

# Towards the production of tikitericin: A novel peptide from *Thermogemmatispora* strain T81

By

Benjamin Peter Baker

Supervisor: Dr Robert Keyzers



A thesis submitted  
to the Victoria University of  
Wellington in partial fulfilment of  
the requirements for the degree  
MSc in Chemistry

Victoria University of Wellington

2016

## Abstract

The utilisation of natural products for treatment of human ailments has been rooted in various cultures for centuries. Extraction of natural products has been essential for the discovery of new drugs and inspiration for synthetic analogues. Since the success of penicillin, microbial natural products have been of interest. Genome mining of *Thermogemmatispora* strain T81, a thermophile from the Taupo Volcanic Zone, found the potential for the production of novel ribosomally synthesised and post-translationally modified peptides (RiPPs). Previous work showed that T81 exhibited antimicrobial activity against a wide variety of extremophilic bacteria. Although the three thiopeptides encoded for in the genome of T81 have not been found, the lanthipeptide tikitericin has recently been isolated and described. Unfortunately tikitericin is produced in low quantities by T81 and bioactivity data has not yet been obtained. Because of its potential antimicrobial activity, different routes to produce it are of interest. The aim of this project was to synthesise tikitericin by solid phase peptide synthesis. MS imaging was also utilised to search for the presence of tikitericin as an antimicrobial agent *in situ*.

# Acknowledgements

*Stay far from timid*

*Only make moves when your hearts in it*

*And live the phrase “sky’s the limit”*

*Biggie Smalls, Sky’s the limit*

Thank you to my supervisor Dr Rob Keyzers. I appreciate that you stayed enthusiastic even during times when I was crestfallen. Without your advice and guidance I would have spent this project meandering aimlessly. Thank you also for all the time you spent proofreading my work. I enjoyed your humour during this project and all of our ridiculous email banter.

Thank you to Associate Professor Peter Northcote. Without you critiquing my presentation during group meetings my first year’s master’s talk would have been a disaster. Thank you also for sharing your incredible knowledge in the field of natural products and NMR, I’ve learnt so much from you.

I would also like to thank Associate Professor Paul Teesdale-Spittle and Dr Joanne Harvey for help and advice in synthetic matters.

Thank you to Emma Aitken, without you doing the work preceding this project I would not have been able to carry out the work described in this thesis. Thank you to Claire Turner for giving me a crash course in synthesis. Without your preliminary work, guidance and advice I would not have been able to do this project as well as I did. I would also like to thank Galen Eakens for taking the time out of his busy schedule to show me the ropes of Solid Phase Peptide Synthesis. It was much appreciated, especially because during that time you were writing up your Phd. Thank you for your patience, explanations and advice.

Next I would like to thank all past and present members of our lab group. Nathaniel Dasyam and Chriselle Braganza for being so hospitable during my summer research scholarships and introducing me to the lab and research. Sarah Moore for all of the help you were during my first year of masters, especially around helping me understand the class concepts. Thank you for your friendship. Also to Vidhiya Damodaran, Danica Carter, Tate Bicknell, Simon Berberich, Andy Woerle, Michael Göttel, Ethan Woolley, Lucile Berthomier, Sarah

Andreassand, Joe Bracegirdle and Phyllida Duncan. Also to the past and present members of the Marine Natural Products group – Dr. Jonathan Singh, Helen Woolner, Taitusi Taufu and Saweon Hong. All of you have made this an unforgettable, fun experience.

I am very grateful to Matthew Stott from GNS Science for all of the work that he has contributed to this and the preceding thesis. Thank you also to his colleague Hannah Peach for providing me with the microbial cultures for IMS. Also, a big thank you to Dr Joanna MacKichan for running bioassays and to Associate Professor Martyn Coles for attempting to acquire a crystal structure. I would like to express gratitude towards Sven Sondhauss for all of the help that you offered towards IMS. Thank you also for putting up with all my jokes about laser beams.

To my fellow postgraduate students, thank you for the memorable times and help whenever I needed advice or just someone to vent to. To Valentine Chan, thank you for helping me initially with Word and Chemdraw. To Vaughan Trounson, without our ridiculously competitive relationship I probably would not have made it this far. To John Park, thank you for all of our hilarious chats at lunch time, it helped me to relax and get my mind off work.

Special thanks to Ian Vorster, you're the real MVP. Without your work towards maintaining the NMRs the whole school would fall apart. Thank you for always being on call to help out.

I would like to thank my brothers and sisters in Christ at Karori Baptist church and Studentlife for an amazing period of growing and walking in the Lord. I am grateful for God's provision of joys, challenges, and grace for growth.

Finally, I would like to thank my family. Thank you to Aunty Maria for sending financial support and encouragement. Thank you to Nathanael and Lisa Dei gratia for providing me with encouragement and prayers. To Manaaki for being the little dude that you are. To Josh and Sarah Baker for taking an interest in my work and encouraging me along the way. Mum, you've been amazing, thank you so much for your understanding when I had to work late. I sincerely appreciate all the wonderful things you do for me, I couldn't have undertaken this without you.

# Table of Contents

Abstract.....	i
Acknowledgements.....	ii
List of Schemes.....	vii
List of Figures .....	ix
Glossary.....	x
Chapter 1.....	1
Introduction .....	1
1.1 Natural products as viable drug candidates .....	1
1.2 Exploration of extremophile natural products .....	5
1.2.1 Untapped potential of microbes.....	5
1.2.2 Extremophiles .....	5
1.2.3 Secondary metabolites from extremophilic organism .....	6
1.2.4 Extremophiles from the Taupo Volcanic Zone .....	7
1.3 Ribosomally produced and posttranslationally modified peptides .....	8
Chapter 2.....	10
Lanthipeptides .....	10
2.1 Introduction .....	10
2.1.1 Biosynthesis .....	11
2.2 Tikitericin: A novel peptide from <i>Thermogemmatispora</i> strain T81 .....	15
2.3 Starting material synthesis.....	16
2.3.1 Developments in synthetic approaches.....	16
2.4 Beginning of the synthesis .....	21
2.4.1 Allylation .....	21
2.4.2 Mesylation .....	22
2.5 Aziridines.....	23
2.5.1 Aziridine formation .....	23
2.5.2 Bonding in aziridines .....	24
2.5.3 Ring-opening processes .....	25
2.5.4 <i>N</i> -substituent effect .....	26
2.5.6 Changing of the aziridine substituent .....	28
2.5.7 Ring opening reaction .....	30

2.6 Troubleshooting of the Deprotection/reprotection of (Me)Lan .....	31
2.6.1 Deprotection/reprotection of (Me)Lan .....	31
2.6.2 Investigation into the effects of DIPEA .....	32
2.6.3 Investigations into chromatographic method and DN's deprotection .....	33
2.7 Concluding remarks .....	37
Chapter 3.....	38
Solid phase peptide synthesis.....	38
3.1 Introduction .....	38
3.2 Historical assessment of peptide synthesis .....	38
3.2.1 Theodor Curtius' contributions to peptide chemistry .....	39
3.2.2 Emil Fischer's contributions to peptide chemistry .....	40
3.2.3 Advancements in protecting group strategies.....	41
3.2.4 Advances in coupling methods .....	43
3.2.5 The development of SPPS .....	46
3.2.5.1 "Quasi-continuous flow" .....	49
3.2.5.2 Fully automated SPPS .....	49
3.3 The attempted total synthesis of tikitericin.....	49
3.3.1 Instruments used for synthesis.....	49
3.3.2 General methodologies and materials of synthesis.....	50
3.3.2.1 Fmoc strategy.....	50
3.3.2.2 Wang resin .....	51
3.3.2.3 Fmoc deprotection.....	51
3.3.2.4 Coupling reagents and protocol.....	52
3.3.2.5 Washing protocol.....	53
3.3.2.6 Kaiser test.....	53
3.3.3 Synthetic steps towards the production of tikitericin .....	54
3.4 Analysis of the product .....	62
3.5 Concluding remarks .....	64
Chapter 4.....	66
Thiopeptides .....	66
4.1 Introduction .....	66
4.2 Thiopeptides produced by T81 .....	68
4.3 MALDI mass spectrometry.....	69
4.3.1 MALDI-based imaging mass spectrometry (IMS).....	70

4.3.2 Mass spectrometry imaging to screen for metabolites from T81 .....	73
4.3.3 Preliminary results of microbial IMS of T81 versus <i>Geobacillus</i> .....	76
4.4 Concluding remarks .....	80
Chapter 5.....	81
Concluding remarks and future directions .....	81
Chapter 6.....	82
Experimental .....	82
6.1 General experimental .....	82
6.2 Synthesis of starting materials.....	83
6.2.1 Synthesis of Lan intermediates and starting material .....	83
6.2.2 Synthesis of MeLan intermediates and starting material.....	87
6.3 SPPS procedures .....	91
6.3.1 General procedure for Fmoc SPPS.....	91
6.3.2 Kaiser test procedure.....	92
6.3.4 Addition of (Me)Lan .....	92
6.3.5 Deprotection of allyl, Fmoc groups and cyclisation for ring formation .....	92
6.3.6 Final cleavage.....	92
6.4 Microbial imaging mass spectrometry.....	93
Appendix A - Microbial culture media components .....	94
Appendix B – SPPS checklist.....	95
Appendix C – Spectra of Lan intermediates.....	100
Appendix D – Spectra of MeLan intermediates .....	119
References .....	138

# List of Schemes

Scheme 2.1 Biosynthetic pathway to lanthipeptide (Me)Lan residues .....	12
Scheme 2.2 Lan synthesis by Michael addition .....	17
Scheme 2.3 Asymmetric lanthionines by desulfurisation.....	17
Scheme 2.4 Vederas' Lan synthesis via a lactone .....	18
Scheme 2.5 Potential lactone ring opening products.....	19
Scheme 2.6 Synthesis of protected Lan derivatives through different R groups .....	19
Scheme 2.7 (Me)Lan synthesis by aziridine ring opening.....	20
Scheme 2.8 Synthesis of orthogonally protected (Me)Lan.....	21
Scheme 2.9 Reaction mechanism of allylation .....	22
Scheme 2.10 Reaction mechanism of mesylation .....	23
Scheme 2.11 Gabriel synthesis .....	23
Scheme 2.12 Reaction mechanism of aziridine formation .....	24
Scheme 2.13 Regioselectivity in ring opening reactions.....	26
Scheme 2.14 Resonance isomers of substituted aziridines .....	27
Scheme 2.15 Postulated reaction mechanism of tritylated aziridine with cysteine .....	28
Scheme 2.16 Reaction mechanism of detritylation and reprotection.....	29
Scheme 2.17 Reaction mechanism of aziridine ring opening .....	30
Scheme 2.18 Reaction mechanism of DN's deprotection/alloc reprotect .....	32
Scheme 3.1 Peptide bond formation.....	38
Scheme 3.2 Emil Fischer's preparation of glycylglycine .....	39
Scheme 3.3 Theodor Curtius' preparation of benzoylglycylglycine.....	39
Scheme 3.4 Theodor Curtius' azide coupling method .....	40
Scheme 3.5 Fischer's acid halide peptide coupling .....	41
Scheme 3.6 Cbz deprotection routes.....	42
Scheme 3.7 Boc deprotection .....	43
Scheme 3.8 Carbodiimide coupling mechanism and racemisation .....	44
Scheme 3.9 Deprotection of Fmoc using piperidine.....	51
Scheme 3.10 Coupling mechanism with DIC and HOBt.....	52
Scheme 3.11 Kaiser test reaction mechanism .....	54
Scheme 3.12 Mechanism of palladium catalysed deallylation .....	56
Scheme 3.13 Solid phase synthesis for forming of ring D.....	58



Scheme 3.14 Solid phase synthesis for forming of ring C.....	59
Scheme 3.15 Solid phase synthesis for forming of ring B.....	60
Scheme 3.16 Solid phase synthesis for forming of ring A.....	61
Scheme 4.1 Proposed biosynthesis of pyridine ring in thiopeptides from two Dha residues <sup>51</sup> .....	68

# List of Figures

Figure 1.1 General RiPP biosynthetic pathway .....	9
Figure 2.1 Bonding in aziridines .....	25
Figure 2.2 Lewis acid interactions .....	26
Figure 2.3 Aziridine activating groups.....	27
Figure 2.4 Cyclic loading protocol .....	34
Figure 2.5 $^1\text{H}$ NMR of 70% fraction (500 MHz, $\text{CDCl}_3$ ).....	35
Figure 2.6 $^1\text{H}$ NMR of 50% fraction (cyan) and starting material (red) (500 MHz, $\text{CDCl}_3$ ) .....	36
Figure 3.2 Common coupling reagents .....	45
Figure 3.1 Mechanical shaker and reaction vessel .....	50
Figure 3.2 Outline of purpose of each protecting group .....	55
Figure 3.3 Principals of SPPS .....	56
Figure 3.4 Duct tape method .....	57
Figure 3.5 $^1\text{H}$ NMR of precipitate (600 MHz, $\text{CD}_3\text{OD}$ ) .....	63
Figure 3.6 COSY NMR of precipitate (600 MHz, $\text{CD}_3\text{OD}$ ).....	64
Figure 4.1 The structural series of thiopeptides .....	67
Figure 4.2 Schematic representation of the MALDI desorption and ionisation process, adapted from Magnet Lab <sup>136</sup> .....	70
Figure 4.3 Steps for IMS adapted from Dorrestein et al. <sup>140</sup> Reproduced with permission of the American Society for Microbiology©, 2016 .....	72
Figure 4.4 T81 (top left), <i>Geobaccillus</i> (top right) and co-culture (bottom).....	74
Figure 4.5 MALDI Target plates Left co-culture, right T81 (top) and <i>Geobaccillus</i> (bottom) .....	75
Figure 4.6 Application of matrix to culture media sections.....	75
Figure 4.7 Target plate positioned in plate holder .....	76
Figure 4.8 Global scan mass spectrum of IMS (900 – 4000 $m/z$ ).....	77
Figure 4.9 Global distribution of selected ions .....	78
Figure 4.10 Calculated isotopic distribution of 23 .....	79

# Glossary

<b><math>\delta</math></b>	Chemical shift (ppm)
<b><math>^1\text{H}</math> NMR</b>	Proton nuclear magnetic resonance
<b><math>^{13}\text{C}</math> NMR</b>	Carbon nuclear magnetic resonance
<b>Alloc</b>	allyloxycarbonyl group
<b><math>\text{CD}_3\text{OD}</math></b>	Deuterated methanol
<b><math>\text{CDCl}_3</math></b>	Deuterated chloroform
<b>COSY</b>	Correlation spectroscopy ( $^1\text{H}$ to $^1\text{H}$ )
<b>d</b>	doublet
<b>dd</b>	doublet of doublets
<b>ddd</b>	doublet of doublet of doublets
<b>DMSO</b>	Dimethyl sulfoxide
<b>DNA</b>	Dioxyribonucleic acid
<b>DNs</b>	2,4-dinitrobenzenesulfonyl group
<b>EtOH</b>	Ethanol
<b>Fmoc</b>	Fluorenylmethyloxycarbonyl group
<b>HRESIMS</b>	High-resolution electrospray ionisation mass spectrometry
<b>HP20</b>	Poly(styrene-divinylbenzene) stationary support
<b>IMS</b>	Imaging Mass Spectrometry
<b><i>in situ</i></b>	On the site
<b><i>in vacuo</i></b>	In a vacuum
<b>IR</b>	Infrared
<b>Lan</b>	Lanthionine
<b><i>m/z</i></b>	mass to charge ratio
<b>MALDI</b>	Matrix-assisted laser desorption/ionization
<b>MeLan</b>	MethylLanthionine
<b>(Me)Lan</b>	MethylLanthionine/Lanthionine
<b>MS</b>	Mass spectrometry
<b>NMR</b>	Nuclear magnetic resonance
<b>NRPS</b>	Non-ribosomal peptide synthetase
<b>PTM</b>	Post translational modification
<b>RIPP</b>	Ribosomally-synthesised and post-translationally modified peptide
<b>RNA</b>	Ribonucleic acid

<b>td</b>	Triplet of doublets
<b>TOF</b>	Time of flight
<b>TVZ</b>	Taupo Volcanic zone

# Chapter 1

## Introduction

### 1.1 Natural products as viable drug candidates

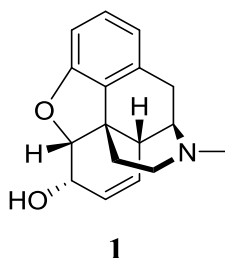
Humans, as a species, has succeeded in its path of evolution through the mastery and utilisation of the wonders that are presented to us rooted in nature. One of the many things that has helped us succeed is the use of the medicinal properties from natural sources. From ancient times, herbs were a prominent source of medicine. Hippocrates, the father of medicine, said “Let food be your medicine and medicine be your food”.<sup>1</sup> The Japanese phrase “ishoku dougen”, which translates to “Food and medicine are of the same origin” is quite appropriate to describe this relationship as we find that the traditional condiments and vegetables used in many cultures have antibacterial and antifungal activity.<sup>2</sup>

The use of medicinal plants was developed in many separate cultures around the world. The earliest recorded evidence of herbal medicine was cuneiform tablets from Mesopotamia dating back to 2200 BC.<sup>3</sup> Recorded are approximately 1000 plant-derived substances including liquorice, poppy juice and myrrh, which are still used today in the alleviation of coughs and colds.<sup>4</sup>

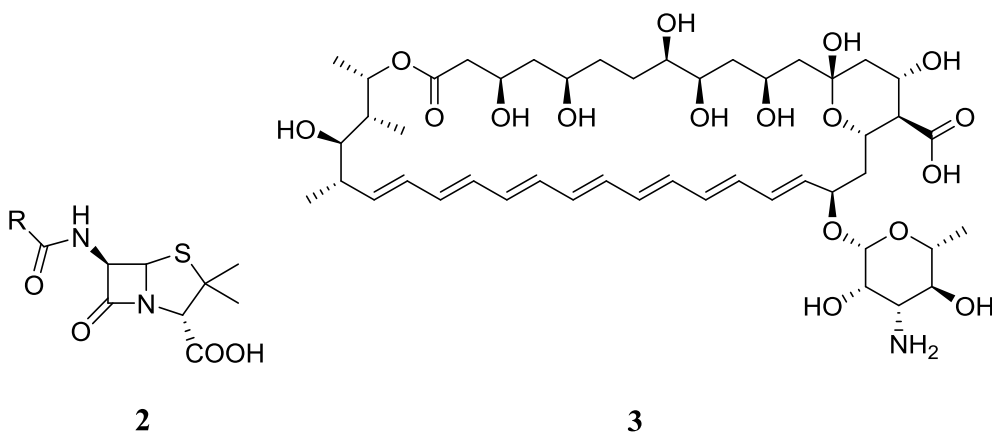
The Ancient Egyptians were, for their time, astute in medical knowledge. While Egyptian medicine started around 2900 BC the best documented example of Egyptian medicine was recorded on the Ebers papyrus dating from 1500 BC.<sup>5</sup> This papyrus described the use of over 700 drugs in which animal products, herbs and human products were utilized. From this, we know that their medical system had a vast amount of prescriptions, surgery and advanced embalming procedures.<sup>6</sup> Juniper berry oil was widely used in their prescriptions and also the embalming process. An analysis of juniper berry oil has found a wide range of terpenoids. These were tested for antibacterial and antifungal activity and were found to be active in both

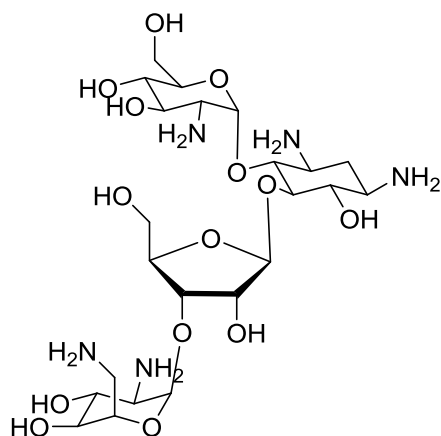
cases.<sup>7</sup> Combined with many other components, this would have been a reason for their successful embalming processes.

The understanding around how various cultures used different plants in their own medicine has contributed towards the field of natural products and drug discovery, as around 120 pharmaceutical products in the last century have been derived from traditional knowledge.<sup>8</sup> One of these is the opioid, morphine. *Papaver somniferum*, the opium poppy, was used from ancient Greece to China as a pain killer.<sup>9</sup> Trade with China brought the poppy to Europe, which led to the isolation of morphine (**1**) by F.W Sertürner in 1815.<sup>10</sup> This was the first pure natural compound isolated from a plant, and due to its success as a commercialised drug, inspired an interest to identify active compounds from other natural remedies.

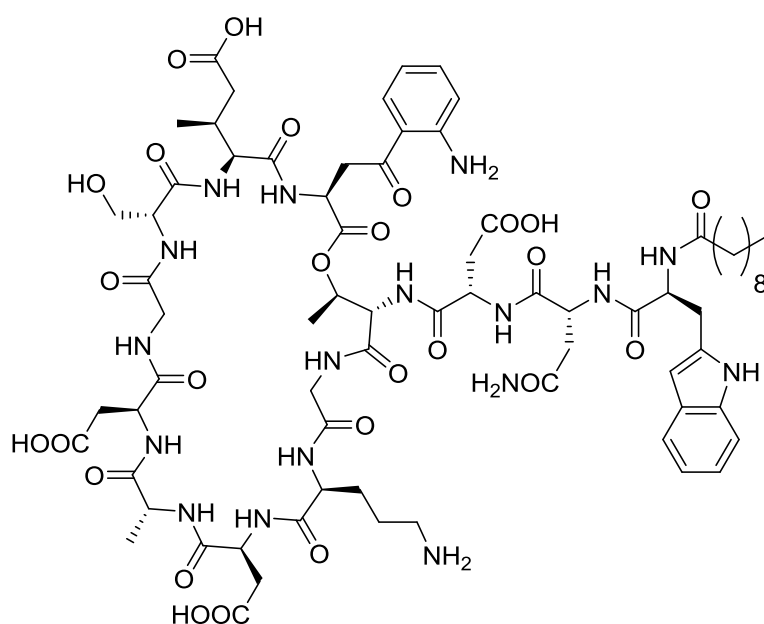


In 1929 Fleming discovered penicillin (**2**), an antibiotic from *Penicillium* mould.<sup>11</sup> This was the first microbial derived, pharmacologically active drug. After this, interest was focussed on microbes as a rich source of bioactives. This led to the discovery of many other pharmacologically active drugs. Examples of these are the antifungal amphotericin B (**3**) from *Streptomyces nodosus*,<sup>12</sup> the antiparasitic paromomycin (**4**) from *Streptomyces krestomucetus*<sup>13</sup> and the antibacterial daptomycin (**5**) from *Streptomyces roseosporus*.<sup>14</sup>





4



5

Compounds produced by organisms can be segregated into two categories, primary and secondary metabolites. Primary metabolites are compounds produced by the organism that are essential for its overall survival. These include amino acids, carbohydrates, lipids and nucleotides.<sup>15</sup> Secondary metabolites are compounds made by organisms that are generally below 3000 Da and are not essential for the overall survival of the organism.<sup>16</sup> Unlike primary metabolites, which follow a relatively universal distribution across all species, secondary metabolites are taxonomically idiosyncratic. Secondary metabolites are essential to give the producing organism a competitive edge, as there is a general consensus that they are produced as a defence against other organisms around them.<sup>17</sup> As secondary metabolites are active against other organisms, this shows that they are able to bind to biological targets and are valuable assets in the exploration for new drugs.

Historically, natural products chemistry has been at the forefront of drug discovery.<sup>18</sup> A recent review by Newman and Cragg studied the influence of natural products on drug discovery.<sup>19</sup> Within the period of 1981 to 2010 there has been 1355 drugs approved by the United States Food and Drug Administration (USFDA). Of these, only 36% were of purely synthetic origin. While only 6% were true secondary metabolites, the other 58% were either derived from natural products, with some synthetic modifications or were fully synthetic but the pharmacophore was of secondary metabolite origin. This analysis demonstrates the importance of natural products research in relation to drug discovery.

Although the influence of secondary metabolites on drug discovery is prominent, this focus faces various challenges, particularly from the pharmaceutical economic environment.<sup>20</sup> Due to competition between different pharmaceutical companies, patent law and investor interests, pharmaceutical companies have focused on maximising profits over research and development. Natural product research essentially ceased in pharmaceutical companies in the early 1990's.<sup>21</sup> This can be attributed to other techniques of discovering drug leads that register activity against biological targets. First, the development of high throughput screening against known biological targets saw the move from natural product extract libraries to synthetic libraries. Second, there was also the development of combinatorial chemistry, which allows chemists to take a known pharmacophore as a starting point and iteratively modify it in a modular synthetic fashion. However, it was established that these synthetic methods were not instilling the same magnitude of chemical diversity that natural products possess.<sup>22</sup> For example, these synthetic bioactives were often lacking multiple chiral centres, heterocyclic substituents or polycyclic units as found in most commercial drugs. This has seen a move from combinatorial chemistry to diversity-oriented synthesis (DOS).<sup>23</sup> DOS describes the synthesis of structurally diverse small molecules that mimic natural products.

However, this does not address the overall problem of drug resistance evolving in bacteria.<sup>24</sup> As all of these methods only work on developing more potent drugs for a specific target, generally a protein, the bacteria is able to evolve resistance to survive the drug. This is why natural product research is important, as it allows us to find new compounds that have different structures and modes of action to combat drug resistant strains of bacteria, and also find drugs which are able to treat currently untreatable diseases.



## **1.2 Exploration of extremophile natural products**

### **1.2.1 Untapped potential of microbes**

Microbes have had billions of years to evolve on this planet.<sup>25</sup> As a result of this, there are more microbial species than vertebrates, invertebrates and plants combined.<sup>26</sup> Although there is a large diversity of microbial organisms, it is estimated that less than 1% of microbes have been cultured in pure liquid media.<sup>27</sup> Pure cultures of microbes are essential to locate secondary metabolites produced by these microbes. Genomic sequencing technologies have advanced a long way, and because of this it has become easy to rapidly sequence a whole microbial genome. This, coupled with the fact that the majority of microbes have small genomes makes genome sequencing a powerful technique for discovering novel microbial natural products. As the gene cassettes of polyketide synthases, ribosomally produced and post translationally modified peptides (RiPP) and nonribosomal peptide (NRP) synthetases (NRPSs) have been studied, sequence homology can be used to identify biosynthetic potential for natural product production with subsequent culturing and isolation allowing these secondary metabolites to be screened from the pure liquid culture. Of course, conventional culturing techniques can only be used to culture a small minority of microbes, so novel culturing methods have to be produced to mimic the environment for successful cultivation.<sup>28</sup>

### **1.2.2 Extremophiles**

Etymologically, the word extremophile comes from the Latin *extremus* which is the superlative of *exter*, meaning ‘being on the outside’. *Phile* comes from the Greek *philos*, meaning love. Thus, when MacElroy first described these organisms in ‘Some comments on the evolution of extremophiles’ he coined this term to describe any organisms that are able to grow and thrive in a hostile environment.<sup>29</sup> Generally bacteria come to mind when the word ‘extremophile’ is used, however extremophiles are found in all three domains of life; (Archaea, Bacteria and Eukaryota). An extreme example of an extremophile is the tardigrade, a micro-animal which can survive the vacuum of outer space, temperatures close to absolute zero and above waters boiling point, large doses of ionizing radiation and very high pressures.<sup>30</sup>

A “normal” environment has to be defined as any environment in which an organism can grow and thrive is “normal” for them and anything in which they would not survive would be considered hostile. Thus, the term extremophile is purely anthropocentric and is defined around conventional growth conditions which is around the temperature of 37 °C, pH 7.4, salinity from

0.9 to 3.0% and 1 atm pressure.<sup>31</sup> Table 1.1 describes the different classes of extremophiles. An extremophile which can survive multiple extreme environments is called a polyextremophile, for example an organism living in an acidic hot pool would be considered both an acidophile and a thermophile.

Type	Definition
Acidophile	An organism with optimum growth at $\text{pH} \leq 3\text{--}4$
Alkaliphile	An organism with optimum growth at $\text{pH} \geq 8$
Halophile	An organism with optimum growth at 2 – 5 M NaCl
Hyperthermophile	An organism with optimum growth $< 80\text{ }^{\circ}\text{C}$
Metalotolerant	An organism which can withstand high concentrations of heavy metal
Piezophile	An organism with optimum growth at high pressure
Thermophile	An organism with optimum growth at 60 – 80 $^{\circ}\text{C}$
Xerophile	An organism which can grow and reproduce in conditions with a low availability to water

Table 1.1 Classification of extremophilic organisms<sup>32</sup>

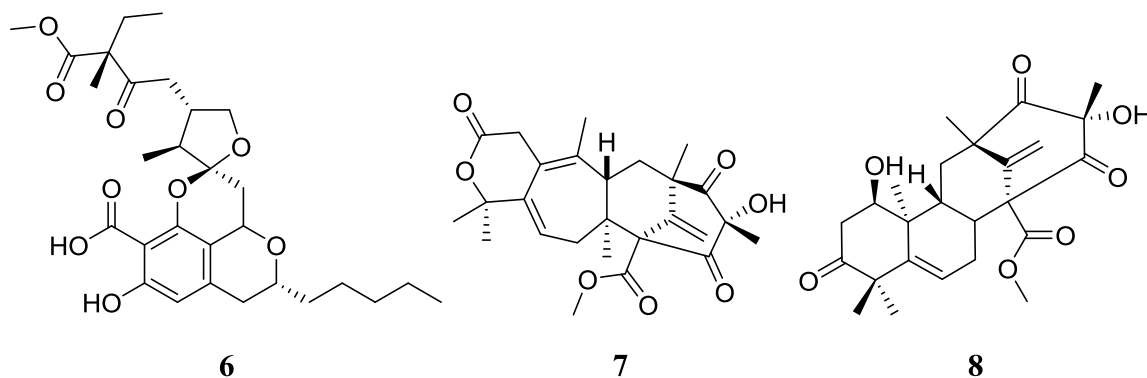
### 1.2.3 Secondary metabolites from extremophilic organism

Extremophiles have to live in environments that are considered abiological and because of this, have evolved to produce biomolecules that help them to maintain metabolic flux.<sup>33</sup> Thus, extremophiles have proven useful in many applications. For example, the use of thermostable enzymes from thermophilic bacteria revolutionised the field of biotechnology. Notably, taq polymerase, a thermal stable DNA polymerase from *Thermus aquaticus*, makes the polymerase chain reaction feasible.<sup>34, 35</sup> Although enzymes from extremophiles have been studied in depth, not much work has been done towards the isolation of secondary metabolites from extremophiles.

Bioprospecting is essentially the systematic exploration of biodiversity to discover new genes, enzymes, microorganisms and bioactive compounds. The secondary metabolites produced by extremophiles, are called extremolites. Bioprospecting of extremophiles has proved successful to find extremolites as described in the review by Wilson and Brimble outlining the finding of extremolites up to 2008.<sup>36</sup>

As an example, Berkeley Pit Lake is a 540 m deep abandoned open pit copper mine. The pit filled with water after the mine was shut, this led to the leaching of heavy metals from the rocks, meaning that the water has a high concentration of metal sulphites (Fe, Cu, Al, Cd, and Zn). The water is also highly acidic, with a pH of 2.5, due to the sulfuric acid used in the mining

process. Due to this harsh environment, it was previously thought that this environment would not be able to support life. However, bioprospecting by Stierle and Stierle has identified over forty different bacteria, protists, fungi, algae and protozoans. A significant fungus found was *Penicillium rubrum*. From this organism they isolated and identified berkelic acid (**6**), a novel spiroketal that exhibits selective activity against ovarian cell line OVCAR-1 with a GI<sub>50</sub> of 91 nM. They also identified berkeleydione (**7**) and berkeleytrione (**8**) that both show inhibitory action against non-small cell lung cancer.<sup>37</sup>



#### 1.2.4 Extremophiles from the Taupo Volcanic Zone

New Zealand sits on the boundary of the Pacific and Indo-Australian plates. This gives New Zealand a wide range of interesting geographical features. The subduction of the Pacific plate under the Indo-Australian plate gives rise to an incredible array of geothermal features in the Taupo Volcanic Zone (TVZ). This includes mud pools, geysers, sulfur pools, geothermally heated soils and hot springs. These features, although great tourist attractions, are harsh environments. These geothermal areas are characterised by having high temperatures, high or low pH, high heavy metal content and high pressures. Thus, any organism living in these environments would be classified as extremophilic.

As previously mentioned, only about 1% of microbes can be cultured by conventional methods, so attempting to culture microbes using traditional techniques and not finding anything is not a very good measure of whether there is anything living in that environment. Biologists have devised another way to determine whether there is life in an extreme environment prior to culturing them. 16S Ribosomal RNA (16S rRNA) is part of the greater ribosomal unit of prokaryotes. It is useful in phylogenetic studies as it is highly conserved within organisms of the same genus and species, but differs between organisms of different genera and species.<sup>38</sup> Thus, bacteria within the same genus will have conserved sequences of 16S rRNA and can be categorised in this way. The sequences are then stored in a database and a sequence of interest

can be compared to the database to find sequence homology and therefore categorise it taxonomically. The Ribosomal Database Project (RDP) currently contains 3,224,600 16S rRNA sequences.<sup>39</sup>

Although geothermal areas are considered as harsh environments, they are also rich in non-organic nutrients such as H<sub>2</sub>, NH<sub>3</sub>, reduced metals and sulfides. There is also vegetation, like moss and ferns, around the geothermal areas that provide decaying material for organotrophic bacteria. This suggests that they could be prime areas for growth of microbes. Stott and coworkers have carried out bioprospecting via analysis of 16S rRNA of soils affected by geothermal steam at Waikite, Tikitere and Ngauruhoe in the TVZ. Phylogenetic analyses of these areas revealed that the composition of bacteria was both rich and diverse. It was dominated with previously uncultured bacteria of the *Actinobacteria*, *Acidobacteria*, *Chloroflexi* and *Proteobacteria* phyla.<sup>40</sup> The majority of these bacteria were also less than 90% similar to any known cultured species.<sup>41</sup>

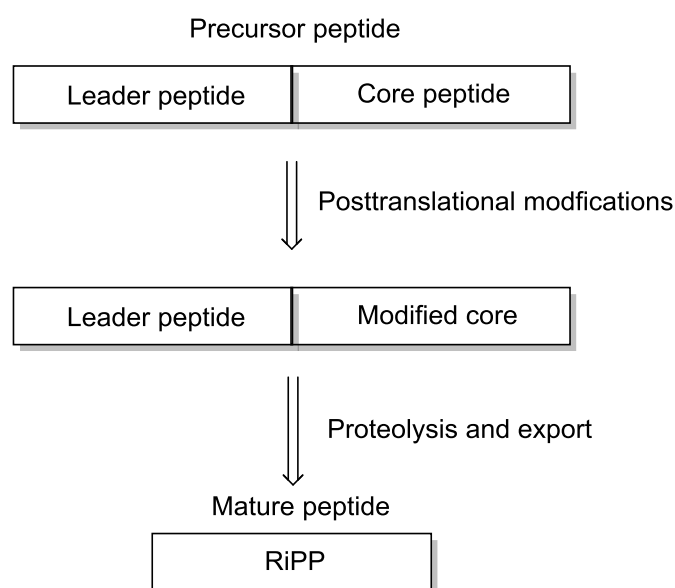
*Thermogemmatispora* strain T81 was isolated from geothermally heated soils in Tikitere. *Thermogemmatispora* is a recently described genus from the phylum *Chloroflexi*. Apart from strain T81, only three other species have been described; *T. onikobensis*, *T. foliorum* and *T. carboxidivorans*.<sup>42,43</sup> T81 is a Gram positive, thermophilic, acidophilic, aerobic, spore forming bacteria. It is able to grow at temperatures between 40 – 70 °C and has optimum growth at 50–60 °C. It is able to grow at pH of 3.1 – 7.1 and has an optimum growth at 4.6 – 6.0.<sup>44</sup> The genome of T81 was sequenced and it was found that it possessed biosynthetic machinery to produce several RiPPs; one lanthipeptide and up to three thiopeptides. Hauch and Stott, through co-culturing experiments, found that T81 could inhibit the growth of other extremophilic bacteria. It has not been determined yet if the RiPPs are responsible for the zone of inhibition generated by T81.

### **1.3 Ribosomally produced and posttranslationally modified peptides**

There are many classes of natural products, the most prevalent being the terpenoids, alkaloids, NRPs and polyketides. Advances in technology have brought a greater ability to sequence genomes and found that there is also a significant number of RiPPs. Defining a peptide as ribosomally produced may sound inconsequential, as peptides are generally ribosomally produced, however as previously mentioned there are also non-ribosomal peptide natural products as common bioactive compounds. The difference between RiPPs and NRPs are quite significant. To synthesise NRPs, their NRPS, which are gene products (enzymes), have to be

synthesised first. The NRPSs are similar to the polyketide synthetases, as in there is a large constituent of enzyme units that form an assembly line. As product is passed down, amino acids are added and modified. The NRPSs, unlike ribosomes, can incorporate non-canonical amino acids into the structure of the peptide. RiPPs on the other hand are limited to the conventional 20 amino acids that ribosomes have access to. Although this may make it seem like the structural diversity is lowered, RiPPs can be heavily post-translationally modified leading to structures resembling NRPs.<sup>45</sup>

The classes of RiPPs, although diverse, share a generic biosynthetic pathway (Figure 1.1). The gene sequence for the precursor peptide is generally clustered with the enzymes responsible for the posttranslational modifications. The gene is transcribed and translated, producing the precursor peptide. The two important components of the precursor peptide are the leader and core peptides. The leader peptide provides recognition sequences for the enzymes that modify the core peptide. Once all modifications are done, the core peptide is now a modified core, after which the leader peptide is cleaved and the modified core is left as the RiPP.<sup>46</sup>



**Figure 1.1 General RiPP biosynthetic pathway**

There are ten distinct compound classes of RiPPs; lanthipeptides, linaridins, proteusins, linear azol(in)e peptides, cyanobactins, thiopeptides, bottromycins, microcins, lasso peptides and microviridins, however only the lanthipeptides (Chapter 2 and 3) and thiopeptides (Chapter 4) will be described here.

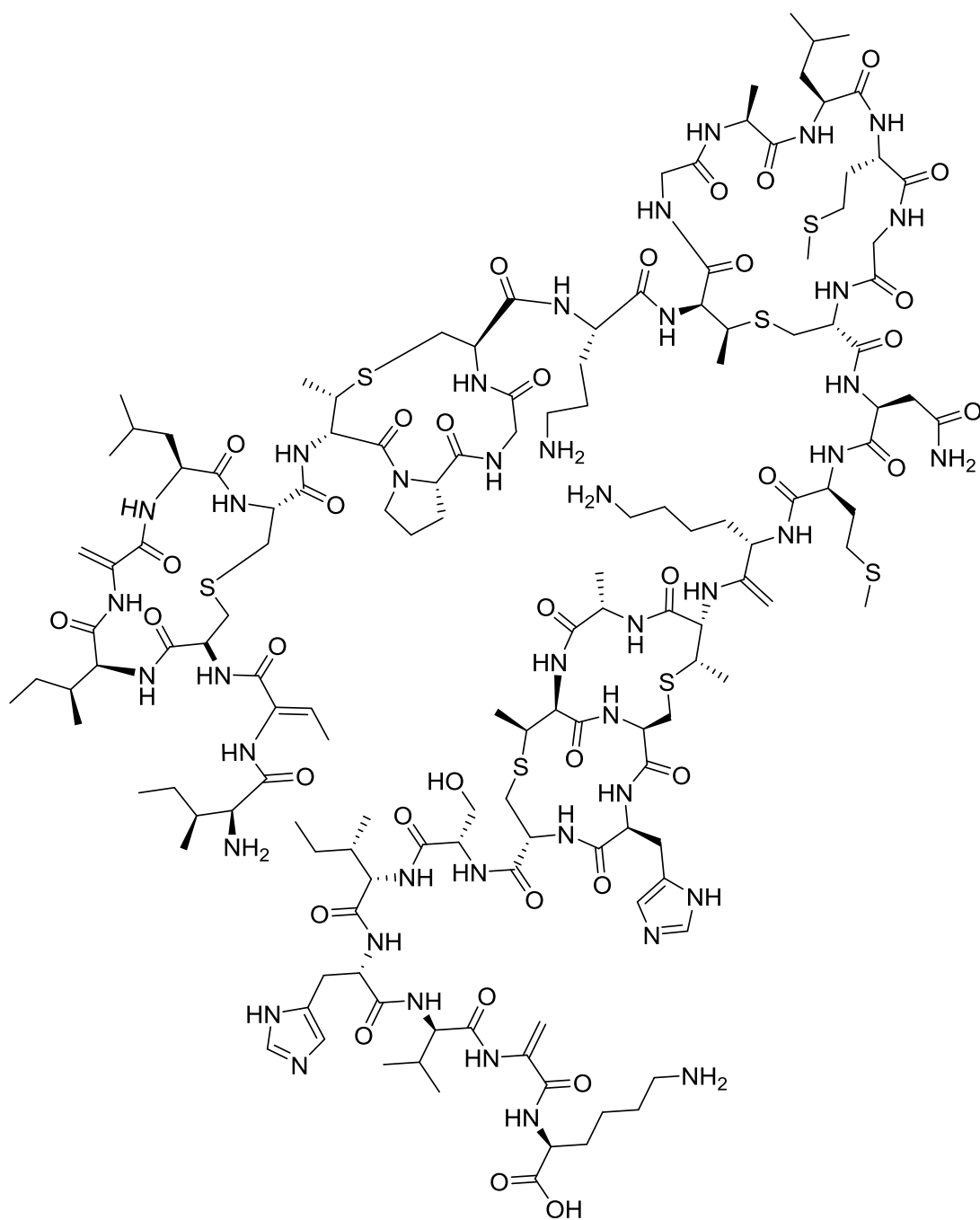
## Chapter 2

### Lanthipeptides

#### 2.1 Introduction

Lanthipeptides are a group of polycyclic RiPPs that are defined by their intramolecular thioether linked amino acid dimers, either lanthionine (Lan) or methyllanthionine (MeLan). A lanthipeptide that exhibits antimicrobial activity is called a lantibiotic. Lantibiotics are known to target small molecules rather than macromolecules. The thioether plays a role in this, as the polycyclic ring system which is installed from the modification of the Lan and MeLan residues imposes a constraint on the structural flexibility of the peptide. This structural constraint helps to increase the affinity for its specific target and also imparts chemical stability.<sup>47</sup>

The prototypic example of a lantibiotic is nisin A (**9**). Nisin A, first described in 1928, has a long history of being used as a safe and effective food preservative. It is produced by *Lactococcus lactis* and inhibits a large range of Gram positive bacteria. The mode of action of nisin A, and the other peptides in its subgroup, is due to the amphiphilic nature of the peptide. Studies with micelles that mimic the cell membrane of bacteria found that nisin can incorporate into the membrane by association of its hydrophilic groups with cellular phospholipid head groups, and its hydrophobic side chains are immersed in the hydrophobic core of the membrane.<sup>48</sup> Nisin docks by binding to lipid II, a small molecule involved in the synthesis of the cell wall in bacteria, and inserts into the membrane forming a pore. This causes a rapid efflux of ions that leads to the depolarisation of the cytoplasmic membrane, disrupting the cellular processes of the bacteria and killing it.<sup>49</sup>

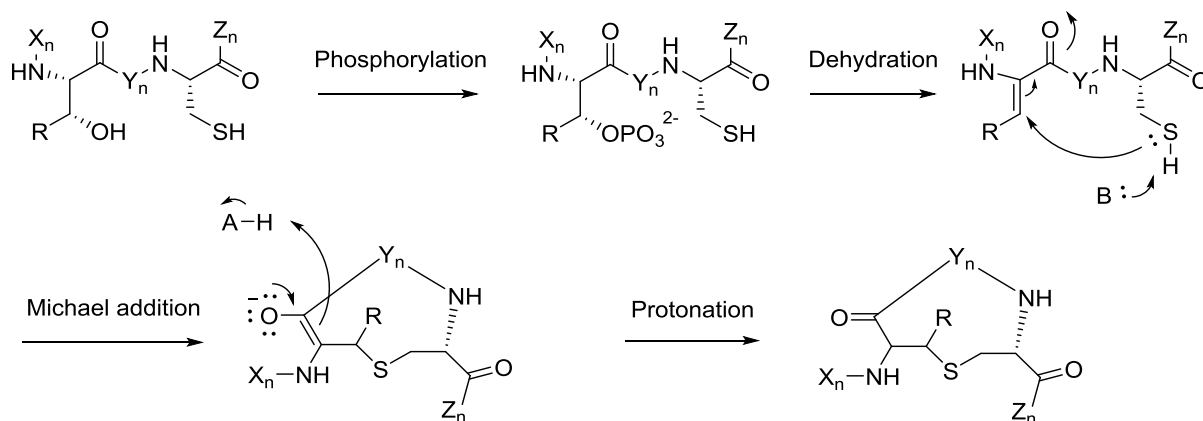


9

### 2.1.1 Biosynthesis

Lan and MeLan residues are not canonical amino acids, therefore they must be installed into the structure during post-translational modifications. Lan groups are a product of a cysteine and a serine while MeLan is made from a cysteine and threonine. The first step in the biosynthesis of (Me)Lan's is the phosphorylation of Ser/Thr producing an activated leaving

group and subsequently the phosphate group is eliminated, forming either dehydroalanine (Dha) or dehydrobutyryne (Dhb). The dehydro amino acids are then involved in a stereospecific Michael addition with the thiol on the cysteine residue, which forms the lanthionine ring. An enolate is formed in the cyclisation, and this goes through an enol-keto tautomerisation to form the final product. (Scheme 2.1).<sup>47</sup> There is potential for two diastereomers of the Lan and four of the MeLan to form in this reaction, although only one of each is generally found in the finalised product.



### Scheme 2.1 Biosynthetic pathway to lanthipeptide (Me)Lan residues

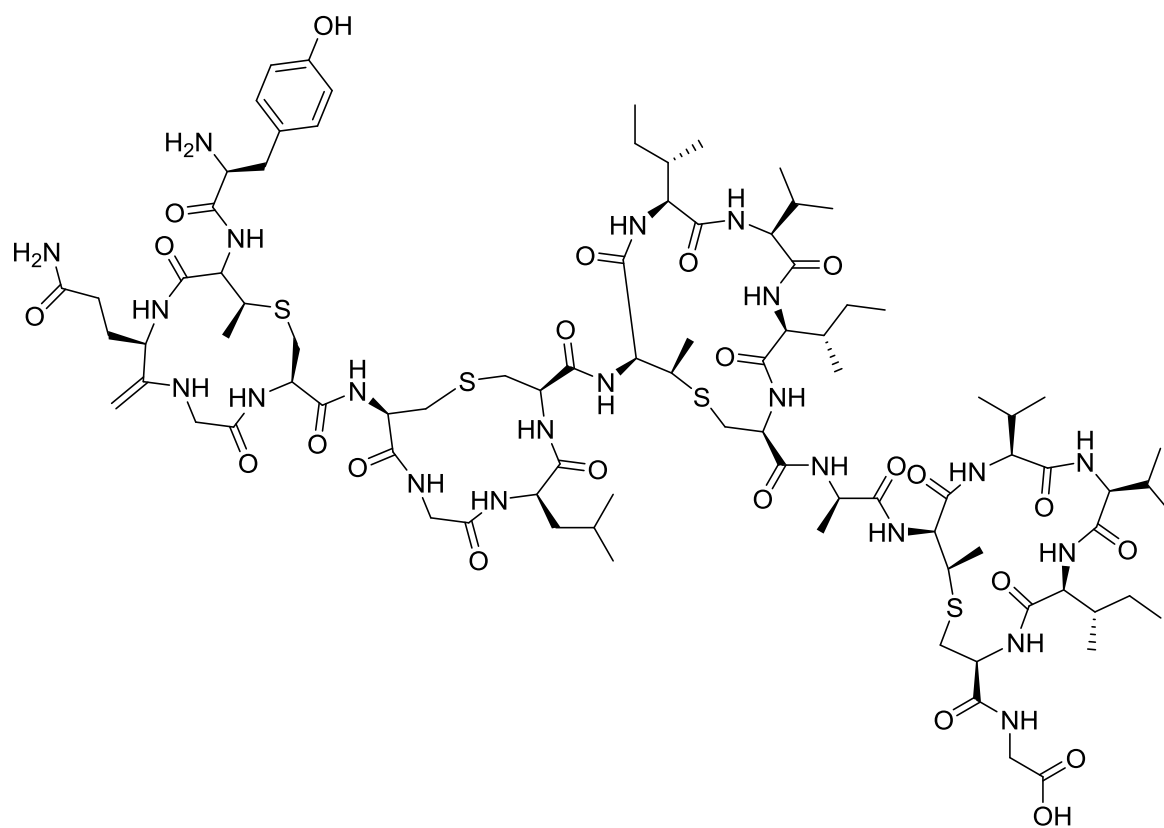
Lanthipeptides are subcategorised in relation to their (Me)Lan biosynthetic synthetases. There are four classes of lanthipeptides. Class I lanthipeptides utilise two synthetases; LanB, the dehydratase responsible for elimination of the hydroxyl group, and LanC which is a cyclase involved in the stereoselective Michael addition. Class II lanthipeptides have a bifunctional synthetase that has incorporated the cyclase and dehydratase into one enzyme complex. Classes III and IV both feature trifunctional enzyme complexes, with an N-terminal phosphatase, a central kinase and a C-terminal cyclase. Class III enzyme complexes are also able to catalyse the formation of carbon-carbon crosslinks, where the double bond of the enol reacts with another Dha and forms a labionine instilling further cyclisation in the structure. Once all the thioether bridges are installed, the modified peptide is then exported out of the producing cell through LanT, an ABC (ATP-binding cassette) transporter. ABC transporters utilise the energy of ATP hydrolysis to transport metabolites across membranes.<sup>50</sup> The leader peptide is then cleaved by a protease. For class I lanthipeptides this is done by LanP. Class II lanthipeptides have a protease domain on LanT that cleaves the leader peptide.<sup>51</sup>

The biosynthesis of lanthipeptides, like most bacterial products, is highly regulated. The genes responsible for regulation of lanthipeptides consist of a receptor histidine kinase (LanK).

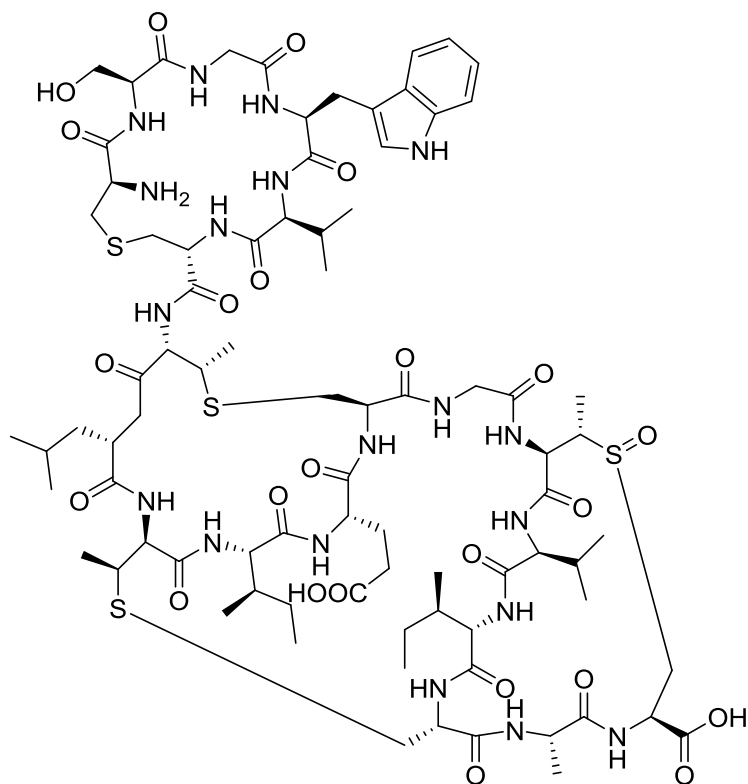


Receptor histidine kinases are transmembrane proteins responsible for signal transduction. When a signal transduction molecule binds to the receptor, it auto-phosphorylates a histidine residue. In regards to lanthipeptide synthesis, the signalling molecule is generally the lanthipeptide itself. The histidine receptor kinase then transfers the phosphate group on to an aspartate residue on the receiver domain of the target protein. LanK transfers the phosphate on to LanR, the response regulator, which in turn activates the biosynthesis of the lanthipeptide.<sup>52</sup> Lanthipeptides can be therefore considered as quorum sensing molecules as they self-regulate their own biosynthesis.<sup>53</sup>

The genes encoding for the thioether (Me)Lan links cannot be used to determine the amino acid sequence of the structure of the core peptide directly, so the structure has to be elucidated *de novo*. This includes both the ring topology, as which double bond the cysteine attacks is not defined leading to multiple possible ring connections, and the Michael addition can result in several diastomeric forms. The topology of the thioethers define the structure of the lanthipeptide; there are linear lanthipeptides like SapT (**10**), bridged lanthipeptides where the thioethers are interlocked as in nisin A (**9**), and globular lanthipeptides in which the MeLans are cross-linked in a compactly folded structure such as actagardine A (**11**).<sup>51</sup> On **11** there is an unusual sulfoxide. Not much is known about the biosynthesis or oxidation process, however it has been shown to be enzyme catalysed.<sup>54</sup>



10



11



Competition experiments with five strains of bacteria were done with T81 on solid growth media to attempt to initiate tikitericin production. Each colony of bacteria, and the zone of inhibition caused by T81, was scraped with a pipette tip and analysed using MALDI MS. From this experiment the expected monoisotopic peak for tikitericin was identified. This showed that tikitericin was produced when T81 was competitively co-cultured with other bacteria. Because T81 failed to produce tikitericin in liquid growth media, to get enough tikitericin, 1200 Petri dish cultures of T81 were scraped and extracted. This only gave ~200 µg of the peptide. Although there was a low mass isolated, this was enough to perform various tests to confirm the linear amino acid sequence and the ring topology. Thermolysin digests, coupled with MS and MSMS analyses, were used to deduce that the peptide ring topology was linear. Further work with GCMS and LCMS using authentic standards established the configurations of the MeLan residues as (2*S*,3*S*,6*R*) and the Lan residue as (2*S*,6*R*).

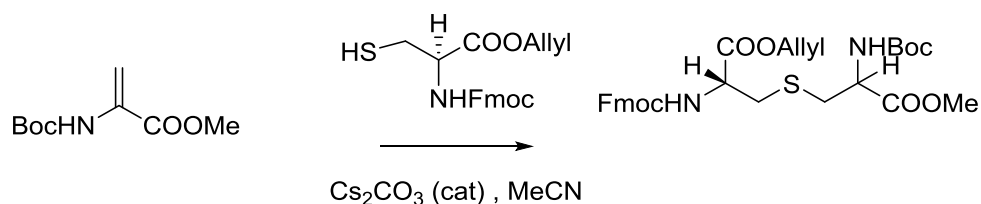
Although T81 shows bioactivity against various strains of extremophilic bacteria *in vitro*, the bioactivity of tikitericin against normal pathogenic bacteria has not been determined as T81 does not produce enough compound. Scraping thousands of plates to extract enough biomass is not a viable strategy to produce more tikitericin, as it is extremely time consuming, low yielding and expensive. Therefore, another method of producing tikitericin had to be devised to determine its bioactivity profile. In this section, the synthesis of starting materials for the solid phase peptide synthesis (SPPS) of tikitericin will be discussed. The aim of the project was to produce a significant amount of the peptide, in pure form, to perform bioassays to elucidate what its potential antibacterial properties are.

## **2.3 Starting material synthesis**

### **2.3.1 Developments in synthetic approaches**

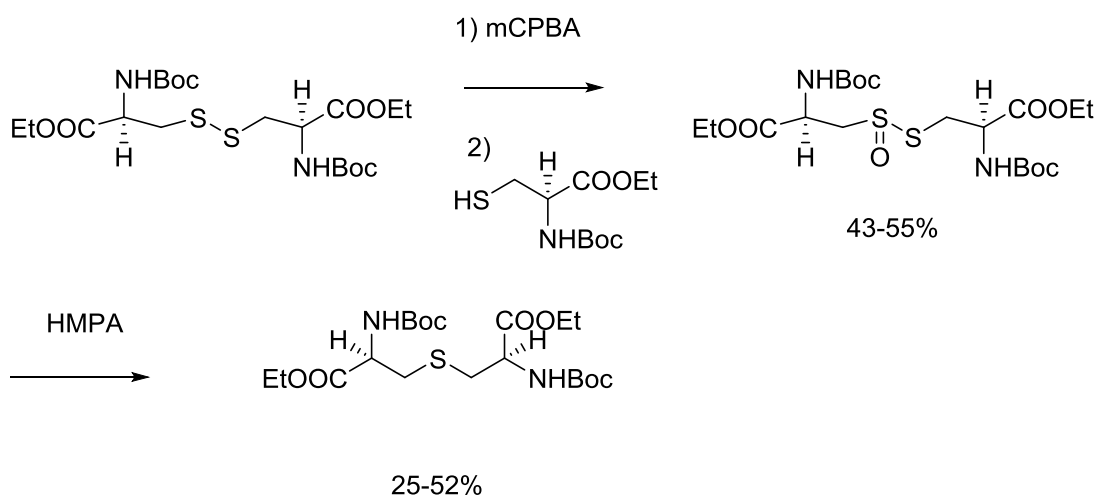
Expectedly, the greatest challenge for synthesising a lanthipeptide is building any lanthionine-containing rings. Unlike the biosynthetic approach, which establishes the (Me)Lan rings after the precursor peptide is synthesised, the synthetic method requires an orthogonally protected (Me)Lan dimer as a starting material. An orthogonally protected (Me)Lan dimer is required for the controlled synthesis of the ring system. This will be discussed in further detail in chapter 3. This can then be used as a building block in SPPS. Previous used synthetic methods will be discussed in the following paragraphs prior to the synthetic strategy carried out for the building blocks used in the synthesis of tikitericin.

The first synthesis of Lan was reported in 1941 by Brown.<sup>57</sup> This was a biomimetic synthesis using Dha and cysteine in a Michael addition to the double bond (Scheme 2.2). However, this synthesis was shown to not be viable by Bradley et al. in 1996.<sup>58</sup> They demonstrated that although the Lan was formed in modest yields (72%), the reaction showed no stereospecificity and two diastereoisomers were formed in a 2:3 ratio. The diastereoisomers were separable by column chromatography, however, and provided an orthogonally protected Lan suitable for SPPS.



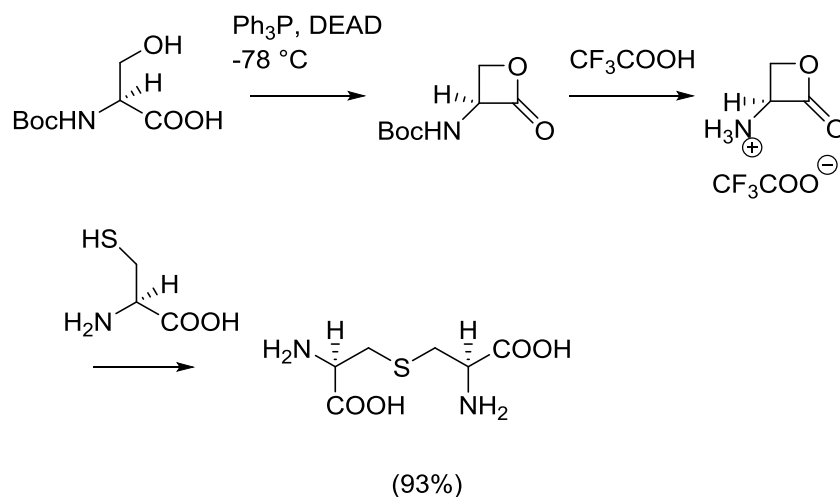
**Scheme 2.2 Lan synthesis by Michael addition**

Desulfurisation to synthesise an orthogonally protected Lan was first reported by Harpp and Gleason.<sup>59</sup> The desulfurisation reaction utilises sulfur dimers to form a thioether while eliminating one sulfur. (*R,R*)-Cysteine derivatives with aminophosphines were used to give the (*R,R*)-Lan. Olsen et al. subsequently adapted this process to form asymmetric Lan residues (Scheme 2.3).<sup>60</sup> Oxidation of the symmetric cysteine derivatives gave a thiosulfinate. A differentially protected cysteine was then used to displace one of the cysteine moieties to give the asymmetric thiosulfinate that was then reacted with HMPA to give the asymmetric Lan. The natures of some of the reactions used are reversible, so the symmetric Lan is also given. The synthesis also suffers from poor yields.



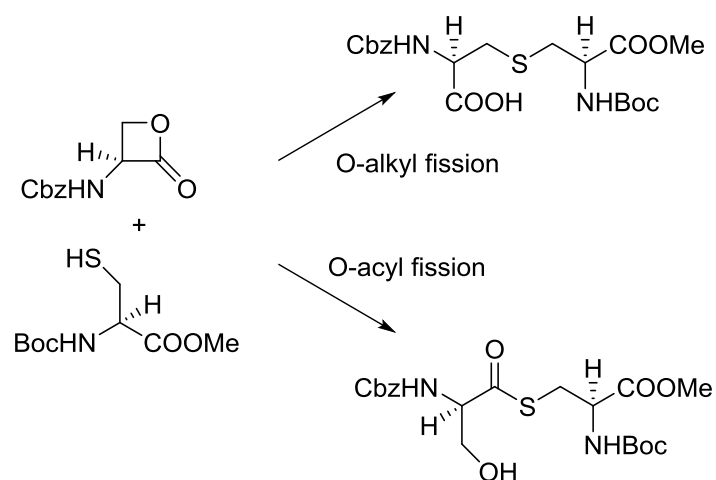
**Scheme 2.3 Asymmetric lanthionines by desulfurisation**

The most successful asymmetric syntheses of Lan have come from the use of ring openings. The most early ring opening reaction for the synthesis of Lan was reported by Vederas and co-workers.<sup>61</sup> The use of Mitsunobu conditions to generate a lactone from Boc-(*S*)-serine followed by deprotection of the Boc group and then ring opening with *R*-cysteine gave the unprotected (*R,R*)-Lan in excellent yield. (Scheme 2.4).

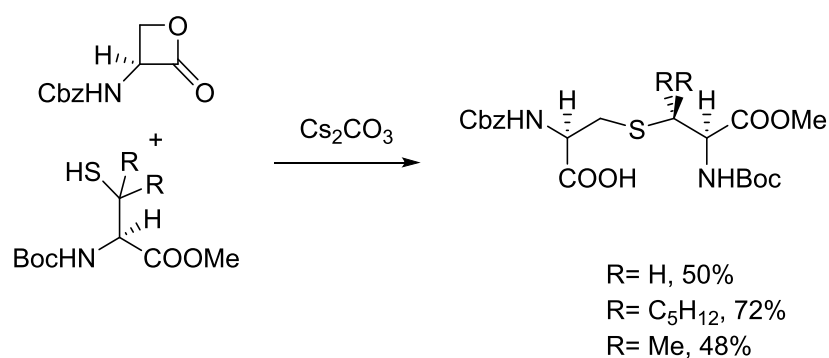


**Scheme 2.4 Vederas' Lan synthesis via a lactone**

There are two potential areas of attack by the thiol to achieve the ring opening of the lactam (Scheme 2.5). The ring opening can either occur in the desired fashion to get the Lan (*O*-alkyl fission) or it can occur with the thiol attacking at the carbonyl (*O*-acyl fission). Shao et al. explored the use of Vederas' method to synthesise an orthogonally protected Lan.<sup>62</sup> A Cbz-*N*-protected serine lactone was reacted with Boc-(*S*)-Cys-OMe for the ring opening reaction. Attempting the reaction with a variety of solvents and bases, they could produce the Lan. However, they found that the reaction also produced the thioester through *O*-acyl fission. The best yield was achieved using  $\text{Cs}_2\text{CO}_3$  as a base. To achieve the lactone ring opening without *O*-acyl fission, bulky groups were substituted on the  $\alpha$ -position of the cysteine. Using  $\beta,\beta$ -pentamethylenecysteine as the nucleophile for the ring opening reaction achieved a high yield without any evidence of *O*-acyl fusion.<sup>63</sup> Further work done by Smith and Goodman showed that the MeLan derivative could also be formed using the  $\alpha$ -methylated lactone (Scheme 2.6). Although these methods do form Lan derivatives, they cannot be tuned to form a Lan that is not substituted. Also Bradley and co-workers found that the reaction conditions could not be tuned so that Fmoc groups could be used, as needed for SPPS.<sup>58</sup>

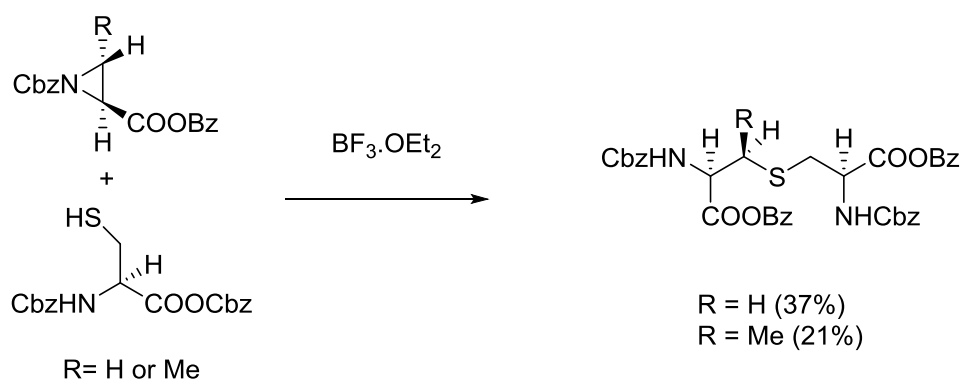


**Scheme 2.5 Potential lactone ring opening products**



**Scheme 2.6 Synthesis of protected Lan derivatives through different R groups**

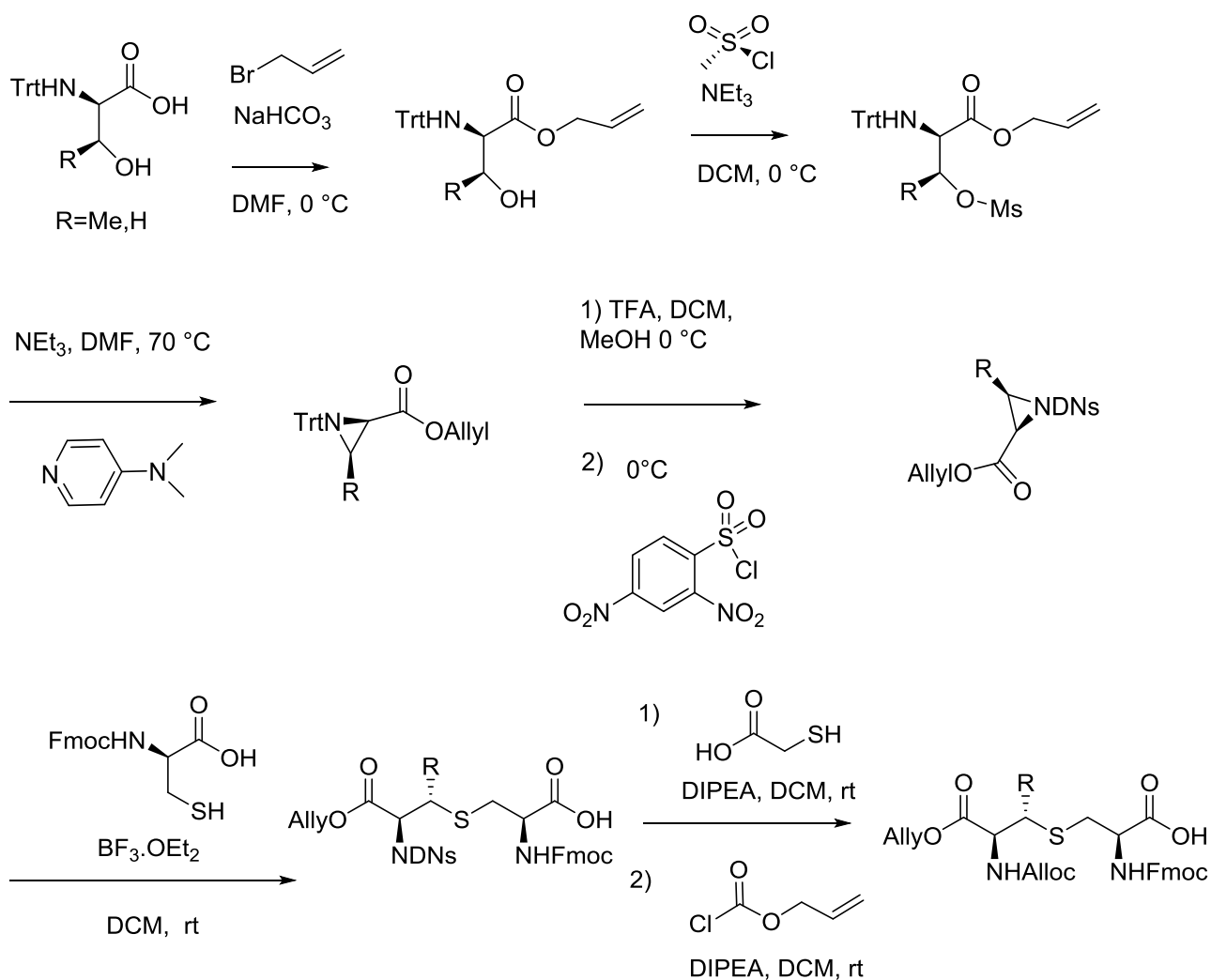
Synthesis of Lan and MeLan by utilising the ring opening of an aziridine has also been attempted. Nakajima first reported this procedure using aziridines derived from serine and threonine.<sup>64</sup> The ring opening was achieved through coupling of a serine derivative catalysed by  $\text{BF}_3 \cdot \text{OEt}_2$  (Scheme 2.7). This reaction, however, gave poor yields.



**Scheme 2.7 (Me)Lan synthesis by aziridine ring opening**

Vederas and co-workers achieved the first synthesis of an orthogonally protected (Me)Lan in the intended diastereomeric form.<sup>65</sup> This is the synthetic approach used for the project described herein (Scheme 2.10). *L*-Thr/Ser-N-Trt carboxylic acid was protected with allyl bromide. The alcohol group was then protected with methanesulfonyl chloride. Elimination of the *O*-Ms group formed an aziridine and then deprotection with trifluoroacetic acid followed by re-protection with 2,4-dinitrobenzenesulfonyl chloride formed the protected aziridine with the right ring opening directing group to get the correct regioselectivity. Coupling with Fmoc-Cys catalysed by  $\text{BF}_3 \cdot \text{OEt}_2$  forms the desired diastereomer. Final deprotection with thioglycolic acid and re-protection with allyl-chloroformate gave the orthogonally protected (Me)Lan suitable for SPPS.





**Scheme 2.8 Synthesis of orthogonally protected (Me)Lan**

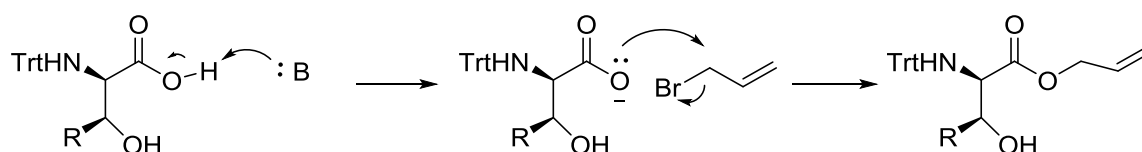
## 2.4 Beginning of the synthesis

In the previous sections synthetic work on production of a (Me)Lan starting material suitable for SPPS has been discussed. Everything hereon is directly correlated to work carried out for this thesis.

### 2.4.1 Allylation

The first step of the synthesis performed was a simple allylation of the carboxylic acid of threonine/serine. The starting materials for this was the triethylamine salt of the tritylated amino acid and allyl bromide. The reaction utilised sodium bicarbonate as a base. The allyl

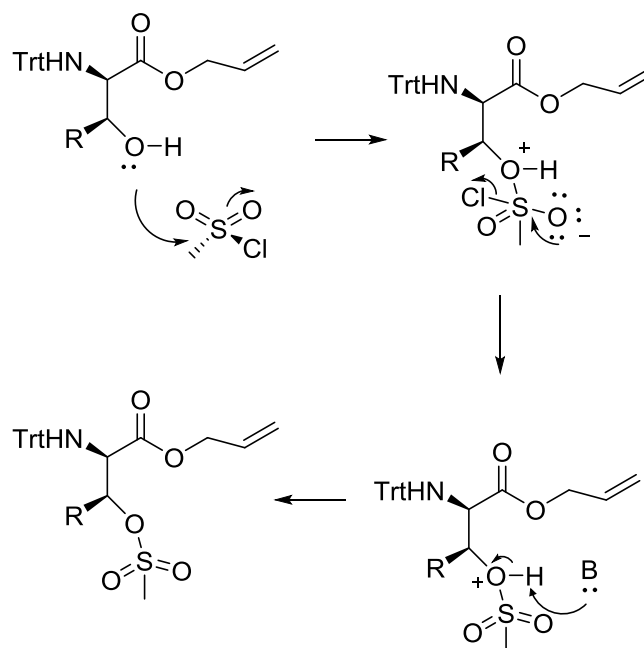
bromide was added to a solution of the amino acid and sodium bicarbonate in DMF and was allowed to warm to room temperature. This was then quenched with water and extracted with ether. The reaction mechanism (Scheme 2.9) is a simple S<sub>N</sub>2 reaction, the carboxylic acid is deprotonated by the base forming an anion. The carboxylate then substitutes for the bromo group and the allylated amino acid is formed. A yield of 58% and 54% was obtained for the serine and threonine respectively.



**Scheme 2.9 Reaction mechanism of allylation**

### 2.4.2 Mesylation

The next step in the synthesis was in preparation for the formation of an aziridine. For this, the alcohol was transformed into a good leaving group through addition of methanesulfonyl chloride to form a mesylate. The allyl amino acid was dissolved in DCM and to this solution was added triethylamine and methanesulfonyl chloride. This was reacted at room temperature and then quenched with a saturated solution of ammonium chloride and extracted with DCM. The reaction mechanism is through substitution of the chloride with the alcohol (Scheme 2.10). The hydroxyl group attacks methanesulfonyl chloride breaking one of the sulfur-oxygen double bonds, forming a negative charge on the oxygen. The double bond reforms and the chloride leaves forming the mesylated allyl amino acid. A yield of 63% and 44% was obtained for the serine and threonine respectively.

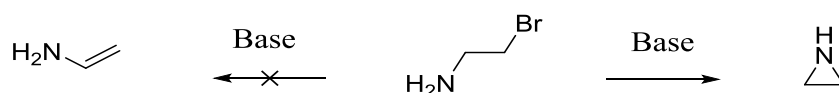


**Scheme 2.10** Reaction mechanism of mesylation

## 2.5 Aziridines

### 2.5.1 Aziridine formation

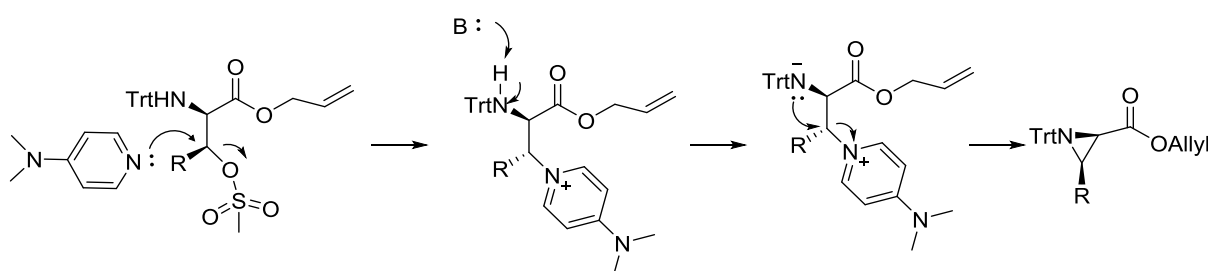
Aziridines have been known since the late 19<sup>th</sup> century, the first one being synthesised by Gabriel in an attempt to synthesise ethenamine by reacting 2-bromoethan-1-amine with base (Scheme 2.11).<sup>66</sup> Marckwald subsequently proposed the correct cyclic structure, suggesting that the chemistry was better explained with the formation of an aziridine than ethenamine.<sup>67</sup> This method of synthesis involving the intramolecular substitution of a leaving group by nitrogen later became known as the Gabriel synthesis and is the general method of synthesis utilised for forming the three-membered heterocycle.



**Scheme 2.11** Gabriel synthesis

In tikitericin, there are three Melans, all being the same diastereomer meaning that during the formation of the aziridine used for the coupling reaction, the correct stereochemistry needs to be conserved so that the single desired diastereomer is formed. The Gabriel synthesis can be

applied to the formation of single enantiomers provided that the starting materials used are enantiopure. 1,2-Amino alcohols, as long as the alcohol is activated for leaving, can form aziridines through a ring-closing reaction, thus the amino acids serine and threonine are convenient starting materials as they are chiral pool reagents.<sup>68</sup> The diastereomer of threonine used for the formation of the aziridine was *L*-threonine (2*S*,3*R*). 4-Dimethylaminopyridine (DMAP) was used in the reaction as a nucleophilic catalyst, and through this interaction the stereochemistry of the methyl group is retained, not inverted (Scheme 2.12). A yield of 72% and 53% was obtained for the serine and threonine respectively.

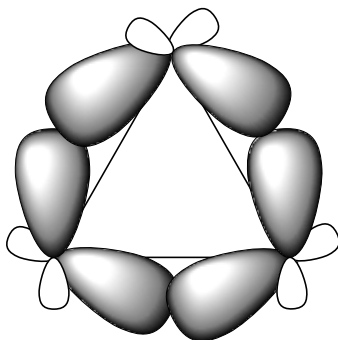


**Scheme 2.12 Reaction mechanism of aziridine formation**

In the previous reaction, the hydroxyl group was mesylated to provide a good leaving group. The reaction proceeds via an  $S_N2$  type mechanism. The first step involves the substitution of DMAP for the mesyl group which inverts the stereochemistry. Triethylamine then extracts a proton from the nitrogen atom. The nitrogen anion is then able to attack the position to substitute for DMAP. This reverts the stereochemistry of the methyl group. This reaction is the same in terms of the serine analogue except the only chiral position is the amine centre.

### 2.5.2 Bonding in aziridines

Aziridine, being a three membered ring, has internal ring bond angles which are formally  $60^\circ$ . This is a large discrepancy from the normal angles established in  $sp^3$ -hybridised bonds. Thus, aziridines are highly strained heterocycles; the B  yer strain for the parent compound is estimated to be  $111 \text{ kJ mol}^{-1}$ .<sup>69</sup> To minimise the strain caused by the small bond angles, the hybridisation of the atoms internal to the ring is changed. The orbitals contributing to the endocyclic bonds digress from being  $sp^3$  hybridised to having a greater p character and overlap outside the axes joining the centre of the nuclei (Fig. 2.1).<sup>70</sup> Thus, the bonding in aziridines can be considered as bent or “banana bonds”.



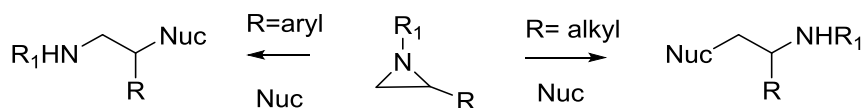
**Figure 2.1 Bonding in aziridines**

The effect that this has on the exocyclic bonds is that they have an increase in s character and consequently a widening of bond angles. This widening of the angles increases the likelihood of ring opening reactions as it is more accessible for nucleophilic attack. The amount of s character a bond has can be determined by NMR through  $^1J_{\text{CH}}$  coupling constants. It was shown that aziridine had a  $^1J_{\text{CH}}$  coupling constant of 166 Hz, which gives an s character of 32.2%. This is in comparison to methane's  $^1J_{\text{CH}}$  coupling constant of 125 Hz, which is about 25% s.<sup>71</sup> Furthermore, the lone pair of the ring nitrogen also has increased s character which decreases the basicity of the cyclic amine.

### 2.5.3 Ring-opening processes

Because of the heterocycle's appreciable B  yer strain, it is susceptible to nucleophilic ring opening reactions, much like epoxides are. This is also due to the electrophilicity of the ring nitrogen. In comparison to epoxides, aziridines are less reactive due to nitrogen's diminished electrophilicity, however they are still able to undergo ring opening under relatively mild conditions.<sup>72</sup>

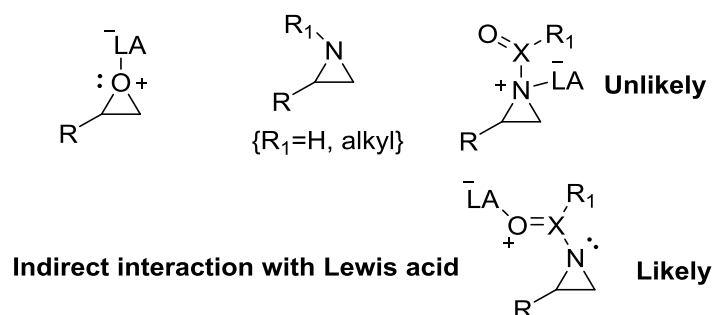
There are a number of factors requiring consideration in the ring opening process, one of these being the regioselectivity of the reaction. Comparable to epoxides, the point of nucleophilic attack is at one of the carbons. If the aziridine is non-symmetrical then there are two points of attack possible leading to two potential products (Scheme 2.13). The point of attack is generally the least substituted carbon, however if there is an aryl group substituted on the carbon then this will be the point of attack due to electronic reasons.<sup>69</sup>



**Scheme 2.13 Regioselectivity in ring opening reactions**

Lewis acids are another factor which can be considered in the ring opening process. Epoxides are Lewis bases and therefore can associate with a Lewis acid via non-bonding electron pairs. This has an effect of weakening the C-O bond that aids in increasing the rate of the ring opening reaction. Aziridines, on the other hand, do not generally have a direct interaction with the Lewis acid. Only when there is a non-oxygenated *N*-substituent will there be a direct interaction with the nitrogen atom (Figure 2.2). This is due to the fact that having an electron withdrawing group in close proximity to the nitrogen will destabilise the positive charge on the nitrogen when interacting with the Lewis acid. As will be discussed in detail below, polar activating *N*-substituents are predominately needed for efficient ring-openings of aziridines. Because of this, the chemistry of aziridines is less dominated by the use of Lewis acids than epoxides, however because polar oxygenated *N*-substituents are desirable for ring openings, Lewis acid activation can still be utilised through the coordination of the oxygen lone pairs to the Lewis acid.<sup>72</sup>

#### Direct interaction with Lewis acid

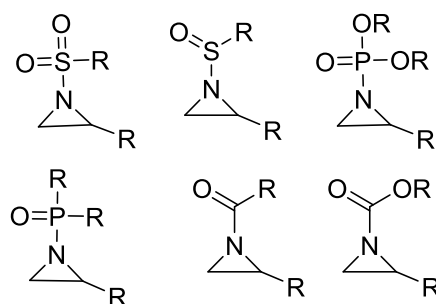


**Figure 2.2 Lewis acid interactions**

#### 2.5.4 *N*-substituent effect

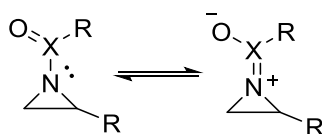
Unlike epoxides, aziridines have an extra valency available on their heteroatom meaning that the chemistry of aziridines is made more complex by available substituents. It is advantageous to utilise substituents that mask the N-H bond, as deprotonation is a competitive side reaction when using highly basic nucleophiles such as carbon-centred ones. Furthermore, the nature of the *N*-substituent places the aziridine into two categories; activated and nonactivated. These classifications determine whether or not the ring-opening proceeds in the absence of a positive charge. An activated aziridine has an *N*-substituent capable of stabilising the negative charge

formed on the nitrogen during the ring-opening reaction while a nonactivated aziridine requires a protonation step for the ring-cleavage so is catalysed using a Brønsted acid. Ring-opening reactions of nonactivated aziridines generally occur by a  $S_N1$  reaction mechanism while activated aziridines generally react via  $S_N2$ .<sup>72</sup> The substituents for nonactivated aziridines are aryl, allyl, alkyl or H.<sup>73</sup> Activating groups are generally oxygenated functional groups such as sulfonyl, sulfinyl, phosphoryl, phosphinyl or carbonyl. (Fig 2.3)



**Figure 2.3 Aziridine activating groups**

Between these groups, there is little possibility to form an imidate-like resonance isomer (Scheme 2.14) through the interaction of the non-bonded electron pairs on the nitrogen and the  $X=O$  bond. This is due to an associated increase in ring strain which would be a higher energy contributor than the non-charged resonance structure.

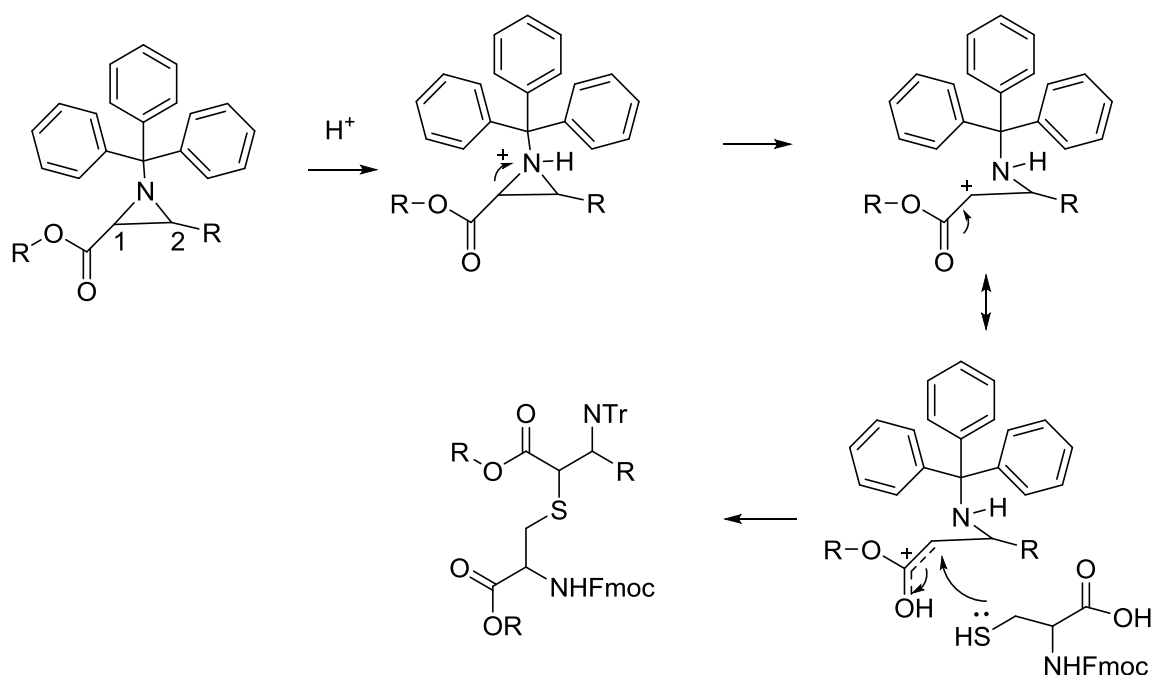


**Scheme 2.14 Resonance isomers of substituted aziridines**

Because there is a low probability of the amidate resonance isomer forming, there must be another reason why these substituents activate the aziridine towards ring-opening. There are both kinetic and thermodynamic reasons that these substituents activate the ring opening. The kinetic reason is that the substituent provides an inductive effect. Because there are polar groups attached they increase polarisation of the C-N bond, which is already polar due to the larger electronegativity of N compared to C. This increases the electrophilicity of the carbon atoms in the heterocycle, increasing their susceptibility towards nucleophilic attack. The thermodynamic reasoning is that the negative charge formed on the nitrogen is able to be stabilised through the N-substituent. Carbonyl substituents are able to form resonance structures to stabilise the negative charge while sulfonamides, phosphonamide and phosphinamide anions are primarily stabilised through inductive effects.<sup>74</sup>

### 2.5.6 Changing of the aziridine substituent

The tritylaziridine is not suitable for the formation of (Me)Lan when reacted with Fmoc-cysteine. This is because the current substituent is non-activating meaning cysteine would substitute at the wrong position. (Scheme 2.15)

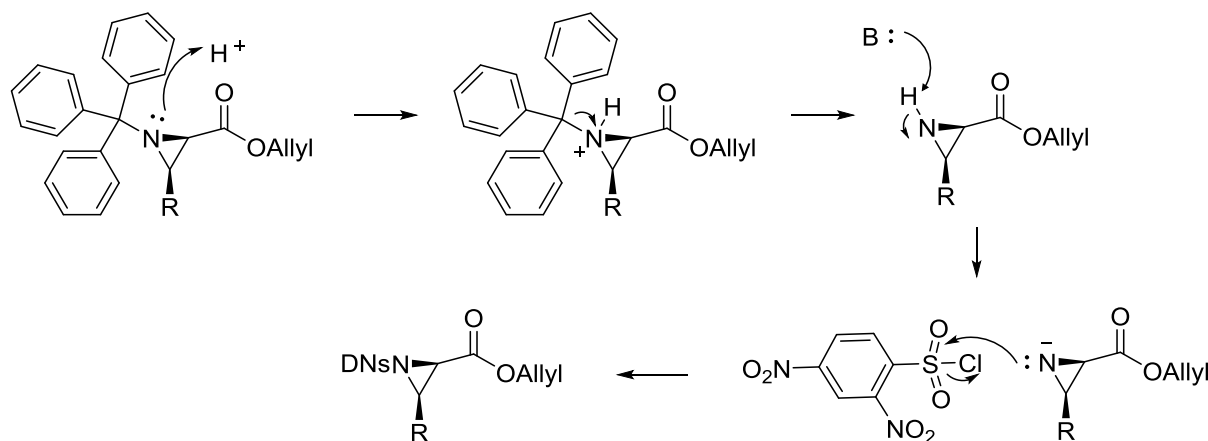


**Scheme 2.15 Postulated reaction mechanism of tritylated aziridine with cysteine**

Because non-activated aziridines go through an  $S_N1$ -type reaction mechanism, the nucleophile will attack where the carbocation formed is best stabilised, in this case at position 1 as it is next to a carbonyl group. The carbonyl is able to stabilise the carbocation by electron donation through resonance. If the carbocation formed at position 2, then there would only be a weak induction from either the methyl group or the hydrogen atoms so the carbocation would be less stabilised, therefore substitution at position 1 is far more favourable. Because of this, it was proposed to change the *N*-substituent from a non-activating one to an activating one. To do this required the removal of the trityl group. This was done using trifluoroacetic acid (Scheme 2.16). The aziridine nitrogen is protonated which forms a positive charge. Because the trityl group is more capable of stabilising the positive charge due to its numerous resonance structures available, the N-Tr bond is cleaved. After this the reaction was worked up, partitioned between water and ether. The ether layer was used to remove the trityl group. The aqueous layer was collected and basified to deprotonate the nitrogen atom of the aziridine to



increase its nucleophilicity. Due to the low solubility of 2,4-dinitrobenzenesulfonyl chlorides in water, ethyl acetate was added to solubilise the 2,4-dinitrobenzenesulfonyl chloride and the reaction proceeded as a biphase. The nitrogen atom of the aziridine attacks the sulfur of 2,4-dinitrobenzenesulfonyl chloride and the chlorine atom leaves, forming the activated aziridine. A yield of 58% and 57% was obtained for the aziridine and methylaziridine respectively.



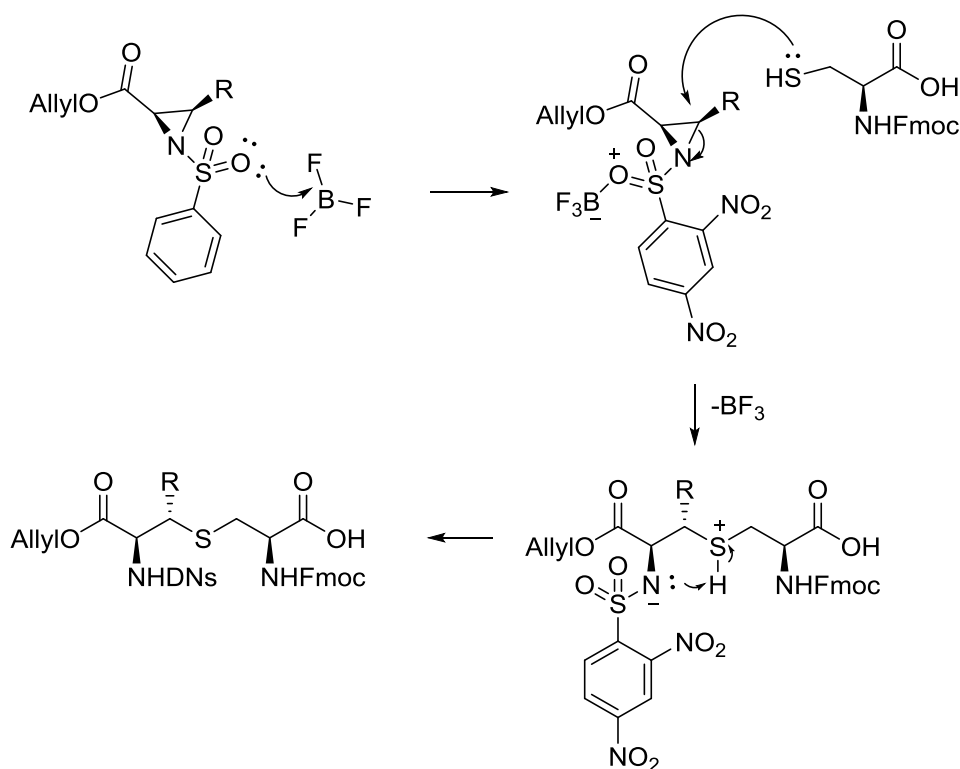
**Scheme 2.16 Reaction mechanism of detritylation and reprotection**

The reason why a sulfonyl protecting group was chosen over the other potential protecting groups is because they are the most widely employed groups for activation of ring opening reactions and therefore their chemistry is best known. In addition to that, the sulfonyl group has a high aptitude for stabilising the negative charge formed during the ring opening and therefore these reactions are particularly facile.<sup>74</sup>

It was decided that 2,4-dinitrobenzenesulfonyl would be used as the sulfonyl group because in comparison to other sulfonyl groups its deprotection is under relatively mild conditions. There were a number of nitrobenzenesulfonyl protecting groups available. The other ones which could have been used were *p*- or *o*-nitrobenzenesulfonyl chloride. Although these would activate the aziridine in the same way, the deprotection reaction tends to go slightly different with them. The way to remove these groups is through nucleophilic aromatic substitution with a thiolate which substitutes *ipso* to the sulfonyl group (see section 2.5.1). There has been evidence that this is less efficient with *p*-nitrobenzenesulfonyl and *o*-nitrobenzenesulfonyl as the nucleophilic aromatic substitution can occur at the position bearing the nitro group leading to the reaction not going forward. The 2,4-dinitrobenzenesulfonyl does not suffer this lack of regioselectivity and hence was the best protecting group for synthesis.<sup>75</sup>

### 2.5.7 Ring opening reaction

Upon the aziridine being activated it was decided that the ring opening reaction would be carried out. This was done by reacting the activated aziridine with Fmoc-cysteine in the presence of boron trifluoride diethyl etherate as a source of  $\text{BF}_3$  as a Lewis acid. Fmoc-cysteine was used in a 4:1 ratio with the aziridine to alleviate the problem that a percentage of the Fmoc-cysteine will most likely be in its oxidised form as a disulphide and hence will not react. The reaction mechanism is as follows (Scheme 2.17). The boron fluoride associates covalently with one of the oxygen atoms on the sulfonyl group forming a positive charge on the oxygen. This increases the electron withdrawing effect of the sulfonyl group, increasing the electrophilicity of the carbon atoms in the aziridine. Fmoc-cysteine reacts with the aziridine in an  $\text{S}_{\text{N}}2$  manner reacting at the less substituted position. In the case of the methyl aziridine there is inversion at the chiral centre. Proton transfer from the sulfur to the nitrogen yields the (Me)Lan shown. A yield of 43% and 38% was obtained for the Lan and MeLan respectively.



**Scheme 2.17 Reaction mechanism of aziridine ring opening**

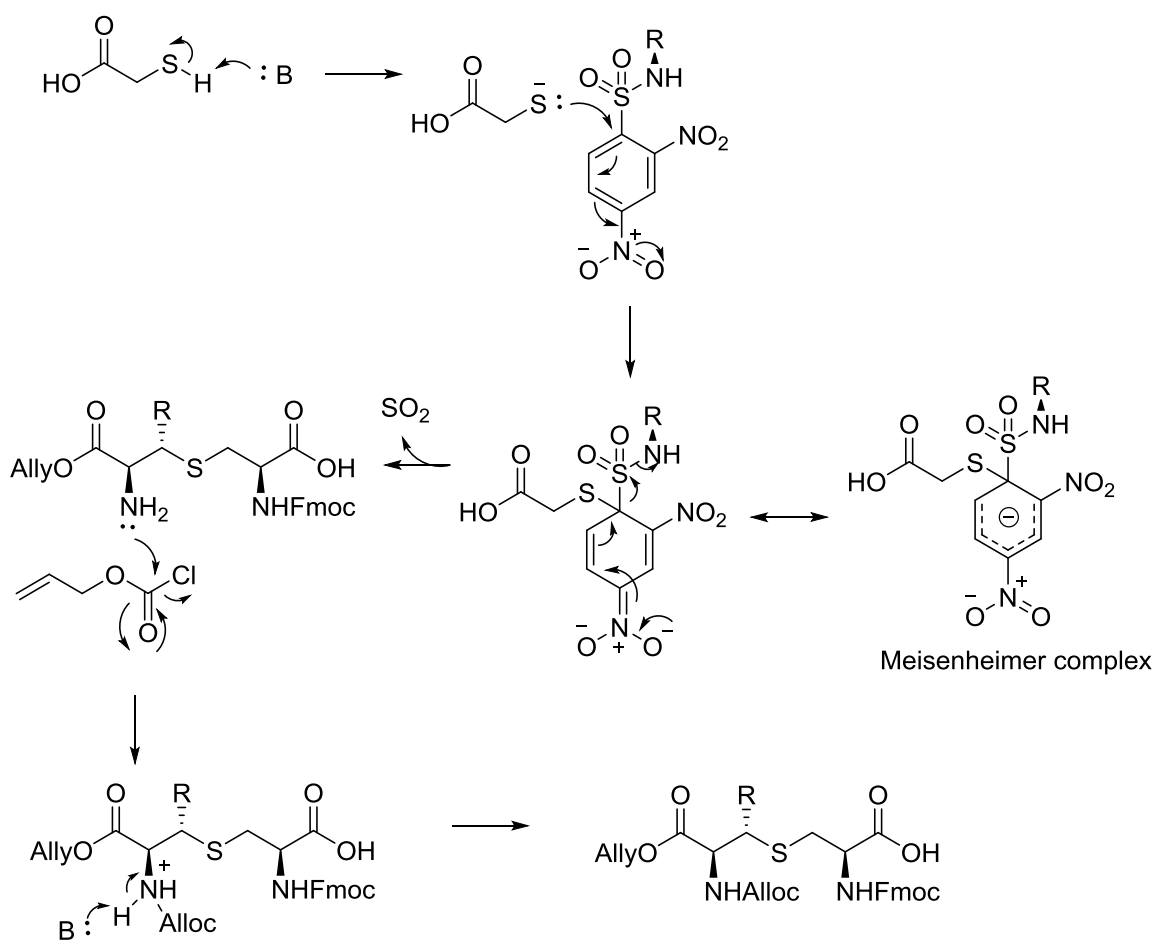
A note on the purification of this and the next (Me)Lan. It was decided that 0.2% acetic acid would be used in the elution solvent for silica gel chromatography. This varies from the 0.1% used by Vederas.<sup>65</sup> The acetic acid is used in the elution to enhance the separation of the (Me)Lan from the starting material. Due to the carboxylic acid now on the structure this will

decrease the effectiveness of the chromatography as a percentage of the (Me)Lan will be in its dissociated form, meaning that streaking will occur throughout the column. Acid in the elution mix prevents the dissociation of the (Me)Lan from occurring due to Le Chatelier's principle, the equilibrium is pushed to the non-dissociated form as there are already protons present. Using 0.1% did not have the desired effect so it was decided that it would be increased to 0.2%, which gave the desired effect of improving chromatographic resolution.

## **2.6 Troubleshooting of the Deprotection/reprotection of (Me)Lan**

### **2.6.1 Deprotection/reprotection of (Me)Lan**

The final reaction of this synthesis was to cleave the 2,4-dinitrobenzenesulfonyl group from the (Me)Lan and reprotect it with the allyloxycarbonyl (Alloc) group. Vederas<sup>65</sup> proposed a one-pot synthesis for this reaction using *N,N*-diisopropylethylamine (DIPEA) as a base. The DN (Me)Lan was dissolved in DCM and reacted with thioglycolic acid for 2 hours and then allyl chloroformate was added with a fresh quantity of base. This was then reacted overnight and purified to yield the allocylated (Me)Lan. The reaction mechanism is as follows (Scheme 2.18). Base deprotonates thioglycolic acid forming a thiolate anion. This then goes through nucleophilic aromatic substitution and attacks the *ipso* position of the thionyl group on 2,4-dinitrobenzenesulfonyl. This forms a Meisenheimer complex as shown. Aromaticity is restored by elimination of the thionyl which breaks down to the free amine and sulfur dioxide. The amine then attacks allyl chloroformate forming an amide. Base then deprotonates the nitrogen atom and the allocylated (Me)Lan is formed.



**Scheme 2.18 Reaction mechanism of DNs deprotection/alloc reprotect**

Setting out optimistically to complete this reaction, it was attempted ten times without yielding anything which resembled Vederas' data. Upon realising that continuing to try this method for this reaction was futile, it was decided that investigating why this reaction was not working was necessary. There were some questions that were needed to be answered to diagnose the issue with this synthesis:

1. Was DIPEA negatively affecting this reaction and not facilitating thioglycolic acid nucleophilic aromatic substitution?
2. Would it be beneficial to separate this "one pot" reaction into two, with a purification step between the deprotection and reprotection?

### 2.6.2 Investigation into the effects of DIPEA

The scenario tested here was is it actually worth using a strong base during this reaction when the majority of the time the only reason for its existence is to sequester liberated protons. Investigations into the purified products of the failed reactions revealed that there was evidence

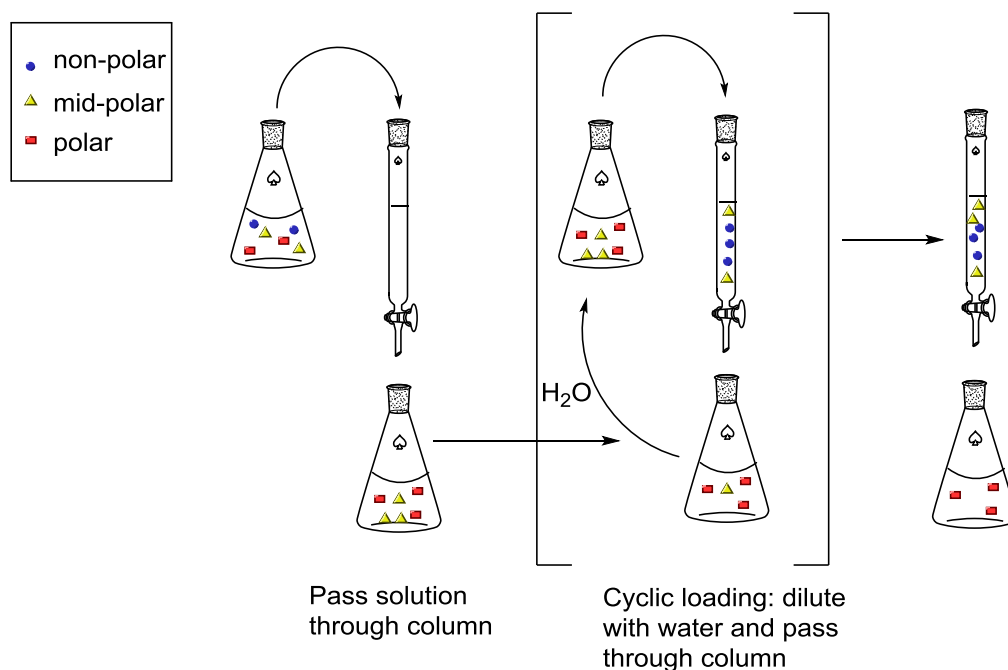
of Fmoc deprotection occurring. It was assumed that DIPEA was culpable for this and that a different, weaker base should be used. It was decided that DIPEA would be replaced with potassium carbonate for this reaction, so the same methodology was carried out. This, however, resulted in the degradation of the DN-(Me)Lan and so it was hypothesised that DIPEA was rather essential for this reaction because it allows the formation of the thiolate for the nucleophilic aromatic substitution.

In reference to the  $pK_a$  values of each species involved in this reaction, DIPEA is 10.75,<sup>76</sup> the acidic proton of Fmoc is 25,<sup>77</sup> the thiol of thioglycolic acid is 10.31 and for reference piperidine is 11.22.<sup>78</sup> This data clearly shows that DIPEA is both sufficient for the formation of the thiolate of thioglycolic acid and also, in comparison with piperidine is able to remove the acidic proton of Fmoc and deprotect it. This is counterintuitive however, as DIPEA is regularly used in peptide couplings. However, these reactions are performed in a short time span, compared to Vederas' method which has the (Me)Lan in the presence of it over a 24 hour period. Piperidine is more adept at the deprotection of Fmoc because it is both strongly basic and also not as sterically hindered as DIPEA. If the (Me)Lan were treated with DIPEA over a long period of time it would be logical to assume that it would have similar effects to that of piperidine. With this knowledge, it was decided that the next action to be taken was to develop a purification step between the steps of deprotection and reprotecting to separate the DIPEA from the (Me)Lan.

### **2.6.3 Investigations into chromatographic method and DN deprotection**

To remove the DIPEA, the most satisfactory manner would be through chromatographic methods. The use of silica gel would not be the most advantageous method as the presence of both an amine and carboxylic acid on the structure would make silica chromatography difficult. Thus, reverse phase chromatography was implemented for this separation using a polystyrene/divinyl/benzene (PSDVB) copolymer, in particular HP-20. The deprotection was executed in the same fashion as Vederas' method. Once it was completed, the reaction mixture was evaporated under reduced pressure. This was then dissolved in methanol and cyclic loaded (Figure 2.4) onto HP-20. The cyclic loading protocol involves adhering all the compounds in the reaction mixture onto the stationary phase. First the methanolic reaction mixture was passed through the stationary phase, and this was collected and diluted with water by 50%. This process was repeated until the methanol concentration was 12.5%. This particular concentration was decided, as the original solution was yellow and this colour had been adsorbed by the stationary phase, leaving the eluent colourless. This method is particularly useful as it addresses the issue of loading a variety of compounds with different polarities to

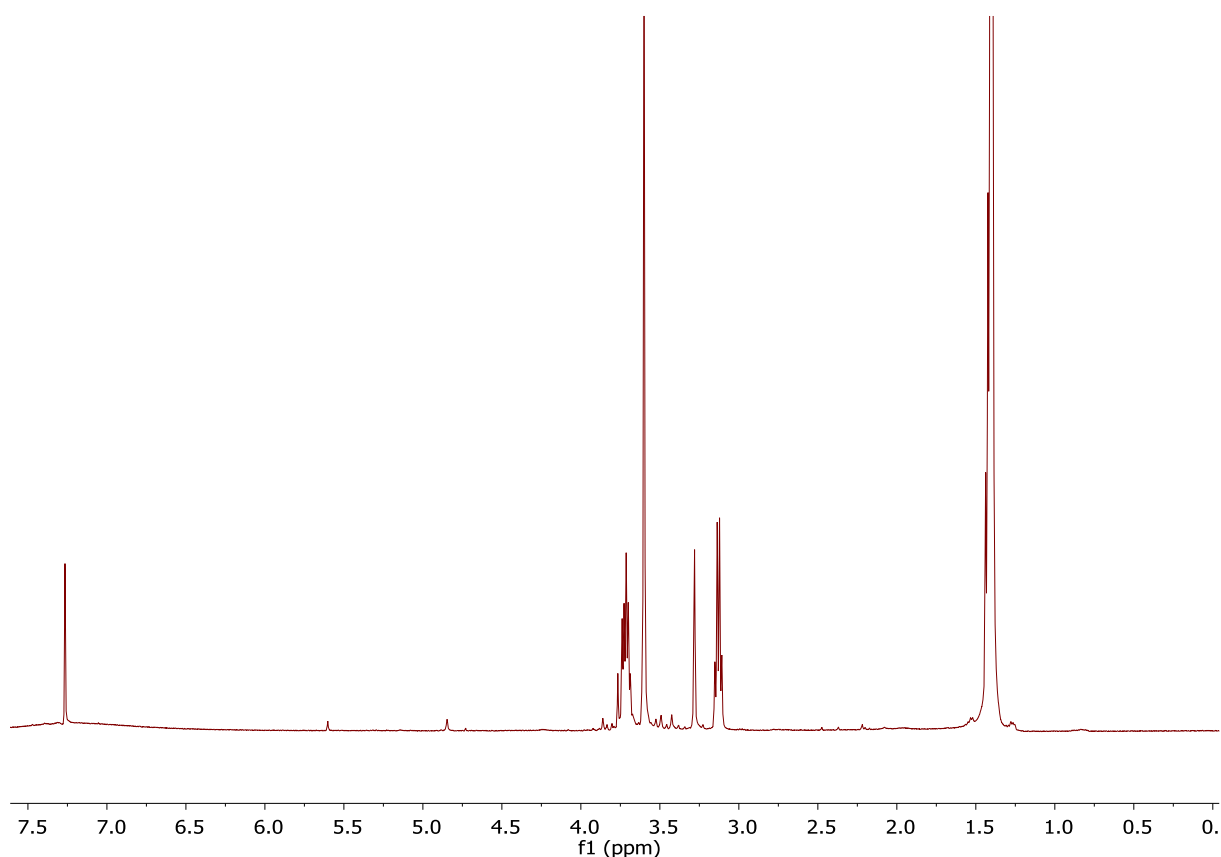
the non-polar stationary phase. As the methanolic reaction mix is first passed through the stationary phase prior to diluting the solution with water, any significant non polar components will be adsorbed to the HP-20. Over time, as the concentration of water is increased compounds of mid and high polarity are absorbed sequentially. The HP-20 was then eluted with 30%, 50% and 70% acetone in water and then 100% acetone.



**Figure 2.4 Cyclic loading protocol**

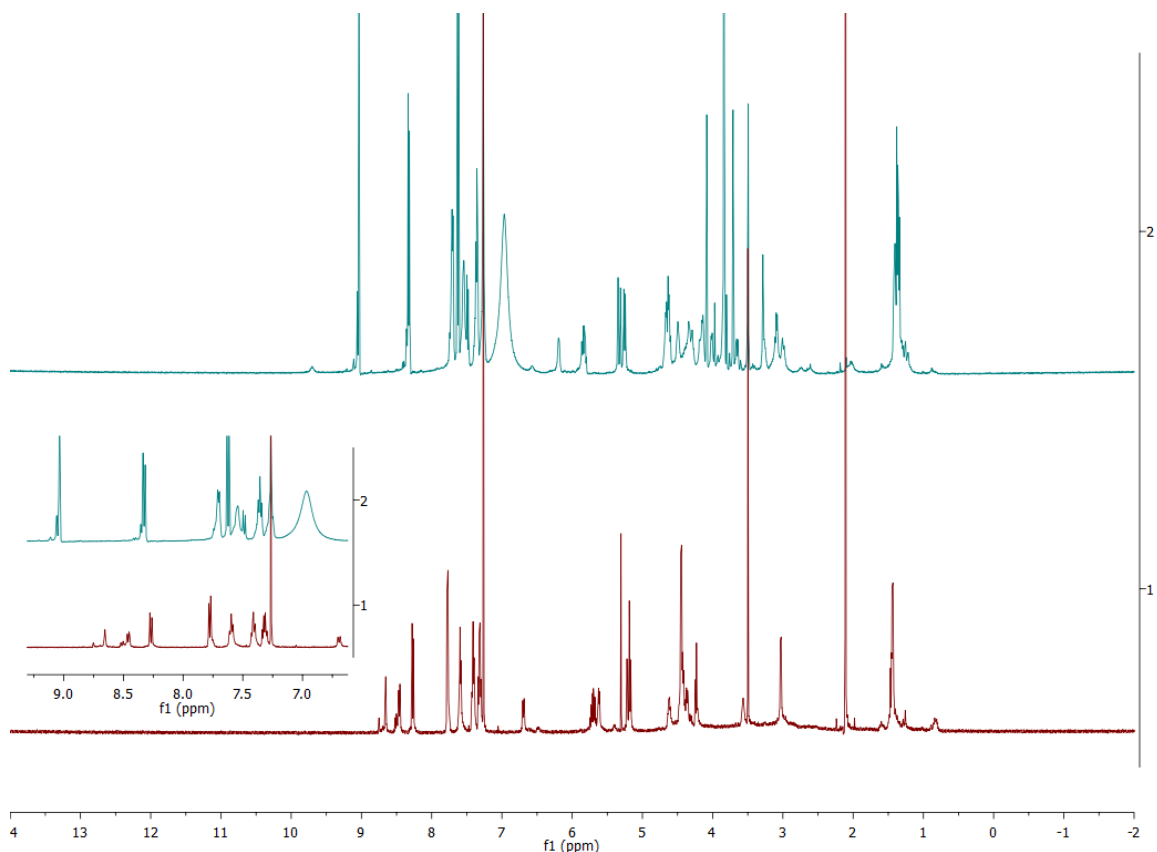
The simplest compound to detect was thioglycolic acid due to its distinct, unpleasant odor. The thioglycolic acid had been successfully removed from the reaction mixture. This was detected in the loading eluent by olfactory analysis and unsurprisingly, given its small, highly polar nature, did not adhere to the stationary phase. The 30 – 100% fractions were evaporated under reduced pressure and analysed by NMR.

The next aspect of interest was whether the DIPEA had adhered to the stationary phase. Confirmation that it had adhered came from the <sup>1</sup>H NMR of the 70% fraction (Figure 2.5). This was an excellent result, as it had clearly been separated from the (Me)Lan intermediate. This shows that this method of separation is sufficient for purification of the reaction between the deprotection and reprotection so the Fmoc group would be safeguarded during the allocylation reaction.



**Figure 2.5  $^1\text{H}$  NMR of 70% fraction (500 MHz,  $\text{CDCl}_3$ )**

In spite of this result, there still was a need to investigate if the Fmoc group was attached to the (Me)Lan derivative and also if the thioglycolic acid was initiating the deprotection of the DN<sub>s</sub> group. The answer came from the 50% acetone fraction. Figure 2.6 shows the  $^1\text{H}$  NMR spectrum of the starting material in comparison to the  $^1\text{H}$  NMR of the 50% acetone fraction. An expanded section of the aromatic region is shown in the bottom left corner. This region is of most concern as both the Fmoc and DN<sub>s</sub> proton resonances are found here. Firstly, it can be seen that the Fmoc protons have barely shifted meaning that the Fmoc group is still intact. The DN<sub>s</sub> aromatic proton resonances can be seen between 8 and 9 ppm. What is expected is if there is a nucleophilic aromatic substitution occurring, as proposed then we should see a significant shift in these resonances should be observed as an electron withdrawing group was replaced with an electron donating group. In the 50% fraction this is exactly what is observed with the DN<sub>s</sub> protons. To make sure that this was not just a product of substitution at one of the NO<sub>2</sub> positions, a literature search was performed searching for 2-[(2,4-dinitrophenyl)thio]acetic acid. Fortunately it had been previously synthesised and the literature data was in agreement with the experimental.<sup>79</sup>



**Figure 2.6  $^1\text{H}$  NMR of 50% fraction (cyan) and starting material (red) (500 MHz,  $\text{CDCl}_3$ )**

Once this method of separation had been established the next investigation to be carried out was if the alloc protection could occur with a base such as potassium carbonate. Using DCM as a solvent the (Me)Lan was dissolved and  $\text{K}_2\text{CO}_3$  and Alloc-Cl were added to the solution and reacted overnight. It was decided that 4 equivalents of Alloc-Cl would be used, instead of the suggested 2 equivalents, to combat the potential for it reacting with the carboxylic acids on 2-[(2,4-dinitrophenyl)thio]-acetic acid and the (Me)Lan itself. The reaction was completed and purified revealing that the desired product had been produced. A yield of 52% and 46% was obtained for the Lan and MeLan respectively.

Chromatography was particularly difficult for this as the  $r_f$  values for 2-[(2,4-dinitrophenyl)thio]-acetic acid and the alloc protected (Me)Lan in 4:1 Pet ether:ethyl acetate was 0.45 and 0.34 giving a difference of 0.1. The issue here was that any lower concentrations of ethyl acetate meant that the 2-[(2,4-dinitrophenyl)thio]acetic acid eluted even slower from the column leading to a large waste of solvent. In other words, this synthetic procedure required further improvement.



## 2.7 Concluding remarks

The synthesis was slightly improved from the original proposed method. The one step which still requires improvement is the chromatography that still proves troublesome. Kan and Fukuyama<sup>80</sup> suggested that removal of the 2-[(2,4-dinitrophenyl)thio]acetic acid from the reaction mix with a liquid-liquid partition using ether and an aqueous NaHCO<sub>3</sub> solution is possible. Ideally the 2-[(2,4-dinitrophenyl)thio]acetic acid is removed in the aqueous layer. If this synthesis is ever attempted again, it is suggested that this separation is attempted prior to chromatography of the final reaction.

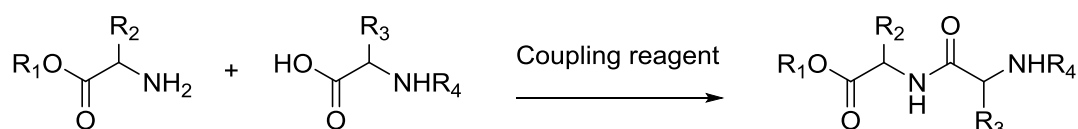
Both the Lan and the MeLan were synthesised by the same reaction method. This was a time consuming process, as the final step of the synthetic scheme had to be done three times for each starting material due to issues with purification, meaning six rounds of dimer starting material were carried out. The optimised yield for the final Lan and MeLan dimers were 52% and 46%, respectively, which is an improvement of yields in comparison to the Vederas method. The overall synthetic yield for the Lan and Melan were 3.4% and 1.3% respectively. The improvement of the current method, as discussed, took a large amount of time meaning that time constraints were imposed on the rest of the project. This, coupled with low quantities of material produced, meant that only one attempt at synthesising tikitericin could be made. With the alloc protected (Me)Lan synthesised, they were then used in the attempted solid phase synthesis of tikitericin which is discussed in Chapter 3, although given the time constraints upon an MSc only one attempt at the total synthesis was feasible.

## Chapter 3

### Solid phase peptide synthesis

#### 3.1 Introduction

Success in peptide synthesis effectively hinges on an immaculate combination of suitable coupling reagents and protecting groups working synergistically. Peptide synthesis can be defined as the formation of a peptide bond between two amino acids (Scheme 3.1). The formation of the amide linkage between the amino acids is called a coupling.<sup>81</sup>



**Scheme 3.1 Peptide bond formation**

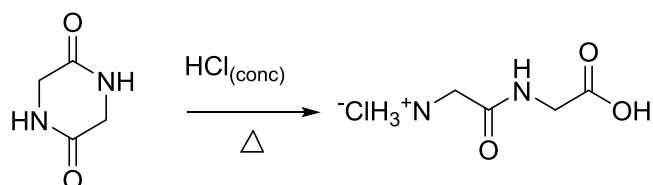
Current methods for peptide synthesis combine two steps. The first step is the activation of the carboxylic acid of one residue by the coupling reagent. The second step is the nucleophilic attack of the amino group of the other residue to form the peptide bond.<sup>81</sup>

Emil Fischer is acknowledged as the pioneer of the field of peptide synthesis, being the first to synthesise a dipeptide in 1901.<sup>82</sup> The field has had a century to evolve since, bringing forth innovations that allow the routine synthesis of peptides between forty to sixty amino acids in length, and proteins up to 200 amino acids in length through the chemoselective ligation of smaller peptide segments.<sup>83</sup> The following section will briefly discuss the historical developments in peptide synthesis, followed by recent methodologies developed. Finally, the total synthesis of tikitericin will be presented with bioassay results and characterisation.

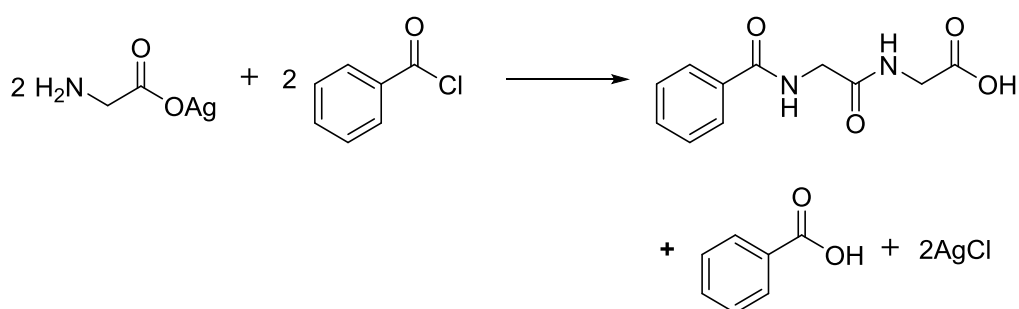
#### 3.2 Historical assessment of peptide synthesis

The beginning of peptide chemistry is considered to be when Emil Fischer synthesised the dipeptide glycylglycine.<sup>82</sup> This was produced through acid hydrolysis by boiling 2,5-diketopiperazine with hydrochloric acid (Scheme 3.2). However, the work of Theodor Curtius had been overlooked as 20 years earlier he had synthesised and characterised a glycylglycine

analogue. This was prepared through the treatment of the silver salt of glycine with benzoylchloride to produce benzoylglycylglycine (Scheme 3.3).<sup>84</sup>



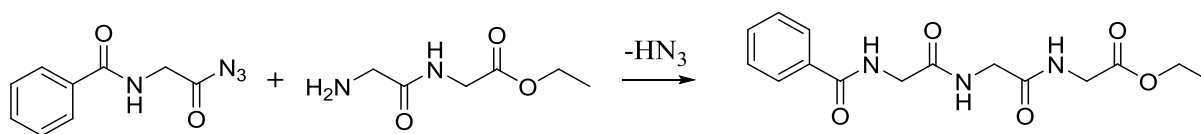
**Scheme 3.2 Emil Fischer's preparation of glycyglycine**



**Scheme 3.3 Theodor Curtius' preparation of benzoylglycylglycine**

### 3.2.1 Theodor Curtius' contributions to peptide chemistry

Curtius was not directly involved in peptide chemistry, but through his developments in inorganic chemistry he did further the field. His studies were directed towards diazo compounds and from this work he was the first to produce hydrazine in 1887,<sup>85</sup> which was to become of great importance in peptide chemistry later. In 1884, it was shown by Kiel that carboxylic hydrazides could be prepared from hydrazine and esters, which could be taken a step further by reacting with nitrous acid to form an azide.<sup>86</sup> Curtius used this idea to synthesise amino acid azides and use them as “activated” amino acids for the establishment of a peptide bond with a second amine group of an amino acid or peptide.<sup>87</sup> Curtius' method was the reverse of current peptide synthesis and consisted of building peptide bonds from the N- to C-termini (Scheme 3.4). He used glycine azide, which was protected at the N-terminus by a benzoyl group, and reacted this with glycyglycine ethyl ester. This process allows one peptide bond to form at a time because the C-terminus of the peptide chain is protected with the ester functional group so the azide can only react with the amine. The peptide elongation is allowed to continue by forming an azide from the ester and subsequent reactions can take place with more ester protected amino acids.



### Scheme 3.4 Theodor Curtius' azide coupling method

This method would have had much greater impact at the time if there was a way to remove the protecting groups utilised. Unfortunately, the first method for removing a benzoyl group came in 1965 when Leopold Horner demonstrated the use of a mercury cathode for electrolytic reduction.<sup>86</sup> Any attempts of removing the acetyl group at the C-terminus would have resulted in the destruction of internal peptide linkages as there was no method known for selective removal. Through much inspiration from Fisher's work, Curtius extended his method of peptide coupling to the synthesis of benzoylated peptides containing both aspartic acid and alanine. Soon after this work, he left the field of peptide chemistry in 1905.<sup>88</sup>

### 3.2.2 Emil Fischer's contributions to peptide chemistry

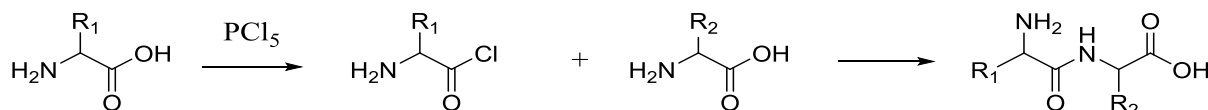
When Fischer entered the field in 1899, only 14 amino acids had been characterised. Prior to beginning peptide synthesis, Fischer submitted himself to studying amino acids as he had concluded that the field would not progress without a good understanding of the building blocks which form peptides. In this year of study he discovered *L*-proline and *L*-hydroxyproline leaving tryptophan, methionine and threonine that would be later discovered through isolation from hydrolysates of protein; asparagine and glutamine would have to be discovered by other methods. All amino acids discovered up to this point had synthetic routes developed for them but they were formed as racemates.<sup>86</sup>

Fischer initiated three goals at the beginning of the 20<sup>th</sup> century, the first being to develop a method to separate a complex mixture of amino acids from a hydrolysate and to quantitatively determine the amounts of each amino acid separated from the hydrolysate. To achieve this the hydrolysate mixture was dissolved in absolute ethanol and HCl was added yielding amino ester hydrochlorides. Sodium hydroxide was added to an aqueous solution of the salt and then saturated the solution with potassium carbonate. He then extracted the aqueous solution with diethyl ether to isolate the free amino esters. Fischer found that each amino ester had a different volatility so he could separate each by fractional distillation.<sup>89</sup>

The second goal Fischer had in mind was to develop a method to separate synthetic amino acids into their pure enantiomers. This was achieved through acylation of the amino group

through either a benzoyl or formyl residue. He then formed diastereomeric salts with optically active bases (generally alkaloids). The diastereomers differed in solubility in various solvents so could be separated through recrystallisation.<sup>89</sup>

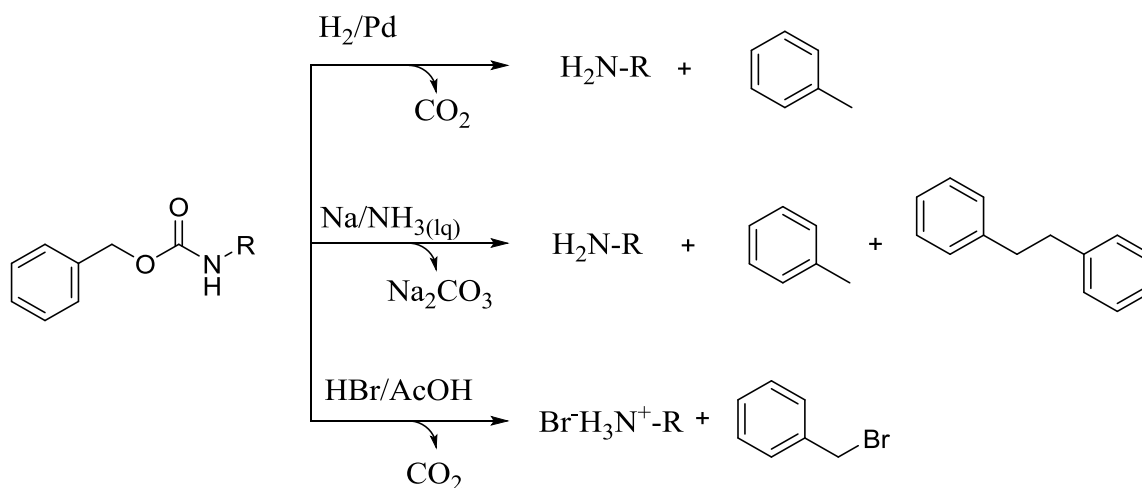
The third goal of Fischer was to develop amino acid derivatives with protecting groups which, once two amino acids were coupled, could be removed without disrupting the peptide bond that was the main issue in Curtius' azide method. However, he was not particularly successful in this area, and instead developed a method of peptide coupling using free amino acids and converting them to acid halides with phosphorus pentachloride. This was then reacted with a non-derivatised amino acid to form the peptide bond (Scheme 3.5).<sup>90</sup> The main problem with this method is that bond formation is not controlled. As the amine of the amino acid halide is not protected, self-condensation of two or more of the activated acid halides as a side reaction is unavoidable.



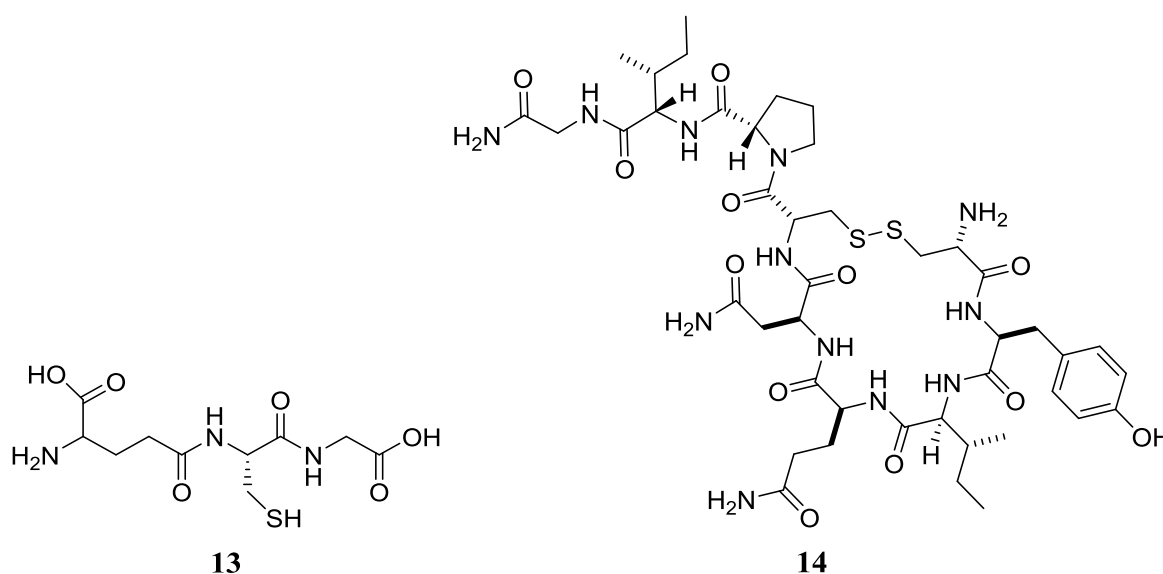
**Scheme 3.5 Fischer's acid halide peptide coupling**

### 3.2.3 Advancements in protecting group strategies

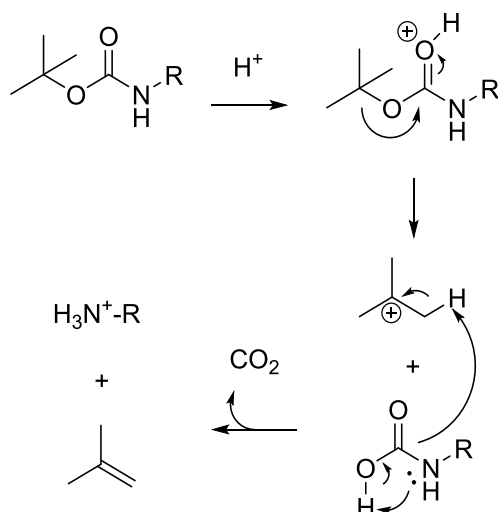
Comparing both Fischer's and Curtius' work, it is apparent that what prevented the field from advancing was a suitable protecting group that is labile in conditions that do not harm the established peptide bonds. In taking into account all of Emil Fischer's brilliant achievements in and outside of peptide chemistry, it is no surprise that one of his former students, M. Bergmann and his colleague L. Zervas, would breathe new life into the field of peptide chemistry by developing the carbobenzoxy group (Cbz) (Scheme 3.6).<sup>91</sup> This was used for the temporary protection of the amino functional group and is stable in the presence of acid. The discovery of the Cbz group inspired many to make numerous small peptides, an example of this being glutathione (**13**).<sup>92</sup> In a momentous effort twenty years later the active hormone oxytocin (**14**) was synthesised by du Vigneaud et al.,<sup>93</sup> leading to du Vigneaud being awarded the Nobel prize in chemistry two years later.



**Scheme 3.6 Cbz deprotection routes**



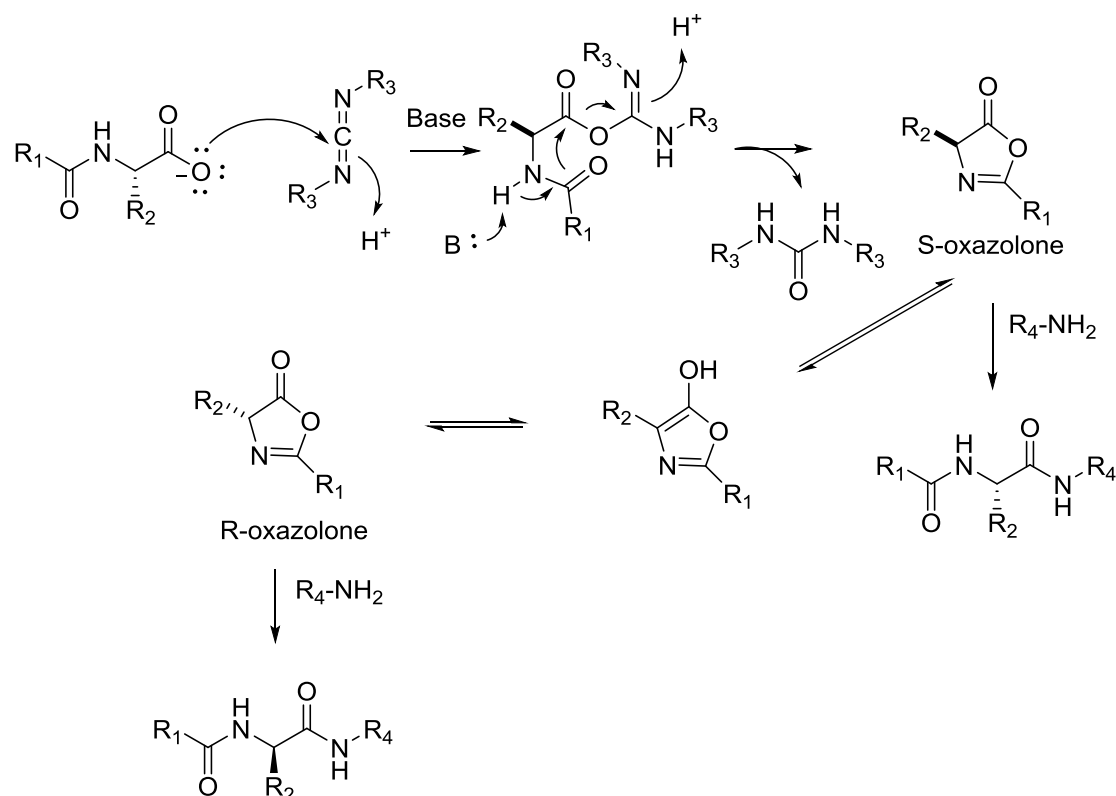
The year 1957 saw the introduction of a new protecting group that was acid labile, the *tert*-butyloxycarbonyl group (Boc) (Scheme 3.7). This new group was developed by F.C McKay and N.F Albertson and was fully orthogonal to the Cbz group, being stable to Birch reduction, hydrogenation and strongly basic conditions.<sup>94</sup> Boc opened up further opportunities for the field, as it was now possible to synthesise more complex peptides in conjunction with the Cbz group. Thus, 1963 saw the total synthesis of the 39-residue peptide  $\beta$ -corticotropin (adrenocorticotrophic hormone) by R. Schwyzler and P. Sieber.<sup>95</sup> This intensive work was executed by synthesising subunits of the peptide with suitable protecting groups and then by coupling the subunits together to complete the full peptide.



**Scheme 3.7 Boc deprotection**

### 3.2.4 Advances in coupling methods

The advancement of protecting groups also saw an effort to increase the number of effective coupling reagents available. The most significant of these additions was the carbodiimides, introduced in 1955 by J.C. Sheehan and G.P. Hess,<sup>96</sup> and independently by H.G. Khorana<sup>97</sup>. Although these reagents made a big impact at the time, issues also arose from their use. This mainly stemmed from the proclivity of carbodiimides to instigate racemisation during the activation step (Scheme 3.7). M. Goodman and L. Levine did studies on this racemisation process using benzoyl-*l*-phenyl-alanine. They followed the reaction polarimetrically and at the point of greatest negative rotation they removed the solvent and purified an oxazolone from the amino ester.<sup>98</sup> Furthermore, rates of reaction were followed spectrophotometrically and they found that the rate of oxazolone racemisation proceeds with first order kinetics while the ring opening reaction to form the non-racemisation product follows second order kinetics and is much slower than the racemisation process. M. Goodman and W.J. McGahren devised the mechanism of racemisation shown in Scheme 3.8.<sup>99</sup> Upon activation by the carbodiimide the oxazolone forms. Through keto-enol tautomerisation the chiral centre can invert and ring opening of the oxazolone yields the undesired enantiomer.

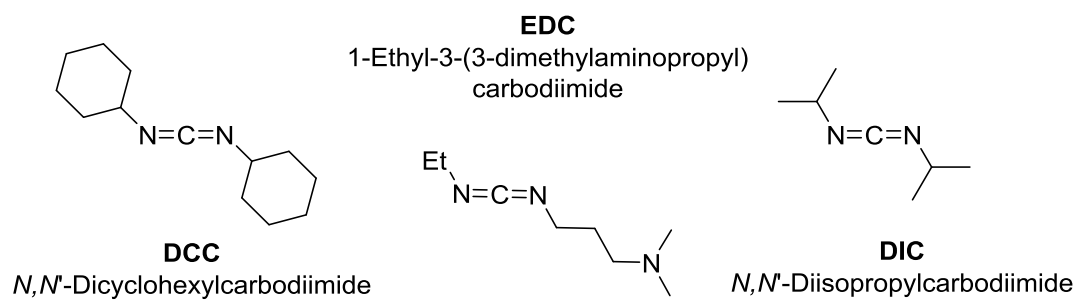


**Scheme 3.8 Carbodiimide coupling mechanism and racemisation**

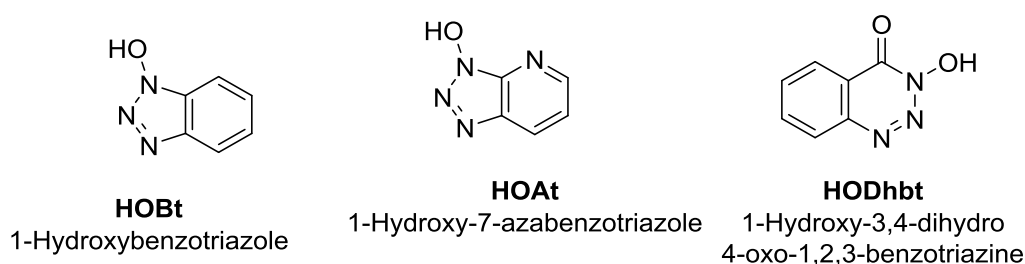
Although the carbodiimides are excellent coupling reagents, this racemisation was a clear issue. Because peptides are generally synthesised for biological purposes, it is important to have the correct stereochemistry at each coupling so steps need to be taken to ensure that the stereochemistry is controlled. The resolution to the issue of racemisation came with the introduction of additives in conjunction with the use of carbodiimides, the first being introduced by W. Koenig and R. Geiger in 1970 who used 1-oxo-2-hydroxydihydrobenzotriazine (HODhbt).<sup>100</sup> Over the years, numerous other coupling reagents and additives have been developed and the most common ones can be seen in Figure 3.2.



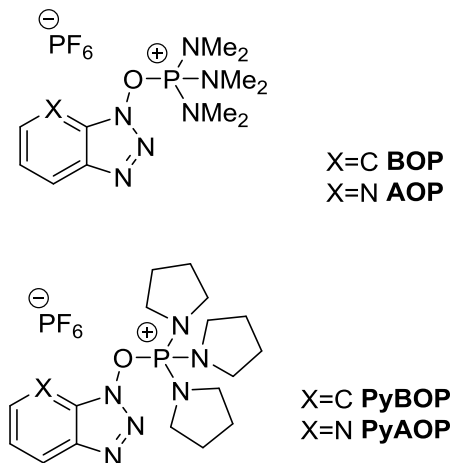
### Carbodiimides



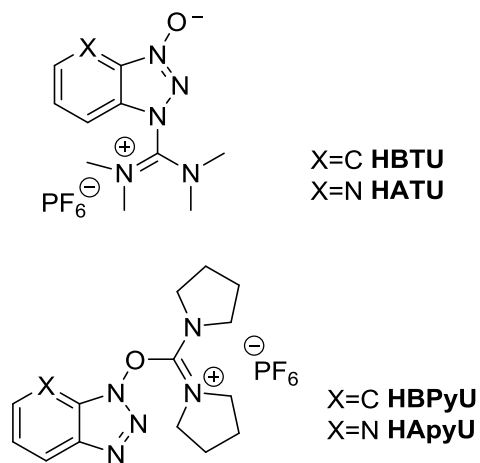
### Coupling additives



### Phosphonium reagents



### Uronium reagents



**Figure 3.2 Common coupling reagents**

Coupling additives have a number of positive influences on the coupling step. First, they prevent the racemisation process induced through activation by the carbodiimide by substituting in for the carbodiimide, an example of this can be seen in section 3.3.2.4 with HOBt and DIC. Coupling additives also positively influence the coupling by improving reaction efficiency. The nature of these additives is electron withdrawing, so when they form the activated ester, as seen in Scheme 3.10, electron density is drawn from the carbonyl carbon

increasing its reactivity towards nucleophiles. The more electron withdrawing heteroatoms in the ring of the coupling additive, the more active the ester is towards the coupling.<sup>101</sup>

### 3.2.5 The development of SPPS

In spite of the fact that solution phase peptide synthesis was proving to be successful, there was a clear pitfall in this method. Solution phase synthesis of peptides is incredibly time consuming, as each successive deprotection and coupling step requires purification. Another issue surrounding this procedure is that it is not suitable for synthesis of long polypeptides as technical difficulties with solubility becomes challenging with each amino acid addition. The resolution to these issues came in 1963 when R. Merrifield introduced Solid Phase Peptide Synthesis (SPPS).<sup>102</sup> This method saves a lot of time, as the growing peptide is tethered to an insoluble support. After each reaction step the reactants are simply removed by filtration and washing. As this is incredibly repetitive, this process can be easily automated.

Merrifield used a styrene-divinylbenzene co-polymer as his solid support and functionalised this by chloromethylation.<sup>102</sup> The resulting resin was treated with the triethylammonium salt of the first protected amino acid in the chain of the proposed peptide to give a benzyl ester. Once the first amino acid was successfully anchored to the resin, the peptide was elongated using *N,N'*-dicyclohexylcarbodiimide as a coupling reagent. Once the peptide was completed, the peptide was cleaved from the resin with liquid hydrogen fluoride (HF). Hydrogen fluoride is not the most suitable reagent for SPPS, as there are a number of side reactions which it can induce during the cleavage step. HF can catalyse Friedel-Crafts reactions between the resin and the aromatic R-groups of the peptide. For peptides containing serine or threonine HF also promotes N-O acyl shifts, replacing the amide linkage with an ester linkage.<sup>103</sup> Tam et al. introduced a protocol using low concentrations of HF in dimethyl sulphide to minimise these side reactions.<sup>103</sup>

1970 saw the development of a new amino-protecting group by L. Carpino and G. Han.<sup>104</sup> The 9-fluorenylmethoxycarbonyl (Fmoc) group is a completely different protecting group to the other amino protecting groups as it is alkali-labile meaning that orthogonal acid-labile protecting groups can be used to protect the amino acid side chains.

Further developments of SPPS have been abundant since it was first invented. As the process is largely repetitive, machine assisted peptide synthesis can easily be done and some have been shown to couple up to 75 residues per day.<sup>105</sup> Furthermore, numerous resins and linkers

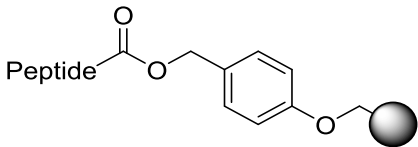
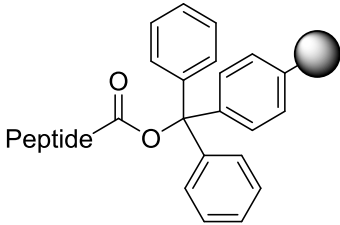
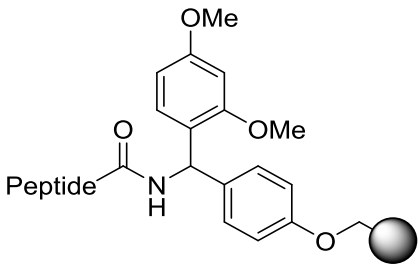
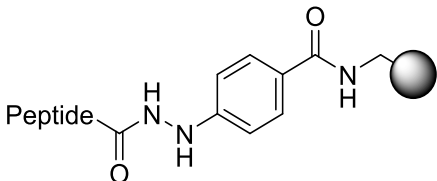
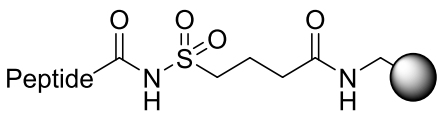
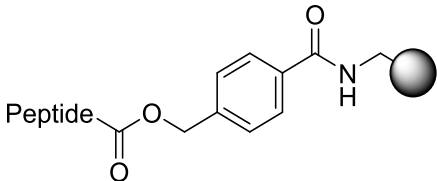
Linker structure	Resin	Cleavage
	Wang resin	95% TFA
	2-Chlorotrityl resin	1% TFA in DCM
	Rink amide resin	95% TFA
	4-Fmoc-hydrazinobenzoyl resin	$\text{Cu}(\text{OAc})_2$ or $\text{RNH}_2$ , $\text{ROH}$ , pyridine
	4-Sulfamylbutyryl resin	(1) $\text{TMSCH}_2\text{N}_2$ (2) $\text{RSH}$
	HMBA resin	$\text{NaBH}_4/\text{EtOH}$

Table 3.1 Resins for Fmoc SPPS - Solid supports are indicated by the bead. These are either crosslinked polystyrene or polyethylene glycol-based polymers. (Adapted from Kimmerlin)<sup>106</sup>

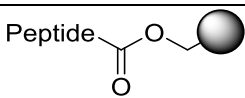
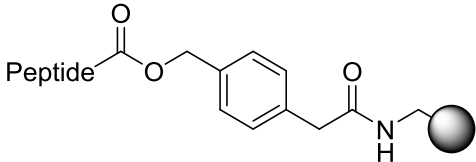
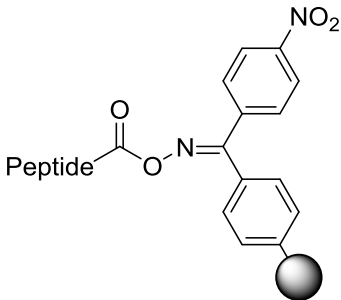
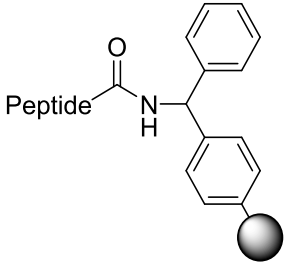
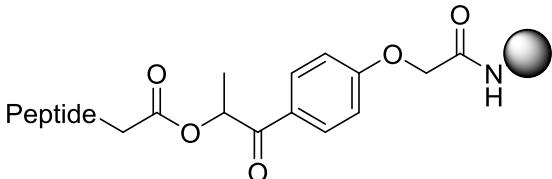
Linker structure	Resin	Cleavage
	Merrifield	HF, TFMSA
	PAM	HF, TFMA
	Oxime	NaOH/dioxane
	MBHA resin	MeOH/DMF/TEA
	Brominated PPOA resin	NH <sub>2</sub> NH <sub>2</sub> /DMF

Table 3.1 continued Resins for Boc SPPS - Solid supports are indicated by the bead. These are either crosslinked polystyrene or polyethylene glycol-based polymers. (Adapted from Kimmerlin)<sup>106</sup>

have been developed for both Boc and Fmoc strategies, these can be seen in Table 3.1 with cleavage strategies for each.

There are a number of different methods implemented in SPPS. Manual SPPS was the first method to be implemented and this will be explained more in-depth in section 3.3. The other two options available are “Quasi continuous flow” and fully automated SPPS.

#### **3.2.5.1 “Quasi-continuous flow”**

“Quasi-continuous flow” is a method of SPPS where the solid support is packed into a column and the reagents and solvents are delivered to the column by a pump. As the solid support has to withstand increased pressure and also retain the same volume, the standard resins are not suitable for this method of synthesis because these swell and contract depending on the solvent.<sup>107</sup> Because of this, composite materials made from rigid supports have been utilised for quasi-continuous flow SPPS. In quasi-continuous flow SPPS, only the Fmoc strategy can be used as the ionic species formed during the cleavage of Boc groups have effects on swelling due to electrostatic forces.<sup>108</sup>

#### **3.2.5.2 Fully automated SPPS**

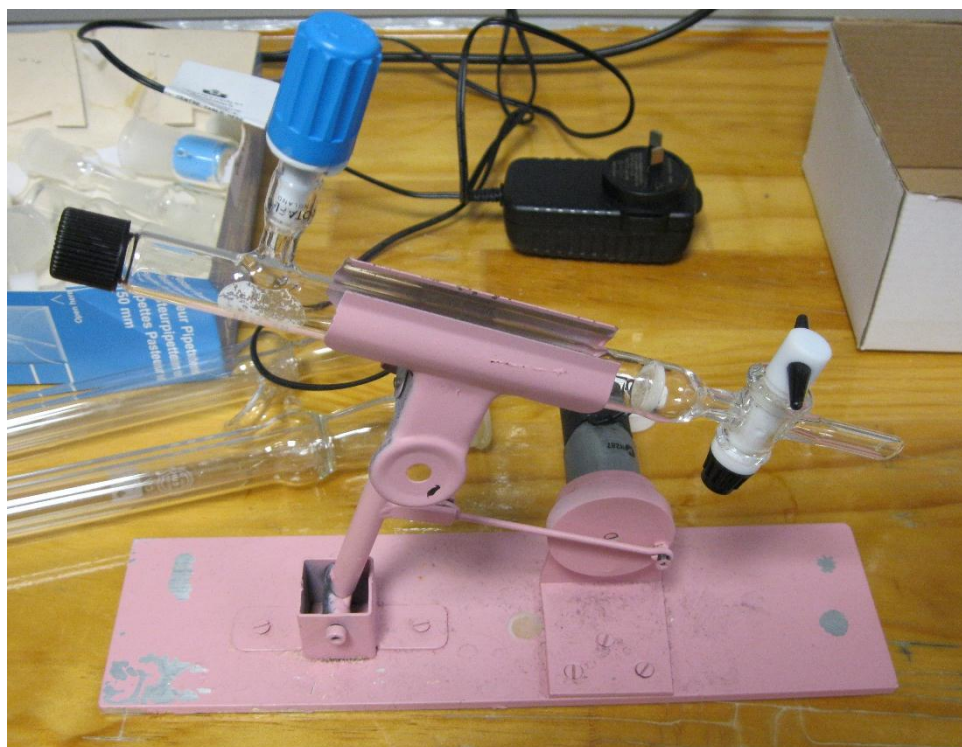
There are numerous fully automated SPPS systems commercially available for both batchwise and continuous flow.<sup>109</sup> Not all machines are built equally though, and there is a broad range available for different needs. Automated SPPS utilising microwave irradiation has been successfully introduced for commercial use as well, these have increased coupling rates and no racemisation is detected when this technique is implemented.<sup>110</sup> When operating under Fmoc SPPS conditions, the automated synthesiser can monitor both coupling and deprotection steps and optimise the time spent on each step for efficiency. This is done by monitoring the chromophores which are released during deprotection (free Fmoc)<sup>111</sup> and also during coupling (HOBt or HOAt).<sup>112</sup>

### **3.3 The attempted total synthesis of tikitericin**

#### **3.3.1 Instruments used for synthesis**

Access to an automated SPPS instrument was not available. A mechanical peptide shaker (Figure 3.1) was custom built at Victoria University of Wellington for manual SPPS. As shown, the shaker consisted of a rotary disc connected to an arm to allow for a rocking motion for gentle agitation of the peptide resin during coupling and deprotection steps. The arm had a clip in place to allow for the securing of the reaction vessel during reaction steps. The reaction

vessel had a glass sinter to allow for drainage of solvents and reactants while retaining the resin. Some reactions needed to be done under inert conditions, thus a stopcock was fitted to allow for introduction of argon. This also allowed for application of sufficient airflow to drain solvent from the resin to dryness.



**Figure 3.1 Mechanical shaker and reaction vessel**

### **3.3.2 General methodologies and materials of synthesis**

#### **3.3.2.1 Fmoc strategy**

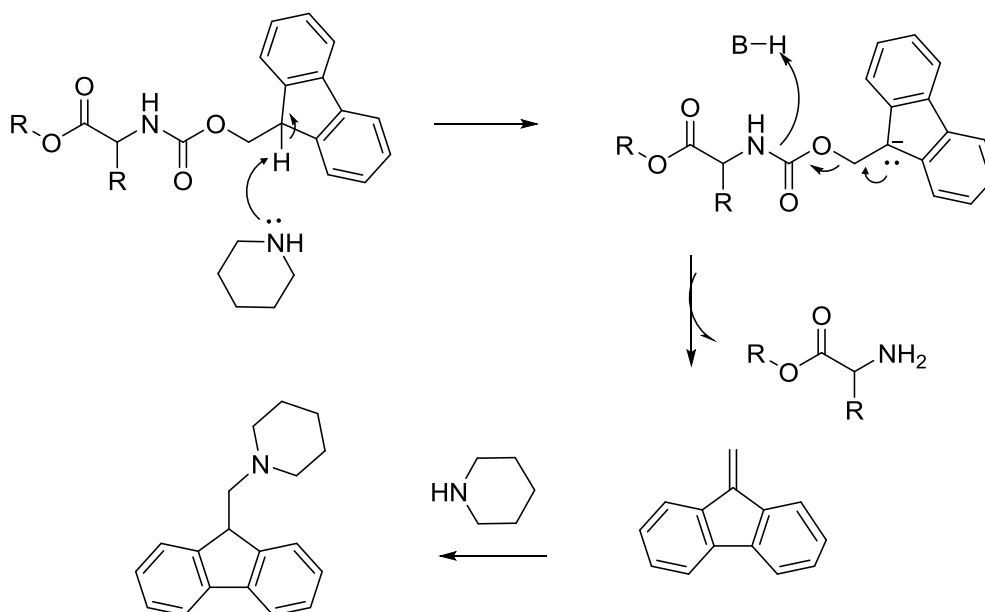
The methods employed during the synthesis of tikitericin are as follows. An Fmoc strategy was utilised in the synthesis. A Boc strategy was not suitable for the synthesis for a number of reasons. The overwhelming reason for this was that we did not have suitable instruments for this, as cleavage protocols for the Boc strategies generally involve the use of HF which requires special vessels that are stable towards its toxicity and corrosiveness. The alternative deprotection protocol for Boc involves the use of TFA, which after repetitive exposure to the peptide has a potential to damage sensitive peptide bonds and can also lead to acid catalysed side reactions.<sup>113</sup> It would also not be compatible with the orthogonality instilled in the (Me)Lan starting materials, as exposure to TFA could have the potential to deprotect the allyl group over continuous exposure to acid. Thus, the Fmoc strategy was chosen because it was both more convenient and also more compatible with the starting materials.

### 3.3.2.2 Wang resin

The resin chosen for use was a Wang resin with  $0.1 \text{ mmol/g}^{-1}$  loading. As will be discussed, there is a potential for dimerization during the ring closing steps of the synthesis. It was suitable that a low molar loading of the resin was used to prevent this from occurring. As leucine is at the C-terminus of tikitericin, the resin came preloaded with Fmoc-leucine. The reason why a preloaded resin was ordered was not only does it save time in the synthesis but also the anchoring of an amino acid to the support can prove difficult. The esterification reaction required for this is difficult due to the fact that it can lead to racemisation, dipeptide formation and low substitution rates, this coupled with the hazards involved in this reaction make this an unattractive endeavour.<sup>114</sup> Fortunately many manufacturers supply preloaded resins and these are quality assessed to ensure that the product is in line with its specifications.

### 3.3.2.3 Fmoc deprotection

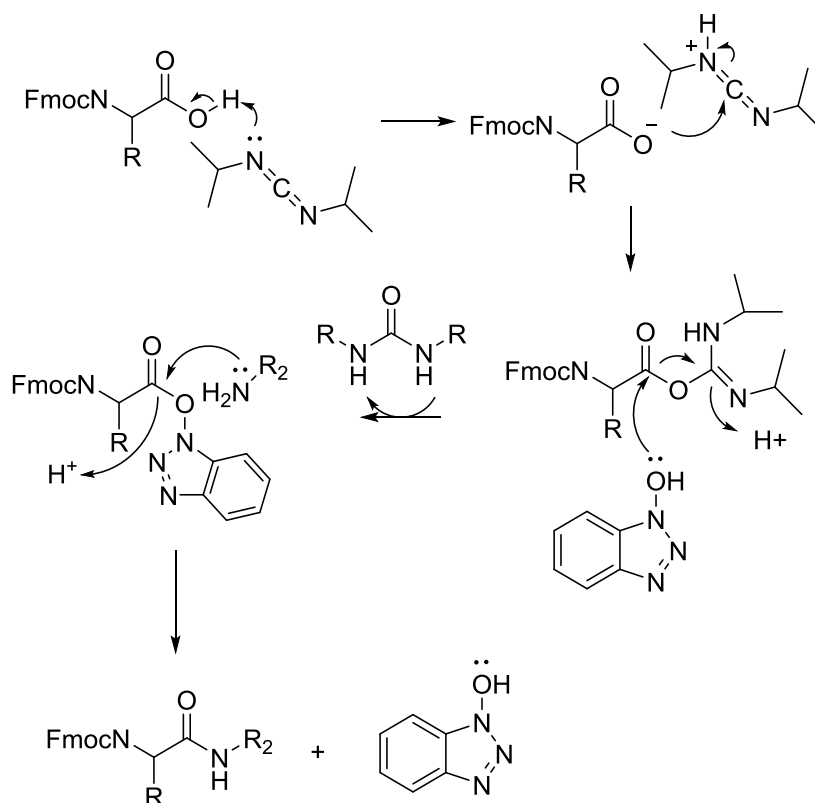
The deprotection protocol that was implemented was 20% piperidine in DMF, which is the general protocol for the Fmoc strategy. Piperidine extracts the acidic proton of the fluorene ring system (Scheme 3.9) and  $\beta$ -elimination proceeds to give a highly reactive dibenzofulvene intermediate. This can cause issues as it can react with protecting groups on the side chains of the amino acids. Piperidine acts as an allyl scavenger in this case and forms a stable adduct with the dibenzofulvene intermediate, preventing side reactions.<sup>114</sup>



**Scheme 3.9 Deprotection of Fmoc using piperidine**

### 3.3.2.4 Coupling reagents and protocol

*N,N'*-Diisopropylcarbodiimide (DIC) was used in tandem with hydroxybenzotriazole (HOBt) as a coupling additive. These two are used together to prevent the racemisation induced by DIC when used alone. The Fmoc-amino acids were prepared in a 1:1 DCM/DMF solvent mixture with the HOBt. This solution was then introduced to the peptide resin, and DIC was introduced after that. This order of addition further prevents racemisation. In the coupling mechanism (Scheme 3.10), DIC extracts a proton from the carboxylic acid of the free Fmoc-amino acid. Nucleophilic attack from the generated carboxylate onto the carbon of the carbodiimide forms an O-acylurea. HOBt reacts with the acylurea, a urea leaves as a by-product and an OBt ester is formed. The OBt essentially enhances the reactivity of the activated ester by stabilising the approach of the amine through hydrogen bonding. The amine reacts with this activated ester and an amide is formed with HOBt leaving.<sup>115</sup>



**Scheme 3.10 Coupling mechanism with DIC and HOBt**



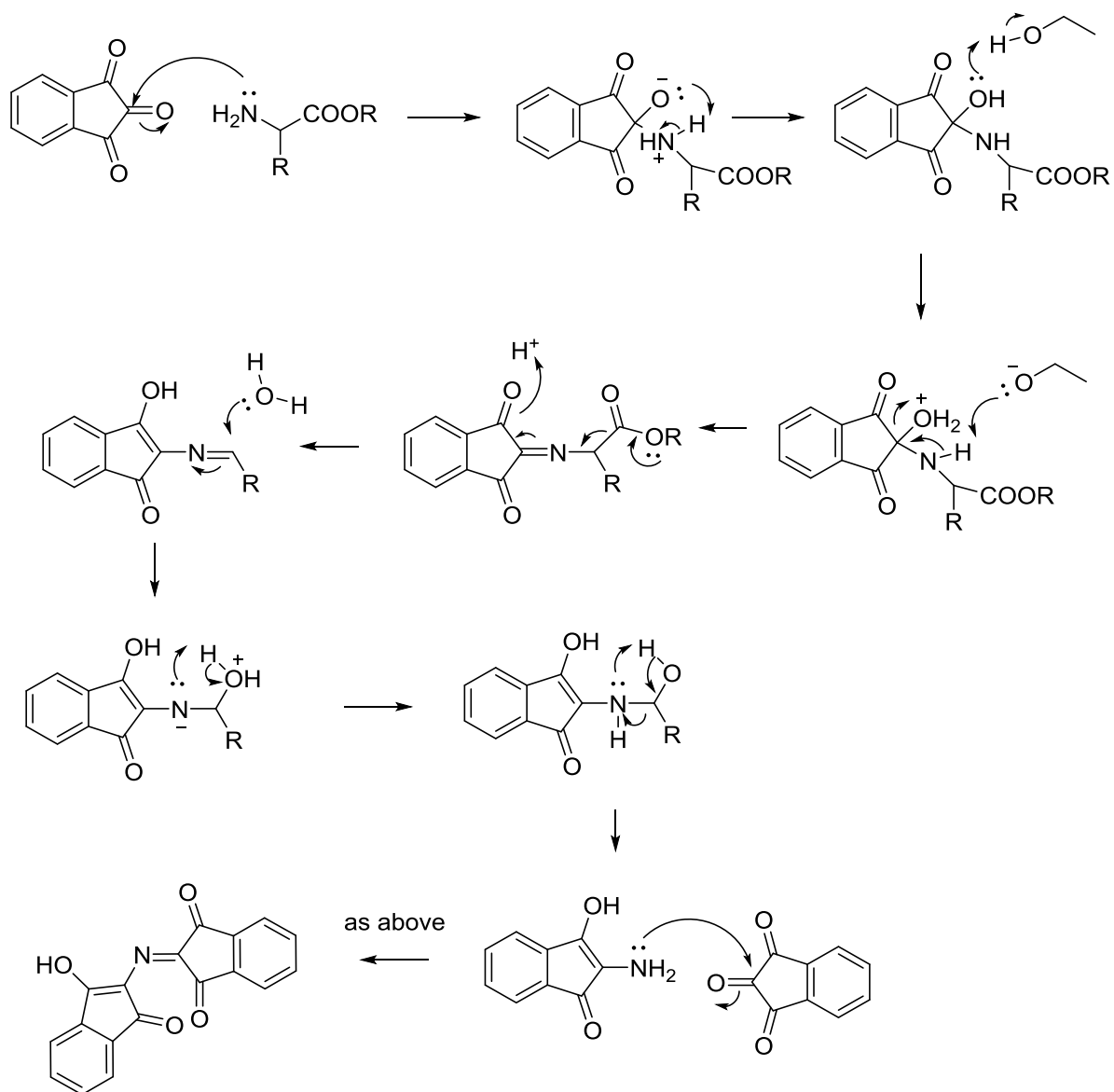
### 3.3.2.5 Washing protocol

Probably the most important step for successful peptide synthesis is the washing step. This is needed to make sure that all reagents and by-products are removed before initiating the next step in the synthesis. This is needed because if there is any piperidine left after the deprotection step, for example, then there is potential for deprotection of the Fmoc and formation of peptide dimers during the coupling step. In general, the resin was washed with DMF and left to soak for five seconds and then drained from the resin. This was repeated five times. For the (Me)Lan residues they were also washed five times with DCM. DCM and DMF was used for washing after the allyl/alloc cleavage and this was repeated ten times each.

DMF was used to rinse the cap of the reaction vessel, this was done to remove any excess reagent and also to clear it of any beads which had been trapped in the thread of the cap.

### 3.3.2.6 Kaiser test

The Kaiser test was developed in 1970 by E. Kaiser.<sup>116</sup> This is a qualitative test for the presence of primary amines, thus it is performed at the end of a coupling step as it gives an indication of whether or not the coupling is complete. The materials involved for this test are ninhydrin in ethanol, potassium cyanide in pyridine and phenol in ethanol. The phenol has been found to increase the rate in which colour is formed and the potassium cyanide creates a reducing environment allowing the reaction to proceed. Ethanol is a source of protons and pyridine also creates a reductive environment and prevents the formation of hydrogen cyanide in the stock solution.<sup>117</sup> After a coupling step, a few resin beads are taken and placed in a small test tube. The Kaiser test solutions are added to the resin and this is heated to 95 °C. If the coupling is incomplete then both the beads and solution will turn blue, which is visible by eye. This is due to the reaction of ninhydrin with a primary amine forming diketohydrindylidenediketohydrindamine (Scheme 3.11).<sup>118</sup> The amine attacks at the central carbonyl of ninhydrin. The oxygen atom of the carbonyl is eliminated leaving as water and an imide is formed. Decarboxylation and rearrangement leads to another imine being formed. Nucleophilic attack at the carbon of the imine and rearrangement forms an alcohol. This leaves as an aldehyde and a primary amine is formed. This condenses with another ninhydrin molecule to form the blue highly conjugated diketohydrindylidenediketohydrindamine.

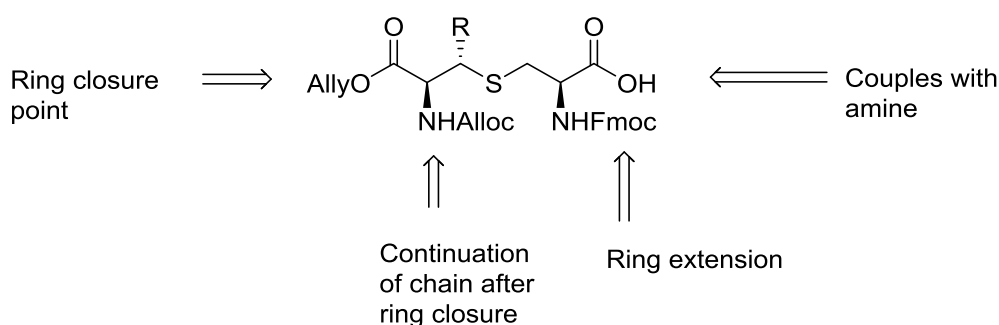


**Scheme 3.11 Kaiser test reaction mechanism**

### 3.3.3 Synthetic steps towards the production of tikitericin

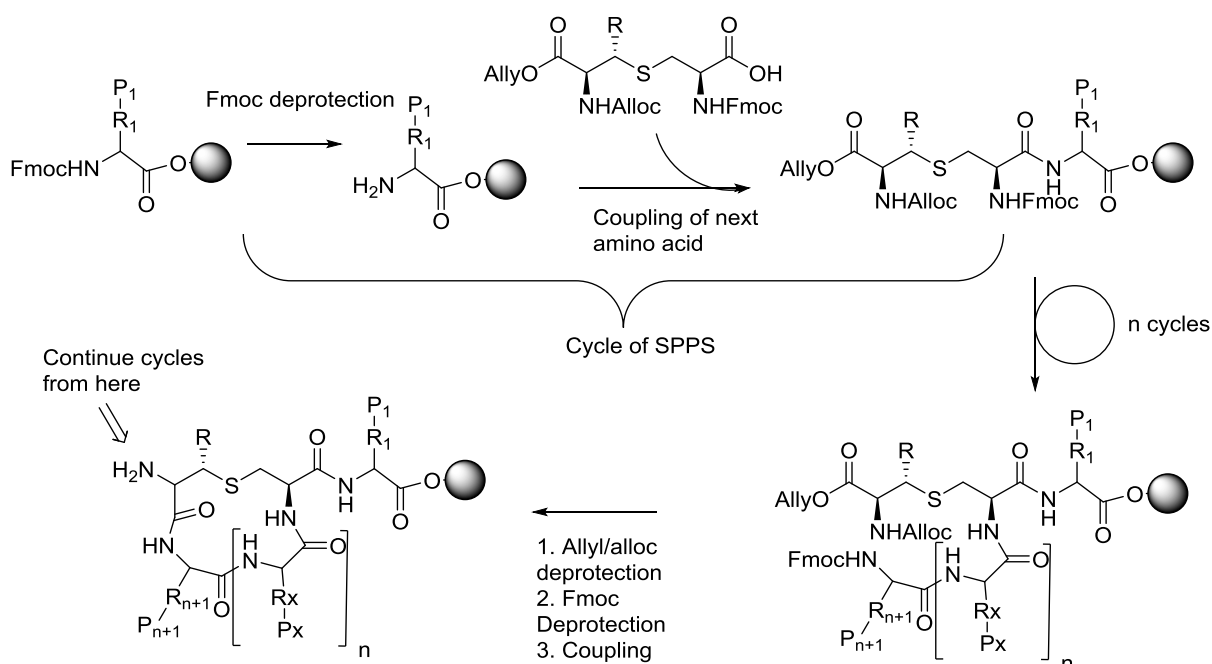
The first step of the synthesis was to soak the resin in DCM for 30 minutes. This helps to swell the resin and is essential for the beginning of the synthesis as it helps increase the coupling efficiency.<sup>119</sup> This step is pertinent as 99% of the coupling sites are found inside the resin beads, so this is needed for opening up the sites for access for reagents.<sup>114</sup> As the Wang resin was preloaded with Fmoc-Leu-OH, an Fmoc deprotection was then performed. The resin was then washed five times with DMF to remove any excess piperidine. After this was completed the first Melan residue was coupled with the resin. A coupling time of 3 hours was used for this as the starting material was limited, so repetition of couplings was unfavourable as it would take

time to recover the starting material. An outline of the use of each protecting group on the (Me)Lan Dimers can be seen in Figure 3.2.



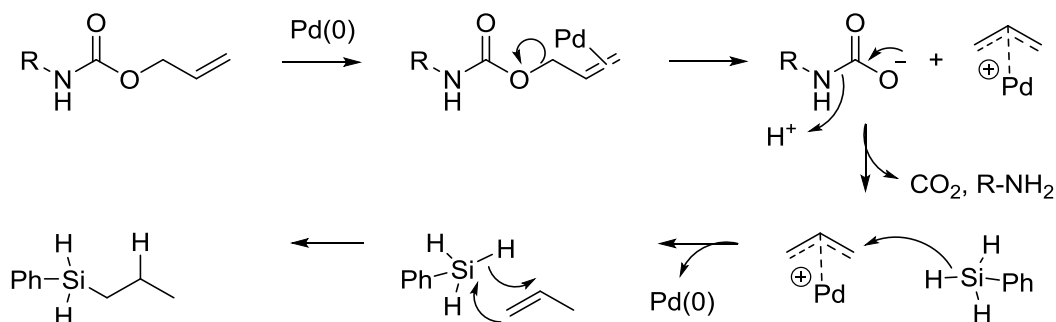
**Figure 3.2 Outline of purpose of each protecting group**

The uses of these protecting groups can be shown in tandem with a simple explanation of the standard protocol from here on in (Figure 3.3). Each amino acid has a side chain that may have potential for reacting with the coupling reagents as well. The amino acids which were used that have this potential are histidine, glutamic acid, aspartic acid, asparagine and glutamine. The acids were protected with a *tert*-butyl group and the amides and histidine was protected with a trityl group. These groups were used as they are acid labile so they are removed with TFA in tandem with the peptide being cleaved from the resin. An SPPS cycle is as follows: Fmoc group is deprotected, the resin was washed and then the next amino acid in the sequence was coupled with the current one. With the (Me)Lan derivatives the chain is built off the amine which is Fmoc functionalised. The desired number of amino acid residues in the ring were coupled and then the allyl and alloc groups were deprotected. The Fmoc group of the last amino acid added was deprotected and then the ring was closed by coupling the primary amine of the previously coupled residue with the recently deprotected carboxyl group.



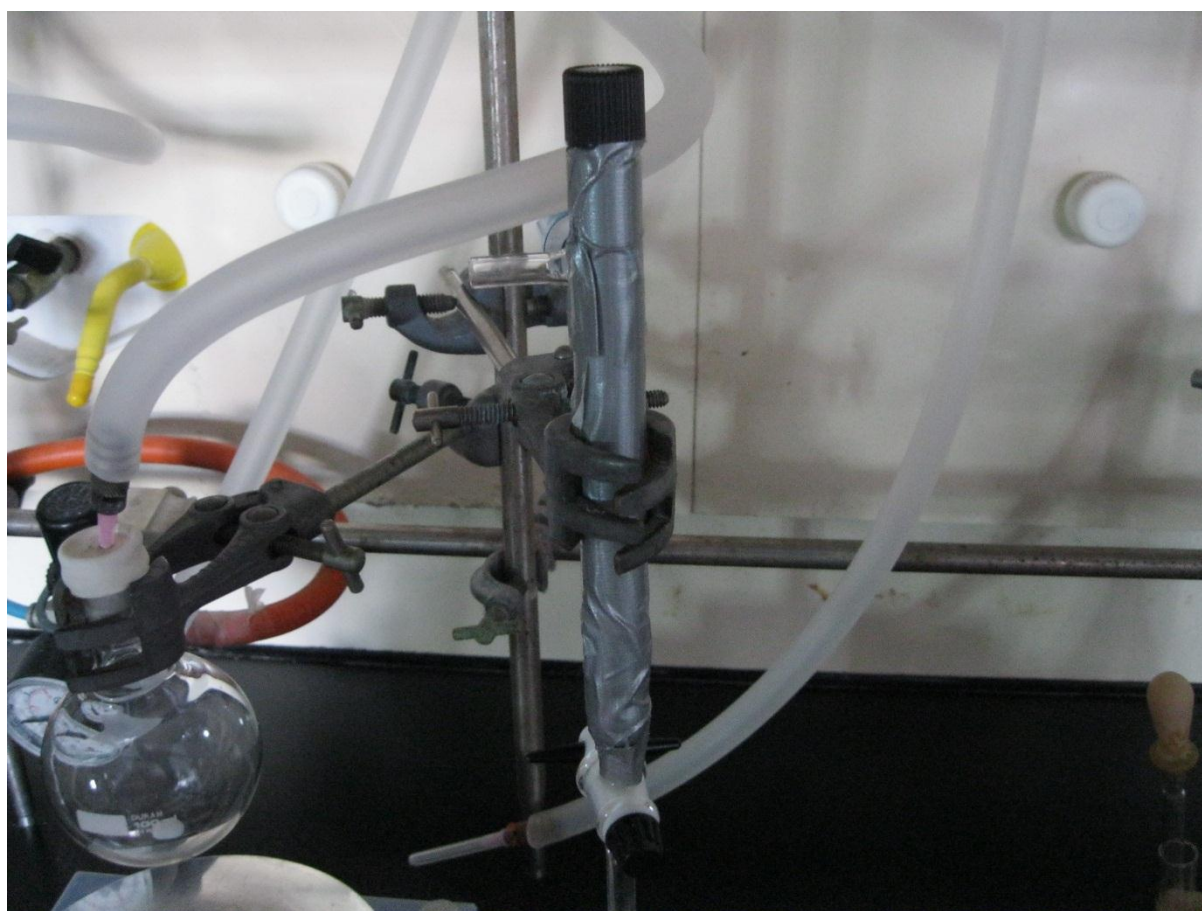
**Figure 3.3 Principles of SPPS**

Once it was confirmed that the coupling had completed by a negative Kaiser test, the next SPPS cycle was performed. The peptide chain was elongated using HOBt/DIC in the following order of amino acids: Fmoc-Gly-OH, Fmoc-Gly-OH, Fmoc-Ala-OH, Fmoc-Ala-OH, Fmoc-Asn(Trt)-OH. The alloc/allyl groups were then deprotected using  $\text{Pd}(\text{PH}_3)_4$  (Scheme 3.12). Phenylsilane was used as a source of hydride and also an allyl scavenger to prevent side reactions with the protecting groups on the side chains. The first step of the mechanism is displacement of the carboxylate forming a carbonate ion and a  $\pi$ -allyl palladium complex. In reference to the alloc, this goes through further rearrangement leading to decarboxylation and a primary amine being exposed. Hydride attack from the phenylsilane on to the  $\pi$ -allyl complex forms propene and regenerates the palladium catalyst. Another molecule of phenylsilane scavenges the propene through a hydroboration type mechanism.<sup>120</sup>



**Scheme 3.12 Mechanism of palladium catalysed deallylation**

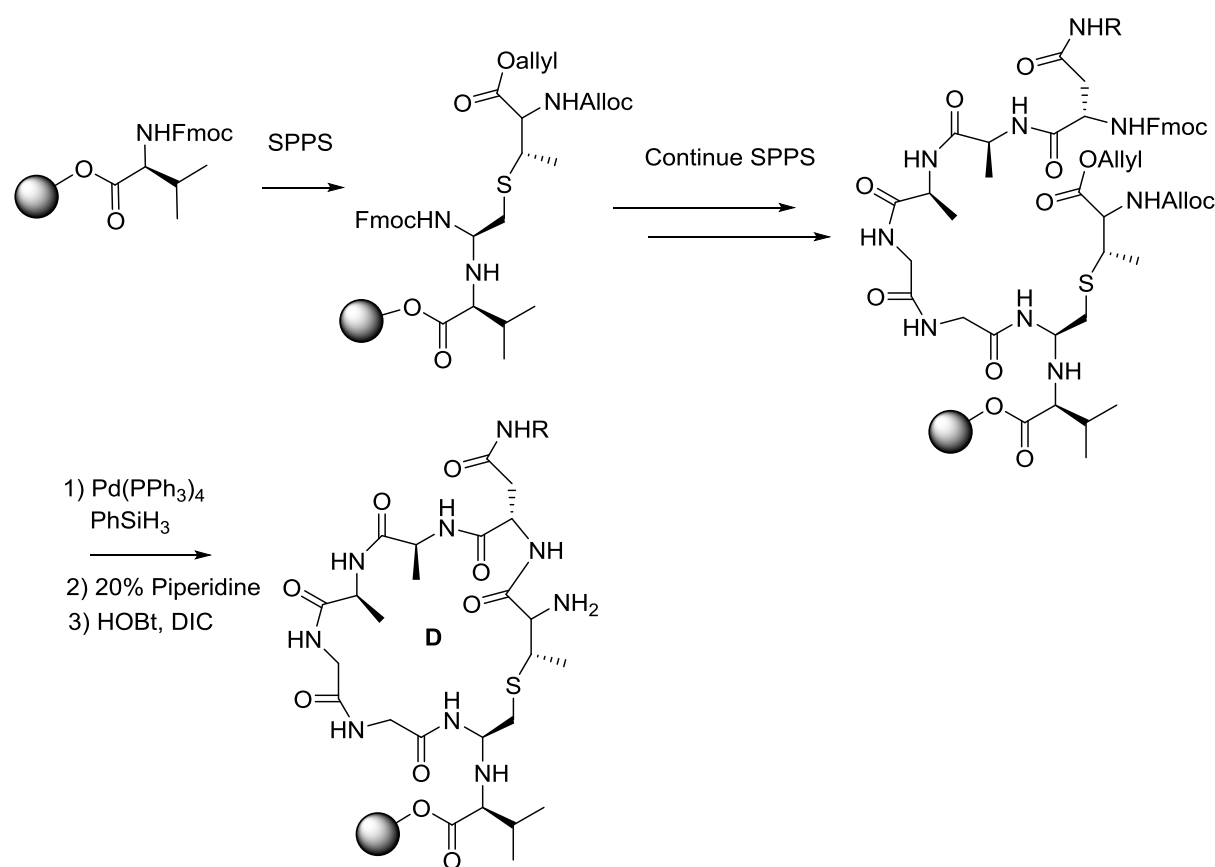
This deallylation reaction needed to be done in the dark as it was light sensitive. Tin foil has never been overly efficient at blocking out light as it is generally difficult to get full coverage over the reaction flask. Because of this it was decided that a new method of protecting the flask from light would be trialled. The most effective method trialled was using tape to cover the reaction vessel (Figure 3.4). The advantage of tape is its adhesive properties meaning that it sticks to the reaction vessel. Silver duct tape was chosen as it had a large degree of opacity. This was cut into small strips and wrapped around the reaction vessel. This was a much more streamlined method compared to tin foil and was effective at blocking out light.



**Figure 3.4 Duct tape method**

The allyl deprotection needed to be done prior to the deprotection of the Fmoc group of the previously coupled peptide. This is because if this was deprotected then there is increased potential for allylation of the amine. This is highly unfavourable because the synthesis would not be able to continue, as the amino group would be capped.<sup>121</sup> This is a potential side reaction because as is seen in Scheme 3.12, the  $\pi$ -allyl palladium complex is electrophilic, so a nucleophilic amine could just as easily attack and form a bond just as the hydride does.

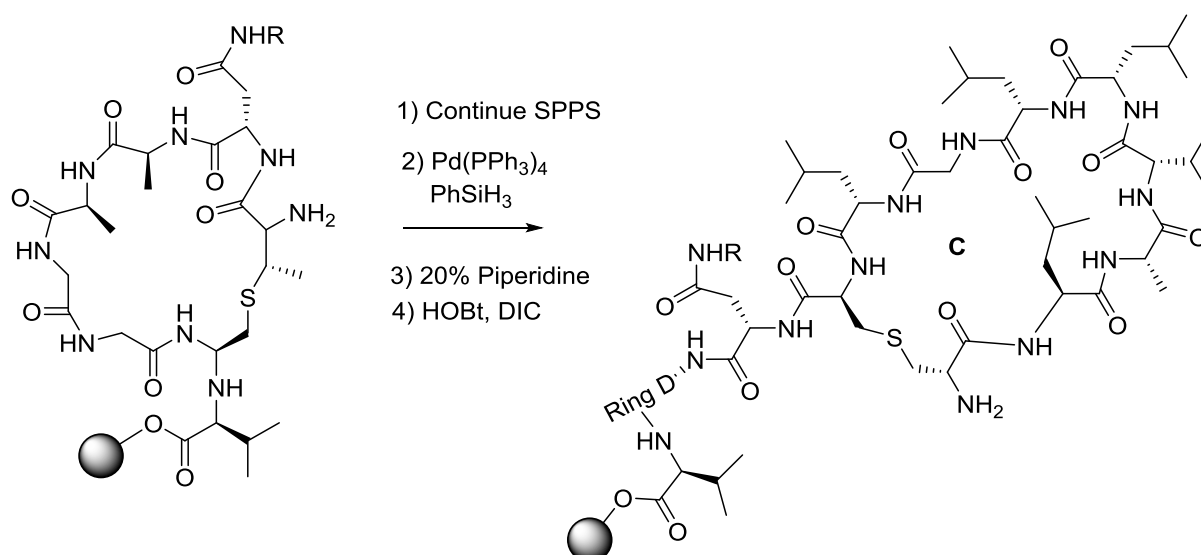
Once the allyl deprotection was completed, the Fmoc group was deprotected and the closure of ring D was attempted under standard coupling procedures. A small sample of resin was taken and the peptide was cleaved off the resin for analysis by HRMS. The molecular formula of  $C_{27}H_{45}N_9O_{10}$  for the closed peptide ring and  $C_{27}H_{47}N_9O_{11}$  for an open ring were both screened for but the ions detected equating to these formulae showed low ion count, and it was assumed that these were just noise. There was no evidence of dimerisation found either. This showed that there was no evidence of the coupling proceeding and it was assumed that the timing was not optimised. Because of this it was decided to increase the equivalents of coupling reagent and the coupling time to overnight. Because of the time pressure, the synthesis was carried on and ring D was assumed to be formed (Scheme 3.13).



**Scheme 3.13 Solid phase synthesis for forming of ring D**

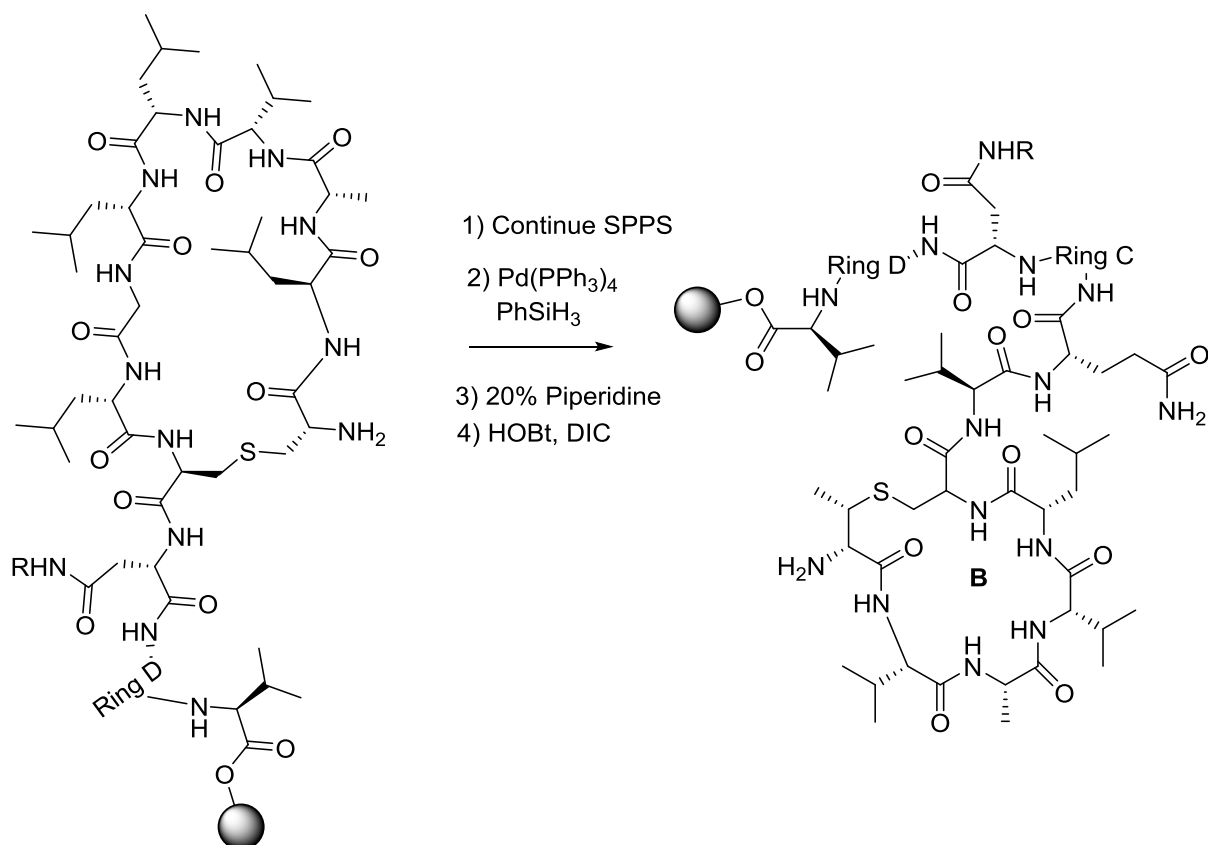
The remaining synthesis was then attempted as follows. Using standard protocols, Fmoc-Asn(Trt)-OH was coupled followed by the Lan. Ring C was then constructed in the following manner. Fmoc-Leu-OH, Fmoc-Gly-OH, Fmoc-Leu-OH, Fmoc-Leu-OH, Fmoc-Val-OH, Fmoc-Ala-OH, Fmoc-Leu-OH were all coupled sequentially. The allyl groups were deprotected followed by the Fmoc and then a coupling to close the ring was done overnight.

To check that the coupling was completed a small sample of resin was cleaved and analysed by LCMS. Both  $C_{71}H_{120}N_{20}O_{21}S_2$  and  $C_{71}H_{122}N_{20}O_{22}S_2$  were screened for and once again only low counts of  $C_{71}H_{120}N_{20}O_{21}S_2$  was detected and this appeared to be noise. Even though this result did not satisfy it was assumed that the equivalents used and coupling time was working for the ring closure, and because of time constraints the synthesis was continued. Thus, ring C was assumed to have been formed (Scheme 3.14).



**Scheme 3.14 Solid phase synthesis for forming of ring C**

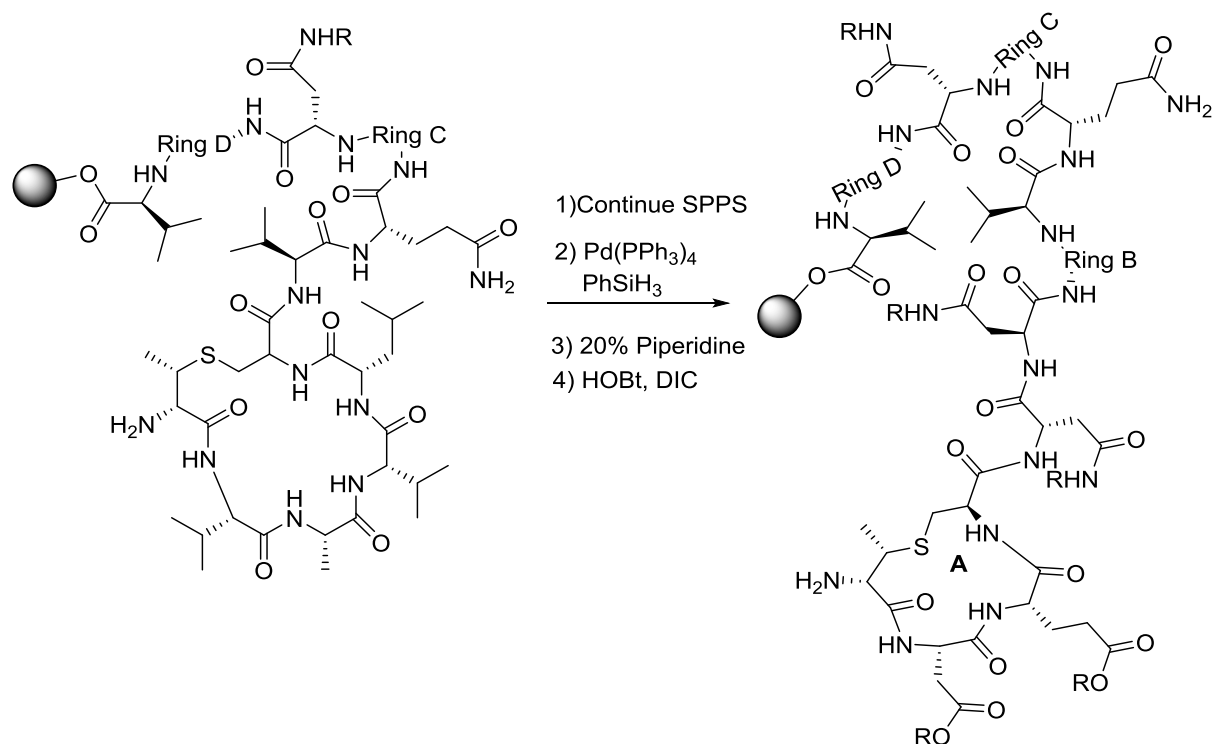
Following on from this, Fmoc-Gln(Trt)-OH was coupled followed by Fmoc-Leu-OH. The next MeLan was coupled into the structure and the standard protocols were used to couple Fmoc-Leu-OH, Fmoc-Val-OH, Fmoc-Ala-OH, Fmoc-Val-OH followed by the same protocol to close the ring. Thus ring B was assumed to be completed (Scheme 3.15).



**Scheme 3.15 Solid phase synthesis for forming of ring B**

Once the closing of ring B had completed, Fmoc-Asn(Trt)-OH was coupled followed by Fmoc-Asn(Trt)-OH. The last MeLan precursor was then added into the structure following the standard protocols. This was followed by coupling Fmoc-Glu(OtBu)-OH and Fmoc-Asp(OtBu)-OH respectively. After the conjectured success of these couplings, the ring was then closed following the standard protocol discussed previously. Thus ring A was assumed to be completed.

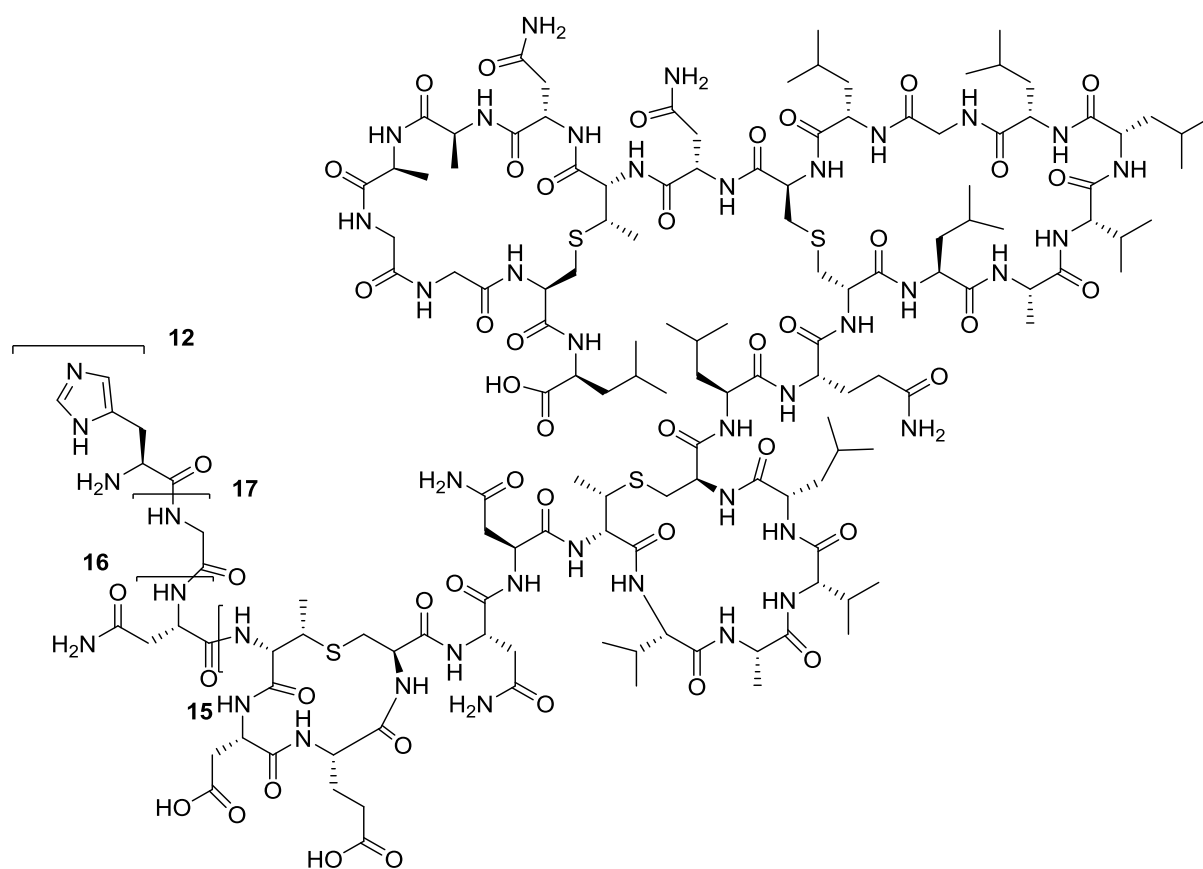




**Scheme 3.16 Solid phase synthesis for forming of ring A**

The next step of the synthesis was cleaving truncated versions of tikitericin as Emma Aitken found mass spectral evidence for the presence of these peptides in cultures of T81 and it is of interest to check the bioactivity of these in comparison to the full peptide of tikitericin.<sup>55, 122, 123</sup> Thus, tikitericin D (**15**) was assumed to be synthesised. Fmoc-Asn(Trt)-OH was then coupled with the remaining resin and tikitericin C (**16**) was assumed to be formed. Fmoc-Gly-OH was coupled with the remaining resin and tikitericin B (**17**) was assumed to be formed. A small amount of resin was taken from each of these steps and the assumed peptides 15-17 were cleaved from the resin. Fmoc-His(Trt)-OH was then coupled with the remaining resin and tikitericin (**12**) was assumed to be formed. The peptide was then cleaved from the remaining resin. The cleavage protocol is as follows. The cleavage cocktail chosen was TFA/phenol/water/ triisopropylsilane (88/5/5/2). This was selected from the Applied Biosystems technical bulletin<sup>124</sup> as it the suggested cleavage cocktail for peptides containing cysteine residues and this was assumed to be compatible with a peptide containing (Me)Lan. The resin was washed five times with DCM and dried via air flow to ensure that there was no DMF left present on the resin. This has to be done as residual DMF can have an inhibitory effect on TFA-acidolysis. The phenol, water and triisopropylsilane were added to a graduated cylinder via micropipette followed by the TFA to ensure proper mixing. This was then added to the reaction vessel containing the peptide resin and agitated for two hours which was

suggested by the technical bulletin. All cleavages of rings and tikitericin B-D were done in sample vials using a magnetic stir bar to agitate. Once the cleavage had finished, the cleavage mixture was filtered into a sample vial and concentrated under reduced vacuum. Ether was then added to the remaining mixture and tikitericin was crash precipitated out by reducing the temperature of the solution by using a liquid nitrogen bath. This was then filtered and transferred to another vial. What was assumed to be tikitericin was revealed to be a white, crystalline solid and 8 mg of sample was isolated.



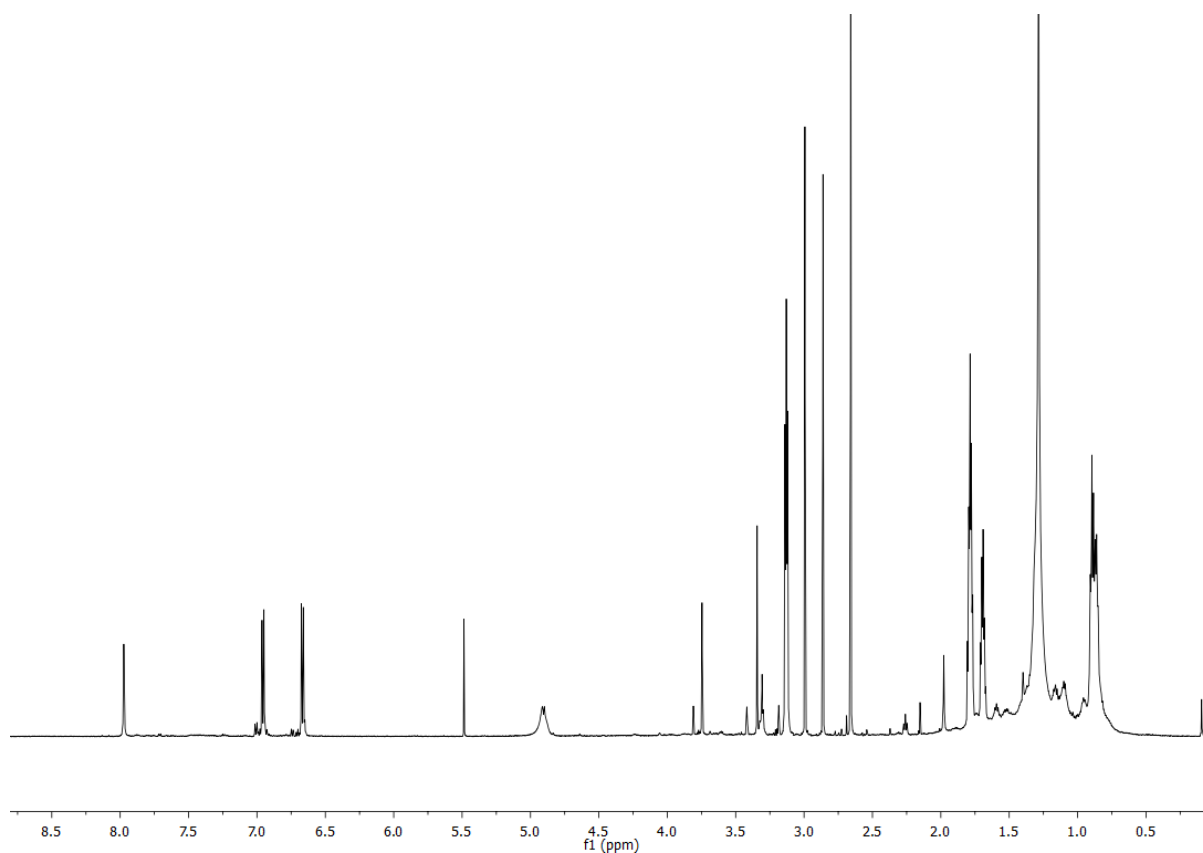
### 3.4 Analysis of the product

Once the cleavage protocol was carried out on the samples, they were analysed with HRMS. The results were not good, as the ions for any of the assumed products could not be detected. This showed that the product had not been synthesised and the endeavour had failed. This was a rather unfortunate result, as a lot of time and resources had been spent towards this.

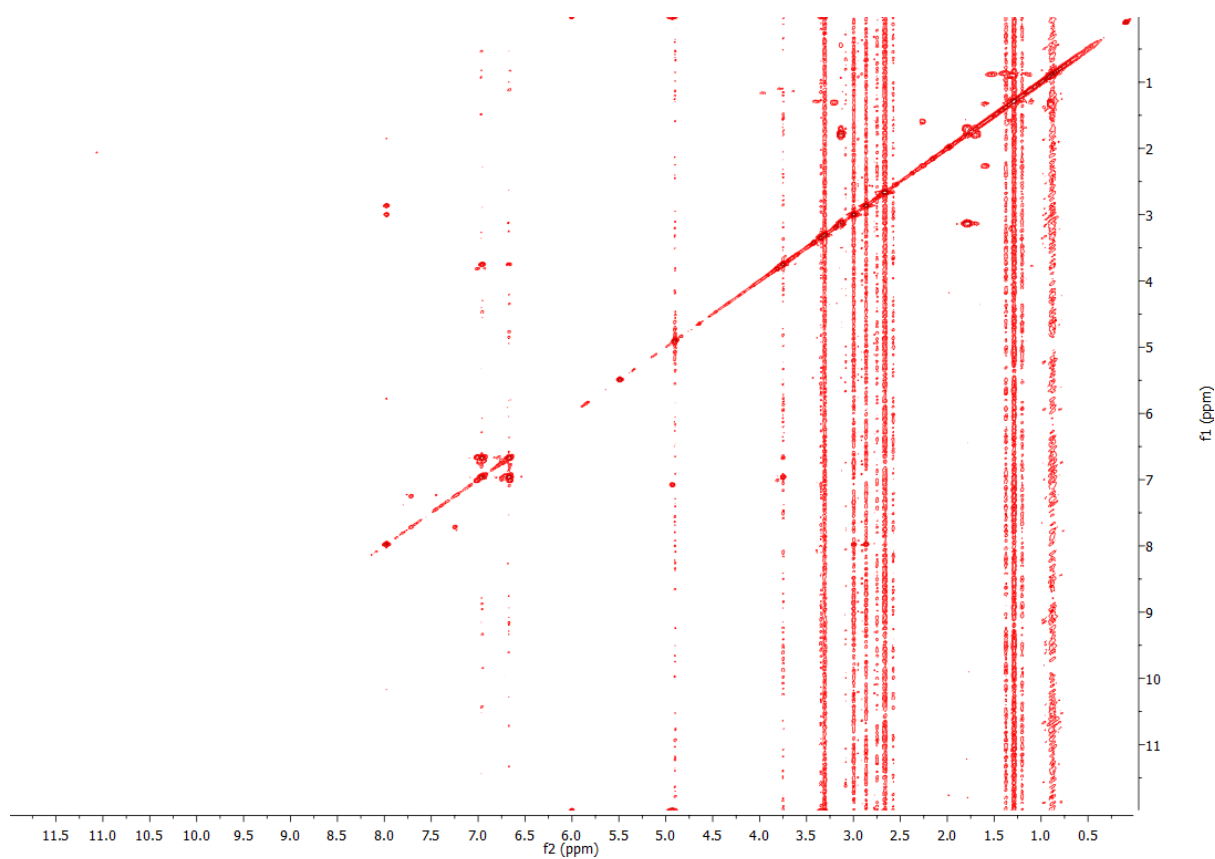
There are some possibilities why this synthesis failed. One potential reason was that the synthesis was not occurring at all to begin with. This could be either due to degradation of the Wang resin that was used or the coupling reagents could have also been degraded. The fact that no positive results during the HRMS screening of the ring closure reactions were obtained supports this hypothesis.

Another possibility is that the coupling reactions were occurring, however they were not being completed. During the whole synthesis no failed Kaiser tests were observed. This could mean that the Kaiser test reagents, which were all premade a significant time before the attempted synthesis, could have potentially degraded and not give the desired result if the coupling reaction had not completed. If the coupling reactions were not complete, then there is a low probability that SPPS would yield a full tikitericin molecule.

Another possibility is that tikitericin was synthesised, but it was destroyed in the process of the cleavage.  $^1\text{H}$  (Figure 3.5) and COSY (Figure 3.6) NMR data were taken of the precipitate revealing a simple structure. Analysis by  $^1\text{H}$  and COSY indicated a full peptide had not been made.



**Figure 3.5  $^1\text{H}$  NMR of precipitate (600 MHz,  $\text{CD}_3\text{OD}$ )**



**Figure 3.6 COSY NMR of precipitate (600 MHz, CD<sub>3</sub>OD)**

The final possibility of what went wrong is that during the cleavage of the allyl group the exposed amine went through an allylation and hence was capped. This would mean that the synthesis would be halted and would not be able to continue.

### 3.5 Concluding remarks

This was the first attempt to synthesise a lanthipeptide at Victoria University of Wellington. It was unsuccessful but these results should prove useful for future attempts to synthesise tikitericin. One thing that would be interesting is for the structure of the compound that precipitated out to be solved, as this has the potential to identify what potentially went wrong in this process. The crystals which have formed could also be observed using X-ray crystallography to confirm the structure. If it is a peptide which could be obtained from part of tikitericin then there is potential that something went wrong in the cleavage of the peptide leading to side reactions.

It is suggested that MALDI be used for the analysis of the rings formed, rather than ESI which was used for the duration of this project. As it was difficult to find a solvent which was able to solvate the rings for analysis, this could be partially a reason why no ions were detected by the HRMS. MALDI has the advantage that solid material can be directly applied to the target plate for detection. Also, as TFA is effective at suppressing ESI signals through its ability to form gas-phase ion pairs with positively-charged analyte ions, ESI should not have been the method of choice for analysing the cleaved off rings.<sup>125</sup>

Unfortunately due to time constraints and low amounts of material this synthesis was only attempted once. If there was more time, there is a likelihood tikitericin could be made. However, due to its large size and complexity, it would be suggested access to an automated SPP synthesizer would be much more suitable to for further attempts as the machines have much greater control and can interpret whether the synthesis is actually going forward; these were luxuries which were not available.

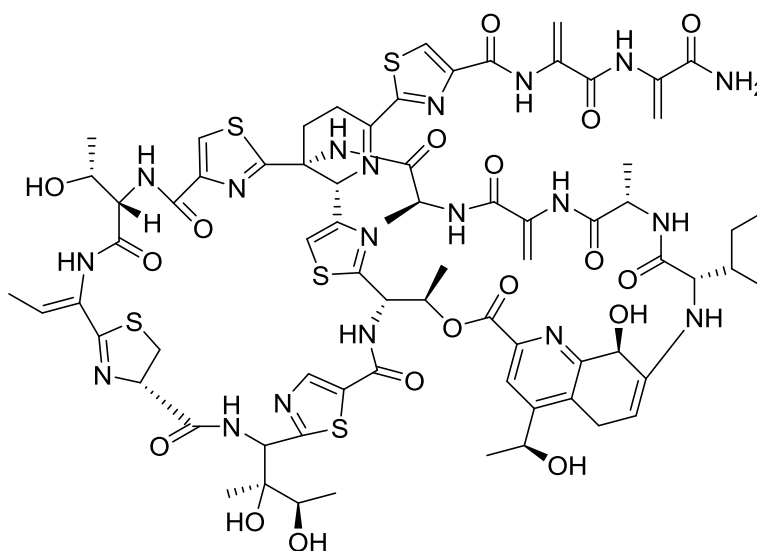
The synthesis of tikitericin is possible and hopefully in the future it will be carried out fully.

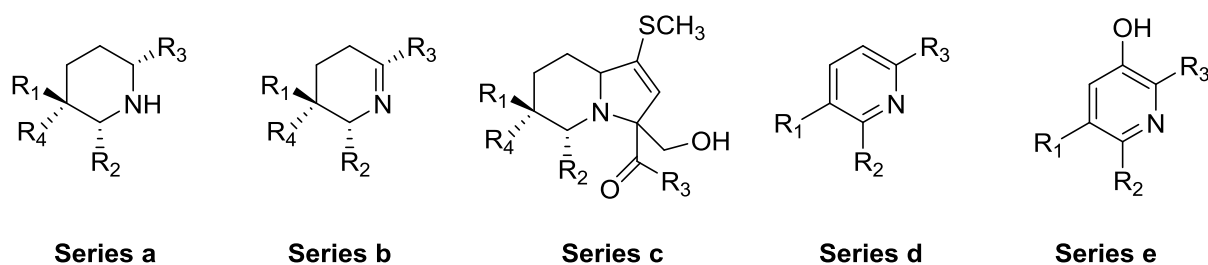
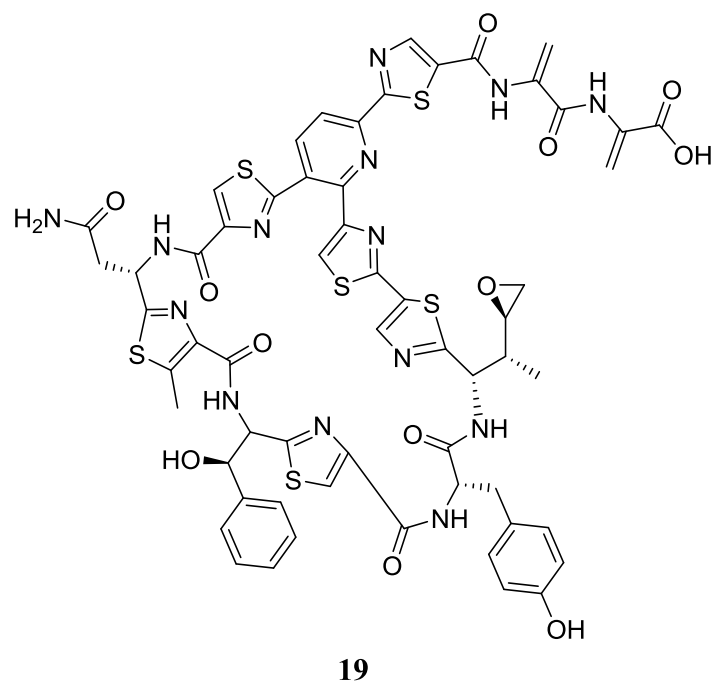
## Chapter 4

### Thiopeptides

#### 4.1 Introduction

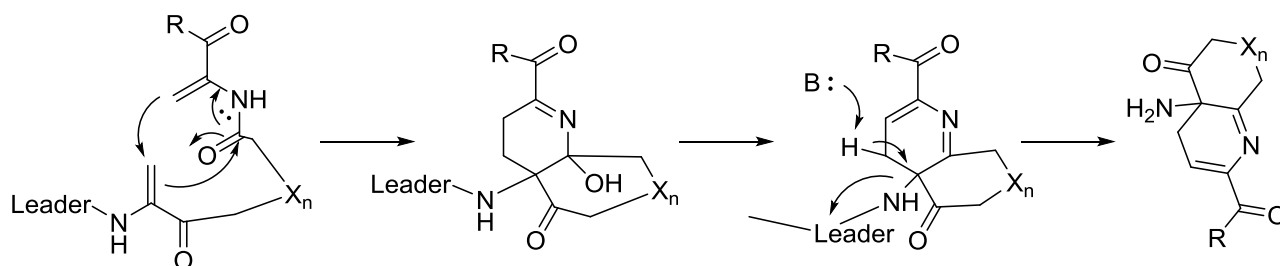
Thiopeptides were established as a class of natural products after the discovery of micrococcin in 1948.<sup>126</sup> Their biosynthetic origin was under debate, until gene sequences were attributed to their production in 2009.<sup>127</sup> Thus, this led to thiopeptides being recognised as RiPPs. Characteristic of thiopeptides is their macrocyclic core, which contains a 6-membered, nitrogenous heterocycle. Thiopeptides are further divided into five sub-families (series a-e), which are defined by the oxidation and substitution pattern of their core heterocycle. The heterocycle can either be a piperidine, dehydropiperidine or pyridine (Figure 4.1). Furthermore, series a – c incorporate into their architecture a second macrocycle, which is formed from a quinaldic acid or an indolic acid derived from tryptophan, an example of this being thiostrepton A (**18**).<sup>128</sup> The most common of the thiopeptides are those belonging to series d, an example of this being thiomuricin A(**19**).<sup>129</sup>





**Figure 4.1 The structural series of thiopeptides**

Thiopeptides are the most heavily post-translationally modified RiPPs known. They are rich in thiazole, oxazoles, Dha and Dhb. The enzymes responsible for the post-translation modifications are clustered together with the precursor peptide (TpdA). There are six proteins responsible for modifications done on the precursor peptide, TpdB-G. TpdB and C share weak similarities with Lan I dehydratases (*see* Chapter 2) and are responsible for the dehydration of Ser and Thr side chains to introduce Dha and Dhb residues. TpdE and G are responsible for the cyclisation of Cys and Ser residues to form thiozoles and oxyzoles. TpdF does not share sequence homology with any proteins of known function and hence, its biochemical role awaits elucidation. TpdD performs the cyclisation of the linear peptide to form the nitrogenous heterocycle. This is formed through a [4+2] cycloaddition between two Dha residues (Scheme 4.1).<sup>51</sup>



**Scheme 4.1 Proposed biosynthesis of pyridine ring in thiopeptides from two Dha residues<sup>51</sup>**

The thiopeptides are of clinical interest as they show potent activity against various drug-resistant pathogens, including vancomycin-resistant *Enterococci*, methicillin-resistant *Staphylococcus aureus* and penicillin-resistant *Streptococcus pneumoniae*. Depending on the thiopeptide, there are two different classes of bacterial targets. Both modes of action are mediated through the inhibition of protein synthesis. The first mode of action is through binding to the 50S ribosomal subunit. This is the area of the ribosome responsible for cycling through new tRNAs, so binding to this area prevents translational elongation. The other mode of action is through binding to elongation factor EF-Tu, which is a protein responsible for ferrying the aminoacylated tRNAs to the ribosome. Binding to EF-Tu by a thiopeptide prevents the binding of tRNAs.<sup>130</sup> The biological activity that thiopeptides exhibit means that they are valuable targets for drug discovery.

## 4.2 Thiopeptides produced by T81

As previously mentioned, after sequencing the genome of T81 Hauch and Stott identified potential gene cluster coding for three precursor peptides and an array of modification enzymes which would lead to the production of up to three thiopeptides. Former VUW MSc student Emma Aitken carried out extensive studies on T81 to identify any thiopeptides produced by T81. Emma Aitken predicted 12 structure from each of the 3 thiopeptides, 36 structures in total, in relation to the post-translational modifications (PTMs) encoded by each gene cluster. However many more structures are possible. A recent review by Mohimani and Pevzner outlines the issue of predicting RiPP structures directly from the gene sequence and PTMs encoded by the gene sequence.<sup>131</sup> Even when the core structure of the RiPP is known and the



types of PTMs can be predicted, multiple PTM sites available on the structure results in thousands of structures being predicted.

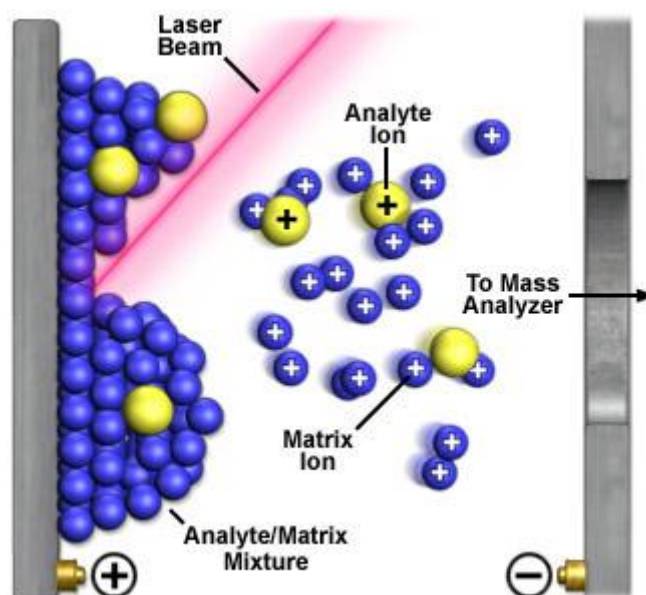
MALDI mass spectrometry was utilised in attempt to locate the thiopeptides. Unlike lanthipeptides, predicting the mass of a thiopeptide is difficult. Although the precursor peptide is known, what remains unknown is which post-translational modifications are being made. Because of this, a number of structures have to be predicted from the modification genes outlined in the gene sequence. Taking into consideration all the possible structures, the thiopeptides were screened for. T81 was co-cultured with other extremophilic bacteria on solid growth media. The thiopeptides were screened for, however, they were unable to be identified.<sup>55</sup>

### **4.3 MALDI mass spectrometry**

Mass spectrometry has been described as the smallest weighing scale in the world.<sup>132</sup> Mass spectrometry is able to identify the mass of large biomolecular complexes, individual biomolecules and small molecules. That, coupled with its high sensitivity, makes it an incredibly useful technique for analysis of molecules.<sup>133</sup>

MALDI is an ionisation technique used in mass spectrometry. It is capable of producing intact molecules over a range of molecular weights through the use of a pulsed laser combined with an energy absorbing (aromatic) matrix (Figure 4.2). The sample is generally dissolved in an organic acid and combined with the matrix compounds and applied to a metal plate. The sample is then introduced to a vacuum to evaporate any solvent, this encourages a crystal lattice to form and any compounds in the sample co-crystallise with the lattice. Next a laser is trained on to the matrix. This is similar to a meteor hitting a planet as in it causes compounds to fly off the surface. The pulsed laser is utilised to irradiate the sample, leading to ablation of the sample and matrix material. The matrix material essentially absorbs the energy from the photons, causing it to become energised. Collisions between the matrix compounds and the analyte causes the analyte to become ionised by protonation or deprotonation which is then accelerated into the mass spectrometer for analysis.<sup>134</sup> The spectrometer measures the mass-to-charge ratio of the analytes through a method called time-of-flight (TOF). The ionised analytes are accelerated by an electric field of a known strength. This gives an equal balance as all ions have the same kinetic energy, but as they have different masses they will all reach different velocities. Therefore, as each ion of a particular mass will reach its own individual velocity, they will reach the detector of a known measured distance at different times and from these experimental parameters the mass-to-charge ratio can be calculated.<sup>135</sup>

The matrix is essential for this process as it absorbs the laser energy. If the pulse laser was directly used on the sample then it would simply destroy it rather than causing ablation and ionisation.<sup>135</sup>



**Figure 4.2 Schematic representation of the MALDI desorption and ionisation process, adapted from Magnet Lab<sup>136</sup>**

#### **4.3.1 MALDI-based imaging mass spectrometry (IMS)**

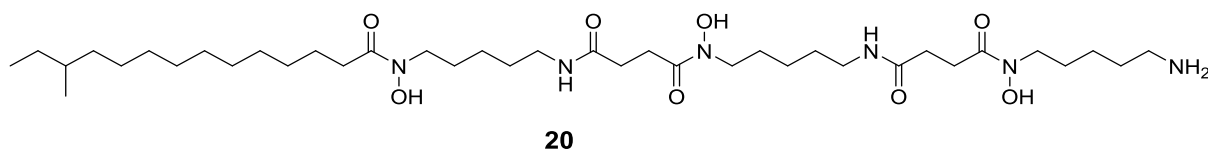
MALDI-based imaging mass spectrometry (IMS) is a technique used to visualise the spatial distribution of proteins, peptides and secondary metabolites produced by an organism. The sample, generally a piece of tissue or bacterial culture, is prepared in the same way as you would prepare a normal MALDI sample for analysis. The sample is then taken and scanned with the MALDI and each point the laser hits is treated as a pixel in an image with an associated mass spectrum.<sup>137</sup>

The uses of this technique are abundant. IMS can be utilised for imaging tissue. An example of this is from Shimma et al.<sup>138</sup> They used IMS to image colon cancer tissue sample. From this they found that an abnormal distribution of phospholipids was associated with cancerous regions of the colon, highlighting such molecules and potential biomarkers for this disease state.

Another example of the imaging of tissue came from Zaima et al.<sup>139</sup> Medaka (Japanese rice fish) gills, liver, brain and eye were taken and subjected to IMS. Analysis of the mass spectra

revealed that each organ had their own unique set of metabolites. This showed that MALDI-based IMS can be used to differentiate between organs.

Dorrestein and co-workers recently demonstrated the use MALDI-based IMS on solid growth media.<sup>137</sup> They co-cultured different species of marine microbes together to image the interactions between them. Essentially, they could visualise the metabolites being produced by bacteria on an agar plate. In this example they were particularly interested in an area of interaction between *Promicromonosporaceae* strain SIO-11 and *Bacillus subtilis* SIO-1. They could see that SIO-11 was clearly inhibiting the motility of the organism SIO-1. When growing these two together in isolation, they saw a metabolite with a mass to charge ratio ( $m/z$ ) of 2869 being secreted at the junction of interaction of these two organisms. Culturing SIO-11 alone showed that this metabolite was produced by SIO-11 and therefore was responsible for the inhibition of motility of SIO-1. They did not have the genome available for SIO-11, but further high resolution mass spectrometry along with tandem MS/MS was able to determine the structure of the protein. This examples shows how powerful this method is, as it can used to determine what organism is producing what metabolite *in vitro*. It also showed the importance of co-culturing, as they found that the promicroferrioxamine (**20**), a hydroxamate siderophore was only produced by SIO-11 in the presence of SIO-1.



As this method clearly shows the potential to identify compounds and give insight into their biological role, Dorrestein and co-workers have published the method of IMS so that it can be more widely used.<sup>140</sup> This process can be seen in Figure 4.3. The first step required for IMS is the culturing of the bacteria. The bacteria of interest is cultured in a petri dish. At this point the metal MALDI target plate can be put in the petri dish and the bacteria cultured on top of it, or the area of interest in the bacterial culture can be cut out and applied to the MALDI target plate. It is then important to take a picture of this, so then it can be overlaid later with the image produced by the MALDI.

The next step is the application of the matrix directly on to the culture media. As previously discussed, a matrix is required for the application of MALDI. The matrix of choice is a 1:1 mixture of 2,5-dihydroxybenzoic acid (DHB) (**21**) and  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) (**22**). This is essentially a broad spectrum matrix as DHB is used for peptides, nucleotides, oligonucleotides and oligosaccharides while CHCA is used for peptides, lipids and

nucleotides and both of these can be used in either positive or negative ion mode.<sup>141, 142</sup> Thus, molecules over a broad mass range can co-crystallise out with the matrix for analysis. The matrix is applied using a sieve so to saturate the sample allowing for sufficient adherence to the target plate and also for high quality crystals to grow for analysis.

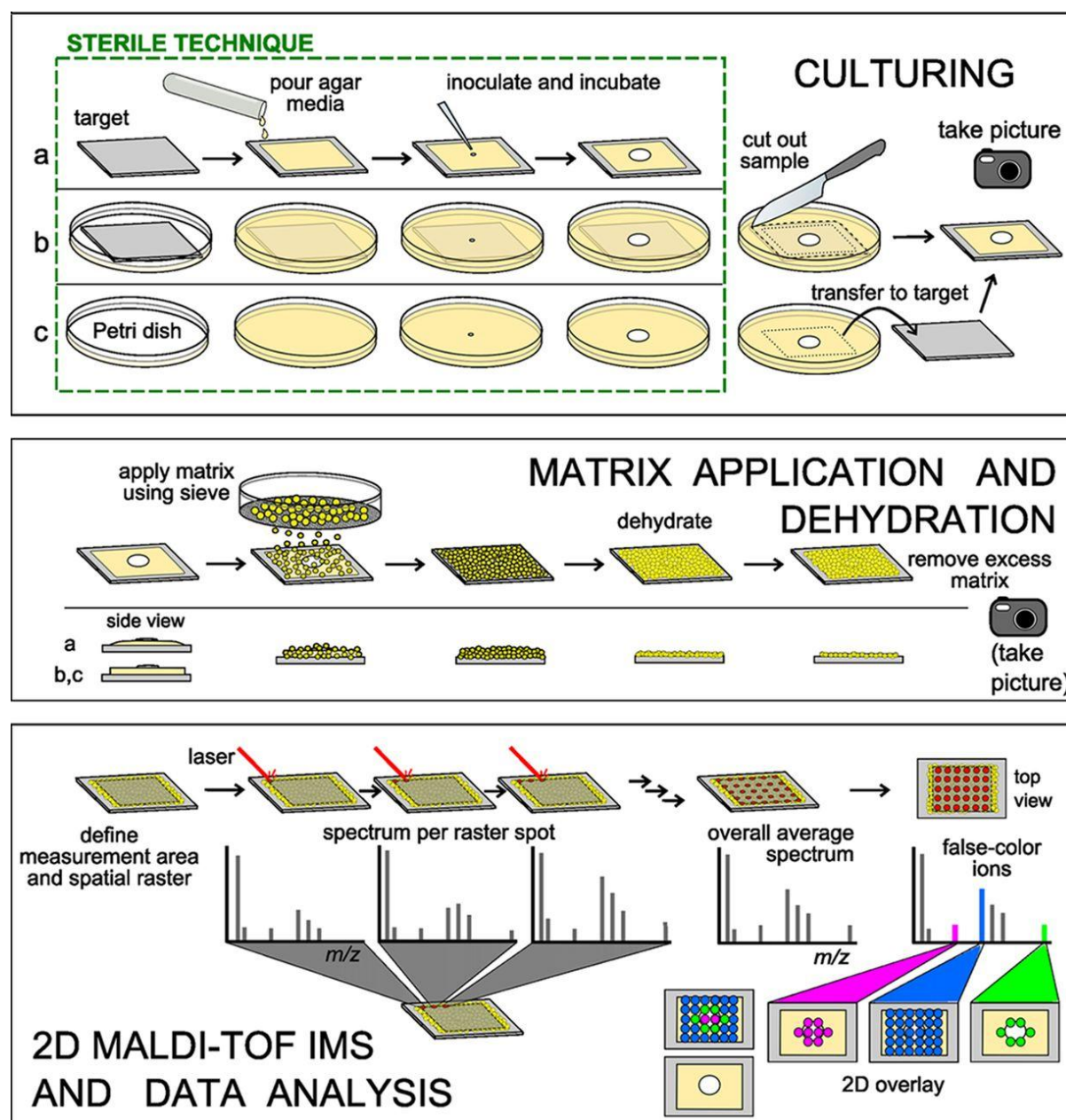
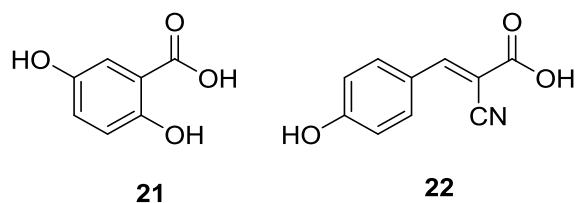


Figure 4.3 Steps for IMS adapted from Dorrestein et al.<sup>140</sup> Reproduced with permission of the American Society for Microbiology©, 2016

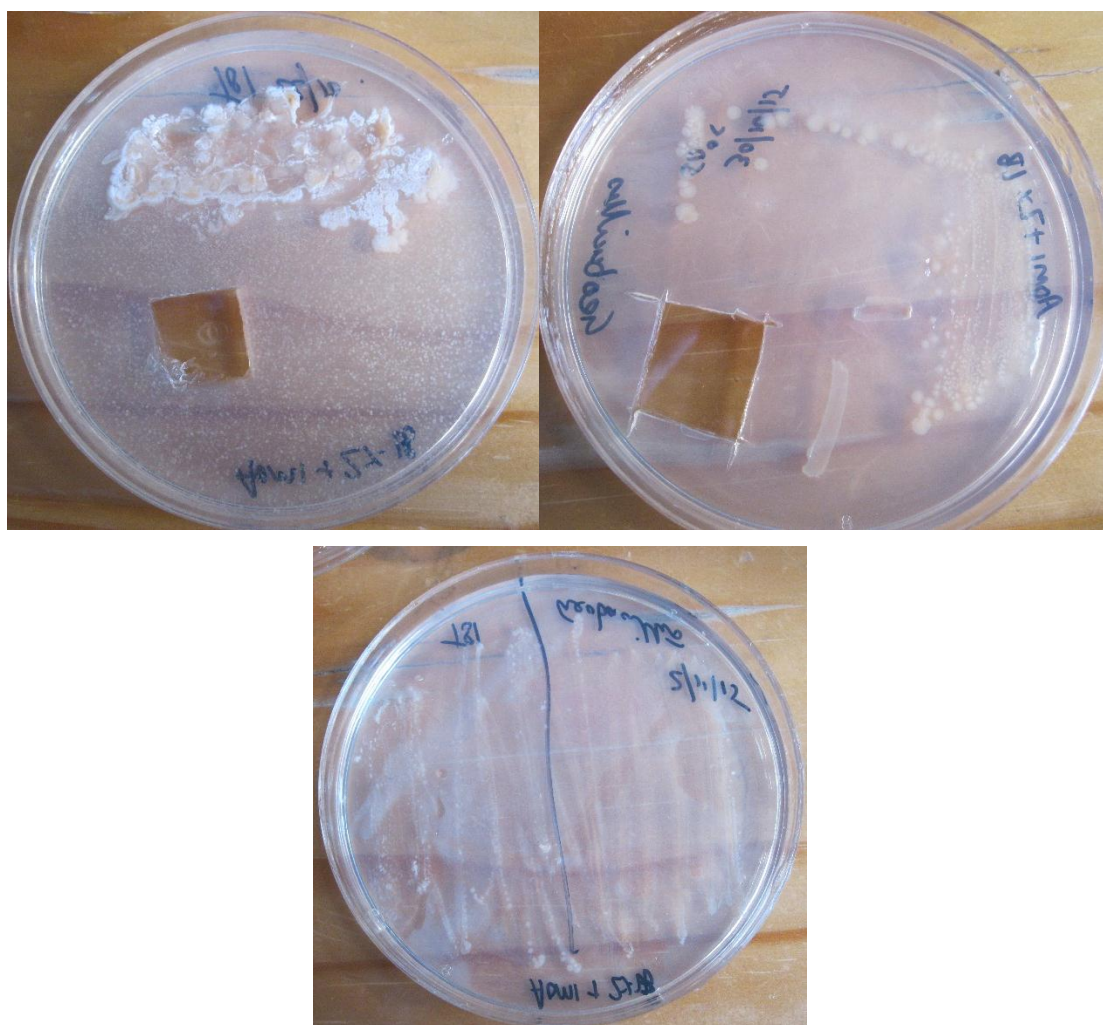


The last step is the dehydration of the sample at 37 °C followed by overnight storage in a desiccator. This step first ensures the formation of the crystal lattice, as dehydration of the sample encourages the crystal growth and hence the co-crystallisation of metabolites of interest as well. As a vacuum is used to dry the sample out, it also allows the testing of the sample in response to vacuum, as the spectrometer is under vacuum as well. If flaking of the sample occurs then it is best not to proceed with the IMS as it could come off and damage delicate instruments inside the spectrometer. It also allows for shrinking of the sample down to 0.4 – 1.0 mm which is the allowed height for the spectrometer. Once all the required steps are taken, the sample can then be subjected to MALDI IMS.

#### **4.3.2 Mass spectrometry imaging to screen for metabolites from T81**

The general method described in the previous method was followed to obtain images of *Thermogemmatispora* strain T81 and a strain of *Geobaccillus* grown both together and separately on culture media. This was done to search for tikitericin and also any thiopeptides.

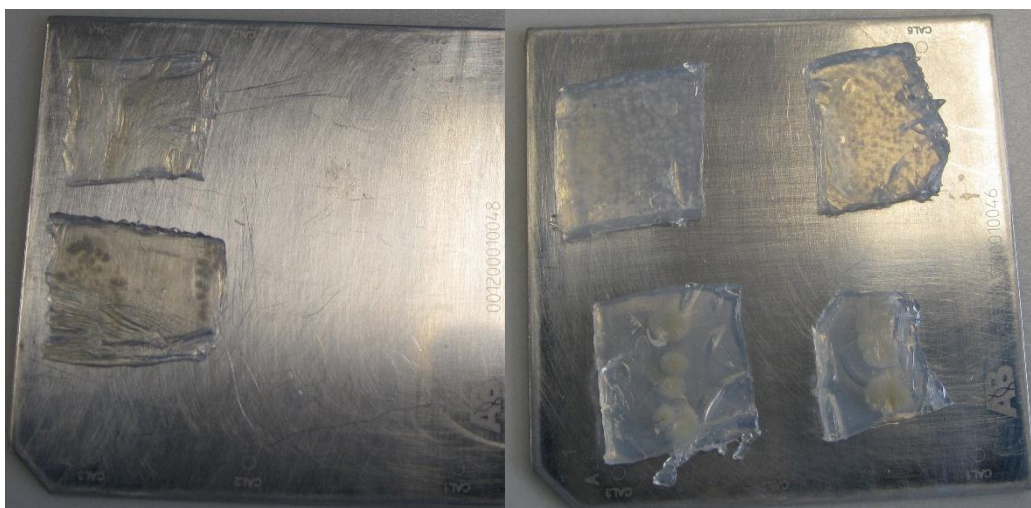
The bacteria were grown on AOM1 Minimal Media (recipe found in Appendix A) at GNS Science. This particular culture media was chosen so that T81 and *Geobaccillus* could be grown together. T81 and *Geobaccillus* were grown in the presence of each other at 60 °C in Petri dishes along with them being cultured alone. These can be seen in Figure 4.4.



**Figure 4.4 T81 (top left), *Geobaccillus* (top right) and co-culture (bottom)**

Sections of the culture media were cut out and transferred to MALDI plates (Figure 4.5). It was difficult to see any obvious zone of inhibition for the co-culture of T81 and *Geobaccillus* so two sections were cut from the middle of plate as it looked like they were growing together there. One problem that was faced was that the agar on both the monoculture plates was quite thick, and this gave us issues with adherence to the metal target plate. The metal target plate we used was unlike normal target plates. In order for the culture media to adhere to the target plate efficiently, the target plate needed to be roughened as suggested by Dorrestein.<sup>140</sup>





**Figure 4.5 MALDI Target plates Left co-culture, right T81 (top) and *Geobaccillus* (bottom)**

As is seen above, two sections were cut of each culture. This was because the MALDI MS that was used can only scan over a  $m/z$  range lower then what was needed. It was therefore decided that one section would be scanned from 0 – 1000  $m/z$  and the other would be scanned 900 – 4000  $m/z$ . The next step was to take these and apply the matrix to them (Figure 4.6). For this, crystals of CHCA and DHB were ground together and applied directly on top of the culture media through a 60  $\mu\text{m}$  nylon mesh. Care had to be taken to give even coverage across the entire section as it was difficult to control where the matrix mix was falling. This gave good coverage over each section of culture media.



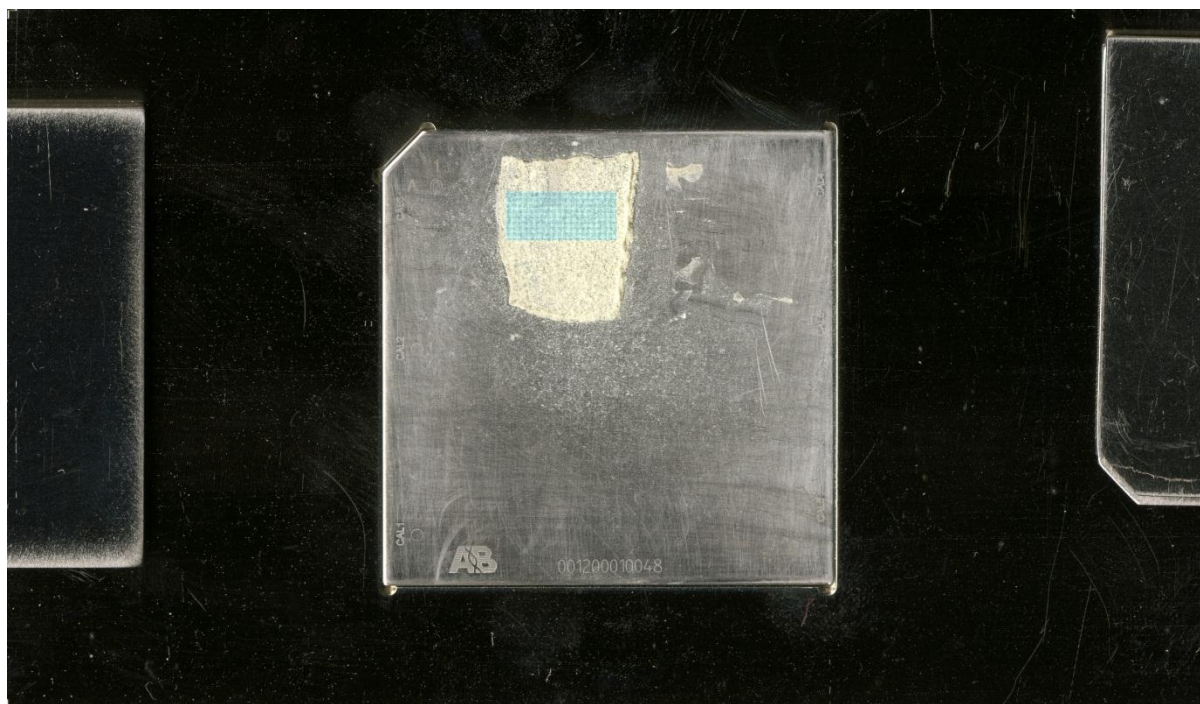
**Figure 4.6 Application of matrix to culture media sections**

These were then taken and incubated at 37 °C for 4 hours to get sufficient dehydration as some of the samples were rather thick. The plates were then transferred to a lyophiliser and desiccated

overnight, the vacuum also tested the samples to ascertain whether they had adhered to the target plate correctly. Unfortunately only one of the samples had adhered to the target plate, but fortunately it was one of the co-culture experiments, which were the primary samples of interest. The only downside to this is that the controls would not be able to be compared to the co-culture therefore direct correlation between metabolites found in the spectra and whether they are being produced by T81 or *Geobaccillus*. However, since the  $m/z$  of interest were known for T81 it can be safely assumed that if a mass match was found it is most likely a metabolite being produced by T81.

#### 4.3.3 Preliminary results of microbial IMS of T81 versus *Geobaccillus*

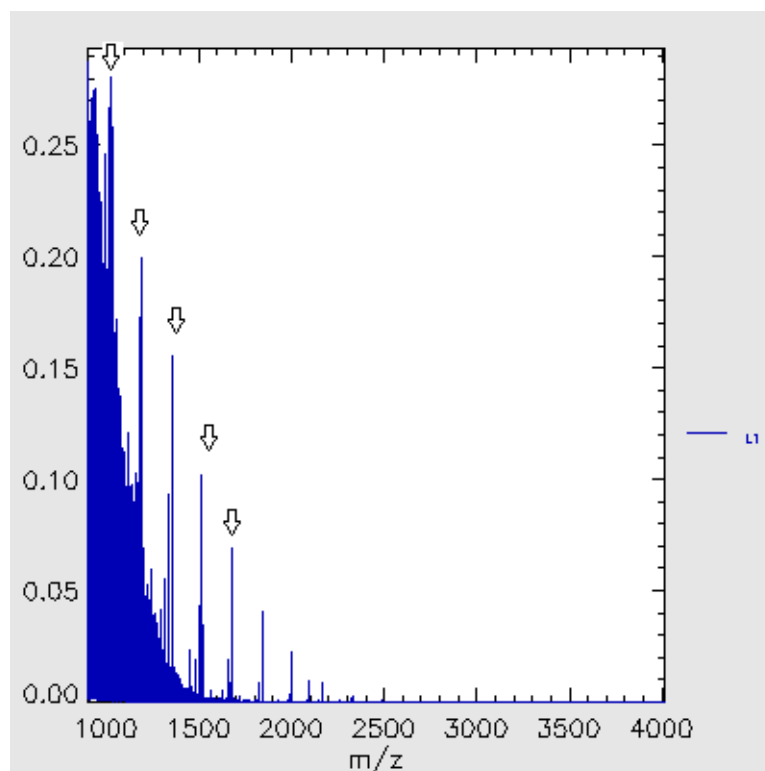
The MALDI target plate was transferred to a MALDI plate holder (Figure 4.7) and an image was taken of it. This was done so that the software could be calibrated to irradiate the region of interest. The area which was imaged scanned from 900 – 4000  $m/z$  can be seen highlighted by blue. The sample was then analysed by MALDI-TOF IMS with a resolution of 100  $\mu\text{m}$  in positive ion mode.



**Figure 4.7 Target plate positioned in plate holder**

The IMS was visualised and normalised with the software Biomap. The global scan mass spectrum of the IMS can be seen in Figure 4.8. Outlined by arrows are peaks which appear to have a spacing characteristic of a polymer.<sup>143</sup> As these peaks are of high ion count, they are most likely pertaining to the culture media.



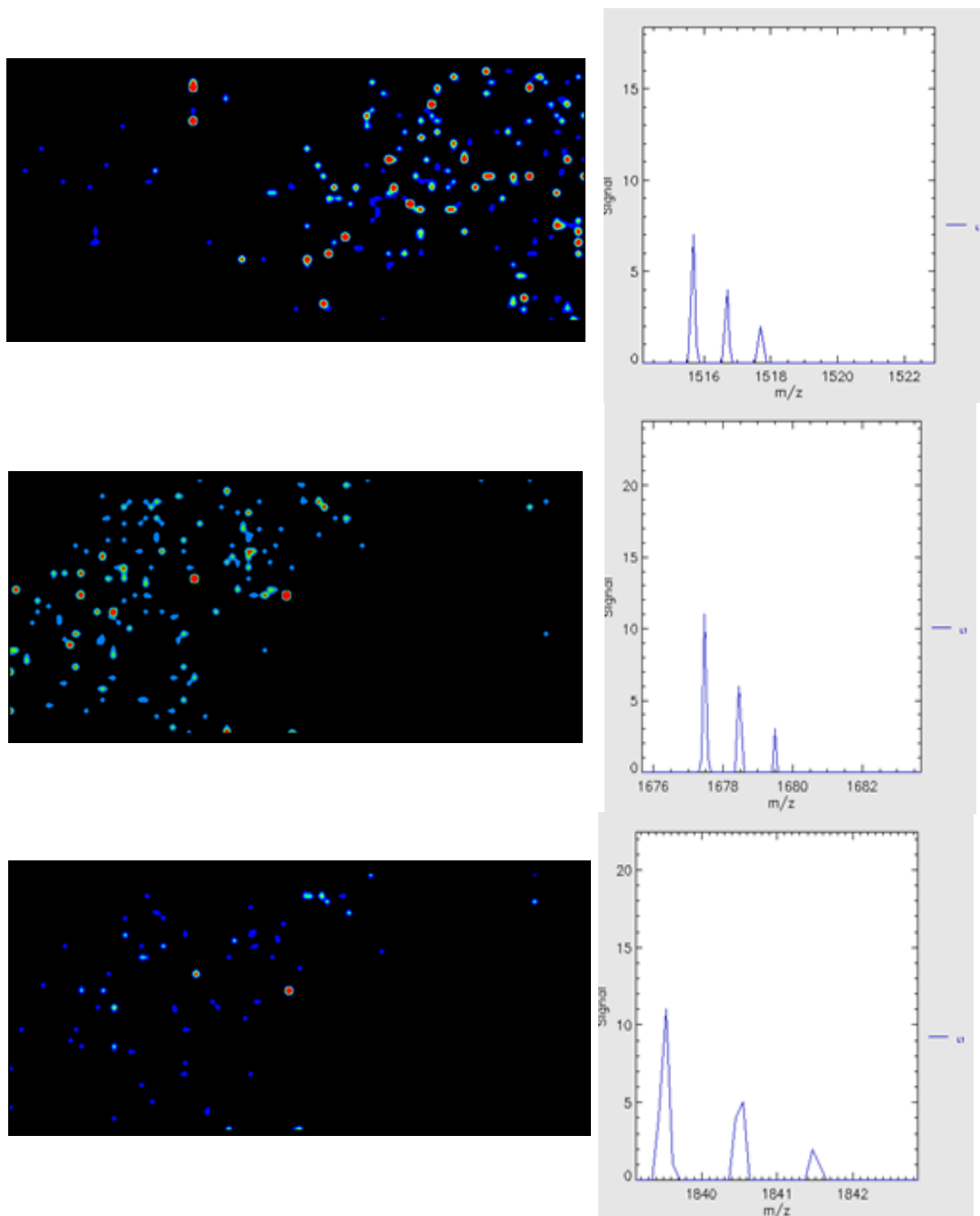


**Figure 4.8 Global scan mass spectrum of IMS (900 – 4000  $m/z$ )**

The first item of interest was whether it was possible to locate tikitericin A – D in the IMS. The ions were searched for, but they could not be located in the IMS. There are two possible reasons why they could not be located. Either T81 does not produce tikitericin in response to *Geobacillus*, or due to the spectrometers low sensitivity in the mass range from 2000 – 4000  $m/z$  and the known low concentration of tikitericin produced by T81 it was unable to be located. Further T81 cultures versus other extremophilic bacteria should be imaged to determine whether the presence of tikitericin as an antimicrobial can be seen *in situ*.

The second item of interest was whether any ion clusters resembling the thiopeptides can be seen and if the ions detected are mass matches to the structures predicted by Emma Aitken.

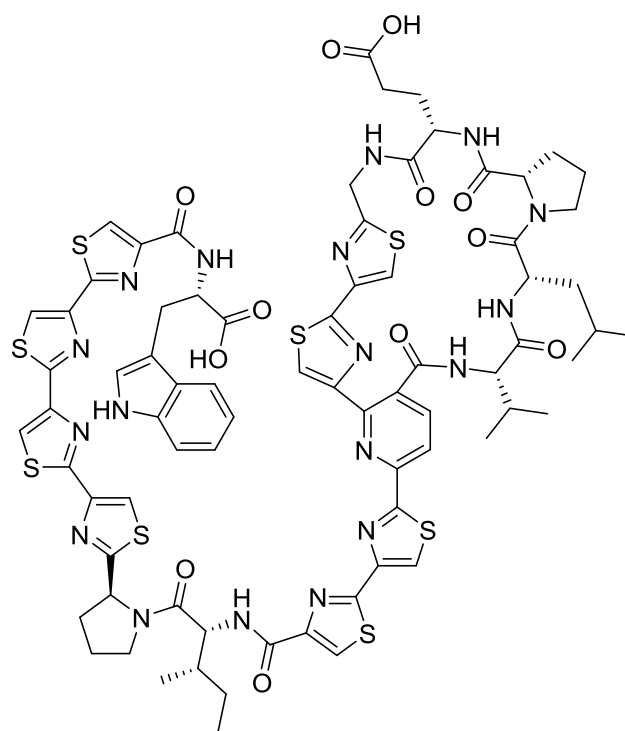
As thiopeptides are difficult to predict the structures of, any ion clusters which looked similar to what would be found were searched for in the global scan mass spectrum and viewed by Biomap. From this, three ion clusters were detected in the image which could be pertaining to thiopeptides. The density and location of these ions can be found in Figure 4.9 along with the mass spectrum associated with each ion.



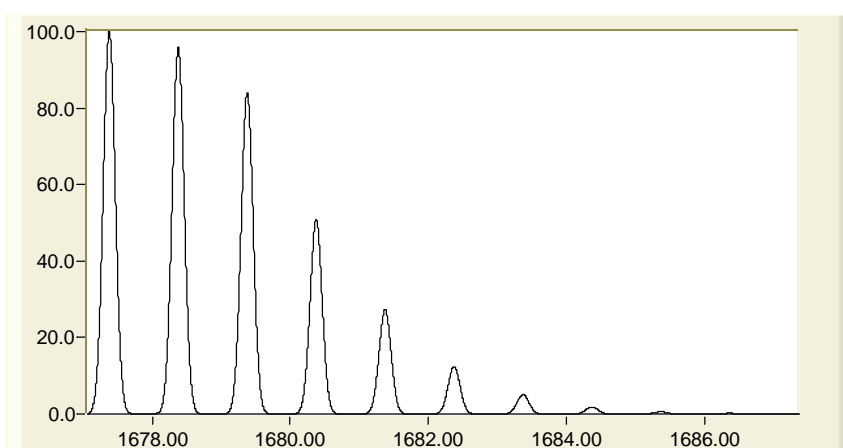
**Figure 4.9 Global distribution of selected ions**

The ion with  $m/z$  1677.47 is reasonably close to one of the structures which Emma Aitken predicted (**23**), which would have a calculated  $m/z$  1677.3689  $[M+H]^+$ , only 0.1 Da difference between predicted and measured. If the surface of the sample was uneven, this generally causes a mass shift which could result in this discrepancy seen. The calculated isotopic distribution of **23** can be seen in Figure 4.10. This is not a full match with the isotopic distribution shown in the experimental mass spectrum. Either way, more data will have to be collected to ascertain whether this is pertaining to **23** or not. The other two images shown do not have a  $m/z$  which

matches any of the predicted structures, however it is interesting to note that the image showing the distribution of the ion  $m/z$  1839.08 is distributed in the same areas as  $m/z$  1677.47. This could potentially be an intermediate product of the biosynthesis of **23** and the difference in masses seen could be due to a tyrosine.



**23**



**Figure 4.10** Calculated isotopic distribution of **23**

## 4.4 Concluding remarks

A co-culture of T81 versus *Geobacillus* was subjected to MALDI IMS to attempt to image tikitericin and their truncated forms as an antimicrobial agent *in situ*. These were undetected though. Due to time constraints and issue in the culturing process it was not possible to image other cocultures, but this target was in mind throughout the project. Because there was no obvious zone of inhibition it is no surprise that tikitericin was not detected.

There is potential that a thiopeptide has been detected, however more work will need to be done to ascertain whether it is the structure shown.

Because the monocultures were too thick, IMS was unable to be obtained for each of them. In the future it would be suggested that the culture media be placed over the MALDI target plate in the initial culturing period and then cut out when being subjected to IMS. This will ensure better adherence to the target plate.

## Chapter 5

### Concluding remarks and future directions

The research described in this thesis is the first example of an attempt at synthesising a lanthipeptide at Victoria University of Wellington. The aim throughout the project was to perform the total synthesis of tikitericin, however a large portion of time was directed towards optimising the synthetic procedures developed by Vederas. The optimisation of his method was successfully carried out, and time constraints meant that only one attempt was possible at synthesising tikitericin. This was unsuccessful.

Preliminary results for IMS were obtained and a thiopeptide was potentially detected. Tikitericin and its truncated forms acting as antimicrobials were not detected by MALDI IMS *in situ*. Future work towards this objective would be to image a range of co-cultures of varying bacteria verse T81. Co-cultures with clear zones of inhibition should be imaged and this will hopefully result in detecting tikitericin.

The most obvious future objective to be completed would be the total synthesis of tikitericin and its precursors to be successfully carried out. This should be carried out with an automated SPP synthesiser. More (Me)Lan precursors need to be made for this and once they are there should be no issues producing tikitericin with an automated SPP synthesiser as it is a much more controlled method compared to manual SPPS. Once tikitericin is synthesised and bioactivity is ascertained, it would be interesting to see the effects of changing the sulfur of the thioether with a different atom, such as carbon, and observing if there is any change in bioactivity or structural stability.

Expressing tikitericin in an exogenous host, such as *E. coli*, would also be another way to produce tikitericin. Van der Donk and co-workers have shown that this is possible with other lanthipeptides.<sup>122</sup> The gene cassette of tikitericin has been isolated and following the procedures outlined by van der Donk it could be expressed in an exogenous host.

## Chapter 6

### Experimental

#### 6.1 General experimental

$^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra of starting materials were recorded on a Varian Unity Inova (500 MHz for  $^1\text{H}$  and 125 MHz for  $^{13}\text{C}$ ) spectrometer.  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra of peptides were recorded on a Varian Direct Drive spectrometer equipped with a triple resonance HCN cryogenic probe operating at 25 K (600 MHz for  $^1\text{H}$  and 150 MHz for  $^{13}\text{C}$ ). The chemical shifts of  $^{13}\text{C}$  and  $^1\text{H}$  spectra were referenced to the residual solvent peaks ( $\text{CDCl}_3$ :  $\delta_{\text{C}}$  77.0,  $\delta_{\text{H}}$  7.26;  $\text{CD}_3\text{OD}$ :  $\delta_{\text{C}}$  49.0,  $\delta_{\text{H}}$  3.31;  $\text{DMSO}-d_6$ :  $\delta_{\text{C}}$  39.5,  $\delta_{\text{H}}$  2.50). All NMR solvents were purchased from Cambridge Isotope Laboratories with the exception of  $\text{CDCl}_3$  which was purchased from Aldrich. ESIMS results were obtained from an Agilent 6530 Q-TOF mass spectrometer with an Agilent 1260 HPLC for solvent delivery utilising a JetStream<sup>TM</sup> electrospray ionisation source in positive ion mode. Water and MeCN solvents were both made up as 5 mM solutions with ammonium formate. IR data was recorded on a Bruker ALPHA FTIR Spectrometer equipped with a Platinum ATR single reflection diamond. Optical rotations were measured at 25 °C using a Rudolph Research Analytical Autopol II polarimeter operating at the emission wavelength of a sodium lamp (589 nm), with concentrations given in 1 mg/ml.

Flash chromatography was performed using a column filled with silica gel 60 (40 – 63  $\mu\text{m}$ ). Purification of products *via* reverse phase chromatography was performed using Supelco Diaion<sup>®</sup> HP20 poly(styrene-divinylbenzene) (PSDVB) resin.

All solvents used were HPLC-grade (purchased from Fischer Scientific) with the exception of hexanes and dichloromethane, which were Optima<sup>®</sup> grade and the *N,N*-dimethylformamide which was purchased from Acros organic. *N,N*-Dimethylformamide was degassed by the standard freeze-pump-thaw technique. Water was distilled prior to use. Dichloromethane used for reactions was purified and degassed by an Innovative Technologies Pure Solve MD5. Solvent compositions used are all reported as % vol/vol. TLC analyses were performed using Machery-Nagel Polygram<sup>®</sup> Sil G/UV<sub>254</sub> plates. TLC plates were visualised under UV light ( $\lambda$  = 350 nm).

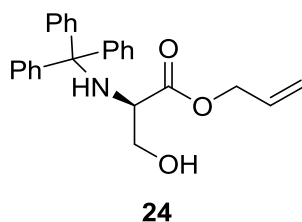
Unless otherwise stated, reactions were performed under an atmosphere of argon (instrument grade) in vacuum-dried glassware using dry solvents and standard syringe techniques. Compounds are named according to IUPAC guidelines.<sup>144</sup> NMR has been assigned in reference to the literature.

Reagents were purchased either from Fischer or Aldrich. Fmoc amino acids were purchased from Intavis, threonine/serine were purchased from Chem-Impex international and Wang resin was purchased from Aapptech.

## 6.2 Synthesis of starting materials

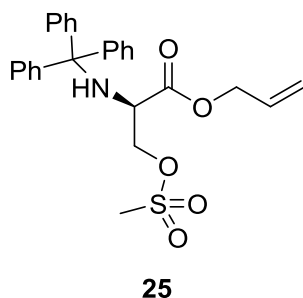
### 6.2.1 Synthesis of Lan intermediates and starting material

#### (2*R*)-*N*-Triphenylmethylserine allyl ester (**24**)



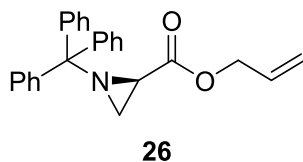
NaHCO<sub>3</sub> (563 mg, 6.7 mmol) was added to a solution of tritylthreonine.triethylamine (1.50 g, 3.35 mmol) in DMF (30 mL). The solution was cooled to 0°C. Allyl bromide (0.35 mL, 4.0 mmol) was added dropwise and the solution was allowed to warm to room temperature overnight. The solution was quenched with water (150 mL) then extracted with Et<sub>2</sub>O (4 x 50 mL). The organic layers were combined, dried over MgSO<sub>4</sub>, filtered and then concentrated *in vacuo*. The crude product used without further purification to yield **24** (1.50 g, 58 %) as a colourless oil.  $[\alpha]_D^{25} = +0.17$ ; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.49 (d, 6H, Trt), 7.26 (t, 6H, Trt), 7.20 (t, 3H, Trt), 5.70 (m, 1H, CH=CH<sub>2</sub>), 5.20 (dq, 1H, *J* = 17.2, 1.5 Hz, CH=CH<sub>2</sub>), 5.18 (dq, 1H, *J* = 10.5, 1.5 Hz, CH=CH<sub>2</sub>), 4.21 and 4.18 (2dd, 2H, *J* = 13.2, 5.9 Hz, CH<sub>2</sub>-CHCH<sub>2</sub>), 3.71 (m, 1H, HR), 3.87 and 3.68 (2dd, 2H, *J* = 10.4, 5.6 Hz, CH<sub>2</sub> α), 3.16 (s, 1H, OH); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 173.1 (CO), 145.6 (C, arom. Trt), 131.5 (CHCH<sub>2</sub>), 128.7, 127.9, 126.6 (C arom. Trt), 118.5 (CHCH<sub>2</sub>), 71.0 (C. aliph. Trt), 65.7 (CH<sub>2</sub>-CHCH<sub>2</sub>), 64.9 (Cβ), 57.8 (Cα). Data consistent with literature.<sup>65</sup>

***N*-Triphenylmethyl-(*R*)-serine(*O*-methanesulfonyl) allyl ester (**25**)**



**24** (1.30 g, 3.35 mmol) was dissolved in DCM (30 mL) and cooled to 0°C. NEt<sub>3</sub> (1.07 mL, 7.71 mmol) was added and after 5 minutes Ms-Cl (0.31 mL, 4.02 mmol) was added dropwise. The solution was allowed to warm to room temperature overnight. The solution was quenched with NH<sub>4</sub>Cl (sat.) solution (35 mL) and extracted with DCM (3 x 50 mL). The organic layers were combined and washed with brine (50 mL), dried over MgSO<sub>4</sub>, filtered and then concentrated *in vacuo* to give a clear foamy oil. The crude product was purified by flash chromatography (SiO<sub>2</sub>, 20% EtOAc in Hexanes) to yield **25** (0.98 g, 63 %) as a white crystalline solid.  $[\alpha]_D^{25} = +9.0$ ; <sup>1</sup>H NMR (500 MHz, DMSO-d<sub>6</sub>) δ 7.41 (d, 6H, Trt), 7.28 (t, 6H, Trt), 7.20 (t, 3H, Trt), 5.70 (m, 1H, CH=CH<sub>2</sub>), 5.21 (dd, 1H, *J* = 5, 1.5 Hz, CH=CH<sub>2</sub>), 5.14 (dd, 1H, *J* = 10.5, 1.5 Hz, CH=CH<sub>2</sub>), 4.19 and 4.15 (2dd, 2H, *J* = 13.2, 5.9 Hz, CH<sub>2</sub>-CHCH<sub>2</sub>), 4.02 (m, 1H, HR), 3.87 and 3.68 (2dd, 2H, *J* = 10.4, 5.6 Hz, CH<sub>2</sub> α), 3.35 (s, 3H, SCH<sub>3</sub>) 3.14 (s, 1H, OH); <sup>13</sup>C NMR (125 MHz, DMSO-d<sub>6</sub>) δ 171.6 (CO), 145.9 (C, arom. Trt), 132.3 (CHCH<sub>2</sub>), 128.8, 128.41, 127.0 (C arom. Trt), 118.5 (CHCH<sub>2</sub>), 70.9 (C. aliph. Trt), 70.6(CH<sub>2</sub>-CHCH<sub>2</sub>), 65.5 (Cβ), 55.9 (Cα), 37.19 (SCH<sub>3</sub>). ESI-MS: C<sub>26</sub>H<sub>27</sub>NO<sub>5</sub> [M+Na]<sup>+</sup> *m/z* found 488.1, calcd 448.1. Data consistent with literature.<sup>65</sup>

**(*R*)-1-Triphenylmethy laziridine-2-carboxylic acid allyl ester (**26**)**

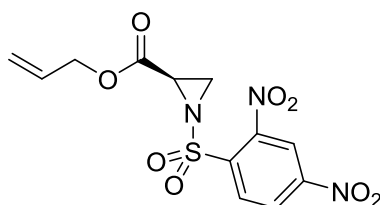


A sample of **25** (0.96 g, 2.06 mmol) was dissolved in DMF (60 mL). NEt<sub>3</sub> (0.34 mL, 2.47 mmol) and DMAP (25 mg, 0.21 mmol) was added and the reaction was heated overnight at 70°C. The reaction was cooled to room temperature, diluted with water (100 mL) and extracted with Et<sub>2</sub>O (4 x 30 mL). The organic layers were combined and washed with brine (30 mL), dried over MgSO<sub>4</sub>, filtered and then concentrated *in vacuo* to give colourless foamy oil. The



crude product was purified by flash chromatography (SiO<sub>2</sub>, 10% EtOAc in Hexanes) to yield **26** (0.550 g, 72%) as a colourless oil.  $[\alpha]_D^{25} = -0.12$ ; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  7.51 (d, 6H,  $J = 5$  Hz, Trt), 7.28 (t, 6H,  $J = 6$  Hz, Trt), 7.22 (t, 3H,  $J = 7$  Hz, Trt), 5.95 (m, 1H, CHCH<sub>2</sub>), 5.36 (dq, 1H,  $J = 17.2, 1.5$  Hz, CHCH<sub>2</sub>), 5.27 (dq, 1H,  $J = 10.5, 1.2$  Hz, CHCH<sub>2</sub>), 4.68 (m, 2H, CH<sub>2</sub>-CHCH<sub>2</sub>), 2.27 (s, 1H H $\alpha$ ), 1.94 (dd, 1H,  $J = 2.7, 6.2$  Hz, HR), 1.42 (dd, 1H,  $J = 6.2$  Hz, 1.6 Hz, H $\alpha$ ); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  171.3 (CO), 143.5 (C<sub>q</sub> arom.), 132.0 (CHCH<sub>2</sub>), 129.3, 127.6, 126.9 (C arom.), 118.5 (CHCH<sub>2</sub>), 74.4 (C<sub>q</sub> aliph.), 65.7 (CH<sub>2</sub>-CHCH<sub>2</sub>), 31.7 (CR), 28.7 (C $\alpha$ ). ESI-MS: C<sub>25</sub>H<sub>23</sub>NO<sub>2</sub> [M+Na]<sup>+</sup>  $m/z$  found 393.1, calcd 393.1. Data consistent with literature.<sup>65</sup>

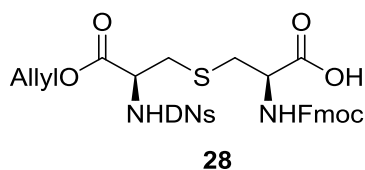
#### Allyl-(*R*)-1-(2,4-dinitrobenzenesulfonyl)aziridine-2-carboxylate (**27**)



**27**

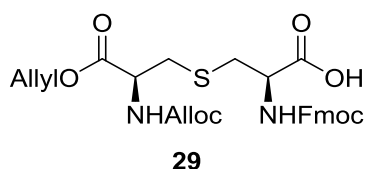
TFA (1.25 mL, 8.7 mmol) was added dropwise at 0 °C to a solution of **26** (0.54 g, 1.45 mmol) in DCM (6 mL) and MeOH (2 mL). The solution was stirred for 2 hours at 0 °C and concentrated *in vacuo*. The resulting residue was partitioned between Et<sub>2</sub>O (40 mL) and H<sub>2</sub>O (20 mL) and the ether layer was extracted with water (3 x 10 mL). The combined aqueous layers were basified to pH 9 with Na<sub>2</sub>CO<sub>3</sub> at 0 °C. EtOAc (40 mL) was added to the aqueous solution followed by 2,4-dinitrobenzenesulfonyl chloride (0.97 g, 3.63 mmol) at 0 °C. The biphasic system was warmed to room temperature and stirred vigorously for 24 h. The two layers were then separated and the aqueous layer was extracted with EtOAc (3 x 10 mL). The organic layers were combined, dried over MgSO<sub>4</sub>, filtered and then concentrated *in vacuo*. The crude product was purified by flash chromatography (SiO<sub>2</sub>, 20% EtOAc in Hexanes) to yield **27** (0.31 g, 58 %) as a light yellow oil.  $[\alpha]_D^{25} = +0.06$ ; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  8.64 (d, 1H,  $J = 2.3$  Hz, DN<sub>s</sub>-H), 8.59 (dd, 1H,  $J = 2.1, 8.6$  Hz, DN<sub>s</sub>-H), 8.51 (d, 1H,  $J = 8.6$  Hz, DN<sub>s</sub>-H), 5.92 (ddt, 1H,  $J = 6.0, 17.4, 22.5$  Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 5.38 (app dq, 1H,  $J = 1.4, 17.3$  Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 5.33 (app dq,  $J = 1.1, 10.4$  Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 4.70 (m, 2H, -CH<sub>2</sub>CH=CH<sub>2</sub>), 3.72 (dd, 1H,  $J = 4.5, 7.0$  Hz, H $\alpha$ ), 3.15 (d, 1H,  $J = 7.1$  Hz, H $\beta$ ), 2.87 (d, 1H,  $J = 4.5$  Hz, H $\beta$ ); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz):  $\delta$  165.8, 150.5, 148.7, 137.0, 133.2, 130.9, 126.9, 120.2, 119.6, 66.8, 38.3, 34.7 ESI-MS: C<sub>12</sub>H<sub>11</sub>N<sub>3</sub>O<sub>8</sub>S [M+H]<sup>+</sup>  $m/z$  found 358.0, calcd 358.0. Data consistent with literature.<sup>65</sup>

**(R)-2-(((fluoren-9-yl)methoxy)carbonylamido)-3-((S)-3-(allyloxy)-2-(2,4-dinitrobenzenesulfonyl)-3-oxopropylthio)propanoic acid (28)**



To a solution of **27** (600 mg, 1.68 mmol) in DCM (10 mL) was added Fmoc-Cys-OH (2.30 g, 6.71 mmol) followed by  $\text{BF}_3 \cdot \text{OEt}_2$  (1.65 mL, 13.40 mmol) at 0 °C. The resulting mixture was warmed to room temperature and stirred for 48h. Removal of solvent *in vacuo* afforded a crude yellow oil, which was purified by flash chromatography ( $\text{SiO}_2$ , 3:7:0.2/EtOAc:Hexanes:HOAc) to yield **28** (0.43 g, 43 %) as a light yellow oil.  $[\alpha]_D^{25} = +0.30^\circ$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  8.70 (d, 1H,  $J = 2.3$  Hz, DNs-H), 8.28 (dd, 1H,  $J = 2.3, 8.8$  Hz, DNs-H), 8.29 (d, 1H,  $J = 8.7$  Hz, DNs-H), 7.77 (d, 2H,  $J = 7.6$  Hz, Fmoc-ArH), 7.77 (d, 2H,  $J = 7.4$  Hz, Fmoc-ArH), 7.61 (br m, 1H, NH), 7.41 (d, 1H,  $J = 6.2$  Hz, NH), 7.32 (t, 2H,  $J = 7.4$  Hz, Fmoc-ArH), 7.30 (t, 2H,  $J = 7.4$  Hz, Fmoc-ArH), 5.70 (ddd, 1H,  $J = 5.0, 11.0, 22.8$  Hz,  $-\text{CH}_2\text{CH}=\text{CH}_2$ ), 5.20 (dd, 1H,  $J = 1.5, 17.3$  Hz,  $-\text{CH}_2\text{CH}=\text{CH}_2$ ), 5.08 (dd, 1H,  $J = 1.2, 10.5$  Hz,  $-\text{CH}_2\text{CH}=\text{CH}_2$ ), 4.41 (d, 2H,  $J = 5.4$  Hz,  $-\text{CH}_2\text{CH}=\text{CH}_2$ ), 4.23 (m, 4H, Fmoc- $\text{CH}_2$ , Fmoc-CH and H $\alpha$ ), 4.02 (m, 1H, H $\alpha$ ), 2.97 (dd, 1H,  $J = 5.7, 13.6$  Hz, H $\beta$ ), 2.91 (dd, 1H,  $J = 4.2, 13.6$  Hz, H $\beta$ ), 2.82 (dd, 1H,  $J = 8.3, 14.0$  Hz, H $\beta$ ), 2.74 (dd,  $J = 9.5, 14.0$  Hz, 1H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  169.1, 156.4, 149.6, 147.6, 143.7, 143.6, 141.0, 139.4, 132.2, 130.6, 127.8, 127.2, 125.1, 120.8, 120.0, 119.7, 67.5, 67.0, 56.4, 47.0, 24.1, 20.8, 14.7 ESI-MS:  $\text{C}_{30}\text{H}_{28}\text{N}_4\text{O}_{12}\text{S}_2$   $[\text{M}+\text{H}]^+$   $m/z$  found 701.1, calcd 701.1. Data consistent with literature.<sup>65</sup>

**(R)-2-(((fluoren-9-yl)methoxy)carbonylamido)-3-((S)-3-(allyloxy)-2-(allyloxycarbonylamino)-3-oxopropylthio)propanoic acid, Lan (29)**

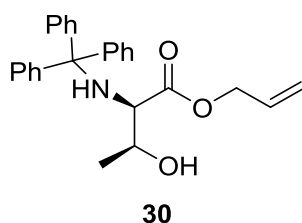


To a solution of **28** (0.10 g, .13 mmol) in DCM (5 mL) was added thioglycolic acid (12  $\mu\text{L}$ , .67 mmol) followed by DIPEA (0.13 mL, 0.80 mmol). The resulting mixture was warmed to room temperature and stirred for 2h. Removal of solvent *in vacuo* afforded a crude orange oil. This

was purified by reverse phase chromatography and the 30 and 50% acetone in water fractions were removed *in vacuo* affording a yellow solid. This was dissolved in DCM and to this solution was added Alloc-Cl (0.17 ml, 1.60 mmol) and K<sub>2</sub>CO<sub>3</sub> (73 mg, 0.60 mmol) at 0 °C. This was allowed to warm to room temperature and stirred overnight. Removal of solvent *in vacuo* afforded a crude yellow oil, which was purified by flash chromatography (SiO<sub>2</sub>, 3: 7: 0.2/EtOAc:Hexanes:HOAc) to yield **29** (.04 g, 52%) as a yellow oil.  $[\alpha]_D^{25} = -5.61^\circ$ ; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  7.76 (d, 2H, *J* = 8.4 Hz, Fmoc-ArH), 7.74 (d, 1H, *J* = 8.4 Hz, NH), 7.60 (d, 2H, *J* = 7.3 Hz, Fmoc-ArH), 7.44 (br s, 1H, NH), 7.39 (t, 2H, *J* = 7.8 Hz, Fmoc-ArH), 7.30 (t, 2H, *J* = 7.3 Hz, Fmoc-ArH), 5.88 (m, 2H, -CH<sub>2</sub>CH=CH<sub>2</sub>), 5.29 (d, 1H, *J* = 17.3 Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 5.26 (d, 1H, *J* = 17.5 Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 5.18 (d, 1H, *J* = 11.1 Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 5.14 (d, 1H, *J* = 11.1 Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 4.60 (d, 2H, *J* = 5.6 Hz, CH<sub>2</sub>CH=CH<sub>2</sub>), 4.41 (d, 2H, *J* = 3.9 Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 4.22 (m, 4H, Fmoc-CH<sub>2</sub>, Fmoc-CH and H $\alpha$ ), 3.99 (br m, 1H, H $\alpha$ ), 3.06 (m, 2H, H $\beta$ ), 2.82 (m, 2H, H $\beta$ ); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz):  $\delta$  174.2, 170.3, 156.3, 156.2, 143.8, 143.6, 141.3, 132.4, 131.2, 127.7, 127.0, 125.1, 120.1, 120.0, 118.0, 67.6, 66.6, 66.2, 58.6, 53.8, 47.1, 29.4; ESI-MS: C<sub>28</sub>H<sub>30</sub>N<sub>2</sub>O<sub>8</sub>S [M+H]<sup>+</sup> *m/z* found 555.1, calcd 555.1. Data consistent with literature.<sup>65</sup>

## 6.2.2 Synthesis of MeLan intermediates and starting material

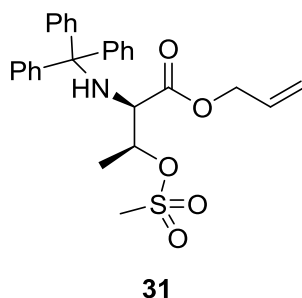
### (2*R*)-*N*-Triphenylmethyl-(*S*)-threonine allyl ester (**30**)



NaHCO<sub>3</sub> (0.55 g, 6.49 mmol) was added to a solution of tritylthreonine.triethylamine (1.50 g, 3.24 mmol) in DMF (30 mL). The solution was cooled to 0°C. Allyl bromide (0.34 mL, 3.88 mmol) was added dropwise and the solution was allowed to warm to room temperature overnight. The solution was quenched with water (150 mL) then extracted with Et<sub>2</sub>O (4 x 50 mL). The organic layers were combined, dried over MgSO<sub>4</sub>, filtered and then concentrated *in vacuo*. The crude product used without further purification to yield **29** (1.40 g, 54 %) as a colourless oil.  $[\alpha]_D^{25} = +32^\circ$ ; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  7.48 (d, 6H, Trt), 7.27 (t, 6H, Trt), 7.19 (t, 3H, Trt), 5.66 (m, 1H, CH=CH<sub>2</sub>), 5.19 (dq, 2H, *J* = 17.2, 1.5 Hz, CH=CH<sub>2</sub>), 5.16 (dq, 1H, *J* = 10.5 Hz, 1.5 Hz, CH=CH<sub>2</sub>), 4.07 and 4.06 (2dd, 1H, *J* = 13.2, 5.9 Hz, CH<sub>2</sub>-CHCH<sub>2</sub>), 3.71 (m, 1H, HR), 3.85 and 3.80 (2dd, 2H, *J* = 10.4, 5.6 Hz, CH<sub>2</sub>  $\alpha$ ), 3.16 (s, 1H, OH) 1.22 (d,

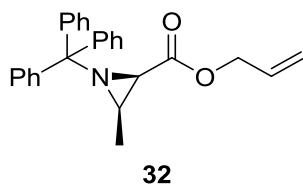
3H, CH<sub>3</sub>, 1.22); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 172.9 (CO), 145.4 (C, arom. Trt), 131.5 (CHCH<sub>2</sub>), 128.9, 127.9, 126.6 (C arom. Trt), 118.9 (CHCH<sub>2</sub>), 70.7 (C. aliph. Trt), 69.8 (CH<sub>2</sub>-CHCH<sub>2</sub>), 65.6 (Cβ), 62.5 (Cα), 18.9 (CH<sub>3</sub>); ESI-MS: C<sub>26</sub>H<sub>27</sub>NO<sub>3</sub> [M+H]<sup>+</sup> *m/z* found 402.2, calcd 402.2. Data consistent with literature.<sup>65</sup>

***N*-Triphenylmethyl-(*RS*)-threonine-(*O*-methanesulfonyl) allyl ester (**31**)**



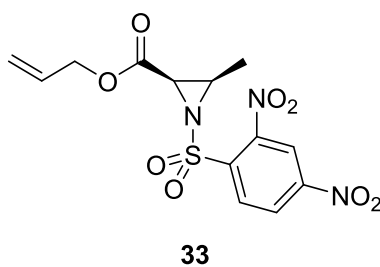
A sample of **30** (1.35 g, 3.35 mmol) was dissolved in DCM (30 mL) and cooled to 0°C. NEt<sub>3</sub> (1.07 mL, 7.71 mmol) was added and after 5 minutes Ms-Cl (0.31 mL, 4.02 mmol) was added dropwise. The solution was allowed to warm to room temperature overnight. The solution was quenched with NH<sub>4</sub>Cl (sat.) solution (35 mL) and extracted with DCM (3 x 50 mL). The organic layers were combined and washed with brine (50 mL), dried over MgSO<sub>4</sub>, filtered and then concentrated *in vacuo* to give a clear foamy oil. The crude product was purified by flash chromatography (SiO<sub>2</sub>, 20% EtOAc in Hexanes) to yield **30** (0.71 g, 44 %) as a white crystalline solid. [α]<sub>D</sub><sup>25</sup> = 9.0; <sup>1</sup>H NMR (500 MHz, DMSO-d<sub>6</sub>) δ 7.41 (d, 6H, Trt), 7.27 (t, 6H, Trt), 7.19 (t, 3H, Trt), 5.66 (m, 1H, CH=CH<sub>2</sub>), 5.15 (dd, 1H, *J* = 5 Hz, 1.5 Hz CH=CH<sub>2</sub>) 5.00 (dd, 1H, *J* = 10.5, 1.5 Hz, CH=CH<sub>2</sub>), 4.04 and 3.78 (2dd, 2H, *J* = 13.2, 5.9 Hz, CH<sub>2</sub>-CHCH<sub>2</sub>), 4.02 (m, 1H, HR), 3.87 and 3.68 (2 dd, 2H, *J* = 10.4 Hz, 5.6 Hz, CH<sub>2</sub> α), 3.35 (s, 3H, SCH<sub>3</sub>) 3.14 (s, 1H, OH) 2.49 (s, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 170.7 (CO), 145.7 (C, arom. Trt), 132.2 (CHCH<sub>2</sub>), 128.9, 128.3, 127.0 (C arom. Trt), 118.4 (CHCH<sub>2</sub>), 79.0 (C. aliph. Trt), 70.9 (CH<sub>2</sub>-CHCH<sub>2</sub>), 65.5 (Cα), 55.0 (CR), 38.2 (SCH<sub>3</sub>), 16.89 (CH<sub>3</sub>). ESI-MS: C<sub>27</sub>H<sub>29</sub>NO<sub>5</sub>S [M+Na]<sup>+</sup> *m/z* found 502.2, calcd 502.2. Data consistent with literature.<sup>65</sup>

**(2*R*,3*R*)-1-Triphenylmethyl-3-methyl-aziridine-2-carboxylic acid allyl ester (32)**



A small amount of **31** (0.67 g, 1.38 mmol) was dissolved in DMF (60 mL). NEt<sub>3</sub> (0.23 mL, 1.67 mmol) and DMAP (17 mg, 0.14 mmol) was added and the reaction was heated overnight at 70°C. The reaction was cooled to room temperature, diluted with water (40 mL) and extracted with Et<sub>2</sub>O (4 x 30 mL). The organic layers were combined and washed with brine (30 mL), dried over MgSO<sub>4</sub>, filtered and then concentrated *in vacuo* to give colourless foamy oil. The crude product was purified by flash chromatography (SiO<sub>2</sub>, 10% EtOAc in Hexanes) to yield **31** (0.50 g, 53%) as a colourless oil. [ $\alpha$ ]<sub>D</sub><sup>25</sup> = -0.07; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  7.52 (d, 6H, *J* = 5 Hz, Trt), 7.28 (t, 6H, *J* = 6 Hz, Trt), 7.21 (t, 3H, *J* = 7 Hz, Trt), 5.93 (m, 1H, CHCH<sub>2</sub>), 5.33 (dq, 1H, *J* = 17.2 Hz, 1.5 Hz, CHCH<sub>2</sub>), 5.25 (dq, 1H, *J* = 10.5, 1.2 Hz, CHCH<sub>2</sub>), 4.68 (m, 2H, CH<sub>2</sub>-CHCH<sub>2</sub>), 2.27 (s, 1H, H $\alpha$ ), 1.94 (dd, 1H, *J* = 2.7, 6.2 Hz, HR), 1.42 (dd, 1H, *J* = 6.2 Hz, 1.6 Hz, H $\alpha$ ) 1.22 (d, 3H, *J* = 1.38 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  170.6 (CO), 144.2 (Cq arom.), 131.7 (CHCH<sub>2</sub>), 129.4, 127.6, 126.8 (C arom.), 118.5 (CHCH<sub>2</sub>), 74.9 (Cq aliph.), 65.3 (CH<sub>2</sub>-CHCH<sub>2</sub>), 36.7 (CR), 34.92 (CH<sub>3</sub>), 28.7 (C $\alpha$ ); ESI-MS: C<sub>26</sub>H<sub>25</sub>NO<sub>2</sub> [M+Na]<sup>+</sup> *m/z* found 406.2 calcd 406.2. Data consistent with literature.<sup>65</sup>

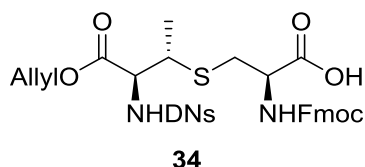
**(2*R*,3*R*)-Allyl 1-(2,4-dinitrobenzenesulfonyl)-3-methylaziridine-2-carboxylate (33)**



TFA (1.25 mL, 8.7 mmol) was added dropwise at 0 °C to a solution of **32** (0.47 g, 1.23 mmol) in DCM (6 mL) and MeOH (2 mL). The solution was stirred for 2 hours at 0 °C and concentrated *in vacuo*. The resulting residue was partitioned between Et<sub>2</sub>O (40 mL) and H<sub>2</sub>O (20 mL) and the ether layer was extracted with water (3 x 10 mL). The combined aqueous layers were basified to pH 9 with Na<sub>2</sub>CO<sub>3</sub> at 0 °C. EtOAc (40 mL) was added to the aqueous solution followed by 2,4-dinitrobenzenesulfonyl chloride (0.82 g, 3.07 mmol) at 0 °C. The biphasic system was warmed to room temperature and stirred vigorously for 24 h. The two

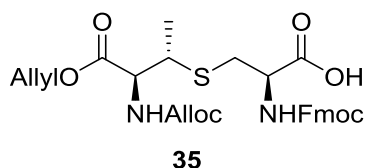
layers were then separated and the aqueous layer was extracted with EtOAc (3 x 10 mL). The organic layers were combined, dried over MgSO<sub>4</sub>, filtered and then concentrated *in vacuo*. The crude product was purified by flash chromatography (SiO<sub>2</sub>, 20% EtOAc in Hexanes) to yield **32** (.26 g, 57 %) as a light yellow oil.  $[\alpha]_D^{25} = +30.3^\circ$ ; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): 8.66 (d, 1H, *J* = 2.0, Ar-H), 8.59 (dd, 1H, *J* = 8.8, 2.4 Hz, Ar-H), 8.55 (dd, 1H, *J* = 8.8, 0.4 Hz, Ar-H), 5.97-5.87 (m, 1H, -CH<sub>2</sub>CH=CH<sub>2</sub>), 5.37 (d, 1H, *J* = 15.6 Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 5.33 (d, 1H, *J* = 10.4 Hz, -CH<sub>2</sub>CH=CH<sub>2</sub>), 4.71-4.68 (m, 2H, -CH<sub>2</sub>CH=CH<sub>2</sub>), 3.81 (d, 1H, *J* = 7.6 Hz, H<sub>α</sub>), 3.48 (qd, 1H, *J* = 5.6, 7.6 Hz, H<sub>β</sub>), 1.46 (d, 2H, *J* = 5.6 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz): δ 164.5, 149.8, 148.4, 137.4, 133.2, 130.9, 126.9, 120.1, 119.3, 66.6, 44.0, 43.3, 12.4; ESI-MS: C<sub>13</sub>H<sub>13</sub>N<sub>3</sub>O<sub>8</sub>S [M+H]<sup>+</sup> *m/z* found 372.0, calcd 372.0. Data consistent with literature.<sup>65</sup>

**(*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamido)-3-((2*S*,3*S*)-4-(allyloxy)-3-(2,4-dinitrobenzenesulfonamido)-4-oxobutan-2-ylthio)propanoic acid (**34**)**



To a solution of **33** (500 mg, 1.3 mmol) in DCM (10 mL) was added Fmoc-Cys-OH (1.85 g, 5.3 mmol) followed by BF<sub>3</sub>·OEt<sub>2</sub> (1.32 mL, 11.0 mmol) at 0 °C. The resulting mixture was warmed to room temperature and stirred for 48h. Removal of solvent *in vacuo* afforded a crude yellow oil, which was purified by flash chromatography (SiO<sub>2</sub>, 3:7:0.2/EtOAc:Hexanes:HOAc) to yield **33** (0.35 g, 38 %) as a light yellow oil.  $[\alpha]_D^{25} = -15.5^\circ$ ; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 8.67 (s, 1H, DN<sub>s</sub>-H), 8.46 (dd, 1H, *J* = 8.4, 2.4 Hz, DN<sub>s</sub>-H), 8.27 (d, 1H, *J* = 8.4 Hz, DN<sub>s</sub>-H), 7.77 (d, 2H, *J* = 7.8 Hz, Fmoc-ArH), 7.60 (m, 2H, Fmoc-ArH), 7.40 (t, 2H, *J* = 7.2 Hz, Fmoc-ArH), 7.31 (t, 2H, *J* = 7.2 Hz, Fmoc-ArH), 6.68 (d, 1H, *J* = 9.0 Hz, NH), 5.70 (m, 1H, CH<sub>2</sub>CH=CH<sub>2</sub>), 5.61 (d, 1H, *J* = 8.4 Hz, NH), 5.22 (d, 2H, *J* = 18.0 Hz, CH<sub>2</sub>CH=CH<sub>2</sub>), 5.18 (d, 2H, *J* = 10.8 Hz, CH<sub>2</sub>CH=CH<sub>2</sub>), 4.60 (app d, 1H, *J* = 5.4 Hz, H<sub>α</sub>), 4.48-4.39 (m, 4H, Fmoc-CH<sub>2</sub> and -CH<sub>2</sub>CH=CH<sub>2</sub>), 4.36 (d, 1H, *J* = 7.2 Hz, H<sub>α</sub>), 4.23 (t, 1H, *J* = 6.6 Hz, Fmoc-CH), 3.49 (br s, 1H, H<sub>β</sub>), 3.03 (dd, 1H, *J* = 13.2, 3.6 Hz, H<sub>β</sub>), 2.14 (dd, 1H, *J* = 13.2, 5.4 Hz, H<sub>β</sub>), 1.44 (d, 3H, *J* = 6.6 Hz, -CH<sub>3</sub>); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz): δ 177.5, 169.1, 155.9, 149.8, 147.7, 143.6, 143.5, 141.3, 139.6, 132.0, 130.6, 127.8, 127.1, 125.0, 120.9, 120.06, 120.04, 67.4, 67.0, 61.7, 54.1, 47.1, 43.6, 33.7, 19.6; ESI-MS: C<sub>31</sub>H<sub>30</sub>N<sub>4</sub>O<sub>12</sub>S<sub>2</sub> [M+Na]<sup>+</sup> *m/z* found 737.1, calcd 737.1. Data consistent with literature.<sup>65</sup>

**(R)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2S,3S)-4-(allyloxy)-3-(allyloxycarbonylamino)-4-oxobutan-2-ylthio)propanoic acid, MeLan (35)**



To a solution of **34** (0.100 g, .14 mmol) in DCM (5 mL) was added thioglycolic acid (12  $\mu$ L, 0.67 mmol) followed by DIPEA (0.13 mL, 0.8 mmol). The resulting mixture was warmed to room temperature and stirred for 2 h. Removal of solvent *in vacuo* afforded a crude orange oil. This was purified by reverse phase chromatography and the 30 and 50% acetone in water fractions were removed *in vacuo* affording a yellow solid. This was dissolved in DCM (5 mL) and to this solution was added Alloc-Cl (0.17 mL, 1.60 mmol) and  $K_2CO_3$  (0.07 g, 0.60 mmol) at 0  $^{\circ}C$ . This was allowed to warm to room temperature and stirred overnight. Removal of solvent *in vacuo* afforded a crude yellow oil, which was purified by flash chromatography ( $SiO_2$ , 3:7:0.2/EtOAc:Hexanes:HOAc) to yield **29** (0.04 g, 46%) as a yellow oil.  $[\alpha]_D^{25} = -0.23^{\circ}$ ;  $^1H$  NMR ( $CDCl_3$ , 500 MHz):  $\delta$  7.77 (d, 2H,  $J = 7.7$  Hz, Fmoc-ArH), 7.62 (d, 2H,  $J = 7.7$  Hz, Fmoc-ArH), 7.61 (br d, 1H,  $J = 8.3$  Hz, NH), 7.40 (t, 2H,  $J = 7.1$  Hz, Fmoc-ArH), 7.32 (m, 3H,  $J = 7.1$  Hz, Fmoc-H and NH), 5.90 (m, 2H,  $-CH_2CH=CH_2$ ), 5.35 (d, 1H,  $J = 17.1$  Hz,  $-CH_2CH=CH_2$ ), 5.27 (d, 1H,  $J = 17.9$  Hz,  $-CH_2CH=CH_2$ ), 5.20 (d, 1H,  $J = 10.2$  Hz,  $-CH_2CH=CH_2$ ), 5.18 (d, 1H,  $J = 10.5$  Hz,  $-CH_2CH=CH_2$ ), 4.58 (d, 2H,  $J = 5.2$  Hz,  $-CH_2CH=CH_2$ ), 4.66 (d, 2H,  $J = 2.7$  Hz,  $-CH_2CH=CH_2$ ), 4.57 (m, 4H, Fmoc- $CH_2$ , Fmoc-CH and H $\alpha$ ), 4.41 (m, 1H, H $\alpha$ ), 4.24 (t, 1H,  $J = 8.5$  Hz, H $\beta$ ), 3.79 (d, 1H,  $J = 13.2$  Hz, H $\beta$ ), 3.01 (t, 1H,  $J = 9.9$  Hz, H $\beta$ );  $^{13}C$  NMR ( $CDCl_3$ , 125 MHz):  $\delta$  170.8, 156.3, 143.7, 141.2, 132.4, 131.2, 127.7, 127.1, 125.1, 120.1, 119.5, 117.9, 67.2, 66.5, 66.4, 58.6, 47.1, 29.1, 19.3; ESI-MS:  $C_{29}H_{32}N_2O_8S$   $[M+H]^+$   $m/z$  found 569.2, calcd 569.2. Data consistent with literature.<sup>65</sup>

## 6.3 SPPS procedures

### 6.3.1 General procedure for Fmoc SPPS

Wang resin (0.180 g, 0.1 mmol/g,  $1.8 \times 10^{-2}$  mmol, 74 – 149  $\mu m$ ) came preloaded with Fmoc-Leu-OH. The desired Fmoc amino acid ( $3.61 \times 10^{-2}$  mmol, 3 equiv to resin loading) and HOBt (8.3 mg,  $3.61 \times 10^{-2}$  mmol, 3 equiv) were dissolved in 50% DCM/DMF (5 mL) and preactivated for 5 minutes. This was transferred to the resin followed by DIC (8.5  $\mu$ L,  $3.61 \times 10^{-2}$ , 3 equiv) and reacted for 30 minutes. The solution was drained from the resin. The completion of

coupling was ascertained by a negative Kaiser test. Subsequent Fmoc deprotection was achieved using 20% piperidine in DMF (5 ml) and reacting for 30 minutes. A washing step (5 x 5 ml DMF) followed both the Fmoc deprotecting and coupling steps.

### **6.3.2 Kaiser test procedure**

The stock solutions for the Kaiser test were made up as follows. Ninhydrin (5 g) in ethanol (100 ml). KCN (2 ml 0.001 M) in pyridine (98 ml). Phenol (80 g) in ethanol (20 ml).

A sample of resin was taken and placed in a small test tube. A few drops of each stock solution were added to the beads. This was warmed in a water bath (95 °C) for 5 minutes. Optical analysis of the beads was implemented to determine the result.

### **6.3.4 Addition of (Me)Lan**

The (Me)Lan ( $3.61 \times 10^{-2}$  mmol, 3 equiv to resin loading) and HOBt (8.3 mg,  $3.61 \times 10^{-2}$ , 3 equiv to resin) were dissolved in 50% DCM/DMF (5 ml) and preactivated for 5 minutes. This was transferred to the resin followed by DIC (8.5  $\mu$ l,  $3.61 \times 10^{-2}$  mmol, 3 equiv to resin) and reacted for 3 hours. The resin was washed (5 x 5 ml DCM and DMF) and completion of coupling was ascertained by a negative Kaiser test.

### **6.3.5 Deprotection of allyl, Fmoc groups and cyclisation for ring formation**

A solution of Pd(PPh<sub>3</sub>)<sub>4</sub> (4 mg, 0.003 mmol, 0.2 equiv to resin) and PhSiH<sub>3</sub> (22  $\mu$ l, 0.018 mmol, 10 equiv to resin) in 50 % DMF/CH<sub>2</sub>Cl<sub>2</sub> (5 mL) was reacted with the resin bound peptide, protected from light, for 2 h under argon. The solution was drained and the resin was washed in the following sequence: 1) DCM (10 x 5 mL), 2) 50% DCM/DMF (10 x 10 mL), 3) DMF (10 x 10 mL). The colour of the resin changed to light yellow after these washings. The Fmoc group was removed with 20% piperidine in DMF (5 mL). The resin was washed with DMF (5 x 5 mL). The cyclization to form the ring was done by adding a solution of HOBt (13.8 mg,  $9 \times 10^{-2}$  mmol, 5.0 equiv to resin) and DIC (22  $\mu$ l,  $9 \times 10^{-2}$  mmol, 5.0 equiv to resin) in 50% DCM/DMF (5 mL) to the linear peptide and reacted overnight.

### **6.3.6 Final cleavage**

The resin was washed with DCM (5 x 5 ml) to remove the presence of DMF. (95: 2: 2: 1) TFA: H<sub>2</sub>O: Phenol: TIPS was mixed in a graduated cylinder and added to the resin. This was agitated for 2 hours. The cleavage mixture was filtered to remove the resin and concentrated under



reduced pressure. Diethyl ether was added followed by cooling in a liquid nitrogen bath to allow precipitation. This afforded an unknown white crystalline solid.

## **6.4 Microbial imaging mass spectrometry**

The culture media area of interest was cut out using a scalpel and transferred to a scratched MALDI target plate and photographed. Matrix (CHCA and DHB purchased from Sigma) was sieved through a 60  $\mu\text{m}$  nylon mesh directly on the culture media sample. This was dehydrated at 37 °C for 4 hours and then lyophilized overnight. The dried sample was photographed again to use in determination of imaging area. The area was chosen with 100  $\mu\text{m}$  resolution and then subjected to MALDI-TOF IMS.

The MALDI-TOF mass spectrometry data were obtained on a Voyager DE pro MALDI-TO mass spectrometer (Applied Biosystems, Foster City, CA, USA) equipped with a 337 nm nitrogen laser operating at 3 Hz. The softwares AB Sciex TOF/TOF™ SeriesExplorer™ Software version 4.1.0 and AB Sciex TOF/TOF Imaging Acquisition Software 1.0 were used, measure across a 0 to 1000  $m/z$  and 800 – 4000  $m/z$  mass range.

The Visualising software used was Biomap.

## **Appendix A - Microbial culture media components**

### **AOM1 Minimal Media**

Bacteria were cultured at pH 6.5 at 60 °C

$(\text{NH}_4)_2\text{SO}_4$  – 4.0 g

$\text{K}_2\text{HPO}_4$  – 0.45 g

$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  – 0.60 g

$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  – 0.005 g

FeEDTA solution – 3 mL

Trace metal solution (Methanotrophs) – 3 mL

Methanogen trace element solution (Wolin) – 1 mL

Phyagel – 1.0 g

Yeast extract – 0.1 g

B Forte Vitamins – 0.1 g

Phytigel – 15 g

$\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$  – 1 g

Distilled  $\text{H}_2\text{O}$  – 500 mL

## Appendix B – SPPS checklist

Step	Mass or mL	Time started and finished	Check
1 Soak resin in CH <sub>2</sub> Cl <sub>2</sub> (30 mins)	180mg/5 mL		
2 Fmoc Cleavage (30 mins)	5 mL 20% Pip/Dmf		
3 Wash	5 x 5 mL DMF		
4 Add MeLan with activator (2 hrs)	30.7 mg 5 x 5 mL		
5 Wash	DCM+DMF		
6 Kaiser test (5 min)			
	5 mL 20% Pip/Dmf		
7 Fmoc Cleavage (30 mins)			
8 Wash	5 x 5 mL DMF		
9 Add Fmoc-glycine (30mins)	16.1 mg		
10 Wash	5 x 5 mL DMF		
11 Kaiser test (5 min)			
	5 mL 20% Pip/Dmf		
12 Fmoc Cleavage (30 mins)			
13 Wash	5 x 5 mL DMF		
14 Add Fmoc-glycine (30mins)	16.1 mg		
15 Wash	5 x 5 mL DMF		
16 Kaiser test (5 min)			
	5 mL 20% Pip/Dmf		
17 Fmoc Cleavage (30 mins)			
18 Wash	5 x 5 mL DMF		
19 Add Fmoc-Alanine (30 mins)	16.8 mg		
20 Wash	5 x 5 mL DMF		
21 Kaiser test (5 min)			
	5 mL 20% Pip/Dmf		
22 Fmoc Cleavage (30 mins)			
23 Wash	5 x 5 mL DMF		
24 Add Fmoc-Alanine (30 mins)	16.8 mg		
25 Wash	5 x 5 mL DMF		
26 Kaiser test (5 min)			
	5 mL 20% Pip/Dmf		
27 Fmoc Cleavage (30 mins)			
28 Wash	5 x 5 mL DMF		
29 Add Fmoc-Asparagine (30 mins)	32.3 mg		
30 Wash	5 x 5 mL DMF		
31 Kaiser test (5 min)			
32 Allyl/Alloc deprotection	22 uL Si/4 mg Pd 5 x 5 mL		
33 Wash	DCM+DMF		
34 Fail Kaiser test (5 min)			

35	Fmoc Cleavage (30 mins)	5 mL 20% Pip/Dmf
36	Wash	5 x 5 mL DMF
37	Add activator (3 hrs)	
38	Wash	5 x 5 mL DMF
39	Add Fmoc-Asparagine (30 mins)	32.3 mg
40	Wash	5 x 5 mL DMF
41	Kaiser test (5 min)	
42	Fmoc Cleavage (30 mins)	5 mL 20% Pip/Dmf
43	Wash	5 x 5 mL DMF
44	Add Lan with activator (2 hrs)	30 mg
45	Wash	5 x 5 mL DMF
46	Kaiser test (5 min)	
47	Fmoc Cleavage (30 mins)	5 mL 20% Pip/Dmf
48	Wash	5 x 5 mL DMF
49	Add Fmoc-leucine (30 mins)	19.1 mg
50	Wash	5 x 5 mL DMF
51	Kaiser test (5 min)	
52	Fmoc Cleavage (30 mins)	5 mL 20% Pip/Dmf
53	Wash	5 x 5 mL DMF
54	Add Fmoc-glycine (30mins)	16.1 mg
55	Wash	5 x 5 mL DMF
56	Kaiser test (5 min)	
57	Fmoc Cleavage (30 mins)	5 mL 20% Pip/Dmf
58	Wash	5 x 5 mL DMF
59	Add Fmoc-leucine (30 mins)	19.1 mg
60	Wash	5 x 5 mL DMF
61	Kaiser test (5 min)	
62	Fmoc Cleavage (30 mins)	5 mL 20% Pip/Dmf
63	Wash	5 x 5 mL DMF
64	Add Fmoc-leucine (30 mins)	19.1 mg
65	Wash	5 x 5 mL DMF
66	Kaiser test (5 min)	
67	Fmoc Cleavage (30 mins)	5 mL 20% Pip/Dmf
68	Wash	5 x 5 mL DMF
69	Add Fmoc-valine (30 mins)	18 mg
70	Wash	5 x 5 mL DMF
71	Kaiser test (5 min)	
72	Fmoc Cleavage (30 mins)	5 mL 20% Pip/Dmf
73	Wash	5 x 5 mL DMF
74	Add Fmoc-Alanine (30 mins)	16.8 mg

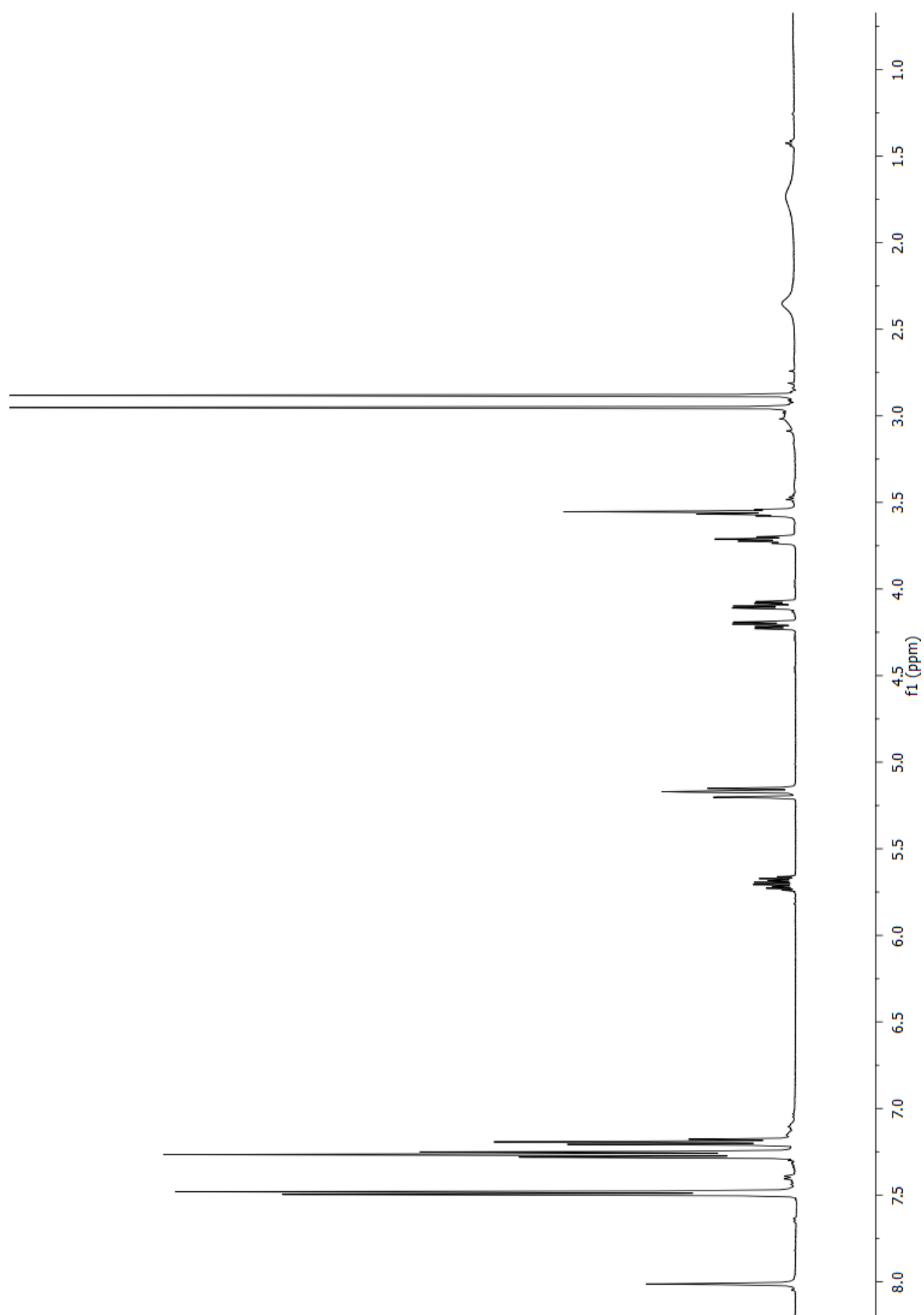
75	Wash	5 x 5 mL DMF
76	Kaiser test (5 min)	
		5 mL 20%
77	Fmoc Cleavage (30 mins)	Pip/Dmf
78	Wash	5 x 5 mL DMF
79	Add Fmoc-leucine (30 mins)	19.1 mg
80	Wash	5 x 5 mL DMF
81	Kaiser test (5 min)	
82	Allyl/Alloc deprotection	22 uL Si/20.8 mg Pd
		5 x 5 ml
83	Wash	DMF+DCM
84	Fail Kaiser test (5 min)	
		5 mL 20%
85	Fmoc Cleavage (30 mins)	Pip/Dmf
86	Wash	5 x 5 mL DMF
87	Add activator (3 hrs)	
88	Wash	5 x 5 mL DMF
89	Add Fmoc-glutamine	33 mg
90	Wash	5 x 5 mL DMF
91	Kaiser test (5 min)	
		5 mL 20%
92	Fmoc Cleavage (30 mins)	Pip/Dmf
93	Wash	5 x 5 mL DMF
94	Add Fmoc-leucine (30 mins)	19.1 mg
95	Wash	5 x 5 mL DMF
96	Kaiser test (5 min)	
		5 mL 20%
97	Fmoc Cleavage (30 mins)	Pip/Dmf
98	Wash	5 x 5 mL DMF
99	Add MeLan with activator (2 hrs)	30.7 mg
100	Wash	5 x 5 mL DMF
101	Kaiser test (5 min)	
		5 mL 20%
102	Fmoc Cleavage (30 mins)	Pip/Dmf
103	Wash	5 x 5 mL DMF
104	Add Fmoc-leucine (30 mins)	19.1 mg
105	Wash	5 x 5 mL DMF
106	Kaiser test (5 min)	
		5 mL 20%
107	Fmoc Cleavage (30 mins)	Pip/Dmf
108	Wash	5 x 5 mL DMF
109	Add Fmoc-valine (30 mins)	18 mg
110	Wash	5 x 5 mL DMF
111	Kaiser test (5 min)	
		5 mL 20%
112	Fmoc Cleavage (30 mins)	Pip/Dmf
113	Wash	5 x 5 mL DMF
114	Add Fmoc-Alanine (30 mins)	16.8 mg

115	Wash	5 x 5 mL DMF
116	Kaiser test (5 min)	
		5 mL 20%
117	Fmoc Cleavage (30 mins)	Pip/Dmf
118	Wash	5 x 5 mL DMF
119	Add Fmoc-valine (30 mins)	18 mg
120	Wash	5 x 5 mL DMF
121	Kaiser test (5 min)	
122	Allyl/Alloc deprotection	22 uL Si/4 mg Pd
		5 x 5 mL
123	Wash	DMF+DCM
124	Fail Kaiser test (5 min)	
		5 mL 20%
125	Fmoc Cleavage (30 mins)	Pip/Dmf
126	Wash	5 x 5 mL DMF
127	Add activator (3 hrs)	
128	Wash	5 x 5 mL DMF
129	Add Fmoc-Asparagine (30 mins)	32.3 mg
130	Wash	5 x 5 mL DMF
131	Kaiser test (5 min)	
		5 mL 20%
132	Fmoc Cleavage (30 mins)	Pip/Dmf
133	Wash	5 x 5 mL DMF
134	Add Fmoc-Asparagine (30 mins)	32.3 mg
135	Wash	5 x 5 mL DMF
136	Kaiser test (5 min)	
		5 ml 20%
137	Fmoc Cleavage (30 mins)	Pip/Dmf
138	Wash	5 x 5 mL DMF
139	Add MeLan with activator (2 hrs)	30.7 mg
		5 x 5 mL
140	Wash	DMF+DCM
141	Kaiser test (5 min)	
		5 ml 20%
142	Fmoc Cleavage (30 mins)	Pip/Dmf
143	Wash	5 x 5 mL DMF
144	Add Fmoc-glutamic acid (30 mins)	23 mg
145	Wash	5 x 5 mL DMF
146	Kaiser test (5 min)	
		5 ml 20%
147	Fmoc Cleavage (30 mins)	Pip/Dmf
148	Wash	5 x 5 mL DMF
149	Add Fmoc-aspartic acid (30 mins)	22.3 mg
150	Wash	5 x 5 mL DMF
151	Kaiser test (5 min)	
152	Allyl/Alloc deprotection	22 uL Si/4 mg Pd
		5 x 5 mL
153	Wash	DMF+DCM

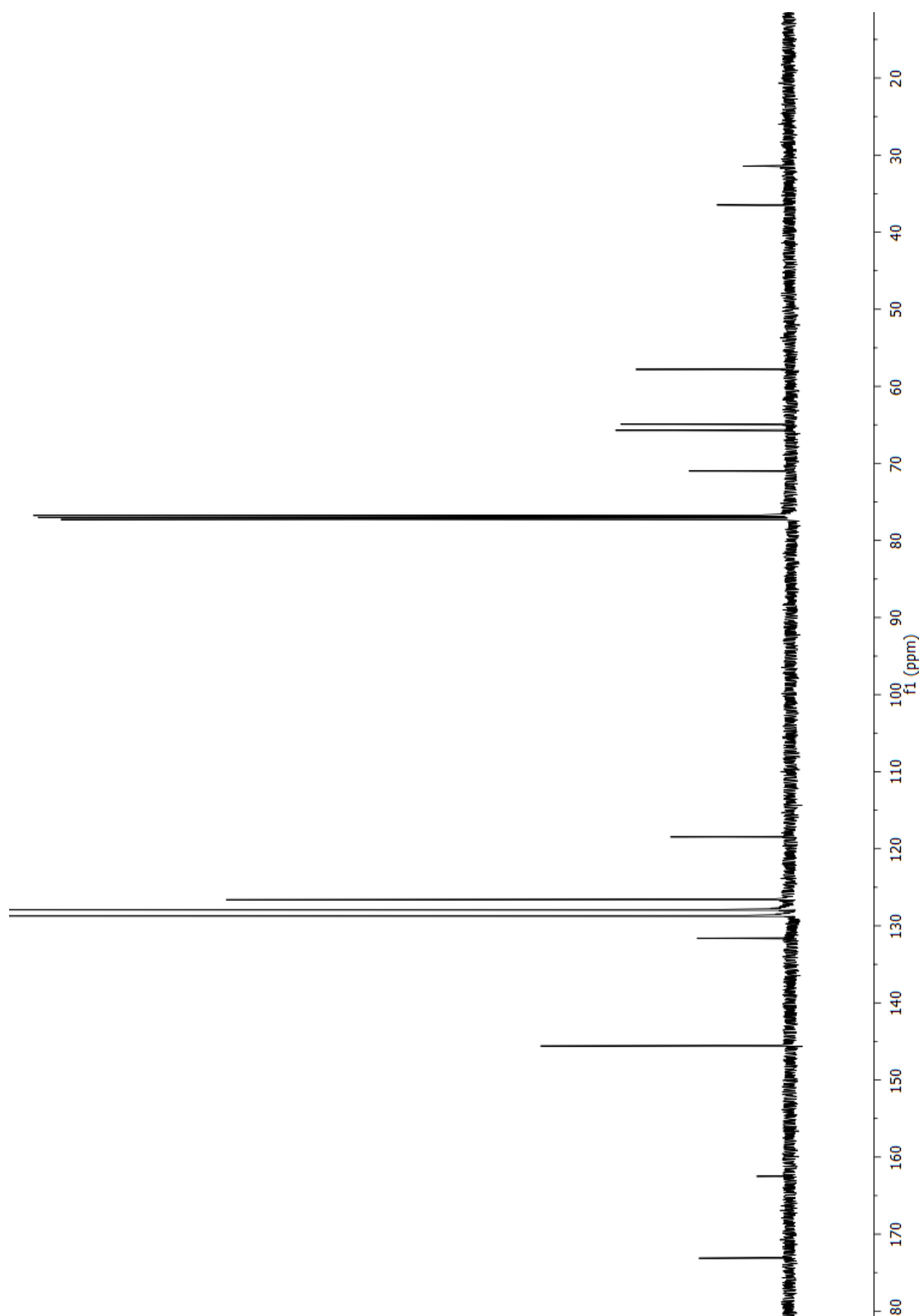
154	Fail Kaiser test (5 min)	5 mL 20%
155	Fmoc Cleavage (30 mins)	Pip/Dmf
156	Wash	5 x 5 mL DMF
157	Add activator (3 hrs)	
158	Wash	5 x 5 mL DMF
159	Add Fmoc-Asparagine (30 mins)	32.3 mg
160	Wash	5 x 5 mL DMF
161	Kaiser test (5 min)	
		5 mL 20%
162	Fmoc Cleavage (30 mins)	Pip/Dmf
163	Wash	5 x 5 mL DMF
164	Add Fmoc-glycine (30mins)	16.1 mg
165	Wash	5 x 5 mL DMF
166	Kaiser test (5 min)	
		5 mL 20%
167	Fmoc Cleavage (30 mins)	Pip/Dmf
168	Wash	5 x 5 mL DMF
169	Add Fmoc-histadine (30 mins)	33.5 mg
170	Wash	5 x 5 mL DMF
171	Kaiser test (5 min)	
		5 mL 20%
172	Fmoc Cleavage (30 mins)	Pip/Dmf
173	Wash	5 x 5 mL DCM
174	Cleave off resin	

## **Appendix C – Spectra of Lan intermediates**

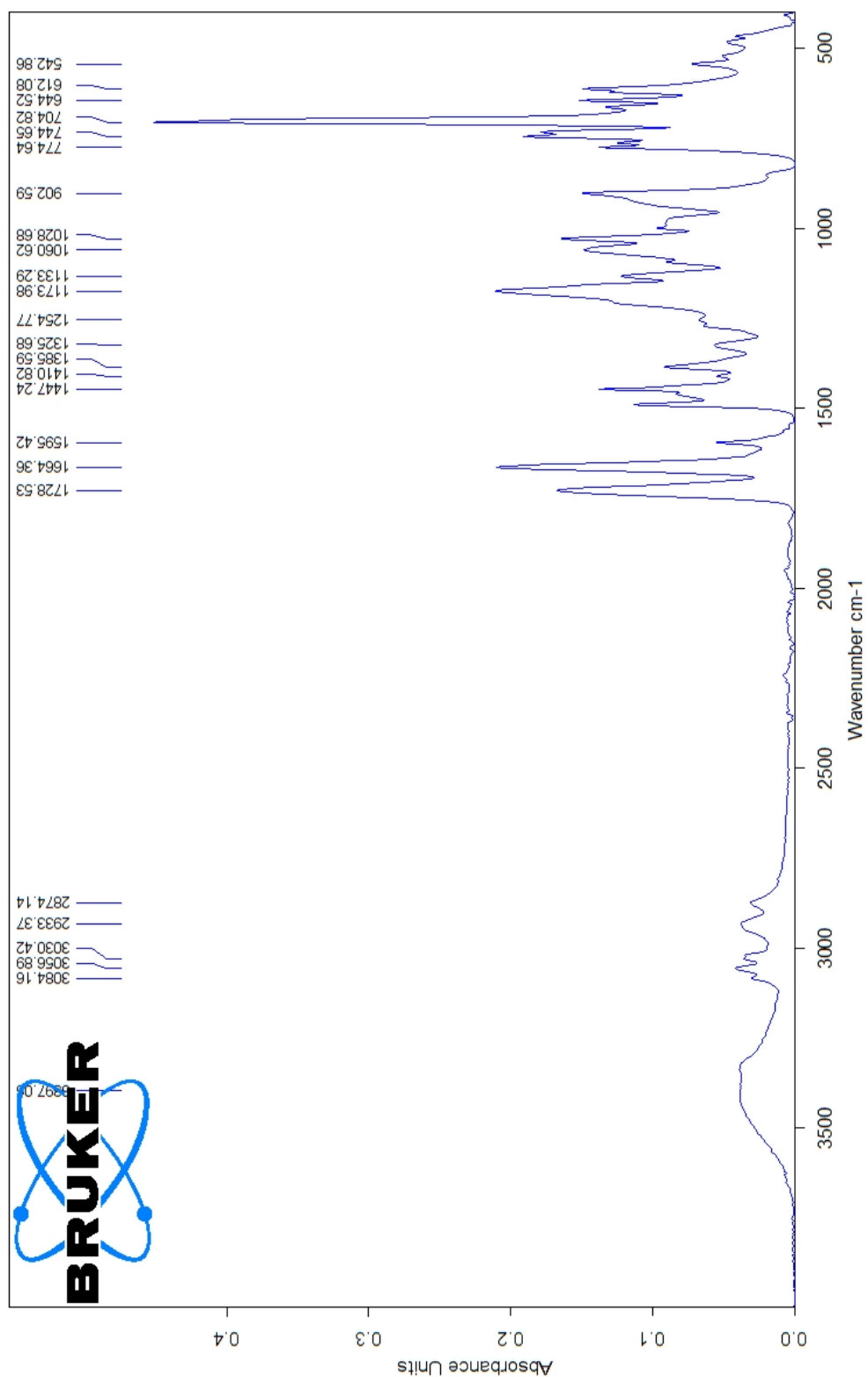




<sup>1</sup>H NMR spectrum (2*R*)-*N*-Triphenylmethylserine allyl ester (**24**) (500 MHz, CDCl<sub>3</sub>)

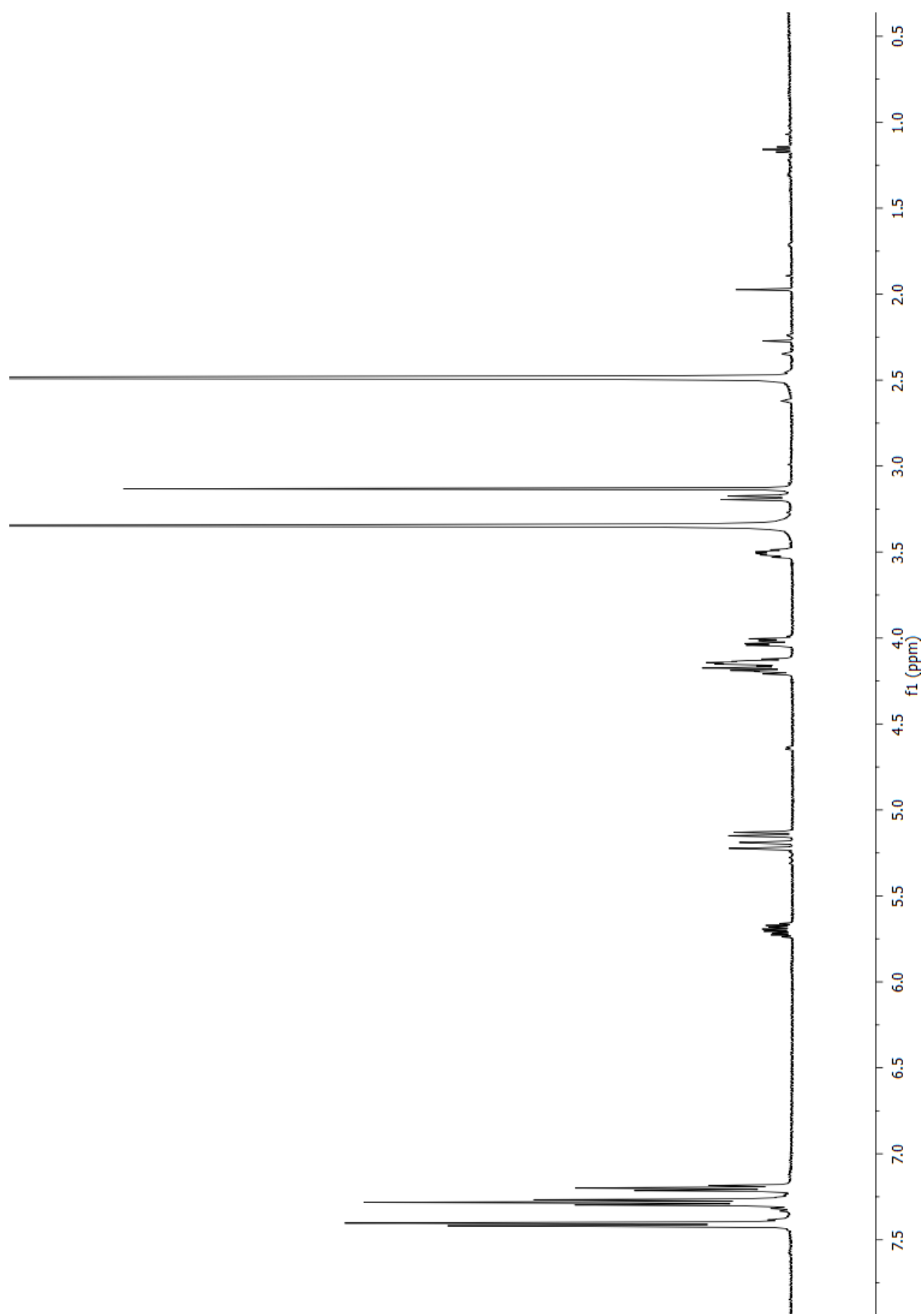


$^{13}\text{C}$  spectrum (2*R*)-N-Triphenylmethylserine allyl ester (**24**) (125 MHz, CDCl<sub>3</sub>)



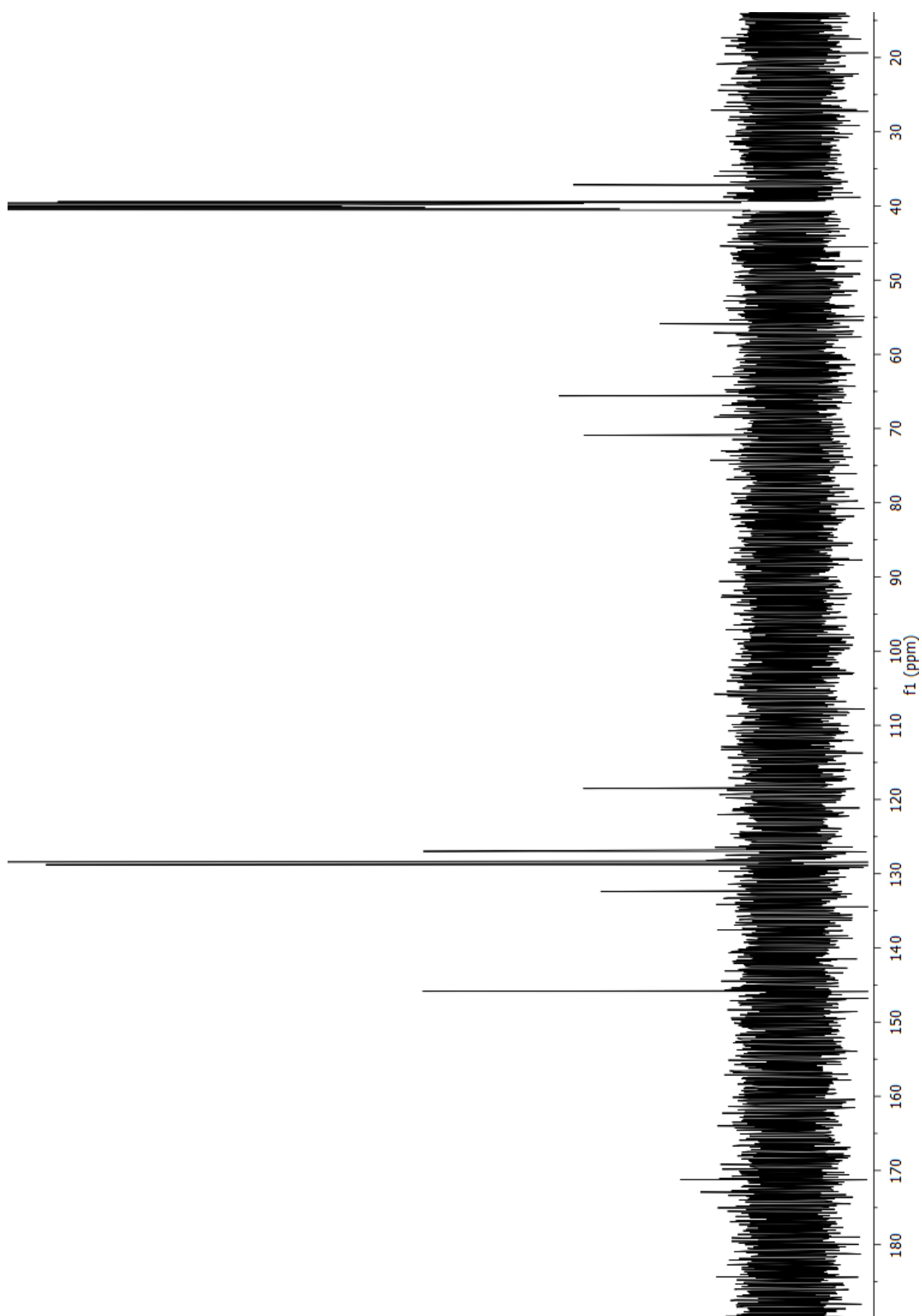
C:\OPUS_7.2\39.1294\MEAS\873	BPB02_56b Allylated serine	Instrument type and / or accessory	15/12/2015
------------------------------	----------------------------	------------------------------------	------------

IR spectrum of (2*R*)-*N*-Triphenylmethylserine allyl ester (**24**) serine

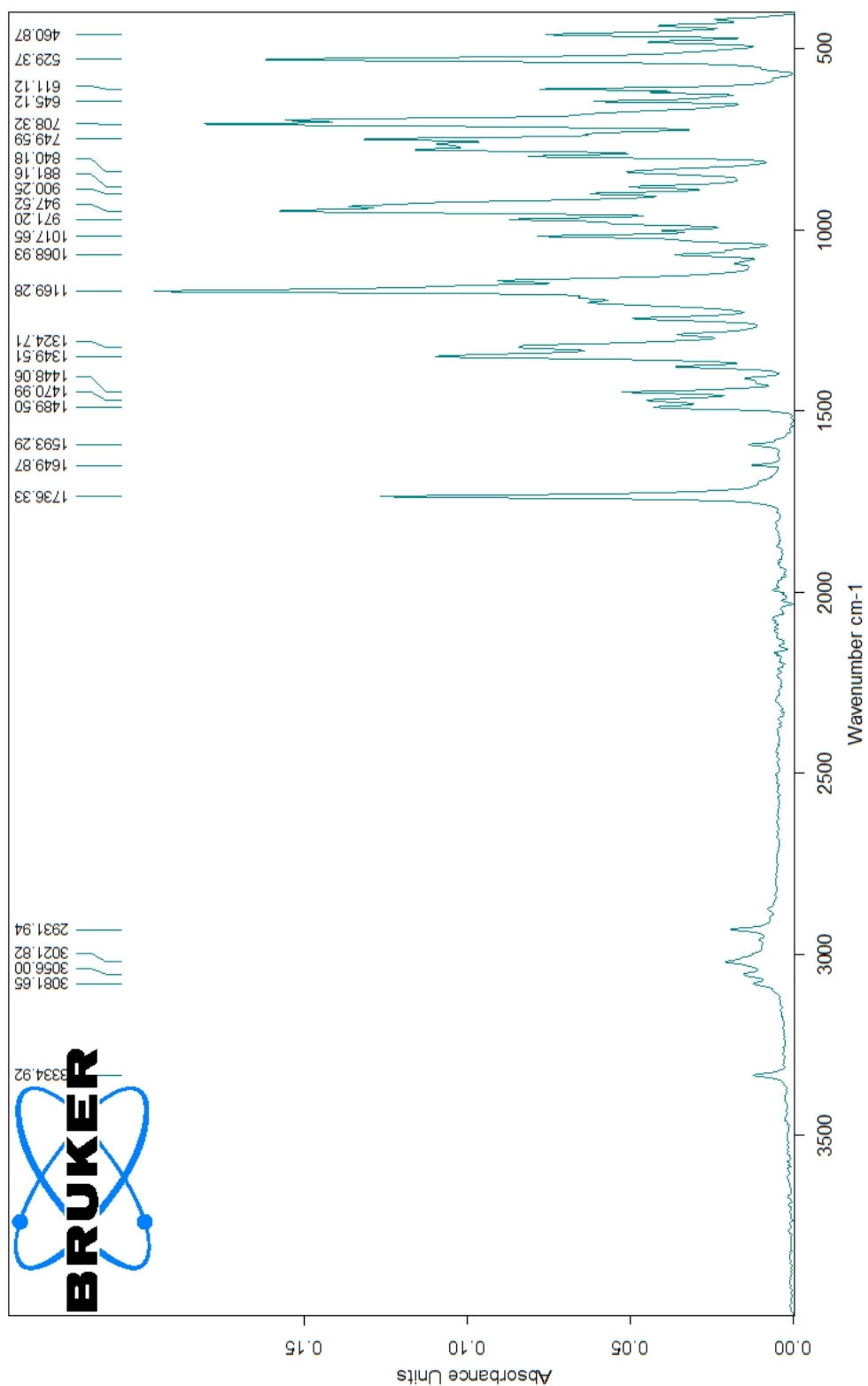


$^1\text{H}$  NMR spectrum of *N*-Triphenylmethyl-(*R*)-serine(O-methanesulfonyl) allyl ester (**25**)

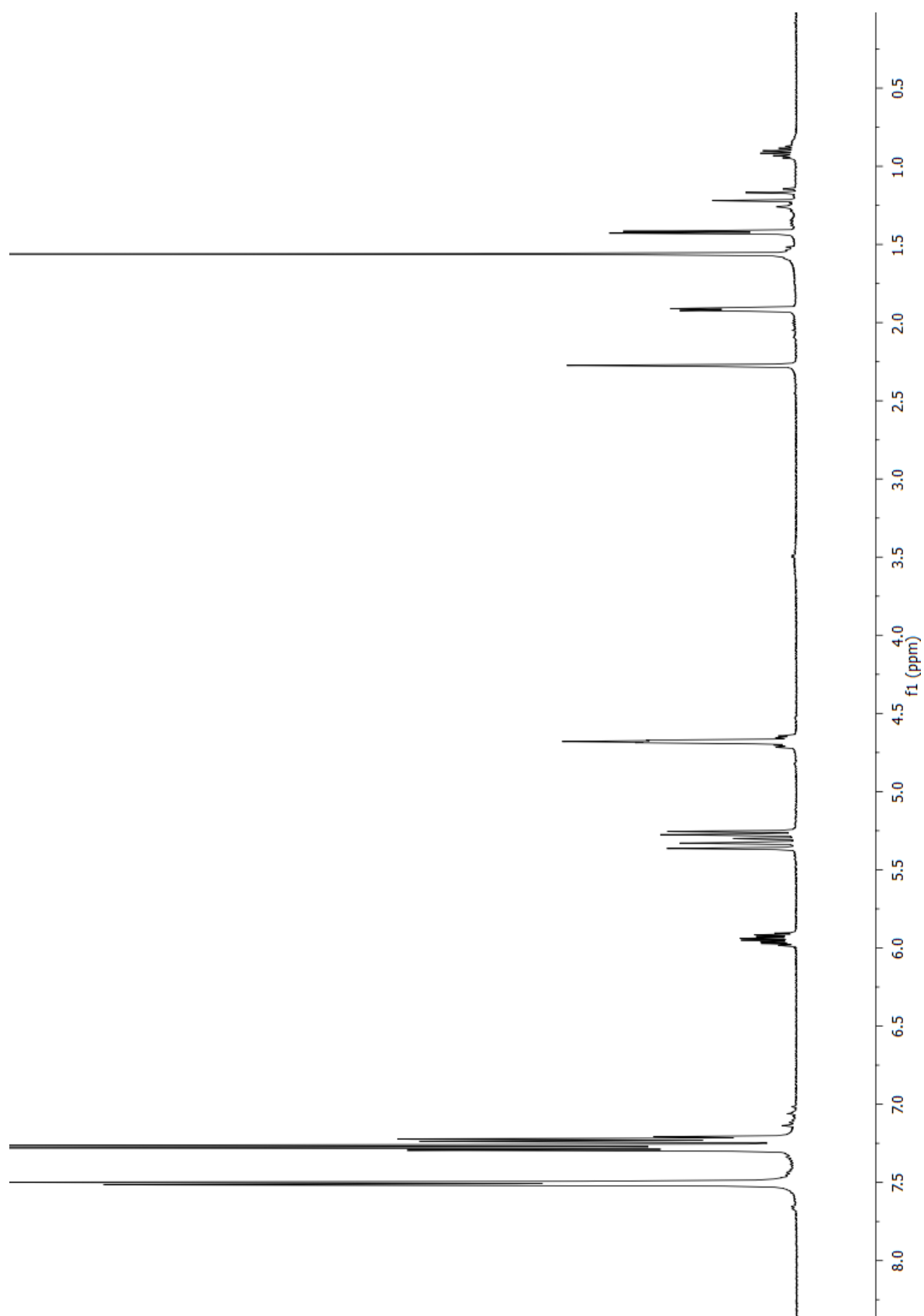
(500 MHz,  $\text{DMSO-d}_6$ )



$^{13}\text{C}$  NMR spectrum of *N*-Triphenylmethyl-(*R*)-serine(*O*-methanesulfonyl) allyl ester (**25**) (125 MHz,  $\text{DMSO-d}_6$ )

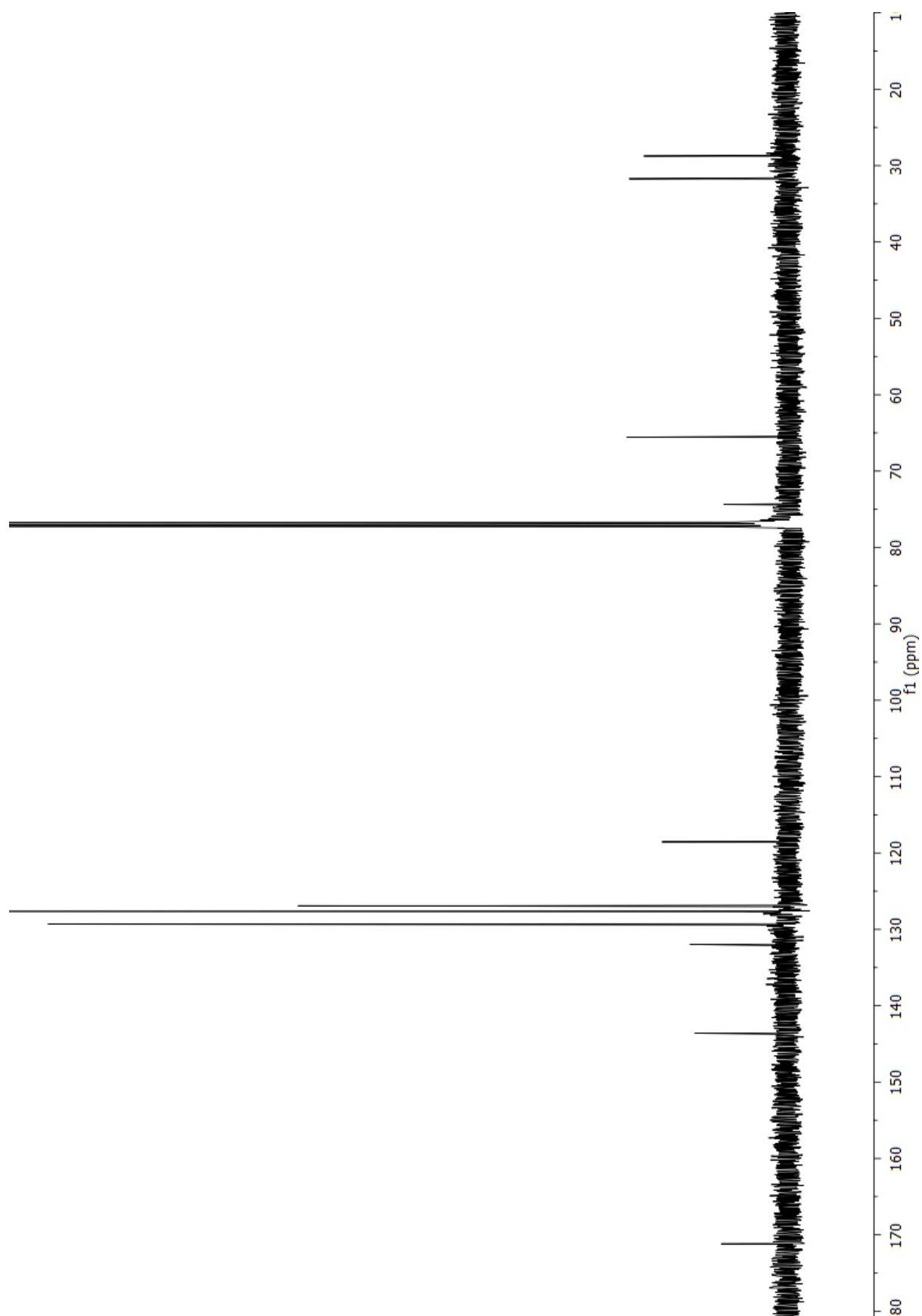


C:\OPUS_7 2.139.1294\MEAS\871	BPB02_60a Mesyl Serine	Instrument type and / or accessory	15/12/2015
-------------------------------	------------------------	------------------------------------	------------



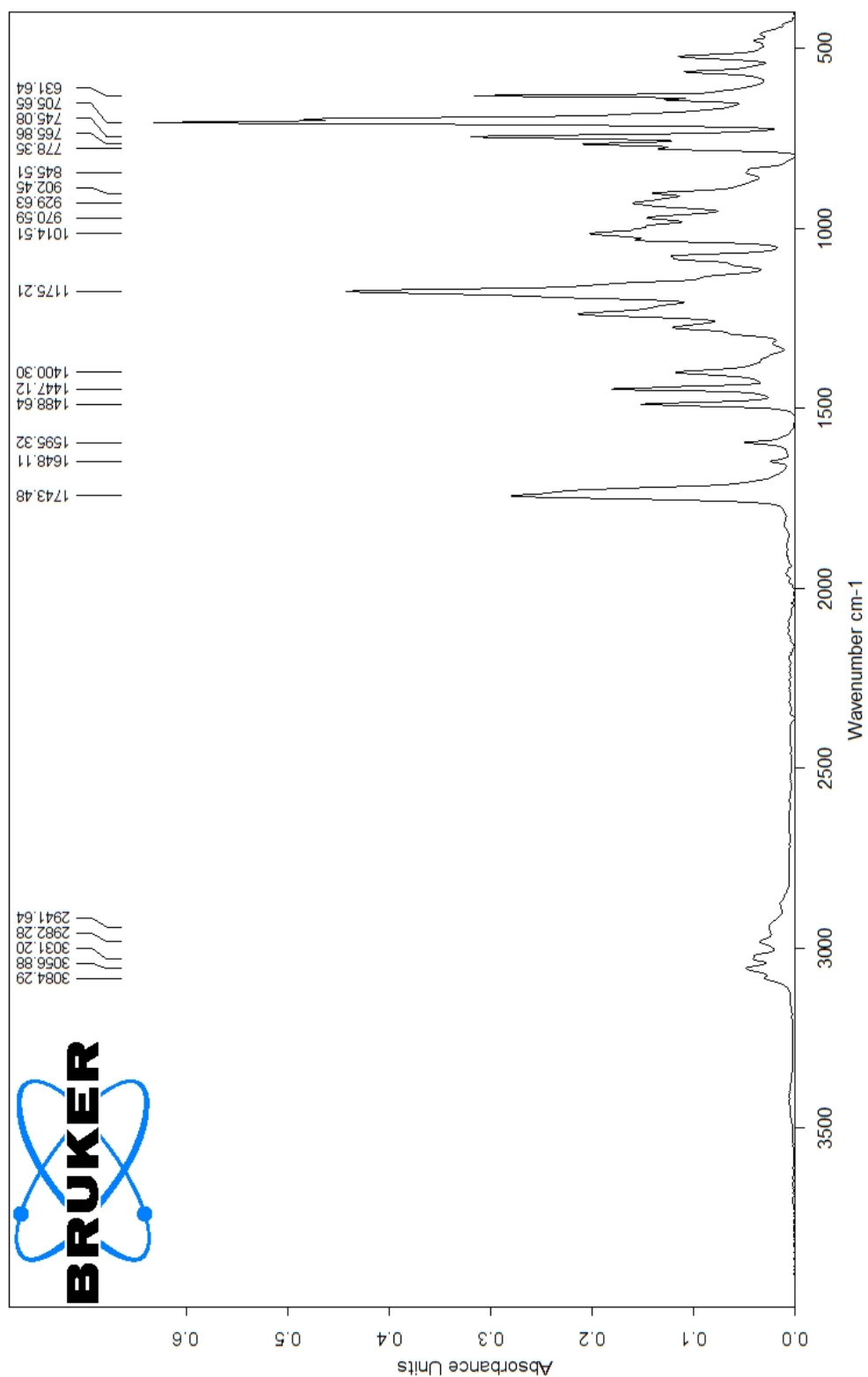
$^1\text{H}$  NMR spectrum of (R)-1-Triphenylmethylaziridine-2-carboxylic acid allyl ester (**26**)

(500 MHz, CDCl<sub>3</sub>)



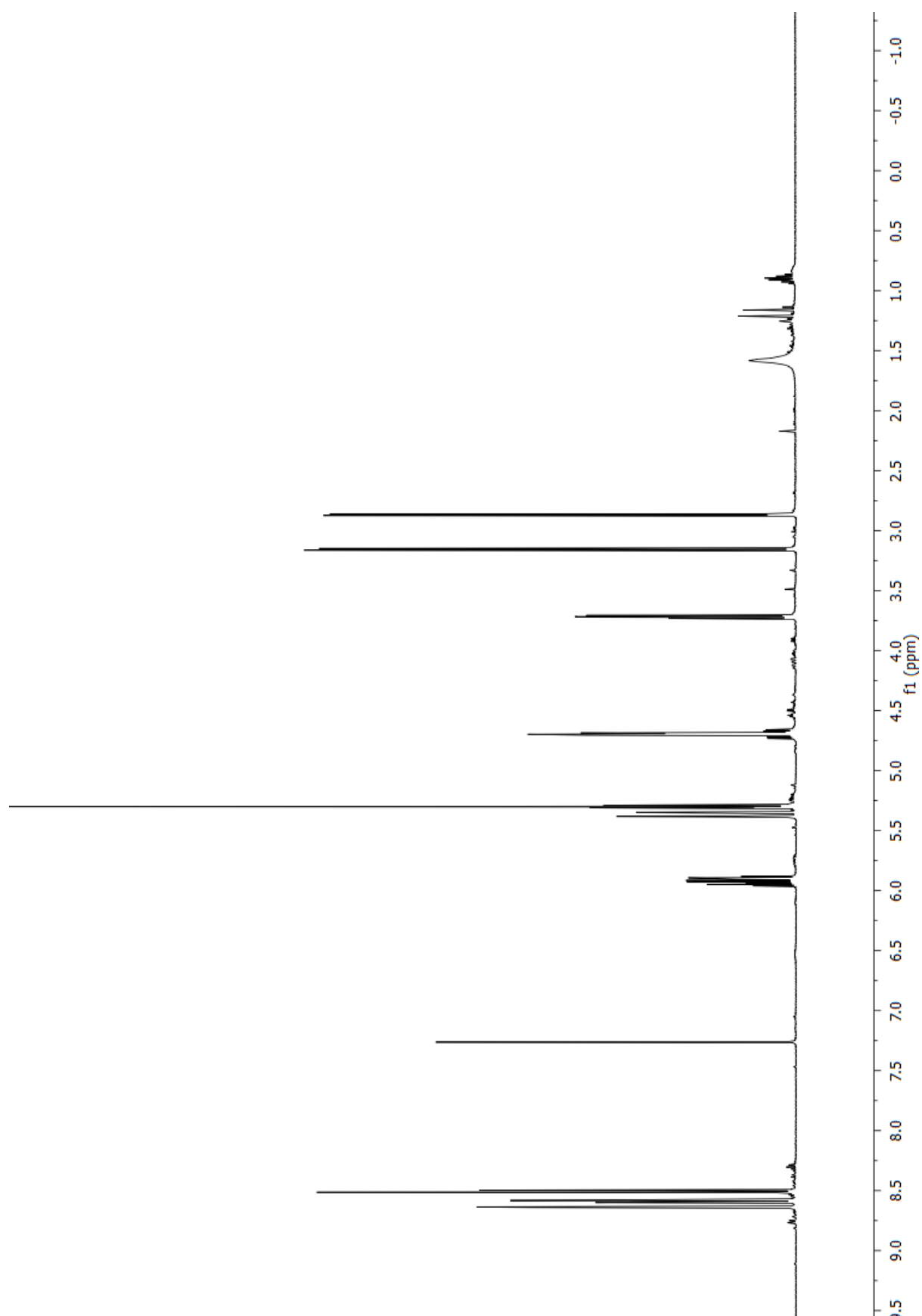
$^{13}\text{C}$  NMR spectrum of (R)-1-Triphenylmethylaziridine-2-carboxylic acid allyl ester (**26**) (125 MHz,  $\text{CDCl}_3$ )



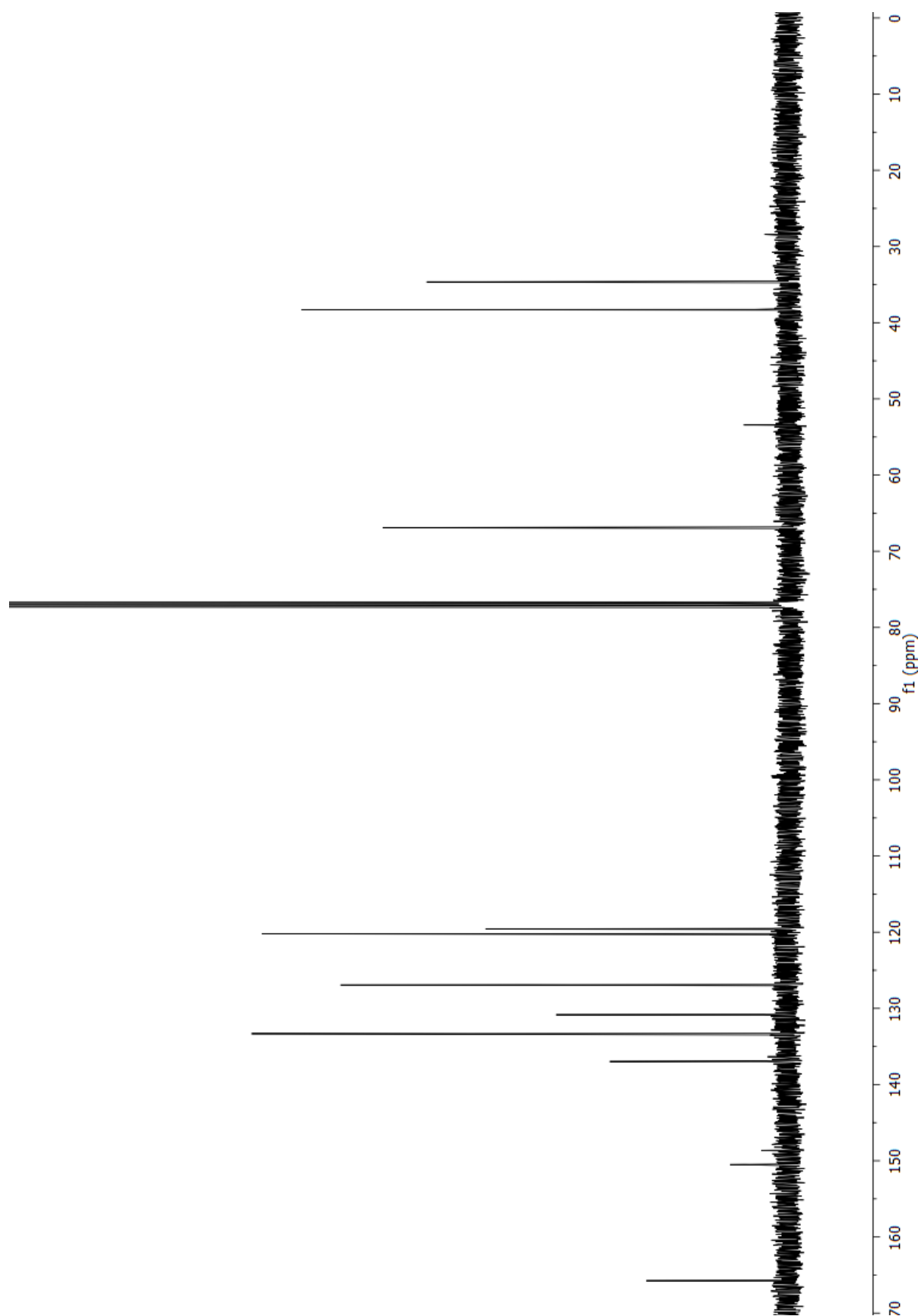


C:\OPUS_7.2.139.1294\MEAS\864	BPB02_65a Serine aziridine first	Instrument type and / or accessory	15/12/2015
-------------------------------	----------------------------------	------------------------------------	------------

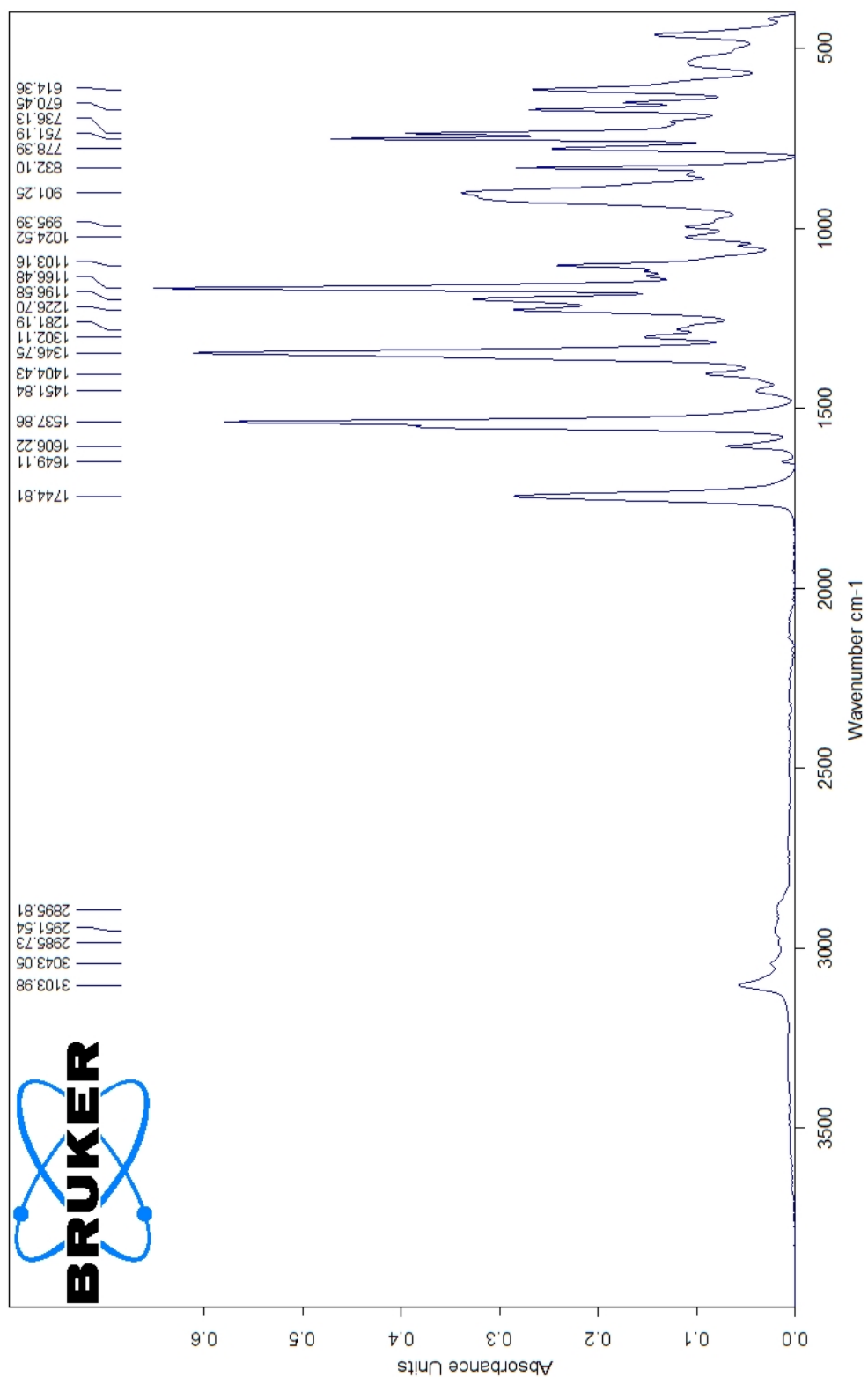
IR spectrum of (*R*)-1-Triphenylmethylaziridine-2-carboxylic acid allyl ester (**26**)



$^1\text{H}$  NMR spectrum of Allyl-(*R*)-1-(2,4-dinitrobenzenesulfonyl)aziridine-2-carboxylate (**27**) (500 MHz,  $\text{CDCl}_3$ )

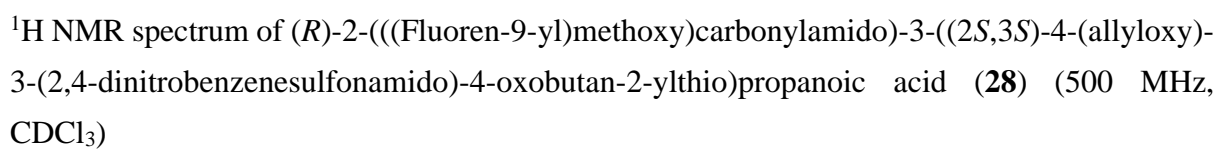


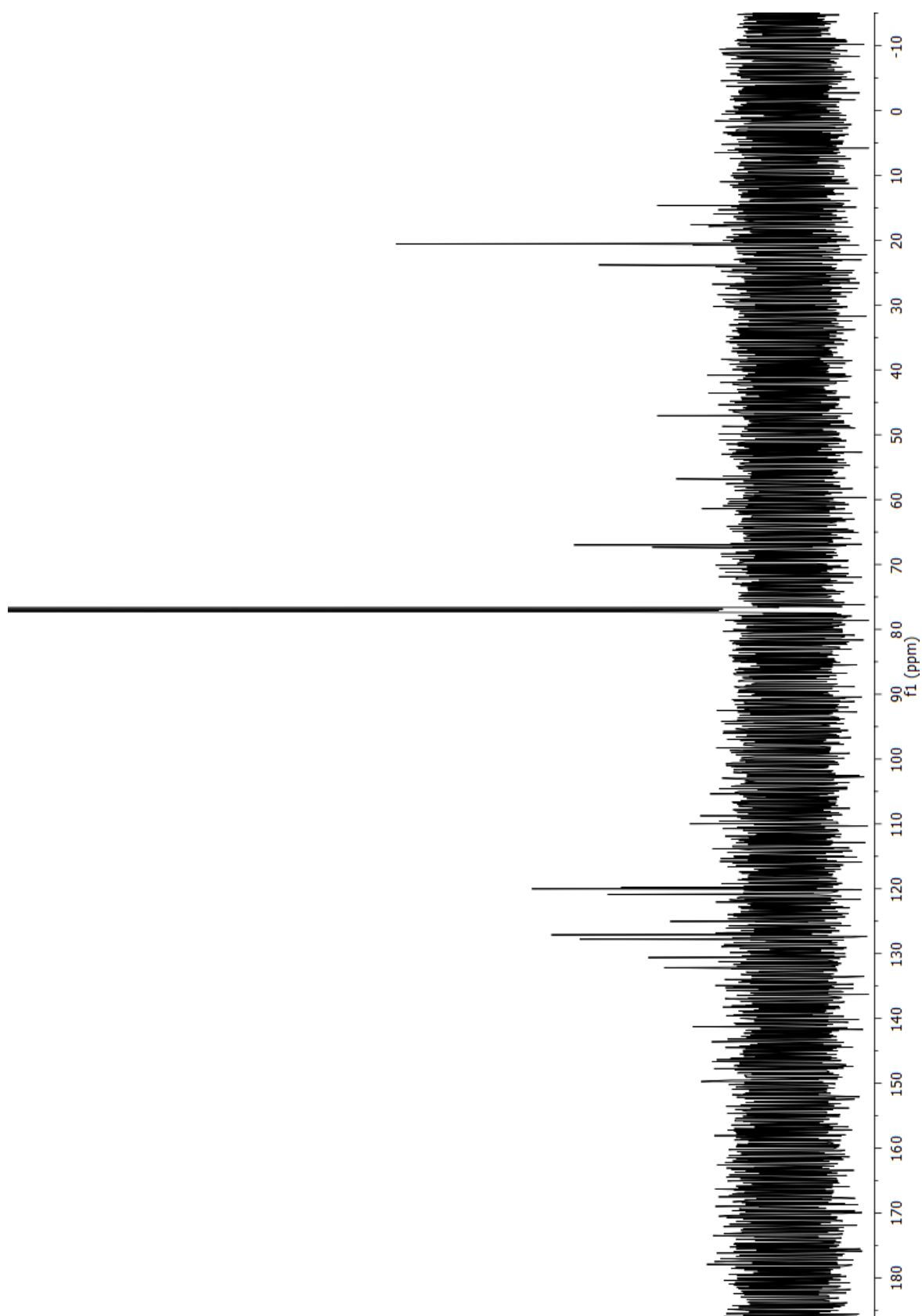
$^{13}\text{C}$  spectrum of Allyl-(*R*)-1-(2,4-dinitrobenzenesulfonyl)aziridine-2-carboxylate (**27**) (125 MHz,  $\text{CDCl}_3$ )



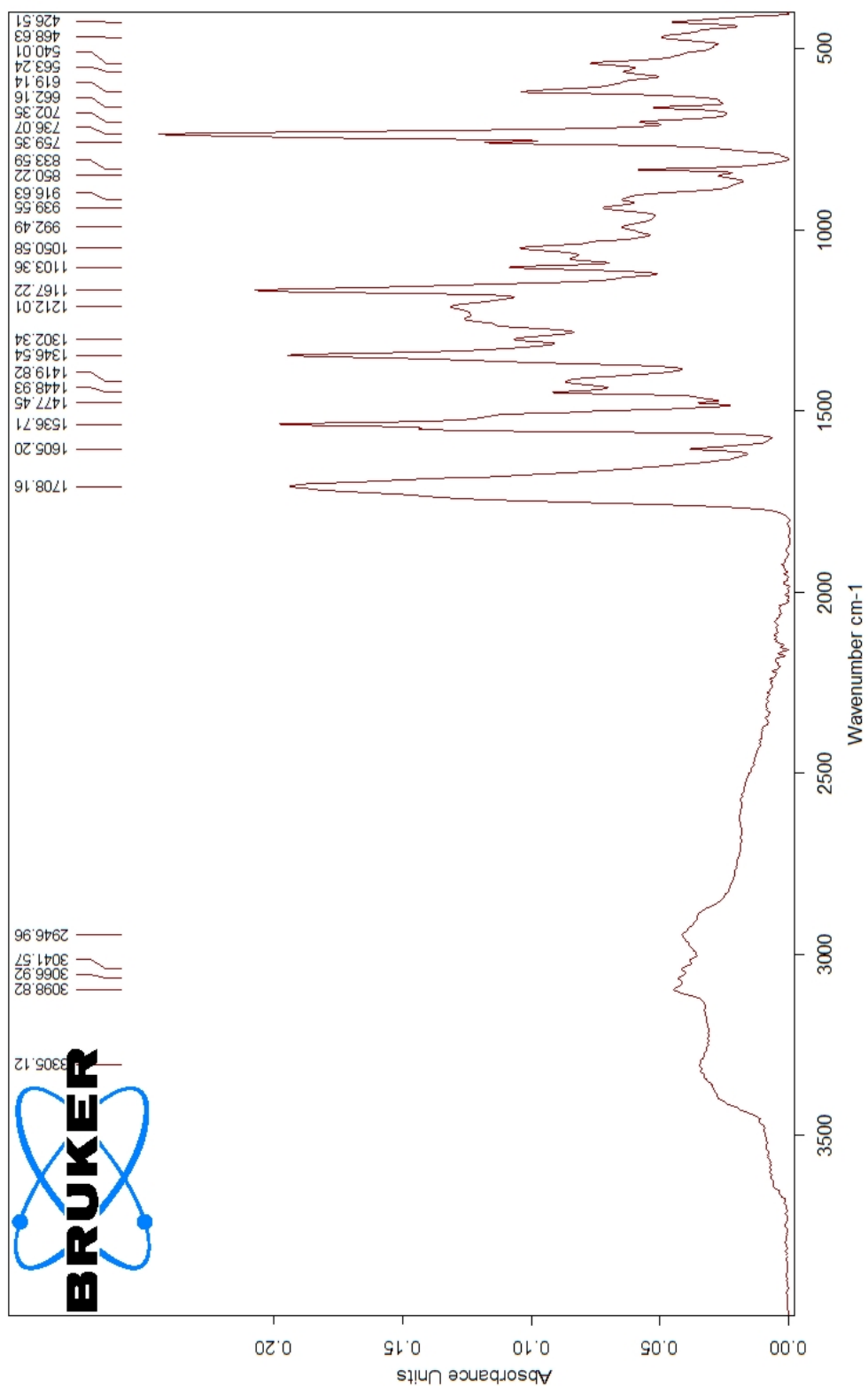
C:\OPUS_7 2.139.1294\MEAS\867	BPB02_85C Second aziridine Serine	Instrument type and / or accessory	15/12/2015
-------------------------------	-----------------------------------	------------------------------------	------------

IR spectrum of Allyl-(*R*)-1-(2,4-dinitrobenzenesulfonyl)aziridine-2-carboxylate (**27**)



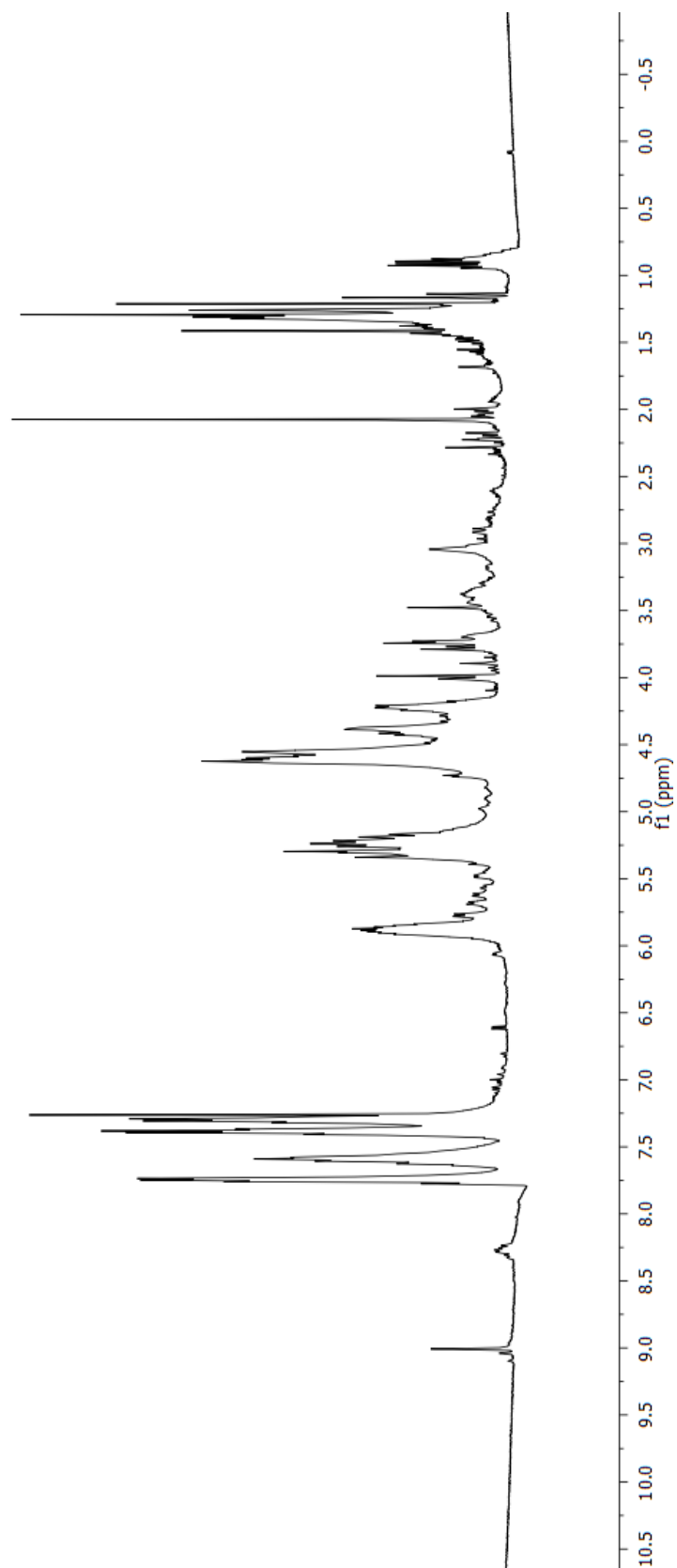


$^{13}\text{C}$  NMR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamido)-3-((2*S*,3*S*)-4-(allyloxy)-3-(2,4-dinitrobenzenesulfonamido)-4-oxobutan-2-ylthio)propanoic acid (**28**) (125 MHz,  $\text{CDCl}_3$ )



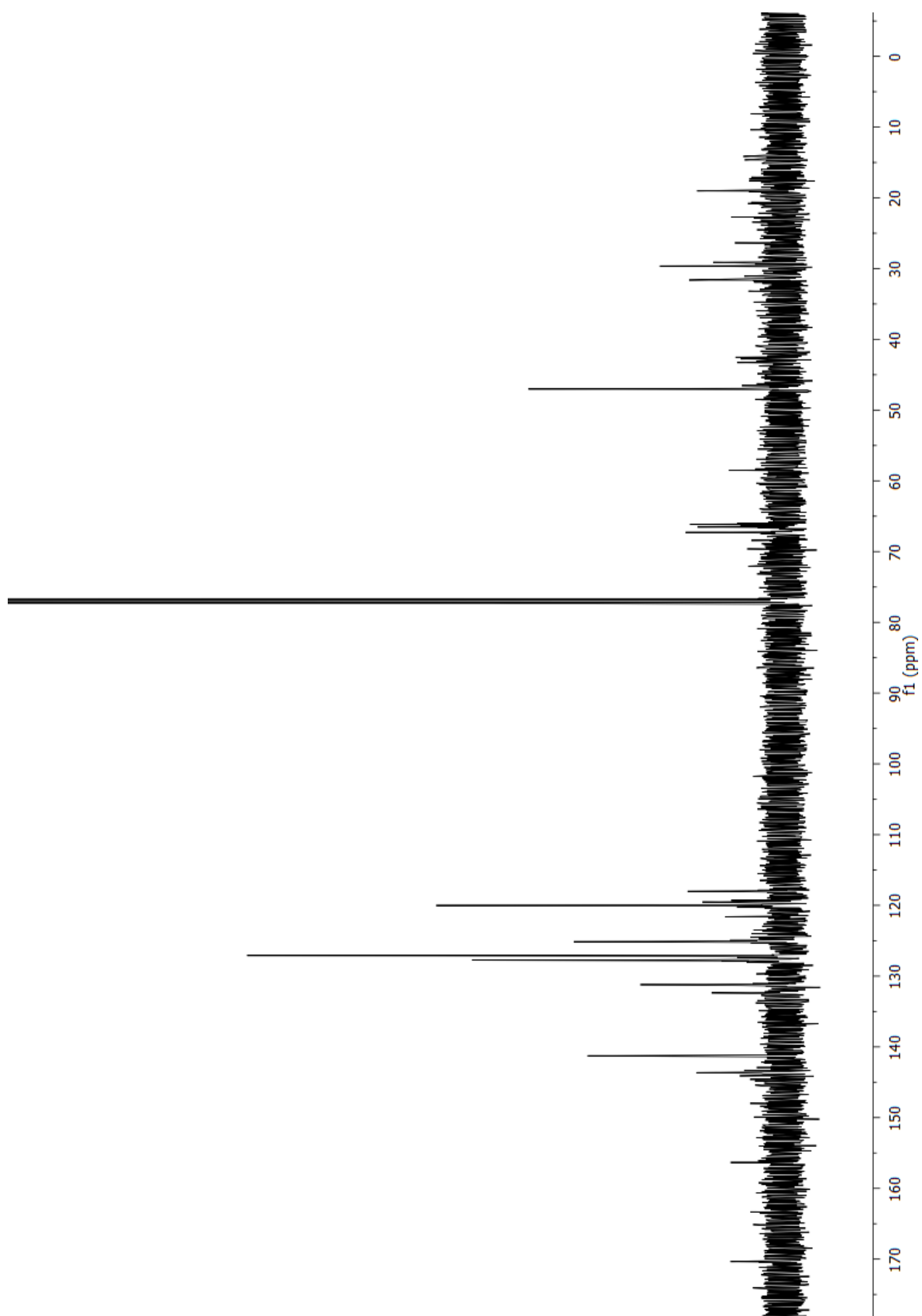
C:\OPUS_7.2.139.1294\MEAS\868	BPB02_74C Lanthionine DNS	Instrument type and / or accessory	15/12/2015
-------------------------------	---------------------------	------------------------------------	------------

IR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamido)-3-((2*S*,3*S*)-4-(allyloxy)-3-(2,4-dinitrobenzenesulfonamido)-4-oxobutan-2-ylthio)propanoic acid (**28**)

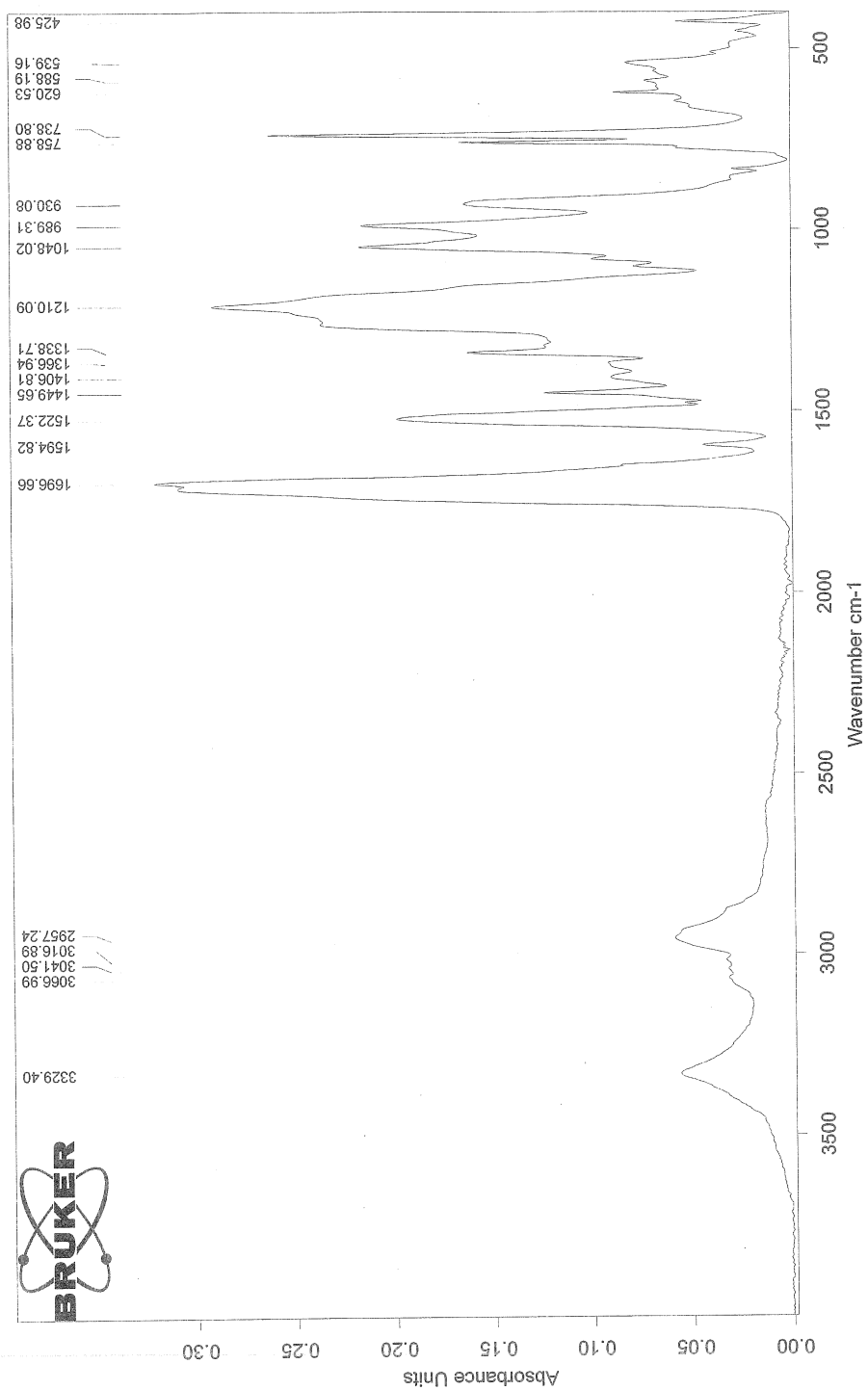


$^1\text{H}$  NMR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2*S*,3*S*)-4-(allyloxy)-3-(allyloxycarbonylamino)-4-oxobutan-2-ylthio)propanoic acid (**29**) (500 MHz,  $\text{CDCl}_3$ )





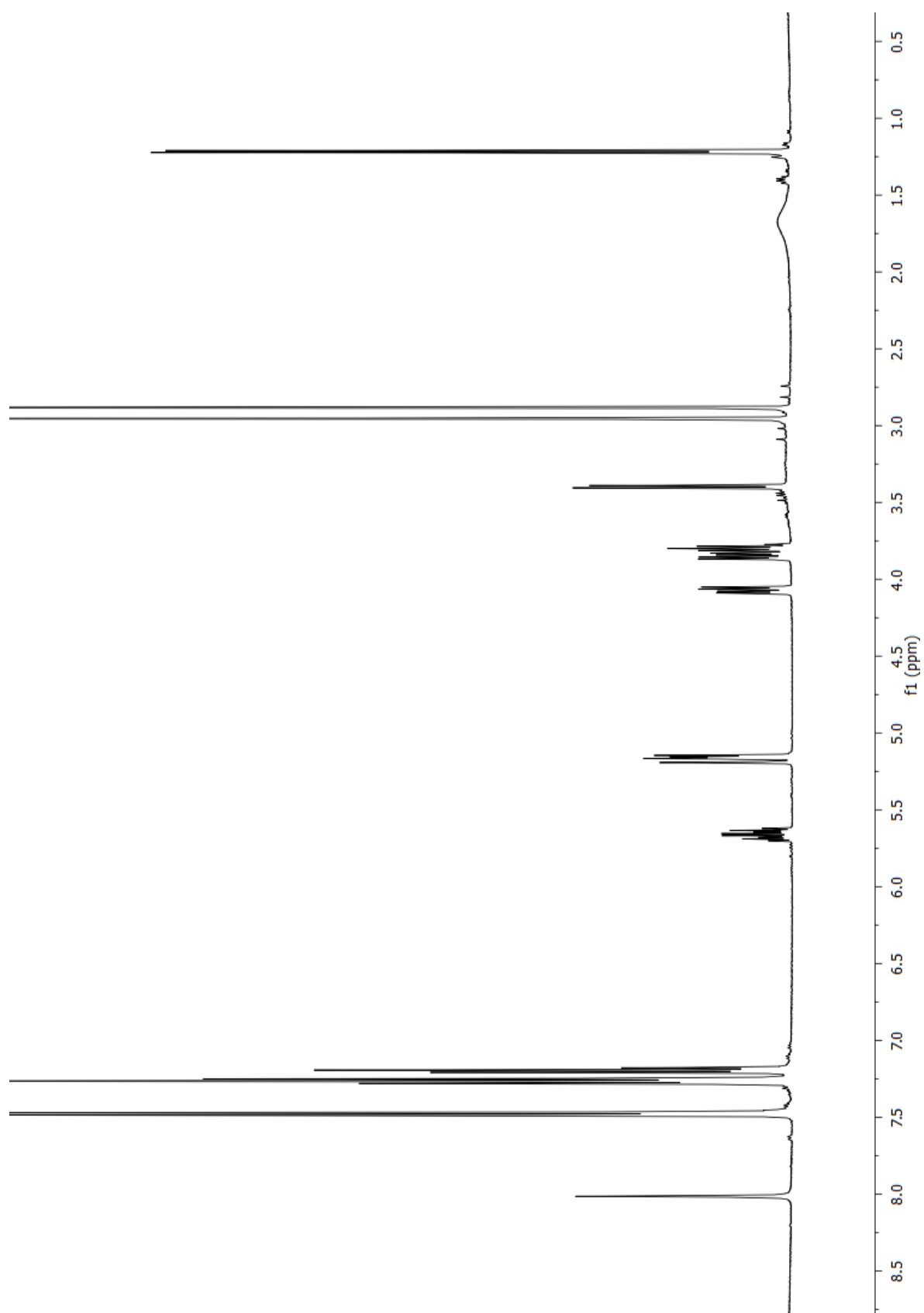
$^{13}\text{C}$  NMR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2*S*,3*S*)-4-(allyloxy)-3-(allyloxycarbonylamino)-4-oxobutan-2-ylthio)propanoic acid (**29**) (125 MHz,  $\text{CDCl}_3$ )



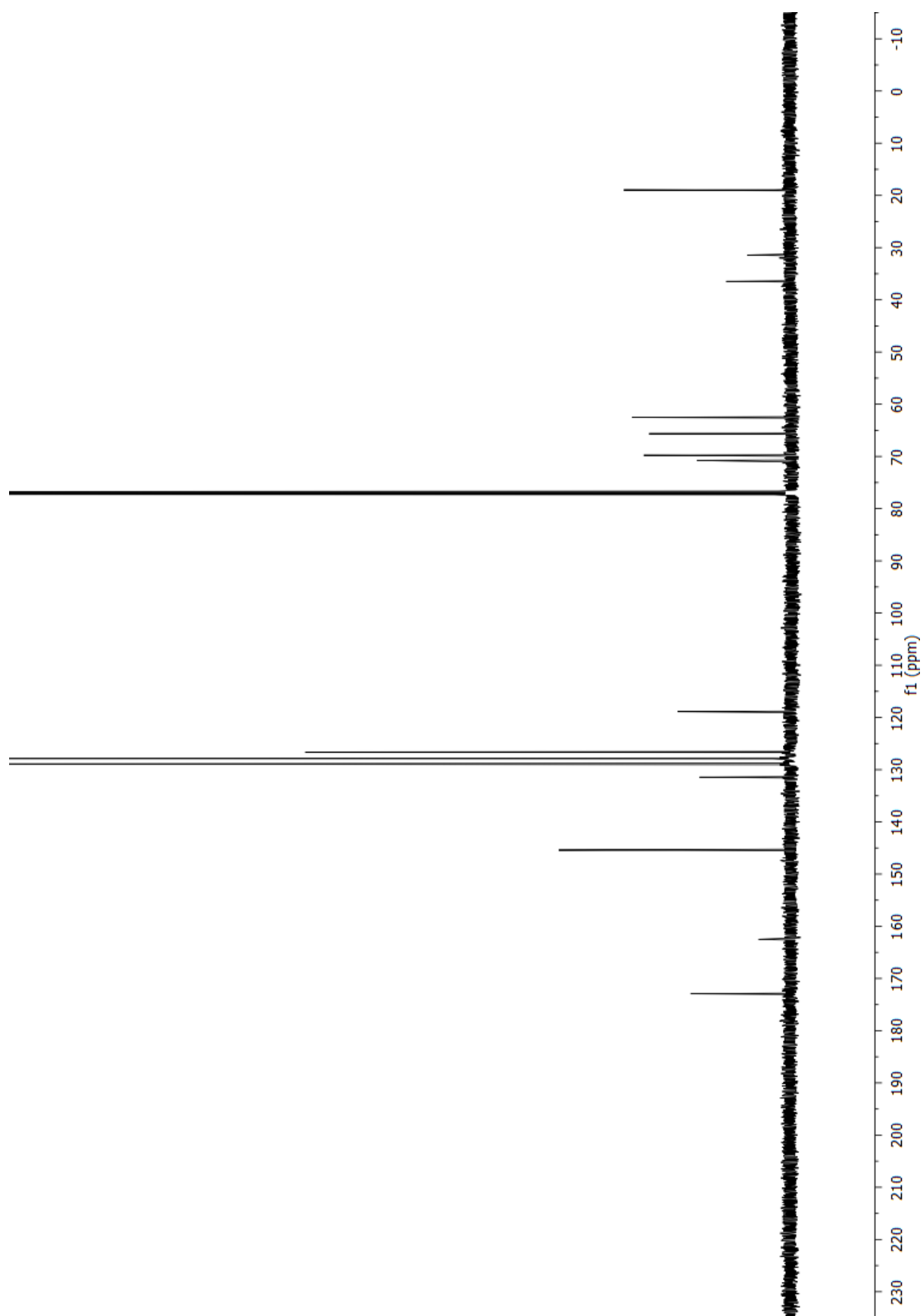
C:\OPUS_7.2.139.1294\MEAS\402	BPB02_80e	Instrument type and / or accessory	7/10/2015
-------------------------------	-----------	------------------------------------	-----------

IR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2*S*,3*S*)-4-(allyloxy)-3-(allyloxycarbonylamino)-4-oxobutan-2-ylthio)propanoic acid (**29**)

## **Appendix D – Spectra of MeLan intermediates**

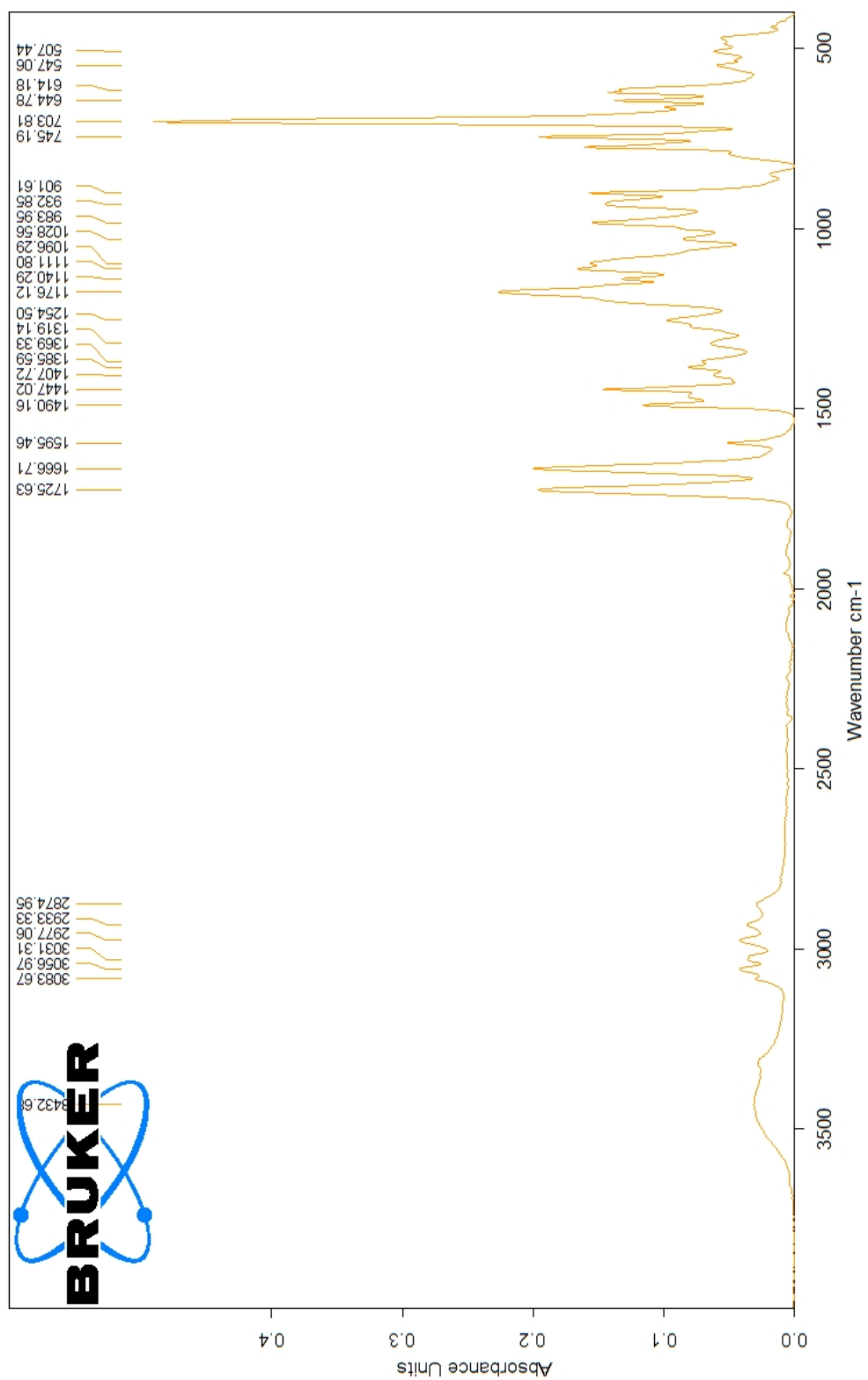


$^1\text{H}$  NMR spectrum of (2*R*)-*N*-Triphenylmethyl-(*S*)-threonine allyl ester (**30**) (500 MHz,  $\text{CDCl}_3$ )

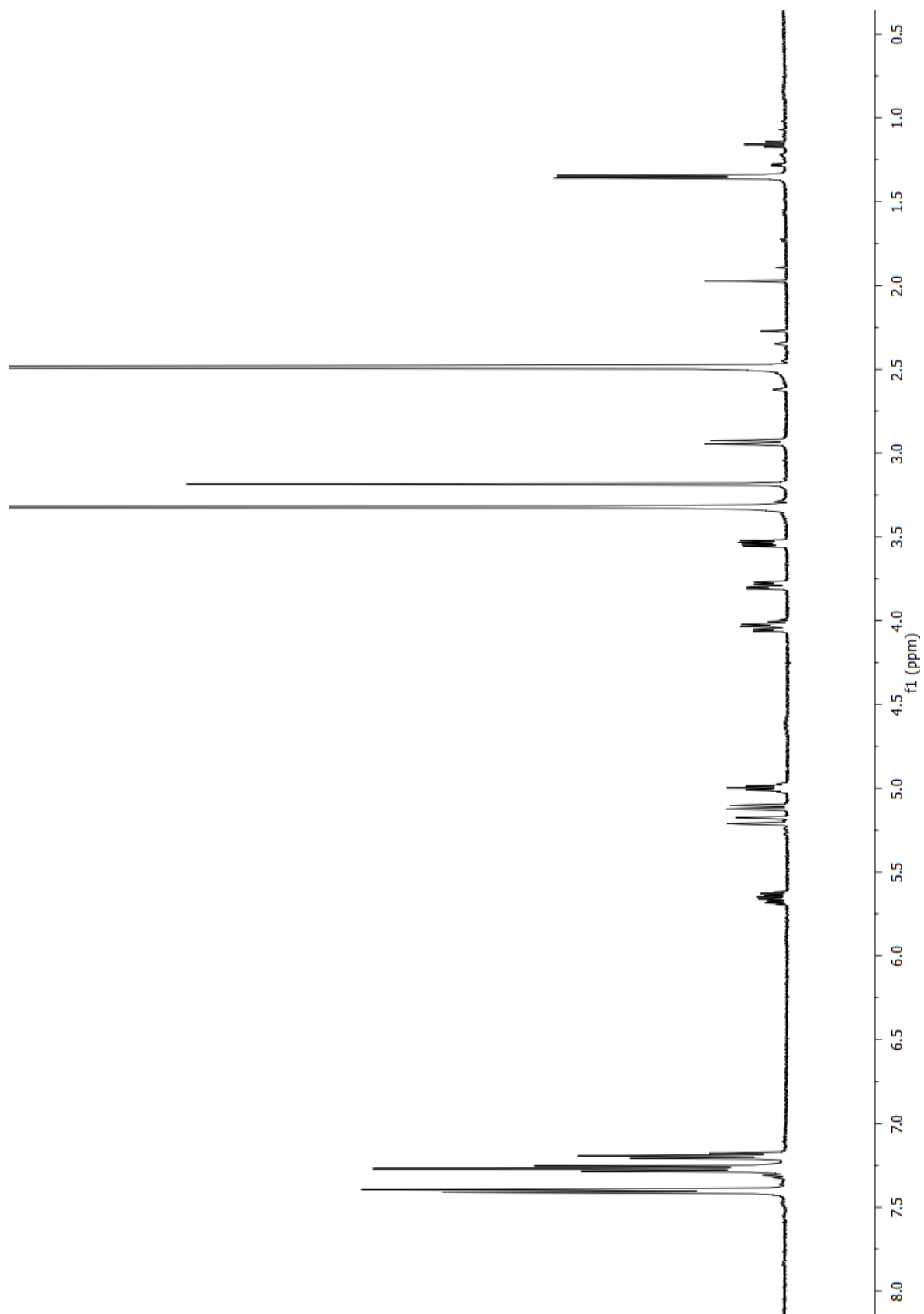


$^{13}\text{C}$  spectrum  $(2R)$ - $N$ -Triphenylmethyl- $(S)$ -threonine allyl ester (**30**) (125 MHz, CDCl<sub>3</sub>)

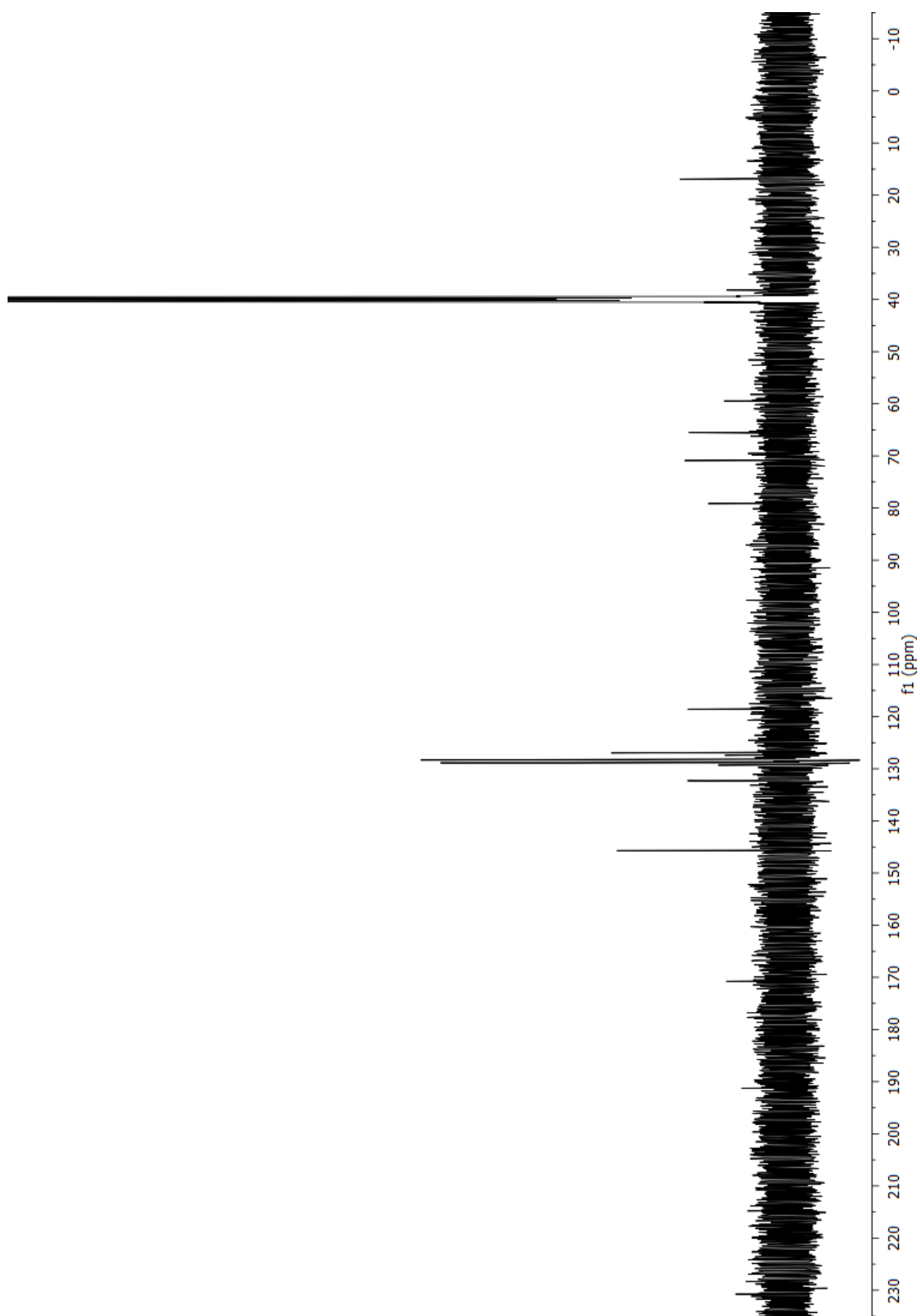
IR spectrum of (2*R*)-*N*-Triphenylmethyl-(*S*)-threonine allyl ester (**30**)



C:\OPUS_7 2.139.1294\MEAS\872	BPB02_55b Allylated threonine	Instrument type and / or accessory	15/12/2015
-------------------------------	-------------------------------	------------------------------------	------------

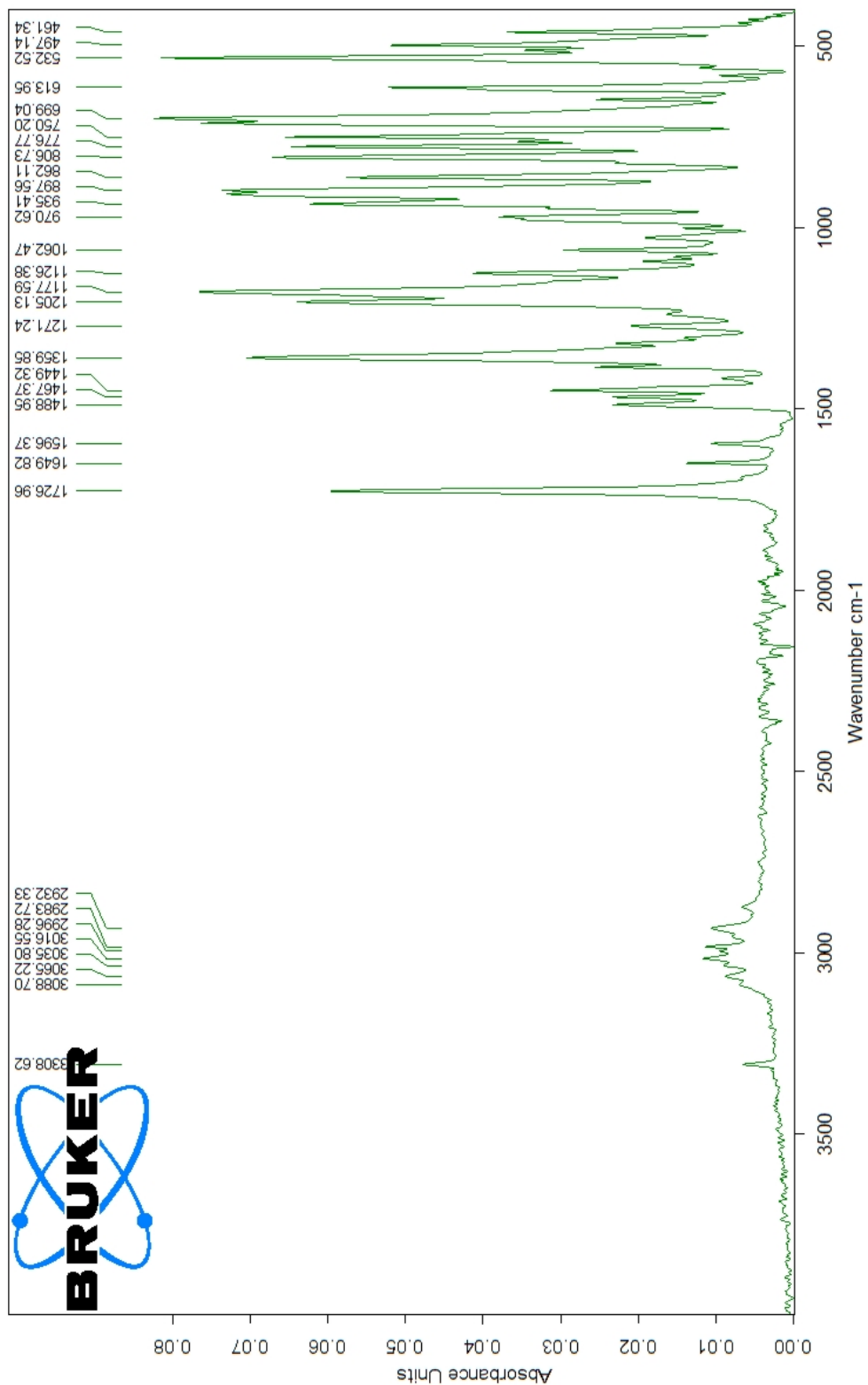


$^1\text{H}$  NMR spectrum of *N*-Triphenylmethyl-*(RS)*-threonine-*(O*-methanesulfonyl) allyl ester (**31**) (500 MHz, DMSO- $\text{d}_6$ )



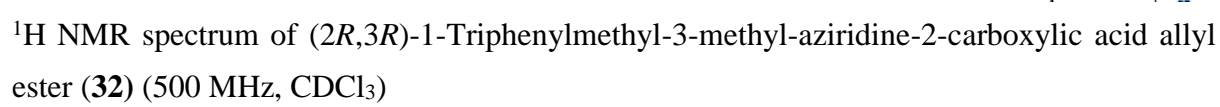
$^{13}\text{C}$  NMR spectrum of *N*-Triphenylmethyl-(*RS*)-threonine-(*O*-methanesulfonyl) allyl ester (**31**) (125 MHz,  $\text{DMSO-d}_6$ )

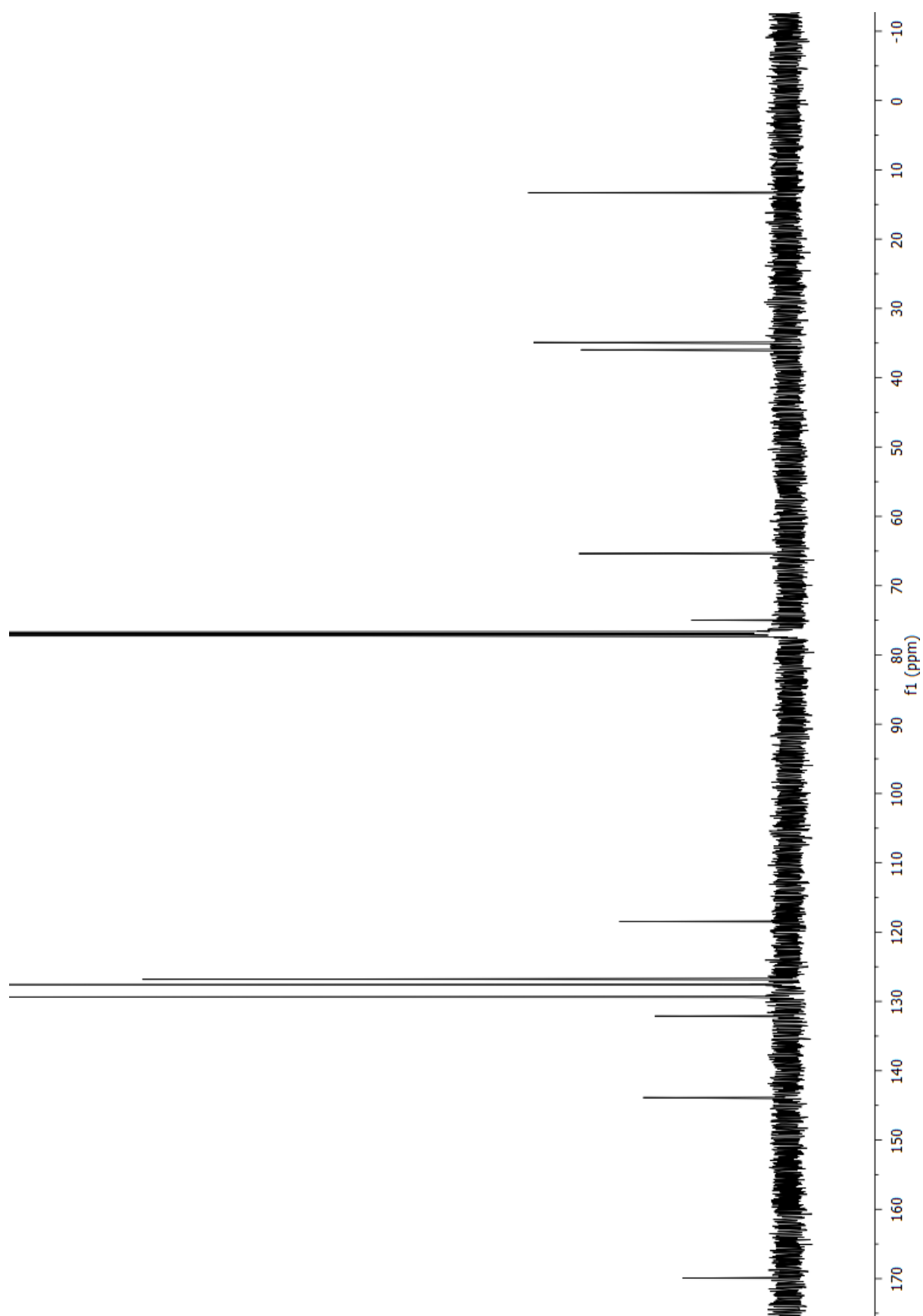




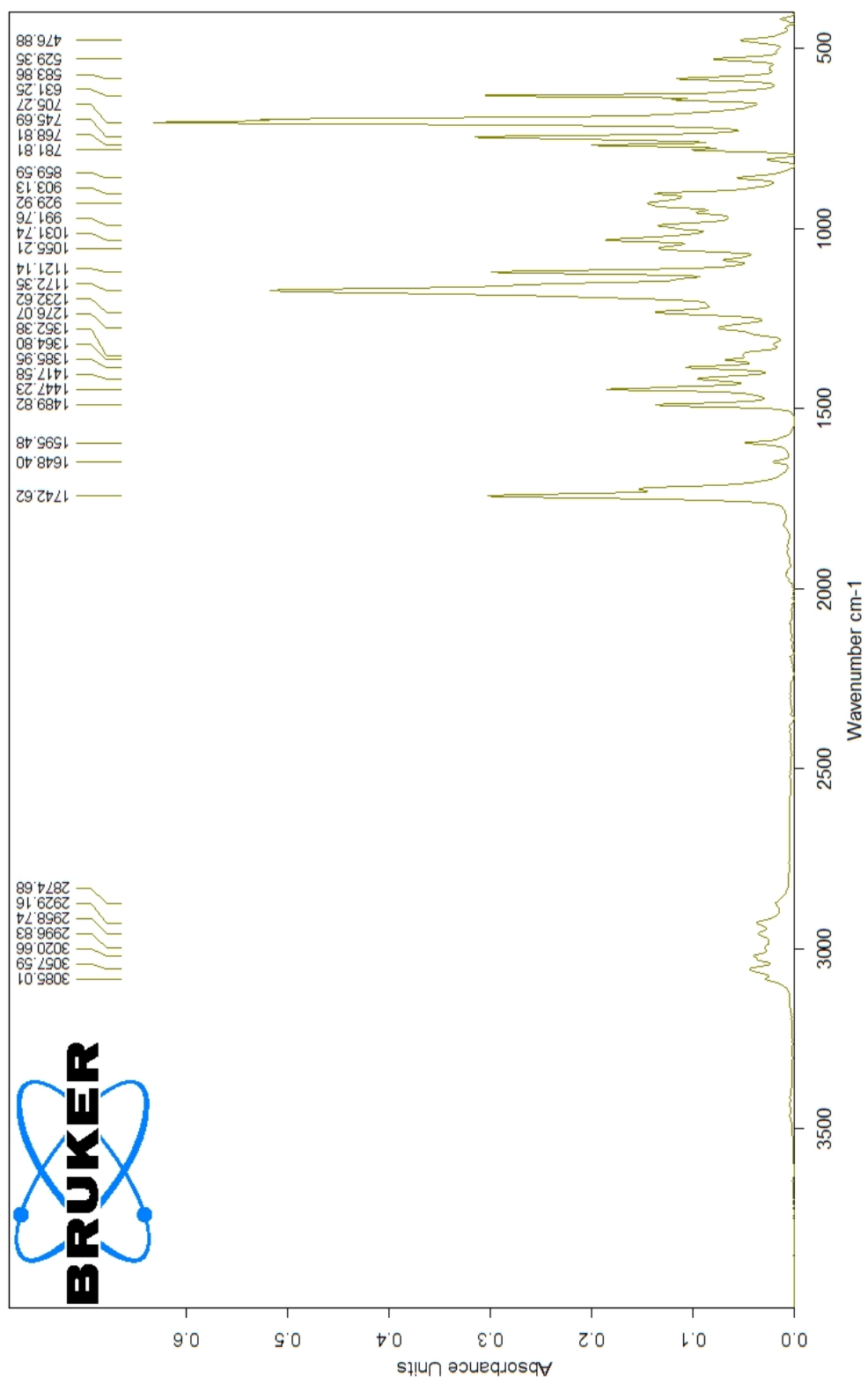
C:\OPUS_7.2.139.1294\MEAS\870	BPB02_59b Mesyl Threonine	Instrument type and / or accessory	15/12/2015
-------------------------------	---------------------------	------------------------------------	------------

IR spectrum of *N*-Triphenylmethyl-(*RS*)-threonine-(*O*-methanesulfonyl) allyl ester (**31**)



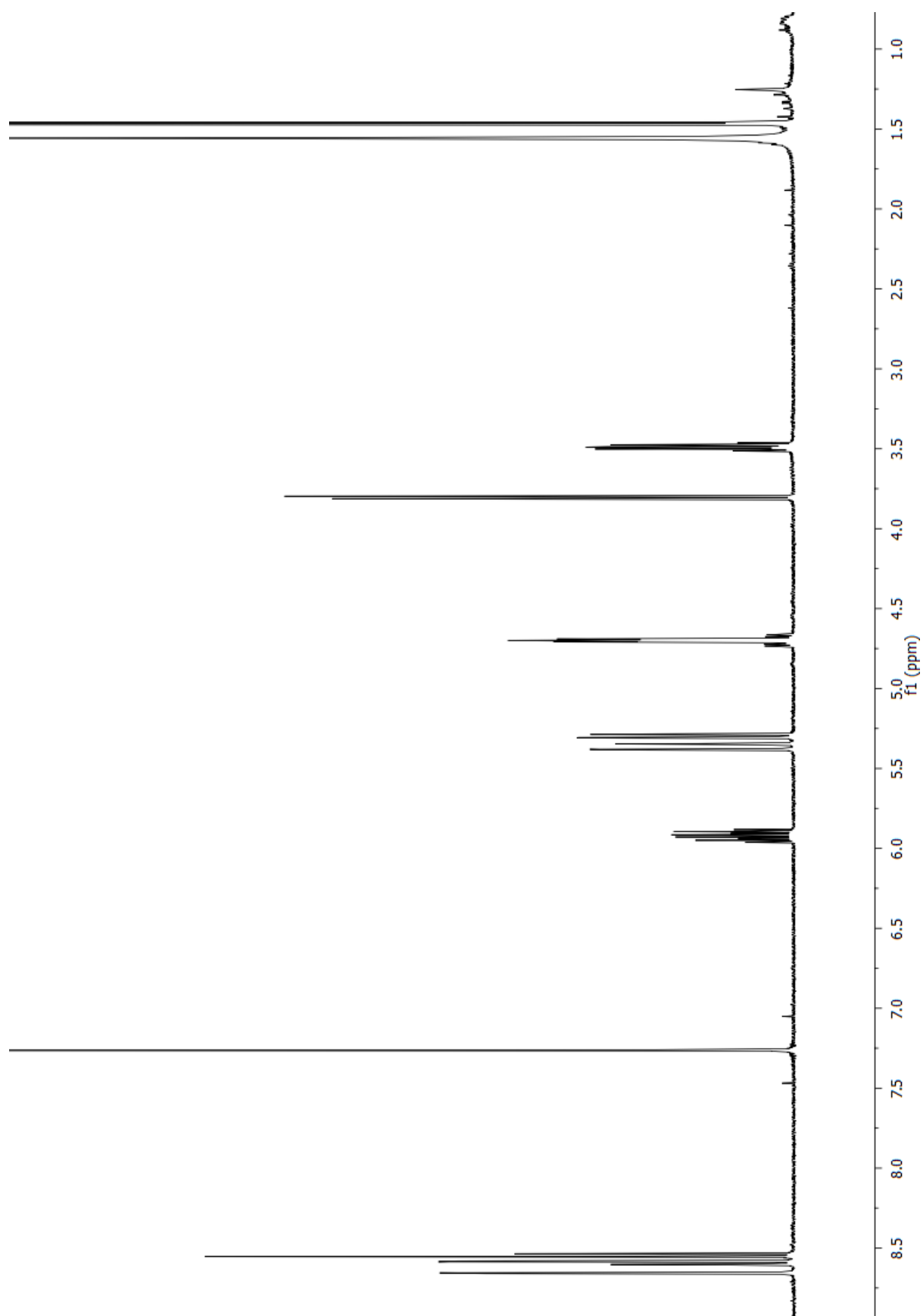


$^{13}\text{C}$  NMR spectrum of (2*R*,3*R*)-1-Triphenylmethyl-3-methyl-aziridine-2-carboxylic acid allyl ester (**32**) (125 MHz,  $\text{CDCl}_3$ )

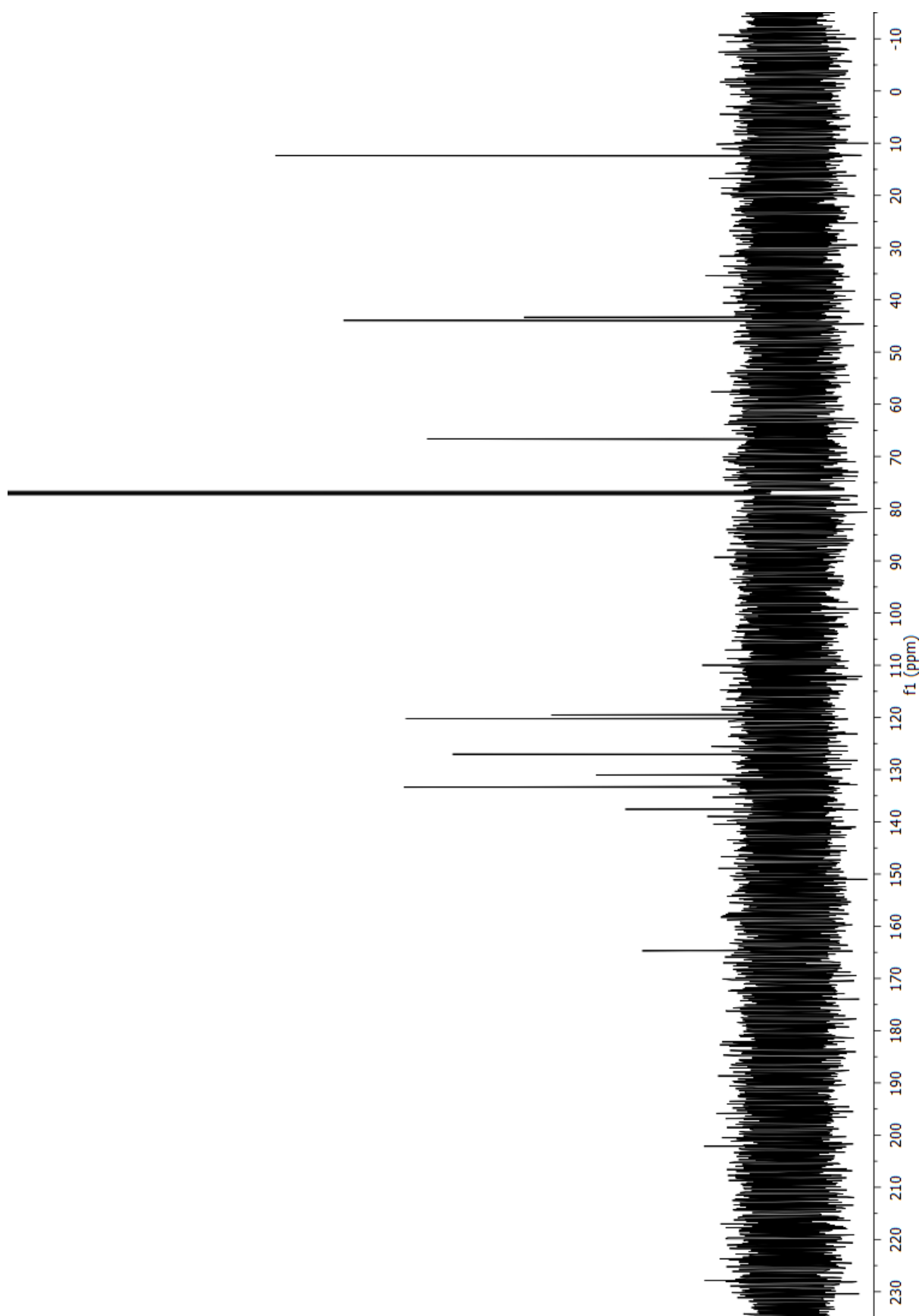


C:\OPUS_7 2.139.1294\MEAS\865	BPB02_63a Threonine aziridine first	Instrument type and / or accessory	15/12/2015
-------------------------------	-------------------------------------	------------------------------------	------------

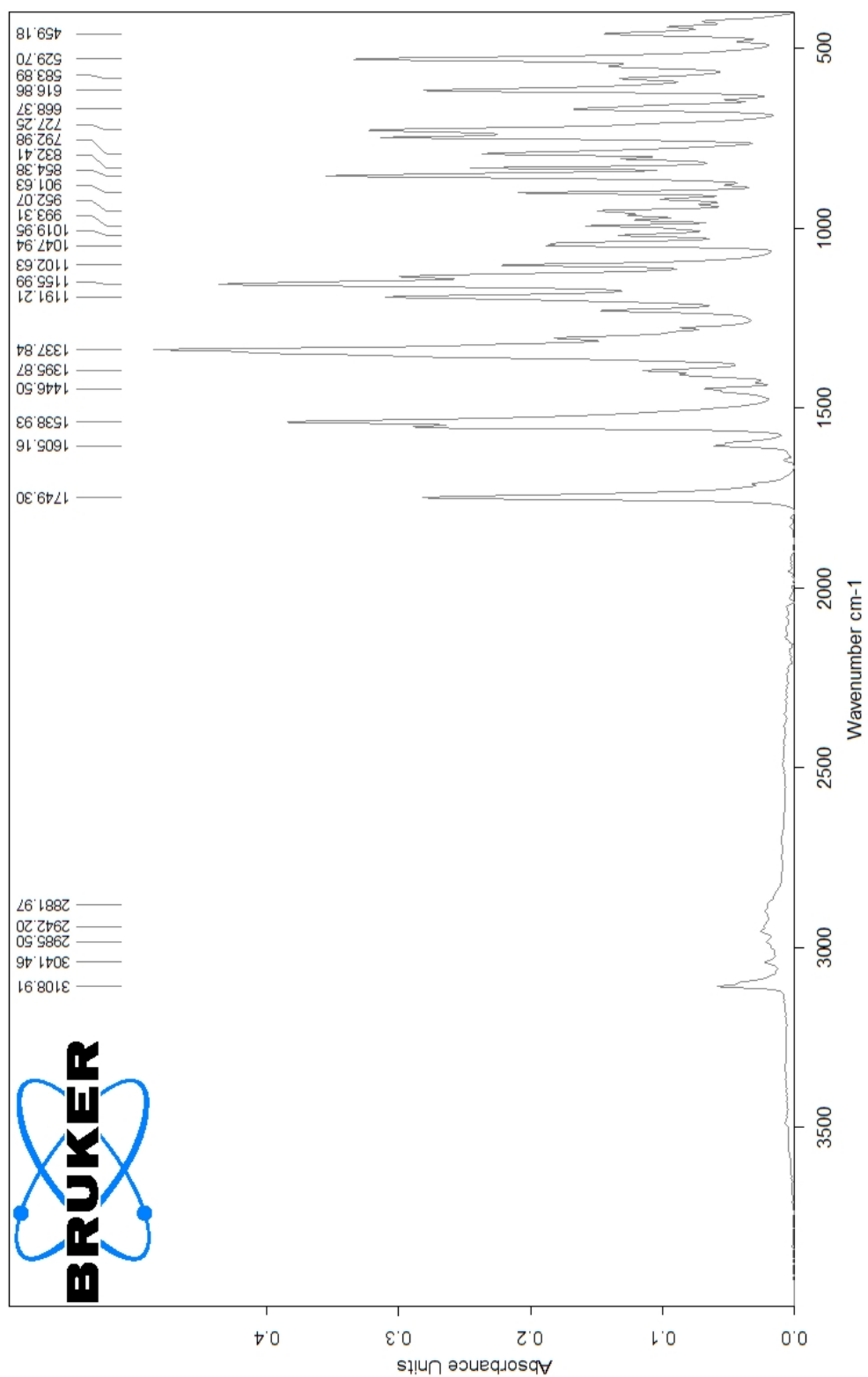
IR spectrum of (2*R*,3*R*)-1-Triphenylmethyl-3-methyl-aziridine-2-carboxylic acid allyl ester  
(32)



$^1\text{H}$  NMR spectrum of (2*R*,3*R*)-Allyl 1-(2,4-dinitrobenzenesulfonyl)-3-methylaziridine-2-carboxylate (**33**) (500 MHz,  $\text{CDCl}_3$ )

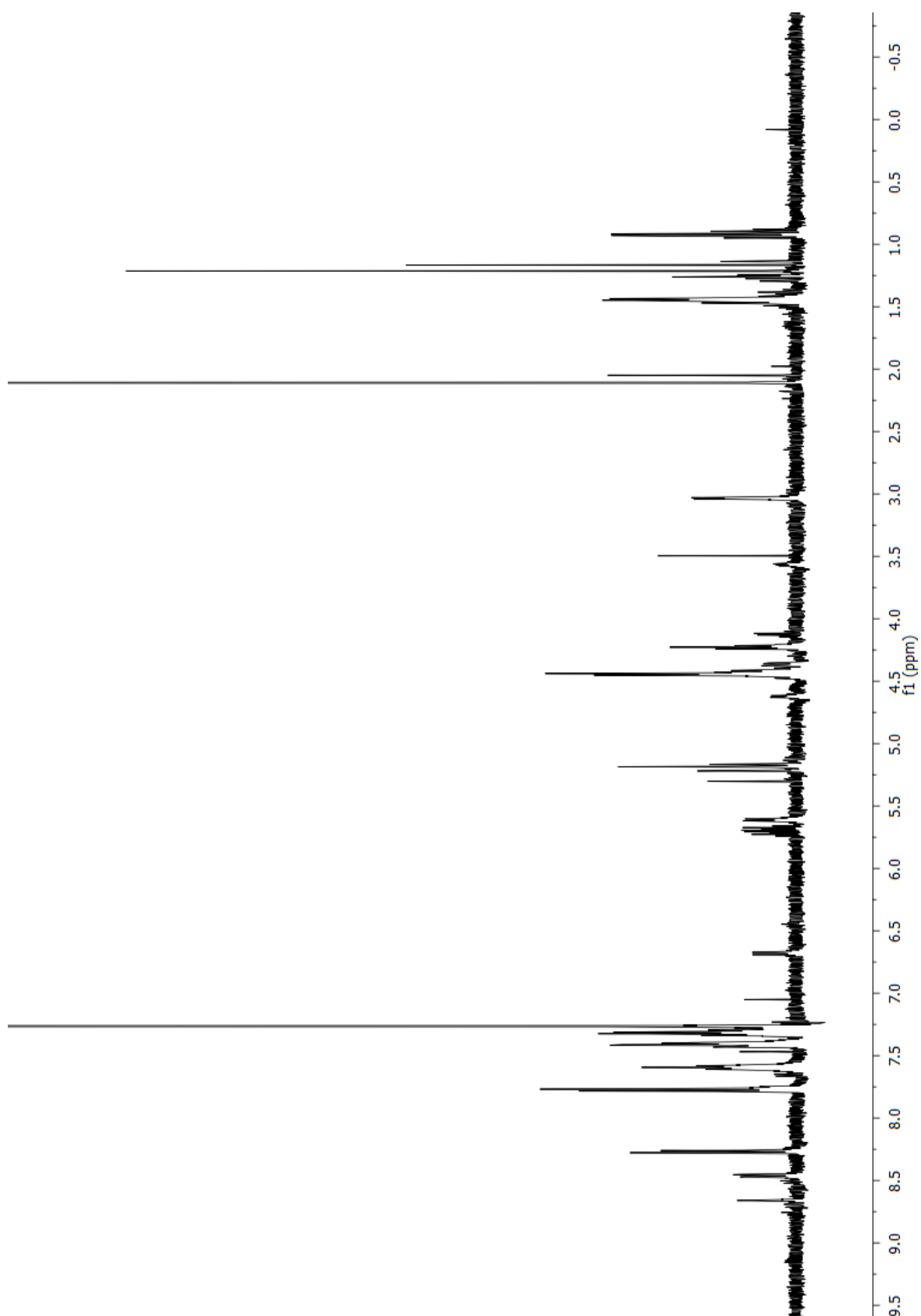


$^{13}\text{C}$  spectrum of (2*R*,3*R*)-Allyl 1-(2,4-dinitrobenzenesulfonyl)-3-methylaziridine-2-carboxylate (**33**) (125 MHz,  $\text{CDCl}_3$ )



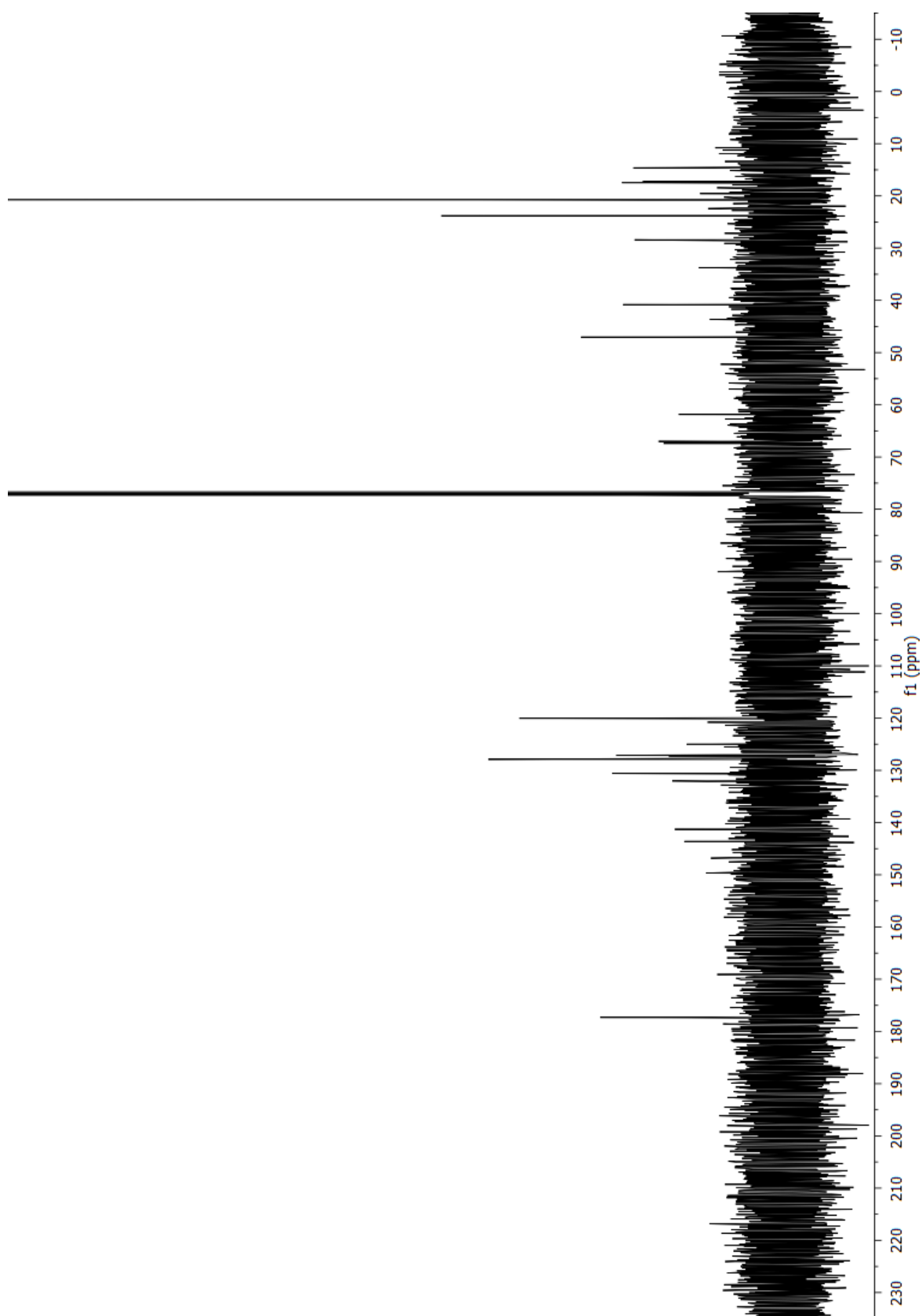
C:\OPUS_7 2.139.1294\MEAS\866	BPB02_84C Second aziridine Threonine	Instrument type and / or accessory	15/12/2015
-------------------------------	--------------------------------------	------------------------------------	------------

IR spectrum of (2*R*,3*R*)-Allyl 1-(2,4-dinitrobenzenesulfonyl)-3-methylaziridine-2-carboxylate  
(33)

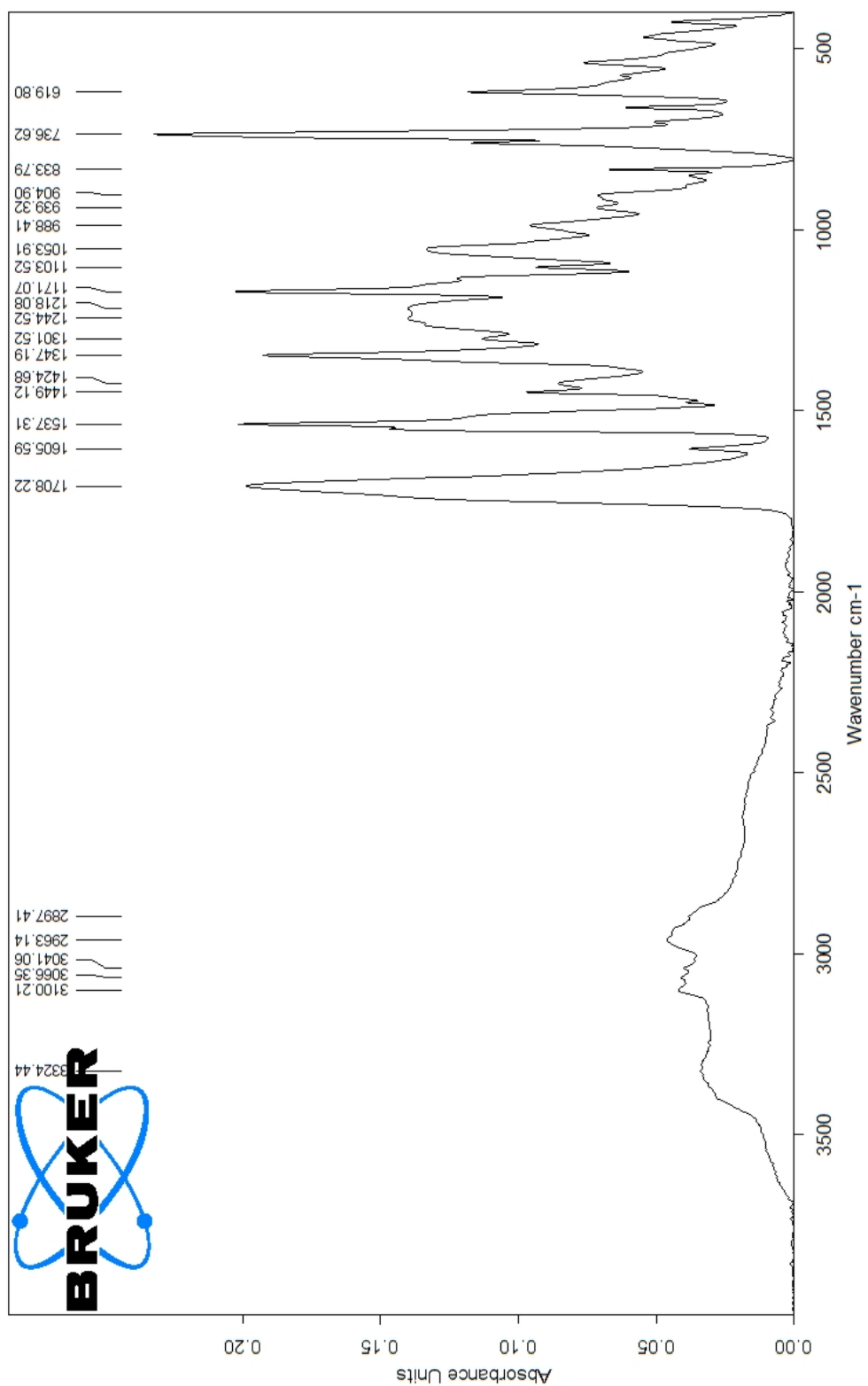


$^1\text{H}$  NMR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2*S*,3*S*)-4-(allyloxy)-3-(2,4-dinitrobenzenesulfonamido)-4-oxobutan-2-ylthio)propanoic acid (**34**) (500 MHz,  $\text{CDCl}_3$ )



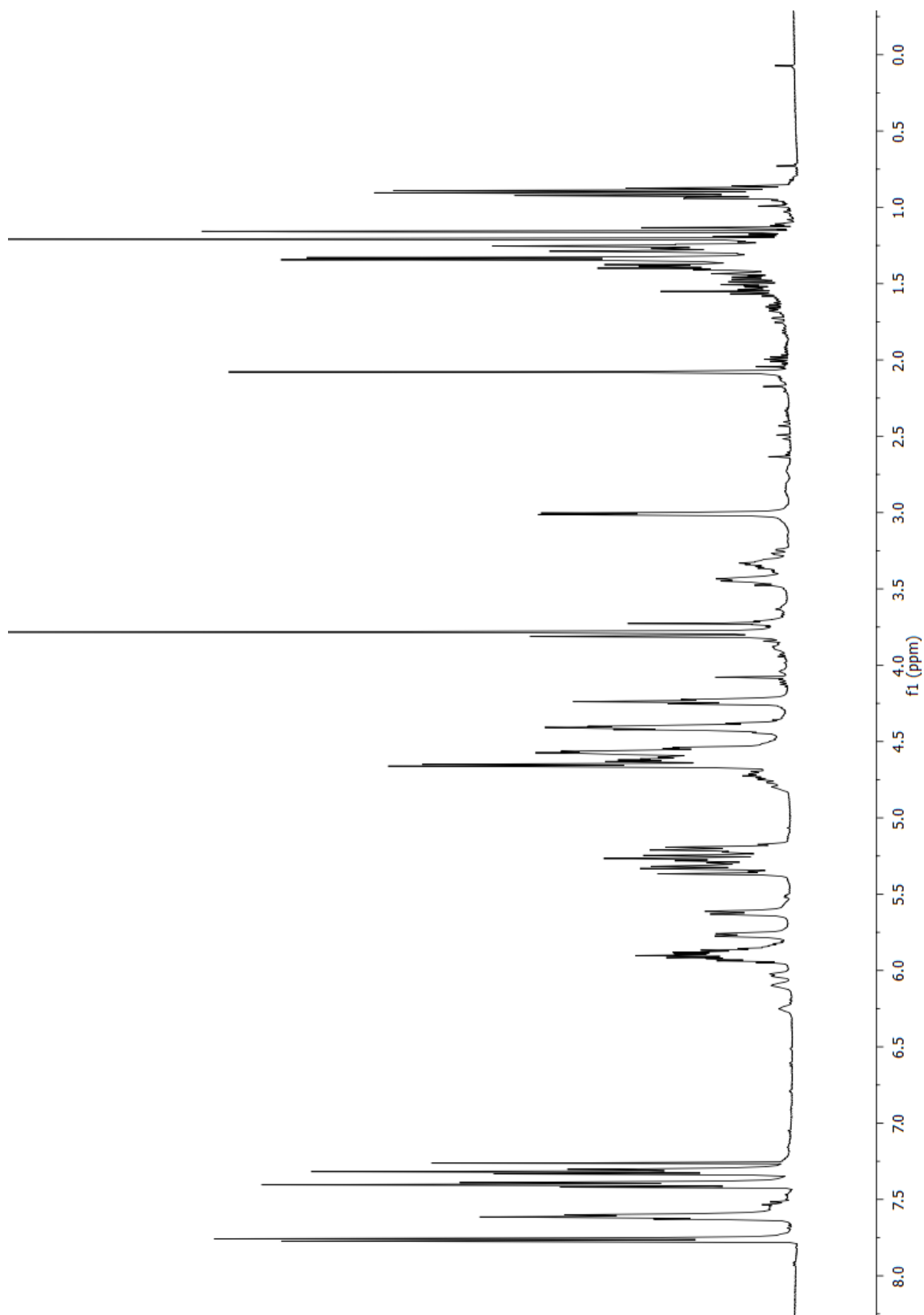


$^{13}\text{C}$  NMR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2*S*,3*S*)-4-(allyloxy)-3-(2,4-dinitrobenzenesulfonamido)-4-oxobutan-2-ylthio)propanoic acid (**34**) (125 MHz,  $\text{CDCl}_3$ )

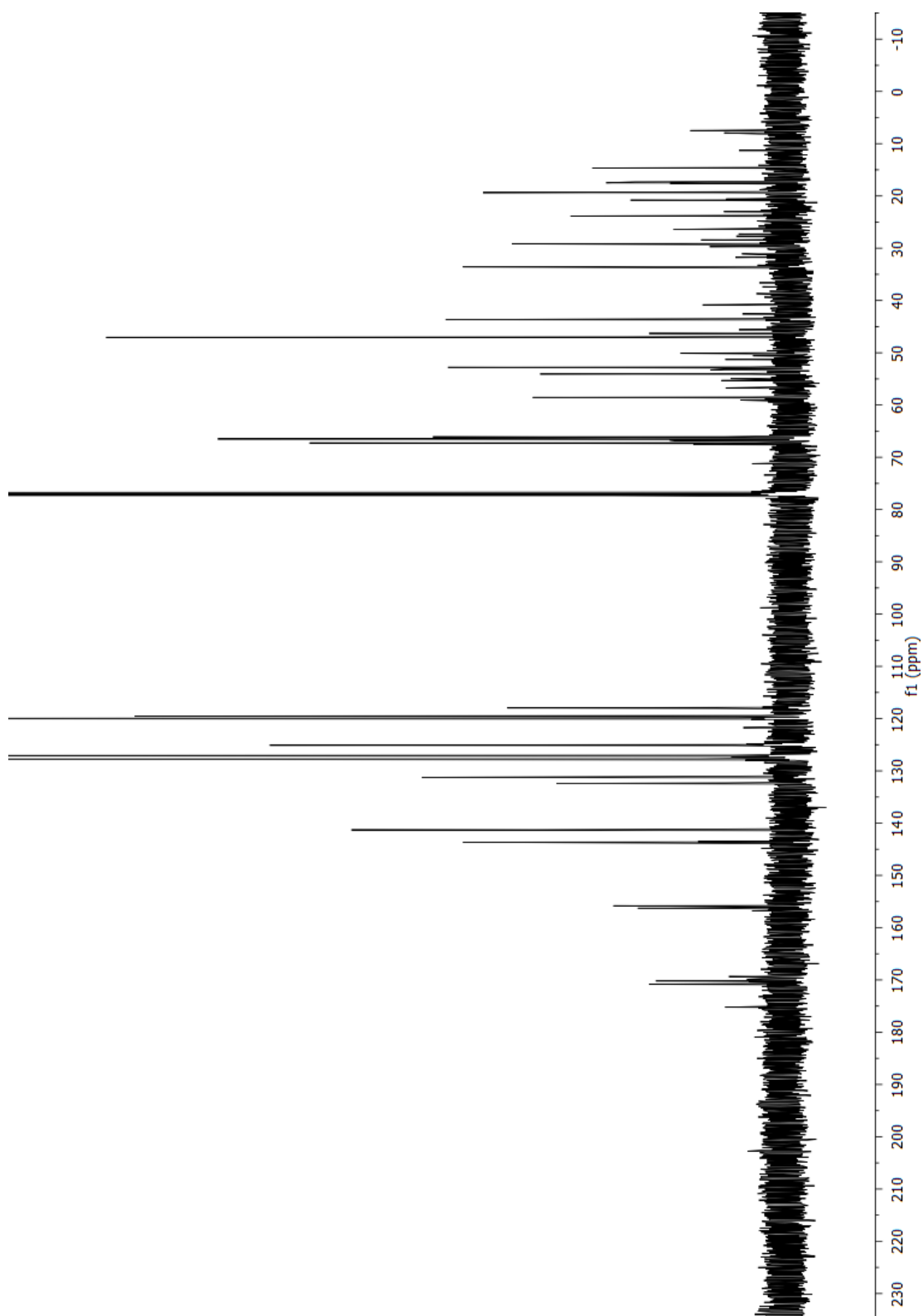


D:\FTIR\BenB\BPB02_78C Melan DNS.0	BPB02_78C Melanthonine DNS	Instrument type and / or accessory	15/12/2015
------------------------------------	----------------------------	------------------------------------	------------

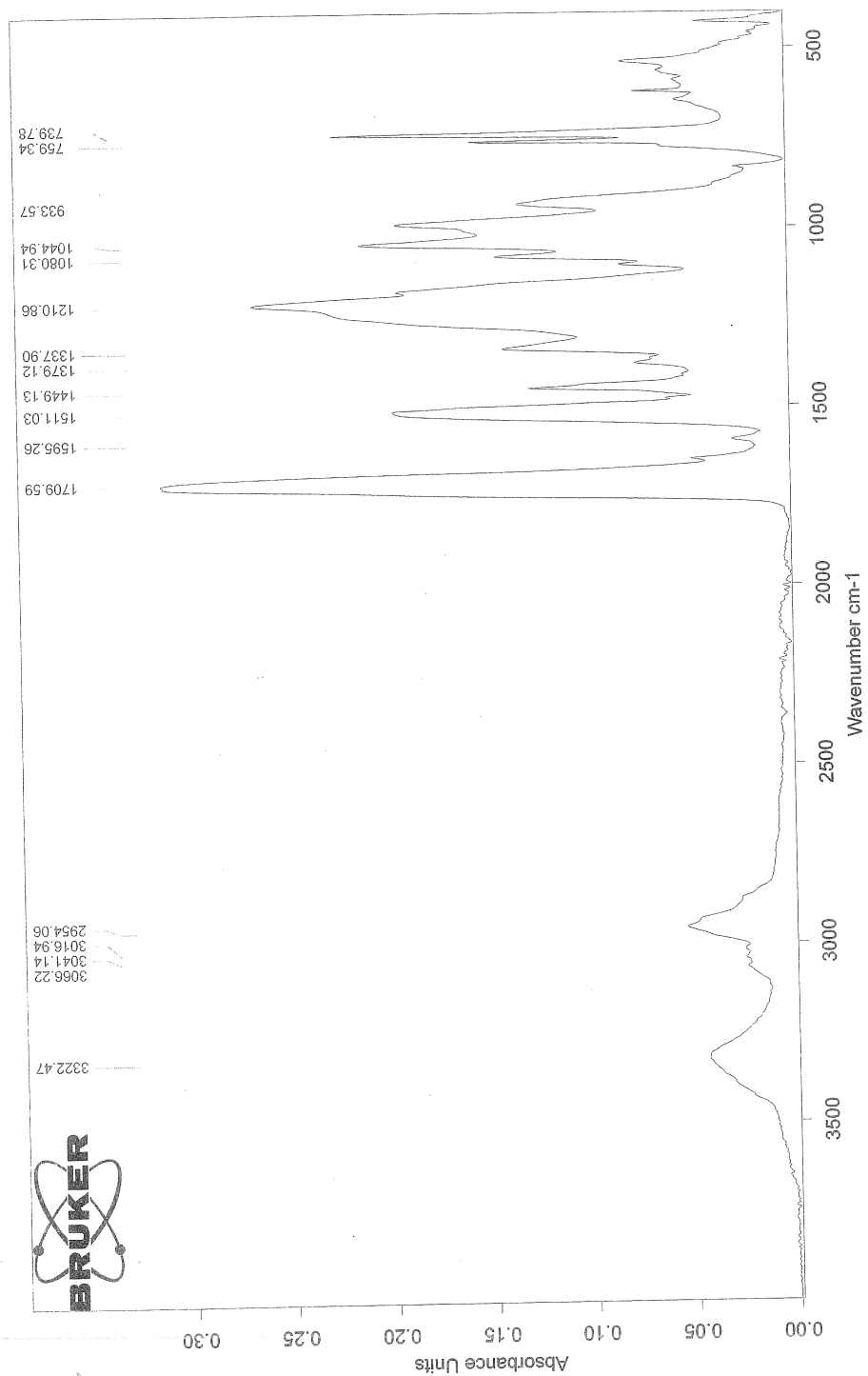
IR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2*S*,3*S*)-4-(allyloxy)-3-(2,4-dinitrobenzenesulfonamido)-4-oxobutan-2-ylthio)propanoic acid (**34**)



$^1\text{H}$  NMR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2*S*,3*S*)-4-(allyloxy)-3-(allyloxycarbonylamino)-4-oxobutan-2-ylthio)propanoic acid (**35**) (500 MHz,  $\text{CDCl}_3$ )



$^{13}\text{C}$  NMR spectrum of (*R*)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2*S*,3*S*)-4-(allyloxy)-3-(allyloxycarbonylamino)-4-oxobutan-2-ylthio)propanoic acid (**35**) (125 MHz,  $\text{CDCl}_3$ )



C:\OPUS_7.2.139\1294\MEAS\1.403	BPB02_83e	Instrument type and / or accessory	7/10/2015
---------------------------------	-----------	------------------------------------	-----------

IR spectrum of (R)-2-(((Fluoren-9-yl)methoxy)carbonylamino)-3-((2S,3S)-4-(allyloxy)-3-(allyloxycarbonylamino)-4-oxobutan-2-ylthio)propanoic acid (**35**)

## References

1. F. Abdullaev and J. Espinosa-Aguirre, *Cancer Detect. Prev.*, 2004, **28**, 426-432.
2. M. A. Huffman, *BioScience*, 2001, **51**, 651-661.
3. J. Wilson, *J. R. Soc. Med.*, 1996, **89**, 135.
4. D. J. Newman, G. M. Cragg and K. M. Snader, *Nat. Prod. Rep.*, 2000, **17**, 215-234.
5. P. Prioreschi, *A History of Medicine: Primitive and ancient medicine*, Horatius Press, 1996.
6. H. Györy, in *Encyclopaedia of the History of Science, Technology, and Medicine in Non-Western Cultures*, ed. H. Selin, Springer Netherlands, 2008, pp. 1508-1518.
7. N. Filipowicz, M. Kamiński, J. Kurlenda, et al., *Phytother. Res.*, 2003, **17**, 227-231.
8. P. Mukeshwar, M. Debnath, S. Gupta, et al., *J. Pharmacognosy Phytother.*, 2011, **3**, 27-37.
9. M. J. Brownstein, *Proc. Natl. Acad. Sci.*, 1993, **90**, 5391-5393.
10. J. Drews, *Science*, 2000, **287**, 1960-1964.
11. A. Fleming, *J. Exp. Pathol.*, 1929, **10**, 226-236.
12. D. Ellis, *J. Antimicrob. Chemother.*, 2002, **49**, 7-10.
13. R. N. Davidson, M. den Boer and K. Ritmeijer, *Trans. R. Soc. Trop. Med. Hyg.*, 2009, **103**, 653-660.
14. V. Miao, M.-F. Coëffet-LeGal, P. Brian, et al., *Microbiology*, 2005, **151**, 1507-1523.
15. S. W. Drew and A. L. Demain, *Annu. Rev. Microbiol.*, 1977, **31**, 343-356.
16. D. H. Williams, M. J. Stone, P. R. Hauck, et al., *J. Nat. Prod.*, 1989, **52**, 1189-1208.
17. M. R. Berenbaum, *Proc. Natl. Acad. Sci.*, 1995, **92**, 2-8.

18. F. E. Koehn and G. T. Carter, *Nat. Rev. Drug Discov.*, 2005, **4**, 206-220.
19. D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2012, **75**, 311-335.
20. J. W.-H. Li and J. C. Vederas, *Science*, 2009, **325**, 161-165.
21. D. J. Newman, *J. Med. Chem.*, 2008, **51**, 2589-2599.
22. T. Henkel, R. M. Brunne, H. Müller, et al., *Angew. Chem. Int. Ed.*, 1999, **38**, 643-647.
23. R. J. Spandl, M. Díaz-Gavilán, K. M. G. O'Connell, et al., *Chem. Rec.*, 2008, **8**, 129-142.
24. A. J. Alanis, *Arch. Med. Res.*, 2005, **36**, 697-705.
25. C. R. Woese, O. Kandler and M. L. Wheelis, *Proc. Natl. Acad. Sci.*, 1990, **87**, 4576-4579.
26. N. R. Pace, *Science*, 1997, **276**, 734-740.
27. C. Schiraldi and M. De Rosa, *Trends Biotechnol.*, 2002, **20**, 515-521.
28. J. Clardy, M. A. Fischbach and C. T. Walsh, *Nat Biotech*, 2006, **24**, 1541-1550.
29. R. D. Macelroy, *BioSyst.*, 1974, **6**, 74-75.
30. D. Schulze-Makuch and J. Seckbach, in *Polyextremophiles*, eds. J. Seckbach, A. Oren and H. Stan-Lotter, Springer Netherlands, 2013, vol. 27, pp. 597-607.
31. F. Canganella and J. Wiegel, *Naturwissenschaften*, 2011, **98**, 253-279.
32. L. J. Rothschild and R. L. Mancinelli, *Nature*, 2001, **409**, 1092-1101.
33. A. A. H. Pakchung, P. J. L. Simpson and R. Codd, *Environ. Chem*, 2006, **3**, 77-93.
34. T. D. Brock and H. Freeze, *J. Bacteriol.*, 1969, **98**, 289-297.
35. G. Schochetman, C.-Y. Ou and W. K. Jones, *J. Infect. Dis.*, 1988, **158**, 1154-1157.
36. Z. E. Wilson and M. A. Brimble, *Nat. Prod. Rep.*, 2009, **26**, 44-71.

37. A. A. Stierle and D. B. Stierle, in *Stud. Nat. Prod. Chem.*, ed. R. Atta ur, Elsevier, 2005, vol. Volume 32, Part L, pp. 1123-1175.
38. P. C. Y. Woo, S. K. P. Lau, J. L. L. Teng, et al., *Clin. Microbiol. Infect.*, 2008, **14**, 908-934.
39. *Ribosomal Database Project*, <http://rdp.cme.msu.edu/>, January 2016.
40. *1000 Springs Project* <http://www.1000springs.org.nz/>.
41. M. B. Stott, M. A. Crowe, B. W. Mountain, et al., *Environ. Microbiol.*, 2008, **10**, 2030-2041.
42. S. Yabe, Y. Aiba, Y. Sakai, et al., *Int. J. Syst. Evol. Microbiol.*, 2011, **61**, 903-910.
43. C. E. King and G. M. King, *Int. J. Syst. Evol. Microbiol.*, 2014, **64**, 1244-1251.
44. M. Vyssotski, J. Ryan, K. Lagutin, et al., *Lipids*, 2012, **47**, 601-611.
45. J. A. McIntosh, M. S. Donia and E. W. Schmidt, *Nat. Prod. Rep.*, 2009, **26**, 537-559.
46. Eric W. Schmidt, *Chem. Biol.*, 2012, **19**, 1501-1502.
47. X. Yang and W. A. van der Donk, *Chem. Eur. J.*, 2013, **19**, 7662-7677.
48. H. W. Van Den Hooven, C. C. M. Doeland, M. Van De Kamp, et al., *Eur. J. Biochem.*, 1996, **235**, 382-393.
49. I. Wiedemann, E. Breukink, C. van Kraaij, et al., *J. Biol. Chem.*, 2001, **276**, 1772-1779.
50. P. M. Jones and A. M. George, *CMLS, Cell. Mol. Life Sci.*, 2004, **61**, 682-699.
51. P. G. Arnison, M. J. Bibb, G. Bierbaum, et al., *Nat. Prod. Rep.*, 2013, **30**, 108-160.
52. J. Dischinger, S. Basi Chipalu and G. Bierbaum, *Int. J. Med. Microbiol.*, 2014, **304**, 51-62.
53. M. Kleerebezem, *Peptides*, 2004, **25**, 1405-1414.
54. S. Boakes, J. Cortés, A. N. Appleyard, et al., *Mol. Microbiol.*, 2009, **72**, 1126-1136.



55. E. J. Aitken, Victoria University of Wellington, 2013.
56. E. Hauch, *Putative lantibiotic gene cluster derived from Thermogemmatispora lignivorax*, GNS Science, 2011.
57. G. B. Brown and V. du Vigneaud, *J. Biol. Chem.*, 1941, **140**, 767-771.
58. J. M. Probert, D. Rennex and M. Bradley, *Tetrahedron Lett.*, 1996, **37**, 1101-1104.
59. D. N. Harpp and J. G. Gleason, *J. Org. Chem.*, 1971, **36**, 73-80.
60. R. K. Olsen, G. D. Kini and W. J. Hennen, *J. Org. Chem.*, 1985, **50**, 4332-4336.
61. L. D. Arnold, R. G. May and J. C. Vederas, *J. Am. Chem. Soc.*, 1988, **110**, 2237-2241.
62. H. Shao, S. H. H. Wang, C.-W. Lee, et al., *J. Org. Chem.*, 1995, **60**, 2956-2957.
63. N. D. Smith and M. Goodman, *Org. Lett.*, 2003, **5**, 1035-1037.
64. K. Nakajima, H. Oda and K. Okawa, *Bull. Chem. Soc. Jpn.*, 1983, **56**, 520-522.
65. W. Liu, A. S. H. Chan, H. Liu, et al., *J. Am. Chem. Soc.*, 2011, **133**, 14216-14219.
66. S. Gabriel, *Ber. Dtsch. Chem. Ges.*, 1888, **21**, 1049-1057.
67. W. Marckwald and O. Frobenius, *Ber. Dtsch. Chem. Ges.*, 1901, **34**, 3544-3558.
68. G. W. Daub, D. A. Heerding and L. E. Overman, *Tetrahedron*, 1988, **44**, 3919-3930.
69. W. McCoull and F. A. Davis, *Synthesis*, 2000, **2000**, 1347-1365.
70. P. B. Karadakov, J. Gerratt, D. L. Cooper, et al., *Comp. Theor. Chem.*, 1995, **341**, 13-24.
71. F. W. Fowler and A. Hassner, *J. Am. Chem. Soc.*, 1968, **90**, 2875-2881.
72. D. Tanner, *Angew. Chem. Int. Ed.*, 1994, **33**, 599-619.
73. J. M. Concellón, P. L. Bernad, J. R. Suárez, et al., *J. Org. Chem.*, 2005, **70**, 9411-9416.
74. J. B. Sweeney, *Chem. Soc. Rev.*, 2002, **31**, 247-258.

75. P. G. M. Wuts and J. M. Northuis, *Tetrahedron Lett.*, 1998, **39**, 3889-3890.
76. S. Boncel, K. Saletra, B. Hefczyc, et al., *Beilstein J. Org. Chem.*, 2011, **7**, 173-178.
77. J. Clayden, N. Greeves and S. Warren, *Organic Chemistry*, OUP Oxford, 2012.
78. R. Williams, W. P. Jencks and F. H. Westheimer, *PKa Data*, Penn State University, 2010.
79. A. A. Farah, S. H. Zheng, S. Morin, et al., *Surf. Sci.*, 2007, **601**, 1677-1683.
80. T. Kan and T. Fukuyama, *Chem. Commun.*, 2004, 353-359.
81. T. I. Al-Warhi, H. M. A. Al-Hazimi and A. El-Faham, *J. SAUDI. CHEM. SOC.*, 2012, **16**, 97-116.
82. E. Fischer and E. Fourneau, *Ber. Dtsch. Chem. Ges.*, 1901, **34**, 2868-2877.
83. L. E. Canne, A. R. Ferre- D'Amare, S. K. Burley, et al., *J. Am. Chem. Soc.*, 1995, **117**, 2998-3007.
84. T. Curtius, *J. prakt. Chem.*, 1882, **26**, 145-208.
85. T. Curtius and F. H. Dellschaft, *J. prakt. Chem.*, 1901, **64**, 419-438.
86. T. Wieland and M. Bodanszky, *The World of Peptides: A Brief History of Peptide Chemistry*, Springer Berlin Heidelberg, 2012.
87. T. Curtius, *J. prakt. Chem.*, 1904, **70**, 57-72.
88. B. Gutte, *Peptides: Synthesis, Structures, and Applications*, Elsevier Science, 1995.
89. H. B. Vickery and C. L. A. Schmidt, *Chem. Rev.*, 1931, **9**, 169-318.
90. E. Fischer, *Ber. Dtsch. Chem. Ges.*, 1905, **38**, 605-619.
91. M. Bergmann and L. Zervas, *Ber. Dtsch. Chem. Ges.*, 1932, **65**, 1192-1201.
92. C. R. Harington and T. H. Mead, *Biochem. J.*, 1935, **29**, 1602-1611.
93. V. du Vigneaud, C. Ressler, J. M. Swan, et al., *J. Am. Chem. Soc.*, 1954, **76**, 3115-3121.

94. F. C. McKay and N. F. Albertson, *J. Am. Chem. Soc.*, 1957, **79**, 4686-4690.
95. R. Schwyzler and P. Sieber, *Nature*, 1963, **199**, 172-174.
96. J. C. Sheehan and G. P. Hess, *J. Am. Chem. Soc.*, 1955, **77**, 1067-1068.
97. H. G. Khorana, *Chem. Ind. (London)*, 1955, **33**, 1087-1088.
98. M. Goodman and L. Levine, *J. Am. Chem. Soc.*, 1964, **86**, 2918-2922.
99. M. Goodman and W. J. McGahren, *Tetrahedron*, 1967, **23**, 2031-2050.
100. W. Koenig and R. Geiger, *Chem. Ber*, 1970, **103**, 2030.
101. Z. J. Kamiński, *Pept. Sci.*, 2000, **55**, 140-164.
102. R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149-2154.
103. J. P. Tam, W. F. Heath and R. B. Merrifield, *J. Am. Chem. Soc.*, 1983, **105**, 6442-6455.
104. L. A. Carpino and G. Y. Han, *J. Am. Chem. Soc.*, 1970, **92**, 5748-5749.
105. M. Schnölzer, P. Alewood, A. Jones, et al., *Int. J. Pept. Protein Res.*, 1992, **40**, 180-193.
106. T. Kimmerlin and D. Seebach, *J Pept Res.*, 2005, **65**, 229-260.
107. E. Bayer, G. Jung, I. Halász, et al., *Tetrahedron Lett.*, 1970, **11**, 4503-4505.
108. I. L. Rodionov, M. Baru and V. Ivanov, *Pept. Res.*, 1992, **5**, 119-125.
109. L. E. Cammish and S. A. Kates, *Fmoc Solid Phase Peptide Synthesis, A Practical Approach*, P.D. White Eds, Oxford Univ. Press, 2000.
110. M. Erdélyi and A. Gogoll, *Synthesis*, 2002, **2002**, 1592-1596.
111. J. E. Fox, R. Newton and C. H. Stroud, *Int. J. Pept. Protein Res.*, 1991, **38**, 62-65.
112. S. A. Salisbury, E. J. Tremeer, J. W. Davies, et al., *J. Chem. Soc., Chem. Commun.*, 1990, 538-540.

113. K. H. Hsieh, M. M. Demaine and S. Gurusidaiah, *Int. J. Pept. Protein Res.*, 1996, **48**, 292-298.
114. M. Amblard, J.-A. Fehrentz, J. Martinez, et al., *Mol. Biotechnol.*, 2006, **33**, 239-254.
115. E. Valeur and M. Bradley, *Chem. Soc. Rev.*, 2009, **38**, 606-631.
116. E. Kaiser, R. L. Colescott, C. D. Bossinger, et al., *Anal. Biochem.*, 1970, **34**, 595-598.
117. E. W. Yemm, E. C. Cocking and R. E. Ricketts, *Analyst*, 1955, **80**, 209-214.
118. C. B. Bottom, S. S. Hanna and D. J. Siehr, *Biochem Educ.*, 1978, **6**, 4-5.
119. A. Thaler, D. Seebach and F. Cardinaux, *Helv. Chim. Acta*, 1991, **74**, 628-643.
120. R. Beugelmans, L. Neuville, M. Bois-Choussy, et al., *Tetrahedron Lett.*, 1995, **36**, 3129-3132.
121. A. Isidro-Llobet, M. Álvarez and F. Albericio, *Chem. Rev.*, 2009, **109**, 2455-2504.
122. N. Garg, W. Tang, Y. Goto, et al., *Proc. Natl. Acad. Sci.*, 2012, **109**, 5241-5246.
123. R. C. Christopher, S. C. Phillip and S. G. Michael, *Curr. Protein Peptide Sci.*, 2005, **6**, 77-84.
124. A. Biosystems, *Cleavage, Deprotection, and Isolation of Peptides after Fmoc Synthesis*.
125. W. Z. Shou and W. Naidong, *J. Chromatogr. B*, 2005, **825**, 186-192.
126. T. Su, *Br. J. Exp. Pathol.*, 1948, **29**, 473.
127. L. C. Wieland Brown, M. G. Acker, J. Clardy, et al., *Proc. Natl. Acad. Sci.*, 2009, **106**, 2549-2553.
128. M. C. Bagley, J. W. Dale, E. A. Merritt, et al., *Chem. Rev.*, 2005, **105**, 685-714.
129. M. J. LaMarche, J. A. Leeds, J. Dzink-Fox, et al., *J. Med. Chem.*, 2012, **55**, 6934-6941.
130. C. T. Walsh, M. G. Acker and A. A. Bowers, *J. Biol. Chem.*, 2010, **285**, 27525-27531.

131. H. Mohimani and P. A. Pevzner, *Nat. Prod. Rep.*, 2016, **33**, 73-86.
132. G. Siuzdak, *The Expanding Role of Mass Spectrometry in Biotechnology*, MCC Press, 2006.
133. K. Chughtai and R. M. A. Heeren, *Chem. Rev.*, 2010, **110**, 3237-3277.
134. M. Karas and F. Hillenkamp, *Anal. Chem.*, 1988, **60**, 2299-2301.
135. A. E. Clark, E. J. Kaleta, A. Arora, et al., *Clin. Microbiol. Rev.*, 2013, **26**, 547-603.
136. *National High Magnetic Field Laboratory*, <https://nationalmaglab.org/user-facilities/icr/techniques/maldi>, 2015.
137. Y.-L. Yang, Y. Xu, R. D. Kersten, et al., *Angew. Chem. Int. Ed.*, 2011, **50**, 5839-5842.
138. S. Shimma, Y. Sugiura, T. Hayasaka, et al., *J. Chromatogr. B*, 2007, **855**, 98-103.
139. N. Zaima, T. Hayasaka, N. Goto-Inoue, et al., *J. Oleo Sci.*, 2009, **58**, 415-419.
140. J. Y. Yang, V. V. Phelan, R. Simkovsky, et al., *J. Bacteriol.*, 2012, **194**, 6023-6028.
141. R. C. Beavis, T. Chaudhary and B. T. Chait, *Org. Mass Spectrom.*, 1992, **27**, 156-158.
142. K. Strupat, M. Karas and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Processes*, 1991, **111**, 89-102.
143. H. Sato, S. Nakamura, K. Teramoto, et al., *J. Am. Soc. Mass. Spectrom.*, 2014, **25**, 1346-1355.
144. H. A. Favre and W. H. Powell, *Nomenclature of Organic Chemistry: IUPAC Recommendations and Preferred Names 2013*, Royal Society of Chemistry, 2013.