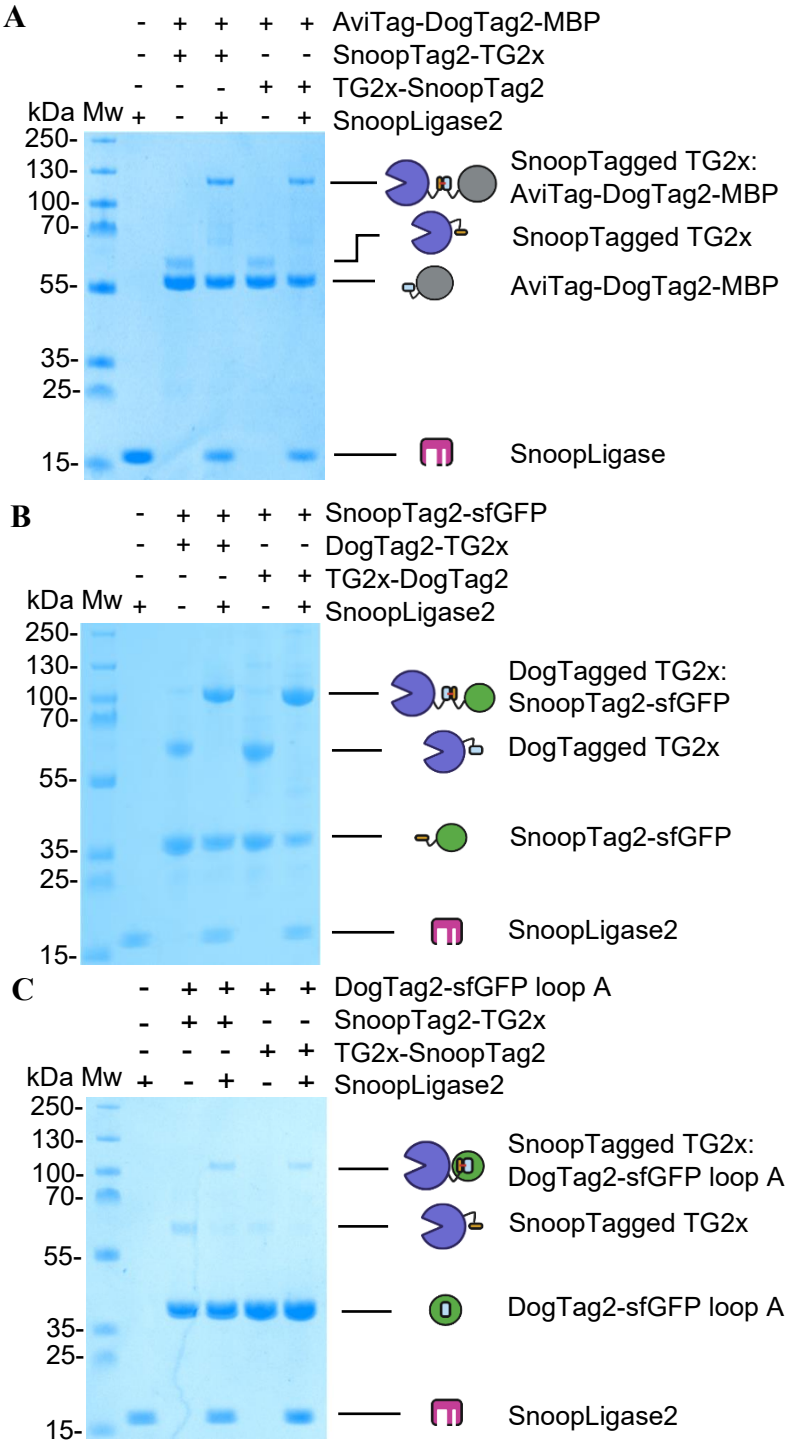


Figure 5.3. Efficient SnoopLigase2-mediated ligation with the tags at either terminus or an internal site. (A-C) - analysis of SnoopLigase2 reactions by 12% SDS-PAGE with Coomassie staining. All samples were reduced with 100 mM DTT before analysis. Mw = molecular weight markers. (A) SnoopTag2 can be ligated at the N- or C-terminus. 1 μM TG2x with SnoopTag2 at the N- or C-terminus was coupled with 7.5 μM AviTag-DogTag2-MBP by incubation with 5 μM SnoopLigase2 for 16 h at 4 °C in PBS pH 7.4. (B) DogTag2 can be ligated at the N- or C-terminus. 10 μM TG2x with DogTag2 at the N- or C-terminus was coupled with 30 μM SnoopTag2-sfGFP by incubation with 20 μM SnoopLigase2 for 16 h at 4 °C in PBS pH 7.4. (C) SnoopLigase2 ligation with tags at an internal site. 2 μM TG2x bearing SnoopTag2 at the N- or C-terminus was coupled to 16 μM DogTag2-sfGFP loop A by incubation with 12 μM SnoopLigase2 for 16 h at 4 °C in PBS pH 7.4.

Near complete depletion of the TG2x-DogTag2 substrate made it possible to isolate the TG2x-DogTag2:SnoopTag2-sfGFP product from unreacted substrates by size-exclusion chromatography (Figure 5.4A,B). The purified ligated product was assayed for transamidase activity and shown to be active, but with decreased activity compared to DogTag2-TG2x (Fig. 5.4C). The observed decrease in activity may be due to sfGFP being a lysine-donor substrate for TG2 as demonstrated in Chapter 3, resulting in decreased coupling of biotin-cadaverine to casein.

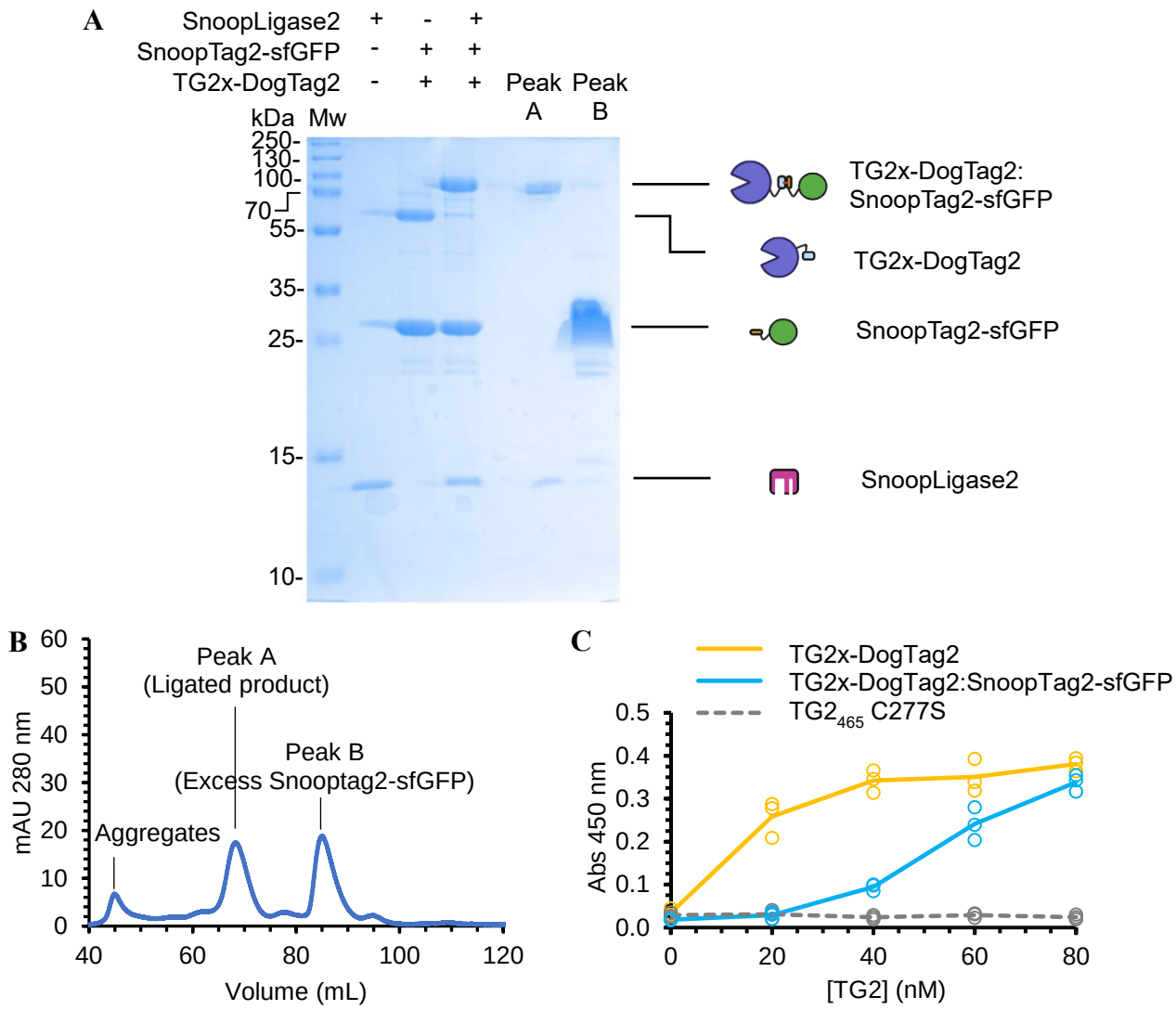


Figure 5.4. Preparation of TG2x:sfGFP ligation product. (A) Analysis of SnoopLigase2 ligation reaction between TG2x-DogTag2 and SnoopTag2-sfGFP by 12% SDS-PAGE with Coomassie staining. 10 µM TG2x-DogTag2, 30 µMSnoopTag2-sfGFP and 20 µM SnoopLigase2 were incubated for 16 h at 4°C in PBS pH 7.4. All samples were reduced with 100 mM DTT before analysis. Mw = molecular weight markers. (B) Size-exclusion chromatography trace of SnoopLigase2 ligation reaction. Reaction sample from (A) was loaded onto a Superdex 200 column. High molecular weight aggregates are commonly observed for both full length and truncated forms of TG2x. The product TG2x-DogTag2:SnoopTag2-sfGFP, non-covalently associated with SnoopLigase2, is separated from unreacted SnoopTag2-sfGFP substrate. (C) Cadaverine assay of TG2:cargo conjugate after isolation in (B). Assay was performed for 30 min at 37 °C with varying concentration of the indicated TG2 variant. Circles show each of the triplicate data-points, with the lines connecting the mean.

5.2.3. TG2-mediated Assembly

I previously demonstrated the propensity for $TG2_{465}$ to conjugate itself to Q-tagged substrates (Chapter 3). Here I explored this behaviour as a method for modular conjugation of cargo to TG2 to a model substrate. Ub(K0) is a ubiquitin variant with all lysine residues replaced with arginine. Qtag was fused to Ub(K0) to create Ub(K0)-Q, which was used as a model substrate. This lysine-deficient substrate was used to eliminate the possibility of formation of Ub(K0)-Q covalent dimers. A transamidation reaction was performed with $TG2_{465}$ and Ub(K0)-Q as a substrate, resulting in formation of a TG2465:Ub(K0)-Q (Figure 5.5A). Adding increased concentrations of Ub(K0)-Q was able to drive increased coupling efficiency, but only to ~50% (Figure 5.5B,C).

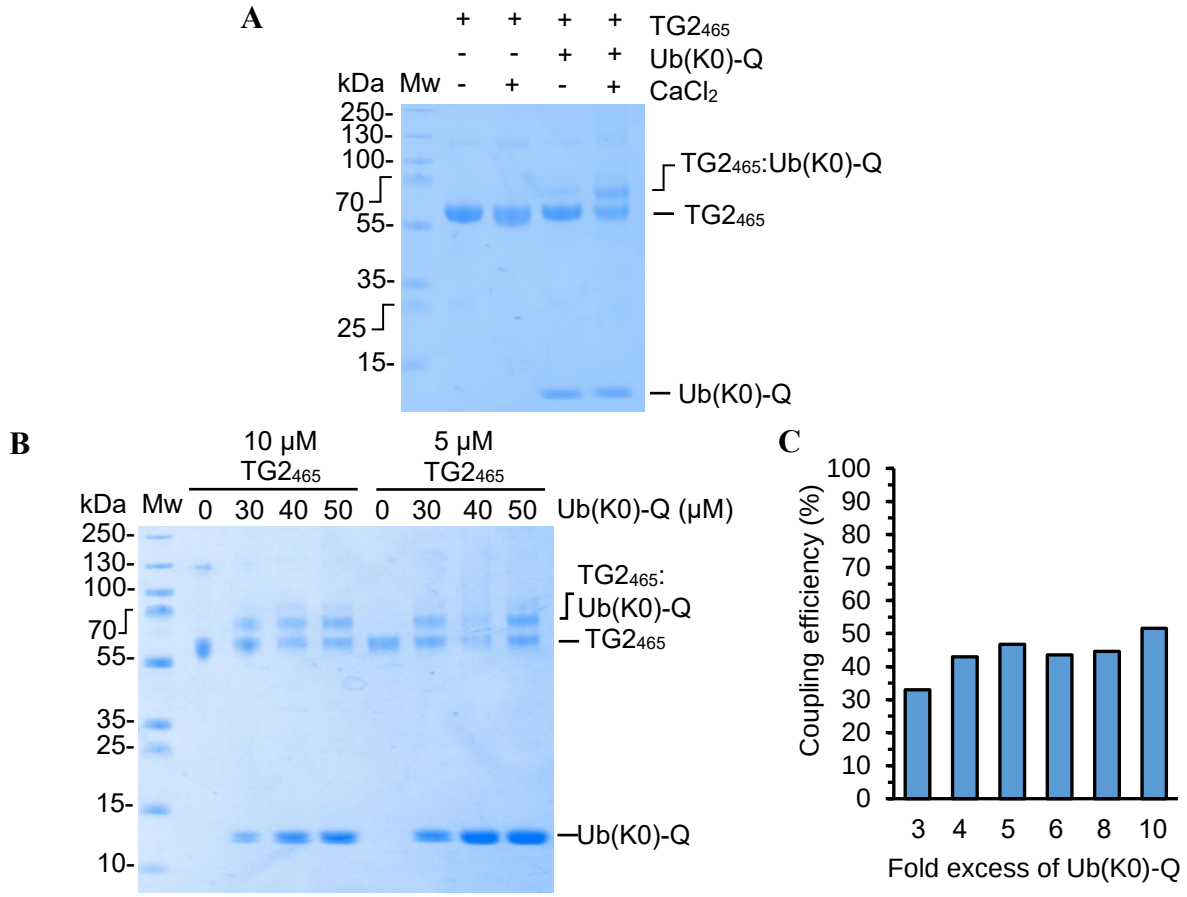


Figure 5.5. Transamidation of a model substrate to TG2$_{465}$. TG2$_{465}$ was assayed with Ub(K0)-Q to investigate the efficiency with which TG2$_{465}$ can conjugate a substrate protein to itself. (A) and (B) - analysis of transamidation by SDS-PAGE with Coomassie staining. All samples were reduced with 100 mM DTT before analysis. Thioester intermidiate was depleted with 100 mM hydroxylamine before analysis. Mw = molecular weight markers. (A) Transamidation reaction of 10 μM of TG2$_{465}$ and 30 μM Ub(K0)-Q incubated in HBS pH 7.4, 2 mM $CaCl_2$ and 1 mM DTT for 1 h at 37 °C. (B) Same as (A) with reactions for a range of concentrations of Ub(K0)-Q. (C) Quantification of coupling efficiency from (B) measured as ratio of product to unreacted TG2$_{465}$ substrate.

In addition to isopeptide bond formation, it has been demonstrated that transglutaminases can catalyse the reverse isopeptidase reaction (Parameswaran et al. 1997, Folk et al. 1967). Isopeptidase activity is a possible explanation for the inability to proceed the above reaction further than 50% complete. Previous work has attempted to separate the isopeptidase and transamidase activities of TG2 by mutation of the active site (Király et al., 2016). Variant W278F was identified in which a tryptophan residue adjacent to the catalytic cysteine is exchanged for phenylalanine; this mutant had unaffected transamidase activity but

has only 35% isopeptidase activity relative to wild-type (Király et al., 2016). I produced $TG2_{465}$ W278F and confirmed transamidase activity was equivalent to $TG2_{465}$ by cadaverine assay (Figure 5.6A). I then performed a transamidation reaction with $TG2_{465}$ W278F and an excess of substrate Ub(K0)-Q. $TG2_{465}$ W278F did not facilitate coupling efficiency beyond 50% (Figure 5.6B,C), suggesting that isopeptidase activity is not the cause of limited transamidation.

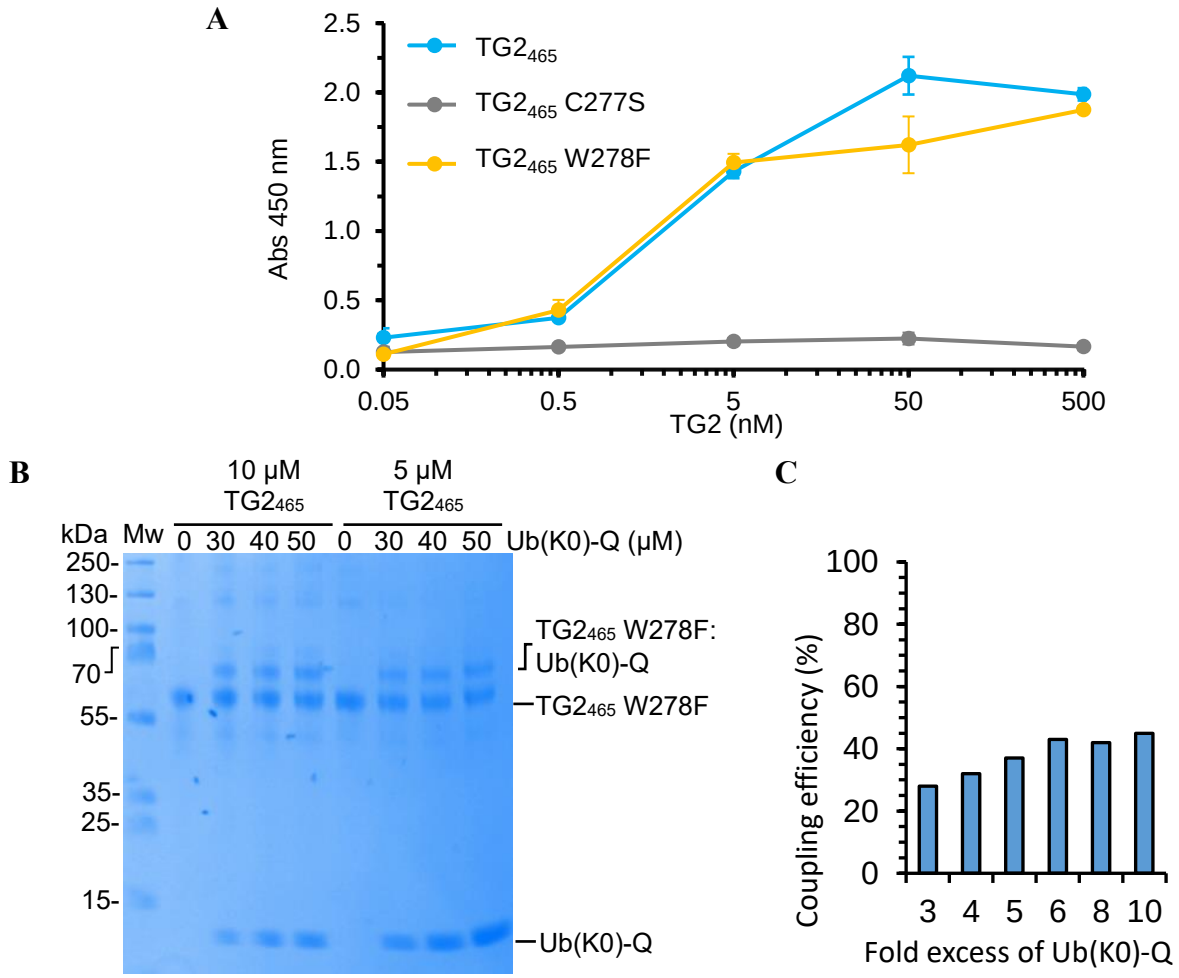


Figure 5.6. Testing transamidation of isopeptidase-impaired mutant $TG2_{465}$ W278F. $TG2_{465}$ W278F was applied for conjugation to Ub(K0)-Q to determine if inhibition of the reverse reaction results in higher coupling efficiency. (A) Cadaverine assay of $TG2_{465}$ W278F performed for 30 min at 37 °C with varying concentration of the indicated TG2 variant. Dots indicate the mean of triplicate data-points, error bars denote ± 1 S.D. (B) Analysis of transamidation by 12% SDS-PAGE with Coomassie staining. 10 µM of $TG2_{465}$ and 30 µM Ub(K0)-Q incubated in HBS pH 7.4 for 1 h at 37 °C. All samples were reduced with 100 mM DTT before analysis. Thioester intermediate was depleted with 100 mM hydroxylamine before analysis. Mw = molecular weight markers. (C) Quantification of coupling efficiency from (B) measured as ratio of product to unreacted $TG2_{465}$ substrate.
in this reaction.

Efficiency of TG2-mediated coupling may be limited by the capacity of TG2-reactive lysine residues to access the active site of $TG2_{465}$ once the enzyme-intermediate complex has been formed between $TG2_{465}$ and Ub(K0)-Q. This hypothesis was explored by performing a transamidation reaction with $TG2_{465}$-K and Ub(K0)-Q to provide a long, flexible lysine donor in the form of the LysTag. $TG2_{465}$-K had equivalent transamidase activity to $TG2_{465}$ (Figure 5.7A) and did not show improved coupling efficiency compared to $TG2_{465}$ (Figure 5.7B,C).

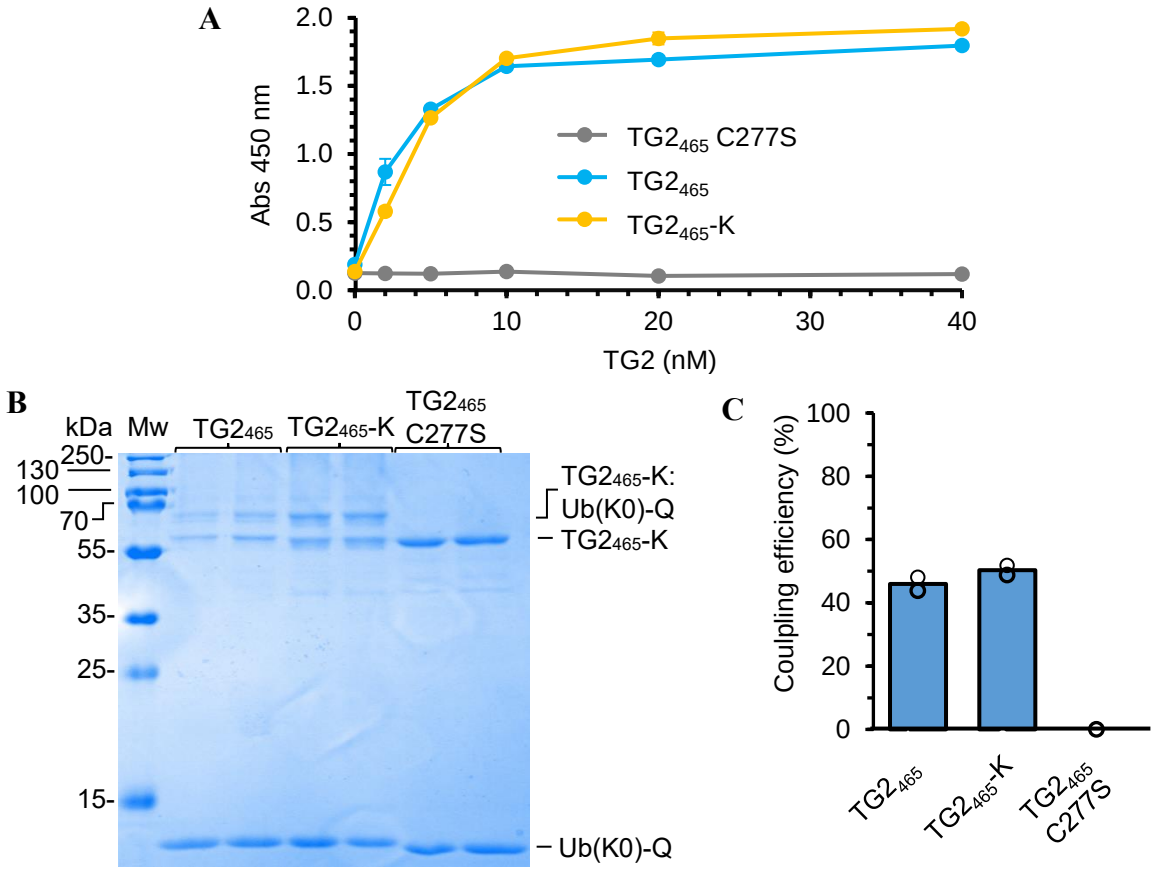


Figure 5.7. Testing transamidation of LysTag-linked $TG2_{465}$. $TG2_{465}$-K was applied for conjugation of Ub(K0)-Q to itself to determine if the presence of a flexible lysine substrate with access to the active site results in higher coupling efficiency. (A) Cadaverine assay of $TG2_{465}$-K was performed for 30 min at 37 °C with varying concentration of the indicated TG2 variant. Dots indicate the mean of triplicate data-points, error bars denote ± 1 S.D. (B) Analysis of transamidation reaction by 12% SDS-PAGE with Coomassie staining. 10 µM of $TG2_{465}$ and 30 µM Ub(K0)-Q incubated in HBS pH 7.4 for 1 h at 37 °C with duplicate samples. All samples were reduced with 100 mM DTT before analysis. Thioester intermediate was depleted with 100 mM hydroxylamine before analysis. Mw = molecular weight markers. (C) Quantification of coupling efficiency from (B) measured as ratio of product to unreacted $TG2_{465}$ substrate. Circles indicate datapoints and the bar denotes the mean.

5.3. Retention of $TG2_{465}$:cargo in Protein Scaffolds and at the cell surface

I previously demonstrated high affinity binding of $TG2_{465}$ to fibronectin and the cell surface (Chapter 3). Here, I used a plate assay to test the ability of a TG2:cargo conjugate to anchor to fibronectin, using model substrate SnoopTag2-sfGFP (Figure 5.8A). TG2x:sfGFP was incubated on a plate coated with fibronectin 45kDa fragment at 37 °C for 1 h then subjected to a series of washes with PBS pH 7.4 + 0.05% (*v/v*) TWEEN 20. Even at the low concentration of 10 nM, TG2x-DogTag2:SnoopTag2-sfGFP bound to the scaffold and very little was lost after 8 washes, demonstrating robust decoration (Figure 5.8B).

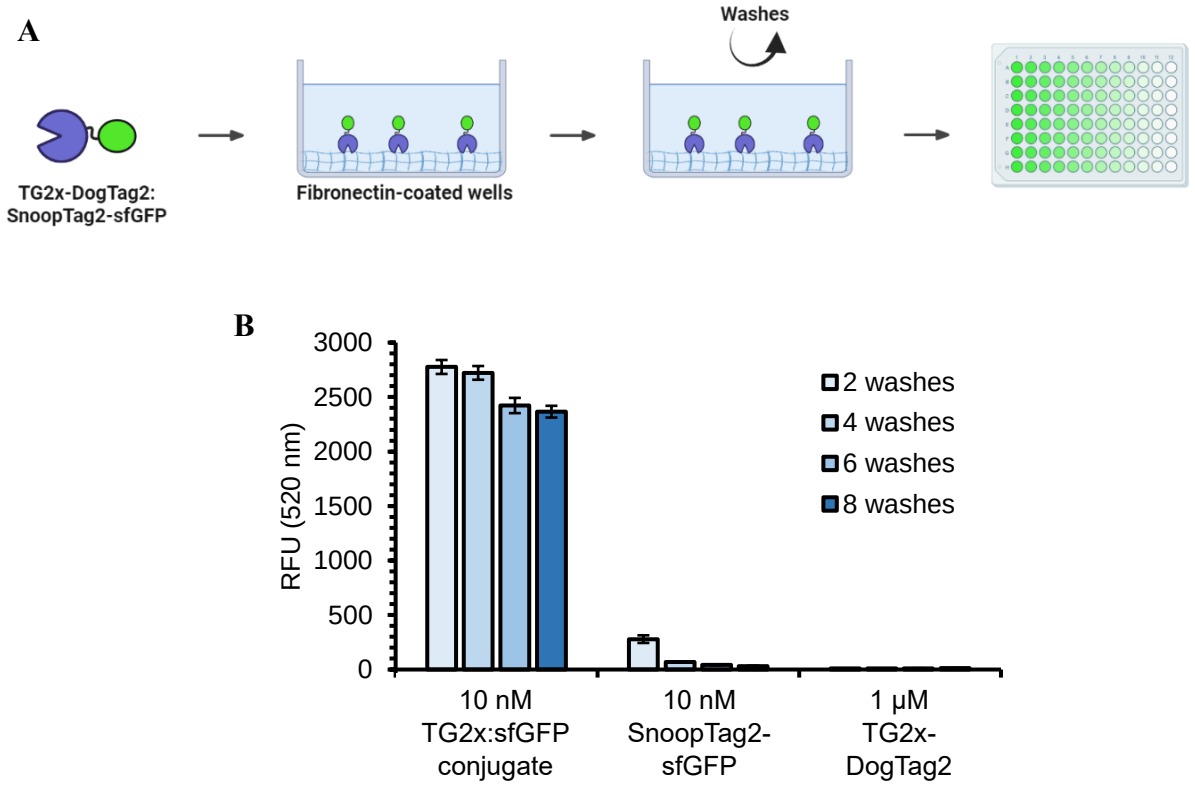


Figure 5.8. Demonstrating robust retention of a TG2:cargo construct in a fibronectin scaffold. TG2x-DogTag2:SnoopTag2-sfGFP was incubated with a fibronectin scaffold and subjected to a series of PBS-T washes followed by detection of sfGFP. (A) Schematic of the plate assay developed for measuring retention of GFP. (B) Plate assay of TG2x-DogTag2:SnoopTag2-sfGFP. All constructs were incubated in wells coated with 45 kDa fibronectin fragment for 1 h in PBS pH 7.4 at 37 C. Washes were performed with PBS pH 7.4 + 0.05% (v/v) TWEEN 20. sfGFP fluorescence was measured with 485 nm excitation and 510 nm emission. RFU = relative fluorescent units. Bars denote mean of three replicates and error bars denote ± 1 S.D. Some error bars are too small to be visible

I then performed a similar assay for binding of a TG2x conjugate to the cell surface (Figure 5.9A). In this case, a TG2x-DogTag2:mRuby-SnoopTag2 was prepared and then incubated with A431 cells for 30 min and then washed 8 times with RPMI media before analysis by flow cytometry. Clear signal was detected for both mRuby and TG2 with an anti-TG2 antibody, demonstrating robust retention of the conjugate at the cell surface. mRuby alone was not detected at the cell surface (Figure 5.9B).

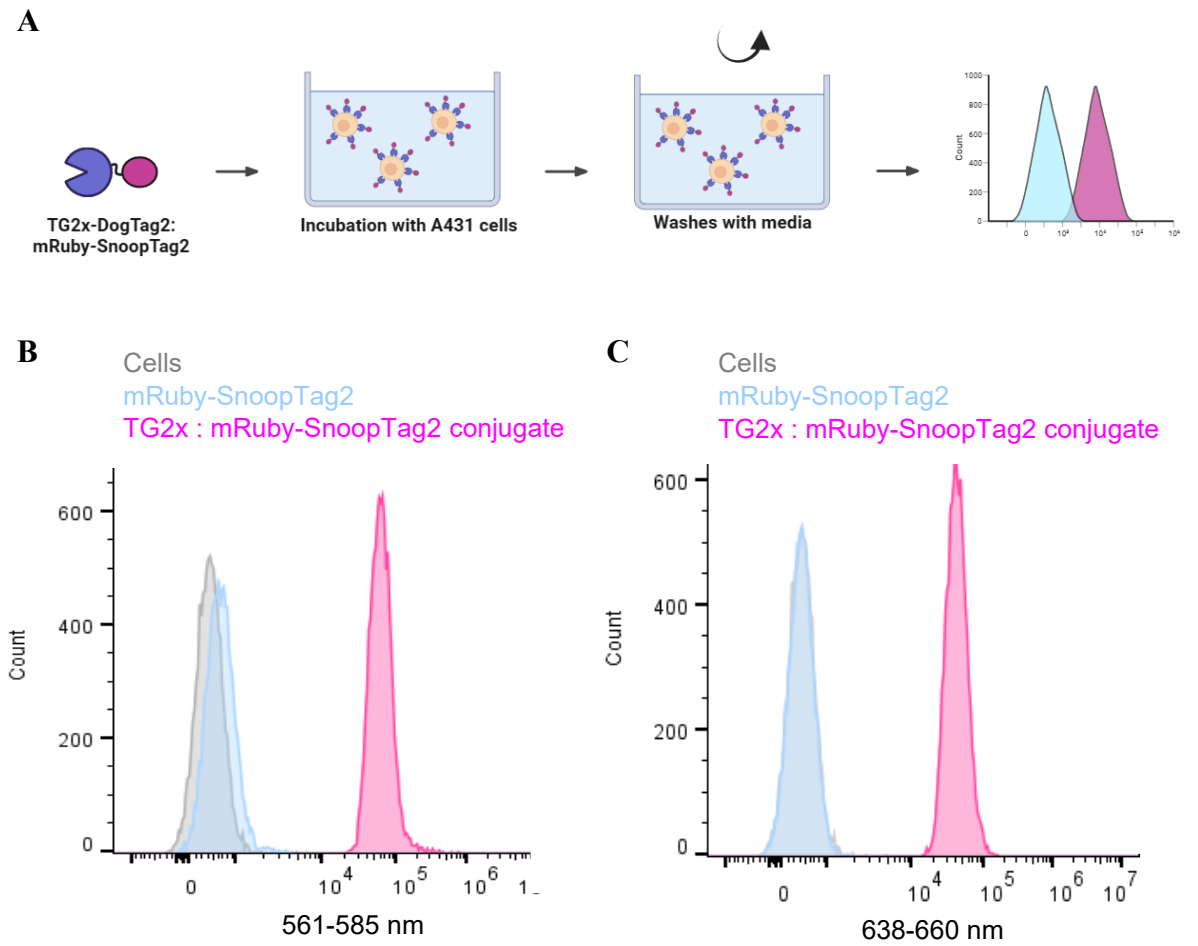


Figure 5.9. Demonstrating robust retention of a TG2:cargo construct at the cell surface. TG2x-DogTag2:mRuby-SnoopTag2 was incubated with 1×10^7 A431 cells and subjected to a series of PBS-T washes followed by detection of mRuby. (A) Schematic of the assay used to detect retention of mRuby. (B-C) 50 nM of protein was incubated with A431 cells for 30 min in HBS pH 7.4 + 1% (*w/v)* BSA at 37 ˚C. After incubation, cells were washed 8 times in HBS pH 7.4 + 1% (*w/v)* BSA, followed by staining with rabbit anti-TG2 primary antibody and anti-rabbit 647 secondary antibody. (B) Cells measured with 561 nm excitation and 585 nm emission to detect mRuby. (C) Cells measured with 638 nm excitation and 660 nm emission to detect TG2.

5.4. Enhanced Activity of a Growth Factor by TG2-mediated Retention

Recombinant TGFα expresses in inclusion bodies in *E. coli*. I expressed TGFα with an N-terminal SnoopTag2 in *E.coli*, solubilised the inclusion bodies in 8 M Urea, and then purified the SnoopTag2-TGFα product (Figure 5.10A). TGFα has three disulfide bonds that are essential for proper folding (Bean and Carr, 1992); these disulfides must be formed during refolding to produce soluble TGFα in active form (Groenen et al., 1994). I refolded the purified SnT2-TGFα using the 'shock dilution' method (Wang et al., 2017b), with glutathione in the buffer to aid disulfide bond formation. SEC analysis of soluble refolded TGFα revealed several peaks, including some small peptide fragments most likely the result of digestion by contaminant protease (Figure 5.10B). The peak corresponding to monomeric SnoopTag2-TGFα was collected and analysed by Mass Spectrometry. Analysis confirmed that the sample was refolded product with the expected mass for SnoopTag2-TGFα with three disulfide bonds (Figure 5.10C).

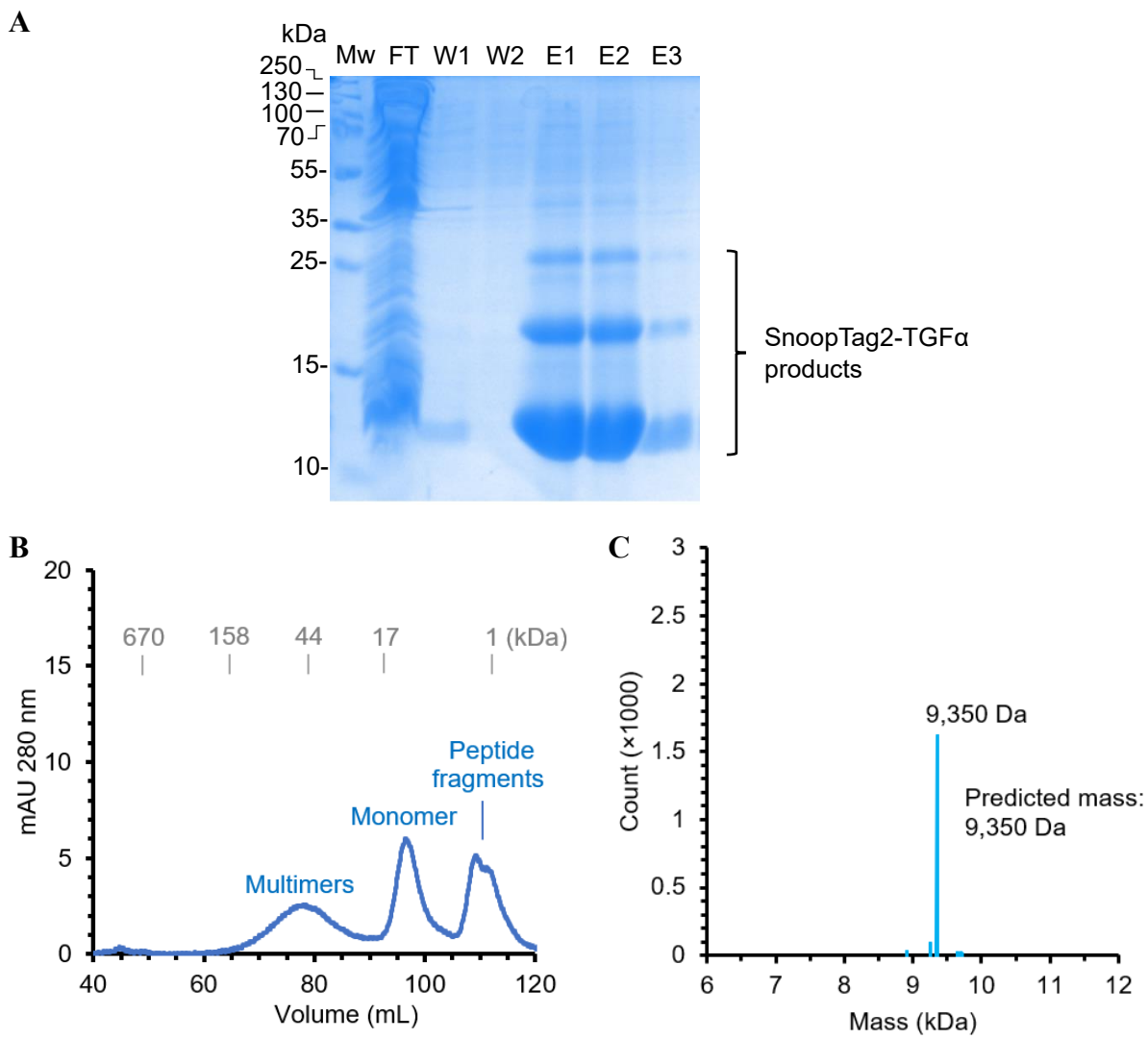


Figure 5.10. Purification and refolding of SnoopTag2-TGFα. (A) Ni-NTA purification of denatured SnoopTag2-TGFα from resolubilised inclusion bodies. Analysis by 18% SDS-PAGE with Coomassie staining. Samples were non-reduced. FT = flow-through, W = washes, E = elution. Column washes performed with 10 mM imidazole, elution performed with 200 mM imidazole. One representative prep shown from two performed in total. (B) Size-exclusion chromatography of SnoopTag2-TGFα after denaturing and refolding shows several peaks of soluble products. (C) Electrospray ionization mass spectrometry detects a 9350 Da product corresponding to SnoopTag2-TGFα with 3 disulfide bonds.

A genetic fusion of TG2x and TGFα was insoluble when expressed in *E. coli* (Figure 5.11A) and, in contrast to SnoopTag-TGFα, aggregated during refolding (Figure 5.11B). Thus, post-translational coupling of TG2x and TGFα after separate expression is required to generate a TG2x-TGFα product. I used SnoopLigase2 to achieve highly efficient ligation of TG2 to TGFα variants (Figure 5.11C). The ligated TG2-DogTag2:SnoopTag2-TGFα product was then purified from excess substrate by size-exclusion chromatography (Figure 5.11D). TGFα (R42A) shows substantially decreased TGFα-induced activation of EGFR (Boyles et al., 2016) and was produced for use as a negative control in TGFα activity assays (Appendix Figure 5).

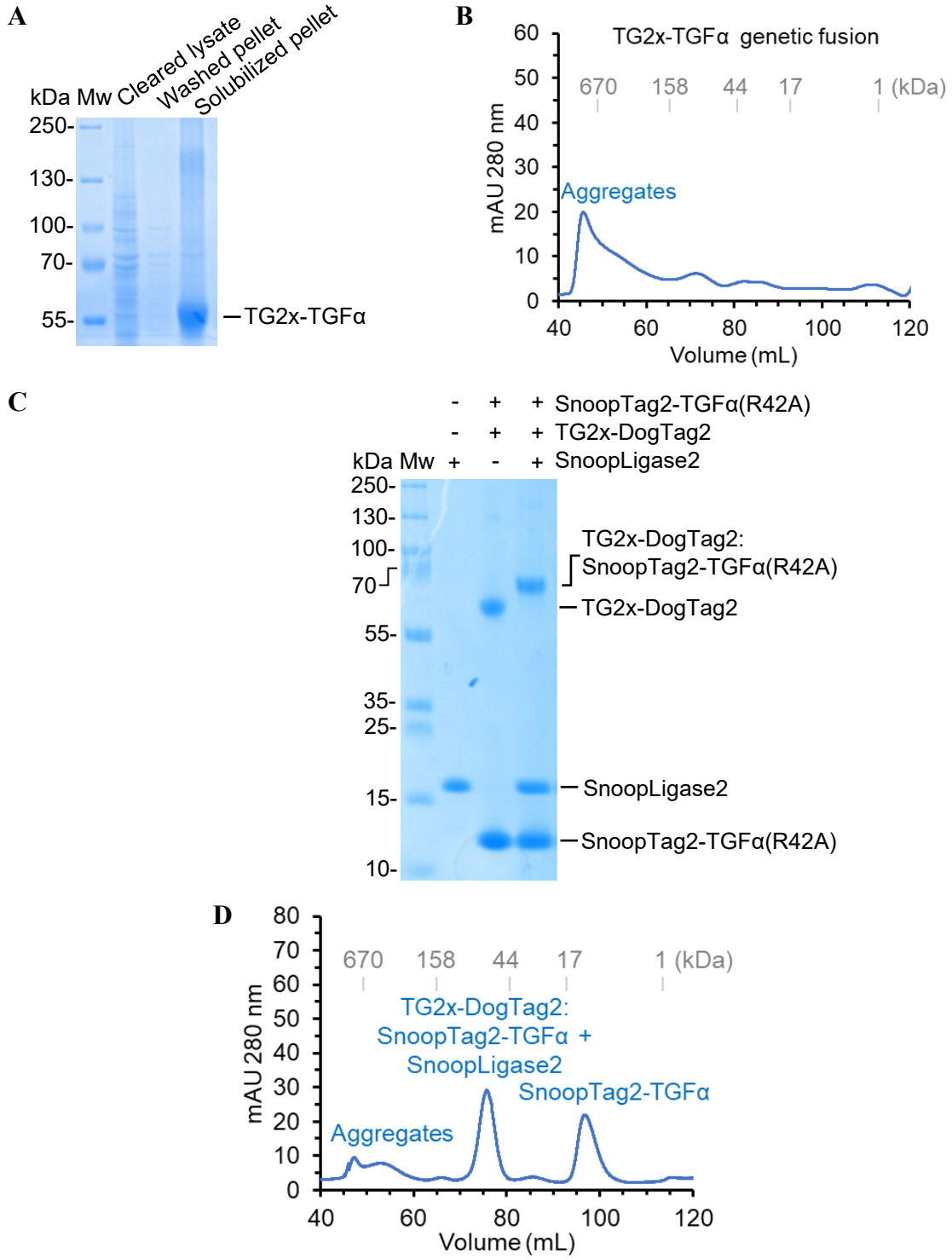


Figure 5.11. TG2x can be efficiently connected to TGFα by ligation but not genetic fusion. (A) Analysis of TG2-TGFα expression in *E. coli*, by 12% SDS-PAGE with Coomassie staining. All samples were reduced with 100 mM DTT before analysis. (B) Size-exclusion chromatography of the TG2-TGFα construct after denaturing and refolding shows incorrectly folded aggregates. (C) Analysis of SnoopLigase2-mediated ligation of TG2x and TGFα(R42A) by 12% SDS-PAGE with Coomassie staining. 10 µM TG2x-DogTag2, 20 µM SnoopTag2-TGFα(R42A) and 15 µM SnoopLigase2 were incubated for 16 h at 4 °C. All samples were reduced with 100 mM DTT before analysis. Mw = molecular weight markers.

(D) Size-exclusion chromatography of the reaction from (C) results in separation of the ligated product from unreacted substrate.

I tested the ability of TG2x-DogTag2:SnoopTag2-TGFα to robustly decorate the cell-surface by incubating the construct with the DU145 prostate cancer cell-line, and subjecting the cells to a series of washes (Figure 5.12A). The DU145 cell line was chosen as it contains a mutation in Akt that confers sensitivity to TGFα signalling. I then used a resazurin assay to measure the effect of TGFα activity on cell proliferation. SnoopTag2-TGFα increased cell proliferation of DU145 cells to a level equivalent to treatment with 10% (*v/v*) fetal bovine serum (FBS) (Fig. 5.12B). When cells were washed after addition of SnoopTag2-TGFα, cell proliferation was no higher than that observed for cells incubated in RPMI alone. Cell proliferation was significantly greater after washing when cells were treated with TG2x-DogTag2:SnoopTag2-TGFα compared with SnoopTag2-TGFα ($p < 0.001$, $n = 3$). Cell proliferation after TG2x-DogTag2:SnoopTag2-TGFα treatment and washes was equivalent to proliferation of cells treated with SnoopTag2-TGFα without washes (ns, $n = 3$). The observed increase in cell proliferation is not due to TG2 itself because TG2x-DogTag2 had no effect on cell proliferation. In addition, the transamidase activity of TG2 is not responsible for increased cell proliferation because the catalytically inactive mutant $TG2_{465}$ C277S-DogTag2:SnoopTag2-TGFα produces equivalent cell proliferation to TG2x-DogTag2:SnoopTag2-TGFα (ns, $n = 3$). Finally, TG2x-DogTag2:SnoopTag2-TGFα(R42A) had no effect on cell proliferation, indicating that the observed increase in cell proliferation by TG2x-DogTag2:SnoopTag2-TGFα is caused by prolonged TGFα activity. These findings suggest that TG2 can decorate the cell surface via non-covalent interactions, resulting in enhanced TGFα activity.

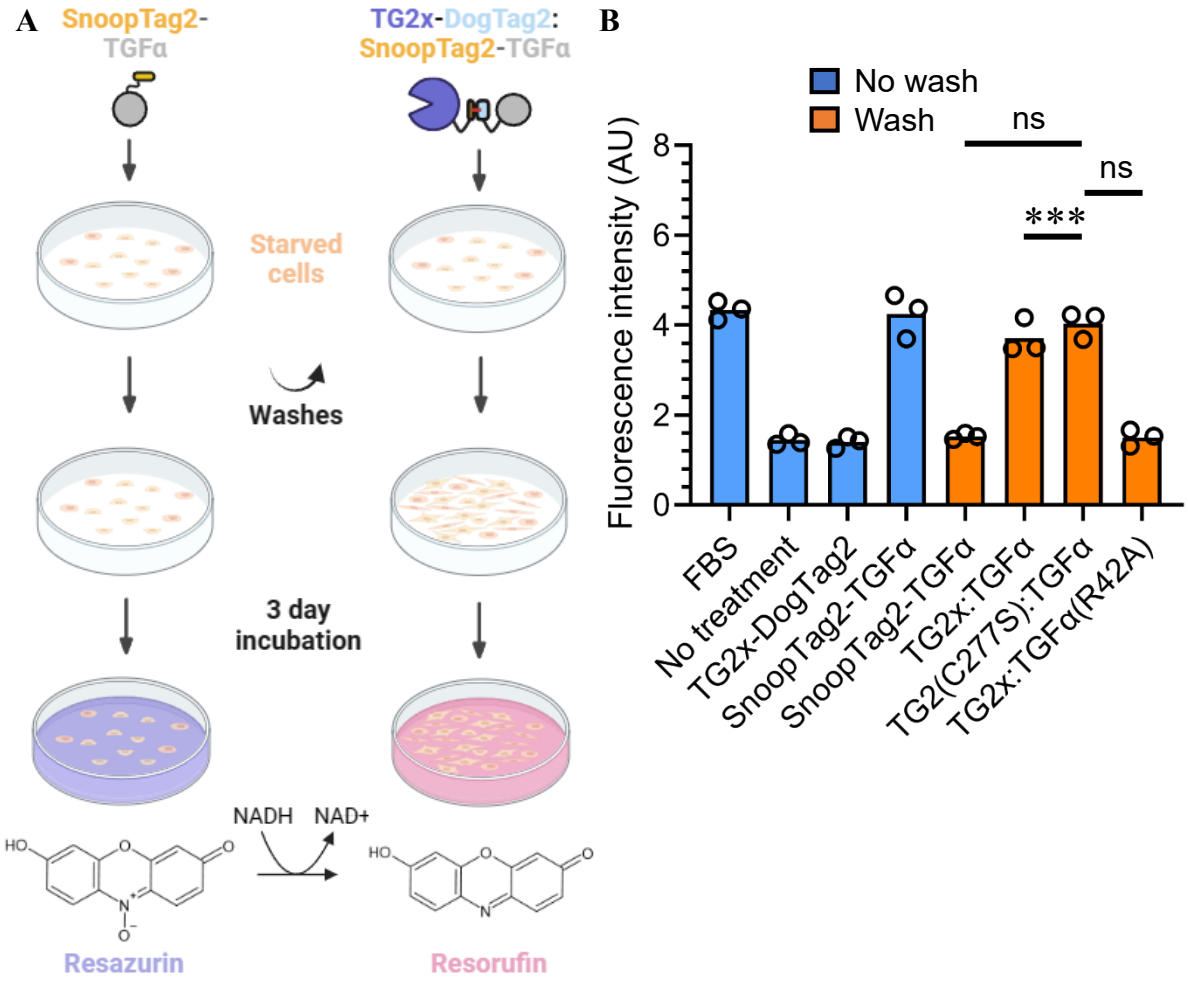


Figure 5.12. TGFα-induced cell proliferation is enhanced by TG2-mediated anchoring to the cell surface. (A) Schematic of the cell assay for testing the ability of TG2x-DogTag2:SnoopTag2-TGFα to decorate the cell surface and allow prolonged TGFα activity. Starved DU145 cells were seeded for 3 h to allow attachment. Once attached, FBS or the indicated protein was added to a final concentration of 100 nM and cells were incubated for 30 min. Selected samples were washed five times with RPMI 1640, and all samples were incubated for 3 days in RPMI 1640. Differences in cell proliferation were measured using the resazurin assay. (B) Non-covalent TG2 anchoring enhanced TGFα-induced proliferation, measured by resazurin assay. Samples in orange were washed after incubation with protein, while blue samples were not washed. Circles show each triplicate data point, with the bar indicating the mean. Selected pairs were compared using an unpaired two-tailed Student's t-test (n = 3, ***p < 0.001). AU = arbitrary units.

DU145 cells present epithelial morphology and upon incubation with TGFα EMT is induced (Qin et al., 2014). During EMT, epithelial cells lose the capacity for cell-cell adhesion and gain migratory and invasive properties (Kalluri and Weinberg, 2009). This process is essential for a range of physiological and pathological processes including embryonic

development, wound healing, tissue regeneration, and cancer progression. A distinct morphological change is also associated with EMT. As cells abolish cell-cell junctions and remodel the cytoskeleton the cell shape becomes spindle-like, allowing cells to move more freely within tissues (Kalluri and Weinberg, 2009).

I incubated DU145 cells with TG2x:TGFα constructs and then stained cells with the actin-binding ligand phalloidin. Cells were analysed by fluorescence microscopy to allow determination of cell shape. Cell shape was analysed in ImageJ and quantified based on the cell aspect ratio. TGFα activity promoted a morphological change from rounded to spindle-shaped cells (Figure 5.13). Cells incubated in only RPMI media had a rounded morphology with a median aspect ratio of 1.7. Cells treated with SnoopTag2-TGFα formed elongated spindle shapes with a significantly larger median aspect ratio of 3.1 ($p < 0.001$, $n = 51$). Similarly, cells washed after addition of TG2x-DogTag2:SnoopTag2-TGFα also showed a significantly larger median aspect ratio ($p < 0.001$, $n = 51$). In contrast, the morphology of cells washed after treatment with SnoopTag2-TGFα was not significantly different to the no treatment condition (ns, $n = 51$).

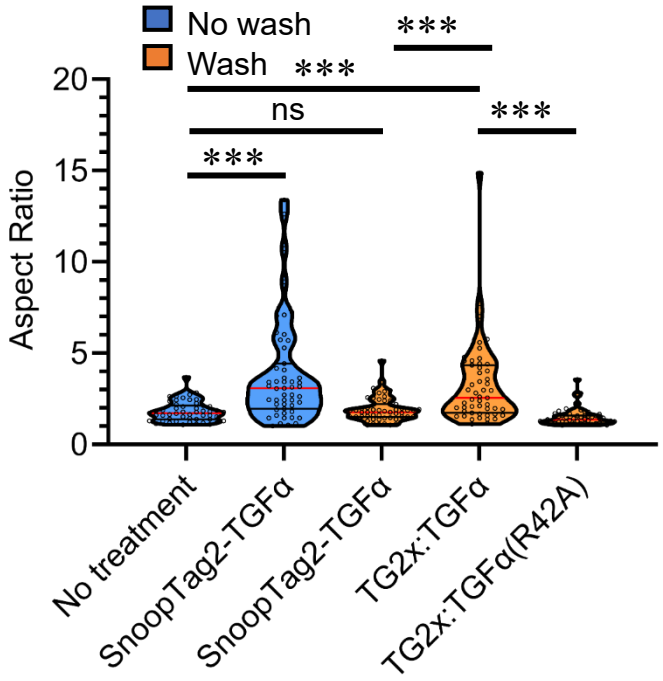


Figure 5.13. TGFα-induced cell morphology changes are enhanced by TG2-mediated anchoring to the cell surface. (A) TG2 anchoring increased TGFα-induced change in cell morphology. DU145 cells were seeded for 3 h to allow attachment. Once attached, FBS or the indicated protein was added to a final concentration of 100 nM and cells were incubated for 30 min. Selected samples were washed five times with RPMI 1640, and all samples were incubated for 3 days in RPMI 1640. Samples in orange were washed with media after 30 min incubation with protein. Blue samples were not washed. Each cell is marked as a black circle. The distribution is marked as a violin plot. The red line denotes the median and the black lines denote the interquartile range. Selected pairs were compared using an unpaired two-tailed Welch's t-test (n = 51, ***$p < 0.001$).

Specific molecular changes associated with EMT include down-regulation of cell adhesion proteins such as E-Cadherin, and remodelling of the cytoskeleton, including upregulation of Vimentin (Kalluri and Weinberg, 2009). I further analyzed TGFα-induced effects by immunostaining DU145 cells for E-Cadherin and Vimentin (Figure 5.14). DU145 cells incubated in RPMI expressed high levels of E-Cadherin and did not express detectable levels of Vimentin. Cells treated with SnoopTag2-TGFα had increased Vimentin expression and decreased E-Cadherin expression, suggesting the cells were in the process of EMT. Washing cells after addition of SnoopTag2-TGFα had no effect on Vimentin and E-Cadherin

expression, while cells treated with TG2x-DogTag2:SnoopTag2-TGFα had upregulated Vimentin and downregulated E-Cadherin.

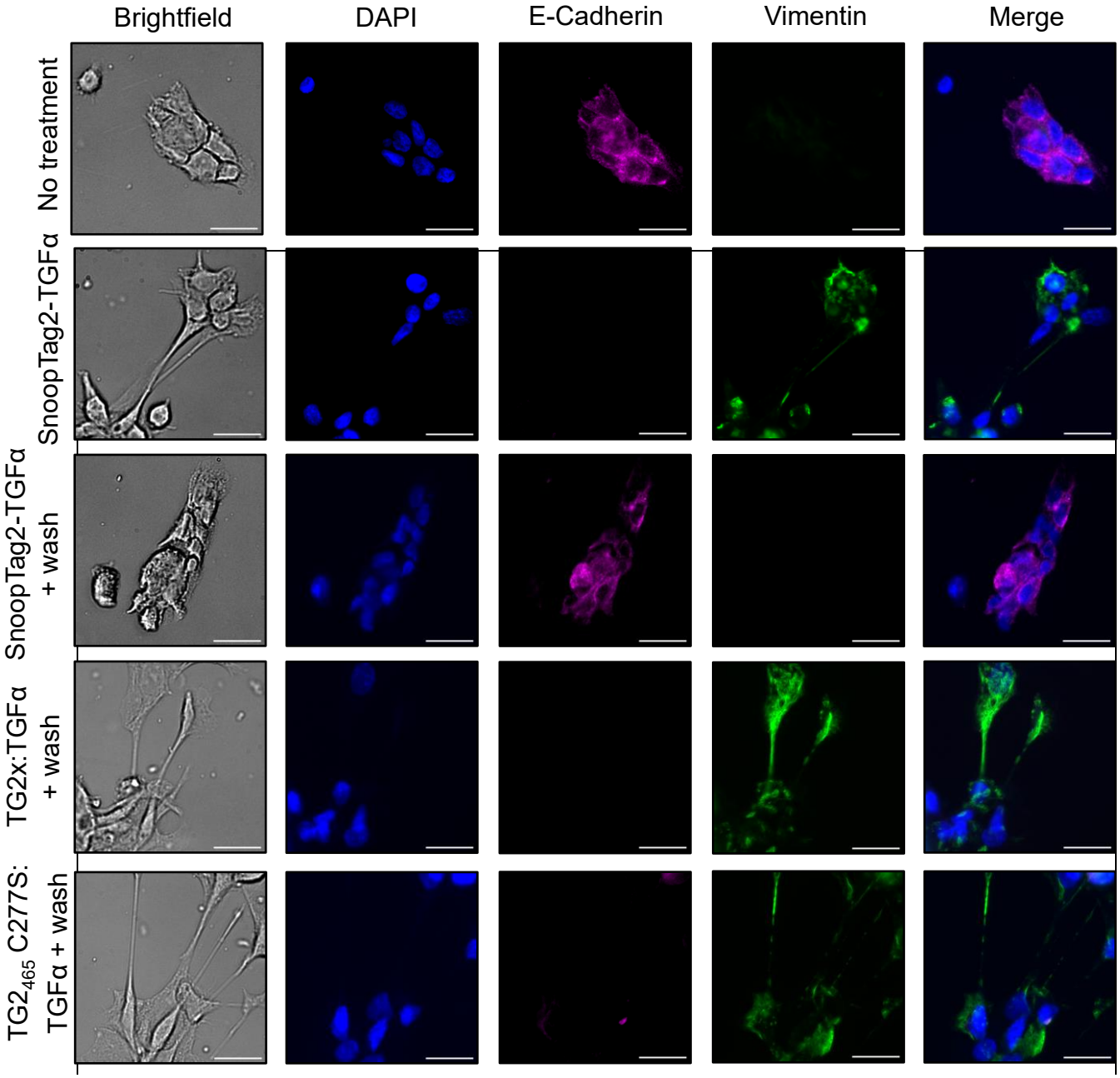


Figure 5.14. TGFα-induced changes in EMT marker expression are enhanced by TG2-mediated anchoring to the cell surface. TG2 anchoring enhanced TGFα-induced increase in Vimentin expression and decrease in E-cadherin expression, detected by fluorescence microscopy. DU145 cells were seeded for 3 h to allow attachment. Once attached, FBS or the indicated protein was added to a final concentration of 100 nM and cells were incubated for 30 min. Selected samples were washed five times with RPMI 1640, and all samples were incubated for 3 days in RPMI 1640. Cells are shown in brightfield (greyscale) and after staining with DAPI (blue) or antibodies to E-cadherin (magenta) and after staining with DAPI (blue) or antibodies to E-cadherin (magenta) and Vimentin (green). The right column is a merge of DAPI, E-cadherin, and Vimentin images. Scale bar = 15 μm.

5.5. Discussion

Both SnoopTag2 and DogTag2 expressed well at both termini and the resulting fusion proteins demonstrated highly efficient conjugation at all termini and an internal loop. DogTag has previously been inserted into proteins at internal sites, demonstrating highly efficient conjugation to DogCatcher from loops and regions of secondary structure (Keeble et al., 2022). Here I demonstrated that SnoopLigase2 can also direct rapid conjugation at an internal site, expanding the scope of SnoopLigase2 to proteins that are not amenable to conjugation at the termini. For example, extracellular conjugation to membrane proteins where both termini are intracellular. These findings demonstrate that SnoopLigase2 is a flexible system, capable of producing conjugates with a range of different architectures. This is particularly valuable for applications where orientation of conjugation is important, for example where one terminus contains a binding site which is integral for the desired application such as in immunotoxins and bispecifics. While highly efficient conjugation to an internal site has been demonstrated, this was in a flexible loop which represents a less challenging site of conjugation than regions of secondary structure. Future work to explore the ability of SnoopLigase2 to couple proteins to structurally complex internal sites would be valuable.

Many human proteins of therapeutic interest are insoluble in *E. coli* and require refolding through highly sensitive protocols which are often established by trial and error. It is highly desirable to be able to conjugate proteins together without having to re-establish such refolding protocols. SnoopLigase2 facilitated creation of a novel TG2:TGFα conjugate not possible by genetic fusion. In addition, many mammalian proteins require expression in mammalian expression systems for correct post-translational modification. Although not explored in this thesis, it has been established that SnoopTag2 and DogTag2 express well in mammalian systems (Keeble et al., 2023).

One disadvantage of the SnoopLigase2 system is its bacterial origin and the persistence of the tags and SnoopLigase2 element in the final product, limiting therapeutic applications. TG2-mediated conjugation of cargo has the benefit over the SnoopLigase2 system of not involving any microbially derived elements. However, coupling efficiency was limited and efforts to improve performance were ineffective. This inefficient coupling makes purification of the product challenging and optimisation is necessary for this to be a useful method of cargo coupling. Future work could explore reaction conditions to find the optimal conditions for transamidation, which are not necessarily the physiological conditions I have used throughout this thesis. This would not be surprising given the multiple layers of regulation of transamidation imparted upon TG2 in the physiological environment and the obvious dangers of highly efficient low specificity transamidation in the extracellular environment. Optimisation experiments have been performed for mTG and activity was found to be increased at higher temperatures and optimal at 50 °C (Liu et al., 2019), but this has not been explored for TG2.

In this chapter I used the SnoopLigase system to generate TG2:cargo conjugates for *in vitro* experiments. SnoopLigase is not an ideal system for the therapeutic application of TG2 due to persistence of bacterial elements in the final product. The stable association of SnoopLigase2 to the product did not appear to impair functional activity of TGFα, but issues with immunogenicity are likely. Previous work found that mice produced anti-SnoopLigase antibodies upon exposure to SnoopLigase-assembled VLPs (Andersson et al., 2019). It may be possible to remove the bound ligase with more stringent washes, but such conditions may impair the function of TG2 and/or the cargo. In any case, the two peptide tags would persist in the final construct and their immunogenicity is currently unknown. Application of TG2 for cell decoration is dependent upon a system to conjugate TG2 to cargo that is either human-derived or scarless. TG2-mediated coupling was explored in Chapter 3, and it may be possible to

improve efficiency to acceptable levels but there will be compatibility issues when proteins have multiple glutamine and lysine donor sites. Inteins are an option for scarless ligation but are not modular enough to be applied widely. Reactivity is highly dependent on the sequence context surrounding the reactive thiols (Pinto et al., 2020) (Sekar et al., 2022). The compatibility of a given protein with the available intein systems is not currently possible to predict in advance.

I have demonstrated that cargo can be robustly displayed in a scaffold or at the cell surface via TG2. There have been many previous efforts to localise therapeutics to the extracellular space (Addi, Murschel and De Crescenzo, 2017) (Battistini et al., 2021). One common method is the use of short peptide tags. RGD-containing peptide motifs have been used for localisation to the cell surface (Gilad et al., 2016), and collagen-binding peptides have been used for localisation to the ECM (Wahyudi et al., 2016). RGD peptides have had success, with low nanomolar affinity of RGD-drug conjugates for integrin demonstrated (Battistini et al., 2021). An RGD-TNFα construct was able to bind to endothelial cells with a picomolar affinity and demonstrate increased cytotoxic activity (Curnis et al., 2004). Success has been more limited for ECM targeting peptides. Many collagen-binding peptides identified in the literature contain a very high proportion of charged residues (Katsumata et al., 2019) (Federico et al., 2015) (Martino et al., 2014) (Lee et al., 2007). Highly charged peptides such as these are likely to be generally sticky and not specific to the ECM. Furthermore, fusions of protein cargo to collagen binding domains have resulted in only micromolar affinity for Type I collagen (Zhao et al., 2007) (Liang et al., 2016) (Andrades et al., 2001). TG2 may be a suitable alternative to collagen-binding peptides for localisation of therapeutic proteins to the ECM, a strategy which is of interest for tissue regeneration applications.

TG2 has known roles in cell adhesion (Shao et al., 2009), so it was not previously known if exogenous TG2 would interfere with induction of EMT. While EMT was unaffected

by exogenous TG2 in these experiments it is possible that TG2 had other undetected effects on DU145 cells. For example, TG2 has been implicated in Wnt/B-catenin signalling via binding to membrane receptor LRP5 which can contribute to cell differentiation (Deasey et al., 2013). Future experiments to test the ability of exogenous TG2 to activate this signalling pathway would be useful. Even if exogenous TG2 has no effect on DU145 cells, this may not be the case for other cell types. TG2 has been shown to influence expression of EMT markers in an ovarian carcinoma cell line which displays mesenchymal morphology (Shao et al., 2009). Thus, it would be beneficial to test cells relevant for wound healing applications such as primary fibroblasts and endothelial cells.

Finally, cells which express high levels of endogenous TG2 may saturate the TG2 binding sites on surface of these cells, leading to limited decoration by exogenous TG2. Tests of cell-surface retention of TG2 with differential levels of endogenous TG2 expression might give interesting insights. In summary, subsequent examination of the effects of TG2 on cells and the propensity for TG2 to bind to a range of cell types is required to establish if TG2 is a broadly useful tool for robust decoration of the cell surface.

Chapter 6 – Summary and General Discussion

The aim of this thesis was to develop a TG2-based tool for protein-protein conjugation. There are a range of protein-protein conjugation tools already available, most of which are of non-human origin and many of which require engineering of all substrates. The substrate specificity of TG2 lends well to a system for conjugation of a specific target protein to a wide range of endogenous proteins to facilitate broad decoration of protein scaffolds. A TG2-based system also has the potential to be entirely human-derived, and thus minimise immunogenicity issues for therapeutic applications. TG2 was engineered with the goal of directing covalent decoration of endogenous protein scaffolds with proteins of interest.

In Chapter 3, I established a protocol for quick and convenient recombinant expression of TG2 in *E. coli*. I used plate assays and flow cytometry to show that these TG2 constructs bind robustly to both the cell surface and fibronectin, and also bind (more weakly) to type I collagen and a basement membrane extract. These non-covalent interactions with ECM components and the cell surface prevent simple removal of TG2 from scaffolds after the reaction. Furthermore, the non-covalent interactions may inhibit the desired conjugation reaction by orienting TG2 in a way that prevents access to the active site. I also established the kinetics of redox regulation of transamidase activity, showing that TG2 is rapidly inactivated in the absence of reducing agent. The strict requirement for reducing conditions limits application, particularly *in* vivo. However, I showed that a C230S mutant is inactivated more slowly in non-reducing conditions than wild-type. Further work is required to determine if the C230S mutant may be capable of performing meaningful levels of conjugation in the extracellular space before becoming inactivated. I also introduced a TG2-reactive peptide tag to the model substrate mClover and found that TG2 had a propensity to conjugate the tag to its own reactive lysine residues. Self-conjugation of TG2 made directing assembly between other protein substrates challenging. I fused TG2-reactive tags to TG2 itself to try to harness self-

conjugation for covalently anchoring TG2 to endogenous protein substrates. However, tagged-TG2 favoured self-conjugation over conjugating itself to other proteins. After exploring an enzymatic approach and an anchoring approach to TG2-based protein coupling, I concluded that the high levels of self-reactivity limit the application of TG2 for directing highly specific conjugation. However, after observing robust non-covalent binding to the cell surface and fibronectin, I decided to develop TG2 as a tool for non-covalent decoration of the extracellular space.

The original goal of this thesis was to develop a tool for covalent conjugation. The presence of reactive lysine residues on TG2 itself interfered with transamidation of lysine residues on target proteins. Mass spectrometry analysis of TG2 self-crosslinked multimers has identified twelve lysine residues as sites of transamidation (Stamnaes et al., 2015). Three of these lysine residues are present in $TG2_{465}$: K380, K429 and K444. To prevent TG2 conjugating itself to the Qtag, these lysine residues could be substituted to arginine. Substitution to arginine would conserve charge whilst abolishing reactivity. Such an approach has been used to prevent E3 ligases from recognising ubiquitin as a substrate (Lim et al., 2005). Alternatively, TG2 could be engineered to moderately increase the specificity of TG2 for lysine residues in flexible regions. The structure of the TG2 active site is well studied (Ha et al., 2018). This information can be used, along with molecular dynamics simulations, to identify a small number of key residues that influence substrate binding. Rational introduction of bulky residues to the active site tunnel could be employed to increase the requirement for substrate flexibility by partially 'blocking' the active site tunnel. Making it harder for substrates to access the active site may increase the substrate specificity such that the lysine residues of $TG2_{465}$ are no longer compatible acyl-acceptors. This active site 'blocking' approach has been used to increase the specificity of fatty acid dehydratase FabA (Bao et al., 2016). As the active site of TG2 is located deep within the core of the protein and shielded from water, minimal changes to the active site

amino acid composition are unlikely to increase the immunogenicity of TG2. Increasing the specificity for acyl donors would inevitably lead to a decrease in the number of reactive residues on protein scaffolds, resulting in less scaffold conjugation. However, if TG2-to-Qtag conjugation was abolished then it may be a worthwhile trade off. Another approach to improving the specificity of TG2 is to engineer the active site to accept glutamate ester as a substrate rather than glutamine. Engineering in this manner would prevent TG2 from conjugating its own glutamine residues or any on its target substrate that are not the intended conjugation site. A glutamate ester could then be incorporated into the Qtag by genetic code expansion. For example, L-glutamic acid γ-benzyl ester has previously been incorporated into proteins by genetic code expansion (Yang et al., 2021).

Another approach for preventing self-conjugation is post-translational modification of TG2 lysine residues. Incubation of TG2 *in vitro* with an NHS ester can be used to form stable amide bonds with the lysine sidechain and render the residue unreactive (Hermanson, 2013). However, there are often unwanted side reactions with other amino acids resulting in sample heterogeneity (Mädler et al., 2009). Another consideration is that lysine modification leads to charge neutralization which may alter structure, stability, and activity of TG2.

Even if TG2 is effectively modified to prevent self-conjugation, there may be issues with multimerization of a Q-tagged cargo protein if the cargo protein contains reactive lysine residues. Genetic engineering or post-translational modification approaches could also be taken to render cargo proteins compatible with the system, but solutions which require modification of the cargo protein may become too inconvenient for the system to be broadly useful.

Another challenge for a TG2-based conjugation tool is the high affinity for scaffold proteins and cell surface. Mutating the fibronectin binding domain helped to decrease affinity for fibronectin 100-fold and slightly reduced the binding to basement membrane extract and cell surface. A similar approach could be taken to decrease affinity of TG2 for heparan sulfate

proteoglycans. One high-affinity heparan sulfate binding site is thought to be in the region of amino acids 202-222 (Wang et al., 2012) (Teesalu et al., 2012). A TG2 variant with the mutations K205A and R209A showed decreased binding to a heparin (Wang et al.,2012). Interactions with β1, β3, α5 and α9 integrins have been demonstrated, but exact binding sites on TG2 are not known (Akimov et al., 2000) (Verderio et al., 2003) (Chen et al., 2010). It has been suggested that these sites are at the termini, so removal of the N-terminal domain of TG2 may decrease cell-surface affinity.

Beyond TG2, there are other human proteins which may be used for protein-protein conjugation to endogenous proteins. In the LACE system described in Chapter 1, Ubc9 has been shown to conjugate thioesters to lysine residues which differ slightly in context from the IKQE substrate motif (Hofmann et al., 2020). This finding suggests that substrate specificity is quite low. If necessary, Ubc9 could be engineered to decrease the specificity requirements for the donor lysine further, to allow coupling to a wide range of endogenous lysine-bearing proteins. However, the requirement for a C-terminal thioester for Ubc9 still limits application.

Alternatively, a non-human protein-protein coupling system could be humanized to decrease its immunogenicity. Such strategies are commonly employed for antibodies and therapeutic enzymes (Waldmann, 2019) (Nyborg et al., 2016) (Fonseca et al., 2021). There are now many computational tools available for predicting peptide binding to MHC class I and class II (Khetan et al., 2022). Sortases are good candidates for this approach because they are well characterised in terms of reaction selectivity (Antos et al., 2016) and crystal structures (Zong et al., 2004).

Aside from repurposing existing technologies, future efforts to develop a human-derived coupling tool should also look toward human proteins which have not been developed for protein conjugation. TG2 is part of a family of 9 known human transglutaminases, most of which can be expressed recombinantly in *E. coli* or Sf9 cells (Hadjivassiliou et al., 2008) (Kim

et al., 2021) (Pietroni et al., 2008). However, these proteins are not as well studied as TG2 and FXIII, so the substrate scope is unknown. It would be valuable to test if any of these transglutaminases has stricter glutamine-donor requirements than TG2 and FXIII.

In Chapter 4, I established a method for modular post-translational conjugation of TG2 and cargo proteins using SnoopLigase2. While SnoopLigase has found successful applications (Andersson et al., 2019) (Chen et al., 2022), the reaction is slow (~24 h to completion) and limited to highly specific reaction conditions which restricts broader adoption. Mutations were introduced into SnoopLigase by Dr. Keeble to generate a second-generation system with improved performance. I tested the second-generation system and found a considerable improvement in reactivity due at least partially to a drastic improvement in solubility of the ligase element in common buffers. I then quantified the activity landscape of SnoopLigase2 in a range of reaction conditions, varying pH, temperature, buffer, and various additives. SnoopLigase2 demonstrated efficient ligation across a wider range of pH values and temperatures than anything currently published amongst other protein-protein conjugation technologies. I also demonstrated that SnoopLigase2 can direct rapid modular coupling at either termini and at an internal site. Such simple and versatile modular coupling has the potential to significantly expand the range of application of biotechnologies to which it is applied, demonstrated in Chapter 5 by assembly of TG2-TGFα constructs. The application of the SnoopLigase2 system to TG2 allowed construction of TG2:cargo constructs that are not possible by genetic fusion. When trying to dissociate SnoopLigase2 from the reaction product, I found that it remains tightly bound in conditions that efficiently dissociate the original SnoopLigase. The stable association of SnoopLigase2 with the SnoopTag2:DogTag2 product may limit certain applications, but there are also interesting applications for post-translational generation of stable heterotrimers. SnoopLigase2 may be used for generating tri-specifics, or

the ligase element could be labelled with a probe for capture or tracking of conjugates without the need to label the tagged proteins.

In Chapter 5, I showed that a TG2 variant can decorate the cell surface with minimal dissociation after mild washes. I then used the SnoopLigase2 system to post-translationally attach TG2 to TGFα, which cannot be genetically fused to TG2, to generate a TG2-TGFα construct. DU145 cells treated with TG2-TGFα showed enhanced TGFα activity compared with TGFα alone. TG2-mediated anchoring to the cell surface enhanced duration of TGFα activity when cells were subjected to washes. Given the role of cell-surface TG2 in a range of physiological processes such as EMT, a thorough analysis of any effects of exogenous surface-bound TG2 on downstream cell signalling pathways would be required to develop this system further. Further work is required to determine if TG2 can robustly bind to other cell types, and to identify the proteins to which TG2 is interacting to establish if TG2 is a broadly useful tool for stable decoration of the cell surface. TG2 affinity for the cell surface may vary between cell types depending on the expression levels of TG2-interacting proteins such as syndecan-4. TG2 affinity may also vary depending on the level of endogenous TG2 already localised to the cell surface. Of particular interest is the ability of TG2 to bind to primary fibroblasts for tissue repair applications.

The persistence of the bacterial SnoopLigase2 components in the final product may limit the potential of this TG2-TGFα construct as a therapeutic. To leave behind fewer microbial elements, SrtA could be used by adding a C-terminal LPXTG motif to TGFα and an N-terminal oligoglycine tag to TG2 (Popp and Ploegh, 2011). Attaching TGFα to the C-terminus of TG2 is preferable to avoid disrupting interactions between the N-terminus of TG2 and fibronectin. Split intein systems offer an alternative to SrtA, but the optimisation to find a suitable split intein pair which produces soluble components and ligates efficiently may be more laborious (Pirzer et al., 2018) (Sekar et al., 2022).

The demonstration of enhanced activity of TGFα by TG2-mediated anchoring may have relevance towards exploring therapeutic applications of TGFα, such as repairing neural tissue after stroke (Alipanahzadeh et al., 2014). Beyond TGFα, TG2-mediated anchoring may have beneficial effects on the activity of other therapeutic proteins. Over 380 protein therapeutics have been FDA-approved and applied to treat a wide range of diseases (Usmani et al., 2017). Such therapeutics include growth factors, interleukins, interferons, lipases, proteases and DNAses. For example, fibroblast growth factor could be anchored to the basement membrane for treatment of diabetic foot ulcers (Liu et al., 2021). Bone morphogenetic proteins could be anchored to tissue scaffolds used in bone grafts (Gillman and Jayasuriya, 2021). DNAse I could be anchored within the lung epithelium to treat cystic fibrosis (Fonseca et al., 2020). Interferons and interleukins could be anchored to the tumour microenvironment for treatment of cancers (Propper and Balkwill, 2022).

Which, if any, of the above applications are suitable for TG2 depends on how well TG2 can perform in different extracellular environments. For example, the profile of proteins present at the cell surface is likely to differ considerably between cell types and tissues, which may lead to differences in the ability of TG2 to bind. Which tissues and diseases are the most suitable for TG2-mediated decoration depends on how well TG2 can bind to the relevant cell types and ECM components. Work needs to be done to test TG2 affinity for a range of cell types and tissues to better identify the most promising applications. Given that TG2 has high affinity for fibronectin and some integrins, it is likely that TG2 will bind to most mammalian cells.

On the other hand, if there are differences in the ability of TG2 to anchor in different tissues, it may be possible to utilise differences between diseased and healthy tissues for targeted decoration of pathogenic tissues with TG2. For example, Tenascin-C is an ECM component which is absent from healthy tissues and upregulated during inflammation

(Udalova et al., 2011). Tenascin-C has previously been targeted for delivery of insulin-like growth factor 1 (IGF-1) to inflamed cartilage for treatment of osteoarthritis (Loffredo et al., 2014). If TG2 demonstrates significantly improved binding for a particular ECM marker of inflammation, such as Tenascin-C, then it may be used to bind preferentially to inflamed tissue for delivery of anti-inflammatory therapy. Furthermore, isoforms of some collagens are upregulated in breast carcinoma (Giussani et al., 2018) and may offer a target for delivery of pro-inflammatory therapeutics to tumours if TG2 is found to bind robustly.

Thought must be given to the roles of TG2 in disease when considering therapeutic application of exogenous TG2. Increased expression and activity of TG2 has long been associated with fibrotic diseases (Johnson et al., 2003) (El Nahas et al., 2004), and inhibition of TG2 transamidation has been shown to attenuate cardiac fibrosis and lung fibrosis in animal models (Wang et al., 2018) (Philp et al., 2018). Furthermore, addition of exogenous TG2 to a cell model of angiogenesis resulted in dose-dependent inhibition of angiogenesis (Jones et al., 2006). Transamidase activity of TG2 at the cell surface has been associated with increased migration of cancer cells (Wang and Griffin, 2013). Finally, there may be a role for TG2 transamidation in activating the immune response in coeliac patients. TG2 has been shown to form covalent complexes with gluten peptides *in vitro* (Stamnaes et al., 2015). These TG2-gluten covalent complexes are effective antigens for B cell-mediated activation of gluten-specific T cells (Lindstad et al., 2021). TG2-specific plasma cells are substantially expanded in coeliac disease patients (Di Niro et al., 2012). The contribution of TG2 to disease pathologies seems to be at least partially through transamidase activity, which could be easily avoided in application of recombinant TG2 by using a transamidase-deficient variant. Transamidase-independent roles in disease may be harder to avoid and should be considered when weighing up the benefits and drawbacks of any therapy.

Another consideration for the application of TG2 is that the extracellular space contains a wide range of proteases, including at least 23 matrix metalloproteinases (Li et al., 2020). Given that endogenous TG2 localises at the cell surface and ECM, it is likely that TG2 is relatively resistant to digestion by extracellular proteases. However, the composition of proteases differs between tissues and disease states (López-Otín and Matrisian, 2007) so it remains to be seen if TG2 is vulnerable to digestion by proteases in certain contexts which may limit its use.

Another possible application for TG2:cargo therapeutics is in functionalised tissue scaffolds. Both natural and synthetic tissue scaffolds can be decorated with growth factors (Zhang et al., 2021) (Jeon et al., 2018). There is interest in designing scaffolds which release growth factors gradually in a way which mimics natural release of growth factors from the ECM (Holl et al., 2021). I showed in Chapter 3 that TG2 only bound weakly to a basement membrane extract, but further work is needed to determine how well TG2 binds to decellularized ECM scaffolds. It would also be interesting to study the release profile of TG2 from scaffolds over a period of days or weeks to see if TG2 could be useful for release of therapeutics from tissue scaffolds. In this thesis I have used a modified TG2 to enhance the activity of a growth factor, and there are a broad range of exciting potential applications for TG2 in the area of therapeutics. Hopefully this work will encourage efforts to characterise the behaviour of exogenous TG2 *in vivo* in a range of tissues to elucidate the most appropriate applications.

Beyond therapeutic applications, TG2-mediated anchoring could be used to assemble photonic biomaterials. Reflectins are a family of proteins found in cephalopods which are expressed in iridophores in the skin to give the skin an iridescent appearance and contribute to adaptive camouflage (Kramer, Crookes-Goodson and Naik, 2007). In response to cell signalling, reflectin aggregates into large multimers with a very high refractive index (Izumi et

al., 2009). Layers of reflectin aggregates have previously been used to produce an iridescent material (Qin et al., 2012). TG2 could be used to anchor reflectin monomers in a protein scaffold, which may be layered with the thickness and number of layers determining the optical effects (DeMartini, Krogstad and Morse, 2013).

In summary, the original aim to develop a tool for protein-protein conjugation was not completely fulfilled. It was established that reactive lysine residues on TG2 pose a significant challenge for utilising TG2 as a conjugating system. I have suggested here several strategies to address this issue. I have also highlighted alternative strategies to develop a similar tool based on other enzymes. I did achieve robust non-covalent decoration of protein scaffolds with TG2x, and demonstrated in a proof-of-concept experiment that TG2-mediated decoration of the cell surface enhances activity of a growth factor. In order to elucidate the potential applications of this TG2-based tool, further work is required to identify the preferred binding partners of TG2x in the ECM and at the cell surface.

References

Abrahmsen, L., Tom, J., Burnier, J., Butcher, K. A., Kossiakoff, A., & Wells, J. A. (1991). Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry*, *30*(17), 4151–4159.

Achyuthan, K. E., & Greenberg, C. S. (1987). Identification of a guanosine triphosphate-binding site on guinea pig liver transglutaminase. Role of GTP and calcium ions in modulating activity. *Journal of Biological Chemistry*, *262*(4), 1901–1906.

Adamiak, K., & Sionkowska, A. (2020). Current methods of collagen cross-linking: Review. *International Journal of Biological Macromolecules*, *161*, 550–560.

Addi, C., Murschel, F., & De Crescenzo, G. (2017). Design and Use of Chimeric Proteins Containing a Collagen-Binding Domain for Wound Healing and Bone Regeneration. *Tissue Engineering. Part B, Reviews*, *23*(2), 163–182.

Aeschlimann, D., Kaupp, O., & Paulsson, M. (1995). Transglutaminase-catalyzed matrix cross-linking in differentiating cartilage: Identification of osteonectin as a major glutaminyl substrate. *Journal of Cell Biology*, *129*(3), 881–892.

Akakura, S., Singh, S., Spataro, M., Akakura, R., Kim, J.-I., Albert, M. L., & Birge, R. B. (2004). The opsonin MFG-E8 is a ligand for the αvβ5 integrin and triggers DOCK180-dependent Rac1 activation for the phagocytosis of apoptotic cells. *Experimental Cell Research*, *292*(2), 403–416.

Akimov, S. S., Krylov, D., Fleischman, L. F., & Belkin, A. M. (2000). Tissue transglutaminase is an integrin-binding adhesion coreceptor for fibronectin. *The Journal of Cell Biology*, *148*(4), 825–838.

Akimov, S. S., & Belkin, A. M. (2001). Cell surface tissue transglutaminase is involved in adhesion and migration of monocytic cells on fibronectin. *Blood*, *98*(5), 1567–1576.

Al-Amer, O. M. (2022). The role of thrombin in haemostasis. *Blood Coagulation & Fibrinolysis*, *33*(3), 145.

Alipanahzadeh, H., Soleimani, M., Soleimani Asl, S., Pourheydar, B., Nikkhah, A., & Mehdizadeh, M. (2014). Transforming Growth Factor-α Improves Memory Impairment and Neurogenesis Following Ischemia Reperfusion. *Cell Journal*, *16*(3), 315–324.

Andersson, A.-M. C., Buldun, C. M., Pattinson, D. J., Draper, S. J., & Howarth, M. (2019). SnoopLigase peptide-peptide conjugation enables modular vaccine assembly. *Scientific Reports*, *9*(1), Article 1.

Andrades, J. A., Wu, L. T., Hall, F. L., Nimni, M. E., & Becerra, J. (2001). Engineering, expression, and renaturation of a collagen-targeted human bFGF fusion protein. *Growth Factors (Chur, Switzerland)*, *18*(4), 261–275.

Antos, J. M., Truttmann, M. C., & Ploegh, H. L. (2016). Recent advances in sortase-catalyzed ligation methodology. *Current Opinion in Structural Biology*, *38*, 111–118.

Bagur, R., & Hajnóczky, G. (2017). Intracellular Ca2+ sensing: Role in calcium homeostasis and signaling. *Molecular Cell*, *66*(6), 780–788.

Banerjee, A., & Howarth, M. (2018). Nanoteamwork: Covalent protein assembly beyond duets towards protein ensembles and orchestras. *Current Opinion in Biotechnology*, *51*, 16–23.

Bao, L., Li, J.-J., Jia, C., Li, M., & Lu, X. (2016). Structure-oriented substrate specificity engineering of aldehyde-deformylating oxygenase towards aldehydes carbon chain length. *Biotechnology for Biofuels*, *9*(1), 185.

Barber, C. J. S., Pujara, P. T., Reed, D. W., Chiwocha, S., Zhang, H., & Covello, P. S. (2013). The Two-step Biosynthesis of Cyclic Peptides from Linear Precursors in a Member of the Plant Family Caryophyllaceae Involves Cyclization by a Serine Protease-like Enzyme *. *Journal of Biological Chemistry*, *288*(18), 12500–12510.

Battistini, L., Bugatti, K., Sartori, A., Curti, C., & Zanardi, F. (2021). RGD Peptide-Drug Conjugates as Effective Dual Targeting Platforms: Recent Advances. *European Journal of Organic Chemistry*, *2021*(17), 2506–2528.

Bean, M. F., & Carr, S. A. (1992). Characterization of disulfide bond position in proteins and sequence analysis of cystine-bridged peptides by tandem mass spectrometry. *Analytical Biochemistry*, *201*(2), 216–226.

Beazley, K. E., Deasey, S., Lima, F., & Nurminskaya, M. V. (2012). Transglutaminase 2-mediated activation of β-catenin signaling has a critical role in warfarin-induced vascular calcification. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *32*(1), 123–130.

Bechtel, T. J., & Weerapana, E. (2017). From structure to redox: The diverse functional roles of disulfides and implications in disease. *PROTEOMICS*, *17*(6), 1600391.

Belkin, A. M., Tsurupa, G., Zemskov, E., Veklich, Y., Weisel, J. W., & Medved, L. (2005). Transglutaminase-mediated oligomerization of the fibrin(ogen) alphaC domains promotes integrin-dependent cell adhesion and signaling. *Blood*, *105*(9), 3561–3568.

Belkin, A. M. (2011). Extracellular TG2: Emerging functions and regulation. *The FEBS Journal*, *278*(24), 4704–4716.

Berillo, D., Yeskendir, A., Zharkinbekov, Z., Raziyeva, K., & Saparov, A. (2021). Peptide-Based Drug Delivery Systems. *Medicina*, *57*(11), 1209.

Bernassola, F., Federici, M., Corazzari, M., Terrinoni, A., Hribal, M. L., De Laurenzi, V., Ranalli, M., Massa, O., Sesti, G., Mclean, W. h. I., Citro, G., Barbetti, F., & Melino, G. (2002). Role of transglutaminase 2 in glucose tolerance: Knockout mice studies and a putative mutation in a MODY patient. *The FASEB Journal*, *16*(11), 1371–1378.

Betting, D. J., Kafi, K., Abdollahi-Fard, A., Hurvitz, S. A., & Timmerman, J. M. (2008). Sulfhydryl-Based Tumor Antigen-Carrier Protein Conjugates Stimulate Superior Antitumor Immunity against B Cell Lymphomas1. *The Journal of Immunology*, *181*(6), 4131–4140.

Bhagawati, M., Terhorst, T. M. E., Füsser, F., Hoffmann, S., Pasch, T., Pietrokovski, S., & Mootz, H. D. (2019). A mesophilic cysteine-less split intein for protein trans-splicing applications under oxidizing conditions. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(44), 22164–22172.

Bhagawati, M., Hoffmann, S., Höffgen, K. S., Piehler, J., Busch, K. B., & Mootz, H. D. (2020). In Cellulo Protein Semi-Synthesis from Endogenous and Exogenous Fragments Using the Ultra-Fast Split Gp41-1 Intein. *Angewandte Chemie (International Ed. in English)*, *59*(47), 21007–21015.

Bosman, F. T., & Stamenkovic, I. (2003). Functional structure and composition of the extracellular matrix. *The Journal of Pathology*, *200*(4), 423–428.

Boyles, J. S., Atwell, S., Druzina, Z., Heuer, J. G., & Witcher, D. R. (2016). Structural basis of selectivity and neutralizing activity of a TGFα/epiregulin specific antibody. *Protein Science*, *25*(11), 2028–2036.

Brooks, D. J., Fresco, J. R., Lesk, A. M., & Singh, M. (2002). Evolution of Amino Acid Frequencies in Proteins Over Deep Time: Inferred Order of Introduction of Amino Acids into the Genetic Code. *Molecular Biology and Evolution*, *19*(10), 1645–1655.

Broun, G. B. (1976). [20] Chemically aggregated enzymes. In *Methods in Enzymology* (Vol. 44, pp. 263–280). Academic Press.

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

Buldun, C. M., Jean, J. X., Bedford, M. R., & Howarth, M. (2018). SnoopLigase Catalyzes Peptide–Peptide Locking and Enables Solid-Phase Conjugate Isolation. *Journal of the American Chemical Society*, *140*(8), 3008–3018.

Buldun, C. M., Khairil Anuar, I. N. A., & Howarth, M. (2021). SnoopLigase-Mediated Peptide-Peptide Conjugation and Purification. *Methods in Molecular Biology (Clifton, N.J.)*, *2208*, 13–31.

Burhan, I., Furini, G., Lortat-Jacob, H., Atobatele, A. G., Scarpellini, A., Schroeder, N., Atkinson, J., Maamra, M., Nutter, F. H., Watson, P., Vinciguerra, M., Johnson, T. S., & Verderio, E. A. M. (2016). Interplay between transglutaminases and heparan sulphate in progressive renal scarring. *Scientific Reports*, *6*(1), 31343.

Cambria, E., Renggli, K., Ahrens, C. C., Cook, C. D., Kroll, C., Krueger, A. T., Imperiali, B., & Griffith, L. G. (2015). Covalent Modification of Synthetic Hydrogels with Bioactive Proteins via Sortase-Mediated Ligation. *Biomacromolecules*, *16*(8), 2316–2326.

Cardoso, I., Stamnaes, J., Andersen, J. T., Melino, G., Iversen, R., & Sollid, L. M. (2015). Transglutaminase 2 interactions with extracellular matrix proteins as probed with celiac disease autoantibodies. *The FEBS Journal*, *282*(11), 2063–2075.

Cardoso, I., Østerlund, E. C., Stamnaes, J., Iversen, R., Andersen, J. T., Jørgensen, T. J. D., & Sollid, L. M. (2017). Dissecting the interaction between transglutaminase 2 and fibronectin. *Amino Acids*, *49*(3), 489–500.

Carvajal-Vallejos, P., Pallissé, R., Mootz, H. D., & Schmidt, S. R. (2012). Unprecedented rates and efficiencies revealed for new natural split inteins from metagenomic sources. *The Journal of Biological Chemistry*, *287*(34), 28686–28696.

Ceresino, E. B., de Melo, R. R., Kuktaite, R., Hedenqvist, M. S., Zucchi, T. D., Johansson, E., & Sato, H. H. (2018). Transglutaminase from newly isolated *Streptomyces* sp. CBMAI 1617: Production optimization, characterization and evaluation in wheat protein and dough systems. *Food Chemistry*, *241*, 403–410.

Chalker, J. M., Gunnoo, S. B., Boutureira, O., Gerstberger, S. C., Fernández-González, M., Bernardes, G. J. L., Griffin, L., Hailu, H., Schofield, C. J., & Davis, B. G. (2011). Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chemical Science*, *2*(9), 1666–1676.

Chen, I., Dorr, B. M., & Liu, D. R. (2011). A general strategy for the evolution of bond-forming enzymes using yeast display. *Proceedings of the National Academy of Sciences*, *108*(28), 11399–11404.

Chen, S.-H., Lin, C.-Y., Lee, L.-T., Chang, G.-D., Lee, P.-P., Hung, C.-C., Kao, W.-T., Tsai, P.-H., Schally, A. V., Hwang, J.-J., & Lee, M.-T. (2010). Up-regulation of fibronectin and tissue transglutaminase promotes cell invasion involving increased association with integrin and MMP expression in A431 cells. *Anticancer Research*, *30*(10), 4177–4186.

Chen, J., Huang, C., Zhao, W., Ren, J., Ji, F., & Jia, L. (2022). SnoopLigase Enables Highly Efficient Generation of C–C-Linked Bispecific Nanobodies Targeting TNF-α and IL-17A. *Bioconjugate Chemistry*, *33*(8), 1446–1455.

Chen, J., Wang, J., Chen, D., Yang, J., Yang, C., Zhang, Y., Zhang, H., & Dou, J. (2013). Evaluation of characteristics of CD44+CD117+ ovarian cancer stem cells in three dimensional basement membrane extract scaffold versus two dimensional monocultures. *BMC Cell Biology*, *14*(1), 7.

Chmura, A. J., Orton, M. S., & Meares, C. F. (2001). Antibodies with infinite affinity. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(15), 8480–8484.

Chu, F., Thornton, D. T., & Nguyen, H. T. (2018). Chemical cross-linking in the structural analysis of protein assemblies. *Methods*, *144*, 53–63.

Clyne, D. H., Norris, S. H., Modesto, R. R., Pesce, A. J., & Pollak, V. E. (1973). Antibody enzyme conjugates the preparation of intermolecular conjugates of horseradish peroxidase and antibody and their use in immunohistology of renal cortex. *Journal of Histochemistry & Cytochemistry*, *21*(3), 233–240.

Cooling, M. T., Rouilly, V., Misirli, G., Lawson, J., Yu, T., Hallinan, J., & Wipat, A. (2010). Standard virtual biological parts: A repository of modular modeling components for synthetic biology. *Bioinformatics (Oxford, England)*, *26*(7), 925–931.

Csosz, E., Bagossi, P., Nagy, Z., Dosztanyi, Z., Simon, I., & Fesus, L. (2008). Substrate preference of transglutaminase 2 revealed by logistic regression analysis and intrinsic disorder examination. *Journal of Molecular Biology*, *383*(2), 390–402.

Csosz, E., Meskó, B., & Fésüs, L. (2009). Transdab wiki: The interactive transglutaminase substrate database on web 2.0 surface. *Amino Acids*, *36*(4), 615–617.

Cui, L., Du, G., Zhang, D., & Chen, J. (2008). Thermal stability and conformational changes of transglutaminase from a newly isolated *Streptomyces hygroscopicus*. *Bioresource Technology*, *99*(9), 3794–3800.

Curnis, F., Gasparri, A., Sacchi, A., Longhi, R., & Corti, A. (2004). Coupling Tumor Necrosis Factor-α with αV Integrin Ligands Improves Its Antineoplastic Activity. *Cancer Research*, *64*(2), 565–571.
Dai, Y., Dudek, N. L., Li,
Q., & Muma, N. A. (2011). Phospholipase C, Ca2+, and calmodulin signaling are required for 5-HT2A receptor-mediated transamidation of Rac1 by transglutaminase. *Psychopharmacology*, *213*(2–3), 403–412.

Dassie, J. P., Liu, X.-Y., Thomas, G. S., Whitaker, R. M., Thiel, K. W., Stockdale, K. R., Meyerholz, D. K., McCaffrey, A. P., McNamara, J. O., & Giangrande, P. H. (2009). Systemic administration of optimized aptamer-siRNA chimeras promotes regression of PSMA-expressing tumors. *Nature Biotechnology*, *27*(9), 839–849.

Davies, G. E., & Stark, G. R. (1970). Use of Dimethyl Suberimidate, a Cross-Linking Reagent, in Studying the Subunit Structure of Oligomeric Proteins*. *Proceedings of the National Academy of Sciences*, *66*(3), 651–656.

Deasey, S., Nurminsky, D., Shanmugasundaram, S., Lima, F., & Nurminskaya, M. (2013). Transglutaminase 2 as a novel activator of LRP6/β-catenin signaling. *Cellular Signalling*, *25*(12), 2646–2651.

DeMartini, D. G., Krogstad, D. V., & Morse, D. E. (2013). Membrane invaginations facilitate reversible water flux driving tunable iridescence in a dynamic biophotonic system. *Proceedings of the National Academy of Sciences*, *110*(7), 2552–2556.

Deng, Y., Wu, T., Wang, M., Shi, S., Yuan, G., Li, X., Chong, H., Wu, B., & Zheng, P. (2019). Enzymatic biosynthesis and immobilization of polyprotein verified at the single-molecule level. *Nature Communications*, *10*(1), 2775.

Dhar, T., & D. Mootz, H. (2011). Modification of transmembrane and GPI-anchored proteins on living cells by efficient protein trans -splicing using the Npu DnaE intein. *Chemical Communications*, *47*(11), 3063–3065.

Di Niro, R., Mesin, L., Zheng, N.-Y., Stamnaes, J., Morrissey, M., Lee, J.-H., Huang, M., Iversen, R., du Pré, M. F., Qiao, S.-W., Lundin, K. E. A., Wilson, P. C., & Sollid, L. M.

(2012). High abundance of plasma cells secreting transglutaminase 2–specific IgA autoantibodies with limited somatic hypermutation in celiac disease intestinal lesions. *Nature Medicine*, *18*(3), Article 3.

Dickerson, E. B., Akhtar, N., Steinberg, H., Wang, Z.-Y., Lindstrom, M. J., Padilla, M. L., Auerbach, R., & Helfand, S. C. (2004). Enhancement of the antiangiogenic activity of interleukin-12 by peptide targeted delivery of the cytokine to alphavbeta3 integrin. *Molecular Cancer Research: MCR*, *2*(12), 663–673.

Ding, W., Pang, X., Ding, Z., Tsang, D. C. W., Jiang, Z., & Shi, B. (2020). Constructing a robust chrome-free leather tanned by biomass-derived polyaldehyde via crosslinking with chitosan derivatives. *Journal of Hazardous Materials*, *396*, 122771.

Dosio, F., Brusa, P., & Cattel, L. (2011). Immunotoxins and Anticancer Drug Conjugate Assemblies: The Role of the Linkage between Components. *Toxins*, *3*(7), Article 7.

Duarte, L., Matte, C. R., Bizarro, C. V., & Ayub, M. A. Z. (2020). Transglutaminases: Part I—origins, sources, and biotechnological characteristics. *World Journal of Microbiology and Biotechnology*, *36*(1), 15.

Duran, R., Junqua, M., Schmitter, J. M., Gancet, C., & Goulas, P. (1998). Purification, characterisation, and gene cloning of transglutaminase from *Streptoverticillium cinnamoneum* CBS 683.68. *Biochimie*, *80*(4), 313–319.

Dziuba, D., Hoffmann, J.-E., Hentze, M. W., & Schultz, C. (2020). A Genetically Encoded Diazirine Analogue for RNA–Protein Photo-crosslinking. *ChemBioChem*, *21*(1–2), 88–93.

Eckert, R. L., Kaartinen, M. T., Nurminskaya, M., Belkin, A. M., Colak, G., Johnson, G. V. W., & Mehta, K. (2014). Transglutaminase Regulation of Cell Function. *Physiological Reviews*, *94*(2), 383–417.

El Nahas, A. M., Abo-Zenah, H., Skill, N. J., Bex, S., Wild, G., Griffin, M., & Johnson, T. S. (2004). Elevated epsilon-(gamma-glutamyl)lysine in human diabetic nephropathy results from increased expression and cellular release of tissue transglutaminase. *Nephron. Clinical Practice*, *97*(3), c108-117.

Emanuelsson, C. S., Boros, S., Hjernoe, K., Boelens, W. C., & Hojrup, P. (2005). Screening for Transglutaminase-Catalyzed Modifications by Peptide Mass Finger Printing Using Multipoint Recalibration on Recognized Peaks for High Mass Accuracy. *Journal of Biomolecular Techniques : JBT*, *16*(3), 197–208.

Esposito, C., & Caputo, I. (2005). Mammalian transglutaminases. *The FEBS Journal*, *272*(3), 615–631.

Evans, T. C., Martin, D., Kolly, R., Panne, D., Sun, L., Ghosh, I., Chen, L., Benner, J., Liu, X.-Q., & Xu, M.-Q. (2000). Protein trans-Splicing and Cyclization by a Naturally Split Intein from the dnaE Gene of Synechocystis Species PCC6803 *. *Journal of Biological Chemistry*, *275*(13), 9091–9094.

Falcone, P., Serafini-Fracassini, D., & Del Duca, S. (1993). Comparative Studies of Transglutaminase Activity and Substrates in Different Organs of *Helianthus tuberosus*. *Journal of Plant Physiology*, *142*(3), 265–273.

Fauser, J., Savitskiy, S., Fottner, M., Trauschke, V., & Gulen, B. (2020). Sortase-Mediated Quantifiable Enzyme Immobilization on Magnetic Nanoparticles. *Bioconjugate Chemistry*, *31*(8), 1883–1892.

Federico, S., Pierce, B. F., Piluso, S., Wischke, C., Lendlein, A., & Neffe, A. T. (2015). Design of Decorin-Based Peptides That Bind to Collagen I and their Potential as Adhesion Moieties in Biomaterials. *Angewandte Chemie (International Ed. in English)*, *54*(37), 10980–10984.

Folk, J. E., & Cole, P. W. (1966). Identification of a functional cysteine essential for the activity of guinea pig liver transglutaminase. *The Journal of Biological Chemistry*, *241*(13), 3238–3240.

Fonseca, C., Bicker, J., Alves, G., Falcão, A., & Fortuna, A. (2020). Cystic fibrosis: Physiopathology and the latest pharmacological treatments. *Pharmacological Research*, *162*, 105267.

Fonseca, M. H. G., Fiúza, T. da S., Morais, S. B. de, Souza, T. de A. C. B. de, & Trevizani, R. (2021). Circumventing the side effects of L-asparaginase. *Biomedicine & Pharmacotherapy*, *139*, 111616.

Freistroffer, D. V., Kwiatkowski, M., Buckingham, R. H., & Ehrenberg, M. (2000). The accuracy of codon recognition by polypeptide release factors. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(5), 2046–2051.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, *31*(13), 3784–3788.

Gaudry, C. A., Verderio, E., Jones, R. A., Smith, C., & Griffin, M. (1999). Tissue transglutaminase is an important player at the surface of human endothelial cells: Evidence for its externalization and its colocalization with the beta(1) integrin. *Experimental Cell Research*, *252*(1), 104–113.

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345.

Gilad, Y., Noy, E., Senderowitz, H., Albeck, A., Firer, M. A., & Gellerman, G. (2016). Synthesis, biological studies and molecular dynamics of new anticancer RGD-based peptide conjugates for targeted drug delivery. *Bioorganic & Medicinal Chemistry*, *24*(2), 294–303.

Gillman, C. E., & Jayasuriya, A. C. (2021). FDA-approved bone grafts and bone graft substitute devices in bone regeneration. *Materials Science and Engineering: C*, *130*, 112466.

Giussani, M., Landoni, E., Merlino, G., Turdo, F., Veneroni, S., Paolini, B., Cappelletti, V., Miceli, R., Orlandi, R., Triulzi, T., & Tagliabue, E. (2018). Extracellular matrix proteins as diagnostic markers of breast carcinoma. *Journal of Cellular Physiology*, *233*(8), 6280–6290.

Gramespacher, J. A., Stevens, A. J., Thompson, R. E., & Muir, T. W. (2018). Improved protein splicing using embedded split inteins. *Protein Science: A Publication of the Protein Society*, *27*(3), 614–619.

Grenard, P., Bates, M. K., & Aeschlimann, D. (2001). Evolution of transglutaminase genes: Identification of a transglutaminase gene cluster on human chromosome 15q15. Structure of the gene encoding transglutaminase X and a novel gene family member, transglutaminase Z. *The Journal of Biological Chemistry*, *276*(35), 33066–33078.

Groenen, L. C., Nice, E. C., & Burgess, A. W. (1994). Structure-function relationships for the EGF/TGF-alpha family of mitogens. *Growth Factors (Chur, Switzerland)*, *11*(4), 235–257.

Grootjans, J. J., Groenen, P. J., & de Jong, W. W. (1995). Substrate requirements for transglutaminases. Influence of the amino acid residue preceding the amine donor lysine in a native protein. *The Journal of Biological Chemistry*, *270*(39), 22855–22858.

Gundersen, M. T., Keillor, J. W., & Pelletier, J. N. (2014). Microbial transglutaminase displays broad acyl-acceptor substrate specificity. *Applied Microbiology and Biotechnology*, *98*(1), 219–230.

Gupta, S., Biswas, A., Akhter, M. S., Krettler, C., Reinhart, C., Dodt, J., Reuter, A., Philippou, H., Ivaskevicius, V., & Oldenburg, J. (2016). Revisiting the mechanism of coagulation factor XIII activation and regulation from a structure/functional perspective. *Scientific Reports*, *6*, 30105.

Guyot, N., Zani, M.-L., Maurel, M.-C., Dallet-Choisy, S., & Moreau, T. (2005). Elafin and Its Precursor Trappin-2 Still Inhibit Neutrophil Serine Proteinases when They Are Covalently Bound to Extracellular Matrix Proteins by Tissue Transglutaminase. *Biochemistry*, *44*(47), 15610–15618.

Ha, H. J., Kwon, S., Jeong, E. M., Kim, C. M., Lee, K. B., Kim, I.-G., & Park, H. H. (2018). Structure of natural variant transglutaminase 2 reveals molecular basis of gaining stability and higher activity. *PloS One*, *13*(10), e0204707.

Hadjivassiliou, M., Aeschlimann, P., Strigun, A., Sanders, D. S., Woodroofe, N., & Aeschlimann, D. (2008). Autoantibodies in gluten ataxia recognize a novel neuronal transglutaminase. *Annals of Neurology*, *64*(3), 332–343.

Hafeez, U., Parakh, S., Gan, H. K., & Scott, A. M. (2020). Antibody–Drug Conjugates for Cancer Therapy. *Molecules*, *25*(20), Article 20.

Han, L., Chen, J., Ding, K., Zong, H., Xie, Y., Jiang, H., Zhang, B., Lu, H., Yin, W., Gilly, J., & Zhu, J. (2017). Efficient generation of bispecific IgG antibodies by split intein mediated protein trans-splicing system. *Scientific Reports*, *7*(1), Article 1.

Hang, J., Zemskov, E. A., Lorand, L., & Belkin, A. M. (2005). Identification of a novel recognition sequence for fibronectin within the NH2-terminal beta-sandwich domain of tissue transglutaminase. *The Journal of Biological Chemistry*, *280*(25), 23675–23683.

Harmand, T. J., Pishesha, N., Rehm, F. B. H., Ma, W., Pinney, W. B., Xie, Y. J., & Ploegh, H. L. (2021). Asparaginyl Ligase-Catalyzed One-Step Cell Surface Modification of Red Blood Cells. *ACS Chemical Biology*, *16*(7), 1201–1207.

Harr, M. W., & Distelhorst, C. W. (2010). Apoptosis and autophagy: Decoding calcium signals that mediate life or death. *Cold Spring Harbor Perspectives in Biology*, *2*(10), a005579.

Hemu, X., Chan, N.-Y., Liew, H. T., Hu, S., Zhang, X., Serra, A., Lescar, J., Liu, C.-F., & Tam, J. P. (2023). Substrate-binding glycine residues are major determinants for hydrolase and ligase activity of plant legumains. *The New Phytologist*, *238*(4), 1534–1545.

Hermanson, G. T. (2013). Chapter 3—The Reactions of Bioconjugation. In G. T. Hermanson (Ed.), *Bioconjugate Techniques (Third Edition)* (pp. 229–258). Academic Press.

Ho, M.-L., Leu, S.-Z., Hsieh, J.-F., & Jiang, S.-T. (2000). Technical Approach to Simplify the Purification Method and Characterization of Microbial Transglutaminase Produced from Streptoverticillium ladakanum. *Journal of Food Science*, *65*(1), 76–80.

Hofmann, R., Akimoto, G., Wucherpfennig, T. G., Zeymer, C., & Bode, J. W. (2020). Lysine acylation using conjugating enzymes for site-specific modification and ubiquitination of recombinant proteins. *Nature Chemistry*, *12*(11), Article 11.

Holl, J., Kowalewski, C., Zimek, Z., Fiedor, P., Kaminski, A., Oldak, T., Moniuszko, M., & Eljaszewicz, A. (2021). Chronic Diabetic Wounds and Their Treatment with Skin Substitutes. *Cells*, *10*(3), 655.

How, M. J., Norman, M. D., Finn, J., Chung, W.-S., & Marshall, N. J. (2017). Dynamic Skin Patterns in Cephalopods. *Frontiers in Physiology*, *8*.

Hummerich, R., & Schloss, P. (2010). Serotonin—more than a neurotransmitter: Transglutaminase-mediated serotonylation of C6 glioma cells and fibronectin. *Neurochemistry International*, *57*(1), 67–75.

Hwang, H.-J., Yang, Y. R., Kim, H. Y., Choi, Y., Park, K.-S., Lee, H., Ma, J. S., Yamamoto, M., Kim, J., Chae, Y. C., Choi, J. H., Cocco, L., Berggren, P.-O., Jang, H.-J., & Suh, P.-G. (2019). Phospholipase C-β1 potentiates glucose-stimulated insulin secretion. *The FASEB Journal*, *33*(10), 10668–10679.

Izoré, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., Di Guilmi, A. M., & Dessen, A. (2010). Structural basis of host cell recognition by the pilus adhesin from Streptococcus pneumoniae. *Structure (London, England: 1993)*, *18*(1), 106–115.

Izumi, M., Sweeney, A. M., DeMartini, D., Weaver, J. C., Powers, M. L., Tao, A., Silvas, T. V., Kramer, R. M., Crookes-Goodson, W. J., Mäthger, L. M., Naik, R. R., Hanlon, R. T., &

Morse, D. E. (2009). Changes in reflectin protein phosphorylation are associated with dynamic iridescence in squid. *Journal of The Royal Society Interface*, *7*(44), 549–560.

James, A. M., Haywood, J., Leroux, J., Ignasiak, K., Elliott, A. G., Schmidberger, J. W., Fisher, M. F., Nonis, S. G., Fenske, R., Bond, C. S., & Mylne, J. S. (2019). The macrocyclizing protease butelase 1 remains autocatalytic and reveals the structural basis for ligase activity. *The Plant Journal: For Cell and Molecular Biology*, *98*(6), 988–999.

Jang, G.-Y., Jeon, J.-H., Cho, S.-Y., Shin, D.-M., Kim, C.-W., Jeong, E. M., Bae, H. C., Kim, T. W., Lee, S.-H., Choi, Y., Lee, D.-S., Park, S.-C., & Kim, I.-G. (2010). Transglutaminase 2 suppresses apoptosis by modulating caspase 3 and NF-κB activity in hypoxic tumor cells. *Oncogene*, *29*(3), Article 3.

Jayachandran, B., Parvin, T. N., Alam, M. M., Chanda, K., & Mm, B. (2022). Insights on Chemical Crosslinking Strategies for Proteins. *Molecules*, *27*(23), Article 23.

Jayadev, R., & Sherwood, D. R. (2017). Basement membranes. *Current Biology*, *27*(6), R207–R211.

Jeon, O., Lee, K., & Alsberg, E. (2018). Spatial Micropatterning of Growth Factors in 3D Hydrogels for Location-Specific Regulation of Cellular Behaviors. *Small*, *14*(25), 1800579.

Jeong, E. M., Lee, K. B., Kim, G. E., Kim, C. M., Lee, J.-H., Kim, H.-J., Shin, J.-W., Kwon, M., Park, H. H., & Kim, I.-G. (2020). Competitive Binding of Magnesium to Calcium Binding Sites Reciprocally Regulates Transamidase and GTP Hydrolysis Activity of Transglutaminase 2. *International Journal of Molecular Sciences*, *21*(3), Article 3.

Jin, X., Stamnaes, J., Klöck, C., DiRaimondo, T. R., Sollid, L. M., & Khosla, C. (2011). Activation of Extracellular Transglutaminase 2 by Thioredoxin. *Journal of Biological Chemistry*, *286*(43), 37866–37873.

Johnsen, S. (2011). *The Optics of Life: A Biologist's Guide to Light in Nature*. Princeton University Press.

Johnson, T. S., El-Koraie, A. F., Skill, N. J., Baddour, N. M., El Nahas, A. M., Njloma, M., Adam, A. G., & Griffin, M. (2003). Tissue Transglutaminase and the Progression of Human Renal Scarring. *Journal of the American Society of Nephrology*, *14*(8), 2052.

Jones, R. A., Kotsakis, P., Johnson, T. S., Chau, D. Y. S., Ali, S., Melino, G., & Griffin, M. (2006). Matrix changes induced by transglutaminase 2 lead to inhibition of angiogenesis and tumor growth. *Cell Death and Differentiation*, *13*(9), 1442–1453.

Kanchan, K., Fuxreiter, M., & Fésüs, L. (2015). Physiological, pathological, and structural implications of non-enzymatic protein-protein interactions of the multifunctional human transglutaminase 2. *Cellular and Molecular Life Sciences: CMLS*, *72*(16), 3009–3035.

Kalluri, R., & Weinberg, R. A. (2009). The basics of epithelial-mesenchymal transition. *The Journal of Clinical Investigation*, *119*(6), 1420–1428.

Kamiya, N., Takazawa, T., Tanaka, T., Ueda, H., & Nagamune, T. (2003). Site-specific cross-linking of functional proteins by transglutamination. *Enzyme and Microbial Technology*, *33*(4), 492–496.

Karpovsky, B., Titus, J. A., Stephany, D. A., & Segal, D. M. (1984). Production of target-specific effector cells using hetero-cross-linked aggregates containing anti-target cell and anti-Fc gamma receptor antibodies. *Journal of Experimental Medicine*, *160*(6), 1686–1701.

Kashiwagi, T., Yokoyama, K., Ishikawa, K., Ono, K., Ejima, D., Matsui, H., & Suzuki, E. (2002). Crystal Structure of Microbial Transglutaminase from Streptoverticillium mobaraense. *Journal of Biological Chemistry*, *277*(46), 44252–44260.

Katsumata, K., Ishihara, J., Mansurov, A., Ishihara, A., Raczy, M. M., Yuba, E., & Hubbell, J. A. (2019). Targeting inflammatory sites through collagen affinity enhances the therapeutic efficacy of anti-inflammatory antibodies. *Science Advances*, *5*(11), eaay1971.

Katt, W. P., Antonyak, M. A., & Cerione, R. A. (2018). Opening up about Tissue Transglutaminase: When Conformation Matters More than Enzymatic Activity. *Med One*, *3*(6), e180011.

Keeble, A. H., Turkki, P., Stokes, S., Khairil Anuar, I. N. A., Rahikainen, R., Hytönen, V. P., & Howarth, M. (2019). Approaching infinite affinity through engineering of peptide–protein interaction. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(52), 26523–26533.

Keeble, A. H., Yadav, V. K., Ferla, M. P., Bauer, C. C., Chuntharpursat-Bon, E., Huang, J., Bon, R. S., & Howarth, M. (2022). DogCatcher allows loop-friendly protein-protein ligation. *Cell Chemical Biology*, *29*(2), 339-350.e10.

Keeble, A. H., Wood, D. P., & Howarth, M. (2023). Design and Evolution of Enhanced Peptide–Peptide Ligation for Modular Transglutaminase Assembly. *Bioconjugate Chemistry*, *34*(6), 1019–1036.

Keillor, J. W., Clouthier, C. M., Apperley, K. Y. P., Akbar, A., & Mulani, A. (2014). Acyl transfer mechanisms of tissue transglutaminase. *Bioorganic Chemistry*, *57*, 186–197.

Keresztessy, Z., Csosz, E., Hársfalvi, J., Csomós, K., Gray, J., Lightowlers, R. N., Lakey, J. H., Balajthy, Z., & Fésüs, L. (2006). Phage display selection of efficient glutamine-donor substrate peptides for transglutaminase 2. *Protein Science: A Publication of the Protein Society*, *15*(11), 2466–2480.

Kerr, C., Szmacinski, H., Fisher, M. L., Nance, B., Lakowicz, J. R., Akbar, A., Keillor, J. W., Lok Wong, T., Godoy-Ruiz, R., Toth, E. A., Weber, D. J., & Eckert, R. L. (2017). Transamidase site-targeted agents alter the conformation of the transglutaminase cancer stem cell survival protein to reduce GTP binding activity and cancer stem cell survival. *Oncogene*, *36*(21), Article 21.

Khetan, R., Curtis, R., Deane, C. M., Hadsund, J. T., Kar, U., Krawczyk, K., Kuroda, D., Robinson, S. A., Sormanni, P., Tsumoto, K., Warwicker, J., & Martin, A. C. R. (2022).

Current advances in biopharmaceutical informatics: Guidelines, impact and challenges in the computational developability assessment of antibody therapeutics. *mAbs*, *14*(1), 2020082.

Kieliszek, M., & Misiewicz, A. (2014). Microbial transglutaminase and its application in the food industry. A review. *Folia Microbiologica*, *59*(3), 241–250.

Kim, C. H., Axup, J. Y., Dubrovska, A., Kazane, S. A., Hutchins, B. A., Wold, E. D., Smider, V. V., & Schultz, P. G. (2012). Synthesis of Bispecific Antibodies using Genetically Encoded Unnatural Amino Acids. *Journal of the American Chemical Society*, *134*(24), 9918–9921.

Kim, J.-K., Bae, H. C., Ro, D. H., Lee, S., Lee, M. C., & Han, H.-S. (2021). Enhancement of Cartilage Regeneration of Synovial Stem Cells/Hydrogel by Using Transglutaminase-4. *Tissue Engineering Part A*, *27*(11–12), 761–770.

Kim, N., Lee, W.-K., Lee, S.-H., Jin, K. S., Kim, K.-H., Lee, Y., Song, M., & Kim, S.-Y. (2017). Inter-molecular crosslinking activity is engendered by the dimeric form of transglutaminase 2. *Amino Acids*, *49*(3), 461–471.

Király, R., Csosz, E., Kurtán, T., Antus, S., Szigeti, K., Simon-Vecsei, Z., Korponay-Szabó, I. R., Keresztessy, Z., & Fésüs, L. (2009). Functional significance of five noncanonical Ca2+-binding sites of human transglutaminase 2 characterized by site-directed mutagenesis. *The FEBS Journal*, *276*(23), 7083–7096.

Király, R., Thangaraju, K., Nagy, Z., Collighan, R., Nemes, Z., Griffin, M., & Fésüs, L. (2016). Isopeptidase activity of human transglutaminase 2: Disconnection from transamidation and characterization by kinetic parameters. *Amino Acids*, *48*(1), 31–40.

Klöck, C., & Khosla, C. (2012). Regulation of the activities of the mammalian transglutaminase family of enzymes. *Protein Science: A Publication of the Protein Society*, *21*(12), 1781–1791.

Kluger, R., Shen, L., Xiao, H., & Jones, R. T. (1996). Systematically Cross-Linked Human Hemoglobin: Functional Effects of 10 Å Spans between Beta Subunits at Lysine-82. *Journal of the American Chemical Society*, *118*(37), 8782–8786.

Krall, J. A., Beyer, E. M., & MacBeath, G. (2011). High- and Low-Affinity Epidermal Growth Factor Receptor-Ligand Interactions Activate Distinct Signaling Pathways. *PLoS ONE*, *6*(1), e15945.

Kramer, R. M., Crookes-Goodson, W. J., & Naik, R. R. (2007). The self-organizing properties of squid reflectin protein. *Nature Materials*, *6*(7), 533–538.

Kumari, P., Bowmik, S., Paul, S. K., Biswas, B., Banerjee, S. K., Murty, U. S., Ravichandiran, V., & Mohan, U. (2021). Sortase A: A chemoenzymatic approach for the labeling of cell surfaces. *Biotechnology and Bioengineering*, *118*(12), 4577–4589.

Kuo, T.-F., Tatsukawa, H., & Kojima, S. (2011). New insights into the functions and localization of nuclear transglutaminase 2. *The FEBS Journal*, *278*(24), 4756–4767.

Lai, T.-S., Lin, C.-J., & Greenberg, C. S. (2017). Role of tissue transglutaminase-2 (TG2)-mediated aminylation in biological processes. *Amino Acids*, *49*(3), 501–515.

Lam, D., Enright, H. A., Cadena, J., Peters, S. K. G., Sales, A. P., Osburn, J. J., Soscia, D. A., Kulp, K. S., Wheeler, E. K., & Fischer, N. O. (2019). Tissue-specific extracellular matrix accelerates the formation of neural networks and communities in a neuron-glia co-culture on a multi-electrode array. *Scientific Reports*, *9*(1), Article 1.

Land, M. F. (1972). The physics and biology of animal reflectors. *Progress in Biophysics and Molecular Biology*, *24*, 75–106.

Le, M., Gohr, C. M., & Rosenthal, A. K. (2001). Transglutaminase participates in the incorporation of latent TGFβ into the extracellular matrix of aging articular chondrocytes. *Connective Tissue Research*, *42*(4), 245–253.

Lee, J.-Y., Choo, J.-E., Choi, Y.-S., Park, J.-B., Min, D.-S., Lee, S.-J., Rhyu, H. K., Jo, I.-H., Chung, C.-P., & Park, Y.-J. (2007). Assembly of collagen-binding peptide with collagen as a bioactive scaffold for osteogenesis in vitro and in vivo. *Biomaterials*, *28*(29), 4257–4267.

Leffler, D. A., & Schuppan, D. (2010). Update on serologic testing in celiac disease. *The American Journal of Gastroenterology*, *105*(12), 2520–2524.

Leker, R. R., Toth, Z. E., Shahar, T., Cassiani-Ingoni, R., Szalayova, I., Key, S., Bratincsák, A., & Mezey, E. (2009). Transforming growth factor alpha induces angiogenesis and neurogenesis following stroke. *Neuroscience*, *163*(1), 233–243.

Levasseur, M. D., Hofmann, R., Edwardson, T. G. W., Hehn, S., Thanaburakorn, M., Bode, J. W., & Hilvert, D. (2022). Post-Assembly Modification of Protein Cages by Ubc9-Mediated Lysine Acylation. *Chembiochem: A European Journal of Chemical Biology*, *23*(20), e202200332.

Lewis Phillips, G. D., Li, G., Dugger, D. L., Crocker, L. M., Parsons, K. L., Mai, E., Blättler, W. A., Lambert, J. M., Chari, R. V. J., Lutz, R. J., Wong, W. L. T., Jacobson, F. S., Koeppen, H., Schwall, R. H., Kenkare-Mitra, S. R., Spencer, S. D., & Sliwkowski, M. X. (2008). Targeting HER2-positive breast cancer with trastuzumab-DM1, an antibody-cytotoxic drug conjugate. *Cancer Research*, *68*(22), 9280–9290.

Li, K., Tay, F. R., & Yiu, C. K. Y. (2020). The past, present and future perspectives of matrix metalloproteinase inhibitors. *Pharmacology & Therapeutics*, *207*, 107465.

Liang, H., Li, X., Wang, B., Chen, B., Zhao, Y., Sun, J., Zhuang, Y., Shi, J., Shen, H., Zhang, Z., & Dai, J. (2016). A collagen-binding EGFR antibody fragment targeting tumors with a collagen-rich extracellular matrix. *Scientific Reports*, *6*, 18205.

Lim, K. L., Chew, K. C. M., Tan, J. M. M., Wang, C., Chung, K. K. K., Zhang, Y., Tanaka, Y., Smith, W., Engelender, S., Ross, C. A., Dawson, V. L., & Dawson, T. M. (2005). Parkin Mediates Nonclassical, Proteasomal-Independent Ubiquitination of Synphilin-1: Implications for Lewy Body Formation. *Journal of Neuroscience*, *25*(8), 2002–2009.

Lin, C.-W., & Ting, A. Y. (2006). Transglutaminase-catalyzed site-specific conjugation of small-molecule probes to proteins in vitro and on the surface of living cells. *Journal of the American Chemical Society*, *128*(14), 4542–4543.

Lin, Y., Hwang, J.-Y., Mangala, L., Fok, J., Merritt, W., Spannuth, W., Nick, A., Vivas-Mejia, P., Deavers, M., Coleman, R., Lopez-Berestein, G., Mehta, K., & Sood, A. (2008). Biological and therapeutic implications of tissue transglutaminase (TG2) silencing in advanced ovarian carcinoma. *Cancer Research*, *68*(9_Supplement), 1202.

Lindstad, C. B., Dewan, A. E., Stamnaes, J., Sollid, L. M., & du Pré, M. F. (2021). TG2-gluten complexes as antigens for gluten-specific and transglutaminase-2 specific B cells in celiac disease. *PloS One*, *16*(11), e0259082.

Liu, S., Cerione, R. A., & Clardy, J. (2002). Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity. *Proceedings of the National Academy of Sciences*, *99*(5), 2743–2747.

Liu, W., Tang, X., Qi, X., Fu, X., Ghimire, S., Ma, R., Li, S., Zhang, N., & Si, H. (2020). The Ubiquitin Conjugating Enzyme: An Important Ubiquitin Transfer Platform in Ubiquitin-Proteasome System. *International Journal of Molecular Sciences*, *21*(8), Article 8.

Liu, Y., Huang, L., Shan, M., Sang, J., Li, Y., Jia, L., Wang, N., Wang, S., Shao, S., Liu, F., & Lu, F. (2019). Enhancing the activity and thermostability of Streptomyces mobaraensis transglutaminase by directed evolution and molecular dynamics simulation. *Biochemical Engineering Journal*, *151*, 107333.

Liu, Y., Liu, Y., Deng, J., Li, W., & Nie, X. (2021). Fibroblast Growth Factor in Diabetic Foot Ulcer: Progress and Therapeutic Prospects. *Frontiers in Endocrinology*, *12*, 744-868.

Loffredo, F. S., Pancoast, J. R., Cai, L., Vannelli, T., Dong, J. Z., Lee, R. T., & Patwari, P. (2014). Targeted delivery to cartilage is critical for in vivo efficacy of insulin-like growth factor 1 in a rat model of osteoarthritis. *Arthritis & Rheumatology (Hoboken, N.J.)*, *66*(5), 1247–1255.

López-Otín, C., & Matrisian, L. M. (2007). Emerging roles of proteases in tumour suppression. *Nature Reviews Cancer*, *7*(10), Article 10.

Lorand, L., & Graham, R. M. (2003). Transglutaminases: Crosslinking enzymes with pleiotropic functions. *Nature Reviews Molecular Cell Biology*, *4*(2), Article 2.

Lorand, L., Velasco, P. T., Murthy, S. N., Wilson, J., & Parameswaran, K. N. (1992). Isolation of transglutaminase-reactive sequences from complex biological systems: A prominent lysine donor sequence in bovine lens. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(23), 11161–11163.

Mädler, S., Bich, C., Touboul, D., & Zenobi, R. (2009). Chemical cross-linking with NHS esters: A systematic study on amino acid reactivities. *Journal of Mass Spectrometry: JMS*, *44*(5), 694–706.

Mahoney, S. A., Wilkinson, M., Smith, S., & Haynes, L. W. (2000). Stabilization of neurites in cerebellar granule cells by transglutaminase activity: Identification of midkine and galectin-3 as substrates. *Neuroscience*, *101*(1), 141–155.

Majumdar, S., Wang, X., Sommerfeld, S. D., Chae, J. J., Athanasopoulou, E.-N., Shores, L. S., Duan, X., Amzel, L. M., Stellacci, F., Schein, O., Guo, Q., Singh, A., & Elisseeff, J. H. (2018). Cyclodextrin Modulated Type I Collagen Self-Assembly to Engineer Biomimetic Cornea Implants. *Advanced Functional Materials*, *28*(41), 1804076.

Mao, H., Hart, S. A., Schink, A., & Pollok, B. A. (2004). Sortase-mediated protein ligation: A new method for protein engineering. *Journal of the American Chemical Society*, *126*(9), 2670–2671.

Marraffini, L. A., DeDent, A. C., & Schneewind, O. (2006). Sortases and the Art of Anchoring Proteins to the Envelopes of Gram-Positive Bacteria. *Microbiology and Molecular Biology Reviews*, *70*(1), 192–221.

Martino, M. M., Briquez, P. S., Güç, E., Tortelli, F., Kilarski, W. W., Metzger, S., Rice, J. J., Kuhn, G. A., Müller, R., Swartz, M. A., & Hubbell, J. A. (2014). Growth factors engineered for super-affinity to the extracellular matrix enhance tissue healing. *Science (New York, N.Y.)*, *343*(6173), 885–888.

Mehta, K., Fok, J., Miller, F. R., Koul, D., & Sahin, A. A. (2004). Prognostic Significance of Tissue Transglutaminase in Drug Resistant and Metastatic Breast Cancer. *Clinical Cancer Research*, *10*(23), 8068–8076.

Mero, A., Spolaore, B., Veronese, F. M., & Fontana, A. (2009). Transglutaminase-Mediated PEGylation of Proteins: Direct Identification of the Sites of Protein Modification by Mass Spectrometry using a Novel Monodisperse PEG. *Bioconjugate Chemistry*, *20*(2), 384–389.

Migneault, I., Dartiguenave, C., Bertrand, M. J., & Waldron, K. C. (2004). Glutaraldehyde: Behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. *BioTechniques*, *37*(5), 790–802.

Milczek, E. M. (2018). Commercial Applications for Enzyme-Mediated Protein Conjugation: New Developments in Enzymatic Processes to Deliver Functionalized Proteins on the Commercial Scale. *Chemical Reviews*, *118*(1), 119–141.

Mishra, P. K., Yoo, C.-M., Hong, E., & Rhee, H. W. (2020). Photo-crosslinking: An Emerging Chemical Tool for Investigating Molecular Networks in Live Cells. *ChemBioChem*, *21*(7), 924–932.

E. Morgan, H., Bruce Turnbull, W., & E. Webb, M. (2022). Challenges in the use of sortase and other peptide ligases for site-specific protein modification. *Chemical Society Reviews*, *51*(10), 4121–4145.

Müller, L. J., Pels, E., Schurmans, L. R. H. M., & Vrensen, G. F. J. M. (2004). A new three-dimensional model of the organization of proteoglycans and collagen fibrils in the human corneal stroma. *Experimental Eye Research*, *78*(3), 493–501.

Muma, N. A., & Mi, Z. (2015). Serotonylation and Transamidation of Other Monoamines. *ACS Chemical Neuroscience*, *6*(7), 961–969.

Murthy, S. N. P., Lukas, T. J., Jardetzky, T. S., & Lorand, L. (2009). Selectivity in the post-translational, transglutaminase-dependent acylation of lysine residues. *Biochemistry*, *48*(12), 2654–2660.

Nakamura, N., Kimura, T., & Kishida, A. (2017). Overview of the Development, Applications, and Future Perspectives of Decellularized Tissues and Organs. *ACS Biomaterials Science & Engineering*, *3*(7), 1236–1244.

Nakaoka, H., Perez, D. M., Baek, K. J., Das, T., Husain, A., Misono, K., Im, M.-J., & Graham, R. M. (1994). Gh: A GTP-Binding Protein with Transglutaminase Activity and Receptor Signaling Function. *Science*, *264*(5165), 1593–1596.

Nanda, N., Iismaa, S. E., Owens, W. A., Husain, A., Mackay, F., & Graham, R. M. (2001). Targeted Inactivation of Gh/Tissue Transglutaminase II. *Journal of Biological Chemistry*, *276*(23), 20673–20678.

Narasimhan, V., Siddique, R. H., Lee, J. O., Kumar, S., Ndjamen, B., Du, J., Hong, N., Sretavan, D., & Choo, H. (2018). Multifunctional biophotonic nanostructures inspired by the longtail glasswing butterfly for medical devices. *Nature Nanotechnology*, *13*(6), 512–519.

Nardacci, R., Lo Iacono, O., Ciccosanti, F., Falasca, L., Addesso, M., Amendola, A., Antonucci, G., Craxì, A., Fimia, G. M., Iadevaia, V., Melino, G., Ruco, L., Tocci, G., Ippolito, G., & Piacentini, M. (2003). Transglutaminase Type II Plays a Protective Role in Hepatic Injury. *The American Journal of Pathology*, *162*(4), 1293–1303.

Neet, K. E., & Koshland, D. E. (1966). The conversion of serine at the active site of subtilisin to cysteine: A 'chemical mutation'. *Proceedings of the National Academy of Sciences*, *56*(5), 1606–1611.

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

Nemes, Z., Adány, R., Balázs, M., Boross, P., & Fésüs, L. (1997). Identification of cytoplasmic actin as an abundant glutaminyl substrate for tissue transglutaminase in HL-60 and U937 cells undergoing apoptosis. *The Journal of Biological Chemistry*, *272*(33), 20577–20583.

Nguyen, G. K. T., Wang, S., Qiu, Y., Hemu, X., Lian, Y., & Tam, J. P. (2014). Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nature Chemical Biology*, *10*(9), Article 9.

Nguyen, G. K. T., Kam, A., Loo, S., Jansson, A. E., Pan, L. X., & Tam, J. P. (2015). Butelase 1: A Versatile Ligase for Peptide and Protein Macrocyclization. *Journal of the American Chemical Society*, *137*(49), 15398–15401.

Nguyen, G. K. T., Qiu, Y., Cao, Y., Hemu, X., Liu, C.-F., & Tam, J. P. (2016). Butelase-mediated cyclization and ligation of peptides and proteins. *Nature Protocols*, *11*(10), 1977–1988.

Nonaka, M., Tanaka, H., Okiyama, A., Motoki, M., Ando, H., Umeda, K., & Matsuura, A. (1989). Polymerization of Several Proteins by Ca2+-Independent Transglutaminase Derived from Microorganisms. *Agricultural and Biological Chemistry*, *53*(10), 2619–2623.

Nonis, S. G., Haywood, J., Schmidberger, J. W., Mackie, E. R. R., Soares da Costa, T. P., Bond, C. S., & Mylne, J. S. (2021). Structural and biochemical analyses of concanavalin A circular permutation by jack bean asparaginyl endopeptidase. *The Plant Cell*, *33*(8), 2794–2811.

Noorbala, M. T., Noorbala, M., Dashti-Rahmatabadi, M. H., Noorbala, M., Noorbala, R., & Mozaffary, B. (2016). Comparison of Hydrogel Produced by Radiation as Applied at the Research Center (Yazd Branch) With MaxGel and Routine Dressing for Second-Degree Burn Repair in Yazd Burn Hospital. *Iranian Red Crescent Medical Journal*, *18*(8), e24384.

Nuijens, T., Toplak, A., Van de Meulenreek, M. B. A. C., Schmidt, M., Goldbach, M., Janssen, D. B., & Quaedflieg, P. J. L. M. (2016). Chemo-enzymatic peptide synthesis (CEPS) using omniligases and selective peptiligases Efficient biocatalysts for assembling linear and cyclic peptides and protein conjugates. *Chimica Oggi-Chemistry Today*, *34*(6A), 16–19.

Nyborg, A. C., Ward, C., Zacco, A., Chacko, B., Grinberg, L., Geoghegan, J. C., Bean, R., Wendeler, M., Bartnik, F., O'Connor, E., Gruia, F., Iyer, V., Feng, H., Roy, V., Berge, M., Miner, J. N., Wilson, D. M., Zhou, D., Nicholson, S., … Baca, M. (2016). A Therapeutic Uricase with Reduced Immunogenicity Risk and Improved Development Properties. *PLOS ONE*, *11*(12), e0167935.

Oberai, A., Ihm, Y., Kim, S., & Bowie, J. U. (2006). A limited universe of membrane protein families and folds. *Protein Science*, *15*(7), 1723–1734.

Ohtsuka, T., Ota, M., Nio, N., & Motoki, M. (2000). Comparison of Substrate Specificities of Transglutaminases Using Synthetic Peptides as Acyl donors. *Bioscience, Biotechnology, and Biochemistry*, *64*(12), 2608–2613.

Oteng-Pabi, S. K., Clouthier, C. M., & Keillor, J. W. (2018). Design of a glutamine substrate tag enabling protein labelling mediated by Bacillus subtilis transglutaminase. *PLOS ONE*, *13*(5), e0197956.

Palomino-Durand, C., Pauthe, E., & Gand, A. (2021). Fibronectin-Enriched Biomaterials, Biofunctionalization, and Proactivity: A Review. *Applied Sciences*, *11*(24), Article 24.

Park, D., Choi, S. S., & Ha, K.-S. (2010). Transglutaminase 2: A multi-functional protein in multiple subcellular compartments. *Amino Acids*, *39*(3), 619–631.

Parker, C. G., & Pratt, M. R. (2020). Click Chemistry in Proteomic Investigations. *Cell*, *180*(4), 605–632.

Patterson, D., Schwarz, B., Avera, J., Western, B., Hicks, M., Krugler, P., Terra, M., Uchida, M., McCoy, K., & Douglas, T. (2017). Sortase-Mediated Ligation as a Modular Approach for the Covalent Attachment of Proteins to the Exterior of the Bacteriophage P22 Virus-like Particle. *Bioconjugate Chemistry*, *28*(8), 2114–2124.

Pavarini, E., Andreani, L. C., Soci, C., Galli, M., Marabelli, F., & Comoretto, D. (2005). Band structure and optical properties of opal photonic crystals. *Physical Review B*, *72*(4), 045102.

Peccoud, J., Blauvelt, M. F., Cai, Y., Cooper, K. L., Crasta, O., DeLalla, E. C., Evans, C., Folkerts, O., Lyons, B. M., Mane, S. P., Shelton, R., Sweede, M. A., & Waldon, S. A. (2008). Targeted Development of Registries of Biological Parts. *PLOS ONE*, *3*(7), e2671.

Philp, C. J., Siebeke, I., Clements, D., Miller, S., Habgood, A., John, A. E., Navaratnam, V., Hubbard, R. B., Jenkins, G., & Johnson, S. R. (2018). Extracellular Matrix Cross-Linking Enhances Fibroblast Growth and Protects against Matrix Proteolysis in Lung Fibrosis. *American Journal of Respiratory Cell and Molecular Biology*, *58*(5), 594–603.

Piersimoni, L., Kastritis, P. L., Arlt, C., & Sinz, A. (2022). Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein–Protein Interactions─A Method for All Seasons. *Chemical Reviews*, *122*(8), 7500–7531.

Pietroni, V., Di Giorgi, S., Paradisi, A., Ahvazi, B., Candi, E., & Melino, G. (2008). Inactive and Highly Active, Proteolytically Processed Transglutaminase-5 in Epithelial Cells. *Journal of Investigative Dermatology*, *128*(12), 2760–2766.

Pinkas, D. M., Strop, P., Brunger, A. T., & Khosla, C. (2007). Transglutaminase 2 undergoes a large conformational change upon activation. *PLoS Biology*, *5*(12), e327.

Pinto, F., Thornton, E. L., & Wang, B. (2020). An expanded library of orthogonal split inteins enables modular multi-peptide assemblies. *Nature Communications*, *11*(1), Article 1.

Piper, J. L., Gray, G. M., & Khosla, C. (2002). High selectivity of human tissue transglutaminase for immunoactive gliadin peptides: Implications for celiac sprue. *Biochemistry*, *41*(1), 386–393.

Plugis, N. M., Palanski, B. A., Weng, C.-H., Albertelli, M., & Khosla, C. (2017). Thioredoxin-1 Selectively Activates Transglutaminase 2 in the Extracellular Matrix of the Small Intestine: implications for celiac disease. *Journal of Biological Chemistry*, *292*(5), 2000–2008.

Pirzer, T., Becher, K.-S., Rieker, M., Meckel, T., Mootz, H. D., & Kolmar, H. (2018). Generation of Potent Anti-HER1/2 Immunotoxins by Protein Ligation Using Split Inteins. *ACS Chemical Biology*, *13*(8), 2058–2066.

Pishesha, N., Bilate, A. M., Wibowo, M. C., Huang, N.-J., Li, Z., Deshycka, R., Bousbaine, D., Li, H., Patterson, H. C., Dougan, S. K., Maruyama, T., Lodish, H. F., & Ploegh, H. L. (2017). Engineered erythrocytes covalently linked to antigenic peptides can protect against autoimmune disease. *Proceedings of the National Academy of Sciences*, *114*(12), 3157–3162.

Pitolli, C., Pietroni, V., Marekov, L., Terrinoni, A., Yamanishi, K., Mazzanti, C., Melino, G., & Candi, E. (2017). Characterization of TG2 and TG1-TG2 double knock-out mouse epidermis. *Amino Acids*, *49*(3), 635–642.

Popp, M. W., Artavanis-Tsakonas, K., & Ploegh, H. L. (2009). Substrate Filtering by the Active Site Crossover Loop in UCHL3 Revealed by Sortagging and Gain-of-function Mutations *. *Journal of Biological Chemistry*, *284*(6), 3593–3602.

Popp, M. W.-L., & Ploegh, H. L. (2011). Making and breaking peptide bonds: Protein engineering using sortase. *Angewandte Chemie (International Ed. in English)*, *50*(22), 5024–5032.

Porzio, O., Massa, O., Cunsolo, V., Colombo, C., Malaponti, M., Bertuzzi, F., Hansen, T., Johansen, A., Pedersen, O., Meschi, F., Terrinoni, A., Melino, G., Federici, M., Decarlo, N., Menicagli, M., Campani, D., Marchetti, P., Ferdaoussi, M., Froguel, P., … Barbetti, F. (2007). Missense mutations in the TGM2 gene encoding transglutaminase 2 are found in patients with early-onset type 2 diabetes. *Human Mutation*, *28*(11), 1150–1150.

Propper, D. J., & Balkwill, F. R. (2022). Harnessing cytokines and chemokines for cancer therapy. *Nature Reviews Clinical Oncology*, *19*(4), Article 4.

Prum, R. O., & Torres, R. (2003). Structural colouration of avian skin: Convergent evolution of coherently scattering dermal collagen arrays. *The Journal of Experimental Biology*, *206*(Pt 14), 2409–2429.

Prum, R. O., Dufresne, E. R., Quinn, T., & Waters, K. (2009). Development of colour-producing beta-keratin nanostructures in avian feather barbs. *Journal of the Royal Society, (*Suppl 2), S253-265.

Qin, G., Dennis, P. B., Zhang, Y., Hu, X., Bressner, J. E., Sun, Z., Crookes-Goodson, W. J., Naik, R. R., Omenetto, F. G., & Kaplan, D. L. (2013). Recombinant reflectin-based optical materials. *Journal of Polymer Science Part B: Polymer Physics*, *51*(4), 254–264.

Qin, W. et al. (2014) 'MicroRNA-124 regulates TGF-α-induced epithelial-mesenchymal transition in human prostate cancer cells', International Journal of Oncology, 45(3), pp. 1225–1231.

Rodolfo, C., Mormone, E., Matarrese, P., Ciccosanti, F., Farrace, M. G., Garofano, E., Piredda, L., Fimia, G. M., Malorni, W., & Piacentini, M. (2004). Tissue Transglutaminase Is a Multifunctional BH3-only Protein. *Journal of Biological Chemistry*, *279*(52), 54783–54792.

Rossin, F., D'Eletto, M., Macdonald, D., Farrace, M. G., & Piacentini, M. (2012). TG2 transamidating activity acts as a reostat controlling the interplay between apoptosis and autophagy. *Amino Acids*, *42*(5), 1793–1802.

Roy, I., Smith, O., Clouthier, C. M., & Keillor, J. W. (2013). Expression, purification and kinetic characterisation of human tissue transglutaminase. *Protein Expression and Purification*, *87*(1), 41–46.

Ruoppolo, M., Orrù, S., D'Amato, A., Francese, S., Rovero, P., Marino, G., & Esposito, C. (2003). Analysis of transglutaminase protein substrates by functional proteomics. *Protein Science: A Publication of the Protein Society*, *12*(6), 1290–1297.

Ryu, Y., & Schultz, P. G. (2006). Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nature Methods*, *3*(4), Article 4.

Saito, G., Swanson, J. A., & Lee, K.-D. (2003). Drug delivery strategy utilizing conjugation via reversible disulfide linkages: Role and site of cellular reducing activities. *Advanced Drug Delivery Reviews*, *55*(2), 199–215.

Sarang, Z., Molnár, P., Németh, T., Gomba, S., Kardon, T., Melino, G., Cotecchia, S., Fésüs, L., & Szondy, Z. (2005). Tissue transglutaminase (TG2) acting as G protein protects hepatocytes against Fas-mediated cell death in mice. *Hepatology*, *42*(3), 578–587.

Sarang, Z., Mádi, A., Koy, C., Varga, S., Glocker, M. O., Ucker, D. S., Kuchay, S., Chishti, A. H., Melino, G., Fésüs, L., & Szondy, Z. (2007). Tissue transglutaminase (TG2) facilitates phosphatidylserine exposure and calpain activity in calcium-induced death of erythrocytes. *Cell Death and Differentiation*, *14*(10), 1842–1844.

Sarkar, N., Clarke, D., & Waelsch, H. (1957). An enzymically catalyzed incorporation of amines into proteins / Nirmal K. Sarkar Donald D. Clarke Heinrich Waelsch Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York, and the New York Psychiatric Institute, New York, N.Y. (U.S.A.). *Chemistry Faculty Publications*.

Scarpellini, A., Huang, L., Burhan, I., Schroeder, N., Funck, M., Johnson, T. S., & Verderio, E. A. M. (2014). Syndecan-4 Knockout Leads to Reduced Extracellular Transglutaminase-2 and Protects against Tubulointerstitial Fibrosis. *Journal of the American Society of Nephrology*, *25*(5), 1013–1027.

Scheu, A. H. A., Lim, S. Y. T., Metzner, F. J., Mohammed, S., & Howarth, M. (2021). NeissLock provides an inducible protein anhydride for covalent targeting of endogenous proteins. *Nature Communications*, *12*(1), Article 1.

Schmidt, M., Toplak, A., Quaedflieg, P. J. L. M., Ippel, H., Richelle, G. J. J., Hackeng, T. M., van Maarseveen, J. H., & Nuijens, T. (2017). Omniligase-1: A Powerful Tool for Peptide Head-to-Tail Cyclization. *Advanced Synthesis & Catalysis*, *359*(12), 2050–2055.

Schmitz, T., Bäuml, C. A., & Imhof, D. (2020). Inhibitors of blood coagulation factor XIII. *Analytical Biochemistry*, *605*, 113708.

Schoene, C., Bennett, S. P., & Howarth, M. (2016). SpyRing interrogation: Analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries. *Scientific Reports*, *6*(1), Article 1.

Schroth, M., Meißner, U., Cesnjevar, R., Weyand, M., Singer, H., Rascher, W., & Klinge, J. (2006). Recombinant Factor XIII Reduces Severe Pleural Effusion in Children after Open-Heart Surgery. *Pediatric Cardiology*, *27*(1), 56–60.

Schumacher, M. A., Dixon, M. M., Kluger, R., Jones, R. T., & Brennan, R. G. (1995). Allosteric transition intermediates modelled by crosslinked haemoglobins. *Nature*, *375*(6526), 84–87.

Sekar, G., Stevens, A. J., Mostafavi, A. Z., Sashi, P., Muir, T. W., & Cowburn, D. (2022). A Conserved Histidine Residue Drives Extein Dependence in an Enhanced Atypically Split Intein. *Journal of the American Chemical Society*, *144*(41), 19196–19203.

Shao, M., Cao, L., Shen, C., Satpathy, M., Chelladurai, B., Bigsby, R. M., Nakshatri, H., & Matei, D. (2009). Epithelial-to-mesenchymal transition and ovarian tumor progression induced by tissue transglutaminase. *Cancer Research*, *69*(24), 9192–9201.

Shi, Q., Kim, S.-Y., Blass, J. P., & Cooper, A. J. L. (2002). Expression in Escherichia coli and purification of hexahistidine-tagged human tissue transglutaminase. *Protein Expression and Purification*, *24*(3), 366–373.

Shi, Y., Liu, K., Zhang, Z., Tao, X., Chen, H.-Y., Kingshott, P., & Wang, P.-Y. (2020). Decoration of Material Surfaces with Complex Physicochemical Signals for Biointerface Applications. *ACS Biomaterials Science & Engineering*, *6*(4), 1836–1851.

Siddique, R. H., Donie, Y. J., Gomard, G., Yalamanchili, S., Merdzhanova, T., Lemmer, U., & Hölscher, H. (2017). Bioinspired phase-separated disordered nanostructures for thin photovoltaic absorbers. *Science Advances*, *3*(10), e1700232.

Siegel, M., Strnad, P., Watts, R. E., Choi, K., Jabri, B., Omary, M. B., & Khosla, C. (2008). Extracellular Transglutaminase 2 Is Catalytically Inactive, but Is Transiently Activated upon Tissue Injury. *PLoS ONE*, *3*(3), e1861.

Siegel, S. D., Liu, J., & Ton-That, H. (2016). Biogenesis of the Gram-positive bacterial cell envelope. *Current Opinion in Microbiology*, *34*, 31–37.

Singh, B., & Coffey, R. J. (2014). From wavy hair to naked proteins: The role of transforming growth factor alpha in health and disease. *Seminars in Cell & Developmental Biology*, *28*, 12–21.

Singh, M., Thakur, M., Mishra, M., Yadav, M., Vibhuti, R., Menon, A. M., Nagda, G., Dwivedi, V. P., Dakal, T. C., & Yadav, V. (2021). Gene regulation of intracellular adhesion molecule-1 (ICAM-1): A molecule with multiple functions. *Immunology Letters*, *240*, 123–136.

Sinz, A. (2017). Divide and conquer: Cleavable cross-linkers to study protein conformation and protein–protein interactions. *Analytical and Bioanalytical Chemistry*, *409*(1), 33–44.

Slaughter, T. F., Achyuthan, K. E., Lai, T. S., & Greenberg, C. S. (1992). A microtiter plate transglutaminase assay utilizing 5-(biotinamido)pentylamine as substrate. *Analytical Biochemistry*, *205*(1), 166–171.

Smith, K., & Rennie, M. J. (2007). New approaches and recent results concerning human-tissue collagen synthesis. *Current Opinion in Clinical Nutrition and Metabolic Care*, *10*(5), 582–590.

Smith, M. E. B., Schumacher, F. F., Ryan, C. P., Tedaldi, L. M., Papaioannou, D., Waksman, G., Caddick, S., & Baker, J. R. (2010). Protein Modification, Bioconjugation, and Disulfide Bridging Using Bromomaleimides. *Journal of the American Chemical Society*, *132*(6), 1960–1965.

Song, H., Chang, W., Lim, S., Seo, H.-S., Shim, C. Y., Park, S., Yoo, K.-J., Kim, B.-S., Min, B.-H., Lee, H., Jang, Y., Chung, N., & Hwang, K.-C. (2007). Tissue transglutaminase is essential for integrin-mediated survival of bone marrow-derived mesenchymal stem cells. *Stem Cells (Dayton, Ohio)*, *25*(6), 1431–1438.

Spirig, T., Weiner, E. M., & Clubb, R. T. (2011). Sortase enzymes in Gram-positive bacteria. *Molecular Microbiology*, *82*(5), 1044–1059.

Spolaore, B., Raboni, S., Satwekar, A. A., Grigoletto, A., Mero, A., Montagner, I. M., Rosato, A., Pasut, G., & Fontana, A. (2016). Site-Specific Transglutaminase-Mediated Conjugation of Interferon α-2b at Glutamine or Lysine Residues. *Bioconjugate Chemistry*, *27*(11), 2695–2706.

Springer, T. A. (1990). Adhesion receptors of the immune system. *Nature*, *346*(6283), Article 6283.

Srinivasarao, M. (1999). Nano-Optics in the Biological World: Beetles, Butterflies, Birds, and Moths. *Chemical Reviews*, *99*(7), 1935–1962.

Stachel, I., Schwarzenbolz, U., Henle, T., & Meyer, M. (2010). Cross-Linking of Type I Collagen with Microbial Transglutaminase: Identification of Cross-Linking Sites. *Biomacromolecules*, *11*(3), 698–705.

Stamnaes, J., Pinkas, D. M., Fleckenstein, B., Khosla, C., & Sollid, L. M. (2010). Redox regulation of transglutaminase 2 activity. *The Journal of Biological Chemistry*, *285*(33), 25402–25409.

Stamnaes, J., Iversen, R., Pré, M. F. du, Chen, X., & Sollid, L. M. (2015). Enhanced B-Cell Receptor Recognition of the Autoantigen Transglutaminase 2 by Efficient Catalytic Self-Multimerization. *PLOS ONE*, *10*(8), e0134922.

Stamnaes, J., Cardoso, I., Iversen, R., & Sollid, L. M. (2016). Transglutaminase 2 strongly binds to an extracellular matrix component other than fibronectin via its second C-terminal beta-barrel domain. *The FEBS Journal*, *283*(21), 3994–4010.

Steffen, W., Ko, F. C., Patel, J., Lyamichev, V., Albert, T. J., Benz, J., Rudolph, M. G., Bergmann, F., Streidl, T., Kratzsch, P., Boenitz-Dulat, M., Oelschlaegel, T., & Schraeml, M. (2017). Discovery of a microbial transglutaminase enabling highly site-specific labeling of proteins. *The Journal of Biological Chemistry*, *292*(38), 15622–15635.

Stern, J., Moraïs, S., Lamed, R., & Bayer, E. A. (2016). Adaptor Scaffoldins: An Original Strategy for Extended Designer Cellulosomes, Inspired from Nature. *mBio*, *7*(2), 10.1128/mbio.00083-16.

Strop, P. (2014). Versatility of Microbial Transglutaminase. *Bioconjugate Chemistry*, *25*(5), 855–862.

Strop, P., Liu, S.-H., Dorywalska, M., Delaria, K., Dushin, R. G., Tran, T.-T., Ho, W.-H., Farias, S., Casas, M. G., Abdiche, Y., Zhou, D., Chandrasekaran, R., Samain, C., Loo, C., Rossi, A., Rickert, M., Krimm, S., Wong, T., Chin, S. M., … Rajpal, A. (2013). Location matters: Site of conjugation modulates stability and pharmacokinetics of antibody drug conjugates. *Chemistry & Biology*, *20*(2), 161–167.

Subramanian, K., Góra, A., Spruijt, R., Mitusińska, K., Suarez-Diez, M., Martins Dos Santos, V., & Schaap, P. J. (2018). Modulating D-amino acid oxidase (DAAO) substrate specificity through facilitated solvent access. *PloS One*, *13*(6), e0198990.

Sugimura, Y., Hosono, M., Wada, F., Yoshimura, T., Maki, M., & Hitomi, K. (2006). Screening for the Preferred Substrate Sequence of Transglutaminase Using a Phage-displayed Peptide Library: Identification Of Peptide Substrates For Tgase 2 And Factor XIIIA *. *Journal of Biological Chemistry*, *281*(26), 17699–17706.

Sung, H.-W., Huang, R.-N., Huang, L. L. H., Tsai, C.-C., & Chiu, C.-T. (1998). Feasibility study of a natural crosslinking reagent for biological tissue fixation. *Journal of Biomedical Materials Research*, *42*(4), 560–567.

Szondy, Z., Sarang, Z., Molnar, P., Nemeth, T., Piacentini, M., Mastroberardino, P. G., Falasca, L., Aeschlimann, D., Kovacs, J., Kiss, I., Szegezdi, E., Lakos, G., Rajnavolgyi, E., Birckbichler, P. J., Melino, G., & Fesus, L. (2003). Transglutaminase 2-/- mice reveal a phagocytosis-associated crosstalk between macrophages and apoptotic cells. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(13), 7812–7817.

Takayama, Y., Kusamori, K., & Nishikawa, M. (2019). Click Chemistry as a Tool for Cell Engineering and Drug Delivery. *Molecules*, *24*(1), Article 1.

Takazawa, T., Kamiya, N., Ueda, H., & Nagamune, T. (2004). Enzymatic labeling of a single chain variable fragment of an antibody with alkaline phosphatase by microbial transglutaminase. *Biotechnology and Bioengineering*, *86*(4), 399–404.

Taki, M., Shiota, M., & Taira, K. (2004). Transglutaminase-mediated N- and C-terminal fluorescein labeling of a protein can support the native activity of the modified protein. *Protein Engineering, Design & Selection: PEDS*, *17*(2), 119–126.

Tang, T. M. S., Cardella, D., Lander, A. J., Li, X., Escudero, J. S., Tsai, Y.-H., & Luk, L. Y. P. (2020). Use of an asparaginyl endopeptidase for chemo-enzymatic peptide and protein labeling. *Chemical Science*, *11*(23), 5881–5888.

Tatsukawa, H., & Hitomi, K. (2021). Role of Transglutaminase 2 in Cell Death, Survival, and Fibrosis. *Cells*, *10*(7), Article 7.

Tatsukawa, H., Fukaya, Y., Frampton, G., Martinez–Fuentes, A., Suzuki, K., Kuo, T.-F., Nagatsuma, K., Shimokado, K., Okuno, M., Wu, J., Iismaa, S., Matsuura, T., Tsukamoto, H., Zern, M. A., Graham, R. M., & Kojima, S. (2009). Role of Transglutaminase 2 in Liver

Injury via Cross-linking and Silencing of Transcription Factor Sp1. *Gastroenterology*, *136*(5), 1783-1795.e10.

Taylor, R. J., Geeson, M. B., Journeaux, T., & Bernardes, G. J. L. (2022). Chemical and Enzymatic Methods for Post-Translational Protein–Protein Conjugation. *Journal of the American Chemical Society*, *144*(32), 14404–14419.

Teesalu, K., Uibo, O., Uibo, R., & Utt, M. (2012). Kinetic and functional characterisation of the heparin-binding peptides from human transglutaminase 2. *Journal of Peptide Science: An Official Publication of the European Peptide Society*, *18*(5), 350–356.

Toplak, A., Teixeira de Oliveira, E. F., Schmidt, M., Rozeboom, H. J., Wijma, H. J., Meekels, L. K. M., de Visser, R., Janssen, D. B., & Nuijens, T. (2021). From thiol-subtilisin to omniligase: Design and structure of a broadly applicable peptide ligase. *Computational and Structural Biotechnology Journal*, *19*, 1277–1287.

Udalova, I. A., Ruhmann, M., Thomson, S. J. P., & Midwood, K. S. (2011). Expression and immune function of tenascin-C. *Critical Reviews in Immunology*, *31*(2), 115–145.

Ueda, A., Umetsu, M., Nakanishi, T., Hashikami, K., Nakazawa, H., Hattori, S., Asano, R., & Kumagai, I. (2020). Chemically Crosslinked Bispecific Antibodies for Cancer Therapy: Breaking from the Structural Restrictions of the Genetic Fusion Approach. *International Journal of Molecular Sciences*, *21*(3), Article 3.

Usmani, S. S., Bedi, G., Samuel, J. S., Singh, S., Kalra, S., Kumar, P., Ahuja, A. A., Sharma, M., Gautam, A., & Raghava, G. P. S. (2017). THPdb: Database of FDA-approved peptide and protein therapeutics. *PloS One*, *12*(7), e0181748.

van den Brûle, F. A., Liu, F. T., & Castronovo, V. (1998). Transglutaminase-mediated oligomerization of galectin-3 modulates human melanoma cell interactions with laminin. *Cell Adhesion and Communication*, *5*(6), 425–435.

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

Verderio, E. A. M., Telci, D., Okoye, A., Melino, G., & Griffin, M. (2003). A novel RGD-independent cel adhesion pathway mediated by fibronectin-bound tissue transglutaminase rescues cells from anoikis. *The Journal of Biological Chemistry*, *278*(43), 42604–42614.

Vowinckel, J., Stahlberg, S., Paulmann, N., Bluemlein, K., Grohmann, M., Ralser, M., & Walther, D. J. (2012). Histaminylation of glutamine residues is a novel posttranslational modification implicated in G-protein signaling. *FEBS Letters*, *586*(21), 3819–3824.

Výborný, K., Vallová, J., Kočí, Z., Kekulová, K., Jiráková, K., Jendelová, P., Hodan, J., & Kubinová, Š. (2019). Genipin and EDC crosslinking of extracellular matrix hydrogel derived from human umbilical cord for neural tissue repair. *Scientific Reports*, *9*(1), Article 1.

Wagner, G. R., & Payne, R. M. (2013). Widespread and Enzyme-independent Nϵ-Acetylation and Nϵ-Succinylation of Proteins in the Chemical Conditions of the Mitochondrial Matrix. *Journal of Biological Chemistry*, *288*(40), 29036–29045.

Wahyudi, H., Reynolds, A. A., Li, Y., Owen, S. C., & Yu, S. M. (2016). Targeting collagen for diagnostic imaging and therapeutic delivery. *Journal of Controlled Release: Official Journal of the Controlled Release Society*, *240*, 323–331.

Wakankar, A. A., Feeney, M. B., Rivera, J., Chen, Y., Kim, M., Sharma, V. K., & Wang, Y. J. (2010). Physicochemical Stability of the Antibody−Drug Conjugate Trastuzumab-DM1: Changes due to Modification and Conjugation Processes. *Bioconjugate Chemistry*, *21*(9), 1588–1595.

Waldmann, H. (2019). Human Monoclonal Antibodies: The Benefits of Humanization. In M. Steinitz (Ed.), *Human Monoclonal Antibodies: Methods and Protocols* (pp. 1–10). Springer.

Wang, H., Wang, L., Zhong, B., & Dai, Z. (2022). Protein Splicing of Inteins: A Powerful Tool in Synthetic Biology. *Frontiers in Bioengineering and Biotechnology*, *10*.

Wang, R., Yang, Z., Luo, J., Hsing, I.-M., & Sun, F. (2017). B12-dependent photoresponsive protein hydrogels for controlled stem cell/protein release. *Proceedings of the National Academy of Sciences*, *114*(23), 5912–5917.

Wang, Y., van Oosterwijk, N., Ali, A. M., Adawy, A., Anindya, A. L., Dömling, A. S. S., & Groves, M. R. (2017). A Systematic Protein Refolding Screen Method using the DGR Approach Reveals that Time and Secondary TSA are Essential Variables. *Scientific Reports*, *7*(1), 9355.

Wang, Z., Collighan, R. J., Pytel, K., Rathbone, D. L., Li, X., & Griffin, M. (2012). Characterization of heparin-binding site of tissue transglutaminase: Its importance in cell surface targeting, matrix deposition, and cell signaling. *The Journal of Biological Chemistry*, *287*(16), 13063–13083.

Wang, Z., & Griffin, M. (2013). The role of TG2 in regulating S100A4-mediated mammary tumour cell migration. *PloS One*, *8*(3), e57017.

Wang, Z., Stuckey, D. J., Murdoch, C. E., Camelliti, P., Lip, G. Y. H., & Griffin, M. (2018). Cardiac fibrosis can be attenuated by blocking the activity of transglutaminase 2 using a selective small-molecule inhibitor. *Cell Death & Disease*, *9*(6), 613.

Weeks, A. M., & Wells, J. A. (2020). Subtiligase-Catalyzed Peptide Ligation. *Chemical Reviews*, *120*(6), 3127–3160.

Wu, Z. P., & Hilvert, D. (1989). Conversion of a protease into an acyl transferase: Selenolsubtilisin. *Journal of the American Chemical Society*, *111*(12), 4513–4514.

Yang, R., Wong, Y. H., Nguyen, G. K. T., Tam, J. P., Lescar, J., & Wu, B. (2017). Engineering a Catalytically Efficient Recombinant Protein Ligase. *Journal of the American Chemical Society*, *139*(15), 5351–5358.

Yang, B., Tang, S., Ma, C., Li, S.-T., Shao, G.-C., Dang, B., DeGrado, W. F., Dong, M.-Q., Wang, P. G., Ding, S., & Wang, L. (2017). Spontaneous and specific chemical cross-linking in live cells to capture and identify protein interactions. *Nature Communications*, *8*(1), Article 1.

Yang, X., Miao, H., Xiao, R., Wang, L., Zhao, Y., Wu, Q., Ji, Y., Du, J., Qin, H., & Xuan, W. (2021). Diverse protein manipulations with genetically encoded glutamic acid benzyl ester. *Chemical Science*, *12*(28), 9778–9785.

Yasueda, H., Kumazawa, Y., & Motoki, M. (1994). Purification and characterization of a tissue-type transglutaminase from red sea bream (Pagrus major). *Bioscience, Biotechnology, and Biochemistry*, *58*(11), 2041–2045.

Yu, Y.-J., Wu, S.-C., Chan, H.-H., Chen, Y.-C., Chen, Z.-Y., & Yang, M.-T. (2008). Overproduction of soluble recombinant transglutaminase from Streptomyces netropsis in Escherichia coli. *Applied Microbiology and Biotechnology*, *81*(3), 523–532.

Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., & Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proceedings of the National Academy of Sciences*, *109*(12), E690–E697.

Zemskov, E. A., Loukinova, E., Mikhailenko, I., Coleman, R. A., Strickland, D. K., & Belkin, A. M. (2009). Regulation of Platelet-derived Growth Factor Receptor Function by Integrin-associated Cell Surface Transglutaminase *. *Journal of Biological Chemistry*, *284*(24), 16693–16703.

Zhang, D., Wang, M., Wu, J., Cui, L., Du, G., & Chen, J. (2008). Two Different Proteases from Streptomyces hygroscopicus Are Involved in Transglutaminase Activation. *Journal of Agricultural and Food Chemistry*, *56*(21), 10261–10264.

Zhang, X., Chen, X., Hong, H., Hu, R., Liu, J., & Liu, C. (2021). Decellularized extracellular matrix scaffolds: Recent trends and emerging strategies in tissue engineering. *Bioactive Materials*, *10*, 15–31.

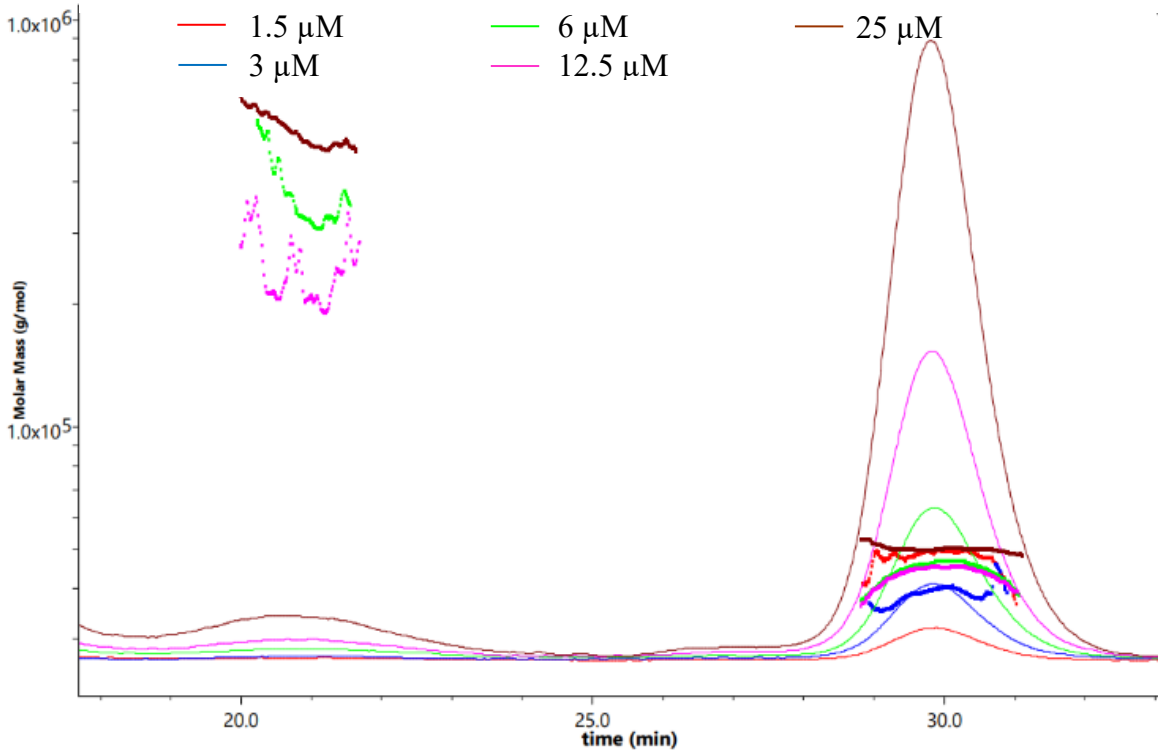
Zhao, L., Li, X., Zhao, J., Ma, S., Ma, X., Fan, D., Zhu, C., & Liu, Y. (2016). A novel smart injectable hydrogel prepared by microbial transglutaminase and human-like collagen: Its characterization and biocompatibility. *Materials Science and Engineering: C*, *68*, 317–326.
Zhao, Q., Wang, Y., Cui, H., & Du, X. (2019). Bio-inspired sensing and actuating materials. *Journal of Materials Chemistry C*, *7*(22), 6493–6511.

Zhao, W., Chen, B., Li, X., Lin, H., Sun, W., Zhao, Y., Wang, B., Zhao, Y., Han, Q., & Dai, J. (2007). Vascularization and cellularization of collagen scaffolds incorporated with two different collagen-targeting human basic fibroblast growth factors. *Journal of Biomedical Materials Research. Part A*, *82*(3), 630–636.

Zhou, H., Xu, J., Liu, X., Zhang, H., Wang, D., Chen, Z., Zhang, D., & Fan, T. (2018). Bio-Inspired Photonic Materials: Prototypes and Structural Effect Designs for Applications in Solar Energy Manipulation. *Advanced Functional Materials*, *28*(24), 1705309.

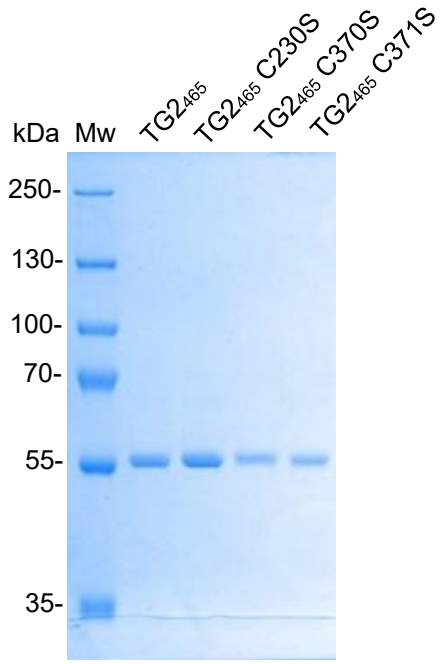
Zong, Y., Bice, T. W., Ton-That, H., Schneewind, O., & Narayana, S. V. L. (2004). Crystal structures of Staphylococcus aureus sortase A and its substrate complex. *The Journal of Biological Chemistry*, *279*(30), 31383–31389.

Appendix

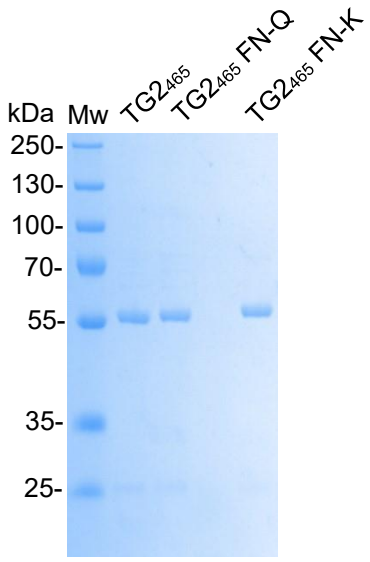


Sample	Molecular weight (kDa)	Uncertainty (%)	Mass fraction (%)
1.5 µM	48.2	3.8	100
3 µM	39.1	12.8	100
6 µM	45.1	5.3	99.4
12.5 µM	44.0	4.8	97.5
25 µM	50.3	0.4	99.1

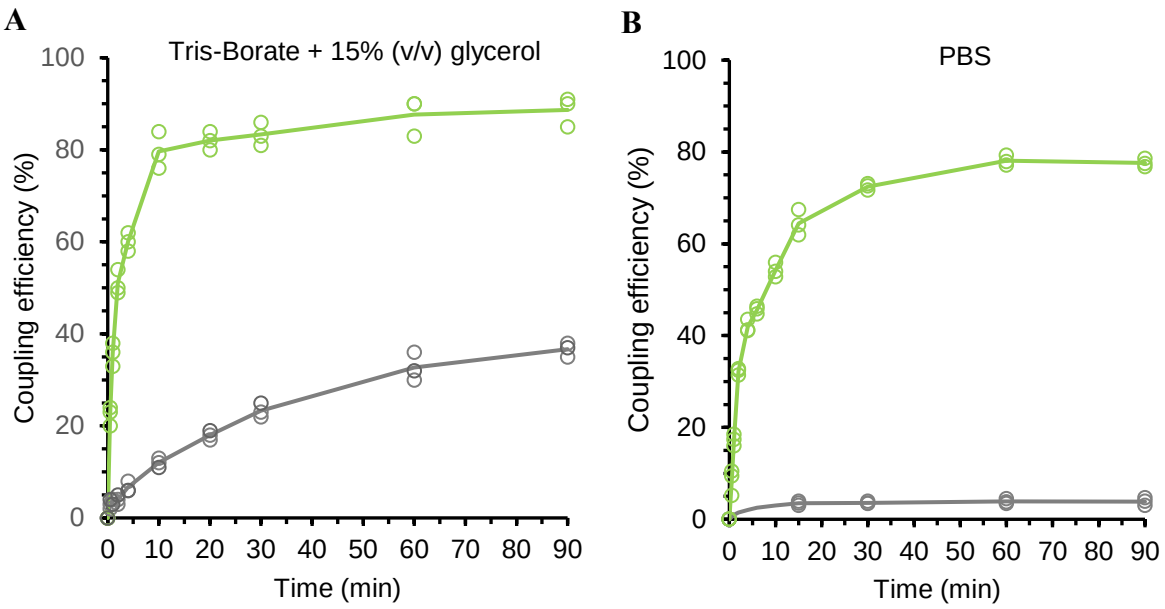
Appendix Figure 1. SEC-MALS of TG2$_{465}$. Running buffer of HBS pH 7.4. The 3 µM sample has high uncertainty so is discounted. The remaining samples have an average molecular weight of 49 kDa, slightly below the expected molecular weight of 54 kDa. This is explained by the elongated shape of TG2$_{465}$.



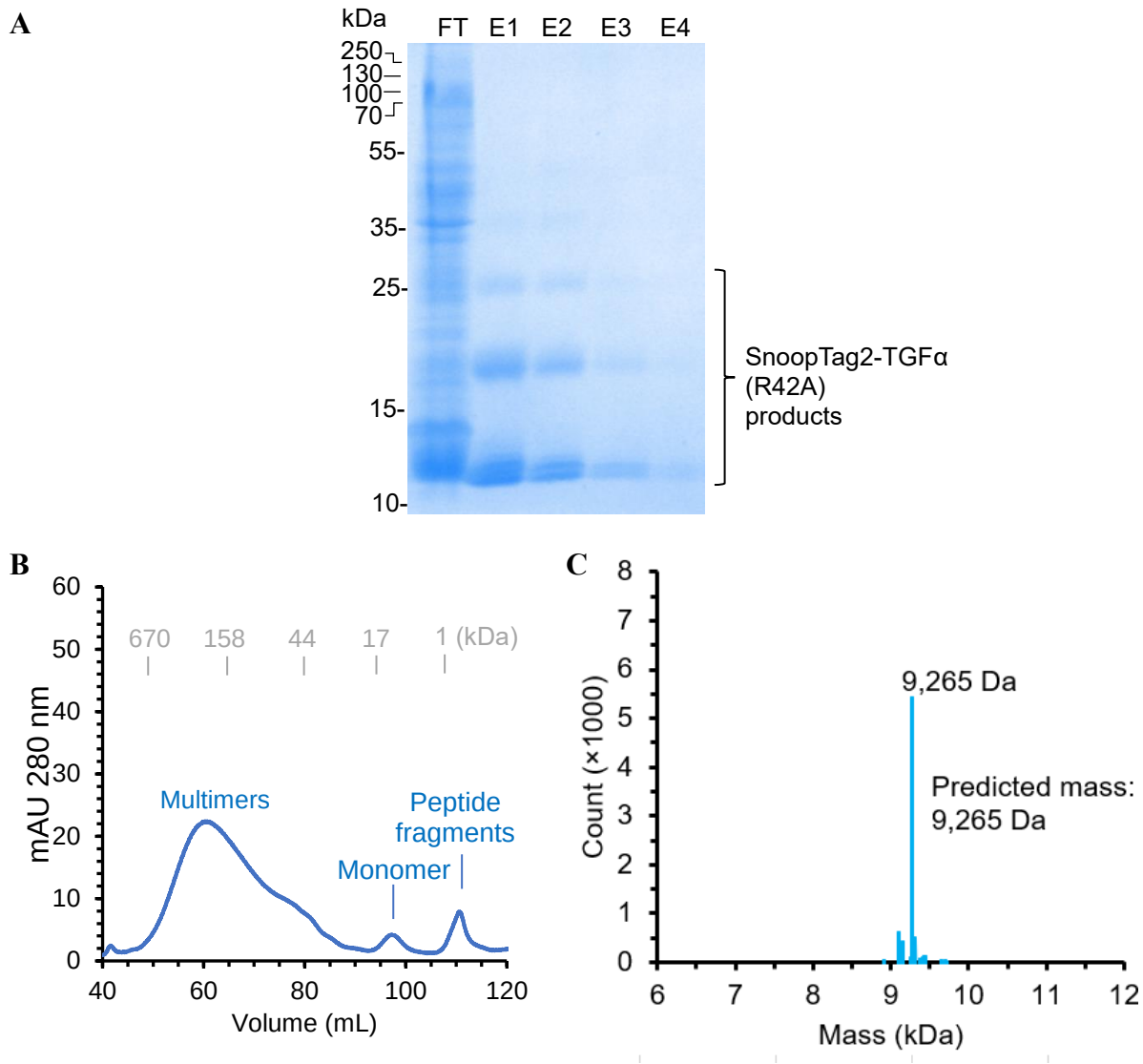
Appendix Figure 2. $TG2_{465}$ redox mutant preparations used for cadaverine assay. 8% SDS-PAGE with Coomassie staining, samples were reduced with 100 mM DTT before analysis. Mw = molecular weight markers.



Appendix Figure 3. Tagged $TG2_{465}$ preparations used for cadaverine assay. 10% SDS-PAGE with Coomassie staining, samples were reduced with 100 mM DTT before analysis. Mw = molecular weight markers.



Appendix Figure 4. Comparison of reactivity for first and second generation SnoopLigase sets. (A) SnoopLigase2 enhances reaction in Tris-borate. 25 μM each of SnoopTagJr/2-sfGFP, SUMO-DogTag(2) and SnoopLigase(2) were incubated for varying times at 21 °C in Tris-borate pH 7.4 with 15% (v/v) glycerol, before SDS−PAGE/Coomassie. Circles show each of the triplicate data points. (B) SnoopLigase2 is functional in the commonly used PBS. As in (A) in PBS pH 7.4. Coupling efficiency is measured as ratio of ligated product to both unreacted substrates. Some data points are overlapping.



Appendix Figure 5. Production of SnoopTag2-TGFα (R42A). (A) Ni-NTA purification of SnoopTag2-TGFα (R42A). FT = flow through, E = elution. Column washes performed with 10 mM imidazole, elution performed with 200 mM imidazole. Analysis by 18% SDS-PAGE with Coomassie staining. Samples were non-reduced. One prep performed. (B) Size-exclusion chromatography of SnoopTag2-TGFα (R42A) after denaturing and refolding shows several peaks of soluble products. (C)) Electrospray ionization mass spectrometry detects a 9265 Da product corresponding to SnoopTag2-TGFα R42A with 3 disulfide bonds