

Justus-Liebig University Giessen

Department 09 - Agricultural Sciences, Nutritional Sciences, and Environmental Management

Dissertation

Submitted in Fulfillment of the Requirements for the Degree of *Doctor rerum naturalium*

(Dr. rer. nat.)

Synthetic Approaches to New Antimicrobial Active Natural Products

by Sandra Semmler

Gießen, April 2022

1st examiner: Prof. Dr. Till F. Schäberle

2nd examinier: Prof. Dr. Peter E. Hammann

Decleration of Originalits / Eigenständigkeitserklärung

Erklärung gemäß der Promotionsordnung des Fachbereichs 09 vom 01.02.2005.

Ich versichere, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Ausführungen, die anderen Schriften wörtlich oder sinngemäß entnommen wurden, sind kenntlich gemacht. Ich habe die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt. Ich stimme zu, dass die vorliegende Arbeit mit einer Anti-Plagiatssoftware überprüft werden darf. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Table of Contents

Declaration of Originality / Eigenständigkeitserklärung	I
Table of Contents	II
Abbreviations	IV
Acknowledgements	VI
Abstract	VIII
1. Introduction	1
2. Objective	3
3. Materials and methods	5
4. Chapter 1 – Globomycin	7
4.1 Introduction	7
4.2 Results and Discussion	13
4.2.1 Molecular Networking and Producer Strain	13
4.2.2 Early ADME tests on Globomycin	16
4.2.3 Structure Based Drug Design (SBDD) and Molecular Docking	16
4.2.4 Synthesis	19
4.2.5 Antibacterial Activity	26
4.3 Summary and Outlook	28
4.4 Experimental	32
4.4.1 Molecular Networking analysis	32
4.4.2 Procedures for the Globomycin Syntheses	32
4.4.3 Activity Tests - Bioassays	58
4.5 Supporting Information	60
4.5.1 Chromatograms and spectra of natural extract	60
4.5.2 Structure Based Drug Design	76
4.5.3 NMR and MS/MS ² data for synthetic compounds	78
5. Chapter 2 – Falcitidin analogs	99
5.1 Manuscript	100
5.2 Supporting Information	114
6. Chapter 3 – Cryopeptides	197

6.1	Introduction.....	197
6.2	Results and Discussion.....	201
6.2.1	Synthesis.....	201
6.2.2	Comparison of Synthetic to Natural Isolated Compounds	204
6.2.3	Activities	212
6.3	Summary and Outlook.....	213
6.4	Experimental.....	215
6.5	Supporting Information	228

Abbreviations

2-CT	2-Chlorotrityl	HR-MS	High resolution-mass spectrometry
2-CTC	2-Chlorotrityl chloride	HSQC	Heteronuclear single quantum correlation
AA	Amino acid	IC ₅₀	Half maximal inhibitory concentration
Ac	Acetyl	IME	Institute for Molecular Biology and Applied Ecology
AcCN	Acetonitrile	iVal	Isovaleroyl
ADME	Absorption, distribution, metabolism, excretion	LC-MS	Liquid chromatography-mass spectrometry
Alloc	Allyloxycarbonyl	LSP A	Lipoprotein signal peptidase II
anhyd.	Anhydrous	M	Molar
arom.	Aromatic	<i>m/z</i>	Mass-to-charge ratio
Aux	Auxiliary	MD	Medical doctor
BHA	β-Hydroxy fatty acid	Me	Methyl
BMS	Borane-methyl sulfide	MedChem	Medicinal chemistry
Bn	Benzyl	MeOH	Methanol
BPC	Base peak chromatogram	Mes	Mesitylene
calcd	Calculated	MH II	Mueller Hinton II
CatB	Cathepsin B	MHC	Mueller Hinton Carbonate
CatL	Cathepsin L	MIC	Minimum inhibitory concentration
c-Hex	Cyclo-hexyl	min	Minutes
CYP	Cytochrome P450	mmol	Milli Mol
Da	Dalton	MS	Mass spectrometry
DAD	Diode array detector	MS/MS	Tandem mass spectrometry
DCM	Dichloromethane	n.a.	Not assigned
<i>de</i>	Diastereomeric excess	n.d.	No data
DEPC	Diethyl cyanophosphonate	n.o.	Not observed
DIC	<i>N,N'</i> -Diisopropylcarbodiimide	NP	Natural product
DIPEA	<i>N,N</i> -Diisopropylethylamine	NRPS	Non-ribosomal peptide synthetases
DMAP	4-Dimethylaminopyridine	OTf	Triflate
DMF	<i>N,N</i> -Dimethylformamide	PA	Phenylacetyl
DMP	Dess-Martin periodinane	PDB	Protein data bank
DMSO	Dimethyl sulfoxide	PE	Petroleum ether
EA	Ethyl acetate	PG	Protecting group
EDC	<i>N</i> -(3-Dimethylaminopropyl)- <i>N'</i> -ethylcarbodiimide	Ph	Phenyl
EIC	Extracted ion chromatogram	qTOF	Quadrupole time-of-flight
ESI	Electrospray ionization	RBM	<i>R</i> -Methylbutanoyl
eV	Electron volt	R _f	Retention factor
FA	Formic acid	rt	Room temperature
Fmoc	FLuorenylmethoxycarbonyl	RP	Reverse phase
Fmoc-OSu	<i>N</i> -(9-Fluorenylmethoxy-carbonyloxy)succinimide	SAR	Structure-Activity Relationship
GLM	Globomycin	sat.	Saturated
h	Hour	SBDD	Structure based drug design
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate	SBM	<i>S</i> -Methylbutanoyl
HD	High definition	SF-1902	Older family name of Globomycin congeners, named after the strain from which it was first isolated
HFIP	Hexafluoroisopropanol	SNAC	<i>N</i> -acetylcysteamine thioester
HMBC	Heteronuclear multiple bond correlation	SPPS	Solid-phase peptide synthesis
HOAt	1-Hydroxy-7-azabenzotriazole	T3P	Propanephosphonic acid anhydride
HPLC	High performance liquid chromatography	TBAF	Tetra- <i>n</i> -butylammonium fluoride
		TBS	<i>Tert</i> -butyldimethylsilyl (also TBDMS)
		TFA	Trifluoroacetic acid

TFE	Trifluoroethanol
THF	Tetrahydrofuran
TIS	Triisopropylsilane
TLC	Thin layer chromatography
Trt	Trityl
UHPLC	Ultra-high performance liquid chromatography
quant.	Quantitatively
UHR	Ultra-high resolution
UHR-MS	Ultra-high resolution-mass spectrometry
UPLC	Ultra-performance liquid chromatography
UV	Ultraviolet
WHO	World health organization
wt	Wild type

Bacteria and fungi

<i>A. baumannii</i>	<i>Acinetobacter baumannii</i>
<i>A. flavus</i>	<i>Aspergillus flavus</i>
<i>C. albicans</i>	<i>Candida albicans</i>
<i>E. coli</i>	<i>Escherichia coli</i>
<i>K. pneumoniae</i>	<i>Klebsiella pneumoniae</i>
<i>M. smegmatis</i>	<i>Mycobacterium smegmatis</i>
<i>M. tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
<i>sp.</i>	Species

Amino acids

Arg; R	Arginine
Asn; N	Asparagine
Asp; D	Aspartic acid
C-I/L/P/F	c-Terminal isoleucine / leucine / proline / phenylalanine
Dhb	Dehydrobutyryne
Dhv	Dehydrovaline
d-V	Dehydrovaline
Gly; G	Glycine
His; H	Histidine
Hyp	Hydroxyproline
Ile; I	Isoleucine
Leu; L	Leucine
Phe; F	Phenylalanine
Pro; P	Proline
Ser; S	Serine
Thr; T	Threonine
Val; V	Valine

NMR

nuclear magnetic resonance	
δ	chemical shift [ppm]
(M)Hz	(Mega) Hertz
1D, 2D, 3D	Dimensional
alkyl	Alkyl
bs	Broad signal
d	Doublet
dd	Doublet of doublet
dq	Doublet of quartet
<i>J</i>	Coupling constant [Hz]
m	Multiplett
ppm	Parts per million
quart	Quaternary
s	Singulet
t	Triplet

Acknowledgements

First, I thank Prof. Dr. Till F. Schäberle, my final supervisor of the thesis, and Prof. Dr. Peter E. Hammann, who supported me as my supervisor for the first year. Thank you both for supporting me throughout this thesis.

A very special thanks goes to Dr. Armin Bauer, for all the input, was it chemical or otherwise and being there whenever questions came up; I would not have been able to finish this without your guidance.

To Prof. Dr. Andreas Vilcinskas as the head of Fraunhofer IME, and to the whole team, especially the group of Natural Products for the support. To Dr. Maria Patras as the head of the analytic and chemistry department of the natural product group and her introducing me to molecular networking and MS/MS interpretation. To Dr. Michael Marnier as head of the screening department for all the screenings he and his team carried out and for the extracurricular activity, be it running, bouldering or diving. To Christoph Hartwig for his technical assistance with any equipment and whenever questions or problems arose. To my colleagues of the analytics and chemistry group, Angelika Brohl, Mona Abdullahi, and Judith Härter for the good atmosphere and helping with purification and everything else. To my fellow PhD students Celine Zumkeller, Markus Oberpaul, Yolanda Kleiner, Mona Bill, and Stephan Brinkmann for the amazing time together, in- and outside the lab. Especially to Stephan and our shared project of Falcitidin analogs. If you would not have found it and took me on, I would not have this amazing opportunity. Thank you for that and also the smooth teamwork. Additionally, to Mona, my climbing partner and all-around sports motivation. To my bachelor student Niclas Kulhanek, and the student assistants, Lysander Wagner, Julian Roth, and especially Victoria Öhler for their help in the lab.

To all the employees of the Justus-Liebig-University, from the chemical department to the NMR service, especially Dr. Heike Hausmann and her team for measuring the NMR samples and for discussions and help. A special thanks to Anja Beneckenstein of the glass blowing unit. Without her and her expertise the custom build peptide vessels would not have been possible. And to Friedemann Dreßler for proof reading, the chemical discussions and the coffee.

To the collaboration of Fraunhofer and Sanofi and later Evotec for the interesting insight into industrial pharmaceutical research and the amazing support throughout my PhD time.

From Sanofi side, as mentioned above Dr. Armin Bauer, a special thanks to Dr. Hans Matter, who not just performed the molecular docking studies, but was also involved in interesting scientific discussions. These meetings have always been one of the most productive and educational meetings during that time and it was something I always looked forward to. To Dr. Michael Kurz for NMR measurements and interpretation. To Dr. Christoph Pöverlein for his help regarding chemical questions and for his synthetic input. And to all involved people of the cooperation who performed tests and provided scientific discussions and feedback. From Evotec side a special thanks to Dr. Sören Schuler for the organization and scientific discussions, as well as the team of Lyon which worked on Globomycin and performed activity tests.

I could go on and on but I will end with the most important support, namely my family and friends.

Abstract

Natural products (NPs) present great chances to identify novel antimicrobial active substances with unprecedented structures. Synthetic access to natural derived compounds is essential in the development process of new drugs. However, usually scarcity of isolated material makes complete activity profiling difficult. Hence, there is the need for synthetic routes to acquire not just the natural product itself, but rather derivatives for the development of in-depth structure-activity relationship (SAR) studies for future drug development. Peptide-based NPs with a molecular weight between “small drugs” and “biologics” are still an underexplored type of NPs, which have the potential to combine the advantage of high selectivity and bioavailability.

Chapter 1 – The known natural product Globomycin (GLM) is a cyclic peptide that shows activity against Gram-negative bacteria (MIC of 6.25 µg/mL for *E. coli* SANK 70569)²⁹. After identifying an in-house producer strain, molecular network analysis of extracts derived from *Streptomyces sp.* HAG010519 revealed a total of 29 natural derivatives of GLM. For eleven derivatives a structural proposal based on MS/MS data has been given. Early ADME (Absorption, distribution, metabolism, excretion) tests were performed and revealed a high metabolic liability of GLM. To address this issue, a rational design approach to develop more stable derivatives was performed. Therefore, structure-based drug design (SBDD) and molecular docking studies were applied to focus the synthesis on the most promising derivatives. A solid-phase peptide synthesis (SPPS) approach was developed, with a macrolactamization between amino acids 5 and 4 (Figure 4-7) for the ring closure. The newly developed synthetic route has overall fewer synthetic steps, a high stereoselectivity for the fatty acid side chain, and gives a two-step procedure to introduce more extraordinary amino acid building blocks like hydroxyprolines. The four successfully synthesized derivatives show noticeable activity against the tested *E. coli* strains, with the highest one of 16 µg/mL, and that the docking scores delivered a good assessment between the compounds.

Chapter 2 – Falcitidin is an inhibitor of cysteine protease falcipain-2. Molecular network revealed over 30 natural analogs of falcitidin. Total synthesis of chosen analogs was achieved using a SPPS approach followed by functional group interconversion of the cleaved peptide acid to the alcohol, followed by Dess-Martin oxidation to the aldehyde. Thereby, access to functionalized pentapeptides was established. *In vitro* testing against selected proteases, as well as falcipain-2 and -3 showed superior inhibitory activity than falcitidin itself.

Chapter 3 – In extracts of *Pedobacter cryoconitis* linear peptides containing dehydro amino acids were detected by metabolomics analyses. The compounds were called cryopeptides and contain two dehydrogenated valines in their structure. This feature can be found in other NPs as well. However, the biochemical basis of this dehydrogenation process of amino acids is not reported yet. To enable further studies in this direction and to investigate which substrate is used, precursor molecules were successfully synthesized. Furthermore, a synthetic route to two Cryopeptides and to two non-natural derivatives was developed for full structure elucidation, activity testing of the molecules and future enzymatic assays.

1. Introduction

Looking at the numbers, the importance of natural products (NPs) in the development and discovery of new and novel drugs is apparent. Roughly 50% of the approved drugs from January 1981 to September 2019 were either i) natural products or derivatives (23.5%) or ii) synthetic drugs based on the structure of a natural product or a mimic thereof (25.7%).¹ In the case of approved antibiotics, over 80% are natural products or are based on their structure.² Secondary metabolites are natural products, which are not essential for the development of an organism but originate in the organism adapting to environmental changes.^{3, 4} Their structural diversity and unique features often do not follow “Lipinski’s rule of five”⁵ and differ vastly, compared to molecules derived from combinatorial chemistry approaches.^{3, 4} Especially microorganisms show a huge potential for the discovery of novel classes of antimicrobial active substances and future drugs based on NPs.^{6, 7} For years, NPs have been a great source for “natural product discovery” for the pharmaceutical industry. Due to an increase of the re-discovery of known molecules, the high costs as well as complex regulations regarding intellectual properties, natural product discovery programs declined over the last couple of years.^{3, 4}

Peptide-based NPs have seen an increase in interest by pharmaceutical industry over the last years.^{8, 9} With less than 50 amino acids,⁹ they bridge the gap between small molecule drugs (<500 Da) and so called “biologics”, drugs consisting of proteins, growth factors or antibodies, all with a high molecular weight (>5000 Da).¹⁰ Protein-based drugs show great selectivity and less side effects but need to be delivered *via* injection. In contrast, small molecule drugs often

¹ D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2020**, *83*, 770–803.

² M. Lakemeyer, W. Zhao, F. A. Mandl, P. Hammann, S. A. Sieber, *Angew. Chem. Int. Ed.* **2018**, *57*, 14440–14475.

³ D. A. Dias, S. Urban, U. Roessner, *Metabolites*, **2012**, *2*, 303–336.

⁴ A. Atanasov *et al.* *Nat Rev Drug Discov*, **2021**, *20*, 200–216.

⁵ C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Advanced Drug Delivery Reviews*, **2001**, *46*, 1–3, 3–26.

⁶ P. Monciardini, M. Iorio, S. Maffioli, M. Sosio, S. Donadio, *Microb. Biotechnol.* **2014**, *7*, 209–220.

⁷ R. Müller, J. Wink, *Int. J. Med. Microbiol.* **2014**, *304*, 3–13.

⁸ T. Dang, R. D. Süßmuth, *Acc. Chem. Res.* **2017**, *50*, 1566–1576.

⁹ G. B. Santos, A. Ganesan, F. S. Emery, *ChemMedChem*, **2016**, *11*, 1–8.

show more side effects but better oral bioavailability.¹⁰ Peptides cover the gap between these two groups of drugs and are an underexplored area in the drug market by now. They emerge as promising candidates to combine the advantages of both drug classes.¹⁰ They can be biochemically divided into two classes of peptides based on their formation: i) ribosomally synthesized and post-translationally modified peptides (RiPPs) and ii) nonribosomally synthesized peptides (NRPs).⁸

Advances in technology as analytical methods, cultivation and genome mining have the potential to reinvigorate the NP-based drug discovery approach.⁴ Higher sensitivity and resolution of analytical methods make it possible to find even small traces of active compounds and to elucidate their structure. The isolation of the active compound oftentimes yields just enough material for basic activity tests and structure elucidation. Due to the scarcity of the isolated compound, synthetic access is still of high importance: Generation of enough material for further testing and researching the mode of action on the one hand and for the development of derivatives to generate a structure-activity-relationship (SAR) on the other.

The collaboration of Fraunhofer with the industrial pharmaceutical partners Sanofi and later Evotec creates a beneficial setting for the discovery and development of novel natural active compounds. The cooperation between academic research and the long-term expertise of pharmaceutical companies support the development of novel microbial active compounds derived from natural products for not just pharmaceutical but also veterinary and agricultural applications.

¹⁰ D. J. Craik, D. P. Fairlie, S. Liras, D. Price, *Chem Bio Drug Des*, **2013**, *81*, 136-147.

2. Objective

Synthetic access to newly found natural products after their discovery and structure elucidation is still indispensable to this date. Computational chemistry and models cannot deliver accurate results on activity but just give an indication, especially comparing derivatives among themselves. Therefore, enough material needs to be synthesized and a synthetic route has to be established in order to not just synthesize the one molecule but rather to be flexible to synthesize derivatives as well. For peptide-based molecules solid-phase peptide synthesis (SPPS) is the method of choice for quick and easy access to peptide scaffolds. Further functionalization can be carried out either on resin or after the cleavage of the peptide from the solid support.

This work is divided in three different projects, each focusing on a specific peptide-based compound in the field of bioactive natural products and their synthesis. Each chapter contains a brief introduction, a results and discussion section, a summary and outlook, and an experimental part as well as the associated supporting information. The materials and methods are combined in one preceding chapter and the references are included chapter wise.

Chapter 1 – Globomycin

Globomycin (GLM, **1.1**) is an underdeveloped literature known antibacterial depsicyclopeptide with promising anti Gram-negative activity. To determine possible weaknesses, ADME studies on GLM will be performed. The results will be incorporated into a rational design approach for new derivatives. Utilizing the published crystallography data, structure-based drug design (SBDD) and molecular docking will be the key methods to narrow down promising structures. Following this, a flexible and robust synthetic route will be developed. It will be concluded by activity tests of the synthesized derivatives.

In a second approach, it is planned to find an in-house producer strain and investigate it to identify additional active natural congeners for an in-detail structure-activity relationship (SAR). Therefore, extracts will be generated and analyzed by UHPLC-MS/MS to create a molecular network in which derivatives will cluster with GLM.

Chapter 2 – Falcitidin analogs

Falcitidin analogs are structurally related to falcitidin, an active inhibitor of the protease Falcipain 2 which represents a promising target in malaria therapy. Falcitidin was found as part of another project in a molecular networking. Its clustering in the molecular network will be explored further for more natural analogs. Following structure elucidation based on isolated compounds and structure hypotheses based on MS data, a flexible synthetic route to natural found molecules as well as future derivatives will be developed and compounds will be synthesized, followed by activity testing. This chapter will be presented as the submitted manuscript.

Chapter 3 – Cryopeptide

The so-called Cryopeptides are small linear pentapeptides, found in the molecular network of *Pedobacter cryoconitis* by Luis J. Linares Otoyá from the research group of Prof. Dr. T. Schäberle from the Justus-Liebig-University. They contain structural interesting dehydrogenated amino acids. This chapter will focus on the synthesis of two natural isolated compounds for comparison and full structure elucidation as well as non-natural derivatives and aminoacyl-*N*-acetylcysteamine thioesters (SNACs) of valine to examine the dehydrogenation process of the amino acids.

3. Materials and methods

All chemicals were purchased from commercial suppliers in the highest quality and used without further purification. The amino acids and resins specifically were purchased from Iris Biotech. Anhydrous solvents were purchased from Acros Organics and were stored under an argon atmosphere.

For monitoring reaction progresses, an 1100 HPLC system with DAD from Agilent (Agilent, Santa Clara, CA, USA) and an Amazon (Bruker, Billerica, MA, USA) ESI Ion trap spectrometer were used. For TLC glass plates from Merck (TLC Silica gel 60 F₂₅₄) were used. Visualization was carried out using UV-light (254 nm), and/or TLC dips, such as: potassium permanganate stain (4.5 g KMnO₄, 30 g K₂CO₃, 4 mL 10 % aqueous NaOH solution, 450 mL H₂O), ninhydrin stain (5 g ninhydrin, 150 mL EtOH), or ceric ammonium molybdate stain (25 g H₃[PMo₁₂O₄₀], 10 g Ce(SO₄)₂, 60 mL conc. H₂SO₄, 940 mL H₂O) followed by gentle heating.

For manual column chromatography glass columns of different diameters (1-8 cm) and silica from Macherey Nagel (silica 60) were used. For automated flash chromatography a Biotage® SP4 from Biotage (Biotage, Uppsala, Sweden) with ISOLUTE Flash Si II columns of different sizes from Interchim were used.

HPLC purifications were done on semi preparative 1100 HPLC and 1200 HPLC systems with DAD from Agilent (Agilent, Santa Clara, CA, USA) and a Gilson (Gilson Incorporated, Middleton, WI, USA) fraction collector. For all purifications the NUCLEODUR® C18 Gravity SB, 3 µm, 250 x 10 mm column was used. Acetonitrile (HPLC grade) was purchased from Fischer scientific and used without further purification. Water was purified using a Sartorius (Sartorius AG, Göttingen, Germany) system.

UHPLC-UHR-MS analysis was performed on a 1290 UHPLC system (Agilent, Santa Clara, CA, USA) equipped with DAD and maXis II™ (Bruker, Billerica, MA, USA) ESI-qTOF-UHR-MS with the gradient: 0 min: 95% A; 0.30 min: 95% A; 18.00 min: 4.75% A; 18.10 min: 0% A; 22.50 min: 0% A; 22.60 min: 95% A; 25.00 min: 95% A (A: H₂O, 0.1% formic acid (FA); B: Acetonitrile, 0.1% FA; Flow: 600 µL/min). Column oven temperature: 45 °C. Column: Acquity UPLC BEH C18 1.7µm (2.1x100 mm) with Acquity UPLC BEH C18 1.7µm VanGuard Pre-Column (2.1x5 mm). Collision induced dissociation was performed at 28.0–35.05 eV using argon at 10⁻² mbar.

The specific rotation of chiral compounds was determined on a digital polarimeter of the model P3000 from Krüss (A.Krüss Optronic GmbH, Hamburg, Germany). Standard wavelength was the sodium D-line with 589 nm. Temperature and concentration (mg/mL) are reported with the determined value.

NMR spectra of synthesized molecules were recorded on an AVANCE III HD 600 spectrometer (600 MHz for ^1H , 151 MHz for ^{13}C), an AVANCE III 400 HD spectrometer (400 MHz for ^1H , 101 MHz for ^{13}C) and an AVANCE III 400 spectrometer (400 MHz for ^1H , 101 MHz for ^{13}C) from Bruker Biospin (Bruker Biospin GmbH, Rheinstetten, Germany). Chemical shifts are reported in ppm and were referenced to the corresponding residual solvent signal (CDCl_3 : $\delta_{\text{C}} = 77.16$ ppm, $\delta_{\text{H}} = 7.26$ ppm; $\text{DMSO-}d_6$: $\delta_{\text{C}} = 39.52$ ppm, $\delta_{\text{H}} = 2.50$ ppm; D_2O : $\delta_{\text{H}} = 4.79$ ppm; CD_3OD : $\delta_{\text{C}} = 49.00$ ppm, $\delta_{\text{H}} = 3.31$ ppm). δ_{C} shifts marked with an * were not observed in the ^{13}C NMR spectrum, but obtained either from HMBC or HSCQ data.

NMR spectra of natural isolated fractions of Falcitidin analogs were measured and interpreted by Michael Kurz (Sanofi, Frankfurt) and were acquired on a Bruker AVANCE 700 spectrometer (700 MHz for ^1H , 176 MHz for ^{13}C) and a Bruker AVANCE 500 spectrometer (500 MHz for ^1H , 126 MHz for ^{13}C). Both instruments were equipped with a 5 mm TCI cryo probe. For structure elucidation and assignment of proton and carbon resonances 1D- ^1H , 1D- ^{13}C , DQF-COSY, TOCSY (mixing time 80 ms), ROESY (mixing time 150 ms), multiplicity edited-HSQC, and HMBC spectra were acquired. ^1H -chemical shifts were referenced to sodium-3-(Trimethylsilyl)propionate-2,2,3,3- d_4 . ^{13}C -chemical shifts were referenced to the solvent signal ($\text{DMSO-}d_6$: $\delta_{\text{C}} = 39.52$ ppm).

NMR spectra of natural isolated fractions of the Cryopeptides were measured on a Bruker Avance III 600 MHz spectrometer (600 MHz for ^1H , 151 MHz for ^{13}C) equipped with a Prodigy cryoprobe (Brucker, Ettlingen, Germany). Chemical shifts are reported in ppm and were referenced to the corresponding residual solvent signal ($\text{DMSO-}d_6$: $\delta_{\text{C}} = 39.52$ ppm, $\delta_{\text{H}} = 2.50$ ppm; CD_3OD : $\delta_{\text{C}} = 49.00$ ppm, $\delta_{\text{H}} = 3.31$ ppm) and interpreted by Dr. Yang Liu.

4. Chapter 1 – Globomycin

This work was carried out in collaboration with the industry partner Sanofi and Evotec, particularly the colleagues Hans Matter and Armin Bauer from Sanofi and Frédéric Jeannot und Pierre Despeyroux from Evotec.

My main contributions to the project are:

- Coordination of the project
- Discussions for the rational design approach and the results of the SBDD / molecular docking (with H. Matter and A. Bauer)
- Identifying MS results of molecular network (with M. Patra)
- Co-development of the synthetic strategies (with S. Schuler and Evotec chemists)
- Syntheses and analytics of derivatives

4.1 Introduction

Infectious diseases caused 20% of deaths worldwide in 2017.¹¹ Even if the mortality caused by infectious diseases is slowly decreasing, the threat of multi-drug resistance is growing.^{12, 13} The WHO addresses the urgency in a list of “Global Priority Pathogens” that include 12 different species.¹⁴ The acronym *ESKAPE* summarizes the six most important ones:

¹¹ Hannah Ritchie, Max Roser, **2018** - "Causes of Death". Published online at OurWorldInData.org. Retrieved from: <https://ourworldindata.org/causes-of-death> [last accessed: 26th March 2021].

¹² <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>, [last accessed: 26th March 2021].

¹³ CDC. “Antibiotic Resistance Threats in the United States”, 2019. Atlanta, GA: U.S. Department of Health and Human Services, CDC; **2019**.

¹⁴ WHO. Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics, **2017**. Can be found under: http://www.who.int/medicines/publications/WHO-PPL-Short_Summary_25Feb-ET_NM_WHO.pdf [last accessed: 26th March 2021].

- E - *Enterococcus faecium* (Gram-positiv)
- S - *Staphylococcus aureus* (Gram-positiv)
- K - *Klebsiella pneumoniae* (Gram-negative)
- A - *Acinetobacter baumannii* (Gram-negative)
- P - *Pseudomonas aeruginosa* (Gram-negative)
- E - *Enterobacter* (Gram-negative)

Antibiotic resistance is a natural evolutionary process, be it an acquired response on environmental change, a coincidental mutation or the intrinsic property of bacteria to be resistant.¹⁵ The mis- and overuse of antibiotics is further accelerating the development of resistances in organisms. Besides the almost instant development of resistances, a lack of new and novel compounds is aggravating the situation.¹⁶ Only 17 new antibiotics were approved and marketed in the US between 2000-2018.¹⁷ In December 2020, 43 antibiotics were in clinical development. Nevertheless, none represents a novel class or a novel mode of action against Gram-negative ESKAPE pathogens.¹⁸ New antibiotics, especially novel antibiotic classes and novel modes of action are severely needed to treat infectious diseases, especially against Gram-negative bacteria.

As Penicillin, one of the first antibacterial natural products, was found in nature,¹⁹ it is not surprising to utilize natural products as a source to discover and develop new antibiotics. Of 21 different classes of antibiotics, 17 are either natural products or derivatives thereof.¹⁷ The diversity and properties of natural products are a great source of novel structures and mode of actions. Over time, the discovery of novel natural products declined though, and the rediscovery of known compounds became more frequent.³ The broadening of the secondary metabolite spectra of examined species is one way to address this problem and the inclusion

¹⁵ J. M. A. Blair *et al.* *Nat Rev Microbiol*, **2015**, 13, 42–51.

¹⁶ G. Annunziato, *Int. J. Mol. Sci.* **2019**, 20, 5844.

¹⁷ M. Lakemeyer, W. Zhao, F. A. Mandl, P. Hammann, S. A. Sieber, *Angew. Chem. Int. Ed.* **2018**, 57, 14440–14475.

¹⁸ <https://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2021/03/tracking-the-global-pipeline-of-antibiotics-in-development>. [Accessed: 26th March 2021].

¹⁹ A. Flemming, *Br J Exp Pathol*, **1929**, 10 (3), 226-236.

of hard to cultivate species. Other alternatives are underexplored natural products, which were found in the past but their potential was not actively exploited due to limited technologies. Those known compounds offer a great starting point for developing new antibiotics and finding novel modes of action. Physico chemical examinations of natural compounds allow for addressing toxicity or stability problems early on. Computer based methods, such as structure based drug design (SBDD) or molecular docking can aid in the development of structural optimization.

Globomycin (**1.1**, GLM) and its naturally occurring congeners, as shown in Figure 4-2, represent a class of such underexplored natural products. This cyclic depsipeptide was first isolated in 1978 from four different *Actinomyces* strains and consists of five amino acids (AA-1 - AA-5) as well as a β -hydroxy fatty acid (BHA) as lipophilic side chain.²⁰ Three main properties make it an attractive starting point for developing a new antibiotic lead structure:

1. Gram-negative activity
2. A novel target, the lipoprotein signal peptidase II (LspA)
3. Chemical structure allows wide optimization

Its activity against Gram-negative bacteria²¹ and *Mycobacterium tuberculosis*²² addresses the most important field of antibiotic research. The Gram-negative activity of GLM originates from inhibition of the signal peptidase LspA, a transmembrane enzyme, which catalyzes the release of signal peptides from bacterial membrane pro-lipoproteins to apo-lipoproteins. Globomycin functions as a non-cleavable peptide, blocking the binding side of LspA and unprocessed pro-lipoproteins accumulate in the cytoplasmic membrane, resulting in cell death (Figure 4-1, highlighted in red).²³

²⁰ M. Inukai, R. Enokita, A. Torikata, M. Nakahara, S. Iwado, M. Arai, *J. Antibiot.* **1978**, *31*, 410-420.

²¹ T. Kiho, M. Nakayama, K. Yasuda, S. Miyakoshi, M. Inukai, H. Kogen, *Bioorg. Med. Chem.* **2004**, *12*, 337-361.

²² N. Banaiee, W. R. Jacobs, J. D. Ernst, *J. Antimicrob. Chemother.* **2007**, *60*, 414-416.

²³ M. W. Weichert *et al.* *Nat Commun.* **2017**, *8*, 15952.

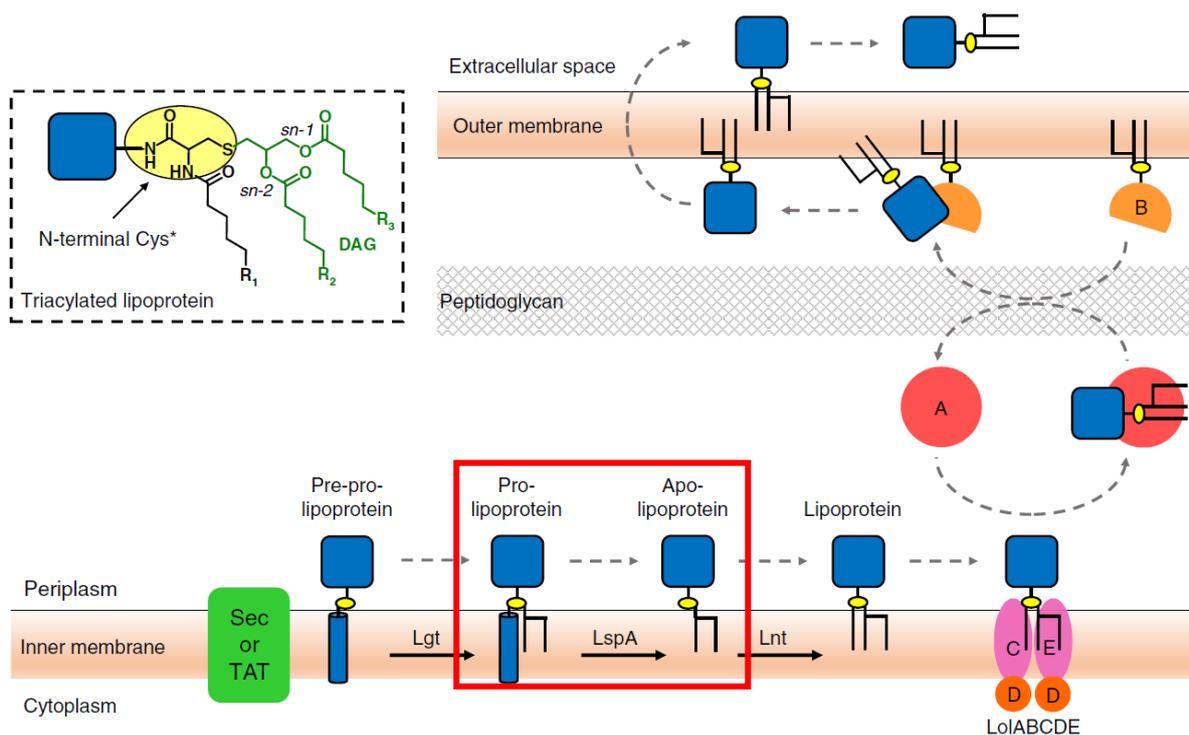


Figure 4-1: Lipoprotein posttranslational processing in Gram-negative bacteria.²³ The by Globomycin (**1.1**) targeted enzyme LspA and its role in the process is highlighted in red. GLM functions as a non-cleavable peptide for LspA and pro-lipoproteins accumulate in the membrane, resulting in cell death.

LspA offers two advantages as an antibiotic target: i) it is exclusively found in bacteria, therefore a low risk of target-based adverse effects in humans can be expected and ii) it has not been addressed by marketed antibiotics yet. Based on the cyclic peptide structure of **1.1**, structural modifications can easily be implemented. The alkyl side chain could be varied in length and shape and different functional residues can be added. The amino acids can be exchanged for other natural or non-natural amino acids or for mimetics. This can aid in the development of a more profound Structure-Activity Relationship (SAR) and improvement of physicochemical properties like metabolic liability, log D values and overall ADME qualities.^{24, 25}

²⁴ F. Hoffmann-La Roche AG, Genentech, Inc., Cyclic Peptide Antibiotics, WO2019/052545 A1, **2019**.

²⁵ K. Garland *et al.* *Bioorg Med Chem Lett*, **2020**, 30 (20), 127419.

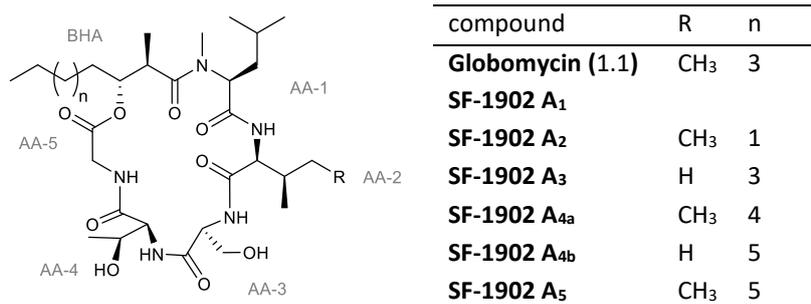


Figure 4-2: Globomycin (1.1) and its literature known naturally occurring congeners.²⁶

The published synthetic approaches (see Figure 4-3 and Table 4-1) differed in two main aspects, namely the ring closure and the stereoselective synthesis of the lipophilic side chain. Sarabia's *et al.* synthesis of Globomycin was a linear approach and focused on solid phase peptide synthesis (SPPS) and macrolactonization after the peptide is cleaved of the resin as the ring closing step. Furthermore, the synthesis of the β -hydroxy acid was explored in two different approaches, both complex multi-step preparations with a moderate overall yield.²⁷ Kiho and co-workers pursued a convergent synthesis based on three fragments. The di- and tripeptide fragments were synthesized in the liquid phase and the ring closing step was based on macrolactamization between serine and threonine. The stereoselective synthesis of the lipophilic side chain was examined in three ways, where the most promising two were three step procedures with high overall yields. One focused on Evans chemistry and the other on a norephedrine based auxiliary to introduce the stereochemistry.²⁶ Both, the Sarabia group as well as Kihos group relied on different protecting groups for each of the hydroxyl groups of threonine and serine, see Table 4-1 and Figure 4-3.

²⁶ T. Kiho, M. Nakayama, K. Yasuda, S. Miyakoshi, M. Inukai, H. Kogen, *Bioorg. Med. Lett.* **2003**, *13*, 2315-2318.

²⁷ F. Sarabia, S. Chammaa, C. García-Ruiz, *J Org Chem*, **2011**, *76* (7), 2132-2144.

Table 4-1: Comparison of Sarabia and Kihos synthetic strategies.^{26, 27}

	Kiho <i>et al.</i>	Sarabia <i>et al.</i>
Style	convergent	linear
Peptide synthesis	liquid; di- and tripeptide fragments	solid phase, linear peptide
Ring closing step	macrolactamization between AA-4 and AA-3	macrolactonization
Protecting groups	Thr(TBS), Ser(Bn)	Thr(TBS), Ser(Bn)
BHA	Evans auxiliary, Norephedrine auxiliary, Mitsunobu based approach	Sulfur ylides Sharpless asymmetric epoxidation

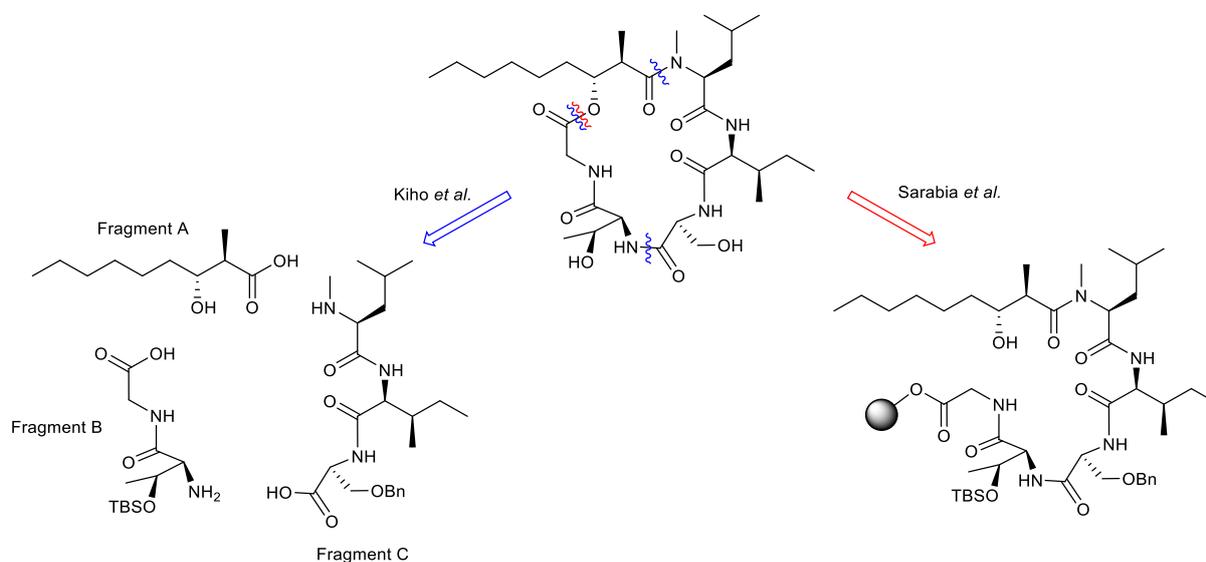


Figure 4-3: Simplified retrosynthetic approach of Kihos's (blue) and Sarabia's (red) work to Globomycin (1.1).^{26, 27}

Kiho *et al.* published activity results of GLM analogues in which they showed initially the importance of the lipophilic side chain and its length, the need for a methylation of the nitrogen on AA-1, and the importance of the serine in 2004.²¹ Besides this first SAR data,

Vogeley *et al.* presented the crystal structure of LspA in complex with Globomycin. This 3D model offers a deeper understanding of molecular interactions and allows for a structure-based optimization approach of Globomycin as a potential lead structure. It identifies LspA as an aspartyl peptidase, with Asp124 and Asp143 located in the catalytic center, which is occupied by the serine of Globomycin.²⁸

4.2 Results and Discussion

The results of this work are divided into four sub chapters, focusing on the following main aspects, with final antibacterial activity tests to conclude the work:

1. Examining the in-house metabolomic library for a producer strain and natural Globomycin analogues.
2. Determining physicochemical properties (ADME / MedChem) and addressing stability problems early on.
3. Evaluating new structural elements and derivatives utilizing structure-based drug design and molecular docking experiments.
4. Combining the advantages of the two known synthetic routes for an easier and more robust synthetic procedure.

4.2.1 Molecular Networking and Producer Strain

To identify natural Globomycin producers in our strain collection, a reverse database search was performed, based on the *m/z* value of the proton adducts and their retention time, referencing to the commercially available Globomycin authentic standard. Using this approach, a producer strain (*Streptomyces sp.* HAG010519) was identified in the in-house metabolomic database of Fraunhofer by Dr. Maria Patras, and it produced Globomycin, SF-1902 A_{2a} and A_{3a} in high titers. Based on the existing data of the producer strain, together we performed a molecular network analysis (see Figure 4-4), which revealed a larger group

²⁸ L. Vogeley *et al.* *Science*, **2016**, 351 (6275), 876-880.

of derivatives (marked in blue) encompassing the congeners SF-1902 A2 - A5 reported in literature (marked in red, Figure 4-4) and new ones (marked in green).

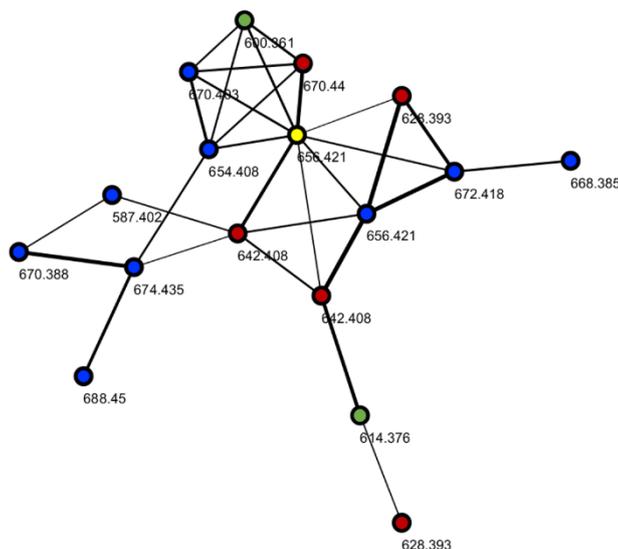
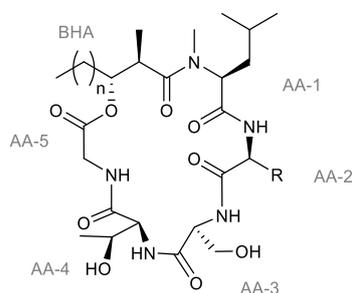


Figure 4-4: Molecular Networking Cluster of the Globomycin family of structural related compounds. Highlighted in yellow is the precursor ion corresponding to Globomycin (**1.1**), in red literature known compounds SF-1902 A2 - A5 and in green novel found congeners that have a similar structure to the known ones. Shown in blue are structural derivatives.

In the following, congeners refer to compounds with the highest structural similarity to Globomycin and the literature known SF-1902 A2 - A5, whereas derivatives differ in a greater structural way from the mother compound e.g., containing different amino acids. In the extract we found 20 new congeners based on the molecular network (nodes in green) and specific searches. For that we deduced structures, derived of Globomycin and its congeners, and searched specifically for them in the chromatogram of the extract. For seven we proposed a structure based on the comparison of their fragmentation pattern to that of Globomycin, shown in Figure 4-5. These congeners differ only in the length of the lipophilic side chain and the residue of AA-2.

The remaining nine compounds present in the molecular network (marked blue) are derivatives with greater structural differences compared to GLM. For four of them we proposed structures based on the fragmentation pattern, mostly modifications of the lipophilic side chain.



compound	AA-2	n	MF
Globomycin (1.1)	<i>allo</i> -Ile	5	C ₃₂ H ₅₇ N ₅ O ₉
SF-1902 A₁			
SF-1902 A_{2a}	<i>allo</i> -Ile	3	C ₃₀ H ₅₃ N ₅ O ₉
SF-1902 A_{2b}	Val	4	C ₃₀ H ₅₃ N ₅ O ₉
SF-1902 A_{3a}	Val	5	C ₃₁ H ₅₅ N ₅ O ₉
SF-1902 A_{3b}	(<i>allo</i>)-Ile	4	C ₃₁ H ₅₅ N ₅ O ₉
SF-1902 A_{4a1}	<i>allo</i> -Ile	6	C ₃₃ H ₅₉ N ₅ O ₉
SF-1902 A_{4a2}	Ile	6	C ₃₃ H ₅₉ N ₅ O ₉
or			
SF-1902 A_{4c}	Leu	6	
SF-1902 A_{4b}	Val	7	C ₃₃ H ₅₉ N ₅ O ₉
SF-1902 A₅	<i>allo</i> -Ile	7	C ₃₄ H ₆₁ N ₅ O ₉
SF-1902 A_{6a}	Val	3	C ₂₉ H ₅₁ N ₅ O ₉
SF-1902 A_{6b}	(<i>allo</i>)-Ile	2	C ₂₉ H ₅₁ N ₅ O ₉
SF-1902 A_{7a}	(<i>allo</i>)-Ile	1	C ₂₈ H ₄₉ N ₅ O ₉
SF-1902 A₈	Val	1	C ₂₇ H ₄₇ N ₅ O ₉

Figure 4-5: Structure of Globomycin (1.1), its natural occurring congeners and the novel found isomers (marked in blue). Proposed structures are based on the comparison of their fragmentation pattern to that of GLM. The stereogenic centers are based on GLM and needs to be confirmed by NMR or crystallography.

For the remaining compounds either the intensity was too low for fragmentation or it was overlapping with other fragmentation patterns. Detailed MS/MS data and chromatograms can be found in the Supporting Information 4.5.1, as well as a detailed structural overview of all congeners (Figure S 4-1, Figure S 4-2).

Our proposed structures are based on the comparison of the fragmentation pattern of new molecules to that of GLM. The specificity of the fragmentation pattern and the distinct fragmentation order of Globomycin (Figure S 4-3) allows for a reliable structural assignment and regiochemical interpretation of the new compounds. Since all derivatives and congeners are very likely synthesized by the same Non-Ribosomal Peptide Synthase (NRPS) gene cluster, it can be assumed that the stereochemistry is conserved. However, this hypothesis needs to be confirmed by structure elucidation *via* NMR or crystallography. In the end, such an approach provides access to naturally occurring GLM analogues.

4.2.2 Early ADME tests on Globomycin

Globomycin (**1.1**) was reported as a Gram-negative antibiotic with a minimal inhibitory concentration (MIC)-value of 6.25 µg/mL for *E. coli* SANK 70569.²⁹ In cooperation with Sanofi we determined MIC values of 8 and 8-16 µg/mL for the *E. coli* strains ATCC 35218 and ATCC 25922, respectively. Our early ADME tests as well as a physicochemical evaluation of Globomycin indicate it to be a promising antibacterial lead structure. Overall, the natural compound shows drug like properties based on its log D value (pH 7.4: 3.95), its solubility of 0.215 mg/mL (phosphate buffer at pH 7.4) and its moderate CYP3A4 inhibition. Its low permeability (in CaCo-2 model of intestinal resorption) can be neglected for an intravenous application. It has a high chemical stability (100%, phosphate buffer pH 7.4, 24 h at 25 °C) and is stable in simulated gastric, intestinal fluids, and in human plasma. The only liability observed from these early *in vitro* eADME studies is a high metabolic liability in human, mouse, and rat liver microsomes, which needs to be addressed for future derivatives. The main metabolite we found in the *in vitro* metabolism study of human liver microsomes indicates a cleavage between AA-3 and AA-4 (Ser and Thr) by amide hydrolysis. Two more metabolites were determined to be hydroxylated at AA-1 (*N*-Me-Leu) and the lipophilic side chain.

4.2.3 Structure Based Drug Design (SBDD) and Molecular Docking

To address the liability of the natural substrate based on the metabolites found in the eADME test, the focus was put on AA-4 and AA-3 as well as the lipophilic side chain. Figure 4-6 shows Globomycin in complex with LspA (PDB: 5DIR). Favorable hydrogen bonds (yellow) are formed between the hydroxyl group of serine and Asp 124 and Asp143, the catalytic dyad, as well as from various carbonyl groups to Arg 116 and Asn 112.

²⁹ Y. Xiao, K. Gerth, R. Müller, D. Wall, *Antimicrob. Agents Chemother.* **2012**, 56, 2014-2021.

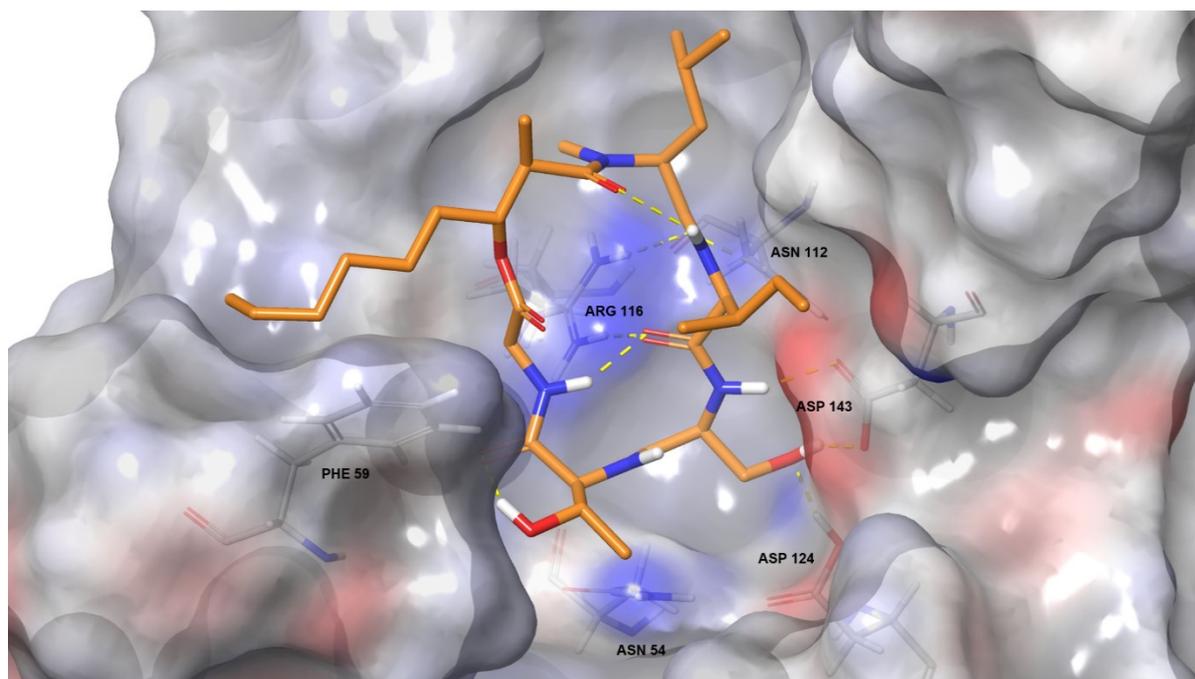


Figure 4-6: Globomycin (in orange) in complex with LspA (PDB 5DIR). Favorable hydrogen bonds are depicted in yellow.

To conserve the residues Asp 124 and Asn 143 of the catalytic center and maintain the interactions between it and serine, we focused on introducing favorable interactions to Asn 54. Therefore, we introduced hydroxyprolines in the place of the *allo*-threonine of GLM, further representing also a complement to the large number of so far described GLM analogues.^{24, 25} The assumption was, that additional hydrogen bonds could be formed and the pocket could be filled out in a better way, due to the steric room of the five-membered ring. To test the optimal position of the hydroxyl group, we exploited 3- and 4-hydroxyproline. Knowing the importance of the lipophilic side chain and the structure of the enzyme, our idea was to utilize Phe 59 for aromatic interactions, specifically sandwich π - π -stacking. During the docking process the most promising derivatives were chosen, based on synthetic availability and docking scores, as presented in Table 4-2. Graphic results of the molecular docking can be found in more detail in the Supporting Information 4.5.2.

Table 4-2: Results of molecular docking and SBDD.

structure	Docking Score [kcal/mol]	Docking Score [kcal/mol]
 Globomycin (1.1)	-99.833	
 1.2	-104.244	
 1.3	-107.569	
 1.4		-108.369
 1.5		-104.718

Based on the docking score of -99 kcal/mol for GLM the results shown in Table 4-2 indicate a higher activity than Globomycin for all derivatives. To study the effect of structural modifications, we only implemented one change for the derivatives **1.2** and **1.3**. The *allo*-threonine was exchanged to 4-Hyp and 3-Hyp, respectively, whereas the length of the lipophilic side chain was kept identical to GLM. Based on the docking scores, the 3-Hyp residue of **1.3** (-107.6 kcal/mol) has a greater effect than the 4-Hyp residue of **1.2** (-104.2 kcal/mol). This is also the case for the comparison of **1.4** and **1.5**, both containing an aromatic side chain but different hydroxyprolines, where the docking score is again higher for the 3-Hyp derivative **1.4**. In order to examine the replacement of the alkyl residue, we introduced an

ethane-phenyl substituent as an aromatic linker to study the effect of an aromatic moiety. Based on the results of the molecular docking, this length seems to be ideal to position the phenyl close enough to Phe 59 for favorable π - π interactions. Comparing compounds **1.2** and **1.5**, which only differ in the type of linker, they show similar docking scores in the range of -104 kcal/mol, suggesting that the effect of the side chain on the affinity is low. This is confirmed by comparing the docking scores of derivatives with longer chains, which vary only marginally. Detailed pictures of the molecular docking of the four chosen structures can be found in the Supporting Information (4.5.2).

4.2.4 Synthesis

Based on the aforementioned findings, a total synthetic route was favored for the four target compounds. In cooperation with Evotec, we planned our synthesis adopting the advantages of SPPS of Sarabia *et al.* and the preparation of the lipophilic side chain based on Kiho's work. The differences and similarities of both synthetic strategies are summarized in Table 4-1 and were discussed in section 4.1. Figure 4-7 exemplifies on Globomycin the retrosynthetic approach of Kiho (blue), Sarabia (red) and the newly developed one in green. Our first tries adopting the macrolactonization procedure of Sarabia *et al.* resulted in very poor yields, so we switched our focus on macrolactamization to close the ring instead. To avoid unnecessary deprotection steps, we implemented a universal protecting group for the hydroxyl functions of amino acid building blocks, simplifying the overall synthesis. Since our first trial of using *t*Bu failed due to deprotection issues, we choose TBS as the universal protecting group for further synthesis. This new synthetic approach combines the advantages of both literature known strategies in one easy and fast procedure, simplifying it further by using only one instead of two different protecting groups for the hydroxyl functions of the amino acids.

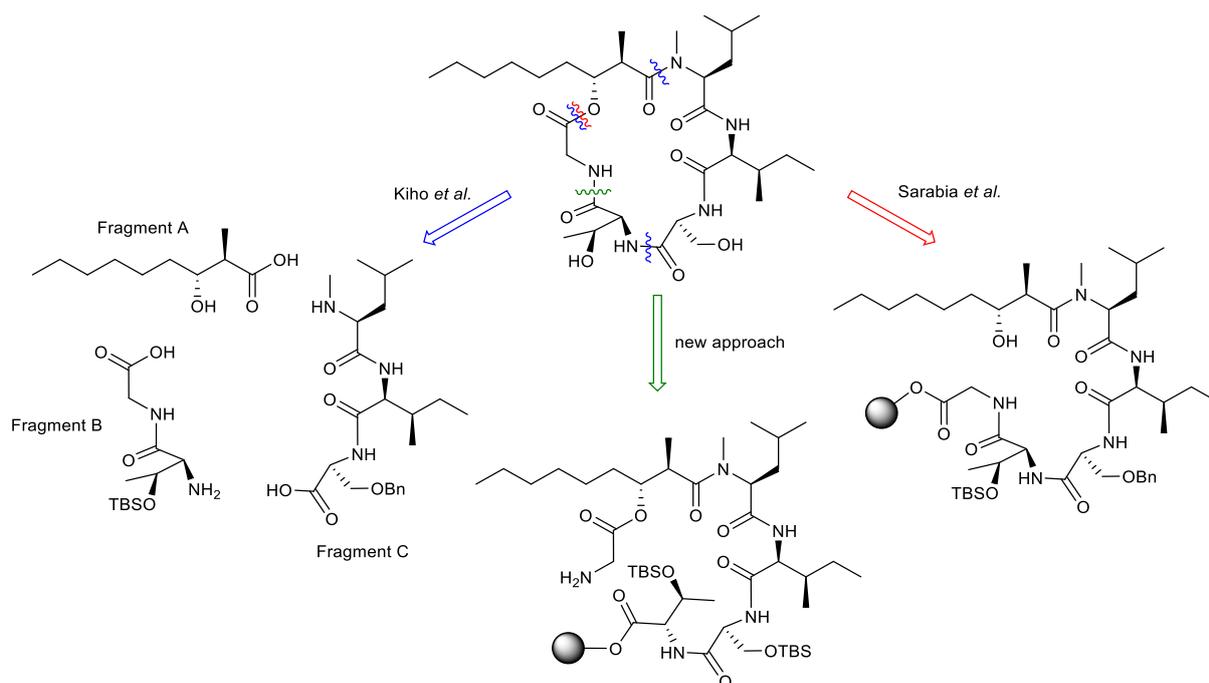
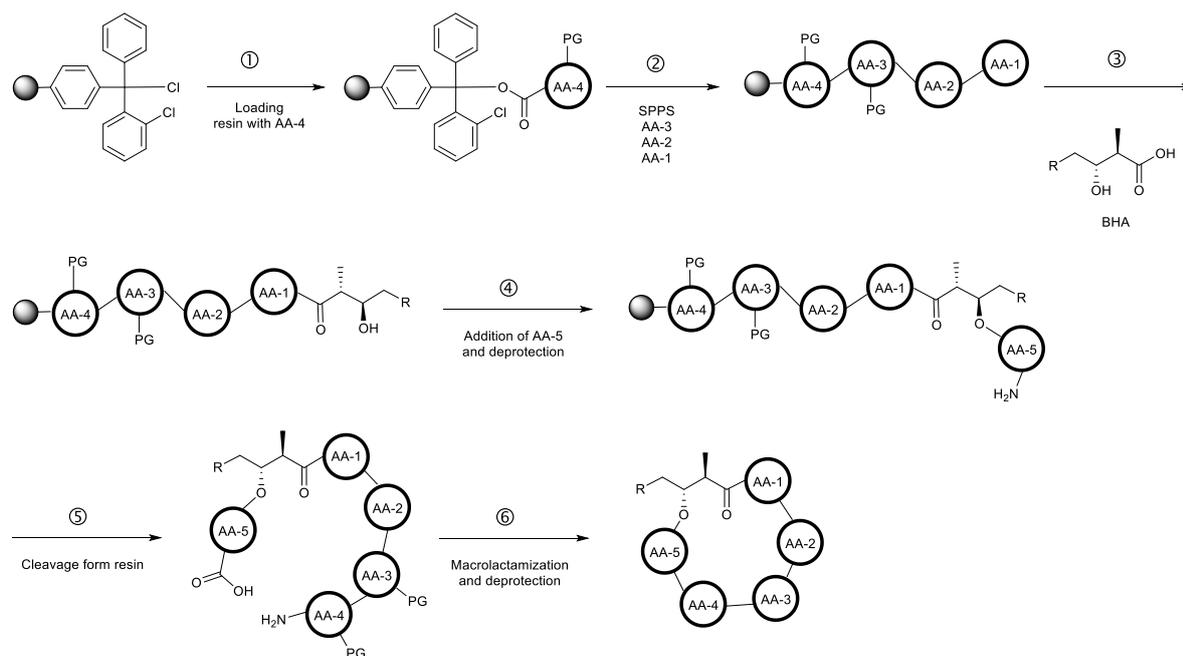


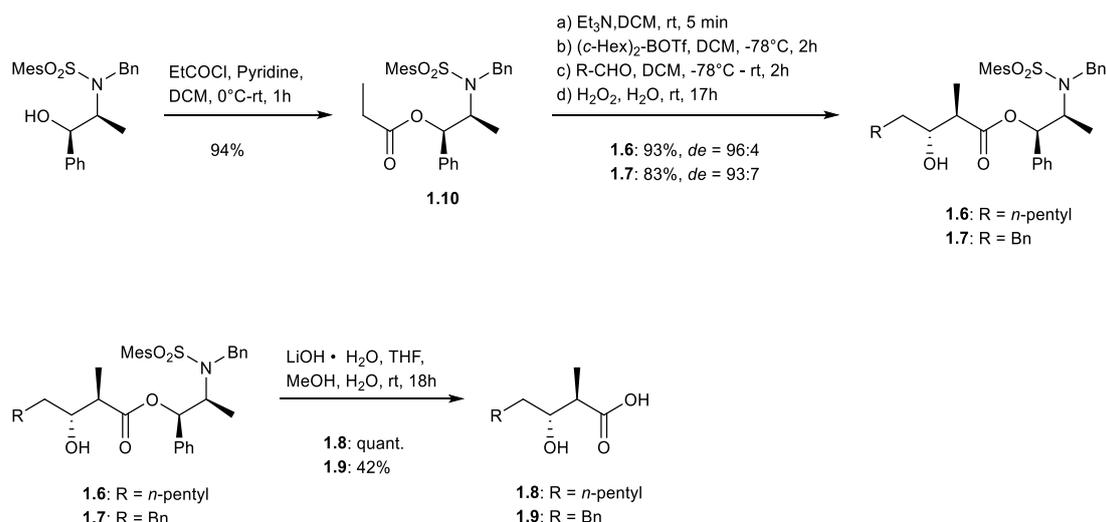
Figure 4-7: Simplified retrosynthetic approach of Kihos (blue) and Sarabias (red) work, as well as the newly developed retrosynthetic approach (green), exemplified on Globomycin.

The generalized Scheme 4-1 gives an overview of the newly developed route. Our strategy can be segmented into six parts: (1) Esterification of the first amino acid to the resin; (2) coupling of amino acids and Fmoc-deprotection; (3) connecting the β -hydroxy acid (BHA); (4) attaching protected glycine and deprotection; (5) cleavage from the resin; (6) ring closure *via* macrolactamization and final deprotection of the hydroxyl groups. Disregarding the Fmoc-deprotection steps, six of nine steps are carried out on solid support, making it an easy and especially fast procedure, avoiding multiple purifications of intermediates.



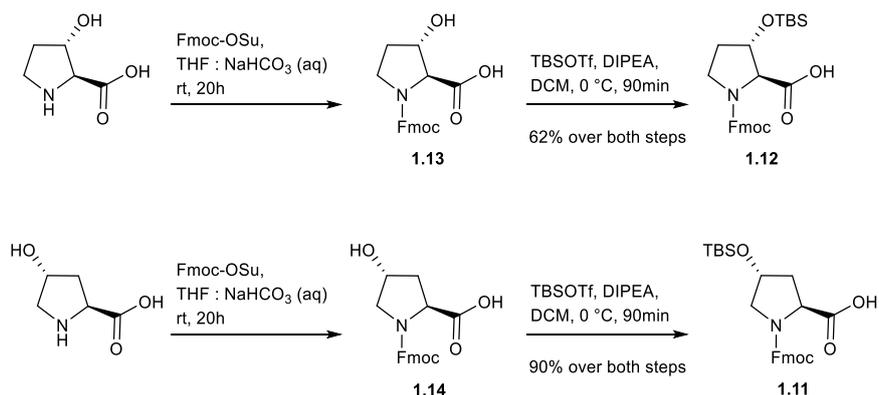
Scheme 4-1: Generalized synthesis of Globomycin and derivatives. It can be divided in six parts. Over 60% of the steps are carried out on solid support, eliminating purifications for intermediates allowing for a fast and easy procedure.

For the overall synthesis the required building blocks, consisting of TBS-protected 3- and 4-hydroxyproline and the chosen β -hydroxy acids, needed to be prepared beforehand. Preparing the BHAs we tried using the Evans as well as the Norephedrine auxiliary based on the procedures by Kiho et al.²⁶ Due to unexpected problems during the hydrolysis of the Evans-aldol product, we chose the norephedrine route, as depicted in Scheme 4-2. For the aldol addition this method offered high diastereoselectivity of 93:7 and 96:4 for the *n*-alkyl (**1.6**) and the aromatic linker (**1.7**), respectively. However, the yields of both compounds over the aldol addition and hydrolysis varied greatly. For the *n*-alkyl linker (**1.8**) a yield of 93% was achieved. The aromatic linker (**1.9**) was obtained in just 34% over the same steps. Especially the hydrolysis with just 42% compared to the quantitative conversion of the alkyl linker affects the overall yield. Part of that could be attributed to the purification *via* crystallization, which was not optimized further. After hydrolysis of the anti-aldol product the norephedrine auxiliary was recovered in yields of 77-89%.



Scheme 4-2: Synthesis of lipophilic side chains based on norephedrine auxiliary.

For the hydroxyproline derivatives, the Fmoc- and TBS-protected amino acid building blocks were synthesized as shown in Scheme 4-3 following standard literature procedures.³⁰ The synthesis was done once in a 6-8 g scale for both building blocks and furnished the desired products in moderate to high yields over both steps. Synthesis of both, Fmoc- and TBS-protected 4-Hyp (**1.11**) could be performed with yields of 90%, whereas the protected 3-Hyp (**1.12**) building block could be obtained only with a yield of 62% over both steps. This could be attributed to steric hindrance due to the proximity of the hydroxy group to the COOH group and the formation of an intramolecular hydrogen bond.



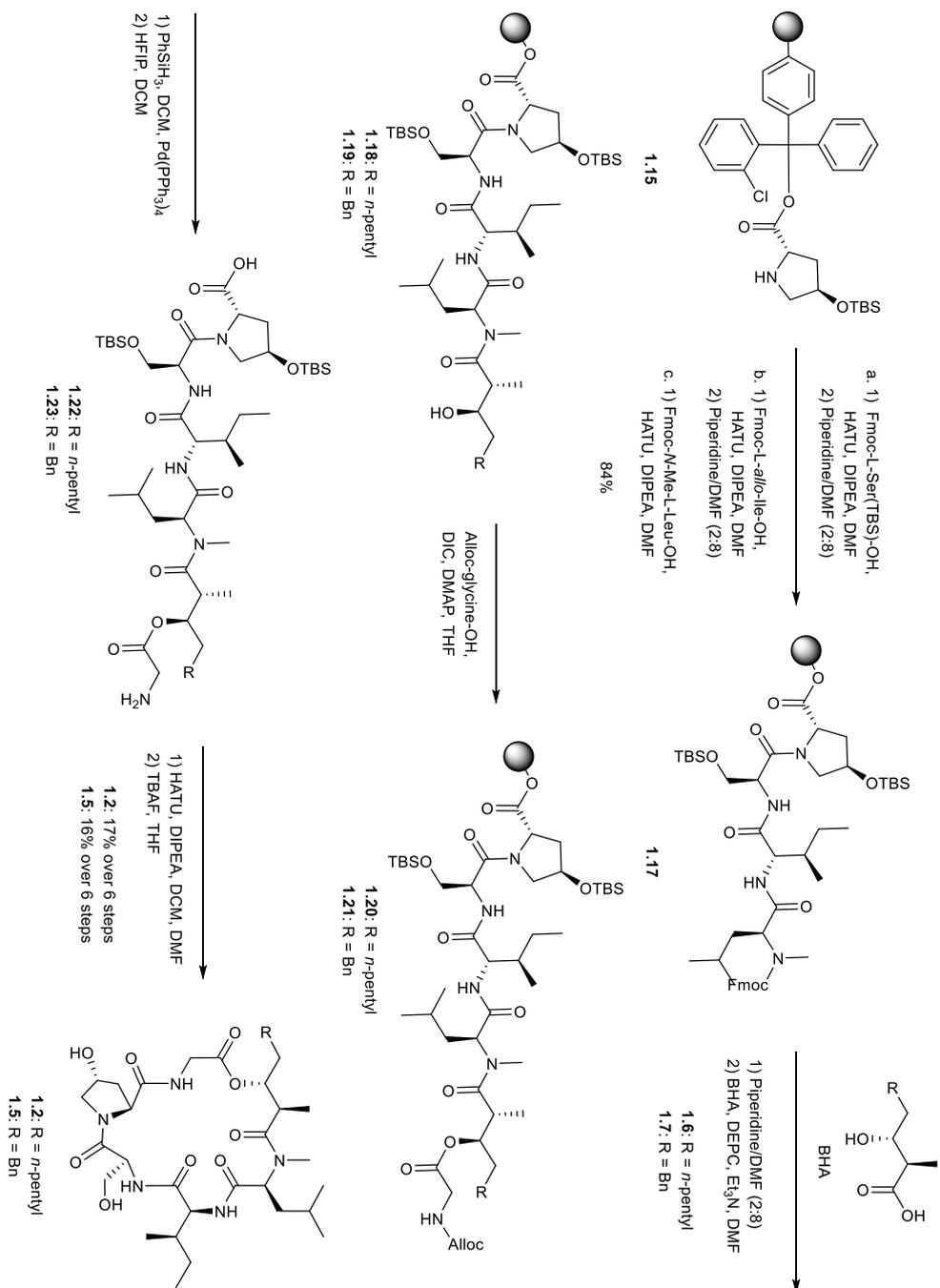
Scheme 4-3: Fmoc- and TBS-protection of 3- and 4-hydroxyproline building blocks.

³⁰ A. Agarkov, S. J. Greenfield, T. Ohishi, S. E. Collibee, S. R. Gilbertson, *J. Org. Chem.* **2004**, *69* (23), 8077–8085.

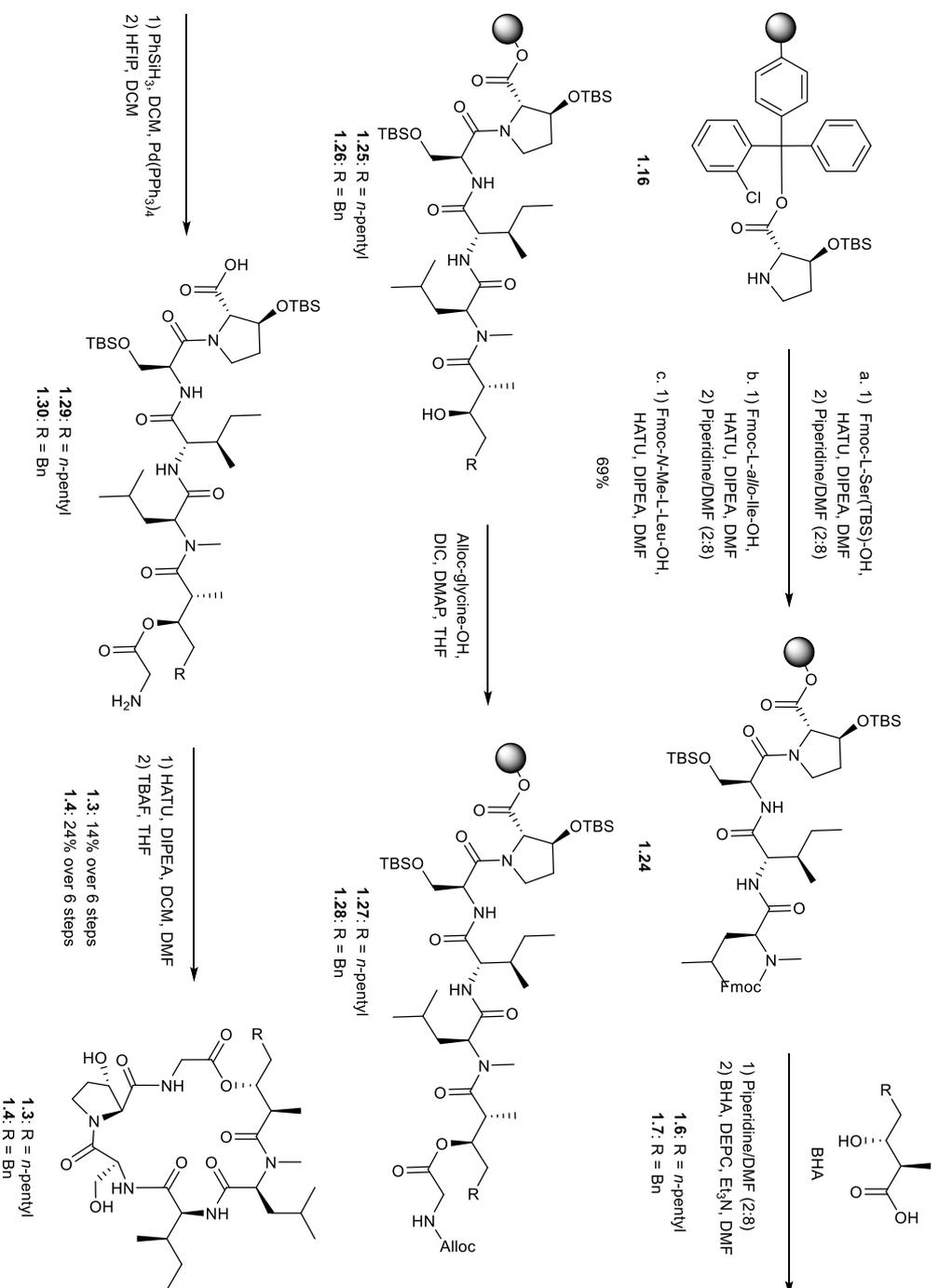
The 2-Chlorotrityl chloride (2-CTC) resin was chosen because of its hyper-acid lability for SPPS. Esterification of the resin with the first amino acid followed standard procedures and the loading was determined by UV/VIS-absorption for each prepared resin.³¹ Since the samples needed to be dried *in vacuo* first and were then stored under argon overnight, the determined loading of the resin could have been a source of error for following calculations. In the presence of the Fmoc-protecting group amino acids and short peptides bound to the 2-CTC resin are very prone to cleavage.³¹ Based on the loading of the purchased 2-CTC resin, the yield was 31% for the 4-Hyp variant (**1.15**) and 23% for the 3-Hyp variant (**1.16**). Those are very low yields for a standard procedure, which is usually known for high yields. Besides an incorrect determination of the loading, steric hindrance of the TBS-protected hydroxyproline could have been a factor for it. This would explain the even lower yield of the 3-Hyp resin where the bulky TBS-protecting group is positioned closer to the resin compared with the 4-Hyp derivative. Following coupling steps based on the estimated loading were successful, so our calculated load seemed appropriate.

The synthesis of **1.2** and **1.5** is depicted in Scheme 4-4. The peptide chain of **1.2** and **1.5** are identical, allowing for a split approach for the synthesis. Based on the determined loading of **53**, the overall yield, for a total of nine steps, not taking the Fmoc-deprotection steps into account, for **1.2** is 14% and for **1.5** it is 13%. The coupling of the first three amino acids was achieved with 84%. The following three steps were carried out on solid support in the same vessel to avoid losing too much material during a transfer. Therefore, we did not determine the yields of those intermediates. Ring closure was directly performed after cleavage of the cyclization precursor from the resin. The protected derivatives were only prepurified by column chromatography before the final deprotection step.

³¹ W. Chan, P. White, *Fmoc Solid Phase Peptide Synthesis Practical Approach*, Oxford University Press, Oxford, 1999.



Scheme 4-4: Synthesis of 1.2 and 1.5. Over 60% of the synthetic steps were carried out on solid support. For **1.2** an overall yield of 14% and for **1.5** 13% was achieved. The ring closure, as one of the final steps, is the macrolactamization of glycine and 4-hydroxyproline. All reactions were carried out at room temperature, detailed conditions can be found in the experimental part (4.4.2).



Scheme 4-5: Synthesis of 1.3 and 1.4. Over 60% of the synthetic steps were carried out on solid support. For **1.3** an overall yield of 10% and for **1.4** 17% was achieved. The ring closure, as one of the final steps, is the macrolactamization of glycine and 3-hydroxyproline. All reactions were carried out at room temperature. Detailed conditions can be found in the experimental part (4.4.2).

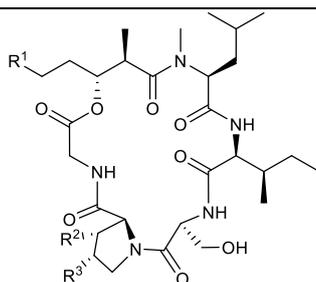
For the derivatives **1.3** and **1.4** (see Scheme 4-5) the overall yield overall steps is 10% and 17%, respectively. The SPPS of the first three amino acids and Fmoc-deprotection is slightly lower with just 69%. The following six steps with yields of 14% and 24% are comparable with the ones from **1.2** and **1.5** with 17% and 16% over the same steps. Estimations based on the prepurified protected derivatives range from 29% to 51% for the deprotection step itself. Considering those, the TBS deprotection and final purification needs to be optimized for further derivatives. Additionally, the determined loading of the attachment of the hydroxyprolines to the resin needs to be considered. The sensitivity of the amino acid-resin bond as long as the Fmoc-group is attached and its storage overnight could have led to an error-prone load, which in turn affects the calculation of the overall yield. During the initial synthesis we focused on the timely delivery of a variety of Globomycin derivatives to establish a SAR. Therefore the yields and synthesis processes were not optimized at this stage.

4.2.5 Antibacterial Activity

Globomycin and the synthetic derivatives **1.2**, **1.3**, **1.4** and **1.5** were tested against a panel of Gram-negative bacteria (*E. coli*, *P. aeruginosa*, *K. pneumoniae* and *A. baumannii*) as well as Gram-positive (*S. aureus*) and a surrogate of *M. tuberculosis* (i.e., *M. smegmatis*) (Table 4-3). Besides the activity for *E. coli* ATCC 25922 $\Delta hldE$, which was performed by Evotec (Sebastien Coyne, PhD and team), all other screenings were performed at the Fraunhofer Institute (Dr. M. Marner and team). By internal standards, activities of $\geq 128 \mu\text{g/mL}$ are considered inactive, 32-64 $\mu\text{g/mL}$ are weakly active, 4-16 $\mu\text{g/mL}$ are moderately active and $\leq 2 \mu\text{g/mL}$ are highly active.

Table 4-3: Antimicrobial activity of Globomycin (1) and derivatives.

	MIC [$\mu\text{g/mL}$]				
	Globomycin	1.2	1.3	1.4	1.5
<i>E. coli</i> ATCC 25922 wt	16	>128	16-32	64	>128
<i>E. coli</i> ATCC 25922 ΔtolC	0.5	1	0.5	1	2
<i>E. coli</i> ATCC 25922 MHC	8-16	16	2	8	32
<i>E. coli</i> ATCC 25922 ΔhldE	0.5	4	1	4	8
<i>P. aeruginosa</i> PA01	>64	>64	>64	>64	>64
<i>P. aeruginosa</i> PA0750	64	64	64-32	>64	>64
<i>K. pneumoniae</i> ATCC 30104	64	>64	>64	>64	>64
<i>A. baumannii</i> ATCC 19606	>64	>64	>64	>64	>64
<i>S. aureus</i> ATCC 33592	>64	>64	>64	>64	>64
<i>M. smegmatis</i> ATCC 607	>64	>64	>64	>64	>64
Docking Score [kcal/mol]	-99.833	-104.244	-107.569	-108.369	-104.718



compounds	R ¹	R ²	R ³
1.2	<i>n</i> -butyl	H	OH
1.3	<i>n</i> -butyl	OH	H
1.4	Ph	OH	H
1.5	Ph	H	OH

Against all other bacteria except *E. coli* all compounds (**1.1** – **1.5**) were either not active or barely active. The derivatives **1.3** and **1.4**, as well as Globomycin (**1**) showed activity against all tested *E. coli* strains. For the wild type (*E. coli* ATCC 25922 wt) Globomycin was still the most active compound, followed by the 3-hydroxyproline compounds **1.3** and **1.4** with a slightly lower activity. The 4-hydroxyproline derivatives **1.2** and **1.5** showed no activity. In MHC the activity was significantly higher compared to the standard medium for all compounds except Globomycin itself, with **1.3** showing the highest activity. The same applies for the ΔhldE variant, where all tested compounds also showed significantly higher activities, with GLM having the best result, followed by the 3-hydroxyproline derivative **1.3**. Both, the ΔhldE variant and the MHC medium affect the stability of the (outer) membrane. The target of GLM, LspA, is located in the inner membrane of the cell (Figure 4-1). The higher activities of all compounds of the ΔhldE variant and the MHC variation, compared to the weaker activities against the wild type indicate that a permeability problem may be the cause. The results of the ΔtolC variant, where all derivatives are highly active, compared to those of the wild type, suggested that the

compounds are actively transported out of the cell by efflux-mechanisms. Again, GLM and **1.3** showed the best results with 0.5 $\mu\text{g}/\text{mL}$. Overall, compounds **1.2**, **1.3**, **1.4** and **1.5** showed moderate to high activity for the MHC and $\Delta hldE$ variants, with the 3-hydroxyproline, *n*-butyl derivative **1.3** being the most active one against all *E.coli* strains, representing a good lead structure for further optimization.

The comparison of the activity results to the docking scores shows a trend and assessments of derivatives is possible. The lower docking scores of **1.2** and **1.5** (~ -104 kcal/mol) and better ones for **1.3** and **1.4** (~ -108 kcal/mol) corresponded to the determined activities. As predicted by docking, the 3-Hyp derivatives **1.3** and **1.4** showed higher activity as the 4-Hyp ones (**1.2** and **1.5**). The docking scores predicted only a marginal effect of the aromatic residue on the activity, comparing **1.3** to **1.4** and **1.2** to **1.5**. The results showed a slight decrease of the activity for the phenyl moiety, comparing compounds **1.3** (16-32 $\mu\text{g}/\text{mL}$) to **1.4** (64 $\mu\text{g}/\text{mL}$). The hypothesis of favorable interactions of Phe 59 to the aromatic residue and π - π -stacking therefor did not work out. This shows the limitation of computer-based methods and that the results are an estimation and testing the compounds is indispensable.

4.3 Summary and Outlook

With just 17 antibiotics approved between 2000 - 2018 and the development of multi-drug resistance, the need for novel antibiotics is more pressing than ever. Natural products offer an abundant diversity of active compounds. They, and structures derived from them, make up over 80% of marketed antibiotics.¹⁷ Besides the strategy to find novel natural products, exploiting underdeveloped substances of the past and using their potential for optimization is another approach. Globomycin (**1.1**, GLM), first reported in 1978,²⁰ is such an underdeveloped substance and presents a very attractive lead structure to develop further based on the following attributes: i) its Gram-negative activity; ii) its structural setup which allows for modification and iii) its unique target LspA, which is exclusively found in bacteria and has not been exploited by other antibiotics yet. Additionally, the crystal structure of GLM in complex with LspA allows for a computer-aided approach,²⁴ like structure-based drug

design (SBDD). The high potential became apparent in 2020 when Garland *et al.* published their work about the optimization of Globomycin analogs with a variety of newly developed, synthesized and tested structures.²⁵

In search of new natural congeners of Globomycin for SAR expansion, we were able to find an in-house producer strain (*Streptomyces sp.* HAG010519). Using molecular networking and specific searches, we found a total of 29 structures related to Globomycin. For seven of 20 found new congeners and for four of nine newly found derivatives we proposed a structure. However, it turned out that all in this way identified derivatives were meanwhile described in patent literature.²⁴

Our early on approach of physiochemical tests to explore possible weaknesses, like stability and toxicity, is indispensable to address possible complications of future drugs as early as possible and to put the focus of the optimization process on where it is needed most. Early ADME tests revealed promising drug-like properties, but a high metabolic liability of the natural substrate (**1.1**). The amide bond cleavage between AA-3 and AA-4 (Ser and Thr) as a major metabolic pathway was identified. We used the published structure-activity relationship (SAR) of GLM,²¹ SBDD and molecular docking to evaluate the binding pocket and to develop new analogues. To address the high liability we exchanged the natural *allo*-threonine for hydroxyprolines, based on the assumption that the steric demand would slow down the amide hydrolysis between it and serine. This structural motif was also chosen based on a possible new hydrogen bond to Asn 54 and filling out the pocket in the catalytic cavity more due to the sterically demanding five-membered ring (Figure 4-6, Supporting Information 4.5.2). To investigate the effect of the lipophilic side chain, we introduced an aromatic residue for more favorable interactions, possible to Phe 59 (Figure 4-6, Supporting Information 4.5.2). The docking scores for the four chosen compounds all ranged higher than the one from GLM. The successful, newly developed synthetic route combines the advantages of the literature known ones,^{26,27} and improving it further with the use of only one protecting group (TBS) instead of the different ones used before (TBS and Bn) and less synthetic steps overall. It offers a high stereoselectivity for the synthesis of the fatty acid side chain, based on the norephedrine auxiliary, with a *dr* of 93:7 and 96:4 for the *n*-alkyl (**1.6**) and the aromatic linker (**1.7**), respectively. The synthesis of the protected amino acid building blocks is performed in just two steps with yields of 62% and 90% for 4-Hyp (**1.11**) and 3-Hyp (**1.12**), respectively,

allowing for an easy procedure to introduce more exotic amino acids. The first step, the attachment of the protected hydroxyprolines to the chosen 2-CTC resin, was not as successful as we expected. With yields of only 23% for the 3-Hyp variant (**1.16**) and 31% for the 4-Hyp variant (**1.15**) and problems during the determination of the loading due to the sensitivity of the amino acid-resin bond, this step needs to be optimized for future syntheses. Over 60% of the synthetic steps were carried out on solid support, making it a fast and easy procedure with the additional benefit of no need for the purification of intermediates. It implements a successful macrolactamization approach and gives moderate to good overall yields of 10-17% over just nine steps, not taking the Fmoc-deprotection steps into account. Our synthetic approach offers a robust and easily modified route to derivatives of Globomycin.

The four derivatives (**1.2** – **1.5**) only show noticeable activity against the tested *E. coli* strains (Table 4-3). Against the wild type, GLM (**1.1**) shows the highest activity with 16 µg/mL, followed by the 3-Hyp derivative **1.3** and **1.4** with 16-32 µg/mL and 64 µg/mL, respectively. The 4-Hyp derivatives **1.2** and **1.5** are not active. All derivatives show significant higher activity for the media variation with MHC and the $\Delta hldE$ and $\Delta tolC$ variants, suggesting a permeability problem and an active efflux mechanism that pumps the compounds out, especially since LspA, the target of Globomycin, is located in the inner membrane of the cell (Figure 4-1). Of all derivatives 3-hydroxyproline, *n*-butyl derivative **1.3** shows the highest activity. The aromatic linker of **1.4** has a negative effect on the activity compared to **1.3**, indicating that there are no interaction to Phe 59 as hoped for. Comparison of the activity results to the calculated docking scores show a good estimation and assessments of derivatives is possible, but it is no guarantee for active compounds and testing is indispensable.

Our results consist with literature and confirm that the strategy to utilize underexplored natural compounds is a highly promising way to new antibacterial drugs with novel mode of actions. The identifications of active natural compounds, literature known or from extracts, paired with early on characterization of ADME parameters and physicochemical properties allows to identify critical aspects for the optimization. These parameters can then be addressed for derivatives, aided by computer-based methods, such as SBDD and molecular docking to narrow down potential lead structures. The herein reported hydroxyproline-based GLM series represents a further diversification of the so far published SAR and the most promising derivative **1.3** can be used as new starting point for a refined SBDD approach. The

presented synthetic route based on SPPS allows for a robust, fast and easy access to new analogues.

To improve our synthetic approach further, the attachment of the amino acid to the resin needs to be optimized. First, it will need to be ascertained if it is the method itself that is responsible for the low yields, if it is due to the steric hindrance of the protected hydroxyprolines or the method to determine the loading of the resin. For the next circle of optimization, our hypothesis that the steric demand of the five-membered ring of the hydroxyprolines will slow down the amide hydrolysis between it and serine needs to be verified by further ADME tests of the synthesized compounds. The metabolic lability in human, mouse, and rat liver microsomes should be carried out for the derivatives, especially for the most promising one **1.3**. Comparison of the results to the ones of GLM will give insight to the stability and possible new weak points. A refined SAR, based on our results as well as considering the findings reported by Garland *et al.*,²⁵ for in-depth optimization of GLM, would offer more insight into the binding of the substrate and key interactions for future derivatization. A new circle of SBDD to find even better candidates would be followed by the synthesis and activity testing

4.4 Experimental

4.4.1 Molecular Networking analysis³²

The UHPLC-qTOF-MS/MS data of the Globomycin producer strain crude extracts were analyzed using molecular networking to allow the variable dereplication of known and unknown metabolites. First, the raw data (.d files) were converted to plain text files (.mgf) containing MS/MS peak lists using MSConvert (ProteoWizardpackage) wherein each parent ion is represented by a list of fragment mass/intensity value pairs (peak picking: vendor MS level = 1–2; threshold: absolute intensity, 1000, most intense). Molecular networking was performed following the established protocols using a cosine similarity cutoff of 0.7. Additionally, ions need a minimum of six shared fragments (tolerance Δ ppm 0.05) with at least one partner ion to be included in the final network. *In silico* fragmented compounds of a commercial database (AntiBase 2017) as well as our in-house reference compound MS/MS database were included in the network as reference substances to narrow down the molecular structure and to highlight compounds of interest. CytoScape was used to visualize the data as a network consisting of nodes and edges, wherein each node represents a parent ion. The edge width represents the cosine similarity score between nodes (thick edges indicate high similarity), and the size of the nodes the relative abundance of the ion in the extract.

4.4.2 Procedures for the Globomycin Syntheses

General Procedures

Coupling of the amino acids and fatty acids³¹

If not noted otherwise, all reactions were carried out in a custom-built solid phase peptide synthesis vessel with a G2 filter and a diameter of 3, 4 or 5 cm at room temperature. Argon

³² Special thanks to Maria Patras for performing the molecular network analysis and helping to identify and elucidate the new structural assignments of the related compounds.

was used for agitation of the resin. 20% Piperidine in DMF and the cleavage cocktail were freshly prepared on the day of use.

The Fmoc-protected amino acid or fatty acid (3 equiv) and HATU (2.9 equiv), dissolved in a small amount of DMF, were added to the swelled resin (1 equiv), followed by DIPEA (6 equiv) and the resin was agitated for 1-5 hours.

Each coupling step was monitored as described in the general method part for the LC-MS sample preparation. Fmoc-deprotection was carried out after each coupling step was completed, as indicated by LC-MS result.

LC-MS and HR-MS sample preparation

The reaction progress of each coupling of the Fmoc-protected amino acids was monitored using LC-MS. For that a few beads of the resin were sampled in a 2 mL SPPS syringe, washed once with DMF and then 2-3 times with DCM. The vessel was closed and 20% HFIP in DCM was added (1-1.5 mL), which changed the color of the beads from yellow-orange to a dark red which again faded over time. The mixture was shaken for 15-30 min and the filtrate was used for LC-MS measurement. For HR-MS measurement the filtrate was concentrated *in vacuo* and dissolved in DMSO.

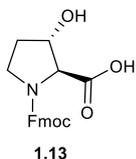
Fmoc-deprotection³¹

The mixture was filtered and the remaining resin was washed five times with DMF. 20% Piperidine in DMF (20-30 mL) was added. After agitation for 2-5 min it was filtrated, rinsed with DMF and the process was repeated four more times. Then the resin was washed with DMF three times.

Cleavage from the resin³¹

To the washed resin, 20% HFIP in DCM was added, coloring the mixture a dark red. The resin was agitated for 10-30 min after which the supernatant was drained and the process was repeated three more times. The combined filtrates were reduced under pressure and then dried further using lyophilization.

Preparation of (2S,3S)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-3-hydroxyproline-2-carboxylic acid (1.13)³⁰



Trans-3-Hydroxyproline (6.10 g, 46.5 mmol) was dissolved in THF and saturated, aqueous NaHCO₃ solution (1:1; 120 mL). Fmoc-OSu (17.27 g, 51.20 mmol) was added and the formed white-orange suspension was stirred at room temperature for 20 h. The reaction progress was monitored by TLC. The mixture was diluted with water and the pH was adjusted to pH = 9 using saturated, aqueous NaHCO₃ solution. The aqueous phase was washed three times with diethyl ether, before its pH was adjusted to pH = 1 using 1 M HCl. The acidic aqueous phase was extracted three times with ethyl acetate. The combined organic layers were washed once with brine, dried over MgSO₄ and concentrated *in vacuo*. The crude product **1.13** (22.59 g) was directly used in the next step, without purification. R_f (PE:EA; 2:1) = 0.53. The NMR data given is a mixture of two rotamers and corresponds to literature.³⁰

¹H-NMR (DMSO-*d*₆, 400 MHz): δ_H [ppm] = 7.94-7.86 (m, 2H, CH_{arom} Fmoc), 7.71-7.61 (m, 2H, CH_{arom} Fmoc), 7.46-7.28 (m, 4H, CH_{arom} Fmoc), 4.40-3.99 (m, 5H, H-2, H-3, CH₂ Fmoc, CH Fmoc), 3.57-3.44 (m, 2H, H-5), 1.98-1.77 (m, 2H, H-4).

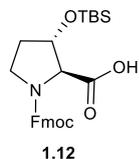
¹³C-NMR (DMSO-*d*₆, 100 MHz): δ_C [ppm] = 172.1, 171.8 (COOH), 154.2, 154.0 (CO Fmoc), 143.80, 143.78, 143.72, 143.66, 140.7, 140.6 (C_{quart} Fmoc), 127.7, 127.1, 125.22, 125.18, 125.1, 120.15, 120.13 (CH_{arom} Fmoc), 73.9, 72.8 (C-2), 68.2, 67.9 (C-3), 67.0, 66.7 (CH₂ Fmoc), 46.61, 46.58 (CH Fmoc), 44.9, 44.3 (C-5), 32.3, 31.3 (C-4).

Additional found signals: δ_H [ppm] = 12.86 (bs), 8.31 (CHCl₃), 5.55, 3.34 (bs, H₂O), 2.59, 1.98 (EA), 1.17 (EA). δ_C [ppm] = 59.7 (EA).

UHR-MS (ESI-TOF) *m/z* calcd for C₂₀H₂₀NO₅: 354.1336 [M+H]⁺; found: 354.1335 [M+H]⁺

Specific rotation [α]_D²⁰ = +16.3° (c = 1.23, CH₃OH)

Preparation of (2S,3S)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-3-((tert-butyl)dimethylsilyloxy)pyrrolidine-2-carboxylic acid (1.12)²⁶



Under argon atmosphere, amino acid **1.13** (10.0 g, 28.3 mmol) was dissolved in anhyd. DCM (600 mL) and cooled to 0 °C. DIPEA (19.0 mL, 109 mmol), followed by TBSOTf (22.0 mL, 95.8 mmol) were added at 0 °C and stirred for 90 min. The reaction was quenched by the addition of methanol and saturated, aqueous NH₄Cl solution and diluted with diethyl ether. The layers were separated and the aqueous phase was extracted twice with diethyl ether. The combined organic layers were washed once with brine, dried over MgSO₄ and concentrated *in vacuo*. The orange syrup was dissolved in methanol (350 mL), saturated, aqueous NaHCO₃ solution (125 mL) and water (65 mL). Potassium carbonate (0.63 g, 4.6 mmol) was added and it was stirred at room temperature for 4 days. It was diluted with diethyl ether and washed once with 10% citric acid (w/v). The aqueous phase was washed three times with diethyl ether and the combined organic layers were washed once with brine, dried over MgSO₄ and concentrated *in vacuo*. Purification of the crude by chromatography (silica, PE:EA, 20:1-0:1) yielded **1.12** as a colorless, foamy solid (8.24 g, 17.6 mmol, 62%). *R_f* (PE:EA; 2:1) = 0.3. The NMR data given is a mixture of two rotamers.

¹H-NMR (CDCl₃, 400 MHz): = δ_H [ppm] = 7.80-7.66 (m, 2H, CH_{arom} Fmoc), 7.64-7.48 (m, 2H, CH_{arom} Fmoc), 7.43-7.21 (m, 4H, CH_{arom} Fmoc), 4.60-4.08 (m, 5H, H-2, H-3, CH₂ Fmoc, CH Fmoc), 3.73-3.60 (m, 2H, H-5), 2.13-1.77 (m, 2H, H-4), 0.92-0.85 (m, 9H, CH₃ TBS), 0.14-0.05 (m, 6H, Si-CH₃ TBS).

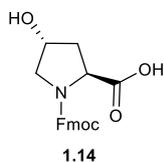
¹³C-NMR (CDCl₃, 100 MHz): δ_C [ppm] = 175.7, 174.2 (COOH), 156.2, 154.8 (CO Fmoc), 144.1, 144.0, 143.9, 141.4, (C_{quart} Fmoc), 127.9, 127.7, 127.2, 127.1, 125.25, 125.21, 125.1, 125.0, 124.9, 120.1, 120.0, (CH_{arom} Fmoc), 76.1, 74.5 (C-3), 68.7, 68.2 (C-2), 68.1, 67.7 (CH₂ Fmoc), 47.3 (CH Fmoc), 45.2, 45.1 (C-5), 33.8, 32.9 (C-4), 25.8, (CH₃ TBS), 18.1 (C_{quart} TBS), -4.76, -4.81 (Si-CH₃ TBS).

Additional found signals: δ_H [ppm] = 1.25. δ_C [ppm] = 60.6 (EA), 14.3 (EA).

UHR-MS (ESI-TOF) *m/z* calcd for C₂₆H₃₄NO₅Si: 468.2201 [M+H]⁺; found: 468.2204 [M+H]⁺

Specific rotation [α]_D²⁰ = -50.7° (c = 1.34, CHCl₃)

Preparation of (2*S*,4*R*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-4-hydroxypyrrolidine-2-carboxylic acid (**1.14**)



Trans-4-Hydroxyproline (8.00 g, 61.0 mmol) was dissolved in THF and saturated, aqueous NaHCO₃ solution (1:1; 280 mL). Fmoc-OSu (22.65 g, 67.14 mmol) was added and the formed white suspension was stirred at room temperature for 20 h. The reaction progress was monitored by TLC. The mixture was diluted with water and the pH was adjusted to pH = 9 using saturated, aqueous NaHCO₃ solution. The aqueous phase was washed three times with diethyl ether, before its pH was adjusted to pH = 1 using 1 M HCl. The acidic aqueous phase was extracted three times with ethyl acetate. The combined organic layers were washed once with brine, dried over MgSO₄ and concentrated *in vacuo*. The crude product **1.14** (22.64 g) was directly used in the next step, without further purification. R_f (PE:EA; 2:1) = 0.45. The NMR data given is a mixture of two rotamers.

¹H-NMR (DMSO-*d*₆, 400 MHz): δ_H [ppm] = 7.93-7.87 (m, 2H, CH_{arom} Fmoc), 7.70-7.64 (m, 2H, CH_{arom} Fmoc), 7.47-7.28 (m, 4H, CH_{arom} Fmoc), 4.46-4.12 (m, 5H, H-2, H-4, CH₂ Fmoc, CH Fmoc), 3.56-3.43 (m, 2H, H-5), 2.31-2.11 (m, 1H, H-3 a), 2.07-1.89 (m, 1H, H-3 b).

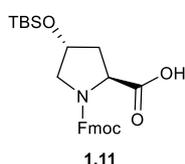
¹³C-NMR (DMSO-*d*₆, 100 MHz): δ_C [ppm] = 174.0, 173.5 (COOH), 154.2, 154.1 (CO Fmoc), 143.7, 143.6, 140.72, 140.70, 140.62, 140.60 (C_{quart} Fmoc), 127.73, 127.69, 127.2, 127.1, 125.30, 125.26, 125.13, 125.10, 120.13, 120.10 (CH_{arom} Fmoc), 68.5, 67.8 (C-4), 67.1, 66.6 (CH₂ Fmoc), 57.8, 57.5 (C-2), 55.0, 54.5 (C-5), 46.6, 46.5 (CH Fmoc), 39.0*, 38.0 (C-3).

Additional found signals: δ_H [ppm] = 5.17, 3.36 (H₂O), 1.17 (EA). δ_C [ppm] = /

UHR-MS (ESI-TOF) m/z calcd for C₂₀H₂₀NO₅: 354.1336 [M+H]⁺; found: 354.1338 [M+H]⁺

Specific rotation $[\alpha]_D^{20} = +36.1^\circ$ (c = 1.55, CH₃OH)

Preparation of (2*S*,4*R*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-4-((*tert*-butyldimethylsilyl)-oxy)pyrrolidine-2-carboxylic acid (**1.11**)



Under argon atmosphere, amino acid **1.14** (61.0 mmol) was dissolved in anhyd. DCM (800 mL) and cooled to 0 °C. DIPEA (48.0 mL, 276 mmol), followed by TBSOTf (46.2 mL, 201 mmol) were added at 0 °C and stirred for

90 min. The reaction was quenched by the addition of methanol and saturated, aqueous NH_4Cl solution and diluted with diethyl ether. The layers were separated and the aqueous phase was extracted twice with diethyl ether. The combined organic layers were washed once with brine, dried over MgSO_4 and concentrated *in vacuo*. The orange syrup was dissolved in methanol (750 mL), saturated, aqueous NaHCO_3 solution (250 mL) and water (130 mL). Potassium carbonate (1.25 g, 9.04 mmol) was added and it was stirred at room temperature for 3 days. It was diluted with diethyl ether and washed once with 10% citric acid (w/v). The aqueous phase was washed three times with diethyl ether and the combined organic layers were washed once with brine, dried over MgSO_4 and concentrated *in vacuo*. Purification of the crude by chromatography (silica, 5 to 50% EA in PE) yielded **1.11** as a colorless, foamy solid (25.56 g, 54.66 mmol, 90%). R_f (PE:EA; 1:1) = 0.24. The NMR data given is a mixture of two rotamers and corresponds to literature.³⁰

$^1\text{H-NMR}$ (CDCl_3 , 400 MHz): δ_{H} [ppm] = 7.80-7.69 (m, 2H, CH_{arom} Fmoc), 7.61-7.51 (m, 2H, CH_{arom} Fmoc), 7.44-7.24 (m, 4H, CH_{arom} Fmoc), 4.57-4.16 (m, 5H, H-2, H-4, CH_2 Fmoc, CH Fmoc), 3.67-3.42 (m, 2H, H-5), 2.33-2.06 (m, 2H, H-3), 0.91-0.84 (m, 9H, CH_3 TBS), 0.11-0.04 (m, 6H, Si- CH_3 TBS).

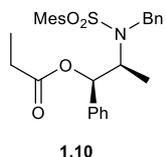
$^{13}\text{C-NMR}$ (CDCl_3 , 100 MHz): δ_{C} [ppm] = 175.7 (COOH), 156.4 (CO Fmoc), 144.2, 143.9, 143.8, 141.5, 141.4, (C_{quart} Fmoc), 127.9, 127.8, 127.23, 127.18, 125.2, 125.1, 125.0, 120.1, 120.0, (CH_{arom} Fmoc), 70.3, 69.7 (C-4), 68.2, 67.7 (CH_2 Fmoc), 58.3, 57.6, (C-2), 55.3, 55.0, (C-5), 47.3, 47.2 (CH Fmoc), 40.1, 38.4 (C-3), 25.8, (CH_3 TBS), 18.1 (C_{quart} TBS), -4.65, -4.74 (Si- CH_3 TBS).

Additional found signals: δ_{H} [ppm] = 5.30 (DCM), 4.12 (EA), 2.05 (EA); 1.26 (EA). δ_{C} [ppm] = 60.6 (EA), 14.4 (EA).

UHR-MS (ESI-TOF) m/z calcd for $\text{C}_{26}\text{H}_{34}\text{NO}_5\text{Si}$: 468.2201 [$\text{M}+\text{H}$]⁺; found: 468.2200 [$\text{M}+\text{H}$]⁺

Specific rotation $[\alpha]_{\text{D}}^{20} = -82.0^\circ$ (c = 1.03, CHCl_3)

Preparation of (1*R*,2*S*)-2-((*N*-benzyl-2,4,6-trimethylphenyl)sulfonamido)-1-phenylpropyl propionate (**1.10**)³³



(1*R*,2*S*)-*N*-Benzyl-*N*-(mesitylenesulfonyl)norephedrine (13.35 g, 31.52 mmol) was dissolved in anhyd. DCM (200 mL) and cooled to 0 °C. Pyridine (3.3 mL, 41 mmol) was slowly added and stirred for 5 min. Propionyl chloride (3.0 mL, 34 mmol) was added dropwise at 0 °C and the reaction was stirred at room temperature for 18 h. It was washed once each with water, 1 M HCl, sat. aqueous NaHCO₃ and brine. Drying over MgSO₄, concentrating *in vacuo* and purification by column chromatography (silica, PE:EA, 4:1) yielded the product **1.10** as white crystals (14.15 g, 29.50 mmol, 94%). *R_f* (PE:EA, 4:1) = 0.52.

¹H-NMR (CDCl₃, 400 MHz): δ_H [ppm] = 7.28-7.23 (m, 2H, *H*_{arom}), 7.21-7.08 (m, 6H, *H*_{arom}), 6.87-6.82 (m, 2H, *H*_{arom}), 6.80 (s, 2H, *H*_{arom}), 5.77 (d, 1H, *J* = 4.0 Hz, *CH*-Ph), 4.64 (d, 1H, *J* = 16.8 Hz, *CH*₂-Ph a), 4.53 (d, 1H, *J* = 16.8 Hz, *CH*₂-Ph b), 3.97 (dq, 1H, *J* = 6.9, 4.0 Hz, *CH*-CH₃), 2.44 (s, 6H, *CH*₃ Mes), 2.20 (s, 3H, *CH*₃ Mes), 2.11 (dq, 1H, *J* = 17.2 Hz, *J* = 7.7 Hz, *CH*₂-CH₃ a), 2.03 (dq, 1H, *J* = 17.3 Hz, *J* = 7.6 Hz, *CH*₂-CH₃ b), 1.04 (d, 3H, *J* = 6.8 Hz, *CH*-CH₃), 0.94 (t, 3H, *J* = 7.4 Hz, *CH*₂-CH₃).

¹³C-NMR (CDCl₃, 100 MHz): δ_C [ppm] = 172.7 (CO), 142.6 (*C*_{quart} Mes), 140.4 (*C*_{quart} Mes), 138.81 (*C*_{quart} Bn), 138.76 (*C*_{quart} Ph), 133.5 (*C*_{quart} Mes), 132.3 (*CH*_{arom} Mes), 128.53, 128.50, 127.9, 127.5, 127.2, 126.1 (*CH*_{arom}), 78.1 (*CH*-Ph), 56.9 (*CH*-CH₃), 48.3 (*CH*₂-Ph), 27.6 (*CH*₂-CH₃), 23.1 (*CH*₃ Mes), 21.0 (*CH*₃ Mes), 12.9 (*CH*-CH₃), 8.9 (*CH*₂-CH₃).

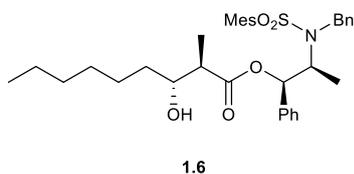
Additional found signals: δ_H [ppm] = 1.49. δ_C [ppm] = /

UHR-MS (ESI-TOF) *m/z* calcd for C₂₈H₃₃NO₄S: 480.2203 [M+H]⁺; found: 480.2201 [M+H]⁺

Specific rotation [α]_D²⁰ = +9.8° (c = 2.05, CHCl₃)

³³ T. Inoue, J. Liu, D. C. Buske, A. Abiko, *J. Org. Chem.* **2002**, 67, 5250-5256.

Preparation of (1*R*,2*S*)-2-((*N*-benzyl-2,4,6-trimethylphenyl)sulfonamido)-1-phenylpropyl (2*R*,3*R*)-3-hydroxy-2-methylnonanoate (1.6)



Under argon **1.10** (10.01 g, 20.87 mmol) and triethylamine (7.0 mL, 50 mmol) were dissolved in anhyd. DCM (150 mL), stirred for 5 min at room temperature and then cooled to -78 °C. 1 M dicyclohexylboron triflate in anhyd. DCM (42 mL, 42 mmol) was added dropwise over a time of 45 min. After the reaction was stirred for 2.5 h at -78 °C *n*-heptanal (5.2 mL, 37 mmol) was added dropwise and stirred for another hour at -78 °C. The reaction was allowed to return to room temperature and was quenched by the addition of pH 7 buffer solution (100 mL) and hydrogen peroxide (40 mL, 35 w%), it was then diluted with Methanol (250 mL). The reaction was stirred for 20 h, concentrated *in vacuo* and the organic and aqueous phases were separated. The aqueous layer was extracted three times with DCM, washed once with brine and was dried over Na₂SO₄. Purification of 31% of the crude by column chromatography (silica, PE:EA, 8:1) was followed by final purification by flash chromatography (0-20% EA in *n*-heptane). It yielded the unwanted *syn*-**1.6** stereoisomer as a colorless syrup (151 mg, 0.254 mmol, overall yield calculated to be: 4%) and the desired *anti*-**1.6** stereoisomer as a colorless syrup (3.57 g, 6.01 mmol, overall yield calculated to be: 93%, *de* = 96:4). *R_f* (*n*-heptane:EA; 4:1) = 0.29.

¹H-NMR (CDCl₃, 400 MHz): δ_H [ppm] = 7.33-7.14 (m, 8H, *H*_{arom}), 6.90-6.83 (m, 4H, *H*_{arom}), 5.83 (d, 1H, *J* = 4.3 Hz, *CH*-Ph Aux), 4.76 (d, 1H, *J* = 16.5 Hz, *CH*₂-Ph a Aux), 4.54 (d, 1H, *J* = 16.5 Hz, *CH*₂-Ph b Aux), 4.12 (dq, 1H, *J* = 4.9, 6.7 Hz, *CH*-CH₃ Aux), 3.66-3.58 (m, 1H, *H*-3), 2.48 (s, 6H, CH₃ Mes), 2.47 (dq, *J* = 6.5, 7.2 Hz, 1H, *H*-2), 2.28 (s, 3H, CH₃ Mes), 1.51-1.20 (m, 10H, CH₂ alkyl), 1.18 (d, 3H, *J* = 7.0 Hz, *CH*-CH₃ Aux), 1.13 (d, 3H, *J* = 7.3 Hz, CH₃-2'), 0.87 (t, 3H, *J* = 6.8 Hz, CH₃-9).

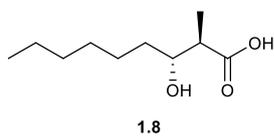
¹³C-NMR (CDCl₃, 100 MHz): δ_C [ppm] = 174.7 (CO), 142.7 (*C*_{quart} Mes), 140.4 (*C*_{quart} Mes), 138.6 (*C*_{quart} Bn), 138.3 (*C*_{quart} Ph), 133.5 (*C*_{quart} Mes), 132.2 (*C*_{arom} Mes), 128.6, 128.5, 128.1, 127.8, 127.3, 126.1, (*C*_{arom}), 78.3 (*CH*-Ph Aux), 73.3 (*C*-3), 56.9 (*CH*-CH₃ Aux), 48.4 (*CH*₂-Ph Aux), 45.6 (*C*-2), 34.6 (CH₂ alkyl), 31.9 (CH₂ alkyl), 29.4 (CH₂ alkyl), 25.5 (CH₂ alkyl), 23.1 (CH₃ Mes), 22.7 (CH₂ alkyl), 21.0 (CH₃ Mes), 14.2 (CH₃-2'), 14.2 (CH₃-9), 13.6 (CH₃ Aux).

Additional found signals: δ_H [ppm] = 2.64 (DMSO). δ_C [ppm] = /

UHR-MS (ESI-TOF) m/z calcd for $C_{35}H_{47}NO_5Na$: 616.3067 $[M+Na]^+$; found: 616.3062 $[M+Na]^+$

Specific rotation $[\alpha]_D^{20} = +35.2^\circ$ ($c = 1.42$, $CHCl_3$)

Preparation of (2*R*,3*R*)-3-Hydroxy-2-methylnonanoic acid (1.8)



The stereoisomer anti-**1.6** (3.525 g, 5.936 mmol) was dissolved in a mixture of MeOH-THF- H_2O (1:1:1; 105 mL) and stirred for 5 min until the syrup was completely dissolved. $LiOH \cdot H_2O$ (1.259 g, 30.00 mmol) was added and it was stirred at room temperature for 20 h. LC-MS indicated complete conversion and the mixture was poured into water and extracted three times with DCM. The combined organic phases were washed once with brine, dried over $MgSO_4$ and concentrated *in vacuo*, which yielded the essentially pure auxiliary ((1*R*,2*S*)-*N*-Benzyl-*N*-(mesitylenesulfonyl)norephedrine) as a white solid (2.54 g, 6.00 mmol, >100%) no purification needed. The aqueous layer was acidified using 1 M HCl to pH = 1 and extracted three times with diethyl ether. The combined organic phases were washed once with brine, dried over $MgSO_4$ and concentrated *in vacuo*. Acid **1.8** was obtained as a slightly yellow oil (1.15 g, 6.11 mmol, >100%), no further purification necessary.

1H -NMR ($CDCl_3$, 400 MHz): δ_H [ppm] = 3.74-3.66 (m, 1H, *H*-3), 2.57 (dq, 1H, $J = 7.0, 6.9$ Hz, *H*-2), 1.61-1.27 (m, 10H, CH_2 alkyl), 1.25 (d, 3H, $J = 7.2$ Hz, CH_3 -2'), 0.89 (t, 3H, $J = 7.1$ Hz, CH_3 -9).

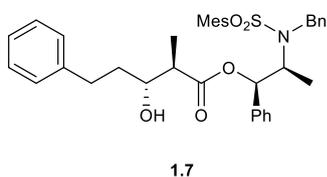
^{13}C -NMR ($CDCl_3$, 100 MHz): δ_C [ppm] = 180.6 (CO), 73.4 (C-3), 45.2 (C-2), 34.78 (CH_2 alkyl), 31.9 (CH_2 alkyl), 29.3 (CH_2 alkyl), 25.5 (CH_2 alkyl), 22.7 (CH_2 alkyl), 14.4 (CH_3 -2'), 14.2 (CH_3 -9).

Additional found signals: δ_H [ppm] = / δ_C [ppm] = /

UHR-MS (ESI-TOF) m/z calcd for $C_{10}H_{21}N_2O_3$: 189.1485 $[M+H]^+$; found: 189.1484 $[M+H]^+$

Specific rotation $[\alpha]_D^{20} = -20.2^\circ$ ($c = 1.49$, $CHCl_3$)

Preparation of (1*R*,2*S*)-2-((*N*-benzyl-2,4,6-trimethylphenyl)sulfonamido)-1-phenylpropyl (2*R*,3*R*)-3-hydroxy-2-methyl-5-phenylpentanoate (1.7)



Under argon **1.10** (1.90 g, 3.96 mmol) and triethylamine (1318 μ L, 9.508 mmol) were dissolved in anhyd. DCM (30 mL), stirred for 5 min at room temperature and then cooled to $-78^\circ C$. 1 M

dicyclohexylboron triflate in anhyd. DCM (7.9 mL, 7.9 mmol) was added dropwise over a time of 45 min. After the reaction was stirred for 2.5 h at -78 °C. 3-Phenylpropionaldehyde (1.0 mL, 7.5 mmol) was added dropwise and stirred for another hour at -78 °C. The reaction was allowed to return to room temperature and was quenched by the addition of pH 7 buffer solution (20 mL) and hydrogen peroxide (7.5 mL, 35 w%), it was then diluted with Methanol (50 mL). The reaction was stirred for 23 h, concentrated *in vacuo* and the organic and aqueous phases were separated. The aqueous layer was extracted three times with DCM, washed once with brine and was dried over Na₂SO₄. Purification of the crude by column chromatography (silica, PE:EA, 5:1) was followed by final purification by flash chromatography (0-20% EA in *n*-heptane). It yielded the unwanted syn-**1.7** stereoisomer as a colorless syrup (158 mg, 0.258 mmol, 7%) and the desired anti-**1.7** stereoisomer as a colorless syrup (2.02 g, 3.29 mmol, 83%, *de* = 93:7). *R_f* (PE:EA; 5:1) = 0.21.

¹H-NMR (CDCl₃, 400 MHz): δ_H [ppm] = 7.31-7.15 (m, 13H, *H*_{arom}), 6.93-6.88 (m, 2H, *H*_{arom}), 6.86 (s, 2H, *H*_{arom} Mes), 5.88 (d, 1H, *J* = 4.4 Hz, *CH*-Ph Aux), 4.72 (d, 1H, *J* = 16.5 Hz, *CH*₂-Ph a Aux), 4.53 (d, 1H, *J* = 16.5 Hz, *CH*₂-Ph b Aux), 4.14 (dq, 1H, *J* = 7.0, 4.3 Hz, *CH*-CH₃ Aux), 3.64 (ddd, 1H, *J* = 9.1, 6.5, 3.0 Hz, *H*-3), 2.84 (ddd, 1H, *J* = 14.0, 9.3, 5.0 Hz, *H*-5a), 2.66 (ddd, 1H, *J* = 13.7, 9.5, 6.9 Hz, *H*-5b), 2.48 (s, 6H, *CH*₃ Mes), 2.47 (dq, 1H, *J* = 7.0, 7.0 Hz, *H*-2), 2.28 (s, 3H, *CH*₃ Mes), 1.77 (dddd, 1H, *J* = 13.5, 10.1, 6.9, 3.3 Hz, *H*-4a), 1.70 (dddd, 1H, *J* = 13.9, 9.3, 4.7, 4.7 Hz, *H*-4b), 1.19 (d, 3H, *J* = 7.0 Hz, *CH*-CH₃ Aux), 1.14 (d, 3H, *J* = 7.1 Hz, *CH*₃-2').

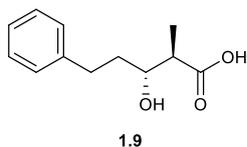
¹³C-NMR (CDCl₃, 100 MHz): δ_C [ppm] = 174.6 (*CO*), 142.7 (*C*_{quart} Mes), 142.0 (*C*_{quart}), 140.3 (*C*_{quart} Mes), 138.4 (*C*_{quart} Bn), 138.2 (*C*_{quart} Ph), 133.4 (*C*_{quart} Mes), 132.2 (*C*_{arom} Mes), 128.6, 128.51, 128.45, 128.1, 127.7, 127.3, 126.1, 126.0 (*C*_{arom}), 78.4 (*CH*-Ph Aux), 72.5 (*C*-3), 56.8 (*CH*-CH₃ Aux), 48.3 (*CH*₂-Ph Aux), 45.6 (*C*-2), 36.4 (*C*-4), 31.9 (*C*-5), 23.0 (*CH*₃ Mes), 21.0 (*CH*₃ Mes), 14.2 (*CH*₃-2'), 13.5 (*CH*₃ Aux).

Additional found signals: δ_H [ppm] = 4.13, 2.05, 1.27 (EA). δ_C [ppm] = 60.5, 21.2, 14.3 (EA)

UHR-MS (ESI-TOF) *m/z* calcd for C₃₇H₄₃NO₅SNa: 636.2754 [M+Na]⁺; found: 636.2755 [M+Na]⁺

Specific rotation [α]_D²⁰ = -20.0° (c = 1.00, CHCl₃)

Preparation of (2*R*,3*R*)-3-hydroxy-2-methyl-5-phenylpentanoic acid (**1.9**)



The stereoisomer anti-**1.7** (987 mg, 1.61 mmol) was dissolved in a mixture of MeOH-THF-H₂O (1:1:1; 30 mL) and stirred for 5 min. LiOH · H₂O (339 mg, 8.08 mmol) was added and it was stirred at room temperature for 48 h. LC-MS showed full conversion and the mixture was poured into water and extracted three times with DCM. The combined organic phases were washed once with brine, dried over MgSO₄ and concentrated *in vacuo* which yielded the crude auxiliary ((1*R*,2*S*)-*N*-Benzyl-*N*-(mesitylenesulfonyl)norephedrine) as yellow crystals. These were collected from different reactions and combined for purification by column chromatography (silica, PE:EA; 5:1). The aqueous layer was acidified using 2 M HCl to pH = 1 and extracted three times with diethyl ether. The combined organic phases were washed once with brine, dried over MgSO₄ and concentrated *in vacuo*. Crystallization from acetonitrile yielded acid **1.9** as colorless crystals (141 mg, 0.677 mmol, 42%).

¹H-NMR (DMSO-*d*₆, 400 MHz): δ_H [ppm] = 7.30-7.12 (m, 5H, *H*_{arom}), 3.60 (ddd, 1H, *J* = 9.4, 6.8, 2.8 Hz, *H*-3), 2.74 (ddd, 1H, *J* = 13.9, 10.0, 4.6 Hz, *H*-5a), 2.55 (ddd, 1H, *J* = 13.8, 10.2, 6.8 Hz, *H*-5b), 2.40 (dq, 1H, *J* = 7.0, 7.0 Hz, *H*-2), 1.67 (dddd, 1H, *J* = 13.5, 10.1, 6.8, 3.2 Hz, *H*-4a), 1.55 (dddd, 1H, *J* = 9.4, 14.0, 9.4, 4.6 Hz, *H*-4b), 0.98 (d, 3H, *J* = 7.2 Hz, CH₃-2').

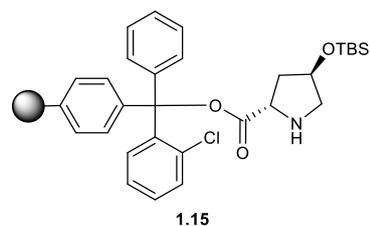
¹³C-NMR (DMSO-*d*₆, 100 MHz): δ_C [ppm] = 176.0 (CO), 142.3 (C_{quart}), 128.3, 128.2, 125.6 (C_{arom}), 70.9 (C-3), 45.8 (C-2), 35.2 (C-4), 31.4 (C-5), 12.6 (CH₃-2').

Additional found signals: δ_H [ppm] = 3.34 (H₂O). δ_C [ppm] = /

UHR-MS (ESI-TOF) *m/z* calcd for C₁₂H₁₆O₃Na: 231.0992 [M+Na]⁺; found: 231.0992 [M+Na]⁺

Specific rotation [α]_D²⁰ = +23.0° (c = 1.65, CHCl₃)

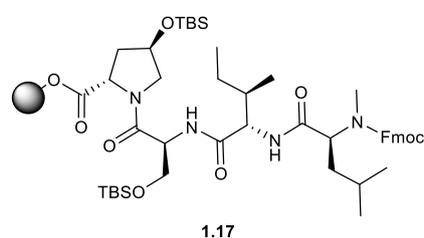
Preparation of 2CT-4-Hyp(TBS)-NH (**1.15**)



2-Chlorotrityl chloride resin (n = 1.55 mmol/g, 4.00 g, 6.20 mmol) was swelled in DCM (30 mL) for 5 min. After removal of the solvent, Fmoc-protected amino acid **1.11** (3.78 g, 8.08 mmol) and DIPEA (3400 μL, 20.00 mmol), dissolved in DCM (40 mL), were added and it was agitated for 45 min. The solvent was filtered

off and the resin was washed twice with DMF. The capping mixture (40 mL), consisting of DCM:MeOH:DIPEA (80:15:5), was added and it was agitated for 10 min. After removal of the solvent, the capping was repeated once more. The resin was washed three times with DMF and a sample was taken for determination of the loading. Fmoc-deprotection was done with 20% piperidine in DMF (30 mL) for 3 min and the process was carried out five times total. The resin was washed six times with DMF, three times with isopropanol and three times with *n*-heptane. It was sucked dry and dried further *in vacuo* for 24 h. Resin **1.15** (4.22 g) was stored under argon at 4 °C. The loading was determined by absorption to be $n = 0.481$ mmol/g (31%).

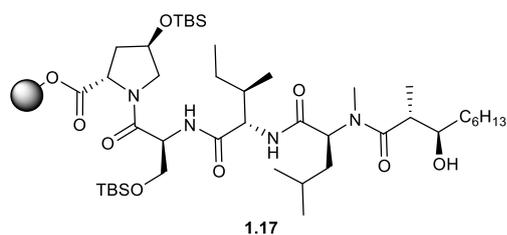
Preparation of 2CT-4-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-Fmoc (1.17)



Resin **1.15** ($n = 0.481$ mmol/g, 4.20 g, 2.02 mmol) was swelled in DMF for 30 min. After removal of the solvent, Fmoc-L-serine(TBS)-OH (2.681 g, 6.071 mmol) and HATU (2.228 g, 5.860 mmol) were added, dissolved in a small volume of DMF, followed by DIPEA (2058 μ L, 12.10 mmol)

and more DMF. The mixture was agitated for 3 h. After Fmoc-deprotection, a solution of Fmoc-L-*allo*-isoleucine-OH (2.137 g, 6.047 mmol) and HATU (2.226 g, 5.854 mmol) in DMF, were added, followed by DIPEA (2058 μ L, 12.10 mmol) and more DMF. The mixture was agitated for 2.5 h. Fmoc-deprotection was followed by the coupling of Fmoc-N-Me-L-leucine-OH (2.225 g, 6.055 mmol), which was dissolved with HATU (2.230 g, 5.865 mmol) in a small amount of DMF. The mixture was agitated for 2.5 h. The supernatant was drained, the resin was washed five times with DMF, three times each with DCM, methanol and diethyl ether. It was sucked dry and dried further *in vacuo* for 12 h. The finished resin **1.17** (5.03 g, $n = 0.402$ mmol/g) was stored under argon at 4 °C.

Preparation of 2CT-4-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-alkyl-BHA (1.18)

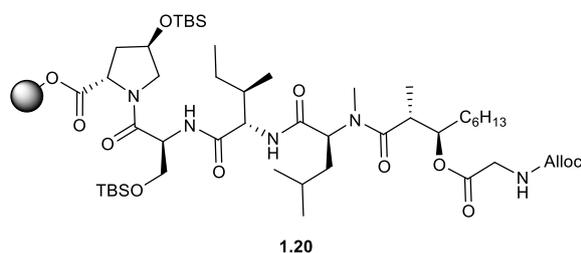


Resin **1.17** ($n = 0.402$ mmol/g, 1.512 g, 0.607 mmol) was swelled in DMF for 60 min. The supernatant was drained and after Fmoc-deprotection the resin was washed twice with anhyd. DMF. The resin was

suspended in a small volume of anhyd. DMF. The solution of acid **1.8** (279 mg, 1.48 mmol) and triethylamine (200 μ L, 1.44 mmol) in DMF was cooled to 0 °C and DEPC (90%, 241 μ L, 1.42 mmol) was added. It was directly transferred to the peptide vessel and the mixture was agitated for 20 h. Due to incomplete conversion, as indicated by LC-MS result, the coupling step was repeated. The mixture was drained, the resin was washed twice with anhyd. DMF and suspended in a small amount of anhyd. DMF. Acid **1.8** (138 mg, 0.733 mmol) was dissolved in anhyd. DMF and triethylamine (100 μ L, 0.721 mmol). It was cooled to 0 °C and DEPC (90%, 121 μ L, 0.714 mmol) was added. It was directly transferred to the peptide vessel and the mixture was agitated for 5 h after which HR-MS indicated almost full conversion. The supernatant was filtered off and the resin was washed five times each with DMF and DCM. The resin coupled peptide **1.18** was used directly in the next step.

UHR-MS (ESI-TOF) m/z calcd for $C_{43}H_{85}N_4O_9Si_2$: 857.5850 $[M+H]^+$; found: 857.5846 $[M+H]^+$

Preparation of 2CT-4-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-alkyl-BHA-Gly-alloc (**1.20**)

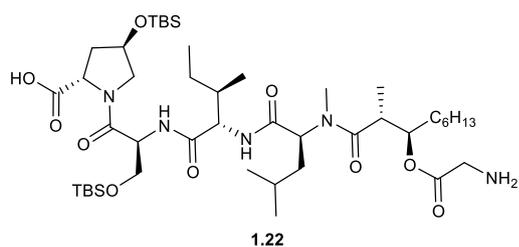


Resin coupled peptide **1.18** (0.607 mmol) was swelled in THF for 10 min, filtered and suspended in anhyd. THF. Alloc-glycine-OH (343 mg, 2.15 mmol), dissolved in anhyd. THF, was added, followed by DIC (335 μ L,

2.16 mmol) and DMAP (27.7 mg, 0.227 mmol). The resin was agitated for 16 h after which the step was repeated. The supernatant was filtered off, the resin was washed once with anhyd. THF and then suspended in anhyd. THF. Alloc-glycine-OH (346 mg, 2.17 mmol), dissolved in anhyd. THF, was added to the resin, followed by DIC (335 μ L, 2.16 mmol) and DMAP (27.2 mg, 0.223 mmol). The mixture was agitated for 8 h after which HR-MS indicated full conversion. The supernatant was drained and the resin was washed once with DCM. The resin coupled peptide **1.20** was directly used in the next step.

UHR-MS (ESI-TOF) m/z calcd for $C_{49}H_{92}N_5O_{12}Si_2$: 998.6276 $[M+H]^+$; found: 998.6276 $[M+H]^+$

Preparation of HOOC-4-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-alkyl-BHA-Gly-NH₂ (**1.22**)

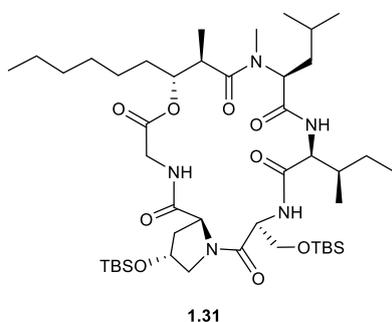


Resin coupled peptide **1.20** (0.607 mmol) was swelled in DCM for 15 min. It was filtered, washed with anhyd. DCM and then suspended in anhyd. DCM. Pd(PPh₃)₄ (83 mg, 0.072 mmol) was added, followed by Phenylsilane (2.1 mL, 17 mmol). The

mixture was agitated for 1 h after which HR-MS indicated full conversion. The solvent was removed and the resin was washed five times with DCM until the solvent ran clear. Cleavage of the peptide from the resin was done as described in the general method part. The linear peptide **1.22** was used without further purification.

UHR-MS (ESI-TOF) m/z calcd for C₄₅H₈₈N₅O₁₀Si₂: 914.6064 [M+H]⁺; found: 914.6065 [M+H]⁺

Preparation of (6*R*,7*R*,10*S*,13*S*,16*S*,20*R*,21*aS*)-13-((*R*)-*sec*-butyl)-20-((*tert*-butyldimethylsilyl)oxy)-16-(((*tert*-butyldimethylsilyl)oxy)methyl)-6-hexyl-10-isobutyl-7,9-dimethyltetradecahydropyrrolo[2,1-*f*][1]oxa[4,7,10,13,16]pentaazacyclononadecine-1,4,8,11,14,17-hexaone (**1.31**)

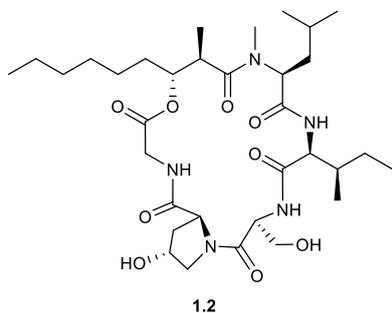


Linear peptide **1.22** (0.607 mmol) was dissolved in DCM:DMF (15:1; 320 mL). HATU (828 mg, 2.18 mmol), followed by DIPEA (612 μ L, 3.60 mmol) were added and the mixture was stirred for 17 h at room temperature after which HR-MS showed complete conversion. It was extracted twice with water, once with brine, dried over MgSO₄ and concentrated

in vacuo. Prepurification by column chromatography (silica, PE:EA; 1:2) yielded **1.31** as slightly yellow needles (329 mg, 0.367 mmol, 60% over 4 steps). R_f (PE:EA; 1:2) = 0.73.

UHR-MS (ESI-TOF) m/z calcd for C₄₅H₈₅N₅O₉Si₂Na: 918.5778 [M+Na]⁺; found: 918.5776 [M+Na]⁺

Preparation of (6*R*,7*R*,10*S*,13*S*,16*S*,20*R*,21*aS*)-13-((*R*)-*sec*-butyl)-6-hexyl-20-hydroxy-16-(hydroxymethyl)-10-isobutyl-7,9-dimethyltetradecahydro-pyrrolo[2,1-*f*][1]oxa[4,7,10,13,16]pentaazacyclononadecine-1,4,8,11,14,17-hexaone (1.2**)**



Under argon peptide **1.31** (296 mg, 0.330 mmol) was dissolved in anhyd. THF (15 mL) in a Teflon® round bottom flask. Acetic acid (1436 μ L, 25.09 mmol) was added, followed by TBAF (18.0 mL, 62.2 mmol) and the mixture was stirred at room temperature for 21 h. It was diluted with EA, washed twice with saturated, aqueous NaHCO₃ solution, once with

brine, dried over MgSO₄ and concentrated *in vacuo*. Purification by semi preparative HPLC (5-50-75-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 3 mL/min) yielded **1.2** as a colorless solid (66 mg, 0.10 mmol, 30%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 8.81 (minor, NH Gly), 8.58 (minor, NH Ser), 8.49 (d, 1H, *J* = 4.6 Hz, NH Ser), 8.21 (d, 1H, *J* = 9.3 Hz, NH *allo* Ile), 8.07 (d, 1H, *J* = 8.3 Hz, NH Gly), 6.68 (minor, NH *allo* Ile), 5.21-5.13 (m, 1H, OH Ser), 5.13-5.08 (m, 1H, OH 4-Hyp), 5.05 (ddd, 2H, *J* = 10.2, 6.4, 3.6 Hz, 3-CH-O BHA), 4.67 (dd, 1H, *J* = 8.5, 4.3 Hz, α -CH 4-Hyp), 4.40 (dd, 1H, *J* = 9.5, 4.6 Hz, α -CH *allo* Ile), 4.26 (d, 1H, *J* = 8.1 Hz, CH₂ a Gly), 4.23 (d, 1H, *J* = 8.3 Hz, CH₂ b Gly), 4.20-4.10 (m, 2H, γ -CH 4-Hyp), 4.02 (ddd, 1H, *J* = 6.7, 6.7, 4.6 Hz, α -CH Ser), 3.61-3.56 (m, 1H, β -CH₂ a Ser), 3.55-3.45 (m, 2H, β -CH₂ b Ser, α -CH *N*-Me Leu), 3.45-3.33 (m, 4H, δ -CH₂ a, b 4-Hyp, α -CH₂ Gly minor), 3.15-3.11 (m, 2H, 2-CH, BHA), 3.10 (s, 3H, N-CH₃), 2.27 (ddd, 1H, *J* = 12.5, 4.9, 4.9 Hz, β -CH₂ a 4-Hyp), 2.13-2.03 (m, 3H, β -CH₂ b 4-Hyp, β -CH₂ a *N*-Me Leu), 1.80-1.71 (m, 2H, β -CH *allo* Ile), 1.71-1.64 (m, 1H, 4-CH₂ a BHA), 1.64-1.54 (m, 2H, β -CH₂ b *N*-Me Leu), 1.52-1.45 (m, 1H, 4-CH₂ b BHA), 1.45-1.38 (m, 2H, γ -CH *N*-Me Leu), 1.38-1.32 (m, 2H, γ -CH₂ a *allo* Ile), 1.32-1.17 (m, 14H, CH₂ BHA), 1.11-1.02 (m, 2H, γ -CH₂ b *allo* Ile), 0.99 (d, 3H, *J* = 7.0 Hz, 2'-CH₃ BHA), 0.94-0.83 (m, 20H, δ -CH₃ *allo* Ile, δ -CH₃ a, b *N*-Me Leu, 9-CH₃ BHA), 0.69 (d, 3H, *J* = 6.9 Hz, γ -CH₃ *allo* Ile).

¹³C-NMR (DMSO-*d*₆, 100 MHz): δ_{C} [ppm] = 176.4 (CO BHA), 171.8 (CO *allo* Ile), 171.0 (CO *N*-Me Leu), 170.7 (CO 4-Hyp), 169.7 (CO Ser), 168.4 (CO Gly), 76.0 (3-CH BHA), 68.5* (α -CH *N*-Me Leu), 66.8 (γ -CH 4-Hyp), 60.7 (β -CH₂ Ser), 59.3 (α -CH 4-Hyp), 54.3 (α -CH Ser), 53.9 (α -CH *allo* Ile), 53.5 (δ -CH₂ 4-Hyp), 41.4 (CH₂ Gly), 40.1* (N-CH₃), 40.1* (2-CH BHA), 38.8* (β -CH₂ 4-Hyp), 38.2 (β -CH₂ *N*-Me Leu), 37.8 (β -CH *allo* Ile), 31.1 (CH₂ BHA), 30.7 (4-CH₂ BHA), 28.8 (CH₂

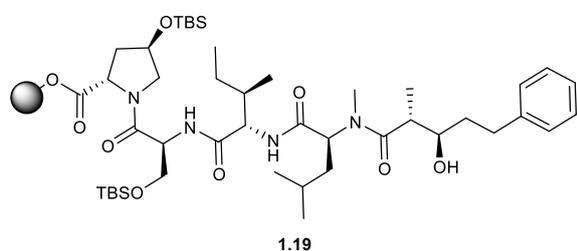
BHA), 25.8 (γ -CH₂ *allo* Ile), 24.6 (γ -CH *N*-Me Leu), 23.2 (CH₂ BHA), 22.8 (δ -CH₃ *a* *N*-Me Leu), 22.0 (CH₂ BHA), 21.7 (δ -CH₃ *b* *N*-Me Leu), 14.7 (2'-CH₃ BHA), 13.9 (γ -CH₃ *allo* Ile), 13.9 (9-CH₃ BHA), 11.6 (δ -CH₃ *allo* Ile).

Additional found signals: δ_{H} [ppm] = 5.75 (DCM), 3.30 (H₂O), 2.74 (DMF), 2.54 (DMSO). δ_{C} [ppm] = 40.4 (DMSO)

UHR-MS (ESI-TOF) m/z calcd for C₃₃H₅₈N₅O₉: 668.4229 [M+H]⁺; found: 668.4233 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{20} = -70.9^\circ$ (c = 1.10, CHCl₃)

Preparation of Pol-4-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-arom-BHA (1.19)

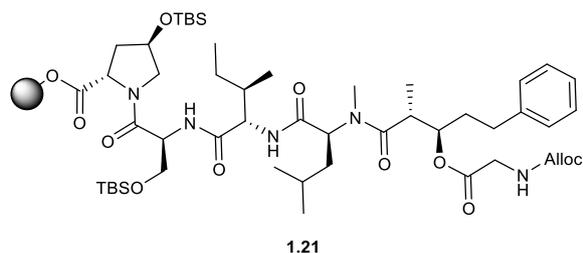


Resin **1.17** (n = 0.402 mmol/g, 1.11 g, 0.446 mmol) was swelled in DMF for 2 h. The supernatant was drained and after Fmoc-deprotection the resin was washed twice with anhyd. DMF. It was suspended in a small

amount of anhyd. DMF. The solution of acid **1.9** (321 mg, 1.54 mmol) and triethylamine (208 μ L, 1.50 mmol) in anhy. DMF was cooled to 0 °C. DEPC (90%, 252 μ L, 1.49 mmol) was added, it was directly transferred to the peptide vessel and the mixture was agitated for 18 h. The step was repeated due to incomplete conversion. Acid **1.9** (104 mg, 0.499 mmol) and triethylamine (69 μ L, 0.50 mmol) were dissolved in anhy. DMF and cooled to 0 °C. DEPC (90%, 84 μ L, 0.50 mmol) was added and it was transferred to the peptide vessel and existing mixture. It was agitated for 20 h after which HR-MS indicated almost full conversion. The supernatant was drained, and the resin was washed five times each with DMF and DCM. The resin coupled peptide **1.19** was directly used in the next step.

UHR-MS (ESI-TOF) m/z calcd for C₄₅H₈₁N₄O₉Si₂: 877.5537 [M+H]⁺; found: 877.5533 [M+H]⁺

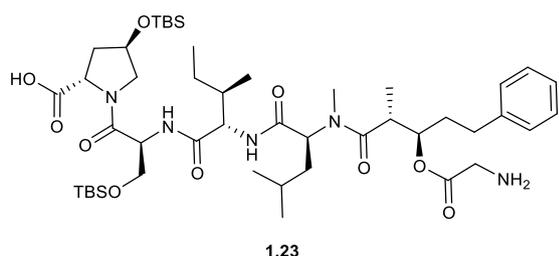
Preparation of Pol-4-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-arom-BHA-Gly-alloc (**1.21**)



Resin coupled peptide **1.19** (0.446 mmol) was swelled in THF for 90 min, filtered and suspended in anhyd. THF. Alloc-glycine-OH (447 mg, 2.81 mmol), dissolved in anhyd. THF, was added, followed by DIC (465 μ L, 3.00 mmol) and DMAP (36.6 mg, 0.300 mmol). The resin was agitated for 20 h, after which HR-MS indicated full conversion. The supernatant was filtered off and the resin was washed once with DCM. The resin coupled peptide **1.21** was directly used in the next step.

UHR-MS (ESI) m/z calcd for $C_{51}H_{88}N_5O_{12}Si_2$: 1018.5963 $[M+H]^+$; found: 1018.5950 $[M+H]^+$

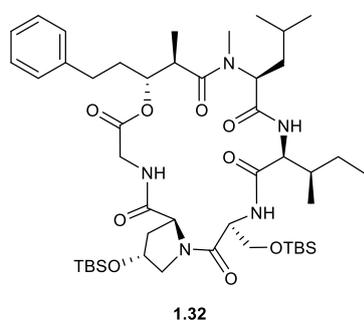
Preparation of HOOC-4-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-arom-BHA-Gly-NH₂ (**1.23**)



Resin coupled peptide **1.21** (0.446 mmol) was swelled in DCM for 15 min. It was drained, washed with anhyd. DCM and then suspended in anhyd. DCM. Pd(PPh₃)₄ (59 mg, 0.051 mmol) was added, followed by Phenylsilane (1481 μ L, 12.00 mmol). The mixture was agitated for 3 h after which HR-MS indicated full conversion. The mixture was filtered and washed five times with DCM until the solvent ran clear. Cleavage of the peptide from the resin was done as described in the general method part. The linear peptide **1.23** was used without further purification.

UHR-MS (ESI-TOF) m/z calcd for $C_{47}H_{84}N_5O_{10}Si_2$: 934.5751 $[M+H]^+$; found: 934.5745 $[M+H]^+$

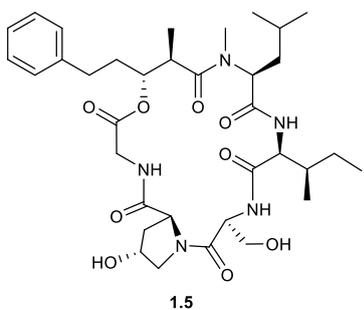
Preparation of (6R,7R,10S,13S,16S,20R,21aS)-13-((R)-sec-butyl)-20-((tert-butylidimethylsilyl)oxy)-16-(((tert-butylidimethylsilyl)oxy)methyl)-10-isobutyl-7,9-dimethyl-6-phenethyltetradecahydropyrrolo[2,1-f][1]oxa[4,7,10,13,16]pentaazacyclononadecine-1,4,8,11,14,17-hexaone (1.32)



Linear peptide **1.23** (0.446 mmol) was dissolved in DCM:DMF (15:1; 256 mL). HATU (571 mg, 1.50 mmol), followed by DIPEA (425 μ L, 2.50 mmol) were added and the mixture was stirred for 16 h at room temperature after which HR-MS showed complete conversion. It was extracted twice with water, once with brine, dried over $MgSO_4$ and concentrated *in vacuo*. Prepurification by column chromatography (silica, PE:EA; 1:2) yielded **1.32** as a colorless oil (169 mg, 0.184 mmol, 41% over 4 steps). R_f (PE:EA; 1:2) = 0.69.

UHR-MS (ESI-TOF) m/z calcd for $C_{47}H_{82}N_5O_9Si_2$: 916.5646 [M+H]⁺; found: 916.5637 [M+H]⁺

Preparation of (6R,7R,10S,13S,16S,20R,21aS)-13-((R)-sec-butyl)-20-hydroxy-16-(hydroxymethyl)-10-isobutyl-7,9-dimethyl-6-phenethyl-tetradecahydropyrrolo[2,1-f][1]oxa[4,7,10,13,16]pentaazacyclonona-decine-1,4,8,11,14,17-hexaone (1.5)



Under argon peptide **1.32** (168 mg, 0.183 mmol) was dissolved in anhyd. THF (15 mL) in a Teflon® round bottom flask. Acetic acid (791 μ L, 13.8 mmol) was added, followed by TBAF (10.0 mL, 34.5 mmol) and the mixture was stirred at room temperature for 22 h. It was diluted with EA, washed twice with saturated, aqueous $NaHCO_3$ solution, once with brine, dried over $MgSO_4$ and concentrated *in vacuo*. Purification by semi preparative HPLC (50-75-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 3 mL/min) yielded **1.5** as a colorless oil (49 mg, 0.071 mmol, 39%).

¹H-NMR (DMSO- d_6 , 600 MHz): δ_H [ppm] = 8.88 (minor, NH Gly), 8.60 (minor, NH Ser), 8.49 (d, 1H, J = 4.2 Hz, NH Ser), 8.21 (d, 1H, J = 9.4 Hz, NH *allo* Ile), 8.01 (d, 1H, J = 8.4 Hz, NH Gly), 7.29-7.14 (m, 8H, CH_{arom} BHA), 6.68 (minor, NH *allo* Ile), 5.14 (ddd, 2H, J = 10.0, 6.0, 3.9 Hz, 3-CH-O BHA), 5.12-5.01 (m, 2H, OH Ser, OH 4-Hyp), 4.67 (dd, 1H, J = 8.5, 4.2 Hz, α -CH 4-Hyp), 4.40

(dd, 1H, $J = 9.5, 4.5$ Hz, α -CH *allo* Ile), 4.27 (d, 1H, $J = 8.5$ Hz, CH₂ a Gly), 4.24 (d, 1H, $J = 8.1$ Hz, CH₂ b Gly), 4.21-4.14 (m, 1H, γ -CH 4-Hyp), 4.02 (ddd, 1H, $J = 7.0, 7.0, 4.4$ Hz, α -CH Ser), 3.62-3.56 (m, 1H, β -CH₂ a Ser), 3.54-3.46 (m, 3H, β -CH₂ b Ser, α -CH *N*-Me Leu), 3.44-3.36 (m, 4H, CH₂ Gly minor, δ -CH₂ a, b 4-Hyp), 3.26-3.21 (m, 2H, 2-CH BHA), 3.12 (s, 3H, N-CH₃ *N*-Me Leu), 2.62-2.57 (m, 2H, 5-CH₂ a, b BHA), 2.26 (ddd, 1H, $J = 12.6, 5.0, 4.8$ Hz, β -CH₂ a 4-Hyp), 2.14-2.04 (m, 3H, β -CH₂ b 4-Hyp, β -CH₂ a *N*-Me Leu), 2.02-1.94 (m, 2H, 4-CH₂ a BHA), 1.83-1.71 (m, 3H, 4-CH₂ b BHA, β -CH *allo* Ile), 1.61 (ddd, 1H, $J = 13.9, 8.2, 5.9$ Hz, β -CH₂ b *N*-Me Leu), 1.48-1.40 (m, 1H, γ -CH *N*-Me Leu), 1.40-1.32 (m, 1H, γ -CH₂ a *allo* Ile), 1.11-1.04 (m, 1H, γ -CH₂ b *allo* Ile), 1.03 (d, 3H, $J = 7.1$ Hz, 2'-CH₃ BHA), 0.88 (d, 6H, $J = 6.3$ Hz, δ -CH₃ a, b *N*-Me Leu), 0.86 (t, 3H, $J = 7.5$ Hz, δ -CH₃ *allo* Ile), 0.70 (d, 3H, $J = 6.9$ Hz, γ -CH₃ *allo* Ile).

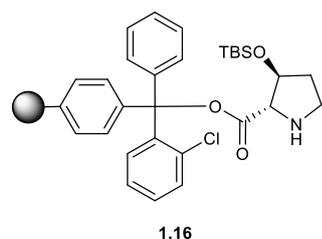
¹³C-NMR (DMSO-*d*₆, 100 MHz): δ_c [ppm] = 176.4 (CO BHA), 171.7 (CO *allo* Ile), 171.0 (CO *N*-Me Leu), 170.7 (CO 4-Hyp), 169.7 (CO Ser), 168.5 (CO Gly), 141.7 (C_{quart} BHA), 128.3, 128.2, 125.8 (CH_{arom} BHA), 75.8 (3-CH-O BHA), 69.0 (α -CH *N*-Me Leu), 66.8 (γ -CH 4-Hyp), 60.7 (β -CH₂ Ser), 59.3 (α -CH 4-Hyp), 54.3 (α -CH Ser), 53.9 (α -CH *allo* Ile), 53.5 (δ -CH₂ 4-Hyp), 41.4 (CH₂ Gly), 40.0 (N-CH₃ *N*-Me Leu), 39.7* (2-CH BHA), 38.7* (β -CH₂ 4-Hyp), 38.1 (β -CH₂ *N*-Me Leu), 37.7 (β -CH *allo* Ile), 32.6 (4-CH₂ BHA), 29.5 (5-CH₂ BHA), 25.8 (γ -CH₂ *allo* Ile), 24.6 (γ -CH *N*-Me Leu), 22.8 (δ -CH₃ a *N*-Me Leu), 21.7 (δ -CH₃ b *N*-Me Leu), 14.7 (2'-CH₃ BHA), 13.9 (γ -CH₃ *allo* Ile), 11.6 (δ -CH₃ *allo* Ile).

Additional found signals: δ_H [ppm] = 8.23 (FA), 5.75 (DCM), 3.31 (H₂O), 2.54 (DMSO). δ_c [ppm] = 40.4 (DMSO)

UHR-MS (ESI-TOF) m/z calcd for C₃₅H₅₄N₅O₉: 688.3916 [M+H]⁺; found: 688.3920 [M+H]⁺

Specific rotation $[\alpha]_D^{20} = -104.6^\circ$ ($c = 0.98$, CHCl₃)

Preparation of 2CT-3-Hyp(TBS)-NH (1.16)

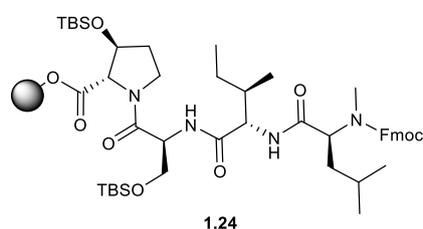


2-Chlorotriptyl chloride resin ($n = 1.55$ mmol/g, 6.01 g, 9.32 mmol) was swelled in DCM (50 mL) for 10 min. After removal of the solvent, the solution of Fmoc-protected amino acid **1.12** (5.62 g, 12.0 mmol) and DIPEA (5.0 mL, 29.4 mmol) in DCM (60 mL) was added and agitated for 1.5 h. The solvent was filtered off and the

resin was washed twice with DMF. The capping mixture (60 mL), consisting of

DCM:MeOH:DIPEA (80:15:5), was added and it was agitated for 10 min. The mixture was drained and the capping was repeated once more. The resin was washed three times with DMF and a sample was taken for determination of the loading. Fmoc-deprotection was done with 20% piperidine in DMF (30 mL) for 3 min and the process was carried out five times total. The resin was washed six times with DMF, three times with isopropanol and three times with *n*-heptane. It was sucked dry and then dried further *in vacuo* for 24 h. Resin **1.16** (6.80 g) was stored under argon at 4 °C. The loading was determined by absorption to be $n = 0.353$ mmol/g (23%).

Preparation of 2CT-3-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-Fmoc (1.24)



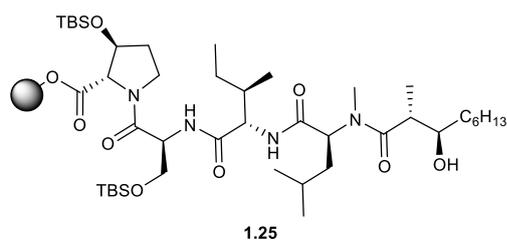
Resin **1.16** ($n = 0.353$ mmol/g, 6.00 g, 2.12 mmol) was swelled in DMF for 30 min. The solvent was removed, and a solution of Fmoc-L-serine (TBS)-OH (2.812 g, 6.368 mmol) and HATU (2.338 g, 6.149 mmol) in DMF was added, followed by DIPEA (2160 μ L, 12.70 mmol) and more DMF.

The mixture was agitated for 1.5 h. After Fmoc-deprotection, Fmoc-L-*allo*-isoleucine-OH (2.244 g, 6.350 mmol) and HATU (2.340 g, 6.154 mmol) were added, dissolved in a small volume of DMF, followed by DIPEA (2160 μ L, 12.70 mmol) and more DMF. The mixture was agitated for 1.5 h. After Fmoc-deprotection, Fmoc-N-Me-L-leucine-OH (2.340 g, 6.368 mmol) and HATU (2.343 g, 6.162 mmol) were dissolved in a small quantity of DMF, and added to the resin. DIPEA (2160 μ L, 12.70 mmol) and more DMF were added and the mixture was agitated for 3 h. The supernatant was drained, the resin was washed five times with DMF, three times each with DCM, methanol and diethyl ether. It was sucked dry and dried further *in vacuo* for 12 h. The finished resin **1.24** (8.65 g, $n = 0.245$ mmol/g,) was stored under argon at 4 °C.

UHR-MS (ESI-TOF) m/z calcd for $C_{48}H_{77}N_4O_9Si_2$: 909.5224 $[M+H]^+$; found: 909.5223 $[M+H]^+$

Preparation of 2CT-3-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-alkyl-BHA (1.25)

Resin **1.24** ($n = 0.245$ mmol/g, 2.05 g, 0.502 mmol) was swelled in DMF for 35 min. The supernatant was drained and after Fmoc-deprotection the resin was washed twice with

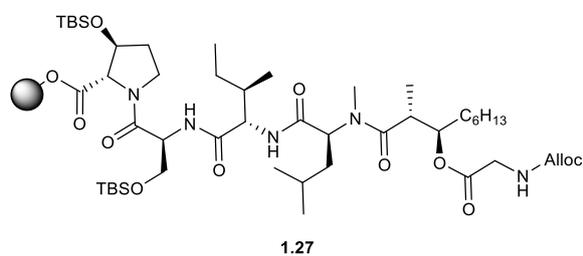


anhyd. DMF. It was suspended in a small amount of anhyd. DMF. Acid **1.8** (284 mg, 1.51 mmol) and triethylamine (208 μ L, 1.50 mmol) were dissolved in anhyd. DMF and cooled to 0 °C. DEPC (90%, 252 μ L, 1.49 mmol) was added, it was directly transferred to

the peptide vessel and the mixture was agitated for 16 h. Due to incomplete conversion, as indicated by LC-MS, the step was repeated. The mixture was filtered off, the resin was washed twice with anhy. DMF and suspended in a small amount of anhyd. DMF. Acid **1.8** (102 mg, 0.542 mmol) and triethylamine (69 μ L, 0.50 mmol) were dissolved in anhy. DMF and cooled to 0 °C. DEPC (90%, 84 μ L, 0.496 mmol) was added and it was transferred to the peptide vessel. The mixture was agitated for 2.5 h, after which HR-MS indicated almost full conversion. The supernatant was drained, and the resin was washed five times each with DMF and DCM. The resin coupled peptide **1.25** was directly used in the next step.

UHR-MS (ESI-TOF) m/z calcd for $C_{43}H_{85}N_4O_9Si_2$: 857.5850 $[M+H]^+$; found: 857.5864 $[M+H]^+$

Preparation of 2CT-3-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-alkyl-BHA-Gly-alloc (**1.27**)

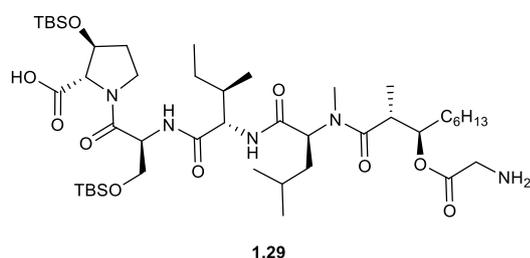


Resin coupled peptide **1.25** (0.502 mmol) was swelled in THF for 10 min, filtered and suspended in anhyd. THF. The solution of Alloc-glycine-OH (242 mg, 1.52 mmol), in anhyd. THF was added, followed by DIC

(232 μ L, 1.50 mmol) and DMAP (19 mg, 0.16 mmol). The resin was agitated for 17 h, after which the step was repeated. The supernatant was drained, the resin was washed once with anhyd. THF and then suspended in anhyd. THF. Alloc-glycine-OH (258 mg, 1.62 mmol), dissolved in anhy. THF was added, followed by DIC (260 μ L, 1.68 mmol) and DMAP (19 mg, 0.16 mmol). The mixture was agitated for 90 min after which HR-MS indicated full conversion. The supernatant was removed and the resin was washed once with DCM. The resin coupled peptide **1.27** was directly used in the next step.

UHR-MS (ESI-TOF) m/z calcd for $C_{49}H_{92}N_5O_{12}Si_2$: 998.6276 $[M+H]^+$; found: 998.6273 $[M+H]^+$

Preparation of HOOC-3-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-alkyl-BHA-Gly-NH₂ (**1.29**)

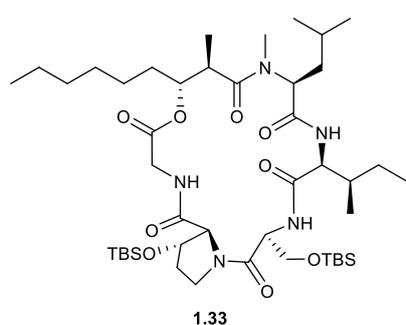


Resin coupled peptide **1.27** (0.502 mmol) was swelled in DCM for 10 min. It was filtered, washed with anhyd. DCM and then suspended in anhyd. DCM. Pd(PPh₃)₄ (58 mg, 0.050 mmol) was added, followed by Phenylsilane (1481 μ L, 12.00 mmol).

The mixture was agitated for 3 h, after which HR-MS indicated full conversion. The mixture was filtered and washed five times with DCM until the solvent ran clear. Cleavage of the peptide from the resin was performed as described in the general method part. Linear peptide **1.29** was used without further purification.

UHR-MS (ESI-TOF) m/z calcd for C₄₅H₈₈N₅O₁₀Si₂: 914.6064 [M+H]⁺; found: 914.6066 [M+H]⁺

Preparation of (6*R*,7*R*,10*S*,13*S*,16*S*,21*S*,21*aS*)-13-((*R*)-*sec*-butyl)-21-((*tert*-butyldimethylsilyl)oxy)-16-(((*tert*-butyldimethylsilyl)oxy)methyl)-6-hexyl-10-isobutyl-7,9-dimethyltetradecahydropyrrolo[2,1-*f*][1]oxa[4,7,10,13,16]pentaazacyclononadecine-1,4,8,11,14,17-hexaone (**1.33**)

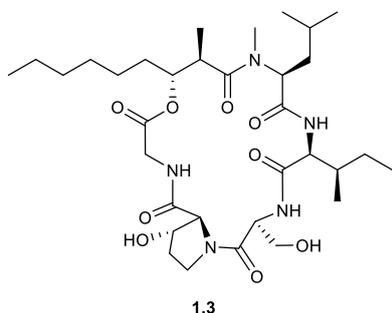


Linear peptide **1.29** (0.502 mmol) was dissolved in DCM:DMF (15:1; 256 mL). HATU (572 mg, 1.50 mmol), followed by DIPEA (425 μ L, 2.50 mmol) were added and the mixture was stirred for 3 h at room temperature, after which HR-MS indicated complete conversion. It was extracted twice with water, once with brine, dried over MgSO₄ and concentrated *in vacuo*.

Prepurification by column chromatography (silica, PE:EA; 1:2) yielded **1.33** as slightly yellow needles (125 mg, 0.139 mmol, 28% over 4 steps). R_f (PE:EA; 1:2) = 0.72.

UHR-MS (ESI-TOF) m/z calcd for C₄₅H₈₆N₅O₉Si₂: 896.5959 [M+H]⁺; found: 896.5956 [M+Na]⁺

Preparation of (6R,7R,10S,13S,16S,21S,21aS)-13-((R)-sec-butyl)-6-hexyl-21-hydroxy-16-(hydroxymethyl)-10-isobutyl-7,9-dimethyltetradecahydro-pyrrolo[2,1-f][1]oxa[4,7,10,13,16]pentaazacyclononadecine-1,4,8,11,14,17-hexaone (1.3)



Under argon peptide **1.33** (125 mg, 0.139 mmol) was dissolved in anhyd. THF (10 mL) in a Teflon® round bottom flask. Acetic acid (601 μ L, 10.5 mmol) was added, followed by TBAF (7.5 mL, 26 mmol) and the mixture was stirred at room temperature for 9 h. It was diluted with EA, washed twice with saturated, aqueous NaHCO₃ solution, once with brine,

dried over MgSO₄ and concentrated *in vacuo*. Purification by semi preparative HPLC (50-75-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 3 mL/min) yielded **1.3** as a colorless solid (47 mg, 0.070 mmol, 51%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 8.79 (minor, NH Gly), 8.69 (minor, NH Ser), 8.56 (d, 1H, *J* = 4.2 Hz, NH Ser), 8.22 (d, 1H, *J* = 9.4 Hz, NH *allo* Ile), 8.10 (d, 1H, *J* = 8.6 Hz, NH Gly), 6.62 (minor, NH *allo* Ile), 5.33 (d, 1H, *J* = 3.6 Hz, OH 3-Hyp), 5.06 (ddd, 1H, *J* = 10.0, 6.3, 3.7 Hz, 3-CH-O BHA), 4.52-4.47 (m, 1H, β -CH 3-Hyp), 4.46-4.42 (m, 1H, α -CH 3-Hyp), 4.42 (dd, 1H, *J* = 9.2, 3.2 Hz, α -CH *allo* Ile), 4.28 (d, 1H, *J* = 8.8 Hz, CH₂ a Gly), 4.25 (d, 1H, *J* = 8.8 Hz, CH₂ b Gly), 3.95 (ddd, 1H, *J* = 7.3, 6.3, 4.4 Hz, α -CH Ser), 3.61-3.55 (m, 1H, β -CH₂ a Ser), 3.55-3.45 (m, 3H, β -CH₂ b Ser, α -CH *N*-Me Leu, δ -CH₂ a 3-Hyp), 3.45-3.35 (m, 2H, δ -CH₂ b 3-Hyp, CH₂ Gly minor), 3.16-3.10 (m, 1H, 2-CH BHA), 3.09 (s, 3H, N-CH₃ *N*-Me Leu), 2.09 (ddd, 1H, *J* = 13.7, 9.4, 5.3 Hz, β -CH₂ a *N*-Me Leu), 1.81-1.70 (m, 4H, β -CH *allo* Ile, γ -CH₂ a, b 3-Hyp), 1.70-1.64 (m, 1H, 4-CH₂ a BHA), 1.61 (ddd, 1H, *J* = 14.0, 8.2, 6.1 Hz, β -CH₂ b *N*-Me Leu), 1.51-1.44 (m, 1H, 4-CH₂ b BHA), 1.44-1.39 (m, 1H, γ -CH *N*-Me Leu), 1.39-1.32 (m, 1H, γ -CH₂ a *allo* Ile), 1.29-1.18 (m, 11H, CH₂ BHA), 1.07 (ddd, 1H, *J* = 13.7, 7.5, 7.5 Hz, γ -CH₂ b *allo* Ile), 0.99 (d, 3H, *J* = 6.9 Hz, 2'-CH₃ BHA), 0.89-0.83 (m, 15H, δ -CH₃ *allo* Ile, 9-CH₃ BHA, δ -CH₃ a, b *N*-Me Leu), 0.70 (d, 3H, *J* = 6.9 Hz, γ -CH₃ *allo* Ile).

¹³C-NMR (DMSO-*d*₆, 100 MHz): δ_{C} [ppm] = 176.4 (CO BHA), 172.0 (CO *allo* Ile), 170.9 (CO *N*-Me Leu), 169.5 (CO 3-Hyp), 168.5 (CO Gly), 168.3 (CO Ser), 75.9 (3-CH BHA), 74.0 (β -CH 3-Hyp), 69.3 (α -CH 3-Hyp), 69.1* (α -CH *N*-Me Leu), 60.4 (β -CH₂ Ser), 55.3 (α -CH Ser), 53.9 (α -CH *allo* Ile), 44.8 (δ -CH₂ 3-Hyp), 41.3 (CH₂ Gly), 40.1* (2-CH BHA, N-CH₃), 38.1 (β -CH₂ *N*-Me Leu), 37.8 (β -CH *allo* Ile), 31.0 (CH₂ BHA), 30.6 (4-CH₂ BHA), 30.2 (γ -CH₂ 3-Hyp), 28.7 (CH₂ BHA), 25.8 (γ -

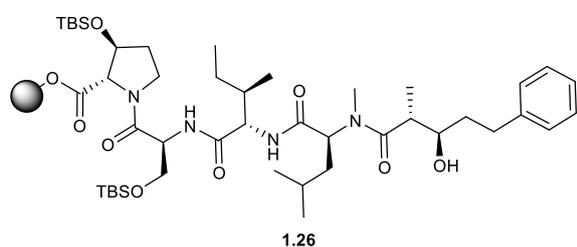
CH₂ *allo* Ile), 24.5 (γ -CH *N*-Me Leu), 23.2 (CH₂ BHA), 22.8 (δ -CH₃ *a N*-Me Leu), 22.0 (CH₂ BHA), 21.7 (δ -CH₃ *b N*-Me Leu), 14.7 (2'-CH₃ BHA), 13.9 (γ -CH₃ *allo* Ile, 9-CH₃ BHA), 11.6 (δ -CH₃ *allo* Ile).

Additional found signals: δ_{H} [ppm] = 3.30 (H₂O), 2.54 (DMSO). δ_{C} [ppm] = 40.4 (DMSO)

UHR-MS (ESI-TOF) m/z calcd for C₃₃H₅₈N₅O₉: 668.4229 [M+H]⁺; found: 668.4232 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{20} = -59.3^{\circ}$ ($c = 1.08$, CHCl₃)

Preparation of 2CT-3-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-arom-BHA (1.26)

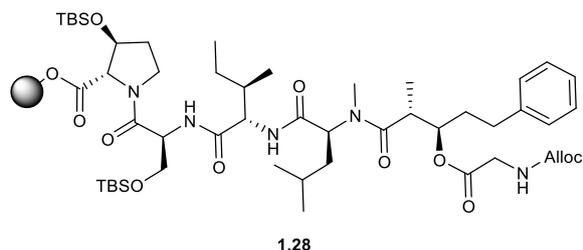


Resin **1.24** ($n = 0.245$ mmol/g, 2.06 g, 0.505 mmol) was swelled in DMF for 1 h. The supernatant was drained and after Fmoc-deprotection the resin was washed twice with anhyd. DMF. The resin was suspended in a

small amount of anhyd. DMF. The solution of acid **1.9** (206 mg, 0.989 mmol) and triethylamine (213 μ L, 1.54 mmol) in anhy. DMF and was cooled to 0 °C. DEPC (90%, 258 μ L, 1.52 mmol) was added, it was directly transferred to the peptide vessel and the mixture was agitated for 19 h. The step was repeated due to incomplete conversion, as indicated by LC-MS result. Acid **1.9** (160 mg, 0.768 mmol) and triethylamine (139 μ L, 1.00 mmol) were dissolved in anhyd. DMF and cooled to 0 °C. DEPC (90%, 168 μ L, 0.992 mmol) was added and it was transferred to the peptide vessel and existing mixture. It was agitated for 8 h, after which HR-MS indicated almost full conversion. The supernatant was drained and the resin was washed five times each with DMF and DCM. The resin coupled peptide **1.26** was directly used in the next step.

UHR-MS (ESI-TOF) m/z calcd for C₄₅H₈₁N₄O₉Si₂: 877.5537 [M+H]⁺; found: 877.5560 [M+H]⁺

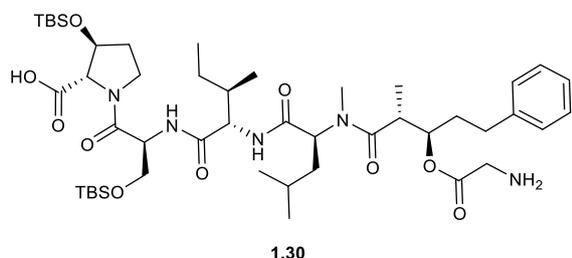
Preparation of 2CT-3-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-arom-BHA-Gly-alloc (**1.28**)



Resin coupled peptide **1.26** (0.505 mmol) was swelled in THF for 10 min, filtered and suspended in anhyd. THF. Alloc-glycine-OH (242 mg, 1.52 mmol), dissolved in anhyd. THF was added, followed by DIC (232 μ L, 1.50 mmol) and DMAP (21.4 mg, 0.175 mmol). The resin was agitated for 7 h, after which the step was repeated. The supernatant was filtered off, the resin was washed once with anhyd. THF and then suspended in anhyd. THF. Alloc-glycine-OH (477 mg, 3.00 mmol), dissolved in anhyd. THF was added to the resin, followed by DIC (465 μ L, 3.00 mmol) and DMAP (36.6 mg, 0.300 mmol). The mixture was agitated for 4 h, after which HR-MS indicated almost full conversion. The supernatant was drained and the resin was washed once with DCM. The resin coupled peptide **1.28** was directly used in the next step.

UHR-MS (ESI) m/z calcd for $C_{51}H_{88}N_5O_{12}Si_2$: 1018.5963 $[M+H]^+$; found: 1018.5959 $[M+H]^+$

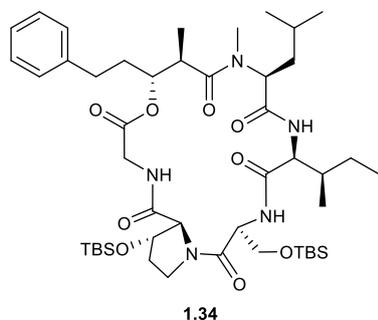
Preparation of HOOC-3-Hyp(TBS)-L-Ser(TBS)-L-*allo*-Ile-N-Me-L-Leu-arom-BHA-Gly-NH₂ (**1.30**)



Resin coupled peptide **1.28** (0.505 mmol) was swelled in DCM for 40 min. It was filtered, washed with anhyd. DCM and then suspended in anhyd. DCM. Pd(PPh₃)₄ (59 mg, 0.051 mmol) was added, followed by Phenylsilane (1481 μ L, 12.00 mmol). The mixture was agitated for 3 h, after which HR-MS indicated full conversion. The mixture was filtered and washed five times with DCM until the solvent ran clear. Cleavage of the peptide from the resin was performed as described in the general method part. Linear peptide **1.30** was used without further purification.

UHR-MS (ESI-TOF) m/z calcd for $C_{47}H_{84}N_5O_{10}Si_2$: 934.5751 $[M+H]^+$; found: 934.5753 $[M+H]^+$

Preparation of (6R,7R,10S,13S,16S,21S,21aS)-13-((R)-sec-butyl)-21-((tert-butyl)dimethylsilyloxy)-16-(((tert-butyl)dimethylsilyloxy)methyl)-10-isobutyl-7,9-dimethyl-6-phenethyltetradecahydropyrrolo[2,1-f][1]oxa[4,7,10,13,16]pentaazacyclononadecine-1,4,8,11,14,17-hexaone (1.34)

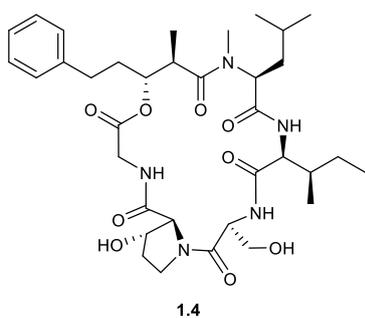


Linear peptide **1.30** (0.505 mmol) was dissolved in DCM:DMF (15:1; 256 mL). HATU (573 mg, 1.51 mmol), followed by DIPEA (425 μ L, 2.50 mmol) were added and the mixture was stirred for 17 h at room temperature, after which HR-MS indicated complete conversion. It was extracted twice with water, once with brine, dried over MgSO_4 and concentrated *in vacuo*.

Prepurification by column chromatography (silica, PE:EA; 1:2) yielded **1.34** as a yellowish foam (376 mg, 0.419 mmol, 83% over 4 steps). R_f (PE:EA; 1:2) = 0.71.

UHR-MS (ESI-TOF) m/z calcd for $\text{C}_{47}\text{H}_{82}\text{N}_5\text{O}_9\text{Si}_2$: 916.5646 $[\text{M}+\text{H}]^+$; found: 916.5659 $[\text{M}+\text{H}]^+$

Preparation of (6R,7R,10S,13S,16S,21S,21aS)-13-((R)-sec-butyl)-21-hydroxy-16-(hydroxymethyl)-10-isobutyl-7,9-dimethyl-6-phenethyl-tetradecahydropyrrolo[2,1-f][1]oxa[4,7,10,13,16]pentaazacyclonona-decine-1,4,8,11,14,17-hexaone (1.4)



Under argon peptide **1.34** (376 mg, 0.419 mmol) was dissolved in anhyd. THF (25 mL) in a Teflon[®] round bottom flask. Acetic acid (1760 μ L, 30.75 mmol) was added, followed by TBAF (22.1 mL, 76.3 mmol) and the mixture was stirred at room temperature for 18 h. It was diluted with EA, washed twice with saturated, aqueous NaHCO_3 solution, once with brine, dried

over MgSO_4 and concentrated *in vacuo*. Purification by semi preparative HPLC (50-75-95% AcCN + 0.1% FA, NUCLEODUR[®] C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 3 mL/min) yielded **1.4** as a colorless oil (83 mg, 0.12 mmol, 29%).

$^1\text{H-NMR}$ (DMSO- d_6 , 600 MHz): δ_{H} [ppm] = 8.90 (minor, NH Gly), 8.77 (minor, NH Ser), 8.58 (d, 1H, J = 4.1 Hz, NH Ser), 8.23 (d, 1H, J = 9.4 Hz, NH *allo* Ile), 8.05 (d, 1H, J = 8.7 Hz, NH Gly), 7.29-7.14 (m, 6H, CH_{arom} BHA), 6.62 (minor, NH *allo* Ile), 5.15 (ddd, 1H, J = 9.9, 6.2, 4.0 Hz, 3-CH-O BHA), 4.49 (d, 1H, J = 3.8 Hz, β -CH 3-Hyp), 4.45-4.41 (m, 2H, α -CH 3-Hyp, α -CH *allo* Ile), 4.30

(d, 1H, $J = 8.5$ Hz, CH_2 a Gly), 4.27 (d, 1H, $J = 8.9$ Hz, CH_2 b Gly), 3.96 (ddd, 1H, $J = 6.4, 6.4, 4.7$ Hz, α -CH Ser), 3.61-3.55 (m, 1H, β - CH_2 a Ser), 3.55-3.46 (m, 3H, β - CH_2 b Ser, δ - CH_2 a 3-Hyp, α -CH *N*-Me Leu), 3.46-3.39 (m, 1H, δ - CH_2 b 3-Hyp), 3.39-3.34 (m, 2H, CH_2 Gly, minor), 3.25-3.19 (m, 2H, 2-CH BHA), 3.12 (s, 3H, N - CH_3 *N*-Me Leu), 2.62-2.56 (m, 2H, 5- CH_2 a,b BHA), 2.10 (ddd, 1H, $J = 13.9, 9.2, 5.0$ Hz, β - CH_2 a *N*-Me Leu), 2.01-1.93 (m, 1H, 4- CH_2 a BHA), 1.83-1.69 (m, 5H, 4- CH_2 b BHA, γ - CH_2 a, b 3-Hyp, β -CH *allo* Ile), 1.62 (ddd, 1H, $J = 13.9, 8.2, 6.2$ Hz, β - CH_2 b *N*-Me Leu), 1.46-1.40 (m, 1H, γ -CH *N*-Me Leu), 1.40-1.33 (m, 1H, γ - CH_2 a *allo* Ile), 1.11-1.05 (m, 1H, γ - CH_2 b *allo* Ile), 1.03 (d, 3H, $J = 6.8$ Hz, 2'- CH_3 BHA), 0.92-0.84 (m, 12H, δ - CH_3 *allo* Ile, δ - CH_3 a, b *N*-Me Leu), 0.70 (d, 3H, $J = 6.9$ Hz, γ - CH_3 *allo* Ile).

^{13}C -NMR (DMSO- d_6 , 100 MHz): δ_c [ppm] = 176.4 (CO BHA), 172.0 (CO *allo* Ile), 170.9 (CO *N*-Me Leu), 169.5 (CO Ser), 168.5 (CO Gly), 168.4 (CO 3-Hyp), 141.7 (C_{quart} BHA), 128.3, 128.2, 125.7 (CH_{arom} BHA), 75.8 (3-CH-O BHA), 74.1 (β -CH 3-Hyp), 69.3 (α -CH 3-Hyp), 69.0 (α -CH *N*-Me Leu), 60.4 (β - CH_2 Ser), 55.3 (α -CH Ser), 53.9 (α -CH *allo* Ile), 44.8 (δ - CH_2 3-Hyp), 41.3 (CH_2 Gly), 40.0* (N - CH_3 *N*-Me Leu), 39.7* (2-CH BHA), 38.1 (β - CH_2 *N*-Me Leu), 37.7 (β -CH *allo* Ile), 32.6 (4- CH_2 BHA), 30.2 (γ - CH_2 3-Hyp), 29.5 (5- CH_2 BHA), 25.9 (γ - CH_2 *allo* Ile), 24.6 (γ -CH *N*-Me Leu), 22.8 (δ - CH_3 a *N*-Me Leu), 21.7 (δ - CH_3 b *N*-Me Leu), 14.7 (2'- CH_3 BHA), 13.9 (γ - CH_3 *allo* Ile), 11.6 (δ - CH_3 *allo* Ile).

Additional found signals: δ_H [ppm] = 8.27 (FA), 5.75 (DCM), 3.34-3.25 (H_2O), 2.54 (DMSO), *OH* Ser and *OH* 3-Hyp were not observed. δ_c [ppm] = 40.4 (DMSO)

UHR-MS (ESI-TOF) m/z calcd for $C_{35}H_{54}N_5O_9$: 688.3916 [M+H] $^+$; found: 688.3917 [M+H] $^+$

Specific rotation $[\alpha]_D^{20} = -52.5^\circ$ ($c = 0.99$, $CHCl_3$)

4.4.3 Activity Tests - Bioassays

Determination of the minimum inhibitory concentrations (MIC) of purified compounds **1.2**, **1.3**, **1.4** and **1.5** were done by micro broth dilution assays in 96 well plates. Each compound was dissolved in DMSO to a 6.4 mg/mL stock solution and tested in triplicates.

For the bacteria *E. coli* ATCC 25922, *E. coli* ATCC 25922 $\Delta tolC$, *P. aeruginosa* PA01, *P. aeruginosa* PA0750, *K. pneumoniae* ATCC 30104, *A. baumannii* ATCC 19606, *S. aureus* ATCC 33592 an overnight culture (37°C, 180rpm) was diluted to $5 \cdot 10^5$ cells/mL in cation

adjusted Mueller Hinton II medium (Becton, Dickinson and Company). For *E. coli* ATCC 25922 MHC the overnight culture was cultivated in regular cation adjusted Mueller Hinton II medium, density adjustment and dilution assays were performed with cation adjusted Mueller Hinton II medium supplemented with NaHCO₃ 44 mM. The pre culture of *M. smegmatis* ATCC 607 was incubated in Brain-Heart Infusion broth (Becton, Dickinson and Company, 48 h, 37°C, 180 rpm) before the cell concentration was adjusted in Mueller Hinton II medium. As positive controls, dilution series (64-0.03 µg/mL) of rifampicin, tetracycline and gentamycin were prepared, except for *M. smegmatis* where isoniazid was used instead of gentamycin as third positive control and for *E. coli* ATCC 25922 $\Delta hldE$, where ciprofloxacin and colistin were used as controls. As negative controls pure cell suspensions were used. After incubation cell growth was assessed by measuring the turbidity with a microplate spectrophotometer at 600 nm (LUMIstar® Omega BMG Labtech). For *M. smegmatis* cell viability was evaluated after 48 h (37°C, 180 rpm, 85% rH) via ATP quantification (BacTiter-Glo™, Promega) according to the manufacturer's instructions.

4.5 Supporting Information

4.5.1 Chromatograms and spectra of natural extract

Table S 4-1: Overview of literature known congeners of Globomycin (1) and new isomers (blue).

compound	MF	[M+H] ⁺ <i>m/z</i>	known structures	new isomers	proposed new structures
Globomycin (1.1)					
SF-1902 A ₁	C ₃₂ H ₅₇ N ₅ O ₉	656.4229	1	0	0
SF-1902 A ₂	C ₃₀ H ₅₃ N ₅ O ₉	628.3916	1	3	SF-1902 A _{2b}
SF-1902 A ₃	C ₃₁ H ₅₅ N ₅ O ₉	642.4073	1	3	SF-1902 A _{3b}
SF-1902 A ₄	C ₃₃ H ₅₉ N ₅ O ₉	670.4386	2	6	SF-1902 A _{4a2} or SF-1902 A _c
SF-1902 A ₅	C ₃₄ H ₆₁ N ₅ O ₉	684.4542	1	1	0
SF-1902 A ₆	C ₂₉ H ₅₁ N ₅ O ₉	614.3760	0	3	SF-1902 A _{6a} SF-1902 A _{6b}
SF-1902 A ₇	C ₂₈ H ₄₉ N ₅ O ₉	600.3603	0	3	SF-1902 A _{7a}
SF-1902 A ₈	C ₂₇ H ₄₇ N ₅ O ₉	586.3447	0	1	SF-1902 A ₈
total			6	20	7

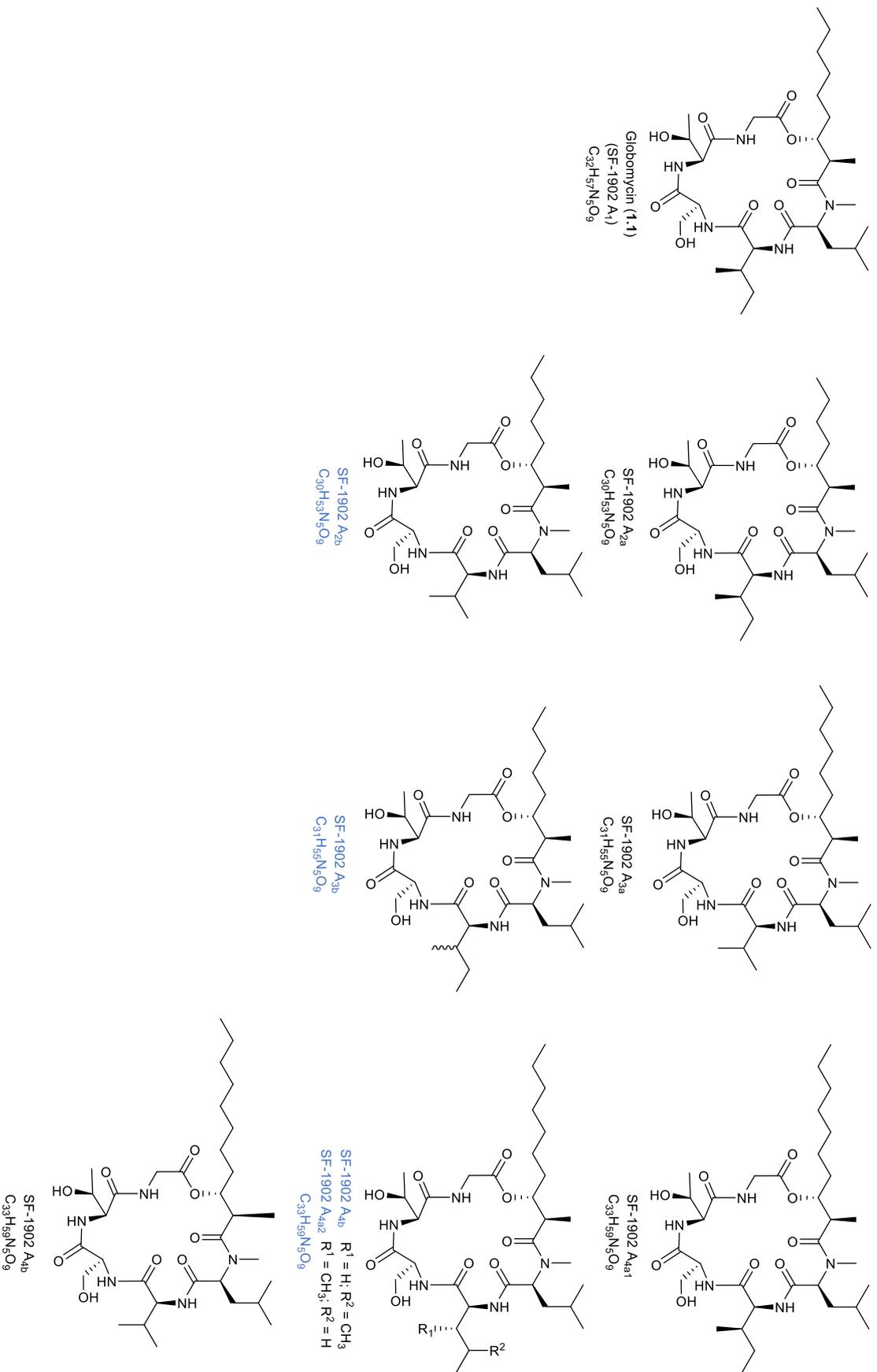


Figure S 4-1: Overview of Globomycin and natural occurring congeners SF-1902 A₁-A₄. New found congeners are shown in blue.

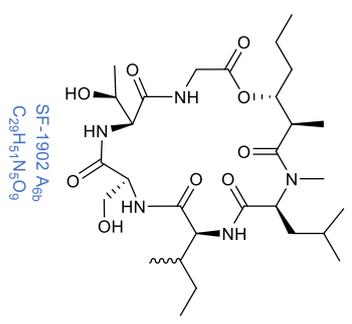
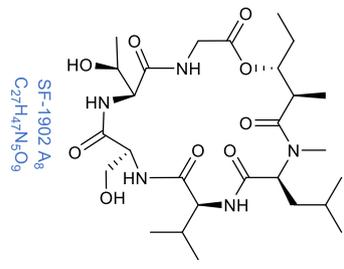
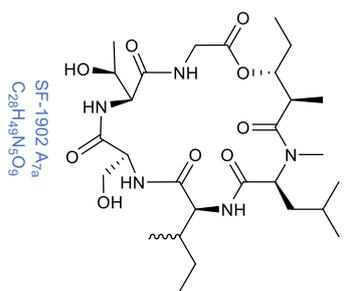
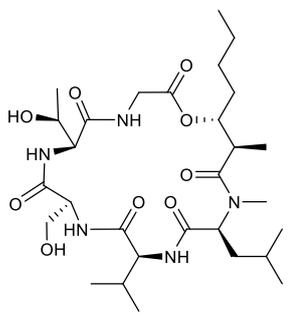
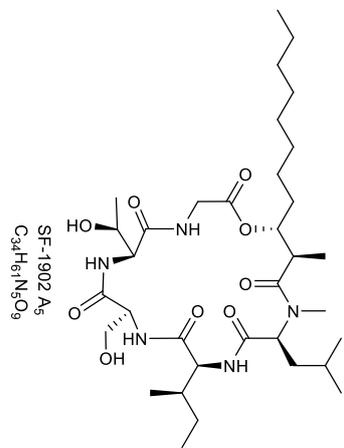


Figure S 4-2: Overview of Globomycin structures and natural occurring congeners SF-1902 A₅-A₈. Novel found congeners are shown in blue.

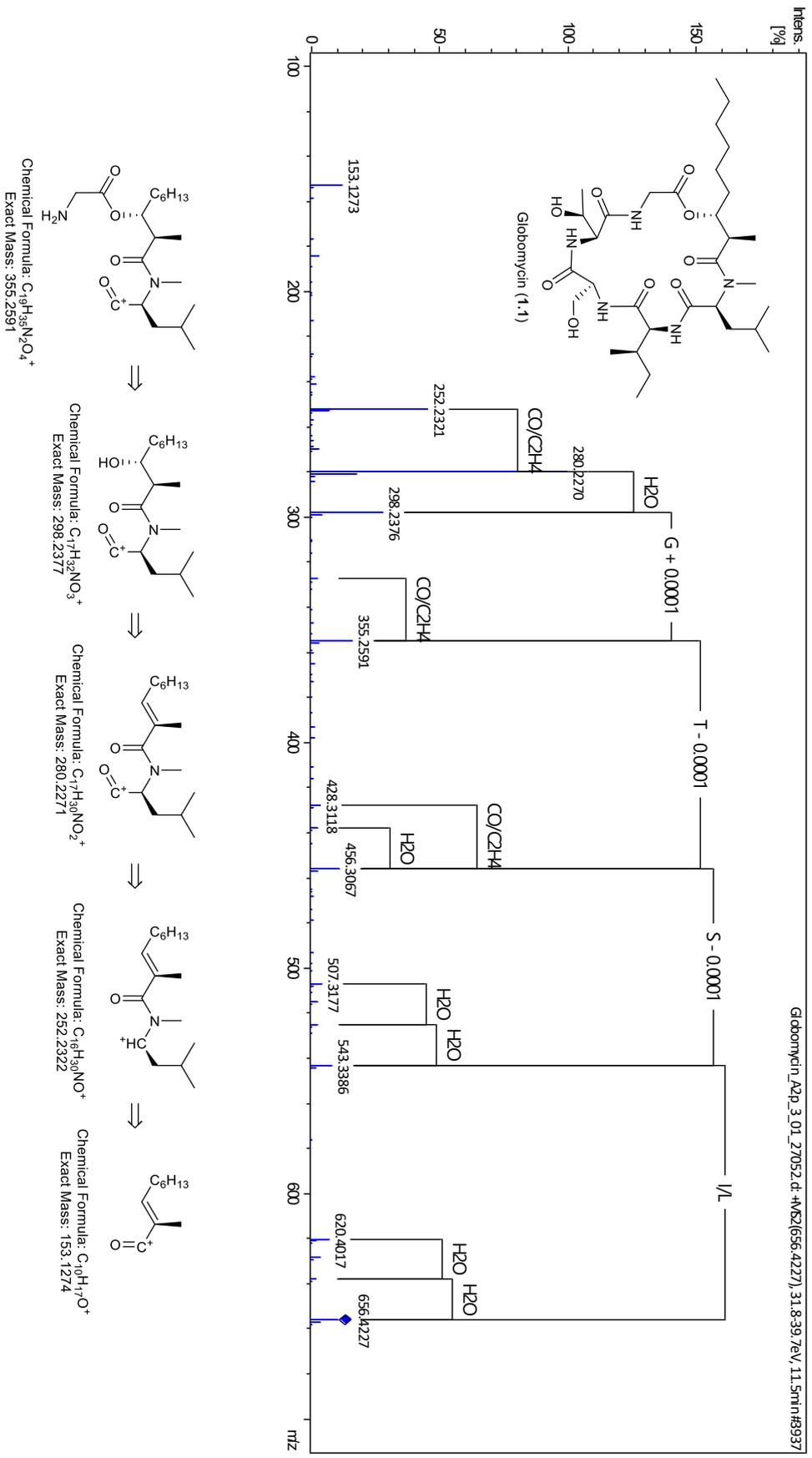


Figure S 4-3: Fragmentation pattern of Globomycin (1.1).

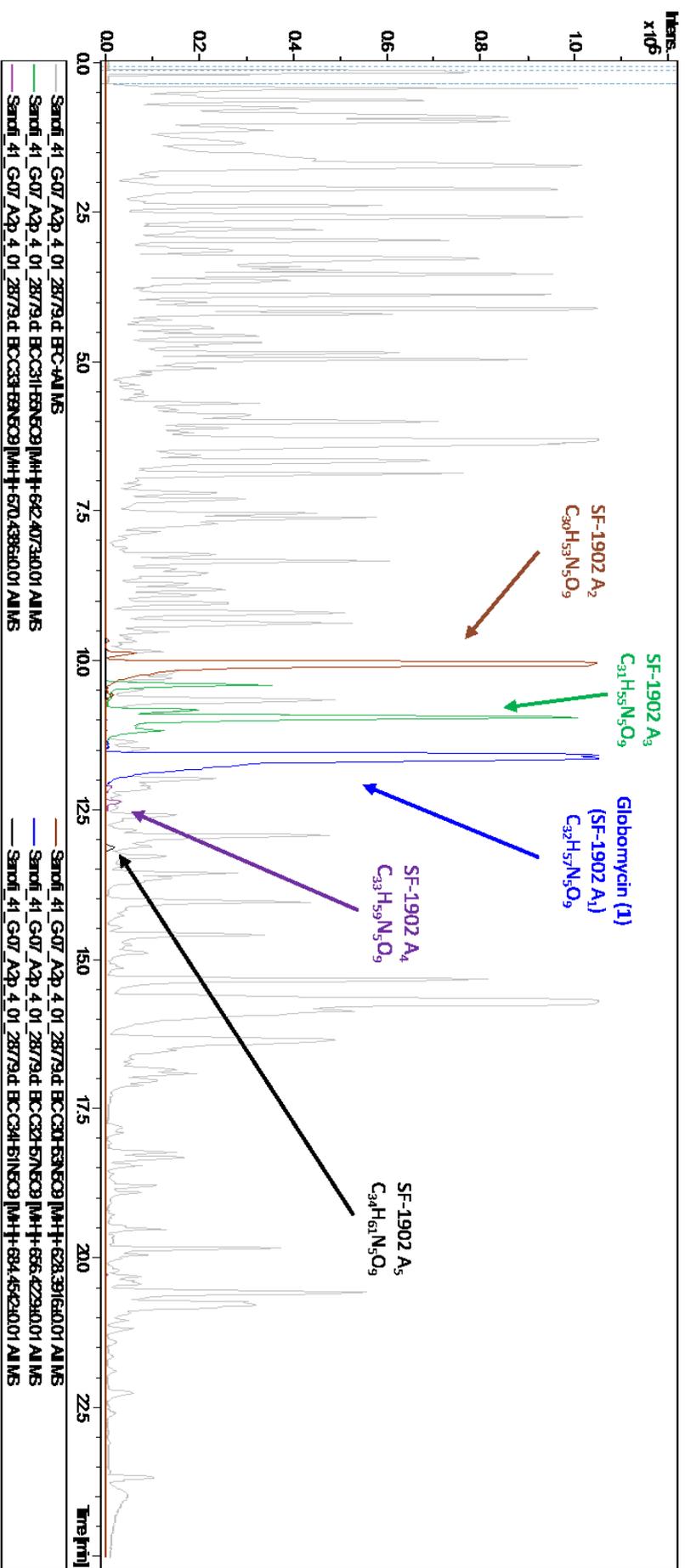


Figure S 4-4: Chromatogram of HAG010519 extract, with extracted chromatograms of literature known congeners of Globomycin. Grey: BPC; brown: EIC of SF-1902 A₂ (628.3916 ± 0.005), green: EIC of SF-1902 A₃ (642.4073 ± 0.005), blue: EIC of Globomycin (1.1) (656.4229 ± 0.005), purple: EIC of SF-1902 A₄ (670.4386 ± 0.005) and black: EIC of SF-1902 A₅ (684.4542 ± 0.005). For SF-1902 A₂ and A₃ a total of 4 isomers each were found, for SF-1902 A₄ a total of 5 isomers were found and for SF-1902 A₅ just the known one.

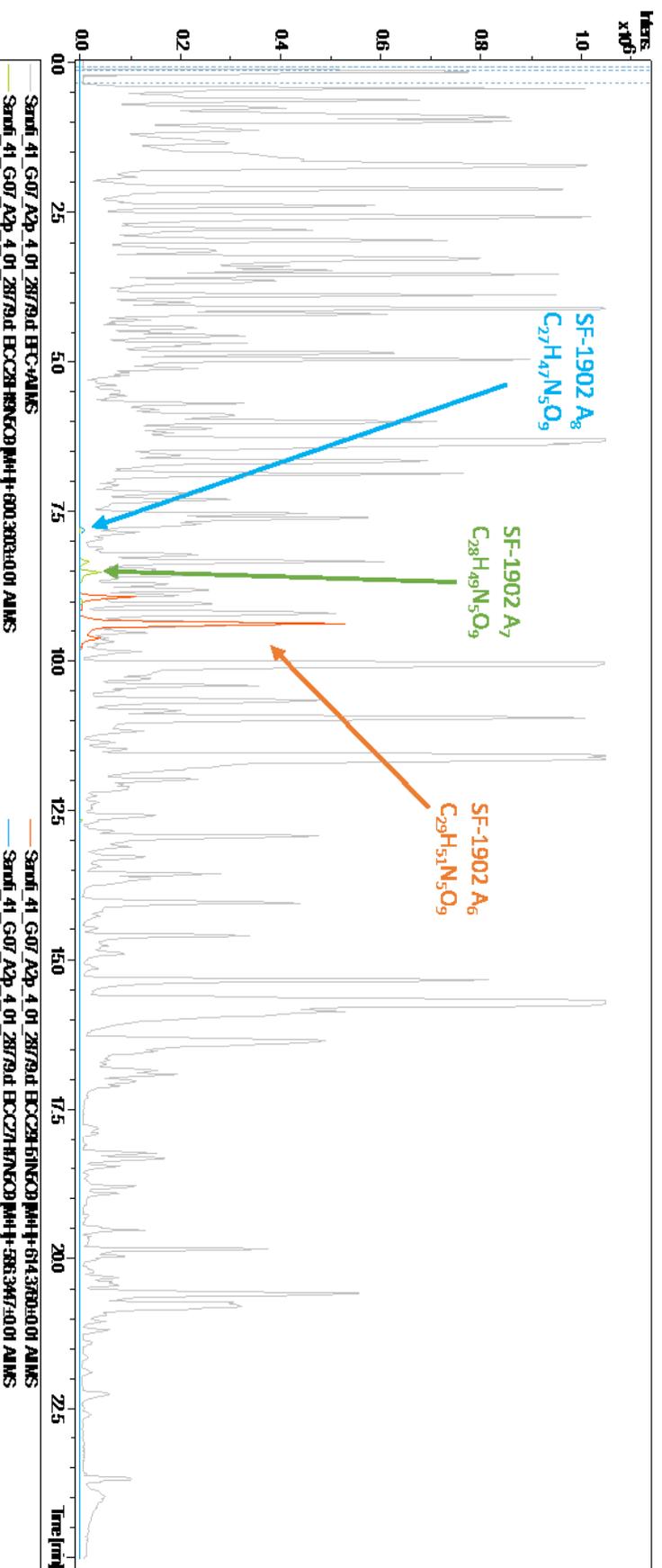


Figure S 4-5: Chromatogram of HAG010519 extract, with extracted chromatograms of novel congeners SF-1902 A₆-A₈ of Globomycin. Grey: BPC; light blue: EIC of SF-1902 A₈ (586.3447 ± 0.005), light green: EIC of SF-1902 A₇ (600.3603 ± 0.005), orange: EIC of SF-1902 A₆ (614.3760 ± 0.005). For SF-1902 A₈ just one isomer was found, for SF-1902 A₇ 3 isomers and for SF-1902 A₆ 6 isomers were found total.

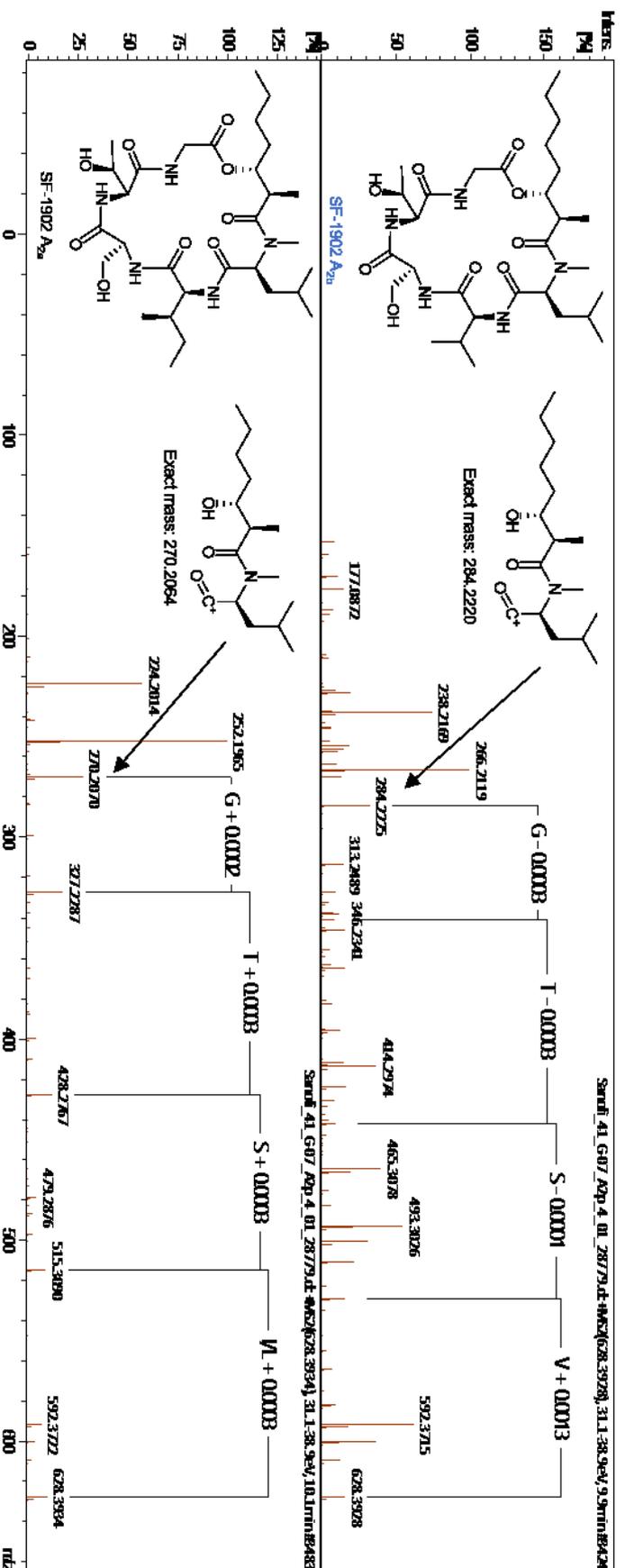
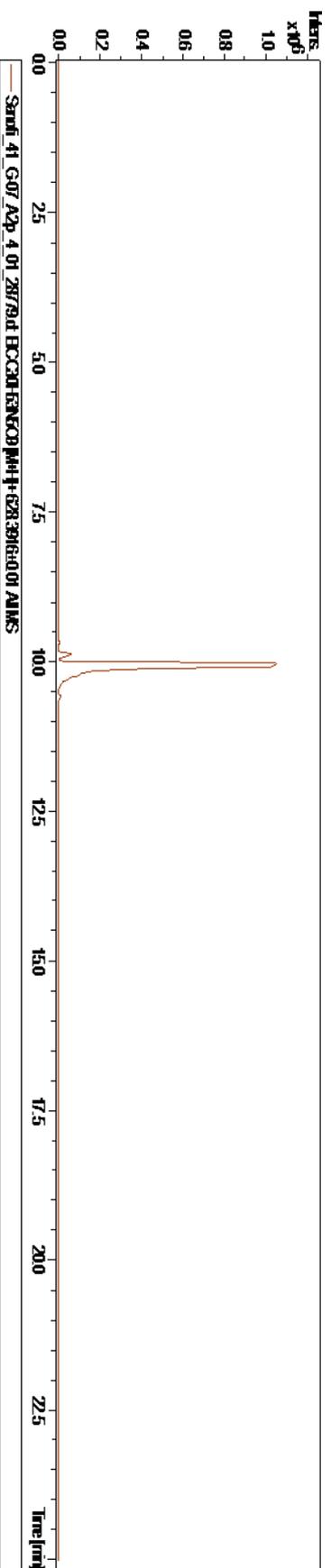


Figure S 4-6: EIC of SF-1902 A2 (628.3916 ± 0.005) and comparison of fragmentation pattern of known SF-1902 A2a and new SF-1902 A2b and its proposed structure.

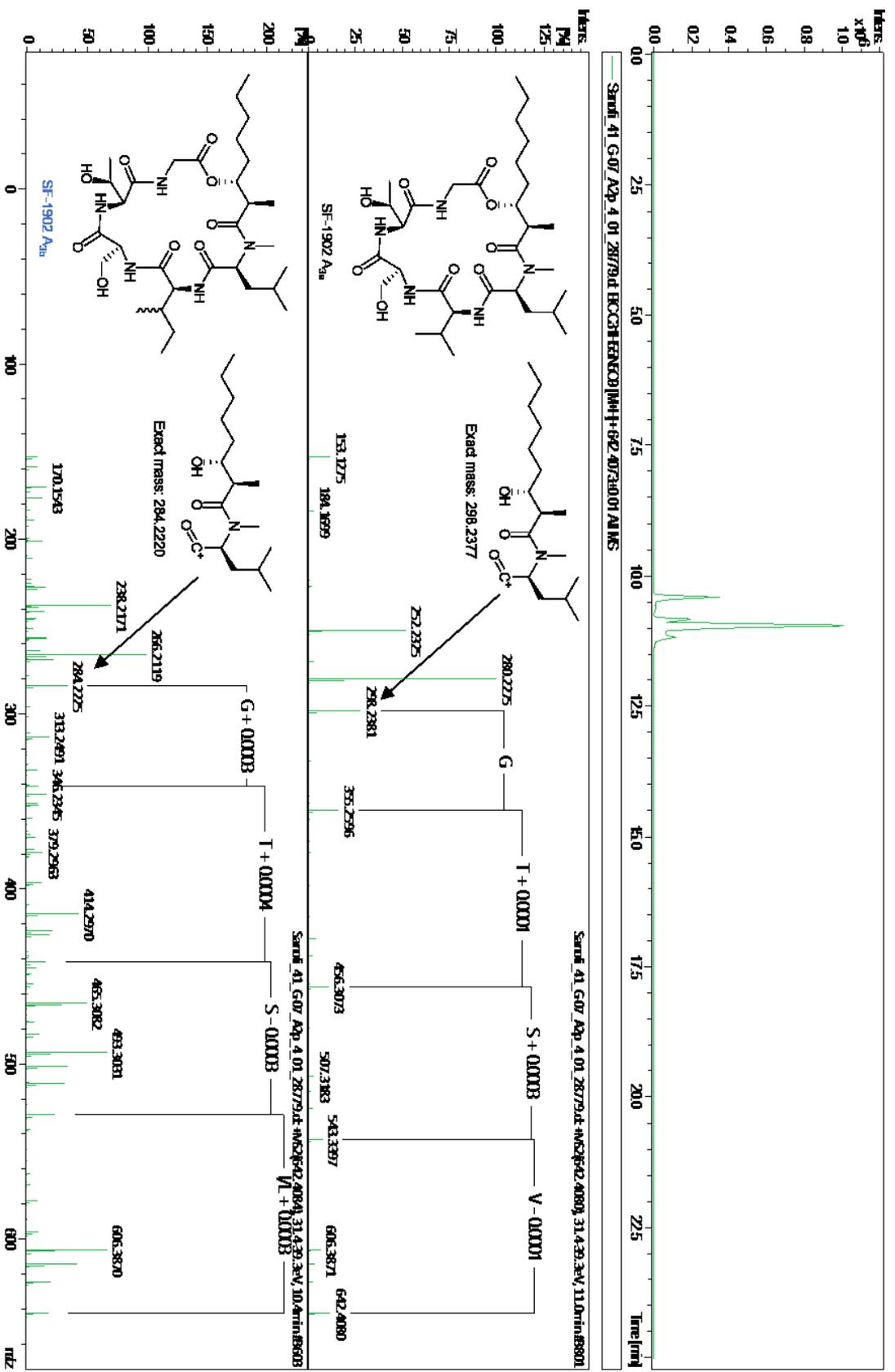


Figure S 4-7: EIC of SF-1902 A3 (642.4073 ± 0.005) and comparison of fragmentation pattern of known SF-1902 A3a and new SF-1902 A3b and its proposed structure.

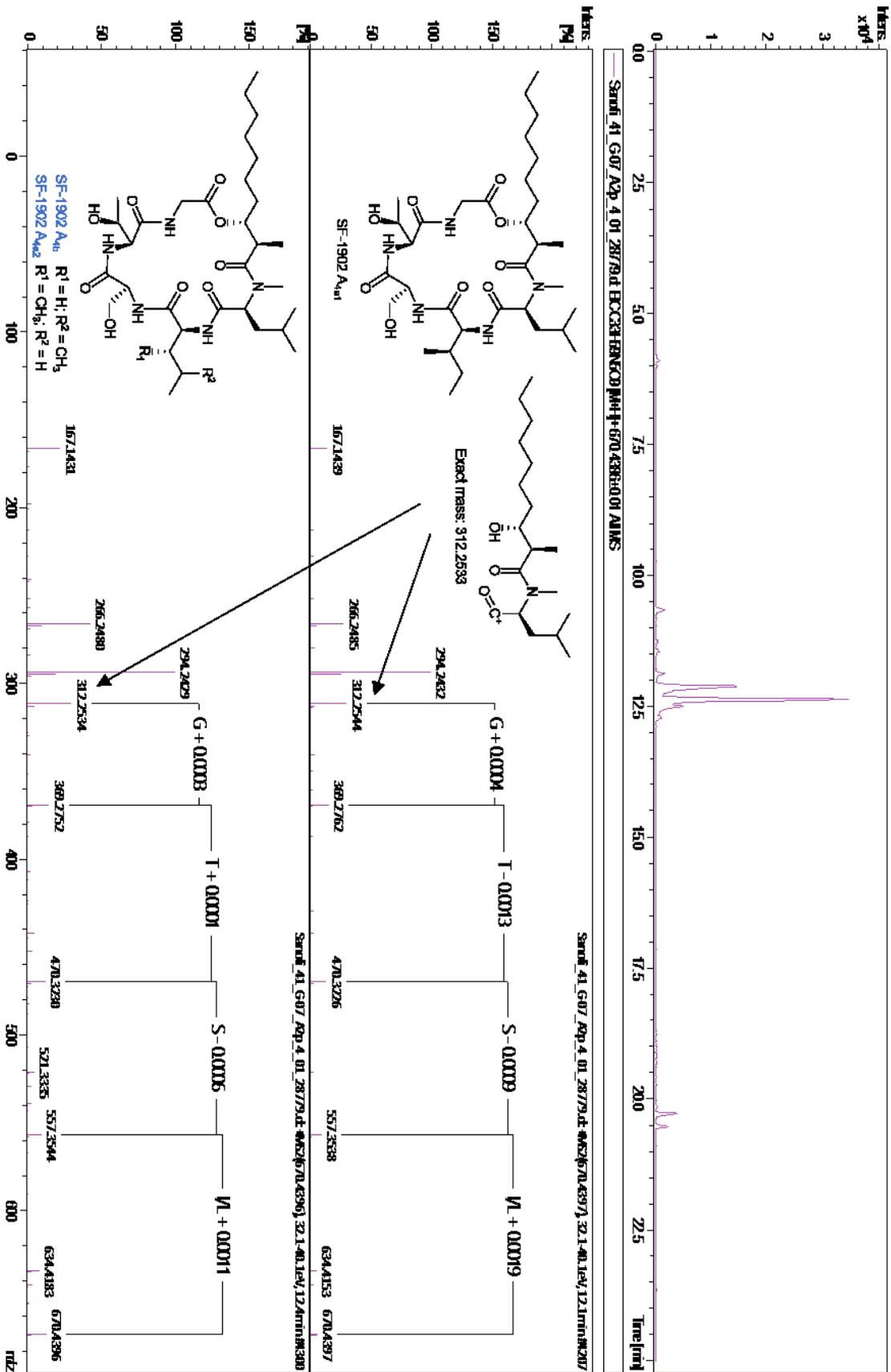


Figure S 4-8: EIC of SF-1902 A4 (670.4386 ± 0.005) and comparison of fragmentation pattern of known SF-1902 A4a1 and new SF-1902 A4b and its proposed structure.

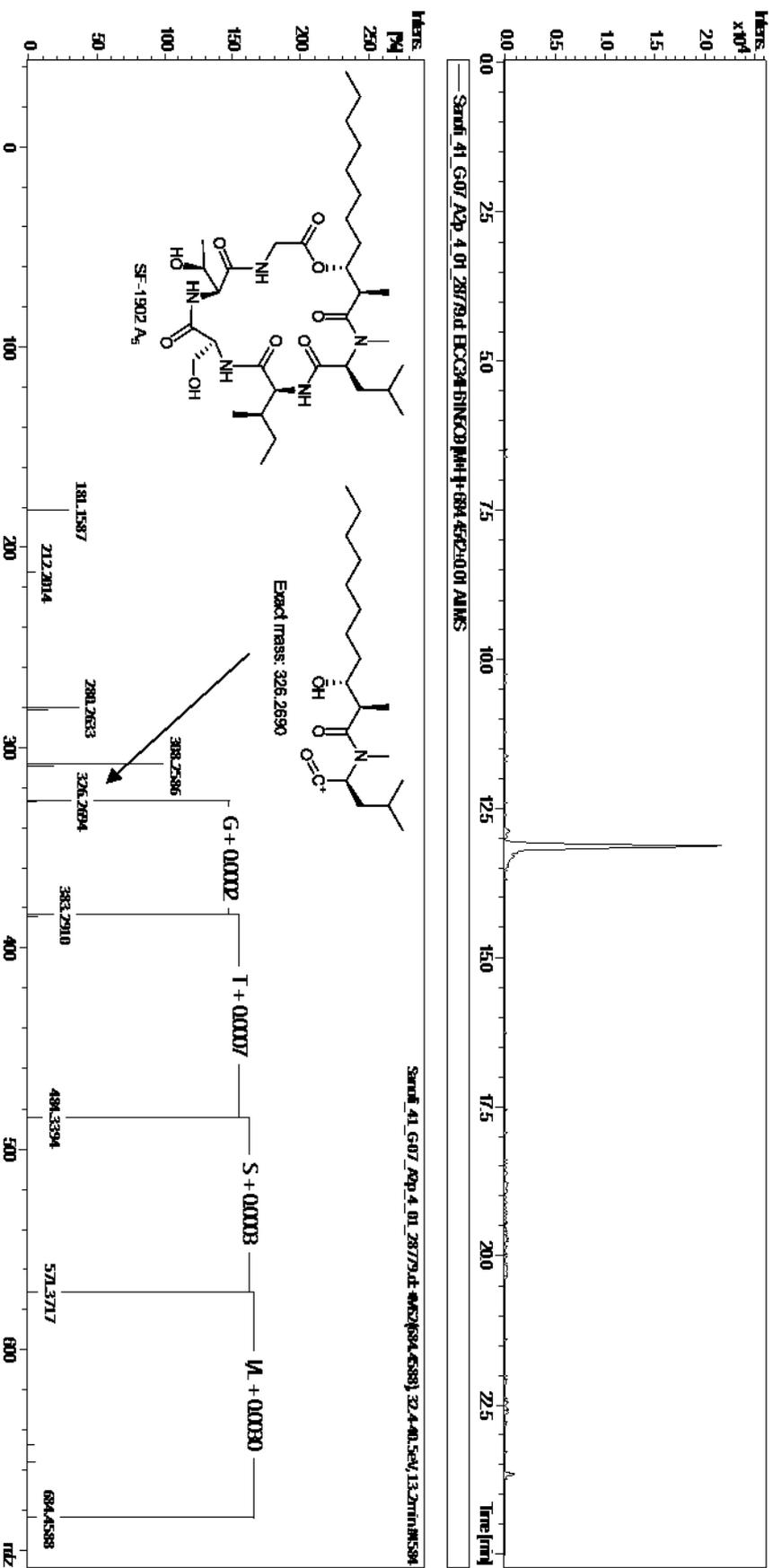


Figure S 4-9: EIC of SF-1902 A5 (684.4542 ± 0.005), the fragmentation pattern and proposed structure.

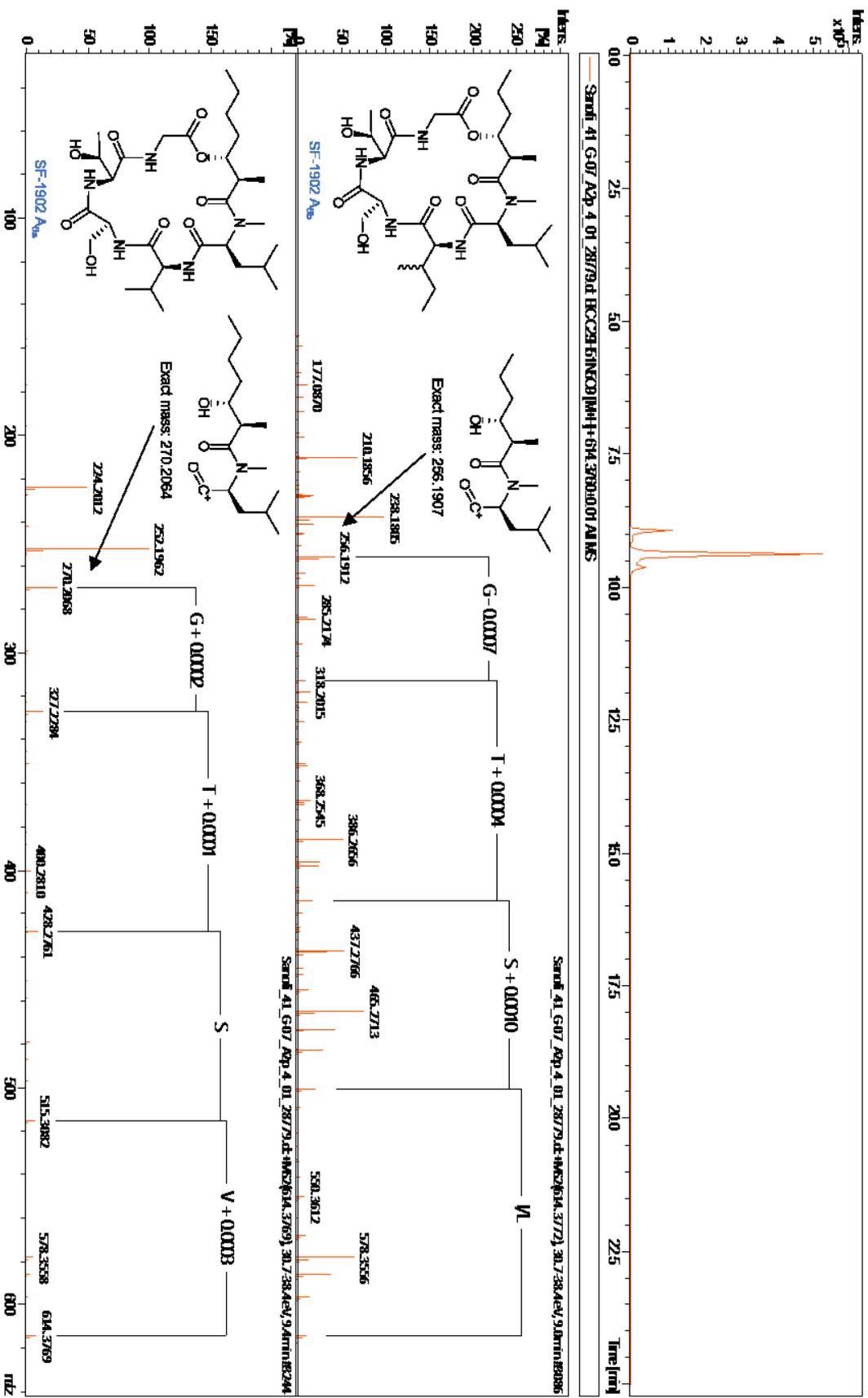


Figure S 4-10: EIC of SF-1902 A6 (614.3760 ± 0.005), the fragmentation pattern of new SF-1902 A6a and A6b and their proposed structure.

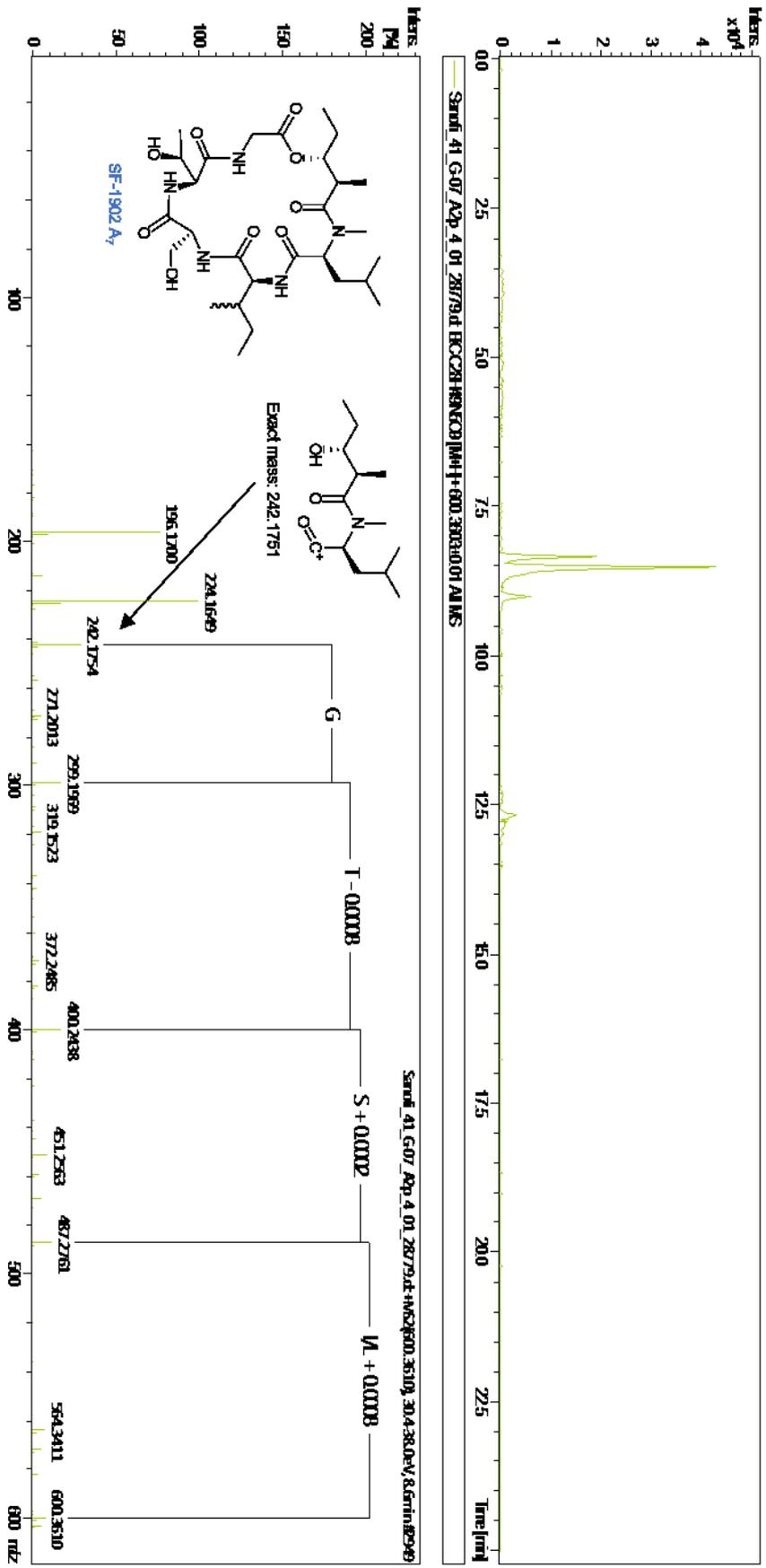


Figure S 4-11: EIC of SF-1902 A₇ (600.3603 ± 0.005), the fragmentation pattern and proposed structure.

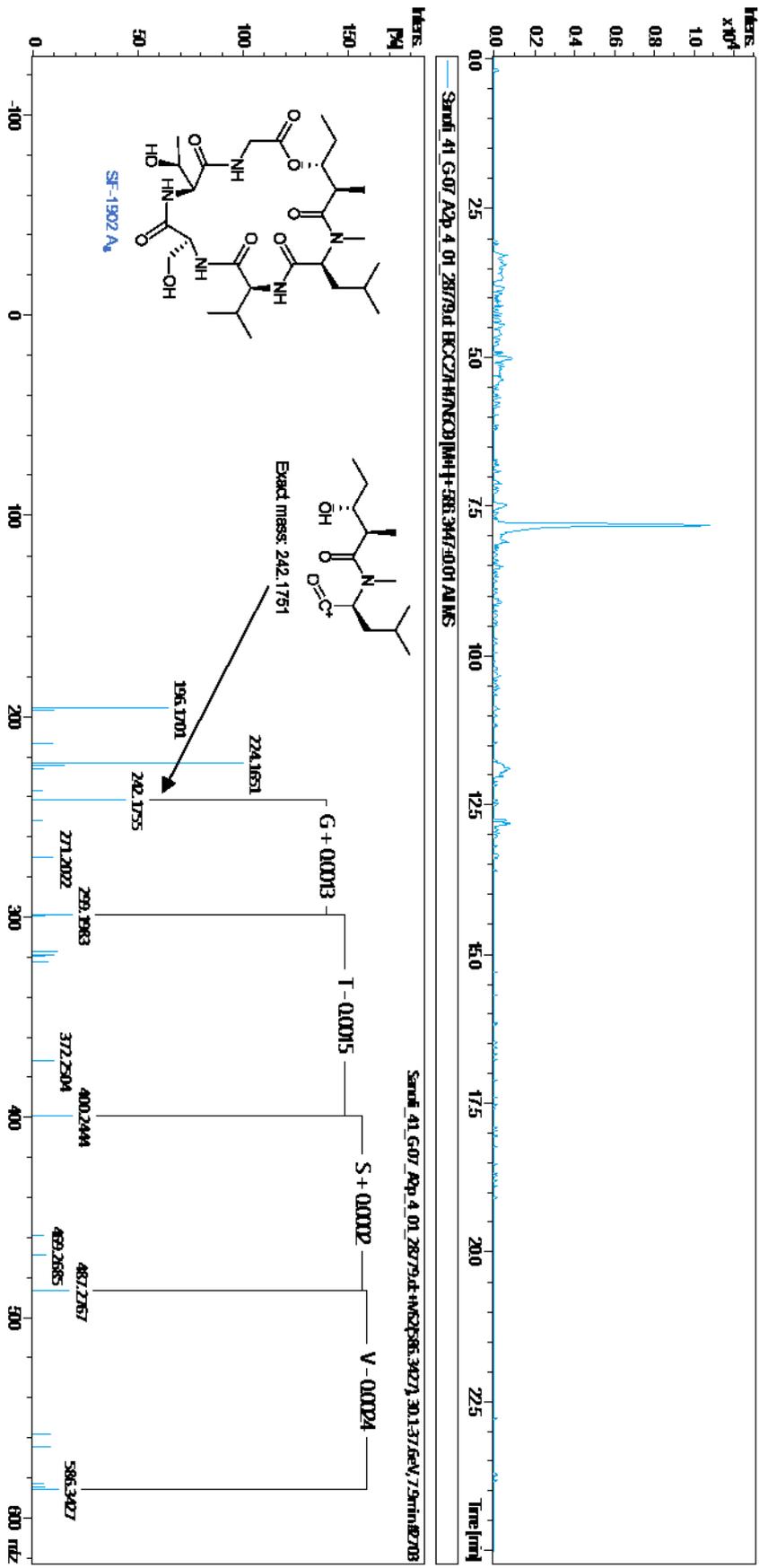


Figure S 4-12: EIC of SF-1902 A8 (586.3447 ± 0.005), the fragmentation pattern and proposed structure.

Table S 4-2: Overview of novel derivatives of GLM (1) based on the molecular network (Figure 4-4) and determined chemical formula.

[M+H] ⁺ <i>m/z</i>	chemical formula	fragmentation pattern
670.403	C ₃₂ H ₅₅ N ₅ O ₁₀	Figure S 4-13
654.408	C ₃₂ H ₅₅ N ₅ O ₉	
587.402	C ₂₉ H ₅₄ N ₄ O ₈	Figure S 4-15
670.388	C ₂₈ H ₅₅ N ₅ O ₁₃	Figure S 4-16
674.435	C ₃₂ H ₅₉ N ₅ O ₁₀	
688.385	C ₃₃ H ₆₁ N ₅ O ₁₀	Figure S 4-18
672.418	C ₃₂ H ₅₇ N ₅ O ₁₀	Figure S 4-19

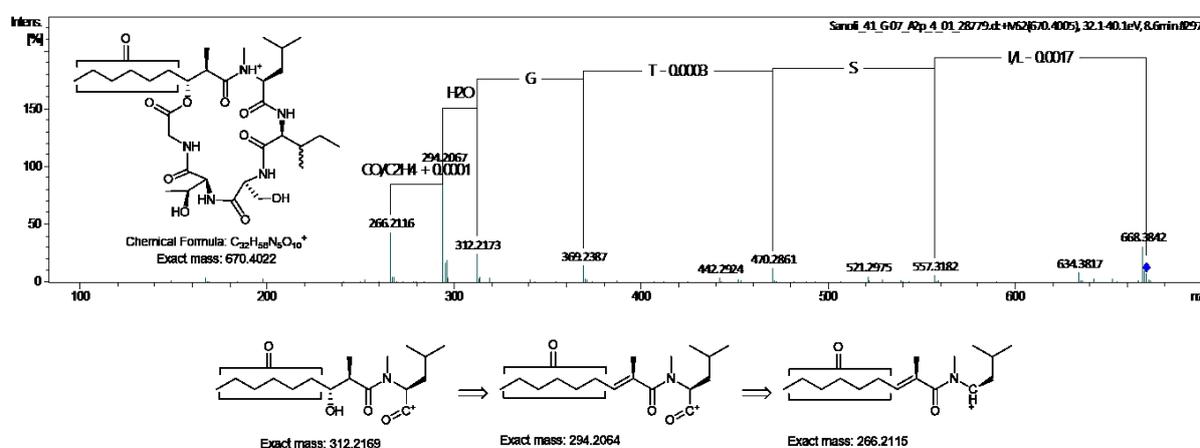


Figure S 4-13: Fragmentation pattern of *m/z* = 670.4005 and its proposed structure.

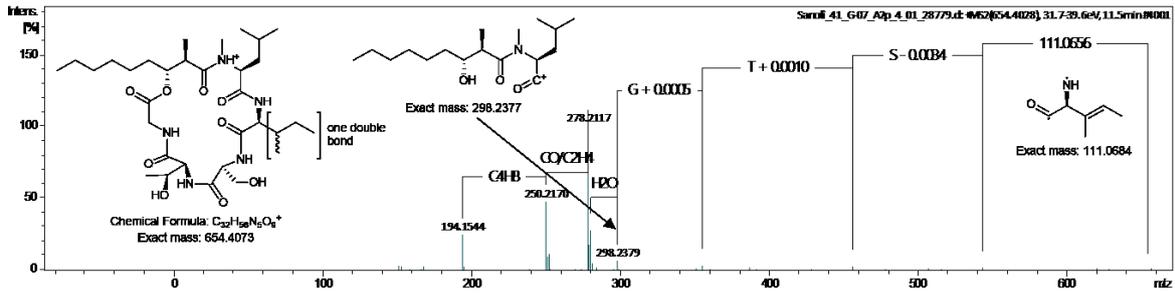
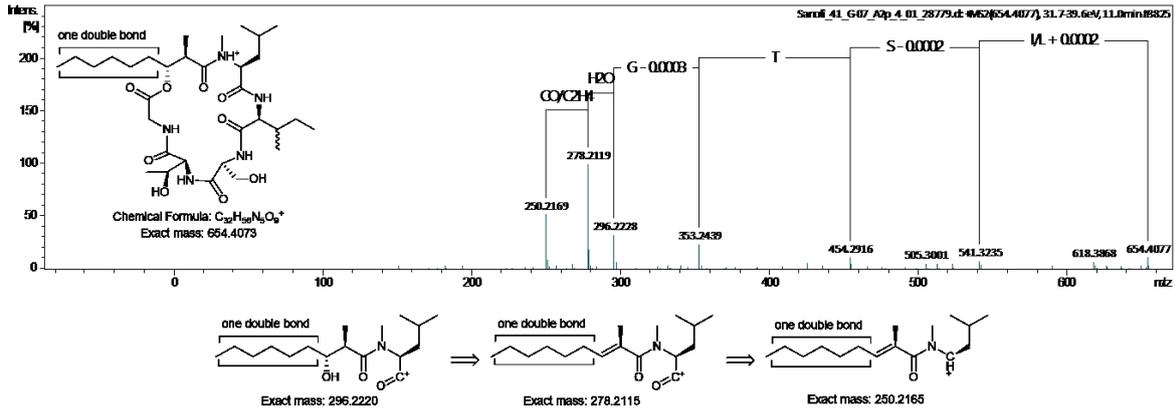


Figure S 4-14: Fragmentation pattern of $m/z = 654.4083$ and its proposed structures.

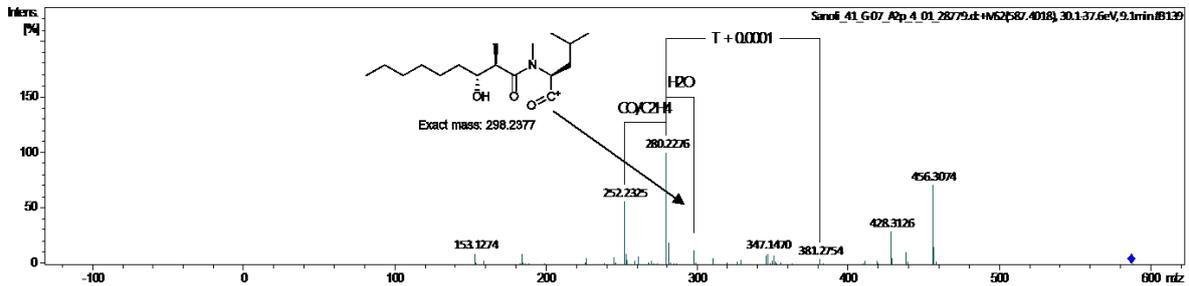


Figure S 4-15: Fragmentation pattern of $m/z = 587.4018$ and proposed structural elements.

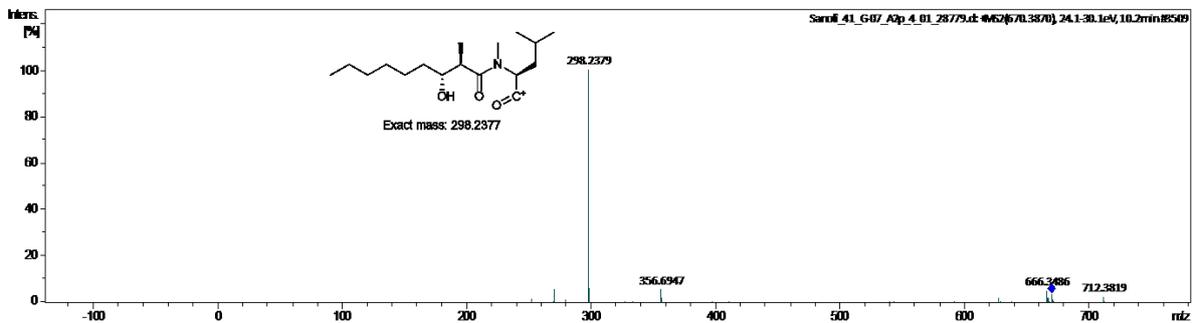


Figure S 4-16: Fragmentation pattern of $m/z = 670.3870$ and proposed structural elements.

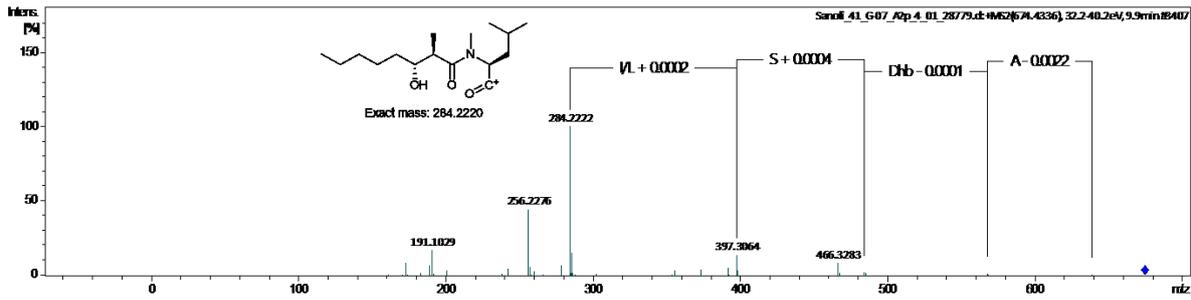


Figure S 4-17: Fragmentation pattern of $m/z = 674.4336$ and proposed structural elements.

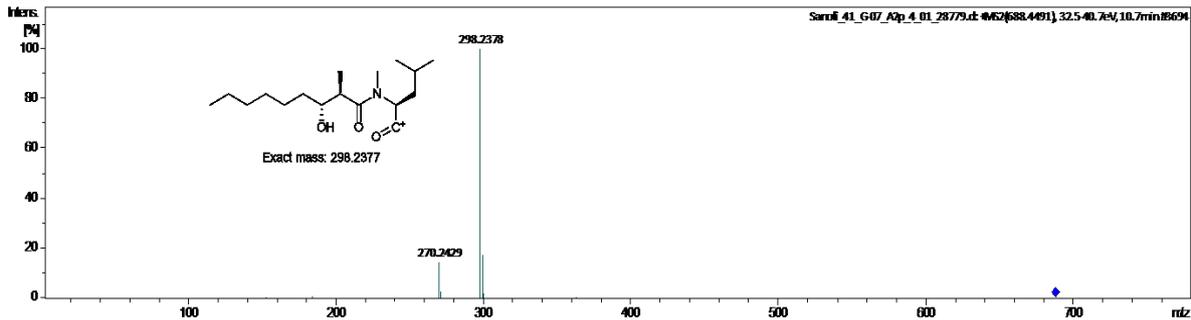


Figure S 4-18: Fragmentation pattern of $m/z = 688.4491$ and proposed structural elements.

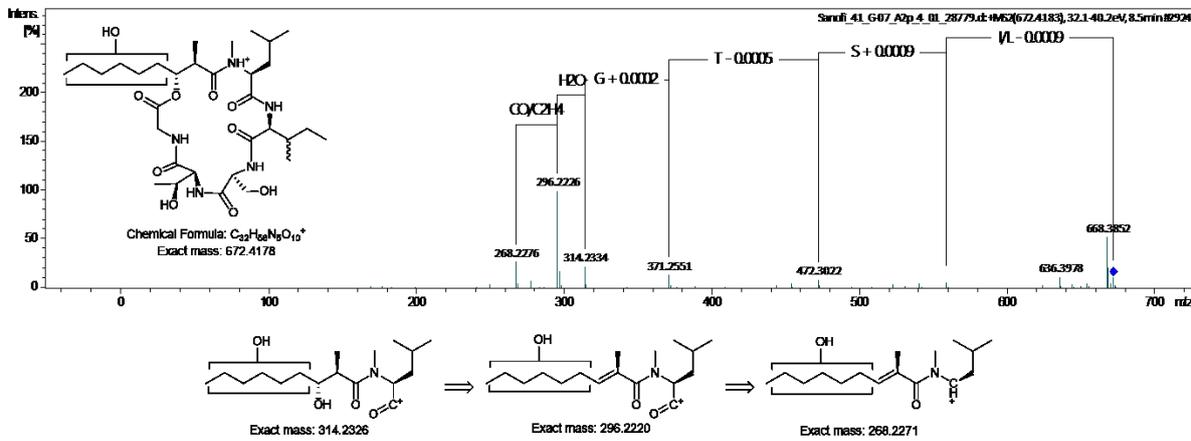


Figure S 4-19: Fragmentation pattern of $m/z = 672.4181$ and its proposed structure.

4.5.2 Structure Based Drug Design

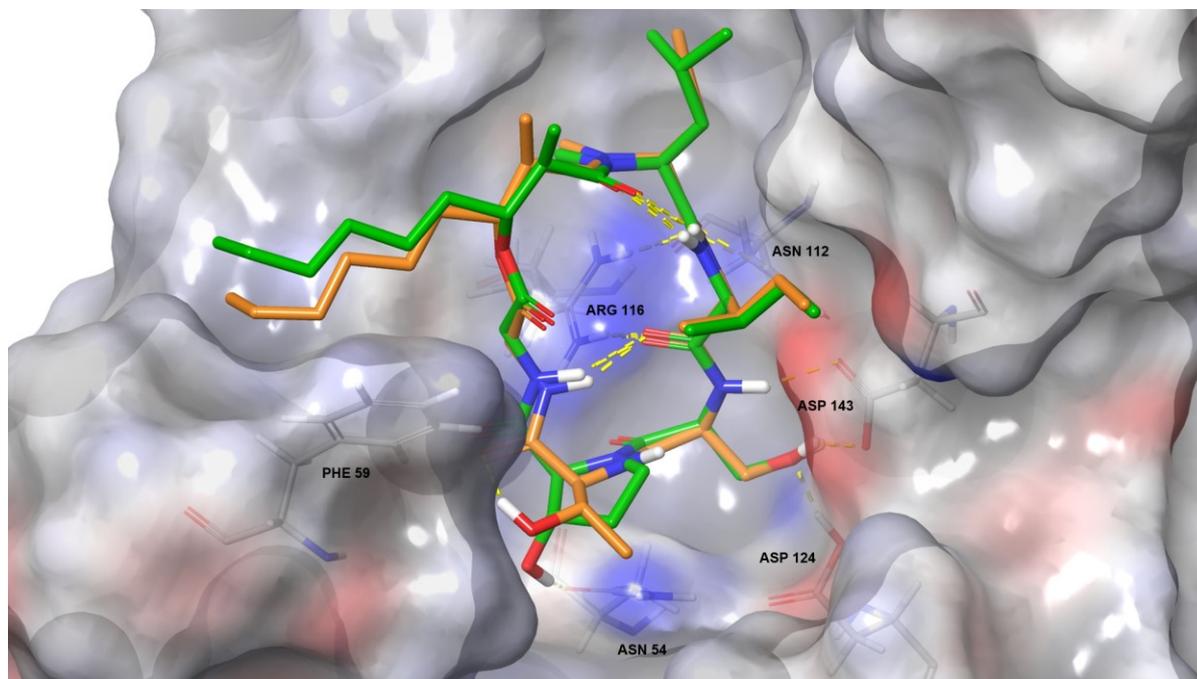


Figure S 4-20: Globomycin (orange) and derivative 1.3 (green) in LspA (PDB 5DIR).

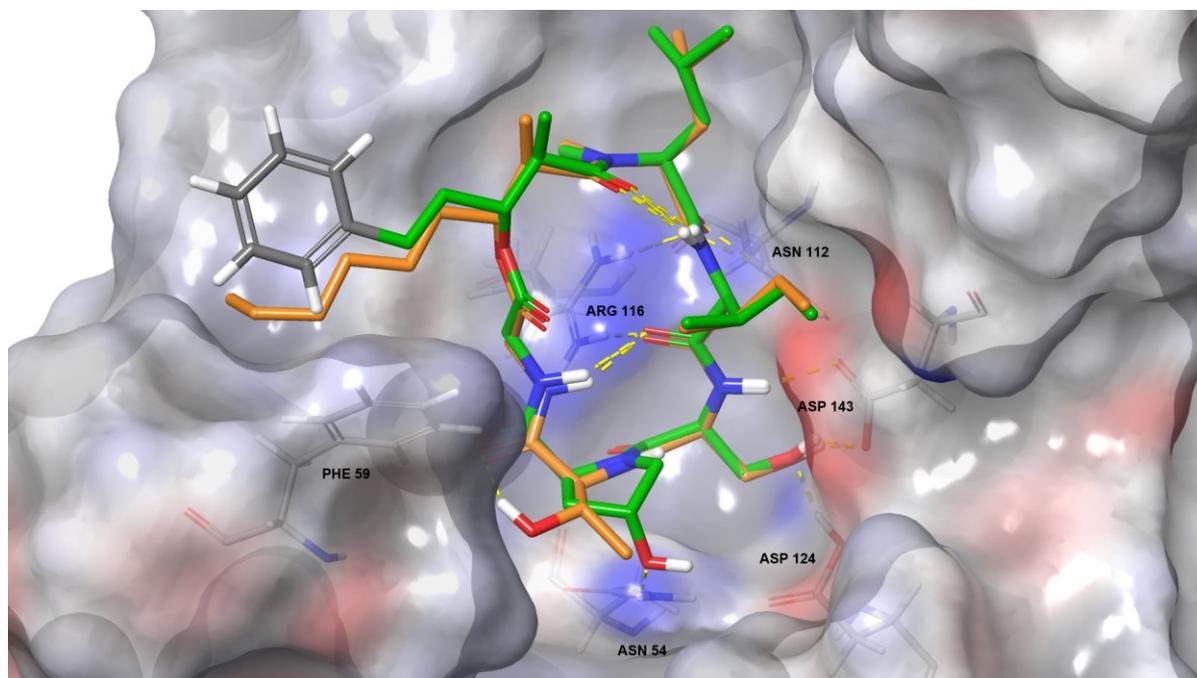


Figure S 4-21: Globomycin (orange) and derivative 1.5 (green) in LspA (PDB 5DIR).

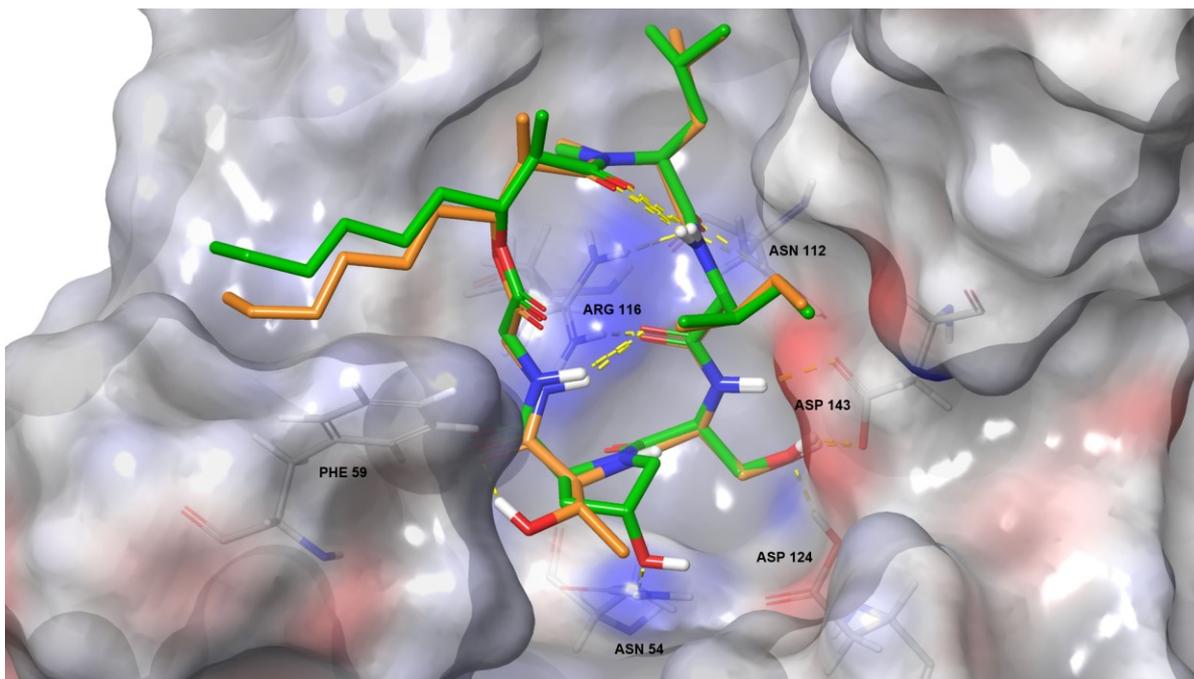


Figure S 4-22: Globomycin (orange) and derivative 1.2 (green) in LspA (PDB 5DIR).

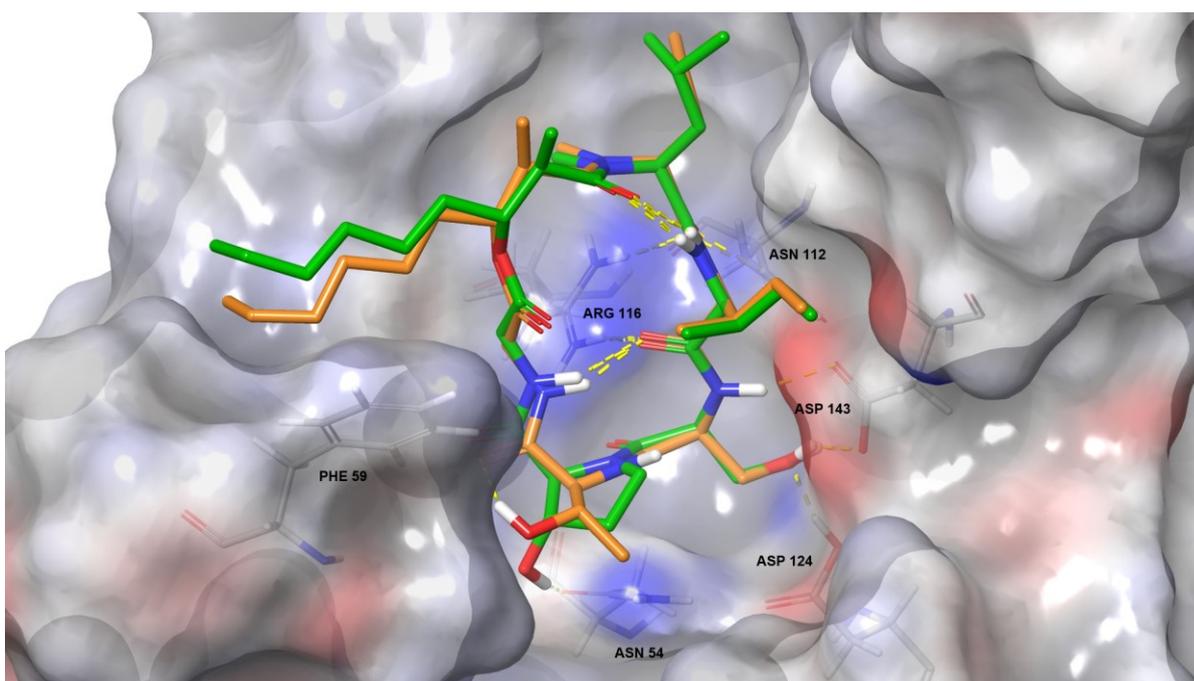


Figure S 4-23: Globomycin (orange) and derivative 1.3 (green) in LspA (PDB 5DIR).

4.5.3 NMR and MS/MS² data for synthetic compounds

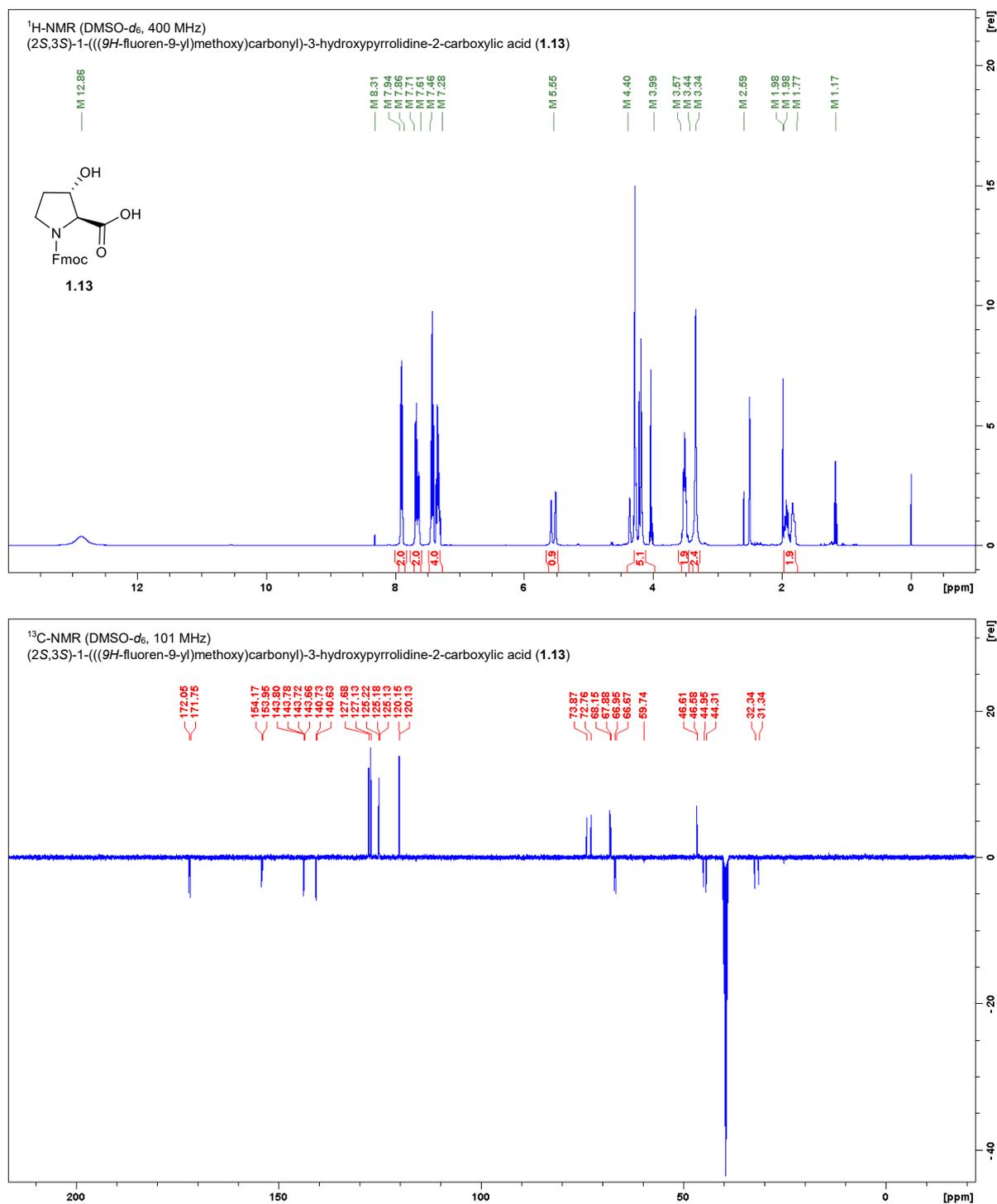


Figure S 4-24: NMR spectra of **1.13** in DMSO-*d*₆.

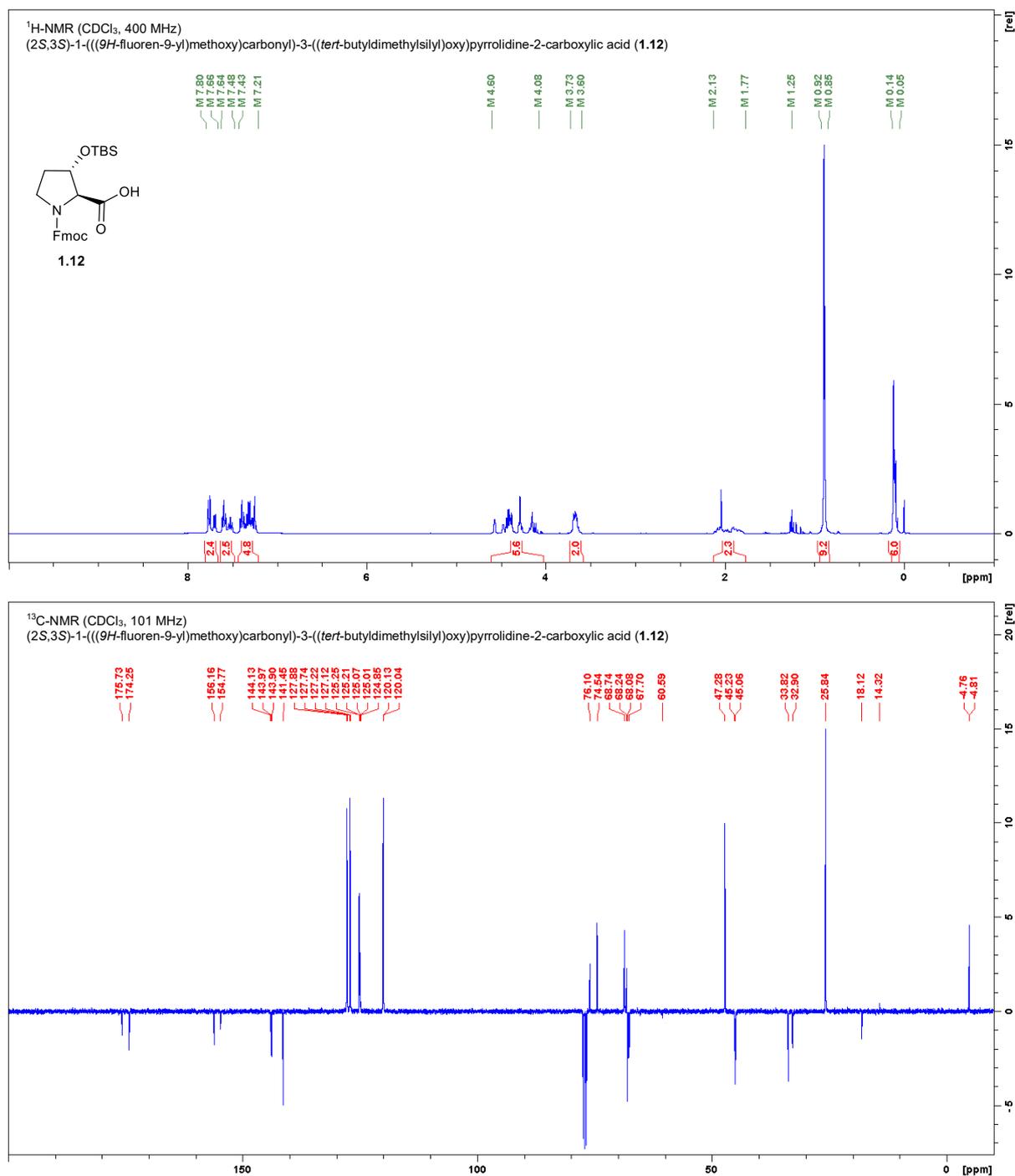


Figure S 4-25: NMR spectra of **1.12** in CDCl₃.

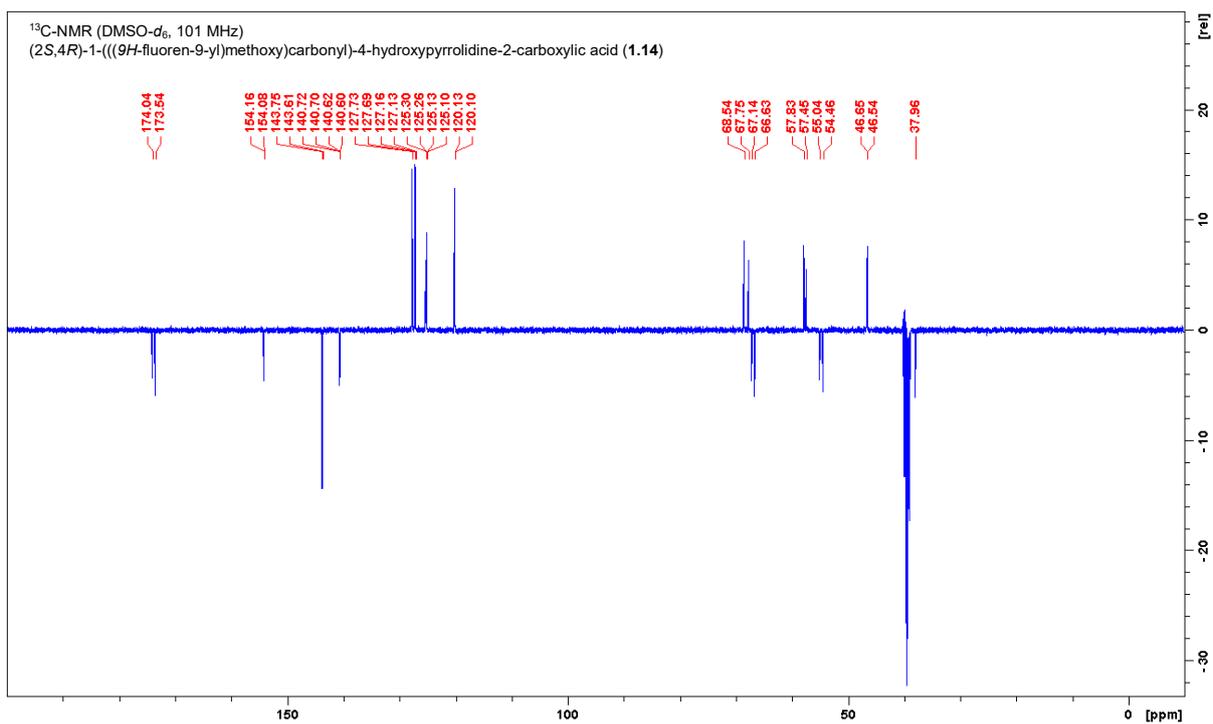
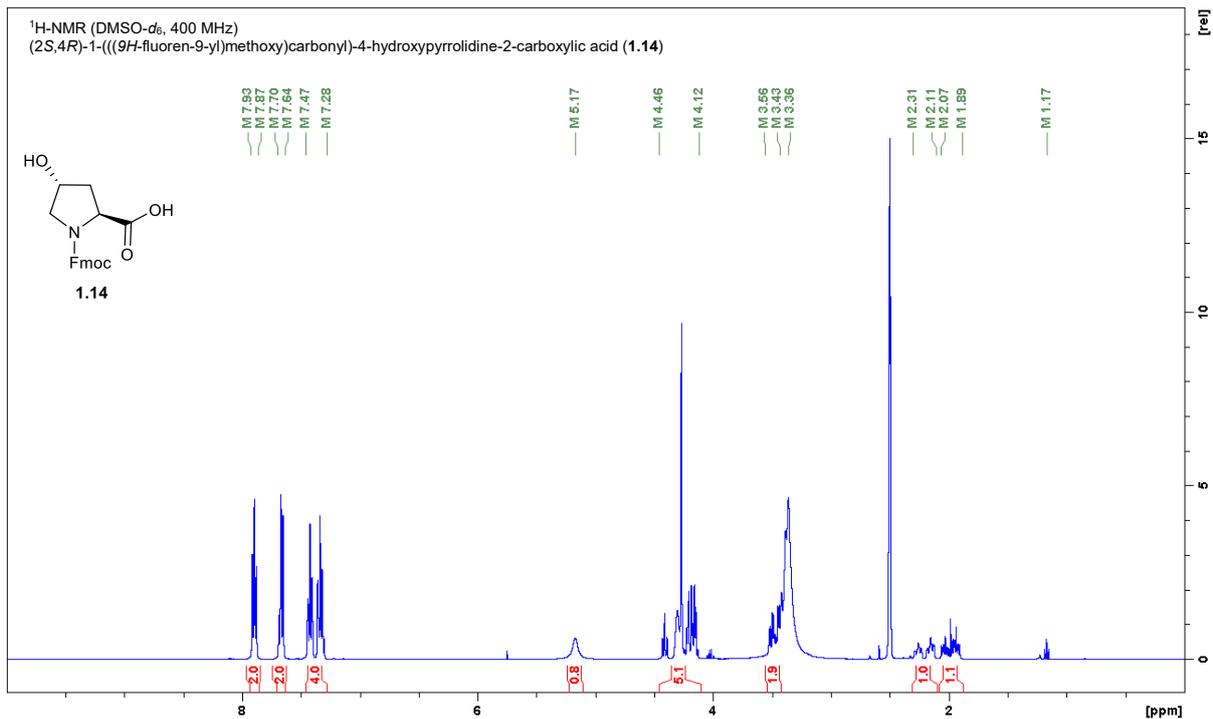


Figure S 4-26: NMR spectra of **1.14** in DMSO-*d*₆.

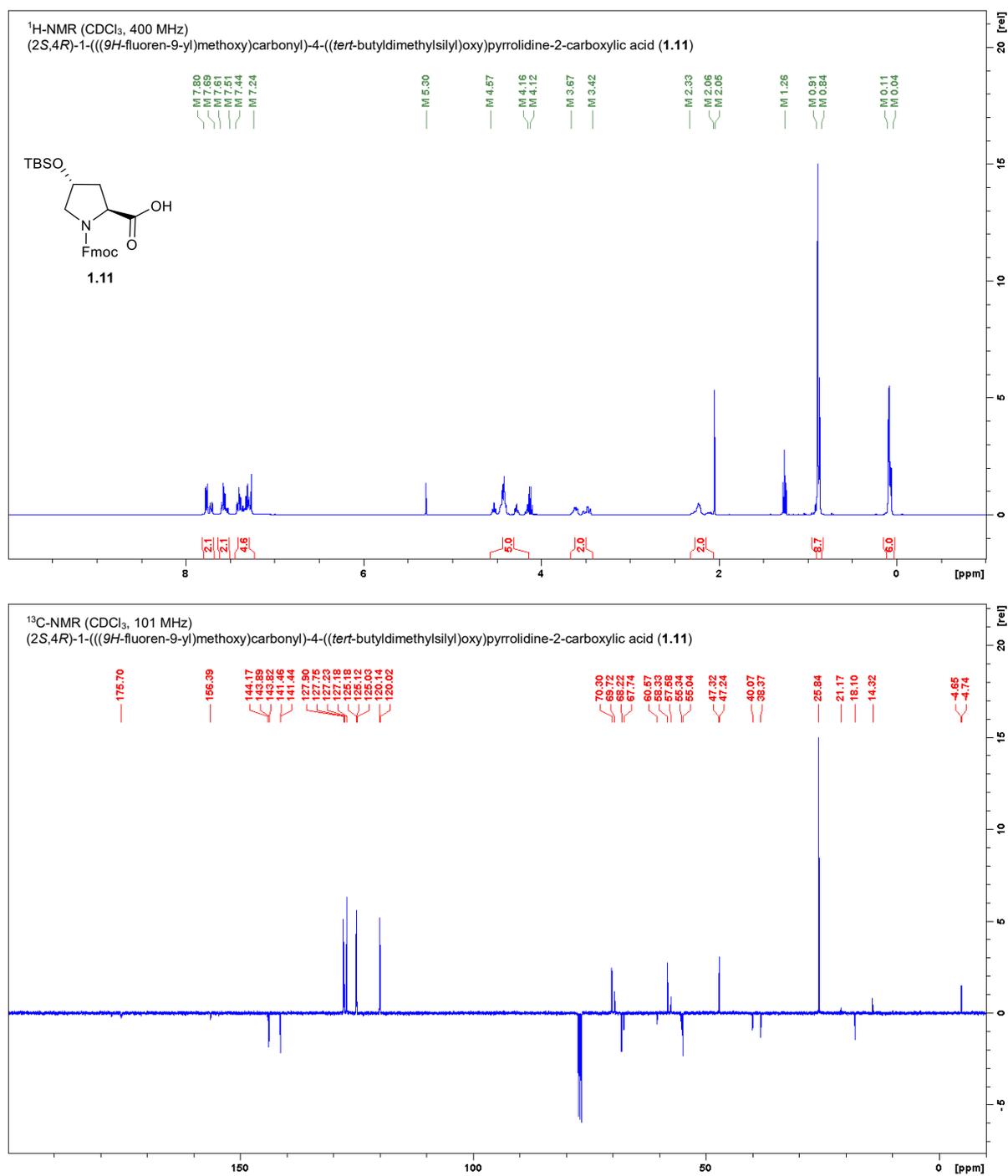


Figure S 4-27: NMR spectra of **1.11** in CDCl₃.

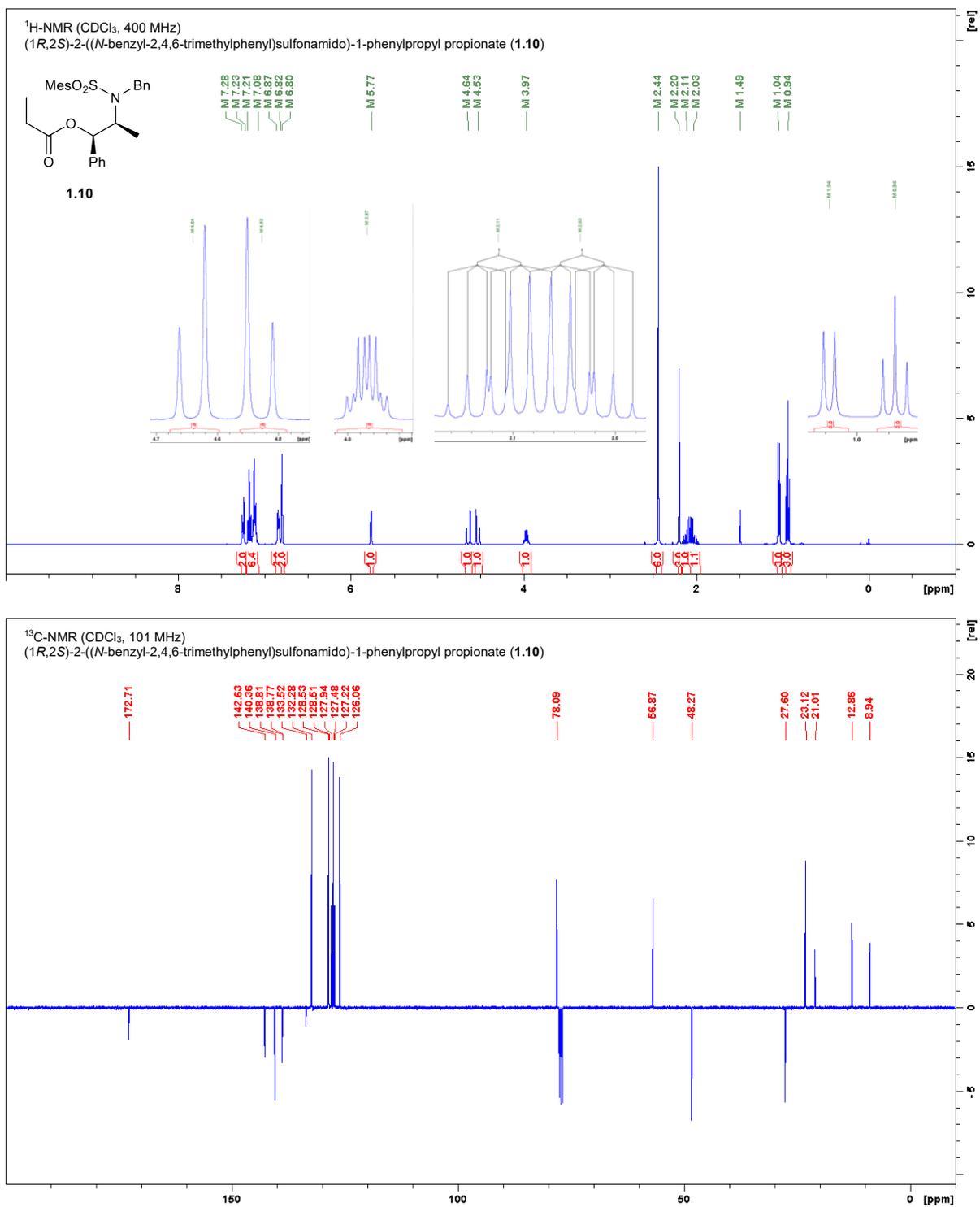


Figure S 4-28: NMR spectra of **1.10** in CDCl₃.

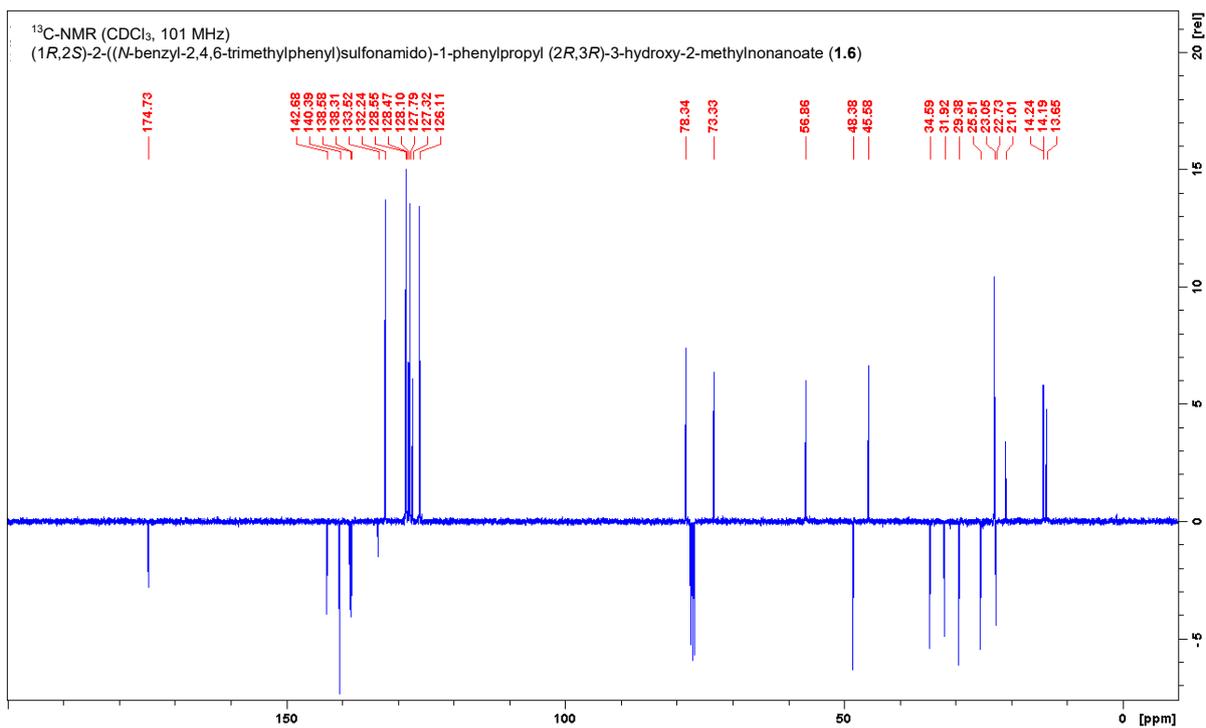
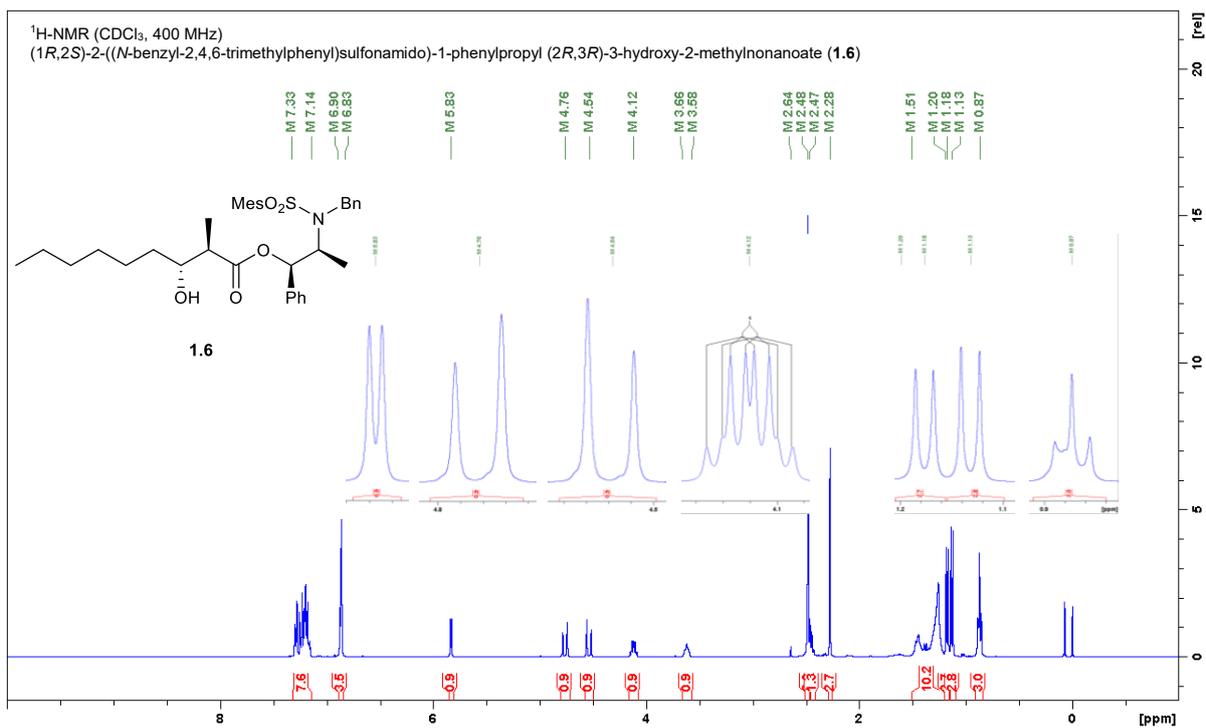


Figure S 4-29: NMR spectra of **1.6** in CDCl₃.

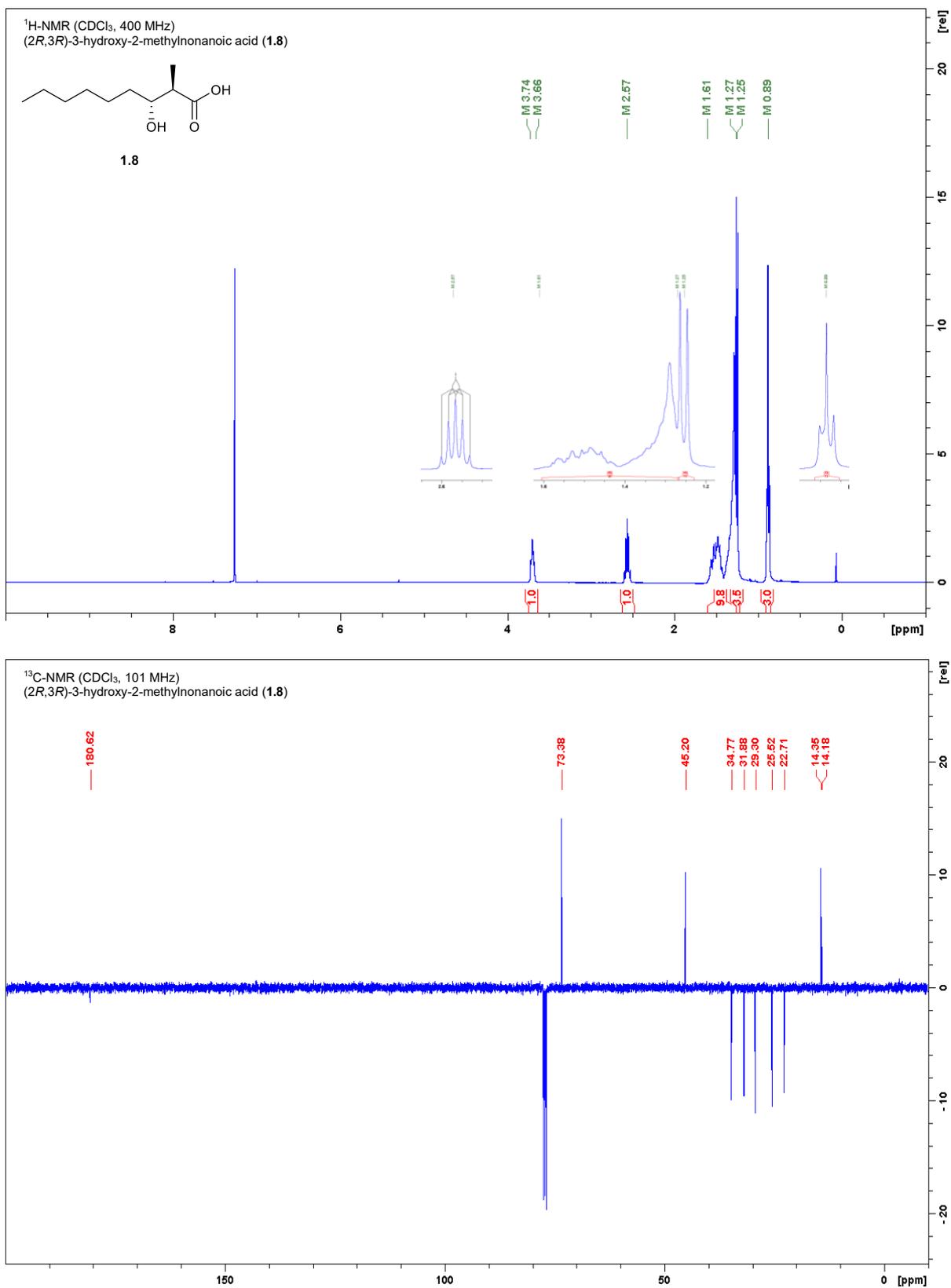


Figure S 4-30: NMR spectra of **1.8** in CDCl₃.

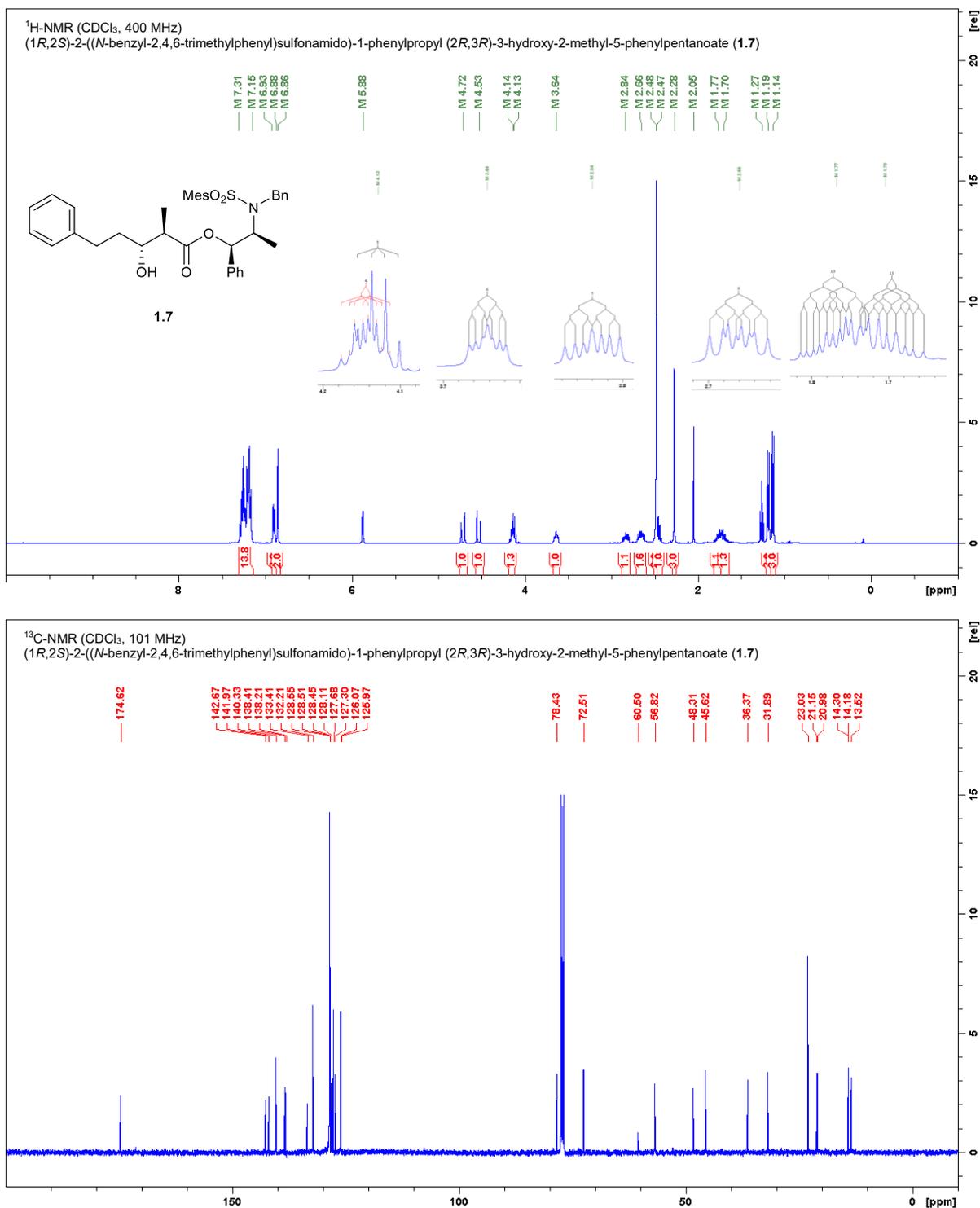


Figure S 4-31: NMR spectra of **1.7** in CDCl₃.

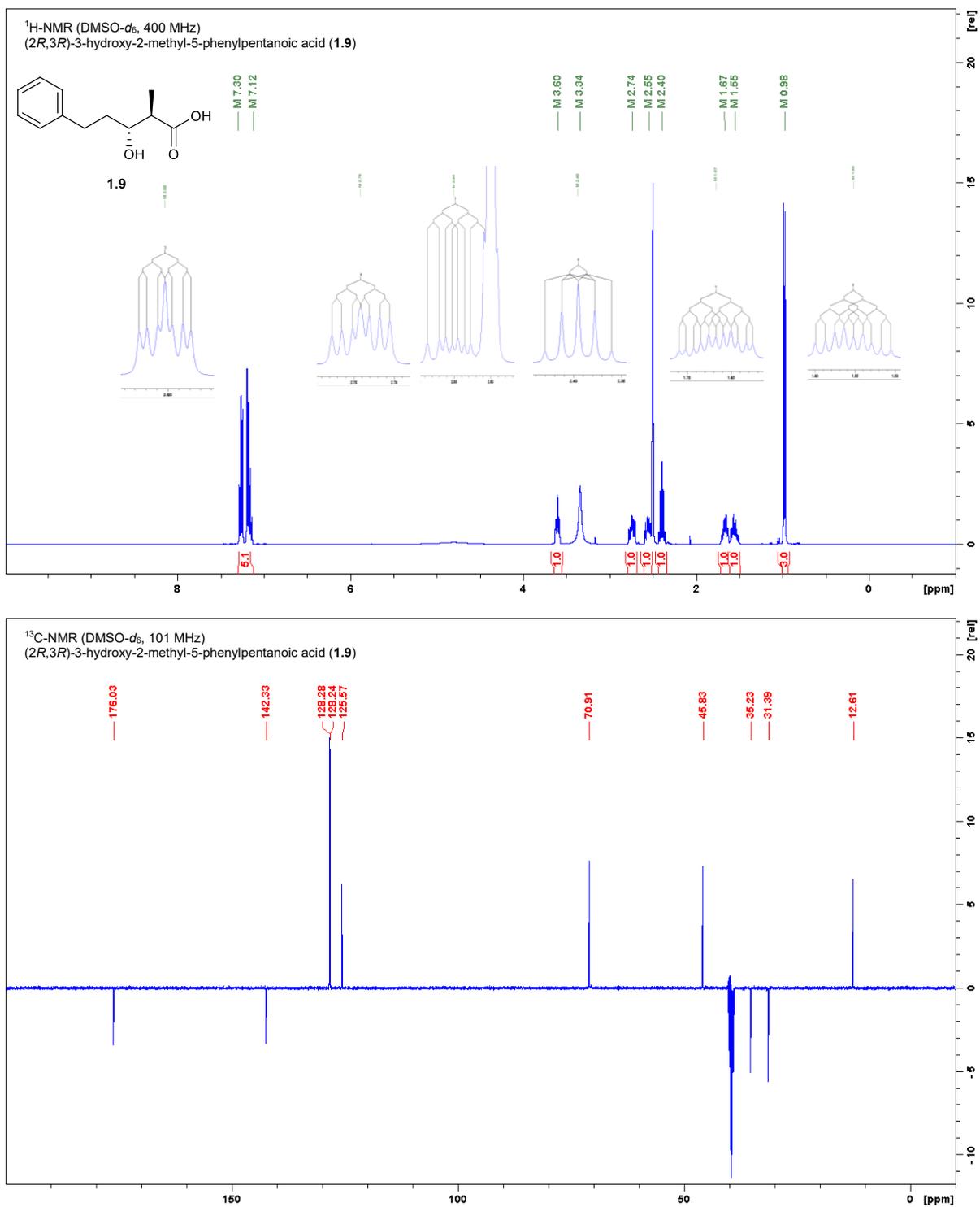


Figure S 4-32: NMR spectra of **1.9** in DMSO-*d*₆.

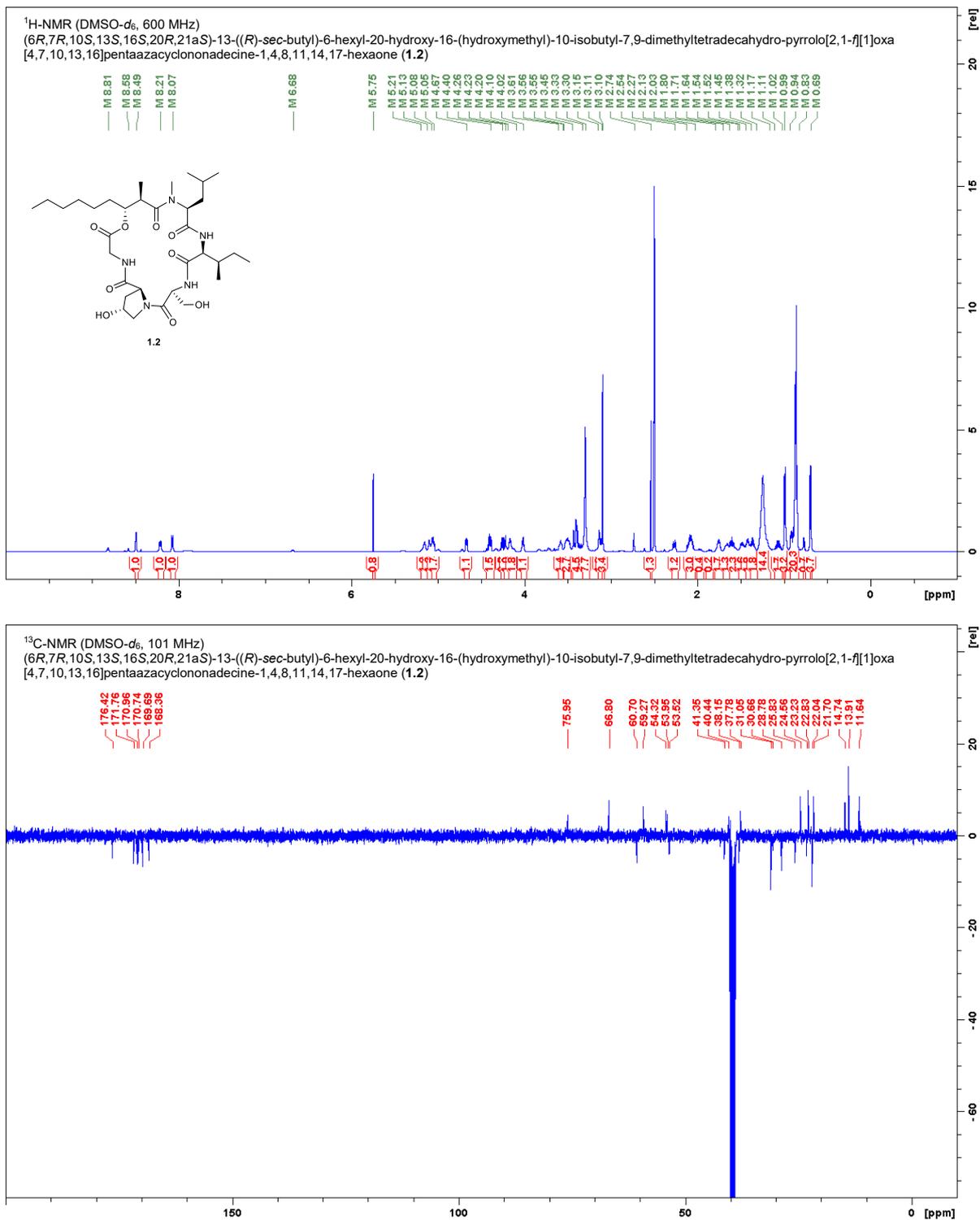


Figure S 4-33: NMR spectra of **1.2** in DMSO-*d*₆.

Table S 4-3: NMR data for 1.2 in DMSO-*d*₆ (600 MHz, 100 MHz).

Position	δ ¹³ C [ppm]	δ ¹ H [ppm]; (m, f, J)
L- <i>N</i> -Me-Leucine		
CO	171.0	-
N-CH ₃	40.1*	3.10 (s, 3H)
α -CH	68.5*	3.55-3.45 (m, 2H, overlay)
β -CH ₂	38.2	2.13-2.03 (m, 3H, overlay) 1.64-1.54 (m, 2H)
γ -CH	24.6	1.45-1.38 (m, 2H)
δ -CH ₃ a	22.8	0.94-0.83 (m, 20H, overlay)
δ -CH ₃ b	21.7	0.94-0.83 (m, 20H, overlay)
L- <i>allo</i> -Isoleucine		
CO	171.8	-
NH	-	8.21 (d, 1H, J = 9.3 Hz) 6.68 (minor)
α -CH	53.9	4.40 (dd, 1H, J = 9.5, 4.6 Hz)
β -CH	37.8	1.80-1.71 (m, 2H)
γ -CH ₂	25.8	1.38-1.32 (m, 2H) 1.11-1.02 (m, 2H)
γ -CH ₃	13.9	0.69 (d, 3H, J = 6.9 Hz)
δ -CH ₃	11.6	0.94-0.83 (m, 20H, overlay)
L-Serine		
CO	169.7	-
NH	-	8.49 (d, 1H, J = 4.6 Hz) 8.58 (minor)
α -CH	54.3	4.02 (ddd, 1H, J = 6.7, 6.7, 4.6 Hz)
β -CH ₂ -OH	60.7	3.61-3.56 (m, 1H) 3.55-3.45 (m, 2H, overlay)
β -OH	-	5.21-5.13 (m, 1H)
L-4-Hydroxyproline		
CO	170.7	-
N	-	-
α -CH	59.3	4.67 (dd, 1H, J = 8.5, 4.3 Hz)
β -CH ₂	38.8*	2.27 (ddd, 1H, J = 12.5, 4.9, 4.9 Hz) 2.13-2.03 (m, 3H, overlay)
γ -CH-OH	66.8	4.20-4.10 (m, 2H)
γ -OH	-	5.13-5.08 (m, 1H)
δ -CH ₂	53.5	3.45-3.33 (m, 4H, overlay)
Glycine		
CO	168.4	-
NH	-	8.07 (d, 1H, J = 8.3 Hz) 8.81 (minor)
CH ₂	41.4	4.26 (d, 1H, J = 8.1 Hz) 4.23 (d, 1H, J = 8.3 Hz) 3.45-3.37 (m, 4H, overlay) minor
BHA		

1-CO	176.4	-
2-CH	40.1*	3.15-3.11 (m, 2H)
2'-CH₃	14.7	0.99 (d, 3H, J = 7.0 Hz)
3-CH-O	76.0	5.05 (ddd, 2H, J = 10.2, 6.4, 3.6 Hz)
4-CH₂	30.7	1.71-1.64 (m, 1H) 1.52-1.45 (m, 1H)
5-8-CH₂	31.1, 28.8, 23.2, 22.0	1.32-1.17 (m, 14H, overlay)
9-CH₃	13.9	0.94-0.83 (m, 20H, overlay)

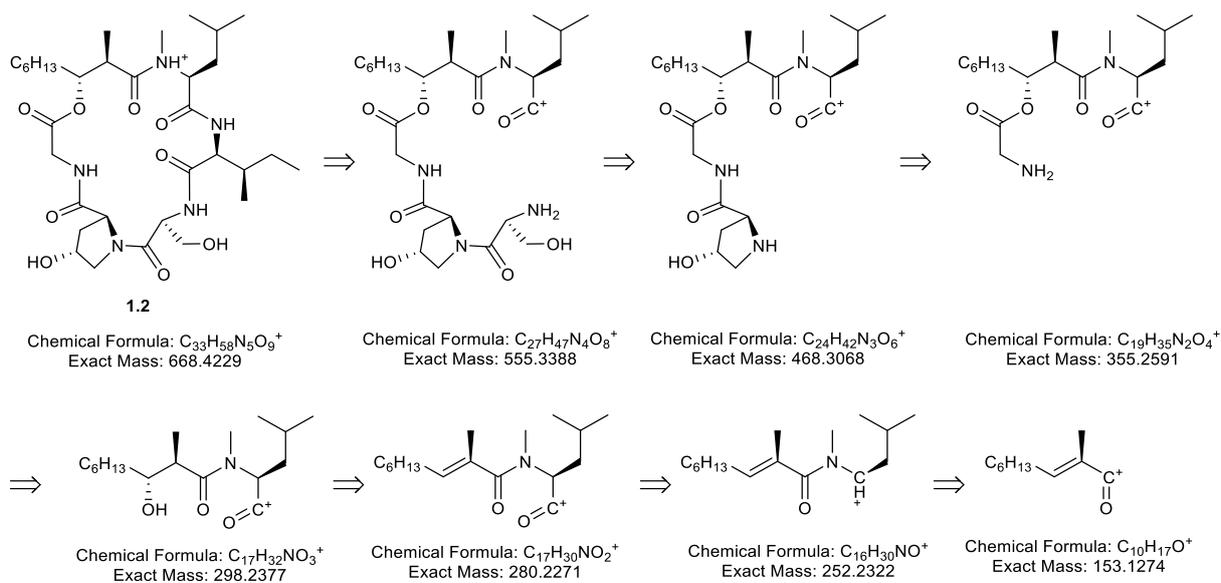
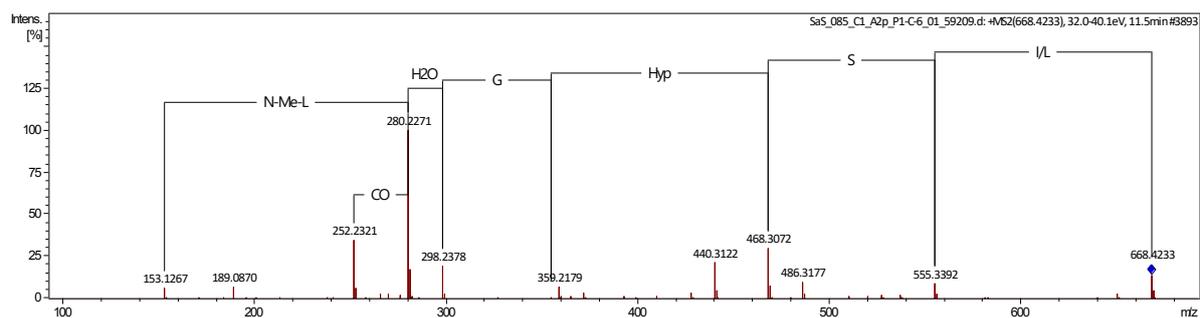


Figure S 4-34: MS² fragmentation of 1.2 and its fragmentation pattern.

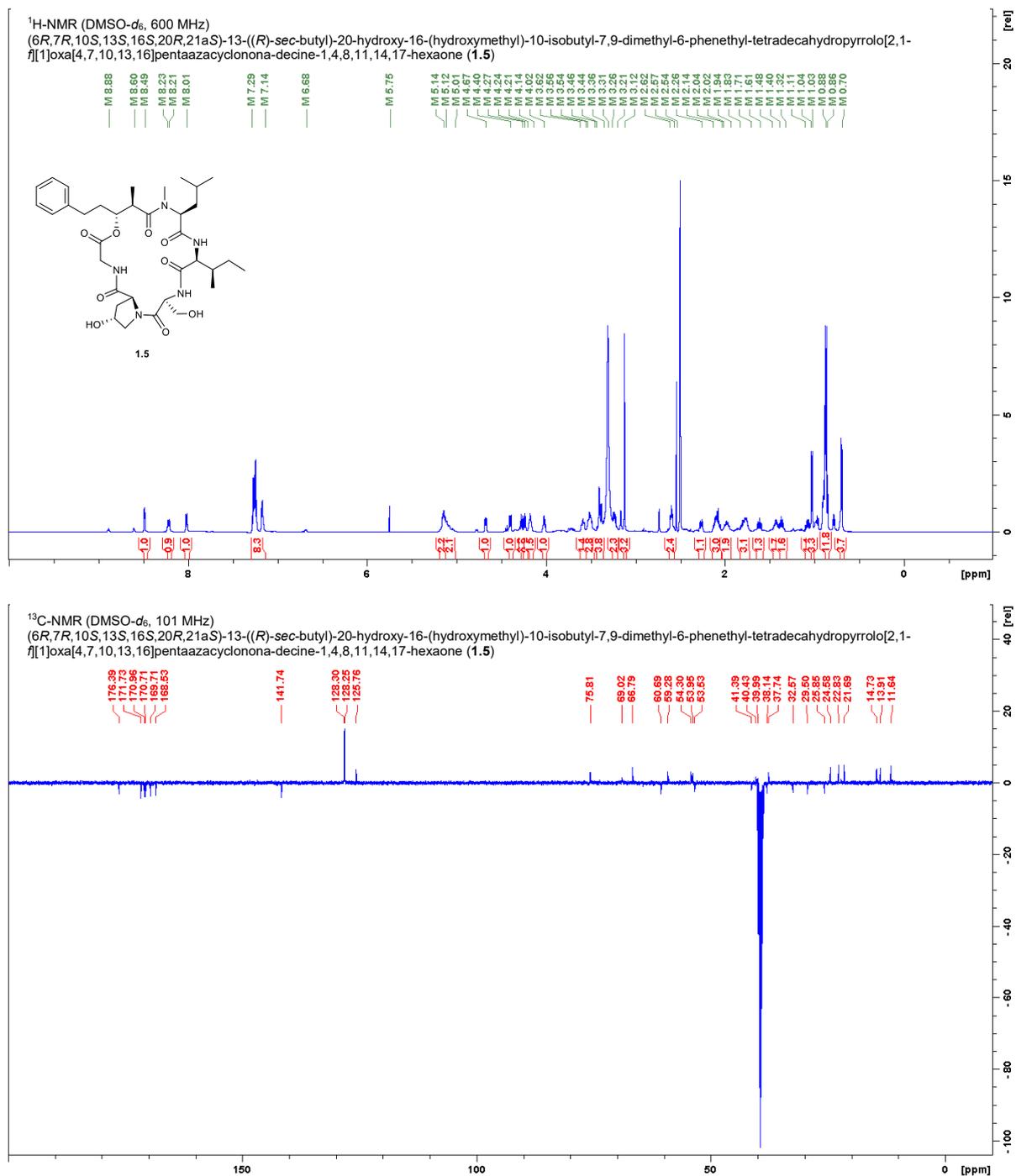


Figure S 4-35: NMR spectra of **1.5** in DMSO-*d*₆.

Table S 4-4: NMR data for 1.5 in DMSO-*d*₆ (600 MHz, 100 MHz).

Position	$\delta^{13}\text{C}$ [ppm]	$\delta^1\text{H}$ [ppm]; (m, f, J)
L- <i>N</i> -Me-Leucine		
CO	171.0	-
N-CH ₃	40.0	3.12 (s, 3H)
α -CH	69.0	3.54-3.46 (m, 3H, overlay)
β -CH ₂	38.1	2.14-.2.04 (m, 3H, overlay)
γ -CH	24.6	1.61 (ddd, 1H, J = 13.9, 8.2, 5.9 Hz)
δ -CH ₃	22.8, 21.7	1.48-1.40 (m, 1H)
		0.88 (d, 6H J = 6.3 Hz)
L- <i>allo</i> -Isoleucine		
CO	171.7	-
NH	-	8.21 (d, 1H, J = 9.4 Hz)
		6.68 (minor)
α -CH	53.9	4.40 (dd, 1H, J = 9.4, 4.5 Hz)
β -CH	37.7	1.83-1.71 (m, 3H, overlay)
γ -CH ₂	25.8	1.40-1.32 (m, 1H)
		1.11-1.04 (m, 1H)
γ -CH ₃	13.9	0.70 (d, 3H, J = 6.9 Hz)
δ -CH ₃	11.6	0.86 (t, 3H, J = 7.5 Hz)
L-Serine		
CO	169.7	-
NH	-	8.49 (d, 1H, J = 4.2 Hz)
		8.60 (minor)
α -CH	54.3	4.02 (ddd, 1H, J = 7.0, 7.0, 4.4 Hz)
β -CH ₂ -OH	60.7	3.62-3.56 (m, 1H)
β -OH	-	3.54-3.46 (m, 3H, overlay)
		5.12-5.01 (m, 2H, overlay)
L-4-Hydroxyproline		
CO	170.7	-
N	-	-
α -CH	59.3	4.67 (dd, 1H, J = 8.5, 4.2 Hz)
β -CH ₂	38.7*	2.26 (ddd, 1H, J = 12.6, 5.0, 4.8 Hz)
		2.14-.2.04 (m, 3H, overlay)
γ -CH-OH	66.8	4.21-4.14 (m, 1H)
γ -OH	-	5.12-5.01 (m, 2H, overlay)
δ -CH ₂	53.5	3.44-3.36 (m, 4H, overlay)
Glycine		
CO	168.5	-
NH	-	8.01 (d, 1H, J = 8.4 Hz)
		8.88 (minor)
CH ₂	41.4	4.27 (d, 1H, J = 8.5 Hz)
		4.24 (d, 1H, J = 8.1 Hz)
		3.44-3.36 (m, 4H, overlay) minor
BHA		
1-CO	176.4	-

2-CH	39.7*	3.26-3.21 (m, 2H)
2'-CH₃	14.7	1.03 (d, 3H, <i>J</i> = 7.1 Hz)
3-CH-O	75.8	5.14 (ddd, 1H, <i>J</i> = 10.0, 6.0, 3.9 Hz)
4-CH₂	32.6	2.02-1.94 (m, 2H)
5-CH₂	29.5	1.83-1.71 (m, 3H, overlay)
C_{quart}	141.7	-
CH_{arom}	128.3, 128.2, 125.8	7.29-7.14 (m, 8H)

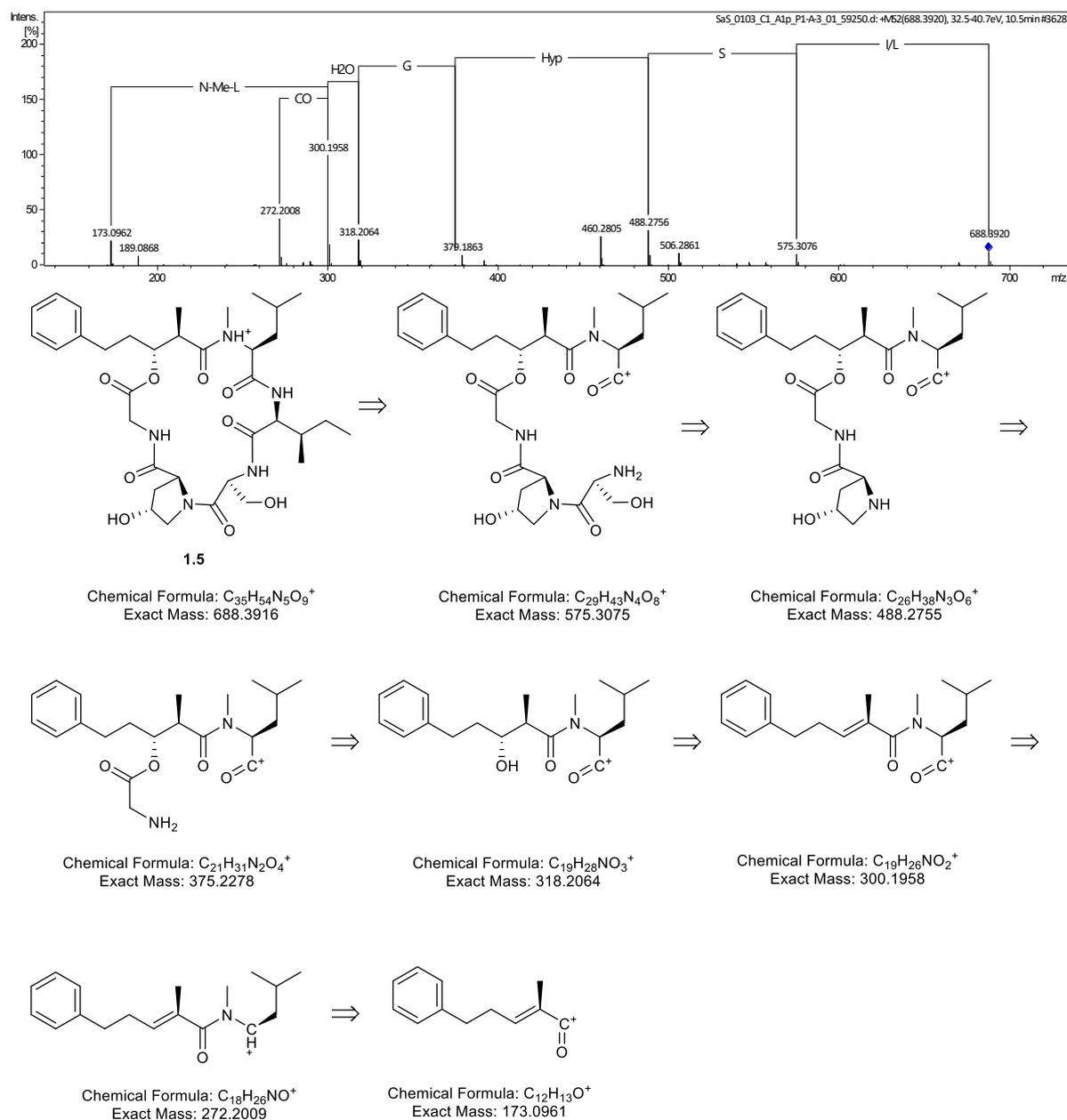


Figure S 4-36: MS² fragmentation of 1.5 and its fragmentation pattern.

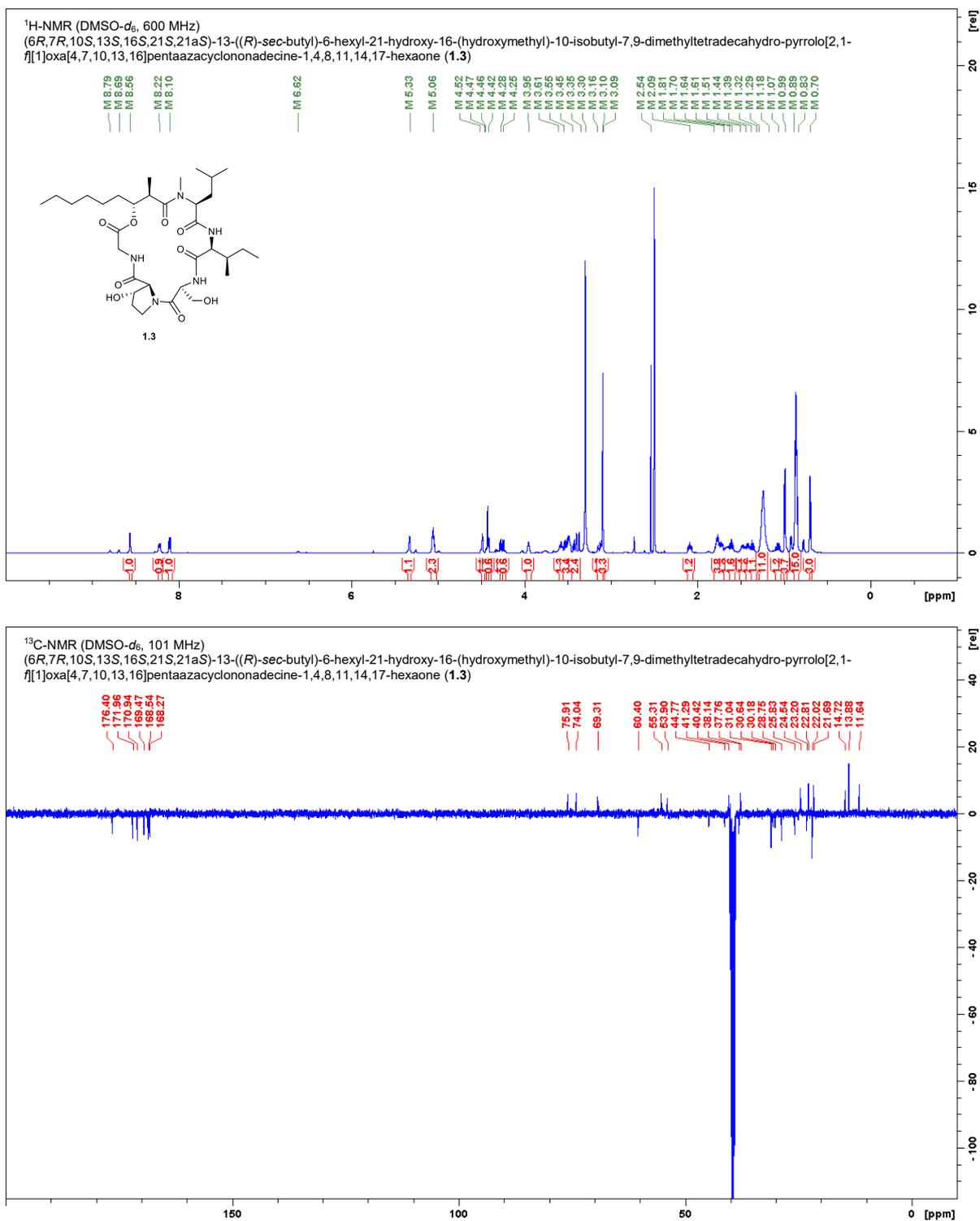


Figure S 4-37: NMR spectra of **1.3** in DMSO-*d*₆.

Table S 4-5: NMR data for 1.3 in DMSO-*d*₆ (600 MHz, 100 MHz).

Position	δ ¹³ C [ppm]	δ ¹ H [ppm]; (m, f, J)
L- <i>N</i> -Me-Leucine		
CO	170.9	-
N-CH ₃	40.1*	3.09 (s, 3H)
α -CH	69.1*	3.55-3.45 (m, 2H, overlay)
β -CH ₂	38.1	2.09 (ddd, 1H, J = 13.7, 9.4, 5.3 Hz) 1.61 (ddd, 1H, J = 14.0, 8.2, 6.1 Hz)
γ -CH	24.5	1.44-1.39 (m, 1H)
δ -CH ₃ a	22.8	0.89-0.83 (m, 15H, overlay)
δ -CH ₃ b	21.7	0.89-0.83 (m, 15H, overlay)
L- <i>allo</i> -Isoleucine		
CO	172.0	-
NH	-	8.22 (d, 1H, J = 9.4 Hz) 6.62 (minor)
α -CH	53.9	4.42 (dd, 1H, J = 9.2, 3.2 Hz)
β -CH	37.8	1.81-1.70 (m, 4H, overlay)
γ -CH ₂	25.8	1.39-1.32 (m, 1H) 1.07 (ddd, 1H, J = 13.7, 7.5, 7.5 Hz)
γ -CH ₃	13.9	0.70 (d, 3H, J = 6.9 Hz)
δ -CH ₃	11.6	0.89-0.83 (m, 15H, overlay)
L-Serine		
CO	168.3	-
NH	-	8.56 (d, 1H, J = 4.2 Hz) 8.69 (minor)
α -CH	55.3	3.95 (ddd, 1H, J = 7.3, 6.3, 4.4 Hz)
β -CH ₂ -OH	60.4	3.61-3.55 (m, 1H) 3.55-3.45 (m, 3H, overlay)
β -OH	-	-
L-3-Hydroxyproline		
CO	169.5	-
N	-	-
α -CH	69.3	4.46-4.42 (m, 1H,)
β -CH-OH	74.0	4.52-4.47 (m, 1H)
β -OH	-	5.33 (d, 1H, J = 3.6 Hz)
γ -CH ₂	30.2	1.81-1.70 (m, 4H, overlay)
δ -CH ₂	44.8	3.55-3.45 (m, 3H, overlay) 3.45-3.35 (m, 2H, overlay)
Glycine		
CO	168.5	-
NH	-	8.10 (d, 1H, J = 8.6 Hz) 8.79 (minor)
CH ₂	41.3	4.28 (d, 1H, J = 8.8 Hz) 4.25 (d, 1H, J = 8.8 Hz) 3.45-3.35 (m, 2H, overlay) minor
BHA		

1-CO	176.4	-
2-CH	40.1*	3.16-3.10 (m, 1H)
2'-CH₃	14.7	0.99 (d, 3H, J = 6.9 Hz)
3-CH-O	75.9	5.06 (ddd, 1H, J = 10.0, 6.3, 3.7 Hz)
4-CH₂	30.6	1.70-1.64 (m, 1H) 1.51-1.44 (m, 1H)
5-8-CH₂	31.0, 28.7, 23.2, 22.0	1.29-1.18 (m, 11H, overlay)
9-CH₃	13.9	0.89-0.83 (m, 15H, overlay)

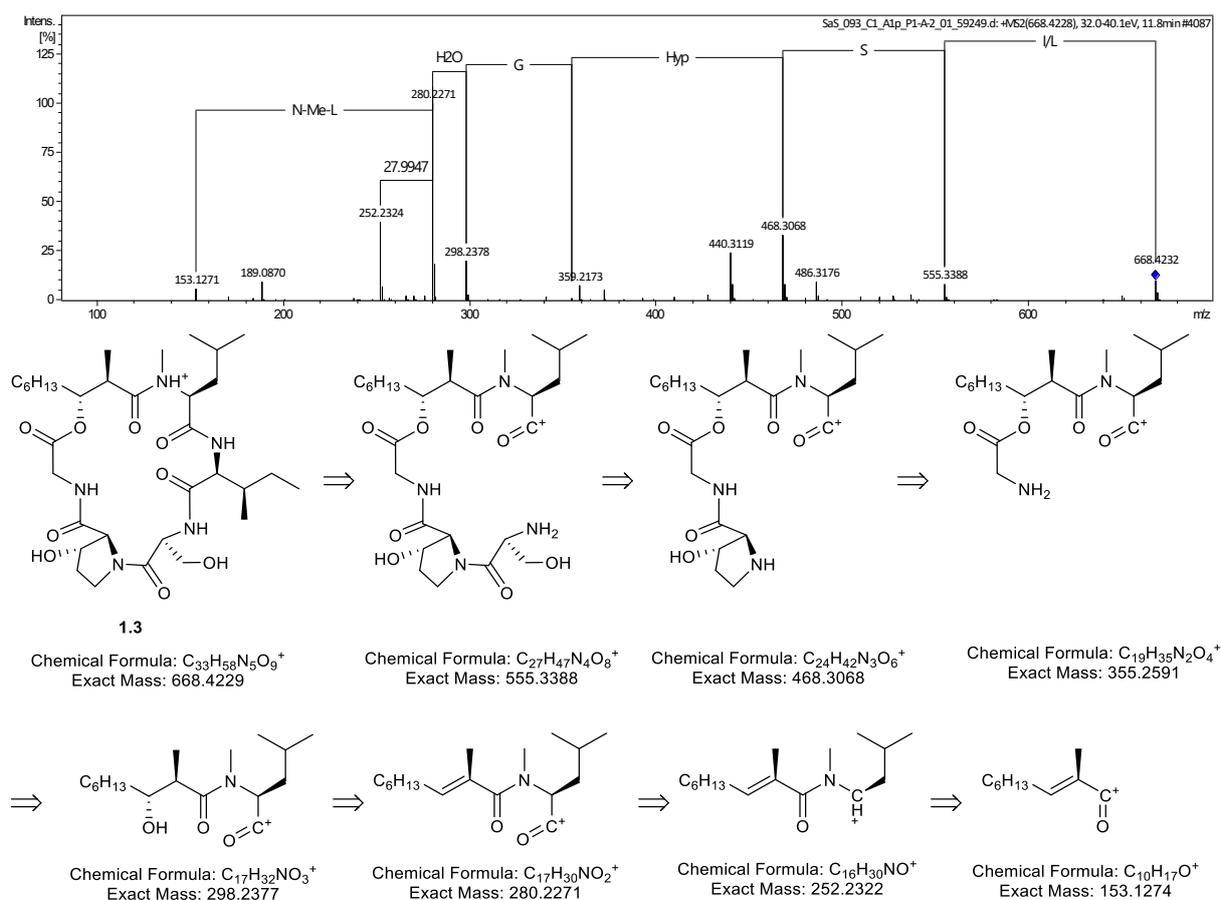


Figure S 4-38: MS² fragmentation of 1.3 and its fragmentation pattern

Table S 4-6: NMR data for 1.4 in DMSO-*d*₆ (600 MHz, 100 MHz).

Position	$\delta^{13}\text{C}$ [ppm]	$\delta^1\text{H}$ [ppm]; (m, f, J)
L- <i>N</i> -Me-Leucine		
CO	170.9	-
N-CH ₃	40.0*	3.12 (s, 3H)
α -CH	69.0	3.55-3.46 (m, 3H, overlay)
β -CH ₂	38.1	2.10 (ddd, 1H, J = 13.9, 9.2, 5.0 Hz) 1.62 (ddd, 1H, J = 13.9, 8.2, 6.2 Hz)
γ -CH	24.6	1.46-1.40 (m, 1H)
δ -CH ₃ a	22.8	0.92-0.84 (m, 12H, overlay)
δ -CH ₃ b	21.7	0.92-0.84 (m, 12H, overlay)
L- <i>allo</i> -Isoleucine		
CO	172.0	-
NH	-	8.23 (d, 1H, J = 9.4 Hz) 6.62 (minor)
α -CH	53.9	4.45-4.41 (m, 2H, overlay)
β -CH	37.7	1.83-1.69 (m, 5H, overlay)
γ -CH ₂	25.9	1.40-1.33 (m, 1H) 1.11-1.05 (m, 1H)
γ -CH ₃	13.9	0.70 (d, 3H, J = 6.9 Hz)
δ -CH ₃	11.6	0.92-0.84 (m, 12H, overlay)
L-Serine		
CO	169.5	-
NH	-	8.58 (d, 1H, J = 4.1 Hz) 8.77 (minor)
α -CH	55.3	3.96 (ddd, 1H, J = 6.4, 6.4, 4.7 Hz)
β -CH ₂ -OH	60.4	3.61-3.55 (m, 1H) 3.55-3.46 (m, 3H, overlay)
β -OH	-	n.o.
L-3-Hydroxyproline		
CO	168.4	-
N	-	-
α -CH	69.3	4.45-4.41 (m, 2H, overlay)
β -CH-OH	74.1	4.49 (d, 1H, J = 3.8 Hz)
β -OH	-	n.o.
γ -CH ₂	30.2	1.83-1.69 (m, 5H, overlay)
δ -CH ₂	44.8	3.55-3.46 (m, 3H, overlay) 3.46-3.39 (m, 1H)
Glycine		
CO	168.5	-
NH	-	8.05 (d, 1H, J = 8.7 Hz) 8.90 (minor)
CH ₂	41.3	4.30 (d, 1H, J = 8.5 Hz) 4.27 (d, 1H, J = 8.9 Hz) 3.39-3.34 (m, 2H, overlay) minor
BHA		

1-CO	176.4	-
2-CH	39.7*	3.25-3.19 (m, 2H)
2'-CH₃	14.7	1.03 (d, 3H, <i>J</i> = 6.8 Hz)
3-CH-O	75.8	5.15 (ddd, 1H, <i>J</i> = 9.9, 6.2, 4.0 Hz)
4-CH₂	32.6	2.01-1.93 (m, 1H)
		1.83-1.69 (m, 5H, overlay)
5-CH₂	29.5	2.62-2.56 (m, 2H)
C_{quart}	141.7	-
CH_{arom}	128.3, 128.2,	7.29-7.14 (m, 5H)
	125.7	

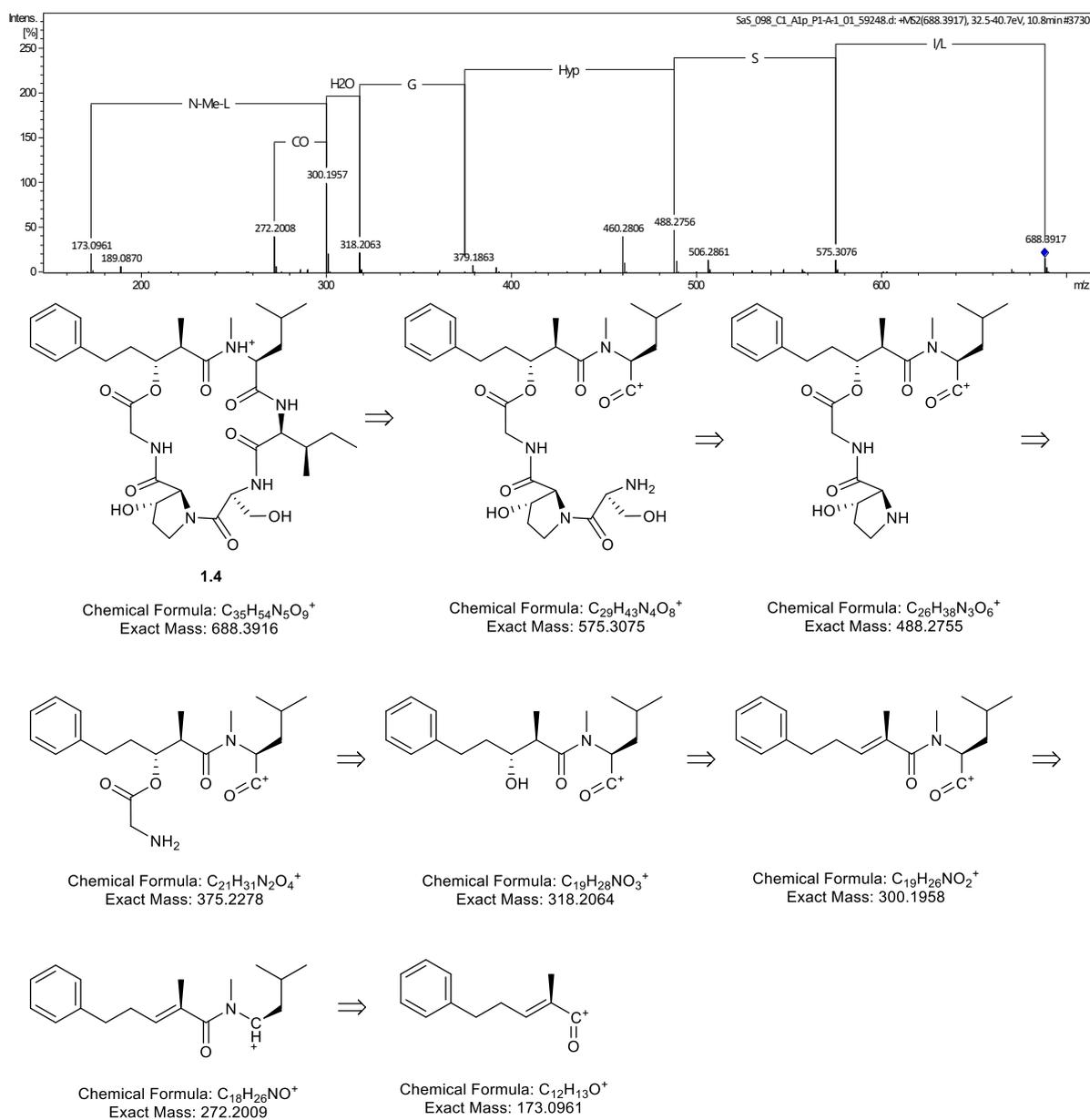


Figure S 4-40: MS² fragmentation of 1.4 and its fragmentation pattern.

5. Chapter 2 – Falcitidin analogs

This work is a joined project with fellow PhD student Stephan Brinkmann. The compound series was firstly identified in the context of the PPP with Evotec and was followed up afterwards as part of a collaboration with the work group of Prof. Dr. Schirmeister from the University of Mainz and Prof. Rosenthal, MD of the University of California, San Francisco.

My main contributions to the project are:

- Input to the structure elucidation of the detected masses of the molecular network
- Development of the synthetic strategy
- Syntheses and analytics of 18 derivatives
- Manuscript drafting

Submitted manuscript in: ACS Chemical Biology (2022):

<https://pubs.acs.org/doi/full/10.1021/acscchembio.1c00861>

5.1 Manuscript

Identification, Characterization, and Synthesis of Natural Parasitic Cysteine Protease Inhibitors: Pentacitidins Are More Potent Falcitidin Analogues

Stephan Brinkmann,[▽] Sandra Semmler,[▽] Christian Kersten, Maria A. Patras, Michael Kurz, Natalie Fuchs, Stefan J. Hammerschmidt, Jenny Legac, Peter E. Hammann, Andreas Vilcinskas, Philip J. Rosenthal, Tanja Schirmeister, Armin Bauer,* and Till F. Schäberle*

Cite This: *ACS Chem. Biol.* 2022, 17, 576–589

Read Online

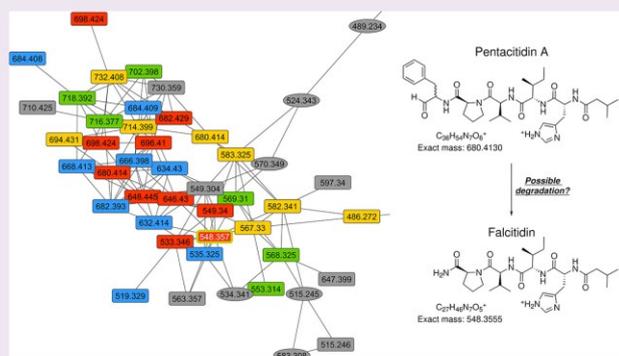
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

Downloaded via PEPPER INC on April 19, 2022 at 08:39:06 (UTC).
See <https://pubs.acs.org/sharingguidelines> for options on how to legitimately share published articles.



ABSTRACT: Protease inhibitors represent a promising therapeutic option for the treatment of parasitic diseases such as malaria and human African trypanosomiasis. Falcitidin was the first member of a new class of inhibitors of falcipain-2, a cysteine protease of the malaria parasite *Plasmodium falciparum*. Using a metabolomics dataset of 25 *Chitinophaga* strains for molecular networking enabled identification of over 30 natural analogues of falcitidin. Based on MS/MS spectra, they vary in their amino acid chain length, sequence, acyl residue, and C-terminal functionalization; therefore, they were grouped into the four falcitidin peptide families A–D. The isolation, characterization, and absolute structure elucidation of two falcitidin-related pentapeptide aldehyde analogues by extensive MS/MS spectrometry and NMR spectroscopy in combination with advanced Marfey's analysis was in agreement with the in silico analysis of the corresponding biosynthetic gene cluster. Total synthesis of chosen pentapeptide analogues followed by in vitro testing against a panel of proteases revealed selective parasitic cysteine protease inhibition and, additionally, low-micromolar inhibition of α -chymotrypsin. The pentapeptides investigated here showed superior inhibitory activity compared to falcitidin.

INTRODUCTION

Parasitic diseases such as malaria and sleeping sickness (human African trypanosomiasis, HAT) are poverty-associated diseases that have a massive impact on human life, especially in tropical areas.^{1,2} Protozoan parasites of the *Trypanosoma* genus, transmitted by tsetse flies, are responsible for HAT.² With fewer than 3000 cases in 2015, the number of reported cases of HAT has reached a historically low level in recent years.² On the contrary, 229 million malaria cases and 409,000 deaths were estimated in 2019.³ *Plasmodium falciparum* is the most virulent malaria parasite, causing nearly all deaths³ as well as most drug-resistant infections.^{4,5} Resistance to most available antimalarials, including artemisinin-based combination regi-

mens that are the standard treatment for falciparum malaria, is of great concern. Hence, there is a great need to discover new active compounds to serve as lead structures for the development of new antimalarial drugs.^{5,4}

Received: October 30, 2021
Accepted: February 25, 2022
Published: March 9, 2022



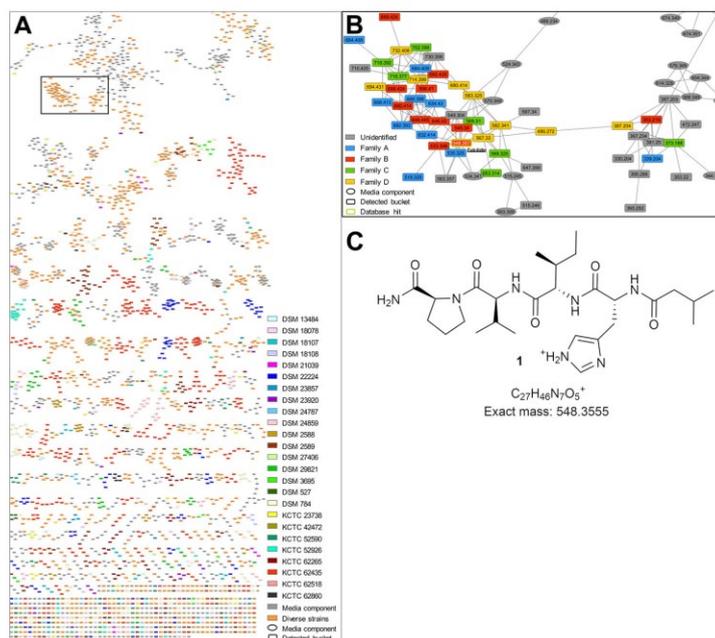


Figure 1. MS²-networking analysis revealed more than 30 natural falcitidin analogues. (a) Complete MS²-network based on extracts of 25 *Chitinophaga* strains. (b) Close view of the falcitidin network color-coded according to the four molecule families A–D. (c) Molecular structure, formula, and calculated mass of falcitidin (**1**), which was putatively annotated for the parent ion of m/z 548.357 [$M + H$]⁺.

Papain-family cysteine proteases represent interesting therapeutic targets against several infectious diseases, including malaria (falcipains) and HAT (rhodesain).⁵ Synthetic or natural product-based small molecule inhibitors were investigated and proved falcipains and rhodesain to be promising targets.^{7,8} However, developing compounds that target these cysteine proteases into drugs has proved challenging, in part due to the difficulty of achieving selectivity.⁹ The acyltetrapeptide falcitidin and its synthetic analogues were described as first members of a new class of cysteine protease inhibitors. In a falcipain-2 assay, the natural compound showed an IC₅₀ value in the low micromolar range, and the optimized synthetic analogues showed sub-micromolar IC₅₀ activity against *Plasmodium falciparum* in a standard blood-cell assay. Instead of reactive groups that covalently-reversibly or covalently-irreversibly bind to active-site cysteines, these compounds share a C-terminal amidated proline.^{10,11}

High structural diversity in combination with diverse biological activities make natural products a valuable source in the search for new lead candidates for drug development. Technological improvements in resolution and accuracy of spectrometry methods followed by complex and large dataset analysis allow the application of metabolomics techniques based on, for example, ultrahigh-performance liquid chromatography in line with high-resolution tandem mass spectrometry (UHPLC-HRMS/MS) for the discovery and characterization of these metabolites present in a given sample.¹² Data visualization and interpretation using tandem mass spectrometry (MS/MS) networks (molecular networking) allow the

automatic annotation of MS/MS spectra against library compounds as well as the identification of signals of interest and their analogues.¹³

In this study, molecular networking was applied to analyze a previously generated dataset based on bacterial organic extracts originated from 25 *Chitinophaga* strains (phylum Bacteroidetes) in which falcitidin was putatively annotated.¹⁴ We aimed to investigate if additional natural analogues were produced by these strains and discovered more than 30 natural analogues. A gene encoding a multimodular nonribosomal peptide synthetase was identified and linked in silico to the production of the pentapeptide aldehydes. After isolation and structure elucidation of two natural aldehyde analogues, total synthesis of all four pentapeptide aldehydes and their carboxylic acid and alcohol derivatives was successfully achieved. In activity testing against a panel of proteases, they performed better than the reference falcitidin and revealed low-micromolar IC₅₀ activity against α -chymotrypsin, falcipain-2, and falcipain-3. Three of them also showed activity against rhodesain.

RESULTS AND DISCUSSION

Discovery of Falcitidin Analogues. Exploration of the genomic and metabolomic potential of the Bacteroidetes genus *Chitinophaga* revealed a high potential to find chemical novelty.¹⁴ In the framework of this previous metabolomic study of *Chitinophaga*, an inhibitor of the antimalarial target falcipain-2, falcitidin (**1**),¹⁰ was detected.¹⁴ Originally wrongly described as myxobacterium-derived tetrapeptides produced

Table 1. Overview of Molecular Features of all Identified Natural Falcitidin Analogues Aligned into Four Protein Families

family	fatty acid residue	amino acid chain					C-terminal group	molecular formula	molecular mass			ion	cpd no.
		1	2	3	4	5			calcd	MS ² network	found		
A	iVal	H	V				COOH	C ₁₇ H ₂₇ N ₄ O ₄	339.2027	339.204	339.2030	[M+H] ⁺	
	iVal	H	V	V	P		CHO	C ₂₇ H ₄₂ N ₆ O ₅	519.3289	519.329	519.3297	[M+H] ⁺	
	iVal	H	V	V	P		CONH ₂	C ₂₆ H ₄₄ N ₇ O ₅	534.3398	534.341	534.3402	[M+H] ⁺	
	iVal	H	V	V	P		COOH	C ₂₆ H ₄₃ N ₆ O ₆	535.3239	535.325	535.3244	[M+H] ⁺	
	iVal	H	V	V	P	L	CHO	C ₃₀ H ₅₂ N ₇ O ₆	632.4130	632.414	632.4148	[M+H] ⁺	
	iVal	H	V	V	P	L	CH ₂ OH	C ₃₂ H ₅₆ N ₇ O ₆	634.4287	634.430	634.4298	[M+H] ⁺	
	iVal	H	V	V	P	F	CHO	C ₃₃ H ₅₂ N ₇ O ₆	666.3974	666.398	666.3982	[M+H] ⁺	15
	iVal	H	V	V	P	F	CH ₂ OH	C ₃₃ H ₅₂ N ₇ O ₆	668.4130	668.413	668.4117	[M+H] ⁺	12
	iVal	H	V	V	P	F	COOH	C ₃₃ H ₅₂ N ₇ O ₇	682.3923	682.393	682.3918	[M+H] ⁺	8
	iVal	H	V	V	P	F	CHO	C ₃₃ H ₅₄ N ₇ O ₇	684.4079	684.408	684.4085	[M+H ₂ O+H] ⁺	15 ^{ac}
B	iVal	H	I				COOH	C ₁₇ H ₂₉ N ₄ O ₄	353.2183	353.219	353.2185	[M+H] ⁺	
	iVal	H	I	V	P		CHO	C ₂₇ H ₄₃ N ₆ O ₅	533.3446	533.346	533.3442	[M+H] ⁺	
	iVal	H	I	V	P		CONH ₂	C ₂₇ H ₄₆ N ₇ O ₅	548.3555	548.357	548.3563	[M+H] ⁺	1
	iVal	H	I	V	P		COOH	C ₂₇ H ₄₅ N ₆ O ₆	549.3395	549.340	549.3397	[M+H] ⁺	17
	iVal	H	I	V	P	L	CHO	C ₃₃ H ₅₂ N ₇ O ₆	646.4287	646.430	646.4305	[M+H] ⁺	
	iVal	H	I	V	P	L	CH ₂ OH	C ₃₃ H ₅₂ N ₇ O ₆	648.4443	648.445	648.4445	[M+H] ⁺	
	iVal	H	I	V	P	F	CHO	C ₃₃ H ₅₄ N ₇ O ₆	680.4130	680.414	680.4131	[M+H] ⁺	2
	iVal	H	I	V	P	F	CH ₂ OH	C ₃₃ H ₅₂ N ₇ O ₆	682.4287	682.429	682.4270	[M+H] ⁺	14
	iVal	H	I	V	P	F	COOH	C ₃₃ H ₅₂ N ₇ O ₇	696.4079	696.410	696.4072	[M+H] ⁺	10
	iVal	H	I	V	P	F	CHO	C ₃₃ H ₅₆ N ₇ O ₇	698.4236	698.424	698.4238	[M+H ₂ O+H] ⁺	2 ^{ac}
C	PA	H	V				COOH	C ₁₇ H ₂₅ N ₄ O ₄	373.1870	373.188	373.1874	[M+H] ⁺	
	PA	H	V	V			COOH	C ₂₄ H ₃₂ N ₅ O ₅	472.2554	not present	472.2560	[M+H] ⁺	
	PA	H	V	V	P		CHO	C ₂₆ H ₄₁ N ₆ O ₅	553.3133	553.314	553.3135	[M+H] ⁺	
	PA	H	V	V	P		CONH ₂	C ₂₆ H ₄₂ N ₇ O ₅	568.3242	568.325	568.3248	[M+H] ⁺	
	PA	H	V	V	P		COOH	C ₂₆ H ₄₁ N ₆ O ₆	569.3082	569.310	569.3101	[M+H] ⁺	
	PA	H	V	V	P	F	CHO	C ₃₀ H ₅₀ N ₇ O ₆	700.3817	not present	700.3806	[M+H] ⁺	16
	PA	H	V	V	P	F	CH ₂ OH	C ₃₀ H ₅₂ N ₇ O ₆	702.3974	702.398	702.3956	[M+H] ⁺	13
	PA	H	V	V	P	F	COOH	C ₃₀ H ₅₀ N ₇ O ₇	716.3766	716.377	716.3765	[M+H] ⁺	9
	PA	H	V	V	P	F	CHO	C ₃₀ H ₅₂ N ₇ O ₇	718.3923	718.392	718.3923	[M+H ₂ O+H] ⁺	16 ^{ac}
	PA	H	I				COOH	C ₂₆ H ₃₂ N ₅ O ₅	387.2027	387.203	387.2033	[M+H] ⁺	
D	PA	H	I	V			COOH	C ₂₅ H ₃₆ N ₅ O ₅	486.2711	486.272	486.2715	[M+H] ⁺	
	PA	H	I	V	P		CHO	C ₃₀ H ₄₃ N ₆ O ₅	567.3289	567.330	567.3300	[M+H] ⁺	
	PA	H	I	V	P		CONH ₂	C ₃₀ H ₄₄ N ₇ O ₅	582.3398	582.341	582.3408	[M+H] ⁺	
	PA	H	I	V	P		COOH	C ₃₀ H ₄₃ N ₆ O ₆	583.3239	583.325	583.3252	[M+H] ⁺	
	PA	H	I	V	P	L	CHO	C ₃₃ H ₅₄ N ₇ O ₆	680.4130	680.414	680.4115	[M+H] ⁺	
	PA	H	I	V	P	L	CH ₂ OH	C ₃₃ H ₅₂ N ₇ O ₆	682.4287	not present	682.4288	[M+H] ⁺	
	PA	H	I	V	P	F	CHO	C ₃₃ H ₅₂ N ₇ O ₆	714.3974	714.399	714.3975	[M+H] ⁺	3
	PA	H	I	V	P	F	CH ₂ OH	C ₃₃ H ₅₄ N ₇ O ₆	716.4136	not present	716.4132	[M+H] ⁺	4
	PA	H	I	V	P	F	COOH	C ₃₃ H ₅₂ N ₇ O ₇	730.3923	not present	730.3939	[M+H] ⁺	11
	PA	H	I	V	P	F	CHO	C ₃₃ H ₅₄ N ₇ O ₇	732.4079	732.408	732.4082	[M+H ₂ O+H] ⁺	3 ^{ac}

^{ac}Nodes found in MS² network, no MS/MS spectra found due to water ion adduct.

by *Chitinophaga* sp. Y23, falcitidin and its synthetic analogues are first members of a new class of cysteine protease inhibitors without a reactive group that covalent-reversibly or covalent-irreversibly binds to the active-site cysteine residue.^{10,11} Based on extracts of all 25 previously cultured *Chitinophaga* strains,¹⁴ here, a complementary molecular networking study was carried out to analyze the chemical profiles (Figure 1a). This facilitated the discovery of more than 30 natural analogues of falcitidin (1, m/z 548.3563 [M+H]⁺, molecular formula C₂₇H₄₆N₇O₅) (Figure 1b,c and Table 1) biosynthesized by *Chitinophaga eiseniae* DSM 22224, *Chitinophaga dinghuensis* DSM 29821, and *Chitinophaga varians* KCTC 52926. Analysis of the corresponding MS/MS spectra led to the assignment of four peptide families that differed in the amino acid chain at position 2 and the acyl residue of the peptide. Members of families A and B shared a fragment ion of m/z 222.1237 [M

+H]⁺ corresponding to the molecular formula C₁₁H₁₆N₃O₂⁺. The fragment can be assigned to the described isovaleroyl (iVal) residue attached to the first amino acid histidine.¹⁰ In contrast, a fragment ion of m/z 256.1084 [M+H]⁺, corresponding to the molecular formula C₁₄H₁₄N₃O₂⁺, was detected in peptides of families C and D (Figures S1–S4). A difference within the acyl moiety was assumed due to the identical number of three nitrogen atoms, also indicating a histidine at position one. Moreover, the MS/MS fragmentation pattern of family members A and B as well as C and D varied in position 2 of the amino acid chain, either carrying valine or isoleucine. Within each peptide family, members differed in chain length from two to five amino acids and their C-terminal functional group (Table 1). Interestingly, a corresponding falcitidin analogue with the unusual C-terminal amidated (CONH₂-group) proline, indicated by a neutral loss of

Table 2. NMR Data Comparison for Pentacitidin A (2) in DMSO-*d*₆ for the Natural Isolated Compound (700 MHz, 176 MHz) and the Synthetic Compound (600 MHz, 151 MHz)^a

position	$\delta^{13}\text{C}$ (ppm)		$\delta^1\text{H}$ (ppm) (m, f, J)		position	$\delta^{13}\text{C}$ (ppm)		$\delta^1\text{H}$ (ppm) (m, f, J)	
	natural isolated	synthetic	natural isolated	synthetic		natural isolated	synthetic	natural isolated	synthetic
phenylalaninal									
CHO	200.3	200.3	9.47, 9.41	9.47 (s, 0.2H) 9.41 (s, 0.2H)	isoleucine	18.5	18.45/18.41	0.86	
NH			8.38	8.36 (d, 0.2H, J = 7.7 Hz) 8.34 (d, 0.2H, J = 7.0 Hz)	CO	170.7	170.7	7.82	7.69–7.63 (m, 1H)
α -CH	59.5	59.7/59.5	4.30	4.35–4.24 (m, 2H, overlay) 4.24–4.16 (m, 2H, overlay)	α -CH	56.3 ^b	56.32/56.29	4.25	4.24–4.16 (m, 2H)
β -CH ₂	33.5	33.6/33.3	3.13, 2.74	3.15–3.07 (m, 1H) 2.91–2.83 (m, 2H, overlay) 2.78–2.69 (m, 2H, overlay)	β -CH	36.9 ^b	36.9/36.8	1.65	1.70–1.60 (m, 2H)
γ -C _{quart}	137.6	137.6			γ -CH ₂	24.0	24.0	1.25	1.30–1.22 (m, 1H)
CH _{anom}	129.3, 128.1, 126.3	129.3, 129.2, 129.1, 128.9, 128.2, 128.1, 126.24, 126.19	7.23, 7.27, n.a.	7.29–7.12 (m, 5H)	γ -CH ₃	15.2	15.2	0.71	0.76–0.67 (m, 6H, overlay)
proline					δ -CH ₃	11.0	11.0	0.74	0.76–0.67 (m, 6H, overlay)
CO	172.0 ^b	172.1/172.0			histidine				
α -CH	59.1 ^b	59.1/59.0	4.29	4.35–4.24 (m, 2H, overlay)	CO	170.0	170.9	8.16	8.00–7.94 (m, 2H, overlay)
β -CH ₂	29.3	29.31/29.28	1.92, 1.54	2.02–1.83 (m, 5H, overlay) 1.83–1.70 (m, 2H, overlay) 1.60–1.51 (m, 1H)	NH			4.71	4.61–4.55 (m, 1H)
γ -CH ₂	24.4 ^b	24.44/24.38	1.79, 1.74	2.02–1.83 (m, 5H, overlay) 1.83–1.70 (m, 2H, overlay)	α -CH	51.7 ^b	52.7	4.71	4.61–4.55 (m, 1H)
δ -CH ₂ N	47.0 ^b	47.00/46.96	3.75, 3.53	3.78–3.68 (m, 1H) 3.57–3.47 (m, 1H)	β -CH ₂	27.9 ^b	29.9*	3.03	2.91–2.83 (m, 2H, overlay) 2.85
valine								2.85	2.78–2.69 (m, 2H, overlay)
CO	169.6	169.58/169.55			γ -C _{quart}	n.a.	134.2		
NH			8.05	8.00–7.94 (m, 2H, overlay)	δ -CH _{anom}	116.7 ^b	n.o.	~7.2	6.76 (s, 1H)
α -CH	55.8	55.7	4.25	4.35–4.24 (m, 2H, overlay)	ϵ -CH _{anom}	133.8	134.5	~8.7	7.50 (s, 1H)
β -CH	29.7	29.7	1.98	2.02–1.83 (m, 5H, overlay)	ϵ -NH				n.o.
γ -CH ₃	19.0	19.03/19.00	0.88	0.90–0.84 (m, 6H, overlay)	isovaleroyl				
					CO	171.6	171.4		
					CH ₂	44.3	44.5	1.96	2.02–1.83 (m, 5H, overlay)
					CH	25.4	25.5	1.89	2.02–1.83 (m, 5H, overlay)
					CH ₃	22.2	22.2	0.80	0.81 (d, 3H, J = 6.4 Hz)
					CH ₃	22.1	22.1	0.77	0.78 (d, 3H, J = 6.4 Hz)

^aDue to the presence of two diastereomers (see above), two sets of signals are obtained. Therefore, for some signals, two chemical shift values are given. n.o., not observed; n.a., not assigned due to line broadening; asterisks denote data obtained from the HMBC or HSQC spectrum; overlay means that the signal overlaps with another signal. ^bBroad signal.

114.0794 Da, was found in each family. Falcitidin (iVal-H-I-V-P-CONH₂) itself belongs to family B. Its analogues with the same neutral loss were assigned to the UHR-ESI-MS ion peaks at m/z 534.3402 [M+H]⁺ (C₂₆H₄₄N₇O₅, family A), m/z 568.3248 [M+H]⁺ (C₂₉H₄₉N₇O₅, family C), and m/z 582.3408 [M+H]⁺ (C₃₀H₄₄N₇O₅, family D). Additionally, within all families, tetrapeptide analogues carrying a C-terminal-unmodified (COOH-group) proline and analogues with an aldehyde moiety (CHO-group) indicated by a neutral loss of 99.0686 Da (C₅H₉NO) were found. Besides these tetrapeptide analogues of falcitidin, truncated di- or tripeptides with C-terminal COOH-groups and larger pentapeptide analogues were predicted based on MS/MS spectra. The spectra of the pentapeptides, which were found in all families, revealed either

phenylalanine or leucine/isoleucine in the C-terminal position. Besides an unmodified C-terminal phenylalanine (COOH-group), additional analogues carrying a C-terminal phenylalaninal (CHO-group) and phenylalaninol (CH₂OH-group) were predicted. The last two analogues were proposed based on the neutral losses of 149.0841 Da (C₉H₁₃NO) and 151.1006 Da (C₉H₁₃NO), respectively. For analogues with a leucine/isoleucine at position 5, only analogues with a CHO- and CH₂OH-group were detected, with neutral losses of 115.1031 Da (C₆H₁₃NO) and 117.1164 Da (C₆H₁₅NO), respectively (Table 1 and Figures S1–S4).

Natural Compound Isolation and Spectroscopic Analysis. In order to confirm the proposed structures of the pentapeptide falcitidin analogues, their functional C-terminal

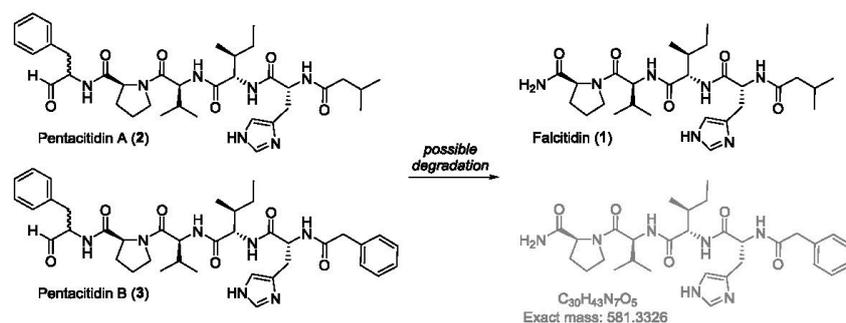


Figure 2. Chemical structures of isolated novel pentacitidin A (2) and B (3) as well as their possible degradation analogues, falcitidin (1) and the corresponding tetrapeptide to pentacitidin B. Black represents synthesized structures; gray represents MS² spectra-confirmed structures.

groups and the unknown acyl chain of peptide families C and D, isolation and spectroscopic analysis of natural pentapeptides with m/z 680.4131 $[M+H]^+$ (2, C₃₆H₅₄N₇O₆), m/z 714.3975 $[M+H]^+$ (3, C₃₉H₅₃N₇O₆), and m/z 716.4132 $[M+H]^+$ (4, C₃₉H₅₄N₇O₆) were achieved. Isolation was performed by the adsorption of the metabolites on XAD16N resin and sequential C18-RP-HPLC followed by C18-RP-UHPLC fractionations starting from a 20 L fermentation of *C. eiseniae* DSM 22224. Based on the analysis of several two-dimensional NMR spectra, including DQF-COSY, TOCSY, ROESY, multiplicity-edited HSQC, and HMBC experiments, the structure of the isolated compound 2 was determined to be the phenylalaninol extended version of falcitidin (iVal-H-I-V-P-F-CHO) determined from the additional aromatic signals $\delta_H = 7.27$ and 7.23 ppm as well as the aldehyde protons at $\delta_H = 9.47$ and 9.41 ppm and $\delta_C = 200.3$ ppm (Table 2 and Figure S5). We proposed the name pentacitidin A. The presence of two signals for several protons and carbon atoms indicated an epimerization of the stereogenic center of the C-terminal phenylalaninol (Figure 2). The analysis of the NMR data of the isolated compound 3 revealed the same peptide sequence. However, instead of the isovaleroyl moiety, 3 contains a phenyl acetyl (PA) residue at the N terminus, leading to the final sequence PA-H-I-V-P-F-CHO (Table 3 and Figure S6). Again, the protons of the terminal aldehyde moiety gave rise to a set of two signals (e.g., δ_H of aldehyde proton = 9.46 and 9.41 ppm), indicating the epimerization of its stereogenic center (Figure 2). Accordingly, we proposed the name pentacitidin B for compound 3. The NMR sample of compound 4 contained a mixture of 4 and aldehyde 3, not allowing for a complete assignment of the NMR signals. The phenylalaninol moiety, as indicated by MS/MS results, was confirmed by the presence of a second methylene group adjacent to C $_{\alpha}$ ($\delta_H = 3.27$ ppm, $\delta_{13C} = 61.8$ ppm), which in the COSY spectrum is correlated with an exchangeable proton at 4.7 ppm (OH). The reduction of the aldehyde function also leads to a significant high-field shift of the H $_{\alpha}$ and the amide proton (Figure S7), verifying the structure of compound 4 to be PA-H-I-V-P-F-CH₂OH (Scheme 1).

Finally, advanced Marfey's analysis was conducted to determine the absolute configuration of the amino acids.¹⁵ Total hydrolysis of a sample containing 3 and 4 followed by chemical derivatization with *N* $_{\alpha}$ -(2,4-dinitro-5-fluorophenyl)-L-valinamide (Marfey's reagent) and UHPLC-MS comparison to reference substrates confirmed the literature-known D-His-L-

Val/Ile-L-Val-L-Pro configuration¹⁰ and the incorporation of L-phenylalanine at position 5 by verifying the presence of L-phenylalaninol (Figure S9). All stereogenic centers were further confirmed by total synthesis and comparison of NMR data to natural isolated compounds (Tables 2 and 3).

Identification of the Biosynthetic Gene Cluster. The publicly available genomes of all three producers, *C. eiseniae* DSM 22224 (FUWZ01000000), *C. dinghuensis* DSM 29821 (QLMA01000000), and *C. varians* KCTC 52926 (JACVFB01000000), were scanned with antiSMASH¹⁶ for NRPS-type biosynthetic gene clusters (BGCs) matching the structural features of the molecules. The number and predicted substrate specificity of the A-domains (Table S1) as well as the precursor supply and post-assembly modifications were taken into account. BGCs congruent to the pentapeptide structure were identified in each case. Furthermore, an epimerization domain to catalyze the conversion of L- to D-amino acids is positioned in agreement with the determined stereochemistry of the molecules. Interestingly, manual search of other publicly available *Chitinophaga* genomes revealed the presence of similar BGCs in genomes of *C. varians* Ae27 (JABAI01000000) and *C. niastensis* DSM 24859 (PYAW00000000) (Figure 3a), the latter also being included in the molecular networking analysis. For *C. niastensis*, however, no production of pentacitidins was observed, which could correlate with the small differences in the A-domain specificity for amino acids 3 (His instead of Val) and 5 (beta-hydroxy-tyrosine (Bht) instead of Phe) (Table S1). The MAFFT alignment¹⁷ of all five BGCs allowed clear cluster boarder prediction and revealed minor variations in the upstream region. In conclusion, pentacitidin biosynthesis is putatively encoded by a single NRPS core gene of 18.5 kbp in length. No additional genes encoding proteins involved in further modifications or transportation were conserved between all five BGCs (Figure 3a). The NRPS starts with a starter condensation (C-starter) domain responsible for the addition of iVal or PA to the peptide core. A terminal reductase domain (TD) at the C-terminal end in four BGCs should be responsible for the reductive release process (Figure 3b).¹⁸ The presence of alcohol and aldehyde pentapeptides suggests a similar reduction as shown for example in the biosynthesis of the siderophore myxochelin.^{19–21} The (peptidyl)acyl thioester attached to the carrier protein could be reduced first to an aldehyde and then to an alcohol via a four-electron reduction during the product release.^{22,23} The

Table 3. NMR Data Comparison for Pentacitidin B (3) in DMSO-*d*₆ for the Natural Isolated Compound (500 MHz, 126 MHz) and the Synthetic Compound (600 MHz, 151 MHz)^a

position	$\delta^{13}\text{C}$ (ppm)		$\delta^1\text{H}$ (ppm) (m, f, J)	
	natural isolated (major)	synthetic	natural isolated (major)	synthetic
phenylalaninal				
CHO	200.3, 200.5	200.3	9.46, 9.41	9.47 (s, 0.2H) 9.41 (s, 0.2H)
NH			8.38, 8.36	8.36 (d, 0.25H, <i>J</i> = 7.7 Hz) 8.34 (d, 0.25H, <i>J</i> = 7.0 Hz)
α -CH	59.5, 59.8	59.1 / 59.0	4.29, 4.17	4.24–4.16 (m, 2H, overlay)
β -CH ₂	33.5, 33.3	33.5 / 33.3	3.12/2.74, 3.09/2.85	3.15–3.06 (m, 1H, overlay) 2.93–2.83 (m, 2H, overlay) 2.78–2.71 (m, 1H, overlay)
γ -C _{quart}	137.7	137.6		
CH _{arom}	129.3, 129.2128.1, 128.2126.2, 126.3	129.3, 129.2, 129.14, 129.09, 129.0, 128.2, 128.1, 128.0, 126.25, 126.20, 126.16	7.21, 7.24, 7.18	7.29–7.13 (m, 10H, overlay)
proline				
CO	172.0	172.1 / 172.0		
α -CH	59.0, 59.1	59.5	4.29, 4.30	4.33–4.24 (m, 2H, overlay)
β -CH ₂	29.3, 29.4	29.32 / 29.29	1.91/1.54, 1.97/1.74	2.02–1.84 (m, 2H, overlay) 1.84–1.69 (m, 2H, overlay) 1.59–1.51 (m, 1H)
γ -CH ₂	24.4, 24.5	24.5 / 24.4	1.78/1.74, 1.91/1.80	2.02–1.84 (m, 2H, overlay) 1.84–1.69 (m, 2H, overlay)
δ -CH ₂ N	46.98, 47.02	47.01 / 46.97	3.75/3.50, 3.75/3.54	3.78–3.68 (m, 1H) 3.57–3.47 (m, 1H)
valine				
CO	169.58	169.61 / 169.58		
NH			7.98, 7.96	7.95 (t, 1H, <i>J</i> = 9.0 Hz)
α -CH	55.8	55.8	4.26	4.33–4.24 (m, 2H, overlay)
β -CH	29.7	29.7	1.97	2.02–1.84 (m, 2H, overlay)
γ -CH ₃	19.0 18.5	19.03 / 19.00 18.5 / 18.4	0.87, 0.85	0.94–0.82 (m, 7H, overlay)
isoleucine				
CO	170.75	170.7		
NH			7.76	7.77–7.71 (m, 1H)
α -CH	56.39	56.40 / 56.37	4.21	4.24–4.16 (m, 2H, overlay)
β -CH	36.79	36.8 / 36.7	1.64	1.69–1.59 (m, 1H)
γ -CH ₂	23.9	23.9	1.22, 0.88	1.27–1.17 (m, 1H) 0.94–0.82 (m, 7H, overlay)
γ -CH ₃	15.3	15.3	0.69	0.75–0.66 (m, 6H, overlay)
δ -CH ₃	11.0	11.0	0.72	0.75–0.66 (m, 6H, overlay)
histidine				
CO	170.8	170.8		
NH			8.29	8.26 (d, 1H, <i>J</i> = 8.1 Hz)
α -CH	52.9	52.9	4.58	4.62–4.55 (m, 1H)
β -CH ₂	n.a.	30.1*	2.89, 2.75	2.93–2.83 (m, 2H, overlay) 2.78–2.71 (m, 1H, overlay)

Table 3. continued

position	$\delta^{13}\text{C}$ (ppm)		$\delta^1\text{H}$ (ppm) (m, f, j)	
	natural isolated (major)	synthetic	natural isolated (major)	synthetic
$\gamma\text{-C}_{\text{quart}}$	n.a.	n.o.		
$\delta\text{-CH}_{\text{arom}}$	n.a.	n.o.	6.75	6.76 (s, 1H)
$\epsilon\text{-CH}_{\text{arom}}$	134.5	134.5	7.49	7.51 (s, 1H)
$\epsilon\text{-NH}$				n.o.
phenylacetyl				
CO	169.9	169.9		
CH_2	42.0	42.0	3.43	3.45 (d, 1H, $J = 14.1$ Hz) 3.42 (d, 1H, $J = 14.1$ Hz)
C_{quart}	136.2	136.2		
CH_{arom}	128.97, 128.05, 126.18	129.3, 129.2, 129.14, 129.09, 129.0, 128.2, 128.1, 128.0, 126.25, 126.20, 126.16	7.17, 7.22, 7.19	7.29–7.13 (m, 10H, overlay)

^an.o., not observed; n.a., not assigned due to line broadening; asterisks denote data obtained from the HMBC or HSQC spectrum; overlay means that the signal overlaps with another signal.

BGC in the genome of *C. varians* Ae27 carries a NAD_binding_4 domain instead of the TD. This domain family is sequence-related to the C-terminal region of the male sterility protein in the arabidopsis species²⁴ and a jojoba acyl-CoA reductase.²⁵ The latter is known to catalyze a similar reduction reaction, with the formation of a fatty alcohol from a fatty acyl substrate.²⁵

Degradation Hypotheses. A BGC congruent to the pentapeptide structure and the presence of the unusual tetrapeptide analogue after NMR study (Figure S5) of its corresponding pentapeptide aldehyde points toward falcitidin and its identified analogues from families A, C, and D with their unusual C-terminal amidated proline to be degradation products (Figure 2). Biochemically, a cleavage by a carboxypeptidase in the presence of ammonia could result in the C-terminal amide group of the proline, which can explain the tetrapeptides in the extract. Chemically, the decomposition of the pentapeptide to the unexpected amide version of the proline could be based on two hypothetical mechanisms: (i) base catalysis by the imidazole moiety of the histidine²⁶ or (ii) an oxidative decomposition via an *N,O*-acetal-like intermediate.²⁷ These hypotheses and their mechanisms will need to be evaluated in further studies. Interestingly, UHPLC control measurements of the pure pentacitidin samples after NMR spectroscopy always revealed the presence of mixtures of the pentapeptide and its corresponding tetrapeptide analogue with the C-terminal amidated proline. The NMR samples were dissolved in DMSO, dried in vacuo, resolved in MeOH, and measured. For example, within the control measurement of pentacitidin B (3), the falcitidin analogue of family D (m/z 582.3408 $[\text{M}+\text{H}]^+$, $\text{C}_{30}\text{H}_{44}\text{N}_7\text{O}_3$) was detected (Figure 2 and Figure S5). However, at the present stage, it also cannot be excluded that falcitidin and pentacitidin analogues are biosynthesized by the same NRPS assembly line and module skipping takes place, which will result in the shorter tetrapeptides.

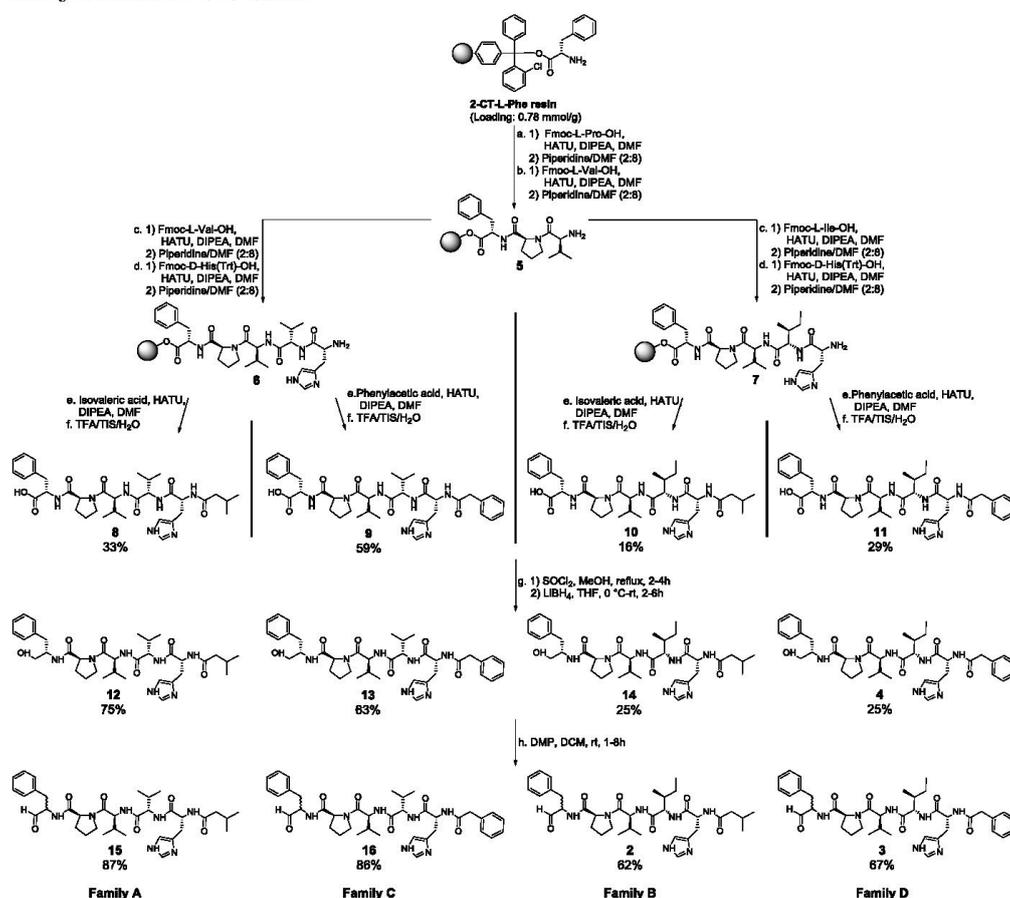
Total Synthesis of Pentapeptide Falcitidin Analogues. The scarcity of natural material and difficulties in isolation due to the co-elution of aldehydes and alcohol analogues made it necessary to synthesize the most promising pentapeptide phenylalanine-aldehydes for biological testing. The application of a split-approach solid-phase peptide synthesis (SPPS)²⁸ followed by functional group interconver-

sion gave access to a variety of falcitidin pentapeptide analogues of families A, B, C, and D (Scheme 1).

Splitting the synthesis after the third amino acid and for the attachment of the two different fatty acids allows the synthesis of all four chosen acid analogues (2, 3, 15, and 16) based on 2-chlorotryl-L-Phe resin (2-CT-L-Phe). The functionalization of the peptide acid yields the corresponding methyl ester, the alcohol, and the aldehyde. Reduction to the alcohol was done via the methyl ester after direct reduction of the acid could not be achieved. The methyl ester was gained by conversion with thionyl chloride in methanol followed by the reduction with LiBH_4 in THF. For the oxidation to the aldehyde, Dess–Martin periodinane (DMP) was chosen as a very mild reagent. We found that small traces of methanol as a stabilizer in DCM as the solvent made the reaction take longer, up to 8 h instead of 1 h, and was responsible for incomplete conversion. Epimerization of the stereogenic center of the aldehydes could not be avoided, nor was it necessary to be avoided based on the data for the natural isolated compounds. They were obtained as diastereomers for the phenylalanine moiety of roughly equal proportions based on NMR data (Figures S20–S23 and Tables S12–S15). Overall yields of SPPS for the four different chains varied between 16 and 59%. The reduction yielded the corresponding alcohols over two steps with 25–75% yield, and oxidation to the aldehydes achieved 62–87% yield (Scheme 1).

For reference purposes, falcitidin (1) was synthesized using the same SPPS approach instead of the literature-known liquid phase method.¹¹ This afforded acid analogue 17 with a yield of 78% and, after amidation, falcitidin (1) with a 41% overall yield (Scheme 2). Additionally, the acid analogue 17 and falcitidin (1) themselves present great functional groups for further derivatization and the introduction of different war heads, like a nitrile or azide group, to increase the potency as a potential inhibitor.^{29,30}

Bioactivity. Falcitidin (1) was previously reported to display an IC_{50} of 6 μM against falcipain-2.¹⁰ Therefore, the falcitidin pentapeptide aldehyde analogues 2, 3, 15, and 16 were tested in a similar in vitro assay against falcipain-2 together with 1 as control. All four aldehydes were active with an IC_{50} of 41.5 μM (15) to 23.7 μM (2). However, the originally described activity of 1 could not be reproduced, with an IC_{50} of >50 μM . This indicated differences in assay

Scheme 1. SPPS Split Approach (a–f) and Functional Group Interconversions (g–h) to Chosen Falcitidin Pentapeptide Analogues of Families A, B, C, and D^a

^aStandard SPPS conditions apply, and reactions were carried out at room temperature if not noted otherwise. Reduction of the acid to the alcohol was achieved via the corresponding methyl ester. The aldehyde was obtained by Dess–Martin oxidation; the stereogenic center of phenylalanine could not be retained. Detailed conditions can be found in the Supporting Information.

sensitivity, which made it difficult to put the aldehyde activities in line with the literature. Aldehydes 3 and 15 also displayed activities against the related *P. falciparum* cysteine protease falcipain-3 (66% sequence identity with falcipain-2), with IC₅₀ values of 45.4 and 42.5 μM, respectively. Inhibition of both falcipains is of importance, since falcipain-3 is able to compensate for the knockout of falcipain-2.³¹ A counter-screen against human cysteine proteases cathepsins B and L and sortase A of *Staphylococcus aureus* as a surrogate for unrelated cysteine proteases revealed no activity for pentapeptide aldehydes 2, 3, 15, and 16 as well as the C-terminal acid (11) and alcohol (4) of 3, thus demonstrating selectivity over these off-targets. Activities with IC₅₀ values of 57.7 μM (2) to 17.1 μM (15) were also observed for the falcipain-homologue cysteine protease of *Trypanosoma brucei rhodesiense* rhodesain.

Most interestingly, aldehydes 3, 15, and 16 displayed higher activities, with IC₅₀s of 3.7 μM (15) and 1.5 μM (16), against α-chymotrypsin, which was tested as a prototype of human serine proteases, while the serine membrane protease matrilysin-2 (TMPS6) was not inhibited (Table 4). To further elucidate these observations, molecular docking studies were performed.

Docking Studies. Due to the high number of rotatable bonds and the associated degrees of conformational freedom in the molecules under elucidation, docking of the full-length peptides is challenging.^{32–34} Hence, truncated tripeptides with an N-terminal acetyl (ace)-cap to avoid a nonpresent charge of full-length inhibitors (2, 3, 4, 11, 15, and 16) were used for docking studies. First, a conventional noncovalent docking was performed. However, aldehydes are known electrophilic

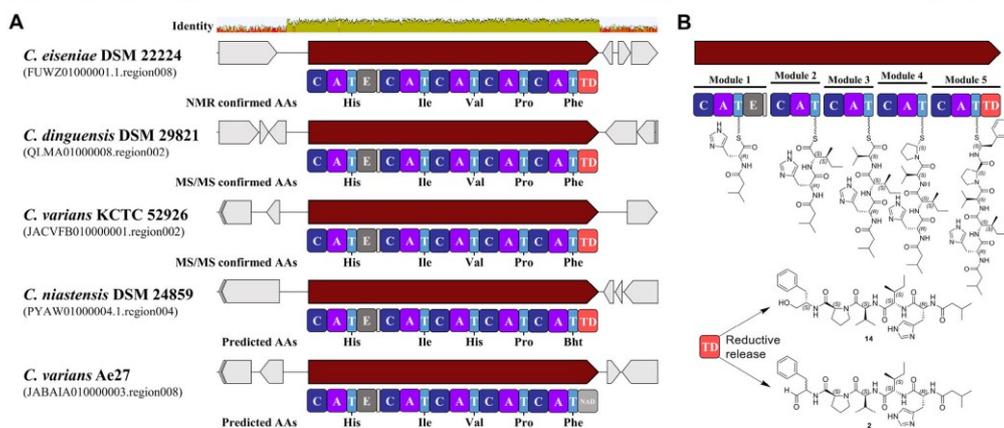


Figure 3. Putative biosynthetic gene clusters for pentacitidin peptides found in various *Chitinophaga* strains. (a) Nucleotide alignment of all five BGCs using MAFFT alignment⁴⁷ allowed clear cluster boarder prediction and revealed minor variations in the upstream region (identity). The single NRPS core gene that is responsible for pentacitidin biosynthesis is shown in dark red, and neighboring genes that do not belong to the pentacitidin BGC are shown in gray. (b) Biosynthetic hypothesis for pentacitidins. AA, amino acids; C, condensation domain; A, adenylation domain; T, peptidyl-carrier protein domain; E, epimerization domain; Bht, beta-hydroxy-tyrosine; TD, terminal reductase domain; NAD₂ binding₄ domain.

Scheme 2. Synthesis of Falcitidin (1) via Acid Analogue 17

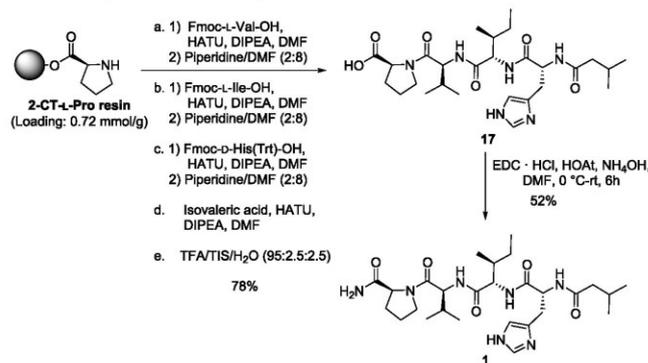


Table 4. Overview of Protease Activity Data^a

no.	peptide sequence	IC ₅₀ (μM)							
		Rhod	falcipain-2	falcipain-3	CatB	CatL	α -CT	SrtA	TMPRSS6
1	iVal H I V P - CONH ₂	-	>50	>50	-	-	-	-	-
17	iVal H I V P - COOH	-	n.d.	n.d.	-	-	-	-	-
11	PA H I V P F COOH	-	n.d.	n.d.	-	-	-	-	-
4	PA H I V P F CH ₂ OH	-	n.d.	n.d.	-	-	-	-	-
2	iVal H I V P F CHO	57.7	33.8	45.4	-	-	-	-	-
15	iVal H V V P F CHO	17.1	42.5	-	-	-	2.2	-	-
3	PA H I V P F CHO	25.4	23.7	42.5	-	-	3.7	-	-
16	PA H V V P F CHO	-	38.2	-	-	-	1.5	-	-

^an.d., not determined; hyphens denote that the compounds were not active.

warheads and able to form covalent-reversible hemithioacetal adducts with catalytic cysteine residues.³⁵ Therefore, the predicted poses with special attention to the distance between

the nucleophilic sulfur of the catalytic cysteine residue and the electrophilic carbon of the aldehyde moiety were elucidated, and an additional covalent docking was performed. The

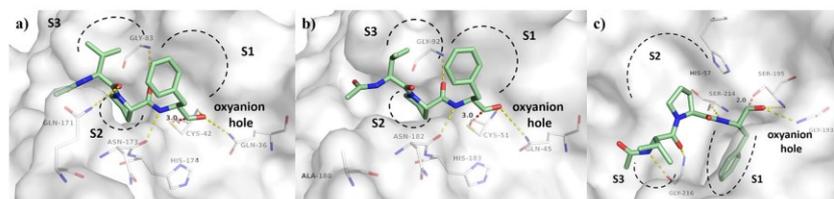


Figure 4. Noncovalent docking predicted binding modes of truncated peptidic protease inhibitor ace-Val-Pro-Phe-aldehyde (green carbon atoms) in complex with falcipain-2 (PDB-ID 3BPF) (a), falcipain-3 (PDB-ID 3BPM) (b), and chymotrypsin (PDB-ID 1AFQ) (c). Proteases are depicted as white transparent surfaces; for clarity, only residues forming polar interactions (yellow dashed lines) and catalytic Cys/Ser and His residues are labeled and depicted as lines. Substrate binding sites S1–S3 are schematically indicated. The distance between nucleophilic sulfur (Cys-42 in falcipain-2 and Cys-51 in falcipain-3) or oxygen (Ser-195 in chymotrypsin) of the catalytic center to the electrophilic carbon atom of the aldehyde is depicted as a dashed red line and labeled with its distance in Å.

tripeptides mimic the orientation of the protease substrates by addressing S3–S1 (Figure 4a,b). Based on these results (Table S16), the low inhibitory potency of the C-terminal alcohol (4) and acid (11) moieties as well as that for falcitidin cannot simply be explained by noncovalent interactions as the docking scores are within a similar range or even higher when compared to the aldehydes. However, based on the predicted binding poses, the covalent reaction between aldehydes and catalytic cysteine residues, which might contribute to higher affinity, seems likely for molecules with proteinogenic L-Phe as P1 residue, with a distance of 3.0 Å between the electrophilic carbon atom and the nucleophilic sulfur of Cys for both falcipain-2 and falcipain-3 as well as the aldehyde oxygen coordinated by hydrogen bonds in the oxyanion hole, but rather unlikely for D-Phe (distances of 7.3 and 5.0 Å for falcipain-2 and falcipain-3, respectively). Additionally, only small differences between covalent and noncovalent binding modes were observed, indicating that no larger conformational changes need to take place during or after reaction. While testing against α -chymotrypsin was conducted to demonstrate selectivity over human off-target serine proteases, the high affinity is reasonable, as inhibitors with aromatic moieties deeply buried in the S1 pocket of the protease were reported previously as well as cleavage preference after Phe.^{36–38} Aldehydes are able to form covalent-reversible hemiacetal adducts with serine;³⁹ moreover, an increased potency is indicated by proximity of the aldehyde to the catalytic Ser-195 residue in the predicted binding modes for both L- and D-Phe at P1 (Figure 4c and Table S16). In conclusion, the higher docking scores for the covalent binding compared to the ones for the noncovalent ones indicate that the literature-known covalent-reversible binding mode in form of the hemithioacetal/hemiacetal adducts takes place for the cysteine/serine residue and the aldehyde moiety of the pentacitidins. Binding of L-Phe seems to be more likely based on the docking results (distance) than D-Phe.

Falcitidin (1) was found to display poor whole-cell activity against chloroquine-sensitive *P. falciparum* strain 3D7 ($IC_{50} > 10 \mu M$), while a structure–activity relationship (SAR) study identified a synthetic trifluoromethyl analogue displaying sub-micromolar IC_{50} activity.¹¹ Together with our results, this allows further development of the SAR, with the aim to increase both potency and selectivity with an improved peptidic recognition sequence. However, no general selectivity issue against serine proteases can be expected, as matrilysin-2 was not inhibited by the compounds. Additionally, the docking

studies revealed a rather unlikely distance for the covalent binding of D-Phe-analogues. With all compounds having been tested as mixtures of D- and L-Phe due to the aldehyde's natural epimerization of its stereogenic center,^{40,41} further analogues with other warheads preventing epimerization such as a nitrile group,^{29,30} for example, should ideally be synthesized and tested in L-configurations only.

CONCLUSIONS

In this study, the metabolic networking analysis of *Chitinophaga* strains led to the discovery of over 30 N-acyl oligopeptides structurally related to falcitidin, an inhibitor of the antimalarial cysteine protease falcipain-2.¹⁰ They were classified into four peptide families (A–D) based on variations in their amino acid chain length, sequence, acyl residue, and C-terminal functionalization. The analysis of MS/MS spectra revealed each family to contain truncated di- or tripeptides and larger pentapeptides, including molecules with classical C-terminal aldehyde moieties, which are supposed to covalently react with the active-site cysteine and serine residues of proteases.⁹ Isolation and structure elucidation of two novel natural pentapeptide aldehydes validated the MS/MS fragmentation pattern analysis. A BGC congruent to the pentapeptide structure was identified *in silico*. Together with the pentacitidin NMR studies, this indicated falcitidin and its tetrapeptide analogues carrying the C-terminally amidated proline to be degradation products from the described pentacitidins. Total synthesis of all four pentapeptide aldehydes and their carboxylic acid and alcohol derivatives was successfully achieved using a solid-phase peptide synthesis (SPPS) split approach followed by functional C-terminal group interconversion and gave access to the most promising aldehyde analogues, allowing their biological profiling. Selective *in vitro* activity against parasitic cysteine proteases rhodesain, falcipain-2, and falcipain-3, together with a low-micromolar IC_{50} inhibition of the serine protease α -chymotrypsin, was observed. This forms the basis for future studies to develop optimized derivatives with increased potency and selectivity against targeted proteases as well as to elucidate the occurrence of falcitidin analogues carrying the unusual C-terminal amidated proline.

EXPERIMENTAL SECTION

General Experimental Procedures. For all UHPLC-QTOF-UHR-MS and MS/MS measurements, a quadrupole time-of-flight spectrometer (LC-QTOF maXis II, Bruker Daltonics, Bremen, Germany) equipped with an electrospray ionization source in line

with an Agilent 1290 infinity II LC system (Agilent Technologies, CA, United States) was used. C18 RP-UHPLC (ACQUITY UPLC BEH C18 column (130 Å, 1.7 μm, 2.1 × 100 mm)) was performed at 45 °C with the following linear gradient (A: H₂O, 0.1% HCOOH; B: CH₃CN, 0.1% HCOOH; flow rate: 0.6 mL min⁻¹): 0 min: 95% A; 0.30 min: 95% A; 18.00 min: 4.75% A; 18.10 min: 0% A; 22.50 min: 0% A; 22.60 min: 95% A; and 25.00 min: 95% A. A 50–2000 *m/z* scan range at a 1 Hz scan rate was used to acquire mass spectral data. The injection volume was set to 5 μL. MS/MS experiments were performed at 6 Hz, and the top five most intense ions in each full MS spectrum were targeted for fragmentation by higher-energy collisional dissociation at 25 eV using N₂ at 10⁻² mbar. Precursors were excluded after two spectra, released after 0.5 min, and reconsidered if the intensity of an excluded precursor increased by a factor of 1.5 or more. Data were analyzed using the Bruker DataAnalysis 4.0 software package. Specific rotation was determined on a digital polarimeter (P3000, A. Krüss Optronic GmbH, Germany). The standard wavelength was the sodium D-line with 589 nm. Temperature, concentration (g 100 mL⁻¹), and solvent are reported with the determined value.

NMR Spectroscopy. The NMR spectra of natural isolated pentacitidins were acquired on a Bruker AVANCE 700 spectrometer (700 MHz for ¹H and 176 MHz for ¹³C) and a Bruker AVANCE 500 spectrometer (500 MHz for ¹H and 126 MHz for ¹³C). Both instruments were equipped with a 5 mm TCI cryoprobe. For structure elucidation and the assignment of proton and carbon resonances, 1D-¹H, 1D-¹³C, DQF-COSY, TOCSY (mixing time 80 ms), ROESY (mixing time 150 ms), multiplicity-edited HSQC, and HMBC spectra were acquired. The NMR spectra of synthesized molecules were recorded on an AVANCE III HD 600 spectrometer (600 MHz for ¹H, 151 MHz for ¹³C) from Bruker Biospin (Bruker Biospin GmbH, Rheinstetten, Germany). ¹H and ¹³C chemical shifts were reported in ppm and referenced to the corresponding residual solvent signal (DMSO-*d*₆: δ_C = 39.52 ppm, δ_H = 2.50 ppm). δ_C shifts marked with asterisks were not observed in the ¹³C NMR spectrum but were obtained either from HMBC or HSQC data.

MS/MS Networking. Molecular networking was performed following established protocols.^{15,42} In brief, parent ions are represented by a list of fragment mass/intensity value pairs within the raw data (*.d files) converted with MScConvert (ProteoWizard package32) into plain text files (*.mgf). These ions are included in the final network once they share at least six fragments (tolerance Δppm 0.05) with at least one partner ion.⁴³ Deposited compounds from an in silico fragmented⁴⁴ commercial database (Antibase 2017⁴⁵) as well as our in-house reference compound MS/MS database were included in the final network to highlight known NPs. A visualization of the network was constructed in Cytoscape v3.6.0.⁴⁶ Edges were drawn between scan nodes with a cosine similarity of >0.7.⁴⁷

Strain Fermentation and Purification of Falcitidin Analogues. A preculture (R2A, 100 mL in a 300 mL Erlenmeyer flask) of *C. eiseniae* DSM 22224 was inoculated from a plate (R2A) and incubated at 28 °C with agitation at 180 rpm for 3 days. A 20 L fermentation in medium 3018 (1 g L⁻¹ yeast extract and 5 g L⁻¹ casitone, pH 7.0) inoculated with 2% (v/v) preculture was carried out in separate 2 L flasks filled with 500 mL of culture volume at 28 °C with agitation at 180 rpm for 4 days. The culture broth was subsequently freeze-dried using a Delta 2-24 LSCplus (Martin Christ Gefrier Trocknungsanlagen GmbH, Osterode am Harz, Germany). The sample was extracted with a one-time culture volume of CH₃OH, evaporated to dryness using rotary evaporation under reduced pressure, and resuspended in 3 L of 10% CH₃OH/H₂O. The extract was loaded onto a XAD16N column (1 L bed volume) and eluted step-wise with 10, 40, 60, 80, and 100% CH₃OH (2 times bed volume each). The 80 and 100% fractions containing falcitidin analogues were pooled, and the sample was adjusted to 200 mg mL⁻¹ in methanol to achieve further separation using preparative C18-RP-HPLC (Synergi 4 μm Fusion-RP 80 Å (250 × 21.2 mm), Phenomenex Inc.) by eluting in a linear gradient increasing from 25 to 75% CH₃CN (+0.1% HCOOH) in 22 min. Fractions of interest were concentrated to 100 mg mL⁻¹ for semipreparative C18-RP-HPLC (Synergi 4 μm Fusion-

RP 80 Å (250 × 10 mm), Phenomenex Inc.) using a linear gradient from 15 to 50% CH₃CN (+0.1% HCOOH) in 29 min. Final purification of the samples of interest (100 mg mL⁻¹) was achieved using UHPLC on a ACQUITY UPLC BEH C18 column (130 Å, 1.7 μm, 100 × 2.1 mm, Waters Corporation), eluting in an isocratic gradient of 27.50% CH₃CN (+0.1% HCOOH) in 18 min. In total, isolation yielded 1.5 mg of 2, 1 mg of 3, and 2 mg of a mixture of 3 and 4.

Pentacitidin A (2). Amorphous, white powder; see Table 2 for ¹H and ¹³C NMR data; positive HR-ESIMS *m/z* 680.4131 [M+H]⁺, calculated mass for C₃₆H₅₄N₇O₆: 680.4130; Δ = 0.15 ppm.

Pentacitidin B (3). Amorphous, white powder; see Table 3 for ¹H and ¹³C NMR data; positive HR-ESIMS *m/z* 714.3975 [M+H]⁺, calculated mass for C₃₉H₅₂N₇O₆: 714.3974; Δ = 0.14 ppm.

Advanced Marfey's Analysis. The absolute configurations of all amino acids were determined by derivatization using Marfey's reagent.¹⁵ Stock solutions of amino acid standards (50 mM in H₂O), NaHCO₃ (1 M in H₂O), and N_α-(2,4-dinitro-5-fluorophenyl)-L-valinamide (L-FDVA, 70 mM in acetone; Sigma-Aldrich, St. Louis, MO, United States) were prepared. Commercially available and synthesized standards were derivatized using molar ratios of amino acid to L-FDVA and NaHCO₃ (1/1.4/8). After stirring at 40 °C for 3 h, 1 M HCl was added to obtain a concentration of 170 mM to end the reaction. Samples were subsequently evaporated to dryness and dissolved in DMSO (final concentration 50 mM). L- and D-amino acids were analyzed separately using C18 RP-UHPLC-MS (A: H₂O, 0.1% HCOOH; B: CH₃CN, 0.1% HCOOH; flow rate: 0.6 mL min⁻¹). A linear gradient of 15–75% B in 35 min was applied to separate all amino acid standards. Total hydrolysis of the peptide sample containing 3 and 4 was carried out by dissolving 250 μg in 6 M DCl in D₂O and stirring for 7 h at 160 °C. The sample was subsequently evaporated to dryness. Samples were dissolved in 100 μL H₂O, derivatized with L-FDVA, and analyzed using the same parameters as described before.

Fluorometric Assays. Rhodocin (Rhod),^{48–50} *Staphylococcus aureus* sortase A (SrtA),⁵¹ and human matriptase-2 (TMPRSS6)⁵² were expressed and purified as published previously, cathepsins B and L (CatB and CatL; human liver, Calbiochem) and α-chymotrypsin (α-CT; Sigma-Aldrich, St. Louis, MO, United States) were purchased. For these proteases except SrtA, fluorescence increase upon cleavage of the fluorogenic substrates was monitored without incubation with a TECAN Infinite F200 Pro fluorimeter (excitation λ = 365 nm; emission λ = 460 nm) in white, flat-bottom 96-well microtiter plates (Greiner Bio-One, Kremsmünster, Austria) with a total volume of 200 μL. Inhibitors and substrates were prepared as stock solutions in DMSO to a final DMSO content of 0.5%. Inhibitors were screened at final concentrations of 20 μM and eventually at 1 μM. IC₅₀ values were determined for compounds for which an inhibition of >50% at a concentration of 20 μM was observed. All assays were performed in technical triplicates and normalized to the activity of DMSO instead of the inhibitors by measuring the increase of fluorescence signal over 10 min. The data were analyzed using GraFit V 5.0.13⁵³ (Erithacus Software, Horley, UK; <http://erithacus.com/grafit/>).

Cbz-Phe-Arg-AMC (Bachem, Bubendorf BL, Switzerland) was used as a substrate for Rhod, CatB, and CatL. The enzymes were incubated at room temperature in enzyme incubation buffer (Rhod: 50 mM sodium acetate (pH 5.5), 5 mM EDTA, 200 mM NaCl, and 2 mM DTT; CatB/L: 50 mM Tris-HCl (pH 6.5), 5 mM EDTA, 200 mM NaCl, and 2 mM DTT) for 30 min. One hundred and eighty microliters of assay buffer (Rhod: 50 mM sodium acetate (pH 5.5), 5 mM EDTA, 200 mM NaCl, and 0.005% Brij35; CatB/L: 50 mM Tris-HCl (pH 6.5), 5 mM EDTA, 200 mM NaCl, and 0.005% Brij35) were added to the 96-well plates; afterward, the respective enzyme in enzyme incubation buffer (5 μL; to yield final concentrations for Rhod, 0.01 μM; CatB, 0.1 μM; and CatL, 0.2 μM) was added followed by 10 μL of DMSO (control) or inhibitor solution in DMSO and, finally, the substrate (5 μL; final concentrations: Rhod, 10 μM; CatB, 100 μM; and CatL, 6.5 μM) was added.⁵⁴

Transpeptidation efficacy of SrtA was examined in vitro as described previously.⁵¹ Briefly, SrtA was diluted in assay buffer (50 mM Tris-HCl (pH 7.50) and 150 mM NaCl) to a final concentration of 1 μ M. The FRET-pair substrate Abz-LPETG-Dap(Dnp)-OH (Genscript, Piscataway, NJ, United States) and the tetraglycine (Sigma-Aldrich, St. Louis, MO, United States) were added at 25 μ M and 0.5 mM, respectively.

α -Chymotrypsin (final concentration, 0.4 μ M) was dissolved in assay buffer containing 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 5 mM EDTA. Suc-Leu-Leu-Val-Tyr-AMC (Bachem, Bubendorf BL, Switzerland) was used as a substrate at a final concentration of 52.5 μ M.⁵⁵

Proteolytic activity of matriptase-2 was measured in a final concentration of 2.5 nM enzyme in 180 μ L reaction buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM CaCl₂, and 0.01% (v/v) T_x-100). After addition of the inhibitors, the reaction was initiated without further incubation by adding the substrate Boc-Leu-Arg-Arg-AMC (Bachem, Bubendorf BL, Switzerland, $K_M = 36.1 \pm 5.8 \mu$ M) to a final concentration of 100 μ M.

The recombinant enzymes falcipain-2 and falcipain-3 were expressed and purified as previously described.⁵⁶ Stock solutions of the compounds, the substrate, and the positive control E-64 (Sigma-Aldrich, St. Louis, MO, United States) were prepared at 10 mM in DMSO. The compounds were incubated in 96-well white flat-bottom plates with 30 nM of recombinant falcipain-2 or falcipain-3 at room temperature in assay buffer (100 mM sodium acetate (pH 5.5) with 5 mM DTT) for 10 min. After incubation, the fluorogenic substrate Z-Leu-Arg-AMC (R&D Systems, Minneapolis, MN, United States) was added at a concentration of 25 μ M in a final assay volume of 200 μ L. Fluorescence was monitored with a Varioskan Flash (Thermo Fisher Scientific Inc., Waltham, MA, United States) with excitation at 355 nm and emission at 460 nm. The IC₅₀ was calculated using GraphPad Prism (GraphPad Software, San Diego, CA, United States) based on a sigmoidal dose-response curve.

Molecular Docking. Molecular docking was performed against falcipain-2 (complex with E-64, PDB-ID 3BPF),⁵⁷ falcipain-3 (complex with leupeptin, PDB-ID 3BPM),⁵⁷ and chymotrypsin (complex with D-leucyl-L-phenylalanyl-p-fluorobenzylamide, PDB-ID 1AFQ).⁵⁸ Conventional noncovalent template-based docking was performed with HYBRID v3.3.0.3 (OpenEye Scientific Software, Santa Fe, NM, US; <http://eyesopen.com>).^{58,59} The receptor was prepared using the make_receptor tool version 3.3.0.3 under default settings for potential field generation around the reference ligand for 3BPM and 1AFQ. As the complexed ligand E-64 of 3BPF does not reach toward S1 of falcipain-2, the potential field was generated using leupeptin from the aligned structure of 3BPM (falcipain-3) for which a similar binding behavior for falcipain-2 and falcipain-3 can be expected.⁶⁰ Eight hundred ligand conformers per molecule for docking were generated using Omega Pose (OMEGA v3.1.0.3, OpenEye Scientific Software, Santa Fe, NM, US; <http://eyesopen.com>).⁶¹ Covalent docking was performed using MOE 2020.09.⁶² Using the covalent reaction of the acetalization from aldehyde and Cys/Ser (for E-64 redocking from epoxide to beta-hydroxythioether), a rigid docking with GB/VI scoring was applied for the initial placement of 50 poses, from which the 10 best-scoring ones were refined using the ASE scoring function. Docking setups were validated by the redocking of crystallographic reference ligands by pose inspection and RMSD calculation (Table S16).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscchembio.1c00861>.

MS/MS-based assignment of peptide sequences; Marfey's analysis; detailed description of all syntheses; ¹H and ¹³C spectra and data of all synthesized compounds and the natural isolated ones; docking results (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Till F. Schäberle – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany; Institute for Insect Biotechnology, Justus-Liebig-University of Giessen, Giessen 35392, Germany; orcid.org/0000-0001-9947-8079; Email: Armin.Bauer@sanofi.com*

Armin Bauer – *Sanofi-Aventis Deutschland GmbH, R&D, Frankfurt am Main 65926, Germany; orcid.org/0000-0002-2171-1333; Phone: +49-641-97-219140; Email: Till.Schaerberle@ime.fraunhofer.de*

Authors

Stephan Brinkmann – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany; orcid.org/0000-0003-4299-1079*

Sandra Semmler – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany*

Christian Kersten – *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany; orcid.org/0000-0001-9976-7639*

Maria A. Patras – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany*

Michael Kurz – *Sanofi-Aventis Deutschland GmbH, R&D, Frankfurt am Main 65926, Germany*

Natalie Fuchs – *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany; orcid.org/0000-0001-6404-676X*

Stefan J. Hammerschmidt – *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany*

Jenny Legac – *Department of Medicine, University of California, San Francisco, California 94143, United States*

Peter E. Hammann – *Evotec International GmbH, Göttingen 37079, Germany*

Andreas Vilcinskas – *Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, Giessen 35392, Germany; Institute for Insect Biotechnology, Justus-Liebig-University of Giessen, Giessen 35392, Germany*

Philip J. Rosenthal – *Department of Medicine, University of California, San Francisco, California 94143, United States*

Tanja Schirmeister – *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany*

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acscchembio.1c00861>

Author Contributions

[†]S.B. and S.S. contributed equally to this work.

Funding

This work was financially supported by the Hessen State Ministry of Higher Education, Research, and the Arts (HMWK) via the state initiative for the development of scientific and economic excellence for the LOEWE Center for Insect Biotechnology and Bioresources. Sanofi-Aventis Deutschland GmbH and Evotec International GmbH contributed in the framework of the Sanofi-Fraunhofer Natural

Products Center of Excellence/Fraunhofer-Evotec Natural Products Center of Excellence.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank J. Glaeser, S. Schuler, Y. Kleiner, and C. Pöverlein for valuable discussions. We thank C. Hartwig, V. Öhler, and the NMR department of the Justus-Liebig-University Giessen for technical assistance. We further thank OpenEye Scientific for free academic software licenses and T. Steinmetzer and coworkers from the Philipps University of Marburg for the kind gift of matriptase-2 coding plasmid for recombinant protein expression.

REFERENCES

- (1) Ashley, E. A.; Pyae Phy, A.; Woodrow, C. J. *Malaria*. *Lancet* **2018**, *391*, 1608–1621.
- (2) Büscher, P.; Cecchi, G.; Jamonneau, V.; Priotto, G. Human African trypanosomiasis. *Lancet* **2017**, *390*, 2397–2409.
- (3) World Health Organization *World malaria report 2020: 20 years of global progress and challenges*; World Health Organization 2020.
- (4) Haldar, K.; Bhattacharjee, S.; Safeukui, I. Drug resistance in Plasmodium. *Nat. Rev. Microbiol.* **2018**, *16*, 156–170.
- (5) Conrad, M. D.; Rosenthal, P. J. Antimalarial drug resistance in Africa: the calm before the storm? *Lancet Infect. Dis.* **2019**, *19*, e338–e351.
- (6) Otto, H.-H.; Schirmeister, T. Cysteine Proteases and Their Inhibitors. *Chem. Rev.* **1997**, *97*, 133–172.
- (7) Ettari, R.; Previti, S.; Tamborini, L.; Cullia, G.; Grasso, S.; Zappalà, M. The Inhibition of Cysteine Proteases Rhodesain and TbCatB: A Valuable Approach to Treat Human African Trypanosomiasis. *Mini-Rev. Med. Chem.* **2016**, *16*, 1374–1391.
- (8) Rosenthal, P. J. Falcipain cysteine proteases of malaria parasites: An update. *BBA-Proteins Proteom.* **2020**, *1868*, 140362.
- (9) Drag, M.; Salvesen, G. S. Emerging principles in protease-based drug discovery. *Nat. Rev. Drug Discov.* **2010**, *9*, 690–701.
- (10) Somanadhan, B.; Kotturi, S. R.; Yan Leong, C.; Glover, R. P.; Huang, Y.; Flotow, H.; Buss, A. D.; Lear, M. J.; Butler, M. S. Isolation and synthesis of falcitidin, a novel myxobacterial-derived acyltetrapeptide with activity against the malaria target falcipain-2. *J. Antibiot.* **2013**, *66*, 259–264.
- (11) Kotturi, S. R.; Somanadhan, B.; Ch'ng, J.-H.; Tan, K. S.-W.; Butler, M. S.; Lear, M. J. Diverted total synthesis of falcitidin acyl tetrapeptides as new antimalarial leads. *Tetrahedron Lett.* **2014**, *55*, 1949–1951.
- (12) Krug, D.; Müller, R. Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. *Nat. Prod. Rep.* **2014**, *31*, 768–783.
- (13) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; de Felicio, R.; Fenner, A.; et al. Molecular networking as a dereplication strategy. *J. Nat. Prod.* **2013**, *76*, 1686–1699.
- (14) Brinkmann, S.; Kurz, M.; Patras, M. A.; Hartwig, C.; Marner, M.; Leis, B.; Billion, A.; Kleiner, Y.; Bauer, A.; Toti, L.; Pöverlein, C.; Hammann, P. E.; Vilcinskas, A.; Glaeser, J.; Spohn, M. S.; Schäberle, T. F. Genomic and chemical decryption of the Bacteroidetes phylum for its potential to biosynthesize natural products. *bioRxiv* **2021**, DOI: 10.1101/2021.07.30.454449.
- (15) Bhushan, R.; Brückner, H. Marfey's reagent for chiral amino acid analysis: a review. *Amino Acids* **2004**, *27*, 231–247.
- (16) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, *47*, W81–W87.
- (17) Katoh, K.; Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780.
- (18) Du, L.; Lou, L. PKS and NRPS release mechanisms. *Nat. Prod. Rep.* **2010**, *27*, 255–278.
- (19) Gaitatzis, N.; Kunze, B.; Müller, R. In vitro reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: Biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 11136–11141.
- (20) Silakowski, B.; Nordsiek, G.; Kunze, B.; Blöcker, H.; Müller, R. Novel features in a combined polyketide synthase/non-ribosomal peptide synthetase: the myxalamid biosynthetic gene cluster of the myxobacterium *Stigmatella aurantiaca* Sga1511. *Chem. Biol.* **2001**, *8*, 59–69.
- (21) Li, Y.; Weissman, K. J.; Müller, R. Myxochelin Biosynthesis: Direct Evidence for Two- and Four-Electron Reduction of a Carrier Protein-Bound Thioester. *J. Am. Chem. Soc.* **2008**, *130*, 7554–7555.
- (22) Chhabra, A.; Haque, A. S.; Pal, R. K.; Goyal, A.; Rai, R.; Joshi, S.; Panjkar, S.; Pasha, S.; Sankaranarayanan, R.; Gokhale, R. S. Nonprocessive 2 + 2e⁻ off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 5681–5686.
- (23) Barajas, J. F.; Phelan, R. M.; Schaub, A. J.; Klierer, J. T.; Kelly, P. J.; Jackson, D. R.; Luo, R.; Keasling, J. D.; Tsai, S.-C. Comprehensive Structural and Biochemical Analysis of the Terminal Myxalamid Reductase Domain for the Engineered Production of Primary Alcohols. *Chem. Biol.* **2015**, *22*, 1018–1029.
- (24) Aarts, M. G. M.; Hodge, R.; Kalantidis, K.; Florack, D.; Wilson, Z. A.; Mulligan, B. J.; Stiekema, W. J.; Scott, R.; Pereira, A. The Arabidopsis MALE STERILITY 2 protein shares similarity with reductases in elongation/condensation complexes. *Plant. J.* **1997**, *12*, 615–623.
- (25) Metz, J. I. M.; Lassner, M. Reprogramming of Oil Synthesis in Rapeseed: Industrial Applications. *Ann. N. Y. Acad. Sci.* **1996**, *792*, 82–90.
- (26) Tan, S.; Li, F.; Park, S.; Kim, S. Transition metal-free aerobic oxidative cleavage of the C–N bonds of α -amino esters. *Org. Chem. Front.* **2019**, *6*, 3854–3858.
- (27) Doering, W. E.; Chanley, J. D. The autoxidation of quinone. *J. Am. Chem. Soc.* **1946**, *68*, 586–588.
- (28) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
- (29) Serim, S.; Haedke, U.; Verhelst, S. H. L. Activity-based probes for the study of proteases: recent advances and developments. *ChemMedChem* **2012**, *7*, 1146–1159.
- (30) Gehringer, M.; Laufer, S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2019**, *62*, 5673–5724.
- (31) Sijwali, P. S.; Rosenthal, P. J. Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 4384–4389.
- (32) Andrusier, N.; Mashiah, E.; Nussinov, R.; Wolfson, H. J. Principles of flexible protein-protein docking. *Proteins* **2008**, *73*, 271–289.
- (33) Pons, C.; Grosdidier, S.; Solernou, A.; Pérez-Cano, L.; Fernández-Recio, J. Present and future challenges and limitations in protein-protein docking. *Proteins* **2010**, *78*, 95–108.
- (34) Knaff, P. M.; Kersten, C.; Willbold, R.; Champanhac, C.; Crespy, D.; Wittig, R.; Landfester, K.; Mailänder, V. From In Silico to Experimental Validation: Tailoring Peptide Substrates for a Serine Protease. *Biomacromolecules* **2020**, *21*, 1636–1643.
- (35) Silva, D. G.; Ribeiro, J. F. R.; de Vita, D.; Cianni, L.; Franco, C. H.; Freitas-Junior, L. H.; Moraes, C. B.; Rocha, J. R.; Burtoloso, A. C. B.; Kenny, P. W.; et al. A comparative study of warheads for design of cysteine protease inhibitors. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 5031–5035.

- (36) Kashima, A.; Inoue, Y.; Sugio, S.; Maeda, I.; Nose, T.; Shimohigashi, Y. X-ray crystal structure of a dipeptide-chymotrypsin complex in an inhibitory interaction. *Eur. J. Biochem.* **1998**, *255*, 12–23.
- (37) Stoll, V. S.; Eger, B. T.; Hynes, R. C.; Martichonok, V.; Jones, J. B.; Pai, E. F. Differences in binding modes of enantiomers of 1-acetamido boronic acid based protease inhibitors: crystal structures of gamma-chymotrypsin and subtilisin Carlsberg complexes. *Biochemistry* **1998**, *37*, 451–462.
- (38) Rawlings, N. D.; Waller, M.; Barrett, A. J.; Bateman, A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **2014**, *42*, D503–D509.
- (39) Cho, S.; Dickey, S. W.; Urban, S. Crystal Structures and Inhibition Kinetics Reveal a Two-Stage Catalytic Mechanism with Drug Design Implications for Rhomboid Proteolysis. *Mol. Cell* **2016**, *61*, 329–340.
- (40) Lyons, B.; Kwan, A. H.; Jamie, J.; Truscott, R. J. W. Age-dependent modification of proteins: N-terminal racemization. *FEBS J.* **2013**, *280*, 1980–1990.
- (41) Kajita, R.; Goto, T.; Lee, S. H.; Oe, T. Aldehyde stress-mediated novel modification of proteins: epimerization of the N-terminal amino acid. *Chem. Res. Toxicol.* **2013**, *26*, 1926–1936.
- (42) Allard, P.-M.; Péresse, T.; Bisson, J.; Gindro, K.; Marcourt, L.; van Pham, C.; Roussi, F.; Litaudon, M.; Wolfender, J.-L. Integration of Molecular Networking and In-Silico MS/MS Fragmentation for Natural Products Dereplication. *Anal. Chem.* **2016**, *88*, 3317–3323.
- (43) Riyanti; Marner, M.; Hartwig, C.; Patras, M.; Wodi, S.; Rieuwpassa, F.; Ijong, F.; Balansa, W.; Schäberle, T. Sustainable Low-Volume Analysis of Environmental Samples by Semi-Automated Prioritization of Extracts for Natural Product Research (SeaPEPR). *Mar. Drugs* **2020**, *18*, 649.
- (44) Allen, F.; Greiner, R.; Wishart, D. Competitive fragmentation modeling of ESI-MS/MS spectra for putative metabolite identification. *Metabolomics* **2015**, *11*, 98–110.
- (45) Laatsch, H. *AntiBase: The Natural Compound Identifier*; 5. Edition; Wiley-VCH, 2017.
- (46) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504.
- (47) Marner, M.; Patras, M. A.; Kurz, M.; Zubeil, F.; Förster, F.; Schuler, S.; Bauer, A.; Hammann, P.; Vilcinskas, A.; Schäberle, T. F.; Glaeser, J. Molecular Networking-Guided Discovery and Characterization of Stechlisins, a Group of Cyclic Lipopeptides from a *Pseudomonas* sp. *J. Nat. Prod.* **2020**, *83*, 2607–2617.
- (48) Previti, S.; Ettari, R.; Cosconati, S.; Amendola, G.; Chouchene, K.; Wagner, A.; Hellmich, U. A.; Ulrich, K.; Krauth-Siegel, R. L.; Wich, P. R.; et al. Development of Novel Peptide-Based Michael Acceptors Targeting Rhodesain and Falcipain-2 for the Treatment of Neglected Tropical Diseases (NTDs). *J. Med. Chem.* **2017**, *60*, 6911–6923.
- (49) Johé, P.; Jaenicke, E.; Neuweiler, H.; Schirmeister, T.; Kersten, C.; Hellmich, U. A. Structure, interdomain dynamics, and pH-dependent autoactivation of pro-rhodesain, the main lysosomal cysteine protease from African trypanosomes. *J. Biol. Chem.* **2021**, *296*, 100565.
- (50) Jung, S.; Fuchs, N.; Johe, P.; Wagner, A.; Diehl, E.; Yuliani, T.; Zimmer, C.; Barthels, F.; Zimmermann, R. A.; Klein, P.; et al. Fluorovinylsulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure-Activity Relationship, Inhibition Mechanism, Metabolism, and In Vivo Studies. *J. Med. Chem.* **2021**, *64*, 12322–12358.
- (51) Barthels, F.; Marincola, G.; Marciniak, T.; Konhäuser, M.; Hammerschmidt, S.; Bierlmeier, J.; Distler, U.; Wich, P. R.; Tenzer, S.; Schwarzer, D.; Ziebuhr, W.; Schirmeister, T. Asymmetric Disulfanylbenzamides as Irreversible and Selective Inhibitors of *Staphylococcus aureus* Sortase A. *ChemMedChem* **2020**, *15*, 839–850.
- (52) Steinmetzer, T.; Schweinitz, A.; Stürzebecher, A.; Dönnecke, D.; Uhland, K.; Schuster, O.; Steinmetzer, P.; Müller, F.; Friedrich, R.; Than, M. E.; Bode, W.; Stürzebecher, J. Secondary amides of sulfonylated 3-aminophenylalanine. New potent and selective inhibitors of matriptase. *J. Med. Chem.* **2006**, *49*, 4116–4126.
- (53) Leatherbarrow, R. J. *GraFit Version 5*; Erithacus Software Ltd.: Horley, U.K., 2006.
- (54) Schirmeister, T.; Kesselring, J.; Jung, S.; Schneider, T. H.; Weickert, A.; Becker, J.; Lee, W.; Bamberger, D.; Wich, P. R.; Distler, U.; et al. Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138*, 8332–8335.
- (55) Ludewig, S.; Kossner, M.; Schiller, M.; Baumann, K.; Schirmeister, T. Enzyme kinetics and hit validation in fluorimetric protease assays. *Curr. Top. Med. Chem.* **2010**, *10*, 368–382.
- (56) Sijwali, P. S.; Brinen, L. S.; Rosenthal, P. J. Systematic optimization of expression and refolding of the Plasmodium falciparum cysteine protease falcipain-2. *Protein Express. Purif.* **2001**, *22*, 128–134.
- (57) Kerr, I. D.; Lee, J. H.; Pandey, K. C.; Harrison, A.; Sajid, M.; Rosenthal, P. J.; Brinen, L. S. Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity. *J. Med. Chem.* **2009**, *52*, 852–857.
- (58) McGann, M. FRED pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2011**, *51*, 578–596.
- (59) McGann, M. FRED and HYBRID docking performance on standardized datasets. *J. Comput. Aid. Mol. Des.* **2012**, *26*, 897–906.
- (60) Lisk, G.; Pain, M.; Gluzman, I. Y.; Kambhampati, S.; Furuya, T.; Su, X.-Z.; Fay, M. P.; Goldberg, D. E.; Desai, S. A. Changes in the plasmodial surface anion channel reduce leupeptin uptake and can confer drug resistance in *Plasmodium falciparum*-infected erythrocytes. *Antimicrob. Agents Chemother.* **2008**, *52*, 2346–2354.
- (61) Hawkins, P. C. D.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **2010**, *50*, 572–584.
- (62) *Molecular Operating Environment (MOE)*; 2020.09; Chemical Computing Group ULC: 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2020. <https://www.chemcomp.com/index.htm>.

5.2 Supporting Information

Supporting Information

Identification, Characterization and Synthesis of Natural Parasitic Cysteine Protease Inhibitors – Pentacitidins are More Potent Falcitidin Analogs

Stephan Brinkmann^{1A}, Sandra Semmler^{1A}, Christian Kerster², Maria A. Patras¹, Michael Kurz³,

Natalie Fuchs², Stefan J. Hammerschmidt², Jenny Legac⁴, Peter E. Hammann⁵, Andreas

Vilcinskas^{1,6}, Philip. J. Rosenthal⁴, Tanja Schirmeister², Armin Bauer^{3,}, Till F. Schäberle^{1,6*}*

- 1 Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Branch for Bioresources, 35392 Giessen, Germany
- 2 Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany
- 3 Sanofi-Aventis Deutschland GmbH, R&D, 65926 Frankfurt am Main, Germany
- 4 Department of Medicine, University of California, San Francisco, 94143 California, United States

S1

5 Evotec International GmbH, 37079 Göttingen, Germany

6 Institute for Insect Biotechnology, Justus-Liebig-University of Giessen, 35392 Giessen,
Germany

λ Authors contributed equally

* Correspondence: Armin.Bauer@sanofi.com; Till.Schaeberle@ime.fraunhofer.de

S2

Contents

Synthesis	14
General procedures	14
Coupling of the amino acids and fatty acid	14
LC-MS and UHR-MS sample preparation	14
Fmoc-deprotection	14
Cleavage from the resin	14
Synthesis of the methyl ester from the corresponding acid	14
Reduction of the methyl ester to the alcohol	15
Dess-Martin oxidation of the alcohol to the aldehyde	15
(3-methylbutanoyl)-D-histidyl-L-isoleucyl-L-valyl-L-proline (17)	15
(S)-1-((3-methylbutanoyl)-D-histidyl-L-isoleucyl-L-valyl)pyrrolidine-2-carboxamide, falcitidin (1)	16
2-Chlorotrityl-L-Phe-L-Pro-L-NH ₂ (5)	17
2-Chlorotrityl-L-Phe-L-Pro-L-Val-L-Val-D-His(Trt)-NH ₂ (6)	17
2-Chlorotrityl-L-Phe-L-Pro-L-Val-L-Ile-D-His(Trt)-NH ₂ resin (7)	18
(3-methylbutanoyl)-D-histidyl-L-valyl-L-valyl-L-prolyl-L-phenylalanine (8)	18
(2-phenylacetyl)-D-histidyl-L-valyl-L-valyl-L-prolyl-L-phenylalanine (9)	19
(3-methylbutanoyl)-D-histidyl-L-isoleucyl-L-valyl-L-prolyl-L-phenylalanine (10)	20
(2-phenylacetyl)-D-histidyl-L-isoleucyl-L-valyl-L-prolyl-L-phenylalanine (11)	21
(S)-N-((S)-1-hydroxy-3-phenylpropan-2-yl)-1-((3-methylbutanoyl)-D-histidyl-L-valyl-L- valyl)pyrrolidine-2-carboxamide (12)	22
(S)-N-((S)-1-hydroxy-3-phenylpropan-2-yl)-1-((2-phenylacetyl)-D-histidyl-L-valyl-L- valyl)pyrrolidine-2-carboxamide (13)	23
(S)-N-((S)-1-hydroxy-3-phenylpropan-2-yl)-1-((3-methylbutanoyl)-D-histidyl-L-isoleucyl-L- valyl)pyrrolidine-2-carboxamide (14)	24

(<i>S</i>)- <i>N</i> ((<i>S</i>)-1-hydroxy-3-phenylpropan-2-yl)-1-((2-phenylacetyl)-D-histidyl-L-isoleucyl-L-valyl)pyrrolidine-2-carboxamide (4)	25
(2 <i>R</i> / <i>S</i>)-1-((3-methylbutanoyl)-D-histidyl-L-valyl-L-valyl)- <i>N</i> ((<i>S</i>)-1-oxo-3-phenylpropan-2-yl)pyrrolidine-2-carboxamide (15)	27
(2 <i>R</i> / <i>S</i>)- <i>N</i> ((<i>S</i>)-1-oxo-3-phenylpropan-2-yl)-1-((2-phenylacetyl)-D-histidyl-L-valyl-L-valyl)pyrrolidine-2-carboxamide (16)	28
(2 <i>R</i> / <i>S</i>)-1-((3-methylbutanoyl)-D-histidyl-L-isoleucyl-L-valyl)- <i>N</i> ((<i>S</i>)-1-oxo-3-phenylpropan-2-yl)pyrrolidine-2-carboxamide (2)	29
(2 <i>R</i> / <i>S</i>)- <i>N</i> ((<i>S</i>)-1-oxo-3-phenylpropan-2-yl)-1-((2-phenylacetyl)-D-histidyl-L-isoleucyl-L-valyl)pyrrolidine-2-carboxamide (3)	30
NMR spectra	31

Table S1. Stachelhaus code for the A-domains found in all pentacitidin BGCs. Bht, beta-hydroxy-tyrosine.	13
Table S2. NMR data for 17 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	32
Table S3. NMR data for 1 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	34
Table S4. NMR data for 8 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	36
Table S5. NMR data for 9 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	38
Table S6. NMR data for 10 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	40
Table S7. NMR data for 11 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	42
Table S8. NMR data for 12 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	44
Table S9. NMR data for 13 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	46
Table S10. NMR data for 14 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	48
Table S11. NMR data for 4 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	50
Table S12. NMR data for 15 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	52
Table S13. NMR data for 16 in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	55
Table S14. NMR data for synthetic pentacitidin A (2) in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	58
Table S15. NMR data for synthetic pentacitidin B (3) in DMSO- <i>d</i> ₆ (600 MHz, 151 MHz).	61
Table S16. Summary of molecular docking results.	63

Figure S1. MS ² -spectra of falcitidin peptide family A.	5
Figure S2. MS ² -spectra of falcitidin peptide family B.	6
Figure S3. MS ² -spectra of falcitidin peptide family C.	7
Figure S4. MS ² -spectra of falcitidin peptide family D.	8
Figure S5. NMR spectra of natural isolated pentacitidin A (2) in DMSO- <i>d</i> ₆ . NMR data can be found in Table 1.	9
Figure S6. NMR spectra of natural isolated pentacitidin B (3) in DMSO- <i>d</i> ₆ . NMR data can be found in Table 2.	10
Figure S7. ¹ H NMR section of natural isolated 3 and 4 in DMSO- <i>d</i> ₆	11
Figure S8. LC-MS chromatogram of the NMR sample of compound 3 before (A) and after NMR study (B).	11
Figure S9. Comparison of the Marfey derivatization products of a mixed sample containing the DCI hydrolysates of 3 and 4 and commercially available amino acid standards (histidine [two extracted ion chromatograms (EICs) with 716.2382 [M+H] ⁺ at 7.8 min and 436.1578 [M+H] ⁺ at 16.6 min corresponding to His+L-FDVA and His+2 L-FDVA, respectively], isoleucine, valine, proline, phenylalaninol [H-Phe-ol]) derivatized with L-FDVA. a) Commercially available L-amino acid standards. b) Commercially available D-amino acid standards. c) Hydrolysate of a mixture of 3 and 4	12
Figure S10. NMR spectra of 17 in DMSO- <i>d</i> ₆	31
Figure S11. NMR spectra of 1 in DMSO- <i>d</i> ₆	33
Figure S12. NMR spectra of 8 in DMSO- <i>d</i> ₆	35
Figure S13. NMR spectra of 9 in DMSO- <i>d</i> ₆	37
Figure S14. NMR spectra of 10 in DMSO- <i>d</i> ₆	39
Figure S15. NMR spectra of 11 in DMSO- <i>d</i> ₆	41
Figure S16. NMR spectra of 12 in DMSO- <i>d</i> ₆	43
Figure S17. NMR spectra of 13 in DMSO- <i>d</i> ₆	45
Figure S18. NMR spectra of 14 in DMSO- <i>d</i> ₆	47
Figure S19. NMR spectra of 4 in DMSO- <i>d</i> ₆	49
Figure S20. NMR spectra of 15 in DMSO- <i>d</i> ₆	51
Figure S21. NMR spectra of 16 in DMSO- <i>d</i> ₆	54
Figure S22. NMR spectra of synthetic pentacitidin A (2) in DMSO- <i>d</i> ₆	57
Figure S23. NMR spectra of synthetic pentacitidin B (3) in DMSO- <i>d</i> ₆	60

n.o. - not observed * - data obtained from HMBC or HSQC spectrum

If not noted otherwise, NMR data and spectra are of synthetic molecules. NMR spectra and data of natural isolated compounds are declared as such.

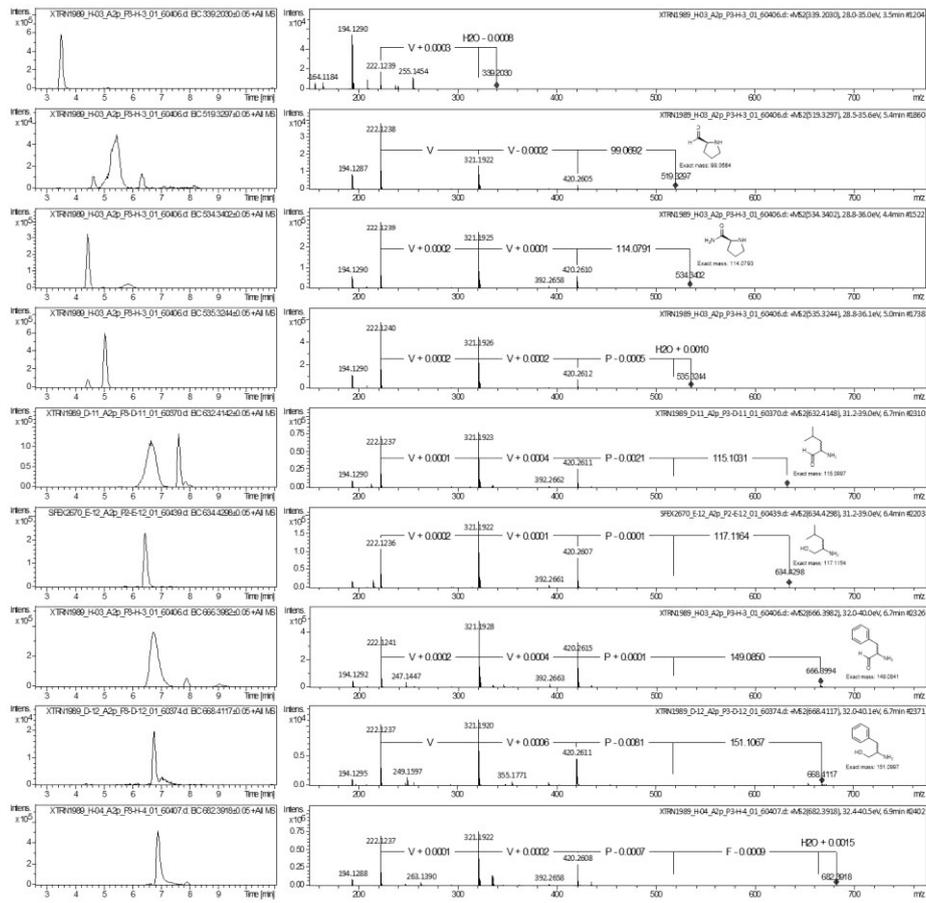


Figure S1. MS²-spectra of falcitidin peptide family A.

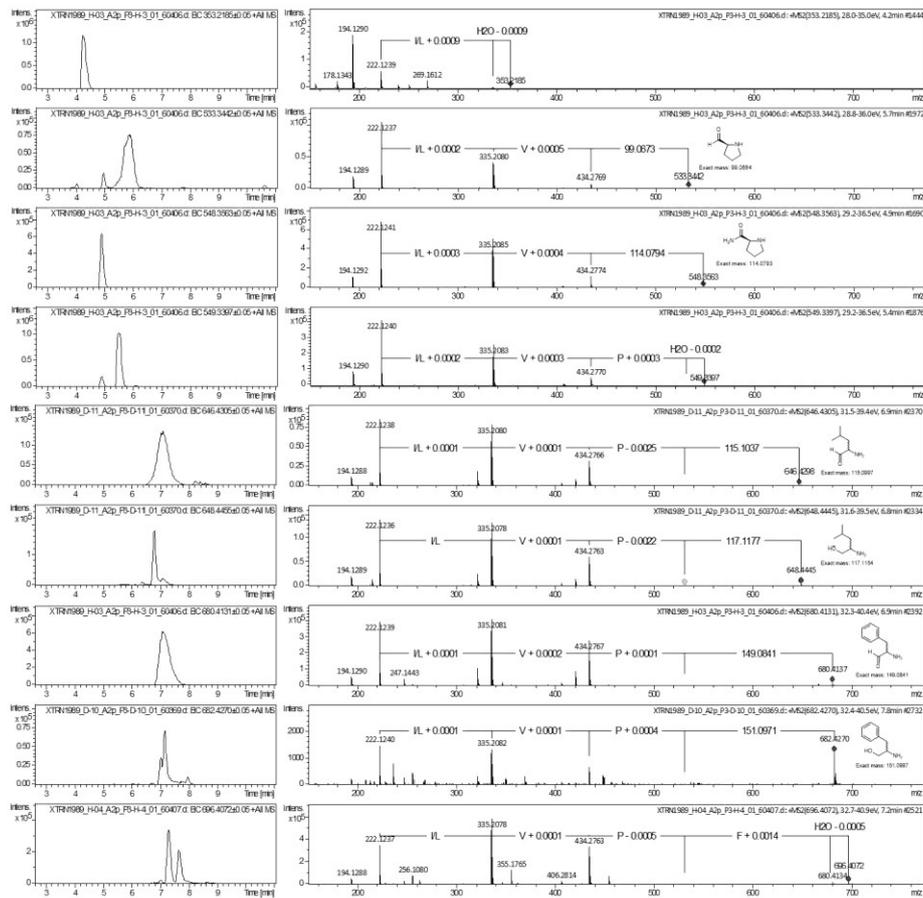


Figure S2. MS²-spectra of falcitidin peptide family B.

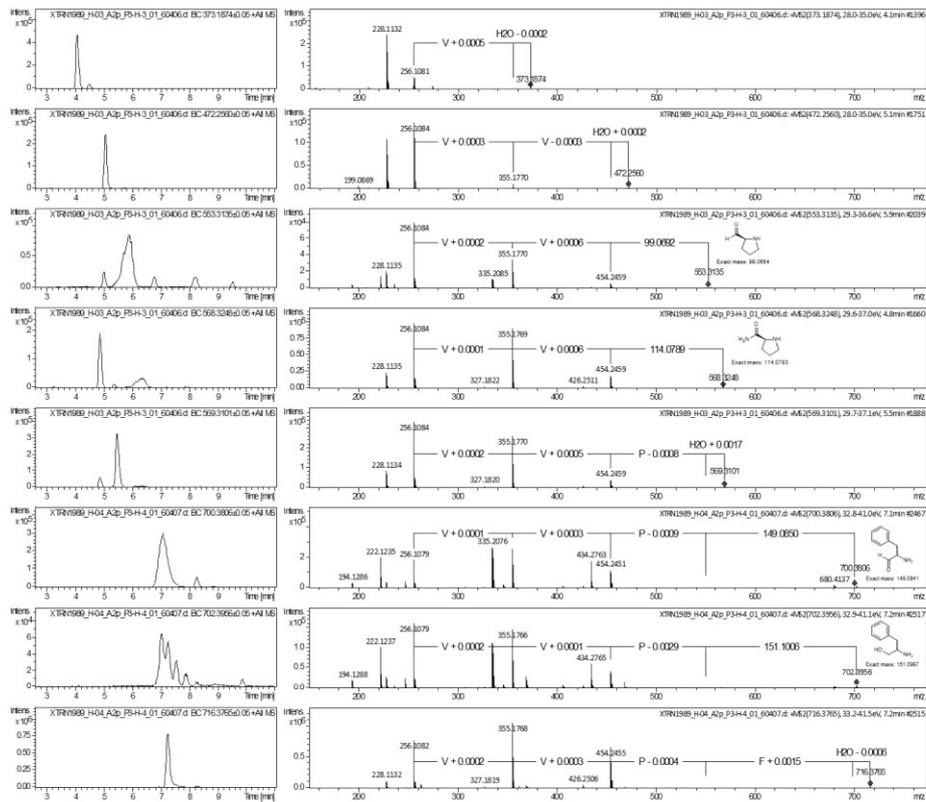


Figure S3. MS²-spectra of falcitidin peptide family C.

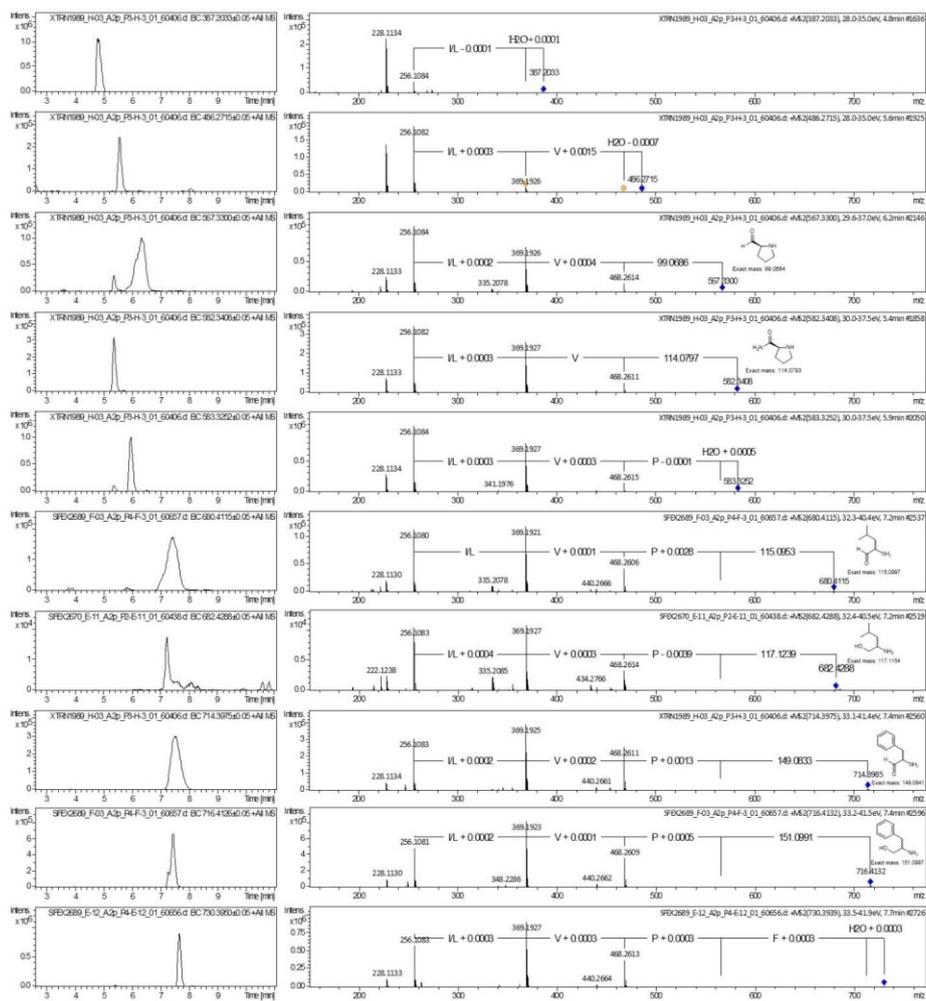


Figure S4. MS²-spectra of falcitidin peptide family D.

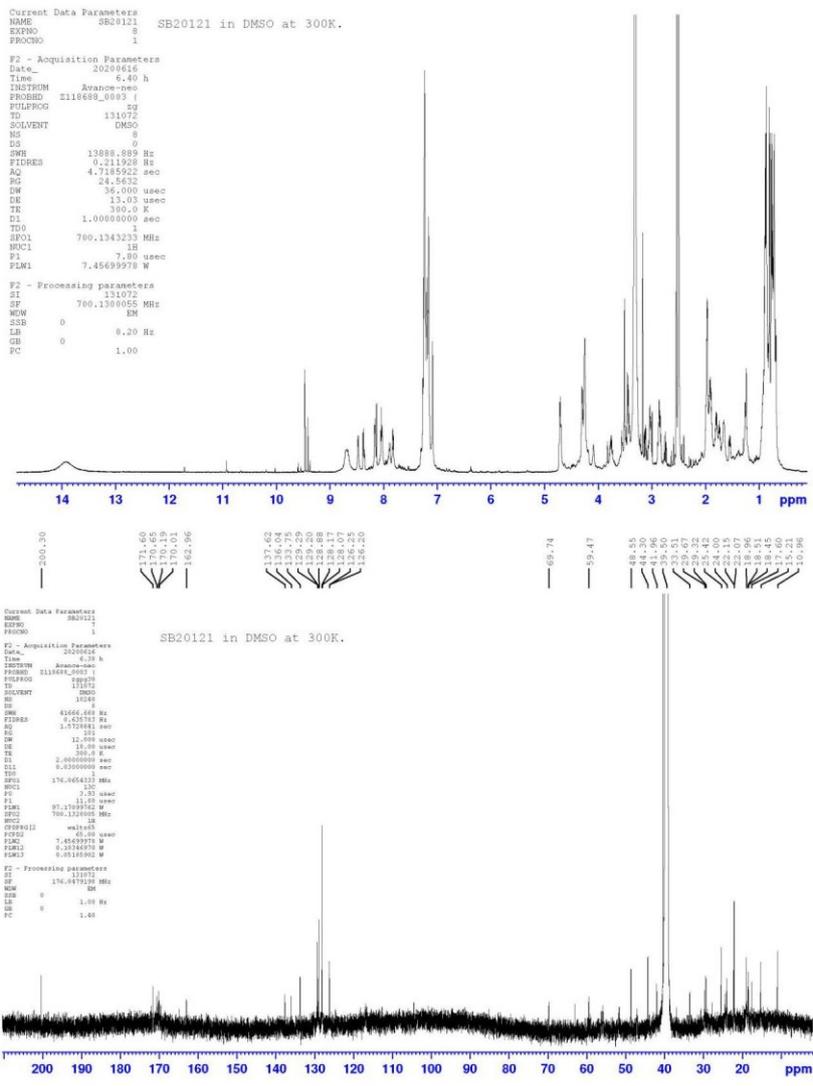


Figure S5. NMR spectra of natural isolated pentacitidin A (2) in DMSO-*d*₆. NMR data can be found in Table 1.

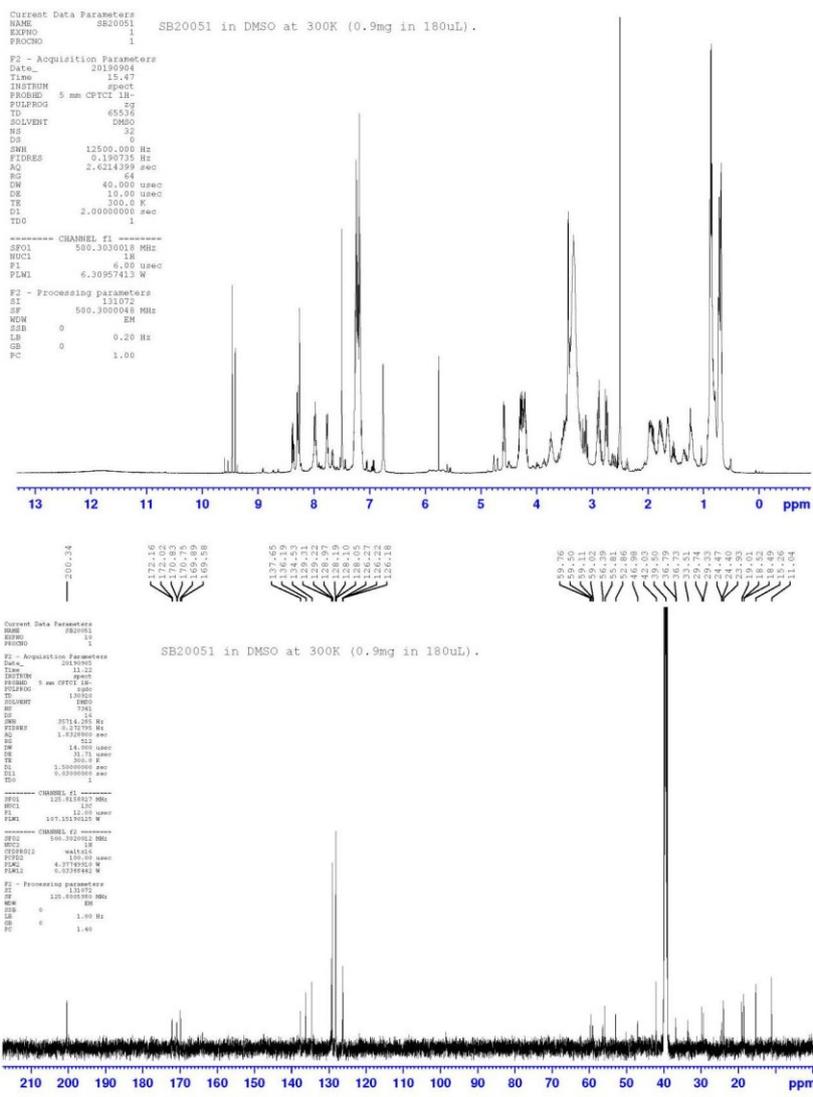


Figure S6. NMR spectra of natural isolated pentacitidin B (3) in DMSO-*d*₆. NMR data can be found in Table 2.

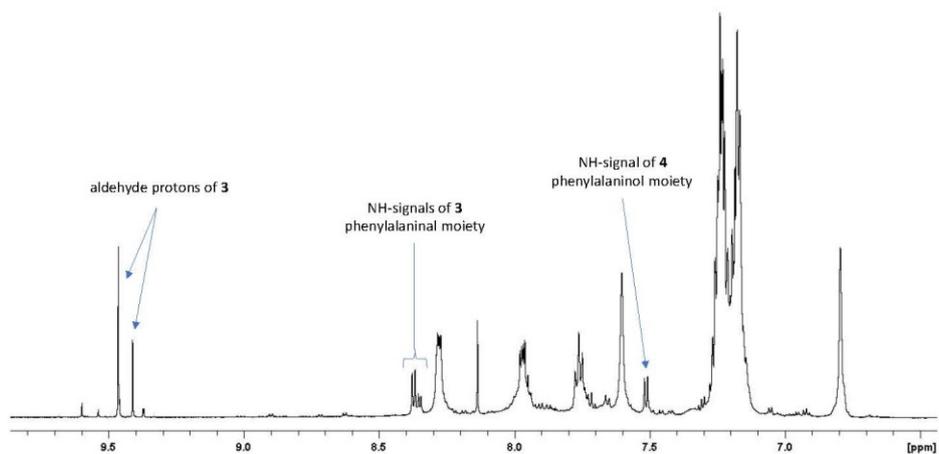
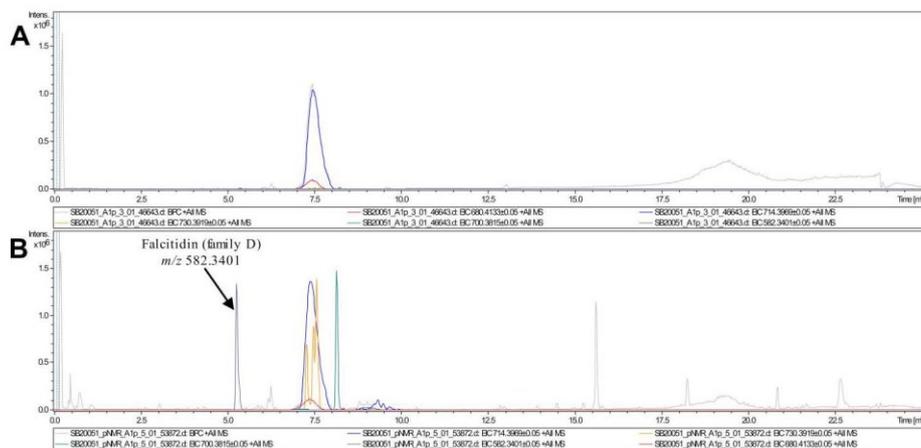


Figure S7. ¹H NMR section of natural isolated **3** and **4** in DMSO-*d*₆.



S13

Figure S8. LC-MS chromatogram of the NMR sample of compound **3** before (A) and after NMR study (B).

S14

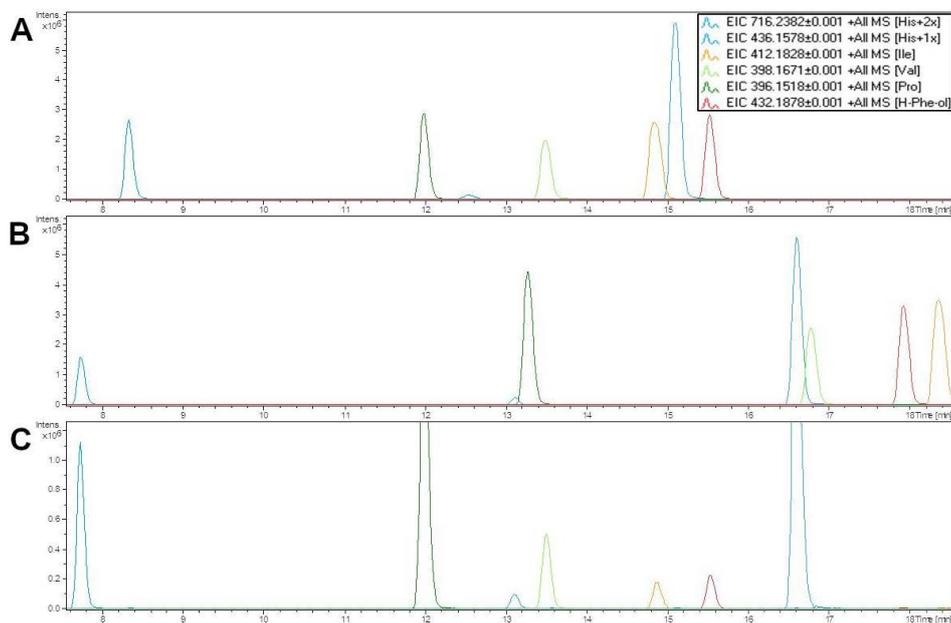


Figure S9. Comparison of the Marfey derivatization products of a mixed sample containing the DCl hydrolysates of **3** and **4** and commercially available amino acid standards (histidine [two extracted ion chromatograms (EICs) with 716.2382 [M+H]⁺ at 7.8 min and 436.1578 [M+H]⁺ at 16.6 min corresponding to His+L-FDVA and His+2 L-FDVA, respectively], isoleucine, valine, proline, phenylalaninol [H-Phe-ol]) derivatized with L-FDVA. **a)** Commercially available L-amino acid standards. **b)** Commercially available D-amino acid standards. **c)** Hydrolysate of a mixture of **3** and **4**.

Table S1. Stachelhaus code for the A-domains found in all pentacitidin BGCs. Bht, beta-hydroxy-tyrosine.

Strain	1 st A-domain (His)	2 nd A-domain (Ile)	3 rd A-domain (Val)	4 th A-domain (Pro)	5 th A-domain (Phe)
<i>C. eiseniae</i> DSM 22224	DSaLiAEVwK (70%)	DAFFIGITFK (90%)	DAfWLGgTFK (90%)	DVQFIAqVVK (90%)	DAmflcGiCK (60%)
<i>C. dinguisis</i> DSM 29821	DSaLiAEVwK (70%)	DAFFIGITFK (90%)	DAfWLGgTFK (90%)	DVQFIAqVVK (90%)	DAifvcAiCK (60%)
<i>C. varians</i> KCTC 52926	DSaLiAEVwK (70%)	DAFFIGITFK (90%)	DAfWfGaTFK (80%)	DVQFvAqVVK (80%)	DAmfvcAICK (60%)
<i>C. niastensis</i> DSM 24859	DSaLiAEVwK (70%)	DAFFIGITFK (90%)	DSaLiAEVwK (70% His)	DVQFfAqVVK (80%)	DAllyGAICK (70% Bht)
<i>C. varians</i> Ae27	DSaLiAEVwK (70%)	DAFFIGITFK (90%)	DAfWLGgTFK (90%)	DVQFvAqVVK (80%)	DAmfvcAICK (60%)

Synthesis

General procedures

Coupling of the amino acids and fatty acid¹

If not noted otherwise, all reactions were carried out in a custom-built solid-phase peptide synthesis vessel with a G2 filter and a diameter of 3, 4 or 5 cm at room temperature. Argon was used for agitation of the resin. 20% Piperidine in DMF and the cleavage cocktail were freshly prepared on the day of use.

The Fmoc-protected amino acid or fatty acid (3 equiv) and HATU (2.9 equiv) dissolved in DMF (5-15 mL) were added to the swelled resin (1 equiv), followed by DIPEA (6 equiv) and the resin was agitated for 1-3 hours.

Each coupling step was monitored as described in the general method part for the LC-MS sample preparation. Fmoc-deprotection was carried out after each coupling step was completed, as indicated by LC-MS result.

LC-MS and UHR-MS sample preparation¹

The reaction progress of each coupling of the Fmoc-protected amino acids was monitored using LC-MS. For that a few beads of the resin were sampled in a 2 mL SPPS syringe, washed once with DMF and then 2-3 times with DCM. The vessel was closed and 20% HFIP in DCM was added (1-1.5 mL), which changed the color of the beads from yellow-orange to a dark red which faded over time. The mixture was shaken for 15-30 min and the filtrate was directly used for LC-MS measurement. For UHR-MS measurement the filtrate was concentrated *in vacuo* and redissolved in AcCN or DMSO.

¹ W. Chan, P. White, Fmoc Solid Phase Peptide Synthesis Practical Approach, Oxford University Press, Oxford, 1999.

Fmoc-deprotection¹

The mixture was filtered and the remaining resin was washed 5 times with DMF. 20% Piperidine in DMF (20-30 mL) was added. After agitation for 2-5 min it was filtrated, rinsed with DMF and the process was repeated four more times. Then the resin was washed with DMF three times.

Cleavage from the resin¹

To the washed resin, the cleavage cocktail consisting of TFA:TIS:H₂O (95:2.5:2.5) (40 mL) was added, coloring the mixture a dark red. The resin was agitated for 25-30 min after which the supernatant was drained and the process was repeated once. The combined filtrates were reduced under pressure and then dried further using lyophilization.

Synthesis of the methyl ester from the corresponding acid

For the esterification of the peptide, the acid (1 equiv) was dissolved in anhydrous methanol and cooled to 0 °C. . Thionyl chlorid (1.5 equiv) was added dropwise and the mixture was stirred under a gentle reflux for 2-4 h. The reaction progress was monitored using LC-MS or UHR-MS. The solvent was evaporated *in vacuo* and the residue was dried further under high vacuum. The crude product was directly used for the synthesis of the alcohol, without further purification.

Reduction of the methyl ester to the alcohol

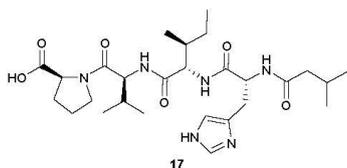
The methyl ester of the peptide (1 equiv) was suspended in anhydrous THF and cooled to 0 °C. LiBH₄ (4 equiv) was added and the reaction was stirred at room temperature for 4-8 h. The reaction progress was monitored periodically using LC-MS and if needed more LiBH₄ (2 equiv) was added at 0 °C. The reaction was quenched using sat. NH₄Cl solution. Water was added until all formed salts were dissolved. The layers were separated and the aqueous layer was extracted with ethyl acetate three times. The combined organic phase was washed once with brine, dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure. Purification by HPLC yielded the product as a white powder or a colorless solid.

Dess-Martin oxidation of the alcohol to the aldehyde

The alcohol (1 equiv) was suspended in anhyd. DCM and DMP (15% in DCM, 3 equiv) was added. The mixture was stirred at room temperature for a total of 2 hours. The reaction progress was monitored every 25 min by UHR-MS. DMP (1 equiv, 0.5 equiv) was added as soon as no more significant conversion of the alcohol to the aldehyde could be observed. At the time a conversion

of over 95% was observed, the reaction was quenched with methanol and concentrated *in vacuo*. Purification by HPLC yielded the product as a colorless solid. Note that the DCM used, must not be stabilized with methanol!. Amylene as a stabilizing reagent worked fine.

(3-methylbutanoyl)-D-histidyl-L-isoleucyl-L-valyl-L-proline (**17**)



The H-L-proline-chlorotrityl resin ($n = 0.72$ mmol/g, 2.70 g, 1.95 mmol) was swelled in DMF for 30 min. After removal of the solvent, Fmoc-(L)-valine-OH (1.987 g, 5.846 mmol) and HATU (2.153 g, 5.662 mmol), dissolved in DMF (5 mL), were added, followed by DIPEA (1986 μ L, 11.68 mmol) and more DMF (15 mL). The mixture was agitated for 1.5 h. After Fmoc-deprotection, Fmoc-(L)-isoleucine-OH (2.062 g, 5.834 mmol) and HATU (2.149 g, 5.652 mmol), dissolved in DMF (5 mL) were added, followed by DIPEA (1986 μ L, 11.68 mmol) and more DMF (15 mL). The mixture was agitated for 1 h. Following Fmoc-deprotection, Fmoc-D-histidine(Trt)-OH (3.621 g, 5.843 mmol) and HATU (2.147 g, 5.647 mmol), solved in DMF (10 mL) were added, then DIPEA (1986 μ L, 11.68 mmol) and more DMF (10 mL). The mixture was agitated for 1.5 h. After Fmoc-deprotection HATU (2.147 g, 5.647 mmol), solved in DMF (3 mL) was added to the resin, then isovaleric acid (642 μ L, 5.844 mmol) and DIPEA (1986 μ L, 11.68 mmol). More DMF (15 mL) was added and the mixture was agitated for 1.5 h. After draining the supernatant, the resin was washed with DMF (3x), DMC (2x) and MeOH (2x). The cleavage was performed as described in the general method section. The crude product was washed three times with ice cold diethyl ether and dried under reduced pressure. Purification of 40% of the crude (1.935 g) using semi preparative HPLC (5-40-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min) yielded **17** as a white wax like film (328.7 mg, 31%, 0.599 mmol, overall yield calculated to be: 78%).

$^1\text{H-NMR}$ (DMSO- d_6 , 600 MHz): δ_{H} [ppm] = 8.97 (d, 1H, $J = 1.1$ Hz, $\epsilon\text{-CH}_{\text{arom}}$ His), 8.15 (d, 1H, $J = 8.4$ Hz, NHHis), 8.05 (d, 1H, $J = 8.3$ Hz, NHVal), 7.84 (d, 1H, $J = 8.8$ Hz, NHlle), 7.35 (s, 1H, $\delta\text{-CH}_{\text{arom}}$ His), 4.74 (ddd, 1H, $J = 8.2, 8.2, 6.6$ Hz, $\alpha\text{-CHHis}$), 4.27 (dd, 1H, $J = 9.0, 9.0$ Hz, $\alpha\text{-CHVal}$), 4.25 (dd, 1H, $J = 9.0, 7.0$ Hz, $\alpha\text{-CHlle}$), 4.17 (dd, 1H, $J = 8.7, 5.0$ Hz, $\alpha\text{-CHPro}$), 3.81 (ddd, 1H, $J = 9.8, 6.4, 6.4$ Hz, $\delta\text{-CH}_2\text{N a Pro}$), 3.56 (ddd, 1H, $J = 10.1, 6.4, 6.4$ Hz, $\delta\text{-CH}_2\text{N b Pro}$),

3.06 (dd, 1H, $J = 15.1, 6.1$ Hz, β -CH₂ a His), 2.86 (dd, 1H, $J = 15.1, 8.7$ Hz, β -CH₂ b His), 2.16-2.09 (m, 1H, β -CH₂ a Pro), 2.02-1.94 (m, 3H, CH₂ iVal, β -CHVal), 1.94-1.85 (m, 2H, CHiVal, γ -CH₂ a Pro), 1.85-1.79 (m, 2H, β -CH₂ b Pro, γ -CH₂ b Pro), 1.69-1.62 (m, 1H, β -CHIle), 1.28-1.20 (m, 1H, γ -CH₂ a Ile), 0.97-0.92 (m, 1H, γ -CH₂ b Ile), 0.91 (d, 1H, $J = 6.8$ Hz, γ -CH₃ a Val), 0.88 (d, 1H, $J = 6.8$ Hz, γ -CH₃ b Val), 0.80 (d, 1H, $J = 6.8$ Hz, CH₃ a iVal), 0.76 (d, 1H, $J = 6.4$ Hz, CH₃ b iVal), 0.74 (t, 1H, $J = 7.4$ Hz, γ -CH₃ Ile), 0.71 (d, 1H, $J = 6.7$ Hz, δ -CH₃ Ile).

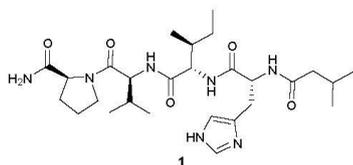
¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_C [ppm] = 173.1 (COOH Pro), 171.7 (CO iVal), 170.7 (CO Ile), 169.8 (CO His), 169.6 (CO Val), 133.7* (*C*_{arom} His), 129.5 (*C*_{quart} His), 116.8 (*C*_{arom} His), 58.5 (α -CH Pro), 56.2 (α -CH Ile), 55.7 (α -CH Val), 51.3 (α -CH His), 46.8 (δ -CH₂-N Pro), 44.3 (CH₂ iVal), 36.9 (β -CH Ile), 29.8 (β -CH Val), 28.7 (β -CH₂ Pro), 27.3 (β -CH₂ His), 25.4 (CH iVal), 24.6 (γ -CH₂ Pro), 24.0 (γ -CH₂ Ile), 22.14 (CH₃ iVal), 22.06 (CH₃ iVal), 18.8 (γ -CH₃ Val), 18.4 (γ -CH₃ Val), 15.2 (δ -CH₃ Ile), 10.9 (γ -CH₃ Ile).

Additional found signals: δ_H [ppm] = Broad peak between δ_H 4.08 to 3.60 (H₂O), 2.99, 2.54 (DMSO). ϵ -NHHis and COOHPro were not observed. δ_C [ppm] = 158.0, 40.4 (DMSO).

UHR-MS (ESI-TOF) m/z calcd for C₂₇H₄₅N₆O₆: 549.3395 [M+H]⁺; found: 549.3401 [M+H]⁺

Specific rotation $[\alpha]_D^{21} = -24.3^\circ$ ($c = 1.44$, CH₃OH)

(*S*)-1-((3-methylbutanoyl)-D-histidyl-L-isoleucyl-L-valyl)pyrrolidine-2-carboxamide, falcitidin (**1**)



Peptide **17** (100.3 mg, 0.1828 mmol) was dissolved in DMF and cooled to 0 °C. HOAt (25.1 mg, 0.184 mmol) and EDC · HCl (36.9 mg, 0.192 mmol) were added and the mixture was stirred for 15 min at 0 °C. Aqueous NH₄OH solution (25%, 220 μ L, 1.41 mmol) was added

dropwise and the yellow mixture was stirred at room temperature for 6 h. It was then concentrated *in vacuo*. Purification of the crude using semi preparative HPLC (5-35-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min) yielded **1** as a colorless syrup (51.7 mg, 0.0944 mmol, 52%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 8.02 (d, $J=8.2$ Hz, *NH*His), 7.99 (d, 1H, $J=8.3$ Hz, *NH*Val), 7.86 (s, 1H, ϵ -*CH*_{arom}His), 7.70 (d, 1H, $J=8.8$ Hz, *NH*Ile), 7.23 (s, 1H, *CONH*₂Pro), 6.90 (s, 1H, δ -*CH*_{arom}His), 6.81 (s, 1H, *CONH*₂Pro), 4.62 (ddd, $J=8.3, 8.3, 6.1$ Hz, α -*CH*His), 4.27 (dd, 1H, $J=8.4, 8.4$ Hz, α -*CH*Val), 4.23 (dd, 1H, $J=8.7, 6.9$ Hz, α -*CH*Ile), 4.21 (dd, 1H, $J=8.5, 4.4$ Hz, α -*CH*Pro), 3.78-3.73 (m, 1H, δ -*CH*₂N a Pro), 3.58-3.52 (m, 1H, δ -*CH*₂N b Pro), 2.92 (dd, $J=14.9, 5.9$ Hz, β -*CH*₂ a His), 2.76 (dd, $J=15.1, 8.8$ Hz, β -*CH*₂ b His), 2.03-1.97 (m, 2H, β -*CH*₂ a Pro, β -*CH*Val), 1.96 (d, $J=6.7$ Hz, *CH*₂ iVal), 1.94-1.86 (m, 2H, γ -*CH*₂ a Pro, *CH* iVal), 1.82-1.74 (m, 2H, γ -*CH*₂ b Pro, β -*CH*₂ b Pro), 1.69-1.61 (m, 1H, β -*CH*Ile), 1.31-1.22 (m, 1H, γ -*CH*₂ a Ile), 0.98-0.91 (m, 1H, γ -*CH*₂ b Ile), 0.90 (d, 1H, $J=6.8$ Hz, γ -*CH*₃ a Val), 0.87 (d, 1H, $J=6.7$ Hz, γ -*CH*₃ b Val), 0.81 (d, $J=6.5$ Hz, *CH*₃ a iVal), 0.78 (d, $J=6.5$ Hz, *CH*₃ b iVal), 0.74 (t, 1H, $J=7.4$ Hz, δ -*CH*₃ Ile), 0.71 (d, 1H, $J=7.0$ Hz, γ -*CH*₃ Ile).

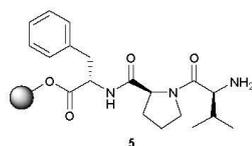
¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_{C} [ppm] = 173.4 (*CONH*₂Pro), 171.5 (*CO* iVal), 170.70 (*CO* Ile), 170.67 (*CO* His), 169.5 (*CO* Val), 134.3 (weak, ϵ -*CH*_{arom}His), 132.3 (weak, γ -*C*_{quart}His), 117.0 (weak, δ -*CH*_{arom}His), 59.2 (α -*CH*Pro), 56.3 (α -*CH*Ile), 55.8 (α -*CH*Val), 52.3 (α -*CH*His), 47.0 (δ -*CH*₂NPro), 44.4 (*CH*₂ iVal), 36.9 (β -*CH*Ile), 29.7 (β -*CH*Val), 29.3 (β -*CH*₂Pro, β -*CH*₂His), 25.5 (*CH* iVal), 24.5 (γ -*CH*₂Pro), 24.0 (γ -*CH*₂Ile), 22.2 (*CH*₃ iVal), 22.1 (*CH*₃ iVal), 19.1 (γ -*CH*₃Val), 18.4 (γ -*CH*₃Val), 15.2 (δ -*CH*₃Ile), 11.0 (γ -*CH*₃Ile).

Additional found signals: δ_{H} [ppm] = 8.14 (FA), 3.17 (MeOH), 2.54 (DMSO). ϵ -*NH*His was not observed. δ_{C} [ppm] = 163.0 (FA), 48.6 (MeOH), 40.4 (DMSO).

UHR-MS (ESI-TOF) m/z calcd for C₂₇H₄₆N₇O₅: 548.3555 [M+H]⁺; found: 548.3556 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{21} = -70.0^{\circ}$ ($c = 1.00$, CH₃OH), the determined specific rotation corresponds to literature.¹⁰

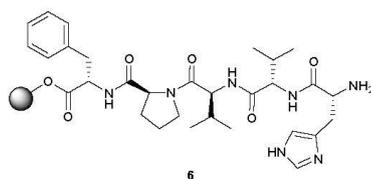
2-Chlorotriptyl-L-Phe-L-Pro-L-NH₂ (5)



H-L-phenylalanine-chlorotriptyl resin ($n = 0.78$ mmol/g, 15.01 g, 11.71 mmol) was swelled in DMF for 30 min. After removal of the solvent, Fmoc-(L)-proline-OH (11.84 g, 35.09 mmol) and HATU (12.89 g, 33.90 mmol) were added, dissolved in a small volume of DMF, followed by DIPEA (12.0 mL, 70.6 mmol) and more DMF. The mixture was agitated for

2 h. After Fmoc-deprotection, Fmoc-(L)-valine-OH (11.91 g, 35.09 mmol) and HATU (12.89 g, 33.90 mmol) were added, dissolved in a small volume of DMF, followed by DIPEA (12.0 mL, 70.6 mmol) and more DMF. The mixture was agitated for 2 h. Following Fmoc-deprotection the resin was washed two times with DCM and dried *in vacuo*, yielding **5** (38.48 g, 11.71 mmol).

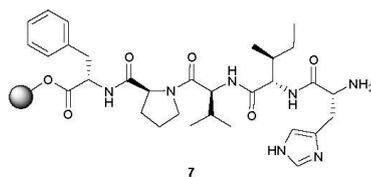
2-Chlorotrityl-L-Phe-L-Pro-L-Val-L-Val-D-His(Trt)-NH₂ (**6**)



Resin bound peptide **5** (19.24 g, 5.855 mmol) was swelled in DMF for 10 min. After removal of the solvent, Fmoc-(L)-valine-OH (5.97 g, 17.59 mmol) and HATU (10.89 g, 28.64 mmol)² were added, dissolved in a small volume of DMF, followed by DIPEA (6.0 mL, 35 mmol)

and more DMF. The mixture was agitated for 2 h. After Fmoc-deprotection, Fmoc-(D)-histidine-(Trt)-OH (10.89 g, 17.57 mmol) and HATU (6.454 g, 16.96 mmol) were added, dissolved in a small amount of DMF, followed by DIPEA (6.0 mL, 35 mmol) and more DMF. The mixture was agitated for 2 h. Since no complete conversion was observed, the supernatant was drained, the resin was washed three times with DMF and the coupling step was repeated. Fmoc-(D)-histidine-(Trt)-OH (3.618 g, 5.838 mmol) and HATU (2.011 g, 5.289 mmol) were added, dissolved in a small volume of DMF, followed by DIPEA (2000 μ L, 11.76 mmol) and more DMF. The mixture was agitated for 1 h, after which complete conversion was observed using UHR-MS. Following Fmoc-deprotection, the resin was washed two times with DCM and dried *in vacuo*, yielding **6** (41.35 g, 5.855 mmol).

2-Chlorotrityl-L-Phe-L-Pro-L-Val-L-Ile-D-His(Trt)-NH₂ resin (**7**)

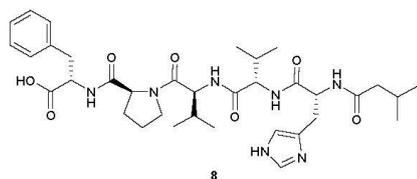


5 (19.24, 5.855 mmol) was swelled in DMF for 20 min. After removal of the solvent, Fmoc-(L)-Ile-OH (6.207 g, 17.56 mmol) and HATU (6.456 g, 16.98 mmol) were added, dissolved in a small volume of DMF, followed by

² Due to miscalculation the equivalents don't adhere to the standard procedure.

DIPEA (6.0 mL, 35 mmol) and more DMF. The mixture was agitated for 1.5 h. After Fmoc-deprotection, Fmoc-(D)-histidine-(Trt)-OH (10.54g, 17.00 mmol) and HATU (6.440 g, 16.96 mmol) were added, dissolved in a small amount DMF, followed by DIPEA (6.0 mL, 35 mmol) and more DMF. The mixture was agitated for 2 h. Following Fmoc-deprotection the resin was washed two times with DCM and dried *in vacuo*, yielding **7** (21.48 g, 5.855 mmol).

(3-methylbutanoyl)-D-histidyl-L-valyl-L-valyl-L-prolyl-L-phenylalanine (**8**)



6 (13.00 g, 1.841 mmol) was swelled in DMF for 10 min. After removal of the solvent, HATU (2.150 g, 5.65 mmol) was added, dissolved in a small volume of DMF, followed by isovaleric acid (643 μ L, 5.85 mmol), DIPEA (2000 μ L, 11.76 mmol) and more DMF. The mixture was

agitated for 3 h, after which LC-MS indicated completeness of the reaction. The supernatant was drained, and the resin was washed with DMF (3x) and DMC (2x). The cleavage was done as described in the general method section. The obtained crude was washed four times with ice cold diethyl ether, decanted and dried under reduced pressure. Purification of 48% of the crude using semi preparative HPLC (5-50-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min) yielded **8** as a white waxy film (196 mg, 0.288 mmol, overall yield calculated to be: 33%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 8.03 (d, 1H, *J* = 8.1 Hz, *NH*His), 8.03 (d, 1H, *J* = 8.1 Hz, ϵ -*NH*His), 7.99 (d, 1H, *J* = 7.9 Hz, *NH*Val 1), 7.98 (d, 1H, *J* = 7.2 Hz, *NH*Phe), 7.91 (s, 1H, ϵ -*CH*_{arom} His), 7.69 (d, 1H, *J* = 9.0 Hz, *NH*Val 2), 7.28-7.17 (m, 5H, δ -*CH*_{arom} Phe), 6.93 (s, 1H, δ -*CH*_{arom} His), 4.63 (ddd, 1H, *J* = 8.5, 8.5, 6.0 Hz, α -*CH*His), 4.38 (ddd, 1H, *J* = 7.6, 7.6, 5.7 Hz, α -*CH*Phe), 4.36 (dd, 1H, *J* = 8.3, 3.7 Hz, α -*CH*Pro), 4.25 (dd, 1H, *J* = 8.3, 8.3 Hz, α -*CH*Val 1), 4.22 (dd, 1H, *J* = 8.9, 6.1 Hz, α -*CH*Val), 3.75 (ddd, 1H, *J* = 9.6, 6.9, 6.9 Hz, δ -*CH*₂N a Pro), 3.53 (ddd, 1H, *J* = 9.4, 7.1, 5.6 Hz, δ -*CH*₂N b Pro), 3.00 (dd, 1H, *J* = 14.2, 5.8 Hz, β -*CH*₂ a Phe), 2.94 (dd, 1H, *J* = 14.6, 6.5 Hz, β -*CH*₂ a His), 2.92 (dd, 1H, *J* = 13.8, 7.5 Hz, β -*CH*₂ b Phe), 2.77 (dd, 1H, *J* = 13.8, 7.5 Hz, β -*CH*₂ b His), 2.00-1.93 (m, 4H, β -*CH*₂ a Pro, β -*CH*Val, *CH*₂ iVal), 1.93-1.89 (m, 2H, β -*CH*Val, *CH* iVal), 1.89-1.83 (m, 1H, γ -*CH*₂ a Pro), 1.83-1.75 (m, 2H, β -*CH*₂

S23

b Pro, γ -CH₂ b Pro), 0.88 (d, 3H, J = 6.8 Hz, γ -CH₃ a Val 1), 0.86 (d, 3H, J = 6.4 Hz, γ -CH₃ b Val 1), 0.81 (d, 3H, J = 6.4 Hz, CH₃ a iVal), 0.78 (d, 3H, J = 6.6 Hz, CH₃ b iVal), 0.73 (d, 3H, J = 6.6 Hz, γ -CH₃ a Val), 0.70 (d, 3H, J = 7.0 Hz, γ -CH₃ b Val).

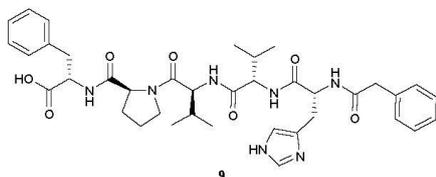
¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_C [ppm] = 172.7 (COOH Phe), 171.5 (CO iVal), 171.4 (CO Pro), 170.7 (CO His), 170.6 (CO Val 2), 169.7 (CO Val 1), 137.3 (γ -C_{quart} Phe), 134.3 (weak, ϵ -CH_{arom} His), 132.2* (γ -C_{quart} His), 129.1, 128.1, 126.4 (δ -CH_{arom} Phe), 117.0* (δ -CH_{arom} His), 59.0 (α -CH Pro), 56.9 (α -CH Val 2), 55.8 (α -CH Val 1), 53.5 (α -CH Phe), 52.3 (α -CH His), 47.0 (δ -CH₂N Pro), 44.4 (CH₂ iVal), 36.7 (β -CH₂ a Phe), 30.7 (β -CH Val), 29.7 (β -CH Val 1), 29.1 (weak, β -CH₂ His), 28.9 (β -CH₂ Pro), 25.4 (CH iVal), 24.3 (γ -CH₂ Pro), 22.2 (CH₃ a iVal), 22.1 (CH₃ b iVal), 19.0 (γ -CH₃ a Val 1), 18.4 (γ -CH₃ b Val 1), 17.6 (γ -CH₃ b Val).

Additional found signals: δ_H [ppm] = 8.14 (FA), Broad peak between 4.00 to 2.73 (H₂O), 3.10, 2.54 (DMSO), 1.18. ϵ -NHHis and COOH Phe were not observed. δ_C [ppm] = 45.7, 8.58.

UHR-MS (ESI-TOF) m/z calcd for C₃₅H₅₂N₇O₇: 682.3923 [M+H]⁺; found: 682.3909 [M+H]⁺

Specific rotation $[\alpha]_D^{21} = -11.2^\circ$ ($c = 1.78$, CH₃OH)

(2-phenylacetyl)-D-histidyl-L-valyl-L-valyl-L-prolyl-L-phenylalanine (**9**)



6 (13.70 g, 1.940 mmol) was swelled in DMF for 10 min. After removal of the solvent, HATU (2.159 g, 5.68 mmol) was added, dissolved in a small volume of DMF, followed by phenylacetic acid (0.799 g, 5.87 mmol), DIPEA (2000 μ L,

11.76 mmol) and more DMF. The mixture was agitated for 3 h, after which LC-MS indicated completeness of the reaction. The supernatant was drained and the resin was washed with DMF (3x), DMC (2x) and MeOH (2x). The cleavage was done as described in the general method section. The obtained crude was washed four times with ice cold diethyl ether, decanted and dried under reduced pressure. Purification of 37% of the crude using semi preparative HPLC (5-50-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min) yielded **9** as a white wax like film (303 mg, 0.423 mmol, overall yield calculated to be: 59%).

S24

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 8.32 (d, 1H, *J* = 8.3 Hz, *NH*His), 7.98 (d, 1H, *J* = 7.2 Hz, *NHPhe*), 7.97 (d, 1H, *J* = 7.5 Hz, *NHVal* 1), 7.82 (s, 1H, ϵ -*CH*_{arom} His), 7.76 (d, 1H, *J* = 8.9 Hz, *NHVal* 2), 7.28-7.15 (m, 10H, δ -*CH*_{arom} Phe, *CH*_{arom} PA), 6.88 (s, 1H, δ -*CH*_{arom} His), 4.63 (ddd, 1H, *J* = 7.7, 7.7, 6.6 Hz, α -*CH*His), 4.39 (ddd, 1H, *J* = 7.6, 7.6, 6.0 Hz, α -*CH*Phe), 4.37 (dd, 1H, *J* = 8.5, 3.4 Hz, α -*CH*Pro), 4.26 (dd, 1H, *J* = 8.3, 8.3 Hz, α -*CHVal* 1), 4.22 (dd, 1H, *J* = 9.0, 6.2 Hz, α -*CHVal* 2), 3.76 (ddd, 1H, *J* = 9.8, 6.8, 6.8 Hz, δ -*CH*₂N a Pro), 3.54 (ddd, 1H, *J* = 9.4, 7.2, 5.4 Hz, δ -*CH*₂N b Pro), 3.46 (d 1H, *J* = 14.3 Hz, *CH*₂ a PA), 3.42 (d 1H, *J* = 14.2 Hz, *CH*₂ b PA), 3.00 (dd, 1H, *J* = 14.1, 5.9 Hz, β -*CH*₂ a Phe), 2.95 (dd, 1H, *J* = 14.8, 5.9 Hz, β -*CH*₂ a His), 2.92 (dd, 1H, *J* = 13.9, 7.7 Hz, β -*CH*₂ b Phe), 2.79 (dd, 1H, *J* = 14.8, 8.4 Hz, β -*CH*₂ b His), 2.00-1.93 (m, 2H, β -*CH*₂ a Pro, β -*CHVal* 1), 1.93-1.89 (m, 1H, β -*CHVal* 2), 1.89-1.84 (m, 1H, γ -*CH*₂ a Pro), 1.84-1.76 (m, 2H, β -*CH*₂ b Pro, γ -*CH*₂ b Pro), 0.88 (d, 3H, *J* = 6.6 Hz, γ -*CH*₃ a Val 1), 0.86 (d, 3H, *J* = 6.7 Hz, γ -*CH*₃ b Val 1), 0.71 (d, 3H, *J* = 6.8 Hz, γ -*CH*₃ a Val 2), 0.67 (d, 3H, *J* = 6.7 Hz, γ -*CH*₃ b Val 2).

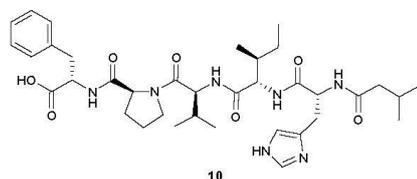
¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_{C} [ppm] = 172.7 (*COOH* Phe), 171.5 (*CO* Pro), 170.7 (*CO* His), 170.6 (*CO* Val 2), 170.0 (*CO* PA), 169.7 (*CO* Val 1), 137.3 (γ -*C*_{quart} Phe), 136.1 (*C*_{quart} PA), 134.3 (ϵ -*CH*_{arom} His), 132.3 (weak, γ -*C*_{quart} His), 129.1, 128.9, 128.1, 128.1, 126.4, 126.2 (δ -*CH*_{arom} Phe, *CH*_{arom} PA), 117.0* (δ -*CH*_{arom} His), 59.0 (α -*CH* Pro), 57.0 (α -*CH* Val 2), 55.9 (α -*CH* Val 1), 53.5 (α -*CH* Phe), 52.6 (α -*CH* His), 47.0 (δ -*CH*₂N Pro), 42.0 (*CH*₂ PA), 36.7 (β -*CH*₂ Phe), 30.6 (β -*CH* Val 2), 29.8 (β -*CH* Val 1), 29.5 (β -*CH*₂His), 28.9 (β -*CH*₂ Pro), 24.3 (γ -*CH*₂ Pro), 19.0 (γ -*CH*₃ a Val 2), 19.0 (γ -*CH*₃ a Val 1), 18.5 (γ -*CH*₃ b Val 1), 17.6 (γ -*CH*₃ b Val 2).

Additional found signals: δ_{H} [ppm] = 8.14 (FA), 2.54 (DMSO), 3.09, 1.17. ϵ -*NH*His and *COOH* Phe were not observed. δ_{C} [ppm] = 8.62.

UHR-MS (ESI-TOF) *m/z* calcd for C₃₈H₅₀N₇O₇: 716.3766 [M+H]⁺; found: 716.3766 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{21} = -12.4^{\circ}$ (c = 1.61, CH₃OH)

(3-methylbutanoyl)-D-histidyl-L-isoleucyl-L-valyl-L-prolyl-L-phenylalanine (**10**)



7 (8.090 g, 2.225 mmol) was swelled in DMF for 30 min. After removal of the solvent, HATU (2.425 g, 6.378 mmol) was added, dissolved in a

S25

small volume of DMF, followed by isovaleric acid (724 μ L, 6.59 mmol), DIPEA (2300 μ L, 13.53 mmol) and more DMF. The mixture was agitated for 2 h, after which LC-MS indicated completeness of the reaction. The supernatant was drained, and the resin was washed with DMF (3x), DMC (2x) and MeOH (2x). The cleavage was done as described in the general method section. The obtained crude was washed four times with ice cold diethyl ether, decanted and dried under reduced pressure. Purification of 35% of the crude using semi preparative HPLC (5-50-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min) yielded **10** as a white wax like film (87.8 mg, 0.126 mmol, overall yield calculated to be: 16%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 8.23 (s, 1H, ϵ -CH_{arom} His), 8.06 (d, 1H, *J* = 8.1 Hz, NHHis), 8.00 (d, 1H, *J* = 8.4 Hz, NHVal), 7.98 (d, 1H, *J* = 7.7 Hz, NH Phe), 7.73 (d, 1H, *J* = 9.2 Hz, NH Ile), 7.28-7.18 (m, 5H, δ -CH_{arom} Phe), 7.05 (s, 1H, δ -CH_{arom} His), 4.66 (ddd, 1H, *J* = 8.1, 8.1, 6.5 Hz, α -CHHis), 4.39 (ddd, 1H, *J* = 7.3, 7.3, 6.2 Hz, α -CH Phe), 4.36 (dd, 1H, *J* = 8.3, 3.9 Hz, α -CH Pro), 4.26 (dd, 1H, *J* = 8.3, 8.3 Hz, α -CH Val), 4.24 (dd, 1H, *J* = 8.5, 7.5 Hz, α -CH Ile), 3.77-3.71 (m, 1H, δ -CH₂N a Pro), 3.56-3.51 (m, 1H, δ -CH₂N b Pro), 3.00 (dd, 1H, *J* = 13.9, 5.9 Hz, β -CH₂ a Phe), 2.97 (dd, 1H, *J* = 15.3, 6.3 Hz, β -CH₂ a His), 2.92 (dd, 1H, *J* = 13.9, 7.9 Hz, β -CH₂ b Phe), 2.79 (dd, 1H, *J* = 15.0, 8.6 Hz, β -CH₂ b His), 2.00-1.92 (m, 4H, CH₂ iVal, β -CH Val, β -CH₂ a Pro), 1.92-1.83 (m, 2H, CH iVal, γ -CH₂ a Pro), 1.83-1.75 (m, 2H, β -CH₂ b Pro, γ -CH₂ b Pro), 1.69-1.62 (m, 1H, β -CH Ile), 1.29-1.22 (m, 1H, γ -CH₂ a Ile), 0.97-0.89 (m, 1H, γ -CH₂ b Ile), 0.87 (d, 3H, *J* = 6.8 Hz, γ -CH₃ a Val), 0.85 (d, 3H, *J* = 6.4 Hz, γ -CH₃ b Val), 0.81 (d, 3H, *J* = 6.6 Hz, CH₃ a iVal), 0.78 (d, 3H, *J* = 6.6 Hz, CH₃ b iVal), 0.74 (t, 3H, *J* = 7.4 Hz, δ -CH₃ Ile), 0.71 (d, 3H, *J* = 7.0 Hz, γ -CH₃ Ile).

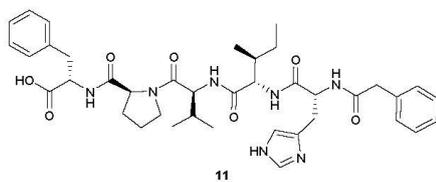
¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_{C} [ppm] = 172.7 (COOH Phe), 171.5 (CO iVal), 171.4 (CO Pro), 170.7 (CO Ile), 170.4 (CO His), 169.6 (CO Val), 137.3 (γ -C_{quart} Phe), 134.1 (ϵ -CH_{arom} His), 131.3 (weak, γ -C_{quart} His), 129.1, 128.1, 126.4 (δ -CH_{arom} Phe), 116.9 (weak, δ -CH_{arom} His), 59.0 (α -CH Pro), 56.2 (α -CH Ile), 55.8 (α -CH Val), 53.5 (α -CH Phe), 52.0 (α -CH His), 47.0 (δ -CH₂N Pro), 44.4 (CH₂ iVal), 36.9 (β -CH Ile), 36.7 (β -CH₂ Phe), 29.8 (β -CH Val), 28.9 (β -CH₂ Pro), 28.6 (β -CH₂ His), 25.4 (CH iVal), 24.3 (γ -CH₂ Pro), 24.0 (γ -CH₂ a Ile), 22.2 (CH₃ a iVal), 22.1 (CH₃ b iVal), 19.0 (γ -CH₃ a Val), 18.4 (γ -CH₃ b Val), 15.2 (γ -CH₃ Ile), 11.0 (δ -CH₃ Ile).

Additional found signals: δ_{H} [ppm] = 8.14 (FA), Broad peak between 4.00 to 2.60 (H₂O), 3.10, 2.54 (DMSO), 1.18. ϵ -NHHis and COOH₂Phe were not observed. δ_{C} [ppm] = 163.0 (FA), 157.7, 45.7, 8.60.

UHR-MS (ESI-TOF) m/z calcd for C₃₆H₅₄N₇O₇: 696.4079 [M+H]⁺; found: 696.4082 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{21} = -34.0^{\circ}$ (c = 2.35, CH₃OH)

(2-phenylacetyl)-D-histidyl-L-isoleucyl-L-valyl-L-prolyl-L-phenylalanine (**11**)



7 (7.994 g, 2.166 mmol) were swelled in DMF for 30 min. After removal of the solvent, HATU (2.425 g, 6.378 mmol) was added, dissolved in a small amount of DMF, followed by phenylacetic acid (0.897 g, 6.588 mmol), DIPEA (2300 μ L,

13.53 mmol) and additional DMF. The mixture was agitated for 2 h, after which LC-MS indicated completeness of the reaction. The supernatant was drained and the resin was washed with DMF (3x), DMC (2x) and MeOH (2x). The cleavage was done as described in the general method section. The obtained crude was washed four times with ice cold diethyl ether, decanted and dried under reduced pressure. Purification of 40% of the crude using semi preparative HPLC (5-50-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min) yielded **11** as a white wax like film (182 mg, 0.249 mmol, overall yield calculated to be: 29%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 8.35 (d, 1H, J = 8.1 Hz, NHHis), 8.20 (s, 1H, ϵ -CH_{arom} His), 7.99 (d, 1H, J = 7.5 Hz, NH₂Phe), 7.98 (d, 1H, J = 8.0 Hz, NHVal), 7.82 (d, 1H, J = 8.6 Hz, NH₂Le), 7.28-7.13 (m, 10H, δ -CH_{arom} Phe, CH_{arom} PA), 7.02 (s, 1H, δ -CH_{arom} His), 4.66 (ddd, 1H, J = 8.2, 8.2, 6.4 Hz, α -CHHis), 4.39 (ddd, 1H, J = 7.8, 7.8, 5.9 Hz, α -CH₂Phe), 4.37 (dd, 1H, J = 8.7, 4.1 Hz, α -CH₂Pro), 4.27 (dd, 1H, J = 8.3, 8.3 Hz, α -CH₂Val), 4.23 (dd, 1H, J = 8.8, 7.0 Hz, α -CH₂Le), 3.78-3.72 (m, 1H, δ -CH₂N a Pro), 3.57-3.51 (m, 1H, δ -CH₂N b Pro), 3.45 (d, 1H, J = 14.2 Hz, CH₂ a PA), 3.42 (d, 1H, J = 14.2 Hz, CH₂ b PA), 3.00 (dd, 1H, J = 13.8, 5.5 Hz, β -CH₂ a Phe), 2.98 (dd, 1H, J = 14.5, 5.9 Hz, β -CH₂ a His), 2.92 (dd, 1H, J = 13.9, 7.9 Hz, β -CH₂ b Phe), 2.82 (dd, 1H, J = 14.9, 8.3 Hz, β -CH₂ b His), 2.00-1.92 (m, 2H, β -CH₂Val, β -CH₂ a Pro), 1.90-1.84 (m, 1H, γ -CH₂ a Pro), 1.84-1.75 (m, 2H, β -CH₂ b Pro, γ -CH₂ b Pro), 1.68-1.61 (m, 1H,

S27

β -CH(Ile), 1.26-1.18 (m, 1H, γ -CH₂ a Ile), 0.94-0.82 (m, 1H, γ -CH₂ b Ile), 0.88 (d, 3H, J = 6.4 Hz, γ -CH₃ a Val), 0.86 (d, 3H, J = 6.6 Hz, γ -CH₃ b Val), 0.73 (t, 3H, J = 7.4 Hz, δ -CH₃ Ile), 0.70 (d, 3H, J = 7.0 Hz, γ -CH₃ Ile).

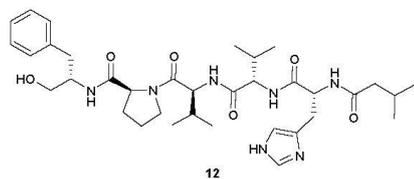
¹³C-NMR (DMSO-*d*₆, 151 MHz): δ _C [ppm] = 172.7 (COOH Phe), 171.4 (CO Pro), 170.6 (CO Ile), 170.2 (CO His), 170.1 (CO PA), 169.7 (CO Val), 137.3 (γ -C_{quart} Phe), 136.1 (C_{quart} PA), 134.1 (ϵ -CH_{arom} His), 131.3* (γ -C_{quart} His), 129.1, 128.9, 128.1, 128.0, 126.4, 126.2 (δ -CH_{arom} Phe, CH_{arom} PA), 116.8 (δ -CH_{arom} His), 59.0 (α -CH Pro), 56.3 (α -CH Ile), 55.8 (α -CH Val), 53.5 (α -CH Phe), 52.2 (α -CH His), 47.0 (δ -CH₂N Pro), 42.0 (CH₂ PA), 36.8 (β -CH Ile), 36.7 (β -CH₂ Phe), 29.7 (β -CH Val), 28.9 (β -CH₂ Pro), 28.8 (β -CH₂ His), 24.3 (γ -CH₂ Pro), 23.9 (γ -CH₂ a Ile), 19.0 (γ -CH₃ a Val), 18.5 (γ -CH₃ b Val), 15.2 (γ -CH₃ Ile), 11.0 (δ -CH₃ Ile).

Additional found signals: δ _H [ppm] = 8.14 (FA), Broad peak between 4.00 to 2.32 (H₂O), 2.54 (DMSO). ϵ -NHHis and COOH Phe were not observed. δ _C [ppm] = 163.0 (FA).

UHR-MS (ESI-TOF) m/z calcd for C₃₉H₅₂N₇O₇: 730.3923 [M+H]⁺; found: 730.3924 [M+H]⁺

Specific rotation [α]_D²¹ = -27.6° (c = 1.45, CH₃OH)

(*S*)-*N*-((*S*)-1-hydroxy-3-phenylpropan-2-yl)-1-((3-methylbutanoyl)-D-histidyl-L-valyl-L-valyl)pyrrolidine-2-carboxamide (**12**)



Peptide **8** (87.5 mg, 0.128 mmol) was dissolved in anhydrous MeOH (20 mL) and thionyl chloride (14.0 μ L, 0.192 mmol) was added at 0 °C. The mixture was stirred under a gentle reflux for 3.5 h and then concentrated *in vacuo*. The crude was

suspended in anhydrous THF and cooled to 0 °C. LiBH₄ (10.5 mg, 0.482 mmol) was added and the reaction was stirred at room temperature. The reaction progress was monitored using LC-MS and after 2 h more LiBH₄ (6.0 mg, 0.28 mmol) was added at 0 °C. After a total of 3 h the reaction was quenched using sat. NH₄Cl solution. Water was added until all formed salts were dissolved. The layers were separated and the aqueous layer was extracted with ethyl acetate three times. The combined organic phases were washed once with brine, dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure. Purification of the crude using semi preparative HPLC (5-

40-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min yielded **12** as a colorless solid (63.9 mg, 0.0957 mmol, 75% over 2 steps).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 7.98 (d, 1H, *J* = 7.9 Hz, *NH*Val 1), 7.98 (d, 1H, *J* = 7.9 Hz, *NH*His), 7.66 (d, 1H, *J* = 9.2 Hz, *NH*Val 2), 7.51 (s, 1H, ϵ -*CH*_{arom} His), 7.51 (d, 1H, *J* = 6.1 Hz, *NH*Phe), 7.27-7.14 (m, 5H, δ -*CH*_{arom} Phe), 6.77 (s, 1H, δ -*CH*_{arom} His), 4.59 (ddd, 1H, *J* = 7.9, 7.9, 6.4 Hz, α -*CH*His), 4.30-4.25(m, 2H, α -*CH*Val 1, α -*CH*Pro), 4.23 (dd, 1H, *J* = 8.2, 6.8 Hz, α -*CH*Val 2), 3.87-3.80 (m, 1H, α -*CH*Phe), 3.76-7.70 (m, 1H, δ -*CH*₂N a Pro), 3.57-3.51 (m, 1H, δ -*CH*₂N b Pro), 3.27 (t, 2H, *J* = 5.2 Hz, *CH*₂OH Phe), 2.90 (dd, 1H, *J* = 14.9, 5.9 Hz, β -*CH*₂ a His), 2.79 (dd, 1H, *J* = 13.7, 6.5 Hz, β -*CH*₂ a Phe), 2.73 (dd, 1H, *J* = 14.8, 8.9 Hz, β -*CH*₂ b His), 2.68 (dd, 1H, *J* = 13.7, 7.1 Hz, β -*CH*₂ b Phe), 2.03-1.97 (m, 1H, β -*CH*Val 1), 1.96 (d, 2H, *J* = 6.6 Hz, *CH*₂ iVal), 1.97-1.84 (m, 4H, β -*CH*Val 2, β -*CH*₂ a Pro, *CH*iVal, γ -*CH*₂ a Pro), 1.81-1.76 (m, 1H, γ -*CH*₂ b Pro), 1.76-1.71 (m, 1H, β -*CH*₂ b Pro), 0.90 (d, 3H, *J* = 6.8 Hz, γ -*CH*₃ a Val 1), 0.88 (d, 3H, *J* = 6.6 Hz, γ -*CH*₃ b Val 1), 0.81 (d, 3H, *J* = 6.4 Hz, *CH*₃ a iVal), 0.78 (d, 3H, *J* = 6.2 Hz, *CH*₃ b iVal), 0.73 (d, 3H, *J* = 6.8 Hz, γ -*CH*₃ a Val 2), 0.70 (d, 3H, *J* = 6.8 Hz, γ -*CH*₃ b Val 2).

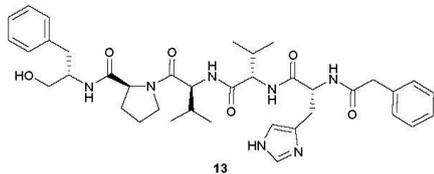
¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_{C} [ppm] = 171.4 (*CO* iVal), 171.1 (*CO* Pro), 171.1 (*CO* His), 170.7 (*CO* Val 2), 169.7 (*CO* Val 1), 138.9 (γ -*C*_{quant} Phe), 134.5 (weak, ϵ -*CH*_{arom} His), 129.1, 128.0, 125.8 (δ -*CH*_{arom} Phe), 61.8 (*CH*₂OH Phe), 59.5 (α -*CH* Pro), 57.0 (α -*CH* Val 2), 55.9 (α -*CH* Val 1), 52.7 (α -*CH* His), 52.2 (α -*CH* Phe), 47.0 (δ -*CH*₂N Pro), 44.5 (*CH*₂ iVal), 36.3 (β -*CH*₂ Phe), 30.6 (β -*CH* Val 2), 29.7 (β -*CH* Val 1), 29.8* (β -*CH*₂ a His), 29.1 (β -*CH*₂ Pro), 25.4 (*CH* iVal), 24.4 (γ -*CH*₂ Pro), 22.2 (*CH*₃ a iVal), 22.1 (*CH*₃ b iVal), 19.0 (γ -*CH*₃ a Val 2), 19.1 (γ -*CH*₃ a Val 1), 18.4 (γ -*CH*₃ b Val 1), 17.6 (γ -*CH*₃ b Val 2).

Additional found signals: δ_{H} [ppm] = 8.15 (FA), Broad peak between 4.00 to 2.58 (*H*₂O), 2.54 (DMSO). ϵ -*NH*His was not observed. δ_{C} [ppm] = δ -*CH*_{arom} His and γ -*C*_{quant} His were not observed.

UHR-MS (ESI-TOF) *m/z* calcd for C₃₅H₅₄N₇O₆: 668.4130 [M+H]⁺; found: 668.4126 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{21} = -28.7^{\circ}$ (*c* = 0.87, *CH*₃OH)

(*S*)-*N*-((*S*)-1-hydroxy-3-phenylpropan-2-yl)-1-((2-phenylacetyl)-D-histidyl-L-valyl-L-valyl)pyrrolidine-2-carboxamide (**13**)



To a solution of peptide **9** (118.1 mg, 0.1650 mmol) in anhydrous MeOH (40 mL) was added thionyl chloride (18.0 μ L, 0.247 mmol) at 0 °C. The mixture was stirred under a gentle reflux for 3.5 h and then concentrated *in vacuo*. The crude

product (107.3 mg, 0.1470 mmol, 89%) was directly used in the next step. For testing the conditions the crude was split into two parts, the syntheses were carried out identical.

The first part of the crude methyl ester (49.5 mg, 0.0678 mmol) was suspended in anhydrous THF and cooled to 0 °C. LiBH₄ (5.8 mg, 0.27 mmol) was added and the reaction was stirred at room temperature. The reaction progress was monitored using LC-MS and after 1 h more LiBH₄ (3.1 mg, 0.14 mmol) was added at 0 °C. After a total of 3 h the reaction was quenched using sat. NH₄Cl solution. Water was added until all formed salts were dissolved. The layers were separated and the aqueous layer was extracted with ethyl acetate three times. The combined organic phases were washed once with brine, dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure.

The second part of the crude methyl ester (68.5 mg, 0.0939 mmol) was suspended in anhydrous THF and cooled to 0 °C. LiBH₄ (8.3 mg, 0.38 mmol) was added and the reaction was stirred at room temperature. The reaction progress was monitored using LC-MS and after 1.5 h more LiBH₄ (4.1 mg, 0.19 mmol) was added at 0 °C. After a total of 2 h the reaction was quenched using sat. NH₄Cl solution. Water was added until all formed salts were dissolved. The layers were separated and the aqueous layer was extracted with ethyl acetate three times. The combined organic phases were washed once with brine, dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure.

Both crudes were combined for purification using semi preparative HPLC (5-40-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min), which yielded **13** as a colorless solid (72.8 mg, 0.104 mmol, 63% over 2 steps).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 8.29 (d, 1H, *J* = 7.9 Hz, *NH*His), 7.96 (d, 1H, *J* = 8.1 Hz, *NH*Val 1), 7.75 (d, 1H, *J* = 9.7 Hz, *NH*Val 2), 7.57 (s, 1H, ϵ -CH_{arom} His), 7.51 (d, 1H, *J*

= 8.4 Hz, *NHPhe*), 7.27-7.14 (m, 10H, δ -*CH*_{arom} Phe, *CH*_{arom} PA), 6.79 (s, 1H, δ -*CH*_{arom} His), 4.60 (ddd, 1H, *J* = 7.9, 7.9, 6.6 Hz, α -*CH*His), 4.29-4.25 (m, 2H, α -*CH*Pro, α -*CH*Val 1), 4.22 (dd, 1H, *J* = 9.0, 6.2 Hz, α -*CH*Val 2), 3.87-3.80 (m, 1H, α -*CH*Phe), 3.76-3.70 (m, 1H, δ -*CH*₂N a Pro), 3.57-3.52 (m, 1H, δ -*CH*₂N b Pro), 3.46 (d 1H, *J* = 14.1 Hz, *CH*₂ a PA), 3.42 (d 1H, *J* = 14.3 Hz, *CH*₂ b PA), 3.28 (t, 2H, *J* = 5.0 Hz, *CH*₂OH Phe), 2.92 (dd, 1H, *J* = 14.7, 6.2 Hz, β -*CH*₂ a His), 2.79 (dd, 1H, *J* = 13.4, 6.2 Hz, β -*CH*₂ a Phe), 2.77 (dd, 1H, *J* = 14.2, 8.7 Hz, β -*CH*₂ b His), 2.68 (dd, 1H, *J* = 13.8, 7.1 Hz, β -*CH*₂ b Phe), 2.03-1.95 (m, 1H, β -*CH*Val 1), 1.95-1.84 (m, 3H, γ -*CH*₂ a Pro, β -*CH*₂ a Pro, β -*CH*Val 2), 1.80-1.71 (m, 2H, γ -*CH*₂ b Pro, β -*CH*₂ b Pro), 0.90 (d, 3H, *J* = 6.8 Hz, γ -*CH*₃ a Val 1), 0.88 (d, 3H, *J* = 6.8 Hz, γ -*CH*₃ b Val 1), 0.71 (d, 3H, *J* = 7.0 Hz, γ -*CH*₃ a Val 2), 0.67 (d, 3H, *J* = 6.8 Hz, γ -*CH*₃ b Val2).

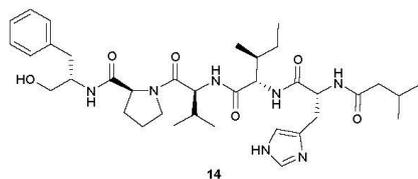
¹³C-NMR (DMSO-*d*₆, 151 MHz): δ _C [ppm] = 171.1 (*CO* Pro), 170.9 (*CO* His), 170.7 (*CO* Val 2), 170.0 (*CO* PA), 169.8 (*CO* Val 1), 138.9 (γ -*C*_{quart} Phe), 136.2 (*C*_{quart} PA), 134.5 (ϵ -*CH*_{arom} His), 132.9* (γ -*C*_{quart} His), 129.1, 129.0, 128.0, 126.2, 125.8 (δ -*CH*_{arom} Phe, *CH*_{arom} PA), 117.0* (δ -*CH*_{arom} His), 61.8 (*CH*₂OH Phe), 59.5 (α -*CH* Pro), 57.1 (α -*CH* Val 2), 56.0 (α -*CH* Val 1), 52.9 (α -*CH* His), 52.2 (α -*CH* Phe), 47.0 (δ -*CH*₂N Pro), 42.0 (*CH*₂ PA), 36.3 (β -*CH*₂ Phe), 30.5 (β -*CH* Val 2), 29.9 (β -*CH*₂ His), 29.7 (β -*CH* Val 1), 29.1 (β -*CH*₂ Pro), 24.4 (γ -*CH*₂ a Pro), 19.1 (γ -*CH*₃ a Val 1), 19.0 (γ -*CH*₃ a Val 2), 18.5 (γ -*CH*₃ b Val 1), 17.6 (γ -*CH*₃ b Val2).

Additional found signals: δ _H [ppm] = 8.15 (FA), 3.66, 3.17 (MeOH), 2.54 (DMSO). ϵ -*NH*His was not observed. δ _C [ppm] = 163.1 (FA), 48.6 (MeOH), 40.4 (DMSO).

UHR-MS (ESI-TOF) *m/z* calcd for C₃₈H₅₂N₇O₆: 702.3974 [M+H]⁺; found: 702.3975[M+H]⁺

Specific rotation [α]_D²¹ = -56.1° (c = 1.07, CH₃OH)

(*S*)-*N*-((*S*)-1-hydroxy-3-phenylpropan-2-yl)-1-((3-methylbutanoyl)-D-histidyl-L-isoleucyl-L-valyl)pyrrolidine-2-carboxamide (**14**)



The peptide **10** (93.8 mg, 0.135 mmol) was dissolved in anhydrous MeOH (20 mL) and thionyl chloride (14.8 μ L, 0.203 mmol) was added at 0 °C. The mixture was stirred under a gentle reflux for 3 h and then concentrated *in vacuo*. The crude

product was suspended in anhydrous THF and cooled to 0 °C. LiBH₄ (12.6 mg, 0.579 mmol) was added and the reaction was stirred at room temperature. The reaction progress was monitored using LC-MS and after 1.5 h more LiBH₄ (6.4 mg, 0.29 mmol) was added at 0 °C. After additional 4 h more LiBH₄ (5.0 mg, 0.23 mmol) was added at 0 °C. After a total of 6.5 h the reaction was quenched using sat. NH₄Cl solution. Water was added until all formed salts were dissolved. The layers were separated and the aqueous layer was extracted with ethyl acetate three times. The combined organic phases were washed once with brine, dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure. Purification of the crude product using semi preparative HPLC (5-45-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 µm, 250 x 10 mm, flow rate: 2 mL/min) yielded **14** as a colorless solid (22.9 mg, 0.0336 mmol, 25% over 2 steps).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_H [ppm] = 7.98 (d, 1H, *J* = 8.3 Hz, *NH*Val), 7.98 (d, 1H, *J* = 8.3 Hz, *NH*His), 7.68 (d, 1H, *J* = 9.4 Hz, *NH*Ile), 7.57 (s, 1H, ε-CH_{arom} His), 7.51 (d, 1H, *J* = 8.3 Hz, *NH*Phe), 7.28-7.14 (m, 5H, δ-CH_{arom} Phe), 6.79 (s, 1H, δ-CH_{arom} His), 4.59 (ddd, 1H, *J* = 8.0, 8.0, 6.5 Hz, α-*CH*His), 4.28 (dd, 1H, *J* = 7.6, 7.6 Hz, α-*CH*Val), 4.26 (dd, 1H, *J* = 8.0, 4.7 Hz, α-*CH*Pro), 4.23 (dd, 1H, *J* = 8.8, 7.0 Hz, α-*CH*Ile), 3.87-3.80 (m, 1H, α-*CH*Phe), 3.74-3.69 (m, 1H, δ-CH₂N a Pro), 3.57-3.51 (m, 1H, δ-CH₂N b Pro), 3.27 (t, 2H, *J* = 5.4 Hz, CH₂OH Phe), 2.89 (dd, 1H, *J* = 14.8, 6.0 Hz, β-CH₂ a His), 2.79 (dd, 1H, *J* = 13.7, 6.5 Hz, β-CH₂ a Phe), 2.73 (dd, 1H, *J* = 14.8, 8.5 Hz, β-CH₂ b His), 2.68 (dd, 1H, *J* = 13.9, 7.1 Hz, β-CH₂ b Phe), 2.03-1.97 (m, 1H, β-*CH*Val), 1.95 (d, 2H, *J* = 6.8 Hz, CH₂ iVal), 1.94-1.84 (m, 3H, β-CH₂ a Pro, γ-CH₂ a Pro, *CH*iVal), 1.80-1.71 (m, 2H, β-CH₂ b Pro, γ-CH₂ b Pro), 1.68-1.62 (m, 1H, β-*CH*Ile), 1.31-1.23 (m, 1H, γ-CH₂ a Ile), 0.97-0.91 (m, 1H, γ-CH₂ b Ile), 0.89 (d, 3H, *J* = 6.8 Hz, γ-CH₃ a Val), 0.87 (d, 3H, *J* = 6.8 Hz, γ-CH₃ b Val), 0.81 (d, 3H, *J* = 6.4 Hz, CH₃ a iVal), 0.78 (d, 3H, *J* = 6.6 Hz, CH₃ b iVal), 0.74 (t, 3H, *J* = 7.5 Hz, δ-CH₃ Ile), 0.72 (d, 3H, *J* = 6.8 Hz, γ-CH₃ Ile).

¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_C [ppm] = 171.4 (CO iVal), 171.0 (CO Pro), 170.9 (CO His), 170.8 (CO Ile), 169.7 (CO Val), 138.9 (γ-C_{quart} Phe), 134.5 (ε-CH_{arom} His), 133.0* (γ-C_{quart} His), 129.1, 128.0, 125.8 (δ-CH_{arom} Phe), 116.9* (δ-CH_{arom} His), 61.8 (CH₂OH Phe), 59.5 (α-*CH*Pro), 56.3 (α-*CH*Ile), 55.9 (α-*CH*Val), 52.6 (α-*CH*His), 52.2 (α-*CH*Phe), 47.0 (δ-CH₂N Pro), 44.5 (CH₂ iVal), 36.8 (β-*CH*Ile), 36.3 (β-CH₂ a Phe), 29.8 (β-*CH*Val), 29.7 (β-CH₂ a His), 29.1 (β-CH₂ Pro), 25.5 (CH iVal), 24.4 (γ-CH₂ Pro), 24.0 (γ-CH₂ a Ile), 22.2 (CH₃ a iVal), 22.1 (CH₃ b iVal), 19.1 (γ-CH₃ a Val), 18.4 (γ-CH₃ b Val), 15.3 (γ-CH₃ Ile), 11.0 (δ-CH₃ Ile).

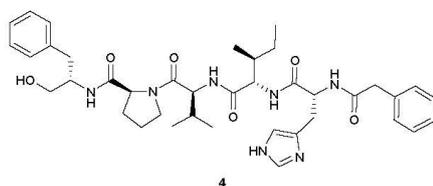
S32

Additional found signals: δ_{H} [ppm] = 8.15 (FA), 3.17 (MeOH), 2.54 (DMSO). ϵ -*NH*His was not observed. δ_{C} [ppm] = /

UHR-MS (ESI-TOF) m/z calcd for $\text{C}_{36}\text{H}_{56}\text{N}_7\text{O}_6$: 682.4287 [M+H]⁺; found: 682.4290 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{21} = -41.7^\circ$ ($c = 1.20$, CH_3OH)

(*S*)-*N*-((*S*)-1-hydroxy-3-phenylpropan-2-yl)-1-((2-phenylacetyl)-D-histidyl-L-isoleucyl-L-valyl)pyrrolidine-2-carboxamide (**4**)



To a solution of peptide **11** (160.0 mg, 0.2192 mmol) in anhydrous MeOH (20 mL) was added thionyl chloride (24.0 μL , 0.329 mmol) at 0 °C. The mixture was stirred under a gentle reflux for 3 h and then concentrated *in vacuo*. The crude was suspended in anhydrous THF and cooled to

0 °C. LiBH_4 (19.4 mg, 0.891 mmol) was added and the reaction was stirred at room temperature. The reaction progress was monitored using LC-MS and after 2.5 h more LiBH_4 (9.3 mg, 0.43 mmol) was added at 0 °C. After a total of 5.5 h the reaction was quenched using sat. NH_4Cl solution. Water was added until all formed salts were dissolved. The layers were separated and the aqueous layer was extracted with ethyl acetate three times. The combined organic phases were washed once with brine, dried over MgSO_4 , filtered and the solvent was evaporated under reduced pressure. Purification of the crude product using semi preparative HPLC (5-40-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μm , 250 x 10 mm, flow rate: 2 mL/min) yielded **4** as a colorless solid (45.1 mg, 0.0630 mmol, 25% over 2 steps).

$^1\text{H-NMR}$ ($\text{DMSO}-d_6$, 600 MHz): δ_{H} [ppm] = 8.27 (d, 1H, $J = 7.9$ Hz, *NH*His), 7.95 (d, 1H, $J = 8.3$ Hz, *NH*Val), 7.75 (d, 1H, $J = 8.8$ Hz, *NH*Ile), 7.53 (s, 1H, ϵ - CH_{arom} His), 7.50 (d, 1H, $J = 8.4$ Hz, *NH*Phe), 7.27-7.14 (m, 10H, δ - CH_{arom} Phe, CH_{arom} PA), 6.77 (s, 1H, δ - CH_{arom} His), 4.59 (ddd, 1H, $J = 7.9, 7.9, 6.6$ Hz, α -*CH*His), 4.27 (dd, 1H, $J = 8.1, 8.1$ Hz, α -*CH*Val), 4.27 (dd, 1H, $J = 8.1, 3.9$ Hz, α -*CH*Pro), 4.21 (dd, 1H, $J = 8.6, 7.2$ Hz, α -*CH*Ile), 3.87-3.80 (m, 1H, α -*CH*Phe), 3.74-3.69 (m, 1H, δ - CH_2N a Pro), 3.57-3.51 (m, 1H, δ - CH_2N b Pro), 3.45 (d, 1H, $J = 13.9$ Hz, CH_2 a PA), 3.42 (d, 1H, $J = 14.1$ Hz, CH_2 b PA), 3.27 (t, 2H, $J = 5.4$ Hz, CH_2OH Phe), 2.90 (dd, 1H, J

= 14.6, 6.1 Hz, β -CH₂ a His), 2.79 (dd, 1H, J = 13.5, 6.3 Hz, β -CH₂ a Phe), 2.75 (dd, 1H, J = 14.4, 7.8 Hz, β -CH₂ b His), 2.68 (dd, 1H, J = 13.8, 7.2 Hz, β -CH₂ b Phe), 2.02-1.95 (m, 1H, β -CHVal), 1.94-1.89 (m, 1H, β -CH₂ a Pro), 1.89-1.82 (m, 1H, γ -CH₂ a Pro), 1.80-1.75 (m, 1H, γ -CH₂ b Pro), 1.75-1.70 (m, 1H, β -CH₂ b Pro), 1.68-1.61 (m, 1H, β -CHIle), 1.27-1.19 (m, 1H, γ -CH₂ a Ile), 0.93-0.81 (m, 1H, γ -CH₂ b Ile), 0.89 (d, 3H, J = 6.8 Hz, γ -CH₃ a Val), 0.87 (d, 3H, J = 6.8 Hz, γ -CH₃ b Val), 0.72 (t, 3H, J = 7.4 Hz, δ -CH₃ Ile), 0.69 (d, 3H, J = 7.2 Hz, γ -CH₃ Ile).

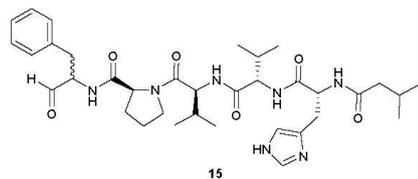
¹³C-NMR (DMSO-*d*₆, 151 MHz): δ _C [ppm] = 171.0 (CO Pro), 170.82 (CO Ile), 170.75 (CO His), 169.9 (CO PA), 169.7 (CO Val), 138.9 (γ -C_{quart} Phe), 136.2 (C_{quart} PA), 134.5 (ϵ -CH_{arom} His), 129.1, 129.0, 128.0, 126.2, 125.8 (δ -CH_{arom} Phe, CH_{arom} PA), 61.8 (CH₂OH Phe), 59.5 (α -CH Pro), 56.4 (α -CH Ile), 55.9 (α -CH Val), 52.9 (α -CH His), 52.2 (α -CH Phe), 47.0 (δ -CH₂N Pro), 42.0 (CH₂ PA), 36.7 (β -CH Ile), 36.3 (β -CH₂ Phe), 30.1 (weak, β -CH₂ His), 29.7 (β -CH Val), 29.1 (β -CH₂ Pro), 24.4 (γ -CH₂ Pro), 23.9 (γ -CH₂ Ile), 19.1 (γ -CH₃ a Val), 18.5 (γ -CH₃ b Val), 15.3 (γ -CH₃ Ile), 11.0 (δ -CH₃ Ile).

Additional found signals: δ _H [ppm] = 8.15 (FA), 3.17 (MeOH), 2.54 (DMSO). ϵ -NHHis was not observed. δ _C [ppm] = 40.4 (DMSO).

UHR-MS (ESI-TOF) m/z calcd for C₃₉H₅₄N₇O₆: 716.4130 [M+H]⁺; found: 716.4121 [M+H]⁺

Specific rotation [α]_D²¹ = -102.2° (c = 2.03, CH₃OH)

(2*R*/*S*)-1-((3-methylbutanoyl)-D-histidyl-L-valyl-L-valyl)-*N*-((*S*)-1-oxo-3-phenylpropan-2-yl)pyrrolidine-2-carboxamide (**15**)



The alcohol **12** (57.9 mg, 0.0867 mmol) was suspended in anhydrous DCM and DMP (15% in DCM, 540 μ L, 0.260 mmol) was added. The mixture was stirred at room temperature and the reaction progress was monitored by UHR-MS. After 2 h

more DMP was added (90.0 μ L, 0.130 mmol). The reaction was quenched with MeOH after 3 h and concentrated *in vacuo*. Purification of the crude product using semi preparative HPLC (5-35-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate:

2 mL/min) yielded **15** as a colorless solid (100% pure: 19.3 mg, 0.0289 mmol; 95% pure: 32.8 mg, 0.0469 mmol, calculated total yield: 87%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 9.47 (s, 0.4H, CHO Phe), 9.41 (s, 0.2H, CHO Phe), 8.36 (d, 0.4H, *J* = 7.5 Hz, NH Phe), 8.34 (d, 0.2H, *J* = 7.2 Hz, NH Phe), 8.01-7.94 (m, 2H, NH Val 1, NH His), 7.68-7.61 (m, 1H, NH Val 2), 7.51 (s, 1H, ϵ -CH_{arom} His), 7.30-7.10 (m, 5H, CH_{arom} Phe), 6.77 (s, 1H, δ -CH_{arom} His), 4.62-4.56 (m, 1H, α -CH His), 4.33-4.25 (m, 2 H, α -CH Phe, α -CH Pro, α -CH Val 1), 4.25-4.15 (m, 2 H, α -CH Phe, α -CH Val 2), 3.78-3.72 (m, 1H, δ -CH₂N a Pro), 3.57-3.47 (m, 1H, δ -CH₂N b Pro), 3.15-3.07 (m, 1H, β -CH₂ Phe), 2.93-2.82 (m, 2H, β -CH₂ a Phe, β -CH₂ a His), 2.78-2.69 (m, 2H, β -CH₂ b Phe, β -CH₂ b His), 2.04-1.84 (m, 7H, β -CH₂ Pro, γ -CH₂ Pro, β -CH Val 1, β -CH Val 2, CH₂ iVal, CH iVal), 1.84-1.71 (m, 2H, β -CH₂ Pro, γ -CH₂ Pro), 1.58-1.52 (m, 1H, β -CH₂ Pro, γ -CH₂ Pro), 0.89 (d, 3H, *J* = 6.8 Hz, γ -CH₃ a Val 1), 0.87 (d, 3H, *J* = 6.6 Hz, γ -CH₃ b Val 1), 0.82 (d, 3H, *J* = 6.4 Hz, CH₃ a iVal), 0.78 (d, 3H, *J* = 6.2 Hz, CH₃ b iVal), 0.75-0.68 (m, 6H, γ -CH₃ a+b Val 2).

¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_{C} [ppm] = 200.3 (CHO Phe), 172.1, 172.0 (CO Pro), 171.4 (CO iVal), 171.0 (CO His), 170.6 (CO Val 2), 169.62, 169.59 (CO Val 1), 137.62, 137.60 (γ -C_{quart} Phe), 134.5 (weak, ϵ -CH_{arom} His), 129.3, 129.2, 128.8, 128.2, 128.1, 127.7, 126.23, 126.18 (δ -CH_{arom} Phe), 59.7, 59.5 (α -CH Phe), 59.1, 59.0 (α -CH Pro), 56.9 (α -CH Val 2), 55.8 (α -CH Val 1), 52.7 (α -CH His), 47.0 (δ -CH₂N Pro), 44.4 (CH₂ iVal), 33.5, 33.3 (β -CH₂ Phe), 30.65, 30.60 (β -CH Val 2), 29.8 (β -CH₂ His), 29.7 (β -CH Val 1), 29.3 (β -CH₂ Pro), 25.4 (CH iVal), 24.5, 24.4 (γ -CH₂ Pro), 22.2 (CH₃ a iVal), 22.1 (CH₃ b iVal), 19.0 (γ -CH₃ a Val 1, γ -CH₃ a Val 2), 18.44, 18.40 (γ -CH₃ b Val 1), 17.6 (γ -CH₃ b Val 2).

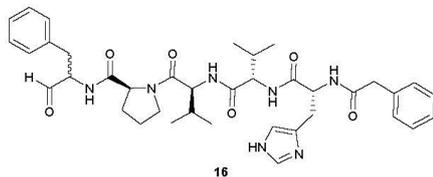
Additional found signals: δ_{H} [ppm] = 8.18 (FA), 3.17 (MeOH). ϵ -NH His was not observed. δ_{C} [ppm] = /, γ -C_{quart} His, δ -CH_{arom} His were not observed.

Due to racemization of the phenylalanine aldehyde, some chemical shifts appear as a double set.

UHR-MS (ESI-TOF) *m/z* calcd for C₃₅H₅₂N₇O₆: 666.3974 [M+H]⁺; found: 666.3976 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{21} = -37.5^{\circ}$ (c = 0.80, CH₃OH)

(2*R*/*S*)-*N*-((*S*)-1-oxo-3-phenylpropan-2-yl)-1-((2-phenylacetyl)-D-histidyl-L-valyl-L-valyl)pyrrolidine-2-carboxamide (**16**)



The alcohol **13** (14.8 mg, 0.0211 mmol) was suspended in anhydrous DCM and DMP (15% in DCM, 131 μ L, 0.0632 mmol) was added. The mixture was stirred at room temperature and the reaction progress was monitored by UHR-MS.

The reaction was quenched with MeOH after 2 h and concentrated *in vacuo*. Purification of the crude product using semi preparative HPLC (5-45-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min) yielded **16** as a colorless solid (12.7 mg, 0.0181 mmol, 86%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 9.46 (s, 0.2H, CHO Phe), 9.41 (s, 0.3H, CHO Phe), 8.36 (d, 0.3H, *J* = 7.5 Hz, NH Phe), 8.34 (d, 0.3H, *J* = 6.8 Hz, NH Phe), 8.28 (d, 1H, *J* = 8.1 Hz, NH His), 7.95 (dd, 1H, *J* = 8.4, 8.4 Hz, NH Val 1), 7.76-7.71 (m, 1H, NH Val 2), 7.50 (s, 1H, ϵ -CH_{arom} His), 7.30-7.12 (m, 10H, δ -CH_{arom} Phe, CH_{arom} PA), 6.76 (s, 1H, δ -CH_{arom} His), 4.62-4.56 (m, 1H, α -CH His), 4.35-4.23 (m, 2 H, α -CH Pro, α -CH Val 1, α -CH Phe), 4.23-4.15 (m, 2 H, α -CH Phe, α -CH Val 2), 3.80-3.70 (m, 2H, δ -CH₂N a Pro), 3.59-3.47 (m, 4H, δ -CH₂N b Pro), 3.46 (d, 2H, *J* = 13.9 Hz, CH₂ a PA), 3.42 (d, 2H, *J* = 14.1 Hz, CH₂ b PA), 3.15-3.06 (m, 1H, β -CH₂ Phe), 2.94-2.82 (m, 2H, β -CH₂ His, β -CH₂ Phe), 2.79-2.71 (m, 2H, β -CH₂ His, β -CH₂ Phe), 2.04-1.84 (m, 4H, β -CH Val 1, β -CH Val 2, γ -CH₂ a Pro, β -CH₂ a Pro), 1.84-1.69 (m, 2H, γ -CH₂ b Pro, β -CH₂ b Pro), 1.58-1.50 (m, 1H, γ -CH₂ b Pro, β -CH₂ b Pro), 0.94-0.80 (m, 6H, γ -CH₃ Val 1), 0.75-0.63 (m, 6H, γ -CH₃ Val 2).

¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_{C} [ppm] = 200.3 (CHO Phe), 172.1, 172.0, 170.9, 170.7, 169.7, 169.6 (CO_{carbonyl}), 169.9 (CO PA), 137.6 (γ -C_{quart} Phe), 136.2 (C_{quart} PA), 134.5 (ϵ -CH_{arom} His), 129.3, 129.2, 129.0, 128.6, 128.2, 128.1, 128.0, 126.25, 126.20, 126.17 (CH_{arom} Phe, CH_{arom} PA), 59.7, 59.5 (α -CH Phe), 59.1, 59.0 (α -CH Pro), 57.1 (α -CH Val 2), 55.8 (α -CH Val 1), 52.9 (α -CH His), 47.0 (δ -CH₂N Pro), 42.0 (CH₂ PA), 33.5, 33.3 (β -CH₂ Phe), 30.0* (β -CH₂ His), 30.6, 30.5 (β -CH Val 2), 29.7 (β -CH Val 1), 29.3 (β -CH₂ Pro), 24.5, 24.4 (γ -CH₂ Pro), 19.0 (γ -CH₃ a Val 2), 18.49 (γ -CH₃ a Val 1), 18.46 (γ -CH₃ b Val 1), 17.6 (γ -CH₃ b Val 2).

S36

Additional found signals: δ_{H} [ppm] = 8.19 (FA), Broad peak between 4.00 to 2.59 (H₂O), 3.17 (MeOH), ϵ -NHHis was not observed. δ_{C} [ppm] = 163.4 (FA), γ -C_{quat} His, δ -CH_{arom} His were not observed.

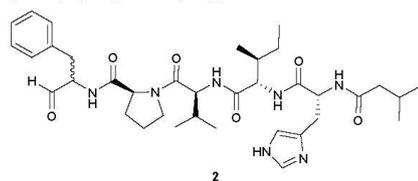
Due to racemization of the phenylalanine aldehyde, some chemical shifts appear as a double set and some integrals aren't accurate due to the broad peak of δ_{H} = 4.00 to 2.59 ppm.

The carbonyl signals 172.1, 172.0, 170.9, 170.7, 169.7, 169.6 were not distinguishable using HMBC data.

UHR-MS (ESI-TOF) *m/z* calcd for C₃₈H₅₀N₇O₆: 700.3817 [M+H]⁺; found: 700.3820 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{21} = -66.7^{\circ}$ (c = 0.75, CH₃OH)

(2*R/S*)-1-((3-methylbutanoyl)-D-histidyl-L-isoleucyl-L-valyl)-*N*-((*S*)-1-oxo-3-phenylpropan-2-yl)pyrrolidine-2-carboxamide (**2**)



Alcohol **14** (19.7 mg, 0.0289 mmol) was suspended in anhydrous DCM and DMP (15% in DCM, 180 μ L, 0.0867 mmol) was added. The mixture was stirred at room temperature and the reaction progress was monitored by UHR-MS. The reaction was

quenched with MeOH after 2 h and concentrated *in vacuo*. Purification of the crude product using semi preparative HPLC (5-45-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min) yielded **2** as a colorless solid (12.1 mg, 0.0178 mmol, 62%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 9.47 (s, 0.2H, CHO Phe), 9.41 (s, 0.2H, CHO Phe), 8.36 (d, 0.2H, *J* = 7.7 Hz, NH Phe), 8.34 (d, 0.2H, *J* = 7.0 Hz, NH Phe), 8.00-7.94 (m, 2H, NH Val, NH His), 7.69-7.63 (m, 1H, NH Ile), 7.50 (s, 1H, ϵ -CH_{arom} His), 7.29-7.12 (m, 5H, δ -CH_{arom} Phe), 6.76 (s, 1H, δ -CH_{arom} His), 4.61-4.55 (m, 1H, α -CH His), 4.35-4.24 (m, 2H, α -CH Phe, α -CH Val, α -CH Pro), 4.24-4.16 (m, 2H, α -CH Phe, α -CH Ile), 3.78-3.68 (m, 1H, δ -CH₂N a Pro), 3.57-3.47 (m, 1H, δ -CH₂N b Pro), 3.15-3.07 (m, 1H, β -CH₂ Phe), 2.91-2.83 (m, 2H, β -CH₂ Phe, β -CH₂ a His), 2.78-2.69 (m, 2H, β -CH₂ Phe, β -CH₂ b His), 2.02-1.83 (m, 5H, β -CH Val, CH₂ i Val, CH i Val, β -CH₂ a Pro, γ -CH₂ a Pro), 1.83-1.70 (m, 2H, β -CH₂ b Pro, γ -CH₂ b Pro), 1.70-1.60 (m,

2H, β -CH Ile), 1.60-1.51 (m, 1H, β -CH₂ Pro), 1.30-1.22 (m, 1H, γ -CH₂ a Ile), 0.98-0.90 (m, 1H, γ -CH₂ b Ile), 0.90-0.84 (m, 6H, γ -CH₃ Val), 0.81 (d, 3H, J = 6.4 Hz, CH₃ a iVal), 0.78 (d, 3H, J = 6.4 Hz, CH₃ b iVal), 0.76-0.67 (m, 6H, δ -CH₃ Ile, γ -CH₃ Ile).

¹³C-NMR (DMSO-*d*₆, 151 MHz): δ _C [ppm] = 200.3 (CHO Phe), 172.1, 172.0 (CO Pro), 171.4 (CO iVal), 170.9 (CO His), 170.7 (CO Ile), 169.58, 169.55 (CO Val), 137.6 (γ -C_{quart} Phe), 134.5 (ϵ -CH_{arom} His), 134.2 (γ -C_{quart} His), 129.3, 129.2, 129.1, 128.9, 128.2, 128.1, 126.24, 126.19 (δ -CH_{arom} Phe), 59.7, 59.5 (α -CH Phe), 59.1, 59.0 (α -CH Pro), 56.32, 56.29 (α -CH Ile), 55.7 (α -CH Val), 52.7 (α -CH His), 47.00, 46.96 (δ -CH₂N Pro), 44.5, (CH₂ iVal), 36.9, 36.8 (β -CH Ile), 33.6, 33.3 (β -CH₂ Phe), 29.9* (β -CH₂ His), 29.7 (β -CH Val), 29.31, 29.28 (β -CH₂ Pro), 25.5 (CH iVal), 24.44, 24.38 (γ -CH₂ Pro), 24.0 (γ -CH₂ Ile), 22.2 (CH₃ a iVal), 22.1 (CH₃ b iVal), 19.03, 19.00 (γ -CH₃ a Val), 18.45, 18.41 (γ -CH₃ b Val), 15.2 (γ -CH₃ Ile), 11.0 (δ -CH₃ Ile).

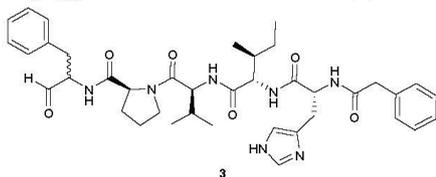
Additional found signals: δ _H [ppm] = 8.19 (FA), 3.17 (MeOH), ϵ -NH His was not observed. δ _C [ppm] = δ -CH_{arom} His was not observed.

Due to racemization of the phenylalanine aldehyde, some chemical shifts appear as a double set.

UHR-MS (ESI-TOF) m/z calcd for C₃₆H₅₄N₇O₆: 680.4130 [M+H]⁺; found: 680.4128 [M+H]⁺

Specific rotation $[\alpha]_D^{21} = -90.4^\circ$ ($c = 0.83$, CH₃OH)

(2*R*/*S*)-*N*-((*S*)-1-oxo-3-phenylpropan-2-yl)-1-((2-phenylacetyl)-D-histidyl-L-isoleucyl-L-valyl)pyrrolidine-2-carboxamide (**3**)



To a solution of alcohol **4** (36.9 mg, 0.0515 mmol) in anhydrous DCM DMP (15% in DCM, 321 μ L, 0.155 mmol) was added. The mixture was stirred at room temperature and the reaction progress was monitored by UHR-MS.

The reaction was quenched with MeOH after 2 h and concentrated *in vacuo*. Purification of the crude product using semi preparative HPLC (5-45-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 2 mL/min) yielded **3** as a colorless solid (24.5 mg, 0.0343 mmol, 67%).

¹H-NMR (DMSO-*d*₆, 600 MHz): δ_{H} [ppm] = 9.47 (s, 0.2H, CHO Phe), 9.41 (s, 0.2H, CHO Phe), 8.36 (d, 0.25H, *J* = 7.7 Hz, NH Phe), 8.34 (d, 0.25H, *J* = 7.7 Hz, NH Phe), 8.26 (d, 1H, *J* = 8.1 Hz, NH His), 7.95 (t, 1H, *J* = 9.0 Hz, NH Val), 7.77-7.71 (m, 1H, NH Ile), 7.51 (s, 1H, ϵ -CH_{arom} His), 7.29-7.13 (m, 10H, δ -CH_{arom} Phe, CH_{arom} PA), 6.76 (s, 1H, δ -CH_{arom} His), 4.62-4.55 (m, 1H, α -CH His), 4.33-4.24 (m, 2H, α -CH Val, α -CH Pro), 4.24-4.16 (m, 2H, α -CH Ile, α -CH Phe), 3.78-3.68 (m, 1H, δ -CH₂N a Pro), 3.57-3.47 (m, 1H, δ -CH₂N b Pro), 3.45 (d, 1H, *J* = 14.1 Hz, CH₂ a PA), 3.42 (d, 1H, *J* = 14.1 Hz, CH₂ b PA), 3.15-3.06 (m, 1H, β -CH₂ Phe), 2.93-2.83 (m, 2H, β -CH₂ His, β -CH₂ Phe), 2.78-2.71 (m, 1H, β -CH₂ His, β -CH₂ Phe), 2.02-1.84 (m, 2H, β -CH Val, β -CH₂ Pro, γ -CH₂ Pro), 1.84-1.69 (m, 2H, β -CH₂ Pro, γ -CH₂ Pro), 1.69-1.59 (m, 1H, β -CH Ile), 1.59-1.51 (m, 1H, β -CH₂ Pro), 1.27-1.17 (m, 1H, γ -CH₂ a Ile), 0.94-0.82 (m, 7H, γ -CH₂ b Ile, γ -CH₃ Val), 0.75-0.66 (m, 6H, δ -CH₃ Ile, γ -CH₃ Ile).

¹³C-NMR (DMSO-*d*₆, 151 MHz): δ_{C} [ppm] = 200.3 (CHO Phe), 172.1, 172.0 (CO Pro), 170.8 (CO His), 170.7 (CO Ile), 169.9 (CO PA), 169.61, 169.58 (CO Val), 137.6 (γ -C_{quat} Phe), 136.2 (C_{quat} PA), 134.5 (ϵ -CH_{arom} His), 129.3, 129.2, 129.14, 129.09, 129.0, 128.2, 128.1, 128.0, 126.25, 126.20, 126.16 (δ -CH_{arom} Phe, CH_{arom} PA), 59.5 (α -CH Pro), 59.1, 59.0 (α -CH Phe), 56.40, 56.37 (α -CH Ile), 55.8 (α -CH Val), 52.9 (α -CH His), 47.01, 46.97 (δ -CH₂N Pro), 42.0 (CH₂ PA), 36.8, 36.7 (β -CH Ile), 33.5, 33.3 (β -CH₂ Phe), 30.1* (β -CH₂ His), 29.7 (β -CH Val), 29.32, 29.29 (β -CH₂ Pro), 24.5, 24.4 (γ -CH₂ Pro), 23.9 (γ -CH₂ Ile), 19.03, 19.00 (γ -CH₃ a Val), 18.5, 18.4 (γ -CH₃ b Val), 15.3 (γ -CH₃ Ile),, 11.0 (δ -CH₃ Ile).

Additional found signals: δ_{H} [ppm] = 8.16 (FA), 3.17 (MeOH), ϵ -NH His was not observed. δ_{C} [ppm] = 59.7, 48.6 (MeOH), δ -CH_{arom} His and γ -C_{quat} His were not observed.

Due to racemization of the phenylalanine aldehyde, some chemical shifts appear as a double set.

UHR-MS (ESI-TOF) *m/z* calcd for C₃₉H₅₂N₇O₆: 714.3974 [M+H]⁺; found: 714.3974 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{21} = -74.6^{\circ}$ (*c* = 0.67, CH₃OH)

NMR spectra

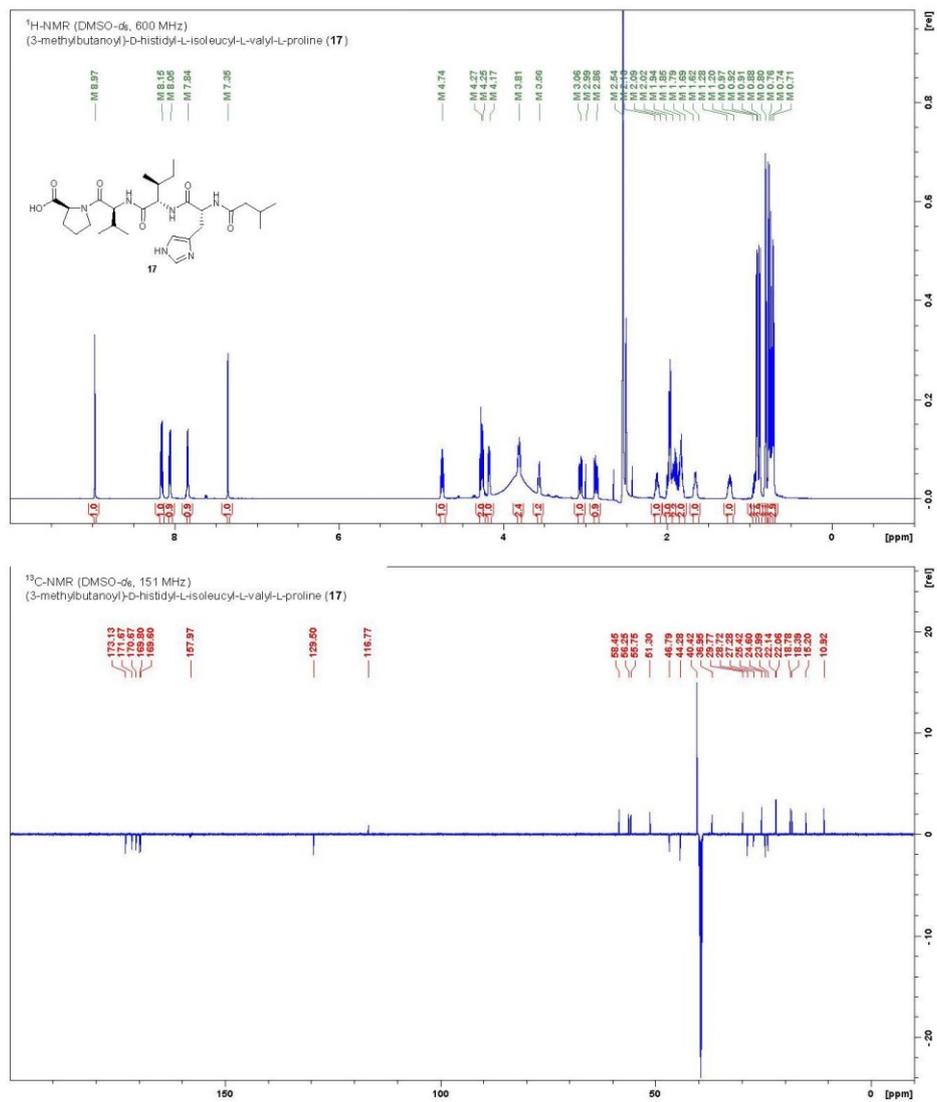


Figure S10. NMR spectra of **17** in DMSO-*d*₆.

Table S2. NMR data for **17** in DMSO- d_6 (600 MHz, 151 MHz).

Position	δ ^{13}C [ppm]	δ ^1H [ppm]; (m, J , J)
L-Proline		
COOH	173.1	n.o.
α -CH	58.5	4.17 (dd, 1H, J = 8.7, 5.0 Hz)
β -CH ₂	28.7	2.16-2.09 (m, 1H) 1.85-1.79 (m, 2H, overlay)
γ -CH ₂	24.6	1.94-1.85 (m, 2H, overlay) 1.85-1.79 (m, 2H, overlay)
δ -CH ₂ N	46.8	3.81 (ddd, 1H, J = 9.8, 6.4, 6.4 Hz) 3.56 (ddd, 1H, J = 10.1, 6.4, 6.4 Hz)
L-Valine		
CO	169.6	-
NH	-	8.05 (d, 1H, J = 8.3 Hz)
α -CH	55.7	4.27 (dd, 1H, J = 9.0, 9.0 Hz)
β -CH	29.8	2.02-1.94 (m, 3H, overlay)
γ -CH ₃	18.8	0.91 (d, 3H, J = 6.8 Hz)
γ -CH ₃	18.4	0.88 (d, 3H, J = 6.8 Hz)
L-Isoleucine		
CO	170.7	-
NH	-	7.84 (d, 1H, J = 8.8 Hz)
α -CH	56.2	4.25 (dd, 1H, J = 9.0, 7.0 Hz)
β -CH	36.9	1.69-1.62 (m, 1H)
γ -CH ₂	24.0	1.28-1.20 (m, 1H) 0.97-0.92 (m, 1H)
γ -CH ₃	10.9	0.74 (t, 3H, J = 7.4 Hz)
δ -CH ₃	15.2	0.71 (d, 3H, J = 6.7 Hz)
D-Histidine		
CO	169.8	-
NH	-	8.15 (d, 1H, J = 8.4 Hz)
α -CH	51.3	4.74 (ddd, 1H, J = 8.2, 8.2, 6.6 Hz)
β -CH ₂	27.3	3.06 (dd, 1H, J = 15.1, 6.1 Hz) 2.86 (dd, 1H, J = 15.1, 8.7 Hz)
γ -C _{quat}	129.5	-
δ -CH _{arom}	116.8	7.35 (s, 1H)
ϵ -CH _{arom}	133.7*	8.97 (d, 1H, J = 1.1 Hz)

ϵ-NH	-	n.o.
Isovaleroyl		
CO	171.7	-
CH₂	44.3	2.02-1.94 (m, 3H, overlay)
CH	25.4	1.94-1.85 (m, 2H, overlay)
CH₃	22.14	0.80 (d, 3H, $J = 6.8$ Hz)
CH₃	22.06	0.76 (d, 3H, $J = 6.4$ Hz)

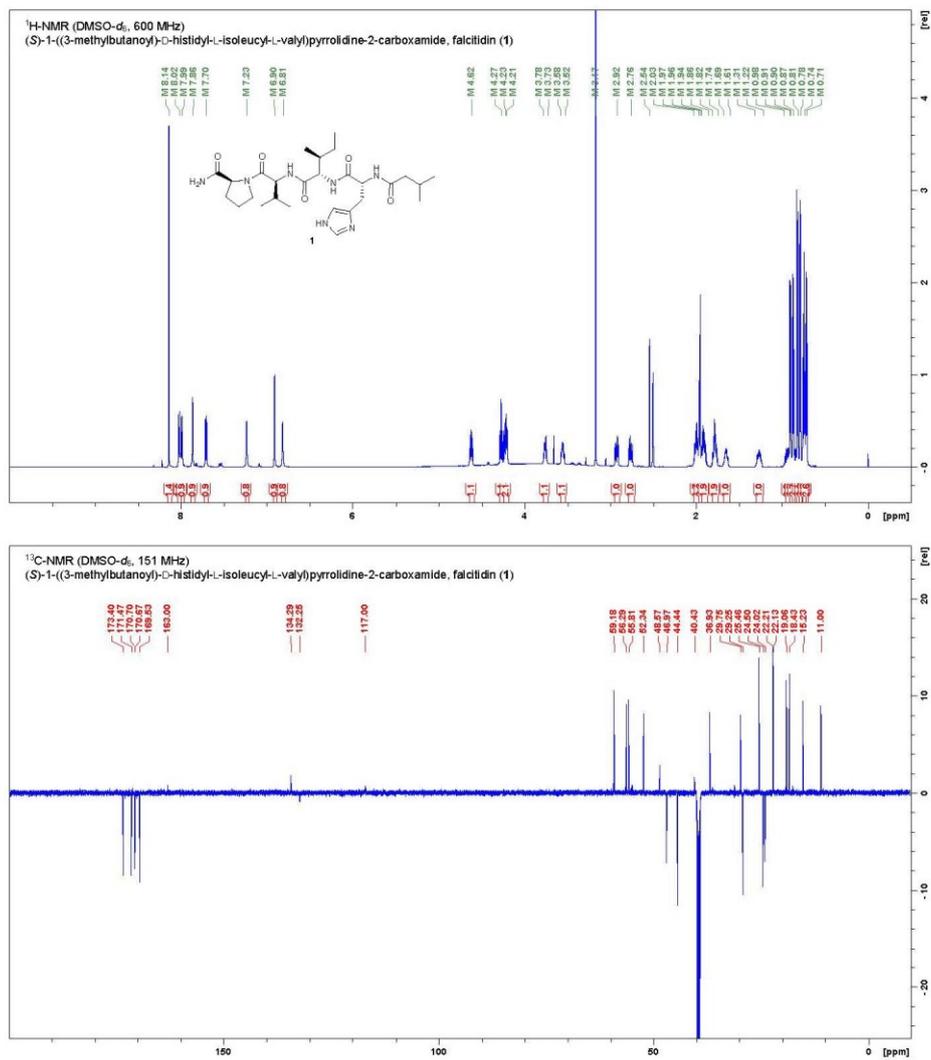


Figure S11. NMR spectra of 1 in DMSO-*d*₆.

Table S3. NMR data for **1** in DMSO-*d*₆ (600 MHz, 151 MHz).

Position	δ ¹³ C [ppm]	δ ¹ H [ppm]; (m, J, J)
L-Proline		
CONH ₂	173.4	7.23 (s, 1H), 6.81 (s, 1H)
α -CH	59.2	4.21 (dd, 1H, <i>J</i> = 8.5, 4.4 Hz)
β -CH ₂	29.3	2.03-1.97 (m, 2H, overlay) 1.82-1.74 (m, 2H, overlay)
γ -CH ₂	24.5	1.94-1.86 (m, 2H, overlay) 1.82-1.74 (m, 2H, overlay)
δ -CH ₂ N	47.0	3.78-3.73 (m, 1H) 3.58-3.52 (m, 1H)
L-Valine		
CO	169.5	-
NH	-	7.99 (d, 1H, <i>J</i> = 8.3 Hz)
α -CH	55.8	4.27 (dd, 1H, <i>J</i> =8.4, 8.4 Hz)
β -CH	29.7	2.03-1.97 (m, 2H, overlay)
γ -CH ₃	19.1	0.90 (d, 1H, <i>J</i> =6.8 Hz)
δ -CH ₃	18.4	0.87 (d, 1H, <i>J</i> =6.7 Hz)
L-Isoleucine		
CO	170.70	-
NH	-	7.70 (d, 1H, <i>J</i> =8.8 Hz)
α -CH	56.3	4.23 (dd, 1H, <i>J</i> =8.7, 6.9 Hz)
β -CH	36.9	1.69-1.61 (m, 1H)
γ -CH ₂	24.0	1.31-1.22 (m, 1H) 0.98-0.91 (m, 1H)
γ -CH ₃	11.0	0.71 (d, 1H, <i>J</i> =7.0 Hz)
δ -CH ₃	15.2	0.74 (t, 1H, <i>J</i> =7.4 Hz)
D-Histidine		
CO	170.67	-
NH	-	8.02 (d, <i>J</i> = 8.2 Hz)
α -CH	52.3	4.62 (ddd, <i>J</i> =8.3, 8.3, 6.1 Hz)
β -CH ₂	29.3	2.92 (dd, <i>J</i> =14.9, 5.9 Hz) 2.76 (dd, <i>J</i> =15.1, 8.8 Hz)
γ -C _{quat}	132.3, weak	-
δ -CH _{arom}	117.0, weak	6.90 (s, 1H)
ϵ -CH _{arom}	134.3, weak	7.86 (s, 1H)

ϵ-NH	-	n.o.
Isovaleroyl		
CO	171.5	-
CH₂	44.4	1.96 (d, $J=6.7$ Hz)
CH	25.5	1.94-1.86 (m, 2H, overlay)
CH₃	22.2	0.81 (d, $J=6.5$ Hz)
CH₃	22.1	0.78 (d, $J=6.5$ Hz)

S45

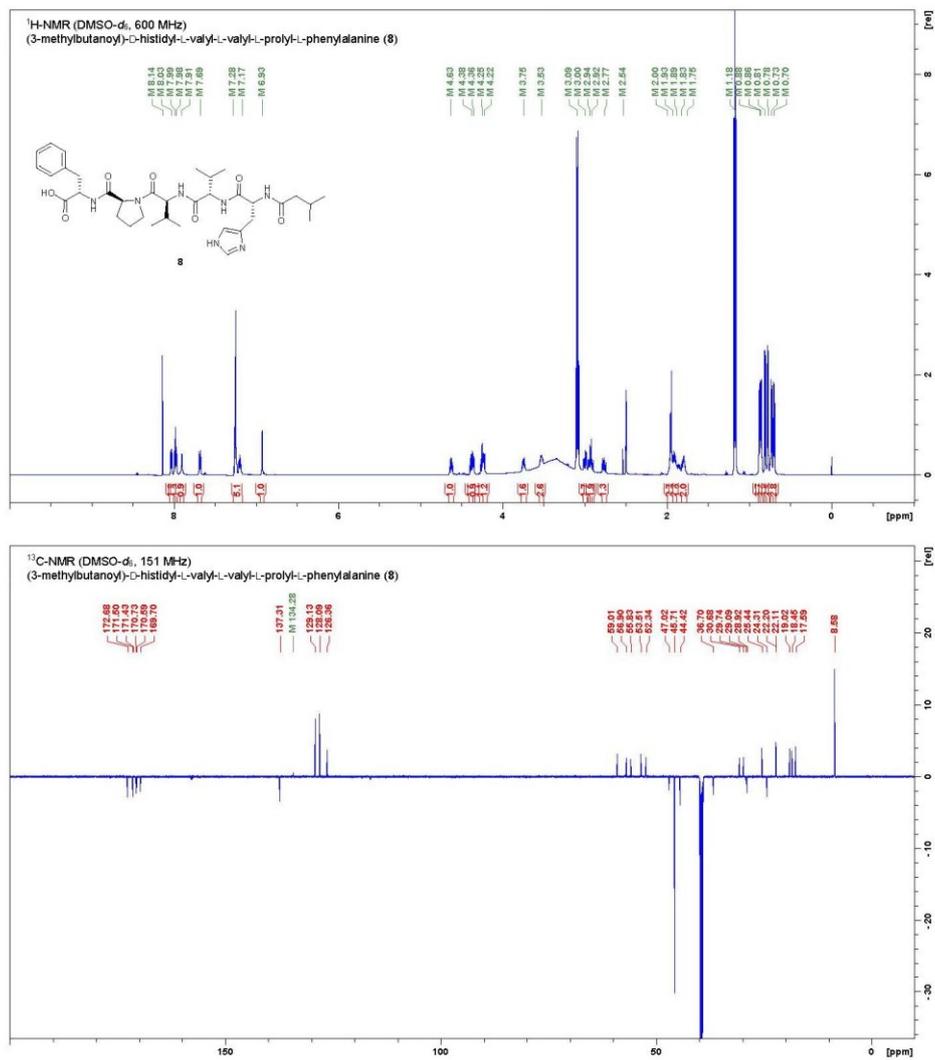


Table S4. NMR data for **8** in DMSO-*d*₆ (600 MHz, 151 MHz).

Position	δ ¹³ C [ppm]	δ ¹ H [ppm]; (m, J, J)
L-Phenylalanine		
COOH	172.7	n.o.
NH	-	7.98 (d, 1H, <i>J</i> = 7.2 Hz)
α -CH	53.5	4.38 (ddd, 1H, <i>J</i> = 7.6, 7.6, 5.7 Hz)
β -CH ₂	36.7	3.00 (dd, 1H, <i>J</i> = 14.2, 5.8 Hz) 2.92 (dd, 1H, <i>J</i> = 13.8, 7.5 Hz)
γ -C _{quart}	137.3	-
CH _{arom}	129.1, 128.1, 126.4	7.28-7.17 (m, 5H)
L-Proline		
CO	171.4	-
α -CH	59.0	4.36 (dd, 1H, <i>J</i> = 8.3, 3.7 Hz)
β -CH ₂	28.9	2.00-1.93 (m, 4H, overlay) 1.83-1.75 (m, 2H, overlay)
γ -CH ₂	24.3	1.89-1.83 (m, 1H) 1.83-1.75 (m, 2H, overlay)
δ -CH ₂ N	47.0	3.75 (ddd, 1H, <i>J</i> = 9.6, 6.9, 6.9 Hz) 3.53 (ddd, 1H, <i>J</i> = 9.4, 7.1, 5.6 Hz)
L-Valine 1		
CO	169.7	-
NH	-	7.99 (d, 1H, <i>J</i> = 7.9 Hz)
α -CH	55.8	4.25 (dd, 1H, <i>J</i> = 8.3, 8.3 Hz)
β -CH	29.7	2.00-1.93 (m, 4H, overlay)
γ -CH ₃	19.0	0.88 (d, 3H, <i>J</i> = 6.8 Hz)
γ -CH ₃	18.4	0.86 (d, 3H, <i>J</i> = 6.4 Hz)
L-Valine 2		
CO	170.6	-
NH	-	7.69 (d, 1H, <i>J</i> = 9.0 Hz)
α -CH	56.9	4.22 (dd, 1H, <i>J</i> = 8.9, 6.1 Hz)
β -CH	30.7	1.93-1.89 (m, 2H, overlay)
γ -CH ₃	19.0	0.73 (d, 3H, <i>J</i> = 6.6 Hz)
γ -CH ₃	17.6	0.70 (d, 3H, <i>J</i> = 7.0 Hz)
D-Histidine		
CO	170.7	-
NH	-	8.03 (d, 1H, <i>J</i> = 8.1 Hz)

S47

α -CH	52.3	4.63 (ddd, 1H, $J = 8.5, 8.5, 6.0$ Hz)
β -CH ₂	29.1, weak	2.94 (dd, 1H, $J = 14.6, 6.5$ Hz) 2.77 (dd, 1H, $J = 13.8, 7.5$ Hz)
γ -C _{quart}	132.2*	-
δ -CH _{arom}	117.0*	6.93 (s, 1H)
ϵ -CH _{arom}	134.3, weak	7.91 (bs, 1H)
ϵ -NH	-	n.o.
Isovaleroyl		
CO	171.5	-
CH ₂	44.4	2.00-1.93 (m, 4H, overlay)
CH	25.4	1.93-1.89 (m, 2H, overlay)
CH ₃	22.2	0.81 (d, 3H, $J = 6.4$ Hz)
CH ₃	22.1	0.78 (d, 3H, $J = 6.6$ Hz)

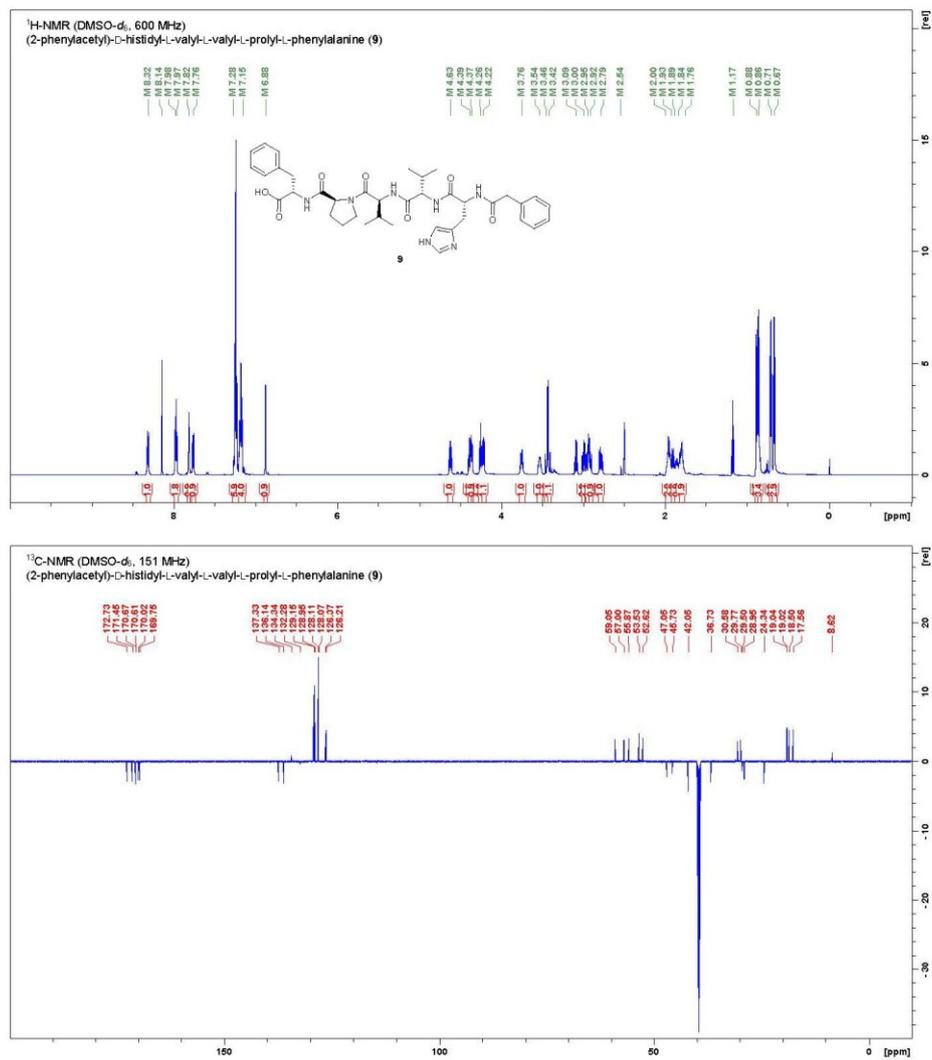


Figure S13. NMR spectra of 9 in DMSO-*d*₆.

Table S5. NMR data for **9** in DMSO-*d*₆ (600 MHz, 151 MHz).

Position	δ ¹³ C [ppm]	δ ¹ H [ppm]; (m, J, J)
L-Phenylalanine		
COOH	172.7	n.o.
NH	-	7.98 (d, 1H, <i>J</i> = 7.2 Hz)
α -CH	53.5	4.39 (ddd, 1H, <i>J</i> = 7.6, 7.6, 6.0 Hz)
β -CH ₂	36.7	3.00 (dd, 1H, <i>J</i> = 14.1, 5.9 Hz)
		2.92 (dd, 1H, <i>J</i> = 13.9, 7.7 Hz)
γ -C _{quart}	137.3	-
CH _{arom}	129.1, 128.9, 128.1, 128.1, 126.4, 126.2	7.28-7.15 (m, 10H, overlay)
L-Proline		
CO	171.5	-
α -CH	59.0	4.37 (dd, 1H, <i>J</i> = 8.5, 3.4 Hz)
β -CH ₂	28.9	2.00-1.93 (m, 2H, overlay)
		1.84-1.76 (m, 2H, overlay)
γ -CH ₂	24.3	1.89-1.84 (m, 1H)
		1.84-1.76 (m, 2H, overlay)
δ -CH ₂ N	47.0	3.76 (ddd, 1H, <i>J</i> = 9.8, 6.8, 6.8 Hz)
		3.54 (ddd, 1H, <i>J</i> = 9.4, 7.2, 5.4 Hz)
L-Valine 1		
CO	169.7	-
NH	-	7.97 (d, 1H, <i>J</i> = 7.5 Hz)
α -CH	55.9	4.26 (dd, 1H, <i>J</i> = 8.3, 8.3 Hz)
β -CH	29.8	2.00-1.93 (m, 2H, overlay)
γ -CH ₃	19.0	0.88 (d, 3H, <i>J</i> = 6.6 Hz)
γ -CH ₃	18.5	0.86 (d, 3H, <i>J</i> = 6.7 Hz)
L-Valine 2		
CO	170.6	-
NH	-	7.76 (d, 1H, <i>J</i> = 8.9 Hz)
α -CH	57.0	4.22 (dd, 1H, <i>J</i> = 9.0, 6.2 Hz)
β -CH	30.6	1.93-1.89 (m, 1H)
γ -CH ₃	19.0	0.71 (d, 3H, <i>J</i> = 6.8 Hz)
γ -CH ₃	17.6	0.67 (d, 3H, <i>J</i> = 6.7 Hz)
D-Histidine		
CO	170.7	-

NH	-	8.32 (d, 1H, $J = 8.3$ Hz)
α -CH	52.6	4.63 (ddd, 1H, $J = 7.7, 7.7, 6.6$ Hz)
β -CH ₂	29.5	2.79 (dd, 1H, $J = 14.8, 8.4$ Hz)
γ -C _{quart}	132.3, weak	-
δ -CH _{arom}	117.0*	6.88 (s, 1H)
ϵ -CH _{arom}	134.3	7.82 (s, 1H)
ϵ -NH	-	n.o.
Phenylacetyl		
CO	170.0	-
CH ₂	42.0	3.46 (d 1H, $J = 14.3$ Hz) 3.42 (d 1H, $J = 14.2$ Hz)
C _{quart}	136.1	
CH _{arom}	129.1, 128.9, 128.1, 128.1, 126.4, 126.2	7.28-7.15 (m, 10H, overlay)

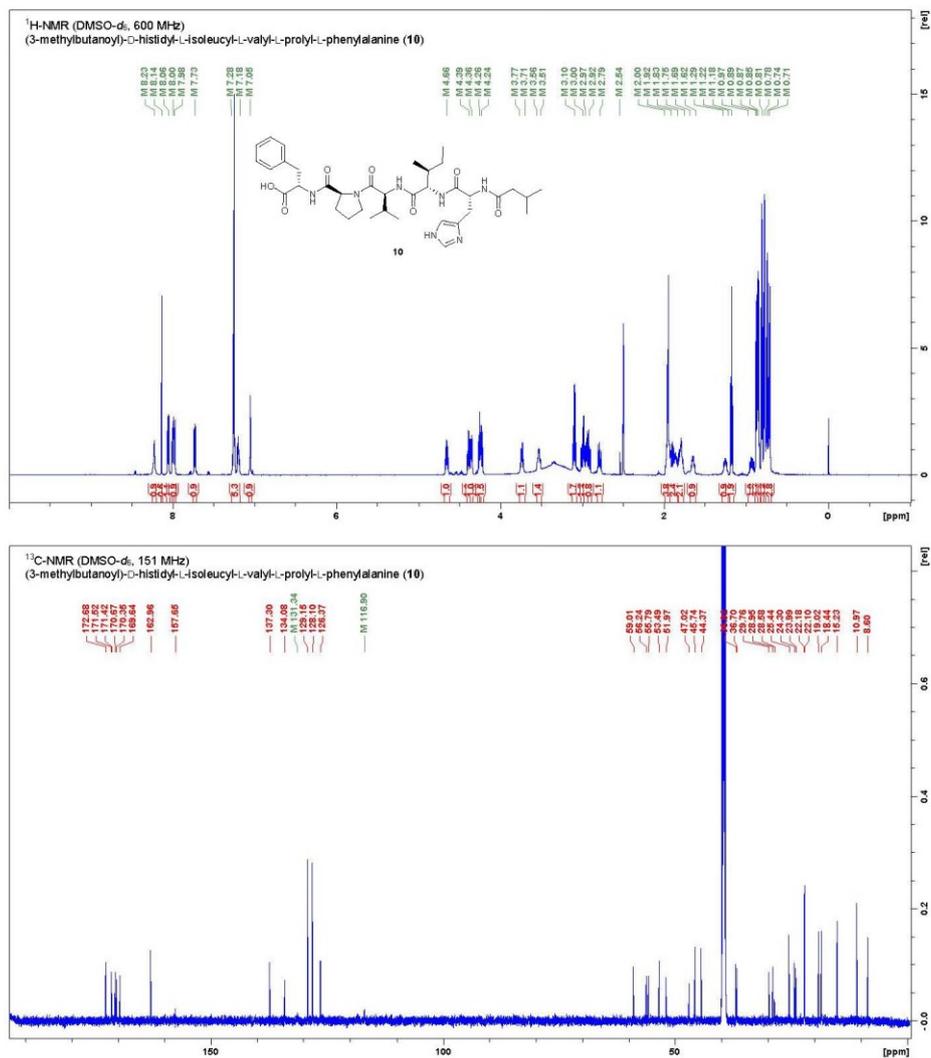


Figure S14. NMR spectra of 10 in DMSO-*d*₆.

Table S6. NMR data for **10** in DMSO- d_6 (600 MHz, 151 MHz).

Position	δ ^{13}C [ppm]	δ ^1H [ppm]; (m, \int , J)
L-Phenylalanine		
COOH	172.7	n.o.
NH	-	7.98 (d, 1H, J = 7.7 Hz)
α -CH	53.5	4.39 (ddd, 1H, J = 7.3, 7.3, 6.2 Hz)
β -CH ₂	36.7	3.00 (dd, 1H, J = 13.9, 5.9 Hz) 2.92 (dd, 1H, J = 13.9, 7.9 Hz)
γ -C _{quart}	137.3	-
CH _{arom}	129.1, 128.1, 126.4	7.28-7.18 (m, 5H)
L-Proline		
CO	171.4	-
α -CH	59.0	4.36 (dd, 1H, J = 8.3, 3.9 Hz)
β -CH ₂	28.9	2.00-1.92 (m, 4H, overlay) 1.83-1.75 (m, 2H, overlay)
γ -CH ₂	24.3	1.92-1.83 (m, 2H, overlay) 1.83-1.75 (m, 2H, overlay)
δ -CH ₂ N	47.0	3.77-3.71 (m, 1H) 3.56-3.51 (m, 1H)
L-Valine		
CO	169.6	-
NH	-	8.00 (d, 1H, J = 8.4 Hz)
α -CH	55.8	4.26 (dd, 1H, J = 8.3, 8.3 Hz)
β -CH	29.8	2.00-1.92 (m, 4H, overlay)
γ -CH ₃	19.0	0.87 (d, 3H, J = 6.8 Hz)
γ -CH ₃	18.4	0.85 (d, 3H, J = 6.4 Hz)
L-Isoleucine		
CO	170.7	-
NH	-	7.73 (d, 1H, J = 9.2 Hz)
α -CH	56.2	4.24 (dd, 1H, J = 8.5, 7.5 Hz)
β -CH	36.9	1.69-1.62 (m, 1H) 1.29-1.22 (m, 1H)
γ -CH ₂	24.0	0.97-0.89 (m, 1H)
γ -CH ₃	15.2	0.71 (d, 3H, J = 7.0 Hz)
δ -CH ₃	11.0	0.74 (t, 3H, J = 7.4 Hz)
D-Histidine		

CO	170.4	-
NH	-	8.06 (d, 1H, $J = 8.1$ Hz)
α -CH	52.0	4.66 (ddd, 1H, $J = 8.1, 8.1, 6.5$ Hz)
β -CH ₂	28.6	2.97 (dd, 1H, $J = 15.3, 6.3$ Hz) 2.79 (dd, 1H, $J = 15.0, 8.6$ Hz)
γ -C _{quat}	131.3	-
δ -CH _{arom}	116.9	7.05 (s, 1H)
ϵ -CH _{arom}	134.1	8.23 (s, 1H)
ϵ -NH His	-	n.o.
Isovaleroyl		
CO	171.5	-
CH ₂	44.4	2.00-1.92 (m, 4H, overlay)
CH	25.4	1.92-1.83 (m, 2H, overlay)
CH ₃	22.2	0.81 (d, 3H, $J = 6.6$ Hz)
CH ₃	22.1	0.78 (d, 3H, $J = 6.6$ Hz)

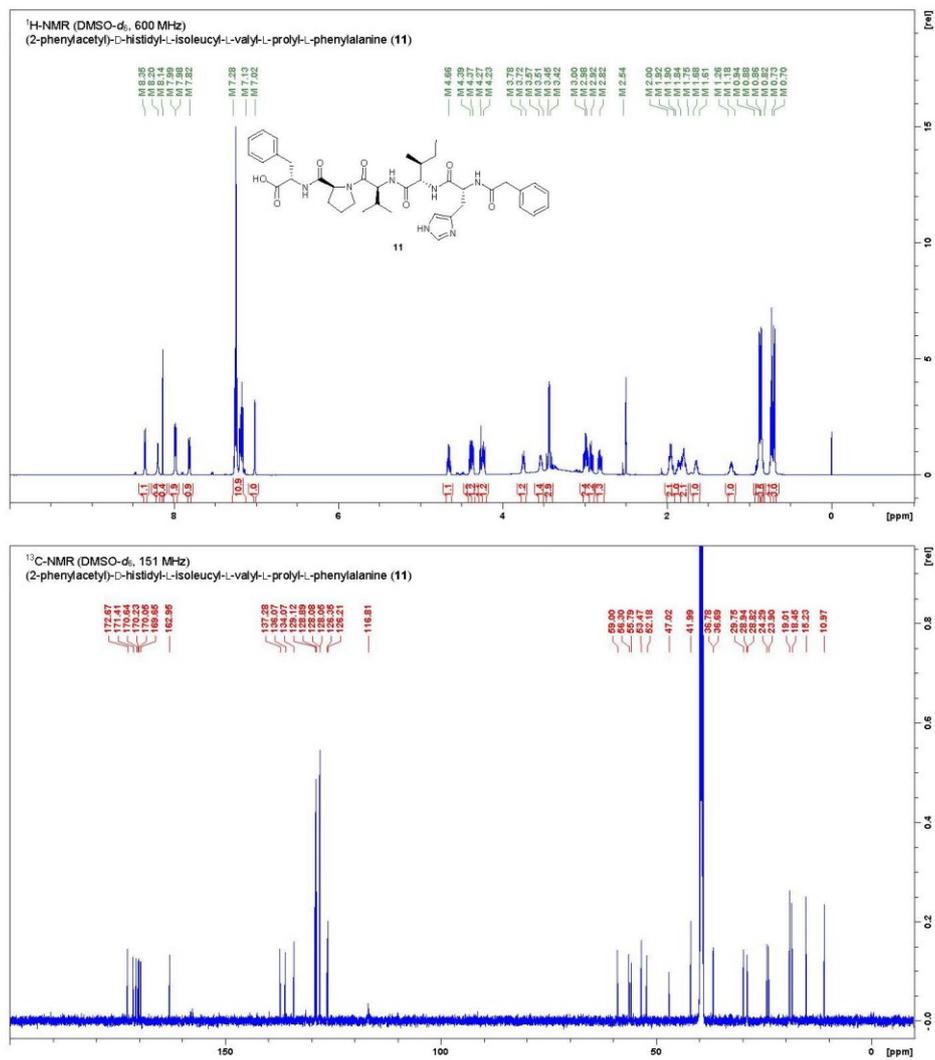


Figure S15. NMR spectra of 11 in DMSO-*d*₆.

Table S7. NMR data for **11** in DMSO- d_6 (600 MHz, 151 MHz).

Position	δ ^{13}C [ppm]	δ ^1H [ppm]; (m, J, J)
L-Phenylalanine		
COOH	172.7	n.o.
NH		7.99 (d, 1H, J = 7.5 Hz)
α -CH	53.5	4.39 (ddd, 1H, J = 7.8, 7.8, 5.9 Hz)
β -CH ₂	36.7	2.92 (dd, 1H, J = 13.9, 7.9 Hz)
γ -C _{quat}	137.3	-
CH _{anom}	129.1, 128.9, 128.1, 128.0, 126.4, 126.2	7.28-7.13 (m, 10H, overlay)
L-Proline		
CO	171.4	-
α -CH	59.0	4.37 (dd, 1H, J = 8.7, 4.1 Hz)
β -CH ₂	28.9	2.00-1.92 (m, 2H, overlay) 1.84-1.75 (m, 2H, overlay)
γ -CH ₂	24.3	1.90-1.84 (m, 1H) 1.84-1.75 (m, 2H, overlay)
δ -CH ₂ N	47.0	3.78-3.72 (m, 1H) 3.57-3.51 (m, 1H)
L-Valine		
CO	169.7	-
NH	-	7.98 (d, 1H, J = 8.0 Hz)
α -CH	55.8	4.27 (dd, 1H, J = 8.3, 8.3 Hz)
β -CH	29.7	2.00-1.92 (m, 2H, overlay)
γ -CH ₃	19.0	0.88 (d, 3H, J = 6.4 Hz)
γ -CH ₃	18.5	0.86 (d, 3H, J = 6.6 Hz)
L-Isoleucine		
CO	170.6	-
NH	-	7.82 (d, 1H, J = 8.6 Hz)
α -CH	56.3	4.23 (dd, 1H, J = 8.8, 7.0 Hz)
β -CH	36.8	1.68-1.61 (m, 1H)
γ -CH ₂	23.9	1.26-1.18 (m, 1H) 0.94-0.82 (m, 1H)
γ -CH ₃	15.2	0.70 (d, 3H, J = 7.0 Hz)
δ -CH ₃	11.0	0.73 (t, 3H, J = 7.4 Hz)
D-Histidine		

CO	170.2	-
NH	-	8.35 (d, 1H, $J = 8.1$ Hz)
α -CH	52.2	4.66 (ddd, 1H, $J = 8.2, 8.2, 6.4$ Hz)
β -CH ₂	28.8	2.98 (dd, 1H, $J = 14.5, 5.9$ Hz) 2.82 (dd, 1H, $J = 14.9, 8.3$ Hz)
γ -C _{quat}	131.3*	-
δ -CH _{arom}	116.8	7.02 (s, 1H)
ϵ -CH _{arom}	134.1	8.20 (s, 1H)
ϵ -NH		n.o.
Phenylacetyl		
CO	170.1	
CH ₂	42.0	3.45 (d, 1H, $J = 14.2$ Hz) 3.42 (d, 1H, $J = 14.2$ Hz)
C _{quat}	136.1	
CH _{arom}	129.1, 128.9, 128.1, 128.0, 126.4, 126.2	7.28-7.13 (m, 10H, overlay)

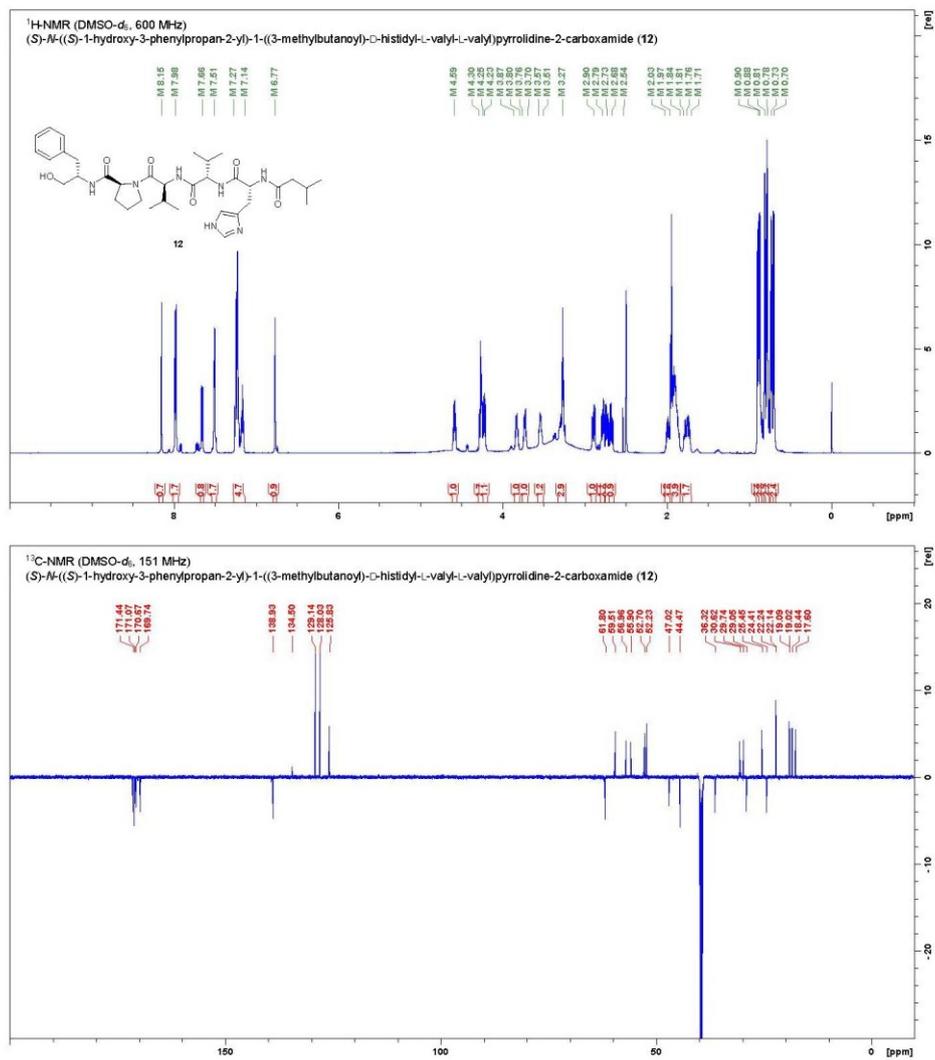


Figure S16. NMR spectra of 12 in DMSO-*d*₆.

Table S8. NMR data for **12** in DMSO- d_6 (600 MHz, 151 MHz).

Position	δ ^{13}C [ppm]	δ ^1H [ppm]; (m, \int , J)
L-Phenylalaninol		
CH ₂ OH	61.8	3.27 (t, 2H, J = 5.2 Hz)
NH	-	7.51 (d, 1H, J = 6.1 Hz)
α -CH	52.2	3.87-3.80 (m, 1H)
β -CH ₂	36.3	2.79 (dd, 1H, J = 13.7, 6.5 Hz) 2.68 (dd, 1H, J = 13.7, 7.1 Hz)
γ -C _{quat}	138.9	-
CH _{arom}	129.1, 128.0, 125.8	7.27-7.14 (m, 5H)
L-Proline		
CO	171.1	-
α -CH	59.5	4.30-4.25(m, 2H, overlay)
β -CH ₂	29.1	1.97-1.84 (m, 4H, overlay) 1.76-1.71 (m, 1H)
γ -CH ₂	24.4	1.97-1.84 (m, 4H, overlay) 1.81-1.76 (m, 1H)
δ -CH ₂ N	47.0	3.76-7.70 (m, 1H) 3.57-3.51 (m, 1H)
L-Valine 1		
CO	169.7	-
NH	-	7.98 (d, 1H, J = 7.9 Hz)
α -CH	55.9	4.30-4.25(m, 2H, overlay)
β -CH	29.7	2.03-1.97 (m, 1H)
γ -CH ₃	19.1	0.90 (d, 3H, J = 6.8 Hz)
γ -CH ₃	18.4	0.88 (d, 3H, J = 6.6 Hz)
L-Valine 2		
CO	170.7	-
NH	-	7.66 (d, 1H, J = 9.2 Hz)
α -CH	57.0	4.23 (dd, 1H, J = 8.2, 6.8 Hz)
β -CH	30.6	1.97-1.84 (m, 4H, overlay)
γ -CH ₃	19.0	0.73 (d, 3H, J = 6.8 Hz)
γ -CH ₃	17.6	0.70 (d, 3H, J = 6.8 Hz)
D-Histidine		
CO	171.1	-
NH	-	7.98 (d, 1H, J = 7.9 Hz)

α -CH	52.7	4.59 (ddd, 1H, $J = 7.9, 7.9, 6.4$ Hz)
β -CH ₂	29.8*	2.90 (dd, 1H, $J = 14.9, 5.9$ Hz) 2.73 (dd, 1H, $J = 14.8, 8.9$ Hz)
γ -C _{quart}	n.o.	-
δ -CH _{arom}	n.o.	6.77 (s, 1H)
ϵ -CH _{arom}	134.5, weak	7.51 (s, 1H)
ϵ -NH	-	n.o.
Isovaleroyl		
CO	171.4	-
CH ₂	44.5	1.96 (d, 2H, $J = 6.6$ Hz)
CH	25.4	1.97-1.84 (m, 4H, overlay)
CH ₃	22.2	0.81 (d, 3H, $J = 6.4$ Hz)
CH ₃	22.1	0.78 (d, 3H, $J = 6.2$ Hz)

Table S9. NMR data for **13** in DMSO- d_6 (600 MHz, 151 MHz).

Position	δ ^{13}C [ppm]	δ ^1H [ppm]; (m, \int , J)
L-Phenylalaninol		
CH ₂ OH	61.8	3.28 (t, 2H, J = 5.0 Hz)
NH	-	7.51 (d, 1H, J = 8.4 Hz)
α -CH	52.2	3.87-3.80 (m, 1H)
β -CH ₂	36.3	2.79 (dd, 1H, J = 13.4, 6.2 Hz) 2.68 (dd, 1H, J = 13.8, 7.1 Hz)
γ -C _{quat}	138.9	-
CH _{arom}	129.1, 129.0, 128.0, 126.2, 125.8	7.27-7.14 (m, 10H, overlay)
L-Proline		
CO	171.1	-
α -CH	59.5	4.29-4.25 (m, 2H, overlay)
β -CH ₂	29.1	1.95-1.84 (m, 3H, overlay) 1.80-1.71 (m, 2H, overlay)
γ -CH ₂	24.4	1.95-1.84 (m, 3H, overlay) 1.80-1.71 (m, 2H, overlay)
δ -CH ₂ N	47.0	3.76-3.70 (m, 1H) 3.57-3.52 (m, 1H)
L-Valine 1		
CO	169.8	-
NH	-	7.96 (d, 1H, J = 8.1 Hz)
α -CH	56.0	4.29-4.25 (m, 2H, overlay)
β -CH	29.7	2.03-1.95 (m, 1H)
γ -CH ₃	19.1	0.90 (d, 3H, J = 6.8 Hz)
γ -CH ₃	18.5	0.88 (d, 3H, J = 6.8 Hz)
L-Valine 2		
CO	170.7	-
NH	-	7.75 (d, 1H, J = 9.7 Hz)
α -CH	57.1	4.22 (dd, 1H, J = 9.0, 6.2 Hz)
β -CH	30.5	1.95-1.84 (m, 3H, overlay)
γ -CH ₃	19.0	0.71 (d, 3H, J = 7.0 Hz)
γ -CH ₃	17.6	0.67 (d, 3H, J = 6.8 Hz)
D-Histidine		
CO	170.9	-

NH	-	8.29 (d, 1H, $J = 7.9$ Hz)
α -CH	52.9	4.60 (ddd, 1H, $J = 7.9, 7.9, 6.6$ Hz)
β -CH ₂	29.9	2.92 (dd, 1H, $J = 14.7, 6.2$ Hz)
		2.77 (dd, 1H, $J = 14.2, 8.7$ Hz)
γ -C _{quat}	132.9*	-
δ -CH _{arom}	117.0*	6.79 (s, 1H)
ϵ -CH _{arom}	134.5	7.57 (s, 1H)
ϵ -NH	-	n.o.
Phenylacetyl		
CO	170.0	-
CH ₂	42.0	3.46 (d 1H, $J = 14.1$ Hz)
		3.42 (d 1H, $J = 14.3$ Hz)
C _{quat}	136.2	-
		129.1, 129.0, 128.0,
CH _{arom}	126.2, 125.8	7.27-7.14 (m, 10H, overlay)

Table S10. NMR data for **14** in DMSO-*d*₆ (600 MHz, 151 MHz).

Position	δ ¹³ C [ppm]	δ ¹ H [ppm]; (m, \int , <i>J</i>)
L-Phenylalaninol		
CH ₂ OH	61.8	3.27 (t, 2H, <i>J</i> = 5.4 Hz)
NH	-	7.51 (d, 1H, <i>J</i> = 8.3 Hz)
α -CH	52.2	3.87-3.80 (m, 1H)
β -CH ₂	36.3	2.79 (dd, 1H, <i>J</i> = 13.7, 6.5 Hz) 2.68 (dd, 1H, <i>J</i> = 13.9, 7.1 Hz)
γ -C _{quat}	138.9	-
CH _{arom}	129.1, 128.0, 125.8	7.28-7.14 (m, 5H)
L-Proline		
CO	171.0	-
α -CH	59.5	4.26 (dd, 1H, <i>J</i> = 8.0, 4.7 Hz)
β -CH ₂	29.1	1.94-1.84 (m, 3H, overlay) 1.80-1.71 (m, 2H, overlay)
γ -CH ₂	24.4	1.94-1.84 (m, 3H, overlay) 1.80-1.71 (m, 2H, overlay)
δ -CH ₂ N	47.0	3.74-3.69 (m, 1H) 3.57-3.51 (m, 1H)
L-Valine		
CO	169.7	-
NH	-	7.98 (d, 1H, <i>J</i> = 8.3 Hz)
α -CH	55.9	4.28 (dd, 1H, <i>J</i> = 7.6, 7.6 Hz)
β -CH	29.8	2.03-1.97 (m, 1H)
γ -CH ₃	19.1	0.89 (d, 3H, <i>J</i> = 6.8 Hz)
δ -CH ₃	18.4	0.87 (d, 3H, <i>J</i> = 6.8 Hz)
L-Isoleucine		
CO	170.8	-
NH	-	7.68 (d, 1H, <i>J</i> = 9.4 Hz)
α -CH	56.3	4.23 (dd, 1H, <i>J</i> = 8.8, 7.0 Hz)
β -CH	36.8	1.68-1.62 (m, 1H)
γ -CH ₂	24.0	1.31-1.23 (m, 1H) 0.97-0.91 (m, 1H)
δ -CH ₃	15.3	0.72 (d, 3H, <i>J</i> = 6.8 Hz)
ϵ -CH ₃	11.0	0.74 (t, 3H, <i>J</i> = 7.5 Hz)
D-Histidine		

CO	170.9	-
NH	-	7.98 (d, 1H, $J = 8.3$ Hz)
α -CH	52.6	4.59 (ddd, 1H, $J = 8.0, 8.0, 6.5$ Hz)
β -CH ₂	29.7	2.89 (dd, 1H, $J = 14.8, 6.0$ Hz) 2.73 (dd, 1H, $J = 14.8, 8.5$ Hz)
γ -C _{quat}	133.0*	-
δ -CH _{arom}	116.9*	6.79 (s, 1H)
ϵ -CH _{arom}	134.5	7.57 (s, 1H)
ϵ -NH	-	n.o.
Isovaleroyl		
CO	171.4	-
CH ₂	44.5	1.95 (d, 2H, $J = 6.8$ Hz)
CH	25.5	1.94-1.84 (m, 3H, overlay)
CH ₃	22.2	0.81 (d, 3H, $J = 6.4$ Hz)
CH ₃	22.1	0.78 (d, 3H, $J = 6.6$ Hz)

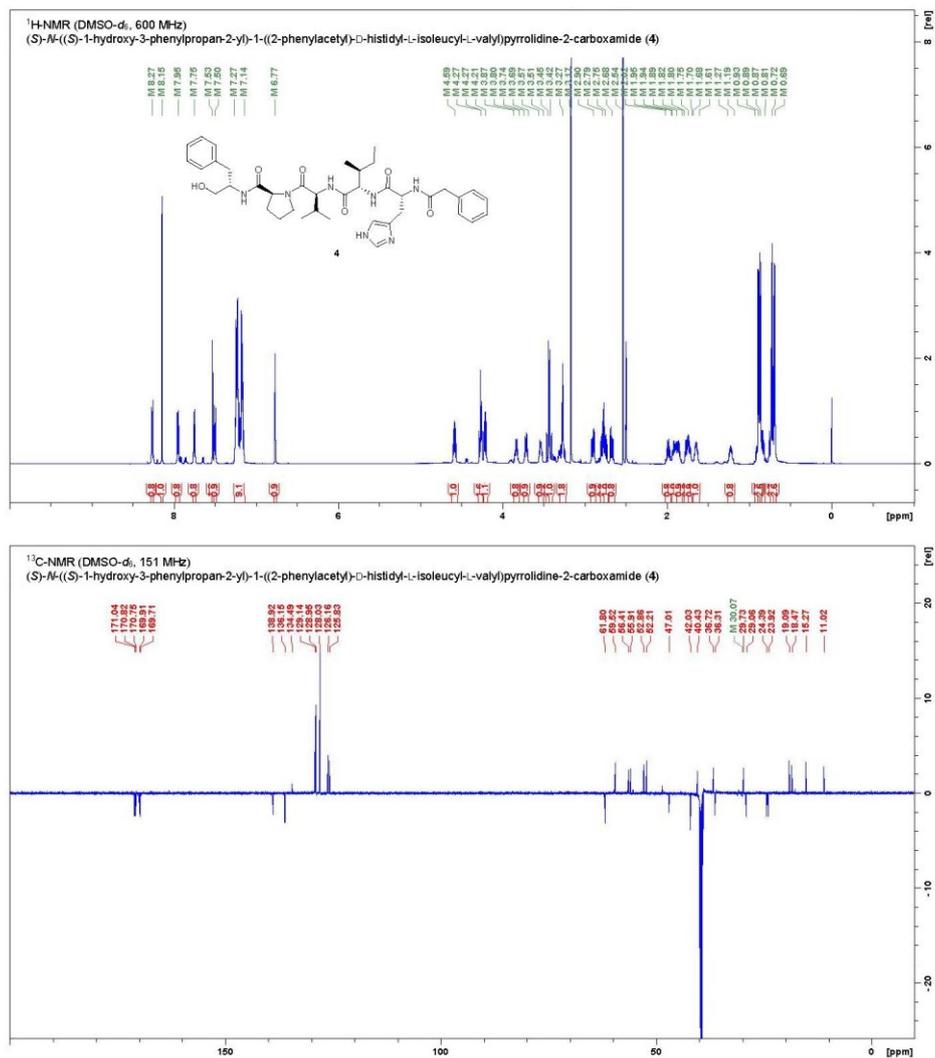


Figure S19. NMR spectra of 4 in DMSO-*d*₆.

Table S11. NMR data for **4** in DMSO-*d*₆ (600 MHz, 151 MHz).

Position	δ ¹³ C [ppm]	δ ¹ H [ppm]; (m, J, J)
L-Phenylalaninol		
CH ₂ OH	61.8	3.27 (t, 2H, <i>J</i> = 5.4 Hz)
NH	-	7.50 (d, 1H, <i>J</i> = 8.4 Hz)
α-CH	52.2	3.87-3.80 (m, 1H)
β-CH ₂	36.3	2.79 (dd, 1H, <i>J</i> = 13.5, 6.3 Hz) 2.68 (dd, 1H, <i>J</i> = 13.8, 7.2 Hz)
γ-C _{quart}	138.9	-
CH _{arom}	129.1, 129.0, 128.0, 126.2, 125.8	7.27-7.14 (m, 10H, overlay)
L-Proline		
CO	171.0	-
α-CH	59.5	4.27 (dd, 1H, <i>J</i> = 8.1, 3.9 Hz)
β-CH ₂	29.1	1.94-1.89 (m, 1H) 1.75-1.70 (m, 1H)
γ-CH ₂	24.4	1.89-1.82 (m, 1H) 1.80-1.75 (m, 1H)
δ-CH ₂ N	47.0	3.74-3.69 (m, 1H) 3.57-3.51 (m, 1H)
L-Valine		
CO	169.7	-
NH	-	7.95 (d, 1H, <i>J</i> = 8.3 Hz)
α-CH	55.9	4.27 (dd, 1H, <i>J</i> = 8.1, 8.1 Hz)
β-CH	29.7	2.02-1.95 (m, 1H)
γ-CH ₃	19.1	0.89 (d, 3H, <i>J</i> = 6.8 Hz)
γ-CH ₃	18.5	0.87 (d, 3H, <i>J</i> = 6.8 Hz)
L-Isoleucine		
CO	170.82	-
NH	-	7.75 (d, 1H, <i>J</i> = 8.8 Hz)
α-CH	56.4	4.21 (dd, 1H, <i>J</i> = 8.6, 7.2 Hz)
β-CH	36.7	1.68-1.61 (m, 1H)
γ-CH ₂	23.9	1.27-1.19 (m, 1H) 0.93-0.81 (m, 1H)
γ-CH ₃	15.3	0.69 (d, 3H, <i>J</i> = 7.2 Hz)
δ-CH ₃	11.0	0.72 (t, 3H, <i>J</i> = 7.4 Hz)

D-Histidine		
CO	170.75	-
NH	-	8.27 (d, 1H, $J = 7.9$ Hz)
α-CH	52.9	4.59 (ddd, 1H, $J = 7.9, 7.9, 6.6$ Hz)
β-CH₂	30.1, weak	2.90 (dd, 1H, $J = 14.6, 6.1$ Hz)
		2.75 (dd, 1H, $J = 14.4, 7.8$ Hz)
γ-C_{quat}	n.o.	-
δ-CH_{arom}	n.o.	6.77 (s, 1H)
ϵ-CH_{arom}	134.5	7.53 (s, 1H)
ϵ-NH	-	n.o.
Phenylacetyl		
CO	169.9	-
CH₂	42.0	3.45 (d, 1H, $J = 13.9$ Hz)
		3.42 (d, 1H, $J = 14.1$ Hz)
C_{quat}	136.2	-
CH_{arom}	129.1, 129.0, 128.0,	7.27-7.14 (m, 10H, overlay)
	126.2, 125.8	

Table S12. NMR data for **15** in DMSO-*d*₆ (600 MHz, 151 MHz).

Position	δ ¹³ C [ppm]		δ ¹ H [ppm]; (m, J, J)
Phenylalaninal			
CHO	200.3		9.47 (s, 0.4H) 9.41 (s, 0.2H)
NH	-		8.36 (d, 0.4H, <i>J</i> = 7.5 Hz) 8.34 (d, 0.2H, <i>J</i> = 7.2 Hz)
α -CH	59.7	59.5	4.33-4.25 (m, 2H, overlay) 4.25-4.15 (m, 2H, overlay)
β -CH ₂	33.5	33.3	3.15-3.07 (m, 1H) 2.93-2.82 (m, 2H, overlay) 2.78-2.69 (m, 2H, overlay)
γ -C _{quat}	137.62	137.60	-
CH _{arom}	129.3, 129.2, 128.8, 128.2, 128.1, 127.7, 126.23, 126.18		7.30-7.10 (m, 5H)
L-Proline			
CO	172.1	172.0	-
α -CH	59.1	59.0	4.33-4.25(m, 2H, overlay) 2.04-1.84 (m, 7H, overlay)
β -CH ₂	29.3		1.84-1.71 (m, 2H) 1.58-1.52 (m, 1H) 2.04-1.84 (m, 7H, overlay)
γ -CH ₂	24.5	24.4	1.84-1.71 (m, 2H) 1.58-1.52 (m, 1H)
δ -CH ₂ N	47.0		3.78-3.72 (m, 1H) 3.57-3.47 (m, 1H)
L-Valine 1			
CO	169.62	169.59	-
NH	-		8.01-7.94 (m, 2H, overlay)
α -CH	55.8		4.33-4.25(m, 2H, overlay)
β -CH	29.7		2.04-1.84 (m, 7H, overlay)
γ -CH ₃	19.0		0.89 (d, 3H, <i>J</i> = 6.8 Hz)
γ -CH ₃	18.44	18.40	0.87 (d, 3H, <i>J</i> = 6.6 Hz)
L-Valine 2			
CO	170.6		-

S71

NH	-		7.68-7.61 (m, 1H)
α -CH	56.9		4.25-4.15 (m, 2H, overlay)
β -CH	30.65	30.60	2.04-1.84 (m, 7H, overlay)
γ -CH ₃	19.0		0.75-0.68 (m, 6H)
	17.6		
D-Histidine			
CO	171.0		-
NH	-		8.01-7.94 (m, 2H, overlay)
α -CH	52.7		4.62-4.56 (m, 1H)
β -CH ₂	29.8		2.93-2.82 (m, 2H, overlay)
			2.78-2.69 (m, 2H, overlay)
γ -C _{quart}	n.o.		-
δ -CH _{arom}	n.o.		6.77 (s, 1H)
ϵ -CH _{arom}	134.5		7.51 (s, 1H)
ϵ -NH	-		n.o.
Isovaleroyl			
CO	171.4		-
CH ₂	44.4		2.04-1.84 (m, 7H, overlay)
CH	25.4		2.04-1.84 (m, 7H, overlay)
CH ₃	22.2		0.82 (d, 3H, $J = 6.4$ Hz)
CH ₃	22.1		0.78 (d, 3H, $J = 6.2$ Hz)

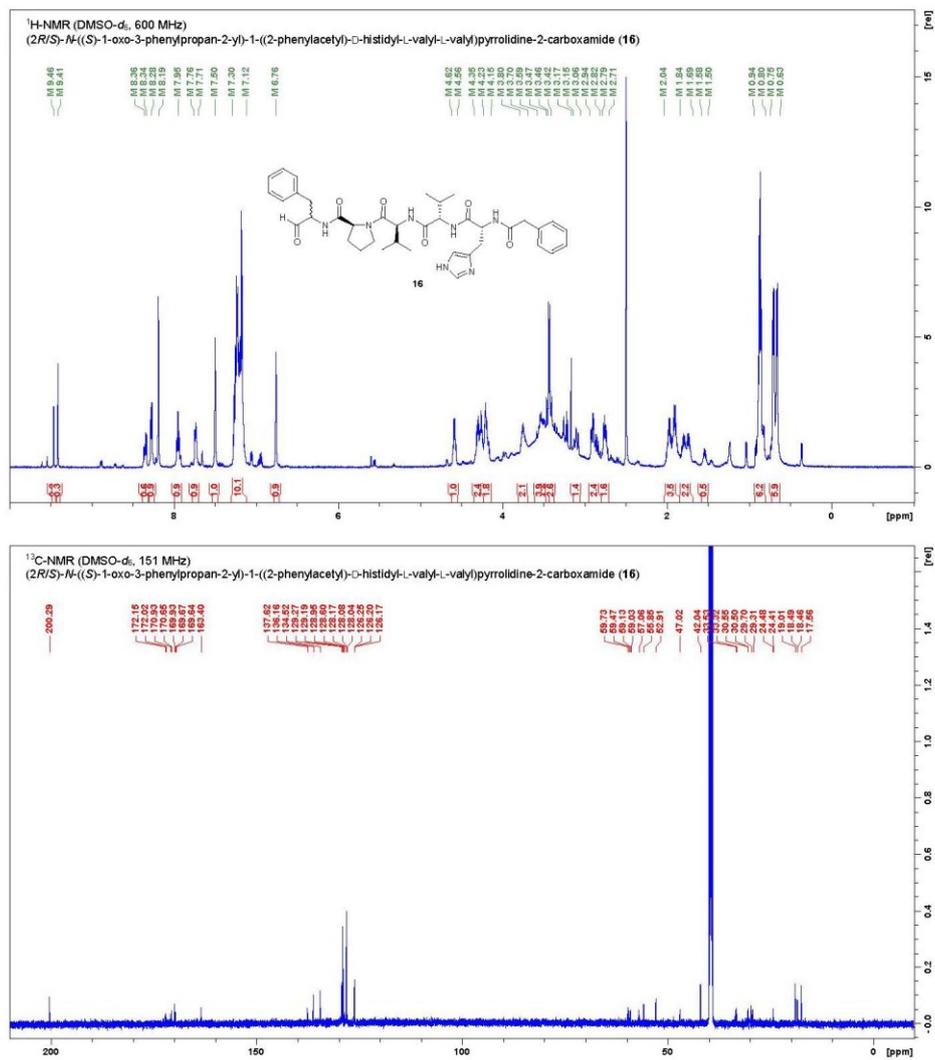


Figure S21. NMR spectra of 16 in DMSO-*d*₆.

Table S13. NMR data for **16** in DMSO-*d*₆ (600 MHz, 151 MHz).

Position	δ ¹³ C [ppm]		δ ¹ H [ppm]; (m, J, J)
Phenylalaninal			
CHO	200.3		9.46 (s, 0.2H) 9.41 (s, 0.3H)
NH	-		8.36 (d, 0.3H, J = 7.5 Hz) 8.34 (d, 0.3H, J = 6.8 Hz)
α -CH	59.7	59.5	4.35-4.23 (m, 2H, overlay) 4.23-4.15 (m, 2 H, overlay) 3.15-3.06 (m, 1H)
β -CH ₂	33.5	33.3	2.94-2.82 (m, 2H, overlay) 2.79-2.71 (m, 2H, overlay)
γ -C _{quart}	137.6		-
CH _{arom}	129.3, 129.2, 129.0, 128.6, 128.2, 128.1, 128.0, 126.25, 126.20, 126.17		7.30-7.12 (m, 10H, overlay)
L-Proline			
CO	not distinguishable		-
α -CH	59.1	59.0	4.35-4.23 (m, 2H, overlay)
β -CH ₂	29.3		2.04-1.84 (m, 4H, overlay) 1.84-1.69 (m, 2H, overlay)
γ -CH ₂	24.5	24.4	2.04-1.84 (m, 4H, overlay) 1.58-1.50 (m, 1H, overlay)
δ -CH ₂ N	47.0		3.80-3.70 (m, 2H) 3.59-3.47 (m, 4H)
L-Valine 1			
CO			-
NH	-		7.95 (dd, 1H, J = 8.4, 8.4 Hz)
α -CH	55.8		4.35-4.23 (m, 2H, overlay)
β -CH	29.7		2.04-1.84 (m, 4H, overlay)
γ -CH ₃	18.49, 18.46		0.94-0.80 (m, 6H)
L-Valine 2			
CO	not distinguishable		-
NH	-		7.76-7.71 (m, 1H)
α -CH	57.1		4.23-4.15 (m, 2 H, overlay)

β -CH	30.6 30.5	2.04-1.84 (m, 4H, overlay)
γ -CH ₃	19.0, 17.6	0.75-0.63 (m, 6H)
D-Histidine		
CO	not distinguishable	-
NH	-	8.28 (d, 1H, $J = 8.1$ Hz)
α -CH	52.9	4.62-4.56 (m, 1H)
β -CH ₂	30.0*	2.94-2.82 (m, 2H, overlay) 2.79-2.71 (m, 2H, overlay)
γ -C _{quart}	n.o.	-
δ -CH _{arom}	n.o.	6.76 (s, 1H)
ϵ -CH _{arom}	134.5	7.50 (s, 1H)
ϵ -NH	-	n.o.
Phenylacetyl		
CO	169.9	-
		3.46 (d, 2H, $J = 13.9$ Hz)
CH ₂	42.0	3.42 (d, 2H, $J = 14.1$ Hz)
C _{quart}	136.2	-
	129.3, 129.2, 129.0,	
	128.6, 128.2, 128.1,	
CH _{arom}	128.0, 126.25,	7.30-7.12 (m, 10H, overlay)
	126.20, 126.17	

Some integrals aren't accurate due to the broad peak of $\delta_{\text{H}} = 4.00$ to 2.59 ppm.

The carbonyl signals 172.1, 172.0, 170.9, 170.7, 169.7, 169.6 were not distinguishable with HMBC data.

S75

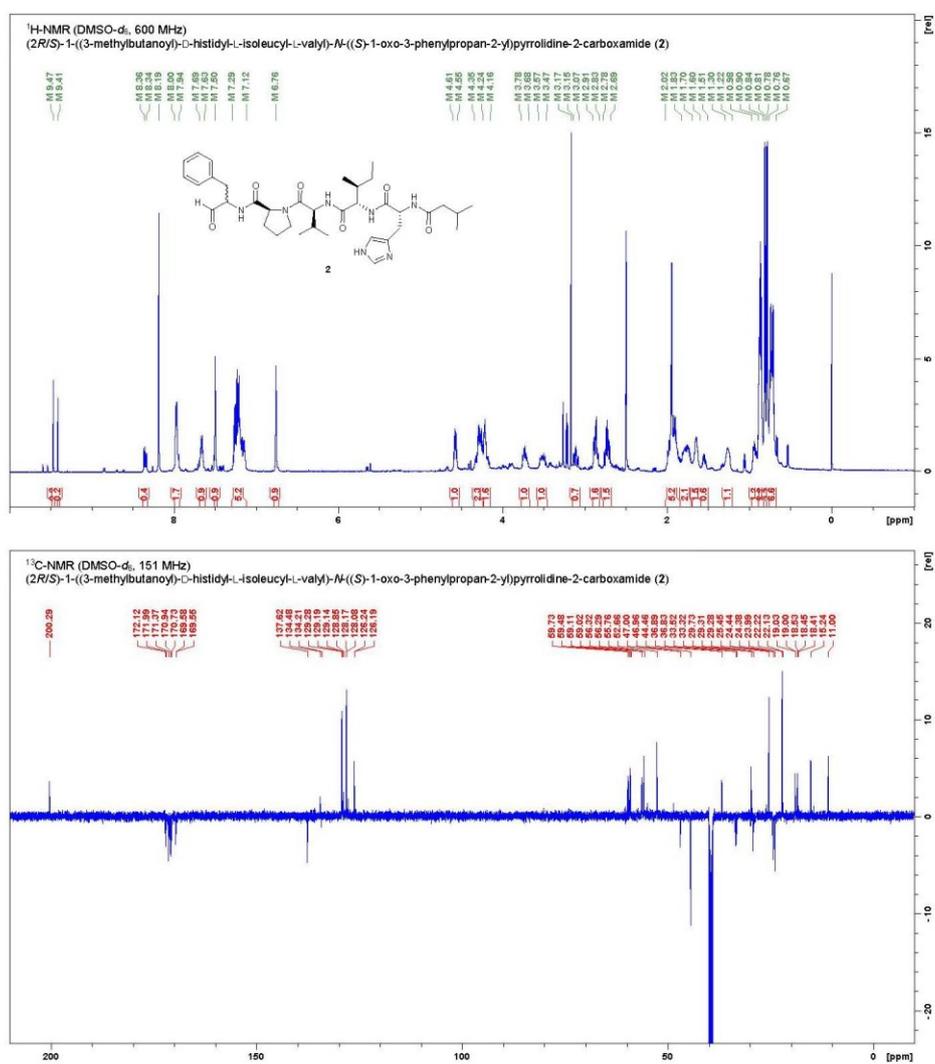


Figure S22. NMR spectra of synthetic pentacitidin A (2) in DMSO-*d*₆.

Table S14. NMR data for synthetic pentacitidin A (**2**) in DMSO-*d*₆ (600 MHz, 151 MHz).

Position	δ ¹³ C [ppm]		δ ¹ H [ppm]; (m, J, J)
Phenylalaninal			
CHO	200.3		9.47 (s, 0.2H) 9.41 (s, 0.2H)
NH	-		8.36 (d, 0.2H, J= 7.7 Hz) 8.34 (d, 0.2H, J= 7.0 Hz)
α -CH	59.7	59.5	4.35-4.24 (m, 2H, overlay) 4.24-4.16 (m, 2H, overlay)
β -CH ₂	33.6	33.3	3.15-3.07 (m, 1H) 2.91-2.83 (m, 2H, overlay) 2.78-2.69 (m, 2H, overlay)
γ -C _{quat}	137.6		-
CH _{arom}	129.3, 129.2, 129.1, 128.9, 128.2, 128.1, 126.24, 126.19		7.29-7.12 (m, 5H)
L-Proline			
CO	172.1	172.0	-
α -CH	59.1	59.0	4.35-4.24 (m, 2H, overlay) 2.02-1.83 (m, 5H, overlay)
β -CH ₂	29.31	29.28	1.83-1.70 (m, 2H, overlay) 1.60-1.51 (m, 1H)
γ -CH ₂	24.44	24.38	2.02-1.83 (m, 5H, overlay) 1.83-1.70 (m, 2H, overlay)
δ -CH ₂ N	47.00	46.96	3.78-3.68 (m, 1H) 3.57-3.47 (m, 1H)
L-Valine			
CO	169.58	169.55	-
NH	-		8.00-7.94 (m, 2H, overlay)
α -CH	55.7		4.35-4.24 (m, 2H, overlay)
β -CH	29.7		2.02-1.83 (m, 5H, overlay)
γ -CH ₃	19.03	19.00	0.90-0.84 (m, 6H, overlay)
γ -CH ₃	18.45	18.41	0.90-0.84 (m, 6H, overlay)
L-Isoleucine			
CO	170.7		-
NH	-		7.69-7.63 (m, 1H)

S77

α -CH	56.32	56.29	4.24-4.16 (m, 2H)
β -CH	36.9	36.8	1.70-1.60 (m, 2H)
γ -CH ₂	24.0		1.30-1.22 (m, 1H) 0.98-0.90 (m, 1H)
γ -CH ₃	15.2		0.76-0.67 (m, 6H, overlay)
δ -CH ₃	11.0		0.76-0.67 (m, 6H, overlay)
D-Histidine			
CO	170.9		-
NH	-		8.00-7.94 (m, 2H, overlay)
α -CH	52.7		4.61-4.55 (m, 1H)
β -CH ₂	29.9*		2.91-2.83 (m, 2H, overlay) 2.78-2.69 (m, 2H, overlay)
γ -C _{quart}	134.2		-
δ -CH _{arom}	n.o.		6.76 (s, 1H)
ϵ -CH _{arom}	134.5		7.50 (s, 1H)
ϵ -NH	-		n.o.
Isovaleroyl			
CO	171.4		
CH ₂	44.5		2.02-1.83 (m, 5H, overlay)
CH	25.5		2.02-1.83 (m, 5H, overlay)
CH ₃	22.2		0.81 (d, 3H, $J = 6.4$ Hz)
CH ₃	22.1		0.78 (d, 3H, $J = 6.4$ Hz)

S78

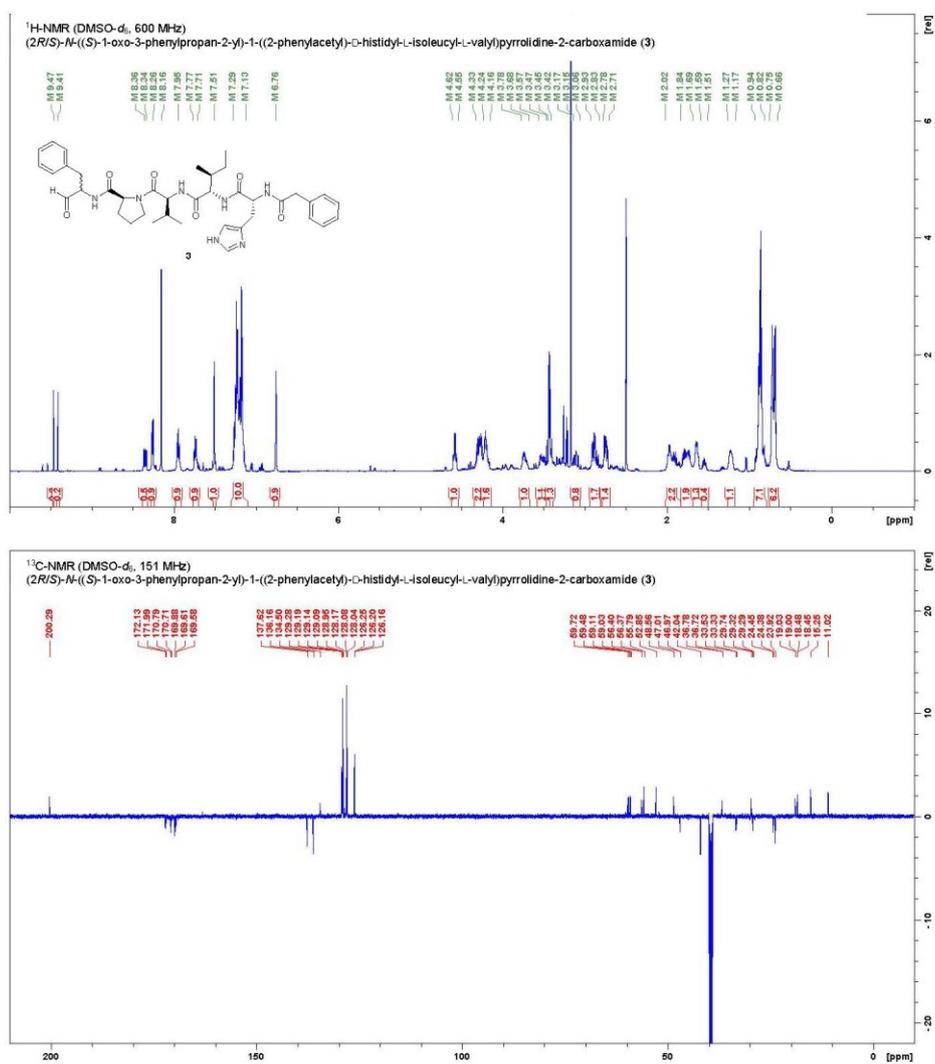


Figure S23. NMR spectra of synthetic pentacitidin B (3) in DMSO-*d*₆.

S79

Table S15. NMR data for synthetic pentacitidin B (**3**) in DMSO-*d*₆ (600 MHz, 151 MHz).

Position	δ ¹³ C [ppm]		δ ¹ H [ppm]; (m, J, J)
Phenylalaninal			
CHO	200.3		9.47 (s, .0.2H) 9.41 (s, .0.2H)
NH	-		8.36 (d, 0.25H, J= 7.7 Hz) 8.34 (d, 0.25H, J= 7.0 Hz)
α -CH	59.1	59.0	4.24-4.16 (m, 2H, overlay) 3.15-3.06 (m, 1H, overlay)
β -CH ₂	33.5	33.3	2.93-2.83 (m, 2H, overlay) 2.78-2.71 (m, 1H, overlay)
γ -C _{quart}	137.6		-
CH _{arom}	129.3, 129.2, 129.14, 129.09, 129.0, 128.2, 128.1, 128.0, 126.25, 126.20, 126.16		7.29-7.13 (m, 10H, overlay)
L-Proline			
CO	172.1	172.0	-
α -CH	59.5		4.33-4.24 (m, 2H, overlay) 2.02-1.84 (m, 2H, overlay)
β -CH ₂	29.32	29.29	1.84-1.69 (m, 2H, overlay) 1.59-1.51 (m, 1H)
γ -CH ₂	24.5	24.4	2.02-1.84 (m, 2H, overlay) 1.84-1.69 (m, 2H, overlay)
δ -CH ₂ N	47.01	46.97	3.78-3.68 (m, 1H) 3.57-3.47 (m, 1H)
L-Valine			
CO	169.61	169.58	-
NH	-		7.95 (t, 1H, J= 9.0 Hz)
α -CH	55.8		4.33-4.24 (m, 2H, overlay)
β -CH	29.7		2.02-1.84 (m, 2H, overlay)
γ -CH ₃	19.03	19.00	0.94-0.82 (m, 7H, overlay)
	18.5	18.4	
L-Isoleucine			
CO	170.7		-
NH	-		7.77-7.71 (m, 1H)

α -CH	56.40	56.37	4.24-4.16 (m, 2H, overlay)
β -CH	36.8	36.7	1.69-1.59 (m, 1H)
γ -CH ₂	23.9		1.27-1.17 (m, 1H)
γ -CH ₃	15.3		0.94-0.82 (m, 7H, overlay)
δ -CH ₃	11.0		0.75-0.66 (m, 6H, overlay)
D-Histidine			
CO	170.8		-
NH	-		8.26 (d, 1H, $J = 8.1$ Hz)
α -CH	52.9		4.62-4.55 (m, 1H)
β -CH ₂	30.1*		2.93-2.83 (m, 2H, overlay)
			2.78-2.71 (m, 1H, overlay)
γ -C _{quart}	n.o.		-
δ -CH _{arom}	n.o.		6.76 (s, 1H)
ϵ -CH _{arom}	134.5		7.51 (s, 1H)
ϵ -NH	-		n.o.
Phenylacetyl			
CO	169.9		-
CH ₂	42.0		3.45 (d, 1H, $J = 14.1$ Hz)
			3.42 (d, 1H, $J = 14.1$ Hz)
C _{quart}	136.2		-
	129.3, 129.2, 129.14,		
	129.09, 129.0, 128.2,		
CH _{arom}	128.1, 128.0, 126.25,		7.29-7.13 (m, 10H, overlay)
	126.20, 126.16		

Table S16. Summary of molecular docking results.

	Falcipain-2		Falcipain-3		Chymotrypsin	
PDB-ID	3BPF		3BPM		1AFQ	
Reference ligand	E-64		Leupeptin		D-Leucyl-L-phenylalanyl- <i>p</i> -fluorobenzylamide	
Docking scores^a	non-covalent (distance) ^b	covalent	non-covalent (distance) ^b	covalent	non-covalent (distance) ^b	covalent
Re-docking [RMSD, Å]	-4.8 (3.5) [1.4] ^d	-15.7	-5.8 (3.6) [1.3]	-14.1	-12.5 (-) [0.8]	-
ace-Val-Pro-Phe-alcohol	-3.2	-	-4.5	-	-7.1	-
ace-Val-Pro-L-Phe-aldehyde	-2.0 (3.0)	-16.0	-2.7 (3.0)	-12.3	-8.4 (2.0)	-10.1
ace-Val-Pro-D-Phe-aldehyde	-2.5 (7.3)	-14.6	-2.5 (5.0)	-13.1	-6.8 (3.4)	-12.3
ace-Val-Pro-Phe-acid	-2.9	-	-3.8	-	-7.5	-
ace-Val-Pro-amide ^e	-2.0	-	-3.3	-	-3.4	-
ace-Val-Val-Pro-amide ^e	-3.6	-	-4.1	-	-4.2	-

^aValues are docking scores in kcal/mol. ^bDescribes distance between electrophilic carbon if present and nucleophilic sulfur of cysteine or oxygen of serine in Å. ^cTruncation of falcitidin.

⁴For 3BPF non-covalent docking, potential field generation, and re-docking the binding mode of aligned leupeptin from PDB-ID 3BPM was used as reference.

S83

6. Chapter 3 – Cryopeptides

This work is a contribution to Luis J. Linares Otoyas project from the research group of Prof. Dr. T. Schäberle from the Justus-Liebig-University

My main contributions to the project are:

- Development of the synthetic strategy
- Synthesis and analytics of
 - Natural products as comparison to isolated ones
 - Derivatives
 - Precursor molecules
- Structural confirmation of the natural isolated compounds based on NMR data

6.1 Introduction

Non-proteinogenic amino acids are common in natural products. They often have a distinct physiological role, like components of bacterial cell walls, and offer unique structural features. One class are α,β -didehydro- α -amino acids (in the following referred to as dehydro amino acids). They are commonly found in bacteria and often show antibacterial properties (Figure 6-1).³⁴

³⁴ D. Siodłak, *Amino Acids*, **2015**, 47, 1-17.

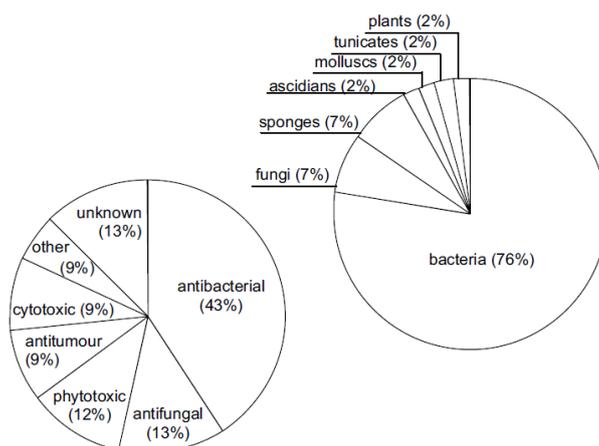


Figure 6-1: Origin and bioactivity of peptides containing dehydro amino acids.

One of the oldest known dehydro amino acids containing peptide is Nisin (**3.1**) that was isolated from *Lactococcus lactis*. It contains dehydroalanine and dehydrobutyrine (Dhb) (Figure 6-1 and Figure 6-2)³⁵ and is nowadays used as a food preservative based on its antibacterial properties. Myxovalargin A (**3.2**), another antibiotic active compound, sets a further example containing dehydrovaline (Dhv) and dehydroisoleucine.³⁶

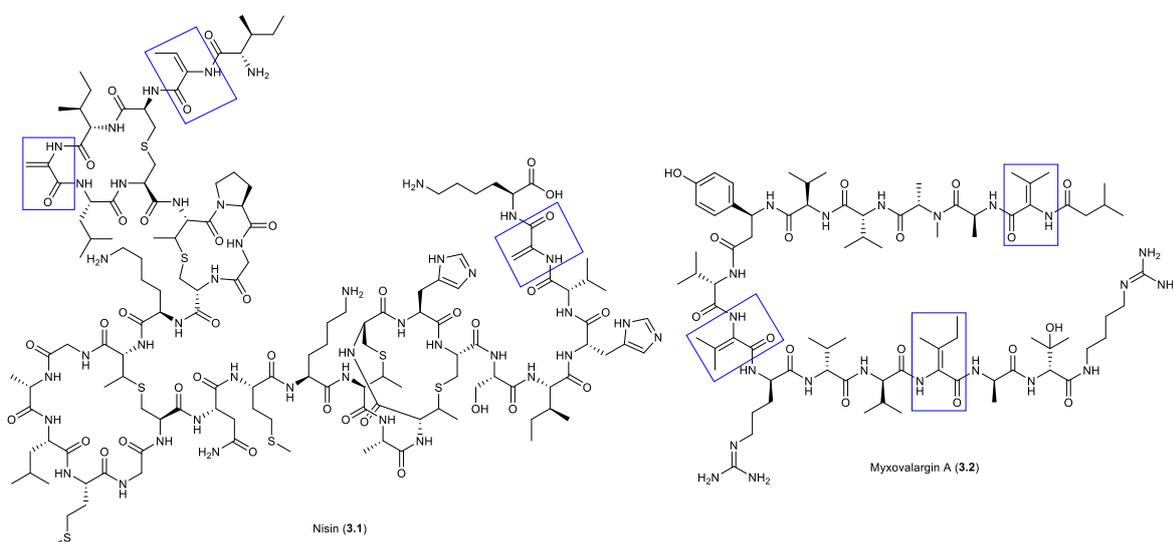


Figure 6-2: Examples of oligopeptides containing dehydro amino acids (highlighted in blue). Nisin (3.1) and Myxovalargin (3.2).

³⁵ A. T. R. Mattick, A. Hirsch, N. J. Berridge, *Lancet*. **1947 Jul 5**, 2 (6462), 5-8.

³⁶ H. Irschik, K. Gerth, G. Höfle, W. Kohl, H. Reichenbach, *J. Antibiot.* **1983**, 36 (12), 6-12.

To date it is not known how and especially when the dehydro amino acids are produced by bacteria and introduced in the peptide chain. Glidopeptine A and its known biosynthetic gene cluster give no indication on when or how the dehydrogenation occurs. The amino acid sequence is consistent with the gene cluster, except for the dehydro moieties. Instead of a coded D-Thr and L-Val, a Dhb and Dhv are incorporated, respectively (Figure 6-3).³⁷

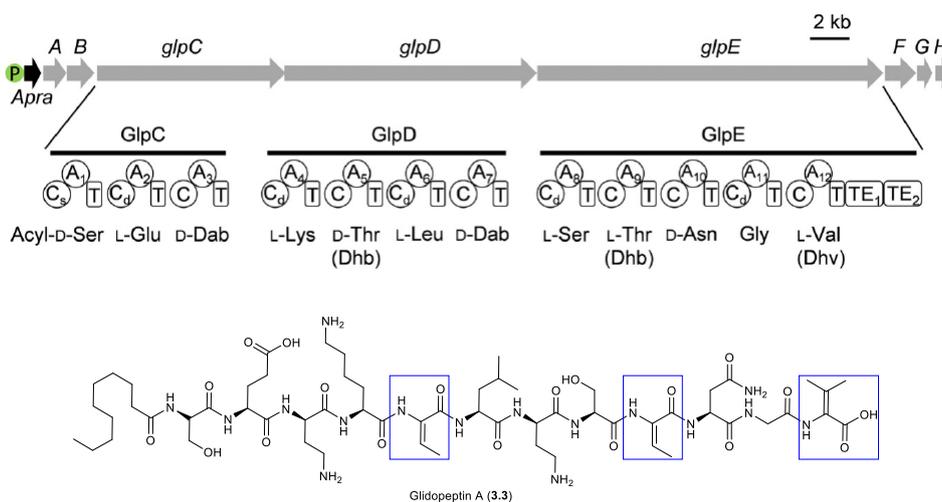


Figure 6-3: Gene cluster and structure of glidopeptin A (3.3). Highlighted in blue are the dehydro amino acids.³⁷

In an extract and a created molecular network of two strains of *Pedobacter cryoconitis* Luis J. Linares Otoyá of the Justus-Liebig-University and research group of Prof. Dr. T. Schäberle could isolate two natural compounds containing dehydro amino acids, **PE2H_3** ($m/z = 713.4343$, molecular formula $C_{36}H_{57}N_8O_7^+$; Figure 6-8 b) and **P13-71** ($m/z = 713.4346$ molecular formula $C_{36}H_{57}N_8O_7^+$; Figure 6-10 b). Analysis of MS² spectra for both substances lead to the assignment of the general peptide sequence Leu-Dhv-Arg-Dhv-Phe-fatty acid (Figure 6-8 d, Figure 6-10 d). They shared a fragment ion of $\sim m/z = 232.1332$ with molecular formula $C_{11}H_{16}N_3O_2^+$ consisting of phenylalanine and the fatty acid, indicating a difference within the acyl moiety with a molecular formula of $C_5H_9O^+$. Proceeding from standard L amino acids, based on data of the biosynthetic gene cluster, three different structures are likely. Containing the general amino acid sequence of Leu-Dhv-Arg-Dhv-Phe and a small fatty acid

³⁷ X. Wang *et al.* *PNAS* **2018**, 115 (18), E4255-E4263

side chain, either an isovaleroyl (iVal) (**3.3**), *S*-methylbutanoyl (SBM) (**3.4**) or an *R*-methylbutanoyl (RBM) (**3.5**) residue are possible (Figure 6-4).

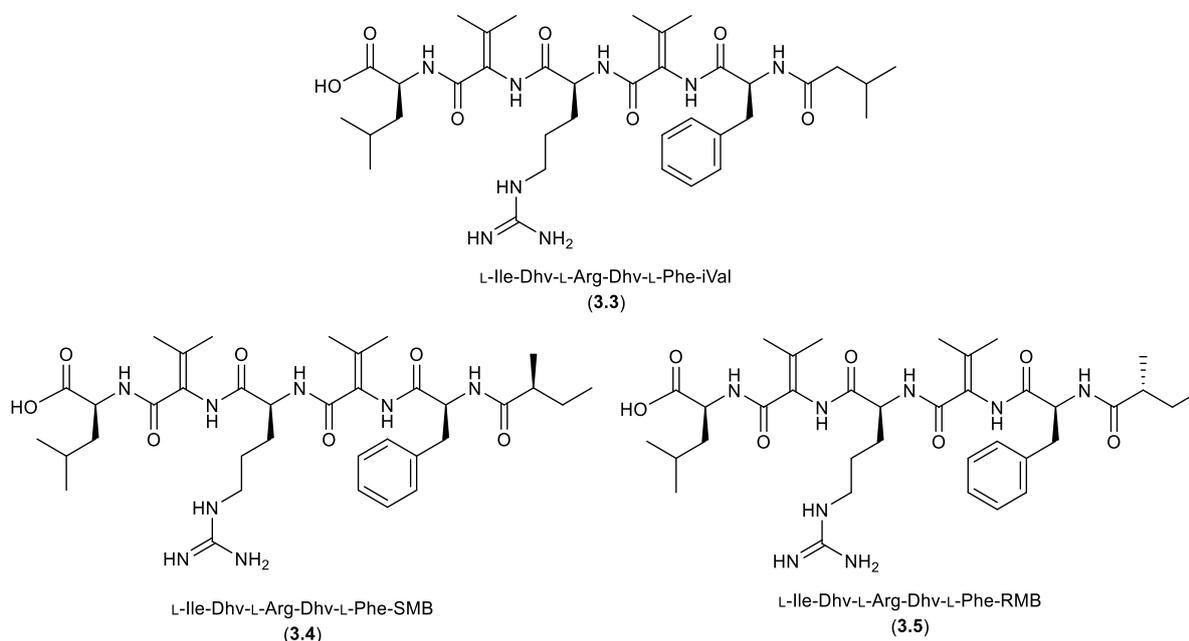


Figure 6-4: Assigned structures for compounds corresponding to ions $\sim m/z = 713.434$.

For both fractions, **P13-71** and **PE2H_3**, structure elucidation by NMR analysis was performed, especially to determine the stereogenic center of the side chain. First results of the NMR spectra, particularly the methyl signals, indicated correspondence of structure **3.3** to fraction **PE2H_3**. The second isolated fraction **P13-71** can either match derivative **3.4** or **3.5**. For final structure elucidation and validation of the assigned structures, as well as for future activity tests, synthetic material is needed.

Since it is not clear when the formation of the dehydrovaline moieties takes place and what the substrate is, three possible scenarios regarding the formation are possible:

- a) Free *D*- or *L*-valine is used as the substrate and is dehydrogenated before it is incorporated in the non-ribosomal peptide (NRP).
- b) During the NRP assembly the growing peptide chain attached to the enzyme is dehydrogenated in position of the valine. The substrate can either have an *R* or *S* configuration, for which we need a mimic of the substrate.

c) The valines of the released peptide are dehydrogenated after the release from the NRPS. Again, the valine can either obtain an *R* or *S* configuration.

For scenario a) commercially available D- and L-valine can be used. For scenario b) a mimic of the growing peptide is needed, for which aminoacyl-*N*-acetylcysteamine thioesters (SNACs) of valine in both configurations were chosen (Figure 6-5).

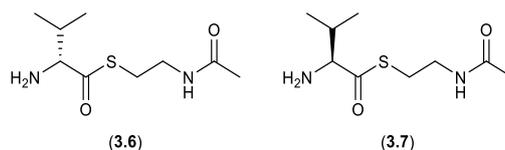


Figure 6-5: Structures of *R*- and *S*-Val-SNAC (3.6, 3.7).

For scenario c) a linear peptide is needed, containing either D- or L-valine instead of the dehydrovalines (Figure 6-6). Based on first NMR data, which hinted that structure **3.3** and its isovaleroyl residue corresponds to the assigned structure of the MS² fragmentation, modified analogues **3.8** and **3.9** were chosen.

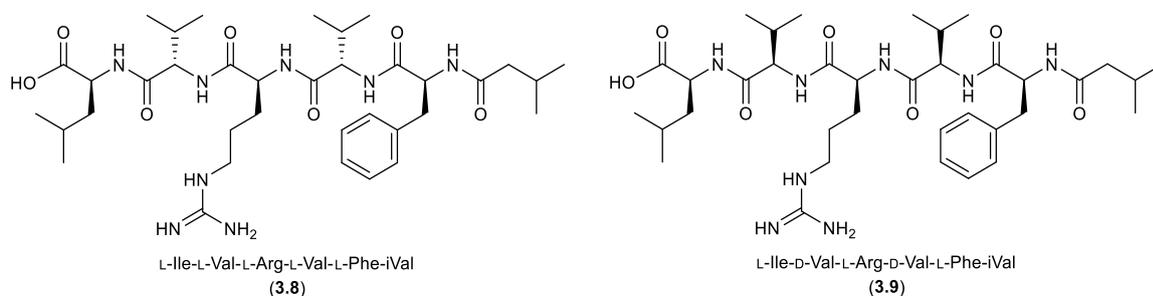


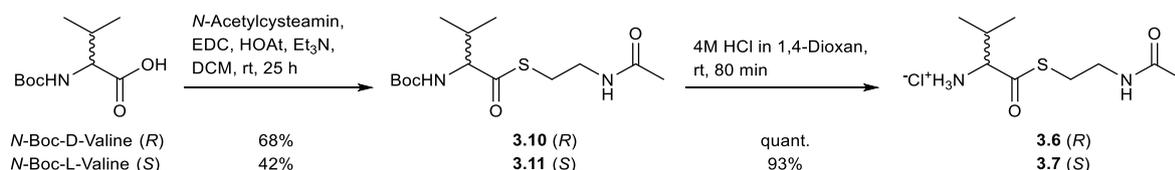
Figure 6-6: Modified analogues corresponding to Structure 3.3 where dehydrovalines are replaced by L-valine (3.8) and D-valine (3.9) and with the iVal side chain residue.

6.2 Results and Discussion

6.2.1 Synthesis

As mimics for testing the dehydrogenation process, both valyl-SNAC enantiomers as shown in Scheme 6-1 were synthesized, following standard peptide coupling and Boc-deprotection

procedures.³⁸ The EDC/HOAt-mediated coupling was carried out in solution and **3.10** and **3.11** were obtained with a 68% and 42% yield, respectively. The deprotection was achieved quantitatively for **3.6** and with 93% for **3.7**. This affords the *R*-enantiomer **3.6** with an overall yield of 68% and 39% for the *S*-enantiomer **3.7**.



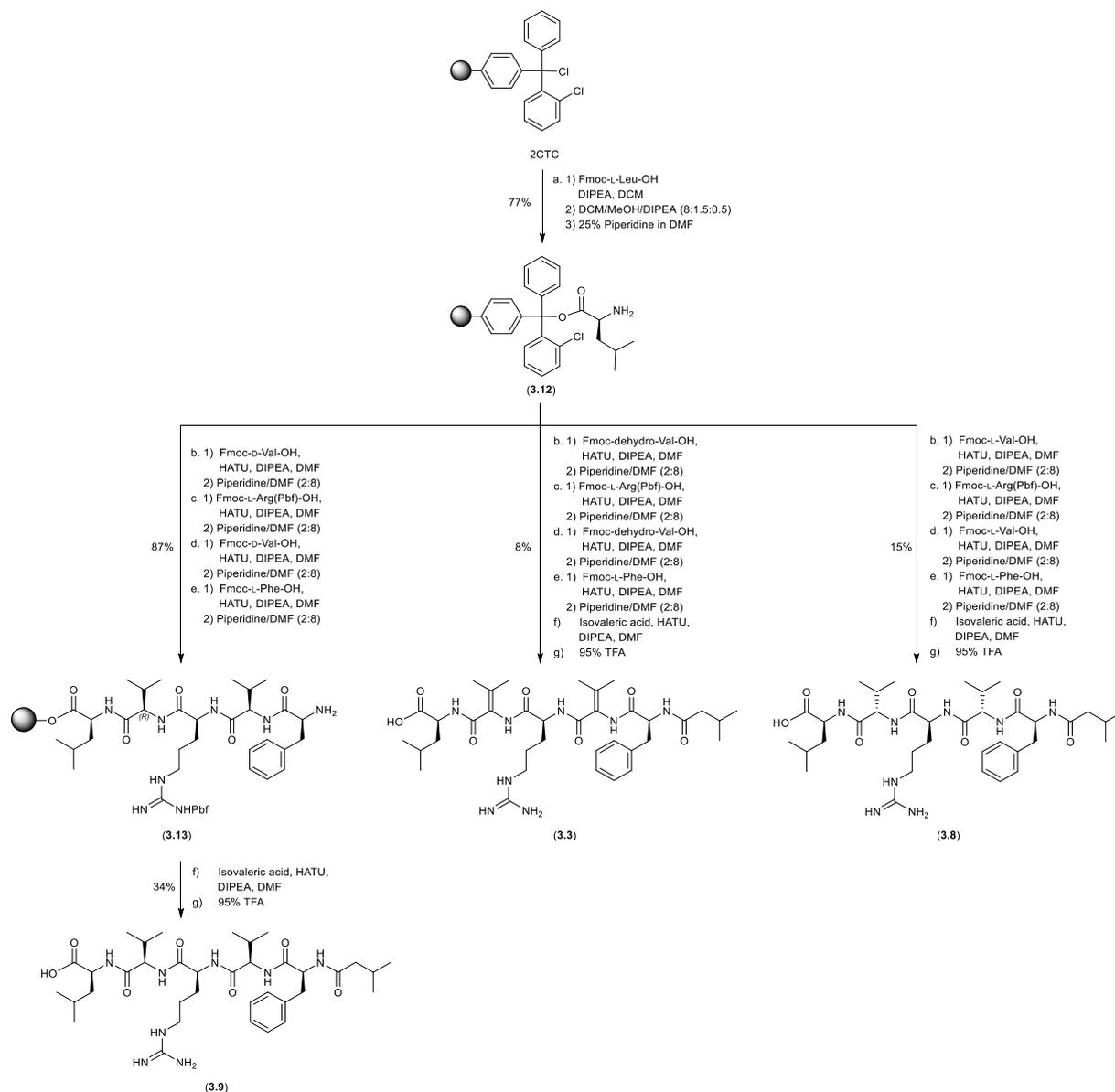
Scheme 6-1: Synthesis of valyl-SNAC enantiomers.

For the synthesis of the linear peptides, compound **3.3** was chosen to compare it to the natural isolated one and its modified analogs **3.9** and **3.8** for the investigation of the dehydrogenation process. Based on previous synthetic approaches towards peptide-based molecules, solid phase peptide synthesis (SPPS) was utilized to synthesize all compounds.³⁹ Due to the different implementation of valine analogues, *D*-Valine, *L*-Valine and dehydrovaline, as the second and fourth amino acid, a split approach was not practicable and a linear synthesis (Scheme 6-2) was developed. The attachment of *L*-Isoleucine to the resin was prepared for **3.9**, **3.3** and **3.8**. The yield, based on the loading and the weight gain was determined to be 77%. For derivative **3.9** the batch was split before the attachment of the fatty acid and a yield of 87% for the coupling of the amino acids and resin bound peptide **3.13** was obtained. The coupling of isovaleric acid and subsequent cleavage of the peptide from the resin yielded **3.9** with 34%. Not taking the attachment of the first amino acid into account, the overall yield is 30%. For Cryopeptide **3.3** with its dehydrovalines the total yield over the same steps is 8% and for **3.8** it is 15%. The work up of these two compounds was more problematic compared to **3.9**, due to the formation of an unsolvable third layer during the extraction. Further analysis revealed the presence of the desired product to be in this layer.

³⁸ D. Niedek *et al.* *J. Org. Chem.* **2020**, *85*(4), 1935-1846.

³⁹ W. Chan, P. White, *Fmoc Solid Phase Peptide Synthesis Practical Approach*, Oxford University Press, Oxford, **1999**.

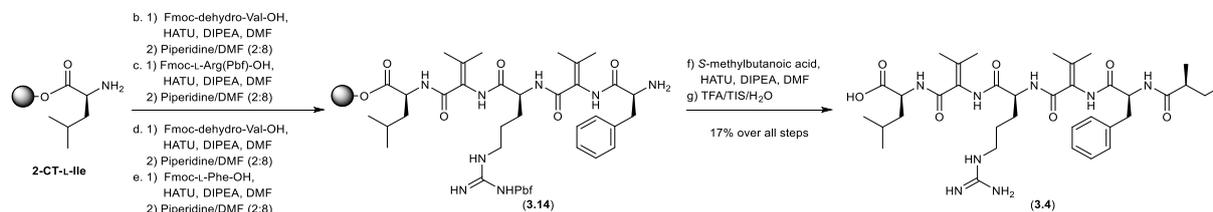
Purification of it remained problematic, but nevertheless the obtained amount was considered sufficient and a further optimization was postponed.



Scheme 6-2: Synthesis for Cryopeptides 3.9, 3.3 and 3.8.

To disclose the stereogenic center of the side chain moiety, which is a methylbutanoyl residue with either an *R* or *S* configuration, linear pentapeptide **3.14** was prepared, after which the batch was split (Scheme 6-3). That allows for a quick and easy derivatization with different fatty acids. The easier obtainable *S*-methylbutanoic acid was the first fatty acid to be used for

derivatization. After the coupling and subsequent cleavage of the peptide from the resin **3.4** was obtained with 17% overall yield.



Scheme 6-3: Synthesis of **3.14** and **3.4**.

Compared to the self-attachment of isoleucine to the resin of the former derivatives **3.3**, **3.8** and **3.9** (Scheme 6-2), for this synthesis of **3.4** commercially available 2-CTC-L-Isoleucine resin was used, expecting the determined yield would be more descriptive. Due to the batch splitting after **3.14**, the resins static behavior when completely dry, and its sticky behavior when damp, the determined yield can still be error-prone. Compared to the synthesis of **3.3**, which only differs in the lipophilic side chain, the yield is more than twice as high. Besides the chosen resin, a different cleavage method and a different work up strategy was adopted as well. The crude was lyophilized directly to improve the yield and to avoid extraction. Overall, the more than 50% higher yield can be attributed more towards changing the work up procedure and not towards the chosen resin (see experimental part for **3.3**, **3.8** and **3.9** for self-prepared resin and **3.14** and **3.4** for bought resin).

6.2.2 Comparison of Synthetic to Natural Isolated Compounds

Derivatives **3.3** and **3.4** were synthesized to confirm the proposed structure of the limited natural isolated products **PE2H_3** and **P13-71**. The stereo information of the fatty acid part of **P13-71** could not be confirmed unambiguously. To elucidate the stereogenic center of it, **3.4** was synthesized and then compared to the data of the natural isolated compound.⁴⁰

⁴⁰ Data interpretation and structure elucidation of natural isolated fractions was performed by Dr. Yang Liu.

The sample of the natural isolated **PE2H_3** is not pure as shown in Figure 6-7. It contains a second ion with $m/z = 715.4496$ $[M+H]^+$ (orange EIC) besides the main ion of $m/z = 713.4348$ $[M+H]^+$ (blue EIC).

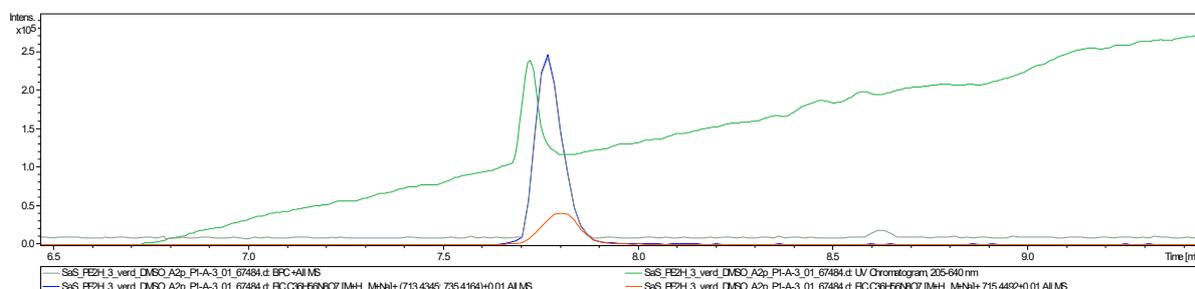
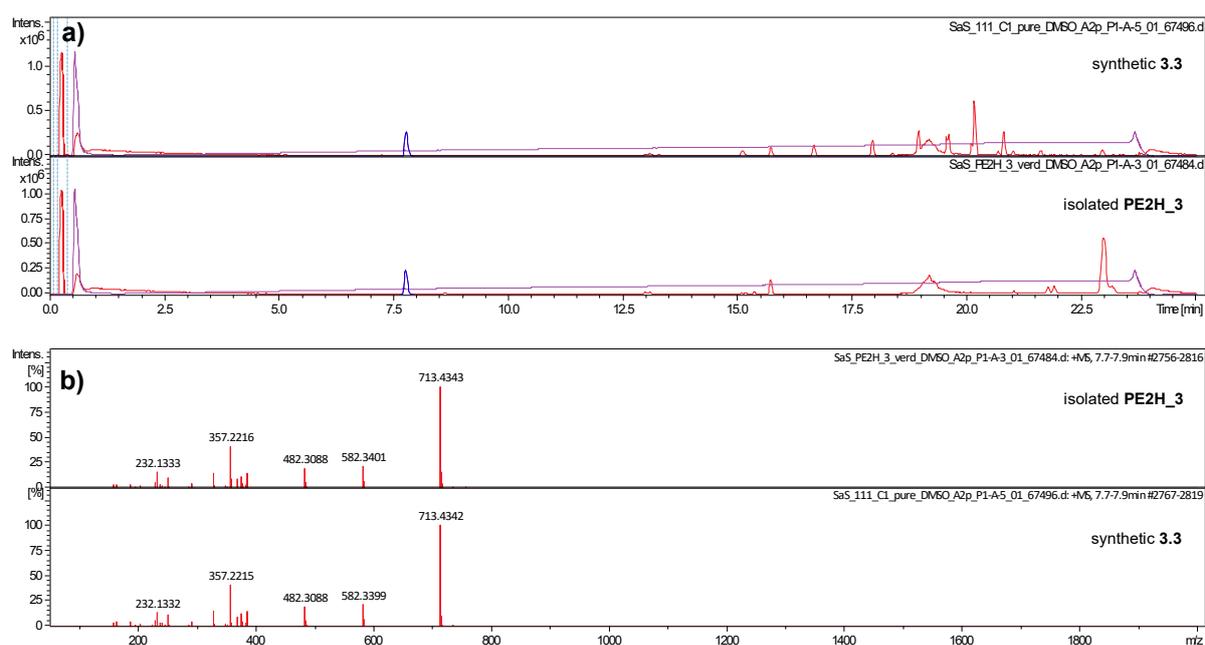


Figure 6-7: Chromatograms of natural isolated PE2H_3. Green: UV chromatogram 205-640 nm; Grey: BPC; Orange: EIC of $m/z = 715.4492 \pm 0.01$; Blue: EIC of $m/z = 713.4346 \pm 0.01$.

Based on MS and MS² data, the comparison of synthetic **3.3** to natural isolated fraction **PE2H_3** confirmed the proposed structure shown in Figure 6-8 based on a) the correlation of the retention time with 7.8 min for both natural and synthetic compound, b) the fragmentation pattern and intensities, which are identical and c) and d) the fragmentation pattern and structural assignment for the parent ion of $m/z = 713.4339$, which is also identical for both compounds.



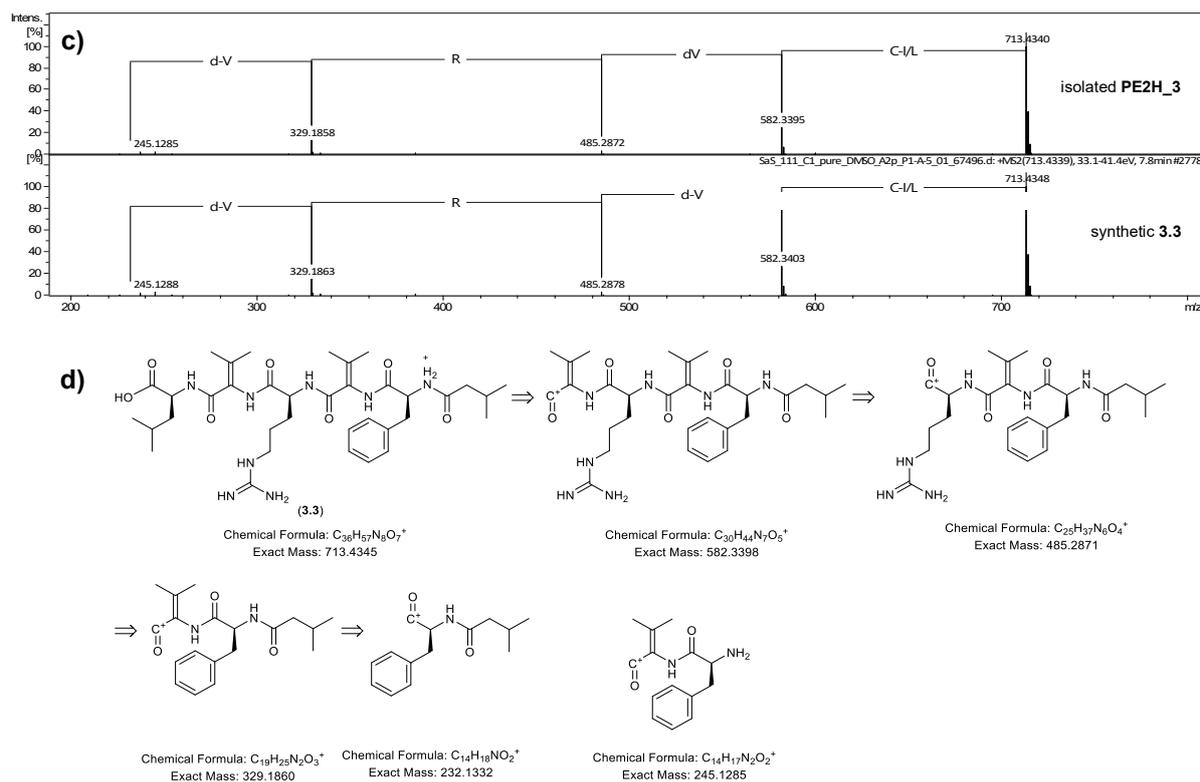


Figure 6-8: Comparison of MS and MS² data of natural isolated compound PE2H_3 and synthetic 3.3. a) Comparison of retention time. b) Comparison of fragmentation pattern. c) Manually annotated fragmentation pattern for parent ion of $m/z = 713.4339$. d) Structural assignment of fragment ions.

The NMR data of the natural isolated compound of fraction **PE2H_3** corresponds to the synthetic **3.3** as well, as shown in Table 6-1, reinforcing the statement of the identical nature of both compounds. The complete analysis and spectra of both compounds can be found in the Supporting Information 6.5.

Table 6-1: NMR data comparison for 3.3 in MeOD for synthetic one (600 MHz, 101 MHz) and natural isolated fraction PE2H_3 (600 MHz, 151 MHz). The data of the natural isolated fraction PE2H_3 is written in blue.

Position	$\delta^{13}C$ [ppm]		δ^1H [ppm]; (m, J, J)	
	synthetic	natural isolated	synthetic	natural isolated
L-Leu				
COOH	n.o.	176.1	n.o.	-
NH	-	-	n.o.	-
α -CH	54.8	52.2	4.34 (dd, 1H, $J = 9.0, 4.9$ Hz)	4.49 (dd, $J = 9.8, 5.1$ Hz)
β -CH ₂	43.6	41.7	1.85-1.52 (m, 6H, overlay)	1.68 (m), 1.62 (m)

γ-CH	26.1	25.8	1.85-1.52 (m, 6H, overlay)	1.73 (m)
δ-CH₃ a	22.7	22.0	0.95 (d, 3H, <i>J</i> = 6.3 Hz)	0.91 (d, <i>J</i> = 6.5 Hz)
δ-CH₃ b	23.7	23.4	0.92 (d, 3H, <i>J</i> = 6.5 Hz)	0.91 (d, <i>J</i> = 6.5 Hz)
<hr/>				
Dhv 1				
CO	167.5*	167.9	-	-
NH	-	-	n.o.	-
α-C_{quart}	141.9*	141.2	-	-
β-C_{quart}	125.2*	124.8	-	-
γ-CH₃ a	21.0	20.9	2.04 (s, 3H)	2.08 (s)
γ-CH₃ b	21.6	21.6	1.75 (s, 3H)	1.80 (s)
<hr/>				
L-Arg				
CO	173.8*	173.4	-	-
NH	-	-	n.o.	-
α-CH	54.8	54.9	4.41 (dd, 1H, <i>J</i> = 7.4, 6.1 Hz)	4.32 (dd, <i>J</i> = 9.0, 5.1 Hz)
β-CH₂	30.0	29.2	2.09-1.92 (m, 4H, overlay) 1.85-1.52 (m, 6H, overlay)	1.96 (m) 1.85 (m)
γ-CH₂	26.1	26.5	1.85-1.52 (m, 6H, overlay)	1.68 (m)
δ-CH₂	42.1	42.0	3.24-3.15 (m, 2H)	3.20 (t, <i>J</i> = 7.0 Hz)
ε-NH	-	-	n.o.	-
ζ-C_{quart}	158.8*	158.5	-	-
η-NH	-	-	n.o.	-
η-NH₂	-	-	n.o.	-
<hr/>				
Dhv 2				
CO	168.4*	168.0	-	-
NH	-	-	n.o.	-
α-C_{quart}	141.5*	141.8	-	-
β-C_{quart}	124.8*	124.9	-	-
γ-CH₃ a	20.9	21.1	2.00 (s, 3H)	2.02 (s)
γ-CH₃ b	21.5	21.6	1.45 (s, 3H)	1.46 (s)
<hr/>				
L-Phe				
CO	173.4*	173.6	-	-
NH	-	-	n.o.	-
α-CH	56.6	56.9	4.62 (t, 1H, <i>J</i> = 7.8 Hz)	
β-CH₂	38.4	38.3	3.10 (dd, 1H, <i>J</i> = 13.8, 7.7 Hz) 2.99 (dd, 1H, <i>J</i> = 12.7, 9.2 Hz)	
γ-C_{quart}	138.2*	137.9	-	-
	130.4,	130.4,		
CH_{arom}	129.7,	129.7,	7.30-7.18 (m, 5H)	7.28 (m, overlay)
	128.0	128.1		
<hr/>				
iVal				
CO	176.1*	176.0	-	-
CH₂	45.9	45.9	2.09-1.92 (m, 4H, overlay)	2.07 (m)
CH	27.4	27.4	2.09-1.92 (m, 4H, overlay)	1.98 (m)
CH₃ a	22.8	22.8	0.87 (d, 3H, <i>J</i> = 6.6 Hz)	0.89 (d, <i>J</i> = 6.6 Hz)
CH₃ b	22.7	22.7	0.82 (d, 3H, <i>J</i> = 6.6 Hz)	0.85 (d, <i>J</i> = 6.6 Hz)

The second ion found in the sample of **PE2H_3** with $m/z = 715.4496$ $[M+H]^+$ (Figure 6-7, orange EIC) has a mass difference of two compared to the ion $m/z = 713.4346$. The associated structure of **3.3** suggests one double bond less, probably of one of the dehydrovalines. The manually annotated fragmentation pattern for parent ion of $m/z = 715.4496$ and its structural assignment (Figure 6-9) reinforce that theory. Nevertheless, the determination on which valine is the dehydrovaline and which is a normal valine was not possible due to the co-elution and co-fragmentation of both ions, the low intensities, and the incomplete stepwise pattern. We found a fragment with $m/z = 329.1860$ that corresponds to the fragment where the N-terminal valine moiety is a dehydrovaline. We also found a fragment ion with $m/z = 237.1324$, which fits a dehydrovaline closer to the C-terminus. Therefore, the position of the dehydrovaline in the linear peptide cannot be determined. It concludes the existence of a derivative containing a valine and dehydrovaline moiety in the same molecule and corresponds to either structural proposal **3.15** or **3.16**.

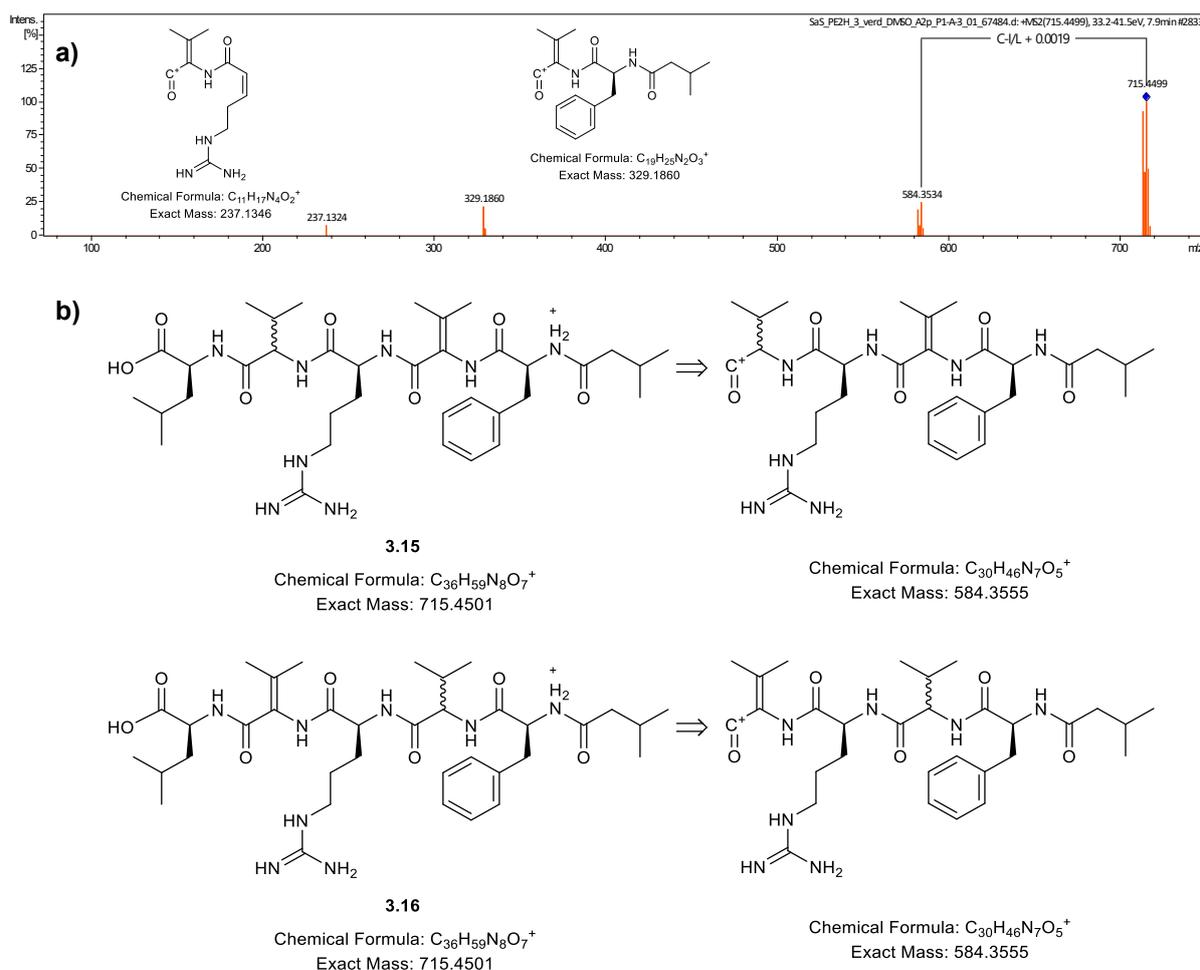


Figure 6-9: MS² data of parent ion $m/z = 713.4339$ of natural isolated PE2H_3. a) Manually annotated fragmentation pattern for parent ion of $m/z = 713.4339$. b) Structural assignment of fragment ions.

The MS and MS² data comparison of synthetic **3.4** to natural isolated fraction **P13-71** confirms the proposed structure shown in Figure 6-10 based on a) the correlation of the retention time with 7.6-7.8 min for both, b) the fragmentation pattern and intensities, which are identical and c) and d) the fragmentation pattern and structural assignment of the parent ion $m/z = 713.4350$, which are identical for both compounds. This is a very good indication that the chosen stereogenic center of the fatty acid is the correct one.

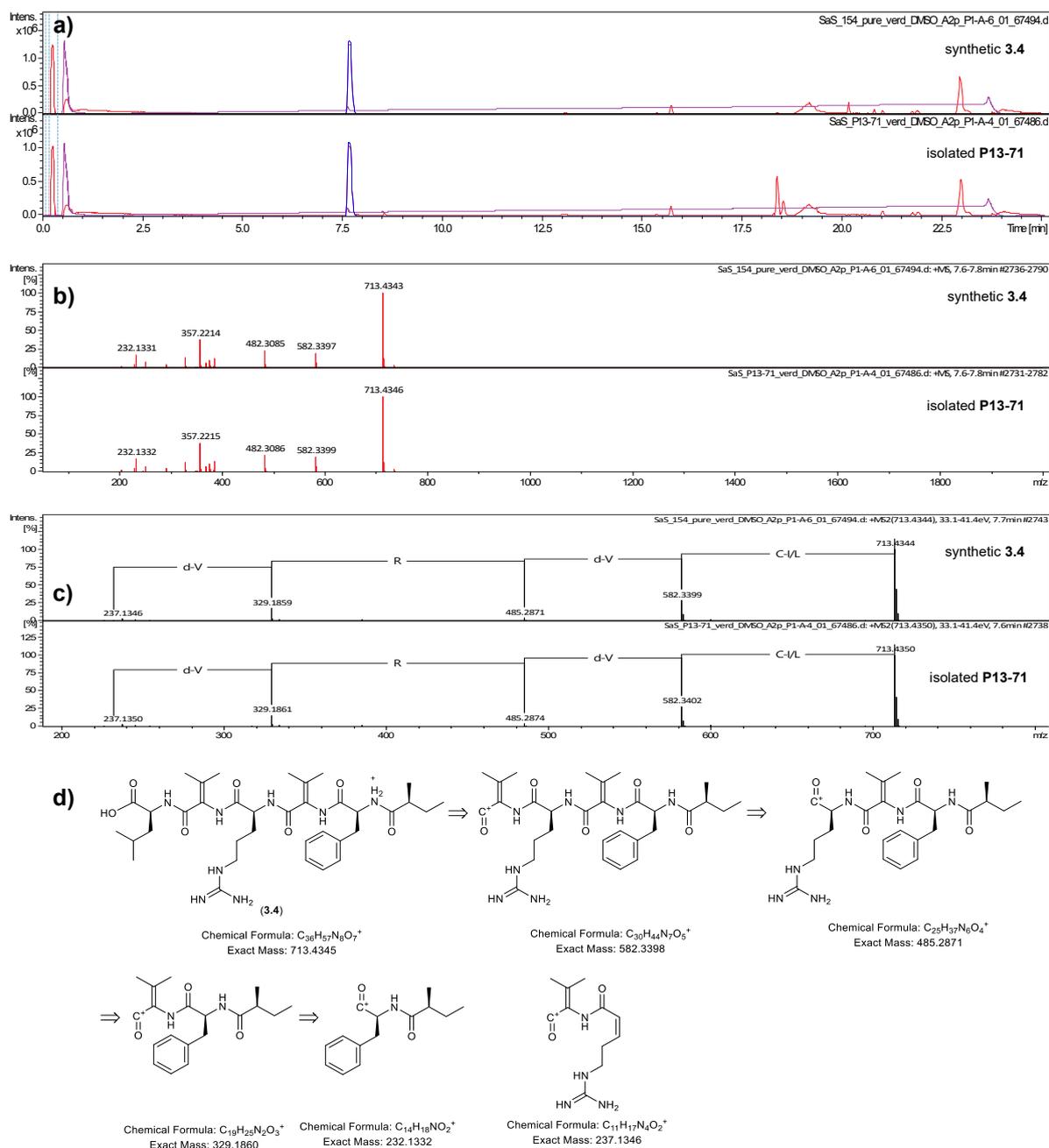


Figure 6-10: Comparison of MS and MS² data of natural isolated compound P13-71 and synthetic 3.4. a) Comparison of retention time. b) Comparison of fragmentation pattern. c) Manually annotated fragmentation pattern for parent ion of m/z = 713.4350. d) Structural assignment of fragment ions.

The NMR data of the natural isolated compound of fraction **P13-71** corresponds to the synthetic **3.4**, as shown in Table 6-2. There are no major deviations regarding the chemical shifts. The signals of the fatty acid side chain, and especially that of the stereogenic center are identical, confirming the chosen *S* configuration combined with the identical fragmentation pattern. Further data and spectra of both compounds can be found in the Supporting Information 6.5.

Table 6-2: NMR data comparison for 3.4 in MeOD for synthetic one (600 MHz, 101 MHz) and natural isolated fraction P13-71 (600 MHz, 151 MHz). The data of the natural isolated fraction P13-71 is written in blue.

Position	$\delta^{13}\text{C}$ [ppm]		$\delta^1\text{H}$ [ppm]; (m, J, J)	
	synthetic	natural isolated	Synthetic	natural isolated
L-Leu				
COOH	176.9	176.2	n.o.	-
NH	-	-	n.o.	-
α -CH	52.9	52.3	4.45 (dd, 1H, $J = 9.4, 5.3$ Hz)	4.48 (dd, $J = 9.8, 5.0$ Hz)
β -CH ₂	42.11	41.7	1.69-1.60 (m, 2H)	1.68 (m), 1.63 (m)
γ -CH	25.9	25.8	1.77-1.69 (m, 3H, overlay)	1.73 (m)
δ -CH ₃ a	22.2	22.0	0.92 (d, 3H, $J = 6.4$ Hz)	0.90 (d, $J = 6.4$ Hz)
δ -CH ₃ b	23.4	23.4	0.91 (d, 3H, $J = 6.4$ Hz)	0.90 (d, $J = 6.4$ Hz)
Dhv 1				
CO	168.0	167.9	-	-
NH	-	-	n.o.	-
α -C _{quart}	141.5	141.3	-	-
β -C _{quart}	124.9	124.9	-	-
γ -CH ₃ a	21.1	20.9	2.07 (s, 3H)	2.08 (s)
γ -CH ₃ b	21.61	21.6	1.79 (s, 3H)	1.80 (s)
L-Arg				
CO	173.44	173.4	-	-
NH	-	-	n.o.	-
α -CH	54.9	54.9	4.34 (dd, 1H, $J = 8.5, 5.7$ Hz)	4.30 (dd, $J = 8.7, 5.4$ Hz)
β -CH ₂	29.4	29.2	2.06-2.00 (m, 1H) 1.90-1.82 (m, 1H)	2.04 (m) 1.86 (m)
γ -CH ₂	26.4	26.5	1.77-1.69 (m, 3H, overlay)	1.71 (m), 1.69 (m)
δ -CH ₂	42.06	42.0	3.21 (t, 2H, $J = 7.2$ Hz)	3.21 (t, $J = 7.1$ Hz)
ϵ -NH	-	-	n.o.	-
ζ -C _{quart}	158.7	158.6	-	-
η -NH	-	-	n.o.	-
η -NH ₂	-	-	n.o.	-
Dhv 2				
CO	167.8	168.0	-	-

NH	-	-	n.o.	-
α -C _{quart}	141.4	141.7	-	-
β -C _{quart}	124.6	124.5	-	-
γ -CH ₃ a	20.9	21.1	2.02 (s, 3H)	2.02 (s)
γ -CH ₃ b	21.57	21.7	1.47 (s, 3H)	1.45 (s)
<hr/>				
L-Phe				
CO	173.38	173.4	-	-
NH	-	-	n.o.	-
α -CH	56.7	56.8	4.59 (t, 1H, $J=7.9$ Hz)	4.58 (t, $J=7.9$ Hz) 3.09 (dd, 1H, $J=13.7$, 7.8 Hz)
β -CH ₂	38.4	38.4	3.10 (dd, 1H, $J=13.7$, 7.7 Hz) 3.02 (dd, 1H, $J=13.7$, 7.9 Hz)	3.01 (dd, 1H, $J=13.7$, 7.8 Hz)
γ -C _{quart}	138.0	137.9	-	-
CH _{arom}	130.4,	130.4,	7.28 (d, 4H, $J=4.6$ Hz) 7.24-7.20 (m, 1H)	7.29 (overlay) 7.22 (m)
	129.6,	129.7,		
	128.0	128.0		
<hr/>				
SMB				
CO	179.9	180.0	-	-
CH	43.1	43.1	2.30-2.23 (m, 1H)	2.27 (m)
CH ₃	17.9	17.9	1.00 (d, 3H, $J=6.8$ Hz)	1.0 (d, $J=6.9$ Hz)
CH ₂	28.2	28.2	1.60-1.52 (m, 1H)	1.56 (m)
			1.40-1.31 (m, 1H)	1.36 (m)
CH ₃	12.3	12.4	0.86 (t, 3H, $J=7.4$ Hz)	0.86 (t, $J=7.4$ Hz)

6.2.3 Activities

The compounds **3.3**, **3.4**, **3.8** and **3.9** were tested against Gram-negative bacteria (*E. coli*, *P. aeruginosa*), against fungi (*A. flavus*, *C. albicans*) as well as Gram-positive (*S. aureus*) and a surrogate of *M. tuberculosis* (*M. smegmatis*) (Table 6-3). The screening was performed at the Fraunhofer Institute (Dr. M. Marner and team). By internal standards, activities of ≥ 128 $\mu\text{g/mL}$ are considered inactive.

Table 6-3: Antimicrobial activity of Cryopeptides.

compounds	Val	Fatty acid
3.9	L-Val	iVal
3.3	Dhv	iVal
3.8	D-Val	iVal
3.4	Dhv	SMB

	MIC [$\mu\text{g/mL}$]			
	3.9	3.3	3.8	3.4
<i>E. coli</i> ATCC 35218	>128	>128	>128	>128
<i>P. aeruginosa</i> ATCC 27853	>128	>128	>128	>128
<i>A. flavus</i> ATCC 9170	>128	>128	>128	>128
<i>C. albicans</i> FH2173	>128	>128	>128	>128
<i>S. aureus</i> ATCC 33592	>128	>128	>128	>128
<i>M. smegmatis</i> ATCC 607	>128	>128	>128	>128

All tested compounds (**3.3**, **3.4**, **3.8** and **3.9**) showed no antibiotic activity as shown in Table 6-3.

As a next step, they were tested against human cysteine proteases cathepsin B (CatB) and L (CatL), as well as the parasitic protease of *Trypanosoma brucei rhodesiense* rhodesain by Dr. T. Schirmeister at the University of Mainz.

The tested compounds (**3.3**, **3.4**, **3.8** and **3.9**) showed no protease activity against cathepsin B or L. Only compound **3.8** showed inhibition of rhodesain with an IC_{50} of 13.7 μM . It needs to be considered that it could have functioned as a substrate as well. First LC-MS results showed no cleavage products, but further investigations are needed.

6.3 Summary and Outlook

Dehydrogenated amino acids are often found in natural products and show promising properties as antibacterial compounds.³⁴ To this date it is not known how and when the dehydrogenation process occurs. The biosynthetic gene cluster, as seen by the example of glidopeptin A, gives no indication on the formation or incorporation of these dehydrogenated

amino acids either.³⁷ Cryopeptides were detected in extracts and subsequent molecular network of *Pedobacter cryoconitis*. The pentapeptides contain two dehydrogenated valines and a fatty acid residue. Due to the scarcity of the natural isolated material and for complete structure elucidation, a solid-phase peptide synthesis (SPPS) was developed to gain access to the natural compounds and derivatives. The synthesis was successful with varying yields between 8-30% for the four pentapeptides (**3.3**, **3.4**, **3.8** and **3.9**). With the elimination of the extraction during the work up the yield was more than doubled, from 8% for **3.3** and 17% for **3.4**. The synthetic compounds **3.3** and **3.4** allowed for the validation of the proposed structures of the natural isolated compounds **PE2H_3** and **P13-71**, respectively. The structure was not only fully elucidated but the amount of material was enough to conduct some antimicrobial screenings. However, neither antibacterial activity for the pentapeptides (**3.3**, **3.4**, **3.8** and **3.9**) nor considerable protease activity could be observed (Table 6-3). To determine the origin of the dehydrogenation process, two mimics, the valyl-SNACs, were successfully synthesized with overall yields of 68% and 39% for the *R*- (**3.6**) and *S*-enantiomer (**3.7**) respectively (Scheme 6-1). A gene encoding a protein that might catalyze this reaction was identified clustered with the NRPS gene putatively encoding the biosynthesis of the cryopeptides. The tests regarding the dehydrogenation process are still underway and results are not available to this date. The enzyme assay needs further optimization and the tests will be carried out as soon as the test is working. Therefore, the determination of the origin of the dehydrogenation could not be completed yet. All needed substrates are synthesized and characterized, and the tests will be conducted as soon as the assay optimization is completed.

6.4 Experimental

General procedures of the Cryopeptide Syntheses⁴¹

Coupling of amino acids and fatty acid:

Unless otherwise noted, all reactions were carried out in a custom-built solid phase peptide synthesis vessel with a G2 filter and a diameter of 3, 4 or 5 cm at room temperature. Argon was used for agitation of the resin. 20% piperidine in DMF and the cleavage cocktail were freshly prepared on the day of use.

The Fmoc-protected amino acid or fatty acid (3.0 equiv) and HATU (2.9 equiv), dissolved in a small amount of DMF, were added to the swelled resin (1.0 equiv). DIPEA (6.0 equiv) was added and the resin was agitated for 1-5 hours.

Each coupling step was monitored as described in the general method part for the LC-MS sample preparation. After each coupling step was complete as indicated by LC-MS, Fmoc-deprotection was carried out using 20% piperidine in DMF.

LC-MS sample preparation

The reaction progress of each coupling of the Fmoc-protected amino acids was monitored using LC-MS. A few beads of the resin were sampled in a 2 mL SPPS syringe, washed once with DMF and then two to three times with DCM. The vessel was closed and 20% HFIP in DCM was added (1-1.5 mL), which changed the color of the beads from yellow-orange to a dark red, which again faded over time. The mixture was shaken for 15-30 min and the filtrate was used for LC-MS measurement.

Fmoc-deprotection

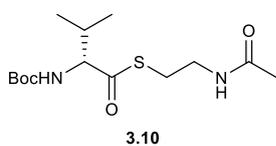
The mixture was filtered and the remaining resin was washed 5 times with DMF. 20% piperidine in DMF (20-30 mL) was added. After agitation for 2-5 min it was filtrated, rinsed with DMF and the process was repeated four more times. Afterwards, the resin was washed with DMF three times.

⁴¹ Lit: W. Chan, P. White, Fmoc Solid Phase Peptide Synthesis Practical Approach, Oxford University Press, Oxford, 1999.

Cleavage from the resin

To the washed resin, 95% TFA was added, coloring the mixture a dark purple. The resin was agitated for 30-60 min after which the supernatant was drained and the process was repeated once more. The combined filtrates were concentrated under reduced pressure and then dried further using lyophilization.

Preparation of S-(2-acetamidoethyl) (R)-2-((tert-butoxycarbonyl)amino)-3-methylbutanethioate (3.10)



N-Boc-D-valine (511 mg, 2.35 mmol), EDC · HCl (496 mg, 2.59 mmol) and HOAt (358 mg, 2.63 mmol) were dissolved in anhydrous DCM. *N*-Acetylcysteamin (95%, 250 μ L, 2.23 mmol) was added, followed by triethylamine (359 μ L, 2.59 mmol). The mixture was stirred at room temperature for 25 h. After dilution with EA, the organic phase was washed three times with citric acid (10% w/v), once with sat. NaHCO₃ solution and once with brine. The organic layer was dried over MgSO₄ and concentrated *in vacuo*. Purification by semi preparative HPLC (32-45% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 3 mL/min) yielded **3.10** as a colorless syrup (506 mg, 1.59 mmol, 68%).

¹H-NMR (CDCl₃, 400 MHz): δ_{H} [ppm] = 5.93 (s, 1H, NH-Ac), 4.98 (d, 1H, *J* = 9.0 Hz, NH-Boc), 4.23 (dd, 1H, *J* = 8.9, 4.6 Hz, α -CH Val), 3.46 (ddd, 1H, *J* = 13.2, 13.2, 6.5 Hz, CH₂-NH a), 3.38 (ddd, 1H, *J* = 13.1, 13.1, 6.2 Hz, CH₂-NH b), 3.04 (ddd, 2H, *J* = 6.4, 6.4, 2.9 Hz, S-CH₂), 2.29-2.19 (m, 1H, β -CH Val), 1.95 (s, 3H, CH₃ Ac), 1.45 (s, 9H, CH₃ Boc), 0.99 (d, 3H, *J* = 6.8 Hz, γ -CH₃ a Val), 0.87 (d, 3H, *J* = 6.8 Hz, γ -CH₃ b Val).

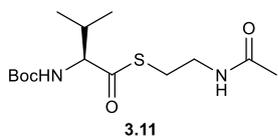
¹³C-NMR (CDCl₃, 100 MHz): δ_{C} [ppm] = 201.9 (CO Val), 170.5 (CO Ac), 155.8 (CO Boc), 80.6 (C_{quart} Boc), 65.8 (α -CH Val), 39.5 (CH₂-NH), 30.9 (β -CH Val), 28.50 (S-CH₂), 28.45 (CH₃ Boc), 23.3 (CH₃ Ac), 19.5 (γ -CH₃ a Val), 17.1 (γ -CH₃ b Val).

Additional found signals: δ_{H} [ppm] = 8.08 (FA), 5.29 (DCM). δ_{C} [ppm] = /

UHR-MS (ESI-TOF) *m/z* calcd for C₁₄H₂₇N₂O₄S: 319.1686 [M+H]⁺; found: 319.1688 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{25} = -12.7^{\circ}$ (c = 1.02, CHCl₃)

Preparation of *S*-(2-acetamidoethyl) (*S*)-2-((tert-butoxycarbonyl)amino)-3-methylbutanethioate (**3.11**)



N-Boc-L-valine (512 mg, 2.36 mmol), EDC · HCl (499 mg, 2.60 mmol) and HOAt (358 mg, 2.63 mmol) were dissolved in anhydrous DCM. *N*-Acetylcysteamin (95%, 262 μ L, 2.34 mmol) was added, followed by triethylamine (359 μ L, 2.59 mmol). The mixture was stirred at room temperature for 22 h. After dilution with EA, the the organic phase was washed three times with citric acid (10% w/v), once with sat. NaHCO₃ solution and once with brine. The organic layer was dried over MgSO₄ and concentrated *in vacuo*. Purification by semi preparative HPLC (32-45% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 3 mL/min) yielded **3.11** as a colorless syrup (312 mg, 0.980 mmol, 42%).

¹H-NMR (CDCl₃, 400 MHz): δ_{H} [ppm] = 6.00 (s, 1H, NH-Ac), 4.98 (d, 1H, J = 8.7 Hz, NH-Boc), 4.23 (dd, 1H, J = 8.9, 4.8 Hz, α -CH Val), 3.47 (ddd, 1H, J = 13.0, 13.0, 6.4 Hz, CH₂-NH a), 3.39 (ddd, 1H, J = 13.0, 13.0, 6.3 Hz, CH₂-NH b), 3.04 (ddd, 2H, J = 6.3, 6.3, 3.6 Hz, S-CH₂), 2.30-2.19 (m, 1H, β -CH Val), 1.97 (s, 3H, CH₃ Ac), 1.46 (s, 9H, CH₃ Boc), 0.99 (d, 3H, J = 6.8 Hz, γ -CH₃ a Val), 0.88 (d, 3H, J = 6.8 Hz, γ -CH₃ b Val).

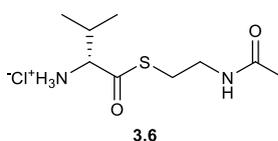
¹³C-NMR (CDCl₃, 100 MHz): δ_{C} [ppm] = 201.9 (CO Val), 170.7 (CO Ac), 155.8 (CO Boc), 80.6 (C_{quart} Boc), 65.8 (α -CH Val), 39.6 (CH₂-NH), 30.9 (β -CH Val), 28.5* (S-CH₂), 28.5 (CH₃ Boc), 23.2 (CH₃ Ac), 19.5 (γ -CH₃ a Val), 17.2 (γ -CH₃ b Val).

Additional found signals: δ_{H} [ppm] = 8.07 (FA), 5.30 (DCM). δ_{C} [ppm] = /

UHR-MS (ESI-TOF) m/z calcd for C₁₄H₂₇N₂O₄S: 319.1686 [M+H]⁺; found: 319.1687 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{25} = +16.3^{\circ}$ (c = 0.80, CHCl₃)

Preparation of (*R*)-1-((2-acetamidoethyl)thio)-3-methyl-1-oxobutan-2-aminium chloride (**3.6**)



Compound **3.10** (480 mg, 1.51 mmol) was dissolved in 4 M HCl in 1,4-dioxane (30 mL) and stirred at room temperature for 70 min in a sealed flask. An argon stream removed residual HCl and the solvent

was removed *in vacuo*. Product **3.6** was obtained as a colorless solid (386 mg, 1.52 mmol, quantitavly) without further purification.

¹H-NMR (DMSO-*d*₆, 400 MHz): δ_{H} [ppm] = 8.63 (bd, 3H, NH₃⁺), 8.18 (t, 1H, *J* = 5.8 Hz, NH-Ac), 4.08 (t, 1H, *J* = 5.1 Hz, α -CH Val), 3.22 (ddd, 2H, *J* = 6.3, 6.3, 6.3 Hz, CH₂-NH) 3.08 (ddd, 1H, *J* = 13.4, 6.7, 6.7 Hz, S-CH₂ a), 3.02 (ddd, 1H, *J* = 13.2, 6.5, 6.5 Hz, S-CH₂ b), 2.26-2.13 (m, 1H, β -CH Val), 1.79 (s, 3H, CH₃ Ac), 0.98 (d, 3H, *J* = 7.4 Hz, γ -CH₃ a Val), 0.96 (d, 3H, *J* = 7.3 Hz, γ -CH₃ a Val).

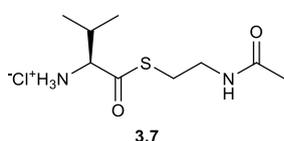
¹³C-NMR (DMSO-*d*₆, 100 MHz): δ_{C} [ppm] = 196.0 (CO Val), 169.4 (CO Ac), 63.4 (α -CH Val), 37.8 (CH₂-NH), 30.0 (β -CH Val), 28.4 (S-CH₂), 22.5 (CH₃ Ac), 18.1 (γ -CH₃ a Val), 17.6 (γ -CH₃ b Val).

Additional found signals: δ_{H} [ppm] = 4.49, 3.56 (1,4-dioxane). δ_{C} [ppm] = 66.4 (1,4-dioxane).

UHR-MS (ESI-TOF) *m/z* calcd for C₉H₁₉N₂O₂S: 219.1162 [M+H]⁺; found: 219.1164 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{24} = -50.7^{\circ}$ (c = 1.40, H₂O)

Preparation of (S)-1-((2-acetamidoethyl)thio)-3-methyl-1-oxobutan-2-aminium chloride (3.7)



The solution of Compound **3.11** (166 mg, 0.521 mmol) in 4 M HCl in 1,4-dioxane (10 mL) was stirred at room temperature for 60 min in a sealed flask. An argon stream removed residual HCl and the solvent was removed *in vacuo*. Product **3.7** was obtained as a colorless solid (123 mg, 0.483 mmol, 93%) without further purification.

¹H-NMR (DMSO-*d*₆, 400 MHz): δ_{H} [ppm] = 8.67 (bd, 3H, NH₃⁺), 8.20 (t, 1H, *J* = 5.4 Hz, NH-Ac), 4.06 (t, 1H, *J* = 5.0 Hz, α -CH Val), 3.22 (ddd, 2H, *J* = 6.2, 6.2, 6.2 Hz, CH₂-NH) 3.07 (ddd, 1H, *J* = 13.3, 6.5, 6.5 Hz, S-CH₂ a), 3.01 (ddd, 1H, *J* = 13.4, 6.8, 6.8 Hz, S-CH₂ b), 2.25-2.13 (m, 1H, β -CH Val), 1.79 (s, 3H, CH₃ Ac), 0.98 (d, 3H, *J* = 6.6 Hz, γ -CH₃ a Val), 0.96 (d, 3H, *J* = 6.7 Hz, γ -CH₃ a Val).

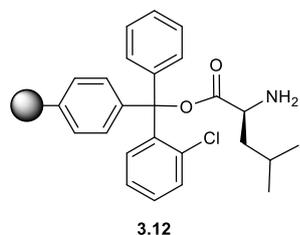
¹³C-NMR (DMSO-*d*₆, 100 MHz): δ_{C} [ppm] = 195.9 (CO Val), 169.4 (CO Ac), 63.5 (α -CH Val), 37.8 (CH₂-NH), 30.0 (β -CH Val), 28.4 (S-CH₂), 22.5 (CH₃ Ac), 18.0 (γ -CH₃ a Val), 17.6 (γ -CH₃ b Val).

Additional found signals: δ_{H} [ppm] = 4.66. δ_{C} [ppm] = /

UHR-MS (ESI-TOF) m/z calcd for $C_9H_{19}N_2O_2S$: 219.1162 $[M+H]^+$; found: 219.1163 $[M+H]^+$

Specific rotation $[\alpha]_D^{24} = +54.0^\circ$ ($c = 1.61$, H_2O)

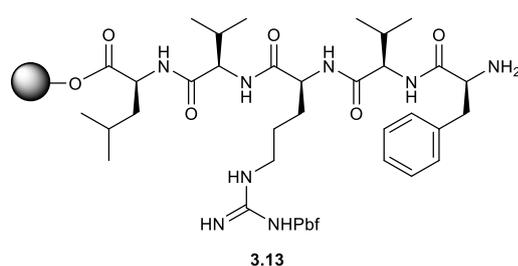
Preparation of 2CT-L-leucine-NH₂ (3.12)



2-Chlorotrityl chloride resin ($n = 1.55$ mmol/g, 5.93 g, 9.19 mmol) was swelled in DCM (50 mL) for 10 min. After removal of the solvent, a solution of Fmoc-L-leucine-OH (4.24 g, 12.0 mmol) and DIPEA (5.1 mL, 30 mmol) in DCM (50 mL) was added and agitated for 1 h. The solvent was drained and the resin was washed twice with DMF.

The capping mixture (50 mL), consisting of DCM/MeOH/DIPEA (80:15:5), was added and it was agitated for 10 min. The mixture was filtered and the capping was repeated once more. Fmoc-deprotection was done with 25% piperidine in DMF (50 mL) for 3 min and the process was performed five times total. The resin was washed four times with DMF, three times with isopropanol and three times with *n*-heptane. It was sucked dry and dried further *in vacuo* for 12 h. Resin **3.12** was stored under argon at 4 °C. The loading was determined by weight gain to be $n = 1.19$ mmol/g (77%).

Preparation of 2CT-L-leucine-D-valine-L-arginine(Pbf)-D-valine-L-phenylalanine-NH₂ (3.13)

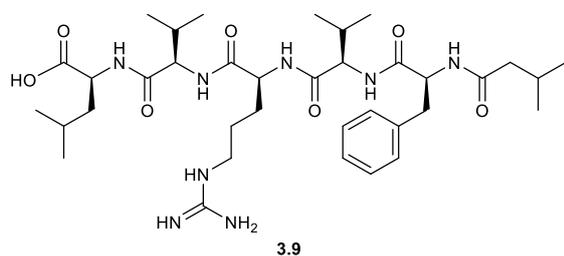


Resin **3.12** ($n = 1.19$ mmol/g, 0.62 g, 0.738 mmol) was swelled in DMF for 30 min. The solvent was removed and a solution of Fmoc-D-valine-OH (0.752 g, 2.22 mmol) and HATU (0.814 g, 2.14 mmol) in DMF was added, followed by DIPEA

(753 μ L, 4.43 mmol) and more DMF. The mixture was agitated for 2 h. After Fmoc-deprotection, Fmoc-L-arginine(Pbf)-OH (1.436 g, 2.213 mmol) and HATU (0.818 g, 2.15 mmol), dissolved in a small amount of DMF, were added, followed by DIPEA (753 μ L, 4.43 mmol) and more DMF. The mixture was agitated for 3 h. After Fmoc-deprotection, Fmoc-D-valine-OH (0.760 g, 2.24 mmol) and HATU (0.812 g, 2.14 mmol) dissolved in DMF, were added, followed by DIPEA (753 μ L, 4.43 mmol) and more DMF. The mixture was agitated for 2.5 h. After Fmoc-deprotection, Fmoc-L-phenylalanine-OH (0.859 g, 2.22 mmol) and HATU

(0.815 g, 2.14 mmol) dissolved in DMF, were added, followed by DIPEA (753 μ L, 4.43 mmol) and more DMF. The mixture was agitated for 1.5 h. After Fmoc-deprotection the supernatant was drained, the resin was washed five times with DMF, three times with isopropanol and three times with *n*-heptane. It was sucked dry and dried further *in vacuo* for 12 h. Resin **3.13** was stored under argon at 4 °C. The loading was determined by weight to be $n = 0.624$ mmol/g (87%).

Preparation of (3-methylbutanoyl)-L-phenylalanyl-D-valyl-L-arginyl-D-valyl-L-leucine (**3.9**)



Resin **3.13** ($n = 0.624$ mmol/g, 1.08 g, 0.674 mmol) was swelled in DMF for 30 min. The solvent was removed and HATU (0.747 g, 1.96 mmol) dissolved in DMF was added, followed by isovaleric acid (221 μ L, 2.01 mmol),

DIPEA (686 μ L, 4.03 mmol) and more DMF. The mixture was agitated for 4 h. The supernatant was drained and the resin was washed three times each with DMF and DMC. The cleavage was performed as described in the general method section. The obtained crude product was dissolved in 20% acetic acid and extracted three times with chloroform. The combined organic phases were dried over MgSO_4 and concentrated *in vacuo*. Purification using semi preparative HPLC (32-50-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 3 mL/min) yielded **3.9** as a colorless powder (152 mg, 0.212 mmol, 34%).

¹H-NMR (MeOD, 600 MHz): δ_{H} [ppm] = 7.30-7.18 (m, 5H, CH_{arom} Phe), 4.67 (t, 1H, $J = 7.7$ Hz, α -CH Phe), 4.36 (dd, 1H, $J = 10.1, 4.5$ Hz, α -CH Leu), 4.29 (t, 1H, $J = 6.6$ Hz, α -CH Arg), 4.24 (d, 1H, $J = 8.4$ Hz, α -CH Val 2), 3.93 (d, 1H, $J = 6.8$ Hz, α -CH Val 1), 3.17-3.11 (m, 2H, δ -CH₂ Arg), 3.03 (dd, 1H, $J = 13.3, 8.1$ Hz, β -CH₂ a Phe), 2.98 (dd, 1H, $J = 13.6, 7.3$ Hz, β -CH₂ b Phe), 2.13-2.00 (m, 4H, β -CH Val 1, β -CH Val 2, CH₂ iVal), 2.00-1.93 (m, 1H, CH iVal), 1.90-1.76 (m, 2H, β -CH₂ Arg), 1.73-1.65 (m, 2H, γ -CH Leu, γ -CH₂ a Arg), 1.65-1.57 (m, 3H, β -CH₂ Leu, γ -CH₂ b Arg), 0.95 (d, 3H, $J = 6.1$ Hz, δ -CH₃ a Leu), 0.94 (d, 3H, $J = 6.1$ Hz, δ -CH₃ b Leu), 0.92 (d, 3H, $J = 6.6$ Hz, γ -CH₃ a Val 2), 0.87 (d, 3H, $J = 6.6$ Hz, CH₃ a iVal), 0.84 (d, 3H, $J = 6.8$ Hz, γ -CH₃ b Val 2), 0.84 (d, 3H, $J = 6.5$ Hz, CH₃ b iVal), 0.74 (d, 3H, $J = 6.8$ Hz, γ -CH₃ a Val 1), 0.70 (d, 3H, $J = 6.8$ Hz, γ -CH₃ b Val 1).

¹³C-NMR (MeOD, 100 MHz): δ_c [ppm] = 179.5 (COOH Leu), 175.3 (CO iVal), 174.4 (CO Val 1), 174.2 (CO Phe), 173.5 (CO Arg), 172.6 (CO Val 2), 158.7 (ζ -C_{quart} Arg), 138.0 (γ -C_{quart} Phe), 130.4, 129.6, 127.9 (C_{arom} Phe), 61.0 (α -CH Val 1), 60.5 (α -CH Val 2), 56.5 (α -CH Phe), 55.4 (α -CH Arg), 54.4 (α -CH Leu), 46.0 (CH₂ iVal), 42.7 (β -CH₂ Leu), 42.2 (δ -CH₂ Arg), 39.3 (β -CH₂ Phe), 31.5 (β -CH Val 2), 30.5 (β -CH Val 1), 30.0 (β -CH₂ Arg), 27.4 (CH iVal), 26.3 (γ -CH Leu), 26.0 (γ -CH₂ Arg), 23.7 (δ -CH₃ a Leu), 22.9 (CH₃ a iVal), 22.7 (CH₃ b iVal), 22.0 (δ -CH₃ b Leu), 20.0 (γ -CH₃ a Val 2), 19.6 (γ -CH₃ b Val 1), 18.8 (γ -CH₃ b Val 2), 18.5 (γ -CH₃ a Val 1).

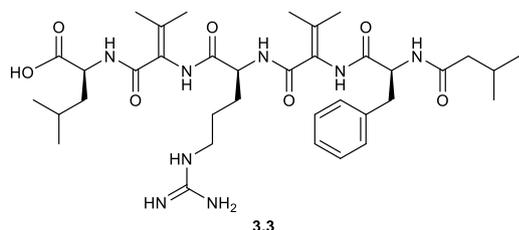
Additional found signals: δ_H [ppm] = 4.88 (H₂O), 4.60 (bs), 2.66 (DMSO), NH signals, COOH Leu and NH/NH₂ Arg were not observed. δ_c [ppm] = 40.4 (DMSO).

The differentiation between Val 1 and Val 2 refers to the spin system and not the position in the molecule. Due to the missing NH signals differentiation was not possible using HMBC data.

UHR-MS (ESI-TOF) m/z calcd for C₃₆H₆₁N₈O₇: 717.4658 [M+H]⁺; found: 717.4660 [M+H]⁺

Specific rotation $[\alpha]_D^{21} = +34.1^\circ$ (c = 1.10, CH₃OH)

Preparation of (2-((S)-5-guanidino-2-(3-methyl-2-((S)-2-(3-methylbutanamido)-3-phenylpropanamido)but-2-enamido)pentanamido)-3-methylbut-2-enoyl)-L-leucine (3.3)



Resin **3.12** (n = 1.19 mmol/g, 0.565 g, 0.672 mmol) was swelled in DMF for 30 min. The solvent was drained and the solution of Fmoc-2,3-dehydrovaline-OH (0.683 g, 2.02 mmol) and HATU (0.741 g, 1.95 mmol) in DMF was added, followed

by DIPEA (685 μ L, 4.03 mmol) and more DMF. The mixture was agitated for 4 h, succeeded by Fmoc-deprotection. Fmoc-L-arginine(Pbf)-OH (1.315 g, 2.027 mmol) and HATU (0.745 g, 1.96 mmol) dissolved in DMF, were added, followed by DIPEA (685 μ L, 4.03 mmol) and more DMF. The mixture was agitated for 4 h after which LC-MS indicated no complete conversion. The coupling step was repeated. Fmoc-L-arginine(Pbf)-OH (0.439 g, 0.677 mmol) and HATU (0.231 g, 0.608 mmol) dissolved in a small amount of DMF, were added to the resin, followed by DIPEA (228 μ L, 1.34 mmol) and more DMF. The mixture was agitated for 3 h. After Fmoc-deprotection, Fmoc-2,3-dehydrovaline-OH (0.685 g, 2.03 mmol) and HATU (0.745 g, 1.96 mmol), dissolved in DMF, were added, followed by DIPEA (685 μ L, 4.03 mmol) and more

DMF. The mixture was agitated for 4.5 h. After Fmoc-deprotection, Fmoc-L-phenylalanine-OH (0.788 g, 2.03 mmol) and HATU (0.746 g, 1.96 mmol), dissolved in DMF, were added, followed by DIPEA (685 μ L, 4.03 mmol) and more DMF. The mixture was agitated for 3 h. After Fmoc-deprotection the supernatant was drained, HATU (0.740 g, 1.95 mmol) dissolved in DMF was added, followed by isovaleric acid (222 μ L, 2.02 mmol), DIPEA (685 μ L, 4.03 mmol) and more DMF. The mixture was agitated for 6 h, after which LC-MS result indicated complete conversion. The supernatant was drained, and the resin was washed three times each with DMF and DMC. The cleavage was done as described in the general method section. The obtained crude product was dissolved in 20% acetic acid and extracted three times with chloroform. During extraction a third, cloudy layer formed, which would not dissolve in either the organic or aqueous phase. Later, LC-MS analysis showed the free acid to be in this 3rd phase but due to its gel like consistency it was not combined with the other organic phases and discarded. The combined organic phases were dried over MgSO₄ and concentrated *in vacuo*. Purification using semi preparative HPLC (28-45-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 3 mL/min) yielded **3.3** as a colorless oil (38 mg, 0.053 mmol, 8%).

¹H-NMR (MeOD, 600 MHz): δ_{H} [ppm] = 7.30-7.18 (m, 5H, CH_{arom} Phe), 4.62 (t, 1H, J = 7.8 Hz, α -CH Phe), 4.41 (dd, 1H, J = 7.4, 6.1 Hz, α -CH Arg), 4.34 (dd, 1H, J = 9.0, 4.9 Hz, α -CH Leu), 3.24-3.15 (m, 2H, δ -CH₂ Arg), 3.10 (dd, 1H, J = 13.8, 7.7 Hz, β -CH₂ a Phe), 2.99 (dd, 1H, J = 12.7, 9.2 Hz, β -CH₂ b Phe), 2.09-1.92 (m, 4H, CH₂ iVal, CH iVal, β -CH₂ a Arg), 2.04 (s, 3H, γ -CH₃ a Dhv 1), 2.00 (s, 3H, γ -CH₃ a Dhv 2), 1.85-1.52 (m, 6H, β -CH₂ b Arg, γ -CH₂ Arg, β -CH₂ Leu, γ -CH Leu), 1.75 (s, 3H, γ -CH₃ b Dhv 1), 1.45 (s, 3H, γ -CH₃ b Dhv 2), 0.95 (d, 3H, J = 6.3 Hz, δ -CH₃ a Leu), 0.92 (d, 3H, J = 6.5 Hz, δ -CH₃ b Leu), 0.87 (d, 3H, J = 6.6 Hz, CH₃ a iVal), 0.82 (d, 3H, J = 6.6 Hz, CH₃ b iVal).

¹³C-NMR (MeOD, 100 MHz): δ_{C} [ppm] = 176.1* (CO iVal), 173.8* (CO Arg), 173.4* (CO Phe), 168.4* (CO Dhv 2), 167.5* (CO Dhv 1), 158.8* (ζ -C_{quart} Arg), 141.9* (α -C_{quart} Dhv 1), 141.5* (α -C_{quart} Dhv 2), 138.2* (γ -C_{quart} Phe), 125.2* (β -C_{quart} Dhv 1), 124.8* (β -C_{quart} Dhv 2), 130.4, 129.7, 128.0 (CH_{arom} Phe), 56.6 (α -CH Phe), 54.8 (α -CH Arg, α -CH Leu), 45.9 (CH₂ iVal), 43.6 (β -CH₂ Leu), 42.1 (δ -CH₂ Arg), 38.4 (β -CH₂ Phe), 30.0 (β -CH₂ Arg), 27.4 (CH iVal), 26.1 (γ -CH₂ Arg, γ -CH Leu), 23.7 (δ -CH₃ b Leu), 22.8 (CH₃ a iVal), 22.7 (δ -CH₃ a Leu, CH₃ b iVal), 21.6 (γ -CH₃ b Dhv 1), 21.5 (γ -CH₃ b Dhv 2), 21.0 (γ -CH₃ a Dhv 1), 20.9 (γ -CH₃ a Dhv 2).

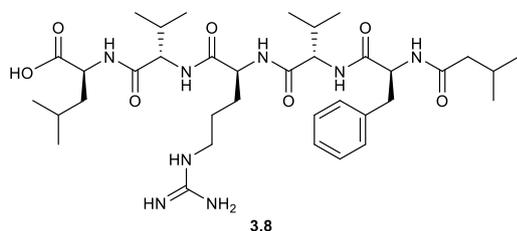
Additional found signals: δ_H [ppm] = 4.87 (H₂O), 3.01, 2.66 (DMSO), NH signals, COOH Leu and NH/NH₂ Arg were not observed. δ_C [ppm] = 40.4 (DMSO), COOH Leu was not observed.

The differentiation between Dhv 1 and Dhv 2 refers to the spin system and not the position in the molecule. Due to the missing NH signals differentiation was not possible using HMBC data.

UHR-MS (ESI-TOF) m/z calcd for C₃₆H₅₇N₈O₇: 713.4345 [M+H]⁺; found: 713.4349 [M+H]⁺

Specific rotation $[\alpha]_D^{21} = +16.7^\circ$ (c = 1.05, CH₃OH)

Preparation of (3-methylbutanoyl)-L-phenylalanyl-L-valyl-L-arginyl-L-valyl-L-leucine (3.8)



Resin **3.12** (n = 1.19 mmol/g, 0.653 g, 0.777 mmol) was swelled in DMF for 30 min. After removal of the solvent, a solution of Fmoc-L-valine-OH (0.793 g, 2.34 mmol) and HATU (0.856 g, 2.25 mmol) in DMF, was added, followed by DIPEA (793 μ L, 4.66 mmol)

and more DMF. The mixture was agitated for 2.5 h, succeeded by Fmoc-deprotection. Fmoc-L-arginine(Pbf)-OH (1.514 g, 2.334 mmol) and HATU (0.859 g, 2.26 mmol) dissolved in DMF, were added, followed by DIPEA (793 μ L, 4.66 mmol) and more DMF. The mixture was agitated for 4.5 h, after which LC-MS indicated incomplete conversion. Fmoc-L-arginine(Pbf)-OH (0.506 g, 0.780 mmol) and HATU (0.276 g, 0.726 mmol) were dissolved in DMF and added to the drained resin. DIPEA (264 μ L, 1.55 mmol) and more DMF followed. The mixture was agitated for 3 h. After Fmoc-deprotection, Fmoc-L-valine-OH (0.790 g, 2.33 mmol) and HATU (0.857 g, 2.25 mmol) dissolved in DMF, were added, followed by DIPEA (793 μ L, 4.66 mmol) and more DMF. The mixture was agitated for 2.5 h, succeeded by Fmoc-deprotection. To the drained resin, Fmoc-L-phenylalanine-OH (0.903 g, 2.33 mmol) and HATU (0.855 g, 2.25 mmol) dissolved in DMF, were added, followed by DIPEA (793 μ L, 4.66 mmol) and more DMF. The mixture was agitated for 2 h. After Fmoc-deprotection, the supernatant was drained, HATU (0.860 g, 2.26 mmol) dissolved in DMF was added, followed by isovaleric acid (256 μ L, 2.33 mmol), DIPEA (264 μ L, 1.55 mmol) and more DMF. The mixture was agitated for 5.5 h. The supernatant was drained and the resin was washed three times each with DMF and DMC. The cleavage was executed as described in the general method section. The obtained crude

product was dissolved in 20% acetic acid and extracted three times with chloroform. During extraction a third, cloudy layer formed, which would not dissolve in either the organic or aqueous phase. Later, LC-MS analysis showed the free acid to be in this 3rd phase but due to its gel like consistency it was not combined with the other organic phases and discarded. The combined organic phases were dried over MgSO₄ and concentrated *in vacuo*. Purification using semi preparative HPLC (28-45-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μm, 250 x 10 mm, flow rate: 3 mL/min) yielded **3.8** as a colorless solid (83.6 mg, 0.12 mmol, 15%).

¹H-NMR (MeOD, 400 MHz): δ_H [ppm] = 7.30-7.15 (m, 5H, CH_{arom} Phe), 4.72 (dd, 2H, *J* = 8.9, 6.2 Hz, α-CH Phe), 4.48 (t, 1H, *J* = 6.8 Hz, α-CH Arg), 4.34 (dd, 1H, *J* = 8.7, 5.4 Hz, α-CH Leu), 4.23 (d, 1H, *J* = 7.0 Hz, α-CH Val 2), 4.21 (d, 1H, *J* = 7.3 Hz, α-CH Val 1), 3.22-3.11 (m, 3H, β-CH₂ a Phe, δ-CH₂ Arg), 2.84, (dd, 1H, *J* = 14.1, 10.3 Hz, β-CH₂ b Phe), 2.18-1.99 (m, 4H, β-CH Val 1, β-CH Val 2, CH₂ iVal), 1.96-1.83 (m, 2H, β-CH₂ Arg), 1.76-1.54 (m, 6H, γ-CH₂ Arg, γ-CH Leu, β-CH₂ Leu, CH iVal), 1.00-0.90 (m, 18H, δ-CH₃ Leu, γ-CH₃ Val 2, γ-CH₃ Val 1), 0.82 (d, 3H, *J* = 6.6 Hz, CH₃ a iVal), 0.74 (d, 3H, *J* = 6.6 Hz, CH₃ b iVal).

¹³C-NMR (MeOD, 100 MHz): δ_C [ppm] = 177.5 (COOH Leu), 175.6 (CO iVal), 173.9 (CO Phe), 173.6 (CO Arg), 173.3 (CO Val 1), 173.1 (CO Val 2), 158.6 (ζ-C_{quart} Arg), 138.6 (γ-C_{quart} Phe), 130.2, 129.5, 127.7 (CH_{arom} Phe), 60.5 (α-CH Val 2), 59.9 (α-CH Val 1), 55.9 (α-CH Phe), 53.9 (α-CH Arg), 53.5 (α-CH Leu), 46.1 (CH₂ iVal), 42.7 (γ-CH₂ Arg), 42.1 (β-CH₂ Leu), 38.6 (β-CH₂ Phe), 32.3 (β-CH Val 1), 31.9 (β-CH Val 2), 30.5 (β-CH₂ Arg), 27.4 (CH iVal), 26.1 (γ-CH Leu), 26.0 (δ-CH₂ Arg), 23.5, 22.7, 22.6, 22.3, 19.8, 19.7 (CH₃ Val 1, Val 2, Leu), 18.7 (CH₃ a iVal), 18.6 (CH₃ b iVal).

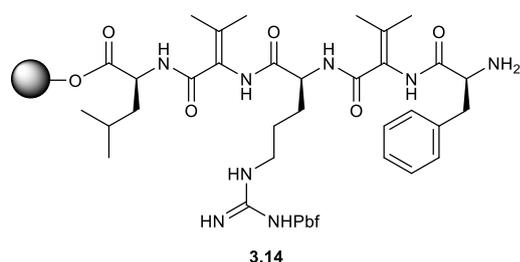
Additional found signals: δ_H [ppm] = 4.88 (H₂O), 2.66 (DMSO), broad peak between 4.99-4.42, NH signals, COOH Leu and NH/NH₂ Arg were not observed. δ_C [ppm] = 40.4 (DMSO).

The differentiation between Val 1 and Val 2 refers to the spin system and not the position in the molecule. Due to the missing NH signals differentiation was not possible using HMBC data.

UHR-MS (ESI-TOF) *m/z* calcd for C₃₆H₆₁N₈O₇: 717.4658 [M+H]⁺; found: 717.4655 [M+H]⁺

Specific rotation [α]_D²¹ = -95.0° (c = 1.00, CH₃OH)

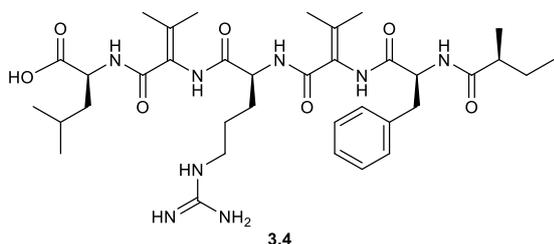
Preparation of 2CT-L-leucine-dehydro-valine-L-arginine(Pbf)-dehydro-valine-L-phenylalanine-NH₂ (**3.14**)



H-L-Ile-2-chlorotrityl resin ($n = 0.74$ mmol/g, 2.106 g, 1.558 mmol) was swelled in DMF for 30 min. The solvent was removed and a solution of Fmoc-2,3-dehydrovaline-OH (1.577 g, 4.674 mmol) and HATU (1.720 g, 4.524 mmol) in DMF was added.

DIPEA (1590 μ L, 9.350 mmol) and more DMF followed. The mixture was agitated for 1 h, succeeded by Fmoc-deprotection. Fmoc-L-arginine(Pbf)-OH (3.037 g, 4.681 mmol) and HATU (1.721 g, 4.526 mmol) were dissolved in DMF and added to the drained resin, followed by DIPEA (1590 μ L, 9.350 mmol) and more DMF. The mixture was agitated for 1 h. After Fmoc-deprotection, Fmoc-2,3-dehydrovaline-OH (1.577 g, 4.674 mmol) and HATU (1.718 g, 4.518 mmol) dissolved in DMF, were added, followed by DIPEA (1590 μ L, 9.350 mmol) and more DMF. The mixture was agitated for 3 h, after which the coupling was repeated due to the incompleteness of the reaction. Fmoc-2,3-dehydrovaline-OH (0.571 g, 1.69 mmol) and HATU (0.624 g, 1.641 mmol) were added as a solution in DMF, followed by DIPEA (575 μ L, 3.38 mmol) and more DMF. The mixture was agitated for 1 h, after which full conversion was observed as indicated by LC-MS. After Fmoc-deprotection, Fmoc-L-phenylalanine-OH (1.816 g, 4.688 mmol) and HATU (1.720 g, 4.524 mmol) dissolved in DMF, were added, followed by DIPEA (1590 μ L, 9.350 mmol) and more DMF. The mixture was agitated for 2 h, after which the coupling was repeated. Fmoc-L-phenylalanine-OH (1.811 g, 4.675 mmol) and HATU (1.719 g, 4.521 mmol) dissolved in a bit DMF, were added to the resin, followed by DIPEA (1590 μ L, 9.350 mmol) and more DMF. The mixture was agitated for 2 h. After Fmoc-deprotection the supernatant was drained, the resin was washed three times with DMF, two times each with isopropanol and *n*-heptane. It was sucked dry and half of resin **3.14** was taken for storage. This was dried *in vacuo* for 18 h, flushed with argon and stored at -8 °C. The rest of resin **3.14** was directly used in the next step.

Preparation of (2-((S)-5-guanidino-2-(3-methyl-2-((S)-2-((S)-2-methylbutanamido)-3-phenylpropanamido)but-2-enamido)pentanamido)-3-methylbut-2-enoyl)-L-leucine (3.4)



Resin **3.14** (0.779 mmol) was swelled in DMF for 30 min. The solvent was drained, HATU (0.861 g, 2.26 mmol) dissolved in DMF was added, followed by *S*-methylbutanoic acid (255 μ L, 2.34 mmol), DIPEA (795 μ L, 4.68 mmol) and more DMF. The mixture was agitated for 1 h. The solvent was removed, and the resin was washed three times each with DMF, isopropanol and *n*-heptane. The cleavage cocktail, consisting of TFA/TIS/H₂O (95:2.5:2.5), was added, coloring the mixture a dark violet. The resin was agitated for 30 min after which the supernatant was drained and the process was repeated once more. The combined filtrates were reduced under pressure and then dried further using lyophilization. Purification of 58% of the crude product using semi preparative HPLC (5-50-95% AcCN + 0.1% FA, NUCLEODUR® C18 Gravity SB, 3 μ m, 250 x 10 mm, flow rate: 3 mL/min) yielded **3.4** as a colorless powder (54.5 mg, 0.0764 mmol, overall yield calculated to be: 17%).

¹H-NMR (MeOD, 600 MHz): δ_{H} [ppm] = 7.28 (d, 4H, J = 4.6 Hz, CH_{arom} Phe), 7.24-7.20 (m, 1H, CH_{arom} Phe), 4.59 (t, 1H, J = 7.9 Hz, α -CH Phe), 4.45 (dd, 1H, J = 9.4, 5.3 Hz, α -CH Leu), 4.34 (dd, 1H, J = 8.5, 5.7 Hz, α -CH Arg), 3.21 (t, 2H, J = 7.2 Hz, δ -CH₂ Arg), 3.10 (dd, 1H, J = 13.7, 7.7 Hz, β -CH₂ a Phe), 3.02 (dd, 1H, J = 13.7, 7.9 Hz, β -CH₂ b Phe), 2.30-2.23 (m, 1H, CH *S*-MBA), 2.06-2.00 (m, 1H, β -CH₂ a Arg), 2.07 (s, 3H, γ -CH₃ a Dhv 1), 2.02 (s, 3H, γ -CH₃ a Dhv 2), 1.90-1.82 (m, 1H, β -CH₂ b Arg), 1.79 (s, 3H, γ -CH₃ b Dhv 1), 1.77-1.69 (m, 3H, γ -CH₂ Arg, γ -CH Leu), 1.69-1.60 (m, 2H, β -CH₂ Leu), 1.60-1.52 (m, 1H, CH₂ a *S*-MBA), 1.47 (s, 3H, γ -CH₃ b Dhv 2), 1.40-1.31 (m, 1H, CH₂ b *S*-MBA), 1.00 (d, 3H, J = 6.8 Hz, CH-CH₃ *S*-MBA), 0.92 (d, 3H, J = 6.4 Hz, δ -CH₃ a Leu), 0.91 (d, 3H, J = 6.4 Hz, δ -CH₃ b Leu), 0.86 (t, 3H, J = 7.4 Hz, CH₂-CH₃ *S*-MBA).

¹³C-NMR (MeOD, 150 MHz): δ_{C} [ppm] = 179.9 (CO *S*-MBA), 176.9 (COOH Leu), 173.44 (CO Arg), 173.38 (CO Phe), 168.0 (CO Dhv 1), 167.8 (CO Dhv 2), 158.7 (ζ -C_{quart} Arg), 141.5 (α -C_{quart} Dhv 1), 141.4 (α -C_{quart} Dhv 1), 138.0 (γ -C_{quart} Phe), 124.9 (β -C_{quart} Dhv 1), 124.6 (β -C_{quart} Dhv 2), 130.4, 129.6, 128.0 (CH_{arom} Phe), 56.7 (α -CH Phe), 54.9 (α -CH Arg), 52.9 (α -CH Leu), 43.1 (CH *S*-MBA), 42.11 (β -CH₂ Leu), 42.06 (δ -CH₂ Arg), 38.4 (β -CH₂ Phe), 29.4 (β -CH₂ Arg), 28.2 (CH₂ *S*-MBA), 26.4 (γ -CH₂ Arg), 25.9 (γ -CH Leu), 23.4 (δ -CH₃ b Leu), 22.2 (δ -CH₃ a Leu), 21.61 (γ -CH₃ b Dhv 1),

21.57 (γ -CH₃ b Dhv 2), 21.1 (γ -CH₃ a Dhv 1), 20.9 (γ -CH₃ a Dhv 2), 17.9 (CH-CH₃ S-MBA), 12.3 (CH₂-CH₃ S-MBA).

Additional found signals: δ_{H} [ppm] = 8.16 (FA), 4.84 (H₂O), 2.66 (DMSO), NH signals, COOH Leu and NH/NH₂ Arg were not observed. δ_{C} [ppm] = 40.4 (DMSO).

Dhv 1 refers to the amino acid on the right side of L-leucine, and Dhv 2 to the amino acid on the right side of L-arginine. Differentiation is based on HMBC data.

UHR-MS (ESI-TOF) m/z calcd for for C₃₆H₅₇N₈O₇: 713.4345 [M+H]⁺; found: 713.4353 [M+H]⁺

Specific rotation $[\alpha]_{\text{D}}^{21} = +63.0^{\circ}$ (c = 1.00, CH₃OH)

6.5 Supporting Information

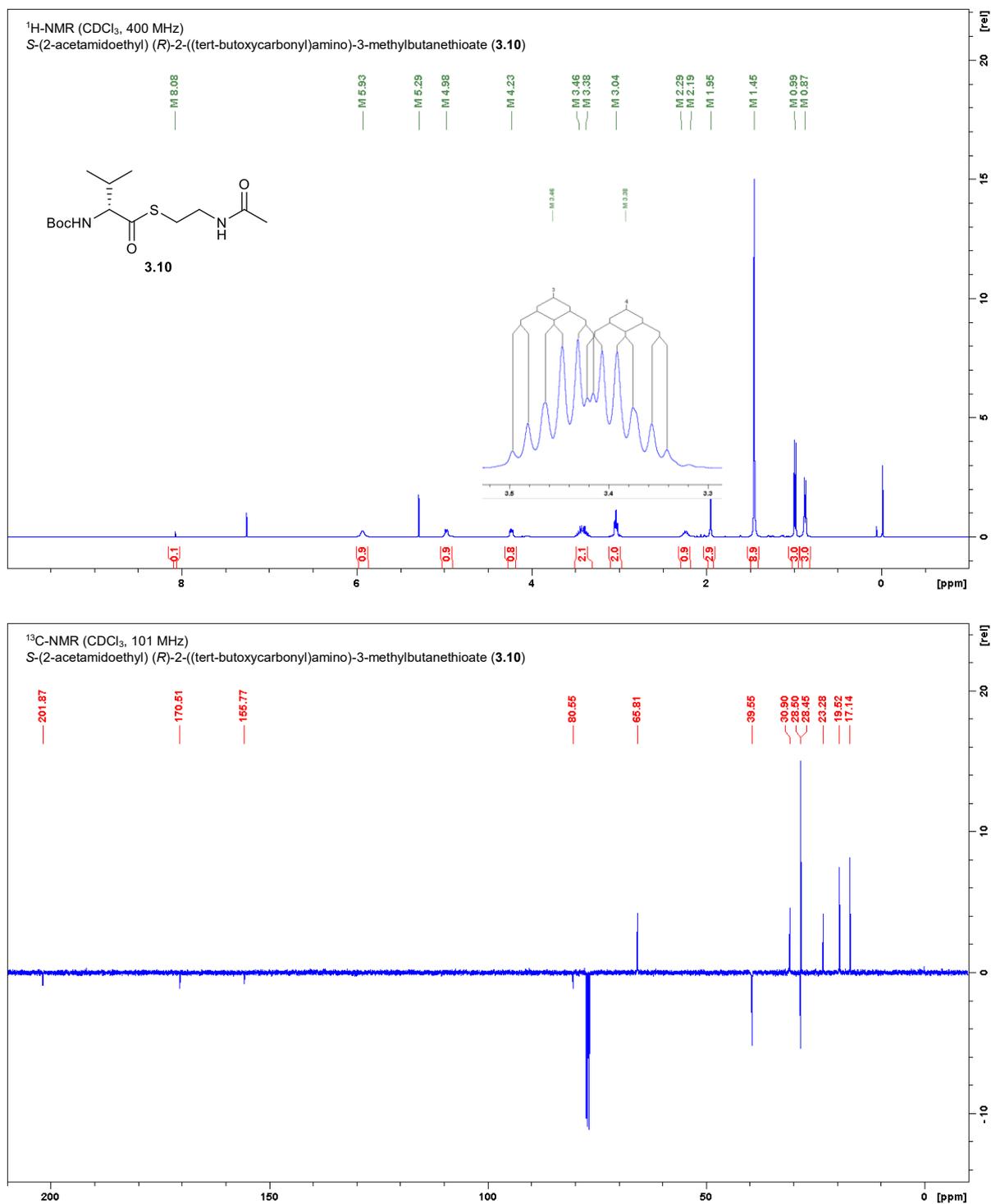


Figure S 6-1: NMR spectra of **3.10** in CDCl₃.

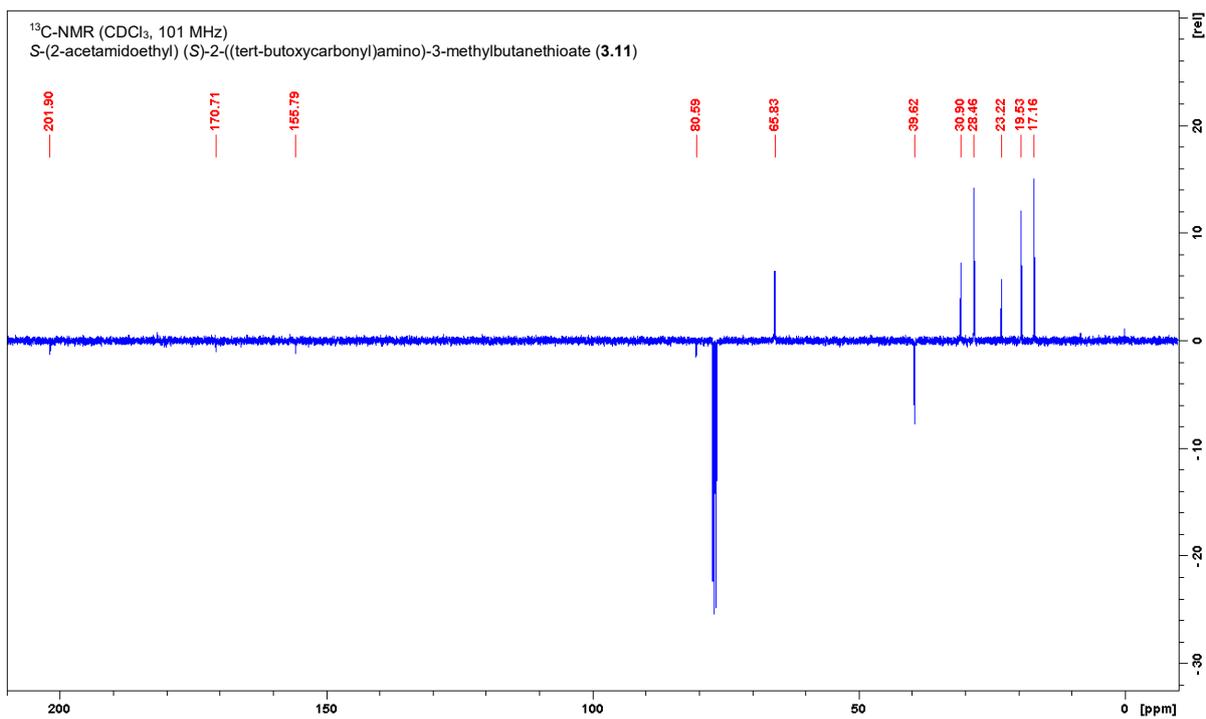
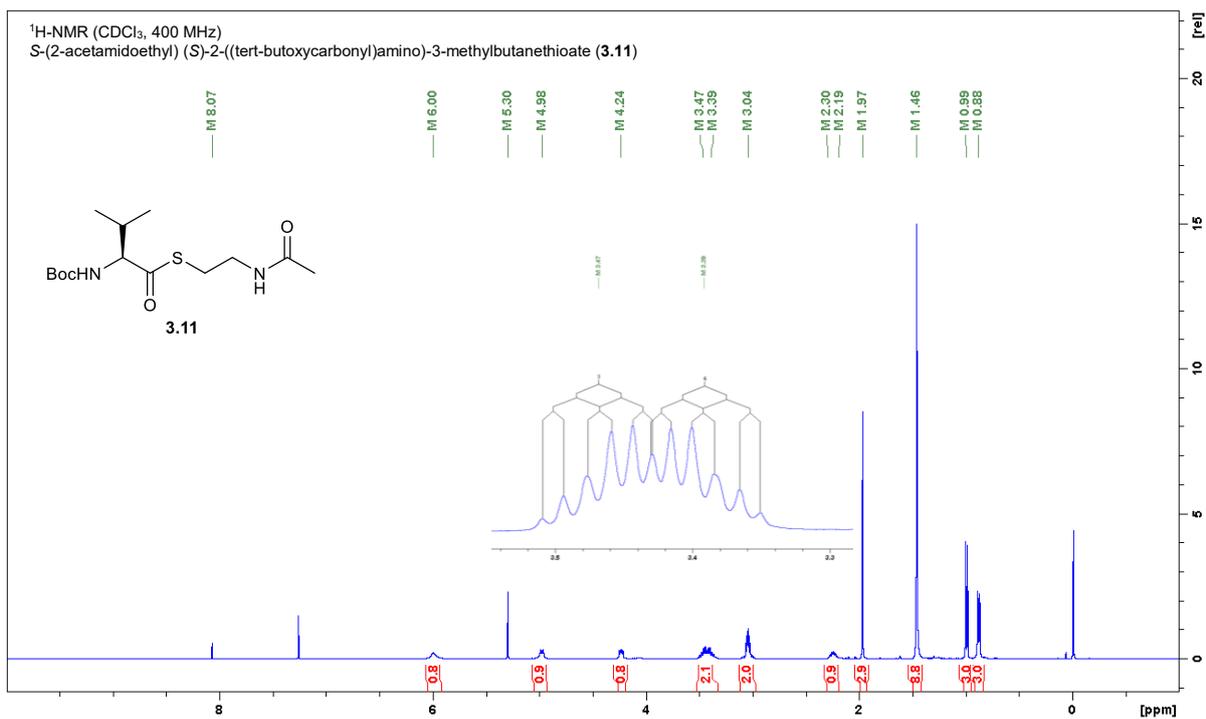


Figure S 6-2: NMR spectra of **3.11** in CDCl₃.

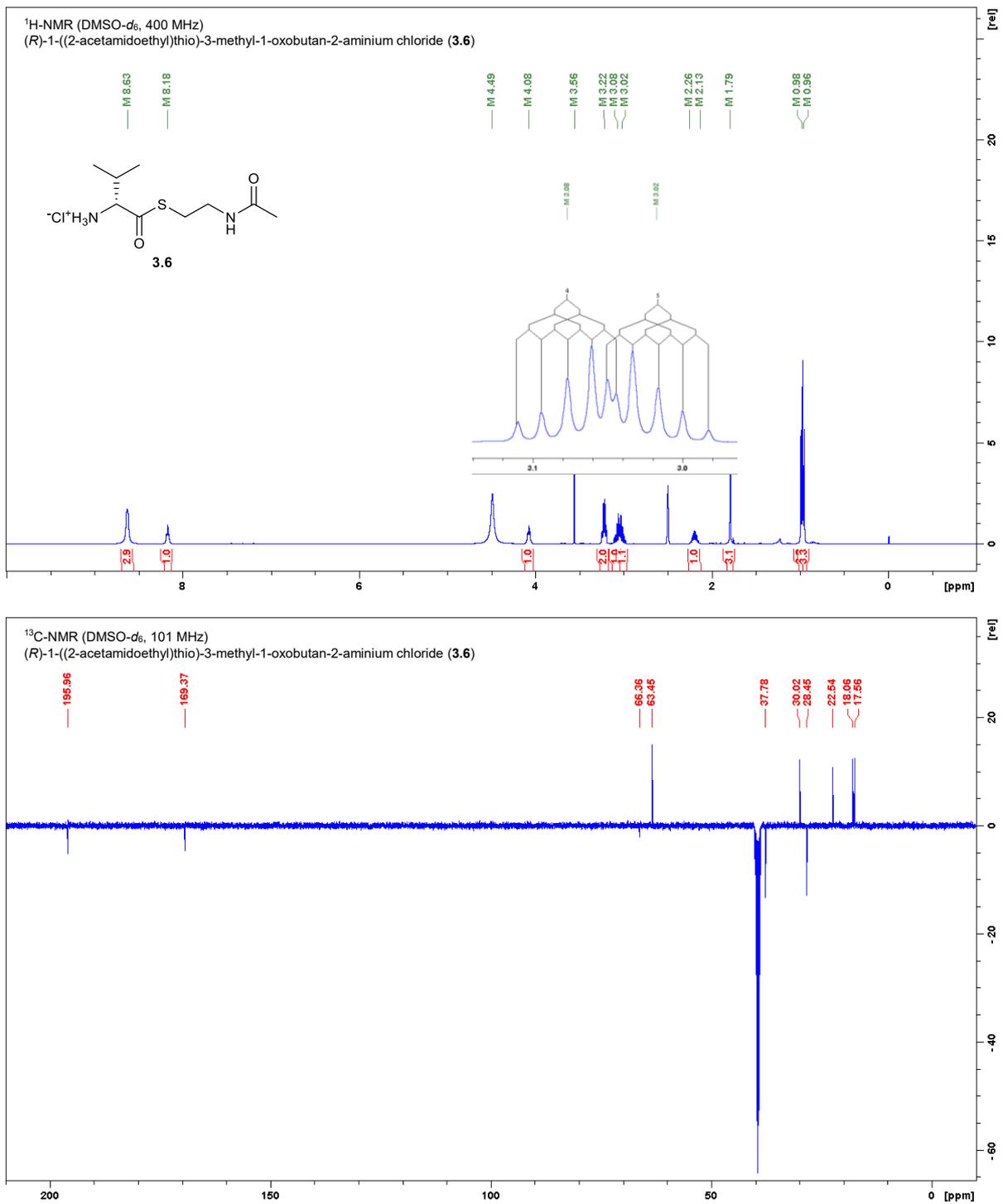


Figure S 6-3: NMR spectra of **3.6** in DMSO-*d*₆.

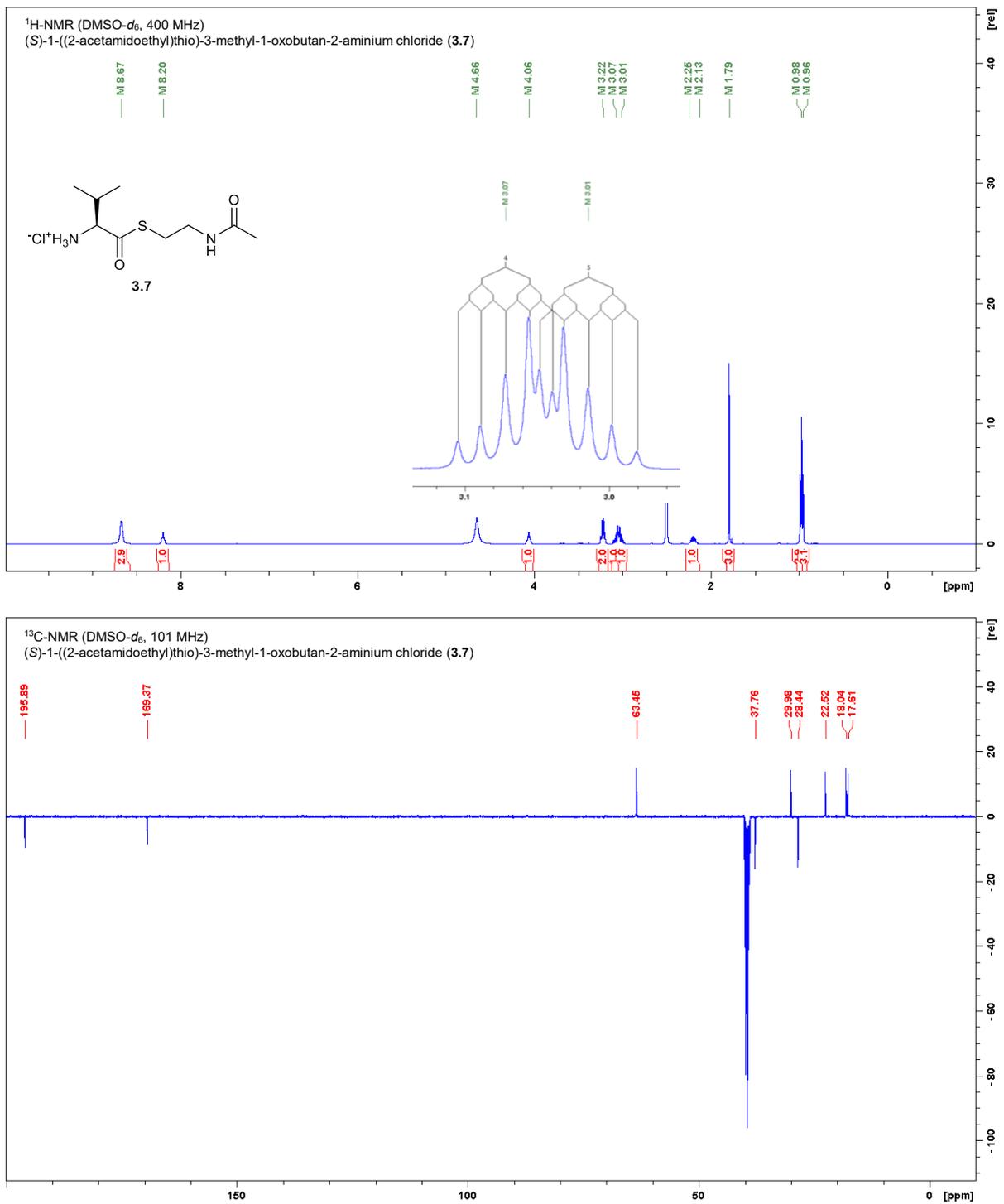


Figure S 6-4: NMR spectra of 3.7 in DMSO-*d*₆.

Table S 6-1: NMR data for 3.9 in MeOD (600 MHz, 101 MHz).

Position	$\delta^{13}\text{C}$ [ppm]	$\delta^1\text{H}$ [ppm]; (m, f, J)
L-Leucine		
COOH	179.5	n.o.
NH	-	n.o.
α -CH	54.4	4.36 (dd, 1H, $J = 10.1, 4.5$ Hz)
β -CH ₂	42.7	1.65-1.57 (m, 3H, overlay)
γ -CH	26.3	1.73-1.65 (m, 2H, overlay)
δ -CH ₃ a	23.7	0.95 (d, 3H, $J = 6.1$ Hz)
δ -CH ₃ b	22.0	0.94 (d, 3H, $J = 6.1$ Hz)
D-Valine 1		
CO	174.4	-
NH	-	n.o.
α -CH	61.0	3.93 (d, 1H, $J = 6.8$ Hz)
β -CH	30.5	2.13-2.00 (m, 4H, overlay)
γ -CH ₃ a	18.5	0.74 (d, 3H, $J = 6.8$ Hz)
γ -CH ₃ b	19.6	0.70 (d, 3H, $J = 6.8$ Hz)
L-Arginine		
CO	173.5	-
NH	-	n.o.
α -CH	55.4	4.29 (t, 1H, $J = 6.6$ Hz)
β -CH ₂	30.0	1.90-1.76 (m, 2H)
γ -CH ₂	26.0	1.73-1.65 (m, 2H, overlay) 1.65-1.57 (m, 3H, overlay)
δ -CH ₂	42.2	3.17-3.11 (m, 2H)
ϵ -NH	-	n.o.
ζ -C _{quart}	158.7	-
η -NH	-	n.o.
η -NH ₂	-	n.o.
D-Valine 2		
CO	172.6	-
NH	-	n.o.
α -CH	60.5	4.24 (d, 1H, $J = 8.4$ Hz)
β -CH	31.5	2.13-2.00 (m, 4H, overlay)
γ -CH ₃ a	20.0	0.92 (d, 3H, $J = 6.6$ Hz)
γ -CH ₃ b	18.8	0.84 (d, 3H, $J = 6.8$ Hz)
L-Phenylalanine		
CO	174.2	-
NH	-	n.o.
α -CH	56.5	4.67 (t, 1H, $J = 7.7$ Hz)
β -CH ₂	39.3	3.03 (dd, 1H, $J = 13.3, 8.1$ Hz) 2.98 (dd, 1H, $J = 13.6, 7.3$ Hz)
γ -C _{quart}	138.0	-
CH _{arom}	130.4, 129.6, 127.9	7.30-7.18 (m, 5H)
Isovaleroyl		

CO	175.3	-
CH₂	46.0	2.13-2.00 (m, 4H, overlay)
CH	27.4	2.00-1.93 (m, 1H)
CH₃ a	22.9	0.87 (d, 3H, <i>J</i> = 6.6 Hz)
CH₃ b	22.7	0.84 (d, 3H, <i>J</i> = 6.5 Hz)

The differentiation between Val 1 and Val 2 refers to the spin system and not the position in the molecule. Due to the missing NH signals differentiation was not possible using HMBC data.

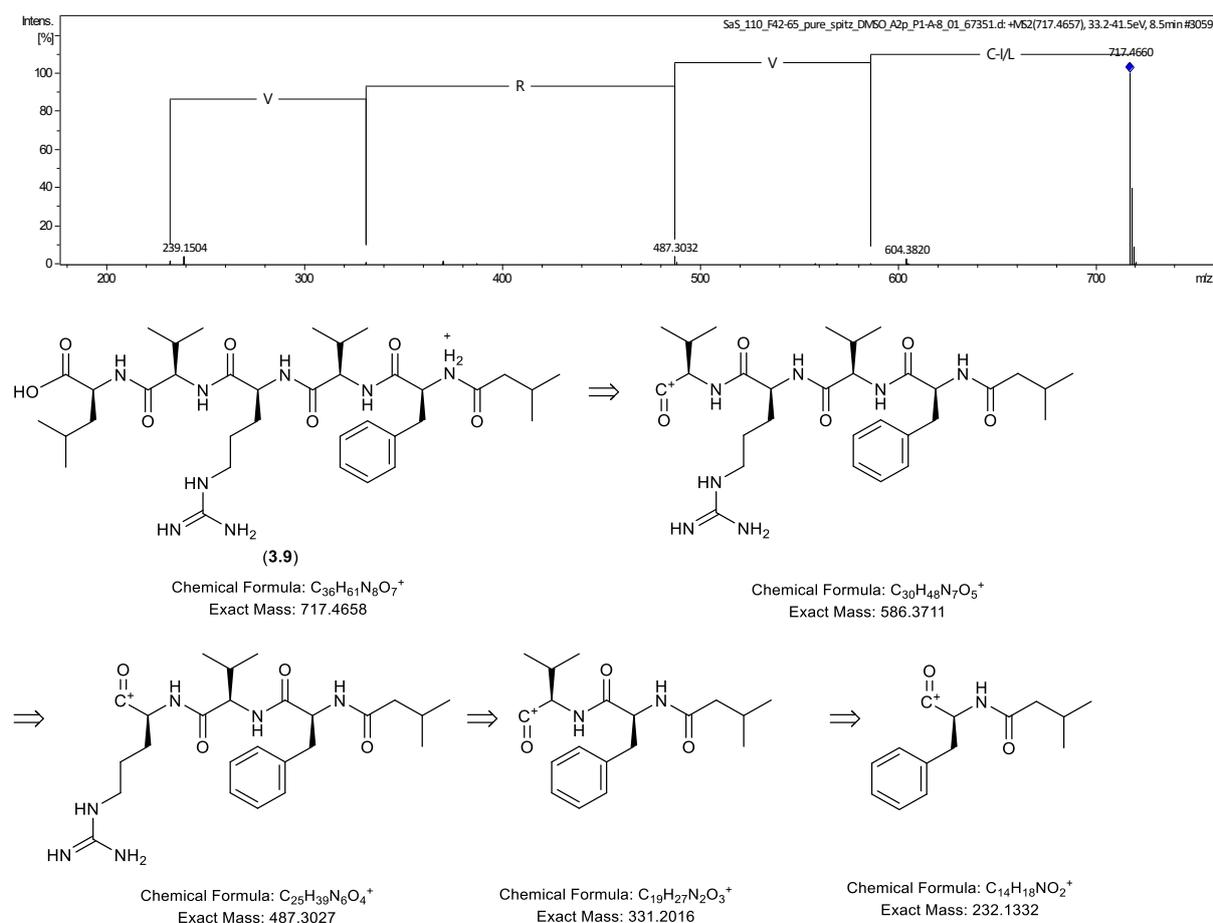


Figure S 6-6: MS² fragmentation of 3.9 and fragmentation pattern.

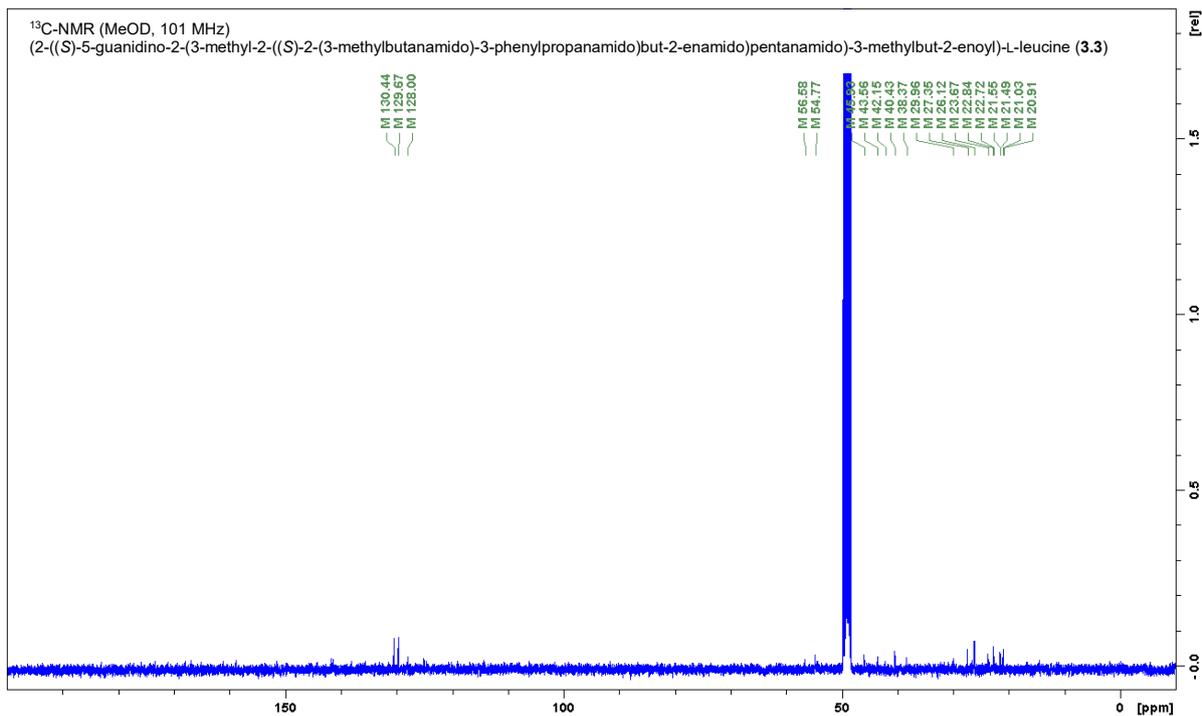
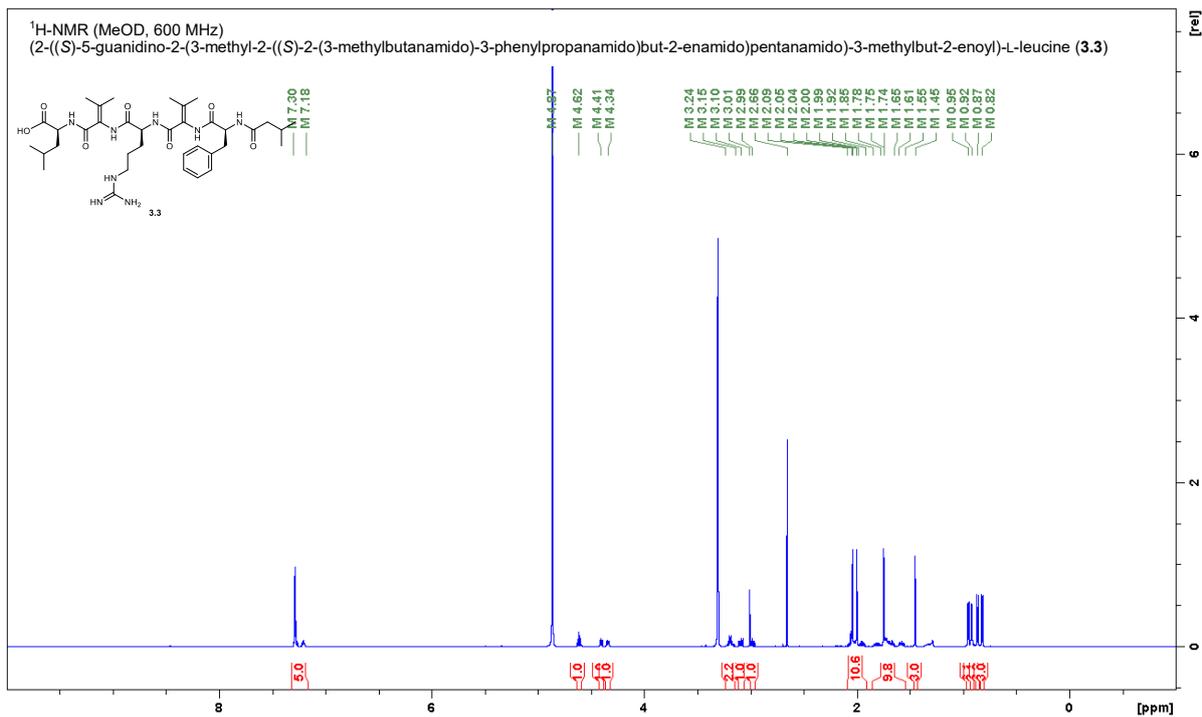


Figure S 6-7: NMR spectra of 3.3 in MeOD.

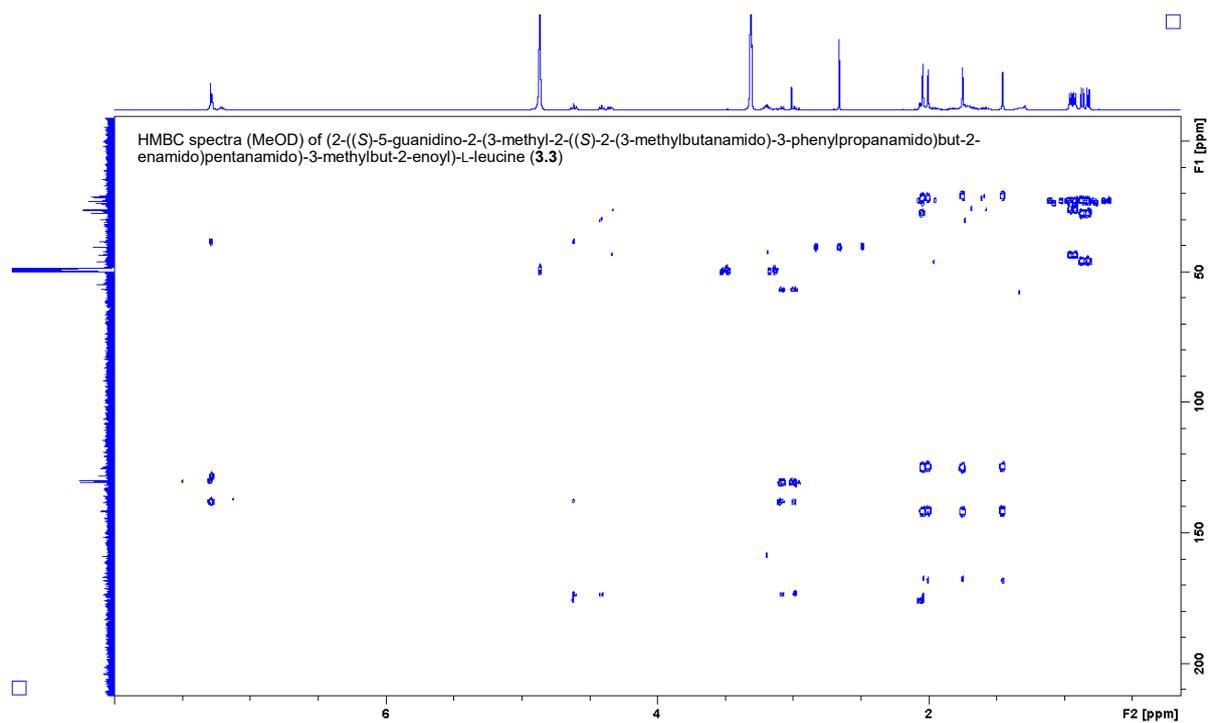


Figure S 6-8: HMBC NMR spectra of **3.3** in MeOD.

Table S 6-2: NMR data for 3.3 in MeOD (600 MHz, 101 MHz).

Position	$\delta^{13}\text{C}$ [ppm]	$\delta^1\text{H}$ [ppm]; (m, f, J)
L-Leucine		
COOH	n.o.	n.o.
NH	-	n.o.
α -CH	54.8	4.34 (dd, 1H, J = 9.0, 4.9 Hz)
β -CH ₂	43.6	1.85-1.52 (m, 6H, overlay)
γ -CH	26.1	1.85-1.52 (m, 6H, overlay)
δ -CH ₃ a	22.7	0.95 (d, 3H, J = 6.3 Hz)
δ -CH ₃ b	23.7	0.92 (d, 3H, J = 6.5 Hz)
Dehydrovaline 1		
CO	167.5*	-
NH	-	n.o.
α -C _{quart}	141.9*	-
β -C _{quart}	125.2*	-
γ -CH ₃ a	21.0	2.04 (s, 3H)
γ -CH ₃ b	21.6	1.75 (s, 3H)
L-Arginine		
CO	173.8*	-
NH	-	n.o.
α -CH	54.8	4.41 (dd, 1H, J = 7.4, 6.1 Hz)
β -CH ₂	30.0	2.09-1.92 (m, 4H, overlay) 1.85-1.52 (m, 6H, overlay)
γ -CH ₂	26.1	1.85-1.52 (m, 6H, overlay)
δ -CH ₂	42.1	3.24-3.15 (m, 2H)
ϵ -NH	-	n.o.
ζ -C _{quart}	158.8*	-
η -NH	-	n.o.
η -NH ₂	-	n.o.
Dehydrovaline 2		
CO	168.4*	-
NH	-	n.o.
α -C _{quart}	141.5*	-
β -C _{quart}	124.8*	-
γ -CH ₃ a	20.9	2.00 (s, 3H)
γ -CH ₃ b	21.5	1.45 (s, 3H)
L-Phenylalanine		
CO	173.4*	-
NH	-	n.o.
α -CH	56.6	4.62 (t, 1H, J = 7.8 Hz)
β -CH ₂	38.4	3.10 (dd, 1H, J = 13.8, 7.7 Hz) 2.99 (dd, 1H, J = 12.7, 9.2 Hz)
γ -C _{quart}	138.2*	-
CH _{arom}	130.4, 129.7, 128.0	7.30-7.18 (m, 5H)
Isovaleroyl		

CO	176.1*	-
CH₂	45.9	2.09-1.92 (m, 4H, overlay)
CH	27.4	2.09-1.92 (m, 4H, overlay)
CH₃ a	22.8	0.87 (d, 3H, <i>J</i> = 6.6 Hz)
CH₃ b	22.7	0.82 (d, 3H, <i>J</i> = 6.6 Hz)

The differentiation between Dehydrovaline 1 and 2 refers to the spin system and not the position in the molecule. Due to the missing NH signals differentiation was not possible using HMBC data.

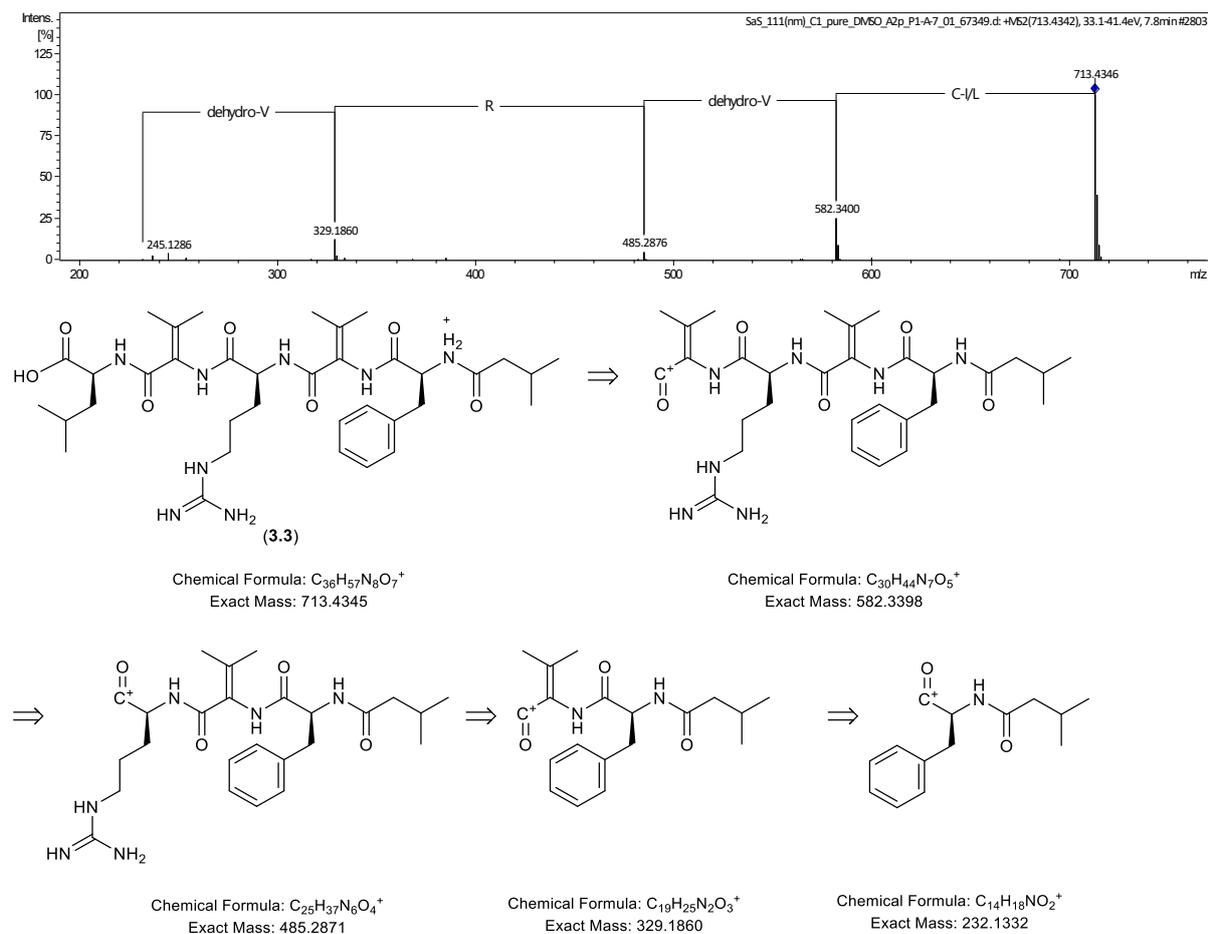


Figure S 6-9: MS² fragmentation of 3.3 and fragmentation pattern.

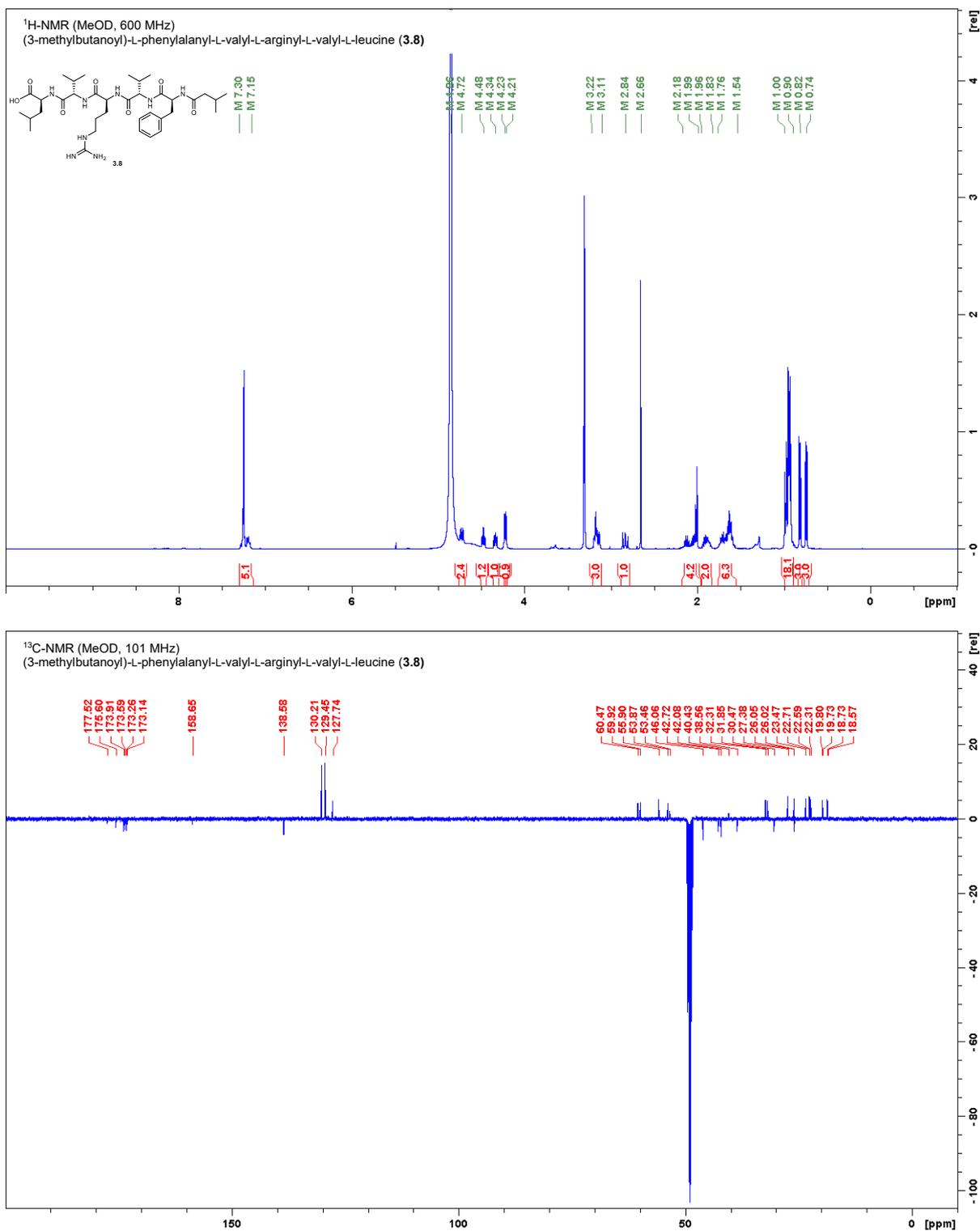


Figure S 6-10: NMR spectra of 3.8 in MeOD.

Table S 6-3: NMR data for 3.8 in MeOD (600 MHz, 101 MHz).

Position	$\delta^{13}\text{C}$ [ppm]	$\delta^1\text{H}$ [ppm]; (m, f, J)
L-Leucine		
COOH	177.5	n.o.
NH	-	n.o.
α -CH	53.5	4.34 (dd, 1H, $J = 8.7, 5.4$ Hz, α -CH Leu),
β -CH ₂	42.7	1.76-1.54 (m, 6H, overlay)
γ -CH	26.1	1.76-1.54 (m, 6H, overlay)
δ -CH ₃ a, b	23.5, 22.7, 22.6, 22.3, 19.8, 19.7 ⁴²	1.00-0.90 (m, 18H, overlay)
L-Valine 1		
CO	173.3	-
NH	-	n.o.
α -CH	59.9	4.21 (d, 1H, $J = 7.3$ Hz)
β -CH	32.3	2.18-1.99 (m, 4H, overlay)
γ -CH ₃ a, b	23.5, 22.7, 22.6, 22.3, 19.8, 19.7Error! Bookmark not defined.	1.00-0.90 (m, 18H, overlay)
L-Arginine		
CO	173.6	-
NH	-	n.o.
α -CH	53.9	4.48 (t, 1H, $J = 6.8$ Hz)
β -CH ₂	30.5	1.96-1.83 (m, 2H)
γ -CH ₂	26.0	1.76-1.54 (m, 6H, overlay)
δ -CH ₂	42.1	3.22-3.11 (m, 3H, overlay)
ϵ -NH	-	n.o.
ζ -C _{quart}	158.6	-
η -NH	-	n.o.
η -NH ₂	-	n.o.
L-Valine 2		
CO	173.1	-
NH	-	n.o.
α -CH	60.5	4.23 (d, 1H, $J = 7.0$ Hz)
β -CH	31.9	2.18-1.99 (m, 4H, overlay)
γ -CH ₃ a, b	23.5, 22.7, 22.6, 22.3, 19.8, 19.7Error! Bookmark not defined.	1.00-0.90 (m, 18H, overlay)
L-Phenylalanine		
CO	173.9	-

⁴² Not distinguishable to CH₃ of Valine 1 and Valine 2

NH	-	n.o.
α -CH	55.9	4.72 (dd, 2H, $J = 8.9, 6.2$ Hz)
β -CH ₂	38.6	3.22-3.11 (m, 3H, overlay) 2.84, (dd, 1H, $J = 14.1, 10.3$ Hz)
γ -C _{quart}	138.6	-
CH _{arom}	130.2, 129.5, 127.7	7.30-7.15 (m, 5H)
<hr/>		
Isovaleroyl		
CO	175.6	-
CH ₂	46.1	2.18-1.99 (m, 4H, overlay)
CH	27.4	1.76-1.54 (m, 6H, overlay)
CH ₃ a	18.7	0.82 (d, 3H, $J = 6.6$ Hz)
CH ₃ b	18.6	0.74 (d, 3H, $J = 6.6$ Hz)

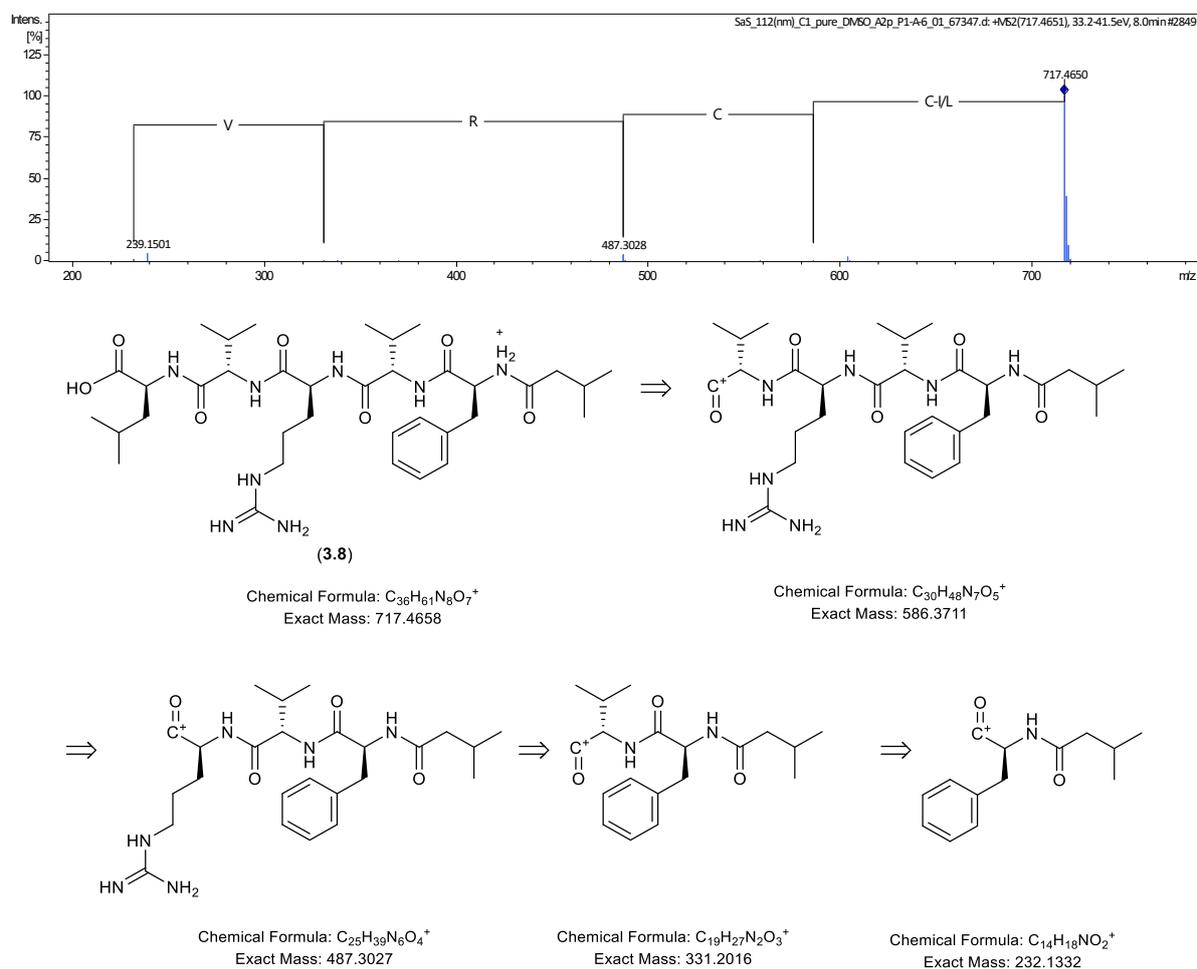


Figure S 6-11: MS² fragmentation of 3.8 and fragmentation pattern.

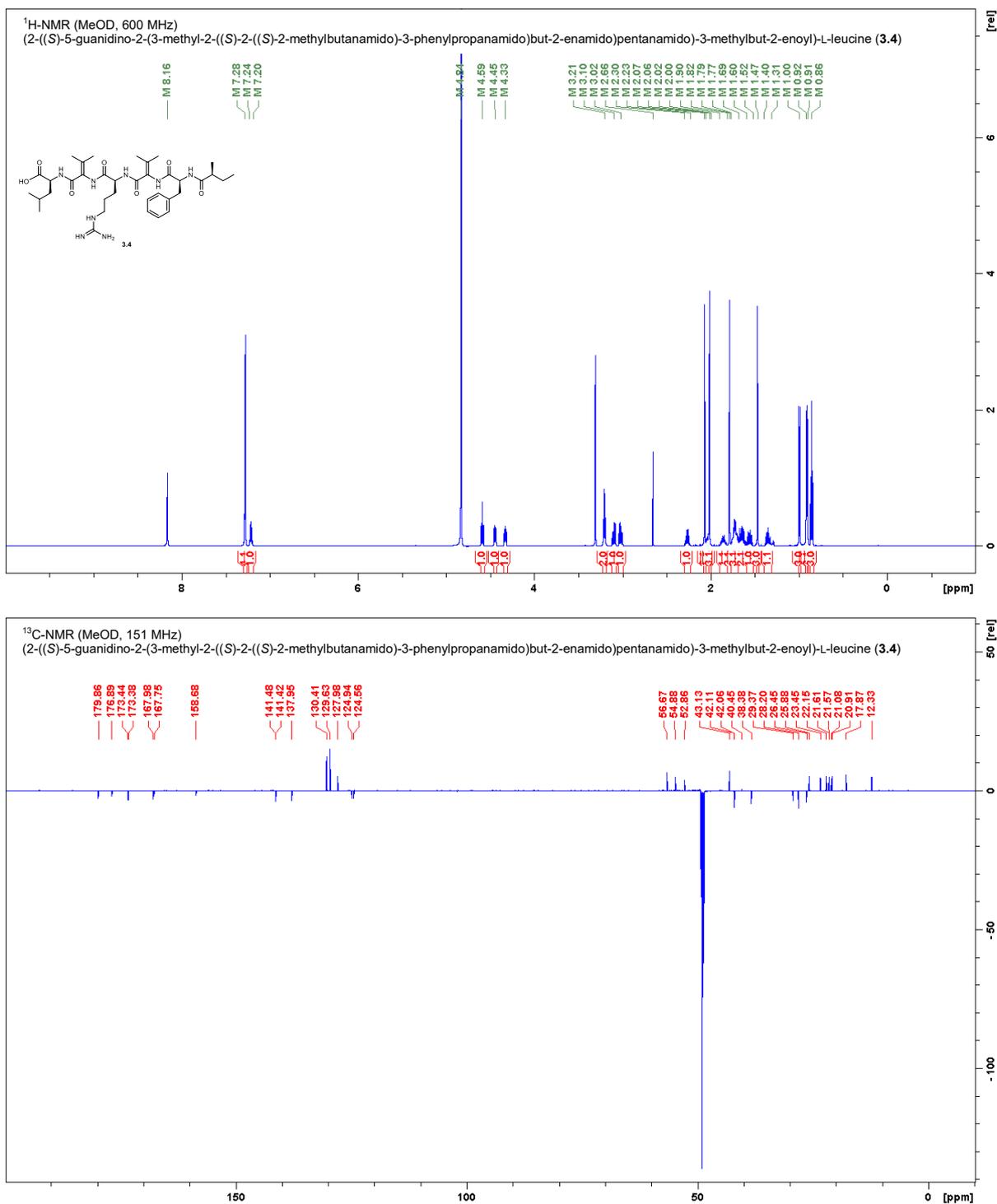


Figure S 6-12: NMR spectra of 3.4 in MeOD.

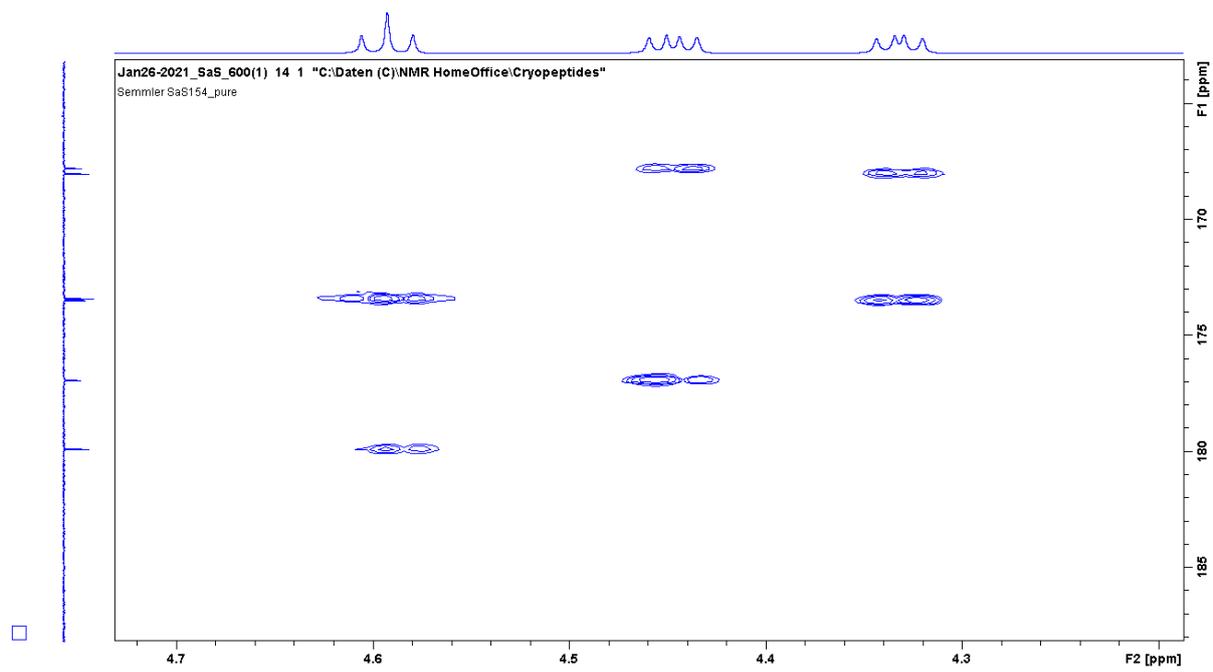
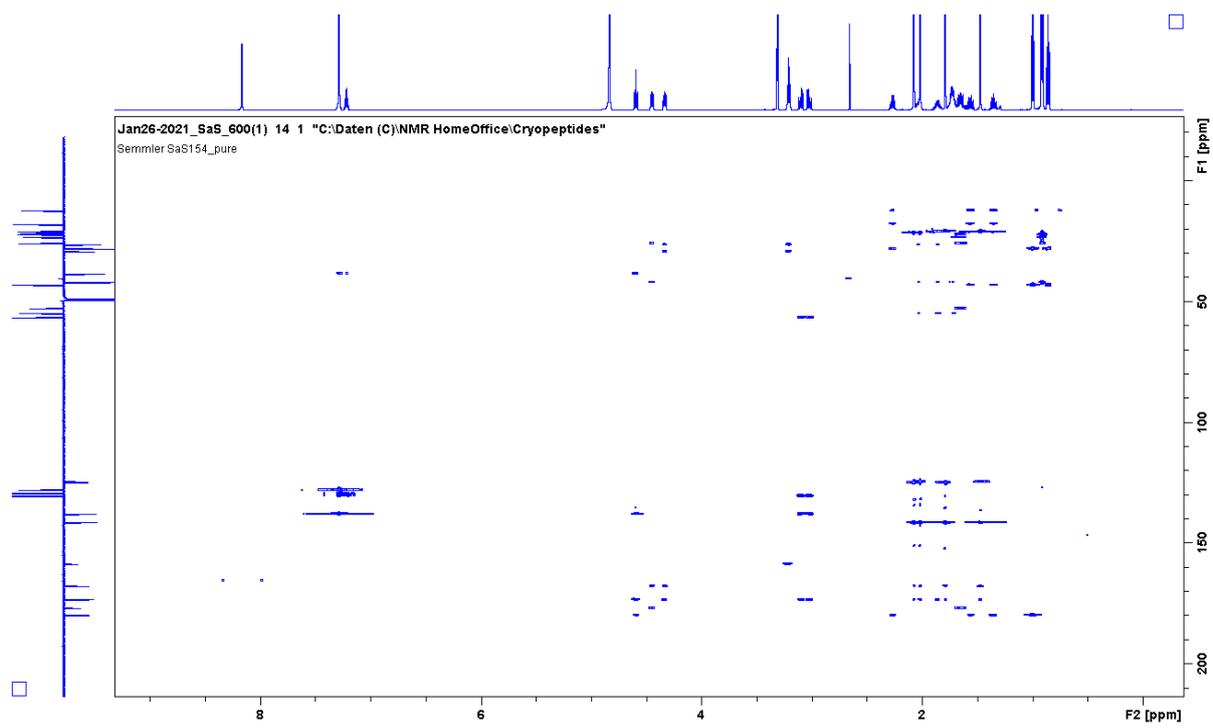


Figure S 6-13: 2D-NMR spectra of 3.4 in MeOD.

Table S 6-4: NMR data for 3.4 in MeOD (600 MHz, 151 MHz).

Position	$\delta^{13}\text{C}$ [ppm]	$\delta^1\text{H}$ [ppm]; (m, f, J)
L-Leucine		
COOH	176.9	n.o.
NH	-	n.o.
α -CH	52.9	4.45 (dd, 1H, $J = 9.4, 5.3$ Hz)
β -CH ₂	42.11	1.69-1.60 (m, 2H)
γ -CH	25.9	1.77-1.69 (m, 3H, overlay)
δ -CH ₃ a	22.2	0.92 (d, 3H, $J = 6.4$ Hz)
δ -CH ₃ b	23.4	0.91 (d, 3H, $J = 6.4$ Hz)
Dehydrovaline 1		
CO	168.0	-
NH	-	n.o.
α -C _{quart}	141.5	-
β -C _{quart}	124.9	-
γ -CH ₃ a	21.1	2.07 (s, 3H)
γ -CH ₃ b	21.61	1.79 (s, 3H)
L-Arginine		
CO	173.44	-
NH	-	n.o.
α -CH	54.9	4.34 (dd, 1H, $J = 8.5, 5.7$ Hz)
β -CH ₂	29.4	2.06-2.00 (m, 1H) 1.90-1.82 (m, 1H)
γ -CH ₂	26.4	1.77-1.69 (m, 3H, overlay)
δ -CH ₂	42.06	3.21 (t, 2H, $J = 7.2$ Hz)
ϵ -NH	-	n.o.
ζ -C _{quart}	158.7	-
η -NH	-	n.o.
η -NH ₂	-	n.o.
Dehydrovaline 2		
CO	167.8	-
NH	-	n.o.
α -C _{quart}	141.4	-
β -C _{quart}	124.6	-
γ -CH ₃ a	20.9	2.02 (s, 3H)
γ -CH ₃ b	21.57	1.47 (s, 3H)
L-Phenylalanine		
CO	173.38	-
NH	-	n.o.
α -CH	56.7	4.59 (t, 1H, $J = 7.9$ Hz)
β -CH ₂	38.4	3.10 (dd, 1H, $J = 13.7, 7.7$ Hz) 3.02 (dd, 1H, $J = 13.7, 7.9$ Hz)
γ -C _{quart}	138.0	-
CH _{arom}	130.4, 129.6, 128.0	7.28 (d, 4H, $J = 4.6$ Hz) 7.24-7.20 (m, 1H)
S-methylbutanoic acid		

CO	179.9	-
CH	43.1	2.30-2.23 (m, 1H)
CH ₃	17.9	1.00 (d, 3H, <i>J</i> = 6.8 Hz)
CH ₂	28.2	1.60-1.52 (m, 1H)
CH ₃	12.3	0.86 (t, 3H, <i>J</i> = 7.4 Hz)

Dhv 1 refers to the amino acid on the right side of L-leucine, and Dhv 2 to the amino acid on the right side of L-arginine. Differentiation is based on HMBC data.

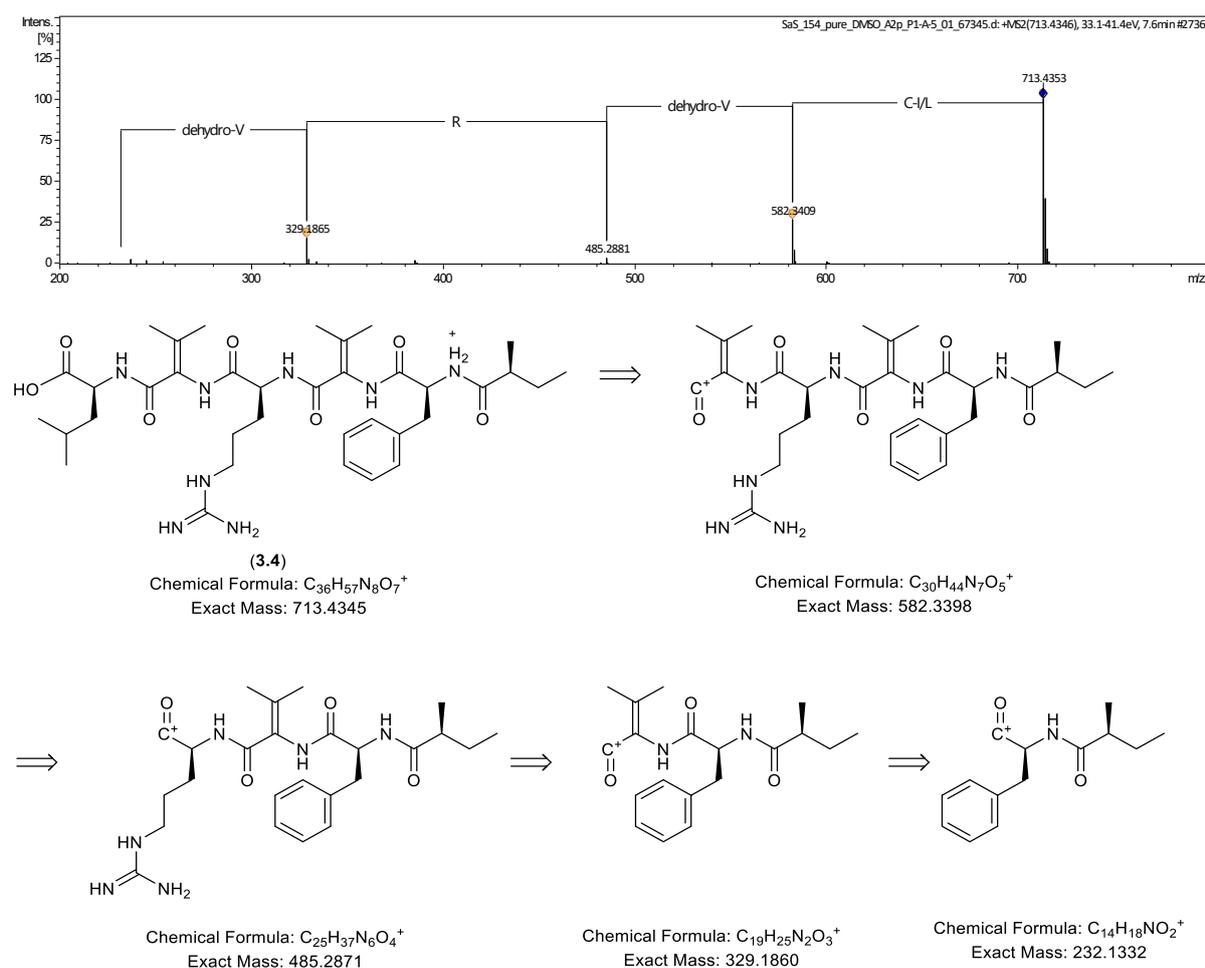


Figure S 6-14: MS² fragmentation of 3.4 and fragmentation pattern.

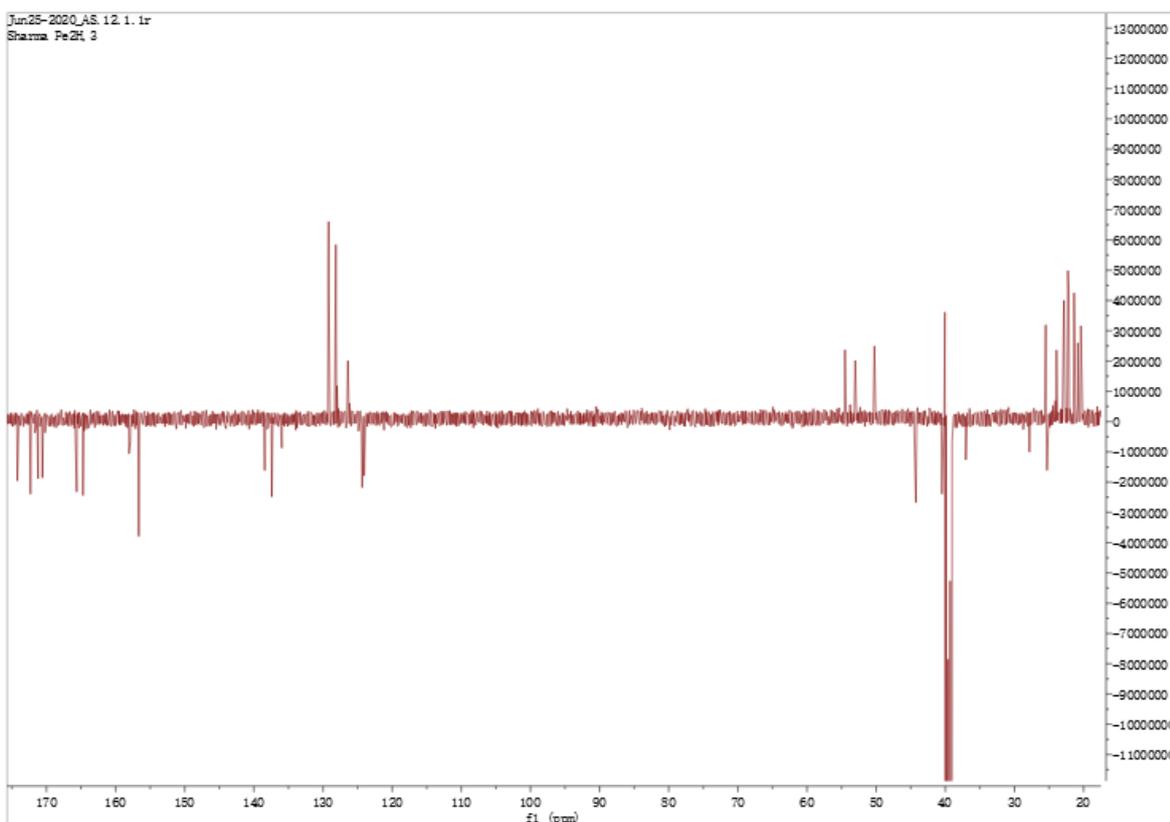
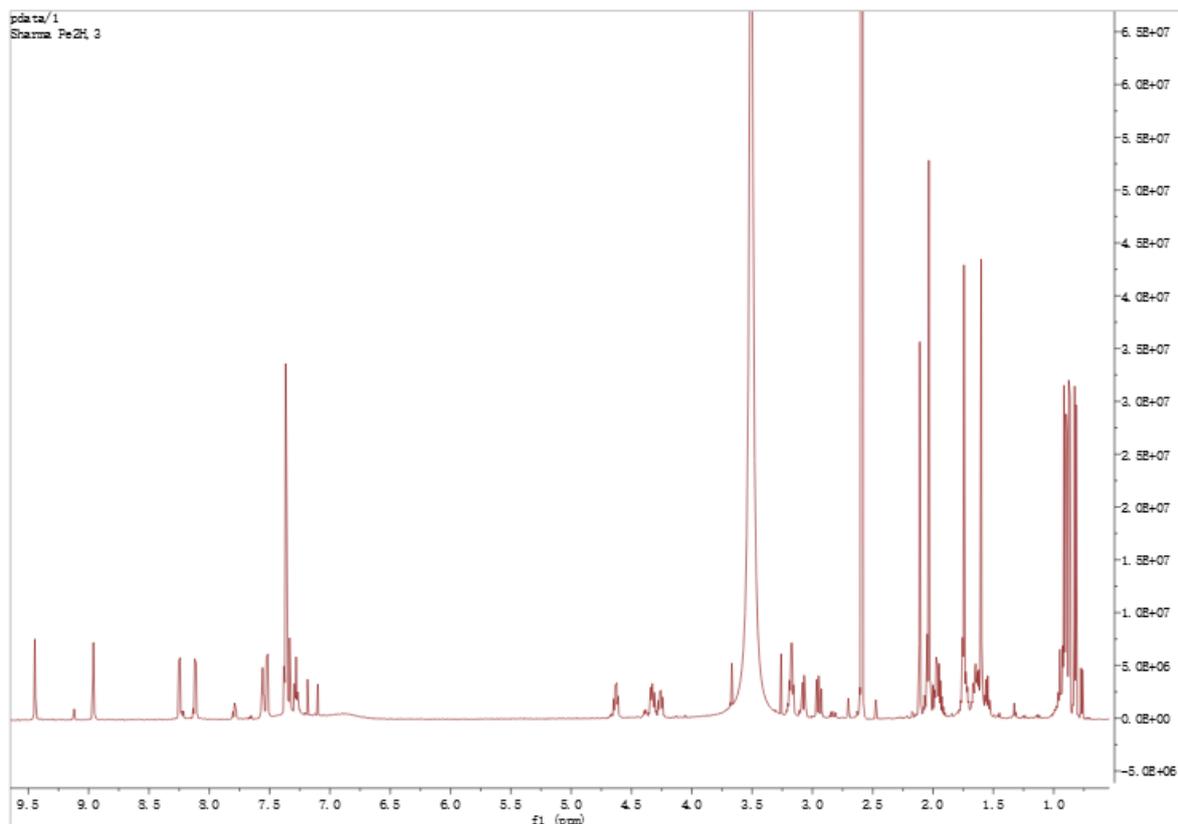


Figure S 6-15: Original NMR spectra of natural isolated fraction PE2H_3 in MeOD (600 MHz, 150 MHz).

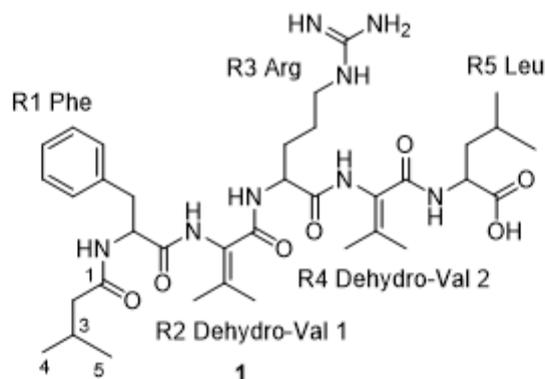


Table 1. ^1H (600 MHz) and ^{13}C (150 MHz) NMR data of **1** (CD_3OD ; δ in ppm)

Unit	Position	δ_{C}	δ_{H} (J in Hz)	HMBC
Starting residue	1	176.01		
	2	45.88	2.07, m	C-1, C-3, C-4, C-5
	3	27.40	1.98, m	C-2, C-4, C-5
	4	22.70	0.85, d (6.6)	C-2, C-3, C-5
	5	22.84	0.89, d (6.6)	C-2, C-3, C-4
R1 Phe	CO	173.58		
	α	56.94	4.57, t (7.8)	C-1, CO, R1- β , R1-1
	β	38.32	3.09, dd (13.8, 7.8) 3.01, dd (13.8, 7.8)	R1- α , R1-1, R1-2
	aromatic	1: 137.89		
		2: 130.39	7.28 (overlapped)	R1- β , R1-1, R1-4
		3: 129.70	7.28 (overlapped)	R1-1, R1-5
		4: 128.06	7.22, m	R1-2, R1-6
	5: 129.70	7.28 (overlapped)	R1-1, R1-3	
	6: 130.39	7.28 (overlapped)	R1- β , R1-1, R1-4	
R2 Dehydro-Val 1	CO	167.99		
	α	141.80		
	β	124.86		
	γ	21.59	1.46, s	R2-CO, R2- α , R2- β , R2- γ'
	γ'	21.07	2.02, s	R2-CO, R2- α , R2- β , R2- γ
R3 Arg	CO	173.41		
	α	54.86	4.32, dd (9.0, 5.1)	R2-CO, R3-CO, R3- β , R3- γ
	β	29.19	1.96, m	R3- α , R3- γ

			1.85, m	
	γ	26.47	1.68, m	R3- δ
	δ	42.04	3.20, t (7.0)	R3- β , R3- γ , R3- ξ
	ξ	158.46		
R4 Dehydro-Val 2	CO	167.86		
	α	141.22		
	β	124.82		
	γ	21.57	1.80, s	R4-CO, R4- α , R4- β , R4- γ'
	γ'	20.91	2.08, s	R4-CO, R4- α , R4- β , R4- γ
R5 Leu	CO	176.12		
	α	52.23	4.49, dd (9.8, 5.1)	R4-CO, R5-CO, R5- β , R5- γ
	β	41.69	1.68, m 1.62, m	R5-CO, R5- α , R5- γ , R5- δ , R5- δ'
	γ	25.79	1.73, m	
	δ	23.39	0.91, d (6.5)	R5- γ , R5- β , R5- δ'
	δ'	21.96	0.91, d (6.5)	R5- γ , R5- β , R5- δ

Table 2. ^1H (600 MHz) and ^{13}C (150 MHz) NMR data of **1** (DMSO- d_6 ; δ in ppm)

Unit	Position	δ_{C}	δ_{H} (J in Hz)	HMBC
Starting residue	1	172.21		
	2	44.26	2.03, m 1.96, m	C-1, C-3
	3	25.47	1.74, m	
	4	21.38	0.89, d (6.6)	
	5	22.84	0.91, d (6.6)	
R1 <u>Phe</u>	NH		8.25, d (6.8)	C-1
	CO	171.20		
	α	54.51	4.63, dd (9.3, 6.8)	R1-CO
	β	38.32	3.07, dd (13.8, 5.9) 2.95, dd (13.8, 9.3)	R1-CO, R1- α , Phe-1, Phe-2
	aromatic	1: 137.89		
		2: 129.20	7.28 (overlapped)	
		3: 129.20	7.36 (overlapped)	
		4: 126.42	7.28, m	
		5: 129.20	7.36 (overlapped)	
		6: 129.20	7.36 (overlapped)	
R2 <u>Dehydro-</u>	NH		9.45, s	R1-CO

Val 1				
	CO	165.62		
	α	135.96		
	β	124.34		
	γ	20.82	1.60, s	R2-CO, R2- α , R2- β , R2- γ'
	γ'	20.40	2.03, s	R2-CO, R2- α , R2- β , R2- γ
R3 <u>Arg</u>	NH		8.11, d (6.9)	R2-CO
	CO	170.59		
	α	53.03	4.26, <u>dd</u> (6.9, 5.0)	
	β	27.83	1.97, m 1.74, m	
	γ	25.29	1.64, m	
	δ	40.49	3.17, q (6.9)	R3- β , R3- γ , R3- ξ
	ξ	156.64		
R4 <u>Dehydro-Val 2</u>	NH		8.96, s	R3-CO
	CO	164.72		
	α	138.43		
	β	124.36		
	γ	21.26	1.74, s	R4-CO, R4- α , R4- β , R4- γ'
	γ'	20.38	2.11, s	R4-CO, R4- α , R4- β , R4- γ
R5 <u>Leu</u>	NH		7.52, d (7.8)	R4-CO
	CO	174.21		
	α	50.26	4.33, <u>ddd</u> (10.1, 7.8, 5.0)	R5-CO
	β	40.06	1.64, m 1.55, m	
	γ	25.49	1.96, m	
	δ	22.27	0.87, d (6.6)	
	δ'	22.16	0.82, d (6.6)	

Figure S 6-16: Original analysis of natural isolated PE2H_3.

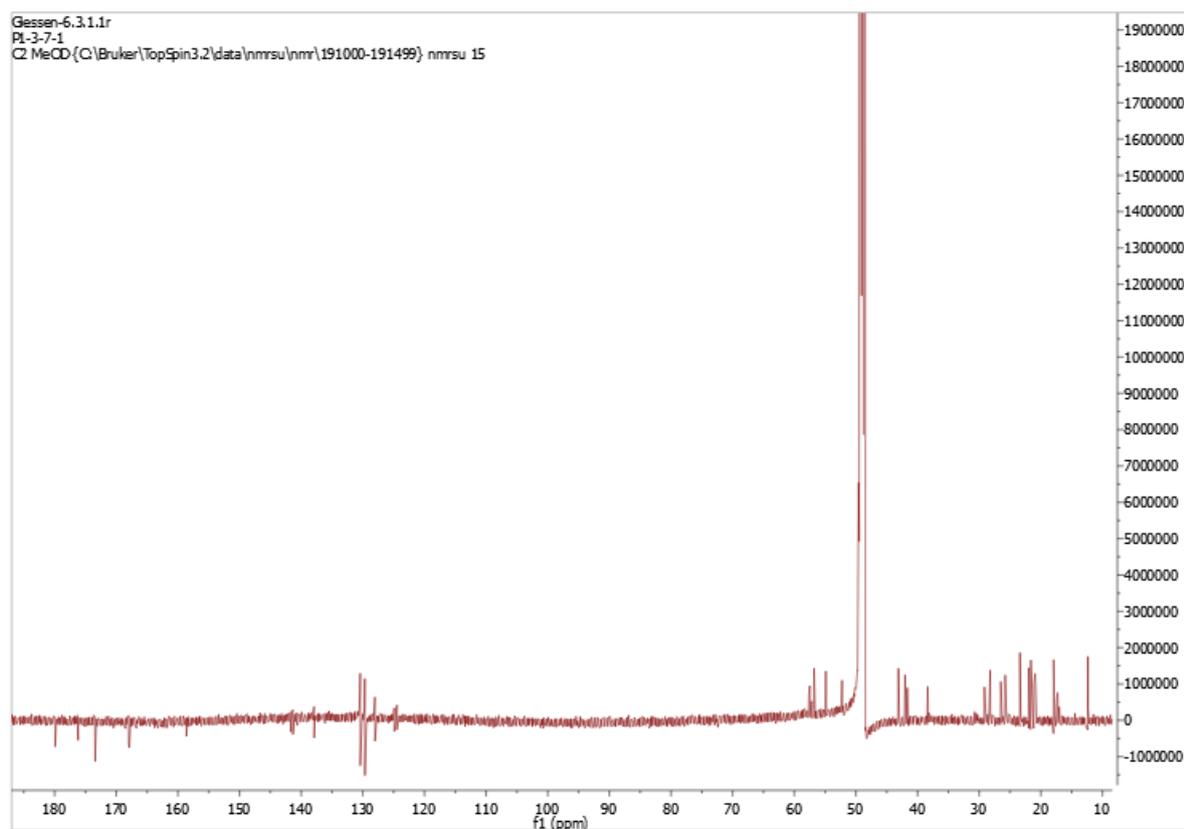
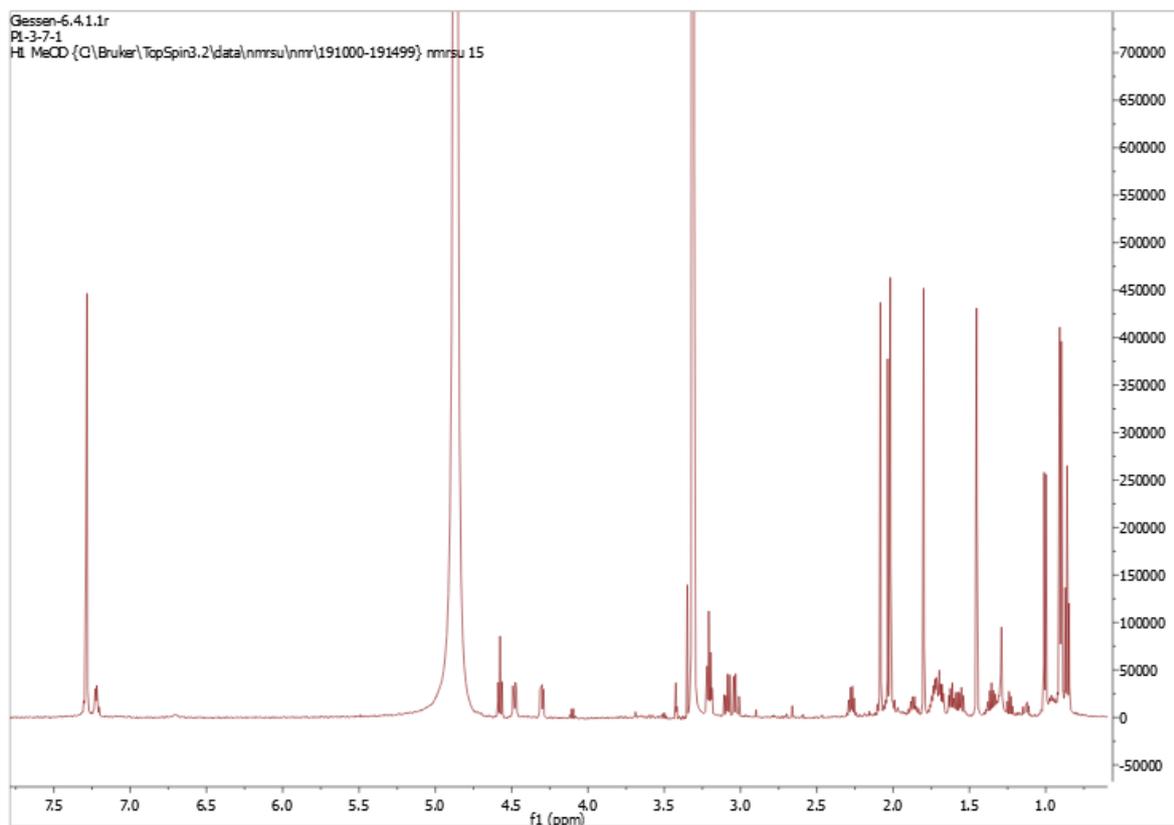


Figure S 6-17: Original NMR spectra of natural isolated fraction P13-71 in MeOD (600 MHz, 150 MHz).

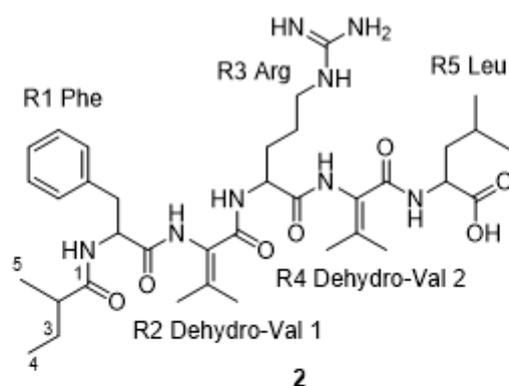


Table 3. ^1H (600 MHz) and ^{13}C (150 MHz) NMR data of **2** (CD_3OD ; δ in ppm)

Unit	Position	ΔC , type	δH (J in Hz)	HMBC
Starting residue	1	180.0, C		
	2	43.11, CH	2.27, m	C-1, C-3, C-4, C-5
	3	28.22 CH ₂	1.56, m 1.36, m	C-1, C-2, C-4, C-5
	4	12.36, CH ₃	0.86, t (7.4)	C-2, C-3
	5	17.91, CH ₃	1.0, d (6.9)	C-2, C-3
R1 Phe	CO	173.42		
	α	56.79	4.58, t (7.9)	C-1, CO, R1- β , R1-1
	β	38.37	3.09, dd (13.7, 7.8) 3.01, dd (13.7, 7.8)	R1- α , R1-1, R1-2
	aromatic	1: 137.89		
		2: 130.43	7.29 (overlapped)	R1- β , R1-1, R1-4
		3: 129.67	7.29 (overlapped)	R1-1, R1-5
		4: 128.04	7.22, m	R1-2, R1-6
	5: 129.67	7.29 (overlapped)	R1-1, R1-3	
	6: 130.43	7.29 (overlapped)	R1- β , R1-1, R1-4	
R2 Dehydro-Val 1	CO	167.99		
	α	141.66		
	β	124.47		
	γ	21.67	1.45, s	R2-CO, R2- α , R2- β , R2- γ'
	γ'	21.11	2.02, s	R2-CO, R2- α , R2- β , R2- γ
R3 Arg	CO	173.42		
	α	54.88	4.30, dd (8.7, 5.4)	R2-CO, R3-CO, R3- β , R3- γ

	β	29.16	2.04, m 1.86, m	R3- δ
	γ	26.50	1.71, m 1.69, m	R3- δ
	δ	42.03	3.21, t (7.1)	R3- β , R3- γ , R3- ξ
	ξ	158.61		
R4 Dehydro- Val 2	CO	167.85		
	α	141.33		
	β	124.91		
	γ	21.62	1.80, s	R4-CO, R4- α , R4- β , R4- γ'
	γ'	20.92	2.08, s	R4-CO, R4- α , R4- β , R4- γ
R5 Leu	CO	176.18		
	α	52.25	4.48, dd (9.8, 5.0)	R4-CO, R5-CO, R5- β , R5- γ
	β	41.65	1.68, m 1.63, m	R5-CO, R5- α , R5- γ , R5- δ , R5- δ'
	γ	25.78	1.73, m	
	δ	21.95	0.90, d (6.4)	R5- γ , R5- β , R5- δ'
	δ'	23.40	0.90, d (6.4)	R5- γ , R5- β , R5- δ

Figure S 6-18: Original analysis of natural isolated P13-71.