Exploring the Interaction Between Fibronectin and Transglutaminase II

Thomas Grant

A thesis submitted to
the University of Ottawa
in partial fulfillment of the requirements of the
Degree Master of Science Chemistry

Department of Chemistry and Biomolecular Sciences
Faculty of Science
University of Ottawa

Abstract

Human transglutaminase 2 (TG2) is a versatile transamidating acyltransferase that is ubiquitously expressed in the human body. It catalyzes a host of functions ranging from crosslinking of glutamine and lysine residues via its transamidating activity in its calcium dependent "open" conformation to acting as a G-protein involved in signal transduction of G-protein Coupled Receptors (GPCRs) in its "closed" conformation.[1] Another intriguing function of TG2 is its role as a cell surface protein where it interacts with a variety of extracellular matrix (ECM) proteins, forming a ternary complex between cell surface receptor β-integrins and the structural protein fibronectin (FN), an interaction important to FN deposition and fibrillogenesis in the ECM.[2] Cell surface TG2 and its interaction with FN has been shown to play a role in tumour cell resistance to chemotherapeutics[3] and increased adhesion of ovarian tumour cells to the ECM.[4] Our objective was to contribute to the elucidation of the non-covalent interaction between TG2 and the 45 kDa Fibronectin gelatin binding domain (45FN) by analyzing potentially key residues on TG2 through alanine site directed mutagenesis (SDM) and Bio-Layer Interferometry (BLI). Additionally, we used Genetic Code Expansion (GCE) to incorporate the UV-photoactivable crosslinking unnatural amino acid (UAA) Azido-L-Phenylalanine (AzF) at these residues. We corroborated that R116 is an important residue for the interaction between 45FN and TG2, based on a three-fold increase in K_D when R116 was mutated to alanine, but we were unable to confirm results that support K30 as being crucial on its own.[5] A crosslinked complex was formed between 45FN and TG2 with AzF incorporated at position F203 using the pULTRA-pCNF orthogonal tRNA/aaRS system. However, more optimization is required for this technique to be viable for analysis of the paired residues via crosslinking MS (XL-MS), due to the low concentrations of crosslinked complex that was formed.

Statement of Work

Dr. Pauline Navals performed *in silico* modeling of the alternative binding pose between 45FN and TG2 and was the lead developer of the rationale behind this study.

Maryam Kettal performed site-directed mutagenesis of F230X and A66X mutants. She also expressed and tested the characteristics of these mutants, this data is included in her Honours thesis, which is cited.

Matt O' Neil expressed F203X TG2 alongside NS3 helicase as shown in Figure 74.

The western blot procedure was followed from the procedure described in the "Resource" section of the Boddy Lab website.

Acknowledgements

I would like to foremost express my gratitude to Dr. Jeffrey Keillor for his patience and encouragement throughout my degree, I was very lucky to have you as my supervisor, and your mentorship helped me in moments of discouragement. Additionally, thank you to the Keillor group for all the laughs, and camaraderie which were a bright part of my time here whether that be playing pool, watching movies or just chatting in the Keillor corner. I will miss our (daily) coffee runs. All the Keillor members have been awesome to work with past and present. Pauline, thank you for getting me started and for your encouragement. Brianna thank you for recommending the best vampire book series out there, it was a blast reading them alongside you! Francis, from day one you have been so helpful and very kind, I really enjoyed working with you in the Biobay by day and defeating Nightlords by night. Sarah, thank you for your uplifting presence, and fun conversations. Alana, thank you for sharing your biochemistry, and music theory knowledge with me, performing with you will always be a nice memory. Lavleen, I really enjoyed our philosophical discussions, and all the laughs as well. Thank you for setting an example of a scientist to be emulated. Ronel and Taylor, best of luck with your studies it was great getting to know you in the short time we had. Daniel, thanks for your friendship and for taking me up on my book recommendations.

Eryn and Matt, thank you very much for helping me with GCE and AzF troubleshooting! I would like also to extend a broad thank you to Marion 2 for all the help with borrowing equipment and reagents. Thanks Tony and Luke for the entertainment in the Biobay. Tony, thanks for the laughs and friendship, gaming with you and Francis helped me out a lot during busy times.

Finally, thank you to my Dad, Nancy, Mom and Camden for your support.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Akt	Protein kinase B
AzF	Azido-L-Phenylalanine
BLI	Bio-Layer Interferometry
DAG	Diacylglycerol
ECM	Extracellular Matrix
EOC	Epithelial Ovarian Cancer
FN	Fibronectin
FRET	Förster Resonance Energy Transfer
GBD	Gelatin binding domain
GCE	Genetic Code Expansion
GPCR	G-protein coupled receptor
HDX-MS	Hydrogen-deuterium mass spectroscopy
HSPGs	Heparin sulphate proteoglycans
hTG2	Human transglutaminase 2
IP_3	inositol (1, 4, 5)-triphosphate
LTBP	Latent TGF-β binding protein
MOE	Molecular operating environment software
PGs	Proteoglycans
PPI	Protein-protein interaction
SDM	Site-directed mutagenesis
SPR	Surface plasmon resonance
TG2	Transglutaminase 2
XL-MS	Crosslinking mass spectroscopy
45FN	45 kDa fibronectin fragment

Chapter 1: Introduction

1.1 Transglutaminase II

Transglutaminases are expressed in a wide number of tissues throughout the human body. They are a family of proteins with diverse functionalities that are still under investigation today, almost six decades since they were first discovered thanks to their transamidating activity in guinea pig liver.[6] There are currently nine transglutaminases identified in the human proteome. Eight of them share a common cysteine-thiol active site, require Ca^{2+} in their activation and catalyze acyl transfer reactions between the γ-carboxamide group of peptide bound glutamine and the ε-amino group of peptide bound lysine forming a ε-(γ-glutamyl)lysine isopeptide bond.[6] Besides these similarities, transglutaminases use their catalytic activity in very different ways to carry out their roles in the body. Factor XIII is a protransglutaminase which is activated by thrombin; it then uses its crosslinking activity to stabilize the fibrin clot, important in wound healing, against environmental damage.[7,8] Transglutaminase 1 is expressed in keratinocytes and plays a crucial role in the late stage of terminal differentiation, it forms the cellular envelope of a highly insoluble body beneath the cell membrane by crosslinking precursor proteins.[9] Transglutaminase 3 is also involved in the development of the cell envelope in the epidermis.[10] Additionally, it is involved in hair follicle development by hardening the inner root sheath through crosslinking protein filaments.[11] These functions represent a fraction of the roles transglutaminases carry, which illustrates the functional diversity of these crosslinking enzymes. Transglutaminase 2 (TG2), specifically human transglutaminase 2 (hTG2), is the subject of our research. It is the most ubiquitously expressed in the family and alongside its primary

crosslinking ability, also functions as a G-protein and as a protein scaffold in the extracellular matrix (ECM) as illustrated in Figure 1.[12]

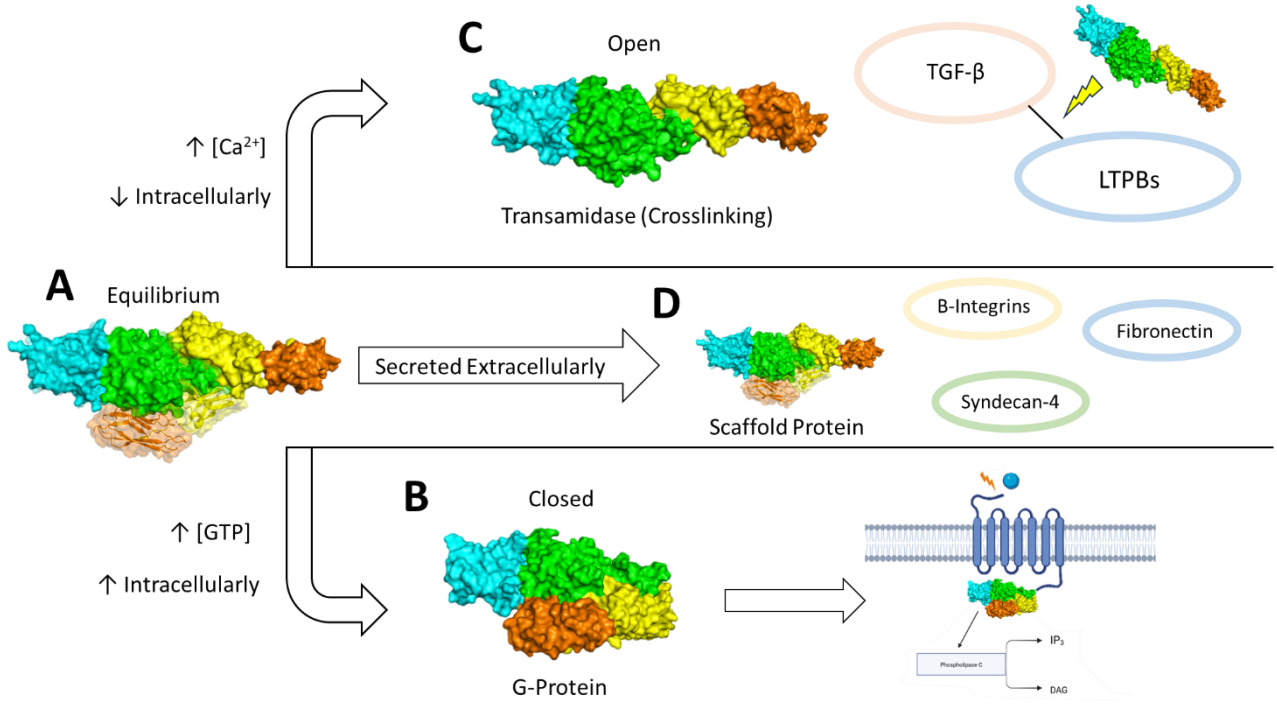


Figure 1. TG2's Diverse Functions. This figure shows a simple schematic of TG2's versatility. (A) TG2 exists in equilibrium between a closed (translucent) and open (opaque) state. This state is affected by effectors like Ca^{2+} and GDP. (B) High concentrations of GDP shift TG2 towards its closed conformation where it acts as a G-protein. (C) High Ca^{2+}concentrations bind allosterically and shift TG2 towards an open conformation where it functions as a transamidase. (D) When TG2 is secreted extracellularly it can behave as a scaffold protein, forming non-covalent interactions with ECM components.

TG2 has been well characterized since its discovery almost 70 years ago. TG2 was the first transglutaminase to be discovered. It was discovered by Heinrich Waelsh in 1957 by the analysis

of soluble protein fractions obtained from the homogenates of the livers of guinea pigs, rats, mice and rabbits. Proteins were found to be crosslinked together in the mixture only in the presence of Ca^{2+} which signalled that a novel type of enzyme was responsible for the crosslinking, rather than another class of enzymes exhibiting glutamo or asparto-transferase activities.[13] TG2 consists of four domains: an N-terminal β-sandwich domain (1-139), a catalytic core domain (140-454), and two C-terminal β-barrel domains (455-591 and 592-687),[14] as illustrated in Figure 1. It is a dynamic protein, existing in an equilibrium between an open and closed conformation. The equilibrium shifts towards open or closed based on which effectors are present in the environment. TG2 usually adopts an open conformation via allosteric activation by Ca^{2+} ions, activating its transamidase or crosslinking function. When Ca^{2+} concentration is low and GTP is present, TG2 primarily exists in its closed conformation and acts as a G protein.[15] TG2's domains work together to ensure its crosslinking and G-protein activity are mutually exclusive by regulating the exposure of the respective active sites to the environment.

As with most transglutaminases, TG2's ability to form crosslinks between proteins and to catalyse the transamidation of glutamine residues with free amines is a staple feature of its functional repertoire. Additionally, TG2 can deamidate glutamine residues converting the primary amide to a carboxylic acid. Inside the cell, TG2 is thought to mainly exist in the closed conformation due to the low level of intracellular Ca^{2+}; this would limit TG2's transamidase activity. Despite this, TG2 still catalyzes intracellular transamidation reactions.[16] One explanation for this is that transient influxes of Ca^{2+} are sufficient to stabilize the open conformation of TG2 and activate its crosslinking activity, especially near the cell membrane.[17] TG2 also catalyzes transamidation reactions in the extracellular matrix where there is a high concentration of Ca^{2+}, but the protein usually exists in an inactive form regulated by the ECM's

oxidative environment through the redox-sensitive cysteine triad Cys^{230}, Cys^{370} and Cys^{371}.[18] Some examples of TG2's physiological crosslinking substrates include: β-crystallin[19], fibronectin[20], type III collagen[21], latent TGF-β binding protein (LTBP)[22] and cytoskeletal proteins such as actin, myosin, and ROCK2.[23] Its covalent modification reactions can inactivate enzymes such as α-ketoglutarate dehydrogenase and glyceraldehyde 3-phosphate dehydrogenase.[24] TG2's crosslinking of extracellular proteins is thought to increase their resistance to environmental factors and also facilitate cell adhesion and motility.[6] When cell-surface TG2 was inactivated in Swiss 3T3 fibroblasts by an anti-TG2 antibody, cell attachment was greatly reduced.[25] Intracellularly, TG2 can regulate processes using its crosslinking activity to apply post-translational modifications to proteins. An example of this is the serotonylation of small G proteins such as Rab4A and RhoA, which activates them, releasing α-granules, critical for the function of platelets.[26] TG2's transamidation ability gives it a versatile role both intracellularly and extracellularly.

TG2's secondary function is to bind and hydrolyze GTP and it was found to be the α-subunit of the G-protein G_h, also known as $G\alpha_h$ - a breakthrough in the field.[27] Over time, TG2 was found to serve as a G-protein for three G-Protein Coupled Receptors (GPCRs): the α_1 adrenergic receptor[28], TPα thromboxane A2 receptor[29] and the oxytocin receptor.[30] The binding site of GTP in TG2 is located in a cleft between the catalytic core and the first β-barrel, where positively charged residues and hydrophobic interactions stabilize the phosphate and guanosine moieties of GTP respectively in the active site.[31] TG2 participates in downstream signalling by activating phospholipase C, which leads to the production of inositol (1,4,5)-triphosphate (IP_3) and diacylglycerol (DAG). IP_3 can activate ion channels and trigger influx of Ca^{2+} into the cell, which regulates a plethora of cellular processes such as muscle contraction, cell proliferation and

of course activation of TG2 transamidation activity.[32] G-protein signalling represents the second major biological function of TG2, but TG2 also has non-enzymatic utility outside of the cell as a scaffold protein.

Scaffold proteins participate in cellular signalling by bringing together two or more proteins that function as signalling enzymes and coordinates their actions.[33] Although most TG2 is present intracellularly, depending on the cell type, 1-20% of TG2 is secreted extracellularly through non-classical mechanisms, possibly through phospholipid-dependent delivery into recycling endosomes.[16,34,35] Outside of the cell, TG2 is known to non-covalently interact with many different proteins including: heparan sulphate proteoglycans (HSPGs) such as syndecan-4, as well as β-integrins and fibronectin (FN).[36] TG2's interaction with HSPGs is thought to be crucial for its translocation to the cell surface.[37] TG2 exists as a partner in a ternary complex between FN and integrins, that participate in activation of integrin signalling pathways important for cell adhesion.[36] Out of all of the diverse activities of TG2, this research project is focused on its role as a scaffold protein in the ECM, specifically its non-covalent interaction with FN.[38]

1.2 The Extracellular Matrix and Fibronectin

The ECM represents non-cellular components that form a highly organized matrix between cells. It can be split into two denominations, the interstitial matrix being most of the space between cells and the pericellular matrix being the matrix adjacent to the cell membrane. All cell types synthesize and secrete macromolecules that make up the specific composition of the ECM, which varies from tissue to tissue. The ECM is home to large variety of macromolecules such as proteoglycans including previously mentioned HSPGs, collagens, laminins, proteases, fibronectin, matricellular proteins and cellular receptors such as integrins.[39] These macromolecules work together to carry out the ECM's functions such as signal transduction and providing structural support for cell proliferation.[40] For example, activation of T-lymphocytes is enhanced by integrin-mediated adhesion to fibronectin or laminin in the ECM.[41] The ECM undergoes constant remodelling, which is important for maintaining normal function. This remodelling is regulated by metalloproteinases and inhibitors of these proteinases.[42] A wide variety of macromolecules are involved in the homeostasis of the ECM.

There are two main classes of macromolecules in the ECM: proteoglycans (PGs) and fibrous proteins. PGs have a wide range of functions; from playing a role in glomerular filtration to generating structures essential for mechanical buffering, fibrous proteins provide structural support and resistance to tensile stresses among other functions.[42] PGs consist of a core protein with one or more glycosaminoglycan (heteropolysaccharide) chains attached. Decorin is a PG present in the ECM that carries one GAG chain. It binds non-covalently to collagen fibrils and supports collagen fibrillogenesis.[39] The lack of decorin in mice results in abnormal fibrillogenesis and skin fragility[43]. In addition, its structural role it has been shown to neutralize TGF-β through binding, which can either inhibit or promote cell growth, depending on the cell

type.[38] Another type of PG are cell surface associated proteins such as syndecans. Syndecans are part of the heparan sulphate proteoglycan family (HSPGs) which include syndecan-1/CD138, syndecan-2/fibroglycan, syndecan-3/N-syndecan and syndecan-4/amphiglycan.[39] Syndecans are single transmembrane proteins that link the actin cytoskeleton and the ECM.[44] Many ECM molecules contain a heparan sulfate binding site that syndecans non-covalently bind to. Syndecan-1 is thought to be a co-receptor with integrin adhesion to collagen[45] and syndecan-4 has been shown to interact with the HepII domain of Fibronectin, with possible regulation by the matricellular ECM protein tenascin-C resulting in cessation of cell cycle progression.[40] TG2 was found to bind heparin and HS in the low nanomolar range and its binding to syndecan-4 is strongly implicated in its RGD-independent cell adhesion process,[46] which will be expounded upon later in this introduction.

Integrins are not PGs, but like syndecans they are important cell-surface receptors for outside-in signal transduction from the ECM and are a partner of TG2 in its scaffolding capacity. Integrins are a heterodimer superfamily of twenty-four combinations of α and β transmembrane subunits, including eighteen types of α subunits and eight types of β subunits. The unique combinations of heterodimers lend specificity for different ECM ligands.[47] The specificity can be grouped into categories like the RGD motif where the ECM ligand binds in an interface between α and β subunits.[48] Integrins use this specificity to regulate fundamental ECM organization such as fibronectin fibrillogenesis, where integrins bind to fibronectin's RGD triad to promote clustering and polymerization.[49] TG2 non-covalently interacts with β1, β3 and β5 integrin subunits and forms ternary complexes with integrin and fibronectin.[34,50] It was found that all TG2 on the surface of human erythroleukemia cells and THP-1 macrophages were bound to the β subunits of integrins.[50,51] These transmembrane cell surface receptors are key in linking

intracellular to extracellular structure, which is composed of not only PGs but also fibrous proteins.

Collagen and fibronectin are two of the most abundant fibrous proteins in the ECM. Collagen is the most abundant fibrous protein; there are twenty-eight different types composed from various polypeptide chains.[39] Residues on collagens are hydroxylated and are secreted into the ECM where their modified peptides are removed, allowing collagen to assemble into fibrils that make up major tissue characteristics such as shape and organization.[40] These fibrils are heterogeneous, made of several types of collagens and also non-collagen proteins, their composition is dependent on developmental stage and tissue type. Collagen can also act as a ligand for cell surface receptors such as integrins.[52] Fibronectin is important for attachment and migration of cells[40] and interacts with a wide variety of proteins including; integrins, syndecans, collagen, fibrin[39] and TG2.[53] This research is focused on the elucidation of the non-covalent interaction of TG2 and fibronectin; therefore, the remainder of this introduction will be centred on fibronectin as well as its interaction with TG2 in physiological and pathogenic contexts.

FN was first discovered in 1948 in plasma and was named cold-insoluble globulin, it was rediscovered in the 1970s as a cell-surface marker for some tumour cells, it was absent in the oncogenic cells. This discovery led to an explosion of research into fibronectin which reduced its impact as an unambiguous marker for cancer, but led to the discovery of integrins, and the role of the ECM in outside-in signalling.[54] The core structure of FN is made of the dimerization of two 250-kDa proteins connected by disulfide bonds in the C-terminus of the peptide. They exist as glycoproteins made of Type I, Type II and Type III repeating subunits.[55] The structure of fibronectin is variable, thanks to mixing of exons by alternative splicing. Fibronectin exists in two major forms. The soluble form is synthesized from liver hepatocytes and exists in plasma; it

is incorporated into the ECM by cell dependent processes. The insoluble form is synthesized by a wide range of cell types and exists in the ECM. It is functional in its fibrillar state, and more heterogeneous due to increased cell type and species specific splicing.[39] FN can be enzymatically digested and research into its domains has discovered binding sites to key ECM proteins. There are many sites for FN adhesion to integrins, one of the major sites being the RGD site previously mentioned. Syndecan binding to FN's HepII domain enhances integrin mediated cell spreading and intracellular signalling.[56] FN also has binding domains for collagen, TG2 and even bacteria such as *Streptococcus pyogenes* and *Staphylococcus aureus*.[57,58] These domains contribute to FN's function as a regulatory protein, in addition to its structural properties for the ECM.

FN is a significant molecule of interest in the ECM both for being the foundational molecule that ignited interest in ECM proteins[54] as well as its plethora of functions, including ECM morphogenesis, and cellular adhesion primarily with integrins. FN is functional in the ECM in its insoluble fibrillar form, which is held together by non-covalent bonds between fibrils.[57] FN self-association is supported through intracellular factors, such as the actin cytoskeleton, via integrins[59] and soluble FN associates initially at focal adhesions.[60] The fibrillogenesis of FN is hypothesized to be supported by mechanical force at these focal adhesions. This hypothesis was supported by experiments where soluble FN molecules were more adhesive for stretched FN fibres, most likely due to exposing the cryptic self-association site on module $FnIII_1$.[57] The assembly of FN has been shown to be necessary for the deposition of subsequent matrix structures like collagen.[61,62] Fibronectin adheres to cells via integrins mainly by its RGD domain, located on $FnIII_{10}$. This domain binds to $\alpha_5\beta_1$, $\alpha_3\beta_1$, $\alpha_8\beta_1$, and $\alpha_v\beta_3$ integrins.[55] An adjacent synergy site PHSRN on $FnIII_9$ is crucial for $\alpha_5\beta_1$ integrin binding.[63] Mechanoregulation of these synergy sites through stretching of FN fibres led to decreased cell

spreading and migration, illustrating FN's importance for cell adhesion.[64] Fibronectin is known to strongly non-covalently associate with TG2 and this interaction is the focus of our research.

TG2 is known to strongly associate to FN,[53] particularly through TG2's N-terminal domain.[65,66] This interaction is important for FN's deposition and fibrillogenesis. It can form a ternary complex with FN and β-integrins as well as bridge FN and syndecan-4, a pathway promoting cell survival during the loss of RGD dependent adhesion to integrins.[34] TG2 increases binding of FN to the cell surface through a ternary complex with $\alpha_5\beta_1$ integrin.[2] TG2 interacts with FN at its gelatin binding domain, consisting of modules $I_6II_{1,2}I_{7\text{-}9}$, which does not overlap with FN's integrin binding modules.[65] This ternary complex has been shown to increase the size of focal adhesions and amplify integrin signalling[50], promote fibronectin matrix assembly[2], and cell migration into tissues containing FN matrices.[51] Another key physiological role for the extracellular TG2 and FN interaction mediated by syndecans is RGD-independent cell adhesion.

The RGD domain on FN has been previously mentioned and is the most important cell surface recognition site for a large portion of integrins, including $\alpha_5\beta_1$, and $\alpha_v\beta_3$ integrin, major integrins for FN cell adhesion.[67] The cell adhesion process can be disrupted by competitive inhibition by RGD peptides, which are released during matrix remodelling of the ECM, an important process for wound healing and angiogenesis.[68] Loss of cell adhesion can lead to anoikis, a form of apoptosis that occurs when anchorage-dependent cells lose their connection to the ECM. TG2 externalization is upregulated during stress and injury conditions and can compensate for the loss of FN cell adhesion through a syndecan-4 and β-1 integrin co-signalling pathway.[69] The interaction between matrix FN-bound TG2 and syndecan-4 leads to activation of the intracellular protein kinase Cα (PKCα) which, through inside-out signalling, binds to the intracellular tail of β1 integrins which promotes focal adhesion formation. Furthermore, PKCα

activates syndecan-2 mediating cytoskeleton formation and cell contraction through the ROCK pathway[67], important for fibronectin fibril deposition[70] and ultimately increased cell survivability. TG2 and FN work together to maintain homeostasis under normal conditions but their interaction can be detrimental in certain pathologies.

1.3 TG2 and Fibronectin in Human Diseases

TG2 is implicated in a wide range of diseases involving both its transamidation and protein scaffolding abilities. TG2's transamidation activity is implicated in celiac disease, it's ability to deamidate gluten derived gliadin residues greatly increases the affinity of these peptides for disease pre-disposing proteins HLA-DQ2 and DQ8. These cell surface proteins display these peptides to immune cells, resulting in an immune response that damages the small intestinal lining in celiac patients.[71] Its role as a transamidating protein in the extracellular matrix is involved in fibrosis pathogenesis. Fibrosis is defined as an abundance of activated fibroblasts and excessive deposition of the ECM, this process is in part driven by TGF-β.[72] ECM proteins such as PGs, collagen and fibronectin act as reservoirs of TGF-β and TG2 regulates its release by crosslinking it to latent TGF- β binding proteins (LTBPs) on the ECM.[3] In addition to fibrosis, TG2 pathology in the ECM is closely related to its scaffold activity and relationship with FN such as in tumour chemotherapeutic resistance and ovarian cancer adhesion to the peritoneal lining.

Tumour cells alter expression of proteins that interact with the ECM to provide signaling important for progression providing resistance to chemotherapeutics. In meningiomas[73] and glioblastomas[74], the expression of extracellular TG2 and fibronectin was increased. The activity of TG2 was increased in both tumour types, by 10-fold in the case of meningiomas compared to normal brain cells. High levels of TG2 expression has been associated with increased chemotherapeutic resistance in breast cancer cells[75] and treating breast and lung cancer cells with FN showed improved survival after a single dose of radiation.[76] When the non-covalent TG2 inhibitor KCC009 was applied to meningioma cell tissue explants and IOMM-Lee cells, the cells failed to assemble dense strands of FN. When IOMM-Lee cells were treated with KCC009 and

15 Gy of radiation, morphological changes consistent with cell death were seen. No visible changes were seen in the absence of TG2 inhibitor and treatment with 15 Gy radiation.[73] Researchers hypothesize that inhibition both of TG2's non-covalent interaction and crosslinking with FN is to blame for the sensitization of these cancer cells to radiation. This hypothesis is supported by TG2's known function as a co-receptor for FN and integrins which leads to outside-in signalling activating anti-apoptotic signalling factors like protein kinase B (Akt).[77] This hypothesis is further supported by evidence such as the integrin α_5 and FN interaction being required for B16 murine melanoma cell metastasis.[78] It is not clear if the loss of formation of fibronectin assembly via KCC009 inactivation of TG2 is caused strictly by interruption of the non-covalent interaction of extracellular TG2 to FN, as opposed to its crosslinking activity, the importance of TG2's scaffolding ability with FN is more prevalent in ovarian cancer.

Epithelial ovarian cancer (EOC) is partly characterized by a unique form of cell metastasis. Metastasis of EOC cells in the intraperitoneal space is facilitated by the EOC cells being directly exposed to the peritoneal fluid and able to freely disseminate as well as adopting a mesenchymal phenotype in response which promotes this dissemination. Dissemination of EOC cells throughout the intraperitoneal space results in adhering and growth as metastatic implants.[4] Knockdown of TG2 in EOC SKOV3 cells resulted in a decrease in EOC adhesion to FN by greater than 50%. Knockdown of TG2 also resulted in a decrease in fibronectin and collagen stimulated chemotaxis. Stable overexpression of TG2 increased adhesion and fibronectin dependent migration of ovarian cancer cells (OV90). The researchers hypothesize that TG2 achieves these effects through its role as a co-receptor with β-integrin and FN. This is supported by observation of diminished expression of β1 integrins on the cell surface when TG2 is downregulated. Clearly, the TG2 and FN interaction is an important part of pathological

processes such as cancer metastasis. The elucidation of this interaction, the goal of this research, could provide novel structural information on the protein-protein interaction between TG2 and FN enabling more informed rational design of inhibitors for this interaction.

1.4 Investigating the TG2 and Fibronectin Interaction

Investigations into the FN and TG2 protein-protein interaction have yielded knowledge on the domains involved and potentially important residues but the precise nature of the interaction is still debated. FN's 45 kDa gelatin domain[65] is known to be the site at which TG2's N-terminal domain non-covalently interacts.[79] The first insight into key residues of TG2 involved in this interaction comes from an analysis of peptide residues 1-7 on the N-terminal domain.[66] Although this research aligns with the N-terminal domain of TG2 interacting with FN it does not consider secondary structure of peptides. Two in-depth studies have investigated the interaction using site-directed mutagenesis and deuterium exchange MS studies[79] as well as *in silico* modelling[80] guided from results of the SDM and deuterium exchange MS. Our lab previously performed *in silico* modelling that corroborated residues previously found to be important as well as identifying a few more residues that could be implicated in FN binding. We aimed to confirm or refute these hypotheses through molecular biology techniques.

The two techniques we used to probe the TG2 FN interaction were Bio-layer Interferometry (BLI) in conjunction with site-directed mutagenesis (SDM) as well as Genetic Code Expansion (GCE) specifically site-specific incorporation of the crosslinking unnatural amino acid (UAA) Azido-L-Phenylalanine (AzF). BLI alongside SDM provides information on how the association of the TG2-FN complex is affected by mutation of key residues. These residues were also targeted for AzF incorporation with the hope of capturing a crosslinked "snapshot" of TG2 and FN in complex through identification of paired residues through digestion and mass spectroscopy analysis.

Chapter 2 will present the SDM experiments and BLI analysis of key residues replaced with alanine. Analysis of single alanine mutations key residues identified by Cardoso *et al.*[5,79] and our

docking studies showed no significant change in association of the 45 kDa FN fragment to TG2 except for the R116A mutation which showed a threefold increase in K_D when measured with BLI. This is consistent with the previous findings that show the mutation of R116A as important for FN association via surface plasmon resonance (SPR) experiments.

Chapter 3 will show our attempts at incorporation of AzF into residues believed to be important for the FN-TG2 interaction. AzF incorporation was attempted with three different expression systems with the yield of mutant protein and incorporation efficiency of AzF increasing. FN-TG2 crosslinking complexes were formed for F203 AzF mutant TG2 and the concentration of this crosslinked complex was increased through use of a different expression system. The crosslinking product was not abundant enough to be digested and analyzed via mass spectroscopy.

Although the precise elucidation of the TG2-FN complex was not achieved, R116A was confirmed to be important for FN-TG2 binding, and the TG2-FN complex was captured using incorporation of the crosslinking UAA AzF. Future research should analyze the H134A among other alanine mutated residues using BLI. Further optimization of the AzF crosslinking assay, expression of protein, or residue chosen for AzF incorporation may increase the yield of crosslinked TG2-FN enough for analysis of paired residues and clear elucidation of the FN TG2 interaction via digestion and mass spectroscopy analysis. This information could be used to guide rational design of protein-protein interaction inhibitors which could reduce the detrimental impacts of FN-TG2 in pathological conditions.

Chapter 2: Exploration of Fibronectin Association to Transglutaminase 2 Using Bio-layer Interferometry

2.1 Introduction

2.1.1 Current Investigations into TG2 and FN Association

The investigations into the nature of the TG2 and FN association began over 35 years ago and much has been uncovered about the minimum domain of FN required to interact with TG2 and the residues that are critical for the interaction between FN and TG2. The 45-kDa gelatin binding domain (GBD) of FN (45FN) ($FNI_6FNII_{1\text{-}2}FNI_{7\text{-}9}$) was identified, via limited chymotrypsin proteolysis, to be the portion of FN that binds to TG2.[53] ELISA studies have shown that this gelatin binding region by itself has the same affinity for erythrocyte TG2 compared to full length FN,[65] although this has been recently challenged.[81] The location on FN essential for TG2 interaction was narrowed down even further with ELISA assays, where the I_8 domain was found to be sufficient for *in vitro* binding to TG2. However, this domain did not result in functional effects when exposed to TG2 expressing cells, while domains $I_{7\text{-}9}$ were sufficient for inducing the same functional effects as the entire 45FN, such as signalling, adhesion, spreading and migration.[80] While the TG2-binding domain on FN has been narrowed down, the precise residues of TG2 that interact with FN are still disputed.

The early studies of TG2's binding site to fibronectin indicated the N-terminal domain of TG2 as the crucial binding site for the gelatin binding domain of FN; however, the pose and

residues involved in this binding are still up for debate. The first examination[66] identified residues 1-7 of the TG2 N-terminal domain as important for binding, through overlaying of FN with denatured TG2 fragments, visualized through immunoblotting. Digestion of the 28-kDa N-terminal domain fragment of TG2 to a 27-kDa fragment lacking the first seven residues abolished binding of full-length FN to the denatured fragments. Furthermore, in an additional study, residues D94 and D97 as well as the β-hairpin loop on residues 88-106 were proposed to be crucial for FN binding.[82] The peptide sequence 88-106 was proposed to be necessary for the binding of FN through affinity chromatography with immobilized TG2 as well as through competition assays of the synthetic peptide sequence competing with TG2 for binding to 45FN. The peptide sequence inhibited binding of TG2 to the GBD of FN by 80%. Additionally, using alanine mutations of D94 and D97, the K_D of the association 45FN and TG2, as measured by SPR, was greatly increased. Although these experiments all shared a common thread, that the N-terminal domain of TG2 is crucial for FN binding, the observations found by these experiments have not yet been replicated.

Cardoso *et al.* also confirmed that the N-terminal domain was crucial for FN binding and identified three residues as important for this reaction: K30, R116 and H134.[79] These residues were determined to be important via hydrogen-deuterium exchange mass spectroscopy (HDX-MS) and SDM experiments analyzed by SPR. HDX-MS measures the exchange of deuterium for hydrogen in amino acid side chains; this exchange can be reduced by a protein-protein interaction, like FN and TG2, by shielding the involved residues from deuterated solvent. HDX-MS experiments identified regions in both the N-terminal and core domain of TG2 where exchange was reduced. These regions in the N-terminus include amino acids 5-12, **41-58**, 59-81, **120-129** and **130-135**, with bolded regions displaying particularly reduced exchange. In the core

domain overlapping regions, residues 195-203, 195-219, 197-219 also showed reduced exchange. Notably, regions 89-100 and 101-114 showed no reduction in exchange, conflicting with the previous identification of the 88-106 region including D94A and D97A as necessary residues for binding to FN.[82] Cardoso *et al*. selected their key residues based on their HDX-MS results, as well as previous HDX-MS and SPR studies involving celiac disease antibodies[5] to analyze using SPR with immobilized 45FN. Selected mutations R116A/H134A, K30E/R116A/H134A, E120A, T58A and D198A all increased the K_D of TG2 with 45FN by 2000, 400, 12, 37 and 86-fold respectively, compared to WT TG2. It is worth noting the K_D was measured using SPR in the previous celiac disease epitope study[5] for residues K30E, H134A, and R116A, which resulted in an increase in K_D of 6, 9, and 25-fold respectively. Interestingly, measurement of the binding affinity of TG2 mutants D94A and D97A resulted in no increase in K_D. Overall, Cardoso *et al.* greatly contributed to the elucidation of the FN binding domain on TG2 by identifying regions and residues important for FN binding, while also challenging previous results.[82]

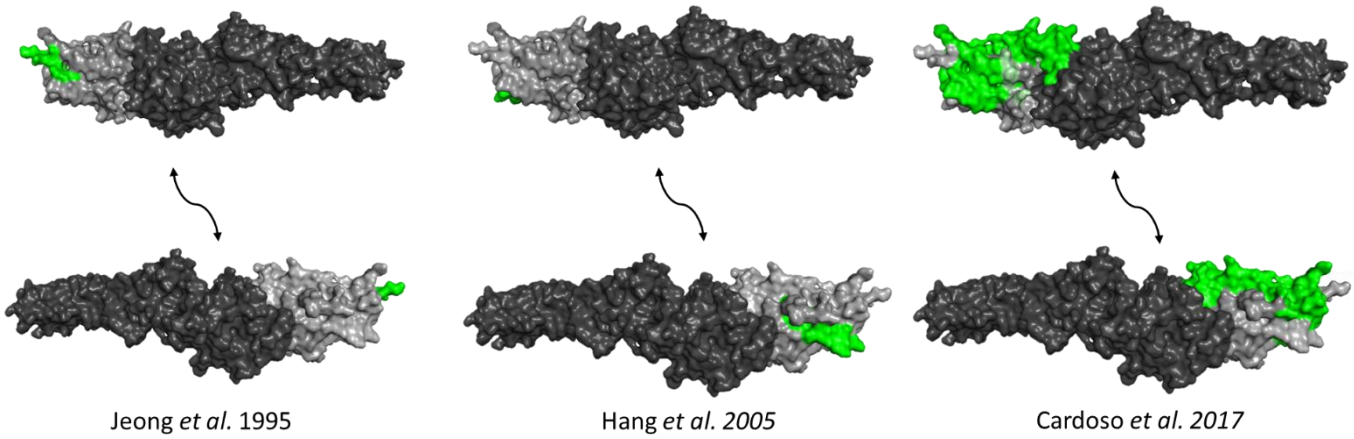


Figure 2. Proposed TG2 Residues Interacting with FN Over Time. Grey region, N-terminal domain of TG2. Green region, residues proposed to bind to FN or 45FN. Black region, core domain and c-terminal β-barrels of TG2. Jeong *et al.*[66] identified residues 1-7 as crucial for binding through immunoblotting of FN with denatured TG2 fragments. Hang *et al.*[82] identified residues 88-106 by competition assays of synthetic TG2 peptide fragments against full TG2 with 45FN. *Cardoso et al.*[79] identified multiple regions as interacting with 45FN using HDX-MS. Notably, each model refuted the other over time, there are little to no overlapping regions.

Soluri *et al.* utilized the new information provided by Cardoso *et al.* to guide *in silico* modelling of the minimum binding domains of FN to TG2.[80] In this same study, they mapped the minimum functional binding domain of FN to TG2 as being the $I_{7\text{-}9}$ subunit, as previously discussed. Soluri *et al.* conducted two docking simulations, one that was minimally guided by experimental data and one that was guided by the key residues identified by Cardoso *et al*. The first experiment used six docking softwares: ClusPro, GrammX, Patch Dock, pyDock, Swarm Dock, and ZDock. The FN domains I_8 and I_9 (PDB ID: 3GXE) were docked without bias onto

both open (PDB ID: 2Q3Z) and closed (PDB ID: 1KV3) TG2. The results were filtered using established experimental data such as the N-terminal domain of TG2 being the principal binding site. Additionally, models based on the open state that were structurally incompatible with the closed state of TG2 were ignored, since it is known that the binding of 45FN is independent of TG2 conformation. The results from this modelling were consistent with the residues implicated by Cardoso *et al.* The authors then ran another simulation using HADDOCK, driven by the seven residues shown by Cardoso *et al.* to significantly reduce the binding to FN (K30, T58, R116, E120, S129, H134 and D198). This resulted in a binding pose that formed four salt bridges, two with K30, one with R116 and one with E120. This study provided a convincing model consistent with experimental evidence of the binding mode of modules I_8 and I_9 of the GBD to TG2.

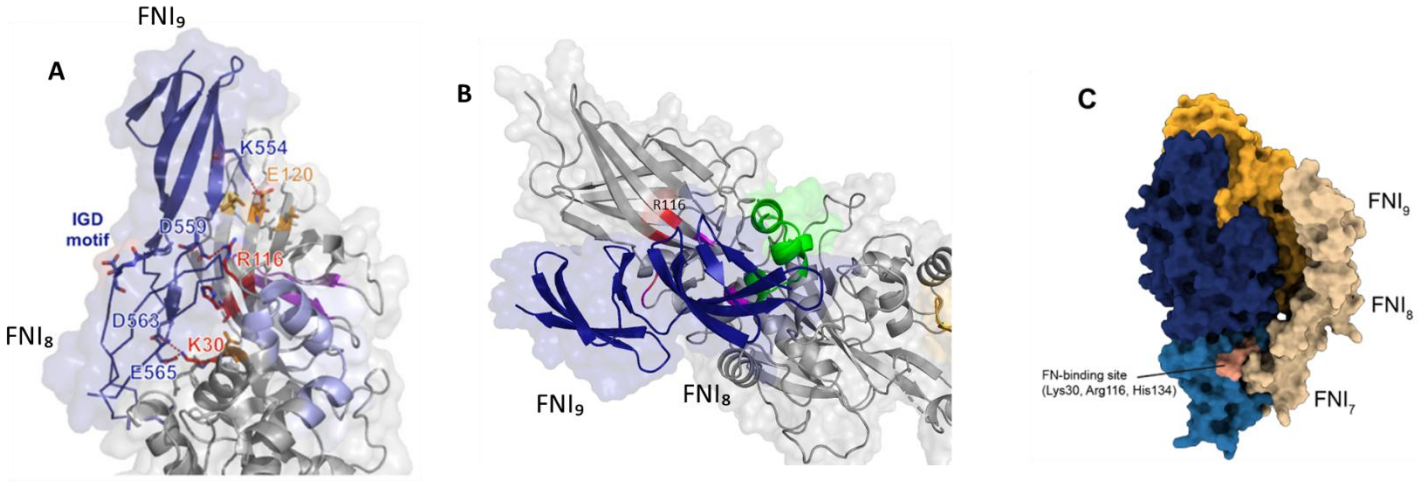
A recent study by Selcuk *et al*.[81] using stretched and relaxed fibronectin fibres has also challenged some preconceptions about FN-TG2 binding. They found that TG2 binds more strongly to relaxed versus stretched fibronectin fibers and that this binding is dependent on the proximity of the N-terminal domain to the C-terminal domain of TG2. In this study, TG2's binding to 45FN was drastically decreased compared to full length FN fibers, contrasting with previous observations.[65] In addition to mechano-regulation experiments, crosslinking-mass spectroscopy (XL-MS) was performed between TG2 with both the 45FN fragment and full length FN. When TG2 was crosslinked with 45FN only one interprotein crosslink was seen, K30 on TG2 to K486 on the FN_7 domain. Selcuk *et al.* used their XL-MS results to guide *in silico* modeling which while greatly differing from Soluri *et al*.'s modelling of the interaction, still implicates K30, R116 and H134 as crucial for the interaction, in line with Soluri *et al.* and Cardoso *et al.* Interestingly, when TG2 was crosslinked with full length FN, K30 crosslinked

with K1837 and K1862 ($FNIII_{14}$) and the C-terminal domain of TG2 formed interprotein crosslinks outside of 45FN, namely with domains $FNIII_{14-15}$ and FNI_2. Considering these findings Selcuk *et al*. suggest that the C-terminal β-barrel domains of TG2 act as a synergy site for FN.

To summarize, it is fair to say that research has been conflicted on the precise residues involved with TG2's interaction with FN. The most recent docking studies have supported the experimental data obtained by HDX-MS and SDM experiments analyzed via SPR, with specifically mutations K30E, H134A and R116A greatly reducing binding.[80,81] Alongside these residues, additional residues may also be involved, including ones in the core domain of TG2, such as D198 according to SDM studies[79], as well as residues in the C-terminal domain according to XL-MS experiments.[81] Evidence of residues of the core domain being involved in FN binding is further supported by the fact that truncated TG2, containing only the N-terminal domain residues (1-143), bound to FN with lower affinity than full length TG2.[79] In light of these conflicting results, our lab conducted its own docking simulations, starting from an unbiased approach, to see if there are plausible binding poses that feature more contact with the core domain of TG2 as well as a potential FN-binding pocket.

Dr. Pauline Navals performed our own docking studies using the Protein-Protein Interaction module of the Molecular Operating Environment software (MOE). She performed a funnel computational analysis where each level of simulation of the TG2/45FN complex would comprise a higher degree of restraint, starting from consideration of the entire TG2 protein, progressing to just its N-terminal domain, to an identified hydrophobic cavity and finally restraining modelling with the three residues deemed to be crucial for TG2/FN binding, R116, H134 and K30. Using protein-protein interaction (PPI) simulations and the previous determined

structural data, she identified a potential small molecule binding site in a hydrophobic pocket created by residues R116 and H134, as well as residues 195-219, which were implicated in 45FN binding by HDX-MS experiments.[79] She used this new potential binding site as a restraint in the funnel computational analysis, which provided a potential pose for 45FN binding (Figure 3). She then used *in silico* protein contact analysis to highlight three residues in addition to R116, H134 and K30 that could be major interactors with this alternative binding pose: L114, E29 and I195. In this thesis project, we will present molecular biology techniques used to explore the potential involvement of these residues, starting with BLI and SDM experiments.



Soluri *et al.* 2019

Alternative Binding Pose (Dr. Navals)

Selcuk *et al.* 2024

Figure 3. In silico docking studies for FNI7-9 and TG2. (A) Figure was borrowed from Soluri *et al.*'s publication[80], docking with FNI_{8-9} residues overlapped with many residues they Cardoso *et al.*[79] showed to be important. (B) Dr. Navals docking includes core residues 195-219 (Green) R116, K30, H134 (Red) as well as additionally identified residues I195, L114 and E29 (Magenta). (C) Figure was borrowed from Selcuk *et al.*[81] and shows TG2 in the closed position. All three docking simulations show TG2's H134, R116 and K30 residues as being involved.

2.1.2 Bio-Layer Interferometry to Identify Key Residues

Bio-layer interferometry (BLI) and SPR are the most common techniques used to study biomolecular interactions. We used BLI to study the association of WT and mutant TG2 protein with 45FN, which represents the GBD. BLI works by having a biocompatible surface coated on the end of a sensor. In our case, this was a Ni-NTA sensor that binds to the C-terminal His tag of our TG2 protein. BLI measures slight changes in reflected light on the sensor surface caused by the formation of molecular complexes with the biomolecule loaded onto the biosensor. The molecule that forms the complex with the attached biomolecule is called the analyte and, in our case, this is 45FN. White light is reflected from two different surfaces, an internal reference layer and the light reflected from the molecular complex layer. Depending on the size and affinity of the analyte, the interference pattern of the white light changes and a wavelength shift can be measured. The time-dependent change of wavelength can be fitted to kinetic parameters like association and dissociation constants.[83] BLI has been used to measure the kinetics of molecules inhibiting protein-protein interactions as well as the interactions themselves.[84]

Our first objective towards exploration of the TG2 and 45FN interaction was to corroborate that the key residues R116, K30 and H134 identified by Cardoso *et al*[5,79] are indeed critical for 45FN binding to TG2. Included in this objective is the testing of residues deemed important in the alternative binding pose suggested by our group's docking experiments, including L114, E29, and I195. Over the course of my thesis work, five out of six residues were replaced with alanine through SDM and the interaction between the resulting mutants and 45FN was measured by BLI and compared to WT TG2. We deemed measurement by BLI to be potentially more representative of the actual binding interaction between 45FN and TG2 than previous SPR studies. The reason for this is due to the different immobilization techniques used

in the different studies. In our BLI studies, the C-terminal His Tag of TG2 is attached to the biosensor, which leaves the N-terminal domain, the principal binding site for 45FN, freely available for protein-protein interactions. In the previous SPR studies, 45FN was stochastically immobilized on the sensor using amide coupling chemistry, which could reduce the freedom of the protein-protein interaction.

The AL5 activity assay[85] was used to assess the integrity of WT and mutant TG2. This assay monitors the formation of a *p*-nitrophenolate product after hydrolysis of the TG2 substrate AL5. TG2's catalytic triad cleaves an ester bond in AL5 releasing *p*-nitrophenolate whose absorbance can be read at 405 nm. This assay is described in the experimental section, and a reaction scheme is shown in the supplementary data. It is important to monitor the structural integrity of TG2 when introducing mutations since loss of tertiary structure could reduce its capacity for use in protein-protein interaction studies with BLI.

The following chapter will describe the creation, expression and analysis of alanine mutated TG2 proteins used in BLI experiments.

2.2 Results and Discussion

2.2.1 Creation and Expression of TG2 Alanine Mutants at Key Residues

To identify which positions were crucial for association of fibronectin to TG2, we performed an alanine scan of the residues identified by Cardoso *et al*[5,79] and previous docking experiments from our lab. These residues are: R116, K30, H134 I195, E29, and L114. To begin our experiments, we performed site directed mutagenesis of these positions illustrated below with agarose gels taken of PCR amplification reactions with forward and reverse primers corresponding to alanine mutations for specific residues.

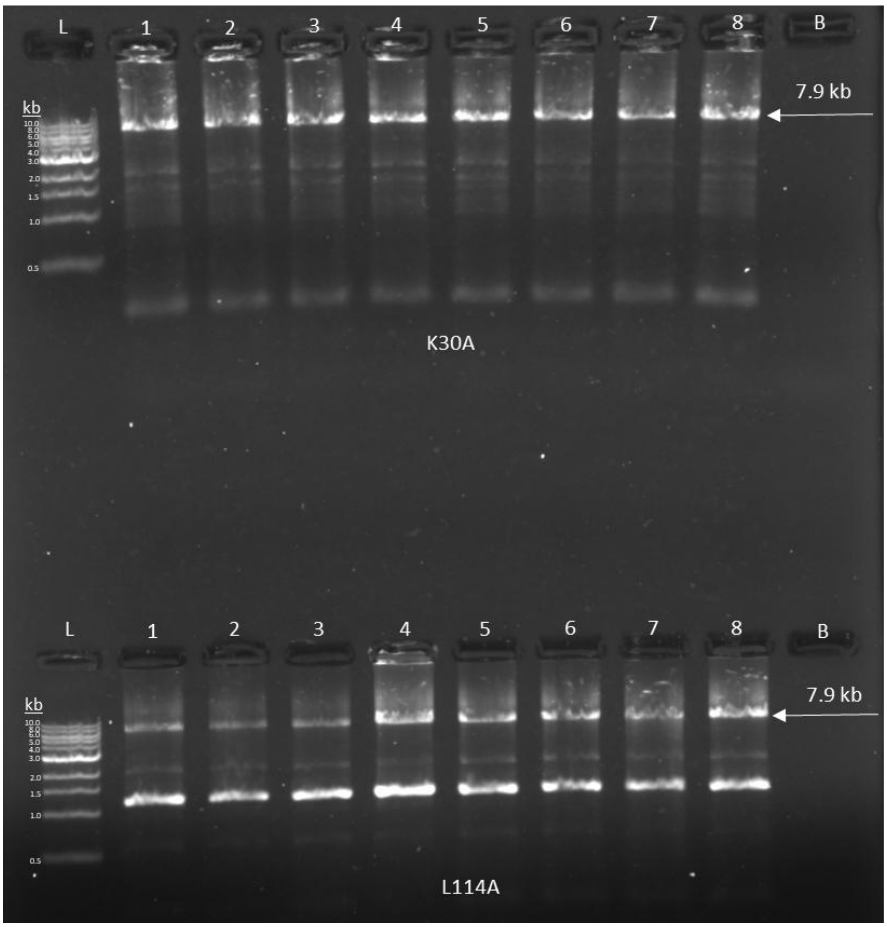


Figure 4. Agarose Gel of SDM PCR Amplification of K30A and L114A. The desired mutant TG2 amplicon is shown at around 7.9 kb and indicated with an arrow.

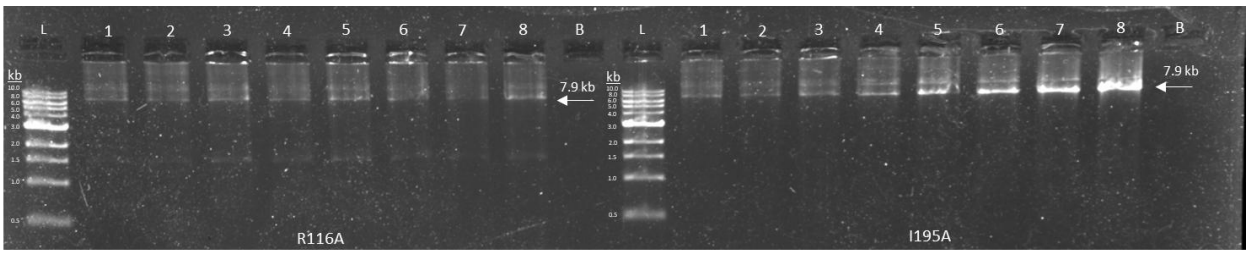


Figure 5. Agarose Gel of SDM PCR Amplification of R116A and I195A. The desired mutant TG2 amplicon is shown at around 7.9 kb and indicated with an arrow.

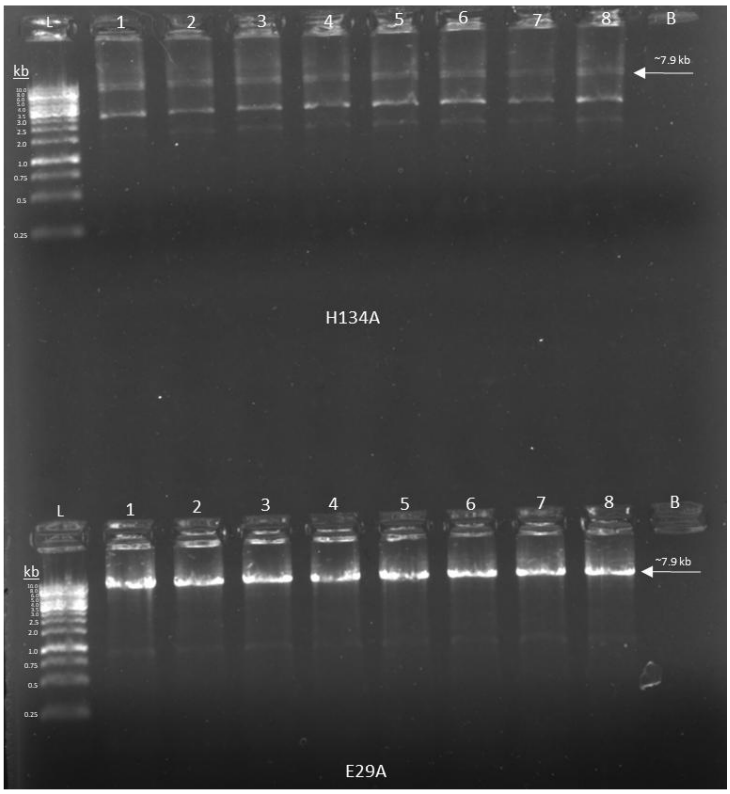


Figure 6. Agarose Gel of SDM PCR Amplification of E29A and H134A. Desired amplicon is shown at around 7.9 kb. The DNA ladder was not resolved very clearly.

Eight identical PCR reactions for each mutant were created and run alongside a blank reaction with no Q5 polymerase added. K30A, L114A, I195A, and E29A all produced bright bands corresponding to the expected kilobase length for the Cterm-pHis-hTG2 plasmid, 7.9 kb. R116A and H134A produced faint but visible bands at around this length. The agarose gel DNA ladder was resolved poorly for the E29A and H134A amplifications, but they were still determined to have bands present near the expected length of 7.9 kb. One positive reaction was chosen for each mutant to continue forward through the rest of the SDM process as described in the experimental section. Upon Sanger sequencing, all mutants were found to have the alanine mutation incorporated, except for the H134A mutant. SDM of the H134A mutant was attempted three different times. The first attempt at mutagenesis yielded no H134A bands on the agarose gel. A fresh WT TG2 template was prepared for the second attempt, which resulted in very faint bands present at around the correct amplicon length. Transformation of the DNA into colonies and sequencing resulted in negative results. For the third attempt, the H134A primers were redesigned to fall within a length of 18-24 bp, 40-60% GC content and T_M's between 54-69 °C as well as having less than a 5 °C difference between forward and reverse primers. Despite these attempts at troubleshooting, the H134A mutant was not obtained.

The next stage of the experiment was the expression of WT and alanine mutated TG2. Three batches of expressions were completed according to the protein expression method described in the materials and methods. Batch one included WT, R116A, and K30A, batch two included WT, E29A and L114A and batch three included WT and I195A. The SDS-PAGEs post Ni-NTA purification are shown below with only one example from the WT expressions shown for simplification.

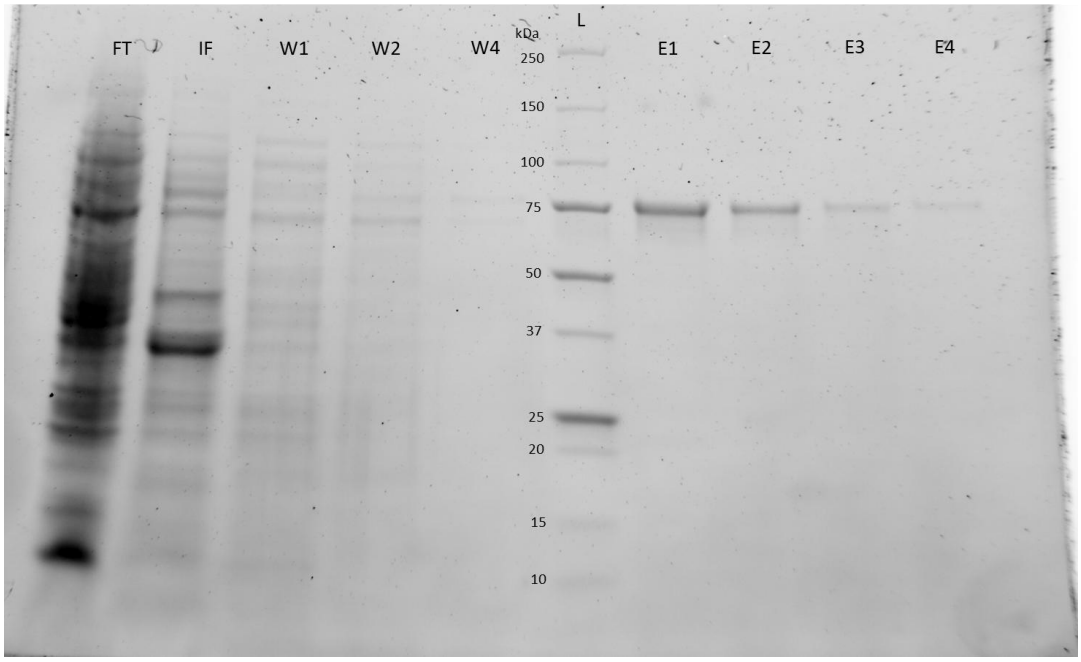


Figure 7. SDS-PAGE of WT TG2 Purification Post Ni-NTA Affinity Chromatography. Lanes are labelled with flowthrough (FT), insoluble fraction (IF), washes (W1, W2, W4) and elution fractions (E1, E2, E3, E4). The expected molecular weight for TG2 is ~78 kDa. This WT expression SDS-PAGE is from the first batch of expressions and is representative of the other two expressions of WT TG2 completed alongside alanine mutated TG2.

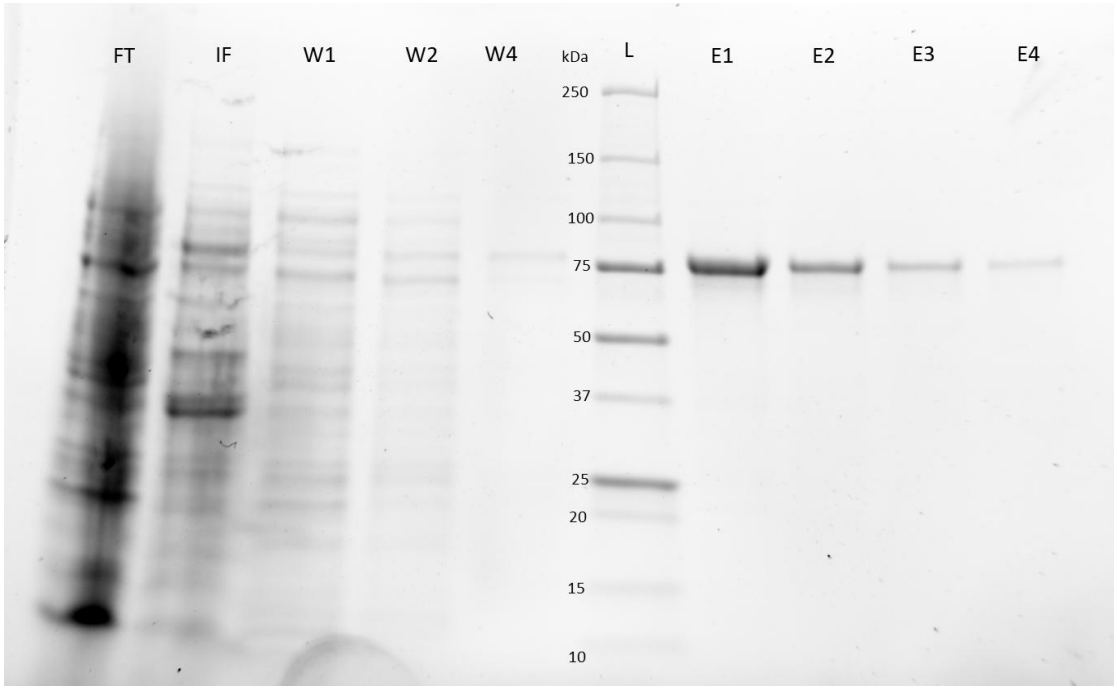


Figure 8. SDS-PAGE of R116A TG2 Purification Post Ni-NTA Affinity Chromatography. Lanes are labelled with flowthrough (FT), insoluble fraction (IF), washes (W1, W2, W4) and elution fractions (E1, E2, E3, E4). The expected molecular weight for TG2 is ~78 kDa.

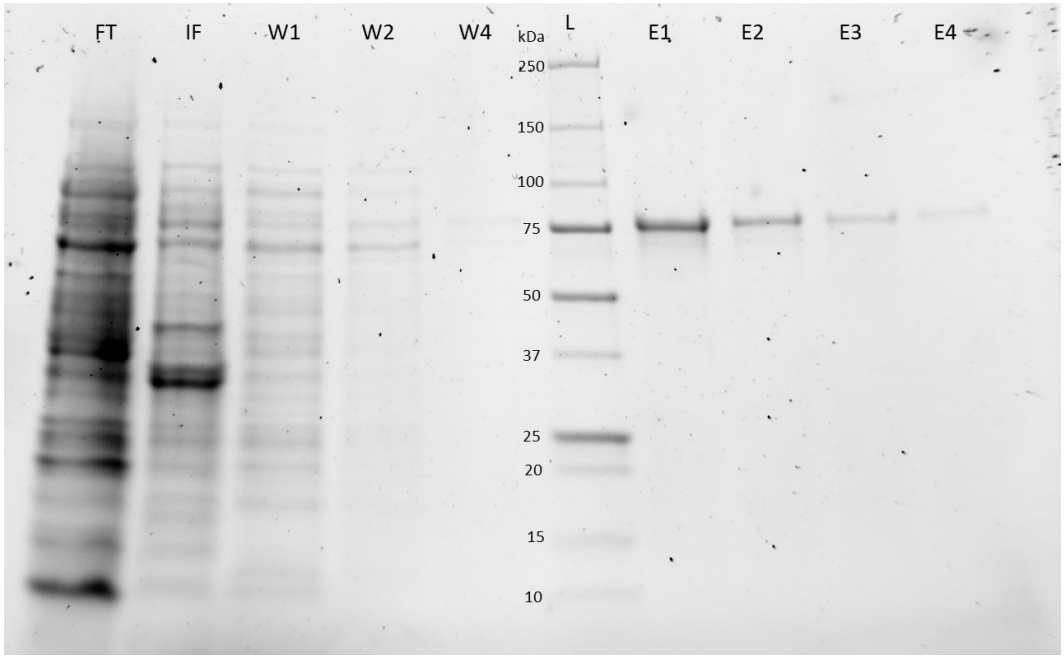


Figure 9. SDS-PAGE of K30A TG2 Purification Post Ni-NTA Chromatography. Lanes are labelled with flowthrough (FT), insoluble fraction (IF), washes (W1, W2, W4) and elution fractions (E1, E2, E3, E4). The expected molecular weight for TG2 is ~78 kDa.

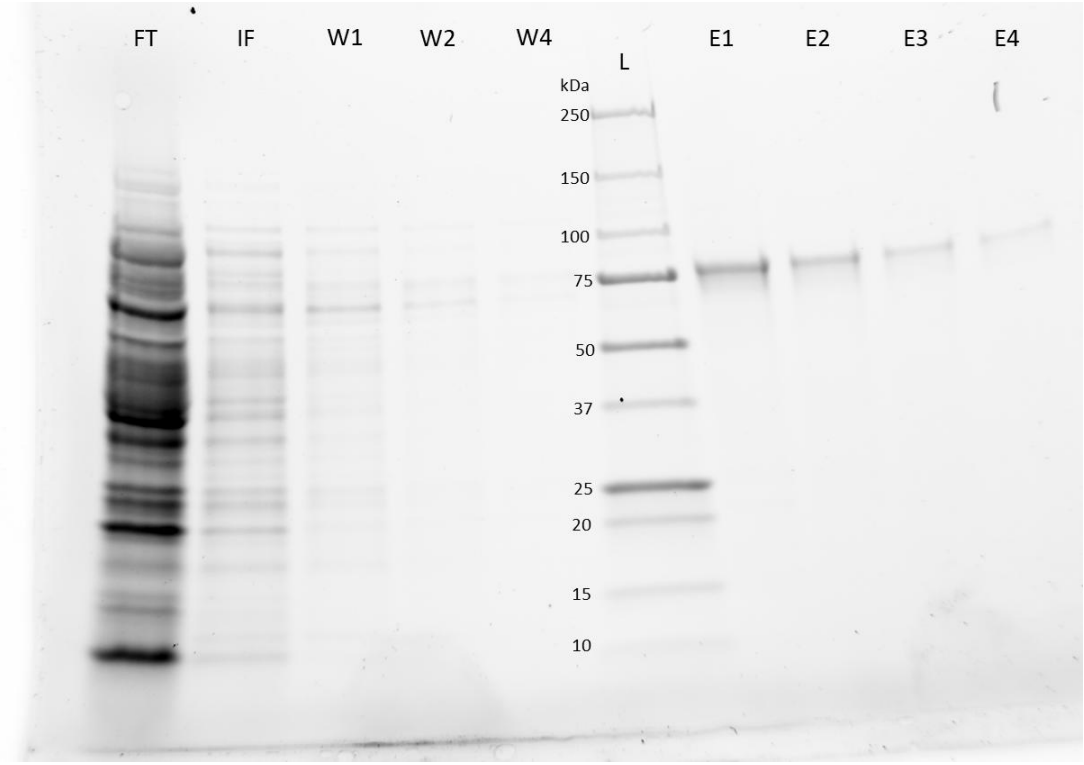


Figure 10. SDS-PAGE of E29A TG2 Purification Post Ni-NTA Chromatography. Lanes are labelled with flowthrough (FT), insoluble fraction (IF), washes (W1, W2, W4) and elution fractions (E1, E2, E3, E4). The expected molecular weight for TG2 is ~78 kDa.

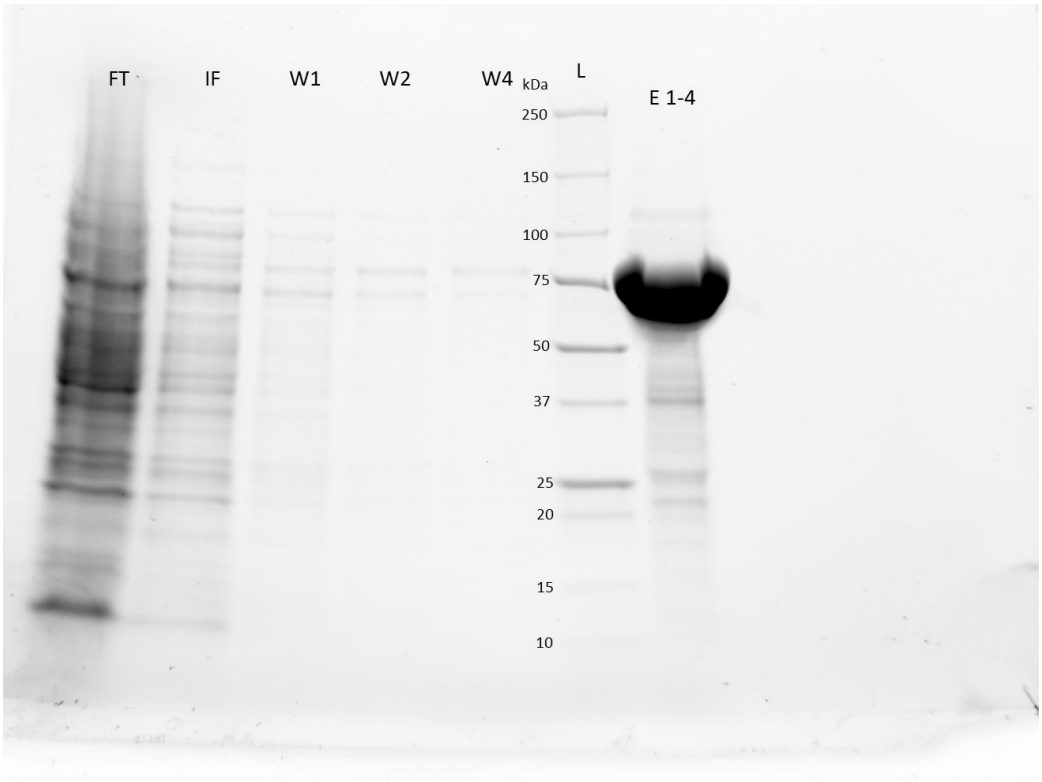


Figure 11. SDS-PAGE of L114A TG2 Purification. Lanes are labelled with flowthrough (FT), insoluble fraction (IF), washes (W1, W2, W4) and elution fractions (E1-4). The expected molecular weight for TG2 is ~78 kDa. The elution sample loaded onto the SDS-PAGE was taken after pooling and concentration with Amicon 30 kDa centrifugation tubes.

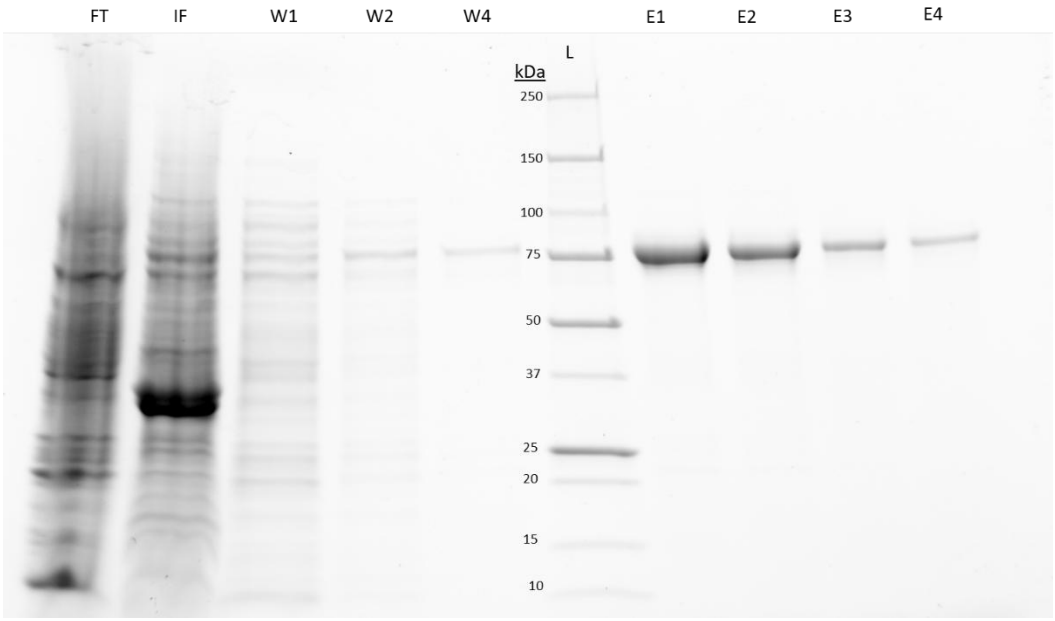


Figure 12. SDS-PAGE of I195A TG2 Purification. Lanes are labelled with flowthrough (FT), insoluble fraction (IF), washes (W1, W2, W4) and elution fractions (E1-4). The expected molecular weight for TG2 is ~78 kDa.

Table 1. Concentration, Yield and Specific Activity for WT and Alanine Mutated TG2 Proteins. The concentration was estimated using the Bradford assay and the specific activity was measured using the AL5 assay. The WT values are averaged from three different expressions while the mutant expressions are from one expression each.

Protein	Concentration (mg/mL) (mL)	Yield (mg/L)	Specific Activity (U/mg)
WT Average	5.58 ± 1.81	9.45 ± 1.49	0.047 ± 0.032
R116A	12.27 ± 0.97 (0.4)	9.82 ± 0.78	0.049 ± 0.002
K30A	2.58 ± 0.39 (1.4)	7.20 ± 0.27	0.055 ± 0.003
E29A	5.88 ± 0.77 (0.79)	9.29 ± 1.22	0.086 ± 0.008
L114A	15.31 ± 2.58 (0.4)	12.25 ± 2.07	0.040 ± 0.004
I195A	17.02 ± 1.69 (0.65)	22.13 ± 2.19	0.035 ± 0.002

The SDS-PAGEs, Figures 9-14 were obtained with samples from the Ni-NTA affinity chromatography process for WT and mutant proteins. The flowthrough is a sample of the liquid passed through the columns after incubation with the Ni-NTA resin. This fraction should contain proteins that have low or no affinity for the Ni-NTA resin. The “IF” lane represents a sample of the insoluble fraction after ultracentrifugation of lysed cells. Proteins that are insoluble will end up in this fraction and including it in the SDS-PAGE allows us to make sure TG2 is not insoluble after expression. Figures 7-12 show an abundance of non-binding proteins in the flowthrough and the insoluble fraction but no excess of protein at a size of 78 kDa indicating that TG2 was not present in the flowthrough or insoluble fraction in large amounts. The elution fractions show a dense band at 78 kDa, the expected molecular weight for TG2. These bands are relatively

isolated in the elution fractions, indicating good purity of the samples. In Figure 11 the elution fractions for the L114A expression were pooled resulting in an overloaded amount of protein in the SDS-PAGE, making it hard to see its molecular weight against the ladder. Despite this we deemed the band to be TG2 due to its abundance over other proteins and relative location on the gel. TG2 was successfully purified according to the SDS-PAGEs.

The concentration and overall yields, presented in Table 1, of the WT and mutant TG2 proteins were measured using the Bradford assay. The data in Table 1 for WT TG2 is averaged from three different expressions while the data for the mutant proteins are from single expressions. The concentrations are quite variable due to the volumes of protein after concentration by Amicon centrifugation leading to different volumes each expression. The yields are useful for direct comparisons, which are very similar to the WT for all alanine mutants. The I195A expression notably had almost double the yield of protein, although it is unclear why this is the case as there was no changes to the expression protocol. The yield of protein is generally high, around 10 mg/L or higher, and enough WT and mutant proteins were expressed for BLI analysis.

The specific activity of the proteins was measured using the AL5 activity assay, which measures the catalytic activity of the TG2 core domain. WT specific activity in U/mg was measured in a wide range for the three expressions as evidenced by the large standard deviation between the three samples. The specific activities of the alanine mutants in Table 1 fell into the range of the WT proteins. It is important to compare the activity of mutant TG2 to WT to have a benchmark for proper folding. If there is a large loss of activity for mutant TG2 it indicates improper folding of the protein and reduces its integrity for use in protein-protein binding

experiments using BLI. These alanine mutated TG2 proteins were deemed to have enough structural integrity to proceed with BLI experiments, which will be discussed in the next section.

2.2.2 Bio-Layer Interferometry to Measure K_D

The WT and mutant TG2 proteins that were created in the previous section were used to perform measurement of the K_D between 45FN and TG2 using BLI. BLI runs were performed according to the experimental section described below, in triplicate for each mutant and WT TG2.

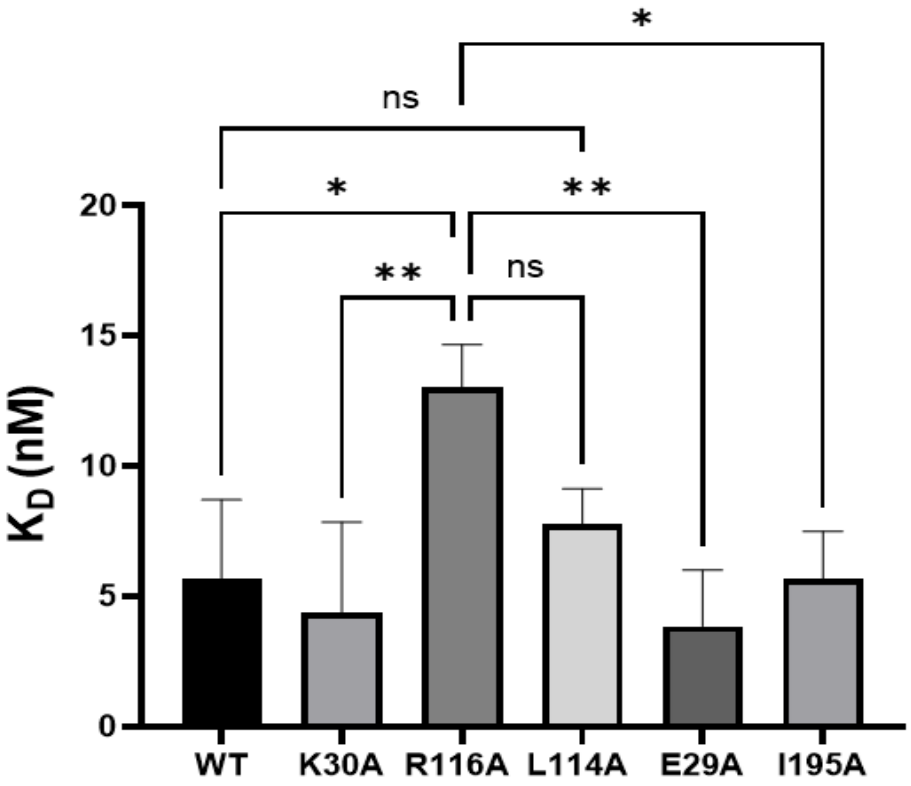


Figure 13. Determination of K_D Using BLI for the Association Between 45FN and Mutant or WT TG2. Measurements were done in triplicate, and values are plotted with the standard deviation. K_D values were determined Figure 13. Determination of KD Using BLI for the Association Between 45FN and Mutant or WT TG2.using global fitting on the BLI software. Results were analyzed on GraphPad Prism using ANOVA analysis. Single asterisks denote a p value ≤ 0.05 while double asterisks denote a p value ≤ 0.01. "ns" denotes a p value of ≥ 0.05 and is not significant.

The BLI results show a significant difference in K_D, illustrated in Figure 13, for only one alanine mutant, R116A. The K_D's of K30A, L114A, E29A and I195A are not significantly different than the WT. The residues implicated in the alternative binding mode modelled by our

group L114A, E29A and I195A did not produce a disruption in the association of 45FN compared to WT TG2, which does not support this hypothetical binding mode. Residues L114 and I195 are changing from a larger to a smaller hydrophobic side chain when replaced with alanine, which is not a major difference. The E29A change, glutamic acid to alanine is more potentially disruptive since it is negatively charged at physiological pH and could form an ionic interaction with an FN residue. Its replacement with an alanine is more indicative that FN does not interact at this position compared to the L114 and I195 mutations.

Table 2. Comparison of K_D Values to Literature Values. Values obtained for the K_D from our group's BLI experiments are compared to Cardoso *et al.* and their SPR experiments.[5]

TG2	Our Study K_D (nM)	Cardoso *et al.* K_D (nM)[5]
WT	5.69 ± 3.02	0.41 ± 0.01
R116A	13.00 ± 1.64	10.40 ± 1.28
K30A	4.38 ± 3.48	
K30E		2.38 ± 0.54
H134A		3.79 ± 0.29

Table 2 shows the comparison between our BLI measurements and the measurements performed by Cardoso *et al*. using SPR. Notably, they report a much lower K_D for the WT TG2 and 45FN association, which leads to their mutants having a bigger difference in K_D from the WT, even though our recorded K_D values for the mutant residues, R116A and K30A compared to K30E are similar. Our research confirms that R116A when mutated alone causes a significant impact on the K_D of TG2 to 45FN. However, there was no impact seen between K30A and WT

TG2 in our study, even though a substantial impact was seen in theirs. On the other hand, these mutants are not directly comparable, since the alanine mutation done for our study abolishes the positive charge of lysine, while replacing it with glutamic acid changes it from positively charged to negatively charged. Therefore, our mutation is a more direct interrogation of the effect the positive charge on lysine has in stabilizing the protein-protein interaction between 45FN and TG2.

Overall, we confirm that R116 mutated alone is important for 45FN interaction with TG2, but our results do not corroborate K30 as having the same effect. In another study by Cardoso *et al*. they found that double and triple mutants of TG2 R116A/H134A and R116A/H134A/K30E resulted in an increase in K_D of 400 and more than 2000-fold respectively., Our study confirms that this effect cannot be achieved by single residue replacements of R116 and K30.

The general difference in K_D measurement, most strongly illustrated by the difference in WT K_D between studies, could be due to the difference in analysis methods. The immobilization technique varies between the studies, in that our study immobilized TG2 to a Ni-NTA biosensor through a C-terminal His-tag, whereas Cardoso *et al*. immobilized 45FN using amide coupling chemistry directly to CM5 sensor chips.[5,79] We believe our study provides more freedom for the protein-protein interaction to occur since the unknown conformation of 45FN is free to associate as needed to the known N-terminal binding domain of TG2. According to that approach taken by Cardoso, 45FN is immobilized reducing its freedom to associate.

2.3 Conclusion

In conclusion, TG2 mutants R116A, K30A, E29A, L114A and I195A were successfully created using SDM. Mutants H134A as well as H134A/R116A were also targets for creation but the SDM experiments and expressions were not successful and therefore were not able to be measured. The measurement of R116A and K30A's effect on 45FN association to TG2 resulted in a less dramatic increase in K_D than previously reported results[5] with no significant increase detected for K30A and a three-fold increase for R116A. Residues E29A, L114A and I195A showed no significant increase in K_D when compared to WT TG2. These results corroborate the finding that R116 is an important residue for the 45FN-TG2 interaction but do not support K30 as being crucial on its own. Furthermore, these results do not support the alternative binding pose suggested by our group, since mutation of residues L114, E29 and I195 resulted in no significant increase in K_D compared to WT TG2. Differences in methods could account for the discrepancy between WT and 45FN K_D measurements with our approach lending more freedom to the 45FN binding partner. Future research using BLI should be focused on analysing the H134 residue, as well as additional residues implicated in the binding pose modelled by Soluri *et al*., such as E120 or D198.[80] More investigation into the alternative binding pose proposed by our group should focus on mutation of residues that could participate in 45FN association by hydrogen bonding or ionic interactions.

2.4 Materials and Methods

2.4.1 General Remarks:

Reagents and chemicals used in solutions were ordered from BioShop Canada unless stated otherwise. SDS-PAGE was performed using Bio-Rad Mini-Protean TGX pre-cast, stain-free gels (4-20% polyacrylamide), and visualized using a Bio-Rad ChemiDoc MP imager. Agarose gels were prepared fresh prior to use, containing 1% agarose (w/v) and 1 μg/mL ethidium bromide. DNA was separated on agarose gel at 105 V in Tris-Acetate-EDTA (TAE) buffer and visualized using a Bio-Rad ChemiDoc MP imager. Bradford assays were performed using the established protocol[86] and measured using a UV-vis spectrophotometer (Varian Cary 100 Bio, 1 cm path length, or Agilent Cary UV-Vis Multicell Peltier.) AL5 activity assays were performed according to the previously established method.[87,88] The specific activity was measured using either a plate reader (Biotek Synergy H1) or UV-vis spectrophotometer (Varian Cary 100 Bio, 1 cm path length). One unit of enzyme activity (U) is defined as the amount of TG2 required to catalyze the formation of 1 μmol of *p*-nitrophenolate per minute. Specific activity is reported as units of activity (U) per milligram of protein.

2.4.2 Site-Directed Mutagenesis of Alanine Mutants

Site-directed mutagenesis was performed using the Cterm-pHis-hTG2 plasmid[89] (Addgene #100719) as a template. Forward and reverse custom primers were synthesized using Integrated DNA technologies with alanine mutations encoded on the forward primer. The sequences of these primers are presented in the appendix. PCR amplification was performed using template with forward and reverse primers using the Q5 High-Fidelity DNA polymerase (New England Biolabs) using BioRad MJ Mini Personal Thermocycler or the BioRad T100 Thermocycler. Amplification of the correct gene was confirmed using agarose gel electrophoresis, DNA was purified (E.Z.N.A Cycle Pure Kit, Omega BioTek) and then the template was digested at 37 °C overnight using Dpn1 (New England Biolabs). After the digestion, the PCR was purified again using the Cycle Pure Kit and phosphorylated with T4 PNK (New England Biolabs) then ligated with T4 DNA Ligase (New England Biolabs). The amplicon as well as a negative control reaction (no Q5 Polymerase at PCR amplification step) were transformed into chemically competent *E. coli* BL21 (DE3) or DH5α cells and plated on LB-Agar plates containing 50 µg/mL kanamycin. If no colonies were present on the negative control plate, five colonies were picked and DNA was extracted (E.Z.N.A Plasmid DNA mini kit, Omega BioTek) then sent for Sanger sequencing (Génome Québec).

2.4.3 Protein Expression

A fresh transformation was performed using plasmids whose sequences were confirmed by Sanger sequencing, coding for either WT, or alanine mutated Cterm-pHis-hTG2 plasmids into E. *Coli* BL21 (DE3) cells. A 5 mL culture of Terrific Broth (TB) was inoculated with a single colony in a culture tube with 50 µg/mL kanamycin and shaken overnight at 37 °C, 250 RPM for about 16 hours using the Innova 4330 shaking incubator. The next day 495 mL of TB containing 50 µg/mL kanamycin was inoculated with the 5 mL overnight culture and the flask was shaken at 37 °C at 250 RPM until and OD_{600} of 0.5-0.8 was reached. The expressions were induced with a final concentration of 1 mM IPTG and were shaken for 20 hours at 18 °C. The next day the induction culture was centrifuged at 4 °C and 4347xg for 30 minutes using the Eppendorf S910R Centrifuge Pellets were frozen at -80 °C for purification on a separate day.

2.4.4 Protein Purification

Cells were resuspended on ice in a total of 12 mL of GST TG2 lysis buffer (20 mM Tris, 150 mM NaCl, 1 mM EDTA, 1 mM TCEP, 0.5% Triton X-100, pH 8.0) or His Tag TG2 Lysis Buffer (50 mM Na_3PO_4, 400 mM NaCl, 5 mM imidazole, 0.5% triton-X100, pH 7.5) and homogenized using the Avestin Homogenizer. For homogenization the resuspended cell volume was split in half, and each half volume was passed through the homogenizer twice at a slow rate of around 1-2 mL/min. Lysate was briefly cooled on ice between passes. The lysate was separated into soluble and insoluble fractions by centrifugation at 4 °C and 11,000 RPM using the Eppendorf S910R centrifuge. Cells and proteins at this point were kept on ice unless stated otherwise. The soluble fraction was incubated with conditioned 0.5 mL of Ni-NTA Resin (Sigma) for 1 hour at 4 °C with shaking. Ni-NTA affinity purification was done at 4 °C using a BIORAD column with four 5 mL washes with TG2 Wash Buffer (50 mM Na_3PO_4, 500 mM NaCl, 60 mM imidazole, pH 7.5) and four 5 mL elution fractions were collected with TG2 Elution Buffer (50 mM HEPES, 100 mM NaCl, 300 mM imidazole, 10% (v/v) glycerol (50%), pH 7.0). Samples from each fraction were mixed in a 1:1 ratio with Laemmli buffer (BioRad) with β-mercaptoethanol and boiled at 95 °C for 5 minutes. Insoluble fraction samples were created by dissolving a small portion of the pellet in 100 µL 8M Urea, then following the same sample treatment. Samples were loaded at 10 µL onto an SDS-PAGE. Elution fractions were concentrated using 10 or 30 kDa Amicon Ultracentrifugation tubes (Millipore). Elution fractions were concentrated to < 1 mL using multiple 20 minute 4000 xg centrifuge spins using the Sorvall ST1R Plus-MD centrifuge. Elution buffer was exchanged with 9 mL TG2 Storage Buffer (20 mM HEPES, 1 mM EDTA, 1 mM TCEP, 20% (v/v) glycerol (50%), pH 7.0) three times and the volume was concentrated to below 1.5 mL. The concentration and specific activity were measured using the Bradford and

AL5 assays respectively. Proteins were then aliquoted and flash frozen using dry ice and acetone, the proteins were stored at -80 °C. Note that buffers were used with and without TCEP addition depending on the purification.

2.4.5 Bio-Layer Interferometry

BLI experiments were run on the ForteBio BLItz^{TM} machine using TG2 immobilized on the Sartorius Octet Ni-NTA biosensor by its C-terminal His Tag. WT or alanine mutant TG2 was loaded onto the sensor at a concentration of 10 µg/mL and the K_D of 45FN (MilliporeSigma) binding was determined using a range of 45FN concentrations (20, 15, 10, 5, 2.5, 1, 0.5, 0.1 and 0 µg/mL) in triplicate. The Ni-NTA sensor was stripped and regenerated at the end of each run and TG2 was freshly prepared and loaded onto the sensor at the start of each run. Buffers used for the BLI run were Assay Running Buffer (1X dPBS, pH 7.1 filter sterilized with a 0.22 µM filter), Stripping Buffer (100 mM Glycine, pH=1.77), Re-equilibration Buffer (0.1% dPBS in MQ H_2O), and Regeneration Buffer (100 mM $NiSO_4$). Each BLI run followed this procedure with 250 µL of the appropriate buffer: Initial Baseline (30 s, Assay Running Buffer), Loading (90 s, Assay Running Buffer with 10 µg/mL TG2), Baseline (30 s, Assay Running Buffer), Association (120 s, Assay Running Buffer with variable concentrations of 45FN), Dissociation (120 s, Assay Running Buffer), Stripping of Ni-NTA (10 s x 3, Stripping Buffer), Re-equilibration of Sensor (10 s × 3, Re-equilibration buffer), Ni-NTA Regeneration (60 s, Regeneration Buffer) and Final Baseline (60 s, Assay Running Buffer). Results were analyzed using the global fitting software for the BLI Blitz pro. These fitting has been used in other studies with TG2 and 45FN.[90] The Ni-NTA biosensor was replaced for each TG2 mutant, including WT. No loading was seen both without TG2 and with 10 µg/mL His-tagged enzyme NylC as negative controls.

Chapter 3: Incorporation of Azido-L-Phenylalanine into Transglutaminase II Towards Crosslinking with Fibronectin

3.1 Introduction

3.1.1 Genetic Code Expansion

Genetic Code Expansion (GCE) is a groundbreaking and versatile technology that has been in development since the discovery of tRNA as the link between mRNA and protein synthesis. It allows researchers to expand beyond the natural 20 amino acids and incorporate unnatural amino acids (UAAs) into proteins. These UAAs can have a variety of functions such as, fluorescence, cross-linking and ligand binding.[91] GCE works by using evolved orthogonal transfer RNAs (tRNAs) and aminoacyl tRNA synthetases (aaRS) that recognize UAAs and can incorporate them into "blank" codons during translation at the ribosome.

Discovery of tRNA as the adaptor molecule that connects mRNA to amino acids during translation has led to manipulation of the types of amino acids tRNAs can transfer. The adaptor hypothesis of tRNA was proven by chemical reduction of Cys-tRNA^{Cys} to Ala-tRNA^{Cys} which showed that the tRNA was the selective moiety in polypeptide synthesis.[91,92] Further modifications of tRNAs were discovered such as Tyr-tRNA^{Tyr} being able to incorporate D-Tyrosine[93] into proteins after enzymatic synthesis of D-Tyr-tRNA^{Tyr}. Additionally, hydroxylation of Phe-tRNA^{Phe} to hydroxyPhe-tRNA^{Phe} resulted in oligomers being synthesized with multiple ester bonds.[94] Another early example of tRNA modification is the generation of N_{ε}-azidobenzoyl-

Lys-tRNA^{Lys} which competed with endogenous Lys residues at multiple residue positions in rabbit reticulocyte lysate.[95] UAAs were being created by modifying the natural amino acid on the tRNA; however, there was a lack of a general system for charging tailored UAA's onto tRNA. A breakthrough in development of *in vitro* incorporation of UAAs by the Shultz lab[96] led to the ability to charge UAA's onto tRNA's using T4 RNA ligase. This allowed charging of UAA's onto suppressor tRNAs, tRNAs that encode for stop codons, such as the amber (TAG), ochre (TAA) and opal (TGA) codons. This breakthrough increased the ease of UAA-tRNA synthesis and introduced site-selectivity into GCE.

Another breakthrough in the field came from the Shultz lab when they were able to fully genetically encode the material required for GCE in *E. coli*.[97] In order to have a functioning system for incorporation of UAAs, a few criteria must be met. Firstly, the UAA-tRNA must not be recognized by endogenous *E. coli* aaRS but must function efficiently in translation. Next, an aaRS paired with the UAA-tRNA must be present that does not recognize other, endogenous tRNAs. This aaRS must also only charge its paired UAA-tRNA with an unnatural amino acid, not an endogenous one. Finally, the UAA must be effectively transported to the cytoplasm when added to growth medium.[98] Shultz found a likely orthogonal suppressor tRNA and aaRS pair that could match this criteria in *E. coli*, the *M. jannashii* tRNA^{Tyr} and TyrRS. By performing directed evolution, with negative or positive selection using amber suppression of toxic or pro-survival genes respectively, they were able to engineer an orthogonal tRNA/aaRS system that selectively incorporates the UAA *O*-methyl-L-tyrosine in place of the amber stop codon in *E. coli*.[98] This breakthrough enhanced the process of GCE from chemoenzymatically synthesizing UAA-tRNAs, to utilizing the endogenous machinery in living cells, increasing the application and adoption of the technology by non-experts in the field.

Evolved tRNA/aaRS pairs have now been synthesized using both prokaryotic and eukaryotic endogenous machinery. The unique properties of UAA's have been used in a wide variety of applications, from enzyme engineering to mimicking epigenetic protein modifications.[99] An example of enzyme engineering comes from Jackson *et al*.[100] where the incorporation of *p*-nitrophenylalanine into nitroreductase in place of its active site Phe124 resulted in an over 30-fold increase in catalytic efficiency. In addition, newly synthesized proteins in mammals can be tagged with UAA's such as Azidohomoalanine (AHA). In a study by Dieterich *et al*.[101] 195 newly synthesized proteins were identified through chemoselective tagging of azide labelled proteins with an alkyne affinity tag after incorporation of AHA. Another application of GCE is to aid research in the function of epigenetic protein modifications by incorporating UAAs that mimic post-translational modifications. One example of this is a study by Park *et al*.[102] where a suppressor tRNA was evolved to incorporate *O*-phosphoserine, the most abundant phosphorylated amino acid in eukaryotes.[91] For our purposes, we aimed to use GCE to incorporate the photoactivable crosslinking UAA, Azido-L-Phenylalanine (AzF) into specific sites on TG2 that are potentially involved in its protein-protein interaction with FN.

3.1.2 AzF Crosslinking to Probe Protein-Protein Interactions

Protein-protein interactions can be investigated using crosslinking methods in tandem with mass spectroscopy (MS). Traditional crosslinking methods usually use a reagent or crosslinker that forms a crosslink between two functional groups between proteins. This crosslinker usually has a defined length and the protein-protein complex can be digested and analyzed via MS. The identification of the paired residues along with the distance restraints of the specific crosslinker can be used to map protein-protein interfaces using *in silico* modelling.[103] Selcuk *et al*.[81], in their XL-MS study of TG2 and FN, utilized a 1:1 mixture of non-deuterated and deuterated disuccinimidyl suberate (DSS-d0/d12) as a crosslinker that specifically reacts with lysine residues. They also induced crosslinking of carboxyl groups and of amine groups with carboxyl groups using non-deuterated and deuterated pimelic dihydrazide with the coupling reagent 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium or DMTMM. These traditional crosslinkers have limitations since they only allow specific functional groups to interact with each other. We aimed to use the UAA AzF as a more specific crosslinker to explore single residue positions in the TG2/FN interface.

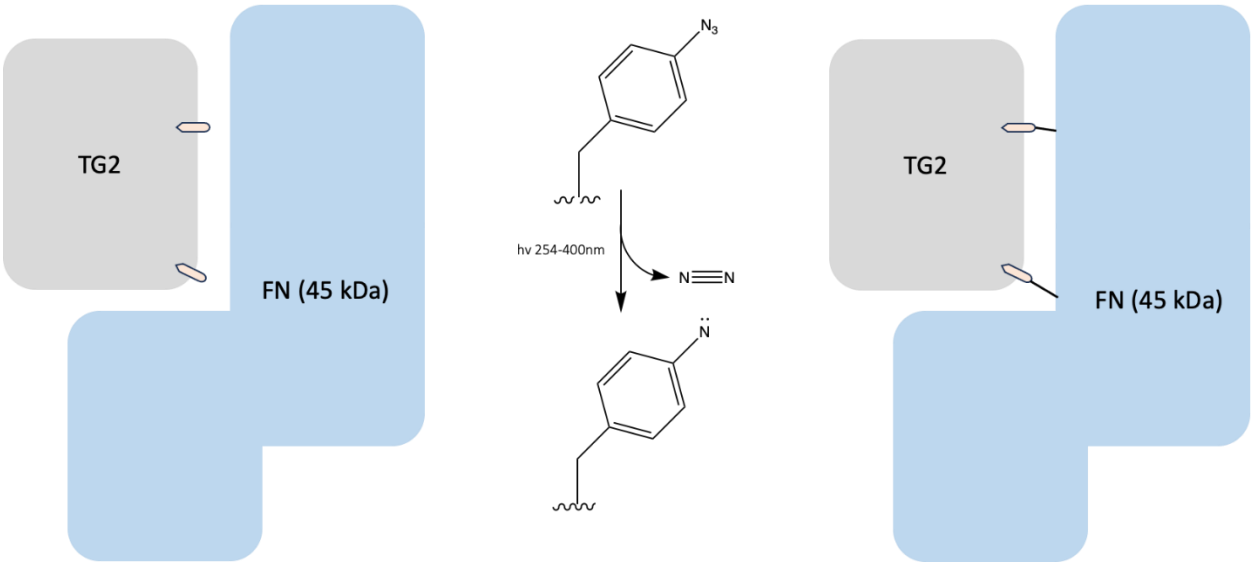


Figure 14. Simple Schematic of AzF Crosslinking Strategy. The creation of a nitrene through UV activation of AzF creates a highly reactive sidechain that binds to any side chain in proximity. We hope to use this to create crosslinked protein-protein complexes between TG2 and 45FN.

Incorporation of AzF through site-specific genetic code expansion can capture snapshots of protein-protein interactions. AzF is photoactivable using 254-400 nm UV light which creates a highly reactive nitrene that can react with the closest residue across from it. This is advantageous over traditional crosslinking methods because the residue on the protein interaction partner opposite from AzF does not need to be a specific functional group to react. AzF crosslinking has been used to study protein binding to substrates, such as bacterial calcium-dependent periplasmic LapG protease to its substrate LapA, a crucial part of bacterial biofilm formation[104] Its reactivity

with a conjugated Cy5 fluorophore has been used in Förster resonance energy transfer (FRET) experiments to investigate the distance dependent interaction of Nsp13 helicase with its dsDNA substrates.[105]

The following chapter describes our attempt to incorporate AzF into residues believed to be important for the TG2 and FN interaction by amber codon suppression. We attempted to monitor the incorporation of AzF in TG2 using a reactive DBCO-Cy5 fluorescent dye and to form a crosslinked protein-protein complex between 45FN and AzF incorporated TG2. We hoped to analyze paired residues by digestion of the covalently bound protein-protein complex and subsequent MS. The residues selected for incorporation of AzF include K30, R116, and H134, chosen due to their proposed involvement in binding 45FN, as described in previously mentioned studies.[5,79,81] We also wanted to study L114, E29 and I195, based on the alternative binding pose proposed by our group. We also attempted AzF incorporation of residues F203 and A66, which are located near the potential hydrophobic pocket in our alternative binding mode. These two residues were included based on their possibility of interfacing with 45FN without being direct binding partners in the interaction, with the hope of improving crosslinking product formation through less disruption of the native non-covalent association. Three different evolved tRNA/aaRS pair expression systems were used to troubleshoot low expression and crosslinking product formation. All of these approaches have been used in other studies to successfully incorporate AzF into proteins expressed in bacterial systems, specifically *E. coli*.[105–107] The following chapter has been organized chronologically and by the expression system used for AzF incorporation, including pDULE-pCNF, pEVOL-pAzF or pULTRA-pCNF.

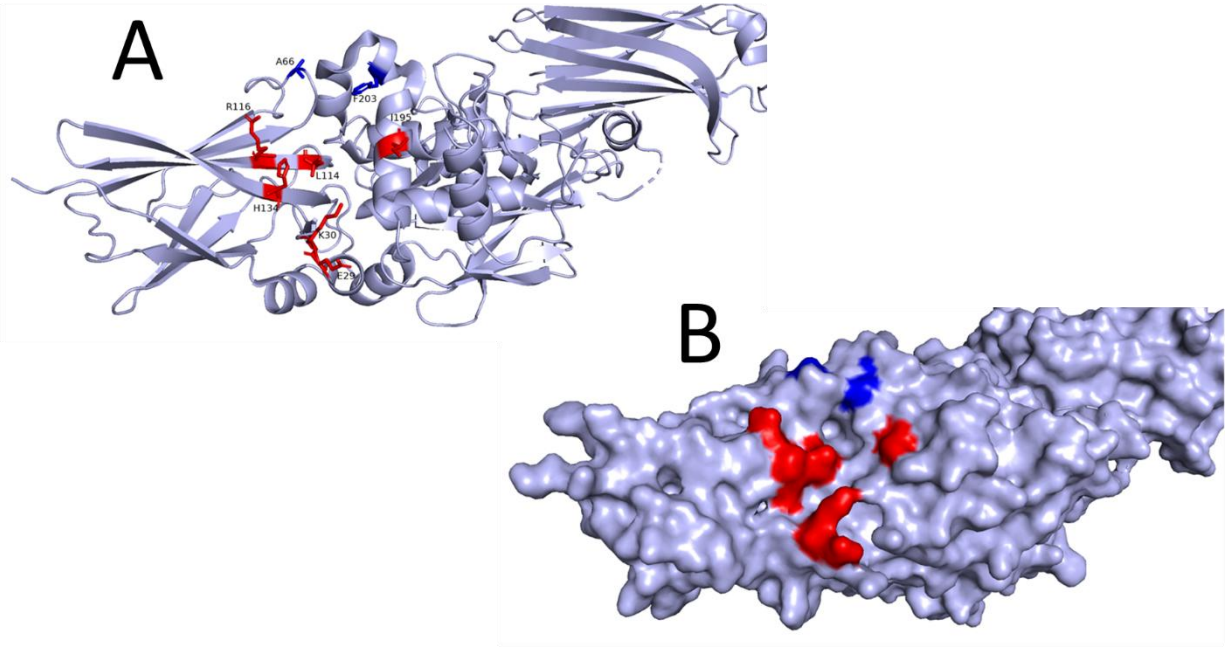


Figure 15. Selected Residues for Incorporation of AzF. (A) Cartoon structure of TG2 with side chains of residues visible. Residues were chosen for being previously established as crucial residues, R116, K30, H134 or as through *in silico* protein contact analysis in Dr. Navals proposed alternative binding pose, L114, E29, I195 these are shown in red. Additional residues F203 and A66 were chosen for not being directly implicated in the alternative binding pose but adjacent, which was hypothesized to improve crosslinking reactions with AzF due to minimal perturbation of the native protein-protein association these are shown in blue. (B) Residues that are surface exposed are visible.

3.2 Results and Discussion

3.2.1 Creation of Amber Mutations of TG2 at Key Residues

To site-specifically guide AzF to residues on TG2 we had to perform SDM to replace selected residues with the amber stop codon (TAG). Going forward, TAG residue replacements will be denoted by "X". We began by mutating six residues, three of which were previously implicated in TG2 FN association[5,79–81], R116X, H134X and K30X, and three of which we found to be implicated in our alternative binding pose, L114X, I195X and E29X. SDM was performed as described in the experimental section unless stated otherwise.

Figures 16-21 show images of agarose gels of the PCR amplification of each residue using a primer with the TAG codon in place of the natural residue. The expected size of the target amplicon is 7.9 kb. Eight PCR reactions were made and ran at a gradient of annealing temperatures on a thermocycler; each position had at least one reaction that contained an amplicon at the expected length for the Cterm-pHis-hTG2 plasmid. One "test" reaction from each agarose gel was chosen to continue forward with alongside the negative control reaction "B" which contained no Q5 polymerase. These reactions were subjected to a PCR clean up and Dpn1 digestion, which cleaves methylated DNA, the original DNA template. This reduces chances of WT plasmid contamination at later steps. The reactions were then phosphorylated with T_4 PNK and subsequently ligated with T_4 DNA Ligase, re-circularizing the linear amplicons into circular plasmids. The test and negative reactions were transformed into BL21 (DE3) cells and plated on 50 μg/mL Kanamycin LB-Agar plates. The Cterm-pHis-hTG2 plasmid contains a kanamycin resistance gene, so circularized plasmid will survive on these plates. If colonies are

seen on the negative control plate, that means that the test plate may be compromised with WT plasmid.

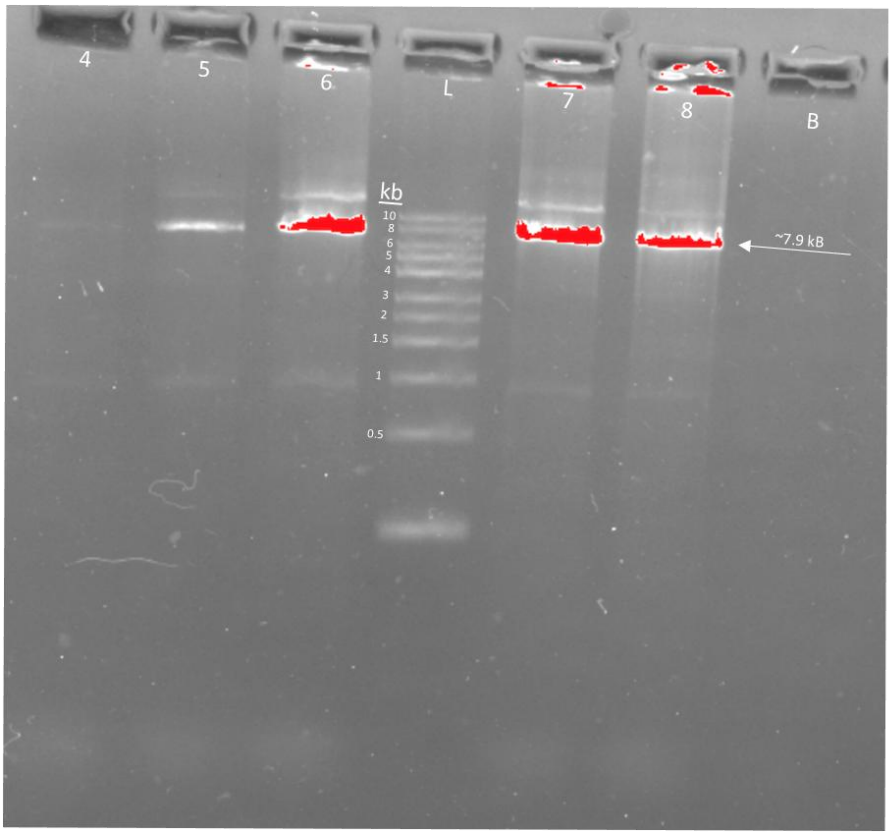


Figure 16. Agarose Gel of PCR Amplification from Site-Directed Mutagenesis of R116X.

Reaction #7 was chosen to continue forward through the SDM protocol.

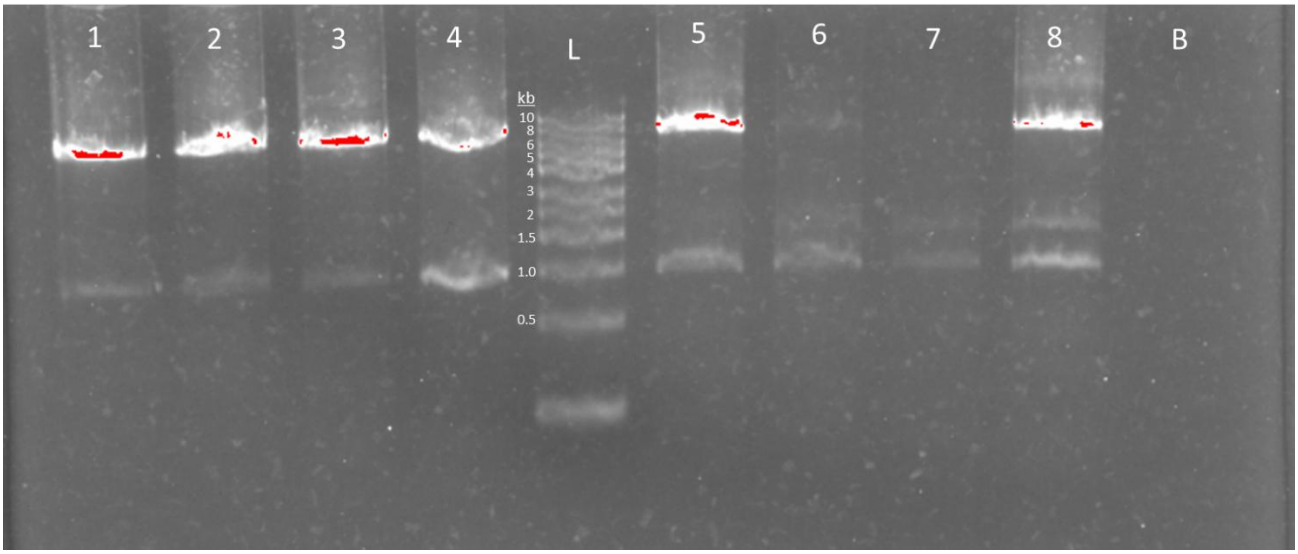


Figure 17. Agarose Gel of PCR Amplification from Site-Directed Mutagenesis of K30X.

Reaction #3 was chosen to continue forward with through the SDM protocol.

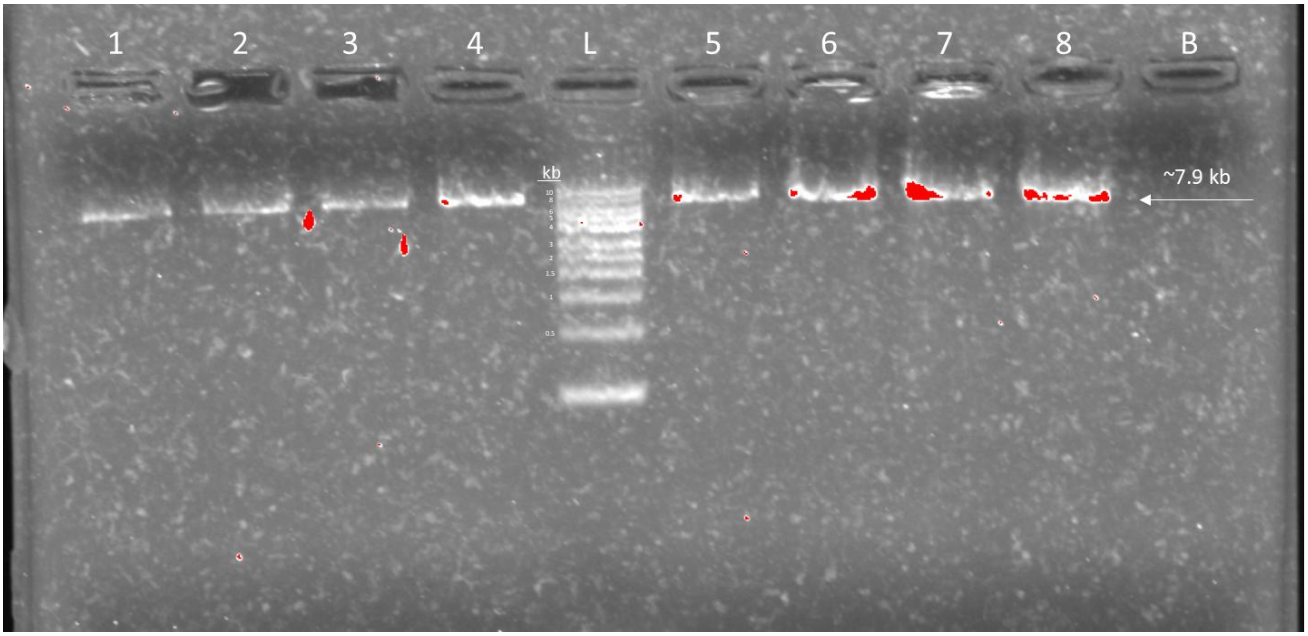


Figure 18. Agarose Gel of PCR Amplification from Site-Directed Mutagenesis of E29X.

Reaction #4 was chosen to continue forward with through the SDM protocol.

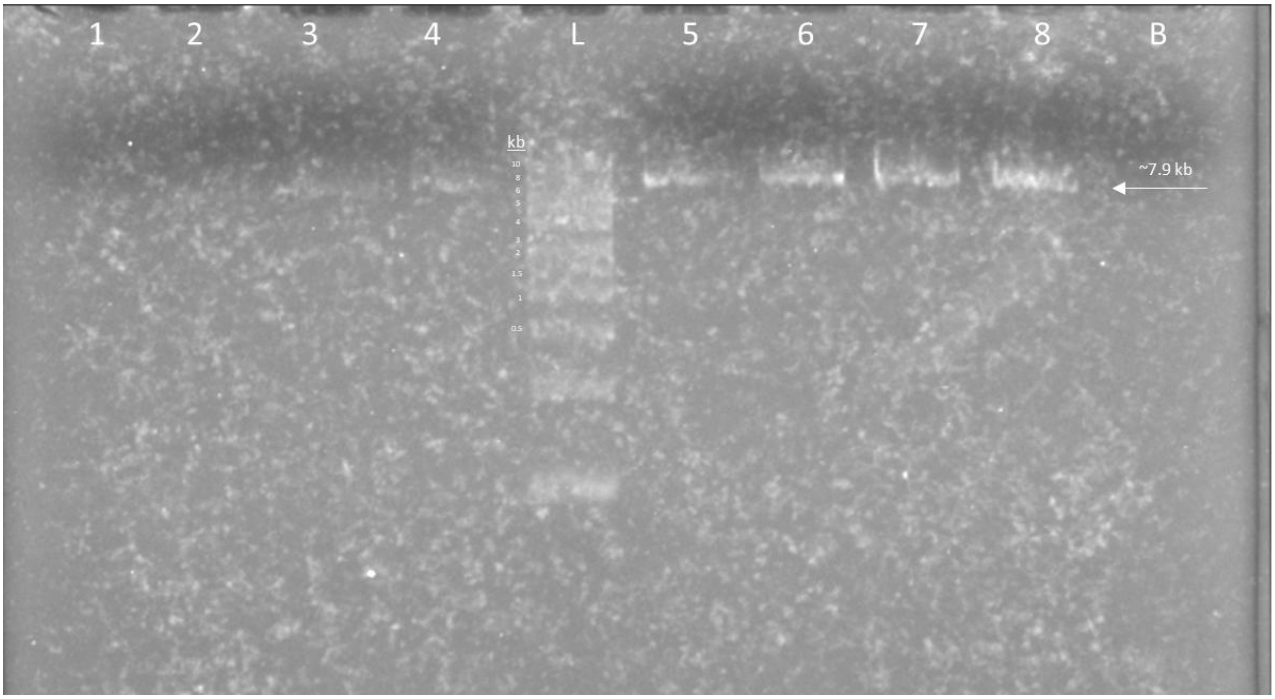


Figure 19. Agarose Gel of PCR Amplification from Site-Directed Mutagenesis of I195X. The gel image did not resolve very well, and it is difficult to see the DNA ladder, nonetheless reaction #6 was brought forward.

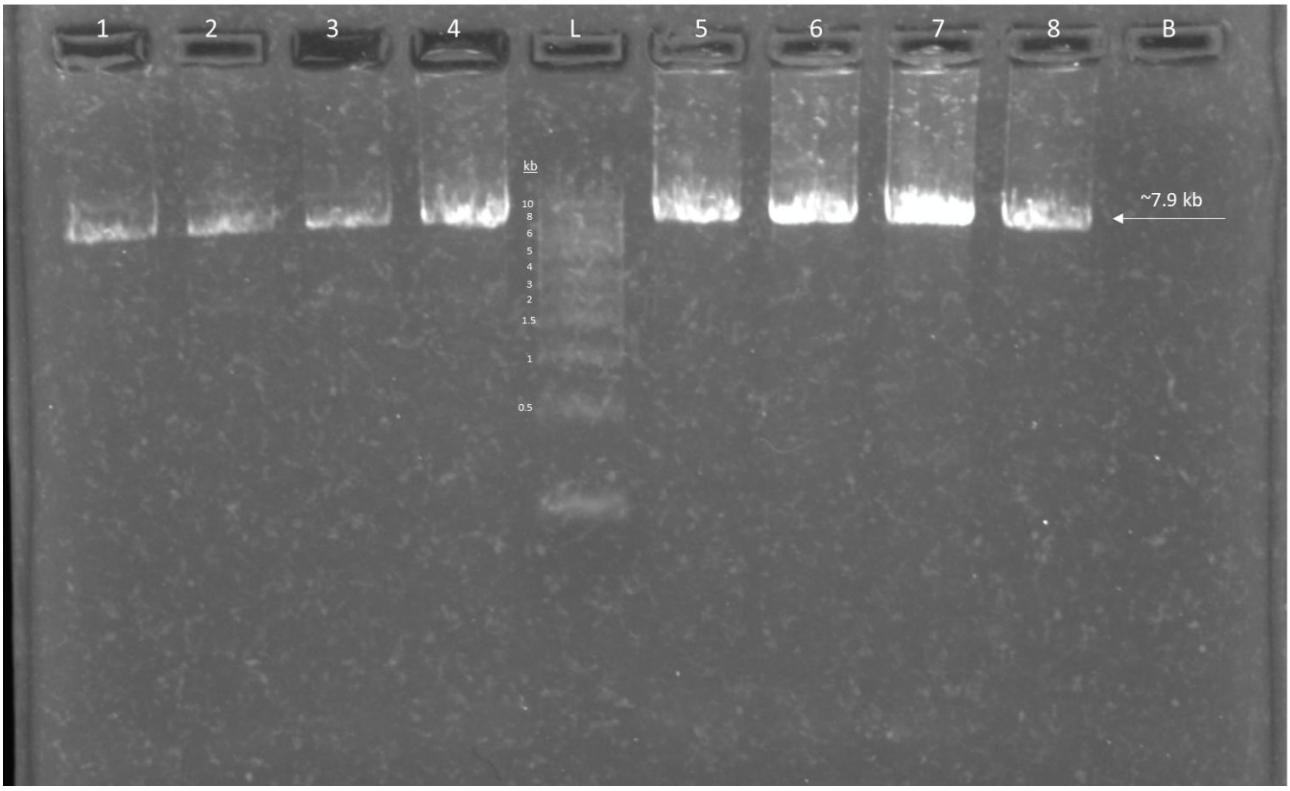


Figure 20. Agarose Gel of PCR Amplification from Site-Directed Mutagenesis of H134X.

Reaction #6 was chosen to continue forward with through the SDM protocol.

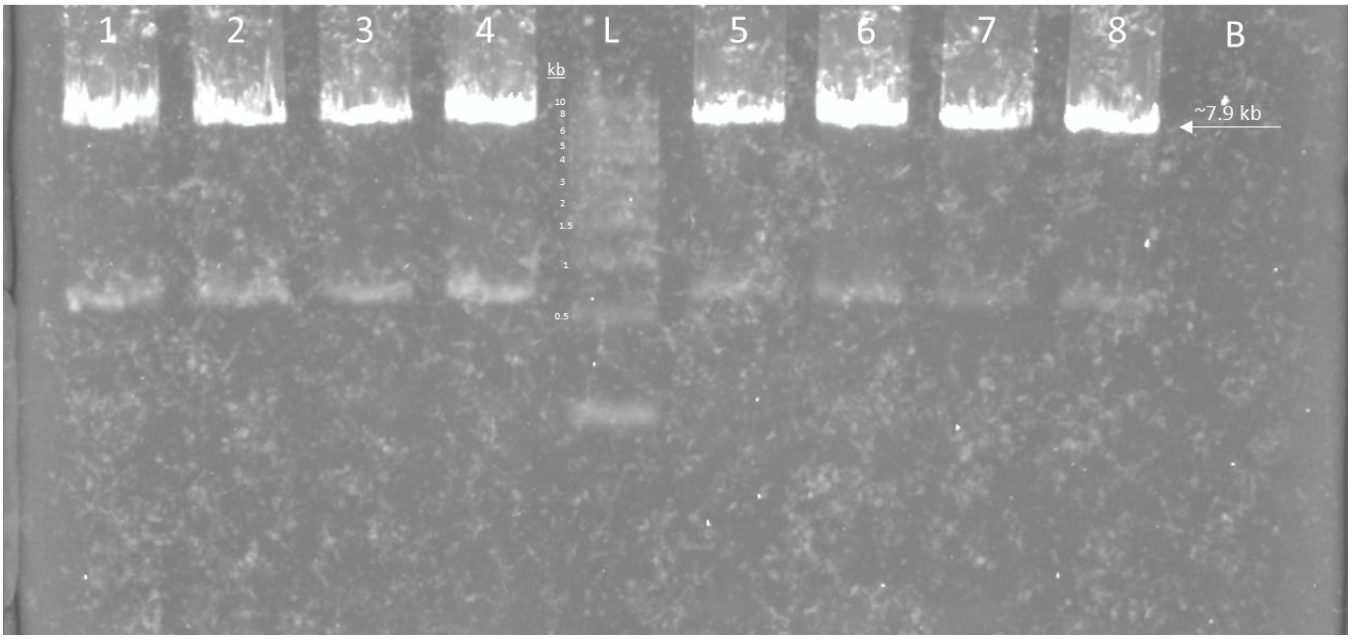


Figure 21. Agarose Gel of PCR Amplification from Site-Directed Mutagenesis of L114X.

Reaction #7 was chosen to be brought forward through the SDM protocol.

All six mutant residue test reactions survived on the Kanamycin LB-Agar plates, with no colonies forming on the negative plates. Five colonies each were picked from each test plate and, the DNA was extracted and sent for Sanger sequencing. Results showed successful replacement of the natural codon for K30, R116, H134, I195, L114 and E29 with the amber codon, TAG. The next stage of the project was to use these mutant Cterm-pHis-hTG2 plasmids in conjunction with a plasmid that expresses an orthogonal tRNA/aaRS pair capable of incorporation of AzF into the amber stop codon in *E. coli*. The first system chosen for this task was the pDULE-pCNF system and the next section will describe our attempts to incorporate AzF using it.

3.2.2 AzF Incorporation Using the pDule-pCNF tRNA System

We first wished to test the expression of WT TG2 without the addition of the pDULE-pCNF system. This is useful because a general baseline of TG2 expression can be made before potentially reducing yields with GCE technology. WT TG2 was expressed according to the general expression protocol described in the experimental section. One change from the general expression protocol is that this expression was done with Rosetta (DE3) *E. coli* cells instead of BL21 (DE3) cells. The Rosetta cell line includes the Rosetta plasmid, a plasmid that expresses eukaryotic tRNAs that match with eukaryotic codons, as present in the hTG2 gene.

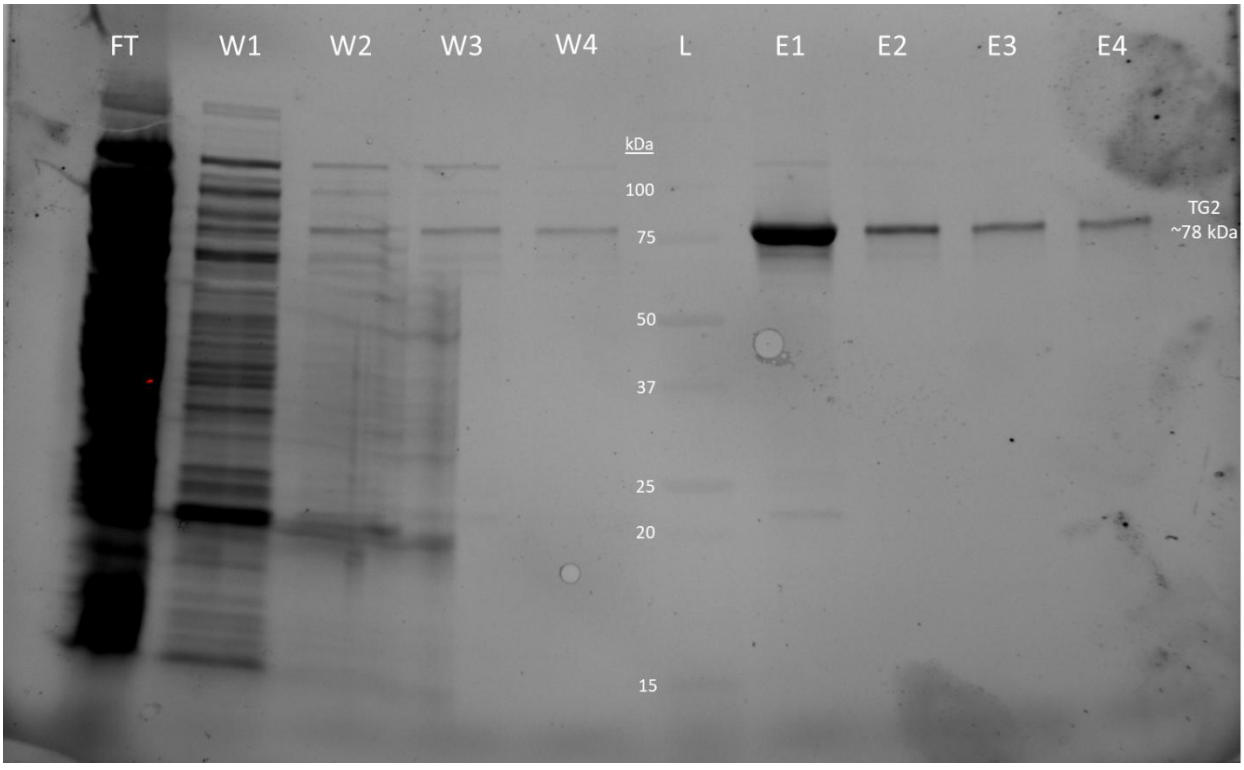


Figure 22. SDS-PAGE of WT TG2 Expression Post Ni-NTA Purification. (FT) represents flowthrough of soluble lysate after incubation with Ni-NTA resin.

WT TG2 is clearly present in the elution fractions at around 78 kDa. The protein appears to be relatively pure, with a few additional bands at a much lower intensity. A yield of 15.74 mg/L and specific activity of 0.107 U/mg were measured for the enzyme obtained from this expression, using the Bradford and AL5 assays described in the experimental section. After establishing a baseline for WT TG2 expression, the next step was to attempt incorporation of AzF using GCE.

The pDULE-pCNF plasmid was co-transformed with either WT or mutant pHis-Cterm-hTG2 plasmids into BL21 (DE3) cells on tetracycline/kanamycin resistant agar plates. A negative control was created with only WT TG2 or with only pDULE-pCNF added. The pDULE-pCNF plasmid carries a tetracycline resistance gene while the Cterm-pHis-hTG2 plasmid carries a kanamycin resistance gene. Negative plates had no colonies and positive plate colonies were used to inoculate expressions of pDULE-pCNF with mutant or WT TG2, as shown below. We started with the expression of WT/pDULE-pCNF, mostly following the general expression method. Exceptions to general method for this expression include: IPTG induction at a final concentration of 2 μM instead of 1 mM due to a typographical error in the protocol used.

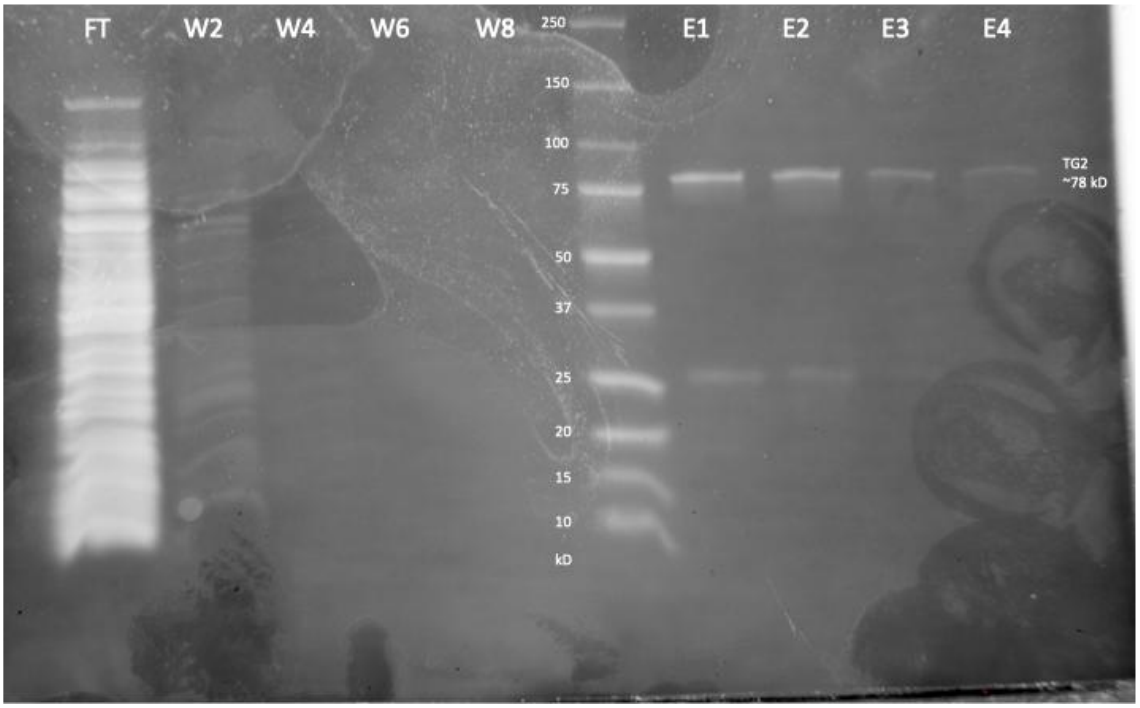


Figure 23. SDS-PAGE of WT/pDULE-pCNF Expression Post Ni-NTA Purification.

The yield for this expression was found to be 2.58 mg/L and the activity was 0.281 U/mg. The yield was reduced with the co-transformation with pDULE-pCNF but there was no decrease in the integrity of the protein, as evidenced by the specific activity. The reduction of the amount of IPTG used for induction may have contributed to the reduced yield of the protein. We next expressed L114X co-transformed with pDULE-pCNF using the general expression protocol for pDULE-pCNF AzF mutants. An exception to the general method for this expression was the formulation of AzF when added to the growth medium. For this expression, AzF was simply dissolved in 0.2 M HCl before being added to the expression culture at induction at a final concentration of 10 mM. Another exception is that the concentration of IPTG used to induce was still 2 μM instead of 1 mM.

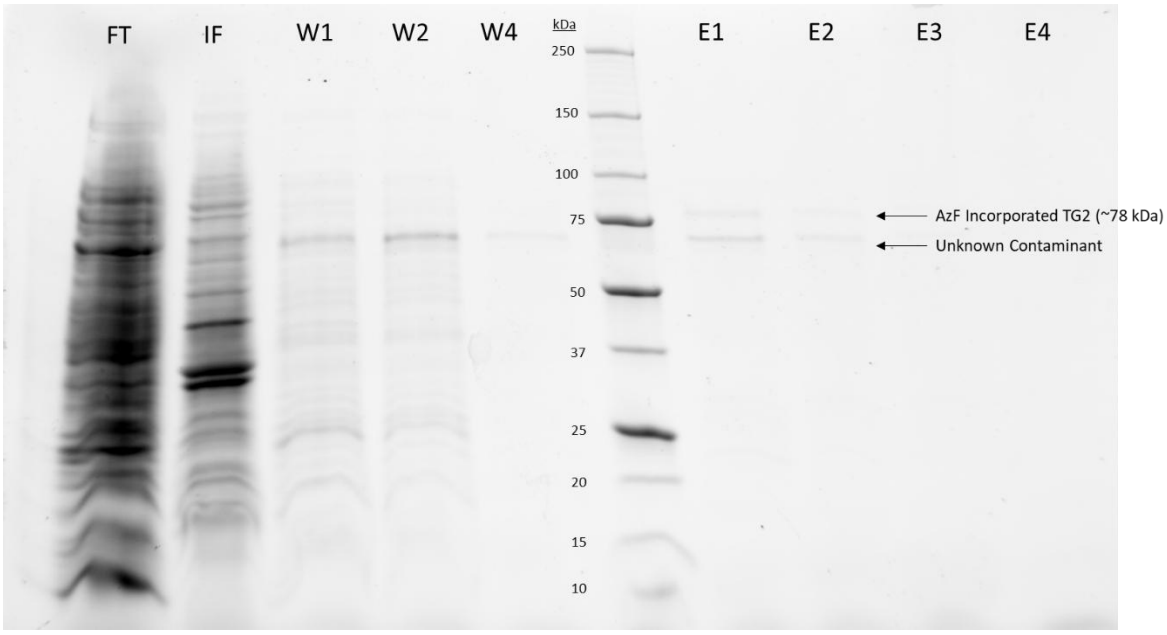


Figure 24. SDS-PAGE of L114X/pDULE-pCNF Expression Post Ni-NTA Purification. (IF) denotes the insoluble fraction which represents insoluble proteins after separation of the cell lysate by centrifugation.

Although TG2 is present on the SDS-PAGE, this expression resulted in a low yield of 0.82 mg/L and almost no activity was detected. It is still unclear whether this decrease in expression was caused by the low concentration of IPTG during induction or from the GCE system itself.

We next tried expression of I195X with the pDULE-pCNF system, where the formulation of AzF added to the induction culture for this expression followed the general method for pDULE-pCNF expression as described in the experimental section. The only exception is that the final concentration of IPTG remained low (2 μM).

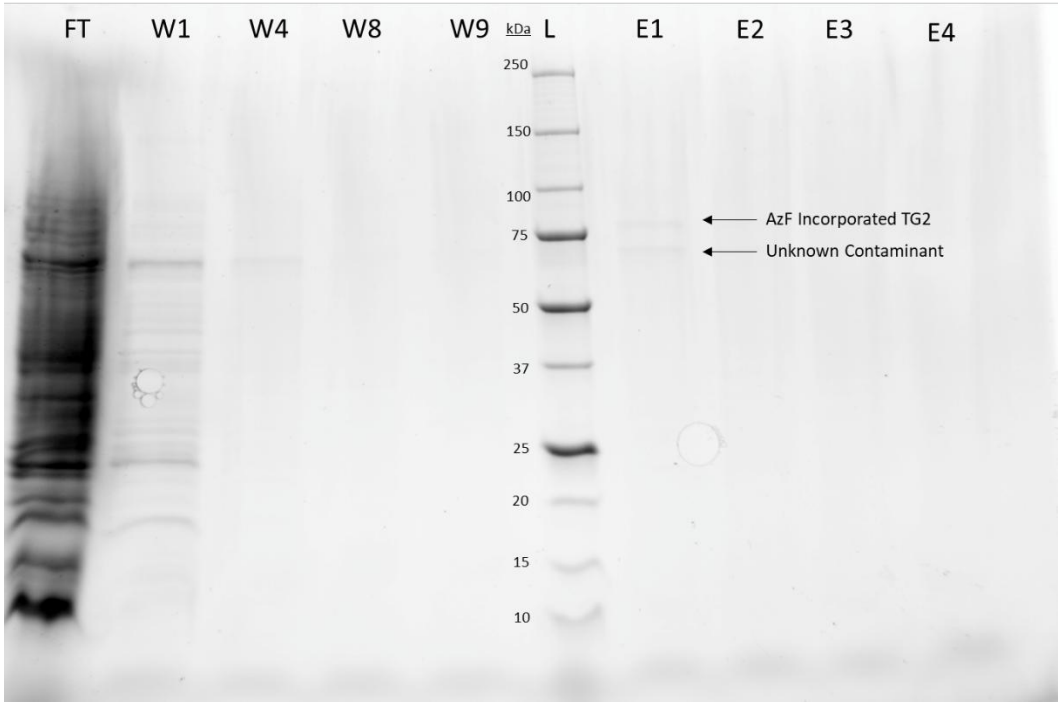


Figure 25. SDS-PAGE of I195X/pDULE-pCNF Expression Post Ni-NTA Purification.

The I195X expression had a low yield of 0.18 mg/L, similar to the L114X expression, which could have been due to the use of the pDULE-pCNF plasmid or the low IPTG induction concentration once again. The activity was also quite low, at 0.015 U/mg.

After this expression, the error with the low IPTG concentration at induction was recognized and for subsequent expressions the final concentration was increased from 2 μM to 1 mM. Additionally, for subsequent expressions, the lysis buffer used was changed to buffer used for TG2 expression in BL21 (DE3) versus expression in Rosetta (DE3) cells. The L114X/pDULE-pCNF expression was repeated following the general pDULE-pCNF expression protocol.

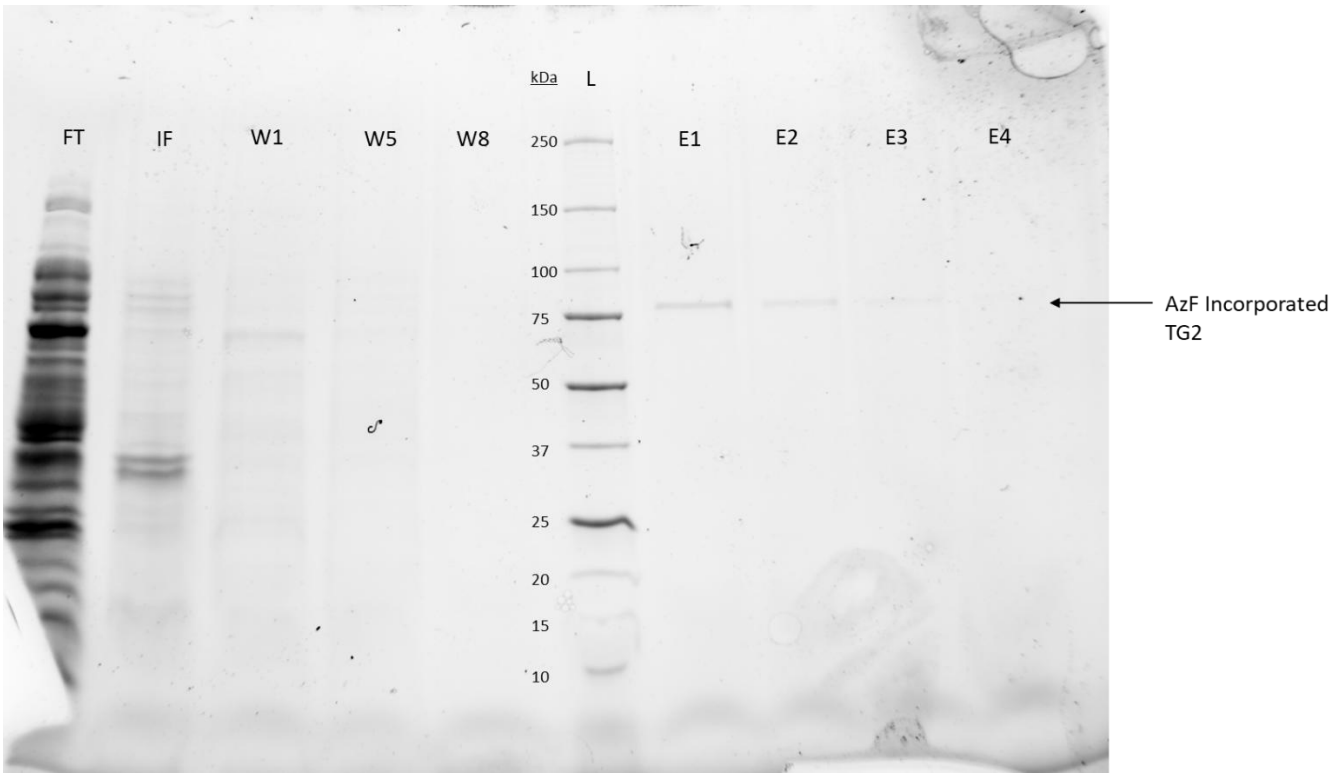


Figure 26. SDS-PAGE of L114X/pDULE-pCNF Expression #2 Post Ni-NTA Purification.

Although, according on the SDS-PAGE shown in Figure 26, it seems as if the level of expression of the L114X mutant was increased by the changes made to the protocol, namely the increase in IPTG concentration and the change in lysis buffer, the Bradford and AL5 assays showed no protein present and no AL5 hydrolysis activity. This was surprising and the expression was repeated under the same conditions to validate the lack of protein present.

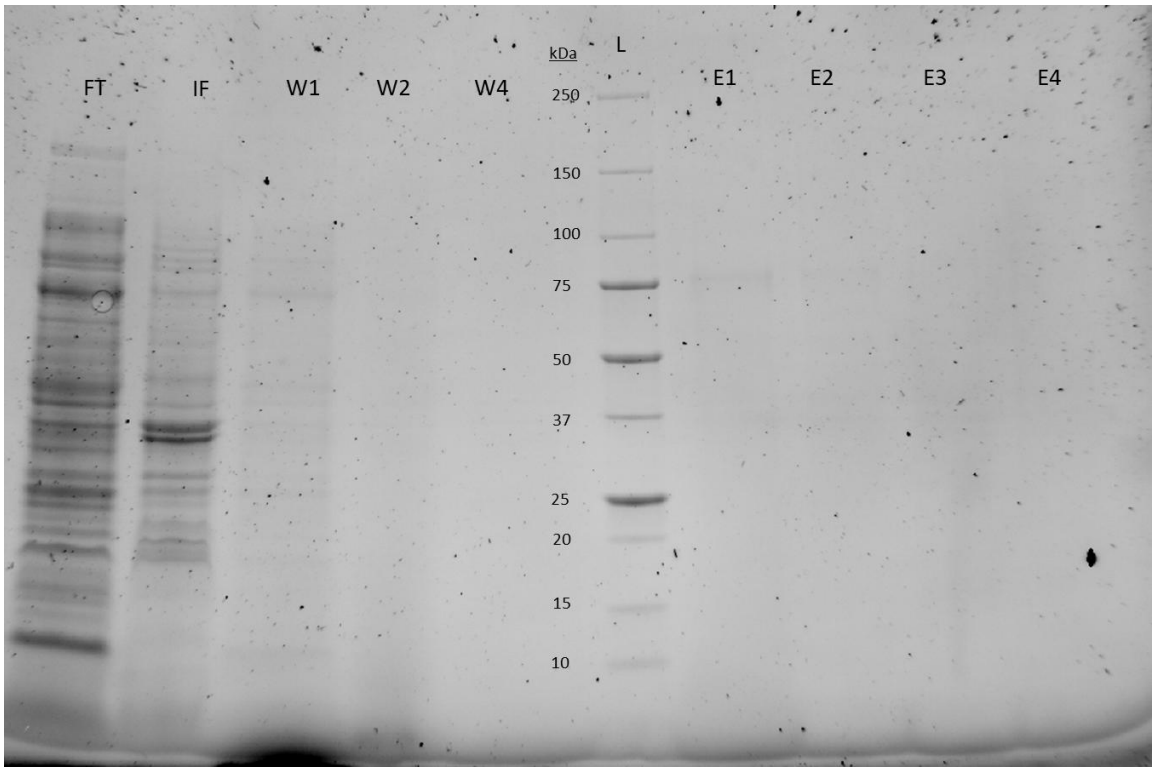


Figure 27. SDS-PAGE of L114X/pDULE-pCNF Expression #3 Post Ni-NTA Purification.

When the experiment was repeated, a very faint TG2 band was seen on the SDS-PAGE but a higher yield of 0.14 mg/L TG2 was measured, and the observed specific activity of 0.054 U/mg was comparable to that of the WT enzyme. We then attempted expressions of two more TAG mutants including, H134X and K30X. The SDS-PAGEs from these expressions are shown below.

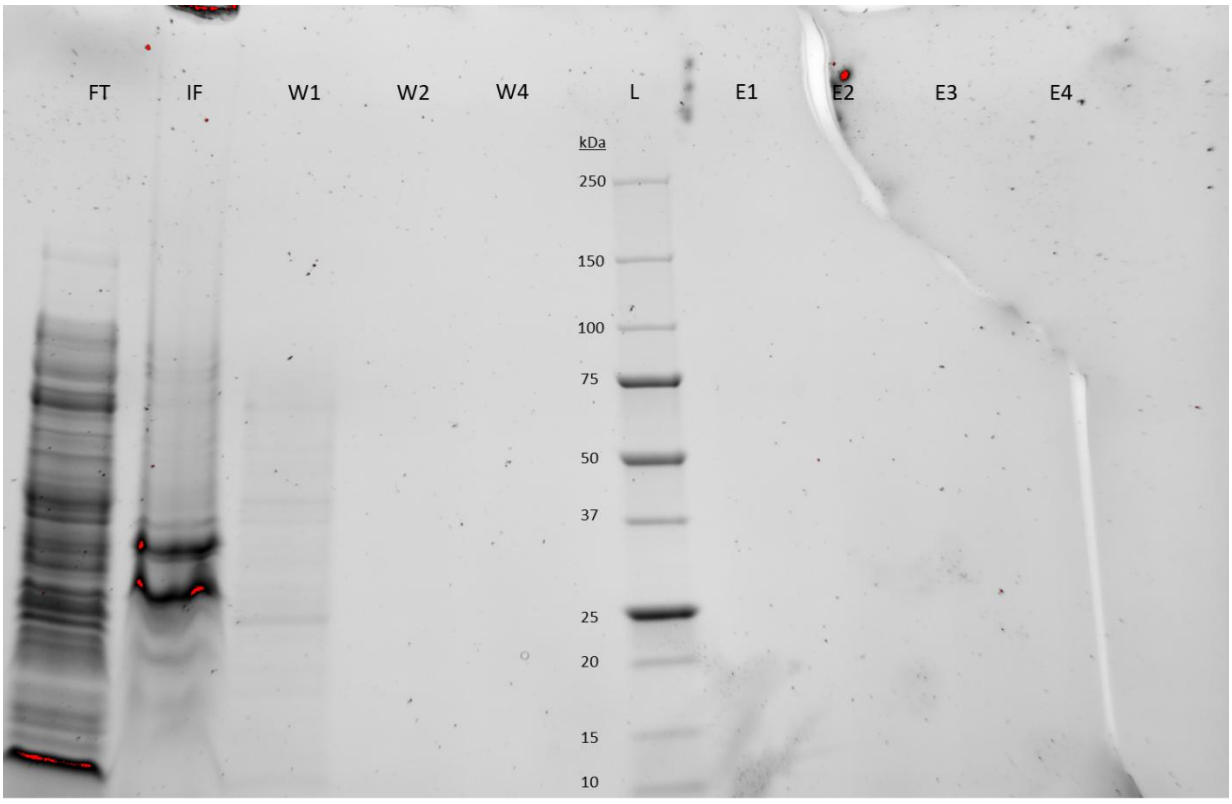


Figure 28. SDS-PAGE of H134X/pDULE-pCNF Expression Post Ni-NTA Purification.

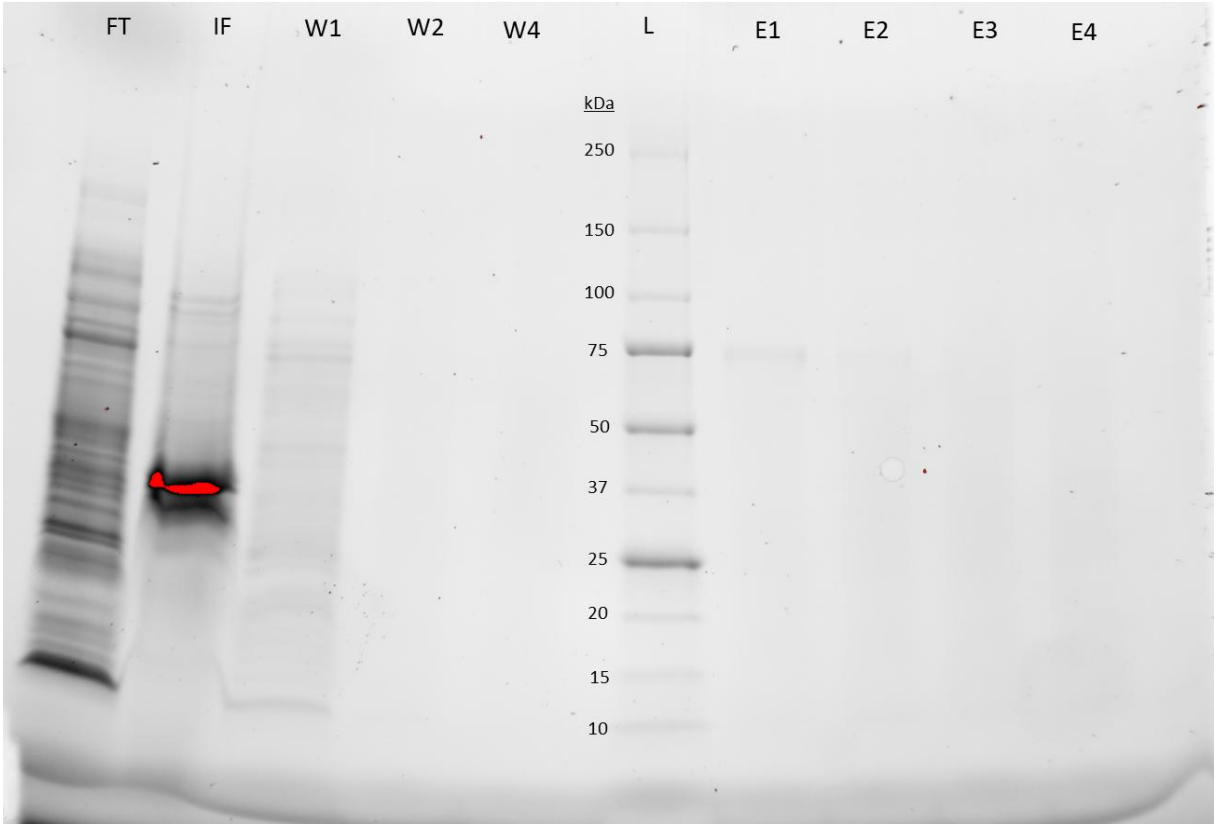


Figure 29. SDS-PAGE of K30X/pDULE-pCNF Expression Post Ni-NTA Purification.

No protein was seen on the SDS-PAGE for the H134X expression, and a very faint band was seen for the K30X expression. For both expressions, no protein or activity were detected by the Bradford and AL5 assays. To conclude our attempts of pDULE-pCNF expressions, another WT/pDULE-pCNF expression was completed following the general method for WT expression.

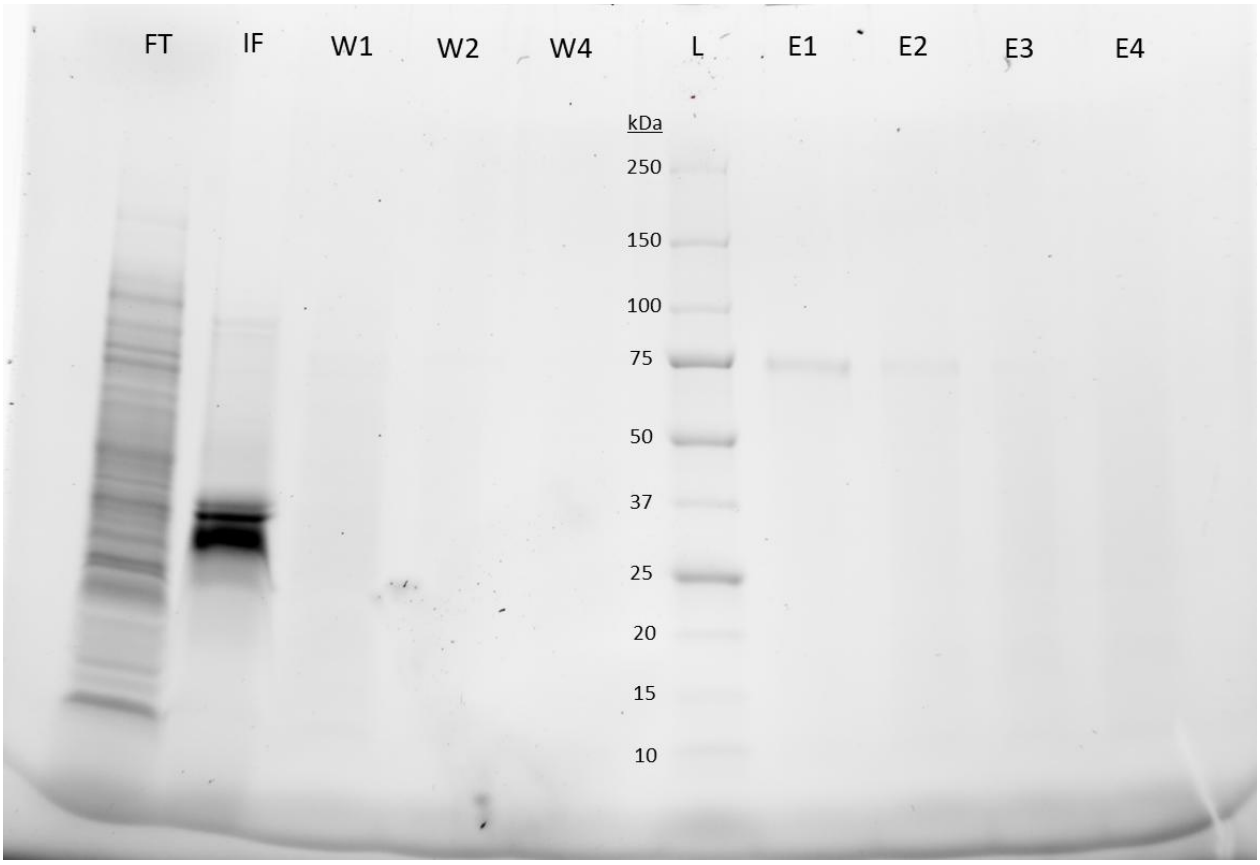


Figure 30. SDS-PAGE of WT/pDULE-pCNF Expression #2 Post Ni-NTA Purification.

The yield detected for this expression was extremely low, at 0.02 mg/L, and the resulting specific activity was very high, 2.42 U/mg potentially due to the very low concentration of protein amplifying the specific activity through the calculations. This WT/pDULE-pCNF expression is in sharp contrast to WT expressed on its own (Figure 22) and the previous WT/pDULE-pCNF expression (Figure 23). Table 3 presents the concentrations, yields and specific activities found for each expression from this section.

Table 3. Summary of pDULE-pCNF Expression Concentrations, Yields and Activities. The concentrations were measured using the Bradford assay and the specific activities were measured using the AL5 assay. BLQ stands for below limit of quantification.

Protein	Concentration (mg/mL) (mL)	Yield (mg/L)	Activity (U/mg)
WT (Fig 22)	5.03 ± 0.78 (1.57)	15.74 ± 2.44	0.107 ± 0.018
WT/pDULE-pCNF #1 (Fig 23)	0.49 ± 0.15 (2.64)	2.58 ± 0.81	0.281 ± 0.023
L114X/pDULE-pCNF #1 (Fig 24)	0.53 ± 0.11 (0.41)	0.81 ± 0.17	0.008 ± 0.006
I195X/pDULE-pCNF #1 (Fig 25)	0.79 ± 0.03 (0.12)	0.16 ± 0.01	0.015 ± 0.001
L114X/pDULE-pCNF #2 (Fig 26)	BLQ	BLQ	BLQ
L114X/pDULE-pCNF #3 (Fig 27)	0.30 ± 0.11 (0.24)	0.13 ± 0.05	0.054 ± 0.017
H134X/pDule-pCNF (Fig 28)	BLQ	BLQ	BLQ
K30X/pDULE-pCNF (Fig 29)	BLQ	BLQ	BLQ
WT/pDULE-pCNF #2 (Fig 30)	0.01 ± 0.05 (2.0)	0.02 ± 0.20	2.42 ± 0.01

Table 3 summarizes the results from the attempt at incorporation of AzF using the pDULE-pCNF system. The overall yields of the proteins expressed with pDULE-pCNF are low

and in many cases activity could not be quantified. A few expression parameters were changed during the troubleshooting of the expression experiment, including fixing the inducing concentration of IPTG by raising it from a 2 μM final concentration to 1 mM, changing the formulation of AzF added at induction, and changing the lysis buffer used. Despite these changes, the expression of WT/pDULE-pCNF appeared to get even worse as shown by the poor yield of the second expression shown in Figure 30 compared to Figure 23.

Overall, the yields were below 1 mg/L for the pDULE-pCNF expressions, except for the first two expressions, which were of WT with and without pDULE-pCNF. The specific activities are high for these first two expressions, but the rest are relatively low. The third expression of L114X had the most activity of these, about half of the WT enzyme from the first expression.

The expressions shown in Figures 25, 26 and 27 were plagued by the problem of low IPTG induction concentration, at 2 μM instead of 1 mM. Interestingly, increasing the concentration did not result in dramatic changes to the yield of protein produced. The formulation of AzF was also changed to follow the protocol used in the Pezacki Lab protocol for adding AzF (Figure 24). No noticeable changes were seen, but this method of preparation should be suitable since it has been established in their laboratory. The lysis buffer was changed from the lysis buffer that is used to purify TG2 from Rosetta (DE3) cells to the lysis buffer used to purify TG2 in BL21 (DE3) cells (Figure 26). No noticeable change was seen in yield, but we decided to keep this change since it is more consistent with the cell type used.

Although some changes were made to the expression protocol that should theoretically increase the yield of protein, especially the increase in IPTG concentration, the yield of protein was not increased, and the contrast between the first expression with WT/pDULE-pCNF (Figure 24) and the last (Figure 30) is large, with a decrease from 2.58 mg/L to 0.02 mg/L. The first

expression, done with only WT TG2, with an inducing IPTG concentration of 1 mM in Rosetta (DE3) cells (Figure 22), suggests that including the pDULE-pCNF plasmid in the expression decreases the yield of protein and that Rosetta (DE3) cells may bolster production of WT TG2. However, Rosetta (DE3) cells cannot be used with GCE expressions since the overloading of antibiotic resistance genes (kanamycin, chloramphenicol and tetracycline) leads to difficulties with growth and selection of positive colonies. At this point, a different method was clearly required, so we chose to change the GCE expression system from pDULE-pCNF to pEVOL-pAzF. Our attempts at AzF incorporation using this new expression system are presented in the next subsection.

3.2.3 AzF Incorporation Using the pEVOL-pAzF tRNA System

A different tRNA system, namely the pEVOL-pAzF expression system, was ordered to hopefully improve the incorporation and expression of TG2 AzF mutant proteins. The main differences between the pEVOL-pAzF and pDULE-pCNF systems is that the pEVOL system has an inducible promoter while the pDULE-pCNF has a constitutive promoter. The constitutive promoter in the pDULE-pCNF constantly produces protein, while the inducible promoter needs to be activated. For pEVOL-pAzF, activation occurs by addition of 0.02% arabinose. pEVOL-pAzF also expresses two copies of the aaRS and one of the suppressor tRNA, compared to just one of each for the pDULE-pCNF. The hypothesis was that more AzF suppressor tRNA and aaRS would be more readily available for incorporation of AzF into TG2 if the induction of TG2 and the evolved tRNA/aaRS pair were synchronized.

We began by incorporation of AzF into K30X, following the general protocol for pEVOL-pAzF expression outlined in the experimental section.

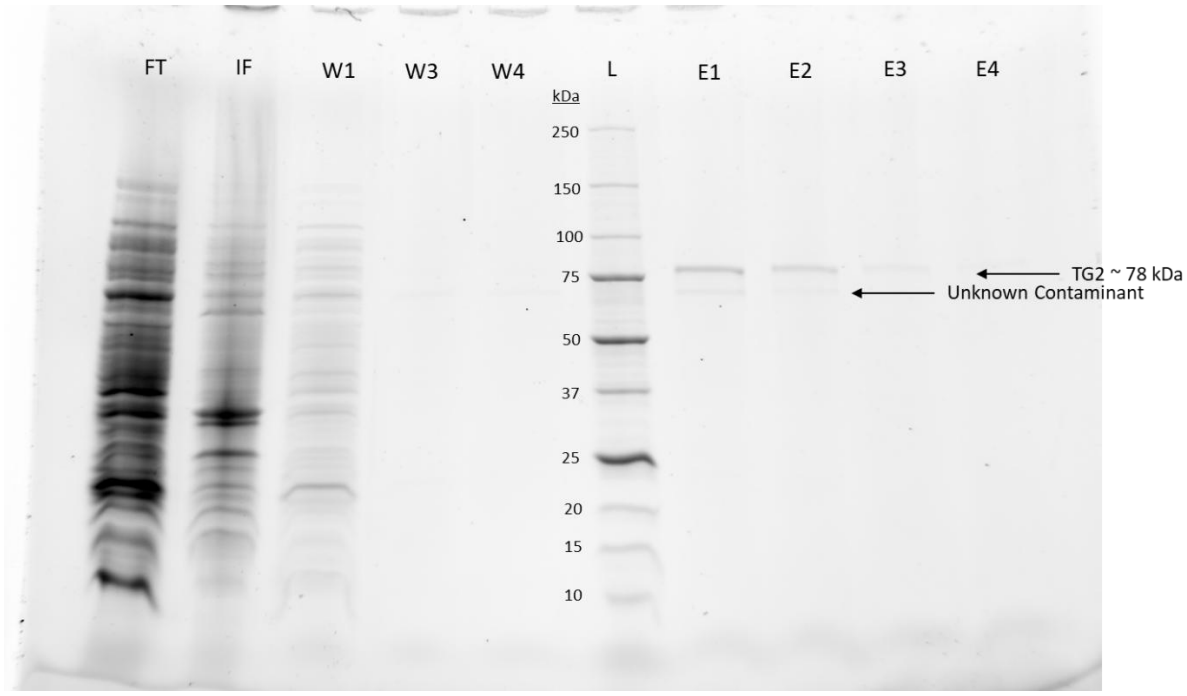


Figure 31. SDS-PAGE of K30X/pEVOL-pAzF Expression Post Ni-NTA Purification.

This expression resulted in a yield of 0.84 mg/L and a specific activity of 0.016 U/mg. The specific activity detected for the K30X/pEVOL-pAzF expression was quite low, indicating that the structure of TG2 may have been perturbed by the K30X mutation. Nonetheless, these results were promising, since expression of K30X/pDULE-pCNF, resulted in no protein or activity detected. We next performed expression of WT/pEVOL-pAzF TG2 alongside L114X/pEVOL-pAzF TG2. Performing these expressions at the same time increases the integrity of comparing between them, since protein expression from day to day can vary significantly.

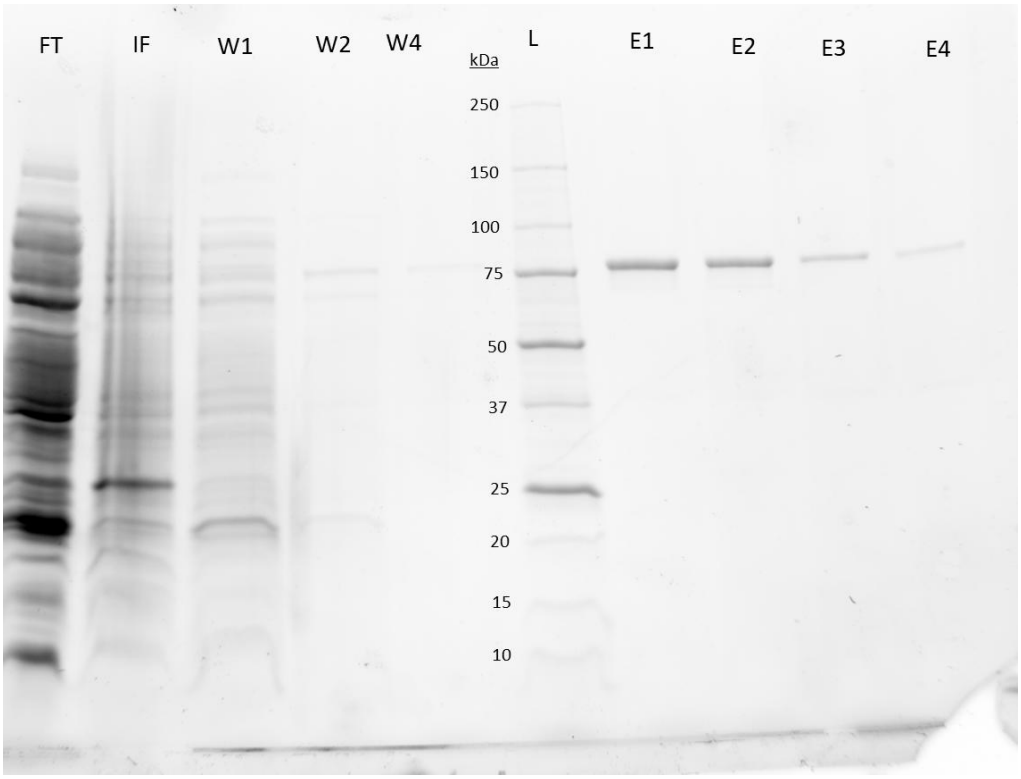


Figure 32. SDS-PAGE of WT/pEVOL-pAzF Expression Post Ni-NTA Purification.

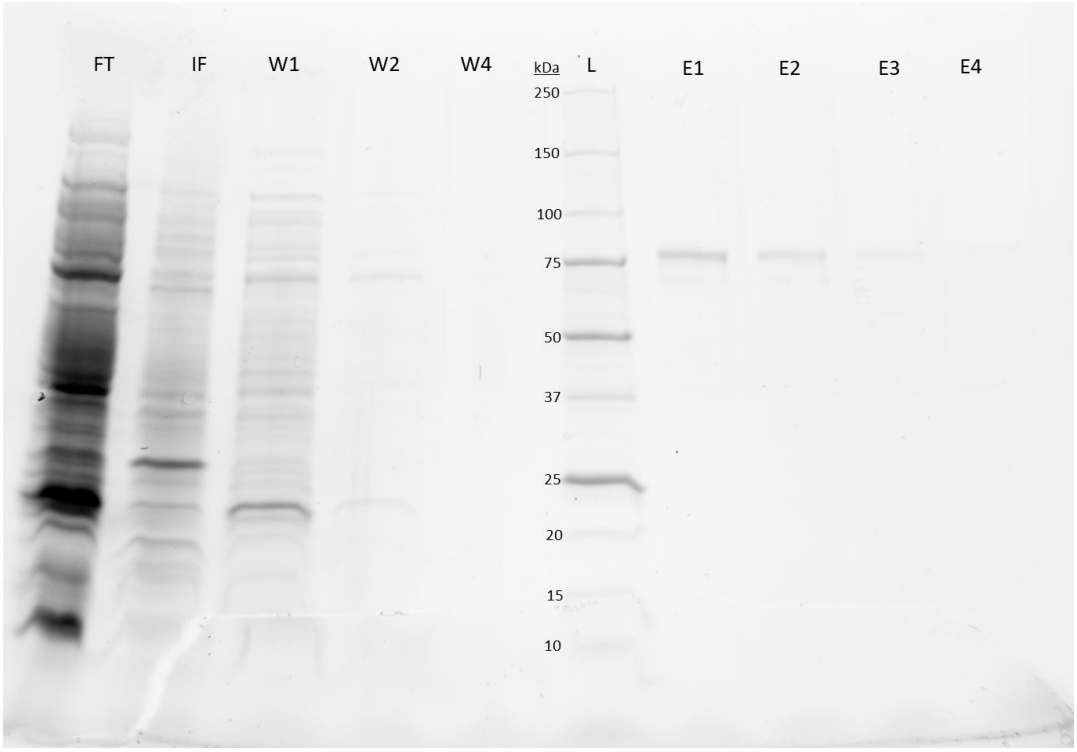


Figure 33. SDS-PAGE of L114X/pEVOL-pAzF Expression Post Ni-NTA Purification.

Expressions of WT and L114X TG2 with the pEVOL-pAzF system resulted in yields of 2.85 mg/L and 1.61 mg/L respectively. This is a dramatic improvement over the yields seen with pDULE-pCNF. The activities of the proteins expressed were 0.180 U/mg and 0.044 U/mg, suggesting that the integrity of L114X may have been affected by the incorporation of AzF. The pEVOL-pAzF plasmid was co-transformed with other mutant plasmids and expression of these mutants were carried out with the pEVOL-pAzF system. Mutants I195X, E29X, K30X and H134X were expressed using the pEVOL-pAzF system.

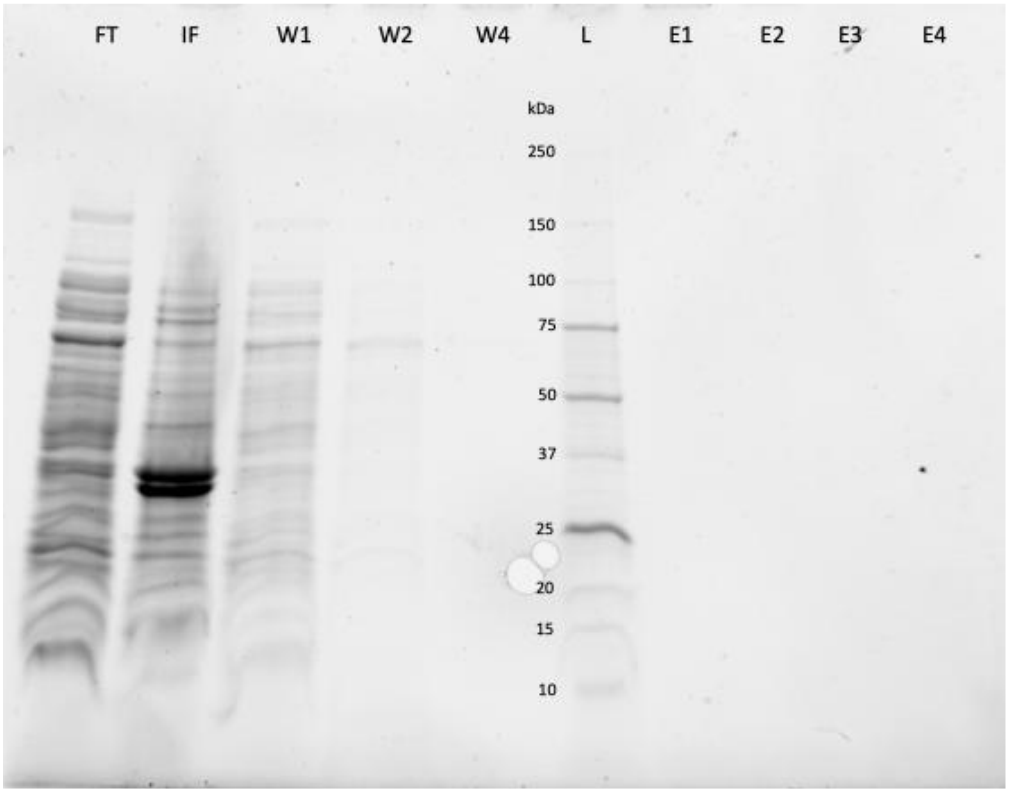


Figure 34. SDS-PAGE of I195X/pEVOL-pAzF Expression Post Ni-NTA Purification.

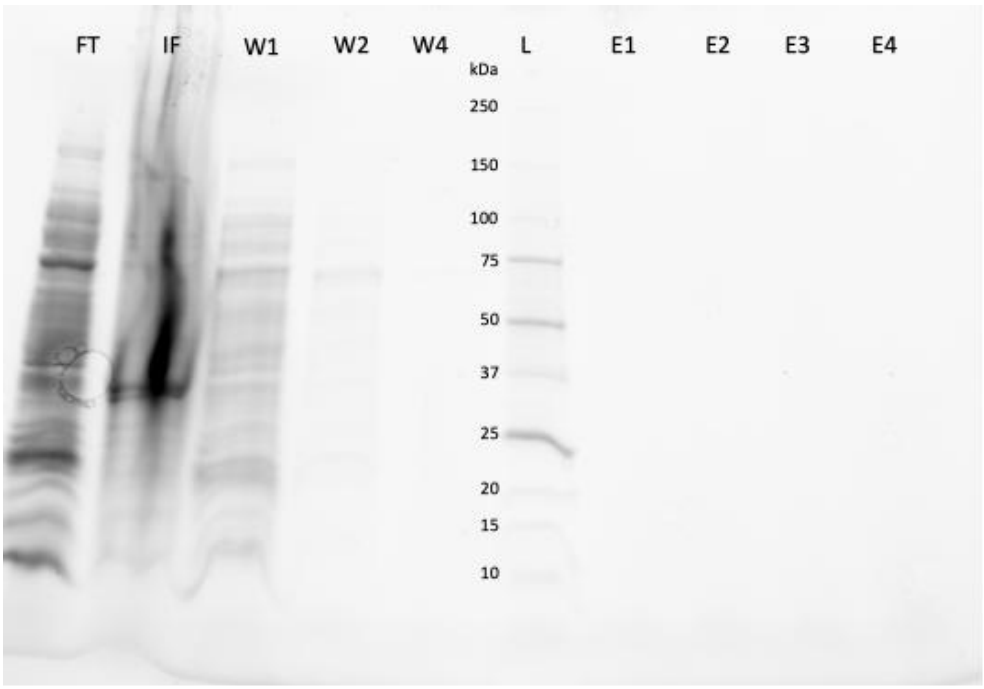


Figure 35. SDS-PAGE of E29X/pEVOL-pAzF Expression Post Ni-NTA Purification.

The expressions of I195X and E29X resulted in no protein expressed. H134X and K30X were expressed alongside each other next.

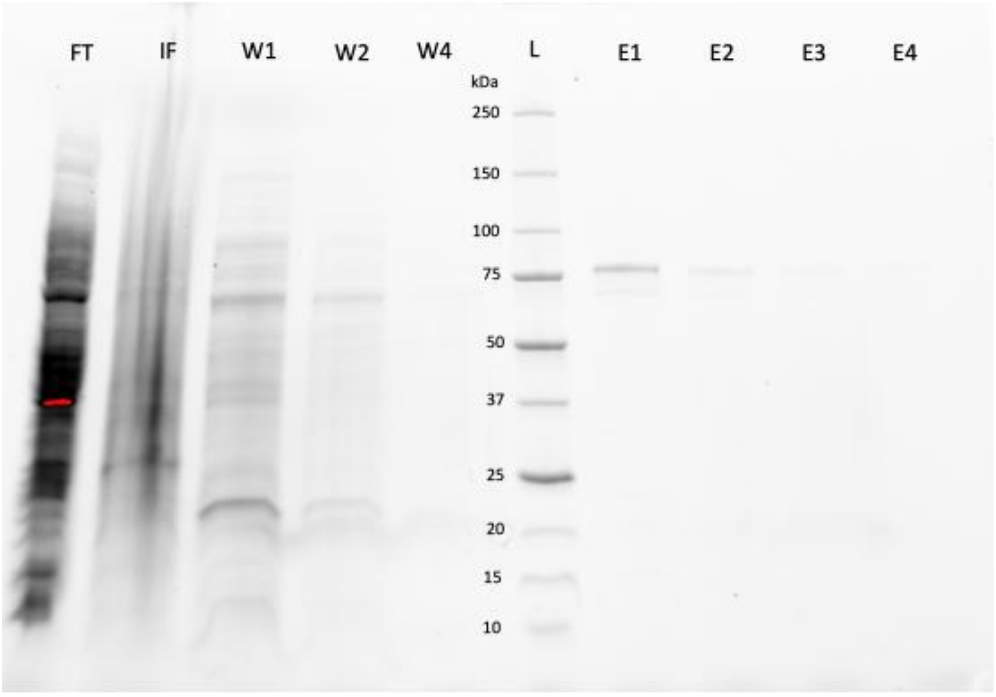


Figure 36. SDS-PAGE of K30X/pEVOL-pAzF Expression #2 Post Ni-NTA Purification.

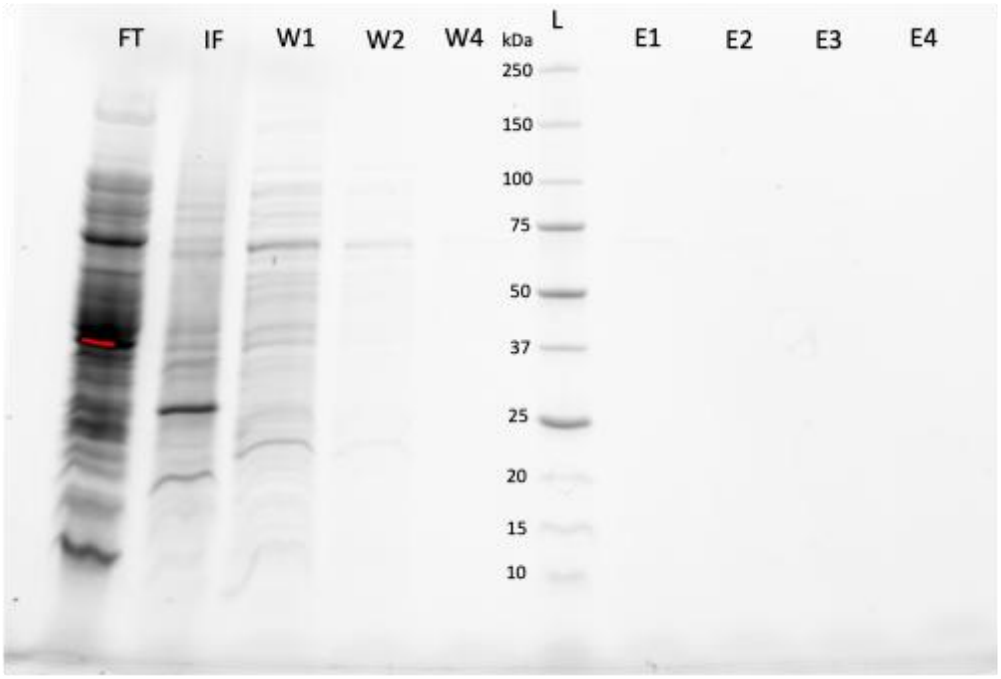


Figure 37. SDS-PAGE of H134X/pEVOL-pAzF Expression Post Ni-NTA Purification.

The re-expression of K30X with the pEVOL-pAzF system, resulted in both a higher yield of 1.93 mg/L and higher specific activity of 0.103 U/mg. In contrast to the first expression, the activity of K30X is comparable to the WT activity. This example shows that there is a high degree of variability in yield and activity between expressions, even when conditions are the same. In Figure 39, the H134X expression resulted in no protein present. We proceeded to then attempt AzF incorporation at position R116.

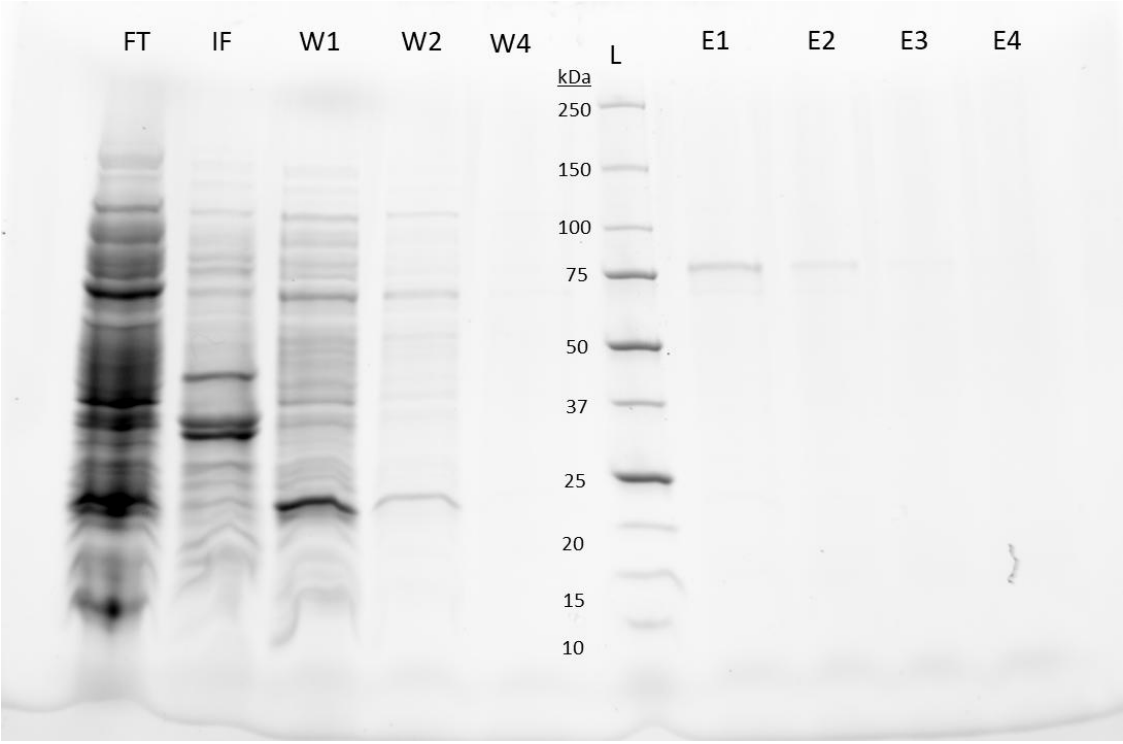


Figure 38. SDS-PAGE of R116X/pEVOL-pAzF Expression Post Ni-NTA Purification.

The R116X expression resulted in a comparable yield of 1.11 mg/L and a specific activity of 0.160 U/mg, indicating that the integrity of R116X is retained after the mutation. After the expression of R116X we once again repeated the expression of K30X, to have another measure to balance the contrasting results seen in expression #1 and #2. This third K30X expression

resulted in a yield of 1.32 mg/L and a specific activity of 0.198 U/mg, corroborating the results found from the second expression.

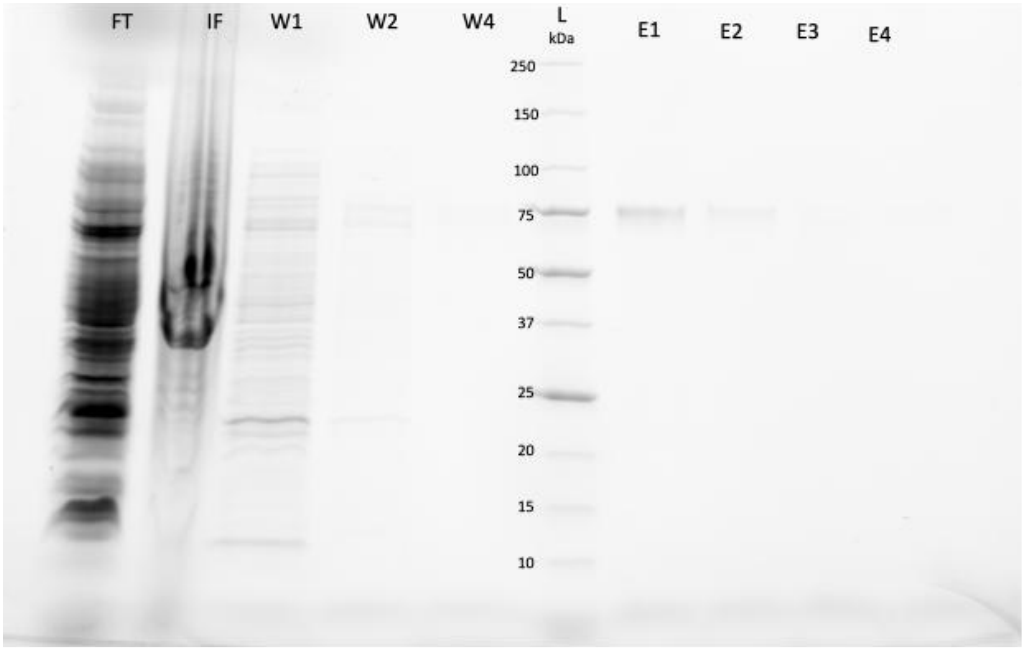


Figure 39. SDS-PAGE of K30X/pEVOL-pAzF Expression #3 Post Ni-NTA Purification.

A DBCO-Cy5 assay was then performed on proteins that were expressed throughout the course of the pEVOL-pAzF portion of the project. this assay theoretically detects incorporated AzF through an SPAAC-reaction with the DBCO portion of the molecule, which is attached to the fluorescent dye Cy5. This dye can be visualized using a gel imager, and the protocol for this experiment is described in the experimental section. NS3 helicase was provided by the Pezacki lab and acts as a positive control for the assay, since it had been previously confirmed by their group to have AzF incorporated. The negative control is ostensibly WT TG2 which should have no possible incorporation of AzF since it is lacking the TAG mutation with no AzF was added to the growth medium.

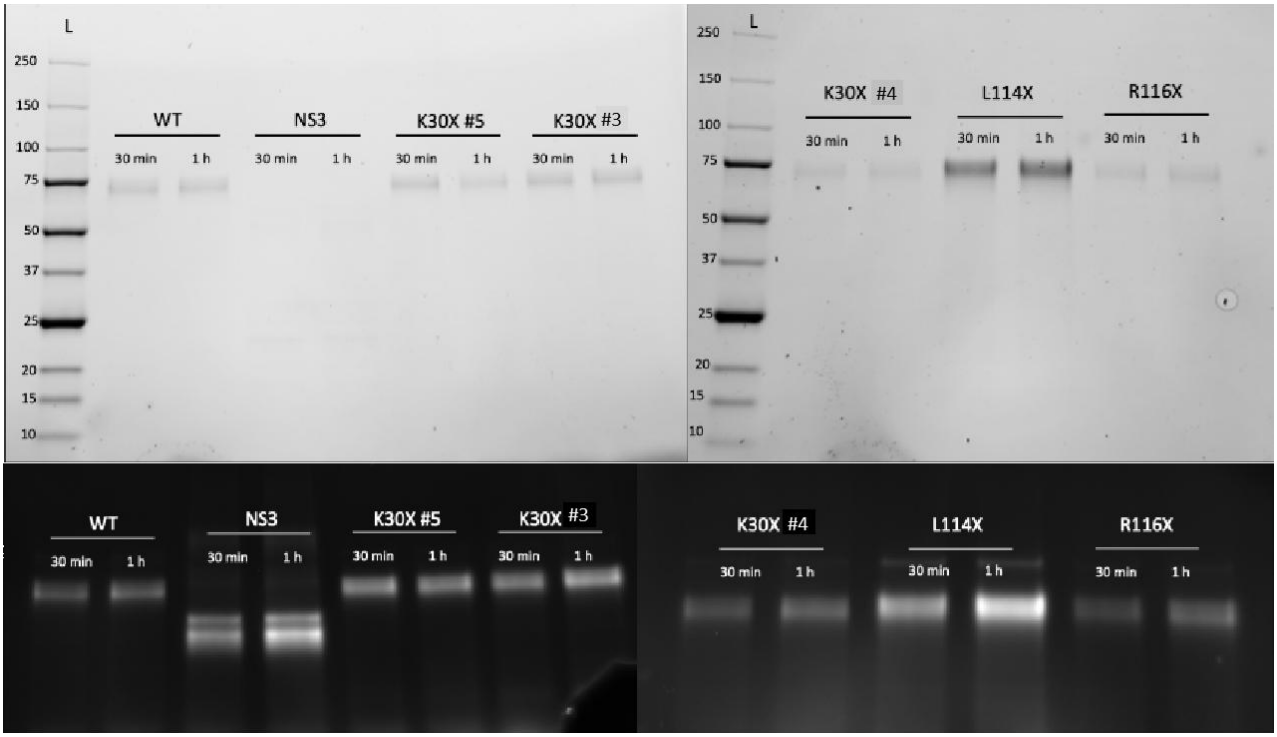


Figure 40. DBCO-Cy5 Assay for WT TG2 and pEVOL-pAzF Mutants. Top, stain-free visualization of SDS-PAGE. Bottom, Cy5 imaging of SDS-PAGE. K30X expression #3 is characterized in this thesis but K30X expressions #4, and #5 were deemed to be redundant to include in detail, however, they were tested in this DBCO-Cy5 assay. The time 30 min or 1 h corresponds to how long the proteins were reacted with DBCO-Cy5.

The brightness of bands on fluorescent imaging of the SDS-PAGE should correspond with the degree of AzF incorporation. If WT contains no AzF, it would appear that R116X also does not contain AzF, according to the similar brightness of its fluorescent band. The bands corresponding to the K30X expressions seem to be slightly brighter but there is a lack of increasing brightness at 1 hour compared to 30 minutes. The positive NS3 helicase control is quite bright and an increase in Cy5 labelling is noted at 1 h versus 30 minutes, which is also seen with the L114X protein. The NS3 helicase appears very strongly on the gel despite its faintness

on the unstained gel. The L114X mutant appears to be very bright but it also seems to have been loaded at a higher concentration when compared to other proteins on the stain-free SDS-PAGE. Considering these concerns, L114X was re-analyzed alone using DBCO-Cy5.

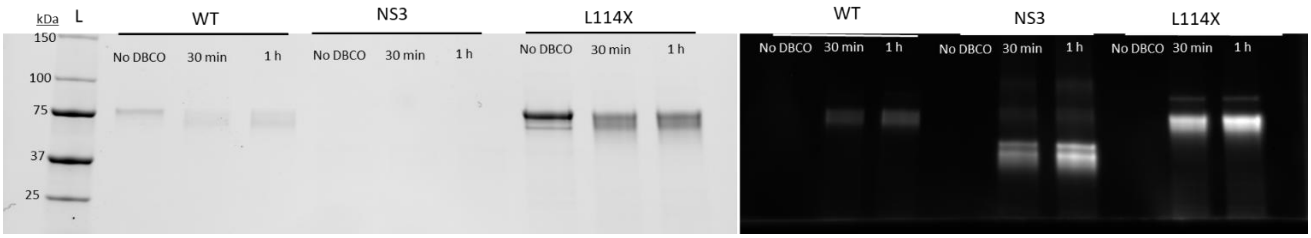


Figure 41. SDS-PAGE of DBCO-Cy5 Assay for WT TG2 and L114X. Left, stain-free imaging. Right, Cy5 imaging. NS3 represents the positive control NS3 helicase which was provided by the Pezacki Lab and has AzF incorporated.

The L114X mutant once again showed substantial brightness, although it is hard to draw concrete conclusions about the incorporation of AzF, due to the difference in loading of the proteins. Proteins were calculated to be loaded onto the SDS-PAGE in equal amounts according to their determined concentration but a large discrepancy between concentrations of proteins on the gel is noted. Nevertheless, L114X was deemed promising enough to continue with crosslinking assays with the 45 kDa Fibronectin fragment (45FN).

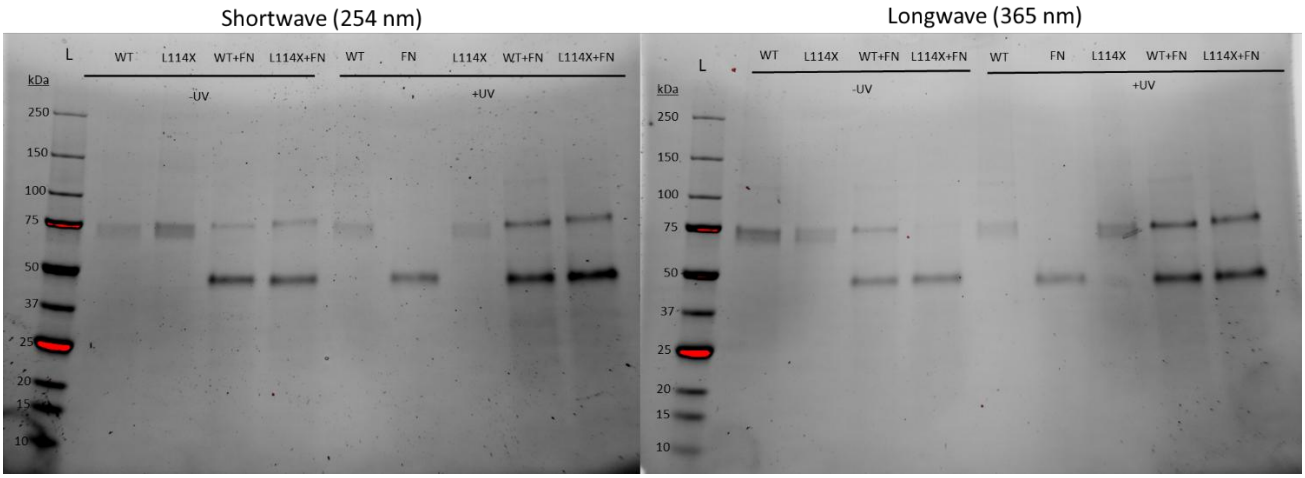


Figure 42. SDS-PAGEs of WT and L114X with 45FN by Short and Longwave Crosslinking. Left, shortwave crosslinking. Right, longwave crosslinking. Both were done in the absence and presence of UV light.

This 45FN WT (present at around 45 kDa) and L114X TG2 crosslinking assay was performed using both short wave (254 nm) and long wave (365 nm) light, according to the protocol described in the experimental section. For successful crosslinking, the 45FN-L114XTG2 complex should appear at around 123 kDa in the "L114X + FN" lanes, for either shortwave or longwave +UV portions of the gel. However, no band is seen in either shortwave or longwave, suggesting formation of the 45FN-L114XTG2 complex did not take place.

At this point in the project, we hypothesized that the crosslinking of AzF to 45-kDa FN may not be working for a few different reasons. Firstly, the efficiency of incorporation of AzF may not have been high enough for sufficient mutant enzyme to be activated to react with 45FN residues. Secondly, the residues we selected for mutation may have been too critical for the non-covalent association between 45FN and TG2, such that mutations of these residues led to too

much disruption for the protein-protein complex to form. Lastly, the crosslinking assay may need to be optimized by changing incubation time, temperature, or UV exposure duration.

The first option we decided to tackle was the second one, residue selection. We chose two new residues to mutate with the amber stop codon (TAG), A66 and F203. These residues are adjacent to our hypothesized fibronectin binding site but are not directly implicated in its association from site-directed mutagenesis experiments.

The mutagenesis done to create the mutated plasmids harbouring these residues is described in the Honours thesis of Maryam Kettal and adapted from that text in this work.[108] We expressed WT, F203X and A66X TG2 alongside each other following the standard protocol using the pEVOL-pAzF system.

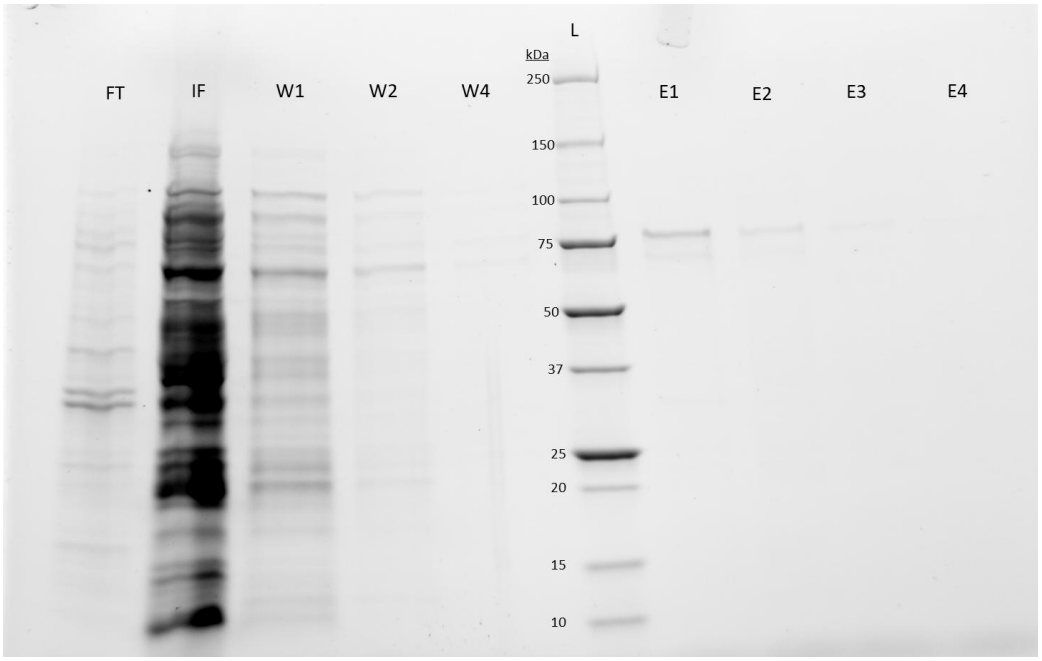


Figure 43. SDS-PAGE of WT/pEVOL-pAzF Expression #2 Post Ni-NTA Purification.

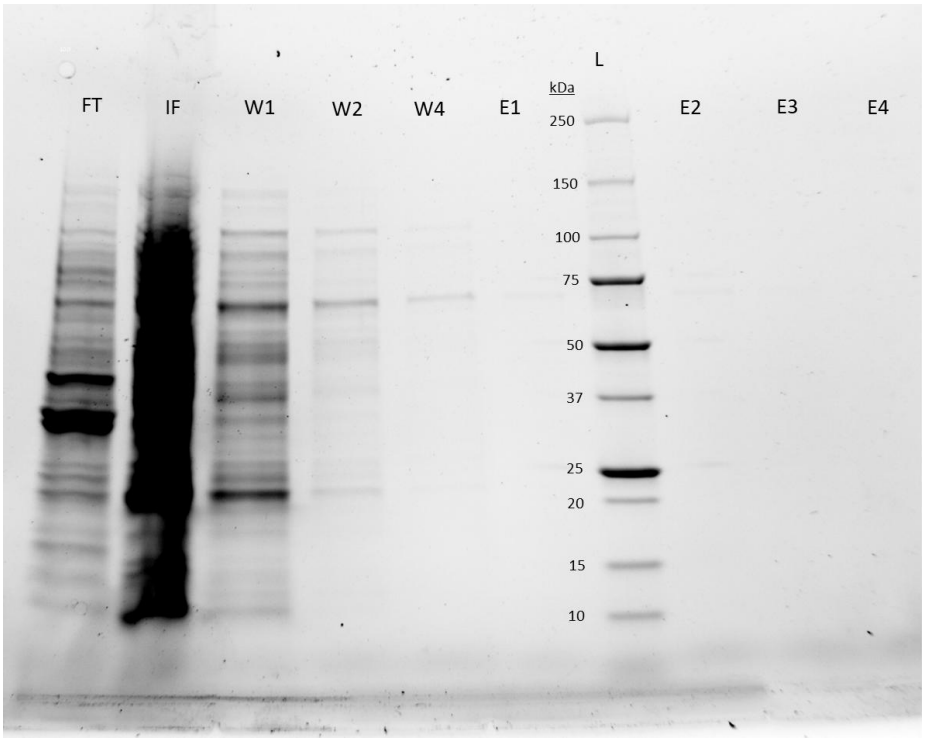


Figure 44. SDS-PAGE of A66X/pEVOL-pAzF Expression Post Ni-NTA Purification.

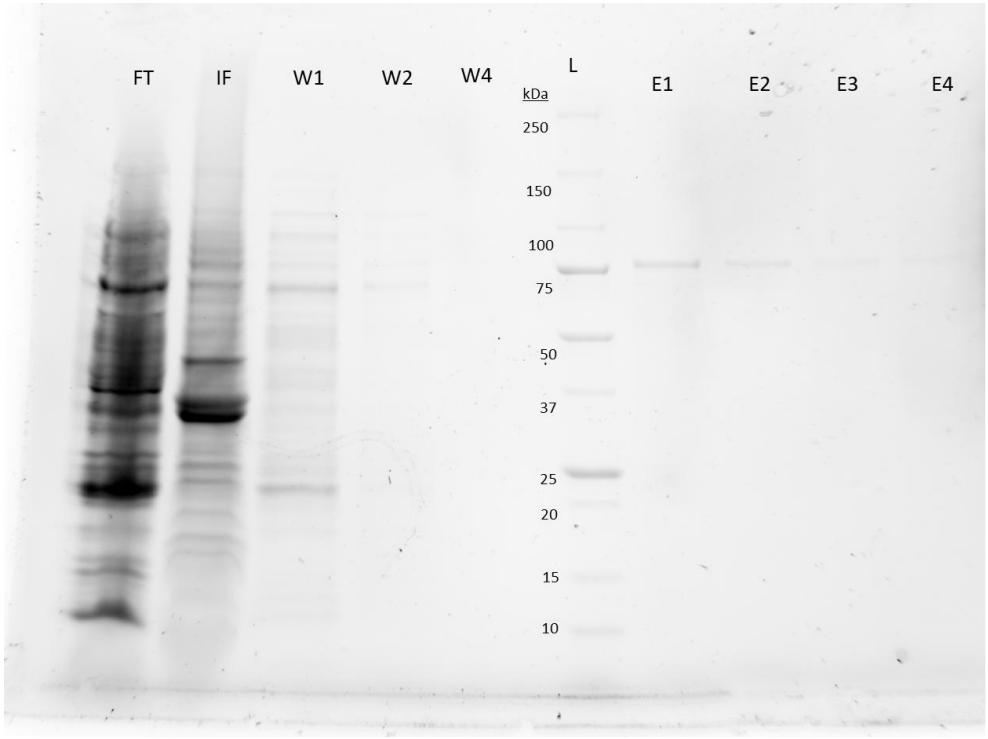


Figure 45. SDS-PAGE of F203X/pEVOL-pAzF Expression Post Ni-NTA Purification.

The WT, A66X and F203X expressions resulted in a yield of 3.24, 1.61 and 4.30 mg/L respectively. The F203X expression was comparable to the WT in terms of yield and activity, with F203X having an activity of 0.060 U/mg and WT having an activity of 0.025 U/mg. The A66X expression had a lower yield and was not active. This could be caused by the A66X mutation disrupting the overall structure of TG2, reducing its AL5 hydrolysis activity. F230X was selected to be advanced to DBCO-Cy5 and crosslinking experiments.

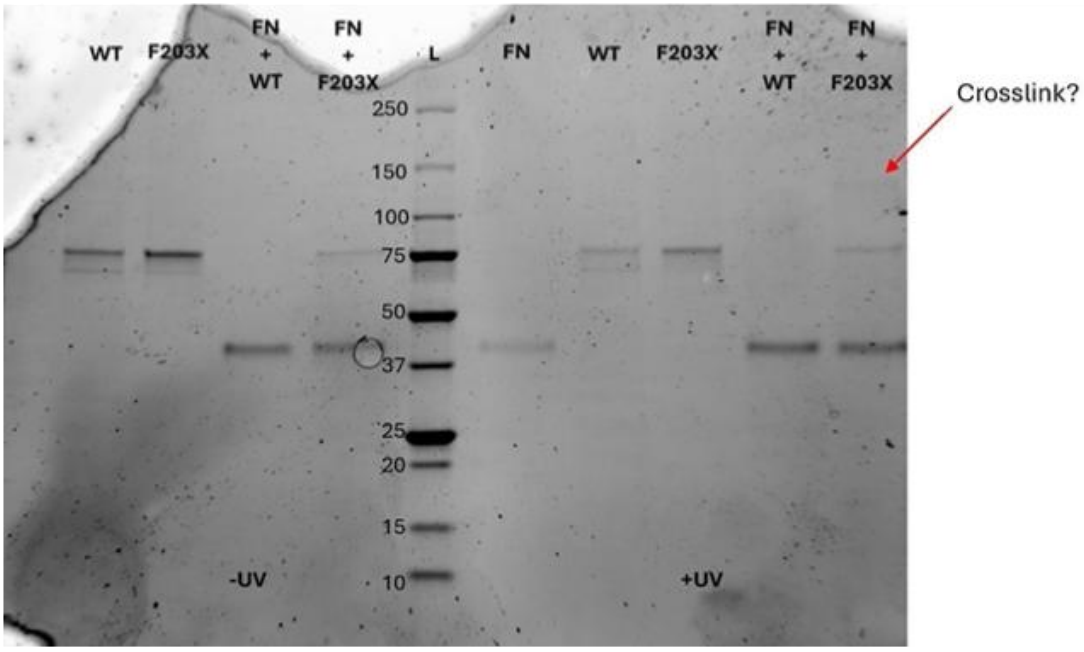


Figure 46. F203X and 45FN Long Wave Crosslinking Experiment.

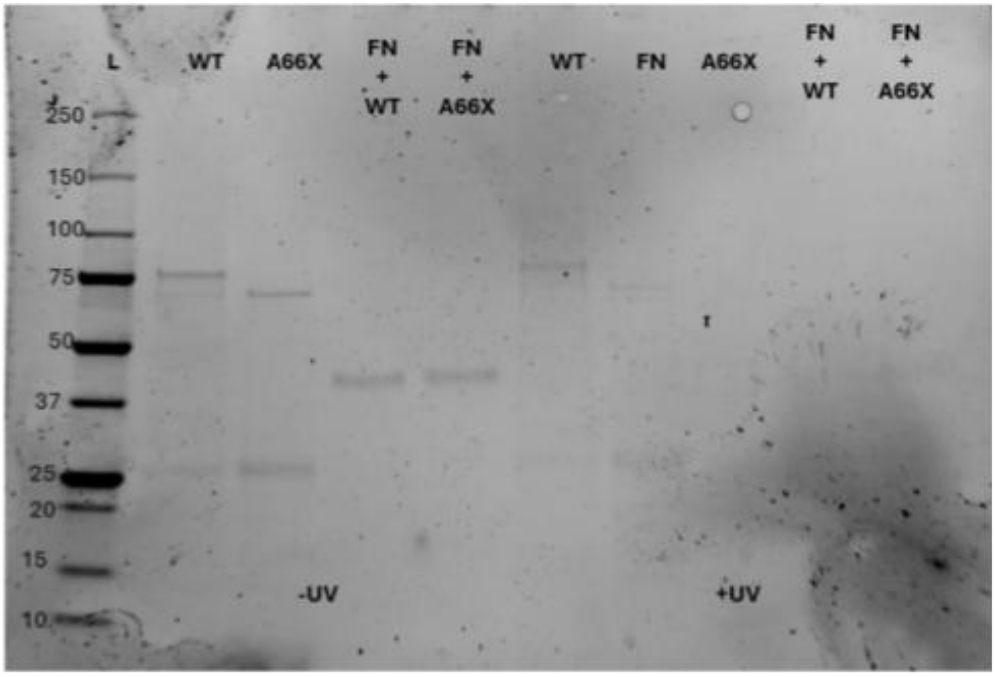


Figure 47. F203X and 45FN Short Wave Crosslinking Experiment.

The initial long and short-wave crosslinking experiments for the F203X mutant were generally inconclusive, due to the disappearance or improper loading of some bands such as the WT in the FN + WT lanes of Figure 48, and multiple lanes in the short-wave experiment. Nonetheless, a band was seen in the +UV portion of the long wave crosslinking gel (Figure 48) with the F230X mutant and FN together. This band appears at approximately the correct molecular weight expected from a crosslinked product (~123 kDa).

Figure 48. DBCO-Cy5 Assay for F203X and WT TG2.

Figure 50 highlights a recurring issue with the DBCO-Cy5 assay, namely that the negative control, WT TG2 should appear dim since there is no possible incorporation of AzF; however, instead it appears bright. The NS3 positive control is also bright, as expected, but the F203X mutant is dim. It is unclear why the F203X is showing up less brightly than the negative control (WT).

Residues F203X and A66X were chosen for investigation to troubleshoot the lack of 45FN-L114XTG2 complex formation after the attempt at crosslinking shown in Figure 44. These residues gave mixed results for expression, crosslinking with 45FN, and incorporation of AzF, as evidenced by the DBCO-Cy5 assay. To continue troubleshooting these issues we decided to test the efficiency of AzF incorporation into F203X by expressing it with and without AzF added to the growth medium at induction. These test expressions were otherwise completed following the general method for pEVOL-pAzF expression, alongside a WT/pEVOL-pAzF expression.

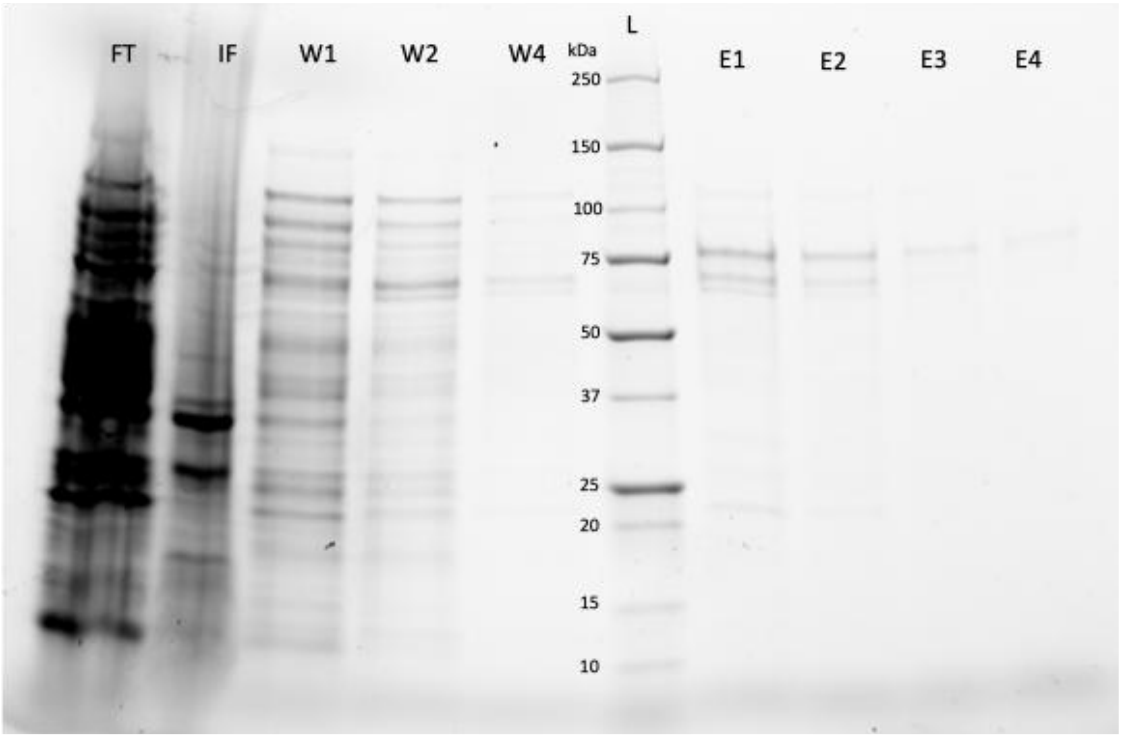


Figure 49. SDS-PAGE of WT/pEVOL-pAzF TG2 Expression #3 Post Ni-NTA Purification.

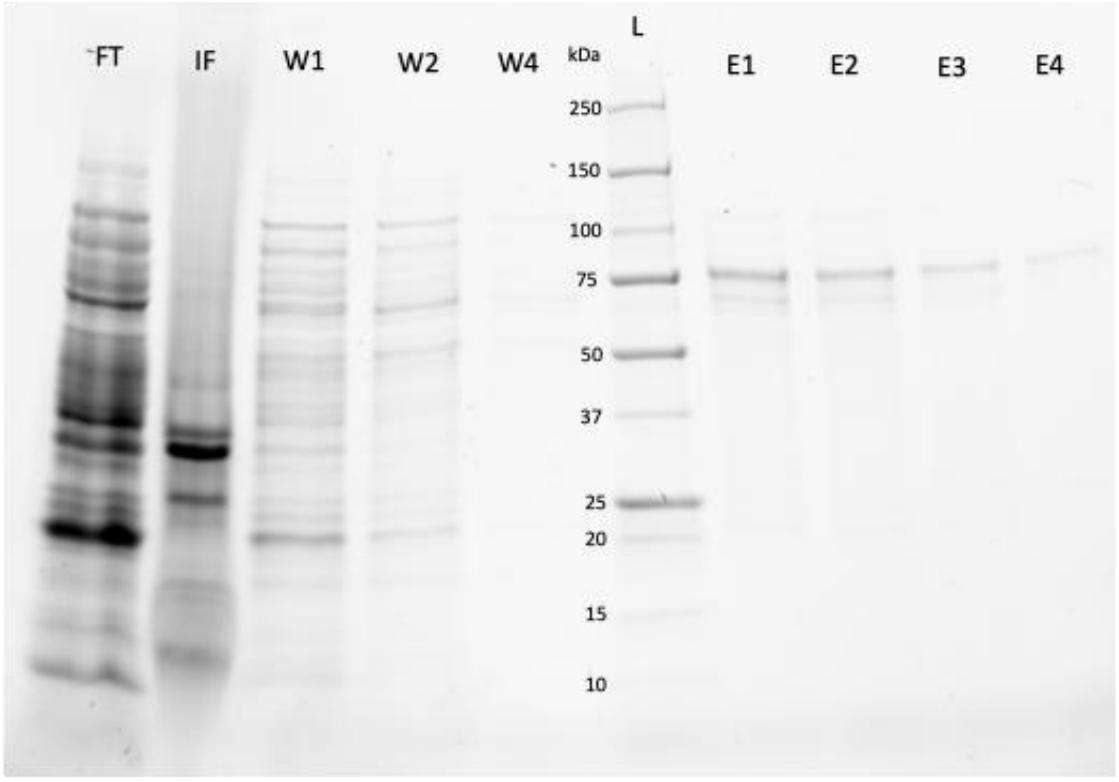


Figure 50. SDS-PAGE of F203X/pEVOL-pAzF Expression with AzF Added at Induction.

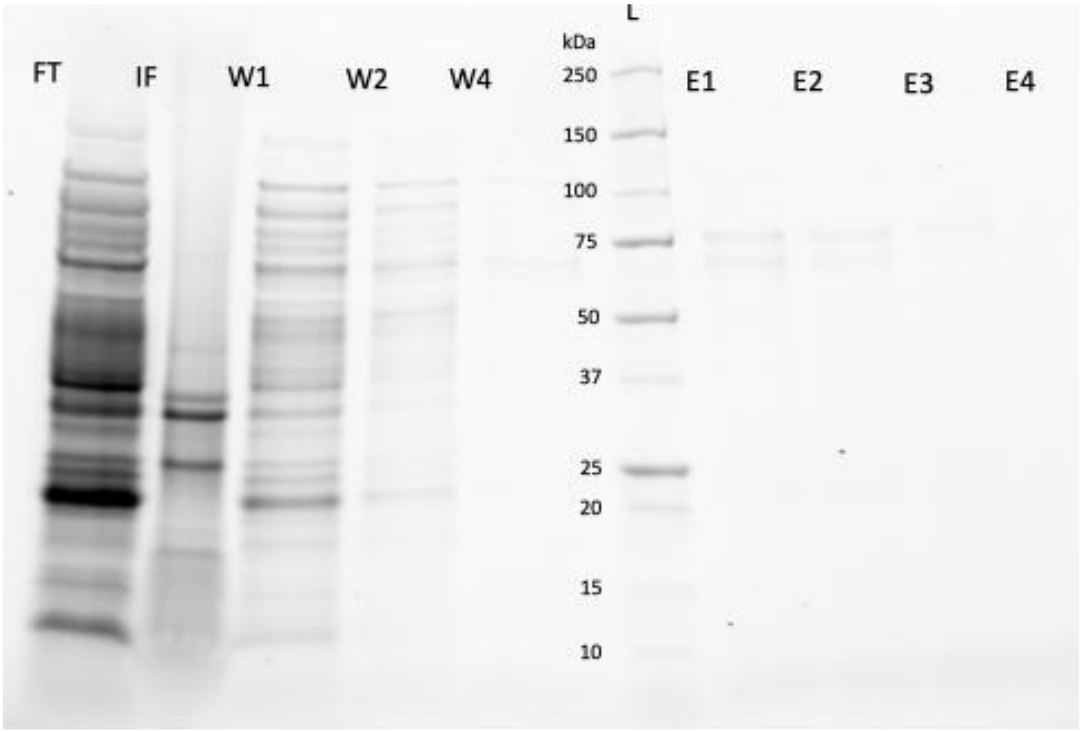


Figure 51. SDS-PAGE of F203X/pEVOL-pAzF Expression without AzF Added at Induction.

Protein was present in all three expressions, including the control experiment where F203X was expressed *without* AzF added to the induction culture. This suggests that the evolved tRNA in the pEVOL-pAzF system is promiscuous for natural amino acids and can sometimes incorporate them at the TAG codon in place of AzF. The F203X +AzF and -AzF proteins were brought forward for further characterization using crosslinking and DBCO-Cy5 assays.

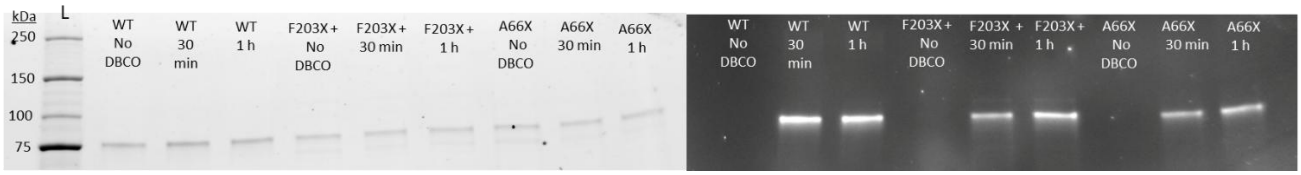


Figure 52. DBCO-Cy5 Assay for F203X/pEVOL-pAzF Expressed with AzF.

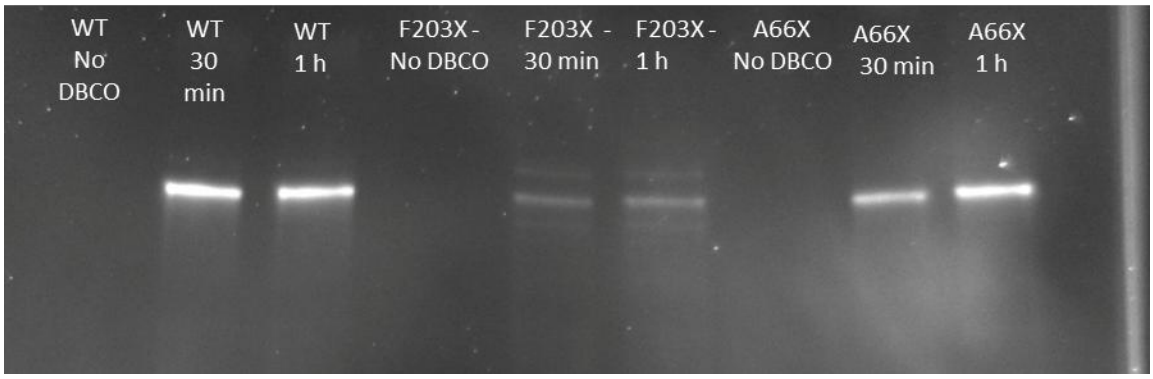


Figure 53. DBCO-Cy5 Assay for F203X/pEVOL-pAzF Expressed without AzF.

In these labelling experiments the proteins were pre-incubated in 8 M urea, prior to incubation with the DBCO-Cy5 dye. This was done to denature the proteins so that tertiary structure would not inhibit DBCO-Cy5 binding. This altered procedure was implemented after seeing the dimness of F203X in Figure 50. It appears the F203X without AzF added to the induction culture had less AzF incorporation due to its comparative dimness. Still, it is hard to draw conclusions due to the bright "negative" control of WT TG2.

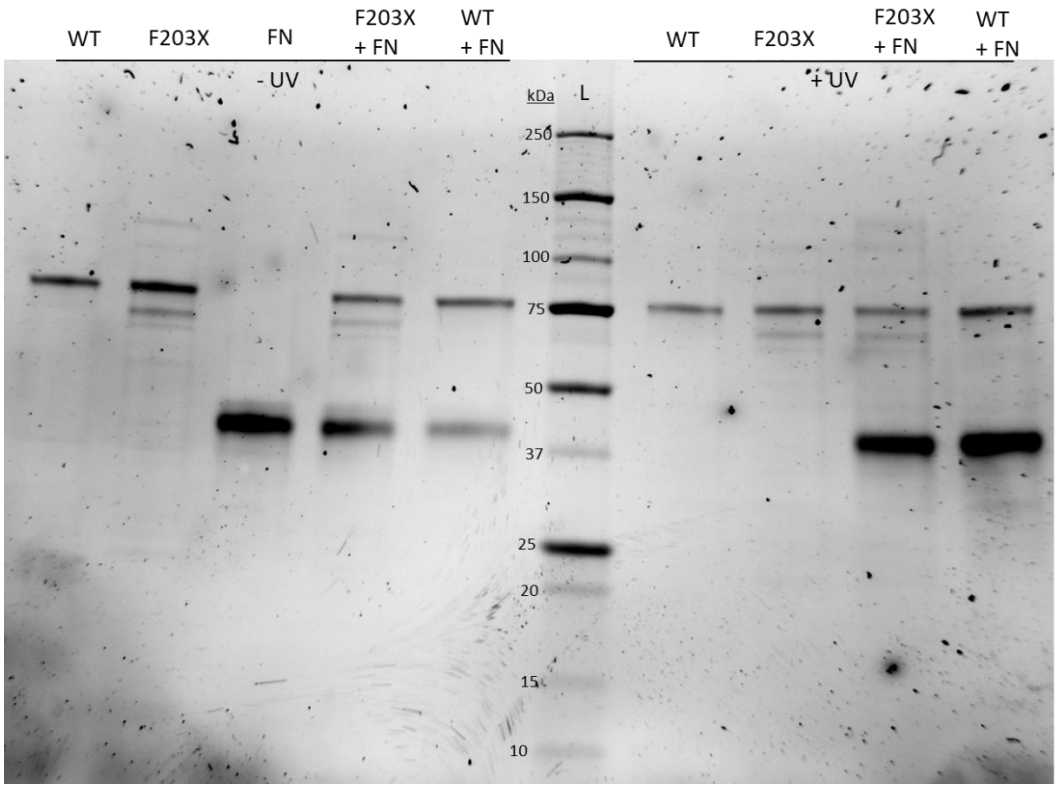


Figure 54. SDS-PAGE of F203X/pEVOL-pAzF Expressed with AzF Short-Wave Crosslinking Assay with 45FN.

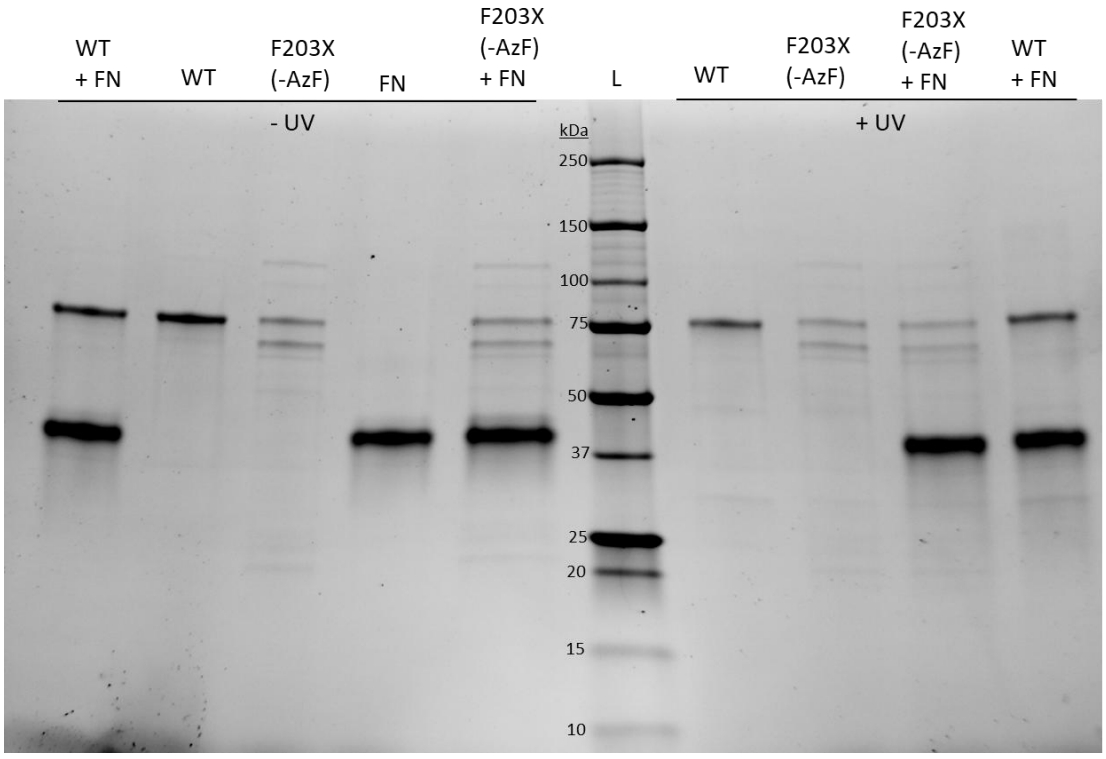


Figure 55. SDS-PAGE of F203X/pEVOL-pAzF Expressed without AzF Short-Wave Crosslinking Assay with 45FN.

The crosslinking assay from Figures 56 and 57 led to some striking results. A band was present at around the expected molecular weight for a crosslinked 45FN-F203XTG2 complex (~123 kDa) in the F203X + FN (+UV) lane for the (+AzF) protein in Figure 54. This band was not present in the (-AzF) protein seen in Figure 57. These results point towards a crosslinked product being formed between F203X TG2 and 45FN. The crosslinking assay was re-done with the F203X (+AzF) protein, and the resulting SDS-PAGE was transferred to a Western Blot with His-Tag detecting antibody, as described in the experimental section.

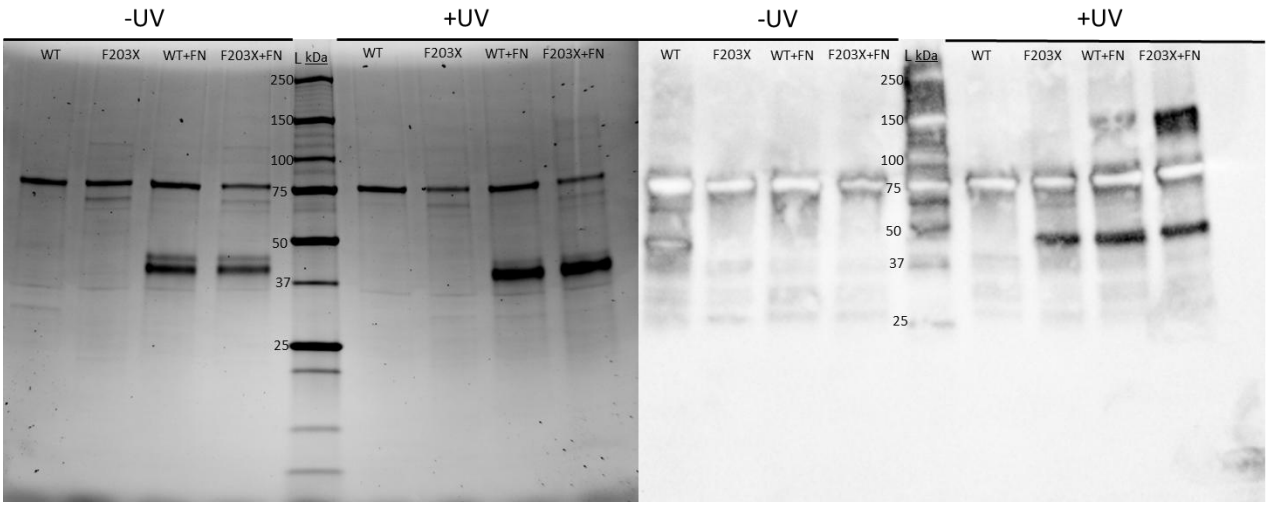


Figure 56. SDS-PAGE and Western Blot of Short-Wave (254 nm) F203X Crosslinking Experiment with 45FN. Left, stain-free SDS-PAGE. Right, western blot visualized with anti-His antibody which only detects His-tagged protein. (FN) represent the 45 kDa FN fragment.

In Figure 56, we can see a smeared band present on the western blot in the F203X + FN lane and a faint band on the WT + FN lane. We hypothesize that the faint band in the WT lane is caused by the dimerization of TG2, hence the molecular weight around 150 kDa. We believe the smeared band is the F203X and fibronectin crosslinked product. This experiment was repeated using long-wave (365 nm) instead of short-wave (254 nm) UV light.

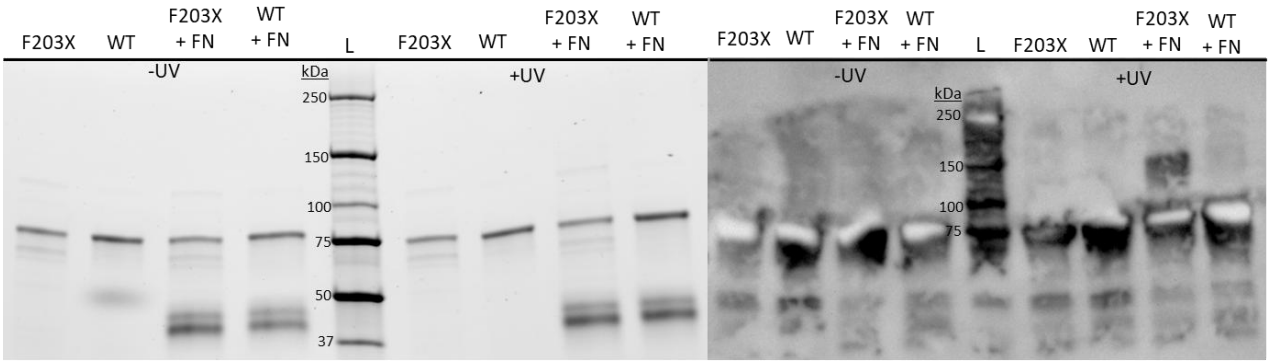


Figure 57. SDS-PAGE and Western Blot of Long-Wave (365 nm) F203X Crosslinking Experiment with 45FN. Left, stain-free SDS-PAGE. Right, western blot visualized with anti-His antibody which only detects His-tagged protein. (FN) represent the 45 kDa FN fragment.

The long-wave UV exposure compared to short-wave seems to have resulted in more specific formation of the crosslinked product, with no WT TG2 dimer forming. It is worth noting that these crosslinking experiments were performed with a higher wattage UV lamp, 6 watts versus 4 watts. The addition of Western Blotting to the crosslinking assay protocol allows visualization of low concentrations of the 45FN and F203XTG2 complex. These results show that we have successfully created a snapshot of F203XTG2 and 45FN in complex.

Table 4. Summary of pEVOL-pAzF Expression Concentrations, Yields and Activities. BLQ stands for below limit of quantification.

Protein	Concentration (mg/mL) (mL)	Yield (mg/L)	Activity (U/mg)
K30X/pEVOL-pAzF #1 (Fig 31)	0.37 ± 0.12 (1.25)	0.84 ± 0.28	0.016 ± 0.014
WT/pEVOL-pAzF #1 (Fig 32)	2.29 ± 0.20 (0.62)	2.84 ± 0.25	0.180 ± 0.006
L114X/pEVOL-pAzF (Fig 33)	1.46 ± 0.26 (0.61)	1.61 ± 0.28	0.044 ± 0.018
I195X/pEVOL-pAzF (Fig 34)	BLQ	BLQ	BLQ
E29X/pEVOL-pAzF (Fig 35)	BLQ	BLQ	BLQ
K30X/pEVOL-pAzF #2 (Fig 36)	0.82 ± 0.10 (1.3)	1.93 ± 0.24	0.103 ± 0.014
H134X/pEVOL-pAzF (Fig 37)	BLQ	BLQ	BLQ
R116X/pEVOL-pAzF (Fig 38)	0.40 ± 0.04 (1.52)	1.11 ± 0.11	0.160 ± 0.002
K30X/pEVOL-pAzF #3 (Fig 39)	0.63 ± 0.08 (1.15)	1.32 ± 0.17	0.198 ± 0.072
WT/pEVOL-pAzF #2 (Fig 43)	3.38 ± 0.41 (0.48)	3.24 ± 0.40	0.025 ± 0.003
A66X/pEVOL-pAzF	2.60 ± 1.82 (0.34)	1.61 ± 1.12	BLQ

(Fig 44)			
F203X/pEVOL-pAzF (Fig 45)	2.88 ± 0.29 (0.82)	4.30 ± 0.43	0.060 ± 0.005
WT/pEVOL-pAzF #3 (Fig 49)	9.78 ± 1.25 (0.2)	3.91 ± 0.50	0.023 ± 0.001
F203X/pEVOL-pAzF +AzF (Fig 50)	1.26 ± 0.17 (2.8)	6.39 ± 0.89	0.063 ± 0.004
F203X/pEVOL-pAzF -AzF (Fig 51)	1.12 ± 0.25 (1.2)	2.43 ± 0.55	0.043 ± 0.009

Table 4 presents the concentrations, yields and activities from the pEVOL-pAzF expressions presented in this section. Overall, we were able to generally increase the yield of protein by using this system over pDULE-pCNF, without making further changes to the established expression protocol. Yields of K30X, L114X, R116X and WT expressions were increased using this system. Variability between expressions, illustrated by the three K30X/pEVOL-pAzF expression, was noted for this system both in terms of yield and activity. It is hard to conclude whether this is from small differences in expressions from batch to batch, variability in the Bradford and AL5 assays, or a different factor entirely. Variability in specific activity is also notable in the decrease from WT/pEVOL-pAzF #1 (Figure 32) to #2 and #3 (Figures 43 and 49); the reason for this is unclear but could be due to removal of TCEP from some protein expression buffers. TCEP can reduce azides to amines, deactivating AzF but it can also maintain the thiol activity of TG2. We continued to use specific activity as a measure of

protein structural integrity despite the variability by prioritizing mutant and WT comparisons between the same batch.

AzF TG2 mutants L114X, K30X and R116X were able to be expressed while H134X, E29X and I195X showed no protein present. Assessment of incorporation using the DBCO-Cy5 assay was generally inconclusive due to two major problems. Firstly, the negative control, WT TG2, sometimes lit up as bright or brighter than TAG mutated proteins with AzF added to the growth medium such as in Figure 48, 52 and 53. Additionally, loading of protein was not equal even though calculations were done using the measured protein concentrations to try and ensure equality as seen in Figures 40 and 41. Despite these qualms L114X showed the most promising evidence of AzF incorporation indicated by increasing fluorescence over increasing reaction time with DBCO-Cy5 and the overall brightest fluorescence of mutant proteins. Therefore, we attempted a crosslinking assay with 45FN under short and longwave conditions. No crosslinked protein-protein complex was seen in Figure 42. We hypothesized that the L114X mutation may be too involved in the 45FN and TG2 native protein-protein association and by mutating it we are disrupting association enough that crosslinking is impaired.

To address this concern, we attempted incorporation of AzF into positions F203 and A66, residues thought to not be crucially involved but still in contact with 45FN according to our alternative binding pose. AzF incorporation of these residues led to high yields for both proteins but loss of activity in A66X. DBCO-Cy5 and crosslinking experiments with F203X were inconclusive although a hint of a 45FN-F203XTG2 crosslinking product was seen in Figure 46. We proceeded to assess whether AzF was being incorporated in F203X by expressing it with and without 0.91 mM AzF added to the growth medium before induction.

These results show that F203X is expressed even in the absence of AzF. Shown in Table 4, the yield of the F203X without AzF added before induction is lower than when AzF was added before induction (Figure 51 versus Figure 52). This suggests that there is a mixed proportion of AzF and endogenous protein present in the F203X expressed with AzF added before induction. Theoretically, if AzF is not present in the solution, the tRNA coding for the amber stop codon, TAG, would not be able to continue translation and TAG would act as a normal stop codon, terminating protein synthesis. This, in combination with the fact that we are expressing TG2 with a C-terminal His tag, should mean that this truncated protein does not show up in the elution fractions post Ni-NTA- affinity chromatography. However, what we observe is full length TG2 being expressed even in the absence of AzF. This suggests that the pEVOL-pAzF tRNA/aaRS system is incorporating endogenous residues at the TAG location to some extent, which means our AzF mutant could exist as a mixture of AzF-incorporated protein and TG2 with an endogenous protein incorporated at the TAG position. Noting this revelation, we continued to characterize the F203X protein expressed with and without AzF added through DBCO-Cy5 and 45FN crosslinking assays.

Remarkably, a band correlating to a 45FN-F203XTG2 crosslinked protein complex was seen when F203X was expressed with AzF but not when AzF was absent (Figure 54). At this point, we added a step to the crosslinking assay, namely the Western Blotting of His tagged proteins, and confirmed that we created a 45FN-F203XTG2 crosslinked complex shown in Figures 56 and 57. Additionally, we found that using longwave UV light resulted in a more selective crosslink product being formed, without WT TG2 dimerization.

In conclusion, the pEVOL-pAzF system increased the yields of AzF TG2 mutants compared to the pDULE-pCNF system, and this system was used to create a 45FN-F203XTG2

protein complex. DBCO-Cy5 assays were generally not useful for measuring AzF incorporation and L114X mutant was not able to form a complex with 45FN, at least under stain-free SDS-PAGE visualization. Expression of F203X TG2 with and without AzF shed light on the promiscuity of the pEVOL-pAzF tRNA resulting most likely in a mixture of endogenous and mutant TG2 being created during expressions with this system. We believe this final point deserved a closer look and spent some time troubleshooting general expression of TG2 with the pEVOL-pAzF system, which is the focus of the next section.

3.2.4 Troubleshooting TG2 Expression

The first troubleshooting experiment was performed to investigate the mixed expression of endogenous and mutant enzyme that was observed with the F203X mutant using the pEVOL-pAzF system. We performed induction of several pEVOL-pAzF mutant plasmids at a small-scale (5 mL) with and without 0.91 mM AzF added to the cultures. The induction cultures were lysed using repeated freeze-thaw cycles and lysates were loaded onto an SDS-PAGE, then transferred to a Western Blot where TG2 was revealed using an anti-His antibody. This antibody detects only His-tagged proteins.

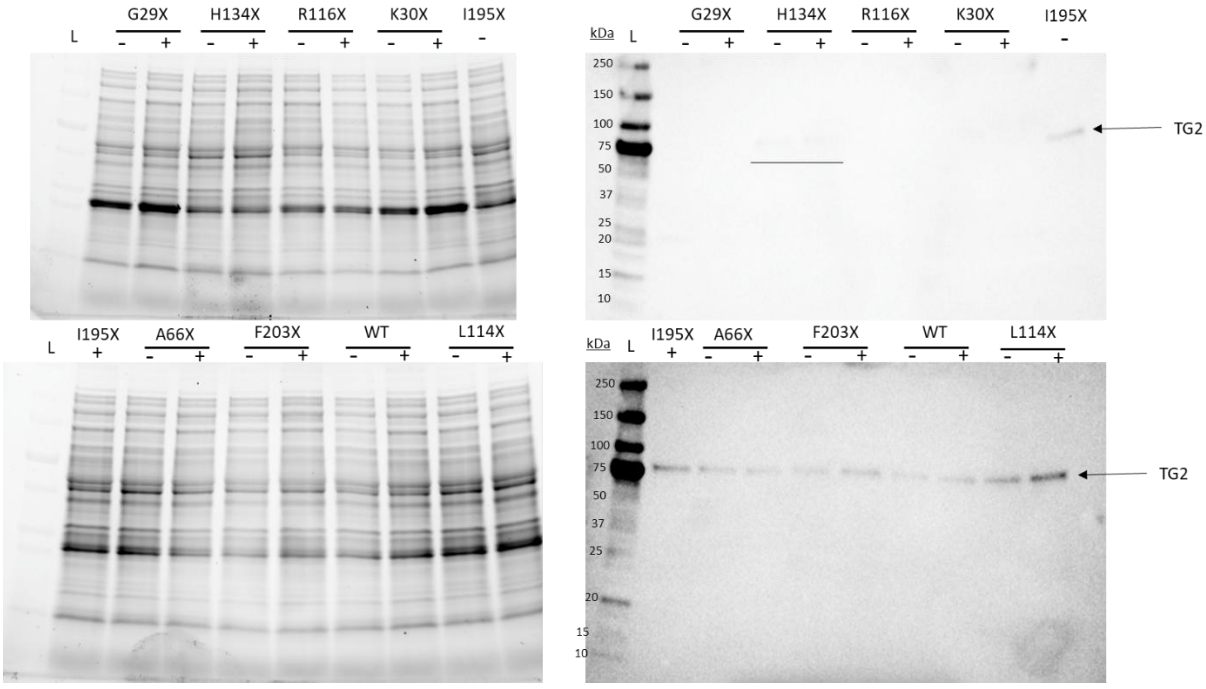


Figure 58. SDS-PAGE and Western Blot of Small-Scale pEVOL-pAzF Expression Testing With and Without AzF Added to Expression Cultures. Gel images on the left are SDS-PAGEs of soluble cell lysate and images on the right are Western Blots of the small-scale expressions of mutants and WT proteins. The underlined portion on the upper right western blot is highlighting faint TG2 bands. (-) and (+) denotes without and with 0.91 mM AzF added to cultures before induction. The western blot was visualized using Anti-His antibodies which detect His-tagged proteins.

The results from the small-scale expression once again highlight the fact that the TAG mutant TG2 is being expressed even in the absence of AzF. We tried to solve this problem by increasing the amount of evolved orthogonal tRNA synthetase pairs available for uptake of AzF and translation corresponding to the TAG codon. This was attempted by pre-transforming competent cells with the pEVOL-pAzF plasmid before transforming WT and mutant TAG TG2

plasmids into the pEVOL-pAzF transformed competent cells before expression. We hoped this would increase the availability of pEVOL-pAzF plasmid and subsequently tRNA for incorporation of AzF during expression.

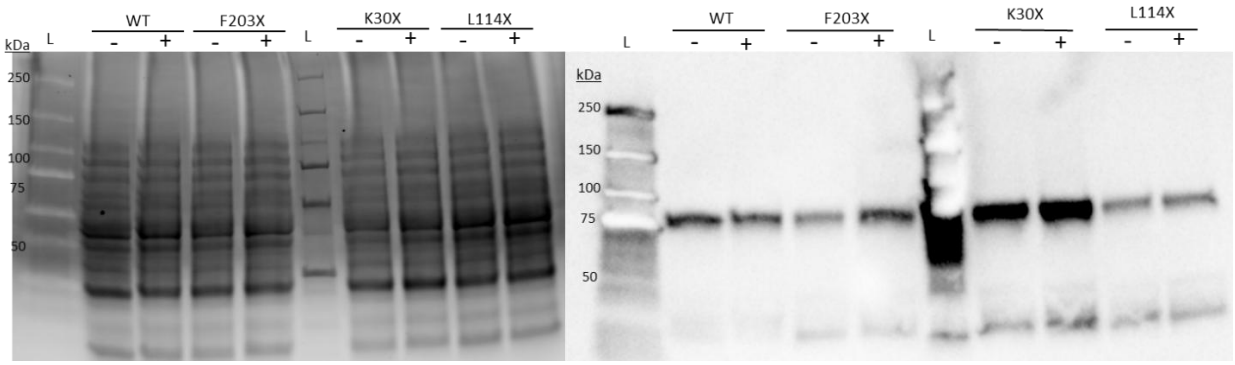


Figure 59. SDS-PAGE and Western Blot of Small-Scale Expression of Pre-transformed pEVOL-pAzF BL21 (DE3) Cells with and without AzF Added to Expression Cultures. These small-scale expression cultures were inoculated with colonies transformed with TG2 plasmids into competent BL21 (DE3) cells that were pre-transformed with pEVOL-pAzF plasmid. (-) and (+) denotes without and with 0.91 mM AzF added to the induction cultures. The western blot was visualized using Anti-His antibodies which detect His-tagged proteins.

For this experiment, we found similar results to the experiment shown in Figure 60, namely no difference between expression with and without AzF for pre-transformed pEVOL-pAzF cells. Another variable in controlling expression was arabinose which induces expression of the evolved tRNA/aaRS orthogonal pair in the pEVOL-pAzF plasmid system. The next experiment took arabinose out of the equation by not adding it to any of the induction cultures.

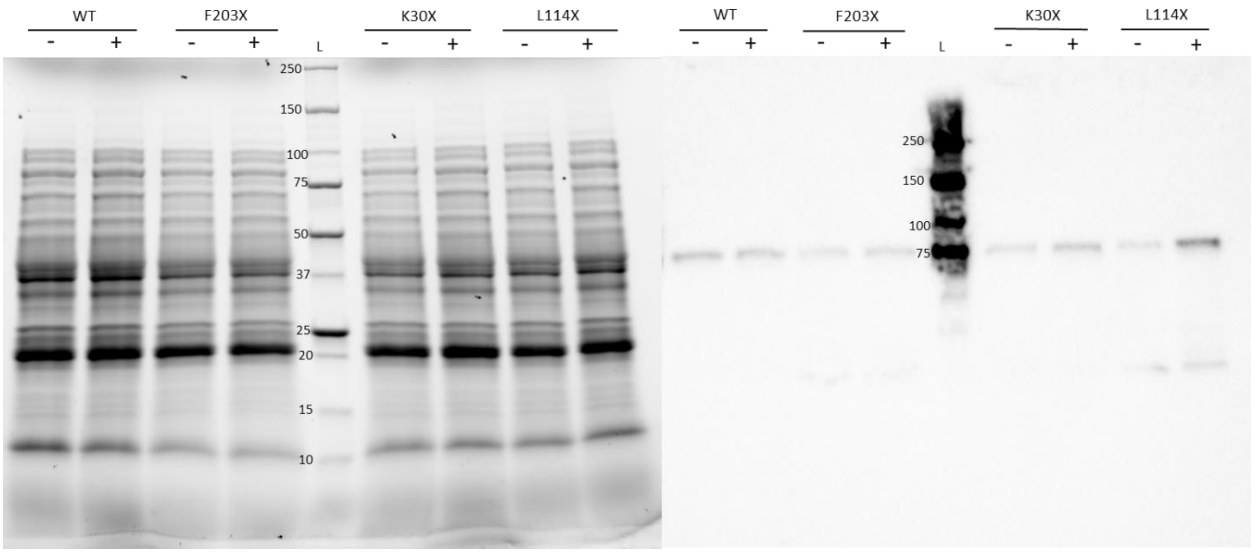


Figure 60. Small-Scale Expression Testing of pEVOL-pAzF TG2 Mutants and WT with and without AzF Added, without Arabinose. (+) and (-) denotes with and without 0.91 mM AzF added to the induction cultures. None of the cultures had Arabinose added. It is worth noting that the (-) WT culture also did not have 1 mM IPTG added. The western blot was visualized using Anti-His antibodies which detect His-tagged proteins.

Similar results were seen in Figure 60 as the results from Figure 61; there is no difference in the amount of TAG mutant TG2 expressed, with or without added AzF, including when the inducer for the GCE machinery, arabinose, is removed.

This experiment highlighted a deeper issue, namely that similar levels of TG2 expression were seen without 1 mM IPTG added to the (-) WT culture as other expressions. This is troubling, since IPTG should be a major inducer for TG2 expression.

To investigate this finding more thoroughly, another expression test was completed at a medium-scale (100 mL) with WT TG2 plasmid in BL21 (DE3) competent cells and WT TG2 with pEVOL-pAzF in BL21 (DE3) competent cells. These expressions were either induced with IPTG or not and TG2 was purified using Ni-NTA affinity chromatography. Samples of pooled

elution fractions after chromatography or samples from the soluble cell lysate were visualized using Anti-His western blotting.

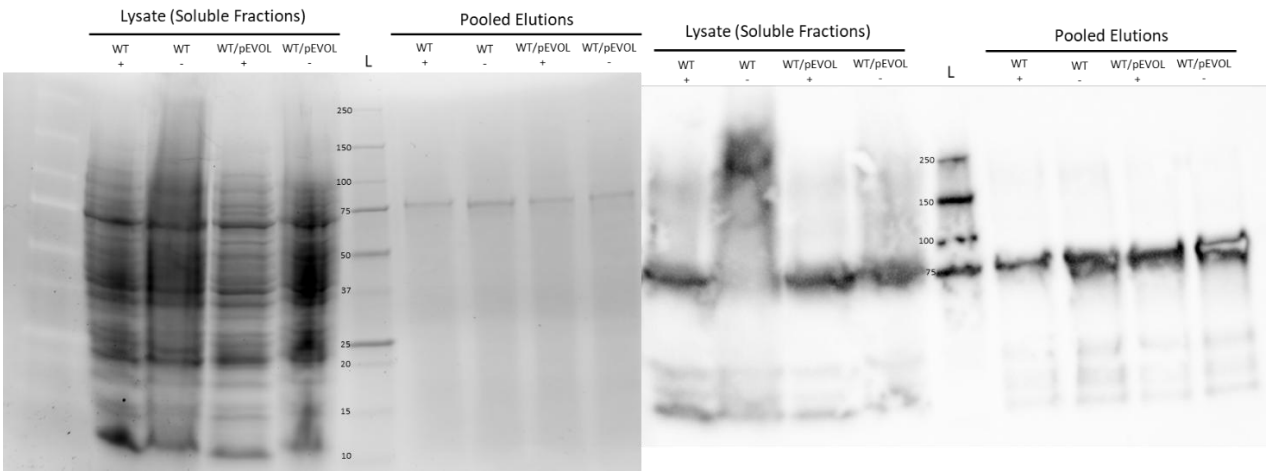


Figure 61. SDS-PAGE and Western Blot of Medium-Scale Expression of WT TG2 and WT/pEVOL-pAzF TG2 with and without 1 mM IPTG Induction. Samples from pooled elution fractions after Ni-NTA purification and soluble cell lysate were visualized using an SDS-PAGE (Left) and Anti-His Western Blot (Right). Expressions were either induced with 1 mM IPTG (+) or had nothing added to the cultures (-). The western blot was visualized using Anti-His antibodies which detect His-tagged proteins.

These troubleshooting experiments were undertaken to assess the results discovered when F203X was expressed with and without 0.91 mM AzF added to the culture, shown in section 3.2.3. In that experiment, F203X was observed to be expressed even in the absence of AzF, albeit at a reduced yield. Subsequent troubleshooting experiments presented in this section illustrate a general lack of control of TG2 expression using the Cterm-pHis-hTG2 and pEVOL-pAzF plasmids concerning induction elements like IPTG, Arabinose and AzF.

In Figure 58, we confirmed that multiple TAG mutant TG2 plasmids continued to express protein even in the absence of 0.91 mM AzF. This suggests that TAG mutants are producing a mixed population of AzF and endogenous amino acids incorporated at the TAG codon. This could be due to promiscuity of the pEVOL-pAzF tRNA, allowing it to bring endogenous amino acids to the TAG codon. Theoretically, TG2 should be truncated if the TAG codon worked as intended, where if no AzF is available for incorporation protein synthesis is stalled.

One attempt at boosting the incorporation of AzF was to increase the amount of tRNA/aaRS present in the system by pre-transforming BL21 (DE3) cells with pEVOL-pAzF before transforming WT or mutant Cterm-pHis-hTG2 into the cells for expression. We re-did the experiment in Figure 58 with these pre-transformed cells and saw no difference in the expression of TG2 with and without AzF, shown in Figure 59. Nonetheless, we decided to keep this change for subsequent expressions since it is established protocol for AzF expression in other laboratories like the Pezacki group.

Another induction element in the pEVOL-pAzF system is arabinose. 0.02% arabinose had been added to the induction cultures to induce the expression of the tRNA/aaRS orthogonal pair. We took this induction element out of the equation in small-scale expressions in Figure 60 to see if expression of WT and mutant TG2 with and without AzF would be affected. We saw no difference in the expression levels. Now, theoretically no AzF suppressor tRNA/aaRS should be present, since no arabinose was added, and the fact that TAG mutants were still expressed at the same level, with and without AzF, leads us to question the effectiveness of the tRNA/aaRS pair in previous expressions. Additionally, in this experiment no IPTG was added to the (-) WT test expression, but TG2 expression was still seen.

To further investigate the lack of control of IPTG on the induction of TG2, we ran an experiment with WT TG2 and WT/pEVOL-pAzF TG2, with and without 1 mM IPTG added, during induction at a medium scale (100 mL). We purified the expressed proteins using Ni-NTA affinity chromatography and assessed the amount of TG2 from the pooled elution fractions and from the soluble cell lysate in Figure 61. Surprisingly, the amount of TG2 expressed in both samples with and without IPTG were the same. This indicates a clear lack of control of TG2 expression with IPTG using the Cterm-pHis-hTG2 expression plasmid.

We also reassessed the Sanger sequencing chromatograms for WT, A66X, F203X, H134X, E29X, K30X, L114X, R116X and I195X Cterm-pHis-hTG2 plasmids used in the expression alongside the pEVOL-pAzF system for the proper mutation present at the desired position. We also checked TAG mutant Sanger sequencing chromatograms for partial contamination with WT codons, which may have explained the excessive readthrough of AzF mutant expression when AzF is not added to the culture and the tRNA/aaRS is not expressed by arabinose. We found no contamination of WT codons. We also sent WT, F203X and R116X Cterm-pHis-hTG2 plasmids for full plasmid sequencing and found that there were no abnormalities in the full plasmids.

Overall, these experiments led to troubling revelations over the efficacy and control of TG2 expression both for AzF incorporation and for general expression of TG2. We confirmed that TAG mutants are at the minimum producing a mixed population of endogenous amino acid and AzF amino acid incorporated at TAG positions. We also confirmed that IPTG and arabinose induction elements have minimal control over the yield of TG2 expressed. Nonetheless, we still achieved formation of an 45FN-F203XTG2 crosslinked complex visible by Anti-His western blotting in Figure 57 and we wished to try one more expression system, pULTRA-pCNF, with

the hope of improving the efficiency of AzF incorporation and subsequently the yield of 45FN-TG2 complexes. The next section summarizes the results from expressions using this system.

3.2.5 AzF Incorporation Using the pULTRA-pCNF System

The pULTRA-pCNF system differs from the pDULE-pCNF system by having an inducible versus constitutive promoter. It differs from the pEVOL-pAzF system by the induction of the evolved tRNA/aaRS orthogonal pair being induced by IPTG instead of Arabinose. This plasmid was gifted to us by the Pezacki group. It has been used to produce AzF mutants in *E. coli*.[109] We aimed to try it for ourselves to see if it would increase the amount of AzF incorporated protein produced and subsequently increase the yield of a 45FN and mutant TG2 crosslinked product.

The first mutant expressed with the pULTRA-pCNF system was the F203X mutant alongside WT TG2 following the protocol for pULTRA-pCNF expression in the material and methods section. This mutant was chosen due to its promising crosslinking results, as shown above in Figure 57.

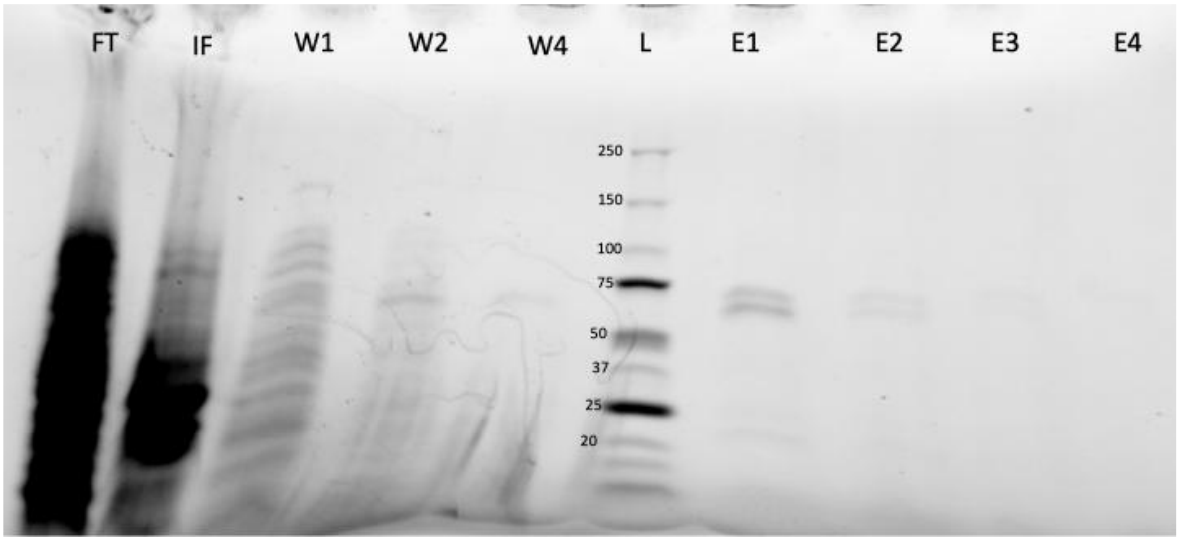


Figure 62. SDS-PAGE of F203X/pULTRA-pCNF TG2 Expression. The wrong concentration of running buffer was used for this SDS-PAGE which is why the protein bands look faint and misshaped.

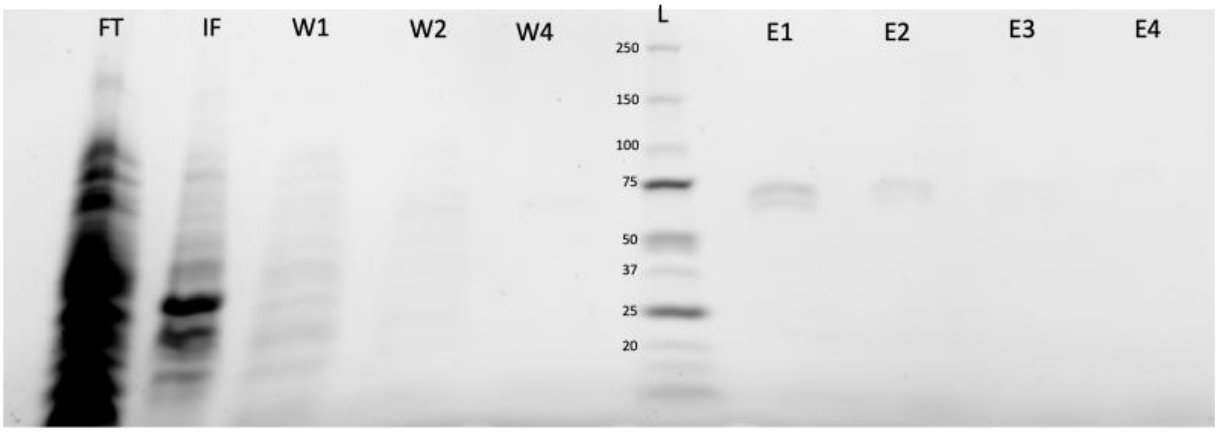


Figure 63. SDS-PAGE of WT/pULTRA-pCNF TG2 Expression. The wrong concentration of running buffer was used for this SDS-PAGE which is why the protein bands look faint and misshaped.

The TG2 bands from the SDS-PAGEs of this expression do not match the expected molecular weight of 78 kDa, but this is most likely because of the wrong concentration of running buffer used for running the SDS-PAGE. These proteins were still brought forward to AL5 activity testing and Bradford protein concentration measurement. The yields of the WT and F203X were 2.92 and 3.24 mg/L respectively, which is comparable to the pEVOL-pAzF expressions. The F230X mutant showed reduced specific activity compared to the Wild-Type, 0.015 versus 0.029 U/mg but the difference was deemed small enough to continue forward with crosslinking and DBCO-Cy5 assays of the F203X mutant.

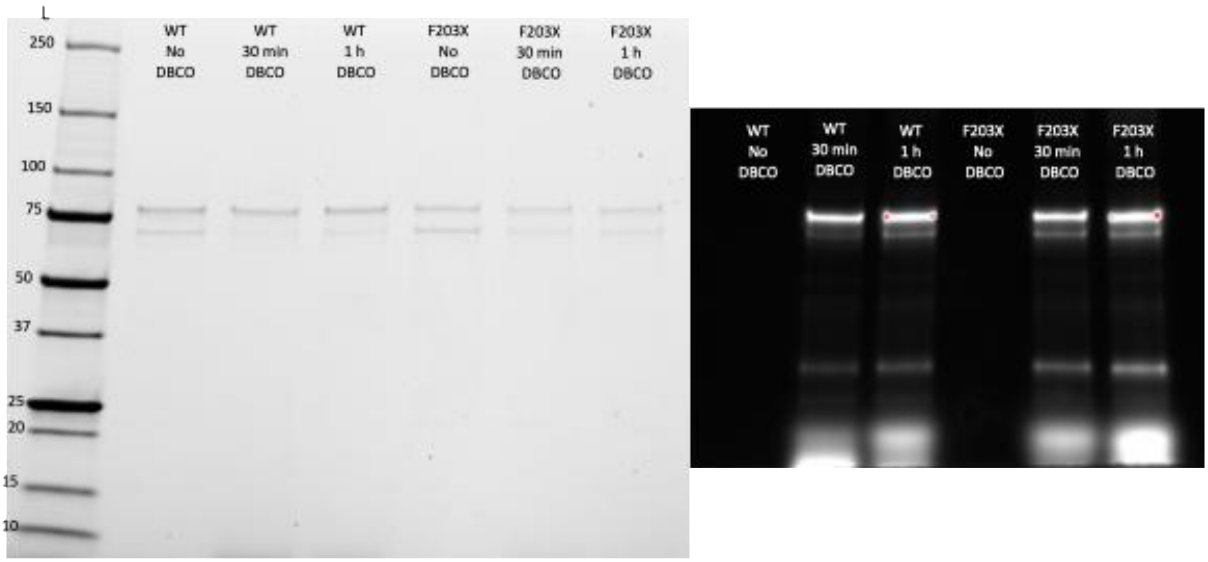


Figure 64. SDS-PAGE WT/pULTRA-pCNF and F203X/pULTRA-pCNF DBCO-Cy5 Assay. The image on the left is the SDS-PAGE visualized under the stain-free conditions while the image on the right is visualized under Cy5 conditions.

Protein bands had the same brightness for both WT and F203X TG2 which is a continuation of the problem of WT TG2 not being an effective negative control for the DBCO-Cy5 assay, due to binding of the DBCO-Cy5 dye to non AzF-incorporated proteins.

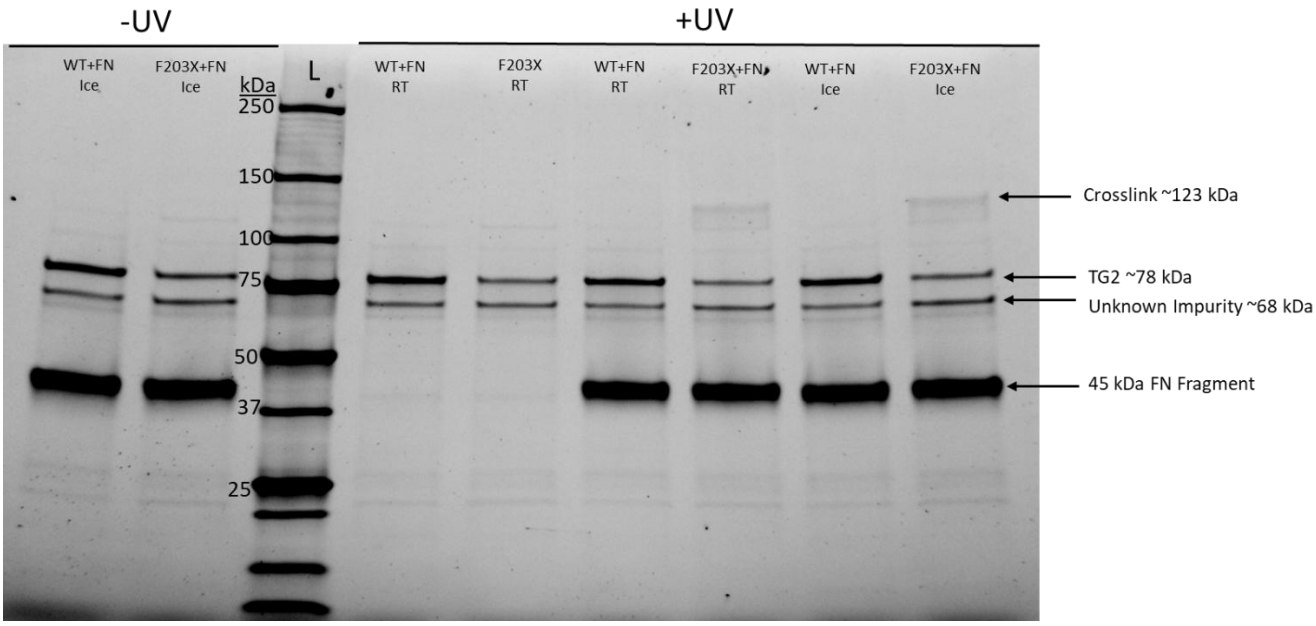


Figure 65. Stain-Free SDS-PAGE of F203X/pULTRA-pCNF and 45FN Long-Wave Crosslinking Assay. Lanes labelled "ice" or "RT" refer to the temperatures the 45FN and WT or F203X proteins were incubated at during the crosslinking assay procedure.

This crosslinking assay, Figure 65, yielded the best evidence of a crosslinked product between F203X and WT TG2 so far, there was a bold band present only for F203X and 45FN lanes, not in the WT TG2 lanes. Notably, the crosslinking band was visible without western blotting necessary.

After these promising results, we expressed the R116X mutant with the pULTRA-pCNF system alongside WT/pULTRA-pCNF.

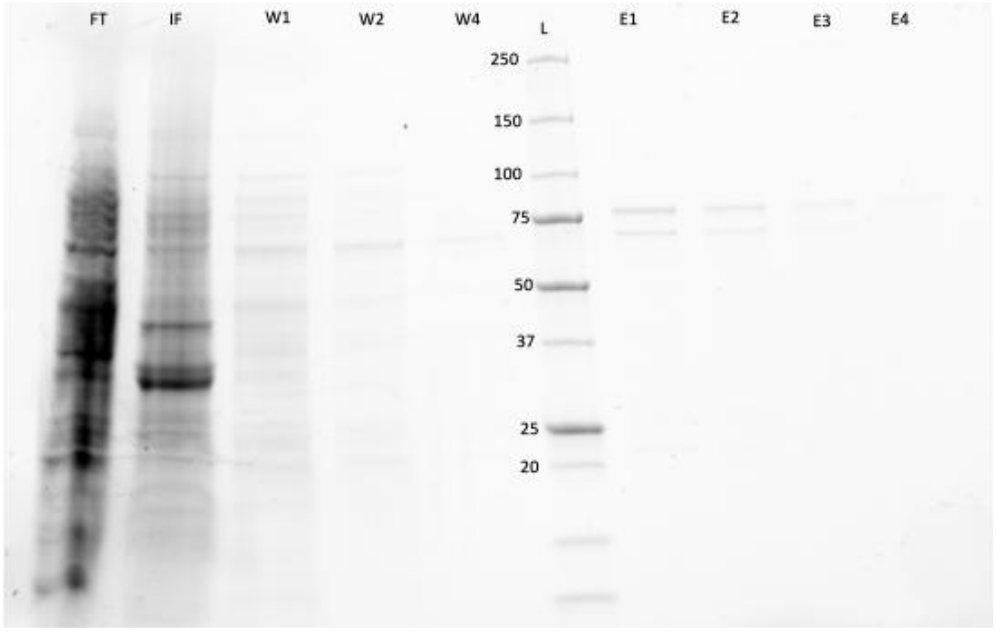


Figure 66. Stain-Free SDS-PAGE of F203X/pULTRA-pCNF and 45FN Long-Wave Crosslinking Assay. Lanes labelled "ice" or "RT" refer to the temperatures the 45FN and WT or F203X proteins were incubated at during the crosslinking assay procedure.

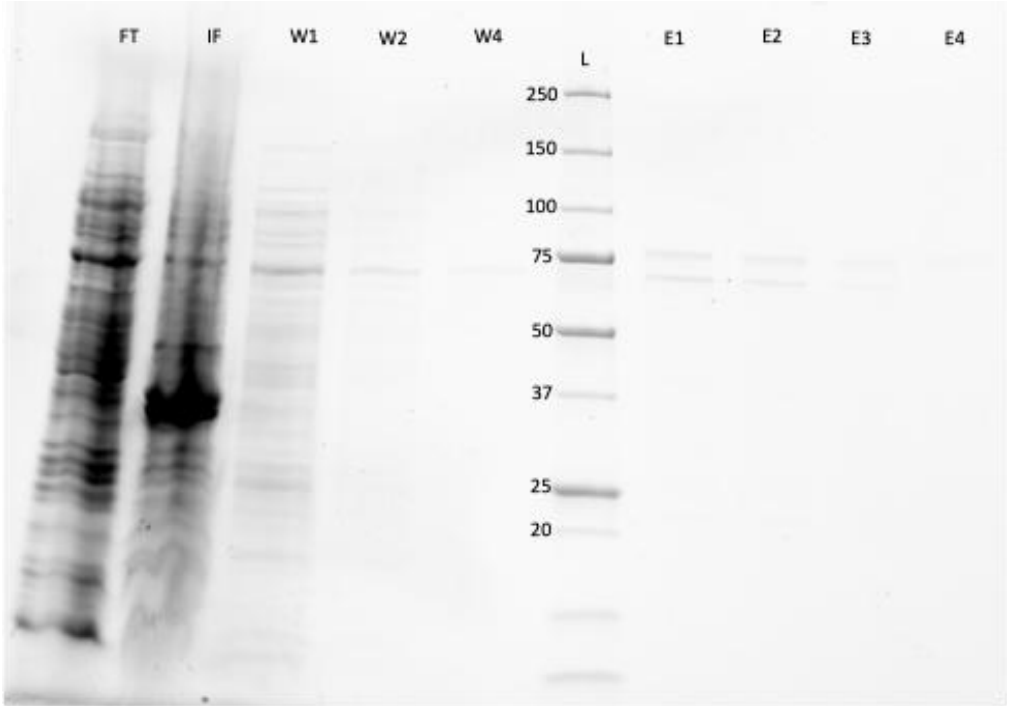


Figure 67. R116X Expression with the pULTRA-pCNF System.

The yields of both the WT and R116X proteins, 4.62 and 2.68 mg/L respectively were both higher than expressions of these proteins using the pEVOL-pAzF plasmid. The specific activities were comparable at 0.017 and 0.015 U/mg respectively. For the DBCO-Cy5 assay, the previous F203X expression as also included as an additional reference.

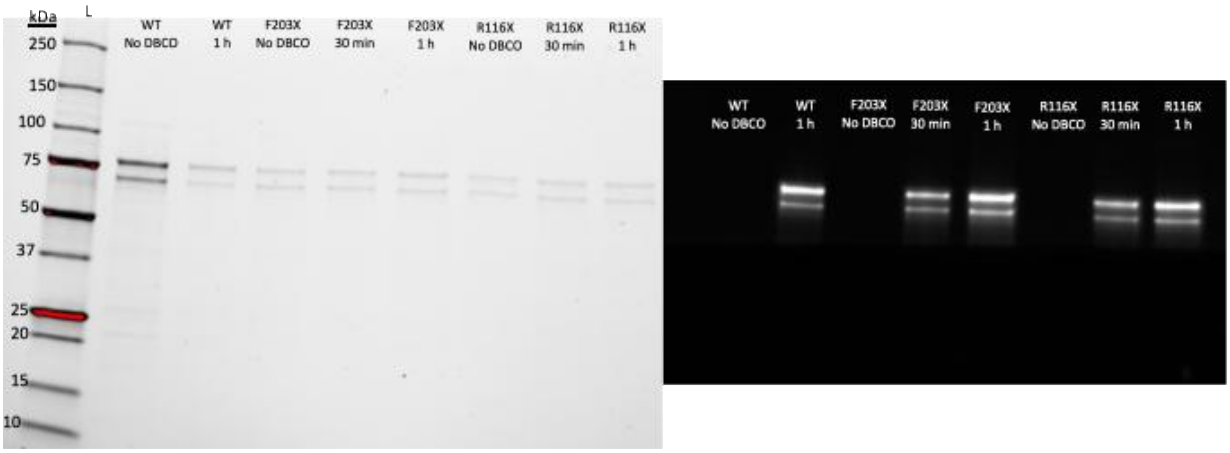


Figure 68. SDS-PAGE of DBCO-Cy5 Assay with R116X/pULTRA-pCNF, F203X/pULTRA-pCNF and WT/pULTRA-pCNF TG2 Expressions.

Once again, similar brightness was seen between the proteins making it difficult to draw conclusions from this assay about AzF incorporation.

Previous crosslinking experiments up to this point were visualized on SDS-PAGEs utilizing Stain-Free technology which is highly sensitive. In order to proceed to the next step, namely in-gel tryptic digestion and MS analysis of paired residues, we consulted with Prof. Jeffrey Smith (Carleton University), an expert collaborator in the field. They informed us that a benchmark for measurement of sufficient protein concentration for tryptic digestion and subsequent MS is a strong band visualized using Coomassie staining on an SDS-PAGE. Therefore, we performed the crosslinking assay of R116X and loaded onto an SDS-PAGE which

was subsequently stained with Coomassie blue and destained as described in the experimental section.

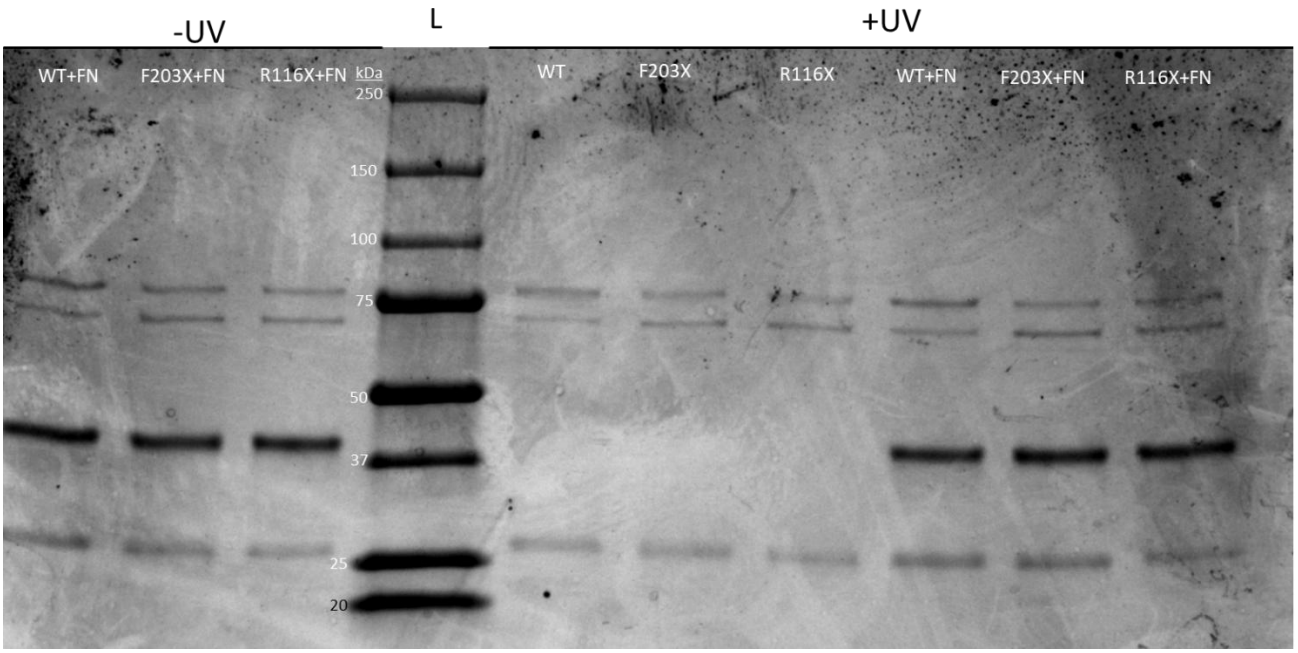


Figure 69. Coomassie stained SDS-PAGE of R116X/pULTRA-pCNF and F203X/pULTRA-pCNF and 45FN Long-Wave Crosslinking Assay.

Unfortunately, no 45FN-R116X TG2 or 45FN-F203X TG2 crosslinked complex was visible using Coomassie staining of the SDS-PAGE. This meant the concentration of the 45FN-F203X TG2 complex was not high enough to justify proceeding with tryptic digestion and MS, according to the criteria established by our expert collaborator.

Table 5. Summary of pULTRA-pCNF Expression Concentrations, Yields and Activities.

Protein	Concentration (mg/mL) (mL)	Yield (mg/L)	Activity (U/mg)
F203X pULTRA-pCNF (Fig 62)	4.57 ± 0.16 (0.39)	3.24 ± 0.11	0.015 ± 0.003
WT pULTRA-pCNF (Fig 63)	2.43 ± 0.40 (0.6)	2.92 ± 0.48	0.029 ± 0.006
WT pULTRA-pCNF #2 (Fig 66)	3.61 ± 0.53 (0.64)	4.62 ± 0.68	0.017 ± 0.005
R116X pULTRA-pCNF (Fig 67)	2.95 ± 0.51 (0.5)	2.68 ± 0.46	0.015 ± 0.002

Expressions using the pULTRA-pCNF system resulted in a slight increase in yield for WT and R116X expressions when compared to the pEVOL-pAzF system. Expressions of F203X were slightly reduced. DBCO-Cy5 assays were still generally inconclusive based on the affinity of the DBCO-Cy5 dye for WT TG2. We hoped that a large increase in AzF incorporation efficiency using the new system could have led to different results with this assay.

The crosslinking assay using F203X TG2 expressed with the pULTRA-pCNF system shown in Figure 65 led to the most promising results seen yet for formation of a 45FN-F203XTG2 complex. The bands representing the complex are clearly visible on an SDS-PAGE visualized using stain-free technology, previously the complex was only clearly visible when using Anti-His antibodies on a western blot in Figures 56 and 57. This suggests that expression with the pULTRA-pCNF system in some way, increased the formation of this complex, perhaps through increased AzF incorporation. However, crosslinking assays of F203X and R116X

visualized by Coomassie staining, Figures 68 and 69, did not result in clearly visible bands for the 45FN-TG2 complex.

While the pULTRA-pCNF system increased the yield of some expressions and increased the formation of the 45FN-F203XTG2 complex, we could not create enough to confidently proceed to the next stage of the project.

3.2.6 Summary of Mutations Across Expression Systems

Across the three expression systems used, five out of eight of the target residues were obtained and further characterized using DBCO-Cy5 and Crosslinking assays. Of the original six residues targeted based on their relevance to the TG2-FN interaction shown in previous publications[79,80], H134, R116, E29, L114, I195 and K30, only three mutant residues were obtained in a concentration useful for further characterization. The residues L114, R116 and K30 were expressed with a yield between 0.84 and 2.68 mg/L (Tables 4 and 5) which were brought forward for characterization. Residues I195, H134 and E29 did not show up on the SDS-PAGE after expression or were purified in very low yields, such as for I195 (0.16 mg/L, Table 3). Of the residues additionally selected for mutation, F203X and A66X, both were obtained at concentrations suitable for characterization, from 1.61-6.39 mg/L (Table 3 and Table 4). Incorporation of non-canonical amino acids can be highly dependant on position and structure of the amino acid being replaced, it is not surprising that some targets were not successful.

Characterization with the DBCO-Cy5 assay was inconclusive across all residues due to issues with non-specific binding to WT TG2, perhaps to surface exposed cysteine residues. This made it difficult to assess the rate of AzF incorporation at these positions. Furthermore, it should be noted that mutant TG2 samples most likely contain a mixed population of endogenous and AzF incorporated TG2 as evidenced by experiments in section 3.2.4. Crosslinking experiments were only shown to be successful with the F203X mutant although crosslinking with Western Blot visualization was not performed on any other of the obtained mutants. F203X forming a crosslinked complex with 45FN points towards our alternative binding pose being viable. It is possible that the conservative mutation from phenylalanine to AzF played a role in the ability of

F203X to form this complex but further research is required to support both this point and our alternative binding pose.

3.3 Conclusion

Our objective was to incorporate the crosslinking UAA, Azido-L-Phenylalanine, using GCE at residues thought to be important for the non-covalent association of 45FN to TG2. We aimed to use this crosslinking residue to create a covalent complex of 45FN and TG2 and analyze this complex through tryptic digestion and subsequent MS analysis of paired residues. We were able to express and purify TG2 from TAG mutated plasmids for positions K30X, R116X, L114X, F203X and A66X, alongside GCE tRNA/aaRS plasmids. We tried the expression with different GCE plasmids, pDULE-pCNF, pEVOL-pAzF and pULTRA-pCNF with the goal of assessing AzF incorporation using a DBCO-Cy5 fluorescent dye and forming a crosslinked complex between 45FN and AzF mutant TG2 which could be digested in-gel with trypsin and subsequently paired residues could be identified with MS.

Multiple expression systems were tested for incorporation of AzF into amber stop codon (TAG) TG2 mutant with varying effects on the yield of protein. The pDULE-pCNF system resulted in overall decreased yields for both the WT and mutant proteins, compared to TG2 expression with the non-GCE Rosetta (DE3) system. The pEVOL-pAzF system showed a dramatic increase in yield from the pDULE-pCNF system. Using this system, we were able to obtain mutant TG2 from Cterm-pHis-hTG2 plasmids mutated at positions K30X, R116X, L114X, F203X and A66X but not for H134X, I195X or E29X. Expression with the pULTRA-pCNF system increased the yield of protein for WT and R116X TG2 but slightly decreased the yield for F203X.

We attempted to characterize the incorporation of AzF using the DBCO-Cy5 fluorescent labelling assay. DBCO-Cy5 should specifically react with AzF, but the results were generally inconclusive due to non-specific labelling of WT TG2. We were able to successfully form a

crosslinked 45FN-F203XTG2 complex using long and shortwave UV light visualized using Anti-His antibodies, which detect His-tagged proteins (Figures 56 and 57). Longwave crosslinking was found to more selectively form the 45FN-F203XTG2 complex. The amount of complex formed was increased by using the pULTRA-pCNF system instead of the pEVOL-pAzF system as evidenced by clear bands visible using stain-free SDS-PAGE visualization (Figure 65). We could not reach concentrations of the complex sufficient to proceed to the next stage of tryptic digestion and MS analysis.

Some problems involving the control of TG2 expression were unveiled through troubleshooting experiments (Section 3.2.4). Induction with and without 1 mM IPTG did not change the amount of WT TG2 expressed (Figure 61). It was also revealed that F203X TG2 was expressed at a significant level even when AzF was not added to the growth medium; this was confirmed by further experiments for other mutants. Theoretically, no TAG mutant TG2 should be expressed without AzF present, only truncated protein. This could be due to promiscuity of the tRNA/aaRS orthogonal pair for endogenous amino acids, most likely resulting in a mixed population of AzF and endogenous amino acid incorporation at TAG mutated positions. It is unclear what is causing these issues, and analysis of the plasmid sequences showed no abnormalities.

Overall, we did achieve the objective of formation of a covalent complex between F203X TG2 and 45FN, but more work is needed to increase the concentration of this complex to a suitable degree for tryptic digestion and MS analysis of paired residues. We recommended using the pULTRA-pCNF expression system for AzF incorporation and either optimizing or using a different method for detection of AzF incorporation, apart from the DBCO-Cy5 assay.

3.4 Materials and Methods

3.4.1 Site-Directed Mutagenesis

The protocol used was identical to the one described in section 2.4.2, apart from the use of different custom primers for TAG mutations, which are presented in the appendix. Some plasmids specified in the results and discussion were sent for full plasmid nanopore sequencing. (Flow Genomics, Plasmidsaurus)

3.4.2 Protein Expression

pDULE-pCNF

Both WT and mutant TG2 Cterm-pHis-hTG2 plasmids were co-transformed into BL21 (DE3) cells with the pDULE-pCNF plasmid[110] (Addgene #85494) and plated on 50 µg/mL kanamycin, 10 µg/mL tetracycline LB-Agar plates. A single colony was picked and used to inoculate 2 mL of TB in a culture tube with appropriate antibiotics. This culture tube was shaken for 7 hours at 37 °C, 250 RPM using Innova 4330 shaking incubator then used to inoculate a flask with 75 mL TB and appropriate antibiotics which was shaken overnight (about 16 hours) at 37 °C, 250 RPM. The next day the 75 mL culture was centrifuged at 4000 × g for 5 minutes in a Sorvall ST1R Plus-MD centrifuge at 4 °C, the pellet was resuspended with ~3.3 mL of TB with appropriate antibiotics and ~1.67 mL of this resuspension was added to 500 mL of TB with appropriate antibiotics. This culture was shaken at 37 °C, 250 RPM until an OD_{600} of 0.5-0.8 was reached. The culture was induced with final concentrations of 1 mM IPTG, and 0.91 mM AzF, the AzF was dissolved in 35 mL TB, 5 mL 0.2M HCl, 10 mL MQ H2O, pH 7.5 before being added to the induction cultures. WT expressions did not have AzF added.

pEVOL-pAzF

Expressions conditions with the pEVOL-pAzF[111] (Addgene #31186) were the same as pDULE-pCNF, except 34 µg/mL chloramphenicol was used in place of 10 µg/mL tetracycline. Arabinose was added at induction at a final concentration of 0.02% (w/v).

pULTRA-pCNF

Expression conditions with the pULTRA-pCNF[109] (Addgene #48215) were generally the same as the pDULE-pCNF, except 50 µg/mL spectinomycin was used in place of 10 µg/mL tetracycline.

3.4.3 Protein Purification

The protocol for protein purification was identical to the one described in section 2.4.4 for all different expression types described in section 3.4.2.

Expressions using the pULTRA-pCNF system were purified using exclusively His-Tag TG2 Lysis Buffer (50mM Na_3PO_4, 400 mM NaCl, 5 mM imidazole, 0.5% triton-X100, pH7.5) and all other buffers contained no TCEP.

3.4.4 Troubleshooting Expressions

Small-Scale Expressions

The general expression protocol for pEVOL-pAzF TG2 was scaled down to a 10-mL scale. AzF mutant or WT TG2 were either co-transformed with pEVOL-pAzF into BL21 (DE3) cells or transformed into BL21 (DE3) cells already containing the pEVOL-pAzF plasmid, depending on the experiment. Culture tubes with 2 mL TB were inoculated with a single colony and appropriate antibiotics and grown overnight at 37 °C, 250 RPM. The next day, 9-mL cultures of TB were inoculated with 1 mL from the pre-culture and grown at 37 °C, 250 RPM to an OD_{600} of 0.4-1.0. Cultures were induced with a final concentration of 1 mM IPTG, 0.91 mM AzF and 0.02% Arabinose depending on the conditions for each experiment described in the results and discussion. Culture tubes were induced at 18 °C for 20 hours and shaken at 250 RPM. The next day induction cultures were centrifuged at 4450 RPM using the Sorvall ST1R Plus-MD centrifuge. And pellets were lysed by resuspension in TG2 Lysis Buffer (20 mM Tris, 150 mM NaCl, 1 mM EDTA, 1 mM TCEP, 0.5% Triton X-100, pH 8.0) and eight freeze/thaw cycles between dry ice and acetone flash freezing and incubation in a 40 °C water bath. Samples were loaded onto an SDS-PAGE as previously described, and His-tagged protein were detected with Anti-His antibody as described in section 3.4.7.

Medium-Scale Expressions

WT TG2 plasmid was either transformed into regular BL21(DE3) cells or BL21 (DE3) cells pre-transformed with pEVOL-pAzF plasmid. Overnight cultures of 5 mL TB were inoculated with single colonies form transformations and grown overnight at 37 °C, 250 RPM. The next day 100-mL TB cultures were inoculated with the 5-mL TB overnight cultures. These cultures were grown at 37 °C, 250 RPM until an OD_{600} of ~0.5 was reached. Flasks were either induced with a

final concentration of 1 mM IPTG or had nothing added. These cultures were shaken overnight at 18 °C, 250 RPM for 20 hours. Protein in each culture was purified according to section 3.4.3 up to Ni-NTA affinity chromatography after which the elution fractions were pooled and loaded on an SDS-PAGE alongside soluble cell lysate samples. This SDS-PAGE was further transferred to a western blot visualized with Anti-His antibody as described in in section 3.4.7.

3.4.5 DBCO-Cy5 Assay

The DBCO-Cy5 (MilliporeSigma) assay was performed by room temperature incubation of 1 μM of WT or AzF mutant TG2 with 2 μM DBCO-Cy5 dye in a 1X PBS solution. These proteins were allowed to react with the DBCO-Cy5 dye for either 30 minutes, or 1 hour. Negative controls had no DBCO-Cy5 added. After incubation, samples were prepared and run on an SDS-PAGE as previously described. SDS-PAGEs were visualized using the stain-free and Cy5 channels on the BioRad Chemidoc imager. Excess dye at the bottom of the SDS-PAGE was removed by cutting off portions.

3.4.6 Crosslinking Assay

The crosslinking assay with 45FN was performed by incubating 2 μM of WT or AzF mutant protein with 4 μM of 45FN for 45 minutes on ice unless stated otherwise. Control samples were also created with only TG2 or only 45FN. The 25 μL sample volumes were split in half and pipetted onto a tinfoil-wrapped 96-well plate. Half of the samples were subjected to either longwave (365 nm) or shortwave (254) UV light with a handheld UV lamp at either 6 watts or 4 watts, depending on the experiment. The other half of the samples were not exposed to UV light but still were covered with a handheld lamp that was not turned on. Samples were prepared and loaded onto either a stain-free SDS-PAGE or an SDS-PAGE stained with Coomassie Blue and ran as previously described. Some crosslinking experiments were visualized with Anti-His antibody according to a western blot protocol in section 3.4.7.

3.4.7 Western Blotting

After running an SDS-PAGE the gel was placed in a petri dish and 10 mL of Transfer Buffer (25 mM Tris, 191 mM Glycine, 20% MeOH) was added. The gel was equilibrated for 1 hr at room temperature. A transfer sandwich was prepared using stacked cathode, filter pad, filter paper, SDS-PAGE, filter paper, filter pad and anode. The components were wetted with Transfer Buffer before stacking. PVDF membrane (Millipore) was cut to match the size of the SDS-PAGE and equilibrated with 100% methanol. The PVDF membrane was placed on top of the SDS-PAGE before closing the transfer sandwich. Air bubbles were removed from the cassette by rolling a glass test tube over the stacked sandwich. The transfer was done using a Mini Trans-Blot ® Electrophoretic Transfer Cell which was filled with Transfer buffer up to the "blotting" line. A frozen blue cooling unit was placed inside the transfer cell to prevent heat build up. The transfer was initiated by running at 40V for 80 minutes. After the transfer was complete, it was carefully removed from the cassette and placed into a petri dish containing 15 mL of Blocking Buffer (100 mL TBS (40 mM Tris-HCl, 150 mM NaCl, pH 7.5), 50 mg/mL powdered milk, 0.1% Tween-20) with the side containing transferred proteins facing upward. Blocking was done overnight on a shaker at 4 °C. After blocking was complete the blocking buffer was decanted, and the membrane was incubated with 5 µL of 1 mg/mL HRP-conjugated His-tag antibody (ThermoFisher) in 5 mL Blocking Buffer. Blocking buffer was carefully decanted and 10 mL of Wash Buffer (100 mL TBS and 0.1% Tween-20) was added to the petri dish on a shaker and shaken for 5 minutes. This step was repeated three times. Detection of the antibody was done using the Millipore Immobilon^{TM} Western Chemiluminescent HRP Substrate (ECL) kit. The ECL mixture was added to the PVDF membrane and equilibrated for 30 seconds. The membrane was imaged using Chemiluminescent detection on a BioRad Chemidoc imager.

Chapter 4: Conclusions and Future Perspectives

TG2's association to 45FN was explored using two different molecular biology techniques: 1) SDM of key residues to alanine mutants, followed by BLI to measure the impact on the protein-protein interaction, and 2) attempts at incorporation of the crosslinking UAA Azido-L-Phenylalanine into TG2 at key residues using GCE technology. SDM and BLI simultaneously corroborated and challenged the impact of key residues identified by Cardoso *et al.*, while not supporting the alternative binding pose proposed by our group. Using this method, more residues that have been previously implicated could be tested and residues that could have stronger interaction based on structure could be chosen, to further confirm or refute our proposed alternative binding pose. AzF incorporation led to the successful formation of a crosslinked 45FN-F203XTG2 complex, but the concentration was not high enough to continue forward with tryptic digestion and MS analysis of paired residues. The incorporation of AzF into TG2 had many issues, such as an ineffective DBCO-Cy5 labelling assay to monitor AzF incorporation, and core issues with protein expression, such as control of IPTG induction and excessive read-through of the TAG codon. If these issues can be addressed, crosslinking-MS studies using UAAs could be a specific and unique tool for monitoring this protein-protein interaction.

The SDM and BLI experiments confirmed that R116 is a crucial residue for the protein-protein interaction of 45FN and TG2 but did not indicate that K30 was crucial when removed on its own. A three-fold increase in K_D was seen for the R116A mutation in our experiment versus WT TG2 and 45FN compared to a 25-fold increase in Cardoso *et al.*'s study.[90] Interestingly, no

significant difference was found when K30A was measured even though Cardoso *et al.* found a 6-fold difference with K30E. We believe our comparison is more representative of K30's role in the protein-protein interaction, since we are removing the positively charged side chain completely while K30E reverses the charge. We also believe our experiment allows more freedom for the native protein-protein interaction to occur since the N-terminal domain of TG2 is immobilized away from the biosensor and 45FN is completely free in solution to bind. Contrastingly, past SPR experiments were based on the use of immobilized 45FN, which may have restricted the ability of the proteins to conform to their native non-covalent interaction. However, this potential problem is not evident, especially since they measured a 13-fold tighter binding affinity of to WT TG2 45FN.

We believe this method's illumination of the TG2/45FN binding mode, and its robustness compared to other studies holds considerable potential. More alanine mutagenesis experiments could reveal whether the immobilization method used in BLI is more representative of the native interaction than stochastic immobilization with SPR. There are also still many key residues that could be studied. One example is the H134 residue, which we were not able to obtain in this work. Cardoso *et al*. report a 9-fold increase with the H134A mutation on SPR and studying this mutation using our method could both support or refute this finding while providing more information on how the two methods compare. Other residues like E120 and T58, also studied by Cardoso *et al.*, are also interesting candidates.[5]

The alanine mutations L114A, I195A and E29A showed no significant difference in FN binding than the WT protein, which does not support our alternative binding pose. Choosing residues implicated in this pose that could have more weight in a non-covalent interaction, like

charged residues instead of hydrophobic residues leucine and isoleucine, may lead to clearer refutation or perhaps confirmation of this binding pose.

For the second method of exploration, AzF incorporation into TG2, some progress was made. The yield of protein produced was overall increased from the first system attempted, pDULE-pCNF, to the final system, pULTRA-pCNF, and a crosslinked 45FN-F203XTG2 was convincingly formed, albeit at a low concentration. Further optimization could push the formation of this complex into quantities suitable for MS analysis of paired residues. Perhaps the crosslinking assay could even be scaled up under the same conditions to produce enough complex.

There were many roadblocks with incorporation, troubleshooting these may lead to more streamlined analysis of AzF mutants. One example of a roadblock was trying to use the DBCO-Cy5 assay to monitor AzF incorporation into TG2. The main issue with this method is that TG2 was labelled brightly even without any AzF added to the culture. It is possible that optimizing the efficiency of AzF incorporation itself could solve this problem by greatly increasing the relative brightness of test compounds, but other methods could also be used. One method uses expression of the protein of interest, in our case TG2, fused to a C-terminal fluorescent protein, which only is expressed when full TG2 is expressed with the non-canonical amino acid.[104] The expression of full-length TG2 could then be monitored by fluorescence, even in impure mixtures.

Besides the labelling assay there was overall trouble expressing TG2 and TG2 mutants. It is worth mentioning that TG2 expression was carried out in the Pezacki Lab alongside successfully purified and AzF incorporated NS3 helicase and still had quite poor results (Supplemental Figure 3), so it is possible that these issues are protein or plasmid specific. The big problems were a lack of IPTG induction control and also excessive read-through of the TAG

codon, described in section 3.2.4. This lack of IPTG induction control was foreshadowed early on during pDULE-pCNF incorporation. When the final concentration of IPTG was increased from 2 μM to 1 mM at induction after realization of an error, the yield of WT or mutant TG2 did not increase.

Plasmids were thoroughly checked through Sanger and full plasmid nanopore sequencing and no abnormalities were seen. Perhaps the Rosetta (DE3) plasmid is crucial for expression of TG2 since it contains rare tRNAs that code for eukaryotic codons present in the TG2 gene. Maybe codon optimization of TG2 for *E. coli* is necessary for AzF incorporation experiments, since GCE machinery cannot be used in tandem with the Rosetta line.

Another consideration is residue selection, picking more conserved residues like phenylalanine and tyrosine for AzF replacement may increase tolerance of the UAA. Perhaps this effect contributed to the success seen with the F203X mutation. Three residues E29X, I195X and H134X were not able to be expressed at all using this system, maybe because of their difference from the UAA. Excessive read-through of the TAG codon points towards issues with the tRNA/aaRS system, suggesting promiscuity for endogenous amino acids. Perhaps using a tRNA/aaRS system evolved from a different tRNA type and organism such as pyrrolysyl-tRNA synthetases from *Methanosarcina mazei*[112] could be used instead of tyrosyl-tRNA synthetases from *Methanocaldococcus jannaschii*, used to derive pEVOL-pAzF and pULTRA-pCNF plasmids.

In conclusion, two unique methods were attempted for exploration of the 45FN/TG2 protein-protein interaction. SDM and BLI are promising techniques and the method can be easily replicated and used to explore more key residues. AzF incorporation was optimized and so far, it appears that the best system to use is the pULTRA-pCNF tRNA/aaRS orthogonal pair, but a

multipronged approach would be necessary for creating a streamlined method for testing AzF incorporation in TG2. A new assay should be attempted instead of DBCO-Cy5 due to its selectivity problem, and core expression issues such as IPTG control would need to be resolved, perhaps through codon optimization of TG2 for *E. coli.* We would recommend tackling these issues first before trying new systems for incorporation. The current protocol for production of the 45FN-F203X TG2 complex could then potentially be scaled up for sufficient amounts to be created for subsequent tryptic digestion and MS analysis of paired residues.

Appendix: Supplemental Figures and Tables

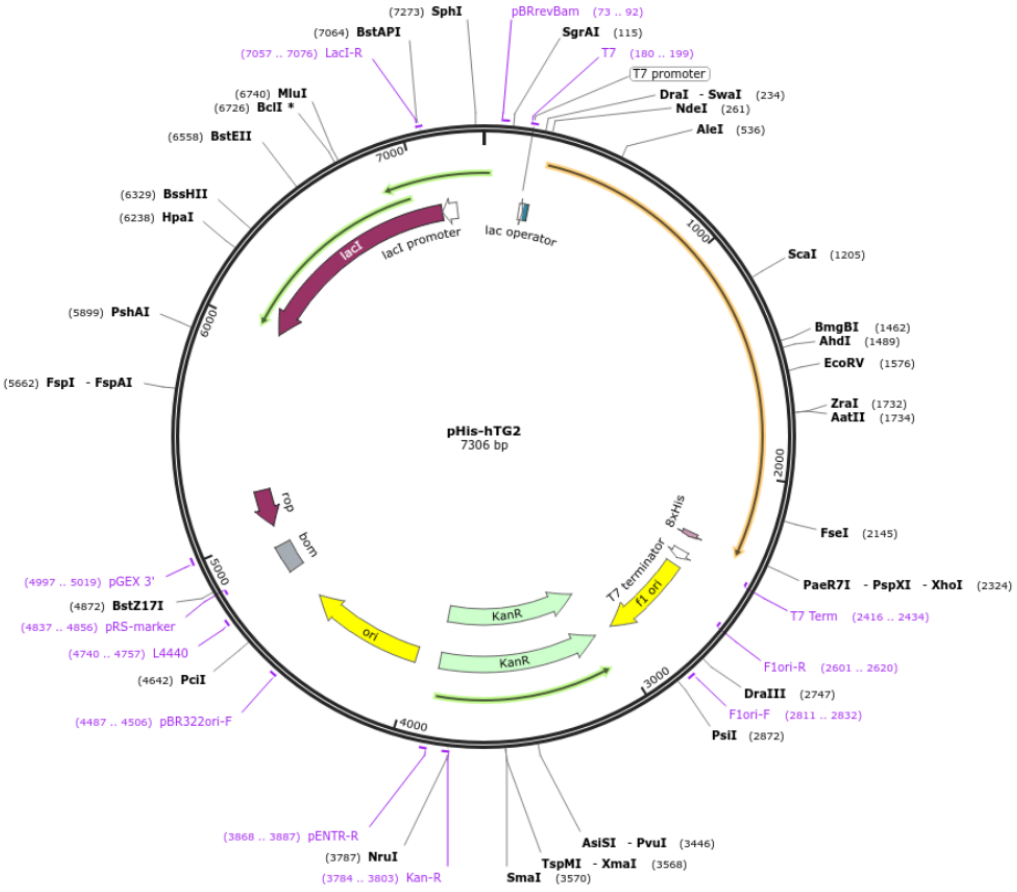


Figure 70. Plasmid Map of Cterm-pHis-hTG2. (Addgene #100719).

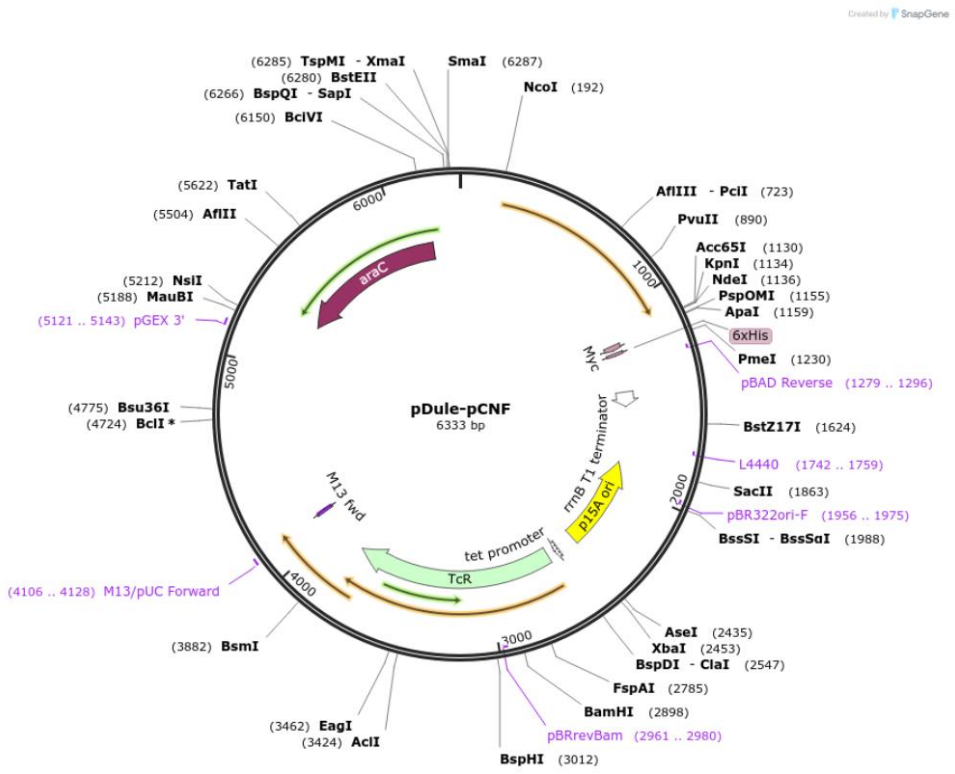


Figure 71. Plasmid Map of pDULE-pCNF. (Addgene #85494).

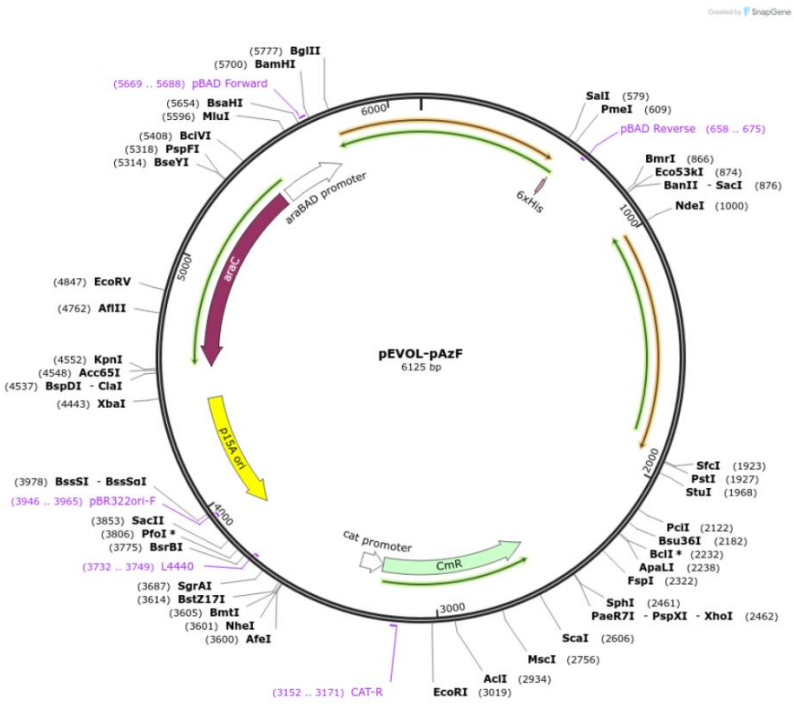


Figure 72. Plasmid Map of pDULE-pCNF. (Addgene #85494).

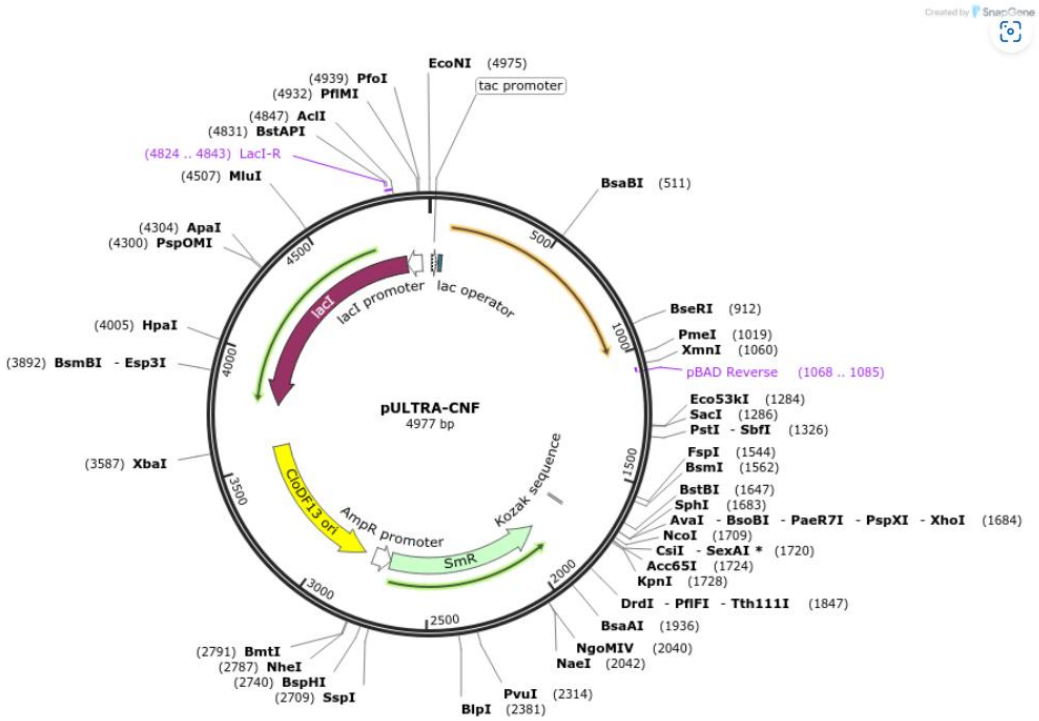


Figure 73. Plasmid Map of pEVOL-pAzF. (Addgene #48215).

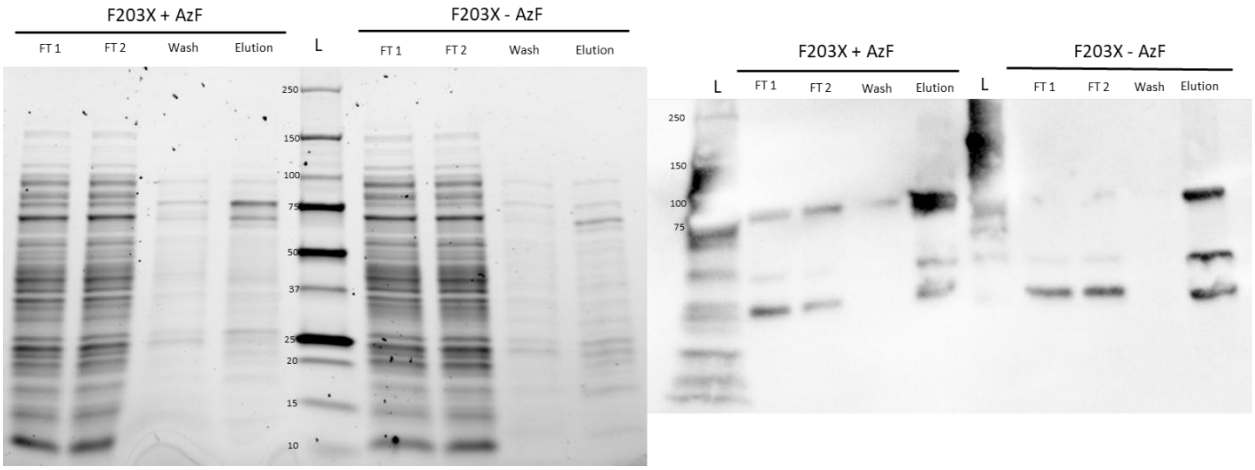


Figure 74. SDS-PAGE and Western Blot of Purification of F203X/pULTRA-pCNF with and without AzF added at induction. Left, SDS-PAGE post Ni-NTA purification. Right, Western Blot of SDS-PAGE using anti-His antibody. Readthrough is still present of the TAG codon and the yield is poor.

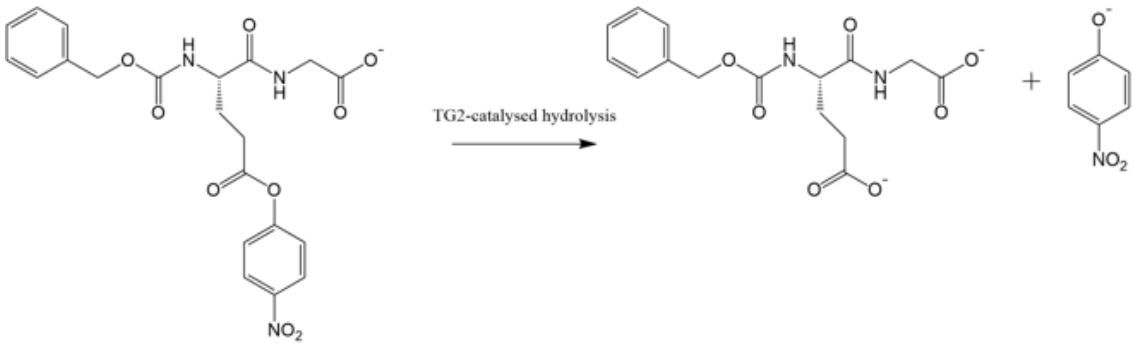


Figure 75. AL5 Activity Assay Reaction Scheme.

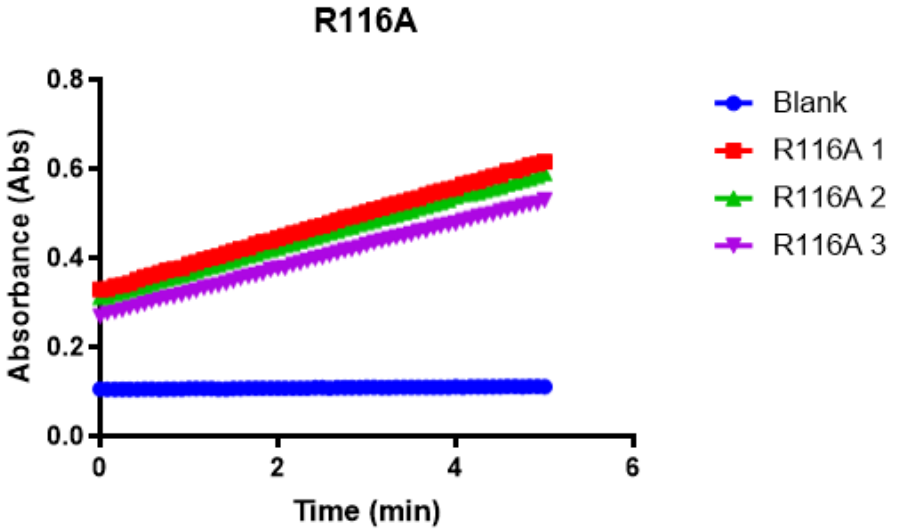


Figure 76. AL5 Activity Assay Curves for R116A. These curves represent the AL5 assay described in the experimental section done with R116A in triplicate and was used to calculate the specific activity shown in Table 1.

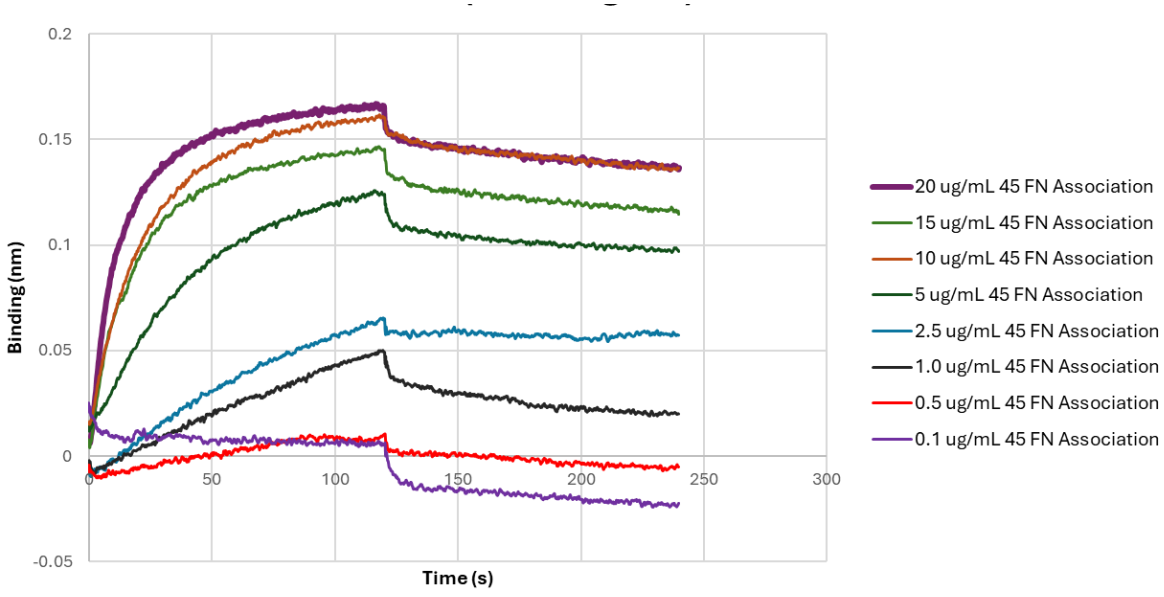


Figure 77. BLI Curves for WT TG2 and 45FN

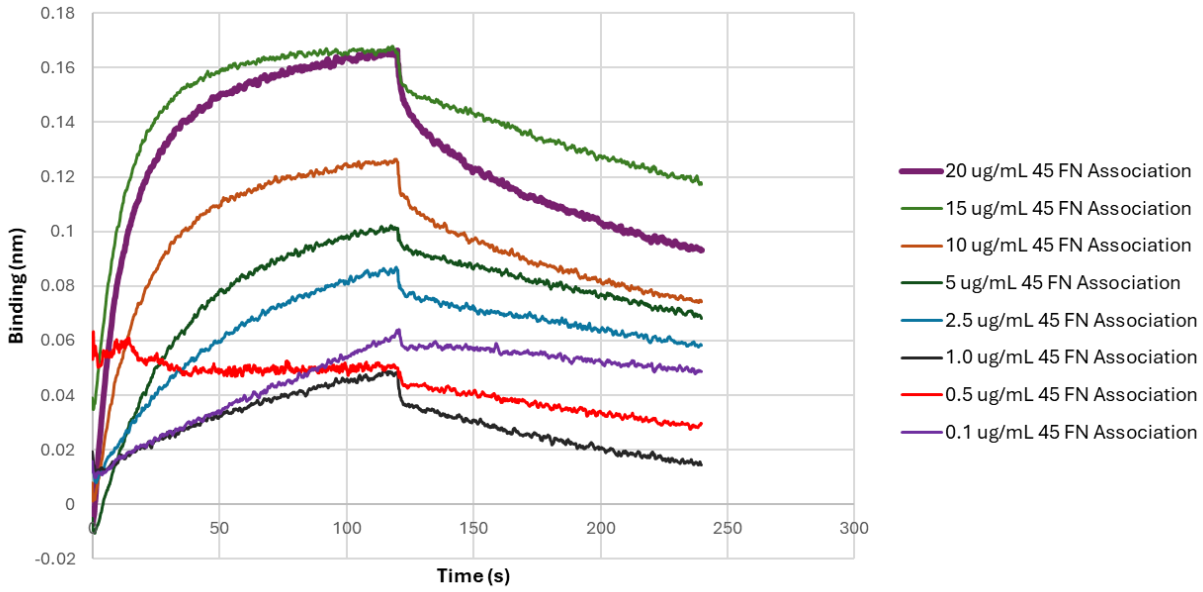


Figure 78. BLI Curves for R116A TG2 and 45FN

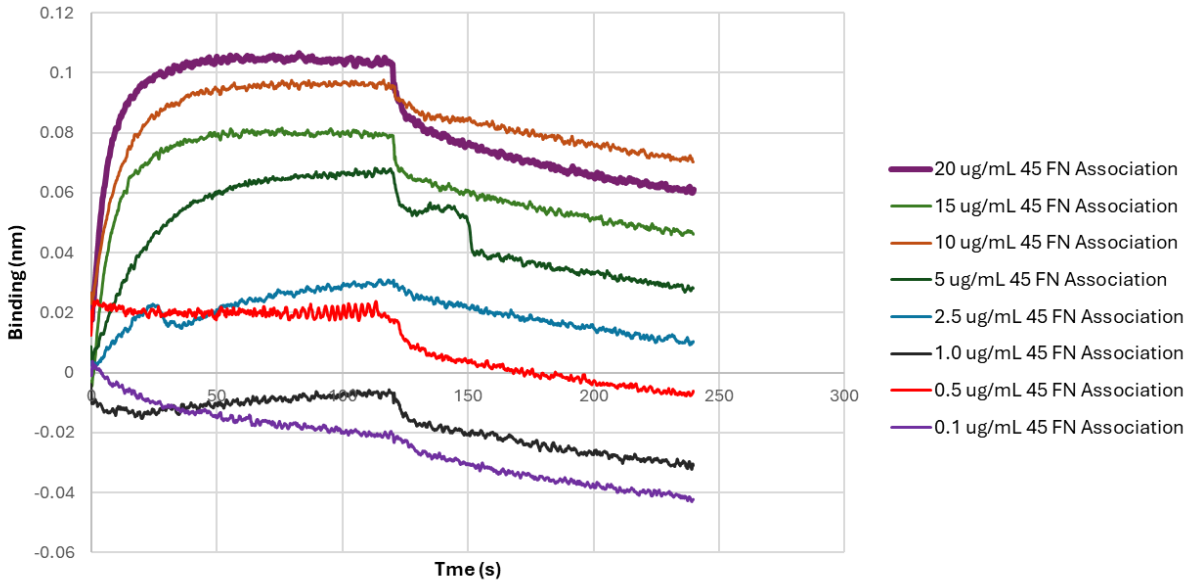


Figure 79. BLI Curves for L114A TG2 and 45FN

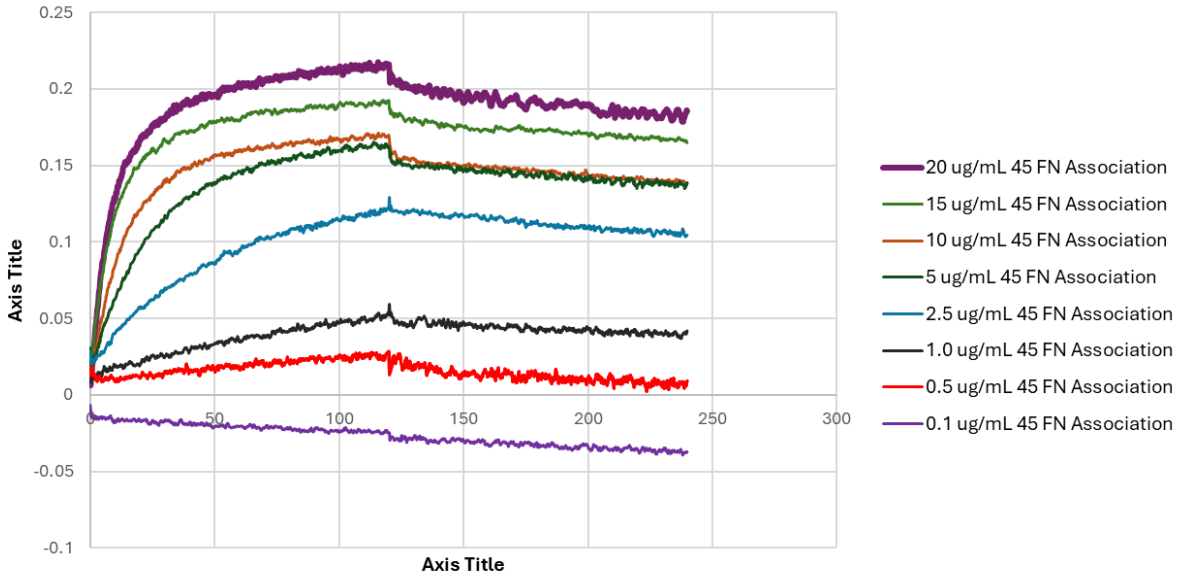


Figure 80. BLI Curves for K30A TG2 and 45FN

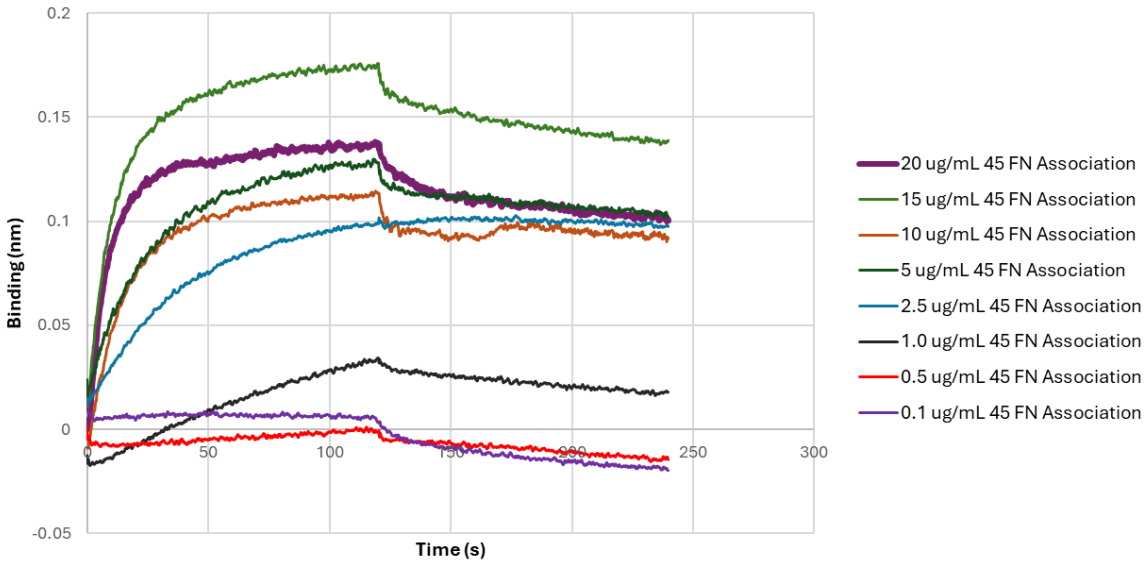


Figure 81. BLI Curves for E29A TG2 and 45FN

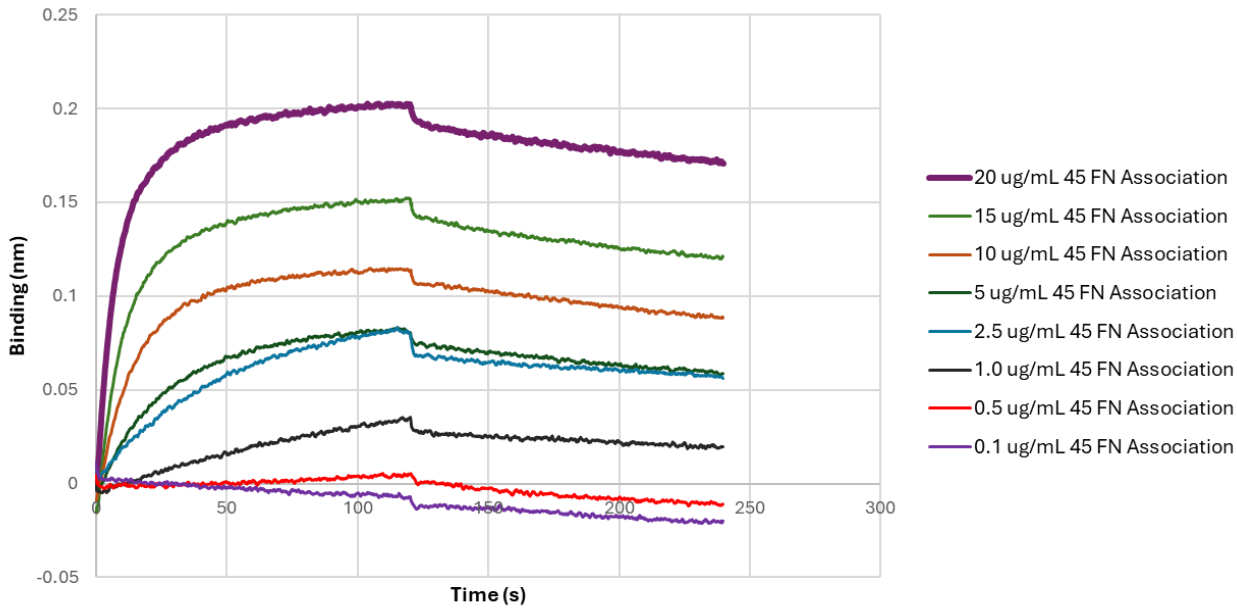


Figure 82. BLI Curves for I195A TG2 and 45FN

Table 6. Forward and Reverse Primer Sequences for Alanine and TAG TG2 Mutants.

Primers were ordered from Integrated DNA Technologies.

Mutation	Forward Primer	Reverse Primer
K30A	5'-GTGCCGAGAGGCGCTGGTGGTGC-3'	5'-AGGTCTGCCGTGTGGTGGTC-3'
L114A	5'-CCCCATCGGCGCGTATCGCCTCAGC-3'	5'-GCGTTAGCCAGGGTGGTGAG-3'
R116A	5'-CGGTCTGTATGCCCTCAGCCTGG-3'	5'-ATGGGGGCGTTAGCCG-3'
I195A	5'-CATCTGCCTGGCGCTTCTAGATGTCAAC-3'	5'-TCTAGGATTCCATCTTCAAACTGC-3'
E29A	5'-CCTGTGTCGGGCGAAGCTGGTG-3'	5'-TCGGCAGTGTGGTGGTCTC-3'
H134A	5'-TGTGCTGGGCGCGTTCATTTTGC-3'	5'-AAGCTAGATCCCTGGTAGCCA-3'
R116X	5'-CGGTCTGTATTAGCTCAGCCTGG-3'	5'-ATGGGGGCGTTAGCCG-3'
K30X	5'-GTGCCGAGAGTAGCTGGTGGTGC-3'	5'-AGGTCTGCCGTGTGGTGGTC-3'
E29X	5'-CCTGTGTCGTAGAAGCTGGTG-3'	5'-TCGGCAGTGTGGTGGTCTC-3'
I195X	5'-CATCTGCCTGTAGCTTCTAGATGTCAA-3'	5'-TCTAGGATTCCATCTTCAAAACTGCC-3'
H134X	5'-TGTGCTGGGCTAGTTCATTTTGC-3'	5'-AAGCTAGATCCTGGTAGCCAG-3'
L114X	5'-CCCCATCGGCTAGTATCGCCTCAG-3'	5'-GCGTTAGCCAGGGTGGTGAG-3'
F203X	5'-CCAAGTAGCTGAAGAACGC-3'	5'-GGTTGACATCTAAAAAAGGATCAGGC-3'
A66X	5'-CGGACCATAGCCTAGCCAGG-3'	5'-GTCACGACACTGAAGGACACTGAAGGTGAG-3'

References

(1) Nurminskaya, M. V.; Belkin, A. M. Cellular Functions of Tissue Transglutaminase. *Int. Rev. Cell Mol. Biol.* **2012**, *294*, 1–97. https://doi.org/10.1016/B978-0-12-394305-7.00001-X.

(2) Akimov, S. S.; Belkin, A. M. Cell-Surface Transglutaminase Promotes Fibronectin Assembly via Interaction with the Gelatin-Binding Domain of Fibronectin: A Role in TGFβ-Dependent Matrix Deposition. *J. Cell Sci.* **2001**, *114* (16), 2989–3000. https://doi.org/10.1242/jcs.114.16.2989.

(3) Soltani, F.; Kaartinen, M. T. Transglutaminases in Fibrosis—Overview and Recent Advances. *Am. J. Physiol.-Cell Physiol.* **2023**, *325* (4), C885–C894. https://doi.org/10.1152/ajpcell.00322.2023.

(4) Satpathy, M.; Cao, L.; Pincheira, R.; Emerson, R.; Bigsby, R.; Nakshatri, H.; Matei, D. Enhanced Peritoneal Ovarian Tumor Dissemination by Tissue Transglutaminase. *Cancer Res.* **2007**, *67* (15), 7194–7202. https://doi.org/10.1158/0008-5472.CAN-07-0307.

(5) Cardoso, I.; Stamnaes, J.; Andersen, J. T.; Melino, G.; Iversen, R.; Sollid, L. M. Transglutaminase 2 Interactions with Extracellular Matrix Proteins as Probed with Celiac Disease Autoantibodies. *FEBS J.* **2015**, *282* (11), 2063–2075. https://doi.org/10.1111/febs.13276.

(6) Griffin, M.; Casadio, R.; Bergamini, C. M. Transglutaminases: Nature's Biological Glues. *Biochem. J.* **2002**, *368* (Pt 2), 377–396. https://doi.org/10.1042/BJ20021234.

(7) Ariëns, R. A. S.; Philippou, H.; Nagaswami, C.; Weisel, J. W.; Lane, D. A.; Grant, P. J. The Factor XIII V34L Polymorphism Accelerates Thrombin Activation of Factor XIII and Affects Cross-Linked Fibrin Structure. *Blood* **2000**, *96* (3), 988–995. https://doi.org/10.1182/blood.V96.3.988.

(8) Komáromi, I.; Bagoly, Z.; Muszbek, L. Factor XIII: Novel Structural and Functional Aspects. *J. Thromb. Haemost.* **2011**, *9* (1), 9–20. https://doi.org/10.1111/j.1538-7836.2010.04070.x.

(9) Yamanishi, K.; Inazawa, J.; Liew, F. M.; Nonomura, K.; Ariyama, T.; Yasuno, H.; Abe, T.; Doi, H.; Hirano, J.; Fukushima, S. Structure of the Gene for Human Transglutaminase 1. *J. Biol. Chem.* **1992**, *267* (25), 17858–17863. https://doi.org/10.1016/S0021-9258(19)37122-4.

(10) Ahvazi, B.; Boeshans, K. M.; Rastinejad, F. The Emerging Structural Understanding of Transglutaminase 3. *J. Struct. Biol.* **2004**, *147* (2), 200–207. https://doi.org/10.1016/j.jsb.2004.03.009.

(11) Lee, J.-H.; Jang, S.-I.; Yang, J.-M.; Markova, N. G.; Steinert, P. M. The Proximal Promoter of the Human Transglutaminase 3 Gene: STRATIFIED SQUAMOUS EPITHELIAL-SPECIFIC EXPRESSION IN CULTURED CELLS IS MEDIATED BY BINDING OF Sp1 AND Ets TRANSCRIPTION FACTORS TO A PROXIMAL PROMOTER ELEMENT (∗). *J. Biol. Chem.* **1996**, *271* (8), 4561–4568. https://doi.org/10.1074/jbc.271.8.4561.

(12) Fesus, L.; Piacentini, M. Transglutaminase 2: An Enigmatic Enzyme with Diverse Functions. *Trends Biochem. Sci.* **2002**, *27* (10), 534–539. https://doi.org/10.1016/S0968-0004(02)02182-5.

(13) Sarkar, N. K.; Clarke, D. D.; Waelsch, H. An Enzymically Catalyzed Incorporation of Amines into Proteins. *Biochim. Biophys. Acta* **1957**, *25*, 451–452. https://doi.org/10.1016/0006-3002(57)90512-7.
(14) Pinkas, D. M.; Strop, P.; Brunger, A. T.; Khosla, C. Transglutaminase 2 Undergoes a Large Conformational Change upon Activation. *PLoS Biol.* **2007**, *5* (12), e327. https://doi.org/10.1371/journal.pbio.0050327.
(15) Navals, P.; Rangaswamy, A. M. M.; Kasyanchyk, P.; Berezovski, M. V.; Keillor, J. W. Conformational Modulation of Tissue Transglutaminase via Active Site Thiol Alkylating Agents: Size Does Not Matter. *Biomolecules* **2024**, *14* (4), 496. https://doi.org/10.3390/biom14040496.
(16) Yao, Z.; Fan, Y.; Lin, L.; Kellems, R. E.; Xia, Y. Tissue Transglutaminase: A Multifunctional and Multisite Regulator in Health and Disease. *Physiol. Rev.* **2024**, *104* (1), 281–325. https://doi.org/10.1152/physrev.00003.2023.
(17) Pavlyukov, M. S.; Antipova, N. V.; Balashova, M. V.; Shakhparonov, M. I. Detection of Transglutaminase 2 Conformational Changes in Living Cell. *Biochem. Biophys. Res. Commun.* **2012**, *421* (4), 773–779. https://doi.org/10.1016/j.bbrc.2012.04.082.
(18) Stamnaes, J.; Pinkas, D. M.; Fleckenstein, B.; Khosla, C.; Sollid, L. M. Redox Regulation of Transglutaminase 2 Activity. *J. Biol. Chem.* **2010**, *285* (33), 25402–25409. https://doi.org/10.1074/jbc.M109.097162.
(19) Groenen, P. J. T. A.; Bloemendal, H.; de JONG, W. W. The Carboxy-Terminal Lysine of αB-Crystallin Is an Amine-Donor Substrate for Tissue Transglutaminase. *Eur. J. Biochem.* **1992**, *205* (2), 671–674. https://doi.org/10.1111/j.1432-1033.1992.tb16827.x.
(20) Fesus, L.; Metsis, M. L.; Muszbek, L.; KoTELiansky, V. E. Transglutaminase-Sensitive Glutamine Residues of Human Plasma Fibronectin Reveled by Studying Its Proteolytic Fragments. *Eur. J. Biochem.* **1986**, *154* (2), 371–374. https://doi.org/10.1111/j.1432-1033.1986.tb09407.x.
(21) Bowness, J. M.; Folk, J. E.; Timpl, R. Identification of a Substrate Site for Liver Transglutaminase on the Aminopropeptide of Type III Collagen. *J. Biol. Chem.* **1987**, *262* (3), 1022–1024. https://doi.org/10.1016/S0021-9258(19)75743-3.
(22) Nunes, I.; Gleizes, P.-E.; Metz, C. N.; Rifkin, D. B. Latent Transforming Growth Factor-β Binding Protein Domains Involved in Activation and Transglutaminase-Dependent Cross-Linking of Latent Transforming Growth Factor-β. *J. Cell Biol.* **1997**, *136* (5), 1151–1163. https://doi.org/10.1083/jcb.136.5.1151.
(23) Csosz, E.; Bagossi, P.; Nagy, Z.; Dosztanyi, Z.; Simon, I.; Fesus, L. Substrate Preference of Transglutaminase 2 Revealed by Logistic Regression Analysis and Intrinsic Disorder Examination. *J. Mol. Biol.* **2008**, *383* (2), 390–402. https://doi.org/10.1016/j.jmb.2008.08.026.
(24) Cooper, A. J. L.; Sheu, K.-F. R.; Burke, J. R.; Onodera, O.; Strittmatter, W. J.; Roses, A. D.; Blass, J. P. Transglutaminase-Catalyzed Inactivation of Glyceraldehyde 3-Phosphate Dehydrogenase and α-Ketoglutarate Dehydrogenase Complex by Polyglutamine Domains of Pathological Length. *Proc. Natl. Acad. Sci.* **1997**, *94* (23), 12604–12609. https://doi.org/10.1073/pnas.94.23.12604.
(25) Verderio, E.; Nicholas, B.; Gross, S.; Griffin, M. Regulated Expression of Tissue Transglutaminase in Swiss 3T3 Fibroblasts: Effects on the Processing of Fibronectin, Cell

Attachment, and Cell Death. *Exp. Cell Res.* **1998**, *239* (1), 119–138. https://doi.org/10.1006/excr.1997.3874.
(26) Walther, D. J.; Peter, J.-U.; Winter, S.; Höltje, M.; Paulmann, N.; Grohmann, M.; Vowinckel, J.; Alamo-Bethencourt, V.; Wilhelm, C. S.; Ahnert-Hilger, G.; Bader, M. Serotonylation of Small GTPases Is a Signal Transduction Pathway That Triggers Platelet α-Granule Release. *Cell* **2003**, *115* (7), 851–862. https://doi.org/10.1016/S0092-8674(03)01014-6.
(27) Gundemir, S.; Colak, G.; Tucholski, J.; Johnson, G. V. W. Transglutaminase 2: A Molecular Swiss Army Knife. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2012**, *1823* (2), 406–419. https://doi.org/10.1016/j.bbamcr.2011.09.012.
(28) Nakaoka, H.; Perez, D. M.; Baek, K. J.; Das, T.; Husain, A.; Misono, K.; Im, M.-J.; Graham, R. M. Gh: A GTP-Binding Protein with Transglutaminase Activity and Receptor Signaling Function. *Science* **1994**, *264* (5165), 1593–1596. https://doi.org/10.1126/science.7911253.
(29) Vezza, R.; Habib, A.; FitzGerald, G. A. Differential Signaling by the Thromboxane Receptor Isoforms via the Novel GTP-Binding Protein, Gh. *J. Biol. Chem.* **1999**, *274* (18), 12774–12779. https://doi.org/10.1074/jbc.274.18.12774.
(30) Baek, K. J.; Kwon, N. S.; Lee, H. S.; Kim, M. S.; Muralidhar, P.; Im, M. J. Oxytocin Receptor Couples to the 80 kDa Gh Alpha Family Protein in Human Myometrium. *Biochem. J.* **1996**, *315 (Pt 3)* (Pt 3), 739–744.
(31) Lee, C. S.; Park, H. H. Structural Aspects of Transglutaminase 2: Functional, Structural, and Regulatory Diversity. *Apoptosis* **2017**, *22* (9), 1057–1068. https://doi.org/10.1007/s10495-017-1396-9.
(32) Bootman, M. D.; Collins, T. J.; Peppiatt, C. M.; Prothero, L. S.; MacKenzie, L.; De Smet, P.; Travers, M.; Tovey, S. C.; Seo, J. T.; Berridge, M. J.; Ciccolini, F.; Lipp, P. Calcium Signalling—an Overview. *Semin. Cell Dev. Biol.* **2001**, *12* (1), 3–10. https://doi.org/10.1006/scdb.2000.0211.
(33) Buday, L.; Tompa, P. Functional Classification of Scaffold Proteins and Related Molecules. *FEBS J.* **2010**, *277* (21), 4348–4355. https://doi.org/10.1111/j.1742-4658.2010.07864.x.
(34) Belkin, A. M. Extracellular TG2: Emerging Functions and Regulation. *FEBS J.* **2011**, *278* (24), 4704–4716. https://doi.org/10.1111/j.1742-4658.2011.08346.x.
(35) Zemskov, E. A.; Mikhailenko, I.; Hsia, R.-C.; Zaritskaya, L.; Belkin, A. M. Unconventional Secretion of Tissue Transglutaminase Involves Phospholipid-Dependent Delivery into Recycling Endosomes. *PLoS ONE* **2011**, *6* (4), e19414. https://doi.org/10.1371/journal.pone.0019414.
(36) Wang, Z.; Griffin, M. TG2, a Novel Extracellular Protein with Multiple Functions. *Amino Acids* **2012**, *42* (2), 939–949. https://doi.org/10.1007/s00726-011-1008-x.
(37) Scarpellini, A.; Germack, R.; Lortat-Jacob, H.; Muramatsu, T.; Billett, E.; Johnson, T.; Verderio, E. A. M. Heparan Sulfate Proteoglycans Are Receptors for the Cell-Surface Trafficking and Biological Activity of Transglutaminase-2. *J. Biol. Chem.* **2009**, *284* (27), 18411–18423. https://doi.org/10.1074/jbc.M109.012948.
(38) Yamaguchi, Y.; Mann, D. M.; Ruoslahti, E. Negative Regulation of Transforming Growth Factor-β by the Proteoglycan Decorin. *Nature* **1990**, *346* (6281), 281–284. https://doi.org/10.1038/346281a0.
(39) Theocharis, A. D.; Skandalis, S. S.; Gialeli, C.; Karamanos, N. K. Extracellular Matrix Structure. *Adv. Drug Deliv. Rev.* **2016**, *97*, 4–27. https://doi.org/10.1016/j.addr.2015.11.001.

(40) Yue, B. Biology of the Extracellular Matrix: An Overview. *J. Glaucoma* **2014**, *23*, S20. https://doi.org/10.1097/IJG.0000000000000108.
(41) Signal Transduction from the Extracellular Matrix. *J. Cell Biol.* **1993**, *120* (3), 577–585. https://doi.org/10.1083/jcb.120.3.577.
(42) Frantz, C.; Stewart, K. M.; Weaver, V. M. The Extracellular Matrix at a Glance. *J. Cell Sci.* **2010**, *123* (24), 4195–4200. https://doi.org/10.1242/jcs.023820.
(43) Danielson, K. G.; Baribault, H.; Holmes, D. F.; Graham, H.; Kadler, K. E.; Iozzo, R. V. Targeted Disruption of Decorin Leads to Abnormal Collagen Fibril Morphology and Skin Fragility. *J. Cell Biol.* **1997**, *136* (3), 729–743. https://doi.org/10.1083/jcb.136.3.729.
(44) Furini, G.; Verderio, E. A. M. Spotlight on the Transglutaminase 2-Heparan Sulfate Interaction. *Med. Sci.* **2019**, *7* (1), 5. https://doi.org/10.3390/medsci7010005.
(45) Vuoriluoto, K.; Jokinen, J.; Kallio, K.; Salmivirta, M.; Heino, J.; Ivaska, J. Syndecan-1 Supports Integrin A2β1-Mediated Adhesion to Collagen. *Exp. Cell Res.* **2008**, *314* (18), 3369–3381. https://doi.org/10.1016/j.yexcr.2008.07.005.
(46) Telci, D.; Wang, Z.; Li, X.; Verderio, E. A. M.; Humphries, M. J.; Baccarini, M.; Basaga, H.; Griffin, M. Fibronectin-Tissue Transglutaminase Matrix Rescues RGD-Impaired Cell Adhesion through Syndecan-4 and B1 Integrin Co-Signaling. *J. Biol. Chem.* **2008**, *283* (30), 20937–20947. https://doi.org/10.1074/jbc.M801763200.
(47) Karamanos, N. K.; Theocharis, A. D.; Piperigkou, Z.; Manou, D.; Passi, A.; Skandalis, S. S.; Vynios, D. H.; Orian-Rousseau, V.; Ricard-Blum, S.; Schmelzer, C. E. H.; Duca, L.; Durbeej, M.; Afratis, N. A.; Troeberg, L.; Franchi, M.; Masola, V.; Onisto, M. A Guide to the Composition and Functions of the Extracellular Matrix. *FEBS J.* **2021**, *288* (24), 6850–6912. https://doi.org/10.1111/febs.15776.
(48) Xiong, J.-P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M.; Goodman, S. L.; Arnaout, M. A. Crystal Structure of the Extracellular Segment of Integrin αVβ3 in Complex with an Arg-Gly-Asp Ligand. *Science* **2002**, *296* (5565), 151–155. https://doi.org/10.1126/science.1069040.
(49) Danen, E. H. J. Integrins: An Overview of Structural and Functional Aspects. In *Madame Curie Bioscience Database [Internet]*; Landes Bioscience, 2013.
(50) Akimov, S. S.; Krylov, D.; Fleischman, L. F.; Belkin, A. M. Tissue Transglutaminase Is an Integrin-Binding Adhesion Coreceptor for Fibronectin. *J. Cell Biol.* **2000**, *148* (4), 825–838. https://doi.org/10.1083/jcb.148.4.825.
(51) Akimov, S. S.; Belkin, A. M. Cell Surface Tissue Transglutaminase Is Involved in Adhesion and Migration of Monocytic Cells on Fibronectin. *Blood* **2001**, *98* (5), 1567–1576. https://doi.org/10.1182/blood.V98.5.1567.
(52) Bella, J.; Hulmes, D. J. S. Fibrillar Collagens. In *Fibrous Proteins: Structures and Mechanisms*; Parry, D. A. D., Squire, J. M., Eds.; Springer International Publishing: Cham, 2017; pp 457–490. https://doi.org/10.1007/978-3-319-49674-0_14.
(53) Turner, P. M.; Lorand, L. Complexation of Fibronectin with Tissue Transglutaminase. *Biochemistry* **1989**, *28* (2), 628–635. https://doi.org/10.1021/bi00428a032.
(54) Longtin, R. Birthday of a Breakthrough: Fibronectin Research Proves Important, But Not As Originally Expected. *JNCI J. Natl. Cancer Inst.* **2004**, *96* (1), 6–8. https://doi.org/10.1093/jnci/96.1.6.

(55) Pankov, R.; Yamada, K. M. Fibronectin at a Glance. *J. Cell Sci.* **2002**, *115* (20), 3861–3863. https://doi.org/10.1242/jcs.00059.
(56) Woods, A.; Couchman, J. R. Syndecans: Synergistic Activators of Cell Adhesion. *Trends Cell Biol.* **1998**, *8* (5), 189–192. https://doi.org/10.1016/S0962-8924(98)01244-6.
(57) Zollinger, A. J.; Smith, M. L. Fibronectin, the Extracellular Glue. *Matrix Biol.* **2017**, *60–61*, 27–37. https://doi.org/10.1016/j.matbio.2016.07.011.
(58) Kuusela, P. Fibronectin Binds to Staphylococcus Aureus. *Nature* **1978**, *276* (5689), 718–720. https://doi.org/10.1038/276718a0.
(59) Wu, C.; Keivenst, V. M.; O'Toole, T. E.; McDonald, J. A.; Ginsberg, M. H. Integrin Activation and Cytoskeletal Interaction Are Essential for the Assembly of a Fibronectin Matrix. *Cell* **1995**, *83* (5), 715–724. https://doi.org/10.1016/0092-8674(95)90184-1.
(60) Christopher, R. A.; Kowalczyk, A. P.; McKeown-Longo, P. J. Localization of Fibronectin Matrix Assembly Sites on Fibroblasts and Endothelial Cells. *J. Cell Sci.* **1997**, *110* (5), 569–581. https://doi.org/10.1242/jcs.110.5.569.
(61) Sottile, J.; Hocking, D. C. Fibronectin Polymerization Regulates the Composition and Stability of Extracellular Matrix Fibrils and Cell-Matrix Adhesions. *Mol. Biol. Cell* **2002**, *13* (10), 3546–3559. https://doi.org/10.1091/mbc.e02-01-0048.
(62) Dallas, S. L.; Sivakumar, P.; Jones, C. J. P.; Chen, Q.; Peters, D. M.; Mosher, D. F.; Humphries, M. J.; Kielty, C. M. Fibronectin Regulates Latent Transforming Growth Factor-Beta (TGF Beta) by Controlling Matrix Assembly of Latent TGF Beta-Binding Protein-1. *J. Biol. Chem.* **2005**, *280* (19), 18871–18880. https://doi.org/10.1074/jbc.M410762200.
(63) Mardon, H. J.; Grant, K. E. The Role of the Ninth and Tenth Type III Domains of Human Fibronectin in Cell Adhesion. *FEBS Lett.* **1994**, *340* (3), 197–201. https://doi.org/10.1016/0014-5793(94)80137-1.
(64) Hubbard, B.; Buczek-Thomas, J. A.; Nugent, M. A.; Smith, M. L. Fibronectin Fiber Extension Decreases Cell Spreading and Migration. *J. Cell. Physiol.* **2016**, *231* (8), 1728–1736. https://doi.org/10.1002/jcp.25271.
(65) Radek, J. T.; Jeong, J. M.; Murthy, S. N.; Ingham, K. C.; Lorand, L. Affinity of Human Erythrocyte Transglutaminase for a 42-kDa Gelatin-Binding Fragment of Human Plasma Fibronectin. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (8), 3152–3156. https://doi.org/10.1073/pnas.90.8.3152.
(66) Jeong, J. M.; Murthy, S. N.; Radek, J. T.; Lorand, L. The Fibronectin-Binding Domain of Transglutaminase. *J. Biol. Chem.* **1995**, *270* (10), 5654–5658. https://doi.org/10.1074/jbc.270.10.5654.
(67) Wang, Z.; Collighan, R. J.; Gross, S. R.; Danen, E. H. J.; Orend, G.; Telci, D.; Griffin, M. RGD-Independent Cell Adhesion via a Tissue Transglutaminase-Fibronectin Matrix Promotes Fibronectin Fibril Deposition and Requires Syndecan-4/2 and A5β1 Integrin Co-Signaling*. *J. Biol. Chem.* **2010**, *285* (51), 40212–40229. https://doi.org/10.1074/jbc.M110.123703.
(68) Verderio, E. A. M.; Telci, D.; Okoye, A.; Melino, G.; Griffin, M. A Novel RGD-Independent Cell Adhesion Pathway Mediated by Fibronectin-Bound Tissue Transglutaminase Rescues Cells from Anoikis*. *J. Biol. Chem.* **2003**, *278* (43), 42604–42614. https://doi.org/10.1074/jbc.M303303200.
(69) Telci, D.; Wang, Z.; Li, X.; Verderio, E. A. M.; Humphries, M. J.; Baccarini, M.; Basaga, H.; Griffin, M. Fibronectin-Tissue Transglutaminase Matrix Rescues RGD-Impaired Cell

Adhesion through Syndecan-4 and B1 Integrin Co-Signaling. *J. Biol. Chem.* **2008**, *283* (30), 20937–20947. https://doi.org/10.1074/jbc.M801763200.
(70) Singh, P.; Carraher, C.; Schwarzbauer, J. E. Assembly of Fibronectin Extracellular Matrix. *Annu. Rev. Cell Dev. Biol.* **2010**, *26* (Volume 26, 2010), 397–419. https://doi.org/10.1146/annurev-cellbio-100109-104020.
(71) Klöck, C.; DiRaimondo, T. R.; Khosla, C. Role of Transglutaminase 2 in Celiac Disease Pathogenesis. *Semin. Immunopathol.* **2012**, *34* (4), 513–522. https://doi.org/10.1007/s00281-012-0305-0.
(72) Budi, E. H.; Schaub, J. R.; Decaris, M.; Turner, S.; Derynck, R. TGF-β as a Driver of Fibrosis: Physiological Roles and Therapeutic Opportunities. *J. Pathol.* **2021**, *254* (4), 358–373. https://doi.org/10.1002/path.5680.
(73) *Tissue transgluaminase 2 expression in meningiomas | springermedicine.com*. https://www.springermedicine.com/tissue-transgluaminase-2-expression-in-meningiomas/21711786 (accessed 2026-01-19).
(74) Yuan, L.; Siegel, M.; Choi, K.; Khosla, C.; Miller, C. R.; Jackson, E. N.; Piwnica-Worms, D.; Rich, K. M. Transglutaminase 2 Inhibitor, KCC009, Disrupts Fibronectin Assembly in the Extracellular Matrix and Sensitizes Orthotopic Glioblastomas to Chemotherapy. *Oncogene* **2007**, *26* (18), 2563–2573. https://doi.org/10.1038/sj.onc.1210048.
(75) Chen, J.; Agarwal, N.; Mehta, K. Chen JS, Agarwal N, Mehta K.. Multidrug-Resistant MCF-7 Breast Carcinoma Cells Contain Deficient Intracellular Calcium Pools. Br Cancer Res Treat 71: 237-247. *Breast Cancer Res. Treat.* **2002**, *71*, 237–247. https://doi.org/10.1023/A:1014461832403.
(76) Cordes, N.; Blaese, M. A.; Plasswilm, L.; Rodemann, H. P.; Van Beuningen, D. Fibronectin and Laminin Increase Resistance to Ionizing Radiation and the Cytotoxic Drug Ukrain in Human Tumour and Normal Cells in Vitro. *Int. J. Radiat. Biol.* **2003**, *79* (9), 709–720. https://doi.org/10.1080/09553000310001610240.
(77) Khwaja, A.; Rodriguez-Viciana, P.; Wennström, S.; Warne, P. H.; Downward, J. Matrix Adhesion and Ras Transformation Both Activate a Phosphoinositide 3-OH Kinase and Protein Kinase B/Akt Cellular Survival Pathway. *EMBO J.* **1997**, *16* (10), 2783–2793. https://doi.org/10.1093/emboj/16.10.2783.
(78) Qian, F.; Zhang, Z.-C.; Wu, X.-F.; Li, Y.-P.; Xu, Q. Interaction between Integrin Alpha(5) and Fibronectin Is Required for Metastasis of B16F10 Melanoma Cells. *Biochem. Biophys. Res. Commun.* **2005**, *333* (4), 1269–1275. https://doi.org/10.1016/j.bbrc.2005.06.039.
(79) Cardoso, I.; Østerlund, E. C.; Stamnaes, J.; Iversen, R.; Andersen, J. T.; Jørgensen, T. J. D.; Sollid, L. M. Dissecting the Interaction between Transglutaminase 2 and Fibronectin. *Amino Acids* **2017**, *49* (3), 489–500. https://doi.org/10.1007/s00726-016-2296-y.
(80) Soluri, M. F.; Boccafoschi, F.; Cotella, D.; Moro, L.; Forestieri, G.; Autiero, I.; Cavallo, L.; Oliva, R.; Griffin, M.; Wang, Z.; Santoro, C.; Sblattero, D. Mapping the Minimum Domain of the Fibronectin Binding Site on Transglutaminase 2 (TG2) and Its Importance in Mediating Signaling, Adhesion, and Migration in TG2-Expressing Cells. *FASEB J.* **2019**, *33* (2), 2327–2342. https://doi.org/10.1096/fj.201800054RRR.
(81) Selcuk, K.; Leitner, A.; Braun, L.; Le Blanc, F.; Pacak, P.; Pot, S.; Vogel, V. Transglutaminase 2 Has Higher Affinity for Relaxed than for Stretched Fibronectin Fibers. *Matrix Biol.* **2024**, *125*, 113–132. https://doi.org/10.1016/j.matbio.2023.12.006.

(82) Hang, J.; Zemskov, E. A.; Lorand, L.; Belkin, A. M. Identification of a Novel Recognition Sequence for Fibronectin within the NH2-Terminal Beta-Sandwich Domain of Tissue Transglutaminase. *J. Biol. Chem.* **2005**, *280* (25), 23675–23683. https://doi.org/10.1074/jbc.M503323200.
(83) Jug, A.; Bratkovič, T.; Ilaš, J. Biolayer Interferometry and Its Applications in Drug Discovery and Development. *TrAC Trends Anal. Chem.* **2024**, *176*, 117741. https://doi.org/10.1016/j.trac.2024.117741.
(84) Orthwein, T.; Huergo, L. F.; Forchhammer, K.; Selim, K. A. Kinetic Analysis of a Protein-Protein Complex to Determine Its Dissociation Constant (KD) and the Effective Concentration (EC50) of an Interplaying Effector Molecule Using Bio-Layer Interferometry. *Bio-Protoc.* **2021**, *11* (17), e4152. https://doi.org/10.21769/BioProtoc.4152.
(85) Cundy, N. J.; Arciszewski, J.; Gates, E. W. J.; Acton, S. L.; Passley, K. D.; Awoonor-Williams, E.; Boyd, E. K.; Xu, N.; Pierson, É.; Fernandez-Ansieta, C.; Albert, M. R.; McNeil, N. M. R.; Adhikary, G.; Eckert, R. L.; Keillor, J. W. Novel Irreversible Peptidic Inhibitors of Transglutaminase 2. *RSC Med. Chem. 14* (2), 378–385. https://doi.org/10.1039/d2md00417h.
(86) Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72*, 248–254. https://doi.org/10.1016/0003-2697(76)90527-3.
(87) Leblanc, A.; Gravel, C.; Labelle, J.; Keillor, J. W. Kinetic Studies of Guinea Pig Liver Transglutaminase Reveal a General-Base-Catalyzed Deacylation Mechanism. *Biochemistry* **2001**, *40* (28), 8335–8342. https://doi.org/10.1021/bi0024097.
(88) Roy, I.; Smith, O.; Clouthier, C. M.; Keillor, J. W. Expression, Purification and Kinetic Characterisation of Human Tissue Transglutaminase. *Protein Expr. Purif.* **2013**, *87* (1), 41–46. https://doi.org/10.1016/j.pep.2012.10.002.
(89) Han, B.-G.; Cho, J.-W.; Cho, Y. D.; Jeong, K.-C.; Kim, S.-Y.; Lee, B. I. Crystal Structure of Human Transglutaminase 2 in Complex with Adenosine Triphosphate. *Int. J. Biol. Macromol.* **2010**, *47* (2), 190–195. https://doi.org/10.1016/j.ijbiomac.2010.04.023.
(90) Sima, L. E.; Yakubov, B.; Zhang, S.; Condello, S.; Grigorescu, A. A.; Nwani, N. G.; Chen, L.; Schiltz, G. E.; Arvanitis, C.; Zhang, Z.-Y.; Matei, D. Small Molecules Target the Interaction between Tissue Transglutaminase and Fibronectin. *Mol. Cancer Ther.* **2019**, *18* (6), 1057–1068. https://doi.org/10.1158/1535-7163.MCT-18-1148.
(91) Shandell, M. A.; Tan, Z.; Cornish, V. W. Genetic Code Expansion: A Brief History and Perspective. *Biochemistry* **2021**, *60* (46), 3455–3469. https://doi.org/10.1021/acs.biochem.1c00286.
(92) Chapeville, F.; Lipmann, F.; Ehrenstein, G. von; Weisblum, B.; Ray, W. J.; Benzer, S. ON THE ROLE OF SOLUBLE RIBONUCLEIC ACID IN CODING FOR AMINO ACIDS*. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48* (6), 1086–1092. https://doi.org/10.1073/pnas.48.6.1086.
(93) Calendar, R.; Berg, P. D-Tyrosyl RNA: Formation, Hydrolysis and Utilization for Protein Synthesis. *J. Mol. Biol.* **1967**, *26* (1), 39–54. https://doi.org/10.1016/0022-2836(67)90259-8.
(94) Fahnestock, S.; Rich, A. Ribosome-Catalyzed Polyester Formation. *Science* **1971**, *173* (3994), 340–343. https://doi.org/10.1126/science.173.3994.340.

(95) Johnson, A. E.; Woodward, W. R.; Herbert, E.; Menninger, J. R. Nepsilon-Acetyllysine Transfer Ribonucleic Acid: A Biologically Active Analogue of Aminoacyl Transfer Ribonucleic Acids. *Biochemistry* **1976**, *15* (3), 569–575. https://doi.org/10.1021/bi00648a018.
(96) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. A General Method for Site-Specific Incorporation of Unnatural Amino Acids into Proteins. *Science* **1989**, *244* (4901), 182–188. https://doi.org/10.1126/science.2649980.
(97) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the Genetic Code of Escherichia Coli. *Science* **2001**, *292* (5516), 498–500. https://doi.org/10.1126/science.1060077.
(98) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the Genetic Code of Escherichia Coli. *Science* **2001**, *292* (5516), 498–500. https://doi.org/10.1126/science.1060077.
(99) Huang, Y.; Zhang, P.; Wang, H.; Chen, Y.; Liu, T.; Luo, X. Genetic Code Expansion: Recent Developments and Emerging Applications. *Chem. Rev.* **2025**, *125* (2), 523–598. https://doi.org/10.1021/acs.chemrev.4c00216.
(100) Jackson, J. C.; Duffy, S. P.; Hess, K. R.; Mehl, R. A. Improving Nature's Enzyme Active Site with Genetically Encoded Unnatural Amino Acids. *J. Am. Chem. Soc.* **2006**, *128* (34), 11124–11127. https://doi.org/10.1021/ja061099y.
(101) Dieterich, D. C.; Link, A. J.; Graumann, J.; Tirrell, D. A.; Schuman, E. M. Selective Identification of Newly Synthesized Proteins in Mammalian Cells Using Bioorthogonal Noncanonical Amino Acid Tagging (BONCAT). *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (25), 9482–9487. https://doi.org/10.1073/pnas.0601637103.
(102) Park, H.-S.; Hohn, M. J.; Umehara, T.; Guo, L.-T.; Osborne, E. M.; Benner, J.; Noren, C. J.; Rinehart, J.; Söll, D. Expanding the Genetic Code of Escherichia Coli with Phosphoserine. *Science* **2011**, *333* (6046), 1151–1154. https://doi.org/10.1126/science.1207203.
(103) Piersimoni, L.; Kastritis, P. L.; Arlt, C.; Sinz, A. Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein–Protein Interactions–A Method for All Seasons. *Chem. Rev.* **2022**, *122* (8), 7500–7531. https://doi.org/10.1021/acs.chemrev.1c00786.
(104) Cooley, R. B.; Sondermann, H. Probing Protein–Protein Interactions with Genetically Encoded Photoactivatable Cross-Linkers. In *c-di-GMP Signaling: Methods and Protocols*; Sauer, K., Ed.; Springer: New York, NY, 2017; pp 331–345. https://doi.org/10.1007/978-1-4939-7240-1_26.
(105) Lundrigan, E.; Hum, C.; Ahmed, N.; Paul Pezacki, J. Monitoring SARS-CoV-2 Nsp13 Helicase Binding Activity Using Expanded Genetic Code Techniques. **2025**. https://doi.org/10.1039/D4CB00230J.
(106) Sapienza, P. J.; Currie, M. M.; Lancaster, N. M.; Li, K.; Aubé, J.; Goldfarb, D.; Cloer, E. W.; Major, M. B.; Lee, A. L. Visualizing an Allosteric Intermediate Using CuAAC Stabilization of an NMR Mixed Labeled Dimer. *ACS Chem. Biol.* **2021**, *16* (12), 2766–2775. https://doi.org/10.1021/acschembio.1c00617.
(107) Ma, X.; Wei, B.; Wang, E. Efficient Incorporation of P-Azido-l-Phenylalanine into the Protein Using Organic Solvents. *Protein Expr. Purif.* **2022**, *200*, 106158. https://doi.org/10.1016/j.pep.2022.106158.

(108) Kettal, M. Deciphering TG2-Fibronectin Interactions via Azide-Based Site-Specific Mutagenesis. Honours Thesis, University of Ottawa, 2025.
(109) Schultz, K. C.; Supekova, L.; Ryu, Y.; Xie, J.; Perera, R.; Schultz, P. G. A Genetically Encoded Infrared Probe. *J. Am. Chem. Soc.* **2006**, *128* (43), 13984–13985. https://doi.org/10.1021/ja0636690.
(110) Miyake-Stoner, S. J.; Miller, A. M.; Hammill, J. T.; Peeler, J. C.; Hess, K. R.; Mehl, R. A.; Brewer, S. H. Probing Protein Folding Using Site-Specifically Encoded Unnatural Amino Acids as FRET Donors with Tryptophan. *Biochemistry* **2009**, *48* (25), 5953–5962. https://doi.org/10.1021/bi900426d.
(111) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of P-Azido-L-Phenylalanine to the Genetic Code of Escherichia Coli. *J. Am. Chem. Soc.* **2002**, *124* (31), 9026–9027. https://doi.org/10.1021/ja027007w.
(112) Fladischer, P.; Weingartner, A.; Blamauer, J.; Darnhofer, B.; Birner-Gruenberger, R.; Kardashliev, T.; Ruff, A. J.; Schwaneberg, U.; Wiltschi, B. A Semi-Rationally Engineered Bacterial Pyrrolysyl-tRNA Synthetase Genetically Encodes Phenyl Azide Chemistry. *Biotechnol. J.* **2019**, *14* (3), 1800125. https://doi.org/10.1002/biot.201800125.