



University of Kentucky
UKnowledge

Theses and Dissertations--Pharmacy

College of Pharmacy

2015

Chemoenzymatic Studies to Enhance the Chemical Space of Natural Products

Jhong-Min Chen

University of Kentucky, jch249@g.uky.edu

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

Chen, Jhong-Min, "Chemoenzymatic Studies to Enhance the Chemical Space of Natural Products" (2015). *Theses and Dissertations--Pharmacy*. 48.
https://uknowledge.uky.edu/pharmacy_etds/48

This Doctoral Dissertation is brought to you for free and open access by the College of Pharmacy at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Pharmacy by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Jhong-Min Chen, Student

Dr. Jürgen Rohr, Major Professor

Dr. Jim Pauly, Director of Graduate Studies

CHEMOENZYMATIC STUDIES TO ENHANCE THE CHEMICAL SPACE OF
NATURAL PRODUCTS

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the College of Pharmacy at the University of Kentucky

By

Jhong-Min Chen

Lexington, Kentucky

Director: Dr. Jürgen Rohr, Professor of Pharmaceutical Sciences

Lexington, Kentucky

2015

Copyright © Jhong-Min Chen 2015

ABSTRACT OF DISSERTATION

CHEMOENZYMATIC STUDIES TO ENHANCE THE CHEMICAL SPACE OF NATURAL PRODUCTS

Natural products provide some of the most potent anticancer agents and offer a template for new drug design or improvement with the advantage of an enormous chemical space. The overall goal of this thesis research is to enhance the chemical space of two natural products in order to generate novel drugs with better *in vivo* bioactivities than the original natural products.

Polycarcin V (PV) is a gilvocarcin-type antitumor agent with similar structure and comparable bioactivity with the principle compound of this group, gilvocarcin V (GV). Modest modifications of the polyketide-derived tetracyclic core of GV had been accomplished, but the most challenging part was to modify the sugar moiety. In order to solve this problem, PV was used as an alternative lead-structure for modification because its sugar moiety offered the possibility of enzymatic *O*-methylation. We produced four PV derivatives with different methylation patterns for cytotoxicity assays and provided important structure-activity-relationship information.

Mithramycin (MTM) is the most prominent member of the aureolic acid type anticancer agents. Previous work in our laboratory generated three MTM analogues, MTM SA, MTM SK, and MTM SDK by inactivating the *mtmW* gene. We developed new MTM analogues by coupling many natural and unnatural amino acids to the C-3 side chain of MTM SA *via* chemical semi-synthesis and successfully made some compounds with both

improved bioactivity and *in vivo* tolerance than MTM. Some of them were consequently identified as promising lead-structures against Ewing's sarcoma.

The potential of selectively generating novel MTM analogues led us to focus on a key enzyme in the biosynthetic pathway of mithramycin, MtmC. This protein is a bifunctional enzyme involved in the biosynthesis of TDP-D-olivose and TDP-D-mycarose. We clarified its enzymatic mechanisms by X-ray diffraction of several crystal complexes of MtmC with its biologically relevant ligands. Two more important post-PKS tailoring enzymes involved in the biosynthesis of the MTM side chains, MtmW and MtmGIV, are currently under investigation. This would not only give us insight into this biosynthetic pathway but also pave the way to develop potentially useful MTM analogues by engineered enzymes.

KEYWORDS: gilvocarcin V, polycarcin V, mithramycin, MtmC, chemical space

Jhong-Min Chen

06.14.2015.

CHEMOENZYMATIC STUDIES TO ENHANCE THE CHEMICAL SPACE OF
NATURAL PRODUCTS

By
Jhong-Min Chen

Dr. Jürgen Rohr

Director of Dissertation

Dr. Jim Pauly

Director of Graduate Studies

06.14.2015

Date

ACKNOWLEDGEMENTS

I started my studies at the University of Kentucky in January, 2011, and the following dissertation was possible because of the support from many people.

First, I deeply appreciate my advisor, Dr. Jürgen Rohr, for his academic direction, kind encouragement, unwavering support, and great patience throughout my graduate studies. I also would like to thank my committee members, Dr. David Watt, Dr. Jon Thorson, and Dr. Steven Van Lanen, for their invaluable advice and assistance in my research. I am very thankful for Dr. David Watt's and Dr. Oleg V. Tsodikov's guidance in medicinal chemistry and enzymology, respectively. In addition, I would like to acknowledge the contributions of Dr. David J. Newman, Dr. Markos Leggas, and Dr. Younsoo Bae for their kind assistance in bioassay experiments.

I am also pleased to have worked with so many great postdoctoral researchers in the Department of Pharmaceutical Sciences, including Dr. Amit Kumar Jha, Dr. Caixia Hou, Dr. Daniel Scott, Dr. Guojun Wang, Dr. Jamie Horn, Dr. Khaled A. Shaaban, Dr. Madan K. Kharel, Dr. Pallab Pahari, Dr. Prithiba Mitra, Dr. Shaimaa Salem, Dr. Xia Yu, and Dr. Xiaodong Liu. In addition, my co-workers have made my studying here a wonderful experience, and I am indebted to Dr. Eric Nybo, Dr. Micah D. Shepherd, Dr. Nidhi Tibrewal, Dr. Sandra Barnard-Britson, Dr. Theresa Downey, Dr. Xiuling Chi, Ashley Arlinghaus, Daiani Savi, Matthew McErlean, Stevi Weidenbach, Tyler Bucci, Wenlong Cai, and Zheng Cui.

Last but certainly not least, my amazing family has always been supportive of my life goals. I would like to express my deep appreciation to my parents, Young-Tsong Chern and Yu-Feng Cheng, and my young sister, Chung-Chun Chen. Without their love, all of my achievements would not have been possible.

Thank you very much for everything that you had done for me, my teachers, co-workers, friends, and family.

TABLE OF CONTENTS

Acknowledgements.....	ii
Table of Contents.....	iii
List of Tables.....	vi
List of Figures.....	vii
List of Abbreviations.....	xi
Chapter 1: Background information.....	1
1.1 Natural products.....	1
1.2 <i>Streptomyces</i> and polyketides.....	3
1.3 Biosynthesis of type-II polyketides.....	7
Type-II polyketide synthase.....	7
Polyketide tailoring enzymes.....	10
Deoxysugar biosynthesis.....	11
Glycosylation.....	12
Methylation.....	14
Reduction.....	15
1.4 Gilvocarcin.....	15
1.5 Mithramycin.....	22
1.6 Summary.....	29
1.7 Specific aims.....	30
Chapter 2: Study and Use of Enzymes for Modification of Polycarcin V.....	32
2.1 <i>In vitro</i> modification of the L-rhamnosyl moiety of polycarcin V.....	32
2.2 Experimental design.....	32
2.3 Results.....	34
<i>In vitro</i> reactions with auxiliary <i>O</i> -methyltransferases.....	34
Structure elucidation of the polycarcin V derivatives.....	36
Cytotoxicity assays.....	41
2.4 Discussion.....	43
2.5 Materials and methods.....	44
General experimental conditions.....	44

Bacterial strains, culture conditions, and plasmids.....	45
Expression and purification of proteins.....	46
Production and purification of polycarcin V.....	47
<i>In vitro</i> methylation reaction conditions and purification of the products.....	48
Cytotoxicity assays.....	49
Chapter 3: Chemical Derivatization of Mithramycin.....	50
3.1 Modification of mithramycin SA with primary amines.....	50
3.2 Experimental design.....	51
3.3 Results.....	53
Modification of mithramycin SA with primary amines.....	53
Cytotoxicity assays.....	54
Structure elucidation of the mithramycin analogues.....	61
3.4 Discussion.....	79
MTM SA (69) and MTM SA methyl ester (72).....	79
MTM SA-Phe (83), MTM SA-Trp (85), and MTM SA-Tyr (87).....	79
MTM SA-His (81) and MTM SA-Lys-2PGs (82).....	80
MTM SA-Ala (78), MTM SA-Gly (80), and MTM SA-Val (88).....	80
Summary.....	81
3.5 Materials and methods.....	81
General materials.....	81
Biosynthesis of MTM SA.....	82
Chemical reaction conditions and purification of the products.....	83
Cytotoxicity assays.....	84
Chapter 4: Structural investigation into MtmC.....	85
4.1 Structural insight into MtmC.....	85
4.2 Experimental design.....	86
4.3 Results.....	86
Homologues of MtmC.....	86
Overview of crystal structures of MtmC with its substrates and co-factors.....	91
MtmC-SAM and MtmC-SAH complexes.....	93

MtmC-SAH-TDP-KOL, MtmC-SAM-TDP, and MtmC-TDP structures.....	93
Methyltransferase activity of wild-type and mutant MtmCs.....	97
4.4 Discussion.....	98
4.5 Materials and methods.....	101
General materials.....	101
Protein expression and purification.....	102
Crystallization, data collection, and crystal structure determination.....	103
Site-directed mutagenesis of MtmC.....	104
Methyltransferase activity assays.....	105
Chapter 5: Summary and future directions.....	108
5.1 Specific aim 1: Study and use of enzymes for modification of polycarcin V.....	108
5.2 Specific aim 2: Chemical derivatization of mithramycin.....	109
5.3 Specific aim 3: Structural investigation into MtmC.....	110
5.4 Conclusions.....	111
Appendices.....	112
References.....	164
Vita.....	172

LIST OF TABLES

Table 1. ¹ H NMR data (500 MHz) for 60 , 61 , 63 , 64	37
Table 2. ¹³ C and gHMBC data (500 MHz) for 60 , 61 , 64	39
Table 3. Cytotoxic activities (GI ₅₀) of various gilvocarcin/polycarcin derivatives in comparison with two established anticancer drugs against human lung and colon cancer cell lines.....	42
Table 4. The IC ₅₀ values of MTM analogues against the A549 cell line <i>in vitro</i>	55
Table 5-1. The cell growth percent (%) data of NCI 60 cell line screen (one-dose) of MTM and its analogues.....	57
Table 5-2. The cell growth percent (%) data of NCI 60 cell line screen (one-dose) of MTM and its analogues.....	59
Table 6. HR-MS data of bioactive MTM analogues.....	62
Table 7. NMR data (500 MHz) of MTM SA methyl ester (72).....	63
Table 8. NMR data (500 MHz) of MTM SA-Ala (78).....	65
Table 9. NMR data (500 MHz) of MTM SA-Gly (80).....	67
Table 10. NMR data (500 MHz) of MTM SA-Lys-2PGs (82).....	69
Table 11. NMR data (500 MHz) of MTM SA-Phe (83).....	71
Table 12. NMR data (500 MHz) of MTM SA-Trp (85).....	73
Table 13. NMR data (500 MHz) of MTM SA-5''-Br-Trp (86).....	75
Table 14. NMR data (500 MHz) of MTM SA-Val (88).....	77
Table 15. MtmC and its homologues.....	88
Table 16. X-ray diffraction data collection and structural refinement statistics.....	107

LIST OF FIGURES

Figure 1. Six categories of natural products.....	3
Figure 2. Examples of drugs biosynthesized <i>via</i> three basic classes of polyketides.....	5
Figure 3. Three basic polyketide synthases (PKS) (a) Type I, (b) Type II, (c) Type III.....	6
Figure 4. Compounds generated from hybrid biosynthetic pathways.	7
Figure 5. The mechanism of the fatty acyl synthase reaction.....	8
Figure 6. The type II PKS is usually an enzyme complex composed of KS α , KS β , ACP, and other proteins that can modify the backbone.....	10
Figure 7. Biosynthetic pathway of deoxysugars.....	12
Figure 8. Two stereochemical types of glycosylation reaction.....	14
Figure 9. The mechanism of methylation catalyzed by methyltransferases.....	15
Figure 10. Diversity of gilvocarcin-type anticancer drugs.....	16
Figure 11. Two proposed mechanisms of enzymatic C-glycosylation.....	18
Figure 12. Proposed <i>in vivo</i> mechanism of gilvocarcin V resulting from binding to DNA (red) and histone H3 (blue).....	19
Figure 13. Proposed biosynthetic pathway of gilvocarcins (2 , 25 , and 26).....	21
Figure 14. Proposed biosynthetic pathway of TDP-D-fucofuranose in <i>E. coli</i> O52 (blue) and <i>S. griseoflavus</i> (red).....	22
Figure 15. Primary members of the aureolic acid family.....	23
Figure 16. The proposed structure of a Mg ²⁺ -coordinated mithramycin dimer complex...	24
Figure 17. Proposed and simplified biosynthetic pathway of mithramycin.....	27
Figure 18. Proposed biosynthetic pathways of the MTM sugars.....	27
Figure 19. Cooperation between MtmGIV and MtmC during the formation of premithramycin A1 and premithramycin A3.....	29
Figure 20. Gilvocarcin V, polycarcin V, and other natural products with L-rhamnopyranose sugar moiety.....	33
Figure 21. Polycarcin V and its derivatives.....	34
Figure 22. HPLC traces of gilvocarcin V and polycarcins.....	35
Figure 23. SDS-PAGE analysis of purified proteins used in this study.....	47
Figure 24. Generation of novel MTM analogues by inactivation of <i>mtmW</i>	51

Figure 25. The chemical mechanism of the coupling reaction between MTM SA and primary amines.....	52
Figure 26. The chemical mechanism of esterification reaction with TMSCHN ₂	52
Figure 27. MTM analogues derived from MTM SA.....	54
Figure 28. The averages of cell growth percent (%) of MTM and MTM analogues.....	61
Figure 29. Multiple sequence alignment of MtmC and its homologues.....	89
Figure 30. Cartoon representation of crystal structures of MtmC with its biologically relevant ligands.....	92
Figure 31. Active site of the ternary complex, MtmC-SAH-TDP-KOL.....	95
Figure 32. The conformational changes of Y79 residue in (A) MtmC-SAM, (B) MtmC-SAM-TDP, (C) MtmC-TDP, and (D) MtmC-SAH-TDP-KOL complexes.....	97
Figure 33. Relative activity of wild-type MtmC and its mutants (Tyr79Ala and Tyr79Phe), relative to that of the wild-type enzyme.....	98
Figure 34. Model of the MtmC-NADPH-TDP-KOL complex.....	100
Figure 35. MtmC sample before purification by FPLC and the 12 nd - 20 th fractions collected from FPLC experiment.....	103
Figure 36. SDS-PAGE gel of the MtmC Tyr79Ala and MtmC Tyr79Phe mutants after dialysis.....	105
Appendix 1. HR-EI-MS spectrum of 2'- <i>O</i> -methyl-polycarcin V (60).....	112
Appendix 2. ¹ H NMR spectrum of 2'- <i>O</i> -methyl-polycarcin V (60).....	113
Appendix 3. gHSQC NMR spectrum of 2'- <i>O</i> -methyl-polycarcin V (60).....	114
Appendix 4. gHMBC NMR spectrum of 2'- <i>O</i> -methyl-polycarcin V (60).....	115
Appendix 5. HR-EI-MS spectrum of 3'- <i>O</i> -methyl-polycarcin V (61).....	116
Appendix 6. ¹ H NMR spectrum of 3'- <i>O</i> -methyl-polycarcin V (61).....	117
Appendix 7. gHSQC NMR spectrum of 3'- <i>O</i> -methyl-polycarcin V (61).....	118
Appendix 8. gHMBC NMR spectrum of 3'- <i>O</i> -methyl-polycarcin V (61).....	119
Appendix 9-1. HR-ESI-MS spectrum of 2',3'-di- <i>O</i> -methyl-polycarcin V (63).....	120
Appendix 9-2. HR-ESI-MS spectrum of 2',3'-di- <i>O</i> -methyl-polycarcin V (63).....	121
Appendix 10. ¹ H NMR spectrum of 2',3'-di- <i>O</i> -methyl-polycarcin V (63).....	122
Appendix 11. HR-EI-MS spectrum of 3',4'-di- <i>O</i> -methyl-polycarcin V (64).....	123

Appendix 12. ¹ H NMR spectrum of 3',4'-di- <i>O</i> -methyl-polycarcin V (64).....	124
Appendix 13. gHSQC NMR spectrum of 3',4'-di- <i>O</i> -methyl-polycarcin V (64).....	125
Appendix 14. gHMBC NMR spectrum of 3',4'-di- <i>O</i> -methyl-polycarcin V (64).....	126
Appendix 15. HR-ESI-MS spectrum of MTM SA methyl ester (72).....	127
Appendix 16. ¹ H NMR spectrum of MTM SA methyl ester (72).....	128
Appendix 17. gHSQC NMR spectrum of MTM SA methyl ester (72).....	129
Appendix 18. gHMBC NMR spectrum of MTM SA methyl ester (72).....	130
Appendix 19. HR-MALDI-TOF-MS spectrum of MTM SA-L-alanine methyl ester (78).....	131
Appendix 20. ¹ H NMR spectrum of MTM SA-L-alanine methyl ester (78).....	132
Appendix 21. ¹³ C NMR spectrum of MTM SA-L-alanine methyl ester (78).....	133
Appendix 22. gHSQC NMR spectrum of MTM SA-L-alanine methyl ester (78).....	134
Appendix 23. gHMBC NMR spectrum of MTM SA-L-alanine methyl ester (78).....	135
Appendix 24. HR-MALDI-TOF-MS spectrum of MTM SA-L-glycine methyl ester (80).....	136
Appendix 25. ¹ H NMR spectrum of MTM SA-L-glycine methyl ester (80).....	137
Appendix 26. ¹³ C NMR spectrum of MTM SA-L-glycine methyl ester (80).....	138
Appendix 27. gHSQC NMR spectrum of MTM SA-L-glycine methyl ester (80).....	139
Appendix 28. gHMBC NMR spectrum of MTM SA-L-glycine methyl ester (80).....	140
Appendix 29. HR-MALDI-TOF-MS spectrum of MTM SA-L-histidine methyl ester (81).....	141
Appendix 30. HR-MALDI-TOF-MS spectrum of MTM SA-L-lysine t-Bu ester <i>N</i> -benzyl carbamate (82).....	142
Appendix 31. ¹ H NMR spectrum of MTM SA-L-lysine t-Bu ester <i>N</i> -benzyl carbamate (82)....	143
Appendix 32. gHSQC NMR spectrum of MTM SA-L-lysine t-Bu ester <i>N</i> -benzyl carbamate (82).....	144
Appendix 33. gHMBC NMR spectrum of MTM SA-L-lysine t-Bu ester <i>N</i> -benzyl carbamate (82).....	145
Appendix 34. HR-MALDI-TOF-MS spectrum of MTM SA-L-phenylalanine methyl ester (83).....	146
Appendix 35. ¹ H NMR spectrum of MTM SA-L-phenylalanine methyl ester (83).....	147
Appendix 36. gHSQC NMR spectrum of MTM SA-L-phenylalanine methyl ester (83)..	148
Appendix 37. gHMBC NMR spectrum of MTM SA-L-phenylalanine methyl ester (83).....	149
Appendix 38. HR-MALDI-TOF-MS spectrum of MTM SA-L-tryptophan methyl ester (85)...	150

Appendix 39. ¹ H NMR spectrum of MTM SA-L-tryptophan methyl ester (85).....	151
Appendix 40. gHSQC NMR spectrum of MTM SA-L-tryptophan methyl ester (85).....	152
Appendix 41. gHMBC NMR spectrum of MTM SA-L-tryptophan methyl ester (85).....	153
Appendix 42. HR-ESI-MS spectrum of MTM SA-L-5"-Br-tryptophan-methyl ester (86).....	154
Appendix 43. ¹ H NMR spectrum of MTM SA-L-5"-Br-tryptophan-methyl ester (86)....	155
Appendix 44. gHSQC NMR spectrum of MTM SA-L-5"-Br-tryptophan-methyl ester (86).....	156
Appendix 45. gHMBC NMR spectrum of MTM SA-L-5"-Br-tryptophan-methyl ester (86)....	157
Appendix 46. HR-MALDI-TOF-MS spectrum of MTM SA-L-tyrosine methyl ester (87).....	158
Appendix 47. HR-MALDI-TOF-MS spectrum of MTM SA-L-valine methyl ester (88).....	159
Appendix 48. ¹ H NMR spectrum of MTM SA-L-valine methyl ester (88).....	160
Appendix 49. ¹³ C NMR spectrum of MTM SA-L-valine methyl ester (88).....	161
Appendix 50. gHSQC NMR spectrum of MTM SA-L-valine methyl ester (88).....	162
Appendix 51. gHMBC NMR spectrum of MTM SA-L-valine methyl ester (88).....	163

LIST OF ABBREVIATIONS

ABCG2	ATP-binding cassette sub-family G member 2
ACoA	acetyl-coenzyme A
ACP	acyl carrier protein
ACN	acetonitrile
AT	acyltransferase
ATP	adenosine triphosphate
ARO	aromatase
CoA	coenzyme-A
CLF	chain-length factor
CTP	cytidine triphosphate
CYC	cyclase
CHM	chromomycin A ₃
DCM	dichloromethane
DH	dehydratase
DIPEA	<i>N,N</i> -diisopropylethylamine
DMSO	dimethyl sulfoxide
<i>E. coli</i>	<i>Escherichia coli</i>
EtOAc	ethyl acetate
ER	enoyl reductase
Enz	enzyme
ESFTs	Ewing sarcoma family of tumors
EWS-FLI1	Ewing sarcoma breakpoint region 1 and Friend leukemia virus integration 1
FAS	fatty acid synthase
FPLC	fast protein liquid chromatography
G1P	α -D-glucose-1-phosphate
G6P	α -D-glucose-6-phosphate
GE	gilvocarcin E

GT	glycosyltransferase
GTP	guanosine triphosphate
GM	gilvocarcin M
GV	gilvocarcin V
H bond	hydrogen bond
HPLC	high-performance liquid chromatography
HR-EI-MS	high resolution-electron ionization-mass sepctrometry
HR-ESI-MS	high resolution-electrospray ionization-mass spectrometry
HTS	high-throughput screening
IC ₅₀	half maximal inhibitory concentration
IPTG	β-D-1-thiogalactopyranoside
KS	β-ketoacyl synthase
KR	ketoreductase
LC-MS	liquid chromatography-mass spectrometry
MAT	malonyltransferase
MCoA	malonyl-coenzyme A
MCAT	malonyl-CoA:ACP transacylase
MeOH	methanol
MS	mass spectrometry
MT	methyltransferase
MTM	mithramycin
MTM SA	mithramycin SA
MTM SA-Ala	mithramycin SA-L-alanine
MTM SA-Glu	mithramycin SA-D-glucosamine
MTM SA-Gly	mithramycin SA-L-glycine
MTM SA-His	mithramycin SA-L-histidine
MTM SA-Lys-2PGs	mithramycin SA-L-lysine t-Bu ester <i>N</i> -benzyl carbamate
MTM SA-Phe	mithramycin SA-L-phenylalanine
MTM SA-5''-Br-Trp	mithramycin SA-L-5''-Br-tryptophan

MTM SA-Trp	mithramycin SA-L-tryptophan
MTM SA-Tyr	mithramycin SA-L-tyrosine
MTM SA-Val	mithramycin SA-L-valine
MTM SDK	mithramycin SDK
MTM SK	mithramycin SK
NADPH	β -nicotinamide adenine dinucleotide phosphate
NDP	nucleotidyl diphosphate
NMR	nuclear magnetic resonance
NRPS	non-ribosomal peptide synthetase
ORF	open reading frames
PCR	polymerase chain reaction
PKS	polyketide synthase
PMC	premithramycinone
PreB	premithramycin B
post-PKS	post-polyketide synthase
PV	polycarcin V
PyBOP	benzotriazol-1-yl-oxypyrrolidinophosphonium hexafluorophosphate
SAH	<i>S</i> -adenosyl-L-homocysteine
SAM	<i>S</i> -adenosyl-L-methionine
SAR	structural-activity-relationship
Sp1	specificity protein 1
TDPG	TDP-D-glucose
TDP	thymidine 5'-diphosphate
TDP-KOL	TDP-4-keto-D-olivose
TMP	thymidine 5'-monophosphate
TMSCHN ₂	trimethylsilyldiazomethane
TRAIL	tumor necrosis factor- α -related apoptosis-inducing ligand
TTP	thymidine triphosphate
UTP	uridine triphosphate

WT

wild type

CHAPTER 1: BACKGROUND INFORMATION

1.1 Natural products

Natural product sources, whether from microbes, plants, fungi, or marine organisms, have long been used by humans to explore for potential medicines against diseases. The history of natural products usage can be traced back to prehistoric times, and their benefits were passed down by word of mouth and then recorded by writing systems of the ancient civilizations, such as the Mesopotamian, Egyptian, Chinese, Indian, and Greek cultures. Today, natural products still provide some of the most potent anticancer and antibiotic agents and offer templates for new drug design and improvement. Among 1,073 approved small-molecule drugs for all diseases worldwide from 1981 to 2010, 686 drugs (64%) were discovered or derived from natural products.^[1] Large polypeptides and vaccines either isolated from an organism/cell line or produced by biotechnological methods in a surrogate host can also be regarded as natural products. The importance of all natural products for drug discovery and development is still greater than just translational natural products if these biomolecules are included among the collection of natural products. Modern medicine is the product of thousands of years' worth of accumulation of natural product research, and natural products will continue to be an important source for new drug discovery in the future.

Natural products may be classified into six categories based on their building blocks and biosynthetic routes: fatty acids and polyketides (**1** and **2**), terpenoids and steroids (**3** and **4**), alkaloids (**5**), phenylpropanoids (**6**), specialized amino acids and peptides (**7**), and specialized carbohydrates (**8**) (**Figure 1**).^[2] In general, bioactive natural products are secondary metabolites that are not directly involved in the normal growth, development, or reproduction of organisms and appear only in a limited number of organisms. They differ from primary metabolites, such as nucleic acids and proteinogenic amino acids, which play a key role in all living systems.

Small-molecule compounds from natural sources play four significant roles in today's pharmaceutical industry.^[3] First of all, chemical space is the space spanned by all energetically stable molecules and chemical compounds in all possible topological isomers, and is a factor in measuring the chemical diversity of a specific group of

compounds. The chemical space of natural products is much larger than that of synthetic compounds and can cover most of the chemical space of the current marketed drugs, especially those that are difficult to produce commercially by synthetic means.^[4] This indicates that chemists could create compounds with improved structural diversity by mimicking some properties of natural products. Natural products could also be used as prototypes or models for chemists to synthesize drugs possessing similar biological relevance with the parent compounds. Moreover, natural products also provide basic compounds for slight modifications to generate more active or less toxic analogues. Our studies on polycarcin V and mithramycin are good examples. Last but not least, natural products could serve as starting compounds for chemical or enzymatic syntheses to generate potent drugs that are not easily obtained by other methods. For example, taxol, a very potent anticancer agent that was isolated only from the bark of the scarce Pacific yew, was synthesized from baccatin III *via* an oxazoline intermediate.^[5] Because baccatin III was found in the leaves of fast-growing yew species, this method prevented precious Pacific yew from extinction and provided an abundant source of starting material for the synthesis of taxol.

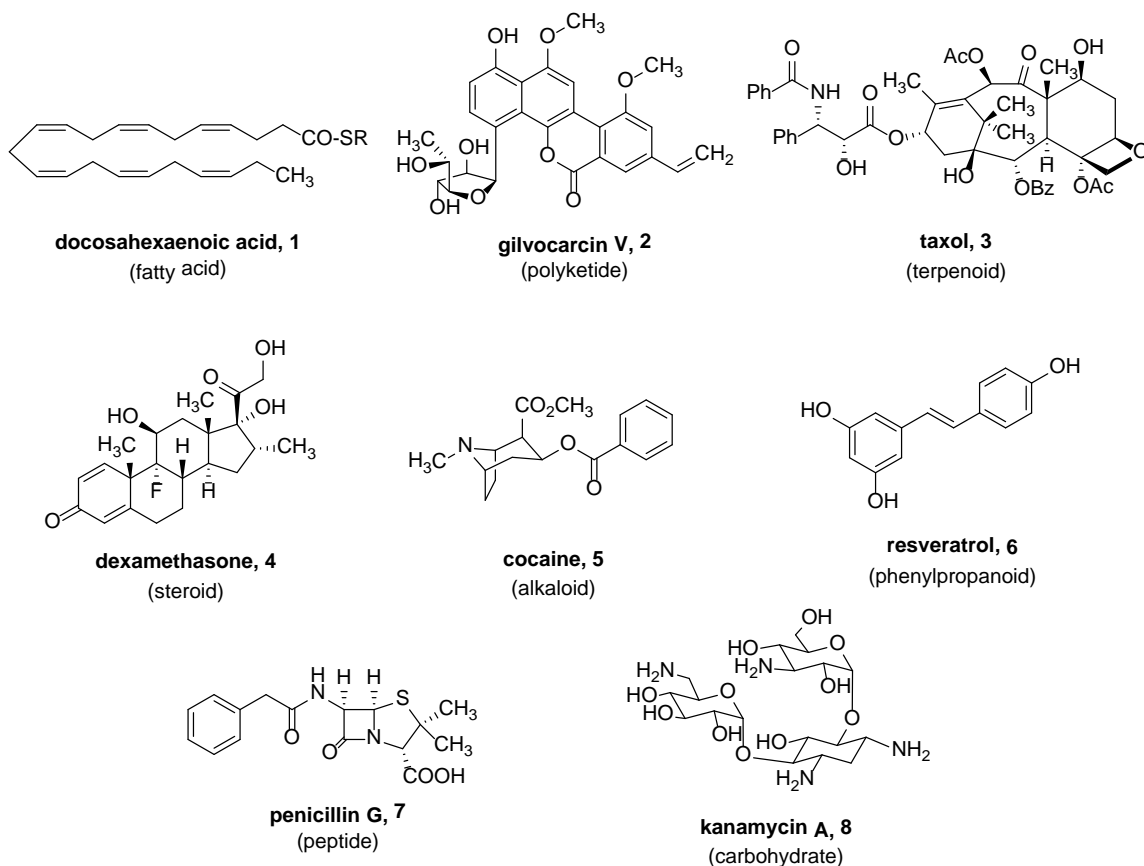


Figure 1. Six categories of natural products

1.2 *Streptomyces* and Polyketides

Streptomyces is the largest genus of Actinobacteria and the type genus of the family Streptomycetaceae. *Streptomyces* are Gram-positive bacteria possessing high-GC-content genomes and usually found in soil as spores but ubiquitous in nature.^[6] Members of the genus *Streptomyces* produce a great number of clinically potential anticancer and antibiotic agents, such as neomycin and chloramphenicol, because of their rich and complex secondary metabolism. Even though some *Streptomyces* spp. infect humans and plants, this genus rarely contains pathogens.^[6] Therefore, like *E. coli*, *Streptomyces* is a good choice for scientists to overexpress specific proteins or produce natural products, such as polyketides.

Polyketides are a large family of natural products that are found in many bacteria, fungi, plants, insects, sponges, and even animals. Because of their unparalleled range of biological activities and enormous commercial value, they are potential lead-structures for new drug discovery to treat many kinds of diseases. Many clinically or commercially important agents are polyketides, including erythromycin A, spinosyn, tetracyclines, and lovastatin. In bacteria, they are biosynthesized from acyl-CoA precursors by polyketide synthases and are reasonably divided into three basic classes based on their biosynthetic pathways (**Figure 2**).^[7] Type I polyketides (**9** and **10**) are often macrolides produced by multifunctional enzymes that are organized into many modules working non-iteratively to complete one cycle of elongation of the polyketide chain (**Figure 3a**). Type II polyketides (**11** and **12**) are often aromatic molecules produced by the iterative action of one or several multienzyme complexes that can carry out many cycles of elongation of the polyketide chain (**Figure 3b**). Type III polyketides (**13** and **14**) are usually small (monocyclic or bicyclic) aromatic molecules produced by acyl carrier protein (ACP)-independent and iterative homodimeric enzymes (**Figure 3c**). In addition, recent literature suggests that the biosynthetic pathways of polyketides are more diverse in both mechanism and structure than the basic three classes, such as type I-type II, type I-type III, FAS-PKS, and PKS-NRPS. (Structure **15 – 17**).^[8]

Despite the enormous success of pharmaceutical polyketides, there is a growing need for the discovery of novel drugs to inhibit the development of drug-resistant pathogens and emerging, infectious microorganisms. However, the structures of polyketides are often too complicated to be efficiently generated by total synthetic strategies. Extensive research on the natural, biosynthetic pathways of polyketides has offered another avenue towards new drug discovery or drug diversification. Using genetic and biochemical information, natural product chemists cannot only generate novel polyketide derivatives but also enhance their chemical space through combinatorial biosynthesis or chemoenzymatic modification. These methods utilize various strategies to modify natural products or their biosynthetic pathways toward the generation of “unnatural” products.

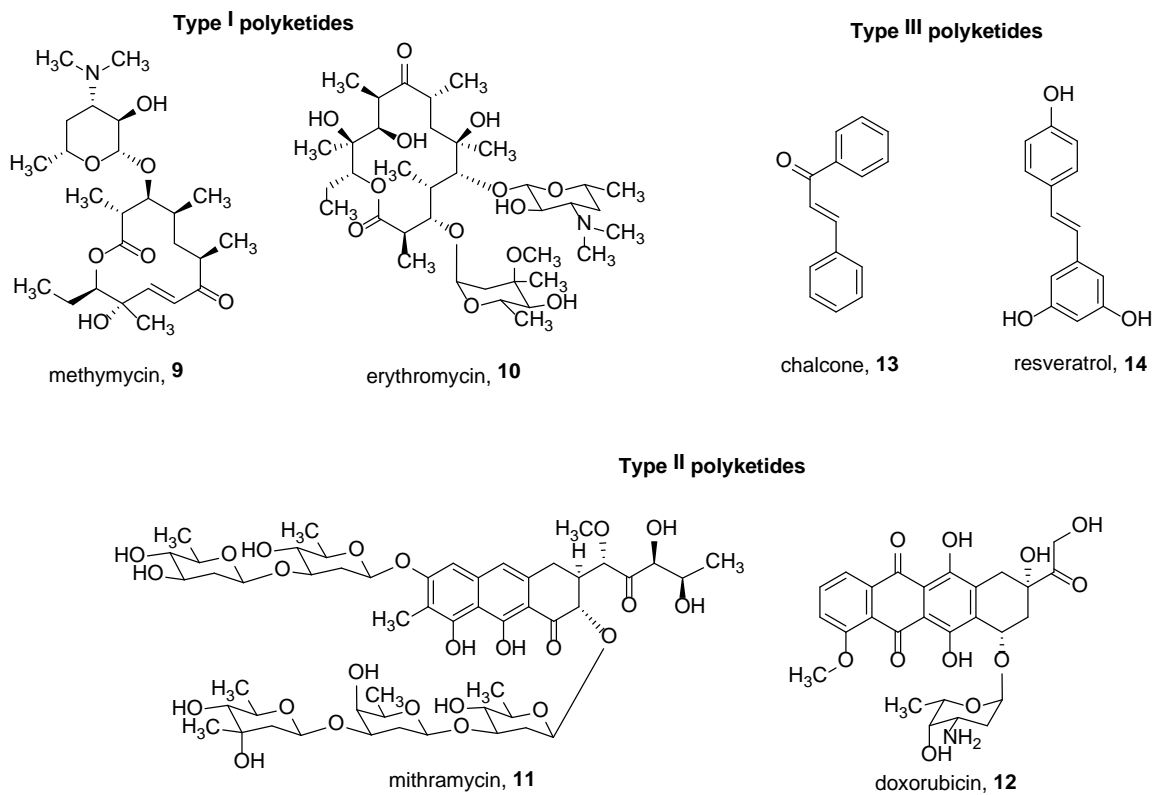
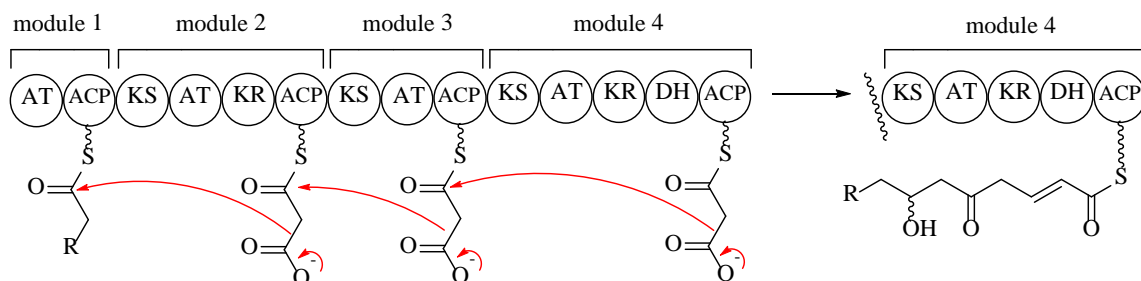
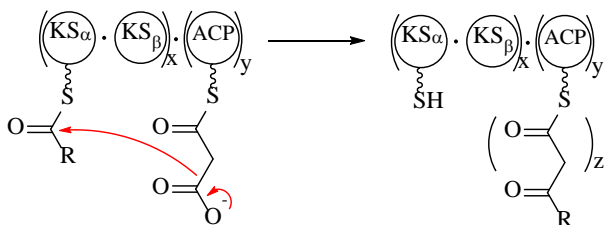


Figure 2. Examples of drugs biosynthesized *via* three basic classes of polyketides.

(a) Type I PKS (noniterative)



(b) Type II PKS (iterative)



(c) Type III PKS (ACP-independent & iterative)

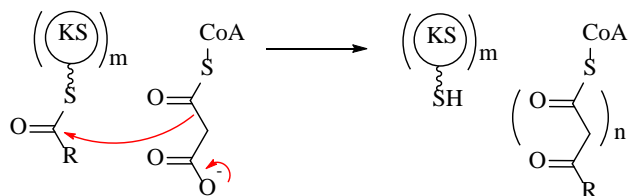


Figure 3. Three basic polyketide synthases (PKS) (a) Type I, (b) Type II, (c) Type III. (AT, acyltransferase; ACP, acyl-carrier protein; KS, β -ketoacyl synthase; KR, ketoreductase; DH, dehydratase; CoA, coenzyme-A.)

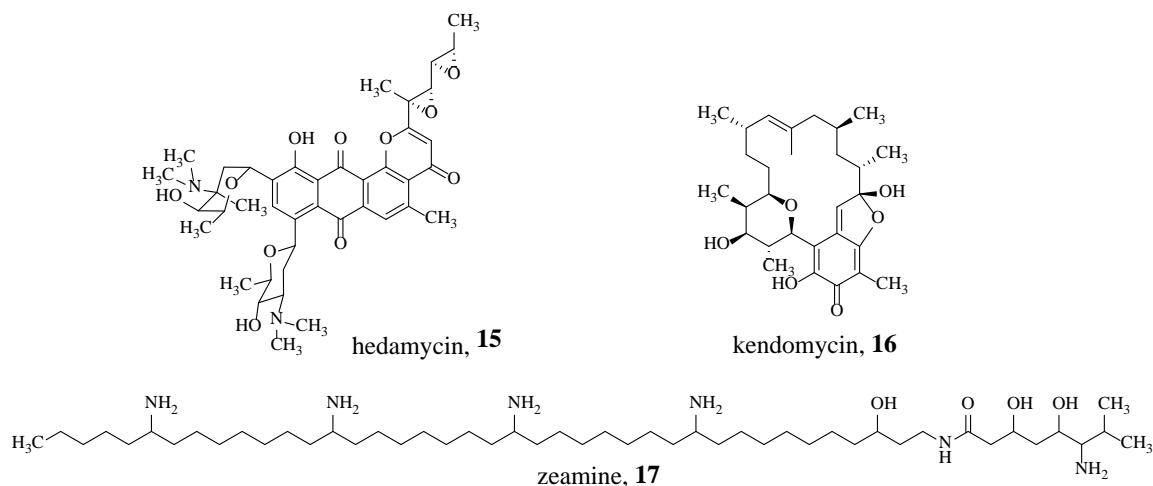


Figure 4. Compounds generated from hybrid biosynthetic pathways.

1.3 Biosynthesis of type-II polyketides

This dissertation focuses on two important type II polyketides, gilvocarcin V and mithramycin. Both of these polyketides are potential anticancer agents, but certain side-effects have seriously limited their clinical utility. This motivated us to try improving these compounds through combinatorial biosynthesis and chemoenzymatic modification. Either way requires a deep fundamental understanding of the biosynthetic processes of the compounds of interest. Therefore, the general information about type II polyketide synthase (PKS) and post-polyketide synthase (post-PKS) tailoring enzymes will be described below.

Type II polyketide synthase

Polyketide biosynthesis shares a common biosynthetic logic with fatty acid biosynthesis. For example, they both originate from simple building blocks such as acetyl-coenzyme A (ACoA) and malonyl-coenzyme A (MCoA) through repetitive decarboxylative Claisen thioester condensations. An activated acyl starter unit is extended with MCoA or MCoA-derived units, and this process typically requires the participation of a β -ketoacyl synthase (KS), an activated acyl carrier protein (*holo*-ACP) and a

malonyl/acyltransferase (MAT/AT).^[9] The elongation cycle is repeated in both pathways until the backbone reaches a predetermined chain length at which point the thioester-bound substrate will be removed from the enzyme by hydrolysis. In fatty acid biosynthesis, every cycle of the elongation reaction catalyzed by the fatty acid synthase (FAS) is followed by a reduction cycle composed of β -keto reduction, dehydration, and enoyl reduction steps to produce a fully saturated backbone (**Figure 5**). The enzymes that participate in this reduction cycle are a ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER).^[2] In contrast to fatty acid biosynthesis, the reduction cycle is optional in the polyketide biosynthesis. The reduction cycle is usually maintained to varying degrees in the biosynthetic pathway of type I polyketides but omitted in the pathways of type II and type III polyketides (**Figure 3**).^[7] Moreover, the polyketide backbones undergo many different kinds of modifications by post-PKS tailoring enzymes to create great structural complexity.

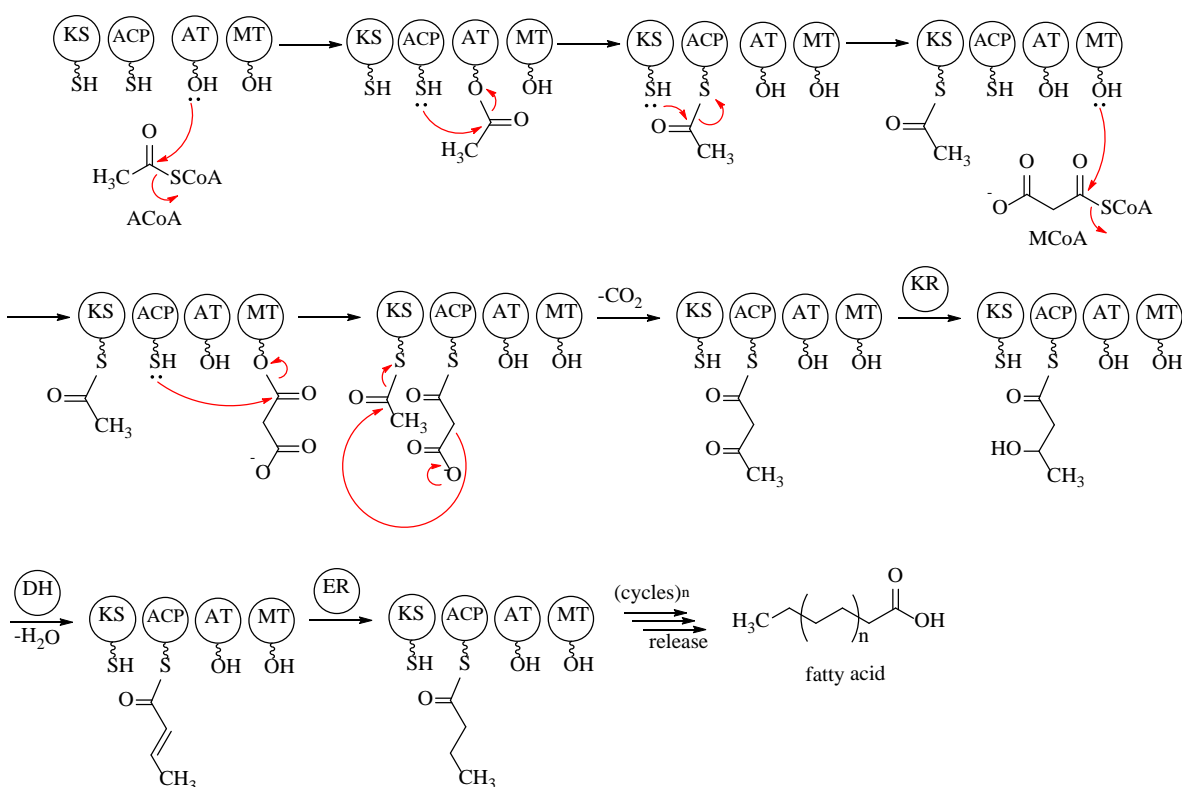


Figure 5. The mechanism of the fatty acyl synthase reaction.

Type II PKSs are only found in prokaryotes, especially in actinomycetes. The typical composition of the minimal enzymes required for the functionality of type II PKS are two ketosynthase units (KS_{α} and KS_{β}) and an acyl carrier protein (ACP). These three proteins form a multi-enzyme complex to assemble the chain of the backbone, and KS_{β} (also named chain length factor, CLF) determines the chain length. The formation of type II polyketide is initiated by the loading of an activated ACP with MCoA by a malonyl-CoA:ACP transacylase (MCAT), borrowed from the native fatty acid biosynthetic pathway, and subsequently this intermediate undergoes decarboxylation to form an acetate ACP complex. The acetate starter unit is transferred from ACP to the KS_{α} subunit. An additional unit of MCoA is loaded onto the ACP which then undergoes decarboxylative Claisen condensation with the acetate primed KS_{α} to catalyze the first elongation step. The product of the extension reaction is found on the ACP, but it will be transferred from ACP back to the KS_{α} for the next round of elongation (**Figure 3**). This process is repeated until the predetermined chain length is achieved. As for more complicated type II PKS, other enzymes, such as ketoreductases, cyclases (CYC), and aromatases (ARO), also interact with the minimal PKS to form a larger enzyme complex and modify the nascent poly- β -keto-thioester intermediate to determine the folding, cyclization, and aromatization pattern (**Figure 6**).^[7] This is the origin of the structural diversity of type II PKS derived polyketide-derived cores.

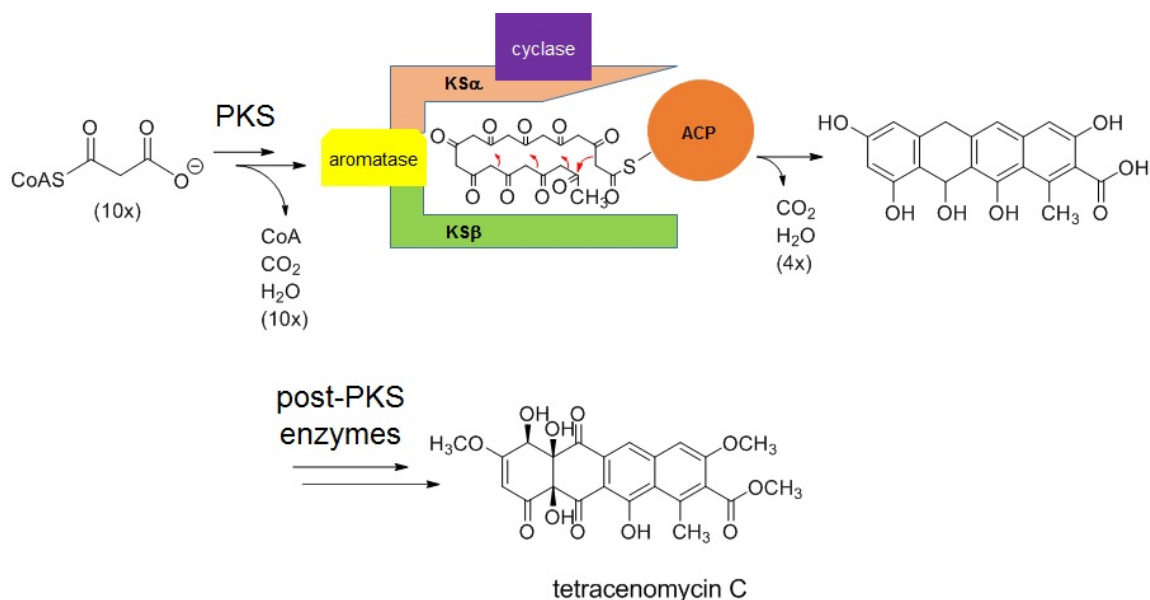


Figure 6. The type II PKS is usually an enzyme complex composed of KS α , KS β , ACP, and other proteins that can modify the backbone.

Polyketide tailoring enzymes

Type II PKS derived cores are further functionalized to obtain higher structural complexity by post-PKS tailoring reactions, including reactions of reduction, oxidation, glycosylation, methylation, halogenation, and addition of deoxysugars (**Figure 6**).^[8a] The physical-chemical properties of polyketides are dramatically changed by this process and become “activated” by these post-PKS enzymes. For example, chromomycin A₃ (CHM), a highly potent anticancer agent from *Streptomyces*, binds to GC-rich regions of DNA to prevent DNA transcription. Degradation studies revealed that CHM analogues lost some of the sugars on the side chains and were less active than the parent compound because the aglycone of CHM did not bind to DNA.^[10] The importance of the post-PKS tailoring enzymes naturally caught natural product chemists’ attention and prompted the development of combinatorial biosynthetic investigations in which thousands of modified polyketides were generated with different pharmacological properties or improved bioactivity profiles.^[11] The post-PKS tailoring reactions related to the topics of this thesis are briefly discussed below.

Deoxysugar biosynthesis

Monosaccharides is extremely important for microorganisms because sugars are not only used to generate energy but also utilized by glycosyltransferases (GTs) in glycosylation reactions to produce secondary metabolites. A majority of secondary metabolites are modified with sugar moieties that are deoxygenated to various degrees prior to their utilization by GTs. The most common sugar in nature is α -D-glucose (**18**), a fully oxygenated sugar, and it is widely used in living organisms as a precursor for the biosynthesis of other sugars. In the beginning, a phosphate group is transferred from ATP to 6-OH of α -D-glucose to generate α -D-glucose-6-phosphate (G6P, **19**) which is interconverted with α -D-glucose-1-phosphate (G1P, **20**) by phosphoglucomutase. Nucleotidyl diphosphate (NDP)-D-glucose synthase is responsible for appending an NDP species to G1P, thereby producing an “activated” NDP-D-glucose (**21**). The NDP used in this reaction could be any nucleotide (NTP), but TTP is the most commonly utilized nucleotide for the deoxysugar biosynthesis in microorganisms. If no further modification is required, the activated TDP-D-glucose (TDPG) could be used directly by GTs to decorate natural products, such as glycopeptide antibiotics.^[12] However, TDP-D-glucose is typically deoxygenated by TDP-D-glucose 4,6-dehydratase to produce the common intermediate for all 6-deoxysugar biosynthetic pathways, TDP-4-keto-6-deoxy-D-glucose (**22**). TDP-4-keto-6-deoxy-D-glucose is a branch point and undergoes further modifications to generate different kinds of deoxysugars by additional deoxysugar biosynthetic enzymes, including dehydratases, epimerases, group transferases, and ketoreductases (**Figure 7**).^[13]

Decorating natural products with deoxysugars significantly changes the physical-chemical properties, and consequently, such modifications are not only useful for structural-activity-relationship (SAR) investigations but also useful for combinatorial biosynthesis. It is possible to construct various deoxysugar genes from the same or different pathways on a single vector and express it by *E. coli* to generate desired NDP-deoxysugars. With substrate-flexible GTs, natural product chemists have rationally designed specific deoxysugar containing natural product analogues.^[14a, 13, 14b]

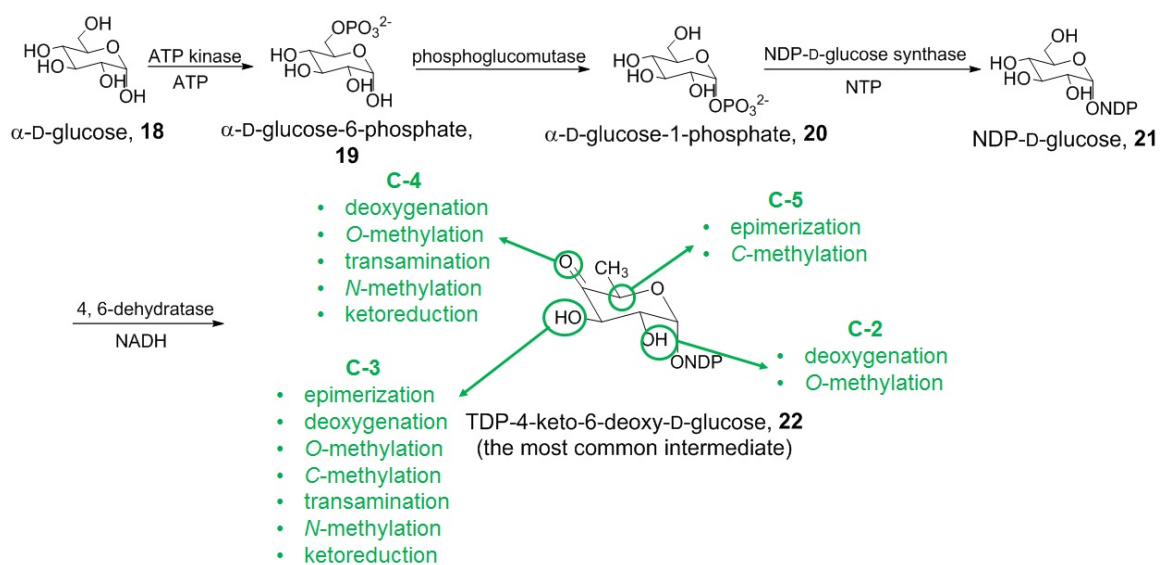


Figure 7. Biosynthetic pathway of deoxysugars.

Glycosylation

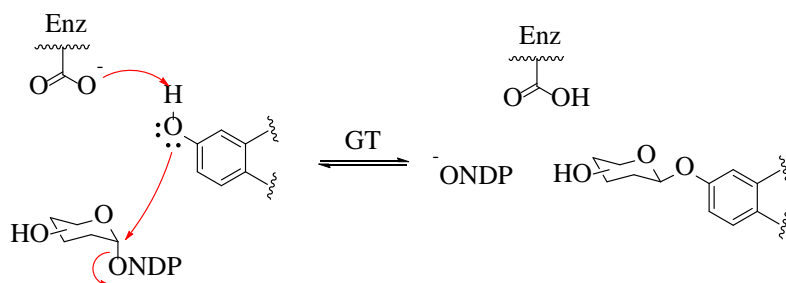
Glycosylation is one of the most important post-PKS tailoring reactions because it can activate marginally active or inactive polyketide-derived cores. This type of reaction is usually carried out by GTs which transfer an NDP-activated sugar (the glycosyl donor) to polyketide-derived cores (the glycosyl acceptor). On occasion, this reaction requires one or more proteins as helper enzymes to assist GTs in transferring sugars efficiently, and many scientists are working to understand the interaction between GTs and their helper enzymes.^[15] Although, glycosylation was always suggested to be a unidirectional reaction, some *O*-glycosylation reactions are reversible, depending on the functionality of the GTs.^[16]

GT transfers sugars to polyketide-derived cores by forming *O*-, *N*-, *S*-, or *C*-glycosidic bonds. *O*-Glycosides are formed frequently and only a minority of *N*-, *S*- and *C*-glycosides exist in secondary metabolites. However, the classification of GTs is based on amino acid sequence similarities rather than the type of glycosidic bonds. Because of the fast-growing genome sequencing techniques, GTs currently are now grouped into more than 90 families, and this number continues to grow.^[17] Even though there is low homology in amino acid sequences of GTs, only two general types of protein folds, GT-A and GT-B, have been

observed for all of the three-dimensional structures of GTs reported thus far. The GTs with GT-A fold have two closely abutting $\beta/\alpha/\beta$ Rossmann domains while the $\beta/\alpha/\beta$ Rossmann domains of the other group with GT-B fold usually face each other and are flexibly linked.^[17] In addition, GTs can be classified into two groups based on the stereochemical outcomes of the reactions that form a new glycosidic bond between the sugar and the polyketide-derived core. The configuration of the anomeric center of the glycosyl donor is either inverted or retained after being transferred to the glycosyl acceptor. The reaction mechanism of inverting GTs is a direct-displacement S_N2 -like reaction in which the nucleophile of the glycosyl acceptor is deprotonated by the GT and then attacks the anomeric carbon to displace the NDP leaving group (**Figure 8a**).^[17] Retaining GTs usually adopt double-displacement mechanism to retain the conformation of the anomeric carbon of the sugar donor. Retaining GTs first attack and tether the glycosyl donor to the enzymes *via* a similar mechanism with inverting GTs. The second displacement occurs when the other catalytic functional group of the enzyme deprotonated the glycosyl acceptor which then can attack the glycosyl donor and form a glycosidic bond (**Figure 8b**).^[17] The stereochemical outcome of a glycosylation cannot be predicted by the folding type of the enzyme, and generally, inverting GTs are much more common in nature.^[17]

GTs plays an important role in combinatorial biosynthesis, and the substrate-flexible GTs are invaluable tools for the generation of novel natural product derivatives. For example, one of the oligosaccharide-forming GTs, LgtC, is both a donor and a glycosyl acceptor-flexible GT. When utilizing unnatural sugars as glycosyl donors or acceptors, it can catalyze the formation of $\alpha(1-2)$, $\alpha(1-3)$, or $\alpha(1-4)$ linkage between two sugars with high-level regio- and stereoselectivity to produce different kinds of disaccharide chains.^[18] This example clearly shows the immense potential of GTs in diversification of natural products.

(a) **Inverting glycosyltransferase**



(b) **Retaining glycosyltransferase**

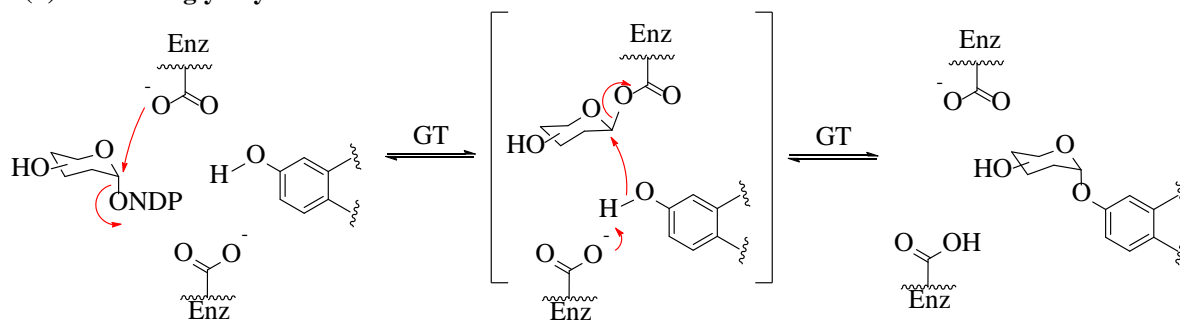


Figure 8. Two stereochemical types of glycosylation reaction

Methylation

The post-PKS tailoring process of polyketides often involves methylation to decorate O, N, S, or C atoms and produce *O*-, *N*-, *S*-, and *C*-tethered methyl groups. This reaction is catalyzed by methyltransferases (MTs) which usually utilize *S*-adenosyl-methionine (SAM, **23**) as a methyl donor in the polyketide biosynthetic pathways (**Figure 9**). Methylation can be involved in the modification of either polyketide-derived cores or sugar moieties, such as L-axenose, L-mycarose, L-nogalose, L-oleandrose, etc. In the sugar biosynthetic pathways, methylation usually occurs prior to the glycosylation step, but sometimes it can happen to sugars that have been already transferred to the aglycones.^[19] Because methyl groups can block H bonds, MTs are also useful enzymes for the combinatorial biosynthesis or SAR investigations.

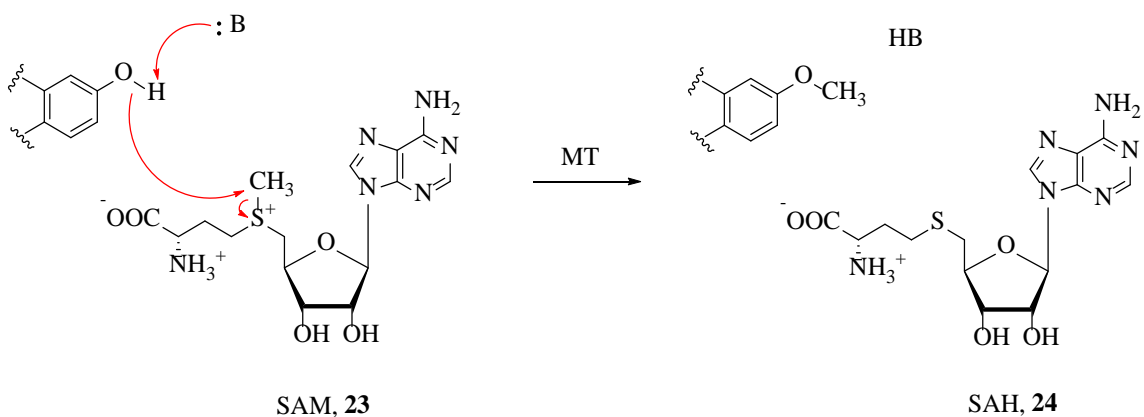


Figure 9. The mechanism of methylation catalyzed by methyltransferases.

Reduction

Reduction is a common reaction occurring in the post-PKS tailoring process. Ketoreductases, which can convert a ketone to a secondary alcohol, are the most frequently found reductases in the polyketide biosynthetic pathways.^[19-20] The required cofactor is usually NADPH, but some reductases may utilize NADH. After the reduction, an unsaturated C-C bond can be introduced into the structure by dehydrogenases to create additional structural complexity. Inactivation of the genes of reductases is also a good strategy toward producing novel polyketide derivatives, especially when the parent compound contains ketone group(s) in the structure. For example, the last step in the biosynthetic pathway of mithramycin (MTM) is reduction of the C-3 side chain with two ketone groups by a ketoreductase, *MtmW*. Inactivation of *mtmW* led to the production of three unnatural mithramycin analogues with better anticancer bioactivity and *in vivo* tolerance than the parent compound as discussed in Section 3.1.^[21]

1.4 Gilvocarcin

Introduction

Gilvocarcin V (GV, **2**), a product of *Streptomyces griseoflavus* Gö 3592 and other *Streptomyces* spp., is the most important member of the gilvocarcin-type aryl-C-

glycosides, a subgroup of the angucycline family of antibiotics that share an oxidatively rearranged polyketide-derived benzo[*d*]naphtho[1,2-*b*]pyran-6-one core with *O*- or *C*-glycosidically linked sugars at various positions.^[14b] This unique natural product was first reported as toromycin without complete structural characterization by Mizuno and coworkers from the culture broth of *S. collinus* in 1980.^[22] Shortly thereafter, other structural-related natural products, such as gilvocarcin M (**25**), gilvocarcin E (**26**), chrysomycin V (**27**), polycarcin V (PV, **28**), ravidomycin V (**29**), FE35A (**30**), Mer1020 dC (**31**), BE-12406A (**32**), were discovered, and all are referred to as gilvocarcin-type aryl-*C*-glycosides (**Figure 10**).^[14b] These structures have variations in the C-8 side chain, usually a single methyl, ethyl, or vinyl functional group, and this is the origin of the M, E, and V abbreviations used for GV-type aryl-*C*-glycosides.

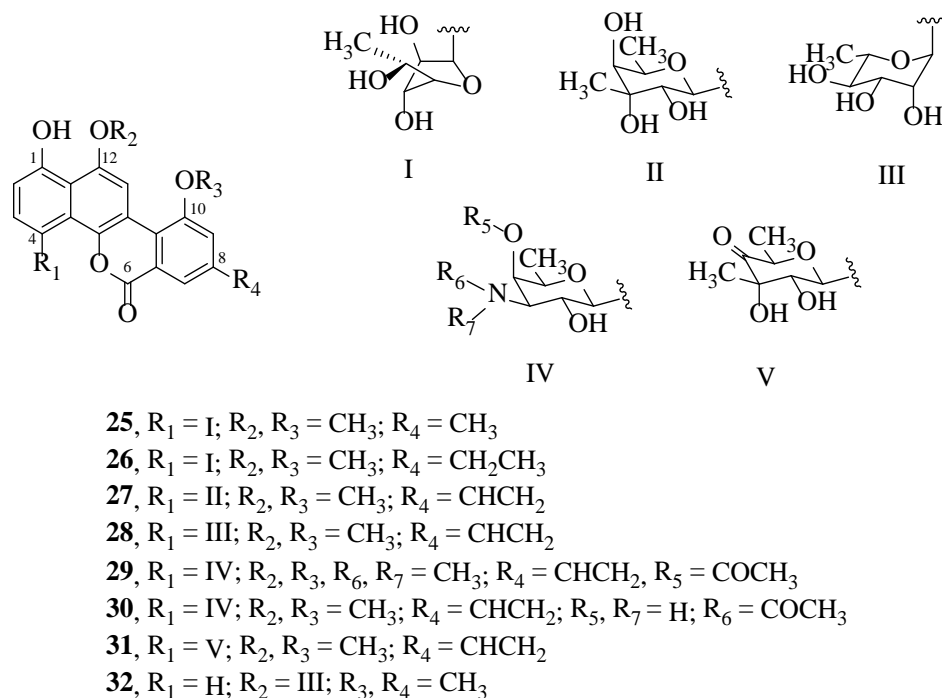


Figure 10. Diversity of gilvocarcin-type anticancer drugs.

In addition, all bioactive members of this family possess a *C*-glycosidically linked 6-deoxyhexose moiety in the C-4 position, and the sugar moiety can be either a furanose (**Figure 10, I**) or pyranose (**Figure 10, II, III, and IV**). It is worth noting that 6-deoxy-L-

hexose (**28** and **32**) is a rare sugar moiety in this family while most of the sugar moieties are in D-form (**25-27** and **29-31**). There are two possible mechanisms of C-glycosylation that have been proposed, O-glycosylation followed by Fries-like O/C rearrangement (**Figure 11**, route A) and direct C-glycosylation (Friedel-Crafts-like reaction) (**Figure 11**, route B).^[23] Even though the exact mechanism remains unsolved, most literature supports the direct reaction of C-glycosylation, because O/C rearrangement is chemically unfavorable for some C-glycosides.^[23-24] The sugar moiety of GV, D-fucofuranose (**Figure 10, I**), is transferred to the polyketide-derived core through C-glycosylation, which results in excellent chemical stability. The stability is because C-glycosides are relatively resistant to spontaneous and enzyme-catalyzed hydrolysis (glycosidase). This fact provides an opportunity to improve the *in vivo* half-life of medicinal O-glycosides by replacing their O-glycosidic bonds with C-glycosidic bonds. Moreover, some scientists had demonstrated that switching an O-glycosyltransferase (O-GT) to a C-glycosyltransferase (C-GT) by engineering the active-site motifs is possible because of the structural similarity of the active sites between the C- and O-GTs.^[24] The engineered C-GT still could recognize the same glycosyl donor and acceptor of the natural O-GT.

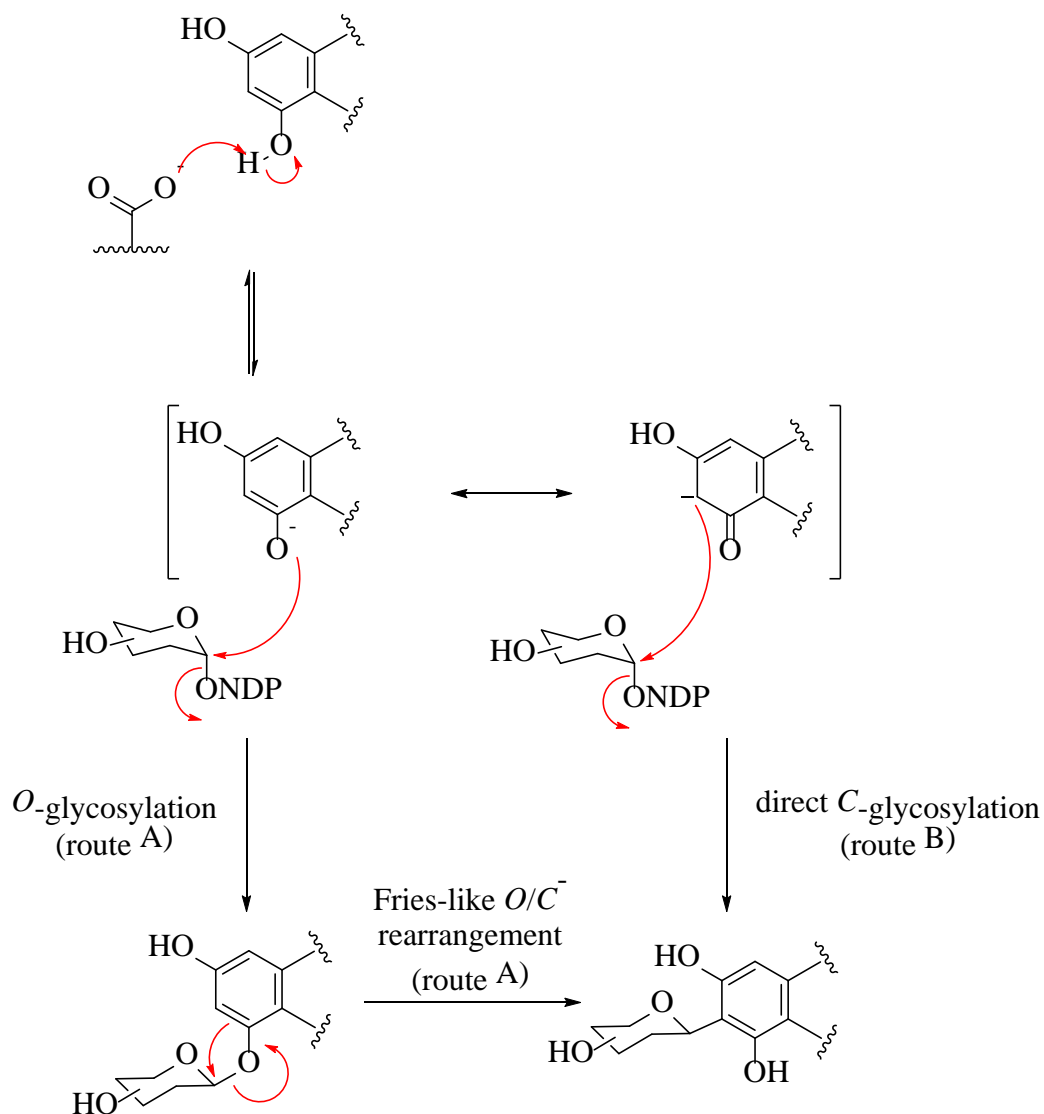


Figure 11. Two proposed mechanisms of enzymatic C-glycosylation.

GV's remarkable antitumor and antibacterial activity *in vivo* with low toxicity is due to a unique, dual mechanism-of-action. First, a photochemical [2+2] cycloaddition of the vinyl side chain with thymine residues of DNA promoted by near-UV or visible blue light results in a covalent binding to DNA, which in turn leads to single-strand scissions.^[25] The bioactivity of GV is also attributed to protein-DNA association resulting from an additional interaction with histone H3, a core component of the histone complex, for which the sugar moiety, D-fucofuranose (**Figure 10, I**), was proposed to form H bonds. The resulting tight

interaction between the histone complex and DNA by these antibiotics leads to growth inhibition of cancer cells (**Figure 12**).^[26]

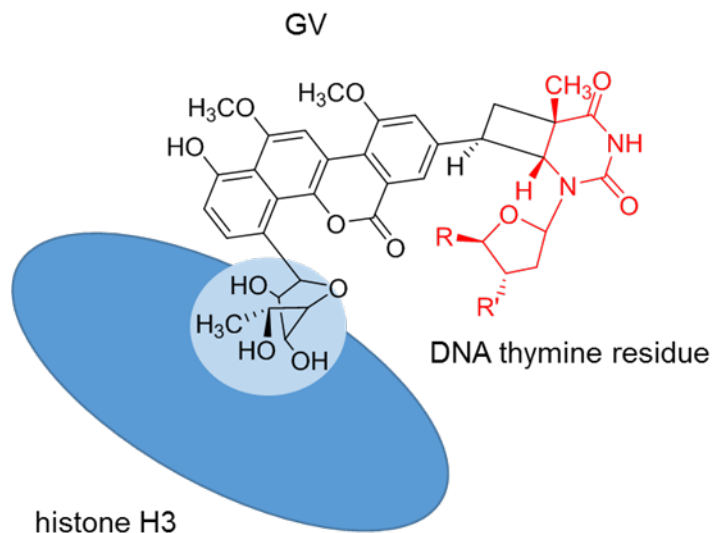


Figure 12. Proposed *in vivo* mechanism of gilvocarcin V resulting from binding to DNA (red) and histone H3 (blue).

Despite intense studies, mainly by researchers of the US National Cancer Institute, a clinical application of GV was never achieved, to a great extent due to its poor solubility and the necessity of activation by light of ~ 400 nm wavelength, which limits tissue penetration.^[27] While modest modifications on the polyketide-derived tetracyclic core can be accomplished through chemo-enzymatic methods, an alternative route for further diversification is through variations in the sugar moieties.^[28a, 28b, 27b, 28c-e] Several chemical or enzymatic approaches to the synthesis of *C*-aryl glycosides and spiro-*C*-aryl glycosides exist, but so far none of them could be used to modify the sugar moiety of GV, D-fucofuranose (**Figure 10, I**), specifically without changing the polyketide-derived tetracyclic core.^[29a, 28e, 29b] Even though different sugar moieties could be transferred to the tetracyclic core *in vivo* by using GilGT, the native glycosyltransferase of the gilvocarcin biosynthetic pathway, or other *C*-glycosyltransferases, such as RavGT and PolGT, from the ravidomycin and polycarcin biosynthetic pathways, respectively, these variations did

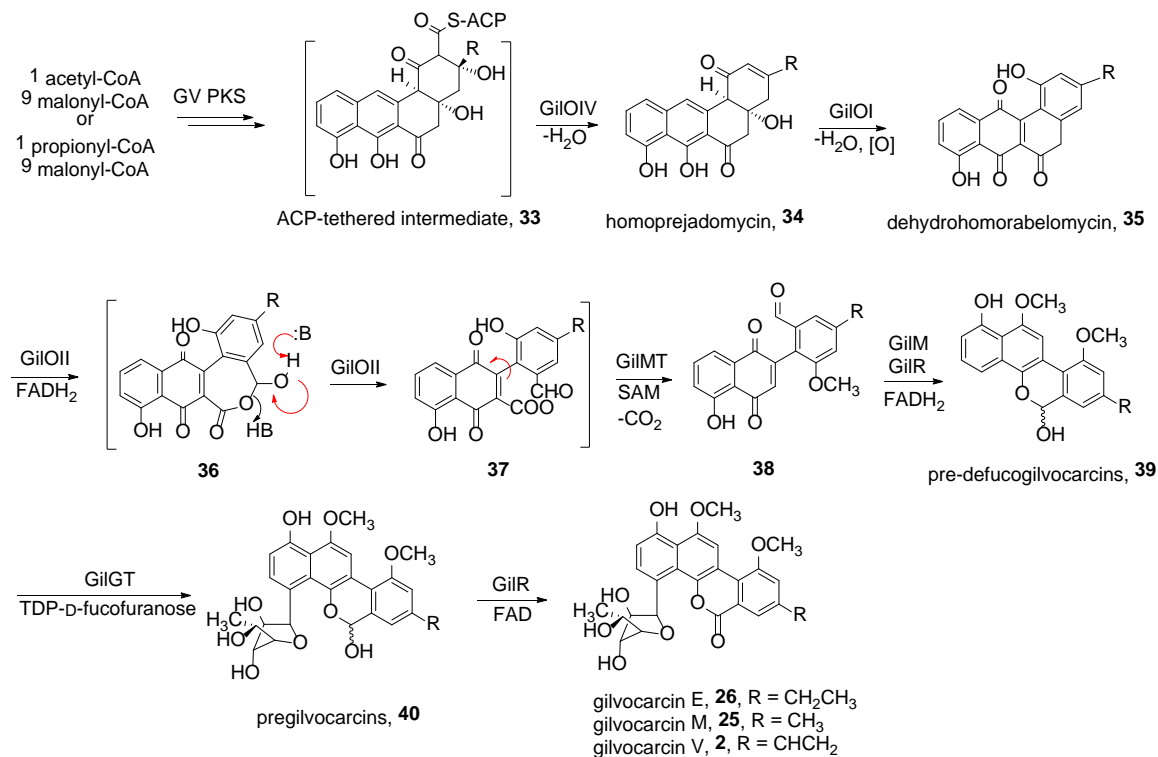
not improve GV, and SAR information regarding the interaction of the deoxysugar moiety and histone H3 is missing.^[27a]

Biosynthetic pathway

The tetracyclic core of GV is biosynthesized *via* the polyketide biosynthetic pathway, particularly a type II PKS (polyketide synthase). Its building blocks are either one unit of acetyl-CoA or propionyl-CoA and 9 units of malonyl-CoA. The first isolation and analysis of biosynthetic gene cluster of GV was carried out with *S. griseoflavus* Gö 3592, and 27 open reading frames (ORFs), including genes encoding the polyketide synthase, post-PKS tailoring enzymes, and deoxysugar biosynthetic enzymes, were discovered.^[30] *S. griseoflavus* Gö 3592 is resistant to genetic modification, but heterologous expression of the *gil* cluster through cosmid cosG9B3 in *S. lividans* TK24 is feasible so that we could confirm the enzymatic function by gene inactivation or *in vitro* assay.^[30]

In the biosynthetic pathway of GV, UWM6 or homo-UWM6 (**33**), which is tethered with ACP, is usually regarded as the first angular-tetracyclic type intermediate, and a multioxygenase complex (consisting of GilOI, GilOII, GilOIII, and GilOIV) plays a key role in the following post-PKS biosynthetic steps (**Figure 13**).^[28d, 31] GilOIV catalyzes 2,3-dehydration to produce homoprejadomycins (**34**) that undergoes a second 4a,12b-dehydration by GilOI to give dehydrohomorabelomycins (**35**). GilOII, a monooxygenase, works consecutively, first adding a 5-OH group and then inserting an O atom between C-5 and C-6, a Baeyer-Villiger reaction that yields structure **36** followed by the C-C bond cleavage to generate structure **37**.^[32] The exact substrate of GilOIII is unclear, but this enzyme is responsible for the vinyl group formation, possibly through hydroxylation and subsequent dehydration on the propionyl-CoA derived ethyl side chain.^[27a] GilMT, a typical SAM-dependent *O*-methyltransferase, is responsible for the decarboxylation and methylation on the hydroxyl group of the phenyl ring. The products (**38**) undergo quinone reduction, hemiacetal formation, and *O*-methylation to give pre-defucogilvocarcins (**39**) with the assistance of a pivotal enzyme, GilM. GilM cooperates with GilR, providing the reduced FADH₂ to help the regeneration of FADH₂ after the reactions.^[33] Inactivation of *gilR* led to the accumulation of pregilvocarcins (**40**) and showed not only that GilR was

responsible for the lactone formation but also that the C-glycosylation reaction occurred prior to this lactone formation.^[34]



p.s. GilOIII is responsible for the formation of the vinyl group, but the substrate is unknown.

Figure 13. Proposed biosynthetic pathway of gilvocarcins (**2**, **25**, and **26**). (GilOIII is responsible for the formation of the vinyl group, but the exact substrate remains unclear.)

The attachment of the furanose sugar moiety is another unclear part of GV biosynthesis. Furanose sugars are rare in polyketide derived natural products.^[27a] Our previous research on the *gil* gene cluster discovered three putative deoxysugar biosynthetic enzymes encoded by the genes *gilD*, *gilE*, and *gilU* and a putative glycosyltransferase, encoded by gene *gilGT*. These enzymes are responsible for the biosynthesis and attachment of the sugar moiety.^[30] *GilD* and *GilE* work together to form TDP-4-keto-6-deoxy-D-glucose (**22**), a common branching point in the biosynthesis of deoxysugars, then *GilU* installs an axial 4-OH group to give TDP-D-fucose (**43**) (**Figure 14**). Presumably an

uncharacterized enzyme carries out ring contraction of TDP-D-fucose and completes the biosynthesis of TDP-D-fucofuranose (**44**) that would be utilized by GilGT as glycosyl donor. This GilGT reaction is a rare C-glycosylation while most of the glycosylated polyketides are O-glycosides.^[27a, 35] With a special furanose moiety attached to the polyketide-derived core through rare C-glycosylation, the investigations on GilGT and SAR information regarding the sugar moiety have a high potential for the discovery of novel GV-type analogues.

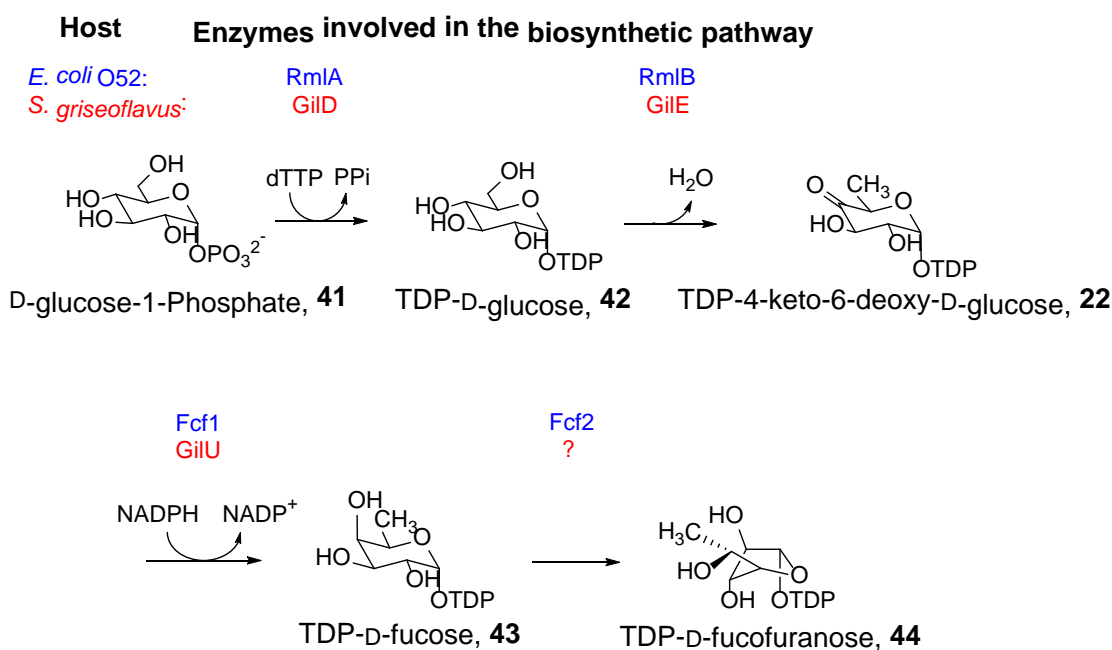


Figure 14. Proposed biosynthetic pathway of TDP-D-fucofuranose in *E. coli* O52 (blue) and *S. griseoflavus* (red).

1.5 Mithramycin

Introduction

Mithramycin (MTM, **11**), the first member of the aureolic acid family, was first discovered as a product of several actinomycetes in the 1950s. MTM possesses excellent anticancer activity and a unique mode of action. All members of the aureolic acid family

are glycosylated aromatic polyketides with yellow color and fluorescence under UV light, and their structures consist of a tricyclic core moiety with a unique dihydroxy-methoxy-oxo-pentyl side chain attached at position C-3.^[21] Moreover, they have two oligosaccharide side chains bound to the aromatic polyketide-derived core with all monosaccharide building blocks belonging to the 2,6-dideoxysugar family and are comprised of different combinations of D-olivose, D-oliose, D-mycarose, L-chromose B, and *O*-methylated or *O*-acetylated derivatives through α -(1, 3) glycosidic bonds. Some compounds, such as MTM, UCH9 (**45**), durhamycin A (**46**), chromomycin A3 (**47**), olivomycin A (**48**), and chromocyclomycin (**49**), have an additional methyl or isobutyl residue at position C-7 (Figure 15).^[36]

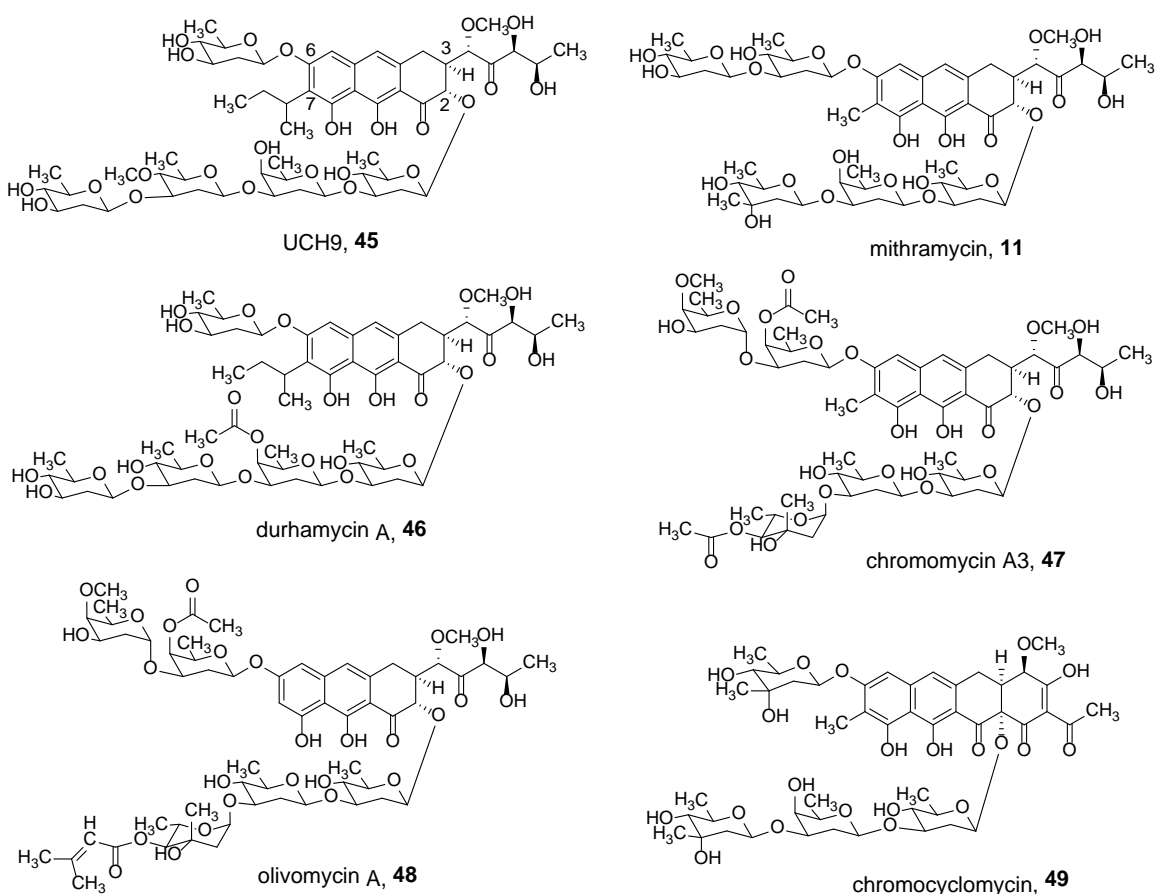


Figure 15. Primary members of the aureolic acid family.

MTM acts by binding the DNA minor groove in regions with high GC content, and this interaction is carried out by forming a dimer under the assistance of a metal ion with a +2 oxidation state, such as Mg^{2+} or Ni^{2+} (**Figure 16**).^[37] The hydroxyl groups on the aglycone of MTM can form many H bonds with the guanine amino protons of DNA.^[36, 38] Moreover, the complex of $Mg^{2+}(MTM)_2$ and DNA can be stabilized by the three side chains of MTM. The disaccharide and dihydroxy-methoxy-oxo-pentyl side chains can interact with DNA phosphate backbone while the trisaccharide side chain wraps across the DNA minor groove.^[39] Sp1 transcription factors have been related to the control of cell growth, survival, and differentiation so that their over-expression usually results in uncontrollable development of tumors.^[40] $Mg^{2+}(MTM)_2$ has been shown to form a complex with promoters that are regulated by transcription factors Sp1, and the transcription of the proto-oncogenes can be shut down.^[36] This gives MTM and other members of the aureolic acid family strong antitumor activity against many types of cancer cell lines. In addition, many investigations have shown that MTM can inhibit calcium resorption in osteoclasts, render cells sensitive to apoptosis mediated by tumor necrosis factor- α -related apoptosis-inducing ligand (TRAIL), exert neuroprotective effects in normal cells, and repress cigarette-smoke induced ABCG2 efflux pumps, one of the markers of cancer stem-like cells.^[41] Therefore, MTM is not only a promising natural product for clinical usage in the anticancer treatment but also an excellent lead-structure for drug discovery.

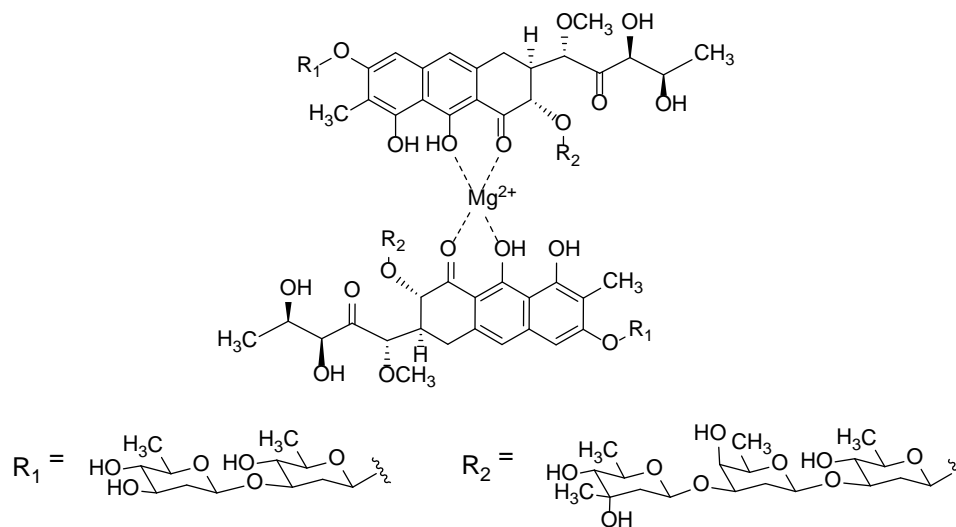


Figure 16. The proposed structure of a Mg^{2+} -coordinated mithramycin dimer complex.

MTM has been used in the treatment of testicular cancer, Paget's bone disease, and hypercalcemia, but its clinical usage is limited because of the unspecific and serious side-effects, such as hepatic, gastrointestinal, bone marrow, and renal toxicities.^[42] The commercial production of MTM was discontinued, and MTM is currently used in basic investigation for structural modification and discovery of new clinical potentials. However, the interest in MTM has recently been renewed from several reports of new uses and bioactivities, especially its antitumor activity to suppress the growth of Ewing sarcoma family of tumors (ESFTs) in xenograft-bearing mice.^[14a] NIH developed a high-throughput screening (HTS) to evaluate more than 50,000 compounds to search for potential candidates that inhibit the Ewing sarcoma breakpoint region 1 and Friend leukemia virus integration 1 (EWS-FLI1) transcription factor and thereby suppress ESFTs. In 2011, it was reported that mithramycin is the only lead-structure which could inhibit ESFTs efficiently and specifically *in vitro* with IC₅₀ between 10 and 15 nM.^[43] This discovery suggested that mithramycin acts on the previously believed undruggable target, EWS-FLI1. Efforts are ongoing to clarify the mechanism of action, improve its bioactivity, and improve *in vivo* tolerance by chemical semi-synthesis or combinatorial biosynthesis.

Biosynthetic pathway

The biosynthetic gene cluster of MTM was isolated and characterized from *S. argillaceus* ATCC 12956 in the 1990s, and most of the genes have been mutated by insertional inactivation to confirm their functions and formulated a biosynthetic pathway (**Figure 17**).^[44] The building blocks of the polyketide-derived core of MTM are one ACoA and nine MCoA, which undergo ten cycles of condensation reaction to generate a linear decaketide, that is aromatized in the first two rings by MtmQ aromatase and then cyclized at the third ring by MtmY cyclase to generate a tricyclic anthrone (**50**). MtmOII oxygenase introduces two hydroxyl groups onto the core prior to the cyclization at the fourth ring by MtmX cyclase, in which a tetracyclic intermediate, 4-demethylpremithramycinone (**51**), is generated. Premithramycinone (PMC, **52**), the last nonglycosylated intermediate in the biosynthetic pathway, is produced through methylation of **51** by MtmMI. Because all members of the aureolic family possess a tricyclic or tetracyclic core moiety derived from

PMC, PMC is a branching point for the biosynthetic pathways of this family, where decoration with different kinds of deoxysugars will occur on this important intermediate.^[36]

The components of the two oligosaccharide side chains of MTM are three D-olivose, one D-oliose, and one D-mycarose. The only difference between D-olivose and D-oliose is the conformation of the hydroxyl group at the position C-4 while D-mycarose has an additional methyl group at the C-3 position. The common intermediate of the deoxysugar biosynthetic pathways, TDP-4-keto-6-deoxy-D-glucose (**22**), is generated from G1P (**41**) by MtmD and MtmE. *mtmV* encodes a TDP-4-keto-6-deoxy-D-glucose-2,3-dehydratase which creates a keto group at the C-3 position. The chemically unstable product of MtmV is reduced by an uncharacterized enzyme to yield TDP-4-keto-2,6-dideoxy-D-glucose (**57**), the branching point for the biosynthesis of all MTM deoxysugars.

MtmU is a reductase, which can directly reduce the keto group of TDP-4-keto-2,6-dideoxy-D-glucose to generate TDP-D-oliose (**58**).^[45] Previous homology searches based on the amino acid sequence revealed that MtmC shared high sequence similarity with many SAM-dependent C-methyltransferases from different biosynthetic pathways of deoxysugars. Subsequently, the recombinant protein was shown to convert **57** to **60** hence functioning as a methyltransferase. However, the *in vitro* assay showed that MtmC also functions as a ketoreductase to reduce the keto group of TDP-4-keto-2,6-dideoxy-D-glucose and produce TDP-D-olivose (**59**). This indicates that MtmC is a bifunctional enzyme and utilized as a reductase and methyltransferase.^[45] The generation of TDP-D-mycarose (**61**) requires two steps – methylation followed by reduction carried out by MtmC and MtmTIII, respectively. TDP-D-mycarose possesses an unstable structure because its hydroxyl group at position C-3 may attack the TDP group to remove the nucleoside by forming 1,2-cyclic phosphate group, by which the production of this sugar could not be detected by UV.^[46] Through an *in vitro* assay with a mixture of MtmC, MtmTIII, MtmGIV with premithramycin A2 (**54**) as a starter, the generation of premithramycin A3 (**55**) confirmed the function of these three enzymes.^[46b]

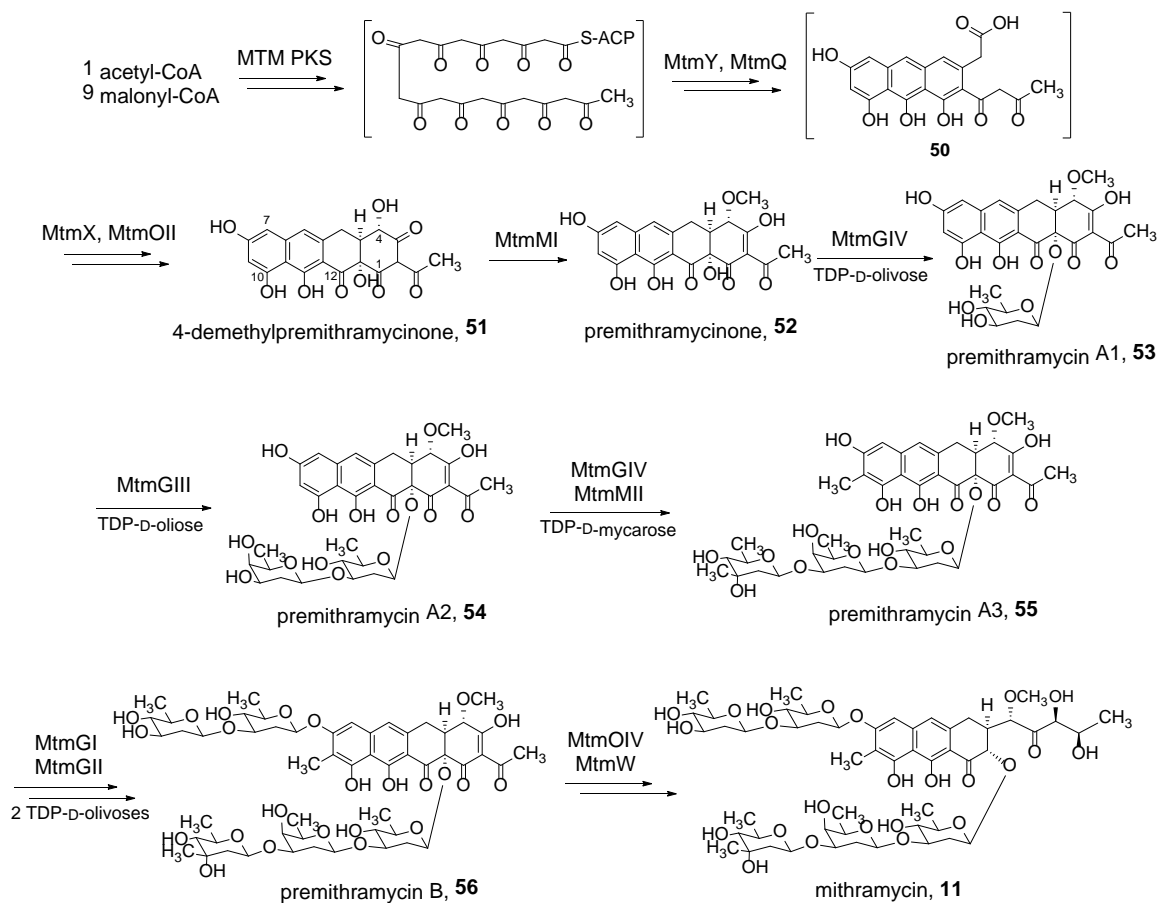


Figure 17. Proposed and simplified biosynthetic pathway of mithramycin.

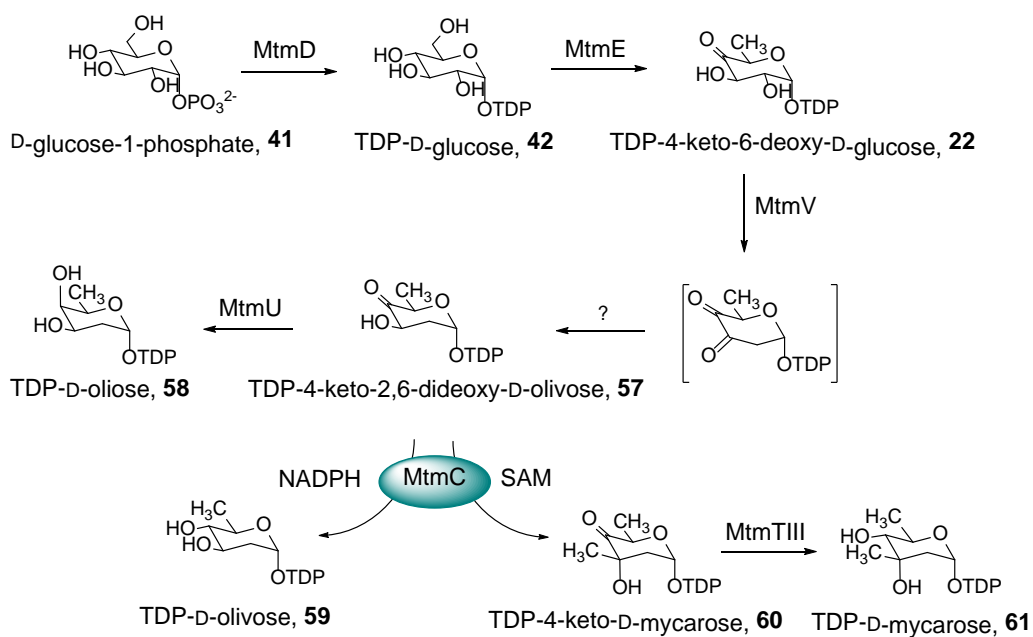


Figure 18. Proposed biosynthetic pathways of the MTM sugars.^[45, 46b]

The biosynthesis of the trisaccharide side chain, which is composed of three different deoxysugars, including D-olivose, D-oliose, and D-mycarose, is formed prior to the disaccharide side chain (**Figure 17**). It is worth noting that only four GTs, MtmGI, MtmGII, MtmGIII, and MtmGIV, are involved in the five glycosylation steps. Our previous investigations demonstrated that MtmGIV is responsible for two different sugar transfers, leading to premithramycin A1 (**53**) and A3 (**55**), through close cooperation with MtmC (**Figure 19**).^[46b] MtmGIII acts between the glycosylation reactions catalyzed by MtmGIV, in which TDP-D-oliose is utilized as a glycosyl donor toward the generation of premithramycin A2 (**54**). Premithramycin A1 can also be recognized as a glycosyl acceptor by CmmGIII from the biosynthetic pathway of chromomycin A3 (**47**).^[36] Premithramycin A2 is decorated with a methyl residue at position C-7 by MtmMII, but the exact substrate of this enzyme remains ambiguous because premithramycin A3 is also recognized as a substrate. The last two glycosylation steps in MTM biosynthesis are achieved by a stepwise deoxysugar transfer of two D-olivose units. *In situ* feeding experiments revealed that MtmGI and MtmGII are responsible for the attachment of the first and second D-olivose moiety, respectively, leading to premithramycin A4 and premithramycin B (PreB, **56**).^[47] Even though PreB possesses almost all of the functional groups in MTM, its bioactivity is much lower than MTM because MTM has a unique pentyl side chain which can interact with the DNA phosphate backbone. The conversion from tetracyclic into tricyclic compounds (from **56** to **11**) is a Baeyer-Villiger oxidation, leading to an oxidative cleavage of the fourth ring by the BVMO, i.e. MtmOIV in the case of MTM biosynthesis.^[48] The last step of the MTM biosynthesis is a reduction catalyzed by MtmW to reduce the keto group at position C-4' in order to stabilize the pentyl side chain.^[21] Because more and more post-PKS tailoring enzymes are identified as being multifunctional and co-dependent on other tailoring enzymes, like MtmGIV and MtmC, there are many intriguing questions waiting for investigation in order to clarify the formation process of all MTM side chains.

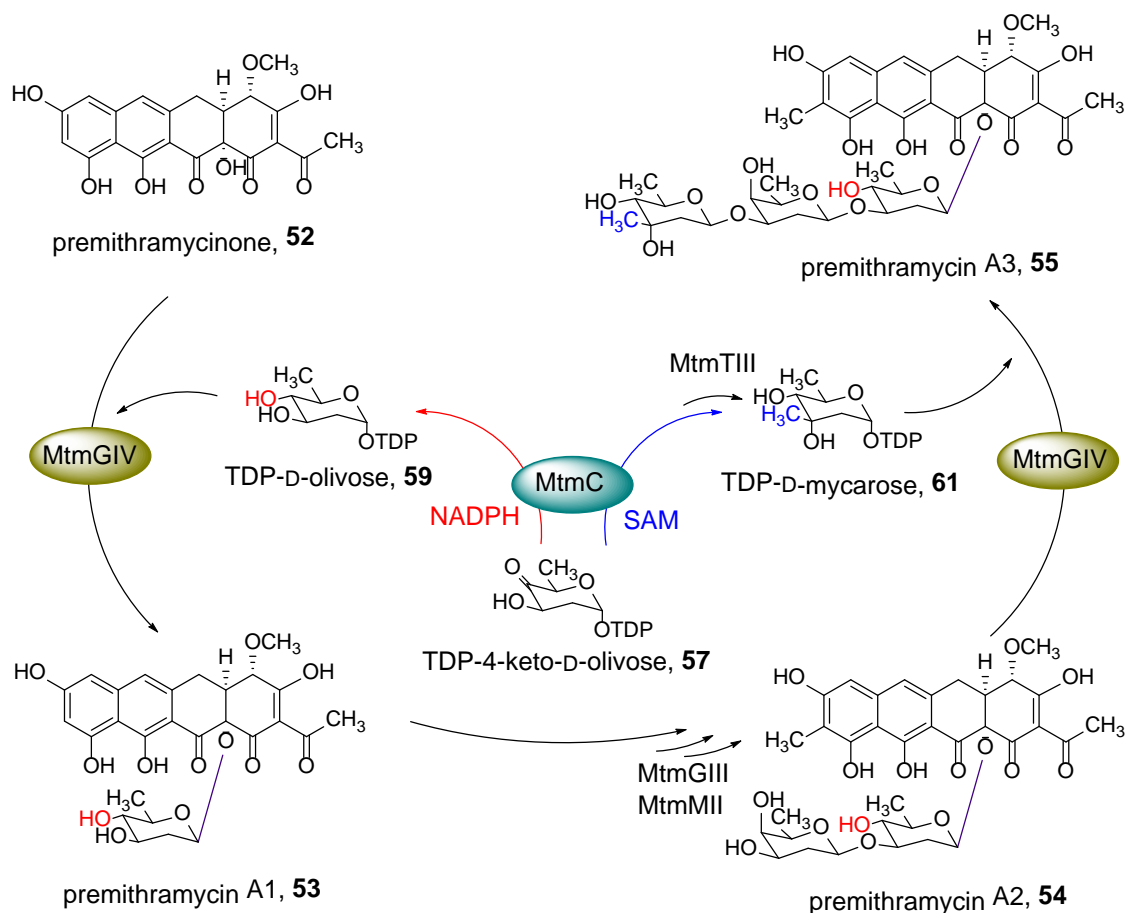


Figure 19. Cooperation between MtmGIV and MtmC during the formation of premithramycin A1 and premithramycin A3.

1.6 Summary

Polyketides are one of the major sources of natural products for new drug discovery with examples found in use throughout the pharmaceutical, agricultural, and chemical industries. However, natural products are not originally produced for human consumption so that they may need to be modified for specific purposes. The most common method is chemical total- and semi-syntheses, but synthetic routes toward large natural products are extremely difficult and may be more complicated when the natural products contain complex structures including sugar moieties. Combinatorial biosynthesis is a potential alternative to chemical synthesis and enables the rational design of natural products for improved bioactivity and *in vivo* tolerance. The understanding of their biosynthetic

pathways is required for this approach, and clarifying the enzymatic mechanisms of the proteins involved in the biosynthetic pathway can benefit protein engineering to generate unnatural products. The overall goal of our laboratory is to elucidate the multi-step biosynthetic pathways of natural products produced by bacteria, fungi, or plants with particular emphasis on enzymatic mechanisms. We generate modified natural product drugs through genetic engineering (pathway engineering and combinatorial biosynthesis), enzymatic or chemical semi-synthesis.

GV and MTM are both type II polyketides with excellent anticancer activities against several cancer cell lines, but their severe drawbacks had limited the clinical usage. GV is only soluble in DMSO and needs UV-visible light for activation so that GV is usually used in the treatment of skin diseases. Many scientists had modified the polyketide-derived core of GV by chemical semi-synthesis and also transferred novel deoxysugars to the core with substrate-flexible GTs, but the SAR information regarding the sugar moiety is still missing. The drawback of MTM is the unspecific and severe toxicity for normal cells, and this has resulted in MTM being withdrawn from the clinics. The recently discovered inhibition of ABCG2 (ATP-binding cassette sub-family G member 2) efflux pump and the cancer-specific hybrid transcription factor, EWS-FLI1, by MTM has attracted scientists to re-investigate this old anticancer antibiotic. This has likewise renewed interest in MTM to enhance the specificity and *in vivo* tolerance. The primary goals of this thesis research is to solve the above-mentioned challenges, and we hope that our research results could contribute useful ideas for further rational drug design of both GV and MTM.

1.7 Specific aims

The objectives of the research are to (a) provide the missing SAR information regarding the sugar moiety of GV; (b) improve the anticancer bioactivity and *in vivo* tolerance of MTM; (c) characterize the enzymatic mechanisms of a bifunctional enzyme involved in the deoxysugar biosynthesis, MtmC. To achieve these goals, the following specific aims were addressed:

Specific aim 1a: Find a model with a sugar moiety that we are able to modify specifically with enzymes.

Specific aim 1b: Evaluate the anticancer activity of GV analogues and use the bioassay data to infer SAR information regarding the sugar moiety.

Specific aim 2a: Carry out large-scale purification of mithramycin SA (MTM SA) and couple several natural and unnatural primary amines to the carboxylic group of MTM SA.

Specific aim 2b: Evaluate the anticancer activity of MTM SA derivatives and identify what kind of side chain on the primary amine may improve the bioactivity as well as the *in vivo* tolerance.

Specific aim 3a: Generate crystals of MtmC complexed with the substrate (TDP-4-keto-2,6-dideoxy-D-olivose (**57**) or TDP) and cofactors (NADPH or SAM) to obtain structural data by X-ray diffraction.

Specific aim 3b: Solve the structure of the crystal complexes and establish the enzymatic mechanisms.

Specific aim 3c: Confirm the proposed enzymatic mechanisms by kinetic studies on the wild-type and mutant MtmC.

CHAPTER 2: STUDY AND USE OF ENZYMES FOR MODIFICATION OF POLYCARCIN V

2.1 *In vitro* modification of the L-rhamnosyl moiety of polycarcin V

The sugar moiety of GV, D-fucofuranose, was shown to interact with histone H3, probably leading to DNA-protein association *in vivo* and resulting in DNA scissions during the replication and transcription processes.^[26] Therefore, it is important to clarify how the sugar moiety forms H bonds with histone H3 in order to enhance the noncovalent interactions and make GV a better anticancer agent. However, TDP-D-fucofuranose is a rare deoxysugar in nature, and its biosynthetic pathway has been only discovered in *E. coli* O52 and *Streptomyces* spp. which can produce gilvocarcins.^[30, 49] Scientists have not yet found any enzymes that can modify the functional groups on this sugar (**Figure 10, I**) because the chemical routes to modify it are complicated and not regio-selective. Therefore, it is very difficult to obtain SAR information regarding the sugar moiety of GV.

In order to solve this problem, we hope to identify a model compound that is composed of the same polyketide-derived core as GV but also with a sugar moiety that is more common than D-fucofuranose. Our previous work has shown that GilGT is a substrate-flexible GT and could be used to transfer L-rhamnopyranose, a common sugar in nature, to the core. The product of the reaction was an isomer of GV, polycarcin V (PV, **56**). Additionally, the following cytotoxicity assay revealed that PV possessed comparable anticancer bioactivity with GV.^[14b] Moreover, PV (**56**) can be readily isolated from wild-type *Streptomyces polyformus* so that it is an ideal alternative model of GV for SAR investigations of the sugar moiety.^[50]

2.2 Experimental design

PV offered the possibility of enzymatic *O*-methylation of the sugar moiety, since many L-rhamnose-*O*-methyltransferases are known. Unlike the rare sugar D-fucofuranose, an L-rhamnopyranose moiety, usually as mono, di-, and per-*O*-methylated derivatives, is commonly found in polyketides, such as in the steffimycin (**57**), elloramycin (**58**) and spinosyn (**59**) pathways (**Figure 20**).^[51] Thus, it was decided to investigate the substrate

flexibility of *O*-methyltransferases from such pathways in order to modify the sugar moiety of PV *in vitro*, and to synthesize mono- (**60-62**), di- (**63-65**), and per-methylated (**66**) PV derivatives (**Figure 21**). It was expected that the methoxy groups might modify the interaction between the sugar moiety and histone H3 (**Figure 12**), thereby providing some SAR information.^[26] Since PV is the natural by-product of the GV producer *S. polyformus*, this compound could be easily produced to serve as a control standard.^[50b] For these reasons, we chose PV as an alternative to GV for initial SAR studies regarding the sugar moiety, expecting some insight into how the three hydroxyl groups of the sugar function.

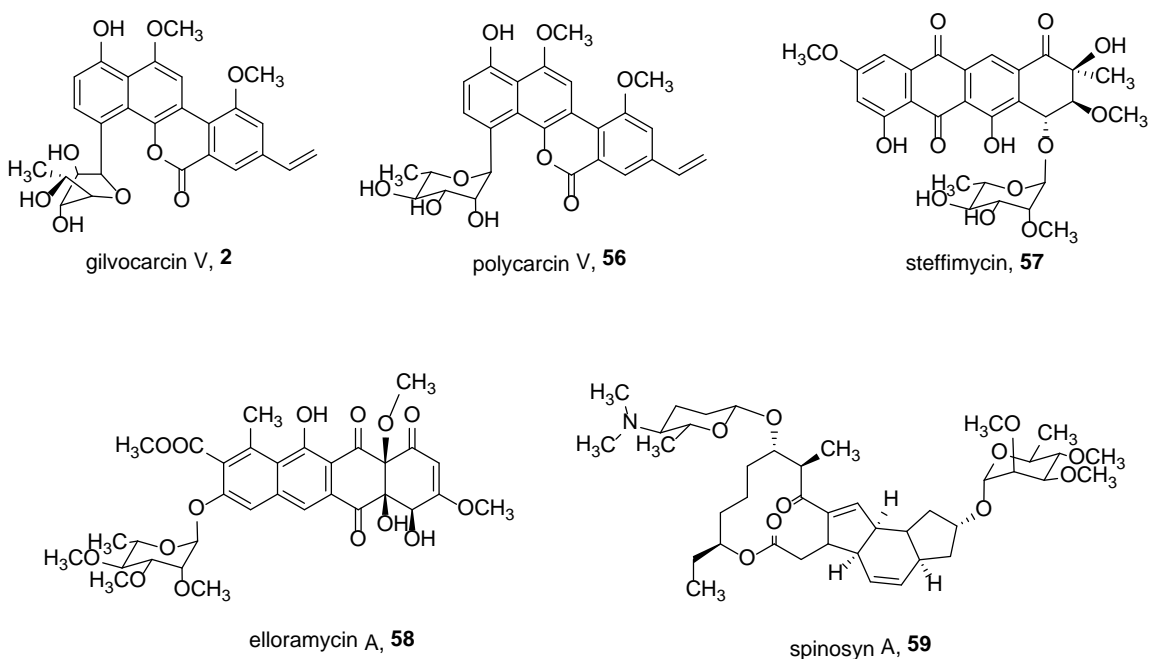


Figure 20. Gilvocarcin V, polycarcin V, and other natural products with L-rhamnopyranose sugar moiety

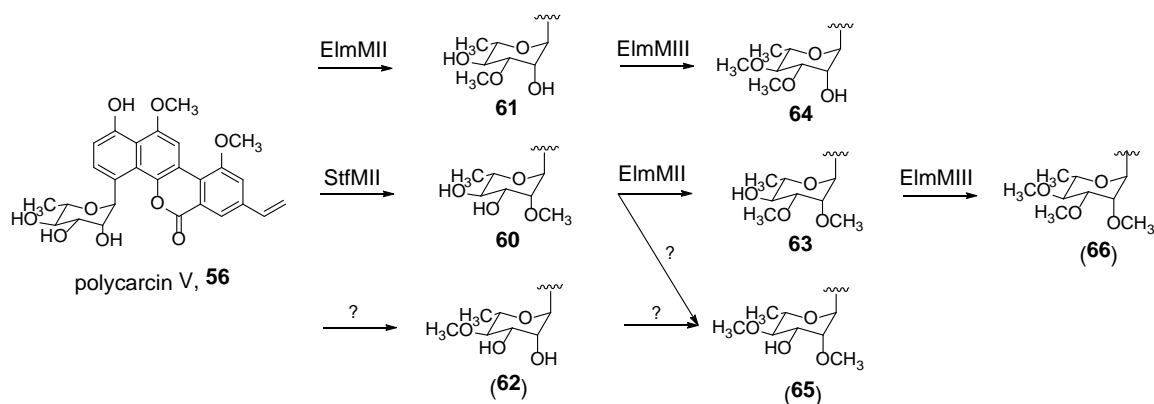


Figure 21. Polycarcin V and its derivatives.

2.3 Results

In vitro reactions with auxiliary *O*-methyltransferases

Three auxiliary *O*-methyltransferases (*O*-MTs), StfMII from the steffimycin producer *S. steffisburgensis* NRRL 3193, and ElmMII and ElmMIII from the elloramycin producer *S. olivaceus* Tü 2353, were used to methylate the L-rhamnopyranose moiety of PV and produce 2'-*O*-methyl-polycarcin V (60), 3'-*O*-methyl-polycarcin V (61), 2', 3'-di-*O*-methyl-polycarcin V (63), and 3', 4'-di-*O*-methyl-polycarcin V (64) (Figure 21 and 22). Other *O*-MTs, for example, one from the spinosyn pathway failed to modify the L-rhamnopyranose moiety of PV, and all attempts failed to find an *O*-methyltransferase able to methylate 4'-OH without the presence of 3'-methoxy group. Thus, 62 and 65 could not be generated. The genes encoding these enzymes were constructed in the vector pET28a(+) to fuse with a *N*-terminal His-Tag and over-expressed by *E. coli* BL21 (DE3). After being purified with the IMAC column, enzymes (40 μ M) and starter compounds (200 μ M) were incubated in a mixture containing 50 mM KH_2PO_4 , 20 mM MgCl_2 , and 2 mM *S*-adenosyl methionine (SAM) at 28 °C for 16 hrs. Low yields of the *in vitro* reactions were the most consistent problem in preparing enough PV derivatives for bioassays. The instability of the enzymes usually led to premature reaction termination or even precipitation of both, enzymes and products. The maximum yield of every *O*-methylation step was about 20%, but the production of 2',3'-di-*O*-methyl-polycarcin V (63) was unexpectedly low, just ~5%. We assumed that this phenomenon might have resulted from the tendency of GV-type

compounds to self-assemble through intermolecular stacking in an aqueous solution.^[52] The studies were further hampered because PV is actually only a small by-product of the GV biosynthesis pathway of *S. polyformus*. The average yield was $\sim 0.5 \text{ mg L}^{-1}$, and required a multi-step purification process including liquid-liquid partition, normal phase silica gel chromatography, and semi-preparative HPLC. Various attempts to increase the production of PV were met at best with moderate success. For example, media variations or adding supplementary scandium, a rare earth element that was shown to stimulate antibiotic production in *Streptomyces* spp. by 2 - 25 fold at low concentrations (10-100 μM) to the culture broth, only moderately increased the average yield of PV. Because of these obstacles, we were not able to produce sufficient 2',3',4'-tri-*O*-methyl-polycarcin V (**66**) for cytotoxicity assay.^[53]

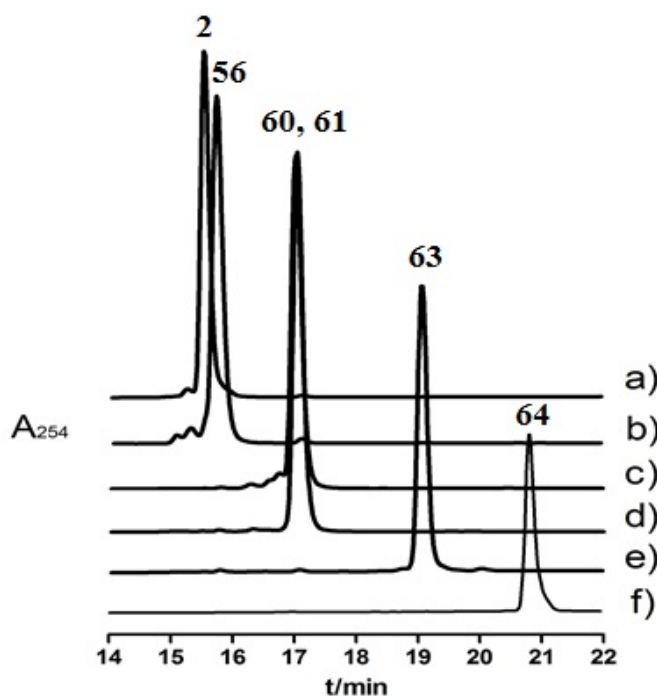


Figure 22. HPLC traces of gilvocarcin V and polycarcins. a) GV, **2** (RT = 15.52 min, AUC = 312,942); b) PV, **56** (15.71 min, 464,750); c) 2'-*O*-methyl-polycarcin V, **60** (17.00 min, 429,135); d) 3'-*O*-methyl-polycarcin V, **61** (17.03 min, 278,817); e) 2', 3'-di-*O*-methyl-polycarcin V, **63** (19.04 min, 61,550); f) 3', 4'-di-*O*-methyl-polycarcin V, **64** (20.85 min, 129,134). A_{254} : absorbance at 254 nm. RT: retention time. AUC: area under curve.

Structure elucidation of the polycarcin V derivatives:

The structures of **60**, **61**, **63**, and **64** were analyzed through NMR spectroscopy (**Tables 1** and **2**) and high resolution mass spectrometry. HR-MS data directly revealed the occurrence of mono- and di-*O*-methylation while two-dimensional gHMBC spectra confirmed the methyl group/s to be transferred to the desired position/s of the sugar moiety. HR-EI-MS data of 2'-*O*-methyl-polycarcin V (**60**) and 3'-*O*-methyl-polycarcin V (**61**) were consistent with the molecular formula of C₂₈H₂₈O₉, indicating [M]⁺ m/z = 508.1717 and 508.1730 (calculated 508.1733) for structure **60** and **61**, respectively. Their ¹H and gHMBC spectra also showed one additional methoxy signal at 2.67 ppm and 3.48 ppm, which could be assigned as 2'- and 3'-methoxy group of the sugar moiety for **60** and **61**, respectively (**Table 1** and **2**). Similarly, for 3',4'-di-*O*-methyl-polycarcin V (**64**), the presence of 3' and 4'-methoxy groups could be confirmed by both HR-EI-MS ([M]⁺ m/z, observed 522.1898, calculated 522.1890) and NMR data which indicated two additional methoxy signals at 3.55 and 3.62 ppm in the ¹H spectrum and correlations with 3' and 4'-proton in gHMBC spectrum (**Tables 1** and **2**). However, because of the poor synthetic yield of derivative **63** (< 5%), we were unable to confirm the structure by 2D NMR and instead relied only on its ¹H NMR spectrum (**Table 1**) and mass spectrometry data to demonstrate the presence of 2' and 3'-methoxy groups. HR-ESI-MS data of **63** was consistent with a molecular formula of C₂₉H₃₁O₉, showing [M+H]⁺ m/z = 523.1961 (calculated 523.1968) and [M-H]⁻ m/z = 521.1809 (calculated 521.1812). The ¹H NMR spectrum showed four methoxy signals, and they could be assigned to two different groups based on the chemical shift values, one (δ_H = 4.26 and 4.27) representing the 10- and 12-methoxy groups of the polyketide-derived tetracyclic core, and the other (δ_H = 3.17 and 3.50) representing the methoxy groups of the sugar moiety.

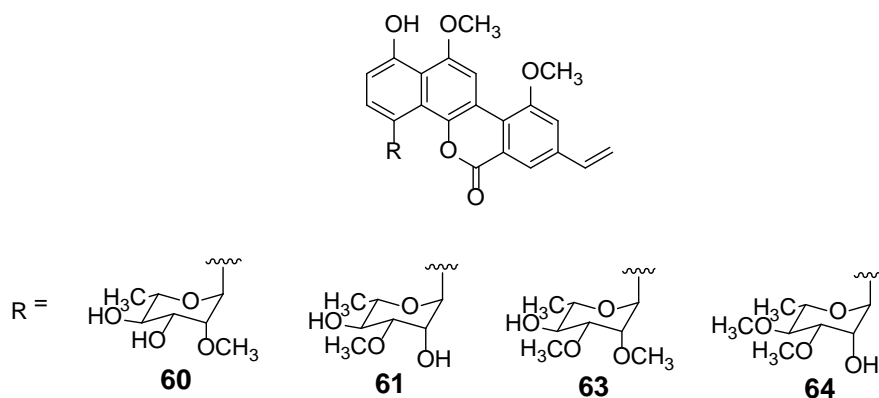


Table 1. ^1H NMR data (500 MHz) for **60**, **61**, **63**, **64**

Compound	60^a	61^a	63^b	64^c
Position	δ_{H} (J in Hz)			
1				
1-OH			9.79, s	9.71, s
2	6.89, d (8.0)	6.88, d (8.0)	6.98, d (8.5)	7.05, d (8.5)
3	7.83, d (8.5)	7.82, d (8.0)	7.97, d (8.5)	7.96, d (8.5)
4				
4a				
4b				
6				
6a				
7	7.83, brs	7.82, brs	8.09, brs	8.09, brs
8				
9	7.40, brs	7.37, brs	7.78, brs	7.37, brs
10				
10-OCH ₃	4.08, s	4.04, s	4.27, s	4.12, s
10a				
10b				
11	8.26, s	8.23, s	8.65, s	8.45, s
12				
12-OCH ₃	4.06, s	4.04, s	4.26, s	4.11, s
12a				
1'	5.81, brs	5.72, brs	5.87, brs	5.92, brs
2'	3.79, d (3.0)	4.48, d (2.5)	3.96, d (3.5)	4.65, d (3.0)
2'-OCH ₃	2.67, s	N/A	3.17, s	N/A
3'	4.03, dd (3.0, 9.5)	3.78, dd (3.0, 9.0)	3.77, dd (3.0, 9.5)	3.94, dd (3.0, 9.5)
3'-OCH ₃	N/A	3.48, s	3.50, s	3.55, s

4'	3.41, t (9.5)	3.57, t (9.0)	3.56, m (4.3, 9.0, 9.0)	3.25, t (9.5)
4'-OCH ₃	N/A	N/A	N/A	3.62, s
5'	3.56, dd (6.0, 9.5)	3.60, dd (6.5, 9.5)	3.47, dd (5.5, 9.0)	3.57~3.62, m
6'	1.47, d (6.0)	1.47, d (5.5)	1.36, d (5.5)	1.43, d (6.5)
1''	6.78, dd (11.0, 17.5)	6.76, dd (11.5, 18.0)	6.97, dd (12.3, 16.3)	6.80, dd (11.0, 18.0)
2''	5.45, d (10.5)	5.43, d (11.0)	5.49, d (11.0)	5.45, d (10.5)
	5.98, d (17.5)	5.96, d (18.0)	6.14, d (18.0)	5.95, d (17.5)

^a The solvent used for NMR experiments was 5% DMSO-*d*₆ in 95% methanol-*d*₄.

^b The solvent used for NMR experiments was acetone-*d*₆.

^c The solvent used for NMR experiments was chloroform-*d*₁.

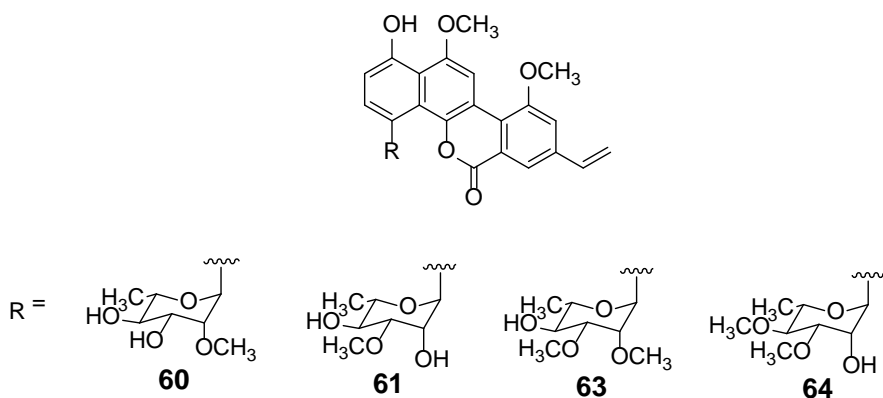


Table 2. $^{13}\text{C}^{\text{a}}$ and gHMBC data (500 MHz) for **60**, **61**, **64**

Compound	60^b		61^b		64^c		
	Position	δ_{C} , mult. ^d	gHMBC	δ_{C} , mult.	gHMBC	δ_{C} , mult.	gHMBC
1		154.8, C	3-H	153.5, C	3-H	153.8, C	3-H
2		113.3, CH		113.1, CH		114.0, CH	
3		130.1, CH	1'-H	130.6, CH	1'-H	130.5, CH	
4		127.1, C	2-H, 1'-H	126.0, C	2-H, 1'-H	110.6, C	2'-H
4a		123.7, C	3-H	123.1, C	3-H	125.7, C	1'-H
4b		143.0, C	11-H	141.9, C	11-H	142.8, C	11-H
6		161.1, C	7-H	160.1, C	7-H	160.1, C	
6a		122.3, C		122.3, C		122.3, C	
7		120.4, CH	9-H	120.3, CH	9-H	120.3, CH	
8		140.3, C	2''-H	139.2, C	2''-H	139.2, C	
9		115.2, CH	7-H	115.2, CH	7-H	115.5, CH	
10		158.8, C	10-OCH ₃	157.5, C	10-OCH ₃	157.5, C	10-OCH ₃
10-OCH ₃		56.8, CH ₃		56.8, CH ₃		57.6, CH ₃	
10a		124.1, C	7-H, 9-H, 11-H	123.0, C	7-H, 9-H, 11-H	123.0, C	
10b		113.4, C		113.4, C		113.4, C	
11		102.8, CH		102.6, CH		103.5, CH	
12		153.4, C	11-H, 12- OCH ₃	152.5, C	11-H, 12- OCH ₃	157.5, C	12-OCH ₃
12-OCH ₃		56.6, CH ₃		56.5, CH ₃		57.6, CH ₃	
12a		116.2, C	2-H, 11-H	115.2, C	2-H, 11-H	115.9, C	11-H
1'		78.4, CH	3-H	78.9, CH	3-H	78.4, CH	
2'		84.9, CH	2'-OCH ₃	69.3, CH		70.2, CH	
2'-OCH ₃		61.7, CH ₃	2'-H	N/A	N/A	N/A	N/A
3'		76.1, CH	2'-H, 4'-H	85.4, CH	3'-OCH ₃	85.7, CH	3'-OCH ₃
3'-OCH ₃		N/A	N/A	56.9, CH ₃		58.3, CH ₃	
4'		74.9, CH	2'-H, 6'-H	73.3, CH	2'-H, 5'-H, 6'-H	83.5, CH	4'-OCH ₃ , 6'-H
4'-OCH ₃		N/A	N/A	N/A	N/A	62.1, CH ₃	
5'		77.8, CH	1'-H, 6'-H	78.0, CH	1'-H, 6'-H	75.1, CH	6'-H

6'	18.7, CH ₃	4'-H, 5'-H	18.6, CH ₃		19.6, CH ₃
1''	136.5, CH	7-H, 9-H	136.4, CH	7-H, 9-H	136.5, CH
2''	117.0, CH ₂		116.9, CH ₂		118.0, CH ₂

^a ¹³C NMR data was inferred from gHSQC and gHMBC spectra.

^b The solvent used for NMR experiments was 5% DMSO-*d*₆ in 95% methanol-*d*₄.

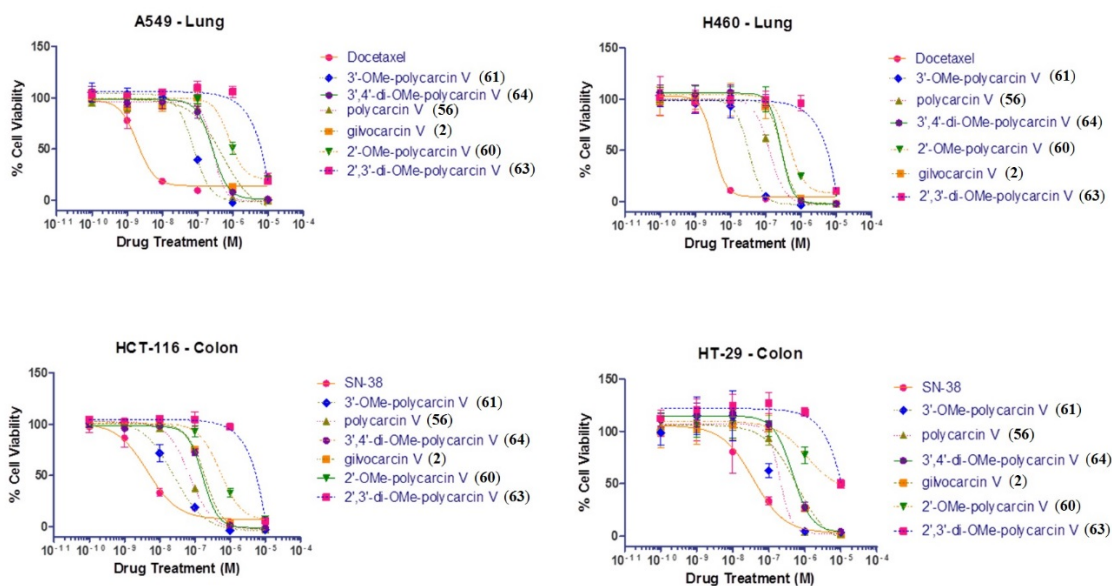
^c The solvent used for NMR experiments was chloroform-*d*₁.

^d multiplicity

Cytotoxicity assays

The bioactivity of the new PV derivatives compared to the natural products GV (**2**) and PV (**56**) were assessed by cytotoxicity assays against cancer cell lines from two human lungs (A549 and H460) and two human colons (HCT-116 and HT-29) (**Table 3**). Docetaxel and SN-38 (active metabolite of irinotecan) were included to compare our analogues to standard clinical treatments for lung and prostate cancers, respectively. The following order of activity (from high to low) was observed: 3'-OMe-PV (**61**) > PV (**56**) > 3',4'-di-OMe-PV (**64**) > GV (**2**) > 2'-OMe-PV (**60**) > 2',3'-di-OMe-PV (**63**).

Table 3. Cytotoxic activities (GI₅₀) of various gilvocarcin/polycarcin derivatives in comparison with two established anticancer drugs against human lung and colon cancer cell lines. (This experiment was carried out by Dr. Jamie Horn. in Dr. Markos Leggas's laboratory)



Compound	Cytotoxicity (GI ₅₀ , nM)			
	A549	H460	HCT-116	HT-29
Docetaxel ^a	1.9	3.3		
SN-38 ^b			4.4	37.4
GV (2)	510.7	287.6	193.6	596.8
PV (56)	281.3	129.4	70.4	203.5
2'-OMe-PV (60)	789.5	452.9	526.6	1140.0
3'-OMe-PV (61)	77.2	33.4	25.3	135.3
2',3'-di-OMe-PV (63)	5.0 x 10 ⁸	2.6 x 10 ⁸	6.4 x 10 ⁸	7.6 x 10 ⁸
3',4'-di-OMe-PV (64)	269.0	273.5	167.4	459.5

^a used as control for human lung cancer cell lines A549 and H460;

^b used as control for human colon cancer cell lines HCT-116 and HT-29

2.4 Discussion

Interestingly, the polyketide-derived benzo[*d*]naphtho[1,2-*b*]pyran-6-one core attached to the L-rhamnose moiety (PV, **56**) was more active than GV (**2**, with D-fucofuranose moiety), possibly because the L-rhamnopyranose might better interact with histone H3 than the D-fucofuranose. Monomethylation had different effects on the lead-structure **56**. *O*-methylation of the 2'-position (2'-*O*-methyl-polycarcin V, **60**) led to decreased activity. While, *O*-methylation of the 3'-position (3'-*O*-methyl-polycarcin V, **61**) increased the activity. This indicated the importance of 2'-OH of L-rhamnopyranose to target interaction, and the necessity of a H-bond donor group in 2'-position. Modifying 3'-OH with nonpolar functional groups (reducing 3'-O to a H-bond acceptor only) seems to improve the interaction between the L-rhamnopyranose sugar moiety and its biological target, presumably histone H3. Although we could not find an enzyme that could selectively *O*-methylate the 4'-position (to yield 4'-*O*-methyl-polycarcin V, **62**) for comparison, it is still reasonable to conclude that 2'-OH is the most important proton donor for the binding of PV to histone H3 because the 4'-OMe of 3', 4'-di-*O*-methyl-polycarcin V (**64**) did not decrease the activity of this compound as much as observed in the other dimethylated polycarcin V, 2', 3'-di-*O*-methyl-polycarcin V (**63**), the weakest PV derivative. We initially hypothesized that two proton donors were required to stabilize the interaction between L-rhamnopyranose and histone H3. However, if this hypothesis were true, 3', 4'-di-*O*-methyl-polycarcin V (**64**) should be much less active. Compound **64** is only marginally weaker than PV. This suggests that the interaction between 2'-OH and histone H3 might be strong enough to overcome potential negative effects resulting from the elimination of other proton donor groups. Another possible explanation is that modifying 4'-OH with nonpolar groups could improve this interaction. According to the HPLC analysis, the structural isomers GV, PV and 2'-*O*-methyl-polycarcin V (**60**), 3'-*O*-methyl-polycarcin V (**61**), respectively, have similar polarity indicated by very close HPLC retention times. On the other hand, 2',3'-di-*O*-methyl-polycarcin V (**63**) is clearly more polar than 3', 4'-di-*O*-methyl-polycarcin V (**64**), because the HPLC retention time of compound **63** significantly differs from that of compound **64** (**Figure 22**). This information indicated that 2'-OH and 4'-OH have different physical and chemical properties, which may contribute or lead to the observed effects on the bioactivity after their methylation. Finally,

the cytotoxicity assays revealed an interesting phenomenon – the colon cancer HT-29 cell line is relatively more resistant against these gilvocarcin-type drugs than all other tested cell lines.

In conclusion, using suitable enzymes we generated four PV derivatives with different combinations of methoxy groups at the sugar moiety for SAR studies. These derivatives may prove useful for further development of new gilvocarcin-type aryl-*C*-glycosides. Despite the inability to obtain 4'-methoxy and 2',4'-dimethoxy PV derivatives (**62** and **65**), we still obtained important information from the data of the cytotoxic assays against four different cancer cell lines. The H-donor/H-acceptor properties of the sugar moiety of gilvocarcin-type aryl-*C*-glycosides likely play an important role in their binding to histone H3. Indeed, it is possible to produce more active derivatives from the natural products by modifying the sugar moieties. Moreover, we noticed somewhat unexpectedly that the new mono-methylated-PV derivatives are partly soluble and di-methylated-PV derivatives are completely soluble in nonpolar solvents. Their solubility is still far from a more desirable water solubility, but it is an improvement compared to the natural products GV and PV, which are only soluble in DMSO. This is the first report of specific enzymatic modification on the sugar moiety of gilvocarcin-type aryl-*C*-glycosides, which provides not only SAR information for further drug discovery and development, but is also another good example of the combinatorial biosynthetic enzymology approach in pharmaceutical sciences.

2.5 Materials and methods

General experimental conditions

All operations were performed under ambient atmosphere. All organic solvents, such as chloroform, dichloromethane, ethyl acetate, and methanol, were purchased from Fisher Scientific Co. (Hampton, NH, USA). All chemicals used for the preparation of culture broth or reaction solution were obtained from Fisher Scientific Co. or Acros Organics Co. (Belgium), except *S*-adenosyl methionine (SAM), which was purchased from Nature Made Nutritional Products Co. (Northridge, CA, USA). The water used in this research was distilled and further purified with Millipore water purification system (Millipore Co.,

Billerica, MA, USA). Analytical thin layer chromatography (TLC) was conducted on silica gel 60-F₂₅₄ from EMD Chemicals Inc. (Darmstadt, Germany), and TLC plates were visualized under UV light at 254 nm. The silica gel (ultra pure, 40-60 μm, 60 Å) used for the flash column chromatography was purchased from Acros Organics Co.

Escherichia coli strains XL1 Blue (Stratagene, La Jolla, CA, USA) and BL21 (DE3) (EMD 4 Biosciences) were used as the hosts for general DNA cloning and over-expression, respectively. Vector pET28a(+) (Novagen, Darmstadt, Germany) was used for protein over-expression. Cultivation, DNA cloning, and transformation in *E. coli* were carried out with standard protocols. The wild type *Streptomyces polyformus* was obtained from Dr. C. Hertweck and Dr. I. Sattler at the Leibniz Institute for Natural Product Research and Infection Biology, Hans-Knöll-Institute (HKI), Beutenbergstr. 11a, D-07745, Jena, Germany.

¹H, gHSQC, and gHMBC spectra were recorded using Varian Inova 500 spectrometer at a magnetic field strength of B₀ 11.74 T (Varian, Inc., Palo Alto, CA, USA). Chemical shifts are quoted in parts per million (ppm) relative to TMS. *J* values are recorded in Hz. All D-containing solvents, such as DMSO-*d*₆, chloroform-*d*₁, methanol-*d*₄, and acetone-*d*₆, were purchased from Cambridge Isotope Laboratories Inc. (Boston, MA, USA) or Sigma-Aldrich Co. (St. Louis, MO, USA). A photodiode array detector (Waters 2996) along with a Micromass ZQ 2000 mass spectrometer (Waters Corporation) equipped with and electrospray ionization (ESI) probe was used to detect the molecular ions and identify the compounds (Waters Co., Milford, MA, USA).

¹H NMR, gHMBC, and gHSQC spectra of PV derivatives, **60**, **61**, **63**, and **64**, were recorded on a Varian Inova 500 spectrometer in either 5% DMSO-*d*₆ in 95% methanol-*d*₄, acetone-*d*₆, chloroform-*d*₁ (**Tables 1 and 2**). MS analysis was carried out by MS facilities at University of Kentucky.

Bacterial strains, culture conditions, and plasmids

Both GV and PV were purified from the wild type *S. polyformus* which was first inoculated on solid M2 medium (4 g/liter glucose, 10 g/liter malt extract, 4 g/liter yeast

extract, 1 g/liter CaCO₃, and 20 g/liter agar) and then transferred to SG liquid media (20 g/liter glucose, 10 g/liter soytone, 2 g/liter CaCO₃, 1 mg/liter cobalt chloride, 175 μM Sc, pH 7.2) for incubation at 28 °C for 7 days. No antibiotics were added when culturing wild type *S. polyformus*. *E. coli* XL1 Blue and BL21 (DE3) were both grown in Lysogeny Broth (LB) supplemented with appropriate antibiotics.

Expression and purification of proteins

The genes encoding StfMII, ElmMII, and ElmMIII were constructed in the vector pET28a(+) purchased from Novagen for protein over-expression. All proteins contain an *N*-terminal 6 × His tag and could be purified with immobilized metal affinity chromatography (IMAC). After amplifying the DNA in *E. coli* XL1 Blue, plasmid isolation was carried out with GeneJet plasmid miniprep kit (Fermentas, Waltham, MA, USA) and then used to transform *E. coli* BL21 (DE3) competent cells for protein over-expression. 1% volume of seed *E. coli* BL21 (DE3) culture with desired plasmid was inoculated into 1 liter of LB supplemented with 50 μg/mL kanamycin (final concentration) and then incubated at 37 °C until OD₆₀₀ value reached 0.4~0.6. The addition of 100 μM β-D-1-thiogalactopyranoside (IPTG, final concentration) induced protein over-expression, and target proteins were obtained after overnight incubation at 18 °C. *E. coli* BL21 (DE3) cell pellets were collected by centrifugation (4000 rpm for 25 min) at room temperature, washed twice with 20 mL lysis buffer (50 mM KH₂PO₄, 300 mM KCl, 10 mM imidazole, pH 8.0), re-suspended in 30 mL lysis buffer for high-pressure French press, and finally centrifuged at 16,500g for 35 min at 4 °C to remove cell debris. The supernatant containing target proteins was loaded on IMAC column (TALON affinity resin, Clontech Laboratories Inc.) that had been equilibrated with 10-fold volume lysis buffer in advance. Impurities and target proteins were eluted by washing buffer (50 mM KH₂PO₄, 300 mM KCl, 20 mM imidazole, pH 8.0) and elution buffer (50 mM KH₂PO₄, 300 mM KCl, 250 mM imidazole, pH 8.0), respectively. After concentrating the elution with Millipore Amicon 30K ultracentrifugal filters, all proteins were desalted with reaction buffer (50 mM KH₂PO₄, 150 mM NaCl, 30% glycerol, pH 7.5) and used immediately for *in vitro* reactions. Protein concentrations were determined with Bradford reagent while their sizes were confirmed

with sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analyses. The measured sizes of purified proteins were in agreement with calculations (**Figure 23**).

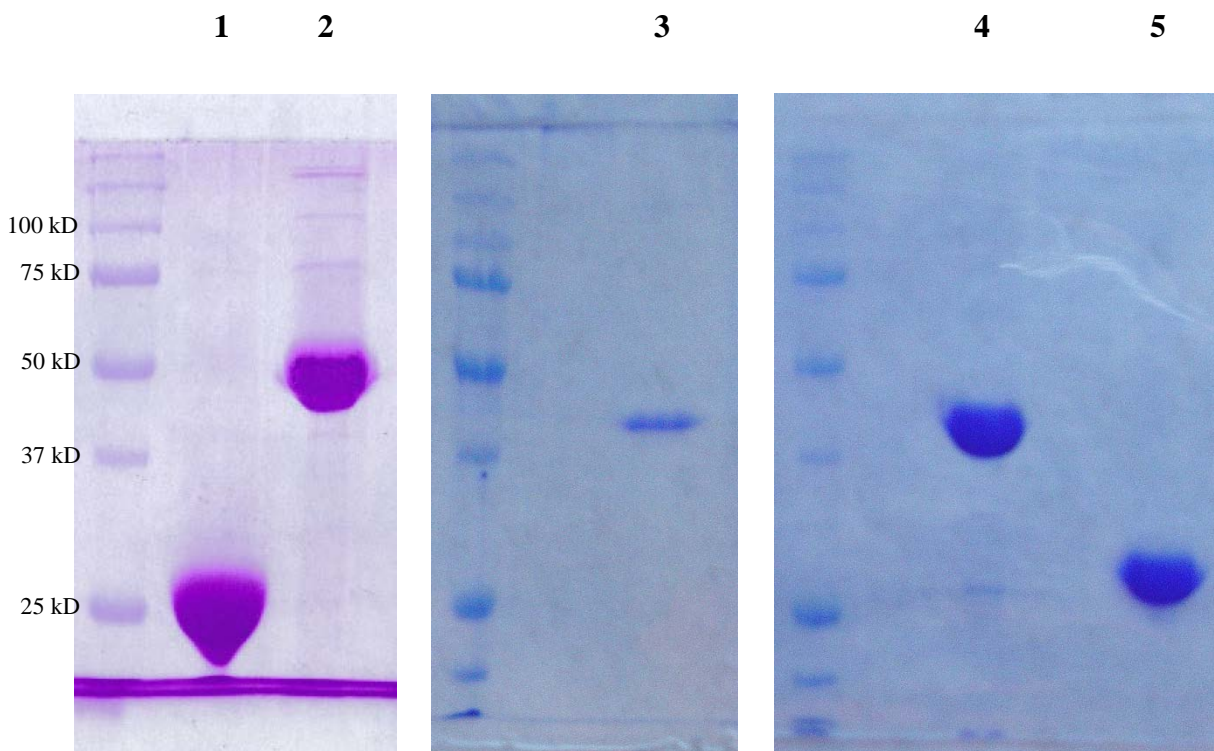


Figure 23. SDS-PAGE analysis of purified proteins used in this study. Lane 2, StfMII (45.61 kD); Lane 3, ElmMII (45.61 kD), Lane 5, ElmMIII (31.21 kD). (Lane 1 and 4 are SAHH and ElmMI, respectively. They were used to increase the yield of *in vitro* reactions but did not work.)

Production and purification of polycarcin V

The spores of the wild type *S. polyformus* on solid M2 medium was used to inoculate 100 liters of liquid SG medium in 250-mL baffled Erlenmeyer flasks (100 mL SG per flask). After 7 days of fermentation at 28 °C with shaking (250 rpm), the culture broth of SG media was centrifuged at 4000 rpm for 25 min to collect cell pellets that were extracted with methanol by sonication. Methanol was then removed from the supernatant under vacuum, and the residue was re-suspended in the aqueous portion of SG media before

extraction with ethyl acetate (EtOAc). After removing EtOAc with rotary evaporator, the extract was passed through normal-phase silica gel column and eluted with 1-liter fractions of 0, 4, and 5% methanol in dichloromethane. Fractions containing gilvocarcins and polycarcins were dried under vacuum and further purified through semi-preparative high-performance liquid chromatography (Waters HPLC system, consisting of a controller, a 2487 Dual λ Absorbance Detector, a 1525EF Binary HPLC Pump, and a column: SunFire™ Prep. C18 5 μ m 10 \times 250 mm; 32.5 min at a flow rate of 2.5 mL/min, UV monitoring absorbance at 254 and 360 nm). Gradient used: solvent A: water with 2% formic acid, solvent B: acetonitrile. Solvent B was increased from 25% to 65% (min 0 to 10), then from 65% to 100% (min 10 to 21), then was decreased back to 25% within 3 minutes, and kept at 25% for the last 8.5 min.

***In vitro* methylation reaction conditions and purification of the products**

All *in vitro* reactions were carried out in the same condition: 500 μ L aliquot containing 50 mM KH_2PO_4 (pH 7.4), 20 mM MgCl_2 , 2 mM *S*-adenosyl methionine (SAM), 200 μ M substrate, 40 μ M methyltransferase, and ddH₂O (doubly distilled water). The reaction solutions were incubated at 28 °C for at least 16 hrs and then extracted with EtOAc three times before further purification by semi-preparative HPLC with the same program, as described above. All compounds submitted to cytotoxic assays were at least 95% pure obtained by one more purification with analytical HPLC and confirmed with LC-MS. (Waters HPLC system, consisting of a controller, a 2996 photodiode array detector, and a 2695 Separations Module, Micromass ZQ, and a column: Symmetry® C18 5 μ m 4.6 \times 250 mm column; 29 min runs at a flow rate of 0.5 mL/min). Gradient used: solvent A: water, solvent B: HPLC-grade acetonitrile. Solvent B was increased from 25% to 100% (min 0 to 15), kept at 100% for 9 min, before it was decreased back to 25% within 2 minutes (min 24 to 26), and kept at 25% for the last 3 min.

Cytotoxicity assay

This cytotoxicity assay was carried out by Dr. Jamie Horn in Dr. Markos Leggas's laboratory at the University of Kentucky. H460 and A549 cells were obtained from ATCC (Manassas, VA). HCT-116 and HT-29 were a gift from Dr. Qing-Bai She at the University of Kentucky. Docetaxel and SN-38 were commercially available from LC-Laboratories and Tocris, respectively. Powdered RPMI-1640 media, McCoy's 5A media, resazurine sodium salt and molecular-grade DMSO were from Sigma-Aldrich. The panel of polycarcin analogues was tested for their cytotoxic activity in human lung and colon cancer cell lines at concentrations ranging from 0.1 to 10,000 nM. All drug solutions were prepared in amber tubes and, all drug treatments were carried out under red light. Lung cancer cells (H460 and A549) were grown in RPMI-1640 (pH 7.4) with 10% FBS (v/v) and 1% streptomycin/penicillin (v/v), whereas the colon cancer cell lines (HCT-116 and HT-29) were cultured in McCoy's 5A (pH 7.4) containing the same supplements. Cells were seeded into 96-well plates at densities yielding exponential growth over a 96 hr experimental time period, specifically at 1,000 (H460), 2,000 (A549 and HCT-116) or 5,000 (HT-29) cells/well in 100 μ L of the appropriate media. The cells were allowed to adhere at 37 °C, and 5% CO₂ for 24 hr. The following day concentrated drug solutions in 1.1% DMSO and 98.9% RPMI-1640 were prepared by serial dilution of 10 mg/mL drug stocks in DMSO and were added to the cells. Cell viability was assessed in a subset of wells following drug/diluent addition and at the end of treatment in all cells (10 μ L of 1 mM resazurine in PBS, pH 7.4). Resazurine was allowed to react for 3 hr at 37 °C (5% CO₂) and fluorescence (590 nm emission/560 excitation) was measured with a spectrophotometer. GI₅₀ parameters and best-fit lines were obtained by nonlinear regression analysis using a dose-response-inhibition equation in Prism 5.04 (GraphPad Prism).

CHAPTER 3: CHEMICAL DERIVATIZATION OF MITHRAMYCIN

3.1 Modification of mithramycin SA with primary amines

Synthetic and semisynthetic approaches toward the generation of unnatural products can be complemented by combinatorial biosynthesis whereby the biosynthetic pathway of the parent compounds are altered in the host through gene inactivation, expression or recombination. MTM is an example of how to utilize combinatorial biosynthesis and chemical modification to generate improved anticancer agents.

The definition of combinatorial biosynthesis is that the biosynthetic pathway of natural products are modified by genetic engineering with nature's biosynthetic machinery in order to produce new and altered structures *in vivo*.^[54] David Hopwood and colleagues are the first group to carry out combinatorial biosynthesis in *Streptomyces* to produce unnatural antibiotics by cloning some or all genes from the biosynthetic pathway of actinorhodin into the producers of medermycin and dihydrogranaticin, respectively.^[55] Our laboratory also used this approach to generate new MTM analogues in 2003, including MTM SK (**67**), MTM SDK (**68**), and MTM SA (**69**), through inactivation of the *mtmW* gene, a gene encoding a ketoreductase carrying out the last step in the MTM biosynthetic pathway.^[21] Compared to the parent compound, MTM SK (**67**) and MTM SDK (**68**) have shortened side chains by one carbon at the 3-position while the side chain of MTM SA (**69**) is the shortest and possesses a carboxylic group. MTM SK and MTM SA are formed through a Favorskii-like rearrangement and retro-aldol-type cleavage, respectively, of the reactive intermediate, mithramycin DK.^[21]

As is well-known, the importance of the C-3 side chain for the anticancer activity had been identified by several *in vitro* and *in vivo* bioassays. MTM SK and MTM SDK showed higher anticancer bioactivity and better *in vivo* tolerance than MTM, but MTM SA was significantly less active.^[56] These results indicate that the C-3 side chain is in part responsible for the interaction between MTM and the DNA-phosphate backbone. One of the reasons of the decreased bioactivity of MTM SA might be that its carboxylic group is too short and carries a negative charge in the physiological environment, a disadvantage for interactions with the naturally negative-charged DNA-phosphate backbone. Nevertheless, its carboxylic acid moiety provides a chance for chemical modification. Our

hypothesis is that the bioactivity of MTM SA can be improved by modifying the C-3 side chain with primary amines through a semi-synthetic approach. This probably can also introduce new functionalities into the C-3 side chain, for example, a new functional group that can interact with a specific protein.^[57]

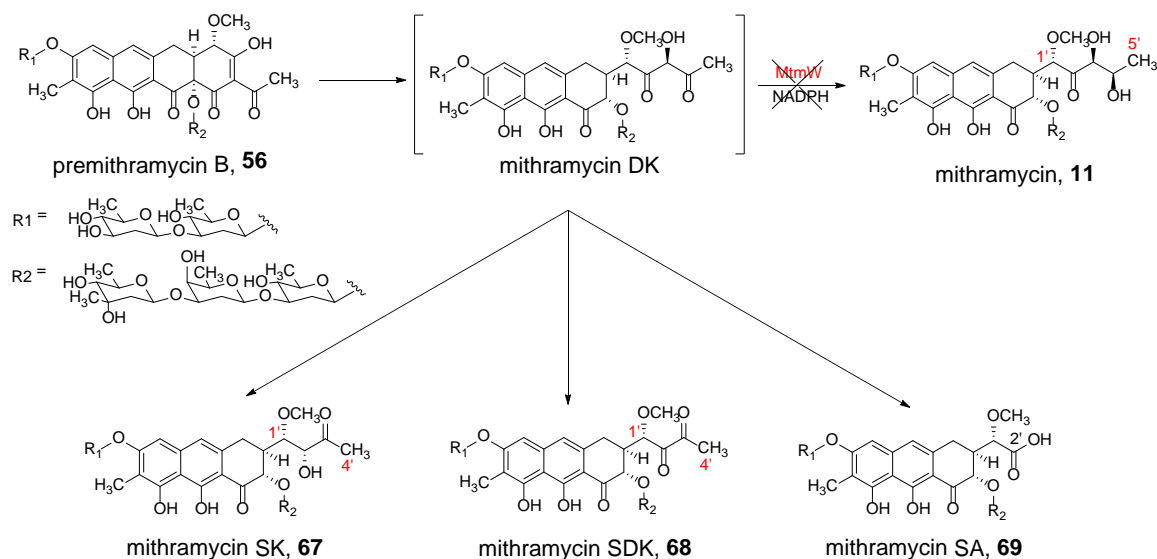


Figure 24. Generation of novel MTM analogues by inactivation of *mtmW*.

3.2 Experimental design

In 2011, Maria N. Preobrazhenskaya and co-workers modified olivomycin SA at the C-3 side chain to generate novel olivomycin analogues.^[57c] Their chemical conditions of the coupling reaction was adopted in this study, but we decided to try more primary amines in order to expand the chemical space of MTM as large as possible. For all the initial trials, 2 mg of MTM SA was used to make sure that the specific primary amines could be transferred to its carboxylic group. The details of the reaction condition are: MTM SA was mixed with 3 equivalents of the desired primary amines for side-chain modification, 3 equivalents of DIPEA (**70**), 2 equivalents of PyBOP (**71**), and the solvent DCM. All reactions started at 4 °C and were checked by LC-MS after 16 hrs. DIPEA was used as a base to deprotonate the carboxylic group of MTM SA while PyBOP was used as a coupling reagent to transfer the desired primary amines onto the deprotonated carboxylic group

(**Figure 25**). In order to confirm that the decreased bioactivity of MTM SA results from the negative-charged carboxylic group, we carried out an esterification reaction to convert the carboxylic acid group to a methyl ester group. The reaction conditions are described as follows: MTM SA was mixed with 2 folds of TMSCHN₂ and 20% methanol in toluene at room temperature and then incubated for 30 min (**Figure 26**). The production of MTM SA methyl ester was confirmed with LC-MS.

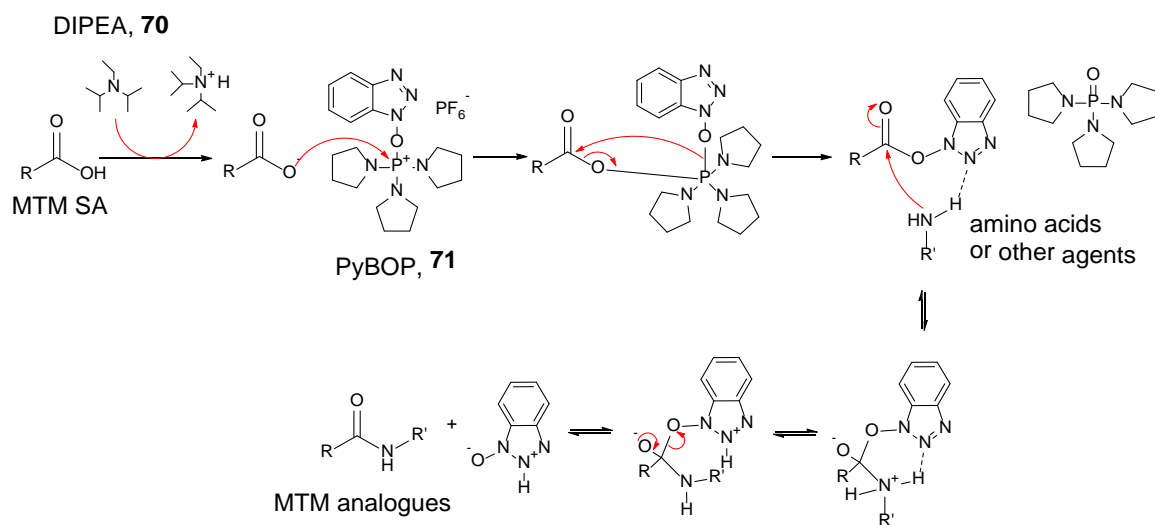


Figure 25. The chemical mechanism of the coupling reaction between MTM SA and primary amines.

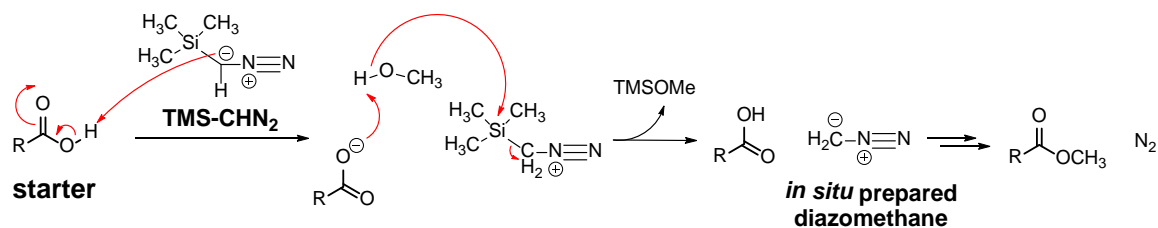


Figure 26. The chemical mechanism of esterification reaction with TMSCHN₂.

3.3 Results

Modification of mithramycin SA with primary amines

We created a library of 16 MTM analogues by coupling a methyl group (**72**) and 15 different compounds with a primary amine functional group to MTM SA (**Figure 27**). The primary amines we used in this study are 1-amino-2-propanone (**73**), cystamine (**74**), *N,N*-dimethylethylenediamine (**75**), D-glucosamine (**76**), methyl hydrazine (**77**), L-alanine methyl ester (**78**), L-cysteine methyl ester (**79**), L-glycine methyl ester (**80**), L-histidine methyl ester (**81**), L-lysine t-Bu ester *N*-benzyl carbamate (**82**), L-phenylalanine methyl ester (**83**), L-serine methyl ester (**84**), L-tryptophan methyl ester (**85**), L-5''-Br-tryptophan methyl ester (**86**), L-tyrosine methyl ester (**87**), and L-valine methyl ester (**88**) (**Figure 27**). All amino acids used in this study possess a methyl group to protect the carboxylic group to increase the efficiency and yield of the products.

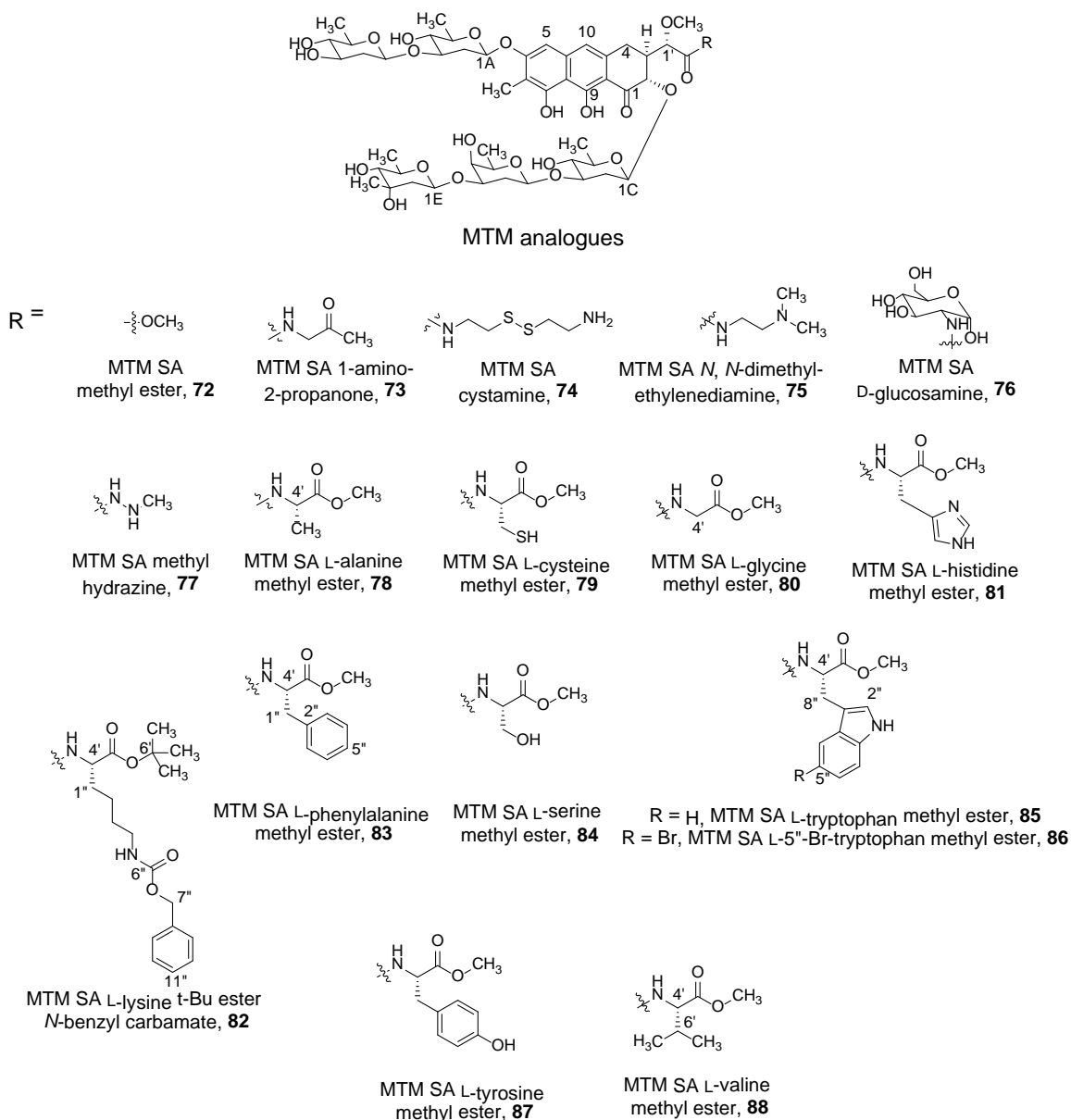


Figure 27. MTM analogues derived from MTM SA

Cytotoxicity assays

The cytotoxicity of some MTM analogues, including structure **73**, **74**, **75**, **77**, **78**, **79**, **80**, **88**, was tested against the A549 non-small cell lung cancer cell line in Dr. Younsou Bae's laboratory at the University of Kentucky's College of Pharmacy (**Table 4**). And, MTM SK (**67**) and MTM SA (**69**) were used as controls for comparison. This bioassay revealed that modifying the C-3 side chain of MTM SA with amino acids, L-alanine, L-

cysteine, L-glycine, and L-valine, could significantly improve the anticancer bioactivity. IC₅₀ values of these four MTM analogues (**78**, **79**, **80**, **88**) are comparable to MTM SK (**67**), but the bioactivity of structure **79** is not as impressive as other analogues. The lower bioactivity of structure **79** is probably due to the polar thiol side chain on the cysteine (pKa = 8.33). Our hypothesis is that the modification of the side chain with non-polar amino acids can yield products with better anticancer activity.

Table 4. The IC₅₀ values of MTM analogues against the A549 cell line *in vitro* (This experiment was carried out by Dr. Daniel Scott in Dr. Bae's laboratory.)

MTM analogues	IC ₅₀ (μM)
MTM SK (67)	0.28 ± 0.11
MTM SA (69)	8.70 ± 0.37
73	N/A ^a
74	N/A ^a
75	2.70 ± 0.22
77	8.80 ± 1.69
78	0.36 ± 0.12
79	1.00 ± 0.19
80	0.55 ± 0.11
88	0.80 ± 0.20

^a Not active.

In order to confirm this hypothesis, we modified the C-3 side chain of MTM SA with more primary amines and also carried out esterification to convert the carboxylic group to an ester (**72**, **76**, **81**, **82**, **83**, **84**, **85**, **86**, and **87**). The amino acids used here were L-histidine, L-lysine, L-phenylalanine, L-serine, L-tryptophan, and L-tyrosine containing protecting group(s) to prevent the coupling reaction occurring at undesired position(s). All products, except MTM SA-5''-Br-Trp (**86**) were shipped to the National Cancer Institute (NCI) at the National Institutes of Health (NIH) for 60-cell line screens (one-dose) in order to identify their medicinal potentials (**Table 5-1** and **5-2**).^[58]

The number reported for the one-dose assay is the cell growth percent of treated cells relative to both the no-drug control and the time zero number of cells. This allows scientists to measure growth inhibition (values between 0% and 100%) as well as lethality (values less than 0%) of their compounds. According to the data of NCI 60 cell line screens (one-dose), the compounds showing the lowest values of the average cell growth percent are MTM SA methyl ester (**72**), MTM SA-Lys-2PGs (**82**), MTM SA-Phe (**83**), and MTM SA-Trp (**85**) (**Figure 28**). The number of the cell growth percent of structure **72**, **83**, and **85** are two folds lower than the parent compound, MTM. While structure **82**, the most promising MTM analogue, is nearly six folds lower than MTM. The anticancer activity of MTM SA-Ala (**78**) is comparable with MTM, but modification of MTM SA with D-glucosamine, L-glycine, L-tyrosine, L-valine (structures **76**, **80**, **87**, and **88**) could not make the products stronger than the parent compound. MTM SA-His (**81**) is the poorest MTM analogue but a good example that supports our hypothesis: modifying the C-3 side chain of MTM SA with non-polar amino acids can generate MTM analogues with better anticancer activity.

Table 5-1. The cell growth percent (%) data of National Cancer Institute (NCI) 60 cell line screen (one-dose) of MTM^a and its analogues. (This experiment was carried out by NCI researchers in NCI.)

	MTM (11)	MTMSA (69)	72	76	78	80
Leukemia cell lines						
CCRF-CEM	3.90	23.77	2.91	2.98	14.48	17.99
HL-60(TB)	3.90	41.33	-33.54	19.37	20.80	41.02
K-562	6.80	11.75	4.16	-4.31	7.10	10.20
MOLT-4	6.20	1.78	-9.54	22.09	0.97	-2.56
RPMI-8226	-14.50	6.67	-17.50	5.17	-5.91	2.86
SR	10.50	10.36	4.74	-34.07	7.81	11.13
Non-small cell lung cancer						
A549/ATCC	2.70	13.73	8.71	8.35	7.05	7.82
EKVX	16.80	31.30	9.91	4.23	21.86	20.21
HOP-62	7.50	20.65	6.33	-11.35	12.98	21.63
HOP-92	-1.50	24.74	-15.60	10.78	3.12	9.61
NCI-H226	10.80	0.86	-41.85	-9.00	-21.48	-10.01
NCI-H23	-28.90	-0.12	-47.33	-3.19	-31.12	-16.76
NCI-H322M	5.90	21.59	12.34	-1.78	11.42	15.73
NCI-H460	2.60	11.19	4.95	-10.85	7.28	9.15
NCI-H522	-44.70	-3.78	-30.26	-11.69	-26.04	-19.00
Colon cancer						
COLO 205	-67.50	-12.53	-37.16	-0.46	-20.51	-17.64
HCC-2998	N/A	-37.76	-32.87	-4.02	-43.97	-42.26
HCT-116	5.30	6.04	3.51	-4.14	4.37	3.52
HCT-15	6.70	38.68	7.16	2.16	19.35	31.45
HT29	0.20	7.12	3.61	4.31	3.99	3.78
KM12	-5.70	7.07	1.50	-7.56	5.08	6.86
SW-620	-25.10	19.12	8.04	-0.61	10.15	12.69
CNS cancer						
SF-268	12.30	29.26	3.86	-10.59	14.07	21.14
SF-295	19.10	11.22	-10.64	-19.08	3.96	6.85
SF-539	-20.20	-23.07	-41.55	0.56	-33.18	-28.66
SNB-19	13.30	10.82	4.99	-12.15	8.48	9.40
SNB-75	-56.90	-55.12	-79.32	-23.55	-63.81	-51.74
U251	9.40	10.09	2.62	-17.57	7.45	9.93
Melanoma						
LOX IMVI	7.00	-6.05	-23.96	-7.56	-27.43	-15.50
MALME-3M	-54.60	-10.40	-58.02	5.81	-49.31	-35.34
M14	-71.50	-8.35	-71.50	5.65	-23.86	-32.07

MDA-MB-435	-62.90	1.21	-62.45	7.66	-39.33	-39.22
SK-MEL-28	N/A	-5.84	-38.04	6.13	-37.75	-17.16
SK-MEL-5	N/A	-17.25	-69.51	-29.80	-67.36	-44.60
SK-MEL-2	-43.90	N/A	N/A	11.21	N/A	N/A
UACC-257	4.40	7.77	-11.21	9.17	-0.39	6.03
UACC-62	-58.60	-68.10	-58.50	-7.62	-52.83	-65.19
Ovarian cancer						
IGROV1	-21.00	35.43	28.01	-10.18	22.76	31.99
OVCAR-3	3.00	10.04	6.49	8.29	5.02	3.71
OVCAR-4	8.20	19.69	3.96	4.39	0.16	2.95
OVCAR-5	-54.90	17.55	18.76	11.31	13.54	16.44
OVCAR-8	N/A	13.55	10.22	-5.61	9.19	12.40
NCI/ADR-RES	44.40	74.05	42.74	10.03	75.42	88.20
SK-OV-3	35.30	23.66	21.00	-3.70	22.63	24.30
Renal cancer						
786-0	2.90	10.27	9.07	-1.29	-4.86	-2.99
A498	N/A	-55.67	-42.88	3.77	-66.47	-74.31
ACHN	3.70	2.89	4.79	1.15	-1.98	-1.39
CAKI-1	17.20	25.66	14.83	-1.65	27.50	43.54
RXF 393	5.90	-57.11	-72.98	-4.38	-70.26	-75.43
SN12C	24.80	15.32	9.04	-10.66	11.49	13.69
TK-10	9.90	24.24	13.66	1.54	13.10	21.04
UO-31	-4.40	8.85	-9.90	-13.47	5.72	10.90
Prostate cancer						
PC-3	10.50	19.70	-23.33	-7.96	8.21	12.93
DU-145	24.70	21.53	10.17	7.19	13.20	17.16
Breast cancer						
MCF7	-3.60	7.44	8.90	-12.84	8.55	7.64
MDA-MB-231/ATCC	14.60	19.95	-8.44	-0.80	18.36	21.50
HS 578T	N/A	-1.09	-10.53	-26.09	-7.37	-6.65
BT-549	-19.90	-41.87	-32.91	-5.36	-32.48	-44.27
T-47D	7.20	12.84	4.00	-4.44	2.98	11.29
MDA-MB-468	N/A	-16.72	-45.55	-20.04	-36.22	-22.30
Mean	-5.52	5.25	-12.57	-3.10	-5.33	-0.79
Delta	65.98	73.35	66.75	30.97	64.93	74.64
Range	115.90	142.15	122.06	56.16	145.68	163.63

^a Tested independently in September, 2014

Table 5-2. The cell growth percent (%) data of National Cancer Institute (NCI) 60 cell line screen (one-dose) of MTM^a and its analogues. (This experiment was carried out in NCI and their researchers.)

	81	82	83	85	87	88
Leukemia cell lines						
CCRF-CEM	50.11	3.09	4.07	5.66	11.08	7.92
HL-60(TB)	53.14	-0.25	-9.07	-8.66	8.20	52.39
K-562	32.09	5.35	5.63	4.37	7.28	5.81
MOLT-4	73.20	-12.61	-5.77	0.76	-1.80	1.35
RPMI-8226	58.17	-5.01	-16.52	-5.61	-5.59	-1.18
SR	36.73	7.17	7.89	4.93	9.70	-8.16
Non-small cell lung cancer						
A549/ATCC	67.64	-25.72	6.81	12.68	4.32	7.37
EKVX	57.59	-25.08	-2.04	3.61	21.48	25.26
HOP-62	46.09	-15.01	4.37	9.74	19.30	6.44
HOP-92	33.97	-82.64	-17.47	-13.87	12.24	-18.92
NCI-H226	50.03	-75.53	-32.96	-38.58	-17.59	-23.39
NCI-H23	34.36	-42.60	-28.50	-20.32	-38.60	-6.89
NCI-H322M	78.18	20.22	7.54	6.45	11.13	24.78
NCI-H460	37.98	-17.59	4.92	6.58	6.99	-6.19
NCI-H522	16.36	-16.04	-30.70	-28.15	-19.54	-20.03
Colon cancer						
COLO 205	24.26	-20.68	-36.70	-39.47	-22.83	13.78
HCC-2998	50.26	-5.55	-45.16	-46.79	-66.97	11.21
HCT-116	40.52	-63.94	1.85	3.71	4.11	4.79
HCT-15	95.18	26.60	5.61	15.91	47.47	23.91
HT29	59.30	2.05	4.24	4.98	3.68	6.32
KM12	58.78	-62.67	-5.45	0.36	5.68	4.63
SW-620	57.51	3.71	6.79	6.98	11.29	16.61
CNS cancer						
SF-268	70.38	-61.51	-2.00	-10.10	20.74	6.33
SF-295	46.78	-89.48	-9.46	-6.35	7.09	-32.53
SF-539	44.05	-32.03	-35.68	-35.30	-27.11	-22.44
SNB-19	50.81	-51.57	4.82	7.09	8.72	8.43
SNB-75	27.11	-94.27	-86.82	-88.46	-52.49	-24.30
U251	38.30	-82.26	2.62	6.13	7.18	-28.07
Melanoma						
LOX IMVI	13.73	-62.85	-18.20	-3.17	-31.25	-76.22
MALME-3M	44.68	-69.38	-54.85	-47.04	-42.98	-12.28
M14	69.79	-69.80	-60.78	-60.18	-20.22	5.18

MDA-MB-435	47.43	-69.79	-69.37	-71.91	-26.78	18.86
SK-MEL-28	41.47	-55.72	-27.03	-27.90	-26.62	-1.86
SK-MEL-5	24.86	-97.67	-78.02	-72.10	-53.27	-74.28
SK-MEL-2	N/A	N/A	N/A	N/A	N/A	N/A
UACC-257	78.78	-5.47	-16.21	-5.00	2.42	3.21
UACC-62	23.32	-62.69	-56.64	-52.75	-53.13	-65.12
Ovarian cancer						
IGROV1	75.11	22.18	25.50	23.41	31.17	32.34
OVCAR-3	66.63	24.77	5.70	4.79	5.52	15.26
OVCAR-4	101.60	43.15	1.44	0.34	7.27	13.44
OVCAR-5	91.60	-77.93	8.60	15.26	10.58	2.67
OVCAR-8	64.91	12.90	7.42	11.00	9.74	12.60
NCI/ADR-RES	106.09	73.20	51.64	76.71	85.64	69.65
SK-OV-3	74.85	-52.69	17.05	18.22	20.74	20.42
Renal cancer						
786-0	63.78	-89.85	0.65	1.54	1.98	-21.25
A498	25.89	-89.88	-40.34	-52.77	-63.17	-83.20
ACHN	68.91	-2.61	-8.69	0.57	4.06	0.32
CAKI-1	90.64	22.64	15.34	19.53	69.45	20.04
RXF 393	54.59	-85.93	-75.04	-85.94	-71.54	-67.98
SN12C	58.21	-7.42	4.42	10.57	13.27	13.20
TK-10	55.81	-9.57	6.76	8.12	23.48	21.70
UO-31	79.65	10.50	-7.79	-6.63	18.06	7.39
Prostate cancer						
PC-3	61.03	3.14	-9.57	-10.39	13.44	18.77
DU-145	55.43	26.52	11.82	12.53	15.76	7.21
Breast cancer						
MCF7	35.00	3.56	3.21	5.98	8.15	6.73
MDA-MB-231/ATCC	71.82	-71.80	-8.99	-14.03	19.52	10.86
HS 578T	34.10	-16.42	-16.74	-13.57	-9.20	-13.93
BT-549	60.60	-79.72	-62.36	-53.44	-36.92	-20.69
T-47D	49.44	7.20	2.51	8.07	4.63	10.17
MDA-MB-468	72.66	-37.55	-36.75	-48.83	-25.52	-3.89
Mean	55.11	-28.45	-13.26	-11.03	-2.04	-1.62
Delta	41.38	69.22	73.56	77.43	69.50	81.58
Range	92.36	170.87	138.46	165.17	157.18	152.85

^a Tested independently in September, 2014

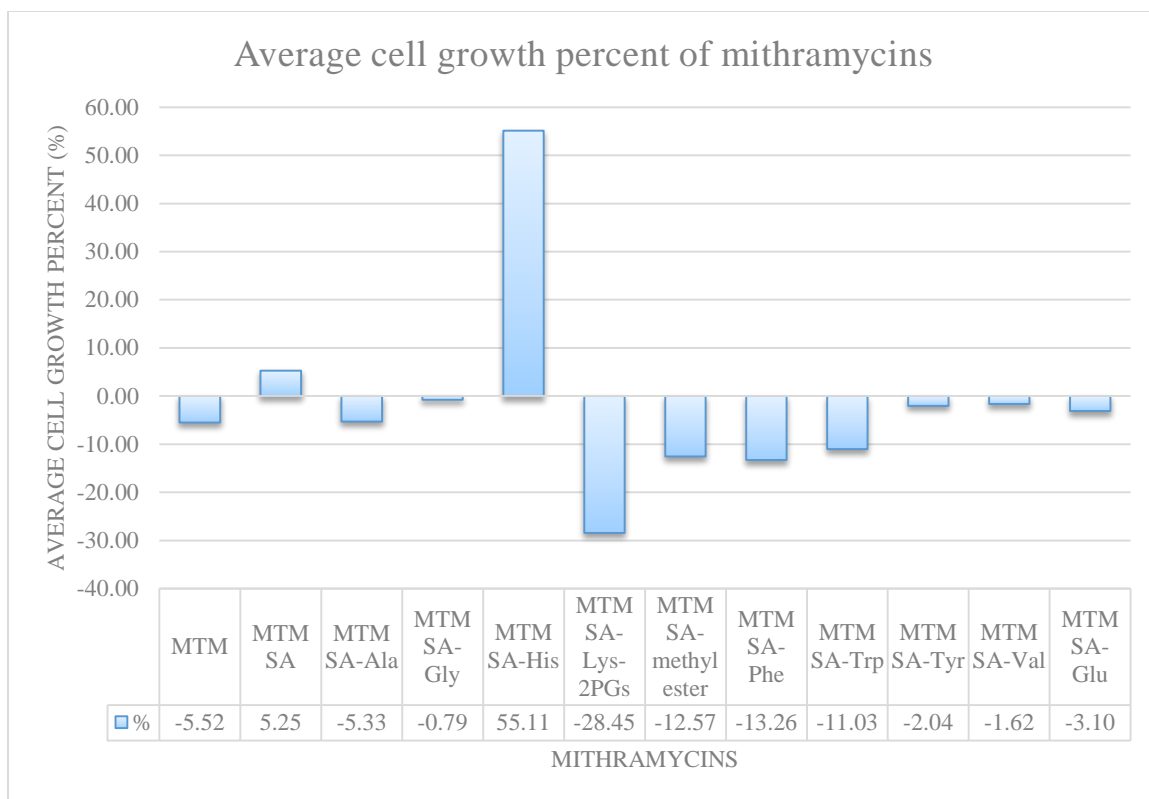


Figure 28. The averages of cell growth percent (%) of MTM and MTM analogues. This column chart is edited based on the data of NCI 60 cell line screen (one-dose). (All 60 cell line screen was carried out by NCI researchers in NCI.)

Structure elucidation of the mithramycin analogues

We did structure elucidation with HR-MS and NMR for ten bioactive MTM analogues (structure **72**, **78**, **80**, **81**, **82**, **83**, **85**, **86**, **87**, and **88**), which showed better anticancer bioactivity than the parent compound, MTM. HR-MS was performed by the University of Kentucky Mass Spectrometry Facility. All observed HR-MS data of MTM analogues matched the calculated molecular weight values and, thereby, were consistent with the occurrence of the coupling reaction with the desired primary amines and the esterification reaction with a methyl group (**Table 6**).

Table 6. HR-MS data of bioactive MTM analogues^a (HR-MS analysis was carried out by MS facilities at the University of Kentucky)

MTM analogue	Molecular formula (HR-MS)	Calculated M.W.	Observed M.W.
72 ^b	[C ₅₀ H ₇₂ O ₂₃ Na] ⁺	1063.4362	1063.4367
78 ^c	[C ₅₃ H ₇₇ NO ₂₄ Na] ⁺	1134.4733	1134.4745
80 ^c	[C ₅₂ H ₇₅ NO ₂₄ Na] ⁺	1120.4577	1120.4558
81 ^c	[C ₅₆ H ₇₉ N ₃ O ₂₄ Na] ⁺	1200.4951	1200.4970
82 ^c	[C ₆₇ H ₉₆ N ₂ O ₂₆ Na] ⁺	1367.6149	1367.6104
83 ^c	[C ₅₉ H ₈₁ NO ₂₄ Na] ⁺	1210.5046	1210.5050
85 ^c	[C ₆₁ H ₈₂ N ₂ O ₂₄ Na] ⁺	1249.5155	1249.5181
86 ^b	[C ₆₁ H ₈₀ N ₂ O ₂₄ Br] ⁻	1303.4284	1303.4283
87 ^c	[C ₅₉ H ₈₁ NO ₂₅ Na] ⁺	1226.4995	1226.5005
88 ^c	[C ₅₅ H ₈₁ NO ₂₄ Na] ⁺	1162.5046	1162.5037

^a Only bioactive MTM analogues were checked with HR-MS.

^b HR-ESI-MS, negative mode, [M – H]⁻

^c HR-MALDI-TOF-MS, positive mode, [M + Na]⁺

The structures of the most active MTM analogues, including structure **72**, **78**, **80**, **82**, **83**, **85**, **86**, and **88**, were further confirmed through ¹H, gHSQC, and gHMBC NMR. The NMR signals of the coupling primary amines and methyl group are marked with bold text and also colored in red. All NMR spectra were recorded by 500 MHz Agilent NMR as described in Chapter 2.5.

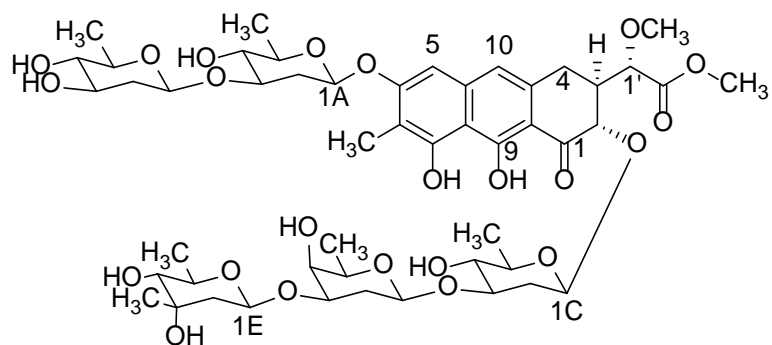


Table 7. NMR data (500 MHz) of MTM SA methyl ester (**72**)^a

Position	δ_H (J in Hz)	δ_C^b , mult. ^c	HMBC
1		(203.5, C) ^d	
2	4.66, d (14.0)	77.7, CH	
3	2.47 – 2.55 (m, 1 H)	45.1, CH	
4	2.64, dd (3.8, 16.3)	28.3, CH ₂	
	3.05, d (17.5)		
4a		(136.5, C) ^d	
5	6.70, s	102.0, CH	112.0, 117.8,
6		160.2, C	
7		112.0, C	
7-CH ₃	2.13, s	8.5, CH ₃	112.0, 156.4, 160.2
8		156.4, C	
8a		108.8, C	
9		(164.5, C) ^d	
9a		109.2, C	
10	6.81, s	117.8, CH	28.3, 102.0
10a		(139.6, C) ^d	
1'	4.34, d (2.0)	78.8, CH	28.3, 45.1, 59.5, 174.2
1'-OCH ₃	3.49, s	59.5, CH ₃	78.8
2'		174.2, C	
2'-OCH₃	3.81, s	52.5, CH₃	174.2
1A	5.30, d (9.0)	97.6, CH	
2A	1.82, m, (12.0)	37.8, CH ₂	
	2.44 (m, 1H)		
3A	3.73 – 3.79 (m, 2H)	80.6, CH	
4A	3.10, t (9.0)	76.2, CH	18.3, 80.6
5A	3.50 – 3.59 (m, 2H)	72.0, CH	
6A	1.35, d (6.5)	18.3, CH ₃	76.2
1B	4.73, dd (1.5, 9.5)	99.6, CH	
2B	1.53 – 1.64 (m, 3H)	40.7, CH ₂	99.6
	2.19, m (1.5, 5.0, 12.5)		78.1

3B	3.69 – 3.73 (m, 2H)	71.9, CH	
4B	2.97, d (9.0)	78.1, CH	18.1
5B	3.32 – 3.39, (m, 2H)	73.6, CH	
6B	1.32, d (5.5)	18.1, CH ₃	73.6, 78.1
1C	5.06, d (9.0)	101.8, CH	
2C	1.53 – 1.64 (m, 3H) 2.59, dd (4.5, 12.0)	38.1, CH ₂	
3C	3.73 – 3.79 (m, 2H)	80.8, CH	
4C	3.05, t (8.8)	76.7, CH	18.5, 73.4, 80.8
5C	3.50 – 3.59 (m, 2H)	73.4, CH	
6C	1.34, d (5.5)	18.5, CH ₃	73.4
1D	4.68, dd (2.0, 10.5)	99.8, CH	
2D	1.80, m (12.0) 1.96 (m, 1H)	33.1, CH ₂	
3D	3.88, m (3.0, 4.8, 12.3)	77.2, CH	
4D	3.69 – 3.73 (m, 2H)	70.5, CH	77.2
5D	3.63 – 3.69 (m, 1H)	72.0, CH	70.5
6D	1.31, d (6.5)	16.8, CH ₃	70.5
1E	4.98, dd (2.0, 9.5)	98.8, CH	77.2
2E	1.53 – 1.64 (m, 3H) 1.92, dd (2.0, 13.5)	45.2, CH ₂	98.8 71.7, 77.7
3E		71.7, C	
3E-CH ₃	1.25, s	27.3, CH ₃	45.2, 71.7, 77.7
4E	2.93, d (9.5)	77.7, CH	18.8, 71.7
5E	3.32 – 3.39 (m, 2H)	73.1, CH	
6E	1.27, d (6.5)	18.8, CH ₃	77.7

^a The solvent used for NMR experiments was methanol-*d*₄.

^b ¹³C NMR data was inferred from gHSQC and gHMBC spectra.

^c multiplicity

^d The numbers in the bracket are cited from the reference.^[59] We did only ¹H, gHSQC, and gHMBC experiments for this compound, and no C-H correlations in the gHMBC spectrum which can indicate the chemical shift values of these carbons.

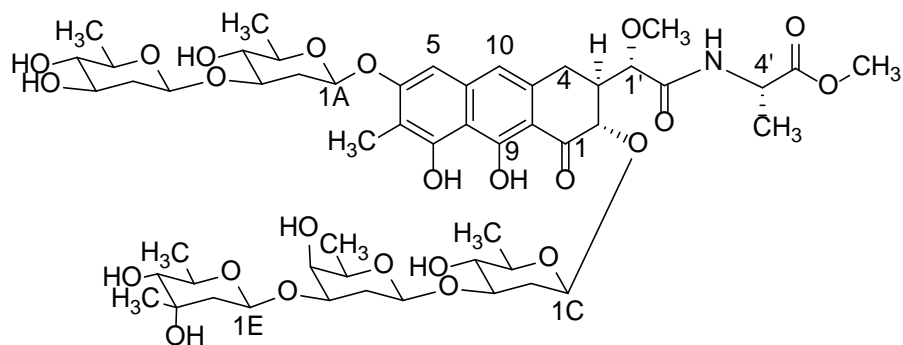


Table 8. NMR data (500 MHz) of MTM SA-Ala (**78**)^a

Position	δ_{H} (<i>J</i> in Hz)	δ_{C} , mult. ^b	HMBC
1		204.2, C	
2	4.42, d (11.5)	78.5, CH	
3	2.45 – 2.56 (m, 2H)	45.0, CH	
4	2.45 – 2.56 (m, 2H) 2.60 – 2.73 (m, 1H)	28.3, CH ₂	
4a		136.4, C	
5	6.34, s	102.1, CH	108.5, 111.8, 118.1
6		160.1, C	
7		111.8, C	
7-CH ₃	2.09, s	8.9, CH ₃	
8		156.7, C	
8a		108.5, C	
9		164.9, C	
9a		109.0, C	
10	6.45, s	118.1, CH	28.3, 102.1, 108.5, 109.0
10a		139.6, C	
1'	4.13, s	81.0, CH	28.3, 45.0, 60.4, 78.5, 174.1
1'-OCH ₃	3.59 (bs, 4H)	60.4, CH ₃	81.0
2'		174.1, C	
4'	4.56, m (7.0)	49.8, CH	17.8, 174.1, 174.4
4'-CH₃	1.52, d (7.0)	17.8, CH₃	49.8, 174.4
5'		174.4, C	
5'-OCH₃	3.75, s	53.1, CH₃	174.4
1A	4.94 – 5.13 (m, 3H)	97.6, CH	
2A	1.75 – 1.87 (m, 2H) 2.36 – 2.45 (m, 1H)	38.2, CH ₂	100.1
3A	3.77 – 3.84 (m, 2H)	81.0, CH	
4A	3.03 – 3.12 (m, 2H)	76.5, CH	18.8, 73.4, 81.0
5A	3.44 – 3.52 (m, 1H)	73.4, CH	
6A	1.33, d (6.5)	18.8, CH ₃	73.4, 76.5

1B	4.65 – 4.78 (m, 2H)	100.1, CH	
2B	1.55 – 1.66 (m, 3H)	40.9, CH ₂	
	2.18 – 2.26 (m, 1H)		72.2, 78.3
3B	3.59 (bs, 4H)	72.2, CH	
4B	2.98, t (9.0)	78.3, CH	18.4, 72.2, 73.8
5B	3.36 – 3.44 (m, 2H)	73.8, CH	
6B	1.34, d (6.5)	18.4, CH ₃	73.8, 78.3
1C	4.94 – 5.13 (m, 3H)	102.4, CH	
2C	1.55 – 1.66 (m, 3H)	38.5, CH ₂	100.3
	2.60 – 2.73 (m, 1H)		
3C	3.77 – 3.84 (m, 2H)	81.2, CH	
4C	3.03 – 3.12 (m, 2H)	76.7, CH	19.0, 73.6, 81.2
5C	3.36 – 3.44 (m, 2H)	73.6, CH	
6C	1.39, d (5.5)	19.0, CH ₃	73.6, 76.7
1D	4.65 – 4.78 (m, 2H)	100.3, CH	81.2
2D	1.75 – 1.87 (m, 2H)	33.3, CH ₂	
	1.96 – 2.02 (m, 1H)		
3D	3.86 – 3.93 (m, 1H)	77.4, CH	
4D	3.68 – 3.72 (bs, 2H)	70.6, CH	17.3, 33.3, 77.4
5D	3.65 (m, 1H)	72.2, CH	70.6
6D	1.31, d (6.5)	17.3, CH ₃	70.6
1E	4.94 – 5.13 (m, 3H)	99.1, CH	77.4
2E	1.55 – 1.66 (m, 3H)	45.4, CH ₂	99.1
	1.93, d (13.0)		71.9, 78.1, 99.1
3E		71.9, C	
3E-CH ₃	1.25, s	27.4, CH ₃	45.4, 71.9, 78.1, 99.1
4E	2.93, d (10.0)	78.1, CH	18.9, 27.4, 71.9, 72.0
5E	3.68 – 3.72 (bs, 2H)	72.0, CH	
6E	1.27, d (6.5)	18.9, CH ₃	72.0, 78.1

^a The solvent used for NMR experiments was methanol-*d*₄.

^b multiplicity

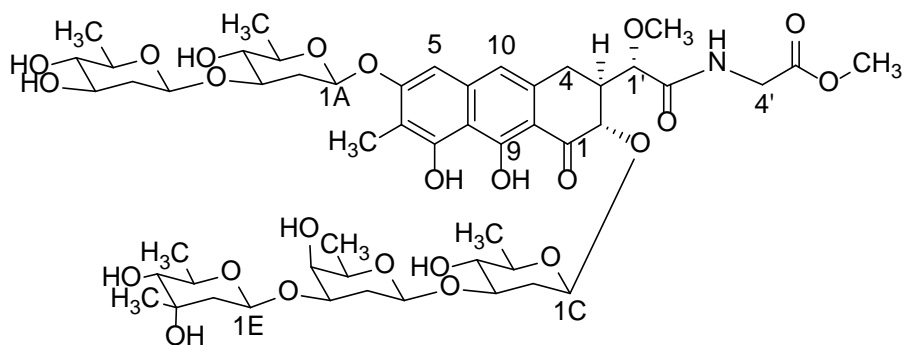


Table 9. NMR data (500 MHz) of MTM SA-Gly (**80**)^a

Position	δ_{H} (<i>J</i> in Hz)	δ_{C} , mult. ^b	HMBC
1		216.2, C	
2	4.36 (bs, 1H)	78.7, CH	
3	2.44 – 2.58 (m, 2H)	45.2, CH	
4	2.44 – 2.58 (m, 2H) 2.82, d (11.0)	28.1, CH ₂	
4a		136.9, C	
5	6.22, s	102.0, CH	
6		159.8, C	
7		111.6, C	
7-CH ₃	2.10, bs	8.9, CH ₃	
8		156.7, C	
8a		108.6, C	
9		164.1, C	
9a		109.1, C	
10	6.48, s	118.2, CH	102.0, 108.6, 109.1
10a		139.5, C	
1'	4.13 (bs, 2H)	81.2, CH	28.1, 45.2, 60.4, 175.2
1'-OCH ₃	3.56 (bs, 4H)	60.4, CH ₃	81.2
2'		175.2, C	
4'	3.99, d (17.5) 4.13 (bs, 2H)	41.8, CH	175.2
5'		171.9, C	
5'-OCH₃	3.80 (bs, 5H)	52.9, CH₃	171.9
1A	4.87 (overlap, 1H)	97.3, CH	
2A	1.74 – 1.82 (m, 2H) 2.36, bs	38.3, CH ₂	100.0
3A	3.80 (bs, 5H)	80.9, CH	
4A	3.07, t (8.5), 2H	76.5, CH	73.3, 80.9
5A	3.36 – 3.48 (m, 3H)	73.3, CH	
6A	1.32, d (6.0), 2H	18.8, CH ₃	73.3, 76.5

1B	4.63 – 4.80 (m, 2H)	100.0, CH	
2B	1.53 – 1.68 (m, 3H) 2.22 (m, 1H)	40.9, CH ₂	
3B	3.56 (bs, 4H)	72.2, CH	
4B	2.98, t (8.8)	78.2, CH	18.4, 72.2, 73.9
5B	3.36 – 3.48 (m, 3H)	73.9, CH	
6B	1.35, d (6.0)	18.4, CH ₃	73.9, 78.2
1C	4.97 – 5.15 (m, 2H)	102.5, CH	
2C	1.53 – 1.68 (m, 3H) 2.62, d (7.5)	38.4, CH ₂	
3C	3.80 (bs, 5H)	81.2, CH	
4C	3.07, t (8.5), 2H	76.7, CH	81.2
5C	3.36 – 3.48 (m, 3H)	73.6, CH	
6C	1.42, bs	18.9, CH ₃	73.6
1D	4.63 – 4.80 (m, 2H)	100.3, CH	
2D	1.74 – 1.82 (m, 2H) 1.89 – 2.01 (m, 2H)	33.3, CH ₂	
3D	3.89, d (11.5)	77.5, CH	
4D	3.71 (bs, 2H)	70.6, CH	77.5
5D	3.63 – 3.68 (m, 1H)	72.2, CH	
6D	1.32, d (6.0), 2H	17.3, CH ₃	70.6, 72.2
1E	4.97 – 5.15 (m, 2H)	99.1, CH	77.5
2E	1.53 – 1.68 (m, 3H) 1.89 – 2.01 (m, 2H)	45.4, CH ₂	99.1 71.9, 78.1
3E		71.9, C	
3E-CH ₃	1.25, s	27.4, CH ₃	45.4, 71.9, 72.0, 78.1
4E	2.93, d (9.5)	78.1, CH	18.9, 72.0
5E	3.71 (bs, 2H)	72.0, CH	
6E	1.27, d (6.5)	18.9, CH ₃	71.9, 72.0, 78.1

^a The solvent used for NMR experiments was methanol-*d*₄.

^b multiplicity

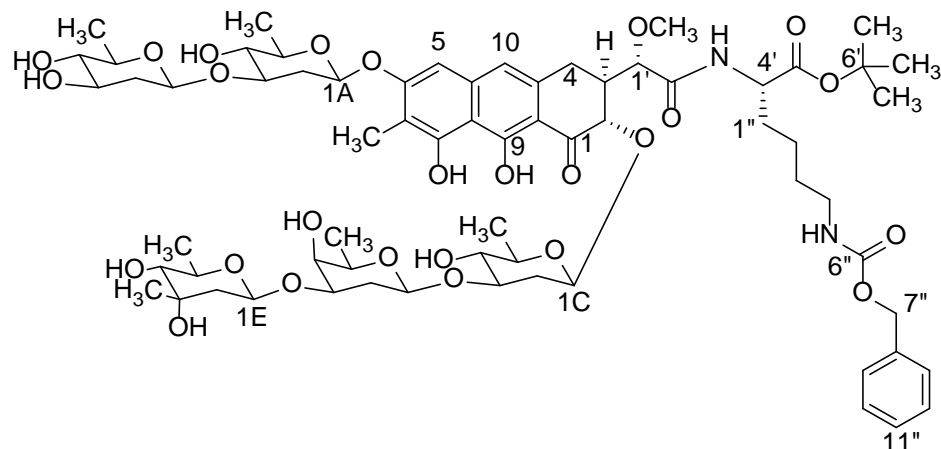


Table 10. NMR data (500 MHz) of MTM SA-Lys-2PGs (**82**)^a

Position	δ_{H} (<i>J</i> in Hz)	$\delta_{\text{C}}^{\text{b}}$, mult. ^c	HMBC
1		(203.5, C) ^d	
2	4.74, d (12.0)	77.9, CH	45.2, 102.0
3	2.58, d (10.5)	45.2, CH	
4	2.72 – 2.80, m (overlap) 2.92 – 3.08 (m, 5H)	28.8, CH ₂	136.2
4a		136.2, C	
5	6.92, bs (overlap)	102.9, CH	28.8, 102.9, 108.9, 109.6, 112.0, 118.2, 140.2, 160.9
6		160.9, C	
7		112.0, C	
7-CH ₃	2.14, s	8.4, CH ₃	112.0, 157.2, 160.9
8		157.2, C	
8a		108.9, C	
9		(164.5, C) ^d	
9a		109.6, C	
10	6.92, bs (overlap)	118.2, CH	28.8, 102.9, 108.9, 109.6, 112.0, 118.2, 140.2, 160.9
10a		140.2, C	
1'	4.17, d (1.5)	81.3, CH	28.8, 45.2, 60.8, 77.9, 172.5
1'-OCH ₃	3.62, s	60.8, CH ₃	81.3
2'		172.5, C	
4'	3.22 – 3.38 (m, 3H)	42.1, CH	157.9
5'		157.9, C	
6'		82.8, C	
6'-CH₃	1.48, s	29.0, CH₃	29.0, 82.8
1''	1.25 – 1.35 (m, 15H) 1.60 – 1.68 (m, 2H)	31.3, CH₂	31.0, 70.0, 72.2, 73.8, 74.0, 76.6, 78.7
2''	1.50 – 1.60 (m, 4H)	25.0, CH₂	33.0
3''	1.25 – 1.35 (m, 15H)	31.0, CH₂	31.0, 70.0, 72.2, 73.8, 74.0, 76.6, 78.7
4''	1.87 – 1.98 (m, 4H)	33.0, CH₂	
6''		158.0, C	

7"	5.05, s	67.0, CH ₂	129.1, 129.3, 129.9, 139.1, 158.0
8"		139.1, C	
9"	7.22 – 7.40 (m, 5H)	129.1, CH	129.1, 139.1
10"	7.22 – 7.40 (m, 5H)	129.9, CH	129.1, 139.1
11"	7.22 – 7.40 (m, 5H)	129.3, CH	129.1, 139.1
1A	5.40, d (9.8)	97.8, CH	
2A	1.74 – 1.86 (m, 2H) 2.43, dd (5.5, 12.5)	38.7, CH ₂	
3A	3.69 – 3.78 (m, 2H)	82.5, CH	
4A	3.04, t (9.3)	76.6, CH	19.1, 73.8, 82.5
5A	3.52 – 3.61 (m, 2H)	73.8, CH	
6A	1.25 – 1.35 (m, 15H)	19.1, CH ₃	31.0, 70.0, 72.2, 73.8, 74.0, 76.6, 78.7
1B	4.63 – 4.72 (m, 2H)	101.1, CH	82.5
2B	1.50 – 1.60 (m, 4H) 2.17, ddd (2.0, 4.5, 12.5)	41.0, CH ₂	101.1
3B	3.52 – 3.61 (m, 2H)	72.5, CH	
4B	2.97, t (8.3)	78.7, CH	18.9, 72.5
5B	3.22 – 3.38 (m, 3H)	74.0, CH	
6B	1.25 – 1.35 (m, 15H)	18.9, CH ₃	31.0, 70.0, 72.2, 73.8, 74.0, 76.6, 78.7
1C	5.14, d (9.0)	102.0, CH	77.9
2C	1.60 – 1.68 (m, 3H) 2.52, dd (4.8, 12.8)	39.0, CH ₂	82.9, 102.0 82.9
3C	3.63 – 3.69 (m, 2H)	82.9, CH	
4C	3.01, t (8.8)	76.8, CH	19.1, 82.9
5C	3.22 – 3.38 (m, 3H)	74.0, CH	
6C	1.25 – 1.35 (m, 15H)	19.1, CH ₃	31.0, 70.0, 72.2, 73.8, 74.0, 76.6, 78.7
1D	4.63 – 4.72 (m, 2H)	101.4, CH	82.9
2D	1.74 – 1.86 (m, 2H) 1.87 – 1.98 (m, 4H)	33.7, CH ₂	77.9, 101.4
3D	3.90, ddd (3.0, 5.0, 12.0)	77.9, CH	
4D	3.69 – 3.78 (m, 3H)	70.0, CH	17.8, 77.9
5D	3.63 – 3.69 (m, 2H)	72.2, CH	
6D	1.25 – 1.35 (m, 15H)	17.8, CH ₃	31.0, 70.0, 72.2, 73.8, 74.0, 76.6, 78.7
1E	4.97, dd (1.8, 9.8)	99.1, CH	45.5, 77.9
2E	1.50 – 1.60 (m, 4H) 1.87 – 1.98 (m, 4H)	45.5, CH ₂	99.1 71.9, 77.9
3E		71.9, C	
3E-CH ₃	1.22, s	28.2, CH ₃	45.5, 71.9, 72.2, 77.9
4E	2.95, d (9.0)	77.9, CH	
5E	3.69 – 3.78 (m, 3H)	72.2, CH	77.9
6E	1.23, d (5.5)	19.5, CH ₃	45.5, 71.9, 72.2, 77.9

^a The solvent used for NMR experiments was acetone-*d*₆.

^b ¹³C NMR data was inferred from gHSQC and gHMBC spectra.

^c multiplicity

^d The numbers in the bracket are cited from the reference.^[59] We did only ¹H, gHSQC, and gHMBC experiments for this compound, and no C-H correlations in the gHMBC spectrum which can indicate the chemical shift values of these carbons.

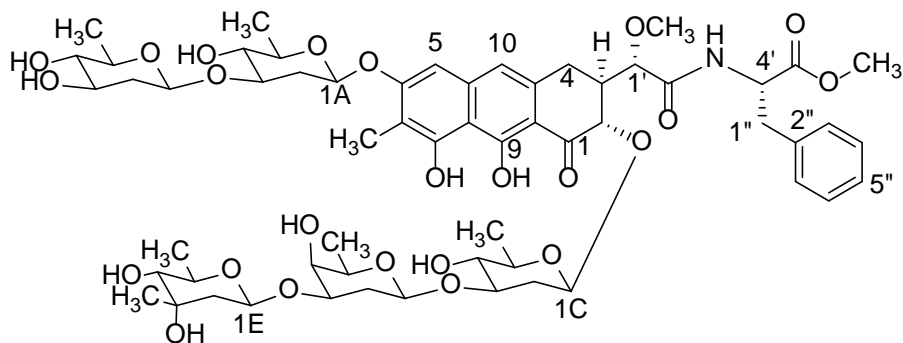


Table 11. NMR data (500 MHz) of MTM SA-Phe (**83**)^a

Position	δ_{H} (<i>J</i> in Hz)	$\delta_{\text{C}}^{\text{b}}$, mult. ^c	HMBC
1		(203.5, C) ^d	
2	4.39, t (11.5)	78.3, CH	44.8, 102.2
3	2.29 – 2.48 (m, 3H)	44.8, CH	
4	2.00 – 2.08 (m, 1H) 2.29 – 2.48 (m, 3H)	27.9, CH ₂	78.3, 108.9, 136.7
4a		136.7, C	
5	6.43, bd (1H, <i>J</i> = 22)	101.9, CH	108.4, 111.8, 118.0, 160.4
6		160.4, C	
7		111.8, C	
7-CH ₃	2.09, s	8.9, CH ₃	111.8, 156.9, 160.4
8		156.9, C	
8a		108.4, C	
9		(164.5, C) ^d	
9a		108.9, C	
10	6.27, bd (1H, <i>J</i> = 20.5)	118.0, CH	27.9, 101.9, 108.4, 108.9, 139.3
10a		139.3, C	
1'	4.05, s	81.0, CH	27.9, 44.8, 60.2, 78.3, 174.2
1'-OCH ₃	3.47, s	60.2, CH ₃	81.0
2'		174.2, C	
4'	4.83, m (overlap)	54.5, CH	
5'		173.1, C	
5'-OCH ₃	3.78, s (overlap, 5H)	52.9, CH₃	173.1
1''	3.24, dd (11.0, 14.0) 3.34 – 3.44 (m, 3H)	37.2, CH₂	54.5, 130.3, 139.0 54.5, 130.3, 139.0
2''		139.0, C	
3''	7.25 – 7.40 (m, 5H)	130.3, CH	37.2, 127.8, 129.6, 130.3, 139.0
4''	7.25 – 7.40 (m, 5H)	129.6, CH	37.2, 127.8, 129.6, 130.3, 139.0

5"	7.25 – 7.40 (m, 5H)	127.8, CH	37.2, 127.8, 129.6, 130.3, 139.0
1A	5.13, bs	98.0, CH	
2A	1.74 – 1.89 (m, 2H) 2.29 – 2.48 (m, 3H)	38.2, CH ₂	81.0, 98.0
3A	3.83 – 3.93 (m, 2H)	81.0, CH	
4A	3.13, t (8.8)	76.5, CH	18.9, 73.2, 81.0
5A	3.54 – 3.67 (m, 3H)	73.2, CH	
6A	1.41, d (5.5)	18.9, CH ₃	73.2, 76.5
1B	4.76, d (10.0)	100.3, CH	81.0
2B	1.52 – 1.68 (m, 3H) 2.23, dd (4.0, 12.0)	40.7, CH ₂	72.0, 100.3 72.0, 78.0
3B	3.54 – 3.67 (m, 3H)	72.0, CH	
4B	2.98, t (9.0)	78.0, CH	18.2, 72.0, 73.8
5B	3.34 – 3.44 (m, 3H)	73.8, CH	72.0, 78.0
6B	1.34, d (6.0)	18.2, CH ₃	73.8, 78.0
1C	5.03, d (9.5)	102.2, CH	
2C	1.52 – 1.68 (m, 3H) 2.61, d (8.5)	38.2, CH ₂	102.2 76.6, 80.6
3C	3.78, s (overlap, 5H)	80.6, CH	
4C	3.07, t (17.5)	76.6, CH	18.9, 73.3, 80.6
5C	3.34 – 3.44 (m, 3H)	73.3, CH	
6C	1.37, d (5.5)	18.9, CH ₃	73.3, 76.6
1D	4.69, d (9.5)	99.8, CH	80.6
2D	1.74 – 1.89 (m, 2H) 1.98 (d, 11.0)	33.1, CH ₂	77.2, 100.3
3D	3.83 – 3.93 (m, 2H)	77.2, CH	
4D	3.66 – 3.75 (m, 2H)	70.3, CH	33.1, 77.2
5D	3.54 – 3.67 (m, 3H)	72.0, CH	17.1, 70.3, 100.3
6D	1.29, d (6.0)	17.1, CH ₃	70.3, 72.0
1E	4.99, d (9.0)	99.0, CH	45.2, 77.2
2E	1.52 – 1.68 (m, 3H) 1.93, d (13.0)	45.2, CH ₂	99.0 27.2, 71.8, 77.7, 99.0
3E		, C	
3E-CH ₃	1.25, s	27.2, CH ₃	45.2, 71.8, 77.7, 99.0
4E	2.93, d (9.5)	77.7, CH	18.9, 71.8
5E	3.66 – 3.75 (m, 2H)	71.8, CH	77.7, 99.0
6E	1.27, d (6.0)	18.9, CH ₃	71.8, 77.7

^a The solvent used for NMR experiments was methanol-*d*₄.

^b ¹³C NMR data was inferred from gHSQC and gHMBC spectra.

^c multiplicity

^d The numbers in the bracket are cited from the reference.^[59] We did only ¹H, gHSQC, and gHMBC experiments for this compound, and no C-H correlations in the gHMBC spectrum which can indicate the chemical shift values of these carbons.

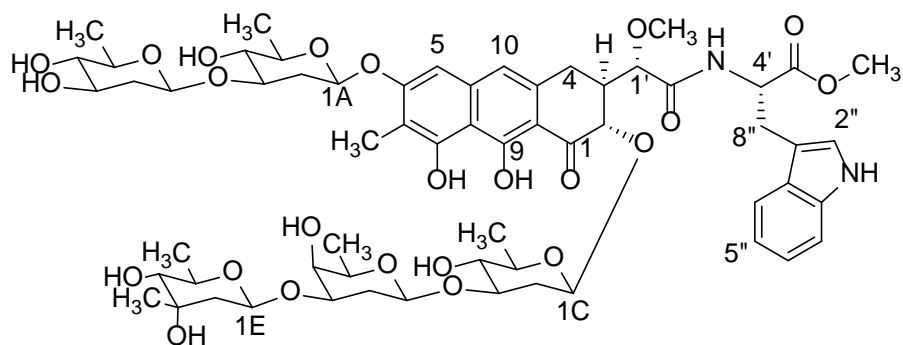


Table 12. NMR data (500 MHz) of MTM SA-Trp (**85**)^a

Position	δ_{H} (J in Hz)	$\delta_{\text{C}}^{\text{b}}$, mult. ^c	HMBC
1		(203.5, C) ^d	
2	4.43, d (11.5)	77.4, CH	44.6, 102.1
3	2.29, d (13.0)	44.6, CH	
4	1.73 – 1.98 (m, 5H) 2.32 – 2.42 (m, 1H)	27.2, CH ₂	77.4, 80.8, 117.9, 135.9
4a		135.9, C	
5	6.56, s	101.7, CH	108.2, 111.3, 117.9, 159.6
6		159.6, C	
7		111.3, C	
7-CH ₃	2.14, s	8.7, CH ₃	111.3, 156.4, 159.6
8		156.4, C	
8a		108.2, C	
9		(164.5, C) ^d	
9a		102.1, C	
10	5.93, s	117.9, CH	27.2, 101.7, 108.2, 139.1
10a		139.1, C	
1'	4.06, s	80.8, CH	27.2, 44.6, 59.9, 77.4, 173.9
1'-OCH ₃	3.47, s	59.9, CH ₃	80.8
2'		173.9, C	
4'	4.95 – 5.02 (m, 2H)	53.1, CH	173.2, 173.9
5'		173.2, C	
5'-OCH₃	3.81, s	52.8, CH₃	173.2
2''	7.21, s (overlap)	124.1, CH	110.9, 112.3, 119.3, 128.5, 137.8
3''		110.9, C	
3''a		128.5, C	
4''	7.66, d (7.5)	119.3, CH	110.9, 112.3, 122.7, 128.5, 137.8
5''	7.21, t (7.5) (overlap)	122.7, CH	110.9, 112.3, 119.3, 128.5, 137.8

6"	7.06, t (7.5)	119.9, CH	110.9, 112.3, 124.1, 128.5, 137.8
7"	7.43, d (8.0)	112.3, CH	119.9, 124.1, 128.5
7"a		137.8, C	
8"	3.37 – 3.43 (m, 2H) 3.50, dd (4.3, 15.3)	28.0, CH₂	53.1, 110.9, 124.1, 128.5 53.1, 110.9, 124.1, 128.5
1A	5.32, d (9.0)	97.7, CH	
2A	1.73 – 1.98 (m, 5H) 2.48, dd (4.0, 10.5)	37.9, CH ₂	97.7 76.0, 80.3
3A	3.65 – 3.78 (m, 4H)	80.3, CH	
4A	3.04, t (9.0)	76.0, CH	18.2, 73.2, 80.3
5A	3.32 – 3.34 (m, 1H)	73.2, CH	
6A	1.44, d (6.0)	18.2, CH ₃	73.2, 76.0
1B	4.65, dd (1.5, 9.5)	99.7, CH	80.3
2B	1.50 – 1.66 (m, 3H) 2.22 (m, 1H)	40.4, CH ₂	71.4, 99.7 71.4, 77.8, 99.7
3B	3.65 – 3.78 (m, 4H)	71.4, CH	77.8
4B	2.98, t (9.0)	77.8, CH	17.6, 71.4, 73.2
5B	3.65 – 3.78 (m, 4H)	73.2, CH	
6B	1.31 – 1.37 (m, 6H)	17.6, CH ₃	73.2, 77.8
1C	5.04, d (9.5)	101.8, CH	
2C	1.50 – 1.66 (m, 3H) 2.56, dd (4.0, 11.0)	37.9, CH ₂	80.5, 101.8 76.0, 80.5, 101.8
3C	3.83 – 3.93 (m, 2H)	80.5, CH	76.0
4C	3.14, t (8.8)	76.0, CH	18.2, 73.4, 80.5
5C	3.37 – 3.43 (m, 2H)	73.4, CH	
6C	1.31 – 1.37 (m, 6H)	18.2, CH ₃	73.2, 76.0
1D	4.78, dd (1.0, 9.5)	100.0, CH	80.5
2D	1.73 – 1.98 (m, 5H) 1.73 – 1.98 (m, 5H)	33.0, CH ₂	77.0, 100.0 70.2, 77.0, 100.0
3D	3.83 – 3.93 (m, 2H)	77.0, CH	100.0
4D	3.65 – 3.78 (m, 4H)	70.2, CH	77.0
5D	3.54 – 3.64 (m, 3H)	71.4, CH	16.8, 70.2
6D	1.29, d (6.5)	16.8, CH ₃	70.2, 71.4
1E	4.95 – 5.02 (m, 2H)	98.8, CH	27.5, 44.8, 77.0
2E	1.50 – 1.66 (m, 3H) 1.73 – 1.98 (m, 5H)	44.8, CH ₂	71.4, 98.8 71.4, 77.4, 98.8
3E		70.0, C	
3E-CH ₃	1.24, s	27.5, CH ₃	44.8, 71.4, 77.4, 98.8
4E	2.92, d (9.5)	77.8, CH	18.2, 71.4
5E	3.54 – 3.64 (m, 3H)	71.4, CH	70.0, 98.8
6E	1.26, d (6.5)	18.2, CH ₃	77.8, 71.4

^a The solvent used for NMR experiments was methanol-*d*₄.

^b ¹³C NMR data was inferred from gHSQC and gHMBC spectra.

^c multiplicity

^d The numbers in the bracket are cited from the reference.^[59] We did only ¹H, gHSQC, and gHMBC experiments for this compound, and no C-H correlations in the gHMBC spectrum which can indicate the chemical shift values of these carbons.

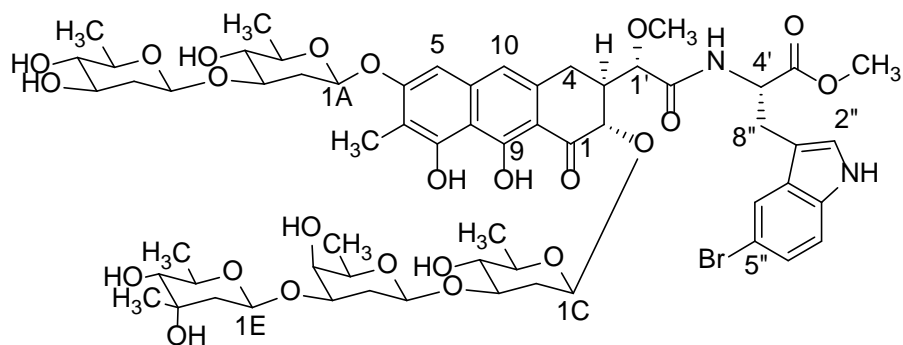


Table 13. NMR data (500 MHz) of MTM SA-5''-Br-Trp (**86**)^a

Position	δ_{H} (<i>J</i> in Hz)	$\delta_{\text{C}}^{\text{b}}$, mult. ^c	HMBC
1		(203.5, C) ^d	
2	4.60 – 4.72 (m, 4H)	77.0, CH	44.3, 101.3
3	2.39 (m, 1H)	44.3, CH	
4	2.09 – 2.14 (m, 1H) 2.62 – 2.73 (m, 1H)	27.3, CH ₂	136.0
4a		136.0, C	
5	6.78, d (7.5)	101.7, CH	108.2, 111.2, 117.3, 159.5
6		159.5, C	
7		111.2, C	
7-CH ₃	2.14, bs	8.4, CH ₃	111.2, 159.5
8		(157.2, C) ^d	
8a		108.2, C	
9		(164.5, C) ^d	
9a		108.7, C	
10	6.27, s	117.3, CH	27.3, 101.7, 108.2, 108.7
10a		140.2, C	
1'	4.10, d (1.5)	80.4, CH	27.3, 44.3, 60.1, 77.0, 171.8
1'-OCH ₃	3.49, s	60.1, CH ₃	80.4
2'		171.8, C	
4'	4.92 – 4.99 (m, 2H)	53.0, CH	27.3, 172.6
5'		172.6, C	
5'-OCH₃	3.76, s	52.6, CH₃	172.6
2''	7.42, s	126.1, CH	111.0, 130.2, 136.0
3''		111.0, C	
3''a		130.2, C	
4''	7.90, d (2.0)	121.9, CH	111.0, 112.4, 124.8, 136.0
5''		112.4, C	
6''	7.32, dd (1.8, 8.8)	124.8, CH	112.4, 114.0, 121.9, 136.0
7''	7.49, d (8.5)	114.0, CH	112.4, 124.8, 130.2
7''a		136.0, C	

8"	3.36 – 3.47 (m, 3H)	27.5, CH₂	53.0, 111.0, 126.1, 130.2, 172.6
1A	5.47, m	97.3, CH	
2A	1.84 – 1.97 (m, 3H) 2.46 – 2.54 (m, 2H)	37.9, CH ₂	81.3, 97.3 81.3
3A	3.83 – 3.92 (m, 2H)	81.3, CH	
4A	3.11, t (9.0)	75.7, CH	18.7, 73.1, 81.3
5A	3.62 – 3.74 (m, 5H)	73.1, CH	
6A	1.41, d (6.0)	18.7, CH ₃	73.1, 75.7
1B	4.78, dd (2.0, 9.5)	100.3, CH	81.3
2B	1.50 – 1.64 (m, 3H) 2.21, ddd (1.5, 4.5, 12.0)	40.3, CH ₂	100.3
3B	3.54 – 3.62 (m, 1H)	71.9, CH	
4B	2.91 – 3.04 (m, 3H)	77.8, CH	18.2, 73.0
5B	3.36 – 3.47 (m, 3H)	73.0, CH	
6B	1.25 – 1.34 (m, 12H)	18.2, CH ₃	73.0, 77.8
1C	5.10, d (9.5)	101.3, CH	
2C	1.50 – 1.64 (m, 3H) 2.46 – 2.54 (m, 2H)	38.3, CH ₂	101.3 81.7
3C	3.62 – 3.74 (m, 5H)	81.7, CH	
4C	2.91 – 3.04 (m, 3H)	76.0, CH	18.2, 73.0, 81.7
5C	3.22 – 3.34 (m, 1H)	73.0, CH	
6C	1.25 – 1.34 (m, 12H)	18.2, CH ₃	73.0, 76.0
1D	4.60 – 4.72 (m, 4H)	100.6, CH	
2D	1.78, dd (12.0, 22.0) 1.84 – 1.97 (m, 3H)	32.8, CH ₂	100.6 77.0
3D	3.83 – 3.92 (m, 2H)	77.0, CH	
4D	3.62 – 3.74 (m, 5H)	69.3, CH	17.1, 32.8, 77.0
5D	3.62 – 3.74 (m, 5H)	71.2, CH	
6D	1.25 – 1.34 (m, 12H)	17.1, CH ₃	69.3, 71.2
1E	4.92 – 4.99 m (2H)	98.4, CH	27.3, 77.0
2E	1.50 – 1.64 (m, 3H) 1.84 – 1.97 (m, 3H)	44.7, CH ₂	98.4 70.7, 77.0
3E		70.7, C	
3E-CH ₃	1.20 – 1.24 (m, 6H)	27.3, CH ₃	44.7, 70.7, 71.2, 77.0, 98.4
4E	2.91 – 3.04 (m, 3H)	77.0, CH	
5E	3.62 – 3.74 (m, 5H)	71.2, CH	77.0
6E	1.20 – 1.24 (m, 6H)	18.7, CH ₃	44.7, 70.7, 71.2, 77.0

^a The solvent used for NMR experiments was acetone-*d*₆.

^b ¹³C NMR data was inferred from gHSQC and gHMBC spectra.

^c multiplicity

^d The numbers in the bracket are cited from the reference.^[59] We did only ¹H, gHSQC, and gHMBC experiments for this compound, and no C-H correlations in the gHMBC spectrum which can indicate the chemical shift values of these carbons.

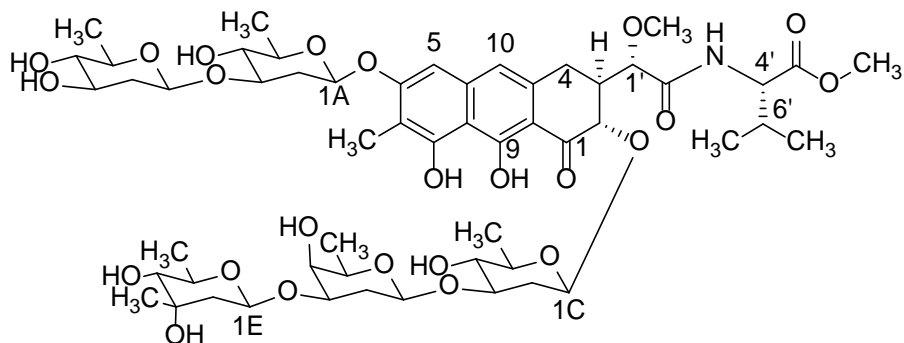


Table 14. NMR data (500 MHz) of MTM SA-Val (**88**)^a

Position	δ_{H} (<i>J</i> in Hz)	δ_{C} , mult. ^b	HMBC
1		202.8, C	
2	4.53, d (10.0)	78.2, CH	
3	2.37 – 2.58 (m, 3H)	45.0, CH	
4	2.37 – 2.58 (m, 3H) 2.81 – 2.90 (m, 1H)	28.5, CH ₂	
4a		136.3 C	
5	6.42 (bs, 2H)	102.0, CH	
6		160.2, C	
7		111.9, C	
7-CH ₃	2.08 (bs, 3H)	8.9, CH ₃	
8		156.8, C	
8a		108.7, C	
9		150.5, C	
9a		109.0, C	
10	6.42 (bs, 2H)	118.0, CH	
10a		139.6, C	
1'	4.17, s	81.3, CH	28.5, 45.0, 60.4
1'-OCH ₃	3.60 (bs, 5H)	60.4, CH ₃	81.3
2'		174.3, C	
4'	4.44, d (3.5)	59.1, CH	173.4
5'		173.4, C	
5'-OCH₃	3.77 (bs, 5H)	52.9, CH₃	173.4
6'	2.32 (m, 1H)	31.8, CH	19.2, 20.1, 59.1
7'	1.05, d (6.5)	20.1, CH₃	19.2, 31.8, 59.1
8'	1.05, d (6.5)	19.2, CH₃	20.1, 31.8, 59.1
1A	5.17, d (7.0)	97.6, CH	
2A	1.74 – 1.88 (m, 2H) 2.37 – 2.58 (m, 3H)	38.2, CH ₂	
3A	3.77 (bs, 5H)	81.0, CH	
4A	3.03 – 3.13 (m, 2H)	76.5, CH	18.8

5A	3.47 – 3.53 (m, 1H)	73.4, CH	
6A	1.26 – 1.44 (m, 5H)	18.8, CH ₃	73.4, 76.5
1B	4.65 – 4.75 (m, 2H)	100.2, CH	
2B	1.52 – 1.65 (m, 3H)	40.9, CH ₂	
	2.17 – 2.24 (m, 1H)		
3B	3.60 (bs, 5H)	72.2, CH	
4B	2.97, t (9.3)	78.3, CH	18.3, 72.2, 73.8
5B	3.34 – 3.42 (m, 2H)	73.8, CH	
6B	1.26 – 1.44 (m, 5H)	18.3, CH ₃	73.8, 78.3
1C	5.06 (bs, 1H)	102.2, CH	
2C	1.52 – 1.65 (m, 3H)	38.4, CH ₂	
	2.58 – 2.67 (m, 1H)		
3C	3.77 (bs, 5H)	81.2, CH	
4C	3.03 – 3.13 (m, 2H)	76.7, CH	18.8, 73.6
5C	3.34 – 3.42 (m, 2H)	73.6, CH	
6C	1.26 – 1.44 (m, 5H)	18.8, CH ₃	
1D	4.65 – 4.75 (m, 2H)	100.2, CH	
2D	1.74 – 1.88 (m, 2H)	33.3, CH ₂	
	1.95 – 2.02 (m, 1H)		
3D	3.85 – 3.93 (m, 1H)	77.5, CH	
4D	3.68 – 3.73 (bs, 2H)	70.6, CH	
5D	3.60 (bs, 5H)	72.2, CH	
6D	1.26 – 1.44 (m, 5H)	17.3, CH ₃	70.6
1E	4.99, d (9.0)	99.1, CH	
2E	1.52 – 1.65 (m, 3H)	45.4, CH ₂	99.1
	1.92, d (13.0)		71.9, 78.1
3E		71.9, C	
3E-CH ₃	1.25, s	27.4, CH ₃	45.4, 78.1, 99.1
4E	2.93, d (9.5)	78.1, CH	18.9, 72.0
5E	3.68 – 3.73 (bs, 2H)	72.0, CH	
6E	1.26 – 1.44 (m, 5H)	18.9, CH ₃	72.0, 78.1

^a The solvent used for NMR experiments was methanol-*d*₄.

^b multiplicity

3.4 Discussion

MTM analogues obtained in this study with the best activity are MTM SA methyl ester (**72**), MTM SA-Lys-2PGs (**82**), MTM SA-Phe (**83**), and MTM SA-Trp (**85**). By comparing these compounds, we could explore some possibilities regarding the MTM structure that may be beneficial for the design of next-generation MTM analogues. These potential ideas are described as follows:

MTM SA (69) and MTM SA methyl ester (72)

The C-3 side chain of MTM is an important functional group for the anticancer bioactivity because it is supposed to interact with the DNA phosphate backbone. Our hypotheses about the poor bioactivity of MTM SA are that the shortened side chain is a disadvantage for the interaction, or that the carboxylic group on the side chain carries a negative charge in the physiological environment which leads to electric repulsion from the negatively charged phosphate groups on the DNA backbone. In order to test the hypotheses, we carried out an esterification reaction on MTM SA in order to mask the carboxylic functional group with a single methyl group. We found that the MTM SA methyl ester was one of the most potent MTM analogues that we had obtained from this study. Therefore, the data were consistent with the hypothesis about electric repulsion instead of side chain length because the C-3 side chain of the MTM SA methyl ester is still the shortest one among all MTM analogues.

MTM SA-Phe (83), MTM SA-Trp (85), and MTM SA-Tyr (87)

MTM SA-Phe and MTM SA-Trp have similar potency as the MTM SA methyl ester, but the anticancer bioactivity of MTM SA-Tyr is lower than MTM. These compounds share a common trait in their structures of the C-3 side chain – they all contain a hydrophobic aromatic ring which may improve the interaction of the C-3 side chain with DNA through hydrophobic stacking with the pyrimidine and purine functional groups of DNA nucleobases. Compared to MTM SA-Phe, MTM SA-Tyr contains one additional hydroxyl group that could be utilized as a proton donor, but its bioactivity is obviously lower than MTM SA-Phe. Thus, this H bond is not necessary for the interaction between the C-3 side chain and DNA.

Moreover, our co-workers' unpublished work has shown that both MTM SA-Phe and MTM SA-Trp possessed higher specificity toward Ewing's sarcoma than parent compound, MTM. Our newest hypothesis is that there is a specific interaction between the C-3 side chain of the two compounds and the protein FLI1 (the DNA-binding domain of the transcription factor EWS-FLI1). The aromatic residues could play a crucial role in the formation of the drug-DNA-transcription factor ternary complex.

MTM SA-His (81) and MTM SA-Lys-2PGs (82)

The pKa value of the imidazole functional group of L-histidine is approximately 6.04. This indicates that in a physiological environment, a relatively small shift in pH value will change the average charge of MTM SA-His. This analogue is non-polar in an environment with a pH value above 6, but its chemical and physical properties may become different when some unknown factors, released by either cancer or normal cells, have made the environment more acidic. Therefore, it is difficult to anticipate the *in vivo* anticancer behavior of MTM SA-His.

In comparison to MTM SA-His, MTM SA-Lys-2PGs (the compound possessing the longest non-polar C-3 side chain among all MTM analogues) showed the strongest anticancer bioactivity in the NCI 60 cell line screen (one-dose). One of the possible explanations is that, like structure **83** and **85**, the *N*-benzyl carbamate could enhance the interaction between the C-3 side-chain and DNA through hydrophobic stacking with the pyrimidine and purine functional groups. Moreover, the non-polarity of the lysine group may offer the C-3 side chain good physical and chemical stability. This is an advantage for the anticancer activity. The last possible idea is that an uncharacterized protease in the cancer cell could hydrolyze the peptide bond on the C-3 side chain of this compound and release the basic amino group of lysine (pKa = 10.67). The positively charged C-3 side chain may enhance lysine's interaction with the DNA phosphate backbone, which is negatively charged in the physiological environment.

MTM SA-Ala (78), MTM SA-Gly (80), and MTM SA-Val (88)

Coupling L-alanine, L-glycine, and L-valine to the C-3 side chain of MTM SA improved the anticancer bioactivity because these amino acids not only converted the

negatively charged carboxylic group to a non-polar functional group in order to enhance its interaction with DNA. Moreover, the only structural difference of the three amino acids is in the side chain at the C β position. L-alanine and L-valine have a methyl and isopropyl group, respectively, while L-glycine does not have any side chain. The approximately equivalent anticancer bioactivity of the coupling products (**78**, **80**, and **88**) indicates that changes in the aliphatic side chains at the C β position do not have any significant difference in the anticancer bioactivity or *in vivo* tolerance.

Summary

In conclusion, MTM is a very potent anticancer agent with special mechanisms. However, its clinical utilization is limited by several side effects. Our laboratory produced three second-generation MTM analogues, MTM SK, MTM SDK, and MTM SA, with combinatorial biosynthesis. Both MTM SK and MTM SDK possess better anticancer activity and *in vivo* tolerance than the parent compound, but MTM SA, the major product accumulated in the culture broth, shows poor bioactivity. Through chemical semi-synthesis, we coupled several natural and unnatural primary amines and amino acids to the C-3 side chain of MTM SA and generated new MTM analogues showing some useful SAR information for further drug design on MTM. Moreover, our unpublished data showed that MTM SA-Trp and MTM SA-Phe may possess better anticancer activity, specificity, and *in vivo* tolerance than MTM SK in the treatment of ESFT cancer cell lines. Therefore, these third-generation MTM analogues can be used as lead-structures to generate the next-generation MTM analogues.

3.5 Materials and methods

General materials

Trimethylsilyldiazomethane (TMSCHN₂), cystamine, methyl hydrazine, L-alanine methyl ester hydrochloride, L-cysteine methyl ester hydrochloride, L-glycine methyl ester hydrochloride, L-serine methyl ester hydrochloride, L-tryptophan methyl ester hydrochloride, L-tyrosine methyl ester hydrochloride, L-valine methyl ester hydrochloride, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluoro-phosphate (PyBOP), DMSO

(molecular biology grade, $\geq 99.9\%$), Dichloromethane (DCM), *N,N*-diisopropylethylamine (DIPEA), were purchased from Sigma-Aldrich (St Louis, MO, USA). 1-amino-2-propanone was purchased from Waterstone Tech (Carmel, IN, USA). *N,N*-dimethylethylenediamine was purchased from TCI America (Portland, OR, USA). L-phenylalanine methyl ester hydrochloride was purchased from Fluka of Sigma-Aldrich (St Louis, MO, USA). L-histidine methyl ester hydrochloride and D-glucosamine was purchased from Acros of Fisher Scientific (Hampton, NH, USA). L-tryptophan 5-Br methyl ester hydrochloride was purchased from CHEM-IMPEX INT'L INC (Wood Dale, IL, USA). L-lysine t-Bu ester *N*-benzyl carbamate hydrochloride was purchased from Merck KGaA (Darmstadt, Germany). Methanol (MeOH), acetonitrile (ACN), celite, C18 RP silica gel, tryptic soy broth (TSB), LB broth, Difco agar, sucrose, potassium sulfate, magnesium chloride, glucose, casamino acids, yeast extract, MOPS, and trace elements were purchased from Fisher Scientific (Hampton, NH, USA). *Streptomyces argillaceus* ATCC 12956 was purchased from ATCC (Manassas, VA, USA).

Biosynthesis of MTM SA

MTM SA was produced by a procedure reported previously.^[60] *S. argillaceus* M7W1 (*mtmW* minus) was inoculated on R5A agar and incubated at 28 °C for 7 days or until spores formed. The spores were then used as seeds to inoculate 100 mL of TSB and allowed to grow for 24 hours in an orbital shaker at 28 °C, 250 rpm. 3 ~ 4 mL of TSB was transferred to 100 mL of R5A (40 flasks) after 24-hr incubation. The cultures in R5A media were grown for 7 days in an orbital shaker at 28 °C, 250 rpm while the production of MTM SA was monitored with HPLC. After 7-day incubation, the cells were collected by centrifugation (4000 rpm for 25 min) at room temperature to separate the cell pellet and culture broth. The cell pellet was resuspended and sonicated in MeOH in order to extract the secondary metabolites, including MTM SA. Following the cell lysis, the cellular debris was filtered off, and MeOH was evaporated from the filtrate to concentrate the extract and combined with the culture broth. *n*-BuOH was used to extract MTM SA from this mixture (three times) and then evaporated after forming an azeotrope with water. The dried extract was re-dissolved in water and loaded onto a 5 × 12 cm C18 reverse-phase column that had been equilibrated with 10 column volumes of water. We purified MTM SA by adopting

the following procedure: washing column with 10% ACN in water, followed by a fractionation of ACN in water from 20% to 50%, and then followed by 100% ACN. After checking every fraction with LC-MS, the fractions containing MTM SA were mixed and dried before further purification with semi-preparative HPLC. We finally obtained approximately 80% pure MTM SA (~ 10 mg/ L) for chemical semi-synthesis.

Chemical reaction conditions and purification of the products

The C-3 side chain of MTM SA was modified with several natural and unnatural primary amines and amino acids through a coupling reaction. After trying several different reaction conditions, we found that PyBOP was the most efficient coupling reagent for this coupling reaction, and DIPEA could deprotonate the terminal carboxylic acid group of MTM SA efficiently. The details of the reaction condition were as follows: 10 mg MTM SA was mixed with 3 folds of desired primary amines for side-chain modification, 3 folds of DIPEA, 2 folds of PyBOP, and the solvent DCM. All reactions were initiated at 4 °C and, the production of MTM analogues was checked with LC-MS after 16 hr. After the reactions were complete, the organic solvent was removed. The dried sample was re-dissolved in MeOH for silica gel open column chromatography to purify MTM analogues with the following procedure: silica gel column was equilibrated with 100% DCM, followed by fractionation of MeOH in DCM from 5% to 65%, and then followed by 100% MeOH. The eluent was collected and checked with normal phase thin layer chromatography (TLC) to identify fractions with the desired products. After we combined all fractions containing the MTM analogues together and removed the organic solvent, semi-preparative HPLC was used to isolate the individual compounds (Waters HPLC system, as described in Chapter 2.5). Gradient used: solvent A: water with 2% formic acid, solvent B: acetonitrile. Solvent B was increased from 25% to 70% (min 0 to 19), then from 70% to 90% (min 19 to 21), then from 90% to 100% (min 21 to 23), and maintained 100% for 8 minutes (min 23 to 31), then decreased back to 25% within 2 minutes (min 31 to 33), and kept at 25% for the last 2.5 min (min 33 to 35.5).

Cytotoxicity assays

Cytotoxicity assay in Dr. Younsoo Bae's laboratory at the University of Kentucky

This cytotoxicity assay was carried out by Dr. Daniel Scott at Dr. Younsoo Bae's laboratory in the University of Kentucky.^[59] A549 cells were cultured as specified from ATCC at 37 °C, 5% CO₂. The cells were added into a 96-well plate (5000 cells/well) and permitted to attach for 24 hr. After 24-hr culture, media were replaced with the side chain modified MTM analogues containing media at differing concentrations ranging from 10 μM to 1 × 10⁻⁴ μM. The cells were incubated with the drug-containing media for 72 h total (n = 8). Cell viability was determined using a resazurin assay that signifies mitochondrial metabolic activity in living cells.^[61] About 10 μL of a 1 mM resazurin solution in phosphate-buffered saline (PBS) was added to the control and analogues-treated cells at the end of the treatment period. Cell viability was determined 3 h later by reading the fluorescence at 560 nm (Ex)/590 nm (Em). The fluorescence signals were quantified using a Spectramax M5 plate reader (Molecular Devices, Sunnyvale, CA, USA) equipped with a SoftMax-Pro software. Cytotoxicity was determined by calculating the half maximal inhibitory concentration (IC₅₀) of each sample.

NCI 60 cell line screen (one-dose) at NCI in the NIH

This 60 cell line screen (one-dose) was carried out by Dr. Jürgen Rohr's collaborators at NCI in the NIH.^[58] All samples were shipped to NCI at 4 °C and then stored at -20 °C before the assay. The information about the methodology of the *in vitro* cancer screen, sample handling and preparation, cell lines used in the screen, and data analysis have been already published on NCI's website (<http://dtp.cancer.gov/branches/btb/ivclsp.html>).

CHAPTER 4: STRUCTURAL INVESTIGATION INTO MTMC

4.1 Structural insight into MtmC

The structure of MTM (**11**) is composed of a tricyclic polyketide-derived core with two alkyl side chains and five sugar moieties (three D-olivoses, one D-oliose, and a D-mycarose) attached as disaccharide and trisaccharide chains. The trisaccharide chain is important for the anticancer activity of MTM because it appears to be evolutionarily optimized for binding DNA in a minor groove of GC-rich regions. Luz E. Núñez and co-workers transformed *S. argillaceus* M7W1 (*mtmW* minus) with pMP3*BII, a sugar plasmid encoding all genes necessary for the biosynthesis of TDP-D-digitoxose, and successfully generated a new MTM analogue with lower anticancer activity but two folds better *in vivo* tolerance than MTM SK, a standard used to evaluate the toxicity of MTM analogues. The replacement of the last sugar moiety on the trisaccharide side chain of this new compound, demycarosyl-3D- β -D-digitoxosylmithramycin SK, with D-digitoxose revealed that modifying this side chain could change the bioactivities of MTM.^[14a] While it is difficult to modify sugar moieties through chemical semi-synthesis, protein engineering of the enzymes involved in the biosynthesis of this trisaccharide side chain is a promising method to generate more MTM analogues. In order to find a better method, we are working to collect more information about the enzymatic mechanism through X-ray crystallography of these enzymes.

Our previous work had demonstrated that MtmC (46.6 kD) is a bifunctional enzyme involved in the biosynthesis of both TDP-D-olivose and TDP-D-mycarose.^[45] MtmC can generate both substrates of MtmGIV by either reducing the C-4 keto group or transferring a methyl group to the C-3 position of the TDP-4-keto-2,6-dideoxy-D-olivose (TDP-KOL) (**Figure 18**).^[46b] Therefore, MtmC was hypothesized to cooperate with another bifunctional enzyme, MtmGIV, in the biosynthesis and positioning of the two sugars by forming a multifunctional enzyme complex. In this, MtmGIII acts by transferring the middle sugar, D-oliose, to premithramycin A1 (**Figure 19**).^[47] With this hypothesis, we started to investigate all three enzymes' interactions with each other using X-ray crystallography. We expect to gain further insights in the protein-protein interactions, the biochemical mechanisms of the reactions, and the substrate channeling inside this proposed

enzymatic complex, MtmC-MtmGIV. The enzymatic mechanism of MtmC, the first piece of the puzzle of this complicated biosynthetic pathway, are described in this chapter.

4.2 Experimental design

MtmC was soluble in *E. coli* BL21 (DE3) and could be purified easily by Ni²⁺ affinity column chromatography, followed by FPLC for further purification. In order to clarify the reductase and methyltransferase functions of this enzyme, we obtained and determined crystal structures of apo-MtmC, binary and trinary complexes of MtmC with its biologically relevant ligands (**Figure 18**). When forming the binary complexes, we tried to generate crystal complexes of MtmC with two different co-substrates (SAM or NADPH), an alternative to the real substrate (TDP), and a co-product of methylation (SAH). Moreover, MtmC was also soaked with one of the co-substrates (SAM and NADPH), one of the substrates (TMP, TDP, and TDP-KOL) toward the generation of trinary complexes. Because MtmC showed high catalytic activity of methylation in our previous studies, we had to soak this protein with SAH and TDP-KOL at the same time, rather than SAM and TDP-KOL. Otherwise, the reaction would happen, and no crystal would be generated. Compared to the methyltransferase function, MtmC needed approximately 16 hours to catalyze the reduction of TDP-KOL.^[45] NADPH might loosely bind to MtmC so that the catalytic activity of MtmC would be lower. In order to clarify this hypothesis, we soaked MtmC with both SAH and NADPH to understand which co-substrate possessed stronger affinity to MtmC. After obtaining the preliminary structural information and identifying the key amino acid residues for catalysis of MtmC, we generated some mutant MtmCs by point mutation for kinetic investigations. Through the comparison between wild-type and mutant MtmCs, we confirmed our inferences about the enzymatic mechanism of MtmC.

4.3 Results

Homologues of MtmC

When doing a BLAST search for MtmC homologues, we obtained a list of many bacterial SAM-dependent C-methyltransferases, and most of them belong to bacterial class

I family of SAM-dependent sugar C-methyltransferases.^[62] The MtmC homologues whose functions had been clarified or postulated are listed in **Table 15**, and the multiple sequence alignment of MtmC and the listed homologues is shown in **Figure 29**. Among these MtmC homologues, TcaB9 from the biosynthetic pathway of *Micromonospora chalcea* for D-tetronitrose in tetrocarcin A is the only homologue whose structure and function was characterized.^[63] Because TcaB9 is a mono-functional C-3' methyltransferase, the structural information of MtmC could not only explain its special bi-functional mechanisms in detail but also reveal divergent structural features of this family of C-methyltransferases responsible for recognition of specific sugar substrates.

Table 15. MtmC and its homologues^a

Protein	Bacterium	Sugar product	Pathway product	Max score
MtmC	<i>Streptomyces argillaceus</i>	D-mycarose/D-olivose	mithramycin	
SnogG2 ^[64] (SnoG)	<i>Streptomyces nogalater</i>	nogalose	nogalamycin	557
SibM ^[65]	<i>Streptomyces sibiricum</i>	sibiromycin	sibiromycin	443
CloU ^[66]	<i>Streptomyces roseochromogenes</i> DS 12.976	5-methyl-L-rhamnose	clorobiocin	425
NovU ^[67]	<i>Streptomyces spheroids (niveus)</i> NCIB 11891	L-noviose ^b	novobiocin	421
CouU ^[68]	<i>Streptomyces rishiriensis</i> DSM 40489	L-noviose ^b	coumermycin A1	409
ORF14/SmtA ^[69]	<i>Amycolatopsis orientalis</i>	L-epivancosamine	chloroeremomycin vancomycin	282
TcaB9 ^[63]	<i>Micromonospora chalcea</i>	D-tetronitrose	tetrocarcin A	280
TiaS2 ^[70]	<i>Dactylosporangium aurantiacum</i> subsp. <i>hamdenensis</i> NRRL 18085	modified D-rhamnose	Tiacumicin B	236
AviG1 ^[71]	<i>Streptomyces viridochromogenes</i> Tü57	L-mycarose	Avilamycin A	196
TylCIII ^[72]	<i>Streptomyces fradiae</i>	L-mycarose	tylosin	185
EryBIII ^[73]	<i>Saccharopolyspora erythraea</i>	L-mycarose	erythromycin	167

^a The homologues are listed in the order of their max score of BLAST with MtmC.

^b Novobiocin and coumermycin A1 contain the same deoxysugar moiety.

CLUSTAL O(1.2.1) multiple sequence alignment

```

AviG1      --MSTTGHSTVIDRCRICDNTELLPVLDLGPQALTGVFPRTRGE--DVPYVPLELVRCSF      56
TylCIIII  MPAVPREDQMIISACRVCGNRELLPVLDLGEQALTGVFPRTREE--TVPSIPLLELVKCSF      58
EryBIIII  -----MGLGKRCICGNHDLSEVLHLGDQALTGVFPNRDE--IVPSVPLELVKCAP      49
TiaS2     -----MHCQICSSNELIHVTSLGHHALPDVVFAPQEAAPFTVYPLDLYVC--      46
ORF14/SmtA -----MSTTSQCRICDGT-VTEFIDFGRQPLSDAFVVP-GDEKKEFFRLATGIC--      48
TcaB9     GHMSHLADVSPPTACRVCVGG-VQEFDLGRQPLSDRFKRP-DELDDEFTYRLAVGRC--      56
NovU      ----MQDIVRPIDECRVCVGHDDWLDVVSFGSTPLAGNLLGDEDDAGGETLFPPLDVVVC--      54
CloU      ----MQDIVRPINKRVCVGHNDWLDVVSFGSTPLAGNLLGKEDAPGGETLFPPLDVVVC--      54
CouU      ----MQDIVRPINERVCVGHNDWLDVVSFGSTPLAGNLLGTEDDPRDETLPPLDVGVV--      54
SibM      ---MTTEIAKSVTTCRVCVGAADWQEVISFSGSVPLANAYLDKPG--PEEAYPLELIS--      52
MtmC     --MDIYGTAARAVTTCRMCGAQDWQEVVDFGFPVPLADSFLEPAASYDDEPRYPLAVVSC--      56
SnoG     --MDNRETVRPVSVCRVCGGNDWQDVVDFGDVPLANGFLSPADSYENRRYPLGLVLS--      56
          *::: * . : * * * *

```



```

AviG1      GGCELVQLRHTADFLMYGEGYGYRSSLNRSMDHLRGKVAAITGLVLDLGGPDLVVDIGS      116
TylCIIII  QGGLVQLRHTPDPGLMYGEGYGYRSGIRPFMIGHLRKVAAIRLVDLGPDDLVLVDIGS      118
EryBIIII  PGCGLVQLRESADFLMYNEGYGYRSGIRPFMINHLHGKVAKLGMVVPVGPDDLVDIGS      109
TiaS2     GRCWLAQIGHAVDPALLFR-DFVYTTESNPASRPHFQRLVDSLVERFQLGTEDLAVDLAS      105
ORF14/SmtA ESCTMVQLMEEVPRDLMFHEAPYPYHSSGSAVMRTHFHELAKRFLTTELTGEDPPIVELGC      108
TcaB9     DSCEMVQLTEEVPRLDMFHEVYYPYHSSGSSVMREHFAMLARDPLATELTGPDPPIVELGC      116
NovU      RRCWLMTLRHVIEPDVLFV-HYRYVASDAGSIIQHMRKLVLCVERIGLTDGDLVVEFGS      113
CloU      RRCWLMTLRHVIEPDVLFV-HYRYVASDAGSIIQHMRKLVLCIGWIGLTDGDLVVEFGS      113
CouU      RRCWLMTLRHVIEPNALFV-HYRYVASDAGSIIQHMRKLVLCIRWVGLTDRDLVVEFGS      113
SibM      RSCRLLSLTAVVDPQALYR-DYAYVTSESETMTHHMRDVVRLCRERI--DAGGLVVELGS      109
MtmC     RSCRLMSLTHVVDPEVLYR-TYPYTTSDSETIKKMHGHVAVCVFERFGIPEGSFVLEIGS      115
SnoG     RACRLMSLTHVVDPEVLYR-DYAYTTPDSEMITQHMRHITALCIRFELPPDSLVELGS      115
          * : : : : : * : : * : : : :

```



```

AviG1      NDGTLAAYPADGPRVLGVDPAAVFAASYPPGVELIPDFFAYDL-----LGGRRAKVVT      171
TylCIIII  NDATLLKAY-PEGPRLVGIDPSGDKFRELYPPHAELIAEYFSRDVFTA-RFGTRRARVIT      176
EryBIIII  NDSTLLRGYLPDAPKLAGFDLVGKFRDLYPPEADLVGTGFFSADAFEE-RYGERRAKVVT      168
TiaS2     NDGTLLEQYA-GRVRLVGVDPSS-TGLAIAKGLPTIVDFFTEATAEKILATYQPAKVIT      163
ORF14/SmtA NDGIMLKAIAEAGVRQLGVFESSGVADLAAAKGIRVRKDFFEATAADIRKNDGPADVIY      168
TcaB9     NDGIMLRTIQEAGVRHLGFEPSSGVAAKAREKIRVVRTDFFEKATADDVRRTEGPANVIY      176
NovU      NTGAHLELFRQAGPRVVGVDPARNLAGVANDRGVETIPAGFTAEVGEEIATRHLARLVY      173
CloU      NTGAHLQLFQQAGMRVVGVDPARNLADIANGRGVETIPAGFTIEVGGELASGHGPARLVY      173
CouU      NTGAHLELFPQQAGMRVVGVDPARNLADIANGRVTTIPAGFTIEVGGQEIITGHPARLVF      173
SibM      NTGNQLQLFQQAGHPVLGVDPAQNLAAIAERRGVPTVADYFNEKVATAIAGEHGRASLIL      169
MtmC     NTGSQLKAFQAGMRTLGI DPARNIAAVANERGIETLPEFFSVDTAALVKKTHGTPQLVL      175
SnoG     NTGRQLMAFREAGMRTLGVDPARNLTDVARNGIETFPDFSHDVARTIRRDHGQARLVL      175
          * . * * : : * : : : :

```



```

AviG1      SIAMFYDLPRPMEFMREVGRLTDDGIWVTEQSYLPAMLHACAYDVVCHHELDYYGLRQI      231
TylCIIII  SIAMFYDLDPPLAFMRDVHDVLADDGIWVMEQSYLPAMLEADAYDIVCHEHLEYALQQI      236
EryBIIII  SIAMFYDLPEPMRFMRDVHDILADDGLWLMEQSYLPMSLDAGAYDVVCHHELEYALAQI      228
TiaS2     TMNAFAHIKELDAFMSGVVTLLAEDGVFVTESSHHLQNVITKMHYDLIHHEHLRFPYSLQPL      223
ORF14/SmtA AANTLCHIPYMSILKGVTTLLGPNGVFVFEDPYLGDIVERTSFDQIYDEHFFFTARSV      228
TcaB9     AANTLCHIPYVQSVLEGV DALLAPDGVFVFEDPYLGDIVAKTSFDQIYDEHFFLFSATS      236
NovU      GRQCFAHIPDIHEVLNGVSALLAPDGLFFVEVPYLVELLKNNQFDITFHEHFSYFSLGSL      233
CloU      GRQCFAHIHDIHEVLNGVSALLAPDGLFFVEVPYLVELLANNQFDITFHEHLSYFSLGSL      233
CouU      GRQCFAHIHVHEVLDGVSALLAPDGLFFVEVPYLVELLANNQFDITFHEHLSYFSLGAL      233
SibM      GRHVFAHIDDLADVLRGVKALLTDDGLFAIEVPYACDLLERVAFDTIYHEHLSYFLISTL      229
MtmC     GRHVFAHIDDVSAVAEGVRDLLGPDLSFAIEVPYLVDMLENERNEFDTIYHEHLSYIGVGS      235
SnoG     GRHVFAHIDDVSDIAAGVRELLSPDGVFAIEVPYVLDLLEKVAFDTIYHEHLSYFTMRSF      235
          : : . * : * : : : * : : : * : : * : :

```

AviG1	EWMAERTGLKVVDAELTPVYGGSLSLVLRARRGSSRQVNEPALARIRAGET-----DLPY	285
TylCIII	EWMAERAGLTVLRAELTDVYGGSLCVTLARASSPHPRDEAGPARIRARETEAKINTMAPF	296
EryBIII	EWMAQRVGLKVVDAEITDVYGGSLCAVLAKQGGSGHPVDEAGLERIRAREAAAKLDTMAPY	288
TiaS2	TVLFARYGLEVFDVEETATQGGSLVYAGRTGVVP-VSA-NVPRLLQETAAGLYSIDTY	281
ORF14/SmtA	QEMARRNGLELVDVERIPVHGGEVRYTLALAGARK-PTE-AVAELLAWEAERKLSFATL	286
TcaB9	QGMAQRQCGFELVDVQRLPFVHGGEVRYTLARQGSRT-PSA-AVAQLLAAEREQELSDMATL	294
NovU	CTLFESHGLRVVDVHTADVHGGSIVVFAAPAAADHEVRP-AVAEMLAEERSQGIAEERTY	292
CloU	CRLFESHGLRVVDVHTADVHGGSIVVFAAPTITAGYEVVP-AVAEMLAEERSQGIAEERTY	292
CouU	CRLFEAHGLRVVDVHTVVDVHGGSIVVFAAPATAGYEVVP-AVAEMLAEERSQGIAEESTY	292
SibM	DRLFTRHGLRAIDVRR LAVHGGSVLVTAARMGSRWETDE-RVGGELIAYEKSERLDTDARY	288
MtmC	VALFRRHGLRVVDVERLAVHGGSILVVFVGLDEGTRATAP-VVEELIALEKERGLYEDATY	294
SnoG	VTLFARHGLRVLDVERFGVHGGSVLVVFVGHEDGPWPERP-SVPELLRVERQRGLYDDATY	294
	: * : . . . ** : . : *	
AviG1	AEFARRTEESRDLLEFLTASRDKGLHTLGYGASTKGNVILQYCGLDETLLEPCIAEVNED	345
TylCIII	EEFARRVEHQRDALRDLDRSRAAGRLTLGYGASTKGNVILQYCGIGERDLPCIGEVSPF	356
EryBIII	EAFARETERQRDQLEFLAKSRAEGKLTGLYGASTKGNVILQYCGLTEQDLPCIGEVSPF	348
TiaS2	QAWTARLMQSKQRLMSLLYGLKAQGATIAIGIGASAKGNTILNYCGIRPDVLDYITDGSEL	341
ORF14/SmtA	ERFAANVKKVKEDLIALLLKLRAEKGRVVGYGATAKSATVTNFCGITPDLVEFISDTTPA	346
TcaB9	RAFAGNVVKIRDELTAALLHRLRAEGRSVVYGATAKSATVTNFCGIGPDLVHVSVDYDTPD	354
NovU	REFADRTERVRAQIRELVRGVVADGKTVAGYGAPTKGSALLAACGLGHQEIFRCSDDTVL	352
CloU	RKFAERTERVRAQIRELVCGVVAEGKSVAGYGAPTKGSALLAACGLGHQEIFRCSDDTAL	352
CouU	QKFAERTERVRAQIRELVRSLVADGKTVAGYGAPTKGSALLTACGLGHQEIFRCSDDTAL	352
SibM	EQFAARVHALRRELTELVRKEAANGKRVAGYGASAKGTTLLNICELTSTEVRFCTDITFPQ	348
MtmC	ERFARHVAEITAEELTSMVRSRAEGKRIAGYGAPAKGNTLLNVCGLTADDFECCDTTEF	354
SnoG	RTFAQRIERVRELPELLRSLVAQKGRIVGYGAPAKGNTILTVCGGLGLKELEYCTDITTEL	354
	:: . : :: * * ** : * . : * : : .	
AviG1	KFGCYTPGTNIPIVSEEEARALEPDDQFLVLPWIYRDAMVARERDFLASGGSLVFPPLPLE	405
TylCIII	KAGRFTPGTGPIVSEEDAKAMRPDQLLVLPWIYREGFVERERDFLAGGGRLVFPPLRLD	416
EryBIII	KSGCYTPGTGPIVSEEEAKSRRPDQLLVLPWIYRDGFVEREQEFLAGGGKLIFFPLPRLE	408
TiaS2	KQGLKAPGSLIPVTADQELLDROPDYALVLTSLNKDIIIRKLTE-KGYQGRFIVAGEVPS	400
ORF14/SmtA	KQGLSPGQHIPVREYREFADDPDYALVLFANWHADEIMNAEQAFRDAGGQWIRYVPNVH	406
TcaB9	KQNRITPGAHIPVRPASAFSDPYDYALLFANWHAEEIMAKEQEFHQAGGRWILYVPEVH	414
NovU	KQGLKILPGSRIPVSPQAGHVPDYLLLANWYAPEIIDNEKEFLENGGRFIVPIPEPR	412
CloU	KQGLKILPGSRIPVSPQAGDHVPDYLLLANWYAPEIISKEKSFLENGGRFIVPIPEPR	412
CouU	KQGLKILPGSRVPIVSPQASDHVPDYLLLANWYASEIITKEKSFLEGGGRFIVPIPEPR	412
SibM	KQGRFVPGTQIPVAPGDVSE-QPDLYLLLANWYAEIILQREAAFLDAGGAFIIPVPEPA	407
MtmC	KQGLVLPGTHIPVRSPEYAKTQADYLLLANWYGEIILAKEGPFADGGRFILPNRPS	414
SnoG	KQGRVLPGTHIPVHAPHAKEHIPDYLLLANWYATEILDKETAFRDNGGRFIVPIPRPS	414
	* ** : * * ** : : * :	
AviG1	VV-----	407
TylCIII	VV-----	418
EryBIII	VV-----	410
TiaS2	VE-----	402
ORF14/SmtA	VS-----	408
TcaB9	IR-----	416
NovU	VISAESTL-	420
CloU	VISAESAL-	420
CouU	VISAESAW-	420
SibM	VYARQKVRQ	416
MtmC	IVPPGE---	420
SnoG	ILTSPPSGS-	422
	:	

Figure 29. Multiple sequence alignment of MtmC and its homologues.

Overview of crystal structures of MtmC with its substrates and co-factors

Through crystallization, we obtained and determined structures of three binary complexes (MtmC-SAM, MtmC-SAH, and MtmC-TDP) and two ternary complexes (MtmC-SAH-TDP-4-keto-D-olivose and MtmC-SAH-TDP) (**Figure 30**). We failed to generate any crystal of apo-MtmC with good quality suitable for X-ray diffraction experiment, likely because of the flexibility of the active site involved in the substrate and co-substrate binding. The overall structures of MtmC and TcaB9 are similar, and the TcaB9 structure was used as a template to determine the structures of MtmC binary and ternary complexes (molecular replacement approach).^[63]

Like TcaB9, the overall structure of MtmC is a monomer with a tripartite fold. The N-terminal domain ranging from residue 1 to 60 is composed of a β -sheet and extensive regions lacking secondary structure where a Zn^{2+} ion was bound tightly to four conserved cysteine residues (Cys13, Cys16, Cys56, and Cys59) (**Figure 30 A**). The conformation type of the central (residues 69-288) and the C-terminal domains (residues 289 – 423) are Rossmann-type fold with a co-substrate and a substrate binding site, respectively. In all of the binary and ternary complexes, the ligands were very clearly resolved in the strong omit $F_o - F_c$ electron density map (**Figure 30**). We found that their structures are similar to each other except a region of the central domain (residues 76 – 84), whose residues can interact with SAM/SAH and the sugar moiety of the substrate. The structural differences in this region are described in detail in the following two sections.

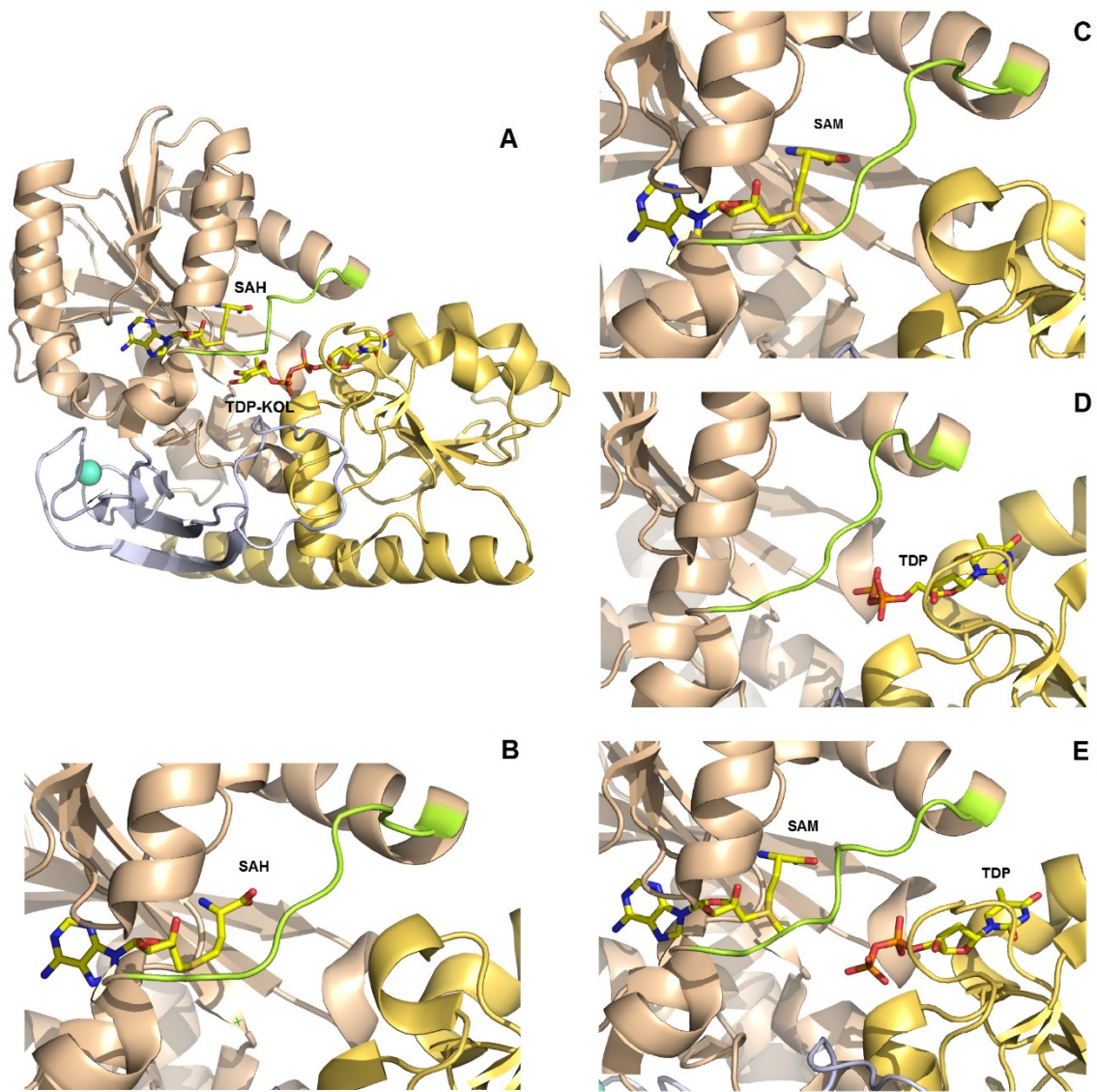


Figure 30. Cartoon representation of crystal structures of MtmC with its biologically relevant ligands. The ligands are shown as sticks and colored based on the type of elements. (C: yellow, O: red, S: orange, P: pink, N: blue, Zn: green cyan) The *N*-terminus is colored blue white; the central domain is colored wheat; the *C*-terminus is colored yellow orange; the sensor loop is colored lemon; the zinc ion is colored green cyan. (A) Overall structure of MtmC-SAH-TDP-4-keto-D-olivose (MtmC-SAH-TDP-KOL). The close-up views of (B) MtmC-SAH, (C) MtmC-SAM, (D) MtmC-TDP, (E) MtmC-SAM-TDP. (The structures of MtmC crystal complexes was resolved by Dr. Tsodikov and Dr. Hou in Dr. Tsodikov's laboratory.)

MtmC-SAM and MtmC-SAH complexes

The conformation of the binary complexes, MtmC-SAM and MtmC-SAH, are similar. In their crystal structures, the central domain contains a conserved binding pocket that is very similar with the one of TcaB9 for the cofactor and its product (**Figure 30 B and C**). However, we observed a striking structural difference in the conformation of a loop region, which is located between residues 76 and 84 of MtmC and will be called as sensor loop. The backbone of the sensor loop in MtmC can switch its conformation based on whether the substrate or co-substrate occupies their respective binding pockets. In contrast, the conformation of this region in TcaB9 does not show any significant changes in its crystal structures of binary or ternary complexes with the co-substrate, substrate, or products.^[63, 74] This difference was more obvious when we compared the binary and ternary complexes of MtmC.

When only SAM/SAH or only TDP is bound to the binding pocket of MtmC, the conformations of the sensor loop allows relatively open access to the active site for ligands from outside (**Figure 30 B, C, and D**). In contrast, the binding of both SAM and TDP to their respective binding pocket makes the sensor loop adopt a conformation closer to the active site in order to catalyze the methyl-transfer reaction (**Figure 30 E**). This idea is supported by the MtmC-SAH-TDP-KOL complex, in which the sensor loop fits in the closest state and sterically blocks the dissociation of the ligands from the active site (**Figure 30 A**).

MtmC-SAH-TDP-KOL, MtmC-SAM-TDP, and MtmC-TDP structures

In order to mimic the reaction intermediate prior to the catalytic methyl transfer by MtmC, we generated two ternary complexes MtmC-SAH-TDP-KOL and MtmC-SAM-TDP to observe a number of interactions between the amino acid residues and the substrate (**Figure A and E**).^[63] These interactions are similar to those observed in the ternary complex TcaB9-SAH-TDP-3-amino-2,3,6-trideoxy-4-keto-D-glucose, but the residue which forms a van der Waals contact between its *C* β -methylene group and the 5-methyl group of the sugar is His178 instead of Asn177 in TcaB9 (**Figure 31**). The 3-OH group of

TDP-KOL is located in a relatively hydrophilic environment and can form H bonds with the side chains of Glu225 and His182 (**Figure 31**). Tyr77, an amino acid residue on the sensor loop, can also form a weak H bond with this hydroxyl group (with an O-O distance of 3.35 Å), but its conformation does not change significantly in different crystal complexes (**Figure 31**). The amino acid residue acting as the catalytic acid for the catalytic methyl transfer in TcaB9 is His225, which forms a hydrogen bond with the 4-keto group of the TDP-KOL while the residue that plays the same role in MtmC is His226 (**Figure 31**). The methylation catalyzed by MtmC occurs with inversion of stereochemistry at the C-3 position of the substrate, and this is supported by the structure of the ternary complex in which the 3-OH group of the sugar points toward the sulfur atom of SAH (**Figure 31**). In contrast, the 3-amino group of TDP-3-amino-2,3,6-trideoxy-4-keto-D-glucose in the ternary complex of TcaB9 points away from the SAH, indicating that methylation carried out by this enzyme happens with stereochemical retention.

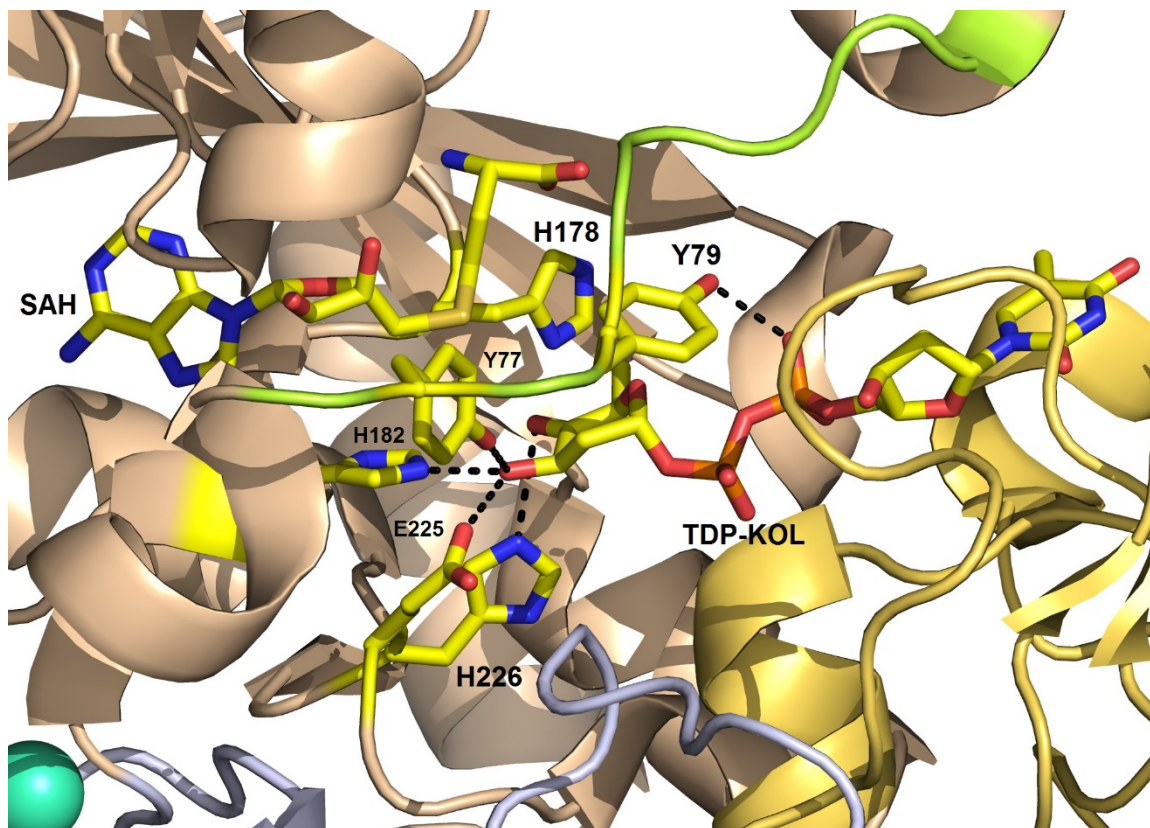


Figure 31. Active site of the ternary complex, MtmC-SAH-TDP-KOL. Amino acid residues involved in the catalysis of methylation are shown and marked here. (C: yellow, O: red, S: orange, P: pink, N: blue, Zn: green cyan) H bond are marked as dashed lines. (The structure of MtmC in complex with the substrate (TDP-KOL) and co-product (SAH) was resolved by Dr. Tsodikov and Dr. Hou in Dr. Tsodikov's laboratory.)

The sensor loop and the Tyr79 residue in it can exhibit a progression of conformation states in the complexes including MtmC-SAM/SAH, MtmC-TDP, MtmC-SAM-TDP, and MtmC-SAH-TDP-KOL. When MtmC only binds to SAM or SAH, Tyr79 points out of the active site (**Figure 32 A**). When only TDP occupies the binding pocket of MtmC, the conformation of the sensor loop remains unchanged, but Tyr79 rotates by 130° where it does not interact with the TDP phosphates (**Figure 32 C**). Compared to the MtmC-SAM and MtmC-SAH complexes, the conformation of the sensor loop in the MtmC-TDP complex is partially disordered, and this indicates that this sensor loop undergoes an order-disorder transition when binding and releasing SAH (**Figure 30 D**). This is probably

because Tyr77, Tyr79, and Thr81 in this loop form a part of the cofactor binding interface, which can be observed in the ternary complexes (**Figure 32**). In the MtmC-SAM-TDP complex, the backbone of the sensor loop around Tyr79 is moved toward the active site for catalysis of the methyl transfer, but Tyr79 points out of the active site probably in order to prevent its aromatic ring from being located in an unfavorable charged environment of the TDP phosphate groups. In this conformation, Tyr79 does not interact with either SAM or TDP (**Figure 32 B**). In the complex MtmC-SAH-TDP-KOL, the sensor loop remains in the same position and conformation while the side chain of Tyr79 moves closer to the substrate (**Figure 32 D**). Its position is different from that in the MtmC-SAM complex by 180° and at an appropriate position to cap the substrate by the steric contact between its phenol ring and the nonpolar sugar-β-phosphate junction as well as the formation of a H bond between the hydroxyl group of Tyr79 and the α-phosphate (**Figure 31**). In this conformation, Tyr79 can also interact sterically with SAH. The conformation of the sensor loop and Tyr79 in this ternary complex, MtmC-SAH-TDP-KOL, is similar to what we have observed in all structures of TcaB9 no matter whether it binds to the co-substrate, substrate, or products. In other words, compared to TcaB9, MtmC shows conformational plasticity of the sensor loop and its residues, especially Tyr79. Besides, we have observed that water molecules fill the empty co-substrate or substrate binding pockets in all of the complexes, except MtmC-SAH-TDP-KOL. Specifically, a water molecule is invariably found in place where the 3-OH group of the TDP-KOL is positioned in the MtmC-SAH-TDP-KOL complex (**Figure 32 A, B, and C**).

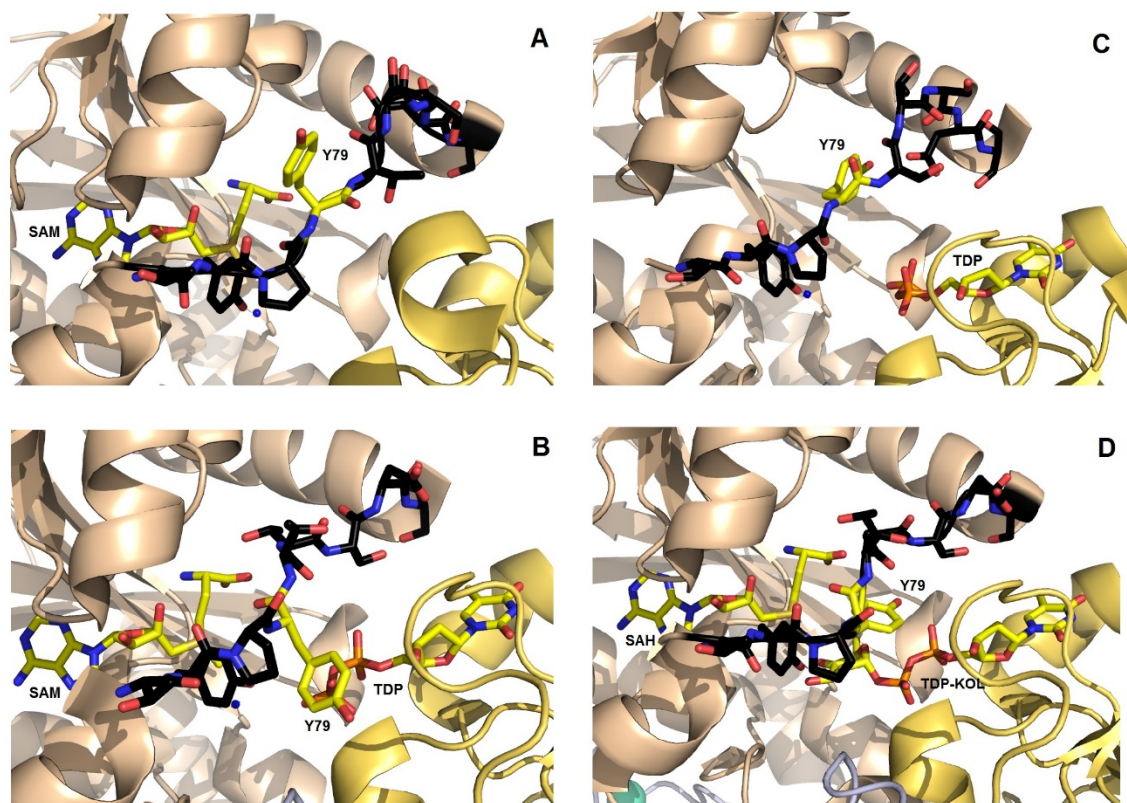


Figure 32. The conformational changes of Y79 residue in (A) MtmC-SAM, (B) MtmC-SAM-TDP, (C) MtmC-TDP, and (D) MtmC-SAH-TDP-KOL complexes. The sensor loop is colored in black, except the Y79 residue that is colored in yellow, i.e. as in the **Figure 30** and **31**. (C: black/yellow, O: red, S: orange, P: pink, N: blue, Zn: green cyan). The water molecule is shown as a sphere with blue color. (The structures of MtmC crystal complexes was resolved by Dr. Tsodikov and Dr. Hou in Dr. Tsodikov's laboratory.)

Methyltransferase activity of wild-type and mutant MtmCs

In order to confirm the importance of Tyr79 for the enzymatic mechanisms of MtmC, we measured the rate of SAH formation upon the generation of TDP-4-keto-D-mycarose (**60**) through *in vitro* assays. In order to understand the importance of Tyr79, we evaluated the relative activity of wild-type MtmC and its mutants instead of measuring their kinetic properties.^[45] We found that both of the MtmC mutants (Tyr79Ala and Tyr79Phe) were approximately 6.25-fold less active than the wild type (**Figure 33**). This data supported our proposed role for Tyr79 residue in positioning the substrate in the appropriate orientation

for catalysis through the interaction between its hydroxyl group and the α -phosphate of the TDP moiety in the substrate (**Figure 31** and **32 D**).

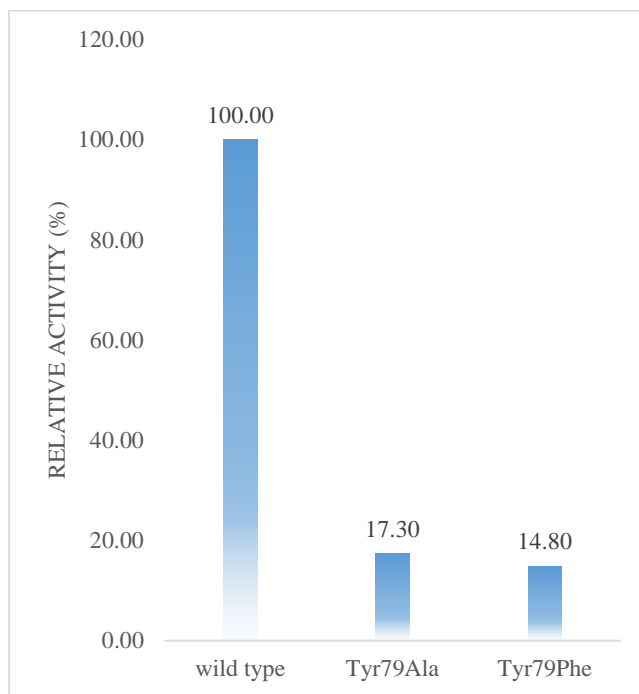


Figure 33. Relative activity of wild-type MtmC and its mutants (Tyr79Ala and Tyr79Phe), relative to that of the wild-type enzyme.

4.4 Discussion

MtmC and MtmGIV are important enzymes with multiple functions involved in the biosynthesis of the trisaccharide side chain of MTM. Understanding of their enzymatic mechanisms is necessary toward generation of novel MTM analogues through protein engineering. As a stepping stone to elucidate how these two enzymes cooperate, we generated and determined the binary and ternary crystal complexes of MtmC with its biologically relevant ligands. According to the amino acid sequence and crystal structure, MtmC possessed a highly conserved TDP-sugar binding pocket in the bacterial class I family SAM-dependent sugar C-methyltransferases. The only access to the sugar moiety without any major protein conformational change is *via* the SAM-binding channel so that MtmC must use the same channel to bind either NADPH or SAM, depending on the

catalytic function. Our structural data and biochemical data show that the sensor loop of MtmC, especially the Tyr79 residue, likely plays an important role in removal of water from the active site, the control of substrate and co-substrate binding, and product release. With the ability to form a bidentate interaction with both the sugar and the TDP moieties of the substrate, Tyr79 is a key amino acid residue for MtmC to discriminate TDP-sugar from TDP or other ligands from the environment. The absence of this interaction in the complex MtmC-SAM-TDP supports this idea. Moreover, the hydrogen bond formation between the hydroxyl group of Tyr79 and the TDP moiety is an important factor for the catalytic turnover because the catalytic turnover of the Tyr79Phe mutant is obviously less active than the wild type. The interaction between Tyr79 and SAM is also probably beneficial for the catalysis.

The active site of TcaB9, a monofunctional methyltransferase, is too constricted for binding of any co-substrate larger than SAM, like NADPH, because its Tyr78, a homologous residue to Tyr79 of MtmC, is always located inside the active site (the “in” conformation). We propose that MtmC has uniquely evolved to accommodate NADPH by swinging the conformation-flexible Tyr79 out of the active site to create more room for NADPH binding (the “out” conformation). Tyr78 of TcaB9 does not undergo this kind of conformational change and thereby it cannot utilize NADPH for reduction reaction. We built a model of MtmC with NADPH binding to the co-substrate pocket in the Tyr79 “out” conformation with the ribose ring of SAM or SAH occupying the same site (**Figure 34**). In this model, the nicotinic acid moiety of NADPH is in place of the phenol ring of Tyr79 in the “in” conformation while its adenine base is exposed to the environment. This is because the co-substrate binding pocket of MtmC is highly conserved to mainly accommodate SAM, like other homologous enzymes in this family. When determining the structure of the crystals grown or soaked with NADPH or NADPH and the substrate, we observed only partial density for NADPH. This reflects a fact that NADPH may be only loosely bound to MtmC, and the affinity is insufficient for building its structure in the active site of MtmC for crystal generation.

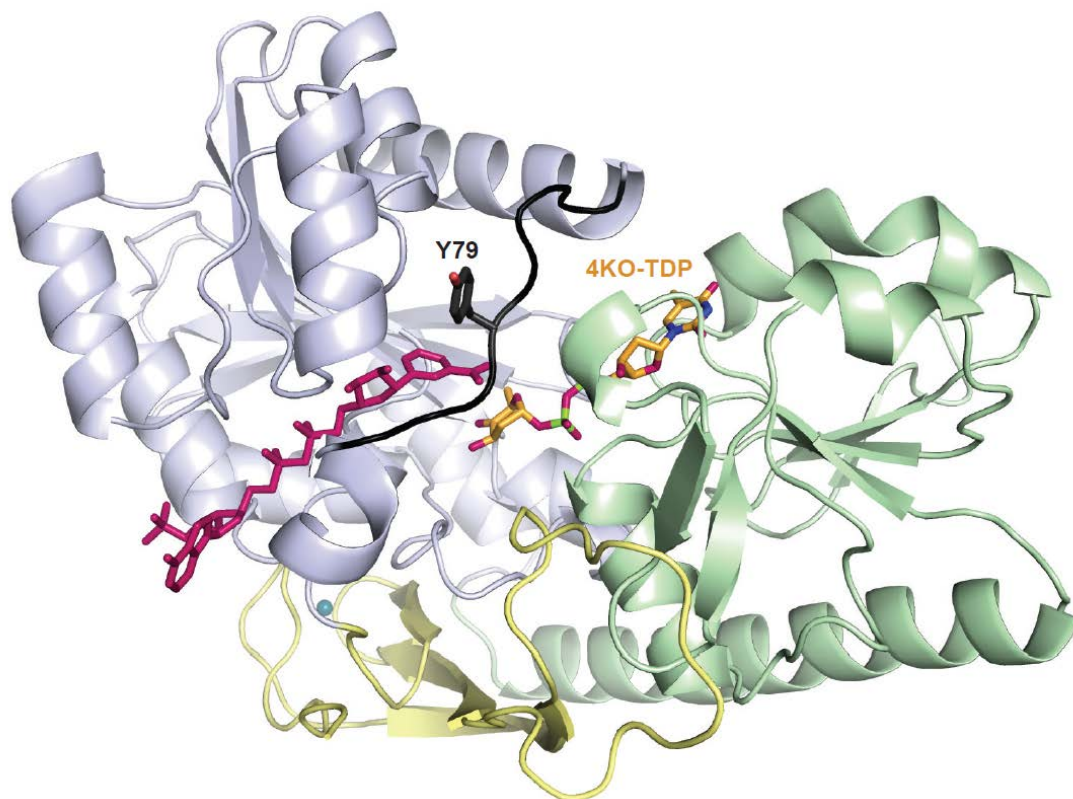


Figure 34. Model of the MtmC-NADPH-TDP-KOL complex. In this model, the conformation of MtmC from the structure of MtmC-SAM was used. (This model was generated by Dr. Tsodikov in Dr. Tsodikov's laboratory.)

In conclusion, MtmC possesses conformational flexibility that its other homologues do not have so that MtmC can utilize both SAM and NADPH as cofactors to catalyze reduction and methylation on TDP-KOL. While both the sugar donors (**59** and **61**) for the biosynthesis of the first and third sugar moieties on the trisaccharide side chain of MTM are generated and transferred by MtmC and MtmGIV, the biosynthesis of the second sugar moiety, D-oliose (**58**), seems independent and not involved in the cooperation between MtmC and MtmGIV. However, there are still many intriguing questions that need to be answered. For example, how is the product of the first MtmC/MtmGIV reaction passed onto MtmGIII and then back to MtmC/MtmGIV? Is it possible that MtmC, MtmGIV, MtmTIII, and MtmGIII are assembled into a large multienzyme complex through uncharacterized protein-protein interactions? We hope that these questions could be answered in the near future.

4.5 Materials and methods

General materials

Thymidine 5'-monophosphate (TMP) disodium salt hydrate, thymidine 5'-diphosphate (TDP) sodium salt, β -nicotinamide adenine dinucleotide phosphate (NADPH) reduced tetra(cyclohexylammonium) salt, MES, ammonium acetate, and PEG 4000 were purchased from Sigma-Aldrich, Co. (St. Louis, MO, USA). NdeI, XhoI, buffers for restriction enzymes, *S*-adenosylmethionine solution (32 mM) was purchased from New England Biolabs, Inc. (Ipswich, MA, USA). Tris-acetate-EDTA (TAE) buffer, tris base, agarose, ethidium bromide, ampicillin sodium salt, and Slide-A-Lyzer[®] dialysis cassette (20 K MWCO, 0.5 – 3.0 mL capacity) were purchased from Thermo Fisher Scientific, Inc. (Hampton, NH, USA). Vector pET22b(+) was purchased from Novagen of Merck KGaA (Darmstadt, Germany). The QuikChange kit for site-directed mutation and Pfu Ultra DNA polymerase were purchased from Agilent Technologies, Inc. (Santa Clara, CA, USA). DNA polymerase for high-GC content DNA was purchased from Clontech Laboratories from Takara Bio, Inc. (Otsu, Shiga, JP). All necessary primers for PCR were synthesized by Integrated DNA Technologies, Inc. (Coralville, IA, USA). 24-well protein crystal growth trays were purchased from Hampton Research Corp. (Aliso Viejo, CA, USA). HisTrap[™] HP column for protein purification was purchased from GE Healthcare of General Electric, Co. (Fairfield, CT, USA). The column installed on FPLC was a size-exclusion Sephacryl S-200 column, also purchased from GE Healthcare. Microscope used to observe crystal formation was purchased from Nikon, Co. (Shinagawa, Tokyo, JP). Liquid nitrogen used to rapidly freeze crystals was purchased from Scott Gross Company, Inc. (Lexington, KY, USA). PCR was carried out in our laboratory with the PCR machine, GeneAmp[®] PCR system 2700, purchased from Applied Biosystems of Thermo Fisher Scientific, Inc. (Hampton, NH, USA). DNA sequencing was carried out by ACGT, Inc. (Wheeling, IL, USA) and GENEWIZ, Inc. (South Plainfield, NJ, USA). The type of the FPLC machine (BioLogic DuoFlow) used for protein purification was purchased from Bio-Rad Laboratories, Inc. (Hercules, California, U.S.A). The UV spectrophotometer, UV-1800, used to determine the protein concentration was purchased from Shimadzu Corp. (Kyoto, JP) X-ray diffraction experiments were carried out by the staff of sectors 21 and

22 of the Advanced Photon Source at the Argonne National Laboratory (Argonne, IL, USA). Other general materials used in this study are described in Chapter 2.5.

Protein expression and purification

The original published sequence of the *mtmC* gene from *S. argillaceus* contained several mistakes; the corrected DNA sequence was deposited into GenBank as entries GUSub26196 and GUSub26197 for the nucleotide and amino acid sequences in 2014 November, respectively. *mtmC* was constructed in the vector pET22b(+) purchased from Novagen for protein over-expression. The protein contains an *N*-terminal 6 × His tag and could be purified with immobilized metal affinity chromatography (IMAC). After amplifying the *mtmC* gene in *E. coli* XL1 Blue, plasmid isolation were carried out with GeneJet plasmid miniprep kit and then used to transform *E. coli* BL21 (DE3) competent cells for protein over-expression. 1% volume of seed *E. coli* BL21 (DE3) culture with desired plasmid was inoculated into 1 liter of LB supplemented with 100 µg/mL ampicillin (final concentration) and then incubated at 37 °C until OD₆₀₀ value reached 0.2~0.3. The *E. coli* culture was transferred to 18 °C for additional 90-minute incubation, followed by the addition of 100 µM β-D-1-thiogalactopyranoside (IPTG, final concentration) to induce protein over-expression, and target proteins were obtained after overnight incubation at 18 °C. *E. coli* BL21 (DE3) cell pellets were collected by centrifugation (4000 rpm for 25 min) at room temperature, washed twice with 20 mL lysis buffer (40 mM Tris, 400 mM NaCl, 10% glycerol, 2 mM β-ME, pH 8.0), re-suspended in 25 mL lysis buffer for sonication, and finally centrifuged at 16,500g for 35 min at 4 °C to remove cell debris. The supernatant containing target proteins was loaded on HisTrap™ HP column that contains Ni²⁺ and had been equilibrated with 10-fold volume lysis buffer in advance. Impurities and target proteins were eluted by washing buffer (40 mM Tris, 400 mM NaCl, 10% glycerol, 2 mM β-ME, 20 mM imidazole, pH 8.0) and elution buffer (40 mM Tris, 400 mM NaCl, 10% glycerol, 2 mM β-ME, 200 mM imidazole, pH 8.0), respectively. We used Slide-A-Lyzer® dialysis cassette (20 K MWCO) to dialyze the protein in dialysis buffer (20 mM Tris, 100 mM NaCl, 10% glycerol, 2 mM β-ME, pH 7.5), followed by concentration with Millipore Amicon Ultra-15 ultracentrifugal filters (30K) to approximately 80 µM. MtmC which was desalted through dialysis was used for *in vitro* assays instead of generating crystals. For

crystallization, the purified protein was further purified with a size-exclusion Sephacryl S-200 column (GE Healthcare) equilibrated in 40 mM Tris-HCl (pH 8.0, adjusted at room temperature), 100 mM NaCl, and 2 mM β -ME. The fractions were checked with SDS-PAGE gel, and samples containing the desired protein were mixed and concentrated by using an Amicon ultracentrifugal filters to 12 mg/mL. The measured sizes of purified proteins were in agreement with the calculated size (**Figure 35**).

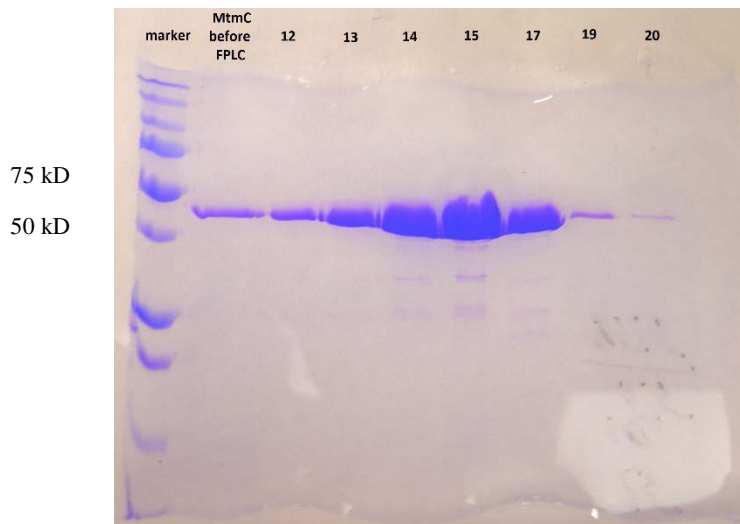


Figure 35. MtmC sample before purification by FPLC and the 12nd - 20th fractions collected from FPLC experiment.

Crystallization, data collection, and crystal structure determination

In the beginning, we followed a sparse incomplete factorial screen (Hampton Research Crystal Screen) to search for an initial crystallization condition by vapor diffusion in hanging drops at 21 °C. After optimizing the initial condition, we set drops containing 1 μ L of MtmC (12 mg/mL) with 1 mM co-substrate or substrate and 1 μ L of the reservoir solution (100 mM MES, pH 5.5, 200 mM ammonium acetate, and 16% PEG 4000). The growing crystals were gradually transferred into the cryoprotectant buffer (100 mM MES, pH 5.5, 200 mM ammonium acetate, 16% PEG 4000, and 20% glycerol) and frozen in liquid nitrogen immediately.

X-ray diffraction data were collected at 100 K at beamlines 21ID-G (for MtmC-SAM-TDP crystals) and 22ID (for other crystals) of the Advanced Photon Source at the Argonne

National Laboratory and then processed with HKL2000.^[75] The structure of the MtmC-SAM-TDP complex was resolved by using molecular replacement with MOLREP with the structure of TcaB9, one of the homologues of MtmC which had been resolved structurally, as a search model.^[76] The obtained structure was then iteratively rebuilt and refined with COOT and REFMAC.^[77] The structures of the co-substrate (SAM) and substrate (TDP) were modeled without uncertainty into strong omit $F_o - F_c$ electron density. The refined structure of MtmC-SAM-TDP complex was used as a starting point to rebuild and refine all other structures of MtmC complexes. The data collection and refinement statistics are listed in **Table 16**. Crystal structures resolved in this study were deposited in the Protein Data Bank (PDB) as entries 4RV9 (MtmC-SAH), 4RVD (MtmC-SAM), 4RVF (MtmC-TDP), 4RVG (MtmC-SAM-TDP), and 4RVH (MtmC-SAH-TDP-KOL).

Site-directed mutagenesis of MtmC

In order to confirm our ideas about the enzymatic mechanism of MtmC, a site-directed mutant of the key amino acid residue for the reaction, Tyr79, to replace it with either Phe or Ala was prepared by Dr. Hou in Dr. Tsodikov's laboratory. MtmC Tyr79Phe and MtmC Tyr79Ala mutants were generated by using the QuickChange kit. We followed the manufacturer's protocol to incorporate the desired mutations into the plasmid, pET22b(+)-*mtmC*. The plasmids containing mutant *mtmC* gene were shipped to the University of Kentucky DNA Sequencing Core for confirmation and used *E. coli* XL1 Blue to generate more plasmids. MtmC mutants was overexpressed in *E. coli* BL21 (DE3) and then purified like the wild-type MtmC for *in vitro* kinetic assay (**Figure 36**).

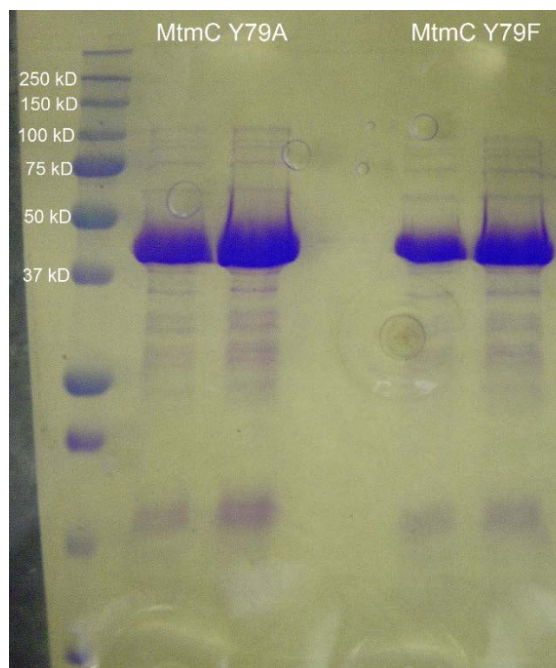


Figure 36. SDS-PAGE gel of the MtmC Tyr79Ala and MtmC Tyr79Phe mutants after dialysis.

Methyltransferase activity assays

The product of the 3-methylation reaction by wild-type MtmC, TDP-4-keto-D-mycarose, could not be observed likely because the axial 3-OH group generated concomitantly with 3-methylation attacked the proximal phosphate to cleave the phosphodiester bond of TDP, resulting in the loss of UV absorbance.^[46b] The cleaved TDP would be mixed with other TDP that originally presented in the sample, and it was difficult to precisely measure the formation of the product. For this reason, we monitored formation of the coproduct SAH. Kinetic assays were conducted for the wild-type and mutant MtmC (Tyr79Phe and Tyr79Ala) by HPLC (Waters 600 system, consisting of a controller, a Waters 996 photodiode array detector, and a Delta 600 pump). A 100 μ L reaction mixture contained 25 mM Hepes (pH 7.5), 50 mM NaCl, 1 mM EDTA, 90 μ M TDP-KOL, 2 mM SAM, and 5 μ M MtmC. The blank control that was used to subtract the background contained the same ingredients but did not include any MtmC. The reactions were initiated by the addition of enzyme and the mixtures incubated at 22 $^{\circ}$ C, to minimize the nonenzymatic degradation of SAM; 50 μ L aliquots were quenched at 60 and 300 min (in the range where the SAH concentration increased linearly over time) with 5 μ L of 1.5 g/mL

trichloroacetic acid (final concentration of 13.6% (w/v)) and then incubated on ice for 10 min. After centrifugation at 13000 rpm for 3 min, 50 μ L of the supernatant was passed through a Phenomenex Kinetex 5 μ m EVO C18 100 Å column (250 mm \times 4.6 mm) and eluted isocratically in 10 mM ammonium formate with 5% methanol (pH 3.0) at a rate of 0.8 mL/min. The area under the chromatogram absorbance peak at 260 nm corresponding to SAH was measured. The rate of conversion of SAM to SAH for the wild-type enzyme and the two mutants was calculated from the 60 and 300 min data points.

Table 16. X-ray diffraction data collection and structural refinement statistics (This table was edited by Dr. Hou and Dr. Tsodikov.)

	MtmC-SAH-TDP-KOL	MtmC-SAM-TDP	MtmC-SAM	MtmC-SAH	MtmC-TDP
Data Collection^a					
space group	<i>I</i> 4 ₁ 22	<i>I</i> 4 ₁ 22	<i>I</i> 4 ₁ 22	<i>I</i> 4 ₁ 22	<i>I</i> 4 ₁ 22
no. of monomers per asymmetric unit	1	1	1	1	1
unit cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	135.4, 135.4, 127.3	134.7, 134.7, 127.8	134.8, 134.8, 129.7	134.2, 134.2, 130.0	134.2, 134.2, 132.3
α , β , γ (deg)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
resolution (Å)	50 – 2.35 (2.39 – 2.35)	50 – 2.3 (2.34 – 2.30)	50 – 2.2 (2.24 – 2.20)	50 – 2.2 (2.24 – 2.20)	50 – 2.7 (2.8 – 2.7)
<i>I</i> / σ	28.2 (2.5)	40.3 (4.7)	39.9 (3.2)	41.8 (3.7)	25.2 (3.3)
completeness (%)	95.7 (98.9)	99.5 (100)	100 (100)	99.0 (100)	100 (100)
redundancy	6.6 (6.5)	9.8 (10)	8.1 (8.2)	9.6 (9.8)	9.7 (10.0)
<i>R</i> _{merge}	0.075 (0.61)	0.077 (0.499)	0.064 (0.625)	0.087 (0.627)	0.129 (0.676)
no. of unique reflections	21204	24771	28993	28774	16056
Structural Refinement					
resolution (Å)	40 – 2.35	40 – 2.3	40 – 2.2	40 – 2.2	40 – 2.7
<i>R</i> (%)	20.6	21.2	21.4	20.3	20.9
<i>R</i> _{free} (%)	23.1	23.8	25.7	24.1	24.8
rmsd from ideal					
bond lengths (Å)	0.005	0.005	0.006	0.006	0.005
bond angles (deg)	1.01	1.08	1.13	1.1	0.97
ramachadran plot ^b (%)					
residues by region)					
most allowed	98.1	97.1	92.2	97.8	90.5
additional allowed	1.9	2.9	7.3	2.2	8.7
generously allowed	0	0	0.6	0	0.3
disallowed	0	0	0	0	0.6 (two residues) ^c

^a Values in parentheses refer to data for the highest-resolution shell.

^b PROCHECK statistics.^[78]

^c The outliers (His178 and Ser82), which are both in or near the active site, fall in the immediate vicinity of the allowed regions of the Ramachandran plot, within the uncertainty of the modest resolution of the MtmC-TDP structure. His178 faces the active site pocket, and its conformation may be somewhat strained by binding of the unnatural TDP ligand; Ser82 is in the sensor loop and is somewhat disordered in this complex.

CHAPTER 5: SUMMARY AND FUTURE DIRECTIONS

In this dissertation, the second and third chapters describe the generation of new natural product analogues with better biological activities by chemoenzymatic methods. The fourth chapter, we focus on the structural determination of MtmC's crystal complexes and the enzymatic actions. The insights we obtained are summarized as follows:

5.1 Specific aim 1: Study and use of enzymes for modification of polycarcin V.

We used polycarcin V (PV) as a model compound for the SAR investigation regarding the sugar moiety. Three heterologous, but substrate-flexible methyl transferases were used to generate four new PV derivatives. According to the cytotoxicity assay, the interaction between the L-rhamnopyranose of PV and histone H3 might be stronger than the D-fucofuranose of GV. In addition, we found that 2'-OH of L-rhamnopyranose is a necessary proton donor for the H bond formation with histone H3. On the other hand, modifying either 3'-OH or 4'-OH with methyl groups might improve the interaction.

Our findings provide a new opportunity for the improvement of GV-type aryl-C-glycosides. Literature shows that the vinyl side chain of gilvocarcin V is not as important as scientists originally proposed because gilvocarcin M (with methyl side chain) still possesses antibiotic and anticancer activities.^[79] In addition, the vinyl group could be replaced by either an epoxide or oxime group without losing anticancer activity.^[27b] In comparison, 2',3'-di-OMe-PV lost its anticancer activity after 2'- and 3'- hydroxyl groups of the sugar moiety were converted to methoxy groups. This reveals an interesting hypothesis – the sugar moiety may be more important for the anticancer activity than the vinyl side chain. Therefore, it is important to understand how the sugar moiety interacts with the proposed target, histone H3, because Akira Matsumoto and colleagues did not provide any details about it.^[26] Looking for the future, scientists need to generate crystal complexes of PV and major histone H3 variants.^[80] The structural insights explored from these crystals may provide helpful information not only to explain the data of our cytotoxicity assay but also to confirm the importance of the sugar moiety of PV. If the anticancer activity of polycarcin V primarily comes from the sugar moiety, the clinical use of GV-type aryl-C-glycosides would not be limited to regional skin diseases. Scientists could replace the vinyl side chain of PV, which needs UV-visible light for activation, with

other functional groups which also help the anticancer behaviors in order to improve its bioactivity and/or water solubility.

5.2 Specific aim 2: Chemical derivatization of mithramycin.

After purifying MTM SA from the culture broth of *S. argillaceus* M7W1, we used PyBOP as the reagent to couple MTM SA with several natural and unnatural primary amines. In addition, we used TMSCHN₂ to transfer a methyl group onto the carboxylic acid group of MTM SA and generated MTM SA methyl ester. According to the data of the cytotoxicity assay and NCI 60 cell line screen, MTM SA methyl ester, MTM SA-Phe, MTM SA-Trp, and MTM SA-Lys-2PGs are the most potential MTM SA analogues. Our hypothesis is that modifying MTM SA with amino acids with aromatic rings could improve the anticancer activity. The possible explanation is that the aromatic rings could enhance the interaction between the C-3 side-chain and DNA through hydrophobic stacking with the pyrimidine and purine functional groups. In addition, MTM SA-Trp and MTM SA-Phe are potential lead structures against Ewing's sarcoma because of their excellent specificity and *in vivo* tolerance.

Looking for the future, we would like to improve the bioactivities of MTM SA-Phe and MTM SA-Trp against Ewing's sarcoma because MTM was reported as the only lead structure, which could act on the previously believed undruggable target (EWS-FLI1).^[43] In order to achieve this goal, we need to understand the mechanism of the two MTM analogues and explain why they can show higher specificity and *in vivo* tolerance than the parent compound. Our hypothesis is that the aromatic ring on their modified C-3 side chains can specifically interact with the FLI1 domain. If our hypothesis is true, the aromatic ring could be further modified to improve the compound-protein interaction. We and our collaborators are working to generate crystal complexes of MTM analogues, FLI1, and short DNA sequences in order to demonstrate this hypothesis. Moreover, the recently discovered MTM analogue, demycarosyl-3D-β-D-digitoxosylmithramycin SK, showed 18 folds higher maximum tolerated dose in animal experiments and better anticancer activity than MTM.^[14a, 81] This is the most potent MTM-type antitumor antibiotic so far. This reveals that modification of the sugar E is a valuable strategy for the development of novel MTM analogues. If we could find a method to modify the sugar E of both MTM SA-Phe

and MTM SA-Trp, these compounds may show higher potential for the clinical treatment against Ewing's sarcoma. Compared to chemical methods, combinatorial biosynthesis and enzymatic modification are more feasible and efficient methods for this work.

5.3 Specific aim 3: Structural investigation into MtmC.

We obtained three binary and two ternary complexes of MtmC with its co-substrate (SAM), substrates (TDP or TDP-4-keto-D-olivose), and co-product (SAH) for X-ray diffraction experiments. The quality of these crystal complexes was sufficient to provide structural information of MtmC in the range of 2.2 – 2.7 Å resolution. Moreover, the ligands were very clearly resolved in the strong omit $F_o - F_c$ electron density map in all of the binary and ternary complexes. The sensor loop of MtmC, especially the Tyr79 residue, possesses conformational flexibility to move “in” and “out” of the active site. In order to confirm the importance of the Tyr79, we generated two MtmC mutants, Tyr79Phe and Tyr79Ala. Both of them showed approximately 6.25 times lower catalytic activity than the wild-type MtmC. Our biochemical data had proved the importance of the hydrogen bond formation between the hydroxyl group of Tyr79 and the α -phosphate group of TDP moiety in the substrate. Moreover, because MtmC possessed a highly conserved TDP-sugar and SAM binding pocket and there is no other channel which can allow NADPH to approach the substrate, MtmC must use the same co-substrate binding pocket to accommodate either NADPH or SAM, depending on the catalytic function. Therefore, we proposed that MtmC has uniquely evolved to accommodate NADPH, a co-factor larger than SAM, by swinging the conformation-flexible Tyr79 out of the active site to create more room for NADPH binding. However, the mechanism of the reduction reaction is still under investigation.

There are many intriguing questions waiting for investigations in order to clarify the formation process of MTM's trisaccharide side chain. Recent literature suggests that post-PKS tailoring enzymes can be multifunctional and co-dependent on other tailoring enzymes. Our previous work also showed that MtmC cooperated with MtmGIV during the formation of premithramycin A1 and premithramycin A3, and MtmGIII acts between the two glycosylation reactions catalyzed by MtmGIV (**Figure 19**).^[45, 46b] Moreover, our unpublished data also indicated that MtmOIV and MtmW worked together during the formation of the C-3 side chain (**Figure 17**). Looking for the future, we would like to

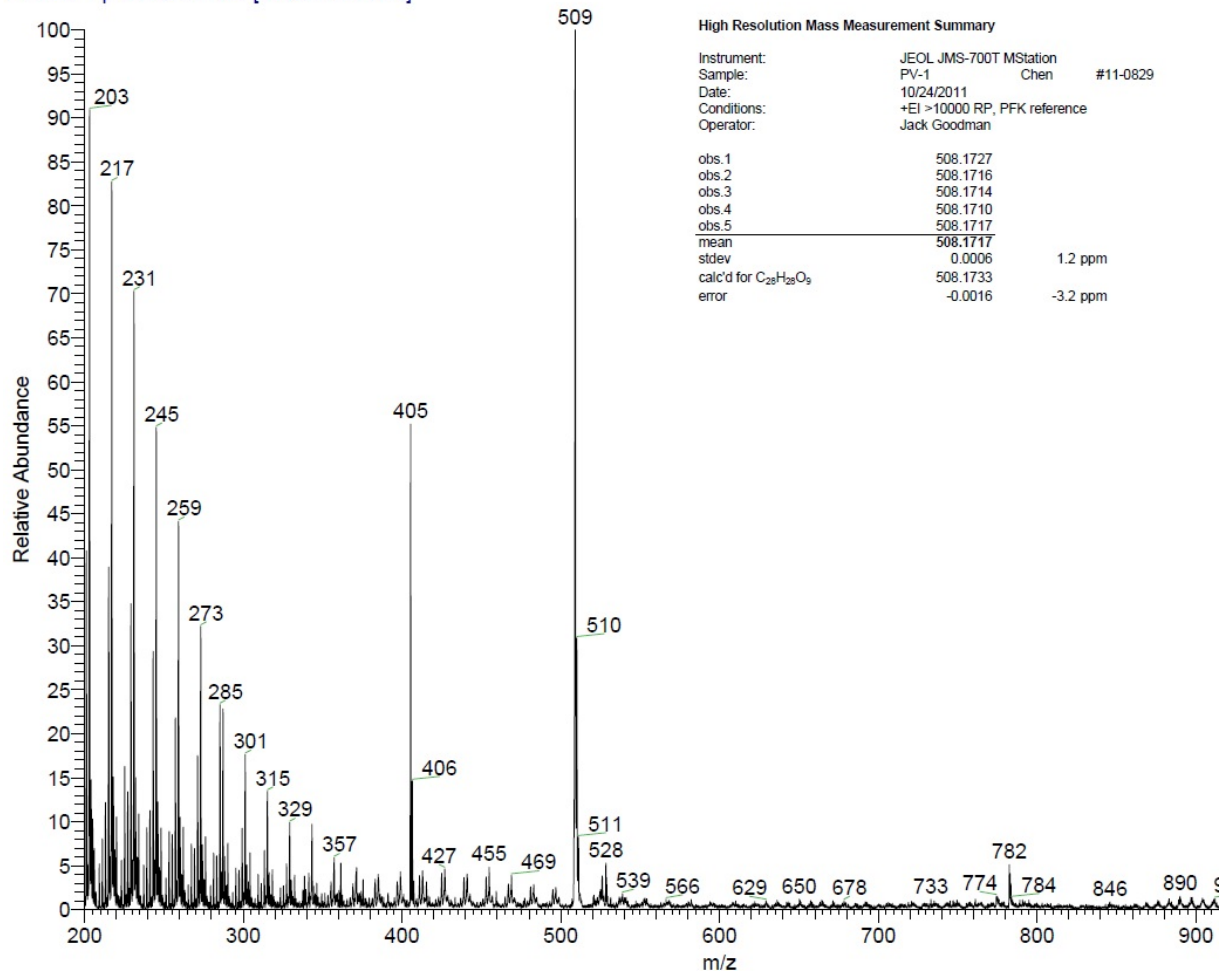
identify the biochemical mechanisms as well as the protein-protein interaction between these post-PKS tailoring enzymes by crystallography and *in vitro* assays. When overexpressing *mtmGIII*, *mtmGIV*, and *mtmW* genes in *E. coli* BL21 (DE3), we found that these proteins formed inclusion bodies. This problems could be resolved by co-expressing chaperones and glycosyltransferases together, and, thereby, chaperones would assist the 3D-structural folding of the glycosyltransferases. In addition, after analyzing the amino acid sequence of MtmW, we found that MtmW contained three hydrophobic amino acids on the *N*-terminal side, but its homologs don't. These are potential membrane anchors, since MtmOIV/MtmW play also an important role to activate the metabolites towards DNA binding. It is likely that the membrane anchor ensures that the MtmOIV/MtmW proteins are near an efflux membrane channel to avoid damage of the DNA of the bacterial cell. The solubility of MtmW could be improved by removing these extra amino acids. With these great progresses, we hope that our questions could be answered in the near future. These information may not only help identify how the substrates and products are transferred between different enzymes but also solve the remaining mystery of MtmC's reduction mechanism in this dissertation. Maybe someday in the future, we would be able to generate more MTM analogues with refined sugar moieties and C-3 side chain by using engineered post-PKS tailoring enzymes.

5.4 Conclusions

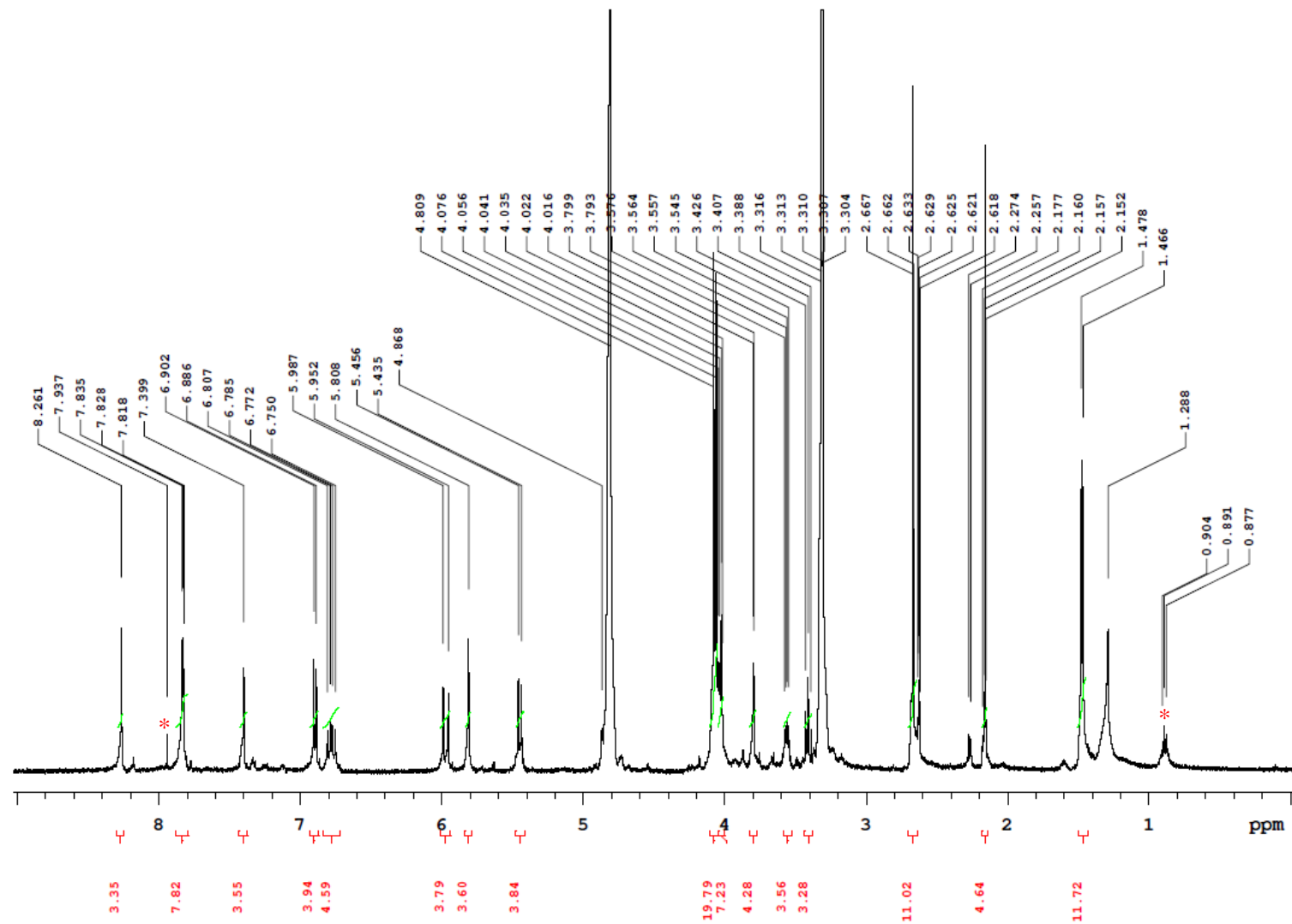
In conclusion, natural products will continue to be an important source for new drug development, especially the small-molecule agents. Their chemical space is larger than synthetic compounds and could be expanded through chemical synthesis and combinatorial biosynthesis in order to improve their bioactivity, specificity, and *in vivo* tolerance. With the high-speed development of synthetic biology, protein engineering, and organelle engineering, scientists will be able to use living cells as a platform to generate new natural products in the future. This is not only a good method to modify natural products efficiently but also a new opportunity to decrease the usage of hazardous chemicals to protect our environment.

APPENDICES

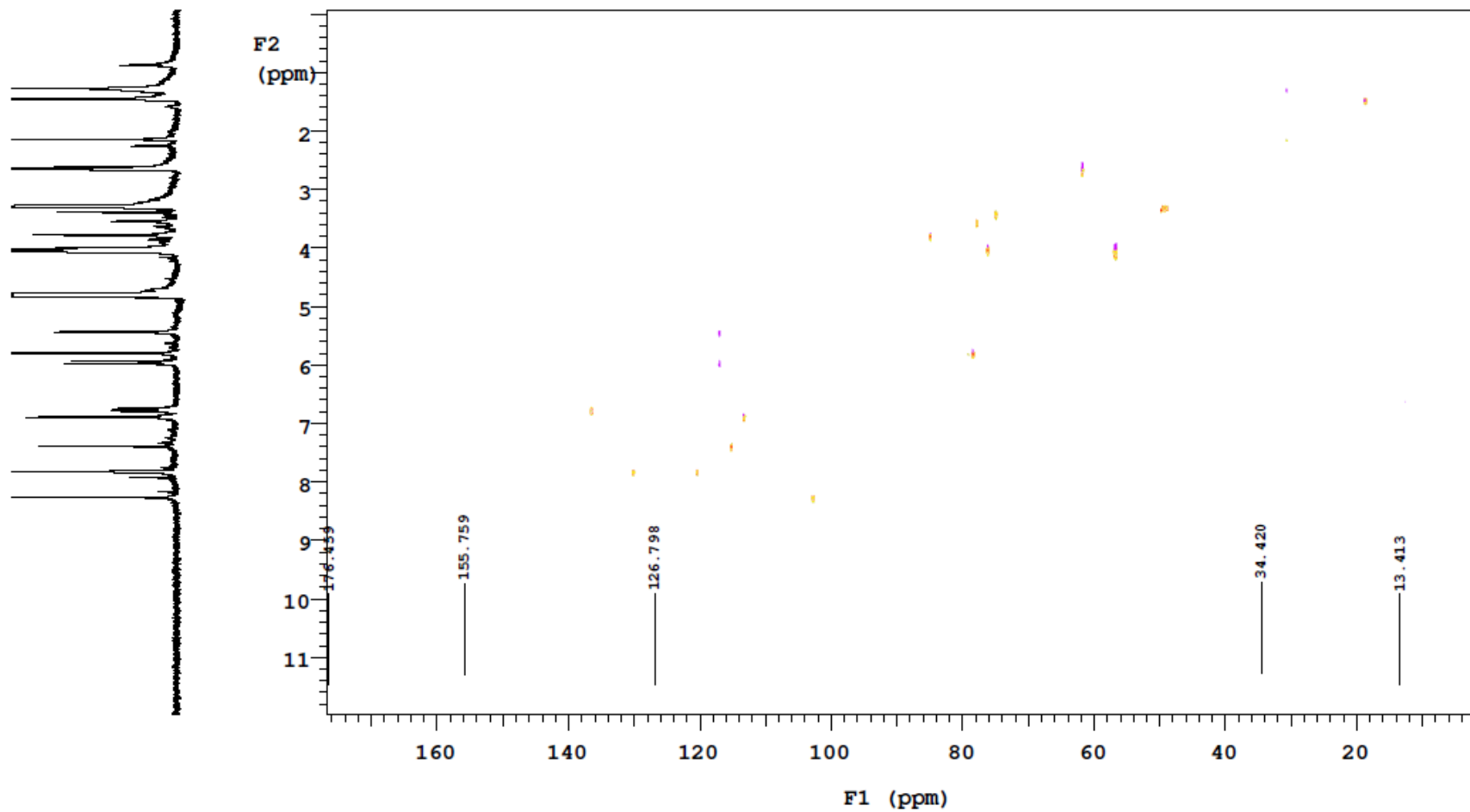
11-0803 #104-115 RT: 2.11-2.33 AV: 12 NL: 4.48E4
 T: ITMS + p ESI E Full ms [200.00-1000.00]



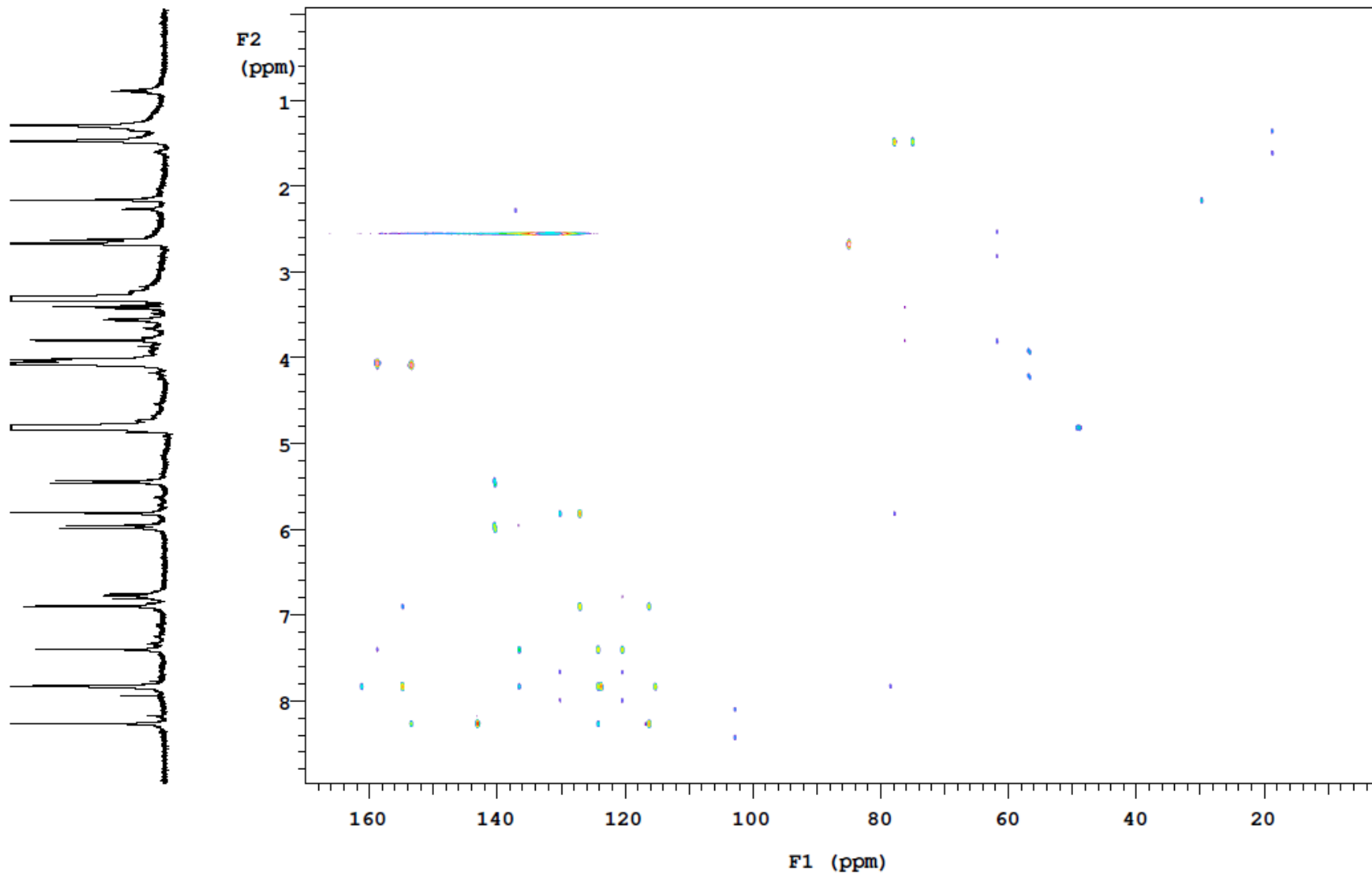
Appendix 1. ESI-MS spectrum (positive mode) and summary of HR-EI-MS data of 2'-*O*-methyl-polycarcin V (**60**)



Appendix 2. ^1H NMR spectrum of 2'-O-methyl-polycarcin V (60) (*: signals from minor impurities)

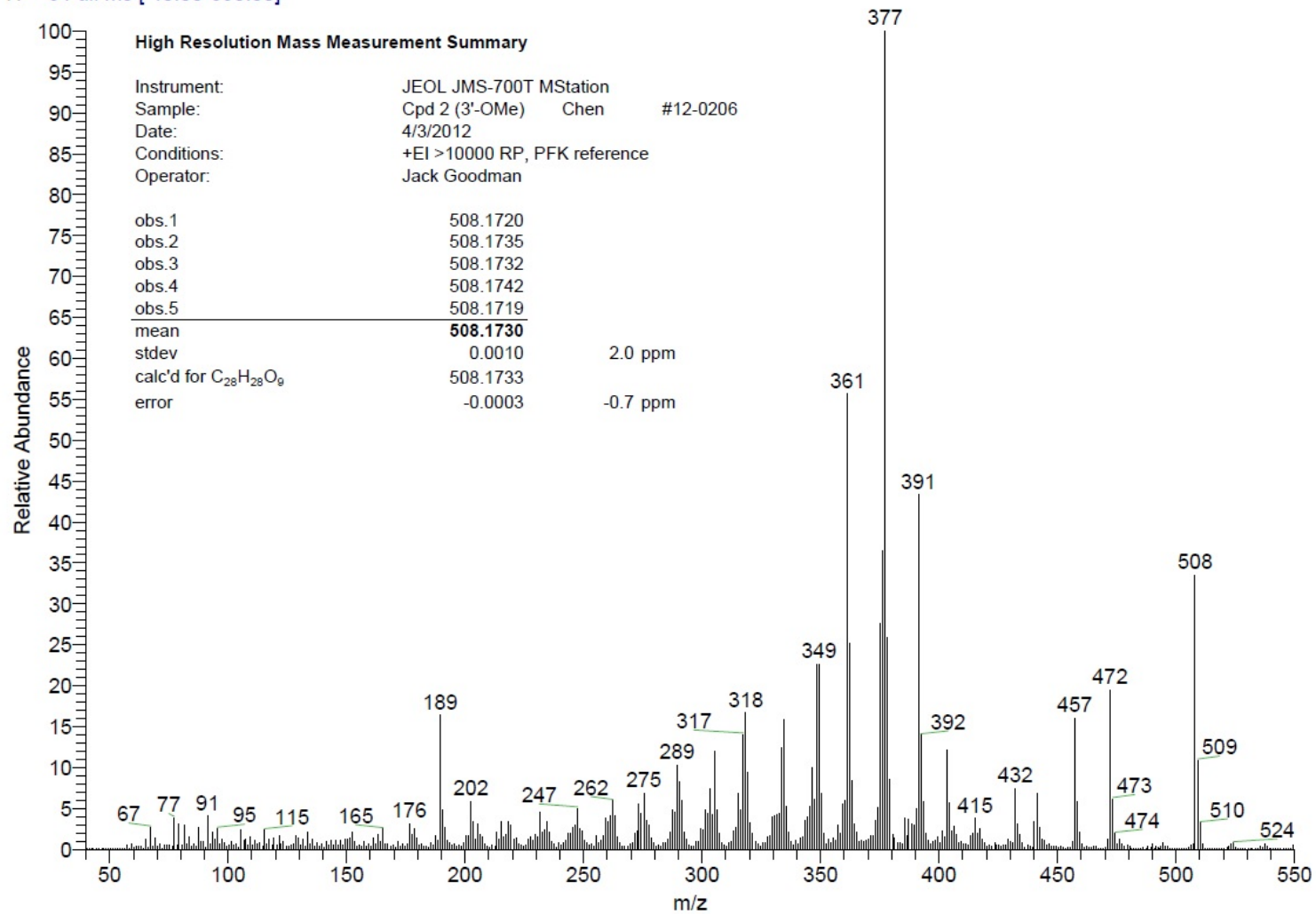


Appendix 3. gHSQC NMR spectrum of 2'-*O*-methyl-polycarcin V (**60**)



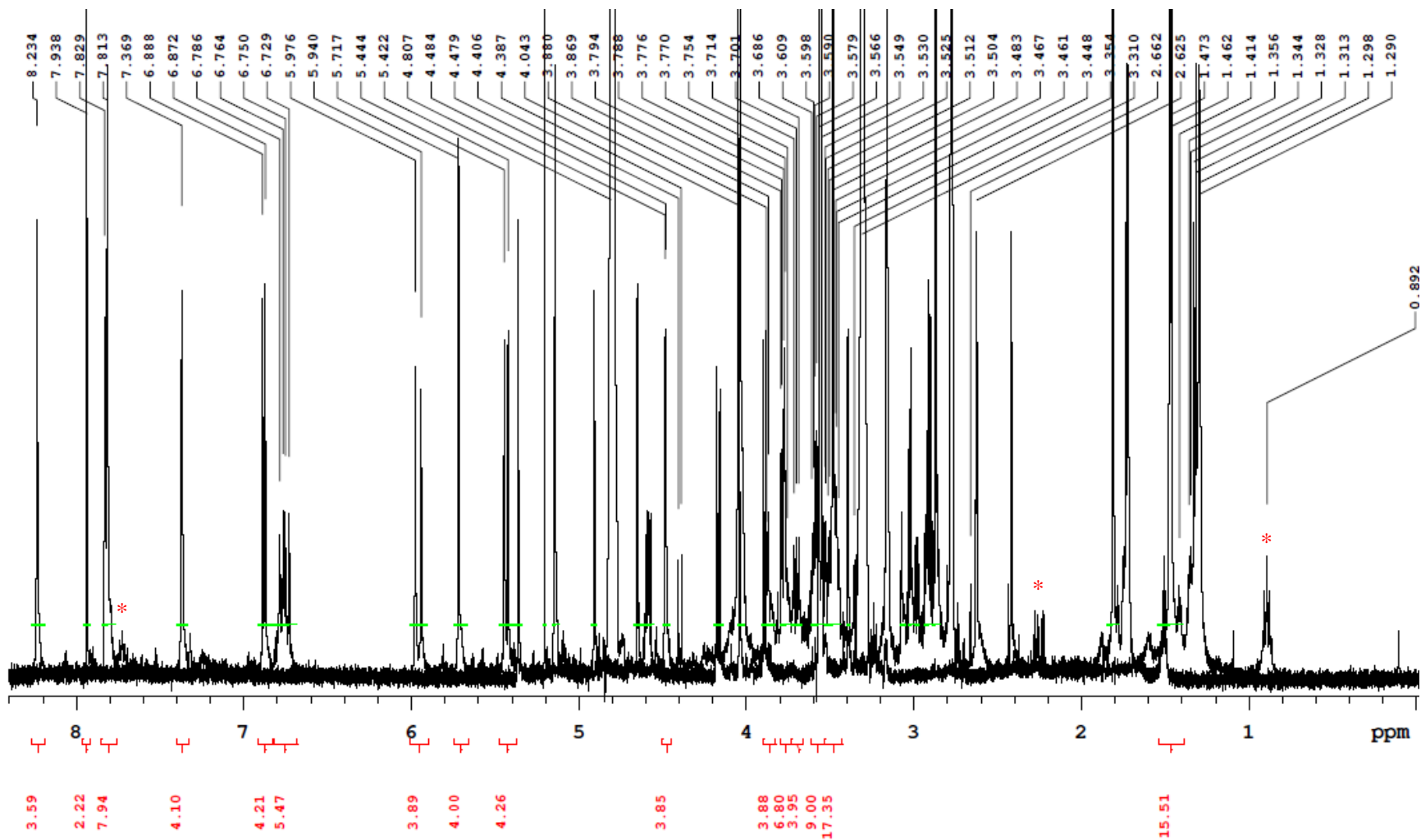
Appendix 4. gHMBC NMR spectrum of 2'-*O*-methyl-polycarcin V (**60**)

12-0206 #254-302 RT: 4.20-5.04 AV: 49 NL: 9.46E4
T: + c Full ms [40.00-650.00]

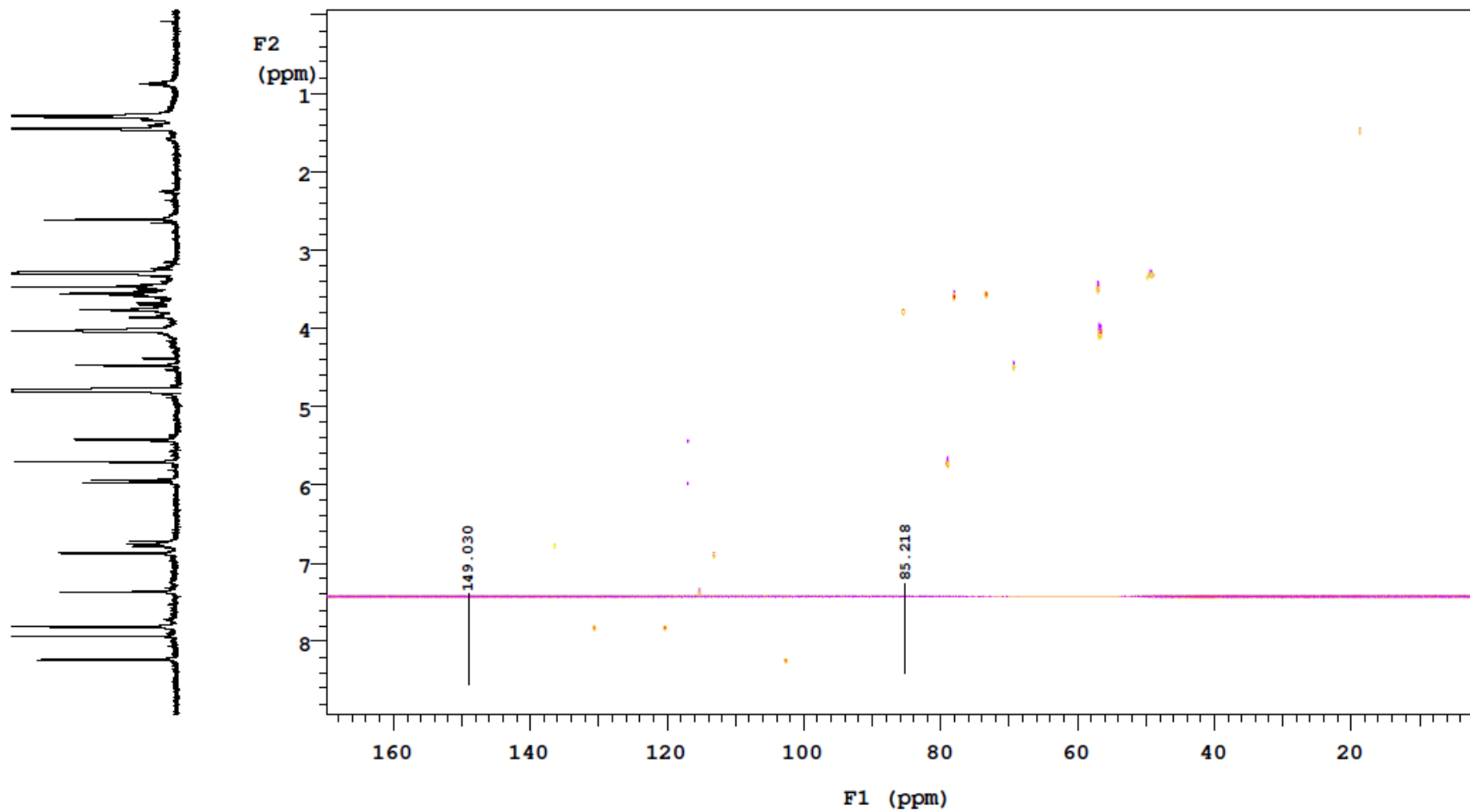


Appendix 5. ESI-MS spectrum (positive mode) and summary of HR-EI-MS data of 3'-O-methyl-polycarcin V (**61**)

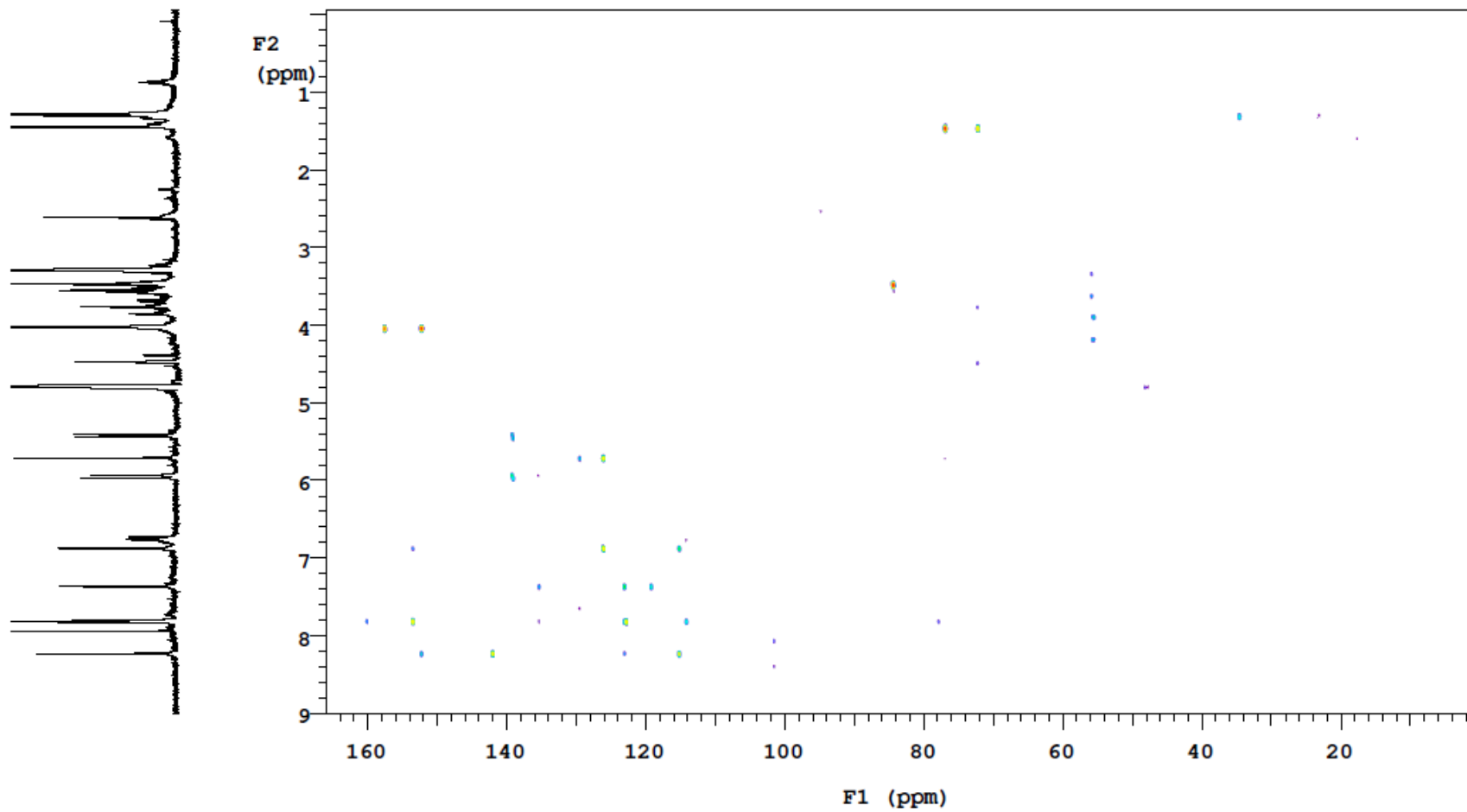
LII



Appendix 6. ¹H NMR spectrum of 3'-O-methyl-polycarcin V (61) (*: signals from minor impurities)

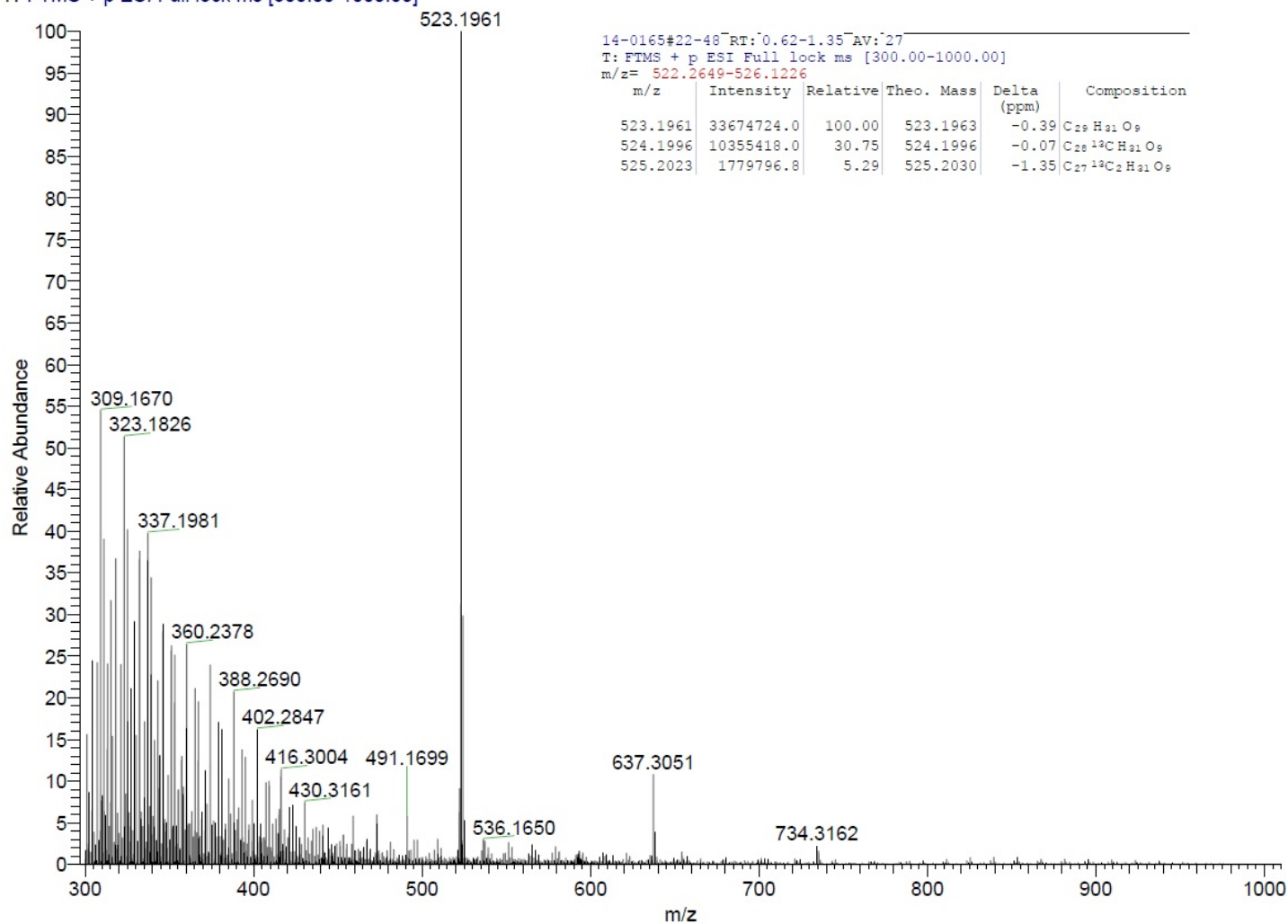


Appendix 7. gHSQC NMR spectrum of 3'-O-methyl-polycarcin V (**61**)



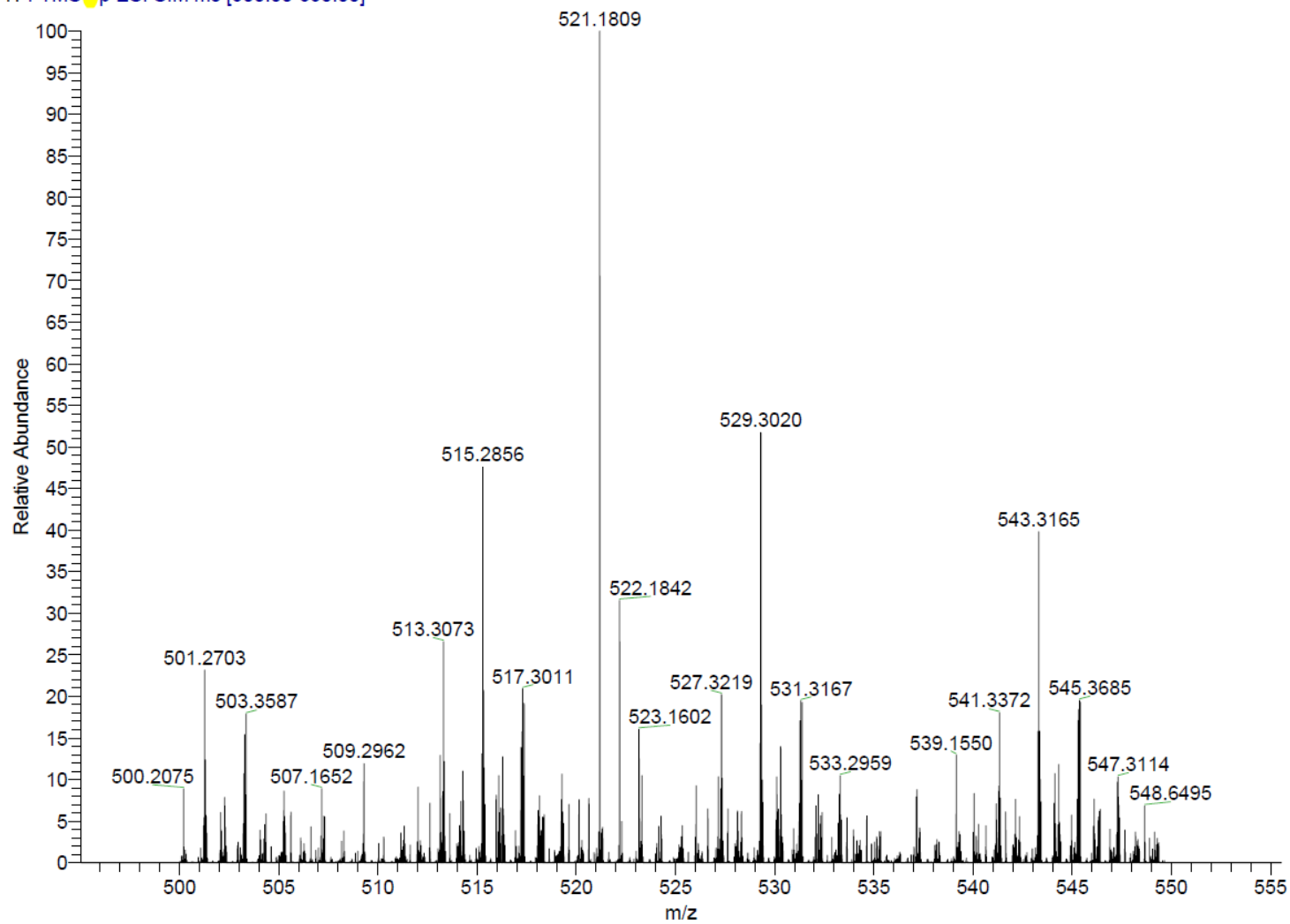
Appendix 8. gHMBC NMR spectrum of 3'-O-methyl-polycarcin V (**61**)

14-0165 #22-48 RT: 0.62-1.35 AV: 27 NL: 3.35E7
T: FTMS + p ESI Full lock ms [300.00-1000.00]

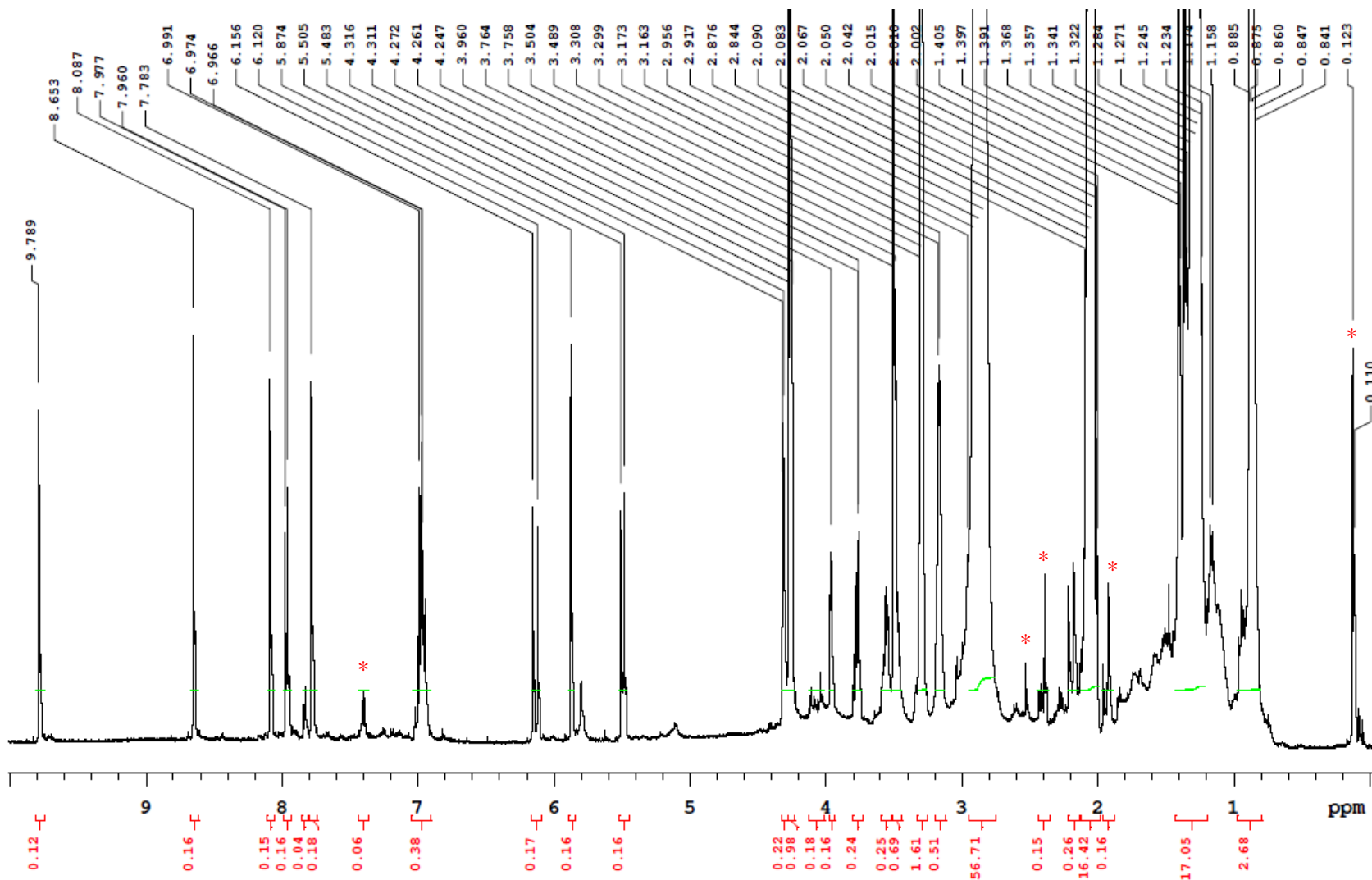


Appendix 9-1. HR-ESI-MS spectrum (positive mode) of 2',3'-di-O-methyl-polycarcin V (**63**)

14-0165 #87-132 RT: 2.57-4.08 AV: 46 NL: 4.25E4
T: FTMS -p ESI SIM ms [500.00-550.00]

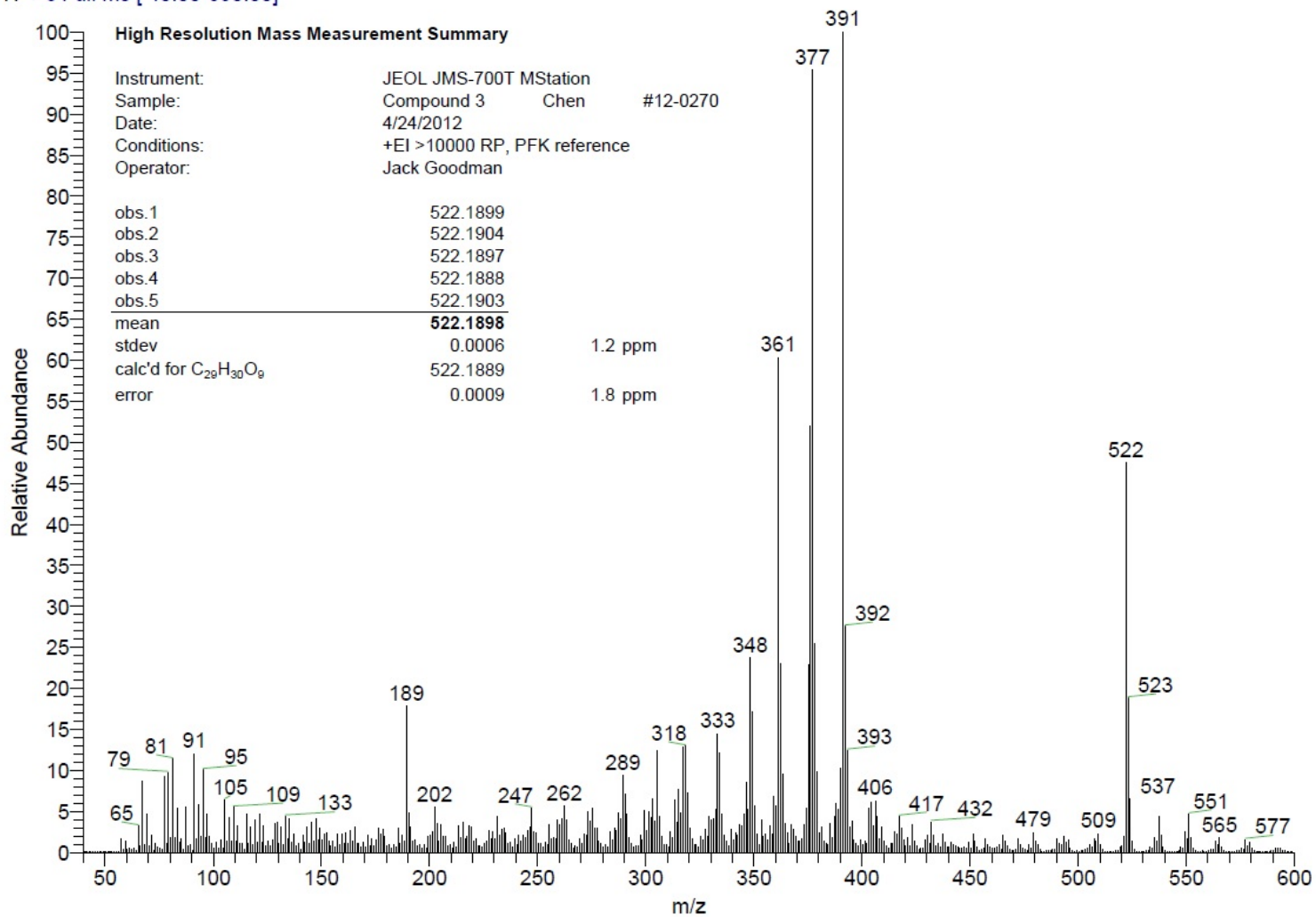


Appendix 9-2. HR-ESI-MS spectrum (negative mode) of 2',3'-di-O-methyl-polycarcin V (**63**)

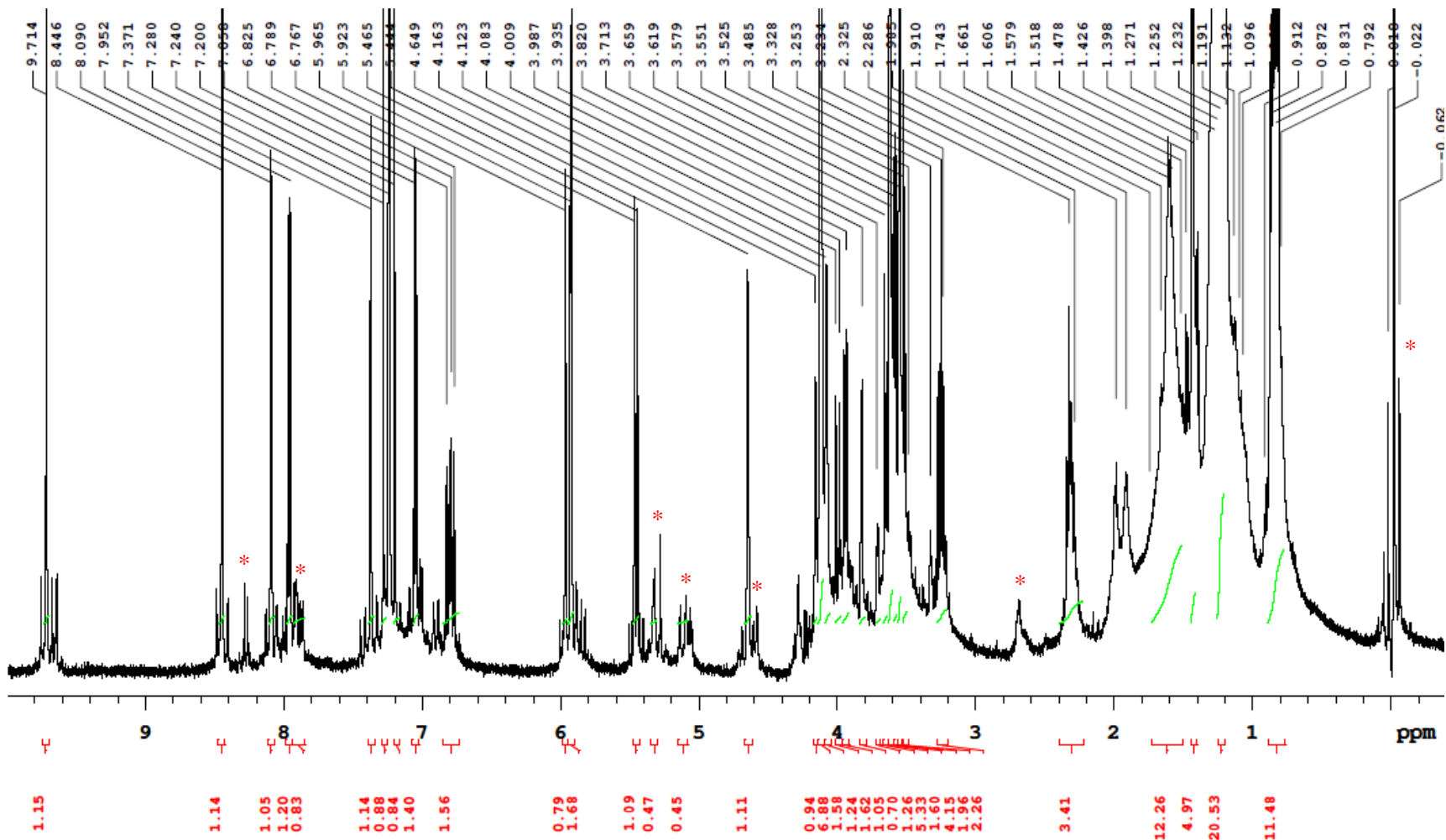


Appendix 10. ^1H NMR spectrum of 2',3'-di-*O*-methyl-polycarin V (63) (*: signals from minor impurities)

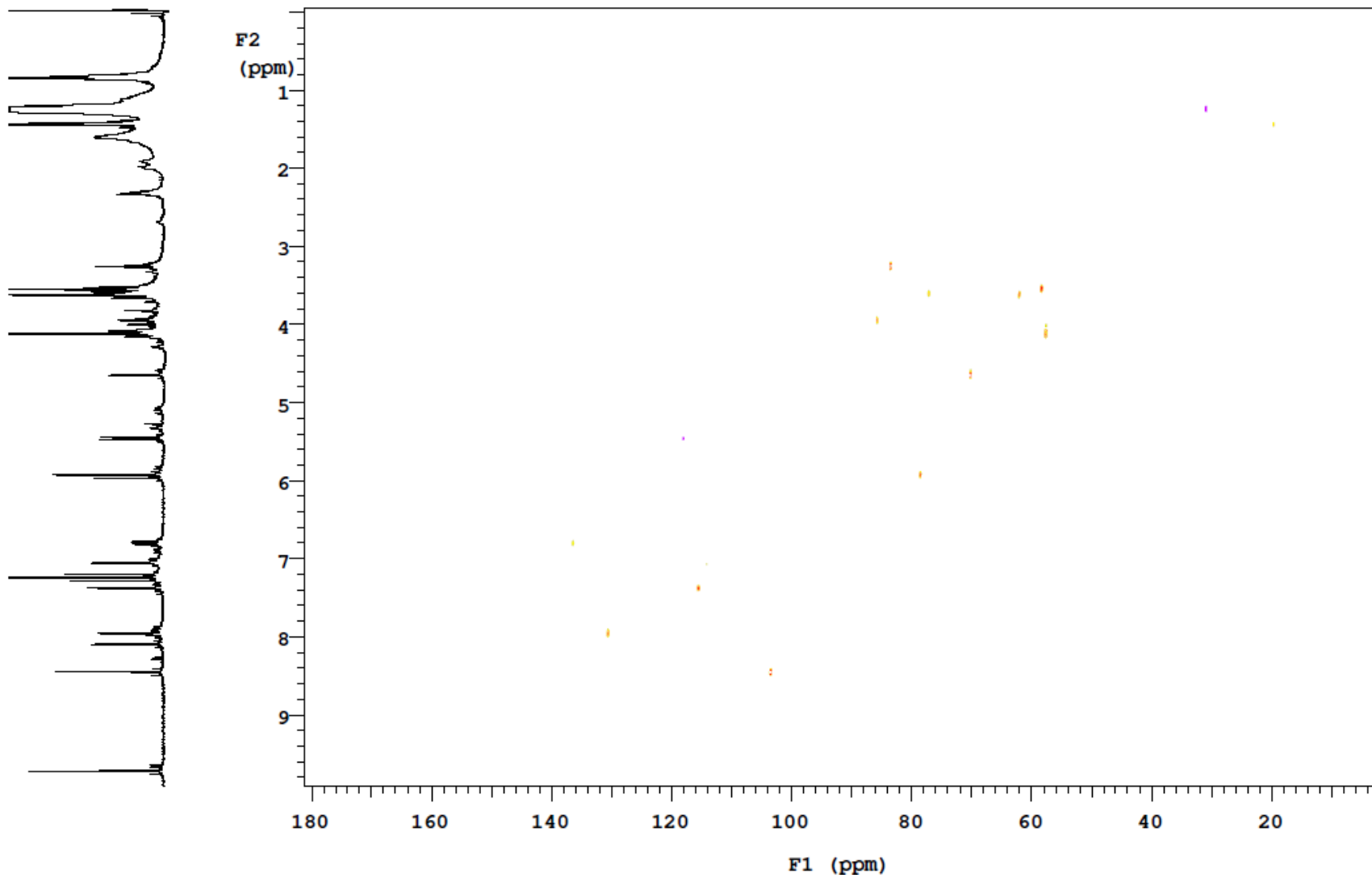
12-0270 #141-242 RT: 2.52-4.41 AV: 102 NL: 5.23E4
T: + c Full ms [40.00-650.00]



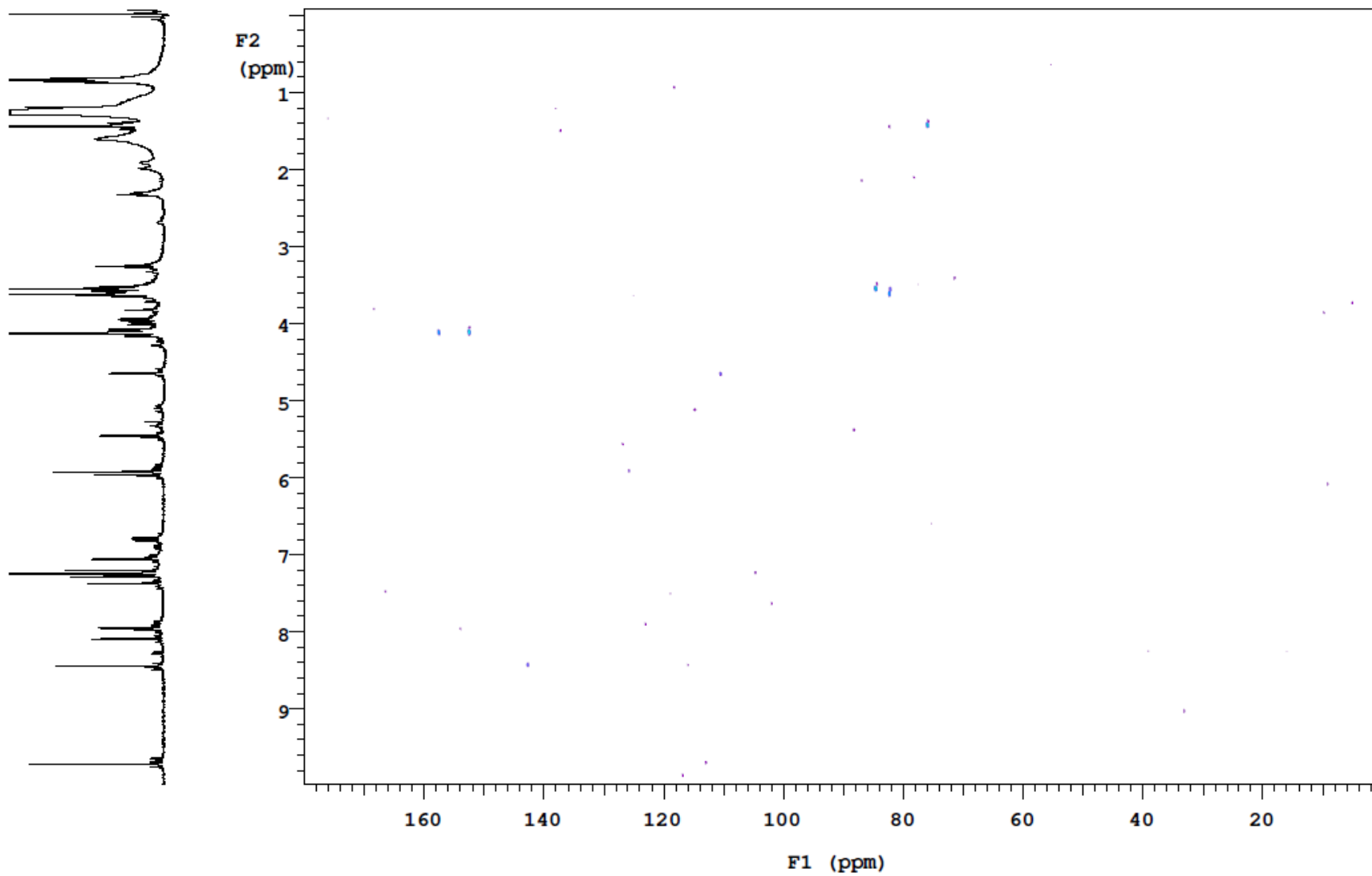
Appendix 11. ESI-MS spectrum (positive mode) and summary of HR-EI-MS data of 3',4'-di-*O*-methyl-polycarcin V (**64**)



Appendix 12. ¹H NMR spectrum of 3',4'-di-O-methyl-polycarcin V (64) (*: signals from minor impurities)

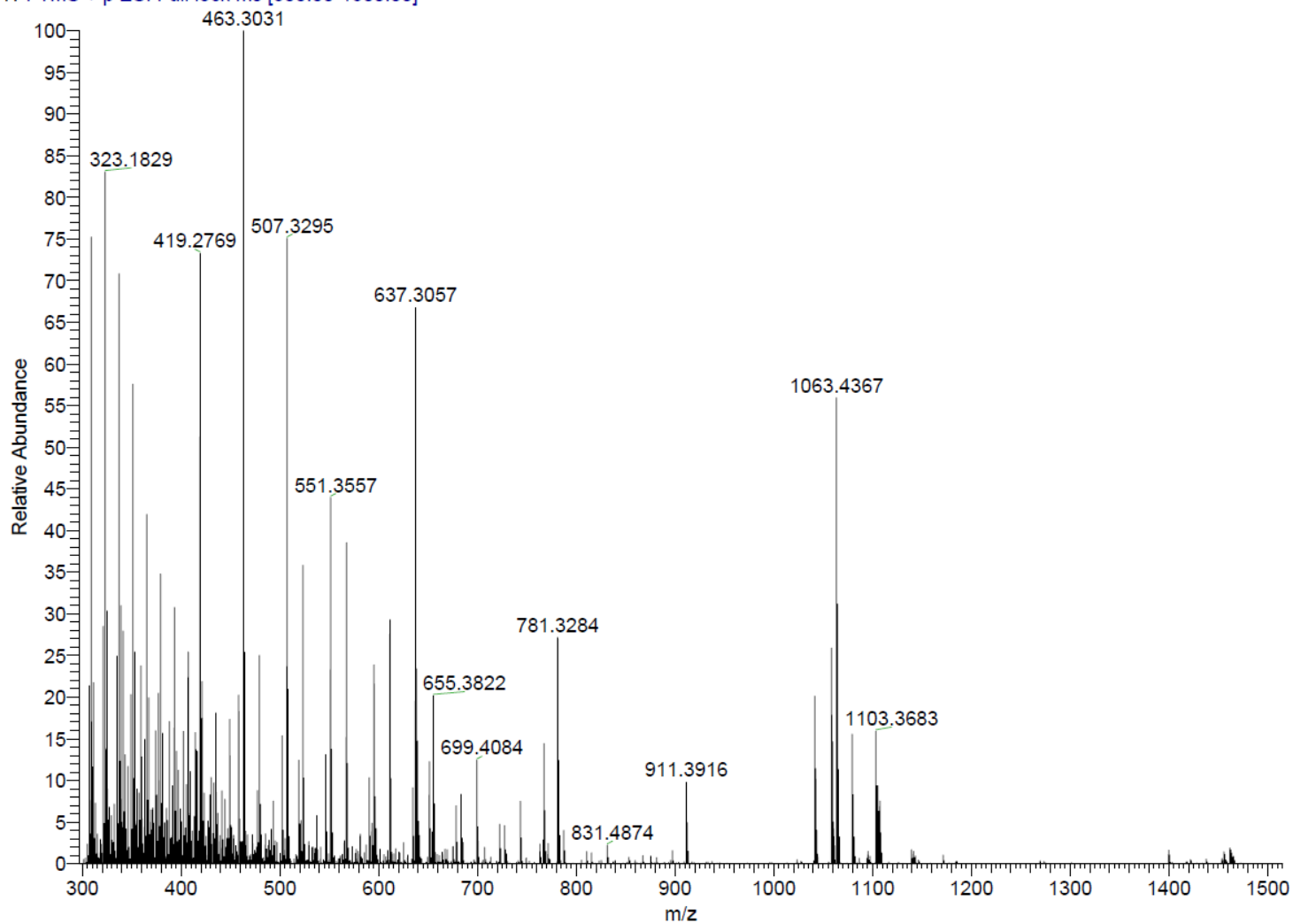


Appendix 13. gHSQC NMR spectrum of 3',4'-di-*O*-methyl-polycarcin V (**64**)

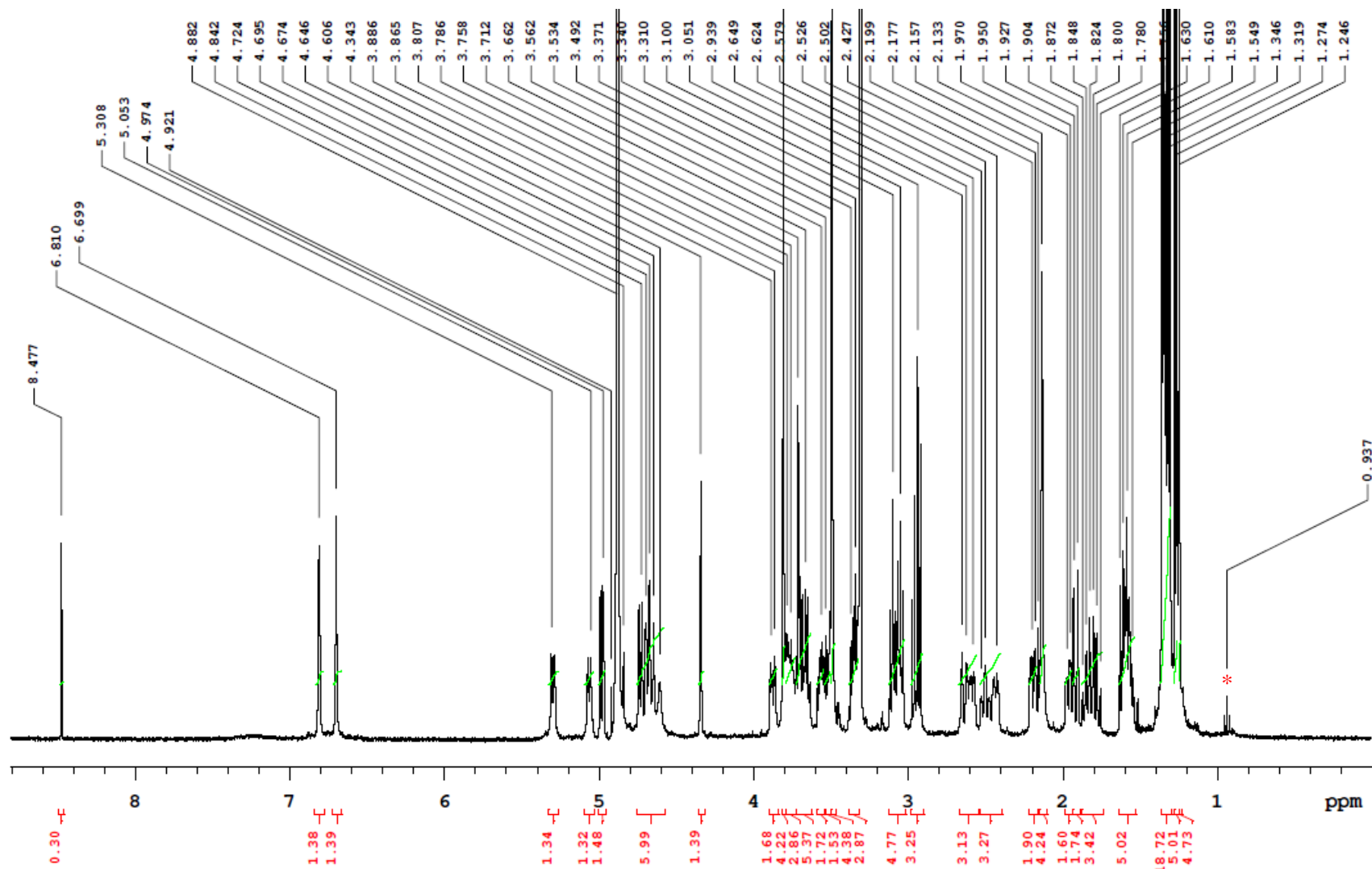


Appendix 14. gHMBC NMR spectrum of 3',4'-di-*O*-methyl-polycarcin V (**64**)

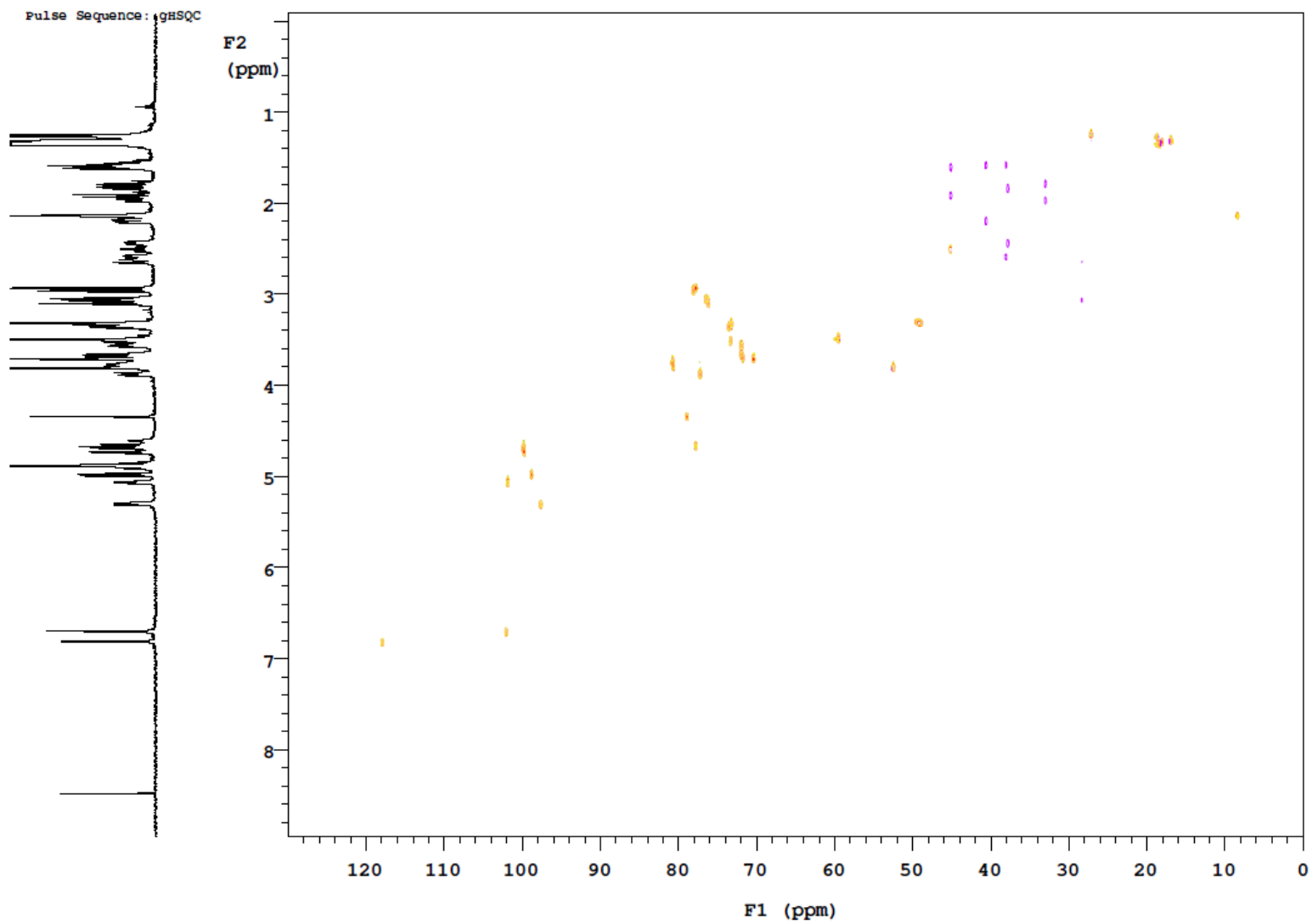
14-0547 #12-28 RT: 0.31-0.73 AV: 17 NL: 6.53E5
T: FTMS + p ESI Full lock ms [300.00-1500.00]



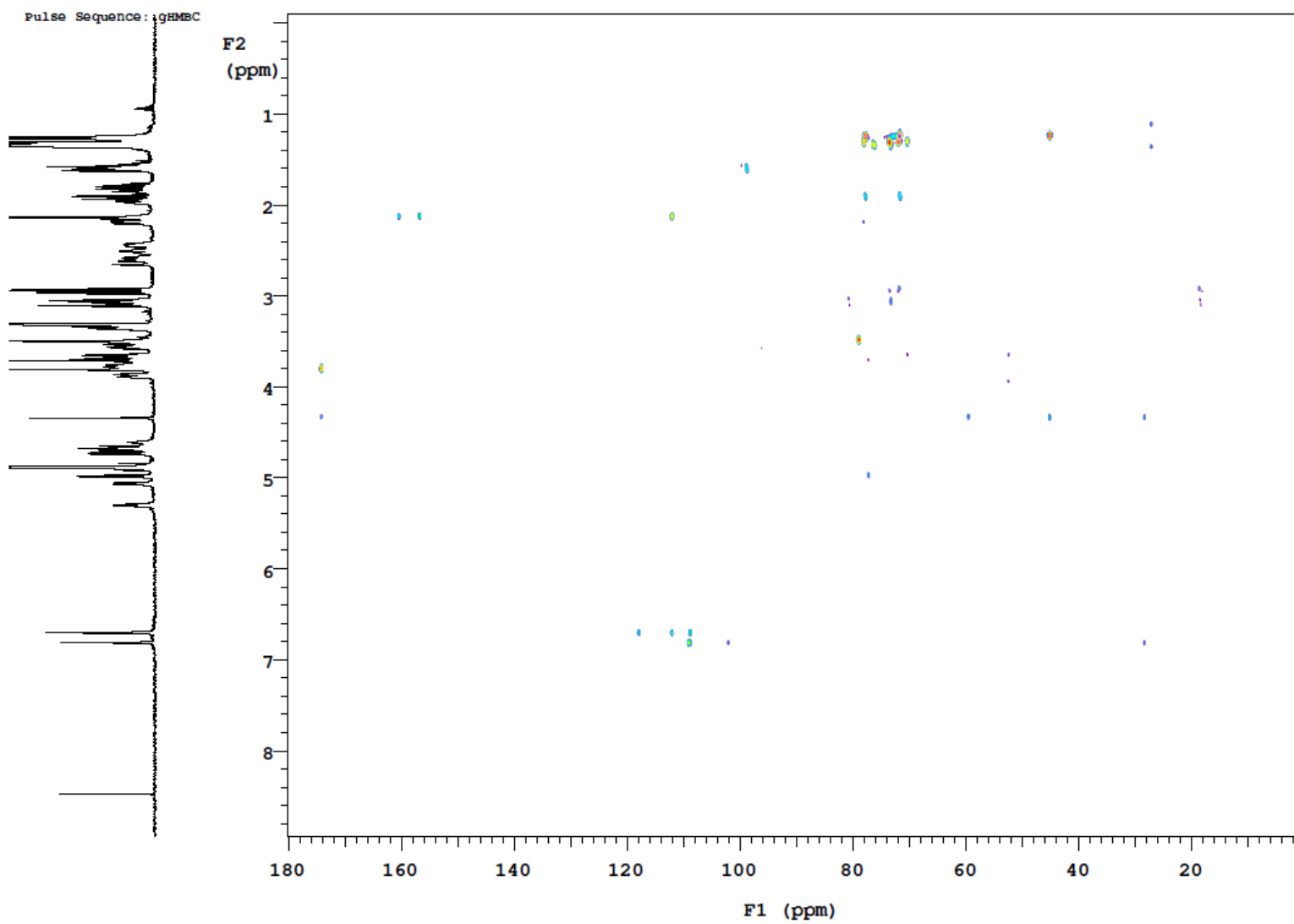
Appendix 15. HR-ESI-MS spectrum of MTM SA methyl ester (72)



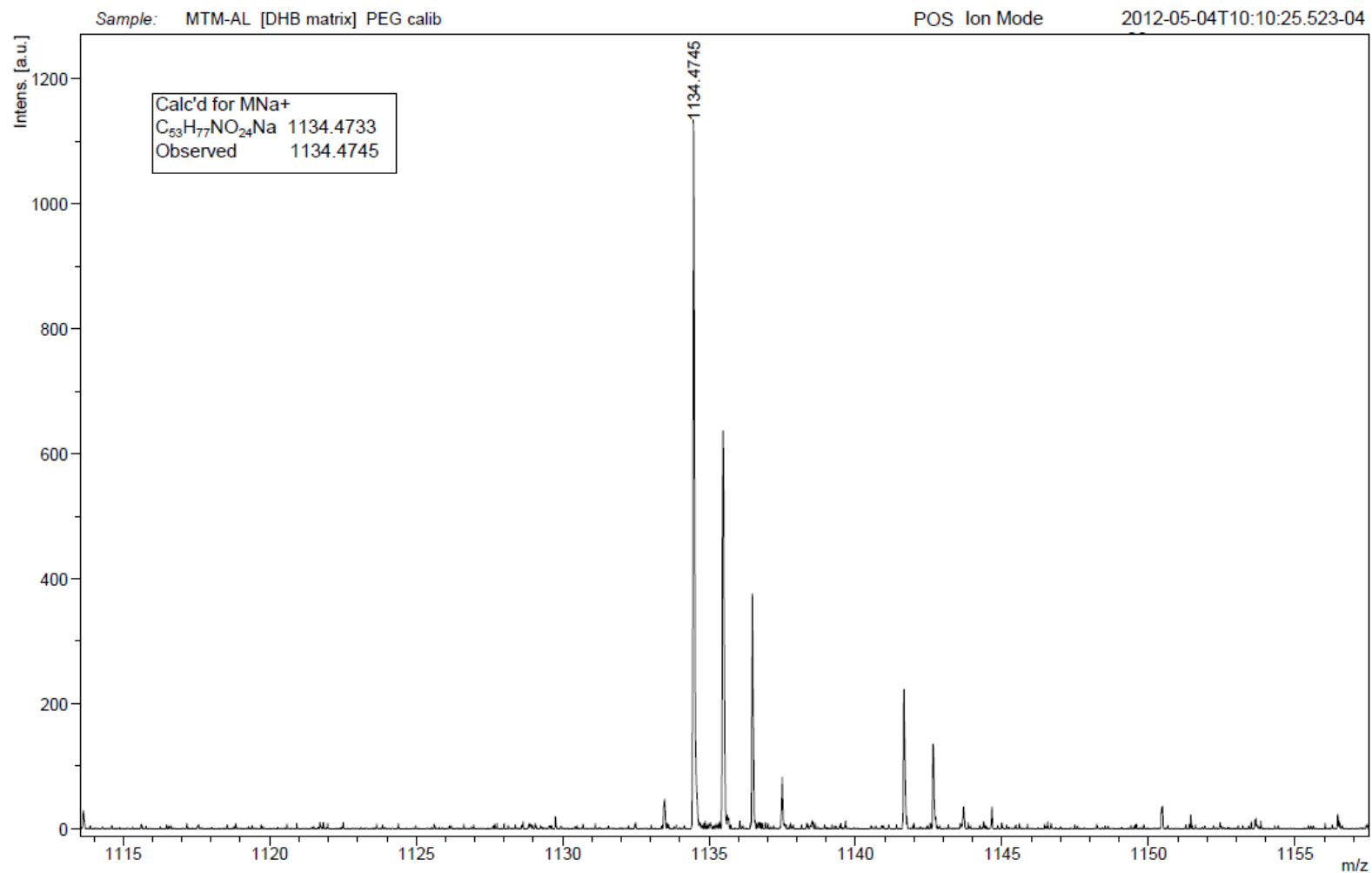
Appendix 16. ¹H NMR spectrum of MTM SA methyl ester (72) (*: signals from minor impurities)



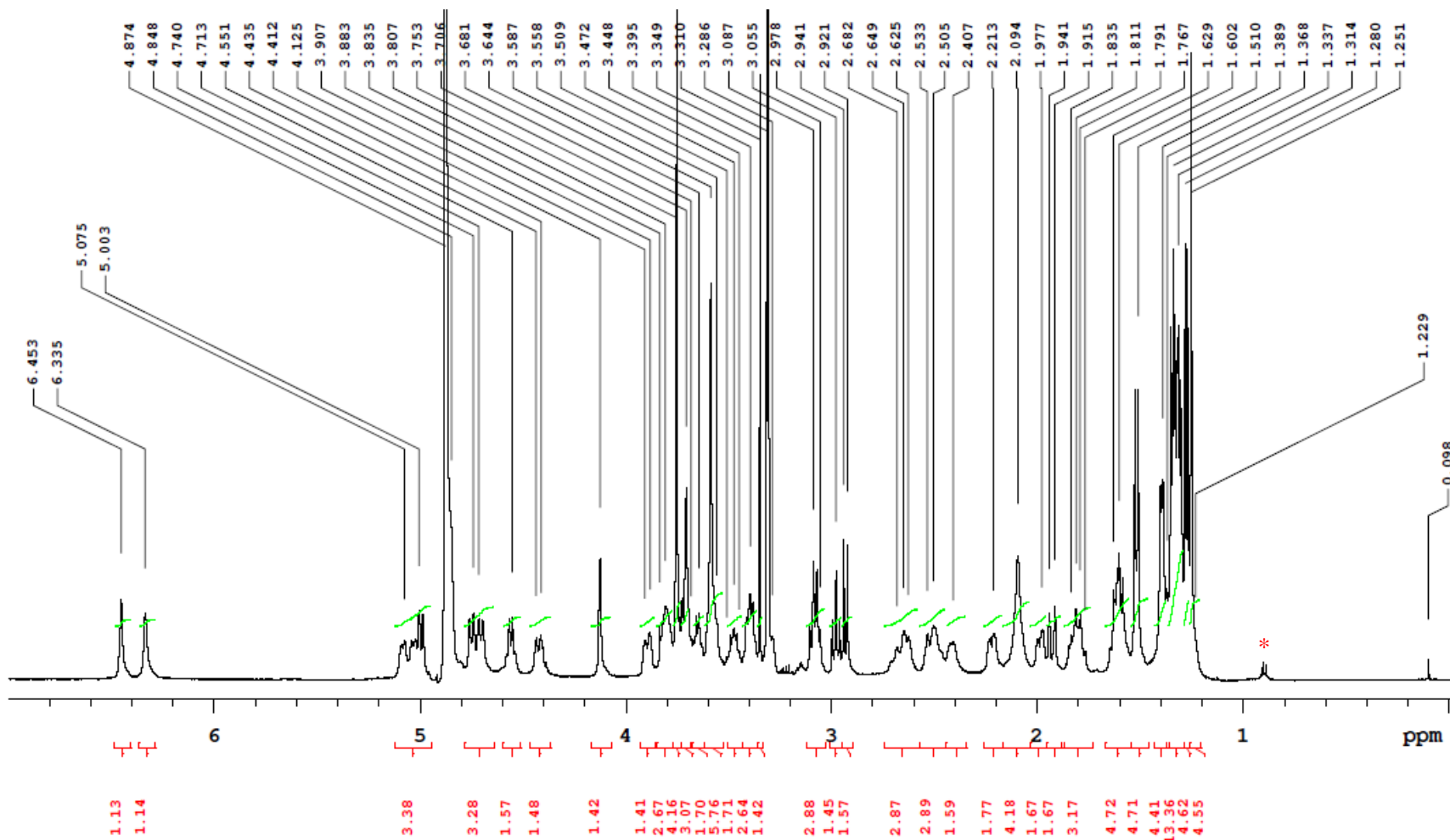
Appendix 17. gHSQC NMR spectrum of MTM SA methyl ester (72)



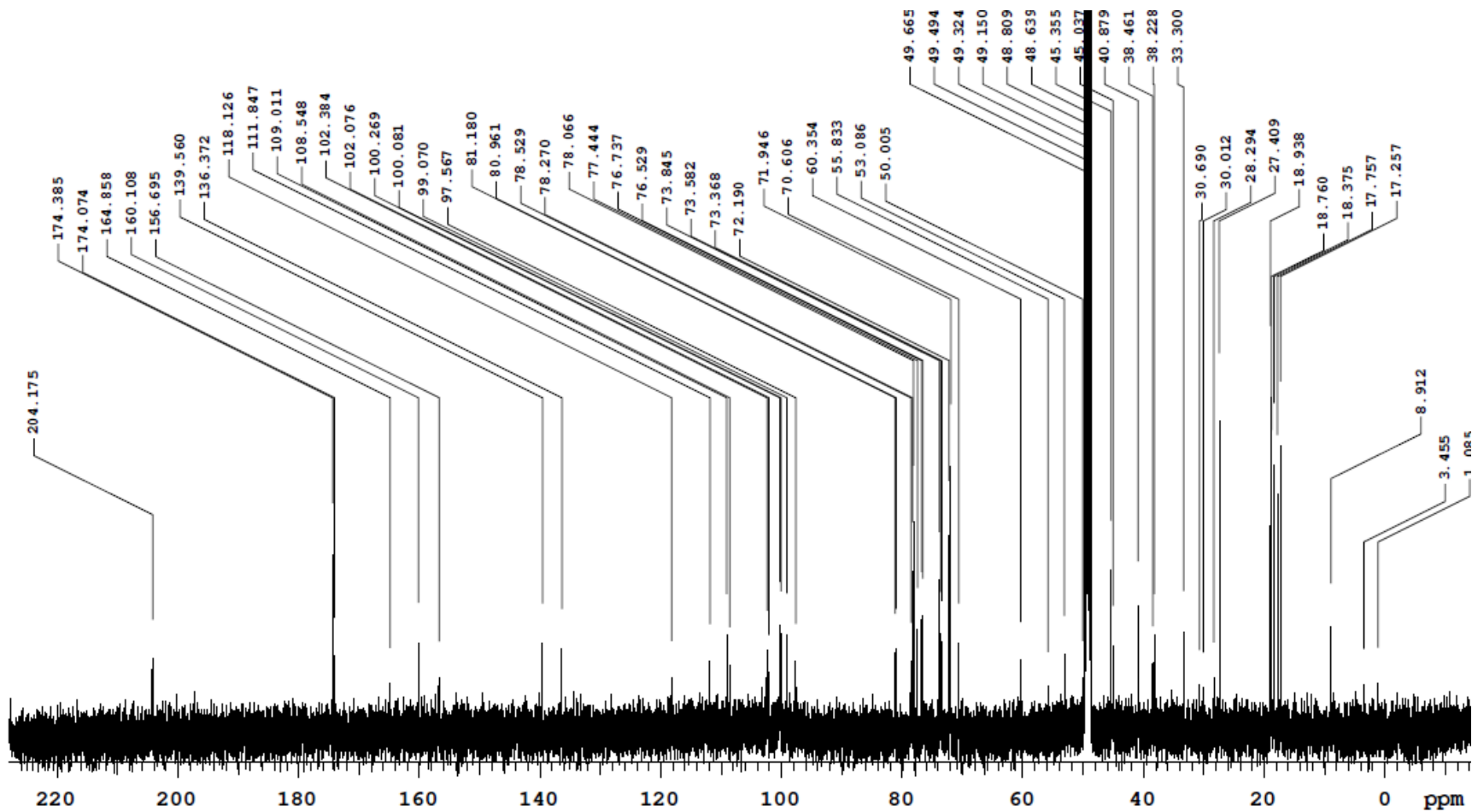
Appendix 18. gHMBC NMR spectrum of MTM SA methyl ester (**72**)



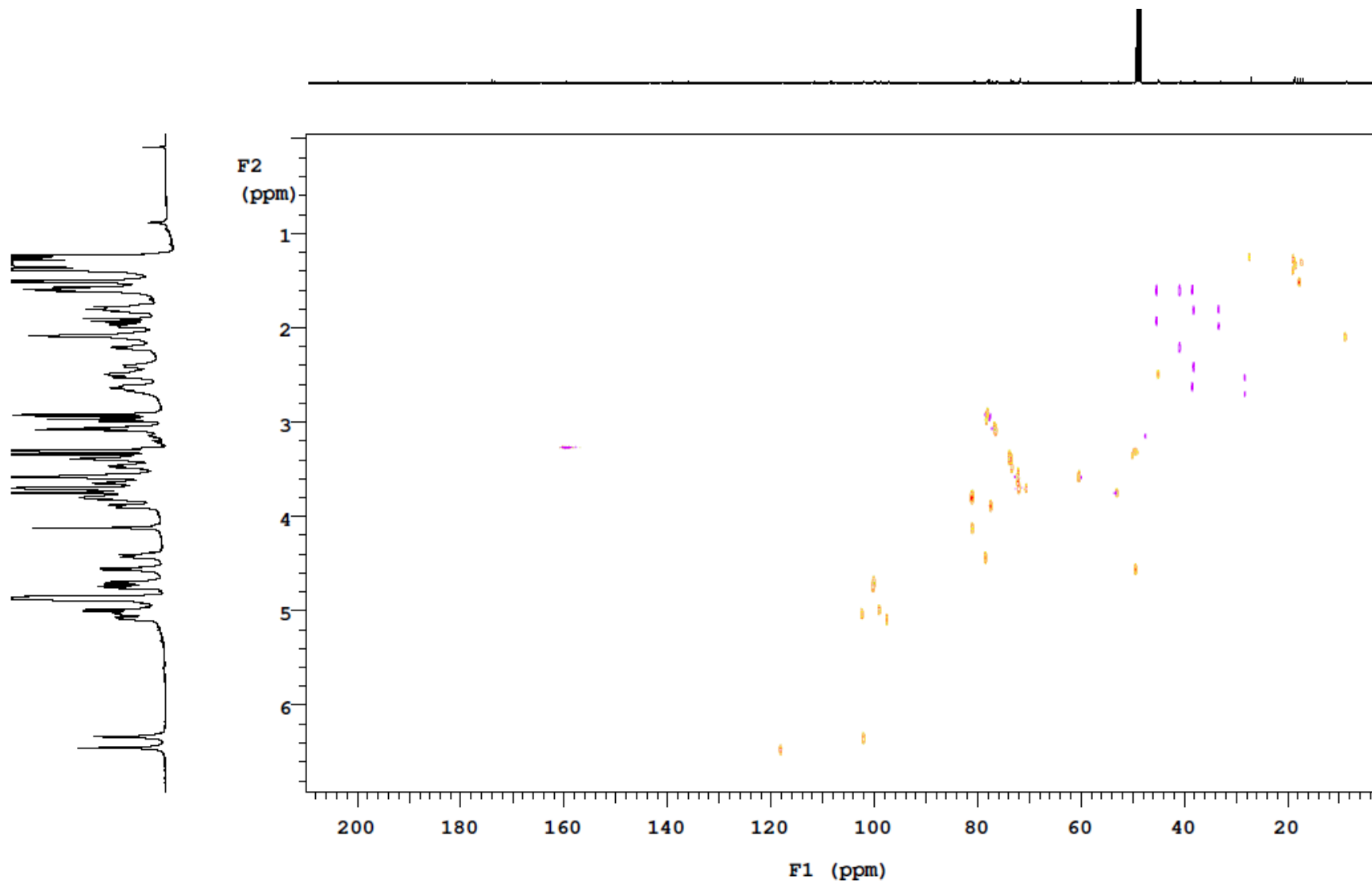
Appendix 19. HR-MALDI-TOF-MS spectrum of MTM SA-L-alanine methyl ester (78)



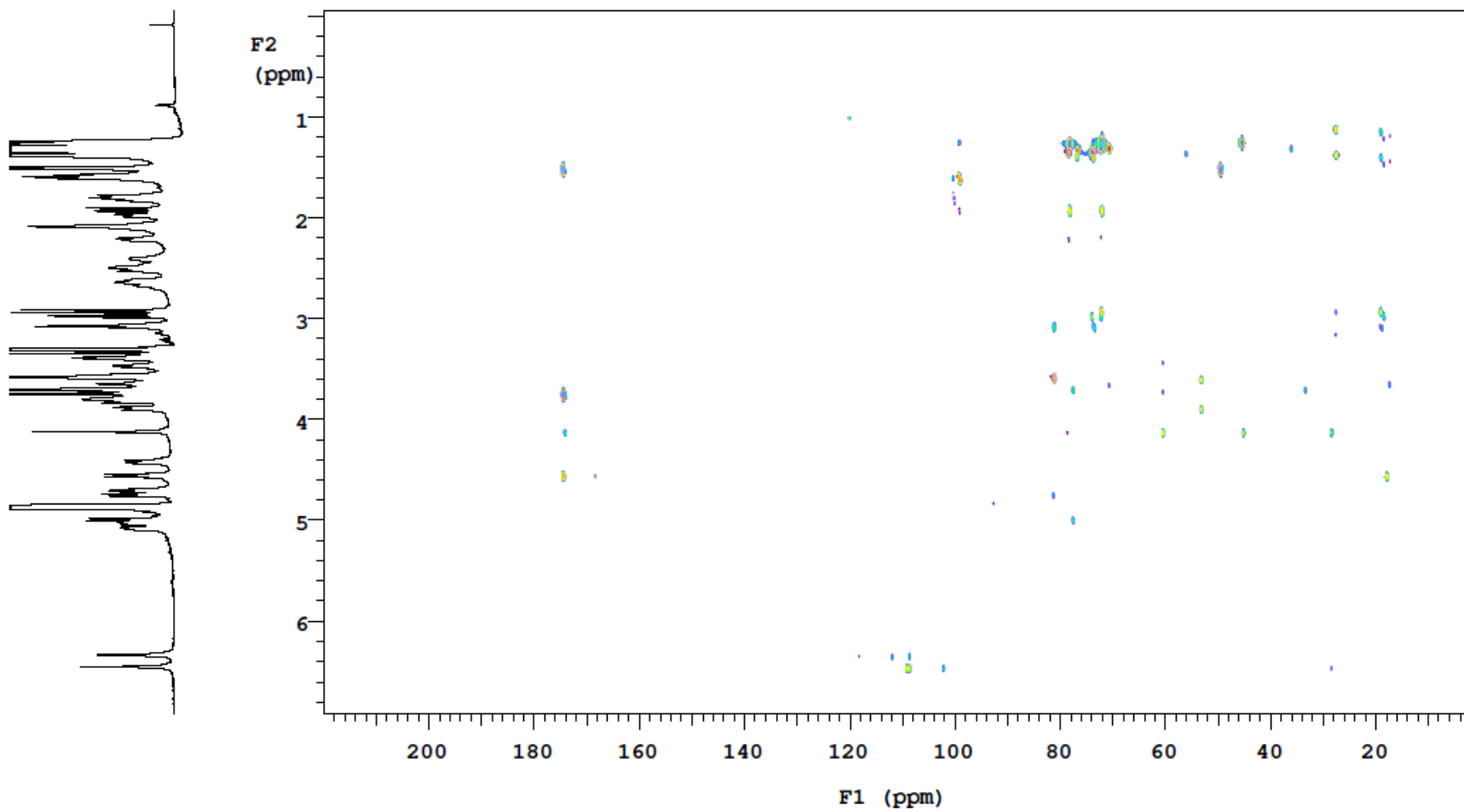
Appendix 20. ^1H NMR spectrum of MTM SA-L-alanine methyl ester (**78**) (*: signals from minor impurities)



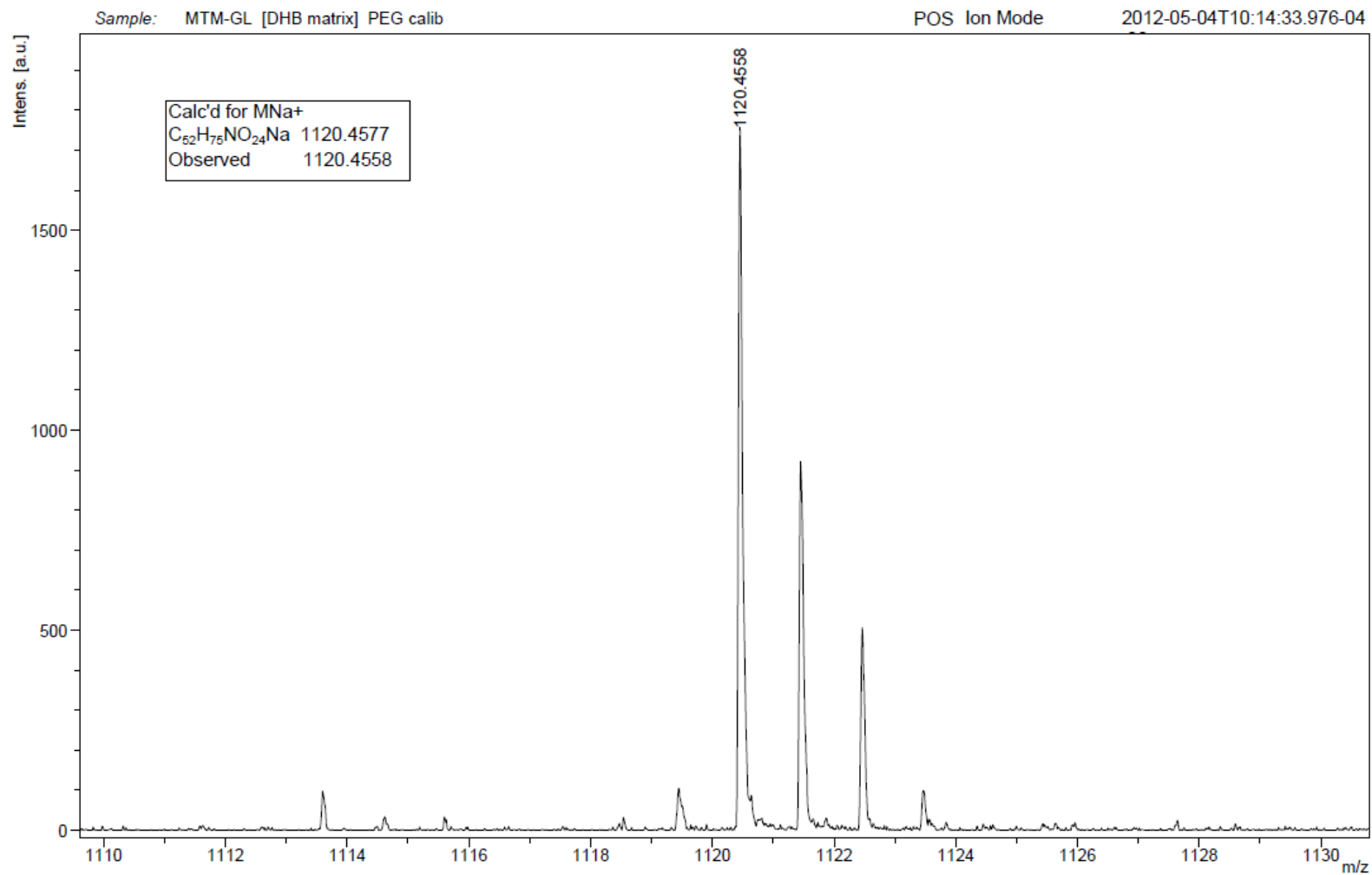
Appendix 21. ^{13}C NMR spectrum of MTM SA-L-alanine methyl ester (78)

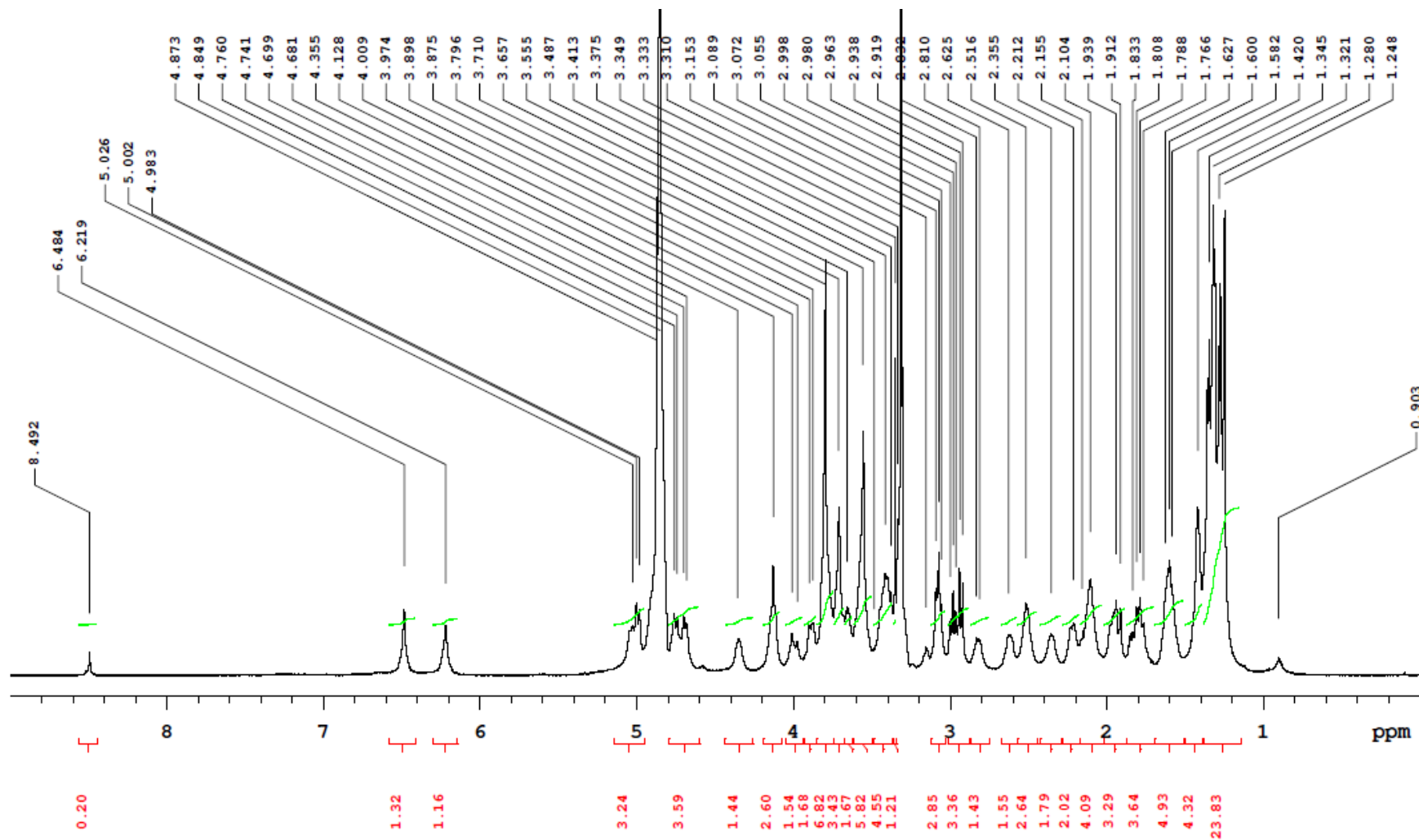


Appendix 22. gHSQC NMR spectrum of MTM SA-L-alanine methyl ester (**78**)

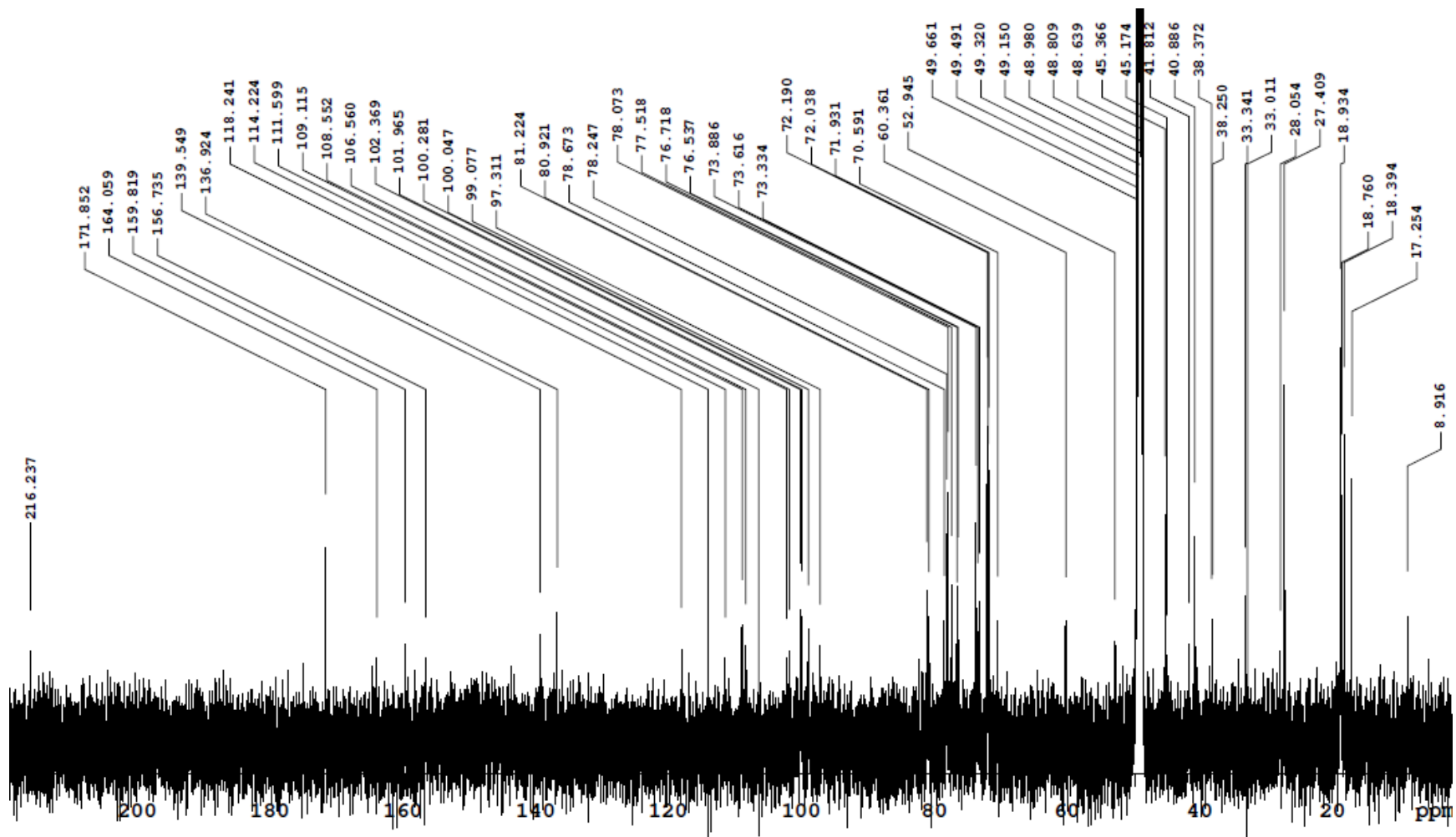


Appendix 23. gHMBC NMR spectrum of MTM SA-L-alanine methyl ester (78)

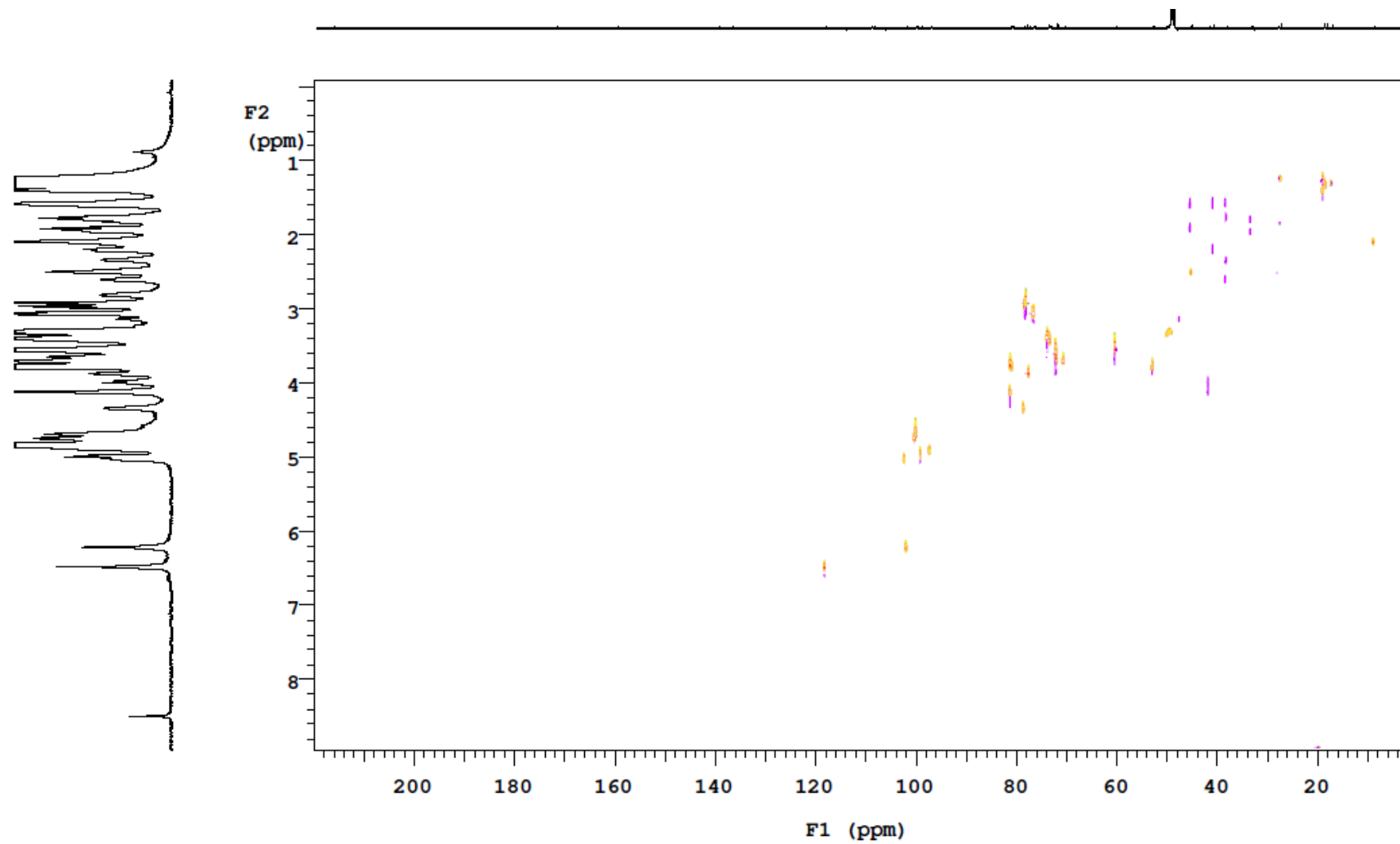
**Appendix 24.** HR-MALDI-TOF-MS spectrum of MTM SA-L-glycine methyl ester (**80**)



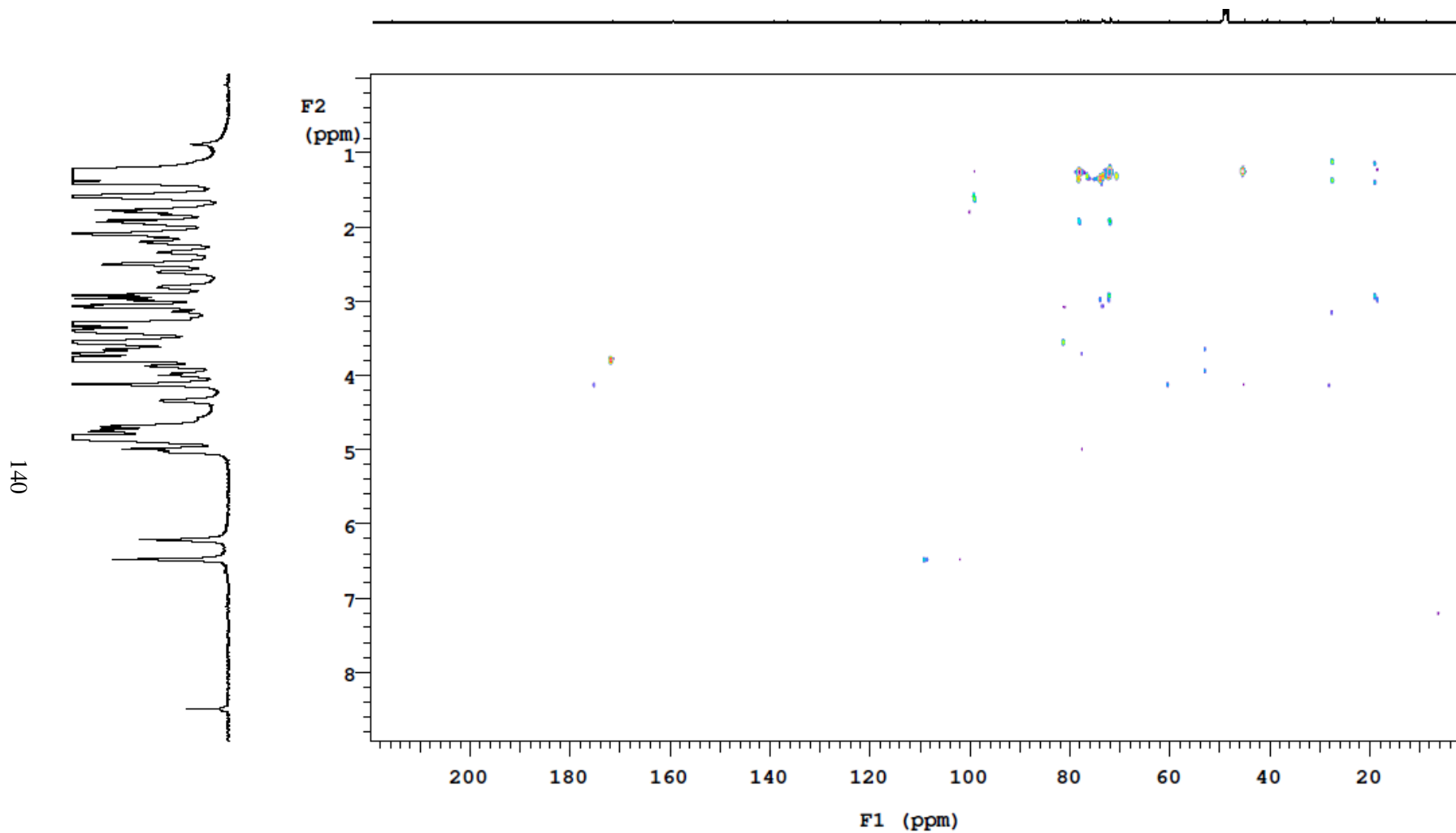
Appendix 25. ^1H NMR spectrum of MTM SA-L-glycine methyl ester (80)



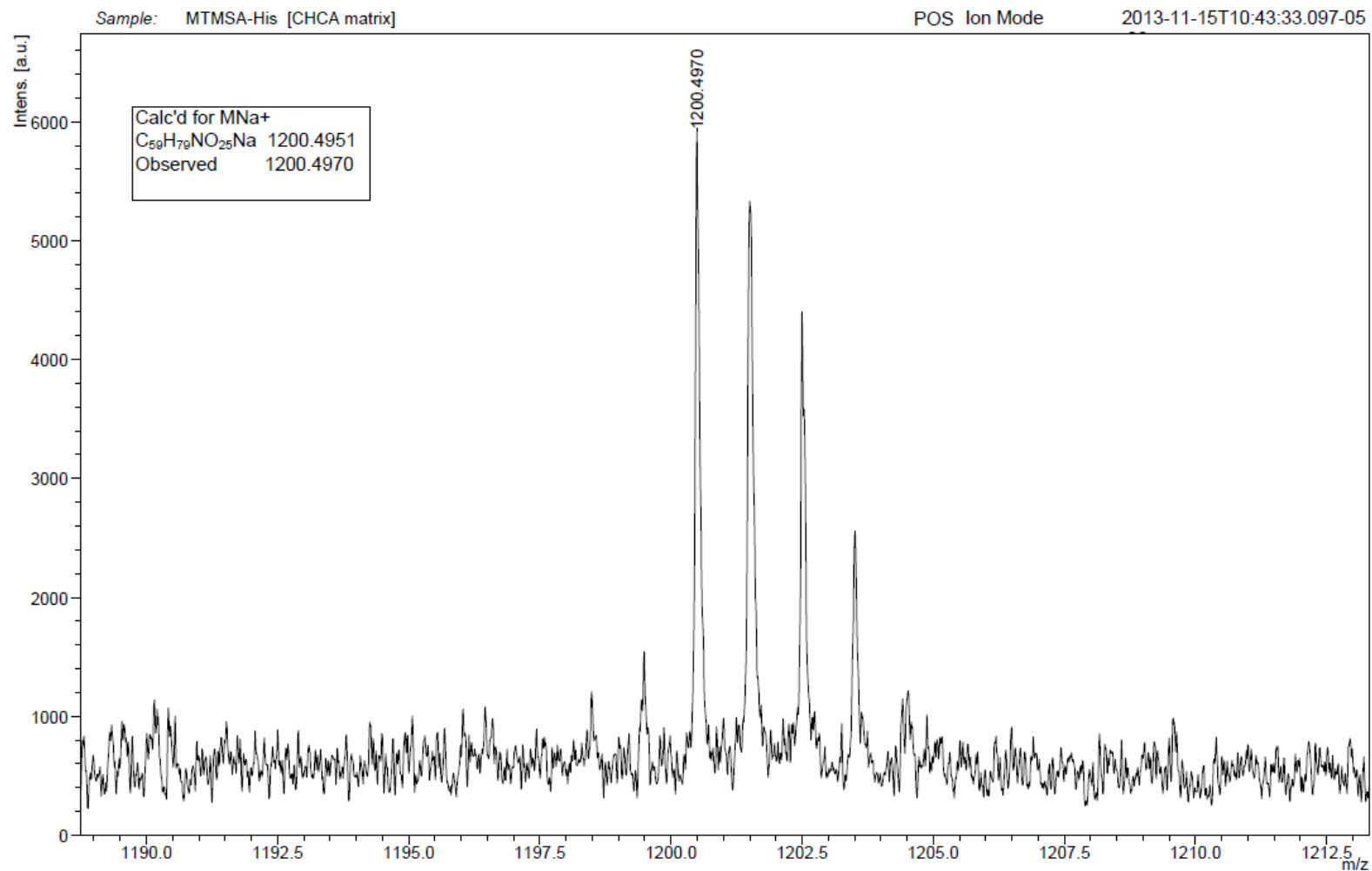
Appendix 26. ^{13}C NMR spectrum of MTM SA-L-glycine methyl ester (80)



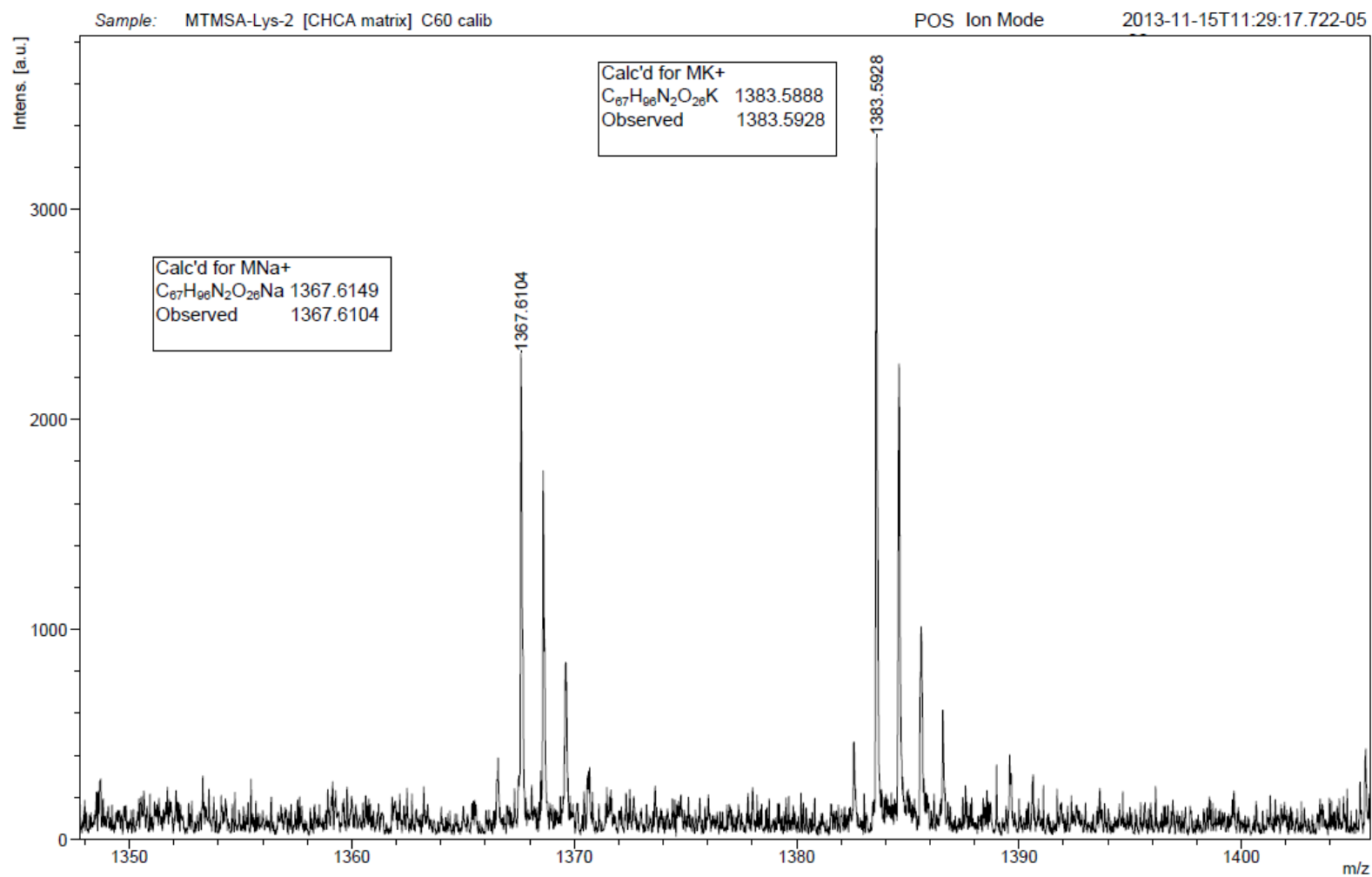
Appendix 27. gHSQC NMR spectrum of MTM SA-L-glycine methyl ester (**80**)



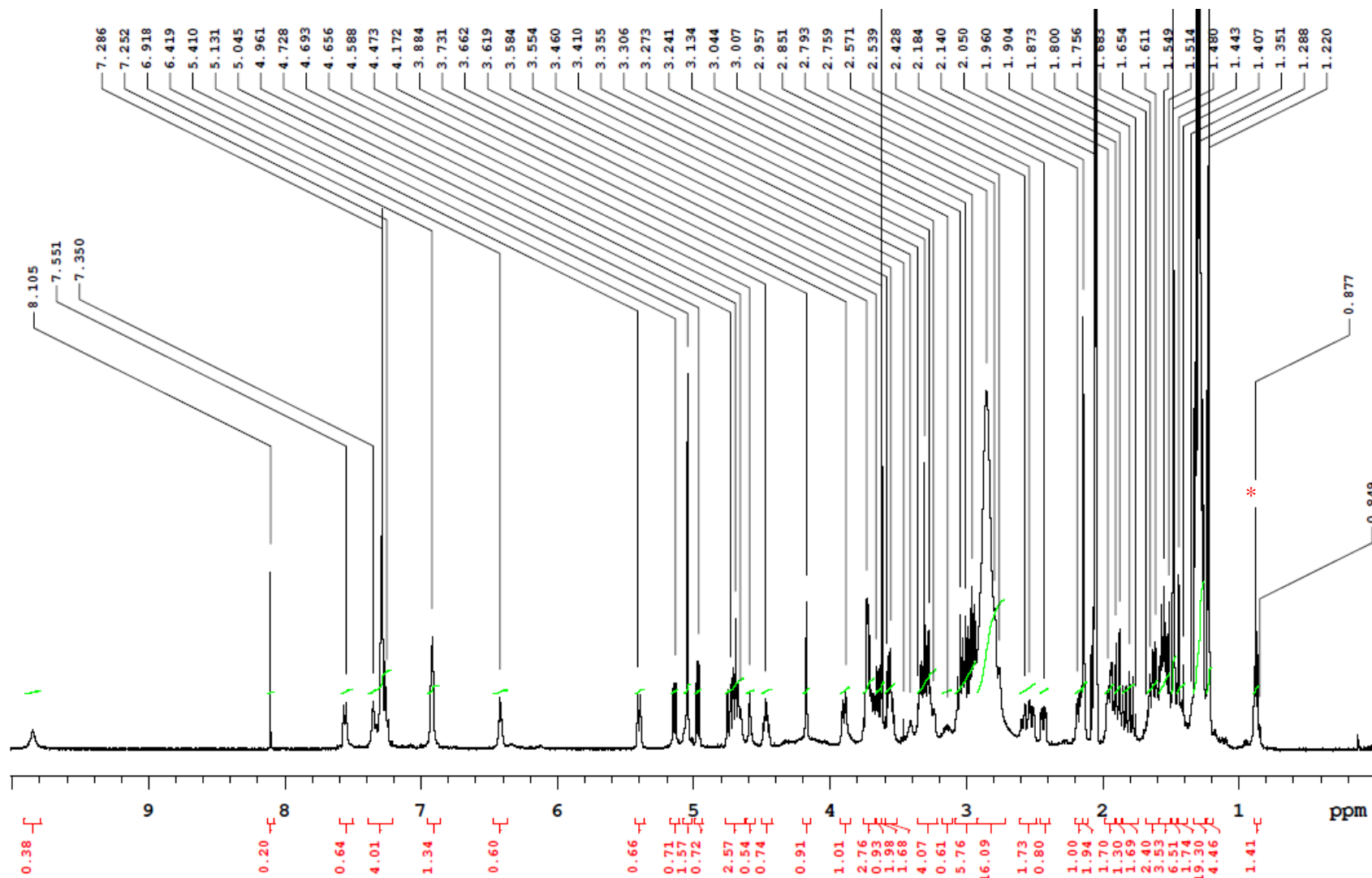
Appendix 28. gHMBC NMR spectrum of MTM SA-L-glycine methyl ester (**80**)



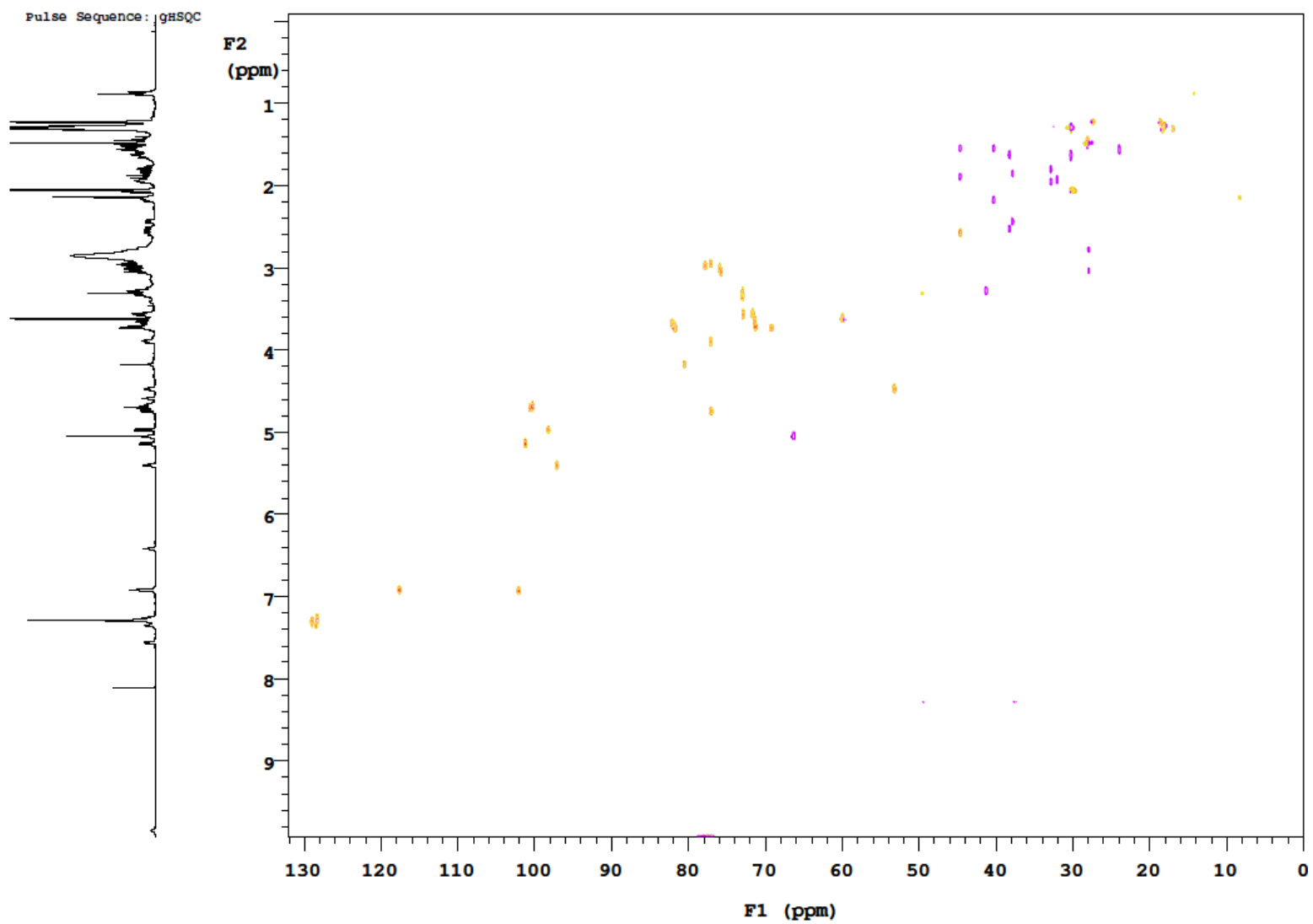
Appendix 29. HR-MALDI-TOF-MS spectrum of MTM SA-L-histidine methyl ester (**81**)



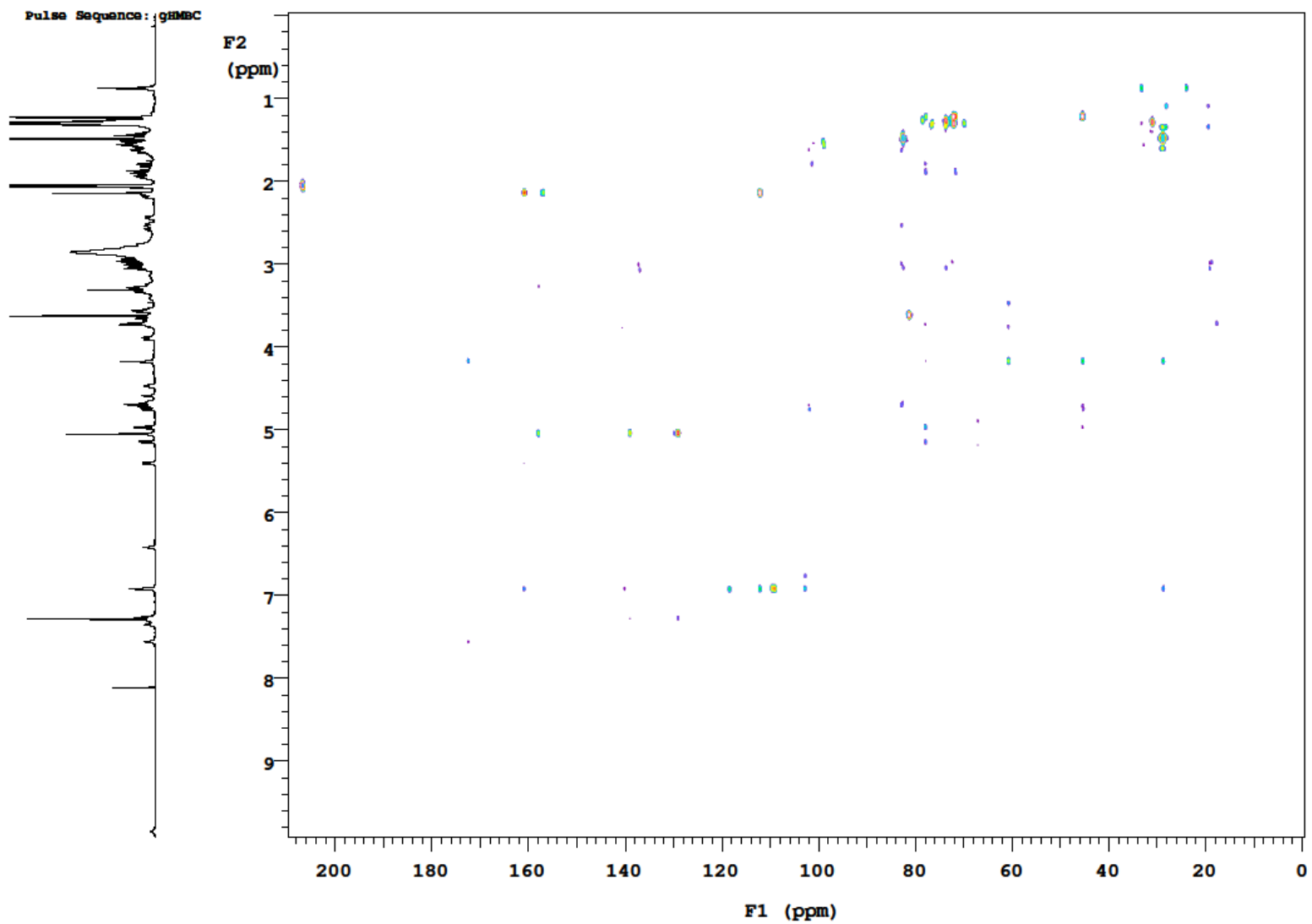
Appendix 30. HR-MALDI-TOF-MS spectrum of MTM SA-L-lysine t-Bu ester *N*-benzyl carbamate (**82**)



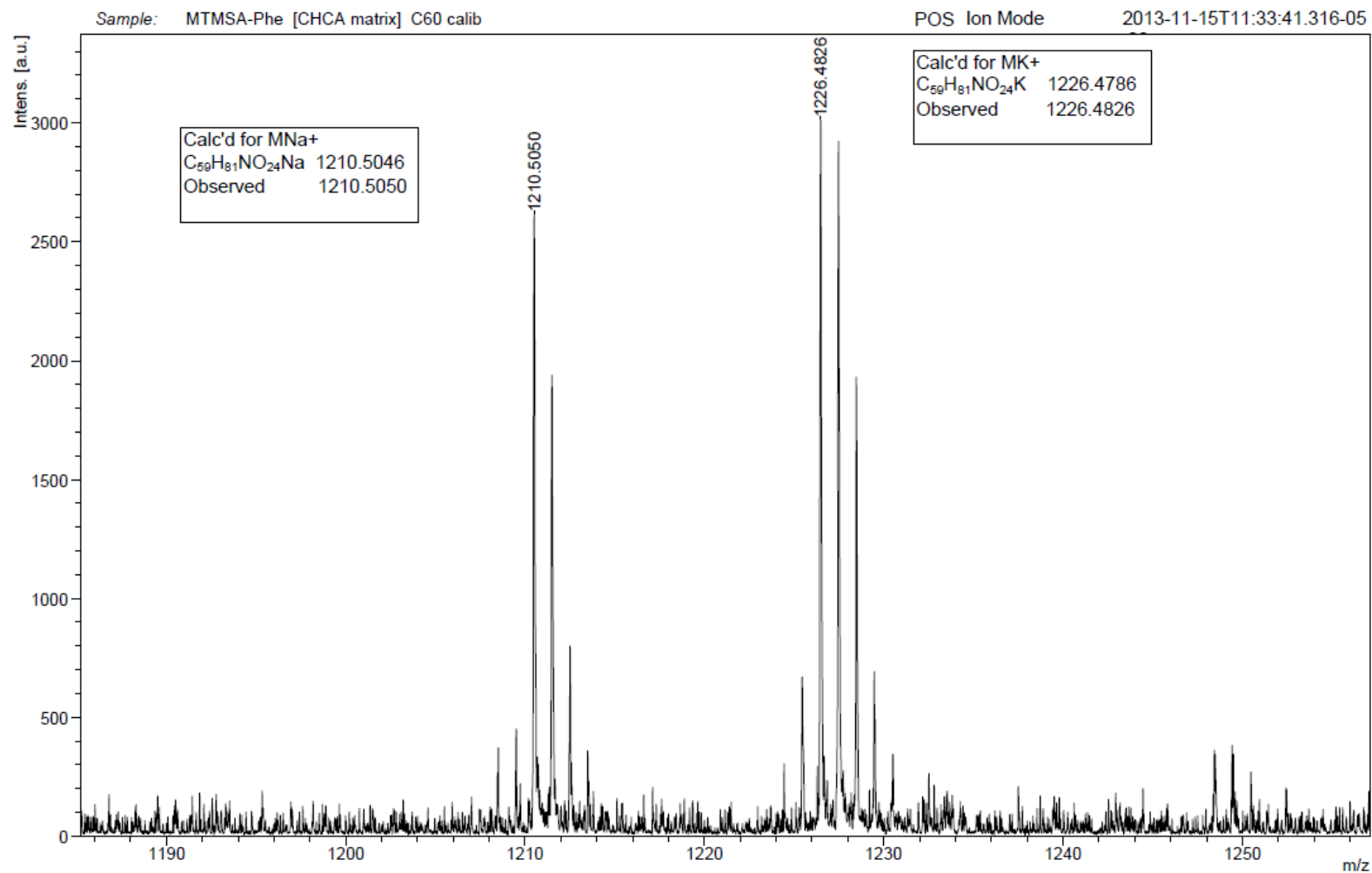
Appendix 31. ^1H NMR spectrum of MTM SA-L-lysine t-Bu ester *N*-benzyl carbamate (**82**) (*: signals from minor impurities)



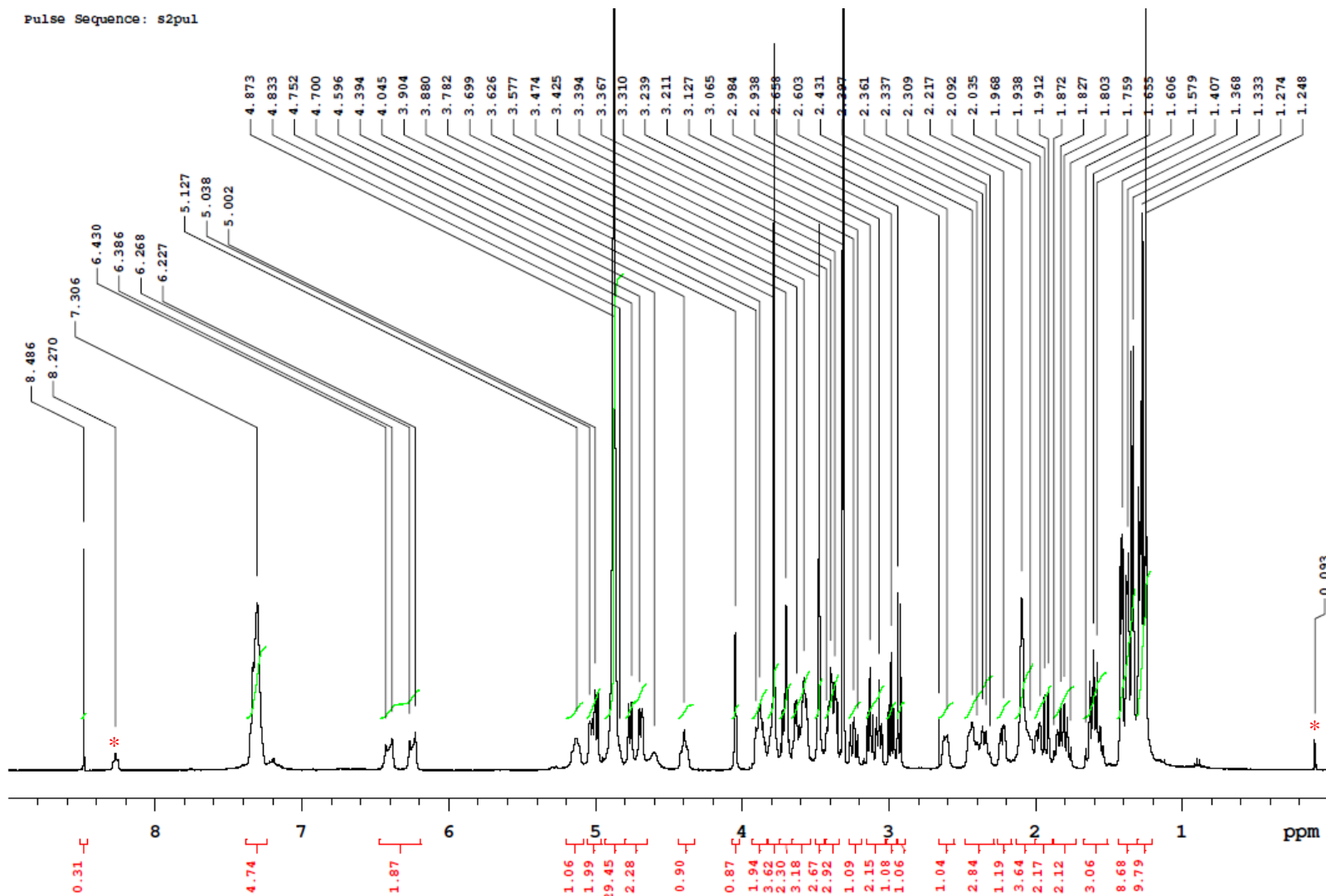
Appendix 32. gHSQC NMR spectrum of MTM SA-L-lysine t-Bu ester *N*-benzyl carbamate (**82**)



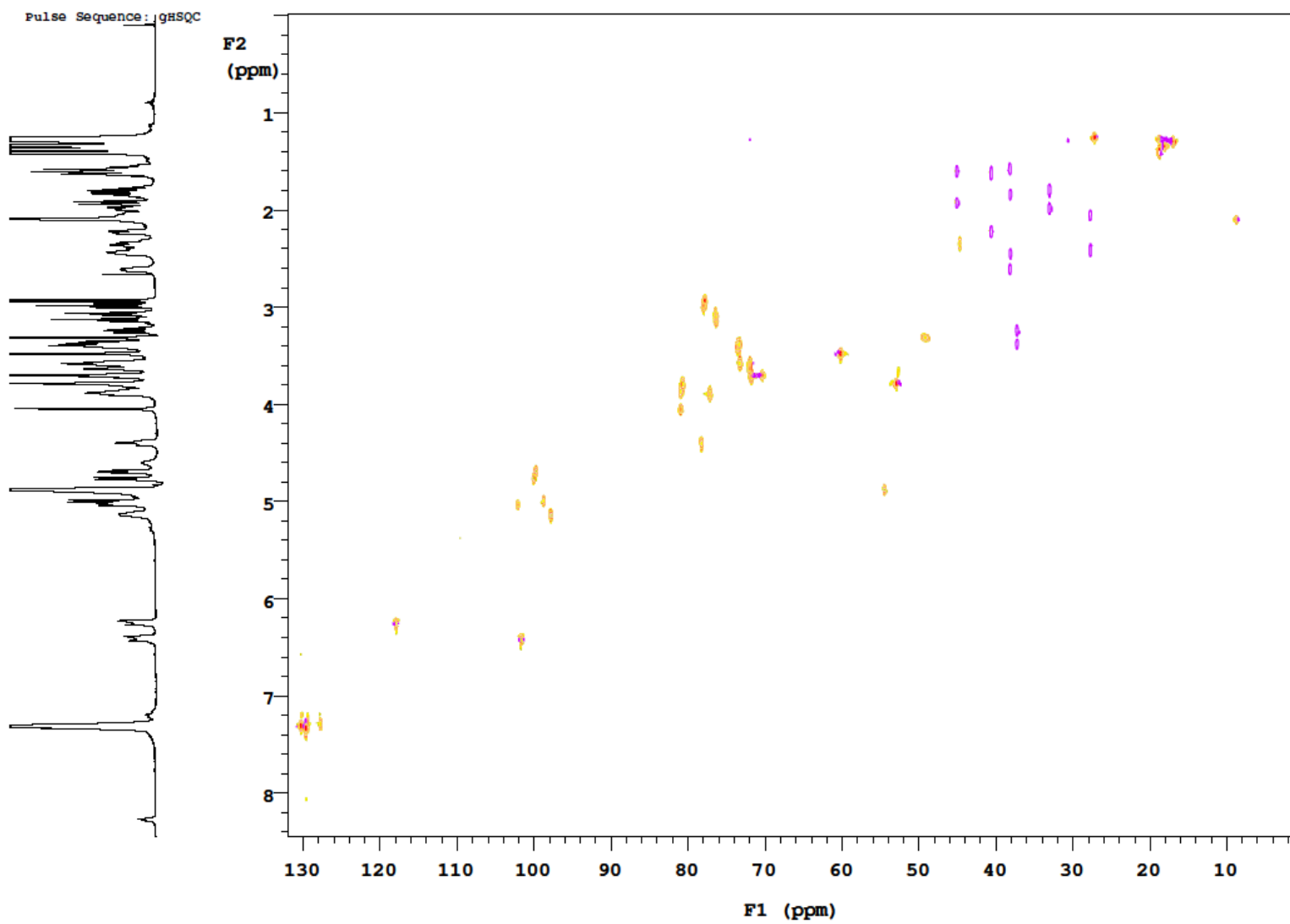
Appendix 33. gHMBC NMR spectrum of MTM SA-L-lysine t-Bu ester *N*-benzyl carbamate (**82**)



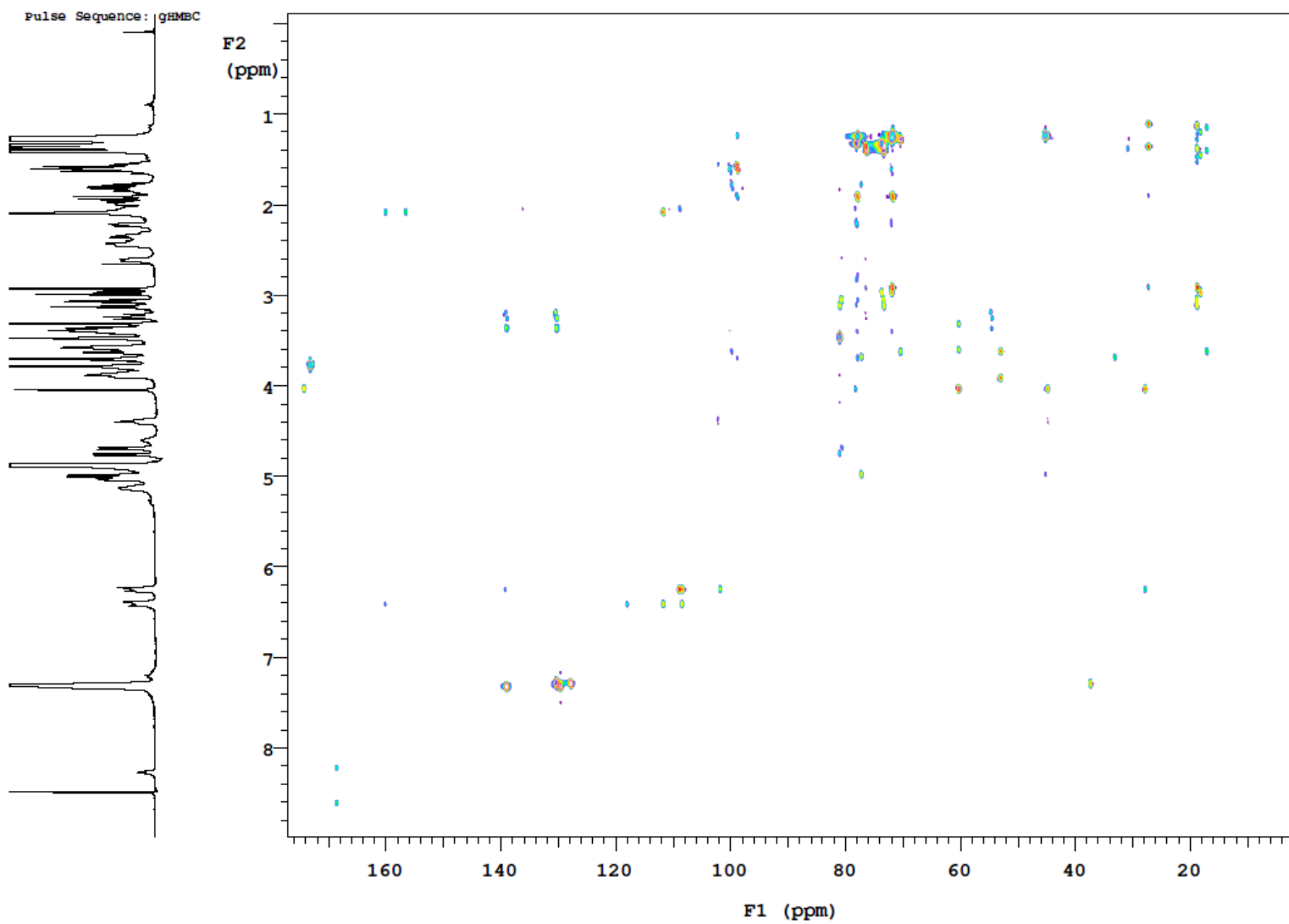
Appendix 34. HR-MALDI-TOF-MS spectrum of MTM SA-L-phenylalanine methyl ester (**83**)



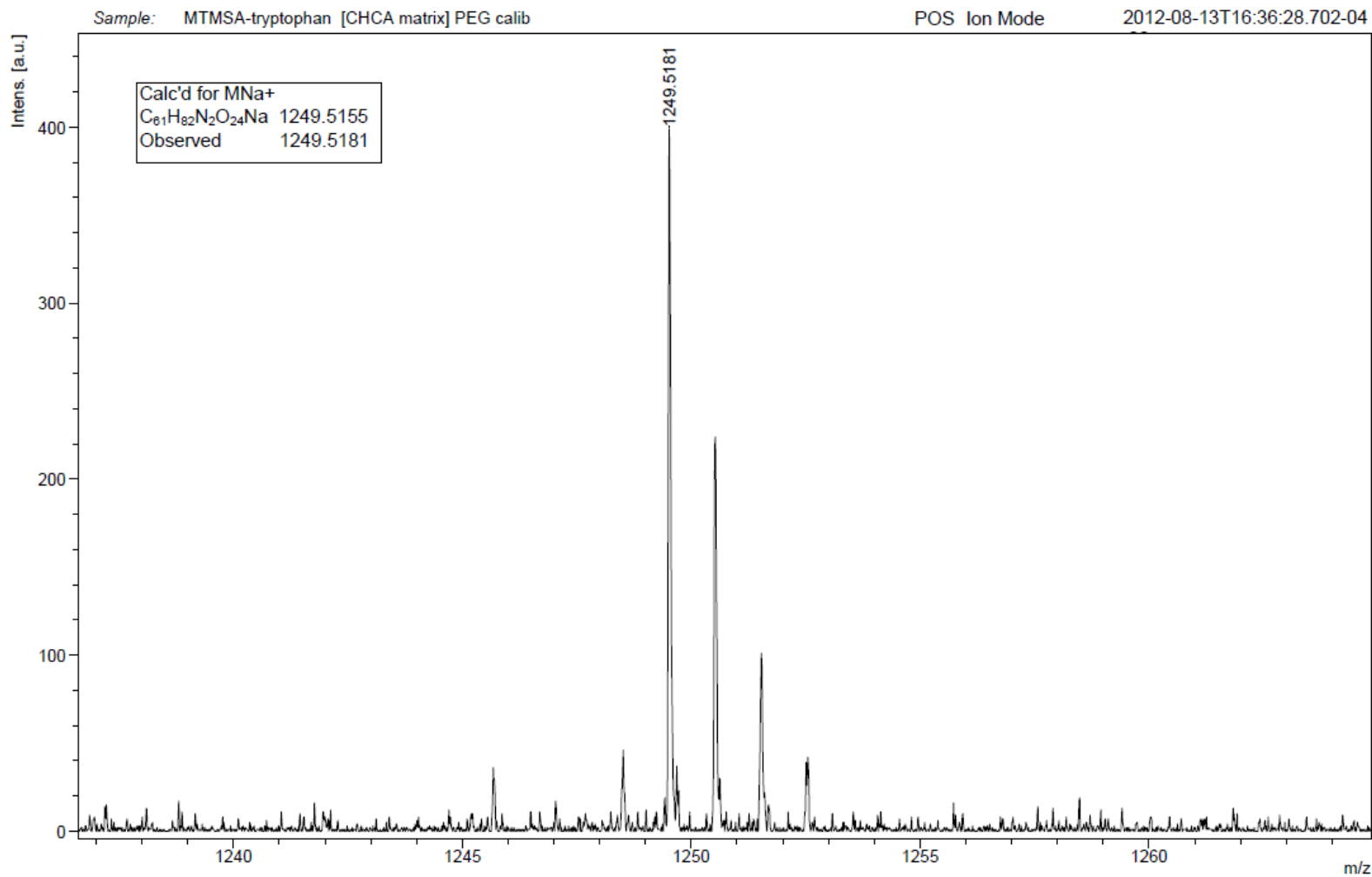
Appendix 35. ^1H NMR spectrum of MTM SA-L-phenylalanine methyl ester (**83**) (*: signals from minor impurities)



Appendix 36. gHSQC NMR spectrum of MTM SA-L-phenylalanine methyl ester (**83**)

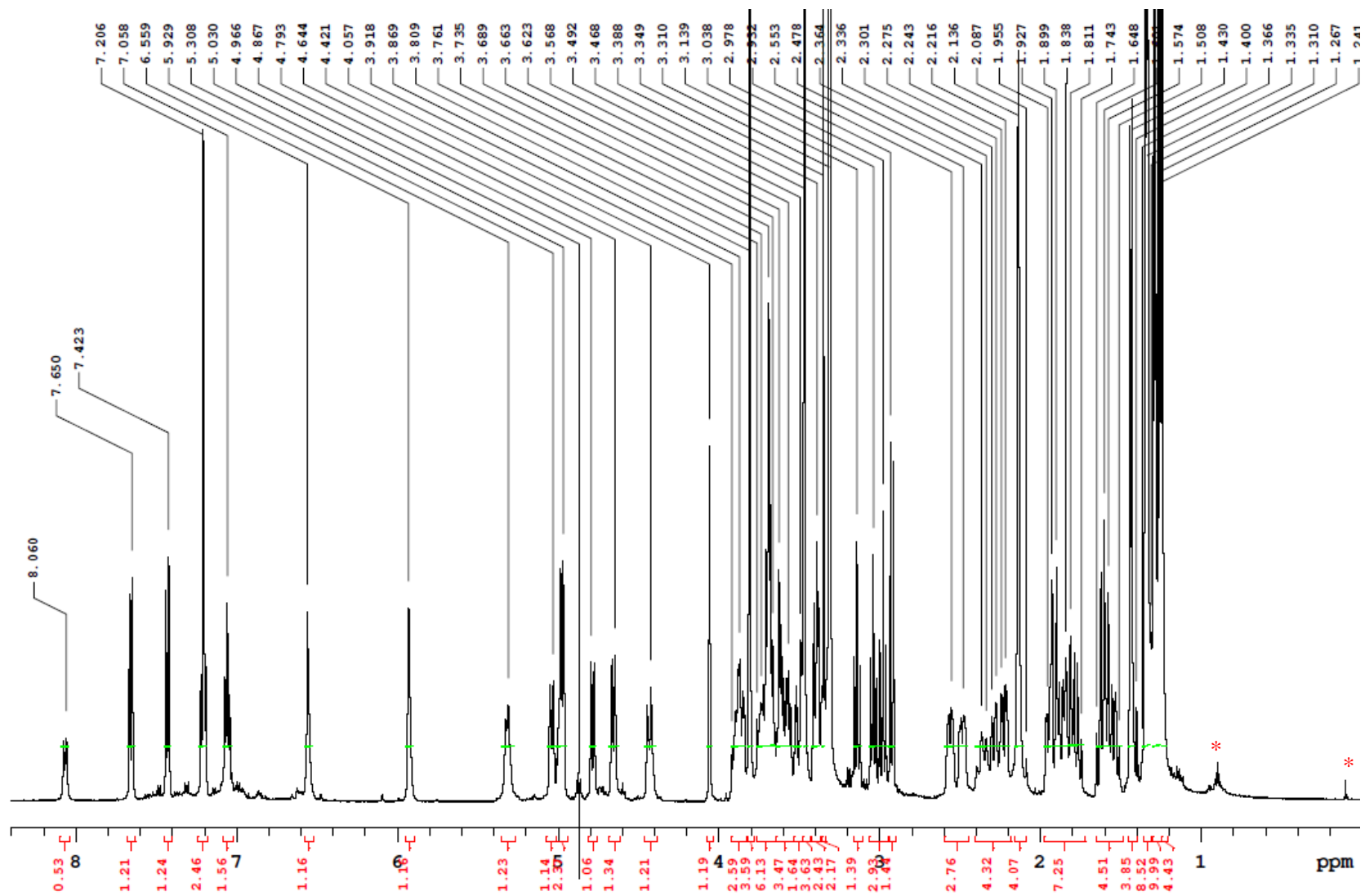


Appendix 37. gHMBC NMR spectrum of MTM SA-L-phenylalanine methyl ester (**83**)

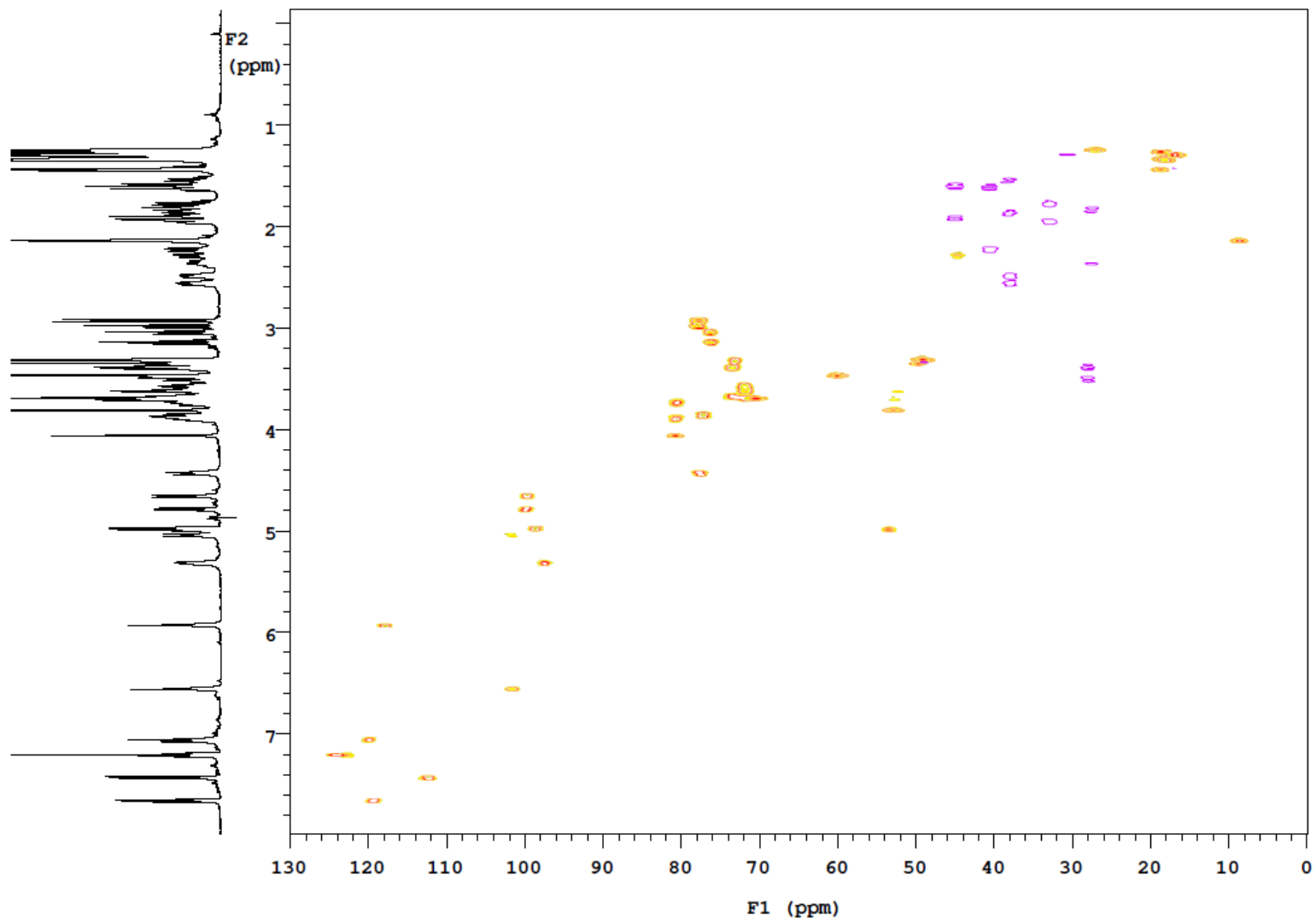


150

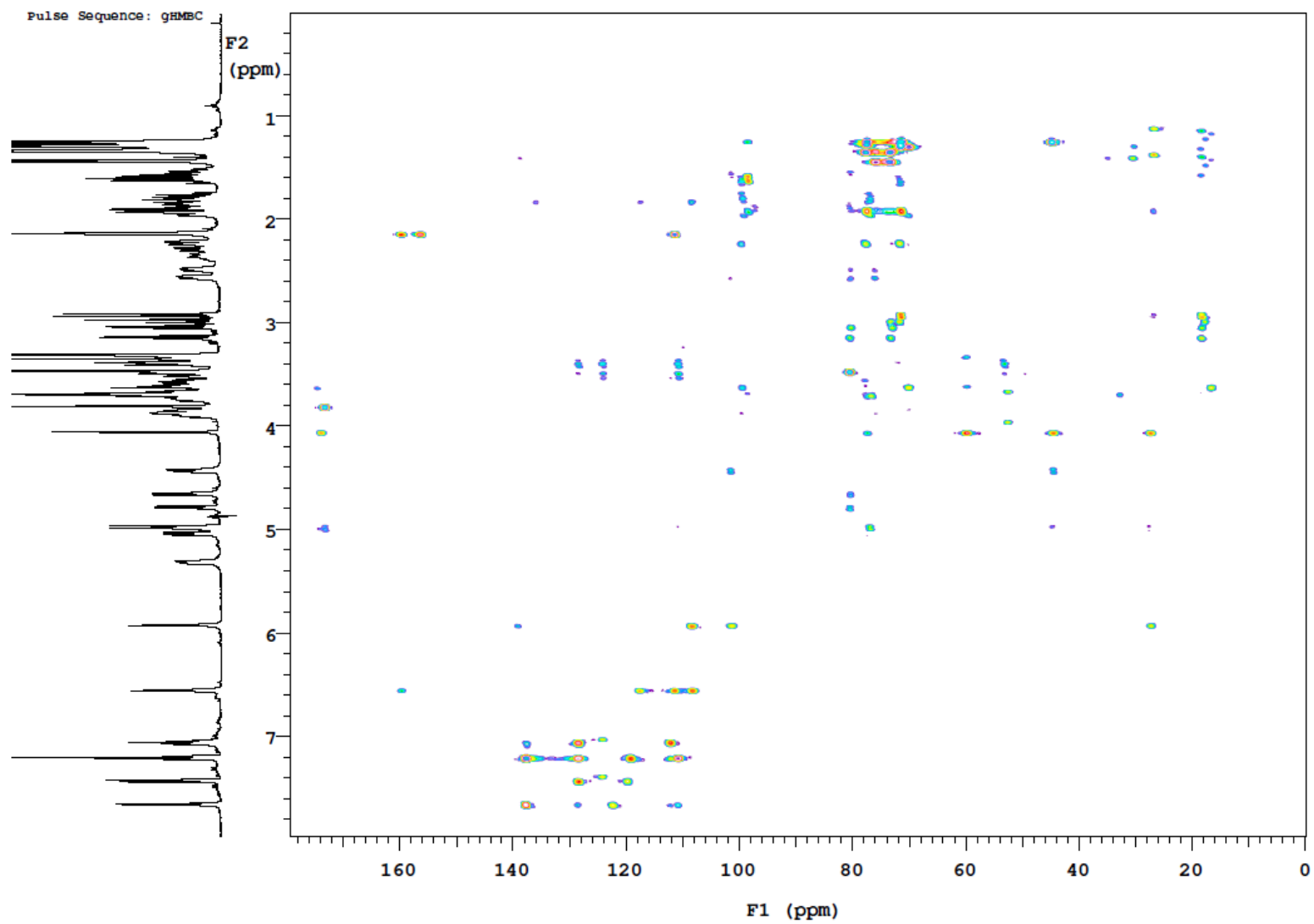
Appendix 38. HR-MALDI-TOF-MS spectrum of MTM SA-L-tryptophan methyl ester (85)



Appendix 39. ¹H NMR spectrum of MTM SA-L-tryptophan methyl ester (85) (*: signals from minor impurities)

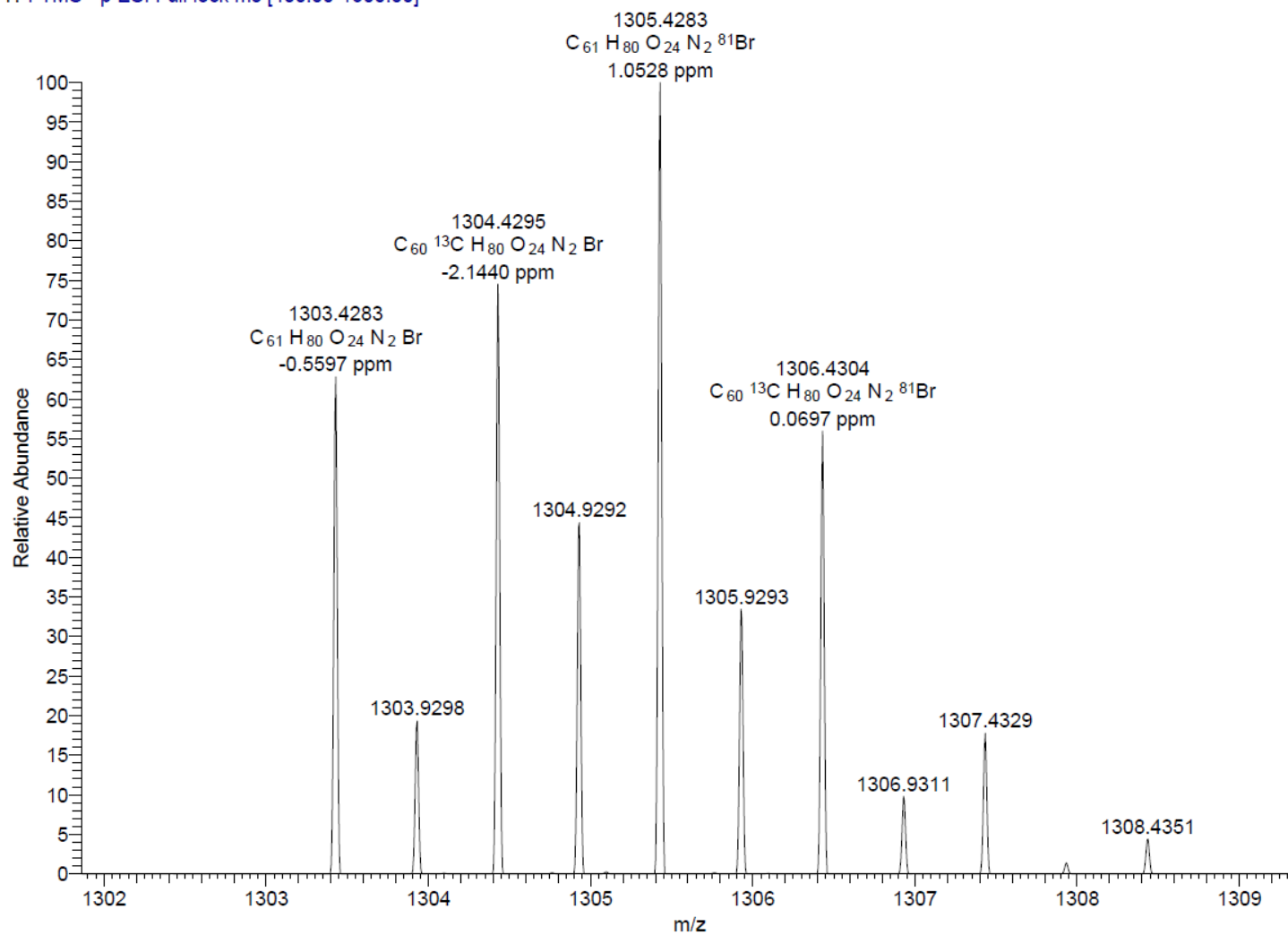


Appendix 40. gHSQC NMR spectrum of MTM SA-L-tryptophan methyl ester (**85**)

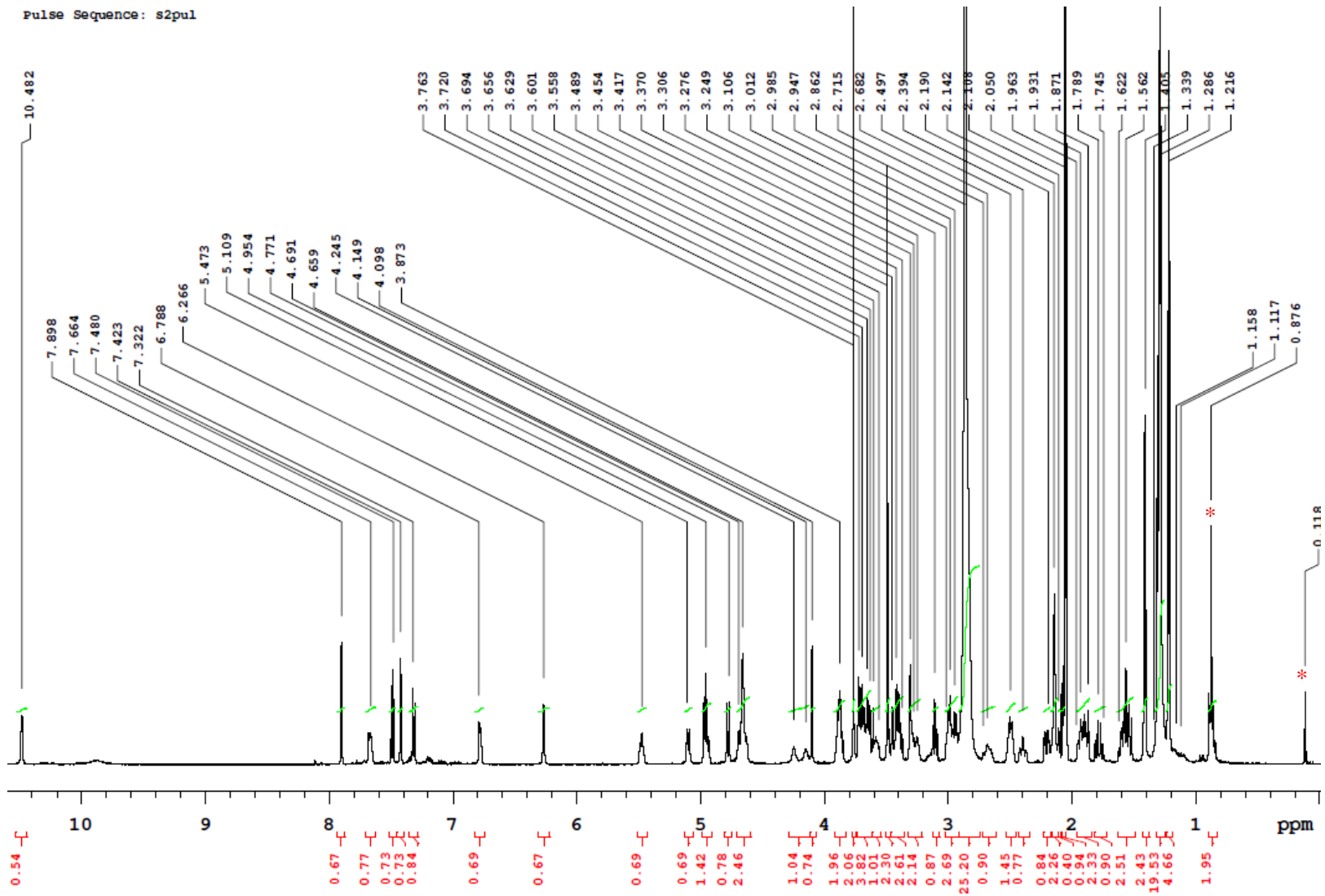


Appendix 41. gHMBC NMR spectrum of MTM SA-L-tryptophan methyl ester (**85**)

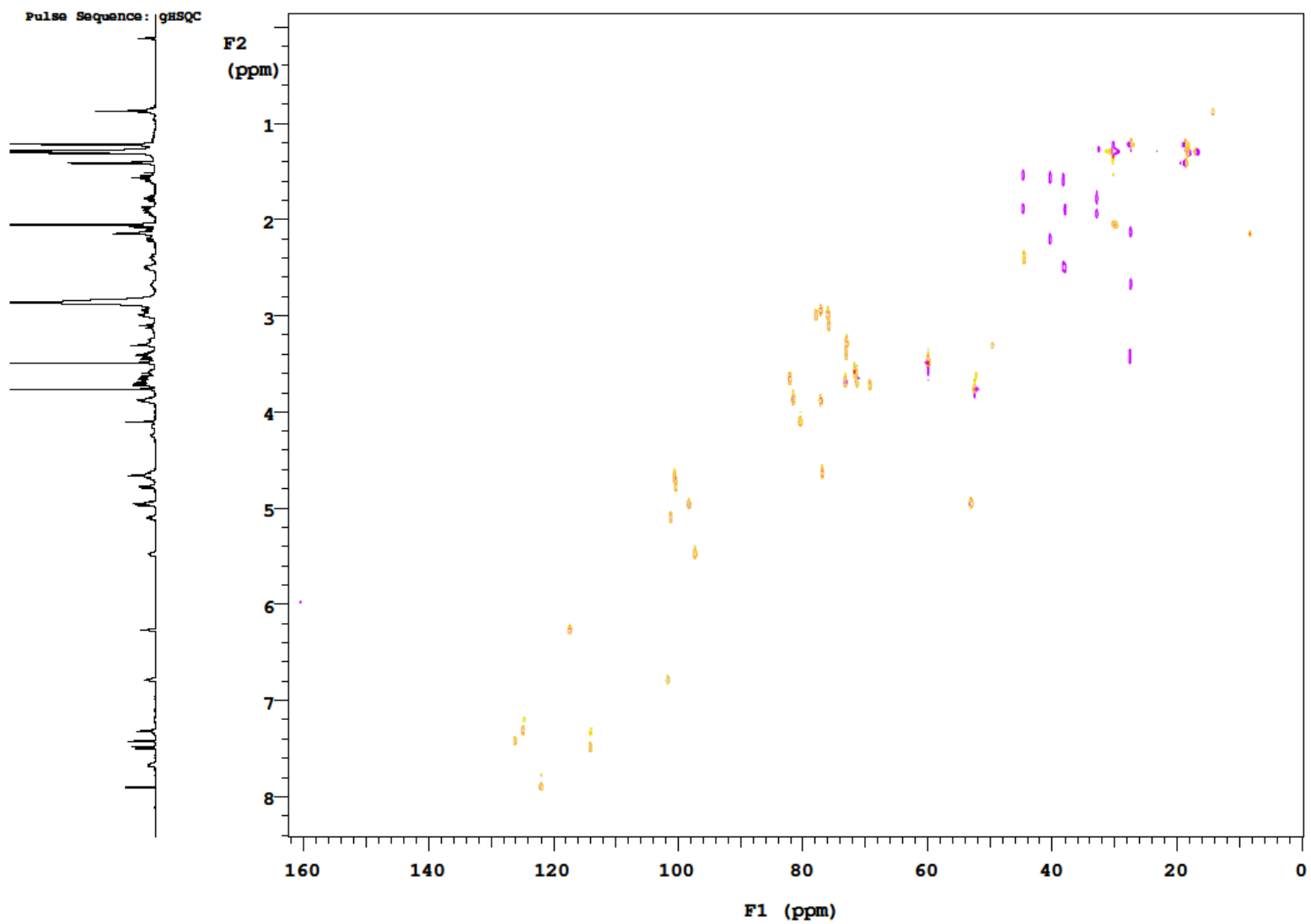
14-0548 #258-276 RT: 6.78-7.25 AV: 19 NL: 8.69E4
T: FTMS - p ESI Full lock ms [400.00-1500.00]



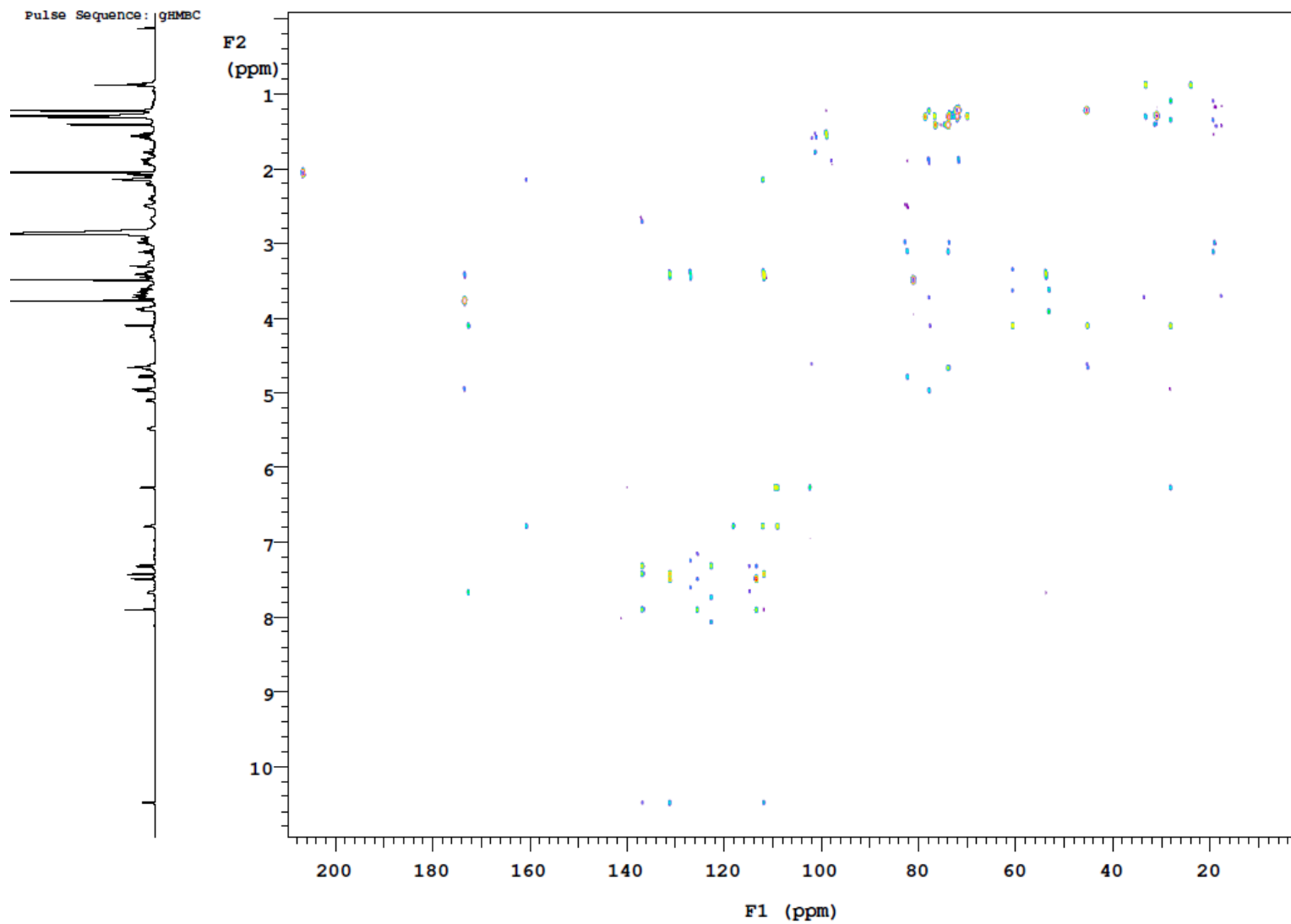
Appendix 42. HR-ESI-MS spectrum of MTM SA-L-5''-Br-tryptophan-methyl ester (**86**)



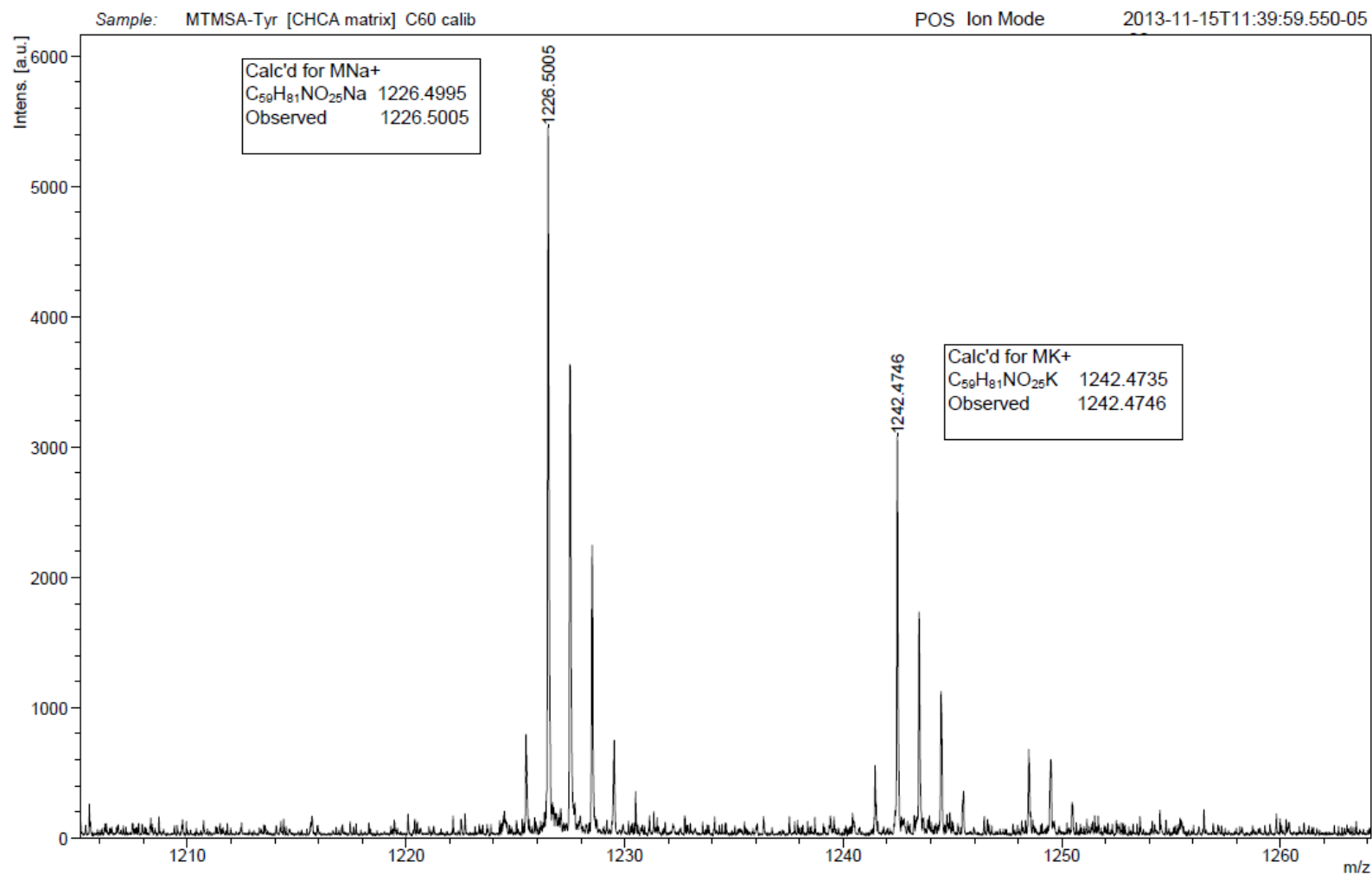
Appendix 43. ^1H NMR spectrum of MTM SA-L-5''-Br-tryptophan-methyl ester (**86**) (*: signals from minor impurities)



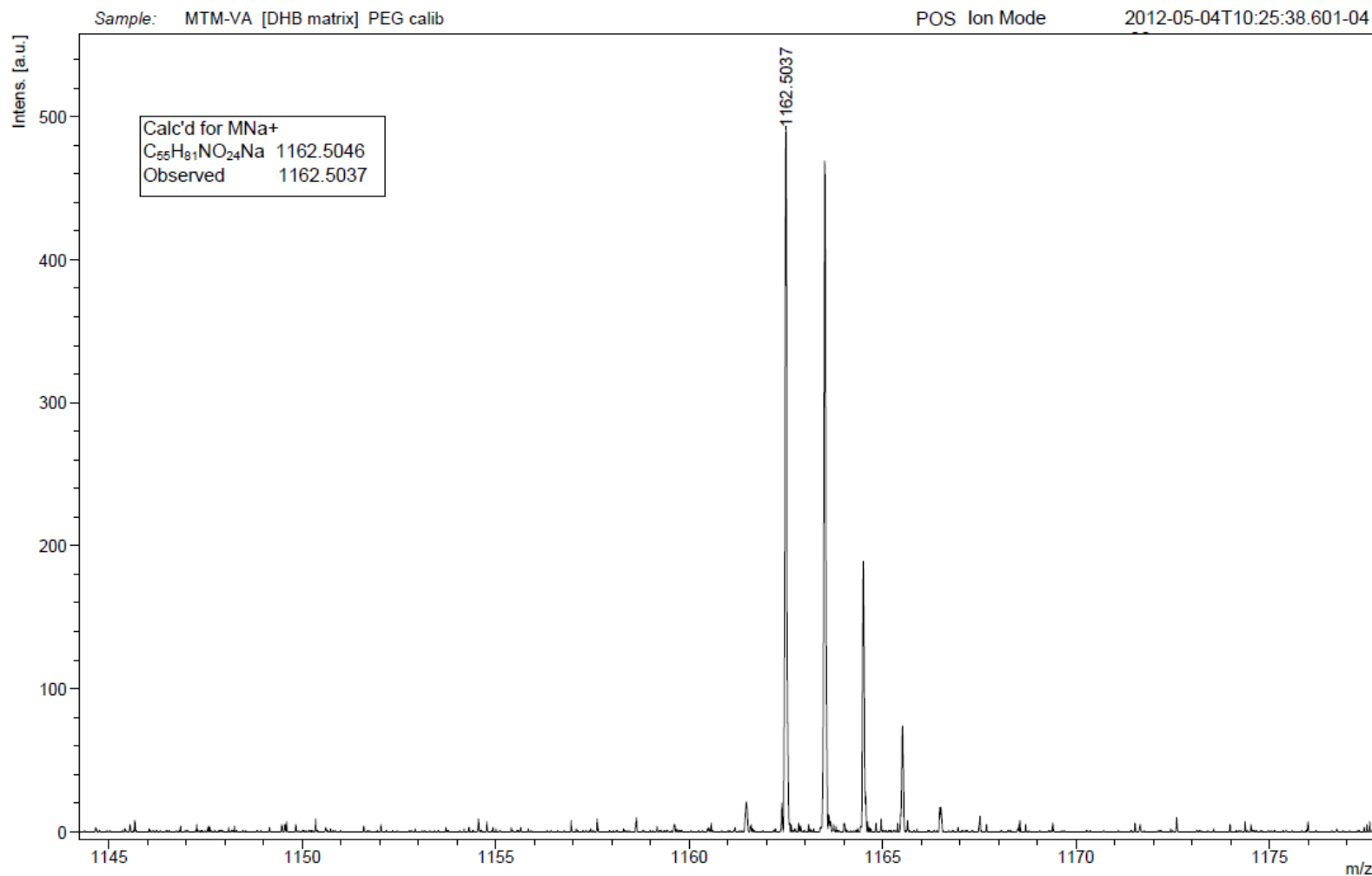
Appendix 44. gHSQC NMR spectrum of MTM SA-L-5''-Br-tryptophan-methyl ester (**86**)



Appendix 45. gHMBC NMR spectrum of MTM SA-L-5''-Br-tryptophan-methyl ester (**86**)



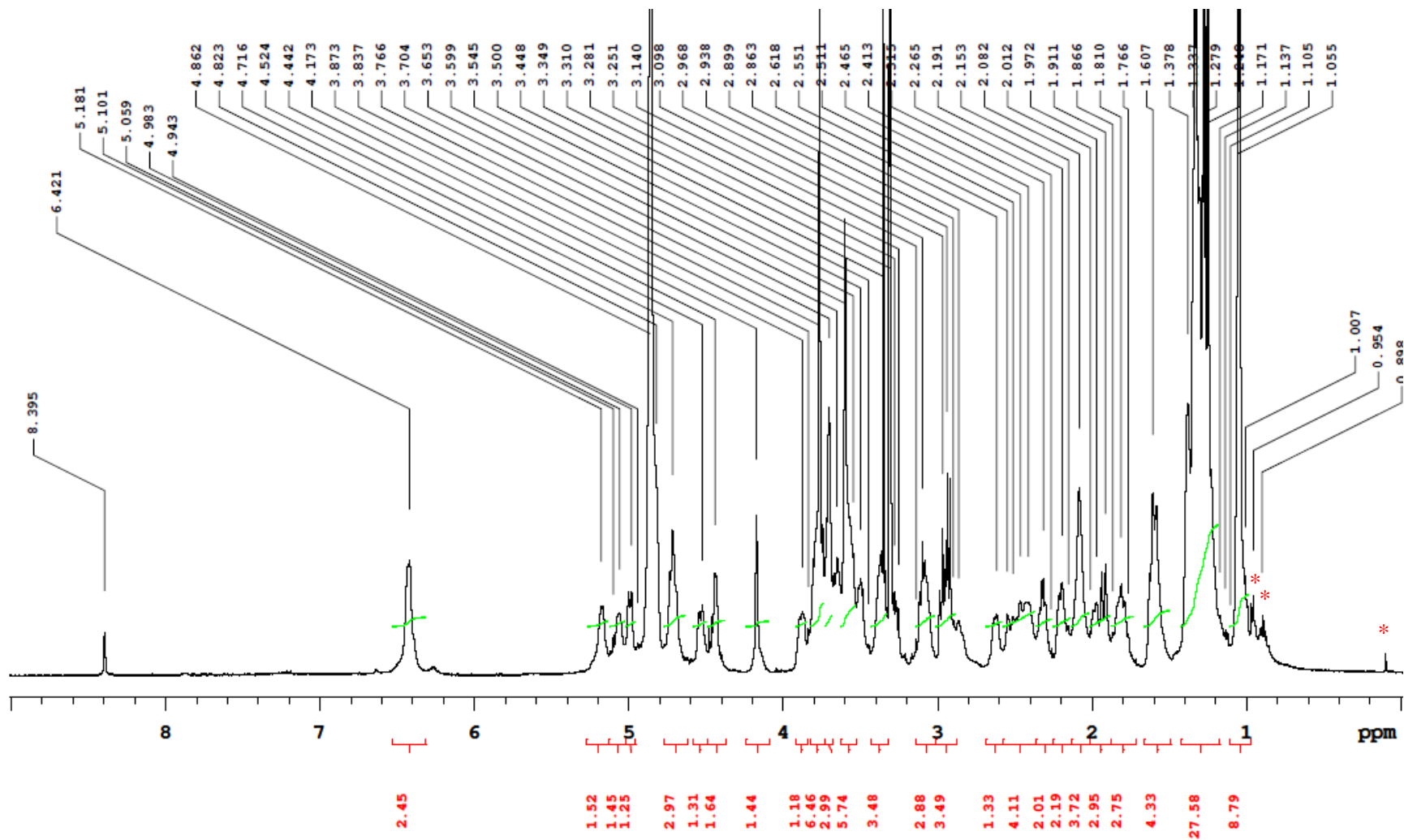
Appendix 46. HR-MALDI-TOF-MS spectrum of MTM SA-L-tyrosine methyl ester (87)



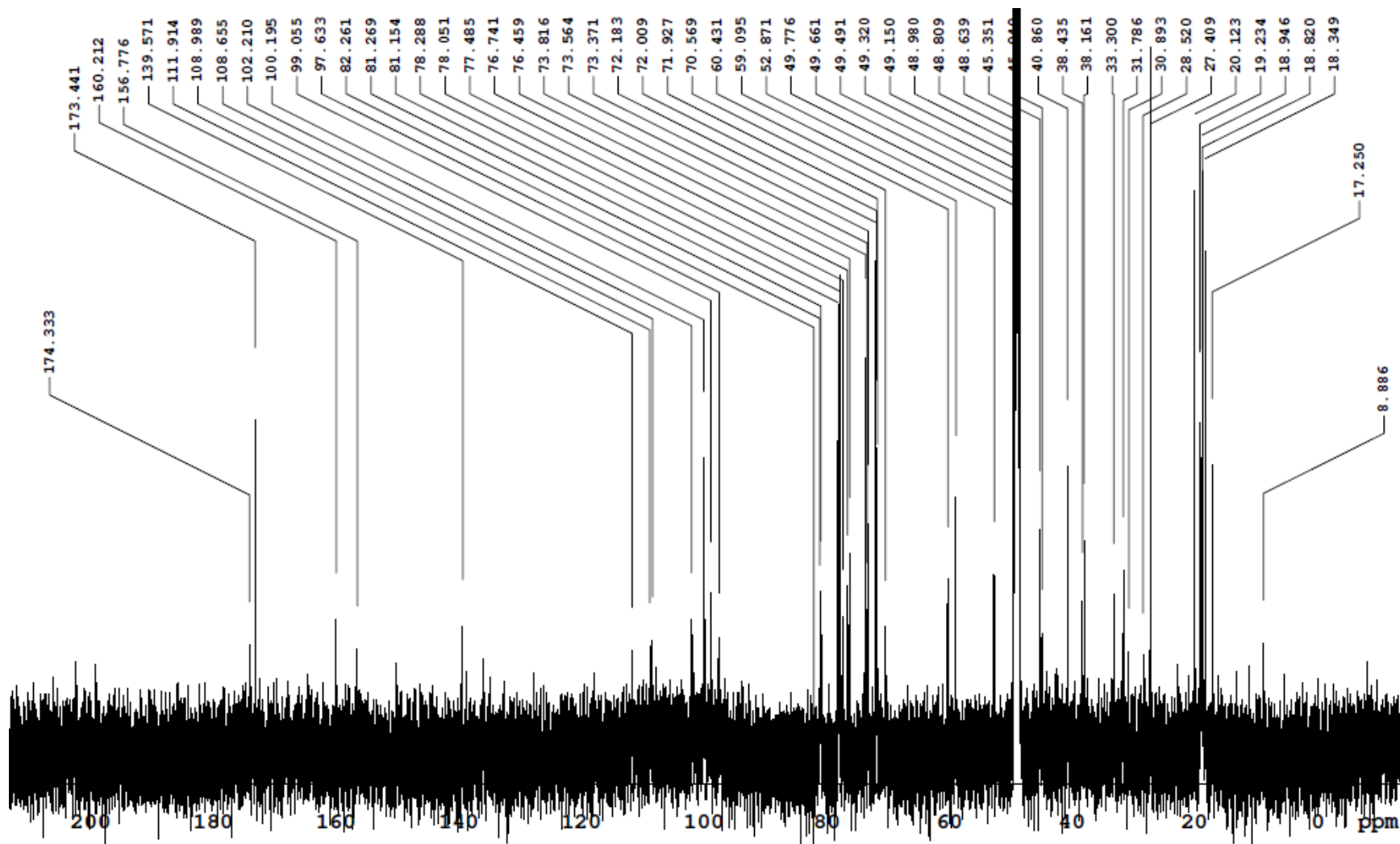
159

Appendix 47. HR-MALDI-TOF-MS spectrum of MTM SA-L-valine methyl ester (**88**)

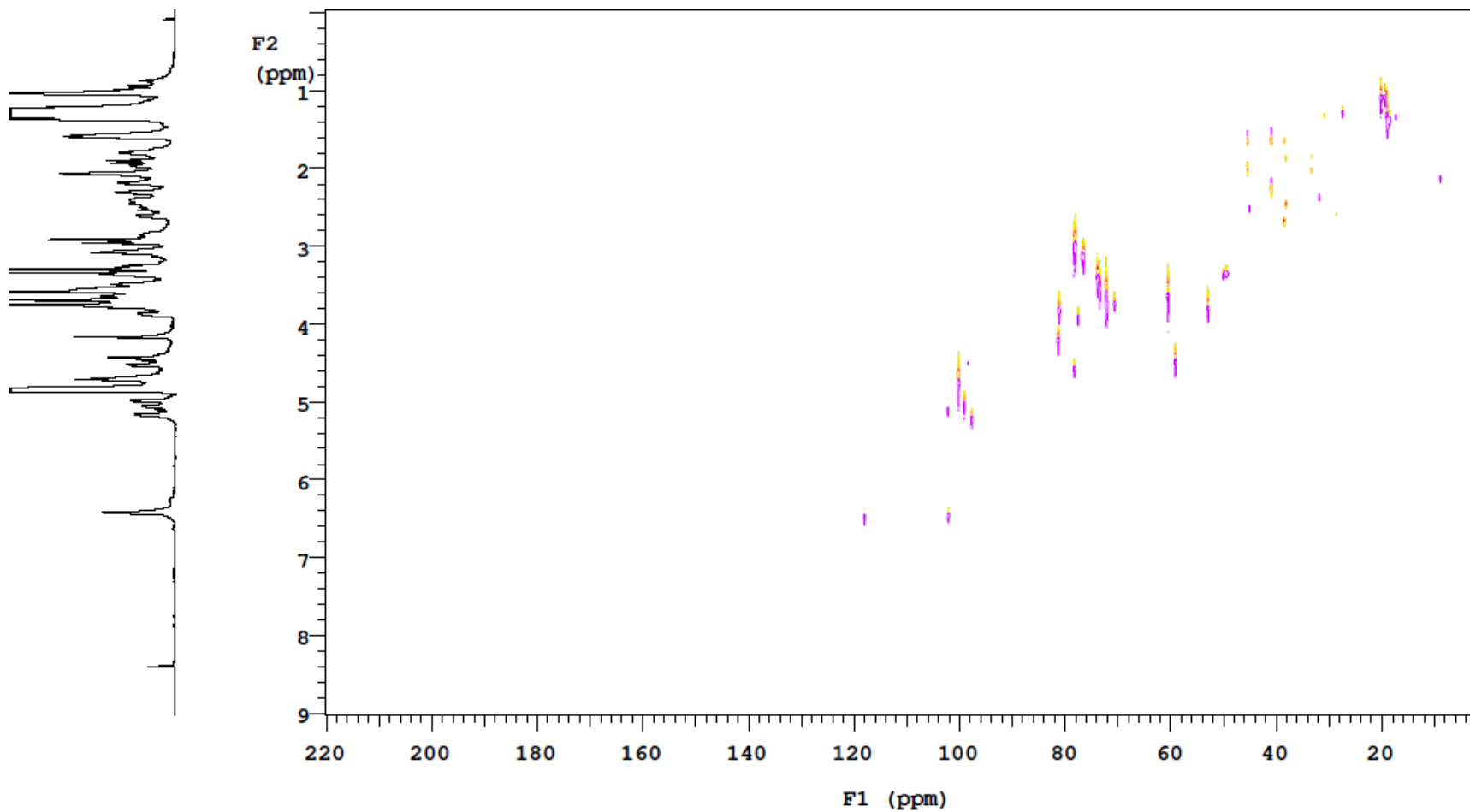
091



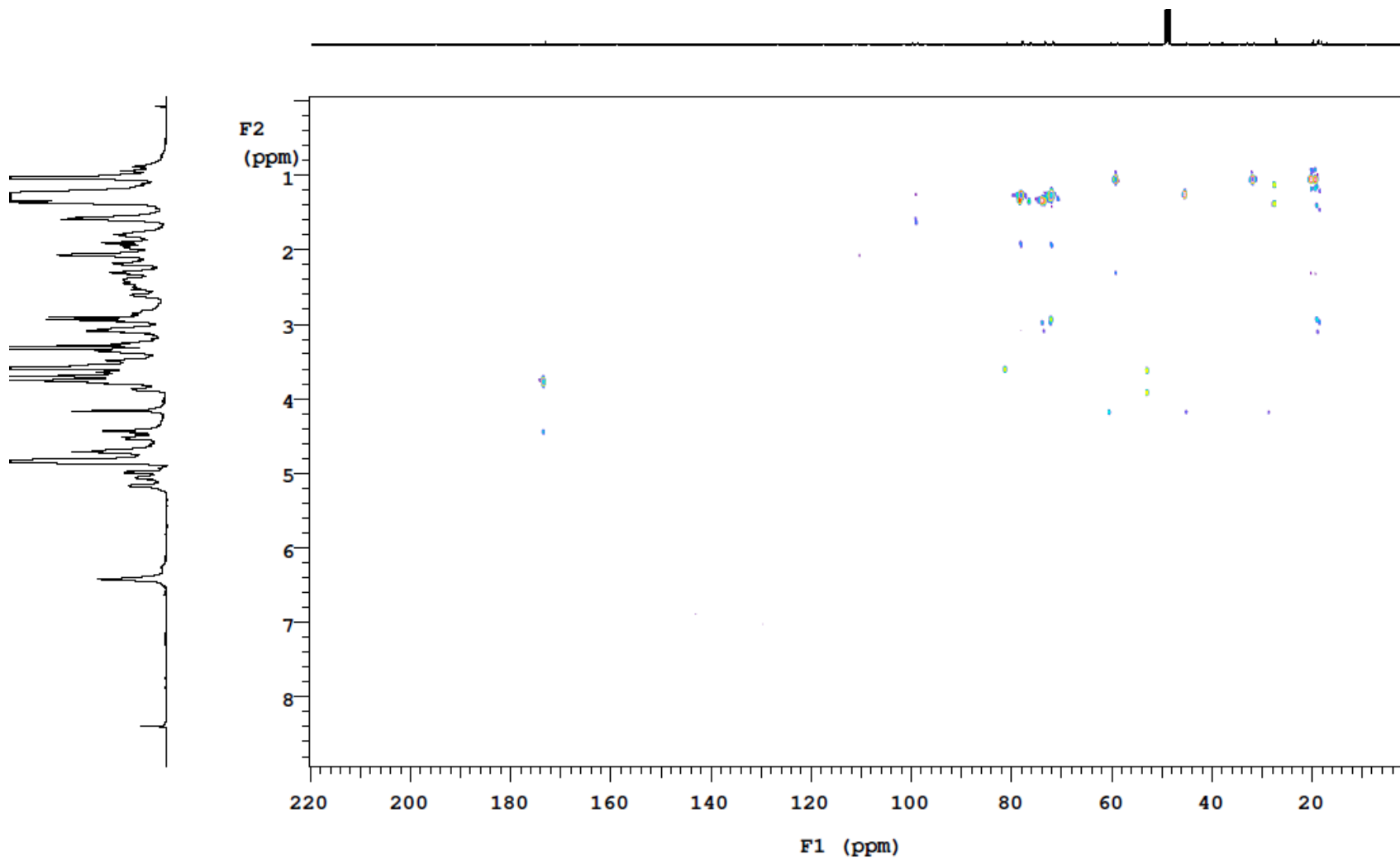
Appendix 48. ¹H NMR spectrum of MTM SA-L-valine methyl ester (88) (*: signals from minor impurities)



Appendix 49. ^{13}C NMR spectrum of MTM SA-L-valine methyl ester (88)



Appendix 50. gHSQC NMR spectrum of MTM SA-L-valine methyl ester (**88**)



Appendix 51. gHMBC NMR spectrum of MTM SA-L-valine methyl ester (**88**)

REFERENCES

- [1] Newman, D. J., Cragg, G. M., Natural products as sources of new drugs over the 30 years from 1981 to 2010, *J. Nat. Prod.*, **2012**, 75, 311-335.
- [2] Dewick, P. M., Medicinal Natural Products, 2nd edition, *John Wiley & Sons, LTD*, **2001**.
- [3] Robbers, J. E., Speedie, M. K., Tyler, V. E., Pharmacognosy and Pharmacobiotechnology, *Williams & Wilkins*, **1996**.
- [4] Dobson, C. M., Chemical space and biology, *Nature*, **2004**, 432, 824-828.
- [5] Kingston, D. G. I., Chaudhary, A. G., Gunatilaka, A. A. L., Middleton, M. L., Synthesis of taxol from baccatin III via an oxazoline intermediate, *Tetrahedron Lett.*, **1994**, 35, 4483-4484.
- [6] Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., Hopwood, D. A., Practical Streptomyces genetics, *John Innes Centre, Norwich, UK*, **2000**.
- [7] Shen, B., Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms, *Curr Opin Chem Biol*, **2003**, 7, 285-295.
- [8] a) Hertweck, C., The biosynthetic logic of polyketide diversity, *Angew. Chem. Int. Ed.*, **2009**, 48, 4688-4716; b) Masschelein, J., Mattheus, W., Gao, L. J., Moons, P., Van Houdt, R., Uytterhoeven, B., Lamberigts, C., Lescrinier, E., Rozenski, J., Herdewijn, P., Aertsens, A., Michiels, C., Lavigne, R., A PKS/NRPS/FAS hybrid gene cluster from *Serratia plymuthica* RVH1 encoding the biosynthesis of three broad spectrum, zeamine-related antibiotics, *PLoS One*, **2013**, 8, e54143; c) Wenzel, S. C., Bode, H. B., Kochems, I., Muller, R., A type I/type III polyketide synthase hybrid biosynthetic pathway for the structurally unique ansa compound kendomycin, *Chembiochem*, **2008**, 9, 2711-2721.
- [9] Garrett, R. H., Grisham, C. M., Biochemistry 3rd edition, *Thomson Brooks/Cole*, **2005**.
- [10] Weymouth-Wilson, A. C., The role of carbohydrates in biologically active natural products, *Nat. Prod. Rep.*, **1997**, 14, 99-110.
- [11] Weissman, K. J., Leadlay, P. F., Combinatorial biosynthesis of reduced polyketides, *Nat. Rev. Microbiol.*, **2005**, 3, 925-936.
- [12] Yim, G., Thaker, M. N., Koteva, K., Wright, G., Glycopeptide antibiotic biosynthesis, *J Antibiot (Tokyo)*, **2014**, 67, 31-41.
- [13] Salas, J. A., Mendez, C., Engineering the glycosylation of natural products in actinomycetes, *Trends Microbiol.*, **2007**, 15, 219-232.
- [14] a) Nunez, L. E., Nybo, S. E., Gonzalez-Sabin, J., Perez, M., Menendez, N., Brana, A. F., Shaaban, K. A., He, M., Moris, F., Salas, J. A., Rohr, J., Mendez, C., A novel mithramycin analogue with high antitumor activity and less toxicity generated by combinatorial biosynthesis, *J. Med. Chem.*, **2012**, 55, 5813-5825; b) Shepherd, M. D., Liu, T., Mendez, C., Salas, J. A., Rohr, J., Engineered biosynthesis of gilvocarcin analogues with altered deoxyhexopyranose moieties, *Appl. Environ. Microbiol.*, **2011**, 77, 435-441.

- [15] a) Borisova, S. A., Liu, H. W., Characterization of glycosyltransferase DesVII and its auxiliary partner protein DesVIII in the methymycin/picromycin biosynthetic pathway, *Biochemistry*, **2010**, *49*, 8071-8084; b) Moncrieffe, M. C., Fernandez, M. J., Spittler, D., Matsumura, H., Gay, N. J., Luisi, B. F., Leadlay, P. F., Structure of the glycosyltransferase EryCIII in complex with its activating P450 homologue EryCII, *J. Mol. Biol.*, **2012**, *415*, 92-101.
- [16] a) Peltier-Pain, P., Marchillo, K., Zhou, M., Andes, D. R., Thorson, J. S., Natural product disaccharide engineering through tandem glycosyltransferase catalysis reversibility and neoglycosylation, *Org. Lett.*, **2012**, *14*, 5086-5089; b) Zhang, C., Griffith, B. R., Fu, Q., Albermann, C., Fu, X., Lee, I. K., Li, L., Thorson, J. S., Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions, *Science*, **2006**, *313*, 1291-1294.
- [17] Lairson, L. L., Henrissat, B., Davies, G. J., Withers, S. G., Glycosyltransferases: structures, functions, and mechanisms, *Annu. Rev. Biochem.*, **2008**, *77*, 521-555.
- [18] Griffith, B. R., Thorson, J. S., A sweet success for substrate engineering, *Nat. Chem. Biol.*, **2006**, *2*, 659-660.
- [19] Olano, C., Mendez, C., Salas, J. A., Post-PKS tailoring steps in natural product-producing actinomycetes from the perspective of combinatorial biosynthesis, *Nat. Prod. Rep.*, **2010**, *27*, 571-616.
- [20] Rix, U., Fischer, C., Remsing, L. L., Rohr, J., Modification of post-PKS tailoring steps through combinatorial biosynthesis, *Nat. Prod. Rep.*, **2002**, *19*, 542-580.
- [21] Remsing, L. L., Gonzalez, A. M., Nur-e-Alam, M., Fernandez-Lozano, M. J., Brana, A. F., Rix, U., Oliveira, M. A., Mendez, C., Salas, J. A., Rohr, J., Mithramycin SK, a novel antitumor drug with improved therapeutic index, mithramycin SA, and demycarosyl-mithramycin SK: three new products generated in the mithramycin producer *Streptomyces argillaceus* through combinatorial biosynthesis, *J. Am. Chem. Soc.*, **2003**, *125*, 5745-5753.
- [22] a) Hatano, K., Higashide, E., Shibata, M., Kameda, Y., Horii, S., Mizuno, K., Toromycin, a new antibiotic produced by *Streptomyces Collinus* subsp-Albescens subsp-Nov, *Agr. Biol. Chem. Tokyo*, **1980**, *44*, 1157-1163; b) Horii, S., Fukase, H., Mizuta, E., Hatano, K., Mizuno, K., Chemistry of toromycin, *Chem. Pharm. Bull.*, **1980**, *28*, 3601-3611.
- [23] Bililign, T., Griffith, B. R., Thorson, J. S., Structure, activity, synthesis and biosynthesis of aryl-C-glycosides, *Nat. Prod. Rep.*, **2005**, *22*, 742-760.
- [24] Gutmann, A., Nidetzky, B., Switching between O- and C-glycosyltransferase through exchange of active-site motifs, *Angew. Chem. Int. Ed.*, **2012**, *51*, 12879-12883.
- [25] a) Alegria, A. E., Zayas, L., Guevara, N., A comparative study of the visible light photochemistry of gilvocarcins V and M, *Photochem. Photobiol.*, **1995**, *62*, 409-415; b) Arce, R., Oyola, R., Alegria, A. E., The photobiological differences of gilvocarcins V and M are not related to their transient intermediates and triplet yields, *Photochem. Photobiol.*, **1998**, *68*, 25-31; c) Elespuru, R. K., Gonda, S. K., Activation of antitumor agent gilvocarcins by visible light, *Science*, **1984**, *223*, 69-71; d) Knobler, R. M., Radlwimmer, F. B., Lane, M. J., Gilvocarcin V exhibits both equilibrium DNA binding and UV light induced DNA adduct formation which is sequence context dependent, *Nucleic Acids Res.*,

- 1992**, *20*, 4553-4557; e) Oyola, R., Arce, R., Alegria, A. E., Garcia, C., Photophysical properties of gilvocarcins V and M and their binding constant to calf thymus DNA, *Photochem. Photobiol.*, **1997**, *65*, 802-810.
- [26] Matsumoto, A., Hanawalt, P. C., Histone H3 and heat shock protein GRP78 are selectively cross-linked to DNA by photoactivated gilvocarcin V in human fibroblasts, *Cancer. Res.*, **2000**, *60*, 3921-3926.
- [27] a) Kharel, M. K., Pahari, P., Shepherd, M. D., Tibrewal, N., Nybo, S. E., Shaaban, K. A., Rohr, J., Angucyclines: Biosynthesis, mode-of-action, new natural products, and synthesis, *Nat. Prod. Rep.*, **2012**, *29*, 264-325; b) Kikuchi, O., Eguchi, T., Kakinuma, K., Kozuka, Y., Shindo, K., Otake, N., Chemical transformation of gilvocarcin V. Modification of the side-chain, *J. Antibiot.*, **1993**, *46*, 985-991.
- [28] a) James, C. A., Snieckus, V., Combined directed remote metalation-transition metal catalyzed cross coupling strategies: the total synthesis of the aglycones of the gilvocarcins V, M, and E and arnottin I, *J. Org. Chem.*, **2009**, *74*, 4080-4093; b) Kharel, M. K., Rohr, J., Delineation of gilvocarcin, jadomycin, and landomycin pathways through combinatorial biosynthetic enzymology, *Curr. Opin. Chem. Biol.*, **2012**, *16*, 150-161; c) Liu, T., Kharel, M. K., Zhu, L., Bright, S. A., Mattingly, C., Adams, V. R., Rohr, J., Inactivation of the ketoreductase gilU gene of the gilvocarcin biosynthetic gene cluster yields new analogues with partly improved biological activity, *Chembiochem*, **2009**, *10*, 278-286; d) Pahari, P., Kharel, M. K., Shepherd, M. D., van Lanen, S. G., Rohr, J., Enzymatic total synthesis of defucogilvocarcin M and its implications for gilvocarcin biosynthesis, *Angew. Chem. Int. Ed.*, **2012**, *51*, 1216-1220; e) Subrahmanyam, A. V., Palanichamy, K., Kaliappan, K. P., Application of an enyne metathesis/Diels-Alder cycloaddition sequence: a new versatile approach to the syntheses of C-aryl glycosides and spiro-C-aryl glycosides, *Chem. Eur. J.*, **2010**, *16*, 8545-8556.
- [29] a) Singh, S., Phillips, G. N., Jr., Thorson, J. S., The structural biology of enzymes involved in natural product glycosylation, *Nat. Prod. Rep.*, **2012**, *29*, 1201-1237; b) Thibodeaux, C. J., Melancon, C. E., Liu, H. W., Unusual sugar biosynthesis and natural product glycodiversification, *Nature*, **2007**, *446*, 1008-1016.
- [30] Fischer, C., Lipata, F., Rohr, J., The complete gene cluster of the antitumor agent gilvocarcin V and its implication for the biosynthesis of the gilvocarcins, *J. Am. Chem. Soc.*, **2003**, *125*, 7818-7819.
- [31] Shepherd, M. D., Kharel, M. K., Zhu, L. L., van Lanen, S. G., Rohr, J., Delineating the earliest steps of gilvocarcin biosynthesis: role of GilP and GilQ in starter unit specificity, *Org. Biomol. Chem.*, **2010**, *8*, 3851-3856.
- [32] Tibrewal, N., Pahari, P., Wang, G., Kharel, M. K., Morris, C., Downey, T., Hou, Y., Bugni, T. S., Rohr, J., Baeyer-Villiger C-C bond cleavage reaction in gilvocarcin and jadomycin biosynthesis, *J. Am. Chem. Soc.*, **2012**, *134*, 18181-18184.
- [33] Tibrewal, N., Downey, T. E., Van Lanen, S. G., Ul Sharif, E., O'Doherty, G. A., Rohr, J., Roles of the synergistic reductive O-methyltransferase GilM and of O-methyltransferase GilMT in the gilvocarcin biosynthetic pathway, *J. Am. Chem. Soc.*, **2012**, *134*, 12402-12405.

- [34] Kharel, M. K., Zhu, L., Liu, T., Rohr, J., Multi-oxygenase complexes of the gilvocarcin and jadomycin biosyntheses, *J. Am. Chem. Soc.*, **2007**, *129*, 3780-3781.
- [35] Thibodeaux, C. J., Melancon, C. E., 3rd, Liu, H. W., Natural-product sugar biosynthesis and enzymatic glycodiversification, *Angew. Chem. Int. Ed.*, **2008**, *47*, 9814-9859.
- [36] Lombo, F., Menendez, N., Salas, J. A., Mendez, C., The aureolic acid family of antitumor compounds: structure, mode of action, biosynthesis, and novel derivatives, *Appl. Microbiol. Biotechnol.*, **2006**, *73*, 1-14.
- [37] Hsu, C. W., Kuo, C. F., Chuang, S. M., Hou, M. H., Elucidation of the DNA-interacting properties and anticancer activity of a Ni(II)-coordinated mithramycin dimer complex, *Biometals*, **2013**, *26*, 1-12.
- [38] Sastry, M., Patel, D. J., Solution structure of the mithramycin dimer-DNA complex, *Biochemistry*, **1993**, *32*, 6588-6604.
- [39] a) Keniry, M. A., Owen, E. A., Shafer, R. H., The three-dimensional structure of the 4:1 mithramycin:d(ACCCGGGT)(2) complex: evidence for an interaction between the E saccharides, *Biopolymers*, **2000**, *54*, 104-114; b) Sastry, M., Fiala, R., Patel, D. J., Solution structure of mithramycin dimers bound to partially overlapping sites on DNA, *J. Mol. Biol.*, **1995**, *251*, 674-689.
- [40] Vizcaino, C., Mansilla, S., Portugal, J., Sp1 transcription factor: A long-standing target in cancer chemotherapy, *Pharmacol. Ther.*, **2015**.
- [41] a) Abbott, B. L., Colapietro, A. M., Barnes, Y., Marini, F., Andreeff, M., Sorrentino, B. P., Low levels of ABCG2 expression in adult AML blast samples, *Blood*, **2002**, *100*, 4594-4601; b) Hall, T. J., Schaeublin, M., Chambers, T. J., The majority of osteoclasts require mRNA and protein synthesis for bone resorption in vitro, *Biochem. Biophys. Res. Commun.*, **1993**, *195*, 1245-1253; c) Osada, N., Kosuge, Y., Ishige, K., Ito, Y., Mithramycin, an agent for developing new therapeutic drugs for neurodegenerative diseases, *J. Pharmacol. Sci.*, **2013**, *122*, 251-256; d) Taylor, D. J., Parsons, C. E., Han, H., Jayaraman, A., Rege, K., Parallel screening of FDA-approved antineoplastic drugs for identifying sensitizers of TRAIL-induced apoptosis in cancer cells, *BMC Cancer*, **2011**, *11*, 470; e) Zhang, M., Mathur, A., Zhang, Y., Xi, S., Atay, S., Hong, J. A., Datrice, N., Upham, T., Kemp, C. D., Ripley, R. T., Wiegand, G., Avital, I., Fetsch, P., Mani, H., Zlott, D., Robey, R., Bates, S. E., Li, X., Rao, M., Schrupp, D. S., Mithramycin represses basal and cigarette smoke-induced expression of ABCG2 and inhibits stem cell signaling in lung and esophageal cancer cells, *Cancer Res.*, **2012**, *72*, 4178-4192.
- [42] a) Brown, J. H., Kennedy, B. J., Mithramycin in the Treatment of Disseminated Testicular Neoplasms, *N. Engl. J. Med.*, **1965**, *272*, 111-118; b) Kennedy, B. J., Metabolic and toxic effects of mithramycin during tumor therapy, *Am. J. Med.*, **1970**, *49*, 494-503; c) Koller, C. A., Miller, D. M., Preliminary observations on the therapy of the myeloid blast phase of chronic granulocytic leukemia with plicamycin and hydroxyurea, *N. Engl. J. Med.*, **1986**, *315*, 1433-1438; d) Ryan, W. G., Mithramycin for Paget's disease of bone, *N. Engl. J. Med.*, **1970**, *283*, 1171; e) Ryan, W. G., Schwartz, T. B., Northrop, G., Experiences in the treatment of Paget's disease of bone with mithramycin, *JAMA*, **1970**, *213*, 1153-1157.

- [43] Grohar, P. J., Woldemichael, G. M., Griffin, L. B., Mendoza, A., Chen, Q. R., Yeung, C., Currier, D. G., Davis, S., Khanna, C., Khan, J., McMahon, J. B., Helman, L. J., Identification of an inhibitor of the EWS-FLI1 oncogenic transcription factor by high-throughput screening, *J. Natl. Cancer. Inst.*, **2011**, *103*, 962-978.
- [44] a) Blanco, G., Fu, H., Mendez, C., Khosla, C., Salas, J. A., Deciphering the biosynthetic origin of the aglycone of the aureolic acid group of anti-tumor agents, *Chem. Biol.*, **1996**, *3*, 193-196; b) Lombo, F., Blanco, G., Fernandez, E., Mendez, C., Salas, J. A., Characterization of *Streptomyces argillaceus* genes encoding a polyketide synthase involved in the biosynthesis of the antitumor mithramycin, *Gene*, **1996**, *172*, 87-91.
- [45] Wang, G., Kharel, M. K., Pahari, P., Rohr, J., Investigating Mithramycin deoxysugar biosynthesis: enzymatic total synthesis of TDP-D-olivose, *Chembiochem*, **2011**, *12*, 2568-2571.
- [46] a) Guyett, P., Glushka, J., Gu, X., Bar-Peled, M., Real-time NMR monitoring of intermediates and labile products of the bifunctional enzyme UDP-apiose/UDP-xylose synthase, *Carbohydr. Res.*, **2009**, *344*, 1072-1078; b) Wang, G., Pahari, P., Kharel, M. K., Chen, J., Zhu, H., Van Lanen, S. G., Rohr, J., Cooperation of two bifunctional enzymes in the biosynthesis and attachment of deoxysugars of the antitumor antibiotic mithramycin, *Angew. Chem. Int. Ed.*, **2012**, *51*, 10638-10642.
- [47] Nur-e-Alam, M., Mendez, C., Salas, J. A., Rohr, J., Elucidation of the glycosylation sequence of mithramycin biosynthesis: isolation of 3A-deolivosylpremithramycin B and its conversion to premithramycin B by glycosyltransferase MtmGII, *Chembiochem*, **2005**, *6*, 632-636.
- [48] Bosserman, M. A., Downey, T., Noinaj, N., Buchanan, S. K., Rohr, J., Molecular insight into substrate recognition and catalysis of Baeyer-Villiger monooxygenase MtmOIV, the key frame-modifying enzyme in the biosynthesis of anticancer agent mithramycin, *ACS Chem. Biol.*, **2013**, *8*, 2466-2477.
- [49] Wang, Q., Ding, P., Perepelov, A. V., Xu, Y., Wang, Y., Knirel, Y. A., Wang, L., Feng, L., Characterization of the dTDP-D-fucofuranose biosynthetic pathway in *Escherichia coli* O52, *Mol. Microbiol.*, **2008**, *70*, 1358-1367.
- [50] a) Cai, X., Ng, K., Panesar, H., Moon, S. J., Paredes, M., Ishida, K., Hertweck, C., Minehan, T. G., Total synthesis of the antitumor natural product polycarcin V and evaluation of its DNA binding profile, *Org. Lett.*, **2014**, *16*, 2962-2965; b) Li, Y. Q., Huang, X. S., Ishida, K., Maier, A., Kelter, G., Jiang, Y., Peschel, G., Menzel, K. D., Li, M. G., Wen, M. L., Xu, L. H., Grabley, S., Fiebig, H. H., Jiang, C. L., Hertweck, C., Sattler, I., Plasticity in gilvocarcin-type C-glycoside pathways: discovery and antitumoral evaluation of polycarcin V from *Streptomyces polyformus*, *Org. Biomol. Chem.*, **2008**, *6*, 3601-3605.
- [51] a) Gullon, S., Olano, C., Abdelfattah, M. S., Brana, A. F., Rohr, J., Mendez, C., Salas, J. A., Isolation, characterization, and heterologous expression of the biosynthesis gene cluster for the antitumor anthracycline steffimycin, *Appl. Environ. Microbiol.*, **2006**, *72*, 4172-4183; b) Kim, H. J., White-Phillip, J. A., Ogasawara, Y., Shin, N., Isiorho, E. A., Liu, H. W., Biosynthesis of spinosyn in *Saccharopolyspora spinosa*: synthesis of permethylated rhamnose and characterization of the functions of SpnH, SpnI, and SpnK, *J. Am. Chem.*

- Soc.*, **2010**, *132*, 2901-2903; c) Patallo, E. P., Blanco, G., Fischer, C., Brana, A. F., Rohr, J., Mendez, C., Salas, J. A., Deoxysugar methylation during biosynthesis of the antitumor polyketide elloramycin by *Streptomyces olivaceus*. Characterization of three methyltransferase genes, *J. Biol. Chem.*, **2001**, *276*, 18765-18774.
- [52] Eguchi, T., Li, H. Y., Kazami, J., Kakinuma, K., Otake, N., Intermolecular stacking of gilvocarcin V tetraacetate as evidenced by nuclear magnetic resonance studies, *J. Antibiot.*, **1990**, *43*, 1077-1081.
- [53] a) Kawai, K., Wang, G., Okamoto, S., Ochi, K., The rare earth, scandium, causes antibiotic overproduction in *Streptomyces* spp, *FEMS Microbiol. Lett.*, **2007**, *274*, 311-315; b) Tanaka, Y., Hosaka, T., Ochi, K., Rare earth elements activate the secondary metabolite-biosynthetic gene clusters in *Streptomyces coelicolor* A3(2), *J. Antibiot.*, **2010**, *63*, 477-481.
- [54] Floss, H. G., Combinatorial biosynthesis--potential and problems, *J. Biotechnol.*, **2006**, *124*, 242-257.
- [55] a) Hopwood, D. A., Malpartida, F., Kieser, H. M., Ikeda, H., Duncan, J., Fujii, I., Rudd, B. A., Floss, H. G., Omura, S., Production of 'hybrid' antibiotics by genetic engineering, *Nature*, **1985**, *314*, 642-644; b) Malpartida, F., Hopwood, D. A., Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host, *Nature*, **1984**, *309*, 462-464.
- [56] Albertini, V., Jain, A., Vignati, S., Napoli, S., Rinaldi, A., Kwee, I., Nur-e-Alam, M., Bergant, J., Bertoni, F., Carbone, G. M., Rohr, J., Catapano, C. V., Novel GC-rich DNA-binding compound produced by a genetically engineered mutant of the mithramycin producer *Streptomyces argillaceus* exhibits improved transcriptional repressor activity: implications for cancer therapy, *Nucleic. Acids. Res.*, **2006**, *34*, 1721-1734.
- [57] a) Tevyashova, A. N., Olsufyeva, E. N., Balzarini, J., Shtil, A. A., Dezhenkova, L. G., Bukhman, V. M., Zbarsky, V. B., Preobrazhenskaya, M. N., Modification of the antibiotic olivomycin I at the 2'-keto group of the side chain. Novel derivatives, antitumor and topoisomerase I-poisoning activity, *J. Antibiot.*, **2009**, *62*, 37-41; b) Tevyashova, A. N., Olsufyeva, E. N., Turchin, K. F., Balzarini, J., Bykov, E. E., Dezhenkova, L. G., Shtil, A. A., Preobrazhenskaya, M. N., Reaction of the antitumor antibiotic olivomycin I with aryl diazonium salts. Synthesis, cytotoxic and antiretroviral potency of 5-aryldiazenyl-6-O-deglycosyl derivatives of olivomycin I, *Bioorg. Med. Chem.*, **2009**, *17*, 4961-4967; c) Tevyashova, A. N., Shtil, A. A., Olsufyeva, E. N., Luzikov, Y. N., Reznikova, M. I., Dezhenkova, L. G., Isakova, E. B., Bukhman, V. M., Durandin, N. A., Vinogradov, A. M., Kuzmin, V. A., Preobrazhenskaya, M. N., Modification of olivomycin A at the side chain of the aglycon yields the derivative with perspective antitumor characteristics, *Bioorg. Med. Chem.*, **2011**, *19*, 7387-7393.
- [58] Shoemaker, R. H., The NCI 60 human tumour cell line anticancer drug screen, *Nat. Rev. Cancer*, **2006**, *6*, 813-823.
- [59] Scott, D., Chen, J. M., Bae, Y., Rohr, J., Semi-synthetic mithramycin SA derivatives with improved anticancer activity, *Chem. Biol. Drug. Des.*, **2013**, *81*, 615-624.

- [60] Scott, D., Rohr, J., Bae, Y., Nanoparticulate formulations of mithramycin analogs for enhanced cytotoxicity, *Int. J. Nanomedicine*, **2011**, *6*, 2757-2767.
- [61] Mann, C. M., Markham, J. L., A new method for determining the minimum inhibitory concentration of essential oils, *J. Appl. Microbiol.*, **1998**, *84*, 538-544.
- [62] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., Basic local alignment search tool, *J. Mol. Biol.*, **1990**, *215*, 403-410.
- [63] Bruender, N. A., Holden, H. M., Probing the catalytic mechanism of a C-3'-methyltransferase involved in the biosynthesis of D-tetronitrose, *Protein Sci.*, **2012**, *21*, 876-886.
- [64] a) Klimishin, D. A., Rabyk, M. V., Fedorenko, V. A., Nogalose methylation in the nogalamycin biosynthesis in *Streptomyces nogales* LV65, *Microbiology*, **2013**, *82*, 162-168; b) Torkkell, S., Kunnari, T., Palmu, K., Mantsala, P., Hakala, J., Ylihonko, K., The entire nogalamycin biosynthetic gene cluster of *Streptomyces nogalater*: characterization of a 20-kb DNA region and generation of hybrid structures, *Mol. Genet. Genomics*, **2001**, *266*, 276-288.
- [65] Li, W., Khullar, A., Chou, S., Sacramo, A., Gerratana, B., Biosynthesis of sibiromycin, a potent antitumor antibiotic, *Appl. Environ. Microbiol.*, **2009**, *75*, 2869-2878.
- [66] a) Freitag, A., Li, S. M., Heide, L., Biosynthesis of the unusual 5,5-gem-dimethyl-deoxysugar noviose: investigation of the C-methyltransferase gene cloU, *Microbiology*, **2006**, *152*, 2433-2442; b) Pojer, F., Li, S. M., Heide, L., Molecular cloning and sequence analysis of the clorobiocin biosynthetic gene cluster: new insights into the biosynthesis of aminocoumarin antibiotics, *Microbiology*, **2002**, *148*, 3901-3911.
- [67] a) Steffensky, M., Muhlenweg, A., Wang, Z. X., Li, S. M., Heide, L., Identification of the novobiocin biosynthetic gene cluster of *Streptomyces spheroides* NCIB 11891, *Antimicrob. Agents Chemother.*, **2000**, *44*, 1214-1222; b) Thuy, T. T., Lee, H. C., Kim, C. G., Heide, L., Sohng, J. K., Functional characterizations of novWUS involved in novobiocin biosynthesis from *Streptomyces spheroides*, *Arch. Biochem. Biophys.*, **2005**, *436*, 161-167.
- [68] Wang, Z. X., Li, S. M., Heide, L., Identification of the coumermycin A(1) biosynthetic gene cluster of *Streptomyces rishiriensis* DSM 40489, *Antimicrob. Agents Chemother.*, **2000**, *44*, 3040-3048.
- [69] Chen, H., Thomas, M. G., Hubbard, B. K., Losey, H. C., Walsh, C. T., Burkart, M. D., Deoxysugars in glycopeptide antibiotics: enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis, *Proc. Natl. Acad. Sci. USA*, **2000**, *97*, 11942-11947.
- [70] Xiao, Y., Li, S., Niu, S., Ma, L., Zhang, G., Zhang, H., Zhang, G., Ju, J., Zhang, C., Characterization of tiacumicin B biosynthetic gene cluster affording diversified tiacumicin analogues and revealing a tailoring dihalogenase, *J. Am. Chem. Soc.*, **2011**, *133*, 1092-1105.
- [71] Weitnauer, G., Muhlenweg, A., Trefzer, A., Hoffmeister, D., Sussmuth, R. D., Jung, G., Welzel, K., Vente, A., Girreser, U., Bechthold, A., Biosynthesis of the orthosomycin

- antibiotic avilamycin A: deductions from the molecular analysis of the avi biosynthetic gene cluster of *Streptomyces viridochromogenes* Tu57 and production of new antibiotics, *Chem. Biol.*, **2001**, 8, 569-581.
- [72] Bate, N., Butler, A. R., Smith, I. P., Cundliffe, E., The mycarose-biosynthetic genes of *Streptomyces fradiae*, producer of tylosin, *Microbiology*, **2000**, 146 (Pt 1), 139-146.
- [73] Gaisser, S., Bohm, G. A., Doumith, M., Raynal, M. C., Dhillon, N., Cortes, J., Leadlay, P. F., Analysis of eryBI, eryBIII and eryBVII from the erythromycin biosynthetic gene cluster in *Saccharopolyspora erythraea*, *Mol. Gen. Genet.*, **1998**, 258, 78-88.
- [74] Bruender, N. A., Thoden, J. B., Kaur, M., Avey, M. K., Holden, H. M., Molecular architecture of a C-3'-methyltransferase involved in the biosynthesis of D-tetronitrose, *Biochemistry*, **2010**, 49, 5891-5898.
- [75] Otwinowski, Z., Minor, W., Processing of X-ray diffraction data collected in oscillation mode., *Methods Enzymol.*, **1997**, 276, 307-326.
- [76] Vagin, A., Teplyakov, A., Molecular replacement with MOLREP, *Acta Crystallogr.*, **2010**, D66, 22-25.
- [77] a) Emsley, P., Cowtan, K., Coot: Model-building tools for molecular graphics., *Acta Crystallogr.*, **2004**, D60, 2126-2132; b) Murshudov, G. N., Vagin, A. A., Dodson, E. J., Refinement of macromolecular structures by the maximum-likelihood method., *Acta Crystallogr.*, **1997**, D53, 240-255.
- [78] Laskowski, R. A., Macarthur, M. W., Moss, D. S., Thornton, J. M., Procheck: A program to check the stereochemical quality of protein structures., *J. Appl. Crystallogr.*, **1993**, 26, 283-291.
- [79] Hou, J., Liu, P., Qu, H., Fu, P., Wang, Y., Wang, Z., Li, Y., Teng, X., Zhu, W., Gilvocarcin HE: a new polyketide glycoside from *Streptomyces* sp, *J Antibiot (Tokyo)*, **2012**, 65, 523-526.
- [80] Tachiwana, H., Osakabe, A., Shiga, T., Miya, Y., Kimura, H., Kagawa, W., Kurumizaka, H., Structures of human nucleosomes containing major histone H3 variants, *Acta Crystallogr D Biol Crystallogr*, **2011**, 67, 578-583.
- [81] Vizcaino, C., Nunez, L. E., Moris, F., Portugal, J., Genome-wide modulation of gene transcription in ovarian carcinoma cells by a new mithramycin analogue, *PLoS One*, **2014**, 9, e104687.

VITA

Jhong-Min Chen

Place of birth:

Keelung City, Taiwan (Republic of China)

Education:

- 2006 B.S., Botany Division, Life Sciences, National Taiwan University, Taipei, Taiwan (GPA: 3.43/4.00)
- 2008 M.S., Drug Discovery Division, Pharmaceutical Sciences, National Taiwan University, Taipei, Taiwan (GPA: 3.77/4.00)

Positions and training:

- 2004-2006 Undergraduate Research, Department of Pharmacy, College of Medicine, National Taiwan University, Taipei, Taiwan. (Advisor: Shoei-Sheng Lee, Ph.D.)
- 2006-2008 Graduate Student (M.S.), Department of Pharmacy, College of Medicine, National Taiwan University, Taipei, Taiwan. (Advisor: Shoei-Sheng Lee, Ph.D.)
- 2008-2010 Research Assistantship, Department of Pharmacy, College of Medicine, National Taiwan University, Taipei, Taiwan. (Advisor: Shu-Wen Lin, Pharm.D.)
- 2011-2015 Graduate Student (Ph.D.), Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, Lexington, KY, USA. (Advisor: Jürgen Rohr, Ph.D.)

Awards:

- 2011-2014 The Scholarship of Government Sponsorship for Overseas Study from the Ministry of Education, Taiwan.

Publications:

1. **Chen, J. M.**, Hou, C., Wang, G., Tsodikov, O. V.*, Rohr, J.*, “Structural insight into MtmC, a bifunctional ketoreductase-methyltransferase involved in the assembly of the mithramycin trisaccharide chain”, *Biochemistry*, 2015, **54**, p.2481-2489.
2. Lassak, J., Keilhauer, E., Fürst, M., Wuichet, K., Gödeke, J., Starosta, A., **Chen, J. M.**, Søggaard-Andersen, L., Rohr, J., Wilson, D., Häussler, S., Mann, M., and Jung, K.*, “Arginine-rhamnosylation as novel strategy to activate translation elongation factor P”, *Nature Chemical Biology*, 2015, **11**, p.266-270
3. **Chen, J. M.**, Shepherd, M. D., Horn, J., Leggas, M., and Rohr, J.*, “Enzymatic methylation and structure-activity-relationship studies on polycarcin V, a gilvocarcin-type antitumor agent”, *ChemBioChem*, 2014, **15**, p.2729-2735
4. Scott, D., **Chen, J. M.**, Bae, Y., and Rohr, J.*, “Semi-synthetic mithramycin SA derivatives with improved anticancer activity”, *Chem. Biol. Drug Des.*, 2013, **81**, p.615-624
5. Lin, S. C., Lin, S. W., **Chen, J. M.**, and Kuo, C. H.*, “Using sweeping-micellar electrokinetic chromatography to determine voriconazole in patient plasma”, *Talanta*, 2010, **82**, p.653-659
6. Lam, S. H., **Chen, J. M.**, Kang, C. J., Chen, C. H., and Lee, S. S.*, “ α -Glucosidase inhibitors from the seeds of *Syagrus romanzoffiana*”, *Phytochemistry*, 2008, **69**, p.1173-1178

Patents:

1. Rohr, J., Scott, D., Leggas, M., **Chen, J. M.**, Tsodikov, O. V., **2015**, “Semi-synthetic mithramycin derivatives with anti-cancer activity”, U.S. Patent Application No. 14/717542, filed May 2015, Patent Pending.
2. Tsodikov, O. V., Rohr, J., Leggas, M., Hou, C., Scott, D., Horn, J., **Chen, J. M.**, Yu, X., Weidenbach, S., **2014**, “Mithramycin analogues and their use”, U.S. Patent Application No. 62/001516, filed May 2014, Patent Pending.
3. Rohr, J., Scott, D., **Chen, J. M.**, and Bae, Y., **2014**, “Semi-synthetic mithramycin derivatives with anti-cancer activity”, WO 2014093652 A1, filed Dec. 12th, **2013**, and issued Jun. 19th, **2014**.

Conference proceedings:

1. Eckenrode, J., Horn, J., **Chen, J. M.**, Rohr, J., Leggas, M.*, “Mithramycin analogs with reduced toxicity for EWS-FLI1 targeting”, *annual meeting of the American Association for Cancer Research (AACR)*, Philadelphia, PA, USA, **2015**
2. **Chen, J. M.**, Hou, C., Wang, G., Tsodikov, O. V.*, Rohr, J.*, “Structural investigations of two bifunctional co-dependent enzymes, MtmGIV and MtmC, involved in the biosynthesis of mithramycin”, *annual meeting of the American Society of Pharmacognosy (ASP)*, Oxford, MS, USA, **2014**
3. Rohr, J.*, Tsodikov, O., Leggas, M., **Chen, J. M.**, Yu, X., Hou, C., Weidenbach, S., Horn, J., “Aureolic acid anticancer drugs – history and future”, *annual meeting of the Society for Industrial Microbiology and Biotechnology (SIMB)*, St. Louis, MO, USA, **2014**
4. **Chen, J. M.**, Wang, G., Rohr, J.*, “The biosynthetic pathway to TDP-D-fucofuranose, the sugar donor substrate for the C-glycosyltransferase involved in gilvocarcin V biosynthesis”, *annual meeting of the American Society of Pharmacognosy (ASP)*, St. Louis, MO, USA, **2013**