

SYNTHESIS AND EVALUATION OF PEPTIDIC PROBES FOR TISSUE
TRANSGLUTAMINASE AND FACTOR XIIIa

AMINA MULANI

Thesis submitted to the
Faculty of Graduate & Postdoctoral Studies
University of Ottawa
in partial fulfillment of the requirements for the
Ph.D. degree in the

Ottawa-Carleton Chemistry Institute
Faculty of Science
University of Ottawa

University of Ottawa
Faculty of Graduate & Postdoctoral Studies

This thesis titled:

Synthesis and evaluation of peptidic probes for tissue transglutaminase and Factor
XIIIa

presented by:

Amina Mulani

was evaluated by a jury composed of the following professors:

Prof. Jeffrey. W Keillor

Supervisor

Prof. Christopher N. Boddy

Jury member

Prof. Roberto A. Chica

Jury member

Prof. Peter H. Buist

Jury member

Professor Stephen L. Bearne

External examiner

Thesis accepted: Wednesday May 21st, 2014

Transglutaminases (TGases) are a group of enzymes that catalyze the formation of an amide bond between the γ -carboxamide group of a glutamine residue and an amine donor, usually an ϵ -amino group of the lysine residue, leading to the formation of ϵ -(γ -glutamyl)lysine crosslinks. Owing to the roles that transglutaminases such as tissue transglutaminase (TG2) and Factor XIIIa (FXIIIa) have been found to play in a wide range of disease states, efforts have been directed towards the study of these proteins. The study of enzymes to better understand their function and mode of action is facilitated through the use of tools such as protein labelling, enzyme inhibition, and substrate analogue kinetic studies among others.

Transition state analogues have been effective inhibitors in the study of enzyme activity. Sulfoxide inhibitors can efficiently mimic transition states leading to the tetrahedral intermediate of an acyl transfer reaction and we discuss the synthesis towards sulfoxide transition state analogue inhibitors of TG2 in chapter 2. Novel sulfoxide compounds were synthesized, though the desired target compounds proved difficult to isolate due to their instability.

Fluorescent probes are effective in protein labelling as a means of discerning activity. This technique was applied in order to elucidate intracellular TG2 activity, which is a topic of controversy. To that end, the synthesis of a fluorescent, TG2-specific, cell permeable probe is discussed in chapter 3. However, preliminary *in vivo* results show that while the probe is cell permeable and fluorescent, it was not TG2-specific. Molecular modelling suggests that the hexa-arginine tag, designed to improve cell permeability, decreases the affinity of the probe for its intended target.

Finally, FXIIIa has become a new addition to the study of transglutaminases in the Keillor group. Given our interest in this enzyme, we had three goals for this work as explained in chapter 4. Firstly, owing to the anticipated high demand for FXIIIa required for later experiments, our primary aim was the development of an optimized method for the expression and purification of recombinant FXIIIa. After evaluating different conditions for FXIIIa expression, the Studier auto-induction ZYP media¹ at 20 °C for 24 h was found to provide the optimal conditions for the expression of recombinant GST-tagged FXIIIa, typically giving a total of 1.5 mg of protein/L of culture.

Secondly, a variety of different peptides were synthesized and tested using a glutamate dehydrogenase (GDH)-based assay to identify a high affinity sequence for a substrate of FXIIIa. The two peptides with the highest affinity for FXIIIa were Ac-DQMMMAF-OH and Ac-DQMML-OH. Testing with TG2 displayed negligible reactivity, confirming their use as orthogonal peptides, results reinforced by modelling studies of the peptides with both FXIIIa and TG2. This discovery represents the first time peptides orthogonal to TG2 with affinity for FXIIIa have been kinetically characterized with both transglutaminase enzymes.

Lastly, our work towards a fluorogenic activity assay by incorporating a coumarin ester through attachment to a glutamic acid residue into a peptide sequence recognized by FXIIIa, will be discussed.

Key words: Enzymes, Transglutaminase, Inhibition, Transition state analogues, Fluorescent probes, High-affinity peptides

Table of Contents

Abstract.....	iii
Table of Contents	v
List of Figures.....	x
List of Schema	xiii
List of Tables	xiv
List of Abbreviations.....	xv
Acknowledgements.....	xix
 CHAPTER 1: Introduction.....	 1
 1.1 Enzymes and catalysis.....	 2
 1.2 Transglutaminases.....	 2
1.2.1 Keratinocyte transglutaminase (Type 1)	5
1.2.2 Tissue transglutaminase (Type 2)	6
1.2.3 Epidermal transglutaminase (Type 3)	6
1.2.4 Plasma transglutaminase (FXIII in humans)	6
1.2.5 Prostate transglutaminase (Type 4)	7
1.2.6 Transglutaminase 6 (Type 6)	7
 1.3 Tissue transglutaminase.....	 7
1.3.1 Effect of GDP and calcium	8
1.3.2 Structure of TG2	9
1.3.3 Physiological Roles of TG2.....	10
1.3.4 Pathological associations of TG2	12
1.3.4.1 TG2 in Celiac disease.....	13
1.3.4.2 TG2 in Alzheimer's disease.....	13
1.3.4.3 TG2 in Huntington's disease	14
1.3.5 Transglutaminase 2 substrates	15
 1.4 Factor XIII.....	 17
1.4.1 Activation of FXIIIA.....	18
1.4.2 Structure of FXIIIA	21
1.4.3 Physiological roles of FXIII	24
1.4.4 FXIII in disease states	25
1.4.5 Substrates of FXIII.....	28

1.5 Research objectives	29
1.5.1 Transition state analogue inhibitors for TG2	30
1.5.2 Fluorescent, cell permeable probe of TG2	31
1.5.3 Studies with FXIIIa	31
CHAPTER 2: Sulfoxide Inhibitors.....	45
2.1 Enzymatic transition states.....	46
2.2 Transition state analogues.....	48
2.2.1 Examples of TS analogues.....	51
2.3 Tissue transglutaminase	52
2.3.1 Irreversible inhibitors of TG2	52
2.3.2 Reversible inhibitors of TG2	54
2.3.3 PNPQLPF as affinity-conferring peptidic sequence for TG2	54
2.3.4 Transition state analogue design for TG2	55
2.4 S-N bond-containing compounds	57
2.4.1 Synthesis of sulfinamides.....	59
2.5 Modelling of the TG2 active site.....	61
2.6 Synthesis of sulfinamide inhibitors.....	64
2.6.1 Synthesis of sulfinamide (2.9) via a methyl sulfinate	64
2.6.2 Synthesis of sulfinamide employing sulfinyl chloride	67
2.6.3 Synthesis of sulfinamide using oxazolidinone	68
2.6.4 Synthesis of aminomethyl sulfoxide (2.10).....	69
2.6.5 Synthesis of aminomethyl sulfoxide via thiol formate	71
2.7 Protolytic cleavage and instability of sulfinamides.....	72
2.8 Conclusions and Perspectives	75
CHAPTER 3: Cell Permeable, Fluorescent TG2 Probe.....	80
3.1 The role of TG2 inside cells	81
3.2 Design of TG2 Probes	83
3.2.1 Fluorophore choice.....	83
3.2.2 Electrophilic warhead	84
3.2.3 TG2 affinity	85

3.2.4	Cell permeable peptides	86
3.2.4.1	HIV-Tat peptide	86
3.2.4.2	Penetratin	87
3.2.4.3	Polyarginine sequences	88
3.2.4.4	Mode of entry	89
3.2.4.5	Various applications of cell permeable peptides	92
3.3	TG2-Specific, fluorescent, cell permeable probes	94
3.4	Synthesis of fluorescent, cell permeable probes	95
3.4.1	Synthesis of Ac-R ₈ K(dansyl)GGK(acryl)LPF-OH	95
3.4.2	Synthesis of Ac-R ₆ K(dansyl)GGK(acryl)LPF-OH	99
3.5	<i>In vitro</i> activity of Ac-R₆K(dansyl)GGK(acryl)LPF-OH with recombinant human TG2	100
3.6	Preliminary <i>in vivo</i> results with SH-SY5Y neuroblastoma cells.....	103
3.7	Modelling of the fluorescent probe with TG2	106
3.8	Conclusions and perspectives	108
CHAPTER 4:	FXIIIa	115
4.1	Factor XIIIa (FXIIIa) as a protein target.....	116
4.1.1	Methods of expression of FXIII	117
4.1.2	Existing methods to test activity of FXIIIa	119
4.1.3	FXIIIa-specific peptides.....	121
4.1.4	FXIIIa-specific activity assay.....	123
4.2	Research goals	125
4.3	Optimization of purification	125
4.4	High affinity peptide sequences for FXIIIa.....	132
4.4.1	Design of high affinity peptide sequences	133
4.4.2	Synthesis of peptides	135
4.4.3	Optimization of FXIIIa activation by thrombin.....	136
4.4.4	Kinetic analysis of FXIIIa substrate peptides	137
4.4.5	Modelling of FXIIIa peptidic sequence with FXIIIa	144
4.4.6	Modelling of peptides with TG2	146

4.5 Synthesis of Ac-DE(O-coumarin)MM-OH	148
4.5.1 First attempt towards activity assay using Asp(OAllyl)-OH	149
4.5.2 Second attempt towards activity assay using FmocGlu(O-coumarin)-OH	150
4.5.3 Third attempt towards activity assay using FmocGlu(OTce)-OH.....	153
4.6 Conclusions and perspectives	156
CHAPTER 5: Perspectives.....	162
CHAPTER 6: Experimental.....	169
6.0 Experimental Instrumentation	170
6.2 Sulfoxide Inhibitors	171
6.2.1 Synthesis of sulfinamide via a methyl sulfinite	171
6.2.2 Synthesis of sulfinamide via a sulfinyl chloride	177
6.2.3 Synthesis of sulfinamide via oxazolidinone	178
6.2.4 Synthesis of amino methyl sulfoxide.....	181
6.2.5 Synthesis of amino methyl sulfoxide via thiol formate	187
6.3 Cell permeable, fluorescent TG2 Probe	188
6.3.1 General procedure for synthesis of Ac-R ₈ K(dansyl)GGK(acryl)LPF-OH	189
6.3.2 Synthesis of Fmoc-K(dansyl)-OH.....	191
6.3.3 Synthesis of Ac-R ₆ K(dansyl)GGK(acryl)LPF-OH.....	196
6.3.4 Kinetic analysis of cell permeable peptide	198
6.3.4.1 GDH assay	198
6.3.4.2 AL5 assay.....	199
6.3.5 Preliminary <i>in vivo</i> results with Ac-R ₆ K(dansyl)GGK(acryl)LPF-OH	200
6.3.6 Computer modelling of Ac-R ₆ K(dansyl)GGK(acryl)LPF-OH with TG2	201
6.4 FXIIIa.....	202
6.4.1 Optimization of expression of GST-FXIIIa.....	203
6.4.1.1 Expression of GST-FXIIIa	203
6.4.1.2 Purification of GST-FXIIIa.....	203
6.4.1.3 Protein concentration.....	204
6.4.1.4 Electrophoresis.....	204
6.4.2 Synthesis of high affinity peptide sequences for FXIIIa.....	205
6.4.3 Optimization of thrombin activation	208
6.4.4 Kinetic analysis of FXIIIa substrate peptides	208

6.4.5	Modelling of peptides with FXIIIa and TG2.....	209
6.4.6	Towards synthesis of Ac-DE(O-coumarin)MM-OH.....	210
6.4.6.1	First attempt towards activity assay using Fmoc-Asp(O-allyl)-OH..	212
6.4.6.2	Second attempt using Fmoc-Glu(O-coumarin)-OH.....	214
6.4.6.3	Third attempt towards activity assay using Fmoc-Glu(OTce)-OH....	220

List of Figures

Figure 1.1: TGase-catalyzed formation of amide bond.	3
Figure 1.2: Ping-pong catalytic cycle of TGase.	3
Figure 1.3: Detailed mechanism of transglutaminases.	5
Figure 1.4: [A] Crystal structure of TG2 in closed conformation. [B] TG2 in open conformation.	10
Figure 1.5: [A] Activation of FXIIIA in the presence of thrombin, fibrin and Ca^{2+} . [B] Activation of FXIII without cleavage of activation peptide at high Ca^{2+} concentration.	20
Figure 1.6: [A] Crystal structure of FXIIIA in an inactive state. [B] FXIIIA ^o in an active conformation.	23
Figure 1.7: Target sulfinamide inhibitors for TG2.....	30
Figure 2.1: Energy profile of catalysis.....	46
Figure 2.2: Molecular electrostatic potential surfaces for inosine, the transition state of PNP and Immucillin-H.....	50
Figure 2.3: TS analogue for MTAN.....	51
Figure 2.4: TS analogue for γ -glutamyl transpeptidase.	52
Figure 2.5: Irreversible inhibitors of TG2.	53
Figure 2.6: Reversible inhibitors of TG2.	54
Figure 2.7: Mechanism of TGase acylation.	56
Figure 2.8: Transition state analogues for TG.	56
Figure 2.9: Sulfenamides, sulfinamides and sulfonamides.	57
Figure 2.10: Fragmentation of peptidic sulfonamides.	58
Figure 2.11: Retrosynthetic schemes of sulfinamide synthesis.	60
Figure 2.12: Active site of TG2 with Cbz-Gln-Gly.	62
Figure 2.13: Charge distribution of the tetrahedral intermediate of acyl-enzyme.	62
Figure 2.14: Charge distribution around hypothetical molecule resembling TS intermediate.	63
Figure 2.15: Charge distribution around sulfinamide molecule.	63
Figure 2.16: Two-step addition-elimination mechanism of sulfinamide hydrolysis.	72
Figure 2.17: Attempted deprotection of <i>tert</i> -butyl phenylalaninate sulfinamide.	73

Figure 2.18: Peptidic sulfinamide reduction via a cyclic intermediate.	74
Figure 2.19: Sulfinic acid derivatives as TS analogues.	75
Figure 3.1: Various electrophilic warheads of TG2.	84
Figure 3.2: Michael addition of acrylic amide to TG2.	85
Figure 3.3: Tat peptide sequences evaluated for permeability.	87
Figure 3.4: Linear vs. branched chains of octa-arginine peptides.	89
Figure 3.5: (left) Fluorescence and confocal microscopy of β -galactoside fused with Tat peptide and (right) concentration (left axis) and enzymatic activity (right axis) of cells treated with Tat- β -galactoside.	92
Figure 3.6: Dual-labelled probe made permeable through incorporation of Tat peptide.	93
Figure 3.7: TG2-specific, cell permeable, fluorescent probe.	94
Figure 3.8: Schematic of the GDH assay with Cbz-Gln-Gly substrate.	100
Figure 3.9: Reciprocal of observed rate vs. Ac-R ₆ -K(dansyl)GGK(acryl)LPF-OH.	101
Figure 3.10: Activity assay measuring release of p-nitrophenolate.	102
Figure 3.11: (left) Phase contrast of retinoic acid-treated cells with 10 μ M cell permeable peptide, (right) under blue light.	104
Figure 3.12: (left) Phase contrast of cells not treated with retinoic acid with 10 μ M of cell permeable peptide, (right) under blue light.	105
Figure 3.13: Cell permeable peptide docked into active site of TG2.	107
Figure 3.14: Non-linear, cell permeable, fluorescent TG2-specific probe.	109
Figure 3.15: Lysine-branched, cell permeable, fluorescent TG2-specific probe.	110
Figure 4.1: GST-FXIIIa plasmid.	118
Figure 4.2: Fluorogenic FRET-based Zedira assay.	120
Figure 4.3: Glutamate dehydrogenase (GDH) assay.	121
Figure 4.4: Alignment of preferred substrate peptide sequences of phage clones for FXIIIa.	122
Figure 4.5: Fluorometric assay for TG2 with 7-hydroxycoumarin.	124
Figure 4.6: Fluorogenic activity assay for FXIIIa shown incorporated into Ac-DEMM-OH as a model peptide.	124
Figure 4.7: Optimization of expression of FXIIIa.	126
Figure 4.8: Different temperatures of expression in TB media for FXIIIa.	127
Figure 4.9: SDS-PAGE gels of FXIIIa expression at different temperatures.	128

Figure 4.10: Total protein obtained with different concentrations of IPTG in TB media.	128
Figure 4.11: SDS-PAGE gels of FXIIIA expression at different IPTG concentrations.	129
Figure 4.12: Total protein obtained with different expression media.	130
Figure 4.13: SDS-PAGE gels of FXIIIA expression with TB and ZYP media.	131
Figure 4.14: GDH assay with glycine ethyl ester as acceptor substrate.	133
Figure 4.15: SDS-PAGE gel of FXIIa with differing quantities of thrombin.	136
Figure 4.16: Rate vs. concentration of Ac-DQMML-OH.	138
Figure 4.17: Rate vs. concentration of Ac-GQSMV-OH.	139
Figure 4.18: Rate vs. concentration of Ac-SQSRV-OH.	139
Figure 4.19: Rate vs. concentration of Ac-DQMHV-OH.	140
Figure 4.20: Rate vs. concentration of Ac-SQMM-OH.	140
Figure 4.21: Rate vs. concentration of Ac-DQMMAF-OH.	141
Figure 4.22: Ratio of $k_{\text{cat}}/K_{\text{m}}$ of peptides with FXIIa.	141
Figure 4.23: Comparison of affinity between FXIIa and TG2.	143
Figure 4.24: (left) Ac-DQMML docked into FXIIa, (right) Ac-DQMMAF docked into FXIIa.	145
Figure 4.25: (left) Ac-DQMML docked into TG2, (right) Ac-DQMMAF docked into TG2.	147
Figure 4.26: Schematic of quenched activity probe for FXIIa.	155
Figure 5.1: TG2-Specific, cell permeable, fluorescent probe.	164
Figure 5.2: Modified (3.16) and branched (3.17) cell-permeable, fluorescent TG2-specific probes.	165

List of Schema

Scheme 2.1: Synthesis of propyl sulfinamide.	60
Scheme 2.2: Synthesis of compound 2.8 via methyl sulfinite.	65
Scheme 2.3: Sulfinamide synthesis via sulfinyl chloride.	67
Scheme 2.4: Synthesis of sulfinamide using oxazolidinone.	68
Scheme 2.5: Synthesis of aminomethyl sulfoxide.	69
Scheme 2.6: Synthesis of aminomethyl sulfoxide via thiol formate.	71
Scheme 3.1: Synthesis of Ac-R ₈ K(dansyl)GGK(acryl)LPF-OH.	96
Scheme 3.2: Synthesis of Fmoc-K(N-dansyl)-OH.	96
Scheme 3.3: Microwave synthesis of Ac-R ₆ K(dansyl)GGK(acryl)LPF-OH.	99
Scheme 4.1: Synthesis of FXIIIa substrate peptides	135
Scheme 4.2: First attempt using Fmoc-Asp(O-allyl)-OH orthogonal to glutamic acid.	149
Scheme 4.3: Second attempt using coumarin directly attached to glutamic acid.	151
Scheme 4.4: Third attempt using trichloroethyl ester as protecting group.	154

List of Tables

Table 2.1: Conditions attempted in synthesis of sulfinamide.	66
Table 2.2: Deprotection conditions for Boc removal.	70
Table 3.1: Reaction conditions for Fmoc-K(N-dansyl)-OH coupling.	97
Table 3.2: Evaluation of arginine coupling conditions.	98
Table 4.1: Peptidic sequences designed for FXIIIa.	134
Table 4.2: Units of thrombin correlated to specific activity of FXIIIa.	137
Table 4.3: Reaction conditions for 7-hydroxycoumarin coupling.	152

List of Abbreviations

Å	ångström
Ac ₂ O	acetic anhydride
ACN	acetonitrile
Boc	<i>tert</i> -butoxycarbonyl
Boc ₂ O	di- <i>tert</i> -butyl dicarbonate
BSA	bovine serum albumin
Cbz	carboxybenzyl
Cbz-Glu-Gly	<i>N</i> -α-benzyloxycarbonyl-L-glutaminylglycine
Da	atomic mass unit (1 Da = 1 g/mol)
DAPI	4'6-diamidino-2-phenylindole
dba	dibenzylideneacetone
DCC	<i>N,N'</i> -dicyclohexylcarbodiimide
DCE	dichlorethane
DCM	dichloromethane
DIC	<i>N,N'</i> -diisopropylcarbodiimide
DIEA	<i>N,N</i> -diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMEM	Dulbecco's Modified Eagle Medium
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
DON	6-diazo-5-oxo-L-leucine
DTT	dithiothreitol
EEDQ	<i>N</i> -ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline
eq.	equivalents
Et	ethyl
EtOAc	ethyl acetate
FAB	fast atom bombardment
FBS	fetal bovine serum
Fmoc	9-fluorenylmethoxycarbonyl
FRET	fluorescence resonance energy transfer
FXIIIa	Factor XIIIa
GDH	glutamate dehydrogenase
GDP/GTP	guanosine diphosphate/triphosphate

GGT	γ -glutamyl transpeptidase
GST	glutathione S-transferase
IC ₅₀	half maximal inhibitory concentration
HATU	1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HBTU	<i>N,N,N',N'</i> -tetramethyl- <i>O</i> -(1 <i>H</i> -benzotriazol-1-yl)uronium hexafluorophosphate
HOBt	hydroxybenzotriazole
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry
IPTG	isopropyl β -D-1-thiogalactopyranoside
k_{cat}	rate constant of catalysis
K_i	inhibitor constant (reversible)
K_I	inhibitor constant (irreversible)
K_m	Michaelis-Menten constant
LB	Luria-Bertani
LCMS	liquid chromatography-mass spectrometry
LiHMDS	lithium bis(trimethylsilyl)amide
LRMS	low-resolution mass spectrometry
M.P.	melting point
MALDI	matrix-assisted laser desorption/ionization
mCPBA	<i>meta</i> -chloroperoxybenzoic acid
Me	methyl
mL	millilitre
MOE	Molecular Operating Environment
mol	mole
MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid
MS	mass spectrometry
MTAN	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase
MW	microwave irradiation
NAD ⁺	nicotinamide adenine dinucleotide
NADH	nicotinamide adenine dinucleotide, reduced form
NBS	<i>N</i> -bromosuccinimide
nBu	<i>n</i> -butyl
NMP	<i>N</i> -methyl-2-pyrrolidone

NMR	nuclear magnetic resonance
O/N	overnight
OAc	acetate
PAGE	polyacrylamide gel electrophoresis
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PBS	phosphate buffered saline
PDB	Protein Data Bank
PNP	purine nucleoside phosphorylase
PyBOP	(benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
RA	retinoic acid
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
Tat	transcription-transactivating
TB	Terrific broth
TBAF	tetra- <i>n</i> -butylammonium fluoride
<i>t</i> Bu	<i>tert</i> -butyl
Tce	trichloroethyl
TFA	trifluoroacetic acid
TG2	tissue transglutaminase
TGase	transglutaminase
THF	tetrahydrofuran
TLC	thin-layer chromatography
TMS	tetramethylsilane
TMSOTf	trimethylsilyl trifluoromethanesulfonate
Tris	tris(hydroxymethyl)aminomethane
TS	transition state
$V_{\max}$	maximum rate of product formation from the enzyme-substrate complex
vs.	versus
ZYP	N-Z amine, yeast, phosphate auto-induction media

To my family,

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Jeff Keillor, for his support, motivation, patience and understanding. The knowledge he shared and encouraged me to develop during my PhD study was invaluable and I am extremely grateful for his consistent guidance. Thus, I would like to express my gratitude to Professor Keillor for accepting me as a student and sharing his lab and expertise.

I would like to thank all of the members of the Keillor group, both past and present, for always being willing to help, and for making our lab such an enjoyable place to work. You never knew the discussions that awaited you each day. I would like to specifically express my gratitude to Drs. Christophe Pardin and Christopher Clouthier for being amazing resources, suggesting useful ideas and providing much-valued support.

I would also like to express my appreciation to the departments of chemistry at the Université de Montréal and the University of Ottawa, without whose support, the work presented herein would not have been possible. I would like to particularly thank Marie-Christine Tang (Université de Montréal) and Sharon Curtis (University of Ottawa) for their support and the skills they shared during instrument training.

Lastly, I would like to extend a huge thanks to my family and friends for their support and encouragement all throughout my PhD and for always having confidence and faith in me.

CHAPTER 1

Introduction

1.1 Enzymes and catalysis

Enzymes are proteins that are specific catalysts, increasing the rate of a reaction without being consumed in the process. They function by lowering the activation energy for a target reaction and generally proceed with a high degree of stereo- and substrate specificities. Many enzymatic reaction rates are millions of times faster than their un-catalyzed reactions.² One example is human carbonic anhydrase, whose reaction rate is typically limited by the diffusion rate of its substrates.³ Enzymes are responsible for the thousands of metabolic processes that sustain life and any malfunction can lead to the development of a number of disease states, highlighting the importance of these biomolecules. The study of enzymes to better understand their function is facilitated through the use of tools such as protein labelling, enzyme inhibition, and substrate analogue kinetic studies among others. An enzyme family of interest in the Keillor group is the transglutaminases, protein cross-linking enzymes that will be discussed in detail below.

1.2 Transglutaminases

Transglutaminases (TGases) are a group of enzymes that catalyze the formation of an amide bond between the γ -carboxamide group of a glutamine residue and an amine donor, usually the ϵ -amino group of a lysine residue, leading to the formation of ϵ -(γ -glutamyl)lysine crosslinks (Figure 1.1).

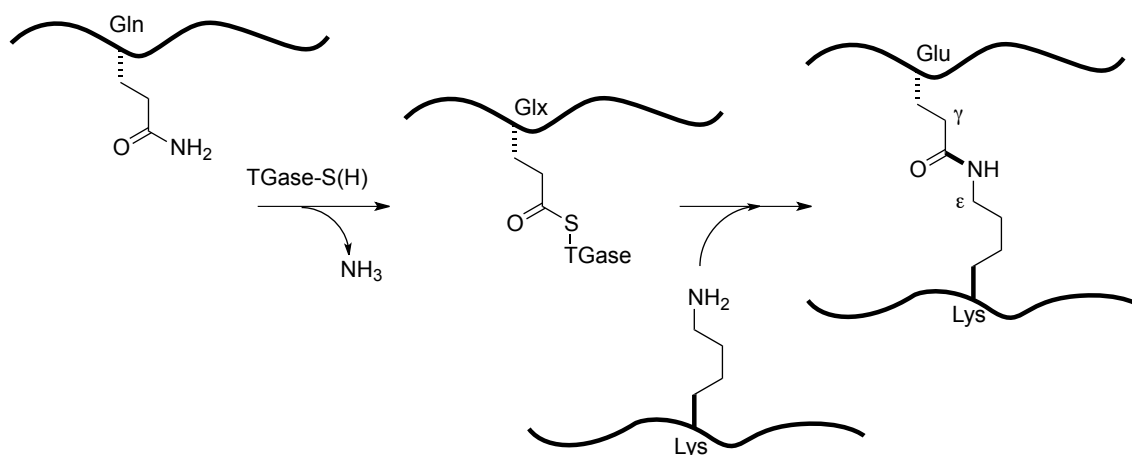


Figure 1.1: TGase-catalyzed formation of amide bond.

The formation of the isopeptide bond is accompanied by the release of ammonia. The mechanism of bond formation is conserved for all transglutaminases, whereby the catalytic acyl-transfer reaction proceeds through a modified ping-pong mechanism (Figure 1.2).^{4,5}

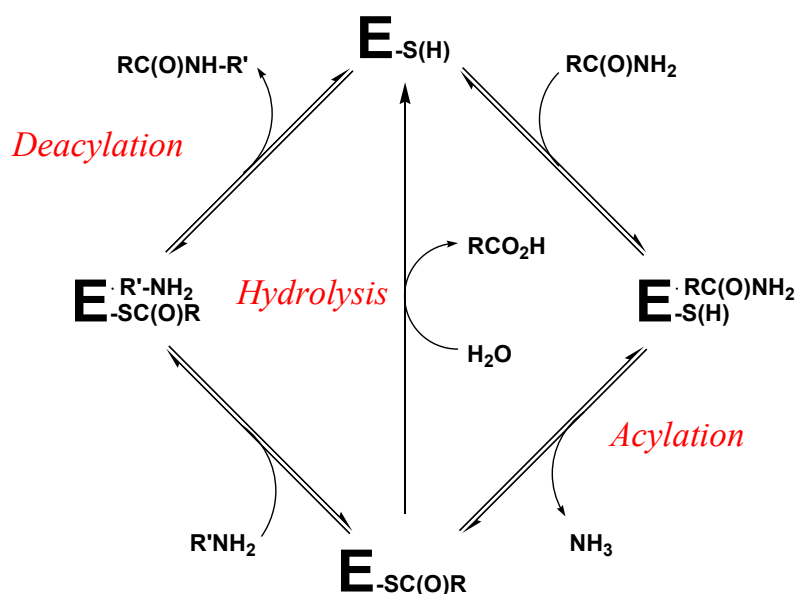


Figure 1.2: Ping-pong catalytic cycle of TGase.

In the first step of the catalytic cycle, the acyl-donor substrate undergoes nucleophilic attack by the active site thiolate, followed by the release of an equivalent of ammonia. The thioester acyl-enzyme intermediate obtained is then subjected to a second nucleophilic attack by the acyl-acceptor substrate, in the rate-determining step. Incorporation of the acyl-acceptor substrate leads to the formation of an amide bond and the regeneration of free enzyme. It is described as a modified ping-pong mechanism because in the absence of an amine donor, the thioester intermediate will be hydrolyzed by water to give the carboxylic acid.

Numerous studies have shown that the active site residues, cysteine, histidine and aspartic acid, form the catalytic triad for TGases and are conserved throughout the enzyme family. The occurrence of amino acids around the catalytic cysteine is also mostly conserved, suggesting the same mechanism. The conserved histidine in proximity to the catalytic cysteine residue suggests its interaction as a general acid and/or base. Additionally, the presence of a tyrosine in proximity to the catalytic cysteine has been proposed to be associated with a stabilization of the charge formed on the tetrahedral intermediate formed during the acylation and de-acylation steps through ion-dipole interactions. This is a hypothesis put forward by the Keillor group, based on modelling experiments as the role of the tyrosine has not been clearly determined.⁶ Khosla and coworkers have also postulated the importance of conserved tryptophan residues in catalysis. They suggest the tryptophan residues form a structure akin to a drawbridge, which can release the transamidated product, once formed.⁷ A detailed mechanism, implicating certain residues, has been proposed by the Keillor group (Figure 1.3).

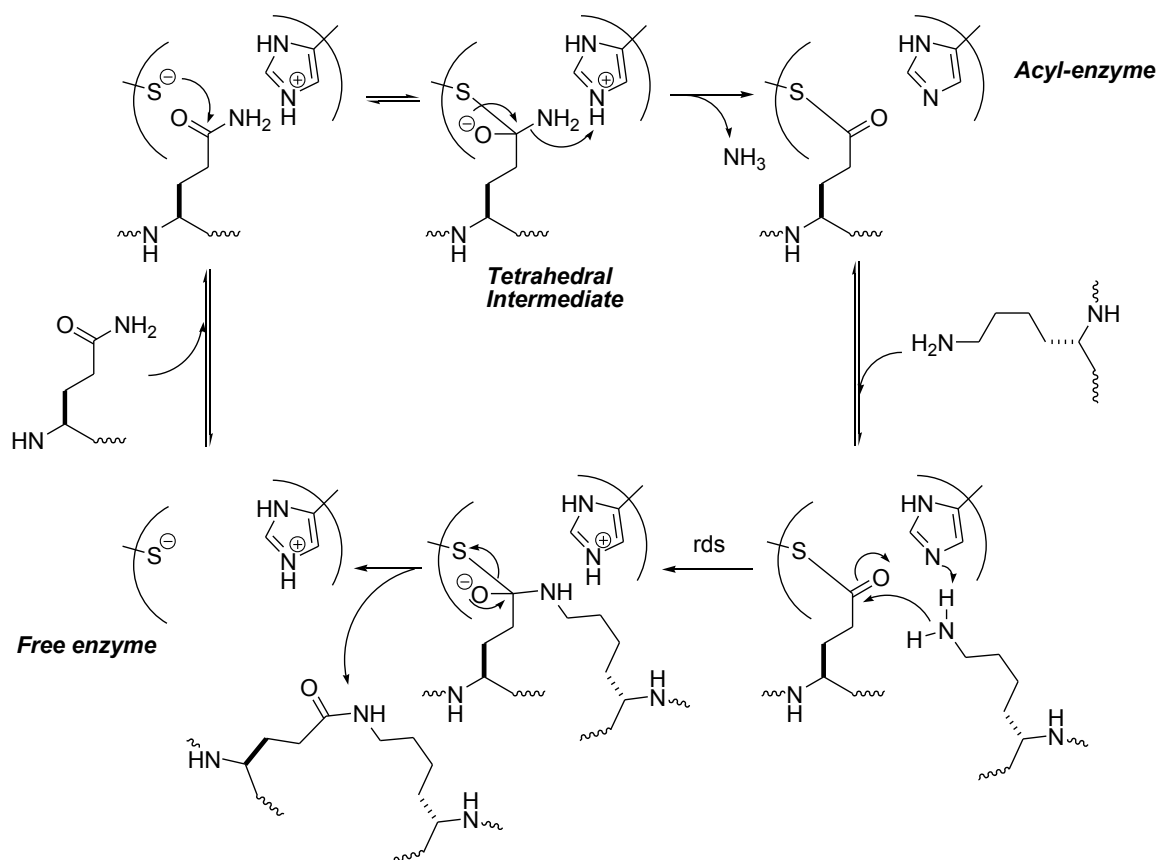


Figure 1.3: Detailed mechanism of transglutaminases.

Despite their similar mechanism, the different transglutaminases have distinct physiological roles and are classified by number. There exist 8 different types of transglutaminases though not all are well characterized. Of TG5 and TG7, little information is known and so they will not be elaborated upon.

1.2.1 Keratinocyte transglutaminase (Type 1)

Keratinocyte TGase is responsible for the cornification of cellular envelopes in hair and nail tissue and is primarily membrane bound.⁸ This monomeric protein contains esterified or thioesterified palmitate and myristate that are responsible for membrane anchorage.⁸

1.2.2 Tissue transglutaminase (Type 2)

Tissue transglutaminase (TG2), also known as cytosolic TGase, is different from other TGases (apart from TG6) in that this protein also binds guanosine-5'-triphosphate (GTP) in addition to acting as a transglutaminase enzyme.⁹ TG2 is expressed in all tissue, notably endothelial cells and the smooth muscle cells in arteries, veins and capillaries.¹⁰ TG2 will be discussed in further detail, as one of the target enzymes of the work presented.

1.2.3 Epidermal transglutaminase (Type 3)

Epidermal TGase is responsible for the cornification of cellular envelopes in the epidermis.⁸ This monomeric protein requires activation by treatment with certain proteases or chaotropic agents and is a characteristic it shares with plasma TGase.⁸

1.2.4 Plasma transglutaminase (FXIII in humans)

Human plasma Factor XIII is one of the best-characterized transglutaminases and is a tetramer composed of two A-chains and two B-chains that are not covalently associated.¹¹ It exists as a proenzyme and requires activation by thrombin before dissociating into active FXIIIa, which has an active role in the stability of blood clots to plasmin degradation. FXIII will be discussed in further detail, as the second target enzyme of this dissertation.

1.2.5 Prostate transglutaminase (Type 4)

Prostate TGase is a dimeric protein implicated in the formation of the postejaculatory vaginal plug in rodents and is also secreted in humans.¹²

1.2.6 Transglutaminase 6 (Type 6)

TG6 has been recently characterized and much like TG2, is allosterically regulated by Ca^{2+} and GTP.¹³ It is abundantly expressed in the central nervous system and is associated with neurogenesis. TG6 is proposed to play a critical role in cortical and cerebellar neurons.¹³

Our topics of study will be focused on TG2 and FXIII. A more exhaustive discussion will be presented for these two enzymes to better understand our interest. Their structure and physiological roles will be described, along with the disease states associated with these enzymes.

1.3 Tissue transglutaminase

TG2 is a monomeric protein of 685-691 amino acids, with a molecular mass of 76 to 85 kDa. The human gene is composed of 13 exons and 12 introns and is found on chromosome 20 (q12).^{14,15} Despite the presence of 17 cysteine residues and 6 potential glycosylation sites, the enzyme is not glycosylated and no disulfide bonds are formed.¹¹

1.3.1 Effect of GDP and calcium

TG2 was originally reported as a GTP-binding protein named the Gh protein,¹⁶ with a molecular weight of 70-80 kDa and implicated in the activation of phospholipase C. It was demonstrated that the Gh protein and TG2 were, in fact, one and the same protein.¹⁷ The ability to bind and hydrolyze GTP with an affinity and rate similar to traditional G proteins distinguishes TG2 from other transglutaminases and suggests that this enzyme participates in signalling pathways.¹⁷⁻¹⁹ Kinetic studies demonstrated that GTP acted as a reversible, noncompetitive inhibitor whose effects could be overcome by high concentrations of Ca^{2+} . It was determined that both GTP and Ca^{2+} function in concert to regulate the transglutaminase function of TG2.¹⁸

From the crystal structure of TG2 complexed with GDP (Figure 1.4 [A]), it was shown that though the transamidation site was similar to other TGases, the guanine nucleotide binding site of TG2 was different from other G proteins. Analysis of the crystal structure suggested negative regulation of transamidation activity by the bound nucleotide. This occurs through blocking of the transamidation active site by loops of the first β -barrel domain.²⁰ Positive regulation of the transamidating activity is achieved through a conformational change triggered allosterically by calcium binding at multiple sites in TG2. Once calcium is bound, the active site is exposed for interaction with glutamyl-containing substrates.^{21,22}

1.3.2 Structure of TG2

Human TG2 consists of 4 domains; an N-terminal β -sandwich, the catalytic α/β -domain containing the catalytic triad and two C-terminal β -barrels. An earlier structure showed TGase bound with GDP²⁰ in a “closed” conformation with the active site inaccessible (Figure 1.4 [A]). More recently, Khosla and coworkers resolved a structure of TG2 in an alternate, extended conformational form termed the “open” form. In this conformation, TG2 is bound with the irreversible inhibitor Ac-P(DON)LPF-NH₂ (Figure 1.4 [B]).⁷ DON is the electrophilic amino acid 6-diazo-5-oxo-L-leucine that mimics an inflammatory gluten peptide. In this structure, the active site is exposed, revealing catalysis takes place in a tunnel, bridged by two tryptophan residues that separate the acyl donor from the acyl-acceptor substrate and stabilize the tetrahedral reaction intermediates.⁷ The two β -barrels have shifted and appear to act as a lid shielding the active site.

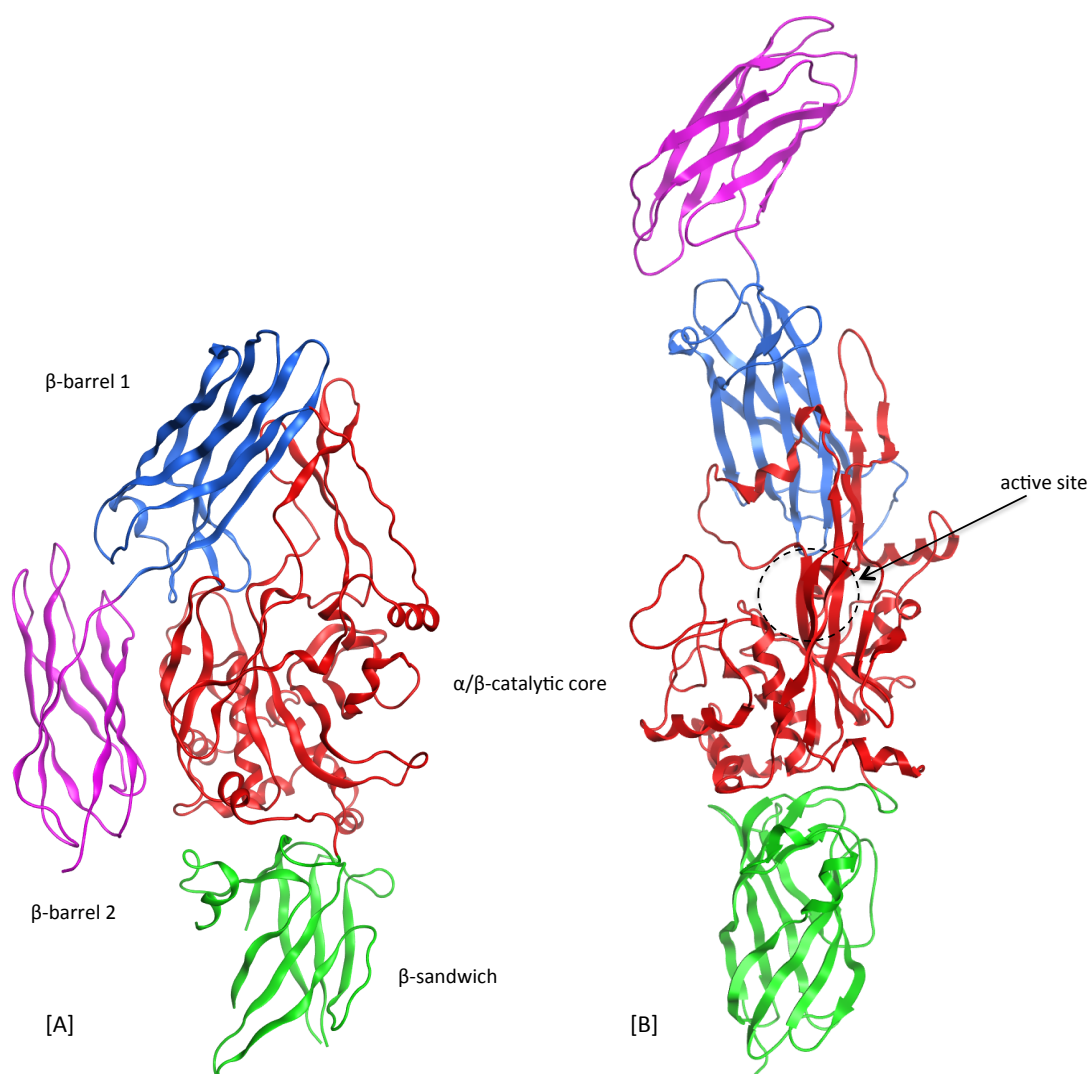


Figure 1.4: [A] Crystal structure of TG2 in a closed conformation (PDB: 1KV3).²⁰ [B] TG2 in an open conformation (PDB: 2Q3Z).⁷ Green = β -sandwich, red = α/β -catalytic core, blue = β -barrel 1, purple = β -barrel 2

1.3.3 Physiological Roles of TG2

TG2 is largely a cytosolic protein where the majority of the protein, approximately 70-80%, is present in the cytoplasm.²⁴ Inside the cell, where concentrations of Ca^{2+} are normally low, the transamidating activity of TG2 remains

dormant and the protein functions as a GTP-binding protein.²⁵ However, TG2 was found to localize on the surface of various cell types as well as in the extracellular matrix,^{26,27} where it displays significant transamidation activity due to the high extracellular Ca^{2+} concentration.²⁵

TG2 has been implicated in a variety of physiological processes including cellular differentiation, cell adhesion, matrix assembly and apoptosis. It was shown that overexpression of TG2 in the human neuroblastoma cell line SK-N-BE(2) exhibited a large increase in the cell death rate, indicating a direct effect of TG2 in phenotypic maturation towards apoptosis.²⁸ Additionally, TG2 has been associated with the formation of protein crosslinks in the apoptotic envelope that prevents leakage of intracellular material, reducing inflammation and autoimmunity.²⁹

It was suggested that the main role of TG2 *in vivo*, in regards to apoptosis, is to ensure that once apoptosis is initiated, it is completed without causing inflammation or apparent tissue injury.³⁰ This is carried out through aiding apoptosis in certain apoptotic cells,^{28,31,32} or promoting the activation of $\text{TGF}\beta$ released by macrophages that promote the death of various cells.²⁵

TG2 was first implicated in cell adhesion and matrix stabilization through the observation that TG2 overexpression on spreading of fibroblast cells led to an increased resistance to their detachment with trypsin.³³ It was later shown that TG2 serves as an adhesion receptor for fibronectin on the cell surface, as well as other extracellular proteins found at the cell surface and in the surrounding extracellular matrix.³⁴ Examples of such proteins would include collagen, laminin and vitronectin, among others.³⁵⁻³⁷

TG2 has been shown to stimulate cell differentiation of various cell types, including neurons, astrocytes and fibroblasts. Through inhibition of TG2 cross-linking activity with monodansylcadaverine, it was found that the decreased amount of protein cross-links affected the differentiation process of NB4 cells to neutrophil granulocytes.³⁸ TG2 activity was also assayed through serial passages in rat glial cells. The decrease of TG2 activity from the 3rd to the 5th passage resulted in cell death and a loss of growth control.³⁹ Finally, Johnson and coworkers demonstrated that TG2 is necessary for neuronal differentiation of human neuroblastoma SH-SY5Y cells.⁴⁰ SH-SY5Y cells overexpressing wild-type tissue transglutaminase spontaneously differentiated into a neuronal phenotype whereas cells that overexpressed inactive TG2 continued to proliferate and exhibit flat polygenic morphology.⁴⁰

1.3.4 Pathological associations of TG2

Since TG2 is important in the stabilization of the extracellular matrix, perturbation of its function has been associated with conditions such as fibrosis and atherosclerosis.¹¹ It has been suggested that TG2 could contribute to the formation of atherosclerotic plaque by catalyzing the incorporation of lipoprotein(a) into plaque through a mechanism which has yet to be determined.⁴¹ TG2 has also been implicated in a variety of diseases, including Celiac disease, as well as neurodegenerative diseases such as Alzheimer's and Huntington's.

1.3.4.1 TG2 in Celiac disease

Celiac disease is counted among the gluten sensitivity diseases; it is an autoimmune disorder triggered by dietary exposure to gliadin, the proteinaceous component of gluten present in wheat, barley and rye. Symptoms will resolve with a gluten-free diet and reoccur upon the resumption of gluten ingestion. Gluten proteins are largely resistant to gastrointestinal proteases, whereby the immunotoxic peptides will accumulate and initiate a deleterious, adaptive and innate immune response in genetically susceptible individuals.⁴² Schuppan and coworkers discovered that TG2 was an autoimmune antigen of celiac disease and gliadin was its preferred substrate.⁴³ Based on *in vivo* experiments, TG2 was thought to be responsible for generating neoepitopes of gliadin through deamidation of its glutamine residues.⁴⁴ The effect of deamidation was shown through a series of mass spectrometry experiments, whereby it led to an increase in antigenicity by stimulating the proliferation of gliadin-specific T-cells.⁴⁵ Currently, a diagnostic method of Celiac's disease involves serological testing for anti-TG2 autoantibodies.⁴⁶

1.3.4.2 TG2 in Alzheimer's disease

Alzheimer's involves progressive neurodegeneration and a decline of cognitive function, associated with the accumulation of fibrillar β -amyloid in extracellular plaque which is also a substrate for TG2.⁴⁷ The hypothesis that amyloid contributes to Alzheimer's was based on observations that the inheritance of mutated amyloid was associated with early onset Alzheimer's along with the report that aggregated β -amyloid was neurotoxic *in vitro*.^{48,49} Through comparisons of normal and Alzheimer's

diseased brains, it was found that TG2 activity was increased in the diseased brain.⁵⁰ This finding suggests that TG2 was a contributing factor in protein aggregate formation.⁵⁰ The crosslinking of β -amyloid protein by TG2 results in the formation of high molecular weight oligomers that resemble proteins isolated from the brains of Alzheimer's patients, indicating the enzyme may be involved, in part, in the formation of these plaque deposits.⁵¹ Lastly, the tau protein, which is a major component in the paired helical filaments that make up neurofibrillary tangles associated with Alzheimer's, is an excellent *in vitro* substrate for TG2.⁵²

1.3.4.3 TG2 in Huntington's disease

Huntington's disease is an autosomal dominant neurodegenerative disorder that is characterized by extensive neuronal cell apoptosis, leading to progressive motor dysfunction and psychiatric disturbances with gradual dementia. It is caused by the expansion of 40 or more cytosine-adenosine-guanine (CAG) repeats in the gene encoding the cytosolic protein, huntingtin (htt).⁵³ Increased length of the CAG repeats correlates with an earlier age at which disease onset occurs.⁵³ Evidence for the involvement of TGase in Huntington's is based on the demonstration of elevated levels of TG2 activity in affected regions of diseased brains.⁵⁴ It was shown that htt was a substrate for TG2 and that htt was found in the aggregates characteristic of Huntington's.⁵⁴ Through comparisons of normal and diseased Huntington's brains, it was shown that TGase-mediated crosslinking of htt was elevated in the diseased brain with significant increases in the cell nuclei.⁵⁵ Additionally, it was found that the ability of polyglutamine peptide sequences to act as TG2 substrates improved with increasing

polyglutamine length and htt containing a pathological length of polyglutamine residues was a better substrate for TG2-mediated polymerization than htt with a shorter, non-pathological glutamine stretch.^{56,57}

Thus, it is the transamidation activity of TG2 that leads to its involvement in diseased states and warrants investigation into the catalytic transglutaminase abilities of this enzyme.

1.3.5 Transglutaminase 2 substrates

Natural substrates of TG2 include fibrin,⁵⁸ actin,⁵⁹ and α_{s1} -casein.⁶⁰ As TG2 is a ubiquitously expressed enzyme, found almost everywhere in tissue, this protein can accept a variety of substrates as the acyl donor. One example is the breakpoint cluster region protein; a multi-domain protein that contains a GTPase activating protein domain for Rac.⁶¹ Activation of TG2 by calcium caused the formation of covalently-linked Bcr, and this protein thus represents a target of TGase activity during extreme stress.⁶¹

Among the non-natural donor substrates of TG2, the most frequently used is Cbz-glutaminy-glycine.⁶² Structural variations of this compound have identified certain characteristics that influence its affinity for TG2.⁶³ It was demonstrated that having a free amine on the glutamine was deleterious to affinity while it was preferable to have an amide bond as opposed to a carboxylic acid on the glutamine. It has previously been observed in our group that protecting the amine with a non-aromatic protecting group (such as Boc) decreased the affinity⁶⁴ and though TG2 will accept an increase in length of the carboxamide side-chain, asparagine is not recognized.⁶⁵

Other peptidic substrates include EAQQIVM, derived from the N-terminal sequence of fibronectin.⁶⁶ It was found that it is primarily the first glutamine residue that is modified by TG2⁶⁶ and this sequence has also been found in other natural TG2 substrates such as fibronectin,⁶⁷ the fibrinogen γ -chain,⁶⁸ plasminogen,⁶⁹ β A3 crystalline chains⁷⁰ and osteonectine.⁷¹

Through screening of known immunogenic peptide substrates, Khosla and coworkers have identified a 33-mer peptide from gliadin peptides consisting of repeating units of PQPQLPY.⁷² The specificity of TG2 towards PQPQLY was found to be appreciably higher than the specificity for its natural substrate.⁷² Based on this backbone, gluten peptide analogs as inhibitors were designed for TG2 with the sequence Ac-PQP(DON)LPF-OH, with a K_i value of 5 μ M.⁷³

Different peptides taken from the TGase substrate domains of pro-elafin, the most abundant component in cornified cell envelopes, were investigated for their competitive inhibitory effect.⁷⁴ A secondary study identified truncated sequences and found that of the various recombinant peptides tested, KVLDGQDP was found to have the best inhibitor effect.⁷⁵

In the Keillor group, Ac-PNPQLPF-OH was envisioned as a TG2 substrate as it conformed to patterns predicted by Fesus and Hitomi.^{76,77} Both scientists had suggested that the asparagine and proline residues around the reactive glutamine were important in TG2 recognition.^{76,77} With a K_m of 6 μ M (Pardin and Keillor, unpublished results), this peptide was validated for *in vivo* reactivity through its use as a donor-substrate sequence in a TG2-mediated cell surface labelling reaction.⁷⁸ The sequence was also modified to include a fluorophore and was used as fluorescent probe for TG2

by replacing the glutamine with a modified lysine.⁷⁹ In place of the ϵ -amino group, an electrophilic acryl amide was incorporated to react irreversibly with the active site cysteine. The sequence was truncated to allow the addition of a spacer attached to a rhodamine molecule and the probe was used to monitor the role of TG2 in arterial rigidification.⁷⁹

In regards to the acceptor substrate, TG2 has been found to be highly promiscuous, generally accepting a wide range of primary alkylamines. This has been demonstrated with multiple amine incorporation assays into proteins using diamines such as putrescine, spermidine and cadaverine as the acyl acceptor. In order to monitor the labelling reaction, the amine is usually bound to a fluorophore, such as in the case of monodansyl cadaverine.^{70,80} Recently in the Keillor group, propargyl amine was shown to be an excellent amine substrate for TG2 and opens the door to site-specific protein modifications.⁸¹

1.4 Factor XIII

FXIII is also known as the blood plasma transglutaminase, owing to its widespread abundance in blood. Additionally, its primary function is involved in blood coagulation. It is a tetrameric molecule composed of two 83 kDa (731 amino acids) A-subunits and two 80 kDa (641 amino acids) B-subunits.^{58,82} The gene coding for the A subunit has been localized to chromosome 6p24-25 and consists of 15 exons separated by 14 introns.⁸³ The gene coding for FXIIIB has been localized to chromosome 1q31-32 with 12 exons separated by 11 introns.⁸⁴ The entire protein is held together noncovalently in a heterologous tetramer of 326 kDa.⁸⁵

Two forms of FXIII exist; blood plasma FXIII, a tetrameric species found in plasma, and cellular FXIII. Secondly, a dimer of the A-subunits in physiological form is present in the cytoplasm of certain cells, particularly platelets, monocytes and macrophages.⁸⁶

The A subunit contains the active site of the enzyme and has 6 potential asparagine-linked glycosylation sites though none have carbohydrate attachments.^{85,87} However, carbohydrates contribute 8.5% of the total molecular weight to the B-subunit.⁸⁸ The B-subunit is a modular protein made up of 10 repeated sushi/glycoprotein-1 domains.⁵⁸ Each sushi domain contains 2 disulfide bridges that sustain its tertiary structure, amounting to a total of 40 cysteine residues, of which 20% are disulfide bridges in the mature B-subunit. The main function of the B-subunit is the stabilization and transport of the hydrophobic A-subunit in the aqueous environment of human plasma.^{89,90}

1.4.1 Activation of FXIII

One characteristic that sets FXIII apart from the rest of the mammalian transglutaminase family (excluding TG3) is the fact that it exists as a proenzyme. The A-subunit of FXIII contains an activation peptide of 37 amino acids that limits access of the substrate to the active site cysteine and is cleaved by thrombin using fibrin as a cofactor in biological systems (Figure 1.5 [A]).^{91,92} However, cleavage of the arginine 37-glycine 38 peptide bond may occur by other serine proteases, including endogenous platelet acid protease and calpain, which were both reported to activate FXIII.^{93,94} Under physiological conditions, a lag phase is observed between thrombin cleavage and

expression of the active site of plasma factor FXIII which represents the time it takes for the B-subunit to dissociate from FXIIIA.⁹⁵ This dissociation of the B-unit is critical to expose the active site cysteine. X-ray crystallographic studies of thrombin-treated FXIIIA have suggested that the activation peptide does not dissociate upon cleavage, but remains associated with the protein.⁹⁶

Similar to TG2, calcium is also necessary for FXIIIa activity. Alkylation experiments of the active site cysteine residue showed that the cysteine remained buried after thrombin cleavage of the propeptide and was only alkylated by iodoacetamide after exposure of the protein to calcium.⁹¹ In fact at high (100 mM) calcium concentrations, it has been demonstrated that the A₂B₂ tetramer of FXIII will dissociate, without prior proteolytic cleavage of the activation peptide, and the non-truncated FXIIIA₂ will transform into the active form (Figure 1.5 [B]).⁹⁷

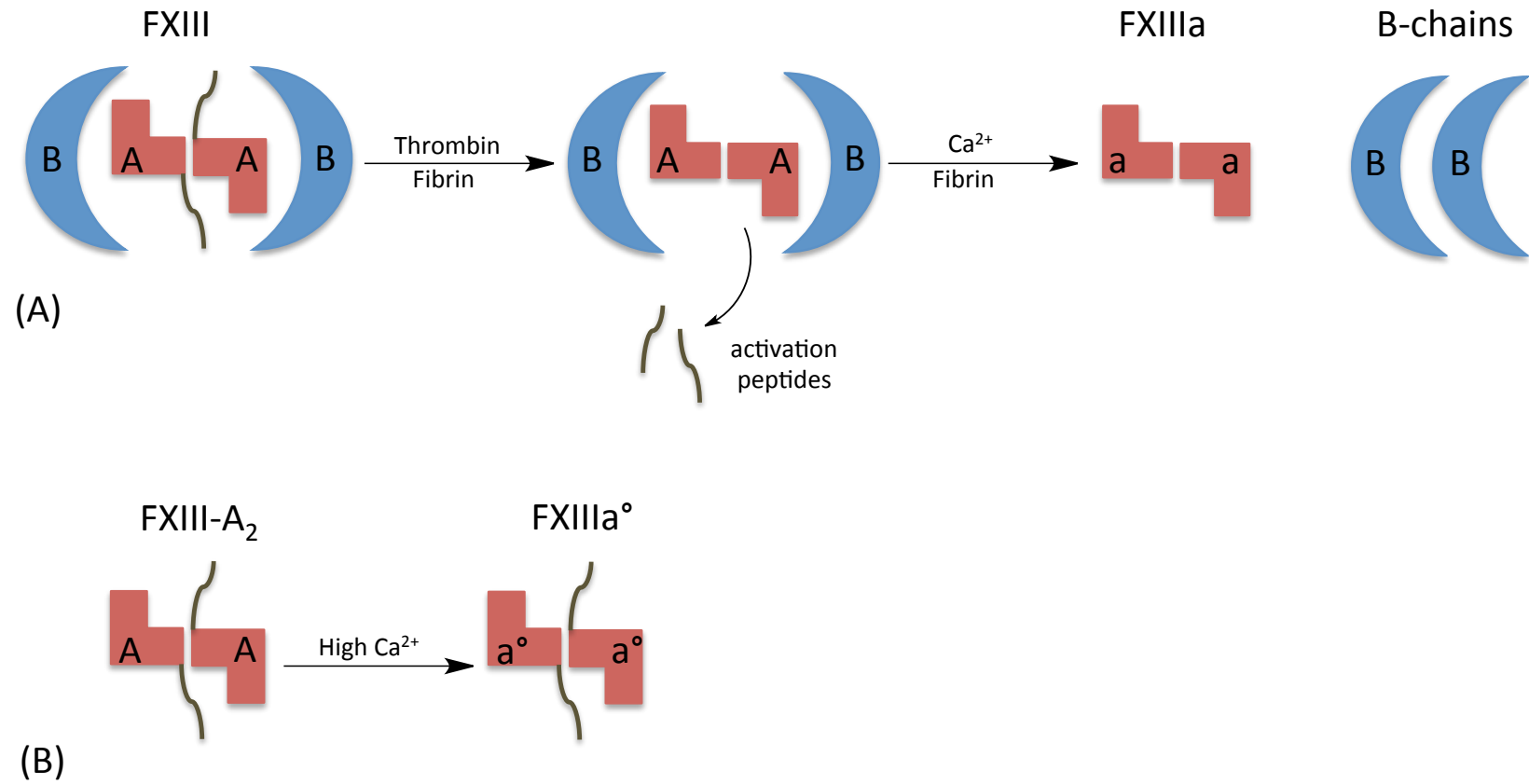


Figure 1.5: [A] Activation of FXIIIa in the presence of thrombin, fibrin and Ca²⁺. [B] Activation of FXIII without cleavage of activation peptide at high Ca²⁺ concentration.⁹⁸

1.4.2 Structure of FXIIIA

There has been much debate about the structure of FXIII in the literature owing to the presence of the propeptide. Like all TGases, FXIIIA is divided into four domains, including the β -sandwich, the α/β -catalytic core and 2 β -barrels.⁹⁶ Initial structural studies of FXIIIA suggested that significant changes did not occur between the non-activated and thrombin and Ca^{2+} activated FXIIIA₂ (Figure 1.6 [A]).⁹⁶ However, chemical modifications of the amino acid side chains suggested that a conformational change did occur. Through alkylation of cysteine and lysine residues followed by mass spectrometry analysis, it was found that additional cysteines were alkylated when the enzyme was activated while a decrease in alkylation of lysine residues was observed.^{99,100} More recently, Muzbek and coworkers postulated that FXIIIA undergoes a conformational change similar to TG2 and confirmed their hypothesis through modelling experiments with homology models of FXIIIA and TG2.¹⁰¹

This supposition was recently reinforced in late 2013 by Glebe and coworkers who reported the first high-resolution structure of FXIIIA in its active state bound to an inhibitor (Figure 1.6 [B]).¹⁰² The enzyme was activated with high concentrations of Ca^{2+} and crystallized in the presence of a Michael acceptor-based inhibitor to mimic the glutamine residue. The combination of the inhibitor and calcium induce a conformational transition in which the two β -barrels swing aside to expose the catalytic center. However, structural rearrangements upon activation were limited to the catalytic core while the β -sandwich and the two β -barrels domains retained their overall fold. From this crystal structure, three calcium binding sites were observed which the authors proposed were necessary to drive the enzyme into the active

state.¹⁰² The close proximity of the catalytic triad to regions of calcium binding had prompted Yee to suggest that calcium ions are part of the regulation of the conformational changes that accelerate catalysis.⁹⁶ This hypothesis was confirmed by Stieler who observed that the catalytic mechanism uses the calcium ions as cofactors to hold the active site in the proper conformation to trigger the formation of the thioester intermediate with glutamine.¹⁰²

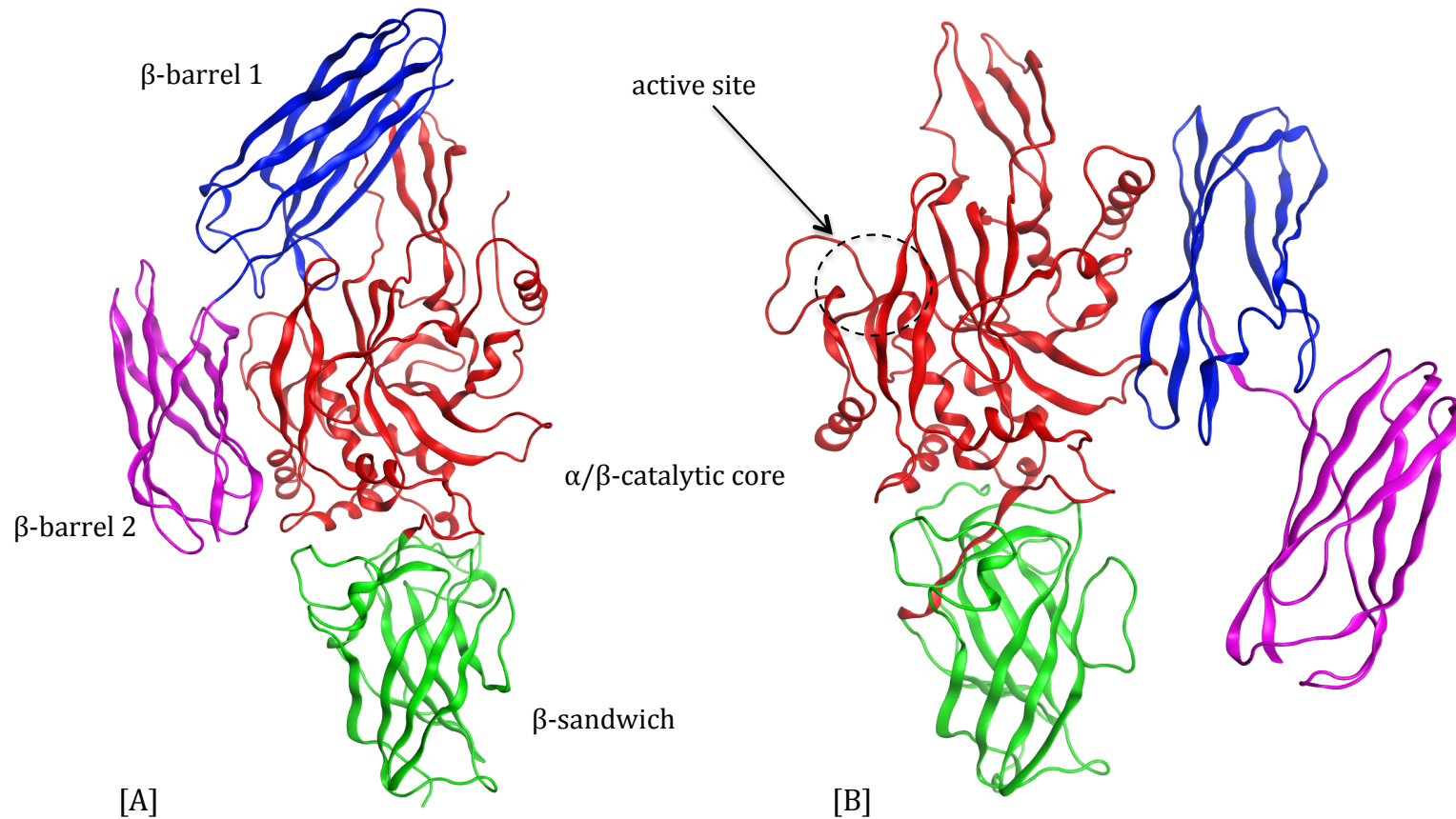


Figure 1.6: [A] Crystal structure of FXIII in an inactive state (PDB: 1GGU).¹⁰³ [B] FXIIIa^o in an active conformation (PDB: 4KTY).¹⁰² Green = β -sandwich, red = α/β -catalytic core, blue = β -barrel 1, purple = β -barrel 2

1.4.3 Physiological roles of FXIII

The primary physiological role of FXIII is maintaining homeostasis by stabilizing the fibrin blood clot, protecting it from fibrinolytic degradation.¹⁰⁴ By catalyzing the formation of ϵ -(γ -glutamyl)lysyl peptide crosslinks in fibrin, the chemical stability and mechanical strength of the blood clot is increased.⁹² FXIIIa is the last enzyme in the blood coagulation cascade and as such, it catalyzes the cross-linking of α_2 -antiplasmin; a highly potent inactivator of plasmin, to the α -chain of fibrin and fibrinogen.^{105,106} In addition to its hemostatic role, FXIII has also been associated with wound healing, angiogenesis, as well as matrix stabilization.

FXIII was implicated to be essential in pregnancy through experiments with FXIIIa-knockout mice. Once pregnant, many mice died and their death was attributed to massive placental hemorrhage and subsequent necrosis that developed in the knockout mice.¹⁰⁷ This observation was similar to reports of spontaneous miscarriage in pregnant humans with FXIII deficiency. These findings led to the implication that maternal FXIII plays a critical role in uterine hemostasis and placenta attachment.¹⁰⁸

FXIII also participates in wound healing, which was inferred from the delay in wound repair in some patients with inherited FXIII deficiency.¹⁰⁹ This participation by FXIII was confirmed through a study on the effect of FXIII deficiency on wound healing in FXIII-deficient mice. FXIII-deficient mice exhibited impaired wound healing demonstrated by a decrease in the percentage of wound closure; effects that were reversed with the addition of FXIII.¹¹⁰ These observations provide evidence for the association of FXIII with wound repair.

As the process of angiogenesis is closely related with hemostasis, it is logical that FXIII would be implicated in this process. Angiogenesis is the process involving remodeling and sprouting of new capillaries from pre-existing blood vessels. Once clot stabilization is achieved, angiogenesis is modulated by proteins and peptide fragments generated from the coagulation and fibronectin system.¹¹¹ FXIII mediates the incorporation of fibronectin and other extracellular matrix proteins into the fibrin clot to prevent further bleeding. Once the clot is anchored to the vessel wall, FXIII facilitates new vessel formation by direct stimulation of endothelial cell proliferation, migration and survival through upregulation of the transcription factors, Egr-1 and c-Jun.¹¹²

Additionally and similar to TG2, FXIII has been implicated in the stabilization and mineralization of the extracellular matrix in bone and cartilage.¹¹³ This observation is supported by the fact that adhesive extracellular glycoproteins such as collagen, vitronectin and the von Willebrand factor are all *in vitro* substrates for FXIII.^{36,114,115}

1.4.4 FXIII in disease states

The earliest reported pathological association of FXIII was made in 1960, where it was noted that the plasma clot of a patient with severe bleeding diathesis was missing FXIII.¹¹⁶ This led to the characterization of FXIII deficiency as a rare autosomal recessive disorder. Later experiments showed that the lack of FXIII results in serious bleeding diathesis, defective wound healing and a high risk of miscarriage in deficient females.⁹⁸

FXIII has also been implicated in atherothrombotic disorders such as myocardial infarction and venous thrombosis. Myocardial infarction arises from the development

of cross-linked, fibrinolysis-resistant, platelet-rich fibrin clots in the arteries while a thrombotic occlusion occurs in the venous system that is low on platelets and is known as venous thrombosis.¹¹⁷

The first evidence of FXIII in myocardial infarction was through analysis of plasma from patients with the disease. It was found that FXIIIa cross-linked fibrin polymers were increased in the plasma and indicated the presence of increased activity of FXIIIa in patients with myocardial infarction.¹¹⁸ Contrastingly, it was also found that FXIII played a role in cardiac remodelling after myocardial infarction. FXIII-deficient mice died within 5 days after myocardial infarction from a ventricular rupture, attributed to the observed cardiac rupture.¹¹⁹ The FXIII-deficient mice exhibited an attenuated inflammatory response, enhanced degradation of the extracellular matrix, impaired collagen synthesis and aggravated cardiac remodelling. These observations led the authors to conclude that FXIII plays a key role in wound healing after myocardial infarction.

The relationship between FXIII and atherothrombotic diseases such as myocardial infarction and venous thrombosis has also been investigated through comparisons of FXIII and a polymorphism of FXIII, namely FXIIIVal34Leu. FXIIIVal34Leu is fairly common with an allele frequency of 0.25-0.3 in the white population.^{93,120,121} The activation of this variant by thrombin occurs more rapidly than that of normal FXIII, and is independent of the interaction between the A and B subunits as both plasma and platelet Leu34 FXIII are activated more rapidly than their Val counterparts.¹²²⁻¹²⁴ Several studies were undertaken that suggested that the Leu allele was protective against myocardial infarction, and this was confirmed by other

groups, who noted that where the prevalence of the Leu allele is the lowest, the highest risk of myocardial infarction was observed.^{93,120,125,126} However, conflicting reports suggested no association between the possession of Leu 34 and a decreased risk of myocardial infarction.¹²³⁻¹²⁶ Ariens suggested that a possible explanation for the discrepancy was the linkage disequilibrium (non-random association of alleles at two or more loci) between Val34Leu and other polymorphisms in the FXIIIA subunit gene.⁸⁴

Similar contradictory results were reported for the FXIIVal34Leu polymorphism in the incidence of venous thrombosis,^{124,125,127-130} indicating that the relationship between FXIIIA and atherosclerosis is neither clear-cut nor simple but requires further investigation and inquiry.

The discovery that cellular FXIII is present in monocytes and macrophages, including tissue macrophages has revealed that FXIII is present in most organs and tissues of the body.⁹⁸ Thus, it is understandable that the ubiquitous nature of this enzyme explains how it could be implicated in both the cause and recovery of atherosclerosis when FXIII has been implicated in both clot formation and wound repair.

The widespread localization of FXIII makes it likely to suppose that this enzyme could be implicated in other physiological processes unrelated to homeostasis or wound repair. Recently, it has been shown that FXIIa activity was increased in obese twins and strongly correlated with measures of adiposity, inflammation and insulin resistance.¹³¹ This finding demonstrates that there is still much to discover about this enzyme.

1.4.5 Substrates of FXIII

FXIII has a wide variety of physiological substrates, including vinculin, osteopontin and plasminogen.¹³²⁻¹³⁴ It has been shown that both zymogen and activated forms of Factor V can form high molecular mass polymers in the presence of FXIIIa, suggesting it contributes both glutamine and lysine residues in the cross-linking reaction, much like fibrin.¹³⁵ Casein was also shown to be a good substrate for FXIII, specifically β -casein was an excellent substrate for FXIIIa while a relatively poor substrate for TG2.¹³⁶

Unnatural substrates for FXIII include peptide substrates, usually derived from protein substrates. One example is the K9-glutamine substrate derived from β -casein with the sequence LGPGQSKVIG.¹³⁷ K9 was commercialized by the Berichrom Corporation, now taken over by Dade Behring in Germany. This sequence was used as the glutamine donor substrate in a photometric assay for FXIIIa activity in blood samples. From the α_2 -antiplasmin inhibitor, the sequence NQEQVSPLTLLK was identified as the site of crosslinking to fibrinogen by FXIIIa.¹³⁸ This peptide was later modified to produce a fluorometric assay for FXIIIa, which is also recognized by TG2, TG3 and TG6 and sold by the German company, Zedira.¹³⁹

In 2006, Hitomi and coworkers identified the first peptidic sequences specific to FXIIIa alone.⁷⁷ Using a phage display peptide library, they identified the primary sequences around the reactive glutamine residue preferred by either TG2 or FXIIIa. The library was screened through avidin affinity purification of phage clones that incorporated a biotin-labeled primary amine in catalytic reactions with TG2 or FXIIIa. The chosen peptides were then expressed as recombinant GST fusion proteins and their

reactivity with TGase was confirmed through incorporation of monodansylcadaverine. The highest affinity peptide for FXIII was identified and consisted of DQMMLPWPAVKL (F11) and is currently available for purchase as a FXIIIa-specific peptide substrate by Zedira.

Similar to TG2, FXIIIa accepts a wide variety of amine donors. This includes lysine-containing proteins such as thrombospondin or myosin.^{140,141} FXIIIa accepts a range of alkyl primary amines as well, including putrescine, monodansylcadaverine and 5-(biotinamido)pentylamine.¹⁴²⁻¹⁴⁴

A study has also been carried out on the acyl acceptor substrate peptides from the FXIII recognition sequence in the γ -fibrin chain. A truncated peptide was identified as GQQHHLGGAKQAGVD and the amino acids around the lysine were exchanged to determine their effect.¹⁴⁵ It was found that FXIII would accept a broad range of substrate peptides, with a proline neighbouring the requisite lysine having the most detrimental effect.¹⁴⁵

1.5 Research objectives

As both TG2 and FXIII have important physiological and pathological roles, their function and activity are a source of interest to the Keillor group. Currently, the mammalian transglutaminase research program involves: 1) mechanistic investigation through capillary electrophoresis of TG2¹⁴⁶ 2) development of potential inhibitors for therapeutic ends or protein labelling for both TG2 and FXIII^{147,148} 3) expression and purification of TG2¹⁴⁹ and FXIII 4) FXIIIa-specific peptides.

1.5.1 Transition state analogue inhibitors for TG2

Various reversible inhibitors have been described for tissue transglutaminase, notably cinnamoyl derivatives from the Keillor group.¹⁵⁰ However, a new class of reversible inhibitors, namely transition state analogues was envisioned as a new generation of high-affinity competitive inhibitors.

Transition state analogue inhibitors are designed from the hypothesis that chemically stable compounds that simulate the geometric and molecular electrostatic characteristics of the transition state will bind to the enzyme tighter than the substrate by a factor approaching the catalytic rate acceleration imposed by the enzyme.

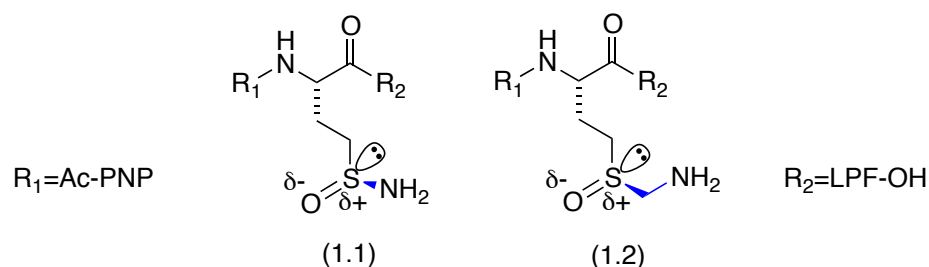


Figure 1.7: Target sulfinamide inhibitors for TG2.

It was hypothesized that the structure of a sulfinamide molecule would resemble the transition state adopted during the acylation step. Thus compounds 1.1 and 1.2 (Figure 1.7) were proposed as analogues of the first and second transition state towards the acylation step of the cross-linking reaction catalyzed by TG2.

The sulfonamide moiety would be incorporated into the Ac-PNPQLPF-OH sequence, in order to take advantage of this peptide's high binding affinity and specificity for TG2. The synthesis towards these target compounds will be discussed in chapter 2.

1.5.2 Fluorescent, cell permeable probe of TG2

As mentioned previously, intracellular TG2 activity consists mainly of GTP binding as the physiological, intracellular concentrations of calcium are low. However, evidence has been presented that TG2 transamidation activity is not limited to the extracellular matrix but also occurs within cells.^{57,151} In order to delineate the activity of intracellular GTP, a cell permeable, TG2-specific, fluorescent probe was synthesized. The synthesis, kinetic characterization and preliminary *in vivo* results of this probe will be discussed in chapter 3.

1.5.3 Studies with FXIIIa

FXIII has recently become a part of the research program of the Keillor group. Through investigation of collagen and fibronectin matrix deposition with one of our inhibitors, it was found that transglutaminase activity was related to osteoblast differentiation, attributed to either TG2 or FXIIIa.¹⁴⁸ However, while current TG2-specific probes exist, there have not been kinetically characterized substrates that are specific to FXIII. Our efforts in identifying a high affinity peptidic sequence for FXIII and applying it towards an activity assay will be discussed in chapter 4. As this goal necessitates the use of large quantities of FXIII, the optimization of expression of a recombinant FXIIIa fusion protein will also be reviewed.

References

- (1) Studier, F. W. *Protein Expression and Purification* **2005**, 41, 207–234.
- (2) Page, M. I.; Jencks, W. P. *Proceedings of the National Academy of Sciences* **1971**, 68, 1678–1683.
- (3) Sly, W. S.; Hu, P. Y. *Annual Review of Biochemistry* **1995**, 64, 375–401.
- (4) E, F. J. *Journal of Biological Chemistry* **1969**, 244, 3707–3713.
- (5) E, F. J.; Mullooly, J. P.; Cole, E. R. *Journal of Biological Chemistry* **1967**, 242, 1838–1844.
- (6) Chica, R. A.; Gagnon, P.; Keillor, J. W.; Pelletier, J. N. *Protein Science* **2004**, 13, 979–991.
- (7) Pinkas, D. M.; Strop, P.; Brunger, A. T.; Khosla, C. *PLoS ONE* **2007**, 5, 2788–2796.
- (8) Greenberg, C. S.; Birckbichler, P. J.; Rice, R. H. *The FASEB Journal* **1991**, 5, 3071–3077.
- (9) Lee, K. N.; Arnold, S. A.; Birckbichler, P. J.; Patterson, M. K., Jr; Fraij, B. M.; Takeuchi, Y.; Carter, H. A. *Biochimica et Biophysica Acta* **1993**, 1202, 1–6.
- (10) Thomazy, V.; Fesus, L. *Cell and Tissue Research* **1984**, 255, 215–224.
- (11) Aeschlimann, D.; Paulsson, M. *Thrombosis and Haemostasis* **1994**, 71, 402–415.
- (12) Ho, K.-C.; Quarmby, V. E.; French, F. S.; Wilson, E. M. *Journal of Biological Chemistry* **1992**, 267, 12660–12667.
- (13) Thomas, H.; Beck, K.; Adamczyk, M.; Aeschlimann, P.; Langley, M.; Oita, R. C.; Thiebach, L.; Hils, M.; Aeschlimann, D. *Amino Acids* **2013**, 44, 161–177.

- (14) Gentile, V.; Davies, P. J. A.; Baldini, A. *Genomics* **1994**, *20*, 295–297.
- (15) Fraij, B. M.; Gonzales, R. A. *Biochimica et Biophysica Acta* **1997**, *1345*, 65–71.
- (16) Im, M. J.; Graham, R. M. *Journal of Biological Chemistry* **1990**, *265*, 18944–18951.
- (17) Nakoaka, H.; Perez, D. M.; Jin Baek, K.; Das, T.; Husain, A.; Misono, K.; Im, M.-J.; Graham, R. M. *Science* **1994**, *264*, 1593–1596.
- (18) Achyuthan, K. E.; Greenberg, C. S. *Journal of Biological Chemistry* **1987**, *262*, 1901–1906.
- (19) Mian, S.; Elaoui, El, S.; Lawry, J.; Gentile, V.; Davies, P. J. A.; Griffin, M. *FEBS Letters* **1995**, *370*, 27–31.
- (20) Liu, S.; Cerione, R. A.; Clardy, J. *Proceedings of the National Academy of Sciences* **2002**, *99*, 2743–2747.
- (21) Király, R.; Csősz, É.; Kurtán, T.; Antus, S.; Szigeti, K.; Simon-Vecsei, Z.; Korponay-Szabó, I. R.; Keresztessy, Z.; Fesus, L. *FEBS Journal* **2009**, *276*, 7083–7096.
- (22) Casadio, R.; Polverini, E.; Mariani, P.; Spinozzi, F.; Carsughi, F.; Fontana, A.; Polverino de Laureto, P.; Matteucci, G.; Bergamini, C. M. *European Journal of Biochemistry* **1999**, *262*, 672–679.
- (23) Jeon, J.-H.; Cho, S.-Y.; Kim, C.-W.; Shin, D.-M.; Kweon, J.-C.; Choi, K.-H.; Park, S.-C.; Kim, I.-G. *Biochemical and Biophysical Research Communications* **2002**, *294*, 818–822.
- (24) Griffin, M.; Casadio, R.; Bergamini, C. M. *Biochemical Journal* **2002**, *368*, 377–396.

- (25) Fesus, L.; Piacentini, M. *Trends in Biochemical Sciences* **2002**, 27, 534–349.
- (26) Martinez, J.; Chalupowicz, D. G.; Roush, R. K.; Sheth, A.; Barsigian, C. *Biochemistry* **1994**, 33, 2538–2545.
- (27) Upchurch, H. F.; Conway, E.; Patterson, M. K., Jr; Maxwell, M. D. *Journal of Cellular Physiology* **1991**, 149, 375–382.
- (28) Melino, G.; Annicchiarico-Petruzzelli, M.; Piredda, L.; Candi, E.; Gentile, V.; Davies, P. J. A.; Piacentini, M. *Molecular and Cellular Biology* **1994**, 14, 6584–6596.
- (29) Melino, G.; Piacentini, M. *FEBS Letters* **1998**, 430, 59–63.
- (30) Fesus, L.; Szondy, Z. *FEBS Letters* **2005**, 579, 3297–3302.
- (31) Oliviero, S.; Amendola, A.; Di Sano, F.; Farrace, M. G.; Fesus, L.; Nemes, Z.; Piredda, L.; Spineda, A.; Piacentini, M. *Molecular and Cellular Biology* **1997**, 17, 6040–6048.
- (32) Rodolfo, C.; Mormone, E.; Matarrese, P.; Ciccocanti, F.; Farrace, M. G.; Garofano, E.; Piredda, L.; Flmia, G. M.; Malorni, W.; Piacentini, M. *Journal of Biological Chemistry* **2004**, 279, 54783–54792.
- (33) Gentile, V.; Thomazy, V.; Piacentini, M.; Fesus, L.; Davies, P. J. A. *Journal of Cell Biology* **1993**, 119, 463–474.
- (34) Gaudry, C. A.; Verderio, E.; Aeschlimann, D.; Cox, A.; Smith, C.; Griffin, M. *Journal of Biological Chemistry* **1999**, 274, 30707–30714.
- (35) Kleman, J.-P.; Aeschlimann, D.; Paulsson, M.; van der Rest, M. *Biochemistry* **1995**, 34, 13768–13775.
- (36) Mosher, D. F.; Schad, P. E. *Journal of Clinical Investigation* **1979**, 64, 781–787.

- (37) Aeschlimann, D.; Thomazy, V. *Connective Tissue Research* **2000**, *41*, 1–27.
- (38) Balajthy, Z.; Csomos, K.; Vamosi, G.; Szanto, A.; Lanotte, M.; Fesus, L. *Blood* **2006**, *108*, 2045–2054.
- (39) Campisi, A.; Renis, M.; Russo, A.; Sorrenti, V.; Di Giacomo, C.; Castorina, C.; Vanella, A. *Neurochemical Research* **1992**, *17*, 1201–1205.
- (40) Tulcholski, J.; Lesort, M.; Johnson, G. V. W. *Neuroscience* **2001**, *102*, 481–491.
- (41) Bowness, M. J.; Tarr, A. H.; Wiebe, R. I. *Thrombosis Research* **1989**, *54*, 357–367.
- (42) Iismaa, S. E.; Mearns, B. M.; Lorand, L.; Graham, R. M. *Physiological Reviews* **2009**, *89*, 991–1023.
- (43) Dieterich, W.; Ehnis, T.; Bauer, M.; Donner, P.; Volta, U.; Riecken, E. O.; Schuppan, D. *Nature Medicine* **1997**, *3*, 797–801.
- (44) Johansen, B. H.; Gjertsen, H. A.; Vartdal, F.; Buus, S.; Thorsby, E.; Lundin, K. E. A.; Sollid, L. M. *Clinical Immunology and Immunopathology* **1996**, *79*, 288–293.
- (45) Molberg, O.; McAdam, S. N.; Korner, R.; Quarsten, H.; Kristiansen, C.; Madsen, L.; Fugger, L.; Scott, H.; Noren, O.; Roepstorff, P.; Lundin, K. E. A.; Sjostrom, H.; Sollid, L. M. *Nature Medicine* **1998**, *4*, 713–717.
- (46) Sblattero, D.; Berti, I.; Trevisiol, C.; Marzari, R.; Tommasini, A.; Bradbury, A.; Fasano, A.; Ventura, A.; Not, T. *American Journal of Gastroenterology* **2000**, *95*, 1253–1257.
- (47) Rasmussen, L. K.; Sorensen, E. K.; Petersen, T. E.; Gliemann, J.; Jensen, P. H. *FEBS Letters* **1994**, *338*, 161–166.
- (48) Golde, T. E.; Eckman, C. B.; Younkin, S. G. *Biochimica et Biophysica Acta* **2000**,

1502, 172–187.

- (49) Chen, S.-Y.; Wright, J. W.; Barnes, C. D. *Brain Research* **1996**, *720*, 54–60.
- (50) Johnson, G. V. W.; Cox, T. M.; Lockhart, J. P.; Zinnerman, M. D.; Miller, M. L.; Powers, R. E. *Brain Research* **1997**, *751*, 323–329.
- (51) Ikura, K.; Takahata, K.; Sasaki, R. *FEBS Journal* **1993**, *326*, 109–111.
- (52) Dudek, S. M.; Johnson, G. V. W. *Journal of Neurochemistry* **1993**, *61*, 1159–1162.
- (53) Group, T. H. D. C. R. *Cell* **1993**, *72*, 971–983.
- (54) DiFiglia, M.; Sapp, E.; Chase, K. O.; Davies, S. W.; Bates, G. P.; Vonsattel, J. P.; Aronin, N. *Science* **1997**, *277*, 1990–1993.
- (55) Karpuij, M. V.; Garren, H.; Slunt, H.; Price, D. L.; Gusella, J.; Becher, M. W.; Steinman, L. *Proceedings of the National Academy of Sciences* **1999**, *96*, 7388–7393.
- (56) Kahlem, P.; Terré, C.; Green, H.; Djian, P. *Proceedings of the National Academy of Sciences* **1996**, *93*, 14580–14585.
- (57) Lesort, M.; Attanavanich, K.; Zhang, J.; Johnson, G. V. W. *Journal of Biological Chemistry* **1998**, *273*, 11991–11994.
- (58) Ichinose, A.; McMullen, B. A.; Fujikawa, K.; Davie, E. W. *Biochemistry* **1986**, *25*, 4633–4638.
- (59) Puszkun, E. G.; Raghuraman, V. *Journal of Biological Chemistry* **1985**, *260*, 16012–16020.
- (60) Ikura, K.; Tsuchiya, Y.; Sasaki, R.; Chiba, H. *European Journal of Biochemistry* **1990**, *187*, 705–711.

- (61) Yi, S.-J.; Groffen, J.; Heisterkamp, N. *BMC Biochemistry* **2011**, 12, 8.
- (62) E, F. J.; I, C. S. *Advances in Enzymology and Related Areas of Molecular Biology* **1973**, 38, 109–191.
- (63) E, F. J. *Advances in Enzymology and Related Areas of Molecular Biology* **1983**, 54, 1–56.
- (64) Pardin, C.; Pelletier, J. N.; Lubell, W. D.; Keillor, J. W. *Journal of Organic Chemistry* **2008**, 73, 5766–5775.
- (65) Ohtsuka, T.; Ota, M.; Nio, N.; Motoki, M. *Bioscience, Biotechnology and Biochemistry* **2000**, 64, 2608–2613.
- (66) Sato, H.; Yamada, N.; Shimba, N.; Takahara, Y. *Analytical Biochemistry* **2000**, 281, 68–76.
- (67) McDonagh, R. P., Jr.; McDonagh, J.; Petersen, T. E.; Thorgensen, H. G.; Skorstengaard, K.; Sottrup-Jensen, L.; Magnusson, S.; Dell, A.; Morris, H. R. *FEBS Letters* 127, 174–178.
- (68) Chen, R.; Doolittle, R. F. *Biochemistry* **1971**, 10, 4487–4491.
- (69) Jensen, P. H.; Schuler, E.; Woodrow, G.; Richardson, M.; Goss, N.; Hojrup, P.; Petersen, T. E.; Rasmussen, L. K. *Journal of Biological Chemistry* **1994**, 269, 15394–15398.
- (70) Berbers, G. A. M.; Feenstra, R. W.; Van Den Bos, R.; Hoekman, W. A.; Bloemendal, H.; de Jong, W. W. *Proceedings of the National Academy of Sciences* **1984**, 81, 7017–7020.
- (71) Hohenadl, C.; Mann, K.; Mayer, U.; Timpl, R.; Paulsson, M.; Aeschlimann, D. *Journal of Biological Chemistry* **1995**, 270, 23415–23420.

- (72) Shan, L.; Molberg, O.; Parrot, I.; Hausch, F.; Filiz, F.; Gray, G. M.; Sollid, L. M.; Khosla, C. *Science* **2002**, *297*, 2275–2279.
- (73) Hausch, F.; Halttunen, T.; Mäki, M.; Khosla, C. *Chemistry & Biology* **2003**, *10*, 225–231.
- (74) Kim, S.-Y.; Park, W.-M.; Jung, S.-W.; Lee, J. *Biochemical and Biophysical Research Communications* **1997**, *233*, 39–44.
- (75) Sohn, J. *Journal of Clinical Investigation* **2003**, *111*, 121–128.
- (76) Keresztessy, Z.; Csősz, É.; Hársfalvi, J.; Csomos, K.; Gray, J.; Lightowlers, R. N.; Lakey, J. H.; Balajthy, Z.; Fesus, L. *Protein Science* **2006**, *15*, 2466–2480.
- (77) Sugimura, Y.; Hosono, M.; Wada, F.; Tohru, Y.; Maki, M.; Hitomi, K. *Journal of Biological Chemistry* **2006**, *281*, 17699–17706.
- (78) Lin, C.-W.; Ting, A. Y. *Journal of the American Chemical Society* **2006**, *128*, 4542–4543.
- (79) Chabot, N.; Moreau, S.; Mulani, A.; Moreau, P.; Keillor, J. W. *Chemistry & Biology* **2010**, *17*, 1143–1150.
- (80) Fink, M. L.; Shao, Y. Y.; Kersch, G. J. *Analytical Biochemistry* **1992**, *201*, 270–276.
- (81) Gnaccarini, C.; Ben-Tahar, W.; Mulani, A.; Roy, I.; Lubell, W. D.; Pelletier, J. N.; Keillor, J. W. *Organic & Biomolecular Chemistry* **2012**, *10*, 5258–5265.
- (82) Ichinose, A.; Davie, E. W. *Proceedings of the National Academy of Sciences* **1988**, *85*, 5829–5833.
- (83) Board, P. G.; Web, G. C.; McKee, J.; Ichinose, A. *Cytogenetics and Cell Genetics* **1998**, *48*, 25–27.

- (84) Web, G. C.; Coggan, M.; Ichinose, A.; Board, P. G. *Human Genetics* **1998**, *81*, 157–160.
- (85) Schwartz, M. L.; Salvatore, P. V.; Hill, R. L.; McKee, P. A. *Journal of Biological Chemistry* **1971**, *246*, 5851–5854.
- (86) Ariens, R. A. S.; Lai, T.-S.; Weisel, J. W.; Greenberg, C. S.; Grant, P. J. *Blood* **2002**, *100*, 743–754.
- (87) Adany, R.; Nemes, Z.; Muszbek, L. *British Journal of Cancer* **1985**, *55*, 421–426.
- (88) Bohn, H.; Haupt, H.; Kranz, T. *Blut* **1972**, *25*, 235–248.
- (89) Cooke, R. D.; Pestell, T. C.; Holbrook, J. J. *Biochemical Journal* **1974**, *141*, 675–682.
- (90) E, F. J.; Finlayson, J. S. In *Advances in Protein Chemistry*; Advances in Protein Chemistry; Elsevier, 1977; Vol. 31, pp. 1–133.
- (91) Curtis, C. G.; Brown, K. L.; B, C. R.; A, D.; Gray, A.; Stenberg, P.; Lorand, L. *Biochemistry* **1974**, *13*, 1–7.
- (92) Lewis, S. D.; Janus, T. J.; Lorand, L.; Shafer, J. A. *Biochemistry* **1985**, *24*, 6772–6777.
- (93) Kohler, H. P.; Stickland, M. H.; Ossei-Gerning, N.; Carter, A.; Mikkola, H.; Grant, P. J. *Thrombosis and Haemostasis* **1998**, *79*, 8–13.
- (94) Ando, Y.; Imamura, S.; Yamagata, Y.; Kitahara, A.; Saji, H.; Murachi, T.; Kannagi, R. *Biochemical and Biophysical Research Communications* **1987**, *144*, 484–490.
- (95) Greenberg, C. S.; Achyuthan, K. E.; Fenton, J. W. *Blood* **1987**, *69*, 867–871.
- (96) Yee, V. C.; Pedersen, L. C.; Bishop, P. D.; Stenkamp, R. E.; Teller, D. C. *Thrombosis Research* **1995**, *78*, 389–397.

- (97) Kristiansen, G. K.; Andersen, M. D. *Journal of Biological Chemistry* **2011**, *286*, 9833–9839.
- (98) Muszbek, L.; Yee, V. C.; Hevessy, Z. *Thrombosis Research* **1999**, *94*, 271–305.
- (99) Sabo, M. T.; Brasher, B. P.; Maurer, M. C. *Biochemistry* **2007**, *46*, 10089–10101.
- (100) Turner, B. T.; Sabo, M. T.; Wilding, D.; Maurer, M. C. *Biochemistry* **2004**, *43*, 9755–9765.
- (101) Komaromi, I.; Bagoly, Z.; Muszbek, L. *Journal of Thrombosis and Haemostasis* **2011**, *9*, 9–20.
- (102) Stieler, M.; Weber, J.; Hils, M.; Kolb, P.; Heine, A.; Büchold, C.; Pasternack, R.; Klebe, G. *Angewandte Chemie International Edition* **2013**, *52*, 11930–11934.
- (103) Fox, B. A.; Yee, V. C.; Pedersen, L. C.; Le Trong, I.; Bishop, P. D.; Stenkamp, R. E.; Teller, D. C. *Journal of Biological Chemistry* **1999**, *274*, 4917–4923.
- (104) McDonagh, J.; McDonagh, R. P., Jr.; Del ge, J. M.; Wagner, R. H. *Journal of Clinical Investigation* **1969**, *48*, 940–946.
- (105) Sakata, Y.; Aoki, N. *Journal of Clinical Investigation* **1980**, *65*, 290–297.
- (106) Tamaki, T.; Aoki, N. *Biochimica et Biophysica Acta* **1981**, *661*, 280–286.
- (107) Koseki-Kuno, S.; Yamakawa, M.; Dickneite, G.; Ichinose, A. *Blood* **2003**, *102*, 4410–4412.
- (108) Iwaki, T.; Sandoval-Cooper, M. J.; Paiva, M.; Kobayashi, T.; Ploplis, V. A.; Castellino, F. J. *American Journal of Pathology* **2010**, *160*, 1021–1034.
- (109) Muszbek, L.; Adany, R.; Mikkola, H. *Critical Reviews in Clinical Laboratory Sciences* **1996**, *33*, 357–421.
- (110) Inbal, A.; Lubetsky, A.; Krapp, T.; Castel, D.; Shaish, A.; Dickneite, G.; Modis, L.;

- Muszbek, L.; Inbal, A. *Thrombosis and Haemostasis* **2005**, *94*, 432–437.
- (111) Browder, T.; Folkman, J.; Pirie-Shepherd, S. *Journal of Biological Chemistry* **2000**, *275*, 1521–1524.
- (112) Dardik, R.; Loscalzo, J.; Inbal, A. *Journal of Thrombosis and Haemostasis* **2005**, *4*, 19–25.
- (113) Nurminskaya, M.; Kaartinen, M. T. *Frontiers in Bioscience* **2006**, *11*, 1591–1606.
- (114) Sane, D. C.; Moser, T. L.; Pippen, A. M. M.; Parker, C. J.; Achyuthan, K. E.; Greenberg, C. S. *Biochemical and Biophysical Research Communications* **1998**, *157*, 115–120.
- (115) Bockenstedt, P.; McDonagh, J.; Handin, R. I. *Journal of Clinical Investigation* **1986**, *78*, 551–556.
- (116) Duckert, F. *Annals New York Academy of Sciences* **1972**, *202*, 190–199.
- (117) Undas, A.; Ariens, R. A. S. *Arteriosclerosis, Thrombosis, and Vascular Biology* **2011**, *31*, e88–e99.
- (118) Francis, C. W.; Connaghan, D. G.; Scott, W. L.; Marder, V. J. *Circulation* **1987**, *75*, 1170–1177.
- (119) Nahrendorf, M.; Hu, K.; Frantz, S.; Jaffer, F. A.; Tung, C.-H.; Hiller, K.-H.; Voll, S.; Nordbeck, P.; Sosnovik, D.; Gattenlohner, S.; Novikov, M.; Dickneite, G.; Reed, G. L.; Jakob, P.; Rozenzweig, A.; Bauer, W. R.; Weissleder, R.; Ertl, G. *Circulation* **2006**, *113*, 1196–1202.
- (120) McCormack, L. J.; Kain, K.; Catto, A. J.; Kohler, H. P.; Stickland, M. H.; Grant, P. J. *Thrombosis and Haemostasis* **1998**, *80*, 523–524.

- (121) Attie-Castro, F. A.; Zago, M. A.; Lavinha, J.; Elion, J.; Rodriguez-Delfin, L.; Guerreiro, J. F.; Franco, R. F. *Thrombosis and Haemostasis* **2000**, *84*, 601–603.
- (122) Ariens, R. A. S.; Philippou, H.; Nagaswami, C.; Weisel, J. W.; Lane, D. A.; Grant, P. J. *Blood* **2000**, *96*, 988–995.
- (123) Wartiovaara, U.; Mikkola, H.; Szoke, G.; Haramura, G.; Karpati, L.; Balogh, I.; Lassila, R.; Muszbek, L.; Palotie, A. *Thrombosis and Haemostasis* **2000**, *84*, 595–600.
- (124) Balogh, I.; Szoke, G.; Karpati, L.; Wartiovaara, U.; Katona, E.; Komaromi, I.; Haramura, G.; Pfliegler, G.; Mikkola, H.; Muszbek, L. *Blood* **2000**, *96*, 2479–2486.
- (125) Franco, R. E.; Reitsma, P. H.; D, L.; Maffei, F. H.; Morelli, V.; M, T.; Araujo, A. G.; Piccinato, C. E.; Zago, M. A. *Thrombosis and Haemostasis* **1999**, *81*, 676–679.
- (126) Gemmati, D.; Serino, M. L.; Ongaro, A.; Tognazzo, S.; Moratelli, S.; Riccardo, R.; Moretti, M.; Scapoli, G. L. *American Journal of Hematology* **2001**, *67*, 183–188.
- (127) Corral, J.; Gonzales-Conejero, R.; Iniesta, J. A.; Rivera, J.; Martinez, C.; Vicente, V. *Haematologica* **2000**, *85*, 293–297.
- (128) Margaglione, M.; Bossonne, A.; Brancaccio, V.; Ciampa, A.; Di Minno, G. *Thrombosis and Haemostasis* **2000**, *84*, 1118–1119.
- (129) Catto, A. J.; Kohler, H. P.; Coore, J.; Mansfield, M. W.; Stickland, M. H.; Grant, P. J. *Blood* **1999**, *93*, 906–908.
- (130) Renner, W.; Koppel, H.; Hoffmann, C. *Thrombosis Research* **2000**, *99*, 35–39.
- (131) Kaye, S. M.; Pietiläinen, K. H.; Kotronen, A.; Joutsu-Korhonen, L.; Kaprio, J.; Yki-Järvinen, H.; Silveira, A.; Hamsten, A.; Lassila, R.; Rissanen, A. *Obesity* **2009**, *20*,

88–94.

- (132) Sorensen, E. S.; Rasmussen, L. K.; Moller, L.; Jensen, P. H.; Hojrup, P.; Petersen, T. E. *Biochemical Journal* **1994**, *304*, 13–16.
- (133) Bendixen, E.; Borth, W.; Harpel, P. C. *Journal of Biological Chemistry* **1993**, *268*, 21962–21967.
- (134) Asijee, G. M.; Muszbek, L.; Kappelmayer, J.; Polgar, J.; Horvath, A.; Sturk, A. *Biochimica et Biophysica Acta* **1988**, *954*, 303–308.
- (135) Huh, M. M.; Schick, B. P.; Schick, P. K.; Colman, R. W. *Blood* **1988**, *71*, 1693–1702.
- (136) Gorman, J. J.; Folk, J. E. *Journal of Biological Chemistry* **1980**, *255*, 419–427.
- (137) Fickenscher, K.; Aab, A.; Stuber, W. *Thrombosis and Haemostasis* **1991**, *65*, 535–540.
- (138) Kimura, S.; Aoki, N. *Journal of Biological Chemistry* **1986**, *261*, 15591–15595.
- (139) Oertel, K.; Hunfeld, A.; Specker, E.; Reiff, C.; Seitz, R.; Pasternack, R.; Dodt, J. *Analytical Biochemistry* **2007**, *367*, 152–158.
- (140) Lynch, G. W.; Slayter, H. S.; Miller, B. E.; McDonagh, J. *Journal of Biological Chemistry* **1987**, *262*, 1772–1778.
- (141) Cohen, I.; Blankenburg, T. A.; Borden, D.; Kahn, D. R.; Veis, A. *Biochimica et Biophysica Acta* **1980**, *682*, 365–375.
- (142) Lee, K. N.; Birckbichler, P. J.; Patterson, M. K., Jr *Clinical Chemistry* **1988**, *34*, 906–910.
- (143) Miraglia, C. C.; Greenberg, C. S. *Analytical Biochemistry* **1985**, *144*, 165–171.
- (144) Lorand, L.; Campbell, L. K. *Analytical Biochemistry* **1971**, *44*, 207–220.

- (145) Kusch, M.; Grundmann, C.; Keitel, S.; Seitz, R.; Konig, H. *Blood Coagulation and Fibrinolysis* **2006**, *17*, 575–580.
- (146) Clouthier, C. M.; Mironov, G. G.; Okhonin, V.; Berezovski, M. V.; Keillor, J. W. *Angewandte Chemie International Edition* **2012**, *51*, 12464–12468.
- (147) Pardin, C.; Roy, I.; Chica, R. A.; Bonneil, E.; Thibault, P.; Lubell, W. D.; Pelletier, J. N.; Keillor, J. W. *Biochemistry* **2009**, *48*, 3346–3353.
- (148) Al-Jallad, H. F.; Myneni, V. D.; Piercy-Kotb, S. A.; Chabot, N.; Mulani, A.; Keillor, J. W.; Kaartinen, M. T. *PLoS ONE* **2011**, *6*, e15893.
- (149) Roy, I.; Smith, O.; Clouthier, C. M.; Keillor, J. W. *Protein Expression and Purification* **2013**, *87*, 41–46.
- (150) Pardin, C.; Roy, I.; Lubell, W. D.; Keillor, J. W. *Chemical Biology & Drug Design* **2008**, *72*, 189–196.
- (151) Smethurst, P. A.; Griffin, M. *Biochemical Journal* **1996**, *313*, 803–808.

CHAPTER 2

Sulfoxide Inhibitors

Introduction

2.1 Enzymatic transition states

Enzymes increase reaction rates, usually at rates of 10^{10} to 10^{15} times faster relative to the same conditions without a catalyst (Figure 2.1).¹ One of the important mechanisms of enzyme catalysis is the differential binding of the transition stateⁱ (TS) of the reaction to the enzyme with greater affinity than the corresponding substrates or products.

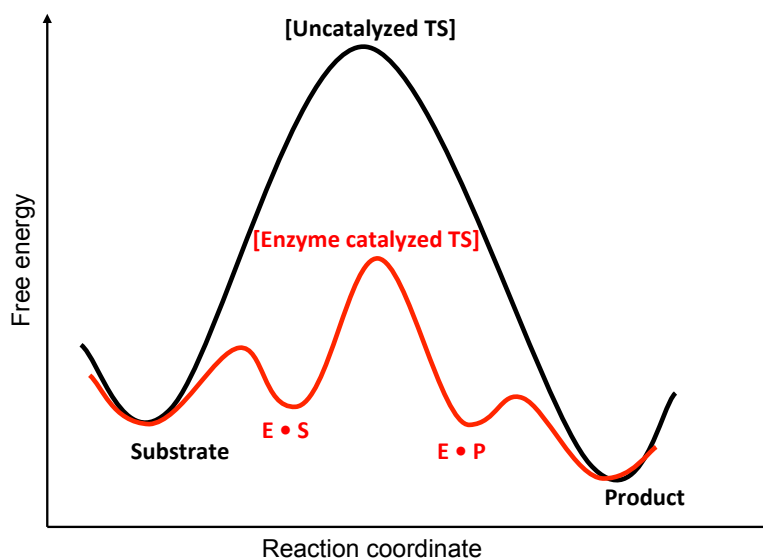


Figure 2.1: Energy profile of catalysis.

The more tightly an enzyme binds its reaction's TS relative to the substrate, the greater the rate of the catalyzed reaction relative to that of the uncatalyzed reaction. So it can be said that catalysis results from preferential binding and thus stabilization of the TS relative to the substrate. In other words, a good substrate won't necessarily be bound with high affinity but may be bound more tightly upon activation to the TS.

ⁱ Transition state will be used herein as a term to refer to the activated complex

Pauling, Wolfenden and Lienhard have all suggested that interactions that preferentially bind the TS will increase its concentration and thus, proportionally increase the reaction rate.²⁻⁴ If an enzyme preferentially binds the TS, then TS analogues - stable molecules that resemble the TS or one of its components - should be potent, competitive inhibitors of the enzyme.

The transition state can be described as the state at which in a reaction, there is equal probability that either the reaction will occur and products will be formed or the species will decompose back into the reactants. The transformation of bond vibrational modes (the stretching and restoring forces) to bond-dissociating translational modes is another way to define the transition state. It occurs within the time span of bond vibration, usually on the order of 10^{-13} s whereas the typical enzyme catalysis rate constant is $1 \times 10^3 \text{ s}^{-1}$, with the delay owing to conformational adjustments to align the reactants in the catalytic site.

It is a fluid balance, whereby the substrate is diffusing in and out of the active site, bond formation is reversible and at the moment when the non-bonding interactions have productively aligned, the angles are correct and there is just enough tension/stretching in the enzyme-substrate complex, the transition state is attained. The fleeting lifetime of the enzymatic TS prevents a thermodynamic equilibrium from forming a tightly bound TS-enzyme complex. Thermodynamic descriptions of the TS complex require bond equilibrium to propagate through both the enzyme and the reactants during the lifetime of the TS; however, since protein conformations are slow, this is impossible during the TS lifetime.⁵

Movement of the enzyme with the substrate is dynamic, constantly sampling different interactions, some of which may lead to the formation of the TS. These continual fluctuations, along with the movements of the enzyme sidechains, loops and domains, all change the enzyme-reactant distances and dynamic interaction energies. Rotation about covalent bonds and flap and domain motions on the order of 10^{-8} to 10^{-3} s form the alignment which is on the cusp to attain the TS.⁶

The TS is obtained when each of the bonds essential to catalysis moves in a coordinated fashion and forms the appropriate contacts necessary to achieve the TS. Computational experiments studying enzyme function have shown that superfamilies of enzymes have conserved functional residues acting as coupled networks that favour catalytic motions more frequently than random (non-catalytic) motions.⁷ However, the momentary geometry that forms the TS does not produce a large-scale conformational change throughout the protein as a tightly bound TS does not exist except in the dynamic sense of a fraction of picosecond in which it actually occurs. In order to capture the rare conformational events that lead to the formation of the TS, transition state analogues have been proposed.

2.2 Transition state analogues

Transition state analogues are competitive inhibitors that take advantage of the bonding interactions present to stabilize the transition state complex during the catalyzed reaction. Although transition state analogues are frequently found to be tight-binding inhibitors, in this chapter we do not use the potency of inhibition as a criterion for being a “transition state analogue”, but rather use this description based

on the way an inhibitor was designed - namely, to resemble the transition state of a given reaction step.

Compared to the forces that bind the TS, the reactants are bound weakly. This is illustrated by bovine purine nucleoside phosphorylase (PNP). The substrate for bovine PNP, inosine, binds fairly weakly to the catalytic site, with a Michaelis constant of 28 μM . The weak binding is reflected in the long hydrogen bond distances observed in the crystal structure of the Michaelis complex.⁸ However, when immucillin-H, a TS analogue was crystallized in the PNP catalytic site, six new hydrogen bonds were observed. This is accompanied by an initial K_i of 41 nM that drops to 23 pM, upon the slow isomerization of the enzyme that causes increased inhibitor binding affinity.⁹ The improved binding of the TS analogue compared to the substrate can be attributed to the reciprocity between the interaction sites on the inhibitor and the reorganization of the protein to keep the configuration that is acquired instantaneously and dynamically in the actual TS.

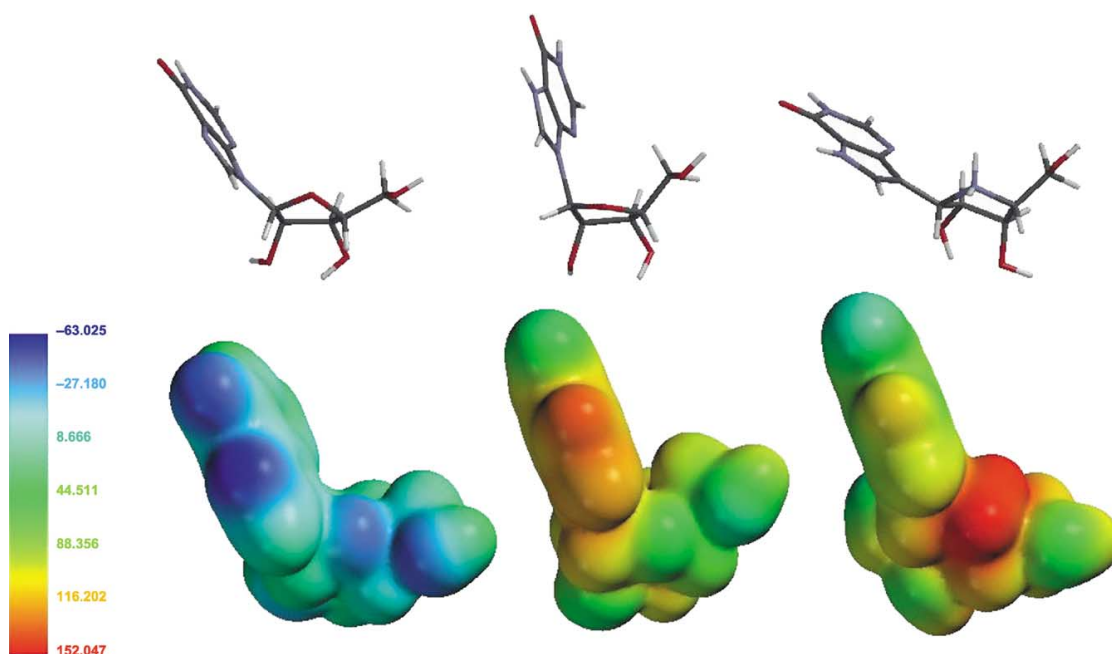


Figure 2.2: Molecular electrostatic potential surfaces for inosine, the transition state of PNP and immucillin-H, respectively. (Schramm, V. L., *Biochimica et Biophysica Acta*, **2002**, 1587, 107-117.) Reproduced with permission.

From Figure 2.2, the similarity of immucillin-H to the transition state can be seen by the molecular electrostatic potential surfaces which compare the substrate, the TS and immucillin-H, respectively.⁹ The similarity of the electrostatic potential of immucillin-H to that of the TS capitalizes on the existing forces used to stabilize the TS whereas the weak binding of inosine can be explained by its inability to make such contacts.

The presence of the TS analogue will induce the protein to adopt a static, thermodynamically equilibrated conformation whereby the release of the analogue would result in a slow and energetically unfavourable expansion of the protein. Owing to this phenomenon, the release rate of TS analogue inhibitors is very slow, ranging

from minutes to days, and this causes tight binding. Dissociation constants in the picomolar range are not uncommon and there have even been analogues with dissociation constants in the femtomolar range described in the literature.¹⁰

2.2.1 Examples of TS analogues

There are a wide variety of transition state analogues that can be found in the literature and examples include immucillin derivatives from the group of Schramm that effectively inhibit 5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase (MTAN). This enzyme is responsible for hydrolyzing its substrates to form adenine and 5-methylthioribose or S-ribosylhomocysteine. One of the most powerful transition state analogues to date includes 5'-*p*-Cl-phenylthio-DADMe-Immucillin-A (2.1) with a K_i value of 47 fM (Figure 2.3).¹⁰ It incorporates features of the substrate, 5'-methylthioadenosine while mimicking the highly dissociated N-ribosyltransferases transition states.

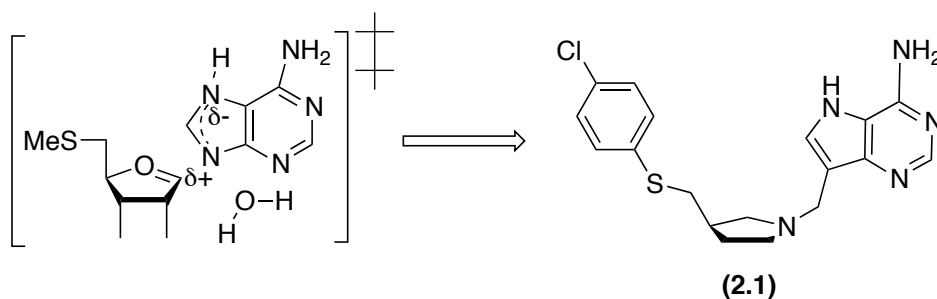


Figure 2.3: TS analogue for MTAN.

Previously, in our group¹¹ we have also experimented with transition state analogues as inhibitors for γ -glutamyl transpeptidase (GGT), an enzyme found in mammals, bacteria and plants, having a key role in the metabolism of glutathione.

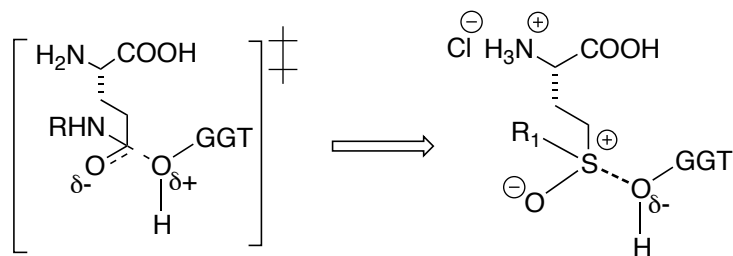


Figure 2.4: TS analogue for γ -glutamyl transpeptidase.

The sulfoxide inhibitor shown here (Figure 2.4) has a K_i value of 53 μM , which is 10-fold lower than the K_m value for its native substrate, glutathione.¹¹ The sulfoxide resembles the transition state in both its geometry and its charge distribution, which would account for its high affinity.¹²

2.3 Tissue transglutaminase

Tissue transglutaminase (TG2) has been identified as an important, therapeutic target due to its role in a number of human disease states including, but not limited to Celiac disease¹³ as well as neurodegenerative diseases such as Alzheimer's,¹⁴ and Huntington's.¹⁵ As a result, a variety of TG2 inhibitors have been synthesized that could serve, not only as useful therapeutics, but also as chemical probes to better understand the function and role of TG2 in the aforementioned disease states.

2.3.1 Irreversible inhibitors of TG2

To date, a number of potential TGase inactivators have been developed based on the possibility of reaction with the active site cysteine, including 1,2,4-thiadiazoles (2.2),¹⁶ dihydroisoxazole derivatives (2.3),¹⁷ dipeptide-bound α,β -unsaturated amides

(2.4),¹⁸ and epoxides (2.5),¹⁸ (Figure 2.5). The majority of these TG2 inhibitors act by targeting the nucleophilic cysteine residue in the TG2 active site, which can be alkylated or acylated by a requisite electrophilic intermediary. While highly effective, many irreversible TG2 inhibitors suffer from a general lack of specificity owing to the reactive nature of the electrophilic warhead groups. As all members of the transglutaminase family of enzymes contain a nucleophilic cysteine residue, these irreversible inhibitors are not specific to a targeted transglutaminase.

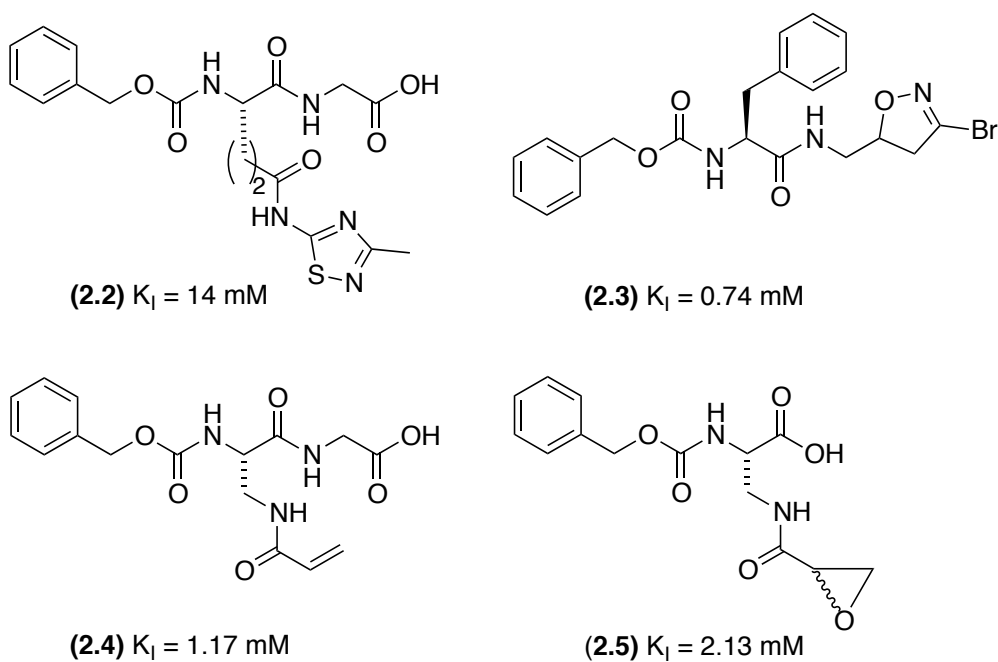


Figure 2.5: Irreversible inhibitors of TG2.

2.3.2 Reversible inhibitors of TG2

Various reversible inhibitors for TG2 have also been synthesized including acylideneoxindole derivatives (2.6),¹⁹ cinnamoyl inhibitors²⁰ (such as 2.7) and a drug-like compound LDN-27219 (2.8) which was found to bind at TG2's GTP-binding site rather at the active site of the enzyme²¹ (Figure 2.6).

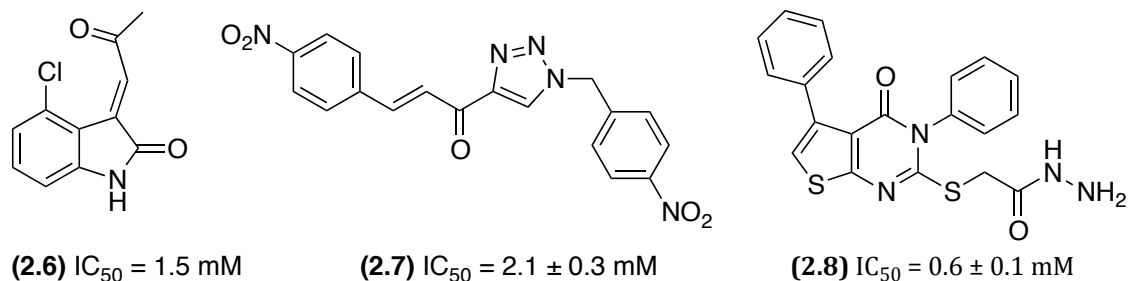


Figure 2.6: Reversible inhibitors of TG2.

Apart from the acylideneoxindole derivatives, the reversible inhibitors 2.7 and 2.8 in Figure 2.6 are specific for TG2 as opposed to other transglutaminases such as Factor XIII. Unfortunately, since the high specificity of these compounds has been linked to their unique conformation, efforts to transfer this specificity to other molecules have proven difficult.

2.3.3 PNPQLPF as affinity-conferring peptidic sequence for TG2

In an effort to improve the specificity of irreversible inhibitors for TG2, research has centred on imbedding electrophilic warhead groups within a peptidic chain that has shown specificity for TG2. Khosla and coworkers identified a truncated peptidic sequence derived from the gluten protein, composed of PQPQLPY,²² that can serve as an excellent substrate for TG2. Kim and coworkers, inspired by sequences from pro-elafin,

a known TG2 substrate for anchoring pro-elafin onto matrix proteins, identified the peptide KVLGDGQDP as a good substrate for TG2.²³ Recently, Hitomi and coworkers, using a phage displayed random peptide library to investigate the primary structures surrounding reactive glutamine residues for TG2, found that specificity for TG2 was based more on peptide conformation rather than sequence.²⁴ Thus, it is not the specific amino acids themselves that confer affinity to TG2 but rather, their overall characteristics.

Inspired by these observations, solid phase synthesis was undertaken to synthesize a variety of peptidic sequences in an effort to find an affinity-conferring peptide. It was determined that the sequence PNPQLPF had excellent affinity for TG2 with a K_m of 6 μ M (Pardin and Keillor, unpublished results). Incorporating an electrophilic moiety in the place of the glutamine residue should produce an irreversible inhibitor specific to TG2. Using this motif, a fluorescent irreversible inhibitor was synthesized for TG2 (RhodB-PGG-K(acryl)LPF-OH) with a K_i of 79 μ M and used as a probe to investigate the role of TG2 in arterial rigidification.²⁵

2.3.4 Transition state analogue design for TG2

Transition state analogue inhibitors are designed from the hypothesis that chemically stable compounds that simulate the geometric and molecular electrostatic characteristics of the transition state will bind to the enzyme tighter than the substrate by a factor approaching the catalytic rate acceleration imposed by the enzyme. From Figure 2.7, it can be seen that there are two transition states leading to and from a tetrahedral intermediate prior to the formation of the acyl-enzyme intermediate. Thus

inhibitors based on the transition states leading to this acyl-enzyme intermediate of TG2 should be effective, by taking advantage of the interactions that occur to stabilize the transition states.

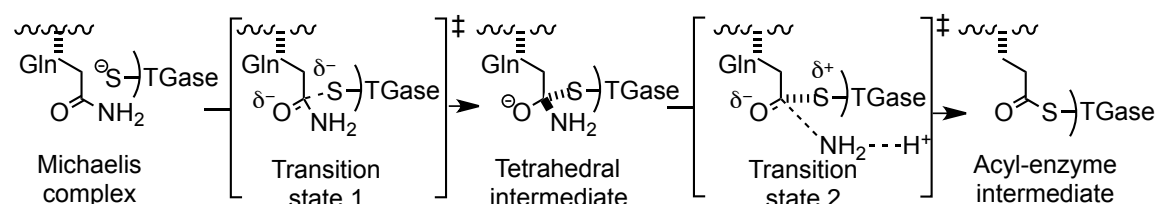


Figure 2.7: Mechanism of TGase acylation.

It was hypothesized that the structure of a sulfinamide molecule would resemble the transition state adopted during the acylation step. Structural studies of aryl sulfinamides have demonstrated that the sulfur atom adopts a tetrahedral-like (as opposed to planar) configuration and the nitrogen atom is flattened slightly to form the trigonal sp^2 structure found in the ground state amide atom.²⁶ Thus, we proposed compounds 2.9 and 2.10 as analogues of the first and second transition states towards the acylation step of the cross-linking catalyzed by TG2.

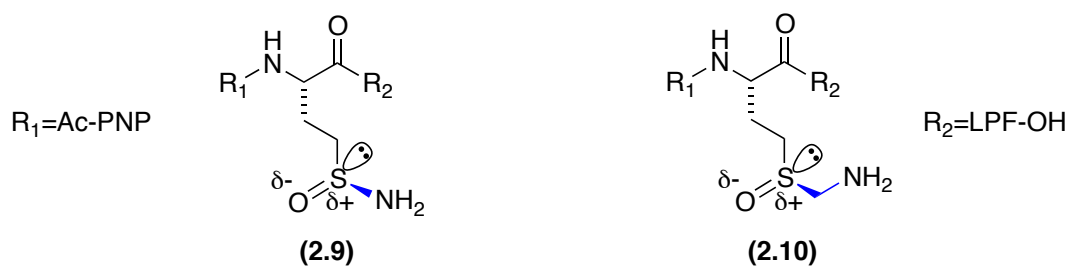


Figure 2.8: Transition state analogues for TG

The R_1 group would be Ac-PNP while the R_2 group would be LPF-OH, which is the amino acid sequence shown to have high affinity and specificity for TG2 relative to other transglutaminases.

These targeted inhibitors should have high affinity and efficacy as competitive inhibitors, since they are based on the transition states leading to the acyl-enzyme intermediate of TG2 and will take advantage of the interactions that stabilize and favour the transition states of the acylation reaction.

2.4 S-N bond-containing compounds

There exist three different types of compounds containing a sulfur-nitrogen bond, namely sulfenamides (2.11), sulfinamides (2.12) and sulfonamides (2.13) (Figure 2.9).

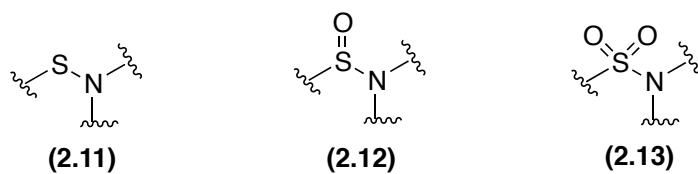


Figure 2.9: Sulfenamides (2.11), sulfinamides (2.12) and sulfonamides (2.13), respectively.

Sulfenamides are organic compounds containing a trivalent nitrogen bonded to a divalent sulfur. They are fairly unstable compounds, owing to the lability of the sulfur-nitrogen bond. It has been noted that the thermal stability of sulfenamides decrease as

the basicity of the sulfenamide nitrogen increases, as the electron-donating ability of the group attached to sulfur increases and when the number of groups attached to the sulfenamide nitrogen decreases.²⁷ This would probably also hold true for sulfinamides (2.12), attached to an inductive electron-donating electronegative atom such as oxygen and a primary amine. Field and coworkers also demonstrated that aromatic sulfenamides prepared from primary aliphatic amines were more reactive than those from dialkylamines, and all attempts to make alkyl sulfenamides were unsuccessful.²⁷

Aryl sulfenamides have also been known to undergo exchange reactions on exposure to light and with gentle heating.²⁸ Cole and coworkers showed that sulfinamides can also undergo exchange reactions at room temperature, suggesting that intermolecular protonation of the sulfinyl group leads to polarization of the sulfur atom and a more basic amine can then displace the original nitrogen.²⁹

Sulfonamides are oxidized sulfur molecules and have been used extensively as antibacterial drugs but have also shown some utility as transition state isosteres for enzyme proteases.³⁰ They contain the requisite tetrahedral geometry but have been found to perform poorly as inhibitors due to their electrostatic effects.³¹ Another drawback is the known tendency of peptidic sulfonamides to fragment, (Figure 2.10).³²

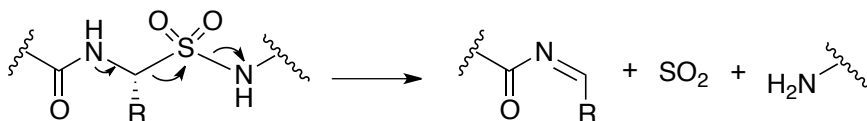


Figure 2.10: Fragmentation of peptidic sulfonamides.

Thus, both sulfenamides and peptidic sulfonamides are labile compounds, owing to their propensity towards S-N bond fissures. In addition to their instability, neither would be appropriate as TS analogues. While sulfenamides and sulfonamides may have tetrahedral geometry, their electrostatic potential is different enough that they do not bear a strong resemblance to the transition state and thus, would not be effective.

2.4.1 Synthesis of sulfinamides

The strong resemblance between our proposed sulfinamide molecules 2.9 and 2.10, and the transition states of TG2 induced us to attempt the synthesis of these potentially unstable molecules. There are a variety of ways to synthesize sulfinamide molecules, including the reaction of sulfinyl chlorides with ammonia,³³ and Brownbridge's method of synthesizing methyl sulfinates combined with the Davis protocol of treatment with LiHMDS and an amine (Figure 2.11).^{34,35} Ellman has reacted sulfinates with a mixture of sodium and ammonia,³⁶ and there is also metal assisted synthesis by Davis whereby metal salts (AgNO_3 , AgOAc , HgCl_2) are dissolved in methanol or ethyl acetate followed by the addition of disulfide and an excess of amine followed by removal of precipitated mercaptide which yields the sulfinamide (Figure 2.11).³⁷

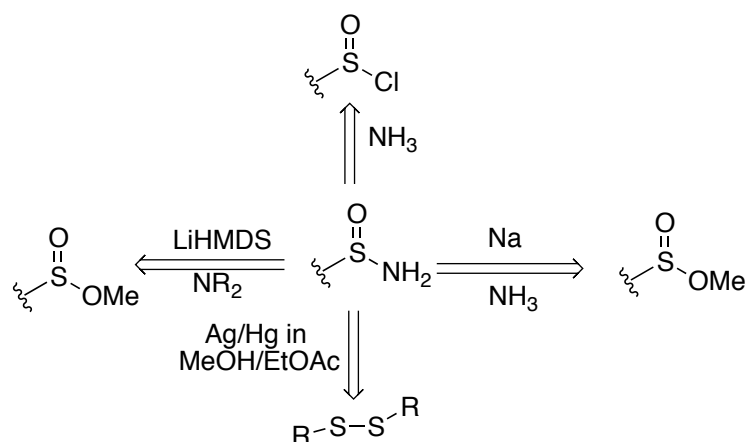
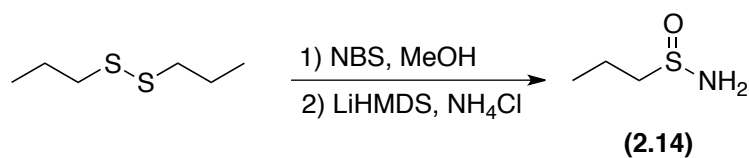


Figure 2.11: Retrosynthetic schemes of sulfinamide synthesis.

Almost all of the examples of sulfinamides in the literature are either aromatic sulfinamides or *tert*-butyl sulfinamides, due to the presumed instability of alkyl sulfinamides. It is very rare to see a compound with a saturated carbon atom adjacent to the sulfinamide moiety. However, a protocol was found by Ruano and coworkers,³⁸ using the Davis method to produce alkyl sulfinamides in adequate yield (40%) as shown in Scheme 2.1.



Scheme 2.1: Synthesis of propyl sulfinamide.

With this synthesis in mind, and knowing that transition state analogues often make highly effective inhibitors, we decided to tackle the synthesis of our proposed sulfinamide compounds 2.9 and 2.10.

Results and Discussion

2.5 Modelling of the TG2 active site

Modelling studies were undertaken previously in collaboration with the Pelletier group using the Insight 2 package, whereby the initial coordinates were taken from a structure of red sea bream tissue TGase.³⁹ The percentage of identity of the catalytic domains of red sea bream TGase with guinea pig liver TGase is greater than 55% while the percentage of homology is ~70%. Guinea pig liver TGase has an identity of 82% with human tissue TGase while the percentage of similarity is 95%. Additionally, all the active site residues of red sea bream tissue TGase are conserved in all sequenced tissue TGases, validating the use of these coordinates for the modelling studies. The donor substrate used in the modelling studies was a Cbz-glutaminyglycine molecule, which was manually positioned into the active site. The model was constructed, minimized, molecular dynamics completed for 10 ps and the structure then reminimized. These modelling studies showed that during the catalytic steps, the tetrahedral intermediate of the acyl enzyme intermediate is stabilized through hydrogen bonding.

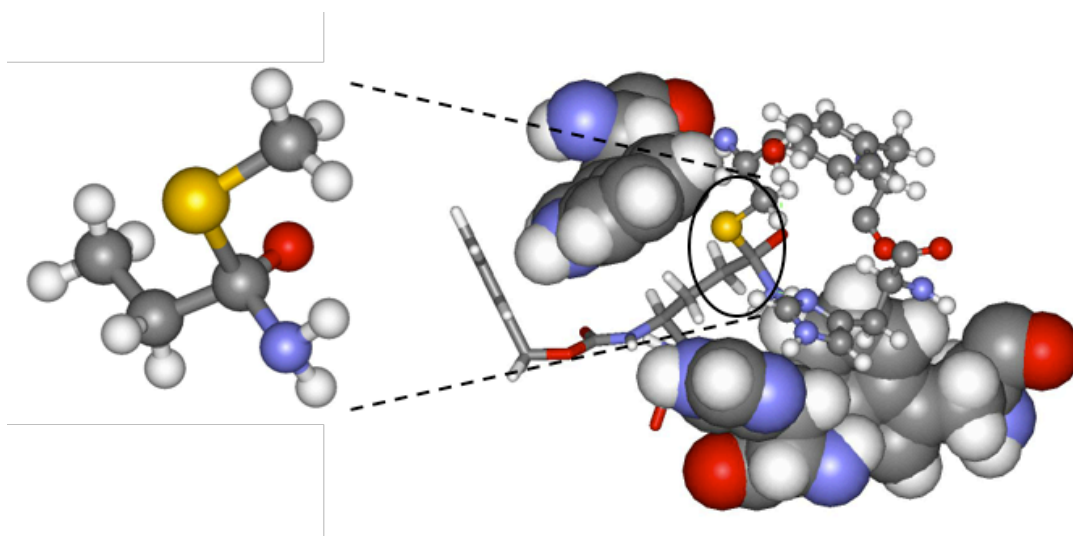


Figure 2.12: Active site of TG2 with Cbz-Gln-Gly.

Since the tetrahedral intermediate is a high-energy species, the transition states leading to its formation and from its decomposition must be similar to it in geometry as well as energy, according to the Hammond postulate.⁴⁰

Thus, using the results of the aforementioned modelling studies, semi-empirical modelling was undertaken using Parameterized Model 3 (PM3) of the anionic tetrahedral intermediate of the acyl-enzyme with the glutamine side chain whereby the sulfur-methyl group is representative of the cysteine bond of transglutaminase to its substrate.

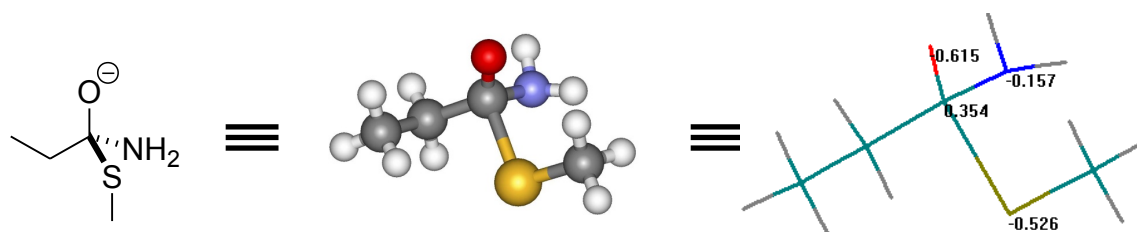


Figure 2.13: Charge distribution of the tetrahedral intermediate of acyl-enzyme.

Modelling resulted in the charge distributions and geometry shown in Figure 2.13 whereby there is a negative charge density around the oxygen atom and a partial positive charge on the carbon atom. If the covalent sulfur-carbon bond is then cleaved and the resulting neutral hypothetical molecule is modelled, fixing the geometry and recalculating the charge distributions, the fictitious molecule is shown to have a high complementarity to the enzyme active site (Figure 2.14). This then serves as a target for mimics that should have high complementarity and affinity for the active site.

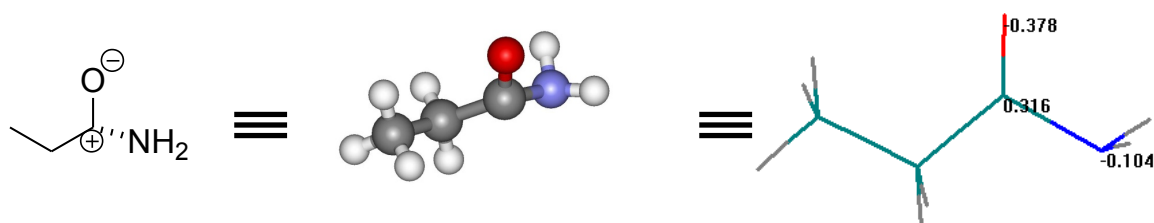


Figure 2.14: Charge distribution around hypothetical molecule resembling TS intermediate.

A sulfinamide derivative was then modelled as an analogue of the tetrahedral intermediate since the central sulfur atom is tetrahedral and retains a partial positive charge (Figure 2.15).

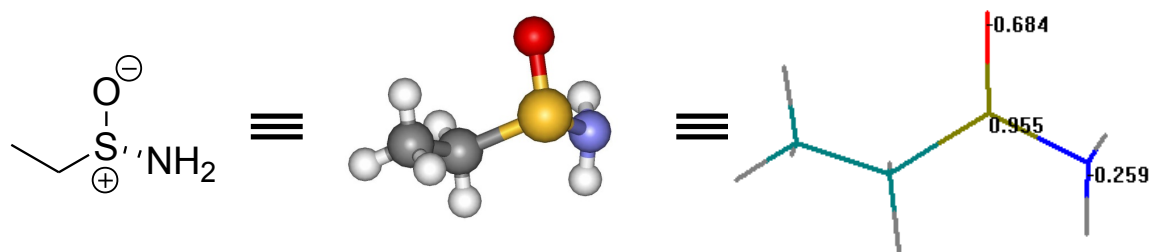


Figure 2.15: Charge distribution around sulfinamide molecule.

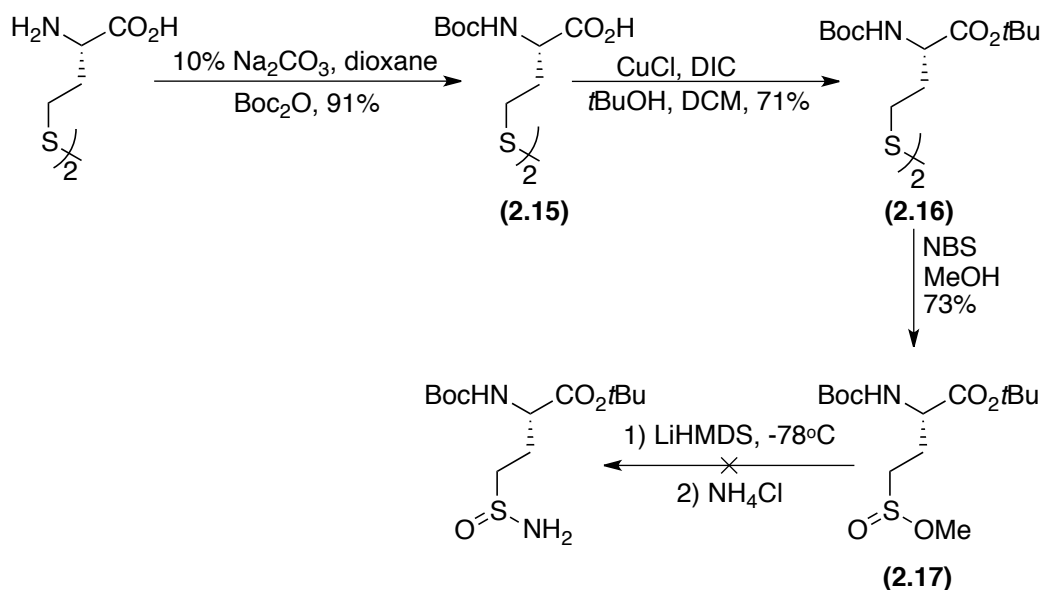
It can be seen that the sulfinamide molecule has a similar charge distribution and geometry as the hypothetical molecule that was designed to have a high affinity for the enzyme active site, thus it should be an effective inhibitor. These calculations validate the choice of our target molecules that include a sulfinamide compound, (2.9) as well as an aminomethyl sulfoxide, (2.10) as analogues of the first and second transition states, respectively, in the acylation step of the cross-linking process.

2.6 Synthesis of sulfinamide inhibitors

Having confirmed the viability of sulfinamide molecules as transition state analogues for TG2, we turned our attention to the difficult synthesis of these unknown compounds. Sulfinamides are not well characterized in the literature, perhaps owing to the instability of this class of compounds and as already mentioned, and are usually found as aromatic or *tert*-butyl sulfinamides.

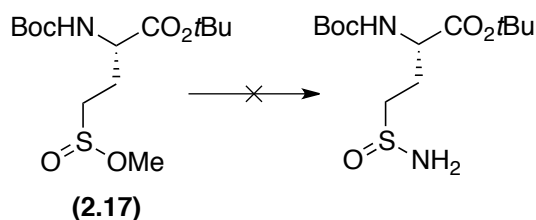
2.6.1 Synthesis of sulfinamide (2.9) via a methyl sulfinite

As we were planning to incorporate the transition state mimic into a peptidic chain having affinity for TG2, Boc and *tert*-butyl esters were chosen as protecting groups for the synthesis of compound 2.9 since they could be removed easily under the same acidic conditions. The amino group could then be selectively protected with an Fmoc-protecting group as per standard Fmoc-based solid phase coupling procedures.²⁵



Scheme 2.2: Synthesis of compound 2.8 via methyl sulfinic acid

Thus, L-homocystine, prepared as per literature sources,⁴¹ was first protected using Boc anhydride, followed by carboxylic acid protection as a *tert*-butyl ester (Scheme 2.2). Once both functional groups were protected, *N*-bromosuccinimide in MeOH was used to convert the disulfide into sulfinic esters as per the Brownbridge protocol.³⁴ This was convenient because it eliminated the oxidation of sulfur as an additional step. The last step of the synthesis was to displace the methanol with an amine group. The literature procedure was confirmed to work in the case of propyl disulfide (Scheme 2.1) but was unsuccessful with the homocystine derivative (Table 2.1).



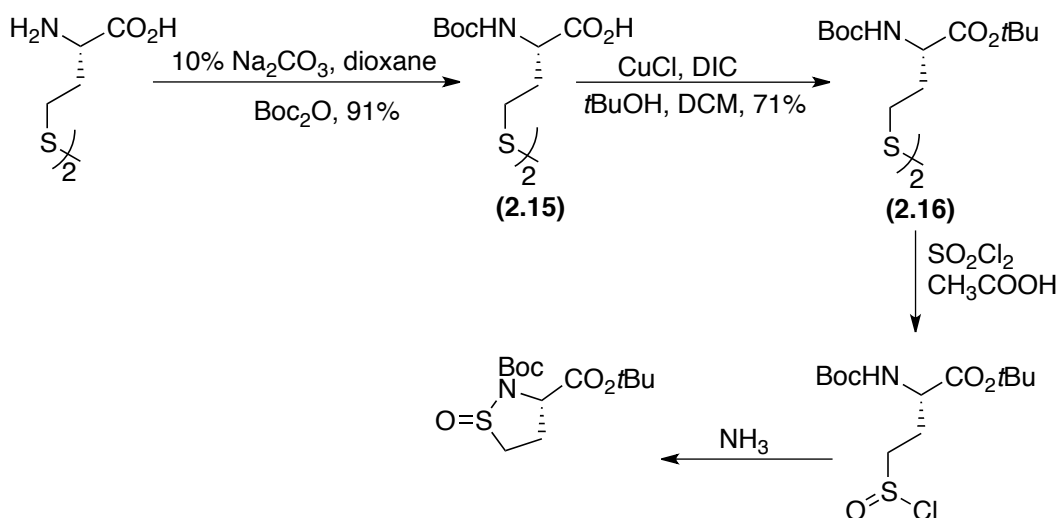
Attempt	Condition	Result
1	1))LiHMDS, -78°C 2) NH ₄ Cl sat.	product decomposition
2	excess NH ₃ in MeOH/THF, reflux	product decomposition
3	3 eq. NaH, THF, excess NH ₃	product decomposition
4	<i>n</i> BuLi, excess NH ₃	product decomposition
5	Fe(NO ₃) ₃ •9H ₂ O, Na, NH ₃ , NH ₄ Cl	product decomposition

Table 2.1: Conditions attempted in synthesis of sulfonamide.

While treatment with the LiHMDS conditions had been employed successfully on a variety of compounds in the literature,³⁸ they were all less functionalized, either with unsubstituted carbon chains (*ex.* propyl disulfide) or if functionalized, the heteroatoms were substituents on an aromatic ring. Numerous variations were attempted that involved inciting the nucleophilic substitution of the methanolate by ammonia (with and without heating) (Table 2.1). When that proved unsuccessful, we tried to improve the nucleophilicity of ammonia by deprotonating it with either (1) sodium hydride, (2) *n*-butyl lithium or (3) pellets of sodium (Table 2.1).³⁶ However, only product decomposition was ever observed. This was confirmed through analysis of the NMR

spectra whereby neither starting material nor product could be conclusively identified. Various attempts to purify the reaction mixtures were unsuccessful as purification conditions could not be determined.

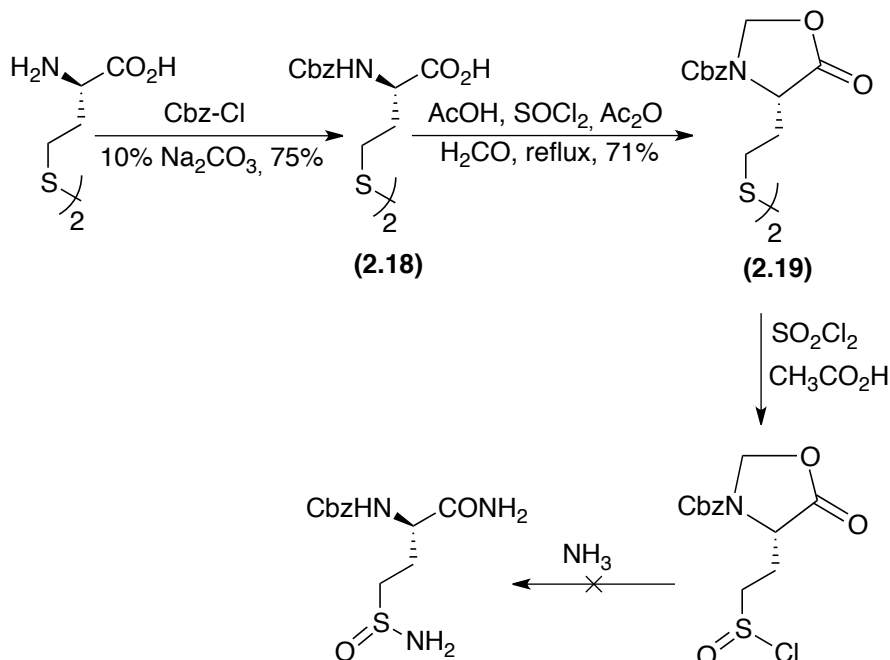
2.6.2 Synthesis of sulfinamide employing sulfinyl chloride



Scheme 2.3: Sulfinamide synthesis via sulfinyl chloride

In another attempt to afford compound 2.9, the same protection scheme was used as previously; however, instead of making a methyl sulfinate, the sulfinyl chloride was made directly (Scheme 2.3). As chlorine is more electrophilic than the methanolate ion, displacement by ammonia should be easier. Given that it was a one-pot reaction and the sulfinyl chloride was not isolated, IR analysis was used to confirm its presence before an excess of ammonia was added. Unfortunately, the compound cyclized to form an isothiazolidine, which was identified through MS-ESI⁺ analysis.

2.6.3 Synthesis of sulfinamide using oxazolidinone



Scheme 2.4: Synthesis of sulfinamide using oxazolidinone.

In a third attempt, the protection scheme was changed to have a Cbz protecting group on the amino moiety (Scheme 2.4). This was done since it has been shown that TG2 has no affinity for a Boc protecting group but will still recognize a Cbz moiety.³⁹ In this way, once we had the sulfinamide, we would be able to test its efficacy immediately without having to undergo any deprotection/reprotection reactions nor solid phase peptide synthesis prior to a preliminary test.

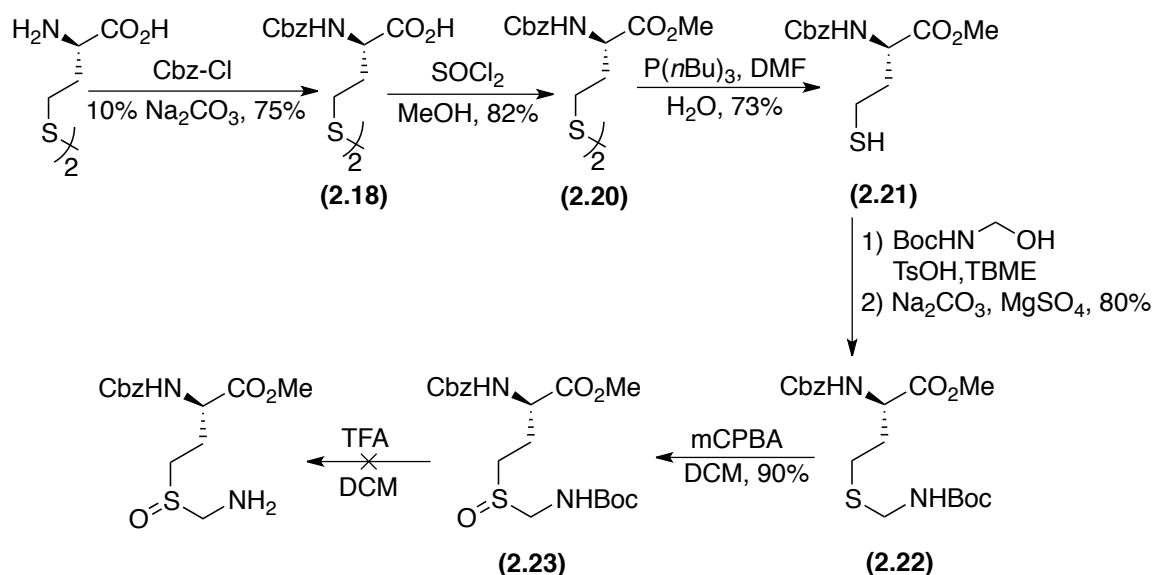
The carboxylic acid was protected using an oxazolidinone that should avoid the formation of the isothiazolidine by-product since the nucleophilic amino group would be embedded in the oxazolidinone ring. Once again, the presence of the sulfinyl chloride was confirmed by IR analysis prior to ammonia addition. It was expected that the basic conditions would cleave the oxazolidinone and afford the C-terminal amide while

simultaneously forming the sulfinamide, as ammonia was added in excess.

Unfortunately, the desired product was not observed; it is unclear if the desired sulfinamide was formed but was unstable and decomposed after formation. Literature precedence points to the sulfinamide moiety being unstable and unable to be isolated, which will be discussed below.^{26,42} It was hypothesized that our amino acid sulfinamides, once formed, were not stable due to the high degree of functionalization of the molecule and immediately broke down.

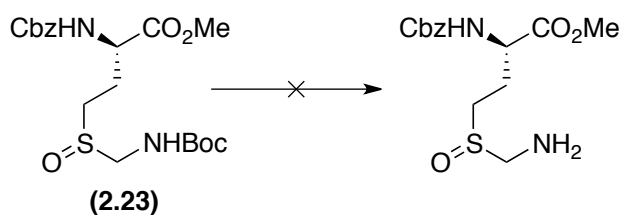
2.6.4 Synthesis of aminomethyl sulfoxide (2.10)

Our second target molecule was an aminomethyl sulfoxide inhibitor (2.10) that would mimic the second transition state of the acylation step, as the ammonia is ejected from the acyl-enzyme intermediate.



Scheme 2.5: Synthesis of aminomethyl sulfoxide

The first step of the synthesis involved a Cbz protection for the amino group, and carboxylic acid protection using a methyl ester. Homocystine was reduced to homocysteine using 2 eq. of tributyl phosphine. The free thiol was reacted with Boc-aminomethanol (prepared according to a literature procedure),^{43,44} using *p*-toluenesulfonic acid as a catalyst to afford the Boc-protected aminomethyl sulfinyl compound 2.22. This was then oxidized with *m*CPBA to yield the Boc-aminomethyl sulfoxide compound. The last step to obtain the inhibitor was a Boc deprotection; however, standard deprotection conditions, namely a 50:50 mixture of TFA:DCM, led to product degradation.



Attempt	Conditions	Result
1	TFA, CH ₂ Cl ₂	product degradation
2	H ₃ PO ₄ , THF	product degradation
3	NEt ₃ , TMSOTf, DCM	product degradation
4	TBAF, THF	product degradation

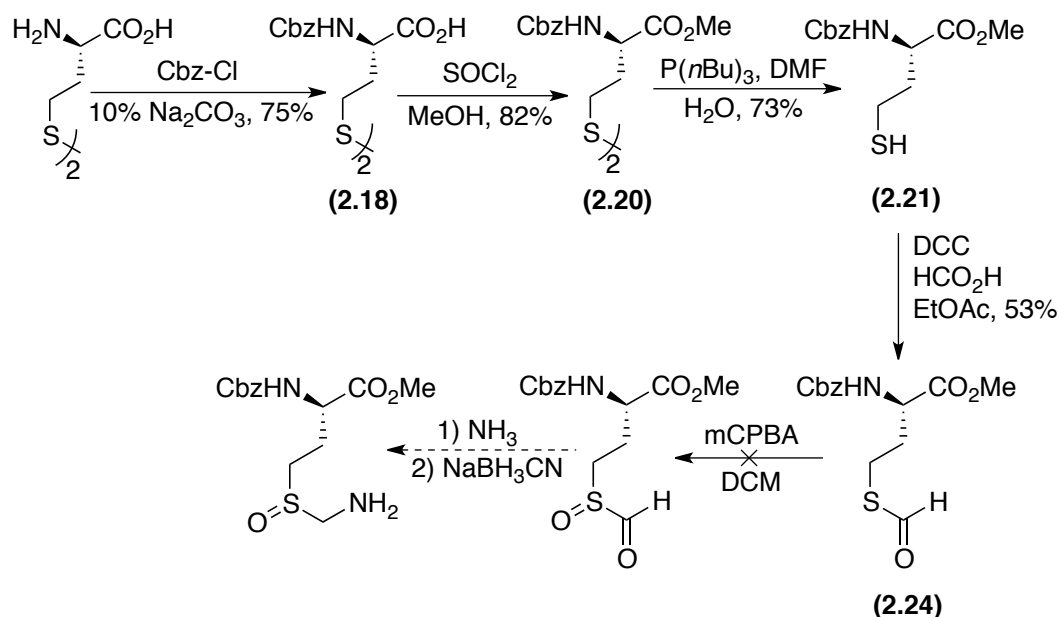
Table 2.2: Deprotection conditions for Boc removal.

Milder deprotection conditions were attempted (Table 2.2), including phosphoric acid in THF⁴⁵ and TBAF in THF.⁴⁶ TMSOTf in conjunction with triethylamine

is used in solid phase synthesis to deprotect the Boc group from lysine without affecting the linkage of the peptide chain to resin;⁴⁷ however, all of these conditions led to product degradation, as identified through the NMR spectrums of the reaction mixtures.

2.6.5 Synthesis of aminomethyl sulfoxide via thiol formate

In a second attempt to synthesize aminomethyl sulfoxide 2.10, we decided to take a different route while using the same protection scheme (Scheme 2.6).



Scheme 2.6: Synthesis of aminomethyl sulfoxide via thiol formate.

After obtaining the homocysteine thiol, it was coupled with formic acid using DCC to make a formate thiolester.⁴⁸ We then attempted to oxidize the sulfur to form a sulfoxide at which point the formate could be subjected to reductive amination conditions to yield the desired amino group, but the oxidation was unsuccessful.

2.7 Protolytic cleavage and instability of sulfinamides

In summary, despite the variety of methods and syntheses attempted, we were unable to synthesize either the sulfinamide (2.9) or the aminomethyl sulfoxide (2.10) transition state analogues. Unfortunately, the desired compounds proved unamenable to isolation due to their instability as highly functionalized molecules. The instability of sulfinamide compounds has been remarked upon in the literature and supports our hypothesis.

Previously in 2007, Karuso and coworkers reported that sulfinamides are susceptible to protolytic cleavage. The acid-catalyzed hydrolysis of sulfinamides results in the cleavage of the S-N bond and loss of the amine residue as a leaving group (Figure 2.16).⁴⁹

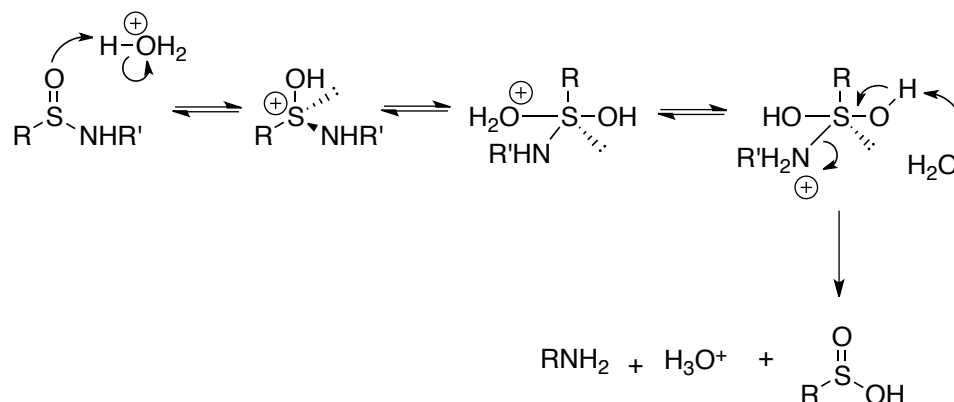


Figure 2.16: Two-step addition-elimination mechanism of sulfinamide hydrolysis

The authors found that sulfinamides containing a secondary as opposed to a tertiary nitrogen atom had a slower rate of hydrolysis and this was hypothesized to be a result of stabilization of intermolecular hydrogen bonding. The large dipole moment of the S-O group leads to the formation of strong hydrogen bonds that could stabilize the molecule and thus lead to a slower hydrolysis rate. The acid-catalyzed hydrolysis of

sulfinamides is said to occur through a two-step addition elimination mechanism (Figure 2.16) though the initial site of protonation and the rate limiting step of the reaction are still unclear and are the subject of contention.⁴⁹ While sulfinamides with primary nitrogen atoms were not expressly studied in this paper, the results can be extrapolated to our primary amine sulfinamide compounds, supporting the facile decomposition of our sulfinamide molecules.

The susceptibility of sulfinamides to acidolytic cleavage would explain the degradation of the aminomethyl sulfoxide analogue during the Boc deprotection. Though it is not a sulfinamide, the protonation of the sulfinyl group would make it more susceptible to nucleophilic attack, possibly by a liberated amine that could then lead to decomposition.

The instability of these sulfinamide molecules and thus the difficulty of isolation have been mentioned in the literature. In their attempt at synthesizing peptide sulfinamide analogues, Merricks and coworkers had synthesized *tert*-butyl phenylalaninate sulfinamide compound (Figure 2.17) whereby the amino group was protected using a phthalimide protecting group.²⁶

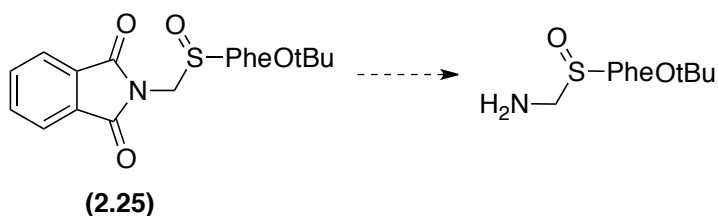


Figure 2.17: Attempted deprotection of *tert*-butyl phenylalaninate sulfinamide

They were unsuccessful in their attempt to remove the phthalimide group and postulated that the desired α -aminosulfinamide was unstable in the mildly acidic

conditions and thus, unable to be isolated. Similarly, Moree and coworkers attempted the synthesis of sulfinamide-containing peptides but were unable to extend the length of the peptide chain, and attributed this to the limited stability of the sulfinamide moiety.⁴² These outcomes parallel our experiences, underlining the difficulty in isolating functionalized, alkyl, primary amine sulfinamide compounds.

Lastly, a recent publication has put forth a mechanism for peptidic sulfinamide reduction through a cyclic intermediate (Figure 2.18).⁵⁰

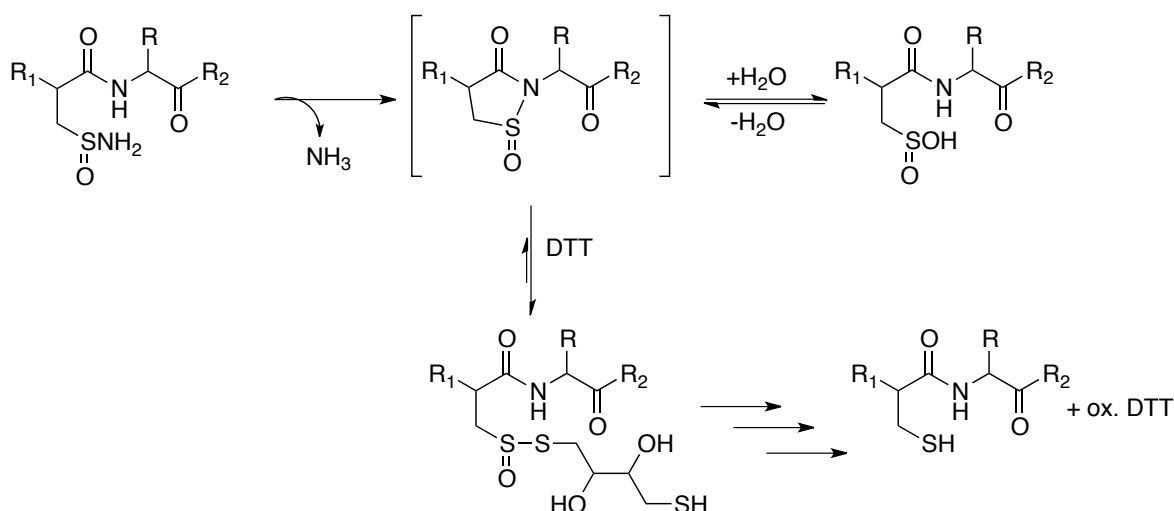


Figure 2.18: Peptidic sulfinamide reduction via a cyclic intermediate.

In the presence of a reducing agent, a disulfide was formed but in the absence of reducing agent, the sulfinamide was hydrolysed to a sulfinic acid, through the involvement of the peptide backbone amide group. This highlights the instability of these molecules, especially peptidic sulfinamides. As our target molecules were highly functionalized peptidic sulfinamides, containing amide bonds, this probably contributed to their intrinsic instability and tendency to decomposition. It can be hypothesized that the peptide backbone of compounds 2.9 and 2.10 participated in

destabilizing the sulfoxide/sulfinamide moiety leading to the degradation of the molecule, especially in the acidic and basic conditions used in the deprotection steps.

2.8 Conclusions and Perspectives

The electronic and spatial similarities between the transition states of the TG2 acylation reactions and compounds 2.9 and 2.10 suggest they would be effective transition state analogues. However, their synthesis proved too difficult and ultimately, we were not able to isolate the desired compounds due to their aforementioned instability. While novel intermediate peptidic sulfoxide compounds were synthesized, our overall targets were not obtained. Though we did expect the synthesis to be difficult, the degree of instability of these sulfinamide compounds was unanticipated.

However, an alternative isostere for the amine group is an alcohol and so it would be interesting to synthesize and test sulfinic acid derivatives 2.26 and 2.27 (Figure 2.19) in terms of their potential to inhibit TG2.

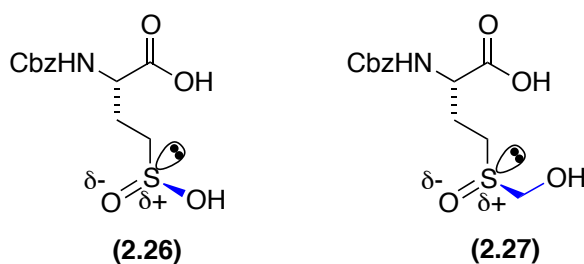


Figure 2.19: Sulfinic acid derivatives as TS analogues

While not as similar to the transition state as a sulfinamide, these compounds should be more stable and less susceptible to hydrolysis or degradation. Thus, this

remains an intriguing avenue to explore, given the utility and efficacy of transition state analogues.

References

- (1) Page, M. I.; Jencks, W. P. *Proceedings of the National Academy of Sciences* **1971**, *68*, 1678–1683.
- (2) Wolfenden, R. *Annual Review of Biophysics and Bioengineering* **1976**, *5*, 271–306.
- (3) Pauling, L. *American Scientist* **1948**, *36*, 50–58.
- (4) Leinhard, G. E. *Science* **1973**, *180*, 149–154.
- (5) Schramm, V. L. *Current Opinion in Structural Biology* **2005**, *15*, 604–613.
- (6) Bruice, T. C. *Accounts of Chemical Research* **2002**, *35*, 139–148.
- (7) Benkovic, S. J. *Science* **2003**, *301*, 1196–1202.
- (8) Mao, C.; Cook, W. J.; Zhou, M.; Federov, A. A.; Almo, S. C.; Ealick, S. E. *Biochemistry* **1998**, *37*, 7135–7146.
- (9) Schramm, V. L. *Biochimica et Biophysica Acta* **2002**, *1587*, 107–117.
- (10) Singh, V.; Evans, G. B.; Lenz, D.; Mason, J. M.; Clinch, K.; Mee, S.; Painter, G. F.; Tyler, P. C.; Furneaux, R. H.; Lee, J. E.; Howell, P. L.; Schramm, V. L. *Journal of Biological Chemistry* **2005**, *280*, 18265–18273.
- (11) Rivard, C. Études cinétiques par HPLC de l'étape d'acylation de la gamma-glutamyl transpeptidase purifiée des reins de rats à l'aide des substrats analogues, Université de Montréal, 2003.
- (12) Lherbet, C.; Gravel, C.; Keillor, J. W. *Bioorganic & Medicinal Chemistry Letters*

- 2004**, *14*, 3451–3455.
- (13) Piper, J. L.; Gray, G. M.; Khosla, C. *Biochemistry* **2002**, *41*, 386–393.
 - (14) Norlund, M. A.; Lee, J. M.; Zainelli, G. M.; Muma, N. A. *Brain Research* **1999**, *851*, 154–163.
 - (15) Dedeoglu, A.; Kubilus, J. K.; Jeitner, T.; Matson, S. A.; Bogdanov, M.; Kowall, N. W.; Matson, W. R.; Cooper, A. J. L.; Ratan, R. R.; Beal, M. F.; Hersch, S. M.; Ferrante, R. J. *The Journal of Neuroscience* **2002**, *22*, 8942–8950.
 - (16) Marrano, C.; de Macédo, P.; Gagnon, P.; Lapierre, D.; Gravel, C.; Keillor, J. W. *Bioorganic & Medicinal Chemistry* **2001**, *9*, 3231–3241.
 - (17) Choi, K.; Siegel, M.; Piper, J.; Yuan, L.; Cho, E.; Strnad, P.; Omary, B.; Rich, K.; Khosla, C. *Chemistry & Biology* **2005**, *12*, 469–475.
 - (18) Marrano, C.; de Macédo, P.; Keillor, J. W. *Bioorganic & Medicinal Chemistry Letters* **2001**, *9*, 1923–1928.
 - (19) Klöck, C.; Jin, X.; Choi, K.; Khosla, C.; Madrid, P. B.; Spencer, A.; Raimundo, B. C.; Boardman, P.; Lanza, G.; Griffin, J. H. *Bioorganic & Medicinal Chemistry Letters* **2011**, *21*, 2692–2696.
 - (20) Pardin, C.; Roy, I.; Lubell, W. D.; Keillor, J. W. *Chemical Biology & Drug Design* **2008**, *72*, 189–196.
 - (21) Case, A.; Stein, R. L. *Biochemistry* **2007**, *46*, 1106–1115.
 - (22) Hausch, F.; Halttunen, T.; Mäki, M.; Khosla, C. *Chemistry & Biology* **2003**, *10*, 225–231.
 - (23) Sohn, J. *Journal of Clinical Investigation* **2003**, *111*, 121–128.
 - (24) Sugimura, Y.; Hosono, M.; Wada, F.; Tohru, Y.; Maki, M.; Hitomi, K. *The Journal of*

- Biological Chemistry* **2006**, 281, 17699–17706.
- (25) Chabot, N.; Moreau, S.; Mulani, A.; Moreau, P.; Keillor, J. W. *Chemistry & Biology* **2010**, 17, 1143–1150.
- (26) Merricks, D.; Sammes, P. G.; Walker, E. R. H.; Henrick, K.; McPartlin, M. M. *Journal of the Chemical Society, Perkin Transactions 1* **1991**, 2169–2176.
- (27) Heimer, N. E.; Field, L. *Journal of Organic Chemistry* **2013**, 35, 3012–3022.
- (28) Bayfield, R. F.; Cole, E. R. *Analytica Chimica Acta* **1966**, 34, 193–200.
- (29) Clarke, V.; Cole, E. R. *Phosphorus, Sulfur, and Silicon and the Related Elements* **1994**, 92, 45–50.
- (30) Hilpert, K.; Ackermann, J.; Banner, D. W.; Gast, A.; Gubernator, K.; Hadvary, P.; Labler, L.; Muller, K.; Schmid, G.; Tschopp, T. B.; van de Waterbeemd, H. *Journal of Medicinal Chemistry* **1994**, 37, 3889–3901.
- (31) Radkiewicz, J. L.; McAllister, M. A.; Goldstein, E.; Houk, K. N. *Journal of Organic Chemistry* **1998**, 63, 1419–1428.
- (32) Gilmore, W. F.; Lin, H.-J. *Journal of Organic Chemistry* **1978**, 43, 4535–4537.
- (33) Cubbage, J. W.; Guo, Y.; McCulla, R. D.; Jenks, W. S. *Journal of Organic Chemistry* **2001**, 66, 8722–8736.
- (34) Brownbridge, P.; Jowett, I. C. *Synthesis* **1988**, 252–254.
- (35) Davis, F. A.; Zhang, Y.; Andemichael, Y.; Fang, T.; Fanelli, D. L.; Zhang, H. *Journal of Organic Chemistry* **1999**, 64, 1403–1406.
- (36) Wakayama, M.; Ellman, J. A. *Journal of Organic Chemistry* **2009**, 74, 2646–2650.
- (37) Davis, F. A.; Friedman, A. J.; Kluger, E. W.; Skibo, E. B.; Fretz, E. R.; Milicia, A. P.; LeMasters, W. C. *Journal of Organic Chemistry* **1977**, 42, 967–972.

- (38) García Ruano, J. L.; Alemán, J.; Fajardo, C.; Parra, A. *Organic Letters* **2005**, *7*, 5493–5496.
- (39) Chica, R. A.; Gagnon, P.; Keillor, J. W.; Pelletier, J. N. *Protein Science* **2004**, *13*, 979–991.
- (40) Hammond, G. S. *Journal of the American Chemical Society* **1953**, *77*, 334–338.
- (41) Butz, L. W.; Vigneaud, du, V. *Journal of Biological Chemistry* **1932**, *99*, 135–142.
- (42) Moree, W. J.; van Gent, L. C.; van der Marei, G. A.; Liskamp, R. M. J. *Tetrahedron* **1993**, *49*, 1133–1150.
- (43) Tsuzuki, Y.; Chiba, K.; Mizuno, K.; Tomita, K.; Suzuki, K. *Tetrahedron: Asymmetry* **2001**, *12*, 2989–2997.
- (44) Fustero, S.; Jiménez, D.; Moscardó, J.; Catalán, S.; del Pozo, C. *Organic Letters* **2007**, *9*, 5283–5286.
- (45) Li, B.; Bemish, R.; Buzon, R. A.; Chiu, C. K. F.; Colgan, S. T.; Kissel, W.; Le, T.; Leeman, K. R.; Newell, L.; Roth, J. *Tetrahedron Letters* **2003**, *44*, 8113–8115.
- (46) Routier, S.; Saugé, L.; Ayerbe, N.; Coudert, G.; Mérour, J.-Y. *Tetrahedron Letters* **2002**, *43*, 589–591.
- (47) Lejeune, V.; Martinez, J.; Cavelier, F. *Tetrahedron Letters* **2003**, *44*, 4757–4759.
- (48) Sprecher, M.; Nov, E. *Synthetic Communications* **1992**, *22*, 2949–2954.
- (49) Piggott, A. M.; Karuso, P. *Tetrahedron Letters* **2007**, *48*, 7452–7455.
- (50) Keceli, G.; Toscano, J. P. *Biochemistry* **2012**, *51*, 4206–4216.

CHAPTER 3

Cell Permeable, Fluorescent TG2 Probe

Introduction

3.1 The role of TG2 inside cells

The transamidating activity of tissue transglutaminase has been widely studied as a function of its extracellular role. TG2 has been implicated in cell adhesion and matrix assembly, cellular differentiation and matrix stabilization.^{1,2} In all of these functions, the primary role of TG2 is in the formation of protein crosslinks, supported by the high extracellular Ca^{2+} concentration.

However, the majority of TG2 is found inside the cell, in the cytoplasm.³ It is presumed that inside the cell, where the concentrations of Ca^{2+} are normally low, the transamidating activity of TG2 remains dormant and the protein functions as a GTP-binding protein.⁴ GTP is a potent inhibitor of TG2 transamidating activity at low Ca^{2+} concentrations.⁵ The combination of effective GTP inhibition at suboptimal (and physiological) Ca^{2+} concentrations is presumed to limit the intracellular transamidating activity of TG2, in so far as *in vitro* data can be extrapolated to *in vivo* conditions.⁶

Evidence has been presented to indicate that TG2 plays a role in signalling events inside the cell as a GTP-binding protein. It has been suggested that TG2 acts as a scaffolding protein in EGF-stimulated cancer cell migration whereby it binds to actin and acts as a scaffold for the recruitment of proteins that influence actin polymerization.⁷ This association is not dependent on transamidating activity. Similarly, transamidating activity is not required for the interaction of TG2 with the hypoxia-inducible factor 1B (HIF 1B), leading to the attenuation of the HIF1 hypoxic response signalling pathway.⁸ Thus, the principal role of intracellular TG2 would imply that its transamidating function is latent and it functions primarily as a scaffold protein.

However, recent publications have reported that intracellular TG2 transamidating activity occurs and is associated to a protective survival function, especially in regards to drug-resistant, metastatic cells. A fairly recent review underlining the transamidating activity of TG2 in cancer cells has demonstrated that cancer cells and cancer cell lines selected for resistance against chemotherapeutic drugs express elevated levels of TG2.⁹ TG2 expression in cancer cells leads to constitutive activation of the focal adhesion kinase, a key mediator in cell-survival signalling pathways.¹⁰ It has also been shown that TG2 is overexpressed in epithelial ovarian cancer and promotes intraperitoneal metastasis.¹¹ However, the exact mechanisms of these associations with TG2 are unclear. TG2 transamidating activity has also been implicated in cystic fibrosis.¹² It is involved in modulating inflammation through functional sequestration of the anti-inflammatory peroxisome proliferator-activated receptor γ in cystic fibrosis.¹² Lastly, Johnson and coworkers have also shown that cytosolic TG2 promotes cell death as a result of its transamidation activity under extreme stress conditions.¹³

Thus, conflicting roles have been presented in regards to the intracellular role of TG2. Physiologically, it would suggest that the cross-linking activity of TG2 would be inhibited by the low concentrations of intracellular calcium but this has not been conclusively proven. It cannot be ruled out that TG2 might interact with other cellular proteins that would result in it acting as a transglutaminase as opposed to a G protein. To address this lack of knowledge about the intracellular role of TG2, it is imperative to develop a cell permeable probe that will detect TG2 transamidating activity, through a reaction with its active site cysteine residue. Therefore, in the hopes of investigating

the intracellular cross-linking activity of TG2 to shed light on its varied roles, our group is interested in synthesizing fluorescent, cell permeable, TG2-specific probes.

3.2 Design of TG2 Probes

The design of our probe to investigate the role of TG2 activity inside the cell will focus on four aspects. Firstly, the probe will need to incorporate a fluorophore for ease of detection. Secondly, the probe should contain an electrophilic warhead so that the reaction with TG2 will lead to covalent labelling. Thirdly, the probe needs to be specific for TG2 to avoid non-specific labelling or reactivity. Fourthly, cell permeability of the probe is necessary to have an accurate measure of intracellular TG2 activity.

3.2.1 Fluorophore choice

By attaching a fluorescent molecule to our probe, this will allow us to visually track the molecule using confocal microscopy. Previous fluorophores used in the group attached to TG2 inhibitors included coumarin, rhodamine, and the dansyl fluorophore.¹⁴⁻¹⁶ Dansyl was chosen for this application both for its ease of coupling along with the fact that it is possible to have indirect detection of the dansyl group with anti-dansyl antibodies. This dual feature of the dansyl fluorophore has been used by the group of Mari Kaartinen to study the role of TG2 and FXIII in osteoblast differentiation.¹⁷

3.2.2 Electrophilic warhead

A covalent linkage between the probe and TG2 is necessary to ensure site specific labelling of the enzyme by the inhibitor and to ensure the inhibitor is not washed out. Previously in the Keillor group, inhibitors bearing various electrophilic warheads were tested against guinea pig liver TGase, chosen for its high homology (80%) with human TGase, validating its use as an animal model. The electrophilic pharmacophores tested for reactivity with the active site cysteine thiol of TG2 included α,β -unsaturated amides (3.1), epoxides (3.2), chloromethyl ketones (3.3) and maleimides (3.4) (Figure 3.1).^{18,19} These were attached to the side chain of Cbz-Gln-Gly, a known substrate of TG2, or a Cbz-Phe scaffold, as shown in Figure 3.1.

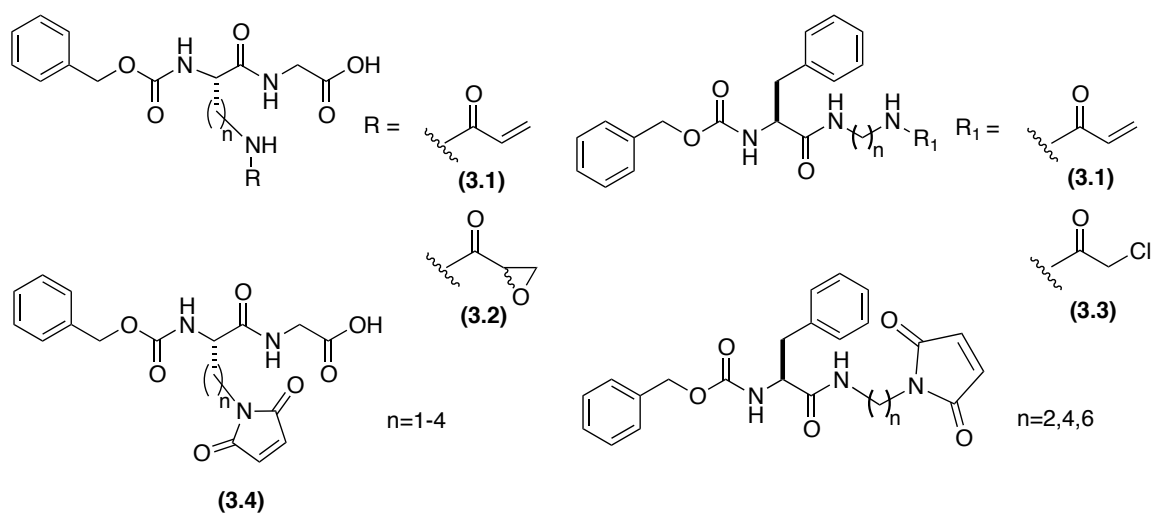


Figure 3.1: Various electrophilic warheads of TG2.

The side chain or spacer length of the analogues was varied between one to six methylene units in order to study the effect of chain length. It was found that the acrylic amide warhead bearing four methylene units in the side chain was the most efficient irreversible inhibitor, with a K_i value of 150 nM.

The greater reactivity of the longer methylene chain could be explained by the hypothesis that longer acyl donor substrates possess the conformational flexibility required to properly position the pendant reactive groups near the active site thiol. It has been noted that TGases are designed to exclude asparagine residues as acyl-donor substrates, reacting specifically with glutamine. This specificity could involve binding of the peptide backbone of the potential acyl-donor substrate at a site sufficiently distant from the site of acylation such that the asparagine β -carboxamide group is unable to acylate the active site thiol residue.¹⁹ However, from these studies, it appears that guinea pig liver tissue TGase is not designed to exclude the longer acyl donor that has more conformational flexibility.

The finding that the acrylic amide with four methylene chains was an effective warhead simplified the synthesis of our probe in that the commercially available lysine amino acid could be modified to contain an acrylic amide. This electrophilic warhead reacts with the active site cysteine through a Michael addition and ensures covalent labelling of TG2 (Figure 3.2).

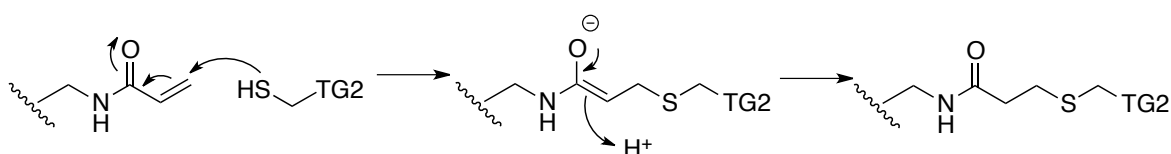


Figure 3.2: Michael addition of acrylic amide to TG2.

3.2.3 TG2 affinity

By including a reactive electrophile, we must assure that the probe is selective for TG2. Embedding the electrophilic acrylic amide in a peptide sequence specific to TG2, as discussed in the previous chapter, will confer TG2 affinity. Ac-PNPQLPF is a

high-affinity sequence for TG2 but the truncated sequence consisting of LPF-OH has also demonstrated selective affinity for TG2 and was used in this study.¹⁵

3.2.4 Cell permeable peptides

Lastly, the probe must be able to enter the cell and this will be realized by linking the acrylic amide warhead in the TG2 affinity sequence to a cell permeable peptide. As the name implies, cell permeable peptides are amino acid chains of varying length that are capable of traversing the cell membrane. This category of peptides includes the HIV-Tat peptide, penetratin, as well as polyarginine sequences, all of which will be discussed below. A wide range of biomolecules such as antigenic peptides,²⁰ peptide nucleic acids,²¹ antisense oligonucleotides,²² full-length proteins,²³ even nanoparticles²⁴ and liposomes²⁵ have been delivered through conjugation with a cell permeable peptide. Otherwise, these biomolecules would have been poorly taken up in cells, if at all, demonstrating the widespread interest in this class of peptides.

3.2.4.1 HIV-Tat Peptide

It was first discovered in 1988 by Frankel and Pabo that the transcription-transactivating (Tat) protein, involved in the replication of human immunodeficiency virus type 1 (HIV-1), could enter cells and translocate into the nucleus.²⁶ By synthesizing and testing various peptidic sequences of the Tat domain from residues 37 to 60 (Figure 3.3), the basic domain of the protein that is responsible for translocation, it was found that the sequence Tat(48-60) was the essential component of the protein for translocation.²⁷

Tat ₄₈₋₆₀	C(Acm)GRKKRRQRRRPPQC	
Tat ₄₃₋₆₀	C(Acm)LGISYGRKKRRQRRRPPQC	
Tat ₃₇₋₆₀	C(Acm)FITKALGISYGRKKRRQRRRPPQC	C(Acm) = Cys-acetimidomethyl
Tat ₃₇₋₅₃	C(Acm)FITKALHISYGRKKRC	

Figure 3.3: Tat peptide sequences evaluated for permeability

The optimal sequence was thus found to be GRKKRRQRRRPP and deletion or substitution of any of the basic charges led to reduced permeability, highlighting the importance of these amino acids.

3.2.4.2 Penetratin

Through experiments conducted with homeoproteins, belonging to the class of transactivating factors first discovered in *Drosophila*, it was observed that the homeodomain Antennapedia was capable of translocating across neuronal membranes and was transported to the nucleus.²⁸ By investigating various peptides taken from the C-terminal third helix of the Antennapedia homeodomain, identified through mutagenesis as crucial to translocation, a 16 amino acid sequence was shown to reproduce the cell permeable behaviour of the entire homeodomain.²⁹ Shorter and/or modified versions were poorly or not at all internalized and this sequence was named penetratin consisting of RQIKIYFQNRRMKWKK. As noted for the Tat peptide, penetratin also contains multiple basic amino acids and interestingly the tryptophan residue was reported to be essential for carrier ability.²⁹

3.2.4.3 Polyarginine sequences

Rothbard *et al.* suggested the importance of the arginine/guanidino moiety, as they found that charge alone was insufficient for cellular uptake.³⁰ They examined the translocation ability of nonamers of arginine, histidine, lysine and ornithine peptides in Jurkat cells and found that after 5 min, all were less effective than the Tat peptide, excluding the nonamer of the arginine peptide that was 20-fold more efficient. This observation led to the implication that guanidinium plays a greater role in facilitating cellular uptake than the charge or backbone structure.

Futaki and coworkers conducted an investigation of linear arginine chains, and found that while a tetramer of arginine residues showed extremely low translocation activity, hexa-arginine and octa-arginine polymers exhibited the maximum internalization and accumulation in the nucleus.³¹ As the chain length further increased, the degree of internalization decreased, suggesting the optimal number of arginine residues is between 6-8. In terms of toxicity, when using a peptide with less than nine arginine residues, HeLa cells had a greater than 95% viability, comparable to the viability observed in similar tests with the Tat peptide.

Branched chains were equally effective, whereby $(R_2)_4$ (3.5), had a charge-dependent manner of internalization similar to the linear R_8 (3.6) peptide (Figure 3.4).³² Though a subtle difference was observed in internalization at 4 °C, the peptides were thought to share a significant part of the internalization pathways and cellular localization sites. A different branched peptide $(RG_3R)_4$ (3.7) (Figure 3.4), did not have a continuous array of arginine residues and so the “positive” patch formed by the residues would be larger and the density of positive charges would be lower than in the

case of the linear (3.7) or (R₂)₄ peptide (3.5) but it was also efficiently delivered. This is suggestive of flexibility in the location of arginine residues in carrier molecules, and leads to the implication that for translocation, linear or continuous arrangement of the arginine residues is not necessary.

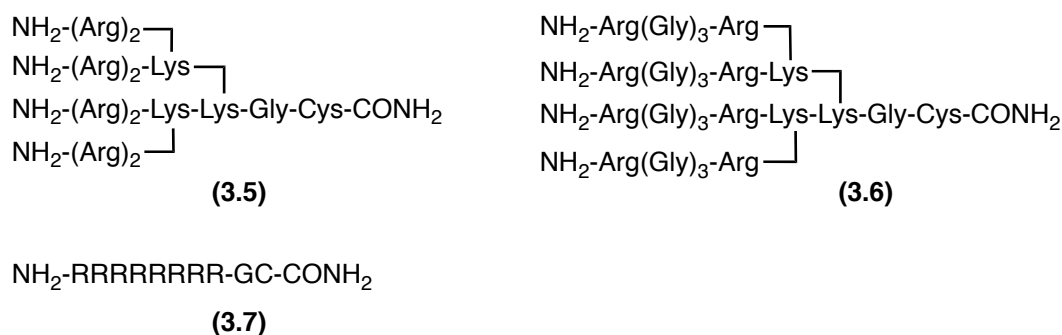


Figure 3.4: Linear vs. branched chains of octa-arginine peptides.

3.2.4.4 Mode of entry

The mechanism of internalization of cell permeable peptides and their cargo has been the subject of controversy. Historically, the Tat peptide had rapid internalization within several minutes and a high extent of nuclear accumulation. Through direct labelling of the peptide with fluorescein or indirect immunofluorescence using a monoclonal antibody against the Tat cluster, it was shown that the Tat peptide was taken up by cells in less than 5 min at concentrations as low as 100 nM.²⁷ It had been found that entry into the cells at 4 °C had almost the same efficiency as at 37 °C and there was low sensitivity to typical clathrin endocytosis inhibitors. Endocytosis is inhibited at low temperatures and so these dual observations led to the assumption that the internalization mechanism of the Tat peptide was separate from endocytosis.²⁷ In addition, structure-activity studies have shown that permeability does not depend

on a specific primary sequence, which implies an independence of receptor recognition.³³ As the polyarginine sequences had the same responses as the Tat peptide, they were also assumed to have an internalization mechanism separate from endocytosis or receptor binding.³¹

More recently, it has been found that the fixation process prior to microscopic observations affects the cellular localization of internalized peptides.³³ When unfixed Jurkat and HeLa cells were observed after incubation with fluorescein-labelled Tat peptide or a nona-arginine sequence, the majority of peptide-derived fluorescence was seen in punctate structures in the cytosol and the internalization was almost completely suppressed at 4 °C. Contrastingly, when the cells were fixed, the punctate structures were lost and the peptides were observed to be diffusely located throughout the cytosol with accumulation in the nucleus. After further deduction and analysis of cells treated at 4 °C, the researchers reasoned that the internalization of peptides observed for cells treated at 4 °C was caused by the cell surface-adsorbed peptide permeating into cells by fixation. To that end, endocytosis could not be ruled out as a mode of entry for the Tat and polyarginine sequences.

There are three types of endocytic pathways, namely clathrin-mediated endocytosis, lipid raft-mediated caveolae endocytosis and macropinocytosis. It was observed that a significant part of the Tat peptide was colocalized with transferrin, used as a marker of endocytosis. Since transferrin is internalized by clathrin-mediated endocytosis, the same mechanism was proposed for the Tat peptide.³³ However, Fittipaldi reported opposing data in the internalization of a fusion protein of Tat and EGFP, an enhanced variant of the green fluorescent protein. The transduction of this

fusion protein was inhibited by methyl-B-cyclodextrin and cytochalasin D, both of which are clathrin-mediated endocytosis inhibitors.³⁴ Additionally, endosomes containing Tat-EFGP were not colocalized with those containing transferrin and so they concluded that the fusion protein was internalized by lipid-raft mediated caveolae endocytosis.³⁴

A third type of endocytosis, macropinocytosis was put forth as the manner of internalization of the Tat-Cre fusion protein whereby acidification of the macropinosomes was critical for cytosolic release of the fusion protein.³⁵ Experiments with polyarginine sequences showed that uptake of an arginine octamer by HeLa cells was suppressed by a macropinocytosis inhibitor and it was deduced that macropinocytosis plays a crucial role in the internalization of peptides.³⁶

Finally, in 2007, Duchardt and coworkers provided evidence that the cell permeable peptides; nona-arginine, Tat and penetratin use all three endocytic pathways simultaneously.³⁷ However, the peptides differ by the extent to which the individual processes contribute to importation.³⁷ Peptide concentration is a key experimental condition when determining the pathway chosen while the cargo attached to the peptide can also play a role. Thus, the mode of entry into cells is both peptide and cargo dependent which can explain the differing results seen by different groups.

3.2.4.5 Various applications of cell permeable peptides

Cell permeable peptides have been used in a variety of ways to deliver numerous compounds into the cell, ranging from proteins and aptamers to small molecules. Some of the applications of cell permeable peptides will be outlined below.

The first example of the Tat peptide being used as a cargo transporter was carried out through fusing the 120-kD β -galactoside (β -gal) protein to the Tat peptide. Dowdy and coworkers demonstrated the ability to deliver a biologically active fusion protein to all tissues in mice, including the brain.³⁸

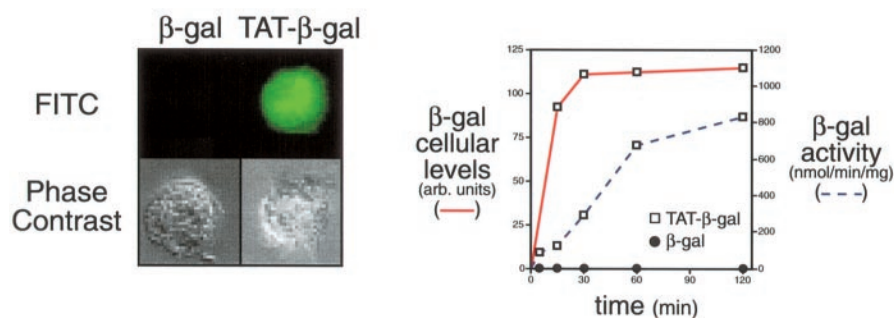


Figure 3.5: (left) Fluorescence and confocal microscopy of β -galactoside fused with Tat peptide and (right) concentration (left axis) and enzymatic activity (right axis) of cells treated with Tat- β -galactoside.³⁸ (Schwarze, S. R.; Ho, A.; Vocero-Akbani, A.; Dowdy, S.F.; *Science* **1999, 285, 1569–1572) Reproduced with permission.**

The work of Dowdy and coworkers is shown in Figure 3.5 where it can be seen that the fused Tat peptide- β -galactosidase protein is present throughout the Jurkat cells. The Tat fused protein was also shown to be active inside HepG2 cells by measuring its activity as a function of concentration.

Enzyme activity has also been modified using cell permeable peptides. Choi and coworkers fused a T-cell antigen to an arginine rich sequence which functioned as the protein transduction domain.³⁹ After transduction into T-cells, it led to an inhibited response in the development of the airway inflammatory responses in mice. Enzymatic activity has also been monitored using a contrast agent consisting of an enzymatically cleavable linker (galactose substrate) inserted between the Tat peptide and a Gd³⁺-based MRI reporter moiety (3.8)(Figure 3.4).⁴⁰

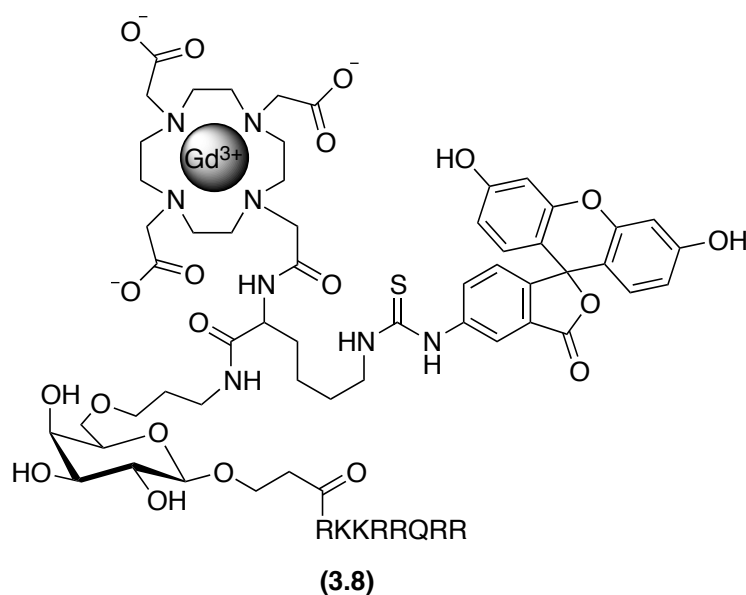


Figure 3.6: Dual-labelled probe made permeable through incorporation of Tat peptide.

Enzymatic activation of the contrast agent following cellular uptake would require enzymatic cleavage of the transporting peptide and then subsequent cellular entrapment of the MRI reporter component that would lead to signal amplification.

3.3 TG2-Specific, fluorescent cell permeable probes

Thus, our goal was to synthesize a fluorescent, TG2-specific, cell permeable probe. The widely used dansyl fluorophore was chosen for its ease of coupling and indirect detection through the use of anti-dansyl antibodies. The dansyl fluorophore will be incorporated into the peptidic sequence through its addition to a lysine ϵ -amino group. As the acrylic amide has previously been shown to be an effective inhibitor of TG2, the electrophilic α,β -unsaturated amide will also be coupled to a lysine residue to keep the flexibility required of the four methylene units. TG2 affinity will be conferred to the probe through the addition of the TG2-specific peptide; LPF-OH. A polyarginine sequence consisting of eight consecutive arginine residues was chosen for its cell permeable properties. The dansyl-modified lysine residue was separated from the acrylic amide-modified lysine residue by a glycine spacer. The polyarginine sequence would be attached to the dansyl-modified lysine residue on the N-terminal side of the glycine spacer while the C-terminal side would have the acrylic amide-modified lysine and the peptidic chain conferring affinity for TG2 (3.9) (Figure 3.7).

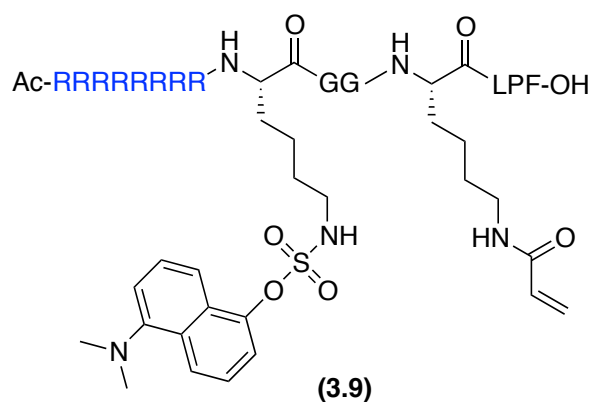


Figure 3.7: TG2-Specific, cell permeable, fluorescent probe.

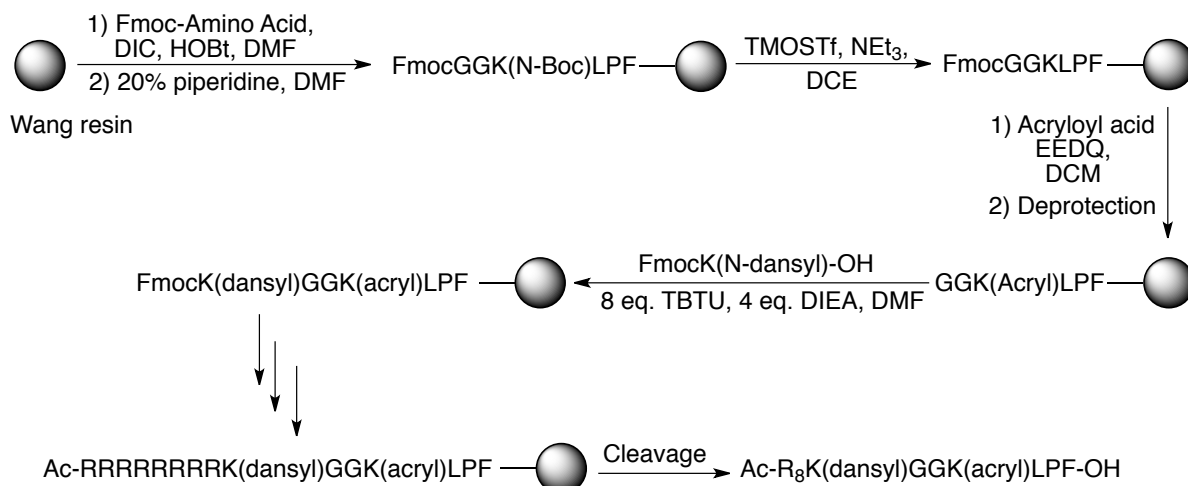
Results/Discussion

3.4 Synthesis of fluorescent, cell permeable probes

The synthesis of the TG2-specific, fluorescent, cell permeable probes were carried out through Fmoc-based solid phase synthesis. The fluorescent dansyl-modified lysine residue was synthesized separately, then incorporated into the peptidic chain while the acrylic amide-modified lysine residue was synthesized on resin.

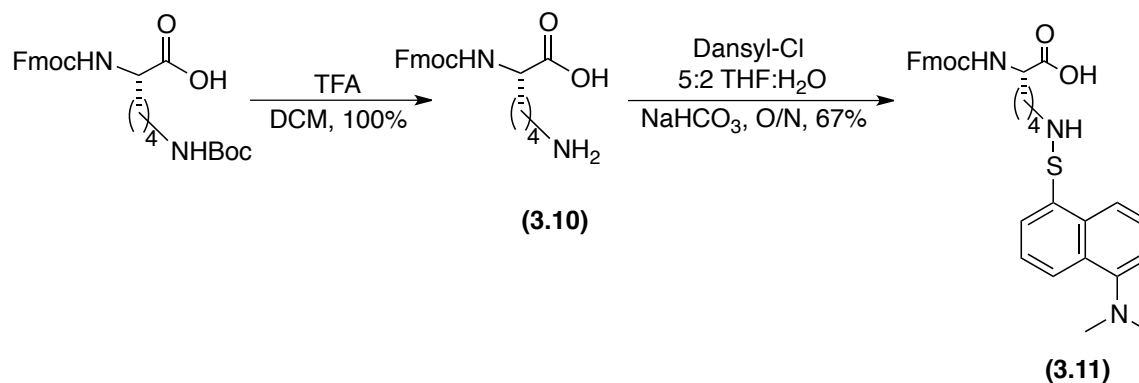
3.4.1 Synthesis of Ac-R₈K(dansyl)GGK(acryl)LPF-OH

The decision to incorporate 8 arginine residues was based on the polyarginine experiments by Futaki and coworkers.³¹ Standard Fmoc-based coupling conditions employing DIC and HOBt were undertaken to couple together the peptidic TG2 affinity sequence of LPF-OH, followed by a Boc-protected lysine and then a di-glycine spacer. The Boc protecting group on the lysine residue was removed using TMSOTf, conditions that were essential in that they removed the Boc protecting group without cleaving the peptide from the resin (Scheme 3.1). Formation of the acrylic amide was undertaken by coupling the free amine to acryloyl acid using EEDQ in the absence of base. Previous work in the group had demonstrated that these conditions led to excellent yields without degradation of the electrophilic acryloyl group.¹⁵ The procedure for Fmoc-based peptide coupling did not affect the integrity of the acryloyl group, allowing the peptide spacer sequence to be added sequentially. A di-glycine spacer was included between the lysine-dansyl residue and the inhibitor affinity sequence to ensure that the fluorophore would not decrease the affinity of the peptide sequence for TG2.



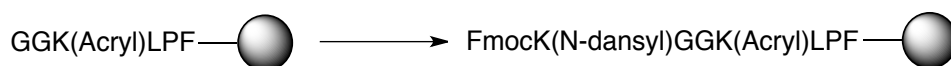
Scheme 3.1: Synthesis of Ac-R₈-K(N-dansyl)GGK(acryl)LPF-OH.

The dansyl-modified lysine residue was synthesized through deprotecting the Boc protecting group of commercially available Fmoc-K(N-Boc)-OH and then reacting the amino acid with dansyl chloride in a mixture of THF and water (Scheme 3.2).



Scheme 3.2: Synthesis of Fmoc-K(N-dansyl)-OH.

It was found that the standard coupling conditions (DIC, HOBT) were unsuccessful for the coupling of the lysine-dansyl residue to the peptide. Several reaction conditions were tried (Table 3.1) to effect this reaction.

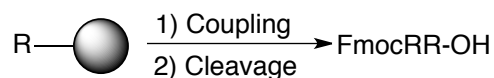


Coupling Reagent	Fmoc-K(N-dansyl)-OH	Activator	Result
3 eq. DIC	3 eq.	3 eq. HOBt	No reaction
3 eq. EEDQ	3 eq.	N/A	No reaction
8 eq. TBTU	4 eq.	4 eq. HOBt + 8 eq. DIEA	Coupling successful

Table 3.1: Reaction conditions for Fmoc-K(N-dansyl)-OH coupling.

The coupling results were verified by the Kaiser test and through verifying the fluorescence (or lack thereof) of a small aliquot of resin beads. Once fluorescence was observed, the coupling was verified using low resolution MS-ESI⁺. The successful conditions for the coupling of the Fmoc-K(N-dansyl)-OH to the rest of the peptide included 8 eq. of TBTU, 3 eq. of HOBt and 4 eq. of DIEA base.

As polyarginine sequences are known to aggregate and minimal yields were observed after the synthesis of the complete peptide, a variety of coupling agents were tested for the arginine coupling reaction to see which would give the highest coupling efficiency (Table 3.2).



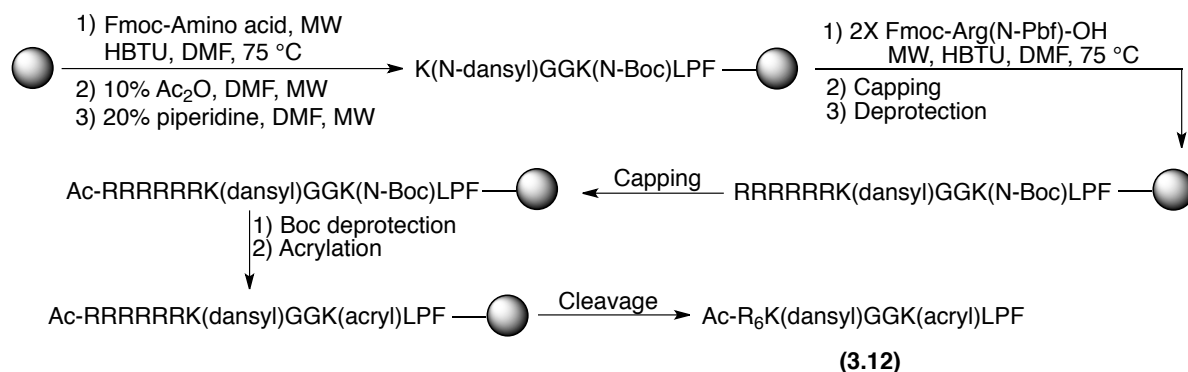
Coupling Reagent	Fmoc-R(Pbf)-OH	HOBt	Yield (%)
4 eq. TBTU	4 eq.	4 eq. + 8 eq. DIEA	20
3 eq HATU	3 eq.	3 eq.	13
3 eq. PyBOP	3 eq.	3 eq.	18
3 eq. EEDQ	3 eq.	N/A	27

Table 3.2: Evaluation of arginine coupling conditions.

The arginine residue was attached to the Wang resin after which various conditions were evaluated to couple the second arginine residue (Table 3.2). The di-arginine peptide was then cleaved using trifluoroacetic acid. The efficiencies of the different coupling conditions were evaluated through comparison of their LC traces. The optimal coupling agent was EEDQ as it had the highest yield of di-arginine product (Table 3.2). The glycine spacer and octa-arginine cell-penetrating sequence were then added using EEDQ as the coupling agent. Lastly, the peptide was cleaved using acidic treatment with trifluoroacetic acid. However, identification of the final product proved elusive, as both LC-MS ESI⁺ and MALDI were unable to detect the peptidic sequence.

3.4.2 Synthesis of Ac-R₆K(dansyl)GGK(acryl)LPF-OH

Futaki and coworkers had demonstrated that the difference in cell permeability and toxicity was minimal between 6-8 arginine residues,³¹ so we decided to reduce the number of arginine residues in our cell-penetrating sequence in the hope of reducing aggregation and perhaps increasing the yield. At this point in time, we had come into possession of a CEM Liberty automated microwave peptide synthesizer machine, which is known to increase the yields and purity of peptide synthesis. It uses microwave irradiation (MW) to increase yields and coupling efficiency with HBTU as an activator. Double coupling was carried out for the arginine residues, whereby the peptide was irradiated twice for thirty minutes at 75 °C (Scheme 3.3).



Scheme 3.3 Microwave synthesis of Ac-R₆K(dansyl)GGK(acryl)LPF-OH (3.12).

Once all the arginine residues had been added and the peptide was capped, Boc deprotection of the lysine residue was undertaken under our standard solid-phase reactor conditions, using TMOSTf and then acrylated using acrylic acid and EEDQ.

After completion of the synthesis, we discovered that the hexamer arginine sequence (3.12) was more soluble and also easily identifiable by ESI and MALDI. The

peptide was purified using reversed phase HPLC and obtained in good yield with 90% purity.

3.5 *In vitro* activity of Ac-R₆K(dansyl)GGK(acryl)LPF-OH (3.12) with recombinant human TG2

The cell permeable peptide was tested as an inhibitor for TG2 using the glutamate dehydrogenase (GDH) assay.⁴¹ Human TG2 was purified according to a procedure published in the Keillor group.⁴² The assay is based on the continuous monitoring of ammonia released during the TG2 reaction, measured by the GDH indicator reaction (Figure 3.7).

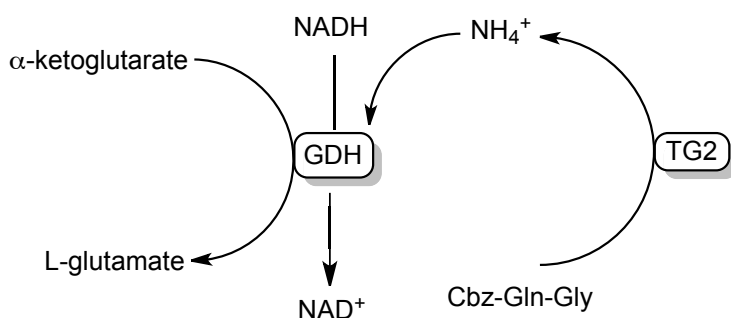


Figure 3.8: Schematic of the GDH assay with Cbz-Gln-Gly substrate

The ammonium produced by the TG2 enzymatic reaction is bound by GDH and is used to effect the reductive amination of α -ketoglutarate, producing L-glutamate, concurrently oxidizing β -NADH to NAD⁺. The disappearance of NADH is followed spectrophotometrically at 340 nm and is coupled to the activity of TG2. The donor substrate used in the assay was Cbz-Gln-Gly, which has been widely used in the literature while the acceptor substrate was *N*-acetyl lysine methyl ester.

Initially, the inhibitor 3.12 was incubated with the enzyme and its effect on TG2 activity was monitored as a function of inhibitor concentration. However, the cell permeable peptide behaved as a reversible inhibitor and was not showing the time-dependent inactivation of TG2 that was expected, had the active site cysteine been covalently bonded by the inhibitor. Thus, it seemed as if the peptide was acting more like a competitive inhibitor and affecting catalysis by blocking substrate binding without actually reacting with the enzyme. The data obtained was plotted as the reciprocal of the observed rate versus inhibitor concentration (Figure 3.9) and a linear trend line was calculated.

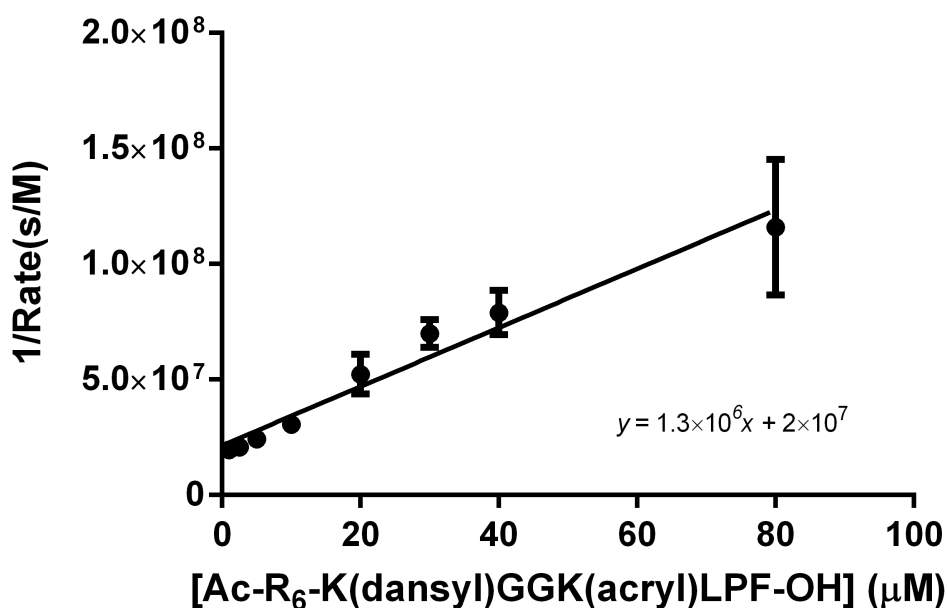


Figure 3.9: Reciprocal of observed rate vs. Ac-R₆-K(dansyl)GGK(acryl)LPF-OH concentration.

The trend line can be inputted into the Equation 3.1 where the x-intercept of the line is the IC₅₀ value of the cell permeable peptide, where the equation assumes that the

peptide is a competitive inhibitor with respect to the substrate. This IC_{50} value was then inputted into Equation 3.2 to obtain a K_i for the inhibitor, which was calculated to be 9 μ M.

Equation 3.1: $\frac{1}{rate} = m[I] + b$

Equation 3.2: $K_i = \frac{IC_{50}}{(1 + \frac{[S]}{K_m})}$

However, the inhibitor was designed to be irreversible with a Michael-type addition reaction (Figure 3.2) occurring between the electrophilic acrylic amide and the nucleophilic cysteine residue in the active site of TG2. To confirm this supposition, the inhibitor was incubated with TG2 in calcium-containing buffer in the absence of substrate for a period of 30 min. Three successive buffer changes over a centrifugal Amicon filter with a 30-kDa cutoff were then undertaken. After the last wash, the activity of the enzyme was evaluated using the substrate *N*-Cbz-L-Glu(γ -*p*-nitrophenyl ester)Gly (3.13), also known as AL5.⁴³ TG2 will cleave the activated ester to release the chromophore, *p*-nitrophenolate (3.15), whose absorbance is measured at 405 nm (Figure 3.10).

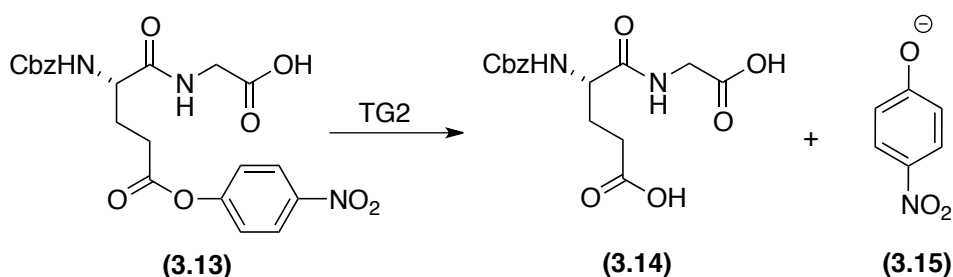


Figure 3.10: Activity (AL5) assay measuring release of *p*-nitrophenolate.

After incubation for 30 min, it was observed that greater and successive concentrations of inhibitor deactivated the enzyme faster, supporting the idea that the

peptide is an irreversible inhibitor of TG2. Unfortunately, a K_i could not be determined because analysis of the kinetic data (not shown) indicated we were not able to capture the initial rate of first order decay, but instead observed the rate plateau after enzyme inactivation. This could be due to an overly long incubation time coupled with the potential loss of TG2 activity due to the multiple washing steps required to prove inhibitor irreversibility. In order to test this hypothesis, the aforementioned experiment should be repeated using shorter incubation times in order to observe the initial decreased rate caused by the cell permeable peptide inhibition of TG2.

Thus, in the presence of TG2 substrate, the peptide was an effective and reversible inhibitor. However, in the absence of competing substrate, and with time, the peptide did bind irreversibly to the enzyme as intended.

3.6 Preliminary *in vivo* results with SH-SY5Y neuroblastoma cells

Once it had been confirmed that the peptide was an irreversible inhibitor, the cell permeability and efficiency of intracellular TG2 labelling by the inhibitor peptide was evaluated. To that end, SH-SY5Y neuroblastoma cells were grown to 70% confluence, after which, the cultures were split. Previous work done by the group of Johnson and coworkers found that treatment with retinoic acid (RA) stimulated TG2 expression and cellular differentiation.⁴⁴ Thus, half of the cells underwent RA treatment whilst the other half was grown under normal conditions. After treating the cells for 8 days, the cells were incubated for 30 min at several concentrations of cell permeable peptide (3.12) at 37 °C and under a CO₂ rich atmosphere. Three successive buffer changes were undertaken after which the cells were fixed using a formaldehyde

solution. The cells were then imaged under a confocal microscope under both blue light and with phase contrast. Images obtained after incubation in the presence of 10 μM of the cell permeable probe showed that 10 μM was the optimal concentration for fluorescence microscopy and so the other images will not be shown.

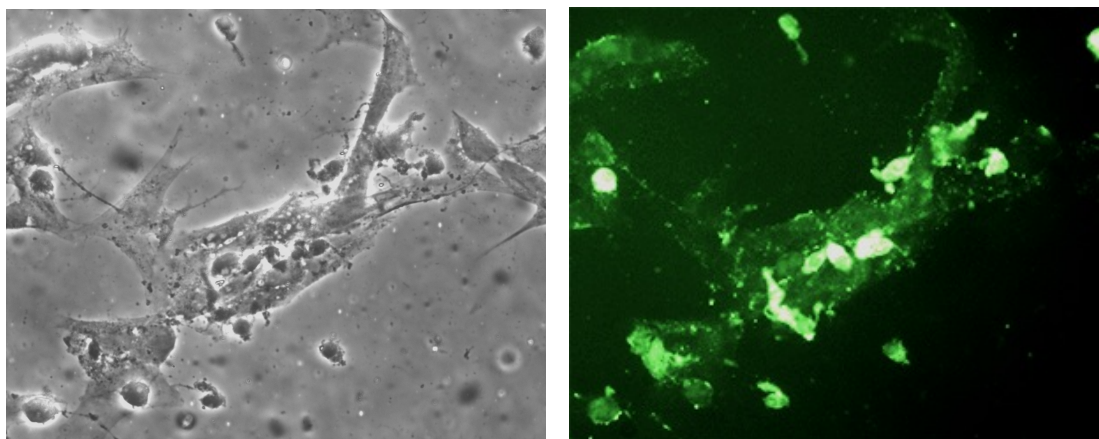


Figure 3.11: (left) Phase contrast of retinoic acid-treated cells with 10 μM cell permeable peptide (3.12), (right) under blue light.

Figure 3.11 shows the cells that were treated with RA, using phase contrast on the left and under blue light on the right. It can be seen that the cells are not robust and in the process of dying, which can be inferred from several observations, including the fact that the cells are clumped together. The cells that are spread out and long are those that have differentiated and are healthy but they are outnumbered by the circular cells, which appear shriveled and small. The inhibitor treatment could be the cause for the apparent cell death though in the cells that are healthy and differentiating, there is a good amount of fluorescence. The observation of fluorescence demonstrates that the probe has entered the cells and effectively labelled what we are presuming to be TG2, as the RA treatment has ensured that TG2 is active and functioning within the cells.

Barring the appropriate controls, it cannot be said with certainty that the probe is labelling TG2 though that was our assumption. Thus, from these images, our initial interpretation was that the probe functioned well, permeated the cell and labelled TG2.

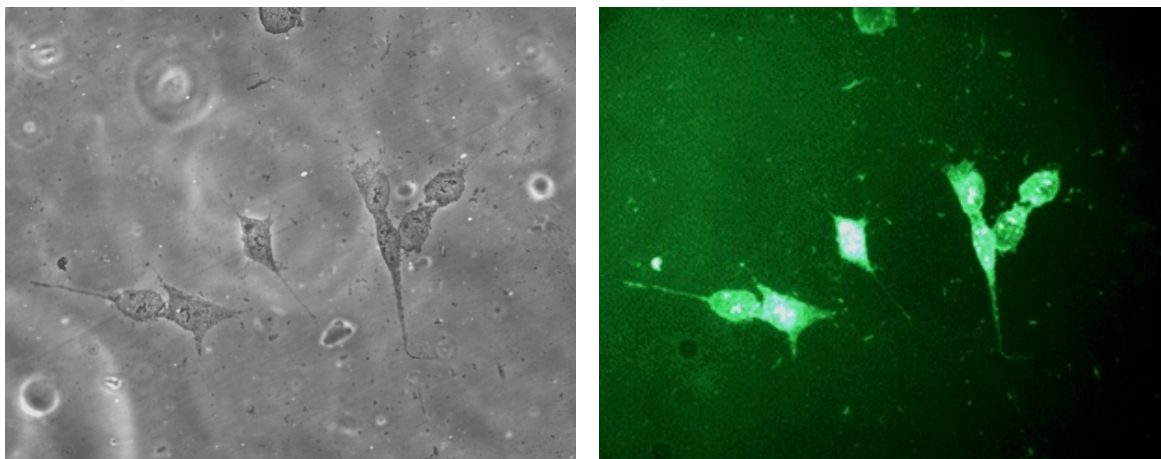


Figure 3.12: (left) Phase contrast of cells not treated with retinoic acid with 10 μ M cell permeable peptide (3.12), (right) under blue light.

Figure 3.12 shows the cells that were not treated with retinoic acid but just grown normally with phase contrast and under blue light. Again, the cells are not robust and are clumped together. The circular shape of the cells indicate those cells that are dead while the cells that are more spread out are the healthy cells. Unexpectedly, there is also fluorescence seen in the healthy cells, which is surprising since TG2 is expressed at very low levels in these cells. This observation suggests that the inhibitor is not only labelling TG2 but other proteins as well. Our probe does not seem to be as specific for TG2 as we had hoped.

From section 3.5 with our kinetic results, we did see that initially, the peptide behaves as a reversible inhibitor and only after incubation with TG2 in the absence of

substrate did the inhibition appear irreversible. We hypothesize that the inhibitor binds reversibly with TG2, leaving the electrophilic acrylic amide free to react and label other proteins, leading to the indiscriminate labelling observed in the cells without retinoic acid treatment. Thus, it seems as if adding the cell permeable tag has affected the affinity of the LPF sequence for TG2.

Our goal of creating a cell permeable peptide was achieved albeit at the expense of TG2 affinity and therefore specificity of labelling. However, it must be said that the peptide is cell permeable, as fluorescence was observed within the cells, but needs to be modified to maintain its affinity for TG2.

3.7 Modelling of the fluorescent probe with TG2

In an attempt to understand the binding mode of the peptide, molecular modelling of the cell permeable peptide with TG2 was undertaken using the MOE (Molecular Environmental Program) from the Chemical Computing Group (Montréal, Canada). A crystal structure of human TG2 crystallized in its open form with an irreversible peptidic inhibitor (PDB: 1KV3) was used for the investigation. The peptide was oriented and placed within the TG2 docking site. The top ranked structures obtained after scoring were retained and evaluated. It can be seen (Figure 3.13) that according to the top scored poses, the peptide has positioned itself with the hexa-arginines binding in a hydrophobic pocket. This pocket will hereby be referred to as the LPF pocket because it is the binding site on the enzyme of the LPF portion of the affinity sequence, PNPQLPF.

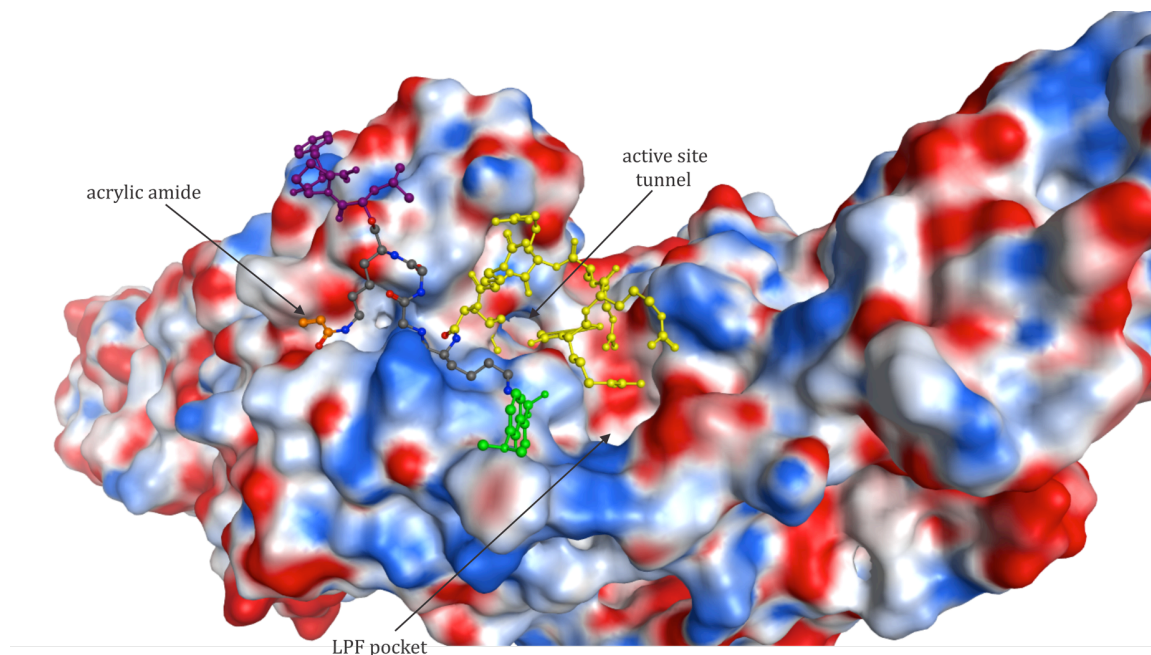


Figure 3.13: Cell permeable peptide docked into active site of TG2. Yellow = polyarginine sequence, green = dansyl fluorophore, orange = acrylic amide, purple = LPF sequence.

This binding mode results in the electrophilic acrylic amide moiety being far from the active site and is actually near to the surface of the protein. The peptide backbone of the inhibitor does block the active site but does so reversibly, explaining the initial results we obtained when we tested the peptide using the GDH assay, with Cbz-Gln-Gly as the substrate.

It was hypothesized that the LPF sequence of the cell permeable peptide would bind in its appropriate pocket, positioning the electrophilic acrylic amide in the active site tunnel. The di-glycine spacer and hexa-arginine sequence would project outwards from the protein surface. The di-glycine spacer was included to minimize interference

of the hexa-arginine and dansyl-modified lysine with the LPF binding but the modelling suggests that the peptide has reversed positions. It is binding in the opposite way we had intended, with the LPF sequence on the surface of the protein. The positively charged arginine residues are attracted to the negatively charged LPF pocket and thus the peptide is re-oriented, with the electrophilic warhead away from the active site cysteine and hovering, unencumbered on the surface of TG2.

Thus, the docking results support the notion of the indiscriminate labelling observed in the SH-SY5Y cells. When the peptide is positioned as is shown in Figure 3.13, where the acrylic amide is positioned near the surface of the protein, it will not react with TG2. Thus, the peptide can dissociate to react with different proteins present in the cell. This method of binding would explain the similar levels of extensive fluorescence seen in both trials of our fixed cells. Through adding the hexa-arginine sequence, we have rendered the probe permeable, though this has negatively affected the specificity of the LPF sequence by influencing its binding mode.

3.8 Conclusions and perspectives

In conclusion, Ac-R₆K(dansyl)GGK(acryl)LPF-OH was synthesized as a cell permeable, TG2-specific, fluorescent probe. *In vivo* cell imaging using SH-SY5Y neuroblastoma cells showed that while the probe was fluorescent and cell-permeable, it was not TG2-specific. Molecular modelling suggests that the probe is positioned in the opposite way intended, with the hexa-arginine sequence binding in the pocket intended for the TG2 affinity-conferring LPF sequence, thus orienting the electrophilic acrylic amide warhead far from the active site cysteine residue.

A re-design of the probe will be necessary in order for it to be effective against TG2. Futaki and coworkers have already shown that cell permeability is not affected with branched-chain polyarginine sequences and a linear arginine sequence is not imperative. To that end, a non-linear version of the cell permeable, TG2-specific probe (3.16) is proposed in Figure 3.14.

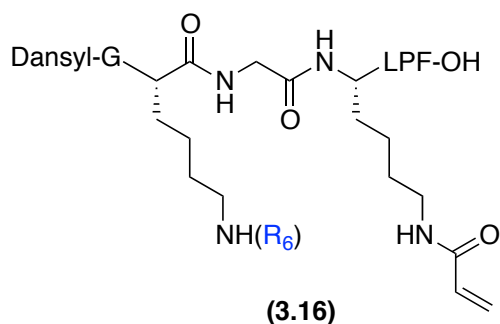


Figure 3.14: Non-linear, cell permeable, fluorescent TG2-specific probe.

By including the hexa-arginine sequence through a flexible lysine linker, this should minimize the interference of the LPF sequence binding to TG2. Thus, the affinity-conferring LPF moiety should be able to bind in its hydrophobic pocket, positioning the acrylic amide for reaction with the cysteine residue while maintaining permeability. The hexa-arginines, as opposed to binding in the LPF pocket, should, due to the methylene flexibility of the lysine linker, stick out from the protein and hover around the surface.

Alternatively, another option could be to branch the arginine residues, by attaching two groups of three arginine residues separately (3.17), as in Figure 3.15.

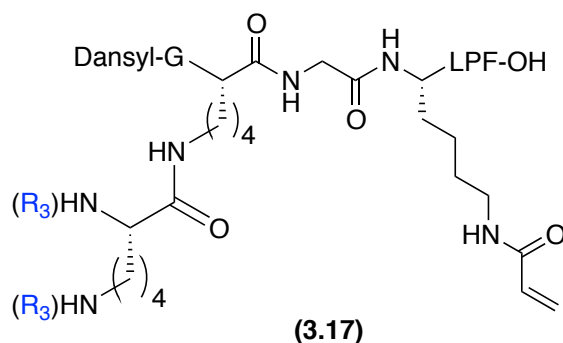


Figure 3.15: Lysine-branched cell permeable, fluorescent TG2-specific probe.

By further increasing the branching, this should ensure the lack of affinity between the hexa-arginine moiety and the LPF pocket. The branched, arginine residues would be too flexible and sterically mobile to constrain itself into the LPF pocket, thus ensuring TG2 affinity.

Hence, these modified and branched versions of the peptide should confer both permeability and TG2 affinity in our efforts to understand TG2 activity *in vivo*. The reconfiguration of peptide attachment remains an avenue to pursue in order to adjust and improve our efforts towards a cell permeable, fluorescent TG2 probe.

References

- (1) Gentile, V.; Thomazy, V.; Piacentini, M.; Fesus, L.; Davies, P. J. A. *Journal of Cell Biology* **1993**, *119*, 463–474.
- (2) Balajthy, Z.; Csomos, K.; Vamosi, G.; Szanto, A.; Lanotte, M.; Fesus, L. *Blood* **2006**, *108*, 2045–2054.
- (3) Griffin, M.; Casadio, R.; Bergamini, C. M. *Biochemical Journal* **2002**, *368*, 377–396.
- (4) Fesus, L.; Piacentini, M. *Trends in Biochemical Sciences* **2002**, *27*, 534–349.
- (5) Jeon, J.-H.; Cho, S.-Y.; Kim, C.-W.; Shin, D.-M.; Kweon, J.-C.; Choi, K.-H.; Park, S.-C.; Kim, I.-G. *Biochemical and Biophysical Research Communications* **2002**, *294*, 818–822.
- (6) Bergamini, C. M.; Collighan, R. J.; Wang, Z. *Advances in Enzymology and Related Areas of Molecular Biology* **78**, 1–46.
- (7) Antonyak, M. A.; Li, B.; Regan, A. D.; Feng, Q.; Dusaban, S. S.; Cerione, R. A. *Journal of Biological Chemistry* **2009**, *284*, 17914–17925.
- (8) Filiano, A. J.; Bailey, C. D. C.; Tucholski, J.; Gundemir, S.; Johnson, G. V. W. *The FASEB Journal* **2008**, *22*, 2662–2675.
- (9) Verma, A.; Mehta, K. *Drug Resistance Updates* **2007**, *10*, 144–151.
- (10) Verma, A.; Wang, H.; Manavathi, B.; Folk, J. Y.; Mann, A. P.; Kumar, R.; Mehta, K. *Cancer Research* **2006**, *66*, 10525–10533.
- (11) Satpathy, M.; Shao, M.; Emerson, R.; Donner, D. B.; Matei, D. *Journal of Biological Chemistry* **2009**, *284*, 15390–15399.
- (12) Maiuri, L.; Luciani, A.; Giardino, I.; Raia, V.; Vilella, V. R.; D'Apolito, M.; Pettoello-

- Mantovani, M.; Guido, S.; Ciacci, C.; Cimmino, M.; Cexus, O. N.; Londei, M.; Quarantino, S. *Journal of Immunology* **2008**, *180*, 7697–7705.
- (13) Milakovic, T.; Tucholski, J.; McCoy, E.; Johnson, G. V. W. *Journal of Biological Chemistry* **2004**, *279*, 8715–8722.
- (14) Gillet, S. M. F. G.; Pelletier, J. N.; Keillor, J. W. *Analytical Biochemistry* **2005**, *347*, 221–226.
- (15) Chabot, N.; Moreau, S.; Mulani, A.; Moreau, P.; Keillor, J. W. *Chemistry & Biology* **2010**, *17*, 1143–1150.
- (16) Keillor, J. W.; Chica, R. A.; Chabot, N.; Vinci, V.; Pardin, C.; Fortin, E.; Gillet, S. M. F. G.; Nakano, Y.; Kaartinen, M. T.; Pelletier, J. N.; Lubell, W. D. *Canadian Journal of Chemistry* **2008**, *86*, 271–276.
- (17) Al-Jallad, H. F.; Myneni, V. D.; Piercy-Kotb, S. A.; Chabot, N.; Mulani, A.; Keillor, J. W.; Kaartinen, M. T. *PLoS ONE* **2011**, *6*, e15893.
- (18) de Macédo, P.; Marrano, C.; Keillor, J. W. *Bioorganic & Medicinal Chemistry* **2002**, 355–360.
- (19) Halim, D.; Caron, K.; Keillor, J. W. *Bioorganic & Medicinal Chemistry Letters* **2007**, *17*, 305–308.
- (20) Shibagaki, N.; Udey, M. C. *Journal of Immunology* **2002**, *168*, 2393–2401.
- (21) Pooga, M.; Soomets, U.; Hallbrink, M.; Valkna, A.; Saar, K.; Resaei, K.; Kahl, U.; Hao, J.-X.; Xu, X.-J.; Wiesenfeld-Hallin, Z.; Hokfelt, T.; Bartfai, T.; Langel, Ü. *Nature Biotechnology* **1998**, *16*, 857–861.
- (22) Astriab-Fisher, A.; Sergueev, D.; Fisher, M.; Ramsay Shaw, B.; Juliano, R. L. *Pharmaceutical Research* **2002**, *19*, 744–755.

- (23) Fawell, S.; Seery, J.; Daikh, Y.; Moore, C.; Chen, L. L.; Pepinsky, B.; Barsoum, J. *Proceedings of the National Academy of Sciences USA* **1994**, *91*, 664–668.
- (24) Lewin, M.; Carlesso, N.; Tung, C.-H.; Tang, X.-W.; Cory, D.; Scadden, D. T.; Weissleder, R. *Nature Biotechnology* **2000**, *18*, 410–414.
- (25) Torchilin, V. P.; Levchenko, T. S.; Rammohan, R.; Volodina, N.; Papahadjopoulos-Sternberg, B.; D'Souza, G. G. M. *Proceedings of the National Academy of Sciences* **2003**, *100*, 1972–1977.
- (26) Frankel, A. D.; Pabo, C. O. *Cell* **1988**, *55*, 1189–1193.
- (27) Vives, E.; Brodin, P.; Lebleu, B. *Journal of Biological Chemistry* **1997**, 16010–16017.
- (28) Gehring, W. J. *Science* **1987**, *236*, 1245–1252.
- (29) Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A. *Journal of Biological Chemistry* **1994**, *269*, 10444–10450.
- (30) Mitchell, D. J.; Kim, D. T.; Steinman, L.; Fathman, C. G.; Rothbard, J. B. *Journal of Peptide Research* **2000**, *56*, 318–325.
- (31) Futaki, S. *Journal of Biological Chemistry* **2000**, *276*, 5836–5840.
- (32) Futaki, S.; Nakase, I.; Suzuki, T.; Zhang; Sugiura, Y. *Biochemistry* **2002**, *41*, 7925–7930.
- (33) Richard, J. P.; Melikov, K.; Vives, E.; Ramos, C.; Verbeure, B.; Gait, M. J.; Chernomordik, L. V.; Lebleu, B. *Journal of Biological Chemistry* **2002**, *278*, 585–590.
- (34) Fittipaldi, A. *Journal of Biological Chemistry* **2003**, *278*, 34141–34149.
- (35) Wadia, J. S.; Stan, R. V.; Dowdy, S. F. *Nature Medicine* **2004**, *10*, 310–315.

- (36) Nakase, I.; Niwa, M.; Takeuchi, T.; Sonomura, K.; Kawabata, N.; Koike, Y.; Takehashi, M.; Tanaka, S.; Ueda, K.; Simpson, J. *Molecular Therapy* **2004**, *10*, 1011–1022.
- (37) Duchardt, F.; Fotin-Mleczek, M.; Schwarz, H.; Fischer, R.; Brock, R. *Traffic* **2007**, *8*, 848–866.
- (38) Schwarze, S. R.; Ho, A.; Vocero-Akbani, A.; Dowdy, S.F.; *Science* **1999**, *285*, 1569–1572.
- (39) Choi, J.-M.; Ahn, M.-H.; Chae, W.-J.; Jung, Y.-G.; Park, J.-C.; Song, H.-M.; Kim, Y.-E.; Shin, J.-A.; Park, C.-S.; Park, J.-W.; Park, T.-K.; Lee, J.-H.; Seo, B.-F.; Kim, K.-D.; Kim, E.-S.; Lee, D.-H.; Lee, S.-K.; Lee, S.-K. *Nature Medicine* **2006**, *12*, 574–579.
- (40) Keliris, A.; Ziegler, T.; Mishra, R.; Pohmann, R.; Sauer, M. G.; Ugurbil, K.; Engelmann, J. *Bioorganic & Medicinal Chemistry* **2011**, *19*, 2529–2540.
- (41) Day, N.; Keillor, J. W. *Analytical Biochemistry* **1999**, *274*, 141–144.
- (42) Roy, I.; Smith, O.; Clouthier, C. M.; Keillor, J. W. *Protein Expression and Purification* **2013**, *87*, 41–46.
- (43) Leblanc, A.; Gravel, C.; Labelle, J.; Keillor, J. W. *Biochemistry* **2001**, *40*, 8335–8342.
- (44) Tulcholski, J.; Lesort, M.; Johnson, G. V. W. *Neuroscience* **2001**, *102*, 481–491.

CHAPTER 4

FXIIIa

Introduction

4.1 Factor XIIIa (FXIIIa) as a protein target

FXIII, also known as blood plasma TGase, exists as an inactive heterotetramer, composed of two A subunits and two B subunits where the active site is contained within subunit A. Exposure to thrombin results in the cleavage of a 4-kDa prosequence from subunit A, which also aids in dissociation of the subunits, creating the active, monomeric form of subunit A, FXIIIa.¹ The enzyme catalyzes the crosslinking of α_2 -antiplasmin α chains, a potent inhibitor of fibrinolysis, into high molecular weight structures. FXIIIa is also responsible for the cross-linking of fibrin to extracellular matrix proteins, resulting in the anchoring of blood clots to the blood vessel wall.^{2,3} Deficiency of FXIIIa results in severe bleeding diathesis and it has also been implicated in atherothrombotic diseases and venous thromboembolism.^{4,5}

Given our interest in this enzyme, we had three goals for this work. Firstly, owing to the anticipated high demand for FXIIIa required for later experiments, our primary aim was the development of an optimized method for the expression of the recombinant subunit, FXIIIa. Secondly, using previously published phage display results as a guide,⁶ we wanted to prepare a series of truncated peptides in order to identify a high-affinity peptidic substrate sequence for FXIIIa that would be orthogonal to the existing sequence for TG2. Lastly, once an orthogonal high affinity FXIIIa substrate sequence was identified, efforts were directed towards modifying the sequence to develop a fluorescent, FXIIIa activity assay.

4.1.1 Methods of expression of FXIII

Several methods are known for expressing and/or purifying FXIII, using several different organisms. Historically, FXIII was first obtained from over-stocked human plasma. Cryoprecipitation was the first step of purification, whereby the plasma was given time to settle at 4 °C and partially pure FXIII was then obtained by centrifugation.⁷ Pure aliquots of the FXIII heterotetramer were obtained by applying the supernatant to a DEAE-cellulose column. From four litres of human plasma, the authors obtained 10 mL of pure Factor XIII with a specific activity of 4000 to 6000 evaluated through the clot solubility test.⁷ The clot solubility test is carried out by determining the least amount of enzyme necessary to form an insoluble clot of fibrin.⁸ Thus, the specific activity was defined as the dilution-fold of the last assay tube containing an insoluble clot, divided by the absorbance at 280 nm of the test solution. The absorbance at 280 nm measures the clottable protein resting in the tube, composed of fibrinogen and synerate.⁸

Later, recombinant expression methods were developed that focused on the expression of only the FXIIIA subunit containing the active site. Amman and coworkers varied the promoters and host-vector systems of 24 different *Saccharomyces cerevisiae* strains. The expression levels of FXIIIA were tested with different media compositions.⁹ Cells were grown in 10 mL of culture broth and after lysis and centrifugation, FXIIIA concentration was measured by applying an enzyme-linked immunosorbent assay to the supernatant. However, due to the limitations of the expression conditions, this study did not attempt the expression of FXIIIA on a production scale but conducted an optimization of expression in yeast on a small scale.

More recently, a method for overexpression in *Escherichia coli* (*E. coli*) was developed for the preparation of FXIIIA protein.¹⁰ *E. coli* is one of the most widely used expression hosts as recombinant protein expression is simple, fast, inexpensive and robust. Lai and coworkers incorporated the gene encoding for FXIIIA into a pGEX-3X vector, which contains the gene for glutathione transferase (GST) at the N-terminus of the protein, along with the ampR gene for ampicillin resistance and the Tac promoter, creating a GST-FXIIIA fusion protein (Figure 4.1).¹¹

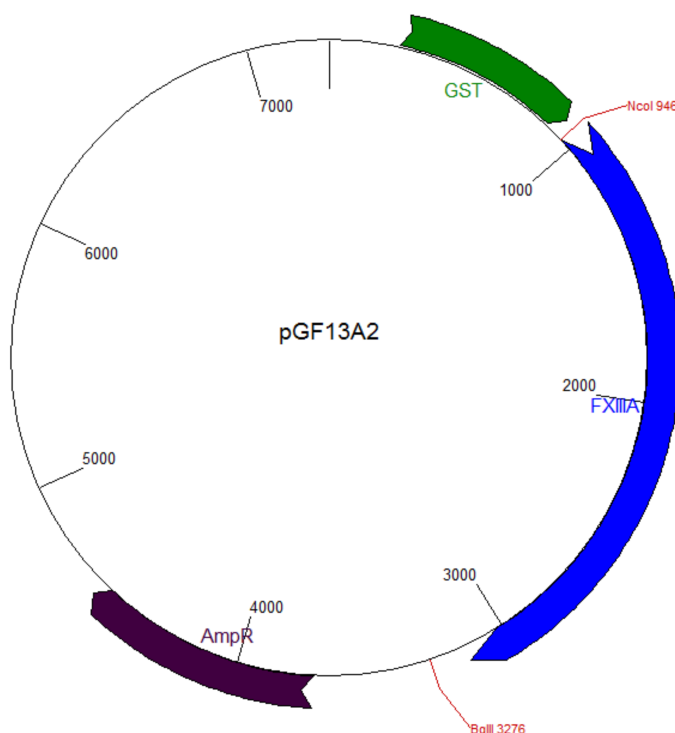


Figure 4.1: GST-FXIIIA plasmid.

The GST tag provides dual advantages, acting as a protein chaperone and expediting purification by affinity chromatography on a glutathione-agarose resin.¹² The expression of the GST-FXIIIA fusion protein was presented as part of a study aimed

at investigating the influence of various mutants on FXIIIA. Purity of the mutant FXIIIA variants expressed ranged from 50-80% although the yield and purity of each mutant was not discussed. Our future work aims at developing high affinity peptide-based probes and inhibitors for FXIIa, requiring large quantities of the protein. This led us to optimize the expression of GST-FXIIIA in order to obtain the highest yield and purity possible. The variables modified included expression temperature, isopropyl β -D-1-thiogalactopyranoside (IPTG) concentration and expression medium.

4.1.2 Existing methods to test activity of FXIIa

Along with our goal to improve the yield of FXIIa purification, an efficient and specific method to test its activity is also required, both for *in vivo* and *in vitro* applications. The earliest methods of detection of FXIIa activity were amine incorporation assays, whereby fluorescent,¹³ radiolabeled¹⁴ or biotinylated¹⁵ amine substrates are covalently attached to a protein using FXIIa; protein-bound and free amines are separated and the protein-linked fraction is measured quantitatively. Though highly sensitive, these amine incorporation assays are cumbersome, time-consuming and difficult to standardize. Since the reproducibility is poor, assays cannot be adapted to clinical chemistry analyzers and the separation step excludes the possibility of a continuous kinetic assay.

The commercially available Zedira assay, based on the reverse “isopeptidase” activity of FXIIa,¹⁶ is widely applied. This detection method uses FRET, whereby the increase of fluorescence results from release of the N-terminal 2-aminobenzoyl

fluorophore, which is internally quenched by the 2,4-dinitrophenyl moiety before transamidation of the peptide substrate (Figure 4.2).

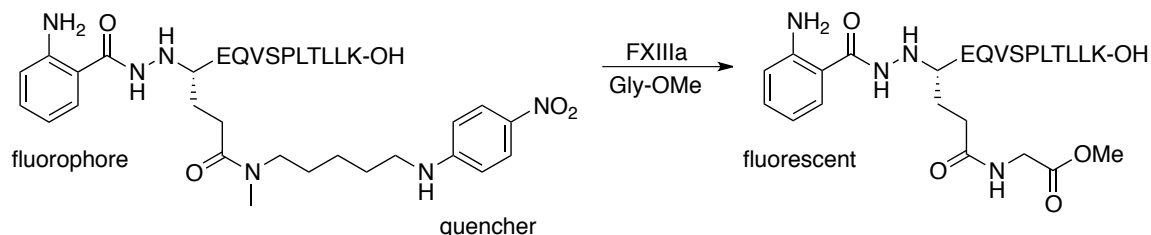


Figure 4.2: Fluorogenic FRET-based Zedira assay.

The fluorophore and quencher are incorporated into a peptidic sequence derived from the α_2 -antiplasmin dodecapeptide NQEQVSPLTLLK, determined to be an excellent substrate for FXIII.¹⁶ While the peptide has high affinity for FXIIIa, the isopeptide was also found to possess cross reactivity with TG1, TG2, TG3 and TG6, and thus is not designed for use in a *selective* FXIIIa activity assay.

Another assay widely applied for transglutaminases is the glutamate dehydrogenase (GDH) assay. This assay is based on the continuous monitoring of ammonia released during the transglutaminase reaction, measured by the GDH indicator reaction (Figure 4.3).¹⁷ The ammonium produced by the FXIIIa enzymatic reaction is bound by GDH and is used to reductively aminate α -ketoglutarate, producing L-glutamate, concurrently oxidizing β -NADH to NAD^+ . The disappearance of NADH is followed spectrophotometrically at 340 nm and is coupled to the activity of FXIIIa.

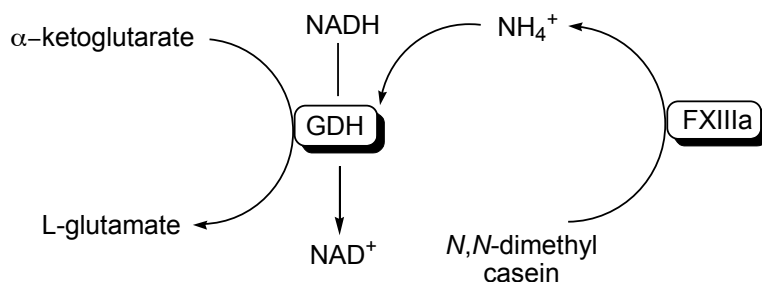


Figure 4.3: Glutamate Dehydrogenase (GDH) assay.

Similarly to the Zedira assay, this assay is not specific to FXIII as casein is also a substrate of TG2.¹⁸

4.1.3 FXIIIa specific peptides

A FXIIIa-specific assay requires a FXIIIa-specific substrate. Recently, Hitomi and coworkers identified a series of FXIIIa-specific peptide sequences derived from an unbiased phage-displayed random peptide library. It was developed as a screening system to elucidate the primary structures surrounding reactive glutamine residues preferred by TG2 and FXIIIa.⁶ Biotin-cadaverine was used in catalytic reactions with either TG2 or FXIIIa and incorporated into the glutamine peptide. This allowed the reactive peptides to be isolated by avidin affinity purification. Isolated peptide sequences were then expressed as GST fusion proteins. Their specificity as glutamine donors was evaluated using a monodansylcadaverine to indicate the degree of TGase-mediated incorporation into the glutamine residue of the peptide sequence. Through their experiments, Hitomi and coworkers identified the sequence of DQMMLPWPVKL (F11) as having the highest reactivity and selectivity for FXIIIa (Figure 4.4).

clone																			
	F23		L	D	Q	I	V	I	P	W	P	T	P	R					
*	F11			D	Q	M	M	L	P	W	P	A	V	K	L				
	F18			D	Q	M	M	L	P	W	P	A	V	K	L				
*	F19			D	Q	W	M	M	A	W	P	S	L	T	L				
	F9			E	Q	Y	Q	L	A	W	P	A	M	Q	S				
	F14			E	Q	Y	Q	L	A	W	P	A	M	Q	S				
	F2			E	Q	Y	Q	L	A	W	P	A	M	Q	S				
*	F16		S	E	Q	H	L	L	K	W	P	P	W	H					
	F7			E	Q	Q	Q	L	S	W	P	P	F	L	P				
*	F12			S	Q	H	P	L	P	W	P	V	L	M	L				
	F26			S	Q	H	P	L	P	W	P	V	L	M	L				
*	F6			S	Q	I	P	M	A	W	P	L	L	S	L				
	F31			T	Q	Y	T	M	T	W	P	F	S	W	R				
	F20			W	Q	Q	Q	M	K	W	P	S	F	T	Y				
*	F13			W	Q	I	P	V	D	W	P	P	L	P	P				
	F29			W	Q	M	Q	L	P	W	P	P	R	L	S				
	F35			G	Q	M	I	L	P	W	P	S	P	R	A				
*	F17			E	Q	F	P	I	A	F	P	R	Y	S	I				
	F21			H	Q	M	V	L	T	Y	P	W	V	P	T				
	F32			D	Q	Y	V	L	T	F	P	W	T	P	R				
	F4			W	Q	M	P	V	N	L	P	I	R	Y	P				
	F15			W	Q	L	K	Y	P	W	I	S	P	V	G				
	F5			H	Q	I	P	I	Q	I	H	S	F	A	M				
*	F28			W	Q	H	K	I	D	L	R	Y	N	G	A				
	F27			W	Q	H	K	I	D	L	R	Y	N	G	A				
	F22			D	Q	S	K	V	W	L	L	P	S	W	K				
	F30		W	H	Y	Q	V	P	Q	W	L	Y	R						
	F25	W	Y	H	P	L	Q	L	T	P	A	P	I						

Figure 4.4: Alignment of preferred substrate peptide sequences of phage clones

for FXIIIa.⁶ (Sugimura, Y.; Hosono, M.; Wada, F.; Tohru, Y.; Maki, M.; Hitomi, K. *Journal of Biological Chemistry* **2006**, *281*, 17699–17706) Reproduced with permission.

Closer inspection of the sequences reveals trends in the various peptidic sequences shown to have affinity for FXIIIa and orthogonality with respect to TG2. These include the presence of a carboxylic acid sidechain (D/E) on the N-terminal side of the reactive glutamine and the presence of a methionine residue on the C-terminal side of the glutamine residue. From the results obtained by Sugimura and coworkers, the negatively charged carboxylic acid N-terminal to the reactive glutamine residue

contributes to the selectivity of the preferred substrates for FXIIIa over TG2. Based on their observations, we decided to incorporate these features in our design of various peptides to test their affinity with FXIIIa.

The efficiency of these peptide substrates was then analyzed kinetically.

$$\text{Rate} = \frac{V_{\max} \times [S]}{(K_m + [S])}$$

In the Michaelis-Menten equation shown above, plotting the observed rate against the substrate concentration results in a hyperbolic curve. The K_m value corresponds to the concentration of substrate at which the observed rate is equal to half of the maximum velocity of the reaction and is an indication of the strength of binding of the substrate to the enzyme. The $V_{\max}$ value represents the maximum rate of product formation from the enzyme-substrate complex. In order to compare reaction rates when different quantities of enzyme are used, the $V_{\max}$ value is divided by the concentration of enzyme to obtain a value known as k_{cat} , which is the rate constant of catalysis. Thus, to compare the selectivity for different peptides, the ratio of k_{cat}/K_m was employed, as it represents the overall catalytic efficiency of the reaction of the enzyme with a substrate.

4.1.4 FXIIIa specific activity assay

Owing to the lack of a highly specific FXIIIa assay, we set out to develop one, based on the incorporation of a fluorogenic residue into a donor substrate peptide sequence selective for FXIIIa.

Previously in the Keillor group, a direct, continuous, fluorometric assay of TG2 was prepared based on 4-(*N*-carbobenzoxyl-L-phenylalanyl-amino)-butyric acid coumarin-7-yl ester (4.1) (Figure 4.5).¹⁹ The progress of the TGase-catalyzed reaction

was monitored as an increase in fluorescence due to the release of the 7-hydroxycoumarin (4.3) through cleavage of the ester bond by TG2 (Figure 4.5). The incorporation of a Cbz group at the N-terminus was enough to afford affinity and selectivity for TG2 as the substrate was unreactive with both FXIIIa and microbial TGase.

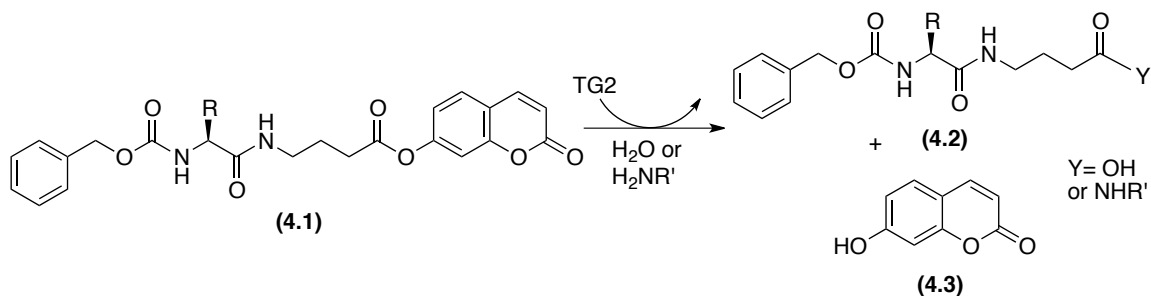


Figure 4.5: Fluorometric assay for TG2 with 7-hydroxycoumarin.

The positive results obtained with this previously developed assay for TG2 inspired us to use the same 7-hydroxycoumarin fluorophore, incorporated into a high affinity peptide that would result in a specific fluorogenic activity assay for FXIIIa (Figure 4.6).

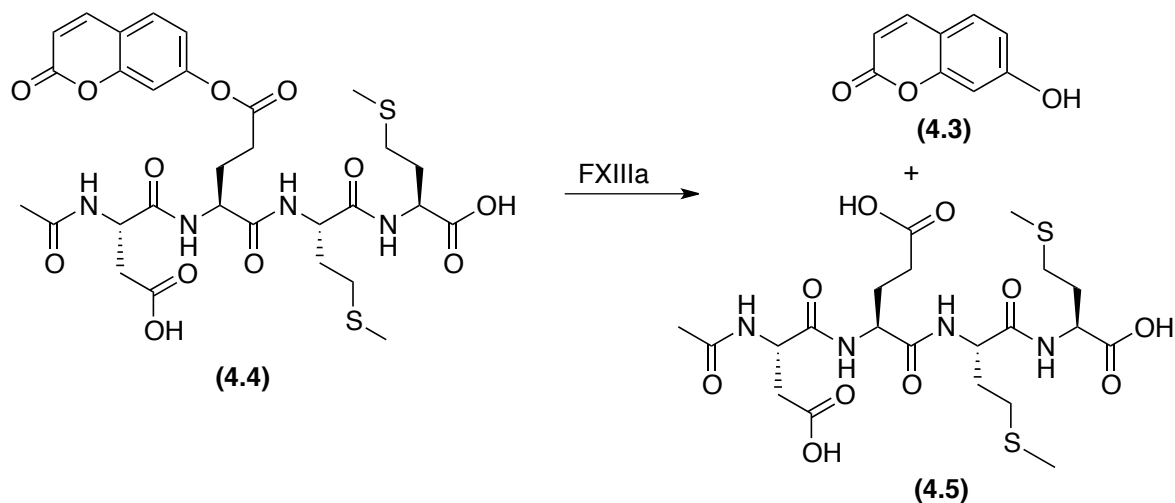


Figure 4.6: Fluorogenic activity assay for FXIIIa shown incorporated into Ac-DEM-OH as a model peptide.

4.2 Research goals

Since a satisfactory expression protocol had not yet been determined for FXIIIa, our first priority was the optimization of purification of FXIIIa. Once we had the enzyme in hand, our interest turned to developing a FXIIIa-specific activity assay. To that end, various peptides were synthesized and tested for affinity for FXIIIa while concurrently working towards the synthesis of a FXIIIa-specific fluorogenic substrate through the modification of glutamic acid.

Results/Discussion

4.3 Optimization of purification

Among the three parameters of expression that we decided to test, namely expression temperature, IPTG concentration and type of media, we decided to start with expression temperature. The optimal temperature was then used in the later experiments, varying IPTG concentration and then expression media (Figure 4.7).

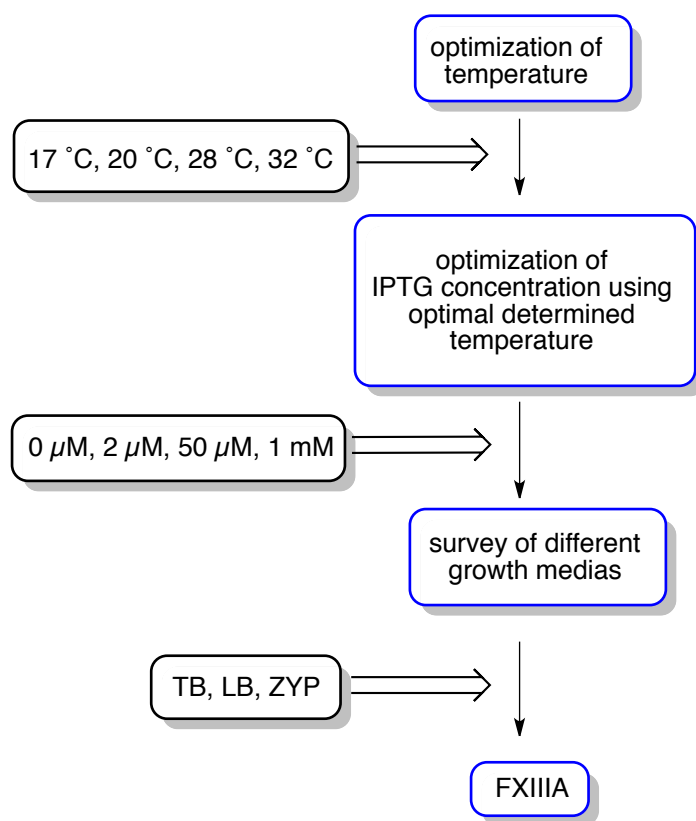


Figure 4.7: Optimization of Expression of FXIIIA.

After transforming *E. coli* BL21 (DE3) cells with the plasmid obtained from Lai and coworkers, we tested different expression temperatures in TB media, ranging from 17 °C to 32 °C without the addition of IPTG (Figure 4.8). The protein purified was a GST-FXIIIA fusion protein with a molecular weight of 110 kDa.

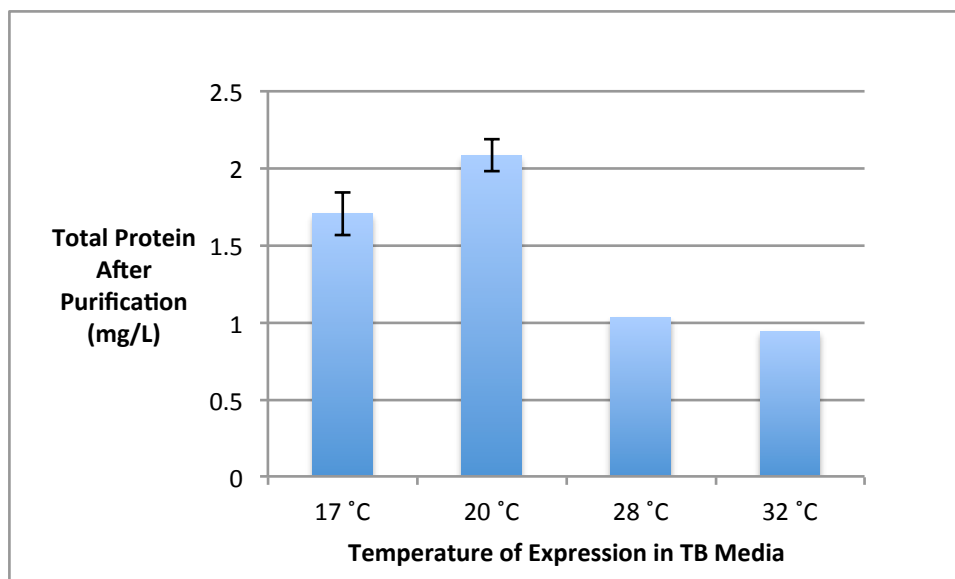


Figure 4.8: Different temperatures of expression in TB media for FXIIIA.

Through analysis of SDS-PAGE gels (Figure 4.9) for purity (after GST-based purification) and the Bradford assay to quantify the amount of protein, it was determined that 20 °C was the optimal temperature to express FXIII. At higher temperatures, namely 28 °C and 32 °C, the amount of FXIII expressed was minimal in comparison to other protein contaminants, such as GST, while at 17 °C little expression was observed. Greenberg and others have reported that during the expression of FXIIIA, *E. coli* produces a 30-kDa FXIII-related protein.^{10,20} During our optimization studies, we tried to minimize the production of these protein contaminants while maximizing the quantity of FXIIIA.

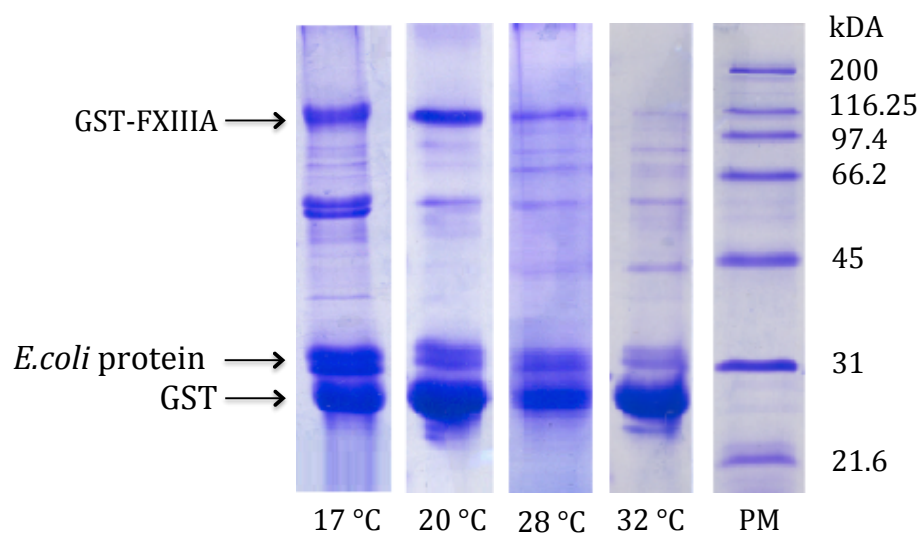


Figure 4.9: SDS-PAGE gels of FXIIIA expression at different temperatures.

Using the optimal temperature of 20°C, the concentration of IPTG necessary was investigated by testing the following IPTG concentrations; 0 μ M, 2 μ M, 50 μ M and 1 mM (Figure 4.10).

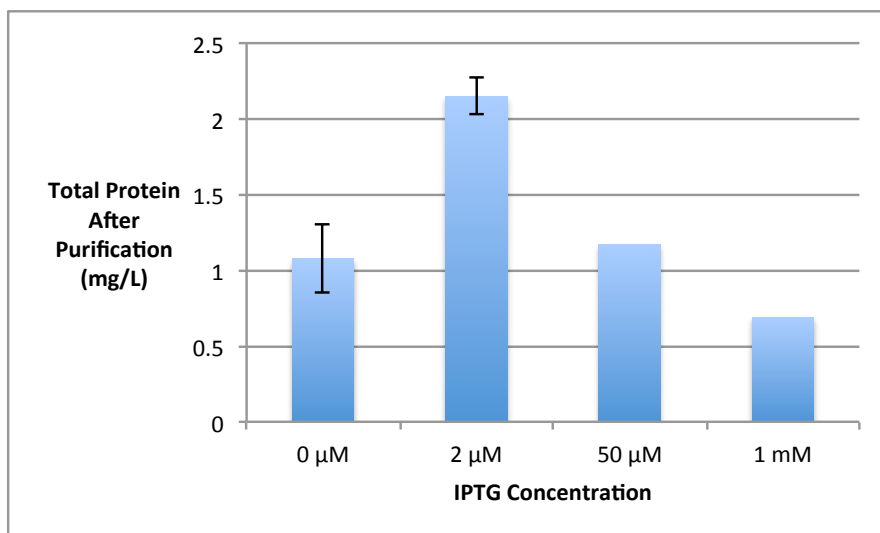


Figure 4.10: Total protein obtained with different concentrations of IPTG in TB media.

It was found that 2 μ M of IPTG was the optimal concentration for FXIIIA expression because at higher concentrations of IPTG, the amount of FXIII in the soluble fraction was minimal when compared to other protein contaminants as seen through SDS-PAGE gels (Figure 4.11).

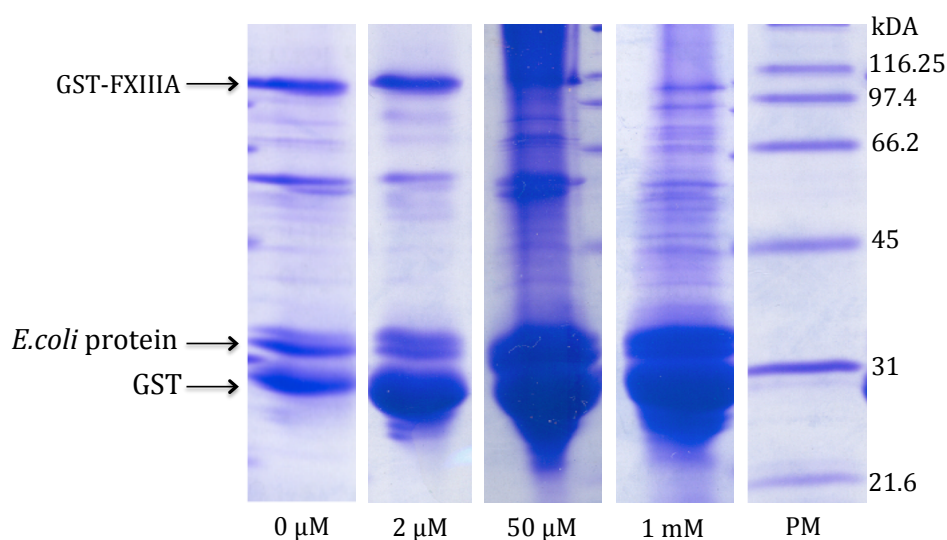


Figure 4.11: SDS-PAGE gels of FXIIIA expression at different IPTG concentrations

The results obtained to this point were similar to the expression conditions used by Lai and coworkers, who expressed the protein at 24 °C in Terrific Broth (TB) media without the addition of IPTG.¹⁰ This already suggested that the protein folded slowly and needed time to express properly. Too high of a concentration of IPTG can cause cellular stress and would probably precipitate an overabundance of FXIIIA which would fold improperly and then be lost, leading to low yields obtained as previously observed with TG2.²¹ Similarly, the higher temperatures of 28 °C and 32 °C would cause the expression machinery to move faster, leading to FXIIIA being expressed too quickly and improperly folding.

The slow expression of FXIIIA with a minimal amount of IPTG led us to consider the ZYP auto-induction medium of Studier as one of our types of expression medium, in addition to LB.²² The principle of auto-induction medium stipulates that by controlling the balance between glucose and lactose in cells, the glucose will prevent induction by lactose but once depleted, auto-induction takes over with lactose. Thus, we expressed FXIIIA in LB medium, using 0 and 2 μ M IPTG along with Studier's ZYP auto-induction medium at 20 °C (Figure 4.12).

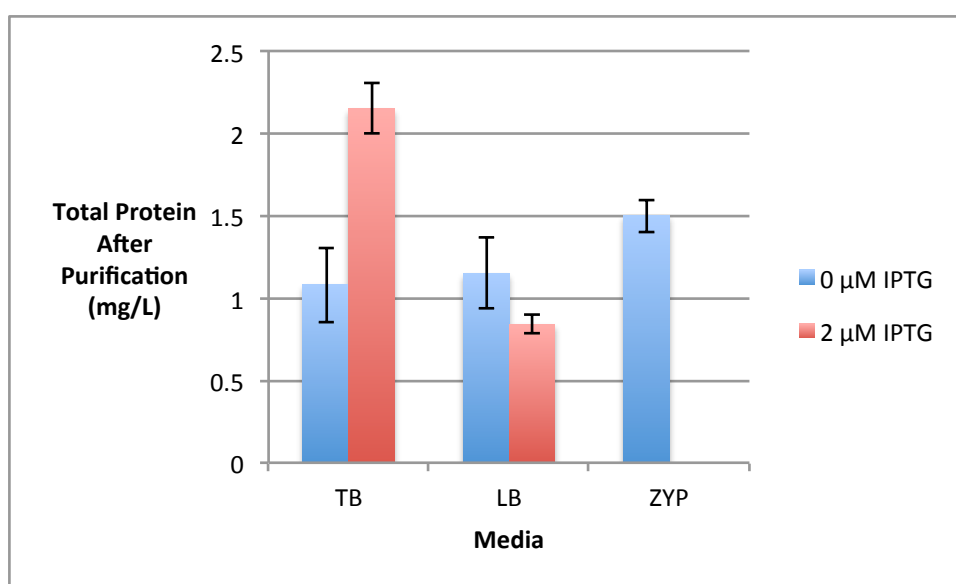


Figure 4.12: Total protein obtained with different expression media.

As shown above, the condition of 2 μ M of IPTG in TB media provided the most total protein (approx. 2.1 mg/L), while with the ZYP auto-induction medium; around 1.5 mg/L of culture of protein was obtained (Figure 4.12). However, the ZYP auto-induction medium provided the purer sample of protein, qualitatively evaluated through SDS-PAGE gels (Figure 4.13). Though it had a lower total amount of protein,

the quantity of FXIIIA was higher with the ZYP auto-induction medium than with 2 μ M of IPTG in TB medium.

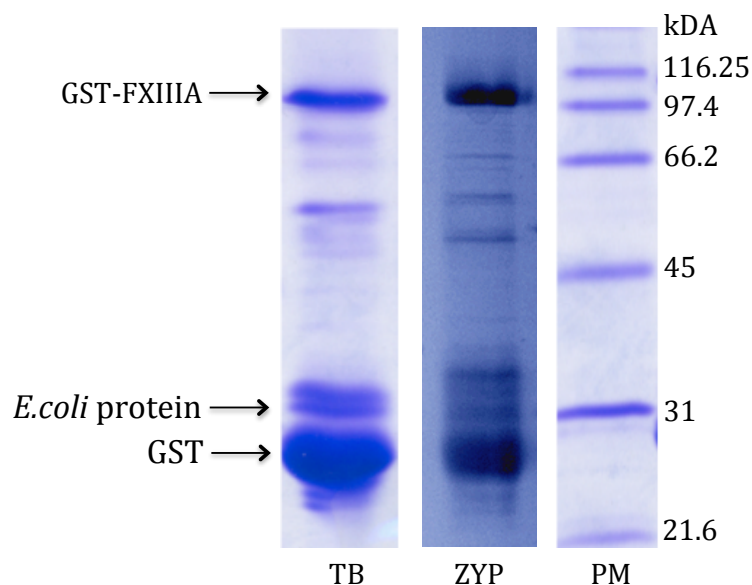


Figure 4.13: SDS-PAGE gel of FXIIIA expression with TB and ZYP media.

In the ZYP medium, the overexpression of FXIIIA was induced by the disappearance of glucose that occurs when the cells have almost reached the end of the growth phase and can now focus on protein expression. Contrastingly, induction with IPTG occurs during the beginning of the exponential growth phase and so the cell's resources are divided between replication and protein expression. This difference could explain why a greater amount of FXIIIA was produced using the ZYP medium as opposed to the TB or LB medium. TB medium was shown to be more efficient than LB medium. This is due to the fact that in the TB medium, additional phosphate ions are added, which would help to buffer pH changes in the medium, while glycerol is a supplementary carbon source in TB medium.

Auto-induction is also more convenient than IPTG induction because auto-inducing medium can simply be inoculated with the expressing bacterial strain and grown to saturation without the need to follow culture growth or add inducer. It has been shown that auto-induction and saturation often occur at a considerably higher cell density at 20 °C rather than 37 °C, and this could perhaps be due to the higher solubility of oxygen at the lower temperature but coincides nicely with the lower temperature required by FXIIIa.²²

Thus, the auto-induction medium at 20 °C provided the purest fraction and greatest quantity of FXIIIa out of the expression conditions investigated.

4.4 High affinity peptide sequences for FXIIIa

Using the phage display results of Hitomi and coworkers, we set out to evaluate whether these sequences could be truncated while still maintaining their affinity for FXIIIa. The F11 peptide was chosen from the Hitomi study and truncated variants were prepared and tested for their affinity for FXIIIa through the use of the glutamate dehydrogenase assay. Truncated peptides were also inspired by sequences from β -casein, a FXIIIa substrate. The truncated peptide was used as the acyl-donor substrate and glycine ethyl ester was used as the acceptor substrate, since it was previously shown to be an effective acyl acceptor (Figure 4.14).²³

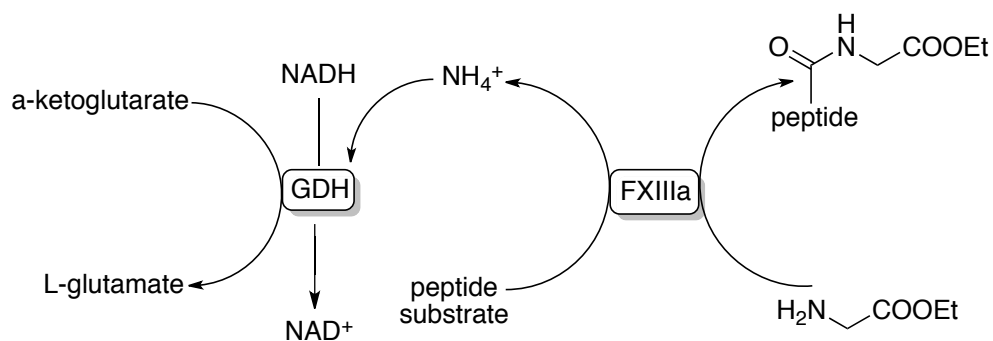


Figure 4.14: GDH assay with glycine ethyl ester as acceptor substrate

The peptides designed were chosen by looking at commonalities in the peptides evaluated by Hitomi⁶ while also trying to avoid sequences recognized by TG2.

4.4.1 Design of high affinity peptide sequences.

A total of 6 different peptides were prepared and tested for their affinity with FXIIIa. Table 4.1 lists the peptides designed as well as the inspiration for their sequence.

Peptide #	Sequence	Model Sequence	Inspiration
4.6	Ac-SQMM-OH	DQMMLPWPAVKL	F11
4.7	Ac-DQMML-OH	DQMMLPWPAVKL	F11
4.8	Ac-DQMHV-OH	DQMMLPWPAVKL	F11
4.9	Ac-DQMMAF-OH	DQMMLPWPAVKL	F11
4.10	Ac-SQSRV-OH	SVLSLSQSKVLPVPE	β -casein
4.11	Ac-GQSMV-OH	LGPGQSKVIG	K9

Table 4.1: Peptidic sequences designed for FXIIIa.

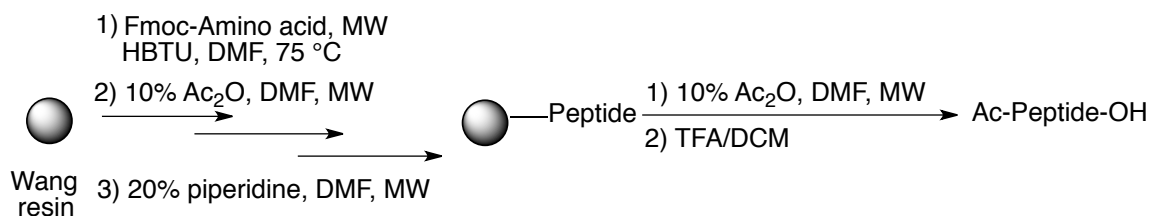
Peptides 4.6-4.8 were chosen to mimic F11, replacing the widely occurring aspartic acid with a serine residue in peptide 4.6, whereas peptide 4.7 was a direct, truncated 5-amino acid version of F11. In peptide 4.8, the second methionine residue was replaced by a histidine to investigate the effect of replacing a hydrophobic residue with an aromatic, hydrophilic amino acid. Peptide 4.9ⁱⁱ was a longer variant, with a total of seven amino acids to check the effect of peptide length. The third methionine was chosen to replace the leucine in F11 since it is a commonly occurring amino acid in that position (in addition to leucine and isoleucine). The phenylalanine in peptide 4.9 was incorporated to keep the aspect of aromaticity but avoiding the use of a tryptophan residue, which is less soluble and also requires a protecting group for the indole nitrogen. Peptide 4.10 was a truncated sequence consisting of 5 amino acids taken

ⁱⁱ Synthesized by Abullah Akbar, PhD Candidate in the Keillor group

from the peptide recognition sequence of β -casein. While casein is a substrate for both FXIII and TG2, β -casein has been shown to have poor recognition by TG2 but is a good substrate for FXIII.²⁴ Lastly, peptide 4.11 was inspired by the K9 sequence used in the Berichrom assay (from the Berichrom Corporation, now taken over by Dade Behring in Germany), that is also based on β -casein.²⁵ However, the lysine amino acid in the K9 sequence was replaced by a methionine residue to avoid cross-linking of the peptide.

4.4.2 Synthesis of peptides

The peptides were synthesized using a CEM Liberty microwave peptide synthesizer, as described in the previous chapter.



Scheme 4.1: Synthesis of FXIIIa substrate peptides

Using the Wang resin, the commercially available amino acids were coupled using microwave irradiation with HBTU as an activator in DMF at 75 °C (Scheme 4.1). After coupling all the residues, the peptide was acetylated and then cleaved using a mixture of TFA and dichloromethane.

4.4.3 Optimization of FXIIIA activation by thrombin

FXIIIA is unique (along with epidermal transglutaminase) among mammalian transglutaminases in that it requires activation through cleavage of a pro-peptide before becoming active. The cleavage is carried out through the serine protease, thrombin. However, if FXIIIA is left in the presence of too much thrombin for too long, a secondary cleavage is mediated by thrombin that renders the enzyme inactive. The secondary cleavage will split the protein into 25.4 kDa and 53.9 kDa fragments. Thus, optimization of this step was carried out, wherein several different amounts of thrombin were added and incubated with our recombinant GST-FXIIIA protein for an hour. The thrombin cleavage site on our recombinant fusion protein is located between the GST tag and FXIIIA protein; thus, we expected to obtain a protein with a molecular weight of 79.2 kDa. From the SDS-PAGE gel of the different samples (Figure 4.15), it can be seen that 5 U of thrombin is enough to cleave the activation peptide and the GST protein under these conditions.

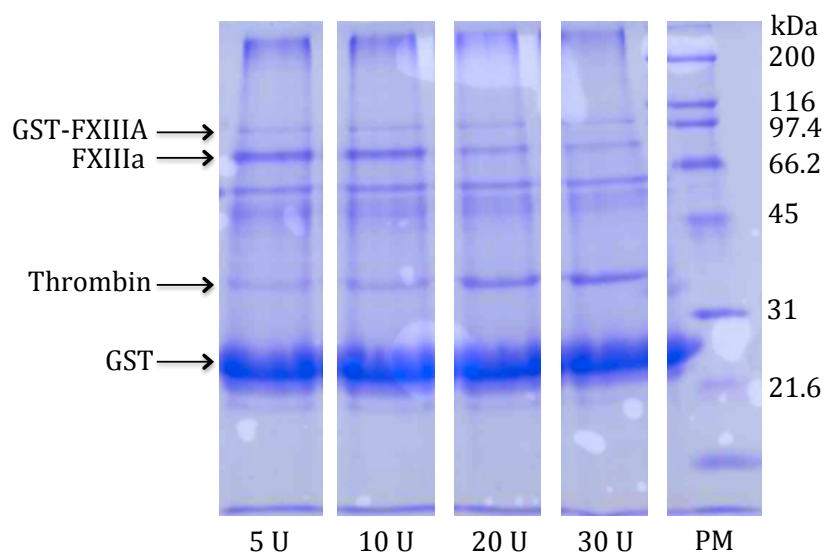


Figure 4.15: SDS-PAGE gel of FXIIIA with differing quantities of thrombin.

The gel shows that 20 U of thrombin proves excessive, as the majority of FXIIIA has been cut at the second thrombin cleavage site, leading to further fragmentation of the protein. At this point, a further study was undertaken to quantitatively refine the amount of thrombin necessary, wherein FXIIIA was incubated with 4, 3, 2, 1 and 0.5 U of thrombin. After one hour, the activity of FXIIa was then measured using the GDH assay with *N,N*-dimethyl casein as a substrate. Through determination of FXIIa activity, evaluated by measuring the decrease in absorbance of NADH, it was found that there was little variation in the specific activities calculated for FXIIa (Table 4.2). Thus, 1 U of thrombin was sufficient for activation of FXIIA.

Units of Thrombin	Specific Activity (U/mg)
0.5	0.12 ± 0.0012
1	0.16 ± 0.0013
2	0.1 ± 0.0018
3	0.15 ± 0.0023
4	0.1 ± 0.0025

Table 4.2: Units of thrombin correlated to specific activity of FXIIa.

4.4.4 Kinetic analysis of FXIIa substrate peptides

Once the optimal quantity of thrombin required for FXIIA activation had been determined, the synthesized peptides were tested for their efficiency as acyl donor substrates with FXIIa. This was carried out using the GDH assay (Figure 4.8), whereby

the decrease in absorbance of NADH was measured as a function of concentration of donor peptide.

However, saturation of FXIIIa activity could not be reached within the solubility limits of the peptides. Hence, the rate of the reaction was then plotted against peptide concentration and a straight line was obtained (Figure 4.16).

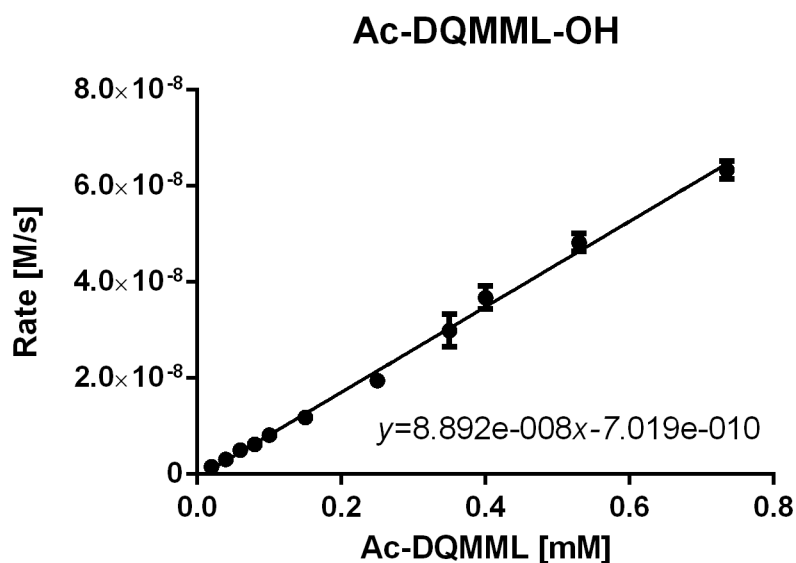


Figure 4.16: Rate vs. concentration of Ac-DQMML-OH.

If we assume this straight line to be representative of the linear portion of a typical Michael-Menten curve, then the slope of this line gives the ratio of $V_{\max}/K_m$ for the specific peptide. This analysis was carried out for the remaining peptides, where a $V_{\max}/K_m$ value was calculated from the linear fit of the data (Figure 4.17-Figure 4.21). Some of the graphs have slightly negative intercepts and this can be attributed to shortcomings in using a coupled assay. Low peptide concentrations result in a minimal

amount of ammonia is being produced which then decreases the turnover of the GDH assay and consequently, the transformation of NADH into NAD⁺.

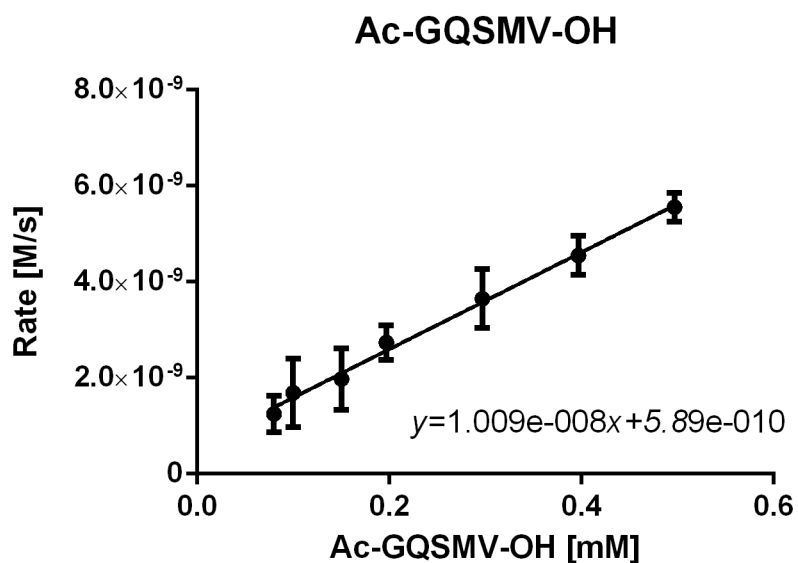


Figure 4.17: Rate vs. concentration of Ac-GQSMV-OH.

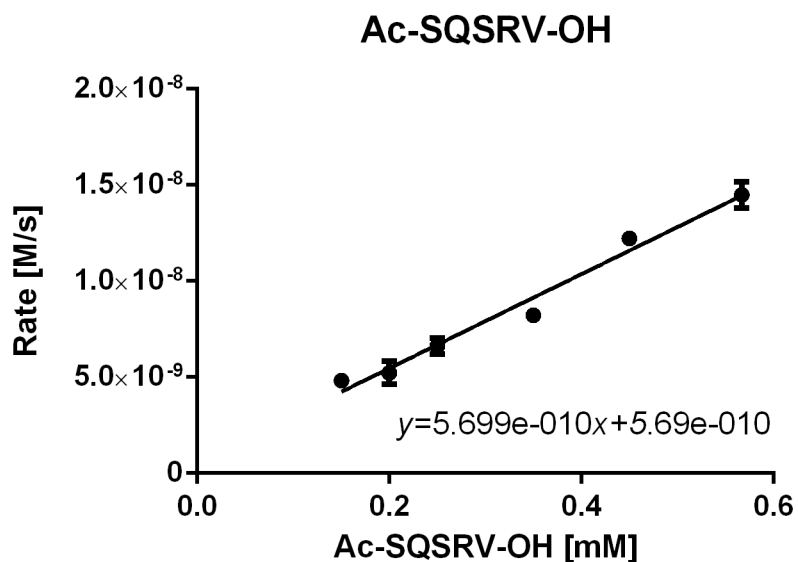


Figure 4.18: Rate vs. concentration of Ac-SQSRV-OH.

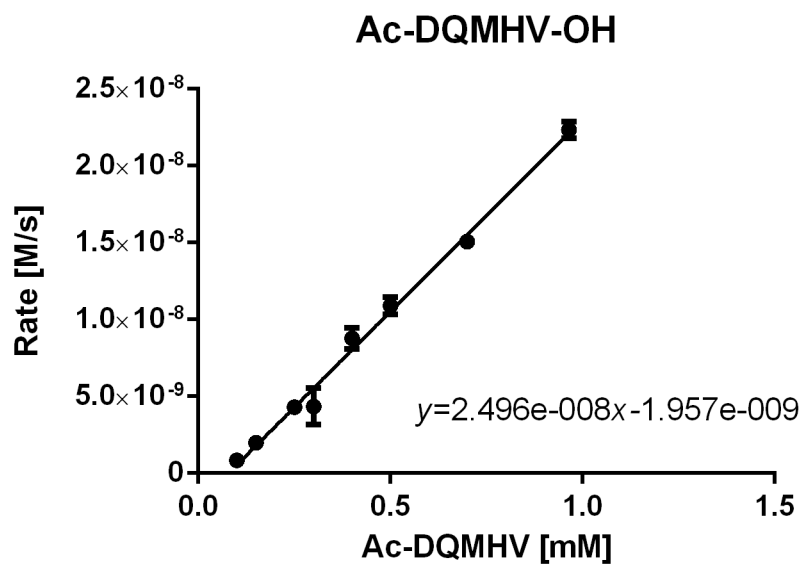


Figure 4.19: Rate vs. concentration of Ac-DQMHV-OH.

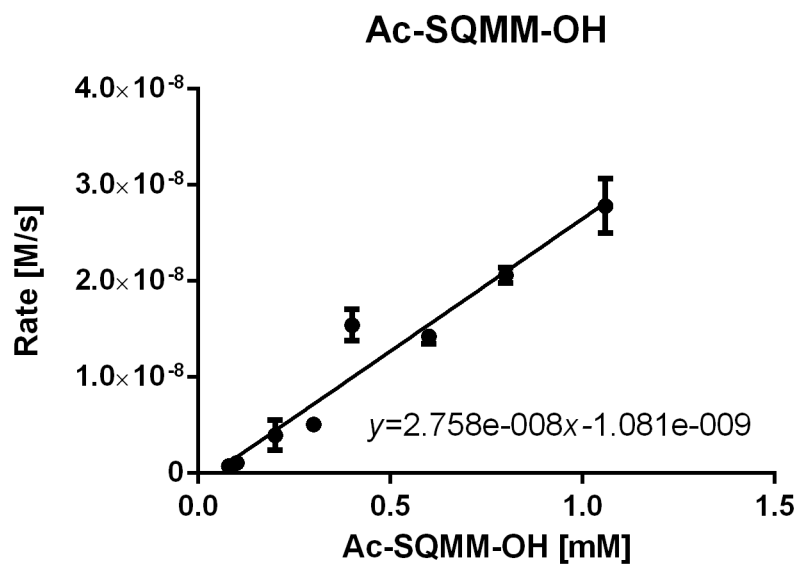


Figure 4.20: Rate vs. concentration of Ac-SQMM-OH.

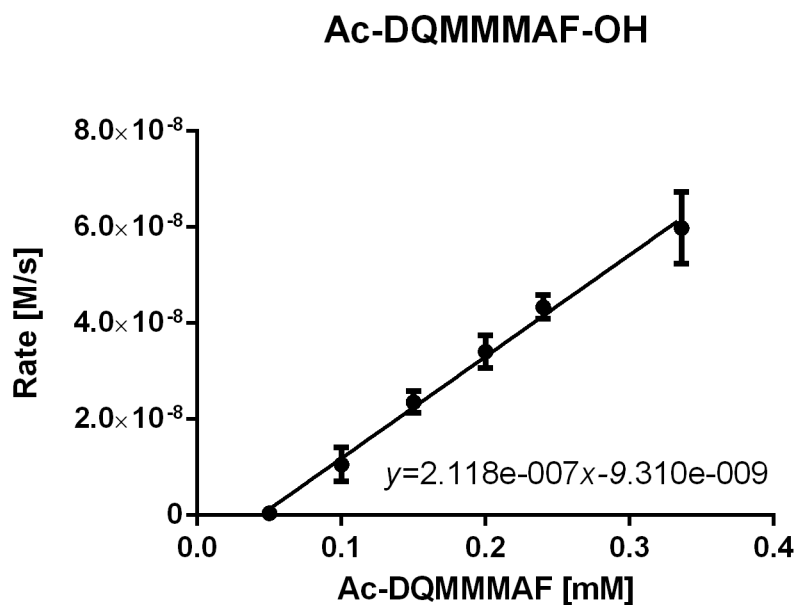


Figure 4.21: Rate vs. concentration of Ac-DQMMMAF-OH

These values were then divided by the appropriate enzyme concentration to obtain the value of k_{cat}/K_m , a measure of catalytic efficiency (Figure 4.22).

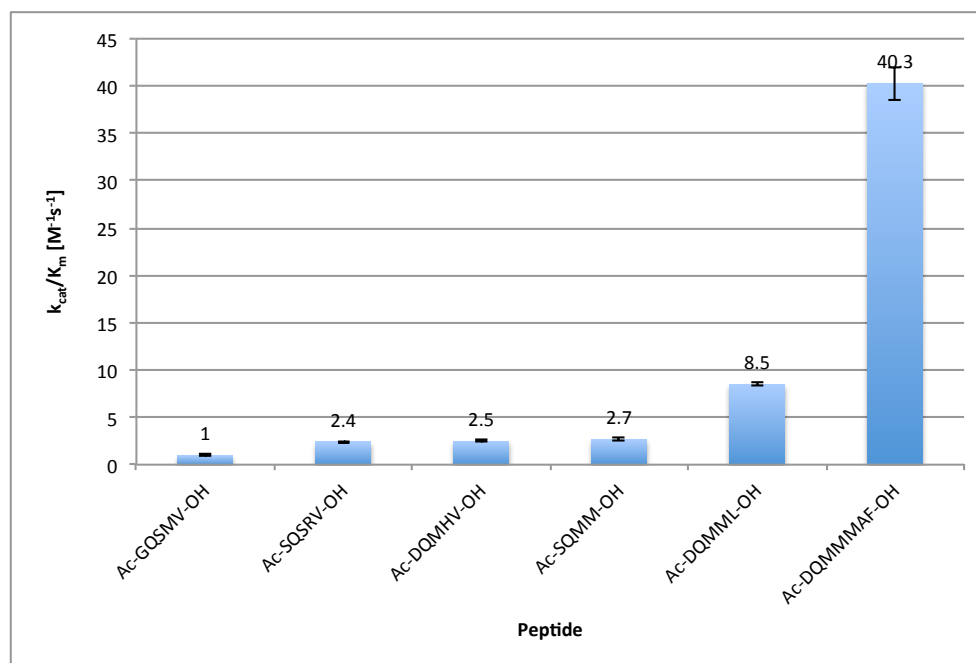


Figure 4.22: Ratio of k_{cat}/K_m of peptides with FXIIIa.

From Figure 4.22, one clear conclusion can be reached. The longer the peptide, the more effectively it is recognized by FXIIIa. The longest peptide, at 7 amino acids, had the highest reactivity of all the peptides tested. The results also lead to the implication that the aspartic acid and methionine residues are important for FXIIIa reactivity and recognition, since the Ac-GQSMV-OH peptide is the least reactive. Comparing peptides Ac-DQMHV-OH and Ac-DQMML-OH, with the latter having almost four times the efficiency of the former, indicates a preference for the MM sequence as both leucine and valine have similar properties. The combination of a hydrophilic amino acid on the N-terminal side of the glutamine residue and a methionine residue on the C-terminal side seem to be ideal for FXIIIa recognition with that being a common trend in peptide reactivity. At this time, it cannot be said whether it is a hydrophilic amino acid or specifically aspartic acid that aids in recognition. Ac-DQMML-OH had higher reactivity than the Ac-SQMM-OH peptide but it remains to be seen whether that was due to the aspartic acid or the added leucine residue. However, it could also be that the aspartic acid helps in providing orthogonality with TG2, by forming unfavourable interactions with the tissue enzyme.

To that end, the two peptides with the highest affinity for FXIIIa, Ac-DQMML-OH and Ac-DQMMAF-OH, were both tested with TG2, using the GDH assay. Analysis was carried out as described above to obtain the ratio of k_{cat}/K_m for the two peptides with TG2.

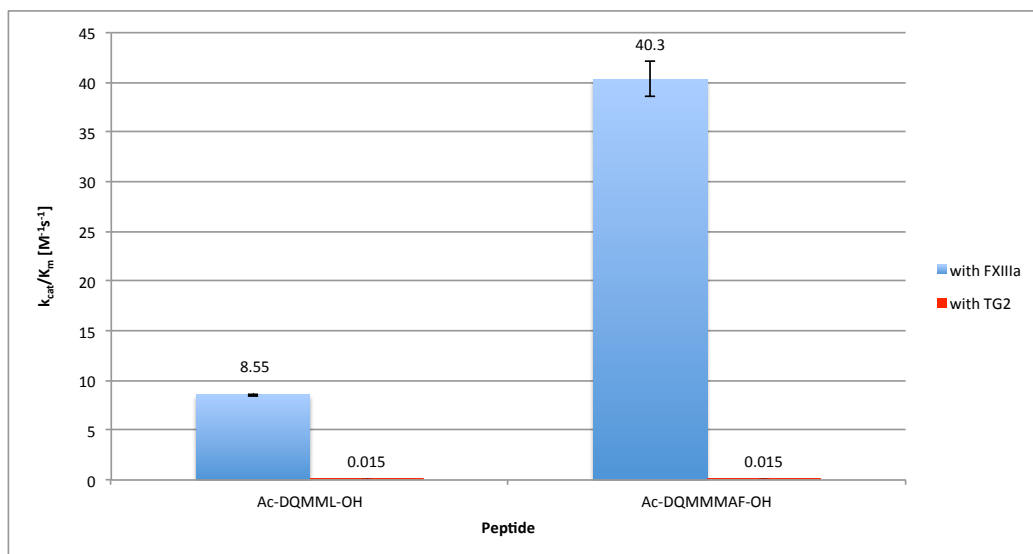


Figure 4.23: Comparison of affinity between FXIIIa and TG2.

From Figure 4.23, it can be seen that both peptides have little recognition for TG2.

Length does not make a difference in terms of binding to TG2 as both peptides had very similar k_{cat}/K_m values with TG2 that were ~ 1000 -fold less than their values with FXIIIa.

Thus, the DQMM motif is orthogonal to TG2 and it remains to be investigated if there exist other peptides with a higher affinity for FXIIIa than the ones identified herein.

Other peptides that should be prepared and tested include Ac-DQMMLPW-OH, which is the truncated version of F11, the best peptide identified by Hitomi. Testing the peptide Ac-SQMMMAF-OH will also help distinguish the level of importance of the aspartic acid residue, both in terms of FXIIIa reactivity along with TG2 recognition.

4.4.5 Modelling of FXIIIa peptidic sequences with FXIIIa

In an attempt to understand the mode of binding of the substrate peptides to FXIIIa and to distinguish the difference between Ac-DQMML-OH and Ac-DQMMMAF-OH, molecular modelling of the peptide structures was undertaken with a homology model of FXIIIa. Numerous groups have indicated that FXIIIa has an open and closed form structure, much like TG2. But at the time of writing, no crystal structure of the open form of FXIIIa had been released.²⁶⁻³⁰ Thus, from the crystal structure of the open form of TG2 (PDB: 1KV3),³¹ a homology model was created for FXIIIa using the MOE (Molecular Operating Environment) program from the Chemical Computing Group. Peptide substrates were minimized using conjugate gradient minimization until a convergence of $0.0001 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-1}$ was reached. The peptide was oriented and placed within the selected docking site using the Triangle Matcher method using the refined, open model of FXIIIa. Docked poses were then scored using the London ΔG scoring algorithm to rank the docking poses. The top ranked structures obtained after scoring were retained and evaluated.

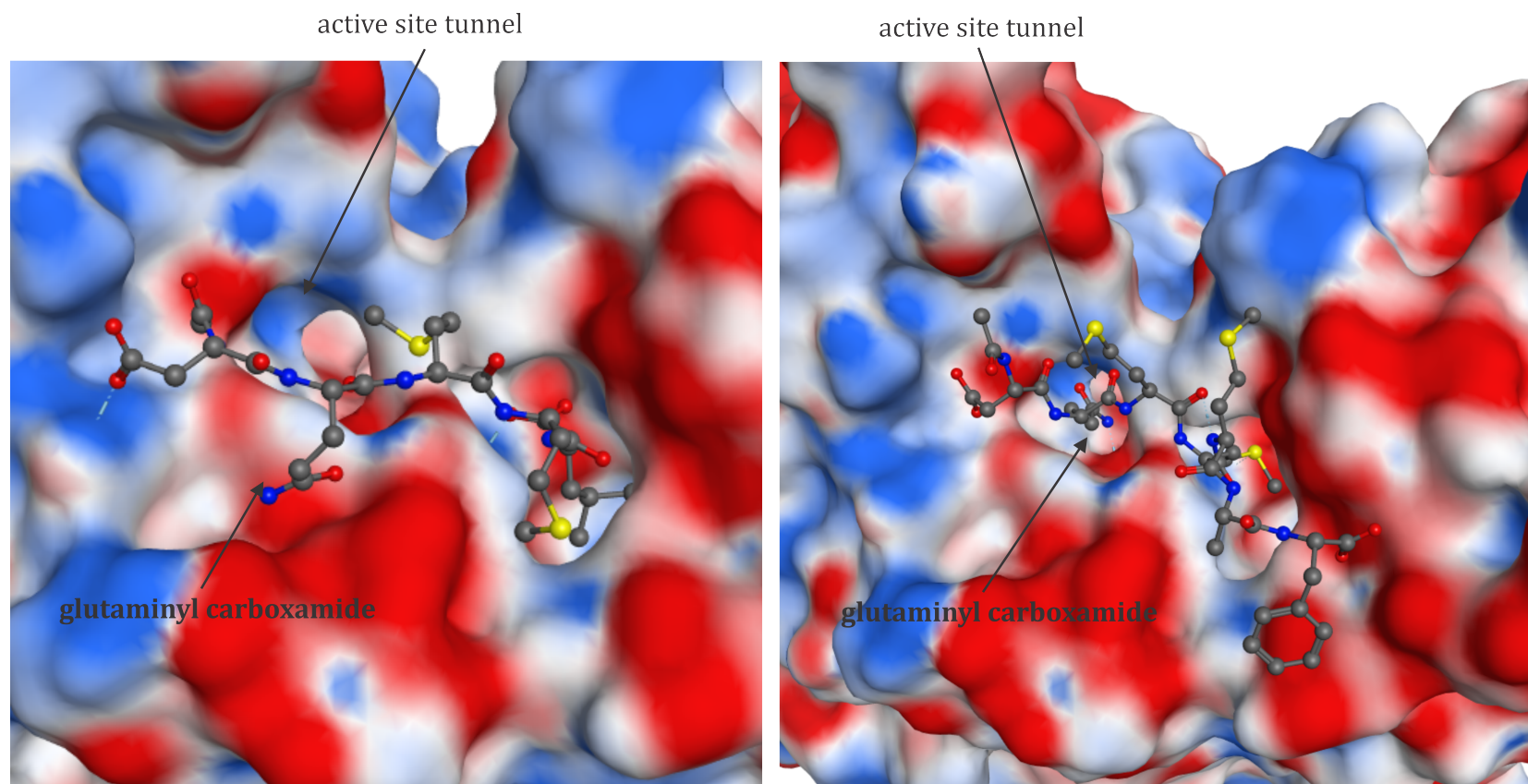


Figure 4.24: (left) Ac-DQMML docked into FXIIIa, (right) Ac-DQMMAF docked into FXIIIa.

From these docking results (Figure 4.24), it can be seen that the longer peptide is a better substrate for FXIIIa. The carboxamide group in Ac-DQMMAF is positioned directly in the active site tunnel and is only 3 Å away from the nucleophilic cysteine residue. This is possible because the third methionine in the sequence occupies a pocket close to the active site that pulls the other two methionine residues to the right. Steric hindrance causes the glutamyl carboxamide to rotate into the active site tunnel. In the Ac-DQMML peptide, the carboxamide is rotated outwards since it doesn't have the steric driving force of the longer peptide, as the two methionine residues are spaced out with the second one occupying a polar/neutral pocket composed of tyrosine, alanine, glutamine and serine residues. These results show a preference for the binding of a methionine residue over a different hydrophobic residue such as leucine, within the polar/neutral pocket adjacent to the active site tunnel. The aspartic acid residue also contributes to preferential binding through a binding interaction in a polar environment, composed of serine and tyrosine residues. This might indicate a preference for the aspartic acid as opposed to serine for increased affinity to FXIIIa. The docking results suggest an explanation for our experimental results as to why the longer peptide chain has higher affinity for FXIIIa.

4.4.6 Modelling of peptides with TG2

To understand the lack of affinity for TG2, both Ac-DQMML-OH and Ac-DQMMAF-OH were also docked onto the open form of TG2.

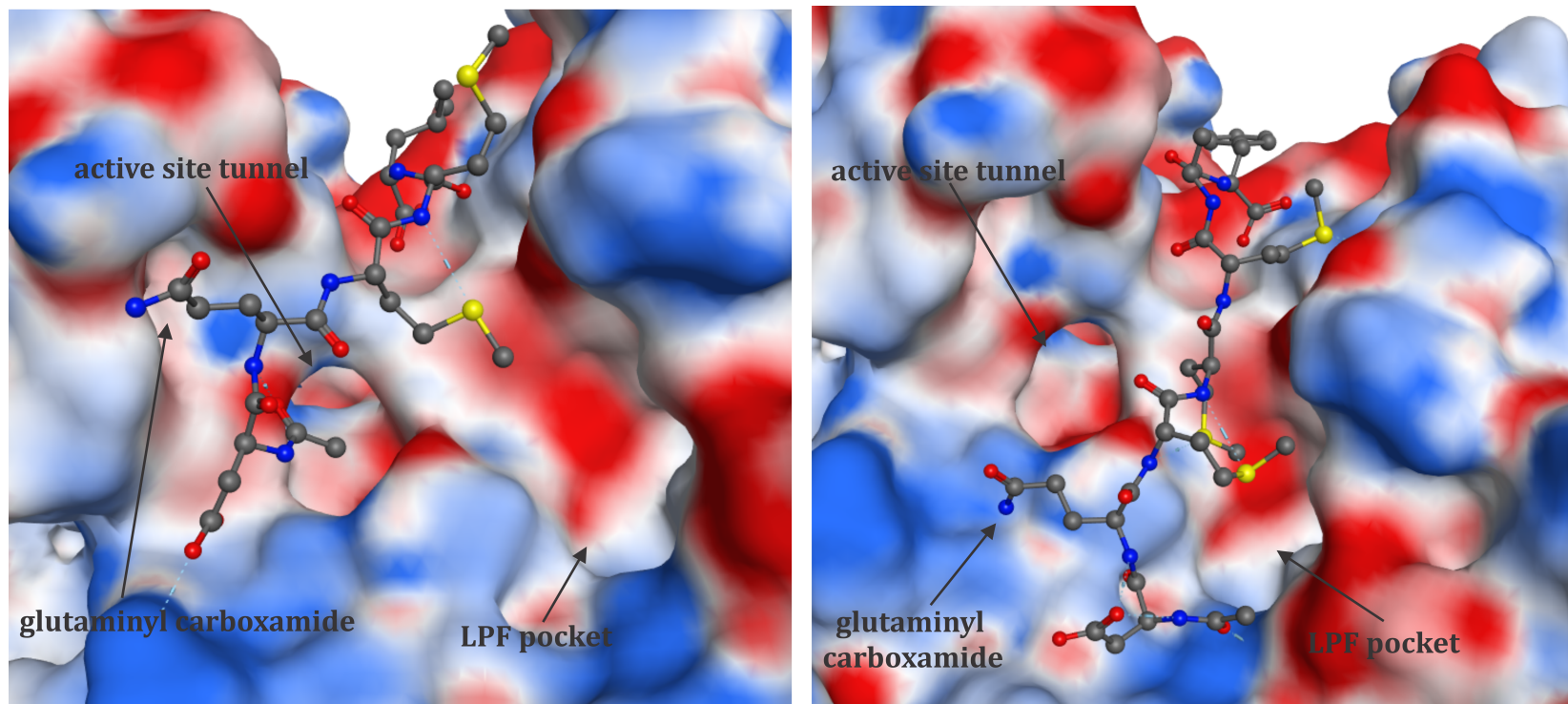


Figure 4.25: (left) Ac-DQMML docked into TG2, (right) Ac-DQMMAF docked into TG2.

The peptide was constrained to bind around the active site of TG2, by manually defining a coordinate box into which the substrate was docked, yet neither peptide binds in an optimal fashion. Both peptides fail to fill the LPF pocket, (so named because it is the binding pocket for the C-terminal LPF sequence of high affinity TG2 substrates) and are positioned more across the surface of the protein (Figure 4.25). While the carboxamide moiety of the Ac-DQMMAF-OH peptide is closer in proximity to the active site cysteine than the shorter Ac-DQMML-OH peptide, neither is in the active site tunnel. The docking results support the supposition that the aspartic acid is deleterious for effective TG2 catalysis. Docking results suggest the electronegative nature of the TG2 binding pocket, coupled with the presence of an aspartic acid residue in the peptide sequences, results in a repositioning of the carboxamide away from the TG2 active site tunnel. This results in the peptides binding more favourably to an electropositive environment near the active site that is not conducive to catalysis. The docking results are therefore consistent with the negligible reactivity of the two peptides observed with TG2. Thus, we report two truncated, kinetically characterized peptidic sequences orthogonal to TG2 with respect to FXIIIa.

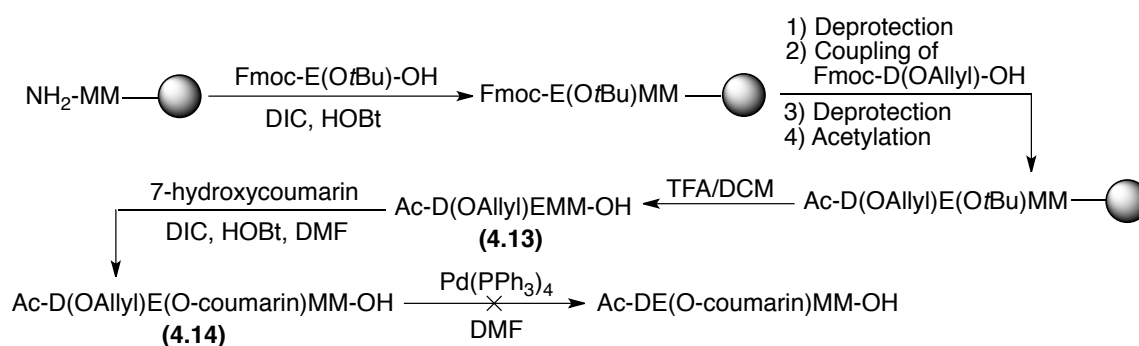
4.5 Towards the synthesis of Ac-DE(O-coumarin)MM-OH

As we were concurrently working towards a fluorogenic activity assay for FXIII, we decided to use the sequence Ac-DQMM-OH as a short, truncated, model sequence for ease of synthesis. The glutamine residue would be replaced with a glutamic acid, with a coumarin ester attached to the carboxylic acid side chain. The general idea behind the synthesis of the substrate was to use orthogonally protected aspartic and glutamic acid

residues. In this way, the glutamic acid could be selectively deprotected and the coumarin fluorophore coupled while leaving the aspartic carboxylic acid side chain untouched.

4.5.1 First attempt towards activity assay using Asp(O-allyl)-OH

Initially, commercially available amino acids were used and the entire peptide was synthesized through Fmoc-based solid-phase synthesis. The initial steps included coupling two methionine residues to Wang resin as per our standard peptide solid phase coupling procedure (Scheme 4.2). Secondly, glutamic acid was coupled with the side chain protected by a *tert*-butyl ester, labile under acidic conditions, the same conditions used to liberate the peptide from resin during the cleavage step.



Scheme 4.2: First attempt using Fmoc-Asp(O-allyl)-OH orthogonal to glutamic acid

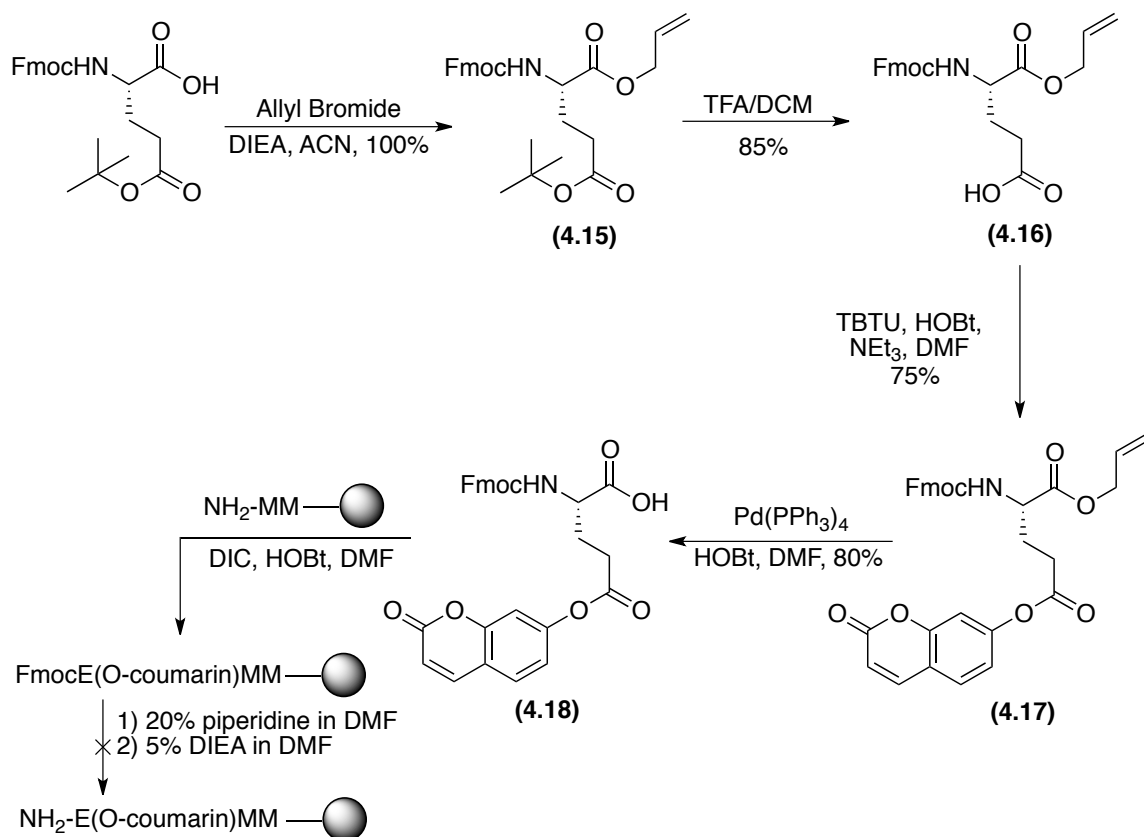
Lastly, aspartic acid was coupled to the glutamic acid peptide, where the aspartic acid side chain was protected using an allyl ester, which is not labile under acidic conditions but rather reducing conditions using a palladium catalyst.³² Thus, after acetylation of the last amino acid, the peptide was cleaved and the *tert*-butyl ester protecting group of

the glutamic acid concurrently removed while the aspartic acid carboxylic acid side chain was still protected with the allyl ester.

The free carboxylic acid was then coupled to 7-hydroxycoumarin and the coupling was verified through ESI-MS⁺. At this point, we tried to deprotect the allyl ester of the aspartic acid in order to obtain our target peptide. Multiple attempts using the tetrakis(triphenyl)phosphine palladium catalyst, freshly prepared or generated *in situ*³³, were unsuccessful, leading us to believe that the methionine sulphurs in our peptide were poisoning the catalyst and halting the deprotection.^{34,35}

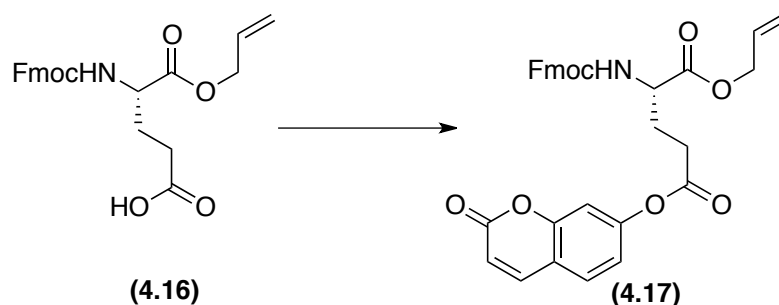
4.5.2 Second attempt towards activity assay using FmocGlu(O-coumarin)-OH

Our second attempt involved adding a coumarin derivatized glutamic acid directly as an unnatural amino acid in the synthesis of the peptide. Fmoc-Glu(OtBu)-OH is commercially available and in order to selectively add the coumarin molecule to the side chain carboxylic acid, it had to be modified through several protecting/deprotecting/reprotecting steps as there are two carboxylic acid moieties on the amino acid (Scheme 4.3).



Scheme 4.3: Second attempt using coumarin directly attached to glutamic acid

We chose the allyl protecting group for the α -carboxylic side chain since it was orthogonal to the *tert*-butyl ester (as previously discussed) and then deprotected the *tert*-butyl ester on the sidechain carboxylic acid using a mixture of trifluoroacetic acid in dichloromethane. Several reaction conditions were explored for the addition of 7-hydroxycoumarin to the carboxylic side chain of glutamic acid (Table 4.3).



Coupling Reagent	7-Hydroxycoumarin	Activator	Yield
1.2 eq DCC	2 eq.	0.3 eq. DMAP	10%
3 eq. DIC	5 eq.	3 eq. HOBT	21%
2 eq. EEDQ	2 eq.	N/A	29%
3 eq. TBTU	3 eq.	3 eq. HOBT + 3 eq. NEt ₃	61%

Table 4.3: Reaction conditions for 7-hydroxycoumarin coupling.

Through a comparison of the LC traces of the various reaction conditions, it was determined that TBTU was the optimal coupling reagent for this reaction (Table 4.3). Lastly, the allyl ester was deprotected using the tetrakis(triphenylphosphine) palladium catalyst.

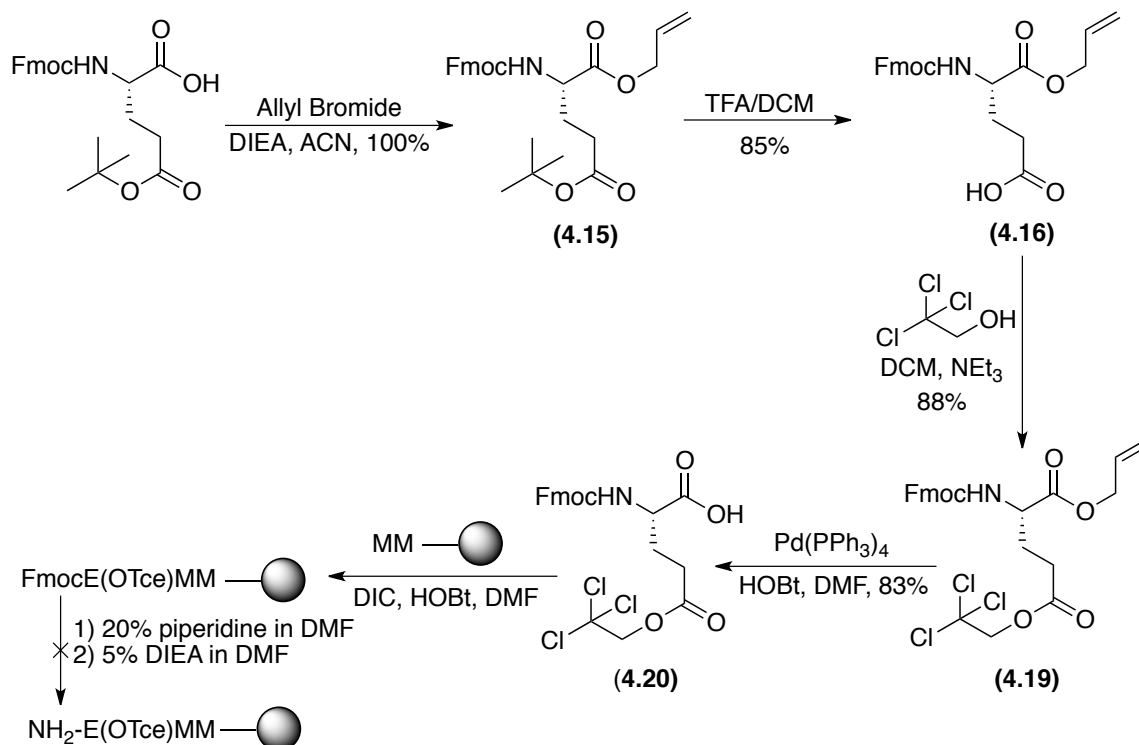
The coumarin-modified glutamic acid was then incorporated into the solid phase synthetic scheme, coupling it to methionine via standard Fmoc-based solid phase coupling procedures. The coupling was confirmed by taking a small aliquot, cleaving it and then testing it for the presence of the tripeptide using ESI-MS⁺ and the next step would then be deprotecting the Fmoc protecting group to continue with the synthesis.

Unfortunately, it was found that the normal Fmoc deprotection conditions (20% piperidine in DMF) were too harsh for the coumarin ester bond and the coumarin was cleaved during the Fmoc deprotection step. We attempted to minimize the deprotection time but we still observed that the majority of our product was lost.

Looking for alternative deprotection conditions, we found that Bogyo and coworkers had effected a survey of mild deprotection conditions for use in Fmoc-based solid phase synthesis with their electrophilic chloromethyl ketone as a modified side chain.³⁵ They found that the very mild conditions of 5% DIEA in DMF for 1 min were enough to effectively remove the Fmoc protecting group while leaving the chloromethyl ketone largely untouched.³⁶ However, with the same conditions, we still observed the cleavage of our coumarin fluorophore, leading us to abandon this method of synthesis.

4.5.3 Third attempt towards activity assay using FmocGlu(OTce)-OH

Our third attempt involved modifying the glutamic acid to have a trichloroethyl ester (Tce) protecting group on the sidechain of the carboxylic acid, which could be deprotected using samarium iodide, and would thus be orthogonal to the *tert*-butyl ester on the aspartic acid (Scheme 3).³⁷ Our goal was that after the entire peptide chain was synthesized, the trichloroethyl ester would be removed without affecting the *tert*-butyl ester protecting group of the aspartic acid. The coumarin molecule could then be coupled to the free carboxylic side chain of the glutamic acid, after which the cleavage of the entire peptide from resin would be carried out.



Scheme 4.4: Third attempt using trichloroethyl ester as protecting group

The trichloroethyl ester group, for which the standard deprotection consists of zinc with acetic acid,³⁸ was unexpectedly found to be cleaved after employing the various Fmoc deprotection conditions used for the previous method. This was verified using ^{13}C -NMR to monitor the characteristic $\text{C}(\text{Cl}_3)$ peak of the trichloro ethyl ester. However, the protected amino acid, Fmoc-Glu(OTce)-OAllyl amino acid smoothly underwent standard Fmoc deprotection conditions (20% piperidine in DMF), whereby neither the allyl protecting group nor the trichloroethyl ester group were affected. Thus, we hypothesize that as a result of attachment to the resin in a peptide chain, the trichloroethyl ester protecting group became susceptible to cleavage by the amine base used in Fmoc deprotections.

There exists an array of different protecting methods we could try, such as the 4- $\{N$ -[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]-amino} benzyl protecting group which is stable to both TFA and piperidine and cleaved with 2% hydrazine in DMF.³⁹ Incorporating this protecting group on the glutamic acid side chain would allow us to complete the entire synthesis and only after acetylation and complete synthesis of the peptide would it be cleaved and the glutamic acid coupled with 7-hydroxycoumarin. However, the lability of the glutamic acid ester bond when incorporated into a peptide is a cause of concern for the future implications of this activity assay.

Changing the linkage to an amide bond as opposed to an ester could solve the issue of bond lability; however, the prohibitive cost of 7-aminocoumarin detracts from this approach. Thus, in order for this assay to be biologically relevant while still maintaining affordability and relative simplicity, a quenched activity assay, using both a fluorophore and quencher through FRET would be more feasible (Figure 4.26). FXIIIa would cleave the quencher from the peptide to release the fluorophore, resulting in an increase in fluorescence.

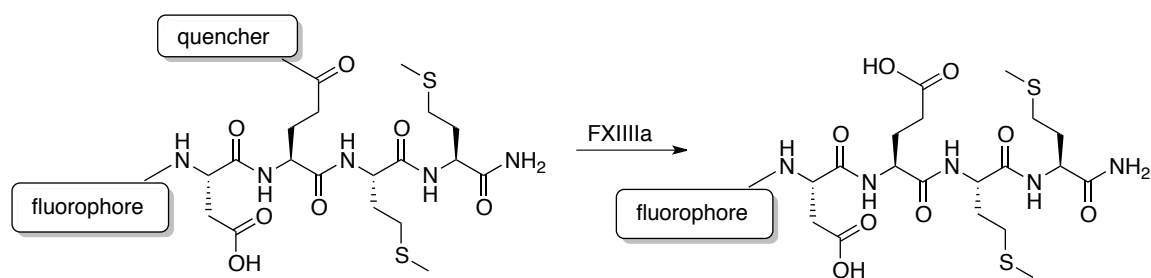


Figure 4.26: Schematic of quenched activity probe for FXIIIa.

4.6 Conclusions and Perspectives

After evaluating different conditions for FXIIIA expression, the Studier auto-induction ZYP media at 20 °C for 24 h was found to be the optimal conditions for the expression of recombinant GST-tagged FXIIIA, whereby a total of 1.5 mg of protein/L of culture was obtained.

A variety of different peptides were synthesized and tested using the GDH assay to identify an orthogonal sequence for FXIIa with respect to TG2. The longest peptide investigated, Ac-DQMMAF-OH, had the greatest affinity for FXIIa, consistent with the trend observed by Hitomi and coworkers, of a hydrophilic amino acid beside the glutamine residue on the N-terminus and a methionine residue on the C-terminus. The two highest affinity peptides, Ac-DQMMAF-OH and Ac-DQMML-OH were tested with TG2 and found to have negligible reactivity, confirming their use as orthogonal peptides. This finding represents the first time peptides orthogonal to TG2 with high affinity for FXIIa have been kinetically characterized with both transglutaminase enzymes.

Molecular modelling of the two peptides using a homology model for FXIIa were consistent with the greater affinity of Ac-DQMMAF-OH for FXIIa. The docking results suggested that the third methionine residue provided a steric pulling effect that caused the glutamine carboxamide moiety to enter the active site tunnel at a distance of 3 Å from the active site cysteine. Docking the peptides into the active site of TG2 were in agreement with our experimental results, with the peptide binding vertically across the groove close to the active site, positioning the carboxamide group away from the active site tunnel.

Lastly, we attempted to synthesize a fluorogenic activity assay by incorporating a coumarin ester through attachment to a glutamic acid residue into a peptide sequence recognized by FXIIIa. As the sequence included an aspartic acid, our approach was to use orthogonal protecting groups and then selectively deprotect the glutamic acid residue to couple the coumarin. When the final deprotection of the aspartic acid residue proved unsuccessful, a coumarin-derivatized glutamic acid residue was coupled directly as an unnatural amino acid. However, the ester bond was labile to the various Fmoc-deprotection conditions attempted and the coumarin fluorophore was cleaved in addition to the Fmoc protecting group. Finally, a different protecting group; trichloroethyl ester, cleaved under reducing conditions, was used for the glutamic acid but also proved susceptible to the mild Fmoc deprotection conditions.

Once incorporated into a peptide, the ester bond is more labile than anticipated which is worrisome for the future biological implications of this assay. To address this lability, a relevant fluorogenic assay for FXIIIa will consist of a quenched activity probe, whereby the cleavage action of FXIIIa will separate the quencher from the fluorophore, both of which are incorporated into a high affinity peptide, thus releasing fluorescence and signalling the catalytic activity of FXIIIa (Figure 4.26).

The fluorophore-quencher pair will be incorporated into a high-affinity peptide for FXIIIa. The importance of the serine residue in recognition has still yet to be determined and so the sequence Ac-SQMMMAF-OH remains to be tested with both FXIIIa and TG2. The peptide Ac-DQMMLPW-OH, a truncated version of Hitomi's peptide with the highest affinity for FXIIIa (F11) will also be investigated. Once an optimal sequence has been found through varying select residues, a fluorogenic activity

assay specific for FXIIIa and orthogonal to TG2 can be synthesized through the use of a fluorophore-quencher pair (Figure 4.26).

References

- (1) Hornyak, T. J.; Bishop, P. D.; Shafer, J. A. *Biochemistry* **1989**, *28*, 7326–7332.
- (2) Bangert, K.; Johnson, A. H.; Christensen, U.; Thorsen, S. *Biochemical Journal* **1993**, *291*, 623–625.
- (3) Ritchie, H.; Lawrie, L. C.; Crombie, P. W.; Mosesson, M. W.; Booth, N. A. *Journal of Biological Chemistry* **2000**, *275*, 24915–24920.
- (4) Ariens, R. A. S.; Lai, T.-S.; Weisel, J. W.; Greenberg, C. S.; Grant, P. J. *Blood* **2002**, *100*, 743–754.
- (5) Duckert, F. *Annals New York Academy of Sciences* **1972**, *202*, 190–199.
- (6) Sugimura, Y.; Hosono, M.; Wada, F.; Tohru, Y.; Maki, M.; Hitomi, K. *Journal of Biological Chemistry* **2006**, *281*, 17699–17706.
- (7) Schwartz, M. L.; Salvatore, P. V.; Hill, R. L.; McKee, P. A. *Journal of Biological Chemistry* **1971**, *246*, 5851–5854.
- (8) Loewy, A. G.; Dunathan, K.; Kriel, R.; Wolfinger, H. L. J. *Journal of Biological Chemistry* **1961**, *236*, 2625–2633.
- (9) Broker, M.; Bauml, O.; Gottig, A.; Ochs, J.; Bodenbenner, M.; Amman, E. *Applied Microbiology and Biotechnology* **1991**, *34*, 756–764.
- (10) Lai, T.-S.; Santiago, M. A.; Achyuthan, K. E.; Greenberg, C. S. *Protein Expression and Purification* **1994**, *5*, 125–132.
- (11) Lai, T.-S.; Achyuthan, K. E.; Santiago, M. A.; Greenberg, C. S. *Journal of Biological*

- Chemistry* **2001**, 269, 24596–24601.
- (12) Harper, S.; Speicher, D. W. In *Methods in Molecular Biology*; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2010; Vol. 681, pp. 259–280.
 - (13) Kusch, M.; Grundmann, C.; Keitel, S.; Seitz, R.; König, H. *Blood Coagulation and Fibrinolysis* **2006**, 17, 575–580.
 - (14) Curtis, C. G.; Brown, K. L.; B, C. R.; A, D.; Gray, A.; Stenberg, P.; Lorand, L. *Biochemistry* **1974**, 13, 1–7.
 - (15) Slaughter, T. F.; Achyuthan, K. E.; Lai, T.-S.; Greenberg, C. S. *Analytical Biochemistry* **1992**, 205, 166–171.
 - (16) Oertel, K.; Hunfeld, A.; Specker, E.; Reiff, C.; Seitz, R.; Pasternack, R.; Dodt, J. *Analytical Biochemistry* **2007**, 367, 152–158.
 - (17) Karpati, L.; Penke, B.; Katona, E.; Balogh, I.; Vamosi, G.; Muszbek, L. *Clinical Chemistry* **2000**, 46, 1946–1955.
 - (18) Ikura, K.; Tsuchiya, Y.; Sasaki, R.; Chiba, H. *European Journal of Biochemistry* **1990**, 187, 705–711.
 - (19) Gillet, S. M. F. G.; Pelletier, J. N.; Keillor, J. W. *Analytical Biochemistry* **2005**, 347, 221–226.
 - (20) Board, P. G.; Pierce, K.; Coggan, M. *Thrombosis and Haemostasis* **1990**, 63, 235–240.
 - (21) Roy, I.; Smith, O.; Clouthier, C. M.; Keillor, J. W. *Protein Expression and Purification* **2013**, 87, 41–46.
 - (22) Studier, F. W. *Protein Expression and Purification* **2005**, 41, 207–234.
 - (23) Cleary, D. B.; Maurer, M. C. *Biochimica et Biophysica Acta (BBA) - Proteins and*

- Proteomics* **2006**, 1764, 1207–1217.
- (24) Gorman, J. J.; Folk, J. E. *Journal of Biological Chemistry* **1980**, 255, 419–427.
 - (25) Fickenscher, K.; Aab, A.; Stuber, W. *Thrombosis and Haemostasis* **1991**, 65, 535–540.
 - (26) Stieler, M.; Weber, J.; Hils, M.; Kolb, P.; Heine, A.; Büchold, C.; Pasternack, R.; Klebe, G. *Angewandte Chemie International Edition* **2013**, 52, 11930–11934.
 - (27) Andersen, M. D.; Faber, J. H. *International Journal of Mass Spectrometry* **2011**, 302, 139–148.
 - (28) Mary, A.; Achyuthan, K. E.; Greenberg, C. S. *Archives of Biochemistry and Biophysics* **1998**, 261, 112–121.
 - (29) Woofter, R. T.; Maurer, M. C. *Archives of Biochemistry and Biophysics* **2011**, 512, 87–95.
 - (30) Komaromi, I.; Bagoly, Z.; Muszbek, L. *Journal of Thrombosis and Haemostasis* **2011**, 9, 9–20.
 - (31) Pinkas, D. M.; Strop, P.; Brunger, A. T.; Khosla, C. *PLoS ONE* **2007**, 5, 2788–2796.
 - (32) Ciommer, M.; Kunz, H. *Synlett* **1991**, 593–595.
 - (33) Tenaglia, A.; Waegell, B. *Tetrahedron Letters* **1988**, 29, 4851–4854.
 - (34) Miyauchi, Y.; Watanabe, S.; Kuniyasu, H.; Kurosawa, H. *Organometallics* **1995**, 14, 5450–5453.
 - (35) Savarin, C.; Srogl, J.; Liebeskind, L. S. *Organic Letters* **2000**, 2, 3229–3231.
 - (36) Kato, D.; Boatright, K. M.; Berger, A. B.; Nazif, T.; Blum, G.; Ryan, C.; Chehade, K. A. H.; Salvesen, G. S.; Bogoy, M. *Nature Chemical Biology* **2005**, 1, 33–38.
 - (37) Pearson, A. J.; Lee, K. *Journal of Organic Chemistry* **1994**, 59, 2304–2313.

- (38) Woodward, R. B.; Heusler, K.; Gosteli, J.; Naegeli, P.; Oppolzer, W.; Ramage, R.; Ranganathan, S.; Vorbruggen, H. *Journal of American Chemical Society* **1966**, *88*, 852–853.
- (39) Conroy, T.; Jolliffe, K. A.; Payne, R. J. *Organic & Biomolecular Chemistry* **2009**, *7*, 2255.

CHAPTER 5

Perspectives

Enzymes are responsible for the thousands of metabolic processes that sustain life and any malfunction can lead to the development of a number of disease states, highlighting the importance of these biomolecules. The study of enzymes to better understand their function is facilitated through the use of tools such as protein labelling, enzyme inhibition, and substrate analogue kinetic studies among others. An enzyme family of interest in the Keillor group is the transglutaminases, particularly tissue transglutaminase (TG2) and Factor XIIIa (FXIIIa). TG2 has been implicated in a variety of diseases including Alzheimer's,¹ Huntington's² and Celiac's disease³ while FXIIIa has been implicated in bleeding diathesis⁴ and atherothrombosis⁵ among others.

With the aim of further investigating TG2 activity, our first goal was to synthesize transition state analogue inhibitors of TG2. The electronic and spatial similarities between the transition states of the TG2 acylation reactions and compounds 2.9 and 2.10 (Figure 2.8) suggest they would be effective transition state analogues.

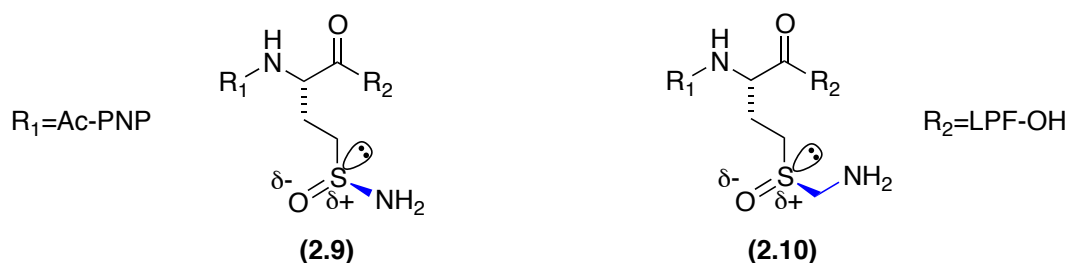


Figure 2.8: Transition state analogues for TG2

However, their synthesis proved difficult and ultimately, we were not able to isolate the desired compounds due to their aforementioned instability. While novel intermediate peptidic sulfoxide compounds were synthesized, our overall targets were not obtained.

Though we did expect the synthesis to be difficult, the degree of instability of these sulfinamide compounds was unanticipated.

An alternative isostere for the amine group is an alcohol and so it would be interesting to synthesize and test sulfinic acid derivatives 2.26 and 2.27 (Figure 2.19) in terms of their potential to inhibit TG2.

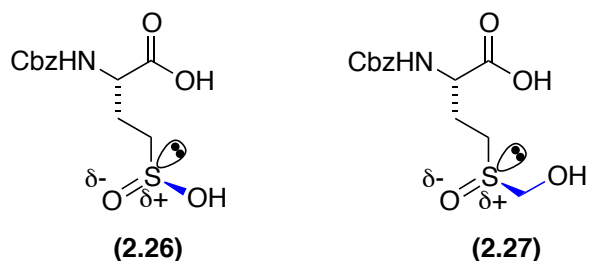


Figure 2.19: Sulfinic acid derivatives as TS analogues

Secondly, conflicting roles have been presented in regards to the intracellular role of TG2.⁶⁻¹⁰ Physiologically, it would seem likely that the cross-linking activity of TG2 would be inhibited in the low concentrations of intracellular calcium but this has not been conclusively proven. In order to investigate the intracellular cross-linking activity of TG2 to shed light on its varied roles, we attempted to synthesize a fluorescent, cell-permeable, TG2-specific probe (compound 3.12, Figure 5.1).

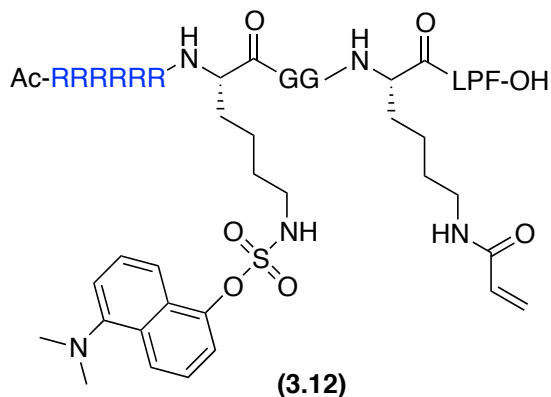


Figure 5.1: TG2-Specific, cell permeable, fluorescent probe.

In vivo cell imaging using SH-SY5Y neuroblastoma cells showed that while the probe was fluorescent and cell-permeable, it was not TG2-specific. Molecular modelling suggests that the probe may be positioned in the opposite way intended, with the hexa-arginine sequence binding in the pocket intended for the TG2 affinity-conferring LPF sequence, thus orienting the electrophilic acrylic amide warhead far from the active site cysteine residue.

Modified (compound 3.16, Figure 5.2) and branched (compound 3.17, Figure 5.2) versions of the peptide should confer both permeability and TG2 affinity in our efforts to understand TG2 activity *in vivo*.

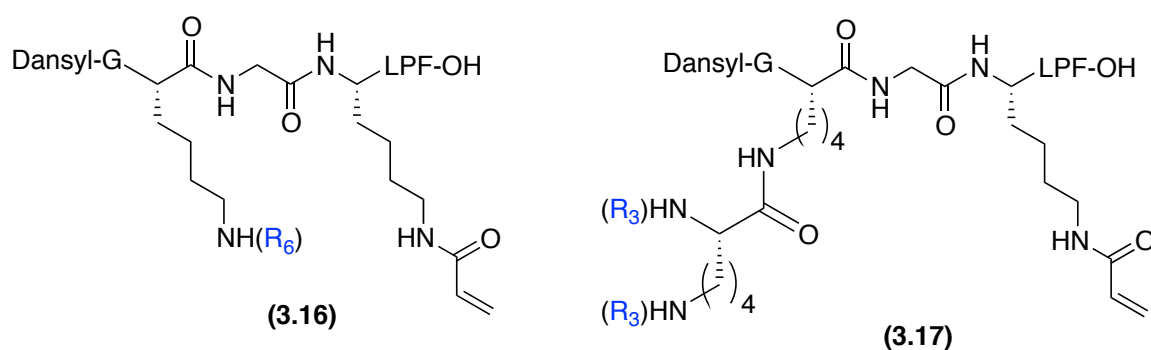


Figure 5.2: Modified (3.16) and branched (3.17) cell-permeable, fluorescent TG2-specific probes.

The reconfiguration of peptide attachment remains an avenue to pursue in order to adjust and improve our efforts towards a cell permeable, fluorescent TG2 probe.

Lastly, FXIII has become a new target transglutaminase in the Keillor group. Given our interest in this enzyme, we had three goals for this work. Firstly, owing to the anticipated high demand for FXIIIa required for later experiments, our primary aim

was the development of an optimized method for the expression of recombinant FXIIIa. Through investigation of different conditions for FXIIIa expression, the Studier auto-induction ZYP media¹¹ at 20 °C for 24 hours provided a total of 1.5 mg of protein/L of culture and comprised the best conditions for the expression of recombinant GST-tagged FXIIIa

Secondly, a variety of different peptides were synthesized and tested using the GDH assay to identify a high affinity sequence for FXIIIa. The longest peptide investigated, Ac-DQMMMAF-OH, had the greatest affinity for FXIIIa. This is consistent with the trend observed by Hitomi and coworkers,¹² where high affinity sequences bear a hydrophilic amino acid N-terminal to the glutamine residue and a methionine residue C-terminal to the glutamine. The two highest affinity peptides, Ac-DQMMMAF-OH and Ac-DQMML-OH were tested with TG2 and found to have negligible reactivity, confirming their use as orthogonal peptides, results supported by modelling studies of the peptides with both FXIIIa and TG2. This finding represents the first time peptides orthogonal to TG2 with affinity for FXIIIa have been kinetically characterized with both transglutaminase enzymes.

Finally, we attempted to synthesize a fluorogenic activity assay by incorporating a coumarin ester through attachment to a glutamic acid residue into a peptide sequence recognized by FXIIIa. However, it was found that once incorporated into a peptide, the ester bond is more labile than anticipated, which is worrisome for the future biological implications of this assay. To address this lability, a relevant fluorogenic assay for FXIIIa will consist of a quenched activity probe whereby the cleavage action of FXIIIa will separate the quencher from the fluorophore, both of which are incorporated into a

high affinity peptide, thus releasing fluorescence and signalling the catalytic activity of FXIIIa (Figure 4.26).

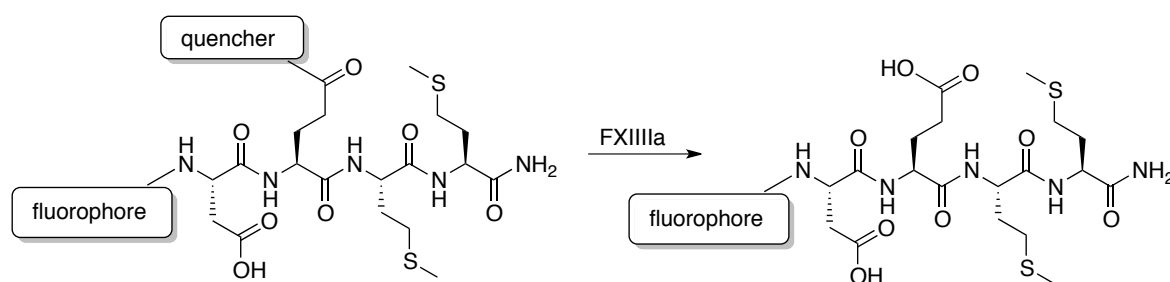


Figure 4.26: Schematic of quenched-activity probe for FXIIIa

The importance of the serine residue in recognition has still yet to be determined and so the sequence Ac-SQMMMAF-OH remains to be tested with both FXIIIa and TG2. The peptide Ac-DQMMLPW-OH, a truncated version of Hitomi's peptide with the highest affinity for FXIIIa (F11)¹² will also be investigated. Once an optimal sequence has been found through varying select residues, a fluorogenic activity assay specific for FXIIIa and orthogonal to TG2 can be synthesized through the use of a fluorophore-quencher pair (Figure 4.26).

References

- (1) Johnson, G. V. W.; Cox, T. M.; Lockhart, J. P.; Zinnerman, M. D.; Miller, M. L.; Powers, R. E. *Brain Research* **1997**, *751*, 323–329.
- (2) DiFiglia, M.; Sapp, E.; Chase, K. O.; Davies, S. W.; Bates, G. P.; Vonsattel, J. P.; Aronin, N. *Science* **1997**, *277*, 1990–1993.
- (3) Dieterich, W.; Ehnis, T.; Bauer, M.; Donner, P.; Volta, U.; Riecken, E. O.; Schuppan, D. *Nature Medicine* **1997**, *3*, 797–801.
- (4) Duckert, F. *Annals New York Academy of Sciences* **1972**, *202*, 190–199.
- (5) Ariens, R. A. S.; Lai, T.-S.; Weisel, J. W.; Greenberg, C. S.; Grant, P. J. *Blood* **2002**, *100*, 743–754.
- (6) Bergamini, C. M.; Collighan, R. J.; Wang, Z. *Advances in Enzymology and Related Areas of Molecular Biology* **78**, 1–46.
- (7) Filiano, A. J.; Bailey, C. D. C.; Tucholski, J.; Gundemir, S.; Johnson, G. V. W. *The FASEB Journal* **2008**, *22*, 2662–2675.
- (8) Maiuri, L.; Luciani, A.; Giardino, I.; Raia, V.; Vilella, V. R.; D'Apolito, M.; Pettoello-Mantovani, M.; Guido, S.; Ciacci, C.; Cimmino, M.; Cexus, O. N.; Londei, M.; Quarantino, S. *Journal of Immunology* **2008**, *180*, 7697–7705.
- (9) Milakovic, T.; Tucholski, J.; McCoy, E.; Johnson, G. V. W. *Journal of Biological Chemistry* **2004**, *279*, 8715–8722.
- (10) Fesus, L.; Thomazy, V.; Falus, A. *FEBS Letters* **1987**, *224*, 104–108.
- (11) Studier, F. W. *Protein Expression and Purification* **2005**, *41*, 207–234.
- (12) Sugimura, Y.; Hosono, M.; Wada, F.; Tohru, Y.; Maki, M.; Hitomi, K. *Journal of Biological Chemistry* **2006**, *281*, 17699–17706

CHAPTER 6

EXPERIMENTAL

6.1 Experimental Instrumentation

The source of solvents and chemical reagents of each chapter will be discussed in the section dedicated to that chapter.

Nuclear magnetic resonance (NMR) spectra were recorded in the solvent indicated with tetramethylsilane (TMS) as internal reference or using the residual proton or ^{13}C signal of the solvent as a reference at ambient temperature. The spectra were recorded on a Bruker Avance 300 Fourier Transform Spectrometer [300 MHz (^1H), 75 MHz (^{13}C)], Bruker Avance 400 Fourier Transform Spectrometer [400 MHz (^1H), 100 MHz (^{13}C)], Bruker AMX-300 Fourier Transform Spectrometer [300 MHz (^1H), 75 MHz (^{13}C) or a Bruker AMX-400 Fourier Transform Spectrometer [400 MHz (^1H), 100 MHz (^{13}C)].

Mass spectra were determined by FAB^+ ionization on an AutoSpec Q spectrometer, a Waters Micromass Q-ToF mass spectrometer or a Bruker Microflex MALDI-TOF mass spectrometer.

Melting points were measured on an EZ-Melt automated melting point apparatus and are uncorrected.

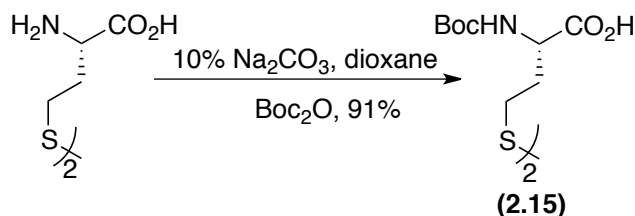
Reactions requiring anhydrous conditions were carried out under a dry nitrogen atmosphere using conventional benchtop techniques.

6.2 Sulfoxide Inhibitors

Reagents, unless otherwise specified, were bought from Sigma-Aldrich Canada Ltd. The common solvents were bought from EMD Chemicals. The compound, L-homocystine, was synthesized according to a literature procedure.¹ Anhydrous solvents (THF, Et₂O, CH₂Cl₂, DMF, CH₃CN, toluene) were obtained by filtration through drying columns on a GlassContour system (Irvine, CA).

6.2.1 Synthesis of sulfinamide via a methyl sulfinate

(2*S*,2'*S*)-4,4'-disulfanediylbis(2-((*tert*-butoxycarbonyl)amino)butanoic acid)
(2.15).



L-Homocystine (0.1 g, 0.37 mmol) was added to 9 mL of a solution of 10% NaCO₃ and dioxane (5:4). Boc anhydride (0.24 g, 1.1 mmol) was added on ice and the solution allowed to warm to room temperature and then stirred overnight. The solution was then neutralized to pH 4 using a 10% citric acid solution and an extraction with ethyl acetate was performed (3 × 15 mL). The organic layers were combined, washed with brine and then dried over magnesium sulfate. The organic solvent was evaporated under vacuum to yield a white solid in 91% yield.

M.P.: 171-175 °C

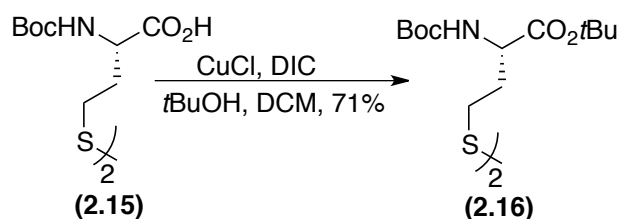
¹H NMR (300 MHz, MeOD) δ (ppm): 4.25-4.19 (m, 2H), 2.87-2.70 (m, 4H), 2.30-2.21 (m, 2H), 2.04-1.93 (m, 2H), 1.45 (s, 18H).

¹³C NMR (100 MHz, MeOD) δ (ppm): 174.3, 156.8, 79.2, 52.3, 34.3, 31.1, 27.4.

HRMS (m/z): [M+Na] Calculated: 491.1498

Found: 491.1487

(2*S*,2'*S*)-di-*tert*-butyl 4,4'-disulfanediylbis(2-((*tert*-butoxycarbonyl)amino)butanoate (2.16)



Both *t*-butanol (0.16 mL, 2.13 mmol) and *N,N*-diisopropylcarbodiimide (DIC) (0.05 mL, 0.32 mmol) were stirred overnight in the presence of a catalytic amount of CuCl (2 mg). The next day, compound **2.15** was added to the solution (0.1 g, 0.23 mmol) along with 5 mL of DCM and the solution refluxed for an additional 3 days. After the third day, the solvent was evaporated and the product purified using column chromatography (9:1 hexanes:ethyl acetate) to yield a white, gummy solid in 71% yield.

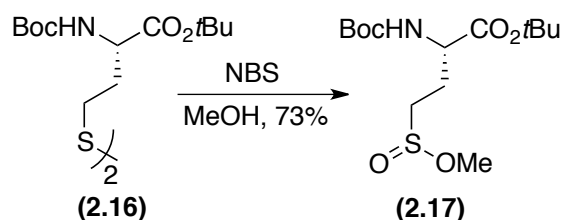
^1H NMR (400 MHz, CDCl_3) δ (ppm): 5.11 (d, 2H, $J = 8$ Hz), 4.26-4.24 (m, 2H), 2.69-2.66 (m, 4H), 2.23-2.17 (m, 2H), 1.47 (s, 18H), 1.44 (s, 18H).

^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 172.0, 156.1, 83.1, 80.7, 70.2, 54.0, 33.7, 29.1, 28.8.

HRMS (m/z): $[\text{M}+\text{Na}]$ Calculated: 587.1492

Found: 587.1517

(2*S*)-*tert*-butyl 2-((*tert*-butoxycarbonyl)amino)-4-(methoxysulfinyl)butanoate
(2.17)



To 10 mL of MeOH was added compound **2.16** (0.15 g, 0.2 mmol) and subsequently *N*-bromosuccinimide (NBS) (0.05 g, 0.3 mmol). The reaction was stirred overnight and the reaction mixture was diluted with DCM. The organic layer was washed with NaHSO_3 (3×10 mL), and NaHCO_3 (3×10 mL). It was then dried over magnesium sulfate, and evaporated under vacuum to yield a clear oil in 73% oil.

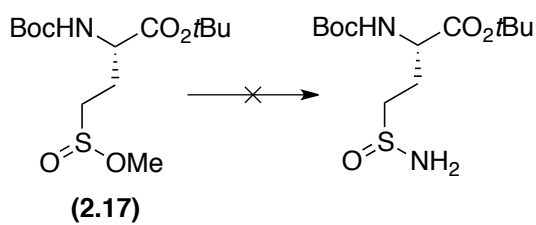
^1H NMR (300 MHz, CDCl_3) δ (ppm): 5.14 (d, 1H, $J = 6$ Hz), 4.27-4.25 (m, 1H), 3.77 (s, 3H), 2.82-2.76 (m, 2H), 2.28-2.23 (m, 1H), 2.05-2.00 (m, 1H), 1.48 (s, 9H), 1.47 (s, 1H)

^{13}C NMR (75 MHz, CDCl_3) δ (ppm): 171.1, 155.2, 83.5, 82.5, 54.4, 53.1, 34.3, 28.2, 27.9.

HRMS (m/z): $[\text{M}+\text{Na}]$ Calculated: 360.14513

Found: 360.14480

Formation of sulfinamide from compound 2.17



Attempt	Condition	Result
1	1))LiHMDS, -78°C 2) NH_4Cl sat.	product decomposition
2	excess NH_3 in MeOH/THF, reflux	product decomposition
3	3 eq. NaH, THF, excess NH_3	product decomposition
4	$n\text{BuLi}$, excess NH_3	product decomposition
5	$\text{Fe}(\text{NO}_3)_3 \cdot 9\text{H}_2\text{O}$, Na, NH_3 , NH_4Cl	product decomposition

Table 2.1: Conditions attempted in synthesis of sulfinamide.

Attempt #1. To a solution of the methyl sulfinate (0.24 g, 0.70 mmol) in 10 mL of dry THF was added a 1 M solution of LiHMDS (0.18 mL, 1.05 mmol) at -78°C. The reaction was warmed to room temperature and stirred overnight. The reaction was quenched with 10 mL of a saturated NH_4Cl solution and stirred for one hour. An

extraction was performed with ethyl acetate (3×15 mL) and the organic phase was dried over magnesium sulfate. Evaporation under vacuum yielded a yellow, gummy solid. Proton NMR revealed decomposition.

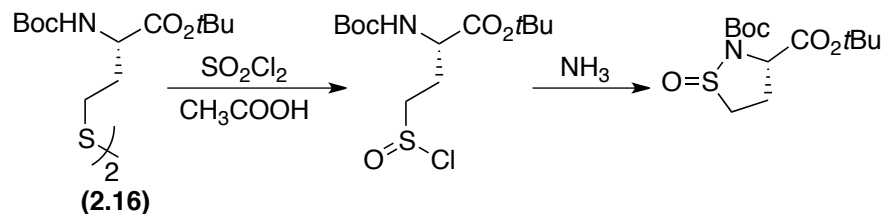
Attempt #2. Methyl sulfinatate (0.1 g, 0.29 mmol) was dissolved in a 5 mL solution of a 50:50 mixture of dry MeOH and THF. To this solution, an excess of ammonia was added using a dry ice condensor attached to an ammonia gas cylinder. The reaction was refluxed for 3 h at which point the solvent was removed under vacuum. An extraction was undertaken using ethyl acetate (3×15 mL) and the organic phase washed with 10% citric acid and water and then dried over magnesium sulfate. Evaporation under vacuum yielded a yellow oil. Proton NMR revealed decomposition.

Attempt #3. To a three-necked round-bottomed flask, equipped with a dry ice condensor, was added dry THF. Ammonia was bubbled through the solvent for 30 min while the solution was on ice. Sodium hydride was then added to the THF/ammonia solution and this was stirred for 5 min to produce a cloudy solution. The methyl sulfinatate (0.180 mg, 0.53 mmol) dissolved in 5 mL of dry THF was added dropwise to the flask. After addition the reaction was refluxed for 3 h after which the solvent was removed under vacuum. An extraction was performed using ethyl acetate (3×15 mL) and the organic phase washed with 10% citric acid, water and brine and then dried over magnesium sulfate. Evaporation under vacuum yielded a yellow oil. Proton NMR revealed decomposition.

Attempt #4. To a three-necked round-bottomed flask, equipped with a dry ice condensor, was added dry THF. Ammonia was bubbled through the solvent for 30 min while the solution was on ice. *n*-BuLi was then added to the THF/ammonia solution and this was stirred for five minutes to produce a cloudy solution. The methyl sulfinatate (0.180 mg, 0.53 mmol) dissolved in 5 mL of dry THF was added dropwise to the flask. After addition, the reaction was refluxed for 3 h after which the solvent was removed under vacuum. An extraction was performed using ethyl acetate (3 × 15 mL) and the organic phase washed with 10% citric acid, water and brine and then dried over magnesium sulfate. Evaporation under vacuum yielded a yellow oil. Proton NMR revealed decomposition.

Attempt #5. To an ammonia solution was added $\text{Fe}(\text{NO}_3)_3 \cdot 9\text{H}_2\text{O}$ (1.5 mg, 0.004 mmol) and Na (0.088 g, 3.8 mmol). The mixture was stirred for one hour and cooled to -48 °C. At this point, a solution of methyl sulfinatate (0.25 g, 0.76 mmol) dissolved in 5 mL of dry THF was added dropwise to the freshly prepared sodium amide solution. After 1 h of stirring, ammonium chloride (0.22 g, 0.004 mmol) was added to the reaction. The reaction was stirred overnight after which ethyl acetate (30 mL) was added to the reaction mixture. The organic phase was washed with 10% citric acid (3 × 20 mL), brine (3 × 20 mL) and then dried over magnesium sulfate. Evaporation under vacuum yielded a yellow oil. Proton NMR revealed decomposition.

6.2.2 Synthesis of sulfinamide via a sulfinyl chloride



A stirred mixture of compound **2.16** (0.09g 0.16 mmol) and acetic acid (0.018 mL, 0.32 mmol) in DCM (5 mL) at -10 °C was treated dropwise with SO₂Cl₂ (0.04 mL, 0.48 mmol). The temperature was maintained at 0-5 °C with stirring for 30 min, at which point the solution was evaporated, leaving a pale, yellow oil. The oil was used without further purification, an IR spectrum being taken to confirm the presence of the sulfinyl chloride. The flask was equipped with a dry ice condensor and ammonia was added dropwise to the flask at -10 °C. The reaction was stirred for 30 min at -10 °C, at which point the condensor was removed and the ammonia left to evaporate in the hood. NMR spectra proved inconclusive to identify the product with only a low resolution mass spectrum analysis supporting the hypothesis of an oxazolidinone ring formed.

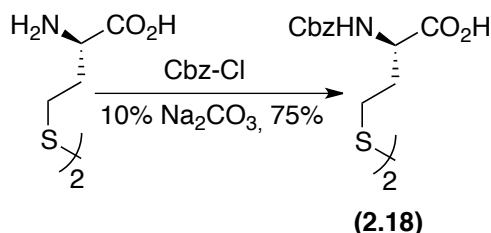
IR (KBr): 769 cm⁻¹

LRMS (m/z): [M+Na] Calculated: 328.1

Found: 328.1

6.2.3 Synthesis of sulfinamide via oxazolidinone

4,4'-disulfanediylbis(2-(((benzyloxy)carbonyl)amino)butanoic acid) (2.18)



To a solution of L-homocysteine (1 g, 0.046 mmol) dissolved in 10% Na₂CO₃ (40 mL) was added benzyl chloroformate (1.3 mL, 0.019 mmol). The reaction was stirred overnight, after which an extraction with ether was performed. The basic layer was collected and acidified to pH 2 and an extraction with ethyl acetate undertaken. The organic layer was collected, dried over magnesium sulfate and then evaporated under vacuum to yield a gummy solid in 75% yield, which was used without further purification.

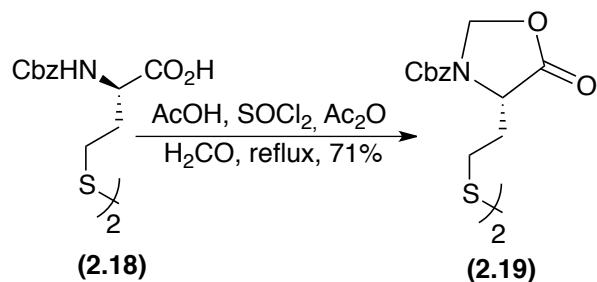
¹H NMR (400 MHz, MeOD) δ (ppm): 7.36-7.295 (m, 10H), 5.11 (s, 4H), 4.36-4.4.33 (m, 2H), 2.82-2.2.72 (m, 4H), 2.29-2.26 (m, 2H), 2.06-2.03 (m, 2H).

¹³C NMR (100 MHz, CDCl₃) δ (ppm): 177.0, 156.6, 135.8, 128.6, 128.5, 128.3, 128.1, 128.0, 69.3, 32.4, 29.7.

HRMS (m/z): [M+H] Calculated: 535.1209

Found: 535.1211

(4*S*,4'*S*)-dibenzyl 4,4'-(disulfanediylbis(ethane-2,1-diyl))bis(5-oxooxazolidine-3-carboxylate) (2.19)



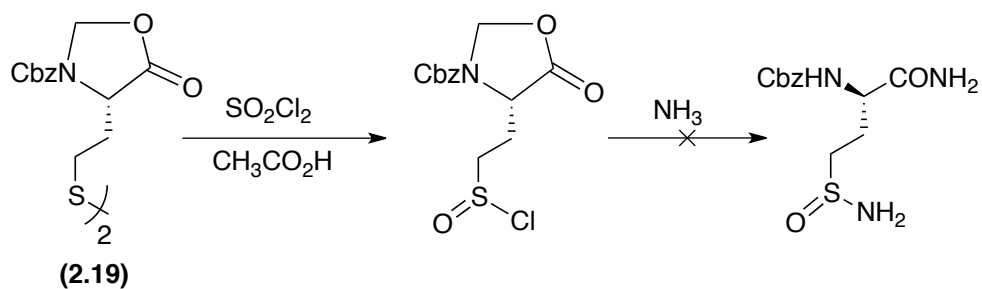
To a stirred solution of compound **2.18** (0.2 g, 0.37 mmol) in 5 mL of acetic acid was added paraformaldehyde (0.13 mg, 4.45 mmol), acetic anhydride (0.17 mL, 1.85 mmol) and thionyl chloride (0.013 mL, 0.17 mmol). The reaction was refluxed overnight and the acetic acid evaporated under vacuum. The product was purified using column chromatography (6:1:3 hexanes:toluene:ethyl acetate) to yield an oil in 71% yield.

¹H NMR (400 Hz, CDCl₃) δ (ppm): 7.42-7.38 (m, 10H), 5.56-5.55 (m, 2H), 5.28-5.27 (m, 2H) 5.18 (s, 4H), 4.5-4.43 (m, 2H), 2.79-2.75 (m, 2H), 2.33-2.30 (m, 2H).

¹³C NMR (100 MHz, CDCl₃) δ (ppm): 170.3, 162.3, 134.5, 129.2, 128.6, 127.7, 127.0, 125.1, 93.4, 60.2, 33.6, 29.8.

LRMS (m/z): [M+H] Calculated: 561.14

Found: 561.2

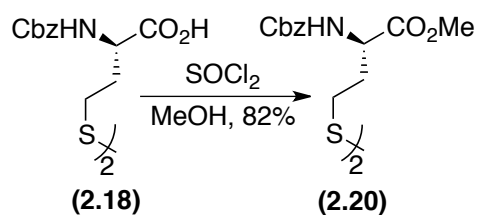


A stirred mixture of compound **2.19** (0.09 g, 0.16 mmol) and acetic acid (0.018 mL, 0.32 mmol) in DCM (5 mL) at -10 °C was treated dropwise with SO₂Cl₂ (0.04 mL, 0.48 mmol). The temperature was maintained at 0-5 °C with stirring for 30 min, at which point the solution was evaporated, leaving a pale yellow oil. The oil was used without further purification, an IR spectrum being taken to confirm the presence of the sulfinyl chloride. The flask was equipped with a dry ice condensor and ammonia was added dropwise to the flask at -10 °C. The reaction was stirred for 30 min at -10 °C, at which point the condensor was removed and the ammonia left to evaporate in the hood. Proton NMR revealed product decomposition.

6.2.4 Synthesis of amino methyl sulfoxide

Dimethyl 4,4'-disulfanediylbis(2-(((benzyloxy)carbonyl)amino)butanoate)

(2.20)



To a 100-mL round-bottomed flask containing 40 mL of MeOH, was dissolved compound **2.18** (1 g, 1.85 mmol). To this solution was added thionyl chloride (0.81 mL, 11.1 mmol) on ice and stirring was continued at room temperature overnight. The solvent was evaporated and the product purified using column chromatography (hexanes:ethyl acetate 7:3) to yield a white solid in 82% yield.

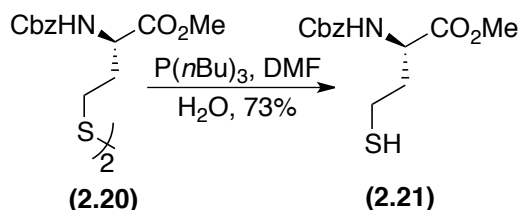
M.P.: 53-57 °C

¹H NMR (300 MHz, CDCl₃) δ (ppm): 7.35 (br. s., 10H), 5.5 (d, 2H), 5.11 (s, 4H), 4.51-4.50 (m, 2H), 3.76 (s, 6H), 2.72-2.69 (m, 4H), 2.27-2.24 (m, 2H), 2.08-2.04 (m, 2H).

¹³C NMR (75 MHz, CDCl₃) δ (ppm): 173.13, 172.02, 136.9, 129.38, 129.1, 129.0, 68.0, 53.7, 53.5, 35.2, 33.2.

HRMS (m/z): [M+Na] Calculated: 587.1498

Found: 587.1514

Methyl 2-(((benzyloxy)carbonyl)amino)-4-mercaptobutanoate (2.21).

In a 100-mL round-bottomed flask, compound **2.20** (0.82 g, 1.45 mmol) was dissolved in 7 mL of DMF followed by the addition of water (0.2 mL). To this solution was added tributyl phosphine (0.71 mL, 2.9 mmol) and the reaction was stirred overnight at room temperature. Water (100 mL) was then added to the reaction mixture and an extraction performed using ethyl acetate (3 × 30 mL). The organic layers were combined, washed with brine, dried over magnesium sulfate and evaporated under vacuum. The product was purified using column chromatography (6:4 hexanes: ethyl acetate) to yield a clear oil in 73% yield.

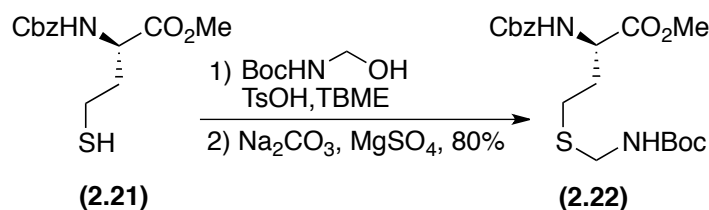
¹H NMR (400 MHz, CDCl₃) δ (ppm): 7.38 (br. s, 5H), 5.36-5.35 (m, 1H), 5.14 (s, 2H), 4.58-4.57 (m, 1H), 3.78 (s, 3H), 2.62-2.58 (m, 2H), 2.20-2.19 (m, 1H), 2.02-1.99 (m, 1H).

¹³C NMR (75 MHz, CDCl₃) δ (ppm): 173.13, 175.9, 136.9, 129.4, 129.1, 127.6, 127.1, 68.0, 53.4, 53.0, 35.2, 33.2.

HRMS (m/z): [M+Na] Calculated: 306.0770

Found: 306.0785

methyl 2-(((benzyloxy)carbonyl)amino)-4-((((tert-butoxycarbonyl)amino)methyl) thio)butanoate (2.22).



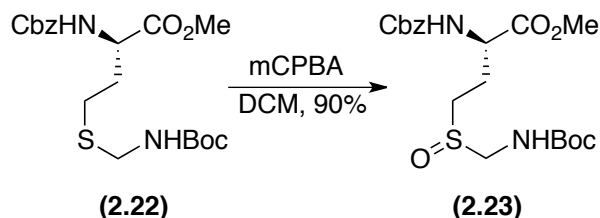
Compound **2.21** (0.58 g, 1.84 mmol) was dissolved in TBME in a 100-mL round-bottomed flask. To the reaction mixture was added N-Boc-aminomethanol (0.32 g, 2.2 mmol) and *p*-toluenesulfonic acid hydrate (0.035 g, 0.18 mmol) and the reaction was stirred overnight at room temperature. The next morning, sodium hydrogen carbonate (0.03 g, 0.37 mmol) and magnesium sulfate (0.11 g, 0.92 mmol) were added to the reaction mixture and it was stirred for another 2 hours. After removal of the solids by filtration, the solvent was evaporated and the product was purified through silica gel chromatography (7:3 hexanes:ethyl acetate) as a transparent oil in 80% yield.

¹H NMR (400 MHz, CDCl₃) δ (ppm): 7.34 (br. s, 5H), 5.1 (s, 2H), 4.67-4.65 (m, 1H), 4.34 (s, 2H), 3.74 (s, 3H), 2.67-2.59 (m, 2H), 2.21-2.17 (m, 2H), 1.44 (s, 9H)

LRMS (m/z): [M+Na] Calculated: 435.1

Found: 435.1

methyl 2-(((benzyloxy)carbonyl)amino)-4-((((tert-butoxycarbonyl)amino)methyl)sulfinyl)butanoate (2.23)



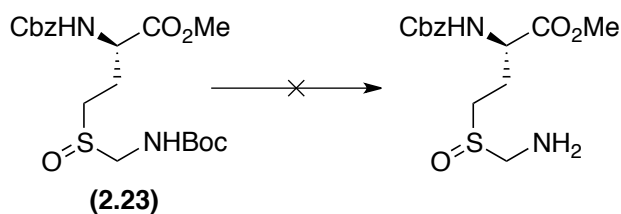
In a 100-mL round-bottomed flask, compound **2.22** (0.1 g, 0.25 mmol) was dissolved in 10 mL of CH₂Cl₂ followed by the addition of mCPBA (0.044 g, 0.25 mmol). The reaction was stirred to completion by monitoring with TLC. A saturated solution of Na₂CO₃ (30 mL) was then added to the reaction mixture and an extraction was performed using ethyl acetate (3 × 30 mL). The organic layers were combined, washed with brine, dried over magnesium sulfate and evaporated under vacuum. The product was purified using silica gel chromatography (6:4 hexanes:ethyl acetate) to yield the product as a gummy solid in 90% yield.

¹H NMR (300 MHz, CDCl₃) δ (ppm): 7.31 (br. s, 5H), 5.95 (br. s, 1H), 5.73 (br. s, 1H), 5.07 (s, 2H), 4.68 (br. s, 1H), 3.72 (s, 3H), 2.76-2.74 (m, 2H), 2.34-2.32 (m, 1H), 2.15-2.12 (m, 1H), 1.44 (s, 9H).

¹³C NMR (75 MHz, CDCl₃) δ (ppm): 172.6, 156.9, 156.2, 136.9, 129.3, 128.8, 83.2, 67.9, 67.1, 59.7, 53.3, 46.2, 37.4, 26.2.

LRMS (m/z): [M+H] Calculated: 429.17

Found: 429.1

Boc deprotection of compound 2.23

Attempt	Conditions	Result
1	TFA, CH ₂ Cl ₂	product degradation
2	H ₃ PO ₄ , THF	product degradation
3	NEt ₃ , TMSOTf, DCM	product degradation
4	TBAF, THF	product degradation

Table 2.2: Deprotection conditions for Boc removal.

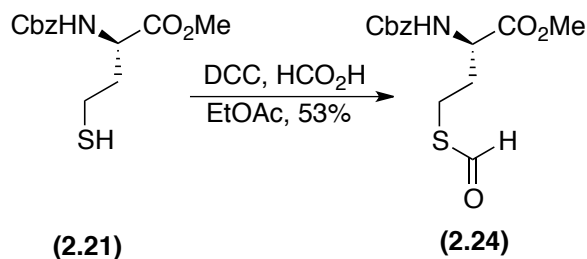
Attempt #1. Compound **2.23** (0.05 g, 0.12 mmol) was stirred in a round-bottomed flask in 10 mL of DCM while trifluoroacetic acid (0.13 mL, 1.2 mmol) was added. The reaction was stirred overnight, at which point the reaction mixture was evaporated using a mixture of cyclohexane and acetone. An NMR spectrum was taken of the crude product and it was found that the reaction conditions led to product degradation.

Attempt #2. In 10 mL of THF, was dissolved compound **2.23** (0.05 g, 0.12 mmol). Phosphoric acid (0.1 mL, 1.2 mmol) was added and the reaction stirred overnight. After the reaction mixture was evaporated under vacuum, an NMR spectrum taken of the crude product showed degradation had occurred.

Attempt #3. To a round-bottomed flask containing compound **2.23** (0.05 g, 0.12 mmol) and 10 mL of dry dichloromethane was added NEt_3 (0.02 mL, 0.12 mmol) and TMSOTf (0.16 mL, 0.12 mmol). The reaction mixture was stirred for two hours at which point the reaction mixture was evaporated under vacuum. An NMR spectrum taken of the crude product showed that it had degraded.

Attempt #4. To 10 mL of THF, compound **2.23** (0.05 g, 0.12 mmol) was added. Tetrabutylammonium fluoride (0.5 mL, 0.18 mmol) was added and the reaction mixture stirred overnight. After the reaction mixture was evaporated under vacuum, an NMR spectrum showed that product degradation had occurred.

6.2.5 Synthesis of amino methyl sulfoxide via thiol formate



An ice-cold solution of dicyclohexylcarbodiimide (DCC) (0.2 g, 0.97 mmol) in dry ethyl acetate was added with stirring to an ice-cold solution of compound **2.21** (0.2 mL, 0.97 mmol) and formic acid (0.037 mL, 0.97 mmol) dissolved in ethyl acetate. Cooling and stirring was continued for two hours and the reaction was monitored by TLC. The precipitated dicyclohexylurea was removed by filtration and the solvent evaporated. The product was then purified using flash chromatography (4 :1 hexanes:ethyl acetate) to yield a clear oil in 53% yield.

¹H NMR (300 MHz, CDCl₃) δ (ppm): 10.10 (s, 1H), 7.37 (br.s, 5H), 5.34 (d, 1H), 5.13 (s, 2H), 4.57-4.53 (m, 1H), 3.77 (s, 3H), 2.64-2.55 (m, 2H), 2.17-2.13 (m, 1H), 2.02-1.97 (m, 1H).

¹³C NMR (75 MHz, CDCl₃) δ (ppm): 187.9, 172.8, 156.7, 136.9, 129.4, 128.9, 128.5, 127.9, 68.1, 61.2, 53.9, 33.6, 23.4.

HRMS (m/z): [M+Na] Calculated: 334.0725

Found: 334.0744

6.3 Cell Permeable, Fluorescent TG2 Probe

All Fmoc-protected amino acids, resins, and coupling reagents were purchased from GL Biochem. Wang resin was purchased from NovaBiochem. All other reagents were obtained from Sigma-Aldrich. Reactions requiring anhydrous conditions were carried out under a dry nitrogen atmosphere employing conventional benchtop techniques. Common solvents were bought from EMD Chemicals whereas anhydrous solvents (THF, Et₂O, CH₂Cl₂, DMF, CH₃CN, toluene) were obtained by filtration through drying columns on a GlassContour system (Irvine, CA). All aqueous solutions were prepared using water purified with a Millipore BioCell system.

Reactor tubes for solid-phase peptide synthesis were obtained from Supelco and shaking was performed on a shaker. All resins were swelled in DMF and washing steps were performed using CH₂Cl₂ and DMF. Purification of all peptides was performed using a preparative HPLC method. Mass spectral data (MS, LCMS) were all obtained using the column, Synergi Polar-RP, 150 × 4.6 mm, 4m (Phenomenex, Torrance, CA). The crude peptide was purified using a preparative Synergi Polar-RP, 100 × 21.20 mm (Phenomenex, Torrance, CA) on a Waters Delta 600 HPLC system.

All buffers and reagents for *in vivo* cell work were purchased from Invitrogen Technologies apart from DMSO (Bioreagent) and retinoic acid (HPLC grade) that were purchased from Sigma Aldrich.

6.3.1 General procedure for synthesis of Ac-R₈K(dansyl)GGK(acryl)LPF-OH:

The peptide was synthesized using standard solid phase Fmoc chemistry. Briefly, the first Fmoc-protected amino acid (5.5 mmol) was coupled to Wang resin (1.1 mmol) using DIC (5.5 mmol) and DMAP (0.11 mmol). The level of loading of the amino acid on the resin after the first coupling step was used as the resin loading capacity for all subsequent steps, as determined by spectroscopic measurement of the UV absorbance of the piperidine dibenzofulvalene adduct. The remaining free hydroxyl functionalities were capped by treating the resin with a mixture of acetic anhydride/pyridine (2:3) and shaken for 2 h. After washing with DMF (3 times with 10 resin volumes), DCM (3 times with 10 resin volumes) and ether (3 times with 10 resin volumes), the Fmoc group was removed by incubating three times with piperidine in DMF (20% v/v; 10x resin volume) for 5 min, followed by washing with DMF (3 times with 10 resin volumes), DCM (3 times with 10 resin volumes) and ether (3 times with 10 resin volumes) in preparation for the next amide coupling. Deprotection was verified by positive Kaiser test on a sample of a few beads.² Then each Fmoc-protected amino acid (1.7 mmol) was coupled to Wang resin preloaded with the necessary carboxyl-terminal amino acid (0.68 mmol) in DMF (5 resin volumes) using HOBt (1.7 mmol) and DIC (1.7 mmol). This operation was realized twice for 30 min. Coupling was verified by negative Kaiser test on a sample of a few beads. In the case of acryloyl acid (1.7 mmol), the coupling reagent was EEDQ (1.7 mmol) and the shaking time was 1 h. The peptide was cleaved from the resin (1 g) by incubating with TFA:DCM (1:1) for 2 h. The peptide was precipitated from the cleavage solution using diethyl ether and

hexane. Peptide sequence purity was performed on an aliquot of resin using the column, Synergi Polar-RP, 150 × 4.6 mm, 4m (Phenomenex, Torrance, CA). The crude peptide was purified using a preparative Synergi Polar-RP, 100 × 21.20 mm (Phenomenex, Torrance, CA) on a Waters Delta 600 Varian (Prep Star) HPLC system.

Boc deprotection protocol

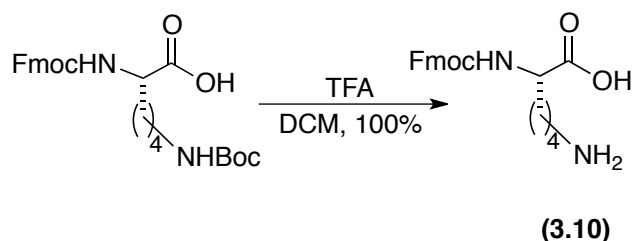
To a reactor containing 1 g of the Wang resin supported Fmoc-peptide (0.68 mmol, according to the loading), were added 30 mL of deprotection mixture, freshly prepared from 470 μ L TEA (2 eq.), 1.09 mL of TMSOTf (0.2 M) and 28.44 mL of anhydrous DCE. The resin was shaken for 10 min then filtered and washed with 5 × 5 mL of DCM, 2 × 5 mL of DIEA 10% in DCM, 3 × 5 mL of DCE. Deprotection was carried out for another 10 min with a fresh deprotection mixture. The resin was filtered then washed with 5 × 5 mL of DCM, 2 × 5 mL of DIEA, 10% in DCM, 2 × 5 mL of DCM, 2 × 5 mL of DMF and 2 × 5 mL of Et₂O. Deprotection was verified by positive Kaiser test on a sample of a few beads.

Acrylation protocol

To the Wang resin supported Fmoc-peptide (1 g, 0.68 mmol) swollen in anhydrous DCM (5 resin volumes) was added acryloyl acid (1.7 mmol) and EEDQ (1.7 mmol), the reaction was shaken 1 h, followed by washing with DMF (3 times with 10 resin volumes), DCM (3 times with 10 resin volumes) and ether (3 times with 10 resin volumes). This operation was repeated twice.

6.3.2 Synthesis of Fmoc-K(Dansyl-OH)

2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-6-aminohexanoic acid (3.10)



To a 50-mL round-bottomed flask containing Fmoc-K(N-Boc)-OH (1 g, 2.13 mmol) was added TFA (7 mL, 85.5 mmol) along with 7 mL of dichloromethane. The reaction was stirred for 2 h at room temperature. The reaction mixture was then evaporated under vacuum and dissolved in a minimal amount of acetone and cyclohexane, which were then evaporated under vacuum. The addition of a minimal amount of acetone and cyclohexane was repeated twice to obtain a pale brown oil in quantitative yield, which was used without further purification.

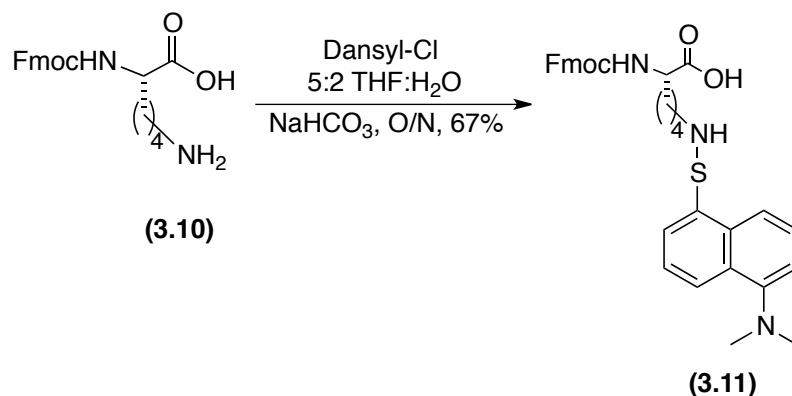
¹H NMR (300 MHz, CDCl₃) δ (ppm): 7.88-7.84 (m, 2H), 7.61-7.59 (m, 2H), 7.42-7.38 (m, 2H), 7.33-7.30 (m, 2H), 4.43-4.40 (m, 2H), 4.31-4.28 (m, 2H), 4.2-4.0 (m, 1H), 2.84-2.63 (m, 2H), 1.45-1.35 (m, 6H).

¹³C NMR (75 MHz, CDCl₃) δ (ppm): 173.4, 160.5, 141.3, 143.5, 127.9, 127.1, 120.0, 67.2, 54.3, 46.9, 31.0, 28.5.

HRMS (m/z): [M+Na] Calculated: 391.1634

Found: 391.1623

(S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-6-(5-(dimethylamino)naphthalene-1-sulfonamido)hexanoic acid (3.11).



To a round-bottomed flask containing compound **3.10** (2.2 g, 6.1 mmol) was added 189 mL of MeOH:H₂O in a 5:2 ratio, NaHCO₃ (1.54 g, 18.3 mmol) followed by the addition of dansyl chloride (1.48 g, 5.5 mmol). The reaction was stirred overnight at which point 1 M HCl (100 mL) was added. An extraction with DCM (4 × 75 mL) was performed. The organic layers were collected and washed with brine, dried over magnesium sulfate, filtered and the solvent evaporated under vacuum to yield a pale yellow solid in 54% yield.

M.P.: 79-82 °C

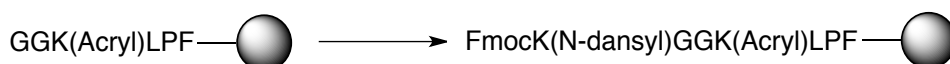
¹H NMR (400 MHz, CDCl₃) δ (ppm): 11.79 (br. s, 1H), 9.40 (d, 1H, *J* = 8 Hz), 8.87 (d, 1H), 8.50 (d, 1H), 7.88-7.84 (m, 2H), 7.77-7.73 (m, 3H), 7.61-7.59 (m, 2H), 7.42-7.38 (m, 2H), 7.33-7.30 (m, 2H), 4.45-4.38 (m, 2H), 4.2-4.0 (m, 1H), 3.32 (s, 6H), 3.13-3.10 (m, 2H), 1.27-1.22 (m, 6H).

^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 174.78, 156.8, 152.67, 144.54, 142.12, 140.58, 134.10, 130.88, 130.29, 129.78, 129.16, 129.12, 128.75, 128.61, 128.23, 127.91, 125.97, 125.84, 123.76, 121.11, 120.82, 119.48, 117.7, 67.36, 54.08, 48.0, 46.80, 43.50, 32.10, 31.76, 22.23.

HRMS (m/z): $[\text{M}+\text{Na}]$ Calculated: 624.2144

Found: 624.2139

Coupling of Fmoc-K(dansyl)-OH to peptide



Coupling Reagent	Fmoc-K(N-dansyl)OH (3.11)	Activator	Result
3 eq. DIC	3 eq.	3 eq. HOBt	No reaction
3 eq. EEDQ	3 eq.	N/A	No reaction
8 eq. TBTU	4 eq.	4 eq. HOBt + 8 eq. DIEA	Coupling successful

Table 3.1: Coupling conditions of Fmoc-K(dansyl)-OH

The coupling of compound **3.11** to the rest of the peptide was investigated using the conditions in Table 3.1. The procedure with DIC and EEDQ was described previously, under the general protocols of solid phase synthesis. In the case of TBTU, all the reagents (4 eq TBTU, 4 eq. HOBt and 8 eq. DIEA) were added to the reactor in DMF (5 resin volumes) with 1 g of resin. The reaction was shaken for 1 hour and this

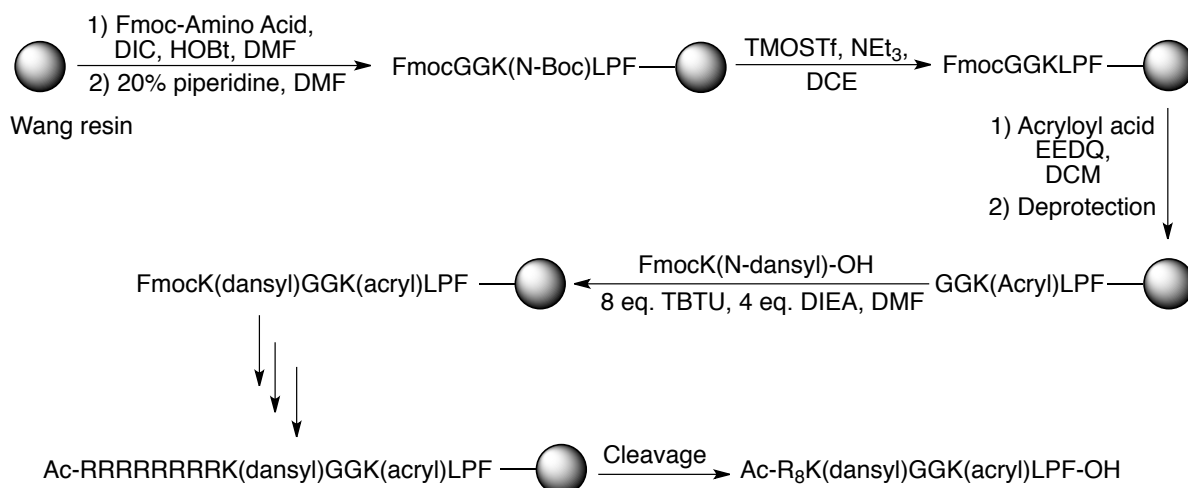
operation was realized twice. After washing with DMF (3 times with 10 resin volumes), DCM (3 times with 10 resin volumes) and ether (3 times with 10 resin volumes), the coupling results were verified by the Kaiser test and through verifying the fluorescence (or lack thereof) of a small aliquot of resin beads. Once fluorescence was observed, the coupling was verified through low resolution MS-ESI⁺, after cleaving 100 mg of peptide from the resin by incubating it for two hours with a mixture of 1:1 TFA:DCM.

FmocK(N-dansyl)GGK(acryl)LPF-OH:

LRMS: [M+H] Calculated: 1201.5

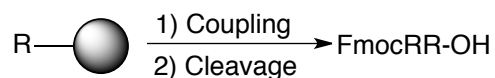
Found: 1201.5

RRRRRRRRK(dansyl)GGK(acryl)LPF-OH



The peptide was synthesized using the general procedures mentioned above, however identification of the final product proved elusive as both LC-MS ESI and MALDI proved unsuccessful at detecting the peptide. To that end, various coupling

conditions were investigated for their arginine coupling efficiency, as shown in Table 6.2.



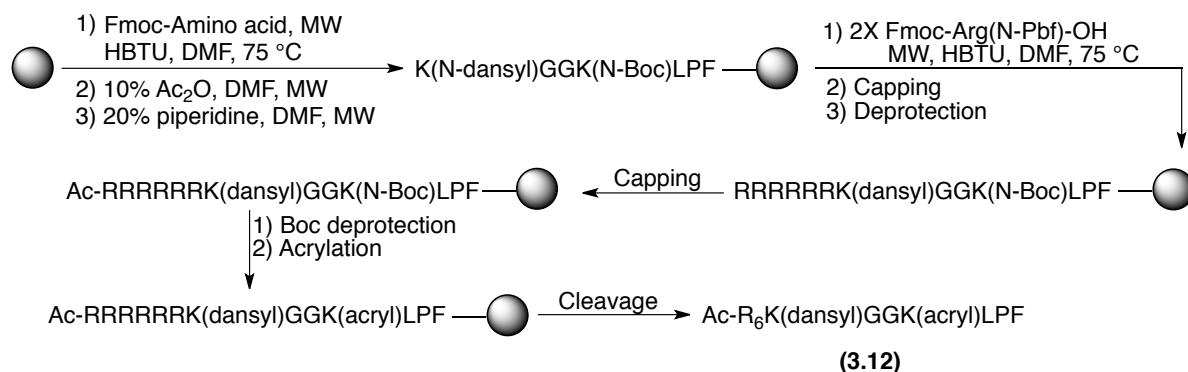
Coupling Reagent	Fmoc-Arg(Pbf)-OH	HOBt	Yield (%)
4 eq. TBTU	4 eq.	4 eq. + 8 eq. DIEA	19.9
3 eq. HATU	3 eq.	3 eq.	12.7
3 eq. PyBOP	3 eq.	3 eq.	18.3
3 eq. EEDQ	3 eq.	N/A	27.1

Table 3.2: Evaluation of arginine coupling conditions.

The general protocol followed for each reaction involved adding the reagents to 100 mg of Wang resin already loaded with arginine as detailed in the general procedure above for the coupling of the first amino acid. The reagents were shaken in 1 mL of DMF for an hour. After washing with DMF (3 times with 10 resin volumes), DCM (3 times with 10 resin volumes) and ether (3 times with 10 resin volumes), the operation was repeated. After the second shaking, the resin was washed with DMF (3 times with 10 resin volumes), DCM (3 times with 10 resin volumes) and ether (3 times with 10 resin volumes) before cleavage from the resin by shaking for two hours in a mixture of 1:1 TFA:DCM. The resin was then washed with DCM (3 times with 10 resin volumes). The peptide was precipitated from the cleavage solution using diethyl ether and

hexane. The precipitated peptide was then recovered through centrifugation at 1100 g for 30 min. As EEDQ was shown to have the highest yield, the 8 arginines were coupled using this procedure however; the product was still not identified according to accepted techniques.

6.3.3 Synthesis of Ac-R₆K(dansyl)GGK(acryl)LPF-OH.



Compound **3.12** was synthesized using an automated Liberty CEM microwave peptide synthesizer (Matthews, NC) by conventional Fmoc-based solid-phase peptide synthesis. Synthesis was carried out on a scale of 1 g of Wang resin with a loading capacity of 0.93 mmol/g (0.25 mmol scale). Each round of peptide elongation consists of (i) Fmoc deprotection with 20% piperidine in DMF for 5 min three times; (ii) washing with DMF; (iii) coupling with *N,N,N',N'*-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU), DIEA, *N*-methylpyrrolidinone (NMP), and the appropriate amino acid in DMF for 11 min; (iv) washing with DMF; (v) acetylation with 10% acetic anhydride in DMF and (vi) washing with DMF. The molar ratio of the amino groups on the resin/ building block/HBTU was 1:4:3.6. After synthesis of the entire peptide, the Boc deprotection and acrylation protocol was carried out as described

previously. The peptide was then cleaved from the resin using 90:5:5 trifluoroacetic acid (TFA)/ dichloromethane (DCM)/thianisole for 120 min. After cleavage, the resin was filtered and washed three times under vacuum using 1:1 cyclohexane/acetone. The peptide was dissolved in a minimal amount of acetone and precipitated from the solution using cold Et₂O with sonication. The precipitated peptide was then recovered through centrifugation at 1100 × g for 30 min.

HPLC analysis and purification.

The crude material was purified by preparative Synergi Polar-RP (100 × 21.20 mm) column, on a Waters Delta 600 HPLC using a percentage of ACN in water as eluant, a flow rate of 6 mL/min and the detector set at 254 nm or 214 nm. The collected fractions were freeze-dried to give the compound. The areas under the peaks were determined using Empower Pro (Waters). The peptide Ac-R₆K(dansyl)GGK(acryl)LPF-OH (3.12) was obtained in an overall 5% yield as a pale yellow powder with 90% purity.

LRMS: [M+H] Calculated: 2013.4

Found: 2013.4

MALDI: [M+H] Calculated: 2013.4

Found: 2014.8

6.3.4 Kinetic analysis of CPP

6.3.4.1 GDH assay

The GDH assay was performed in triplicate using 200 mM MOPS buffer (pH 7.2), 1 mM EDTA, 500 μ M reduced form of nicotinamide adenine dinucleotide (NADH), 10 mM α -ketoglutarate, 0.96 U of glutamate dehydrogenase (GDH), varying concentrations of the cell permeable peptide (3.12) (0-80 μ M), and 10 mM *N*-acetyl lysine methyl ester as an acyl-acceptor substrate in a final volume of 180 μ L. The reaction solution was preincubated for 5 min at 37 °C prior to initiation of the reaction by the addition of TG2 (0.007 U in a volume of 20 μ L) to give a final volume of 200 μ L. For the corresponding blank solution, the TG2 solution was replaced with water. The assays were performed in a 96-well plate, and the decrease in absorbance due to the oxidation of NADH was followed against a blank at 340 nm in a Synergy HT Multi-Mode Microplate Reader (Winooski, VT) thermostated at 37 °C. After a brief lag during which sufficient ammonia is produced to saturate GDH, linear slopes of absorbance versus time were measured. This lag phase is intrinsic to all continuous enzyme-coupled assays, for which it is difficult to measure true initial rates. However, the brevity of this lag in this assay allows for the accurate measurement of approximate initial rates, which are still a valid indicator of substrate turnover. These approximate initial rates were corrected for path length using the instrument software. The data obtained (blank-corrected) was plotted as the reciprocal of the observed rate versus inhibitor concentration (Figure 3.8) and a linear trend line was calculated.

The trend line can be inputted into the Equation 3.1 where the x-intercept of the line is the IC₅₀ value of the cell permeable peptide. This IC₅₀ value was then substituted into Equation 3.2 to obtain a K_i for the inhibitor, which was calculated to be 9 μM.

$$\text{Equation 3.1: } \frac{1}{\text{rate}} = m[I] + b \qquad \text{Equation 3.2: } K_i = \frac{IC_{50}}{(1 + \frac{[S]}{K_m})}$$

6.3.4.2 AL5 assay

This activity assay was performed using 200 mM MOPS (pH 7), 100 mM CaCl₂, and 20 mM EDTA with varying concentrations of the cell permeable peptide (3.12) (0 μM to 535 μM) in a total volume of 260 μL. The reaction solution was preincubated for 5 min at 37 °C prior to initiation of the reaction by the addition of TG2 (0.015U in 40 μL) to give a final volume of 300 μL. For the corresponding blank solution, the TG2 solution was replaced with water. The reaction solutions were then incubated for 30 min at 37 °C. The solutions were then added to an Amicon centrifugal filter with a 30-kDa cutoff and centrifuged three times at 955 × g for 4 min, successively washing the filter each time. After the third wash, the sample left in the filter was added to a 96-well plate and the substrate *N*-Cbz-L-Glu(γ-*p*-nitrophenyl ester)Gly (20 μL of a 1.4 mM solution) was added at a concentration of 150 μM in the well. The increase in absorbance at 415 nm was followed against a blank in a Synergy HT Multi-Mode Microplate Reader thermostated at 24 °C. Linear slopes of absorbance versus time were obtained. However, analysis of the kinetic data (not shown) indicated we were not able to capture the initial rate of first order decay, but instead observed the rate plateau after enzyme inactivation.

6.3.5 Preliminary *in vivo* results with Ac-R₆K(dansyl)GGK(acryl)LPF-OH

SH-SY5Y neuroblastoma cells were obtained from the lab of Professor Gail Johnson. Approximately 7×10^5 cells/well were transferred to a 6-well culture plate (Corning), containing 2 mL of DMEM (Dulbecco's Modified Eagle Medium)/F12 medium supplemented with 10% FBS (fetal bovine serum), 1% antibiotics (penicillin and streptomycin), 1% L-glutamine and 1% of a non-essential amino acids solution in 6 well plates and grown at 37 °C under 5% CO₂. The cells were divided into two groups, so that three wells underwent additional retinoic acid (RA) treatment. The treatment consisted of a solution of retinoic acid dissolved in DMSO being added to each well to a final concentration of 20 µM. The remaining three wells underwent no treatment as a control. The cells were grown to 70-80% confluence before treatment with compound **3.12**.

Once ready, the SH-SY5Y cells were incubated with 10 µM of compound **3.12** at 37 °C under 5% CO₂ for 30 min. The medium containing compound **3.12** was then removed from the cells and the cells were washed with a 10% PBS solution. Supplemented DMEM/F12 medium (2 mL) was added to the cells and they were incubated for 20 min at 37 °C under 5% CO₂. The medium was then removed and the cells washed with 10% PBS (phosphate buffered saline) solution. This procedure was repeated twice, whereby the supplemented DMEM/F12 medium was added to the cells, incubation was carried out for 20 minutes at 37 °C under 5% CO₂ and then washed with 10% PBS solution. A solution of 4% paraformaldehyde was then added to the cells and they were incubated at room temperature for 20 min in the dark. The cells were then washed twice with a 10% PBS solution and were mounted on glass slides using 25 µL of

mounting media. They were dried at room temperature for 2 hours and then stored at 4 °C overnight. The confocal microscopic images were taken with an Olympus BX-60 confocal microscope using a DAPI filter, magnified at . The DAPI filter comprises an excitation filter that allows light between 360-370 nm, and a long pass 420 nm emission filter.

6.3.6 Computer modelling of Ac-R₆K(dansyl)GGK(acryl)LPF-OH with TG2

The crystal structure of TG2 in the open form (PDB: 1KV3) was refined and hydrogen atoms were added at the normal ionization state of amino acids at pH 7.0 while partial charges were fixed to the AMBER99 atom types. The final structure was minimized using conjugate gradient minimization until a convergence of 0.001 kcal•mol⁻¹•Å⁻¹ was reached. To prevent the introduction of steric clashes between residue side-chains, model backbones were tethered by limiting the deviation of the backbone to no more than 0.5 Å from the initial position. Compound **3.12** was built using the Builder module of MOE, protonated and atomic potentials (or partial charges) were fixed to the AMBER99 force field as assigned in the MOE software. The catalytic domain was manually selected and contained the residues important for TG2 catalytic function. The peptide was oriented and placed within the selected TG2 docking site with the Triangle matcher method using the refined open form of TG2. This method uses pre-established docking protocols developed in the Keillor lab. Docked poses were then scored using the London ΔG scoring algorithm to rank the docking poses. Protein solvation was accounted for via implicit solvation. The top ranked structures obtained after scoring were retained and evaluated.

6.4 FXIIIa

All Fmoc-protected amino acids, resins, and coupling reagents were purchased from GL Biochem, Wang resin was purchased from NovaBiochem. All other reagents were obtained from Sigma-Aldrich or BioShop. Reactions requiring anhydrous conditions were carried out under a dry nitrogen atmosphere employing conventional benchtop techniques. Common solvents were bought from EMD Chemicals whereas anhydrous solvents (THF, Et₂O, CH₂Cl₂, DMF, CH₃CN, toluene) were obtained by filtration through drying columns on a GlassContour system (Irvine, CA). All aqueous solutions were prepared using water purified with a Millipore BioCell system.

Reactor tubes for solid-phase peptide synthesis were obtained from Supelco and shaking was performed on a shaker. All resins were swelled in DMF and washing steps were performed using CH₂Cl₂ and DMF (EMD Chemicals). Purification of all peptides was performed using a preparative HPLC method. Mass spectral data (MS, LCMS) were all obtained using the column, Synergi Polar-RP, 150 × 4.6 mm, 4m (Phenomenex, Torrance, CA). The crude peptide was purified using a preparative Synergi Polar-RP, 100 × 21.20 mm (Phenomenex, Torrance, CA) on a Waters Delta 600 HPLC system.

6.4.1 Optimization of expression of GST-FXIIIA

6.4.1.1 Expression of GST-FXIIIA

BL21(DE3) cells were transformed with the expression plasmid pGF13A2, obtained from Lai and coworkers.³ Cells were taken from a frozen stock and grown overnight in 10 mL of Terrific Broth (TB)/ Luria-Bertrani(LB) or Studier's NZ-amine, yeast, phosphate (ZYP) auto-induction medium containing 100 µg/mL ampicillin. This bacterial suspension was then used to inoculate 1 L of the appropriate fresh medium and the culture was incubated at 37 °C with shaking (240 rpm). When the optical density at 600 nm reached approximately 0.6, IPTG was added (for TB and LB media) and the culture was incubated for an additional 20 h at varying temperatures (17-32 °C) with shaking at 240 rpm. Cells were then harvested by centrifugation (30 min, 2000 × g, 4 °C) and the resulting pellet was resuspended in 30 mL of lysis buffer (20 mM Tris, 150 mM NaCl, 1 mM EDTA, 1 mM TCEP and 15% glycerol at pH 8.0). The cells were lysed with a cell disrupter (Constant Systems) and the crude lysate was clarified by centrifugation (60 min, 44000g, 4 °C) and filtration (0.22-µm pore size).

6.4.1.2 Purification of GST-FXIIIA

All purification steps were carried out at 0-4 °C. To the lysate was added 1 mL of Glutathione Sepharose 4B resin that had been previously equilibrated with lysis buffer. The slurry was mixed gently (100 rpm on a rotary shaker) for 2 h and then loaded onto a gravity filtration column (Bio-Rad Econo-Column chromatography column 1 × 10 cm). This column was then washed with 10 mL of lysis buffer containing 0.5 % Triton X-100, 10 mL of lysis buffer and 10 mL of elution buffer (50 mM Tris and 15% glycerol at pH

7.5). These combined fractions were retained as flow-through, while the FXIIIA remained bound on the column. To the column was added 5 mL of GST elution buffer (50 mM Tris and 10 mM glutathione at pH 7.5) that was allowed to stand for 25 minutes. FXIIIA was eluted from the column, and another 3 mL of GST elution buffer was added to the column and allowed to stand for 25 minutes. At this point, the buffer was eluted from the column and the combined fractions with FXIIIA were concentrated using a 50-kDa Amicon filter with the lysis buffer. The concentrated FXIII filtrate was aliquoted into micro-centrifuge tubes (250 μ L/tube), flash-frozen on dry ice and stored at -80 °C.

6.4.1.3 Protein concentration

Total protein concentration was determined using the Bio-Rad protein assay, a method based on the Bradford assay, using BSA (bovine serum albumin) as a standard. Colour development took place over 20 min and absorbance values were measured using a Varian Cary 100 Bio and compared to standard curves performed daily.

6.4.1.4 Electrophoresis

Separation of proteins by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was conducted using the Bio-Rad Mini-PROTEAN® Tetra Cell electrophoresis system according to procedures of Laemmli⁴ using minigels (86 × 68 mm) prepared with 10% acrylamide resolving gel at pH 8.8, 4% acrylamide stacking gel at pH 6.8 and 2.7% crosslinker concentrations. The protein sample (15 μ L) was mixed with 5 μ L of loading buffer (50 mM Tris-HCl (pH 6.8), 100 mM DTT, 2% w/v SDS,

0.1 % w/v bromophenol blue, and 10% v/v glycerol) and boiled for 5 min prior to loading. Protein bands were revealed using 250R Coomassie blue staining followed by drying. The SDS-PAGE broad-range molecular weight markers (Bio-Rad) were used as standards.

6.4.2 Synthesis of high affinity peptide sequences for FXIIIa

Peptides were synthesized using an automated peptide synthesizer (Liberty CEM Microwave Peptide Synthesizer) by conventional Fmoc-based solid-phase peptide synthesis (SPPS). Synthesis was carried out on a scale of 1 g of Wang resin with a loading capacity of 0.93 mmol/g (0.25 mmol scale). Each round of peptide elongation consists of (i) Fmoc deprotection with 20% piperidine in DMF for 5min three times; (ii) washing with DMF; (iii) coupling with *N,N,N',N'*-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU), *N,N*-diisopropylethylamine (DIEA), *N*-methylpyrrolidinone (NMP), and the appropriate amino acid in DMF for 11 min; (iv) washing with DMF; (v) acetylation with 10% acetic anhydride in DMF and (vi) washing with DMF. The molar ratio of the amino groups on the resin/ building block/HBTU was 1:4:3.6. The peptide was then cleaved from the resin using 90:5:5 trifluoroacetic acid (TFA)/ dichloromethane (DCM)/thioanisole for 120 min. After cleavage, the resin was filtered and washed three times under vacuum using 1:1 cyclohexane/acetone. The peptide was dissolved in a minimal amount of acetone and precipitated from the solution using cold Et₂O with sonication. The precipitated peptide was then recovered through centrifugation at 1100 × g for 30 min. The crude material was purified by preparative Synergi Polar-RP (100 × 21.20 mm) column, on a Waters Delta 600 HPLC

using a percentage of ACN in water as eluant, a flow rate of 10 mL/min and the detector set at 254 nm or 214 nm. The collected fractions were freeze-dried to give the desired peptide whereby small aliquots tested demonstrated purities of at least 90% for each peptide as determined by calculating the area under the peak using Empower Pro.

Ac-SQMM-OH (4.6)

Obtained as a white powder in an overall 8% yield. Purification conditions were 5 to 80% ACN in 25 min. Retention time: 3.09 min.

HRMS (m/z): [M+Na] Calculated: 560.1825 Found: 560.1843

Ac-DQMML-OH (4.7)

Obtained as a white powder in an overall 13% yield. Purification conditions were 35 to 90% ACN in 25 min. Retention time: 5.17 min.

HRMS (m/z): [M+Na] Calculated: 677.2639 Found: 677.2565

Ac-DQMHV-OH (4.8)

Obtained as a white powder in an overall 9% yield. Purification conditions were 5 to 80% ACN in 25 min. Retention time: 2.61 min.

HRMS (m/z): [M+Na] Calculated: 693.2642

Found: 693.2653

Ac-DQMMMAF-OH (4.9)

Obtained as a white powder in an overall 10% yield. Purification conditions were 40 to 95% ACN in 25 min. Retention time: 3.46 min.

HRMS (m/z): [M+Na] Calculated: 936.3156

Found: 936.3088

Ac-SQSRV-OH (4.10)

Obtained as a white powder in an overall 11% yield. Purification conditions were 15 to 80% ACN in 25 min. Retention time: 4.62 min.

HRMS (m/z): [M+Na] Calculated: 618.3211

Found: 618.3187

Ac-GQSMV-OH (4.11)

Obtained as a white powder in an overall 5% yield. Purification conditions were 5 to 80% ACN in 25 min. Retention time: 4.51 min.

HRMS (m/z): [M+Na] Calculated: 561.2343

Found: 561.2383

6.4.3 Optimization of thrombin activation

The optimization of thrombin activation of FXIIIa was carried out by incubating a solution of varying units of thrombin (5-30 units) in 200 mM MOPS, 10 mM CaCl_2 , 100 mM α -ketoglutarate, 100 mM glycine methyl ester, 500 μM NADH, 4.8 units of GDH and 142 nM FXIIIa in a total volume of 600 μL . Each sample was incubated for 1 hour at 37 $^\circ\text{C}$ at which point the reactions were quenched by placing the samples on ice. An SDS-PAGE gel was then undertaken with the samples and the results compared.

To further and quantitatively refine the amount of thrombin necessary for FXIII activation, an activity test using the GDH assay was performed with various amounts of thrombin. Varying units of thrombin (0.5-5 units) in 200 mM MOPS (pH 7.0), 10 mM CaCl_2 , 100 mM α -ketoglutarate, 100 mM glycine methyl ester, 500 μM NADH, 4.8 units of GDH and 142 nM FXIIIa in a total volume of 600 μL were incubated for 1 hour at 37 $^\circ\text{C}$. The specific activity of FXIIIa was measured by monitoring the decrease in absorbance of NADH at 340 nm after the addition of *N,N*-dimethylcasein to a final concentration of 700 μM . One unit of enzyme activity (U) is defined as the amount of FXIIIa that catalyzes the formation of 1 μmol of NAD^+ from NADH per minute and is reported as units per milligram of protein.

6.4.4 Kinetic analysis of FXIIIa substrate peptides

The peptides were analyzed kinetically through the use of the GDH assay, where the release of ammonia catalyzed by FXIIIa leads to the formation of NAD^+ from NADH and a concurrent decrease in absorbance at 340 nm. The peptides served as the

glutamine donor while the buffer contained glycine ethyl ester as the amine acceptor substrate. The standard assay was performed in 200 mM MOPS (pH 7.0), 10 mM CaCl_2 , 100 mM α -ketoglutarate, and 500 μM NADH, 0.96 units GDH, varying concentrations of peptide and fixed concentrations of an acyl acceptor substrate (100 mM glycine ethyl ester) in a final volume of 180 μL were preincubated for 3 min at 37 °C. The enzyme reactions were initiated by adding 16.6 μg of activated FXIIIa (in a volume of 20 μL) to give a final volume of 200 μL . For the corresponding blank solution, the TGase solution was replaced by water. The increase in absorbance at 415 nm was followed against a blank in a Synergy HT Multi-Mode Microplate Reader thermostated at 37 °C. Linear slopes of absorbance versus time were obtained. All experiments were performed in triplicate whereby all peptides except Ac-DQMMAF were tested using FXIIIa obtained from the same purification. The activity of FXIIIa used to test Ac-DQMMAF was normalized by testing it with Ac-DQMML-OH and comparing the rates.

The linear slopes (blank-corrected) were transformed into initial reaction rates by dividing by the extinction coefficient of NADH ($6220 \text{ M}^{-1}\text{cm}^{-1}$). The rate of the reaction was then plotted against peptide concentration and a straight line was obtained for each peptide.

6.4.5 Modelling of peptides with FXIIIa and TG2

From the crystal structure of the open form of TG2 (PDB: 1KV3), a homology model was created for FXIIIa using MOE (Molecular Operating Environment) program version 2011.10 from the Chemical Computing Group. The sequence of FXIIIa was

directly aligned with the known TG2 open form template, with an approximate 37% sequence identity between FXIIIa and TG2 using the BLOSUM Alignment matrix. The sequence of FXIIIa was then threaded onto TG2 and known structural information from the TG2 template was transferred into the rough FXIIIa structure, which would become the primary template. Twenty-five intermediate models for the open form FXIIIa structure were generated for which small loop-based or side chain-based modifications were carried out to the primary template and each modification was coarsely energy-minimized for steric interaction. The final model generated, based on the cartesian average of the twenty-five intermediate models was validated relative to steric and atomic clashes as well as rotamer outliers to confirm model quality. The model was refined and hydrogen atoms were added at the normal ionization state of amino acids at pH 7.0 while partial charges were fixed to the AMBER99 atom types. The final structure was minimized using conjugate gradient minimization until a convergence of $0.001 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-1}$ was reached. To prevent the introduction of steric clashes between residue side-chains, model backbones were tethered by limiting the deviation of the backbone to no more than $0.5 \text{ \AA}$ from the initial position. The FXIIIa peptides were built using the Builder module of MOE, protonated and atomic potentials (or partial charges) were fixed to the AMBER99 force field as assigned in the MOE software. The structure was then minimized using conjugate gradient minimization until a convergence of $0.001 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-1}$ was reached. The catalytic domain was manually selected and contained the residues important for FXIIIa catalytic function. The peptide was oriented and placed within the selected FXIIIa docking site with the Triangle matcher method using the refined, open model of FXIIIa. Docked poses were

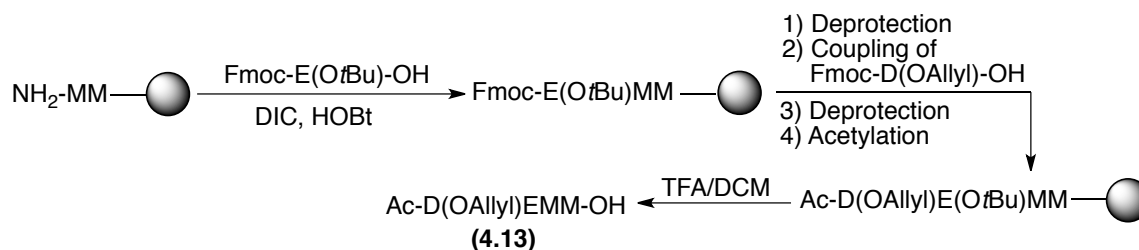
then scored using the London ΔG scoring algorithm to rank the docking poses. Protein solvation was accounted for via implicit solvation. The top ranked structures obtained after scoring were retained and evaluated.

For TG2, the crystal structure (PDB: 1KV3) was refined and hydrogen atoms were added at the normal ionization state of amino acids at pH 7.0 while partial charges were fixed to the AMBER99 atom types. The final structure was minimized using conjugate gradient minimization until a convergence of $0.001 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-1}$ was reached. To prevent the introduction of steric clashes between residue side-chains, model backbones were tethered by limiting the deviation of the backbone to no more than $0.5 \text{ \AA}$ from the initial position. The catalytic domain was manually selected and contained the residues important for TG2 catalytic function. The peptide was oriented and placed within the selected TG2 docking site with the Triangle matcher method using the refined, open form of TG2. Protein solvation was accounted for via implicit solvation. Docked poses were then scored using the London ΔG scoring algorithm to rank the docking poses. The top ranked structures obtained after scoring were retained and evaluated.

6.4.6 Attempted Synthesis of Ac-DE(O-coumarin)MM-OH

6.4.6.1 First attempt towards activity assay using Fmoc-Asp(OAllyl)-OH

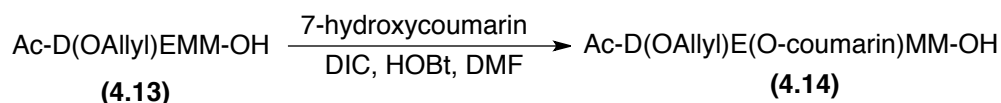
Ac-D(OAllyl)E(OtBu)MM-OH (4.13)



The synthesis of this peptide was carried out as per the general Fmoc-based solid phase coupling techniques outlined for the synthesis of Ac-R₈K(dansyl)GGK(acryl)LPF-OH, using Wang resin. The peptide was cleaved from the resin (1 g) by incubating with TFA:DCM (1:1) for 2 h. The peptide was precipitated from the cleavage solution using diethyl ether and hexane.

LRMS (m/z): [M+H]: Calculated: 551.1

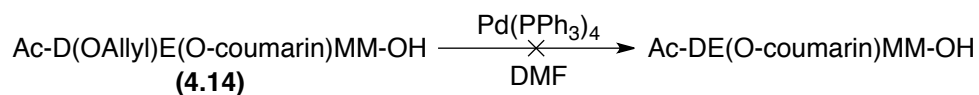
Found: 551.2

Ac-D(OAllyl)E(O-coumarin)MM-OH (4.14)

To a 100-mL round-bottomed flask containing Ac-D(OAllyl)E(OtBu)MM-OH crude peptide (0.245 g, 0.4 mmol) was added DIC (0.186 mL, 1.2 mmol), HOBT (0.165 g, 1.2 mmol) and 7-hydroxycoumarin (0.329 g, 2.0 mmol). The reagents were solubilized in 10 mL of DMF and the reaction stirred overnight at 45 °C. The following morning, both water and DCM were added to the reaction mixture and an extraction performed four times with DCM. The organic layer was washed three times with 0.1 M HCl, 4 times with sat. NaHCO₃:sat Na₂CO₃ in a ratio of 50:50, once with water and lastly, once with brine. The organic phase was dried over magnesium sulfate, filtered and then removed under rotary evaporation. The residue was then dissolved in a minimum of acetone and precipitated with ether.

LRMS (m/z): [M+H] Calculated: 694.2

Found: 694.3

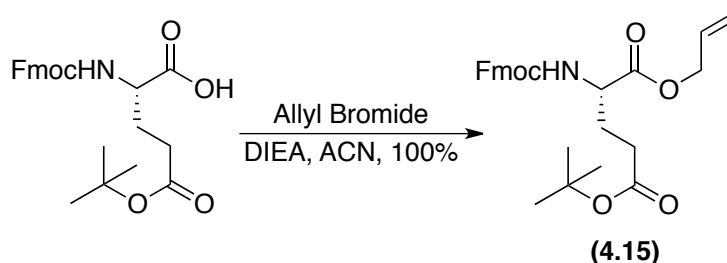
Ac-DE(O-coumarin)MM-OH

The Pd(PPh₃)₄ catalyst was created *in situ* using Pd(dba)₂ and PPh₃ in dry dichloromethane, under an inert atmosphere. After formation of the catalyst, the crude

Ac-DE(O-coumarin)MM-OH peptide was added to the solution and the reaction mixture was stirred for two hours, with monitoring by TLC. The starting material was never deprotected and it was concluded that the methionine sulfurs were poisoning the catalyst. A second attempt involved using freshly prepared $\text{Pd}(\text{PPh}_3)_4$ from the method of Malatesia *et al.*,⁵ but the same result was obtained.

6.4.6.2 Second attempt using Fmoc-Glu(O-coumarin)-OH

1-allyl 5-*tert*-butyl 2-(((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)pentanedioate (4.15)



To a 250-mL round-bottomed flask containing 100 mL of acetonitrile, Fmoc-Glu(OtBu)-OH (2 g, 4.7 mmol) was added. Allyl bromide (11.2 mL, 129.4 mmol) was then added, followed by DIEA (1.75 mL, 9.21 mmol). The reaction mixture was stirred overnight at which point 100 mL of dichloromethane was added to the flask. The organic phase was washed with 3 × 30 mL of 0.1 M HCl, 3 × 30 mL of saturated NaHCO_3 solution, and then 3 × 30 mL of brine. The organic phase was then dried over magnesium sulfate and evaporated to yield a viscous, slightly yellow oil in quantitative yield that was used without further purification.

HRMS (m/z): [M+Na] Calculated: 488.2049 Found: 488.2046

(4.15) $\xrightarrow[\text{85\%}]{\text{TFA/DCM}}$ (4.16)

To a 250-mL round-bottomed flask containing 15 mL of dichloromethane, compound **4.15** (2.29 g, 4.9 mmol) was added. Trifluoroacetic acid (15 mL, 19 mmol) was added and stirring at room temperature was continued until complete disappearance of the starting material was observed by TLC. The reaction mixture was then evaporated and washed three times with a 1:1 mixture of cyclohexane and acetone. After the last wash, the residue was dissolved in acetone and hexane was

added while sonicating the flask. The desired product precipitated as a white powder in 94% yield.

M.P.: 117-122 °C

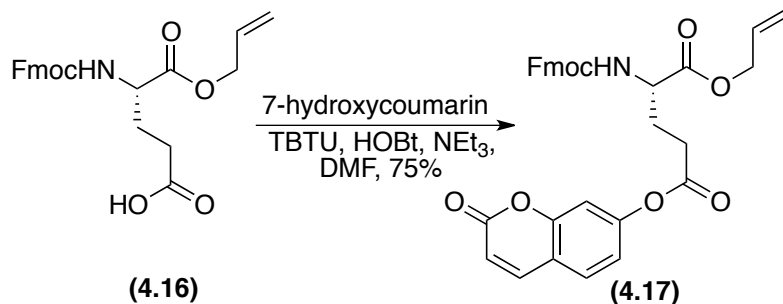
¹H NMR (400 MHz, MeOD) δ (ppm): δ 7.78 (d, 2H), 7.57 (dd, 2H), 7.38 (t, 2H), 7.30 (t, 2H), 5.90-5.86 (m, 1H), 5.47 (d, 1H), 5.19 (d, 1H), 4.60 (d, 2H), 4.40-4.35 (m, 2H) 4.30-4.26 (m, 1H), 4.22-4.19 (m, 1H), 2.34-2.29 (m, 2H), 2.13-2.09 (m, 1H), 1.97-1.94 (m, 1H)

¹³C NMR (100 MHz, CDCl₃) δ (ppm): 178.1, 177.7, 171.7, 156.5, 143.9, 143.7, 141.33, 131.1, 127.8, 127.1, 125.1, 124.9, 120.0, 119.7, 119.3, 67.2, 66.3, 53.9, 53.3, 47.2, 29.9, 27.4, 27.1

HRMS (m/z): [M+Na] Calculated: 432.1423

Found: 432.1390

1-allyl 5- (2-oxo-2H-chromen-7-yl)-2-((((9H-fluoren-yl)methoxy)carbonyl)amino)pentanedioate (4.17)



To a 100-mL round-bottomed flask containing 7 mL of DMF, compound **4.16** (0.2 g, 0.49 mmol), TBTU (0.47 g, 1.47 mmol), HOBT (0.23 g, 1.47 mmol) and 7-hydroxycoumarin (0.12 g, 0.74 mmol) were added and the reaction mixture stirred overnight. Water (50 mL) was then added and an extraction was performed using ethyl acetate. The organic phase was collected and washed with brine, dried over magnesium sulfate and evaporated under vacuum. The product was purified using flash column chromatography with an eluent of 65:45 hexanes:ethyl acetate which afforded a white powder in 75% yield.

M.P.: 119-123 °C

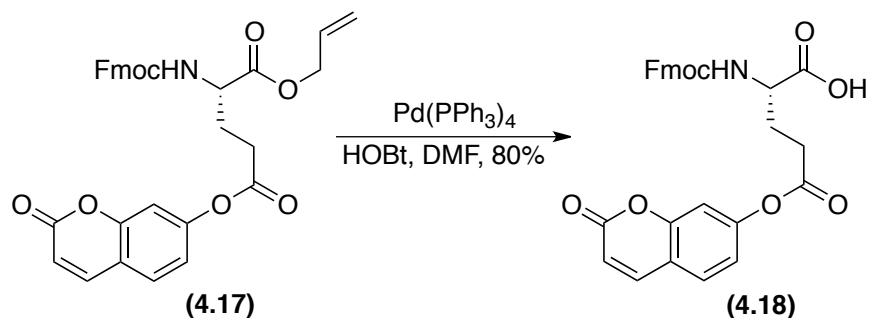
¹H NMR (400 MHz, (CD₃)₂CO) δ (ppm): 7.95 (d, 1H), 7.84 (d, 2H), 7.68-7.62 (m, 3H), 7.41-7.36 (m, 2H), 7.32-7.27 (m, 2H), 7.17-7.11 (m, 2H), 6.37 (d, 1H), 6.00-5.89 (m, 1H), 5.31 (d, 1H), 5.21 (d, 1H), 4.64-4.61 (m, 2H), 4.45-4.41 (m, 1H), 4.36 (d, 2H), 4.23 (t, 1H), 2.85-2.81 (m, 2H), 2.35-2.31 (m, 1H), 2.15-2.10 (m, 1H)

^{13}C NMR (100 MHz, $(\text{CD}_3)_2\text{CO}$) δ (ppm): 171.4, 170.6, 159.4, 156.4, 154.8, 153.4, 144.2, 144.1, 143.3, 141.3, 132.4, 129.1, 127.7, 127.1, 127.1, 125.3, 120.0, 118.4, 117.5, 116.9, 115.8, 110.1, 66.4, 65.3, 53.4, 47.2, 30.1, 27.2

HRMS (m/z): $[\text{M}+\text{Na}]$ Calculated: 576.1635

Found: 576.1642

2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-5-oxo-5-((2-oxo-2H-chromen-7-yl)oxy)pentanoic acid (4.18)



To a 100-mL round-bottomed flask containing 15 mL of dry dichloromethane, compound **4.17** (0.43 g, 0.78 mmol), HOBT (0.12 g, 0.78 mmol) and $\text{Pd}(\text{PPh}_3)_4$ (0.09 g, 0.078 mmol) were added and the reaction mixture stirred under an inert atmosphere until the complete disappearance of starting material was observed. The reaction mixture was filtered and the dichloromethane was evaporated under vacuum. The product was purified using flash column chromatography (6:3:1 hexanes:toluene:ethyl acetate) to obtain an off-white powder in 80% yield.

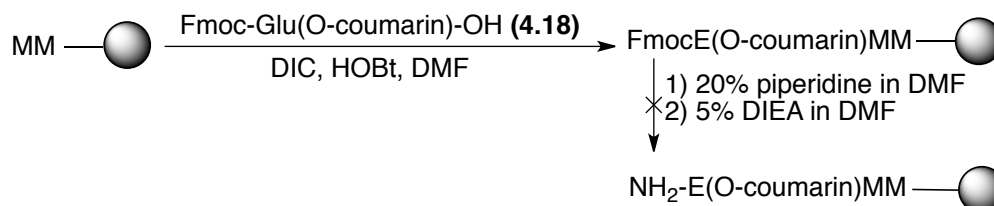
M.P.: 124-128 °C

^1H NMR (400 MHz, $(\text{CD}_3)_2\text{CO}$) δ (ppm): 8.01 (d, 1H), 7.73 (d, 2H), 7.61-7.65 (m, 3H), 7.46-7.42 (m, 2H), 7.35-7.32 (m, 2H), 7.20-7.17 (m, 2H), 6.37 (d, 1H), 4.64-4.61 (m, 2H), 4.45-4.41 (m, 1H), 4.36 (d, 2H), 4.23 (t, 1H), 2.85-2.81 (m, 2H), 2.35-2.31 (m, 1H), 2.15-2.10 (m, 1H)

^{13}C NMR (100 MHz, $(\text{CD}_3)_2\text{CO}$) δ (ppm): 171.4, 170.6, 159.4, 156.4, 154.8, 153.4, 144.2, 144.1, 143.3, 141.3, 129.1, 127.7, 127.1, 127.1, 125.3, 120.0, 117.5, 116.9, 115.8, 110.4, 67.1, 51.5, 47.2, 30.1, 27.6

HRMS (m/z): $[\text{M}+\text{Na}]$ Calculated: 536.1232

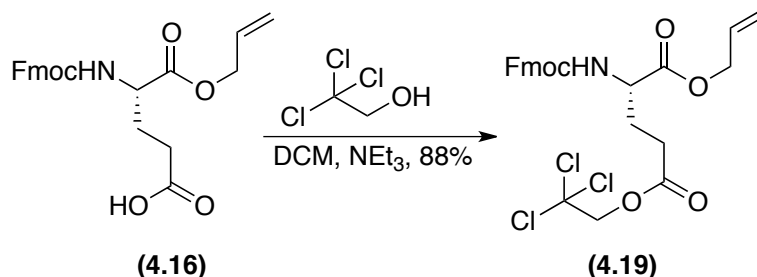
Found: 536.1311



The solid phase synthesis was carried out according to standard coupling procedures already mentioned. It was found that standard deprotection conditions resulted in the cleavage of the ester bond to release the coumarin fluorophore. Modified deprotection conditions were attempted using 5% DIEA in DMF incubated with 1 g of resin for a total of 1 min. These modified conditions also resulted in cleavage of the ester bond, releasing the coumarin molecule.

6.4.6.3 Third attempt towards activity assay using Fmoc-Glu(OTce)-OH

(*S*)-1-allyl 5-(2,2,2-trichloroethyl) 2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)pentanedioate (4.19**)**



To a 100-mL round-bottomed flask containing compound **4.16** (1 g, 2.4 mmol), was added DMAP (0.15 g, 1.2 mmol), DCC (0.66 g, 3.2 mmol) and 2,2,2-trichloroethanol (0.74 mL). The reaction was stirred overnight in 50 mL of DCM at room temperature. It was then filtered and the filtrate washed three times with 0.1 M HCl. The organic phase was dried over magnesium sulfate and then evaporated under vacuum. The product was purified using column chromatography with an eluent of 6:4 hexanes:ethyl acetate and obtained as a cream-white solid in 88% yield.

M.P.: 68-71 °C

¹H NMR (400 MHz, CDCl₃) δ (ppm): 7.76 (d, 2H), 7.65-7.63 (m, 2H), 7.35 (t, 2H), 7.28 (t, 2H), 5.94-5.85 (m, 1H), 5.3 (d, 1H), 5.19 (d, 1H), 4.60 (d, 2H), 4.35-4.30 (m, 2H), 4.28-

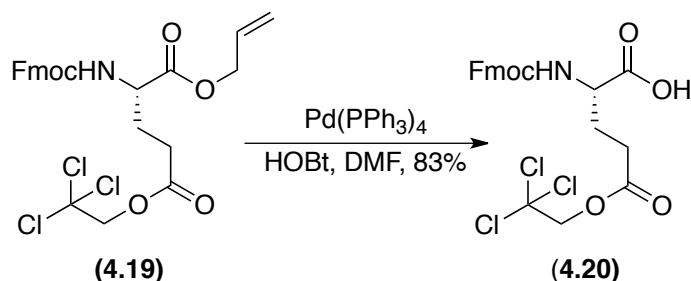
4.24 (m, 1H), 4.22-4.17 (m, 1H), 2.40-2.37 (m, 2H), 2.19-2.10 (m, 1H), 1.96-1.88 (m, 1H).

^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 171.5, 171.2, 143.8, 143.6, 141.3, 141.3, 131.3, 127.8, 127.1, 125.1, 120.1, 120.1, 119.3, 99.6, 94.8, 74.1, 67.2, 66.4, 64.5, 53.2, 47.1, 30.6, 29.9, 27.5.

HRMS (m/z): $[\text{M}+\text{Na}]$ Calculated: 562.0567

Found: 562.0558

(*S*)-2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-5-oxo-5-(2,2,2-trichloroethoxy)pentanoic acid.



To a 100-mL round-bottomed flask containing 15 mL of dry dichloromethane, compound **4.19** (0.43 g, 0.78 mmol), HOBT (0.12 g, 0.78 mmol) and $\text{Pd(PPh}_3)_4$ (0.09 g, 0.078 mmol) were added and the reaction mixture stirred under an inert atmosphere until the complete disappearance of starting material was observed. The reaction mixture was filtered and the dichloromethane was evaporated under vacuum. The

product was purified using flash column chromatography (6:4 hexanes:ethyl acetate) and obtained in 83% yield as a white powder.

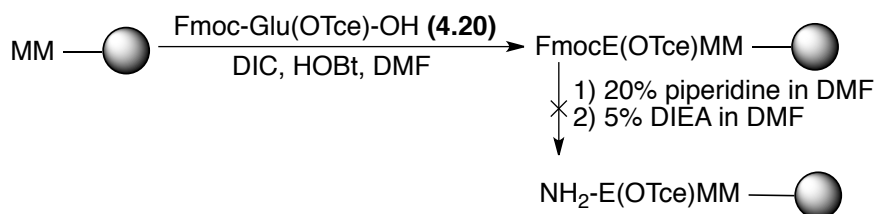
M.P: 107-110 °C

^1H NMR (400 MHz, CDCl_3) δ (ppm): 7.73 (d, 2H), 7.56-7.51 (m, 2H), 7.35 (t, 2H), 7.28 (t, 2H), 4.7 (s, 2H), 4.60 (d, 2H), 4.35-4.30 (m, 1H), 4.22-4.17 (m, 1H), 2.61-2.53 (m, 2H), 2.35-2.30 (m, 1H), 2.27-2.23 (m, 1H).

^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 176.9, 171.2, 156.2, 143.8, 141.3, 141.33, 132.5, 128.8, 127.1, 125.0, 120.0, 120.0, 115.7, 94.7, 74.2, 67.2, 53.1, 47.1, 30.0, 29.93, 27.3.

HRMS (m/z): $[\text{M}+\text{Na}]$ Calculated: 522.0254

Found: 522.0251



The solid phase synthesis was carried out according to standard coupling procedures already mentioned. Compound **4.20** was coupled to the resin using the standard conditions. Deprotection of the Fmoc protecting group using standard conditions of 20% piperidine in DMF (3 × 5 min) led to cleavage of the trichloroethyl

ester. The presence/absence of the trichloroethyl ester was confirmed by ^{13}C NMR, monitoring the distinctive $\text{C}(\text{Cl}_3)$ peak. Milder deprotection conditions were attempted, namely 5% DIEA in DMF for 1 min, but this also proved unsuccessful.

References

- (1) Butz, L. W.; Vigneaud, du, V. *Journal of Biological Chemistry* **1932**, 99, 135–142.
- (2) Kaiser, E.; Colescott, R. L.; Bossinger, D.; Cook, I. P. *Analytical Biochemistry* **1970**, 34.
- (3) Lai, T.-S.; Achyuthan, K. E.; Santiago, M. A.; Greenberg, C. S. *Journal of Biological Chemistry* **2001**, 269, 24596–24601.
- (4) Laemmli, U. K. *Nature* **1970**, 227, 680–685.
- (5) Malatesia, L.; Angoletta, M. *Journal of Chemical Society* **1957**, 1186–1188.