

Design and Synthesis of Protein Chemical Crosslinkers: A Modular Approach

Kayla Monique Downey

Bachelor of Science (Molecular and Drug Design)

Supervisors:

Assoc. Prof. Tara Pukala

Prof. Andrew Abell

Thesis submitted for the degree of Master of Philosophy



THE UNIVERSITY
of ADELAIDE

21st November 2017

School of Physical Sciences

The University of Adelaide

Contents

Abstract	iii
Declaration	v
Acknowledgements	vi
Abbreviations	vii
List of Figures	viii
List of Schemes	x
1 Introduction	1
1.1 Protein Structure	1
1.2 Methods for Determining Protein Structure	2
1.2.1 Mass Spectrometry and Protein Structure	4
1.3 Chemical Crosslinking Mass Spectrometry	5
1.3.1 The CXMS Workflow	7
1.3.2 Advantages and Challenges of CXMS	9
1.3.3 Crosslinker Design	13
1.4 Mass Spectrometry	17
1.4.1 The Ion Source	18
1.4.2 Mass Analysers	20
2 Design of a modular synthetic protocol for crosslinker diversification	25
2.1 Introduction	25
2.1.1 Aims	26
2.2 Design of a modular synthetic protocol	26
2.2.1 Retrosynthetic analysis of the target crosslinker	28
2.3 Crosslinker 1a	32
2.3.1 Synthesis of Crosslinker 1a	32
2.4 Conclusions and Future Work	38

2.5	Experimental	40
2.5.1	General Methods	40
2.5.2	Synthesis and Characterisation	42
3	Synthesis of cleavable crosslinkers and crosslinking of <i>Staphylococcus aureus</i> biotin protein ligase	47
3.1	Introduction	47
3.1.1	Cleavable bonds	48
3.1.2	<i>Staphylococcus aureus</i> biotin protein ligase	51
3.1.3	Aims	53
3.2	Crosslinker 1b	54
3.2.1	Synthesis of Crosslinker 1b	56
3.2.2	Dissociation of the cleavable bond using CID	62
3.3	Crosslinker 1c	63
3.3.1	Synthesis of Crosslinker 1c	65
3.3.2	Dissociation of the cleavable bond using CID	69
3.4	Crosslinking of <i>SaBPL</i>	71
3.5	Conclusions and Future Work	74
3.6	Acknowledgements	75
3.7	Experimental	75
3.7.1	Mass Spectrometry	75
3.7.2	<i>SaBPL</i>	76
3.7.3	General Synthetic Methods	77
3.7.4	Synthesis and Characterisation	78
	Summary	82
	Bibliography	84
	A NMR Spectra	95

Abstract

The study of protein structure and interactions is pivotal in understanding the function and malfunction of complex biological systems. The structures of some proteins are unable to be determined using traditional high resolution biophysical techniques, requiring the development of amenable low resolution alternatives. Chemical Crosslinking Mass Spectrometry (CXMS) is one technique which can be used to probe protein structure through the formation of covalent linkages between protein residues. The formation of these links is facilitated by chemical crosslinking reagents.

Widespread use of the CXMS technique has been hampered primarily by analytical challenges pertaining to the detection and identification of crosslinked species using Mass Spectrometry (MS). Attempts to mitigate the challenges have been made by modifying the structure of chemical crosslinkers through the addition of functional groups such as affinity tags, isotope labels and cleavable bonds. Crosslinkers combining more than one type of functional group (combination crosslinkers) present the most promising targets for CXMS applications, combining the benefits of each functional group. However, combination crosslinkers are not commercially available, thus necessitating in-house synthesis. Incorporating more than one functionality also results in more complex molecular structures and synthetic processes, making the crosslinkers difficult to adapt to suit a particular experiment. Consequently, the use of combination crosslinkers has been limited to date to a small number of studies.

The research presented in this thesis describes the development of a modular chemical crosslinker design and corresponding synthetic protocol for the synthesis of combination crosslinkers. The modular crosslinker structure can be readily modified to include a range of functional groups using a small number of different reactions, including amide coupling and O-alkylation, and commercially available starting materials such as Boc-serine, from a minimum of five synthetic steps. The utility of the synthetic process was validated through the synthesis of a crosslinker containing an alkyne functional group, which can be used to attach a biotin affinity tag through alkyne-azide Huisgen Cyclisation.

Synthesis of two custom designed combination crosslinkers utilising alkyne tags and cleavable bonds is also described. The function of the cleavable bonds was established using collision induced dissociation processes within the mass spectrometer. Ensuring that a crosslinker is effective in probing quaternary structure and protein-protein interactions is essential as the investigation of these structures is a major goal of CXMS. Therefore, a crosslinking assay using *Staphylococcus aureus* biotin protein ligase, which forms homodimers when substrate bound, was also developed using the commercially available crosslinkers Disuccinimidyl Suberate (DSS) and Dithiobis(succinimidyl) Propionate (DSP), to enable the efficacy of crosslinkers synthesised using the modular synthetic protocol to be determined.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the Universitys digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Kayla Monique Downey

21st November 2017

Acknowledgements

I would like to thank my supervisors, Dr. Tara Pukala and Professor Andrew Abell, for their constant support and encouragement over the last two years. Thank you for being great guides, putting up with some tears, and for laughing with (or at) me. I would also like to thank the Pukala and Abell Groups (special mention to the ground floor Badger labs) and the people in chemistry with whom I have formed great friendships, for the company, laughs and guidance. You have all made these two years colourful and flavoursome, like a good soup.

I am grateful to the School of Physical Sciences Staff, and to the Technical Staff of the Badger and Johnson buildings for administrative and technical assistance throughout this journey.

I acknowledge and thank Steven Polyak and Louise Sternicki for their contribution to this thesis regarding the work with Biotin Protein Ligase. Your contributions are greatly appreciated.

This thesis would not exist without the tireless support of my family and friends. Thank you to my Mum and Dad, your endless love, support and nagging has guided me to this point. Thank you for everything. You have always believed in me, know that it is appreciated. To my Nanna, thank you for your love, encouragement and pride in all of my achievements. Your phone calls and text messages before every university event showed me how much you care, I will always be grateful for it. I am thankful everyday for my patient and understanding friends. To them I present the source of cancelled plans and unsociable behaviour, I promise now I will try to remember to reply to text messages.

Finally, thank you to my year 10 science teacher Clive Dobson and year 11 chemistry teacher Jane Nykke, for instilling a passion for science, chemistry and learning that will remain with me for life.

Abbreviations

Arom	Aromatic Chemical Shift
ATP	Adenosine Triphosphate
Boc	tert-Butyloxycarbonyl
BPL	Biotin Protein Ligase
CID	Collision Induced Dissociation
Cryo-EM	Cryo-electron Microscopy
CXMS	Cross Linking Mass Spectrometry
CXL	Chemical Cross Linking
Da	Daltons
kDa	kilo-Daltons
DC	Direct Current
DDA	Dithiodiglycolic Acid
DEA	Diethylamine
DCM	Dichloromethane
DIC	N,N-diisopropylcarbodiimide
DMF	Dimethylformamide
DNA	Deoxyribonucleic Acid
DSP	Dithiobis(succinimidyl) Propionate
DSS	Disuccinimidyl Suberate
EDC-HCl	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide Hydrochloride
ECD	Electron Capture Dissociation
EI	Electron Ionisation
eq	Equivalents
ESI	Electrospray Ionisation
FAB	Fast Atom Bombardment
Fmoc	Fluorenylmethyloxycarbonyl
h	hour(s)
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HDX	Hydrogen/Deuterium Exchange

HDX-MS	Hydrogen/Deuterium Exchange Coupled to Mass Spectrometry
HRMS	High Resolution Mass Spectrometry
HTH	Helix-Turn-Helix
HOBt	1-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
IMMS	Ion Mobility Mass Spectrometry
IR	Infrared
MALDI	Matrix Assisted Laser Desorption Ionisation
MCP	Microchannel Plate
MeO	Methyl Ester Protecting Group
MeOH	Methanol
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MS/MS/MS	3-Stage Tandem Mass Spectrometry
m/z	Mass to Charge Ratio
nESI	Nano-Electrospray Ionisation
NHS	N-hydroxy Succinimide
NMR	Nuclear Magnetic Resonance Spectroscopy
PBS	Phosphate-Buffered Saline
Q-ToF	Quadupole-Time of Flight
R _f	Retention Factor
RF	Radio Frequency
Rxn	Reaction
SaBPL	<i>Staphylococcus aureus</i> Biotin Protein Ligase
SAXS	Small Angle X-ray Scattering
SOCl ₂	Thionyl Chloride
TFA-NHS	N-trifluoroacetoxy Succinimide
TFAA	Trifluoroacetic Anhydride
THF	Tetrahydrofuran
ToF	Time of Flight
TLC	Thin Layer Chromatography
WT	Wild Type

List of Figures

1.1	Levels of protein structure	2
1.2	Simple crosslinker structure	6
1.3	Chemical crosslinking mass spectrometry workflow	8
1.4	Crosslink types	9
1.5	MS/MS fragmentation of type 2 crosslinked species	12
1.6	Isotope labelled crosslinker BSG- d_4	13
1.7	Cleavable crosslinkers DSSO and DSP	14
1.8	Biotin tagged crosslinker BCCL1	16
1.9	Bifunctionalised, and trifunctionalised crosslinkers EGS- d_{12} and CB-DPS	17
1.10	Schematic diagram of the Waters QToF2 mass spectrometer	18
1.11	Diagrammatic representation of the ESI process	20
1.12	The quadrupole mass analyser	21
1.13	Peptide fragmentation in positive mode using CID	23
1.14	Peptide fragmentation in negative mode using CID	24
2.1	Reactive group types	27
2.2	General crosslinker 1	28
2.3	Retrosynthesis of the biotinylated general crosslinker	29
2.4	Retrosynthesis of biotin azide	29
2.5	Retrosynthesis of general crosslinker 1	30
2.6	Retrosynthesis of dimethyl ester 5	31
2.7	Retrosynthesis of hydrochloride salt 6	31
2.8	Crosslinker 1a	32
2.9	NMR comparison of 10 and 11	38
3.1	Sequencing of individual peptides from a type 2 crosslink	49
3.2	Marker ion formation from gas phase dissociation of DSP	50
3.3	X-ray crystal structure of <i>Sa</i> BPL with lysines highlighted	52
3.4	Distances between the two lysine 99 residues and the two lysine 100 residues of the <i>Sa</i> BPL homodimer to be probed using the synthesised crosslinkers	53

3.5	Crosslinker 1b	54
3.6	Crosslinker 1b CID dissociation products	55
3.7	NMR comparison of 13 and 14	60
3.8	Diester 14 CID dissociation products	62
3.9	MS/MS spectrum of diester 14	63
3.10	Crosslinker 1c	63
3.11	Crosslinker 1c CID dissociation products	64
3.12	¹ H COSY spectrum of 16	67
3.13	Monoester 16 CID dissociation products	70
3.14	MS/MS spectrum of monoester 16	70
3.15	Commercially available crosslinkers DSS and DSP	71
3.16	SDS-PAGE analysis of <i>Sa</i> BPL crosslinking reaction mixtures	73
A.1	¹ H NMR spectrum of diester 5 in CDCl ₃	95
A.2	¹ H NMR spectrum of monoester 7 in DMSO- <i>d</i> ₆	96
A.3	¹ H NMR spectrum of diester 10	96

List of Schemes

2.1	Synthetic scheme for Crosslinker 1a	33
2.2	<i>O</i> -alkylation of Boc-L-Serine	34
2.3	One pot N-Boc cleavage and esterification of 3	34
2.4	HATU mediated peptide coupling of momomethyl adipate to 6	35
2.5	Methyl ester hydrolysis of 9	35
2.6	Functionalisation of 10 with NHS	37
3.1	Synthetic scheme for Crosslinker 1b	56
3.2	Amide coupling to form monoacid 12	57
3.3	Esterification of 12 to form 14	59
3.4	Hydrolysis of 14 to form 15	60
3.5	Synthetic scheme for Crosslinker 1c	65
3.6	Amide coupling to form monoester 16	66
3.7	Fmoc deprotection and amide coupling to form diester 18	69