

This is to certify that the
thesis entitled


IONIC CROSS-LINKING REAGENTS AND TANDEM MASS
SPECTROMETRY FOR MAPPING STRUCTURES OF
PROTEINS AND PROTEIN COMPLEXES

presented by

YALI LU

has been accepted towards fulfillment
of the requirements for the

M.S. degree in Chemistry


Major Professor's Signature

4/16/08
Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

**IONIC CROSS-LINKING REAGENTS AND TANDEM MASS SPECTROMETRY
FOR MAPPING STRUCTURES OF PROTEINS AND PROTEIN COMPLEXES**

By

YALI LU

A THESIS

**Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of**

MASTER OF SCIENCE

Department of Chemistry

2008

A

rea

ord

the

syn

sub

MS

frag

was

link

each

of t

iden

prod

cha

link

cle

bac

ans

ABSTRACT

IONIC CROSS-LINKING REAGENTS AND TANDEM MASS SPECTROMETRY FOR MAPPING STRUCTURES OF PROTEINS AND PROTEIN COMPLEXES

BY

YALI LU

A series of novel 'fixed charge' sulfonium ion-containing 'amine reactive' cross-linking reagents were synthesized, characterized, and initially applied to three model peptides, in order to develop a 'targeted' multistage tandem mass spectrometry based approach for the identification and characterization of protein-protein interactions. Among them, symmetrical S-Methyl 5,5'-dipentanoylhydroxysuccinimide sulfonium iodide (or methyl sulfate) exhibits the best stability to hydrolysis and alcohoholysis. Under low energy CID-MS/MS conditions, peptide ions containing this cross-linker are shown to exclusively fragment via facile cleavage of the bond directly adjacent to the 'fixed charge', which was found to be independent of the precursor ion charge state. Inter-molecular cross-linked peptides are readily identified by the formation of two characteristic peptide ions, each containing an additional mass modification corresponding to the remaining portion of the fragmented cross-linking reagent. Peptides containing 'dead-end' cross-links are identified based on the observation of a characteristic neutral loss. MS³ of the initial products formed from these reactions can be used to provide information required for characterization of the peptide sequences and modification site(s) involved in the cross-linking reactions. MS/MS of intra-molecular cross-linked peptides results in initial cleavage of the cross-linker, followed by immediate further fragmentation of the peptide backbone, thereby allowing detailed structural information to be obtained by direct analysis of the MS/MS product ion spectrum.

com

Gav

Stat

thro

mer

valu

exp

mar

link

Tar

ins

pro

goe

Ra

di

de

ha

du

ACKNOWLEDGMENTS

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. First I would like to express my deep thanks to my advisor, Dr. Gavin E. Reid, for his guidance and support during my research and study at Michigan State University. His perpetual energy and enthusiasm in research had motivated me throughout this work. I also wish to express my sincere thanks to other committee members, Dr. Gary J. Blanchard, Dr. Babak Borhan, and Dr. A Daniel Jones for their valuable advice and friendly help through extensive discussions and interesting explorations around my work.

I would like to give my special thanks to Ms. Marina Tanasova who taught me many things in organic synthesis and helped me realize the preparation of ionic cross-linking reagents. This project could not have been completed without the efforts of Ms. Tanasova. The previous and current group members of the Reid research group have inspired me in research and life through our interactions during long hours in the lab, providing me a convivial place to work. Thanks. In particular, my special appreciation goes to Amanda Palumbo, Jennifer M. Froelich, Suchitra Kaplinghat, and Katherine Rank for their friendship, encouragement, and helpful discussions along this journey.

Finally, I wish to thank my family and friends for helping me get through the difficult times, and for all the emotional support and caring they provided. I owe my deepest thanks to my husband, Gaochuang Yang, whose love, support and understanding has always been along with me. Thanks all who had accompanied me and supported me during this not easy but rewarding process.

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	viii
LIST OF SCHEMES.....	xiv
CHAPTER ONE	
INTRODUCTION	1
1.1 Protein Structure and Function	1
1.2 Current Methods Employed for Protein Structure Analysis	2
1.3 Chemical Cross-Linking	2
1.3.1 Cross-Linking Reagents	4
1.3.2 Possible Cross-linked Products from a Cross-Linking Reaction	5
1.3.3 Challenges associated with the identification and characterization of chemically cross-linked peptides	8
1.3.4 Affinity Cross-Linking Reagents	8
1.3.5 Isotope Labeled Cross-Linking Reagents	9
1.3.6 Cleavable Cross-Linking Reagents	11
1.4 Tandem Mass Spectrometry (MS/MS)	14
1.4.1 Quadrupole Mass Analyzer	15
1.4.2 The Quadrupole Ion Trap	17
1.4.3 Multiple Stage Tandem Mass Spectrometry (MS ⁿ) Realized in Ion Trap	18
1.5 Gas Phase Fragmentation Reactions of Protonated Cross-linked Peptide Ions.....	19
1.5.1 Mechanisms for the Fragmentation of Protonated Peptide Ions	20
1.5.2 Controlling Peptide Fragmentation Reactions via the use of Fixed Charge Derivatization	23
1.6 Aims of this Thesis	25
CHAPTER TWO	
EXPERIMENTAL.....	27
2.1 Materials	27
2.2 Mass Spectrometry	28
2.3 Synthesis of Ionic Cross-Linking Reagent S-Methyl 2-acetoyle,5'-pentanoyldihydroxysuccinimide sulfonium (Compound I in Scheme 2.1)	28
2.3.1 Carboxymethylmercaptopropionic Acid (1)	29
2.3.2 Thio 2-acetoyle-5'-pentanoyldihydroxysuccinimide (2)	31
2.3.2 S-Methyl 2-acetoyle-5'-pentanoyldihydroxysuccinimide Sulfonium (I)	34
2.4 Synthesis of Ionic Cross-Linking Reagent S-Methyl 3-propionoyl,5'-pentanoyldihydroxysuccinimide sulfonium (Compound II in Scheme 2.6).....	36
2.4.1 1-Carboxyethylmercaptopropionic Acid (3).....	36
2.4.2 Thio 3-propionoyl-5'-pentanoyldihydroxysuccinimide (4)	37
2.4.3 S-Methyl 3-propionoyl-5'-pentanoyldihydroxysuccinimide Sulfonium (II) ...	38

2.5 Synthesis of Ionic Cross-Linking Reagent S-Methyl 5,5'-dipentanoylhydroxysuccinimide sulfonium (Compound III in Scheme 2.8)	40
2.5.1 5-Mercaptopentanoic Acid (5).....	41
2.5.2 5,5'-Thiodipentanoic Acid (6).....	42
2.5.3 5,5'-Thiodipentanoylhydroxysuccinimide (7).....	44
2.5.4 S-methyl 5,5'-Thiodipentanoylhydroxysuccinimide (III)	45
2.6 Peptide Cross-Linking Reactions.....	46

CHAPTER THREE

ALKYL CHAIN LENGTH EFFECTS ON THE STABILITY AND FRAGMENTATION BEHAVIOR OF IONIC CROSS-LINKING REAGENTS.....

3.1 Introduction.....	48
3.2 Cross-Linking Reagent S-Methyl 2-acetoxy,5'-pentanoyldihydroxysuccinimide sulfonium (Compound I in Scheme 2.1).....	50
3.2.1 Stability of I following Reaction with Iodomethane.....	50
3.2.2 Transesterification and Hydrolysis Reactions of I.....	52
3.2.3 Supplementary Evidence from 1-Carboxymethylmercaptopropionic Acid (Compound 1 in Scheme 2.1).....	55
3.3 Cross-Linking Reagent S-Methyl 3-propionoyl,5'-pentanoyldihydroxysuccinimide sulfonium (Compound II in Scheme 2.6).....	57
3.3.1 Stability in Iodomethane	57
3.3.2 Stability under Aqueous Conditions	57
3.3.3 Supplementary Evidence from 1-Carboxyethylmercaptopropionic Acid (Compound 3 in Scheme 2.6).....	58
3.3.4 Suspected E2 Elimination Reactions Might Lead to Unintended Cleavage on Cross-Linker.....	60
3.4 Cross-Linking Reagent S-Methyl 5,5'-dipentanoylhydroxysuccinimide sulfonium (Compound III in Scheme 2.8).....	61
3.4.1 Stability in Aqueous Conditions and Methanol	61
3.4.2 Comparison among Three Ionic Cross-Linking Reagents and Their Intermediates	62
3.5 Cross-Linked Neurotensin	63
3.5.1 Cross-Linking Reaction with BID-NHS	63
3.5.2 Cross-Linking Reaction with Ionic Cross-Linking Reagent I and II	66

CHAPTER FOUR

IDENTIFICATION OF CROSS-LINKED PEPTIDES FOR PROTEIN INTERACTION STUDIES USING A SYMMETRICAL IONIC CROSS-LINKING REAGENT AND MASS SPECTROMETRY.....

4.1 Introduction.....	71
4.2 Cross-Linking Reaction with Neurotensin	73
4.2.1 Cross-Linking Reactions in DMF and PBS	73
4.2.2 Inter-Molecular Cross-Linked Neurotensin	74
4.2.3 Mono-Linked Neurotensin	76
4.3 Cross-Linking Reaction with Insulin-Like Growth Factor I (57-70)	78
4.4 Cross-Linking Reaction with VTMAHFWNFGK (MWK).....	83

4.4.1 Intra-Molecular Cross-Linked MWK	84
4.4.2 Mono-Linked MWK	85
4.4.3 Double Hydrolyzed Mono-Link.....	87
4.4.4 Inter-Molecular Cross-Link	89
4.5 Summary	91
REFERENCES	93

LIST OF TABLES

TABLE		PAGE
1.1	Reactive cross-linker groups and their functional group targets. Adapted from Reference 7.....	4

1

LIST OF FIGURES

FIGURE	PAGE
1.1 Cross-linking methodology flowchart by the bottom-up approach. Adapted from Reference 6.....	3
1.2 (A) Classification of cross-linked peptides into Type 0, Type 1 and Type 2. (B) Multiple modifications by the combinations of Type 0, 1 and 2. Adapted from Reference 11.....	7
1.3 Typical stability diagram for quadrupole ion trap. The larger balls represent high m/z ions whereas smaller balls represent the low m/z ions. Reproduced and modified from Reference 29.....	18
2.1 Characterization of (1) by nanoelectrospray quadrupole ion trap mass spectrometry and CID MS/MS. (A) MS and (B) CID-MS/MS of m/z 191.1.....	30
2.2 Characterization of (2) by nanoelectrospray quadrupole ion trap mass spectrometry and MS/MS. (A) Mass spectrum of (2), due to the poor protonation of diester, the $[M+Na]^+$ adduct is the most intensive peak; (B) CID MS/MS of $[M+Na]^+$ adduct, neutral loss of 115 was observed; (C) CID MS/MS of $[M+NH_4]^+$ adduct, yielding $[M+H]^+$ and m/z at 272^+ fragment ions; (D) CID MS ³ of $[M+H]^+$ coming from $[M+NH_4]^+$, giving rise to the fragment at m/z of 272^+ , by neutral loss of 115.....	33
2.3 Characterization of ionic cross-linking reagent I by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The methyl sulfonium ion is at m/z 401^+ ; (B) CID-MS/MS product ion spectrum of M^+	35
2.4 Characterization of (3) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The pseudo molecular ion $[M-H]^-$ at m/z 205.1 ⁻ and the dimer at m/z 410.9 ⁻ ; (B) CID-MS/MS of selected parent ion at m/z 205.1 ⁻	37
2.5 Characterization of (4) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) Mass spectrum of (2), $[M+NH_4]^+$ adduct is the most abundant peak due to poor protonation of the diester; (B) CID MS/MS of $[M+NH_4]^+$ adduct, neutral loss of NH_3 and further loss of 115 was observed.....	38

2.6	Characterization of ionic cross-linking reagent II by nanoelectrospray quadrupole ion trap mass spectrometry. (A) Mass spectrum of methyl sulfonium ion at m/z 415 ⁺ ; (B) CID-MS/MS product ion spectrum of M ⁺ ion.....	39
2.7	Characterization of phenacylsulfonium ion derivative of (II) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) Mass spectrum of phenacylsulfonium ion as an alternative for compound II, showing M ⁺ at m/z 519 ⁺ ; (B) CID-MS/MS yields the same product ion at m/z 198.....	40
2.8	Characterization of (5) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The pseudomolecular ion [M-H] ⁻ at m/z 133.1 ⁻ and the dimer at m/z 266.8 ⁻ ; (B) CID-MS/MS of selected parent ion at m/z 132.9 ⁻	42
2.9	Characterization of compound (6) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The pseudo molecular ion [M-H] ⁻ at m/z 233.2 ⁻ and the dimer at m/z 466.9 ⁻ ; (B) CID-MS/MS of selected parent ion at m/z 233.1 ⁻	43
2.10	Characterization of (7) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The pseudomolecular ion [M+H] ⁺ at m/z 429.1 ⁺ , the NH ₄ ⁺ and Na ⁺ adducts are present; (B) The CID-MS/MS of selected parent ion at m/z 428.9 ⁺ , m/z at 314.0 is the main fragment, which corresponds to 314.2 fragment in source in panel A.....	45
2.11	Characterization of ionic cross-linking reagent III by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The methyl sulfonium ion at m/z 443 ⁺ ; (B) CID-MS/MS product ion spectrum of M ⁺	46
3.1	Characterization of dimethyl sulfonium 8 by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The dimethyl sulfonium ion at m/z 260 ⁺ , peak at m/z 404 corresponds to ammonium adduct of reactant 2 and m/z 272 is a fragment generated in source from it; (B) CID-MS/MS product ion spectrum of M ⁺	51
3.2	Nanoelectrospray quadrupole ion trap mass spectrum of sulfonium I methyl sulfate in 50 % MeOH + 50 % H ₂ O + 1 % AcOH MS buffer ...	53
3.3	Nanoelectrospray quadrupole ion trap mass spectra of sulfonium I methyl sulfate stored in dry DMF, 100 % CH ₃ CN was employed as MS solvent; (A) 0 min; (B) 45 min; (C) 75 min; (D) 2 hr; (E) 3 hr; (F) 4 hr	54

3.4	Nanoelectrospray quadrupole ion trap mass spectra from methylation of 1 with iodomethane. Similar mechanism as indicated in Scheme 3.2 to form dimethyl sulfonium. (A) The dimethyl sulfonium ion at m/z 163 ⁺ , peak at m/z 101 corresponds to a fragment in source; (B) CID-MS/MS product ion spectrum of ion at m/z 163 ⁺	56
3.5	Nanoelectrospray quadrupole ion trap mass spectra of sulfonium II iodide. (A) In 100 % CH ₃ CN buffer, small peak of dimethyl sulfonium (m/z = 260, compound 8 in Scheme 3.2) observed; (B) in 50 % CH ₃ CN + 50 % H ₂ O buffer, a small peak resulting from methyl transesterification (m/z = 318, compound 9 in Scheme 3.3) was observed.....	58
3.6	Nanoelectrospray quadrupole ion trap mass spectra from methylation of 3 with iodomethane. (A) Expected product 12 at m/z 221 is dominant peak, small amount of dimethyl sulfonium (m/z = 163) (Compound 11 in Scheme 3.5) by-product observed. (B) CID-MS/MS product ion spectrum of ion at m/z 221 ⁺	59
3.7	Nanoelectrospray quadrupole ion trap mass spectra of sulfonium III methylsulfate. (A) In 100 % CH ₃ CN buffer, small peak of methyl ester sulfonium (m/z = 360, analogous to compound 9 in Scheme 3.3) observed; (B) in 50 % CH ₃ CN + 50 % H ₂ O buffer; (C) in 50 % AcCN + 50 % H ₂ O + 1 % AcOH buffer; (D) in 50 % MeOH + 50 % H ₂ O + 1 % AcOH buffer.....	62
3.8	Low energy CID MS/MS of BID-NHS cross-linked neurotensin. (A) Precursor ion in the 5+ charge state, product ion at m/z 861 corresponds to the loss of benzyl cations; (B) precursor ion in the 4+ charge state, most abundant product ion at m/z 856 corresponds to peptide backbone cleavage between I and L residues; (C) precursor ion in the 3+ charge state, dominant product ion resulted from neutral loss of NH ₃ from N or R side chains; (D) precursor ion in the 2+ charge state, dominant product ion resulted from neutral loss of NH ₃ from N or R side chains.....	65
3.9	Low energy CID MS/MS spectra of ionic cross-linked neurotensin. (A) From cross-linker II; (B) from cross-linker I. The arrow indicates the precursor ions: (A) m/z 706.9, z = 5; (B) m/z 704.2, z = 5.....	68
3.10	Low energy CID MS ³ or MS/MS of neurotensin. (A) MS ³ of S modified neurotensin from cross-linker I, * indicates S _I modification; (B) MS ³ of S modified neurotensin from cross-linker II, ** indicates S _{II} modification (C) MS/MS of unmodified neurotensin in 2+ charge state.....	69

3.11	Low energy CID MS ³ or MS/MS of neurotensin. (A) MS ³ of ITP modified neurotensin, both I and II give same product ions; (B) MS/MS of unmodified neurotensin in 3+ charge state	70
4.1	Nano-electrospray quadrupole ion trap mass spectra of III cross-linked neurotensin. (A) Cross-linking reaction in DMF at 75 min with 1:2 molar ratio of cross-linker to peptide; (B) cross-linking reaction in PBS buffer at 90 min with 2.5:1 molar ratio of cross-linker to peptide. M: unmodified neurotensin, [2M+4H+ITP-S _{III}] ⁵⁺ (m/z 712.4): inter-molecular cross-link in 5+ charge [M+2H+ITP-S _{III} N] ³⁺ (m/z 667.6): unhydrolyzed mono-link in 3+ charge state, [M+2H+ITP-S _{III} OH] ³⁺ (m/z 635.3): hydrolyzed mono-link in 3+ charge state.....	74
4.2	Low energy CID MS/MS of inter-molecular cross-linked neurotensin by III prepared in DMF. (A) Precursor ion in 5+ charge state (m/z 712.5); (B) precursor ion in 4+ charge state (m/z 890.5). The arrow indicates peak of precursor ions.....	75
4.3	CID MS ³ product ion spectra of modified neurotensin in Figure 4.2 formed from inter-molecular cross-links. (A) [M+2H+ITP] ³⁺ ion; (B) [M+2H+S _{III}] ²⁺ ion; (C) [M+H+ITP] ²⁺ ion. ITP modification is labeled as #, S _{III} modification is labeled as *.....	76
4.4	Low energy CID MS/MS spectra of mono-linked neurotensin by III prepared in PBS buffer. (A) Unhydrolyzed mono-link, m/z 667.7, z = 3; (B) hydrolyzed mono-link, m/z 635.3, z = 3. The unassigned arrow indicates peaks of precursor ions	77
4.5	Low energy CID MS/MS spectra of hydrolyzed mono-link in 2+ charge state prepared by reaction of neurotensin with III in PBS buffer. The unassigned arrow indicates peak of precursor ions, m/z 952.3, z = 2.....	78
4.6	Nanoelectrospray quadrupole ion trap mass spectrum of III cross-linked IGF. Reaction performed in PBS buffer. M: unmodified IGF; [M+H+ITP-S _{III}] ²⁺ : intra-molecular cross-link in 2+ charge state; [M+H+ITP-S _{III} OH] ²⁺ : hydrolyzed mono-link in 2+ charge state.....	80
4.7	CID MS/MS spectrum of doubly charged intra-molecular cross-linked IGF. ITP modification is labeled as #, S _{III} modification is labeled as *..	81
4.8	Low energy CID MS/MS and MS ³ spectra of hydrolyzed mono-linked IGF. (A) CID MS/MS of hydrolyzed mono-linked IGF, unassigned arrow indicates peak of precursor ions (m/z 863.8, z = 2); (B) CID MS ³ product ion spectrum of [M+H+ITP] ²⁺ ion from hydrolyzed mono-linked IGF. ITP modification is labeled as #.....	82

4.9	Nanoelectrospray quadrupole ion trap mass spectrum of III cross-linked MWK. Reaction performed in PBS buffer. M: unmodified MWK; $[M+2H+ITP-S_{III}OH]^{3+}$ (m/z 523.3): hydrolyzed mono-link in 3+ charge state; $[M+2H+ITP-S_{III}N]^{3+}$ (m/z 555.9): unhydrolyzed mono-link in 3+ charge state; $[2M+4H+ITP-S_{III}]^{5+}$ (m/z 578.5): inter-molecular cross-link in 5+ charge; $[M+H+2ITP-S_{III}OH]^{3+}$ (m/z 600.3): double hydrolyzed mono-link in 3+ charge state; $[2M+3H+ITP-S_{III}]^{4+}$ (m/z 722.8): inter-molecular cross-link in 4+ charge; $[M+H+ITP-S_{III}]^{2+}$ (775.9): intra-molecular cross-link in 2+ charge state; $[M+H+ITP-S_{III}OH]^{2+}$ (m/z 784.3): hydrolyzed mono-link in 2+ charge state.....	84
4.10	CID MS/MS spectrum of intra-molecular cross-linked MWK. ITP modification is labeled as #, S_{III} modification is labeled as *.....	85
4.11	Low energy CID MS/MS spectra of mono-linked MWK by III prepared in PBS buffer. (A) Unhydrolyzed mono-link in 3+ charge state; (B) hydrolyzed mono-link in 3+ charge state; (C) unhydrolyzed mono-link in 2+ charge state; (B) hydrolyzed mono-link in 2+ charge state.....	86
4.12	CID MS ³ product ion spectra of ITP modified MWK in Figure 4.11 formed from hydrolyzed mono-link. (A) $[M+2H+ITP]^{3+}$ ion; (B) $[M+H+ITP]^{2+}$ ion. ITP modification is labeled as #; the unassigned arrow indicates the precursor ion, m/z 710.4, z = 2.....	87
4.13	Low energy CID MS/MS and MS ³ product ion spectra of double hydrolyzed mono-link of MWK. (A) CID MS/MS product ion spectrum of double hydrolyzed mono-link of MWK; (B) CID MS ³ product ion spectrum of $[M+H+ITP+ITP-S_{III}OH]^{3+}$ ion formed from double hydrolyzed mono-link in Figure 4.13A. The unassigned arrow indicates the precursor ion: (A) m/z 600.2, z = 3; (B) m/z 550.8, z = 3..	88
4.14	CID MS ⁴ product ion spectrum of double ITP modified MWK ion formed from double hydrolyzed mono-link in Figure 4.13. ITP modification is labeled as #; the unassigned arrow indicates the precursor ion, m/z 501.5, z = 3.....	89
4.15	Low energy CID MS/MS spectra of inter-molecular cross-linked MWK by III. (A) Precursor ion in 5+ charge state, m/z 578.2; (B) precursor ion in 4+ charge state, m/z 722.6. The arrow indicates peak of precursor ions.....	90
4.16	CID MS ³ product ion spectra of modified MWK in Figure 4.15A formed from inter-molecular cross-links. (A) $[M+2H+ITP]^{3+}$ ion; (B)	

$[M+2H+S_{III}]^{2+}$ ion. ITP modification is labeled as #, S_{III} modification is labeled as *	91
---	----

LIST OF SCHEMES

SCHEME	PAGE
1.1 Nomenclature for peptide fragment ions. Adapted from Reference 38..	21
1.2 Fragmentation mechanism at aspartyl-prolyl bond within SuDPG cross-linked peptides. Adapter from Reference 27.....	22
1.3 Fixed charge derivatization of methionine sulfonium ion derivatized peptide ions. Adapted and modified from Reference 47.....	24
1.4 Gas-phase fragmentation behavior of the sulfonium ion derivative of methionine-containing peptides. Adapted and modified from Reference 47.....	24
2.1 The synthesis route of ionic cross-linking reagent S-methyl 2-acetoxy-5'-pentanoyldihydroxysuccinimide sulfonium (I).....	29
2.2 Fragmentation scheme of (1) anions. The spectrum is displayed on Figure 2.1B. Square brackets indicate ion-molecule complexes.....	31
2.3 Fragmentation of Na ⁺ adduct of (2). The spectrum is displayed on Figure 2.2B. Square brackets indicate ion-molecule complexes.....	34
2.4 Fragmentation of protonated (2). The spectrum is displayed on Figure 2.2D. Square brackets indicate ion-molecule complexes. The protonated (2) is supposed to have the similar fragmentation pathway as Na ⁺ adduct of (2) for the same neutral loss of 115 observed but no proton transfer observed to give product ion at m/z of 116 due to NHS has lower affinity to proton than that of Na ⁺	34
2.5 Fragmentation of sulfonium ion (I). The spectrum is displayed on Figure 2.3B. Fragment of m/z at 198 ⁺ is the main process; otherwise the pathway for the neutral loss of 115 is also present.....	35
2.6 Synthesis route of ionic cross-linking reagent S-Methyl 3-propionoyl-5'-pentanoyldihydroxysuccinimide sulfonium (II).....	36
2.7 Fragmentation scheme of (3) anions. The spectrum is displayed on Figure 2.4B. Square brackets indicate ion-molecule complexes.....	37
2.8 Synthesis of ionic cross-linking reagent S-Methyl 5,5'-dipentanoyldihydroxysuccinimide sulfonium (III).....	41

2.9	Fragmentation mechanism of (5) anion. The spectrum is displayed on Figure 2.8B. Square brackets indicate ion-molecule complexes.....	42
2.10	Fragmentation mechanism of (6) anion. The spectrum is displayed on Figure 2.9B. Square brackets indicate ion-molecule complexes.....	44
3.1	Dimethyl sulfonium 8 obtained from methylation of 2 with iodomethane.....	51
3.2	Mechanism of dimethyl sulfonium formation by reaction of iodomethane with I.....	52
3.3	Structures of methyl transesterification and hydrolysis products of I....	53
3.4	Mechanism of hydrolysis of I, methanol has the similar process as a nucleophile for methyl transesterification.....	55
3.5	Methylation reaction of 1 with iodomethane and potential mechanisms for gas phase fragmentation of dimethyl sulfonium product.....	56
3.6	Methylation reaction of 3 with iodomethane and potential mechanisms for gas phase fragmentation of methylated product 12.....	59
3.7	Potential fragmentation mechanisms of II intermolecular cross-linked peptides.....	60
3.8	Structure of the cross-linker N-benzyliminodiacetylhydroxysuccinimid (BID-NHS). Adapted from Reference 26.....	63
3.9	Proposed Proposed mechanism for the formation of benzyl cations from BID-NHS cross-linked peptide. Adapted from Reference 26.....	64
3.10	Proposed fragmentation mechanism of I and II intermolecular cross-linked peptides. The resulting mass additions (relative to the single peptide) following cleavage on the C-S bond are indicated, where S denotes a sulfide modification and ITP denotes protonated iminotetrahydropyran modification; I and II indicate cross-linking reagent.....	67
4.1	Proposed fragmentation mechanism of III cross-linked peptides. The resulting mass additions (relative to the single peptide) following cleavage on the C-S bond are indicated, where S denotes a sulfide modification and ITP denotes protonated iminotetrahydropyran modification. (A)Intramolecular cross-link; (B) unhydrolyzed and hydrolyzed mono-link; (C) intermolecular cross-link.....	72

CHAPTER ONE

INTRODUCTION

1.1 Protein Structure and Function

The study of proteins has fascinated biochemists and chemists for over a century. Understanding the compositions, three-dimensional structures, and chemical activities of proteins are among some of the most pressing scientific issues, which ultimately aim at R&D of new drugs for human healthcare.

Proteins perform a wide variety of roles. The most important biological functions of proteins involve the biochemical catalysts-function as enzymes; storage and transport by binding other molecules; providing support and shape to cells; decoding information in the cell; serving as hormones or receptors for hormones, etc [1]. The key to appreciating how different proteins function in different ways lies in an understanding of protein structure. Protein structure is defined at four levels [1]. Primary structure describes the linear sequence of amino acid residues in a protein molecule. Secondary structure relates to the regularities maintained by hydrogen bonds between amide hydrogens and carbonyl oxygens of the peptide backbone. α helices and β sheets are two major secondary structures. The shape of the fully folded polypeptide chain defines the tertiary structure, which is stabilized by non-covalent interactions between amino acid side chains within the molecule and with water molecules surrounding it. Proteins can also work together to achieve a particular function by forming transient or stable multi-subunit complexes, resulted in the highest level of organization, quaternary structure. Identification of the

interacting partners and further mapping of the specific interaction sites are crucial to understand how proteins function together as biological assemblies [1].

1.2 Current Methods Employed for Protein Structure Analysis

Numerous techniques have been developed to probe protein-protein interactions. X-ray crystallography and multidimensional NMR are currently primarily applied due to their very high structural resolution. X-ray crystallography uses the diffraction pattern produced by bombarding a single crystal with beam of collimated X-rays to solve the crystal structure; therefore, a single crystal has to be obtained for X-ray studies. Alternatively, nuclear magnetic resonance (NMR) spectroscopy is capable of three-dimensional structure determination of proteins in solution. However, the limitations of NMR spectroscopy result from the low inherent sensitivity and from the high complexity and information content of NMR spectra. Other methods such as yeast two-hybrid screens and protein micro arrays provide very high throughput but structural resolution is low [2], large numbers of non-specific interactions are detected as well which provide little information regarding the specific interacting sites between the protein subunits [3].

1.3 Chemical Cross-Linking

Because of its high sensitivity and rapid analysis, chemical cross-linking combined with mass spectrometry is a promising approach to study the low-resolution structure of protein topology and protein interactions [4]. Cross-linkers can covalently link interacting regions within a protein or interacting regions between the individual components of multi-protein complexes. Most importantly, non-covalent protein-protein interactions,

which may be transient or dependent on specific physiological conditions, can be captured into long-lived covalent complexes [2]. Typically in a bottom-up strategy, cross-linked proteins or protein complexes are separated by one-dimensional gel electrophoresis (SDS-PAGE), and then subjected to enzymatic in-gel digestion, chromatographic fractionation and mass spectrometry analysis [5]. Ultimately the assignment of distance constraints within a single protein or protein complexes can be employed to provide information regarding the protein-protein interactions and three-dimensional structure of the protein or protein complex. This strategy is represented schematically in Figure 1.1 [6].

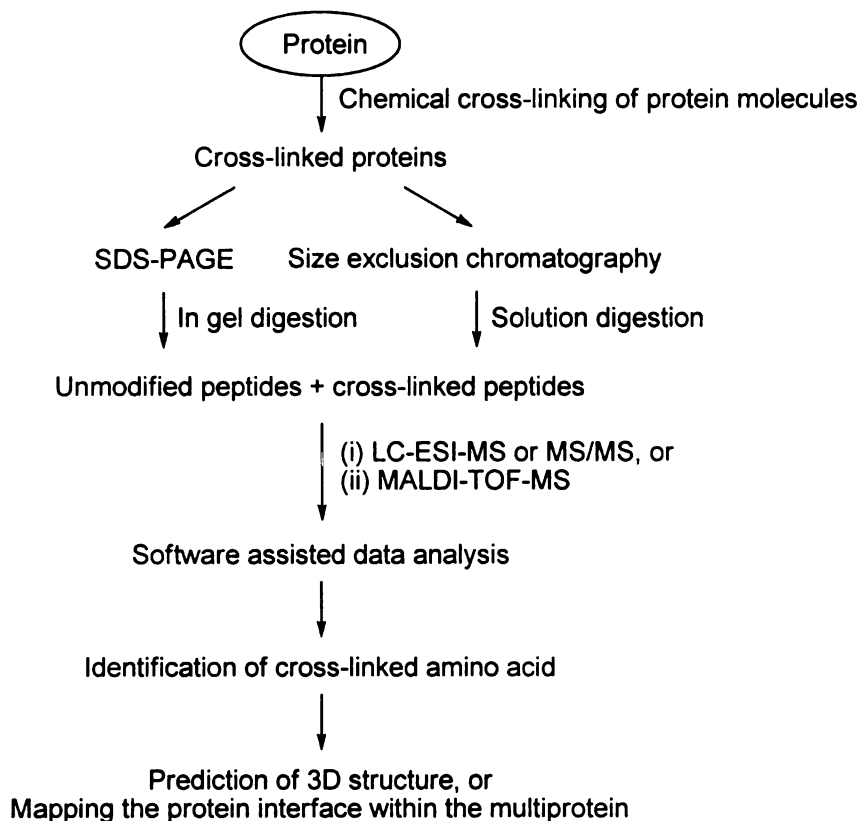


Figure 1.1 Cross-linking methodology flowchart by the bottom-up approach. Adapted from Reference 6

1.3.1 Cross-Linking Reagents

Cross-linking reagents contain two or more reactive ends to specific functional groups on proteins or other molecules. A large number of reagents have been described in the literature, and many are now commercially available. Major target functional groups for cross-linking reactions include primary amines, sulfhydryls, carbonyls, carbohydrates and carboxylic acids. These are summarized in Table 1.1 [7].

Table 1.1 Reactive cross-linker groups and their functional group targets. Adapted from Reference 7

Reactive Group	Target Functional Group	Reactive Group	Target Functional Group
Aryl Azide	Nonselective (or primary amine)	Maleimide	Sulfhydryl
Carbodiimide	Amine/Carboxyl	NHS-ester [*]	Amine
Hydrazide	Carbohydrate (oxidized)	PFP-ester ^{**}	Amine
Hydroxymethyl /Phosphine	Amine	Psoralen	Thymine(photoreac- tive intercalator)
Imidoester	Amine	Pyridyl Disulfide	Sulfhydryl
Isocyanate	Hydroxyl (non-aqueous)	Vinyl Sulfone	Sulfhydryl, amine, hydroxyl
Carbonyl	Hydrazine	Carbonyl	Hydrazine

^{*} N-hydroxysuccinimide ester

^{**} Pentafluorophenyl ester

Cross-linkers can be either homobifunctional, with two identical reactive groups at both reactive sites, or heterobifunctional, containing different reactive groups. Otherwise, relatively newly developed tri-functional cross-linking reagents incorporate a third functional group for reacting with a third target group, or for affinity purification of cross-linked products; a biotin moiety is commonly used for such purpose [8, 9]. Usually, cross-linking reagents have a spacer carbon chain; however zero-length cross-linkers may be employed to catalyze the formation of a chemical bond between two target groups without introducing an intervening linker. For example, carbodiimides are applied to mediate the direct formation of amide bonds between a carboxyl group and an amine group [10].

Besides the reactive functional groups of cross-linking reagents, the spacer arm length, water solubility or cell membrane permeability and additional modules for the separation and purification of cross-linked products are significant characteristics that should be taken into account for a specific application. The spacer arm length defines the distance relationship between two target groups on amino acid residues of a protein or protein complexes, which indicates the folding pattern of a protein or interacting sites between two subunits, i.e. protein-protein interactions. The solubility of cross-linking reagents in aqueous solution expands their applications to large proteins and protein complexes with poor solubility in organic solvents. The specific tags contained in cross-linking reagents that are used to facilitate the identification of cross-linked products will be discussed in the following section.

1.3.2 Possible Cross-linked Products from a Cross-Linking Reaction

Following the chemical cross-linking reaction, a variety of products may be produced. As shown in Figure 1.2 [11], after proteolysis even for a single cross-linking reaction, three distinct types of cross-linked peptides might be formed:

- (1) dead-end modified peptides (Type 0) with only one of the reactive groups of cross-linker reacted with the protein and the other one hydrolyzed. A type 0 modification doesn't provide any information related to the distance constraints between two amino acid residues but may indicate reactive groups exposed on the protein surface;
- (2) intramolecular cross-linked peptides (Type 1) with both reactive groups of cross-linker reacted with two amino acids on the same peptide chain. Type 1 cross-link can be applied to map low three-dimensional structure of proteins.
- (3) intermolecular cross-linked peptides (Type 2) with both reactive groups of cross-linker reacted with two amino acids on the two different peptide chains after proteolysis. If these two peptide chains come from two proteins within the complex, the important information about interacting sites will be obtained.

Combinations of type 0, 1, or 2 modifications give rise to a large number of possible cross-linked products, whereas the yield of multiple cross-linking reactions are usually low. The results from double modifications are also presented in Figure 1.2, where α chain can be either same as or different from β chain.

1.3.3 Challenges associated with the identification and characterization of chemically cross-linked peptides

As described above, despite the relative straightforwardness of cross-linking approaches, identification of cross-linked products by mass spectrometry is not straightforward due to the high complexity of the resultant peptide mixtures.

For the cross-linking reaction of protein or protein complexes, enzymatic digestion is usually performed prior to mass spectrometry analysis [bottom-up]. This process increases the complexity because cross-linked peptides are present along with a large number of unmodified peptides. The well known term, ‘looking for a needle in a haystack’, is often used to describe the challenges associated with the identification of cross-linked peptides resulting from this complexity.

To date, a number of techniques have been developed to facilitate the unambiguous assignment of cross-linked peptides, involving the enrichment of cross-linker-containing species by specific affinity tags [8, 9, 12-14] or via the introduction of discriminating properties such as differential isotope labels [15-22], specific ‘solution-phase’ [22-25] or ‘gas-phase’ [26, 27] cleavage sites, within the cross-linking reagent itself, or within the cross-linked peptides.

1.3.4 Affinity Cross-Linking Reagents

Affinity enrichment methods have typically incorporated a biotin functional group into the cross-linker, allowing their affinity purification by avidin affinity column or on avidin beads to reduce the complexity of the mixtures for subsequent mass spectrometry analysis.

1.3.5 Isotope Labeled Cross-Linking Reagents

Differential isotope labeling strategies enable low-abundance cross-linked peptides to be easily detected by their distinctive mass shift or isotopic patterns following mass spectrometry analysis. This can be realized by incorporation of the isotope label within the protein or peptide, or within the cross-linker contained in the protein or peptide.

1.3.5.1 Isotope Labeling on Polypeptide Chain

One approach involves the introduction of two ^{18}O atoms to the C-terminal carboxyl group during proteolytic digestion in ^{18}O enriched water [16]. Thus, the number of C-terminal carboxyl groups resulting from digestion can be calculated by their characteristic mass shifts. Intermolecular cross-linked peptides (Type 2, Figure 1.2) are readily distinguished by a characteristic 8 Da shift compared with peptides formed upon proteolysis in natural abundance water [14]. However, non-modified, dead-end and intramolecular cross-linked peptides are not able to be distinguished, as they will all exhibit the same 4 Da mass increment, similar to that for unmodified peptides.

Another approach described by Chen and coworkers [17] involves reductive dimethylation of primary amino groups within a protein, followed by enzymatic hydrolysis and derivatization of the newly formed N-termini with a 1:1 (w/w) mixture of 2,4-dinitrofluorobenzene- d_0/d_3 . Due to the incorporation of two dinitrophenyl groups, intermolecular cross-linked peptides will be distinguished by a characteristic 1:2:1 isotope pattern in the mass spectra from the 1:1 isotope pattern of other cross-linked types and non-modified peptides. For visualizing intermolecular cross-linked peptides, a mixed isotope cross-linking (MIX) strategy was designed [16] by mixing 1:1 uniformly ^{15}N -

1.3.5 Isotope Labeled Cross-Linking Reagents

Differential isotope labeling strategies enable low-abundance cross-linked peptides to be easily detected by their distinctive mass shift or isotopic patterns following mass spectrometry analysis. This can be realized by incorporation of the isotope label within the protein or peptide, or within the cross-linker contained in the protein or peptide.

1.3.5.1 Isotope Labeling on Polypeptide Chain

One approach involves the introduction of two ^{18}O atoms to the C-terminal carboxyl group during proteolytic digestion in ^{18}O enriched water [16]. Thus, the number of C-terminal carboxyl groups resulting from digestion can be calculated by their characteristic mass shifts. Intermolecular cross-linked peptides (Type 2, Figure 1.2) are readily distinguished by a characteristic 8 Da shift compared with peptides formed upon proteolysis in natural abundance water [14]. However, non-modified, dead-end and intramolecular cross-linked peptides are not able to be distinguished, as they will all exhibit the same 4 Da mass increment, similar to that for unmodified peptides.

Another approach described by Chen and coworkers [17] involves reductive dimethylation of primary amino groups within a protein, followed by enzymatic hydrolysis and derivatization of the newly formed N-termini with a 1:1 (w/w) mixture of 2,4-dinitrofluorobenzene- d_0/d_3 . Due to the incorporation of two dinitrophenyl groups, intermolecular cross-linked peptides will be distinguished by a characteristic 1:2:1 isotope pattern in the mass spectra from the 1:1 isotope pattern of other cross-linked types and non-modified peptides. For visualizing intermolecular cross-linked peptides, a mixed isotope cross-linking (MIX) strategy was designed [16] by mixing 1:1 uniformly ^{15}N -

labeled and unlabeled proteins to form a mixture of homodimers. Cross-linked peptides from an intermolecular origin exhibit triplet/quadruplet MS peaks, unique from doublet peaks of all other peptide species.

Identification of intermolecular cross-linked products by introducing isotope labels on the protein or peptide is feasible for the study of protein-protein interactions; however intramolecular cross-link can't be distinguished by this method.

1.3.5.2 Isotope Labeled Cross-Linking Reagents

Isotope-coding of cross-linking reagents allows straightforward mass spectrometric detection of peptides carrying the linker even in complex enzymatic mixtures. By reacting with 1:1 (w/w) mixtures of stable isotope-labeled and non-labeled cross-linking reagents, cross-linked peptides can be easily detected by their distinctive 1:1 isotope pattern [19, 20]. Isotope-labeled cross-linkers could be easily combined with other strategies for further identification of cross-link types [21, 22]. Seebacher *et al.* [21] have developed a new integrative method employing isotope-labeled cross-linking reagents in the presence of isotope-labeled water. In this way, deadend cross-linked peptides could be discriminated by an additional 2 Da splitting from hydrolysis of the cross-linking reagent's free reactive ester in [^{16}O] H_2O /[^{18}O] H_2O (1:1) buffer. Similarly, employing isotope labeled cross-linking reagents followed by ^{18}O labeling could further distinguish intermolecular cross-link from all the cross-linked products.

Isotope-labeling technologies have been widely employed to assist in finding cross-linked peptides in mixtures. However, in a spectrum with high complexity, the characteristic isotopic pattern of cross-linked peptides will often be masked because of

their small mass difference, or due to the low abundance of cross-linked products. Moreover, the use of only one isotope labeling process is not capable to identify all the cross-linked types at once.

1.3.6 Cleavable Cross-Linking Reagents

To improve the ability to identify and sequence cross-linked peptides, cleavable cross-linking reagents have been developed. The cleavage reaction can be performed in solution by hydrolysis [22], or by the use of reducing agents in the case of disulfide containing cross-linking reagents [23]. Alternatively, cleavage can be realized in the gas phase during an MS/MS experiments if labile bonds are contained within the cross-linker.

1.3.6.1 Chemically Cleavable Cross-Linking Reagents

The thiol-cleavable cross-linking reagent 3,3'-dithio-*bis*(succinimidylpropionate) (DTSSP) has been applied to a number of proteins and protein complexes [23-25]. After reduction of the cross-linker disulfide bond, intermolecular cross-linked peptides give rise to two separated components, each containing a reduced linker. Intra-molecular cross-links yield two reduced linker halves on the same peptide, whereas deadend type only have one present. Different cross-link types are distinguished by their characteristic mass shifts before and after reduction.

Another interesting example involves cleavable cross-linkers combined with isotope-coding strategies [22]. The isotope-labeled cleavable cross-linker is applied with its unlabeled counterpart to provide additional discriminating information because of the distinct isotope-pairs thus formed, allowing unambiguous identification of cross-linked

products. However, the application of cleavable cross-linkers by a chemical reaction relies on mass spectrometry analysis before and after the cleavage reaction takes place. Thus, the identification of cross-linked products might be complicated if the corresponding signals of reduced cross-link species are not observed, or if they overlap with non-cross-linked peptides.

The solution phase enrichment/cleavage methods and isotope labeling methods involve several steps of chemical reactions prior to and/or following the cross-linking process. Thus unintended by-reactions should be avoided in those reactions for unambiguous identification of cross-links.

To simplify the problem, more effort is being placed on the use of gas phase dissociation methodologies for realizing purification, discrimination and subsequent amino acid sequence analysis of cross-linked products.

1.3.6.2 Gas Phase Cleavable Cross-Linking Reagents (using MS/MS)

Low-energy collision induced dissociation (CID) tandem mass spectrometry (MS/MS) based methods have been developed for the cleavage and identification of cross-link containing peptides in the gas phase. The cleavage site can be incorporated on a side-chain of the cross-linking reagent, resulting in the formation of a stable marker ion upon MS/MS, while maintaining the cross-linked peptide linkages [26]. Alternatively, the cleavage site may be incorporated directly into the cross-linker spacer chain, resulting in cleavage of the cross-link upon MS/MS [27]. By forming a stable marker ion, cross-linked peptides can be rapidly identified by precursor ion scanning MS/MS methods, while the incorporation of a single gas-phase cleavable bond within the linker region

allows further the use of MSⁿ analysis for further amino acid sequence of each peptide following the initial cleavage reaction. Both strategies, though still needing improvement, are promising and could serve as basis for the second generation of cross-linking reagents [4].

Back et al. [26] have described the use of a bifunctional lysine reactive cross-linker, N-benzyliminodiacetylhydroxysuccinimid (BID), which yields a stable benzyl cation marker ion under MS/MS conditions, to successfully identify inter- and intra-molecular cross-linked peptides. However, formation of the benzyl cation at low m/z results in an inability to employ this strategy in quadrupole ion trap instruments, due to the low mass cut off inherent to this instrumentation. In addition, the benzyl cation is only observed as a dominant product for certain charge states of a protonated peptide.

Affinity enrichment techniques have been combined together with gas phase cleavable elements for the identification of cross-linked peptides, i.e. tri- and multi-functional group cross-linking reagents. For example, Bruce *et al.* have introduced a novel cross-linker strategy, termed protein interaction reporter (PIR), with the addition of an affinity tag, a hydrophilic group, a photocleavable group and low-energy MS/MS cleavable bonds [14, 28]. Upon CID MS/MS, PIR intermolecular cross-linked peptides fragment at two cleavage sites within the linker giving rise to two separated peptides each with an additional mass and a reporter ion. However, PIR is a large molecule with a spacer arm chain length of nearly 43 Å [27], making it less informative to determine the specific interaction points of protein-protein interactions.

To reduce the cross-linker length, Soderblom *et al.* [27] have developed a set of single site gas phase cleavable cross-linking reagents that can be selectively fragmented

in the source region of the mass spectrometer. The designed cleavage site at an aspartyl-prolyl bond is thought to be mediated by proton transfer from the aspartyl side chain to the basic amine of the adjacent prolyl residue to realize the preferential cleavage.

Regardless of whether the cleavage site is incorporated on the cross-linker side-chain or in the cross-linker spacer chain, the gas-phase cleavage is required to take place prior to cleavage of the peptide backbone. However, the mechanisms responsible for the gas-phase fragmentation reactions that give rise to the product ions of interest are typically highly dependent on the charge state and amino acid composition of the peptide (i.e., proton mobility), thus the product ions required for crosslink identification may only be observed in low abundance from total peptide ions or many dissociation channels [26-28].

To obtain further insights into this issue, the mechanisms those are likely to be responsible for the fragmentation of protonated cross-linked peptide ions by tandem mass spectrometry are briefly described below. Furthermore, a brief discussion of the principles of operation of the instrumentation in which these tandem mass spectrometry measurements are performed is also given.

1.4 Tandem Mass Spectrometry (MS/MS)

Tandem mass spectrometry (MS/MS) is a method involving selection of a precursor ion in a first stage of mass analysis, an intermediate reaction event (typically dissociation) followed by a second stage of mass analysis. These experiments can be realized in space, in instruments containing distinct mass analyzer regions, such as a triple quadrupole, or in time, by performing a sequence of analysis steps in an ion storage device such as an

ion trap. For in space MS/MS experiments, three main scan modes, product ion scan (daughter scan), precursor ion scan (parent scan), and neutral loss scan, are available [29]. However, mass analysis involving more than two stages can not be realized by this type of instrument. The identification of cross-linked peptides often requires multiple stage tandem mass spectrometry to be employed. The quadrupole ion trap could satisfy such purpose. To understand the principles how a quadrupole ion trap works, we can first consider the general operating principles of the quadrupole mass analyzer.

1.4.1 Quadrupole Mass Analyzer

The quadrupole mass analyzer separates ions based on achieving a stable trajectory in oscillating electric fields according to the ions' m/z ratio. Four parallel hyperbolic or cylindrical rods are employed with a variable amplitude oscillating AC (RF) voltage and variable DC voltage superimposed to pairs of opposite rods.

When ions are accelerated along the z -axis into the space between the quadrupole rods, they are also subjected to accelerations along the x and y axis resulting from forces induced by the electric fields, which are a function of the position of the ion from the center of the rods, i.e. x and y . If the values of x and y never reach the quadrupole field radius r_0 , the ion will have a stable trajectory. The motion of ions within a quadrupole mass analyzer is described by a second-order differential equation named after the 19th century French physicist E. Mathieu: the Mathieu equation. The solutions to the equation described in following formulas define the regions of stability and instability for ions as a function of the operating conditions in a quadrupole mass analyzer,

$$a_u = a_x = -a_y = \frac{8zeU}{m\omega^2 r_0^2}$$

$$q_u = q_x = -q_y = \frac{4zeV}{m\omega^2 r_0^2}$$

where U is the amplitude of the DC potential, V is the amplitude of the AC potential, $\omega = 2\pi\nu$ is the angular frequency (ν is the frequency of the AC potential), and t is time.

a and q values define whether the ions have stable trajectories when passing through the quadrupole electric field at given instrumental conditions. Therefore after rearranging these equations, the operating parameters for the quadrupole mass analyzer are derived.

$$U = a_u \frac{m\omega^2 r_0^2}{z8e}$$

$$V = q_u \frac{m\omega^2 r_0^2}{z4e}$$

As shown above, for a given m/z, x and y can be determined as a function of U and V during a time span. More straightforwardly, a combination of U and V values determines whether an ion at a specific m/z has a stable trajectory through the quadrupole, which is defined by solutions of a_u and q_u values. Thus a mass spectrum is

obtained by scanning a constant U/V ratio, allowing only one m/z ion at a time to be stable and pass through the quadrupole, whereas others are unstable and then lost.

1.4.2 The Quadrupole Ion Trap

The quadrupole ion trap is based on the same principle as the quadrupole mass analyzer, where an oscillating electric field is generated in three-dimensions resulting from the application of an RF potential to a hyperbolic ring electrode, while two hyperbolic end cap electrodes are held at ground.

The solution of the Mathieu equation can also be used to describe the ion motion in three dimensional ion traps. Since the DC potential applied to the end caps is usually zero, the solution for a_u is equal to zero whereas q is given by the following equation,

$$q_u = -2q_r = \frac{8zeV}{m(r_0^2 + 2z_0^2)\omega^2}$$

where r_0 and z_0 are the distances from the centre of the trap to the ring electrode and end cap electrodes, respectively, V is the amplitude of the RF potential, and ω is the angular velocity of the RF potential applied a fixed frequency.

To have a stable trajectory, ions must have motions such that their coordinates never reach or exceed r_0 and z_0 , which are defined by a_u and q_u values shown as stability diagram in Figure 1.6. According to the equation above, high m/z ions are represented by larger balls with low q values and low m/z ions by small balls with high q values under the same RF potential. Since there is no DC voltage applied, the ion trap is operated

along the $V(q)$ axis. Thus ions with different m/z values are present together inside the trap as long as their q values are less than the $q = 0.908$ at the boundary.

To acquire a mass spectrum in the quadrupole ion trap, the amplitude of the applied RF potential is increased while simultaneously applying a supplementary (auxiliary) radio frequency to the end cap electrodes at a frequency corresponding to a q value of 0.86. As ions of different m/z successively achieve a q value of 0.86, they are excited and become resonantly ejected. The maximum RF operating voltage and the q -eject value determines the highest m/z value that can be analyzed.

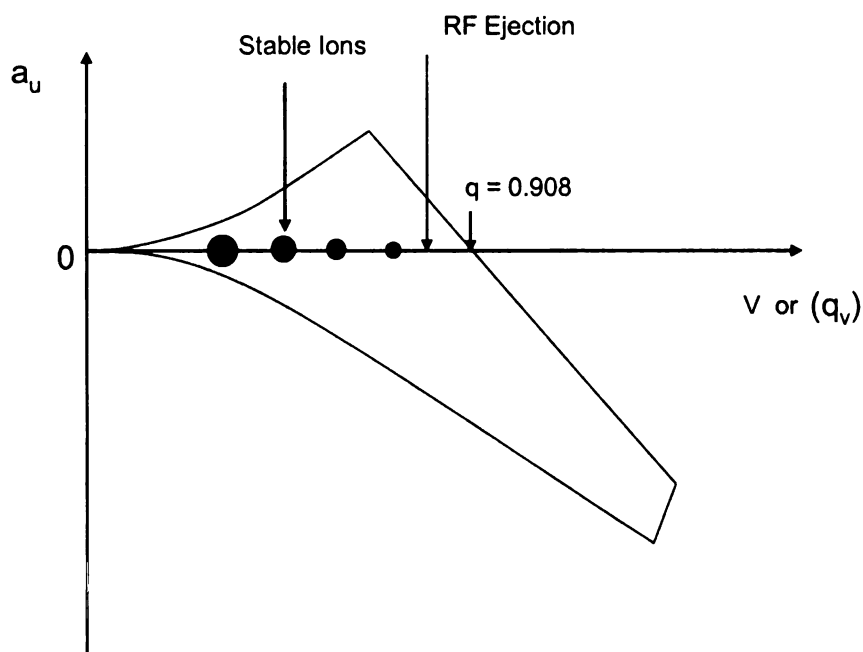


Figure 1.3 Typical stability diagram for quadrupole ion trap. The larger balls represent high m/z ions whereas smaller balls represent the low m/z ions. Reproduced and modified from Reference 29

1.4.3 Multiple Stage Tandem Mass Spectrometry (MS^n) Realized in Ion Trap

The main advantage of the ion trap is the ability to perform multistage tandem mass spectrometry (MSⁿ) experiments. This can be achieved by first selecting a precursor ion by expelling all other ions via the application of supplementary AC frequencies at their secular frequencies. Then the isolated ions are subjected to resonance excitation via the application of a supplementary AC frequency at the ions secular frequency of motion, resulting in energetic collisions with the inert He bath gas present in the trap. The product ion spectrum is then obtained by resonant ejection. This time-dependant tandem mass spectrometry can be repeated to provide MSⁿ spectra if one of the product ions is selected for further fragmentation.

1.5 Gas Phase Fragmentation Reactions of Protonated Cross-linked Peptide Ions

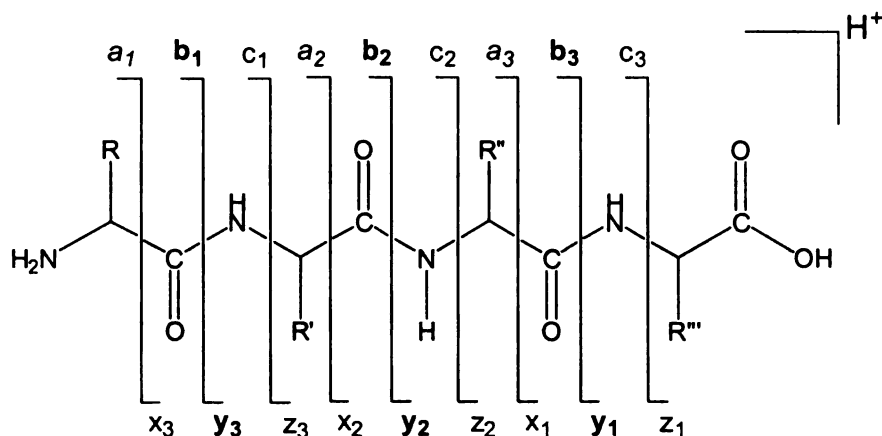
Proteins may be identified by ‘fingerprinting’ analysis of the masses obtained from their proteolytically derived peptides [30]. Most commonly however, protein identification is achieved by tandem mass spectrometry of selected peptide ions to obtain peptide sequence information, which is interpreted by either database analysis [31, 32] or *de novo* sequence analysis methods [33, 34]. The successful sequence analysis of protonated peptide ions relies heavily on the formation of a complete series of product ions resulted from the cleavage along the peptide backbone. However, other product ions resulting from cleavage at the amino acid side chains, multiple cleavage of the peptidic chain or involving rearrangements are also observed, complicating the mass spectrum and hindering peptide sequence analysis [35]. An understanding of mechanisms of protonated peptide fragmentation and how they control the types of fragment ions observed is a key

to increasing the accuracy of protein identification and then used for the determination of post-translational modifications [36, 37] or modifications involving peptide cross-linking.

1.5.1 Mechanisms for the Fragmentation of Protonated Peptide Ions

The nomenclature for peptide fragment ions formed from dissociations along the peptide backbone is shown in Scheme 1.1 [38]. Product ions corresponding to cleavage of the α C-C, the C-N amide bond and the N- α C bond, are termed a-, b-, and c-types ions if the charge is retained on the N-terminal or x-, y- and z- type ions if the C-terminal retains the charge. Collision induced dissociation (CID) is the most commonly used technique for energy transfer required for gas phase fragmentation reactions. Under the low energy CID conditions employed in triple quadrupole and quadrupole ion trap mass spectrometers, b- and y-type ions are typically the dominant product ions produced.

Protonation is typically required to occur at each of the amide bonds along the peptide backbone for a complete sequence information under low energy activation conditions. The mechanism for the cleavage process is generally described using the “mobile proton model” described by Dongre *et al.* [39] and extended by Kapp *et al.* [40], which defines 3 categories; “non mobile” when the total number of protons \leq the number of Arg residues, “mobile” when the total number of protons $>$ the total number of basic residues (i.e. combined number of Arg, Lys and His), or “partially mobile” when the number of Arg residues \leq the total number of protons \leq the total number of basic residues [40, 41].



Scheme 1.1 Nomenclature for peptide fragment ions. Adapted from Reference 38

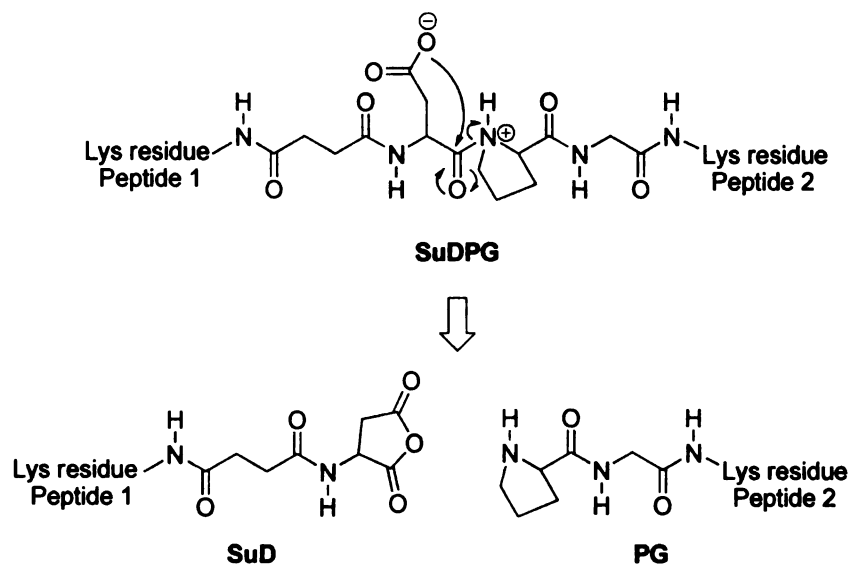
Under “non-mobile” proton conditions, higher energies are required to mobilize the proton to initiate “charge-remote” fragmentation pathways (i.e. no proton involvement) [40, 42]. Whereas under “mobile” or “partially-mobile” proton conditions, fragmentation of most protonated peptides requires the involvement of a proton at the cleavage site, i.e. the cleavages are “charge-directed” [43, 44].

During a “charge-directed” process, after the proton is transferred to the cleavage site, dissociation of the amide bond is initiated by nucleophilic attack from an adjacent nucleophile, e.g. amide carbonyl, resulting in an ion-molecule complex, followed by further dissociation without or with intermolecular proton transfer giving b- and/or y-type ions.

As mentioned above, BID cross-linked neurotensin was observed to yield extensive benzyl cations only in its 5+ charge state [26], indicating that a “mobile-proton” condition was required for CID MS/MS experiments. However, even when a mobile proton was present, other processes, including cleavages along the peptide

backbone, as well as cleavage of the amide bond in the cross-linker chain were also observed.

The cross-linking reagent (SuDPG) developed by Soderblom *et al.* [27] is based on the incorporation of an aspartyl-prolyl bond within the linker region for specific gas-cleavage site. As shown in Scheme 1.2, cleavage is initiated by transferring a labile proton from the aspartyl side chain to the basic backbone amine of the adjacent prolyl residue, and then the carboxy anion attacks the carbonyl carbon, resulting in a cyclized anhydride and a free prolyl residue. However the Kapp *et al.* study indicated that fragmentation of Asp-Pro bonds is enhanced only for peptides under non-mobile conditions [40]. As a result, realizing preference cleavage in both BID and SuDPG cross-linked peptides is limited by their charge states.



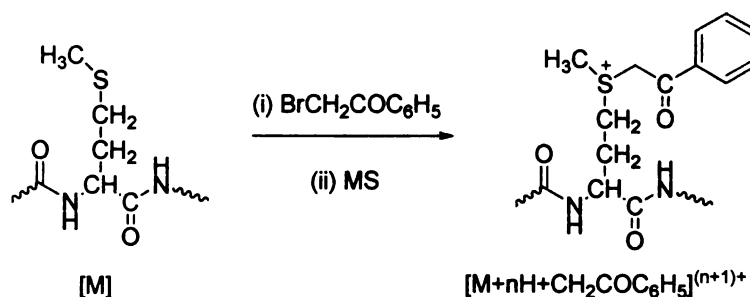
Scheme 1.2 Fragmentation mechanism at aspartyl-prolyl bond within SuDPG cross-linked peptides. Adapter from Reference 27

1.5.2 · Controlling Peptide Fragmentation Reactions via the use of Fixed Charge Derivatization

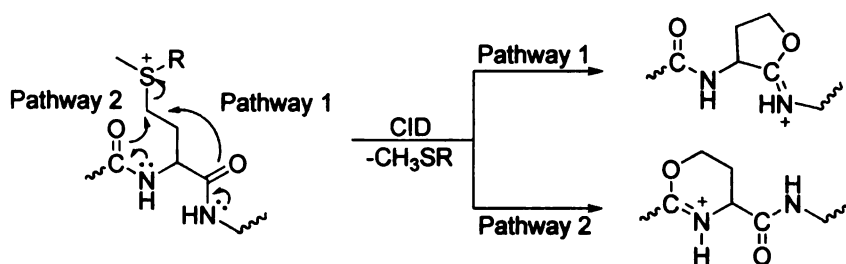
The mechanistic limitations associated with the fragmentation of protonated cross-linked peptide ions has led to consideration of alternate strategies for controlling or directing the fragmentation of cross-linked peptides toward the selective formation of analytically useful product ions.

A number of different methods for controlling the fragmentation of non cross-linked peptide ions, involving the introduction of a fixed charge to the peptide, have recently been developed. Recently, the Reid group has described a novel chemical derivatization strategy for selective MS/MS based peptide identification, characterization and quantitative analysis, involving ‘fixed charge’ sulfonium ion derivatization of specific functional groups (e.g., the side chains of methionine (Scheme 1.3) or cysteine containing amino acids) within a peptide of interest [45-48]. From an earlier study on the gas-phase fragmentation behavior of phenacyl sulfonium ion derivatized methionine containing peptides, the exclusive fragmentation pathway observed corresponded to selective cleavage adjacent to the site of the fixed charge within the peptide ion resulting in the neutral loss of phenacyl methyl sulfide. Furthermore, this cleavage was found to occur independently of the precursor ion charge state or amino acid composition (i.e., proton mobility) [46]. Based on molecular orbital calculations on a simple peptide model, together with experimental studies on multistage dissociation of methionine derivatized sulfonium ion containing peptide, a further investigation in the mechanism responsible for this selective fixed charge side-chain fragmentation indicates that S_N2 reactions occur, involving the N- and C- terminal amide bonds adjacent to the methionine side chain,

resulting in the formation of stable cyclic five- and six-membered ring product ions respectively (Scheme 1.4) [46]. These stable structured product ions were then subjected to further dissociation by MS³ to obtain additional structural information. Extensive studies reveal that though selective fragmentation at the fixed charge containing side chain is independent of amino acid composition and charge state of the peptide ion, loss of the side-chain fragment either as a neutral or charged species was observed highly dependent on the proton mobility of the precursor ion and the identity of sulfonium substitute [48].



Scheme 1.3 Fixed charge derivatization of methionine sulfonium ion derivatized peptide ions. Adapted and modified from Reference 47



Scheme 1.4 Gas-phase fragmentation behavior of the sulfonium ion derivative of methionine-containing peptides. Adapted and modified from Reference 47

The studies above provide a means to form characteristic product ions in high relative abundance, thereby enabling side chain fixed charge derivative containing peptides to be selectively identified from within complex mixtures, resulting in mixture simplification and improved dynamic range for qualitative and/or quantitative proteome analysis

Based on the studies of gas-phase fragmentation reactions of methionine side chain fixed charge sulfonium ion containing peptides, the incorporation of 'fixed charge' sulfonium ion derivatives into chemical cross-linking reagents would enable this strategy to be extended toward the development of improved MS/MS based approaches for the identification of specific interaction sites in proteins and multi-protein complexes. In the first stage of MS/MS, ionic cross-linked peptides could be identified by either a product ion scan or a neutral loss scan from their characteristic fragmentation pattern on the cross-linker, then following MS³, cleavages along the peptide backbone will take place giving rise to a series of product ions containing information about the peptide sequence and the modification sites. Thus the cross-linking information could be used to derive a set of distance constraints for structure modelling.

1.6 Aims of this Thesis

To develop improved chemical cross-linking and tandem mass spectrometry methodologies for the analysis of protein structures and protein-protein interactions, the aims of this thesis are:

1. Synthesis of a series of novel water soluble 'fixed charge' sulfonium ion containing cross-linking reagents.

2. Evaluation of the multistage gas-phase fragmentation reactions of cross-linked peptide ions formed by reaction with these reagents, using synthetic peptides and standard proteins and protein complexes.

CHAPTER TWO

EXPERIMENTAL

2.1 Materials

All chemicals were analytical reagent (AR), or of a comparable or higher grade and used without further purification. Thioglycolic acid, 3-Mercaptopropionic acid, N,N'-Dicyclohexylcarbodiimide and Phenacyl bromide were purchased from Fluka (Switzerland). 5-Bromovaleric acid and Thiourea were from Sigma-Aldrich (St. Louis, MO, USA). N-hydroxysuccinimide was purchased from Pierce (Rockford, IL, USA). Sodium hydroxide, Sodium phosphate dibasic (crystal), Potassium phosphate monobasic (crystal), Dimethyl sulfate and N,N'-Dimethyl formamide (DMF) were purchased from Spectrum Chemicals (Gardena, CA, USA). Sodium chloride, Sodium sulfate anhydrous, Hydrochloric acid and Ethyl acetate were purchased from Columbus Chemical Industries (Columbus, WI, USA). Potassium chloride, Sulfuric acid and Ethyl ether were purchased from Jade Scientific (Canton, MI, USA). Glacial acetic acid, Dichloromethane, Chloroform, Methanol (anhydrous) and Ethanol were purchased from Mallinckrodt Chemicals (Phillipsburg, NJ, USA). Iodomethane was purchased from EMD Chemicals (San Diego, CA, USA), and Acetonitrile was from Fisher Scientific (Hampton, NH, USA).

The peptide VTMAHFWNFGK was obtained from Auspep (Parkville, Australia), Neurotensin was purchased from Sigma-Aldrich (St. Louis, MO, USA), and Insulin-like growth factor I (57-70) was purchased from American Peptide Company (Sunnyvale, CA).

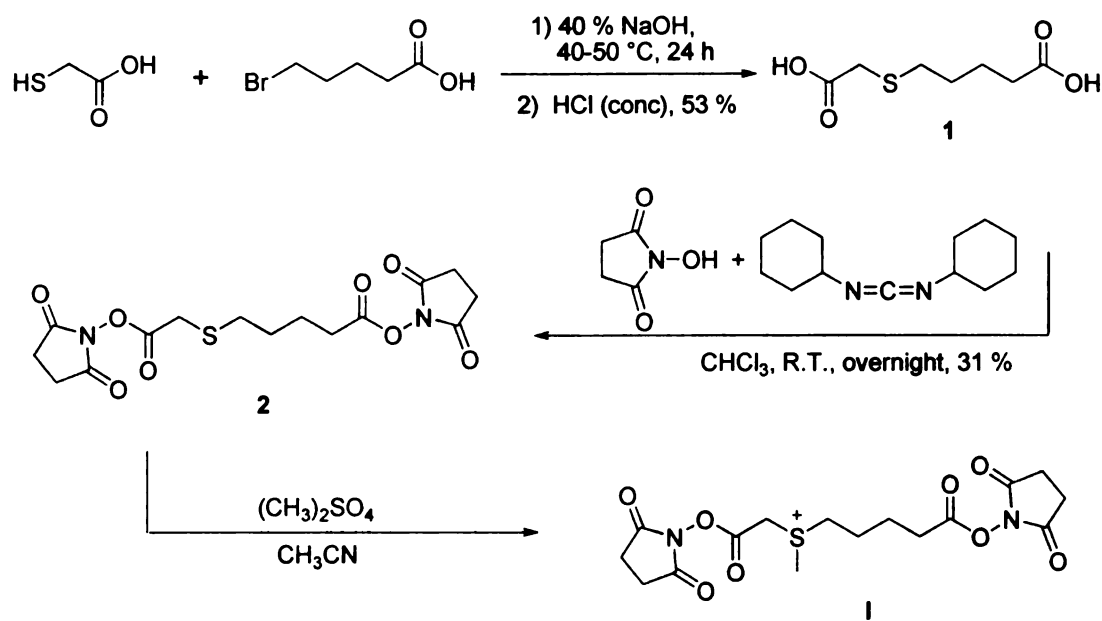
Water was deionized and purified by a Barnstead nanopure diamond purification system (Dubuque, Iowa, USA). DMF was dried over 3 Å molecular sieves (Spectrum Chemicals) and filtered. All reactions were performed in oven dried glassware.

¹H-NMR spectra were obtained on Varian Inova 300 MHz or 500 MHz instruments and are reported in parts per million (ppm) relative to the solvents resonances (δ), with coupling constants (*J*) in Hertz (Hz).

2.2 Mass Spectrometry

A Thermo model LCQ Deca quadrupole ion trap mass spectrometer (San Jose, CA), equipped with nanoelectrospray ionization was used for all analysis reported. Samples were introduced into the mass spectrometer at a flow rate of 0.5 μL / min. The capillary voltage was set to 2.5 KV and capillary temperature was in the range of 150 to 200 °C, other parameters were obtained as a result of auto tune function optimized for each individual compound. Low energy CID MS/MS and MSⁿ experiments were performed using helium as a collision gas, at an activation *q* value of 0.25 and an activation time of 30 ms. Collision energies were individually optimized for each compound of interest.

2.3 Synthesis of Ionic Cross-Linking Reagent S-Methyl 2-acetoxy,5'-pentanoyldihydroxysuccinimide sulfonium (Compound I in Scheme 2.1)



Scheme 2.1 The synthesis route of ionic cross-linking reagent S-methyl 2-acetoxy-5'-pentanoyldihydroxysuccinimide sulfonium (**I**)

2.3.1 Carboxymethylmercaptopropionic Acid (1)

Following the method by Rabinovich, et al. [49], a freshly prepared solution of 5-bromovaleric acid (4.0 g, 22.1 mmol) in 40 % NaOH (10 ml) was added dropwise to an ice bath cold solution of thioglycolic acid (2.1 g, 22.8 mmol) in 40 % NaOH (10 ml). The resulting reaction mixture was stirred at 40-50 °C for 24 hours. After the heating was terminated, the product mixture was acidified with concentrated hydrochloric acid to pH=1 and repeatedly extracted with dichloromethane (50 ml \times 5). Extracts were combined, dried over anhydrous Na_2SO_4 , filtered and evaporated under reduced pressure to give a compound (**1**) as a white solid in 53 % (2.3 g) yield. ^1H NMR (500 MHz, CDCl_3): δ 1.64-1.77 (m, 4H), 2.38 (t, 2H, $J = 7$ Hz), 2.67 (t, 2H, $J = 7$ Hz), 3.23 (s, 2H), 10.18 (s, broad, 2H); ^{13}C NMR (125 MHz, CDCl_3): δ 23.48, 28.04, 32.14, 33.35, 33.38, 176.61 and 179.73.

The product was dissolved in methanol:H₂O 50:50 v/v to a concentration of 20 μ M and characterized by negative ion mode nanoelectrospray mass spectrometry and MS/MS as described above. The deprotonated pseudo-molecular ion $[M-H]^-$ at m/z 191.2 (Figure 2.1A) was selected for fragmentation by CID-MS/MS, the fragments obtained were consistent with the structure, and the fragmentation pathways are demonstrated in Scheme 2.2.

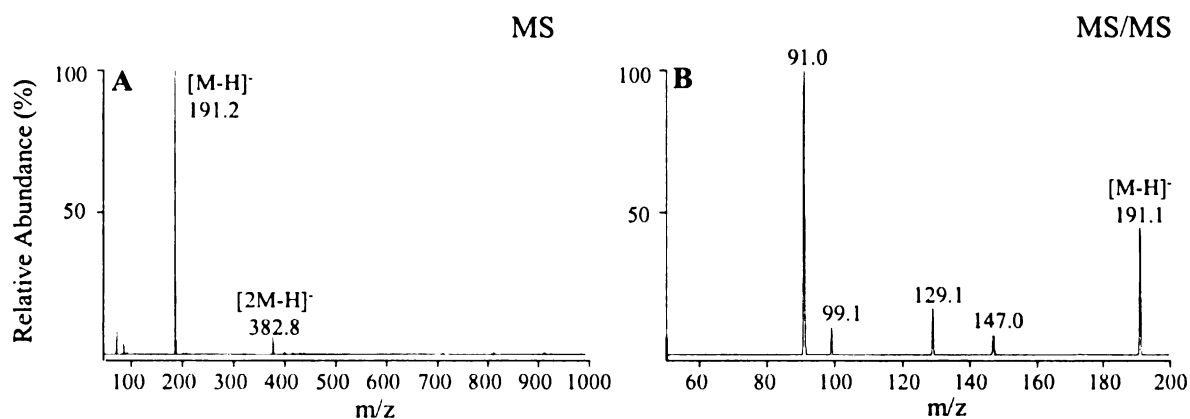
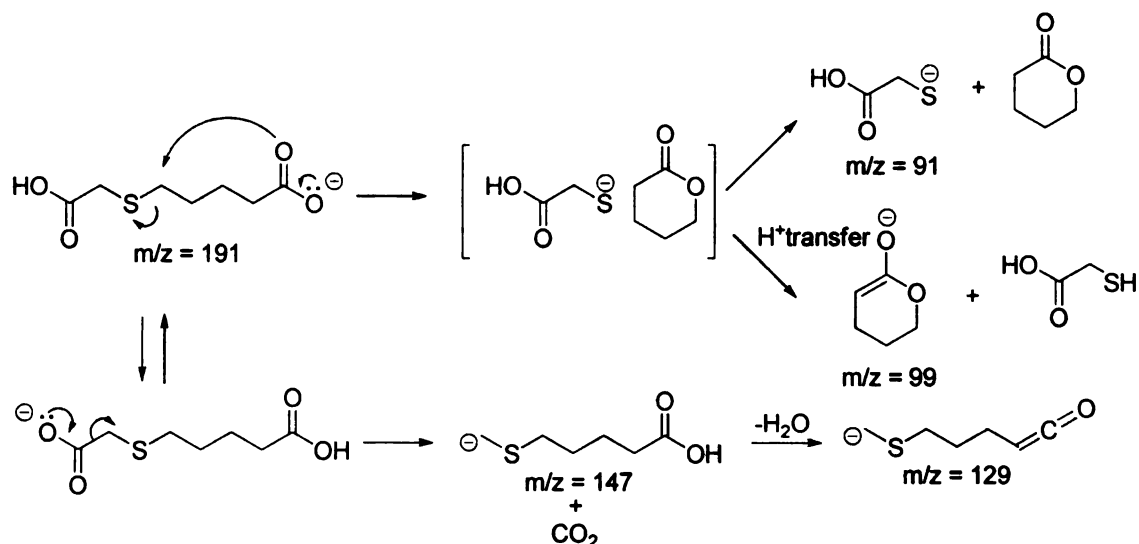


Figure 2.1 Characterization of (1) by nanoelectrospray quadrupole ion trap mass spectrometry and CID MS/MS. (A) MS and (B) CID-MS/MS of m/z 191.1



Scheme 2.2 Fragmentation scheme of (1) anions. The spectrum is displayed on Figure 2.1B. Square brackets indicate ion-molecule complexes

2.3.2 Thio 2-acetyl-5'-pentanoyldihydroxysuccinimide (2)

Following methods analogous to other N-hydroxysuccinimide esters [50-52], compound **1** (0.8 g, 4.2 mmol) and N-hydroxysuccinimide (NHS) (1.2 g, 10.4 mmol) were dissolved in 3:1 v/v mixture of chloroform and dichloromethane (DCM) (40 ml) and stirred for 5 min at room temperature. N,N'-dicyclohexylcarbodiimide (DCC) (2.2 g, 10.6 mmol) was then added and the mixture stirred overnight. After filtration of dicyclohexylurea (DCU) precipitate and solvent removal under reduced pressure, the oily residue was dissolved in a minimum amount of ethyl acetate. The remaining DCU was precipitated and removed by filtration. Following rotary evaporation of the ethyl acetate, the residue was dissolved in DCM, washed with 1M NaOH and H₂O then evaporated to near dryness. The residue was then recrystallized with ethyl ether containing a trace amount of acetone to give 0.5 g (31 %) of compound (**2**). ¹H NMR (500 MHz, CDCl₃): δ 1.71-1.77 (m, 2H), 1.81-1.87 (m, 2H), 2.62 (t, 2H, *J* = 7 Hz), 2.74 (t, 2H, *J* = 7 Hz), 2.80

(s, broad, 8H), 3.45 (s, 2H); ^{13}C NMR (125 MHz, CDCl_3): δ 23.40, 25.53, 25.54, 27.41, 29.98, 30.24, 31.53, 165.82, 168.27, 168.86 and 169.17.

For mass spectrometry, compound **2** was dissolved in acetonitrile to a concentration of 10 μM . In the positive ion mode, the abundance of the pseudo molecular $[\text{M}+\text{H}]^+$ ion was much lower than the Na^+ and K^+ adducts due to poor protonation of the diester. The dimer adducts are also observed. The peak of m/z at 272.2 in Figure 2.2A corresponds to a fragment generated in the ionization source from precursor ions of ammonium adduct $[\text{M}+\text{NH}_4]^+$, which is also a product ion present in the MS/MS and MS^3 spectra in Figure 2.2C and 2.2D. The dominant fragments of $[\text{M}+\text{Na}]^+$ and $[\text{M}+\text{H}]^+$ ions were from a neutral loss of 115; the fragmentation pathway is displayed in Scheme 2.3 and 2.4.

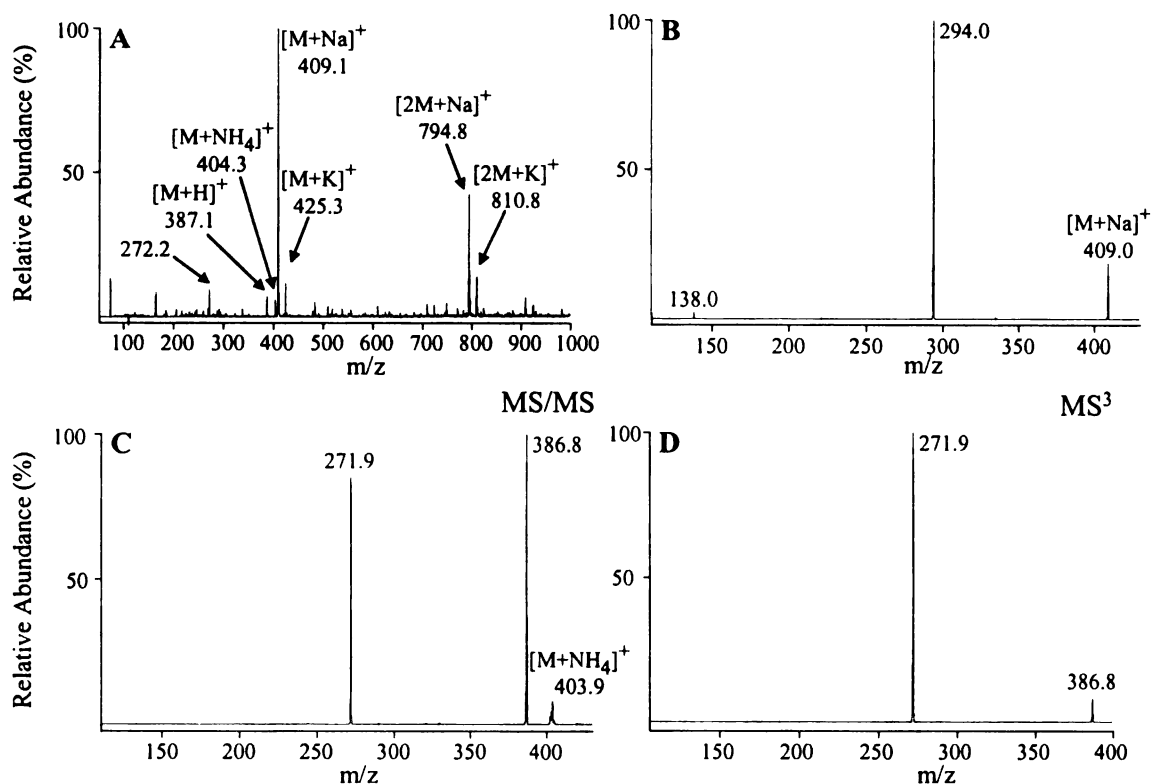
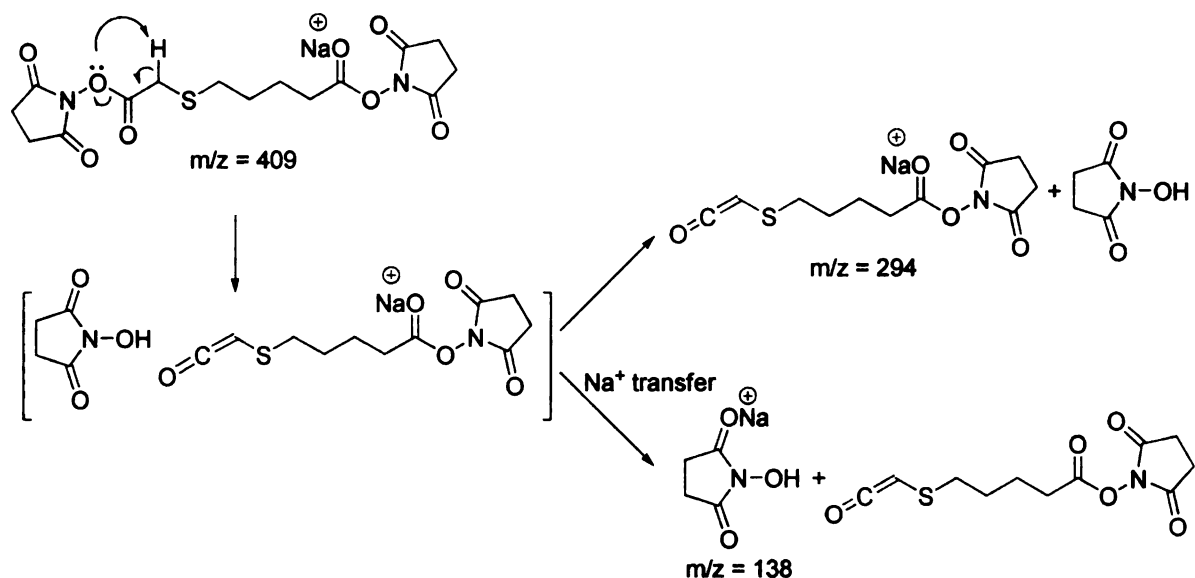
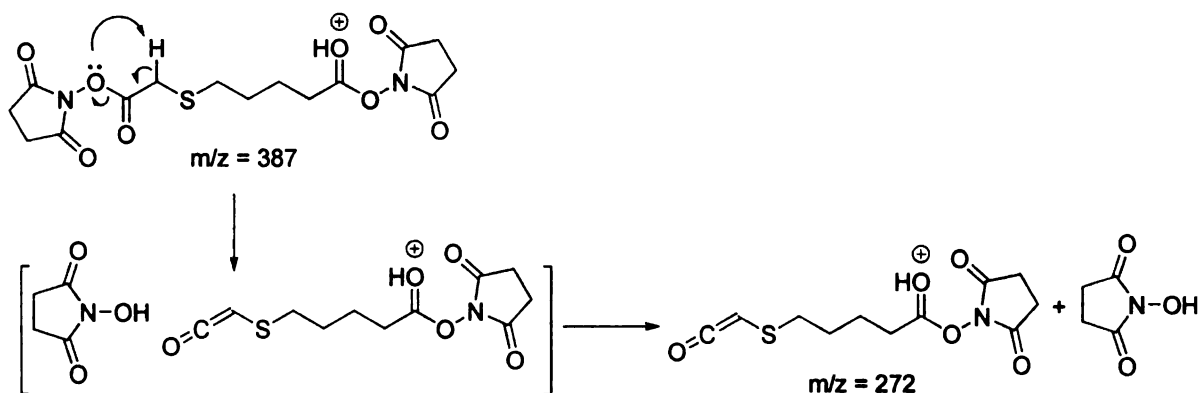


Figure 2.2 Characterization of (2) by nanoelectrospray quadrupole ion trap mass spectrometry and MS/MS. (A) Mass spectrum of (2), due to the poor protonation of diester, the $[M+Na]^+$ adduct is the most intensive peak; (B) CID MS/MS of $[M+Na]^+$ adduct, neutral loss of 115 was observed; (C) CID MS/MS of $[M+NH_4]^+$ adduct, yielding $[M+H]^+$ and m/z at 272^+ fragment ions; (D) CID MS³ of $[M+H]^+$ coming from $[M+NH_4]^+$, giving rise to the fragment at m/z of 272^+ , by neutral loss of 115



Scheme 2.3 Fragmentation of Na^+ adduct of (2). The spectrum is displayed on Figure 2.2B. Square brackets indicate ion-molecule complexes



Scheme 2.4 Fragmentation of protonated (2). The spectrum is displayed on Figure 2.2D. Square brackets indicate ion-molecule complexes. The protonated (2) is supposed to have the similar fragmentation pathway as Na^+ adduct of (2) for the same neutral loss of 115 observed but no proton transfer observed to give product ion at m/z of 116 due to NHS has lower affinity to proton than that of Na^+

2.3.2 S-Methyl 2-acetyl-5'-pentanoyldihydroxysuccinimide Sulfonium (I)

A mixture of compound **2** (193 mg, 0.5 mmol) and dimethyl sulfate (2.52 g, 20 mmol) in 2 ml acetonitrile was allowed to react at room temperature for 4 days. Following freeze drying, a dark brown oily residue was obtained. Figure 2.3 shows the MS and MS/MS spectra obtained by analysis of the sulfonium (**I**) methylsulfate. The fragmentation pathways are displayed in Scheme 2.5, where two competitive routes are observed giving rise to product ions at m/z of 198^+ and 286^+ respectively.

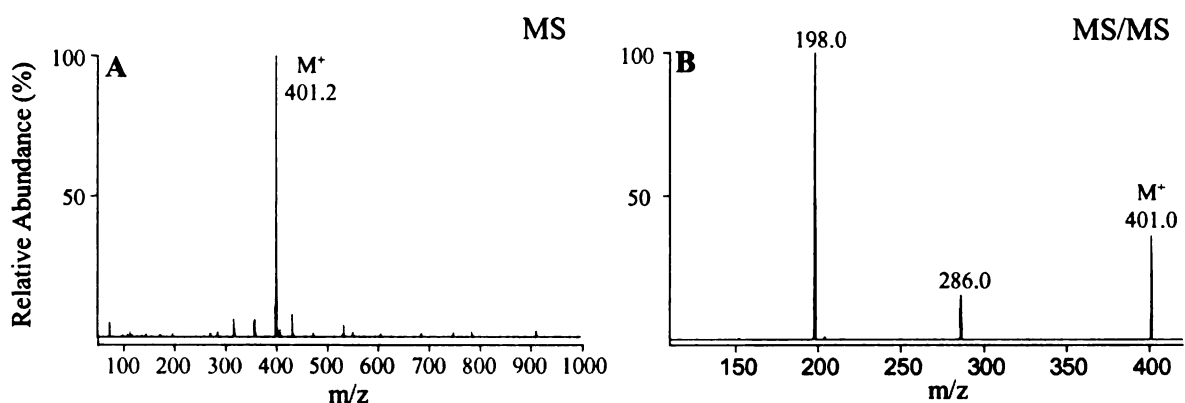
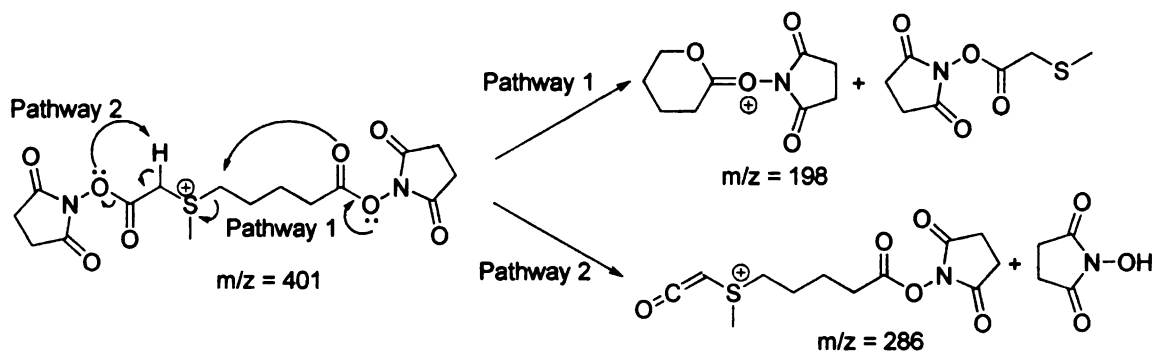
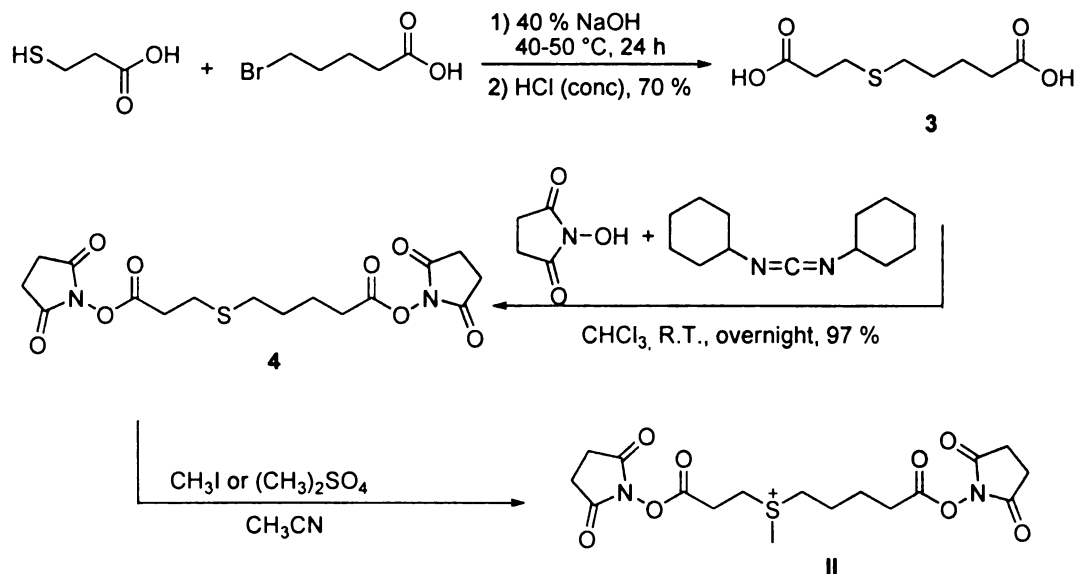


Figure 2.3 Characterization of ionic cross-linking reagent **I** by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The methyl sulfonium ion is at m/z 401⁺; (B) CID-MS/MS product ion spectrum of M^+



Scheme 2.5 Fragmentation of sulfonium ion (**I**). The spectrum is displayed on Figure 2.3B. Fragment of m/z at 198^+ is the main process; otherwise the pathway for the neutral loss of 115 is also present

2.4 Synthesis of Ionic Cross-Linking Reagent S-Methyl 3-propionoyl,5-pentanoyldihydroxysuccinimide sulfonium (Compound II in Scheme 2.6)



Scheme 2.6 Synthesis route of ionic cross-linking reagent S-Methyl 3-propionoyl-5'-pentanoyldihydroxysuccinimide sulfonium (II)

2.4.1 1-Carboxyethylmercaptopropionic Acid (3)

This acid was synthesized following the method described for the 1-carboxymethylmercaptopropionic acid (1), from a solution 5-bromovaleric acid (2.0 g, 11.0 mmol) in 40 % NaOH (5 ml) and solution of 3-mercaptopropionic acid (1.2 g, 11.3 mmol) in 40 % NaOH (5 ml) in 70 % (1.6 g) yield. ^1H NMR (500 MHz, CDCl_3): δ 1.60-1.67 (m, 2H), 1.69-1.75 (m, 2H), 2.37 (t, 2H, $J = 7$ Hz), 2.54 (t, 2H, $J = 7$ Hz), 2.64 (t, 2H, $J = 7$ Hz), 2.76 (t, 2H, $J = 7$ Hz), 11.45 (s, broad, 2H); ^{13}C NMR (125 MHz, CDCl_3): δ 23.62, 26.45, 28.61, 31.62, 33.46, 34.68, 178.25 and 179.84.

The characterization by mass spectrometry and MS/MS is shown in Figure 2.4; the fragments obtained were consistent with the referring structure (Scheme 2.7).

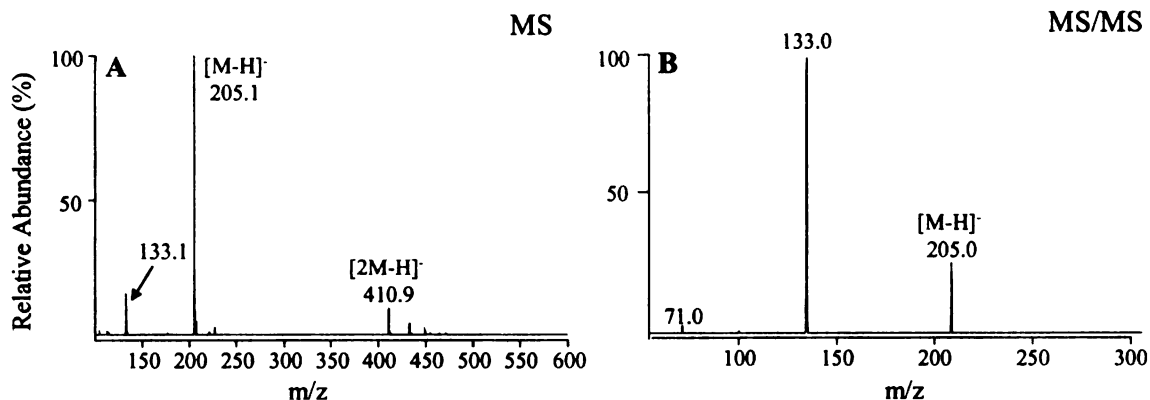
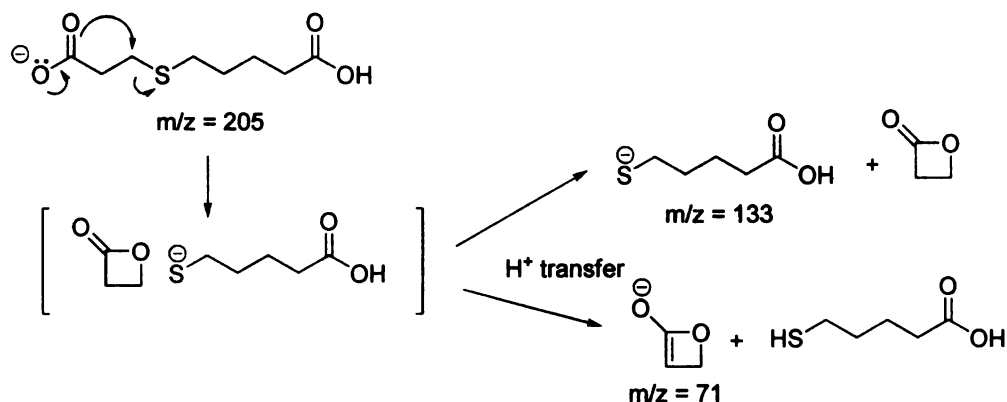


Figure 2.4 Characterization of (3) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The pseudo molecular ion $[M-H]^-$ at m/z 205.1⁻ and the dimer at m/z 410.9⁻; (B) CID-MS/MS of selected parent ion at m/z 205.1⁻



Scheme 2.7 Fragmentation scheme of (3) anions. The spectrum is displayed on Figure 2.4B. Square brackets indicate ion-molecule complexes

2.4.2 Thio 3-propionoyl-5'-pentanoyldihydroxysuccinimide (4)

This diNHS-ester was synthesized and purified in the same manner as thio 2-acetoyl-5'-pentanoyldihydroxysuccinimide (2). (3) (1 g, 4.8 mmol), N-hydroxysuccinimide (1.4 g, 12.2 mmol) and DCC (2.5 g, 12.1 mmol) were dissolved in chloroform (30 ml) and stirred overnight. After purification a light yellow solid (4) was

obtained in 97 % (1.9 g) yield. ^1H NMR (500 MHz, CDCl_3): δ 1.67-1.73 (m, 2H), 1.80-1.86 (m, 2H), 2.55-2.62 (m, 4H), 2.80-2.90 (m, 12H); ^{13}C NMR (125 MHz, CDCl_3): δ 23.49, 25.50, 26.09, 28.17, 30.37, 31.35, 31.96, 167.13, 168.25, 168.94 and 169.12.

Following analysis by ESI-MS, the singly protonated ion was observed at m/z 400.9, however due to poor protonation, the $[\text{M}+\text{NH}_4]^+$ adduct at m/z 418.2 was observed as the most abundant ion. The ion at m/z 286.0 in Figure 2.5A was found to correspond to a source CID fragment of the $[\text{M}+\text{NH}_4]^+$ adduct, by analysis of its CID-MS/MS product ion spectrum in Figure 2.5B. The MS^3 spectrum (data not shown) of the $[\text{M}+\text{H}]^+$ ion (m/z 400.8) formed by fragmentation of the $[\text{M}+\text{NH}_4]^+$ precursor gave rise to the product at m/z 286 as well the neutral loss of 115 Da, similar to that seen for (2) in Scheme 2.4.

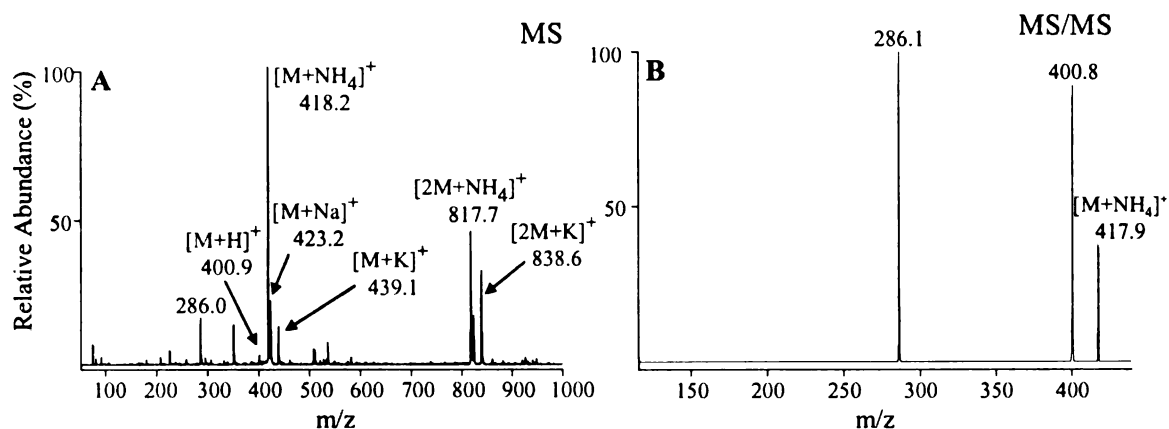


Figure 2.5 Characterization of (4) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) Mass spectrum of (2), $[\text{M}+\text{NH}_4]^+$ adduct is the most abundant peak due to poor protonation of the diester; (B) CID MS/MS of $[\text{M}+\text{NH}_4]^+$ adduct, neutral loss of NH_3 and further loss of 115 was observed

2.4.3 S-Methyl 3-propionoyl-5'-pentanoyldihydroxysuccinimide Sulfonium (II)

The methylation of (**4**) can be performed by reacting with dimethyl sulfate as methods mentioned above, from a starting substrate of 200 mg (0.5 mmol). Alternatively, methylation can be achieved by stirring of (**4**) with iodomethane (2.1 g, 15 mmol) in 2 ml acetonitrile at room temperature in the dark for 24h. A light yellow solid was obtained after freeze drying. The two procedures gave the methyl sulfonium ion with different counter ions. Both ions exhibited identical fragmentation behavior following low energy CID MS/MS. Figure 2.6 demonstrates positive mode mass spectrometry of molecular ion at m/z 415 and MS/MS spectra obtained by analysis of the sulfonium (**II**) iodide. A little different from the fragmentation process of (**I**) as shown in Scheme 2.5, pathway 1 is still the predominant mechanism for (**II**) resulting in 6-membered ring oxonium at m/z of 198 and complementary ion at m/z of 218, however the neutral loss of 115 was too weak to be observed.

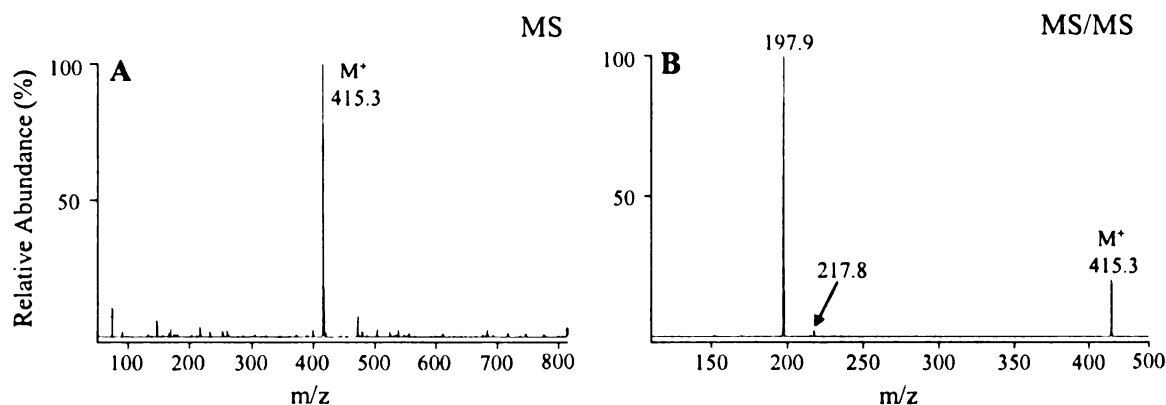


Figure 2.6 Characterization of ionic cross-linking reagent **II** by nanoelectrospray quadrupole ion trap mass spectrometry. (A) Mass spectrum of methyl sulfonium ion at m/z 415⁺; (B) CID-MS/MS product ion spectrum of M^+ ion

Phenacyl bromide is another commonly used alkylating reagent yielding a phenacyl substituted sulfonium ion. A mixture of (**4**) (200 mg, 0.5 mmol) and phenacyl bromide

(1.99 g, 10 mmol) in acetonitrile (5 ml) was allowed to react at room temperature in the dark for 5 days. After freeze drying, the residue was characterized by mass spectrometry and MS/MS without separation of excess reagent. The obtained sulfonium ion has different m/z value from that of methylation products, but still gave the same main fragment at m/z 198 (complementary ion at m/z of 322) as shown in Figure 2.7B, product ions at m/z of 161 and 289 can be derived only from a rearrangement due to the present of the phenacyl group.

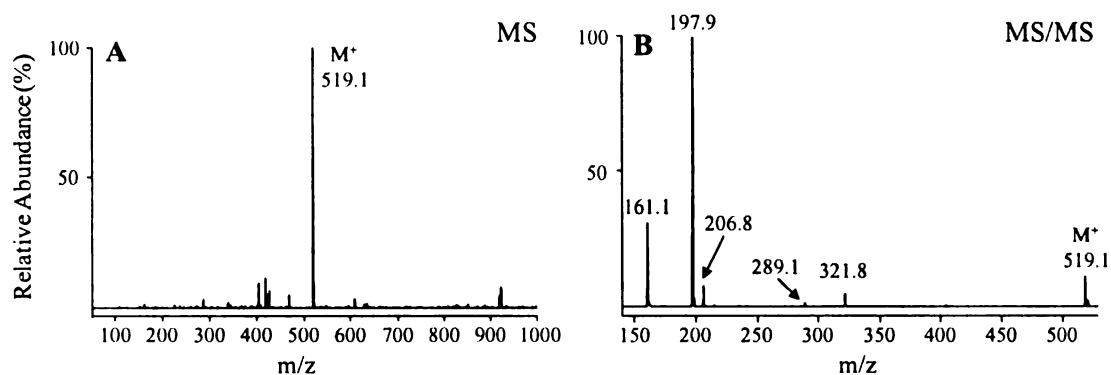
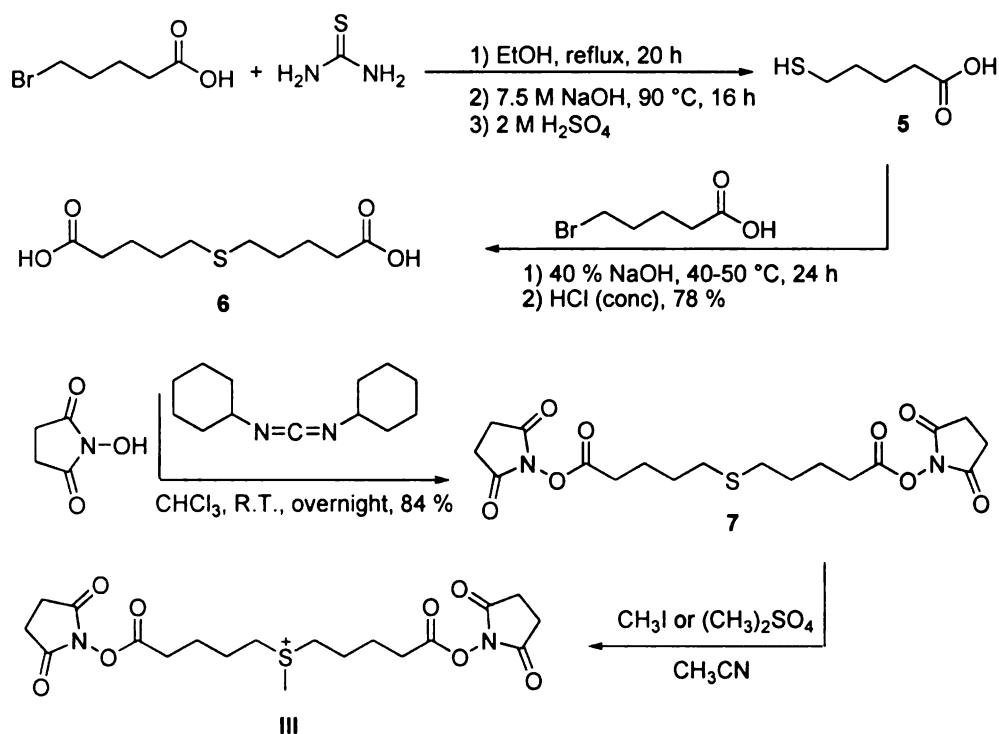


Figure 2.7 Characterization of phenacylsulfonium ion derivative of (II) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) Mass spectrum of phenacylsulfonium ion as an alternative for compound II, showing M^+ at m/z 519⁺; (B) CID-MS/MS yields the same product ion at m/z 198

2.5 Synthesis of Ionic Cross-Linking Reagent S-Methyl 5,5'-dipentanoylhydroxysuccinimide sulfonium (Compound III in Scheme 2.8)



Scheme 2.8 Synthesis of ionic cross-linking reagent S-Methyl 5,5'-dipentanoylhydroxysuccinimide sulfonium (**III**)

2.5.1 5-Mercaptopentanoic Acid (**5**)

The procedure was adapted from that of Jessing [53] for 5-bromovaleric acid. 5-bromovaleric acid (2.2 g, 12.1 mmol) and thiourea (1.4 g, 18.4 mmol) were dissolved in ethanol (25 ml) and refluxed for 20 h. The solvent was removed under reduced pressure and 7.5 M NaOH (aq) (25 ml, 187 mmol) was added. The mixture was stirred for an additional 16 h at 90 °C. Then with cooling on an ice bath, 2M H₂SO₄ (aq) was added slowly under stirring until pH = 1. The product was extracted with CH₂Cl₂ (2×100 mL), combined extracts were dried with MgSO₄ and concentrated by rotary evaporation to give the title acid (**5**) as a colorless oil in quantitative yield. The product was then used without further purification. ¹H NMR (300 MHz, CDCl₃): δ 1.32(t, 1H, *J* = 7.8 Hz), 1.59-

1.73 (m, 4H), 2.32 (t, 2H, $J = 7.5$ Hz), 2.49 (q, 2H, $J = 6.9$ Hz), 8.95 (s, broad, 1H); ^{13}C NMR (75 MHz, CDCl_3): δ 23.21, 24.06, 33.07, 33.35 and 179.49.

The characterization by mass spectrometry and MS/MS analysis is shown in Figure 2.8, the fragmentation behavior of deprotonated $[\text{M-H}]^-$ is consistent with the referring structure giving the product ion at m/z of 99 from a neutral loss of H_2S as displayed in Scheme 2.9.

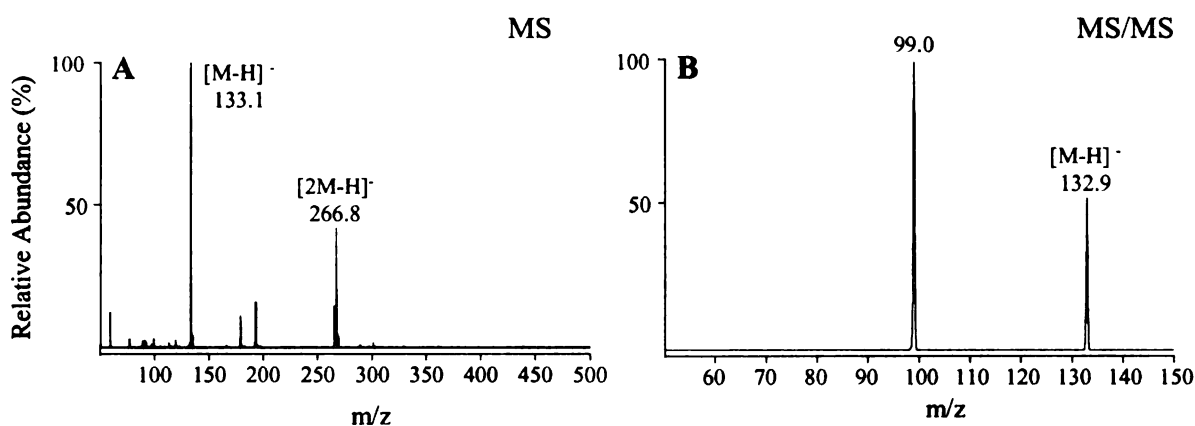
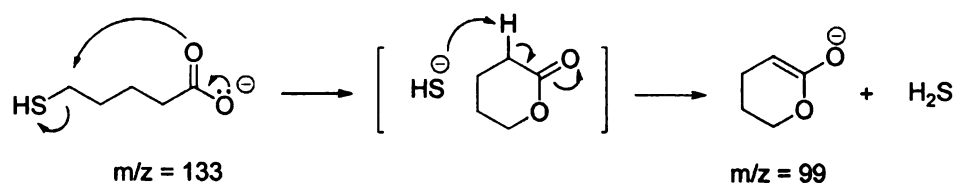


Figure 2.8 Characterization of (5) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The pseudomolecular ion $[\text{M-H}]^-$ at m/z 133.1⁻ and the dimer at m/z 266.8⁻; (B) CID-MS/MS of selected parent ion at m/z 132.9⁻



Scheme 2.9 Fragmentation mechanism of (5) anion. The spectrum is displayed on Figure 2.8B. Square brackets indicate ion-molecule complexes

2.5.2 5,5'-Thiodipentanoic Acid (6)

The acid (6) was obtained by the method described for the 1-carboxymethylmercaptopropionic acid (1), from 5-bromovaleric acid (3.4 g, 18.8 mmol) dissolved in 40 % NaOH (8 ml) and freshly prepared (5) (2.5 g, 18.7 mmol) dissolved in 40 % NaOH (8 ml) in 78 % (3.4 g) yield. ^1H NMR (300 MHz, CDCl_3): δ 1.49-1.67 (m, 8H), 2.26 (t, 4H, $J = 7.2$ Hz), 2.44 (t, 4H, $J = 7.2$ Hz); ^{13}C NMR (75 MHz, CDCl_3): δ 23.84, 28.80, 31.42, 33.39 and 176.79.

The characterization by mass spectrometry and MS/MS is shown in Figure 2.9. The deprotonated pseudo-molecular ion $[\text{M}-\text{H}]^-$ at m/z 233 (Figure 2.9A) was selected for fragmentation by CID-MS/MS, the fragments obtained were consistent with the referring structure, and the fragmentation pathways are demonstrated in Scheme 2.10.

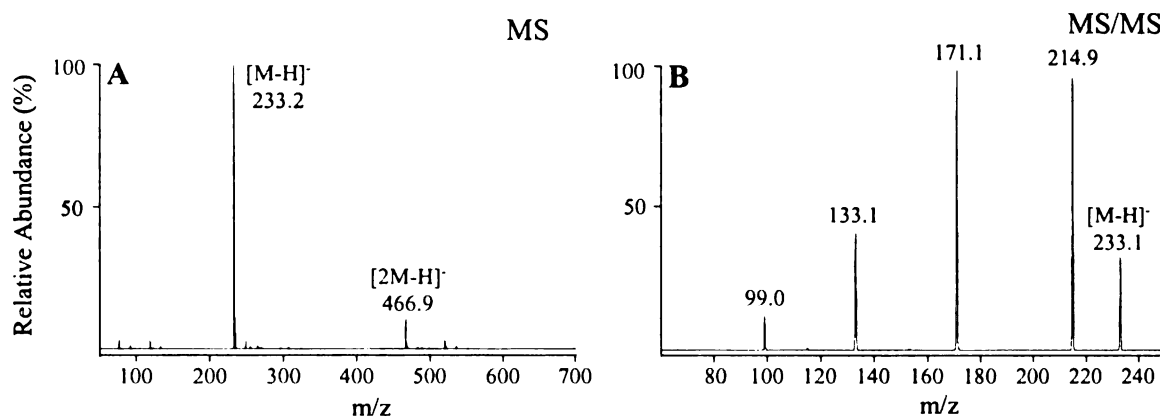
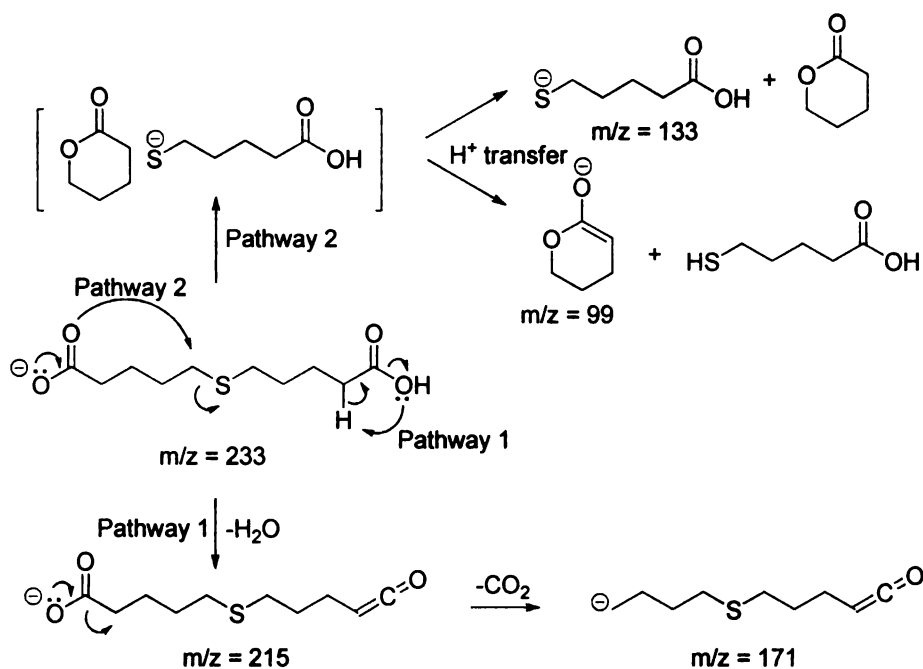


Figure 2.9 Characterization of compound (6) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The pseudo molecular ion $[\text{M}-\text{H}]^-$ at m/z 233.2 $^-$ and the dimer at m/z 466.9 $^-$; (B) CID-MS/MS of selected parent ion at m/z 233.1 $^-$



Scheme 2.10 Fragmentation mechanism of (6) anion. The spectrum is displayed on Figure 2.9B. Square brackets indicate ion-molecule complexes

2.5.3 5,5'-Thiodipentanoylhydroxysuccinimide (7)

This diNHS-ester was synthesized and purified in the same manner as thio 2-acetoxy-5'-pentanoyldihydroxysuccinimide 2. Thiodiacid (6) (1.17 g, 5.0 mmol), N-hydroxysuccinimide (1.44 g, 12.5 mmol) and DCC (2.58 g, 12.5 mmol) were dissolved in a 10:1 v/v mixture of chloroform and dichloromethane (55 ml) and allowed to react at room temperature overnight. The final crystallized product (7) was obtained as a light yellow solid in 84 % (1.8 g) yield. 1H NMR (300 MHz, $CDCl_3$): δ 1.63-1.71 (m, 4H), 1.76-1.84 (m, 4H), 2.50 (t, 4H, $J = 7.5$ Hz), 2.60 (t, 4H, $J = 7.5$ Hz), 2.78 (s, 8H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 23.60, 25.50, 28.35, 30.42, 31.03, 168.32 and 169.14.

Upon characterization by mass spectrometry and MS/MS, the singly charged $[M+H]^+$ ion at m/z 429.1 was the most abundant ion, the NH_4^+ and Na^+ adducts were also

present. The peak of m/z at 314.2 corresponds to an in source fragment of the $[M+H]^+$ ion, via the neutral loss of 115, a similar fragmentation pathway as that of (**2**) in Scheme 2.3 but no proton transfer observed.

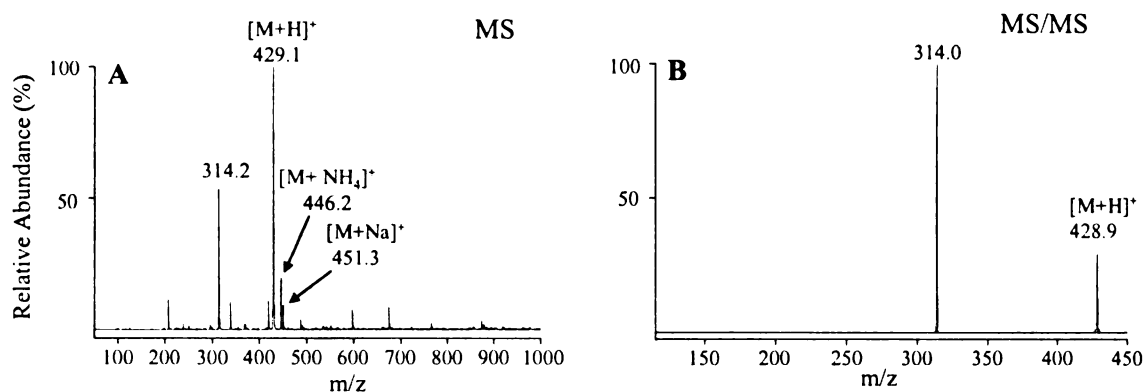


Figure 2.10 Characterization of (**7**) by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The pseudo molecular ion $[M+H]^+$ at m/z 429.1⁺, the NH_4^+ and Na^+ adducts are present; (B) The CID-MS/MS of selected parent ion at m/z 428.9⁺, m/z at 314.0 is the main fragment, which corresponds to 314.2 fragment in source in panel A

2.5.4 S-methyl 5,5'-Thiodipentanoylhydroxysuccinimide (**III**)

The methylation of (**7**) was performed by reacting with either iodomethane or dimethyl sulfate as methods mentioned above, from a starting substrate of 0.5 mmol (214 mg) scale. Both methods gave the same sulfonium ion with different counter ions; identical fragmentation behavior was also observed following low energy CID MS/MS. Figure 2.11 demonstrates positive mode mass spectrometry of molecular ion at m/z 433 and MS/MS spectra obtained by analysis of the sulfonium (**III**) methylsulfate. The sulfonium ion (**III**) was observed giving the same main product ion, 6-membered ring oxonium at m/z of 198, as shown in Figure 2.11B. Fragment at m/z 131 could be derived from a further fragmentation of complementary ion of 198.

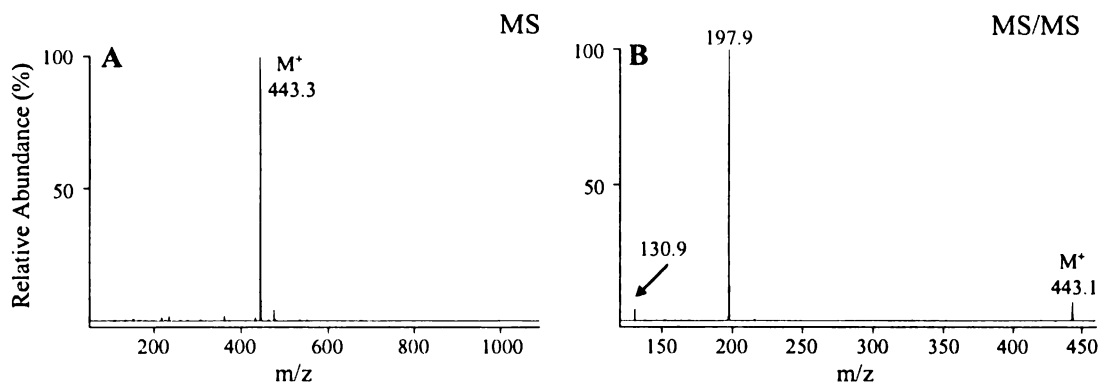


Figure 2.11 Characterization of ionic cross-linking reagent **III** by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The methyl sulfonium ion at m/z 443⁺; (B) CID-MS/MS product ion spectrum of M⁺

The sulfonium **III** iodide was purified by precipitation from dichloromethane/hexanes. ¹H NMR (500 MHz, CDCl₃): δ 1.89-1.91 (m, 8H), 2.68 (t, 4H, J = 6.5 Hz), 2.78 (s, 8H), 3.02 (s, 3H), 3.53 (t, 4H, J = 7.5 Hz); ¹³C NMR (125 MHz, CDCl₃): δ 22.51, 22.66, 23.00, 25.53, 29.99, 40.34, 168.21 and 169.73.

2.6 Peptide Cross-Linking Reactions

Peptides were dissolved in 40 μL of either DMF or phosphate buffered saline (PBS, 800 mg NaCl, 217 mg Na₂HPO₄•7H₂O, 20 mg KCl and 20 mg KH₂PO₄ per 100 ml, pH 7.5) to a concentration of 0.6 mM. The cross-linker solutions were prepared in DMF immediately before use. 4 μL of cross-linker solution was added to the 40 μL peptide solution and the reaction allowed to proceed at room temperature for 30 min-24 h. Cross-linking reactions were optimized by varying the cross-linker, cross-linker concentration and reaction time.

The product solutions were directly diluted after reaction with a solution of acetonitrile:water:acetic acid 60:40:1 v/v to a concentration of 10 μM for

nanoelectrospray when DMF was used as the reaction buffer. When PBS buffer was used, the products were desalted by sep-pak purification with elution performed by applying 2 ml of 40 %, 60 % and 80 % acetonitrile aqueous solution containing 0.05 % formic acid successively, then the product was subjected to mass spectrometry analysis without further dilution [20, 26].

CHAPTER THREE

ALKYL CHAIN LENGTH EFFECTS ON THE STABILITY AND FRAGMENTATION BEHAVIOR OF IONIC CROSS-LINKING REAGENTS

3.1 Introduction

An ionic cross-linking reagent that could have preference for cleavage under low energy CID MS/MS conditions has to satisfy a number of prerequisites: (1) the reagent must be stable when subjected to cross-linking reaction conditions, as minimum requirements, it must exhibit certain stability in aqueous solutions at mild pH values; (2) it must contain an appropriate linker length for detecting specific interaction sites within protein complexes, usually a short (4-8 Å) spacer arm is used in intra-molecular cross-linking studies, while intermolecular cross-linking is favored with a cross-linker containing a long spacer arm [7]; (3) the energy required for cleavage within the linker region should be lower than that for peptide backbone cleavage.

On the basis of previous experience, introducing a fixed charge could realize selective cleavage at a site adjacent to the fixed charge within a peptide ion [45-48]. Back et al. proposed to quaternarize the amine of BID in order to have a reaction channel independent of the charge states of cross-linked peptide for yielding marker ions [26]. From a previous study in our laboratory of the low energy CID-MS/MS spectra obtained from sulfonium and ammonium fixed-charge containing peptides [45], selective loss of dialkylsulfide from sulfonium derivatives is an energetically favored process; however the loss of trimethylamine from ammonium ion derivatives appears to be highly dependent on proton mobility. In a further investigation into the electronic effects of

sulfonium substitute [54], the positive charge on sulfur produces a strong dipole with resultant weakening C-S bond, and the ability of the sulfur atom to accept the electron pair of C-S bond enabled the sulfonium to be a better leaving group than an ammonium. These results suggest the incorporation of fixed charge sulfonium ions into chemical cross-linking reagents for approaches aimed at selective gas phase fragmentation.

The reactive groups of cross-linking reagents generally target primary amines or sulfhydryl groups in proteins. NHS esters are highly reactive amine-specific cross-linking reagents [55-57], targeted to the free N-termini and ϵ -amino groups in lysine side chains of proteins to create stable amide bonds. Due to the large number of lysine residues present on the surface of most proteins [2], it is likely that two lysine side chains will be close to each other at a protein-protein interface. Because of their low abundance, cysteines are less common on the surface of proteins, and because they are easily subject to oxidation reactions forming disulfide bonds, the sulfhydryl group is not as an attractive target functional group for the novel ionic cross-linking reagents, though it could also be employed. The application of NHS esters is usually limited by their poor water solubility for large scale protein complex studies, resulting in the use of the more expensive sulfo-NHS esters. However, simple members of ionic sulfonium salts typically exhibit high solubilities in water [54], thus sulfonium containing cross-linkers are expected to have an advantage of substantial solubility in both organic and aqueous solutions.

Cross-linking reactions often employ the use of a series of cross-linking reagents with different linker region lengths to map the comprehensive topology of protein and protein complex. Our initial design of cross-linking reagents has a general structure with NHS pentanoate at one side and an NHS acetate or propanoate at the other side of the

sulfonium sulfur. The third sulfur substituent could be either a methyl or phenacyl group, dependent on which alkylating agent was to be employed. NHS pentanoate forms a pentanamide bond following cross-linking reaction with primary amine in protein. When subject to CID MS/MS, nucleophilic attack from the carbonyl group on the linker would result in the formation of a protonated six-membered iminotetrahydropyran product ion upon cleavage on the C-S bond. This process is expected to have a low-energy transition state compared to the energy required for amide bond cleavage based on the previous theoretical calculations and experimental results described by Amunugama *et al.* [47].

In this chapter, the reaction characteristics as well as the application of these first two versions of the ionic cross-linking reagents (**I** and **II** in Scheme 2.1, 2.6) are discussed in detail. Experiments based on the multistage MS/MS of cross-linked neurotensin were carried out to provide evidence to support the formation of a protonated six-membered iminotetrahydropyran product ion as the dominant mechanism.

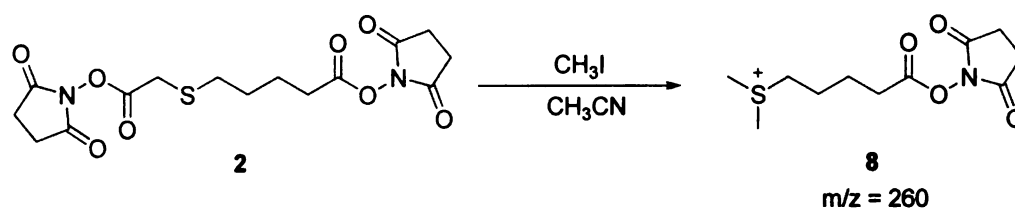
3.2 Cross-Linking Reagent S-Methyl 2-acetoxy,5'-pentanoyldihydroxysuccinimide sulfonium (Compound I in Scheme 2.1)

3.2.1 Stability of I following Reaction with Iodomethane

Iodomethane is a commonly used methylation agent, and an example of moderately soft electrophilic center [58], thus the methylation with iodomethane tends to occur at the soft nucleophile center such as a dialkylsulfide. Together with its volatility, iodomethane is expected to be a suitable alkylating agent to prepare sulfonium from di-NHS ester

intermediates. However, a reaction from **2** and iodomethane was not able to give sulfonium **I** but a dimethyl substitute sulfonium as shown in Scheme 3.1.

Figure 3.1 demonstrates the positive mode mass spectrometry of the molecular ion at m/z 260 and the MS/MS spectra obtained by analysis of the sulfonium (**8**) iodide. The proposed 6-membered ring oxonium product ion observed at m/z of 198 from the neutral loss of dimethyl sulfide from (**8**) was consistent with the referring structure.



Scheme 3.1 Dimethyl sulfonium **8** obtained from methylation of **2** with iodomethane

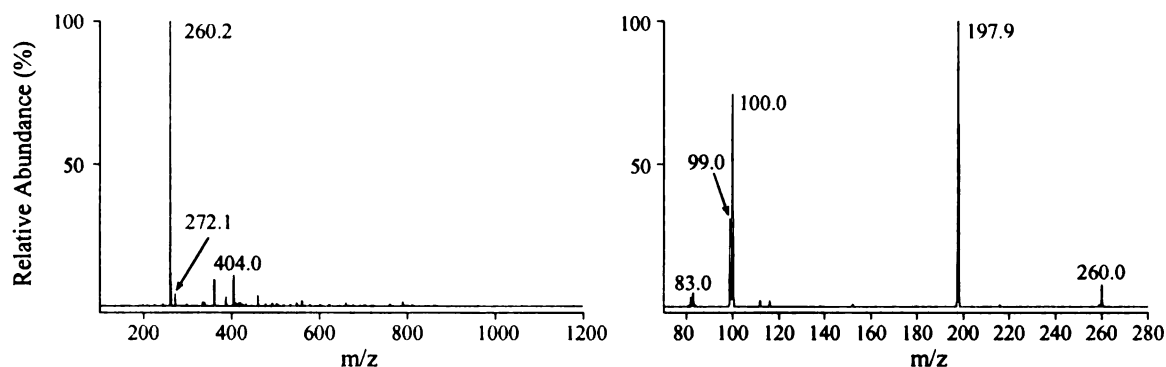
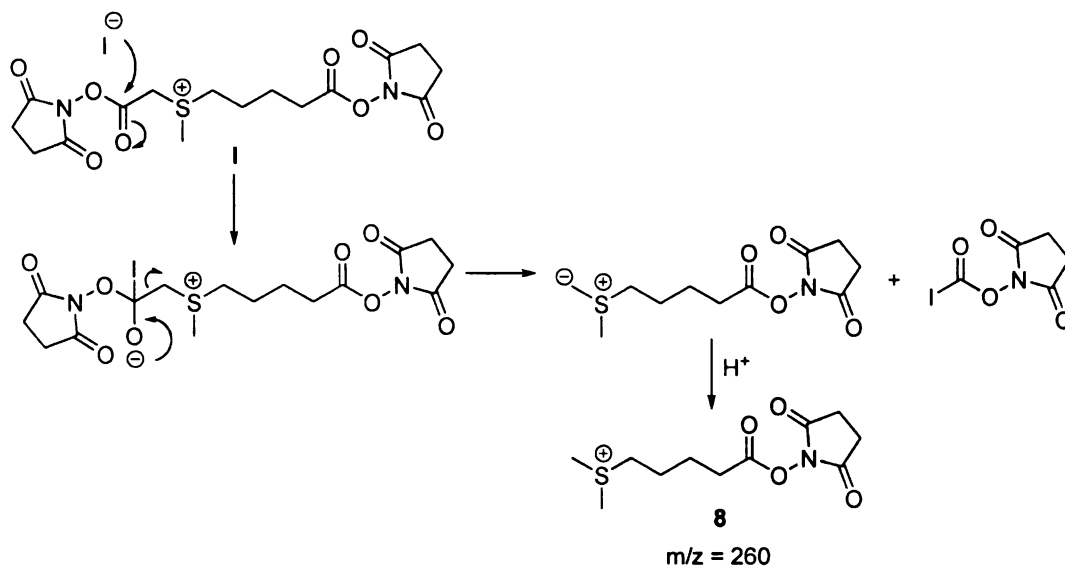


Figure 3.1 Characterization of dimethyl sulfonium **8** by nanoelectrospray quadrupole ion trap mass spectrometry. (A) The dimethyl sulfonium ion at m/z 260⁺, peak at m/z 404 corresponds to ammonium adduct of reactant **2** and m/z 272 is a fragment generated in source from it; (B) CID-MS/MS product ion spectrum of M^+

Iodide is generally a rather poor nucleophile at carbon centers to initiate a nucleophilic addition reaction at the carbonyl group, however when the carbonyl carbon is only one methylene group apart from a sulfonium positive charge center, it becomes

more electron deficient and hence becomes more easily subject to iodide nucleophilic attack, leading to reactions giving rise to **8** shown in Scheme 3.2.



Scheme 3.2 Mechanism of dimethyl sulfonium formation by reaction of iodomethane with **I**

Cross-linking reagent **I** was not successfully prepared by methylation with iodomethane due to the electron deficient carbonyl carbon. Presumably due to the steric hindrance resulting from the bulky NHS ester group, phenacyl bromide could also not be employed to alkylate **2**. The close arrangement of the NHS ester group and sulfur center in space not only brings difficulty to the preparation of sulfonium from **2** but also has great impact on the stability of sulfonium NHS ester functional group which has tendency to undergo hydrolysis and transesterification reactions [7].

3.2.2 Transesterification and Hydrolysis Reactions of **I**

I was easily subject to transesterification and hydrolysis because of the electron deficient carbonyl carbon on the short chain length side of the sulfonium ion crosslinking reagent. When 50 % MeOH + 50 % H₂O + 1 % AcOH was employed for analysis of **I** by MS, no molecular ion was observed. Instead, dominant methyl transesterification product (m/z 318) and significant hydrolysis product (m/z 304) were observed as shown in Figure 3.2 and Scheme 3.3.

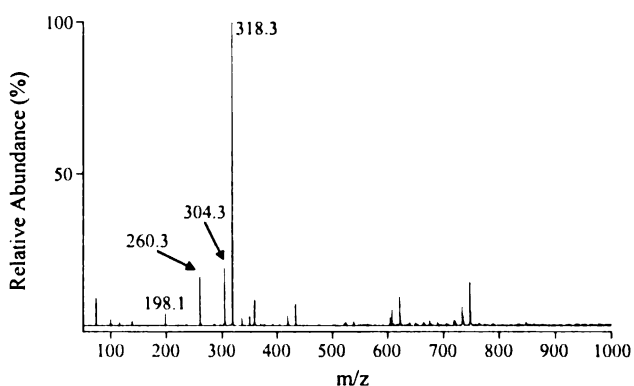
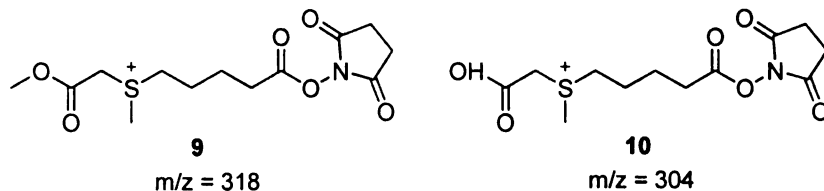


Figure 3.2 Nanoelectrospray quadrupole ion trap mass spectrum of sulfonium **I** methyl sulfate in 50 % MeOH + 50 % H₂O + 1 % AcOH MS buffer



Scheme 3.3 Structures of methyl transesterification and hydrolysis products of **I**

Transesterification and hydrolysis were also observed when **I** was stored in dry DMF solutions or when diluted in 100 % AcCN. A 3 mM stock solution of **I** in dry DMF was diluted to 20 μ M by AcCN at different time and immediately injected to the mass spectrometer. A notable increase in the relative abundance of **9** and **10** and a decrease of **I**

was observed during a 4-hour period storage in DMF (Figure 3.3), giving evidence for the instability of **I** even when potential nucleophiles were minimized. Trace amounts of water could initiate substantial hydrolysis reactions in a relative short time following the proposed mechanism in Scheme 3.4, while a similar mechanism for methanol could result in methyl esterification.

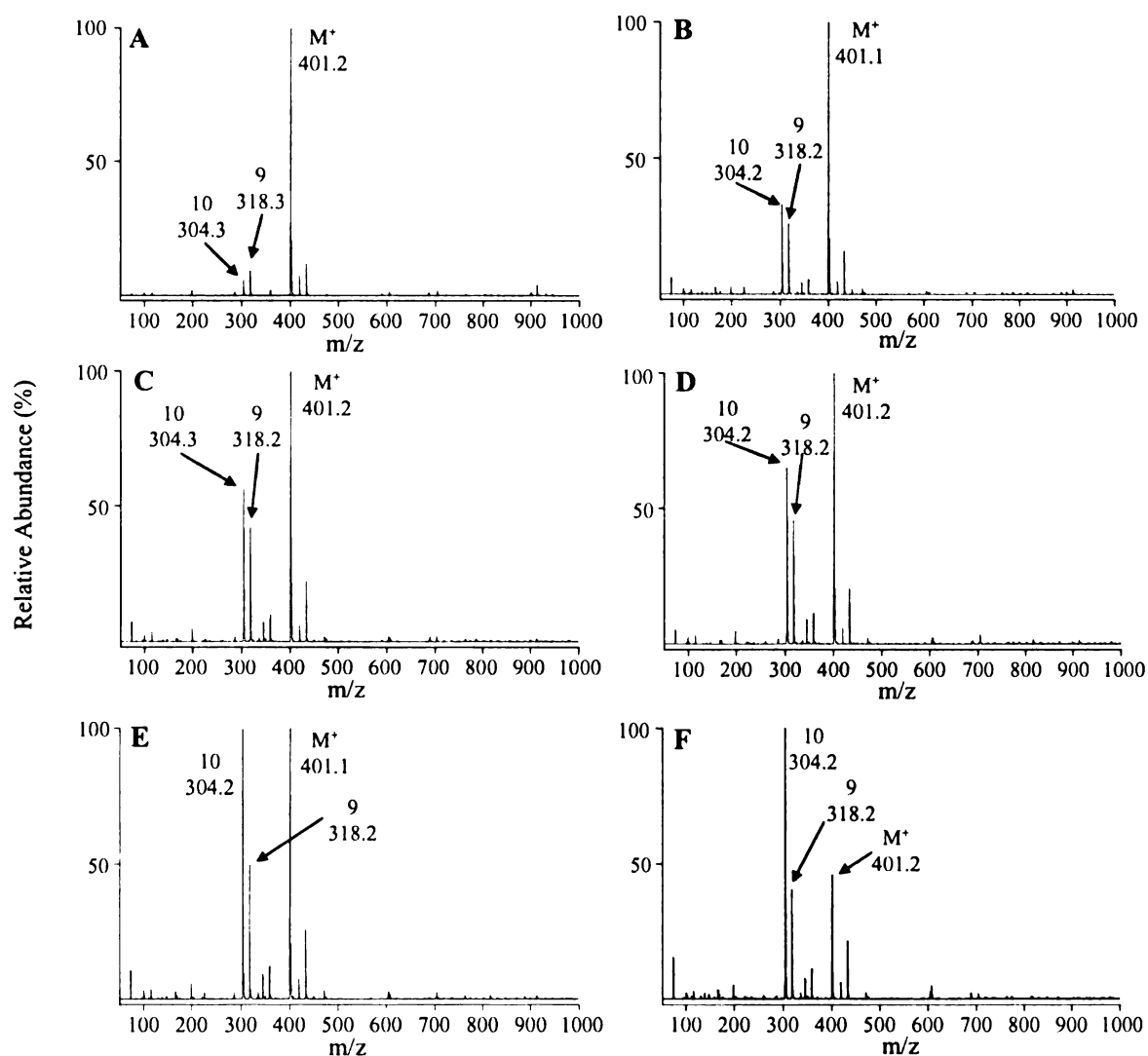
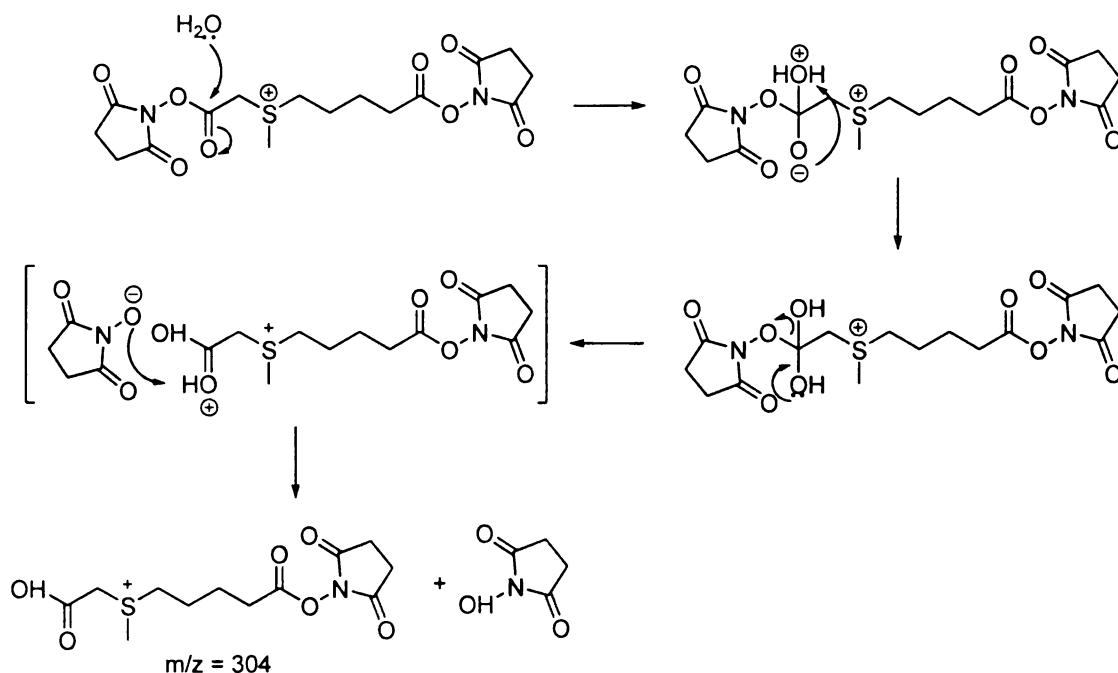


Figure 3.3 Nanoelectrospray quadrupole ion trap mass spectra of sulfonium **I** methyl sulfate stored in dry DMF, 100 % CH₃CN was employed as MS solvent; (A) 0 min; (B) 45 min; (C) 75 min; (D) 2 hr; (E) 3 hr; (F) 4 hr



Scheme 3.4 Mechanism of hydrolysis of **I**, methanol has the similar process as a nucleophile for methyl transesterification

3.2.3 Supplementary Evidence from 1-Carboxymethylmercaptopropionic Acid (Compound **1** in Scheme 2.1)

Additional evidence for a highly reactive carbonyl carbon one methylene group from the sulfonium positive charge center was obtained from the reaction of **1** with iodomethane under the same conditions as those employed for **2** (Figure 3.4 and Scheme 3.5).

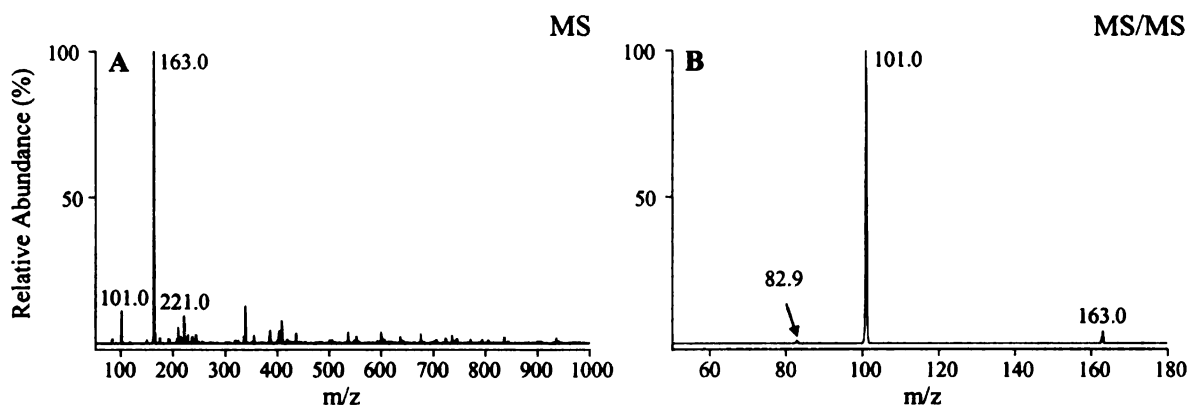
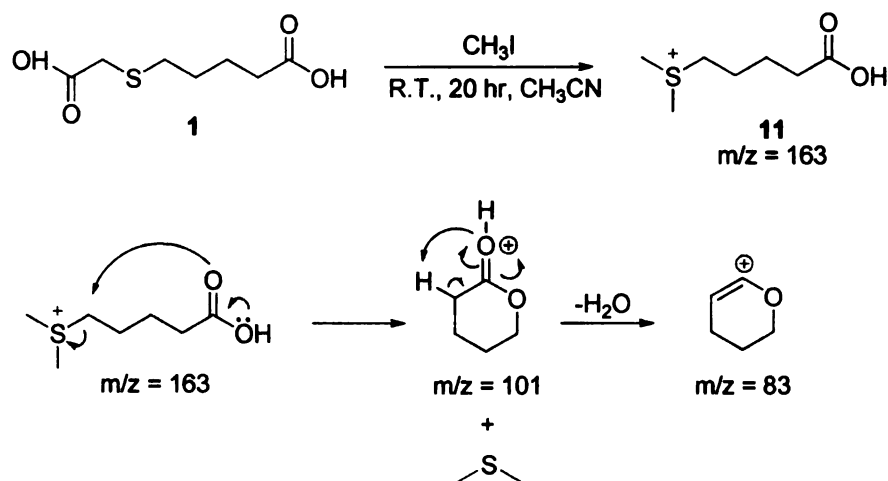


Figure 3.4 Nanoelectrospray quadrupole ion trap mass spectra from methylation of **1** with iodomethane. Similar mechanism as indicated in Scheme 3.2 to form dimethyl sulfonyl. (A) The dimethyl sulfonyl ion at m/z 163⁺, peak at m/z 101 corresponds to a fragment in source; (B) CID-MS/MS product ion spectrum of ion at m/z 163⁺



Scheme 3.5 Methylation reaction of **1** with iodomethane and potential mechanisms for gas phase fragmentation of dimethyl sulfonyl product

Based on the results described above, cross-linking reagent **I** would be expected to undergo fragmentation via a charge-directed mechanism resulting in formation of a dominant six-membered ring product ion. However, the highly reactive carbonyl carbon limits the application of this reagent under aqueous buffer conditions.

3.3 Cross-Linking Reagent S-Methyl 3-propionoyl,5-pentanoyldihydroxysuccinimide sulfonium (Compound II in Scheme 2.6)

3.3.1 Stability in Iodomethane

The only difference in the structure of **II** and **I** is the inclusion of an additional methylene group at the short NHS ester end. Because of this methylene group, the electron density of the carbonyl carbon is not so strongly withdrawn by the positive charge on the sulfur, making it much less reactive than that of **I**. As a result, **II** is quite stable when present together with iodomethane and no significant dimethyl sulfonium could be observed.

3.3.2 Stability under Aqueous Conditions

The addition of one methylene group improves the stability of **II** not only in the presence of excess iodomethane, but also under aqueous conditions. The mass spectra of freshly prepared **II** in 100 % CH₃CN or in 50 % CH₃CN + 50 % H₂O MS solvent are quite similar (Figure 3.5).

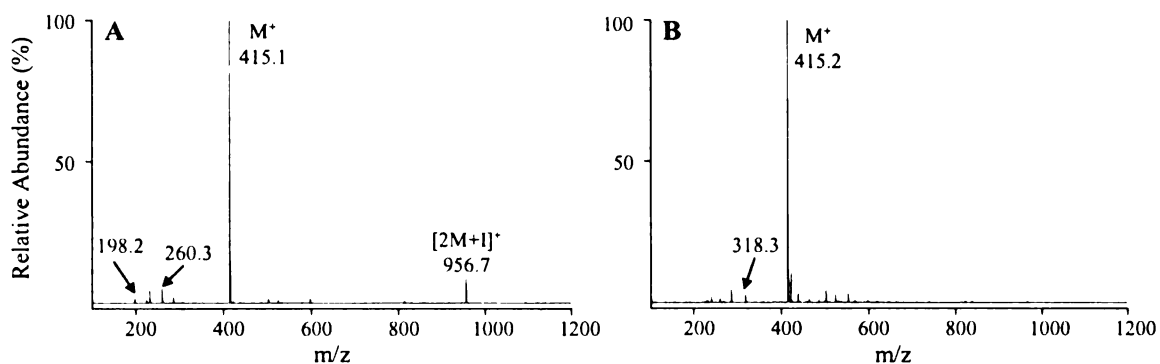


Figure 3.5 Nanoelectrospray quadrupole ion trap mass spectra of sulfonium **II** iodide. (A) In 100 % CH₃CN buffer, small peak of dimethyl sulfonium (m/z = 260, compound **8** in Scheme 3.2) observed; (B) in 50 % CH₃CN + 50 % H₂O buffer, a small peak resulting from methyl transesterification (m/z = 318, compound **9** in Scheme 3.3) was observed

3.3.3 Supplementary Evidence from 1-Carboxyethylmercaptopropionic Acid (Compound **3** in Scheme 2.6)

To further demonstrate the improvement of stability brought by adding one methylene group between the carbonyl carbon and the sulfur atom, similar methylation experiment with iodomethane was carried out for **3**. On the basis of our previous experience, we could expect methylated **3** (compound **12**, m/z = 221) as the main product rather than the dimethyl sulfonium (compound **11**, m/z = 163). As shown in Figure 3.6, m/z 163 was observed at only low abundance compared to m/z 221.

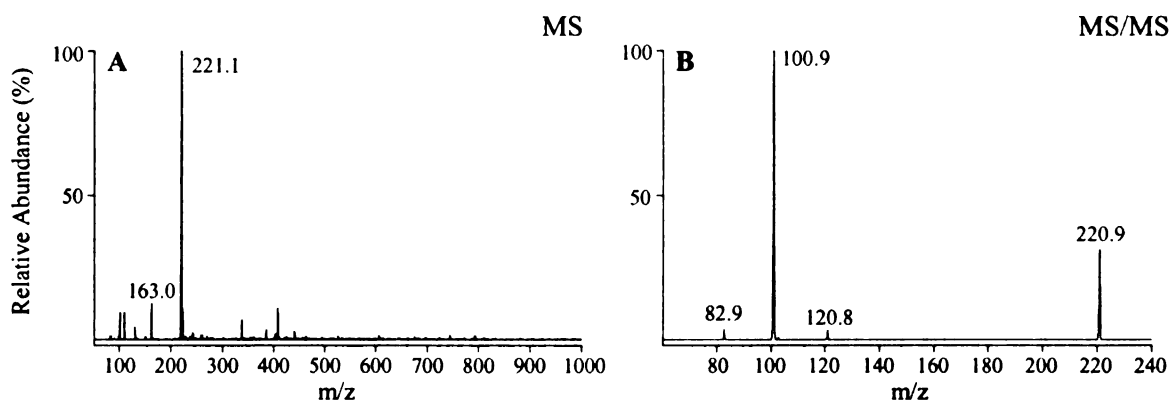
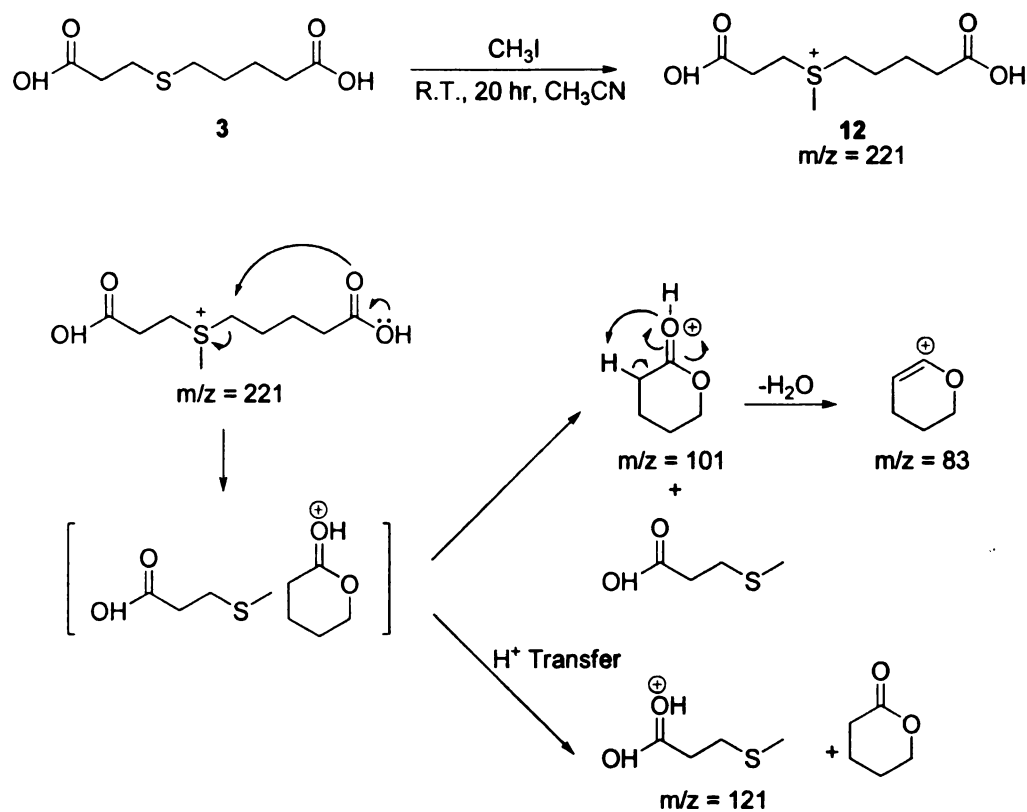


Figure 3.6 Nanoelectrospray quadrupole ion trap mass spectra from methylation of **3** with iodomethane. (A) Expected product **12** at m/z 221 is dominant peak, small amount of dimethyl sulfonium (m/z = 163) (Compound **11** in Scheme 3.5) by-product observed. (B) CID-MS/MS product ion spectrum of ion at m/z 221⁺

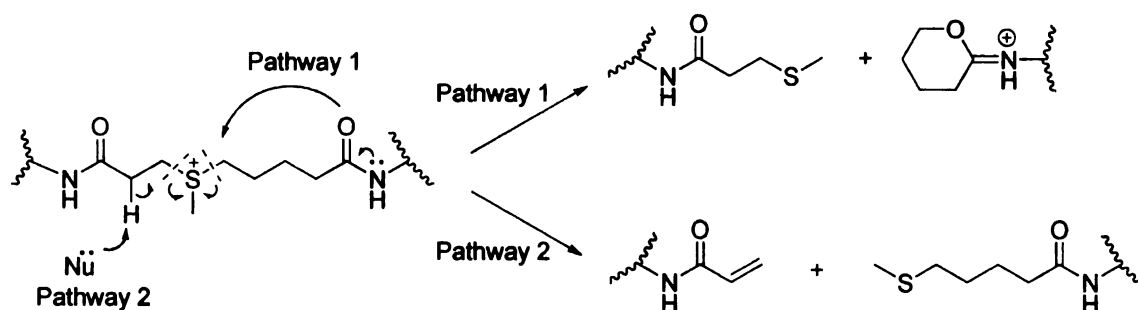


Scheme 3.6 Methylation reaction of **3** with iodomethane and potential mechanisms for gas phase fragmentation of methylated product **12**

The improved stability of cross-linking reagent **II** was experimentally demonstrated above. The addition of one methylene group also alleviates steric hindrance on sulfur brought by the large NHS ester group. As a result, phenacyl bromide could be employed to alkylate **4** (Scheme 2.6) to yield a phenacyl substituted sulfonium ion (Figure 2.7).

3.3.4 Suspected E2 Elimination Reactions Might Lead to Unintended Cleavage on Cross-Linker

The additional methylene group makes it possible, however, for E2 elimination reactions to occur. During this process, a proton from the methylene group adjacent to the carbonyl on the short chain is transferred to a nucleophile, which could come from the same or other molecules. As illustrated in Scheme 3.7, for the case of an intermolecular cross-linked peptide, pathway 2 would result in intramolecular proton transfer and bond cleavage to yield alternative product ions.



Scheme 3.7 Potential fragmentation mechanisms of **II** intermolecular cross-linked peptides

Based on the previous study by Amunugama *et al.* [47], the formation of protonated six-membered tetrahydropyran product ion via pathway 1 is predicted to be the energy favored fragmentation process. However, the possibility of cleavage on the other C-S

bond due to the E2 mechanism is not able to be excluded. The coexistence of two fragmentation pathways of **II** cross-linked peptides would complicate the resultant MS/MS spectra. To simplify the identification of cross-linked peptides, a symmetrical ionic cross-linking reagent **III** was developed.

3.4 Cross-Linking Reagent S-Methyl 5,5'-dipentanoylhydroxysuccinimide sulfonium (Compound **III in Scheme 2.8)**

III is analogous to **I** and **II**, but contains 4 methylene groups on both chains. On the basis of our previous experiments, we could expect **III** to be most stable among three reagents. Under CID-MS/MS conditions the symmetrical structure results in formation of the same products upon cleavage of either C-S bond in **III**.

3.4.1 Stability in Aqueous Conditions and Methanol

The hydrolysis reaction of NHS-ester is a major competing reaction of the NHS-ester acylation reaction when carrying out cross-linking reactions. Consequently, stability of cross-linking reagents in aqueous buffers must be considered. Similar experiments as described above were performed using freshly prepared **III** in water or/and methanol containing MS solvents. No obvious hydrolysis or transesterification was observed from **III** in water or methanol, even under acidic conditions (Figure 3.7). Because of the solubility and stability of this reagent under both aqueous and non-aqueous conditions, **III** has the potential to cross-link a large range of proteins under varying conditions.

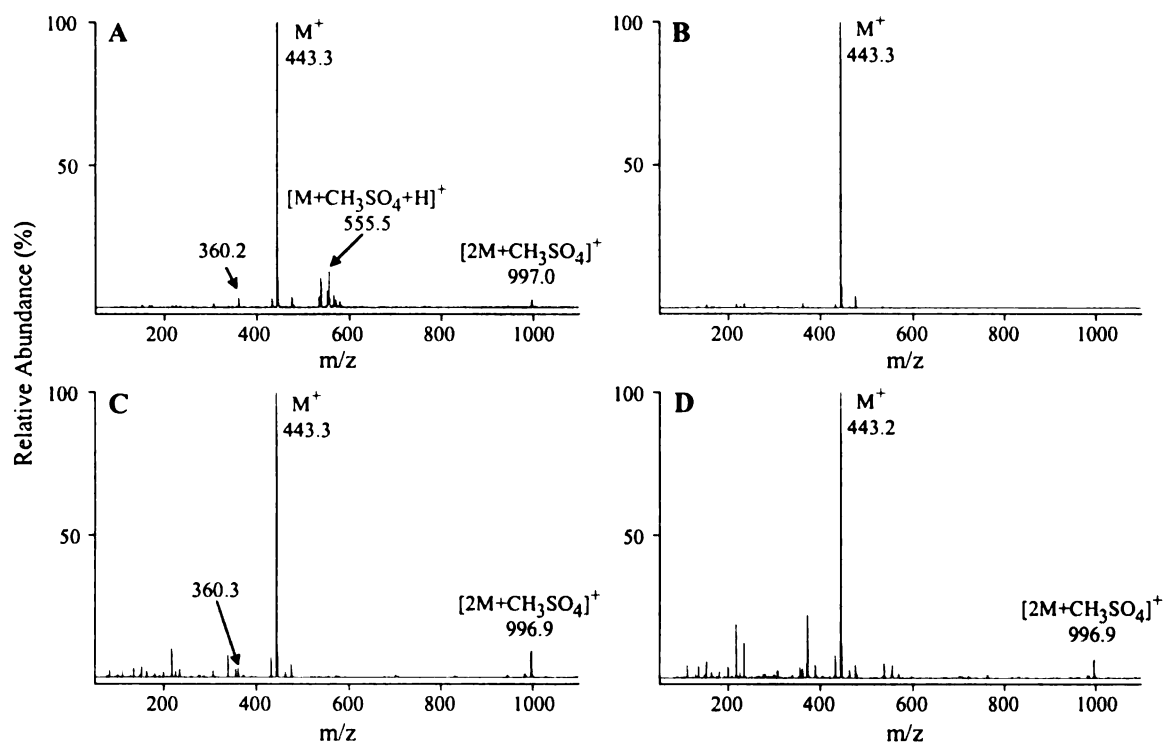


Figure 3.7 Nanoelectrospray quadrupole ion trap mass spectra of sulfonium **III** methylsulfate. (A) In 100 % CH₃CN buffer, small peak of methyl ester sulfonium ($m/z = 360$, analogous to compound **9** in Scheme 3.3) observed; (B) in 50 % CH₃CN + 50 % H₂O buffer; (C) in 50 % AcCN + 50 % H₂O + 1 % AcOH buffer; (D) in 50 % MeOH + 50 % H₂O + 1 % AcOH buffer

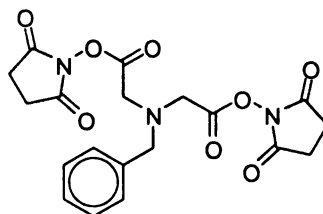
3.4.2 Comparison among Three Ionic Cross-Linking Reagents and Their Intermediates

The structure of the three cross-linking reagents and their synthetic intermediates only differ in the length of the carbon chain on one side of the sulfur atom. For the sulfonium compounds, the methylene group in the shorter chain is more reactive due to electron withdrawing effects resulting from the vicinal carbonyl group and positive charge centers. One consequence is that this reagent is more vulnerable to nucleophilic attack. Although **III** was found to satisfy all the principal requirements for a gas phase

cleavable cross-linking reagent, **I** and **II** could also be useful for varying the arm chain length in some specific applications and reaction conditions.

3.5 Cross-Linked Neurotensin

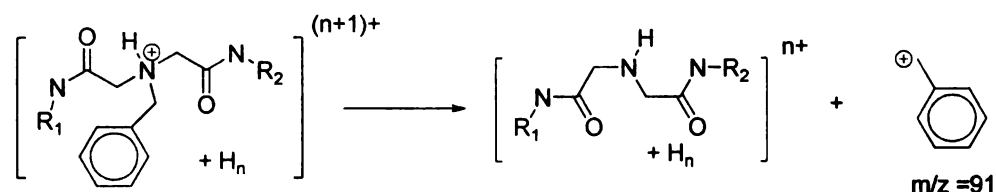
Neurotensin (pELYENKPRRPYIL, MW 1672.9) contains only one lysine residue at position six, with a cyclized N-terminal glutamic acid. Under cross-linking reaction conditions, the primary amine at this lysine residue is the only available reactive group toward NHS-esters. Consequently, only intermolecular cross-link and/or mono-link products could be observed from a cross-linking reaction. The use of a gas phase cleavable cross-linking reagent to aid the identification of cross-linked neurotensin has previously been described by Back *et al.* using their BID-NHS reagent (Scheme 3.8), which could yield a low mass specific marker ion under low energy CID conditions [26]. As a comparison, cross-linking reactions with **I**, **II** and BID-NHS were performed following a method modified from that of Back *et. al.*, using a 1:1 molar ratio of cross-linking reagent to neurotensin in DMF, a reaction time of 30 min, and a 20-fold excess of water added after reaction for 15 min to hydrolyze unreacted NHS-ester groups.



Scheme 3.8 Structure of the cross-linker N-benzyliminodiacetylhydroxysuccinimide (BID-NHS). Adapted from Reference 26

3.5.1 Cross-Linking Reaction with BID-NHS

The proposed fragmentation reaction from BID-NHS intermolecular cross-linked peptides is illustrated in Scheme 3.8.



Scheme 3.9 Proposed mechanism for the formation of benzyl cations from BID-NHS cross-linked peptide. Adapted from Reference 26

Similar to that described by Back, this reaction was only found to occur during the fragmentation of cross-linked neurotensin in its 5+ charge state (Figure 3.8A). At other charge states, the dominant processes were peptide backbone cleavage or side chain neutral losses that were not able to give diagnostic information for identification (Figure 3.8B-D).

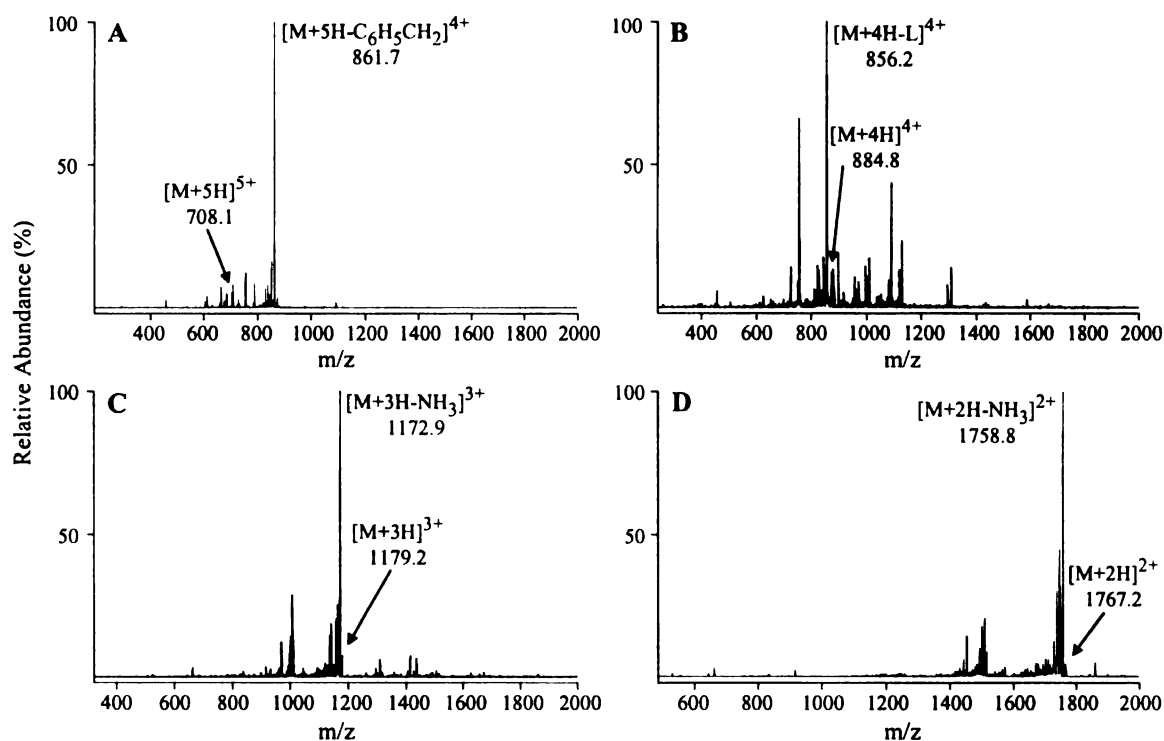


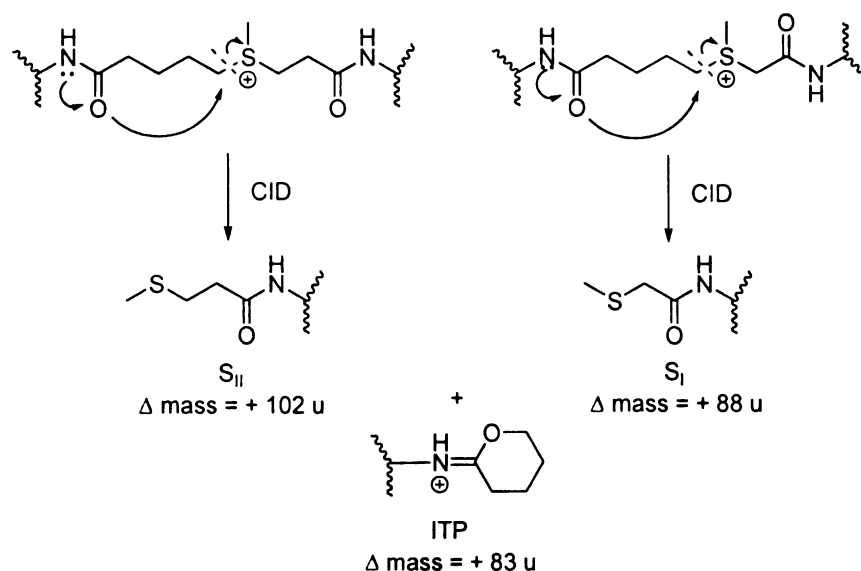
Figure 3.8 Low energy CID MS/MS of BID-NHS cross-linked neurotensin. (A) Precursor ion in the 5+ charge state, product ion at m/z 861 corresponds to the loss of benzyl cations; (B) precursor ion in the 4+ charge state, most abundant product ion at m/z 856 corresponds to peptide backbone cleavage between I and L residues; (C) precursor ion in the 3+ charge state, dominant product ion resulted from neutral loss of NH_3 from N or R side chains; (D) precursor ion in the 2+ charge state, dominant product ion resulted from neutral loss of NH_3 from N or R side chains

A “mobile proton” condition is required to realize the formation of expected marker ions from BID-NHS cross-linked peptides. In the case of neurotensin, 4 ionizing protons are expected to be ‘sequestered’ on the side chains of arginine, lysine and histidine residues in the peptide sequence. Thus, to initiate cleavage between the nitrogen and the benzyl moiety, a fifth proton would be required. In order to realize the intended cleavage independent of the charge state of the cross-linked peptides, Back proposed the possibility of quaternarizing the amine in the linker chain. However, presumably due to severe steric hindrance of the NHS-ester and benzyl functional groups, fixing the charge

on the amine next to the desired leaving group could not be achieved here directly from BID-NHS.

3.5.2 Cross-Linking Reaction with Ionic Cross-Linking Reagent I and II

The proposed fragmentation reactions of **I** and **II** intermolecular cross-linked peptides are illustrated in Scheme 3.9. The labile C-S bond at the long chain length side of reagents as indicated by dashed lines allows for preferential cleavage to produce two unique peptides, each containing an additional mass modification corresponding to the remaining portion of the cross-linker fragments. Because **I** and **II** only differ by one methylene group, fragmentation of **I** and **II** intermolecular cross-linked peptides are expected to generate one peptide with an additional mass of 83 u for the ITP (iminotetrahydropyran) modification, and another peptide with either a 88 u or 102 u sulfide modification (S). Tandem mass spectrometry of individual peptides at higher energy CID then allows peptide identification and determination of the sites of modifications via interpretation of the y- and b-type ions generated. Note that the feature of ITP which distinguishes it from other neutral modifications is that it results in the addition of one charge to the product ion.



Scheme 3.10 Proposed fragmentation mechanism of **I** and **II** intermolecular cross-linked peptides. The resulting mass additions (relative to the single peptide) following cleavage on the C-S bond are indicated, where S denotes a sulfide modification and ITP denotes protonated iminotetrahydropyran modification; I and II indicate cross-linking reagent

The cross-linked neurotensin in its 5+ charge state, formed by reaction with either **I** or **II**, were analyzed by MS/MS (Figure 3.9). As expected, the fragmentation occurred exclusively at the desired C-S bonds and the peptide chains remained intact to generate simple MS² spectra. Two peaks emerged: one corresponding to the triply charged neurotensin for the protonated ITP modification ($[\text{MH}_2 + \text{ITP}]^{3+}$), which is the same product ion from **I** or **II**, and the other corresponding to doubly charged neurotensin for the S modification ($[\text{MH}_2 + \text{S}]^{2+}$).

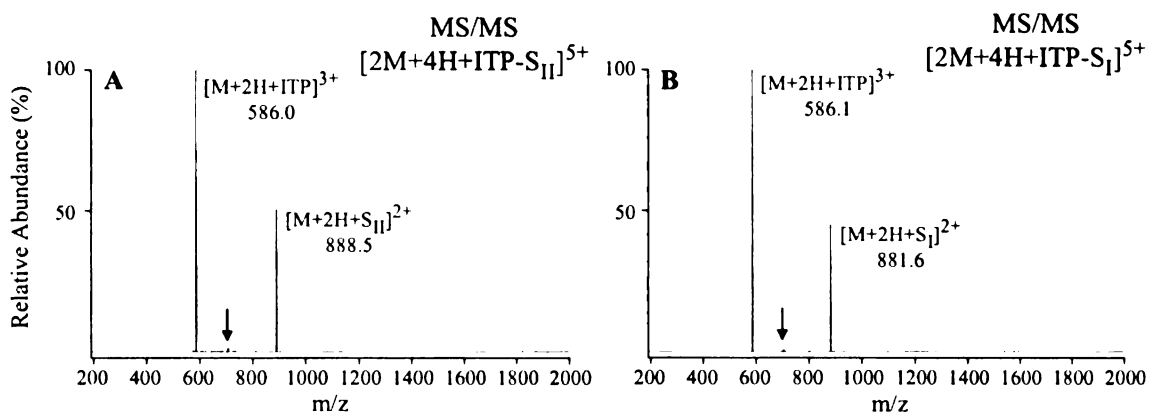


Figure 3.9 Low energy CID MS/MS spectra of ionic cross-linked neurotensin. (A) From cross-linker II; (B) from cross-linker I. The arrow indicates the precursor ions: (A) m/z 706.9, $z = 5$; (B) m/z 704.2, $z = 5$

To obtain additional sequence information, the two peaks resulting from the MS/MS fragmentation reactions were selected and subjected to MS³. The sequence of neurotensin and the sites of modification were then identified from the product ion spectra. As shown in Figure 3.10A and 3.10B, the observation of a series of b- and y-type ions provided the sequencing information for the S modified neurotensin. In particular, the existence of fragment ions at y_7 which do not contain the modification and y_8 -NH₃ which does confirmed that the modification occurred at Lys 6 of the peptide. The MS/MS spectrum of unmodified neurotensin in its 2+ charge state is shown in Figure 3. It is noticeable that the S modifications did not significantly affect the fragmentation pattern.

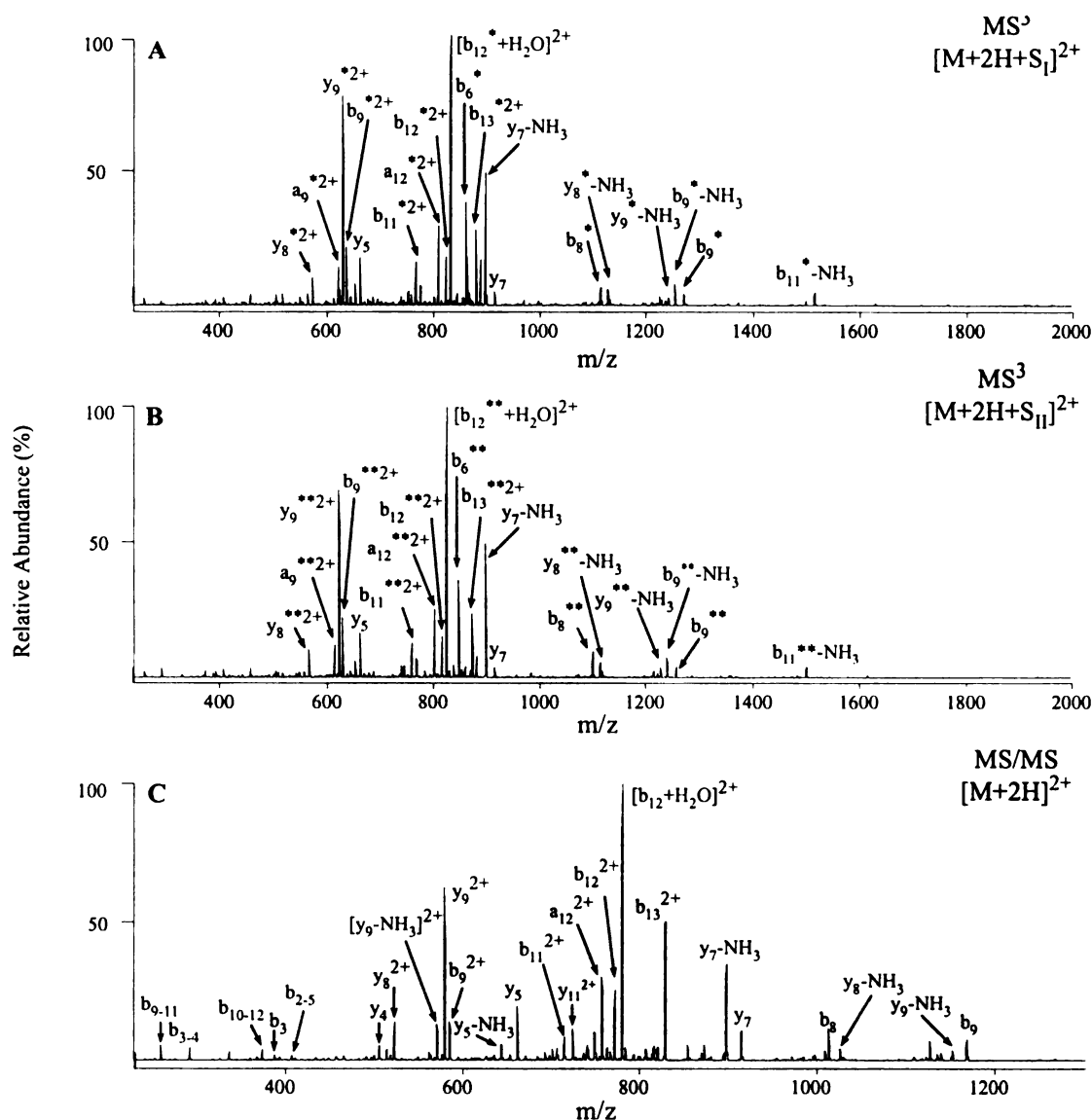


Figure 3.10 Low energy CID MS³ or MS/MS of neurotensin. (A) MS³ of S modified neurotensin from cross-linker I, * indicates S_I modification; (B) MS³ of S modified neurotensin from cross-linker II, ** indicates S_{II} modification (C) MS/MS of unmodified neurotensin in 2+ charge state

Similarly, the triply charged ITP modified precursor ions at m/z 586 were isolated and fragmented (Figure 3.11A). The fragmentation pattern of ITP modified neurotensin was observed to be quite similar to that of unmodified neurotensin in 2+ charge state instead of its 3+ charge state (Figure 3.11B), presumably due to the extra charge being

fixed on the ITP group, and thus does not change the proton mobility conditions of the peptide.

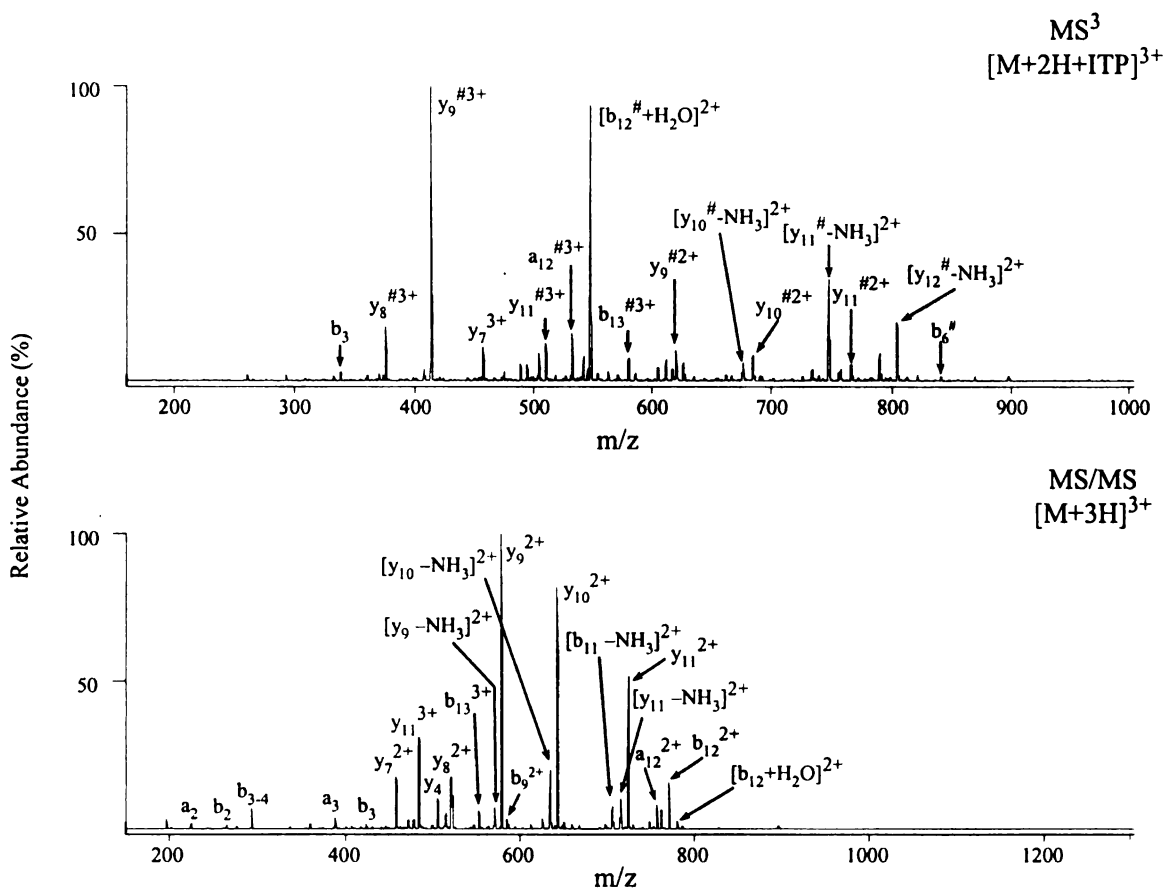


Figure 3.11 Low energy CID MS³ or MS/MS of neurotensin. (A) MS³ of ITP modified neurotensin, both **I** and **II** give same product ions; (B) MS/MS of unmodified neurotensin in 3+ charge state

The capability of our ionic cross-linking reagents and methodology was experimentally demonstrated from the preliminary work of **I** and **II**. Exclusive cleavage on desired C-S bonds under low energy CID makes identification of cross-linked peptides straightforward. However, the instability of these reagents under aqueous conditions limited their utility.

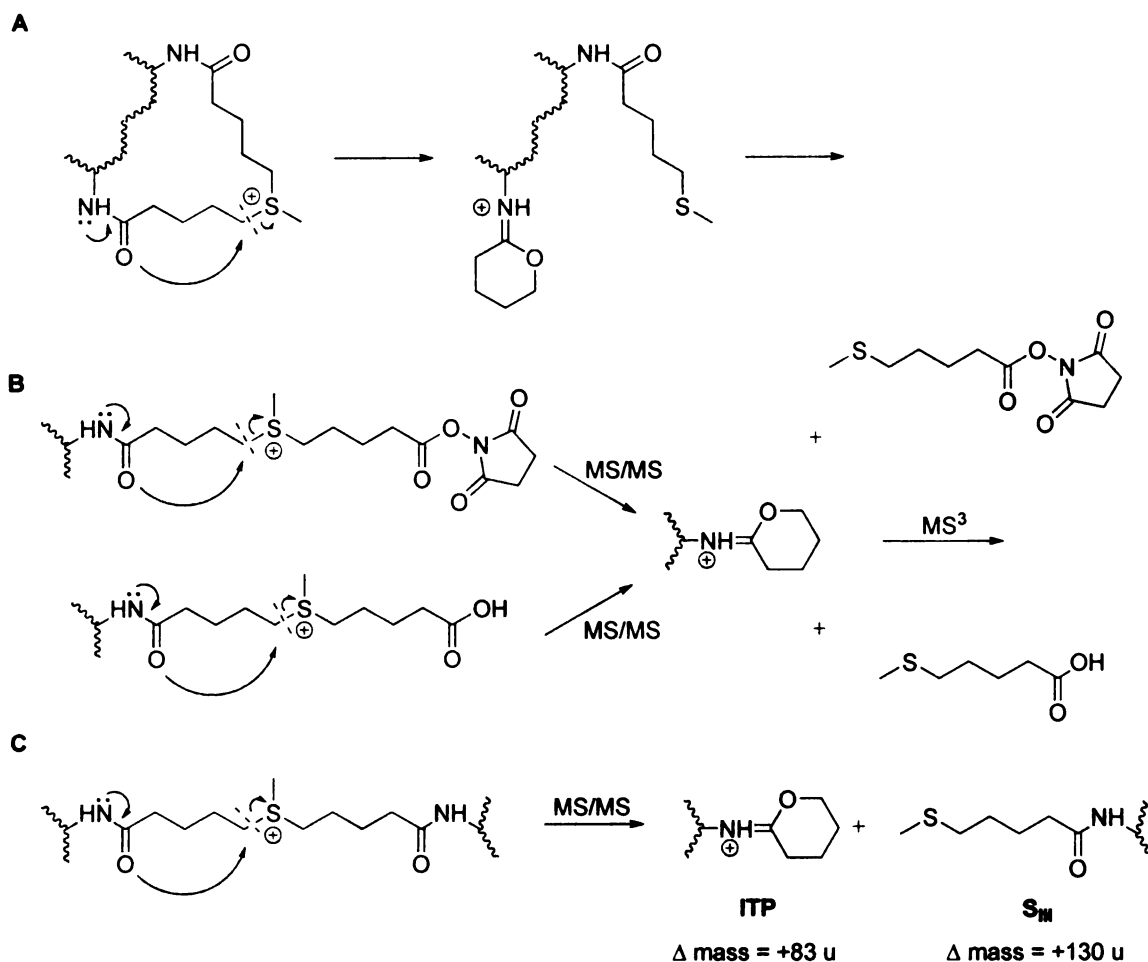
CHAPTER FOUR

IDENTIFICATION OF CROSS-LINKED PEPTIDES FOR PROTEIN INTERACTION STUDIES USING A SYMMETRICAL IONIC CROSS-LINKING REAGENT AND MASS SPECTROMETRY

4.1 Introduction

Due to its improved stability under aqueous conditions, cross-linking reactions with **III** could be optimized in PBS buffer; one of most commonly used reaction buffers for such purpose [11, 14, 23]. Generally, there are three main cross-link types generated from a single cross-linking reaction as described earlier, mono-link (type 0), intra-peptide cross-link (type 1), and inter-peptide cross-link (type 2). Usually mono-link is the modification of peptide with one reactive group of cross-linking reagent and the other one hydrolyzed. However from our previous experience there is still some unhydrolyzed mono-link observed under typical conditions. Thus two types of mono-link are described in this work: hydrolyzed and unhydrolyzed.

Potential mechanisms for fragmentation of the three cross-link types generated by reaction with **III** are depicted in Scheme 4.1. For an intra-peptide cross-link (Figure 4.1A), the desired C-S bond cleavage is expected to be the energetically favored process. However, no change of m/z will be observed during the event allowing further fragmentation to directly occur to yield b- and y-type ions. Due to the symmetrical structure of **III**, two C-S bonds are expected to be cleaved with similar efficiencies, though only one is displayed in Scheme 4.1A. Therefore it should be noted that a given b or y product ion may be observed containing both ITP and S modifications.



Scheme 4.1 Proposed fragmentation mechanism of **III** cross-linked peptides. The resulting mass additions (relative to the single peptide) following cleavage on the C-S bond are indicated, where S denotes a sulfide modification and ITP denotes protonated iminotetrahydropyran modification. (A) Intramolecular cross-link; (B) unhydrolyzed and hydrolyzed mono-link; (C) intermolecular cross-link

For the unhydrolyzed and hydrolyzed mono-link (Scheme 4.1B), it is predicted that the major cleavage takes place in the C-S bond closer to the peptide chain due to the amide nitrogen is greater nucleophile compared to either the ester or acid oxygen.

Similar to **I** and **II** described earlier, two separated peptide chains would be generated upon low energy CID from an inter-molecular cross-linked peptide of **III** (Scheme 4.1 C): one corresponding to ITP modification with an additional mass of 83 u

and the other corresponding to S modification with an additional mass of 130 u. If this is a homo-dimer cross-linked peptide, i.e. the two peptides connected together are the same; single ITP and S modified product ions will emerge in the MS/MS scan due to symmetrical structure of **III**. Otherwise for hetero-dimer cross-linked peptides, a pair of ITP and S modified product ions will be generated. Further tandem mass spectrometry (MS^3) of the individual product ions allows the identification of the peptide sequence and modification sites.

4.2 Cross-Linking Reaction with Neurotensin

4.2.1 Cross-Linking Reactions in DMF and PBS

Neurotensin was dissolved in solvents to a concentration of 0.6 mM for all the cross-linking reactions as mentioned before. The cross-linking of neurotensin in DMF, carried out at 1:2 molar ratio of cross-linker to peptide and a reaction time at 75 min, without hydrolysis, was found to give the highest relative yield of inter-molecular cross-link (Figure 4.1A). However, prominent unmodified neurotensin peaks were observed indicating an incomplete reaction. For cross-linking in PBS, a first attempt was made to optimize reaction conditions by varying molar ratios of cross-linker to peptide at 1:2, 2.5:1 and 5:1 and by using reaction times of 45, 90 and 120 min for each reaction ratio. The best relative yield of mono-link products (Figure 4.1B) was obtained at a 2.5:1 molar ratio of cross-linker to peptide for 90 min. Hydrolyzed mono-link product appeared and was observed the most abundant cross-linking products when the reaction was conducted in PBS buffer at the referring conditions but the inter-molecular cross-link still could not

be improved by this strategy. Though unmodified neurotensin peaks predominated in the mass spectra of product mixtures from cross-linking reactions, signals of cross-linked products were strong enough for subsequent fragmentation and unambiguous identification.

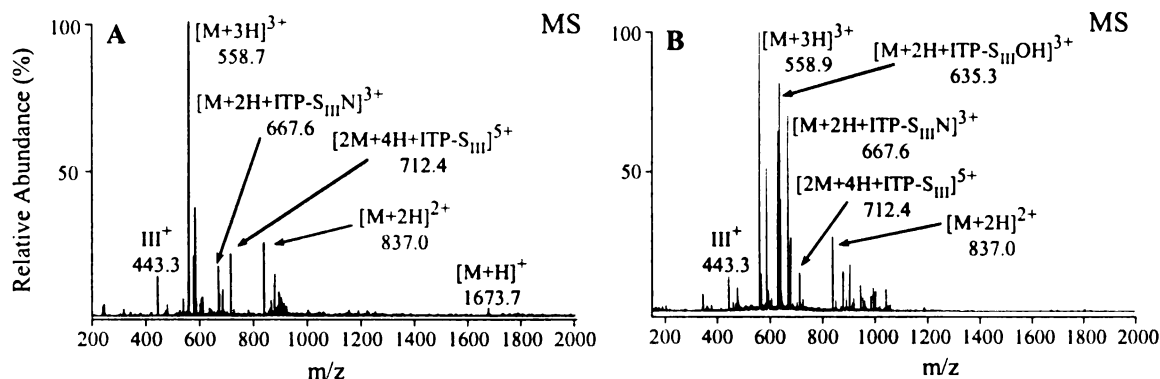


Figure 4.1 Nano-electrospray quadrupole ion trap mass spectra of **III** cross-linked neurotensin. (A) Cross-linking reaction in DMF at 75 min with 1:2 molar ratio of cross-linker to peptide; (B) cross-linking reaction in PBS buffer at 90 min with 2.5:1 molar ratio of cross-linker to peptide. M: unmodified neurotensin, [2M+4H+ITP-S_{III}]⁵⁺ (m/z 712.4): inter-molecular cross-link in 5+ charge [M+2H+ITP-S_{III}N]³⁺ (m/z 667.6): unhydrolyzed mono-link in 3+ charge state, [M+2H+ITP-S_{III}OH]³⁺ (m/z 635.3): hydrolyzed mono-link in 3+ charge state

4.2.2 Inter-Molecular Cross-Linked Neurotensin

The majority of inter-molecular cross-linked neurotensin were present in its 5+ charge state due to two basic residues contained in each peptide chain and one fixed charge on the cross-linker (Figure 4.1). Some +4 charge state was also observed. The dissociation of inter-molecular cross-links was characterized by low energy CID MS/MS. As indicated in Figure 4.2, MS/MS gave simple fragmentation patterns for both charge states due to specific cleavage of the cross-linker. Thus, we can conclude that the

fragmentation of inter-molecular ionic cross-linked neurotensin is not affected by the charge state.

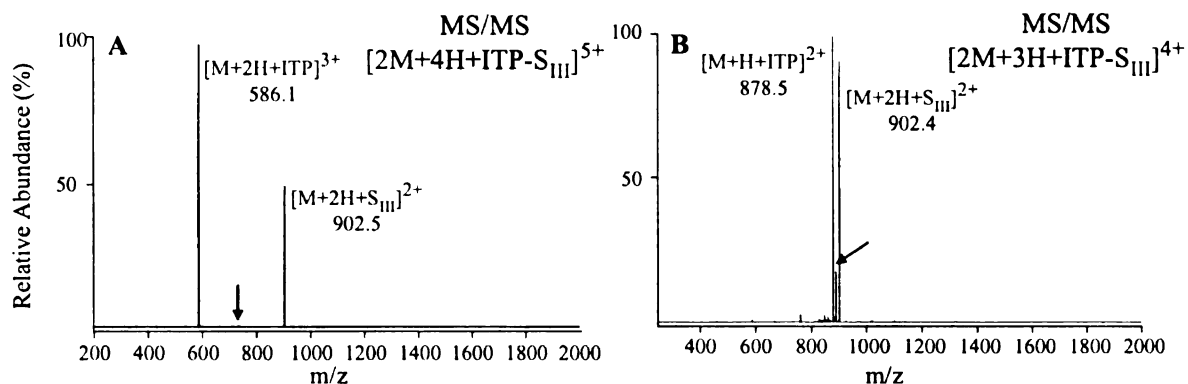


Figure 4.2 Low energy CID MS/MS of inter-molecular cross-linked neurotensin by **III** prepared in DMF. (A) Precursor ion in 5+ charge state (m/z 712.5); (B) precursor ion in 4+ charge state (m/z 890.5). The arrow indicates peak of precursor ions

The two product ions from the +5 charge state MS/MS spectrum were then selected and subjected to MS^3 . As shown in Figure 4.3, the observation of a series of b- and y-type ions provided information on the sequence and modification site for the peptides. In particular, the presence of y_8 ions containing ITP or S tags and the presence of y_7 ions without any modification provided evidence that the modification had occurred at ϵ -amine of Lys 6.

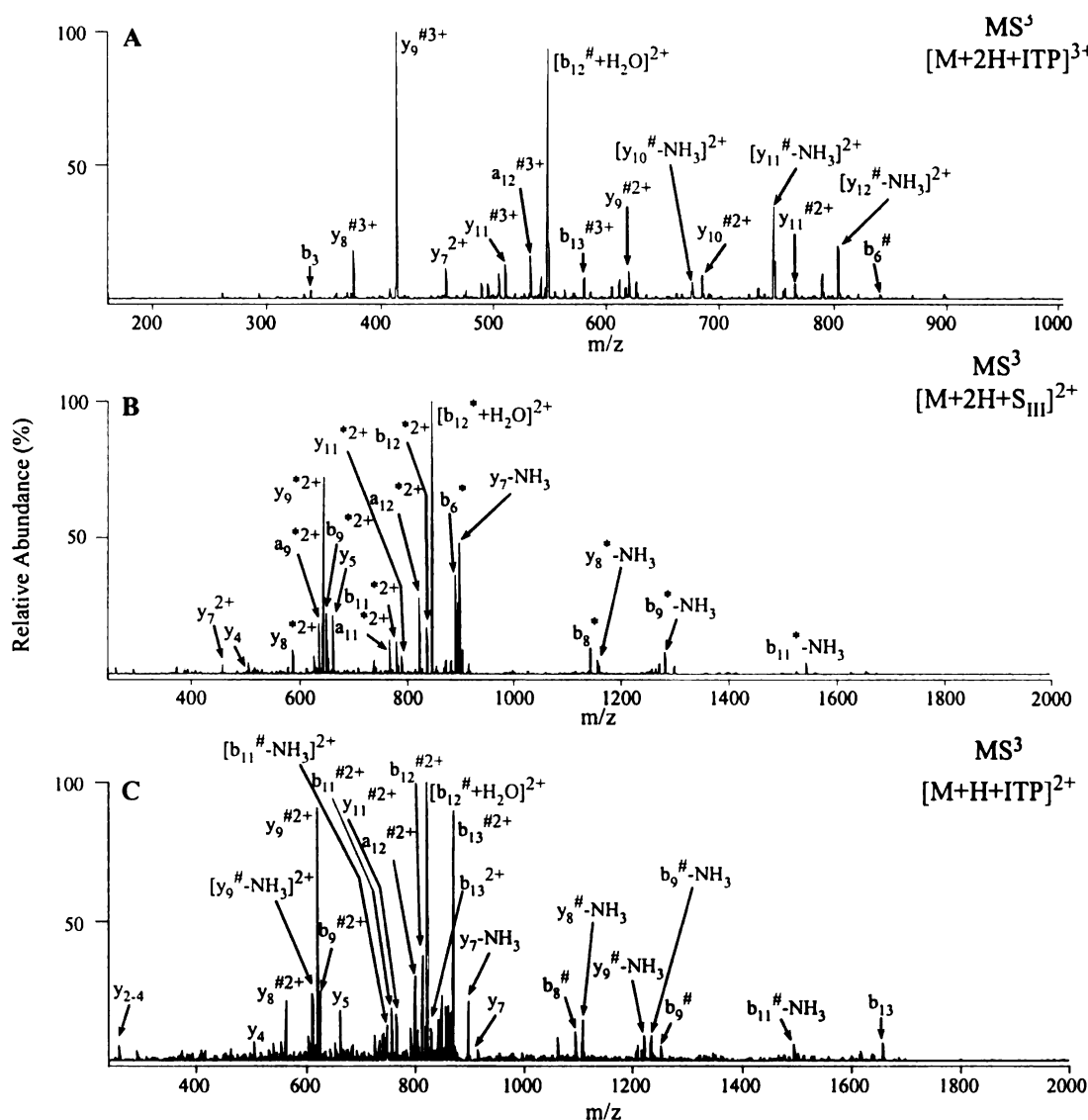


Figure 4.3 CID MS^3 product ion spectra of modified neurotensin in Figure 4.2 formed from inter-molecular cross-links. (A) $[M+2H+ITP]^3+$ ion; (B) $[M+2H+S_{III}]^{2+}$ ion; (C) $[M+H+ITP]^{2+}$ ion. ITP modification is labeled as #, S_{III} modification is labeled as *

4.2.3 Mono-Linked Neurotensin

The majority of cross-linked products existed in mono-link types when cross-linking was performed in PBS buffer (Figure 4.1B). Consistent with the prediction shown in Scheme 4.1B, a product ion containing an ITP tag is the dominant fragment from both unhydrolyzed and hydrolyzed mono-links in 3+ charge state (Figure 4.4). It

appears that the carbonyl group of cross-linker-lysine amide bond on the mono-link is a much stronger nucleophile than NHS-ester carbonyl or free carboxylic acid under the experimental conditions, thus much more reactive to initiate the fragmentation reactions. However another pathway involving nucleophilic attack from the carboxylic acid carbonyl was observed as a minor pathway, as evidenced by a minor S modified product ion formed by fragmentation of the hydrolyzed mono-link (Figure 4.4B).

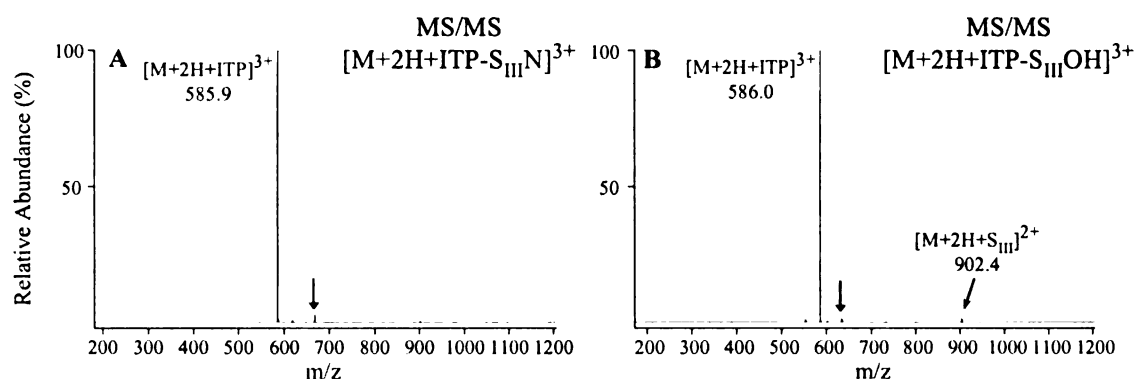


Figure 4.4 Low energy CID MS/MS spectra of mono-linked neurotensin by **III** prepared in PBS buffer. (A) Unhydrolyzed mono-link, m/z 667.7, $z = 3$; (B) hydrolyzed mono-link, m/z 635.3, $z = 3$. The unassigned arrow indicates peaks of precursor ions

Distinct fragmentation behavior was detected in CID MS/MS of the doubly charged hydrolyzed mono-link (Figure 4.5) where an S tagged neurotensin product ion dominated the spectrum. Thus, it appears that charge states can have some influence in the fragmentation behavior of mono-linked neurotensin. However, no matter which C-S bond cleaves, it is still the energetically favorable mechanism and takes place prior to peptide backbone cleavage.

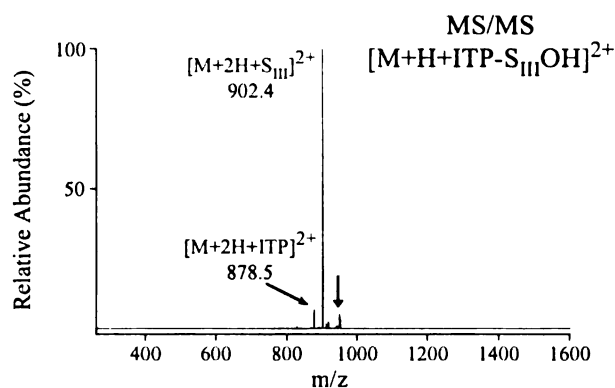


Figure 4.5 Low energy CID MS/MS spectra of hydrolyzed mono-link in 2+ charge state prepared by reaction of neurotensin with **III** in PBS buffer. The unassigned arrow indicates peak of precursor ions, m/z 952.3, $z = 2$

The product ion spectra obtained by MS³ dissociation of the doubly and triply charged ITP modified ions, and the doubly charged S modified ions in Figure 4.4 and 4.5 were found to yield the similar product ion patterns as those observed in Figure 4.3 from the inter-molecular cross-link (data not shown).

The cross-linker **III** has been successfully applied to cross-linking reactions with neurotensin in both DMF and PBS medium. The experimental capability and methodology was demonstrated when inter-molecular cross-link and mono-linked peptides were unambiguously identified from low energy CID MS/MS and through MS³. Simple MS/MS patterns make the identification of cross-link types straightforward and the modification attached on the peptide product ions do not affect the ability to generate b- and y-type ions for further characterization of the peptide sequence and modification sites. The capability of this novel cross-linking reagent will be discussed more in following sections.

4.3 Cross-Linking Reaction with Insulin-Like Growth Factor I (57-70)

Insulin-like growth factor I (57-70) (IGF, ALLETYCATPAKSE, MW 1495.7), contains two primary amine functional groups: one is on the lysine side chain at position 12 and the other is on the N terminus. With two reactive groups available toward NHS-esters, three cross-linked products might be observed from a cross-linking reaction. Cross-linking of IGF was performed in PBS buffer at 2.5:1 molar ratios of cross-linker (15 mM, 4 μ L in DMF) and peptide (0.6 mM, 40 μ L in PBS buffer) for 90 min at room temperature. The majority of cross-linking led to intra-molecular cross-link type, although a minor hydrolyzed mono-link was observed and no significant inter-molecular product (Figure 4.6).

When two reactive amine groups are present in the same peptide chain, there are some questions should be considered before further investigation of MS/MS and MS³ behavior of cross-linked products. One question is whether two amine groups have same possibility to form a mono-link or involved in an inter-molecular cross-link? A related problem might be that for an intra-molecular cross-link, the cleavage of two C-S bonds is random or preferential due to the amide bond formed toward different residues or at different places.

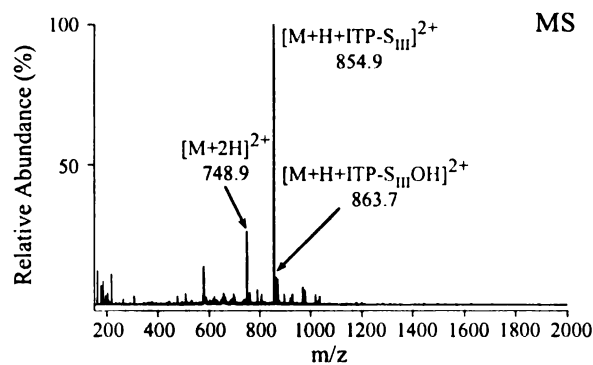


Figure 4.6 Nanoelectrospray quadrupole ion trap mass spectrum of **III** cross-linked IGF. Reaction performed in PBS buffer. M: unmodified IGF; $[M+H+ITP-S_{III}]^{2+}$: intra-molecular cross-link in 2+ charge state; $[M+H+ITP-S_{III}OH]^{2+}$: hydrolyzed mono-link in 2+ charge state

To address these questions, a first attempt was made to characterize the fragmentation of intra-molecular cross-linked IGF. Consistent with the prediction shown in Scheme 4.1A, direct b- and y-type product ions were generated upon CID MS/MS (Figure 4.7). For an intra-molecular cross-linking occurring at the N-terminus and Lys 12, we could expect that all the resultant b-type ions are modified with an ITP or S tag while y-type ions starting from y_3 should be modified as well. And besides precursor ion, only the b-type ions starting from b_{12} could contain both tags. However, a discrepancy was observed between the theoretical prediction and the experimental results. In particular, y_3 , y_5 and y_6 ions were observed without modification. Otherwise b_7 , b_8 , b_9 , b_{11} and y_8 ions were identified containing both ITP and S tags. Based on those results, it appears that some cross-linking had taken place at residue 7, i.e., cysteine. The reactivity of NHS-esters toward sulphydryl functional group was seldom reported before; further evidence is required to open up this new channel for the structural analysis of proteins.

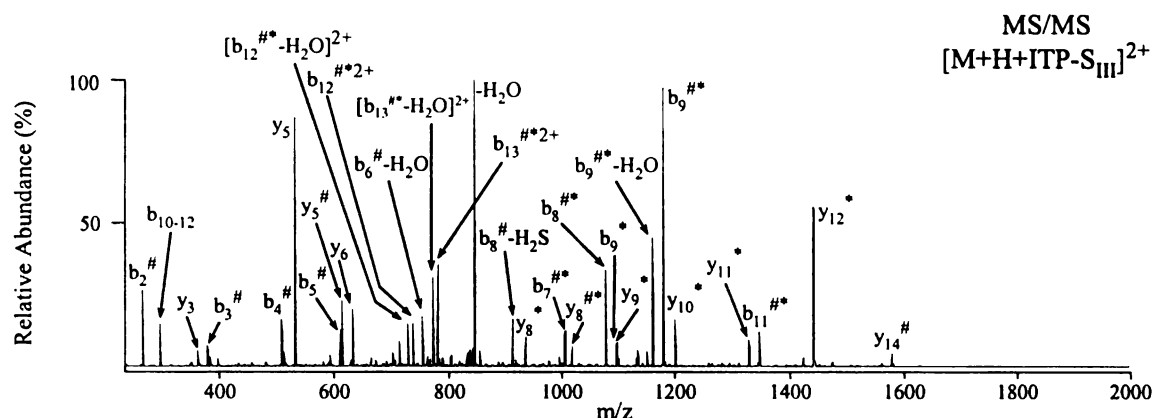


Figure 4.7 CID MS/MS spectrum of doubly charged intra-molecular cross-linked IGF. ITP modification is labeled as #, S_{III} modification is labeled as *

Although C-S bond cleavage is not able to be detected directly from MS/MS spectrum; the generated b- and y-type ions provided us hints regarding the fragmentation mechanism within the linker. Based on the observations that singly tagged b-type ions tended to contain ITP, and y-type ions with single S modification tended to be present, it appears that cleavage didn't occur randomly at each C-S bond but was instead preferred at the N-terminal amide side. The controlling factors responsible for this bias are not clear.

For the IGF hydrolyzed mono-link, similar results to those described above were also obtained by CID MS/MS and MS³. As shown in Figure 4.8A, ITP modified peptide ions were observed as dominant fragment ions from the doubly charged precursor. Based on the experience of mono-linked neurotensin, a minor peak corresponding to S modified peptide ions was also observed. However the appearance of a y₅ ion implies that some peptide backbone cleavage had also occurred during MS/MS. This is probably due to enhanced cleavage at the N-terminal side of proline [40, 59]. Thus we could have confidence by the facts that even when compared to preferential N-terminal proline

fragmentation, desired C-S bond cleavage within the linker region is still the dominant pathway.

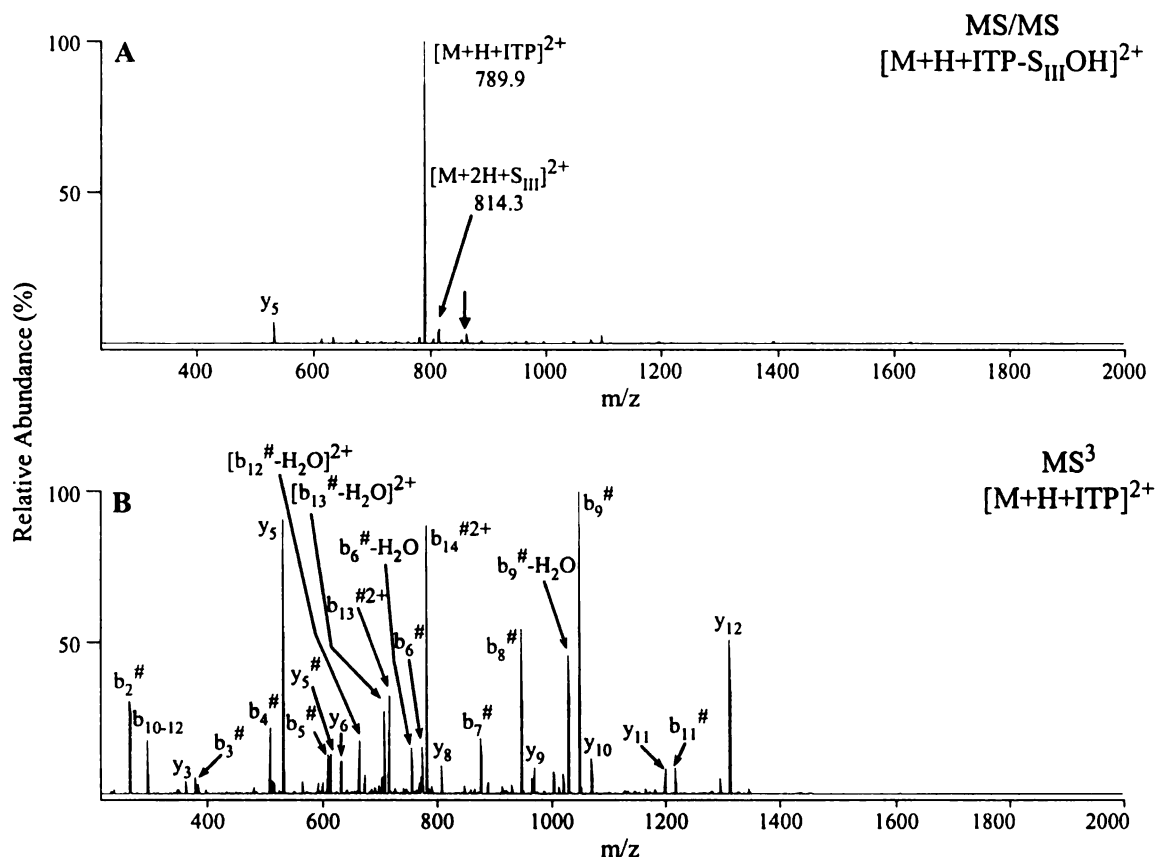


Figure 4.8 Low energy CID MS/MS and MS³ spectra of hydrolyzed mono-linked IGF. (A) CID MS/MS of hydrolyzed mono-linked IGF, unassigned arrow indicates peak of precursor ions (m/z 863.8, z = 2); (B) CID MS³ product ion spectrum of $[M+H+ITP]^{2+}$ ion from hydrolyzed mono-linked IGF. ITP modification is labeled as #

MS³ spectrum obtained by dissociation of doubly charged ITP modified peptide (Figure 4.8 B) was observed to be very similar to that from fragmentation of intra-molecular cross-linked IGF due to the fact that they are at the same proton conditions and the attached ITP or S tag doesn't affect the peptide backbone cleavage. This observation might be utilized to assist in identifying peptide sequence and modification sites by

comparing m/z difference between ITP and S tags, especially when a data base searching program is employed.

Since there are three reactive functional groups capable of cross-linking reactions, one related problem is their relative reactivity toward NHS-esters. The presence of a modified y_5 ion could only come from mono-link at Lys 12; otherwise the absence of modified y_8 to y_{13} ions implies no significant mono-link occurred at Cys 7. On the basis of this observation, a mono-link at N-terminal is responsible for formation of the majority of modified b-ions and unmodified y-ions. A much smaller modified y_5 peak indicates that the N-terminal mono-link was the main modification site.

IGF is an ideal model to characterize the dissociation of intra-molecular cross-link. However the factors responsible for formation of the intra-molecular cross-link as a major product instead of other cross-link types is not clear. A better control of cross-linking reactions needs further investigation into the reaction conditions as well as peptide compositions.

4.4 Cross-Linking Reaction with VTMAHFWNFGK (MWK)

Synthetic peptide VTMAHFWNFGK (MWK, MW 1337) contains two primary amine functional groups: one is on the N terminus and the other on the C terminus lysine side chain. Thus modifications on the N-terminal will be observed from all b-type ions while modifications of the lysine side chain would be detected on all y-type ions. MWK was cross-linked in PBS buffer at room temperature, following the conditions employed for cross-linking of IGF. From a single cross-linking reaction, three types of cross-linked products: intra-molecular, inter-molecular cross-link and mono-link were all identified as

shown in Figure 4.9. Multiple cross-linking reactions were also observed, where two cross-linker molecules had reacted with two primary amine groups on the peptide chain by one of their NHS-esters followed by hydrolysis of the other NHS-ester, resulting in double hydrolyzed mono-link.

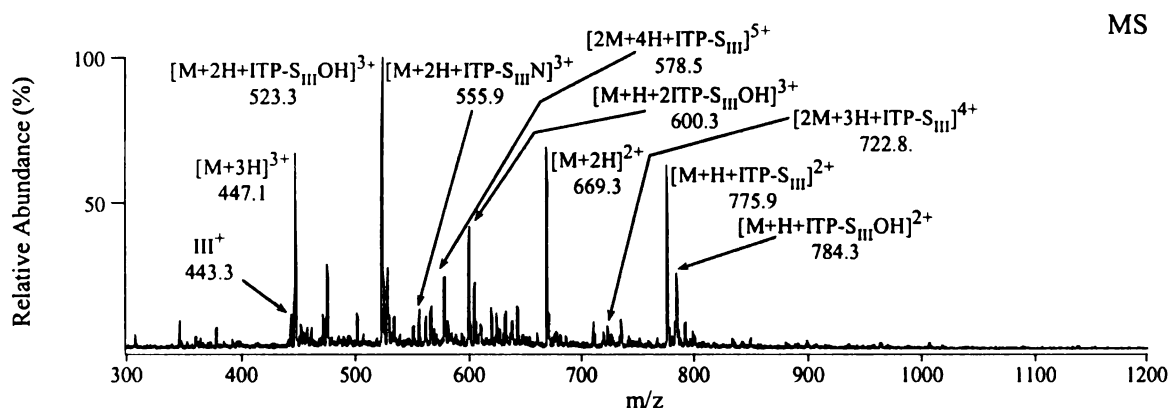


Figure 4.9 Nanoelectrospray quadrupole ion trap mass spectrum of **III** cross-linked MWK. Reaction performed in PBS buffer. M: unmodified MWK; $[\text{M}+2\text{H}+\text{ITP-S}_{\text{III}}\text{OH}]^{3+}$ (m/z 523.3): hydrolyzed mono-link in 3+ charge state; $[\text{M}+2\text{H}+\text{ITP-S}_{\text{III}}\text{N}]^{3+}$ (m/z 555.9): unhydrolyzed mono-link in 3+ charge state; $[2\text{M}+4\text{H}+\text{ITP-S}_{\text{III}}]^{5+}$ (m/z 578.5): inter-molecular cross-link in 5+ charge; $[\text{M}+\text{H}+2\text{ITP-S}_{\text{III}}\text{OH}]^{3+}$ (m/z 600.3): double hydrolyzed mono-link in 3+ charge state; $[2\text{M}+3\text{H}+\text{ITP-S}_{\text{III}}]^{4+}$ (m/z 722.8): inter-molecular cross-link in 4+ charge; $[\text{M}+\text{H}+\text{ITP-S}_{\text{III}}]^{2+}$ (775.9): intra-molecular cross-link in 2+ charge state; $[\text{M}+\text{H}+\text{ITP-S}_{\text{III}}\text{OH}]^{2+}$ (m/z 784.3): hydrolyzed mono-link in 2+ charge state

4.4.1 Intra-Molecular Cross-Linked MWK

Intra-molecular cross-linked MWK is observed dominantly in a 2+ charge state. The majority of mono-links and inter-molecular cross-links were observed in 3+ and 5+ charge states. CID-MS/MS of the intra-molecular cross-linked MWK peptide, shown in Figure 4.10, indicated that ITP and S tags could be attached on the same b- or y-ion, providing evidence that cleavage had occurred at both C-S bonds. The chance of cleavage of either bond was not equal, however, as ITP tags were mainly found on y-ions and S

tags were mainly on b-ions. This implies that the favorable cleavage pathway is associated with nucleophilic attack from the carbonyl group of the lysine-linked amide bond.

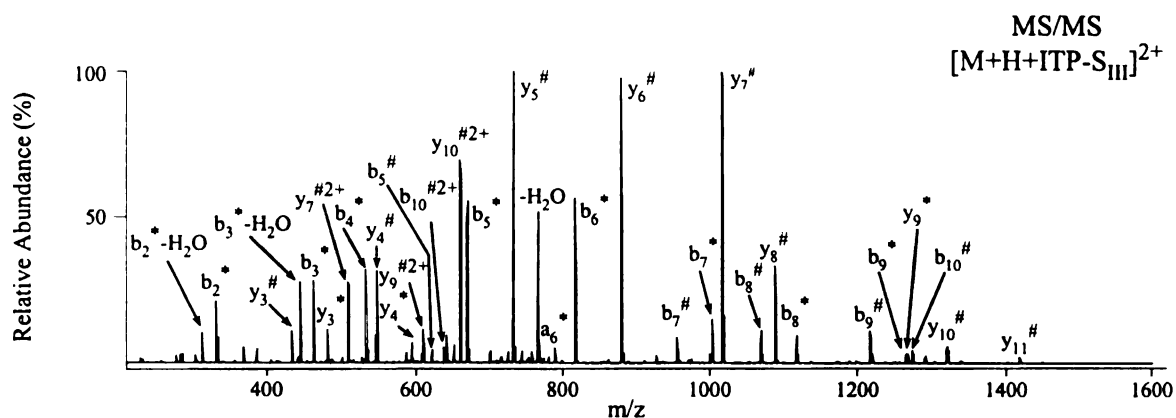


Figure 4.10 CID MS/MS spectrum of intra-molecular cross-linked MWK. ITP modification is labeled as #, S_{III} modification is labeled as *

4.4.2 Mono-Linked MWK

As cross-linking was performed in PBS buffer, the majority of unreacted NHS-ester functional group underwent hydrolysis to yield a hydrolyzed mono-link product. However, the hydrolysis did not go to completion as a small amount of unhydrolyzed mono-link was still observed.

CID MS/MS of the triply (Figure 4.11 A and B) or doubly (Figure 4.11 C and D) charged NHS-ester or hydrolyzed mono-links all resulted in neutral losses of a dialkyl sulfide (either as the methyl(pentanoylhydroxysuccinimide) sulfide or as the methyl(pentanoic acid) sulfide) as the dominant fragmentation pathway, resulting in ITP modified peptide ions.

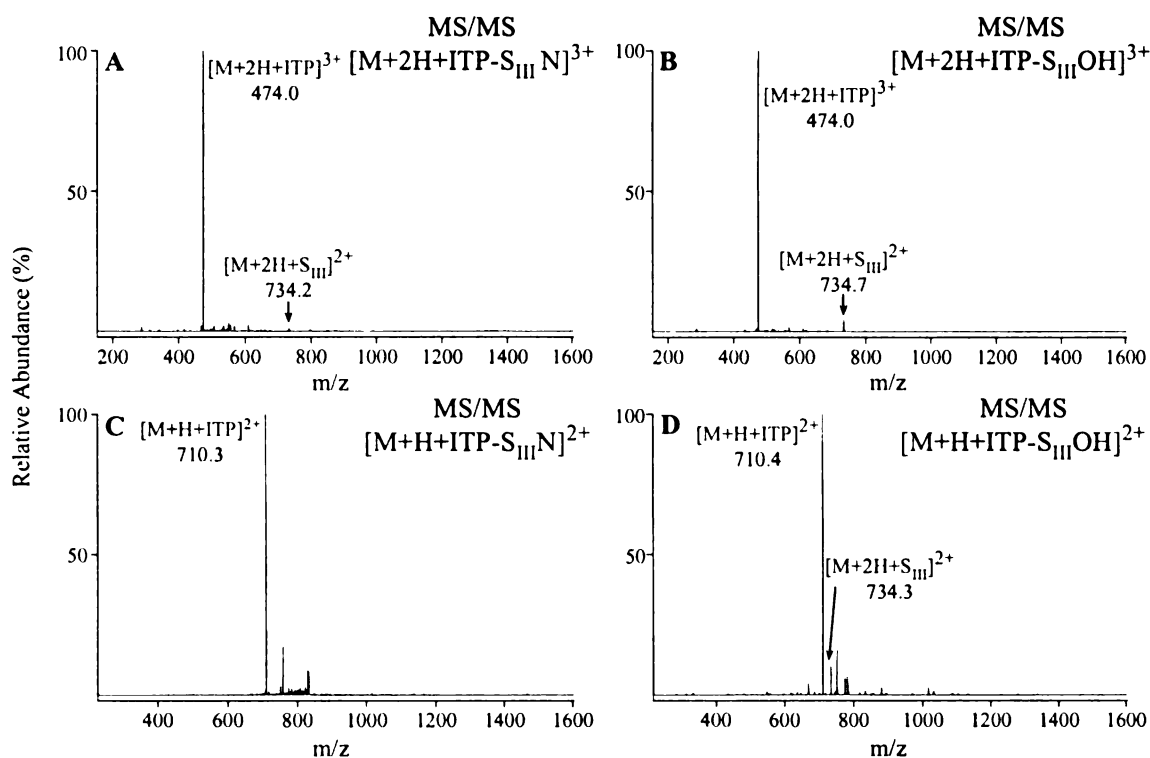


Figure 4.11 Low energy CID MS/MS spectra of mono-linked MWK by **III** prepared in PBS buffer. (A) Unhydrolyzed mono-link in 3+ charge state; (B) hydrolyzed mono-link in 3+ charge state; (C) unhydrolyzed mono-link in 2+ charge state; (B) hydrolyzed mono-link in 2+ charge state

The spectra obtained by MS³ dissociation of the triply and doubly charged ITP modified peptides are shown in Figure 4.12 A and B, respectively. Distinct fragmentation behaviors were observed from these two charge states. It can be seen from Figure 4.12 A that b-ions are dominantly present in their ITP modified forms, while most y-ions are unmodified, indicating that the monolink for the 3+ charge state was primarily from reaction at the N-terminus of the peptide. The data in Figure 4.12 B, however, indicated no significant difference in the relative intensity of the b- and y-type product ions with or without the modification, indicating no preference for formation of the mono-link at either the N-terminal or C-terminal lysine. The results are consistent with a decreased

proton affinity for the lysine residue when the cross-linking reaction had occurred at its amine group.

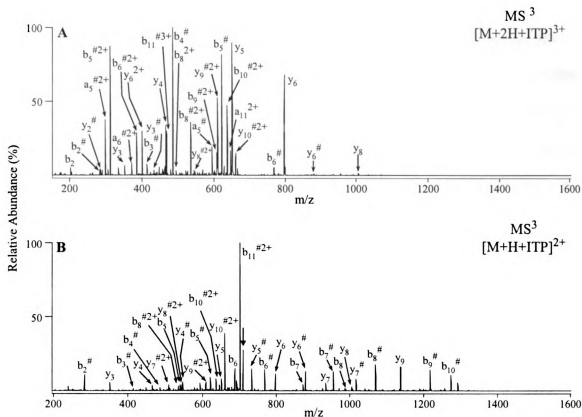


Figure 4.12 CID MS³ product ion spectra of ITP modified MWK in Figure 4.11 formed from hydrolyzed mono-link. (A) [M+2H+ITP]³⁺ ion; (B) [M+H+ITP]²⁺ ion. ITP modification is labeled as #; the unassigned arrow indicates the precursor ion, m/z 710.4, z = 2

4.4.3 Double Hydrolyzed Mono-Link

Although the cross-linking reaction was carried out at a low cross-linker to peptide ratio, a double hydrolyzed mono-link was observed, as shown in Figure 4.9. A peptide which is multiply mono-linked does not provide any information on spatial relationship

between two reactive functional groups, but it could yield important information on the surface accessibility of target functional groups in a protein cross-linking experiment.

A particular feature of the double mono-link, which distinguishes it from single modifications, is its multistage MS/MS behavior. Similar to that described earlier, the neutral loss of dialkylsulphide (5-(methylthio)pentanoic acid) was the dominant MS/MS fragmentation pathway from the 3+ doubly mono-linked peptide precursor ion of MWK. However, further dissociation of the ITP modified product from this reaction resulted in a second loss of 5-(methylthio)pentanoic acid, indicative of the presence of the second mono-link. Thus, the identification of the peptide sequence and modification sites in this case required the use of an MS⁴ experiment (Figure 4.14).

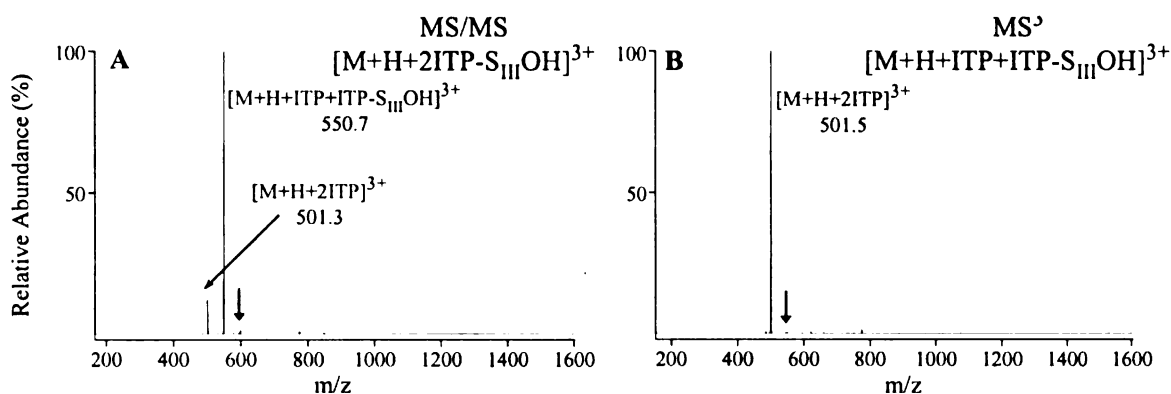


Figure 4.13 Low energy CID MS/MS and MS³ product ion spectra of double hydrolyzed mono-link of MWK. (A) CID MS/MS product ion spectrum of double hydrolyzed mono-link of MWK; (B) CID MS³ product ion spectrum of $[M+H+ITP+ITP-S_{III}OH]^{3+}$ ion formed from double hydrolyzed mono-link in Figure 4.13A. The unassigned arrow indicates the precursor ion: (A) m/z 600.2, z = 3; (B) m/z 550.8, z = 3

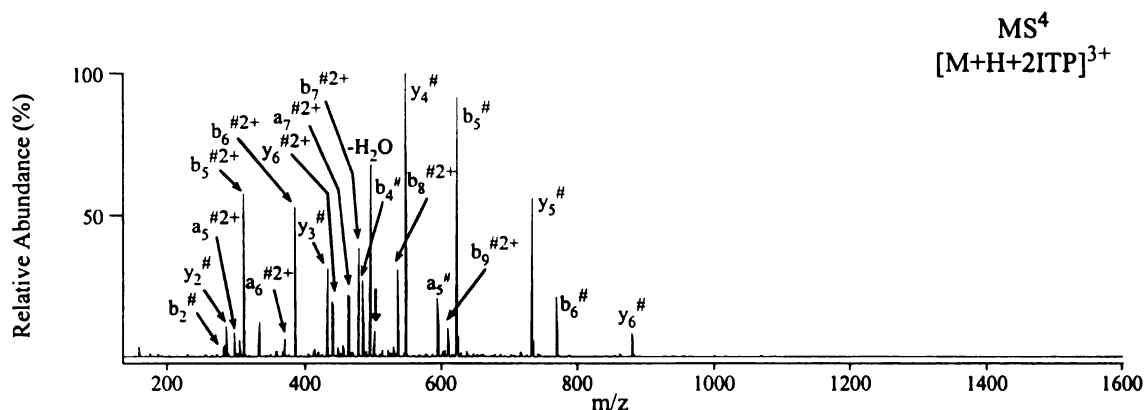


Figure 4.14 CID MS⁴ product ion spectrum of double ITP modified MWK ion formed from double hydrolyzed mono-link in Figure 4.13. ITP modification is labeled as #; the unassigned arrow indicates the precursor ion, m/z 501.5, z = 3

4.4.4 Inter-Molecular Cross-Link

Characterization of the dissociation of inter-molecular cross-linked peptides not only provides us information enabling peptide identification and determination of modification sites, but also allows for investigation into the factors influencing bond cleavage adjacent to the ‘fixed charge’ within the linker. With two reactive amine groups available in MWK for cross-linking, there could be three possible structures of the inter-molecular cross-linked product. Only two product ions will appear upon low energy CID MS/MS of inter-molecular cross-links, corresponding to ITP and S modifications. However, each peak could be representative of peptides modified at two different sites, N-terminal or C-terminal lysine residue, which could be differentiated by two series of b- and y-type ions observed upon MS³.

As described earlier, the cleavage of C-S bond adjacent to the ‘fixed charge’ was found to be independent of the charge state of the cross-linked peptide (Figure 4.15 A and B). Two product ions were generated in each spectrum, containing ITP and S

modifications. Interestingly, MS³ fragmentation of the ITP modified triply charged product ion (Figure 4.16A) resulted in the formation of a different product ion spectrum to that obtained from the mono-linked peptide (Figure 4.12A). For this peptide, more y-type ions were observed containing the ITP tag compared to those from mono-links, indicating a greater extent of lysine side chain modification.

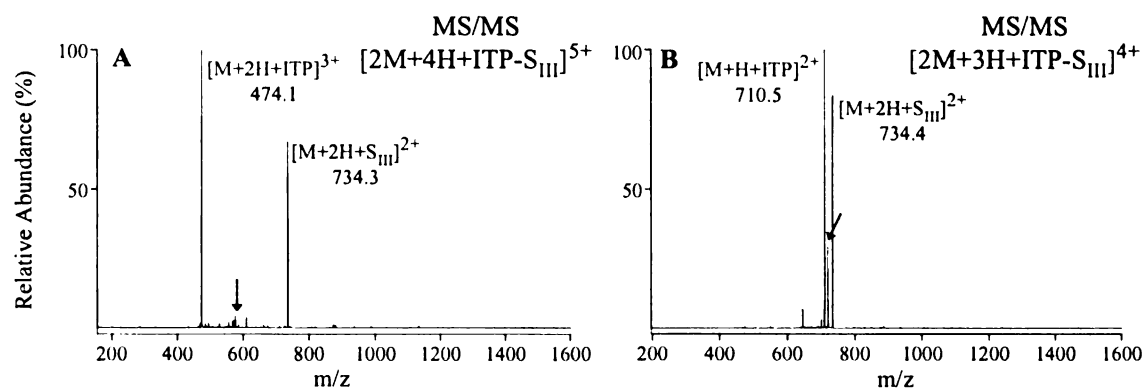


Figure 4.15 Low energy CID MS/MS spectra of inter-molecular cross-linked MWK by **III**. (A) Precursor ion in 5+ charge state, m/z 578.2; (B) precursor ion in 4+ charge state, m/z 722.6. The arrow indicates peak of precursor ions

The product ion spectrum obtained by MS³ dissociation of the complementary S modified peptide ion formed from the inter-molecular cross-link is shown in Figure 4.16B. Large numbers of product ions were generated providing extensive sequence information. Thus the fragment ions generated by MS³ of both ITP and S modified peptides supported the assignment of the peptide sequence.

The inter-molecular cross-link described here involved two identical peptides cross-linked together; whereas proteolysis of cross-linked proteins or protein complexes would usually produce heterogeneous cross-linked peptide dimers. However, MWK still provides us an ideal model to examine the fragmentation behavior of different cross-

linked products and different charge states for the identification of peptide sequence and modification sites.

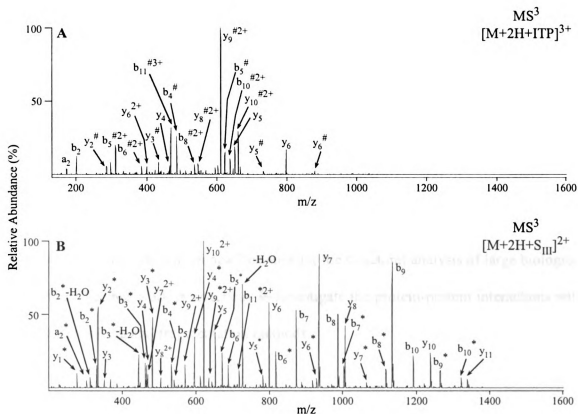


Figure 4.16 CID MS³ product ion spectra of modified MWK in Figure 4.15A formed from inter-molecular cross-links. (A) [M+2H+ITP]³⁺ ion; (B) [M+2H+S_{III}]²⁺ ion. ITP modification is labeled as #, S_{III} modification is labeled as *

4.5 Summary

The multistage tandem mass spectrometry of cross-linked peptides presented here illustrate that peptides containing ionic cross-linking reagent **III** could be applied to effectively distinguish between mono-linked, intra-, and inter-molecular cross-links by their distinct fragmentation patterns. Upon low energy CID MS/MS, the mono-link

experiences a neutral loss giving rise to an intact peptide chain modified with an ITP tag. MS/MS of inter-molecular cross-linked peptides gives rise to two separate peptide chains attached with ITP or S tags. From further fragmentation events, the peptide sequence and modification site(s) could be unambiguously identified. Sequence ions are directly generated from CID MS/MS of an intra-molecular cross-link, by searching for specific mass additions corresponding to an ITP or S tag on the b- and y-type product ions. The dominant cleavage of the C-S bond adjacent to the 'fixed charge' within the linker region is experimentally demonstrated to be independent of the charge state of the cross-linked peptide. The solubility and improved stability under aqueous conditions make this symmetrical ionic cross-linking reagent feasible to be applied to large scale proteins and protein complexes, opening up new horizons for the structural analysis of large biological assemblies. Further work is underway to investigate the protein-protein interactions with these cross-linking and fragmentation methods.

REFERENCES

1. Horton H.R.; Moran L.A.; Ochs R.S.; Rawn J.D.; Scrimgeour K.G. Principles of Biochemistry, 3rd edition, Pearson Education, Inc., Prentice Hall, **2002**, Page 51.
2. Trakselis, M.A.; Alley, S.C.; Ishmael, F.T. Identification and mapping of protein-protein interactions by a combination of cross-linking, cleavage, and proteomics. *Bioconj. Chem.* **2005**, *16*, 741-750.
3. Bader, G.D.; Hogue, C.W.V. Analyzing yeast protein-protein interaction data obtained from different sources. *Nature Biotechnol.* **2002**, *20*, 991-997.
4. Sinz, A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* **2006**, *25*, 663-682.
5. Young, M.M.; Tang, N.; Hempel, J.C.; Oshiro, C.M.; Taylor, E.W.; Kuntz, I.D.; Gibson, B.W.; Dollinger, D. High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5802-5806.
6. Chakravarti, B.; Lewis, S.J.; Chakravarti, D.N.; Raval, A. Three dimensional structures of proteins and protein complexes from chemical cross-linking and mass spectrometry: a biochemical and computational overview. *Current Proteomics* **2006**, *3*, 1-21.
7. Cross-Linking Reagents Technical Handbook, Pierce, **2005**, Page 3.
8. Alley, S.C.; Ishmael, F.T.; Jones, A.D.; Benkovic, S.J. Mapping protein-protein interactions in the bacteriophage T4 DNA polymerase holoenzyme using a novel trifunctional photo-cross-linking and affinity reagent. *J. Am. Chem. Soc.* **2000**, *122*, 6126-6127.
9. Trester-Zedlitz, M.; Kamada, K.; Burley, S.K.; Fenyo, D.; Chait, B.T.; Muir, T.W. A modular cross-linking approach for exploring protein interactions. *J. Am. Chem. Soc.* **2003**, *125*, 2416-2425.

10. Sinz, A. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *J. Mass Spectrom.* **2003**, *38*, 1225-1237.
11. Schilling, B.; Row, R.H.; Gibson, B.W.; Guo, X.; Young, M.M. MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 834-850.
12. Sinz, A.; Kalkhof, S.; Ihling, C. Mapping protein interfaces by a trifunctional cross-linker combined with MALDI-TOF and ESI-FTICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1921-1931.
13. Ahrends, R.; Kosinski, J.; Kirsch, D.; Manelyte, L.; Giron-Monzon, L.; Hummerich, L.; Schulz, O.; Spengler, B.; Friedhoff, P. Identifying an interaction site between MutH and the C-terminal domain of MutL by crosslinking, affinity purification, chemical coding and mass spectrometry. *Nucl. Acids Res.* **2006**, *34*, 3169-3180.
14. Chowdhury, S.M.; Munske, G.R.; Tang, X.; Bruce, J.E. Collisionally activated dissociation and electron capture dissociation of several mass spectrometry-identifiable chemical cross-linkers. *Anal. Chem.* **2006**, *78*, 8183-8193.
15. Reynolds, K.J.; Yao, X.; Fenselau, C. Proteolytic ^{18}O labelling for comparative proteomics: evaluation of endoprotease Glu-C as the catalytic agent. *J. Proteome Res.* **2002**, *1*, 27-33.
16. Back, J.W.; Notenboom, V.; de Koning, L.J.; Muijsers, A.O.; Sixma, T.K.; de Koster, C.G.; de Jong, L. Identification of cross-linked peptides for protein interaction studies using mass spectrometry and ^{18}O labeling. *Anal. Chem.* **2002**, *74*, 4417-4422.
17. Chen, X.; Chen, Y.H.; Anderson, V.E. Protein cross-links: universal isolation and characterization by isotopic derivatization and electrospray ionization mass spectrometry. *Anal. Biochem.* **1999**, *273*, 192-203.
18. Taverner, T.; Hall, N.E.; OHair, R.A.J.; Simpson, R.J. Characterization of an antagonist interleukin-6 dimer by stable isotope labeling, cross-linking, and mass spectrometry. *Biol. Chem.* **2002**, *277*, 46487-46492.

19. Müller, D.R.; Schindler, P.; Towbin, H.; Wirth, U.; Voshol, H.; Hoving, S.; Steinmetz, M.O. Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal. Chem.* **2001**, *73*, 1927-1934.
20. Pearson, K.M.; Pannell, L.K.; Fales, H.M. Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 149-159.
21. Seebacher, J.; Mallick, P.; Zhang, N.; Eddes, J.S.; Aebersold, R.; Gelb, M.H. Protein cross-linking analysis using mass spectrometry, isotope-coded cross-linkers, and integrated computational data processing. *J. Proteome Res.* **2006**, *5*, 2270-2282.
22. Petrotchenko, E.V.; Olkhovik, V.K.; Borchers, C.H. Coded cleavable cross-linker for studying protein-protein interaction and protein complexes. *Mol. Cell Proteomics* **2005**, *4*, 1167-1179.
23. Bennett, K.L.; Kussmann, M.; Björk, P.; Godzwon, M.; Mikkelsen, M.; Sørensen, P.; Roepstorff, P. Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping—A novel approach to assess intermolecular protein contacts. *Protein Sci.* **2000**, *9*, 1503-1518.
24. Back, J.W.; Sanz, M.A.; de Jong, L.; de Koning, L.J.; Nijtmans, L.G.J.; de Koster, C.G.; Grivell, L.A.; van der Spek, H.; Muijsers, A.O. A structure for the yeast prohibitin complex: Structure prediction and evidence from chemical crosslinking and mass spectrometry. *Prot. Sci.* **2002**, *11*, 2471-2478.
25. Peterson, J.J.; Young, M.M.; Takemoto, L.J. Probing α -crystallin structure using chemical cross-linkers and mass spectrometry. *Molecul. Vision* **2004**, *10*, 857-866.
26. Back, J.W.; Hartog, A.F.; Dekker, H.L.; Muijsers, Anton O.; de Koning, L.J.; de Jong, Luitzen A new cross-linker for mass spectrometric analysis of the quaternary structure of protein complexes. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 222-227.
27. Soderblom, E.J.; Goshe, M.B. Collision-induced dissociative chemical cross-linking reagents and methodology: Applications to protein structural characterization using tandem mass spectrometry analysis. *Anal. Chem.* **2006**, *78*, 8059-8068.

28. Tang, X.; Munske, G.R.; Siems, W.F.; Bruce, J.E. Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions. *Anal. Chem.* **2005**, *77*, 311-318.
29. Hoffmann, E.D.; Stroobant, V. Mass Spectrometry Principles and Applications, 2nd edition, John Wiley and Sons, New York, Page 136.
30. Henzel, W.J.; Billeci, T.M.; Stults, J.T.; Wong, S.C.; Grimley, C.; Watanabe, C. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *195*, 5011-5015.
31. Eng, J.K.; McCormack, A.L.; Yates, J.R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976-989.
32. Mann, M.; Wilm, M. Error-Tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* **1994**, *66*, 4390-4399.
33. Johnson, R.S.; Martin, S.A.; Biemann, K.; Stults, J.T.; Watson, J.T. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal. Chem.* **1987**, *59*, 2621-2625.
34. Hunt, D.F.; Yates, J.R.; Shabanowitz, J.; Winston, S.; Hauer, C.R. Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 6233-6237.
35. Simpson, R.J.; Connolly, L.M.; Eddes, J.S.; Pereira, J.J.; Moritz, R.L.; Reid, G.E. Proteomic analysis of the human colon carcinoma cell line (LIM 1215): Development of a membrane protein database. *Electrophoresis* **2000**, *21*, 1707-1732.
36. Chait, B.T.; Kent, S.B.H. Weighing naked proteins: Practical, high-accuracy mass measurement of peptides and proteins. *Science* **1992**, *257*, 1885-1894.
37. Lee, S.-W.; Berger, S.J.; Martinovic, S.; Pasa-Tolic, L.; Anderson, G.A.; Shen, Y.; Zhao, R.; Smith, R.D. Direct mass spectrometric analysis of intact proteins of the yeast

large ribosomal subunit using capillary LC/FTICR. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5942-5947.

38. Biemann, K.; Papayannopoulos, I.A. Amino acid sequencing of proteins. *Acc. Chem. Res.* **1994**, *27*, 370-378.

39. Dongre, R.A.; Jones, J.L.; Somogyi, A.; Wysocki, V.H. Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: evidence for the mobile proton model. *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.

40. Kapp E.A.; Schütz, F.; Reid, G.E.; Eddes, J.S.; Moritz, R.L.; O'Hair, R.A.J.; Speed, T.P.; Simpson, R.J. Mining a tandem mass spectrometry database to determine the trends and global factors influencing peptide fragmentation. *Anal. Chem.* **2003**, *75*, 6251-6264.

41. Reid, G. E.; Roberts, K. D.; Kapp, E. A.; Simpson, R. J. Statistical and mechanistic approaches to understanding the gas-phase fragmentation behavior of methionine sulfoxide containing peptides. *J. Proteome Res.* **2004**, *3*, 751-759.

42. Lioe, H.; O'Hair, R.A.J. A novel salt bridge mechanism highlights the need for nonmobile proton conditions to promote disulfide bond cleavage in protonated peptides under low-energy collisional activation. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1109-1123.

43. O'Hair, R.A.J. The role of nucleophile-electrophile interactions in the unimolecular and bimolecular gas-phase ion chemistry of peptides and related systems. *J. Mass Spectrom.* **2000**, *35*, 1377-1381.

44. Schlosser, A.; Lehmann, W.D. Five-membered ring formation in unimolecular reactions of peptides: a key structural element controlling low-energy collision-induced dissociation of peptides. *J. Mass Spectrom.* **2000**, *35*, 1382-1390

45. Roberts, K.D.; Reid, G.E. Leaving group effects on the selectivity of the gas-phase fragmentation reactions of side chain fixed-charge-containing peptide ions. *J. Mass Spectrom.* **2007**, *42*, 187-198.

46. Reid, G.E.; Roberts, K.D.; Simpson, R.J., O'Hair, R.A.J. Selective identification and quantitative analysis of methionine containing peptides by charge derivatization and tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1131-1150.

47. Amunugama, M.; Roberts, K.D.; Reid, G.E. Mechanisms for the selective gas-phase fragmentation reactions of methionine side chain fixed charge sulfonium ion containing peptides. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1631-1642.
48. Sierakowski, J.; Amunugama, M.; Roberts, K.D.; Reid, G.E. Substituent effects on the gas-phase fragmentation reactions of sulfonium ion containing peptides. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1230-1238.
49. Rabinovich, M.S.; Levimov, M.M.; Kulikova, G.N.; Yakushina, L.M.; Verkhovtseva, T.P.; Meller, F.M. An investigation in the field of the synthesis of precursors and fragments of antibiotics VII. Carboxy derivatives of mercaptoacetic acid. *Zhurnal Obshchei Khimii* **1962**, *32*, 1167-1172.
50. Anderson, G.W.; Zimmerman, J.E.; Callahan, F.M. The use of esters of N-hydroxysuccinimide in peptide synthesis. *Metab. Clin. Exp.* **1964**, *13*, 1026-1031.
51. Zhang, J.; Jing, B.; Tokutake, N.; Regen, S.L. Transbilayer complementarity of phospholipids. A look beyond the fluid mosaic model. *J. Am. Chem. Soc.* **2004** *126*, 10856-10857.
52. Numata, M.; Koumoto, K.; Mizu, M.; Sakurai, K.; Shinkai, S. Parallel vs. anti-parallel orientation in a curdlan/oligo(dA) complex as estimated by a FRET technique. *Org. Biomol. Chem.* **2005**, *3*, 2255-2261.
53. Jessing, M.; Brandt, M.; Jensen, K.J.; Christensen, J.B.; Boas, U. Thiophene backbone amide linkers, a new class of easily prepared and highly acid-labile linkers for solid-phase synthesis. *J. Org. Chem.* **2006**, *71*, 6734-6741.
54. Oae, S., Organic Chemistry of Sulfur. Plenum Press, New York, **1977**, page 485.
55. Anderson, G.W.; Zimmerman, J.E.; Callahan, F.M., The use of activated esters in peptides synthesis. *Metabol. Clin. Exp.* **1964**, *13*, 1026-1031.
56. Bragg, P.D.; Hou, C., Subunit composition, function, and spatial arrangement in the (calcium-magnesium ion)-activated adenosinetriphosphatases of Escherichia coli and Salmonella typhimurium. *Arch. Biochem. Biophys.* **1975**, *167*, 311-321.

57. Lomant, A.J.; Fairbanks, G., Chemical probes of extended biological structures: synthesis and properties of the cleavable protein cross-linking reagent [35S]dithiobis(succinimidyl propionate). *J. Mol. Biol.* **1976**, *104*, 243-261.
58. Pearson, R.G.; Sobel, H.; Songstad, J., Nucleophilic reactivity constants toward methyl iodide and *trans*-[Pt(py)₂Cl₂]. *J. Am. Chem. Soc.* **1968**, *90*, 319-326.
59. Tang, X.J.; Thibault, P.; Boyd, R.K. Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation. *Anal. Chem.* **1993**, *65*, 2824-2834.