DESIGNING CONFORMATIONAL CONTROL OF HUMAN TISSUE TRANSGLUTAMINASE

Thesis by
Alexandria H. T. Berry

In Partial Fulfillment of the Requirements for the degree
of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California
2014
(Defended February 20th, 2014)

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Stephen Mayo, for accepting me into his lab and providing me with the opportunity and support to accomplish the research described here. Steve provided excellent scientific advice and insight, and always came through when needed most in all aspects of my graduate student career.

I especially would like to thank past and current members of the Mayo lab who educated me in the use of the equipment in the lab and provided constant scientific and emotional support throughout, especially Toni Lee, Matthew Moore, Alex Nishtal, Jennifer Keefe, Roberto Chica, Bernardo Sosa Padilla Arujo, Tim Wannier, Emzo de los Santos, Kurt Mou, Gene Kym, Seth Lieblich, Jan Kostecki, Samy Hamdouche, Jackson Cahn, and Mohsen Chitsaz. Extra special thanks are owed to Rhonda DiGiusto for being a friend, advocate, and extremely supportive individual professionally and personally. Marie Ary is also owed a big thank you for all her help over the years, especially as an editor, and thanks are also owed to Lilian Porter. I would also like to thank my undergraduate advisor, Nigel Richards, who made it possible for me to attend Caltech and was the inspiration for my desire to pursue this course.

I would also like to thank many other Caltech/Protabit individuals outside the Mayo lab for their support and teaching over the years, especially Paul Patterson, Pamela Bjorkman, David Tirrell, Bil Clemons, Rob Phillips, Paul Chang, Heidi Privett, Alison Ross, David Akopian, Inderjit Nangiana, Robert Graham, Maja Bialecka-Fornal, and especially Zakary Singer, without whom it would have been impossible for me to complete this thesis and who has been my best friend these past five years.

I also need to thank my family, especially my brother for always being my buddy and having my back, my mother for sacrificing herself constantly so she could support me and my goals my entire life, and my father, who I always strive to make proud. I thank God for the grace and strength to be here and do this work, and for getting me through. JMJ.

ABSTRACT

Understanding the mechanisms of enzymes is crucial for our understanding of their role in biology and for designing methods to perturb or harness their activities for medical treatments, industrial processes, or biological engineering. One aspect of enzymes that makes them difficult to fully understand is that they are in constant motion, and these motions and the conformations adopted throughout these transitions often play a role in their function.

Traditionally, it has been difficult to isolate a protein in a particular conformation to determine what role each form plays in the reaction or biology of that enzyme. A new technology, computational protein design, makes the isolation of various conformations possible, and therefore is an extremely powerful tool in enabling a fuller understanding of the role a protein conformation plays in various biological processes.

One such protein that undergoes large structural shifts during different activities is human type II transglutaminase (TG2). TG2 is an enzyme that exists in two dramatically different conformational states: (1) an open, extended form, which is adopted upon the binding of calcium, and (2) a closed, compact form, which is adopted upon the binding of GTP or GDP. TG2 possess two separate active sites, each with a radically different activity. This open, calcium-bound form of TG2 is believed to act as a transglutaminse, where it catalyzes the formation of an isopeptide bond between the sidechain of a peptide-bound glutamine and a primary amine. The closed, GTP-bound conformation is believed to act as a GTPase. TG2 is also implicated in a variety of biological and pathological processes.

To better understand the effects of TG2's conformations on its activities and pathological processes, we set out to design variants of TG2 isolated in either the closed or open conformations. We were able to design open-locked and closed-biased TG2 variants, and use these designs to unseat the current understanding of the activities and their concurrent conformations of TG2 and explore each conformation's role in celiac disease models. This work also enabled us to help explain older confusing results in regards to this enzyme and its activities. The new model for TG2 activity has immense implications for our understanding of its functional capabilities in various environments, and for our ability to understand which conformations need to be inhibited in the design of new drugs for diseases in which TG2's activities are believed to elicit pathological effects.

TABLE OF CONTENTS

Acknowledgements		iii
Abstract		iv
Table of Contents		v
Tables and Figures		vi
Abbreviations		x
Chapter 1	*Introduction*	1
Chapter 2	*Stabilization of tissue transglutaminase in open or closed conformations using computational multi-state protein design*	32
Chapter 3	*The addition of TG2 designs with distinct conformations and activities elicit differential expression of celiac-related genes in cell culture*	83
Appendix A	*Site-saturation mutagenesis of the active site of a designed Kemp elimination enzyme*	127
Appendix B	*Exploring cyanovirin binding to Ebola Glycoprotein 1,2 to facilitate future antiviral designs*	151

TABLES AND FIGURES

Figure 1-1. Human tissue transglutaminase consists of four domains 13

Figure 1-2. TG2 adopts two conformational states 14

Figure 1-3. A close up of the GTPase active site of TG2 in the closed conformation 15

Figure 1-4. The transglutaminase reaction of TG2 16

Figure 1-5. The transglutaminase active site of TG2 in the open conformation 17

Figure 1-6. In the open conformation, the GTPase active site is torn in half 18

Figure 1-7. The hinge region of TG2 in the closed and open conformations 19

Figure 1-8. Model for the activities and conformations of TG2 20

Figure 2-1. Design regions of TG2 54

Figure 2-2. Irreversible transglutaminase inhibitor 55

Figure 2-3. Open-locked hinge design residue Y315 56

Figure 2-4. TRIAD predicted structure of Y315A 56

Figure 2-5. TRIAD predicted structure of Y315E 57

Figure 2-6. TRIAD predicted structure of Y315E_L508T 57

Figure 2-7. TRIAD predicted structure of Y315E_L508W 58

Figure 2-8. Open-locked WT design residues of the β2-core interface 59

Figure 2-9. TRIAD predicted structure of I178H_N181H 60

Figure 2-10. TRIAD predicted structure of I178H_N181H_S253K_I255H 60

Figure 2-11. Native gels of open-locked designs 61

Figure 2-12. Native gel of C277A_Y315E 61

Figure 2-13. Closed-locked WT design residues in closed TG2 62

Figure 2-14. Closed-locked WT design residues in open TG2 62

Figure 2-15. TGase active site residues with closed design residues in closed 63

Figure 2-16. TGase active site residues with closed design residues in open 63

Figure 2-17. Native gels of closed-biased designs 64

Figure 2-18. TRIAD predicted structure of L312Y 65

Figure 2-19. TRIAD predicted structure of L312Y_F316T 65

Figure 2-20. Transglutaminase activity assay of open-locked and WT TG2 66

Figure 2-21. Transglutaminase activity assay of open-locked Y315E without calcium 67

Figure 2-22. Kinetic data from transglutaminase assay for all designs and WT 68

Table 2.1. Kinetic data from transglutaminase assay for all designs and WT 69

Figure 2-23. Transglutaminase active site residues in open and closed TG2 70

Figure 2-24. Overlay of the catalytic triad in open and closed TG2 70

Figure 2-25. Overlay of the catalytic triad of closed TG2 and TG3 71

Figure 2-26. Inhibition of transglutaminase activity by GTP as a % of variants activity 72

Figure 2-27. Inhibition of transglutaminase activity by GTP for WT and L312Y 73

Figure 2-28. Inhibition of transglutaminase activity by GTP for WT and Y315E_L508W 74

Figure 2-29. GTPase assay of WT, L312Y, L312Y_F316T, and Y315E_L508W TG2 75

Table 2-2. GTPase assay of WT, L312Y, L312Y_F316T, and Y315E_L508W TG2 75

Table 3-1. List of primers for qPCR analysis 100

Table 3-2. Normalized expression of NFKB1A 101

Figure 3-1. Normalized expression of NFKB1A in experiment 1 102

Table 3-3. Statistically significant differences in NFKB1A, experiment 1, 3 hours 103

Table 3-4. Statistically significant differences in NFKB1A, experiment 1, 12 hours 103

Figure 3-2. Normalized expression of NFKB1A in experiment 2 104

Table 3-5. Statistically significant differences in NFKB1A, experiment 2, 3 hours 105

Table 3-6. Statistically significant differences in NFKB1A, experiment 2, 12 hours 105

Table 3-7. Statistically significant differences in NFKB1A, experiment 2, 24 hours 106

Table 3-8. Normalized expression of SELE 107

Figure 3-3. Normalized expression of SELE in experiment 1 108

Table 3-9. Statistically significant differences in SELE, experiment 1, 3 hours 109

Table 3-10. Statistically significant differences in SELE, experiment 1, 12 hours 190

Figure 3-4. Normalized expression of SELE in experiment 2 110

Table 3-11. Statistically significant differences in SELE, experiment 2, 3 hours 111

Table 3-12. Statistically significant differences in SELE, experiment 2, 12 hours 111

Table 3-13. Statistically significant differences in SELE, experiment 2, 24 hours 112

Table 3-14. Normalized expression of RHOB 113

Figure 3-5. Normalized expression of RHOB in experiment 1 114

Figure 3-6. Normalized expression of RHOB in experiment 2 115

Table 3-15. Normalized expression of ITGB1 116

Figure 3-7. Normalized expression of ITGB1 in experiment 1 117

Figure 3-8. Normalized expression of ITGB1 in experiment 2 118

Table 3-16. Normalized expression of FN1 119

Figure 3-9. Normalized expression of FN1 in experiment 1 120

Figure 3-10. Normalized expression of FN1 in experiment 2 121

Figure 3-11. Normalized relative expression of SELE after 3 hours, experiment 2, 6 pairs 122

Table 3-17. Statistically significant differences in SELE after 3 hours, experiment 2, 6 pairs 122

Figure A-1. The Kemp elimination of 5-nitrobenzisoxazole 141

Figure A-2. Design structure of HG2 142

Figure A-3a. HG2 active site residues for site saturation mutagenesis 143

Figure A-3b. HG2 active site residues for site saturation mutagenesis, cont. 144

Table A-1. Summary of site saturation mutagenesis results 145

Figure A-4. Activity screen of active site saturation mutants for HG2/S265T 146

Table A-2. Kinetic data for the HG2/S265T mutants more active than "WT" in screen 147

Figure A-5. Calculated PHOENIX score vs. k_{cat}/K_m for HG2/S265T active site mutants 148

Figure B-1. WT CVN monomer bound to mannose moieties 163

Figure B-2. Structure of Ebola glycoprotein GP1,2 164

Figure B-3. Gene construct and domains of Ebola Zaire GP1,2 165

Figure B-4. Gene assembly protocol scheme 166

Figure B-5. Gene assembly of GP1,2 167

Figure B-6. GP1,2 purified from *Sf9* cells 168

Figure B-7. SDS-PAGE gels of Ebola Zaire GP1,2 under reduced and unreduced conditions 169

Figure B-8. Liquid chromatography and TOF-MS of unreduced GP1,2, species 1 170

Figure B-9. Liquid chromatography and TOF-MS of unreduced GP1,2, species 2 171

Figure B-10. Liquid chromatography and TOF-MS of reduced GP1,2, species 1 172

Figure B-11. Liquid chromatography and TOF-MS of reduced GP1,2, species 2 173

Figure B-12. Liquid chromatography and TOF-MS of reduced GP1,2, species 3 174

Figure B-13. CD wavelength scan of GP1,2 175

Figure B-14. GP1,2 colored by secondary structure 176

ABBREVIATIONS

AIDS	acquired immunodeficiency syndrome
APC	antigen presenting cell
CD	Celiac disease
cDNA	coding Deoxyribonucleic acid
CPD	computational protein design
CVN	Cyanovirin-N
Da	Daltons
DNA	Deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DSC	dansylcadaverine
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EBM-2	Endothelial Cell Basal Medium-2
EBOV	Ebola virus
ECM	extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGM-2	Endothelial Growth Media-2
ELISA	enzyme-linked immuneobsorbent assays
FN1	Fibronectin 1
FXIII	Factor XIII
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GDP	guanosine diphosphate
GGEL	γ-glutamyl-ε-lysine[Nε-(γ-L-glutamyl)-L-lysine]
Gln	Glutamine
GP	Ebola glycoprotein gene
GP1	Ebola glycoprotein subunit 1
GP1,2	Ebola glycolprotein 1,2 with deleted mucin domain
GP2	Ebola glycoprotein subunit 2
GTP	guanosine triphosphate
GTP*	γ-^{32}P-GTP
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	human immunodeficiency virus
HUVECs	Human umbilical vein endothelial cells
IgA	Immunoglobulin A
ITGB1	Integrin, Beta 1
IκBα	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha
$K_{1/2}$	the protein concentration that provides half the maximal rate

kb	kilobases
k_{cat}	catalytic constant or turnover number
kDa	kilodaltons
K_m	Michaelis constant
k_{max}	the maximal rate constant with saturating protein
k_{obsd}	observed rate constant
LB	Luria-Bertani broth
LC-MS	liquid chromatography-mass spectrometry
MD	molecular dynamics
MHC	major histocompatibility complex
NBX	transition state of the 5-nitrobenzoxizole Kemp elimination reaction
NBZ	5-nitrobenzoxizole
NF-κB	Nuclear factor κB
NFKBIA	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha
NMC	N,N-dimethylated casein
nt	nucleotide
PAGE	polyacrylamine gel electrophoresis
PCR	polymerase chain reaction
Pro	Proline
qPCR	quantitative polymerase chain reaction
RHOB	Ras homolog family member B
RNA	Ribonucleic acid
RNAi	RNA interference
s	seconds
SARS	severe acute respiratory syndrome
SDHA	Succinate dehydrogenase complex, subunit A
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SELE	Selectin E
Sf9	*Spodoptera frugiperda* insect cells
sGP	Ebola secreted glycoprotein
TAX	10A xylanase
TBP	TATA box binding protein
TG1	transglutaminase type 1
TG2	human tissue transglutaminase or type II transglutaminase
TG3	transglutaminase type 3
TG4	transglutaminase type 4
TG5	transglutaminase type 5
TG6	transglutaminase type 6
TG7	transglutaminase type 7
TNF-α	Tumor necrosis factor α

U	units
U	units
UV	ultraviolet
V	Volts
V_{max}	maximal rate of reaction
WT	wild type
βME	β-mercaptoethanol

C h a p t e r 1

INTRODUCTION

Proteins are dynamic

One of the major strengths of protein crystallography is that it allows us to view protein structures with atomic detail, and from this structural information deduce functional relevance and biological mechanisms. However, these static images can give the viewer the wrong impression: proteins are in constant motion, sampling an array of conformations, with some movements miniscule, others dramatic. These small conformational vibrations or radical reorganizations can be critical for biological function or sometimes enormously detrimental. For example, hemoglobin, responsible for delivering oxygen to tissues, undergoes a conformational change upon the binding of oxygen to the first of its four subunits that translates to an increase in oxygen affinity in the remaining subunits [1]. Conversely, misfolded proteins called prions are infectious agents that can cause disease by acting as a conformational template for existing proteins, effecting a cascade where they induce the normally folded proteins to adopt the prion's misfolded and disease-associated state [2].

Therefore, a full understanding of the dynamics of a protein is essential for teasing out its mechanism of action, and if necessary, how to perturb any unwanted behavior.

Human tissue transglutaminase undergoes a very large conformational change

One such enzyme whose biological activity is intimately intertwined with its conformational inclinations is human tissue transglutaminase, also known as human type II transglutaminase (TG2). TG2 consists of four domains: an N-terminal β-sandwich, a core, and two β-barrel domains (Figure 1-1). TG2 exists in two distinct conformations, closed and open, and each conformation was believed to catalyze a separate reaction; the closed conformation was believed to act as a GTPase, while the open was believed to act as the transglutaminase [3-5].

In the closed conformation, the transglutaminase active site is sequestered in a grove between the core and the interface of the β-barrel domains, which was believed to prevent its activity in the

closed conformation (Figures 1-2, 1-3). The GTPase site is formed on the opposite face of the protein at the core-β1-barrel interface [5]. TG2 assumes the closed conformation upon GTP binding, and therefore this is the conformation that was assumed to be adopted during the GTPase reaction. The closed conformation is also the preferred conformation.

In the extended, open conformation, which occurs upon calcium binding and exposes the transglutaminase active site residues at the core just above the β1-barrel, TG2 is believed to catalyze an acyltransferase reaction relatively non-specifically. This results in the formation of an amide bond between a peptide-bound γ-carboxamide of glutamine and the ε-amino group of lysine or another primary amine (Figure 1-4) [6]. The first step of this reaction occurs when a thiol residue (C227 - part of a stereotypical catalytic triad) attacks a gluamine's carboxamide, releasing ammonia and forming a covalently-linked intermediate (Figure 1-5). A cross link is formed when a lysine sidechain or another primary amine attacks this intermediate. The resulting γ-glutamyl-ε-lysine[N^{ε}-(γ-L-glutamyl)-L-lysine] (GGEL) isopeptide linkage can occur between glutamine and lysine residues in different peptides, stabilizing these polymers. Alternatively, TG2 can proceed only halfway through the reaction, deamidating a glutamine residue to form a glutamate (Figure 1-4) [6]. In the open conformation, the GTPase active site is torn in half as the residues making up this active site reside on both the core and the β1-barrel. Therefore, upon assuming the open conformation, TG2 was believed to be incapable of acting as a GTPase (Figure 1-6) [3].

The transition from closed to open predominately involves the movement of random coils. Therefore, most of TG2's secondary structural elements remain in their original conformations in both the open and closed structures, with one notable exception: residues 309 through 319. These residues make up the hinge of TG2 and undergo a dramatic secondary structural change upon the shift from open to closed. They form a β-strand in the closed conformation and an α-helix in the open conformation (Figure 1-7) [3, 5].

TG2 is regulated by a number of effectors

TG2 is present in most mammalian tissues, and exists both intracellularly and extracellularly [7, 8]. Therefore, it is essential that its activities be tightly regulated, especially given the promiscuity of the transglutaminase activity (see Figure 1-8) [9, 10].

TG2's GTPase activity is inhibited by Ca^{2+} [11] with an estimated IC_{50} ranging from 300 nM to 6 μM [12, 13], and GTP inhibits the transglutaminase activity with a dissociation constant of 1.6 μM [14]. TG2 binds six Ca^{2+} ions with an apparent overall dissociation constant of 90 μM [11, 15]. The GTPase reaction has a K_m of approximately 130 μM [4, 16]. The concentration of Ca^{2+} required for *in vivo* transglutaminase activity, however, is less well defined. In permeabilized cells, 10-100 μM Ca^{2+} is required for measurable TG2 activity.

These results led to the belief that TG2 is a latent enzyme with respect to transglutaminase activity under normal intracellular conditions [17], where Ca^{2+} concentrations are typically below 1μM and GTP is readily available [9]. However, later studies demonstrated that TG2 could be activated *in situ* by physiological increases in calcium concentration to approximately the nanomolar range, suggesting that intracellular factors may increase TG2's calcium sensitivity [18].

Extracellularly, where the concentration of Ca^{2+} ions is high and GTP is low, a more amenable environment exists for the transglutaminase activity of TG2. However, TG2 remains largely inactive as a transglutaminase extracellularly under normal conditions due to inactivation via oxidation [19]. Due to the high redox potential of the extracellular matrix, a reversible disulfide bridge forms between C371 and C370, which alters the backbone of TG2 and induces the adoption of the open conformation, which, surprisingly, inactivates TG2 by a mechanism that is not yet clear [20, 21].

The mechanism by which the extracellular transglutaminase activity of TG2 is activated is still under investigation, but known activators include the protein cofactor thioredoxin-1 [21], interferon-γ [21, 22], and cellular stresses like tissue injury [19, 23].

How are these activities and conformations biologically relevant?

TG2 is found ubiquitously in mammalian tissues (hence its name), including the liver, brain, intestine, and erythrocytes [24, 25]. TG2 is located both intracellularly and extracellularly [8]. Though primarily in the cytosol [26], TG2 can be found in the extracellular matrix [27], associated with the plasma membrane [4, 28], in the nucleus [29-31], and in mitochondria [32, 33].

TG2's diverse locations and activates have led it to be implicated in a wide variety of physiological processes. For example, extracellular TG2's transglutmainase activity plays a role in wound healing [19, 34], while intracellular TG2 acting as a GTPase has been shown to be the functional α-subunit of a G-protein involved with α_1 adrenergic receptor signaling [4], as well as with other receptor types [35, 36]. However, the biological importance of TG2 as a G-protein is still being determined [37, 38]. TG2 has also been shown to both promote apoptosis and promote cell survival, depending on the circumstances. TG2 is involved in cell adhesion [39, 40], matrix stabilization [41-43], cell proliferation and migration [10, 44], neuronal growth and regeneration [45-47], and as a transcriptional regulator of a myriad of processes involving stress induced nuclear localization [48-57].

Tissue transglutaminase is a unique member of a large protein family

The transglutaminase family consists of 9 members: TG1 through TG7, factor XIII, and band 4.2 [58, 59]. All but band 4.2 can act as transglutaminases, and these all require Ca^{2+} binding for this activity [58]. As mentioned above, TG2 is the most widely expressed of all the transglutaminases, and is found in almost every tissue of the body [9, 10]. The other transglutaminases are more specifically expressed, for instance, TG1, TG3, and TG5 are expressed in keratinocytes and/or hair follicles and are involved in differentiation [60-66]. TG4 is localized in the prostate gland and its role in human biology is not yet totally understood [67]. Factor XIII is found in plasma, platlets, monocytes, and macrophages and is involved in functions such as fibrin stabilization, wound healing, and angiogenesis [68-76].

TG1 through TG7 and band 4.2 exist as monomers [66, 68, 77]. In the plasma, factor XIII exists as a tetramer with two potentially active FXIII-A subunits and two inhibitory/carrier FXIII-B subunits (FXIII-A_2B_2) [68]. In the cytoplasm of platelets and monocytes, factor XIII is a dimer (FXIII-A_2) [68].

Crystal structures have been determined for only three of these transglutaminases: factor XIII-A_2, TG2, and TG3 [3, 5, 78-80]. TG2 is the only transglutaminase confirmed to act as a GTPase and the only transglutaminase crystallographically shown to adopt the elongated open conformation. As mentioned above, all three of these transglutaminases require Ca^{2+} binding for activation [58]. Although TG2 adopts an open conformation upon Ca^{2+} binding and has been shown to bind six Ca^{2+}

ions, no Ca^{2+} was found bound to the open crystal structure, which was achieved by covalent inhibition by a peptide substrate in the transglutaminase active site [3, 11].

Factor XIII is activated by thrombin and by the binding of a single Ca^{2+} ion [81, 82]. Thrombin cleaves off an N-terminal peptide by hydrolyzing the R37-G38 peptide bond that, in addition to the binding of calcium, causes the inhibitory FXIII-B subunits to dissociate from the complex, resulting in the active dimer FXIII-A_2* [83-86]. The structure adopted by the active form of factor XIII, FXIII-A_2*, is still being debated. Hydrogen deuterium exchange experiments and chemical modification of side chains studies have suggested conformational changes occur when the enzyme transitions from inactive FXIII-A_2 to active FXIII-A_2* [87-89]. However, the crystal structure of thrombin activated and Ca^{2+}-bound FXIII-A_2 still adopted the closed conformation and did not appear significantly different from the non-activated form of FXIII-A_2 [90]. This implies that any conformational changes in FXIII-A_2 may be fleeting and unstable.

In order to be activated, TG3 binds three Ca^{2+} ions [80] and undergoes a cleavage event in the region equivalent to the hinge region in TG2, resulting in two globular domains that remain associated with each other in the closed conformation: a 50 kDa amino terminus with the catalytic site and a 27 kDa carboxy terminus [66, 91]. Crystal structures of TG3 in its active conformation bound to 3 Ca^{2+} ions have been solved [80, 92], as well as structures for the inactive form of the enzyme [80]. Surprisingly, the active and calcium-bound form of TG3 remains in the closed conformation, but activation causes a loop consisting of residues 320 to 325 to move and open a cavity through the enzyme towards the active site residues, although further rearrangements are believed to be necessary to accommodate substrates [93]. TG3 is not believed to adopt an open, elongated conformation similar to that of TG2.

TG2 is the most nonspecific recognizer of glutamine donor substrates of all transglutaminases

In addition to being the only transglutaminase definitively shown to adopt an open, elongated conformation, TG2 is also the least discriminating in its choices of transglutaminase substrates. TRANSDAB is a database of all known transglutaminase substrate proteins and interaction partners [94]. TG2 has by far the largest number of known substrate proteins, with the current count being 155, and has long been known to be extremely nonspecific [9, 10]. Factor XIII-A is much more specific in its recognition of glutamine residues than TG2 [94] and has a much smaller amount of

known substrates at 28, although the two glutamine donors fibrin and α_2-PI are significantly better substrates than the others listed [68, 95, 96]. TG3 acts on 16 known substrates and is considered highly specific [92]. The other transglutaminases are also extrememly specific: TG1 has 14 known substrates, TG4 has 3, and TG5 has 4.

It is interesting to note that among TG2, factor XIII, and TG3, substrate promiscuity increases with observed increasing conformational flexibility upon the binding of Ca^{2+}. TG2 is the most promiscuous and adopts a fully elongated conformation upon calcium binding. Factor XIII has been shown to undergo conformational changes, although not observed in the crystal structure bound to calcium, and has much stricter substrate preferences. TG3, which has been shown to undergo a flexing to open a channel to the active site and remain in the closed conformation when bound to calcium, exhibits the highest specificity of the three.

Truncation studies of TG2 have produced confusing results as to which domains are necessary and sufficient to preform TG2's activities

As mentioned above and as shown in Figure 1-3, the crystal structure of TG2 bound to GDP revealed that the GTPase active site comprises residues from the core domain and the β-barrel 1 domain [5]. This result is surprising considering data obtained from two deletion studies. A domain-deletion study found that the core domain alone exhibited a 50% increase in GTPase activity when compared to the full-length protein, while a variant with the N-terminal β-sandwich and the core, but lacking both β-barrel domains, showed no significant difference in activity relative to full-length TG2 [97]. A second study using successive C-terminal deletion analysis to attempt to pinpoint the GTP binding site found that deleting off the two β-barrel domains resulted in a 34-fold increase in GTPase activity relative to WT [98]. All of these truncated variants of TG2 lacked all but two (K173 and F174) of the GTPase active site residues later identified by X-ray crystallography (Figure 1-3).

The same two papers investigated these deletions' effects on the transglutaminase activity. The domain-deletion study found that the loss of the two β-barrel domains caused a 30% loss in activity relative to full-length TG2, while the core domain alone lacked any transglutaminase activity [97]. The successive C-terminal deletion analysis found that the first variant with the deletion of β-barrel 1 and half of β-barrel 2 and a second variant with the deletion of both β-barrels resulted in a 95-99%

loss in transglutaminase activity relative to full-length protein [98]. Again, this is a surprising result considering that all the active site residues involved in catalysis and substrate and calcium binding identified biochemically and in the open conformation crystal structure reside in the core and were neither deleted in any of these truncation variants, nor do they directly interact with any of those residues deleted [3, 11].

The transglutaminase activity of TG2 plays a critical role in celiac disease

Celiac disease (CD) is an autoimmune condition affecting the small intestine, and is triggered in genetically predisposed individuals by the ingestion of gluten – proteins present in wheat and other grains. In CD, antibodies attack the small intestinal lining, causing the formation of lesions characterized by increased epithelial cell proliferation and decreased differentiation, as well as the infiltration of T cells into the mucosa [99, 100]. These cellular events lead to the classic symptoms of villious atrophy and crypt hyperplasia, which cause diarrhea and impair the patient's ability to absorb nutrients [101, 102]. CD can also present extraintestinally, in symptoms including osteoporosis, dermatitis herpetiformis, and neurological disorders, such as gluten ataxia [103].

Gluten is an insoluble heterogeneous mixture of storage proteins known as gliadins and glutenins, both of which contain repetitive sequences enriched in proline (Pro; 15%) and glutamine (Gln; 35%). These residues render the gluten resistant to proteases, and the resulting 30-40 residue-length peptides remain intact when they are released into the gut lumen [104, 105].

At this point in celiac disease, tissue transglutaminase plays its first known role. Intestinal TG2 deamidates the gluten peptides at specific Gln residues, increasing their affinity for the mutant histocompatibility class II molecule HLA-DQ2 (present in 90% of patients), or HLA-DQ8 (10% of patients), on the surface of antigen-presenting cells (APCs) [106, 107]. The peptide is then presented to a DQ2/8-restricted $CD4^+$ T cell, which elicits the T cell's proliferation and the deleterious immune activation characterizing CD [108-110].

CD is characterized by anti-TG2 antibodies

TG2 itself is the primary autoanitgen in celiac disease [111]. Anti-TG2 antibodies produced by celiac patients are highly specific and detection of their presence is the primary means of diagnosis [111-114]. Upon the ingestion of gluten in an afflicted individual, anti-TG2 antibodies are secreted in the small intestinal mucosa [115], accumulating immediately below the small-bowel epithelial membrane around mucosal blood vessels [116, 117]. Here, the autoantibodies bind extracellular TG2 on fibroblasts on the small intestine's basement membrane [112, 116, 118]. Interestingly, in this same region, the subepithelial layer of the jejunal mucosa, studies have demonstrated an increase in TG2 expression and transglutaminase activity in the celiac patient [119, 120]. On a related note, TG3 has also been identified as the autoantigen of dermatitis herpetiformis, a blistering skin disease associated with celiac disease as mentioned above [65].

CD anti-TG2 autoantibodies affect CD symptoms and TG2 transglutaminase activity

Surprisingly, these anti-TG2 autoantibodies were found not to ameliorate symptoms in celiac patients, but to affect pathogenic symptoms. Celiac autoantibodies have been shown to induce proliferation and inhibit differentiation of epithelial cells [121], possibly contributing to the development of crypt hyperplasia, which is characterized by precisely that. Anti-TG2 celiac autoantibodies have also been shown to induce apoptosis both in neuronal cells and trophoblasts, possibly contributing to the neurological and pregnancy complications observed in CD [122, 123]. Celiac autoantibodies also enhance lymphocyte binding to the endothelium and transendothelial migration, accompanied by an up-regulation in the lymphocyte adhesion molecule E-selectin [124], which is interesting in light of reports that in untreated celiac disease, lymphocyte adhesion molecules including E-selectin are up-regulated in the intestinal mucosal, possibly contributing to lymphocyte migration into the lamina propria from the endothelial barrier [125-127]. Total celiac IgA and anti-TG2 celiac IgA have been reported to inhibit angiogenesis and increase vascular permeability, whereas commercially available anti-TG2 antibody, CUB7402, or non-celiac IgA did not [124, 128-132], possibly contributing to the disorganization reported in the vascular architecture of the CD lesion [133].

In angiogenesis, new blood vessels are formed from old ones, and processes playing a key role, such as matrix stabilization, are carefully regulated [8]. TG2 has been implicated in numerous

studies to play a role in matrix stabilization relevant to angiogenesis, promoting angiogenesis during wound healing [134] and inhibiting it during tumor formation by increasing the accumulation of extracellular matrix proteins [135, 136], presumably because angiogenesis generally requires localized destabilization of the extracellular matrix [137]. In addition, TG2 expression is downregulated in endothelial cells during capillary morphogenesis [138].

Interestingly, it was reported that the anti-angiogenesis effects induced by celiac disease anti-TG2 autoantibodies, such as increased endothelial cell macromolecule permeability and tubule length, could be reversed upon the addition of the TG2 transglutaminase inhibitor R281 [124, 129]. The above findings led to the hypothesis that the anti-TG2 autoantibodies could perhaps be activating the transglutaminase activity of TG2, and that this activation could lead to further disease progression by increasing TG2's deamidation of gluten. Initial studies using liquid phase assays showed that celiac anti-TG2 autoantibodies had a moderate inhibitory effect on transglutaminase activity [139-141]. More recent studies either more closely mimicking *in vivo* conditions [142], or *in situ* with live endothelial cells [124, 129] demonstrated that celiac patient autoantibodies or celiac-derived monoclonal anti-TG2 antibodies actually enhanced transglutaminase activity, while commercially available or non-CD IgA did not. In the first study [142], this activation was lost once the celiac patient removed gluten from their diet. The latter two studies [124, 129] showed that this activation was lost upon the addition of transglutaminase inhibitors.

Celiac anti-TG2 autoantibodies recognize a conformational epitope consisting of residues from three domains

Early attempts to identify epitopes recognized by CD autoantibodies on TG2 consisted of fragmentation studies. One such study found that both the N- and C-terminal domains of TG2 were important for the binding celiac patient antibodies [143]. A second, similar study found that both the N- and C-termini harbored recognized epitopes, but that the core needed to also be present with either terminus in order for binding. The core itself was not found to be sufficient [144]. A third study found that the core and N-terminus were important for recognition [145]. Fragmentation studies are not ideal, however, since the epitopes recognized by anti-TG2 celiac antibodies are

known to be conformationally dependent, indicated by their inability to bind linearized TG2 peptides in phage display [146] or there decreased ability to bind in Western blot or paraffin-embedded tissues [147].

Recently, a group working on determining the Ca^{2+} binding sites on TG2 noticed that one of their calcium binging-site knock-outs destroyed celiac IgA's ability to bind TG2. However, the presence of calcium or GDP had no effect on celiac IgA's ability to bind wild-type TG2. After further investigation, they were able to define a single TG2 epitope recognized by IgA from 18 different celiac patients that resided close to this calcium binding site. This epitope consists of residues from the N- and C-termini and the core, in agreement with the fragmentation studies mentioned above, and is only completely assembled in the closed conformation [148]. This led them to hypothesize that the binding of celiac antibodies to TG2 may stabilize the closed conformation, and by doing so confer the gain in catalytic activity observed in the above-mentioned studies. However, the idea that the closed form of TG2 is the more active is contrary to the long-held view that the open, extended conformation of TG2 was the transglutaminase active form.

Conformational protein design is an excellent tool for studying conformational roles in biology

The multifunctional nature of TG2 complicates our ability to determine its role in cell physiology and pathology. Firstly, it is desirable to confirm or disprove assumptions about TG2's conformations' affects on its activities made based on crystal structures. In addition, it would be useful to isolate both of TG2's activities from each other in order to pinpoint which activities and conformations of TG2 contribute to various biological processes. Unfortunately, traditional methods used to do this are not tractable for TG2. Inhibition is a very common method used to selectively block enzymatic activity, but small molecule inhibitors can have undesirable secondary effects. For example, transglutaminase inhibitors have been known to force the adoption of the open conformation of the enzyme [3, 20]. Active site mutational knock-outs are undesirable in the case of TG2 because (1) they assume our model for TG2 activities and their corresponding conformations is correct, and (2) mutations to the catalytic cysteine of the transglutaminase active site (C277) are also known to cause TG2 to adopt the open conformation [149-151].

Using computational protein design (CPD) to design variants of TG2 that isolate each conformation would be one way to get around many of these challenges. The Mayo lab has already demonstrated

CPD's ability to preferentially stabilize a protein in two distinct conformation states [152]. In this work, CPD was used to stabilize the integrin I domain in either the open or closed conformation. Experimental results showed the open-form variants to be much more active than the WT integrin I, whereas the closed-form variants exhibited no activity, demonstrating that the open conformation is the physiologically relevant conformation for activation by ligand binding.

In CPD, the design algorithm samples amino acid types and conformations at the designated positions in each fold simultaneously. Pairwise interaction energies between amino acid sidechains and between sidechains and backbone are calculated using a molecular mechanics force field with terms for van der Waals interactions, atomic solvation, electrostatics, and hydrogen bonding [153-157]. An overall score for each sequence in the context of a particular fold is calculated from these interaction energies, and the sequence with the minimum energy is found using a sequence optimization algorithm [158-160].

The goal of this thesis was to use CPD to construct variants of TG2 locked into a single conformational state, *i.e.*, stabilized in either the closed form (closed-locked) or the open form (open-locked), and then use these designs to gain insight into the biology of this enzyme.

If the model depicted in Figure 1-8 were correct, a closed-locked TG2 variant would eliminate transglutaminase activity without the use of a small molecule inhibitor that may have unwanted side effects. In addition, a closed-lock design should maintain GTPase activity, isolating it from the transglutaminase activity and the open conformation. A closed-locked design would allow us to test the assumption that the closed conformation is inactive as a transglutaminase and could be used to validate the role of the GTPase reaction alone in physiology and disease.

An open-locked TG2 variant would allow us to test the hypothesis that the open form of the enzyme is the transglutaminase active/GTPase inactive form. If this hypothesis is correct, the open-locked version of the enzyme should be constitutively active with respect to transglutaminase activity and should exhibit no GTPase activity. Again, an open-locked TG2 could then be used to investigate the role of the transglutaminase reaction in biology separate from the GTPase activity. Additionally, in order to achieve the crystal structure for the open form of the enzyme, an inhibitor was necessary [3]. An open-locked form of the enzyme could potentially be used to crystallize the apo form of the enzyme.

Given that (1) other TG2's seem to be active in a more closed form, (2) truncation studies have shown that the β-barrels were unnecessary for GTPase activity and necessary for transglutaminase activity, and (3) celiac anti-TG2 autoantibodies can increase transglutaminase activity and seem to bind a closed form of the enzyme, the real possibility existed that the long-held model for the activities and conformations of TG2 was incorrect. Designing such open and closed-locked TG2s enabled us to investigate this hypothesis.

Chapter 2 of this thesis describes the testing of the existing model for TG2's activities, activation, inhibition, and corresponding conformation by the design and characterization of open-locked and closed-biased variants of TG2. Chapter 3 investigates if conformationally distinct forms of TG2 with differing levels of transglutaminase activity can elicit differing effects in cell culture, and therefore attempts to further investigate the hypothesis that the stabilization of the closed conformation and the resulting increase in transglutmainase activity of TG2 led to pathogenic effects observed in CD. Appendix A details the active site saturation mutagenesis of a designed Kemp elimination enzyme from work previously done in our group in an attempt to improve its activity and to further optimize our design parameters. We set out to accomplish this so that the predictions made by our computational methods better match that which is observed *in vitro* and therefore improve our model, furthering our long term goal of designing *de novo* enzymes with k_{cat}/K_m values closer to those found in nature. Appendix B reports initial work of the expression of Ebola Glycoprotein 1,2 for use in antiviral design.

Tables and Figures

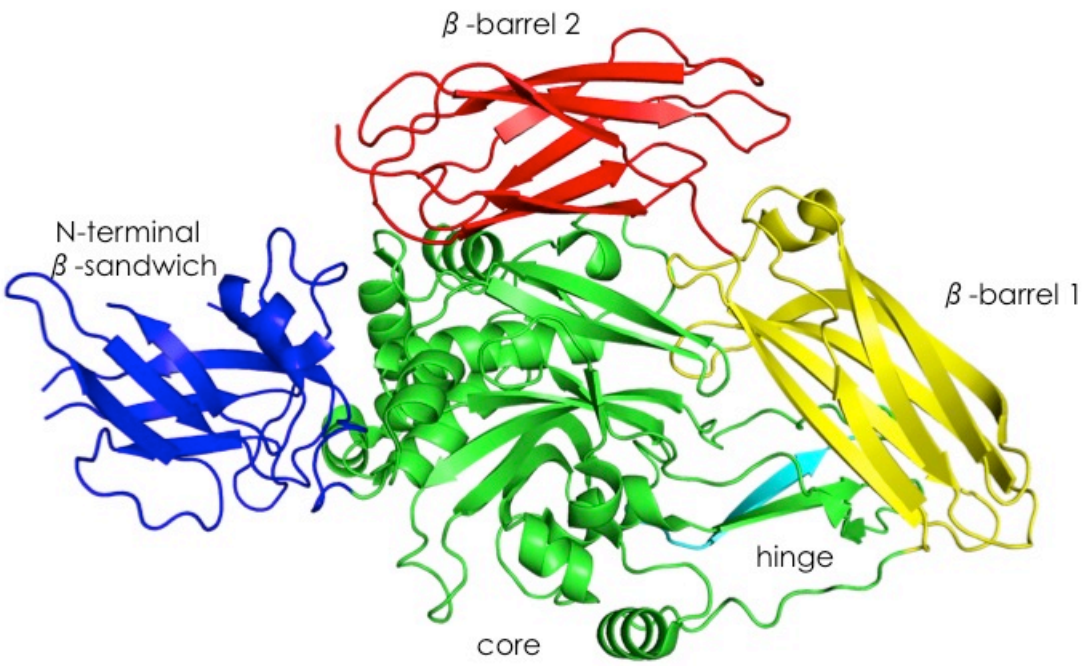


Figure 1-1. Human tissue transglutaminase consists of four domains: (1) the N-terminal β-sandwich (blue), (2) the core (green), (3) β-barrel 1 (yellow), (4) and β-barrel 2 [32]. The hinge region is shown in cyan [5].

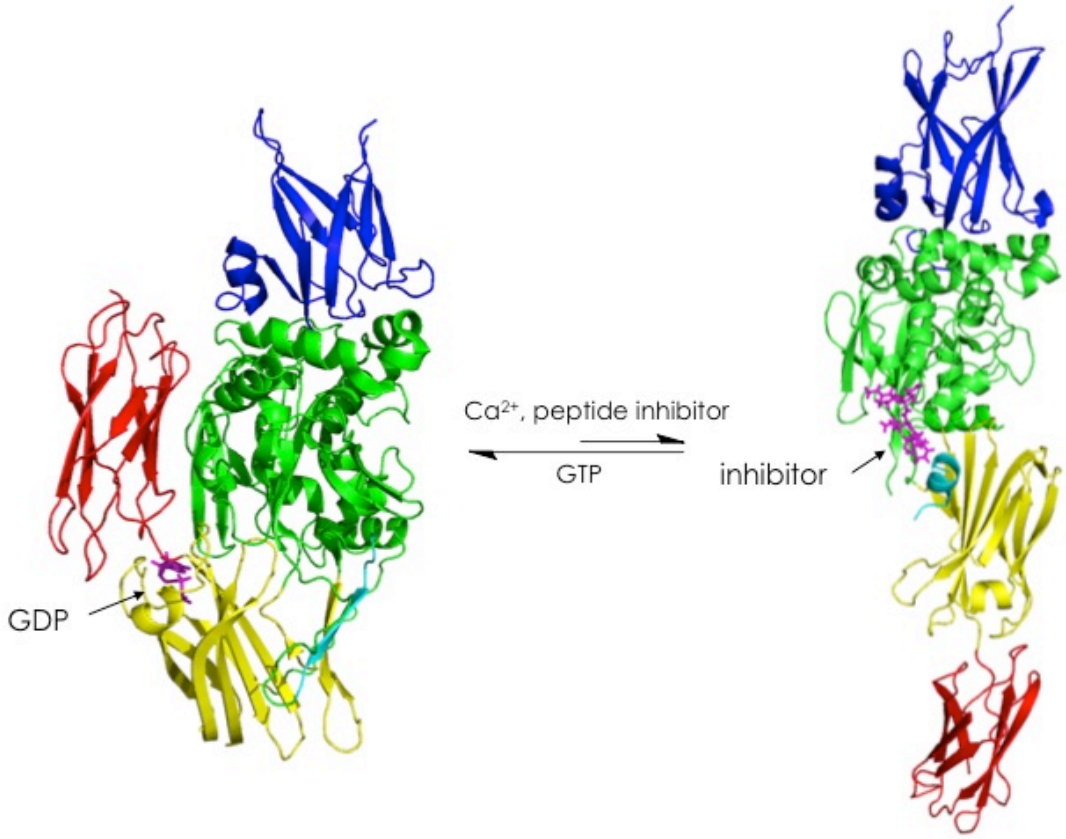


Figure 1-2. TG2 adopts two conformational states. The closed state of TG2 (left) occurs upon GTP binding. The open conformation (right) exposed the transglutaminase active site and occurs upon calcium binding. The closed structure shows GDP bound to the GTPase active site [5] and the open structure shows the covalently-bound peptide inhibitor Ac-P(DON)LPF-NH_2 bound in the transglutaminase site [3].

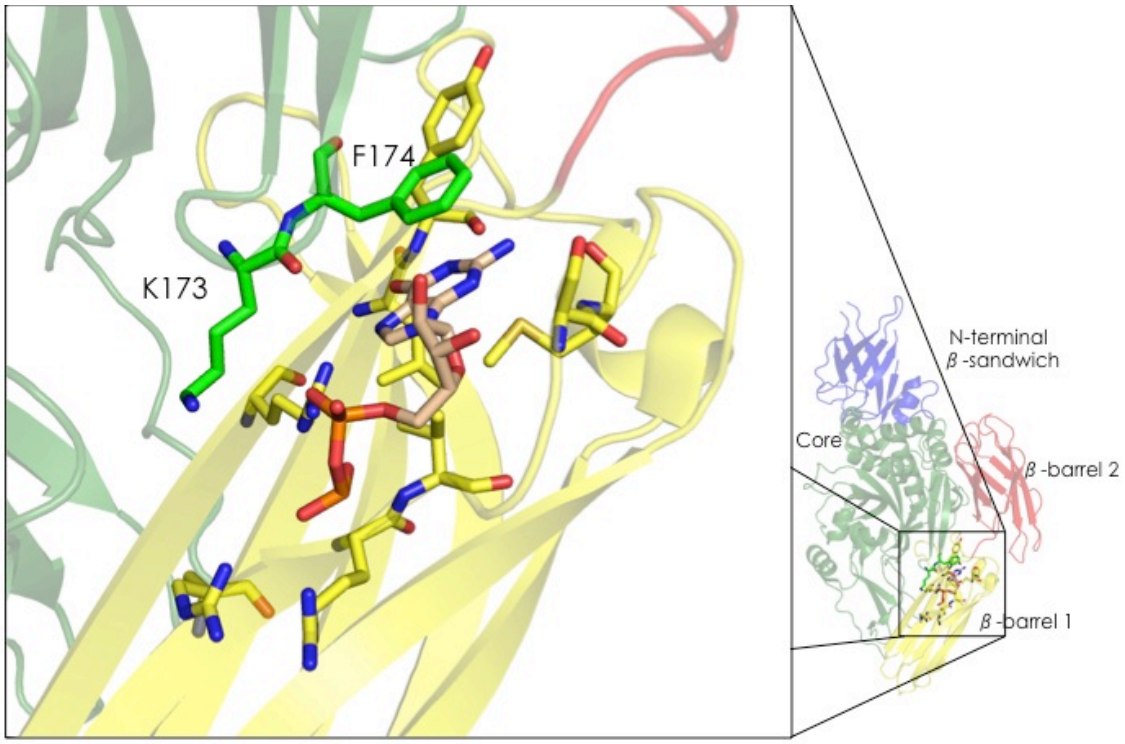


Figure 1-3. A close up of the GTPase active site of TG2 in the closed conformation, with active site residues pictured in green when from the core domain and yellow when from β-barrel 1. GDP s pictured in salmon [5].

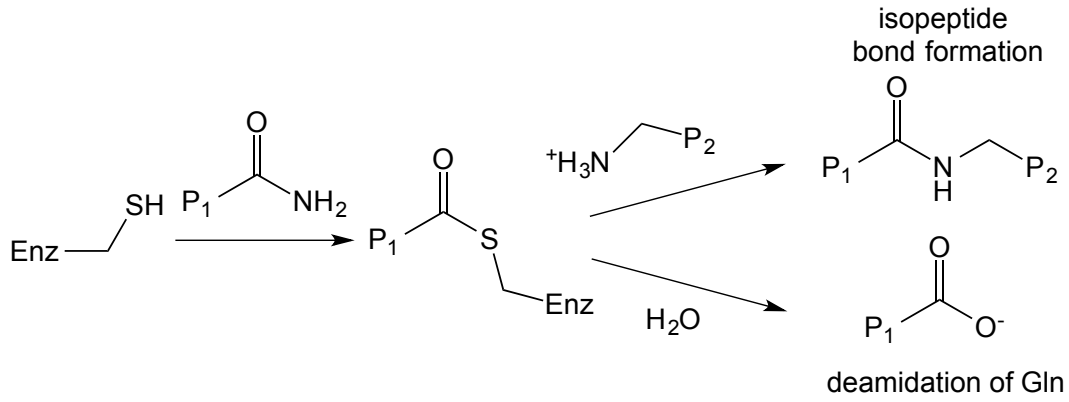


Figure 1-4. TG2 catalyzes the Ca^{2+}-dependent formation of an amide bond between the γ-carboxamide of a glutamine residue on a protein and the ε-amino group of a lysine or another primary amine. Alternatively, covalent substrate-enzyme intermediate can be hydrolyzed by water in a deamidation reaction to convert the glutamine to a glutamate. This reaction is of particular interest in celiac disease, in that when TG2 deamidates gluten, it increases gluten's affinity for the mutant histocompatibility class II molecules, resulting in the immune response characteristic of celiac disease [106, 107].

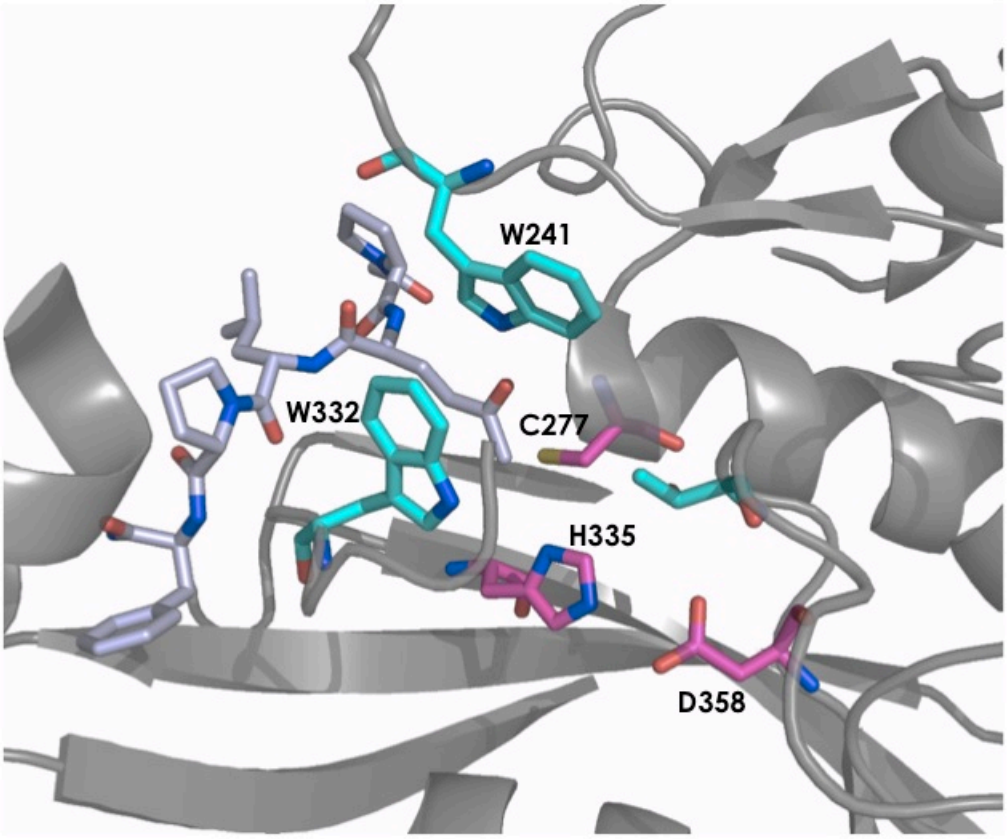


Figure 1-5. The transglutaminase active site of TG2 in the open conformation [3]. C277 is the catalytic residue, and is part of the catalytic triad along with D358 and H335. D358 hydrogen bonds to H335, increasing the pKa of the H335, and allowing it to deprotonate C277. W241 and W332 line the catalytic tunnel and stabilize the enzyme-thiol intermediate.

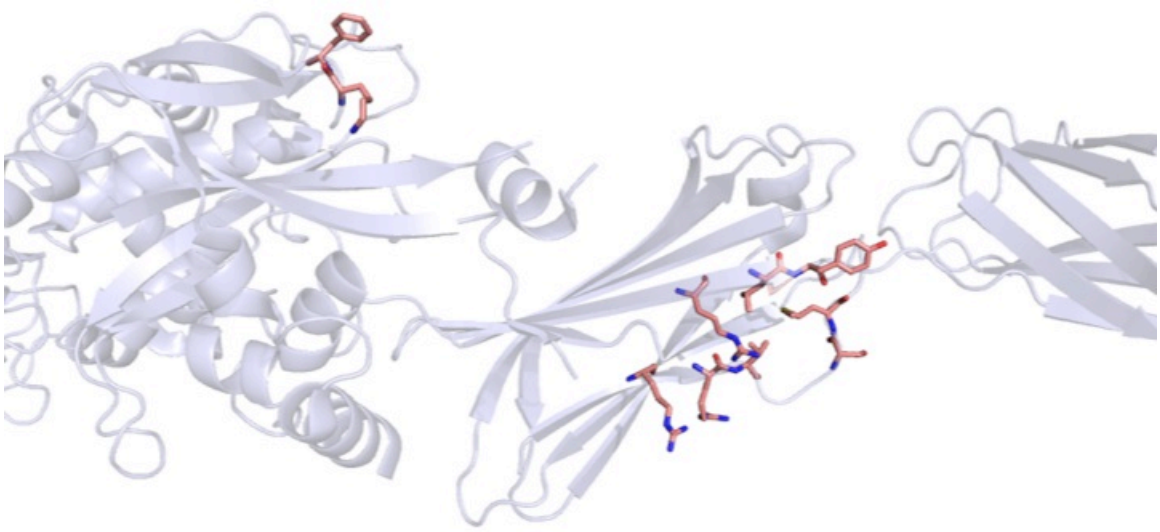


Figure 1-6. In the open conformation, the GTPase active site is torn in half as the residues making up this active site reside on both the core and the β1-barrel. Therefore, upon assuming the open conformation, TG2 was believed to be incapable of acting as a GTPase [3]. Active site residues shown in pink.

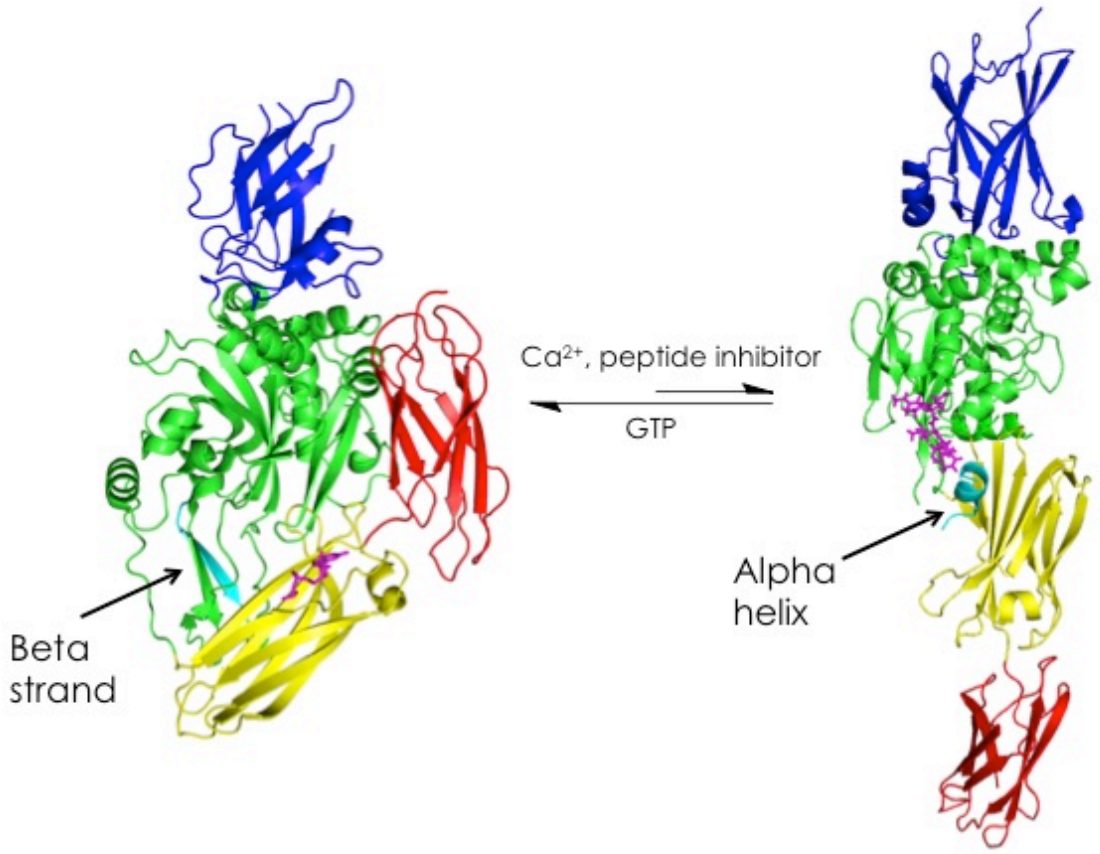


Figure 1-7. The hinge region of TG2 (residues 309 through 319) forms a β-stand in the closed conformation [5], and an α-helix in the open [3].

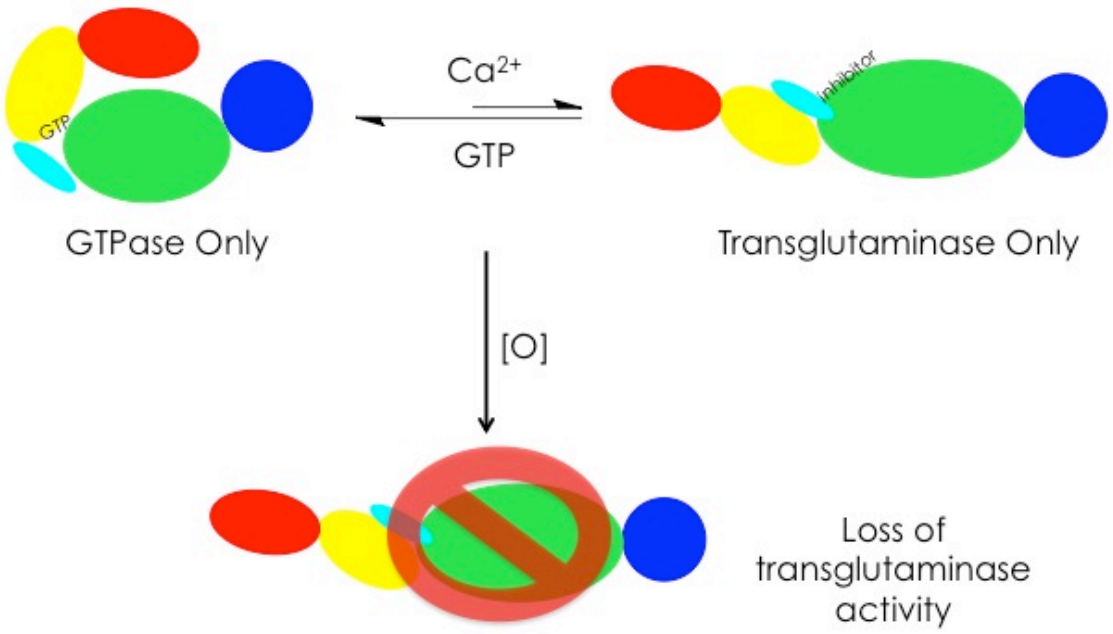


Figure 1-8. Multiple effectors regulate TG2's activities and conformations. Ca^{2+} inhibits GTPase activity and drives the adoption of the open conformation [11], while GTP inhibits transglutaminase activity and leads to the adoption of the closed conformation [14]. Oxidation of TG2 causes the formation of a disulfide bridge between C370 and C371, locking the enzyme in the open conformation and inactivating it [20, 21].

References

1. Perutz, M.F., *Stereochemistry of cooperative effects in haemoglobin.* Nature, 1970. **228**(5273): p. 726-39.

2. Prusiner, S.B., *Novel proteinaceous infectious particles cause scrapie.* Science, 1982. **216**(4542): p. 136-44.

3. Pinkas, D.M., et al., *Transglutaminase 2 undergoes a large conformational change upon activation.* PLoS Biol, 2007. **5**(12): p. e327.

4. Nakaoka, H., et al., *Gh: a GTP-binding protein with transglutaminase activity and receptor signaling function.* Science, 1994. **264**(5165): p. 1593-6.

5. Liu, S., R.A. Cerione, and J. Clardy, *Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity.* Proc Natl Acad Sci U S A, 2002. **99**(5): p. 2743-7.

6. Folk, J.E., *Transglutaminases.* Annu Rev Biochem, 1980. **49**: p. 517-31.

7. Fesus, L., et al., *Transglutaminase induction by various cell death and apoptosis pathways.* Experientia, 1996. **52**(10-11): p. 942-9.

8. Park, D., S.S. Choi, and K.S. Ha, *Transglutaminase 2: a multi-functional protein in multiple subcellular compartments.* Amino Acids, 2010. **39**(3): p. 619-31.

9. Klock, C., T.R. Diraimondo, and C. Khosla, *Role of transglutaminase 2 in celiac disease pathogenesis.* Semin Immunopathol, 2012. **34**(4): p. 513-22.

10. Gundemir, S., et al., *Transglutaminase 2: a molecular Swiss army knife.* Biochim Biophys Acta, 2012. **1823**(2): p. 406-19.

11. Kiraly, R., et al., *Functional significance of five noncanonical Ca2+-binding sites of human transglutaminase 2 characterized by site-directed mutagenesis.* FEBS J, 2009. **276**(23): p. 7083-96.

12. Lai, T.S., et al., *Identification of two GTP-independent alternatively spliced forms of tissue transglutaminase in human leukocytes, vascular smooth muscle, and endothelial cells.* FASEB J, 2007. **21**(14): p. 4131-43.

13. Datta, S., M.A. Antonyak, and R.A. Cerione, *Importance of Ca(2+)-dependent transamidation activity in the protection afforded by tissue transglutaminase against doxorubicin-induced apoptosis.* Biochemistry, 2006. **45**(44): p. 13163-74.

14. Begg, G.E., et al., *Mutation of a critical arginine in the GTP-binding site of transglutaminase 2 disinhibits intracellular cross-linking activity.* J Biol Chem, 2006. **281**(18): p. 12603-9.

15. Bergamini, C.M. and M. Signorini, *Calcium dependent reversible inactivation of erythrocyte transglutaminase by acrylamide.* Biochem Int, 1988. **17**(5): p. 855-62.

16. Lai, T.S., et al., *Regulation of human tissue transglutaminase function by magnesium-nucleotide complexes. Identification of distinct binding sites for Mg-GTP and Mg-ATP.* Journal of Biological Chemistry, 1998. **273**(3): p. 1776-81.

17. Smethurst, P.A. and M. Griffin, *Measurement of tissue transglutaminase activity in a permeabilized cell system: its regulation by Ca2+ and nucleotides.* Biochem J, 1996. **313 (Pt 3)**: p. 803-8.

18. Zhang, J., et al., *Modulation of the in situ activity of tissue transglutaminase by calcium and GTP.* Journal of Biological Chemistry, 1998. **273**(4): p. 2288-95.

19. Siegel, M., et al., *Extracellular transglutaminase 2 is catalytically inactive, but is transiently activated upon tissue injury.* PLoS One, 2008. **3**(3): p. e1861.

20. Stamnaes, J., et al., *Redox regulation of transglutaminase 2 activity.* J Biol Chem, 2010. **285**(33): p. 25402-9.

21. Jin, X., et al., *Activation of extracellular transglutaminase 2 by thioredoxin.* Journal of Biological Chemistry, 2011. **286**(43): p. 37866-73.

22. Diraimondo, T.R., C. Klock, and C. Khosla, *Interferon-gamma activates transglutaminase 2 via a phosphatidylinositol-3-kinase-dependent pathway: implications for celiac sprue therapy.* J Pharmacol Exp Ther, 2012. **341**(1): p. 104-14.

23. Ientile, R., D. Caccamo, and M. Griffin, *Tissue transglutaminase and the stress response.* Amino Acids, 2007. **33**(2): p. 385-94.

24. Johnson, G.V., et al., *Transglutaminase activity is increased in Alzheimer's disease brain.* Brain Research, 1997. **751**(2): p. 323-9.

25. Fesus, L. and M. Piacentini, *Transglutaminase 2: an enigmatic enzyme with diverse functions.* Trends Biochem Sci, 2002. **27**(10): p. 534-9.

26. Zemskov, E.A., et al., *The role of tissue transglutaminase in cell-matrix interactions.* Front Biosci, 2006. **11**: p. 1057-76.

27. Aeschlimann, D., O. Kaupp, and M. Paulsson, *Transglutaminase-catalyzed matrix cross-linking in differentiating cartilage: identification of osteonectin as a major glutaminyl substrate.* J Cell Biol, 1995. **129**(3): p. 881-92.

28. Korner, G., et al., *Bovine aortic endothelial cell transglutaminase. Enzyme characterization and regulation of activity.* Biochem J, 1989. **262**(2): p. 633-41.

29. Lesort, M., et al., *Distinct nuclear localization and activity of tissue transglutaminase.* J Biol Chem, 1998. **273**(20): p. 11991-4.

30. Singh, U.S., J.W. Erickson, and R.A. Cerione, *Identification and biochemical characterization of an 80 kilodalton GTP-binding/transglutaminase from rabbit liver nuclei.* Biochemistry, 1995. **34**(48): p. 15863-71.

31. Peng, X., et al., *Interaction of tissue transglutaminase with nuclear transport protein importin-alpha3.* FEBS Lett, 1999. **446**(1): p. 35-9.

32. Rodolfo, C., et al., *Tissue transglutaminase is a multifunctional BH3-only protein.* Journal of Biological Chemistry, 2004. **279**(52): p. 54783-92.

33. Malorni, W., et al., *The adenine nucleotide translocator 1 acts as a type 2 transglutaminase substrate: implications for mitochondrial-dependent apoptosis.* Cell Death Differ, 2009. **16**(11): p. 1480-92.

34. Upchurch, H.F., et al., *Localization of cellular transglutaminase on the extracellular matrix after wounding: characteristics of the matrix bound enzyme.* J Cell Physiol, 1991. **149**(3): p. 375-82.

35. Vezza, R., A. Habib, and G.A. FitzGerald, *Differential signaling by the thromboxane receptor isoforms via the novel GTP-binding protein, Gh.* Journal of Biological Chemistry, 1999. **274**(18): p. 12774-9.

36. Baek, K.J., et al., *Oxytocin receptor couples to the 80 kDa Gh alpha family protein in human myometrium.* Biochem J, 1996. **315 (Pt 3)**: p. 739-44.

37. Zhang, J., et al., *Novel bimodal effects of the G-protein tissue transglutaminase on adrenoreceptor signalling.* Biochem J, 1999. **343 Pt 3**: p. 541-9.

38. Nanda, N., et al., *Targeted inactivation of Gh/tissue transglutaminase II.* J Biol Chem, 2001. **276**(23): p. 20673-8.

39. Gaudry, C.A., et al., *Cell surface localization of tissue transglutaminase is dependent on a fibronectin-binding site in its N-terminal beta-sandwich domain.* Journal of Biological Chemistry, 1999. **274**(43): p. 30707-14.

40. Gaudry, C.A., et al., *Tissue transglutaminase is an important player at the surface of human endothelial cells: evidence for its externalization and its colocalization with the beta(1) integrin.* Exp Cell Res, 1999. **252**(1): p. 104-13.

41. Aeschlimann, D., D. Mosher, and M. Paulsson, *Tissue transglutaminase and factor XIII in cartilage and bone remodeling.* Semin Thromb Hemost, 1996. **22**(5): p. 437-43.

42. Aeschlimann, D. and V. Thomazy, *Protein crosslinking in assembly and remodelling of extracellular matrices: the role of transglutaminases.* Connect Tissue Res, 2000. **41**(1): p. 1-27.

43. Jones, R.A., et al., *Reduced expression of tissue transglutaminase in a human endothelial cell line leads to changes in cell spreading, cell adhesion and reduced polymerisation of fibronectin.* Journal of Cell Science, 1997. **110 (Pt 19)**: p. 2461-72.

44. Mehta, K., J.Y. Fok, and L.S. Mangala, *Tissue transglutaminase: from biological glue to cell survival cues.* Front Biosci, 2006. **11**: p. 173-85.

45. Mahoney, S.A., et al., *Stabilization of neurites in cerebellar granule cells by transglutaminase activity: identification of midkine and galectin-3 as substrates.* Neuroscience, 2000. **101**(1): p. 141-55.

46. Eitan, S. and M. Schwartz, *A transglutaminase that converts interleukin-2 into a factor cytotoxic to oligodendrocytes.* Science, 1993. **261**(5117): p. 106-8.

47. Eitan, S., et al., *Recovery of visual response of injured adult rat optic nerves treated with transglutaminase.* Science, 1994. **264**(5166): p. 1764-8.

48. Takeuchi, Y., et al., *Nuclear translocation of tissue type transglutaminase during sphingosine-induced cell death: a novel aspect of the enzyme with DNA hydrolytic activity.* Z Naturforsch C, 1998. **53**(5-6): p. 352-8.

49. Campisi, A., et al., *Glutamate-evoked redox state alterations are involved in tissue transglutaminase upregulation in primary astrocyte cultures.* FEBS Lett, 2004. **578**(1-2): p. 80-4.

50. Lesort, M., et al., *Distinct nuclear localization and activity of tissue transglutaminase.* Journal of Biological Chemistry, 1998. **273**(20): p. 11991-4.

51. Mann, A.P., et al., *Overexpression of tissue transglutaminase leads to constitutive activation of nuclear factor-kappaB in cancer cells: delineation of a novel pathway.* Cancer Res, 2006. **66**(17): p. 8788-95.

52. Tatsukawa, H., et al., *Role of transglutaminase 2 in liver injury via cross-linking and silencing of transcription factor Sp1.* Gastroenterology, 2009. **136**(5): p. 1783-95 e10.

53. Filiano, A.J., et al., *Transglutaminase 2 protects against ischemic insult, interacts with HIF1beta, and attenuates HIF1 signaling.* FASEB J, 2008. **22**(8): p. 2662-75.

54. Filiano, A.J., et al., *Transglutaminase 2 protects against ischemic stroke.* Neurobiol Dis, 2010. **39**(3): p. 334-43.

55. Jang, G.Y., et al., *Transglutaminase 2 suppresses apoptosis by modulating caspase 3 and NF-kappaB activity in hypoxic tumor cells.* Oncogene, 2010. **29**(3): p. 356-67.

56. Balajthy, Z., et al., *Tissue-transglutaminase contributes to neutrophil granulocyte differentiation and functions.* Blood, 2006. **108**(6): p. 2045-54.

57. Ahn, J.S., et al., *Tissue transglutaminase-induced down-regulation of matrix metalloproteinase-9.* Biochem Biophys Res Commun, 2008. **376**(4): p. 743-7.

58. Griffin, M., R. Casadio, and C.M. Bergamini, *Transglutaminases: nature's biological glues.* Biochem J, 2002. **368**(Pt 2): p. 377-96.

59. Satchwell, T.J., et al., *Protein 4.2: a complex linker.* Blood Cells Mol Dis, 2009. **42**(3): p. 201-10.

60. Candi, E., et al., *Biochemical, structural, and transglutaminase substrate properties of human loricrin, the major epidermal cornified cell envelope protein.* Journal of Biological Chemistry, 1995. **270**(44): p. 26382-90.

61. Candi, E., et al., *Transglutaminase 5 cross-links loricrin, involucrin, and small proline-rich proteins in vitro.* Journal of Biological Chemistry, 2001. **276**(37): p. 35014-23.

62. Candi, E., R. Schmidt, and G. Melino, *The cornified envelope: a model of cell death in the skin.* Nat Rev Mol Cell Biol, 2005. **6**(4): p. 328-40.

63. John, S., et al., *Epidermal transglutaminase (TGase 3) is required for proper hair development, but not the formation of the epidermal barrier.* Plos One, 2012. **7**(4): p. e34252.

64. Hitomi, K., et al., *Analysis of epidermal-type transglutaminase (TGase 3) expression in mouse tissues and cell lines.* Int J Biochem Cell Biol, 2001. **33**(5): p. 491-8.

65. Sardy, M., et al., *Epidermal transglutaminase (TGase 3) is the autoantigen of dermatitis herpetiformis.* J Exp Med, 2002. **195**(6): p. 747-57.

66. Kim, H.C., et al., *Protransglutaminase E from guinea pig skin. Isolation and partial characterization.* Journal of Biological Chemistry, 1990. **265**(35): p. 21971-8.

67. Williams-Ashman, H.G., et al., *Transamidase reactions involved in the enzymic coagulation of semen: isolation of -glutamyl- -lysine dipeptide from clotted secretion protein of guinea pig seminal vesicle.* Proc Natl Acad Sci U S A, 1972. **69**(8): p. 2322-5.

68. Komaromi, I., Z. Bagoly, and L. Muszbek, *Factor XIII: novel structural and functional aspects.* J Thromb Haemost, 2011. **9**(1): p. 9-20.

69. Muszbek, L., V.C. Yee, and Z. Hevessy, *Blood coagulation factor XIII: structure and function.* Thromb Res, 1999. **94**(5): p. 271-305.

70. Muszbek, L., et al., *The involvement of blood coagulation factor XIII in fibrinolysis and thrombosis.* Cardiovasc Hematol Agents Med Chem, 2008. **6**(3): p. 190-205.

71. Katona, E., et al., *A simple, quick one-step ELISA assay for the determination of complex plasma factor XIII (A2B2).* Thromb Haemost, 2000. **83**(2): p. 268-73.

72. Inbal, A., et al., *Impaired wound healing in factor XIII deficient mice.* Thromb Haemost, 2005. **94**(2): p. 432-7.

73. Dardik, R., J. Loscalzo, and A. Inbal, *Factor XIII (FXIII) and angiogenesis.* J Thromb Haemost, 2006. **4**(1): p. 19-25.

74. Noll, T., et al., *Effect of factor XIII on endothelial barrier function.* J Exp Med, 1999. **189**(9): p. 1373-82.

75. Hirahara, K., et al., *Suppressive effect of human blood coagulation factor XIII on the vascular permeability induced by anti-guinea pig endothelial cell antiserum in guinea pigs.* Thromb Res, 1993. **71**(2): p. 139-48.

76. Nurminskaya, M. and M.T. Kaartinen, *Transglutaminases in mineralized tissues.* Front Biosci, 2006. **11**: p. 1591-606.

77. Lorand, L. and R.M. Graham, *Transglutaminases: crosslinking enzymes with pleiotropic functions.* Nat Rev Mol Cell Biol, 2003. **4**(2): p. 140-56.

78. Weiss, M.S., H.J. Metzner, and R. Hilgenfeld, *Two non-proline cis peptide bonds may be important for factor XIII function.* FEBS Lett, 1998. **423**(3): p. 291-6.

79. Yee, V.C., et al., *Three-dimensional structure of a transglutaminase: human blood coagulation factor XIII.* Proc Natl Acad Sci U S A, 1994. **91**(15): p. 7296-300.

80. Ahvazi, B., et al., *Three-dimensional structure of the human transglutaminase 3 enzyme: binding of calcium ions changes structure for activation.* EMBO J, 2002. **21**(9): p. 2055-67.

81. Hornyak, T.J. and J.A. Shafer, *Role of calcium ion in the generation of factor XIII activity.* Biochemistry, 1991. **30**(25): p. 6175-82.

82. Lewis, S.D., et al., *Regulation of formation of factor XIIIa by its fibrin substrates.* Biochemistry, 1985. **24**(24): p. 6772-7.

83. Ortner, E., et al., *Sensitive and selective detection of free FXIII activation peptide: a potential marker of acute thrombotic events.* Blood, 2010. **115**(24): p. 5089-96.

84. Radek, J.T., et al., *Association of the A subunits of recombinant placental factor XIII with the native carrier B subunits from human plasma.* Biochemistry, 1993. **32**(14): p. 3527-34.

85. Schroeder, V., et al., *Factor XIII activation peptide is released into plasma upon cleavage by thrombin and shows a different structure compared to its bound form.* Thromb Haemost, 2007. **97**(6): p. 890-8.

86. Lorand, L., et al., *Human plasma factor XIII: subunit interactions and activation of zymogen.* Methods Enzymol, 1993. **222**: p. 22-35.

87. Sabo, T.M., P.B. Brasher, and M.C. Maurer, *Perturbations in factor XIII resulting from activation and inhibition examined by solution based methods and detected by MALDI-TOF MS.* Biochemistry, 2007. **46**(35): p. 10089-101.

88. Turner, B.T., Jr., et al., *Mapping of factor XIII solvent accessibility as a function of activation state using chemical modification methods.* Biochemistry, 2004. **43**(30): p. 9755-65.

89. Turner, B.T., Jr. and M.C. Maurer, *Evaluating the roles of thrombin and calcium in the activation of coagulation factor XIII using H/D exchange and MALDI-TOF MS.* Biochemistry, 2002. **41**(25): p. 7947-54.

90. Yee, V.C., et al., *Structural evidence that the activation peptide is not released upon thrombin cleavage of factor XIII.* Thromb Res, 1995. **78**(5): p. 389-97.

91. Kim, I.G., et al., *The deduced sequence of the novel protransglutaminase E (TGase3) of human and mouse.* Journal of Biological Chemistry, 1993. **268**(17): p. 12682-90.

92. Ahvazi, B., et al., *Roles of calcium ions in the activation and activity of the transglutaminase 3 enzyme.* J Biol Chem, 2003. **278**(26): p. 23834-41.

93. Ahvazi, B., K.M. Boeshans, and F. Rastinejad, *The emerging structural understanding of transglutaminase 3.* J Struct Biol, 2004. **147**(2): p. 200-7.

94. Csosz, E., B. Mesko, and L. Fesus, *Transdab wiki: the interactive transglutaminase substrate database on web 2.0 surface.* Amino Acids, 2009. **36**(4): p. 615-7.

95. Sakata, Y. and N. Aoki, *Cross-linking of alpha 2-plasmin inhibitor to fibrin by fibrin-stabilizing factor.* J Clin Invest, 1980. **65**(2): p. 290-7.

96. Lorand, L., K. Konishi, and A. Jacobsen, *Transpeptidation mechanism in blood clotting.* Nature, 1962. **194**: p. 1148-9.

97. Iismaa, S.E., et al., *The core domain of the tissue transglutaminase Gh hydrolyzes GTP and ATP.* Biochemistry, 1997. **36**(39): p. 11655-64.

98. Lai, T.S., et al., *C-terminal deletion of human tissue transglutaminase enhances magnesium-dependent GTP/ATPase activity.* Journal of Biological Chemistry, 1996. **271**(49): p. 31191-31195.

99. Sollid, L.M. and B. Jabri, *Is celiac disease an autoimmune disorder?* Curr Opin Immunol, 2005. **17**(6): p. 595-600.

100. Koning, F., et al., *Pathomechanisms in celiac disease.* Best Pract Res Clin Gastroenterol, 2005. **19**(3): p. 373-87.

101. Dicke, W.K., H.A. Weijers, and J.H. Van De Kamer, *Coeliac disease. II. The presence in wheat of a factor having a deleterious effect in cases of coeliac disease.* Acta Paediatr, 1953. **42**(1): p. 34-42.

102. Jabri, B. and L.M. Sollid, *Tissue-mediated control of immunopathology in coeliac disease.* Nat Rev Immunol, 2009. **9**(12): p. 858-70.

103. Hernandez, L. and P.H. Green, *Extraintestinal manifestations of celiac disease.* Curr Gastroenterol Rep, 2006. **8**(5): p. 383-9.

104. Wieser, H., *Chemistry of gluten proteins.* Food Microbiol, 2007. **24**(2): p. 115-9.

105. Shan, L., et al., *Structural basis for gluten intolerance in celiac sprue.* Science, 2002. **297**(5590): p. 2275-9.

106. Quarsten, H., et al., *HLA binding and T cell recognition of a tissue transglutaminase-modified gliadin epitope.* Eur J Immunol, 1999. **29**(8): p. 2506-14.

107. Sollid, L.M. and E. Thorsby, *HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis.* Gastroenterology, 1993. **105**(3): p. 910-22.

108. Nilsen, E.M., et al., *Gluten specific, HLA-DQ restricted T cells from coeliac mucosa produce cytokines with Th1 or Th0 profile dominated by interferon gamma.* Gut, 1995. **37**(6): p. 766-76.

109. Nilsen, E.M., et al., *Gluten induces an intestinal cytokine response strongly dominated by interferon gamma in patients with celiac disease.* Gastroenterology, 1998. **115**(3): p. 551-63.

110. Alaedini, A. and P.H. Green, *Narrative review: celiac disease: understanding a complex autoimmune disorder.* Ann Intern Med, 2005. **142**(4): p. 289-98.

111. Dieterich, W., et al., *Identification of tissue transglutaminase as the autoantigen of celiac disease.* Nat Med, 1997. **3**(7): p. 797-801.

112. Maki, M., *The humoral immune system in coeliac disease.* Baillieres Clin Gastroenterol, 1995. **9**(2): p. 231-49.

113. Dieterich, W., et al., *Autoantibodies to tissue transglutaminase as predictors of celiac disease.* Gastroenterology, 1998. **115**(6): p. 1317-21.

114. Fasano, A., et al., *Prevalence of celiac disease in at-risk and not-at-risk groups in the United States: a large multicenter study.* Arch Intern Med, 2003. **163**(3): p. 286-92.

115. Caja, S., et al., *Antibodies in celiac disease: implications beyond diagnostics.* Cell Mol Immunol, 2011. **8**(2): p. 103-9.

116. Korponay-Szabo, I.R., et al., *In vivo targeting of intestinal and extraintestinal transglutaminase 2 by coeliac autoantibodies.* Gut, 2004. **53**(5): p. 641-8.

117. Salmi, T.T., et al., *Endomysial antibody-negative coeliac disease: clinical characteristics and intestinal autoantibody deposits.* Gut, 2006. **55**(12): p. 1746-53.

118. Rantala, I., et al., *Periodate-lysine-paraformaldehyde as fixative for the study of duodenal mucosa. Morphologic and immunohistochemical results at light and electron microscopic levels.* Acta Pathol Microbiol Immunol Scand A, 1985. **93**(4): p. 165-73.

119. D'Argenio, G., et al., *Human serum transglutaminase and coeliac disease: correlation between serum and mucosal activity in an experimental model of rat small bowel enteropathy.* Gut, 1989. **30**(7): p. 950-4.

120. Esposito, C., et al., *Expression and enzymatic activity of small intestinal tissue transglutaminase in celiac disease.* Am J Gastroenterol, 2003. **98**(8): p. 1813-20.

121. Halttunen, T. and M. Maki, *Serum immunoglobulin A from patients with celiac disease inhibits human T84 intestinal crypt epithelial cell differentiation.* Gastroenterology, 1999. **116**(3): p. 566-72.

122. Cervio, E., et al., *Sera of patients with celiac disease and neurologic disorders evoke a mitochondrial-dependent apoptosis in vitro.* Gastroenterology, 2007. **133**(1): p. 195-206.

123. Di Simone, N., et al., *Anti-tissue transglutaminase antibodies from celiac patients are responsible for trophoblast damage via apoptosis in vitro.* Am J Gastroenterol, 2010. **105**(10): p. 2254-61.

124. Myrsky, E., et al., *Celiac disease IgA modulates vascular permeability in vitro through the activity of transglutaminase 2 and RhoA.* Cell Mol Life Sci, 2009. **66**(20): p. 3375-85.

125. Ensari, A., et al., *Time-course of adhesion molecule expression in rectal mucosa of gluten-sensitive subjects after gluten challenge.* Clin Exp Immunol, 1993. **92**(2): p. 303-7.

126. Jelinkova, L., et al., *Increased levels of circulating ICAM-1, E-selectin, and IL-2 receptors in celiac disease.* Dig Dis Sci, 2000. **45**(2): p. 398-402.

127. Di Sabatino, A., et al., *Increased expression of mucosal addressin cell adhesion molecule 1 in the duodenum of patients with active celiac disease is associated with depletion of integrin alpha4beta7-positive T cells in blood.* Hum Pathol, 2009. **40**(5): p. 699-704.

128. Myrsky, E., et al., *Coeliac disease-specific autoantibodies targeted against transglutaminase 2 disturb angiogenesis.* Clin Exp Immunol, 2008. **152**(1): p. 111-9.

129. Caja, S., et al., *Inhibition of transglutaminase 2 enzymatic activity ameliorates the anti-angiogenic effects of coeliac disease autoantibodies.* Scand J Gastroenterol, 2010. **45**(4): p. 421-7.

130. Martucciello, S., et al., *RhoB is associated with the anti-angiogenic effects of celiac patient transglutaminase 2-targeted autoantibodies.* J Mol Med (Berl), 2012. **90**(7): p. 817-26.

131. Castellanos-Rubio, A., et al., *Angiogenesis-related gene expression analysis in celiac disease.* Autoimmunity, 2012. **45**(3): p. 264-70.

132. Kalliokoski, S., et al., *Celiac Disease-Specific TG2-Targeted Autoantibodies Inhibit Angiogenesis and in Mice by Interfering with Endothelial Cell Dynamics.* PLoS One, 2013. **8**(6): p. e65887.

133. Cooke WT, H.G., *Coeliac Disease*1984, New York: Churchill Livingstone.

134. Haroon, Z.A., et al., *Tissue transglutaminase is expressed, active, and directly involved in rat dermal wound healing and angiogenesis.* FASEB J, 1999. **13**(13): p. 1787-95.

135. Haroon, Z.A., et al., *Tissue transglutaminase is expressed as a host response to tumor invasion and inhibits tumor growth.* Lab Invest, 1999. **79**(12): p. 1679-86.

136. Jones, R.A., et al., *Matrix changes induced by transglutaminase 2 lead to inhibition of angiogenesis and tumor growth.* Cell Death Differ, 2006. **13**(9): p. 1442-53.

137. Ingber, D.E., *Mechanical signaling and the cellular response to extracellular matrix in angiogenesis and cardiovascular physiology.* Circ Res, 2002. **91**(10): p. 877-87.

138. Bell, S.E., et al., *Differential gene expression during capillary morphogenesis in 3D collagen matrices: regulated expression of genes involved in basement membrane matrix assembly, cell cycle progression, cellular differentiation and G-protein signaling.* J Cell Sci, 2001. **114**(Pt 15): p. 2755-73.

139. Esposito, C., et al., *Anti-tissue transglutaminase antibodies from coeliac patients inhibit transglutaminase activity both in vitro and in situ.* Gut, 2002. **51**(2): p. 177-81.

140. Dieterich, W., et al., *Autoantibodies of patients with coeliac disease are insufficient to block tissue transglutaminase activity.* Gut, 2003. **52**(11): p. 1562-6.

141. Barone, M.V., et al., *Humoral immune response to tissue transglutaminase is related to epithelial cell proliferation in celiac disease.* Gastroenterology, 2007. **132**(4): p. 1245-53.

142. Kiraly, R., et al., *Coeliac autoantibodies can enhance transamidating and inhibit GTPase activity of tissue transglutaminase: dependence on reaction environment and enzyme fitness.* J Autoimmun, 2006. **26**(4): p. 278-87.

143. Seissler, J., et al., *Autoantibodies from patients with coeliac disease recognize distinct functional domains of the autoantigen tissue transglutaminase.* Clin Exp Immunol, 2001. **125**(2): p. 216-21.

144. Sblattero, D., et al., *The analysis of the fine specificity of celiac disease antibodies using tissue transglutaminase fragments.* Eur J Biochem, 2002. **269**(21): p. 5175-81.

145. Nakachi, K., et al., *Epitopes recognised by tissue transglutaminase antibodies in coeliac disease.* J Autoimmun, 2004. **22**(1): p. 53-63.

146. Di Niro, R., et al., *Characterizing monoclonal antibody epitopes by filtered gene fragment phage display.* Biochem J, 2005. **388**(Pt 3): p. 889-94.

147. Sulkanen, S., et al., *Tissue transglutaminase autoantibody enzyme-linked immunosorbent assay in detecting celiac disease.* Gastroenterology, 1998. **115**(6): p. 1322-8.

148. Simon-Vecsei, Z., et al., *A single conformational transglutaminase 2 epitope contributed by three domains is critical for celiac antibody binding and effects.* Proc Natl Acad Sci U S A, 2012. **109**(2): p. 431-6.

149. Begg, G.E., et al., *Mechanism of allosteric regulation of transglutaminase 2 by GTP.* Proc Natl Acad Sci U S A, 2006. **103**(52): p. 19683-8.

150. Gundemir, S. and G.V. Johnson, *Intracellular localization and conformational state of transglutaminase 2: implications for cell death.* PLoS One, 2009. **4**(7): p. e6123.

151. Ruan, Q., et al., *The Differential Effects of R580A Mutation on Transamidation and GTP Binding Activity of Rat and Human Type 2 Transglutaminase.* Int J Clin Exp Med, 2008. **1**(3): p. 248-59.

152. Shimaoka, M., et al., *Computational design of an integrin I domain stabilized in the open high affinity conformation.* Nat Struct Biol, 2000. **7**(8): p. 674-8.

153. Dahiyat, B.I. and S.L. Mayo, *Protein design automation.* Protein Sci, 1996. **5**(5): p. 895-903.

154. Dahiyat, B.I. and S.L. Mayo, *Probing the role of packing specificity in protein design.* Proc Natl Acad Sci U S A, 1997. **94**(19): p. 10172-7.

155. Lazaridis, T. and M. Karplus, *Effective energy function for proteins in solution.* Proteins, 1999. **35**(2): p. 133-52.

156. Street, A.G. and S.L. Mayo, *Pairwise calculation of protein solvent-accessible surface areas.* Fold Des, 1998. **3**(4): p. 253-8.

157. Dahiyat, B.I., D.B. Gordon, and S.L. Mayo, *Automated design of the surface positions of protein helices.* Protein Sci, 1997. **6**(6): p. 1333-7.

158. Desmet, J., et al., *The dead-end elimination theorem and its use in protein side-chain positioning.* Nature, 1992. **356**(6369): p. 539-42.

159. Desmet, J., J. Spriet, and I. Lasters, *Fast and accurate side-chain topology and energy refinement (FASTER) as a new method for protein structure optimization.* Proteins, 2002. **48**(1): p. 31-43.

160. Allen, B.D. and S.L. Mayo, *Dramatic performance enhancements for the FASTER optimization algorithm.* J Comput Chem, 2006. **27**(10): p. 1071-5.

C h a p t e r 2

STABILIZATION OF TISSUE TRANSGLUTAMINASE IN OPEN OR CLOSED CONFORMATIONS USING COMPUTATIONAL MULT-STATE PROTEIN DESIGN

Abstract

Human type II transglutaminase (TG2) is an enzyme that exists in two dramatically different conformational states, and each was believed to have a unique activity. In the open, extended form, adopted upon calcium binding, the transglutaminase active site is exposed, which was believed to allow TG2 to catalyze the formation of an isopeptide bond between the sidechain of a peptide-bound glutamine and a primary amine. Upon GTP binding to a separate GTPase active site, TG2 adopts a compact closed conformation, which partially obstructs the transglutaminase active site and therefore was believed to only allow for GTPase activity. TG2 has been linked to celiac disease, as well as many other cellular processes, both physiological and pathological. However, the multi-functional nature of TG2 makes it difficult to determine the relevance of each activity and conformation separate from each other. In this project, we used computational multi-state and single-state protein design to engineer TG2 variants adopting either the open or closed conformation. These open-locked and closed-biased TG2 variants allowed us to characterize the catalytic capabilities of these conformations. Specifically, our designs allowed us to determine three previously unknown characteristics of TG2: (1) the open conformation of TG2 is insufficient for transglutaminase activity, and therefore that the Ca^{2+}-dependence of this reaction is derived from more than driving conformational change, (2) the closed-biased designs of TG2 have higher transglutaminase activity than the open-locked designs, and in some cases, WT, and (3) not only can both conformations act as a GTPase, but the open-locked TG2 is significantly more active than the WT or closed-biased designs. These discoveries have important implications for the celiac drug design and for understanding the catalytic capabilities and possible effects of TG2 under various biological conditions.

Introduction

Human type II transglutaminase [1] is a member of a large family of enzymes that catalyze the Ca^{2+}-dependent crosslinking of enzymes [2, 3]. This reaction, known as the transglutaminase or transamidation reaction, catalyzes the formation of an isopeptide bond between the sidechain of a peptide-bound glutamine and a primary amine from another protein, *i.e.*, a lysine residue [2, 4, 5]. This cross-linking reaction is used to modulate the strength of various biological structures. Other members of this enzyme family include TG3 and factor XIII, which are involved in differentiation in hair follicles and stabilizing fibrin and therefore blood clots, respectively [6-17]. TG2 is expressed in most mammalian tissues and is found both intercellularly and extracellularly [18-20] and is implicated in a wide variety of biological processes including differentiation and apoptosis [21-23], wound healing [24, 25], and neuronal growth and regeneration [26-28].

Unlike other transglutaminases, TG2 also acts as a GTPase and has been shown to adopt two dramatically different conformational states, each believed to have its own activity [29-31]. TG2 consists of four domains: (1) the N-terminal β-sandwich, (2) the core domain, (3) β-barrel 1, and (4) β-barrel 2. In the open conformation, these domains are aligned linearly resulting in an elongated form of the enzyme [29]. In the closed conformation, β-barrels 1 and 2 fold onto the core domain, much like the folding of a jack-knife, resulting in a compact conformation. Surprisingly, the secondary structure of these domains remain relatively unchanged in each conformation, and the majority of backbone rearrangement occurs in the region of the protein between the core and β-barrel 1 [30]. This region largely consists of random coils, save for a motif comprising residues 309 through 319, known as the hinge. In the closed form of the enzyme, these residues form a β-strand, while in the open conformation they adopt an α-helix.

In the closed conformation, the transglutaminase active site is wedged into a grove between the core and the interface of the β-barrel domains [30]. This was believed to block its activity. Conversely, the GTPase active site is located in the interface between the core and β-barrel 1 that forms in the closed conformation on the opposite side of TG2 relative to the transglutaminase active site. The GTPase active site residues were deduced from the closed structure of TG2 bound to GDP [30]. TG2 assumes the closed conformation upon GTP binding and is the preferred conformation of the enzyme. Intracellularly, TG2 uses this activity to act as an α-subunit of a G-protein [31].

The open conformation of TG2 occurs upon the binding of six Ca^{2+} ions to the core domain [29, 32]. However, the only crystal structure of the open conformation has no calcium ions bound to it

and was achieved through the covalent binding of a peptide inhibitor to the transglutaminase site [29]. In the open conformation, the transglutminase active site is exposed at the C-terminal region of the core domain of the protein, just above β-barrel 1. This exposure was believed to result in its activation. All of the transglutaminase active site residues are located on the core, and consist of a catalytic triad with an active site cysteine (C277), which attacks a glutamine's carboxamide, releasing ammonia. If water hydrolyzes the covalent thiol-intermediate, the reaction is a deamidiation and the glutamine is converted to a glutamate. If not, the reaction proceeds as described above, and a primary amine attacks the intermediate, resulting in an amide bond [33]. The GTPase active site is torn in two upon the formation of the open conformation, since the GTPase active site residues reside on both the core domain and the β-barrel 1 domain. Therefore, it was long assumed that the open conformation was incapable of acting as a GTPase [29].

As alluded to above, Ca^{2+} inhibits TG2's GTPase activity [32, 34, 35], and GTP inhibits the transglutaminase activity [36]. This inhibition/activation was believed to be due to the large-scale conformational change brought about by the binding of calcium or GTP (i.e., the binding of GTP causes the closed conformation, and therefore blocks transglutaminase activity and forms the GTPase site, and the binding of calcium causes the open conformation, and therefore prevents the GTPase activity and exposes the transglutaminase site, leading to its activity). The transglutaminase activity can also be inactivated by the oxidation of TG2, resulting in the formation of a disulfide bond between C371 and C370, which causes the enzyme to adopt the open conformation [37]. Therefore, under normal conditions intracellularly, where calcium concentrations are low and GTP concentrations are higher, TG2 is believed to act as a GTPase, not a transglutaminase [38, 39]. Extracellularly, where calcium concentrations are high and GTP low, TG2 primarily acts as a transglutaminase or in receptor signaling [40-49]. Due to the high redox potential of the extracellular matrix, however, extracellular TG2 may normally be oxidized and can be activated by various effectors [25, 37, 50-53], although the mechanism of how this occurs and when it is necessary is not yet clear.

As mentioned above, all transglutaminases require Ca^{2+} binding for transglutaminase activity [2]. Of the 9 members of the transglutaminase family, only 3 have been crystallized: TG2, TG3, and factor XII [29, 30, 54-56]. All three exhibit strong structural similarity. However, each of these three binds a different number of Ca^{2+} ions and has a different conformational response to this binding. TG2 binds 6 Ca^{2+} ions which causes the enzyme to adopt the open conformation, yet the open crystal structure had no Ca^{2+} ions bound [29, 32]. Of the three, only TG2 has been shown to adopt this elongated form. As detailed in Chapter 1, TG2 is also the most promiscuous in regards to

its glutamine substrate preferences of the three crystallized transglutaminases [38, 57, 58]. Factor XIII binds a single Ca^{2+} ion [59], and biochemical studies have shown that it may undergo some conformational changes during activation [60-62]. However, the crystal structure of thrombin activated and Ca^{2+}-bound factor XIII still adopted the closed conformation and did not appear significantly different from the non-activated crystal structure [55], implying that such conformational changes may be fleeting and unstable. As discussed in Chapter 1, Factor XIII has much stricter substrate preferences than TG2 [10, 58, 63, 64]. However, TG3 is the most discerning of all three, having the fewest number of known glutamine substrates [58, 65]. Interestingly, TG3, which binds three Ca^{2+} ions in the crystal structure, undergoes a cleavage event in the hinge region of the protein between the core domain and the β-barrel 1 domain, resulting in two globular domains that remain associated with each other in the closed conformation, even upon the addition of calcium [9, 53, 66]. Therefore, TG3 is not believed to adopt an open conformation. Substrates are believed to be able to access the active site due to two conformational adjustments: (1) calcium binding causes a loop to move and open a cavity though the enzyme to the active site and, (2) a flexing of the two subunits to accommodate the substrates [67].

As discussed in Chapter 1, the role of each domain in TG2's two activities is not entirely clear due to results from truncation studies that conflict with our crystallographic structural understanding of the active sites. For instance, although the GTPase active site comprises residues from the core domain and the β-barrel 1 domain, deletion of both β-barrel domains in two separate studies had little effect or actually increased GTPase activity relative to the full-length protein [68, 69]. All of these truncated variants lacked all but two of the active site residues.

In addition, these same two studies found that the loss of the β-barrel domains resulted in anywhere from a 30% to 99% loss in transgluaminase activity [68, 69], which is puzzling considering that the transglutaminase active site residues and the calcium binding sites are all located on the core of the protein and do not interact directly with the deleted domains.

Among its many biological roles, TG2 has also been implicated in pathological events. In celiac disease, for example, TG2 deamidates gluten peptides, the trigger of celiac disease, which increase their binding affinity for mutant major histocompatibility complex (MHC) class II molecules, which elicits the autoimmune response [70-77]. TG2 is also the primary autoantigen of celiac disease [72], and the anti-TG2 autoantibodies that arise in these patients have been shown to be able to elicit pathogenic responses all on their own [78-86]. Recently, studies found that autoantibodies against TG2 isolated from celiac patients increase TG2 transglutaminase activity and bind a motif that

exists in the closed conformation [81, 83, 87, 88], suggesting that the stabilization of this conformation leads to the increase in activity, and that this increase in activity may lead to the pathogenic effects observed upon the addition of these antibodies to various endothelial cell models.

If the closed conformation is in fact more active than the open form of the enzyme, it would fly in the face of what was for years believed about TG2, but would help account for the results observed in the truncation studies and the ability of TG3 and factor XIII to acts as transglutaminases despite failing to adopt an open, extended conformation. Since the promiscuity of these three transglutaminases increases with the increasing ability to adopt an open conformation, perhaps such opening is more important for allowing more flexibility in the binding of substrates, and less essential for transglutaminase activity itself. This calls in to question the hypothesis that calcium's only role is to cause the enzyme to adopt the open conformation and the assumption that this open conformation is transglutaminase active of its own merit.

One way to test the existing model for the activities and their activation, inhibition, and corresponding conformations of TG2 would be to engineer conformationally isolated variants of TG2. One method to accomplish such a goal is computational protein design (CPD).

To selectively stabilize one TG2 conformation over the alternate, a new variation of CPD will be used in addition to traditional design methods: multi-state design (MDS) [89, 90]. In MSD, the design process considers multiple structural states while determining a protein sequence. Here, possible amino acid sequences are suggested, and the suitability of each sequence is determined by calculating an overall energy score acquired from combining the individual energies of rotamer optimization calculations performed on each of the considered states [90].

During an MSD calculation, each structural state is assigned a single processor in a computer cluster. The job of each processor is to hold the pairwise energy matrix for its assigned state in memory, and to use this information to evaluate the energies of candidate sequences in that state only. An additional processor has the job of identifying sequences to be scored, sending these sequences to the other processors, collecting the energy results, and determining the overall fitness score from this information [71, 90]. MSD can be used to maximize specificity for a target fold by stabilizing the desired conformation (positive design) and destabilizing the undesirable conformation (negative design). MSD selects sequences maximizing the transfer of free energy of a design from a desired conformation to a group of unwanted conformations [91]. As mentioned previously, crystal structures of both the open and closed conformations of TG2 have been solved

[29, 30]. MSD can therefore be used to design an open-locked variant of TG2 by stabilizing the open conformation (positive design) while destabilizing the closed (negative design), and vice versa to design a closed-locked variant.

Open-locked and closed-locked versions of TG2 would allow us to characterize the catalytic capabilities of each conformation, and by understanding the biological conditions in which the enzyme resides in in various tissues, predict which activities the enzyme may be catalyzing in these scenarios. Understanding which conformations catalyze which reaction also has important implications for drug design; if the closed form of TG2 is catalyzing the deamidation of gluten in celiac disease, drugs should be designed to bind this conformation, and not the open. Therefore, the goal of this chapter was to use computational multi-state protein design to engineer and then characterize TG2 variants locked in either the open or closed conformation.

Methods

Design of open-locked and closed-biased TG2

To design variants of TG2 in open-locked and closed-biased conformations, two areas of the protein were selected for residue design (Figure 2-1). These regions were chosen because they show a high level of change in environment or structure with the shift between the closed and open forms. The two regions were: (a) the β2-core interface, and (b) the hinge. In the β2-core interface, residues go from being solvent exposed in the open form to buried in the closed. In the hinge, residues adopt a β-strand in the closed form and an α-helix in the open. Residues located in either active site (GTPase or transglutaminase), as well as those known to be involved in Ca^{2+} binding were excluded from all design calculations. The ResClass tool in the CPD program PHOENIX was used to characterize residues' degrees of solvent exposure [92]. All design calculations were carried out using the newly developed CPD program TRIAD and its multi-state design (MSD) capability. These calculations determined the optimal sequence (A_0) for each conformational state (s). These calculations were carried out in a manner similar to those described in Shimoaka *et al*. [93], except that occlusion-based solvation and the more versatile optimization algorithm FASTER was used [90, 93].

The MSD algorithm that was used in the design of closed-biased and open-locked TG2 is described elsewhere [89, 90]. Two calculations will be run for each of the three design regions: (1) an MSD calculation to stabilize the closed conformation of TG2 (closed-locked, PDB ID code: 1KV3), and (2) an MSD calculation to stabilize the open form of TG2 (open-locked, PDB ID code: 2Q3Z). In these calculations, stabilization of the positive design state (ρ) and destabilization of the negative design state (η) will be achieved using the MSD optimization algorithm to determine the sequence that minimizes the scoring function $\sigma(A)$:

$$\sigma(A) = \Delta E_{\rho}(A) - W\Delta E_{\eta}(A) \quad (1)$$

W a weighing factor used to control the relative contributions of ρ-state stabilization versus η-state destabilization. $\Delta E_s(A)$ is the difference in energy between sequence A in conformation s and the optimal sequence A_0 for conformation s that was determined by the initial single state design calculations:

$$\Delta E_s(A) = E_s(A) - E_s(A_0) \quad (2)$$

The open-locked TG2 designs were generated by focusing on the β2-core interface and the hinge region. Initial hinge designs were performed in TRIAD using the design residues 309 through 319, the rosetta force field, floating residues within 6 Å of the design residues, *i.e.*, their amino acid identities were retained, but their side chains could sample alternate conformations, allowing all residues at design sites, with "rama," the secondary structural term varied from 0.4 to 1.6, and the weighing factor, *W*, ranging from 0.05 to 0.5.

After being informed by these designs for the hinge region, Y315 was selected as the initial design residue. This residue was selected because it favors β-secondary structure over α, while the hinge is an α-helix in the open conformation, and because this residue forms a hydrogen bond with the backbone of E329 in the closed structure. To disrupt these interactions, we used the information gleaned from our initial designs and rational design and created to mutants: Y315E and Y315A.

To accommodate the Y315E mutation, MSD calculations were performed exactly as indicated above, except rama was defined as 0.8, Y315 was forced to glutamate, L508 was designated a design residue and allowed to sample all residues, residues R506, Y528, L566, and V568 were designated as design positions or restricted to WT, and *W* was defined as 0, making this calculation a single state design essentially.

For open-locked designs derived from mutations at the β2-core interface, residues I178, N181, S253, I255, G256, E557, V665, N667, and K677 were designated as design positions. MSD calculations were conducted as described above for the initial hinge designs, except rama was defined as 0.8, only polar residues were allowed at design positions, and *W* for the negative design was set at 0.05.

To design closed-biased TG2, residues L312, E314, and F316 were designated as design residues and were allowed to sample Val, Ile, Cys, Tyr, Thr, Phe, Trp, His, and Ser, which were chosen specifically for their preference for β-secondary structure over α. MSD calculations were carried out as they were for the initial hinge designs to favor the open structure as detailed above, except float positions were designated to be all residues within 9 Å of the design positions and the crosseta force field was used.

Assembly and cloning of human TG2 gene

Oligonucleotides of approximately 40 nucleotides (nt) in length optimized for expression in *Escherichia coli* (*E. coli*) were designed for gene assembly of N-terminally His-6 tagged human TG2 using the program DNAWorks [94]. The final gene was 1.2 kb in length.

An assembly reaction was conducted using 2.5 U of KOD Hot Start DNA Polymerase in 1X Buffer for KOD Hot Start Polymerase (Novagen), 1.5 mM $MgSO_4$, 0.2 mM of each dNTP, and 0.01 μM of each assembly oligonucleotide to a final volume of 50 μL. The PCR program consisted of 94°C for 2 min, 25 cycles of 94°C for 30s, 68.2°C for 30s, 68°C for 30s, followed by a final 10 min extension period at 68°C. After an initial run of this program, the program was repeated after the reactions were spiked with an additional 2.5 U of KOD Hot Start Polymerase and 0. 01 μmol of each dNTP.

An amplification reaction was then run using 2 U of KOD Hot Start DNA Polymerase in 1X Buffer for KOD Hot Start Polymerase (Novagen), 1.5 mM $MgSO_4$, 0.2 mM of each dNTP, 1 μM forward and 1 μM reverse amplification primers, and 1 μL of the assembly reaction at a final volume of 50 μL. Amplification PCR conditions consisted of 10 min at 94°C, 25 cycles of 94°C for 30s, 65°C for 30s, 68°C for 120s, and a final 10 min extension at 68°C.

The assembled TG2 gene was digested using NdeI and BamHI (New England Biolabs) and ligated into the *E. coli* expression vector pET11a (Novagen).

Site-directed mutagenesis

Mutagenesis primers were designed using Agilent Technologies Quikchange Primer Design tool. Mutants of TG2 were generated using a protocol based on the Quikchange Site-Directed Mutagenesis Kit (Stratagene). Briefly, 50 ng of template DNA was added to a 50 μl reaction mixture containing 5 μl of 10x reaction buffer, 2.5 U of *PfuTurbo* DNA polymerase, 0.08 mM dNTP mix, (Stratagene), and 125 ng of each mutagenesis primer. PCR was carried out on Mastercycler Personal Thermocycler (Eppendorf) with cycling conditions of 95°C for 30 sec, and then 18 cycles of 95°C for 30 sec, 52°C for 1 min, and 68°C for 14 min and 40 sec. Completed mutagenesis reactions were then *Dpn*I (New England Biolabs) digested with 40 U of the enzye for 2 hours at 37°C.

Protein expression and purification

LB/ampicillin starter cultures were inoculated from a glycerol stock of BL-21 (DE3) Gold *E. coli* containing the expression vector pET11a (Novagen) with the gene for the variant of interest and grown overnight at 37°C with shaking. 4 mL of starter culture were added to 100 mL of Overnight Express Instant TB media (Novagen) with ampicillin and grown at 37°C with shaking. After 3 hours, cultures were grown at 16°C for 24 hours with shaking.

Cells were harvested by centrifugation at 5000 x g for 10 minutes at 4°C. Each 100 mL pellet was resuspended in 8 mL of lysis buffer (50 mM sodium phosphate, pH 8.0, 0.3 M NaCl, 2.5 mM imidazole, 1x CelLytic B (Sigma-Aldrich), 5 mM β-mercaptoethanol (βME), 0.2 mg/mL lysozyme (Sigma-Aldrich), and 1 U Benzonase endonuclease (Merck)). The lysate was centrifuged at 19000 x g for 45 at 4°C and the supernatant collected.

His-Select spin columns (Sigma-Aldrich) were equilibrated with 600 µL of equilibrium buffer (50 mM sodium phosphate, pH 8.0, 0.3 M NaCl, 2.5 mM imidazole, and 5 mM βME). The crude supernatant was applied to the columns and the columns were washed with 1200 µL of wash buffer (50 mM sodium phosphate, pH 8.0, 0.3 M sodium chloride, 5 mM imidazole, and 5 mM βME). The purified protein was eluted by the addition of 300 µL of elution buffer (50 mM sodium phosphate, pH 8.0, 0.3 M sodium chloride, 250 mM imidazole and 5 mM βME).

Eluted protein was buffer exchanged into storage buffer (50 mM HEPES, pH 7.0, 10% glycerol, 100 mM NaCl, 1 mM EDTA, and 5 mM DTT) with 0.5 mL 50K Amicon Ultra centrifugal filters (Millipore).

Protein concentration determination

Protein concentrations were determined by UV absorbance denaturation in 8 M guanidinium hydrochloride for 5 minutes at a dilution of 10x to 100x. The extinction coefficient for each protein at 280 nm was calculated based on the number of tryptophans and tyrosines in the protein using ExPASy's ProtParam tool.

Native PAGE to monitor conformation

Native gel electrophoresis experiments were performed similar to those previously used to monitor the conformational state of TG2 [29, 36, 95]. 5 μM TG2 was incubated for one hour at room temperature in preincubation buffer (75 mM imidazole, 0.5 mM EDTA, 5 mM DTT, pH 7.2) in each of four separate conditions: (1) no effectors, (2) 500 μM GTP and 1 mM $MgCl_2$, (3) 5 mM $CaCl_2$, and (4) 25 μM peptide inhibitor (Figure 2-2) and 5 mM $CaCl_2$. The peptide inhibitor was a gift from the Child Health and Development Institute. After the incubation, the samples were diluted 2-fold with 2x sample buffer (0.125 M Tris, pH6.8, 20% glycerol, 0.005% (w/v) bromophenol blue, and 20 mM DTT). Diluted samples were then added to Bio-Rad, 4-20% Mini-PROTEAN® TGX™ precast gels and run at 4°C for 150 minutes at 125 V in Tris-glycine native running buffer. Gels were stained with Coomassie blue and destained with destain solution (20% methanol, 10% acetic acid).

Transglutaminase activity assay of TG2 catalyzed incorporation of dansylcadaverine into N,N-dimethylated casein

The TG2 catalyzed iospeptide bond formed between dansylcadaverine (DSC), (Sigma 30432), and one of the Gln residues of N,N-dimethylated casein (NMC), (Sigma C9801), was assayed by measuring the increase in dansyl fluorescence at 535 nm with an excitation wavelength of 350 nm at room temperature on a Safire^2 microplate reader (Tecan) based on previous methods [96, 97]. Reactions contained 20 nM TG2, 25 mM HEPES, pH 7.4,250 mM NaCl, 2 mM $MgCl_2$, 0.5 mM $CaCl_2$, 1 mM DTT, 0.05% Pluronic F-127, and 20 μM NMC. Each of the 4 casein proteins that make up NMC has a minimum of 14 Gln residues. The DSC concentrations ranged from 0 and 1.5625 to 100 μM. Reactions were initiated upon the addition of enzyme. The rate of reaction versus the corresponding substrate concentration was fit to the Michaelis-Menten equation (equation 3),

$$v = V_{max}[S]/K_m + [S] \quad (3)$$

where v is the initial reaction rate, V_{max} is the maximal rate of reaction, K_m is the Michaelis constant, and [S] is the substrate concentration. The experiment to test if an open-locked design (Y315E) of

TG2 could be active without Ca^{2+} was conducted exactly as above, but without the addition of $CaCl_2$.

Inhibition of transglutaminase activity by GTP

The inhibition of TG2's transglutaminase activity was conducted based on previous published protocols [97, 98]. Reactions contained 20 nM TG2, 25 mM HEPES, pH 7.4, 250 mM NaCl, 2 mM $MgCl_2$, 0.5 mM $CaCl_2$, 1 mM DTT, 0.05% Pluronic F-127, 80 μM N,N-dimethylated casein (NMC), (Sigma C9801), and 100 μM dansylcadaverine (DSC), (Sigma 30432). GTP was added to the reaction mixture at concentrations ranging from 0 to 500 μM. Reactions were initiated upon the addition of enzyme. Reaction progress was monitored at room temperature by measuring the increase in fluorescence intensity using an excitation wavelength of 350 nm and an emission wavelength of 535 nm on a Safire^2 microplate reader (Tecan).

Single turnover kinetic analysis for the GTPase reaction

The GTPase reactions were performed at room temperature with γ-^{32}P-GTP (GTP*) in 50 mM Tris-HCl, pH 7.5, 4 mM $MgCl_2$, 4 mM EDTA, 10% (v/v) glycerol, and 1 mM DTT as in [68]. Single turnover kinetics studies were performed as in [99]. Briefly, reactions were initiated upon the addition of GTP* to a final concentration of 1.68 nM and excess protein ($\geq$ 10 μM). At various time points, aliquots were removed from the reaction and quenched in 0.75 M potassium phosphate, pH 3.3. The extent of the reaction was analyzed by thin-layer chromatography (PEI cellulose F) using 1 M formic acid and 0.5 M lithium chloride as the running buffer. The ratios of product to substrate were then quantified with a phosphorimager (Storm 840; GE Healthcare).

The reaction time courses were fit to equation 4, in which Frac(S) is the fraction of substrate GTP* remaining at each time point, a is the end point fraction of substrate, b is the initial fraction of substrate, and k_{obsd} is the observed rate constant of the reaction [99].

$$\text{Frac(S)} = a + (b - a)\exp(-k_{obsd}t) \quad (4)$$

The reaction time courses were fit to equation 2 assuming an endpoint (a) of 0.001. The k_{obsd} for each reaction was then plotted against the given enzyme concentration and fit to the Michaelis-Menten-like equation 5 to plot k_{obsd}, versus the concentration of TG2 to determine both k_{max}, the maximal rate constant with saturating protein, and $K_{1/2}$, the protein concentration that provides half the maximal rate.

$$k_{obsd} = k_{max} \text{ x } [TG2]/(K_{1/2} + [TG2]) \quad (5)$$

Results and Discussion

Design of open-locked TG2

After several rounds of hinge-focused design, Y315 was selected as the primary design residue for engineering open-locked TG2. This residue is located in the hinge, which forms an α-helix in the open conformation and a β-strand in the closed (Figure 2-3). In addition to being part of the hinge, Y315 was targeted for design because tyrosine strongly favors β-secondary structure over α, and because Y315 forms a hydrogen bond with the backbone of E329 in the closed structure (Figure 2-3, B). Therefore, we chose to force Y315 to an alanine or glutamate residue, both of which strongly favor α over β and would disrupt the hydrogen bond interaction in the closed conformation (Figures 2-4 and 2-5). However, the Y315E mutation was predicted to be forced into a hydrophobic pocket in the open conformation. Therefore, subsequent rounds of design were conducted to accommodate the Y315E mutation in this hydrophobic pocket that surrounds Y315 in the native open conformation, which resulted in the Y315E_L508W and Y315E_L508T designs. The threonine mutation in Y315E_L508T was predicted to make a favorable interaction with the sidechain of Y315E, while L508T's CH_3 group mimicked the native leucine's sidechain facing into the structure (Figure 2-6). The L508W in Y315E_L508W was predicted to accommodate the Y315E mutation by forcing the sidechain of Y315E to hydrogen bond with a native arginine residue (Figure 2-7).

The β2-core interface design resulted in a 9-fold mutant (I178H, N181H, S253K, I255H, G256R, E557T, V665R, N667T, and K677T). The WT residues are solvent-exposed in the open conformation and buried in the closed (Figure 2-8). The polar mutations were predicted to destabilize the closed conformation and favor the open. The initial two (Figure 2-9) and four (Figure 2-10) mutations were made (I178H_N181H and I178H_N181H_S253K_I255H, and respectively) along with the hinge designs were monitored for their conformation using native gel electrophoresis.

WT TG2, when run on a native gel, adopted both the closed (high mobility) and open (low mobility) conformations. Upon the addition of GTP, WT TG2 adopted the closed conformation and the open conformation was not observed. When WT TG2 is run in the presence of calcium, the open conformation is observed, but usually a loss of protein occurs due to self-crosslinking and resulting precipitation. The presence of an irreversible inhibitor in addition to calcium, however, inactivates the enzyme and forces it into the open conformation, resulting in a strong band of the open conformation (Figure 2-11). These results agreed with previous findings [29].

According to our native gels, all six open-locked designs were successful and adopted only the open conformation for all four conditions tested (Figure 2-11). No high-mobility band was observed in any condition for any design. Some designs showed a reduction in the amount of protein visible under the third condition (+ Ca^{2+}), as with WT, indicating the enzyme was still active and was also self-crosslinking. Indeed, upon the addition of calcium ions to some samples, precipitation was observed (data not shown). Open-locked, active site knock out C277A_Y315E was also shown to adopt only the open conformation under all conditions (Figure 2-12).

Design of closed-biased TG2

The design residues selected for engineering a closed-stabilized TG2 (L312, E314, and F316) were all located on the hinge (Figures 2-13 and 2-14) and were distal from the tranglutaminasae active site residues in both conformations (Figures 2-15 and 2-16). The closed stabilized calculations resulted in 23 suggestions that were made and run on native gels (Figure 2-17). Most showed an increase in preference for the closed state in the first condition, where the enzyme is run on the native gel without any additives (see E314T and L312I_E314T, for example). Some showed no large change from WT behavior for all for conditions (L312V_F316V). Others (such as L312Y and F316T) had bands for both the open and closed conformations in the third condition (+ Ca^{2+}, which drives WT to adopt the open conformation). Some designs exhibited an intermediate band (see L312Y condition 3).

The most dramatic changes in conformational preferences were found in designs E314T_F316T, L312V_F316T, and especially L312Y_F316T. L312Y_F316T appears to prefer the closed conformation in conditions one through three, unlike WT. All designs still adopted the open conformation upon inhibition by the covalent inhibitor, however, implying that the binding of the inhibitor forces the adoption of the open conformation. These results led us to conclude that while several of our designs showed a marked preference over WT for the closed conformation, yet were still able to adopt the open conformation when bound to inhibitor, we had designed "closed-biased" variants of TG2 rather than closed-locked.

Transglutaminase activity of the open-locked designs of TG2

All open-locked design showed significantly less activity than WT TG2 (see Figures 2-20, 2-22 and Table 2-1). The most active open-locked design suffered a 70% loss in activity relative to WT TG2. This result was unexpected, given that the open form of the enzyme is the one adopted by the WT enzyme in the presence of calcium ions and was believed to be the transglutaminase active conformation. The expected outcome was that these open-locked designs would be more active than WT TG2 because the initial opening of the enzyme was accomplished and maintained. WT TG2 was determined to have a V_{max} of 1.6 ± 2.5% FI/s. Y315E was the most active open-locked design with a V_{max} of 0.49 ± 8.6% FI/s, and the least active open-locked design had a V_{max} of 0.10 ± 23% FI/s.

Role of calcium ions in transglutaminase activity

The role of calcium in transglutaminase activity is not fully understood. As discussed in the introduction, calcium is necessary for transglutaminase activity and causes the enzyme to adopt the open conformation. But what effect of Ca^{2+} binding is making it necessary for transglutaminase activity? Is driving the adoption of the open conformation calcium's only role in achieving transglutaminase activity? And, therefore, would an open-locked version of TG2 still need calcium in order to be active? Or, is calcium not only involved in large-scale conformational regulation, but also plays some unknown role in catalysis? In order to investigate these questions, we tested the transglutaminase activity of the most active open-locked design, Y315E, in the absence of calcium ions. If calcium's only role in transglutaminase activity is to force the adoption of the open conformation, and this conformation is active in and of itself, Y315E should have been active despite the absence of calcium ions. However, under these conditions, Y315E without calcium had no observable transglutaminase activity (Figure 2-21). Therefore, the Ca^{2+} dependence of TG2's transglutaminase activity arises from more than the calcium-induced large-scale conformational opening of TG2.

What Ca^{2+} is doing to promote activity besides causing the enzyme to adopt the open form is unclear. Perhaps smaller conformational changes induced by the binding of these ions, as seen in TG3, where loops surrounding bound Ca^{2+} exhibit small conformational movements, also occur in TG2, and these small changes induce activity. The oxidation of TG2, which is inactivating, also causes the adoption of the open conformation, and the mechanism of this inactivation is not fully

understood. It is clear, however, that merely being "open" is not sufficient for TG2 to exhibit transglutaminse activity.

Transglutaminase activity of the closed-biased designs of TG2

Because the adoption of the closed conformation was believed to be the mechanism by which GTP inhibits TG2's transglutaminase activity, we expected that the closed-biased designs would be significantly less active than WT and open-locked TG2. However, all of the closed-biased designs assayed for transglutaminase activity were more active than any of the open-locked designs, except closed-biased E314T, which was less active than 4 open-locked designs (see Figure 2-22 and Table 2-1). In addition, four of the closed-biased designs were more active than WT TG2: L312Y, L312V, L312Y_F316T (the "most closed"), and F316T. Of all of the designs, open and closed, only L312I_E314T had a significant change in apparent K_m, although the error in this value was large.

The increase in TG2 activity of some closed-biased designs was a surprising result, and came on the heels of the discovery that the open-locked TG2s exhibited decreased transglutaminase activity. Taken together, these results were the opposite of what was expected, causing us to look back at other confusing results in regards to TG2 transglutaminase activity.

As mentioned in Chapter 1 and this chapter's introduction, several studies have cast doubt on the current model of transglutaminase activity. Two truncation studies showed that deletion of the C-terminal β-barrels resulted in a loss in transglutaminase activity despite these domains being very distal from the transglutaminase active site in the assumed active open conformation [68, 69]. Both factor XIII and TG3 are able to catalyze the transglutaminase reaction albeit more specifically than TG2, despite being shown to adopt the closed conformation in the presence of calcium [38, 57, 58] [55, 59]. The active site residues of TG2 adopt largely the same conformation in both the open sand closed conformations (see Figures 2-23 and 2-24), and TG2's active site residues in the closed conformation also align well to the active site residues of TG3 in the active conformation bound to calcium (Figure 2-25) [29, 30, 53]. Additionally, celiac patients' autoantibodies against TG2 have been shown to increase transglutaminase activity of TG2 and bind a motif that is complete only in the closed conformation [81, 83, 87, 88]. These results led the authors to propose that TG2 might be more active as a transglutaminse in the closed conformation.

Our results support this hypothesis and help to clarify the previously mentioned observations of the apparent need for the β-barrel domains and those results that called into question the necessity of the open conformation for transglutaminase activity. However, since all designs still adopted the open conformation upon inhibitor binding, it may be possible that this open conformation is necessary for activity in that it helps to accommodate substrates, and that this explains the relative promiscuity of TG2 for glutamine substrates compared to its less open relations TG3 and factor XIII. Perhaps an equilibrium between the two structures exist during catalysis, and the enzyme cycles through the conformations during the various steps of the transglutaminase reaction. It is also possible that a shift towards favoring the closed conformation helps the enzyme move through a rate-limiting step. Single molecule experiments to monitor the motion of the enzyme during catalysis and a deamidase assay as in [87] would further our understanding of the role of each conformation in the mechanism of the TG2-catalyzed transglutaminase reaction.

Inhibition of transglutaminase activity by GTP

No large differences in the inhibition of transglutaminase activity by GTP were observed between WT TG2 and closed-biased L312Y when analyzed by the percent of the variant's own activity lost (Figure 2-26) or by the percent of WT activity lost (Figure 2-27). However, open-locked Y315E_L508W lost less of a percent of its own activity, relative to WT (Figure 2-26) and was more active than WT at high GTP concentrations (Figure 2-28), implying that it was less susceptible to inhibition by GTP.

These results were expected, in the sense that closed-biased L312Y and WT TG2 were not expected to behave too differently given they adopt the same conformation in the presence of GTP. Also, open-locked Y315E_L508W was expected to be less susceptible to GTP inhibition given its active site is disrupted upon the adoption of the open conformation. However, Y315E_L508W was still susceptible to inhibition by GTP to the point of retaining only approximately 20% of its uninhibited activity at 500 μM GTP; the same concentration of GTP used in the native gels in condition two where Y315E_L508W still adopted the open conformation (see Figure 2-11).

The mechanism by which GTP was thought to inhibit transglutaminase activity was by forcing the adoption of the closed conformation, thereby obscuring access to the transglutaminase active site. The susceptibility of open-locked TG2 to inhibition by GTP despite it being shown that the enzyme

remains open at high GTP concentrations and the fact that closed-biased designs showed at times increased transglutaminase activity relative to WT call this theory into question, and suggest that some alternative mechanism of inhibtion may be at work. One study suggested that a set of conserved tryptophan residues in the core of the protein were important in GTP inhibition of transglutaminase activity, and may be involved in "allosteric relay machinery" that propagates inhibitory effects from GTP binding to the transglutaminase active site [100]. Repeating the study detailed in the above paper with open-locked TG2 to investigate if these tryptophan residues would still be involved in GTP inhibition of transglutaminase activity would be enlightening.

Single turnover kinetic analysis of the GTPase of WT, open-locked, and closed-biased TG2

Single turnover kinetics, where enzyme is saturating, were used to analyze the GTPase activity of WT TG2 and our designs because it enabled us to avoid complications arising from rate limiting product dissociation. Due to time constraints, the determination of conditions for the WT assay and an initial assay of WT plus open-locked Y315E_L508W, closed-biased and most active L312Y, and most closed closed-biased, and more active than WT protein L312Y_F316T were the only experiments accomplished. Therefore, this assay needs to be further optimized and repeated.

However, initial results were dramatic and unexpected (Figure 2-29 and Table 2-2). GTPase activity increased with increasing preference for the open conformation, with open-locked Y315E_L508W being the most active GTPase and most closed closed-biased L312Y_F316T being the least active. As detailed in Chapter 1, the binding of GTP induces the adoption of the closed conformation, and the residues which make up the GTPase active site in the crystal structure of closed TG2 involve residues in the core and β-barrel 1 domain, which are torn appart in the open structure. Therefore, we expected that the closed-biased TG2s would be more active as a GTPase than WT, and that the open-locked TG2 would be the least active. These results indicate the exact opposite.

This result corroborates the results of the two truncation studies mentioned in the introduction, where the core of the protein minus the β-barrel domains was found to be more active as a GTPase than full length protein [68, 69]. These results suggest that the closed form of TG2 may do more to regulate GTPase activity than promote it, and call into question the current understanding of the

GTPase active site. A crystal structure of open-locked TG2 bound to GDP would be useful in furthering our understanding of the mechanism of TG2 as a GTPase.

Conclusions

In summary, we were successful in designing open-locked and closed-biased variants of TG2 using computational protein design. These designs allowed us to characterize the activities of these conformations, and resulted in conclusions about the behavior of TG2 opposite to the current understanding, but which clarified and supported previous studies. The existing model for TG2 conformational behavior and catalytic activity was that the open form of TG2, adopted upon calcium binding, catalyzed only the transglutaminase reaction and the closed form, adopted upon the binding of GTP, catalyzed only the GTPase reaction. Our designs allowed us to determine three previously unknown characteristics of TG2: (1) the open conformation of TG2 is insufficient for transglutaminase activity, and therefore that the Ca^{2+}-dependence of this reaction is derived from more than driving conformational change, (2) the closed-biased designs of TG2 have higher transglutaminase activity than the open-locked designs, and in some cases, WT, and (3) not only can both conformations act as a GTPase, but the open-locked TG2 is significantly more active than the WT or closed-biased designs.

These results have immense implications for the understanding of TG2's behavior in biology and the targeting of this enzyme for disease therapy. These results also explain previously confusing truncation studies, help explain the various substrate specificities among the transglutaminase family, and lend credence to the idea that celiac anti-TG2 antibodies could be increasing transglutaminase activity by stabilizing the closed conformation, and therefore this conformation should be used as a model for designing small molecule transglutaminse inhibitors for drug development.

Acknowledgements

We thank Heidi Privett and Chang of Protabit for constant support and information about using TRIAD, much of it when the program was very new. Special thanks goes to David Akopian of the Shan lab at Caltech who provided training, kinetic analysis advice and experimental design, and shared his bench space for the radiometric GTPase experiments.

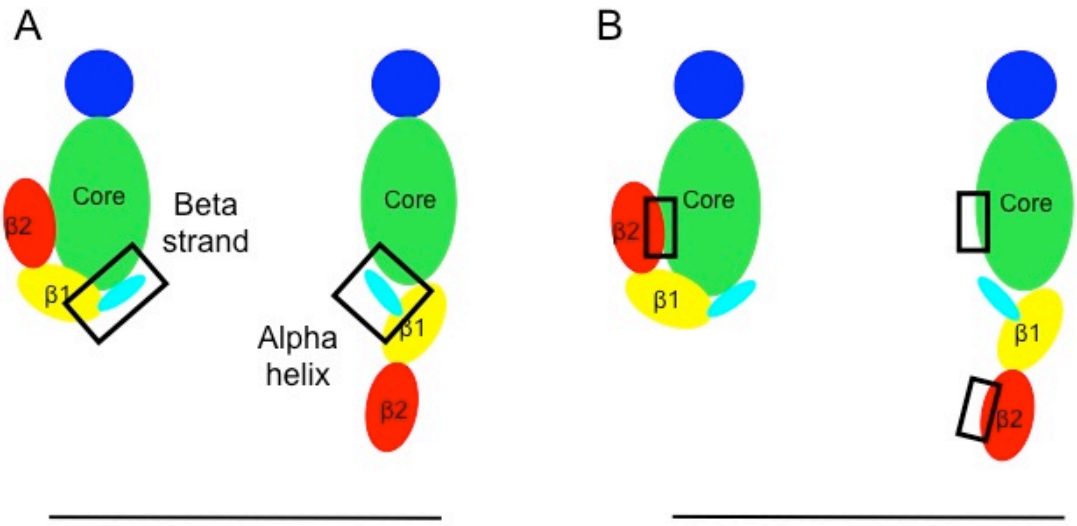


Figure 2-1. Design regions of TG2. The domains of TG2 are illustrated in cartoon form for the open and closed conformations (right and left respectively). The two regions to be targeted as design regions are (A) the β2-core interface, and (B) the hinge.

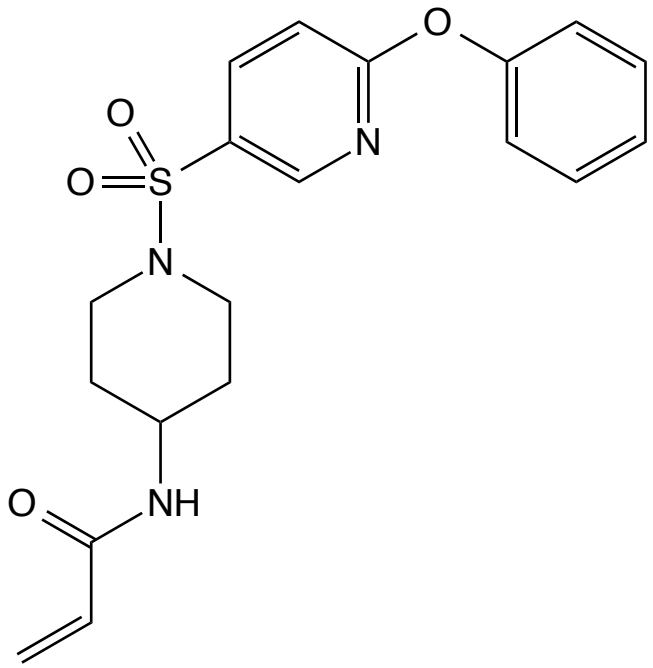


Figure 2-2. Irreversible transglutaminase inhibitor of TG2 provided by the CHDI.

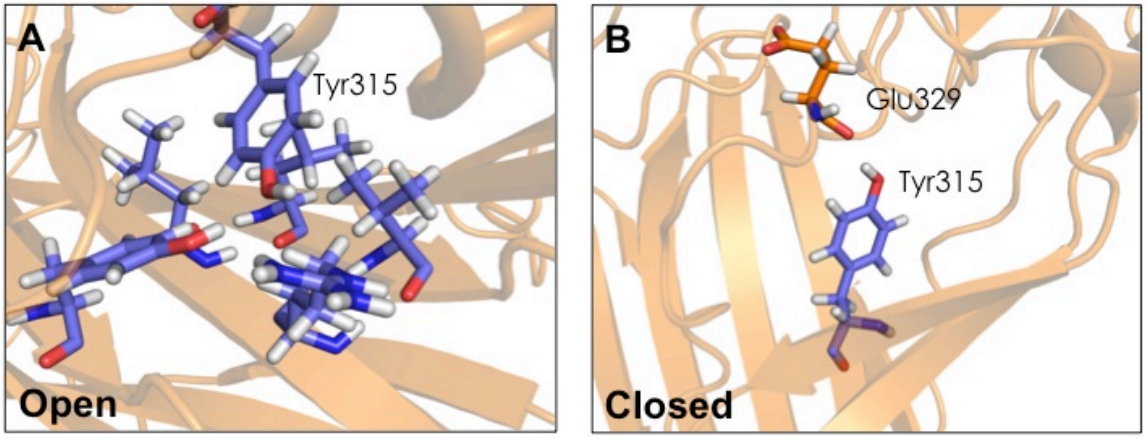


Figure 2-3. Open-locked design hinge residue Tyr315 for in the WT crystal structures. (A) Y315 in the open form of TG2 is nestled in a pocket on the β1-barrel [29]. (B) In closed TG2, Y315 forms a hydrogen bond with the backbone of E329 [30].

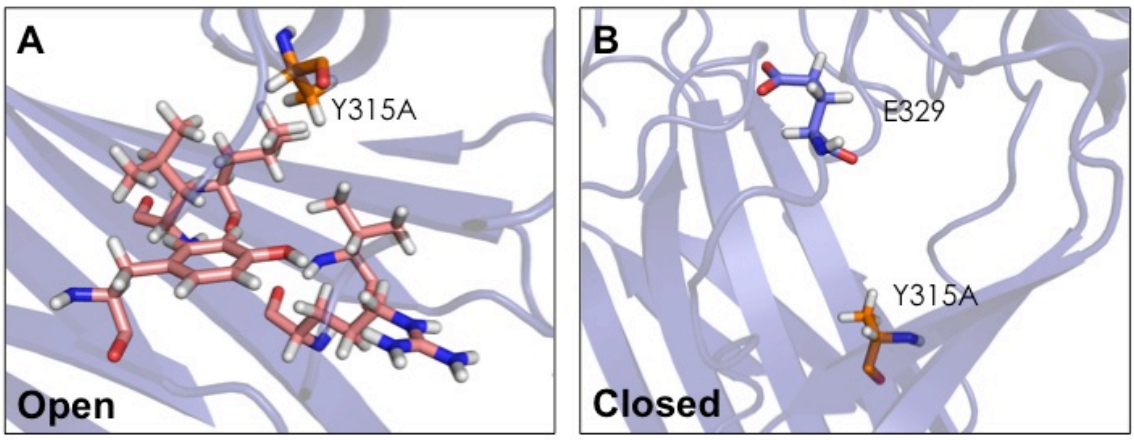


Figure 2-4. TRIAD predicted structure for open-locked design Y315A. (A) In the open structure, the alanine mutation does not fill the pocket normally occupied by the Y315 sidechain. (B) In the closed structure, alanine is unable to form the hydrogen bond and leaves a whole in the closed form of the enzyme.

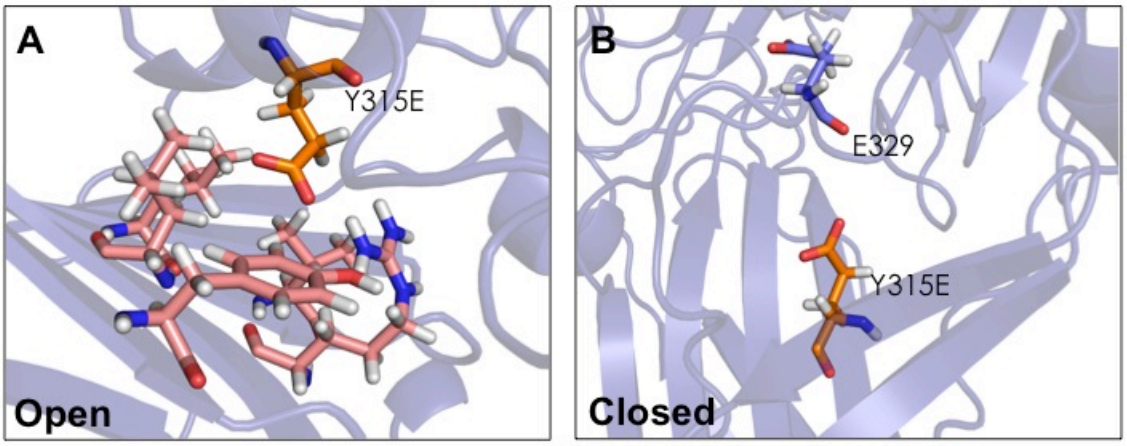


Figure 2-5. TRIAD predicted structure for open-locked design Y315E. (A) In the open structure, the glutamine residue is forced into the hydrophobic pocket. (B) In the closed structure, the sidechain of Y315E makes an unfavorable interaction with the backbone of E329.

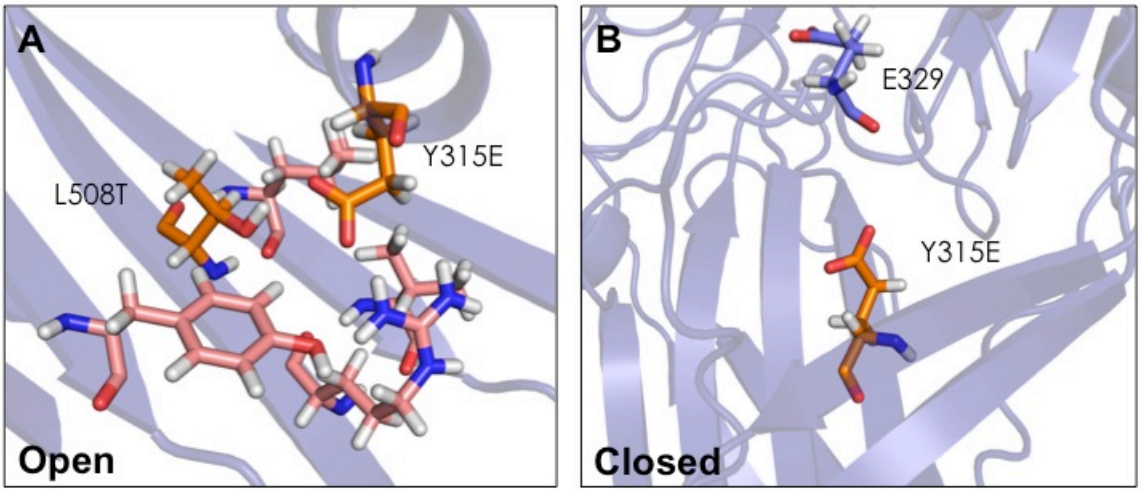


Figure 2-6. TRIAD predicted structure for open-locked design Y315E_L508T. (A) In the open form, the L508T mutation is predicted to make a favorable interaction with the Y315E sidechain to help accommodate it in the pocket. The CH_3 group of L508T mimics the WT L508 sidechain. A WT R506 residue is also predicted to flip inwards toward Y315E. (B) In the closed structure, the sidechain of Y315E makes an unfavorable interaction with the backbone of E329.

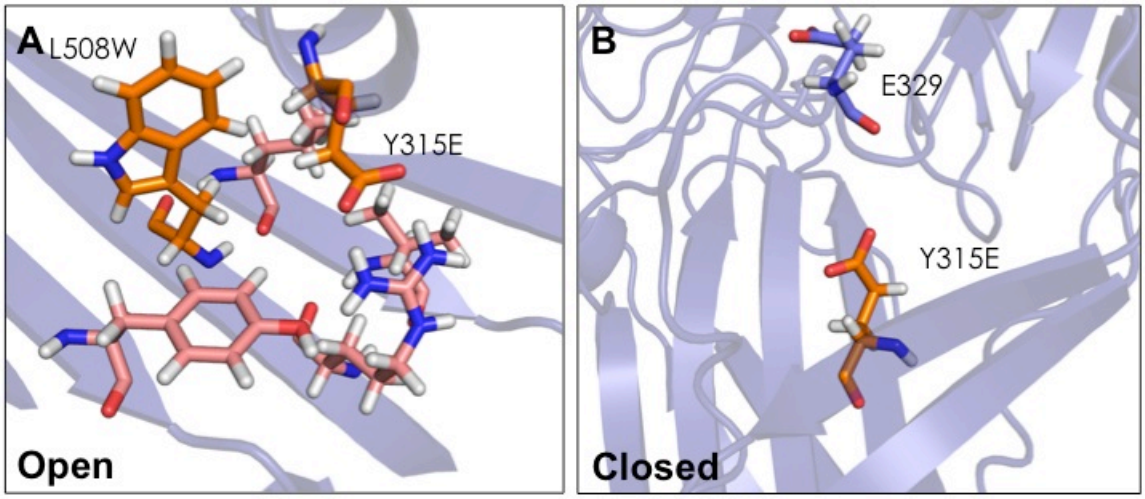


Figure 2-7. TRIAD predicted structure for open-locked design Y315E_L508W. (A) In the open form, L508W is predicted to force the sidechain of Y315E to interact with R506. (B) In the closed structure, the sidechain of Y315E makes an unfavorable interaction with the backbone of E329 .

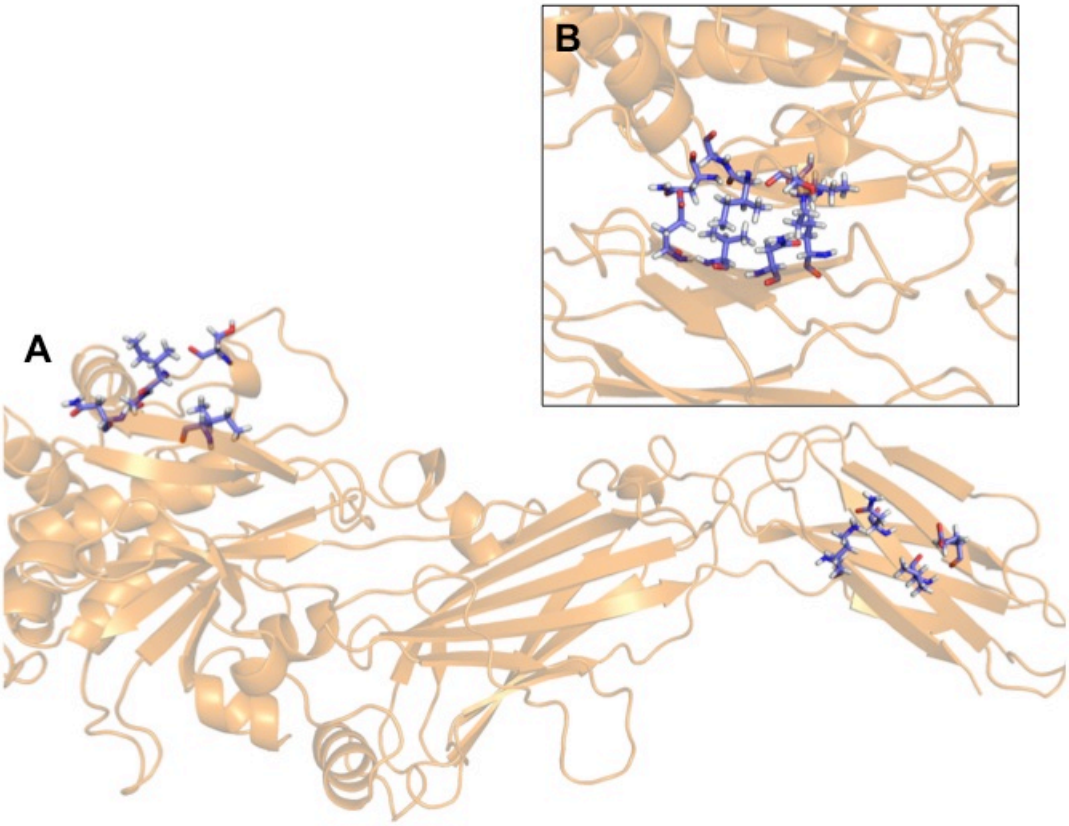


Figure 2-8. Open-locked WT design residues for the β2-core interface in the WT structure. (A) In the open conformation, the design residues are solvent exposed on both the core of the protein and on β barrel 2 [29]. (B) In the closed conformation, these residues are buried [30].

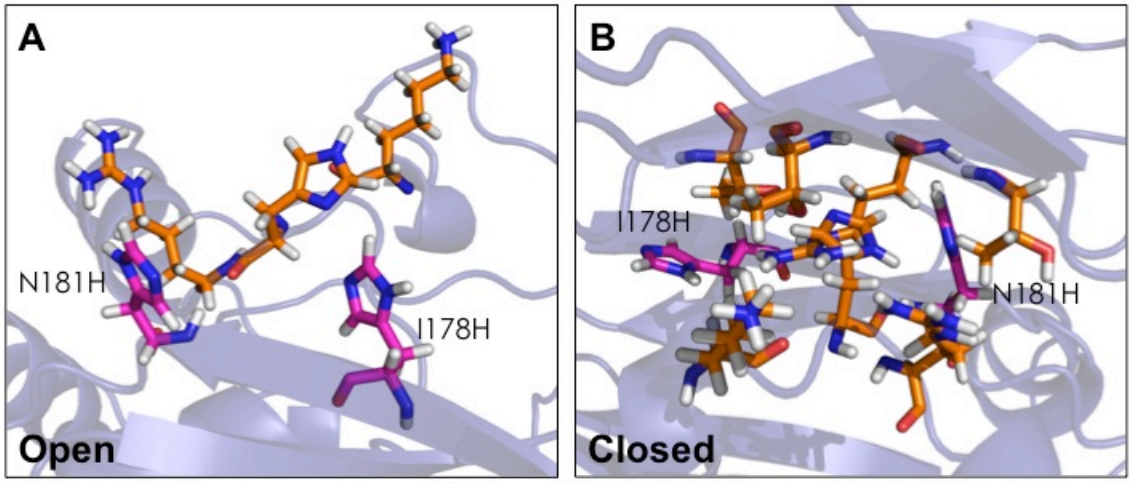


Figure 2-9. TRIAD prediction of open-locked core design residues I178H and N181H for the β2-core interface in the open conformation (A) and the closed conformation (B).

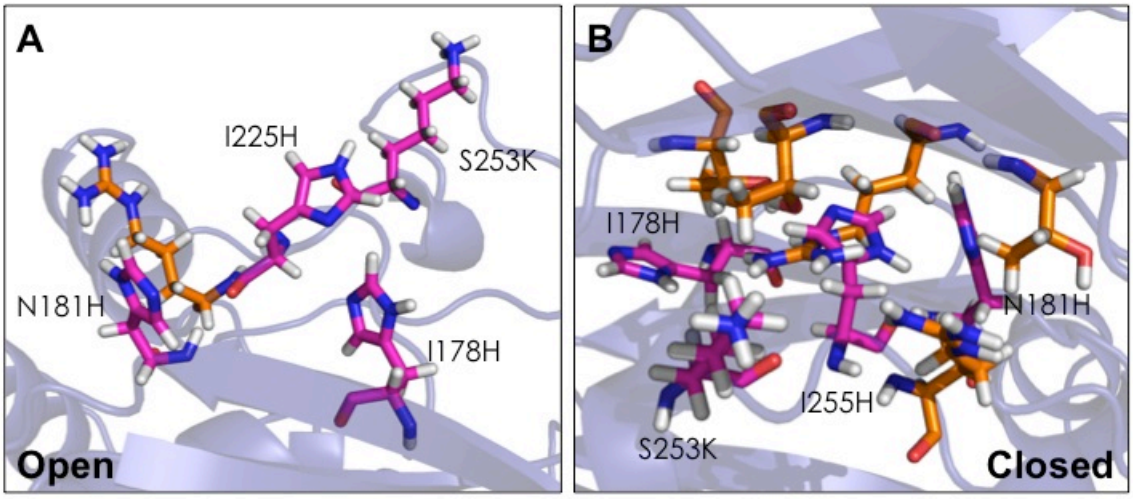


Figure 2-10. TRIAD prediction of open-locked core design residues I178H, N181H, S253K, and I255H for the β2-core interface in the open conformation (A) and the closed conformation (B).

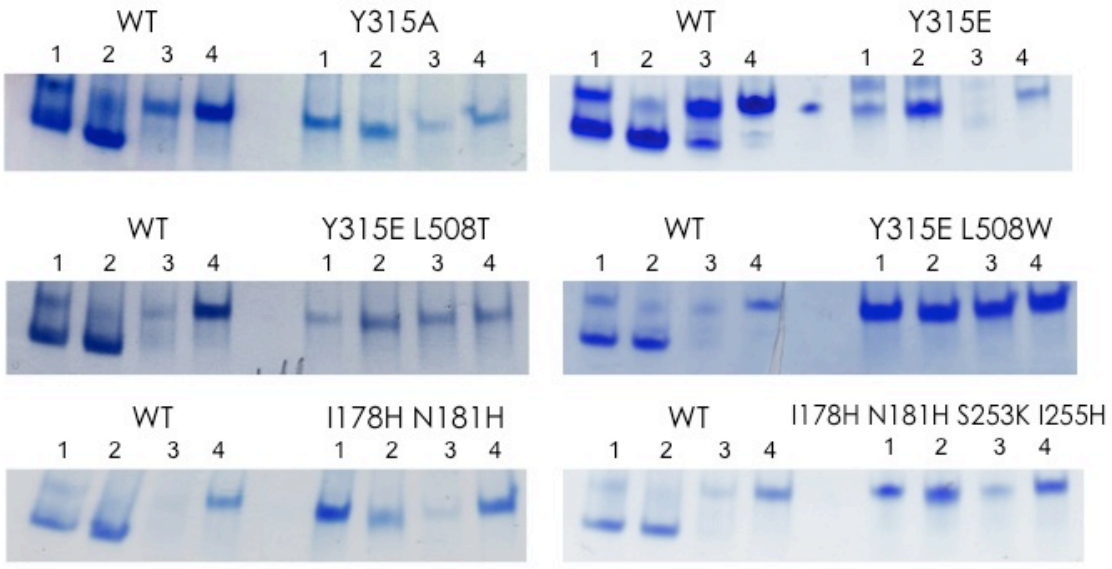


Figure 2-11. Native gels of open-locked designs showing that all designs adopt the open conformation under all conditions. Set of four conditions on the left of each gel are for the WT protein, and the right side shows the results for the indicated design. The conditions are as follows: (1) purified WT TG2 or design only, (2) protein with 500 µM GTP and 1 mM $MGCl_2$, (3) protein with 5 mM $CaCl_2$, and (4) protein with 5 mM $CaCl_2$ and 25 µM irreversible peptide inhibitor.

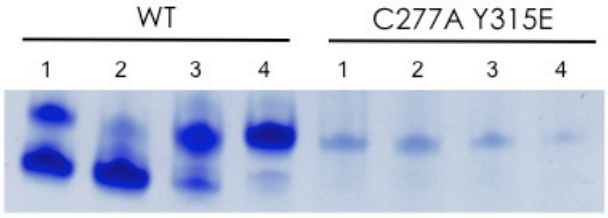


Figure 2-12. Native gel of WT and open-locked and active site knock-out C277A_Y315E. The conditions are as follows: (1) purified WT TG2 or design only, (2) protein with 500 µM GTP and 1 mM $MGCl_2$, (3) protein with 5 mM $CaCl_2$, and (4) protein with 5 mM $CaCl_2$ and 25 µM irreversible peptide inhibitor.

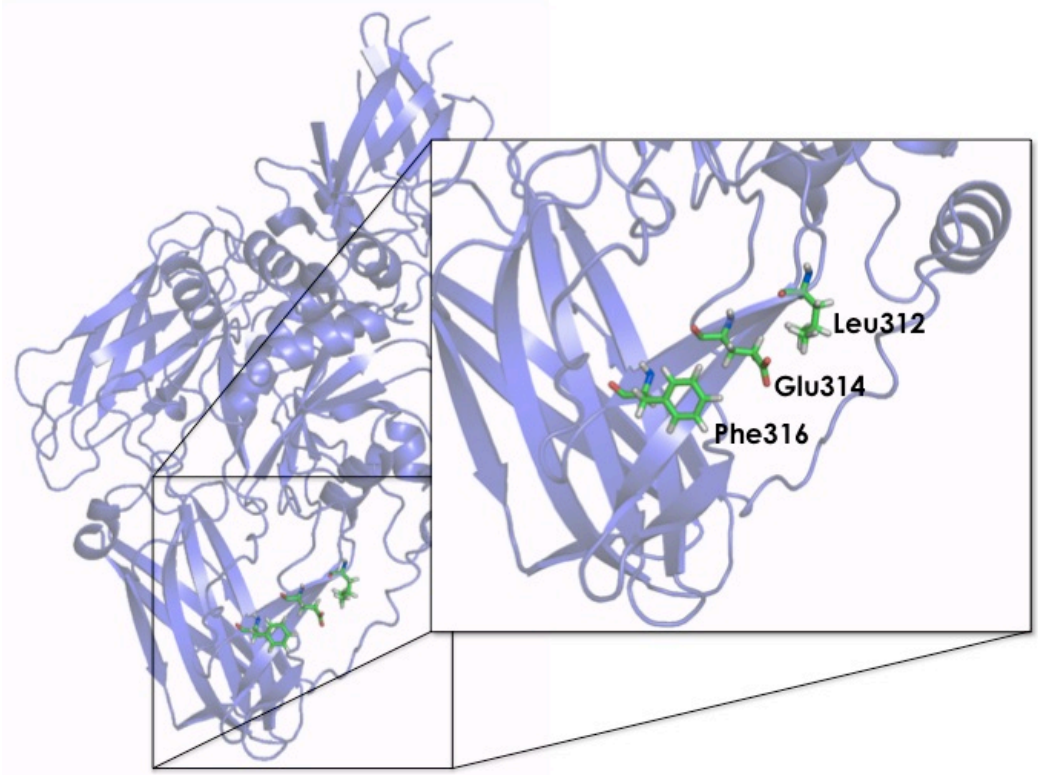


Figure 2-13. Closed-locked WT design residues in the closed form of TG2: L312, E314, and F316 [30].

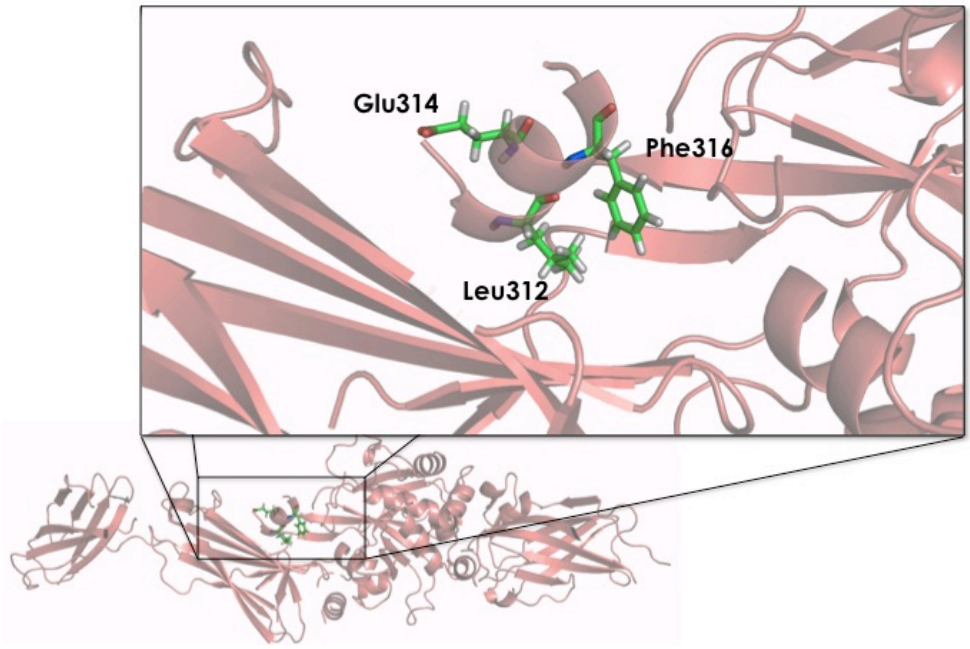


Figure 2-14. Closed-locked WT design residues in the open form of TG2 [29].

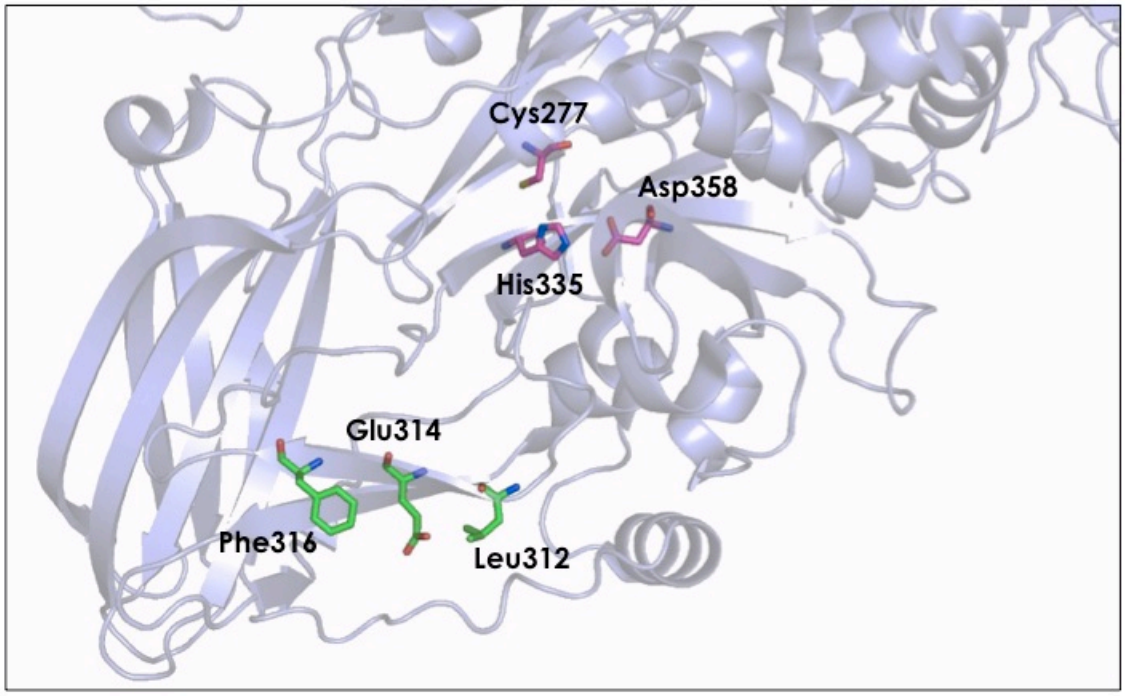


Figure 2-15. Closed-locked WT design residues in the closed conformation of TG2 are located distally from the transglutaminase active site [30].

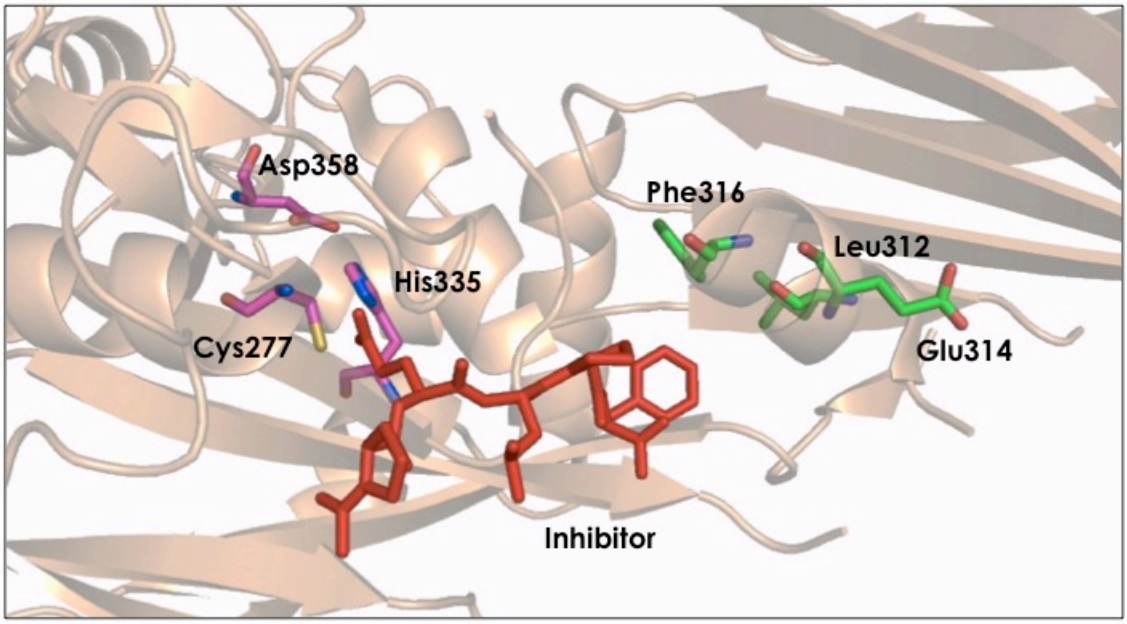


Figure 2-16. Closed-locked WT design residues in the open conformation of TG2 are located on the hinge just below the end of the inhibitor on the opposite side of the hinge [29].

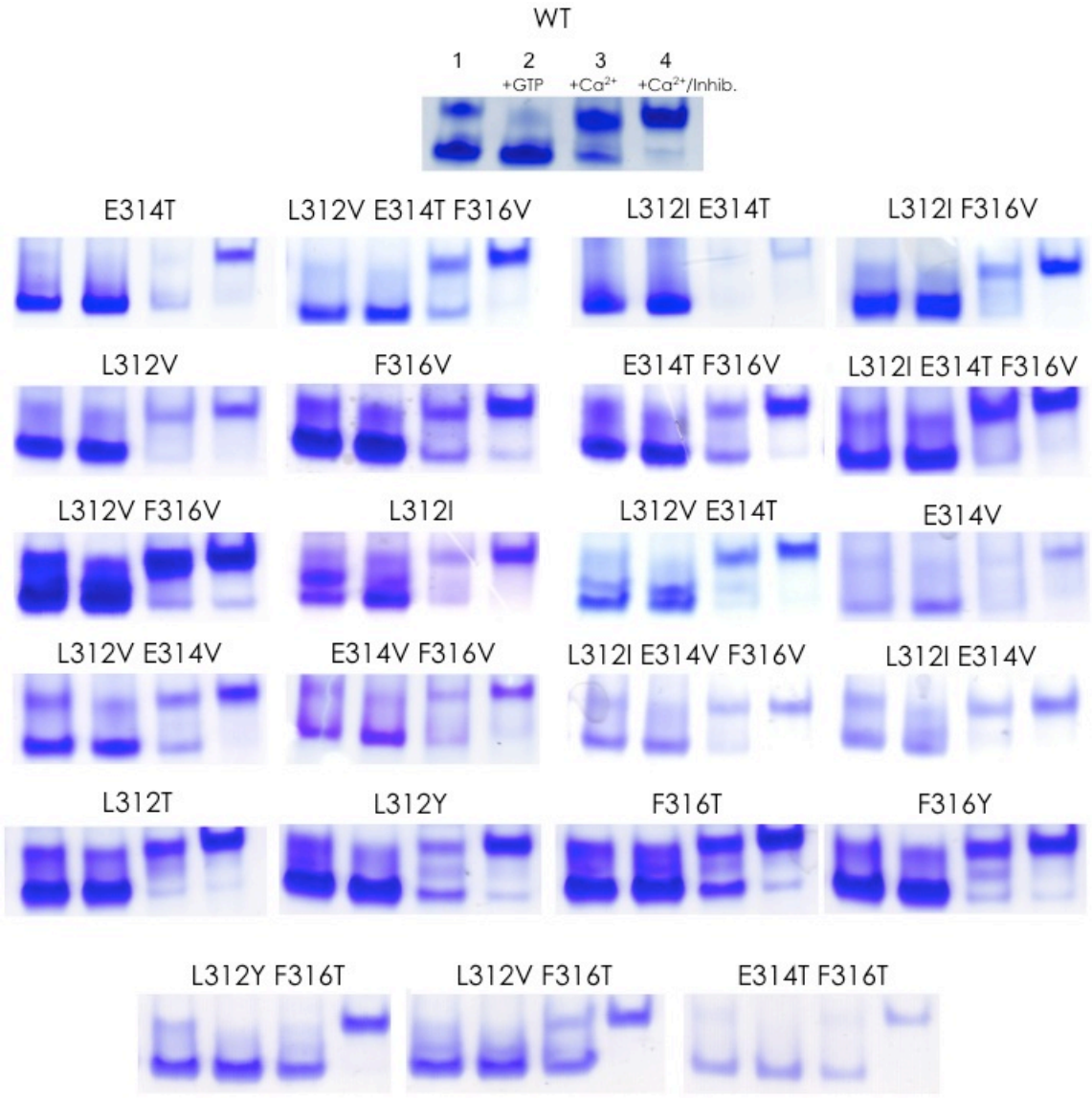


Figure 2-17. Native gels of closed-biased designs. Conditions 1 through 4 are the same for each indicated TG2 variant. The conditions are as follows: (1) purified WT TG2 or design only, (2) protein with 500 μM GTP and 1 mM $MGCl_2$, (3) protein with 5 mM $CaCl_2$, and (4) protein with 5 mM $CaCl_2$ and 25 μM irreversible peptide inhibitor.

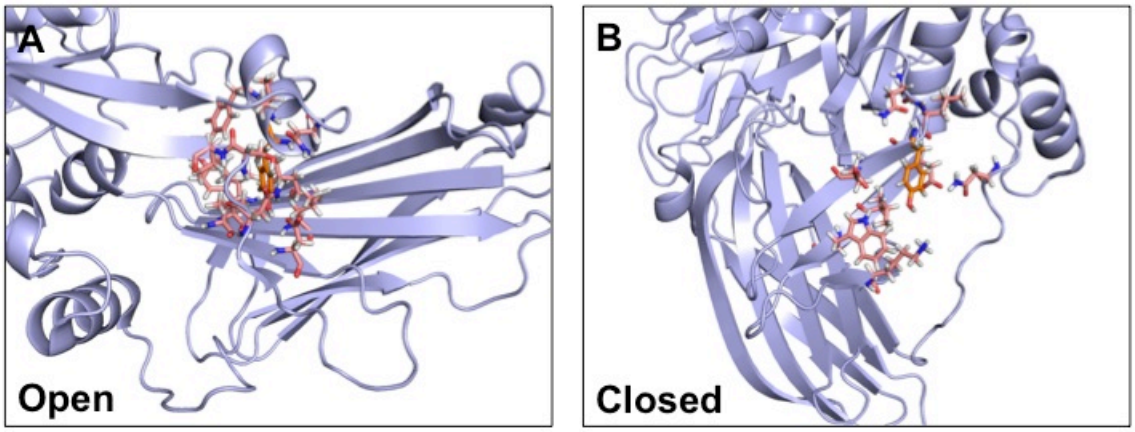


Figure 2-18. TRIAD prediction of closed-biased hinge design residue L312Y in orange in the open conformation (A) and the closed conformation (B).

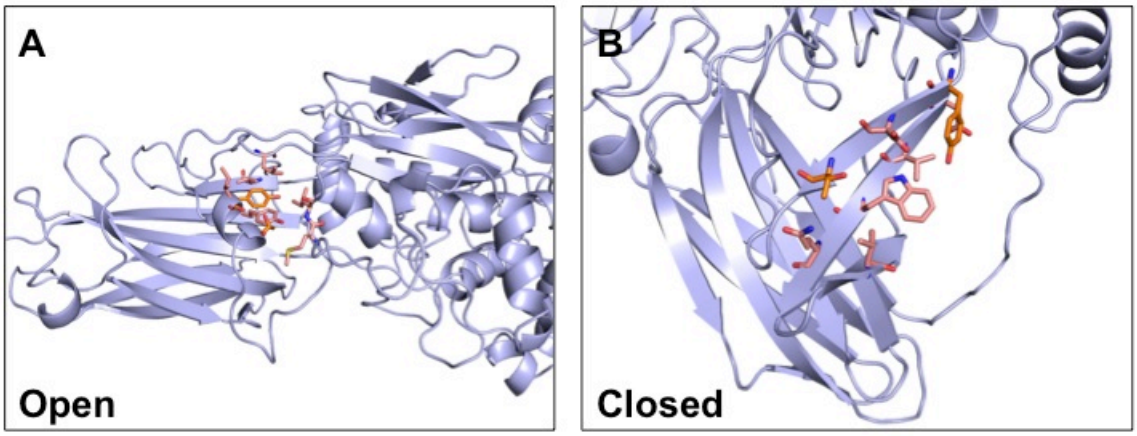


Figure 2-19. TRIAD prediction of closed-biased hinge design residues L312Y_F316T in orange in the open conformation (A) and the closed conformation (B).

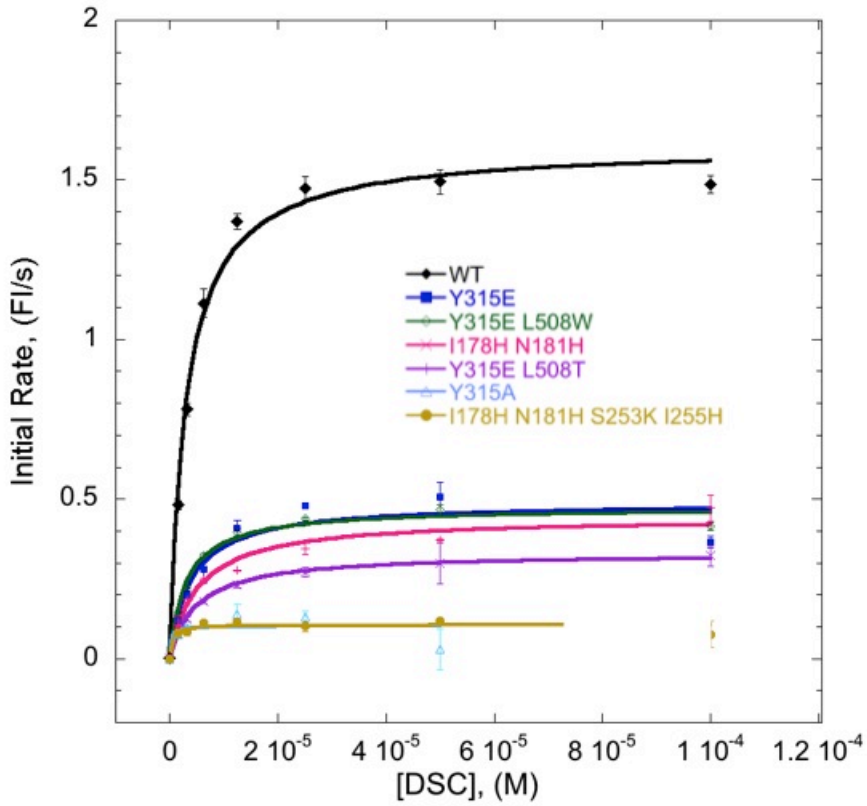


Figure 2-20. Transglutaminase activity assay of TG2 catalyzed incorporation of dansylcadaverine into N,N-dimethylated casein for WT and open-locked designs. WT protein had much greater activity than any of the open-locked designs.

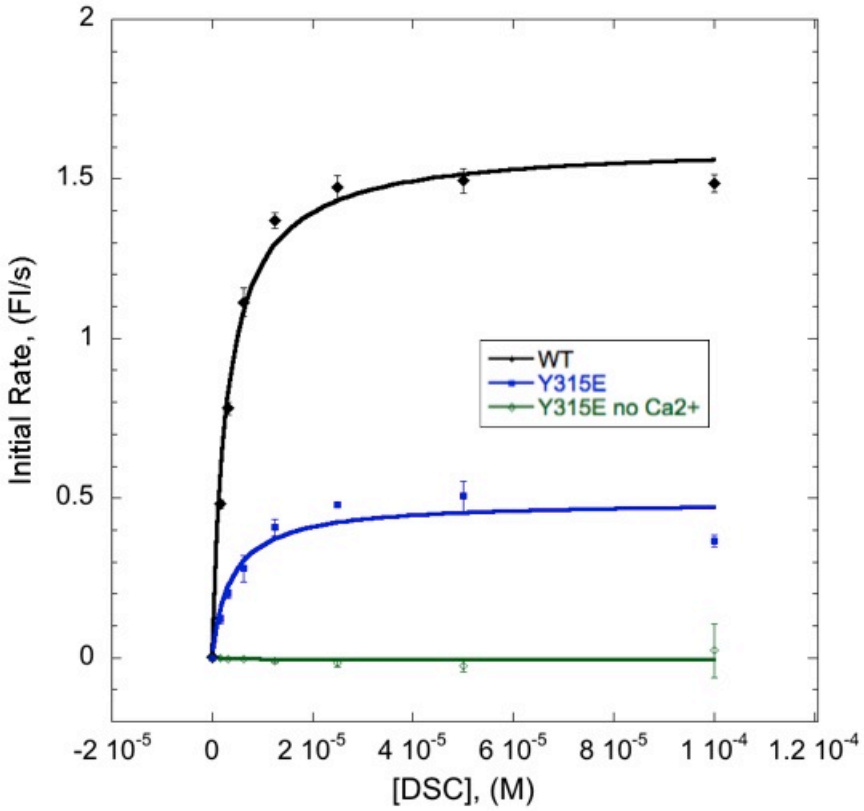


Figure 2-21. Transglutaminase activity assay of TG2 catalyzed incorporation of dansylcadaverine into N,N-dimethylated casein for WT and open-locked design Y315E with and without calcium. The adoption of the open conformation alone is not sufficient for transglutaminase activity.

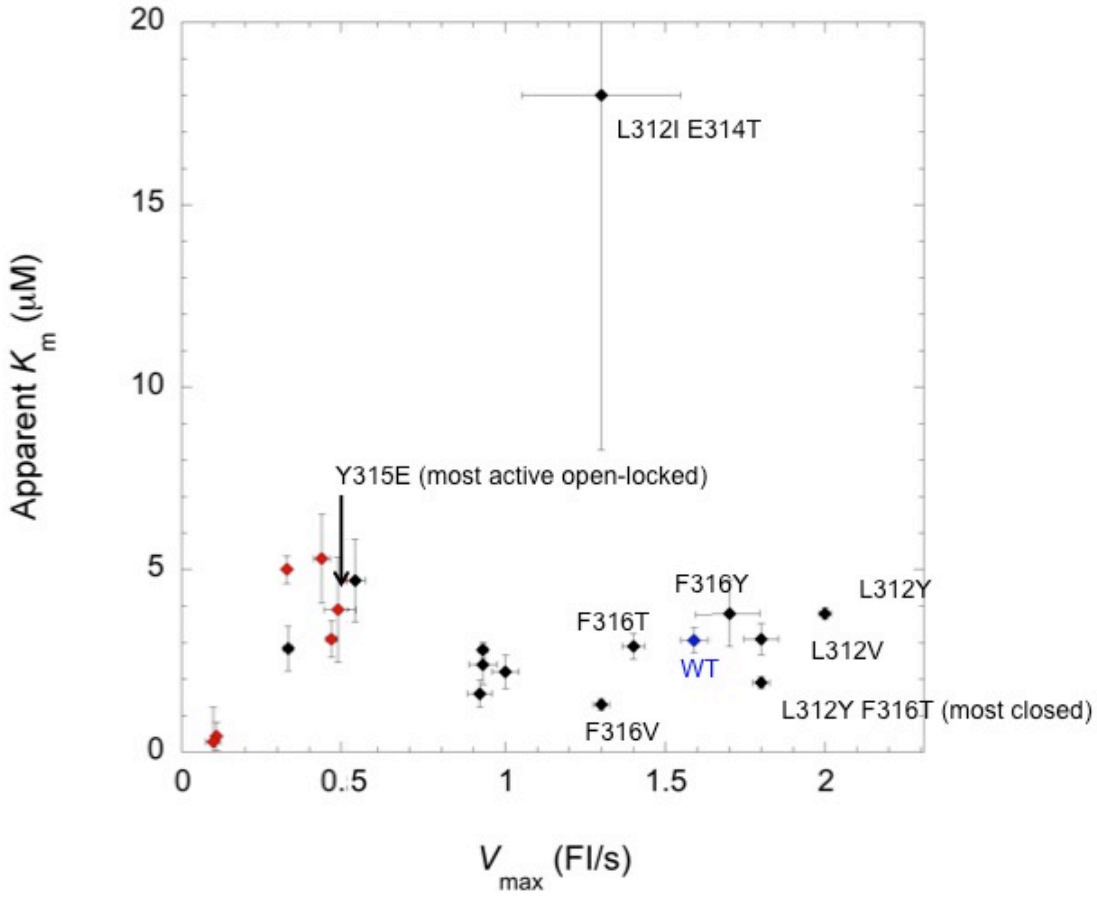


Figure 2-22. Kinetic data for the transglutaminase activity assay of TG2 catalyzed incorporation of dansylcadaverine into N,N-dimethylated casein for WT and all designs. Open-locked residues are in red, closed-biased in black, and WT in blue.

Table 2-1. Kinetic data for the transglutaminase activity assay of TG2 catalyzed incorporation of dansylcadaverine into N,N-dimethylated casein for WT and all designs, in the order of decreasing activity.

TG2 Variant	V_{max} (FI/s)	Apparent K_m (μM)
L312Y	2.0 ± 0.98%	3.8 ± 4.2%
L312V	1.8 ± 3.0%	3.1 ± 14%
L312Y F316T	1.8 ± 1.5%	1.9 ± 7.9%
F316Y	1.7 ± 5.6%	3.8 ± 24%
WT	1.6 ± 2.5%	3.1 ± 11%
F316T	1.4 ± 2.5%	2.9 ± 12%
L312I E314T	1.3 ± 19%	18 ± 54%
F316V	1.3 ± 1.9%	1.3 ± 12%
E314T F316V	1.0 ± 4.2%	2.2 ± 21%
L312I E314T F316V	0.93 ± 4.7%	2.4 ± 23%
L312I F316V	0.93 ± 1.6%	2.8 ± 7.6%
L312V F316T	0.92 ± 4.1%	1.6 ± 23%
L312V E314T F316V	0.53 ± 6.0%	4.7 ± 24%
Y315E	0.49 ± 8.6%	3.9 ± 37%
Y315E L508W	0.47 ± 3.5%	3.1 ± 16%
I178H N181H	0.44 ± 5.9%	5.3 ± 23%
Y315E L508T	0.33 ± 2.0%	5.0 ± 7.6%
E314T	0.32 ± 4.8%	2.7 ± 23%
I178H N181H S253K I255H	0.11 ± 7.8%	0.43 ± 91%
Y315A	0.10 ± 23%	0.28 ± 340%

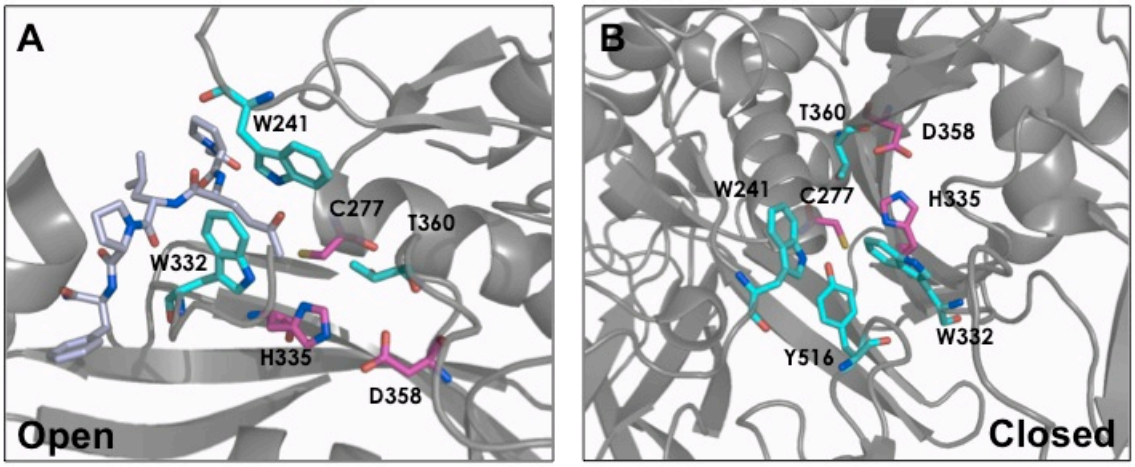


Figure 2-23. The active site residues of the transglutaminase site of TG2 assume similar conformations in both the open and closed forms of the enzyme [29, 30].

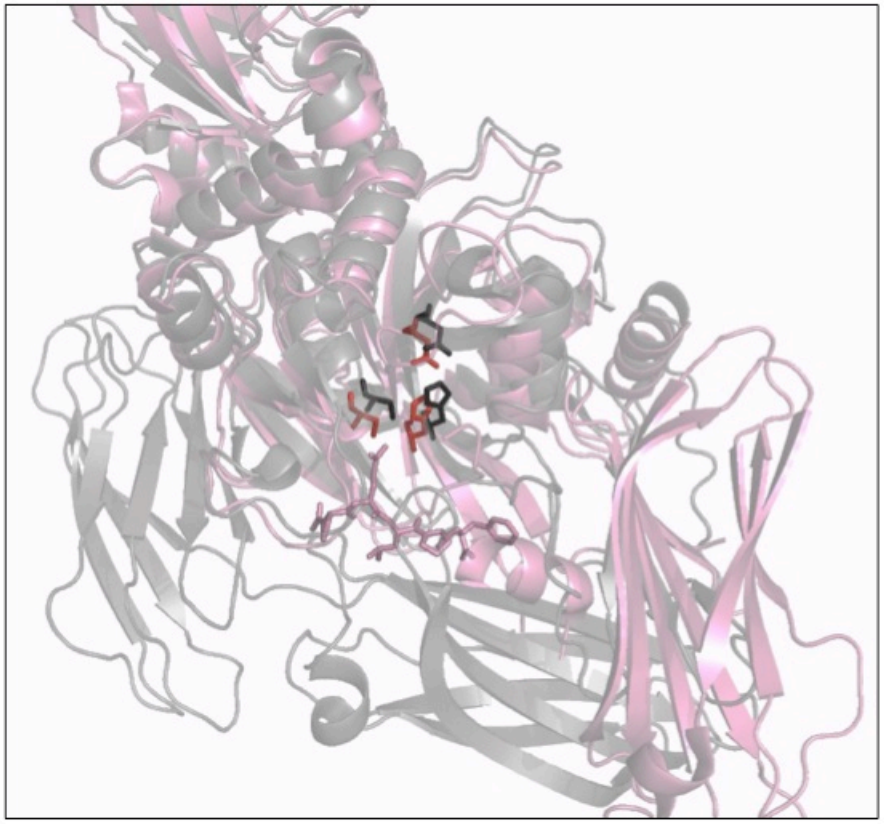


Figure 2-24. The catalytic triad overlays well in both the open and closed conformations. The open conformation is shown in pink with its catalytic triad in red. The closed conformation is shown in gray and its catalytic triad is shown in black [29, 30].

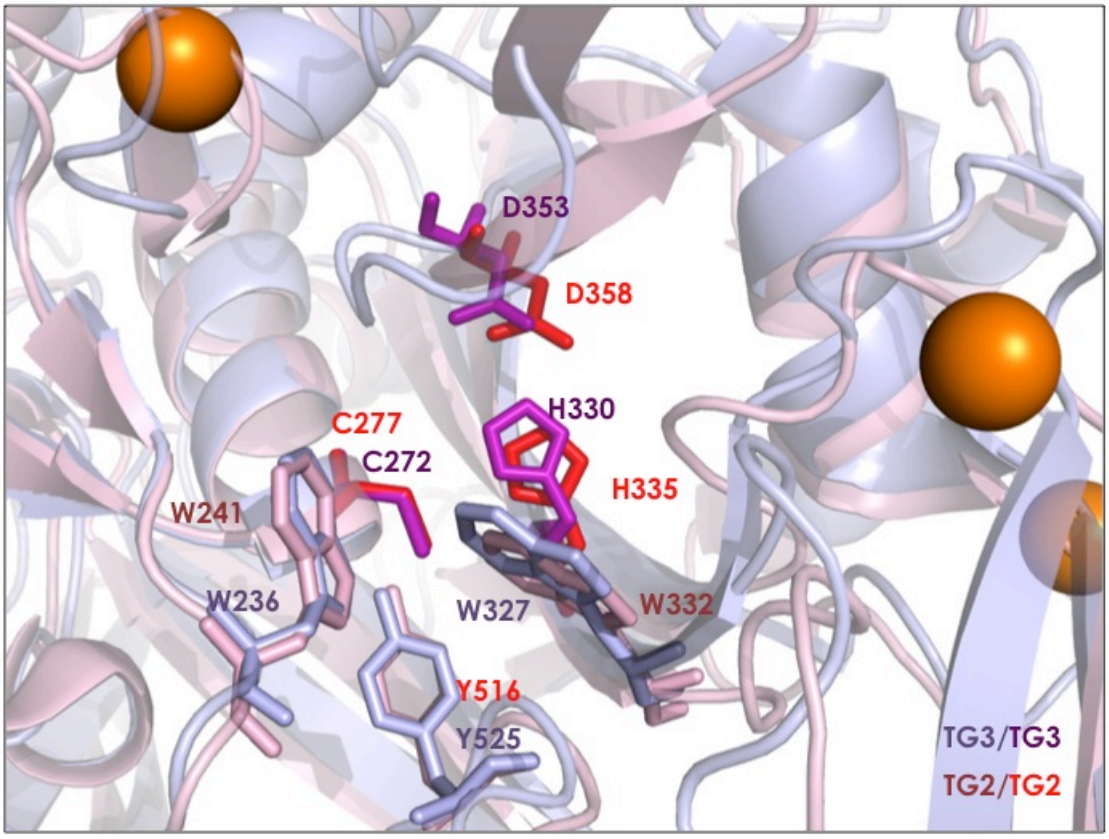


Figure 2-25. Overlay of the transglutaminase active sites of the closed form of TG2 (pink protein, red active site residues) and the closed, active, and calcium-bound form of TG3 (purple protein, magenta active site residues) [30, 53].

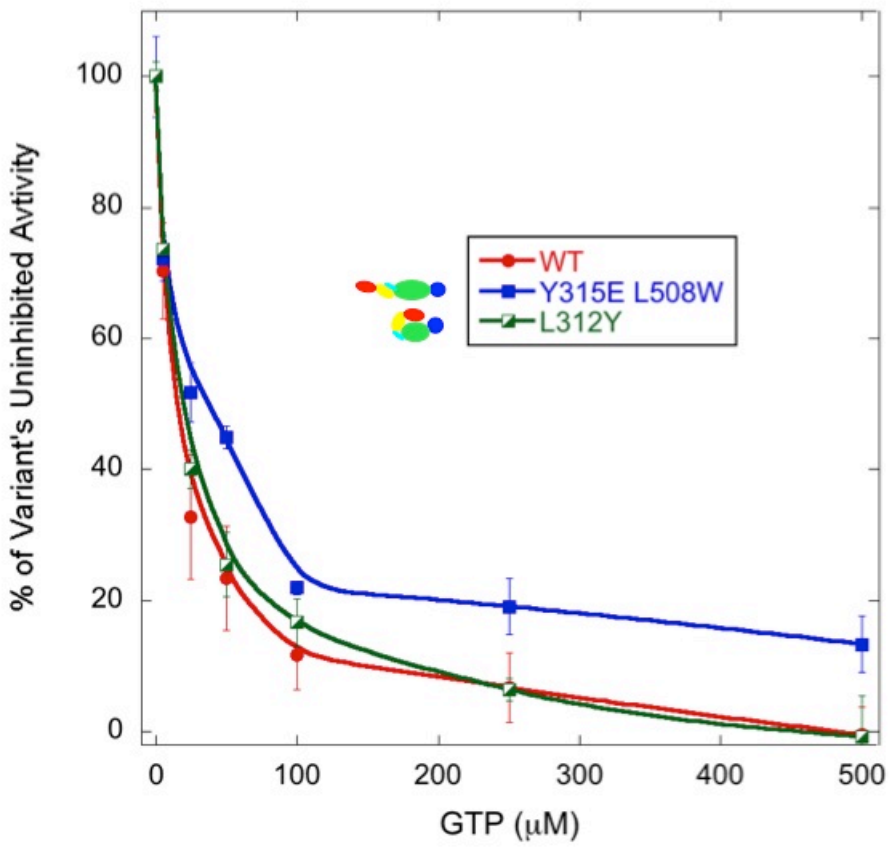


Figure 2-26. Inhibition of transglutaminase activity by GTP for WT, open-locked Y315E_L508W, and closed-biased L312Y as shown as percent of each variants own uninhibited activity.

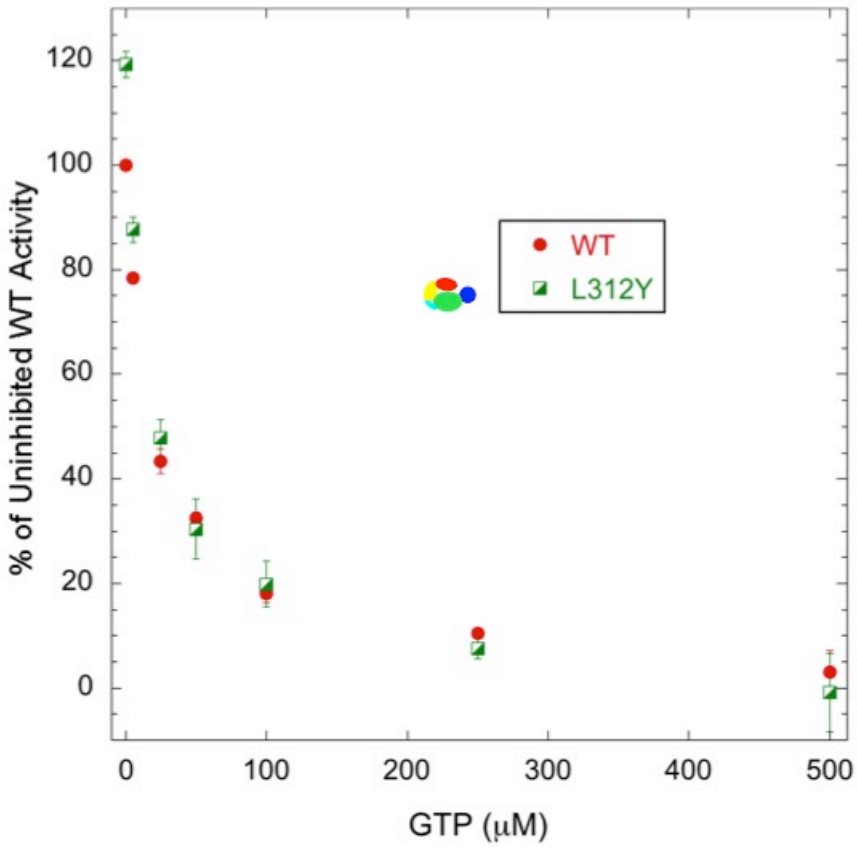


Figure 2-27. Inhibition of transglutaminase activity by GTP for WT and closed-biased L312Y as shown as percent of WT's uninhibited activity.

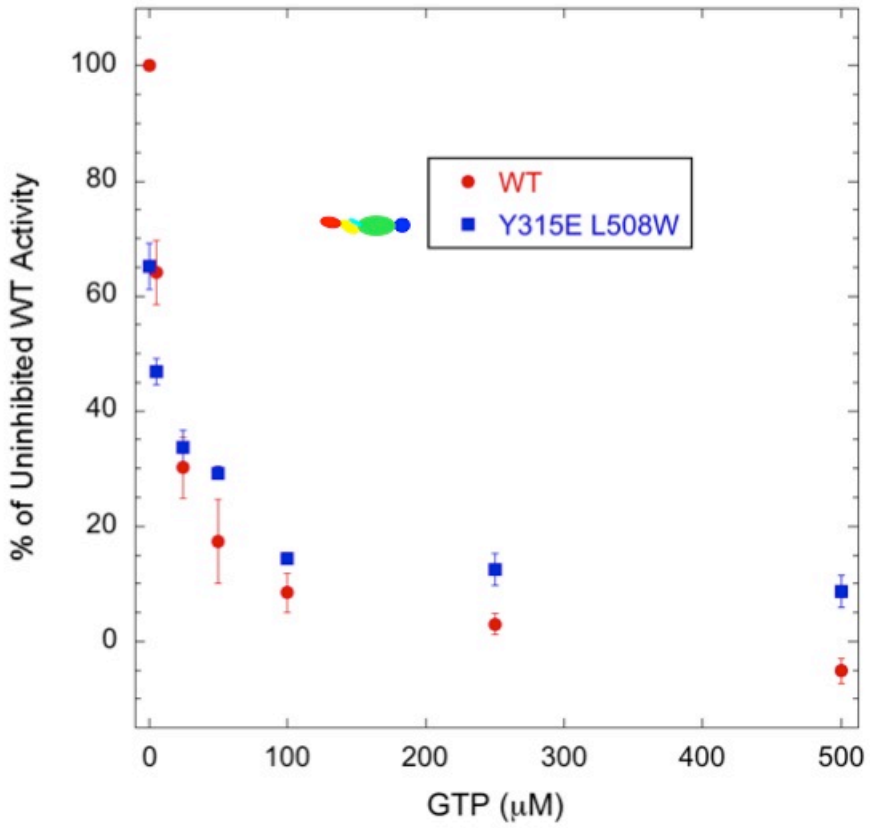


Figure 2-28. Inhibition of transglutaminase activity by GTP for WT and open-locked Y315E_L508W as shown as percent of WT's uninhibited activity.

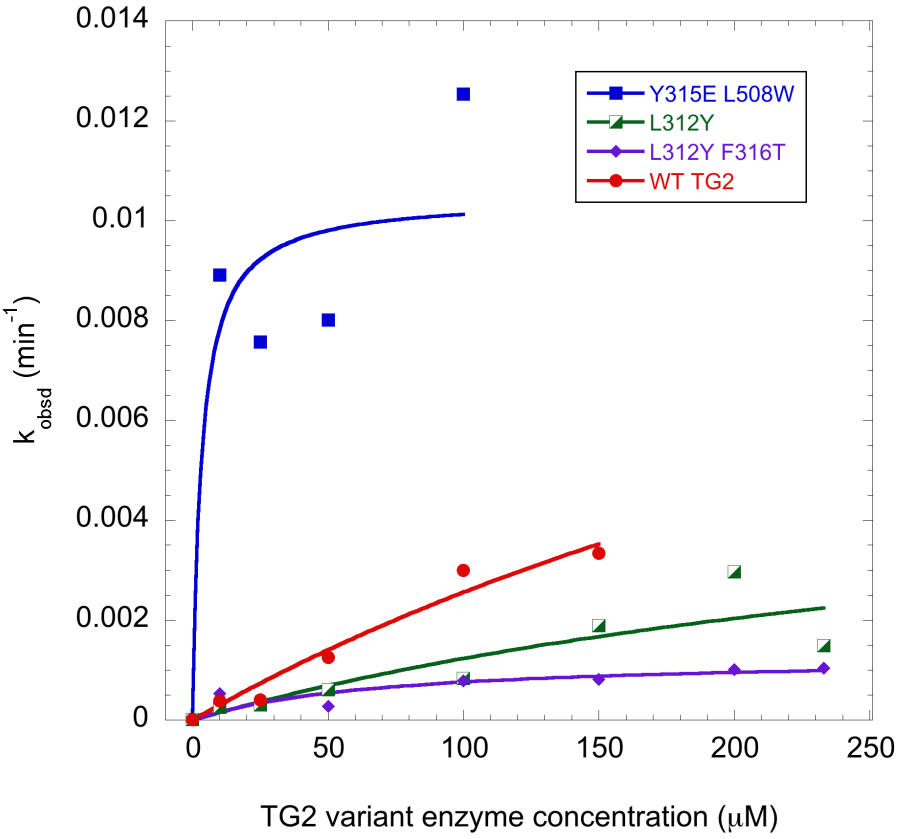


Figure 2-29. Single turnover kinetic analysis for the GTPase reaction for WT TG2, open-locked Y315E_L508W, closed-biased and most transglutaminase active L312Y, and most closed-biased L312Y_F316T. These concentration dependencies are fit to equation 5 to determine k_{max}, the maximal rate constant with saturating protein, and $K_{1/2}$, the protein concentration that provides half the maximal rate.

Table 2-2. Summary of rate and equilibrium constants for the single turnover GTPase reaction of TG2 variants.

TG2 Variant	k_{max} (min^{-1})	$K_{1/2}$ (μM)
WT	0.014 ± 0.014	460 ± 550
Y315E L508W	0.01 ± 0.0020	3.4 ± 4.9
L312Y	0.0058 ± 0.0076	370 ± 720
L312Y F316T	0.0013 ± 0.00034	68 ± 51

References

1. Ai, L., et al., *Ataxia-Telangiectasia, Mutated (ATM)/Nuclear Factor kappa light chain enhancer of activated B cells (NFkappaB) signaling controls basal and DNA damage-induced transglutaminase 2 expression.* J Biol Chem, 2012. **287**(22): p. 18330-41.

2. Griffin, M., R. Casadio, and C.M. Bergamini, *Transglutaminases: nature's biological glues.* Biochem J, 2002. **368**(Pt 2): p. 377-96.

3. Satchwell, T.J., et al., *Protein 4.2: a complex linker.* Blood Cells Mol Dis, 2009. **42**(3): p. 201-10.

4. Lorand, L. and R.M. Graham, *Transglutaminases: crosslinking enzymes with pleiotropic functions.* Nat Rev Mol Cell Biol, 2003. **4**(2): p. 140-56.

5. Folk, J.E. and S.I. Chung, *Transglutaminases.* Methods Enzymol, 1985. **113**: p. 358-75.

6. John, S., et al., *Epidermal transglutaminase (TGase 3) is required for proper hair development, but not the formation of the epidermal barrier.* Plos One, 2012. **7**(4): p. e34252.

7. Hitomi, K., et al., *Analysis of epidermal-type transglutaminase (TGase 3) expression in mouse tissues and cell lines.* Int J Biochem Cell Biol, 2001. **33**(5): p. 491-8.

8. Sardy, M., et al., *Epidermal transglutaminase (TGase 3) is the autoantigen of dermatitis herpetiformis.* J Exp Med, 2002. **195**(6): p. 747-57.

9. Kim, H.C., et al., *Protransglutaminase E from guinea pig skin. Isolation and partial characterization.* Journal of Biological Chemistry, 1990. **265**(35): p. 21971-8.

10. Komaromi, I., Z. Bagoly, and L. Muszbek, *Factor XIII: novel structural and functional aspects.* J Thromb Haemost, 2011. **9**(1): p. 9-20.

11. Muszbek, L., V.C. Yee, and Z. Hevessy, *Blood coagulation factor XIII: structure and function.* Thromb Res, 1999. **94**(5): p. 271-305.

12. Muszbek, L., et al., *The involvement of blood coagulation factor XIII in fibrinolysis and thrombosis.* Cardiovasc Hematol Agents Med Chem, 2008. **6**(3): p. 190-205.

13. Katona, E., et al., *A simple, quick one-step ELISA assay for the determination of complex plasma factor XIII (A2B2).* Thromb Haemost, 2000. **83**(2): p. 268-73.

14. Inbal, A., et al., *Impaired wound healing in factor XIII deficient mice.* Thromb Haemost, 2005. **94**(2): p. 432-7.

15. Dardik, R., J. Loscalzo, and A. Inbal, *Factor XIII (FXIII) and angiogenesis.* J Thromb Haemost, 2006. **4**(1): p. 19-25.

16. Noll, T., et al., *Effect of factor XIII on endothelial barrier function.* J Exp Med, 1999. **189**(9): p. 1373-82.

17. Hirahara, K., et al., *Suppressive effect of human blood coagulation factor XIII on the vascular permeability induced by anti-guinea pig endothelial cell antiserum in guinea pigs.* Thromb Res, 1993. **71**(2): p. 139-48.

18. Johnson, G.V., et al., *Transglutaminase activity is increased in Alzheimer's disease brain.* Brain Research, 1997. **751**(2): p. 323-9.

19. Fesus, L. and M. Piacentini, *Transglutaminase 2: an enigmatic enzyme with diverse functions.* Trends Biochem Sci, 2002. **27**(10): p. 534-9.

20. Park, D., S.S. Choi, and K.S. Ha, *Transglutaminase 2: a multi-functional protein in multiple subcellular compartments.* Amino Acids, 2010. **39**(3): p. 619-31.

21. Chiocca, E.A., P.J. Davies, and J.P. Stein, *Regulation of tissue transglutaminase gene expression as a molecular model for retinoid effects on proliferation and differentiation.* J Cell Biochem, 1989. **39**(3): p. 293-304.

22. Piacentini, M., et al., *The expression of "tissue" transglutaminase in two human cancer cell lines is related with the programmed cell death (apoptosis).* Eur J Cell Biol, 1991. **54**(2): p. 246-54.

23. Nemes, Z., Jr., et al., *Identification of cytoplasmic actin as an abundant glutaminyl substrate for tissue transglutaminase in HL-60 and U937 cells undergoing apoptosis.* J Biol Chem, 1997. **272**(33): p. 20577-83.

24. Upchurch, H.F., et al., *Localization of cellular transglutaminase on the extracellular matrix after wounding: characteristics of the matrix bound enzyme.* J Cell Physiol, 1991. **149**(3): p. 375-82.

25. Siegel, M., et al., *Extracellular transglutaminase 2 is catalytically inactive, but is transiently activated upon tissue injury.* PLoS One, 2008. **3**(3): p. e1861.

26. Mahoney, S.A., et al., *Stabilization of neurites in cerebellar granule cells by transglutaminase activity: identification of midkine and galectin-3 as substrates.* Neuroscience, 2000. **101**(1): p. 141-55.

27. Eitan, S. and M. Schwartz, *A transglutaminase that converts interleukin-2 into a factor cytotoxic to oligodendrocytes.* Science, 1993. **261**(5117): p. 106-8.

28. Eitan, S., et al., *Recovery of visual response of injured adult rat optic nerves treated with transglutaminase.* Science, 1994. **264**(5166): p. 1764-8.

29. Pinkas, D.M., et al., *Transglutaminase 2 undergoes a large conformational change upon activation.* PLoS Biol, 2007. **5**(12): p. e327.

30. Liu, S., R.A. Cerione, and J. Clardy, *Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity.* Proc Natl Acad Sci U S A, 2002. **99**(5): p. 2743-7.

31. Nakaoka, H., et al., *Gh: a GTP-binding protein with transglutaminase activity and receptor signaling function.* Science, 1994. **264**(5165): p. 1593-6.

32. Kiraly, R., et al., *Functional significance of five noncanonical Ca2+-binding sites of human transglutaminase 2 characterized by site-directed mutagenesis.* FEBS J, 2009. **276**(23): p. 7083-96.

33. Folk, J.E., *Transglutaminases.* Annu Rev Biochem, 1980. **49**: p. 517-31.

34. Lai, T.S., et al., *Identification of two GTP-independent alternatively spliced forms of tissue transglutaminase in human leukocytes, vascular smooth muscle, and endothelial cells.* FASEB J, 2007. **21**(14): p. 4131-43.

35. Datta, S., M.A. Antonyak, and R.A. Cerione, *Importance of Ca(2+)-dependent transamidation activity in the protection afforded by tissue transglutaminase against doxorubicin-induced apoptosis.* Biochemistry, 2006. **45**(44): p. 13163-74.

36. Begg, G.E., et al., *Mechanism of allosteric regulation of transglutaminase 2 by GTP.* Proc Natl Acad Sci U S A, 2006. **103**(52): p. 19683-8.

37. Stamnaes, J., et al., *Redox regulation of transglutaminase 2 activity.* J Biol Chem, 2010. **285**(33): p. 25402-9.

38. Klock, C., T.R. Diraimondo, and C. Khosla, *Role of transglutaminase 2 in celiac disease pathogenesis.* Semin Immunopathol, 2012. **34**(4): p. 513-22.

39. Smethurst, P.A. and M. Griffin, *Measurement of tissue transglutaminase activity in a permeabilized cell system: its regulation by Ca2+ and nucleotides.* Biochem J, 1996. **313 (Pt 3)**: p. 803-8.

40. Aeschlimann, D., D. Mosher, and M. Paulsson, *Tissue transglutaminase and factor XIII in cartilage and bone remodeling.* Semin Thromb Hemost, 1996. **22**(5): p. 437-43.

41. Aeschlimann, D., O. Kaupp, and M. Paulsson, *Transglutaminase-catalyzed matrix cross-linking in differentiating cartilage: identification of osteonectin as a major glutaminyl substrate.* J Cell Biol, 1995. **129**(3): p. 881-92.

42. Jones, R.A., et al., *Reduced expression of tissue transglutaminase in a human endothelial cell line leads to changes in cell spreading, cell adhesion and reduced polymerisation of fibronectin.* Journal of Cell Science, 1997. **110 (Pt 19)**: p. 2461-72.

43. Kojima, S., K. Nara, and D.B. Rifkin, *Requirement for transglutaminase in the activation of latent transforming growth factor-beta in bovine endothelial cells.* J Cell Biol, 1993. **121**(2): p. 439-48.

44. Janiak, A., E.A. Zemskov, and A.M. Belkin, *Cell surface transglutaminase promotes RhoA activation via integrin clustering and suppression of the Src-p190RhoGAP signaling pathway.* Mol Biol Cell, 2006. **17**(4): p. 1606-19.

45. Singh, U.S., et al., *Role of transglutaminase II in retinoic acid-induced activation of RhoA-associated kinase-2.* EMBO J, 2001. **20**(10): p. 2413-23.

46. Verderio, E.A., et al., *A novel RGD-independent cel adhesion pathway mediated by fibronectin-bound tissue transglutaminase rescues cells from anoikis.* J Biol Chem, 2003. **278**(43): p. 42604-14.

47. Akimov, S.S. and A.M. Belkin, *Cell-surface transglutaminase promotes fibronectin assembly via interaction with the gelatin-binding domain of fibronectin: a role in TGFbeta-dependent matrix deposition.* J Cell Sci, 2001. **114**(Pt 16): p. 2989-3000.

48. Lorand, L., J.E. Dailey, and P.M. Turner, *Fibronectin as a carrier for the transglutaminase from human erythrocytes.* Proc Natl Acad Sci U S A, 1988. **85**(4): p. 1057-9.

49. Akimov, S.S., et al., *Tissue transglutaminase is an integrin-binding adhesion coreceptor for fibronectin.* J Cell Biol, 2000. **148**(4): p. 825-38.

50. Jin, X., et al., *Activation of extracellular transglutaminase 2 by thioredoxin.* Journal of Biological Chemistry, 2011. **286**(43): p. 37866-73.

51. Diraimondo, T.R., C. Klock, and C. Khosla, *Interferon-gamma activates transglutaminase 2 via a phosphatidylinositol-3-kinase-dependent pathway: implications for celiac sprue therapy.* J Pharmacol Exp Ther, 2012. **341**(1): p. 104-14.

52. Ientile, R., D. Caccamo, and M. Griffin, *Tissue transglutaminase and the stress response.* Amino Acids, 2007. **33**(2): p. 385-94.

53. Ahvazi, B., et al., *Three-dimensional structure of the human transglutaminase 3 enzyme: binding of calcium ions changes structure for activation.* EMBO J, 2002. **21**(9): p. 2055-67.

54. Weiss, M.S., H.J. Metzner, and R. Hilgenfeld, *Two non-proline cis peptide bonds may be important for factor XIII function.* FEBS Lett, 1998. **423**(3): p. 291-6.

55. Yee, V.C., et al., *Structural evidence that the activation peptide is not released upon thrombin cleavage of factor XIII.* Thromb Res, 1995. **78**(5): p. 389-97.

56. Yee, V.C., et al., *Three-dimensional structure of a transglutaminase: human blood coagulation factor XIII.* Proc Natl Acad Sci U S A, 1994. **91**(15): p. 7296-300.

57. Gundemir, S., et al., *Transglutaminase 2: a molecular Swiss army knife.* Biochim Biophys Acta, 2012. **1823**(2): p. 406-19.

58. Csosz, E., B. Mesko, and L. Fesus, *Transdab wiki: the interactive transglutaminase substrate database on web 2.0 surface.* Amino Acids, 2009. **36**(4): p. 615-7.

59. Hornyak, T.J. and J.A. Shafer, *Role of calcium ion in the generation of factor XIII activity.* Biochemistry, 1991. **30**(25): p. 6175-82.

60. Sabo, T.M., P.B. Brasher, and M.C. Maurer, *Perturbations in factor XIII resulting from activation and inhibition examined by solution based methods and detected by MALDI-TOF MS.* Biochemistry, 2007. **46**(35): p. 10089-101.

61. Turner, B.T., Jr., et al., *Mapping of factor XIII solvent accessibility as a function of activation state using chemical modification methods.* Biochemistry, 2004. **43**(30): p. 9755-65.

62. Turner, B.T., Jr. and M.C. Maurer, *Evaluating the roles of thrombin and calcium in the activation of coagulation factor XIII using H/D exchange and MALDI-TOF MS.* Biochemistry, 2002. **41**(25): p. 7947-54.

63. Sakata, Y. and N. Aoki, *Cross-linking of alpha 2-plasmin inhibitor to fibrin by fibrin-stabilizing factor.* J Clin Invest, 1980. **65**(2): p. 290-7.

64. Lorand, L., K. Konishi, and A. Jacobsen, *Transpeptidation mechanism in blood clotting.* Nature, 1962. **194**: p. 1148-9.

65. Ahvazi, B., et al., *Roles of calcium ions in the activation and activity of the transglutaminase 3 enzyme.* J Biol Chem, 2003. **278**(26): p. 23834-41.

66. Kim, I.G., et al., *The deduced sequence of the novel protransglutaminase E (TGase3) of human and mouse.* Journal of Biological Chemistry, 1993. **268**(17): p. 12682-90.

67. Ahvazi, B., K.M. Boeshans, and F. Rastinejad, *The emerging structural understanding of transglutaminase 3.* J Struct Biol, 2004. **147**(2): p. 200-7.

68. Iismaa, S.E., et al., *The core domain of the tissue transglutaminase Gh hydrolyzes GTP and ATP.* Biochemistry, 1997. **36**(39): p. 11655-64.

69. Lai, T.S., et al., *C-terminal deletion of human tissue transglutaminase enhances magnesium-dependent GTP/ATPase activity.* Journal of Biological Chemistry, 1996. **271**(49): p. 31191-31195.

70. Sollid, L.M. and B. Jabri, *Is celiac disease an autoimmune disorder?* Curr Opin Immunol, 2005. **17**(6): p. 595-600.

71. Koning, F., et al., *Pathomechanisms in celiac disease.* Best Pract Res Clin Gastroenterol, 2005. **19**(3): p. 373-87.

72. Dieterich, W., et al., *Identification of tissue transglutaminase as the autoantigen of celiac disease.* Nat Med, 1997. **3**(7): p. 797-801.

73. Quarsten, H., et al., *HLA binding and T cell recognition of a tissue transglutaminase-modified gliadin epitope.* Eur J Immunol, 1999. **29**(8): p. 2506-14.

74. Sollid, L.M. and E. Thorsby, *HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis.* Gastroenterology, 1993. **105**(3): p. 910-22.

75. Nilsen, E.M., et al., *Gluten induces an intestinal cytokine response strongly dominated by interferon gamma in patients with celiac disease.* Gastroenterology, 1998. **115**(3): p. 551-63.

76. Nilsen, E.M., et al., *Gluten specific, HLA-DQ restricted T cells from coeliac mucosa produce cytokines with Th1 or Th0 profile dominated by interferon gamma.* Gut, 1995. **37**(6): p. 766-76.

77. Alaedini, A. and P.H. Green, *Narrative review: celiac disease: understanding a complex autoimmune disorder.* Ann Intern Med, 2005. **142**(4): p. 289-98.

78. Halttunen, T. and M. Maki, *Serum immunoglobulin A from patients with celiac disease inhibits human T84 intestinal crypt epithelial cell differentiation.* Gastroenterology, 1999. **116**(3): p. 566-72.

79. Cervio, E., et al., *Sera of patients with celiac disease and neurologic disorders evoke a mitochondrial-dependent apoptosis in vitro.* Gastroenterology, 2007. **133**(1): p. 195-206.

80. Di Simone, N., et al., *Anti-tissue transglutaminase antibodies from celiac patients are responsible for trophoblast damage via apoptosis in vitro.* Am J Gastroenterol, 2010. **105**(10): p. 2254-61.

81. Myrsky, E., et al., *Celiac disease IgA modulates vascular permeability in vitro through the activity of transglutaminase 2 and RhoA.* Cell Mol Life Sci, 2009. **66**(20): p. 3375-85.

82. Myrsky, E., et al., *Coeliac disease-specific autoantibodies targeted against transglutaminase 2 disturb angiogenesis.* Clin Exp Immunol, 2008. **152**(1): p. 111-9.

83. Caja, S., et al., *Inhibition of transglutaminase 2 enzymatic activity ameliorates the anti-angiogenic effects of coeliac disease autoantibodies.* Scand J Gastroenterol, 2010. **45**(4): p. 421-7.

84. Martucciello, S., et al., *RhoB is associated with the anti-angiogenic effects of celiac patient transglutaminase 2-targeted autoantibodies.* J Mol Med (Berl), 2012. **90**(7): p. 817-26.

85. Castellanos-Rubio, A., et al., *Angiogenesis-related gene expression analysis in celiac disease.* Autoimmunity, 2012. **45**(3): p. 264-70.

86. Kalliokoski, S., et al., *Celiac Disease-Specific TG2-Targeted Autoantibodies Inhibit Angiogenesis and in Mice by Interfering with Endothelial Cell Dynamics.* PLoS One, 2013. **8**(6): p. e65887.

87. Kiraly, R., et al., *Coeliac autoantibodies can enhance transamidating and inhibit GTPase activity of tissue transglutaminase: dependence on reaction environment and enzyme fitness.* J Autoimmun, 2006. **26**(4): p. 278-87.

88. Simon-Vecsei, Z., et al., *A single conformational transglutaminase 2 epitope contributed by three domains is critical for celiac antibody binding and effects.* Proc Natl Acad Sci U S A, 2012. **109**(2): p. 431-6.

89. Allen, B.D., *Development and validation of optimization methods for the design of protein sequences and combinatorial libraries*, 2009, California Institute of Technology.

90. Allen, B.D. and S.L. Mayo, *An efficient algorithm for multistate protein design based on FASTER.* J Comput Chem, 2010. **31**(5): p. 904-16.

91. Havranek, J.J. and P.B. Harbury, *Automated design of specificity in molecular recognition.* Nat Struct Biol, 2003. **10**(1): p. 45-52.

92. Ary, M.L.a.P., H.K., *PHOENIX User Manual*2009: California Institute of Technology.

93. Shimaoka, M., et al., *Computational design of an integrin I domain stabilized in the open high affinity conformation.* Nat Struct Biol, 2000. **7**(8): p. 674-8.

94. Hoover, D.M. and J. Lubkowski, *DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis.* Nucleic Acids Res, 2002. **30**(10): p. e43.

95. Murthy, S.N., et al., *Interactions of G(h)/transglutaminase with phospholipase Cdelta1 and with GTP.* Proc Natl Acad Sci U S A, 1999. **96**(21): p. 11815-9.

96. Case, A., et al., *Development of a mechanism-based assay for tissue transglutaminase--results of a high-throughput screen and discovery of inhibitors.* Anal Biochem, 2005. **338**(2): p. 237-44.

97. Case, A. and R.L. Stein, *Kinetic analysis of the action of tissue transglutaminase on peptide and protein substrates.* Biochemistry, 2003. **42**(31): p. 9466-81.

98. Bergamini, C.M., M. Signorini, and L. Poltronieri, *Inhibition of erythrocyte transglutaminase by GTP.* Biochim Biophys Acta, 1987. **916**(1): p. 149-51.

99. Peluso, P., et al., *Role of SRP RNA in the GTPase cycles of Ffh and FtsY.* Biochemistry, 2001. **40**(50): p. 15224-33.

100. Murthy, S.N., et al., *Conserved tryptophan in the core domain of transglutaminase is essential for catalytic activity.* Proc Natl Acad Sci U S A, 2002. **99**(5): p. 2738-42.

Chapter 3

THE ADDITION OF TG2 DESIGNS WITH DISTINCT CONFORMATIONS AND ACTIVITIES ELICIT DIFFERENTIAL EXPRESSION OF CELIAC-RELATED GENES IN CELL CULTURE

Abstract

Celiac disease is an autoimmune condition affecting the small intestine in response to the ingestion of gluten in affected individuals. The primary autoantigen in celiac disease is type II transglutaminase (TG2), and these anti-TG2 antibodies themselves have been shown to induce anti-angiogenic and other pathogenic effects in endothelial models similar to those seen in celiac patients, including the up-regulation of genes like E-selectin. In celiac disease, TG2's transglutaminase activity contributes to pathogenesis by catalyzing the deamidation of gluten, which facilitates the binding cascade that leads to the autoimmune response. Recently, celiac autoantibodies have been shown to increase the transglutaminase activity of TG2 and bind a motif on the enzyme that is fully assembled only in the closed conformation, suggesting that the stabilization of the closed conformation may be leading to this increase in activity. Indeed, the results of Chapter 2 support this hypothesis. The goal of this chapter was to test if the increase in E-selectin expression and the other anti-angiogenic effects induced by celiac anti-TG2 antibodies are derived from their increasing of transglutaminase activity. In order to accomplish this, we utilized the design variants detailed in Chapter 2 to investigate if conformationally distinct forms of TG2 with differing levels of transglutaminase activity could elicit differing effects in cell culture, and to further investigate the hypothesis that the stabilization of the closed conformation and the resulting increase in transglutmainase activity of TG2 led to pathogenic effects observed in CD. Using qPCR, we found that the addition of TG2 variants, including a transglutaminase knockout, to HUVEC culture did increase the activation of NF-κB (which is activated in celiac disease) and increased E-selectin expression. The increases in E-selectin among closed-biased and WT TG2 did increase with increasing transglutaminase activity, but open-locked knocked-down and knocked-out versions of TG2 also increased NF-κB activiation and E-selectin expression, making it impossible to conclude from this initial study if increased transglutaminase activity alone is responsible for these aberrations. Further studies investigating possible transglutaminase activity independent mechanisms of pathway activation are therefore indicated.

Introduction

The transglutaminase activity of TG2 plays a critical role in celiac disease

Celiac disease (CD) is an intestinal autoimmune condition triggered in genetically predisposed individuals by the ingestion of gluten. [1, 2]. In CD, TG2 deamidates gluten peptides in the small intestine, increasing their binding affinity for mutant major histocompatibility complex (MHC) class II molecules, which elicits the autoimmune response [3-8]. The primary autoantigen in celiac disease is also TG2 [3], and a primary hallmark of CD is the secretion of anti-TG2 antibodies into the small intestinal mucosa where they bind extracellular TG2 on fibroblasts surrounding mucosal blood vessels [9-13].

CD anti-TG2 antibodies affect angiogenesis

Chapter 1 details the ability of CD anti-TG2 autoantibodies to cause pathogenic symptoms in endothelial models [14-22]. One such effect is the ability of CD antibodies to enhance lymphocyte binding to the endothelium and their transendothelial migration, which is complemented by a more than 2-fold up-regulation in expression of E-selectin, a lymphocyte adhesion molecule [17]. This up-regulation in E-selectin [17] was determined by adding celiac antibodies at a concentration of 1 μg/mL to human umbilical vein endothelial cells (HUVECs) for 24 hours, and then analyzing them for E-selectin protein expression by flow cytometry and immunofluorescent staining. These effects are relevant and suggestive because increased levels of adhesion molecules crucial to the inflammatory response and lymphocyte migration, such as E-selectin, are also observed in untreated celiac disease [23-25].

Many other studies have investigated celiac total IgA and purified celiac anti-TG2 autoantibodies' ability to inhibit angiogenesis and increase vascular permeability [17-22]. One study [20] used microarray technology to analyze the expression of a set of angiogenesis and endothelial cell biology genes in HUVECs after a 24 hour treatment with either total IgA from celiac or control patients, or monoclonal control or anti-TG2 celiac miniantibodies, also at a concentration of 1 μg/mL. Compared to control, total IgA up-regulated the expression of four genes, and down-regulated the expression of six. Of these ten genes, three were selected for RT-PCR analysis and confirmed for differential expression: integrin beta 1 (ITGB1) decreased by a factor of 0.73; fibronectin 1 (FN1) increased by a factor of 1.2; and ras homolog gene family member B (RHOB)

increased by a factor of 1.26. RHOB was found to be similarly up-regulated in gene expression by monoclonal anti-TG CD antibodies, and RhoB protein expression was also increased by a factor of approximately 1.45 by both total IgA and monoclonal anti-TG2 CD antibodies.

CD anti-TG2 antibodies increase TG2 transglutaminase activity

CD anti-TG2 autoantibodies have been reported to increase the transglutaminase activity of TG2 [17, 19, 26]. In the celiac patient, an increase in transglutaminase activity could potentially increase the enzyme's ability to deamidate gluten, creating a positive feedback loop where the disorder reinforces itself by increasing TG2's ability to create more binding partners for the mutant MHC class II molecules [19].

Because the celiac antibodies were found to be transglutaminase activating, it was believed that the antibodies might be stabilizing a more active form of TG2 [26]. Recently, the TG2 epitope recognized by these celiac autoantibodies was identified to be a conformationally dependent one comprising residues from three different domains: the N-terminus, the core, and β-barrel 2 [27]. Because of their differing locations, the residues are only spatially close to each other when the enzyme adopts the closed conformation. Therefore, if the antibodies are increasing the activity of TG2, they may be doing so by stabilizing the closed conformation of the enzyme, which was traditionally believed to be the transglutaminase inactive form.

Nuclear factor-κB is activated in the intestinal mucosa in celiac disease

Nuclear factor-κB (NF-κB) is a transcription factor that regulates immune and inflammatory responses [28, 29], and is activated in biopsy specimens of celiac patient small intestinal mucosa [30]. This activation is highest in untreated celiac patients, and decreases upon treatment, yet remains higher than control subjects [30].

Suggestively, TG2 expression and activity is closely connected with NF-κB activation. Studies investigating chemoresistance of various cancers showed both tumor cells overexpressing TG2 and the activation of TG2 in such cells led to the activation of NF-κB [31]. In addition, both the

inhibition of transglutaminase activity and the knock-down of TG2 expression by RNAi inhibited NF-κB activation [31, 32].

TG2 gene and protein expression is tightly correlated with NF-κB activity [32, 33]. In fact, TG2 and NF-κB have been shown to exist in a self-reinforcing feedback loop, where TG2 activates NF-κB, and in return NF-κB induces further TG2 expression [33-35]. On a related note, TNF-α, an inducer of NF-κB , has also been shown to up-regulate E-selectin [36, 37].

The mechanism by which TG2 activates NF-κB is still not fully understood, but TG2 has been shown to activate NF-κB by inducing the depletion of inhibitory subunit α of NF-κB (IκBα or NFKB1A), and thereby preventing it from inhibiting NF-κB [38-40]. The proposed mechanism of this depletion of IκBα is the crosslinking of IκBα by TG2, resulting in its polymerization and inability to inhibit NF-κB [38]. This theory is contested, however, by a study showing that the overexpression of WT TG2 and the active site knockout TG2 mutant C277S were both able to activate NF-κB in mammary epithelial cells, although they also concluded that this activation was produced by the depletion of IκBα, and that IκBα associated in complexes with TG2-C277A [41].

TG2 stabilized in various conformations with varying levels of TGase activity are a new tool

The above studies found that autoantibodies against TG2 isolated from celiac patients increase TG2 transglutaminase activity and bind a motif that exists in the closed conformation, suggesting that the stabilization of this conformation leads to the increase in activity. This hypothesis is supported by the findings detailed in Chapter 2 that the transglutaminase activities of the open-locked designs of TG2 are significantly decreased compared to WT and closed-biased TG2 designs. If the increase in adhesion protein expression and the anti-angiogenic effects of adding anti-TG2 autoantibodies to the media of cultured HUVECs are mediated by the binding of these antibodies to extracellular TG2, and the resulting increase in activity, it is conceivable that the addition of TG2 to the media of HUVECs could induce a similar response.

In fact, adding TG2 to the media of HUVECs has been shown to elicit anti-angiogenic effects: a previous study showed the addition of WT TG2 to HUVEC culture at 50 μg/mL resulted in the accumulation of extracellular matrix (ECM), leading to the anti-angiogenic characteristic of suppressed endothelial tubule formation [42] similar to that seen in investigations of CD autoantibodies' anti-angiogenic effects [17-20, 22, 27].

The previous chapter describes the design of variants of TG2 preferring either the open or the closed conformation. Having such designs available provided the unique opportunity to investigate if conformationally distinct forms of TG2 with differing levels of transglutaminase activity could elicit differing effects in cell culture, and to further investigate the hypothesis that the stabilization of the closed conformation and the resulting increase in transglutmainase activity of TG2 led to pathogenic effects observed in CD. In this chapter, using qPCR, we probed both the activation of NF-κB and the expression-level response of genes shown to be affected by the addition of celiac anti-TG2 autoantibodies to the addition of various TG2 variants to HUVEC cell culture, similar to the methods used in the celiac autoantibody studies detailed above. The variants used included the transglutaminase inactive and open-conformation-adopting C277A_Y315E, open-locked Y315E_L508W (V_{max} 0.47 FI/s) with low transglutaminase activity, WT TG2 (V_{max} 1.6 FI/s), and closed-biased L312Y (V_{max} 2.0 FI/s) and L312Y_F316T (V_{max} 1.8 FI/s) with higher than WT transglutaminase activity.

Methods

Protein expression and purification

LB/ampicillin starter cultures were inoculated from a glycerol stock of BL-21 (DE3) Gold *E. coli* containing the expression vector pET11a (Novagen) with the gene for the variant of interest and grown overnight at 37°C with shaking. 4 mL of starter culture were added to 100 mL of Overnight Express Instant TB media (Novagen) with ampicillin and grown at 37°C with shaking. After 3 hours, cultures were grown at 16°C for 24 hours with shaking.

Cells were harvested by centrifugation at 5000 x g for 10 minutes at 4°C. Each 100 mL pellet was resuspended in 8 mL of lysis buffer (50 mM sodium phosphate, pH 8.0, 0.3 M NaCl, 2.5 mM imidazole, 1x CelLytic B (Sigma-Aldrich), 5 mM β-mercaptoethanol (βME), 0.2 mg/mL lysozyme (Sigma-Aldrich), and 1 U Benzonase endonuclease (Merck)). The lysate was centrifuged at 19000 x g for 45 at 4°C and the supernatant collected.

His-Select spin columns (Sigma-Aldrich) were equilibrated with 600 µL of equilibrium buffer (50 mM sodium phosphate, pH 8.0, 0.3 M NaCl, 2.5 mM imidazole, and 5 mM βME). The crude supernatant was applied to the columns and the columns were washed with 1200 µL of wash buffer (50 mM sodium phosphate, pH 8.0, 0.3 M sodium chloride, 5 mM imidazole, and 5 mM βME). The purified protein was eluted by the addition of 300 µL of elution buffer (50 mM sodium phosphate, pH 8.0, 0.3 M sodium chloride, 250 mM imidazole and 5 mM βME).

Eluted protein was buffer exchanged into storage buffer (50 mM HEPES, pH 7.0, 10% glycerol, 100 mM NaCl, 1 mM EDTA, and 5 mM DTT) with 0.5 mL 50K Amicon Ultra centrifugal filters (Millipore).

Protein concentration determination

Protein concentrations were determined by UV absorbance denaturation in 8 M guanidinium hydrochloride for 5 minutes at a dilution of 10x to 100x. The extinction coefficient for each protein at 280 nm was calculated based on the number of tryptophans and tyrosines in the protein using ExPASy's ProtParam tool.

Cell culture

Human umbilical vein endothelial cells (HUVECs) (Invitrogen and Lonza) were cultured in Endothelial Growth Media-2 (EGM-2 Medium; Lonza). EGM-2 consists of Endothelial Cell Basal Medium-2 (EBM-2; Lonza) and the endothelial cell growth factors in the EGM-2 Bulletkit (Lonza). Cells were cultured on type I collagen-coated plates at 37°C in a 5% CO_2 incubator.

Gene expression assay of treated cell cultures

HUVECs were plated at 10,000 cells/cm^2. After 20 hours, new media was added to the culture wells. The enzymes under investigation were WT, transglutaminase active site knockout C227A_Y315E, closed biased L312Y or L312Y_F316T, or open-locked Y315E_L508W. All enzymes were added to the media at a final concentration 5 μg/mL. Tumor necrosis factor α (TNF-α) (Sigma-Aldrich) was used as a positive control for NF-κB activation and therefore NFKB1A expression [43] and was added at a final concentration of 10 ng/mL. Three negative controls consisted of either the addition of storage buffer without enzyme, the active site knockout C227A_Y315E, or untreated cells. Time points were investigated by adding the enzyme or control effectors either when the new media was added, 12 hours after, or 21 hours after, resulting in samples that were treated for 24, 12, or 3 hours, respectively. Total RNA was extracted using the RNAeasy Kit (Qiagen) according to the manufacturer's protocol. Experiments were done with a biological replicate. The experiment was performed twice, the first time with cells from Invitrogen, the second time with cells from Lonza. All experiments were performed on passages 2 through 5.

qPCR

The comlementary DNA (cDNA) was synthesized using the iScript cDNA Synthesis Kit (BioRad) using 0.5 μg of RNA. Quantitative PCR was performed on a BioRad CFX96 Real Time System with $1/50^{th}$ of the cDNA reaction, single gene PrimeTime qPCR 0.5 μM/0.25 μM final primer/probe mixtures from IDT (see Table 3-1), and SsoFast Probes Supermix (BioRad) in a total volume of 10 μL. All probes used the fluorophore FAM. The cycling conditions consisted of an initial

denaturation for 30 s at 95°C, followed by 39 cycles of 5 s at 95°C and 10 s at 60°C. Each biological replicate was performed with a technical replicate. NF-κB activation was monitored by measuring the changes in expression of NFKB1A as described previously [43]. The expression level of genes was determined by the relative quantity method with the control expression level being that found in untreated cells, and normalized using the geometric mean of the housekeeping genes GAPDH, SDHA, and TBP.

Statistical analysis

Data are presented as mean ± standard error of mean (SEM). Statistical analysis was performed using the Student's *t* test with the Bonferroni correction, where the $p<0.05$ was considered significant. The Bonferroni correction for these data sets resulted in a modified *p* value 0.002 for a two-tailed test.

Results

When the Bonferroni correction was used to analyze the data with 21 pairwise comparisons, both NFKB1A and SELE were affected relative to buffer by various proteins in observations deemed statistically significant ($p<0.05$). However, no statistically significant changes in FN1, ITGB1, or RHOB relative expression were observed in this study (see Tables 3-14 through 16 and Figures 3-5 through 3-10). Results were qualitatively reproducible between experiments using different HUVEC sources (Figures 3-1 through 3-10).

Closed-biased L312Y_F316T and L312Y, as well as TNF-α were the only proteins whose effects on NFKB1A were determined to be statistically significant. After the 3-hour treatment, L312Y_F316T and L312Y showed increases in NFKB1A expression relative to buffer ranging from 4-fold to almost 6-fold (Table 3-2 and Figures 3-1, 3-2). In experiment 2, an increase in NFKB1A expression relative to buffer was seen for L312Y after 12 hours, and for L312Y_F316T after 24 hours (Table 3-2 and Figure 3-2).

TNF-α, C277A_Y315E, and Y315E_L508W all caused an increase in the mean relative expression level of NFKB1A, but due to larger errors, were deemed statistically insignificant, although the response to TNF-α after 12 hours in experiment 1 was determined to be significant (Table 3-2 and Figures 3-1, 3-2).

The response in SELE expression was large and observed to be statistically significant for every variant of TG2 added in at least one condition, including even the transglutaminase knockout C277A_Y315E (see Tables 3-8, 3-9, 3-10, 3-11, 3-12, and 3-13 and Figures 3-3, 3-4). After 12 hours of treatment, the increase in expression of SELE fell off dramatically, but remained elevated relative to buffer. After 3 hours of treatment, and in some cases 12 hours, the mean relative expression level of SELE increased by a very large amount in both experiments in response to TNF-α, Y315E_L508W, and even transglutaminase knockout C277A_Y315E, but again, were deemed statistically insignificant due to large errors.

In order to compare the relative expression levels of SELE among TG2 variants, I focused on experiment 2 and limited the number of pairwise comparisons to 6, so that the Bonferroni correction became less restrictive (Figure 3-11, Table 3-17). When restricted to the pairwise comparisons among closed stabilized L312Y_F316T and L312Y, WT TG2, and open-locked Y315E_L508W, all comparisons were statistically significant save between L312Y and Y315E_L508W. All three

designs exhibited a larger normalized expression levels than WT. Among WT and the closed designs, increases in SELE expression correlated with increasing transglutaminase activity, with WT having the lowest level of expression and L312Y the highest (Table 3-8b). Surprisingly, Y315E_L508W, with the lowest transglutaminase activity and highest GTPase activity in vitro had the highest mean level of expression of all four variants, although the difference between L312Y and Y315E_L508W was statistically insignificant.

Discussion

The purpose of this study was to investigate perturbations arising in gene expression levels of targets previously found to be associated with celiac disease or the addition of celiac anti-TG2 antibodies in response to the addition of conformationaly distinct forms of TG2 with differing levels of transglutaminase activity. NFKB1A expression, which correlates with NF-κB activation, and E-selectin expression, exhibited consistent and statistically significant up-regulation. RHOB, FN1, and ITGB1, whose expression has previously been shown to be affected by CD antibodies, showed no statistically significant responses.

NF-κB is a transcriptional regulator critical in the inflammatory response that can be activated by TG2, and in a positive-feedback loop, can induce the expression of TG2 [31, 33, 35, 44]. NF-κB is activated in the celiac gut [30], and celiac patients exhibit an increase in TG2 expression and transglutaminase activity [45, 46]. Studies have indicated that the binding of celiac antibodies to closed TG2 may be stabilizing a more active form of the enzyme, resulting in an increase in transglutaminase activity [17, 19, 26, 27]. We hypothesized that such an increase in transglutaminase activity could be a source of the induction of NF-κB activation in celiac disease, and therefore, this study investigated whether conformationaly distinct forms of TG2 with increasing levels of transglutaminase activity could induce increasing levels of NF-κB activity when added to the media of cultured HUVECs. NF-κB activation was quantitated by measuring the expression of the IκB-α inhibitory subunit (NFKB1A), which is expressed in response to NF-κB activation [43].

The modest increases in NFKB1A expression in response to the most transgltuaminase active designs (closed-biased L312Y and 312Y_F316T) indicate that activation of NF-κB may be occurring in response to higher extracellular levels of TG2 and/or transglutmainase activity.

At this point, however, we are unable to conclude that increased transglutaminase activity or the closed conformation of TG2 are solely responsible for any possible increases in NF-κB activity, due to observed increases in the mean relative expression of NFKB1A in response to the transglutaminase knockout C277A_Y315E, and the transglutaminase knock-down and open-locked Y315E_L508W, although these data have large errors, rendering them statistically insignificant.

It is not entirely unexpected that TG2 variants with little to no transglutaminase activity would elicit NF-κB activation given that the overexpression of C277S TG2, another transglutaminase knockout,

has been shown to elicit a similar level of NF-κB activation to that elicited by WT TG2 when overexpressed in mammary epithelial cells [41]. However, the result remains confusing since inhibiting transglutaminase activity has been shown to inhibit NF-κB activation [31, 32]. Therefore, further investigation by traditional means of measuring NF-κB activation, such as an electrophoretic mobility shift assay, are warranted and should include TG2 variants with decreased or eliminated transglutaminase activity, as well as celiac anti-TG2 autoantibodies. Such traditional means of measuring NF-κB activation may give more precise results.

During the inflammatory response in celiac disease, lymphocytes adhere to the vascular endothelium as they migrate to the sites of chronic inflammation in the lamina propria of the small intestine [17, 23]. Integral to lymphocytes' ability to bind the endothelium are adhesion molecules, which are up-regulated in celiac disease [23-25]. E-selectin is one such adhesion molecule whose levels in the sera of celiac patients have been found to be more than 10-fold higher than the levels in healthy controls [24]. One study showed that celiac IgA induces the up-regulation in expression of E-selectin, whereas non-celiac IgA did not [17].

As mentioned above, evidence suggests that celiac anti-TG2 autoantibodies may be binding the closed-conformation of TG2, possibly stabilizing a more transglutaminase-active form of the enzyme, suggesting that this increased activity may be the source of the pathogenic effects derived from these antibodies [17, 19, 26, 27]. If this hypothesis were true, we expected the closed-biased, more transglutaminase active variants of TG2 (L312Y and L312Y_F316T) to result in the largest increase in E-selectin (SELE) expression, whereas the less active open-locked Y315E_L508W was expected to yield the lowest increase in SELE expression.

In experiment 2, the relative expression levels of SELE among WT and closed-biased L312Y_F316T and L312Y did increase with increasing transglutaminase activity: (1) WT, TGase V_{max} 1.6 FI/s, 101-fold increase in SELE expression relative to untreated, (2) closed-biased L312Y_F316T, V_{max} 1.8 FI/s, 146-fold increase, (3) closed-biased L312Y, V_{max} 2.0 FI/s, 230-fold increase (See chapter 2 and table 3-8b). This indicates that greater levels of transglutaminase activity may lead to more SELE expression and therefore greater anti-angiogenic effects.

However, variants with little to no transglutaminase activity still yielded large increases in SELE expression in both experiments. Open-locked Y315E_L508W, with a transglutaminase V_{max} of 0.47 FI/s gave the largest fold increase in SELE (404-fold) in experiment 2, similar to what was observed upon treatment with TNF-α, a known inducer of SELE expression [36]. Transglutaminase

knockout and open-locked C277A_Y315E also yielded large increases in SELE expression in both experiments, although when the Bonferroni correction for 21 comparisons were made, this increase was determined to be statistically significant for only one condition (see Tables 3-8 through 3-13 and Figures 3-3 and 3-4). These results cast serious doubt on the hypothesis that increased transglutaminase activity alone is the cause of the pathogenic effects observed. Extracellular TG2 has been shown to activate various pathways by both transglutaminase dependent [47] and independent mechanisms [48, 49], and our results indicate that the open conformation itself may be conducive to a binding event that induces SELE expression via some unknown pathway. Also of note is that Y315E_L508W exhibits the largest GTPase activity of all the designs added, although C277A_Y315E's GTPase activity is unknown. If this increase in GTPase results from the adoption of the open conformation, this activity may also have some unknown effect on SELE expression.

TG2 has previously been reported to activate NF-κB [33-35], and TNF-α, a cytokine involved in the immune and inflammatory responses which also activates NF-κB, has also been shown to up-regulate E-selectin [36, 37]. E-selectin expression also has been reported to be induced by NF-κB binding to E-selectin's regulatory elements [50]. Another study demonstrated that celiac anti-TG2 antibodies could induce NF-κB activity by binding Toll-like receptor 4 (TLR4) in 293T cells [51]. Perhaps the increase in E-selectin resulting from the addition of CD anti-TG2 antibodies results from up-regulation via NF-κB activation mediated by the binding of CD antibodies to TLR4, and adding TG2 to the media can bring about a similar effect by activating NF-κB. As mentioned in the introduction, the mechanism by which increased amounts of TG2 can induce NF-κB are not yet fully understood, and may be independent of transglutaminase activity.

An interesting follow-up study would be to repeat the experiment, but block NF-κB activation and look to see if this blocks any up-regulation in E-selectin expression in response to the addition of TG2. If the observed increase in E-selectin expression in this study in response to transglutaminase active variants is decreased in the repeated experiment in cells with blocked NF-κB activation, but still elevated relative to untreated cells, this could indicate that increased transglutaminase activity could still contribute to the pathogenic effects induced by celiac autoantibodies, but by a separate pathway. This hypothesis would further be supported if E-selectin expression were unchanged in cells treated with transgltuaminase knockouts when NF-κB activation is blocked. This method would only work, however, if the proposed effects induced by increasing transglutaminase activity do not rely upon NF-κB activation as well.

Additionally, another type of assay may be beneficial to look at the anti-angiogenic effects of various TG2 designs compared to CD antibodies. One option is the tubule-length formation assay in HUVECs, where decreased tubule length formation is an indication of inhibited angiogenesis. This assay could be illuminating because studies showed that reduced tubule formation in response to anti-TG2 antibodies could be reversed by the addition of transglutaminase inhibitors [17, 19].

The possibility exists that simply adding TG2 to the media of HUVECs does not correctly mimic the events that occur when CD anti-TG2 antibodies bind extracellular TG2 associated with other proteins on the surface of the cell, and that the addition of TG2 to the media only activated NF-κB, and therefore increased E-selectin, via the positive feedback loop between TG2 and NF-κB. Also, although celiac autoantibodies were found to increase transgltuaminase activity in culture, the TG2 added to the media of HUVECs may not be active, although previous studies have shown that exogenous TG2 added to HUVEC media retained transglutaminase activity [42]. The stability of TG2 in the media of HUVECs is also unknown, and any increase in TG2 activity might be fleeting and therefore insufficient to elicit a response when compared to that of TG2 associated with the membrane of cells, or antibodies associated with endogenous TG2. This could explain why statistically significant responses in FN1, ITGB1, and RHOB expression were not observed. Repeated addition of TG2 variants to the cells, similar to the method used in [42], may be able to overcome short TG2 lifetimes in the media. Measuring the amount of endogenous *in situ* TG2 activity via fluorescein cadaverine incorporation as preformed in [42] would be useful in confirming and quantifying relative amounts of TG2 activity and the lifetime of the enzyme. It is also possible that increased transglutaminase activity plays no role in the pathogenic effects of celiac antibodies, but this is doubtful considering results that found that the anti-angiogenic and increased lymphocyte adhesion and transendothelial migration effects of CD antibodies could be reversed by the addition of transglutaminase inhibitors [17, 19], although the control of inhibitor alone added to cells was not conducted.

Knock-down of endogenous TG2 and transfection and expression of designs in HUVECs represents an alternative means to investigate the effects of differing levels of activity of TG2 on gene expression, but may be undesirable since TG2 is involved in GTPase signal transduction intracellularly, and therefore may alter the internal biology of the cell. This is especially relevant given the dramatic differences in GTPase activity observed among the TG2 designs. The open-locked variant used in this study (Y315E_L508W) has a much higher GTPase activity than WT or the two closed-biased designs studied here (see Chapter 2), therefore it would be interesting to test the GTPase activities of other open-locked designs to attempt to find one whose GTPase activity is

not that different than WT. This may not be possible given the potential for the conformation to be the source of the increase in GTPase activity. However, even if an open-locked TG2 with similar GTPase activity to WT can be designed to negate this inequality, a second undesirable complication arises from the potential for the change in conformational preference to interfere with TG2's interactions with other proteins. Due to these concerns, and because the effects of celiac anti-TG2 antibodies are believed to be mediated solely by their binding to extracellular TG2, we chose to simply add the designs to the cell media in this study.

Conclusions

In summary, the addition of all design variants of TG2 to HUVEC culture did increase NFKB1A and especially E-selectin expression. Because TG2 also has been shown to induce NF-κB activation via a pathway that may be independent of transglutaminase activity, it is possible that addition of TG2 to the media may bring about NF-κB activation, and the associated up-regulation of E-selectin. Therefore, future experiments should attempt to isolate the effects of transglutaminase activity by blocking NF-κB activation. The increases in E-selectin expression were strong, fleeting, and for closed-biased and WT TG2, correlated with transglutaminase activity, indicating that increased transglutaminase activity may be pathogenic. This finding supports the hypothesis that the stabilization of the closed form of TG2 may increase transglutaminase activity, and this may cause pathogenic effects. However, the transglutaminase knock-down and knockout variants induced at times even larger increased in E-selectin expression, which in some cases was indistinguishable from TNF-α-induced E-selectin up-regulation. These two variants adopt the open conformation, and it is possible that this conformation or a possible increase in GTPase activity, and not transglutaminase activity, may contribute to the effects observed in some unknown way.

In addition to conducting the tubule-formation assay to compare the responses of live cells to various levels of transglutaminase activity and conformation, we believe a fuller picture of the cell's state after exposure to our variants is warranted. Because E-selectin gene expression and NF-κB activation were increased in this study as a response to our TG2 variants by mechanisms that are not fully understood, we believe that this experiment should be repeated and RNA-Seq conducted to profile the transcriptome in response to various TG2 designs. The results from such a study may be able to indicate if the different designs are activating different pathways or help determine the mechanism by which we see these changes in expression and activation.

Acknowlegements

Special thanks goes to Zakary Singer of the Elowitz lab who trained me and allowed me to use one of the Elowitz lab's hoods to conduct all cell culture experiments, and who also advised me on tissue culture and qPCR. Thanks also to Lauren LeBon from the Elowitz lab who provided me with the HUVECs from Invitrogen.

Tables and Figures

Table 3-1. List of primers for qPCR analysis.

Gene	Primer Direction	Sequence
GAPDH	F	5'-TGTAGTTGAGGTCAATGAAGGG-3'
	R	5'-ACATCGCTCAGACACCATG-3'
SDHA	F	5'-CTGCCCCTTGTAGTTGGTG-3'
	R	5'-GTGGACGTCACGAAGGAG-3'
TBP	F	5'-TCGTGGCTCTCTTATCCTCAT-3'
	R	5'-CAGTGAATCTTGGTTGTAAACTTGA-3'
RHOB	F	5'-GGTTTCTTTTCCCTCTCCTTGT-3'
	R	5'-GCACTCGTCATTTGCATACG-3'
SELE	F	5'-TGTAGCTGAAGTTTCCCAGTG-3'
	R	5'-CTGGCTTCAGTGGACTCAA-3'
ITGB1	F	5'-GGTCAATGGGATAGTCTTCAGC-3'
	R	5'-GTAGCAAAGGAACAGCAGAGA-3'
NFKBIA	F	5'-GAGTCAGAGTTCACGGAGTTC-3'
	R	5'-CATGTTCTTTCAGCCCCTTTG-3'
FN1	F	5'-GTCCTTGTGTCCTGATCGTTG-3'
	R	5'-AGGCTGGATGATGGTAGATTG-3'

Table 3-2. Normalized expression of NFKB1A under different treatment conditions after 3, 12 and 24 hours' exposure, with standard error of means. (A) Experiment 1. (B) Experiment 2. Values are relative to untreated cells.

A	3 hours	12 hours	24 hours
Buffer	0.813 ± 0.0250	0.860 ± 0.0959	1.22 ± 0.120
TNFα	10.5 ± 1.42	6.67 ± 0.124	4.95 ± 0.774
C277A Y315E	5.85 ± 0.345	3.73 ± 0.0864	2.68 ± 0.193
Y315E L508W	8.14 ± 2.19	3.11 ± 0.218	1.82 ± 0.441
WT	3.71 ± 0.243	2.36 ± 0.0689	2.65 ± 0.302
L312Y F316T	3.87 ± 0.089	2.62 ± 0.139	2.61 ± 0.199
L312Y	5.12 ± 0.220	2.94 ± 0.825	2.94 ± 0.217

B	3 hours	12 hours	24 hours
Buffer	0.847 ± 0.103	1.07 ± 0.100	1.34 ± 0.014
TNFα	9.31 ± 2.82	7.56 ± 1.22	9.12 ± 0.587
C277A Y315E	4.19 ± 3.45	4.48 ± 0.273	3.94 ± 0.632
Y315E L508W	6.86 ± 0.441	4.26 ± 0.292	3.76 ± 0.200
WT	3.99 ± 0.176	2.13 ± 0.270	3.25 ± 0.093
L312Y F316T	4.47 ± 0.0983	3.24 ± 0.0864	3.21 ± 0.046
L312Y	5.03 ± 0.102	4.44 ± 0.0886	3.52 ± 0.744

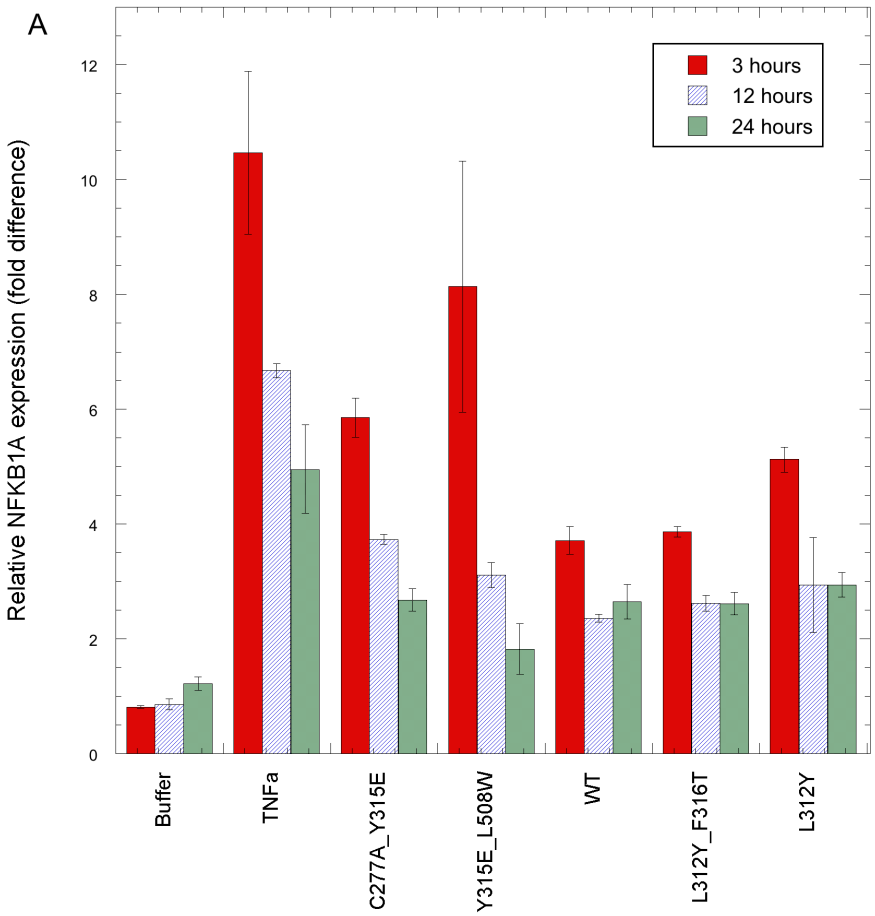


Figure 3-1. Normalized relative expression of NFKB1A in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 1 (Invitrogen). Values are relative to untreated cells, error bars indicate standard error of means.

Table 3-3. Matrix of relative differences in NFKB1A expression determined to be significant after 3 hours of treatment in experiment 1 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							
C277A Y315E							
Y315E L508W							
WT							
L312Y F316T							X
L312Y							
NFKB1 1_3hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

Table 3-4. Matrix of relative differences in NFKB1A expression determined to be significant after 12 hours of treatment in experiment 1 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							X
C277A Y315E							
Y315E L508W							
WT						X	
L312Y F316T							
L312Y							
NFKB1 1_12hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

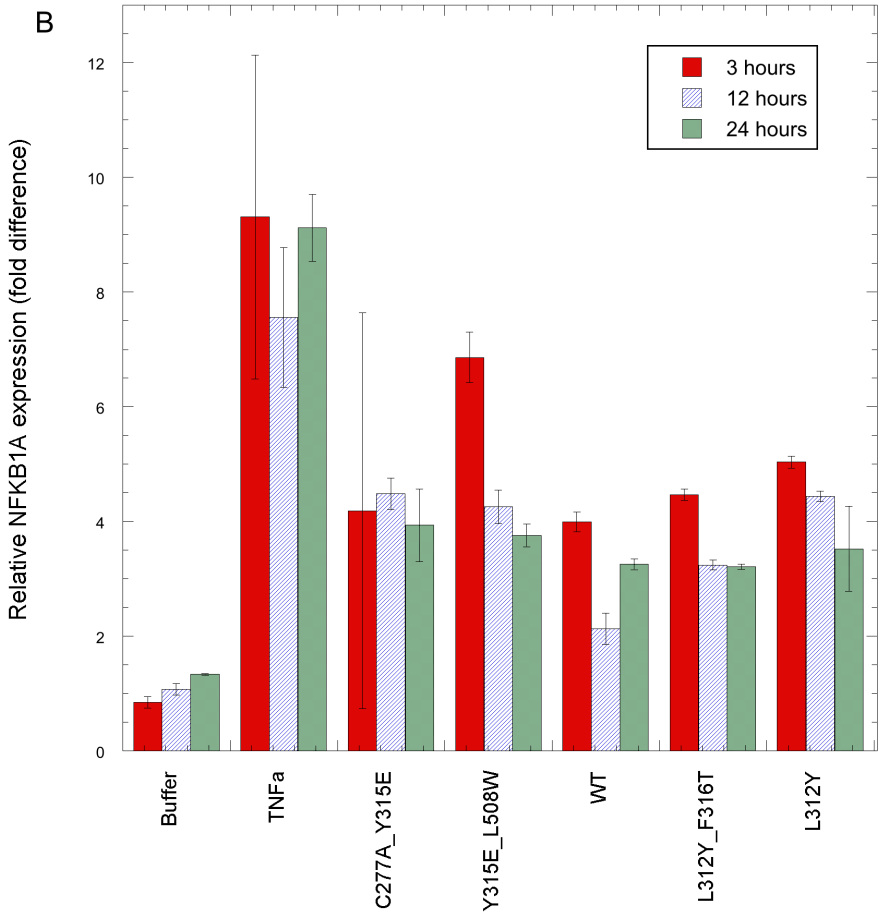


Figure 3-2. Normalized relative expression of NFKB1A in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 2 (Lonza). Values are relative to untreated cells, error bars indicate standard error of means.

Table 3-5. Matrix of relative differences in NFKB1A expression determined to be significant after 3 hours of treatment in experiment 2 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							
C277A Y315E							
Y315E L508W							
WT							
L312Y F316T							X
L312Y							X
NFKB1A 2_3hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

Table 3-6. Matrix of relative differences in NFKB1A expression determined to be significant after 12 hours of treatment in experiment 2 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							
C277A Y315E							
Y315E L508W							
WT							
L312Y F316T							
L312Y							X
NFKB1A 2_12hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

Table 3-7. Matrix of relative differences in NFKB1A expression determined to be significant after 24 hours of treatment in experiment 2 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							
C277A Y315E							
Y315E L508W							
WT							
L312Y F316T							X
L312Y							
NFKB1A 2_24hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

Table 3-8. Normalized expression of SELE under different treatment conditions after 3, 12 and 24 hours' exposure, with standard error of means. (A) Experiment 1. (B) Experiment 2. Values are relative to untreated cells.

A	3 hours	12 hours	24 hours
Buffer	0.861 ± 0.0537	1.25 ± 0.143	1.85 ± 0.123
TNFα	245 ± 33.7	150 ± 8.15	91.3 ± 5.69
C277A Y315E	164 ± 9.65	13.3 ± 0.356	3.05 ± 0.0977
Y315E L508W	287 ± 67.4	11.7 ± 1.69	2.84 ± 0.0903
WT	55.4 ± 4.18	8.29 ± 0.288	4.80 ± 0.545
L312Y F316T	63.5 ± 1.94	12.9 ± 1.35	4.73 ± 0.413
L312Y	119 ± 1.67	12.4 ± 2.47	2.34 ± 0.188

B	3 hours	12 hours	24 hours
Buffer	0.588 ± 0.155	1.13 ± 0.241	2.24 ± 0.0231
TNFα	386 ± 39.3	413 ± 28.9	130 ± 24.1
C277A Y315E	95.9 ± 15.9	32.7 ± 1.96	9.61 ± 0.986
Y315E L508W	404 ± 24.4	21.1 ± 0.548	5.70 ± 0.235
WT	101 ± 2.77	14.4 ± 1.74	6.45 ± 0.175
L312Y F316T	146 ± 3.18	16.6 ± 0.0742	7.63 ± 0.111
L312Y	230 ± 6.26	17.4 ± 0.177	5.24 ± 1.88

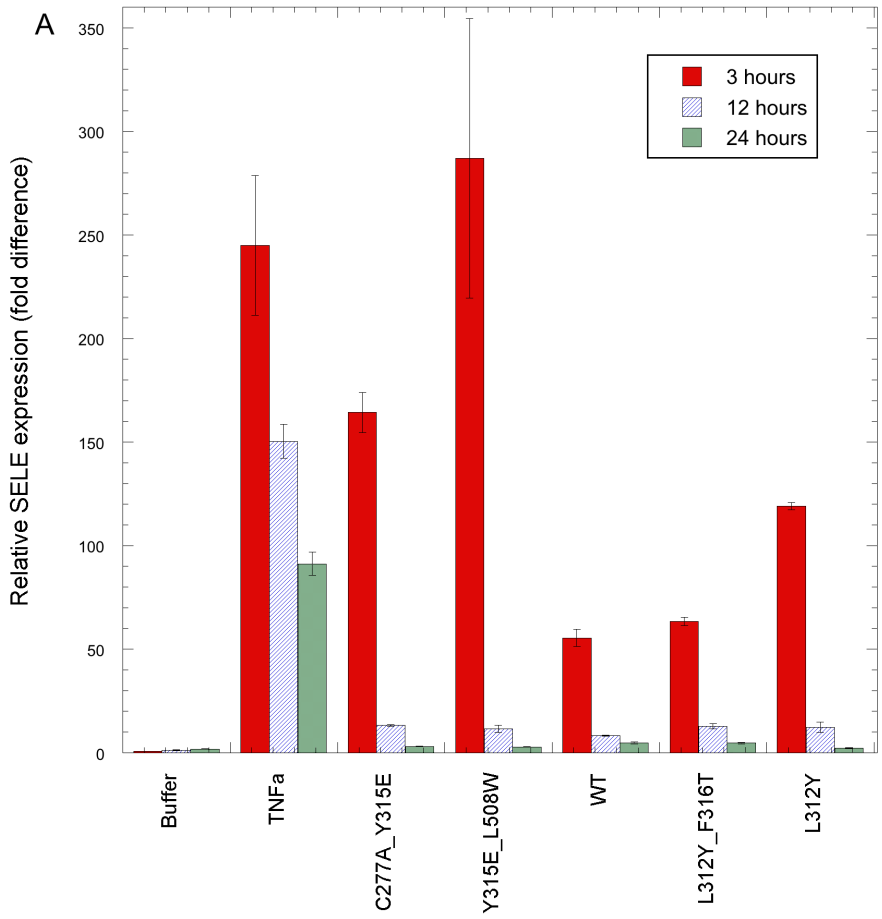


Figure 3-3. Normalized relative expression of SELE in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 1 (Invitrogen). Values are relative to untreated cells, error bars indicate standard error of means.

Table 3-9. Matrix of relative differences in SELE expression determined to be significant after 3 hours of treatment in experiment 1 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							
C277A Y315E							
Y315E L508W							
WT							
L312Y F316T							X
L312Y							X
SELE 1_3hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

Table 3-10. Matrix of relative differences in SELE expression determined to be significant after 12 hours of treatment in experiment 1 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							
C277A Y315E							X
Y315E L508W							
WT							
L312Y F316T							
L312Y							
SELE 1_12hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

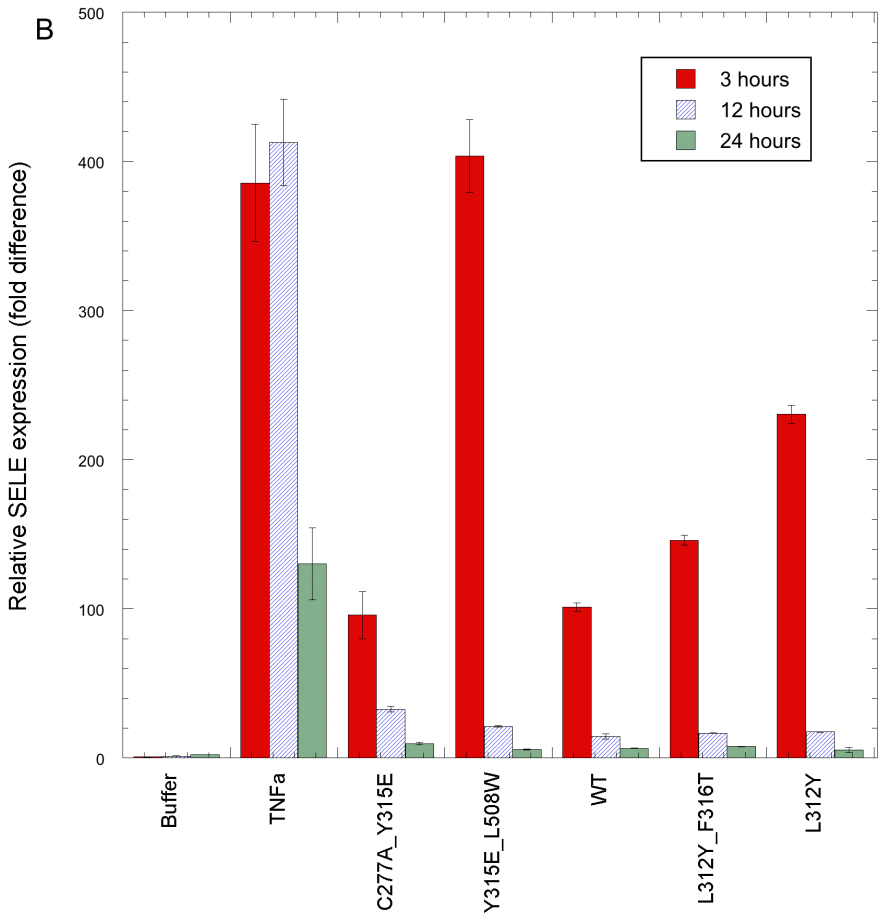


Figure 3-4. Normalized relative expression of SELE in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 2 (Lonza). Values are relative to untreated cells, error bars indicate standard error of means.

Table 3-11. Matrix of relative differences in SELE expression determined to be significant after 3 hours of treatment in experiment 2 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							
C277A Y315E							
Y315E L508W							
WT							X
L312Y F316T							X
L312Y							X
SELE 2_3hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

Table 3-12. Matrix of relative differences in SELE expression determined to be significant after 12 hours of treatment in experiment 2 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							
C277A Y315E							
Y315E L508W							X
WT							
L312Y F316T							X
L312Y							X
SELE 2_12hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

Table 3-13. Matrix of relative differences in SELE expression determined to be significant after 24 hours of treatment in experiment 2 using the Bonferroni correction where the $p<0.05$ was considered significant.

Buffer							
TNFα							
C277A Y315E							
Y315E L508W							
WT							X
L312Y F316T							X
L312Y							
SELE 2_24hr	L312Y	L312Y F316T	WT	Y315E L508W	C277A Y315E	TNFα	Buffer

Table 3-14. Normalized expression of RHOB under different treatment conditions after 3, 12 and 24 hours' exposure, with standard error of means. (A) Experiment 1. (B) Experiment 2. Values are relative to untreated cells.

A	3 hours	12 hours	24 hours
Buffer	1.02 ± 0.0208	1.15 ± 0.161	1.03 ± 0.0720
TNFα	2.46 ± 0.192	1.87 ± 0.0569	1.36 ± 0.303
C277A Y315E	1.80 ± 0.0401	1.45 ± 0.0430	1.23 ± 0.0539
Y315E L508W	2.39 ± 0.396	1.40 ± 0.166	1.40 ± 0.0309
WT	1.49 ± 0.0961	1.21 ± 0.0114	1.44 ± 0.180
L312Y F316T	1.37 ± 0.0315	1.21 ± 0.0811	1.41 ± 0.0930
L312Y	1.68 ± 0.0376	1.25 ± 0.332	1.42 ± 0.107

B	3 hours	12 hours	24 hours
Buffer	0.921 ± 0.262	1.03 ± 0.0364	1.15 ± 0.0217
TNFα	2.19 ± 0.243	1.82 ± 0.185	2.29 ± 0.273
C277A Y315E	1.94 ± 0.361	1.35 ± 0.0658	1.46 ± 0.137
Y315E L508W	1.87 ± 0.0595	1.41 ± 0.0192	1.55 ± 0.153
WT	1.45 ± 0.0559	1.64 ± 0.220	1.49 ± 0.036
L312Y F316T	1.61 ± 0.0841	1.29 ± 0.0259	1.33 ± 0.120
L312Y	1.55 ± 0.0544	1.43 ± 0.0904	1.26 ± 0.321

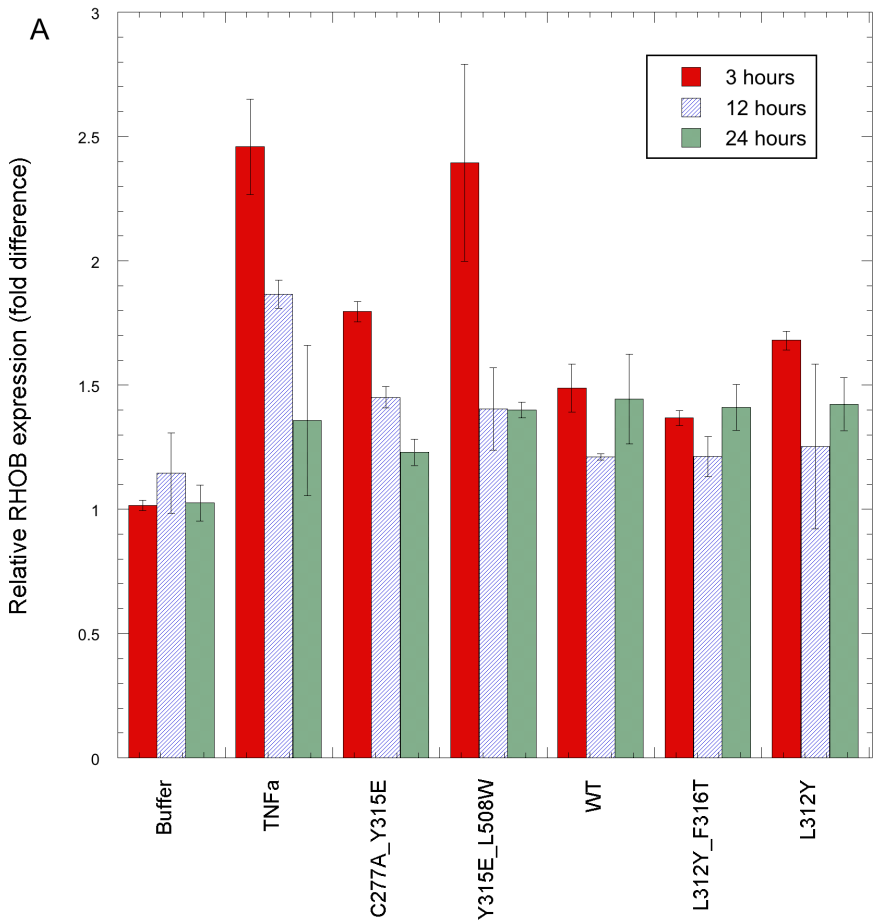


Figure 3-5. Normalized relative expression of RHOB in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 1 (Invitrogen). Values are relative to untreated cells, error bars indicate standard error of means.

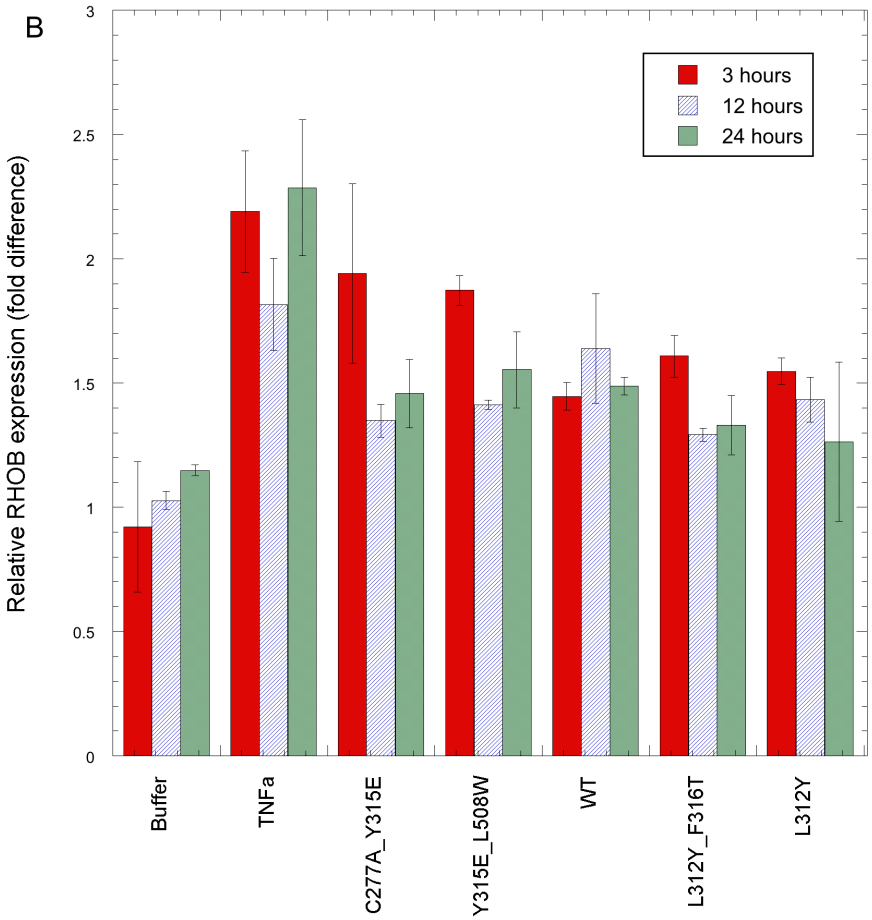


Figure 3-6. Normalized relative expression of RHOB in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 2 (Lonza). Values are relative to untreated cells, error bars indicate standard error of means.

Table 3-15. Normalized expression of ITGB1 under different treatment conditions after 3, 12 and 24 hours' exposure, with standard error of means. (A) Experiment 1. (B) Experiment 2. Values are relative to untreated cells.

A	3 hours	12 hours	24 hours
Buffer	1.05 ± 0.0295	0.997 ± 0.111	1.25 ± 0.0833
TNFα	0.360 ± 0.0522	1.19 ± 0.0248	1.16 ± 0.0582
C277A Y315E	1.09 ± 0.0244	1.26 ± 0.0321	1.41 ± 0.0313
Y315E L508W	1.12 ± 0.141	1.07 ± 0.0772	1.59 ± 0.0369
WT	1.07 ± 0.0529	1.11 ± 0.00972	1.65 ± 0.191
L312Y F316T	0.818 ± 0.146	1.27 ± 0.0466	0.708 ± 0.0566
L312Y	1.07 ± 0.0186	1.26 ± 0.159	1.55 ± 0.115

A	3 hours	12 hours	24 hours
Buffer	0.833 ± 0.134	0.867 ± 0.0672	0.932 ± 0.0135
TNFα	0.608 ± 0.0593	0.959 ± 0.0555	1.09 ± 0.135
C277A Y315E	0.494 ± 0.0940	1.07 ± 0.0622	1.13 ± 0.0947
Y315E L508W	0.928 ± 0.0414	1.07 ± 0.0136	1.17 ± 0.0630
WT	0.876 ± 0.0427	0.763 ± 0.0889	1.02 ± 0.0333
L312Y F316T	0.925 ± 0.0370	0.955 ± 0.00399	1.01 ± 0.0230
L312Y	0.885 ± 0.0422	0.992 ± 0.00541	0.723 ± 0.0671

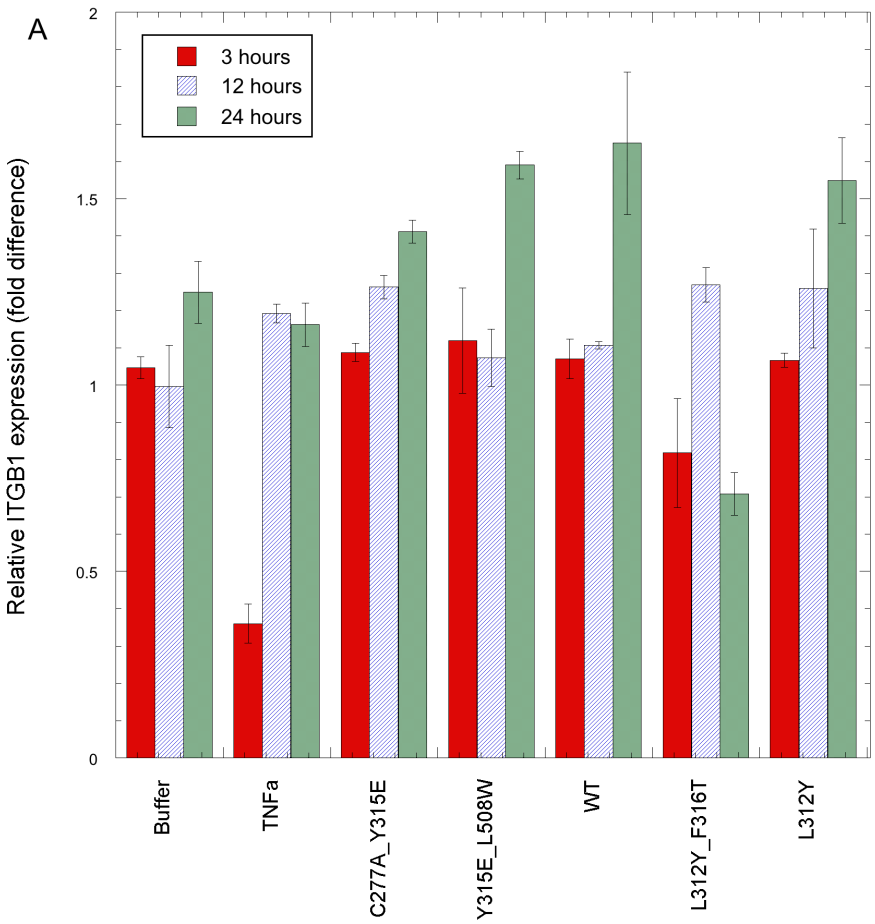


Figure 3-7. Normalized relative expression of ITGB1 in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 1 (Invitrogen). Values are relative to untreated cells, error bars indicate standard error of means.

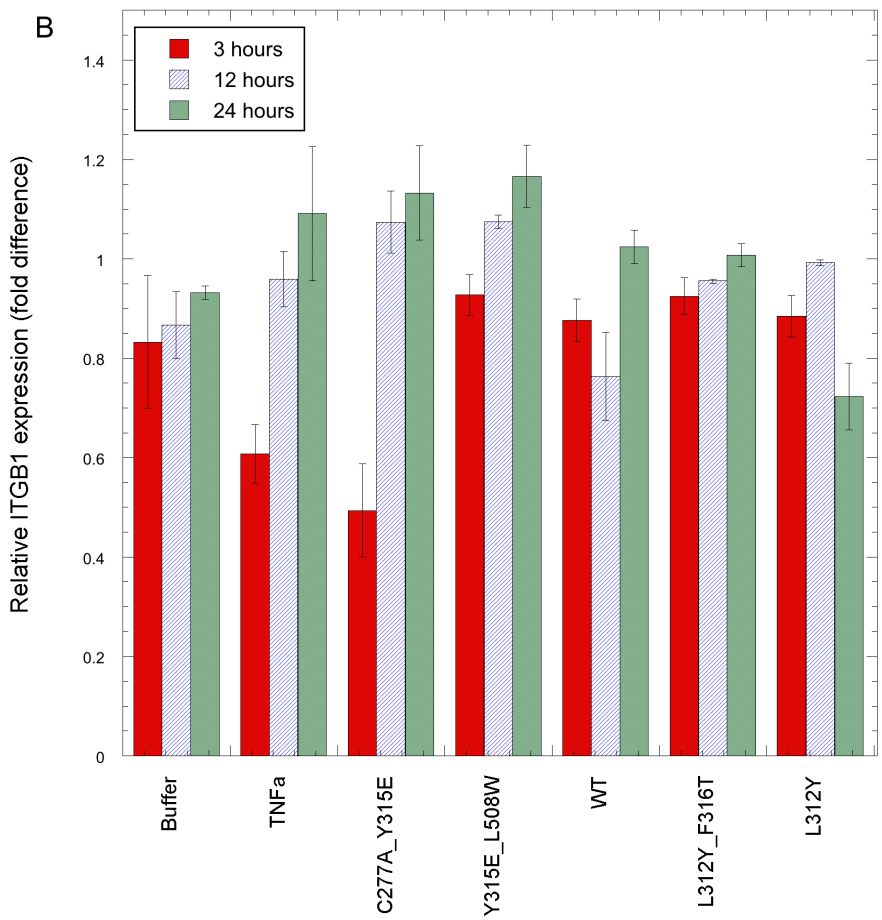


Figure 3-8. Normalized relative expression of ITGB1 in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 2 (Lonza). Values are relative to untreated cells, error bars indicate standard error of means.

Table 3-16. Normalized expression of FN1 under different treatment conditions after 3, 12 and 24 hours' exposure, with standard error of means. (A) Experiment 1. (B) Experiment 2. Values are relative to untreated cells.

A	3 hours	12 hours	24 hours
Buffer	0.803 ± 0.0161	0.799 ± 0.0871	1.04 ± 0.0609
TNFα	0.755 ± 0.0574	0.892 ± 0.0330	0.973 ± 0.0409
C277A Y315E	1.04 ± 0.0615	0.938 ± 0.0147	1.04 ± 0.0381
Y315E L508W	0.986 ± 0.128	0.832 ± 0.0463	1.33 ± 0.0680
WT	0.944 ± 0.0414	0.970 ± 0.0292	1.26 ± 0.160
L312Y F316T	1.01 ± 0.0402	0.960 ± 0.0358	1.30 ± 0.0816
L312Y	1.03 ± 0.0253	1.06 ± 0.161	1.39 ± 0.115

A	3 hours	12 hours	24 hours
Buffer	0.966 ± 0.170	1.28 ± 0.0879	1.30 ± 0.0415
TNFα	0.813 ± 0.0805	1.04 ± 0.101	1.22 ± 0.141
C277A Y315E	0.571 ± 0.133	1.10 ± 0.0561	1.19 ± 0.152
Y315E L508W	1.40 ± 0.0787	1.10 ± 0.0753	1.56 ± 0.0370
WT	1.27 ± 0.0646	0.960 ± 0.117	1.47 ± 0.110
L312Y F316T	1.39 ± 0.0542	1.15 ± 0.0172	1.40 ± 0.0179
L312Y	1.28 ± 0.0252	1.18 ± 0.0529	0.889 ± 0.101

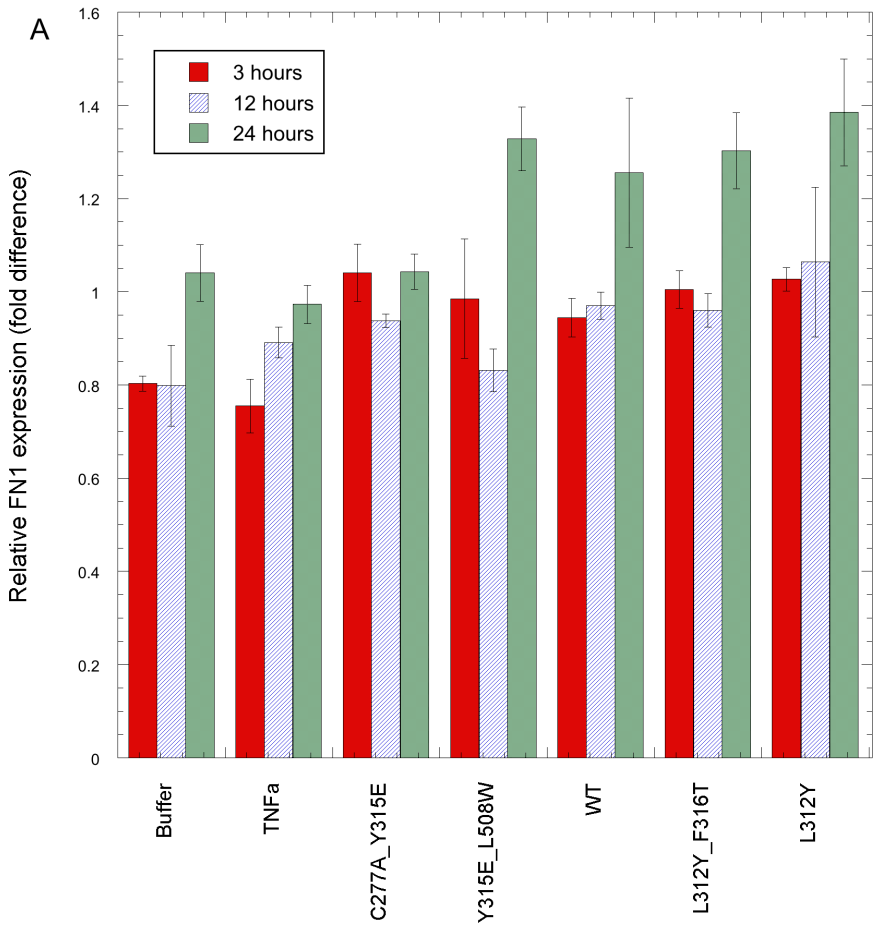


Figure 3-9. Normalized relative expression of FN1 in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 1 (Invitrogen). Values are relative to untreated cells, error bars indicate standard error of means.

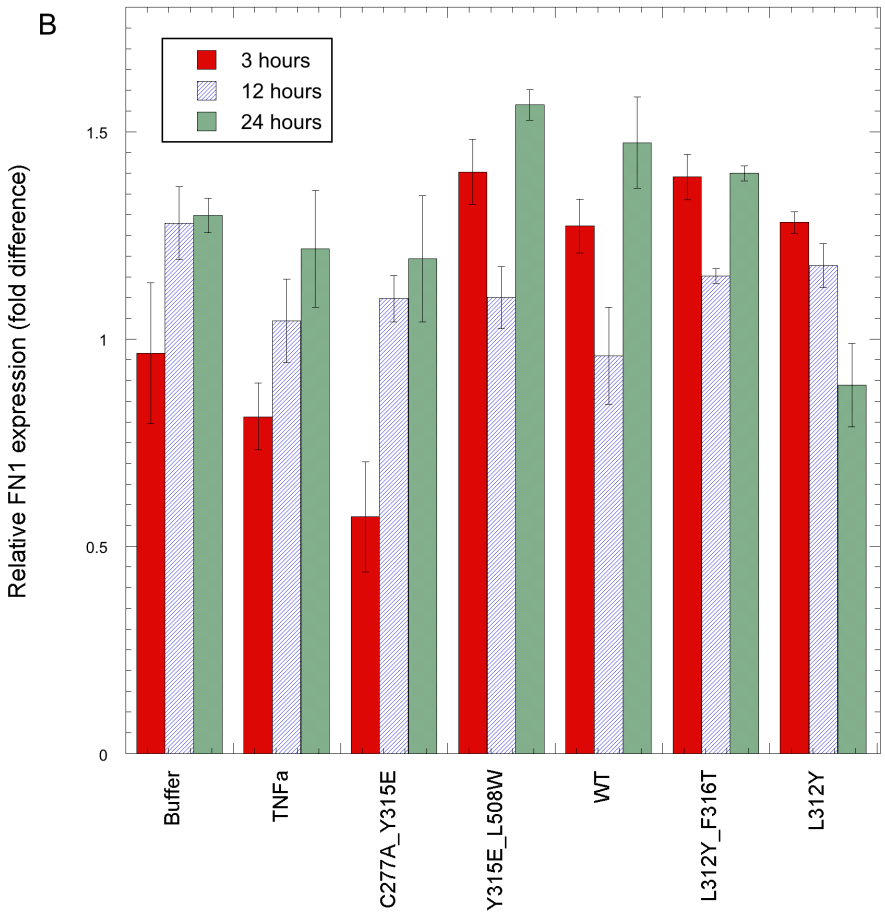


Figure 3-10. Normalized relative expression of FN1 in HUVECs treated with indicated proteins for 3 hours [52], 12 hours (blue), or 24 hours (green) in experiment 2 (Lonza). Values are relative to untreated cells, error bars indicate standard error of means.

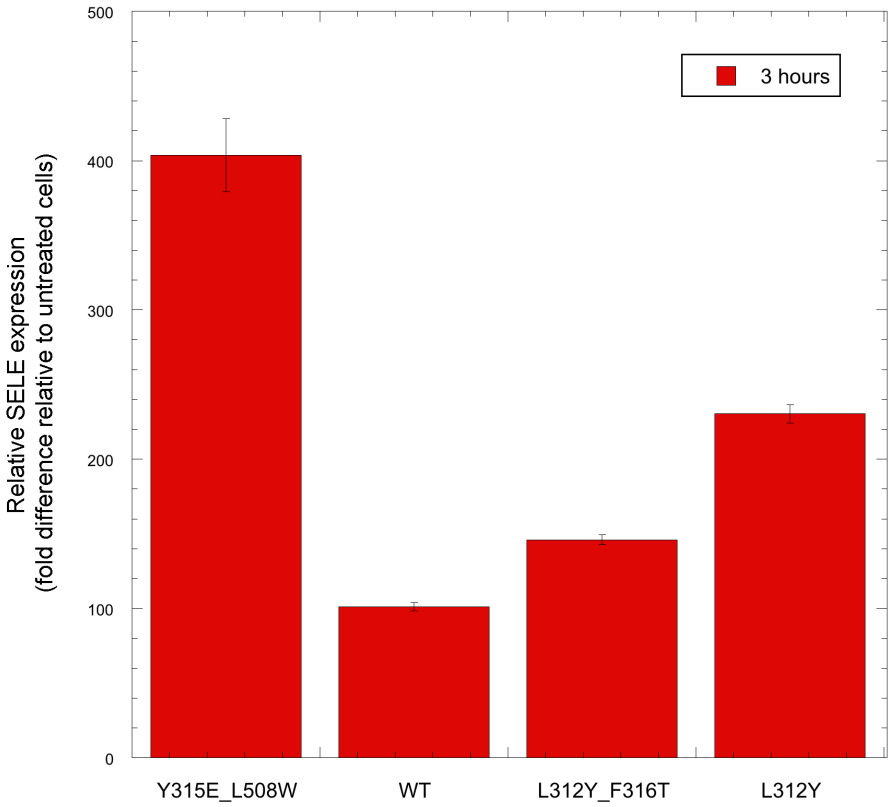


Figure 3-11. Normalized relative expression of SELE in HUVECs after 3 hours of treatment with the indicated variant of TG2 in experiment 2. Values are relative to untreated cells.

Table 3-17. Matrix of the pairwise comparisons of relative differences in SELE expression determined to be statistically significant in experiment 2 when limited to the three designs indicated and WT TG2.

Y315E L508W				
WT	X			
L312Y F316T	X	X		
L312Y		X	X	
	Y315E L508W	WT	L312Y F316T	L312Y

References

1. Sollid, L.M. and B. Jabri, *Is celiac disease an autoimmune disorder?* Curr Opin Immunol, 2005. **17**(6): p. 595-600.

2. Koning, F., et al., *Pathomechanisms in celiac disease.* Best Pract Res Clin Gastroenterol, 2005. **19**(3): p. 373-87.

3. Dieterich, W., et al., *Identification of tissue transglutaminase as the autoantigen of celiac disease.* Nat Med, 1997. **3**(7): p. 797-801.

4. Quarsten, H., et al., *HLA binding and T cell recognition of a tissue transglutaminase-modified gliadin epitope.* Eur J Immunol, 1999. **29**(8): p. 2506-14.

5. Sollid, L.M. and E. Thorsby, *HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis.* Gastroenterology, 1993. **105**(3): p. 910-22.

6. Nilsen, E.M., et al., *Gluten specific, HLA-DQ restricted T cells from coeliac mucosa produce cytokines with Th1 or Th0 profile dominated by interferon gamma.* Gut, 1995. **37**(6): p. 766-76.

7. Nilsen, E.M., et al., *Gluten induces an intestinal cytokine response strongly dominated by interferon gamma in patients with celiac disease.* Gastroenterology, 1998. **115**(3): p. 551-63.

8. Alaedini, A. and P.H. Green, *Narrative review: celiac disease: understanding a complex autoimmune disorder.* Ann Intern Med, 2005. **142**(4): p. 289-98.

9. Maki, M., *The humoral immune system in coeliac disease.* Baillieres Clin Gastroenterol, 1995. **9**(2): p. 231-49.

10. Caja, S., et al., *Antibodies in celiac disease: implications beyond diagnostics.* Cell Mol Immunol, 2011. **8**(2): p. 103-9.

11. Korponay-Szabo, I.R., et al., *In vivo targeting of intestinal and extraintestinal transglutaminase 2 by coeliac autoantibodies.* Gut, 2004. **53**(5): p. 641-8.

12. Salmi, T.T., et al., *Endomysial antibody-negative coeliac disease: clinical characteristics and intestinal autoantibody deposits.* Gut, 2006. **55**(12): p. 1746-53.

13. Rantala, I., et al., *Periodate-lysine-paraformaldehyde as fixative for the study of duodenal mucosa. Morphologic and immunohistochemical results at light and electron microscopic levels.* Acta Pathol Microbiol Immunol Scand A, 1985. **93**(4): p. 165-73.

14. Halttunen, T. and M. Maki, *Serum immunoglobulin A from patients with celiac disease inhibits human T84 intestinal crypt epithelial cell differentiation.* Gastroenterology, 1999. **116**(3): p. 566-72.

15. Cervio, E., et al., *Sera of patients with celiac disease and neurologic disorders evoke a mitochondrial-dependent apoptosis in vitro.* Gastroenterology, 2007. **133**(1): p. 195-206.

16. Di Simone, N., et al., *Anti-tissue transglutaminase antibodies from celiac patients are responsible for trophoblast damage via apoptosis in vitro.* Am J Gastroenterol, 2010. **105**(10): p. 2254-61.

17. Myrsky, E., et al., *Celiac disease IgA modulates vascular permeability in vitro through the activity of transglutaminase 2 and RhoA.* Cell Mol Life Sci, 2009. **66**(20): p. 3375-85.

18. Myrsky, E., et al., *Coeliac disease-specific autoantibodies targeted against transglutaminase 2 disturb angiogenesis.* Clin Exp Immunol, 2008. **152**(1): p. 111-9.

19. Caja, S., et al., *Inhibition of transglutaminase 2 enzymatic activity ameliorates the anti-angiogenic effects of coeliac disease autoantibodies.* Scand J Gastroenterol, 2010. **45**(4): p. 421-7.

20. Martucciello, S., et al., *RhoB is associated with the anti-angiogenic effects of celiac patient transglutaminase 2-targeted autoantibodies.* J Mol Med (Berl), 2012. **90**(7): p. 817-26.

21. Castellanos-Rubio, A., et al., *Angiogenesis-related gene expression analysis in celiac disease.* Autoimmunity, 2012. **45**(3): p. 264-70.

22. Kalliokoski, S., et al., *Celiac Disease-Specific TG2-Targeted Autoantibodies Inhibit Angiogenesis and in Mice by Interfering with Endothelial Cell Dynamics.* PLoS One, 2013. **8**(6): p. e65887.

23. Ensari, A., et al., *Time-course of adhesion molecule expression in rectal mucosa of gluten-sensitive subjects after gluten challenge.* Clin Exp Immunol, 1993. **92**(2): p. 303-7.

24. Jelinkova, L., et al., *Increased levels of circulating ICAM-1, E-selectin, and IL-2 receptors in celiac disease.* Dig Dis Sci, 2000. **45**(2): p. 398-402.

25. Di Sabatino, A., et al., *Increased expression of mucosal addressin cell adhesion molecule 1 in the duodenum of patients with active celiac disease is associated with depletion of integrin alpha4beta7-positive T cells in blood.* Hum Pathol, 2009. **40**(5): p. 699-704.

26. Kiraly, R., et al., *Coeliac autoantibodies can enhance transamidating and inhibit GTPase activity of tissue transglutaminase: dependence on reaction environment and enzyme fitness.* J Autoimmun, 2006. **26**(4): p. 278-87.

27. Simon-Vecsei, Z., et al., *A single conformational transglutaminase 2 epitope contributed by three domains is critical for celiac antibody binding and effects.* Proc Natl Acad Sci U S A, 2012. **109**(2): p. 431-6.

28. Hayden, M.S. and S. Ghosh, *NF-kappaB, the first quarter-century: remarkable progress and outstanding questions.* Genes Dev, 2012. **26**(3): p. 203-34.

29. Chen, L.F. and W.C. Greene, *Shaping the nuclear action of NF-kappaB.* Nat Rev Mol Cell Biol, 2004. **5**(5): p. 392-401.

30. Maiuri, M.C., et al., *Nuclear factor kappa B is activated in small intestinal mucosa of celiac patients.* J Mol Med (Berl), 2003. **81**(6): p. 373-9.

31. Mann, A.P., et al., *Overexpression of tissue transglutaminase leads to constitutive activation of nuclear factor-kappaB in cancer cells: delineation of a novel pathway.* Cancer Res, 2006. **66**(17): p. 8788-95.

32. Kim, D.S., et al., *Reversal of drug resistance in breast cancer cells by transglutaminase 2 inhibition and nuclear factor-kappaB inactivation.* Cancer Res, 2006. **66**(22): p. 10936-43.

33. Ai, L., et al., *Ataxia-Telangiectasia, Mutated (ATM)/Nuclear Factor kappa light chain enhancer of activated B cells (NFkappaB) signaling controls basal and DNA damage-induced transglutaminase 2 expression.* J Biol Chem, 2012. **287**(22): p. 18330-41.

34. Wu, K., et al., *Genotoxic effect and nitrative DNA damage in HepG2 cells exposed to aristolochic acid.* Mutat Res, 2007. **630**(1-2): p. 97-102.

35. Wu, Z.H. and S. Miyamoto, *Many faces of NF-kappaB signaling induced by genotoxic stress.* J Mol Med (Berl), 2007. **85**(11): p. 1187-202.

36. Mackay, F., et al., *Tumor necrosis factor alpha (TNF-alpha)-induced cell adhesion to human endothelial cells is under dominant control of one TNF receptor type, TNF-R55.* J Exp Med, 1993. **177**(5): p. 1277-86.

37. Bevilacqua, M.P., et al., *Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils related to complement regulatory proteins and lectins.* Science, 1989. **243**(4895): p. 1160-5.

38. Lee, J., et al., *Transglutaminase 2 induces nuclear factor-kappaB activation via a novel pathway in BV-2 microglia.* J Biol Chem, 2004. **279**(51): p. 53725-35.

39. Park, S.S., et al., *Transglutaminase 2 mediates polymer formation of I-kappaBalpha through C-terminal glutamine cluster.* J Biol Chem, 2006. **281**(46): p. 34965-72.

40. Kim, D.S., et al., *I-kappaBalpha depletion by transglutaminase 2 and mu-calpain occurs in parallel with the ubiquitin-proteasome pathway.* Biochem Biophys Res Commun, 2010. **399**(2): p. 300-6.

41. Kumar, S. and K. Mehta, *Tissue transglutaminase constitutively activates HIF-1alpha promoter and nuclear factor-kappaB via a non-canonical pathway.* PLoS One, 2012. **7**(11): p. e49321.

42. Jones, R.A., et al., *Matrix changes induced by transglutaminase 2 lead to inhibition of angiogenesis and tumor growth.* Cell Death Differ, 2006. **13**(9): p. 1442-53.

43. Bottero, V., et al., *Monitoring NF-kappa B transactivation potential via real-time PCR quantification of I kappa B-alpha gene expression.* Mol Diagn, 2003. **7**(3-4): p. 187-94.

44. Ghosh, S., M.J. May, and E.B. Kopp, *NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses.* Annu Rev Immunol, 1998. **16**: p. 225-60.

45. D'Argenio, G., et al., *Human serum transglutaminase and coeliac disease: correlation between serum and mucosal activity in an experimental model of rat small bowel enteropathy.* Gut, 1989. **30**(7): p. 950-4.

46. Esposito, C., et al., *Expression and enzymatic activity of small intestinal tissue transglutaminase in celiac disease.* Am J Gastroenterol, 2003. **98**(8): p. 1813-20.

47. Kojima, S., K. Nara, and D.B. Rifkin, *Requirement for transglutaminase in the activation of latent transforming growth factor-beta in bovine endothelial cells.* J Cell Biol, 1993. **121**(2): p. 439-48.

48. Balklava, Z., et al., *Analysis of tissue transglutaminase function in the migration of Swiss 3T3 fibroblasts: the active-state conformation of the enzyme does not affect cell motility but is important for its secretion.* J Biol Chem, 2002. **277**(19): p. 16567-75.

49. Akimov, S.S., et al., *Tissue transglutaminase is an integrin-binding adhesion coreceptor for fibronectin.* J Cell Biol, 2000. **148**(4): p. 825-38.

50. Schindler, U. and V.R. Baichwal, *Three NF-kappa B binding sites in the human E-selectin gene required for maximal tumor necrosis factor alpha-induced expression.* Molecular and Cellular Biology, 1994. **14**(9): p. 5820-31.

51. Zanoni, G., et al., *In celiac disease, a subset of autoantibodies against transglutaminase binds toll-like receptor 4 and induces activation of monocytes.* PLoS Med, 2006. **3**(9): p. e358.

52. Rodolfo, C., et al., *Tissue transglutaminase is a multifunctional BH3-only protein.* Journal of Biological Chemistry, 2004. **279**(52): p. 54783-92.

A p p e n d i x A

SITE-SATURATION MUTAGENESIS OF THE ACTIVE SITE OF A DEISGNED KEMP ELIMINATION ENZYME

Abstract

Enzymes are biological super catalysts. They are highly specific, can function at biological pH and temperature, and accelerate reactions with rate enhancements of 10^5–10^{17}. Enzymes have practical applications as medicinal therapeutics and as catalysts for industrial processes, but enzymes available for such uses have historically been limited to what is found in nature. Therefore, the *de novo* design of enzymes with novel activities has become a major scientific goal. Previous work in the Mayo lab resulted in the successful design of a *de novo* enzyme catalyzing the Kemp elimination: HG2/S265T. This enzyme was designed using the protein design software PHOENIX and molecular dynamics (MD) simulations, and has a k_{cat}/K_m of 430 $M^{-1}s^{-1}$. Although quite an accomplishment, this value is still far lower than the 10^7 $M^{-1}s^{-1}$ values typically found in nature. To further increase the activity of this enzyme and to identify weaknesses in our computational design methods, we performed site-saturation mutagenesis on the active site of HG2/S265T. The resulting 209 mutants were expressed and screened for Kemp elimination activity. Of the 209 mutants screened, 22 were exhibited higher activity than WT protein. After purification and kinetic characterization, four were found to be more active than HG2/S265T: K50M, M172I, M84I, and K50Q. K50M saw the largest overall increase in activity, with a k_{cat}/K_m of 2.5 x 10^3. Using these results to optimize our design parameters so that the predictions made by our computational methods better match that which is observed *in vitro* will improve our model, furthering our long term goal of designing *de novo* enzymes with k_{cat}/K_m values closer to those found in nature.

Introduction

Enzymes

All living organisms must be able to self-replicate and selectively catalyze chemical reactions. Biological systems employ enzymes, which are usually specialized proteins, as catalysts for their necessary reactions (the exception being catalytic RNA molecules). Enzymes are far superior to synthetic and inorganic catalysts: they bind their substrates with high specificity, function in an aqueous environment (often under mild pH and temperature conditions), and accelerate chemical reactions several orders of magnitude (rate enhancements typically range from 10^5 to 10^{17}) [1].

Enzymes are of practical interest in medicine and industry. Many diseases are caused by the lack of a necessary enzyme or by some aberrant enzymatic activity. Consequently, enzymes or molecules that inhibit them are often used as therapeutics. The special properties of these molecules also make them very attractive in the chemical industry. Their high specificity, ability to yield products of a single enantiomer, activity at low temperatures, and impressive rate enhancements make enzymes much more cost effective than inorganic and synthetic catalysts, which often require extreme temperatures to accelerate reactions minimally and with low specificity.

Enzymes available for use as drugs and to catalyze industrial processes have historically been limited to what can be found in nature. Protein engineering has been used to improve the stability or specificity of natural enzymes, but the *de novo* design of enzymes with novel activities is much more challenging and has become a major scientific goal.

Enzyme Design

Early work to design a *de novo* enzyme focused on the Kemp elimination, not because it is an important reaction for industry or medicine, but because it is a well-characterized reaction with no natural protein catalyst and is a realistic first time *de novo* enzyme design goal [2-4]. The Kemp elimination is a simple reaction involving the base-catalyzed abstraction of a proton from the C3 position of 5-nitrobenzisoxazole to yield 2-cyano-4-nitrophenolate (Figure A-1). These design attempts were based on the work of Thorn *et al.*, who created a catalytic Kemp elimination antibody, 34E3, by generating charge complementarity between an antibody and a benzimdazolium hapten. 34E3 has a k_{cat}/K_m of 5.5×10^3 $M^{-1}s^{-1}$ [5]. X-ray crystallography of 34E3 showed that

hydrogen bonding, π stacking, and van der Waals interactions positioned the substrate for catalysis. Specifically, the catalytic glutamate residue formed a bidentate salt bridge with the charged guanidinium group of the hapten substrate. Also, the hapten was stacked between the sidechains of two aromatic residues, a tryptophan and a tyrosine [6].

Enzyme Design in the Mayo Lab

Past work in the Mayo lab also focused on designing a *de novo* enzyme to catalyze the Kemp elimination [7]. To achieve this goal, a design strategy dubbed the protein design cycle was used. The design cycle consists of three main steps: (1) design, where protein sequences predicted to assume the desired three-dimensional structure are calculated using computational methods, (2) evaluation, where the sequences predicted by the design calculations are analyzed and tested both computationally and *in vitro* to see if they possess the desired characteristics, and (3) adjustment, where the results of the evaluations are used to modify the design procedure to garner improved results [7, 8].

Initial attempts at a de novo Kemp elimination enzyme produced HG1. This design was based on the scaffold 10A xylanase (TAX) and was performed using ORBIT, a computational protein design software program created in the Mayo lab. Unfortunately, HG1 was inactive. X-ray crystallography showed that the active site was very similar to the design; however, structural analysis and molecular dynamics (MD) simulations also indicated that the active site was flexible and very solvent exposed with many water molecules present. To avoid these problems, future designs positioned the active site more deeply within the protein scaffold.

A second round of design, this time using the design software PHOENIX (also developed in the Mayo lab), yielded HG2, with a k_{cat}/K_m of 122 $M^{-1}s^{-1}$. The design for HG2 included a more buried catalytic aspartate at position 127, a tryptophan residue at position 265 to π stack with the substrate, a lysine at position 50 to interact with the substrate nitro group, and a serine at position 265 to hydrogen bond with the isoxazole oxygen (Figure A-2). MD simulations indicated that although S265 did not make the desired hydrogen bond contact to the substrate, the serine was preventing water from entering the active site by pushing the substrate closer to the base.

MD was also successful in predicting the catalytic activity of three additional designs in two alternate scaffolds, and subsequently became an integral part of our design procedure. The mutant HG2/S265T was predicted to have improved activity by making the desired hydrogen bonding contact to the substrate and by creating better active site packing due to its larger size. HG2/S265T was constructed and found to have a k_{cat}/K_m of 430 $M^{-1}s^{-1}$, three-fold higher than any Kemp elimination enzyme designed thus far [7, 9].

Site-Saturation Mutagenesis of the HG2 Active Site

The following describes the site-saturation mutagenesis on the 11 residues surrounding the substrate in the active site of HG2. Each of the 11 residues was mutated to every other possible amino acid iteratively, resulting in a total of 209 mutants. These mutants were screened for Kemp elimination activity. The purpose of this experiment is two-fold. First, we hope to improve HG2's activity by discovering which mutations will result in a better enzyme. Secondly, these results will inform us as to how well our computational methods are able to accurately predict the optimal sequence or the outcome of a mutation for this design. In particular, these data will be used as a stringent screen of an MD protocol being developed in the Mayo lab to predict the activity of all 209 mutants that will be produced. Design parameters will then be optimized to reduce any disagreement between what is predicted and what is observed.

Methods

Choosing Active Site Residues for Mutagenesis

Active site residues were selected for mutagenesis by examining the active site of the HG-2 design. Every residue surrounding the transition state of the 5-nitrobenzoxizole Kemp elimination reaction (NBX) in the design was chosen for mutagenesis. Eleven residues were chosen in all: Ala21, Met42, Glu46, Met84, Met172, Met237, and Leu236, as well as those interacting with the substrate: Trp44, Lys50, Asp127, and Ser265 (or Thr265) (Figures A-3a and b).

Megaprimed Site-Directed Mutagenesis

Site-directed mutagenesis was performed using the megaprimed, ligase-free, PCR-based, site-directed mutagenesis method described in Tseng *et al.* [10]. This method is specialized for high-throughput mutagenesis because it only requires one mutagenic primer. For each mutant, one mutagenic forward primer and one universal flanking primer are used in an initial amplification reaction to generate complementary megaprimers that are then used to amplify the mutated plasmid as one would in a typical QuikChange reaction (Stratagene). The melting temperatures of the mutagenic forward primers were calculated using the equation:

$$T_m = 81.5 + 0.41(\%GC) - 675/N - \%mismatch \quad (1)$$

For the majority of the mutants, the generation of megaprimers followed by the amplification of mutated template was performed in a single reaction. PCR mixtures contained 25 µL 22 ng of template HG2/S265T in pET11a (Novagen), 100 nM of each primer, 200 µM of each deoxymucleotide triphosphate (dNTP), 1.25 U of *PfuTurbo* DNA polymerase, and *PfuTurbo* DNA polymerase buffer (Stratagene). The PCR program consisted of 95°C for 5 min; 18 cycles to generate megaprimer consisting of 95°C for 1 min, 60°C for 1 min, and 68°C for 2 min. Amplification of mutated template was then achieved by 18 cycles of 95°C for 1 min and 68°C for 20 min; a final extension at 68°C for 10 min completed the reaction.

Generation of megaprimers and amplification of mutated template was verified via agarose gel electrophoresis. If a reaction yielded megaprimer but failed to amplify template, a second reaction was run using the same mixture components as above except the primers were replaced by 1.25 µl

of the PCR products from the initial reaction. The PCR was then run under the following conditions: 95°C for 5 min; 18 cycles of 95°C for 1 min and 68°C for 20 min; and a final extension at 68°C for 10 min.

Upon completion of the PCR reactions, 3.5 µl of 10x NEBuffer 4, 3.18 µl Diluent Buffer B, and 66.4 U (3.32 µl) of *Dpn*I (New England Biolabs) were added to each reaction mixture and incubated at 37°C for 2 hours.

Transformation and Plating

Transformation and plating were performed using a Tecan Freedom EVO liquid handling robot by dividing the mutagenic primers over three 96-well plates. Mutagenesis and transformation was performed on one plate at a time. PCR reactions were transferred to 96-well PCR plates and 2 µl of each reaction was transformed into 20 µl of BL21 DE3 chemically competent cells by heatshock at 42°C for 45 sec in the robot thermocycler. 100 µl of LB (Luria-Bertani broth) was then added to each reaction, and the cultures were shaken at 1800 rpm in 30-sec intervals for 2 hours to keep cells suspended.

Cells were then plated onto 48 well LB/agar ampicillin qtrays and allowed to grow overnight at 37°C. Two colonies were picked for each mutant and grown for 14 hours at 37°C in 200 µl of LB/ampicillin with 10% glycerol in 96-well culture plates.

The plates were replicated into glycerol stocks, and the original culture plate was sent to Agencourt Biosciences for off-site mini-prepping and sequencing. If both colonies yielded wild-type sequences or undesired mutations, another colony from that mutant's well was picked, mini-prepped (Qiagen), and sequenced. If this attempt still yielded the incorrect sequence, or if no colonies had grown in the first place, the mutagenesis procedure was repeated for that mutant along with the next plate of mutants.

Protein expression and purification

The expression and purification of HG2 and its mutants was conducted exactly as described previously [11]. Briefly, for the screening of HG2 mutants, 200 µL LB/ampicillin grown overnight at 37°C with shaking after inoculation from frozen BL-21 (DE3) *E.coli* stocks. The full 200 µL of overnight culture was added to 5 mL of Overnight Express Instant TB media (Novagen) with ampicillin in 24-well culture plates (Whatman) with Bugstopper Venting Capmats (Whatman). Cultures were grown at 37°C with shaking, and then moved to 18°C with shaking to be grown overnight. Cells were harvested by centrifugation at 5000 x g for 10 minutes at 4°C. Each 100 mL pellet was resuspended in 8 mL of lysis buffer (50 mM sodium phosphate, pH 8.0, 0.3 M NaCl, 2.5 mM imidazole, 1x CelLytic B (Sigma-Aldrich), 5 mM β-mercaptoethanol (βME), 0.2 mg/mL lysozyme (Sigma-Aldrich), and 1 U Benzonase endonuclease (Merck)). The lysate was centrifuged at 19000 x g for 45 at 4°C and the supernatant collected.

For full-scale purification, proteins were expressed in 1 L LB/ampicillin cultures in BL-21 (DE3) *E.coli* inoculated with 10 mL overnight starter cultures grown at 37°C. The 1 L cultures were grown at 37°C until reaching OD_{600} ~ 0.6. Expression was induced upon the addition of 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown at 18°C for 18 hours. Cells were harvested by centrifugation at 5000 x g for 10 min at 4°C. The pellet was resuspended in 10 mL of equilibrium buffer (50 mM sodium phosphate with 0.3 M sodium chloride, pH 8.0), 1 U Benzonase, and both RNAse and Lysozyme and lysed by sonication (Misonix). The lysed cells were centrifuged at 15,000 x g for 40 min at 4°C. The supernatant was incubated at 4°C for 1 hour with 2 mL of affinity gel suspension (HIS-select) (Sigma-Aldrich) that had been equilibrated with 10 bed volumes of equilibrium buffer (50 mM sodium phosphate with 0.3 M sodium chloride, pH 8.0). The resin was washed in a gravity column with 10 bed volumes of Wash I buffer (50 mM sodium phosphate with 0.3 M sodium chloride, pH 8.0, 10 mM imidazole), and then again with 10 bed volumes of Wash II buffer (50 mM sodium phosphate with 0.3 M sodium chloride, pH 8.0, 20 mM imidazole). 3 mL of elution buffer (50 mM sodium phosphate with 0.3 M sodium chloride and 250 mM imidazole, pH 8.0) was then used to elute the protein. The eluate was concentrated and buffer exchanged using a Amicon Ultra -15 filter 10,000 MWCO centrifugal concentrators into storage buffer (50 mM sodium citrate pH 5.5, 150 mM NaCl).

Protein concentration determination

Protein concentrations were determined by UV absorbance denaturation in 8 M guanidinium hydrochloride for 5 minutes at a dilution of 10x to 100x. The extinction coefficient for each protein at 280 nm was calculated based on the number of tryptophans and tyrosines in the protein using ExPASy's ProtParam tool.

Kemp elimination activity screening

Cell lystates from the small scale purifications described above were used to screen for Kemp elimination by added 20 μL of lysate to 180 μL of assay buffer (25 mM HEPES, pH 7.25, 100 mM NaCl, and 0.8 mM 5-nitrobenzisoxazole (5-NBZ)). Reactions were conducted in black 96-well microtiter plate with clear bottom (Greiner) and the production of the phenolate product was monitored by an increase in absorbance at 380 nm using a Safire^2 microplate reader (Tecan). All mutants' activity were normalized to WT.

Kinetic measurements

Michaelis-Menten parameters were determined using 5 μM purified enzyme and monitoring phenolate production at 380 nm (ε = 15,800 $cm^{-1}M^{-1}$) [2] using a Shimadzu UV 1601 spectrophotometer at 27°C. Assays were conducted in assay buffer (25 mM HEPES, pH 7.25, 100 mM NaCl, and 0.031 to 1 mM 5-nitrobenzisoxazole (5-NBZ)). Initial rates were calculated from the first 300 sec of the reaction. The rate of reaction versus the corresponding substrate concentration was fit to the Michaelis-Menten equation (equation 2),

$$v = V_{max}[S]/K_m + [S] \quad (2)$$

where v is the initial reaction rate, V_{max} is the maximal rate of reaction, K_m is the Michaelis constant, and [S] is the substrate concentration.

PHOENIX energy calculations

PHOENIX energy calculations were conducted to score the mutants as was previously reported [7, 11]. Briefly, calculations were conducted using single state design based on the HG2/S265T design structure using occlusion-based solvation and the optimization algorithm FASTER [12]. Residues 16, 17, 46, 47, 50, 79, 87, 90, 170, 207, 209, 239, 275, and 276 were floated, allowing the WT sidechains to sample alternative conformations using a backbone independent library. One design position was designated per calculation and allowed to mutate to every possible amino acid. The calculations yielded a predicted structure file and a score for each mutant.

Results and Discussion

Site-Directed Mutagenesis

Mutagenesis reactions were initially intended to be carried out in 96-well PCR plates using Phusion HS polymerase (New England Biolabs) on the robot using the calculated heating setting on the thermocycler. These first attempts failed, however, and conditions were optimized by hand to what is listed above (*PfuTurbo* DNA polymerase, etc.). Additional experiments demonstrated that doing mutagenesis by hand using PCR tubes gave better results than running these reactions on the robot in PCR plates. Running reactions by hand also allowed for more flexibility when certain reactions had to be repeated from previous 96-well plates, or when a second stage of PCR was required to generate amplified template. Therefore, mutagenesis reactions were carried out by hand using PCR tubes. Further optimizations are necessary to successfully complete these reactions using the automated site-directed mutagenesis method.

HG2 Mutants

In all, 206 of the original 209 mutants were obtained (98.6%). Table 1 summarizes the number of mutants obtained for each culture plate. In total, four sets of culture plates were used. The average percent of mutants obtained from wells with colonies was 84.2%. Of the third colonies picked from a well after the first two sequenced at Agencourt did not yield the desired result, 28% from plate 1 and 50% of those from plate 2 were the correct sequence.

Third colonies were picked only for the first and second plate. The remaining mutants after three sets of plates were repeated from the mutagenesis step and sent on a fourth plate. This yielded a high success rate: 88.0% of these mutants were obtained. Based on these results, sending two colonies for sequencing appears to be optimal. This may suggest that if the correct sequence is not obtained within the first two colonies sequenced, the mutagenesis reaction has something inherently wrong with it.

The third plate was the most troublesome; only 63.3% of the mutants with colonies were correct, and the overall transformation was very poor. Two issues may have contributed to this: (1) 49 of the 88 reactions were mutants that had failed on previous plates, and (2) the amount of cells ran out during transformation, leaving some wells with very little competent bacteria. The failed reactions

were repeated using the two-round method of PCR, and gave much better results. Use of the two-round method for mutants that had trouble forming amplified template was generally helpful.

The mutagenesis reactions and transformations were repeated for the remaining three mutants of the original 209 (A21D, M84G, and M237S) until they were obtained.

Kemp elimination activity screening and kinetic characterization

Of the 209 mutants screened, 22 were exhibited higher activity than WT protein (Figure A-4). Six of the 22 mutations were also found to increase HG2/S265T activity in screens in the Hilvert lab, which used directed evolution to discover mutations that could improve the enzymes activity [13]. Since this was a screen using cell lysates, uncorrected for expression levels, 21 of these mutants were purified and kinetically characterized. Of those initial 22 mutants, 21 plus WT (HG2/S265T), scaffold (1GOR), and knockout D127V were purified and assayed. Of these 21 point mutants, four were found to be more active than HG2/S265T: K50M, M172I, M84I, and K50Q (Table A-2). Only one of these mutations, K50M, was used in HG3.17, the most active HG2/S265T evolved mutant found by the Hilvert lab [13].

The contradictory results between the screen and the assays are probably due to differential levels of expression between HG2/S265T and the mutants.

PHOENIX characterization of active site saturation mutants

The overall energy score for each mutant was relatively well anti-correlated to the k_{cat}/K_m value (Figure A-5). However, as expected, HG2/S265T was scored as the most stable of the sequences. Indicating that our method for predicting which sequence would best catalyze the Kemp elimination is definitely flawed. This is due to probably several factors, a few of which being: (a) inaccuracies the force field itself, (b) the scoring is based on a static picture of the enzyme and movement may play a part in catalysis, (c) inaccuracies or fundamentally flawed method for how we define the stabilization of the transition state, and especially (d) the recent finding that the substrate is flipped in the HG3.17 structure, indicating that for many of these mutants the orientation of the substrate may be defined incorrectly [13].

The flipped orientation of the substrate is very relevant, considering that the above authors found that in this configuration, K50Q assumes the designed role of S265T, and donates a hydrogen bond to the phenoxide leaving group [13]. This flipping of the substrate also makes it difficult to predict how the other beneficial mutations are enhancing the catalysis. Therefore, redoing the calculations in the newer program TRIAD with the new orientation of the substrate may yield a better anti-correlation between activity and energy score.

Multi-state design (MSD) and molecular dynamics have also been used to better model movement in catalysis and protein function [7, 14]. These techniques could also be used to improve our ability to predict which mutations would result in the best catalyst.

Conclusions

Designing the *de novo* Kemp elimination enzyme, HG2/S265T, with a k_{cat}/K_m of 430 $M^{-1}s^{-1}$ is already an incredible accomplishment. HG3.17 has a k_{cat}/K_m of 2.3 x 10^5 [13]. In this study, we were able to improve the k_{cat}/K_m of 2.5 x 10^3 using only computational design and active site saturation mutagenesis. Further studies should test the activities of mutants containing a combination of the most active mutants. However, the rate enhancement and efficiency of HG2/S265T, HG2/S265T/K50M, and HG3.17 does not approach those seen in natural enzymes (which typically have k_{cat}/K_m values in the neighborhood of 10^7 $M^{-1}s^{-1}$ [15]). The upper limit of k_{cat}/K_m imposed by the diffusion-controlled limit is 10^8 to 10^9 $M^{-1}s^{-1}$ [1]. Ideally, we would like to be able to design enzymes with k_{cat}/K_m values close to these, but clearly our current design technique needs to be further optimized to achieve such a goal. These 209 mutants will hopefully not only allow us to improve the activity of HG2/S265T, but also identify weaknesses in both our already moderately successful PHOENIX and MD design protocols. Optimizing parameters so that predictions match what is observed will improve our models' accuracy, allowing us to design better enzymes.

Acknowledgements

Special thanks goes to Christoffer Norn for conducting most of the large-scale purifications and preforming the large-scale kinetic assays on most of the mutants. Special thanks also goes to Alex Nisthal for help using the robot for the high-throughput aspects of this project.

Figures and Tables

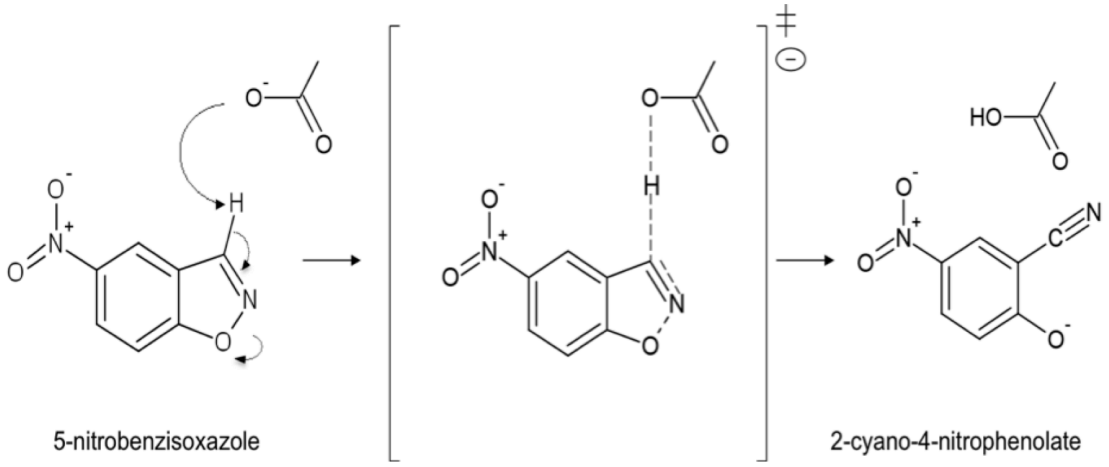


Figure A-1. The Kemp elimination of 5-nitrobenzisoxazole. A base abstracts a proton from the C3 position of 5-nitrobenzisoxazole to yield 2-cyano-4-nitrophenolate.

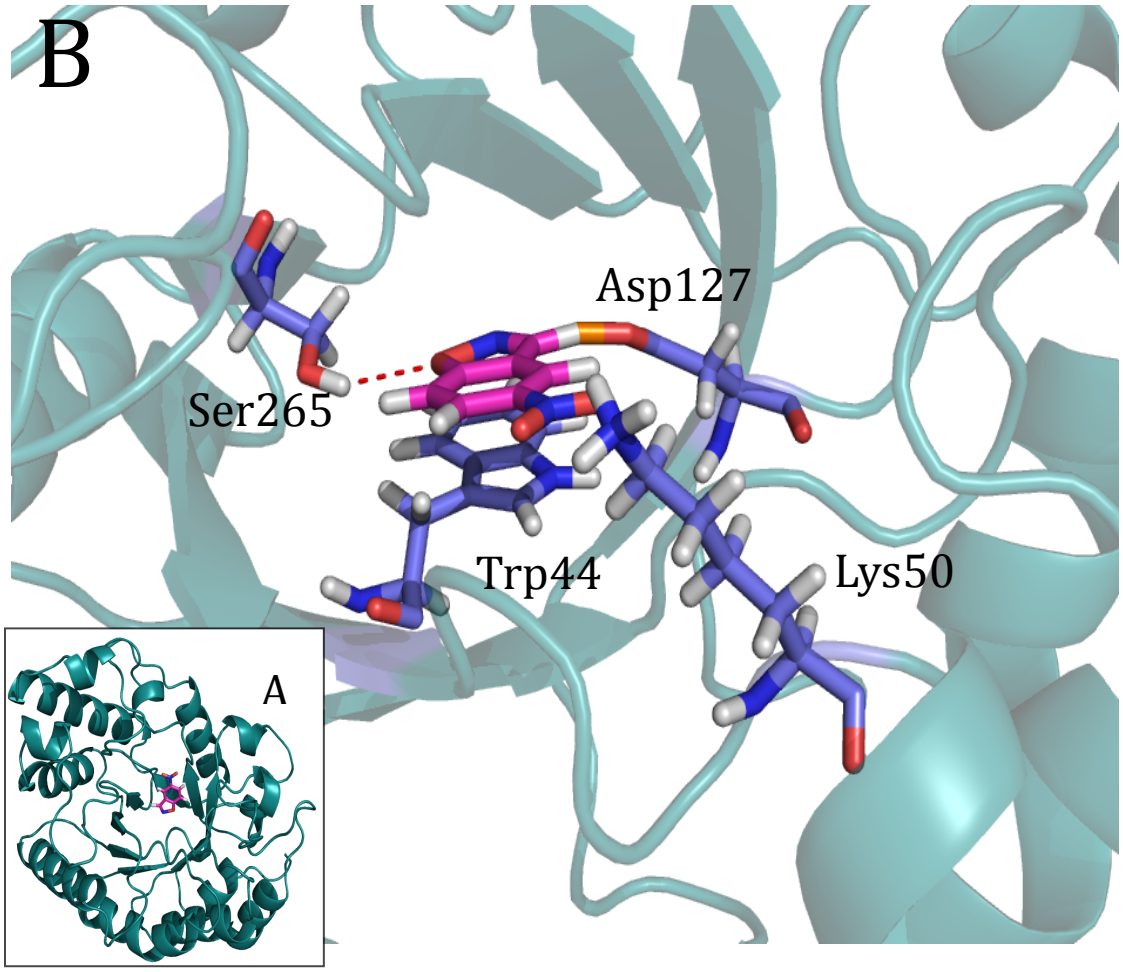


Figure A-2. Design structure of HG2. (A) shows the transition state, NBX (magenta), placed in the design active site of HG-2 within the structure of the scaffold, TAX. (B) shows the active site design of HG-2 with NBX surrounded by Ser265, Trp44, Lys50, and the catalytic residue Asp127.

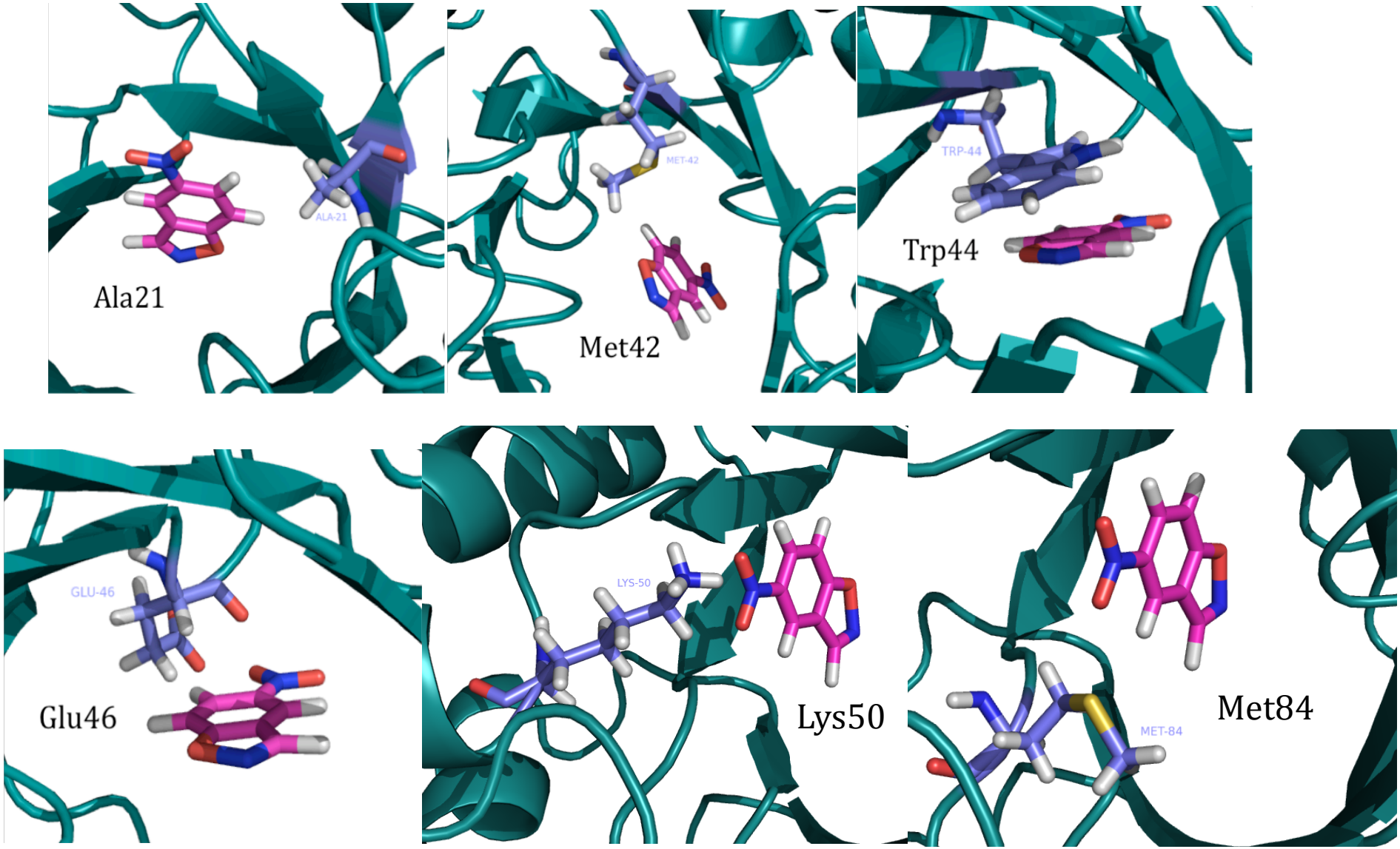


Figure A-3a. HG-2 active site residues chosen for site-saturation mutagenesis

Images are from the design structure of HG-2. Residues for mutagenesis are shown in purple, and the transition state, NBX is shown in magenta. Trp44 π-stacks with NBX and Lys50 interacts with NBX's nitro group.

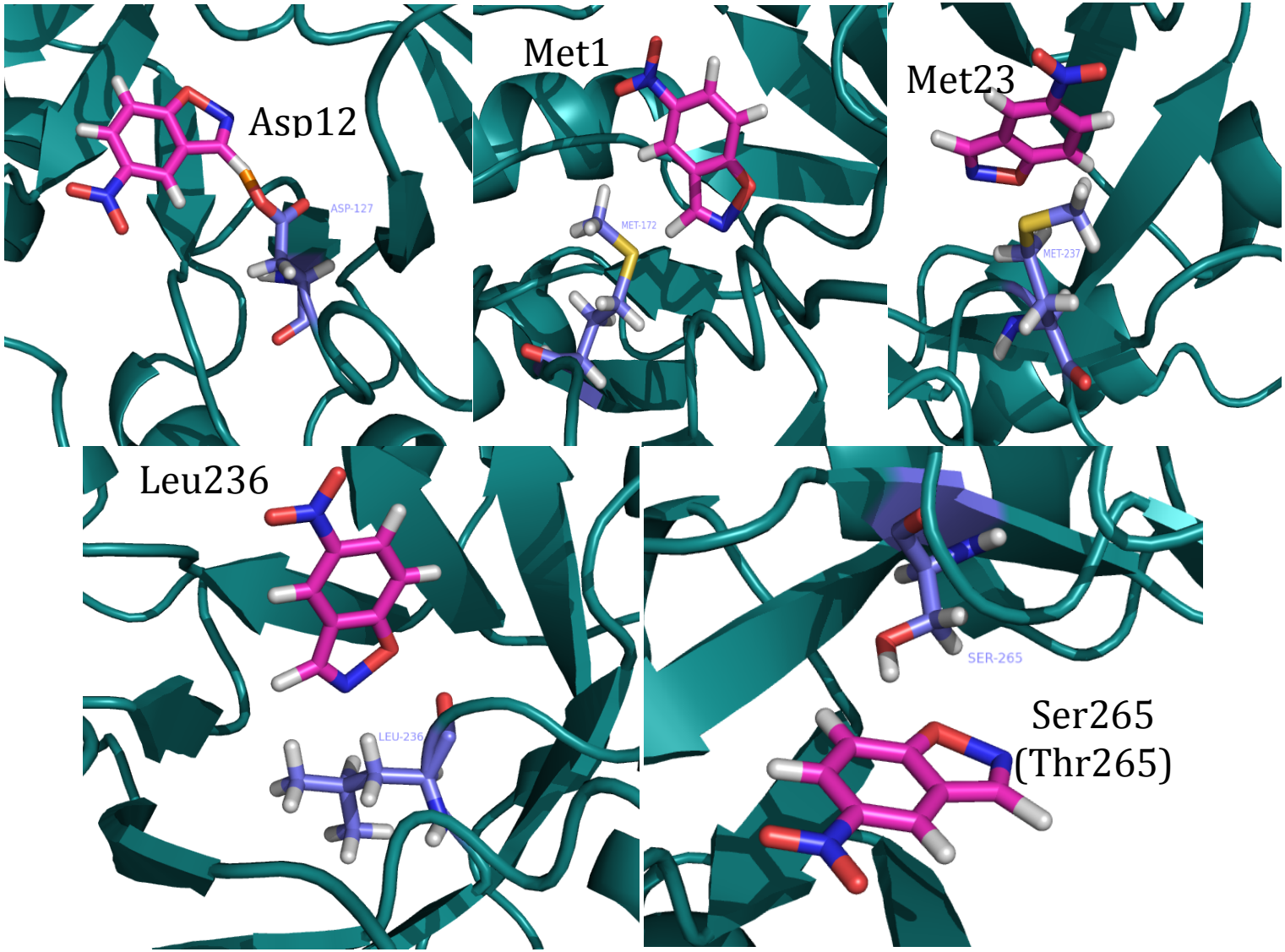


Figure A-3b. HG-2 active site residues chosen for site-saturation mutagenesis, continued

Images are from the design structure of HG-2. Residues for mutagenesis are shown in purple, and the transition state, NBX, is shown in magenta. Asp127 is the catalytic base. Ser265 was meant to hydrogen bond with the isoxazole oxygen (Thr265 in HG-2 S265T).

Table A-1. Summary of site-saturation mutagenesis results

Plate	Total # Rxns	# of Rxns w/ Colonies	# Mutants Achieved After Agencourt Seq.	# Mutants Achieved After 3rd Seq.	% Mutants Achieved After Agencourt Seq.	% Mutants From Colonies	% Mutants Achieved From Total Rxns
1	95	92	67	74	72.8	80.4	77.9
2	76	76	72	74	94.7	97.4	97.4
3	88	60	38	N/A	63.3	63.3	43.2
4	25	23	22	N/A	95.7	95.7	88.0
AVG					81.6	84.2	76.6

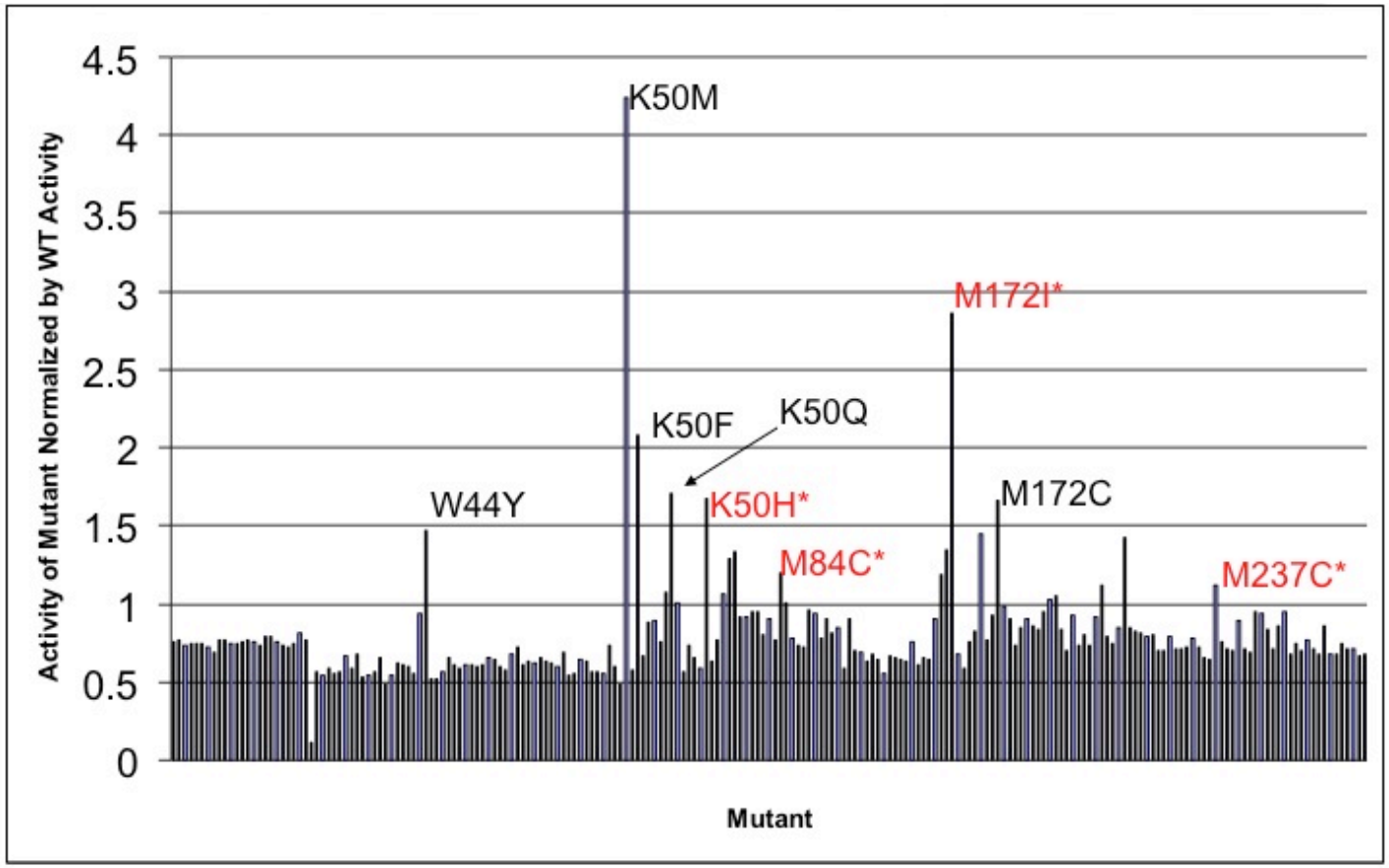


Figure A-4. Activity screen of site saturation mutants for HG2/S265T. Mutations also found by Hilvert lab in red.

Table A-2. Kinetic data for the HG2/S265T mutants found in the screen to be more active than "WT."

Mutant	Normalized Scrren Activity	V_{max} (M/s)	k_{cat} (s^{-1})	K_m (M)	$k$cat/Km ($s^{-1}M^{-1}$)
K50M	4.24	1.79E-05 ± 1.86E-06	3.58E+00 ± 3.71E-01	1.45E-03 ± 2.24E-04	2.47E+03 ± 4.60E+02
M172I	2.86	2.99E-05 ± 6.27E-06	5.99E+00 ± 1.25E+00	2.84E-03 ± 7.58E-04	2.11E+03 ± 7.17E+02
M84I	1.33	1.45E-05 ± 1.67E-06	2.90E+00 ± 3.35E-01	1.58E-03 ± 2.66E-04	1.84E+03 ± 3.75E+02
K50Q	1.70	1.87E-05 ± 2.78E-06	3.74E+00 ± 5.56E-01	2.30E-03 ± 4.55E-04	1.63E+03 ± 4.03E+02
HG2 S265T	1	1.30E-05 ± 2.48E-06	2.59E+00 ± 4.97E-01	2.90E-03 ± 7.06E-04	8.93E+02 ± 2.77E+02
M172C	1.67	1.59E-05 ± 1.04E-05	3.17E+00 ± 2.07E+00	4.40E-03 ± 3.40E-03	7.21E+02 ± 7.30E+02
K50H	1.68	8.76E-06 ± 1.47E-06	1.75E+00 ± 2.95E-01	2.66E-03 ± 5.78E-04	6.59E+02 ± 1.81E+02
M172L	1.35	1.76E-05 ± 5.90E-06	3.52E+00 ± 1.18E+00	6.00E-03 ± 2.29E-03	5.86E+02 ± 2.98E+02
M172V	1.18	4.79E-06 ± 8.63E-07	9.57E-01 ± 1.73E-01	2.36E-03 ± 5.64E-04	4.05E+02 ± 1.21E+02
M172T	1.44	2.82E-06 ± 2.73E-07	5.63E-01 ± 5.47E-02	1.41E-03 ± 2.06E-04	4.00E+02 ± 7.03E+01
M84L	1.29	5.63E-06 ± 2.88E-06	1.13E+00 ± 5.75E-01	3.21E-03 ± 2.04E-03	3.51E+02 ± 2.86E+02
L236M	1.06	3.25E-06 ± 6.20E-07	6.50E-01 ± 1.24E-01	2.49E-03 ± 6.21E-04	2.61E+02 ± 8.21E+01
M237C	1.12	3.67E-06 ± 8.40E-07	7.33E-01 ± 1.68E-01	3.10E-03 ± 8.90E-04	2.37E+02 ± 8.70E+01
M84C	1.19	2.61E-05 ± 4.78E-05	5.23E+00 ± 9.56E+00	2.46E-02 ± 4.66E-02	2.12E+02 ± 5.58E+02
L236D	1.42	4.13E-07 ± 4.17E-08	8.26E-02 ± 8.33E-03	5.78E-04 ± 1.15E-04	1.43E+02 ± 3.20E+01
W44Y	1.47	2.55E-05 ± 5.47E-05	5.10E+00 ± 1.09E+01	3.65E-02 ± 8.03E-02	1.39E+02 ± 4.28E+02
M84V	1.07	9.02E-07 ± 1.20E-07	1.80E-01 ± 2.39E-02	1.55E-03 ± 3.02E-04	1.16E+02 ± 2.73E+01
K50F	2.08	2.42E-06 ± 6.08E-07	4.85E-01 ± 1.22E-01	4.42E-03 ± 1.31E-03	1.10E+02 ± 4.26E+01
L236V	1.03	9.60E-07 ± 2.07E-07	1.92E-01 ± 4.14E-02	3.27E-03 ± 8.76E-04	5.88E+01 ± 2.02E+01
M84P	1.01	1.33E-07 ± 8.11E-09	2.65E-02 ± 1.62E-03	4.73E-04 ± 6.12E-05	5.60E+01 ± 8.02E+00
L236Q	1.12	2.91E-07 ± 7.58E-08	5.82E-02 ± 1.52E-02	2.10E-03 ± 7.43E-04	2.78E+01 ± 1.22E+01
1GOR	N/A	8.90E-08 ± 4.41E-08	1.78E-02 ± 8.81E-03	3.20E-03 ± 1.97E-03	5.57E+00 ± 4.41E+00
D127V	0.82	8.84E-09 ± 1.83E-09	1.77E-03 ± 3.65E-04	3.60E-04 ± 1.72E-04	4.91E+00 ± 2.55E+00

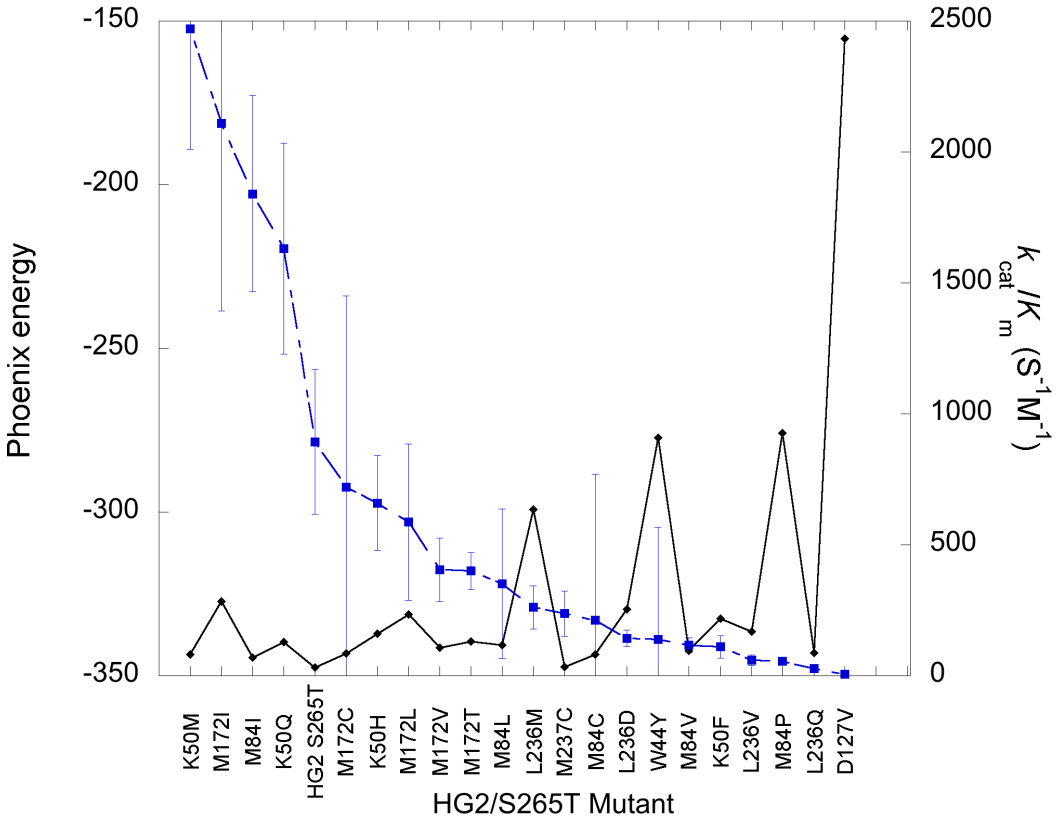


Figure A-5. Calculated PHOENIX energy (black) versus k_{cat}/K_m (blue) of the HG2/S265T active site saturation mutants.

References

1. Nelson, D.L.a.C., M.M., ed. L.P.o. Biochemistry2005, New York: W.H. Freeman and Company.

2. Casey, M.L., et al., *Physical Organic-Chemistry of Benzisoxazoles .1. Mechanism of Base-Catalyzed Decomposition of Benzisoxazoles.* Journal of Organic Chemistry, 1973. **38**(13): p. 2294-2301.

3. Kemp, D.S. and M.L. Casey, *Physical Organic-Chemistry of Benzisoxazoles .2. Linearity of Bronsted Free-Energy Relationship for Base-Catalyzed Decomposition of Benzisoxazoles.* Journal of the American Chemical Society, 1973. **95**(20): p. 6670-6680.

4. Kemp, D.S., D.D. Cox, and K.G. Paul, *Physical Organic-Chemistry of Benzisoxazoles .4. Origins and Catalytic Nature of Solvent Rate Acceleration for Decarboxylation of 3-Carboxybenzisoxazoles.* Journal of the American Chemical Society, 1975. **97**(25): p. 7312-7318.

5. Thorn, S.N., et al., *Large Rate Accelerations in Antibody Catalysis by Strategic Use of Haptenic Charge.* Nature, 1995. **373**(6511): p. 228-230.

6. Debler, E.W., et al., *Structural origins of efficient proton abstraction from carbon by a catalytic antibody.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(14): p. 4984-4989.

7. Privett, H.K., et al., *Iterative approach to computational enzyme design.* Proc Natl Acad Sci U S A, 2012. **109**(10): p. 3790-5.

8. Dahiyat, B.I. and S.L. Mayo, *Protein design automation.* Protein Sci, 1996. **5**(5): p. 895-903.

9. Rothlisberger, D., et al., *Kemp elimination catalysts by computational enzyme design.* Nature, 2008. **453**(7192): p. 190-5.

10. Tseng, W.C., et al., *A novel megaprimed and ligase-free, PCR-based, site-directed mutagenesis method.* Anal Biochem, 2008. **375**(2): p. 376-8.

11. Privett, H.K., *An interative Approach to De Novo Computational Enzyme Design and the Successful Application to the Kemp Elimination*, in *Chemistry*2009, California Insitute of Technology. p. 216.

12. Allen, B.D. and S.L. Mayo, *An efficient algorithm for multistate protein design based on FASTER.* J Comput Chem, 2010. **31**(5): p. 904-16.

13. Blomberg, R., et al., *Precision is essential for efficient catalysis in an evolved Kemp eliminase.* Nature, 2013. **503**(7476): p. 418-21.

14. Havranek, J.J. and P.B. Harbury, *Automated design of specificity in molecular recognition.* Nat Struct Biol, 2003. **10**(1): p. 45-52.

15. Lad, C., N.H. Williams, and R. Wolfenden, *The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 5607-10.

Appendix B

EXPLORING CYANOVIRIN BINDING TO EBOLA GLYCOPROTEIN 1,2 TO FACILITATE FUTURE ANTIVIRIAL DESIGNS

Abstract

Many deadly diseases are caused by enveloped viruses, including AIDS, smallpox, hantavirus disease, and Ebola. New fears of viral breakouts due to bioterrorism or increased international travel have made the development of broad-spectrum antivirals a medical priority. Cynaovirin-N (CVN) is a lectin with potent antiviral activity against many enveloped viruses. CVN binds high-mannose glycosylation sites on viral envelope glycoproteins, blocking their ability to effectively bind to host cell receptors and initiate membrane fusion, thus preventing infection. Ebola glycoprotein 1,2 (GP1,2) is the viral envelope protein on Ebola virus (a biohazard level 4 pathogen) that initiates membrane fusion. The mucin-like domain of GP1,2 is the primary viral determinant of Ebola pathogenicity. CVN exhibits antiviral activity against Ebola, and it has been shown that such activity most likely originates from CVN's ability to bind GP1,2. Our goal is to design CVN variants with improved antiviral activity that could potentially be used as broad-spectrum antiviral therapeutics. Toward this goal, we plan to develop *in vitro* CVN binding assays to a wide range of viral envelope glycoproteins, thereby providing a safe way to estimate the antiviral activities of the variants. Thus far, we have expressed GP1,2 with the mucin-like domain in insect cells. Future work will include the development and optimization of other GP1,2-CVN assays as well as assays involving other viral envelope glycoproteins.

Introduction

Enveloped Viruses

A virus is an infectious agent and intracellular parasite, and often causes disease in its animal host. Viruses are unable to reproduce themselves, and therefore must infect a host cell and commandeer its molecular machinery in order to replicate. An infectious virus particle, or virion, consists of nucleic acids making up the viral genome, a surrounding protein capsid, and in some cases, an outer phospholipid bilayer envelope. Viruses with these outer membranes are known as enveloped viruses. During replication of an enveloped virus, the capsid assembles around the viral genome and buds from the host cell. During budding, part of the host cell membrane coats the protein, becoming the viral envelope. Within the envelope are heavily glycosylated virus-encoded proteins, which bind to host cell membrane receptors, facilitating the membrane fusion event that occurs when the newly-budded virus infects another host cell. These glycoproteins are known as fusion proteins, and their heavy glycosylation acts to cloak them from normal immune system responses [1].

Ebola Virus

Enveloped viruses cause a myriad of human diseases, such as AIDS (acquired immunodeficiency syndrome), influenza, small pox, herpes simplex, SARS (severe acute respiratory syndrome), and Ebola [2, 3]. Ebola virus (EBOV) is a member of the *Filovirdae* family and causes hemorrhagic fever with mortality rates of 50-90% [2]. Five species of the virus exist: Bundibugyo, Sudan, Côte d'Ivoire, Reston, and Zaire. Ebola Zaire is the most lethal of these, with mortality rates of 90% [4]. Patients infected with Ebola exhibit headache, nausea, malaise, high fevers, aberrant coagulation and vascular permeability causing bleeding, bruising and rashes [5]. Death results from multiple organ failure within 6-9 days after the first appearance of symptoms [6].

Ebola patients mainly receive palliative care [5]. Several preliminary vaccine trials in primates have been successful [7-10], but no approved vaccines, post-exposure treatments or cures for humans exist [5]. Ebola outbreaks have increased more than four-fold since 1994, and Ebola has been identified as a virus that may be used as a biological weapon [11]. Therefore, the development of vaccines and therapies for Ebola has become a high priority.

Cyanovirin-N

Cyanovirin-N (CVN) is an 11 kDa lectin exhibiting potent antiviral activity against many enveloped viruses, and is a promising candidate for broad-spectrum antiviral therapy [12, 13]. CVN was isolated from the cyanobacterium *Nostoc ellipsosporum* in a screen conducted to find compounds with anti-human immunodeficiency virus (HIV) activity [13]. In addition to being a potent antiviral against HIV [13-16], CVN has been shown to have activity against influenza [17, 18], hepatitis C [19], herpesvirus 6, measles [20], and Ebola [12, 21]. In each of these cases, CVN binds high-mannose glycosylation sites on the viral envelope glycoproteins, blocking their ability to effectively bind to host cell receptors and initiate membrane fusion, thus preventing infection.

CVN has two binding sites, a "high affinity" and "low affinity" site (Figure B-1) [22], and specifically binds α(1-2) linked oligomannose moieties within Man-8 or Man-9 glycosylation sites [22-25]. Such high mannose glycosylation is often found on viral envelope proteins, but is uncommon in mammalian oligosaccharides. This may allow CVN to specifically target viruses and infected host cells rather than healthy host tissue, an important characteristic for a potential therapeutic. CVN is small and stable [13, 26-28]. In non-human primate studies, CVN has already been shown to be an effective prophylactic against rectal and vaginal SIV/HIV-1 transmission, exhibiting only a low level of toxicity [14, 15]. Treatment with CVN can also cause HIV and influenza to evolve to eliminate N-linked glycosylation sites on their envelope glycoproteins [16, 17, 29, 30]. However, loss of glycosylation exposes the protein to the adaptive immune system, and has been shown to make such viruses more susceptible to antibody neutralization [31]. Such characteristics make CVN a potentially very valuable drug.

Ebola Glycoprotein 1,2

The Ebola virus genome contains 7 genes from which 8 proteins are synthesized. The *GP* gene directs the synthesis of two proteins via a transcriptional editing event: sGP and GP1,2. sGP, a 364-residue secreted glycoprotein whose function remains unclear, is the primary gene product. The synthesis of the secondary protein, GP1,2, occurs upon addition of a non-templated adenosine within a length of 7 adenosines in the coding region [32, 33]. GP1,2 is a 676-residue transmembrane-linked glycoprotein that is post-translationally cleaved by furin into the subunits

GP1 and GP2. These subunits remain covalently linked via a single disulphide-bond [34]. GP1 serves as the viral envelope protein that causes viral attachment to host cells, whereas GP2 mediates virus-host membrane fusion [35-37]. GP1,2 assumes a trimeric complex on the envelope (Figure B-2).

In addition to its role in membrane fusion, GP1,2 synthesis has also been shown to be the main determinant of Ebola pathogenicity [38-42]. Specifically, this cytotoxicity was mapped to the heavily glycosylated mucin-like domain of GP1,2. Synthesis of Zaire GP1,2 in human embryonic kidney 293 cells caused cell rounding and detachment within 24 hours after transfection, and eventually resulted in cell death, whereas the expression of sGP and GP1,2(Δmuc), a GP1,2 mutant with a deletion of the mucin-like domain, caused no such effects. Addition of cycloheximide (a protein synthesis inhibitor) to cells expressing GP1,2 blocked the cytotoxicity. Expression of Zaire GP1,2 in human saphenous vein explants resulted in the erosion of the endothelial cell layer and exposure of the underlying basement membrane. These explant experiments were repeated in human and cynomolgus monkey vessels using the GP1,2 from Ebola Reston, which is highly lethal in many primates, but not humans. While the monkey vessels showed similar damage to those human vessels expressing Zaire GP1,2, the human vessels expressing Reston GP1,2 remained unaffected [38].

For these reasons, GP1,2 is the main target for vaccine and antiviral therapy design. Moreover, it has been shown that CVN binds to GP1,2, and this binding is the probable mechanism providing CVN with its anti-Ebola activity [21].

A goal of our lab is to engineer a CVN variant with increased antiviral activity that could be used as a broad-spectrum antiviral therapeutic or prophylactic. Several variants with improved antiviral activity have been constructed, and work on a “lectibody” (a CVN-Fc variant designed to induce an immune response) is underway [43]. So far, testing of these CVN variants has been limited to HIV and influenza systems. We would like to expand to other viral models. However, due to their level 4 biohazard rating, it is impossible for us to work with small pox virus and Ebola virus in our facilities. In collaboration with Irene Maier, a postdoc in our lab, I am instead developing *in vitro* CVN binding assays with Ebola GP1,2, which will provide a safe way to estimate the antiviral activity of our designed variants. Data from these assays will then be used to improve upon existing designs to create a therapeutic CVN variant with superior avidity that retains its broad-spectrum capabilities.

Methods

Construct Design

The sequence for Ebola Zaire GP1,2 was obtained from the NCBI (National Center for Biotechnology) database. Recombinant Ebola Zaire GP1,2 gene was designed to include residues 33 through 632, excluding residues 1-32 (the signal peptide) and 633-676 (the membrane-proximal external region, the transmembrane region and the cytoplasmic tail). An N-terminal His-6 tag, TEV cleavage site and GGSGG-linker were added to the N-terminus of the protein (Figure B-3). A single adenosine was added to the 7-long adenosine span in the gene sequence to ensure the synthesis of GP1,2 over sGP.

Gene Assembly and Cloning

40-nucleotide long overlapping oligonucleotides were used to construct the GP1,2 gene. We employed the recursive PCR (polymerase chain reaction) method as described in Stemmer *et al.* [44], except the assembly reaction was divided into three stages: (1) an initial stage in which the gene was divided into quarters and each quarter was assembled individually, (2) a repeat assembly on each quarter in which the reaction was spiked with dNTPs and polymerase, and (3) a final stage in which the 4 previously assembled quarters were combined and assembled into the full-length gene (Figure B-4). The assembled gene was subsequently cloned into baculovirus transfer vector pAcGP67A (BD Biosciences).

Protein Expression and Purification

All recombinant baculovirus generation, protein expression, and protein purification was performed at the Protein Expression Center (Caltech). Recombinant baculoviruses were generated using purified pAcGP67A with the Ebola Zaire GP1,2 gene insert in *Spodoptera frugiperda* (Sf9) insect cells. Protein was expressed in 1 L Sf9 cultures infected with recombinant baculovirus. Protein was purified via Ni-NTA affinity chromatography and was stored in 10 mM sodium phosphate, 140 mM sodium chloride at pH 7.4.

Verification of Cleavage Event to Form GP1 and GP2

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and liquid chromatography-mass spectrometry (LC-MS) were both run on reduced and unreduced samples of purified GP1,2 to determine if the cleavage event that separates the protein into GP1 and GP2 occurred. LC-MS was conducted at the Protein/Peptide MicroAnalytical Laboratory (Caltech). An in-gel trypsin digestion with reduction and alkylation of Coomassie stained bands followed by analysis by mass spectrometry was also conducted on purified GP1,2 at the Proteome Exploration Laboratory (Caltech).

Circular Dichroism (CD)

CD data were obtained using an Aviv DS spectropolarimeter and the temperature was controlled with a thermoelectric unit. Samples contained 3.7 μM protein in buffer containing 2.5 mM sodium phosphate and 35 mM sodium chloride at pH 7.4. Wavelength scans were performed in triplicate in 1 nm steps from 250 nm to 200 nm with an averaging time of 3 s.

Results and Discussion

Gene Assembly and Expression

Gene assembly yielded an expected PCR product of 1.8 kilobases. The method of dividing the gene assembly oligonucleotides into sets of quarters and running an initial assembly reaction on each yielded much better results than performing an equivalent number of assembly reactions while including all oligonucleotides (Figure B-5). These results demonstrated that smaller gene pieces assembled better, whereas running more assembly reactions helped minimally.

To avoid deletion of the mucin domain, which is necessary for expression in human cells, yet retain glycosylation, an insect cell line was chosen for Ebola GP1,2 expression as previously described [45, 46]. 1 L cultures yielded about 8 mg of soluble protein (Figure B-6).

Verification of Cleavage Into GP1 and GP2

Because GP1 and GP2 are linked only by a single disulfide bond, SDS PAGE with reducing conditions should yield two separate bands. The molecular weight of unglycosylated GP1,2 is expected to be 67.737 kilodaltons (kDa), with GP1 at 53.075 kDa., and GP2 at 14.662 kDa. Reduced GP1,2 SDS PAGE showed a shift of on the order of 12 kDa down from its corresponding position on the gel in its unreduced form (Figure B-7). In addition, bands appear around 21 kDa when GP1,2 is run under reduced conditions.

LC-MS was also done on reduced and unreduced samples of GP1,2. Under unreduced conditions, two species were observed: one at 70.834 kDa and one at 71.550 kDa (Figures B-8 and B-9). The molecular weight of mannose is 180.156 Daltons (Da). The difference in weight between the two species is 716 Da (equivalent to ~ 4 mannoses). These two species may represent two forms of GP1,2 with slightly different glycosylation. The reduced sample yielded three species: one with a small peak and a molecular weight of 70.846 kDa, a species at 14.663 kDa, and a third at 24.264 kDa (Figures B-10, B-11, and B-12).

The 14.663 kDa peptide corresponds to GP2, which has a predicted molecular weight of 14.662 kDa. The peptide with a weight of 70.846 kDa probably corresponds to the full-length GP1,2, as observed under unreduced conditions. The second species of full-length GP1,2 (71.550 kDa) does

not appear under reduced conditions. The third species at 24.264 kDa in reduced conditions appears to be a fragment of GP1. Perhaps full-length GP1 has trouble flying during mass spectrometry and therefore remains undetected.

To further investigate this, a trypsin digestion followed by mass spectrometry was also conducted, but unfortunately the protein prep used for this experiment was unusually unstable and had degraded, yielding a reduced gel with many bands, making the results inconclusive. However, several bands were still analyzed, and large coverage of the peptide sequence was found, leading one to believe that the entire GP1,2 was being expressed.

CD Analysis of GP1,2

CD analysis of GP1,2 (Figure B-13) showed that the protein exhibits both α- and β-secondary structural elements, which is in agreement with the published crystal structure (Figure B-14). Proteins with both α- and β-secondary structure often show two negative CD bands at 222 and 208-210 nm and one positive band near 190-195 nm similar to those of all-α proteins. Sometimes a single, broad minimum may appear between 210 and 220 nm due to overlapping of various α-helix and β-sheet bands [47]. GP1,2 exhibited a negative band at about 207 nm with a broad shoulder at ~220 nm.

Future Work

Structure of GP1,2

To further support that GP1,2 is correctly folded, CD will be used to monitor a thermal denaturation of the protein. If the protein is folded, it should be possible to determine the melting temperature of the protein by observing the loss of current secondary structure.

Future work will also include attempting to crystallize our Ebola GP1,2 construct. The current crystal structure of GP1,2 was obtained from a construct without the mucin-like domain [11]. Solving a crystal structure of our recombinant GP1,2 would not only definitively verify whether or not the protein was folded correctly, it would also provide the first atomic resolution structure of the toxic mucin-like domain.

Repeating the trypsin digestion and mass spectrometry on a new batch of GP1,2 will also be done to identify the bands obtained when running SDS-PAGE on reduced and unreduced samples of GP1,2. This experiment will confirm whether or not the protein has been processed correctly. If it has not, the enzyme furin (New England Biolabs) will be used to cleave GP1,2 after expression and purification.

Generation of an Unglycosylated GP1,2

To crystallize our GP1,2 protein and to generate a negative control for our CVN binding studies, we will attempt to fully deglycosylate GP1,2. The current crystal structure of GP1,2 was obtained after deglycosylation of GP1,2 by the enzyme PNGase F [11]. PNGase F removes N-linked glycosylation from proteins and is often used for this purpose on proteins intended for crystallization screens to help increase sample homogeneity and promote crystal contacts. However, the serine- and threonine-rich mucin-like domain contains many predicted regions of O-linked glycosylation. Indeed, incubating our GP1,2 with PNGase F results in a smear on an SDS-PAGE gel. There is no enzyme comparable to PNGase F for the removal of O-linked glycosylation, so another strategy will have to be used. Unfortunately, expression of GP1,2 in *E. coli* to avoid glycosylation has previously been shown to be impossible [42]. One option is to construct another mutant of GP1,2 with the mucin-like domain deleted. This construct would only be useful as a negative control for binding studies. To get a crystal structure of the mucin-like

domain, one possible option is to mutate out all the predicted sites of O-linked glycosylation. Unfortunately, the mucin-like domain, which spans 151 residues, includes 21 serines and 27 threonines, so such a tactic might not be feasible.

Optimizing Storage Buffer Conditions

The batches of GP1,2 from our expression system have proven somewhat unstable. We store the protein at 4°C in 10 mM sodium phosphate and 140 mM sodium chloride at pH 7.4. Two expressions have been done; the first broke down after approximately two months under these conditions, and the second degraded within three weeks. Flash freezing the protein stops degradation, but upon thawing the protein breaks down within a few days. To avoid this, storage buffer conditions need to be optimized.

TEV Cleavage and ELISA

We plan to investigate GP1,2 binding to CVN using enzyme-linked immunosorbent assays (ELISA). In these experiments, GP1,2's His-tag will be cleaved off, and the protein will be immobilized. CVN will then be allowed to bind to GP1,2. A primary antibody will then bind the His-tag of CVN. The secondary antibody, which will be conjugated to horseradish peroxidase, will then bind the primary antibody, providing a mechanism for readout. Such an assay will further characterize the binding of CVN and subsequent designed variants to GP1,2. These data will then be used to improve the design of CVN antivirals.

New Viral Proteins and CVN Engineering

The ultimate goal of this project is to engineer a CVN variant that can be used as a broad spectrum antiviral. These BIAcore and ELISA assays developed using GP1,2 will be used to characterize the binding ability of CVN in conjunction with binding and neutralization studies using HIV and influenza glycoproteins. In addition to these glycoproteins, we wish to expand our binding assays to glycoproteins from other viruses that could become a threat, such as smallpox and hantavirus.

Having a wide range of assays with which to test our CVN variants will expand our ability to engineer as universal an antiviral as possible.

Acknoledgements

Thanks goes to Inderjit Nangiana of the Caltech Protein Expression Center for expression and purification of Ebola glycoprotein 1,2. Also, I would like to thank Dr. Robert Graham of the Caltech Proteome Exploration Laboratory for trypsin digests and mass spectrometry experiments, and the CaltechProtein/Peptide MicroAnalytical Laboratory for performing mass spectrometry experiments.

Figures

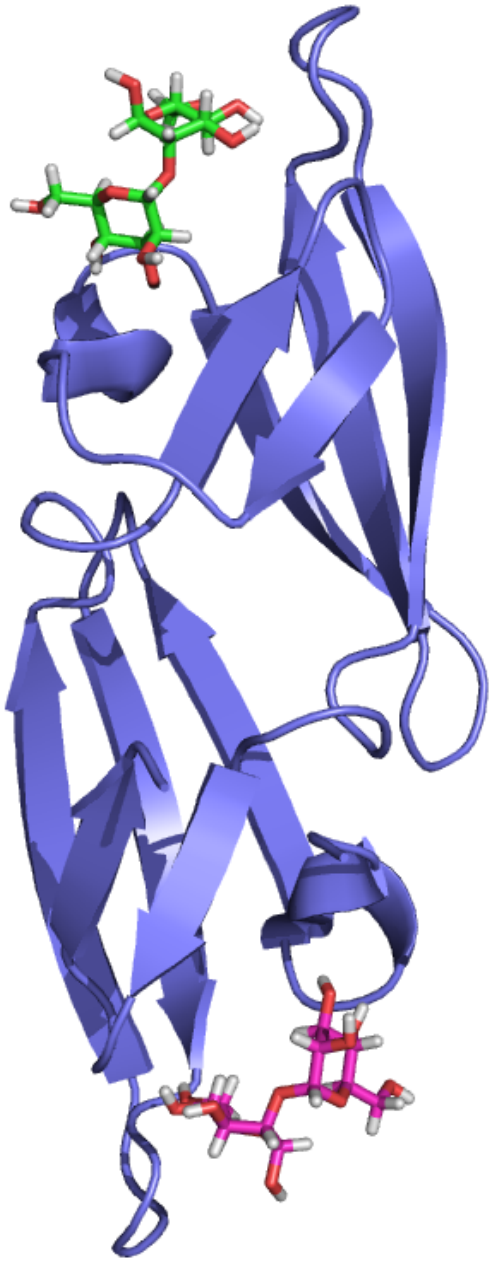


Figure B-1. WT CVN monomer bound to mannose moieties [48]. Carbohydrates in green are bound to the low affinity binding site, and those in magenta are bound to the high affinity site. CVN monomer is shown in purple.

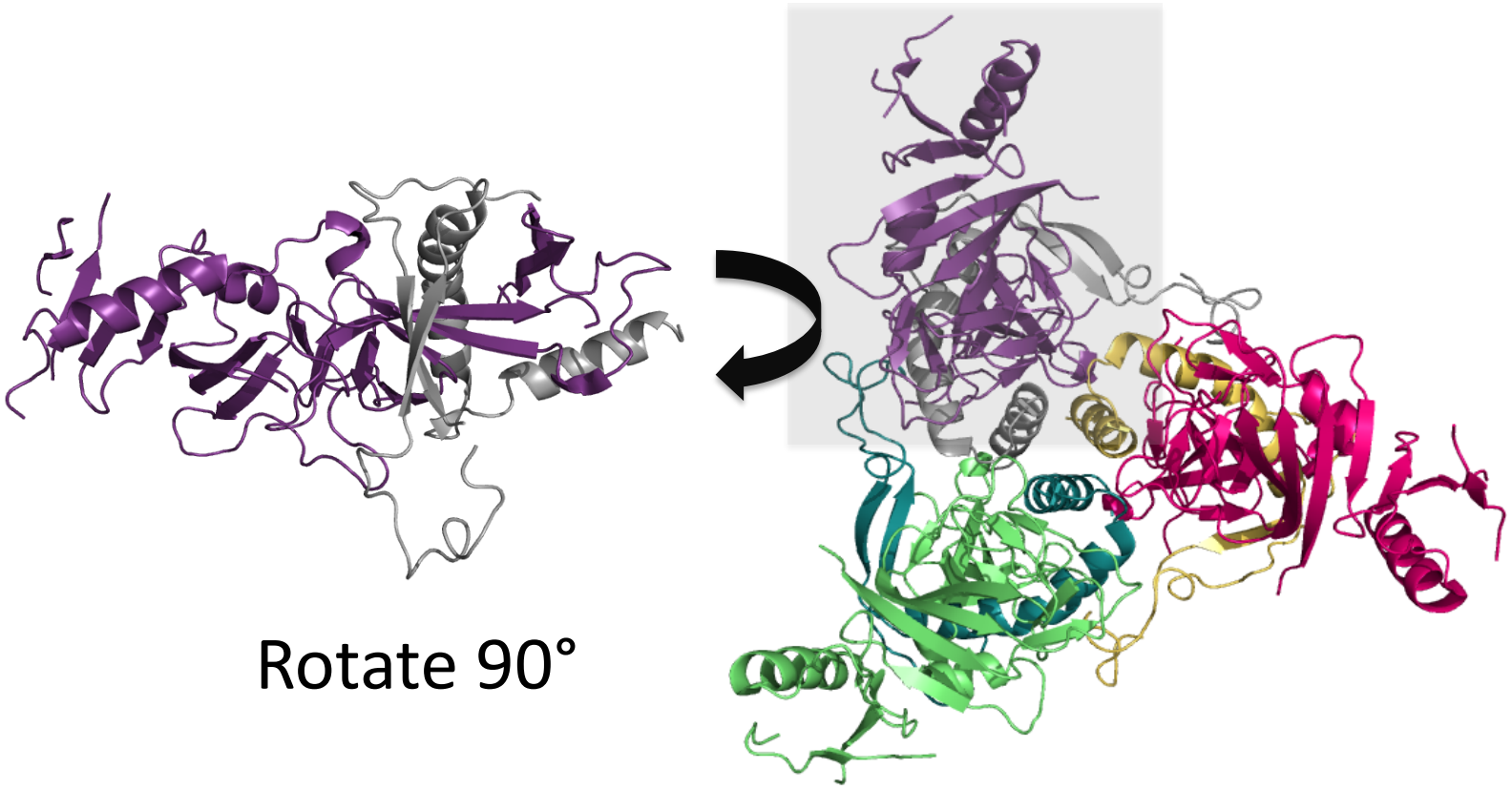


Figure B-2. Structure of Ebola glycoprotein GP1,2 [11]. (Left) Ebola GP1,2 consists of two peptides, GP1 and GP2, joined together by a single disulphide bond. GP1 is shown in purple and GP2 in silver. (Right) GP1,2 exists as a trimer on the surface of the viral envelope.

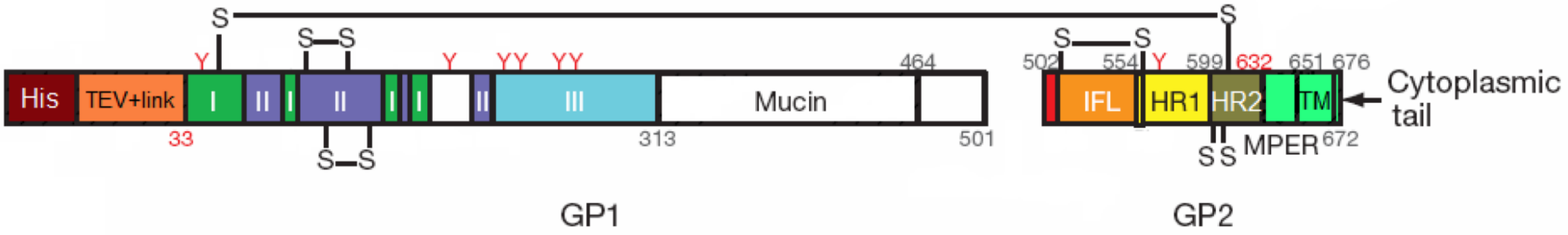


Figure B-3. Gene construct and domains of Ebola Zaire GP1,2. From left to right: dark red His, N-terminal His-tag; peach TEV+link, TEV cleavage site and GGSGG-linker; green I, GP1 base; purple II, GP1 head; cyan III, GP1 glycan cap; white Mucin, mucin-like domain; orange IFL, internal fusion loop; yellow HR1, heptad repeat 1; gold HR2, heptad repeat 2; light green MPER, membrane-proximal external region; light green TM, transmembrane domain. All domain shown in light green (residues 633-676) were deleted from the gene. The red Y's indicate N-linked glycosylation sites (Adapted from Figure 1 of Lee et al. [11]).

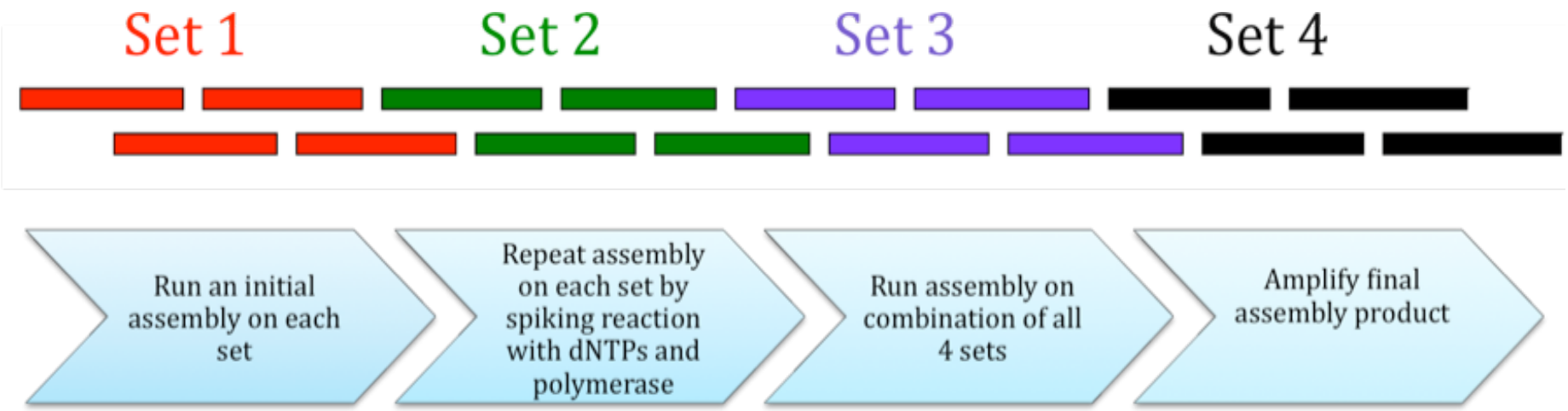


Figure B-4. Gene assembly protocol scheme. Oligonucleotides are represented by rectangular boxes and are divided into four quarters or sets by color. An initial assembly reaction was run on each set individually (round 1) and repeated (round 2). The individual assemblies were then combined for a final assembly reaction (round 3), and then the final, full-length assembly product was amplified.

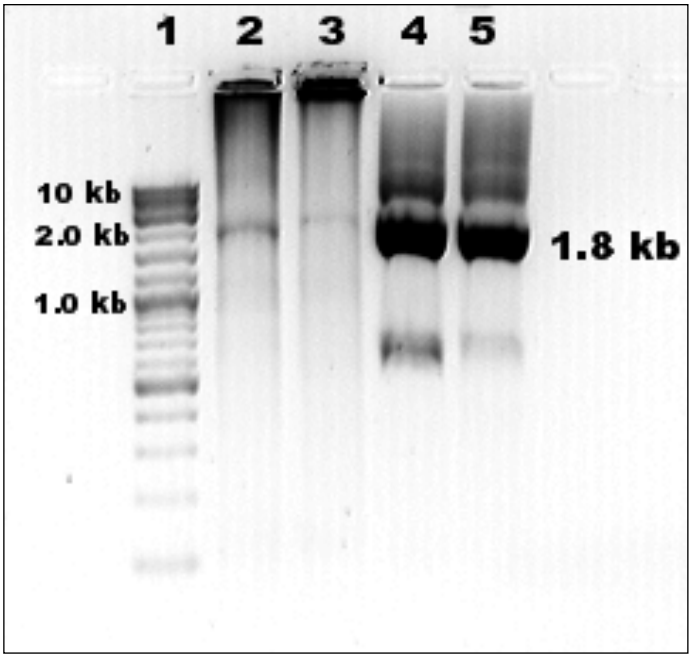


Figure B-5. Gene assembly of GP1,2. Dividing the assembly reaction into 4 subsets to run 2 initial assembly reactions, and then mixing these assembled quarters together for a final assembly yielded much better results than running all assemblies with all oligonucleotides present. Agarose gel lanes contain as follows: lane 1: NEB 2-log ladder; lanes 2 and 3: amplification of PCR product where all sets (1-4) were together for all 3 rounds of gene assembly (initial assembly, repeated assembly, and final assembly); lanes 4 and 5: amplification of PCR products where sets 1-4 were separate for the first 2 rounds of assembly, and together on the final round.

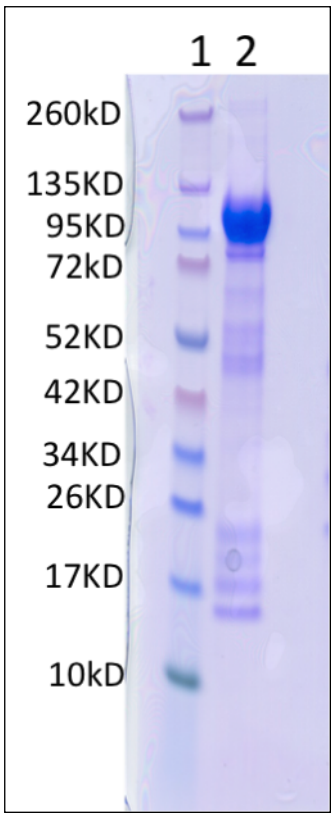
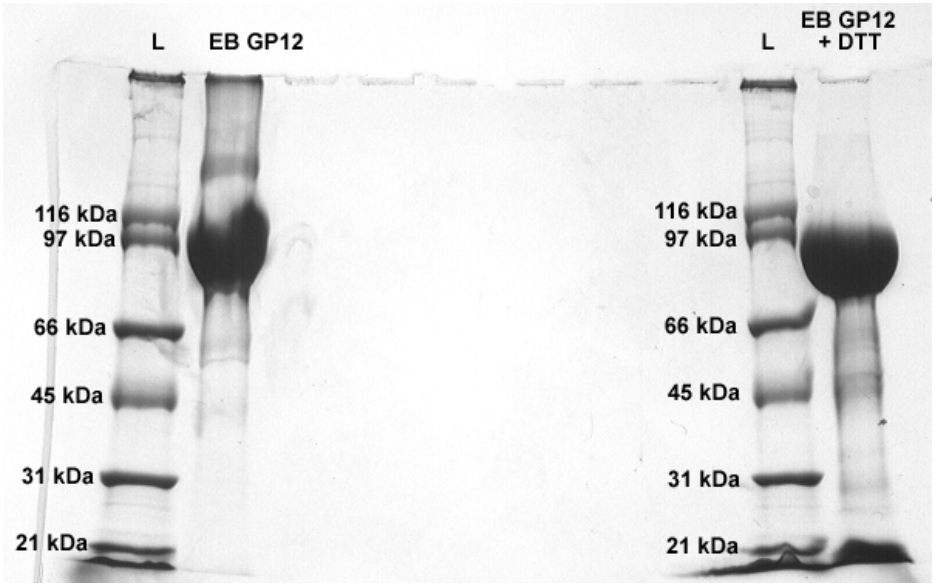


Figure B-6. GP1,2 purified from *Sf9* cells. Reduced SDS page gel run by Inderjit Nangiana at the Protein Expression Center shows purified GP1,2 running at approximately 95 kilodaltons in lane 2.

A



B

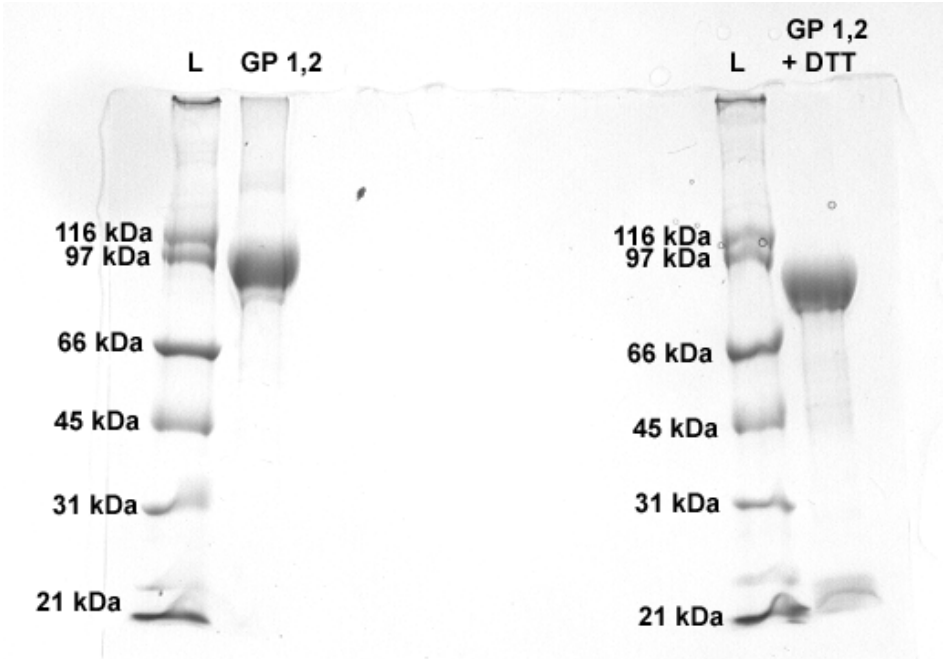


Figure B-7. SDS-PAGE gels of Ebola Zaire GP1,2 under reduced and unreduced conditions. Ebola GP1,2 shifts down the gel about 10 kDa as determined by graphing the logs of the distances traveled of each band versus its molecular weight. Bands appear around 21 kDa under reduced conditions. (A) was obtained running 50 μg of GP1,2, whereas in (B), 20 μg was run.

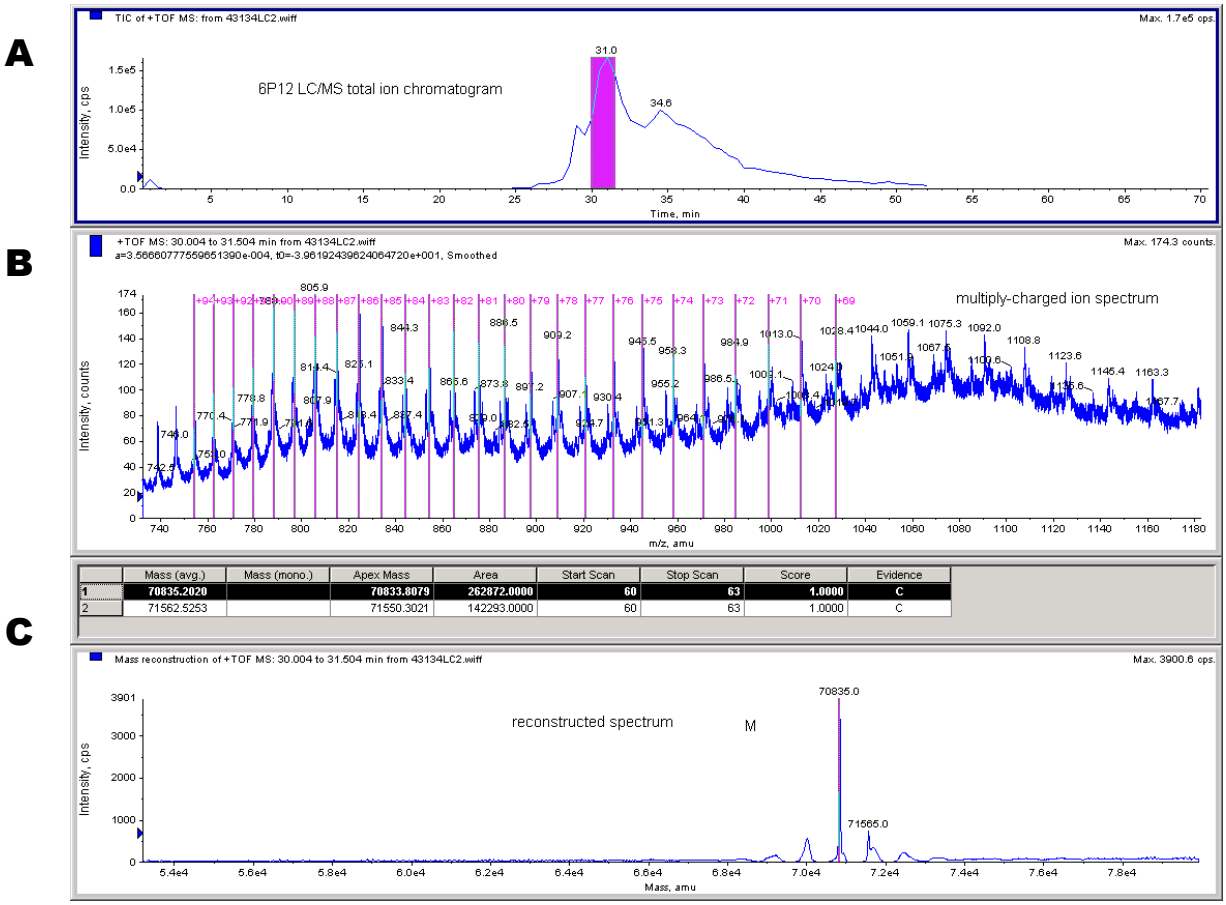


Figure B-8. Liquid chromatography and time of flight mass spectrometry (TOF-MS) of unreduced GP1,2, species 1. (A) The elution profile of a LC of the unreduced sample of GP1,2. (B) The multiply-charged ion spectrum of the sample peak highlighted in A analyzed with TOF-MS. (C) The multiply-charged ion peaks in B were reconstructed to yield a single peak with mass value of 70.833 kDa.

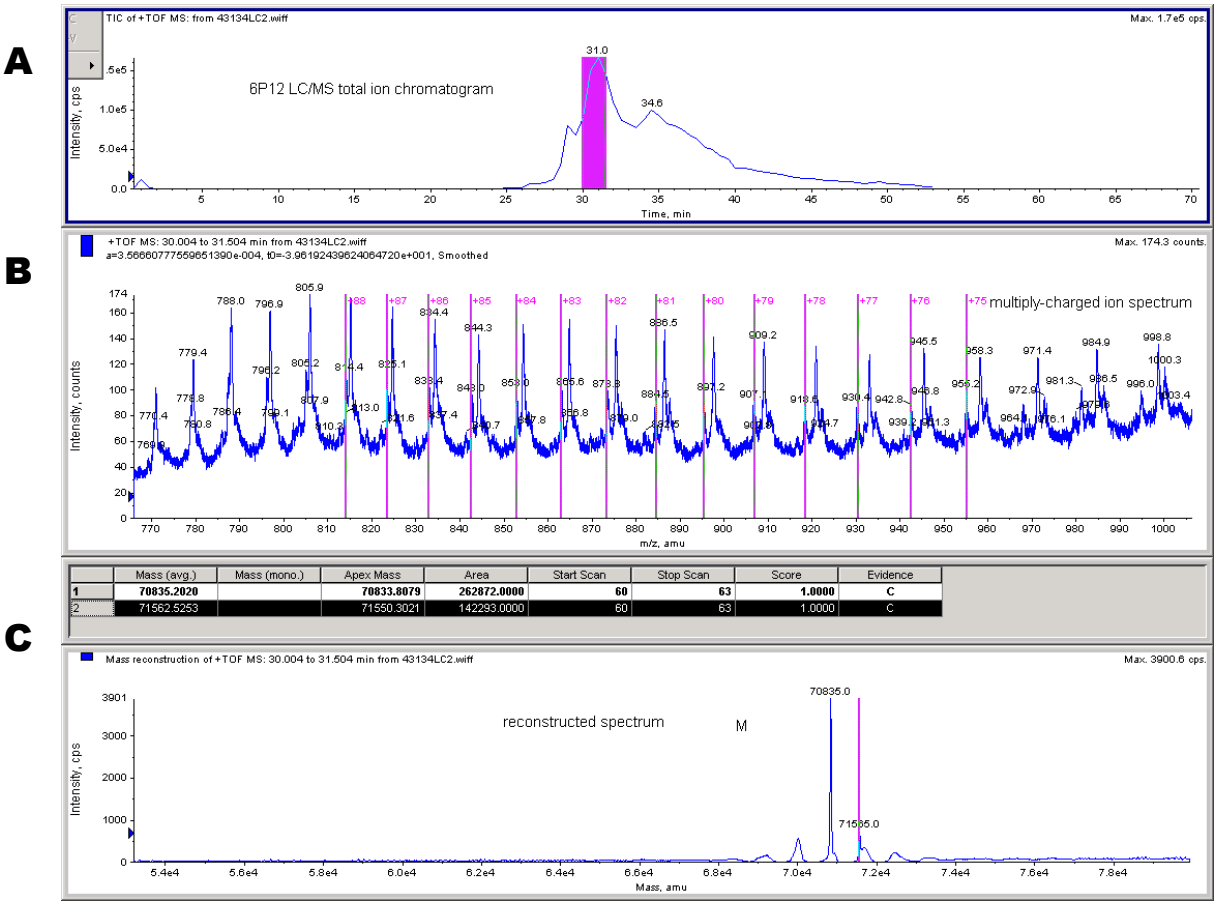


Figure B-9. Liquid chromatography and time of flight mass spectrometry of unreduced GP1,2, species 2. (A) The elution profile of a LC of the unreduced sample of GP1,2. (B) The multiply-charged ion spectrum of the sample peak highlighted in A analyzed with TOF-MS. (C) The multiply-charged ion peaks in B were reconstructed to yield a single peak with mass value of 71.550 kDa.

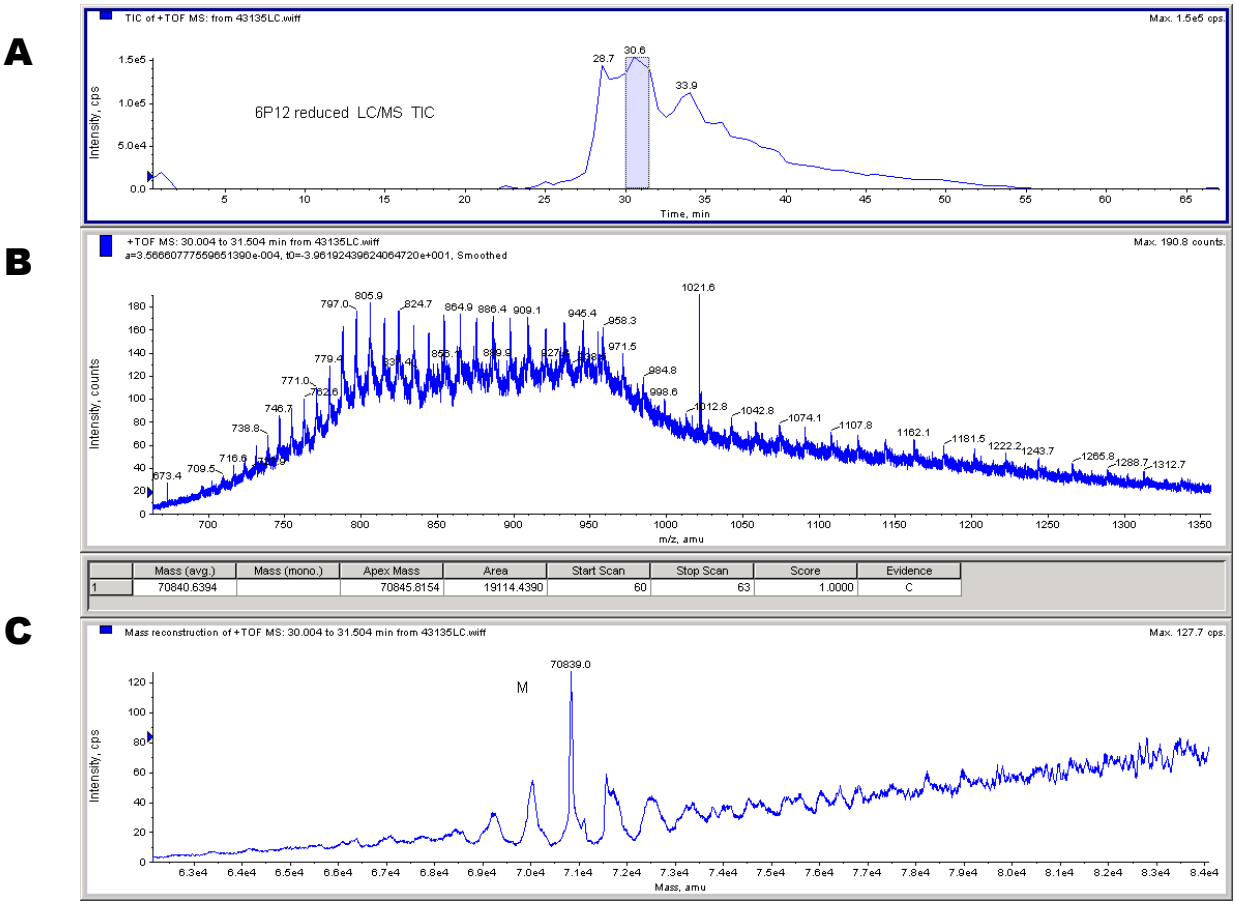


Figure B-10. Liquid chromatography and time of flight mass spectrometry of reduced GP1,2, species 1. (A) The elution profile of a LC of the reduced sample of GP1,2. (B) The multiply-charged ion spectrum of the sample peak highlighted in A analyzed with TOF-MS. (C) The multiply-charged ion peaks in B were reconstructed to yield a single peak with mass value of 70.846 kDa.

A

B

C

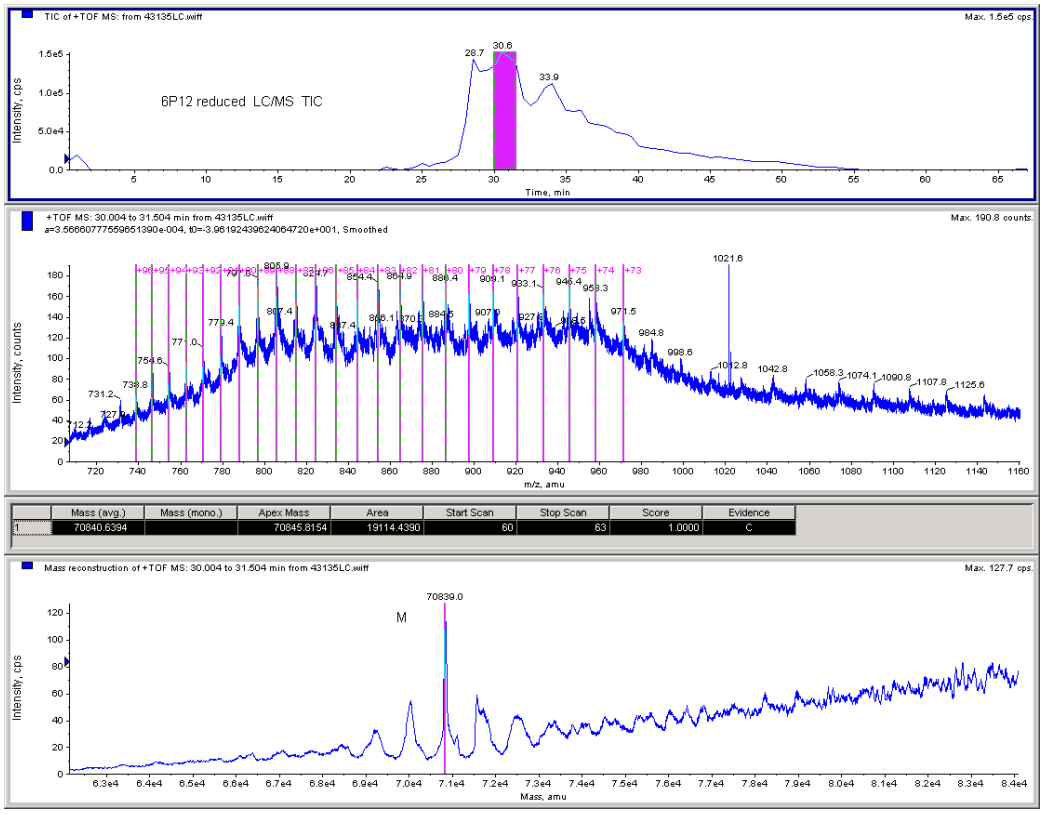


Figure B-11. Liquid chromatography and time of flight mass spectrometry of reduced GP1,2, species 2. (A) The elution profile of a LC of the reduced sample of GP1,2. (B) The multiply-charged ion spectrum of the sample peak highlighted in A analyzed with TOF-MS. (C) The multiply-charged ion peaks in B were reconstructed to yield a single peak with mass value of 14.663 kDa.

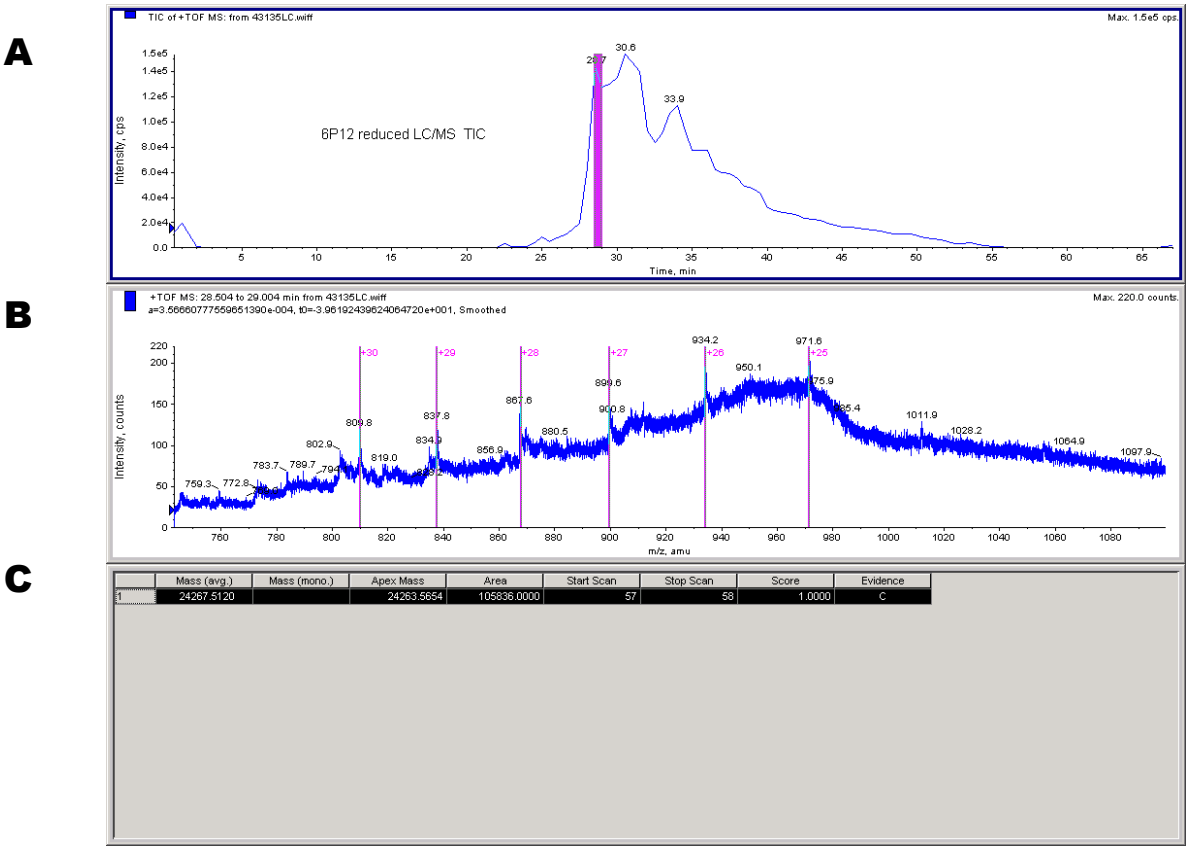


Figure B-12. Liquid chromatography and time of flight mass spectrometry of reduced GP1,2, species 3. (A) The elution profile of a LC of the reduced sample of GP1,2. (B) The multiply-charged ion spectrum of the sample peak highlighted in A analyzed with TOF-MS. (C) The mass of this peak was determined to be 24.264 kDa.

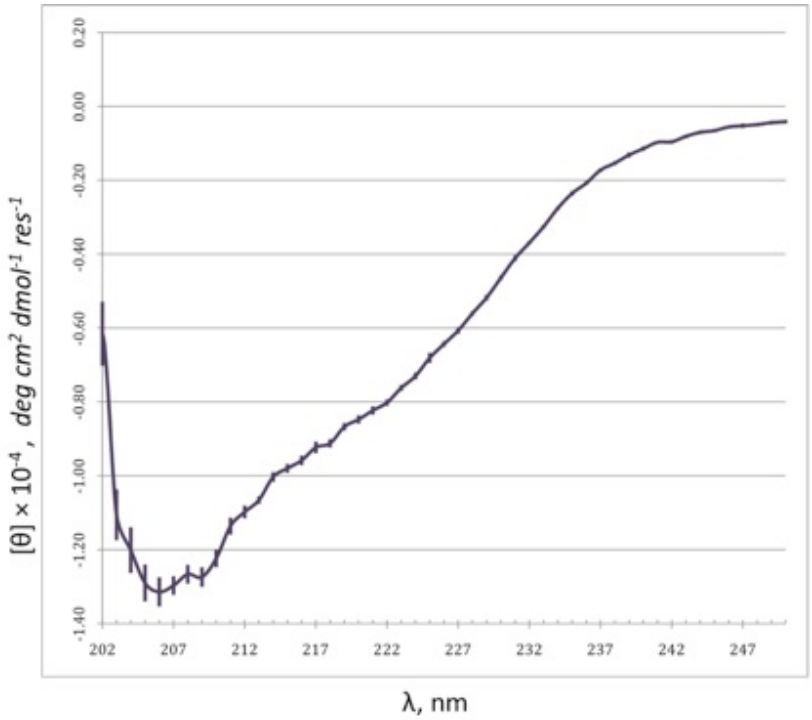


Figure B-13. CD wavelength scan of GP1,2. Samples contained 3.7 μM protein in buffer containing 2.5 mM sodium phosphate and 35 mM sodium chloride at pH 7.4. GP1,2 exhibited a negative band at about 207 nm with a broad shoulder at ~220 nm, consistent with a structure containing α- and β-secondary structure.

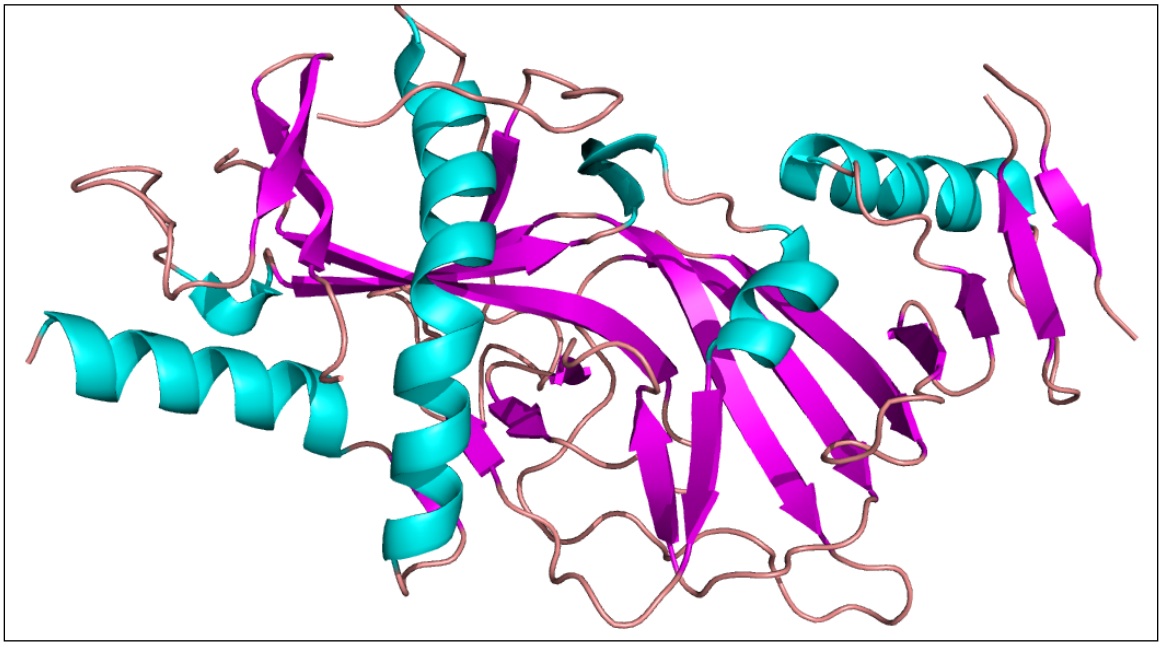


Figure B-14. GP1,2 colored by secondary structure. The crystal structure of GP1,2 [11] shows both α- (cyan) and β- (magenta) secondary structural elements overlapping, consistent with the obtained CD wavelength scan in Fig. 13.

References

1. Lodish, H., et al., *Molecular Cell Biology*. 6th ed, ed. K. Ahr. 2008, New York: W. H. Freeman and Company.

2. Fields, B.N., D.M. Knipe, and P.M. Howley, *Fields virology*. 5th ed. 2007, Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

3. Strauss, J.H. and E.G. Strauss, *Viruses and human disease*. 2nd ed. 2008, Amsterdam; Boston: Elsevier/Academic Press.

4. Sullivan, N.J., et al., *Development of a preventive vaccine for Ebola virus infection in primates.* Nature, 2000. **408**(6812): p. 605-609.

5. Sullivan, N.J., et al., *Correlates of protective immunity for Ebola vaccines: implications for regulatory approval by the animal rule.* Nature Reviews Microbiology, 2009. **7**(5): p. 393-400.

6. Peters, C.J., et al., in *Fields virology*, B.N. Fields, D.M. Knipe, and P.M. Howley, Editors. 1996, Lippincott-Raven Press: Philadelphia.

7. Jones, S.M., et al., *Live attenuated recombinant vaccine protects nonhuman primates against Ebola and Marburg viruses.* Nature Medicine, 2005. **11**(7): p. 786-790.

8. Hampton, T., *Vaccines against Ebola and Marburg viruses show promise in primate studies.* Jama-Journal of the American Medical Association, 2005. **294**(2): p. 163-164.

9. Kobinger, G.P., et al., *Chimpanzee adenovirus vaccine protects against Zaire Ebola virus.* Virology, 2006. **346**(2): p. 394-401.

10. Sullivan, N.J., et al., *Immune protection of nonhuman primates against Ebola virus with single low-dose adenovirus vectors encoding modified GPs.* Plos Medicine, 2006. **3**(6): p. 865-873.

11. Lee, J.E., et al., *Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor.* Nature, 2008. **454**(7201): p. 177-U27.

12. Barrientos, L.G. and A.M. Gronenborn, *The highly specific carbohydrate-binding protein cyanovirin-N: Structure, anti-HIV/Ebola activity and possibilities for therapy.* Mini-Reviews in Medicinal Chemistry, 2005. **5**(1): p. 21-31.

13. Boyd, M.R., et al., *Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: Potential applications to microbicide development.* Antimicrobial Agents and Chemotherapy, 1997. **41**(7): p. 1521-1530.

14. Tsai, C.C., et al., *Cyanovirin-N inhibits AIDS virus infections in vaginal transmission models.* Aids Research and Human Retroviruses, 2004. **20**(1): p. 11-18.

15. Tsai, C.C., et al., *Cyanovirin-N gel as a topical microbicide prevents rectal transmission of SHIV89.6P in macaques.* Aids Research and Human Retroviruses, 2003. **19**(7): p. 535-541.

16. Balzarini, J., et al., *Mutational pathways, resistance profile, and side effects of cyanovirin relative to human immunodeficiency virus type 1 strains with N-glycan deletions in their gp120 envelopes.* Journal of Virology, 2006. **80**(17): p. 8411-8421.

17. Smee, D.F., et al., *Treatment of influenza A (H1N1) virus infections in mice and ferrets with cyanovirin-N.* Antiviral Research, 2008. **80**(3): p. 266-271.

18. O'Keefe, B.R., et al., *Potent anti-influenza activity of cyanovirin-N and interactions with viral hemagglutinin.* Antimicrobial Agents and Chemotherapy, 2003. **47**(8): p. 2518-2525.

19. Helle, F., et al., *Cyanovirin-N inhibits hepatitis C virus entry by binding to envelope protein glycans.* Journal of Biological Chemistry, 2006. **281**(35): p. 25177-25183.

20. Dey, B., et al., *Multiple antiviral activities of cyanovirin-N: Blocking of human immunodeficiency virus type 1 gp120 interaction with CD4 and coreceptor and inhibition of diverse enveloped viruses.* Journal of Virology, 2000. **74**(10): p. 4562-4569.

21. Barrientos, L.G., et al., *Cyanovirin-N binds to the viral surface glycoprotein, GP(1,2) and inhibits infectivity of Ebola virus.* Antiviral Research, 2003. **58**(1): p. 47-56.

22. Bewley, C.A. and S. Otero-Quintero, *The potent anti-HIV protein cyanovirin-N contains two novel carbohydrate binding sites that selectively bind to man(8) D1D3 and Man(9) with nanomolar affinity: Implications for binding to the HIV envelope protein gp120.* Journal of the American Chemical Society, 2001. **123**(17): p. 3892-3902.

23. Bolmstedt, A.J., et al., *HIV-inhibitory natural products part 71 - Cyanovirin-N defines a new class of antiviral agent targeting N-linked, high-mannose glycans in an oligosaccharide-specific manner.* Molecular Pharmacology, 2001. **59**(5): p. 949-954.

24. Shenoy, S.R., et al., *Multisite and multivalent binding between cyanovirin-N and branched oligomannosides: Calorimetric and NMR characterization.* Chemistry & Biology, 2002. **9**(10): p. 1109-1118.

25. Shenoy, S.R., et al., *Selective interactions of the human immunodeficiency virus-inactivating protein cyanovirin-N with high-mannose oligosaccharides on gp120 and other glycoproteins.* Journal of Pharmacology and Experimental Therapeutics, 2001. **297**(2): p. 704-710.

26. Mori, T., et al., *Functional homologs of cyanovirin-N amenable to mass production in prokaryotic and eukaryotic hosts.* Protein Expression and Purification, 2002. **26**(1): p. 42-49.

27. Mori, T., et al., *Recombinant production of cyanovirin-N, a potent human immunodeficiency virus inactivating protein derived from a cultured Cyanobacterium.* Protein Expression and Purification, 1998. **12**(2): p. 151-158.

28. Colleluori, D.M., et al., *Expression, purification, and characterization of recombinant cyanovirin-N for vaginal anti-HIV microbicide development.* Protein Expression and Purification, 2005. **39**(2): p. 229-236.

29. Hu, Q.X., N. Mahmood, and R.J. Shattock, *High-mannose-specific deglycosylation of HIV-1 gp120 induced by resistance to cyanovirin-N and the impact on antibody neutralization.* Virology, 2007. **368**(1): p. 145-154.

30. Witvrouw, M., et al., *Resistance of human immunodeficiency virus type 1 to the high-mannose binding agents cyanovirin N and concanavalin A.* Journal of Virology, 2005. **79**(12): p. 7777-7784.

31. Reitter, J.N., R.E. Means, and R.C. Desrosiers, *A role for carbohydrates in immune evasion in AIDS.* Nature Medicine, 1998. **4**(6): p. 679-684.

32. Sanchez, A., et al., *The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(8): p. 3602-3607.

33. Sanchez, A., et al., *Biochemical analysis of the secreted and virion glycoproteins of Ebola virus.* Journal of Virology, 1998. **72**(8): p. 6442-6447.

34. Jeffers, S.A., D.A. Sanders, and A. Sanchez, *Covalent modifications of the Ebola virus glycoprotein.* Journal of Virology, 2002. **76**(24): p. 12463-12472.

35. Volchkov, V., et al., *Processing of the Ebola virus glycoprotein by the proprotein convertase furin.* Proceedings of the National Academy of Sciences, 1998. **95**(10): p. 5762-5767.

36. Takada, A., et al., *A system for functional analysis of Ebola virus glycoprotein.* Proc Natl Acad Sci U S A, 1997. **94**(26): p. 14764-9.

37. Wool-Lewis, R.J. and P. Bates, *Characterization of Ebola virus entry by using pseudotyped viruses: Identification of receptor-deficient cell lines.* Journal of Virology, 1998. **72**(4): p. 3155-3160.

38. Yang, Z., et al., *Identification of the Ebola virus glycoprotein as the main viral determinant of vascular cell cytotoxicity and injury.* Nature Medicine 2000. **6**(8): p. 886-889.

39. Simmons, G., et al., *Ebola virus glycoproteins induce global surface protein down-modulation and loss of cell adherence.* Journal of Virology, 2002. **76**(5): p. 2518-2528.

40. Ray, R.B., et al., *Ebola virus glycoprotein-mediated anoikis of primary human cardiac microvascular endothelial cells.* Virology, 2004. **321**(2): p. 181-188.

41. Francica, J.R., M.K. Matukonis, and P. Bates, *Requirements for cell rounding and surface protein down-regulation by Ebola virus glycoprotein.* Virology, 2009. **383**(2): p. 237-247.

42. Das, D., et al., *Differential expression of the Ebola virus GP(1,2) protein and its fragments in E-coli.* Protein Expression and Purification, 2007. **54**(1): p. 117-125.

43. Keeffe, J.K., *Engineering Cyanovirin-N for Enhanced Viral Neutralization*, in *Chemistry*. 2009, California Institute of Technology: Pasadena. p. 200.

44. Stemmer, W.P.C., et al., *Single-Step Assembly of a Gene and Entire Plasmid from Large Numbers of Oligodeoxyribonucleotides.* Gene, 1995. **164**(1): p. 49-53.

45. Prehaud, C., et al., *Recombinant Ebola virus nucleoprotein and glycoprotein (Gabon 94 strain) provide new tools for the detection of human infections.* Journal of General Virology, 1998. **79**: p. 2565-2572.

46. Kawar, Z., et al., *N-Glycan processing by a lepidopteran insect alpha 1,2-mannosidase.* Glycobiology, 2000. **10**(4): p. 347-355.

47. *Circular Dichroism and the Conformational Analysis of Biomolecules*, ed. G.D. Fasman. 1996, New York: Plenum Press.

48. Bewley, C.A., *Solution structure of a cyanovirin-N : Man alpha 1-2Man alpha complex: Structural basis for high-affinity carbohydrate-mediated binding to gp120.* Structure, 2001. **9**(10): p. 931-940.